

FUSÃO DE IMAGENS LANDSAT E MINERAÇÃO DE DADOS GEOGRÁFICOS PARA MAPEAR O CERRADO DOS MUNICÍPIOS DE PARAOPEBA E CAETANÓPOLIS – MG.

Luciano Teixeira de Oliveira ¹
Luis Marcelo Tavares de Carvalho ¹
Fausto Weimar Acerbi Júnior ^{1,2}

¹ Universidade Federal de Lavras – UFLA/DCF/LEMAF
Caixa Postal 3037 – 37200-000 – Lavras-MG, Brasil
oliveiralt@yahoo.com.br
passarinho@ufla.br
fausto@ufla.br

² Wageningen University and Research Centre – WUR/GRS
P.O. Box 47 – 4700 AA – Wageningen, The Netherlands
fausto.acerbi@wur.nl

Abstract. The main objective of this paper was to map the fragments of the cerrado biome in the municipalities of Paraopeba and Caetanópolis. An ETM+ image was enhanced to a spatial resolution of 15 meters, using the non-decimated biorthogonal wavelet transform with Antonini 7/9 filter. The input datasets for the classifiers were composed by the ETM+ spectral bands, NDVI, Tasseled Cap, spectral unmixing fraction images, Isodata spectral clustering, distance from rivers and an elevation model. Decision trees using the GINI algorithm were used constructed to classify this data set. The classification accuracy was 95,46% and the Kappa coefficient was 0,9481, generating a map of the cerrado physiognomies with good reliability. The rules used in the multivariate decision trees separated well the characteristics of each studied class. According to our results the two municipalities still have 6,637.00 hectares of forest, 14,950.00 hectares of savanna and 9,820.00 hectares of native fields.

Palavras-chave: cerrado, image classification, image fusion, decision trees, wavelets, cerrado, classificação de imagens, árvores de decisão, ondaletas.

1. Introdução

O município de Paraopeba e Caetanópolis, por apresentarem no centro-oeste de Minas Gerais e, por possuírem uma paisagem altamente fragmentada, tendo na matriz agrária vários plantios de eucalipto provenientes de diversos tratamentos silviculturais e estágios de crescimento diferentes, como também, o predomínio de pastagens, foram selecionados para aplicar a metodologia proposta por Oliveira (2004).

Esta situação observada nestes municípios é bem representativa de regiões que sofreram uma alta intervenção agropecuária, o que dificulta a separação das fitofisionomias naturais dos demais elementos da paisagem, devido a alta complexidade da nova paisagem.

Neste sentido, objetivo principal deste trabalho é de classificar e quantificar a vegetação da Região de Paraopeba e Caetanópolis, para avaliar a eficiência da metodologia proposta por Oliveira (2004), utilizando o nível um de classificação.

2. Material e métodos

2.1. Processamento das imagens

Utilizou-se neste trabalho, uma imagem do período seco, devido a presença de nuvens na época úmida, esta escolha foi feita devido a não se ter observado diferença significativa entre

os conjuntos de atributos da imagem de uma época para a temporal, no nível – 1 de detalhamento nos resultados de Oliveira (2004).

A imagem gerada pelo sensor ETM+ (época seca) foi corrigida, usando a transformação polinomial de 1o grau e a reamostragem por vizinho mais próximo (Lillesand & Kiefer, 1999). Esta foi adquirida com nível máximo de correção (1G), ou seja, com correção para a elevação do terreno. A correção valeu-se de parâmetros como efemérides, altitude do satélite, sistema de imageamento do sensor e modelo de referência da Terra para o georeferenciamento usando apenas um ponto de controle (INPE, 2003). Este ponto localiza-se nas coordenadas S 19° 22' 59" de latitude, W 44° 29' 46" de longitude referente a um cruzamento entre talhões de Eucalyptus sp da Empresa Vallourec & Mannesmann Florestal, situada no Município de Paraopeba. O sistema de projeção e Datum utilizados foram, respectivamente, UTM e WGS-84.

Na imagem aplicou-se a transformação Ondaleta bi-ortogonal não decimada com filtro de Antonini 7/9, o procedimento de estruturação do conjunto de atributos foi baseado nos resultados de Oliveira (2004), para uma classificação não detalhada, nível – 1, para o mapeamento do Bioma cerrado.

Gerou-se para esta região um conjunto de imagens derivadas do sensor ETM+ (época seca) por meio de NDVI, MLME, Tasseled Cap (Crist e Cicone, 1984), classificação Isodata com 25 classes, “buffer” dos rios e modelo de elevação digital, seguindo os mesmos critérios para cada processamento conforme Oliveira (2004).

