



20 e 21 de outubro
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP

Avaliação da influência do uso de rotulação hierárquica em classificação de imagens de sensoriamento remoto

Eliana Pantaleão¹, Luciano Vieira Dutra²

¹Programa de Doutorado em Computação Aplicada – CAP, INPE

²Divisão de Processamento de Imagens - INPE

{elianap, dutra}@dpi.inpe.br

Abstract. *This paper evaluates the influence of using hierarchical labeling in the classification accuracy of remotely sensed images. In hierarchical labeling, structured labels are associated to the training samples to be used in supervised classification. Two kinds of classifiers are analyzed: parametric and non-parametric. Tests show they behave differently when class structures are used to guide the classification steps. In this work, the classifiers tested were maximum likelihood and support vector machines. Kappa coefficient was used to evaluate the results.*

Resumo. *Este artigo avalia a influência do uso de rotulação hierárquica na acurácia da classificação de imagens de sensoriamento remoto. Com o uso de rotulação hierárquica, estruturas de rótulos são associadas às amostras de treinamento usadas na classificação. Dois tipos de classificadores são analisados: paramétricos e não-paramétricos. Testes mostram que eles se comportam de forma diferente quando estruturas de classes são usadas para guiar os passos de classificação. Neste trabalho, os classificadores testados foram máxima verossimilhança e máquinas de vetores-suporte. O coeficiente Kappa foi utilizado para avaliar os resultados.*

Palavras-chave: *sensoriamento remoto, classificação de imagens, hierarquia de classes.*

1. Introdução

A área de sensoriamento remoto ótico compreende a análise e interpretação de um conjunto de medidas da radiação eletromagnética (EMR) que é refletida ou emitida por um alvo e observada e registrada a distância por um instrumento que não está em contato com o alvo [MATHER, 2004]. Essas medidas podem ser registradas por plataformas orbitais, como ilustrado na Figura 1.

No caso da observação da superfície da Terra, o resultado dessas medidas são matrizes de valores digitais, as quais são normalmente chamadas de imagens. A geração de mapas temáticos através da interpretação dos dados existentes nessas imagens é o interesse principal do sensoriamento remoto. Para atingir esse objetivo, diversos métodos foram desenvolvidos, entre eles a classificação supervisionada dessas imagens.

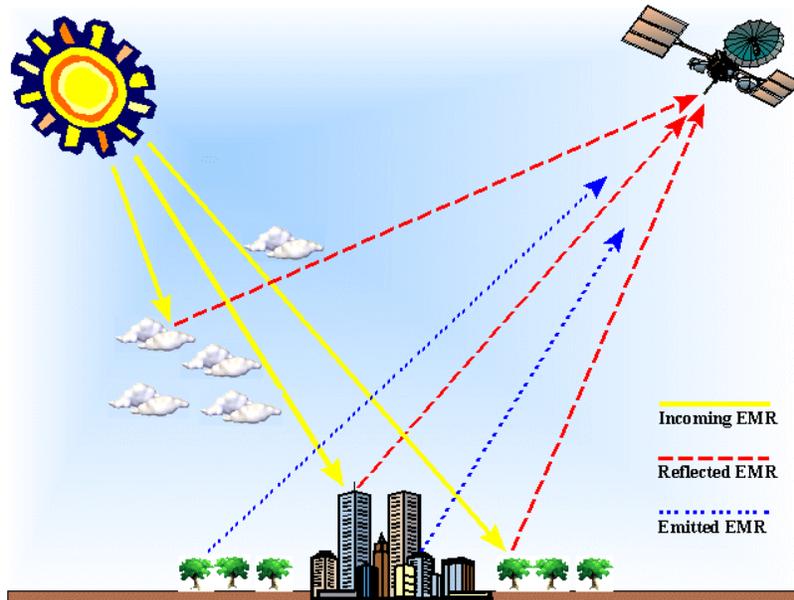


Figura 1. Exemplo de atividade de sensoriamento remoto [USC,2003].

Conforme expõe Jensen (2007), quando se realiza uma classificação supervisionada em imagens de sensoriamento remoto, a identidade e localização de alguns dos tipos de cobertura do solo, como área urbana, agricultura, floresta, água e outros, são conhecidos *a priori* através da combinação de diversas fontes. Entre essas, pode-se citar o trabalho de campo, a interpretação de imagens aéreas, a análise de mapas previamente obtidos e a experiência pessoal do interpretador. O analista procura localizar feições específicas nos dados de sensoriamento remoto que representem exemplos homogêneos destes tipos de cobertura do solo conhecidos. Essas áreas são normalmente chamadas de amostras de treinamento, pois suas características (em especial as espectrais) são utilizadas para treinar o algoritmo de classificação que realizará o mapeamento do restante da imagem.

Este artigo propõe uma forma de rotular as amostras de treinamento de forma hierárquica e analisa a influência do aprendizado hierárquico na acurácia da classificação de imagens de sensoriamento remoto. São analisados dois tipos de classificadores: os paramétricos e os não-paramétricos.

