

SEGMENTAÇÃO AUTOMÁTICA DE ÁREAS URBANAS EM IMAGENS DE SENSORIAMENTO REMOTO

Anderson G. Marco¹, Adriana B. Bruno², Francisco A. Rodrigues³, Luciano da F. Costa¹ e Odemir M. Bruno¹

¹ Universidade de São Paulo
Instituto de Física de São Carlos
anderson.marco@gmail.com; {luciano,bruno}@ifsc.usp.br

² Universidade de São Paulo
Instituto de Arquitetura e Urbanismo
adrianabb@sc.usp.br

³ Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
francisco@icmc.usp.br

Resumo – A adoção de ferramentas e métodos de visão computacional aplicados as imagens de sensoriamento remoto pode trazer um grande benefício para os estudos de urbanismo. Neste trabalho é apresentado um arcabouço computacional para auxiliar a aquisição e análise de imagens de sensoriamento remoto, provindas do Google Earth®. O sistema proposto se conecta ao Google Earth e compõe uma imagem especificada pelas coordenadas de latitude e longitude informadas pelo usuário. O sistema também segmenta automaticamente a imagem em áreas edificadas e não-edificadas. O processo de segmentação é realizado por meio de métricas do histograma cromático. É apresentado um experimento com imagens de duas cidades do interior paulista, São Carlos e Rio Claro, a fim de validar o sistema e determinar quais as melhores métricas de histogramas devem ser utilizadas para a tarefa.

Palavras-chave – sensoriamento remoto, segmentação de imagens, análise de histograma.

Abstract – Computer vision methods applied in remote sensing images can bring important benefits to the field of urbanism. In this work, a computer framework to aid acquiring and analyzing remote sensing images from Google Earth® is presented. The proposed system connects to the Google Earth and extracts an image specified by the latitude and longitude coordinates. The image is automatically segmented into edified and non-edified regions. The segmentation process is carried out by the analysis of the image chromatic histogram. An experiment showing results from the segmentation of two São Paulo State's towns are presented. The comparison of the histogram metrics and the color channels and its segmentation performance are discussed.

Keywords – remote sensing image, image segmentation, histogram analysis.

1 Introdução

A cidade, mais do que uma aglomeração de pessoas, é o espaço no qual se constituem complexas relações entre a população, as instituições públicas e as privadas, possibilitando a concretização de idéias firmadas em conceitos como desenvolvimento, progresso e modernidade, conceitos esses arraigados no consenso geral da humanidade e que estão atrelados ao ideário de trazer benefícios à mesma. Portanto a compreensão desse espaço urbano se faz vital para uma melhor sobrevivência humana uma vez que uma urbanização planejada pode retornar em significativos benefícios para seus habitantes, assim como problemas de ordem territoriais, econômicos e sociais, bem como seus efeitos secundários, como a baixa qualidade de vida, precariedades na infra-estrutura, criminalidade, sistemas de trânsito mal resolvidos, podem ser evitados ou então minimizados mediante um bom planejamento da atuação no ambiente urbano.

Nesse contexto, consolida-se o urbanismo, ciência relacionada com o estudo, ordenamento, controle e planejamento da cidade e da urbanização como de caráter eminentemente multidisciplinar. Áreas como a geografia, ciências sociais, ecologia, entre outras contribuem com o estudo por meio de seus conhecimentos, assim como novas tecnologias provenientes de áreas

como física e computação podem trazer uma nova perspectiva no campo em questão. Uma vez que a materialidade construída da cidade é reflexo do modo cultural de viver da população e entendendo a cultura como algo que se transforma ao longo do tempo, a cidade acaba por ser concebida como espaço vivo e em constante mudança. Logo, voltadas ao dinamismo que caracteriza o processo de produção do espaço urbano, imagens de sensoriamento remoto são um excelente recurso no sentido de monitorar e acompanhar esse desenvolvimento urbano. Nesse sentido, imagens de sensoriamento remoto vêm sendo utilizadas como um recurso adicional a fim de subsidiar o estudo do espaço urbano. Ao conjugá-las com técnicas computacionais de reconhecimento de padrões e análise de imagens percebeu-se um adicional no seu potencial como fonte de pesquisa. Desse modo, inicia-se em 2007 uma pesquisa sobre o tema [1–3] destinada a verificar relações entre a morfologia urbana e imagens de sensoriamente remoto, da qual esse trabalho é decorrente. Notavelmente, a aplicação de ferramentas computacionais em imagens de sensoriamento remoto visando estudos urbanos permite uma análise quantitativa e objetiva, a qual pode ser utilizada para a investigação de aspectos relacionados às configurações espaciais dos bairros, relacioná-los com a qualidade de vida, bem como monitorar a expansão da malha urbana e o índice de ocupação da cidades, fornecendo subsídios às políticas de planejamento e gestão de cidades. Além disso, a identificação nas imagens de satélite de outros elementos, como por exemplo, a distinção de áreas urbanas e áreas verdes, pode ser utilizada para o acompanhamento do avanço da mancha urbana sobre áreas críticas de cinturões verdes, o monitoramento de áreas de preservação ambiental [4], ou mesmo o simples mapeamento de vazios urbanos presentes em áreas subutilizadas de determinadas cidades. Esses estudos visam contribuir com a ciência urbana promovendo melhorias na interpretação da imagem visual.