As frações de mistura foram obtidas levando-se em consideração a teoria do simplex (Correia, 1982, Aguiar, 1991; Schowengerdt, 1997; Mather, 1999; Tso & Mather 2001), obtendo-se os pixels puros dos extremos da distribuição do espaço amostral, Vermelho x infravermelho **Figura 1**.

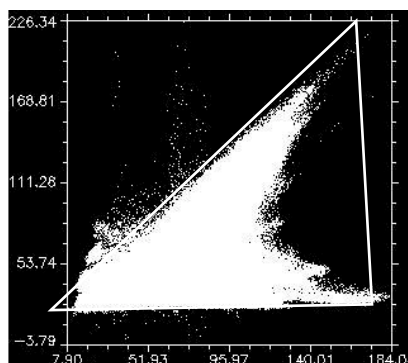


Figura 1. Localização dos extremos do triângulo no espaço amostral vermelho x infravermelho, onde se extraiu os Pixels puros de solo, sombra e vegetação, nas imagens Landsat época seca.

A hidrografia foi utilizada também para fazer a correção dos limites municipais obtidos do IBGE, os quais detectou-se imperfeições no confrontamento das divisas dos rios com o real traçado dos mesmos. Gerando-se a Delimitação da área de estudo.

2.2. Amostragem da verdade terrestre no campo

Para esta região seguiu o critério estabelecido por Ribeiro & Walter (1998), selecionado o nível 1 de classificação, a qual divide a o Bioma do cerrado como: Florestal, Savânica e Campestre.

Nesta fez-se um sobrevôo na região com caminhamentos terrestres, em locais onde se observou na imagem manchas de vegetação (**Figura 2**), onde possibilitou a coleta das amostras de campo, utilizando a referência de um GPS Garmin III+ sem correção diferencial.

Com a sistematização do conjunto de amostras da vegetação e do uso do solo, procedeu-se a aleatorização das amostras de treinamento, aferição e acuracidade, estabelecendo 20% para cada agrupamento de amostras conforme **Tabela 1**.

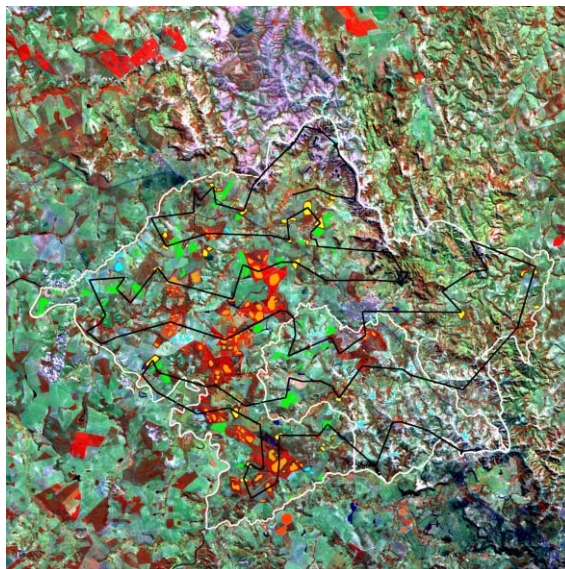


Figura 2. Composição colorida, RGB - 453 (Landsat) época seca, com caminhamento aéreo em preto e limites municipais em branco, onde o maior município pertencente a Paraopeba e o menor município pertencente a Caetanópolis.

Tabela 1. Amostragem Nível – 1.

Amostras	Florestal	Savânica	Campestre	Eucaliptal	Pastagem	Cultura	Solo	Água	Sombra
Amostragem	11318	23575	12064	46230	55290	5542	22515	5842	1706
Treinamento	2000	2000	2000	2000	2000	1000	2000	1000	400
Aferição	2000	2000	2000	2000	2000	1000	2000	1000	400
Acuracidade	2000	2000	2000	2000	2000	1000	2000	1000	400

2.3. Classificação da região

Para a classificação utilizou-se um algoritmo Gini do programa Cart 5.0, plicando-se, uma pré-análise entre as árvores uni e multivariada, onde selecionou-se a árvore multivariada devido a esta ter tido um menor número de regras e um menor custo para a classificação, e um maior Kappa, observado na **Tabela 2** (Breiman et al., 1984).

Tabela 2. Amostragem Nível – 1.

