

Pesquisas em Geociências

<http://seer.ufrgs.br/PesquisasemGeociencias>

**Reconhecimento e classificação da cobertura vegetal a partir de informações
provenientes do laser scanning empregando a função
discriminante linear de Fischer**

Marcos Benedito Schimalski, Edson Aparecido Mitishita, Anselmo Chaves Neto
Pesquisas em Geociências, 36 (2): 141-148, set./dez., 2009.

Versão online disponível em:

<http://seer.ufrgs.br/PesquisasemGeociencias/article/view/17858>

Publicado por

Instituto de Geociências



Portal de Periódicos
UFRGS

UNIVERSIDADE FEDERAL
DO RIO GRANDE DO SUL

Informações Adicionais

Email: pesquisas@ufrgs.br

Políticas: <http://seer.ufrgs.br/PesquisasemGeociencias/about/editorialPolicies#openAccessPolicy>

Submissão: <http://seer.ufrgs.br/PesquisasemGeociencias/about/submissions#onlineSubmissions>

Diretrizes: <http://seer.ufrgs.br/PesquisasemGeociencias/about/submissions#authorGuidelines>

Data de publicação - set./dez., 2009.

Instituto de Geociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

Reconhecimento e classificação da cobertura vegetal a partir de informações provenientes do laser *scanning* empregando a função discriminante linear de Fischer

Marcos Benedito SCHIMALSKI¹, Edson Aparecido MITISHITA² & Anselmo Chaves NETO³

1. Universidade do Contestado, Curso de Engenharia Florestal, Caixa Postal 01, CEP 89.460-000, Canoinhas, SC, Brasil. E-mail: mbs@cni.unc.br

2. Universidade Federal do Paraná, Setor de Ciências da Terra, Departamento de Geomática, Caixa Postal 19.001, CEP 81.531-990, Curitiba, PR, Brasil. E-mail: mitishita@ufpr.br

3. Universidade Federal do Paraná, Setor de Ciências Exatas, Departamento de Estatística, Caixa Postal 19.081, CEP: 81.531-990, Curitiba, PR, Brasil. E-mail: anselmo@ufpr.br

Recebido em 10/2008. Aceito para publicação em 11/2009.

Versão online publicada em 14/06/2010 (www.pesquisasemgeociencias.ufrgs.br)

Resumo - A evolução do espaço rural vinculado a ampliação da fronteira agrícola, necessária as atividades do agro- negócio e da silvicultura, exercem forte pressão sobre os recursos naturais impactando, muitas vezes, de forma negativa as intervenções exercidas por tais atividades. Neste contexto, o conhecimento da realidade física necessária ao planejamento das intervenções, pode ser obtido através do emprego do Sensoriamento Remoto, cujos produtos podem auxiliar os gestores na orientação das atividades expansionistas e conservacionistas. Contudo, para a verificação da legislação ambiental brasileira, além de informações planimétricas, necessita-se de informações referentes à Geomorfologia. O sistema *laser scanning* apresenta vantagens no fornecimento destas informações além de possibilitar a separação da cobertura vegetal através de produtos derivados de suas observações, entre eles a imagem de intensidade. O objetivo geral deste trabalho foi verificar a classificação da cobertura vegetal a partir das observações brutas obtidas pelo laser *scanning*, empregando a Função Discriminante Linear de Fischer. Para possibilitar a classificação foram gerados parâmetros adicionais oriundos das observações laser: densidade de pontos, variância altimétrica e a altura da cobertura vegetal, além da informação referente à intensidade. São apresentadas as etapas de trabalho e os resultados obtidos no processo para a área de estudo localizada na represa do Capivari em Curitiba no Estado do Paraná.

Palavras-chave: sensoriamento remoto, tipologias vegetais, LIDAR, análise multivariada.