2. Rotulação hierárquica

Neste trabalho, propõe-se a utilização de uma estrutura de classes como a mostrada na Figura 2. Essa estrutura é uma árvore de classes em que cada filho representa uma subclasse da classe mãe (superclasse).

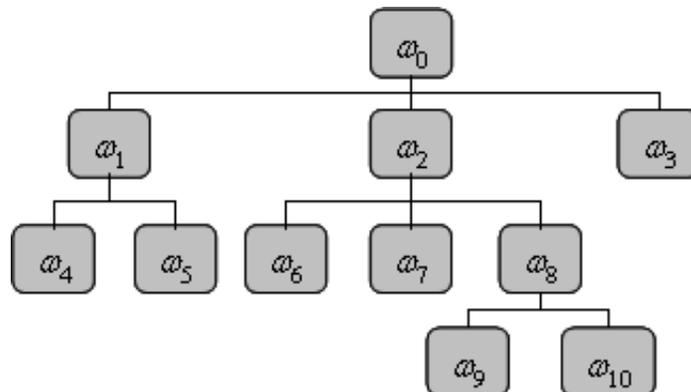


Figura 2. Estrutura de classes

Com o uso da rotulação hierárquica, cada amostra de treinamento receberá, ao invés de um rótulo simples, como “ w_4 ”, um rótulo estruturado que representa o caminho, na estrutura de classes, desde a raiz até a classe que se deseja rotular, como por exemplo “ w_1-w_4 ” ou “ $w_2-w_8-w_{10}$ ”. Não é necessário que o último rótulo seja uma folha, também é possível usar “ w_2-w_8 ” ou “ w_2 ”. A raiz da árvore, a classe w_0 , representa a classe “universo” e não possui qualquer papel na tarefa de classificação, sendo, por isso, excluída do rótulo composto.

A ideia é utilizar estas amostras hierarquicamente, ou seja, realizar inicialmente a classificação para o primeiro nível da estrutura e utilizar cada resultado obtido na classificação dos demais níveis. Em outras palavras, o classificador aprende hierarquicamente cada uma das classes da estrutura. Apenas padrões associados à superclasse serão apresentados para a classificação em suas subclasses. Como é possível utilizar rótulos que não terminam em folhas, algumas amostras podem ser descartadas nas últimas etapas da classificação.

Essa estrutura pode ser fornecida pelo usuário ou gerada por algum algoritmo de análise de amostras. No primeiro caso, sua estrutura será semântica, ou seja, o usuário cria esta estrutura conforme sua visão da cena e dos objetos que deseja classificar. Desta forma, além das classes em si, o usuário fornece mais informações sobre os objetos existentes na cena, de acordo com sua percepção, na forma de uma estrutura que vai do geral para o específico. No segundo caso, a estrutura reflete a organização das classes no espaço de atributos considerado. Por exemplo, se o usuário está interessado em quatro classes de objetos, as estruturas possíveis estão mostradas na Figura 3. A classificação tradicional pode ser simulada pela utilização da estrutura da Figura 3a.

As vantagens da utilização de estruturas hierárquicas para auxiliar na classificação foram observadas, entre outros, por Shahbaba e Neal (2007), Dumais e Chen (2000) e Koller e Sahami (1997). De acordo com a proposta deste trabalho, no fim da tarefa de classificação, os pixels são associados com os rótulos compostos, o que permite a visualização da imagem classificada em diferentes níveis da estrutura de classes.

Todos os algoritmos utilizados neste trabalho foram implementados em IDL [ITT 2008], dentro do programa SACI [PANTALEAO e DUTRA 2008].

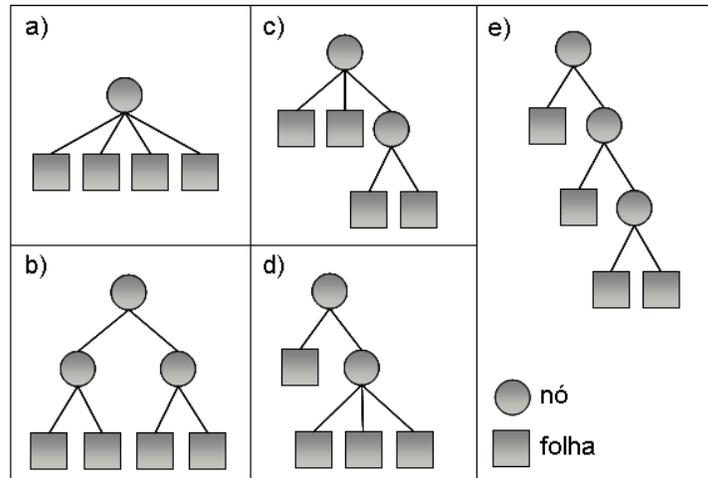


Figura 3. Possíveis estruturas para 4 classes de interesse

3. Classificadores paramétricos e não-paramétricos

Um classificador supervisionado é denominado paramétrico quando utiliza parâmetros estatísticos na estimação do modelo das classes dos dados a partir das amostras de treinamento. Ou seja, na fase de treinamento, o classificador extrai parâmetros estatísticos das amostras e utiliza algum método para estimar os parâmetros dos dados originais. Essas estimativas são utilizadas na classificação dos padrões desconhecidos, ou seja, no restante da imagem.

