

Tool Development for Manual Classification by Decision Tree

Fernanda Seidl ¹
Ricardo Pontes Bonfiglioli ¹

1 Fundação de Ciências, Aplicações e Tecnologias Espaciais - FUNCATE
Avenida Dr. João Guilhermino - 429 - 13° Andar - 12.210-131 - São José dos Campos - SP,
Brasil
fernanda.seidl@funcate.org.br
ricardo.bonfiglioli@funcate.org.br

Abstract. Due to difficulties in classifying images with high spectral correlation, there was the need to develop a tool for manual classification by decision tree based on features extracted from polygons related to an image of interest, so that these polygons can be separated by spectral, spatial and textural properties. For that, it was also necessary to develop a tool for feature extraction, which can be associated with the polygons table, allowing the polygons classification according to the conditions set by the user with the features. Therefore, it was got a classification result to confirm the value of these tools. It was observed the importance of the functionality increase for image analysis, such as features graphics generation and histogram of an interest area.

Keywords: features extraction, classification, decision tree, classes conditions.

1. Introduction

Cerrado is a Brazilian biome which presents a very complex process of visual interpretation even for supervised classifiers, being necessary to obtain spatial, spectral and textural features associated with regions (segments), from various products like Normalized Difference Vegetation Index (NDVI), Linear Spectral Mixture Models (MLME) and the image. The features obtained may be interrelated in order to create decision rules for generating the final legend of the project.

In order to reduce the complexity of classification of this biome in *TerraClass Cerrado* project (Rodrigues, 2014) - partially developed in Amazon Regional Center (Portuguese: *Centro Regional da Amazônia*; CRA) on National Institute for Space Research (Portuguese: *Instituto Nacional de Pesquisas Espaciais*; INPE), it was proposed the development of a classification method by decision tree and features extraction, which can be used in classification.

These tools will be available in the versions *TerraAmazon*¹ 4.5.0 and *TerraView*² 4.3.0. *TerraAmazon* is a system used to monitor the Amazon deforestation (Freitas, 2007). These systems use *Terralib*³ library, an open source GIS classes and functions library (Câmara, 2014).

2. Methodology

The proposed tools have been developed in the DIP plugin (Digital Image Processing), based on *TerraLib* 4.3.0. The diagram in Figure 1 shows the necessary procedures for classification by decision tree.

¹ TerraAmazon: <http://terraamazon.org>

² TerraView: <http://www.dpi.inpe.br/terraview>

³ TerraLib Library: <http://www.terralib.org>

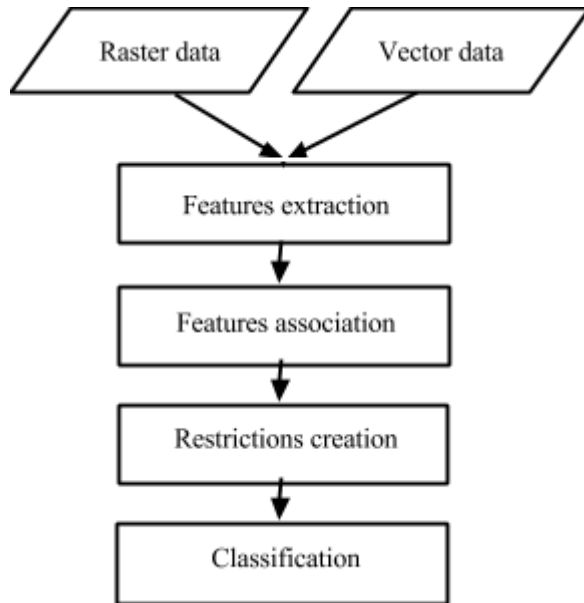


Figure 1 - A classification process.

A raster and a polygonal vector data Geo-referenced from the same region are necessary as input data. The features selection (Figure 2) to be used for characterizing the polygons is done with the choice of features based on the region/biome as appropriate to the project.