Uma das dificuldades iniciais do projeto era a lenta construção de um banco de amostras decorrente do processo manual de aquisição das imagens do software Google Earth, o que acarretava em dispêndio de considerável tempo utilizado pelo especialista para a execução da tarefa, e consequentemente, reduzia a possibilidade de estudos urbanísticos mais abrangentes. Como forma de sanar essa demanda foi proposto um arcabouço visando a coleta automática das imagens de satélite, bem como a classificação automática de áreas edificadas.

A utilização das imagens de sensoriamento remoto se apresenta como uma relevante fonte de dados para a leitura e compreensão de padrões morfológicos urbanos. A morfologia urbana é o resultado da interação de elementos constituintes da cidade, como sistema viário, lotes, quadras, áreas verdes e edificações, que por sua vez variam em função dos modos de uso e ocupação do solo, bem como das condições geográficas existentes. Nas imagens de satélite, a morfologia urbana se traduz em uma superfície de textura, definida pela organização dos pixels [5,6], que pode ser mais ou menos complexa de acordo com os arranjos espaciais do território que representa.

A organização dos pixels conjugada com o atributo cor, são elementos que se relacionam à representação de áreas edificadas e não edificadas da superfície terrestre, objetos de estudo desse artigo, portanto ambos os atributos serão utilizados como fontes para a análise. A análise de textura está relacionada com a distribuição e arranjo dos pixels, sejam eles cromáticos ou em níveis de cinza de uma imagem. Neste trabalho, a análise de textura foi realizada por meio da análise da distribuição do histograma e de medidas estatísticas obtidas por meio deste.

Este trabalho apresenta um arcabouço computacional que se conecta com o Google Earth e adquire imagens relacionadas às áreas estabelecidas pelo usuário e as classifica. O processo de segmentação é realizado pelos atributos de textura e cor, que são analisados por medidas de histogramas e classificados pelo método K-médias. Nas próximas sessões, além de detalhar o funcionamento do sistema será apresentado um experimento de classificação do espaço urbano de duas cidades do interior paulista: São Carlos e Rio Claro.

2 Arcabouço

O arcabouço desenvolvido possibilita a automatização tanto do processo de aquisição das imagens do Google Earth, como a segmentação das mesmas. Desenvolvido em Java, e operacional em multi-plataformas, facilita a realização de experimentos urbanos em imagens de cidades. Em seu processo de aquisição, o usuário informa ao arcabouço as coordenadas geográficas (latitude e longitude) do canto superior esquerdo e inferior direito da área a ser adquirida. Para tanto, conecta-se com o Google Earth e como este limita a devolução das imagens em um tamanho de 600x600 pixels, o sistema varre a área limitada pelas coordenadas, devolvendo um conjunto de n imagens fragmentadas. O processo de concatenação dessas imagens requer uma efetiva atenção do sistema, a fim de garantir uma justa sobreposição dos pixels das bordas das imagens. O resultado é uma imagem única, com a qualidade de 150 dpi pronta para ser utilizada. (veja Fase 1 da Figura 1). A partir de então se inicia a Fase 2 (veja Figura 1) do arcabouço direcionado ao módulo de segmentação. Utilizando as metodologias análise de histograma e o algoritmo estatístico de k-médias [7], que serão descritas na próxima sessão, o arcabouço segmenta automaticamente a imagem de satélite classificando-a em dois grupos, o primeiro contendo as áreas edificadas e o segundo as áreas não-edificadas (entre as quais se incluem tanto vegetações densas como os cinturões verdes, assim como as áreas rurais e pequenas espaços gramados ou arborizados na região urbana). O processo de segmentação tem início quando o arcabouço fragmenta a imagem em uma malha. Cada elemento da malha é analisado, classificado e rotulado. O tamanho da malha determina a resolução e a granulosidade do resultado. Se por um lado, em malhas grossas, cada elemento apresenta regiões maiores e com mais informações para fornecer uma melhor análise de textura, elas também podem ter diferentes classes contidas no mesmo elemento, prejudicando o resultado. Nas malhas finas, a resolução do resultado é melhor e as classes mais definidas, porém se a região for demasiadamente pequena, a informação será pobre para classificar a textura (Para este trabalho foi adotada empiricamente uma janela de 10x10).