Para isso, é necessário assumir alguma hipótese sobre a distribuição dos dados originais. No caso do classificador de máxima verossimilhança (MAXVER) utilizado neste trabalho, a hipótese é a normalidade, ou seja, os dados possuem distribuição Gaussiana. Em geral, acredita-se que essa hipótese é verdadeira para dados de sensoriamento remoto, em que as classes apresentam grande concentração em torno da média. Sendo assim, os parâmetros estimados são as médias, variâncias e covariâncias entre os canais para cada classe.

O MAXVER é um algoritmo de classificação pontual, ou seja, classifica um pixel de cada vez. Cada pixel é representado por seu vetor de atributos, que corresponde aos valores digitais medidos nas várias faixas de frequência, de acordo com o sensor utilizado. A decisão de atribuir uma classe a um determinado padrão é baseada na minimização média de perdas, ou seja, na probabilidade estimada da pertinência do pixel à classe, comparada com a probabilidade de pertinência às demais classes [THEODORIDIS e KOUTROUMBAS 2006].

Um problema que pode ocorrer com este tipo de classificador é quando o número de amostras de treinamento é pequeno em relação à dimensionalidade dos dados. Nessa situação, tem-se o Fenômeno de Hughes ou Maldição da Dimensionalidade, que ocorre devido à impossibilidade da obtenção de estimativas confiáveis de um número crescente de parâmetros da distribuição para um número fixo de amostras. Ou seja, quando se aumenta a dimensionalidade dos dados, apesar de se adicionar informação ao classificador, a acurácia diminui. Nesse caso, é necessário aplicar alguma técnica para redução da dimensionalidade sem perda significativa de informação.

Já a classificação não-paramétrica não assume qualquer hipótese sobre a distribuição dos dados. Em geral, esses métodos procuram localizar regiões de baixa densidade de dados entre as classes conhecidas, estimando separadores entre elas. Alguns exemplos de classificadores não-paramétricos são árvores de decisão (DT), redes neurais (RN) e máquinas de vetores suporte (SVM).

O método de SVM [Vapnik 1995] emprega algoritmos de otimização para localizar os limites ótimos entre as classes. Essas fronteiras ótimas podem ser generalizadas para padrões desconhecidos com um número mínimo de erros, entre todas as fronteiras possíveis, desta forma minimizando a confusão entre as classes. Neste trabalho, foi utilizado o SVM pontual, com estratégia multiclasse um-para-todos, *kernel* RBF.

De acordo com Huang et al. (2002), o SVM não é afetado pela Maldição da Dimensionalidade. Isso é confirmado por diversos experimentos, como Melgani e Bruzzone (2004) que utilizaram SVM com sucesso para a classificação de imagens hiperespectrais, sem a necessidade da redução da dimensão dos dados.

4. Experimentos e resultados

Inicialmente, a influência da classificação hierárquica foi testada na imagem sintética mostrada na Figura 4. Esta imagem contém quatro quadrados de mesmo tamanho (256 x 256 pixels), nas cores azul, verde, vermelho e amarelo. Cada uma das regiões não é homogênea, contendo ruído Gaussiano. O ruído é principalmente visível na região amarela, em que podem ser observados diversos pontos verdes.

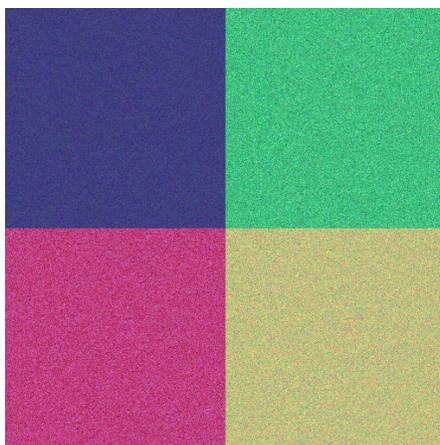


Figura 4. Imagem sintética.

Esta imagem é bastante apropriada para este estudo devido às características de seus três canais, R, G e B, cuja visualização encontra-se na Figura 5. Pode-se observar que no canal R, as classes azul e verde (parte superior) podem ser facilmente discriminadas das classes vermelho e amarelo (parte inferior). Algo diferente ocorre no canal G, em que as classes da esquerda (azul e vermelho) possuem grande dissimilaridade com as da direita (verde e amarelo). O canal B é irrelevante para a classificação, uma vez que as classes só são distinguíveis pela textura, e este atributo não é utilizado nos testes. A média de cada classe por canal pode ser vista na Tabela 1, em que se pode observar a mesma média para todas as classes no canal B.

