

Estudo comparativo de dois métodos para a contagem de árvores em mosaicos de fotografias obtidas por VANT

Dyeison Cesar Mlenek¹
Dr. Jorge Antonio Silva Centeno¹
Luís Fernando da Silva¹
Luis Guilherme Cosiaki¹

1Universidade Federal do Paraná - UFPR
Caixa Postal 19001 - 81531-970 - Curitiba - PR, Brasil
dyeison.mlenek@rumotopografia.com.br
centeno@ufpr.br
luis.silva@rumotopografia.com.br
luis.cosiaki@rumotopografia.com.br

Abstract. The problem of extracting features in digital images has been, for a long time, a very large challenge in various areas of engineering. For example, Kovalevskii (1967) used the dynamic programming algorithm to extract tipografados characters and manuscripts. The use of small format digital cameras aboard unmanned aerial vehicles UAV is a popular option for the mapping of tree plantations as a tool for forest management. This article describes a comparative study of two methods for counting trees in mosaics of photographs obtained by VANT. The first method is the well-known classification supervised and the second is a method based on segmentation and classification of segments after the image be transformed to a most advantageous color system that allows the derivation of an index of vegetation. The results show that both methods are capable of producing equivalent results, but the method of segmentation is faster and easy to apply. The results presented by both methods show that the methods used are not totally accurate, but rather tools for an estimate close to reality. In general the methods pose a good estimate for the solution of the problem, and can the undertaking in question employ both. With that concluded that the results are satisfactory.

Keywords: Remotes sensing, image classification, segmentation, forestry.

Resumo. O uso de câmaras digitais de pequeno formato a bordo de Veículos Aéreos Não Tripulados VANT é uma opção muito procurada para o mapeamento de plantações de árvores como ferramenta para o manejo florestal. Neste artigo é descrito um estudo comparativo de dois métodos para a contagem de árvores em mosaicos de fotografias obtidas por VANT. O primeiro método é a conhecida classificação supervisionada e o segundo é um método baseado na segmentação e classificação de segmentos após a imagem ser transformada para um sistema de cores mais vantajoso que permite a derivação de um índice de vegetação. Os resultados mostram que ambos métodos são capazes de produzir resultados equivalentes, porém o método de segmentação é mais rápido e de fácil aplicação. Os resultados apresentados pelos dois métodos mostram que os métodos utilizados não são totalmente precisos, mas sim ferramentas para uma estimativa próxima da realidade. Em geral, os métodos representam uma boa estimativa para a solução do problema, e pode a empresa em questão empregam ambos. Com isso concluir-se que os resultados são satisfatórios.

Palavras-chave: Sensoriamento remoto, segmentação, classificação supervisionada, manejo florestal.

1. Introdução

Um dos recentes avanços no sensoriamento remoto que tem introduzido grandes alterações na análise de imagens é o aumento da resolução espacial. Mais recentemente, este fator foi ainda mais aprofundado com a disponibilidade de plataformas tipo VANT (veículo aéreo não tripulado) que permitem obter imagens desde plataformas voando a baixa altura e com alta resolução espacial. A utilização de VANT's tem-se popularizado devido a seu baixo custo e à possibilidade de que estas plataformas podem ser empregadas por um vasto público.

Entre os setores mais interessados no uso de VANT's estão a agricultura e a silvicultura, pois estes setores se interessam em levantamentos atualizados de grandes campos de cultivo. Por exemplo, as empresas produtoras de papel ou madeira estão interessadas em monitorar o

estágio de crescimento de suas plantações de maneira a otimizar suas atividades de corte e prever o volume produzido.

Dentro desta realidade, a motivação para esta pesquisa é contribuir para o manejo florestal usando dados de VANT's. É assim apresentada uma metodologia para a identificação e contagem de árvores em uma plantação de reflorestamento com base em imagens VANT, considerando que no sul do Brasil existem grandes plantações para a indústria de papel e madeira. Estas indústrias estão interessadas em conhecer o volume total de madeira produzida, bem como a posição das árvores, para aplicar métodos de agricultura de precisão.

Mas este interesse não é local, pois outros países como, por exemplo, Canadá ou Finlândia também tem fortes indústrias baseadas no reflorestamento e a silvicultura. Tradicionalmente levantamentos expeditos e imagens de sensoriamento remoto orbital são usados para esta atividade. Visitas à campo são indesejadas pois elas envolvem o risco de acidentes. O uso de imagens orbitais de alta resolução esbarram em alguns problemas como a demora na aquisição da imagem, eventual cobertura de nuvens e alto custo, porém é uma alternativa que oferece a grande vantagem de dispor de imagens na região do infravermelho próximo do espectro.

