Chapter IV
Image Mining:
Detecting Deforestation Patterns Through Satellites

Marcelino Pereira dos Santos Silva
Rio Grande do Norte State University, Brazil

Gilberto Câmara
National Institute for Space Research, Brazil

Maria Isabel Sobral Escada
National Institute for Space Research, Brazil

ABSTRACT

Daily, different satellites capture data of distinct contexts, which images are processed and stored in many institutions. This chapter presents relevant definitions on remote sensing and image mining domain, beyond referring to related work on this field and to the importance of appropriate tools and techniques to analyze satellite images and extract knowledge from this kind of data. The Amazonia deforestation problem is discussed, as well INPE's effort to develop and spread technology to deal with challenges involving Earth observation resources. An image mining approach is presented and applied on a case study, detecting patterns of change on deforested areas of Amazonia. The purpose of the authors is to present relevant technologies, new approaches and research directions on remote sensing image mining, demonstrating how to increase the analysis potential of such huge strategic data.

INTRODUCTION

Motivation

Data acquisition and storage technology progress has led to a huge amount of data stored in repositories, which grow fast. Among increasing and relevant data acquired and processed, there is a strategic segment: satellite images, also known as remote sensing images.

The search for less expensive and more efficient ways to observe Earth motivated man in develop-

ing remote sensing satellites. They are currently the most significant source of new data about the planet, and remote sensing image databases are the fastest growing archives of spatial information. The variety of spatial and spectral resolutions for remote sensing images ranges from IKONOS 1-meter panchromatic images to the next generation of polarimetric radar imagery satellites. Given the widespread availability of remotely sensed data, many government and private institutions have built large remote sensing image archives.

The US National Satellite Land Remote Sensing Data Archieve, managed by USGS EROS Data Center, hosts 1.400 terabytes of satellite data gathered during 40 years. Satellites, like Terra and Aqua (NASA), generate 3 terabytes of images every day. The Brazil's National Institute for Space Research (INPE) holds more than 130 terabytes of image data, covering 30 years of remote sensing activities which are available on a database with free online access.

Actual society problems demand smart exploration of the vast and growing remote sensing data. There is a need for understanding relevant data and use it effectively and efficiently. Although valuable information is contained in image repositories, the volume and complexity of this data makes difficult (generally impossible) for human beings extract strategic information (knowledge) without appropriate tools (Piatetsky-Shapiro, Djeraba, Getoor, Grossman, Feldman & Zaki, 2006).

Data mining research has enabled powerful tools, new technologies and challenging techniques for relevant data domains. However, large image datasets need specific analysis resources and smart techniques and methodologies. The availability of huge remote sensing image repositories demands appropriate resources to explore this data.

A vast remote sensing database is a collection of landscape snapshots, which supplies a single opportunity to understand how, when and where changes occurred in the world. When such rich data is not analyzed, or it is done inefficiently, relevant information to understand complex processes and help solving challenging problems is wasted.

General Perspective and Objectives of the Chapter

In this chapter, which extends previous work (Silva, Câmara, Souza, Valeriano, & Escada, 2005), the authors intend to present relevant definitions on remote sensing and image mining domain, beyond presenting related work on this field and the importance of appropriate tools and techniques to explore satellite images and extract strategic knowledge from this kind of data.

They also discuss the Amazonia deforestation problem to demonstrate, through an image mining process, the strength of this approach to identify patterns and fight against the increase of affected areas in this forest. Developed technologies to support the process will be presented, providing an overview of methodologies, tools and techniques involved in research efforts.

Future trends and conclusion will bring reflection elements to consider classical and new mining resources to deal with challenging demands, citing limitations and also revealing directions to new research initiatives and relevant problems.

REMOTE SENSING AND IMAGE MINING

Broad Definitions

The first operational remote sensing satellite (LANDSAT-1) was launched in 1972, since then there has been a large worldwide experience in data gathering, processing and analysis of remotely sensed data. According to Canada Centre for Remote Sensing (2003), *remote sensing* is the science (and to some extent, art) of acquiring information about the Earth surface without actually being in contact with it. In