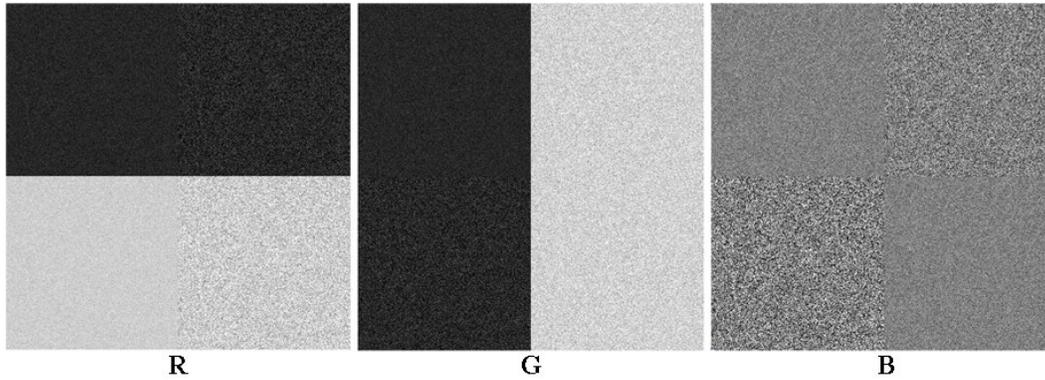


Figura 5. Canais RGB da imagem sintética.

Tabela 1. Média das classes por canal.

Classe	Localização	R	G	B
Azul	Quadrante superior esquerdo	64.00	63.98	128.01
Verde	Quadrante superior direito	64.02	191.98	128.01
Vermelho	Quadrante inferior esquerdo	192.02	63.99	128.03
Amarelo	Quadrante inferior direito	191.83	191.94	128.01

Primeiramente foi realizada a classificação tradicional. Foram coletadas 4 amostras com 10x10 pixels, cada uma correspondendo a uma das classes Azul, Verde, Vermelho e Amarelo. Essas amostras foram submetidas ao treinamento dos classificadores MAXVER e SVM, e os resultados da classificação podem ser vistos na Figura 6. As acurácias obtidas foram, respectivamente, 99.8448% e 99.7882%, com valores de Kappa iguais a 0.9979 e 0.9972. As amostras de teste utilizadas para a estimativa da acurácia foram quadrados de 254x254, ou seja, eliminadas apenas as bordas de cada região.

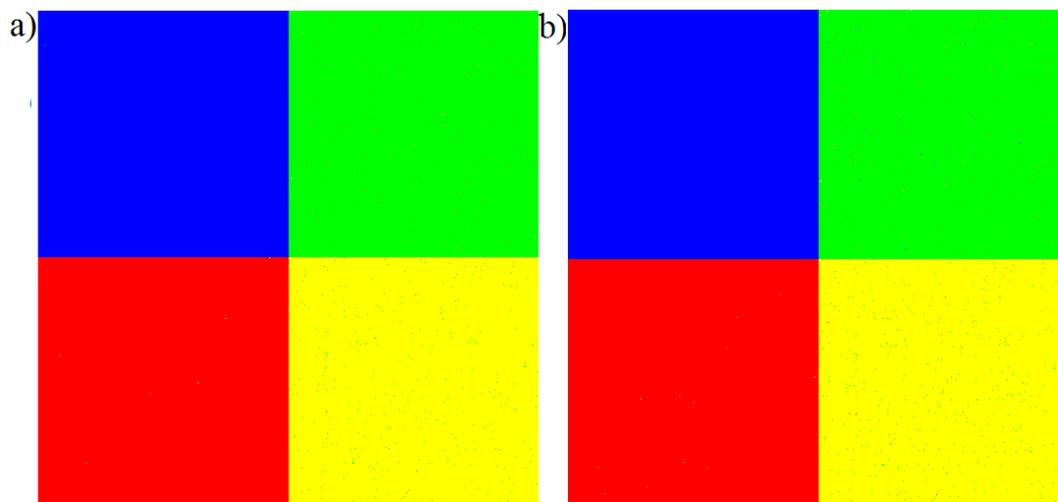


Figura 6. Imagens classificadas por (a) MAXVER e (b) SVM.

A seguir, as amostras foram rotuladas hierarquicamente. Devido às características dos canais mostradas na Figura 5 e confirmadas pelas médias na tabela 1, optou-se por utilizar duas variações da estrutura de classes da Figura 3b, que podem ser vistas na Figura 7. Os resultados das 4 classificações são mostrados na Figura 8 e seus valores de Kappa, na Tabela 2.