O uso de VANT's transportando câmaras digitais surge como opção interessante, porém uma grande limitação é a carência de imagens no infravermelho quando se usam câmaras convencionais. Em alguns trabalhos científicos são mostradas alternativas para compensar esta limitação, como em Ohta et al (1980), Almeer (2012), Ponti (2013) ou Yang et al. (2015).

Neste artigo é descrito um estudo comparativo de dois métodos para a contagem de árvores em mosaicos de fotografias obtidas por VANT. O primeiro método é conhecido classificação supervisionada e o segundo é um método baseado na segmentação e classificação de segmentos.

2. Metodologia

Para obtenção do número de árvores na imagem, dois diferentes métodos foram empregados, sendo o primeiro o método da classificação supervisionada, que requer alta intervenção humana, e um segundo baseado na segmentação das árvores baseada nas diferenças de cor e tamanho. Os tópicos a seguir apresentam as etapas da metodologia empregadas para obtenção dos resultados.

2.1 Classificações supervisionada

O método da classificação supervisionada aplicado é o da Máxima Verossimilhança Gaussiana. Este método demanda participação do usuário, quem fornece as amostras de treinamento para o classificador. Esta tarefa consome tempo para a identificação das amostras.

O algoritmo de classificação é baseado na minimização do erro cometido através da lei e classificação mostrada na equação 1. O pixel pertence àquela classe que maximiza a probabilidade que associa esta classe ao pixel. Para um conjunto de classes finito:

$$x \in A \text{ se } p(A|x) > p(B|x) \text{ com } A \neq B \quad (1)$$

$$\text{Onde } p(A|x) = \frac{p(x|A) \cdot p(A)}{p(x)}$$

Sendo $p(x)$ igual para todas as classes, este termo pode ser desprezado durante a classificação. Da mesma maneira é comum assumir probabilidades *a priori* ($p(A)$) constantes. Finalmente, assumindo uma distribuição Gaussiana para "n" bandas espectrais, tem-se para cada classe a função densidade de probabilidade mostrada na Equação 2.

$$p(x|A) = \frac{1}{(2\pi)^{n/2} |\Sigma_A|^{1/2}} \exp\{(x - m_A) \Sigma_A^{-1} (x - m_A)^t\} \quad (2)$$

A lei de decisão depende das estatísticas das classes, vetor de médias m_A e a matriz variância-covariância Σ_A . Cabe ao usuário informar ao computador estas estatísticas identificando visualmente e demarcando com o mouse na tela regiões cobertas por cada tipo de cobertura ou alvo de interesse.

Após a modelagem da distribuição das classes no espaço multiespectral, o seguinte passo é aplicar o algoritmo de decisão a todos os pixels na imagem. Nesta fase de rotulação uma classe é adotada para cada pixel na imagem com base no seu vetor de leituras ou valores digitais (x).

O produto é uma imagem temática, no formato *raster*. Para as seguintes etapas do processo esta imagem foi transformada ao formato vetorial. Separou-se para isto os pixels da classe “árvores” e gerou-se uma imagem binária. Os polígonos extraídos foram a seguir analisados e contados.

Nesta fase foi possível descartar polígonos com área superior à média da área das copas das árvores da plantação. Cabe aqui ressaltar que, por se tratar de um reflorestamento, as árvores são muito parecidas, pois foram plantadas na mesma data e são indivíduos clonados. Cada polígono gerado como feição de interesse teve sua área calculada e pixels e depois transformadas a metros usando a resolução espacial média da imagem. Na imagem, o tamanho médio de uma copa de árvore era de 400 pixels, o que levou a fixar limites inferior e superior de 100 e 1000 pixels.

2.2 Segmentação

Como maneira alternativa de extrair as árvores visíveis na imagem considerou-se reduzir a intervenção humana aplicando métodos de segmentação. A experiência da seleção de amostras na imagem *RGB* mostrou a dificuldade que se enfrenta quando não se dispõe de imagens do infravermelho próximo. Existem dificuldades de separar a vegetação de árvores de solo e grama.

