Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril 2005, INPE, p. 1175-1181.

Aerial photo land cover classification of cerrado physiognomies: detailed or accurate maps

Ernesto Camelo de Castro

UCG – Universidade Católica de Goiás
Caixa Postal 86, 74605-010, Goiânia, GO, Brasil
ernesto@ucg.br

Abstract. The Cerrado Biome in Brazil has a remarkable diversity of vegetation types and physical environments so endangered by the new trends in the country's economic development that it has been included in the Conservation International list of 25 biodiversity "hotspots" on Earth. The Cerrado Biome landscape diversity makes it a very puzzling subject for landscape ecology studies and remote sensing data are the best tool nowadays for land cover data survey. This work uses high-resolution black and white panchromatic aerial photos in an experiment to evaluate the possibilities of this data source for the cerrado physiognomic types identification and detection of human influences on the landscape pattern. The results show that many abuses might have being done with mapping works assumed correct by the simple fact that they use aerial photo coverage where an appropriate accuracy assessment work is not performed.

Kewords: Aerial photo, cerrado physiognomies identification, accuracy assessment, foto aerea, identificação de fisionomias de cerrado, determinação de acurácia.

1. Introduction

The Cerrado, a sub-humid tropical biome located in Brazil, has a remarkable diversity of vegetation types closely related to each other and to the physical environment, and is experiencing substantial new development from advanced agricultural technologies and industry. Only about 20% of the original 1,783,169 km^2 of the Cerrado was estimated preserved in 1999 (Mittermeier et al. 1999). Dias (1990) compiled results from sources back to 1985 to show that, at that time, only 7% of the natural Cerrado landscapes were protected by law, compared to 37% of completely anthropogenic landscapes and 56% with some sort of management for production, including low technology pastures managing for cattle ranching. These and many other environmental threatening facts made the Conservation International include the Cerrado in the list of 25 biodiversity "hotspots" on Earth where biologically rich areas are under the greatest threat of destruction (Mittermeier et al. 1999).

In this scenario remote sensing and geographic information systems (GIS) are considered to be fundamental tools for conservation biology studies (Lunetta 1998, Haines-Young et al. 1993). A very important issue that is not clear so far is what is the scale of resolution and level of detail for landscape patterns classification achievable by the available remote sensing data sources. Despite de fact that the scale issue in remote sensing, GIS and landscape ecology has been of great concern in the scientific society (Quattrochi and Goodchild 1997, Lowel and Jaton 1999, Congalton and Green 1999, Withers and Meentemeyer 1999, Wu and Qi 2000), yet, many misleading studies have being done based on land cover maps created from these data sources without appropriate accuracy. The set up of hierarchical levels of land cover for the cerrado with its association with each remote sensing data type might help to reach a more trustful mapping standard (Castro 2002).

The present study intends to clarify some of the issues present on land cover maps crated based on medium to high-resolution aerial photos. These issues are discussed based on an experiment made on the São Vicente River Valley contained within the Terra Ronca State Park (PETR – "Parque Estadual de Terra Ronca") boundaries.

2. Objectives

Our main objective is to evaluate the efficiency of the photo interpretation methodology, using advanced geographic information system technology, for the landscape pattern definition of the cerrado. The fundamental question is: how the cerrado's landscape patterns can be characterized through medium to high-resolution panchromatic aerial photography or similar scale remote sensing data. It is important to mention that we are also looking for detection, as much as possible, for low impact human interference in the landscape pattern.

The specific questions addressed during the experiment were:

- How the land cover pattern of the cerrado can be detected at these spatial and spectral scales with the use of on screen digitizing advantages;
- How the accuracy varies with physiognomic levels of details;
- What kind of human influences can be detected at an acceptable level of accuracy;

This is only the initial work for a more comprehensive study of the applicability of the most common remote sensing data scales for the Cerrado Biome.

3. Area of study

The state law number 10.879 from 1989 created the Terra Ronca State Park (PETR), aiming the preservation of the fauna, flora and water springs of around 57.000 hectares in São Domingos County, State of Goiás, where natural phenomena generated one of the largest speleological complex in Brazil. Another factor that motivated the creation of the park was that the region has one of the last large areas of the Cerrado Biome, where the low level of human impact allows the observance of its characteristic large diversity of physiognomic types.

