APPLYING SATELLITE IMAGES CLASSIFICATION ALGORITHMS FOR SOIL COVER AND GEORESOURCES IDENTIFICATION IN NOVA LIMA, MINAS GERAIS - BRAZIL

Vladimir Diniz Vieira Ramos
Universidade Federal de Minas Gerais
Departamento de Cartografia
Av. Antônio Carlos, 6626, Belo Horizonte, Minas Gerais. CEP 31270-901
vladbh@terra.com.br

Pedro Pina
Instituto Superior Técnico – Portugal
Departamento de Engenharia de Minas e Georecursos
Alameda de St. António dos Capuchos, 1 Cep: 1169-047 Lisboa, Portugal
ppina@mail.ist.utl.pt

ABSTRACT - This paper applies image classification techniques of images from ASTER and Landsat. The images are from Belo Horizonte and Nova Lima municipal areas specially the borders between them which presents expressive changes in dwelling profile and strong mining activity. This location is a metropolitan county area under an increasing influence of the principal city Belo Horizonte, it is going through a starting process of conurbation, since the expansion of the south neighborhoods and condominiums. The cities expansions are accompanying the road BR040, pathway to Ouro Preto city. The presented procedures were developed by the SPRING processing resources and images classification, that is a free geographical information software developed by The National Institute of Spatial Research - Brazil. The objective is to verify the ASTER and Landsat images capacity to classify the ground cover and the region georesources through the SPRING algorithms and compare the results. Through bands with spatial resolution of 15 and 30 meters of Aster and 3 bands of Landsat ETM of 30 meters, it was tested the SPRING classifiers efficiency from a simple composition and another from the main component. The bands from 15 to 30 meters were re-sampled in SPRING to allow the images classification and the components generation. The methodology is based on classifiers application on both images ASTER and Landsat. It was applied two supervised classification methods in each image, the MAXVER and the BATTACHARA. The Main Components method was utilized on ASTER images in order to get information from the largest number of bandwidths, summarizing the information in three bandwidths to use in the classification. A RGB543 composition was adjusted on Landsat images a to compare the efficiency of both images. While comparing the results of each composition for both methods defined by the methodology, the classifier using the segmentation process and the Principal Component generated bands has obtained the best result according the report generated by SPRING classification samples. The Principal Components RGB image showed the

ability to identify subclasses of all classes defined in the methodology, which hindered the gathering of samples for classification, but has show the good ASTER pictures quality to identify a more detailed level of soil occupation. Thus, the ASTER images potential to map and monitor the use and occupation of various regions was confirmed. It was associated with the Principal Components Process, both in monitoring the use and occupation and in the detection of georesources. The combination of classification methods and Principal Components has been offering greater capacity to distinguish among classes, which makes possible to identify subclasses in all classes defined by the methodology. This will make us able to study how to identify these more detailed subclasses. The growing need for more detailed classification plans has assigned to classifying tools a huge responsibility in databases construction, thus we must be aware to this method capacity and limitations. Tools must be built to provide more reliable databases. Understanding the spatial phenomena behavior with greater precision allows us to manage and direct actions that minimize the consequences of human changes over the environment and allows us to manage the available natural resources in a sustainable way.

1. INTRODUCTION

Geoprocessing has been a tool of great importance to resources management and administration at different scales. Understanding the behavior and distribution of surface resources and surface phenomena enables planning actions that lead to the full development of a region, and also contributes to minimize the impacts that economic development brings to the area and to its available natural resources.

This research is linked to studies of FARO-ALFA Network - Advanced Training in Ornamental Rocks and Geoprocessing. The network is coordinated by the University of Bologna and includes the IGC-UFMG participation. Among the participating universities we had this one opportunity to study and exchange knowledge, particularly with the Instituto Superior Técnico – Portugal (IST). One of the goals of the network is to explore geoprocessing techniques and methodologies that provide information for Ornamental Rocks localization and characterization, as well as providing proposals for environmentally sustainable exploration. As part of our collaboration in the study group, we made knowledge exchanges about our procedures for images classification in IST and some results for the application of classifiers algorithms are presented in this paper. The classifiers are employed in the management of a county envirocnmental landscape characterized by extensive exploration activity and significant territorial transformation.