Para contornar esta dificuldade optou-se por utilizar um índice de cor que salientasse o contraste da vegetação usando as bandas do visível. Em um trabalho pioneiro, Ohta et al. (1980) realizaram um estudo de imagens reais obtidas com câmaras fotográficas (*RGB*). Com o intuito de entender as variações de cores para fins de segmentação, Ohta et al. (1980) aplicaram a transformação das componentes principais a estas diversas imagens e constataram que existe uma tendência para a orientação das componentes. Ou seja, os eixos das componentes principais, calculados a partir dos valores digitais nas bandas *RGB* tendem a ser paralelas. Com base nesta constatação, Ohta et al. (1980) propuseram um sistema alternativo, conforme mostra a equação 3,4 e 5.

$$I_1 = \frac{1}{3}(R + G + B) \quad (3)$$

$$I_2 = \left(\frac{-1}{4}R + \frac{1}{2}G - \frac{1}{4}B\right) \quad (4)$$

$$I_3 = \left(\frac{1}{2}R - \frac{1}{2}G\right) \quad (5)$$

Analisando a segunda componente deste sistema nota-se o cálculo do contraste entre a banda do verde e as outras duas. Simplificando esta componente chega-se à relação mostrada na equação 6.

$$I_2 = (2G - R - B) \quad (6)$$

Esta componente foi utilizada como entrada no processo de segmentação por limiarização (*thresholding*). O problema central neste caso é a escolha do limiar mais adequado. Este problema será melhor discutido na seção de resultados. A limiarização produz uma imagem binária segundo a regra mostrada na equação 7.

$$J > \text{limiar} = 255 \quad (7)$$
$$J < \text{limiar} = 0$$

A binarização da imagem resulta em agrupamentos de pixels com características similares em termos de valor digital. Porém, as regiões não estão devidamente identificadas. Por isso, o seguinte passo foi a rotulação das regiões aplicando o método de rotulação de componentes conexos (Haralick e Shapiro, 1991).

A rotulação de componentes conexos delimita regiões conexas, que potencialmente podem corresponder aos objetos de interesse, neste caso as árvores. As regiões obtidas variam no tamanho, pois vão desde pixels isolados a grandes agrupamentos de vegetação rasteira, passando pelas copas de árvore. Para separar as copas de árvore, foi utilizado o critério de área.

Na etapa final, realizou-se a comparação do sucesso destas duas abordagens. Para isto foram aplicados os coeficientes de qualidade propostos por Lee et al (2003). Como a classe de interesse é única, são consideradas três possibilidades no mapa temático resultante. A primeira corresponde aos elementos corretamente classificados (TP - *true positive*). Existem dois tipos de erro nesta classificação: Árvores não detectadas pelo algoritmo (*false negative* – FN) e outros objetos que não são árvores, mas que o classificador considerou como árvore (*false positive* – FP). Com tais valores, analisou-se a qualidade do resultado a partir da qualidade estatística.

As fórmulas utilizadas são mostradas nas equações 8 -10:

$$\text{Completeness} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{Completeness} = \frac{TP}{TP+FP} \quad (9)$$

$$\text{Quality} = \frac{TP}{TP+FN+FP} \quad (10)$$

3. Experimentos e Resultados

Os experimentos foram realizados usando dados de um levantamento de uma plantação de Eucaliptos no município de Telêmaco Borba em Paraná, Brasil. Para o levantamento fotogramétrico foi usado o sistema da *SenseFly (eBee)* e as imagens foram processadas para obter ortofotos e finalmente um mosaico. A câmara digital utilizada é a Canon S510. A altura de voo foi de 145m acima do terreno, em média, o que resultou em pixel no terreno de 5,5cm.

As árvores de eucalipto são uma espécie clonada, com altura média de um metro acima do terreno e copa em torno de 70 a 120cm. Na figura 1 é mostrado um recorte deste mosaico. À esquerda é apresentado o recorte contendo várias árvores, solo, troncos caídos e vegetação rasteira. À direita é mostrada uma ampliação da região marcada no primeiro recorte, mostrando com maior detalhe as características das árvores em termos espaciais e de cor.

Para medir o sucesso de cada método e poder utilizar as métricas de qualidade, foram inicialmente contadas visualmente todas as árvores presentes no recorte selecionado. Estas árvores foram admitidas como verdade de campo, sendo os métodos propostos comparados com esses resultados. Ao Total foram admitidas 301 árvores presentes no recorte.