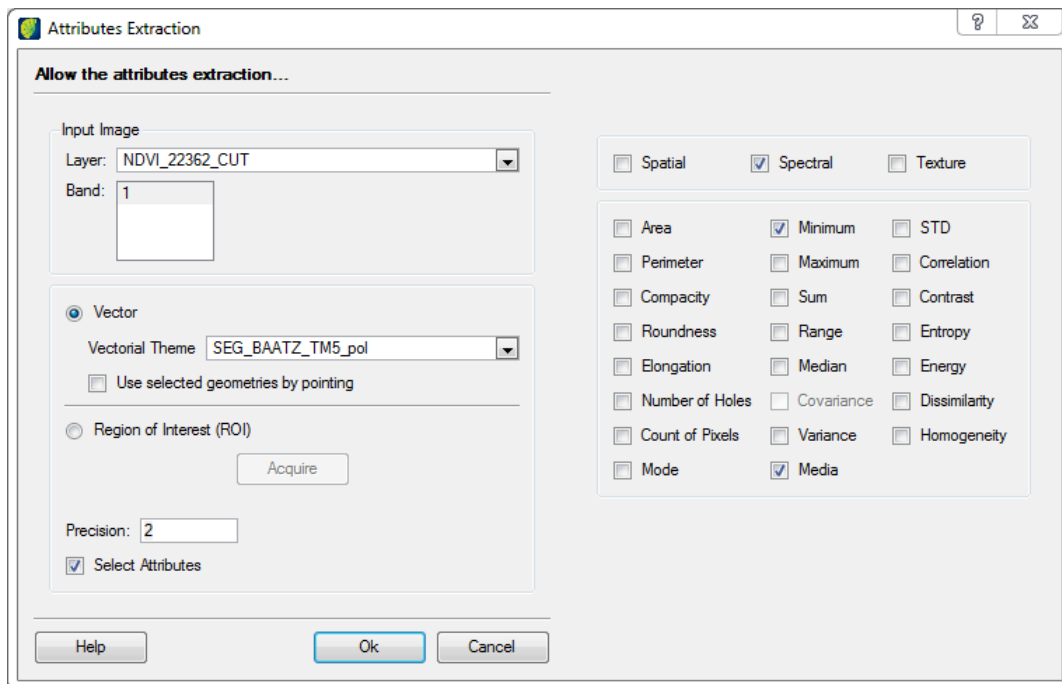


Figure 2 - Features selection.

The features present in the extraction tool are shown in Table 1.

Table 1 - Spatial, spectral and textural features.

Spatial	Spectral	Texture
Area	Mode	Variance
Perimeter	Minimum Value	Media
Compacity	Maximum Value	Standard Deviation
Roundness	Sum	Correlation
Elongation	Range	Contrast
Number of Holes	Median	Entropy
Count of Pixels	Covariance	Energy
	Variance	Dissimilarity
	Media	Homogeneity
	Standard Deviation	

For each polygon the values of each feature are calculated. The highlighted features in Table 1 were implemented; others were already available in *TerraLib*. The formulas for the spatial features and the texture variance are found in the ITT-VIS (2008) and Hall-Beyer (2012), respectively. After analysis of the results (Figure 3), it is possible to associate automatically the features to the polygons in the database.

object_id	Min_value_B1	Mean_B1
1	105.00	46.78
2	169.00	147.82
3	151.00	114.74
4	108.00	67.17
5	101.00	43.73
6	50.00	31.14
7	85.00	29.53
8	108.00	49.31
9	79.00	39.46
10	145.00	92.73
11	139.00	117.94
12	105.00	31.17

Centralize polygon when selected
 Save table
 Join columns to original layer
 Create new layer
 Suffix:

Figure 3 - Analysis of result

In the next step, in the classification by decision tree tool, the user defines the conditions of classes and sub classes, based in the experience acquired on the image analysis and in the calculated features of the previous step. This is shown in the example of Figure 4.

