

## Segmentação de imagens aéreas de alta-resolução para a detecção de telhados de edifícios

Vanessa Jordão Marcato <sup>1</sup>  
Aluir Porfírio Dal Poz <sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista - UNESP  
Faculdade de Ciências e Tecnologia - FCT  
Caixa Postal 468 - 19060-900 – Presidente Prudente - SP, Brasil  
vanessamarcato@yahoo.com.br, aluir@fct.unesp.br

**Abstract.** In this paper, a method is proposed for building roofs detection using high-resolution aerial images. The difficulty of this task is to obtain a single region for each building roof from a segmentation technique. Thus, to solve this problem initially some preprocessing techniques are applied in the image. First, the RGB image is transformed into the YUV color space, followed the use of U and V components for generating a two-dimensional histogram. This histogram passes through a process of smoothing by the average and then is applied a watershed technique to separate significant regions of the image in this histogram. The result of this processing is used to segment the image using a pixel-based classification. In the following an operator to fill holes is applied to eliminate small noises present in the regions. The contours of segmented regions are obtained together with their corresponding areas. Finally, the contour of the largest region is conserved, since the building roof is isolated in the image and it is the main object of a pre-selected region. The algorithms used in this work are available in the MVTec Halcon 11.0 software. Some preliminary experiments were made and showed that the method is promising.

**Palavras-chave:** building roof, image processing, aerial image, telhado de edifício, processamento de imagens, imagem aérea.

### 1. Introdução

A extração de objetos que compõem as cenas urbanas (edificações, rodovias, vegetação, entre outros), em imagens aéreas, é um problema de bastante interesse para as Geociências. As pesquisas na área de extração de feições têm crescido nos últimos anos. Em particular, a extração de telhados de edifícios tem sido estudada nas últimas quatro décadas. Até meados da década de 1990 as imagens aéreas eram as fontes usuais de dados utilizadas para a extração. No final dessa mesma década outras fontes de dados (por exemplo, as imagens de satélites de alta-resolução e os dados de varredura a Laser) passaram a ser utilizadas.

A extração de telhados de edifícios é um problema difícil no âmbito do reconhecimento de objetos, o que está relacionado com a complexidade e a variabilidade da cena (Sohn (2004)). Nesse contexto, diversos trabalhos veem sendo desenvolvidos nessa área, envolvendo o uso de imagens aéreas, dados Laser ou a combinação de ambos os dados. Muller e Zaum (2005), propuseram detectar edifícios em imagens aéreas com uso da segmentação da imagem por crescimento de regiões e posterior separação de regiões entre edifícios e vegetação com base em atributos geométricos e fotométricos associados às regiões previamente segmentadas. Katartzis e Sahli (2008) propõem um método automático para detecção de telhados em imagens aéreas. A abordagem é baseada em um modelo de interpretação de imagem estocástico, o modelo de MRF (*Markov Random Field* - Campos Aleatórios de Markov), que combina informações de contexto 2D e 3D da imagem. Lafarge et al. (2008) desenvolveram uma metodologia para a extração de edifícios a partir de Modelos Digitais de Superfície (MDS), em que as partes dos edifícios são extraídas como objetos retangulares. Estes objetos são posteriormente agregados para gerar modelos 3D completos de edificações. Kabolizade et al. (2010) propõe um modelo snake que foca na extração de telhados de edifícios em imagens aéreas coloridas e dados Laser. Galvanin e Dal Poz (2012) utilizaram dados provenientes de varredura a LASER em conjunto com o MRF para a extração de contornos de telhados. Em Chen et al. (2012) é proposto um método para detecção de edifícios com uso de dados LASER e

imagem QuickBird. Os dados LASER são utilizados para gerar um MDSn (Modelo Digital de Superfície normalizado) e, assim serem obtidos os candidatos a representar edifícios, em seguida, a vegetação é filtrada com base em informações da imagem. Awrangjeb et al. (2013) apresentam um método para reconstrução automática de telhados 3D através da integração de dados LASER com imagem multiespectral. As informações provenientes de ambos os dados são utilizadas para a classificação de retas detectadas na imagem.