Abstract - BEACH CLASSIFICATION OF THE VEGETAL COVERING FROM RAW DATA PROCEEDING FROM LASER SCANNING USING FISCHER'S LINEAR DISCRIMINATE FUNCTION. The evolution of rural space and the agriculture activity enlargement required for agribusiness and silviculture has put strong pressure on natural resources causing, in most of cases, negative effects on interventions done by these activities. In this context, the physical reality knowledge and the intervention planning can be obtained from the Remote Sensing which products can aid authorities to guide the expansionist and conservationist activities. However, in order to check the Brazilian environmental legislation and plan metric data, information related to Geomorphology is necessary too. The laser scanning system presents some advantages to offer this information as well as to separate the vegetal covering through derivate products from its observations such as image of intensity. The general goal of this study was testing the classification of vegetal covering based on rough observations obtained by laser scanning, using Fischer's Linear Discriminate Function. In order to classify it, additional parameters from laser observations such as density of points, altimetry variance, vegetal covering height and information related to intensity were added to. The steps of the work are presented as well as the results of the research carried out at Capivari dam located in Curitiba, State of Paraná.

Keywords: remote sensing, vegetal recovering, LIDAR, multivariate analysis.

1. Introdução

A legislação ambiental brasileira apresenta parâmetros norteadores à ocupação do espaço físico que impõem restrições a alterações nas formações naturais remanescentes, tratando da utilização racional dos recursos naturais mediante a definição do uso do solo. Estes parâmetros delimitam áreas de preservação permanente (APP) através da definição de feições naturais específicas. Citam-se como exemplos a delimitação de áreas de encosta, mata de galeria, topo de morro e declividades pré-definidas (BRASIL, 1965).

Além de informações referentes ao uso do solo faz-se essencial à caracterização da geomorfologia e mais especificamente a representação do relevo para verificação da legislação. Para tanto, técnicas de mensuração direta e indireta são empregadas para a obtenção de informações de relevo. Geralmente, as técnicas de mensuração indireta quando comparadas com as técnicas diretas em campo, apresentam vantagens relacionadas com o menor custo e menor tempo de execução das principais etapas de trabalho envolvidas no processo. Dentre as formas indiretas, a utilização de técnicas automáticas de geração de Modelo Digital do Terreno (MDT) com base em pares estereoscópicos, provenientes de coberturas aéreas, são as mais comumente empregadas para a extração da altimetria (Baltsavias, 1999). Alternativamente, o ALS (*Laser Scanning* Aéreo) é uma inovação tecnológica bastante importante para o Sensoriamento Remoto. Trata-se de uma tecnologia relativamente nova que possibilita a obtenção automática de coordenadas tridimensionais de pontos sobre a superfície da Terra em um curto período de tempo (Dalmolin & Santos, 2003). Esse sistema possui a vantagem de determinar uma grande quantidade de pontos posicionados tridimensionalmente na superfície terrestre, porém, apresenta a desvantagem de não direcionar o raio para o levantamento de somente entidades de interesse, acarretando o desenvolvimento de técnicas para a classificação de pontos, bastante específicas, como por exemplo, a separação de pontos nas elevações e no terreno.

A separação do Modelo Digital de Superfícies (MDS) e do MDT a partir dos dados obtidos pelo ALS é uma tarefa que exige a utilização de algoritmos que funcionam corretamente em determinadas situações. Os filtros para a separação de pontos *laser* podem ser agrupados

de acordo com suas características em: morfológicos, por triangulação, modelagem de superfície e segmentação (Kobler *et al.*, 2007). Geralmente, os algoritmos de filtragem apresentam bons resultados de classificação em áreas com relevo mais suave. Entretanto, em áreas inclinadas e cobertas por vegetação surgem problemas na classificação, tornando as áreas rurais e, especificamente, as florestais, como críticas na separação das feições da superfície e do terreno.

Para que as informações provenientes do *laser* auxiliem na verificação da legislação ambiental e possibilitem a extração de parâmetros satisfatórios para as aplicações na área florestal faz-se necessário obter o MDT e o MDS classificados de acordo com a cobertura vegetal (Holmgren, 2003).

Em função do exposto, o problema a ser abordado por este trabalho é a classificação do uso do solo empregando somente as observações brutas obtidas pelo ALS. São objetivos deste estudo (i) determinar as variáveis a serem empregadas na classificação estatística; (ii) obter amostras representativas das tipologias referentes ao uso do solo; (iii) classificar a cobertura vegetal empregando a função discriminante linear (FDL) de Fischer; (iv) comparar os resultados obtidos com a FDL e com a classificação baseada em fotointerpretação.

