

Investigação do processo de segmentação multiresolução utilizando o critério de ponderação de formas e cores aplicadas às imagens de áreas urbanas de alta resolução espacial do satélite Ikonos

Jaqueline Vicente Matsuoka¹
Vitor Haertel¹

¹ Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia
Universidade Federal do Rio Grande do Sul
Av. Bento Gonçalves, 9500 - Campus do Vale
Caixa Postal 15044
91501-970 - Porto Alegre - RS - Brasil
Fone: (051) 3316-6221 - Fax: (051) 3316-6657
jaquematsuoka@gmail.com
victor.haertel@ufrgs.br

Abstract: The more recent developments in sensor technology have made available to the international community a new generation of high spatial resolution image data. The analysis of these data requires, however, the development of new methodologies, particularly with regard to the classification procedure. An approach that has shown promising results consists in performing a previous segmentation of the image followed by a classification process that takes the image segments as the patterns to be labeled rather than individual pixels. In this study we investigate the image segmentation methodology as implemented in the *e-Cognition* software package. We start by combining the 1 m spatial resolution panchromatic data acquired by the IKONOS II sensor system with the 4 m spatial resolution multispectral data acquired by the same system. The combined image merges the thematic information captured by the multispectral bands with the higher spatial resolution provided by the panchromatic band. Segmentation procedures were then applied to this combined image, using the criteria implemented in the *eCognition* software, showing promising results.

Palavras-chave: remote sensing, image processing, urban areas, multiresolution segmentation, sensoriamento remoto, processamento de imagens, áreas urbanas, segmentação multiresolução.

1. Introdução

Segmentação é a subdivisão de uma imagem em regiões homogêneas, conforme um determinado critério. Durante muitos anos a segmentação de imagens foi alvo de muitas pesquisas na área de análise de imagens. Poucas abordagens conseguem obter resultados qualitativamente convincentes, robustos e aplicáveis as configurações operacionais. Uma razão é que a segmentação de uma imagem em um dado número de regiões é um problema que admite um número muito grande de soluções possíveis. O alto número de graus de liberdade deve ser reduzido para uma ou poucas soluções que satisfaçam dados requerimentos. Outra razão é que algumas regiões de interesse são extremamente heterogêneas, surgem então, ambigüidades e a necessidade de distinguir informações não são diretamente avaliadas. Tais requerimentos dizem respeito à qualidade, desempenho – tamanho do conjunto de dados e tempo de processamento.

Existem muitos métodos para obter objetos da imagem, cada um com suas vantagens e desvantagens. Alguns são totalmente automáticos e outros são apenas semi-automáticos. As recentes pesquisas dividem os métodos de segmentação de imagens em dois grupos: métodos de conhecimento dirigido (*top-down*) e métodos de dados dirigidos (*bottom-up*). A diferença básica entre ambas as abordagens é: o método *top-down* usualmente leva a resultados locais porque ele somente identifica *pixels* ou *regiões* que a descrição do modelo

encontra, e os métodos *bottom-up* fazem a segmentação da imagem completa. Eles agrupam *pixels* para agrupamentos espaciais com certos critérios de homogeneidade e heterogeneidade.

O procedimento básico do software *eCognition* é segmentação multiresolução pelo grande conhecimento livre da extração de primitivas de objetos da imagem. Isto foi desenvolvido para produzir objetos de imagens de diferentes resoluções e alta qualidade. A segmentação multiresolução é essencialmente um procedimento de otimização heurística, que minimiza localmente a média da heterogeneidade dos objetos da imagem para uma dada resolução sobre a cena inteira.

Este trabalho mostra e analisa como é realizado o processo de segmentação multiresolução no software *eCognition* e quais são as suas vantagens, para isso foram realizados diferentes testes com diferentes imagens de alta resolução do satélite IKONOS, obtidas através de fusões das bandas PAN e multispectrais.

2. Metodologia

2.1. Imagens Ikonos

Com 1 m de resolução no modo PAN e PSM os dados imagem coletados pelo sistema sensor a bordo do satélite IKONOS oferece a mais fina resolução espacial entre as imagens orbitais atualmente disponíveis. Ele permite discriminar objetos de 1 m² de área ou maior.

