

Classificação Orientada ao Objeto com o Uso de Imagens Sintéticas

Rodrigo Pereira Lersch
Benamy Turkienicz

Universidade Federal do Rio Grande do Sul – UFRGS
Rua Sarmento Leite, 320 salas 306/307 – POA/RS (Cep:90020-150)
rodrigolersch@gmail.com; benamy.turkienicz@gmail.com

Abstract. Orbital images with high spatial resolution started to be used fairly recently, at the beginning of this century. New techniques had to be developed as to adjust required data for digital classification procedures. The academia is engaged in the development of these techniques since orbital images were available whereas municipal administrations, in general, did not make significant progress. This lack of knowledge has generated sheer difficulties to some administrations to specify needed images for further analytical work. One example of this problem is the purchase of synthetic merged images with no awareness that sensors work with specific bands representing ranges of the electromagnetic spectrum, an essential component of the process of pattern recognition. This study investigate the possibility of using synthetic images in the digital classification processes. Processes of object-oriented classification were compared to the pixel by pixel process classification, utilizing transformed images. Methods of post-classification based on the filtering of the product of classification are proposed (GIS procedure). All methods investigated were quantified in order to estimate the viability of using synthetic images of Quick-Bird sensor for identifying features of the urban environment. These procedures aim to contribute to urban studies, especially in the automated surveying of the urban morphology.

Palavras-chave: SIG, Sensoriamento Remoto, Classificação de Imagens Orientada ao Objeto, Urbanismo, Levantamento Urbano.

1. Introdução

O uso de Sensoriamento Remoto para identificação e caracterização de fenômenos urbanos tem se tornado mais eficaz nos últimos anos. Isto se deve ao surgimento de sensores com resolução espacial cada vez maior, como: Ikonos e Quick Bird. Durante as décadas de 1980 e 1990 já se dispunha de sensores com resolução espacial média: 10 a 30 m, como LandSat, CBERS ou SPOT. No entanto esta resolução ainda era inviável para identificação da maioria dos elementos presentes nas áreas urbanas, tais como: edificações, praças ou lotes vazios.

Junto com o grande avanço dos sistemas sensores surge um novo problema. As técnicas tradicionais de classificação pixel a pixel já se tornam pouco acuradas para classificar imagens de alta resolução espacial, nas quais um mesmo alvo passa a ser representado por uma grande variação de valores dos contadores digitais. Isto se deve a alta resolução espacial e radiométrica destas novas imagens. Para solucionar este problema uma das técnicas que vem tomando cada vez mais espaço no meio acadêmico e profissional são os processos de classificação orientados ao objeto. Na classificação orientada ao objeto o espaço é segmentado em regiões, as quais possuem um determinado padrão de comportamento espectral. Posteriormente estas regiões são então classificadas segundo padrões espectrais, padrões de forma e padrões de textura. A Textura neste caso é determinada por padrões de distribuição dos contadores digitais segundo um determinado modelo estatístico (entropia, similaridade, desvio padrão, dentre outros).

A maioria dos estudos urbanos realizadas atualmente recorre ao uso de Sistemas de Informações Geográficas (SIG) e Sensoriamento Remoto. Em Sensoriamento Remoto se constata, principalmente a utilização de imagens de alta resolução espacial. Este uso é por vezes feito de forma mais amadorística, utilizados imagens on-line, obtidas de *softwares* como o Google Earth, ou utilizando-se de imagens originais. Em Países onde existe a falta de

recursos, tanto técnicos como financeiros, o uso de imagens como as obtidas de forma on-line torna-se uma solução em relação à falta de dados mais adequados. Contudo este produto só é válido para utilização via classificação visual. No Brasil prefeituras de cidades pequenas não possuem recursos para adquirir imagens de alta resolução espacial, nestes casos o uso de imagens on-line pode contribuir como uma solução alternativa. Muitas vezes, porém, municípios de médio e grande porte contam com imagens dos sensores Ikonos ou Quick Bird. Predominantemente estas imagens são vendidas a estes municípios como imagens sintéticas em composições, que predominantemente, utilizam os três canais do espectro visível (RGB321). O presente estudo visa avaliar se tais imagens sintéticas podem ser uma fonte de dados viável para utilização dos processos de classificação digital. Ressalta-se que tais imagens possuem a resolução espacial maior, oriunda do canal pancromático e estão fusionadas as três bandas multiespectrais abrangendo o espectro do visível. A informação espectral original dos canais multiespectrais, encontra-se alterada. Soma-se a isto a inexistência do canal referente ao infravermelho.