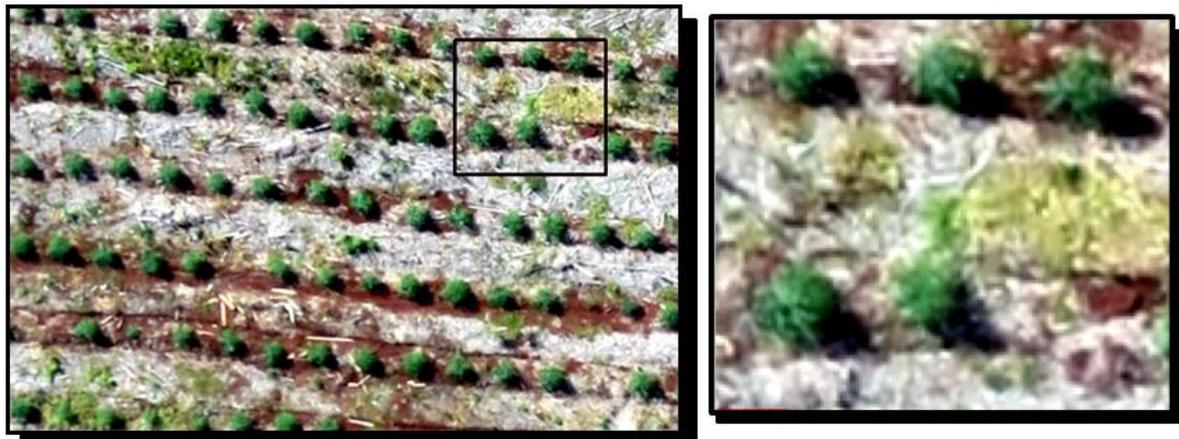


Figura 1. Recorte dos dados de entrada. (a) recorte. (b) ampliação do recorte.

Na figura 2 são apresentadas três imagens. Na primeira (a) são mostradas as árvores identificadas visualmente. Na segunda imagem é mostrado o resultado da classificação supervisionada pela Máxima Verossimilhança. Na última imagem é apresentado o resultado da segmentação.

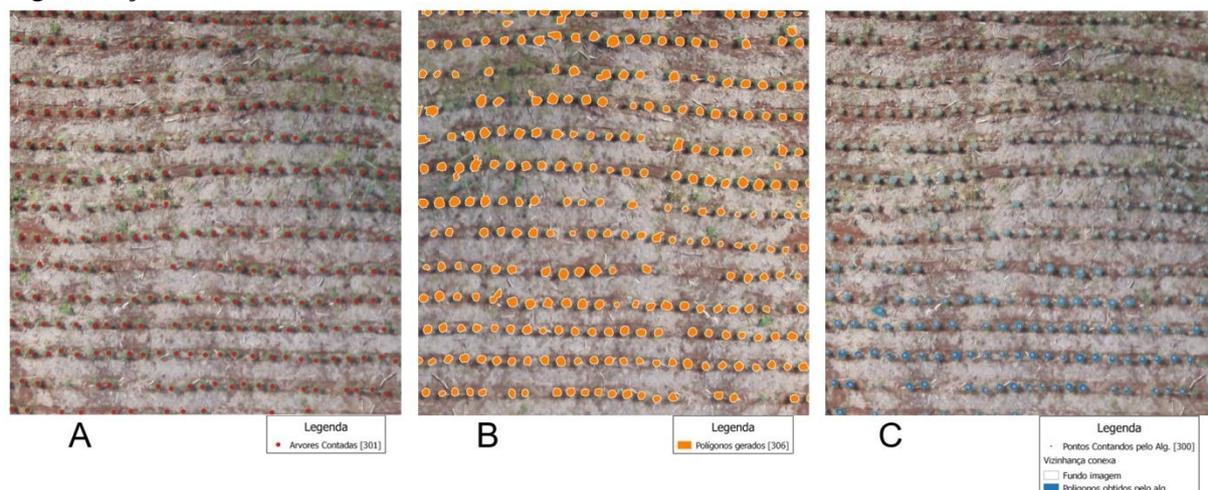


Figura 2. Comparativo entre os métodos aplicados. (A) verdade de campo. (B) Máxima Verossimilhança. (C) Segmentação por Ohta et al (1980).

Com o método de classificação supervisionada foram identificados 306 objetos, enquanto com a segmentação 300. Consta-se que ambos os métodos chegaram próximos do valor verdadeiro, porém o método da segmentação é melhor. Nota-se também que o método de classificação supervisionada superestimou a quantidade de árvores, provavelmente pela confusão de copas de árvores com outros tipos de vegetação.

