



AUTORES AUTHORS	PALAVRAS CHAVES/KEY WORDS		AUTORIZADA POR/AUTHORIZED BY	
	KNOWLEDGE-BASED SYSTEM CROP AREA ESTIMATION BLACKBOARD ARCHITECTURE		Roberto Pereira da Cunha Remote Sensing Director	
AUTOR RESPONSÁVEL RESPONSIBLE AUTHOR		DISTRIBUIÇÃO/DISTRIBUTION		REVISADA POR /REVISED BY
Flávio Roberto D. Velasco		☐ INTERNA / INTERNAL ☒ EXTERNA / EXTERNAL ☐ RESTRITA / RESTRICTED		
CDU/UDC			DATA / DATE	
621.376.5			October, 1989	
TÍTULO/TITLE	PUBLICAÇÃO Nº PUBLICATION NO		ORIGEM ORIGIN	
	INPE-4990-PRE/1555		DPI	
A KNOWLEDGE-BASED INFORMATION BASED FOR CROP AREA ESTIMATION		PROJETO PROJECT		
		ANIMA		
AUTORES/AUTHORSHIP	Nº DE PAG. NO OF PAGES		ÚLTIMA PAG. LAST PAGE	
	08		07	
		VERSÃO VERSION		Nº DE MAPAS NO OF MAPS
RESUMO - NOTAS /ABSTRACT - NOTES				
For reasons of accuracy and timeliness, it is interesting that a system for crop production estimation be as automatic as possible, making use of digital processing of satellite images. This paper discusses some mandatory as well desirable requirements for such a system and presents the architecture of a computer information system designed to estimate crop area based on satellite images. To meet the requirements, it is proposed a "knowledge-bases system" to estimate crop area through the integration of satellite image information with maps (thematic and conventional) and tabular census data. The information needed is extracted by independent "knowledge sources" (KS) (experts) and communicated and integrated through a global structured memory area ("blackboard model").				
OBSERVAÇÕES /REMARKS				
Trabalho apresentado no IV Simpósio Latinoamericano sobre Sensores Remotos, de 20 a 24 de novembro de 1989, Bariloche, Argentina.				

A KNOWLEDGE-BASED INFORMATION SYSTEM
FOR CROP AREA ESTIMATION*

Flávio Roberto Dias Velasco
Leila Maria Garcia Fonseca
Guaraci José Erthal
Leonardo Sant'Anna Bins
Antonio Miguel Vieira Monteiro
Instituto de Pesquisas Espaciais - INPE
Caixa Postal 515, 12201 - S. J. dos Campos, S. Paulo, Brasil

ABSTRACT

For reasons of accuracy and timeliness, it is interesting that a system for crop production estimation be as automatic as possible, making use of digital processing of satellite images. This paper discusses some mandatory as well as desirable requirements for such a system and presents the architecture of a computer information system designed to estimate crop area based on satellite images. To meet the requirements, it is proposed a "knowledge-based system" to estimate crop area through the integration of satellite image information with maps (thematic and conventional) and tabular census data. The information needed is extracted by independent "knowledge sources" (KS) (experts) and communicated and integrated through a global structured memory area ("blackboard model").

RESUMO

Por razões de precisão e presteza, é interessante que sistemas para estimativa de produção de safras agrícolas sejam o mais automatizados possível, para tanto fazendo uso do processamento digital de imagens de satélites. Este trabalho discute

* This work has been supported by SID Informatica, Projeto Estra.

alguns requisitos obrigatórios e desejáveis de tais sistemas e apresenta a organização de um sistema de informação computadorizado para a estimativa de áreas agrícolas baseado em imagens de satélite. Para satisfazer os requisitos, é proposto um "sistema baseado em conhecimento" para estimativa de área através da integração da informação da imagem de satélite com mapas (temáticos e convencionais) e dados censitários tabulares. A informação desejada é extraída por "fontes de conhecimento" (subsistemas especialistas) independentes e comunicada e integrada através de uma área de memória global e estruturada ("modelo do quadro-negro").

1. INTRODUCTION

Accurate and timely crop production estimation is fundamental for a country like Brazil, whose economy is mostly based upon the agricultural sector. This information is essential for planning the production storage, transportation, processing of grain crops; making marketing decisions and determining national agricultural policies. In any of these fields, any decision taken based on incorrect information might represent a waste of resources and irremediable losses.

