Morphology-based Building Detection from Airborne Lidar Data

[This Paper was the Winner of the 2008 BAE Systems Award Given at the ASPRS 2008 Annual Conference]

Xuelian Meng, Le Wang, and Nate Currit

Abstract

The advent of Light Detection and Ranging (lidar) technique provides a promising resource for three-dimensional building detection. Due to the difficulty of removing vegetation, most building detection methods fuse lidar data with multi-spectral images for vegetation indices and relatively few approaches use only lidar data. However, the fusing process may cause errors introduced by resolution and time difference, shadow and high-rise building displacement problems, and the geo-referencing process. This research presents a morphological building detecting method to identify buildings by gradually removing non-building pixels. First, a ground-filtering algorithm separates ground pixels with buildings, trees, and other objects. Then, an analytical approach removes the remaining non-building pixels using size, shape, height, building element structure, and the height difference between the first and last returns. The experimental results show that this method provides a comparative performance with an overall accuracy of 95.46 percent as in a study site in Austin urban area.

Introduction

Automatic building detection from high spatial resolution remote sensing images has gained wide attention for urban intra-structure analysis, GIS data production, three-dimensional urban visualization, and hazard damage evaluation (Cho *et al.*, 2004; Peng and Liu, 2005; Peng *et al.*, 2005). Traditional urban land-cover and land-use analysis is mainly based on pixel-level classification instead of the classification of urban elements. Successful building detection allows urban researchers to analyze urban phenomena based on meaningful urban elements and their characteristics by simulating human logical analysis (Jensen, 2000).

Building detection requires high-resolution aerial photographs or satellite images with resolution higher than 5 meters, which can express buildings as objects instead of mixed pixels (Jensen, 2000; Jin and Davis, 2005). Traditional color or multi-spectral images commonly present shadows and high-rise building displacement problems and make building detection troublesome. The newly-developed lidar technique measures the actual three-dimensional coordinates of the reflectance from ground objects instead of a snap-shot from a view angle; hence, the lidar data is free from the influence of sunshine and the high-rise building displacement problem, and produces building-height information as a bonus (Zhang *et al.*, 2006). Additionally, the unique difference between the first and last return in vegetated areas helps to remove most vegetation. Nevertheless, some regularly shaped objects appear similar in shape to buildings in two-dimensional images and are hard to determine by visual interpretation. Researchers can solve this problem by automatically detecting buildings from three-dimensional lidar data. Since this research chooses lidar as the main data source, the following review will focus on lidar-related building detection.

Many techniques have been developed to detect buildings from lidar data, and the most popular approach is to fuse lidar data with multi-spectral images. Popular approaches include morphology (Cho *et al.*, 2004), image segmentation (Rottensteiner and Briese, 2003; Cho *et al.*, 2004), classification (Forlani *et al.*, 2006), and methods based on vegetation indices (Sohn and Dowman, 2003 and 2007; Rottensteiner *et al.*, 2003). Relatively few methods have been developed to identify buildings only from lidar data because of the difficulty to fully remove vegetation (Zhang *et al.*, 2006; Morgan and Tempfli, 2000). Developing comparative techniques to identify buildings only from lidar data is worthwhile because fusing lidar data with other resources introduces errors in the building detection process. These errors have three main causes. First, in order to obtain NDVI, researchers often use multi-spectral images with a different resolution, such as Ikonos with a resolution of 4 meters; while interpolated images from airborne lidar data for building detection can usually reach 0.5 to 1 meters (Sohn and Dowman, 2003). The fusing process usually produces a resolution in between, which sacrifices lidar resolution (Tullis and Jensen, 2003). Secondly, the fusing process, such as registration and resampling, introduces errors in building detection. On the other hand, building or mountain shadows, clouds, and the building displacement problem due to the camera view angle may cause errors in building detection. These errors counteract the advantage of fusing lidar data with other resources, and reduce building identification accuracy. Third, the cost and availability of

Xuelian Meng and Nate Currit are with the Geography Department, Texas State University-San Marcos, 601 University Drive, San Marcos, TX 78666 (xm1001@txstate.edu; currit@txstate.edu).