O resultado da classificação pelo método da Máxima Verossimilhança foi também analisado em termos de acurácia aplicando os índices de qualidade propostos por Lee et al (2003). Os resultados são: *Completeness*=98,36; *Correctness*=96,78 e *Quality*=95,25.

Embora a proposta original fosse reduzir a participação humana no processo, isto não foi totalmente possível para atingir os resultados mostrados acima. Existem pontos no processo onde a participação humana é necessária e fundamental para o sucesso do processo. A primeira situação diz respeito ao limiar utilizado para binarizar a imagem. Dada a imagem calculada (I_2) pela equação 7, alguns valores de limiar foram estudadas. A figura 3 a seguir representa os diferentes valores utilizados.

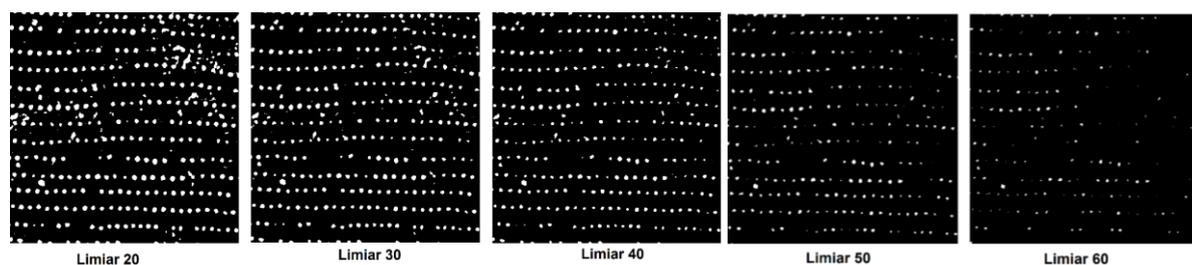


Figura 3. Limiares estudados.

Nota-se que quanto mais baixo o limiar utilizado, mais ruídos a imagem fica, consequentemente mais feições não verdadeiras são empregadas na contagem final das árvores, elevando o número de falsos negativos. Da mesma maneira quanto mais alto o limiar, menos feições verdadeiras são classificadas, elevando o número de falsos positivos. Ao analisar as imagens, verifica-se que duas opções são mais condizentes com a realidade, o Limiar de 50 e 40. Logo as duas imagens geradas foram objetos de análises.

A segunda variável a ser fornecida pelo usuário é a área das feições. Dentro dessa variável, duas hipóteses foram levantadas, um como a imagem do limiar 40 e outra com o limiar de 50. A partir das imagens binarizadas, foram executadas uma série de tentativas com diferentes áreas. A tabela a seguir, apresenta essas tentativas, onde um campo é referente a área da feição e outro é a contagem retornada pelo algoritmo.

Tabela 1. Limiares estudados com as respectivas áreas geradas.

Limiar 40		Limiar 50	
Área mínima	Contagem	Área mínima	Contagem
50	297	20	259
45	300	15	267
40	304	10	272
30	311	5	284
25	319	4	286
20	320	3	289
15	325	2	294
10	330	1	320
5	332	0,5	320
1	349	0,01	320

A série com valores de limiar 50 tende à 320 árvores à medida que a área das feições diminui, pressupondo-se que 320 é número máximo de feições detectadas pelo algoritmo. Já com o limiar 40, a mesma área 1 apresenta um número superior de contagem do que a série de limiar 50. A figura 4 representa os valores obtidos. Analisando os resultados encontrados das duas séries com o valor absoluto (301) de árvores, verifica-se que duas opções são condizentes com a realidade. Sendo uma hipótese com limiar 40 e área 45, e outra com limiar 50 e área 2.

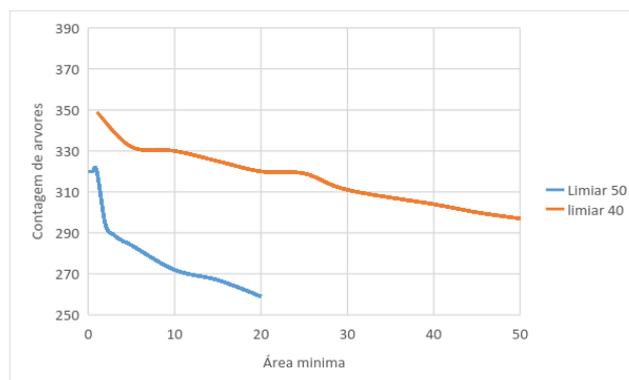


Figura 4 - Limiares estudados com as respectivas áreas geradas.

