

# VI CONGRESSO IBÉRICO de Agro-Engenharia

5 a 7 de Setembro | 2011  
Universidade de Évora | Portugal



## Detecção de vegetação arbórea através de segmentação e classificação orientada a objecto de imagens multiespectrais de alta resolução (Quickbird).

A.M.O. Sousa<sup>a1</sup>, P. Mesquita<sup>2</sup>, A.C. Gonçalves<sup>a3</sup>, J.R. Marques da Silva<sup>a4</sup>, L.L. Silva<sup>a5</sup>

<sup>a</sup>Universidade de Évora, Departamento de Engenharia Rural, Escola de Ciências e Tecnologia, Apartado 94, 7002-554 Évora

<sup>b</sup>Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Grupo Água, Solo e Clima Apartado 94, 7002-554 Évora, Portugal,

asousa@uevora.pt<sup>1</sup>, paulomesquita00@gmail.com<sup>2</sup>, acag@uevora.pt<sup>3</sup>, jmsilva@uevora.pt<sup>4</sup>, llsilva@uevora.pt<sup>5</sup>

### Resumo

O principal objectivo deste estudo foi avaliar o potencial de imagens de alta resolução espacial (Quickbird) utilizando o método de segmentação multi-resolução e classificação digital orientada a objecto na 1) identificação de manchas por espécie florestal obtendo mapas de ocupação do solo e 2) identificação de área de coberto arbóreo por espécie para a região de Portel (Alentejo). Procedeu-se ao cálculo de bandas sintéticas tais como o índice de vegetação NDVI e as compentes principais para aumentar a capacidade de separação entre as classes no processo de segmentação e classificação. Ambos os resultados foram validados através do calculo da matriz de confusão e do coeficiente *Kappa*. Os resultados foram bastantes satisfatórios, para o mapa de ocupação do solo, obtendo-se um *Kappa* de 82,5% e para o mapa de área de coberto arbóreo de 80% e ambos com precisão global acima de 88%. Podemos concluir que a metodologia desenvolvida e os dados de alta resolução do satélite Quickbird produzem resultados satisfatórios para a identificação e separação das espécies florestais. A baixa resolução espectral deste satélite é compensada pela elevada resolução espacial e pela utilização de outros parâmetros de forma e de textura dos objectos individuais e entre objectos.

**Palavras-chave:** Detecção remota, espécie florestal, ocupação do solo, segmentação multi-resolução, classificação orientada a objecto.

### Abstract

The main objective of this study was to evaluate the potential of high spatial resolution images (Quickbird) using the method of segmentation multi-resolution and object-oriented classification 1) identification of forest areas by species obtaining maps of land-cover and 2) identification of area of tree cover area in the region of Portel (Alentejo). Proceeded to the calculation of synthetic bands, such as the vegetation index NDVI and the principal component analysis, to increase the capacity of separation between classes in the process of segmentation and classification. Both results were validated by calculating the confusion matrix and Kappa coefficient. The results were satisfactory enough, for the map of land-cover was obtained a kappa of 82.5% and to tree cover map was 80%, for both the overall accuracy are above 88%. We can conclude that the methodology and data from high spatial resolution data from Quickbird satellite produce satisfactory results for the identification and separation of forest species. The low spectral resolution of this satellite is compensated by the high spatial resolution and by the use of others parameters, like form and texture of the individual objects and between objects.

**Keywords:** Remote sensing, forest types, landuse, multi-resolution segmentation; object-oriented classification

## 1. INTRODUÇÃO

A floresta em Portugal ocupa 37,5% do território nacional, apresentando diferentes taxas de arborização nas várias regiões do País. A distribuição segundo as principais espécies, indica o pinheiro bravo (*Pinus pinaster*) como a espécie florestal predominante, com 27,2% da ocupação, equivalente a 885 mil hectares. Segue-se os Eucaliptos (*Eucalyptus globulus*) com 22,7%, o que corresponde a cerca de 739 mil hectares de floresta e o Sobreiro (*Quercus suber*) com 22%, o que corresponde perto de 716 mil hectares (IFN5, 2010). O sector florestal e a sua gestão é de elevada importância à escala mundial. Pela relevância das funções económicas, ambientais, sociais e culturais a ela associadas.