Nesta pesquisa utilizou-se para teste uma imagem sintética do sensor Quick Bird na composição RGB321. Não é possível saber se a fusão foi feita utilizando-se uma transformação IHS (*Intensity-Hue-Saturation*) ou por Componentes Principais. É mais provável que seja esta última já que a imagem resultante conserva o formato de 16 bits. A imagem cobre boa parte do bairro Menino Deus na cidade de Porto Alegre, situado no Estado do Rio Grande do Sul. Aplicou-se sobre este produto o processo de classificação orientada ao objeto e comparou-se este resultado ao obtido com uma classificação pixel a pixel, via Máxima Verossimilhança Gaussiana. O objetivo é avaliar o potencial de reconhecer edificações contidas nesta região, utilizando os diferentes processos de classificação digital que posteriormente são filtradas por técnicas de SIG, segundo a metodologia proposta. Posteriormente foi utilizado um mapa contendo a verdade terrestre das edificações para aferir e quantificar os resultados obtidos em cada técnica e etapa.

2. Metodologia

Para realizar o processo de classificação foi necessário optar por um *software* que contasse com o processo de classificação orientada ao objeto, para os quais existem poucas opções. Assim o *software* deve contar com rotinas de segmentação e de classificação por regiões. Um dos *softwares* mais promissores para este tipo de classificação é o *software* comercial eCognition. Contudo como se visa testar o aproveitamento de imagens sintéticas, não faria muito sentido à utilização de um programa de alto custo, que seria inviável para a aquisição por prefeituras de municípios de médio, ou mesmo grande porte. Com base neste fator se optou pelo SPRING desenvolvido pelo Instituto Nacional de Pesquisas Espaciais (INPE). Embora o Spring não conte com o mesmo número de ferramentas enfocando aspectos como textura e forma dos segmentos é um *software* que apresenta um modelo de segmentação bastante satisfatório, baseado em crescimento de regiões, além de ser de uso livre, o que o torna acessível para qualquer instituição ou órgão público.

Primeiramente a imagem foi decomposta em três canais: azul, verde e vermelho. Estes canais já não correspondem às janelas espectrais do sensor original. Posteriormente é realizada a segmentação pelo processo de crescimento de regiões. Neste processo agrupamentos de pixels que possuem similaridade são definidos em uma mesma região. Processo feito de forma automática pelo sistema. O operador só define o número mínimo de pixels por segmento (polígono) e a similaridade. Segundo o critério de similaridade as regiões adjacentes vão sendo agrupadas. O processo é repetido até que nenhum outro agrupamento possa ser feito (INPE, 2012). Nos testes realizados se optou por um fator de similaridade igual a cinco e áreas com no mínimo vinte pixels. Com o uso destes parâmetros se verificou que

todos os alvos que se desejava classificar apresentavam os limites dos segmentos bem definidos. Porém sem gerar segmentos demasiados para definir um mesmo alvo (*oversegmentation*).

O próximo passo consiste na classificação utilizando-se das regiões segmentadas. Para isto foi escolhido classificador distância de Bhattacharya. Este possibilita um processo de classificação supervisionada. O analista define assim os segmentos que vão corresponder às classes identificadas. A distância de Bhattacharya mede o poder de discriminação estatística entre um par de classes espectrais. Ela mede a distância média entre as distribuições de probabilidades de classes espectrais (INPE, 2012). O classificador pode ser definido pela equação 1 (Moreira, 2005):

$$B(p_i, p_j) = \frac{1}{2}(m_1 - m_2)^T \sum (m_i - m_j) + \frac{1}{2} \ln \frac{|\sum (m_i - m_j)|}{|\sum i|^{\frac{1}{2}} |\sum j|^{\frac{1}{2}}} \quad (1)$$

Onde, B = Distância de Bhattacharya; p_i e p_j = pixels nas classes i e j; m_i e m_j = médias das classes i e j; T = matriz transposta; ln = logaritmo neperiano; i e j = classes dentro do contexto.

