Near Ground Level Sensing for Spatial Analysis of Vegetation

Tom Sauer, John Rasure, Charlie Gage

University of New Mexico
Department of Electrical and Computer Engineering
Albuquerque, NM, 87131

Abstract

Measured changes in vegetation indicate the dynamics of ecological processes and can identify the impacts from disturbance. Traditional methods of vegetation analysis tend to be slow because they are labor intensive; as a result, these methods are often confined to small local area measurements. Scientists need new algorithms and instruments that will allow them to efficiently study environmental dynamics across a range of different spatial scales. Presented is a new methodology that address this problem. This methodology includes the acquisition, processing and presentation of near ground level (NGL) image data and its corresponding spatial characteristics. The systematic approach taken encompasses a feature extraction process, a supervised and unsupervised classification process, and a region labeling process yielding spatial information.

1. Introduction

1.1. Motivation

During the 1990's NASA will establish a new remote sensing system, the Earth Observation System (EOS), with a variety of sensors and resolutions. Interpretation of the data at different resolutions will require ground level validation and correlation studies that quantify the heterogeneity of the environment over the range of spatial scales. Both transect sampling (NGL sensing) and remote sensing (satellite sensing) provide data that can identify changes in landscape[1]. Changes in species populations represent shifts in community organization that typically show temporal and spatial variation. Changes in organization among species can occur randomly or in response to governing biotic and abiotic factors[2]. These types of changes can not be detected accurately at the satellite sensing level, and currently the NGL methods used to determine change are typically labor intensive and slow. Thus, there exists a need to develop a new methodology to analyze the NGL sensed data.

This new methodology, also should provide the scientist with information that correlates satellite imagery with NGL imagery. For example, the spectral signature for a pixel in a satellite image provides a single, integrated measure of the ecological patterns within the ground surface area represented by the pixel. The same pixel value may be the result of diverse ground conditions. Without finer resolution imagery it is impossible to determine whether this signature

corresponds to a uniform cover of vegetation or various combinations of vegetation patterns. For the ecological community to make full use of remotely sensed data, it is critical to provide a way to relate the integrated reflectance values to the variety of vegetation patterns that occur at different scales.

1.2. Summary of the NGL Methodology

The NGL sensing system provides absolute and relative measurements of ground level vegetation. The NGL measurement process starts with the acquisition of 35mm color slide images of field plots. The field plot images, ranging in resolution from 1mm to 1cm, and varying in size from 0.5 m^2 to 10 m^2, are obtained using a camera gimbal mounted on a boom. Each rectangular plot is corner marked for later spatial registration. The NGL image is digitized with a high resolution, 4000 by 6000 pixels, slide scanner, and then image analysis is performed using a workstation based software system called Khoros (see Appendix).

The NGL images are comprised of only three spectral bands in the visual region of the spectrum: red, green and blue (RGB). Since the NGL images do not contain spectral information in the infrared region, the image processing analysis that allows differentiation between plant species and the differentiation of above ground biomass and bare ground is more difficult.
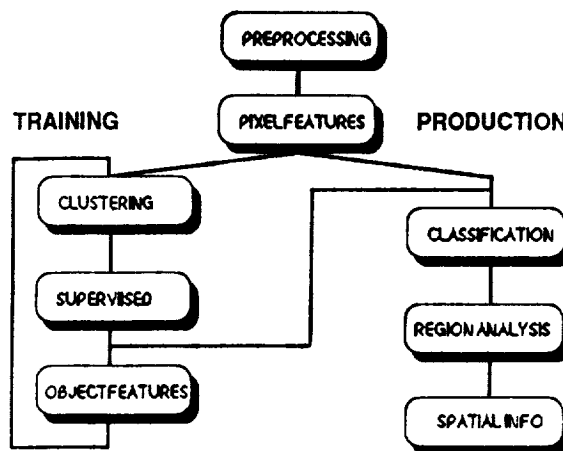

Figure 1. Block diagram of NGL methodology.

The NGL image analysis employs four related components: preprocessing, feature extraction, classification, and region analysis, see Figure 1. Preprocessing can involve warping the image to achieve spatial registration, median filtering to reduce noise, and image cropping. Once image preprocessing is complete, pixel features can be extracted. Pixel features include spectral information, local statistical measures, and various texture measures such as Hurst fractal dimension or the Laws' texture metrics. Representation of the RGB triples in other color spaces often allows for a better segmentation. Local statistical operations that can be computed using Khoros include: mean, variance, contrast, second angular moment, zero order entropy, and dispersion. Spatial feature extraction can be performed on one, all, or a ratio of the preprocessed RGB data bands.

The original spectral bands (RGB) and the feature bands of data are then combined to produce a multiband image. The concept of a multiband image is analogous to a multispectral image; each pixel in the image now contains many elements or attributes (spectral reflectance and features). Since a pixel contains many elements it can be thought of as a vector, where each element in the vector represents a different attribute associated with the pixel. This data organization lends itself to general classification methods.