2. Materiais e Métodos

2.1 Localização da área

O local escolhido para a realização do trabalho encontra-se próximo à represa do rio Capivari, no município de Curitiba, com uma área de 9,32 Km². As coordenadas limitantes em metros (referidas ao Datum WGS84 e em projeção Universal Transversa de Mercator) para a área são mínimas: E - 713350 N - 7217500 e máximas: E - 717235 N - 7219900.

A região escolhida caracteriza-se em um ambiente rural, apresentando variações na cobertura vegetal e também pelo número reduzido de edificações.

2.2 Dados

O conjunto de dados empregado é procedente da varredura a *laser* que foi executada pelo LACTEC - Instituto de Tecnologia para o

Desenvolvimento. Além das observações brutas geradas pelo ALS foi cedida uma aerofoto em escala 1:20.000 cuja cobertura coincide com a varredura a *laser*. A figura 1 apresenta a aerofoto utilizada no estudo.



Figura 1. Aerofoto mostrando a área de estudo.

Para o processamento das observações brutas foram desenvolvidas rotinas computacionais no programa Matlab 2007. O arquivo com as observações brutas obtidas pelo ALS contém 9.745.737 pontos, sendo que as observações armazenadas consistem das coordenadas (E, N, H) no Sistema de Projeção UTM (Universal Transversa de Mercator) com adição da intensidade de resposta para cada ponto. A densidade média obtida foi de 1,05 ponto por metro quadrado.

2.3. Fotointerpretação

A primeira fase do trabalho consistiu na definição das classes temática que compõe a cobertura do solo a ser classificado. Como não foi possível a realização de inspeções com base em levantamentos a campo, procedeu-se a fotointerpretação da aerofoto sendo identificadas as seguintes classes temáticas: -mata nativa (Formação Natural Secundária em Estágio Avançado de Regeneração); -Capoeira (Formação Natural Secundária em Estágio Médio de Regeneração); -Capoeirinha (Formação Natural Secundária em Estágio Inicial de Regeneração); -Campo (Agropecuária); -Rodovia, e -Água.

Com auxílio da fotointerpretação foi então

possível selecionar 7 amostras georeferenciadas de cada classe na nuvem de pontos *laser*. Para facilitar o processamento computacional, a nuvem de pontos foi armazenada na forma de uma grade regular, com espaçamento horizontal de 10 m. Em cada célula da grade foram armazenadas as coordenadas UTM (E,N) para o centro da célula, a intensidade média, a variância altimétrica, a densidade de pontos e a altura da cobertura vegetal. A altura da cobertura vegetal foi obtida a partir da diferença entre o MDS e o MDT. A grade gerada apresentou 93.235 células contendo os quatro parâmetros mencionados em cada uma das células, ou seja, a intensidade média, variância altimétrica, densidade de pontos e altura da cobertura vegetal.

2.4. Processamento da classificação

Com a grade referente à varredura *laser* criada, procedeu-se a classificação empregando a estatística multivariada. Segundo Johnson (1998) os métodos multivariados são úteis para auxiliar na análise de grande quantidade de informações provenientes de um grupo de variáveis obtidas a partir de certo número de experimentos. A importância e o uso dos métodos multivariados aumentam com o número de variáveis a ser medidas, bem como, com o aumento do número de experimentos.

No escopo dos métodos multivariados existe uma parte de reconhecimento e classificação de padrões, que engloba a Função Discriminante Linear de Fischer, a Função Discriminante Quadrática e a Regressão Logística, entre outras técnicas de reconhecimento e classificação de padrões. O reconhecimento é feito ajustando-se, de forma adequada, funções aos vetores observados de dois ou mais grupos considerados distintos. A função abaixo representa uma função discriminante (Hair, 2005):

$$Z_{jk} = a + \sum_{i=1}^n W_i X_{ik}$$

Onde:

Z_{jk} - é o escore discriminante da função discriminante j para o objeto k ;

a - é o intercepto;

W_i - é o peso discriminante para a variável independente i ,

X_{ik} - variável independente i para o objeto k .

