Image Understanding Using Artificial Intelligence Technology

Dave Kargel
Geographer

Will Brandt
Engineer

PRC, Inc.
111 West Port Plaza, Suite 327
St. Louis, MO 63146

Abstract

An artificial neural network approach was evaluated in multispectral image processing applications, including general land cover classification and land use feature identification. The performance of an artificial neural network was compared to that of a standard statistical classification technique available from a commercial image processing package (ERDAS). Landsat Thematic Mapper 30 meter multispectral data of St. Louis, Missouri were utilized for the classification. Seventeen distinct land cover types were identified in the St. Louis study area. Supervised classifications were performed with both the neural network and the standard statistical methods. The results show that the neural network is more responsive to cultural material with skewed spectral distributions and sub-pixel spatial resolution (e.g. asphalt roads). The neural network's probabilistic output also enables discrepancy resolution capabilities for the higher level task of land use feature identification. A model was developed to identify roads from the multispectral classifications. The neural network based model performed much better than the model based on the standard statistical classification technique.

1. Introduction

The PRC Image Understanding Using Artificial Intelligence Technology Independent Research and Development project, initiated in 1991, is presently evaluating the utility of neural network and expert system technologies in management, classification and feature extraction functions related to multispectral images. The long term goal will be to build a prototype capable of handling numerous types of images and performing the various functions and analyses now performed manually.

As users of geographic data better understand the value of remote imagery, the demand for automated image processing and feature extraction capabilities increases. Many government agencies and private firms have poured money into research and development efforts addressing image understanding applications of particular concern to them. Many of the applications currently utilize techniques combining computer manipulation and human interpretation of the imagery.

Currently, the application of spectral and spatial enhancement functions, derivation of spectral signature ranges, and coordinate transformations for georeferencing are all performed by computer algorithms. However, the choice of computer algorithms is still dependent on the direct interface with a professional image interpreter who is knowledgeable of the user application and the corresponding relevance of computer algorithms.

Artificial intelligence can be used to encode the knowledge base and decision rules implemented by professional image interpreters. Many of the repetitive decisions made in image interpretation may be delegated to artificial intelligence based automation. The added automation will increase the speed, precision, accuracy and adaptability of the image interpretations. The resulting data will be cheaper in terms of time, money, and necessary expertise.

This report contains three sections. The following section discusses the approach we have taken in automating the image understanding task. This includes a specific discussion of critical areas in the image understanding task and how they may be automated. The next section reports the most resent research results. The results are divided based on the critical areas addressed. The final section contains our conclusions.

2. Automation Methodology

2.1 Image Inventory

An image inventory capability is an essential part of an automated image understanding system. A properly indexed image inventory system allows the correct images to be chosen from inventory for each different application query received. The images must be inventoried by the query criteria used to select images for each application. This synergistic set of criteria includes spectral resolution, spatial resolution, temporal issues and the georeferencing system and of the inventoried imagery.

We are convinced that the task of selecting the imagery required for an application can be most effectively automated by implementing an expert system based approach. The expert system is given a rule base and knowledge set incorporating the decision process of the professional image interpreter. This selection process will allow the automated image understanding system to be nearly as flexible as a human interpreter. Figure 1 shows a simplistic expert system architecture.

2.2 Image Georeferencing

The inclusion of an automated image georeferencing process will allow proper image registration when an application calls for a specific coordinate system. The georeferencing process will be applied as an automated point matching algorithm from a newly acquired image to a previously georeferenced image stored in inventory [TON89; PERL89]. An expert system can then apply the proper transformation process and select any coordinate system required for an application. The georeferencing information for an image can be stored automatically in the image inventory. This will facilitate the georeferencing and the image selection processes.

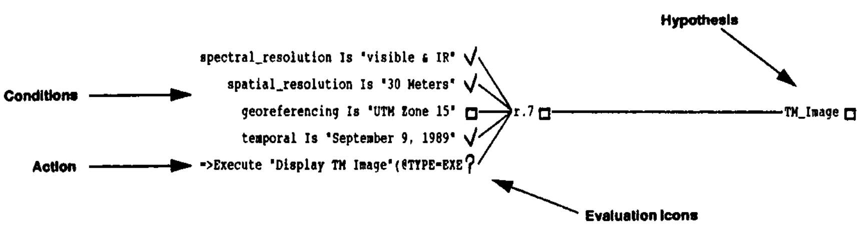

Figure 1. Expert system architecture

2.3 Spectral Image Processing

PRC will automate the spectral classification process using an integrated neural network and expert system based solution. While an expert system is an artificial intelligence architecture which represents expert knowledge as an explicit rule set and knowledge base, a neural network architecture represents expert knowledge as a generalization of examples presented to it. This is beneficial when the development of explicit rules would be too complex or constricting. PRC will integrate these two artificial intelligence architectures into a synergistic spectral classification solution.