Le Wang is with the Department of Geography, State University of New York at Buffalo, Wilkeson Quad, Buffalo, NY 14261 (lewang@buffalo.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 75, No. 4, April 2009, pp. 437–442.

0099-1112/09/7504–0437/$3.00/0

other high-resolution images is critical for data production in industry. The purpose of this research is to develop an alternative building detecting method based only on lidar data. Users can easily adapt this method with other resource when extra images are available.

Morphology-based Building Detection

In order to identify buildings, we use a method that first removes ground pixels using a ground filtering algorithm, and subsequently removes the remaining non-building pixels using size, shape, height, building element structure, and the height difference between the first and last returns based on a morphological and analytical approach, as shown in Figure 1. The major procedures can be grouped into two steps: ground filtering and removal of the remaining non-building elements.

Ground Filtering

Ground filtering is a critical first step for removing non-building objects from lidar data (Morgan and Tempfli, 2000; Zhang *et al.*, 2006). We apply a multi-directional ground filtering algorithm on an image interpolated from the lidar first return in order to identify ground and generate a Digital Elevation Model (DEM) based on the results (Meng *et al.*, 2009). This filter provides good performance in dense urban areas with complex ground covering such as trees, shrubs, short walls, and vehicles. The raw lidar data of the first and last returns are irregularly distributed point clouds. A two-dimensional mesh is overlaid on the point clouds, and the elevation of the nearest point is assigned to the elevation of the cell. Then, the ground filter gradually removes non-ground pixels based on thresholds of slopes, elevation difference between the point and its local minimum elevation, and the elevation difference between the point and the nearest ground point. After ground filtering, we generate a DEM to facilitate the calculation of object heights.

Removing Non-building Pixels Based on Morphology Operators

After removing ground pixels, the remaining pixels are mainly composed of vegetation, buildings, and other objects like short walls, and vehicles. Buildings differ from other objects in size, shape, height, and texture. Elevation differences, based on the differences between lidar first and last returns, are important indicators of tall vegetation and are used to separate buildings from trees. Specifically, our method utilizes three pixel-based operations to break non-building blocks into smaller fragments. A building recovering process then refills holes in buildings for pixels wrongly removed in the previous process. Finally, this research converts building candidates into vectors, and then an object-based analysis further removes non-building parcels based on area and compactness thresholds. The following section provides details about the procedures in the dashed frame in Figure 1.

Two primary characteristics commonly used to remove vegetation from building candidates are object heights and the elevation differences between the first and last lidar returns. After the ground filtering process, object heights can be derived by the difference between the first lidar return and the bare earth model DEM generated from the ground filtering results. Objects that are shorter than the minimum building height are removed from the building candidates, such as short walls, shrubs, and vehicles. Besides, the large elevation difference between the first and last lidar returns is a critical indicator for tall vegetation. The active lidar sensors send out laser beams to ground and accept multiple returns depending on the distances the laser beams penetrate. The first return records the reflectance from ground and object surfaces, and the last return documents the deepest points the laser beams reach. Laser lights possess a relatively strong ability to penetrate vegetations and hardly pass hard man-made objects such as building roofs. Theoretically, the elevation differences among ground or building areas should be 0; however, these elevation differences commonly present small values in practice due to the accuracy of lidar measurement and data preprocessing. As a result, we apply an elevation difference threshold to filter out vegetation with large differences.

Height and elevation difference in vegetation areas remove large portion of non-building pixels; however remaining non-building fragments, especially in dense stands of trees, may present similar sizes and morphological characteristics to buildings, making them difficult to separate from buildings. Additionally, trees attached to buildings create irregular building shapes that are difficult to remove. Hence, this research designed an operator based on building element structure to break dense vegetation into smaller fragments and detach remaining linear trees from buildings. Building structure in neighboring pixels may differ depending on spatial resolution. The following illustration applies to images with resolutions equal or higher than one meter, which is often the occasion for urban lidar data. In a high-resolution image, buildings are not supposed to be linear lines, since buildings are usually wider than one meter. Given a three by three neighborhood, as in Figure 2, if both the pixels above and below the pixel (*i, j*) are non-building pixels, or both the pixels on the left and right sides of the