It has long been acknowledged that remotely sensed data can play a major role in the monitoring of agricultural crops, in the estimation of planted area and crop yield. Brazil, through the Institute for Space Research, since 1975 has been implementing operational systems and conducting experiments to use this source of agricultural information. It has, since then, accumulated a vast amount of experience in mapping of several crops like sugar cane, irrigated rice and wheat. Some of the crops, like sugar cane, were mapped for the entire country. Most of the experiments, however, relied on visual image interpretation. Also the crops were either winter or semi-perennial crops.

Most of the annual crops in Brazil are grown in the summer. Except for wheat, which is a winter crop, the other crops of economic significance (soy bean, rice and corn) are planted in October/December and harvested in March/April. Summer crops are very important not only for food production but also for export.

The problem of crop production estimation can be subdivided into two subproblems: crop area estimation and crop yield estimation. Crop production is then estimated as the product of these two quantities. Satellite images can give a significant contribution to the solution of both problems. The present work is restricted to the first problem and proposes a system for crop area mapping. Crop area is obtained as a by product of the mapping.

For some crops, the mapping can be performed through visual analysis of satellite images. The visual mapping, however, has some limitations in what refers to the comprehensiveness and the timeliness of the estimation because, besides being a repetitive and tiresome work, it depends on image products (transparencies and photos) which take more time to produce.

The conventional method for digital crop mapping is the (per pixel) statistical image classification. Although it has been used successfully in several experiments, several limitations that are intrinsic to the method make it difficult to employ the method in large areas for operational systems. These difficulties are

due to several causes like the similarities of spectral signatures of different ground cover, poor ground information, the need that the training conditions be as close as possible of the classification conditions, etc.

To overcome the limitations of the traditional approach, various knowledge-based information systems have recently been proposed for the mapping and classification of satellite images (Goldberg et al. 1987, Wharton, 1987, Mason et al. 1987). This paper discusses some requirements for information systems designed to estimate crop area for Brazilian summer crops and proposes a knowledge-based system intended to meet these requirements.

The organization of this paper is as follows. Section 2 discusses some requirements for an operational system for crop area estimation. Section 3 proposes an organization for an information system intended to meet the requirements described in Section 2. Finally, Section 4 describes the development strategy and some of the objectives of the proposed system.

2. SYSTEM REQUIREMENTS

There are some mandatory requirements for a system intended to estimate crop planted area, specially in the case of summer crops. Basically, the system should:

- i. be integratable with existing crop statistics gathering systems;
- ii. not depend on images from a specific date;
- iii. work satisfactorily even in the presence of incomplete or not totally correct data;
- iv. require few computer and human resources.

Requirements (i) and (iv) are necessary to guarantee the usability of the system, while requirements (ii) and (ii) are derived from existing restriction on the availability of data. Each of these requirements is discussed below.

Integration with existing systems

One should be able to use the information provided by the system in conjunction with information obtained through conventional means. This implies, for example, that area estimates be produced respecting the same boundaries used in the conventional system based in stratification and sampling.

Satellite image independence

For every crop there is a period in its growth when it is easier to distinguish it from the other crops or other objects in the satellite image. However, the summer in most agricultural areas of Brazil is also the rainy season. It means that it is very unlikely to obtain an image completely free of clouds at the most appropriate time. Therefore, most probably, the coverage of the region of interest will have to be obtained with several images taken at different dates.

Incomplete or incorrect data

Errors in the input data of the system (maps, census data and even images) are inevitable or are too costly to be avoided. The system should be able to analyse the data and subject them to consistency checks. Also in many situations the data are missing or lacking the desired precision. In this case the system should be able to infer the desired information or replace it with information obtained by an equivalent source.

Economy of resources

It is a truism to say that material and human resources are usually scarce. This requirement, however, is hard to achieve when it involves the processing of

satellite images because of the sheer volume of the data. More specifically, the system should be able to map a state like Parana or São Paulo with a mini-computer or super micro-computer because these will be the computing resources available. Also the system should be able to operate and should not require special knowledge either in computing or remote sensing.