The use of neural network technology will improve the classification and the interpretation of spectral signatures. Figure 2 shows a very simplistic neural network architecture. A neural network architecture consists of a large network of simple processing elements or nodes, represented by circles, which process information in response to examples. The input layer can accept pixel values from different spectral bands of raw satellite imagery. The hidden layer stores features it learns from the input layer. Finally, the output layer contains nodes representing landuse classes into which the raw pixel values will be classified. The knowledge in a neural network is distributed throughout the network in the form of internode connections and weighted links which form the inputs to the nodes. The link weights serve to enhance or inhibit the input stimuli values which are then added together at the nodes. If the sum of all the inputs to the node exceeds some threshold value, the node executes and produces an output which is passed on to other nodes or is used to produce some output response [NEUR91; CARP92; JONE87].

Neural network classification techniques provide several useful advantages over standard classification techniques. Unlike the standard classification techniques, which attempt to fit each classification scheme into a standard distribution curve, a neural network is able to handle classes with skewed distributions. These features give neural networks a fault tolerant behavior which is generally referred to as "graceful degradation". Graceful degradation allows such systems to operated successfully in noisy environments [BISC92], such as, a mountainous area with associated areas of bright reflectance and shadows. A neural network is also able to provide classifications with associated confidence levels for all possible material classes (e.g., water-55%, asphalt-45%) [NEUR91]. This can be useful information for automating the problem of discrepancy resolution in pixel classification. For example, if an area is identified as asphalt but includes a water pixel in the middle, it may be analyzed for discrepancies. If the classification probabilities for the water pixel are equal to the ones shown above, the 45% asphalt classification will be accepted as the true classification.

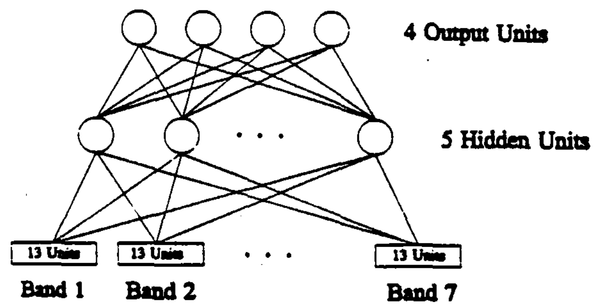

Figure 2. Neural network architecture

An expert system will be used to optimize the neural networks classifying the pixels. By referring to the information stored in the image inventory, the expert system can allow the neural network to compensate its signatures for different imaging conditions. For example, an image taken at noon on a sunny, summer day will have brighter streets than an image taken at dawn on a rainy, winter day. This will allow for complete automation of the spectral classification process by incorporating into a system, the flexibility associated with the human interpreter.

Such a system will also be equipped to learn signatures for materials it was not trained for. The system operator will follow the steps for standard supervised classifications by identifying representative pixels in the image from each target class. Then using the information stored in the image inventory and the expert system, the system will build an image independent spectral signature for each target class compensating for the atmospheric conditions of the imagery. These signatures can then be included in the fully automated spectral classification portion of an image understanding system.

2.4 Feature Extraction

The automation of the feature identification process involves the integration of the spectral classification information and low-level spatial feature detection. The low-level spatial features, such as lines, edges, and regions, will be detected by corresponding spatial image processing algorithms. Many such general low-level spatial image processing algorithms have been developed and proven successful [THOM87; FUA87; TON89]. The particular features to be detected will be decided by the expert system based on the present application query.

The expert system will then initiate a high-level feature identification process. This process, illustrated in Figure 3, will integrate the material classification information for individual pixels and the low-level feature groups they correspond to [HUER87]. For example, an asphalt pixel corresponding to a linear feature is identified as a road. This can be handled by a technique called raster modelling which combines different data sources and derives new classifications for pixels based on a structured spatial processing language [TOML90].

To identify the more complex high-level features, such as airports, it will be useful for the expert system to convert the simpler features to vector format. In the vector format, data is stored as discrete objects located at discrete locations in coordinate space, unlike raster data which generalizes objects into a predetermined grid format. In the vector format, the expert system is able to build features based on rules of topology and areal composition, such as shape and texture [HARW87; HARV92]. These types of rules can distinguish between an orchard and a woodland, with an orchard having a uniform pattern of trees and grassy areas and a woodland having a more random pattern.