Para cada elemento da malha é realizada uma análise de textura cromática. Ela é feita por meio da análise de histograma. Foram consideradas as seguintes medidas obtidas dos histogramas neste trabalho: média, variância, limiar de Otsu e entropia. Por

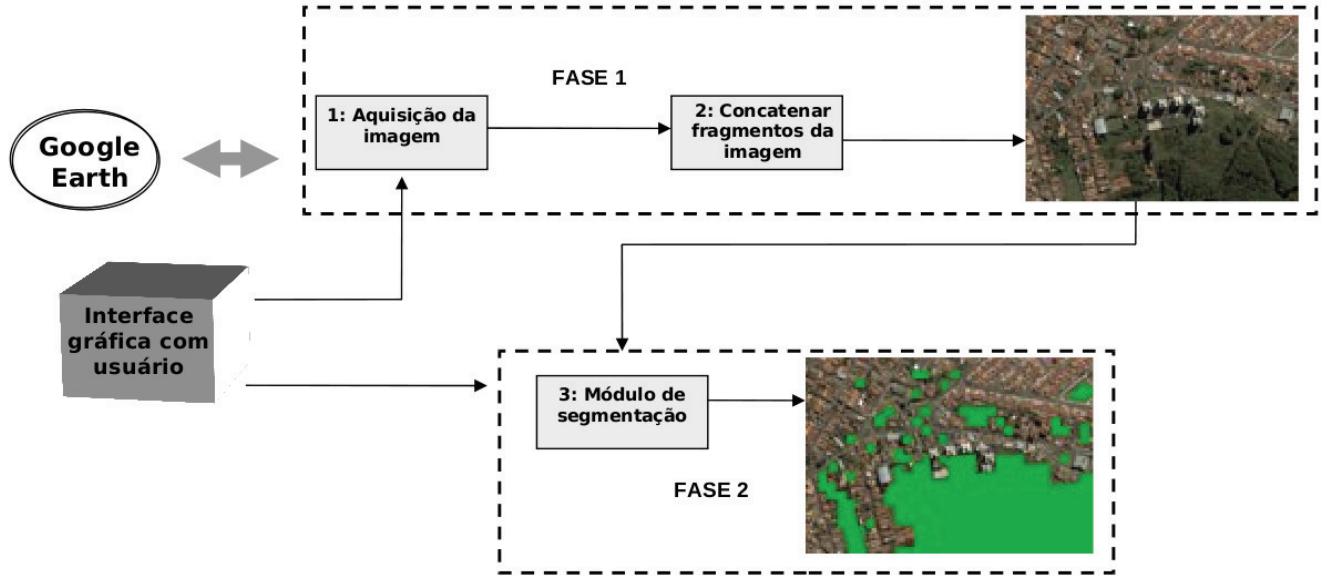


Figura 1: Diagrama mostrando o funcionamento do arcabouço.

fim, com base no vetor de característica calculado de cada amostra é realizada a classificação. O sistema utiliza o sistema não-supervisionado k-médias. Dessa forma, introduz-se apenas o número de classes que se pretende obter (no caso aqui, de valor 2) e o método segmenta a imagem em duas classes sendo edificadas e não-edificadas (áreas verdes, lagos e terrenos não-edificados).

3 Algoritmo de Segmentação

O método de segmentação proposto neste trabalho é baseado na análise de histogramas. A imagem é dividida em janelas de 10x10 pixels, e para cada janela é calculado um vetor de características com base no seu histograma. Os vetores de características são classificados pelo algoritmo k-médias e com isso as regiões são rotuladas em edificadas e não-edificadas. Nesta seção são apresentados os métodos de análise de histograma e classificador k-médias.

3.1 Análise de Histograma Colorido

Sendo a imagem uma função $f(x,y)$, o histograma é uma função mostrando o número de pixels da imagem que possuem uma determinada intensidade [8], dada por:

$$\sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \delta(f(x,y), i), \quad \text{onde } \delta(j, i) \text{ é a função de Kronecker}$$

$$\delta(j, i) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}$$

Com base no histograma foram extraídas quatro métricas para caracterizar as janelas das imagens: média, desvio padrão, entropia e limiar de Otsu. A média e o desvio padrão são as mais simples, fundamentadas em momentos de primeira ordem da análise do histograma. A métrica é responsável por medir a tendência dos níveis de cinza e o desvio padrão, o quando o histograma varia. As equações a seguir calculam respectivamente a média e o desvio padrão de um histograma:

$$\mu = \sum_{i=0}^{G-1} ip(i), \quad \sigma^2 = \sum_{i=0}^{G-1} (i - \mu)^2 p(i).$$