However, the region has a relatively old presence of modern human economic activity detected everywhere by subsistence agriculture and ranching activities based on family labor. Therefore, most of the land cover types have some sort of human interference, like the use of the cerrado for cattle grazing, with regular fire setting to favor the more open cerrado physiognomic types that have more dense grass cover.

The São Vicente River Valley delimited our study area in its upper section where it is contained between the “Serra Geral” and the “Serra do Calcário”. The springs from the base of the “Serra Geral” drains this 12,974 ha basin into a cavern on the “Serra do Calcário” base. Although most of it is contained within the PETR boundaries many private properties inside the park were not paid off yet. Therefore, there is still a considerable amount of human impact inside the park boundaries. On the other hand, this area had never had a very intense economic activity. Most of the ranching was based on maintenance of open cerrado areas for pastures with very primitive techniques like the use of fire, which is a sort of cerrado's natural component. On the upper part of the valley, on the top of the “Serra Geral”, there is a relatively narrow fringe of area within a different socio-economic development situation. Large properties on a plain where intensive agriculture was developed since the eighties decade cut down most of the natural vegetation to the ground, sometimes with no sign of it within the sight range. Even eucalyptus and pine forestry had been practiced in the region.

In 1997, an aerial photo coverage was contracted by the Environmental State Agency (AAEG - Agencia Ambiental do Estado de Goiás), to support the studies for the park settlement and management. Surprisingly, the management plan on development nowadays for the park is using Landsat images instead of the photo coverage.

4. Methodology

The work started using the black and white panchromatic aerial photo coverage of the PETR performed in 1997 by the Environmental State Agency (AAEG). The original product was on printed copies at roughly 1:10,000 map scale. These photos were scanned at 600 dpi resolution to allow good quality zooming and make easy the process of land cover on screen classification.

The digital versions of the photos were imported into ERDAS Imagine ver. 8.5 (ERDAS 1999) for geometric correction prior to the interpretation process. Reference points were collected on a 1999 Landsat ETM image. A first order polynomial model with nearest neighbor re-sampling method was used for the geometric correction using at least 9 reference points. The points were located the most scattered as possible on the photo and root mean squared error (RMSE) of the model was never larger then 80 pixels. This was considered very good taking that the photo resolution was around 42 cm the RMS error was never larger then 1 pixel from the reference source.

After the geo-correction two mosaics were made with the photos inside the São Vicente watershed to facilitate the on screen interpretation. The watershed boundaries were determined with a stereoscopy work on the printed version of the photos. Actually, stereoscopy was also used to help solving doubts during the interpretation process.

The mosaics were opened on ArcView ver. 3.2 (ESRI, Inc. 1999) where the on-screen interpretation was performed. An average 1:5,000 map scale was kept while digitizing the patches of land cover types. Zooming tools and stereoscopy on the printed copies were used to aid the careful and meticulous work of interpretation.

Table 1 – Land cover classes.

	Original	**Simplified**
Natural	*Gallery forest*	*Gallery forest*
	Dry forest	*Dry Forest*
	Cerradão	
	Palmeiral	*Palmeiral / Vereda*
	Vereda	
	Cerrado denso	*Cerrado denso*
	Cerrado típico	*Cerrado típico*
	Cerrado ralo	*Cerrado ralo*
	Cerrado rupestre	*Cerrado rupestre*
	Campo rupestre	
	Campo sujo seco	*Campo sujo*
	Campo sujo húmido	
	Campo de murunduns	
	Campo limpo seco	*Campo limpo*
	Campo limpo húmido	
	Rivers	-
Human made	*Housing*	*Anthropogenic*
	Roads	
	Ranching management	
	Water dams	
	Other human made	
	Intensive pasture	*Pastures*
	Extensive pasture	
	Agriculture	*Agriculture*
	Small crops	
	Resting ag. areas	*Resting ag. areas*

The original 26 land cover classes used for the interpretation were divided into 10 human made classes and 16 "natural" classes (**Table 1**). What I mean for "natural" classes, are actually classes with non-detectable human interferences. Non-detectable human interferences are those that the extent of the human interference cannot be easily determined only by looking to the vegetation physiognomy as explained in the study area section above.