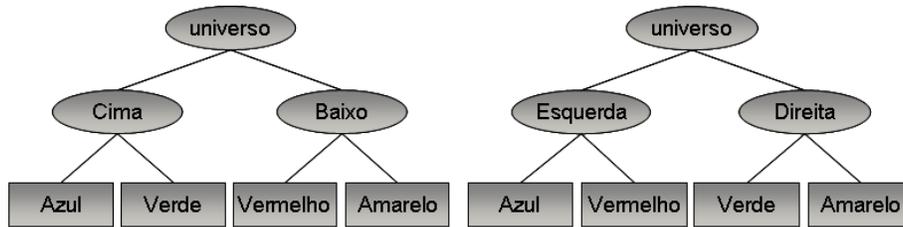


Figura 7. Estruturas de classe utilizadas nos testes com rotulação hierárquica.

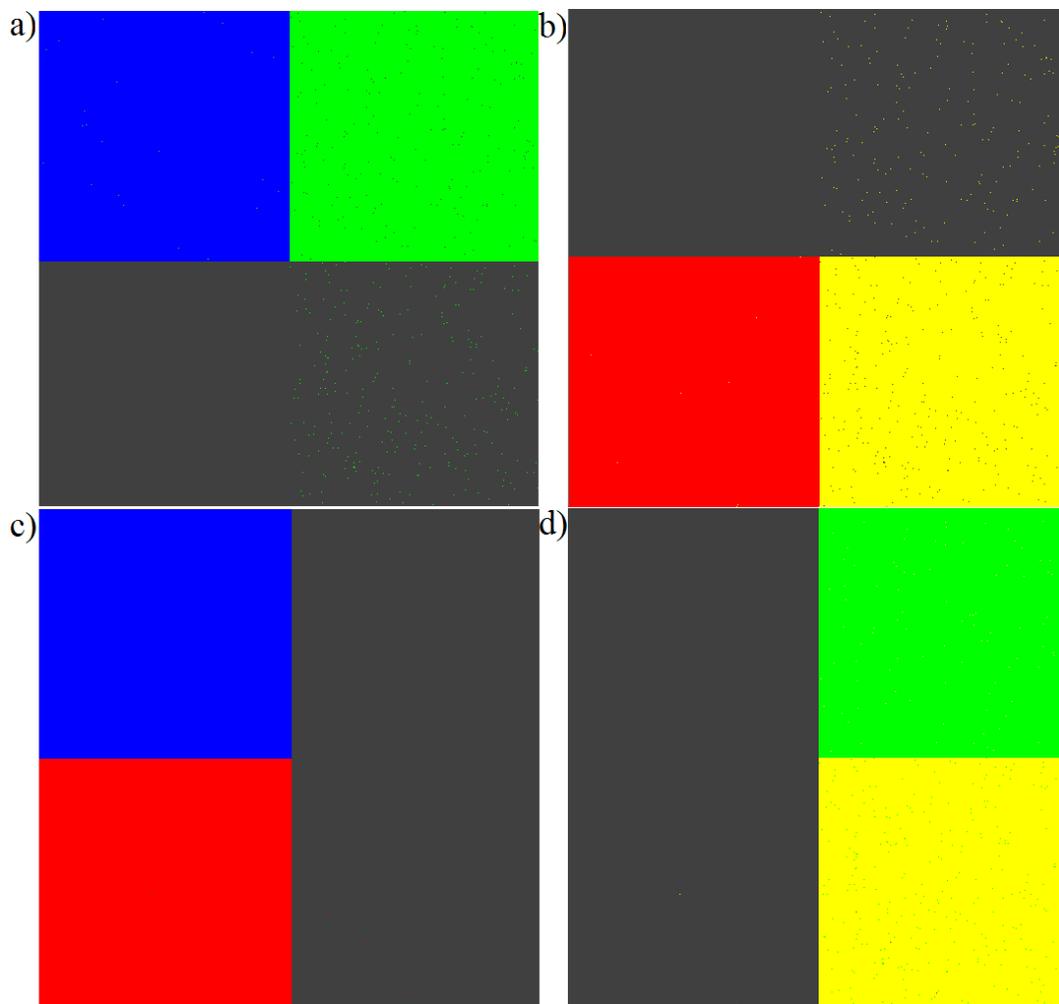


Figura 8. Resultados das classificações MAXVER:
 a) Cima b) Baixo c) Esquerda d) Direita.

Tabela 2. Kappa para os resultados das classificações da imagem Sintética.

Classificador	Hierarquia	Kappa nível 1	Kappa nível 2 classe 1	Kappa nível 2 classe 2
MAXVER	Cima/Baixo	0.9963	0.9997	0.9999
MAXVER	Esquerda/Direita	1.0000	1.0000	0.9940
SVM	Cima/Baixo	0.9967	0.9998	0.9997
SVM	Esquerda/Direita	0.9998	1.0000	0.9923

Todas as classificações foram realizadas com as mesmas amostras de treinamento, apenas excluídas as que não se aplicavam à classificação corrente, conforme o nível na estrutura de classes. Para a classe Cima, por exemplo, foram agrupadas as amostras das classes Azul e Verde, e o mesmo foi realizado para as demais classes da hierarquia.