The overall classification process first involves a one-time training phase that produces a mapping that is then used in the production phase. The training phase is a two part process, unsupervised classification followed by supervised classification. The unsupervised classification portion of the training phase is used to, (1) reduce the complexity and the dimensionality of the multiband image, and (2) determine the inherent structure of the data based solely on reflectance and texture measurements, which are unconstrained by external knowledge of the data. The supervised classification step allows an analyst to map the *clusters* determined by the unsupervised classifier to specific desired classes. The motivation for using the unsupervised classifier first is to reduce the complexity associated with the supervised classification. It has been found that combining the two types of classifiers in this manner produces relatively accurate decision boundaries, and therefore near minimal classification error[3][4].

After a single pass of the training process, object features are obtained that can be added to the original multiband image. This multiband image is used as a new input to the training process, see Figure 1. Object features such as geometric moments, fractal dimension and morphology supplement the pixel features used in the previous pass to produce a more accurate classification. The final result of the training phase will produce data vectors that represent the different cluster means and variances.

The second phase of the classification process uses the results, cluster means and variances, obtained in the training phase to classify other images that fit in the same representative set used in the training phase. Algorithms as simple as a minimum distance classifier, or as robust as the approximated likelihood ratio detector are available in Khoros. The classification process is followed by spatial analysis. Percent coverage of above ground biomass and individual plants is calculated. This information is the basis for the time series analysis that then can be correlated with changes seen at the remote sensing level.

This methodology has been applied to the analysis of images for the Sevilleta Long Term Ecological Research, LTER, project. The National Science Foundation LTER program supports research on long-term ecological phenomena at a national network of sites. One major goal of the LTER project is to study long-term trends in natural ecosystems that have not previously been systematically monitored. The NGL methodology is capable of analyzing image data acquired from large transect plots and small fertilizer plots. A time series analysis of this data can be accurately tracked and eventually correlated with changes seen at the global level. This methodology provides a link that will allow ecological phenomena that occur on large time scales to be investigated.

2. Theory of the NGL Classification System

2.1. Image Preprocessing

Image preprocessing for the NGL images acquired at the Sevilleta LTER site only require geometric correction and image cropping. Median filtering was originally used to reduce noise artifacts, but the final spatial measurements exhibit distortion caused by the smoothing effects of

13

the filter.

Since the NGL data are captured using a 35mm camera gimbal mounted on a boom, the image will be distorted because of camera position and terrain topology. The corner markers in the image provide tie points that will allow the image to be warped back to the correct geometry. Since the acquisition system uses 35mm slides, the xy pixel ratio in the image is 2/3, and must be corrected back to a 1/1 pixel ratio when the slides are digitized. The Khoros interactive image editor allows a user to select the image corner points and record the xy locations as source tie points. The user must then specify the distance between the tie points. From this information the destination tie points may be computed. The following table and equations describe the destination tie point computations

$$\alpha = \frac{L}{(x'^{2}+y'^{2})^{\frac{1}{2}}} \tag{1}$$

Where : $\quad x' = k(x_{p_1}-x_{p_2}) \quad y' = k(y_{p_1}-y_{p_2})$

p_1 *is some tie point*

p_2 *is some other tie point*

k is a pixel aspect ratio constant

L is the actual distance between p_1 *and* p_2 *in meters*

α *is the new coordinate position translation factor*

Source tie points	Destination tie points
(x_1,y_1)	$(0,0)+(x_1,y_1)$
(x_2,y_2)	$(\frac{1}{\alpha},0)+(x_1,y_1)$
(x_3,y_3)	$(\frac{1}{\alpha},\frac{1}{\alpha})+(x_1,y_1)$
(x_4,y_4)	$(0,\frac{1}{\alpha})+(x_1,y_1)$

The next step in the registration process is to use the four source and destination tie point pairs to compute the coefficients for two first-order equations that will be used to perform the image registration. In some cases, a wide angle lens causes severe image distortion. This requires the use of more than four tie points, resulting in a higher order warping polynomial. The original image is warped using the computed polynomial equation and bilinear interpolation, and then cropped using the destination tie points.

2.2. Feature Extraction

Texture measures are typically computed on a single band image, necessitating the reduction of the RGB image to a single band image. This reduction is performed using a color quantizer that reduces the image from 16.7 million colors (3 bands) to 256 colors (1 band). Alternatively, the RGB image can be converted to the HSV (Hue, Saturation, and Value) color space with the spatial measures computed on the value band.

Simple statistical parameters based on local area measurements over a small moving window are commonly used to provide texture information[5]. The statistical parameters, mean and variance, are based on the central moments and are used to provide an indication of how uniform or regular a region is. Contrast provides a measure of the dissimilarity of the intensity values in the image, and angular second moment yields a measure of uniformity or homogeneity of the gray level values. An indication of the texture nonuniformity is provided by a measure of the entropy. Texture measures based on these statistical parameters did not yield any new information that aided in the classification process. For this reason, features based on simple statistical parameters were not used.

