Using an Expert System to Update Forest Maps

Yves Voirin, Froduald Kabanza, Goze Benie, Eric Beaudry and Kami Rousseau
Université de Sherbrooke
Sherbrooke, Québec, Canada
{yves.voirin, kabanza, goze.benie, eric.beaudry, kami.rousseau}@usherbrooke.ca

Abstract

An important application for satellite images is to monitor environmental temporal changes. The comparison of an old image to a recent one allows the detection of change that occurred during the period between which both images were taken. To date, change identification is essentially done by human experts. Experts can use image processing software to quickly obtain an approximate map, however obtaining an accurate one remains a lengthy process which requires knowledge about the field, sensors, image processing, and image interpretation. In this paper, we present an emerging application that uses multi-agent and expert-system technology to help users in updating forest maps. This is achieved by assisting them in three tasks: selection of the appropriate images and ancillary data for updating a map; processing of the selected image and data to generate interpretable output (e.g., classifying them into thematic regions); interpretation of the output.

Keywords: Expert systems, image analysis, forest mapping
ISESS Submission Category: P-II Case Study, Decision Support System (DSS)

1 Introduction

Many environmental applications in remote sensing deal with the problem of the identification of zones that have changed overtime. For instance, in forest management, they are used to monitor burned areas or the regeneration of previously cut areas. In this paper, we present an emerging expert system that is aimed at assisting in the identification of forest changes due to natural phenomena (e.g., zones damaged by fire, wind or flood), forest activities (e.g., cut or regeneration zones), epidemics (e.g., caused by insects) and phenology (e.g., defoliation).

Traditionally, the detection of changes is completed in three steps. The first step is to acquire satellite images. This acquisition requires expertise in order to select images from satellite sensors that use the appropriate spectral band for a forest application. Indeed, a sensor may be effective for urban areas but it may not necessarily be suitable for forest regions, because objects in these two situations do not have the same spectral and geometrical features. Another constraint in the selection of images is cost: a high spatial resolution image usually gives a lot of details, but costs more money and more time for analysis. On the other hand, a too low resolution image may not reveal enough details to detect changes. Hence, to choose the right image, one needs minimal knowledge in remote sensing principles (particularly knowledge about spectral bands provided by sensors and about spectral bands that are necessary for the detection of a specific object). We have represented some of this knowledge into a rule-based *image data selection assistant (IDSA)* that is integrated in our image analysis tool.

The second step is image segmentation (also called classification). It consists in dividing the image into different zones (called segments or classes). Each segment corresponds to a certain theme of interest, that is, the nature of changes. For instance, we may have zones damaged by fire, zones damaged by wind, zones of cuts, or zones of regeneration. In each case, we have subclasses relative to the importance of change. Cuts may be subdivided into several classes, such as clear cuts or selective cuts. A sequence of processes with the raw data is necessary before the stage of segmentation. First, noise must be filtered out (for example with SAR images, the signal may be perturbed by the speckle). After filtering, geometrical corrections may be applied to reduce the distortions caused by the movement of satellite and the ellipsoid of the Earth, so each pixel of the image must correspond to a precise geographical area. This may be followed by atmospheric corrections to reduce atmospheric perturbations (gas in the atmosphere may interfere with the satellite signal; clouds and cloud shadows also interfere with the signal).

The last operation is the actual segmentation, which attempts to group adjacent regions of the image sharing similar characteristics and to identify them. For all these operations, commercial tools (such as ENVI, PCI Geomatics, ESRI ARCGIS, and IML Sepia) as well as public tools such as (GRASS and Mapserver) are available, and are not limited to forest images. Within a particular tool, each of these operations is often supported by many different alternative algorithms, so that the user must select and apply one depending on the image attributes (e.g., its spatial resolution) and the types of changes he or she wants to detect (e.g., some segmentation algorithms perform better on regions covered with vegetation than on urban terrains). For example, choosing the wrong filtering algorithm may jeopardize the quality of the segmentation process completed later in the flow sequence of image processing, even if at that stage the correct segmentation algorithm is used. Thus it requires a degree of expertise to know which method to apply, depending on the situation. However, image processing methods change constantly, so the problem of selecting the appropriate method is real even for experts.