Apesar de a classificação clássica (sem rotulação hierárquica) já apresentar altíssima acurácia, podemos observar um pequeno aumento no valor de Kappa. O melhor resultado ocorreu com a utilização da estrutura Esquerda/Direita, com o classificador Maxver. Este classificador teve bom desempenho neste exemplo devido à peculiaridade do espaço de atributos (Figura 9).

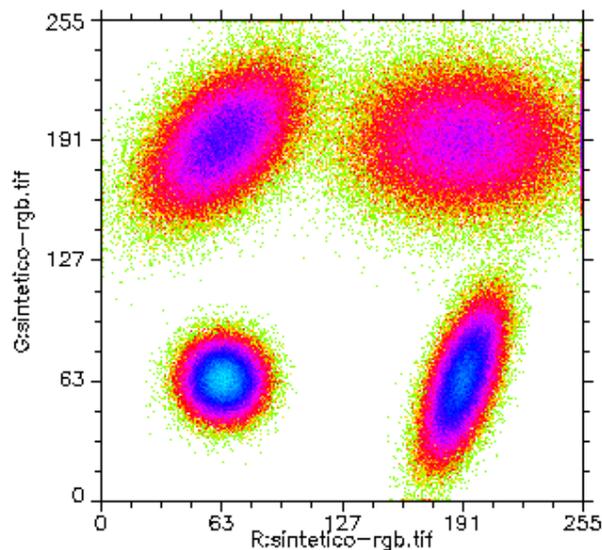


Figura 9. Classes Verde, Amarelo, Azul e Vermelho no espaço RxG (vermelho x verde).

A seguir, o mesmo tipo de experimento foi repetido para uma imagem de sensoriamento remoto. A imagem escolhida foi da área da Floresta de Tapajós, no estado do Pará (Figura 10). A imagem do satélite IRS (*Indian Remote Sensing Satellite*) – P6 (RESOURCESAT-I) possui resolução espacial de 23,5 metros e foi adquirida no dia 10 de setembro de 2009. Foram utilizados os canais azul, verde, vermelho e infravermelho próximo.

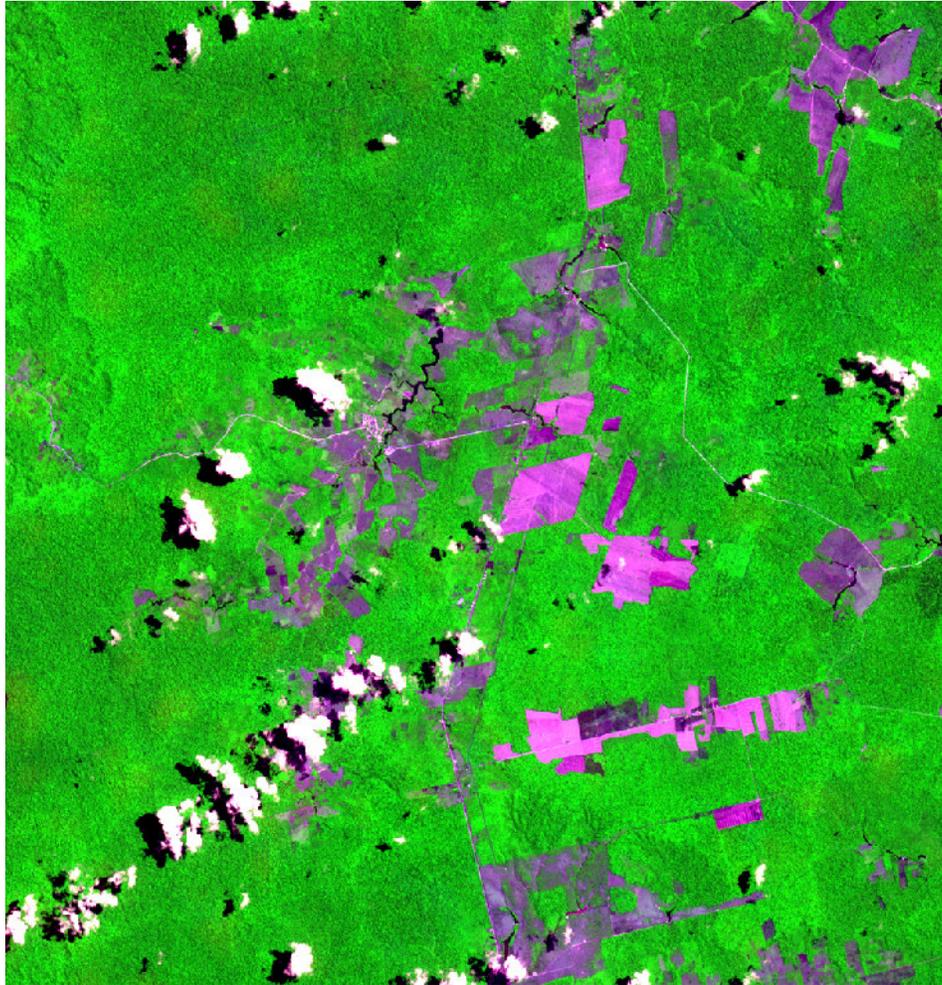


Figura 10. Região de estudo no Pará.