Ainda segundo Hair (2005) a análise discriminante tem por finalidade: determinar se existem diferenças estatisticamente significativas entre os perfis de escore médio em um conjunto de variáveis para dois (ou mais) grupos definidos preliminarmente; determinar quais das variáveis independentes explicam o máximo de diferenças nos perfis de escore médio dos dois ou mais grupos; estabelecer procedimentos para classificar objetos em grupos, com base em seus escores em um conjunto de variáveis independentes; e, estabelecer o número e a composição das dimensões de discriminação entre grupos formados a partir do conjunto de variáveis independentes.

A análise discriminante canônica ou FDL de Fischer cria uma função de classificação através de combinações lineares das variáveis originais. Essa função é criada de forma que contenha todas as informações a respeito do conjunto de variáveis originais (Johnson, 1998).

A FDL de Fischer não necessita da suposição de que as diversas populações sejam normais multivariadas, porém assume que as matrizes de covariâncias são iguais.

Sejam $g = 6$ classes de uso do solo e $p = 4$ variáveis explicativas (intensidade média, densidade de pontos, variância altimétrica e altura da cobertura vegetal).

Seja \bar{x}_i o vetor médio amostral de cada uma das cinco classes identificadas na foto-interpretção:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

e \bar{x} o vetor médio global para as cinco classes de uso do solo:

$$\bar{x} = \frac{\sum_{i=1}^g \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^g n_i}$$

A matriz Soma de Produtos Cruzados entre os grupos populacionais é:

$$\hat{B}_0 = \sum_{i=1}^g (\underline{x}_i - \bar{x})(\underline{x}_i - \bar{x})'$$

A matriz W é:

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$$

E, a matriz $W^{-1}\hat{B}_0$ pode ser decomposta, de acordo com o Teorema da Decomposição Espectral no produto das matrizes dos autovetores e autovalores, ou seja, $W^{-1}\hat{B}_0 = P\Lambda P'$.

Portanto, os seus p autovetores são: e_1, e_2, \dots, e_p . Assim, o $\min\{g - 1; p\}$ é 4 e tem-se 4 discriminantes para os g grupos, cujas expressões são:

$$\hat{Y}_i = \hat{e}_i \underline{x} \quad i = 1, 2, 3, 4.$$

Seja \underline{x} é composto pelas $p = 4$ variáveis.

Foram selecionadas sete amostras de cada classe, o que totalizou 42 amostras. Portanto, tem-se uma matriz de dados de ordem 42×4 e cuja distribuição é mostrada na figura 2.

A imagem mostrada na figura 2 é a imagem de intensidade gerada a partir das observações ALS. Deve ser destacado que não foram obtidas amostras na área de classificação.

Empregando a formulação demonstrada foram determinadas quatro funções discriminantes lineares e, usando-as, calculou-se os centróides (vetores médios) para cada classe temática, considerando-se as 42 amostras obtidas.

Para classificar a grade da varredura *laser* os valores armazenados em cada célula foram aplicados nas quatro funções discriminantes e obteve-se, também, centróides dos $g = 6$ grupos.

As coordenadas das células foram usadas para calcular a distância euclidiana de cada uma, a cada um dos seis centróides. A menor das seis distâncias determina a classe a que pertence à célula processada. A classificação foi codificada de forma crescente, sendo esta informação armazenada em cada célula.

Para esta codificação foram empregados números inteiros no intervalo de 1 a 6. A grade classificada possibilitou então criar uma imagem temática referentes as classes elencadas.

Para a verificação da eficiência da metodologia proposta foram totalizadas as áreas

referentes a cada uma das seis classes presentes na imagem classificada e com auxílio da fotointerpretação foram determinadas as áreas destas classes na imagem de intensidade. Estes valores permitiram reconhecer os percentuais de acerto para cada tipologia.