Our system includes an algorithm selection assistant (ASA) which assesses the image characteristics and the image processing task in order to propose the appropriate image function at every stage. For example, when the user selects an image processing function for segmentation, ASA is activated to validate that selection as a background task. If the selection is not appropriate, a dialog is initiated with the user to propose alternative methods with explanations of their relevance. ASA is also configurable to select algorithms without initiating a dialog with the user.

The final step is validating the segmentation completed in the previous step. The validation consists of confirming the identification of every region and by eliminating the possible errors. To automatically update a map, the segmentation result is supposedly sufficient for the identification of changes. However, because of noise in data and intrinsic errors in image processing algorithms, forest managers still rely a great deal on human interpreters who use such automatically-segmented images, complemented with ground knowledge. A human interpreter not only uses spectral and visual information from images, but also his or her knowledge about the region, experience, and common sense. A natural question is "how to develop an effective assistant for helping users during the interpretation step ". Concerning this question, we added an expert-system based classified image interpretation assistant (CI2A) to our system in order to validate segmentations made by traditional algorithms.

The remainder of this paper is organized as follows. In the next section, we describe the overall architecture of our system. We then provide more details on each of the three assistants. We subsequently discuss some of the lessons learned at the current stage of our implementation. We conclude with a discussion on related work and our ongoing research and implementation efforts.

2 System Architecture

SITI software is written mainly in Java (with a few basic libraries in C++) and has five main modules (Figure 1): interfaces, an image processing library, IDSA, ASA and CI2A. Figure 2 shows a snapshot of SITI GUI, with a loaded image and on the bottom its histogram view. The GUI menu includes typical commands for image processing tools.

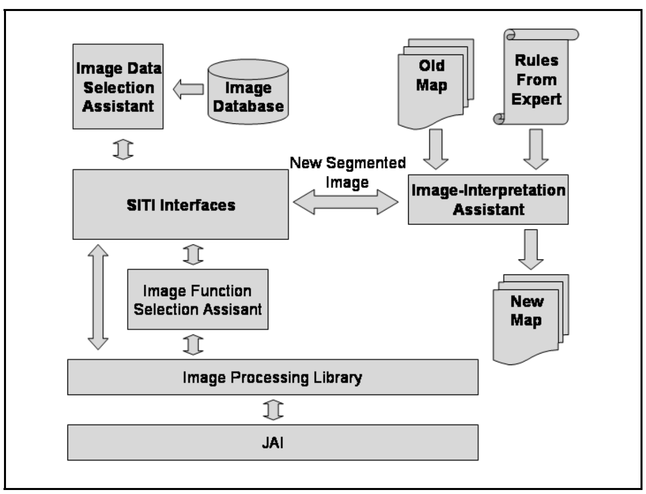


Figure 1: SITI Architecture

The interfaces consist of graphic user interfaces (GUIs) that implement interactions with the user and internal functions that implement communication between the different components.

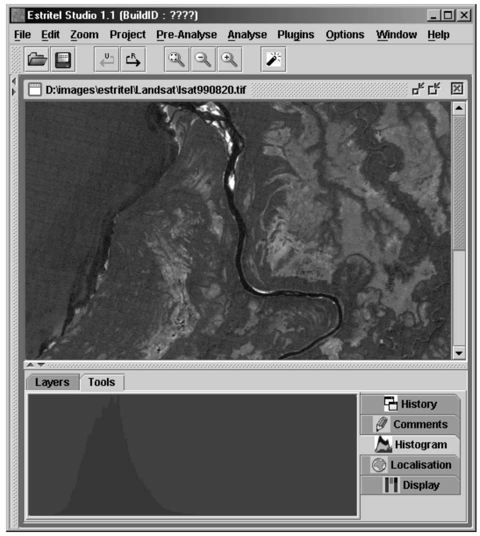


Figure 2: SITI GUI

The image Processing Library is an abstract layer on top of the Java Advanced Imaging (JAI) library from Sun Microsystems. The library is composed of custom functions for advanced processing, which are discussed in the section below on ASA. We mainly use JAI for Image I/O, since it supports major image formats like TIFF and GeoTIFF. We developed custom JAI operators for filtering, segmentation, and texture analysis.