Although numerous texture measures have been proposed to characterize the spatial texture features in an image, good results have been obtained using a set of spatial convolution masks proposed by K. I. Laws[6]. The Laws' texture masks are comprised of a set of 5 by 5 masks that are convolved over the entire image[7]. The masks are intended to be sensitive to visual structures such as edges, ripples, and spots.

Each of the Laws' texture masks are derived from a set of five basic vectors. There are a total of 25 possible masks, each formed by multiplying two of the five vectors together. They are designed to act as matched filters for certain types of quasiperiodic variations commonly found in textured regions.

Various texture masks were tried in order to achieve good discriminating power between adjacent regions in the image. The set of texture masks that provided the best results include the L5E5 and E5L5 masks. The L5E5 and E5L5 masks are constructed by multiplying the L5 and E5 vectors, yielding the following texture masks:

$$L5E5 = \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \qquad E5L5 = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

The L5E5 mask tends to detect edges arising from horizontal changes in texture, while the E5L5 mask detects texture changes in the vertical direction.

Once the spatial texture features are extracted by convolving each mask with the gray level image, an additional feature selection step is used to reduce the dimensionality of the classification process. This involves a 50% blending of the two texture bands into one texture band that contains the information extracted by each of the texture masks. By using one texture band, the overall weighting of the texture features relative to the spectral image information is reduced. This provides a more representative weighting of the original spectral information relative to the spatial texture information in the classification process.

2.3. Classification

As was mentioned above, the classification process is a two phase process. The first phase is considered the training phase, while the second phase implements the actual classifier and is referred to as the production phase. The training phase is further broken into two parts, unsupervised classification and supervised classification. The unsupervised training determines the inherent structure of the data, unconstrained by external knowledge about the vegetation patterns, while the supervised training imposes the analyst's knowledge of the vegetation patterns to constrain the results. The final objective of the classification process is to reduce a large data set

15

(the multiband image) into a few classes in a single band image.

2.3.1. Training Phase

The goal of the training phase is to produce an ensemble of data that characterizes a representative set of NGL images, so that other images with similar characteristics can be *automatically* classified (the production phase). In the unsupervised classification portion of the training phase the algorithm maps areas on the ground that have similar texture and spectral reflectance characteristics to the same cluster. The resulting clusters assigned to the image pixels therefore represent different *classes* that may or may not correspond to the classes of ground objects that we are ultimately interested in mapping. A good example of such a situation is the mapping of shadow areas and wet or dark soil areas. The analyst may want to ultimately consider both of these classes as bare ground, but each may represent a separate cluster as produced by the unsupervised classifier. The output of the unsupervised classifier is a single band pseudo colored image that represents a map of the clustered pixel vectors, the cluster centers (means), and variances. The mean and variance data represent the ensemble of data that characterizes a specific set of NGL images.

Image data that represents specific areas to be classified are submitted to the unsupervised classifier. The unsupervised classifier is implemented as a clustering algorithm that will determine the *natural* groupings of clusters of the data in K-dimensional feature space. The cluster centers represent an estimate of the probability density function. The cluster centers are then assigned to classes during the supervised classification. The determination of clusters is accomplished by the K-means clustering algorithm[8].

The K-means algorithm is a partitional algorithm that attempts to minimize the sum of squared errors in its cluster assignments. The similarity measure used is the Euclidean distance. The K-means algorithm partitions the data space by using a search method where patterns are moved from one cluster to another until all patterns belong to a cluster. Each cluster is identified by a single cluster center (mean) and cluster variance. Since the K-means algorithm uses the Euclidean distance as a similarity measure, it is vital that the features previously determined are weighted so as not to bias the results produced by K-means. The performance of K-means is improved if the feature pixel vectors are orthogonal. In practice, however, this is rarely the case. Therefore, it is best to *over-cluster* the pixel vectors resulting in a less refined classification.

Experiments show that the number of clusters produced by K-means should be about four to seven times the final number of classes desired. The cluster centers provide the location in K-dimensional space for each cluster, while the variance describes the size and orientation of each cluster. This information is used in the supervised classification described below.

The output of the unsupervised classifier provides a mapping of pixels in the original image to different clusters. The clusters produced by the unsupervised classifier are usually not the desired classes; the object of the supervised classifier is to map each cluster to a desired class. The supervised classification process is performed manually using the Khoros image editor. The Khoros image editor allows the analyst to display both the clustered image and the original image. Cluster numbers in the clustered image can then be assigned to specific desired classes. The resulting mapping of clusters to specific desired classes will be used in the production phase of the classification process.

Often the data in a cluster may need dividing because it is spread over multiple desired classes. The P(m,L) fractal algorithm can be used to help determine the splitting of the clusters.

The P(m,L) distribution is obtained from the unsupervised classification image data. Frequency distributions for each class m, are determined for a series of different window sizes, L. The resulting probability distribution should provide valuable information describing the aggregation of classified pixels in the NGL image.

The P(m,L) probability density function has moments that vary with the measurement scale. This scale dependent characteristic of the moments provides a framework for transforming plant coverage estimates from one scale to another. It has been found that natural landscapes often exhibit consistent changes in the fractal dimensions over a range of moments[9]. This provides a way of measuring the degree of relationship from one scale to another.