Uma interessante utilização destes dados consiste na combinação de imagens adquiridas no modo PANCROMÁTICO, P&B, com 1 m de resolução, com imagens multispectrais coloridas de 4 m de resolução, para a geração de imagens coloridas com 1m de resolução, combinando então as vantagens dos dois tipos de imagens. Como este sistema sensor adquire sistematicamente as imagens no modo PAN e MS para todas as áreas, esta fusão e o produto PSM podem ser gerados para todas as imagens adquiridas pelo IKONOS.

As imagens de alta resolução espacial revolucionaram o modo de utilizar as imagens de satélites, pois se anteriormente um *pixel* continha vários objetos, agora, um objeto é composto de vários *pixels*. Os algoritmos de interpretação mudam os níveis de detalhamento também.

São inúmeras as aplicações potenciais que vão se tornar reais utilizando as imagens de alta resolução espacial, em quase todos os ramos que já usam imagens de satélites, e muitos outros que sofriam da limitação da resolução, desde a gestão do espaço geográfico complexo e extenso até a análise pontual.

2.2. Segmentação Multiresolução de Objetos das Imagens

A segmentação é utilizada em análise de imagens para extração de objetos de interesse. Porém, em muitos casos as cenas que contém estes objetos podem ser bastante heterogêneas. Quando um procedimento de segmentação segue um critério geral de homogeneidade, isto é, sem levar em consideração o grau de heterogeneidade da cena, torna-se mais difícil a extração de objetos de interesse. O procedimento de segmentação do *eCognition* utiliza a abordagem orientada à objetos, a qual visa encontrar primitivas de objetos para servir como informações adicionais para outras segmentações ou futuras classificações.

Para a finalidade de segmentação multiresolução foi desenvolvida um novo procedimento, o qual permite a extração de grande parte do conhecimento das primitivas de objetos de imagens homogêneas em algumas resoluções escolhidas, especialmente levando em consideração o contraste do local. Pode ser aplicado a um grande número de tipos de dados; trabalha em um número arbitrário de canais simultaneamente e com altas variações de texturas ou baixo contraste como imagens de radar e VHR.

Com a finalidade de receber material sem nenhum pré-processamento para a análise de imagens orientada a objetos alguns critérios foram seguidos para o desenvolvimento da segmentação multiresolução:

- Um processo de segmentação deve produzir segmentos altamente homogêneos para a separação ótima e a representação de regiões da imagem;
- Como cada problema de análise de imagem trata as estruturas de uma determinada escala espacial baseada em dados específicos, o tamanho médio dos objetos da imagem deve ser adaptado à escala de interesse;
- Quase todos os atributos de objetos das imagens (tom, textura, forma e relações de regiões adjacentes), são mais ou menos dependentes da escala. Assim, somente estruturas de escalas semelhantes são de qualidade comparável e tem atributos comparáveis. Por isso, os objetos da imagem resultantes devem ser aproximadamente da mesma magnitude;
- O procedimento de segmentação deve ser aplicável a um número grande de diferentes tipos de dados;
- O resultado da segmentação deve ser reproduzível;
- O processo de segmentação deve ser computacionalmente eficiente, mesmo para grandes conjuntos de dados, como ocorre em imagens de satélite.

Para que as técnicas de segmentação sejam satisfatórias, é necessário que elas reproduzam a análise feita pelo olho humano, do contrário nenhum processo de segmentação será convincente o bastante. Para atingir este objetivo, é preciso realizar a manipulação do contraste local da imagem, assim como a segmentação da imagem em regiões com dimensões aproximadamente similares, produzindo então objetos da imagem com a mínima suavização da borda, mesmo em imagens com homogeneidade espectral e com alto grau de textura.

A maioria do processo faz a segmentação levando em conta as características globais do espaço. Porém, as características globais, dependem de um subconjunto particular da cena. A mesma área da imagem que é parte de subconjuntos diferentes será segmentada diferentemente por meio de tais procedimentos, sempre dependendo das características espaciais do respectivo subconjunto.