As imagens de satélite de resolução espacial moderada como as imagens Landsat tem sido utilizadas para produção de cartografia florestal, no entanto nos últimos anos as imagens de alta resolução espacial vieram trazer outro tipo de detalhe a estes estudos para escalas regionais. A obtenção de mapas espaciais locais com elevada precisão só é possível com o surgimento das imagens de muito alta resolução espacial (Aplin, 2003). Espera-se que estas imagens sejam um desafio às fotografias aéreas e aos dados recolhidos por sensores aerotransportados, tendo em conta as suas características geométricas, operacionais e de custos. Novas questões tem sido colocadas quanto às metodologias a utilizar na classificação destes novos dados.

O método de segmentação e classificação orientada a objecto, através do processo de segmentação, permite utilizar como unidade de classificação não o pixel individual, mas sim um conjunto de pixels, tornando assim possível o incremento de mais variáveis para o processo de classificação. Para além da informação espectral, é possível considerar outros parâmetros bem presentes nas imagens de alta resolução espacial, tais como a forma, textura e relação entre objectos vizinhos.

O objectivo deste estudo é avaliar o potencial de imagens de alta resolução espacial (Quickbird) utilizando o método de segmentação multi-resolução e classificação digital orientada a objecto na 1) identificação de manchas por espécie florestal obtendo cartas de ocupação do solo e 2) identificação de área de coberto arbórea para a região de Portel.

## 2. MATERIAL E MÉTODOS

Foi utilizada imagem de alta resolução espacial do satélite Quickbird da *Digital Globe*. As imagens foram adquiridas no modo “*Ortho-ready Pan-Sharpned*” georreferenciada no sistema de coordenadas geográficas WGS84. A imagem adquirida cobre uma área de aproximadamente 225 Km<sup>2</sup>, correspondendo à fusão da banda pancromática com as 4 bandas espectrais, b1-azul (0,45-0,52µm), b2-verde (0,52-0,60µm), b3-vermelho (V) (0,63-0,69µm) e b4- infravermelho próximo (IVP) (0,76-0,90µm), resultando as 4 bandas com uma resolução espacial de 0,70 metros e resolução radiométrica de 11bits. A área de estudo tem uma dimensão aproximada de 1,2 km por 1,2 km, próximo da Vila de Portel, distrito de Évora, pequena área cortada da imagem adquirida.

O pré-processamento da imagem foi realizado no *software* comercial ENVI 4.8 (ITT), este compreendeu a correcção geométrica da imagem através de orto-rectificação e georreferenciação. O processo de orto-rectificação foi realizado através do Modelo Digital de Elevação (MDE) do satélite *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER) com 30 m de resolução espacial, após este passo a imagem foi georeferenciada para o sistema de projecção Hayford Gauss Militar através de pontos com

coordenadas conhecidas. Após a correcção geométrica a imagem ficou com erro global de 0,6 m. Em seguida procedeu-se ao cálculo de bandas sintéticas, que entram no processamento da imagem juntamente com as bandas originais. As banda sintéticas calculadas foram o índice de vegetação *Normalized Difference Vegetation Index* (NDVI) (Huete *et al.*, 2002), calculado de acordo com a equação (1) e a primeira componente resultante do método de Análise de Componentes Principais (ACP) (Lillesand *et al.*, 2004) realizada com base nas quatro bandas originais.

$$NDVI = \frac{IVP-V}{IVP+V} \quad (1)$$

A metodologia utilizada na obtenção do mapa de coberto arbóreo com base na imagem Quickbird divide-se em duas etapas. A primeira, classifica o ocupação do solo de acordo com as classes de ocupação existentes e a segunda extrai o coberto arbóreo para cada uma das classes de interesse resultantes da primeira etapa.

A obtenção do mapa de ocupação do solo baseou-se fundamentalmente no uso de filtros de médias aplicados sobre as bandas (NDVI, IVP, ACP), segmentação da imagem em objectos e classificação dos mesmos através do método do vizinho mais próximo (Mallinis *et al.*, 2008). A aplicação de filtros às bandas referidas, teve como objectivo diminuir pequenas diferenças espectrais locais por forma a permitir a identificação de objectos de maiores dimensões mais homogéneos e assim o isolamento de áreas com a mesma ocupação.