Once the moment bands have been determined, they will be appended to the multiband image containing the spectral and texture bands. This image will then be reprocessed by the training phase. The result of this iterative processing will produce statistics (cluster means and variances) that better describe the desired classes.

2.3.2. Production Phase

The object of the this phase is to take the mapping obtained in the training phase and allow unsupervised classification of subsequent images that are considered to be in the same representative data set as used during training. It is required that the same feature extraction process is performed on the new images as was performed on the training set. The unsupervised classifier used in this phase is the approximated likelihood ratio detector (ALRD). The ALRD uses the cluster centers, cluster variances, and cluster to class mapping to classify new images. This robust unsupervised algorithm is not limited to detecting whether a pixel vector belongs to a single class. A pixel vector can be assigned to multiple classes and through a thresholding test determine to which class the pixel best belongs. If a pixel vector does not have a high enough probability to belong to any class, then it is considered an outlier, and thus unclassifiable. This algorithm uses the ratio of the distance of a pixel vector to a cluster center to each diagonal element of the covariance matrix (variance elements), to determine to which class a pixel belongs. In other words, the algorithm computes the probability density function of all clusters that belong to a class and then determines if a data point has a high enough probability to belong to that class. The diagonal of the covariance matrix is computed by Equation (1) while Equations (2) and (3) perform the likelihood ratio test.

$$diag\,(C_i) = \begin{bmatrix} \dfrac{1}{N_i}\sum\limits_{l=1}^{N_i}(X_{0il}-m_{0i})^2 \\ \dfrac{1}{N_i}\sum\limits_{l=1}^{N_i}(X_{1il}-m_{1i})^2 \\ . \\ . \\ . \\ \dfrac{1}{N_i}\sum\limits_{l=1}^{N_i}(X_{nil}-m_{ni})^2 \end{bmatrix} = \begin{bmatrix} \sigma_{0i}^2 \\ \sigma_{1i}^2 \\ . \\ . \\ \sigma_{ni}^2 \end{bmatrix} \qquad (2)$$

where N_i is the number of points in the ith class, m_i is the cluster mean, l is an index into the list of data points belonging to ith cluster, and n is the dimensionality of the vector.

$$P_i = \sum_{j=0}^{D} \frac{||X_{ji}-m_i||^2}{K\sigma_{ji}^2} \qquad (3)$$

17

P_i is the likelihood ratio of the ith class, D is the dimension of the unclassified data vector X, and K is a tuning parameter.

$$class_X = \begin{cases} no\ class\ \text{if}\ P_i > 1 \\ class\ i\ \text{if}\ P_i < 1\ minimum\,(P_i) \end{cases} \qquad (4)$$

The diagonal of the covariance matrix and the cluster means are computed during the training process. The similarity measure used by the ALRD is the same as that used in the K-means algorithm, the Euclidean distance. The tuning factor adjusts the likelihood ratio, which either increases or decreases the number of outliers detected.

The ALRD is used rather than the minimum distance classifier because it allows for outlier or unclassifiable pixels. This reduces the size of the training set because it eliminates the need to classify every possible pixel vector. The ALRD also uses both the size and orientation of the classes in K-dimensional feature space to aid in classifying new pixels.

2.4. Region Analysis

The final step in the NGL image analysis is the calculation of class and region moments [10]. For example, in the case of a two class image (above ground biomass and bare ground), the area calculations result in percent vegetation cover. More detailed information can be obtained by labeling the individual objects in the two class image and then calculating moments.

Labeling of individual objects is based on the splitting and merging of regions, where the decision metric is the gradient between eight-connected neighbors. The labeling algorithm uses either the difference between the gray levels of adjacent pixels or the Euclidean distance between adjacent pixel vectors as the gradient value. If the gradient value is less than a threshold the regions are merged. The moment calculations (standard, central, and invariant) on the resulting labeled regions give detailed spatial information on each object. This information provides the analyst with the necessary information to track individual plant changes over time.

The region analysis algorithm generates two images; (1) an axis image that contains a cross for each region with the cross centroid located at the center of the object, and (2) a region outline image or contour image. Overlaying the outline image upon the original RGB image or the axis image, provides the analyst with a means of visual interpretation and verification. This assists in the time series analysis since it allows the analyst to visually track the vegetation changes.

3. Discussion of a Specific Example

The NGL methodology has been applied to helping ecologists at the Sevilleta LTER site track the vegetation change in both transect and fertilizer plots. The transect plot dimensions are usually 10m by 5m, and fertilizer plots are usually 1m by 0.5m. The following example will illustrate the results obtained by using the NGL methodology on transect images. A representative image of a transect area in the field is used as the training pattern, then another image is classified based on the results from the training. In this example, vegetation is segmented from all other matter, thus a two class problem.