Before the accuracy assessment work the land cover map created with the aerial photo interpretation was simplified to better achieve our main objective, which is related to the land cover types of importance for landscape pattern analysis of the cerrado. Therefore, from the natural classes, "rivers" were masked out for not being a vegetation type of class. During the interpretation process I realized that there was too much uncertainty on some classes that were right away merged and considered not clearly identifiable with this kind of remote sensing data source. This action brought down the number of natural classes to 9. The original

10 human related classes were also simplified to 4 classes of really different types of land cover with different kind of impact to the cerrado ecosystem (**Table 1**).

The resolution of the 1:10,000 original source can be considered 0.5 mm of map measurement (Goodchild 1993). That would give us a 5 m aerial photo resolution. However, with the digitizing process used, where the photos were scanned at a 600 dpi resolution, and zooming in to 1:2,500 scale during the interpretation, we could say that our resolution was increased to 1.25 m. Lunetta (1998) considers that the spatial resolution of the remote sensed data needs to be at least one order of magnitude greater than the minimum mapping unit. Considering that a 1.6 m^2 does not define a land cover type I assumed the Goodchild's 5 m resolution was more than one order of magnitude greater than any land cover patch mapped. Hence, I took the interpretation process actual resolution as a factor of accuracy improvement.

The random points for verification were generated with ERDAS Imagine ver. 8.5. A 5 m pixel resolution image was generated from the land cover map created from the aerial photo interpretation. A multinomial distribution method as described by Congalton and Green (1999, pg.18) was used to calculate the recommended number of points for each class, resulting on 2591 points.

The points were scanned thru, overlaid on the photos, for assessment of those with clear classification to reduce the number of points that needs to be checked in the field. Three fieldwork trips were programmed also to optimize the use of the network of roads and tracks and reach as good as possible, one point on each polygon of the land cover map and check at least 75 points from each class. Finally 1288 points were checked for accuracy in 21 days of fieldwork. The points were found using a global positioning system (GPS Garmin® eTrex Vista).

5. Results

The first result of the accuracy assessment was an overall accuracy of 77% with Khat = 0.73, Z = 52.5. That is way bellow the "85% cutoff level between acceptable and unacceptable result" (Congalton and Green 1999). Users' accuracies were good (above 85%) only for dry forest, palmeiral, campo limpo, agriculture and resting agriculture classes. That was not what was expected after the above described thorough work of interpretation.

Analyzing the data and the processes some experiments were performed to bring this result to a wider acceptable level of accuracy. I used a simple hierarchical scaling strategy (Wu 1999) to identify patch types to be combined and simplify my map and evaluate the accuracy improvement. The steps for simplification were chosen based on the classes' similarities taken from their participation on each other in the error matrix and the qualitative justification for their misclassification as explained bellow.

Taking into consideration that pastures in this region are kept in a physiognomic state very similar to "cerrado ralo" and that it presented the lowest users accuracy (36%) in the error matrix, all pastures polygons were reclassified into "cerrado ralo". This procedure improved the overall accuracy to 79%, with Khat = 0.76, Z = 58.6.

"Cerrado típico" and "cerrado denso" were the next class chosen for combination, first because they had both users and producers low accuracies. Also, because those low accuracies can be easily explained. I used Ribeiro and Walter's (1998) definitions for those classes and the boundaries between the two are really confusing. They define "cerrado denso" as a cerrado physiognomy with 50% to 70% canopy cover of arboreal species from 5 to 8 meters tall and "cerrado típico", 20% to 50% canopy cover of arboreal species 3 to 6 meters tall. Not even high-resolution aerial photos are suitable to tell apart the height differences used as boundaries for those definitions. Moreover, Nature do not play the game by human defined rules and insist in mixing very well trees from many statures together in the cerrado, even more when canopy cover is not very dense. The overall accuracy improved to 83%, with khat = 0.79, Z = 63.1, after the combination of those two classes.