A entropia de Shannon [4] é uma medida clássica adotada em teoria da informação que determina o grau de embaralhamento ou desorganização. A equação a seguir é utilizada para calcular a entropia de um histograma:

$$H = - \sum_{i=0}^{G-1} p(i) \log_2[p(i)].$$

O algoritmo de Otsu [9] é tradicionalmente utilizado como método de binarização automática. O algoritmo analisa o histograma e localiza o ponto mais adequado para limiarizar a imagem em dois níveis de profundidade. O método é baseado no cálculo da variância das duas classes obtidas. Para cada possível limiar é calculada a variância intra e inter-classes. O valor eleito é aquele na qual é otimizada a mínima variância intra-classe e a máxima variância entre-classes [9]. Neste trabalho o limiar obtido pelo método de Otsu é utilizado com métrica de histograma.



Figura 2: Mosaico gerado pelo sistema de aquisição para a cidade de São Carlos (livre de escala).

3.2 Algoritmo estatístico de k-médias [7]

K-médias é um método não supervisionado de análise de aglomerados o qual visa partitionar n observações para K aglomerados, em que cada observação pertence ao aglomerado com a distância mais próxima de sua média. O algoritmo k-médias consiste

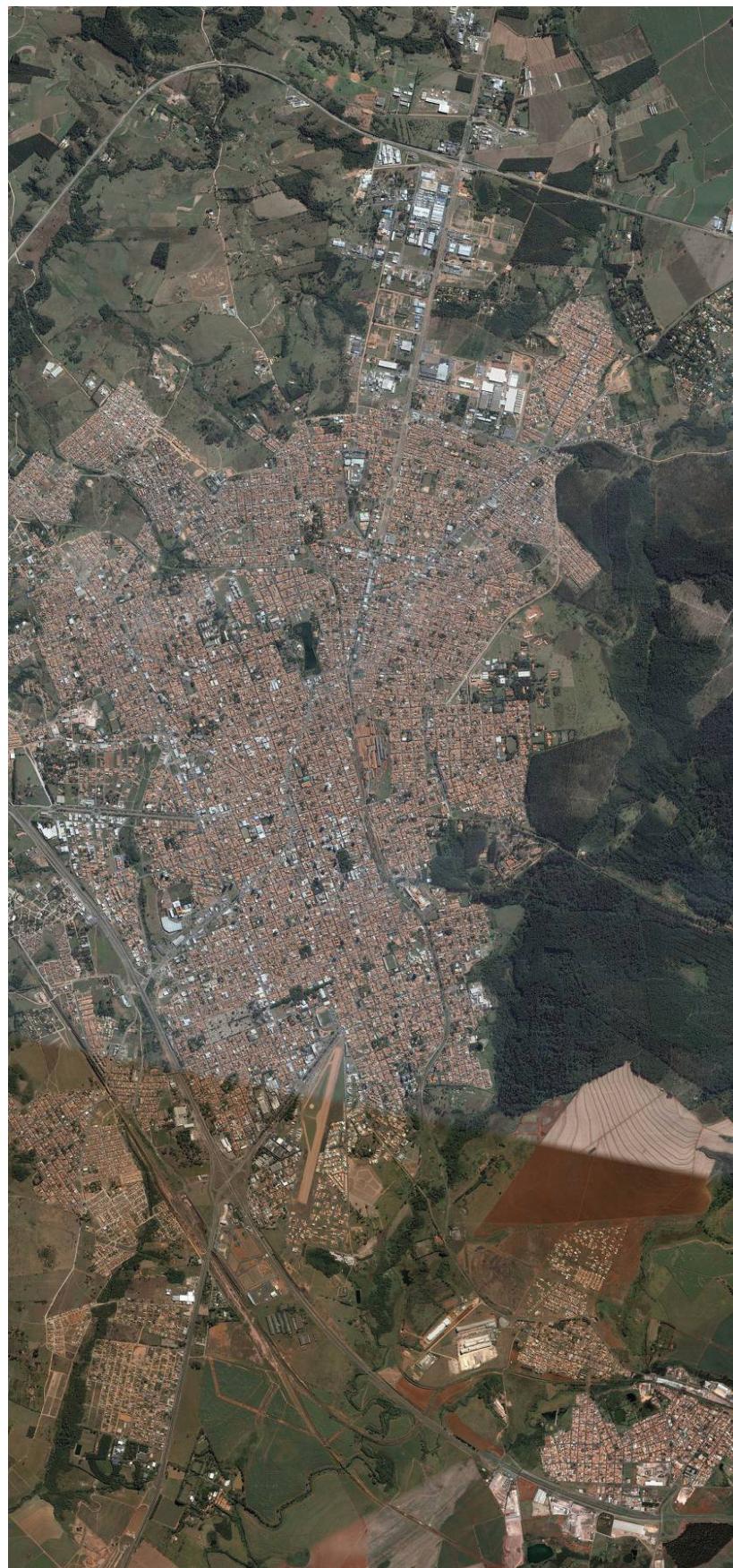


Figura 3: Mosaico gerado pelo sistema de aquisição para a cidade de Rio Claro (livre de escala).

em um processo iterativo onde são informados os números de aglomerados (K) e também um conjunto inicial 1 de sementes μ_i^1 , onde i corresponde a cada aglomerado. A determinação do conjunto de observações baseia-se na proximidade de cada elemento do cluster μ como mostra a equação a seguir:

$$G_i^t = \{x_j : \|x_j - \mu_i^t\| \leq \|x_j - \mu_{q_i}\|\}, q_i = \{y : y \neq i : 1 < y < k\}$$

$$\mu_i^{t+1} = \frac{1}{n(G_i^t)} \sum_{x_j \in G_i^t} x_j$$

Depois disso, a aparência de cada cluster é calculada baseada em G_i^t . O processo continua até que o conjunto G^t e G^{t+1} se tornem iguais.

4 Experimento

