

MEDIÇÃO DE COBERTURA VEGETAL USANDO ANÁLISE COMPUTACIONAL DE IMAGENS

MEASUREMENT OF VEGETABLE COVER USING COMPUTER IMAGE ANALYSIS

Leandro Inácio Gonçalves¹
Luigi Francesco Mazzini Passerino²
Ricardo Pires³

Data de entrega dos originais à redação em: 12/12/2014
e recebido para diagramação em: 28/04/2015.

Atualmente, nas grandes cidades, dá-se importância crescente à presença de áreas verdes em proporção adequada. Para isso, pode ser necessário medir-se esta proporção periodicamente, de forma confiável e rápida. O objetivo deste trabalho é a medição automática da cobertura vegetal presente em áreas das quais se têm imagens aéreas, como as do serviço Google Maps. Para isso, foi usada a biblioteca computacional OpenCV e a classificação de trechos de imagens por meio de uma máquina de vetores de suporte. Os resultados obtidos foram satisfatórios.

Palavras-chave: Áreas verdes. Cobertura Vegetal. Medição. OpenCV. Máquina de Vetores de Suporte.

Currently, in large cities, growing importance is given to the presence of green areas in proper proportion. For this, it may be necessary to measure up this proportion periodically, in a reliable and fast way. The objective of this study is the automatic measurement of green areas present in areas of which aerial images are available, such as in the Google Maps service. To this goal, the OpenCV computational library was used along with a classification process carried by means of a support vector machine. The results obtained are satisfactory.

Keywords: Green Areas. Vegetable Cover. Measurement. OpenCV. Support Vector Machine.

1 INTRODUÇÃO

Atualmente, nas grandes cidades, devido a diversos fatores, dá-se importância crescente à presença de áreas verdes em proporção adequada (MEUNIER, 2014). Para isto, pode ser necessário medir-se esta proporção periodicamente, de forma confiável e automática.

Uma forma de medição de cobertura vegetal que vem sendo usada há décadas é a que se vale de dados registrados por satélites. Exemplos deste tipo são apresentados em (STOW; CHEN, 2002; NICHOL; WONG, 2007; CHEN; CHENG; WANG, 2013). Eles usam dados tais como a intensidade de radiação em frequências fora do espectro visível. Em particular, em (CHEN; CHENG; WANG, 2013), é usada a biblioteca OpenCV (OPENCV, 2014) para a manipulação de imagens no processo de medição, associada ao uso de um método estatístico.

Imagens aéreas se tornaram mais facilmente disponíveis com o advento do serviço Google Maps (GOOGLE, 2014), de livre acesso, o qual fornece imagens aéreas detalhadas de muitas cidades pelo mundo, incluindo cidades brasileiras. As imagens disponíveis livremente por este serviço não possuem todos os dados que seriam obtidos por um satélite, mas têm resolução adequada ao objetivo deste trabalho, uma vez que a resolução com que áreas de várias grandes cidades são disponibilizadas permite a identificação até de árvores isoladas.

Tendo-se em mãos uma imagem aérea, uma solução para a medição da cobertura vegetal presente na área correspondente seria uma avaliação subjetiva, feita por um ser humano. Este poderia quadricular a imagem,

contar os quadrados preenchidos predominantemente por vegetação e dividir este número pelo número total de quadrados, obtendo, assim, a proporção buscada. Evidentemente, para a medição da cobertura vegetal em uma grande cidade, usando-se imagens de boa resolução que cubram toda a área de interesse, a quantidade de trabalho seria enorme e demandaria um longo tempo. Justifica-se, assim, a busca por uma solução automatizada.

Uma maneira automática simples de se medir a cobertura vegetal numa imagem digital seria usando-se um programa computacional que lesse cada ponto (*pixel*) da imagem, contasse o número de pontos verdes e dividisse este número pelo número total de pontos da imagem. Mas, um ponto isolado identificado como sendo verde pode não fazer parte de vegetação. Ele pode fazer parte da pintura de um imóvel, de um carro ou de um outro objeto artificial qualquer. Por isto, o reconhecimento de uma área como sendo de vegetação deve se dar não pela avaliação de pontos isolados, mas avaliando-se conjuntos de pontos vizinhos. Um programa deve analisar, então, pequenas regiões da imagem e avaliar se o padrão de distribuição de tons de verde e de sombra em cada região corresponde àquele típico de uma área vegetal.

