Individual Tree Identification from High Resolution MEIS Images

François A. Gougeon
Petawawa National Forestry Institute
fgougeon@pnfi.forestry.ca

Abstract

Relative to conventional aerial photographs, MEIS images make better use of the colour spectrum and can easily be registered to geographic coordinates. In the short term, this enables on-screen computer-aided image interpretation and compatibility with geographical information systems. In the long term, there is the potential to leave most or all of the interpretation task to computers. If this is to be successful with high resolution images (0.1-1m/pixel), we have to achieve the automatic isolation of individual trees from each other and from the background vegetation, followed by a tree by tree species identification. The trees' geographic positions, crown areas, and heights would also be available to create detailed forest inventories. This paper describes on-going work at PNFI on the computer analysis of medium (1-10m/pixel) and high spatial resolution MEIS images, shows preliminary results, and discusses future avenues such as image understanding.

Identification des arbres à partir d'images MEIS à haute résolution

François A. Gougeon
Institut forestier national de Petawawa
fgougeon@pnfi.forestry.ca

Résumé

Par rapport aux photographies aériennes conventionnelles, les images MEIS font une meilleure utilisation du spectre des couleurs et peuvent facilement être alignées à des coordonnées géographiques. À court terme, ceci facilite une interprétation à l'écran, assistée par ordinateur, et une compatibilité avec les systèmes d'information géographique. À long terme, il pourrait être possible de laisser la majorité ou toute cette tâche d'interprétation aux ordinateurs. Pour réussir cela à partir d'images à haute résolution (0.1-1m/pixel), nous devons parvenir à un isolement automatique des arbres entre eux et par rapport à la végétation environnante, suivie d'une identification à l'arbre près de leurs espèces. La position géographique, la superficie de leur couronne, et la hauteur des arbres seront aussi disponibles pour créer des inventaires détaillés de la forêt. Cet article décrit les travaux en cours à l'IFNP sur l'analyse des images MEIS de moyenne (1-10m/pixel) et haute résolution spatiale, présente des résultats préliminaires, et discute de directions futures telle la compréhension des images.

Pages 117-128 in Leckie, D.G.; Gillis, M.D., eds. Intl. Forum on Airborne Multispectral Scanning for Forestry and Mapping (with Emphasis on MEIS), Val-Morin, Québec, Canada, April 13-16, 1992. Inf. Rep. PI-X-113, Forestry Canada, Petawawa Natl. For. Inst.

Individual Tree Identification from High Resolution MEIS Images

François A. Gougeon
Petawawa National Forestry Institute
fgougeon@pnfi.forestry.ca

Introduction

The Multispectral Electro-optical Imaging Sensor (MEIS) is a Canadian remote sensing system capable of acquiring high quality digital images from an aircraft. These images make a more complete use of the colour spectrum than aerial photographs and can be automatically registered to geographic coordinates. Using standardized image enhancements and interpretation keys, they lend themselves well to on-screen computer-aided image interpretation. An interpreter workstation with a link to a geographical information system could facilitate the production or updates of a forest inventory (see Figure 1).

Some of the short term objectives of MEIS research at the Petawawa National Forestry Institute (PNFI) are to develop the tools and methodology needed by this new approach and to compare it with conventional photo interpretation and inventory production techniques. However, in the long term and in view of the increasingly precise information needed for forest resource management, automatic ways of gathering forest details will be needed. To do this, it is necessary to isolate individual trees from one another and from the background vegetation, followed by tree by tree species identification. A tree's geographic position and its crown area are thus immediately available. Tree height can be obtained because stereo pairs can be specified as part of MEIS image sets. This may lead to forest inventories on a tree by tree basis, either as a goal in itself, or as a necessary step towards accurate forest stand inventories.

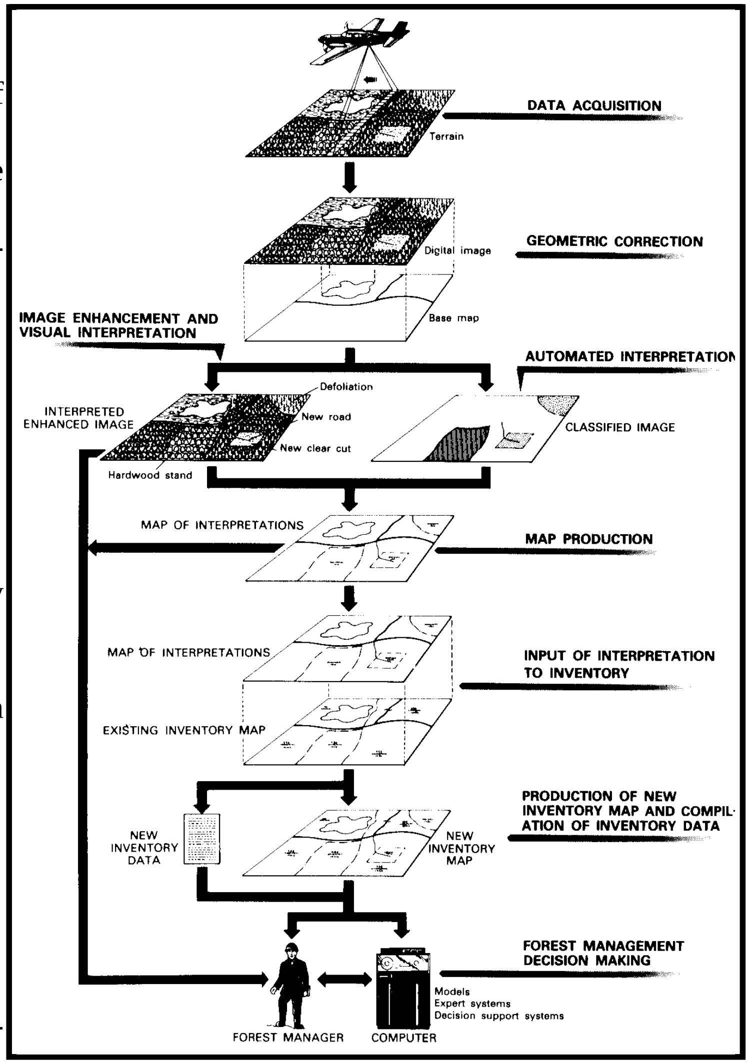

Figure 1 - MEIS-based forest inventory concept (after Leckie 1990)

The automatic isolation and identification of trees in high spatial resolution MEIS images (30-70 cm/pixel) constitute the main subjects of this paper. However, because digital remote sensing is still usually associated with low resolution satellite imagery (10-80 m/pixel), it is worth reviewing some of the problems that are encountered in image analysis as spatial resolution gets higher and higher. Different spatial resolutions usually imply different statistical views of the data and of the entities to recognize; they require different approaches to the image analysis problem. Results of medium resolution (1-10 m/pixel) image classifications and segmentation will be given. Also, as the need to create new image analysis methods can lead to using an image understanding paradigm, some work in aerial image understanding is briefly reviewed.