O processo entre a segmentação e a classificação foi iterativo, onde em cada etapa foram isoladas as classes melhor delimitadas. Após a obtenção do mapa de ocupação do solo, constituído pelas classes Vegetação de zonas húmidas, Eucaliptal, Montado, Plantações jovens e Água, foi efectuada a extração da área de coberto arbóreo para as classes Eucaliptal, Montado e Plantações jovens através da aplicação de limiares à banda correspondente à primeira componente principal. Às restantes classes, não foram aplicados quaisquer procedimentos, por não enquadrarem o objectivo de cartografar o coberto arbóreo para vegetação de zonas húmidas.

A obtenção dos mapas finais deu-se através de segmentação da imagem em objectos e respectiva classificação em classes de ocupação do solo dos mesmos, toda esta fase do trabalho foi realizada através do *software* comercial *Definiens Developer 8*.

A imagem foi classificada em dois níveis, um de menor escala aplicado a toda a área de estudo onde foram identificadas as unidades espaciais existentes com base nas suas características espectrais, obtendo-se o mapa de ocupação do solo. No segundo nível, de maior escala, para cada uma das áreas extraídas no primeiro nível foi realizada a identificação e extracção de coberto arbóreo, ou seja, obteve-se duas classes: i) coberto arbóreo para cada classe e ii) outra classe designada neste trabalho por solo que inclui alguma vegetação arbustiva e solo. A Tabela 1 ilustra as classes consideradas para cada nível.

**Tabela 1.** Classes de ocupação do solo definidas em cada nível de segmentação.

Nível 1	Nível 2
Vegetação de zonas húmidas	Vegetação de zonas húmidas
Montado	Sobreiro + Azinheira
	Solo
Eucaliptal	Eucaliptal
	Solo
Plantações jovens	Plantações jovens
	Solo

O processo de segmentação divide a imagem em regiões homogéneas constituídas por vários pixels com base em parâmetros definidos pelo utilizador (Castillejo-González, *et al.*, 2009), este processo pode ser realizado através de vários algoritmos distintos. Neste estudo a segmentação da imagem foi realizada através de dois algoritmos presentes no *software* Definiens Devoleper 8. O primeiro algoritmo é a segmentação multi-resolução, que foi utilizado para segmentar as diferentes unidades de ocupação do solo indicadas no nível 1. O segundo, baseado na segmentação por diferença de contraste (*Contrast Split Segmentation - CSS*) que foi utilizado na expansão da classe Vegetação de zonas húmidas por forma a incluir nesta classe todas as áreas fortemente afectadas pela presença de água, incluindo vegetação erbácia, e na delimitação do coberto arbóreo realizada no segundo nível. A segmentação multi-resolução agrega pixels ou objectos com base em critérios de homogeneidade, que são uma combinação de critérios espectrais e de forma (Definiens, 2010). A definição destes critérios foi realizada de forma empírica até que os objectos resultantes sejam concordantes com as classes pretendidas. O algoritmo CSS divide a imagem em objectos escuros e claros baseado num limiar que maximiza o contraste entre objectos resultantes (Definiens, 2010), este limiar óptimo é detectado de forma automática a partir de limites definidos pelo utilizador para uma dada banda. Uma das características deste algoritmo é a possibilidade de classificar de forma imediata os objectos resultantes. Esta característica faz com que o seu uso seja adequado no âmbito deste trabalho, a sua utilização na extracção da área de coberto arbóreo foi realizada através deste algoritmo dado que o contraste entre esta e solo/vegetação arbustiva é bastante elevado no índice de vegetação NDVI.

O processo de classificação foi realizado com recurso a áreas de treino e através do método de classificação do vizinho mais próximo. Este método é não-paramétrico e baseia-se na distância euclidiana da observação a ser classificada à área de treino mais próxima (Hubert-Moy *et al.*, 2001). As variáveis utilizadas no processo de classificação foram seleccionadas de entre variáveis estatísticas calculadas para cada uma das bandas processadas tais como: médias, desvios-padrão e enviesamento para cada objecto. O processo de selecção das variáveis a utilizar baseou-se em observações do espaço bidimensional entre variáveis e na maximização da distância de separação entre classes.