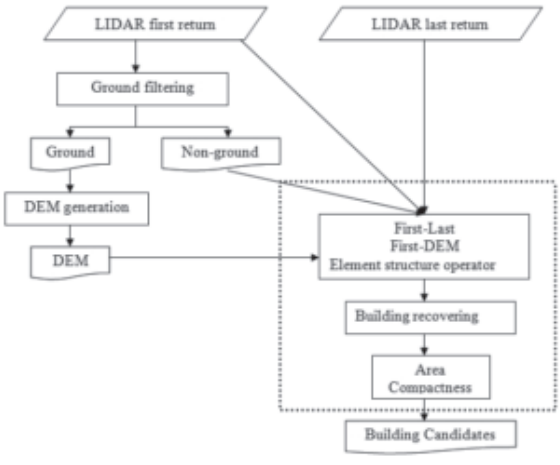

Figure 1. Flow chart of the morphology-based building detection.

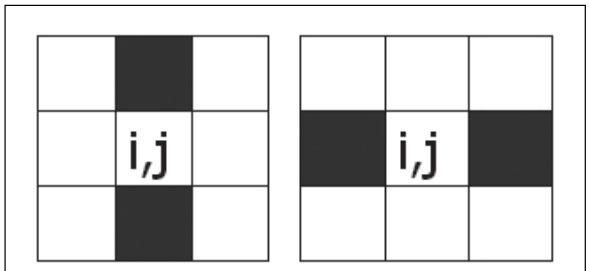

Figure 2. Building elementary structure filtering.

pixel (*i, j*) are non-building pixels; then, the center pixel (*i, j*) is removed from building candidates.

Unlike most methods, this research designed a building recovering process to fill the holes in buildings caused by previous processes before converting building candidates into objects. In the process of elevation difference filtering between the first and last returns, some pixels within building polygons are wrongly removed due to sensor errors, preprocessing, or narrow objects on top of roofs such as power lines, antennas, and vegetation. These building holes increase the difficulty to separate buildings with vegetation based on morphology indices, such as roundness or compactness. These noisy pixels within building polygons randomly appear within buildings and usually contain few pixels. Given a nine by nine neighborhood, a vegetation fragment is likely to contain several non-building pixels. For each non-building pixel, this filter compares the elevation of the pixel with the average value of its neighboring building candidate pixels. If the elevation difference is lower than the threshold, then the pixel is added to the building candidate. Results prove that the recovering process effectively refills holes within buildings but has little influence on vegetation fragments. Finally, this research converted building candidate pixels into vectors and further removes non-building parcels based on area and compactness thresholds. Equation 1 illustrates the compactness calculation; the output based on object analysis is the final building candidate map:

$$\text{Compactness} = \text{area} \times \text{peremeter}^2/(4) \quad (1)$$

Study Site

The study site is located on the campus of the University of Texas at Austin and covers an area of 0.67 km^2. The site is on a hillside covered by dense urban features such as bridges, vehicles, trees, flagpoles, fences, shrubs, short walls, and buildings with various sizes and shapes. A Optech ALTM lidar sensor collected the raw lidar data in 2000. A sampling process based on the nearest principle converts the point clouds from the first and last returns to images with a 0.5-meter resolution, and the pixel values are the elevation values of the nearest points searched within certain distance in raw lidar data. Figure 3 shows the first returns of the study site with the year-2000 white building ground truth polygon data that is modified from the year-2003 building vector data based on a field trip on 24 March 2007.