Aerial Image Understanding

Image processing tools are often defined as those which produce improved images from the multispectral images acquired by remote sensing. Filtering and enhancement processes are examples of such tools. Image analysis tools produce thematic results, which may be stored in image form, creating a description of the input image in terms of classes of interest. Image classifications or image segmentation processes followed by segment identification are examples of such tools. Image understanding processes or systems attempt to create a complete description of the physical world depicted by the remotely sensed images (see Figure 2). Historically, most of the work in image understanding has involved recognizing things in known, human-made, environments. However, the complexity found in high resolution MEIS images of forests, as well as the weaknesses encountered with image analysis approaches, may force us to use an image understanding paradigm.

The image understanding systems most applicable to our work are those that have dealt with extracting information from digitized aerial photographs. Let us briefly outline the most interesting ones and review some of their salient features. A system developed at the University of Kyoto, Japan (Nagao and Matsuyama 1980), analyses digitized aerial photographs of suburban areas. It first segments the image on the basis of colour homogeneity, then compiles various characteristics for each segment, and finally analyses segment relationships using a rule-based expert system (ES). The ES uses a "blackboard architecture", in which partial results of various processes are posted and decisions as to the existence and nature of various objects can be taken. For example, an object such as a house is usually made up of various segments in a particular arrangement. This system is typical of a bottom-up approach to image understanding, where the information extraction process is carried out from the detection of simple features in an image, to their agglomeration into higher level features, to the interpretation of these features in terms of known objects, without reference to an explicit model of the world (except as hard-coded in the expert rules) and without the capability of going back to the image for further processing or analysis.

In contrast, the Image Reasoning Program (Selfridge and Sloan 1981) developed at the University of Rochester uses a top-down approach to image understanding. An Appearance Model Expert, tapping into an Appearance Model of the world, decides on the main recognition goals to be achieved and passes requests, in the form of subgoals, to an Operator Expert which has access to image processing operators. The Operator Expert decides which operators to use in order to accomplish one of these subgoals (its main goal for the moment) and dynamically modifies or tunes various parameters of the image processing operators until this goal is achieved. Of course, partial results are always fed back to the higher levels but the control of the image understanding process is inherently driven from the top.

The ACRONYM system (Binford et al. 1980) was developed at MIT to recognize from aerial photos various types of airplanes parked at an airport. It uses an intermediate approach where information from low level image processing is agglomerated to an intermediate state and compared with information brought down from higher level, model-based processes. More precisely, the bottom-up work consists of extracting edges and line segments from the image and relating them to each other to get ribbon-like features. The top-down work starts with a world model that describes airplane parts and airport features in terms of groups of 3-D cylinders and uses the known position of the camera to generate various possible ribbon configurations depending on possible plane types. The intermediate level tries to match the possible ribbon configurations with the actual ribbon configuration found in the image. Still, the ACRONYM system is basically using top-down control of the image understanding process; it is model-driven rather than data-driven.

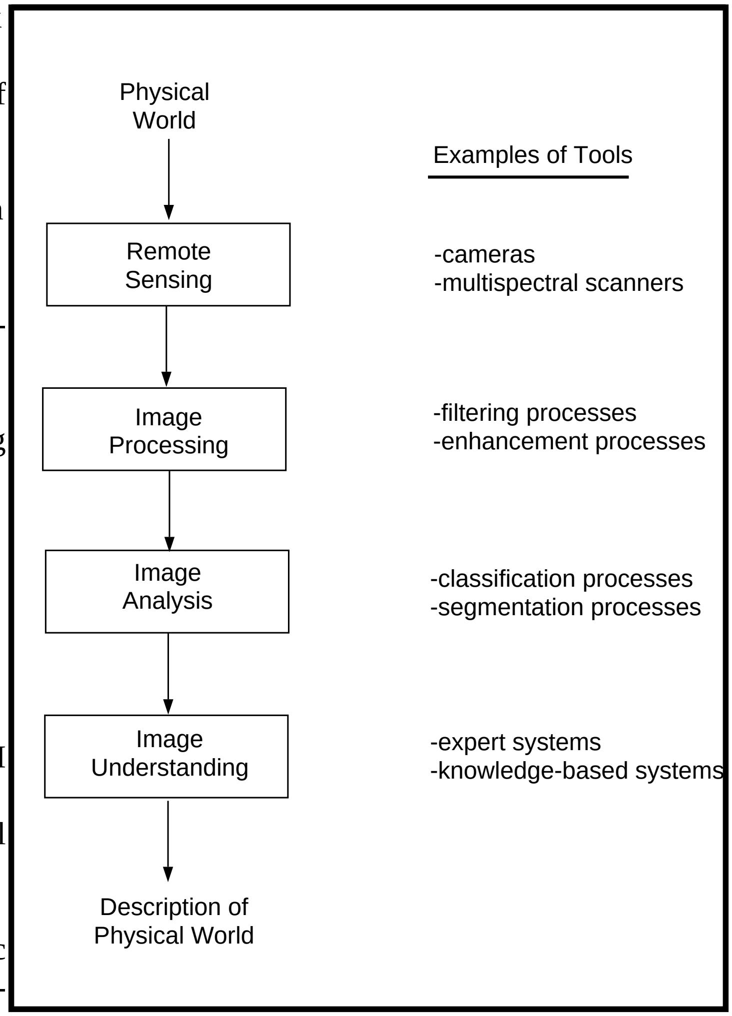

Figure 2 - The fields and tools of computer vision

The SIGMA system from Tohoku University, Japan (Matsuyama 1987), is a second generation system built from the experiences acquired at the University of Kyoto and the knowledge that an ACRONYM -type world model can be a definite advantage. Again, it tries to understand aerial photos of suburban areas. It combines three expert systems: the low level vision expert, the model selection expert, and the geometric reasoning expert. It first uses a bottom-up approach to recognize the obvious objects in the scene and connects them with appropriate models. For example, the model of a typical suburban house has a front lawn, a backyard, a driveway, etc. It then uses the models as guides to what possible related objects could be found in the image and returns to the image with more precise and more localized operators, making various attempts to find these missing features. For example, a house and a front lawn were found initially, but the model expects a driveway to be also present, so efforts will be made to find such a driveway. These attempts to find missing features are made using various image processing operators and the system accumulates evidences from all the partial results as to the presence or absence of the desired feature (e.g., the driveway). This system is thus both data-driven and model-driven.