Para a realização do experimento, o sistema extraiu imagens do Google Earth referentes às cidades de São Carlos e Rio Claro. O mosaico gerado para as duas cidades São Carlos e Rio Claro é apresentado nas Figuras 2 e 3 respectivamente. No experimento as imagens foram segmentadas em regiões edificadas e não-edificadas. O processo de segmentação foi realizado utilizando diferentes parâmetros do algoritmo, tais como a combinação de métricas de histogramas e sua análise isolada e combinação de canais de cores e/ou níveis de cinza. Para avaliar a acurácia do sistema foi realizada uma comparação do mesmo com uma imagem segmentada manualmente por um especialista. A Figura 4 mostra a imagem de alta resolução da cidade de São Carlos, uma determinada região ampliada e sua correspondente, classificada pelo especialista em áreas edificadas e não-edificadas.



Figura 4: Mapa da cidade de São Carlos e o recorte de uma determinada região (a), com o processamento de áreas edificadas e não-edificadas realizados por especialista (b).

5 Resultados

A Figura 5 apresenta a distribuição dos padrões reconhecidos num gráfico de scatter para as variáveis. Para a redução dimensional foi utilizado o modelo de componentes principais (PCA) [7]. Neste modelo, ao invés de utilizar a classificação k-médias diretamente nos dados, houve um transformação linear projetando os dados no melhor arranjo de componentes. Cada um dos gráficos apresenta as janelas de 10x10 pixels classificadas pelo especialista. Para cada janela foram extraídas as métricas (Entropia, Média, Otsu e Desvio Padrão), e sua combinação dos canais R, G, B e todos os canais (R+G+B+Gray). Neste gráfico um pequena região da cidade de São Carlos foi utilizada (contendo apenas 12000 janelas) a fim de permitir a visualização

dos dados. No gráfico, o vermelho (1) representa as áreas interpretadas como edificadas pelo especialista e o azul (0) as não-edificadas. Como pode ser observado, as métricas apresentam um bom potencial de separação das classes, agrupando os dados com certa distinção. Entre as métricas, a entropia e a média apresentaram os melhores potenciais de discriminação.