A exactidão de ambas as classificações foi realizada com base em pontos aleatórios sobre a área de estudo. Para a validação do primeiro mapa, com as classes de ocupação do solo, foram considerados 60 pontos e para o segundo mapa, com o coberto arbóreo, foram considerados 120 pontos onde apenas foi tido em conta a presença ou não de vegetação. Para cada ponto foi atribuída a classe real, existente no terreno, por interpretação visual. Esta informação foi comparada com os resultados de ambas as classificações através de uma matriz de confusão e do coeficiente *Kappa*.

A matriz de confusão (Congalton *et al.*, 1983; Stehman, 1997) apresenta o número de pixels classificados correctamente em oposição ao número de pixels previstos para cada classe no decorrer da classificação. Esta matriz permite obter informação da precisão global do respectivo classificador, bem como a precisão de classificação de cada classe, através do cálculo do coeficiente *Kappa*, dado pela seguinte equação:

$$Kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})} \quad (2)$$

onde  $N$  é o valor total de observações incluídas na matriz,  $r$  o número de linhas da matriz de confusão,  $x_{ii}$  o número de observações na linha e coluna  $i$ ,  $n_{i+}$  número de observações na linha  $i$  e  $x_{+i}$  o número total de observações na coluna  $i$  (Fonseca e Fernandes, 2004). Este

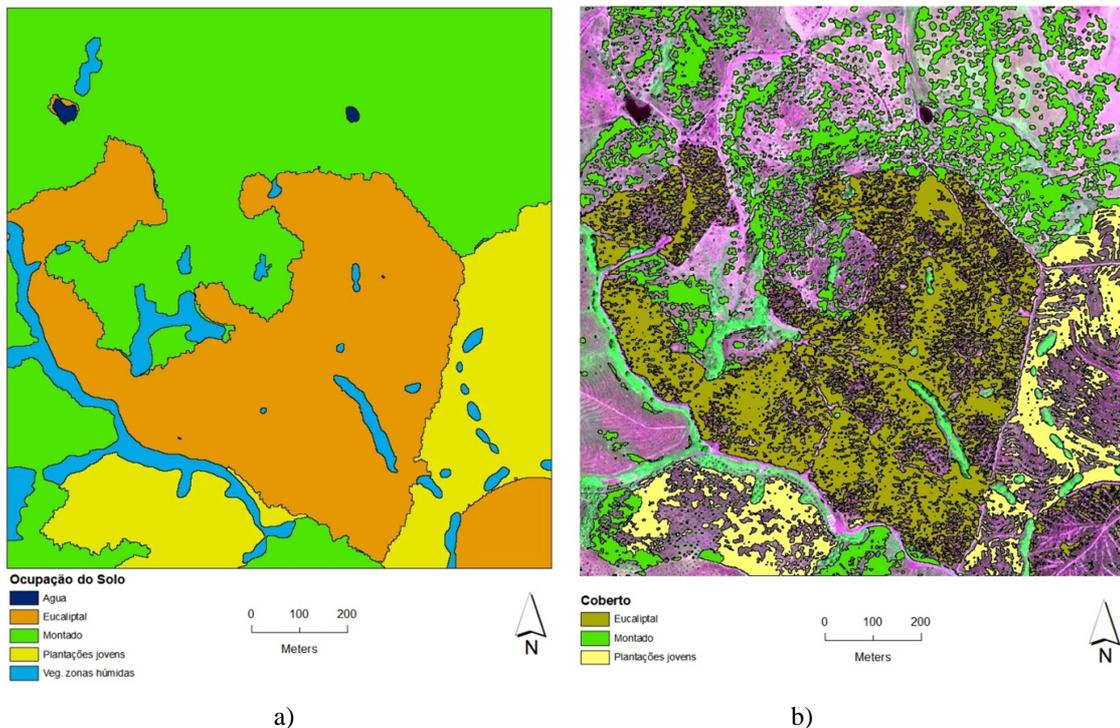
coeficiente é uma estatística que indica o grau de concordância entre o resultado da classificação digital e a realidade no campo.

Foram calculados os erros de comissão e de omissão para se perceber a dificuldade de identificação das classes de ocupação do solo, bem como da área de coberto arbóreo para a classificação orientada a objecto com imagem de alta resolução espacial.

### 3. RESULTADOS E DISCUSSÃO

O processo de segmentação multi-resolução com imagens de alta resolução espacial permite trabalhar a diferentes escalas, onde com a menor escala é possível identificar padrões espectrais e espaciais por forma a isolar as espécies e com a grande escala a identificação do coberto arbóreo dentro de cada área por espécie.