Another system mixing bottom-up and top-down image understanding control was developed in Germany (Nicolin and Gabler 1987). The bottom-up process, which uses a multi-resolution pyramid of data and region- and edge-based segmentation algorithms, generates "clues"of what could be found in the image. The top-down process, which uses structural analysis to attempt to group segments, generates "expectations". These clues and expectations are used by the bidirectional control mechanism to generate hypotheses and test them, and thus accumulate evidence.

Over several years, a group at Carnegie Mellon University has developed different image understanding systems grouped under the umbrella name of one of their early system: MAPS - the Map Assisted Photointerpretation System (McKeown 1987). Most of these systems deal with images in the context of providing information to Geographic Information Systems (GISs) or of providing information in a synergistic way with GISs. They differ in conceptual approaches or in the artificial intelligence tools being used (e.g., their resolution mechanism or their knowledge representation). For the more recent systems, the original MAPS is seen as a collection of databases (terrain, map, landmark, and image), each with sophisticated access methods. One system of particular interest is the rule-based expert system named SPAM - System for Photointerpretation of Airports using MAPS. It uses region-based, texture-based, and depth-based segmentation and other feature extraction processes, in combination with knowledge about airports and the information provided by MAPS, to identify runways and buildings in aerial images of the Washington airport. Some of the most recent work of this group deals with combining multispectral and panchromatic stereo imagery for cartographic feature extraction of urban areas. In general terms, it can be viewed as information extraction and image understanding by fusing data from various sources, each providing a different perspective on reality (see paper by D. McKeown, this volume).

The image understanding (and/or analysis) research closest to mine is the doctoral and postdoctoral work of A. Pinz in Austria (Pinz 1989, Pinz and Bischof 1990). The goals are the same: the isolation and identification of trees in high resolution data acquired from airplanes. Pinz uses an expert system which controls a database of image analysis routines and a multilevel, multiparadigm representation of the forestry environment (i.e. using various AI knowledge representation schemes) to isolate individual trees. The trees are located by finding local maxima in the image (possibly tree tops) and examining how the spectral values fall off from this assumed tree center. Tree species identification (and, possibly in future, their damage levels) is obtained by classifications using feedforward backpropagation neural networks, a new classification approach popular in the artificial intelligence community. Pinz also succeeded in "crossing" various neural networks together using an approach that could be described as "neuron transplantations". The results are very encouraging, with 85-90% accuracy separating five species (spruce, beech, fir, pine, and larch).

Problems with High Resolution Image Analysis

In the present paper, high resolution images are defined as being made of multispectral data acquired from an airborne sensor for which the spatial resolution of each pixel at ground level is of the order of 0.1-1 m/pixel. Similarly, medium resolution images imply 1-10 m/pixel and low resolution images, typically acquired by satellites, 10-100 m/pixel.

At this point in time the better known image analysis techniques in digital remote sensing have been developed for satellite images. For example, a typical maximum likelihood classifier classifies images (with or without training) on a pixel by pixel basis, comparing the multispectral value of a given pixel with that of various classes represented by their mean multispectral values and their covariance matrices (an indication of the spread of values around the mean values). The pixel is assigned to the class to which its has the highest likelihood of belonging. The two major drawbacks of this approach are that: a) unless the multispectral data are augmented by derived textural features (e.g., by co-occurence or variogram approaches), a pixel-based classifier does not take neighboring pixels into consideration when making its decisions; and b), the pixel values forming a class are assumed to be "normally" distributed around their mean (i.e. forming a Gaussian, normal, or bell-shaped distribution). Although these constraints are reasonable for the analysis of low resolution satellite images, they make the classifier less and less effective as the spatial resolution of the data gets higher.

In the medium resolution range the texture within a forest stand, which is readily visible, can become as important as the spectral or colour component in recognizing that stand. So methods that take into consideration neighborhoods or areas in making their decisions (rather than single pixels) are usually more effective. In addition, the presence of distinctive sunlit and shaded areas within forest stands makes their class signatures (mean vectors and covariance matrices) inherently skewed or bimodal, or normal with such large spreads and with significant overlaps between each class that classifiers based on a normality assumption are ineffective (see Figure 3). Fortunately, in addition to the development of new classification methods, there are ways to make these old standby classifiers more effective.

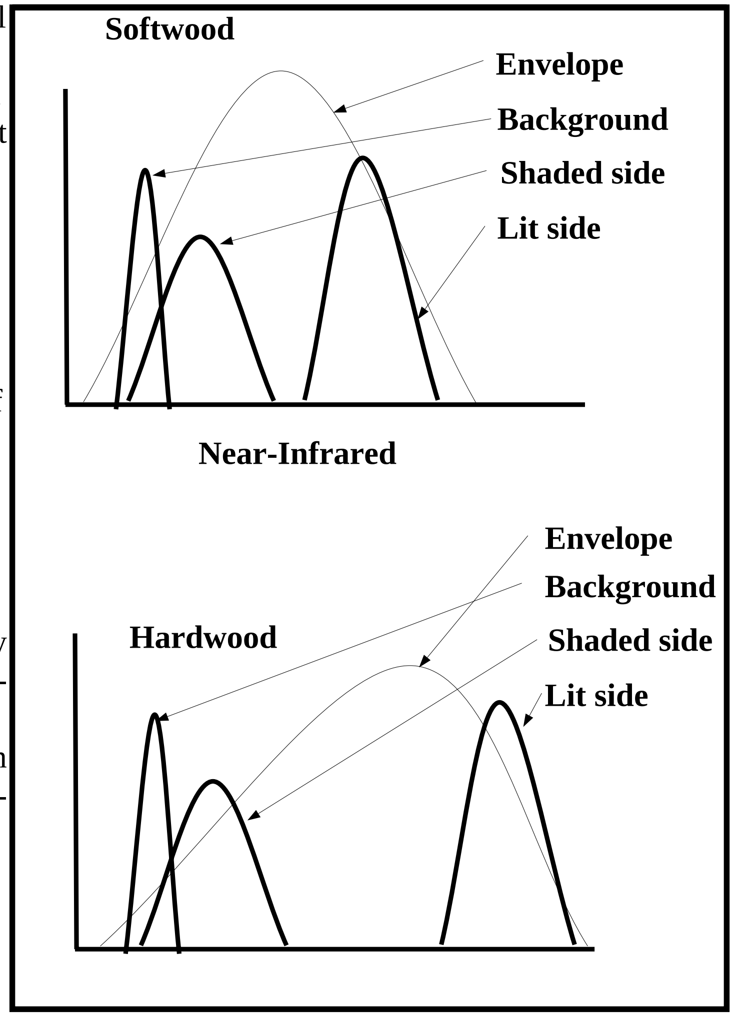

Figure 3 - When inherently multimodal distributions (sunlit and shaded regions) are assumed unimodal by a classifier, class signatures (hardwood vs softwood here) become very wide and hard to differentiate and the pixel-based classifier becomes ineffective.