3 Image Data Selection Assistant (IDSA)

In order to detect large area cuts, an inexperienced person may order a low spatial resolution Landsat MSS image, with a pixel size of 30 meters by 30 meters, thinking it is sufficient to obtain a highly accurate result. He or she is not aware that a better spatial resolution image (e.g. Ikonos) would allow detection of roads and the border of cuts with greater accuracy. We know that it requires a road to transport the machinery to perform the cuts, and that this road is a important key for the accurate identification of a cut.

The goal of IDSA is to assist the user in choosing images and ancillary data to be used for image classification or interpretation. It locates images for map creation or updating by applying knowledge rules based on the spectral and spatiotemporal properties of the images available in its database. For the time being, the knowledge base is relates to forest mapping and contains 50 rules. Initially, all images in the desired time period are considered valid; subsequently basic rules are used to eliminate images that are not within the region of interest; advanced rules are utilized to validate the request, while others eliminate those sensors that cannot be used for the selected theme (e.g., MSS images cannot be used to create a map with a scale larger than 1/125 000). If some necessary images cannot be found to fulfill the request, IDSA will suggest images to acquire.

From the user perspective, IDSA is an intelligent “File Open”. Images are stored as files, but their properties (called image metadata) are stored in an SQL image database, linked to SITI using Java Database Connection (JDBC). IDSA uses an expert system to reason about the database content and to provide assisted access to it. The current implementation uses Java Expert System Shell (JESS) engine (Friedman, 2003). In addition to production rules about image metadata and their usage, IDSA knowledge base has production rules that reason about facts from user-profiles to teach them about image metadata. IDSA knowledge base is structured into different small knowledge bases, organized depending on themes. For instance, production rules specific to image data from a region are in a knowledge base for the region. Tutoring rules are also structured and stored separately. This rule structure significantly reduces the overhead in JESS pattern matching, as the relevant knowledge base is loaded in dynamically.

4 Image Processing Assistant (ASA)

ASA is activated from Pre-Analysis and Analysis menus, by selecting 'Assisted Algorithm Selection', leading to a GUI for specifying selection criteria and having an assisted access to the list of available algorithms (Figure 5).

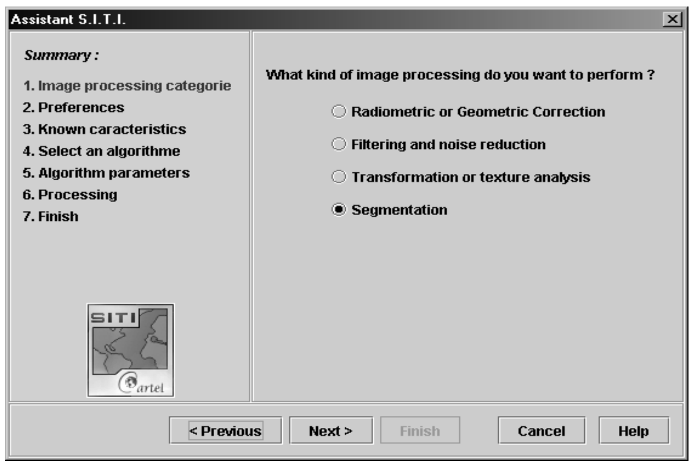


Figure 5: ASA Main GUI

For instance, these are some of the segmentation functions: K-Means, SEM, EM, Fussy SEM, Markov random fields, Watershed, Watermerge, Splitmerge, maximum likelihood. Each of these functions implements a particular technique in image segmentation, which can be easily recognized by readers who have advanced knowledge in image processing.

Anticipating a deployment of our system, in which image processing algorithms are distributed across multiple servers, ASA uses a multi-agent auction metaphor to recommend image processing algorithms that fit with the requirements. ASA is wrapped into a buyer agent (it "buys" image processing services), whereas each possible algorithm is wrapped into supplier agent. Given a service request from the buyer, for example segmenting the image currently loaded, all segmentation agents make a bid and the winner is selected. A bid is computed as a weighed sum of the estimate CPU time for processing of an image (which is estimated based on a formula that takes into account the image size and the $O(n)$ complexity of the algorithm where n is the image size).