This example begins with a representative transect plot image, Figure 2, that has been spatially registered and cropped.
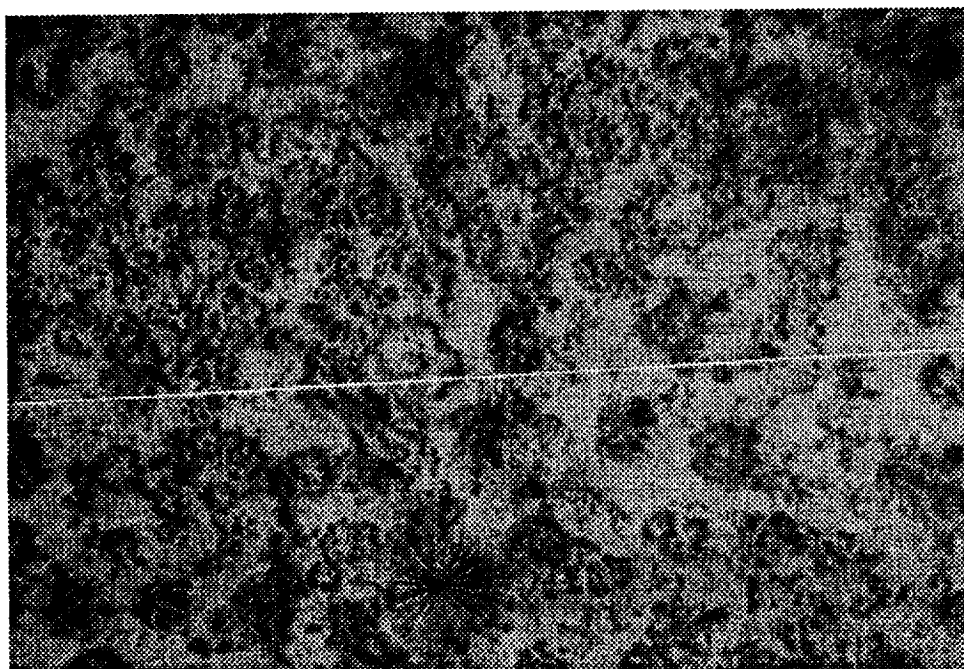
Figure 2. Representative transect plot image.

The next step in the process is to compute pixel features. Pixel features are computed using the Laws' texture metrics. Since these metrics only work an a single band image, the original RGB image must be compressed down to a single band. This is performed using the color quantizer method. Two different Laws' texture kernels, E5L5 and its transpose L5E5, are convolved with the one band image producing two single band images that are blended together producing another single band image, shown in Figure 3.

The texture band is then appended to the end of the original RGB image. This new multi-band image is used in the classification training process. The K-means algorithm, produces a single band cluster number image shown in Figure 4.

Figure 5 illustrates a plot of the distribution of the cluster centers. Each row of impulses represent a different set of cluster center values. This plot gives a visual interpretation of the correlation between different cluster centers.

The next step is the supervised classification phase of the training. Cluster numbers are assigned to specific classes using the Khoros interactive image editor. The result of the supervised classification is shown in Figure 6.

At this point the training can stop if all the clusters have been mapped to the desired classes. Otherwise, object features are computed and the system is retrained. In this example all clusters have been mapped to the two desired classes. This ends the training phase.
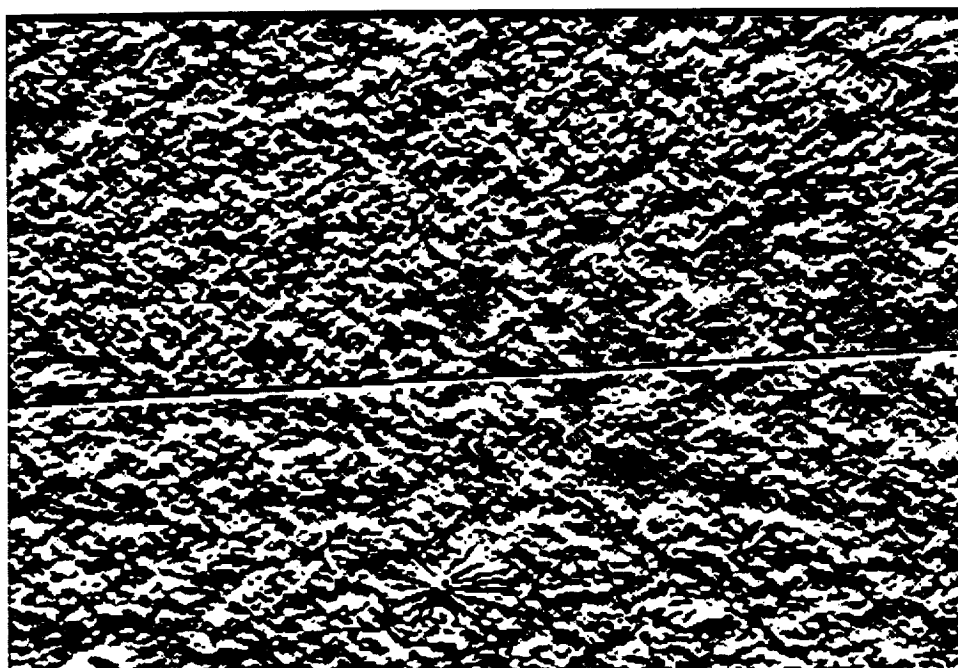
Figure 3. Laws' texture band.
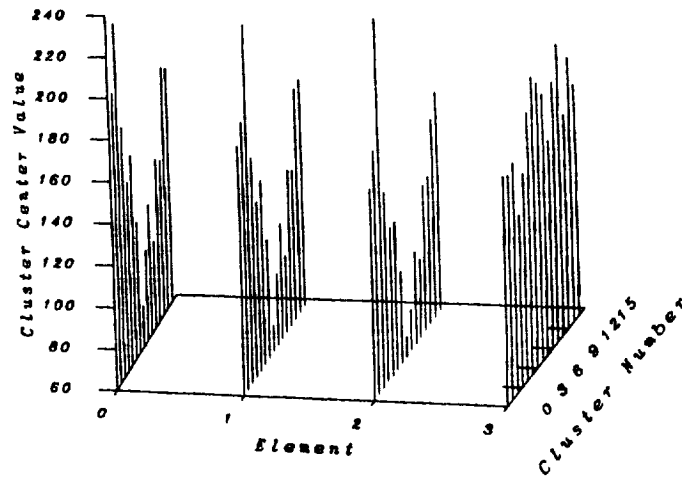