In high resolution images, the problem is compounded by the fact that we may not even be dealing with the recognition of the same entities as in the low and medium resolutions. The recognition of forest stands can potentially be replaced by the recognition of individual trees, possibly followed later by their regrouping into stands. The individual trees themselves may exhibit a lot of texture and structure within their crown, correlated or not with the texture and structure exhibited by the entire stand (and with view and sun angles). New image analysis and even probably image understanding methods need to be developed for images of this type.

Possible Approaches for Medium Resolution MEIS Images

a) Ways to Still Make Use of Maximum Likelihood Pixel-based Classifiers

Because most commercial image analysis software packages contain a pixel-based maximum likelihood classification process and because the digital remote sensing community is familiar with this type of classifier, having used it extensively to analyse satellite images, it was felt worth investigating ways to make use of this classifier with medium resolution MEIS data. The solution resides in finding ways to eliminate factors that cause the high variability in spectral signatures. With the realization that there is no need to classify every pixel in an image if tree species is our only concern, two approaches were tested , both based on masking part of the image before the classification process (Gougeon and Moore, 1988).

In the first approach, a mask was used that allowed only the pixels corresponding to the lit side of trees to reach the classification process. This eliminated pixels on the shaded side of the trees, as well as the pixels representing the understorey or the ground that were in the shadow of dominant trees. In the second approach, the mask was further refined to allow only the brightest pixel of each tree to reach the classification process. In the case of coniferous trees, this brightest pixel generally corresponds to a location near the top of the tree; that section of the tree is the most directly lit and is less likely to contain shaded areas. This approach was thus named the "tree top" classification approach.

These two approaches, as well as a conventional full image (unmasked) pixel by pixel classification, were tried on a 1.2 m/pixel MEIS image of forest plantations. Table 1 shows the classification accuracies for four coniferous species, while Figure 4 shows on a tree-by-tree basis the results of the tree top classifier.

b) Image Segmentation by Region Growing

At medium resolution, approaches that take texture directly into account have more potential than pixel-based maximum-likelihood classifiers or modifications thereof. That is true even when some indirect texture measurements are introduced to that classifier (Teillet et al. 1981). One such

Table 1 - Supervised classification results of medium resolution MEIS data.

Class	Unmasked	Lit side mask	Tree top mask
Red Pine	52%	54%	63%
Red Spruce	60%	52%	77%
White Pine	35%	48%	56%
White Spruce	42%	59%	65%
Average accuracy	47%	53%	65%

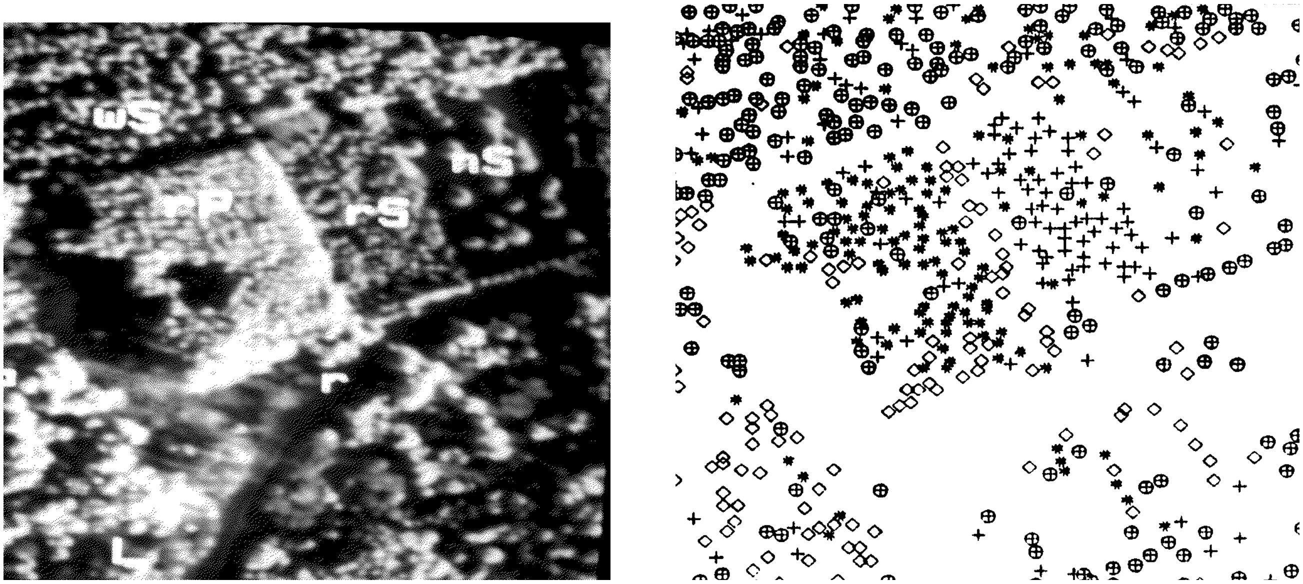

Figure 4 - MEIS image of forest plantations at 1.2m/pixel and results of the tree top classification approach
(∗ = Red Pine, + = Red Spruce, ◊ = White Pine and ⊕ = White Spruce)

approach is to segment the image into homogeneous regions (presumably forest stands) and then to recognize or identify these regions. A recent example of a texture-based segmentation process is that of St-Onge and Cavayas (1991). Another, combining more precise multispectral and textural signatures in a region growing segmentation process, is that of Gougeon and Wong (1986).

A region growing segmentation process starts by finding core areas on the image that are typical of potentially larger regions, extracts their parameters, and tries to grow these core areas so that they fill the whole region of which they are typical. Parameters control the growth so that, generally, it stops after encountering areas of a different nature. For example, from a core area typical of a particular forest stand, a region should grow to encompass all of the stand. On medium resolution MEIS images, such a process should deal directly with the significant texture present and preferably have precise multispectral signatures and metrics that do not make any assumption as to the distribution of pixel values (i.e., no normality assumption). Figure 5 shows the results obtained on a 1.2 m/pixel MEIS image of forest plantations at PNFI using the segmentation procedure of Gougeon and Wong (1986).

Once such an image has been segmented into spectrally and texturally homogeneous regions, there is still a need to identify the exact content of these regions, based on the regions' spectral and textural signatures. This identification process can be affected by parameters such as sensor response variations, atmospheric conditions, sun elevation and azimuth, sensor view angle, topography, forest stand density, and tree shapes and sizes. A study was done to see if some of these effects could recognized, quantified and compensated for in the spectral and textural signatures (Gougeon and Wong, 1987). An analysis method was developed and partial results are promising. It led to the concept of using a knowledge-based system to identify regions resulting from the segmentation process (see Figure 6).