Um trabalho que usa este tipo de técnica, mas não para medição de vegetação, é (CHAPELLE; HAFFNER; VAPNIK, 1999). Seu objetivo é o de, por meio de um programa, classificar uma imagem a ele apresentada como pertencendo a uma dentre várias classes pré-estabelecidas, tais como "aviões" e "ursos". Para isto, é usada a biblioteca OpenCV e uma máquina de vetores de suporte, mais

1 Consultor nas áreas de Risco, Inteligência de negócios e Integração de dados, pelo SAS Brasil.

2 Graduando em Tecnologia em Sistemas Eletrônicos.

3 Professor na área de Eletrônica no Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

conhecida pela sua sigla em inglês, SVM, de *Support-Vector Machine* (STEINWART; CHRISTMANN, 2008). A SVM é uma ferramenta de classificação. Para ser usada, ela deve passar, inicialmente, por uma fase de treinamento, quando lhe são apresentados, pelo usuário, vários elementos pertinentes ao problema a ser resolvido e lhe é indicada a classe à qual pertence cada um destes elementos. Com as informações que lhe são fornecidas no treinamento, a SVM calcula uma fronteira matemática ótima entre aquelas classes. Esta fronteira pode ser uma reta, um plano ou um hiperplano, dependendo do número de dimensões envolvidas no problema. Finalmente, quando lhe é apresentado algum novo elemento (que não fazia parte do conjunto de treinamento), a SVM o localiza em relação à fronteira calculada na fase de treinamento e, com isto, classifica-o numa das classes existentes. Em (CHAPELLE; HAFFNER; VAPNIK, 1999), na fase de treinamento da SVM, são apresentadas imagens digitais pertencentes àquelas classes (“ursos”, “aviões” etc.) e é indicada pelo usuário a classe à qual pertence cada uma daquelas imagens. Cada imagem digital é lida pela OpenCV e, dela, é obtido um histograma contendo a distribuição dos pontos da imagem pelas cores possíveis, no padrão HSV de representação de cores (SHARMA, 2003). Apenas as duas primeiras componentes, H e S, das coordenadas HSV são usadas, porque a componente V é afetada apenas pela luminosidade. Cada histograma é fornecido como dado de entrada para a SVM, bem como a indicação da classe à qual pertence. Após o fornecimento dos histogramas de um conjunto de imagens de treinamento, a SVM passa a classificar corretamente a maioria das novas imagens que lhe são fornecidas. Assim, a SVM consegue associar cada classe de imagens a uma classe de padrões de histogramas de cores.

O objetivo do presente trabalho é a automação da medição da proporção de cobertura vegetal presente em áreas urbanas, a partir de imagens digitais como aquelas disponibilizadas no serviço Google Maps. É usada aqui a biblioteca OpenCV, bem como uma máquina de vetores de suporte e histogramas de cores, usando ideias como aquelas de (CHAPELLE; HAFFNER; VAPNIK, 1999). Espera-se que os histogramas de pequenas regiões da imagem possam revelar a presença ou não de vegetação em cada região, porque áreas com vegetação têm padrões característicos de distribuição de cores

(especialmente de tons de verde) e de luzes sombra, devidos à folhagem.

2 DESENVOLVIMENTO

O desenvolvimento prático se iniciou obtendo-se, no serviço Google Maps, uma imagem aérea do bairro da Vila Andrade, na cidade de São Paulo, uma imagem do bairro de Vila Bastos, em Santo André (SP) e uma imagem de área coberta totalmente por floresta na Amazônia, conforme as figuras de 1 a 7. Da imagem original da Vila Andrade (figura 2), foram geradas outras imagens (figura 1, figura 3, figura 4 e figura 5), por edição gráfica. Uma destas alterações consistiu na inserção de quadras poliesportivas (figura 1), pintadas de verde, copiadas de outros locais da mesma imagem e colocadas em lugar de áreas cobertas por vegetação na imagem original, para se verificar se o método a ser usado é capaz de distinguir um piso pintado de verde de área com vegetação. Outra das alterações, ao contrário, consistiu em se copiar uma área coberta por vegetação na imagem original de forma que ela cobrisse outras áreas que, originalmente, não eram vegetação (figura 3, figura 4 e figura 5). O uso de versões alteradas de uma mesma imagem, ao invés de se usarem imagens de locais diferentes, justifica-se pelo fato de que, tendo-se pequenas alterações controladas de cobertura vegetal de uma imagem para outra, a verificação do correto funcionamento do método aqui empregado fica facilitado. A figura 6, do bairro de Vila Bastos, no município de Santo André (SP), foi usada por se tratar de imagem de uma área muito urbanizada, com cobertura vegetal consideravelmente menor do que na Vila Andrade. Finalmente, a figura 7 é uma imagem de uma área totalmente coberta por floresta na Amazônia.