Nesse contexto, o presente trabalho corresponde a uma parcela já desenvolvida de um projeto que visa extrair automaticamente contornos de telhados de edifícios com a integração de dados LASER e dados de imagem. Nesse trabalho então é apresentado um método de segmentação de imagem aérea de alta-resolução que permite a detecção de telhados isolados de edifícios. Devido principalmente à variações de iluminação nas diferentes faces do telhado, a segmentação de telhados em uma única região costuma ser difícil, o que caracteriza a principal motivação desse trabalho. Também, ao final são obtidos os contornos das regiões correspondentes aos telhados de edifícios.

## 2. Método

Nesse trabalho foram utilizados operadores disponíveis no software Halcon 11 da MVTec. O processo de detectar superfícies inclinadas, como o caso do telhado de edifício, em imagens é bastante desafiador devido as diferentes reflexões da luz no momento da tomada da imagem, que causam uma alta descontinuidade em intensidade/cor em uma única estrutura (Cote e Saeedi, 2013). Dessa forma, com vista a solucionar possíveis problemas relacionados a segmentação de um mesmo telhado de edifício em várias regiões, optou-se por trabalhar com o espaço de cor YUV (Netravali e Haskell, 1995), mais próximo da percepção humana e que torna possível remover a influência da componente intensidade.

Em um primeiro momento, foi realizada a transformação das componentes RGB para o espaço de cor YUV (Devo e Mandowara (2012)). A componente Y determina a intensidade da cor e as componentes U e V são a cor propriamente dita. A seguir, na Figura 1, pode-se observar uma imagem RGB e suas componentes YUV.

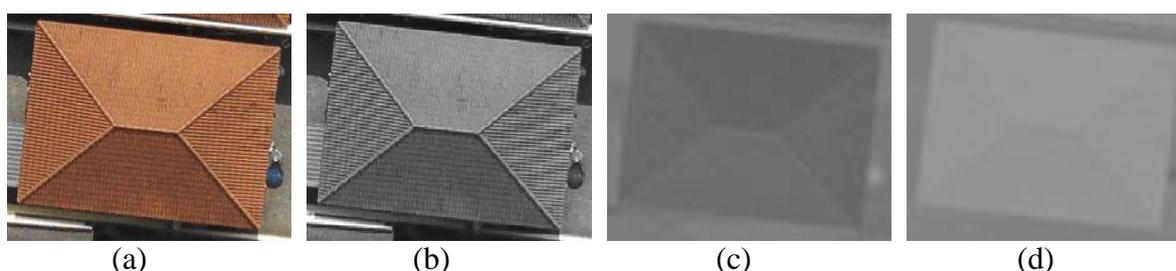


Figura 1. Espaço de cor YUV. (a) Imagem RGB; (b) componente Y; (c) componente U e; (d) componente V.

Na sequência, as componentes U e V são usadas para gerar um histograma bidimensional. Os tons de cinza da componente U são interpretados como índices do eixo linha e os tons de cinza da componente V como índices do eixo coluna. O valor de cinza de um ponto  $P(g_1, g_2)$  da imagem de saída indica a frequência de combinação  $(g_1, g_2)$ . A seguir, a imagem de saída passa por um processo de suavização pela média, com uma máscara  $3 \times 3$ .

O próximo passo consiste em determinar quais e quantas são as regiões significativas da imagem nesse histograma. Para tanto, foi utilizado o operador *watershed* (Gonzalez e Woods, 2000) com uso de limiar. Inicialmente, esse operador segmenta regiões pelo método *watershed* (Gonzalez e Woods, 2000) e em seguida funde regiões que são separadas umas das outras por uma *watershed* menor que um limiar estipulado. Dessa forma, esse limiar funciona como um critério de fusão entre as regiões segmentadas. Supondo  $B_1$  e  $B_2$  sendo os menores

valores de cinza entre duas regiões vizinhas e  $W$  o menor valor de cinza da *watershed* que separa essas duas regiões, duas regiões são fundidas se  $\max\{W-B_1, W-B_2\} < Limiar$ .