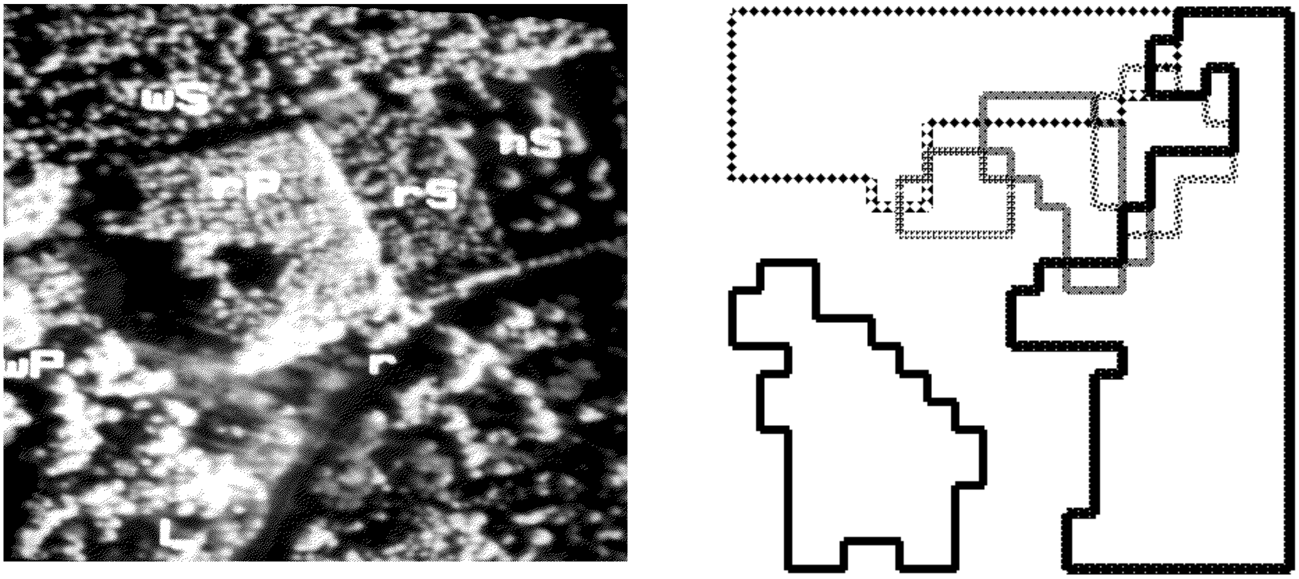

Figure 5 - MEIS image of forest plantations at 1.2 m/pixel and results of a multispectral and textural segmentation

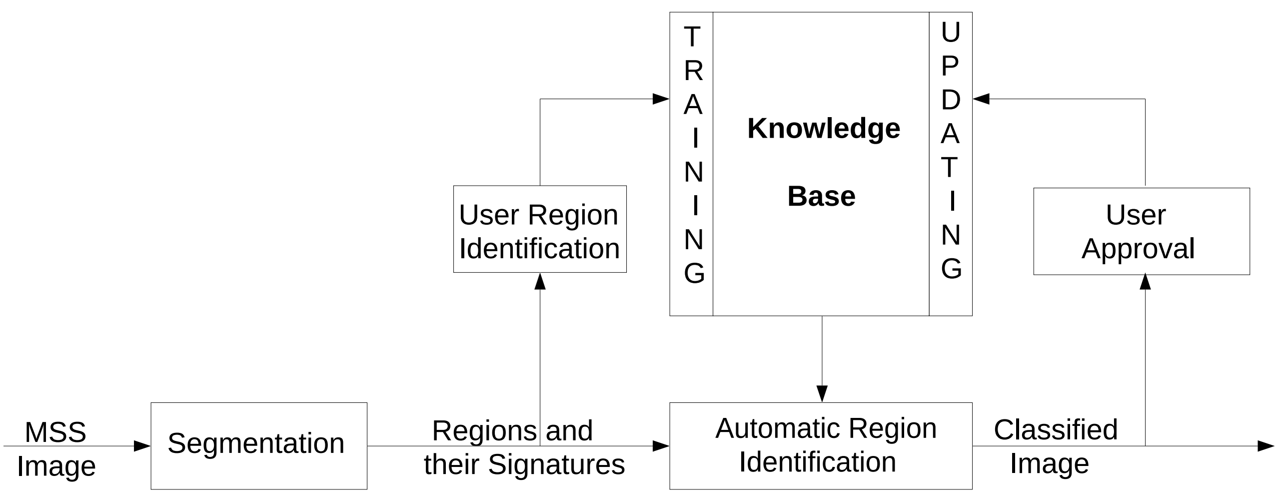

Figure 6 - A possible knowledge-based segment identification system

Ongoing Research Work on High Resolution MEIS Images

a) Individual Tree Isolation

High resolution images, with their detailed look at the forest, force us to think in terms of individual trees and the data within them and around them rather than consider forest stands. Individual trees are the logical entities (or objects) for the computer to analyze in such images, and if we were to identify these trees, we could always recreate the desired forest stands later. The first step towards identifying individual trees is to be able to isolate them from the background vegetation and from each other. This is difficult even when there is substantial shadow.

Part of the difficulty is that, even though the shaded areas may help to isolate the lit parts of trees, one is often left with tree clusters and groupings where the end of one tree and the beginning another is hard to assess. Also, even if a given tree is distinct enough, its crown area (an important forestry measurement) may be difficult to ascertain. Parts of the tree itself may be shaded and other vegetation may lie in the tree’s shadow. Distinctiveness will suffer.

Over the past three years, various approaches have been tried to isolate individual trees. Given a certain viewpoint and sun elevation, it is possible to assume a well defined shape for conical coniferous tree species (e.g., a circular or crescent shape). However, tests done, assuming such a shape and using a generalized Hough transform (Ballard 1981) to move the data into a parameter space suited to recognize that shape, have not been successful. A morphological approach (Mouchot et al. 1989), typically more independent of fixed shapes and relying on eroding and redilating image entities, also had poor success. Various neural network approaches (Zeidenberg 1990) are still being investigated in collaboration with l’Université du Québec à Hull. A "valley following" technique, succeeded by heuristic rules to complete the isolation process, is showing great potential. Its description follows.