Figure 4. Cluster number image.
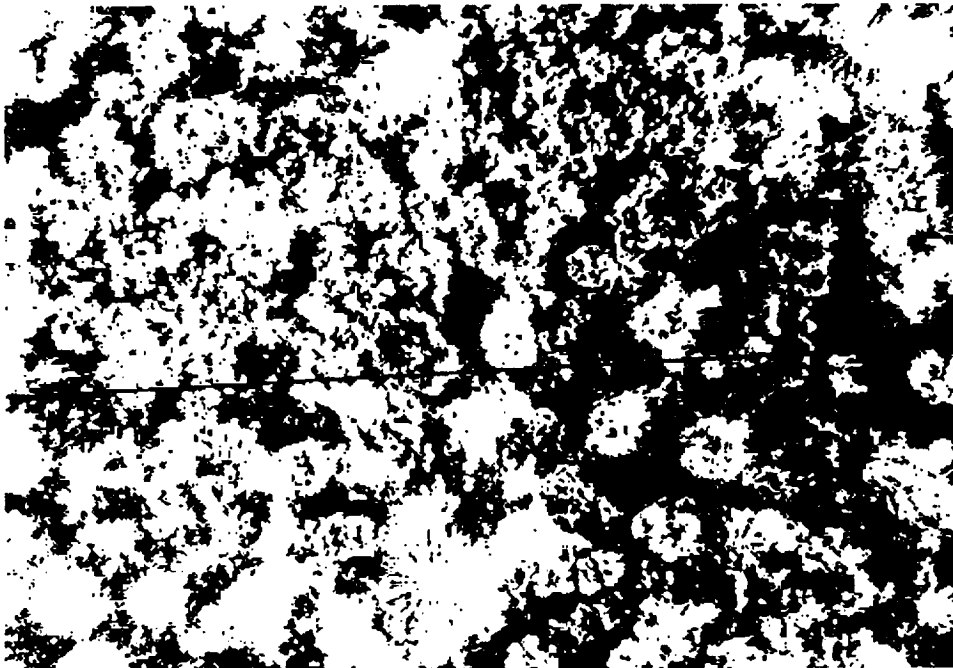
Figure 5. Plot of cluster centers.

Figure 6. Resulting class image after training phase.

21

With the results produced by the training phase, other images from the same representative set can be classified using the approximated likelihood ratio detector. The new image to classify is shown in Figure 7.
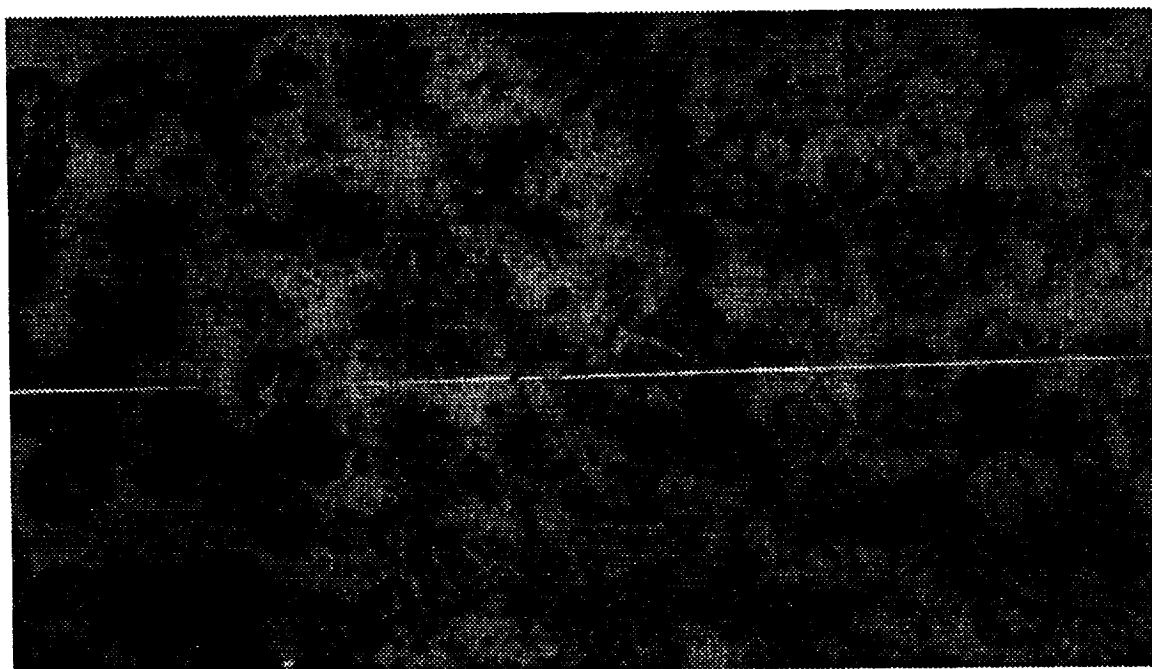