Foram realizados processos de classificação utilizando limiares de aceitação de 95, 99 e 99,9%, verificou-se que o uso dos dois primeiros níveis deixavam muitos segmentos sem classificação, e muitos destes correspondiam a edificações. Devido a este resultado se tomou o limiar de 99.9% de aceitação. Como efeito colateral se terá um aumento significativo nos erros de inclusão. Como haverá uma filtragem na pós-classificação existe a possibilidade de reduzir o erro acumulado. Foram definidas treze classes de uso do solo, sendo sete pertencentes a edificações e as demais a outros elementos, são elas:

- Telhado de Fibrocimento
- Telhado Cerâmico
- Telhado de Polímero ou Fibra de Vidro
- Quatro tipos de Terraços (sendo 4 tipos básicos de cobertura e revestimento)
- Solo exposto
- Asfalto
- Canteiros
- Vegetação Arbórea
- Vegetação Rasteira
- Sombra

Os segmentos utilizados como amostras para definir as classes foram convertidos em polígonos independentes para posterior utilização, como amostras, no processo de classificação pixel a pixel. Portanto os dois métodos de classificação utilizaram exatamente o mesmo conjunto de amostras, o que possibilita a comparação entre ambos os métodos. Para classificação pixel a pixel utilizou a Máxima Verossimilhança Gaussiana (MVG), classificador este amplamente conhecido e utilizado em reconhecimento de padrões. A função decisão da MVG é um caso particular do Teorema de Bayes. Pode ser descrita de uma forma mais simplificada pela equação (equação 2):

$$G_i(x) = -\ln |\Sigma_i| - (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \quad (2)$$

Os dados utilizados como verdade terrestre para quantificação das acurácias foi obtido, em realidade, por classificação visual da mesma imagem aqui trabalhada. Foi escolhida esta região por se conhecer bem a área, o que acabou por dispensar trabalho de campo para

elaboração do plano de informação (PI) referente à verdade terrestre. Este PI foi digitalizado seguindo o princípio de delimitar apenas as partes visíveis dos telhados e terraços, desconsiderando-se assim áreas sombreadas sobre estas superfícies, ou copas de árvores que por ventura estavam sobrepondo estas estruturas. Esta medida foi tomada porque se busca aqui avaliar a acurácia potencial dos processos de classificação digital (e de pós-processamento por SIG) e não criar uma metodologia para inferir uma determinada classe onde não exista a informação espectral contida nos contadores digitais da imagem.

Em ambiente de SIG todas as classes pertencentes a edificações foram agregadas a uma única classe denominada edificações, sendo as demais descartadas. Este primeiro resultado foi quantificado utilizando-se de verdade terrestre, tanto a MVG, quanto o produto da classificação por regiões.

Um segundo conjunto de resultados foi gerado, utilizando-se pós-classificação em SIG. Primeiramente foram filtrados os dados por um *buffer* baseado na malha viária, já que as coberturas de fibrocimento tendem a apresentar confusão com o asfalto. Para avenidas se definiu 22 metros e para vias comuns 14 metros, este intervalo é suficiente para cobrir as vias e boa parte das calçadas. O *Buffer* foi utilizado para cortar erros de classificação, neste caso asfalto classificado como edificações, tanto as provenientes da MVG quanto da classificação por regiões.

