

EMPREGO DOS ÍNDICES DE AUTOCORRELAÇÃO ESPACIAL NA AVALIAÇÃO DA SEGMENTAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO

Espindola, G.¹

1- INPE, Instituto Nacional de Pesquisas Espaciais, www.inpe.br - giovana@dpi.inpe.br

RESUMO – Os algoritmos de segmentação têm sido amplamente usados na extração de informações de imagens de sensoriamento remoto. Visto que segmentar uma imagem significa agrupar pixels vizinhos em regiões com base em critérios de similaridade, a obtenção de resultados satisfatórios na classificação orientada a objetos é diretamente influenciada por bons resultados na etapa de segmentação. O desempenho dos algoritmos de segmentação é fortemente dependente de parâmetros internos que, em geral, podem ser manipulados pelo usuário. Em consequência, a avaliação de resultados da segmentação não é uma tarefa trivial, sendo importante o estudo de técnicas confiáveis de avaliação da qualidade dos algoritmos e de seus parâmetros. Assim, este trabalho discute o uso de indicadores de autocorrelação espacial na avaliação da qualidade dos algoritmos de segmentação através da mensuração de características como homogeneidade interna dos segmentos e contraste entre vizinhança. As hipóteses do trabalho foram avaliadas comparando-se os resultados da segmentação de imagens LANDSAT e CBERS.

ABSTRACT – Segmentation algorithms have been often used for extracting information in remote sensing images. Given that segmentation consists in a process where the pixels of an image are grouped into homogeneous contiguous areas based on similarity criteria, segmentation algorithms have many advantages over pixel-based image classifiers. Their performance is strongly dependent on ad-hoc parameters provided by the user, and for that reason it is very important to devise reliable techniques to evaluate the quality of segmentation algorithms and their parameters. As a result, evaluation of segmentation results is non trivial. This paper discusses the use of spatial autocorrelation indicators as a tool for evaluating the quality of segmentation algorithms. The general idea is that a good segmentation has two qualities from a spatial statistics viewpoint: homogeneity into regions and enhanced neighborhood regions. The stated hypotheses were validated by comparing different segmentation results on LANDSAT and CBERS images.

PALAVRAS CHAVE: remote sensing, image processing, ESDA.

1 INTRODUÇÃO

As altas taxas de desflorestamento da Amazônia e seus possíveis impactos negativos no meio ambiente têm motivado uma série de estudos científicos que buscam entender as causas e as proporções do desflorestamento. Nesta linha, o Instituto Nacional de Pesquisas Espaciais (INPE) vem, desde 1989, produzindo estimativas anuais de taxas de desflorestamento da Amazônia Legal. Tais estimativas são produzidas pelo Projeto PRODES Digital, que utiliza técnicas de processamento digital de imagens visando à geração de um banco de dados geográfico multitemporal (INPE, 2004b).

Para Alves (2001), o mapeamento de áreas desflorestadas é tradicionalmente feito através da análise de imagens do satélite LANDSAT. No entanto, segundo dados do *The United States Geological Survey* (USGS, 2004), desde maio de 2003, o satélite LANDSAT-7 apresenta problemas na aquisição de imagens ETM+. Tal problema tem afetado toda a comunidade de usuários que têm optado pelo uso de imagens TM do satélite LANDSAT-5. Não obstante, este último, lançado em 1984, já está além do seu tempo de operação previsto, devendo ser desligado nos próximos anos. A lacuna deixada pelos satélites da família LANDSAT deverá ser preenchida por diversos outros satélites que sejam compatíveis com a demanda dos usuários no que diz respeito à qualidade e a acessibilidade dos dados,

podendo-se citar, o EOS Terra, o EOS Aqua e o CBERS-2.

Dos produtos gerados pelo Projeto PRODES, os mapas temáticos de desflorestamento podem ser citados como os de maior importância, dado que são as bases para a obtenção das taxas de desflorestamento bruto e anual. Todavia, mesmo contemplando uma rotina de segmentação e classificação de imagens, a metodologia do PRODES não possui uma rotina de validação dos dados gerados, uma vez que os possíveis erros dos mapas temáticos são corrigidos na etapa de edição matricial.