Results and Discussion

This research experiment applied a 0.2-meter threshold for the elevation difference between the first and last return, a 2 meter height threshold to remove off-terrain objects lower than the minimum building height, and a 40 m^2 area threshold to remove objects smaller than the minimum building size. Figure 4 is the result of final building candidates overlaid on top of the first return. The accuracy assessment shows 95.46 percent of the overall accuracy, and a 0.89 Kappa value. The overall accuracy shows the percentage of the pixels that are correctly identified, which is obtained by dividing the number of correctly identified ground and non-ground pixels with the total pixel number. Kappa value considers the performance in different classes based on a confusion matrix (Congalton, 1991; Congalton and Green, 1999; Jensen, 2005). Table 1 shows the error matrix and Figure 5 is the classified error spatial distribution analysis. On the right side of Figure 5, the top two symbols represent pixels that are correctly identified. The third symbol means the errors of labeling building pixels as non-building (false negative), and the fourth symbol demonstrates the errors of labeling non-building pixels as building (false positive). The results show that this method successfully

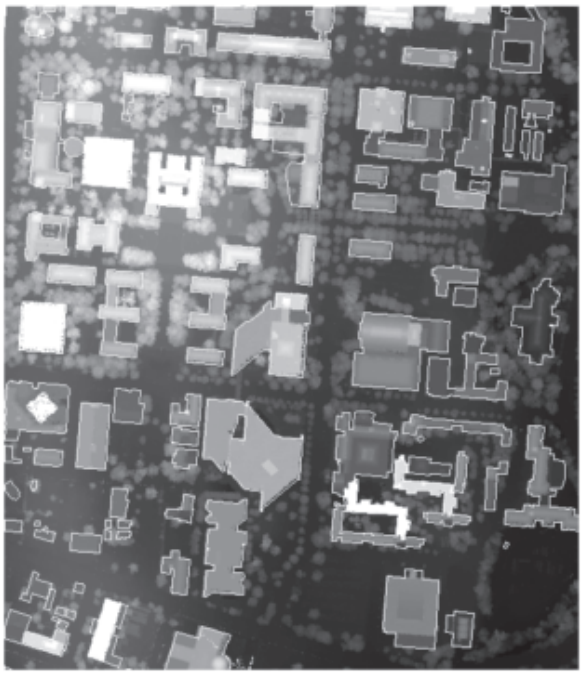

Figure 3. The lidar first return and building ground truth data overlay.

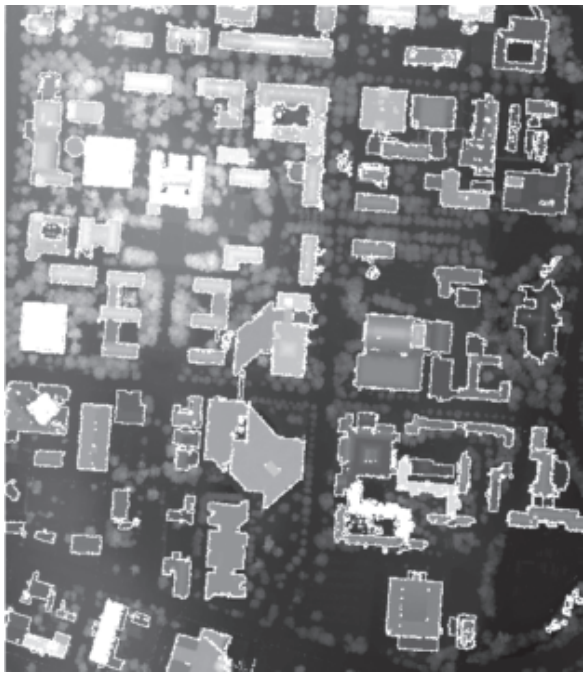

Figure 4. Building detection results.

TABLE 1. ERROR MATRIX IN PERCENTAGE FOR BUILDING DETECTION RESULTS

		Ground Truth Data	
		Ground	Building
Detection Results	Ground	68.75	2.78
	Building	1.76	26.71

identified most buildings. Few buildings that are smaller than the minimum building threshold are removed from the building candidates, which are mainly storage outbuildings. Some large and dense vegetation parcels are troublesome to remove as the vectors present similar size and compactness.