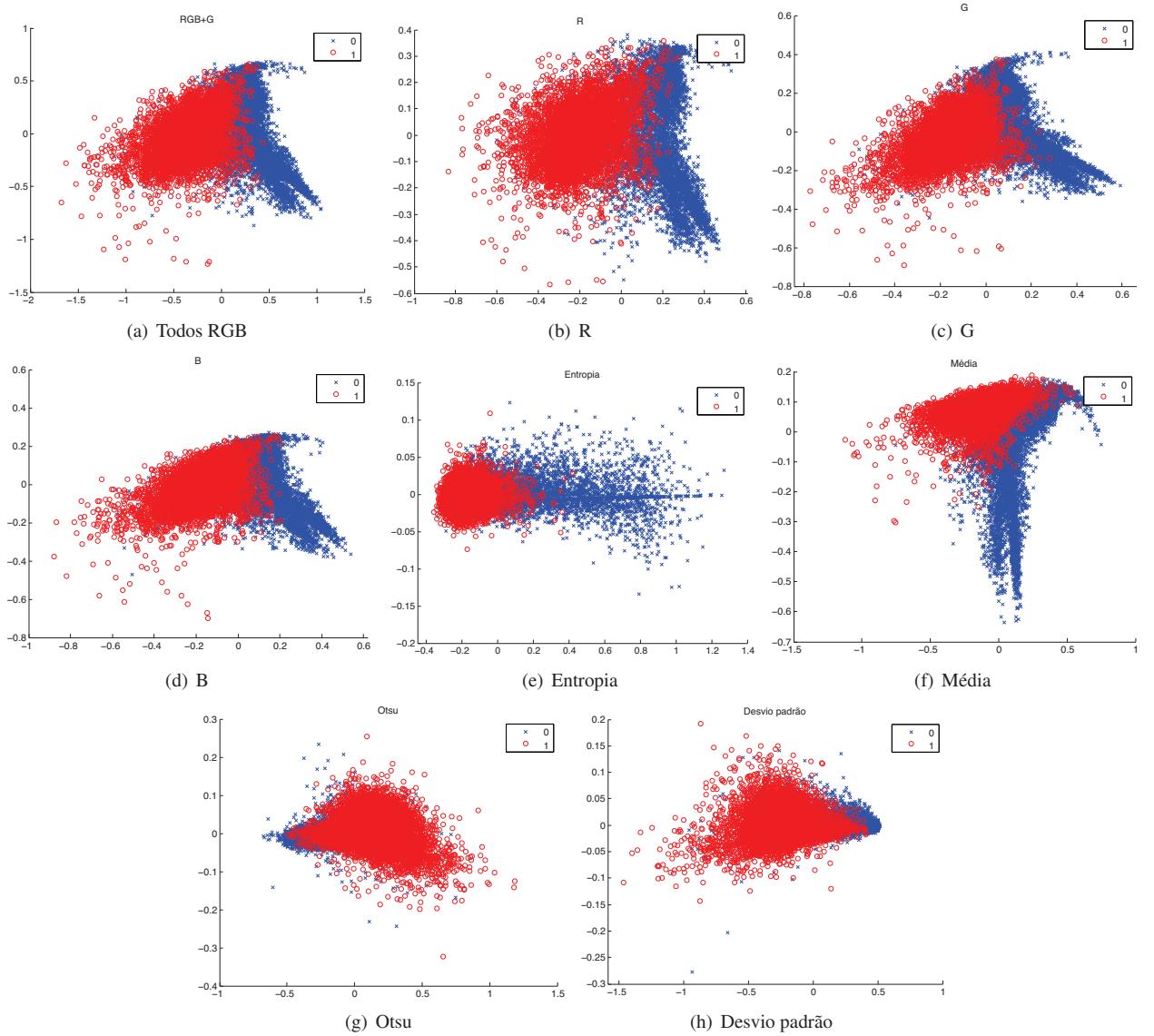


Figura 5: Gráficos de dispersão, mostrando as duas primeiras componentes, confrontando a segmentação realizada pelo especialista e pelas das métricas aplicadas nos canais de cor RGB.

A Figura 6 mostra a região ampliada da Figura 4 e o resultado da segmentação da mesma por diferentes métodos. A Figura 6c exibe a combinação das métricas (média, desvio padrão, entropia e Otsu) nos canais cromáticos R, G e B e também no nível de cinza. As Figuras 6d, 6e e 6f apresenta a mesma combinação usando apenas um dos canais cromáticos, R, G e B respectivamente. A Figura 6b, apresenta a segmentação realizada pelo especialista, a qual é utilizada para validar e analisar os resultados obtidos pelo sistema. Nas Figuras 6g, 6h, 6i e 6j, são realizadas a segmentação de apenas uma métrica (respectivamente, entropia, média, desvio padrão e Otsu) nos três canais cromáticos.

As Tabelas 1 e 2, mostram os erros de segmentação obtidos para a segmentação da cidade de S. Carlos. Como erros são considerados a classificação incorreta da área edificada e da não-edificada. Na primeira tabela os resultados combinam as quatro métricas nos canais cromáticos e níveis de cinza e na segunda é realizada uma análise de cada métrica nos canais cromáticos. Considerando os dois erros (edificada e não-edificada) o melhor resultado foi obtido por todos os métodos aplicados nos canais R+G+B+Cinza. O segundo melhor resultado foi obtido pela análise do canal azul (B). Considerando a segunda tabela, o melhor método foi da entropia e o pior resultado foi do desvio padrão. Os resultados da segmentação da cidade de Rio Claro são apresentados nas Tabelas 3 e 4. Observando a primeira tabela, os melhores resultados foram do canal azul e em seguida utilizando todos os canais mais os níveis de cinza. Este resultado confirma a importância do canal azul na segmentação das regiões edificadas. Na segunda tabela, o melhor resultado foi novamente a entropia.