Neste sentido, este trabalho tem por objetivo utilizar índices de autocorrelação espacial para avaliar a qualidade da segmentação gerada pelo Projeto PRODES. Das técnicas de análise exploratória de dados espaciais são usadas neste trabalho, a matriz de proximidade espacial e os índices de Moran e de Geary.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Os Satélites LANDSAT-5 e CBERS-2

O satélite LANDSAT-5 foi lançado em 11 de março de 1984 tendo abordo o sensor Thematic Mapper (TM). Este sensor fornece imagens em 7 bandas espectrais com uma resolução espacial de 30 m, para cenas com tamanho de 170 x 183 km e resolução temporal de 16 dias (USGS, 2004).

Lançado em 21 de outubro de 2003, o satélite CBERS-2 apresenta as mesmas características técnicas do CBERS-1 (INPE, 2004a). Analogamente, o satélite CBERS-2 acomoda os sistemas ópticos CCD – Câmera Imageadora de Alta Resolução, IRMSS – Imageador por Varredura de Média Resolução e WFI – Imageador de Amplo Campo de Visada. Em sua órbita, o satélite cruza o equador sempre na mesma hora local, demorando 26 dias para retornar ao mesmo ponto de cobertura da Terra. A câmera CCD do satélite CBERS fornece imagens em 5 bandas espectrais, com resolução de 20 m. A Tabela 1 apresenta, comparativamente, as bandas do CCD/CBERS-2 (INPE, 2004) e do TM/LANDSAT-5 (USGS, 2004).

Bandas	CCD/CBERS-2	TM/LANDSAT-5
PAN	0,51 – 0,73 μm	-----
B1	0,45 – 0,52 μm	0,45 – 0,52 μm
B2	0,52 – 0,59 μm	0,52 – 0,60 μm
B3	0,63 – 0,69 μm	0,63 – 0,69 μm
B4	0,77 – 0,89 μm	0,76 – 0,90 μm
B5	-----	1,55 – 1,75 μm
B7	-----	2,08 – 2,35 μm
Resolução	20 m	30 m

Tabela 1. Comparação entre as bandas do CCD/CBERS-2 e TM/LANDSAT-5.

2.2 O Projeto PRODES

Para Duarte et al. (2003), o Projeto de Estimativa do Desflorestamento Bruto da Amazônia – Projeto PRODES Digital – tem como objetivo mapear a extensão do desflorestamento bruto da Amazônia Brasileira através do processamento digital de imagens TM/LANDSAT-5. Tal processamento engloba, prioritariamente, as técnicas de modelo linear de mistura espectral, segmentação de imagens por crescimento de regiões e classificação por regiões, todas implementadas e disponíveis no software SPRING (Sistema para Processamento de Informações Georeferenciadas), desenvolvido pelo INPE (2004c).

2.3 Técnicas de Processamento Digital de Imagens

2.3.1 Modelo Linear de Mistura Espectral

O modelo linear de mistura espectral (MLME) visa estimar para cada pixel da imagem a proporção dos componentes solo, vegetação e sombra, a partir da resposta espectral nas bandas do sensor, gerando as imagens fração solo, vegetação e sombra (Shimabukuro e Smith, 1991). Este modelo pode ser descrito segundo a equação a seguir:

$$r_i = a \cdot vege_i + b \cdot solo_i + c \cdot sombra_i + e_i \quad (1)$$

onde,

r_i , é a resposta do pixel na banda i ;

a, b, c , são respectivamente as proporções de vegetação, solo e sombra;

$vege_i, solo_i, sombra_i$, são respectivamente as respostas espectrais das componentes vegetação, solo e sombra;

e_i , é o erro na banda espectral i .