Nível 1	Custo Relativo	Número de Folhas	Número de Árvores	Coeficiente Kappa	Variância
Árvore univariada	0,042	185	54	0,9519	-0,00056
Árvore multivariada	0,036	117	49	0,9578	-0,00064

Devido ao fato do algoritmo de árvore de decisão ser não paramétrico, trabalhando com regras para separar as classes, fez-se a comparação das resoluções selecionando a resolução que gerou o menor índice de árvore e o menor número de regras, para se proceder a escolha das amostras para a classificação, sendo utilizado os valores fixados na **Tabela 1**.

2.4. Análise da acuracidade

Para verificar a acuracidade dos classificadores, utilizou-se as amostras de acuracidade da Tabela 1, com as quais se gerou a matriz de confusão e o coeficiente Kappa (Colganton e Green, 1999; Tso e Mather 2001; Gaboardi, 2002).

Também se verificou a acuracidade de cada classe em separado pelo método proposto por Kalensky and Scherk (1975) citados por Carvalho (2001), o qual leva em consideração os erros de comissão e omissão de cada classe.

$$A_i = \frac{N_i}{(N_i + E_i)}$$

Onde:

A_i – É a acuracidade no mapeamento da classe i ;

N_i – É o número de pixels classificados corretos dentro de cada classe i ;

E_i – É o número de comissões e omissões da classe i .

3. Resultados e discussão

Com a utilização dos procedimentos propostos por Oliveira (2004), foi possível mapear as fitosionomias do Bioma Cerrado no municípios da região de Paraopeba e Caetanópolis (**Figura 2**), com uma precisão geral de 95,46% e um coeficiente Kappa de 0,9481, o qual segundo a classificação de Landis e Koch (1977), é considerado como excelente.

