

Teaching while selecting images for satellite-based forest mapping

Froduald Kabanza and Kami Rousseau

Département d'informatique

Université de Sherbrooke

Sherbrooke, Québec, J1K2R1, Canada

{froduald.kabanza, kami.rousseau}@usherbrooke.ca

Abstract

Satellite images are increasingly being used to monitor environmental temporal changes. The general approach is to compare old images to recent ones acquired from a satellite in order to detect changes that occurred during the period between which these images were taken. An important step in this overall process is the acquisition of image data that will best allow assessing the temporal changes. This acquisition requires some expertise in order to select images from satellite sensors that use the appropriate spectral band for a forest application. For instance, a sensor suitable for classifying urban images will not necessarily be appropriate for forest mapping, because buildings reflect electromagnetic waves differently from trees and hence show a different spectrum with the same sensor. In this paper, we present an emerging Image Data Selection Assistant (IDSA) that uses an expert system combined with intelligent tutoring to help users in choosing images for updating forest maps, while at the same time teaching them how to best select images depending on the task.

Keywords: tutoring system, expert-system, image processing, image acquisition

1 Introduction

The identification of zones that changed overtime is an important problem in environment monitoring. This problem arises for example in forest mapping when monitoring burned areas or the regeneration of previously cut areas. The problem also occurs in agriculture when assessing plant growth or damage by insects or diseases. Another application deals with assessing damaged regions by natural disasters, such as flooding. Traditionally, the detection of changes is done in three steps. First, one has to acquire satellite images for the region of interest. Then, one segments or classifies the images into different zones (called segments or classes), such that 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 zone of regeneration). Finally, the classified image is interpreted and compared to old images to determine changes that have occurred and produce an update map.

The acquisition of images requires some expertise in order to select images from satellite sensors that use the appropriate spectral band for a forest application. For instance, a sensor suitable for classifying urban images will not necessarily be appropriate for forest mapping, because buildings reflect electromagnetic waves differently from trees and hence show a different spectrum with the same sensor. Another constraint in the selection of images is cost. A high spatial resolution image usually gives better results, but costs more money and perhaps more time in 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 the physics of images (specifically which spectral bands are covered by the satellite sensors, and which spectral bands match best the kind of details one want to assess in the forest mapping).

We have been trying to represent some of this knowledge into a rule-based *image data selection assistant* (IDSA). IDSA is a component of a more complex system for image processing, called SITI (for « *Système Intelligent de Traitement d'Images* », that is, the French translation of “*intelligent image processing system*”). SITI was developed by a group of researchers at the “Centre d'Applications et de Recherches en Télédétection”

(CARTEL) at the Université de Sherbrooke in Canada. It is written mostly in Java (with a few basic libraries in C++) and uses Java Advanced Imaging (JAI) library from Sun Microsystems for basic image processing and home-made functions for advanced image operations such as segmentation. Figure 1 shows a snapshot of SITI GUI, with a loaded image and its histogram view on the bottom. The tool allows to load in an image and segment it (this involves a number of specialized image processing algorithms which can all be selected by the user either directly or automatically through the use of an algorithm selection assistant, such as noise filtering, geometric correction, and the actual image classification algorithms).