Dos procedimentos usados na estimação das proporções das componentes de um dado pixel, o mais aceito atualmente utiliza o método dos mínimos quadrados para a solução de um sistema de equações lineares. A descrição do método dos mínimos quadrados para o modelo linear de mistura espectral é detalhada em Shimabukuro e Smith (1991).

2.3.2 Segmentação de Imagens

O processo de segmentação consiste em subdividir uma imagem em regiões homogêneas considerando alguns de seus atributos, como por exemplo, o nível de cinza dos pixels e a textura, visando caracterizar a representatividade dos objetos da cena (Bins et al., 1996). A segmentação pressupõe a geração de objetos internamente homogêneos e estatisticamente distintos de seus vizinhos, que serão gerados segundo critérios de descontinuidade ou de similaridade.

O algoritmo de segmentação proposto por Bins et al. (1996), é baseado na tradicional técnica de crescimento de regiões que consiste em um processo de iteratividade no qual regiões são agrupadas a partir de pixels individuais. Tal implementação propõe um agrupamento de regiões baseado no conceito de região vizinha mais similar. Neste método são definidos dois limiares, o primeiro é o de similaridade, abaixo do qual duas regiões são consideradas similares e o segundo é o de área, valor de área mínima para que uma região seja individualizada.

2.3.3 Análise de Dados de Área

No estudo de dados agregados por área, a análise da autocorrelação espacial visa identificar a estrutura de correlação espacial que melhor descreva os dados, sendo que a magnitude da autocorrelação pode ser estimada, por exemplo, através do índice global e local de Moran (Câmara et al., 2003). Estas estatísticas dependem da variabilidade espacial dos dados de área, estimada a partir da matriz de proximidade espacial. Neste sentido, a autocorrelação espacial mede o quanto o valor observado de um atributo em uma região é independente dos valores desta mesma variável nas localizações vizinhas.

Para Bailey e Gatrell (1995), a definição da matriz de proximidade leva em consideração as medidas de proximidade espacial entre áreas adjacentes, sendo que estas medidas podem ser a distância ao centróide da área, as adjacências ou o tamanho das fronteiras de área para uma dada configuração espacial.

No contexto da análise exploratória dos dados, a caracterização da dependência espacial é um aspecto

fundamental, pois mostra como os valores estão correlacionados no espaço (Câmara et al., 2003). Assim, pode-se usar como função de estimação de uma variável o Índice Global de Moran, expresso pela seguinte fórmula:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (z_i - \bar{z}) \cdot (z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2} \quad (2)$$

onde,

n , é o número de áreas;

z_i , é o valor do atributo considerado na área i ;

\bar{z} , é o valor médio do atributo na região de estudo;

w_{ij} , são os elementos da matriz normalizada de proximidade espacial.

Desta forma, o índice de Moran pode variar de -1 a +1, sendo que valores positivos indicam uma correlação direta e valores negativos indicam uma correlação inversa. Todavia, é importante que se estabeleça uma validade estatística para este índice, sendo que a estimação da significância de tal índice é feito, comumente, através do teste da pseudo-significância.

A visualização da dependência espacial pode ser materializada pelo diagrama de espalhamento de Moran, construído com base nos valores dos atributos normalizados. Assim, constrói-se um gráfico bidimensional de z (valores normalizados) por wz (média dos vizinhos), que indica pontos de associação espacial positiva e pontos de associação espacial negativa, de acordo com a dispersão dos pontos no gráfico.

Para o cálculo do índice global de Moran, assume-se a hipótese de estacionariedade de primeira e segunda ordem, desta forma, quando o dado é não estacionário a função de autocorrelação continua decaindo mesmo após ultrapassar a distância onde há influências locais, neste caso são indicados outros índices, como o índice global de Geary, expresso pela fórmula:

$$c = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (z_i - z_j)^2}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n z_i^2} \quad (3)$$

onde,

n , é o número de áreas;

z_i , é o valor do atributo considerado na área i ;

z_j , é o valor do atributo considerado na área j ;

w_{ij} , são os elementos da matriz normalizada de proximidade espacial.