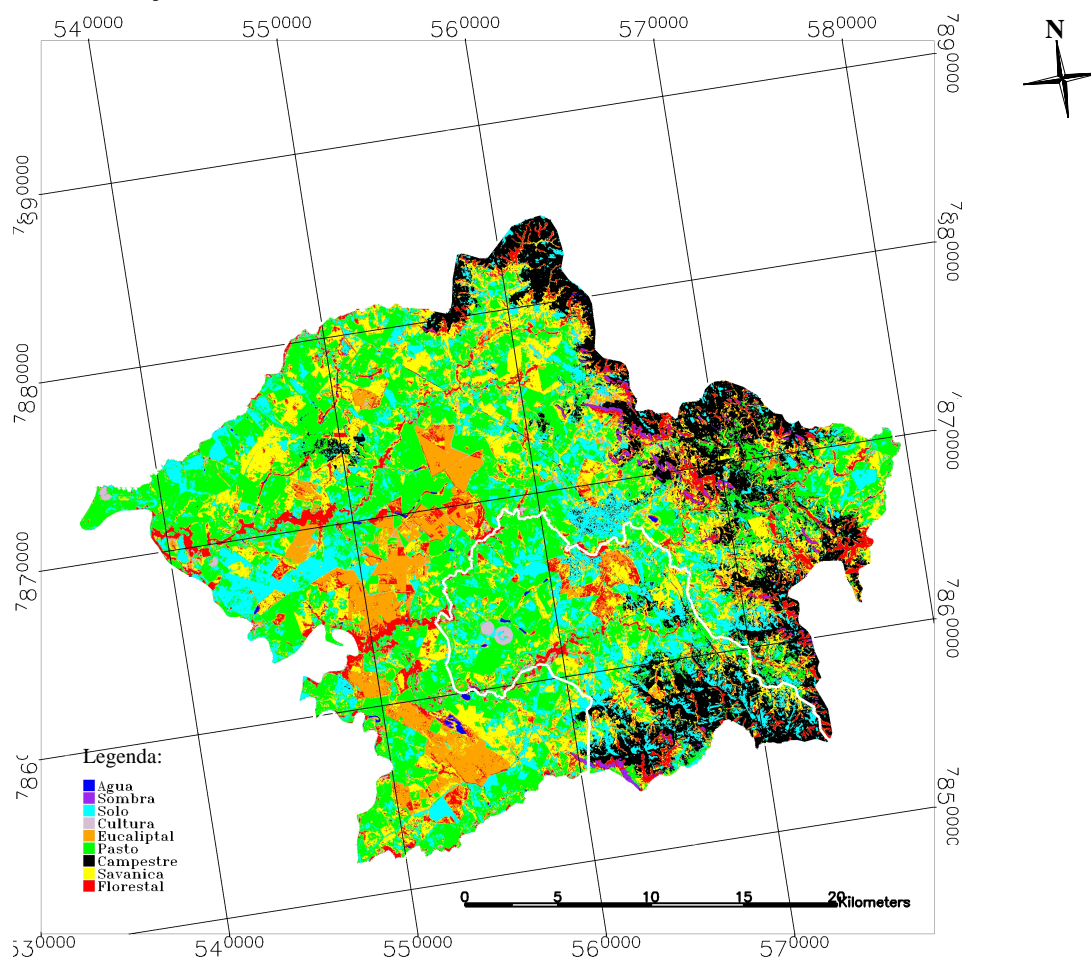


Figura 2. Resultado da Classificação do uso da terra nos Municípios de Paraopeba e Caetanópolis.

O resultado da precisão individual de cada classe (**Tabela 3**), dá uma idéia do erro que está ocorrendo em cada classe. Este resultado observado pela análise qualitativa e quantitativa da imagem gerada corrobora com a afirmação observada por Oliveira (2004), da viabilidade da aplicação do procedimento para a classificação do bioma cerrado no Centro Oeste mineiro.

Tabela 3. Matriz de confusão para Imagem Landsat Época Seca com 95,46% de acurcacidade e 0,9481 de Coeficiente Kappa.

Classes	Sombra	Água	Cultura	Pasto	Solo	Campestre	Savanica	Florestal	Eucaliptal	Total
Sombra	392	9	0	0	0	0	0	12	0	413
Água	5	990	0	0	0	0	0	0	0	995
Cultura	0	0	989	3	0	0	0	3	4	999
Pasto	0	0	0	1977	31	2	4	0	6	2020
Solo	0	0	0	8	1953	13	0	0	0	1974
Campestre	0	0	0	3	13	1978	0	0	0	1994
Savanica	0	0	0	2	3	6	1875	48	28	1962
Florestal	3	1	5	2	0	1	78	1872	241	2203
Eucaliptal	0	0	6	5	0	0	43	65	1721	1840
Total	400	1000	1000	2000	2000	2000	2000	2000	2000	14400
Precisão	93,1	98,5	97,9	96,8	96,5	98,1	89,8	80,3	81,2	95,5

O melhor desempenho da árvore multivariada sobre a univariada, pode ter sido devido à complexidade do sistema ao qual as duas árvores foram submetidas, evidenciando a necessidade de se fazer uma análise prévia da classificação.

Observou-se também uma maior confusão na **Tabela 3**, dos plantios de Eucalipto com a Vegetação Florestal, que pode ter ocorrido devido a influência da alta variabilidade dos plantios observado na **Figuras 2**.

Os resultados do gráfico da **Figura 3** permitiram observar a fragmentação existente nestes dois municípios, os quais possuem apenas 30% de vegetação nativa de um total de 78.247 ha, mostrando a influência marcante da agropecuária no processo de substituição do uso solo.

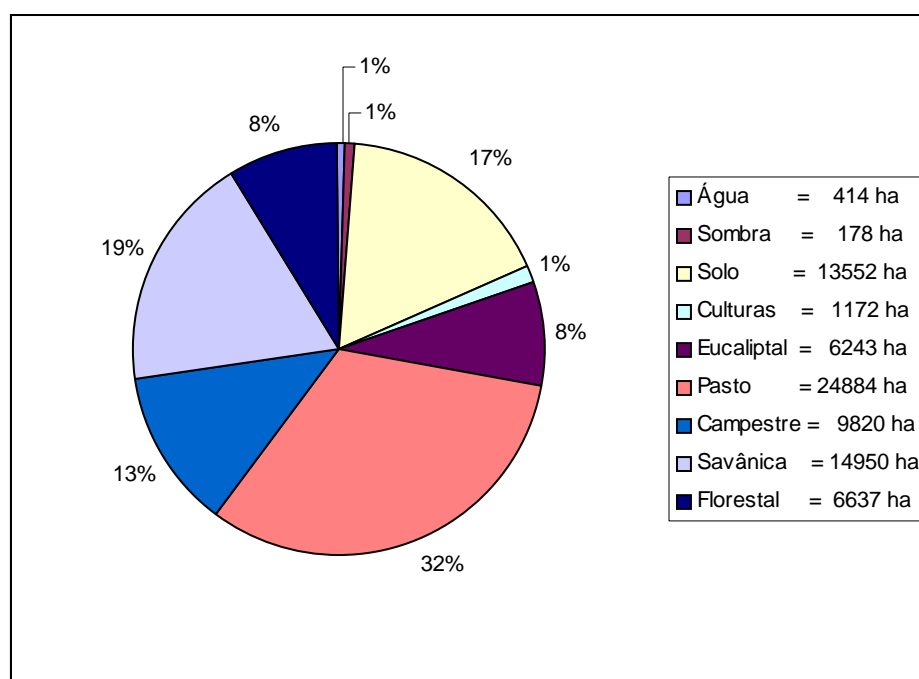


Figura 3. Área e percentual de cada classe nos municípios de Paraopeba e Caetanópolis.