Todos os processos foram desenvolvidos dentro de ambiente de SIG. Para a segunda forma de filtragem foi utilizado o produto da classificação por regiões, já convertido na classe única denominada edificações. Primeiramente se converteu esta imagem contendo a classe edificações para o formato vetorial. O segundo passo consistiu na eliminação dos arcos internos, que separavam polígonos vizinhos. Com este procedimento se eliminou polígonos muito pequenos que anteriormente representavam classes diferentes de telhado ou terraço. Agora estas divisões não fazem mais sentido, pois, todas pertencem à nova classe edificação. Com a eliminação dos arcos internos o conjunto de todos os polígonos passa a ser representados por um único código identificador (ID); é desejável, porém, que polígonos que representam ilhas isoladas possam ser identificados isoladamente, para tanto se reorganizou o arquivo vetorial resultante, de forma que, cada polígono isolado passa a ser representado por um ID específico. Todas estas operações reduziram o número de polígonos (segmentos) originalmente classificados de 3738 para 939. Finalmente procedeu-se o novo processo de filtragem baseado especificamente no tamanho dos polígonos resultantes. Como os polígonos contíguos foram agrupados em um único grande polígono se pode inferir que em polígonos, que são ilhas isoladas, áreas muito pequenas possam representar, possivelmente, erros de classificação e não terraços ou telhados de edificações. Dessa forma foram cortados polígonos com áreas menores do que 20 m², depois áreas menores que 30 e 50 m². Todos estes resultados foram então quantificados com a utilização do PI referente à verdade terrestre. Assim erros de inclusão e omissão foram medidos. Empiricamente já é possível saber que quanto mais se aumentar o tamanho das áreas de exclusão mais erros de inclusão serão cortados, da mesma forma os erros de omissão vão aumentar com o aumento da área excluída. A quantificação vai definir em que tamanho de área excluída acontece uma diminuição significativa nos erros de inclusão com um aumento pequeno nos erros de omissão, de forma que haja um ganho na acurácia geral da classificação. Este método de filtragem aqui adotado foi utilizado apenas no produto da classificação por regiões. Os processos de classificação pixel a pixel geram muitas áreas classificadas onde existem apenas um, ou poucos, pixels isolados dentro de uma segunda classe espectral. Esta característica tornaria inviável o método aqui desenvolvido para a classificação pixel a pixel. No caso do processo de classificação por regiões o método foi satisfatoriamente bem sucedido, já que foi pensado em função das características espaciais deste tipo classificação.

3. Resultados e Discussão

Muitos estudos tem demonstrado que os processos de classificação orientada ao objeto se mostram mais eficazes que os processos pixel a pixel, quando aplicados em áreas urbanas. Esta melhor eficácia da classificação orientada ao objeto pode ser encontrado na pesquisa desenvolvida por Batista e Haertel (2007), nos anais do XIII SBSR. Os resultados obtidos aqui confirmam o que já era esperado, um desempenho bem mais satisfatório da classificação por regiões. Mesmo antes dos processos de filtragem o produto gerado pela classificação por regiões apresentou um desempenho melhor do que o processo pixel a pixel.

Existem 220782 m² ocupados por edificações na área de estudo, segundo o PI de verdade terrestre. Desta área 163028 m² foram identificadas pelo processo pixel a pixel, uma acurácia de 73%. A classificação por regiões pura classificou corretamente 187847 m², acerto de 85%. Registrou, portanto, um erro de omissão de 15% contra 27% do processo de classificação pixel a pixel. Os erros de inclusão consistem em pixels que foram classificados como pertencentes à classe em questão, mas que em realidade não o são. Em termos de erros de inclusão o processo de classificação por regiões apresentou uma vantagem ainda maior do que no processo pixel a pixel. Enquanto o processo pixel a pixel obteve 50,4% de erros de inclusão a classificação por regiões obteve um erro de 35%. Os processos de filtragem, abordados a seguir, visam reduzir este tipo de erro. A tabela 1 mostra a quantificação dos processos de classificação e filtragem.

Tabela 1. Quantificação dos processos de classificação e filtragem.