Geoprocessing is an important tool in the management and development of environmental and anthropic activities. It is also known by one of its main components, the GIS (Geographic Information Systems). From these complex systems it is possible identify and space a wide range of physical and social phenomenon, as long as this tool allows one to organize and systemize the data, allowing detailed information analysis that resulting in planning studies, and further, predictive studies. According to Xavier da Silva (1999), since the information is systematized and specialized, it comes up the ability to review, and data gets a geographical nature.

The space is at constant transformation, as it suffers the influence of various physical or social factors. According to Corrêa (2002), if the space is a social product, the actions that change it occurs by various social agents, linked directly or indirectly, and these agents may change by following trends from inside or outside, causing new local axis and thus drastic changes in soil use.

In the case of intensive spatial changes, Belo Horizonte Metropolitan Region (RMBH) is, in Minas Gerais Estate, the best example. In different periods of time the RMBH was affected in different areas of expansion, but the focus of this work is the most current axis of expansion: the south route toward the city of Nova Lima.

It is extremely important to define monitoring and interpretation methods for spatial changes, especially in the significant municipalities like Nova Lima. because only then it is possible to define a sustainable management to minimize the impacts of urban sprawl of Belo Horizonte on Nova Lima and the county self expansion.

The Geoprocessing tools allow a systematic analysis of the information. One of the databases sources that allow such assessments at different scales is the remote sensors from satellite images, of different spatial and spectral resolutions.

On the Remote Sensing progress, Landsat images were of great assistance to build databases and analysis from raster images, especially for the Brazilian territory. Meanwhile, a new product is drawing attention from researchers due to the radiometric ability and spatial resolution: the ASTER images.

This papers main objective is to compare the efficiency of ASTER and Landsat images in the land cover classification and the identification of Nova Lima georesources. It is an objective to compare the classification algorithms in soil cover studies and to establish which SPRING classification software algorithm presents the best result in interpreting the images. The parameters will be able to be defined: what type of image and the best methodology to identify the cover classes the interest soil, using the municipality of Nova Lima area as a benchmark.

2. DESCRIPTION OF AREA

Nova Lima county borders Belo Horizonte county, its headquarters is 22 km from the Belo Horizonte (Minas Gerias Estate capital). Its area is about 429 km^2 and the population density is 149.6 habitants/km^2 (Figure 1).

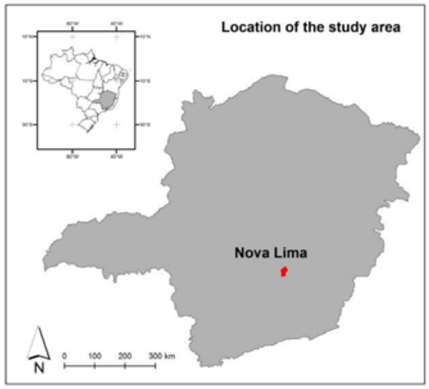

Figure 1: Studied Area Localization

The city has a tradition of mining and was founded in 1891. Costa and Pacheco (2004) explain that the core city of Nova Lima was consolidated during the setting of the mining activity, where the principal activity was the gold extraction. This situation has creating a larger social segmentation with the occurrence of several social classes.

According to its location like a metropolitan county area, it is under an increasing influence of the principal city Belo Horizonte. A process of conurbation is starting due to the expansion of the south neighborhoods south condominiums, which are following the axis of the road BR040, output to Ouro Preto city. This expansion is reflected in its population growth (Table 1) and also on its development (Table 2).