A seguir, na Figura 2 pode ser observado um exemplo de imagem representando um histograma bidimensional produzido a partir das componentes U e V apresentadas na Figura 1 (c) e (d). Também estão representadas as regiões significativas encontradas pelo método watershed com uso de limiar. O valor de limiar utilizado foi igual a 50.

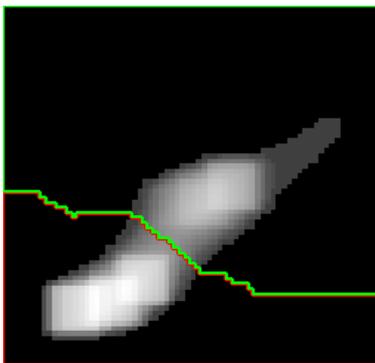


Figura 2. Histograma bidimensional e as *watersheds*.

As *watersheds* encontradas delimitam regiões significativas da imagem no histograma, por exemplo, na Figura 2 foram encontradas duas regiões significativas (uma região delimitada pela linha verde e a outra pela vermelha). Esse resultado é então utilizado para segmentar a imagem a partir de um operador de classificação por pixel de duas dimensões. As dimensões consideradas são as componentes U e V da imagem.

Dessa forma, para cada ponto da imagem são considerados dois valores de cinza (um da componente U e outro da V) e então é feita a comparação com as *watersheds* encontradas a partir do histograma. Assim, seja P um ponto com coordenadas P (L, C),  $g(l)$  o valor de cinza na posição (L, C) da componente U e  $g(c)$  o valor de cinza na posição (L, C) da componente V, busca-se no histograma com base nas *watersheds* encontradas em qual das regiões o ponto deve ser agregado.

Segmentada a imagem, é aplicado um operador de preenchimento de buracos. Esse operador elimina pequenos ruídos através da comparação do pixel com sua vizinhança.

Também são calculadas as áreas de cada uma das regiões segmentadas na imagem. Por fim, conserva-se o contorno da região de maior área, uma vez que espera-se que o telhado de edifício seja o principal objeto de uma região pré-selecionada. Portanto, esse método depende da prévia delimitação de uma região que contém o edifício de interesse. Isso pode ser realizado, por exemplo, através da detecção do edifício em um MDS obtido a partir de dados de varredura laser, com posterior projeção dele no plano da imagem, para possibilitar o isolamento de uma sub-imagem.

### 3. Resultados

Para a realização dos experimentos foi utilizada uma imagem aérea de alta-resolução da cidade de Presidente Prudente/SP. Foram feitos seis experimentos, que podem ser observados a seguir.

O primeiro experimento, Figura 3, mostra um telhado de edifício simples que apresenta diferenças em texturas e iluminação nas suas diferentes faces (Figura 3(a)).

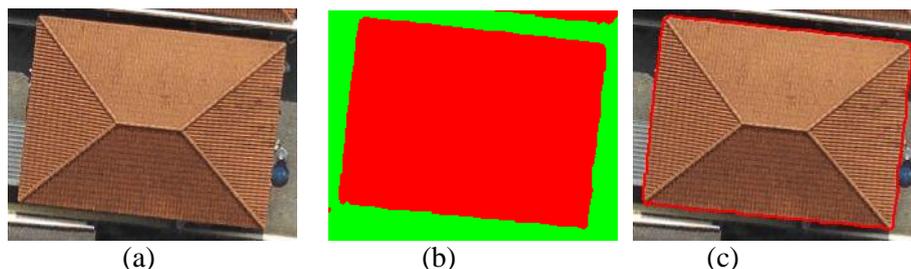


Figura 3. Experimento 1. (a) Imagem RGB; (b) imagem segmentada; e (c) Contorno do telhado de edifício.