CLASSIFICAÇÃO	Área – Classificada (m ²)	Acurácia (m ² e %)	Erros de Omissão	Erros de Inclusão
1. Verdade Terrestre	220782			
2.1. Clas. Pixel a Pixel MAXVER	328814	163028 (73%)	27%	165786 (50,4%)
2.2. Clas. Pixel a Pixel MAXVER -Filt. Vias	264852	162638 (73%)	27%	102214 (31%)
3.1. Clas. Or. Ao Objt.	291917	187847 (85%)	15%	104070 (35%)
3.2. Clas. Or. Ao Objt. - filt. Vias	260904	187422 (84,9%)	15,10%	73481 (28,1%)
3.3. Clas. Or. Ao Objt. - filt. < 20m ² e Vias	258236	196524 (84,4%)	15,60%	71712 (24,5%)
3.4. Clas. Or. Ao Objt. - filt. < 30m ² e Vias	256340	185957 (84,2%)	15,80%	70382 (24,1%)
3.5. Clas. Or. Ao Objt. - filt. < 50m ² e Vias	254142	185490 (83,4%)	16,60%	68988 (23%)

Como o método aqui desenvolvido, baseado no corte de polígonos que representam áreas muito pequenas só é viável na classificação por regiões, o produto da MVG só foi filtrado pelo *buffer* referente ao sistema viário. Este procedimento não modificou os erros de omissão, mas, reduziu os erros de inclusão para 31%, mais de 19% de redução em relação ao resultado original. Na classificação orientada ao objeto a filtragem pelo sistema viário causou um aumento insignificante nos erros de omissão e uma diminuição de quase 7% nos erros de inclusão. A este resultado foi adicionada a filtragem para áreas menores do que 20 m². Esta segunda filtragem causou uma nova redução de 3,6% nos erros de inclusão com um aumento novamente menor nos erros de omissão, na casa dos 0,4%. Esta tendência se manteve, na medida em que foi aumentado o tamanho da área a ser excluída. Houve sempre um ganho mais significativo nos erros de inclusão com um incremento sempre de menor magnitude nos erros de omissão. A figura 1 contém um gráfico que demonstra o comportamento da acurácia com os progressivos processos de filtragem. No eixo x, se obtém as etapas de filtragem (mesmas da tabela 1), e no y o erro percentual.

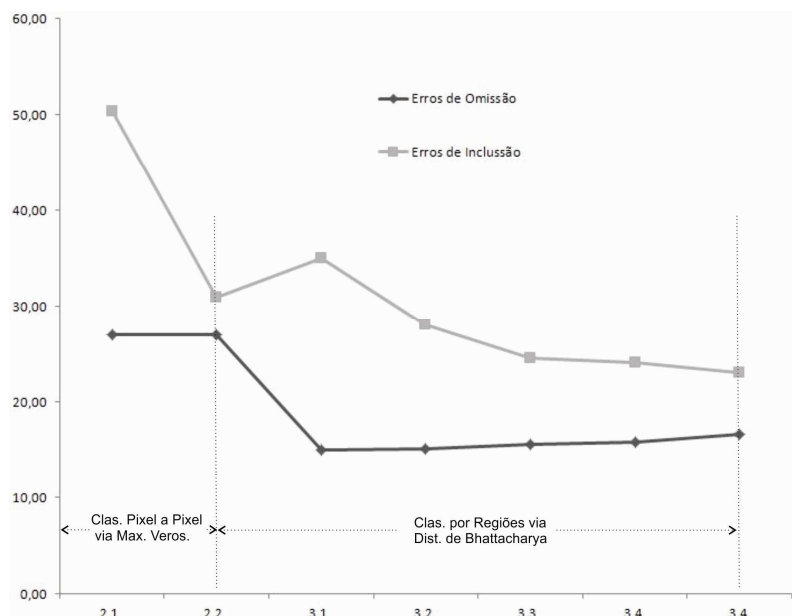


Figura 1. Comportamento da acurácia com os processos de filtragem.

Nesta figura os estágios 2.1 e 2.2 correspondem à classificação pixel a pixel via MVG, os demais estágios correspondem à classificação por regiões via distância de Bhattacharya. Como se observa no gráfico a classificação pixel a pixel só apresentou um resultado melhor do que a classificação por regiões para os erros de inclusão, no estágio 2.2 em relação ao 3.1, ou seja, a MVG já filtrada pelas vias em relação à classificação por regiões pura.