Para evolução de uma técnica de segmentação é necessário, ainda, definir medidas e comparabilidade dos resultados do processo de segmentação. Dada a definição de heterogeneidade e um tamanho médio para os objetos da imagem, o critério utilizado para o desenvolvimento da segmentação multiresolução e o de que a heterogeneidade média dos pixels deve ser minimizada.

2.2.2. Cálculo do Critério de Heterogeneidade

O processo de segmentação multiresolução é uma técnica de fusão de regiões *bottom-up*. A formação de um objeto é feita a partir de um pixel. Através de etapas subsequentes, pequenas regiões da imagem são fundidas em uma grande região. Durante todo o processo de agrupamento, o procedimento de otimização minimiza a heterogeneidade ponderada nh resultando em objetos da imagem, onde n é o tamanho do segmento e h é uma definição de heterogeneidade arbitrária. O objeto é incrementado até que o menor incremento exceda o limiar definido pelo parâmetro de escala, quando isso ocorre o processo é interrompido. Assim, a segmentação é um procedimento de otimização local.

O *eCognition* usa a soma dos desvios padrões dos valores espectrais em cada banda (*layer*), as quais são ponderadas de acordo com a importância para a segmentação, para descrever a heterogeneidade espectral e da cor:

$$h = \sum_c w_c \cdot \sigma_c \quad (01)$$

onde w_c é o peso atribuído a cada banda.

Em alguns casos o uso da minimização de heterogeneidade espectral leva a segmentos ramificados ou objetos com bordas quebradas, o efeito pode ser mais forte em dados altamente texturados. Por isso, em muitos casos a mistura dos critérios de heterogeneidade espectral e espacial é usada a fim de reduzir o desvio de formas suavizadas e compactadas. Heterogeneidade de uma forma compacta é dada pela fórmula abaixo:

$$h = \frac{l}{\sqrt{n}} \quad (02)$$

onde l é o fator de comprimento da borda e n é o número de pixels que forma o objeto.

Uma outra forma de descrever a heterogeneidade da forma:

$$h = \frac{l}{b} \quad (03)$$

onde b é o menor comprimento da borda possível dado por um retângulo limite do objeto.

A aplicação desta definição de heterogeneidade otimiza a suavização das formas dos objetos resultantes. A mistura dos critérios de forma e espectral resulta em objetos mais compactos e mostram bordas mais suavizadas. Embora não seja tão homogêneo com respeito à cor, eles são mais satisfatórios ao olho humano.

A fim de encontrar o melhor resultado para a segmentação, o usuário pode definir alguns parâmetros, tais como escala, um único *layer* a ser utilizado ou mesmo uma mistura dos *layers* e o peso para cada um deles. A comparação dos valores de uma fusão com o parâmetro de escala define o critério de parada do algoritmo. Como mencionado anteriormente, o parâmetro de escala é a medida para a variação máxima em heterogeneidade que pode ocorrer na fusão de dois objetos. Internamente, esse valor é elevado ao quadrado e serve como limiar que faz com que o algoritmo seja encerrado.

O critério de heterogeneidade é dividido em duas partes:

1. *Critério para tonalidade (cores, espectral)*: é a variação na heterogeneidade que ocorre quando a fusão de dois objetos na imagem é descrita pelo peso dado os desvios padrões dos valores espectrais (bandas, *layers*);

2. *Critério da forma*: é o valor encontrado pelo melhor modelo, escolhido entre dois, para descrever as formas idealmente.