Na Fig. 1 a) pode ser observado o mapa com a delimitação e identificação das classes de ocupação do solo e na Fig. 1 b) o mapa com o coberto arbóreo para cada classe identificada na Fig. 1 a).



**Figura 1. a)** Mapa de ocupação do solo, **b)** mapa de coberto arbóreo sobre uma composição colorida em falsa cor (RGB-b3b4b2).

Calcularam-se as matrizes de confusão para ambos os mapas, foram utilizados 60 pontos de validação para a avaliar a precisão do mapa de ocupação do solo e 117 (foram eliminados 3 pontos) para o mapa com o coberto arbóreo (Tabela 2 e 3).

*Tabela 2. Matriz de confusão para o mapa de ocupação do solo.*

	<b>Eucaliptal</b>	<b>Montado</b>	<b>Plantações jovens</b>	<b>Veg. Zonas húmidas</b>	<b>Total</b>	<b>Erro de Comissão (%)</b>
<b>Eucaliptal</b>	23	0	0	1	24	4,17
<b>Montado</b>	2	20	0	4	26	23,08
<b>Plantações jovens</b>	0	0	8	0	8	0,00
<b>Veg. zonas húmidas</b>	0	0	0	2	2	
<b>Total</b>	25	20	8	7	60	
<b>Erro de Omissão (%)</b>	8,00	0,00	0,00	71,43		
<b>Kappa = 82,5%</b>						
<b>Precisão global = 88,3%</b>						

*Tabela 3. Matriz de confusão para a área de coberto arbóreo.*

	<b>Coberto arbóreo</b>	<b>Solo</b>	<b>Total</b>	<b>Erro de Comissão (%)</b>
<b>Coberto arbóreo</b>	50	9	59	15,30
<b>Solo</b>	3	55	58	5,20
<b>Total</b>	53	64	117	
<b>Erro de Omissão (%)</b>	5,70	14,10		
<b>Kappa = 80%</b>				
<b>Precisão global = 89,7%</b>				

Para ambas os objectivos deste trabalho obteve-se resultados bastante satisfatórios com a metodologia aplicada e com as imagens de alta resolução espacial. Com o valor do coeficiente *Kappa* para o mapa de ocupação do solo de 82,5% e para o coberto arbóreo de 80%. Os resultados confirmam-se com o cálculo da precisão global que apresenta valores elevados de concordância com a realidade no terreno. Podemos assim dizer que a detecção, delineamento e identificação automática das classes de ocupação do solo com dados de detecção remota torna-se cada vez mais importante no sector florestal. Estes dados podem tornar-se uma mais valia na realização de inventário florestal, gestão de povoamentos florestais e contribuição de informação precisa para a estimativa de biomassa.

Analisando os erros de omissão e comissão para o mapa de ocupação do solo, verifica-se um elevado erro de omissão (71,43%) para a classe de Vegetação em zonas húmidas. Este erro possivelmente deve-se à presença de pequenas áreas de vegetação erbácea, que ficou classificada como Montado, que por não apresentar um sinal espectral de vegetação vigorosa tão intenso como as restantes áreas da classe de vegetação em zonas húmidas, não foi identificada e incluída dessa classe. Outro factor que influencia este erro é a data da imagem (Bunting e Lucas, 2006), mês de Maio. Neste mês a vegetação erbácea ainda não se apresenta seca, se a imagem fôr adquirida na época seca provavelmente esta dificuldade de segmentação e classificação diminui. Pelas mesmas razões pode-se analisar um erro de comissão de 23,08% para a classe de Montado.

Analisando a matriz de confusão para o mapa de coberto arbóreo não se encontram erros de omissão e comissão elevados, no entanto ha uma pequena confusão entre o que é arbóreo e solo, que se deve há dificuldade de atribuir a classe real aos pontos quando estes estão na bordadura das copas.

Para confirmar estes resultados e avaliar os métodos utilizados com estes dados, pretendemos aumentar a área de estudo e aumentar o número de pontos utilizados para a validação.