Foram coletados 1593 pixels para treinamento e 818 pixels para teste, verificados em trabalho de campo realizado em setembro de 2009 e rotulados conforme a hierarquia de classes mostrada na Figura 11.

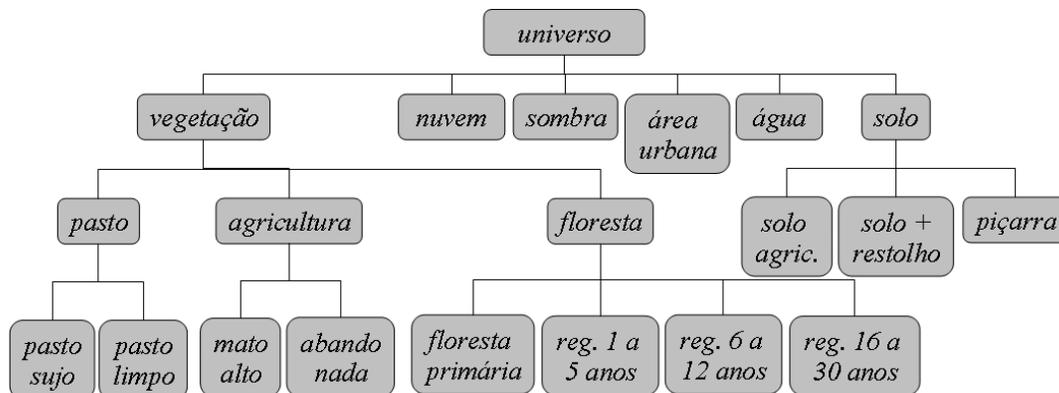


Figura 11. Hierarquia de classes utilizada.

Duas classificações foram realizadas para cada classificador (MAXVER e SVM), utilizando níveis diferentes da árvore de classes da Figura 11. A primeira foi realizada para todas as folhas da hierarquia, no modo tradicional (Figura 12) e obteve os valores de Kappa respectivamente 0.5220 e 0.6129. A segunda foi feita hierarquicamente conforme a estrutura de classes apresentada e o resultado do primeiro nível pode ser visto na Figura 13.

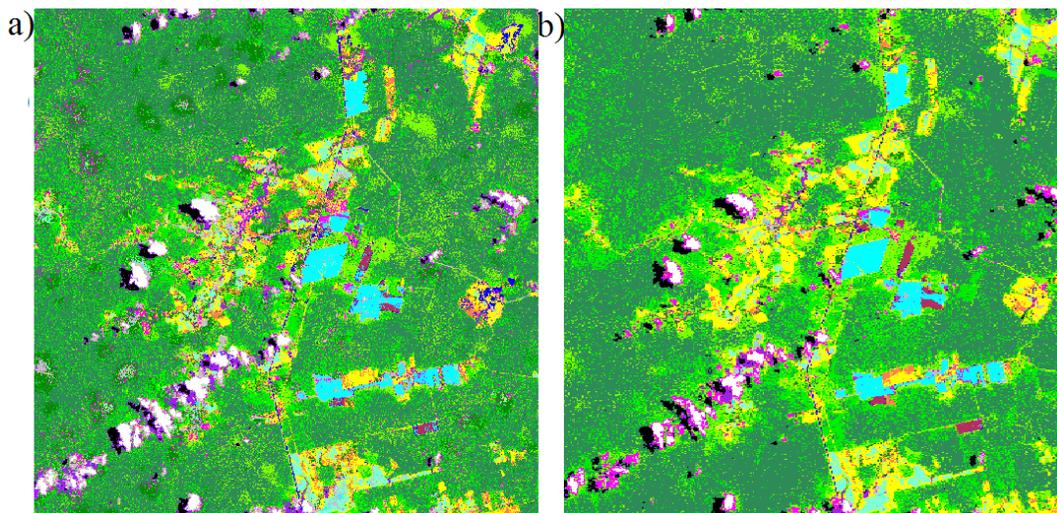


Figura 12. Classificação tradicional com a) MAXVER e b) SVM.