5 Image Interpretation Assistant

The role of CI2A is to verify the thematic class of each region in a segmented image, using an older one covering the same region. Like IDSA, CI2A is based on JESS engine. Each region is translated according to an *image object* (Sester, 2000), providing a basis for using image region properties as facts in JESS.

An object is a structure representing a region with its attributes, such as geometry (perimeter, main axes, orientation, and compactness), texture (entropy, homogeneity), spectrum (reflectance, vegetation indexes), a class label resulting from segmentation (e.g., bare soil, water, forest or cuts) and spatial relationships (objects inside the region, neighbor objects, and close objects). Within JESS an image object is simply manipulated as a fact. Given an object O, and its attribute $a_1, \ldots, a_n$, the corresponding fact (or predicate) is represented as (O $a_1, \ldots, a_n$). Since satellite images can have a huge size (8000 x 8000 pixels is typical), the number of facts may be too important. An interesting functionality of IDSA and ASA is to anticipate image data needs by CI2A, so as for example to segment and classify images in advance. Processed images are stored in a permanent database, on which CI2A is expected to base its facts.

In addition to region attributes, we have facts that do not relate directly to images, such as regions cut over the last 5 years, regions already known to have caught fire and so on. All such relevant information is in a database. Moreover, JESS can prompt questions when interacting with the user to ask about facts. For instance, it may ask “is there a road in this region?”. With this knowledge and region attributes, we defined rules for forest mapping. Approximately 70% of them were acquired from the literature (Hayes and Sader, 2001) and 30% from forest experts.

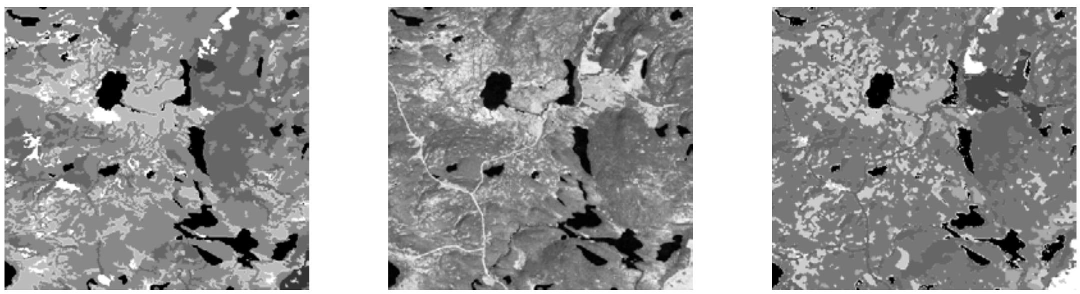
For instance, cuts are indicated by low vegetation indexes (for example, the infra-red band is very low) and an increase in brightness. We also know that forest managers build roads near cut areas. A cut area can also be found near another cut zone. If an area on the old map was covered with forest, but now has low vegetation indexes, then it must be a cut area. All such knowledge can be expressed using expert system production rules. To illustrate, the following two JESS rules specify that if area on the old map is covered with forest, but now has low vegetation index (NDVI), then it must be a cut area (the last rule has a .8 certainty factor).

```
(defrule low-ndvi
?f<-(object (number ?number)(oldclass  mixed|
dense|open|shrub)(ndvi ?ndvi))
(test (< ?ndvi ?*ndvilow*))
=>
(assert (object ?number has a low value of  ndvi)))

(defrule possible-clear-cut
?f<-(object ?number has a low value of  ndvi)
?f2<-(object ?number has a road in proximity)
=>
(probClassSeg "store" cut 0.8))
```

Labrador Scenario

We show an example of image interpretation for the region in Labrador, Canada. At first, the user loads the old map (Figure 6.a) and the recent image (Figure 6.b). Each image has a size of 360 x 335 pixels, with a resolution of 30 x30 meters per pixel. The recent image is then classified using a segmentation algorithm (Figure 6.c).



(a) (b) (c)

Figure 6: available data

The next step, the user defines the legend of the map (Figure 7). Then the system can reason about the thematic name of each class, such as shrub or cut. Figure 8 is the output of classification validation by CI2A. For pretty printing, the legend and the map are displayed in black and white; the actual palette is in color.