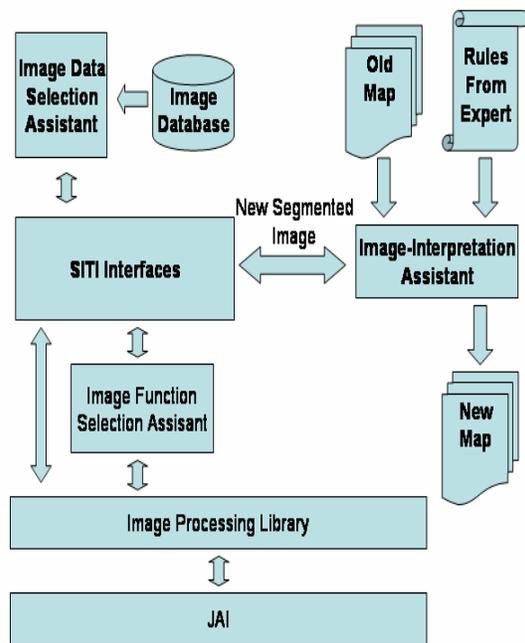


Figure 1 (a): SITI architecture

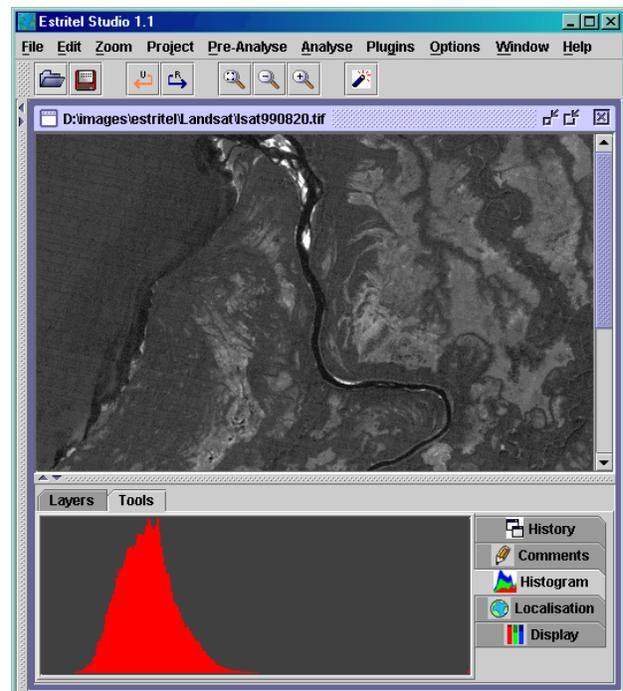


Figure 1 (b): SITI GUI

From the user perspective, IDSA is an intelligent “File Finder”. 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. 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. In this paper, we report on this integration of an expert system and an intelligent tutoring system into the image selection module of SITI. Our goal is not to replace the competence of a human expert. Our goal is rather to make the user's daily work much easier and to reduce the costs associated with map updating.

IDSA is still in the early stages of development, but we already have a prototype that works enough to enable us sharing some insights at the current stage of development. In the next section, we discuss the most related work. We then describe the IDSA modules, illustrating them with two simple scenarios. The focus is only on the selection of image data, not on the overall architecture of SITI, hence we will not discuss the processing that goes on after image data have been selected.

2 Related Work

The need for tools to assist in selecting image data is a recognized problem that is addressed differently in many applications. The simplest approach is to use image catalogs sorted by acquisition dates, with searchable directories providing a slightly higher level of organization by grouping images by themes [6]. Intelligent database systems, for example using data mining, provide even much higher flexibility [9]. Still the user has no tutoring assistance on the relevance of recommendations made by the system. However, the use of assistance in data selection has been applied in other domains such as geographic information systems [13].

Some aspects of selecting appropriate image processing algorithms at different stages can be formulated as a planning problem. Choosing for example the wrong filtering algorithm may jeopardize the quality of the segmentation process done later in the flow sequence of image processing even if at that stage the right algorithm is selected. Recent work use AI planning techniques by approaching this problem as a problem of automatically generating sequences or flows of image processing routines [12].

Of course there exists documentation related to image data selection from which users can learn. For example, the Belgian Federal Science Policy Office (BFSPPO) created the "Telsat guide" to select imaging sensors [14]. They also provide another tool to select images by sensor, from a database. However, all these tools are kept

separate. In a similar vein, the Canada Centre for Remote Sensing provides Internet tutorials for image processing in general [5], including one tailored to assisting users in finding the mode of RADARSAT-1 images most appropriate to extract thematic information or to generate a digital elevation model [11]. However, these tools can not be used to select images from a database on the fly during an automated image processing task.

Efforts have also been made to use Geographical Information Systems (GIS) in K-12 classrooms for active learning [2]. This may be an interesting area for further development and deployment of computer programs similar to our own.

On the other hand, the use of expert systems in computer-aided education is widespread and covers a range area of applications. One of the earliest such applications is GUIDON project in medicine, which integrated tutoring production rules into MYCIN to teach student to diagnose infectious diseases given a patient's history, physical examination, and laboratory results [3,4]. Another, more recent example, is SlideTutor which uses an expert system to teach student on the diagnosis of skin inflammatory diseases [7]. The validity of production rules to model cognition in intelligent tutoring systems (ITS) has also been demonstrated by a quite large number of prototype applications built around Anderson's ACT-R cognitive architecture [1]. This framework encompasses much more than a traditional expert system inference engine. In particular, it has several concurrent modules reminiscent of human cognition, using a different principle for modeling production rules and a different approach for firing them, compared to a traditional expert system engine. Production rules have utility measures which can be learned over time through the concurrent executions of different modules depending on how and how much those rules are involved. Another interesting feature of ACT-R is the clear distinction between goals from other facts, which allows to reason about the intentions and actions of a system being modeled. Nevertheless, ACT-R remains essentially a symbolic, production-rule based system, so are the many ITS applications based on it.