A partir das hipóteses levantadas, os valores para os falsos negativos e falsos positivos foram contados. Para o limiar de 50 estes são: FP=50 e FN=25. Para o limiar de 40 tem-se FP=21; FN=20.

Para o limiar 50, onde aumentou o valor espectral das feições e diminuiu a área, os erros encontrados foram maiores em virtude dos objetos menores classificados como árvores, e também pela segmentação de algumas árvores, sendo que nesse caso dois pontos foram criados para uma só árvores. A combinação de limiar 40 e área 45 apresentou resultados melhores. A partir da contagem os valores da qualidade foram calculados para as duas hipóteses. Nota-se que os melhores resultados foram obtidos com limiar 40 e área 45.

Tabela 2. Qualidade dos resultados gerados por vizinhança conexa.

	Limiar 40 área 45	Limiar 50 área 2
Completeness	93,47 %	90,3 %
Correctness	93,76 %	91,76%
quality	88,01%	83,84%

Ao comparar os resultados dos métodos (tabela 2), nota-se que o método de classificação apresenta resultados mais precisos e mais próximos da realidade. Porém o método de classificação demanda maior tempo e trabalho, sendo que todo processo para análise de 1 fotografia demorou em torno de 30 minutos. Outro agravante ao método é quantidade de softwares utilizados para gerar os resultados, com isso o custo de implantação se torna maior.

O método de segmentação apresentou ganhos quanto ao custo de implantação, sendo que apenas um software livre é aplicado, além do tempo de processamento, que é muito curto.

4. Conclusão

Os resultados apresentados por ambos os métodos demonstram que os métodos utilizados não são totalmente precisos, e sim ferramentas para uma estimativa próxima da realidade.

Os resultados encontrados com o método de classificação apresentam valores mais próximos dos verdadeiros, porém esbarram no demorado e individual processo para geração de apenas uma fotografia.

Com o algoritmo de vizinhança conexas, as estimativas são inferiores aos valores da classificação, porém há um ganho considerando se analisarmos o tempo de resposta entre ambos os métodos. Outro critério a se analisar é o custo de implantação, sendo o algoritmo mais barato em virtude do software utilizado e do tempo de resposta.

Os resultados encontrados com o algoritmo podem ser melhorados caso se ache um equilíbrio perfeito entre o limiar para a binarização e a área das zonas detectadas pelo algoritmo. Com as análises apresentadas nota-se que há necessidade de certos cuidados ao verificar os

resultados encontrados, sendo que alguns resultados apresentam contagem muito próxima da realidade, porém tais valores são julgados imprecisos se analisarmos a qualidade pela matriz de confusão. A análise por qualidade, observando-se os valores falsos positivos e negativos é fundamental para um correto parecer sobre os métodos.

Em geral os métodos apresentam uma boa estimativa para a solução do problema, podendo ambos serem empregados pela empresa interessada. Com isso conclui-se que os resultados encontrados são satisfatórios.

Agradecimentos

Os autores agradecem à KLABIN S.A. pela disponibilidade dos dados usados nesta pesquisa.

Referências Bibliográficas

Almeer, M. H. (2012): **Vegetation Extraction from Free Google Earth Images of Deserts Using a Robust BPNN Approach in HSV Space**, in: International Journal of Advanced Research in Computer and Communication Engineering, Vol. 1, Issue 3, 134–140.

Haralick, R.M. e Shapiro, L.G., **Computer and robot vision**. Addison-Wesley Longman Publishing Co., Inc. 1991.

Lee, D.S., Shan, J., Bethel, J.S., 2003. **Class-guided building extraction from IKONOS imagery**. Photogrammetric Engineering and Remote Sensing 69 (2), 143–150.

Ohta, Y-I. , Kanade, T. e Sakai, T., **Color information for region segmentation**. Computer Graphics and Image Processing 13, 222-241 (1980).

Ponti, M. P. (2013): **Segmentation of Low-Cost Remote Sensing Images Combining Vegetation Indices and Mean Shift**, in: IEEE Geoscience and Remote Sensing Letters, Vol. 10, Issue 1, 67–70.

Yang, W., Wang, S., Zhao, X., Zhang, J. e Feng, J., (2015): **Greenness identification based on HSV decision tree**, in: INFORMATION PROCESSING IN AGRICULTURE 2.