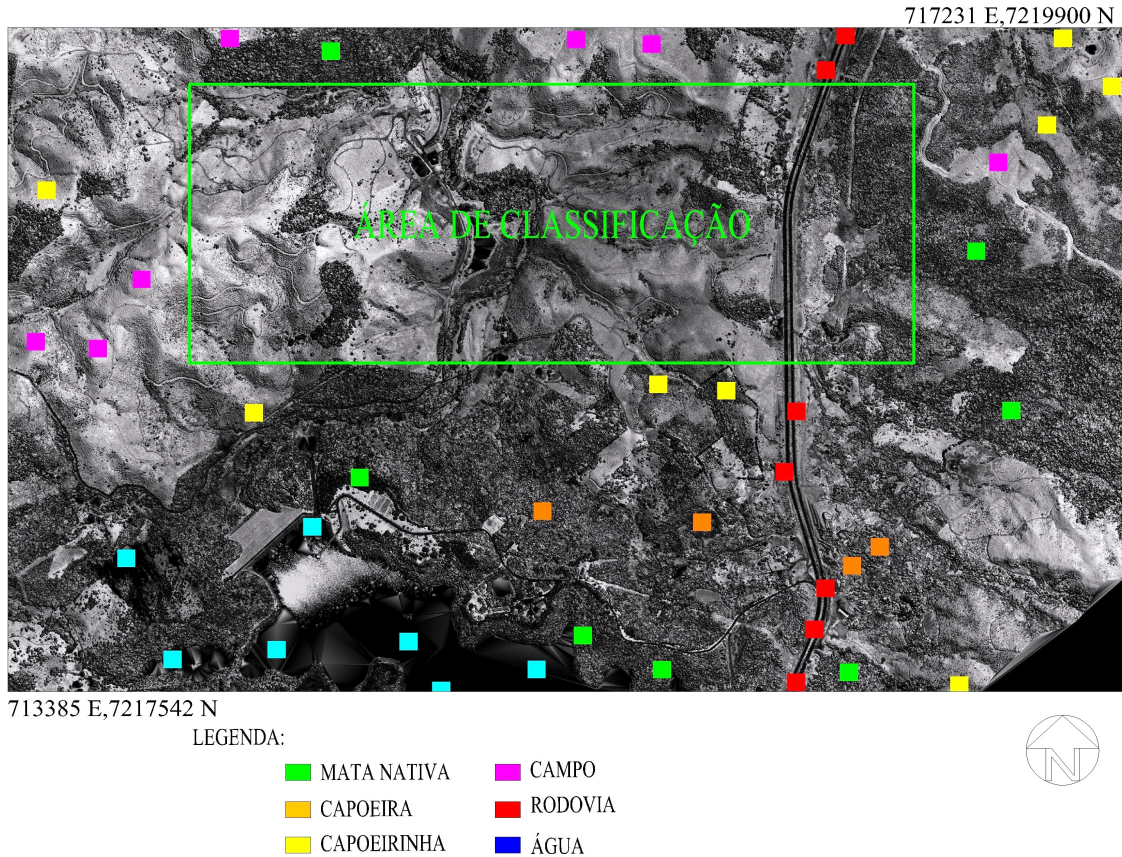


Figura 2. Distribuição das amostras na imagem de intensidade.

3. Resultados

3.1. Geração do modelo digital do terreno aproximado

A seleção de células da grade *laser* próximas ao terreno foi feita empregando-se os seguintes critérios de separação: variância altimétrica no interior da célula da grade menor do que 10; densidade de pontos no interior da célula da grade menor do que 0,95.

Esses critérios foram obtidos através da análise das células próximas do solo e sobre as

elevações, sendo separadas 62.933 das 93.235 células da grade *laser*. Porém, os critérios utilizados para separação de células próximas do solo não foram eficientes, uma vez que com a análise visual do MDT, constatou-se a permanência de células acima do terreno. A figura 3 apresenta o MDT para a área de estudo.

Deve-se destacar, no entanto, que o enfoque deste trabalho é possibilitar a separação das coberturas vegetais e que esta ineficiência pode interferir no resultado da classificação, uma vez que a altura da cobertura vegetal é uma variável que será empregada na classificação.