Em termos gerais os resultados obtidos pelo sistema foram satisfatórios. Considerando todas as métricas e todos os canais obteve cerca de 14% de erro de áreas edificadas e 7% de não-edificadas em S. Carlos e 9% de erro de áreas edificadas e 8%

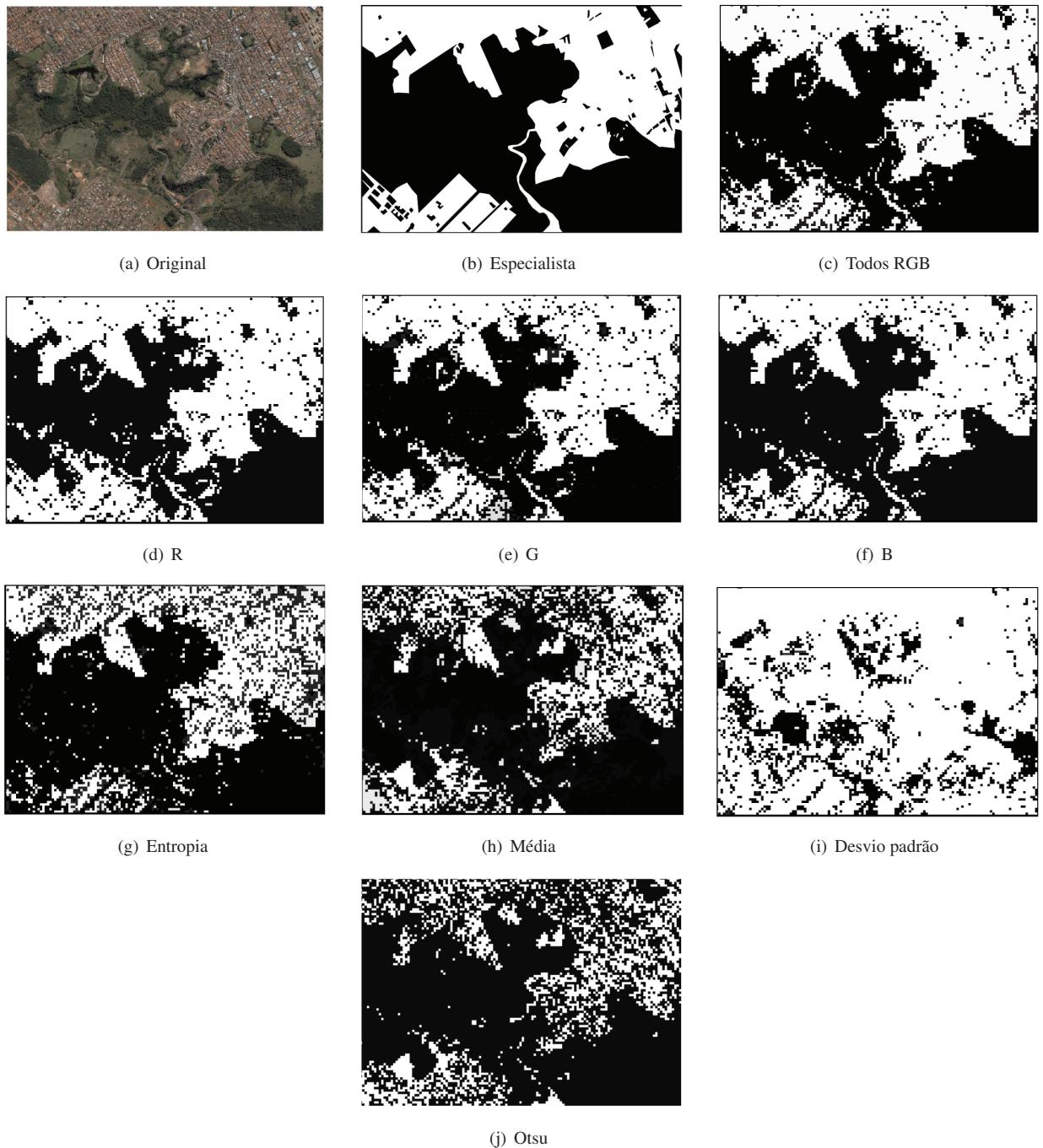


Figura 6: A região de São Carlos e os exemplos de resultados obtidos.

de não-edificadas em Rio Claro. Estes resultados são adequados para a segmentação automática com o propósito de análise de características urbanas.