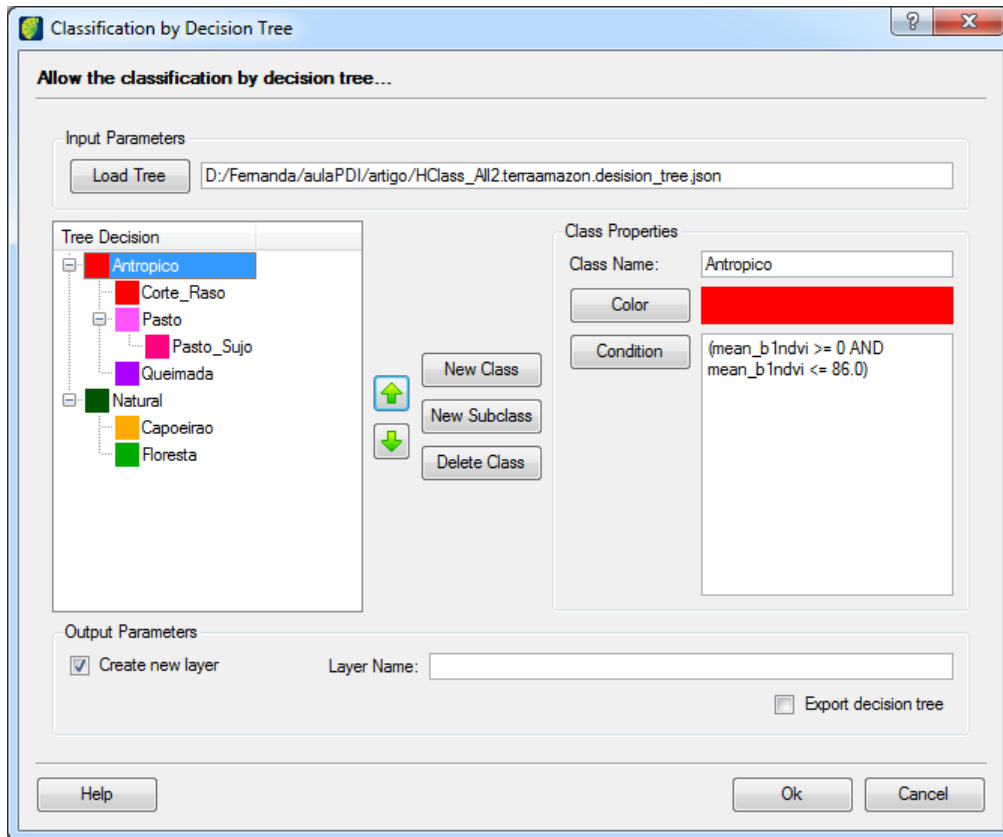


Figura 4 - Classification by decision tree.

After defining the decision tree, the polygons classification is performed according to the classes already defined, generating a vectorial data using a color legend associated with their classes. The decision tree can be exported for use in future classifications.

3. Results and Discussion

From the clipping of a Landsat TM 5 image, taken from orbit 223/062 (Figure 5) in a region of Pará state, northern of Brazil, the MLME and NDVI products were obtained. With these products the features of average and minimum value were extracted, as shown in Figure 3.



Figure 5 - Clipping of an image to classification

Classes and their conditions were defined based on the features and image analysis, as shown in Figure 6.









	Antrópico	(mean_b1ndvi >= 0 AND mean_b1ndvi <= 86.0)
	Corte_Raso	(min_value_b1mlme >= 146)
	Pasto	(min_value_b1mlme < 146)
	Pasto_Sujo	(min_value_b1ndvi >= 82 AND min_value_b1ndvi <= 140)
	Queimada	(min_value_b1ndvi < 82)
	Natural	(mean_b1ndvi >= 86.1 AND mean_b1ndvi <= 158.8)
	Capoeirão	(min_value_b1ndvi >= 82 AND min_value_b1ndvi <= 140)
	Floresta	(min_value_b1ndvi > 140)

Figure 6 - Tree Decision

After the creation of the decision tree, classification of polygons was performed, obtaining the result shown in Figure 7.