Table 1: Population Growth

Period	Growth Rate %
1970-80	21,27
1980-91	27,11
1991-96	8,7
1996- 2001	12,88

Source: IBGE

Table 2: Urbanization Degree

Years	% Urban Population	% Rural Population
1970	80,54%	19,46%
1980	85,03%	14,98%
1991	84,04%	15,96%
2001	97,91%	2,09%

Source: IBGE

3. MATERIALS AND METHODS

The remote sensing database is formed by two scenes, an ASTER 2002 orbit/point 131/602 covering Nova Lima county area, and a Landsat ETM 2007 orbit/point 218/74, the area of study corresponding to the coordinates -44° 01’, -20° 13’ / -43° 47’, -19° 55’.

The ASTER scenes are Level 1B Data type products that, according to Abrams and Hook (2002), are products of the same type of Level 1A Data, but with applied radiometric and geometric parameters and are stored in HDF file with the metadata.

The software used is the SPRING that, according to INPE (2004), has the following characteristics: runs as a database without geographical borders and supports large data volumes (without limitations of scale, projection and zone). It manages both vector data and raster matrix data, it performs the remote sensing data integration and provides a powerful working environment through a combination of menus and windows with a spatial language easily programmable by the user.

Through bands with spatial resolution of 15 and 30 meters of Aster and 3 bands of Landsat ETM of 30 meters, the SPRING classifiers efficiency was tested from a simple composition and from the main component. Using the SPRING, the bands were resampled from 15 meters to 30 meters to allow the images classification of images and the components generation.

According to Crósta (1992), the Analysis by Principal Components (APC), also known as Transformation by Principal Components or Transformed by Karhunen-Loève, it is well fitted to process images of multispectral sensors. Originally developed in the context of electrical engineering with the purpose of processing electronic signals to remove noise, the APC was incorporated in remote sensing. The results of APC are difficult to generalize for certain applications. The results are dependent of the scene spectral characteristics. Crósta (op. cit) explains that the results obtained in an area do not necessarily are repeated in another:

> *The generation of Principal Components is a spectral rotation technique that reduces or removes spectrum redundancy. While considering the two-dimensional attributes space of two images A and B, it specifies the distribution centroid position, the distribution gravity center, not the geometric center. This distribution center is a point where an average pixel of intensity, if any exists, would be positioned. As soon as the point is known, it is made a calculation of the spread along each axis, the variance. It is measured by the standard deviation square of the intensities histogram for each band and gives a measure of the contrast of each spectral band. The correlation between each pair of image bands is specified, using the covariance parameter. This concerns two specific variances and describes how the information contained in a pair of bands is common to both. The*

value of the covariance may be positive, when data are positively correlated and negative when they are negatively correlated.

It is important that the new bands do not present any correlation between them. It generates a new set of images whose has main bands, much smaller in number, that presents a concentration of total information in the original channels.

The classes were defined from previous experience with classification, area knowledge and high-resolution images available for classification validation. It was used a Quickbird image taken in 2007 of the Nova Lima County. It was also considered the study objective in the area, the characteristics and history of formation to define the classes group.

It was also applied a band equalization. To achieve the equalization it was used the SPRING arithmetic tool C= gain*A+offset, with 127 for average and 40 for standard deviation, which are the best parameters to highlight Landsat images. This can be expressed by the equation:

Is= (Ie-average (Ie) *A + average (Iref)
Is= Ie*gain + offset

Where:

A= $(Vref/Ve)_{(1/2)}$
Is= output image
Offset = Average (Iref) – $(Vref/Ve)_{(1/2)}$ *average (Ie)
Average (Iref) = 127
σ (Iref) = 40
σ(Iref) = $(Vref)_{(1/2)}$

Classes defined for classification:

- Urban area or amended - urban areas, mining areas, urban roads, outcrops, areas degraded by human action.
- Dense vegetation - area with large number of trees, forest remnants, spring areas, riparian forest, parks, reforestation areas.
- Underwood - area with a predominance of shrubs and herbaceous vegetation, pasture areas, cultivation areas, typical areas of Cerrado.
- Water and Shadow – Water mirror areas or shadow predominance.
- Exposed soil - areas where the soil structure is exposed due to natural processes or human action.

To test the images in soil cover classification and region georesources identification it was defined two types of SPRING classifiers. These classifiers were tested in the two images. The classification was made from bands collected from the Principal

Components results and from the RGB 543 Landsat composition, both by supervised methods, with the former using a segmentation procedure preceding the classification process and a pixel-to-pixel classification.