4. Conclusão

A utilização do conjunto de atributos usados neste trabalho permitiu o mapeamento das fitofisionomias do bioma cerrado com uma boa confiabilidade;

As regras utilizadas na árvore multivariada com o algoritmo Gini conseguiram separar bem as características de cada classe estudada.

O nível 1 da classificação usado, conseguiu atender a necessidade do mapeamento da região.

Os municípios possuem 6.637 ha de formação florestal, 14.950 ha de formação savânica e 9.820 ha de formação campestre.

5. Referências Bibliográficas

Aguiar, A.P.D. **Utilização de atributos derivados de proporções de classes dentro de um elemento de resolução de imagem (“pixel”) na classificação multispectral de imagens de sensoriamento remoto.** 1991. 227 p. (INPE-5306-TDI/496). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1991.

Breiman, L.; Friedman, J.H.; Olshen, R.A. **Classification and regression trees.** Belmont: Chapman & Hall, 1984. 358 p.

Carvalho, L.M.T. **Mapping and monitoring forest remnants: a multi-scale analysis of spatio-temporal data.** 2001. 140 p. (ISBN 90 – 5808 – 545 – 7). Thesis (Doctor in Remote Sensing) - Wageningen University, Wageningen.

Correia, V.R.M. **Estudo das medidas de qualidade para estimação de proporções de classes em elementos de resolução de imagens.** 1982. 118 p. (INPE-2712-TDL). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1982.

Colgaton, R.G.; Green, K. **Assessing the accuracy of remotely sensed data: principles and practices.** New York: Lewis Publishers, 1999. 137 p.

Crist, E.P.; Cicone, R.C. A physically – based transformation of thematic mapper data – the TM Tasseled Cap. **IEEE Transactions on Geoscience and Remote Sensing**, Los Alamitos, v. 22, n. 3, p. 256-262, 1984.

Gaboardi, C. **Utilização de imagem coerência SAR para classificação do uso da terra: Floresta Nacional do Tapajós.** 2002. 137 p. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2002.

Instituto Nacional de Pesquisas Espaciais. Divisão de Processamento de Imagens (INPE-DPI). **SPRING, Manual do usuário [on line].** São José dos Campos, 2003.

Koukoulas, S.; Blackburn, G. A. Introducing new indices for accuracy evaluation of classified images representing semi-natural woodland environments. **Photogrammetric Engineering & Remote Sensing**, Los Alamitos, v. 67, n. 4, p. 499 -510, 2001.

Landis, J. R.; Koch, C. H. **The measurement of observer agreement for categorical data.** **Biometrics**, Washington, v. 33, n. 3, p. 159-174, 1997.

Lillesand, T.M.; Kiefer, R.W. **Remote sensing and image interpretation.** 4. ed. USA: John Wiley, 1999. 724 p.

Mather, P.M. **Computer processing of remotely-sensed images: an introduction.** 2. ed. Nottingham, UK: John Wiley, 1999. 292 p.

Oliveira, L.T.; **Fusão de imagens de sensoriamento remoto e mineração de dados geográficos para mapear as fitofisionomias do Bioma Cerrado.** 2004. 131p. (CDD – 621.3678 – 526.982) Dissertação (Mestrado em Manejo Ambiental) – UFLA. Lavras. 2004.

Ribeiro, J.F.; Walter, B.M.T. Fitofisionomias do bioma Cerrado. In: Sano, S.; Almeida, S. P. (Ed.) **Cerrado: ambiente e flora.** Planaltina: EMBRAPA-CPAC, 1998. p. 89-169.

Schowengerdt, R.A. **Models and methods for image processing.** 2. ed. New York: Academy Press, 1997. 522 p.

Tso, B.; Mather, P.M. **Classification Methods for remotely sensed data.** New York: Taylor & Francis, 2001. 332 p.