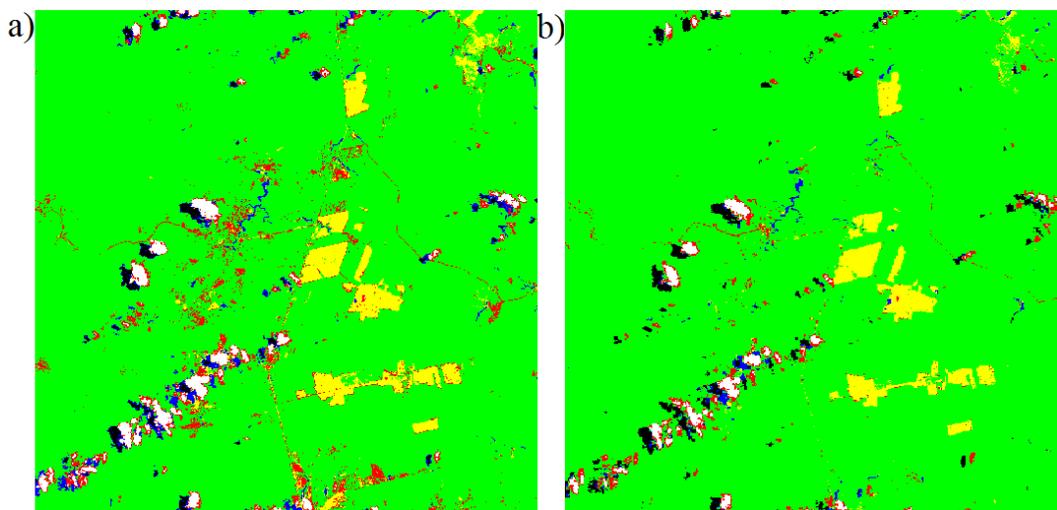


Figura 13. Primeiro nível da classificação hierárquica com a) MAXVER e b) SVM.

Neste experimento, a rotulação hierárquica funcionou melhor para o classificador SVM, pois o espaço de atributos desta imagem para estas classes é muito mais complexo, como pode ser visto na Figura 14. Os valores de Kappa encontrados para o nível 1 foram de 0.9098 e 0.9339 respectivamente e, para o nível 2, 0.7677 e 0.8124 para a classe vegetação e 1.000 e 0.9719 para a classe solo.

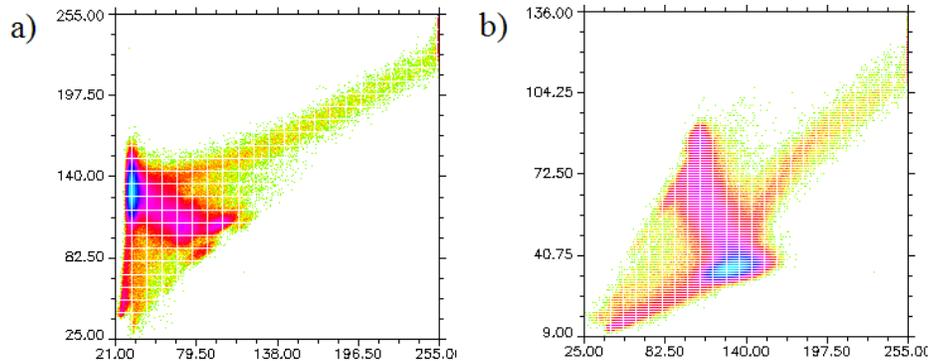


Figura 14. Espaço de atributos para a imagem de Tapajós, canais a) RxG e b) GxIR.

Referências

- Dumais, S., Chen, H. (2000) "Hierarchical Classification of Web Content". In: Proceedings of the 23rd Annual international ACM SIGIR Conference on Research and Development in Information Retrieval (Athens, Greece, July 24 - 28, 2000). SIGIR '00. ACM, New York, NY, 256-263.
- Huang, C.; Davis, L.S.; Townshend, J.R.G. (2002) "An Assessment of Support Vector Machines for Land Cover Classification". Intern. J. of Remote Sensing, vol.23, n4.
- ITT (2008). Visual information solutions. <http://www.itvis.com/>.
- Jensen, John R. (2007) "Remote Sensing of the Environment: An Earth Resource Perspective". 2. ed., Upper Saddle River, NJ: Prentice Hall.
- Koller, D., Sahami, M. (1997) "Hierarchically Classifying Documents Using Very Few Words". In: Proceedings of the Fourteenth International Conference on Machine Learning (July 08 - 12, 1997). D. H. Fisher, Ed. Morgan Kaufmann Publishers, San Francisco, CA, 170-178.
- Mather, P. M. (2004) "Computer processing of remotely-sensed images: an introduction". 3. ed. NY: John Wiley & Sons.
- Melgani, F.; Bruzzone, L. (2004) "Classification of Hyperspectral Remote Sensing Images with Support Vector Machines". IEEE Transactions on Geoscience and Remote Sensing, vol.42, n.8, agosto.
- Pantaleão, E.; Dutra, L. V. (2008) "Feature Selection with SACT". VIII WORCAP.
- Shahbaba, B., Neal, R.M. (2007) "Improving Classification When a Class Hierarchy is Available Using a Hierarchy-Based Prior". Bayesian Analysis, vol.2, n1, pp.221-238.
- Theodoridis, S.; Koutroumbas, K. (2006). Pattern Recognition. Academic Press, San Diego, 3. ed.
- USC (University of South Carolina) (2003) "Introductory digital image processing: The remote sensing process". Disponível em: <<http://www.cas.sc.edu/geog/rsrab/Rsc/fmod1.html>>. Acessado em novembro, 2008.
- Vapnik, V. N. (1995) "The nature of statistical learning theory". Springer-Verlag.