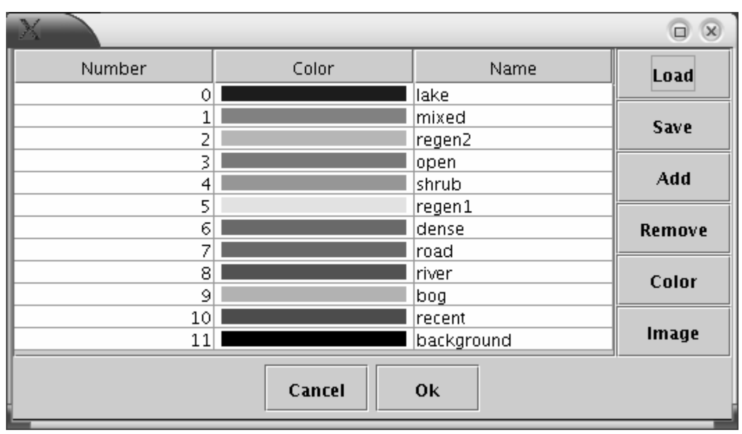


Figure 7: legend editor

For this example, a set of 2506 objects has been found. The system created 37 291 facts and fired 45 182 rules. The interpretation process lasted 2.07 min on a Pentium III (1 Ghz) computer with 384 Mg (RAM).

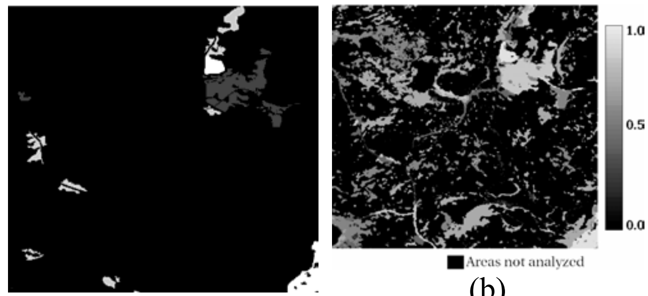


Figure 8 : results with CI2A

The output of CI2A is a map of changed areas and a map of certainty factors. For this example, the assistant finds out that some regions in the image do not actually correspond to disturbed areas (black areas in figure 8.a). On the other hand, other zones are changed on the ground. Dark grey areas are the recent cuts confirmed by CI2A. The brighter areas are the regeneration sites. The second map (figure 8.b) represents the degree of certainty associated with each region.

Experiments

We compared CI2A to other approaches of map updating (see (Coppin, and Bauer, 1996) for a survey and table 1 for a summary). A partially automated method is one by which a user employs image analysis tools but has to provide key parameters (e.g., threshold value for removing the noise, identification a set of areas in the scene to train/supervise a classification algorithm).

Table 1: classical updating approaches.

Type	Change detection	Identification	Input data
Partially automated methods	Image differencing	statistical classification	Old and new image
	Principal components analysis	statistical classification	Old and new image
	Cross-correlation analysis	statistical classification	Old map and new image
Human interpretation	Visual detection	Visual identification	Old map and new image

Human interpretation is still the method of reference, given that human experts achieve so far more accurate results; however this is also more expensive and is time-consuming. On the image of figure 6, CI2A gives only 2% of false alarms and a good identification rate of 86 % (average). Experts would reach 0% of false alarms and 90% of accuracy. On the other hand, the best semi-automatic method (image differencing) gives 7% of false alarms and 80% of good identification, where “false alarm” means the rate of misclassified zones and “good identification” represents the rate of correct classification. We performed several other tests on 5 different set of images (including Figure 6). Table 2 gives the average results for a human expert, CI2A and the semi-automatic methods. Thus CI2A strongly reduces the errors and the average identification rate is more accurate than a conventional updating method.