3 MATERIAL E MÉTODOS

Conforme proposto em INPE (2004), as imagens utilizadas neste estudo correspondem a uma área representativa da floresta Amazônica, da época de seca (correspondente geralmente aos meses de julho, agosto e setembro), com cobertura mínima de nuvens e com boa qualidade radiométrica. Tais imagens (Figura 1) correspondem a uma cena CCD/CBERS-2 (167/113) de 09/08/2004 e parte de uma cena TM/LANDSAT-5 (227/68) de 30/07/2004.

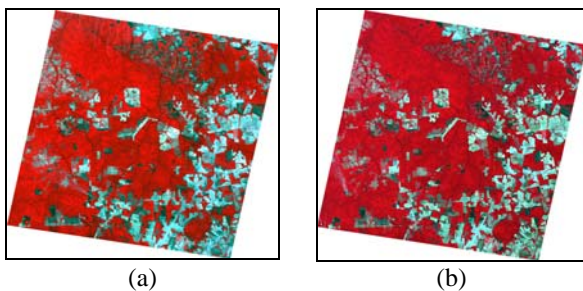


Figura 1. Composição Colorida 4(R)3(G)2(B), (a) imagem CCD 167/113 de 09/08/2004 e (b) imagem TM 227/68 de 30/07/2004.

Tendo como base a metodologia apresentada por Mello et al. (2003), este trabalho foi desenvolvido de acordo com a seguinte seqüência:

a. Seleção do Material: seleção das imagens do ano de interesse com cobertura de nuvens mínima;

b. Georreferenciamento das Imagens: segundo Crósta (1993), as imagens geradas por sensores remotos são sujeitas a uma série de distorções espaciais, não possuindo, portanto precisão cartográfica quanto ao posicionamento dos objetos, superfícies ou fenômenos nelas representados. Desta forma, neste trabalho, as imagens foram georreferenciadas utilizando-se pontos de controle que foram usados no ajuste de um polinômio de mapeamento de 1º grau. Após isso, fez-se uma análise da qualidade do georreferenciamento através dos pontos de teste que deram uma idéia não tendenciosa da precisão desta etapa.

c. Modelo Linear de Mistura: neste trabalho foram geradas as imagens fração solo, sombra e vegetação,

sendo que apenas as duas primeiras foram utilizadas nas etapas subsequentes.

d. Reamostragem: a imagem georreferenciada foi reamostrada para a resolução espacial de 60 x 60 m.

e. Segmentação das Imagens Fração-Sombra e Solo: a imagem fração solo foi segmentada pelo método de crescimento de regiões, utilizando limiares variados de similaridade e área.

g. Avaliação da Segmentação: foram calculados índices de autocorrelação espacial, como os índices globais de Geary e Moran.

4 RESULTADOS

A metodologia deste trabalho contempla a avaliação da segmentação das imagens fração-solo e sombra do sensor CCD/CBERS-2 (bandas 234) e TM/LANDSAT-5 (bandas 345). Para cada uma das respectivas imagens segmentadas foram criados polígonos a partir dos segmentos, sendo que, para cada um dos polígonos foram extraídos os atributos de média dos níveis de cinza dos pixels.

A primeira etapa da metodologia avaliou a distinção entre a vizinhança dos segmentos através do cálculo do índice global de autocorrelação espacial. A presença da tendência (não-estacionariedade de primeira ordem) dos dados foi visualizada através de gráficos dos atributos para os eixos leste-oeste e norte-sul, sendo que, para a construção destes gráficos foram usadas as coordenadas geográficas dos centróides dos polígonos.

Para cada uma das segmentações, com limiares de similaridade e área de 8 e 16, respectivamente, foram calculados e comparados os índices globais de Geary e de Moran (Tabela 2). Sabendo-se que a segmentação é tanto melhor quanto menor for o índice global de Moran, verificou-se, pela análise da tabela 2, que o melhor resultado foi obtido com a imagem fração-solo do sensor TM/LANDSAT-5 e o pior com a imagem fração-solo do sensor CCD/CBERS-2.