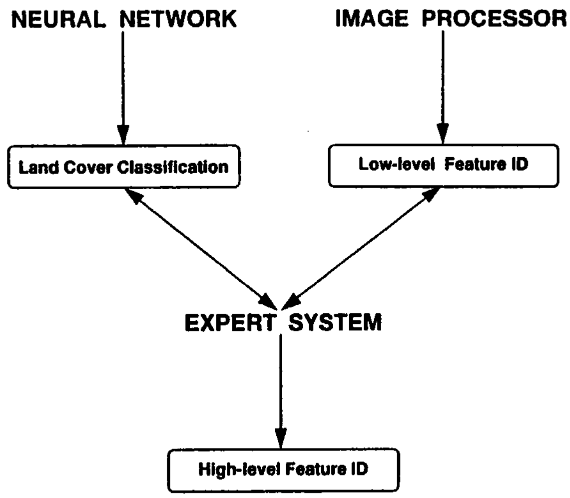

Figure 3. Expert system integration of spectral and spatial experts

3. Image Understanding Research Activities

3.1 Spectral Image Processing

Our research in the automation of spectral image processing examined the use of a neural network for the classification of water, pavement, and vegetation from a September 9, 1989 Landsat Thematic Mapper (TM) 30 meter multispectral image of St. Louis, Missouri, USA. The spectral classification was automated with a uniquely configured back-propagation neural network that was developed using NeuralWorks Professional II/Plus by NeuralWare, Inc., a commercial neural network builder. The neural network contains six spectral class

10 17 255 50 0	-1 0 2 0 -1	0 0 255 0 0
20 50 255 20 0	-1 0 2 0 -1	0 0 255 0 0
15 20 255 20 0	-1 0 2 0 -1	0 0 255 0 0
Input image values (first band)	Line filter	Output image values

Figure 4. Spatial enhancement: line detector (vertical, bright)

inputs (blue, green, red, near-infrared, mid-infrared-band 5, and mid-infrared-band 7), nine hidden units, and 17 land cover class outputs (main channel, shallow channel, deep lake, shallow pond, lake edge, runway, building, trees, grass, sand, pavement [1-2], clay, vegetation [1-4]).

The neural network was trained with a file containing 1000 pixel values from each land cover class. The resulting spectral classification was visually verified against an April 10, 1990, 1 inch to 400 foot panchromatic aerial photograph of the study area. The aerial photograph provides a good representation of the land cover within the St. Louis image. The neural network identifies all the target land cover classes. It also identifies more of the mixed pixel pavement surfaces than the ERDAS supervised classification technique.

3.2 Low-level Feature Extraction

The PRC research and development efforts focusing on automated low-level feature extraction have concentrated on line detection algorithms. The basic line detection algorithms applied were derived from work presented by Dr. Jezching Ton [TON89]. The Ton algorithms were chosen because of their direct application to road detection, reportedly high success rate, and straight forward development.

While Ton compiled his computer algorithms in Pascal, our approach was to develop batch files in ERDAS, a commercial image processing package. The line detection algorithms were tested on the TM image described in the previous section. The basic Ton algorithms involve the application of a special set of spatial enhancement techniques called convolution filters. Figure 4 shows this application. The filters are designed to detect bright lines in Landsat TM data in eight different directions. The line filter in Figure 4 is designed to detect vertical lines. A two band image is derived as output from the filter application. The first band is composed of the highest value returned for each pixel from all of the applied filters. The second band is the direction associated with the filter returning the highest value for each pixel. The two band output image can them be interpreted as line segments with associated intensities and directions. In Figure 4, the bright pixels in the center of the input image are highlighted from the background pixels in the output image.

The line detection algorithms, written by Regional Research and Development Service (RRDS) at Southern Illinois University at Edwardsville (SIUE), were composed of ERDAS batch files and C code. The algorithms were then applied to Band 3 of the Landsat TM image of St. Louis. This band is highly responsive to cultural features. The results of the line detection algorithm on the St. Louis data are encouraging. A visual inspection based on the airphotograph mentioned earlier shows that virtually all of the major lines and most of the minor lines are detected. It is believed that the low resolution of Landsat TM data is responsible for the difficulty in detecting some lines representing roads narrower than 30 meters. We feel confident that the line detection algorithms we have developed provide accuracy levels high enough to be used in an automated feature extraction system. The true test will be the accuracy levels achieved by the high-level feature identification algorithms that utilize the line detection results as input.