Com o aumento cada vez maior da área a ser excluída a tendência é que as duas curvas, vistas no gráfico tendam a se encontrar (erros de inclusão e omissão). Neste caso a relação de filtragem por tamanho passara a não ser mais compensatória. Acredita-se que a partir dos 100 m² de área ocupada por edificações a relação de filtragem já comece a não ser vantajosa.

No último estágio foram cortadas áreas abaixo de 50m². Se comparado ao produto original houve uma redução de 12% nos erros de inclusão, de 35% para 23% de área classificada indevidamente. O consequente aumento dos erros de omissão para esta mesma faixa foi bastante baixo 1,6 %, o que comprova que a relação nas filtragens em áreas com até 50 m² é muito satisfatória, já que o ganho geral de acurácia é de 10,4%. A figura 2 contém um gráfico mostrando a redução dos erros de inclusão em relação ao produto puro e aos respectivos estágios finais de filtragem.

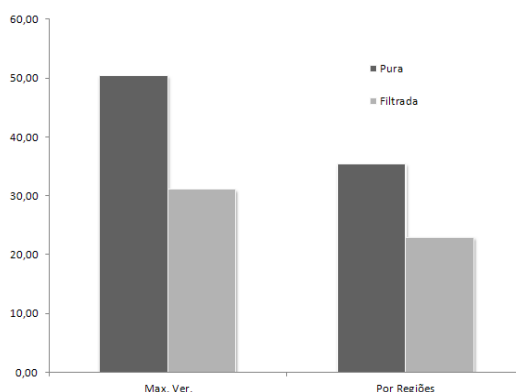


Figura 2. Erros de Inclusão antes e depois dos processos de filtragem.

A figura 2 consegue demonstrar de forma mais clara que embora o processo de classificação por regiões acumule menos erros de inclusão que o pixel a pixel o processo final

de filtragem (vias e áreas com menos de 50 m²) conseguiu melhorar substancialmente a acurácia final da classificação. A figura 3 mostra um mapa contendo um pedaço da área de estudo e os respectivos erros de inclusão, omissão e área corretamente classificada, após o processo de filtragem final.

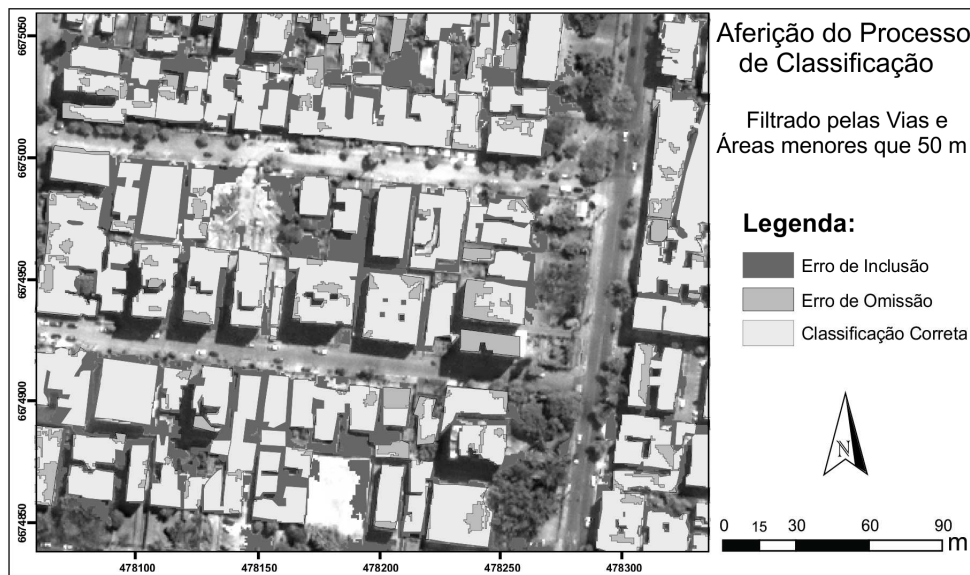


Figura 3. Pedaço da área de estudo com a área classificada e os erros calculados.