The contribution of our paper is not in the use of production-rule based system for intelligent tutoring *per se*. It is rather in demonstrating the feasibility of automated tutoring for the problem of image data selection which, as far as we know, has not been addressed in this way. We demonstrate the availability of expert knowledge for the

image data selection problem and the validity of its modeling by production rules. This is significant in that it paves the way for improved user-interfaces for image processing applications, which is an area with a growing potential, not only in remote sensing (our current domain of interest), but also for example in medical imagery. Once the validity of knowledge representation in this domain is established with classical production rules, an interesting next step is to adopt a more sophisticated architecture of cognition like ACT-R.

3 IDSA Modules

IDSA uses Java Expert System Shell (JESS) engine [10]. It contains six main modules (Figure 2): the image database, the user model, the expert rules, the question bank, the thematic expert and the tutor. 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 reduces significantly the overhead in the expert system's pattern matching, for the relevant knowledge base is loaded in dynamically.

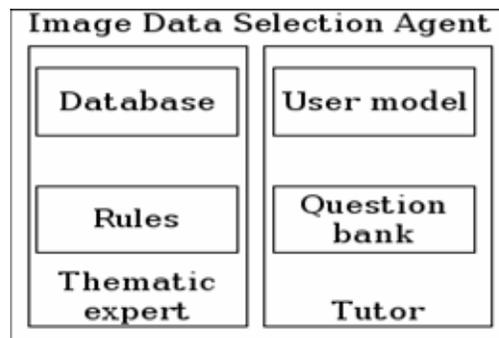


Figure 2: IDSA modules

3.1 Thematic expert

All the images' metadata (coordinates, sensor, file location) are stored in a remotely accessible **database**. We developed our prototype using **Microsoft Access** via RMI-JDBC. The database server is responsible for finding all images in a requested time-frame. IDSA decides whether or not to keep an image based on its coordinates. Basic

image requests use geographical coordinates and dates (e.g., getting an image taken over region for a given date). Advanced requests, for which the use of the expert system is necessary, may also involve the theme (e.g., forestry) or the desired scale of the output map.

The goal of the expert system is to assist the user in choosing images and ancillary data to be used for image classification or interpretation. It finds images for map creation or update 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 about forest mapping and contains 50 rules. Initially, all images in the desired time period are considered valid; then some basic rules are used to eliminate images that are not within the region of interest; some advanced rules are involved to validate the request, while others eliminate those sensors that can not 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, the system will suggest images to acquire. The system tries to find as many images as possible for a given location, in order to have information about temporal variability. It also tries to avoid suggesting too many different sensors and spectral bands, to reduce the risk of adding too much noisy signal [11]. When the system has finished processing the request, the images that it found are available on CD or on a networked image server.

3.2 Tutor

Users would not trust the results if they were unable to verify their accuracy. Some users consider the reasoning process to be just as important as the final answer. Therefore the expert system produces a trace of its reasoning, with a justification for each intermediate step in its decision process. Other intelligent data selection programs do that, e.g. the NASA Goddard's Intelligent Information Fusion System (IIFS) [12]. But these traditional expert systems, like IIFS, stop at this point. In addition, IDSA includes a tutoring system integrated with the data selection module, explaining each step of the reasoning process to the user, helping him learn about the information that can be extracted from the images.

After the thematic expert has completed the processing, it sends its trace to the tutor module, which chooses a theme from all concepts used in it. The tutor verifies the user's comprehension of technical terms by asking him questions, including in some cases multiple choice questions, with regard to the image data he is trying to acquire and to what we plan to do with them. The tutor keeps track of the users' answers in a user profile to avoid asking again the same questions. Rather than giving a sample of production rules that are not necessary easily understood with the underlying syntax and image processing technical concepts they convey, we opt for illustration with some simple scenarios.

4 Illustrations

The Canada Centre for Remote Sensing provided us with images to validate the system. These images are from the regions of Goose Bay, Labrador, and Saguenay-Lac-Saint-Jean in Québec, Canada. In addition, the database contains radar images from the RADARSAT sensor, as well as optical images from the MSS, TM, ETM+, HRV, AVHRR, IKONOS and Polder sensors. We also have airborne hyperspectral CASI images, digitalized aerial photographs, digital elevation models and forest maps. We now illustrate IDSA through two different scenarios.