Figura 1 - Bairro de Vila Andrade, São Paulo, com 6 quadras poliesportivas inseridas



Figura 2 - Bairro de Vila Andrade, São Paulo, imagem original



Figura 3 - Bairro de Vila Andrade, São Paulo, 1 região verde inserida



Figura 4 - Bairro de Vila Andrade, São Paulo, 2 regiões verdes inseridas



Figura 5 - Bairro de Vila Andrade, São Paulo, 17 regiões verdes inseridas



Figura 6 - Bairro Vila Bastos, Santo André



Figura 7 - Área totalmente coberta por floresta, na Amazônia

Na seção 1, afirmou-se que uma pessoa poderia medir a cobertura vegetal em uma imagem quadriculando-a e contando os quadrados preenchidos por vegetação. Neste trabalho, este método é usado na geração dos resultados padrão. Espera-se que os resultados gerados automaticamente pelo método aqui desenvolvido fiquem próximos destes resultados padrão.

Aqui, o treinamento da SVM usa pequenos retângulos, com dimensões de 78x33 pontos, pertencentes às imagens obtidas. Foi montado um banco de dados contendo dezenas daqueles retângulos, amostrados de diferentes imagens. Este banco de dados contém retângulos com vegetação em diferentes condições de luminosidade e com diferentes tons de verde e diferentes padrões de sombra. Retângulos não correspondentes a vegetação têm cores e características mais variadas do que os anteriores. Alguns correspondem a asfalto, outros a telhados, a calçadas e a outros elementos artificiais da paisagem urbana, em diferentes condições de luminosidade.

Foi desenvolvido um programa, escrito na linguagem C++ e usando a biblioteca OpenCV. Ele lê cada retângulo e, como em (CHAPELLE; HAFFNER; VAPNIK, 1999), monta um histograma de cores do retângulo, usando apenas as componentes H e S do trio HSV de cada pixel. Foram feitos experimentos usando os seguintes números de faixas de valores de H e de S nos histogramas: 8, 12, 16 e 20. As faixas dos histogramas têm larguras uniformes.

Na fase de treinamento, os histogramas de retângulos com vegetação foram apresentados à SVM da OpenCV como pertencentes a uma classe rotulada como "+1". Os histogramas de retângulos sem vegetação foram apresentados à SVM como pertencentes a uma classe "-1".

Após o treinamento, na fase de classificação, o programa recebe uma imagem completa e a fragmenta em retângulos. O histograma de cada retângulo é montado e submetido à SVM, para ser classificado como +1 ou -1. Finalmente, o número de histogramas classificados como +1 é dividido pelo número total de retângulos da imagem, obtendo-se, assim, a proporção de vegetação da imagem.

3 RESULTADOS

Os resultados obtidos aplicando-se este método são apresentados na tabela 1:

Tabela 1 – Resultados da aplicação do método

Imagem	Número de faixas	Área verde calculada (%)	Padrão (%)	Diferença
Imagem1.jpg bairro Vila Andrade, 6 quadras inseridas	8	32.25%	19.50%	12.75%
	12	25.50%	19.50%	6.00%
	16	24.00%	19.50%	4.50%
	20	24.00%	19.50%	4.50%
Imagem2.jpg bairro Vila Andrade, original	8	32.50%	22.25%	10.25%
	12	26.00%	22.25%	3.75%
	16	24.50%	22.25%	2.25%
	20	25.00%	22.25%	2.75%
Imagem3.jpg bairro Vila Andrade, 1 região verde inserida	8	34.25%	22.75%	11.50%
	12	28.75%	22.75%	6.00%
	16	27.25%	22.75%	4.50%
	20	27.50%	22.75%	4.75%
Imagem4.jpg bairro Vila Andrade, 2 regiões verdes inseridas	8	36.50%	29.00%	7.50%
	12	31.75%	29.00%	2.75%
	16	29.50%	29.00%	0.50%
	20	30.25%	29.00%	1.25%
Imagem5.jpg bairro Vila Andrade, 17 regiões verdes inseridas	8	66.50%	70.75%	-4.25%
	12	65.50%	70.75%	-5.25%
	16	64.25%	70.75%	-6.50%
	20	64.50%	70.75%	-6.25%
Imagem6.jpg bairro Vila Bastos, Santo André	8	0.50%	1.75%	-1.25%
	12	0.00%	1.75%	-1.75%
	16	0.00%	1.75%	-1.75%
	20	0.25%	1.75%	-1.50%
Imagem7.jpg área totalmente coberta por floresta, na Amazônia	8	99.25%	99.75%	-0.50%
	12	99.75%	99.75%	0.00%
	16	99.75%	99.75%	0.00%
	20	99.75%	99.75%	0.00%

Nela, são apresentados resultados para sete imagens, com variados graus de cobertura vegetal. Cada imagem foi fragmentada em retângulos e foram calculados histogramas de cores, usando-se diferentes números de faixas para os valores de H e de S: 8, 12, 16 e 20.