## 4. CONCLUSÕES

Neste estudo o método de segmentação multi-resolução e classificação orientada a objecto com imagens de alta resolução produz resultados bastante satisfatórios na identificação das espécies florestais presentes na área de estudo. Utilizar como elemento de classificação o objecto em vez do pixel fornece mais informação espectral, espacial e de textura na associação dos elementos às classes, no entanto lança o desafio de como usar esta informação de forma mais eficiente (Mallinis *et al.*, 2008).

Este estudo demonstra que a segmentação multi-resolução é capaz de delinear o padrão dependente da escala, da heterogeneidade e da estrutura da floresta. Esta ferramenta é também um potencial para melhorar o delineamento de inventário caracterizando a complexidade da estrutura da floresta através de escalas múltiplas espaciais e temporais (Lamonca *et al.*, 2008).

A detecção, delineamento e identificação automática do coberto arbóreo com dados de detecção remota de alta resolução espacial torna-se cada vez mais importante no sector florestal. Estes dados podem contribuir com a área coberta e a distribuição espacial de espécies florestais de forma mais rápida e com menor custo, podendo torna-se uma mais valia na realização de inventário florestal, gestão de povoamentos florestais e estimativa de biomassa.

No entanto as imagens Quickbird, apesar da sua elevada resolução espacial, poderão apresentar dificuldades na identificação e separação entre algumas espécies florestais devido à sua limitada resolução espectral, principalmente pela ausência de uma banda de infravermelho médio.

## 5. AGRADECIMENTOS

Agradecemos ao Engenheiro João Roma e aos Técnicos do Departamento de Engenharia Rural da Universidade de Évora, Srº Custódio Alves e Srº José Condeças pela ajuda na realização de levantamento de pontos de controlo para a georreferenciação da imagem bem como pelo apoio na realização do trabalho de campo.

## 6. BIBLIOGRAFIA

- Aplin, P. (2003) Remote Sensing: base mapping. Progress in Physical Geography, Vol.27 (2), 275-283.
- Bunting, P., Lucas, R. (2006) The delineation of tree crowns in Australian mixed species forests using hyperspectral compact airborne spectrographic imager (CASI) data. Remote Sensing of Environment, Vol. 101, 230-248.
- Castillejo-González, I. L., López-Granados, F., Garcia-Ferrer, A., Peña-Barragán, M., Jurado-Expósito, M., Orden, M. S., González-Audicana, M. (2009) Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using Quickbird imagery. Computers and Electronics in Agriculture, Vol. 68, 207-215.
- Congalton, R.G. Oderwald R. G., Mead, R. A. (1984) Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. Photogrammetric Engineering and remote sensing, Vol. 49, 1671-1678.
- Definiens. eCognition 8.0.1. Reference Book. Munich, Germany, Definiens AG, 2010.
- Fonseca, A.D., Fernandes, J.C. Detecção Remota. Editora Lidel, 2004.

- Hubert-Moy, L., Cotonnec, A. Le Du, L. Charlin, A., Perez, P. (2001) A comparison of classification procedures of remotely sensed data applied on different landscape units. *Remote Sensing of Environment*, Vol. 75 (2), 174-187.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., Ferreira, L.G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, Vol. 83, 195-213.
- IFN5 (2010) Relatório final do 5º Inventário Florestal Nacional. Direção Nacional de Gestão Florestal.
- Johansen, K. , Phinn, S. ( 2006b) Mapping structural parameters and species composition of riparian vegetation using IKONOS and Landsat ETM+ data in Australian tropical Savannahs. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, 71-80.
- Lamonca, A., Corona, P., Barbati, A. (2008) Exploring forest structural complexity by multi-scale segmentation of VHR imagery. *Remote Sensing of Environment*, Vol. 112, 2839-2849.
- Lillesand, T. M., Kiefer, R. W., Chipman, J.W. (2004) *Remote Sensing and image interpretation* (5<sup>th</sup> ed.). New York, USA: Wiley.
- Mallinis, G., Koutsias, N., Tsakiri-Strati, M., Karteris, M. (2008) Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *Photogrammetric Engineering and Remote Sensing*, Vol. 63, 237-250.
- Song, C., Dickinson, M.B., Su, L., Zhang, S., Yaussey, D. (2010) Estimating average tree crown size using spatial information from Ikonos and Quickbird images: Across-sensor and across-site comparisons. *Remote Sensing of Environment*, Vol. 114, 1099-1107.
- Stehman, S.V. (1997) Selecting an interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, Vol. 24, N. 6, 1219-1236.