3.3 High-level Feature Extraction

The first high-level feature extraction research has focused on road identification algorithms. These algorithms attempt to integrate the results received by the spectral land cover classifications and the low-level line detection algorithms. The road identification algorithm is a raster model developed to replicate a simple expert system based decision algorithm.

The road extraction algorithm was applied to the St. Louis TM data. As mentioned earlier, a neural network classification includes weights from zero to one representing the probability of the pixel corresponding to each class. The classes with the two highest probability for each class were chosen as input to the road detection algorithm. These two levels of neural network material classification were then recoded to pavement and non-pavement materials. This information was then compared with the low-level line feature outputs in a simple rule based model. The output of the model depicts major and minor roads within St. Louis.

The output was compared with truth data to obtain quantitative accuracy levels. The truth data was digitized from the panchromatic airphotograph of the study area. The airphotograph was manually digitized and rectified to the UTM coordinate system of the TM image. The final digitized truth data were then registered and overlaid onto the results of the high-level road extraction algorithm. This overlay provided a pixel based accuracy metric for the road extraction results. This showed accuracy levels of 96% for major roads within the control area and 91% for secondary roads. The omission of most road pixels appears to be due to the poor resolution of the Landsat TM data and because of the time of year the St. Louis image was taken. The TM image was taken in early September, when the trees were still fully canopied. The truth data photograph was taken early the following April when the canopies were minimal. The airphotograph reveals that many of the roads omitted by the road detection algorithm had trees over them and would have been obscured by foliage in the September Landsat imagery.

4. Conclusion

Our research has proven the basic approach taken in image understanding. The neural network land cover classification algorithm has been effectively demonstrated in general land cover classification and specifically in classification of material associated with cultural features. The image processing based low-level feature extraction technique has been proven effective and efficient in extracting linear features. The expert system integrated high-level feature extraction algorithm proves to be more effective than traditional computerized classification techniques and as effective as manual image interpretation. PRC intends to continue with research and development efforts aimed at producing a prototype image understanding system based on the approach defined in this report.

Bibliography

BISC92 Bischof, H., Schneider, W., and Pinz, A. "Multispectral Classification of Landsat-Images Using Neural Networks." IEEE Transactions on Geoscience and Remote Sensing, May 1992, pp 482-490.

CARP92 Carpenter, Gail and Grossberg, Stephen. "A Self-Organized Neural Network For Supervised Learning, Recognition, and Prediction."IEEE Communications, September 1992, pp 38-49.

FUA87 Fua, Pascal & Hanson, Andrew. "Using Generic Geometric Models or Intelligent Shape Extraction." Proceedings: DARPA Image Understanding Workshop, February 1987, pp 227-241.

HARV87 Harvey, Wilson, Diamond, Matthew & McKeown, David Jr. "Tools for Acquiring Spatial and Functional Knowledge in Aerial Image Analysis." Proceedings: DARPA Image Understanding Workshop, February 1987, pp

857-873.

HARW87 Harwood, David. "Interpretation Aerial Photographs by Segmentation and Search." Proceedings: DARPA Image Understanding Workshop, February 1987, pp 507-520.

HUER87 Huertas, A., Cole, W. & Nevatia, R. "Detecting Runways in Aerial Images." Proceedings: DARPA Image Understanding Workshop, February 1987, pp 273-339.

JONE87 Jones, William and Hoskins, Josiah. "Back-Propagation: A Generalized Data Learning Rule." Byte, October 1987, pp 155-162.

NEUR91 "NeuralWorks Professional II/Plus and NeuralWorks Explorer Neural Computing." Pittsburgh: NeuralWare, Inc., 1991.

PERL89 Perlant, Frederic and McKeown, David. "Scene Registration in Aerial Image Analysis." Proceedings: DARPA Image Understanding Workshop, May 1989, pp309-331.

THOM87 Thompson, D. & Mundy, J. "Model-Directed Object Recognition on the Connection Machine." Proceedings: DARPA Image Understanding Workshop, February 1987, pp98-106.

TOML90 Tomlin, C. Dana. Geographic Information Systems and Cartographic Modeling. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1990.

TON89 Ton, Jezching. "A Knowledge Based Approach to Landsat Image Interpretation." Doctoral Dissertation, Michigan State University: Department of Computer Science, 1989.