Tabela 1: Resultados das métricas nos canais cromáticos e níveis de cinza, aplicados à cidade de São Carlos.

São Carlos		
	edificada	não edificada
R+G+B+gray	14,155	7,055
R	11,205	10,9
G	15,41	7,45
B	16,26	5,805
gray	24,695	5,31

Tabela 2: Resultados das métricas individuais nos canais cromáticos, aplicados à cidade de São Carlos.

São Carlos		
	edificada	não edificada
media	48,59	4,645
entropia	38,725	4,92
desvio	4,57	58,235
otsu	57,33	4,43

Tabela 3: Resultados das métricas nos canais cromáticos e níveis de cinza, aplicados à cidade de Rio Claro.

Rio Claro		
	edificada	não edificada
R+G+B+gray	9,19	8,385
R	9,83	8,52
G	9,025	9,195
B	7,29	9,75
gray	15,6	6,1

Tabela 4: Resultados das métricas individuais nos canais cromáticos, aplicados à cidade de Rio Claro.

Rio Claro		
	edificada	não edificada
media	34,47	50,295
entropia	28,67	4,125
desvio	5,375	39,155
otsu	34,07	7,15

6 Conclusão

Neste trabalho foi apresentado um arcabouço computacional para extrair imagens do Google Earth, mediante coordenadas informadas pelo usuário e segmentá-las automaticamente classificando as áreas edificadas e não-edificadas. O sistema surgiu de um projeto iniciado em 2007, o qual utiliza visão computacional para quantificar e analisar atributos urbanos em imagens de sensoriamento remoto.

O processo de segmentação adotou a análise de histogramas de canais cromáticos e níveis de cinza. Nos resultados foram comparados a combinação de quatro métricas (média do histograma, desvio padrão, entropia e limiar de Otsu) aplicadas nos canais cromáticos e também analisadas independentemente. Os resultados mostraram que a melhor métrica extraída da análise do histograma é a entropia e a informação cromática, em especial do canal azul (modelo RGB), que apresentou relevância na caracterização das regiões edificadas.

Os resultados também demonstraram a viabilidade da segmentação automática de regiões edificadas e não-edificadas. Deste modo, o arcabouço proposto será de grande valia para o desenvolvimento de pesquisas de urbanismo em imagens de sensoriamento remoto.

Agradecimentos

Odemir M. Bruno agradece ao CNPq (Procs. #303746/2007-1 e #504476/2007-6). Anderson G. Marco agradece ao CNPq (Proc. #504476/2007-6) pela bolsa de iniciação científica.

References

- [1] A. R. Backes, A. B. Bruno, M. N. B. Filho and O. M. Bruno. “Dimensão Fractal Volumétrica aplicada à imagens urbanas de sensoriamento remoto”. In *IV Workshop de Visão Computacional*, pp. 196–200, Bauru, 2008.
- [2] A. R. Backes, A. B. Bruno, M. N. B. Filho and O. M. Bruno. “Dimensão fractal aplicada em imagens de satélite de áreas urbanas”. In *III Workshop de Visão Computacional*, volume 1, pp. 52–57, São José do Rio Preto, 2007.
- [3] A. R. Backes, A. B. Bruno, M. N. B. Filho and O. M. Bruno. “Análise da complexidade de texturas em imagens urbanas utilizando dimensão fractal”. In *IX Simpósio Brasileiro de Geoinformática*, pp. 215–220, Campos do Jordão, 2007.
- [4] A. L. Barbieri, G. Arruda, F. A. Rodrigues, O. M. Bruno and L. da F. Costa. “An entropy-based approach to automatic image segmentation of satellite images”. eprint arXiv:0911.1759, [physics.comp-ph], 2009.
- [5] M. Pesaresi. “Textural Classification of Very High-Resolution Satellite Imagery; Empirical Estimation of the Interaction Between Window Size and Detection Accuracy in Urban Environment”. In *ICIP (1)*, pp. 114–118, 1999.
- [6] J. L. L. Ramos, R. A. R. Ramos, H. Santos and D. Fernandes. “Urban areas identification through clustering trials and the use of neural networks”. In *14TH European Colloquium on Theoretical and Quantitative Geography*, Tomar, Portugal, September 9-13 2005.
- [7] R. O. Duda, P. E. Hart and D. G. Stork. *Pattern Classification*. John Wiley & Son, 2001.
- [8] D. Casanova. “Identificação de espécies de vegetais através da análise de textura foliar”. Master’s thesis, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2008.
- [9] N. Otsu. “A threshold selection method from gray-level histograms”. *IEEE Trans. Sys., Man., Cyber*, vol. 9, pp. 62–66, 1979.