Na Figura 3 (b) verifica-se que o método apresentou um bom resultado, uma vez que segmentou o telhado de edifício em uma única região e, dessa forma, foi capaz de lidar com as diferenças em textura e iluminação presentes na estrutura do telhado. Nota-se que o contorno extraído (Figura 3 (c)) delimitou de forma adequada os limites do telhado de edifício.

O próximo experimento (Figura 4) apresenta um telhado de edifício que possui diferenças em textura e uma área de sombra (Figura 4 (a)). Nota-se (Figura 4 (b)) que o método conseguiu delimitar de forma adequada o telhado de edifício, no entanto alguns detalhes do telhado que possuíam grandes diferenças em cor, como por exemplo nas áreas correspondentes às regiões de sombra, foram segmentados em pequenas regiões internas a grande região que corresponde ao telhado edifício.

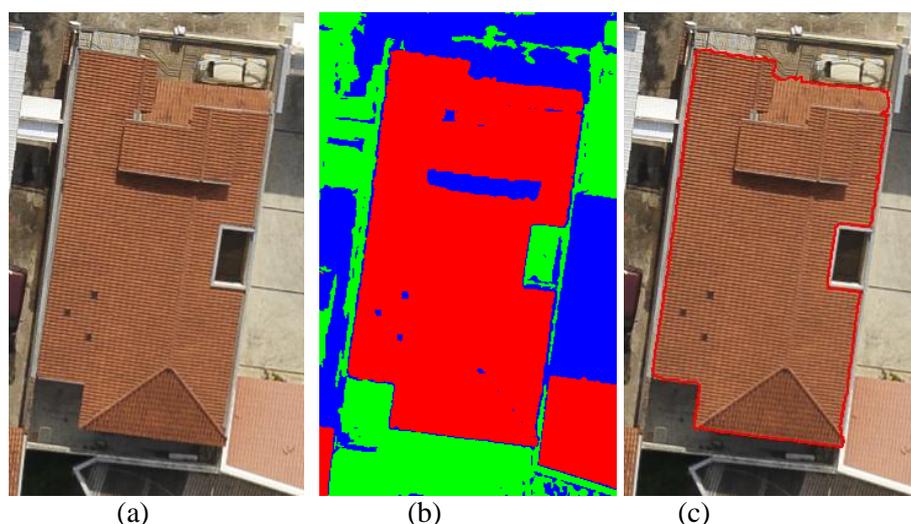


Figura 4. Experimento 2. (a) Imagem RGB; (b) imagem segmentada; e (c) contorno do telhado de edifício.

O que pode ser constatado nesse experimento (Figura 4 (c)) é que apesar de existirem essas pequenas regiões internas, no momento da seleção pela maior área somente conserva-se o correspondente contorno, o que permite então desconsiderar essas regiões menores.

O Experimento 3 (Figura 5) mostra um telhado de edifício que possui diferenças em textura nas suas diferentes faces.