The technique is conveniently described by a geographical analogy. At a conceptual level, trees, as depicted in a grey level image, can be visualized as mountains (i.e., clumps of pixels with higher values than their surrounding). The technique consists of following the valleys that invariably exist between them. During a first pass through the image, "potholes" or "lakes" are located. They correspond to points that are local minima. In subsequent passes, starting from these potholes, valleys are followed upstream in an incremental fashion to the highest point of the valley (hopefully without going up any mountain), and then, downstream through the adjacent valley to a neighbouring pothole. In all of these iterations, the continuation of a valley, as defined at the pixel level, is the next pixel bordered on each side by pixels of higher value. In other words, we are looking for the next pixel up or down in a typical V-shaped valley. To save time and avoid confusion, large, "almost flat", "sea level plateaus" were previously eliminated from this process by a simple pre-classification. The "valley following" technique isolates individual trees rather well (at least visually), as seen in Figure 7. However, a careful examination of Figure 7 will reveal that in a lot of cases the trees are not completely surrounded by valleys. This problem is alleviated by creating various rules describing how and when to possibly close a tree contour. This, of course, is not a trivial problem. The creation of these heuristic rules (i.e., rules of thumb) constitute the bulk of the individual tree isolation endeavour and will be reported in an upcoming article.

At this point in time, the approach is deemed to be very promising. However, it has only been

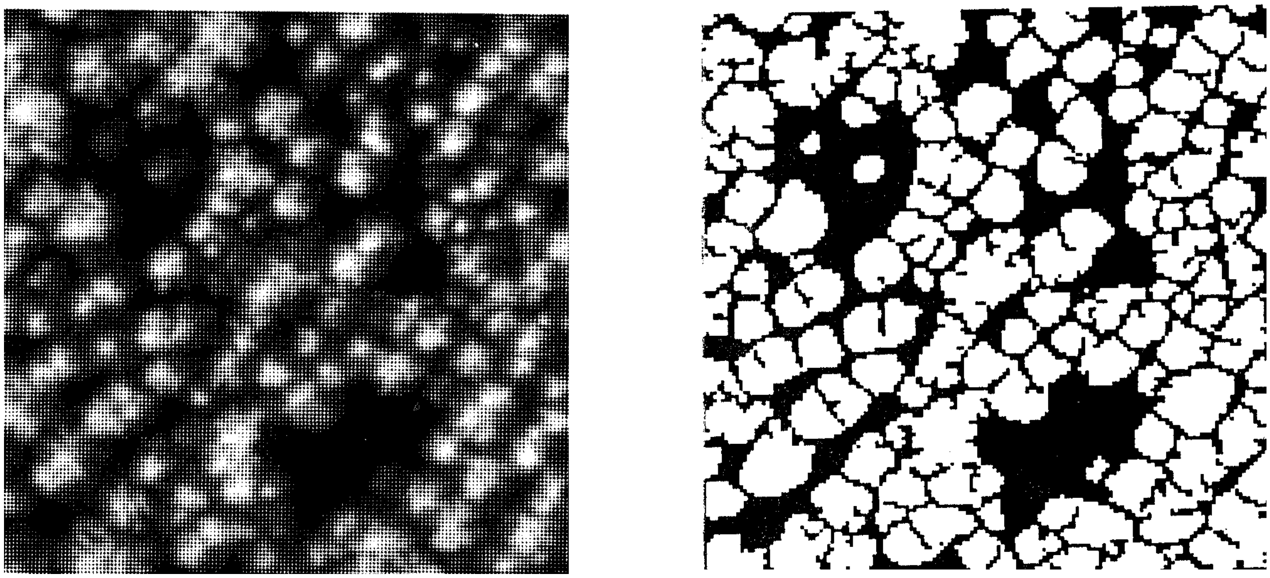

Figure 7 - White Spruces from PNFI's Hudson Plantation and the preliminary results of the "Valley Following" tree isolation process (with 31 cm/pixel data from the near-infrared channel)

tried on a MEIS image of a coniferous forest, and not on a hardwood or mixed forest. In addition, the imagery was taken in late fall where low sun elevation favours the production of shadows and shaded areas. This approach would no doubt fail with a rather flat canopy. Its success in hardwood forests is uncertain, but will be tested soon.

b) Various Approaches to Multispectral Classification of Individual Trees

As mentioned before, high resolution MEIS images (from low altitude flights) force us to think in terms of individual tree isolation and identification, rather than forest stand delineation. Similarly, it is more appropriate to think in terms of object-based (i.e., based on the trees) rather than pixel-based classification processes to identify species. This implies the creation of new classification techniques, whether based on the multispectral characteristics of the trees or on other characteristics such as texture, structure, and context.

In the past few years various approaches (some traditional, some novel) have been examined to facilitate the multispectral classification of individual trees. They differ mainly in the way that signatures are acquired from the trees. These techniques and their signature definitions are listed in Table 2. The first three classification schemes rely on only one multispectral vector per tree (i.e., one number per channel) to make their decisions. This follows the concept that representing the spectral characteristics of a tree by a single value per channel is sufficient to classify that tree. These methods do not take advantage of the fact that the tree itself may cover an area of 25-150 pixels. The other four classification schemes are variations of the "Tree Colour Line " approach. It is based on the idea that, independent of the illumination situation (i.e., lit or shaded tree sides), a tree species has a certain ubiquitous colour and certain light penetration properties related to its structure. The spread of a single tree's pixels in multispectral space can be described at a first level of approximation by a line through that space. This multidimensional line can be described by its slope and intercept in its various two dimensional projections (as in "si_sign", in Table 2) or by a multidimensional line, obtained by finding the principal component of the data, anchored by the mean of the data (as in "mpc1_sign"). At a

Tree Average (ave_sign):

The mean multispectral vector of all pixels contained in the tree.

Average Lit Side (lit_sign):

The mean multispectral vector of all pixels found on the lit side of the tree. For the sake of simplicity, the lit side is defined as all pixels with values in the near-infrared band above the mean value of that band.

Tree Top (tt_sign):

The multispectral vector of the most brilliant pixel in the near infrared channel, often corresponding to an area near the tree top of conifers.

Slope & Intercept (si_sign):

Based on the "Tree Colour Line" approach , these signatures consist of the slopes and intercepts of the tree colour line as all the bands are compared with the near infrared band one by one (sometimes refered to as the "Multiple 2D Tree Colour Line" approach).

Mean vector & first principal component (mpc1_sign):

The "Tree Colour Line" approach in multiple dimensions - the distribution of pixels for each tree is represented by its first eigenvector (its colour line direction) and its mean vector (to anchor the line in space).

Mean vector & first principal component & eigenvalues (mpc1ev_sign):

As above, but eigenvalues are added to describe the spread of the distribution in various directions.

Mean and covariance matrix (mcov_sign):

The mean and covariance matrix of the distribution of pixels in multispectral space for each tree.

Table 2 - The various multispectral classification schemes
and their signature definitions.