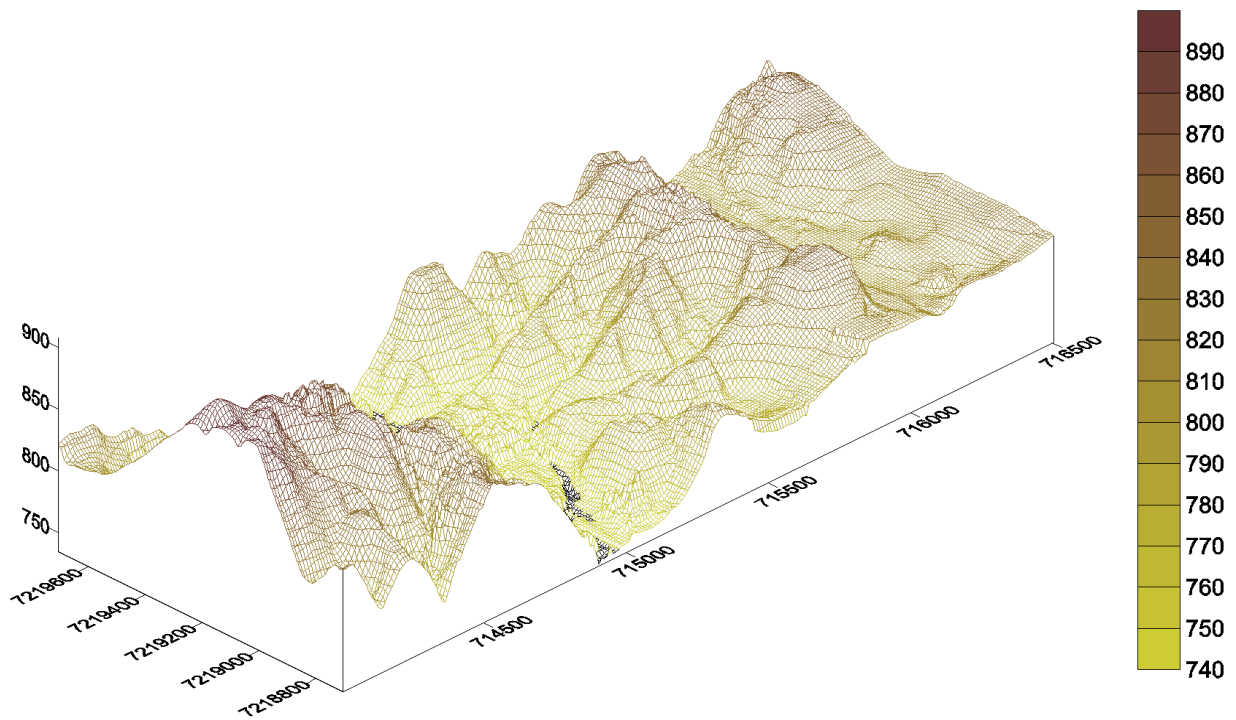


Figura 3. Modelo Digital do Terreno para a área de estudo.

3.2. Funções discriminantes

A partir das amostras obtidas, referentes a cada uma das seis classes, foram determinadas as funções discriminantes:

$$Y_1 = 0,5946561.X_1 + 0,2125909.X_2 + 0,0060188.X_3 - 0,027042.X_4;$$

$$Y_2 = -0,065099.X_1 + 0,2786487.X_2 + 0,105008.X_3 + 0,019193.X_4;$$

$$Y_3 = -0,020442.X_1 + 1,0050732.X_2 - 0,01028.X_3 - 0,008115.X_4;$$

$$Y_4 = -0,005563.X_1 + 0,0215538.X_2 - 0,00694.X_3 + 0,1376384.X_4.$$

Sendo:

- X_1 - intensidade média;
- X_2 - densidade de pontos;
- X_3 - variância altimétrica;
- X_4 - altura da cobertura vegetal.

Y_n - funções discriminantes.

Com o auxílio das funções discriminantes calcularam-se as coordenadas dos centróides para cada classe referente ao uso do solo. O quadro 1 apresenta os centróides para cada classe.

Quadro 1. Coordenadas dos centróides calculados para cada classe temática.