Table 2: results for each approach (average)**.**

Method	False alarm (%)	Good identification (%)
Image differencing	7	80
Principal components analysis	7	79
Cross-correlation analysis	6	53
Human interpretation	0	90
CI2A	2	86

7 Related Work

The need for tools to assist in selecting image is a recognized problem for several applications. Diverse applications approach it differently. The simplest case involves image catalogs sorted by acquisition dates. Searchable directories provide a slightly higher level of organization by grouping images by themes. Image search engines are by far the most utilized approach (Cahumeau and Dimap, 1999): the user can search for images using keywords, however he or she is not provided assistance in determining whether the image is appropriate for the task at hand or not. Intelligent database systems, for example utilizing data mining provide even much higher flexibility (Fraser and Smith, 1999). The user, however, is not provided tutoring assistance on the relevance of recommendations made by the system. The use of assistance in data selection has also been applied in other domains, including geographic information systems (Lanter, 1992; Guntig, 1994).

A lot of works exist for updating map with satellite images, but only few approaches have used an expert system. Goldberg *et al.* (1985) and Fung *et al.* (1993) gave a possible application of a knowledge based system. The both approaches were too complicate to implement in a real situation. The knowledge of many fields in remote sensing was necessary, that is why no experiments have been performed.

8 Conclusions

In his 1993 survey of expert systems in remote sensing, Tsatsoulis notes that one of the major limitations is the integration of too much information (Tsatsoulis, 1993). This observation is still true today, perhaps even more so. In addition to the complexity of knowledge that has to be integrated in an expert system for remote sensing image analysis, we now have an increasing complexity in the number of image processing algorithms that have to be used, raw images from much more available sensing devices, with different levels of resolution and image ancillary data.

Contrary to previous approaches that use one expert system to interpret images, our approach separates the task into data acquisition, image-processing algorithm selection, and image interpretation, each of them being assisted by a separate expert system. Also, CI2A validates classified images rather than trying to classify them. With the current implementation, the major bottleneck is in the inference engine JESS. We are currently improving this model by attempting to replace JESS by decision trees. Later on, we also plan on experimenting with existing techniques for learning decision trees based upon patterns in the way users process images using SITI. To that end, new metadata could be included in the IDSA database (e.g., histogram and texture parameters). The inclusion of these metadata would help in choosing images, much like the cloud cover information normally included with optical images. That way, when suggesting a purchase, IDSA would show the user similar images to test the efficiency of the processing algorithms, before actually buying those images.

Acknowledgement

This work was supported by the Canadian Natural Science and Engineering Research Council through its Strategic Research Program. We are also grateful to our colleagues Dong Chen He, Kalifa Goita, Mathieu Vorron and Ko Fung for helpful input.

References

Chahuneau, F. Dimap: Satellite image metadata standard based on XML. *In Proceedings of XML 99*, Philadelphia, December 1999.

Coppin, P. R. and Bauer, M.E.: Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Environement*, 13, 207-234, 1996.

Fraser, D. and Smith, D. Rising. Above a Quagmire of Search Protocols. *In Proceedings of the EO/GEO-99 Web and Internet Workshop,* Washington Area, USA. 1999.

Friedman-hill, E. JESS in Action. *Manning Publications*, 2003.

Fung, K., Goodenough, D., and Ryerson, R. The Photo-Interpretation Keys Expert System (PIKES). *International Geoscience and Remote Sensing Symposium*, Tokyo (Japan), 177-182, 1993.

Goldberg, M., Goodenough, D., Alvo, M., and Karam, G. A Hierarchical Expert System for Updating Forestry Maps with LANDSAT Data. *IEEE Transactions on Geosciences and Remote Sensing*, 73, 1054-1063, 1985.

Guting, H., An Introduction to Spatial Database Systems, *Very Large DataBase Journal,* 3, 357-399, 1994.

Hayes, J. and Sader, A. Comparison of Change-Detection Techniques for Monitoring Tropical Forest Clearing and Vegetation Regrowth in a Time Series. *Photogrammetric Engineering and Remote Sensing*, 67(9), 1067-1075, 2001.

Lanter, D., *Intelligent Assistants for Filling Critical Gaps in GIS: A Research Program*, Department of Geography, University of California at Santa Barbara, Report 92-4, 1992.

Sester, M.Knowledge acquisition for the automatic interpretation of spatial data. *International journal of Geographical Information Science*, 14(1), 1-24, 2000.

Tsatsoulis, C. Expert Systems in Remote Sensing Applications, *IEEE Geoscience and Remote Sensing Society Newsletter*, June, 7-15, 1993.