Imagem	MORAN	GEARY
C234_SOLO 5361 polígonos	0,416	0,771
C234_SOMBRA 5118 polígono	0,296	0,516
L345_SOLO 3614 polígono	0,177	1,061
L345_SOMBRA 2857 polígono	0,211	0,562

Tabela 2. Índices de Moran e Geary.

Foram geradas ainda mais duas segmentações para a imagem fração-solo do sensor CCD/CBERS-2, com valores de similaridade e de área respectivamente igual a 4 e 16, para a primeira e 16 e 16 para a segunda. Para essas segmentações foram calculados e comparados os índices globais de Geary e de Moran (Tabela 3). Comparativamente, pela análise da tabela 3, foi possível verificar que a segmentação com limiar de similaridade igual a 16, da imagem fração-solo do sensor CCD/CBERS-2 foi a que obteve o melhor resultado.

Imagem	MORAN	GEARY
C234_SOLO_4 10695 polígonos	0,511	0,564
C234_SOLO_8 5361 polígono	0,416	0,771
C234_SOLO_16 1875 polígono	0,207	0,979

Tabela 3. Índices de Moran e Geary.

Para a etapa de avaliação da homogeneidade interna dos segmentos foram escolhidos, aleatoriamente, cinco polígonos da imagem segmentada originada a partir da imagem fração-solo das bandas 234 do sensor CCD/CBERS-2 (Figura 02). Em cada um dos polígonos foram criadas grades regulares para o cálculo os indicadores globais de autocorrelação espacial (Tabela 04). Com base nos resultados foi possível verificar que dentre os polígonos avaliados, o mais homogêneo é o 04.

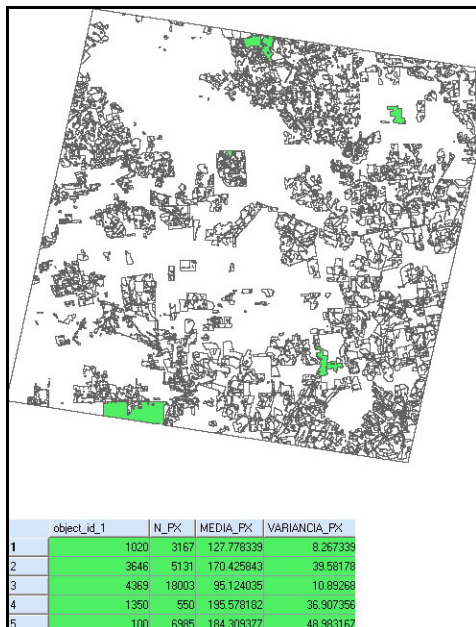


Figura 2. Representação dos polígonos com a tabela de atributos.

Polígono	Moran
POL_01 750 pontos	0,645
POL_02 160 pontos	0,223
POL_03 1265 pontos	0,392
POL_04 4512 pontos	0.843
POL_05 1762 pontos	0,183

Tabela 4. Índices de Moran para a homogeneidade interna.

5 CONCLUSÕES

A aplicação dos indicadores de autocorrelação espacial na avaliação da qualidade da segmentação foi realizada com sucesso para as imagens frações dos sensores CCD/CBERS-2 e TM/LANDSAT-5. Desta forma, os resultados encontrados corroboram a hipótese de que os índices de autocorrelação espacial são úteis à avaliação da qualidade dos resultados da segmentação.

O emprego de tais índices torna-se uma ferramenta prática e eficiente que permite a avaliação dos resultados da segmentação, uma vez que permitem aos usuários de SIG a opção por um parâmetro quantitativo e objetivo de avaliação dos resultados.