Figure 7. Transect image to be classified.

The same image preprocessing and feature extraction is performed on this image as on the training image, resulting in a multiband image. The result of the unsupervised classification is shown in Figure 8.

The final step is to perform region and class analysis to determine the desired spatial information. Figure 9 illustrates the result of the analysis procedure. This image shows the size of the regions by outlining them and the orientation by the crosses in each outlined region. In this example the percent coverage of vegetation is 38.43%.

This example illustrates the applicability of the NGL methodology for the Sevilleta LTER project. The system is planned for production use by the end of the year. The ecologists see this approach as critical to the successful and timely analysis of the thousands of transect and fertilizer plot images required by the project.
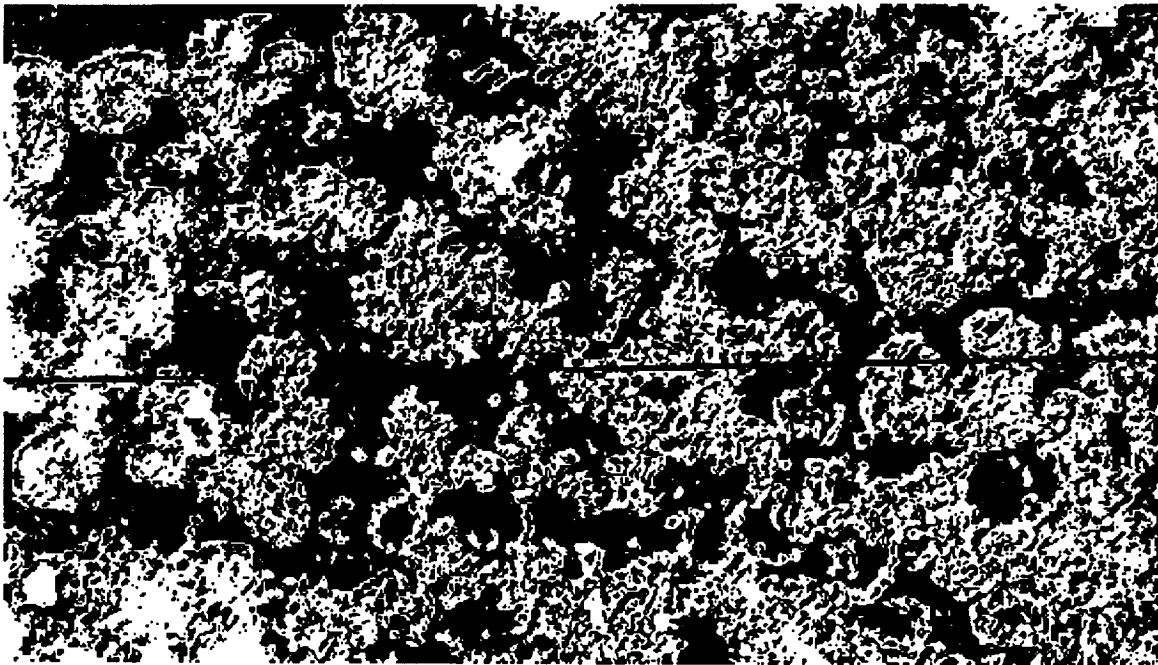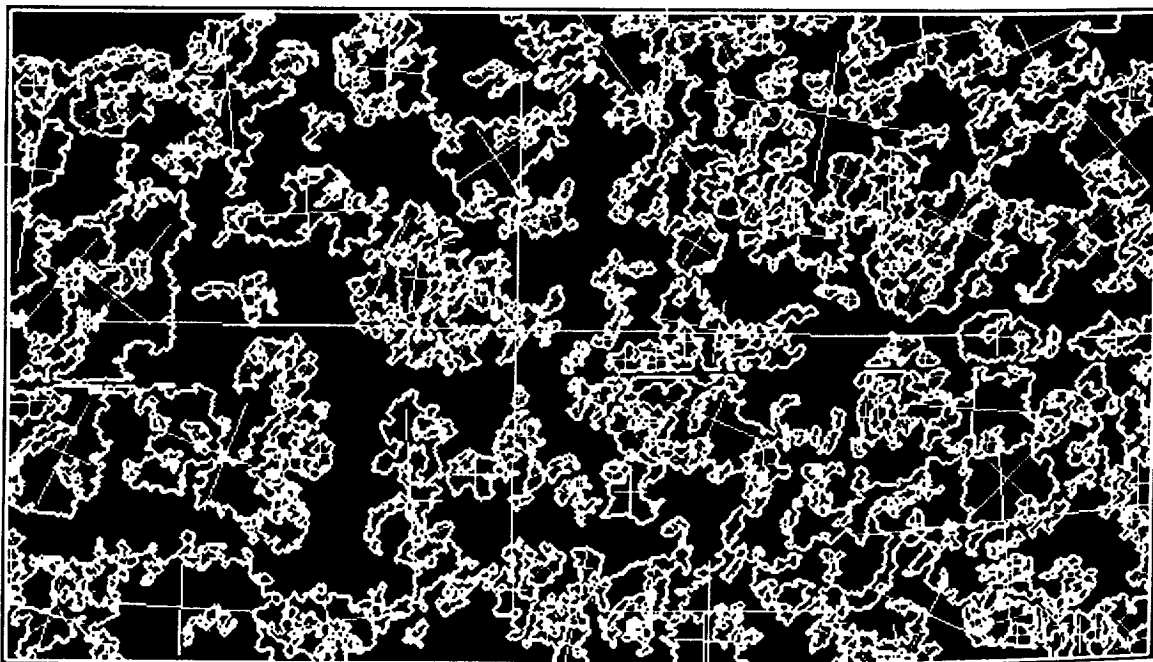
Figure 8. Resulting image after classification.