The segmentation process is a combination of smaller regions that have a certain similarity in their spectral behavior. The SPRING parameters were set with similarity of 15 and area of 30 pixels to create the segmentation. It was followed by the choice of targeted images classification method and Battacharya. According to INPE (2004), the Battacharya measure of distance is used in this regions classifier to measure the statistical separability between a pair of spectral classes, i.e., it measures the average distance between the distributions of spectral classes probabilities.

Another method also used in two images was the Maxver, or Maximum Likelihood, which is a process of pixel-to-pixel classification that considers the weighted distances between the average classes gray levels, using statistical parameters, INPE (2004).

4. RESULTS AND DISCUSSION

From the Principal Components method were defined 3 PCs (Principal Components) where most of the information from the 9 bands of Aster were concentrated (Table 3). From these 3 PCs it was created one of the images RGB composition.

Table 3: Principal Component self-values generated

PC	Self-values	Percentage
P1	**14148.45**	**84.82**
P2	**1552.52**	**9.31**
P3	**665.00**	**3.99**
P4	158.16	0.95
P5	53.89	0.32
P6	42.46	0.25
P7	26.14	0.16
P8	18.74	0.11
P9	14.85	0.09

The second image is the product of an ETM RGB 543 Landsat composition, where the bands were highlighted from gain values and pre-defined offset.

The SPRING produces reports about the classes obtained in the classification process. From these reports and from the visual analysis final results generated by the classification process it is possible to answer which method produces the best result within the classes defined in the methodology.

While comparing the results of each composition for the two methods defined by methodology, the classifier that used the segmentation process and the Principal Component generated bands brings the best results according to the report generated by ponent generated bands brings the best results according to the report generated by SPRING classification samples. Through a visual analysis of the classification result, this composition was the best in identifying the occupation diversity of the land area (Table 4).

Table 4: Performance Comparison

Image	General Performance	Average Confusion
RGB Principal Components Aster - Battachara	**99.8%**	**0.11%**
RGB Principal Components Aster - Maxver	97.08%	2.92%
RGB Landsat - Battachara	97.47%	2.53%
RGB Landsat - Maxver	93.91%	0.32%

In the classification process we are able to see that the exposed soil classes, water and shade and dense vegetation are those with less confusion or even none if compared to the other ones. However, the vegetation classes have high levels of confusion with the urban area and the degraded areas. This mixture varies from method to method, being much smaller while using the segmented image, and increasing while using the pixel-to-pixel method, which has hardened the classification.

The Principal Components RGB image showed the ability to identify subclasses of all classes defined in the methodology, which hindered the samples gathering for classification. Otherwise, it has shown the good performance of Aster pictures to identify a more detailed level of soil occupation. Thus, it is confirmed the ASTER images potential to map and monitor the usage and occupation of various regions, while associated with the Principal Components Process, not only in monitoring the usage and occupation but in the detection of georesources (Figures 2 and 3).