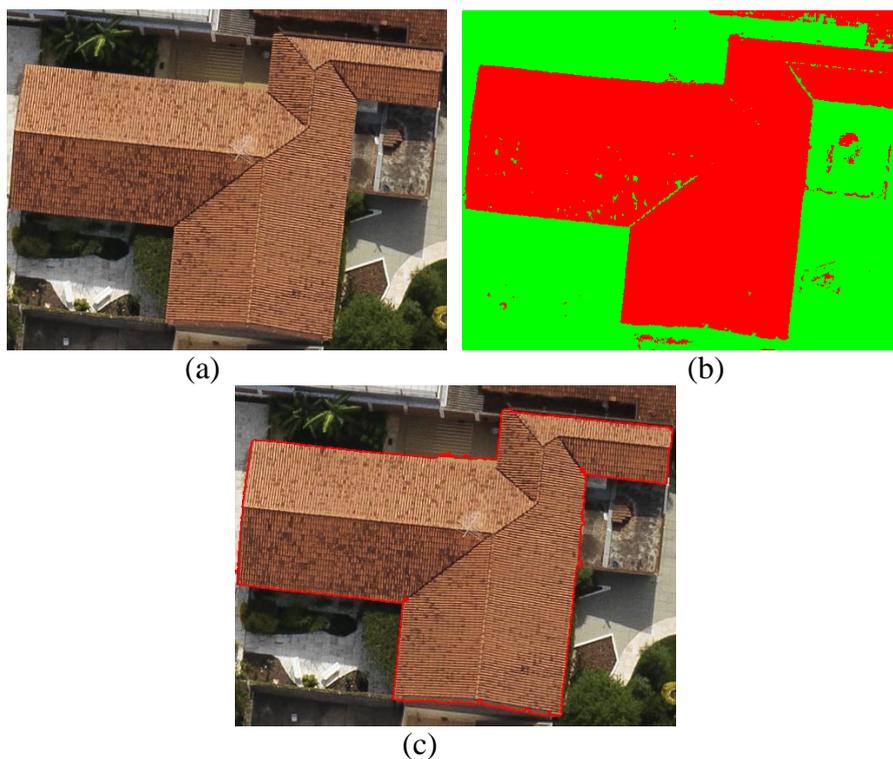


Figura 5. Experimento 3. (a) Imagem RGB; (b) imagem segmentada; e (c) contorno do telhado de edifício.

Nota-se (Figura 5 (c)) que o resultado foi bastante satisfatório, o qual permitiu uma adequada delimitação do contorno do telhado de edifício.

Os Experimentos 4 e 5 baseiam-se em telhados de edifícios bastante parecidos que apresentam sombras causadas por estruturas próximas, como no caso do experimento 4 (Figura 6 (a)) e uma árvore no experimento 5 (Figura 7 (a)).

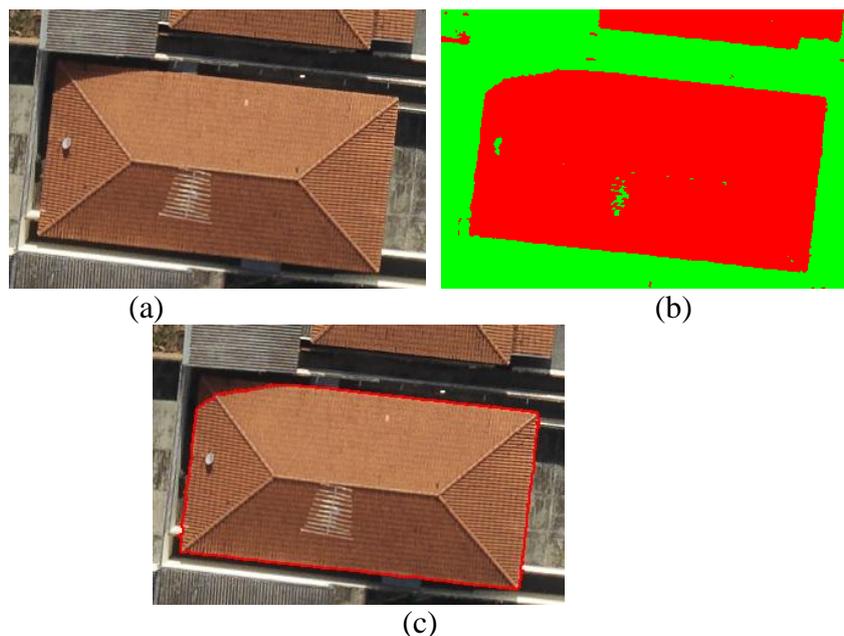


Figura 6. Experimento 4. (a) Imagem RGB; (b) imagem segmentada; e (c) contorno do telhado de edifício.