Para encontrar a heterogeneidade global f , são levados em conta os critérios de cor e de forma. Primeiramente é calculada a heterogeneidade da cor, por (04):

$$h_{color} = \sum_c w_c (n_{merge} \cdot \sigma_{merge}^c - (n_{obj1} \cdot \sigma_c^{obj1} + n_{obj2} \cdot \sigma_c^{obj2})) \quad (04)$$

e então a heterogeneidade da forma (07), a qual consiste de dois sub-critérios, compactidade (05) e suavidade(06):

$$h_{cmpct} = n_{merge} \cdot \frac{l_{merge}}{\sqrt{n_{merge}}} - (n_{obj1} \cdot \frac{l_{obj1}}{\sqrt{n_{obj1}}} + n_{obj2} \cdot \frac{l_{obj2}}{\sqrt{n_{obj2}}}) \quad (05)$$

$$h_{smooth} = n_{merge} \cdot \frac{l_{merge}}{b_{merge}} - \left(n_{obj1} \cdot \frac{l_{obj1}}{\sqrt{n_{obj1}}} + n_{obj2} \cdot \frac{l_{obj2}}{\sqrt{n_{obj2}}} \right) \quad (06)$$

$$h_{shape} = w_{cmpct} \cdot h_{cmpct} + (1 - w_{cmpct}) \cdot h_{smooth} \quad (07)$$

onde n é o tamanho do objeto, l é o perímetro e b é o perímetro do retângulo limite do objeto.

O critério de *suavidade* é usado para otimizar objetos com bordas suaves. Para citar um exemplo, esse critério deve ser usado quando se trabalha com dados muito heterogêneos, como imagens de radar. Já o critério de *compacidade* é usado para otimizar objetos compactos. Ou seja, deve ser usado quando em objetos de imagem um tanto quanto compactos, mas separados de objetos não compactos por um fraco contraste. Esse critério pode ser utilizado para áreas urbanas, onde a resposta espectral é parecida, porém com formas bem diferentes.

E finalmente é calculado f , pela equação (08):

$$f = w \cdot h_{color} + (1 - w) \cdot h_{shape} \quad (09)$$

onde w é o peso dado para cor (contrário a forma) com $0 \leq w \leq 1$.

2.2.3. Vizinhança de pixels plana e diagonal

O *eCognition* faz a distinção entre dois tipos de definição de vizinhança de pixels: vizinhança plana de 4 e vizinhança diagonal de 8, como mostram as **figuras 1 e 2** respectivamente.

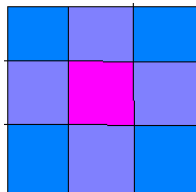


Figura 1 – Vizinhança plana de 4.

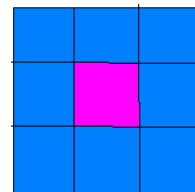


Figura 2 – Vizinhança diagonal de 8.

3. Realização dos testes e análise dos resultados

Para realização dos testes foram utilizadas imagens do satélite IKONOS de alta resolução fusionadas. Na fusão utilizou-se as bandas 1, 2 e 3 com resolução de 4 metros e a pancromática com resolução de 1m para aumentar a acurácia espacial, portanto as imagens resultantes da fusão possuem resolução de 1 metro.

Foram realizados alguns testes com diferentes imagens e diferentes parâmetros, como serão mostrados a seguir. Cada imagem tem sua particularidade, algumas tem formas bem definidas e cores diferentes o que facilita o processo.

No primeiro teste foi utilizada a imagem mostrada na **figura 3**. Pode-se notar que a imagem possui algumas formas bem definidas, porém a resposta espectral (cor) pode gerar alguma confusão no processo de segmentação.

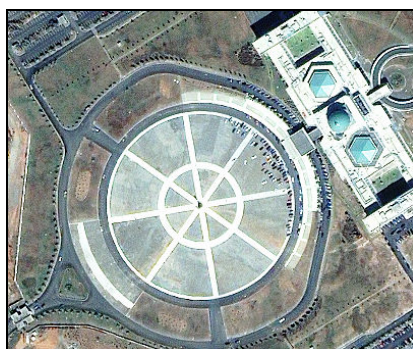


Figura 3 – Imagem de Abuja – Nigéria

Utilizando a configuração *default* do *eCognition*, como mostra a **tabela 1**, tem-se uma segmentação ruim. Isso ocorre porque o parâmetro de escala utilizado é muito pequeno para os objetos existentes na imagem. E também porque o valor dado à forma deveria ter sido maior que o valor dado a cor, devido a grande semelhança das respostas espectrais dos objetos. Finalmente o critério de compacidade e suavidade estão corretos, estes não estão totalmente incorretos, pois o maior peso foi dado a suavidade, ou seja, os objetos que possuem similaridade espectral e não são tão compactos são identificados, porém isso pode causar a segmentação de áreas que na realidade compõem uma só. Observando a imagem na **figura 4**, pode-se observar que foram criados objetos muito pequenos, assim os polígonos são mostrados um a um individualmente.