4. Conclusões

Os resultados aqui obtidos comprovam que é viável a utilização de imagens sintéticas provenientes do sensor Quick Bird para os processo de classificação digital, quando utilizando-se da classificação orientada ao objeto. Já que erros como de omissão de 15% e 35% para inclusão podem ser considerados aceitáveis para muitos estudos. No caso da classificação pixel a pixel pura erros de inclusão de mais de 50% demonstram que a utilização deste tipo de classificação é desaconselhável.

Os processos de filtragem aqui desenvolvidos podem ser aplicados em qualquer tipo de classificação orientada ao objeto, e não necessariamente em imagens sintéticas. O resultado final desta filtragem mostrou-se muito adequado para maioria dos estudos urbanos, já que após a filtragem o erro máximo obtido não passou de 23% para inclusão. Condicionando um aumento nos erros de omissão de 15 para 16,6%. O pequeno aumento registrado nos erros de omissão se mostrou muito aceitável, visto a grande diminuição nos erros de inclusão. Em se tratando de informações comuns e fáceis de desenvolver em SIG utilizou-se apenas o sistema viário, no entanto diversas informações disponíveis podem ser utilizadas como, por exemplo, informações provenientes dos planos diretores dos municípios. Muitos destes dados já se encontram especializados e em formato digital, cabe ao pesquisador à seleção e uso adequado destas informações. Foi dado um ênfase maior, nesta pesquisa, a filtragem por tamanho de segmentos classificados. Esta última metodologia pode ser a melhor contribuição deste estudo, já que não foi encontrado um procedimento semelhante na literatura, soma-se a isto a fato de o método ter se mostrado eficaz e de fácil aplicação para redução de erros de omissão.

Agradecimentos

Neste estudo foi utilizado um recorte sobre uma imagem sintética do sensor Quick Bird que cobre toda área urbana do município de Porto Alegre – RS. Esta imagem foi gentilmente cedida pelo DEMAB (Departamento Municipal de Habitação), órgão este pertencente a

prefeitura do Município de Porto Alegre. Agradecemos a este departamento pela utilização sem fins lucrativos da imagem. Por fim agradecemos a CAPES por financiar este estudo e outras atividades desenvolvidas pelo primeiro autor em diversas atividades realizadas na Universidade Federal do Rio Grande do Sul.

Referências Bibliográficas

Centeno, J.; Souza G.. **Uso dos fractais na análise da fragmentação de uma cena urbana através de imagem de alta resolução.** In: XIV Jornadas de Jóvenes Investigadores, 2006, Campinas. Anais do XIV Jornadas de Jóvenes Investigadores, 2006. v. 1. p. 1-9.

eCognition. **User's Guide 4.** Definiens Imaging. Muchen, Germany, 2004.

Fukunaga, K. **Introduction to Statistical Pattern Recognition.** Ed. Boston: Academic Press, 2nd edition, 1990.

Batista, M; Haertel, V. **Classificação hierárquica orientada a objeto em Imagens de alta resolução espacial empregando atributos espaciais e espectrais.** XIII Simpósio Brasileiro de Sensoriamento Remoto, Anais, Florianópolis, Brasil, 21–26 de abril 2007, INPE, p. 489-497.

INPE, Instituto Nacional de Pesquisas Espaciais. **Sistema de Processamento de Informações Georreferenciadas – SPRING 5.2 – Manual do Usuário.** São José dos Campos, SP. 2012.

Leonardi, F.; Almeida C.; Fonseca L.; Camargo F. **Avaliação Comparativa entre Classificação Supervisionada por Regiões e Orientada a Objeto para Imagens de Alta Resolução Espacial: Cbers 2B-HRC e QuickBird.** XIV Simpósio Brasileiro de Sensoriamento Remoto, Anais, Natal, Brasil, 25-30 abril 2009, INPE, p. 981-988.

Moreira, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação,** Ed. UFV, 3ed. 320p, 2005.