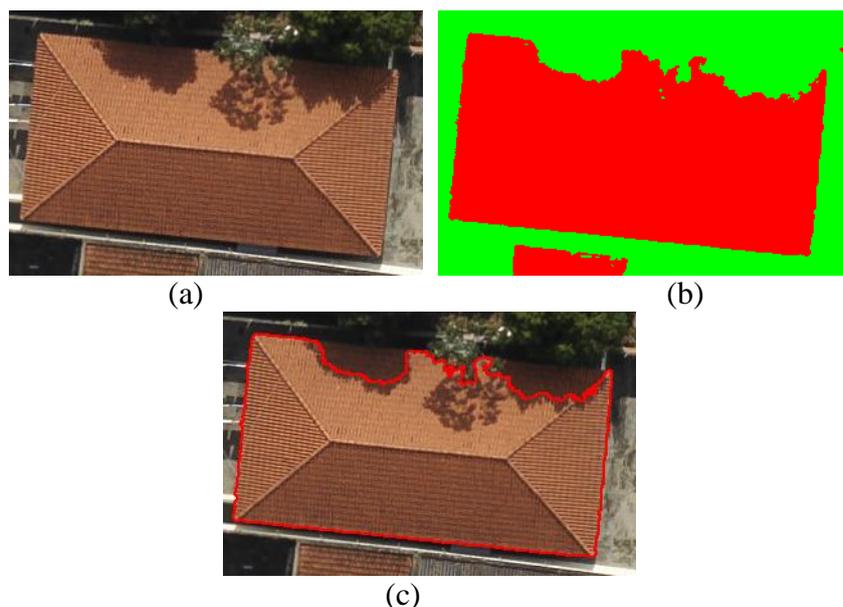


Figura 7. Experimento 5. (a) Imagem RGB; (b) imagem segmentada; e (c) Contorno do telhado de edifício.

Pode-se observar (Figura 6 (b) e Figura 7 (b)) que as sombras atrapalham o processo de segmentação, o que acaba por prejudicar o resultado final (Figura 6 (c) e Figura 7 (c)) e, dessa forma, os telhados não são bem delimitados.

O último experimento (Figura 8) apresenta um telhado na cor cinza com pequenas áreas de sombra.

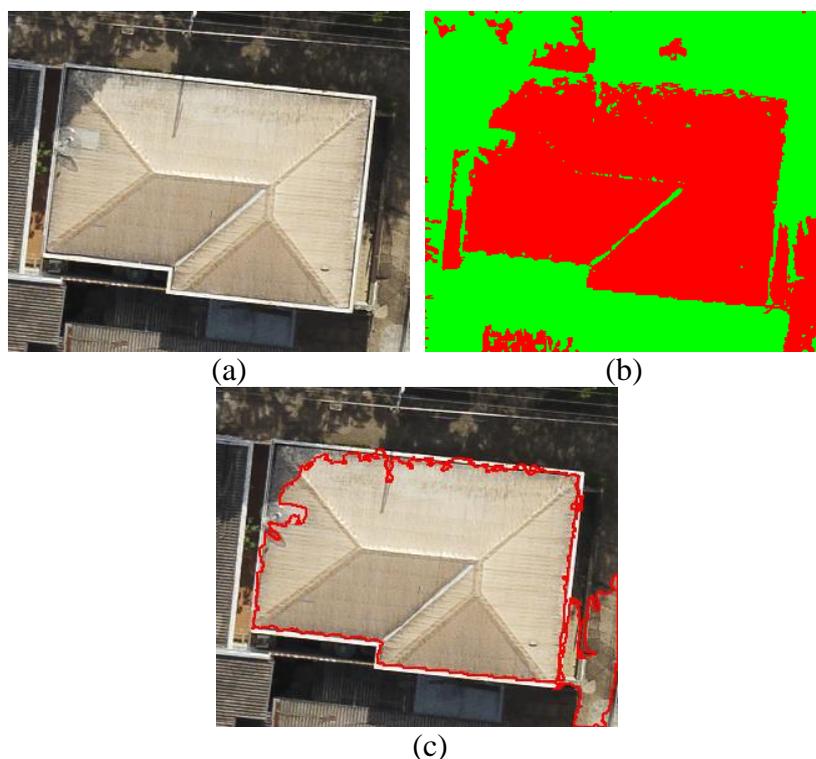


Figura 8. Experimento 6. (a) Imagem RGB; (b) imagem segmentada; e (c) contorno do telhado de edifício.