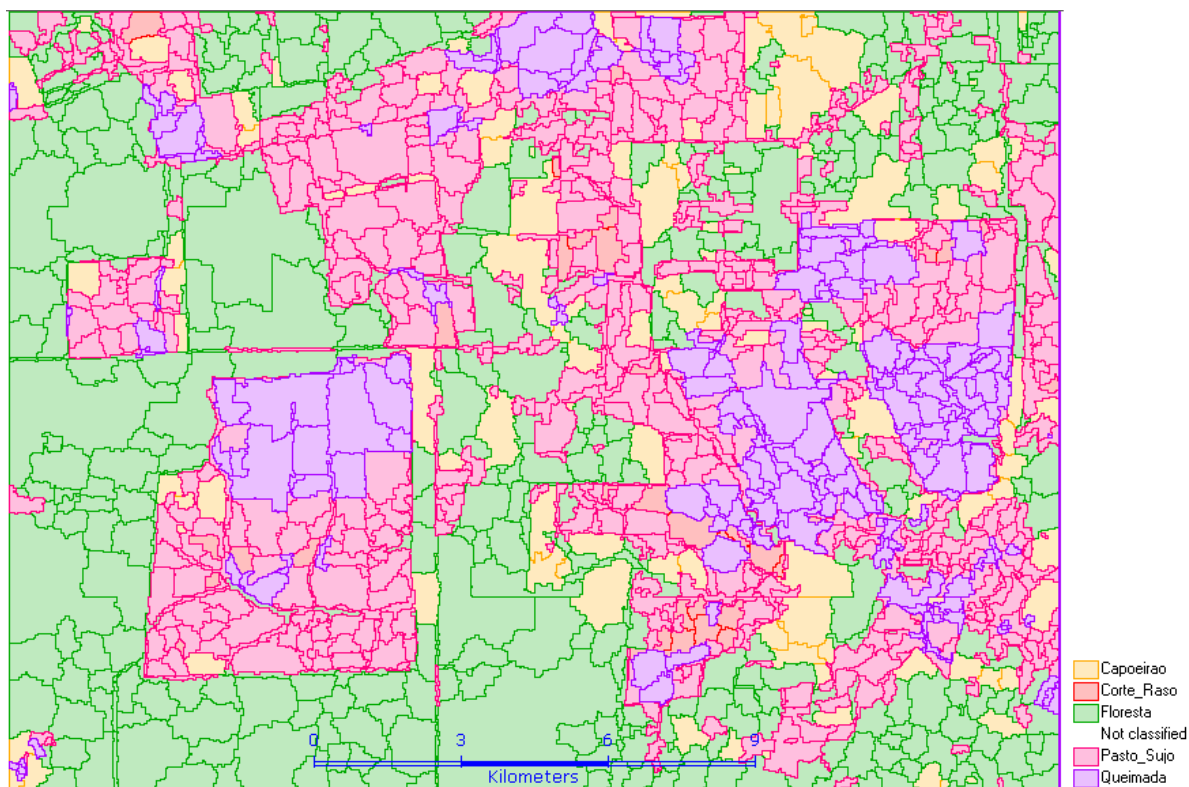


Figure 7 - Result of classification

To establish the level of accuracy of this example, it would be interesting to compare it with other classifications already obtained for this region.

4. Conclusion

The process was very efficient, but requires a wide knowledge of the calculations of the features and of how each feature is related to polygons and image characteristics.

In order to favor the analysis of regions and their features in future work, the visualization of the graph of the features and the histogram of their regions will be added in the tool of extraction features. In the next stage of the tool development, the possibility of adding user-defined operations to devise new features is being planned.

Referencics

Câmara, G.; Vinhas, L.; Ferreira, K. R.; Queiroz, G. R.; Souza, R. C. M.; Monteiro, A. M. V.; Carvalho, M. T.; Casanova, M. A.; Freitas, U. M. TerraLib: An Open Source GIS Library for Large-scale Environmental and Socio-economic Applications. In: HALL, G. B. (ed.). **Open Source Approaches for Spatial Data Handling**. Berlin: SPRINGER, 2008.

Freitas, U. M.; Ribeiro, V. O.; Queiroz, G. R.; Petinatti, M. R. e Abreu, E. S. The Amazon Deforestation Monitoring System: A Large Environmental Database Developed on TerraLib and PostgreSQL. **OSGeo Journal**. Vol. 3 - FOSS4G 2007 Proceedings, Dec. 2007.

Hall-Beyer, M. **GLCM Texture Tutorial**. Available at: <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>, Access: 19 aug. 2014.

ITT-VIS; ENVI **Feature Extraction Module User's Guide**. 2008. Available at: http://www.exelisvis.com/portals/0/pdfs/envi/feature_extraction_module.pdf. Access: 19 aug. 2014.

Rodrigues, N. **Uso e cobertura da terra no Cerrado serão mapeados**. 2014. Available at: <https://www.embrapa.br/busca-de-noticias/-/noticia/1577803/uso-e-cobertura-da-terra-no-cerrado-serao-mapeados->; Access: 10 sept. 2014.

Acknowledgments

The authors would like to thank Luis Waldyr Rodrigues Sadeck and Professor Thales Sehn Körting, PhD, both from INPE, for their help and support.