Classe Temática	Y_1	Y_2	Y_3	Y_4
Nativa	7,2592	5,4938	6,0119	0,6224
Capoeira	8,2927	0,0000	0,9735	1,2414
Capoeirinha	7,2315	0,0000	0,8902	0,2620
Campo	10,9732	0,0000	0,3070	0,9222
Água	1,4915	0,0000	-0,0413	-0,0121
Rodovia	7,9909	0,0000	0,5598	-0,0336

Com as funções discriminantes ajustadas e os centróides de cada classe definidos, procedeu-se a classificação individual para cada célula na área de classificação. Os valores referentes aos parâmetros intensidade, variância altimétrica, densidade e altura da cobertura vegetal (X_1, X_2, X_3, X_4) foram aplicados a cada uma das funções discriminantes, retornando valores (Y_1, Y_2, Y_3, Y_4).

Estes valores foram então utilizados para o cálculo da distância euclidiana em relação aos

centróides de cada classe, retornando seis distâncias. A menor das seis distâncias caracteriza a qual classe pertence à célula selecionada. A codificação da classe era então introduzida na célula processada. Para a geração do mapa temático referente à classificação foram adotados os seguintes valores: mata nativa, valor 6; capoeira, valor 5; capoeirinha, valor 4; campo, valor 3; água, valor 2; rodovia: valor 1.

A imagem temática obtida com a classificação estatística da grade *laser* é apresentada na figura 4.

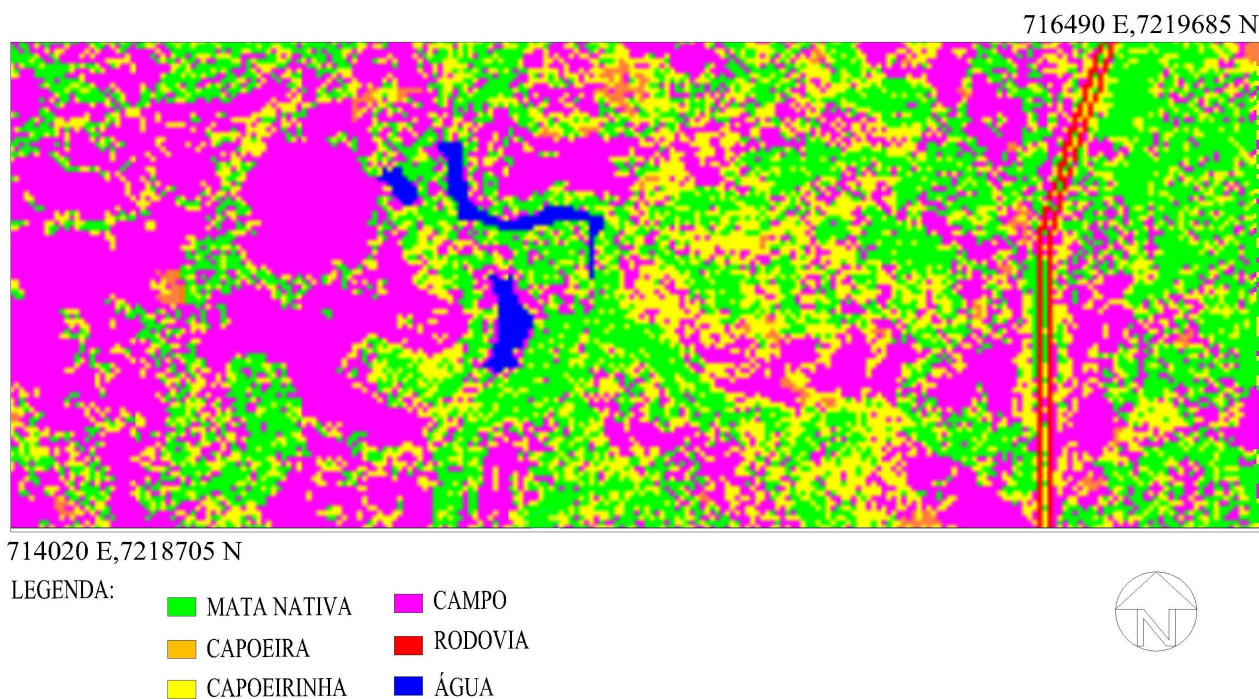


Figura 4. Imagem classificada para a área selecionada.