Na Figura 8 (b) verifica-se que o telhado de edifício de um modo geral foi bem delimitado, no entanto por ser um telhado de cor acinzentada que possui cor bastante próxima

a calçamentos de cimento, nota-se que houve uma confusão que acabou por segmentar parte de uma calçada na mesma região que inclui o telhado de edifício.

#### 4. Conclusões

Esse trabalho teve por objetivo desenvolver um método de detecção de telhados de edifício em imagem aérea de alta-resolução e, assim, obter os respectivos contornos.

O método de segmentação empregado proporcionou bons resultados. Deve-se ressaltar o benefício do uso do espaço de cor YUV, que foi capaz de contornar o problema relacionado às diferenças em iluminação presentes em uma mesma estrutura de telhado de edifício. Também, o uso do método *watershed* possibilitou a separação adequada dos objetos presentes na cena. Vale ressaltar que o método *watershed* trabalha com detecção de descontinuidades, o que favorece o emprego desse método na detecção de contornos de telhados de edifícios caracterizados como um tipo de descontinuidade.

Em trabalhos futuros pretende-se trabalhar com a regularização dos telhados em casos envolvendo sombras ou até mesmo desenvolver um método para tratá-las. Também, será trabalhado o problema de localização automática do edifício na imagem, visando sua posterior detecção pelo método apresentado.

#### Agradecimentos

Os autores gostariam de agradecer à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo financiamento dessa pesquisa através de uma bolsa de doutorado (Processo nº 2012/22332-2).

#### Referências Bibliográficas

Awrangjeb, M.; Zhang, C. Fraser, C. S. Automatic extraction of building roofs using LIDAR data and multispectral imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 83, p. 1-18, 2013.

Cheng, L.; Gong, J.; Chen, X.; Han, P. Building boundary extraction from high resolution imagery and LIDAR data. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 37, p. 693-698, 2008.

Cote, M., Saedi, P. Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution. **IEEE Transactions Geoscience and Remote Sensing**, v. 51, p. 313-328, 2013.

Devi, M. S e Mandowara. Extended Performance Comparison of Pixel Window Size for Colorization of Grayscale Images using YUV Color Space. In: Nirma University International Conference on Engineering (NUiCONE), 2012, Gujarat, India. **Proceedings**...IEEE Publisher, 2012, p. 1-5. On-line. ISBN 978-1-4673-1720-7. Disponível em: < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6493197>>. Acesso em: 23 out. 2014.

Galvanin, E. A. S.; Dal Poz, A. P. Extraction of building roof contours from Lidar using a Markov-Random-Field-Based approach. **IEEE Transactions Geoscience and Remote Sensing**, v.50, p. 981-987, 2012.

Gonzalez, R. C.; Woods, R. E. **Digital Image Processing**, Addison-Wesly publishing company, 2000.

Kabolizade, M., Ebadi, H., Ahmadi, S. An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data. **Computers, Environment and Urban Systems**, v. 34, p. 435-441, 2010.

Katartzis, A.; Sahli, H. A Stochastic Framework for the Identification of Building Rooftops using a single Remote Sensing Image. **IEEE Transactions Geoscience and Remote Sensing**, v. 46, p. 259-271, 2008.

Lafarge, F.; Descombes, X.; Zerubia, J.; Pierrot-Deseilligny, M. Automatic building extraction from DEMs using an object approach and application to the 3D-city modeling. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 63, p. 365-381, 2008.

Müller, S.; Zaum, D. W. Robust building detection in aerial images. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 36, p. 143-148, 2005.

Netravali, A. N., Haskell, B. G. **Digital Pictures: Representation, Compression and Standards**. Springer US, 1995. 706p.

Sohn, G. Extraction of buildings from high-resolution satellite data and lidar. In: XX ISPRS CONGRESS, Istanbul, Turkey, 2004. **Proceedings...** 2004, 7p. On-line. Disponível em:  
< <http://www.isprs.org/proceedings/XXXV/congress/comm3/papers/420.pdf> >. Acesso em: 23 out. 2014.