4.1 Labrador Scenario

A user needs images covering the region of Labrador (north-east Canada). For this, he begins by selecting "File", then "Open using IDSA" from SITI main GUI. IDSA then loads in the user's profile and the image knowledge base. After checking the profile IDSA notices that this is an expert user, so it deactivates novice tutoring production rules; as a consequence, novice tutoring interaction is skipped, so the user is prompted to specify his image request (Figure 3(a)); he gives Labrador coordinates, the period of interest (beginning and end dates), what the image is aimed for (forest mapping) and the scale (1/20000).

IDSA can now begin its inference. It finds out that some images in the Labrador database are not georeferenced, that is, they have not been assigned exact geographical coordinates. Since images should normally be georeferenced before being loaded and processed, the user is warned. She is asked if the aforementioned

images are to be kept; she confirms (Figure 3(b)). IDSA displays a list of relevant images from which she can open any by clicking on it (Figure 3(c)).

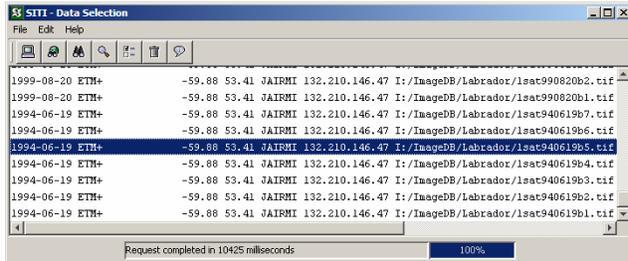


Figure 3 (a): making an image request



Figure3 (b): validating selected images

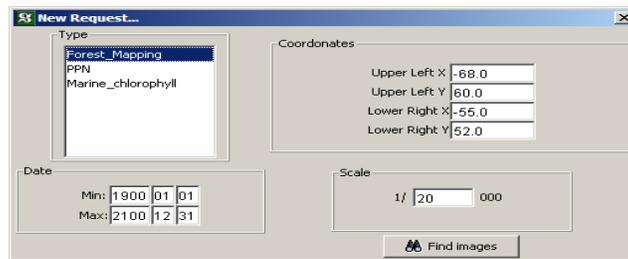


Figure 3 (c): returned list of images

4.2 Ontario Scenario

A beginner has an old map of south Ontario (from -83°W , 48°N to -78°W , 45°N) and want recent images of the region to update the map. He makes a request (as with the previous user, the request includes region coordinates and what he wants the data for); however, IDA finds no map matching the request specifications and no image. So it recommends purchasing Landsat images. This failure and the lack of evidence of database experience in the user profile fire production rules that activate IDSA tutoring knowledge base, switching IDSA into a tutoring interaction. A few database multiple-choice questions are asked and the student fails (Figure 4(a)). IDSA suggests him taking some database query reading material (Figure 4(b)).

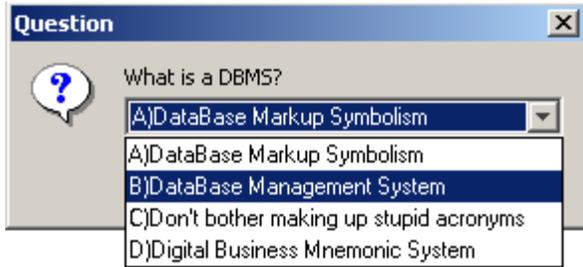


Figure 4(a): IDSA tutor asking a question

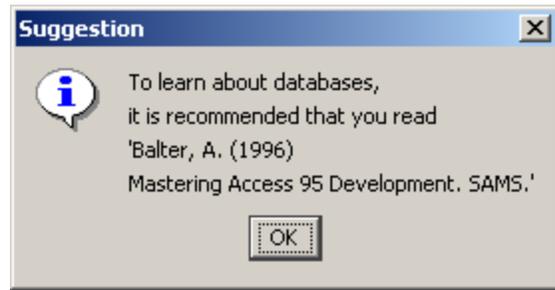


Figure 4(b): IDSA tutor making a reading recommendation

Below is IDSA trace, which can be consulted by the user.

Retrieving all images from the specified time period
 Removed all images outside the region of interest
 Removed all images from sensor MSS
 Removed non-georeferenced images
 Run finished!
 Data Selection agent will now print its trace:
 Getting the images from the database
 Cannot use MSS imagery (scale larger than 1/125000)
 Removed all MSS images
 No map was found. Asking for user guidance
 No map: user wants to use cartography rules
 Non-georeferenced images found. Ask user for guidance
 Non-georeferenced images found; they were removed
 No data available for selected region
 End of trace

IDSA only finds Landsat MSS images for the zone of interest. These images were eliminated because they are not appropriate for mapping at a scale finer than 1/125000. After a look at the trace, the user realizes he can get the data he wants by changing the scale in his request to 1/250000. He reformulates his request, using a scale of 1/250000, in the hope of keeping the MSS images. While processing the request, IDSA realizes that the choice

of scale is unusual. It asks the students if he made a mistake (Figure 5). IDSA finally gives a list of image data.