Tabela 1 – Parâmetros *default* do *eCognition*

Parâmetros	
Escala	10
Cor	0.8
Forma	0.2
Suavidade	0.9
Compacidade	0.1

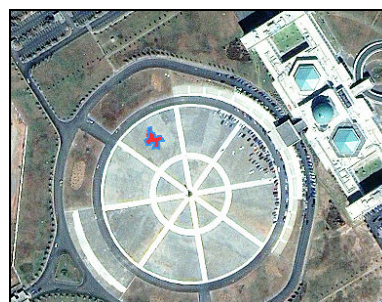


Figura 4 – Teste 1: Imagem segmentada 1 com parâmetros incorretos

Alterando os valores para os parâmetros como mostra a **tabela 2**, obtém-se a imagem segmentada ilustrada na **figura 5**. Utilizando estes valores na segmentação geraram-se objetos maiores, assim como pequenas variações espectrais não foram segmentadas, isso ocorreu porque a forma dos objetos teve um peso maior no processo de segmentação. Ambos os testes utilizaram a vizinhança de quatro, ou seja, somente os pixels das vizinhanças planas (esquerda, direita, acima e abaixo) foram considerados como vizinhos. Pode-se observar que o resultado obtido na segunda segmentação foi consideravelmente melhor que na primeira.

Tabela 2 – Parâmetros utilizados no processo de segmentação 2

Parâmetros	
Escala	40
Cor	0.4
Forma	0.6
Suavidade	0.5
Compacidade	0.5

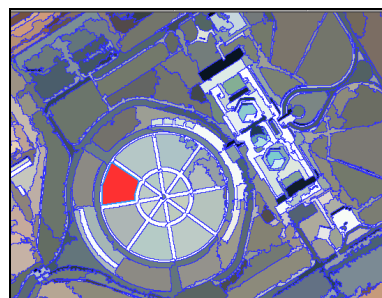


Figura 5 – teste1: Imagem segmentada 2 com parâmetros corretos e com polígonos.

A seguir, são analisados os resultados obtidos pela segmentação feita com os diferentes tipos de vizinhança. Na vizinhança de 4 obtém-se uma segmentação boa, porém as feições curvilíneas e diagonais na maioria das vezes são perdidas parcialmente. Já na vizinhança de 8, podemos observar que tais feições são detectadas com boa exatidão.

Comparando as imagens das **figuras 6 e 7**, pode-se notar nas regiões marcadas por círculos vermelhos, que algumas feições detectadas na imagem da **figura 6**, com vizinhança de 8, foram perdidas na imagem da **figura 7**, com vizinhança de 4.

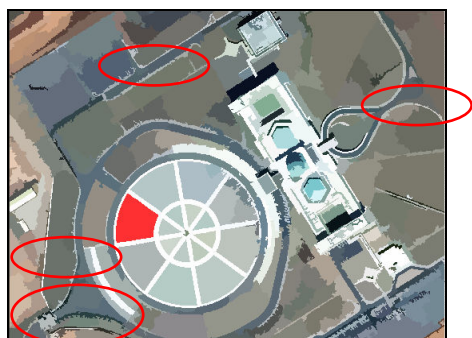


Figura 6 – Teste 1: Imagem segmentada 3 com parâmetros corretos, vizinhança de 8.