A terceira coluna da tabela mostra o resultado da cobertura vegetal, calculado pelo programa aqui desenvolvido, como a proporção de retângulos que foram classificados como vegetação pela SVM.

A quarta coluna da tabela mostra o resultado padrão de cobertura vegetal, calculado subjetivamente, conforme descrito anteriormente.

A última coluna mostra, para cada caso, a diferença entre o resultado padrão e o resultado calculado automaticamente.

Verifica-se que, em quase todos os casos, a diferença entre o resultado padrão e o calculado pelo

programa foi de menos de 10%. Em geral, a medição automática gera resultados proporcionais à cobertura vegetal avaliada subjetivamente.

Os piores resultados ocorreram para casos em que se usaram 8 faixas nos histogramas. Provavelmente, isto se deve à resolução insuficiente destes histogramas,

o que pode acarretar numa classificação com variações superiores a 10%. Os casos em que se usaram 12, 16 ou 20 faixas nos histogramas, em geral, apresentaram resultados melhores do que os casos de 8 faixas, mas nenhum deles se mostrou sistematicamente superior aos demais.

4 CONCLUSÃO

Dos resultados obtidos, conclui-se que o método aqui aplicado funcionou a contento, especialmente se forem usados números de faixas a partir de 12 nos histogramas de cores. Desta forma, podem ser obtidos resultados razoavelmente precisos para a medição de cobertura vegetal, usando-se imagens aéreas disponíveis ao público pelo serviço Google Maps,

não se exigindo, portanto, dados especiais obtidos por satélites.

Como trabalhos futuros, pode-se avaliar o efeito nos resultados do uso de maior número de faixas nos histogramas (acima de 20 faixas). Também pode-se avaliar o efeito do uso de retângulos com dimensões diferentes das aqui usadas. E, ainda, pode-se buscar estender este método à identificação de outras características da área avaliada, tais como: área pavimentada, área coberta por água (rios, lagos, enchentes), classificação da vegetação pelo tipo (distinguindo tons de verde entre si) e variação temporal da área coberta por vegetação.

REFERÊNCIAS

MEUNIER, Isabelle. **Por que as Áreas Verdes são tão Importantes para uma Cidade?** Disponível em: <<http://www.ufrpe.br/>

artigo_ver.php?idConteudo=1259>. Acesso em: 10 de dezembro de 2014.

STOW, Douglas A.; CHEN, Dong Mei. Sensitivity of Multitemporal NOAA AVHRR Data of an Urbanizing Region to Land-Use/Land-Cover Changes and Misregistration. **Remote Sensing of Environment**, volume 80, número 2, páginas: 297–307, Elsevier, 2002.

NICHOL, J.; WONG, M. S. Remote Sensing of Urban Vegetation Life Form by Spectral Mixture Analysis of High-Resolution Ikonos Satellite Images. **International Journal on Remote Sensing**, volume 28, número 5, páginas: 985–1000, Taylor & Francis Online, Janeiro de 2007.

CHEN, Renxi; CHENG, Li; WANG, Chengfang. Vegetation Extraction Based on the Visual Characteristics of Plants from Remote Sensing Image. **Seventh International Conference on Image and**

Graphics (ICIG), 2013, páginas: 243–247, IEEE, Qingdao, Shandong, China, Julho de 2013.

OPENCV. **OpenCV** (Open Source Computer Vision). Disponível em: <<http://www.opencv.org>>. Acesso em: 10 de dezembro de 2014.

GOOGLE. **Google Maps**. Disponível em: <<http://www.google.com/maps>>. Acesso em: 10 de dezembro de 2014.

CHAPELLE, O.; HAFFNER, P.; VAPNIK, V.N. Support Vector Machines for Histogram-Based Image Classification. *Neural Networks*, **IEEE Transactions** on, 10(5):1055–1064, setembro de 1999.

STEINWART, Ingo; CHRISTMANN, Andreas. **Support Vector Machines**. Springer Publishing Company, Incorporated, 1st edition, 2008.

SHARMA, Gaurav. **Digital Color Imaging Handbook**. CRC Press, Boca Raton, Florida, Estados Unidos, 2003.