That could be considered an almost reasonable overall accuracy. I was able to reach 85% (khat = 0.81, Z= 65.8) accuracy, combining classes like resting agriculture with "cerrado ralo" and "campo sujo" with "campo limpo" (**Figure 1**). However, the question at that point is more related to the objective of the land cover mapping than to improve the accuracy in two percentage points.

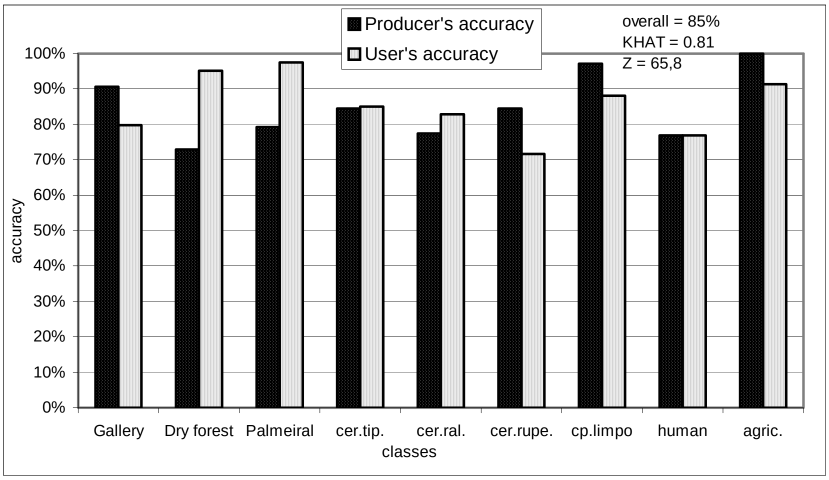

Figure 1 – Best overall accuracy after class combinations.

6. Discussion

The use of remote sensing sources for land cover mapping is becoming a very common use. However, not much attention is given to the accuracy of those maps. As mentioned before, the AAEG technicians, who had the photo coverage preferred to use a land cover map created from a Landsat image for the management plan for the park.

After the high impact human activities like the agro-forestry developed in the region, I was expecting to be able to have at least pastures accurately mapped and account for this minimum of human interference in the landscape pattern of this region. Surprisingly, it was the first class in need to be merged for a considerable improvement in the accuracy of the interpretation. I believe that this is not necessarily the case for other cerrado regions where more advanced technology ranching practices are dominant, where exotic more productive grass species are planted, soil acidity corrected and recurrent cerrado plants are effectively controlled. As mentioned before, in our area of study the ranching economic activity is still very primitive. It is practiced with native grass species maintenance and other cerrado plant control done predominantly with the use of regular fire setting.

I believe also that this low technology ranching economy plays an important role in making it more difficult to separate "cerrado típico" and "cerrado denso" in this region. Some conservation experiments in preserved areas of cerrado have shown that when there are soil conditions for dense vegetation cerrado development, if the area is kept protected from fire recurrence, the more open types of cerrado vegetation give place to denser higher canopy cover. These observations confirm published theories about fire importance in the cerrado physiognomies distribution (Ribeiro and Walter 1998, Coutinho 1992, Eiten 1990).

In the introductory section of this paper some numbers were provided from previous studies looking for the evaluation of the Cerrado conservation stage. Those evaluations were made for the entire biome and most certainly did not use aerial photo interpretation to reach those numbers. We have just seen with the present experiment that not even the most common kind of pastures can be separated from natural open cerrado physiognomies with reasonable accuracy. Therefore, how could we have an actual view of the impact that modern human activities are making to the cerrado environment.

My interest in landscape ecology of the cerrado is to understand the relevance of its complex landscape pattern on its most important processes. However, exactly because of this ecosystem landscape complexity, with its wide range of physiognomic forms, which masks out much of the exogenous interferences, we have to be more careful with conclusions taken from studies of the cerrado landscape pattern.

A land cover map with 26 classes could be produced with this methodology where the best of advanced technology and careful meticulous one-man interpretation was used with a lot of work hours. That could convince most of the managers and decision makers. Our analysis showed however that a map with reasonable overall accuracy produced from this aerial photo interpretation could only have about 10 classes. The best accuracy result was on agriculture related areas (**Figure 1**). Without this class, not even the 85% threshold would have been reached as overall accuracy. Hence, is this data worth work with for landscape pattern analysis of the cerrado physiognomies? That is our next step. I will evaluate the effect of the hierarchical level change of this data onto the landscape indices of the study area.