3.3 Geração e análise da imagem classificada

Para analisar o acerto da metodologia proposta foi feita a comparação entre os resultados obtidos pela classificação estatística com os resultados obtidos através da vetorização da imagem de intensidade auxiliada pela fotointerpretação da aerofoto. Neste caso, optou-

se por trabalhar com a imagem de intensidade evitando a necessidade de retificação da fotografia aérea. Deve-se destacar, no entanto, que a classificação foi manual e, portanto, podem ocorrer pequenas variações de um intérprete para outro. A partir da totalização de áreas (em hectares) pelos dois processos, elaborou-se o quadro 2.

Quadro 2. Índice de acerto e as áreas totalizadas através da vetorização e classificação estatística.

Classe Temática	Vetorização (ha)	Classificação (ha)	Diferença (ha)	Acerto (%)
Nativa	23,09	35,52	-12,43	65,00
Capoeira	65,54	46,82	18,72	71,43
Capoeirinha	32,14	84,56	-54,96	38,01
Campo	128,67	82,83	45,84	64,37
Água	2,96	2,53	0,43	85,47
Rodovia	1,94	2,08	-0,14	93,27

A análise do quadro 2 possibilita determinar que o acerto médio geral para a classificação estatística foi de 69,60 %.

Os piores resultados foram obtidos na classificação da capoeirinha, campo e mata nativa. Este valor baixo no acerto da classificação pode estar relacionado com a variável altura da cobertura vegetal e, conseqüentemente, com a escolha de pontos próximos do solo.

Além disso, destaca-se o acerto geral da água e da estrada, simultaneamente 85,47 e 93,27%. A água, um caso específico que tem a particularidade de absorver quase totalmente a energia incidente proveniente do feixe laser.

4. Conclusões e recomendações

Após a análise dos resultados obtidos pode-se concluir que as variáveis escolhidas: intensidade média, variância altimétrica e altura da cobertura vegetal possibilitaram a classificação da nuvem de pontos gerada pelo ALS empregando a FDL de Fischer.

A fotointerpretação da imagem possibilitou a definição das classes temáticas e obtenção de amostras referentes a estas classes, minimizando a necessidade de inspeções de campo. A classificação proposta é promissora e merece um estudo mais aprofundado visando o estabelecimento de uma metodologia que possibilite uma maior acurácia na separação das classes temáticas.

A comparação da classificação obtida com o emprego das funções discriminantes com a vetorização manual detectou erros, provavelmen-

te, relacionados à variável altura da cobertura vegetal. Esta variável poderá ser mais bem determinada através da utilização de uma metodologia (filtro) mais eficiente para a separação dos pontos do solo e das elevações. Contudo, o acerto médio obtido reforça o potencial que a metodologia proposta apresenta.

A classificação a partir dos dados brutos provenientes do *laser scanning* foi satisfatória. Destaca-se, porém, que talvez sejam necessárias além das informações *laser*, outras fontes de informações espectrais, como aerofotos e imagens de satélite de alta resolução.

Referências

- Baltsavias, E. P. 1999. A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54: 83-94.
- Brasil. *Lei n° 4.741*, 15 de setembro de 1965.
- Dalmolin, Q. & Santos, D. R. 2003. *Sistema laser scanner: conceitos e princípios de funcionamento*. Curitiba, Ed. da UFPR, 94p.
- Hair, J.F. 2005. *Análise multivariada de dados...* Porto Alegre, Bookman, 5ª ed, 688p.
- Holmgren, J. 2003. Estimation of Forest Variables using Airborne Laser Scanning. Swedish. Umea, 95 p. Doctor Thesis, Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Sciences.
- Johnson, D. E. 1998. Applied multivariate methods for data analysts. Duxbury Press, U.S.A.
- Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K., Dzeroski, S. 2007. Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. *Remote Sensing of Environment*, 108: 9-23.