Figure 9. Region outline image.

23

4. Conclusion

This paper presents a new methodology for ground level vegetation analysis. It emphasizes an integrated approach using existing algorithms and introduces a new classifier, the approximated likelihood ratio detector. Some of the techniques used in the analysis of the NGL imagery include preprocessing, feature extraction, classification, and region analysis. The goal is to allow the scientist to correlate information obtained at the satellite sensing level with more detailed information contained in the NGL imagery. Existing techniques based on satellite imagery do not provide enough detailed information for a complete vegetation analysis. The approach presented here provides a means of accurately tracking and quantifying the vegetation changes across a range of different scales. Future development of this system includes the integration of spatial results of NGL images into GIS.

Appendix

The Khoros system integrates multiple user interface modes, code generators, instructional aids, data visualization, and information processing to produce a comprehensive image processing research tool. This system can easily be tailored to other application domains because the tools of the system can modify themselves as well as the system. This attribute is important in a system that is designed to be extensible and portable.

The Khoros infrastructure consists of three major components: a high level user interface specification, methods of software development embedded in a code generation tool set, and an interoperable data exchange format and algorithm library. These basic facilities have been used to build a set of applications for performing image processing research, algorithm development, and data visualization. One of the most powerful features of the system is its high-level, abstract visual language.

Khoros is a successful demonstration of how development programming, end-user applications programming, information processing, data display, instruction, documentation, and maintenance can be integrated to build a state-of-the-art image/data processing and visualization software environment.

References

[1] R.A. Schowengerdt, *Techniques for Image Processing and Classification in Remote Sensing*, Academic Press, New York, 1983.

[2] J.W. Brunt and W. Conley, "Behavior of a Multivariate Algorithm for Ecological Edge Detection", *Ecological Modeling*, 49, 1990, pp. 179-203.

[3] D. Hush and J. Salas, "Classification with Neural Networks: A Comparison", *Proc. 11th Annual ISE Conference*, Albuquerque, NM, May 1989.

[4] D. Hush, "Classifiers: Results for I4-4I Case with 8 Dimensions", *Sandia National Laboratory Report*, Department 9133, January 5, 1989.

[5] R.M. Haralick, "Statistical and Structural Approaches to Texture", *Proceedings of the IEEE*, vol. 67, no. 5, May 1979, pp. 786-804.

[6] J.Y. Hsiao and A.A. Sawchuk, "Supervised Textured Image Segmentation Using Feature Smoothing and Probabilistic Relaxation Techniques", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 12, December 1989, pp. 1279-1292.

[7] K.I. Laws, "Rapid Texture Identification", *Proc. SPIE*, vol.238, 1980, pp. 376-380.

[8] J.T. Tou and R.C. Gonzales, *Pattern Recognition Principles*, Addison-Wesley Publishing Co., Reading, Massachusetts, 1974.

[9] B.T. Milne, "Spatial Aggregation and Neutral Models in Fractal Landscapes", Submitted to *The American Naturalist*, December 1989.

[10] M.D. Levine, *Vision In Machine and Man*, McGraw Hill, New York, 1985.