Besides the mandatory requirements, it is desirable that the system should be expandable to provide other kinds of information. It should be easy to add new functions to the system in a natural way. For example, it should be possible to add crop yield models, that together with area, could give estimates of the total agricultural production. Also it should be easy to accomodate new data sources like new satellite images. The system should also evolve into an integrated agricultural and land use information system. This would not only improve the usability of the system but also protect the investment made in its construction.

3. SYSTEM PROPOSAL

To satisfy the mandatory as well the desirable requirements, it is proposed a "knowledge-based system" for crop area determination. The central idea of this system is to integrate all existing data sources, like satellite images, maps (topographic and others), census data and meteorological data. The knowledge-based information system for crop area estimation is organized into two large modules, that act sequentially: the Segmenter and the Interpreter. Figure 3.1 shows the overall structure of the system.

As an intermediate product, the system produces a "segmented image" which is a non-pictorial representation containing of a list of regions with possibly different land cover. The other input to the Interpreter are models of spectral response and geometric properties of objects that may be present on the scene and tables containing historic and current data of the region under analysis.

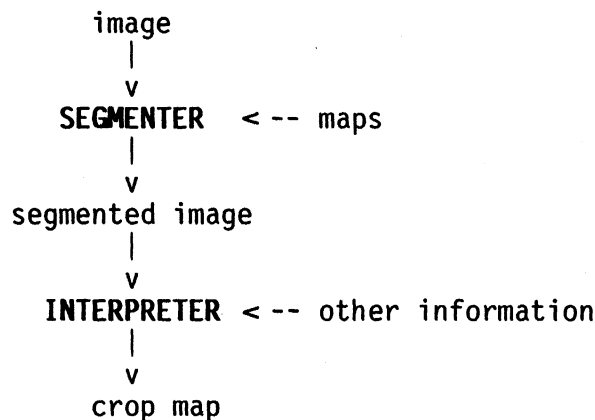


Figure 3.1. Crop area estimation information system.

The Segmenter is composed of three subsystems: the naïve segmenter, the border and line finder and the informed segmenter, as shown in Figure 3.2.

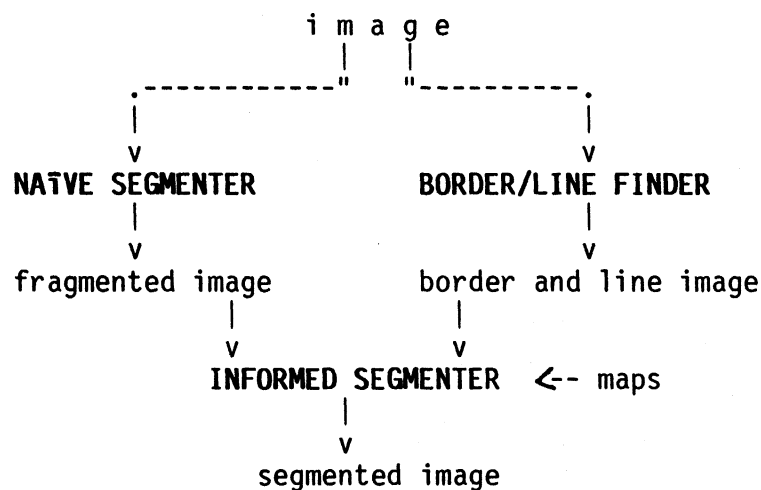


Figure 3.2. Structure of the Segmenter module.

The Naïve Segmenter and the border and line finder use only information present in the image. The segmentation obtained by the Naïve Segmenter is refined by the Informed Segmenter that will use both the output of the Border/line finder and information contained in maps of the region.

Both the Informed Segmenter and the Interpreter will be organized as blackboard systems. A blackboard system is composed of several independent "experts", integrated through a global structured memory area - the "blackboard" (Erman et al., 1980). The blackboard model was chosen over the classical expert system structure, composed of working memory, inference engine and knowledge base because of two inflexibilities of the classical model (Engelmore and Morgan, 1988), in which:

- i. The ordering of the rules determines the control structure;
- ii. The nature of the inference engine determines the knowledge representation.