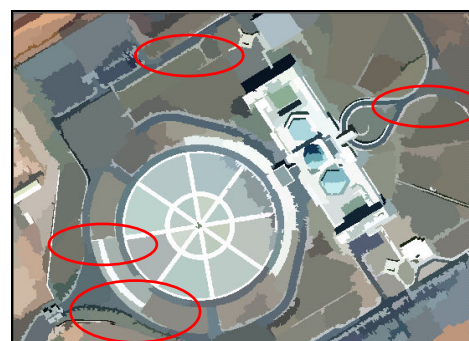


Figura 7 – Teste1: Imagem segmentada 4 com parâmetros corretos, vizinhança de 4.

Em imagens bastante heterogêneas como a mostrada na **figura 8**, o *software* também apresenta um bom desempenho, desde que os parâmetros exigidos sejam atribuídos corretamente, de acordo com a metodologia descrita anteriormente.

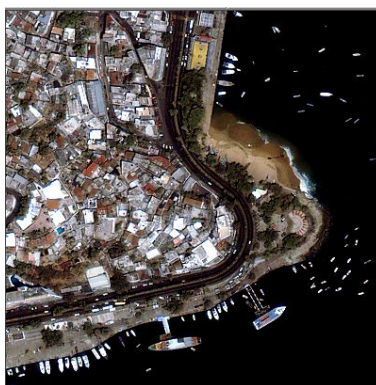


Figura 8 – Imagem de Acapulco – México

Ao realizar-se a segmentação multiresolução na imagem da **figura 8**, com os parâmetros mostrados na tabela 3, obtém-se um excelente resultado, até mesmo os menores objetos, como os barquinhos atracados no porto, são reconhecidos, como se pode observar pelas **figuras 9 e 10**.

Tabela 3 – Parâmetros utilizados na segmentação 1 e 2 do teste 2

Parâmetros	
Escala	35
Cor	0.4
Forma	0.6
Suavidade	0.8
Compacidade	0.2



Figura 9 – Teste 2: Segmentação 1 (Vizinhança de 4)



Figura 10 – Teste2: Segmentação 2 (Vizinhança de 8)

4. Considerações Finais e Conclusões

Atualmente a evolução das imagens multiespectrais exige que as aplicações de processamento de imagens sejam cada vez mais dinâmicas. À medida que as características de resolução espectral, espacial e radiométrica das imagens evoluem, os *softwares* de processamento digital e interpretação devem evoluir com elas, para que todas as informações possam ser retiradas e analisadas.

A utilização do critério de homogeneidade levando em consideração somente às características espectrais das imagens deixa muito a desejar, pois com o surgimento das imagens de alta resolução atributos espaciais como forma e textura adquiriram crescente importância no processo de interpretação. Assim, um *software* com as características do *eCognition*, que utiliza ambos os critérios e ainda o grau de heterogeneidade e as bandas que estão sendo utilizadas no processo os resultados da segmentação se tornam consideravelmente mais precisos e confiáveis.

Após a realização de alguns testes utilizando o *software eCognition* e imagens multiespectrais fusioandas nas bandas 1,2,3 e pan do satélite IKONOS, pode-se concluir que:

- O processo de segmentação utilizando a mistura dos critérios de cor (espectral) e forma (compacidade e suavidade) consegue reconhecer um grande número de objetos, mesmo em imagens heterogêneas;
- A segmentação multiresolução permite que imagens extremamente heterogêneas e com baixo contraste sejam segmentadas, obtendo bons resultados;
- Até mesmo objetos muito pequenos existentes na imagem, conseguem ser identificados, devido ao privilégio que o usuário possui de poder fornecer o parâmetro de escala;
- As feições curvilíneas, como estradas e rodovias, são melhores identificáveis devido a utilização do tipo de vizinhança 8;

Os resultados obtidos com a realização dos testes utilizando as imagens de alta resolução do IKONOS foram excelentes.

Bibliografia

- GONZALEZ, R. C.; WOODS, R. E. (2000). *Digital Image Processing*. Reading. Addison-Wesley, 509 p.
- SONKA, M. et al. (1999). *Image Processing, Analysis, and Machine Vision*. Pacific Grove, ITP, 770 p.
- GOMES, J.; VELHO, L. (2002). *Computação Gráfica: Imagem*. Rio de Janeiro, IMPA / SBM.