Signatures	Classifier	Average Overall Accuracy		
ave_sign	ML	67	(61.3 - 75.5)	(12 sets)
lit_sign	ML	**73.5**	**(70.2 - 79.8)**	(10 sets)
tt_sign	ML	62.1	(52.1 - 66.7)	(10 sets)
si_sign	ML	66.7	(51.2 - 70.3)	(10 sets)
mpc1_sign	ML	63.7	(54.7 - 75.0)	(10 sets)
mpc1ev_sign	ML	63.2	(55.2 - 69.6)	(10 sets)
mcov_ca_sign	ML	64.3	(56.5 - 69.7)	(10 sets)
ave_sign	NN (HN=5)	70.0	(66.7 - 72.8)	(5 sets)
lit_sign	NN (HN=5)	72.5	(70.4 - 74.5)	(5 sets)
tt_sign	NN (HN=5)	67.4	(63.8 - 71.1)	(5 sets)
si_sign	NN (HN=5)	62.9	(60.7 - 70.3)	(5 sets)
mcov_sign	NN (HN=15)	65.0	(56.6 - 77.8)	(5 sets)

Table 3 - Comparison of various multispectral classification schemes
for trees individually delineated on a high resolution (36 cm) MEIS image.

second level of approximation, the spread of a single tree's pixels in multispectral space can be described as taking the form of a cigar. This cigar, the spread of the data around the tree's colour line, can be parametrized by its eigenvalues (as in "mpc1ev_sign"), or by its covariance matrix (as in "mcov_sign"). The latter is theoretically more precise, but it implies the availability of more data to be properly defined. In small trees, it is not obvious that there is always enough pixels to define that matrix with enough precision for species separation. A detailed description of these signatures is beyond the scope of this paper.

Most of these tree signatures were tested with both maximum likelihood and feedforward backpropagation neural network classifiers. From five species, 201 trees randomly split with equal probability between training groups and testing groups were used. Because of the low number of trees available, which led to classification replication instabilities, most experiments were repeated from five to 12 times with randomly selected tree groups. Preliminary results in the form of average overall accuracies are shown in Table 3.

These preliminary results can be interpreted in at least two ways. On the positive side, it is reassuring to know that there are quite a few possible ways of getting classifications of high resolution MEIS images. In addition, it is encouraging that some of these procedures are easily implemented on typical commercial image analysis systems (i.e., Table 3, first section). These procedures are faring just as well, if not better, than the new algorithms designed specifically for tree-based classification (i.e., Table 3, second section). Moreover, resorting to neural networks may not be necessary (Table 3, third section). On the negative side, not one of these methods is giving great results, even though we are dealing with only five species, albeit five very spectrally close species. Although not shown here, it should also be noted that one particular species, white pine, either due to its nature or to the lack of a sufficient number of training trees, has been bringing these accuracy figures substantially down. More research is needed and various other forestry situations should also be examined. However, one should keep in mind that in any classification into classes with subtle differences, complete accuracy is impossible. Obtaining accuracies of the order of 75-85% would be a reasonable goal, especially considering conventional aerial photo interpretation accuracies.

The interpretation of the low accuracies registered here (50-80%) and of the rather large variations encountered as experiments are repeated with different randomly selected trees, should also be tempered by the fact that these are preliminary results, obtained with what appears to be a less than sufficient number of trees. Additional trees are in the process of being added to these experiments. Tests similar to that of a production environment will be possible when the tree isolation programme is completed and can directly feed the classifiers.

Additional experiments are underway with various textural parameters and their synergistic effects when added to the multispectral signatures. At this spatial resolution, texture within a single tree crown can be considered to be a manifestation of the interaction between the tree's physical structure and the lighting situation. However, much of the existing methods of describing texture may not be appropriate due to the lack of significant statistics available from within a single tree crown. Future experiments will deal more directly with various structural parameters, and their potential synergistic effects with multispectral and textural parameters. As the number of parameters to combine in making a decision about tree species identity increases, neural net classifiers, capable of making decisions in more complex decision spaces, will have to replace the ubiquitous maximum likelihood classifier. Ultimately, though, if many multispectral, textural and structural approaches are developed, combining them in the hope of getting a synergistic effect will start to be counterproductive. A more direct way of selecting the appropriate and necessary parameters from the different domains will be needed. This leads us to the idea of possibly using an expert system or a knowledge-based system to make these decisions on a case by case (i.e. tree by tree) basis.

Towards Automatic Interpretation of Aerial Images

Although an image analysis paradigm may restrict us to think in terms of the synergy among properly weighted signatures gathered in different domains (e.g., multispectral, textural, structural, and contextual), an image understanding paradigm may provide us with more freedom to select the proper tools as needed and accumulate differentiating evidence gradually in a logical fashion. It is a well known fact that, given a certain number of classes of interest, image analysis tools and domains will differ in their ability to separate them. Given the knowledge (acquired by experimentation) of which tools are more useful for each case and given a prior automatic isolation of the individual trees in a MEIS image, it should be possible to create an expert system that will attempt to identify a given tree species on a tree by tree basis using a selected subset of these tools. Such a system is depicted in Figure 8.

With the addition of parallax information from MEIS stereo pairs, this system leads to an individual tree inventory, where tree positions, species, crown areas, and heights are tabulated. Although an individual tree inventory "per se" may not be what is needed, I believe that it may be a necessary intermediate step in obtaining accurate forest inventories. The individual tree-based information produced in such a way can always be recombined into the desired forest stands afterward. In fact, that may help solve the general problem of differing inventory needs for different applications. Although it is not obvious how recombination will take place, some researchers are already addressing the issue (Eldridge N.R., pers. comm.). A possible approach might involve making decisions on stand boundaries using various image segmentation tools at different lower spatial resolution levels of a multi-resolution pyramid. In any