7. Acknowledgements

This project was accomplished with supports from WWF - Brasil and Ford Foundation through the program "Natureza e Sociedade" (BRZ NT 432/2001) and from Universidade Católica de Goiás (res.# 645). I also thanks the Agencia Ambiental do Estado de Goiás for providing the aerial photos and access to the Terra Ronca State Park infrastructure.

8. References

Castro, E. C. Human influence detection at different landcover hierarchical levels on the Cerrado Biome. In: 17th Annual Symposium of the International Association of Landscape Ecology – IALE United States Regional Chapter. Landscapes in Transition: cultural drivers and natural constraints, 2002, Lincoln, NE. **Proceedings**… Lincoln: US-IALE, 2002, p. 81-82. URL: < http://www.calmit.unl.edu/usiale2002/>

Congalton, R. G. and Green, K. **Assessing the Accuracy of Remotely Sensed Data: principles and practices**. Boca Raton, FL.: Lewis Publishers. 1999. 137 p.

Coutinho, L. M. O cerrado e a ecologia do fogo. **Ciência Hoje**, p. 130-138, maio, 1992. Volume especial Eco-Brasil.

Dias, B. F. S. Conservação da Natureza no Cerrado Brasileiro. In: Pinto, M. N. (org.). **Cerrado: caracterização, ocupação e perspectivas**. p. 583-640, Brasília: Editora Universidade de Brasília, 1990.

Eiten, G. Vegetação do Cerrado. In: Pinto, M. N. (org.). **Cerrado: caracterização, ocupação e perspectivas**. p. 583-640, Brasília: Editora Universidade de Brasília, 1990.

ERDAS, Inc. **ERDAS Imagine v.8.4**, Atlanta, GE: ERDAS, Inc., 1999.

ESRI, Inc. **ArcView 8.3.**, Redlands, CA: Environmental Systems Research Institute, Inc., 1999.

Goodchild, M. F. Data models and data quality: problems and prospects. In: Goodchild, M. F., Parks, B. O. and Steyaert, L. T. (editors). **Environmental modeling with GIS**, New York, NY: Oxford University Press, 1993.

Lunetta, R. S. Applications, project formulations, and analytical approach. In: Lunetta, R. S. and Elvidge, C. D. (ed.). **Remote Sensing Change Detection: environmental monitoring methods and applications**, Chelsea, MI: Ann Arbor Press, 1998.

Lowel, K. and Jaton, A. **Spatial accuracy assessment: land information uncertainty in natural resources**. Chelsea, MI: Sleeping Bear Press, Inc, 1999.

Mittermeier, R. A.; Myers, N.; Gil, P. R. and Mittermeier, C. **Hotspots: earth's biologically richest and most endangered terrestrial ecoregions**. Mexico City: Conservation International, 1999.

Quattrochi, D. A. and Goodchild, M. F. **Scale in remote sensing and GIS**, Boca Raton, FL: CRC Press, 1997.

Ribeiro, J. F. e Walter, B. M. T. Fitofisionomias do Bioma Cerrado. In: Sano, S. M. e Almeida, S. P. (ed.). **Cerrado: ambiente e flora**, Planaltina, DF: EMBRAPA-CPAC, 1998.

Haines-Young, R.; Green, D. R. and Cousins, S. H. **Landscape Ecology and GIS**, London: Taylor & Francis, 1993.

Withers, M. A. and Meentemeyer, V. Concepts of Scale in Landscape Ecology. In: Klopatek, J. M. and Gardner, R. H. (eds.) **Landscape Ecological Analysis: Issues and Applications**, New York, NY: Springer-Verlag Inc., 1999. (205--252)

Wu, J. Hierarchy and scaling: extrapolating information along a scaling ladder. **Canadian Journal of Remote Sensing,** v. 25, p. 367-380, 1999.

Wu, J. and Qi, Y. Dealing with scale in landscape analysis: An overview. **Geographic Information Sciences,** v. 6, p. 1-5, 2000.