By contrast, in the blackboard model the knowledge is divided into modules (KS), with separate inference engines. In this way there is no intrinsic requirement for homogeneity in the many different representations and control strategies. The communication between various KSs is made through a global memory area: the blackboard. The blackboard is usually divided into levels corresponding to different abstraction levels. Also in the blackboard model control is neither top-down or bottom-up; it is opportunistic: a KS is activated when it is needed, through a scheduling mechanism.

Some of the knowledge sources that will be present in the Segmenter and Interpreter are:

- i. Spectral Response Experts, that know about factors that influence the spectral response (like topography, illumination)
- ii. Geometry Experts, which include the Shape Expert, the Texture Expert, and the Spatial Arrangement Expert.
- iii. Map Experts, that know how to analyze various types of maps, like topographic maps, soil maps, etc.
- iv. Land Cover Experts, that know facts about all possible objects that may be present in the image crop and other, like how the spectral response varies with

time, for each time of the year, which is the likely spectral and geometric characteristics of the objects, etc.

Image independence will be achieved using images from several dates. In this way, a crop will be mapped based on its "spectral history", i. e. how its satellite image appearance evolved with time, rather than on spectral characteristics of a specific date. The system will store models of how the spectral responses of the several objects likely to be present in the scene varies with time and other factors. These models together with information obtained from independent sources will give the system the ability to map crops even with incomplete or erroneous data.

4. CONCLUSIONS

The construction of a system like the one proposed in the previous section represents a great effort, likely to take several years of work. The strategy adopted is first to construct a prototype for a limited region and with some of the experts and then to incrementally add functionality to the system. The blackboard model chosen is ideally suited for this time of development.

The main objectives of the prototype system are:

i. To develop a system to estimate planted area of summer crops that will be able to integrate satellite image information with several existing data sources.

The prototype system will provide the motivation to conduct studies in the direction of automating the analysis of satellite images. The knowledge obtained in the construction of this system will enable researchers to design more ambitious systems both in the area of crop mapping and in other applications such as environment monitoring.

ii. To study the effectiveness of the various data sources in the task of mapping summer crops.

It is difficult to tell a priori which information sources are more relevant for deriving area information for the crops. The system will permit to experiment with various data sources. It should be possible, for example, to run the system without some of the information.

iii. To develop models that relate the spectral response of objects in the scene to other variables.

It is not totally understood the interrelationships existing among the various factors that affect the spectral response, i. e., how the crop "looks" in the image, specially for summer crops in tropical latitudes. These models, once built, will help greatly in the understanding of satellite images and, therefore, increase their usefulness.

iv. To help to specify a comprehensive system for agricultural and land use information.

The prototype system to be developed will cover only a relatively small region. It is expected, however, that the information to be gained in designing and implementing such system will enable the developers to specify a larger, more comprehensive system for agricultural and land use information in a region, state or even country level. More specifically, it should be possible to answer questions like these: how much computing power is needed for a large system? how much training is required to operate such a system? how accurate is the information provided? how this accuracy is dependent on the several factors involved?

Several parallel studies have to be conducted in order to provide for the information needed in the system. One of these is "model development" which will cover basically models that characterize, for example, (i) summer crops spectral responses from the time the crops are planted until they are harvested; (ii) spectral responses of other objects that may be present in the image and (iii) how the spectral responses are affected by meteorological factors.

5. REFERENCES

Engelmore, R. and Morgan, T., editors, Blackboard Systems, Addison-Wesley, 1988.

Erman, L. D., Hayes-Roth, F., Lesser, V. R. and Reddy, D. R., The Hearsay-II speech understanding system: integrating knowledge to resolve uncertainty, ACM Computing Surveys 12(2):213-253, 1980.

Goldberg, M., Goodenough, D. G., Alvo, M. and Karam, G., "A hierarchical expert system for updating forestry maps with Landsat data", Proceedings of the IEEE, vol. 73(6):1054-1063, June 1985.

Mason, D. C., Corr, D. G., Cross, A., Hogg, D. C., Lawrence, D. H., Petrou, M. and Tailor, A. M., "The use of digital map data in the segmentation and classification of remotely-sensed images", International Journal of Geographical Information Systems, vol. 2(3):195-215, 1988.

Wharton, S. W., "A spectral-knowledge-based approach for urban land-cover discrimination", IEEE Transactions on Geoscience and Remote Sensing, vol. GE-25(3):272-282.