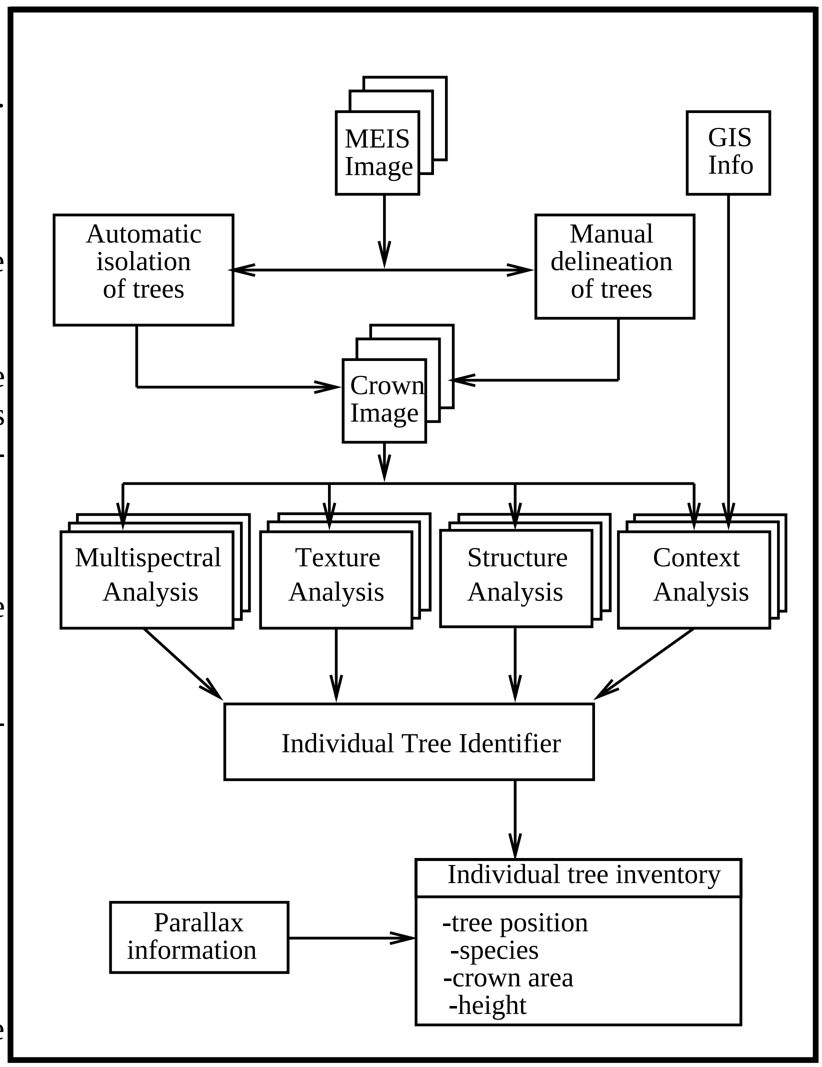

Figure 8 - Individual tree inventory from high resolution MEIS images

case, keep in mind that this individual tree inventory would be acquired almost automatically (i.e., it is not labour intensive) and need not be stored as such if not needed. Additionally, you can rest assured that if an individual tree inventory is available as an intermediate product, it will eventually be put to good use.

Conclusion

In the short term, most of the research work in the digital remote sensing project at PNFI is directed towards solving problems posed by the computer-assisted interpretation of standardized MEIS image enhancements. Essentially, the aim is an improved, computerized version of what is currently being done with conventional aerial photos. This approach makes good use of some of the advantages of the digital imagery produced by the MEIS sensor, namely: the capability to register these images precisely to geographic coordinates, thus to facilitate the interaction with Geographic Information Systems; the better use of colour information due to the multispectral nature of these images; and, the possibility of an automatic interpretation of things such as lakes and roads. This should already provide substantial improvements over conventional ways of doing forest inventories and be an effective way of introducing this new technology.

On the other hand, for digital MEIS data to fulfill its full potential, more automation of the interpretation process is needed. Longer term research towards the complete automation of the inventory process is also being undertaken at PNFI. This automation (or semi-automation) will most likely be achieved by making use of various image analysis tools to identify species on an individual tree basis. The proper path is still unclear, but seems to fall along the lines of either combining summarily the information provided by all of these tools, or selecting the appropriate tools as needed, in an ongoing refinement process, using an expert system. In any case, because of the differences from satellite imagery, a variety of new image analysis tools for medium and high resolution MEIS data needs to be developed and tested.

References

Ballard, D.H. 1981. Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition. 13(2): 111-122.

Binford, T.O.; Brooks, R.A.; Lowe, D.G. 1980. Image understanding via geometric models. Pages 364-369 in Proc. 5th Int. Conf. Pattern Recognition, Atlantic City, N.J. Dec 1-4, 1980, IEEE Press.

Gougeon, F.A.; Moore, T. 1988. Classification individuelle des arbres à partir d'images à haute résolution spatiale, Pages 185-196 in Proc. 6e congrès de L'association québécoise de télédétection , Sherbrooke, Québec. May 4-6, 1988.

Gougeon, F.A.; Wong, A.K.C. 1986. Spectral and textural segmentation of multispectral aerial images. Pages 291-300 in Proc. 10th Can. Symp. Rem. Sens., Edmonton, Alberta. May 5-8, 1986.

Gougeon, F.A.; Wong, A.K.C. 1987. Towards knowledge-based spectral and textural signature recognition, Pages 565-578 in Proc. 11th Can. Symp. Rem. Sens., Waterloo, Ontario. June 22-25, 1987.

Leckie, D.G. 1990. Advances in remote sensing technologies for forest surveys and management. Can.

J. For. Res., 20:464-483.

Mouchot, M.-C.; McCullough, G.; Fabbri, A.; Dupont, O.; Kushigbor, C. 1989. Application de la morphologie mathématique à l'étude des réseaux hydrologiques complexes. Pages 51-71 in Proc. 6e congrès de L'association québécoise de télédétection, Sherbrooke, Québec. May 4-6, 1988.

Matsuyama, T. 1987. Knowledge-based image understanding systems and expert systems for image processing. IEEE Trans. Geosci. Rem. Sens. GE-25(3):305-316.

McKeown, D. M., Jr. 1987. The role of artificial intelligence in the integration of remotely sensed data with geographic information systems. IEEE Trans. Geosci. Rem. Sens. GE-25(3):330-348.

Nagao, M.; Matsuyama, T. 1980. A structural analysis of complex aerial photographs. New York, Plenum.

Nicolin, B.; Gabler, R. 1987. A knowledge-based system for the analysis of aerial images. IEEE Trans. Geosci. Rem. Sens. GE-25(3):317-329.

Pinz, A.J.; Bischof, H. 1990. Constructing a neural network for the interpretation of the species of trees in aerial photographs. Pages 755-757 in Proc. 10th Int. Conf. Pattern Recognition, Atlantic City. N.J. June 16-21, 1990.

Pinz, A. 1989. Final results of the Vision Expert System VES: finding trees in aerial photographs. Pages 90-111 in Pinz, A.,ed., Wissensbasierte Mustererkennung (Knowledge-Based Pattern Recognition); OCG-Schriftenreihe 49, Oldenbourg.

Selfridge, P.G.; Sloan, K.R.,Jr. 1981. Reasoning about images: using meta-knowledge in aerial image understanding. Pages 755-757 in Int. Joint. Conf. on Artificial Intelligence, Vol. 2.

St-Onge, B.A.; Cavayas, F. 1991. La segmentation d'images de forêt de haute résolution à l'aide de mesures de texture basées sur le demi-variogramme. Pages 219-225 in Proc. 7ème congrès de L'association québécoise de télédétection, Montréal, Québec. October 23-25, 1991.

Teillet, P.M.; Goodenough, D.G.; Guindon, B.; Meunier, J.-F.; Dickinson, K. 1981. Digital analysis of spatial and spectral features from airborne mss of a forested area. Pages 883-903 in 15th Int. Symp. Rem. Sens. Env. May 1981.

Zeidenberg, M. 1990. Neural networks in artificial intelligence. Ellis Horwood Ltd. Chichester, England. 286 p.