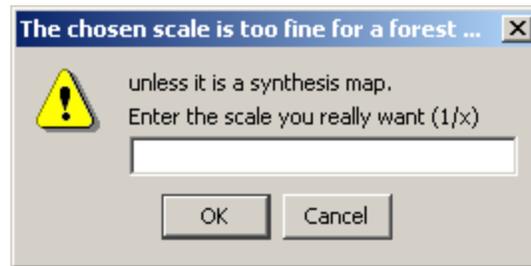


Figure 5: IDSA reacts to an incoherent request

5 Conclusion

The above scenarios illustrate the deductive reasoning capability of IDSA, combining image data selection and some tutoring interaction. So far the production rules in the system only covers the topic of forest mapping, with fifty rules. The inclusion of new themes is being considered (atmosphere, urban environment).

As with other expert systems, one of the challenges we are facing is with the bookkeeping of production rules. Keeping the rules in many separate files in general makes it easier to maintain them (by sensor, by theme, and by tutoring function). This, however, tends to increase inconsistency in the knowledge base. For example, many rules used for Landsat ETM+ are similar to those used for Landsat TM. If those rules were not in the same file, it would be easy for them to become inconsistent. In that case, a given request might accept TM images and reject ETM+, which is clearly not desirable. At this stage, we have not yet formalized the management of production rules; we store them hierarchically, but the hierarchies are determined empirically. As the knowledge base grows, we will have to adopt a formal rule management system.

We also think of experimenting with mechanisms to learn new rules by finding users' patterns in the processing of images using SITI. To that end, we first plan adopting the ACT-R architecture of cognition [1]. Using JESS, we were able to confirm the relatively easy specification of production rules to model expert knowledge for image date selection. This expert knowledge will serve as a start point to model ACT-R production rules. In this process, we also plan to add new metadata 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 some similar images, to test the efficiency of the processing algorithms, before actually buying those images. The successful use of acquired data will then be used as feedback to the learning modules of ACT-R (this is done by updating utility measures that are associated to production rules).

6 Acknowledgements

IDSA was developed within a research project supported by Canada Centre for Remote Sensing and the Canadian natural Science and Engineering Research Council through the strategic research partnership program. We would also like to thank our colleagues in the project who contributed to IDSA, particularly Eric Beaudry.

References

- [1] J.R. Anderson and K. Gluck. What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. M. Carver (Eds.) *Cognition & Instruction: Twenty-five years of progress*, 227-262, 2001.
- [2] K. Bjork. *Emerging technologies for active learning*. Leadership and the new technologies perspectives. No. 10. July/August 1999.
- [3] J. Clancey. Knowledge-Based Tutoring - The GUIDON Program. Cambridge, MA: MIT Press, 1987.
- [4] J. Clancey. Guidon. *Journal of Computer-based Instruction* 10:8-14,1983.
- [5] CCRS, *Learning resources*. 2003. http://www.ccrs.nrcan.gc.ca/ccrs/learn/learn_e.html
- [6] F. Chahuneau. Dimap : Satellite image metadata standard based on XML. *In Proceedings of XML 99*, Philadelphia, December 1999.
- [7] S. Crowley and P. Medvedeva. A General Architecture for Intelligent Tutoring of Diagnostic Classification Problem Solving. *Proc AMIA Symposium*, 2003: 185-189.
- [8] I. Cyr and Th. Toutin. *RADARSAT-1 Stereo Advisor on CCRS Web*. Canadian Journal of Remote Sensing , 2001 , Vol. 27 , No 1 , pp. 62-66
- [9] D. Fraser and D. Smith. Rising above a Quagmire of Search Protocols. *In Proceedings of the EO/GEO-99*

Web and Internet Workshop, Washington Area, USA. 1999.

[10] E. Friedman-Hill. *JESS In Action*. Manning Publications, 2003.

[11] M-C. Girard. *Traitement des Données de Télédétection*, Série environnement, Dunod, Paris, 1999.

[12] A. Lansky. Web-Based Planning for Image Processing. In Proceedings of the AAAI Spring Symposium on Acquisition, Learning, and Demonstration: Automating Tasks for Users, 1996.

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

[14] Telsat. *Satellite and Sensors Search*. Belgian Federal Science Policy Office, 2003.

<http://telsat.belspo.be/satellites/Satellites.asp>