6 REFERÊNCIAS

- Alves, Diógenes S. O Processo de Desmatamento na Amazônia. In: **Parcerias Estratégicas**. n° 12. Setembro, 2001. p. 259-275.
- Bailey, T.; Gatrell, A. **Interactive Spatial Data Analysis**. London, Longman Scientific and Technical, 1995.
- Becker, Bertha K. **Amazônia**. 5 ed. São Paulo: Editora Ática, 1997. 112 p.
- Bins, L. S.; Fonseca, L. G.; Erthal, G. Satellite Imagery Segmentation: a region growing approach. In: Simpósio Brasileiro de Sensoriamento Remoto, 8., Salvador, abr. 1996. **Anais**. São José dos Campos: INPE, 1996. Disponível na biblioteca digital URLib: <sid.inpe.br/deise/1999/02.05.09.30>. Acesso em: 02 nov. 2004.
- Bins, L. S.; Fonseca, L. G.; Erthal, G. Um Método de Classificação Não-Supervisionada por Regiões. In: Proceeding of the Sixth Brazilian Symposium on Graphic

Computation and Image Processing, Recife. 1993. p 65-68.

Câmara, G.; Fucks, S.; Carvalho, M.; Monteiro, A. **Análise Espacial de Dados Geográficos**. São José dos Campos. INPE, 2003 (3a. edição, revista e ampliada). Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise>>. Acesso em: 15 nov. 2004.

Crosta, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto**. Campinas: IG/Unicamp, 1993.

Duarte, V.; Shimabukuro, Y.; Aulicino, L. Metodologia para Padronizar e Atualizar o Banco de Dados do Projeto "Prodes Digital". In: Simpósio Brasileiro de Sensoriamento Remoto, 11., 2003, Belo Horizonte. **Anais**. São José dos Campos: INPE, 2003. p. 2705-2712. Disponível na biblioteca digital URLib: <tid.inpe.br/sbsr/2002/19.314>. Acesso em: 30 out. 2004.

Escada, M. Isabel B. **Evolução de Padrões da Terra na Região Centro-Norte de Rondônia**. 2003. 157 p. (INPE – 10209-TDI/899). Tese (Doutorado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2003. Disponível na biblioteca digital URLib: <sid.inpe.br/jeferson/2003/06.30.13.31>. Acesso em: 17 jul. 2004.

Instituto Nacional de Pesquisas Espaciais (INPE). **CBERS**. São José dos Campos. 2004a. Disponível em: <<http://www.cbers.inpe.br>>. Acesso em: 09 nov. 2004.

Instituto Nacional de Pesquisas Espaciais (INPE). **PRODES Digital**. São José dos Campos. 2004b. Disponível em: <<http://www.obt.inpe.br/prodesdigital/metodologia.html>>. Acesso em: 10 set. 2004.

Instituto Nacional de Pesquisas Espaciais (INPE). **SPRING**. São José dos Campos. 2004c. Disponível em: <<http://www.dpi.inpe.br/spring>>. Acesso em: 01 set. 2004.

Mello, E.; Moreira, J.; Santos, J.; Shimabukuro, Y.; Duarte, V.; Souza, I.; Barbosa, C.; Cartaxo, R.; Paiva, J. Técnicas de Modelo da Mistura Espectral, Segmentação e Classificação de Imagens TM-LANDSAT para o Mapeamento do Desflorestamento da Amazônia. In: Simpósio Brasileiro de Sensoriamento Remoto, 11., 2003, Belo Horizonte. **Anais**...São José dos Campos: INPE, 2003. p. 2705-2712. Disponível na biblioteca digital URLib: <tid.inpe.br/sbsr/2002/19.314>. Acesso em: 30 out. 2004.

Shimabukuro, Y. E.; Smith, J. A. The Least-Squares Mixing Models to Generate Fraction Images Derived from Remote Sensing Multispectral Data. In: **IEEE Transactions on Geoscience and Remote Sensing**, vol. 29, p. 16-20, 1991.

The United States Geological Survey (USGS). **Landsat Project**. Disponível em: <<http://landsat.usgs.gov>>. Acesso em: 14 out. 2004.