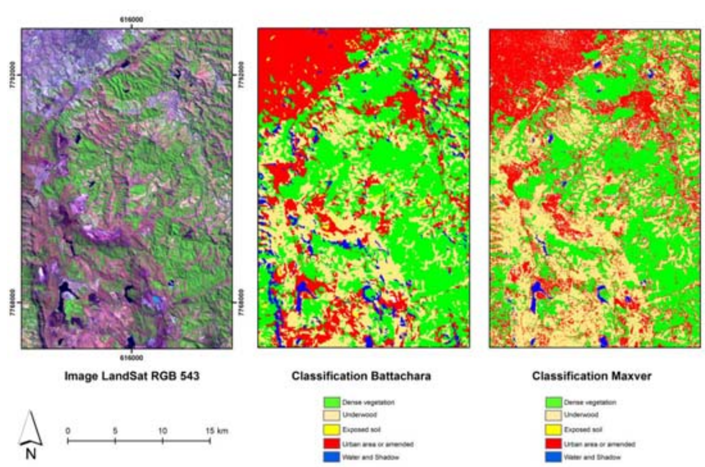

Figure 2. Landsat Classifying Comparison

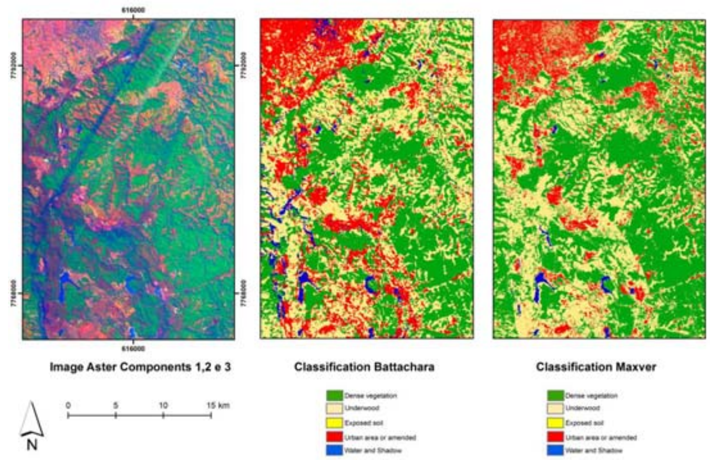

Figura 3. Aster Classifying Comparison.

5. CONCLUSION

Used as tools for acquiring databases, satellite images have shown to be, every day, a great source of information as they progress their spectral and temporal spatial resolution. In this paper, the Aster images have demonstrated their ability using the SPRING classification methods. Within an area of great complexity and achieving an acceptable accuracy, it was possible to identify the land occupation of two counties on its main elements.

The combination of classification methods with the Principals Components has been offering greater capacity to distinguish among classes, which makes possible to identify subclasses in all classes defined by the methodology. This enables us to study how to identify these more detailed subclasses.

Once we have verified the method benefits for the identification of land occupation in a certain area, it may be possible to replicate it in other areas where there are images of the same pattern. It is important to consider that the whole classification process was done in a certain image type, which was acquired in a specific date and has also received an enhancement treatment, pointing out that all these aspects vary from image to image.

The growing need for a more detailed classification plan has assigned the classifying tools a huge responsibility in databases construction, thus we must be aware to this method capacity and limitations. We must build tools to provide more reliable databases, because understanding the spatial phenomena behavior with greater precision allows us to manage and direct actions that minimize the consequences of human changes over the environment and allows us to manage the available natural resources in a sustainable way.

6. REFERENCES

ABRAMS, Michael HOOK, Simon, **ASTER User Handbook Version 2**, Jet Propulsion Laboratory 4800 Oak Grove Dr. Pasadena, CA 91109 Bhaskar Ramachandran EROS Data Center Sioux Falls, SD 57198

COSTA, Geraldo. M.; PACHECO, P. D. Planejamento urbano no ambiente metropolitano: o caso do município de Nova lima na Região Metropolitana de Belo Horizonte. In: VII SEMINÁRIO DA HISTORIA DA CIDADE E DO URBANISMO, 7, 2004, Niterói**. Anais do VII Seminário da História da Cidade e do Urbanismo**, Niterói: Escola de Arquitetura e Urbanismo da UFF, 2004. p. 1-12.

CORRÊA, Roberto Lobato. **O Espaço urbano**. São Paulo: Ática, 2002. 94p.

CRÓSTA, Álvaro Penteado. **Processamento Digital de Imagens de Sensoriamento Remoto**. Campinas, SP: IG/UNICAMP, 1992. 170p

IBGE, **Pesquisa de Informações Básicas Municipais** – 2001.

INPE, INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **SPRING:** manual do usuário. São José dos Campos: INPE. 2004

XAVIER-DA-SILVA, Jorge. **Acesso a dados e transformações preparatórias à análise ambiental**. Rio de Janeiro: Lageop, 1999. 12 p. Apostila do Curso de Especialização em Geoprocessamento - Midia CD-rom).