To further analyze the major factors that lead to the 4.54 percent misclassification, we examined the major error types by comparing them to the high-resolution Google Earth™ images as in Figure 6. The first type of error is erroneously removing small buildings. These errors occur because the estimation of the minimum building size is based on residential or commercial buildings. In this result, most of the building structures removed are storage outbuildings, green houses, outside equipment. As an example, Figure 6a is a fitness storage room based on the field trip. The second type of error originates from errors in ground truth data or differences in the algorithmic building definition and the ground truth data. As shown in Figure 6b, there are three error clusters where building pixels are labeled as non-building pixels. The upper cluster is a concrete surface with the same height as the ground, but is higher than the underground building.

Similarly, ground within the buildings in Figure 6d and 6f are surfaces with the same heights as the outside ground surfaces, but are wrongly digitized as buildings. The lower cluster in Figure 6b is a concrete surface for an open dining area, which is elevated around 30 centimeters above ground. While unable to view buildings three-dimensionally, the digitizer for building vectors are more likely to identify the surface as part of the building. This means that the ground truth data and the building detecting algorithms may have different definitions about buildings for some complex structures. The third type of error is due to the sensor errors, preprocessing, and narrow objects on top of roofs such as power lines, antennas, and vegetation. Figure 6c is a high-rise building with many small antennas or sensors on the top of the building, which cause the error in the process of the lidar first and last return elevation difference filtering. Figure 6b is a roof with a south edge as low as the bushes and ground, which may be identified as ground in the ground filtering process. The fourth type of error is caused by dense vegetation that is hard to penetrate by laser beams. Figure 6e is an example of dense tree area that is difficult to remove.

Overall, we present a morphological building detection method to gradually remove non-building objects from building candidates based on lidar data. We tested the method on a study site with various ground features, such as bridges, vehicles, trees, fences, short walls, and buildings from small storage rooms to large irregularly shaped man-made structures. Without the aid of other image resources, this method was able to produce a competitive overall accuracy of 95.46 percent and 0.89 Kappa value based on a study site in Austin, Texas. Within the 4.54 percent errors, most mislabeled pixels distribute around building boundaries instead of confusion with other stand-alone objects, and we observed four major components that cause those errors.

Conclusions

The advent of lidar technique provides a promising resource to detect three-dimensional buildings. Current lidar-based methods commonly work on a combination of lidar and

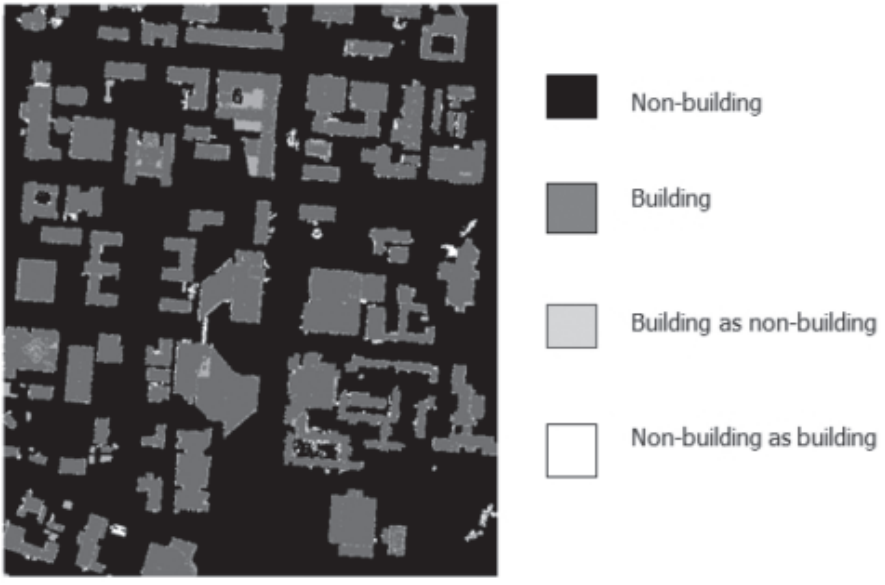

Figure 5. Error spatial distribution analysis for the building detection. The upper two on the right side are correctly identified points. The third symbols represents the error of labeling building pixels as non-building, and the fourth color demonstrates the error of labeling non-building pixels as building.

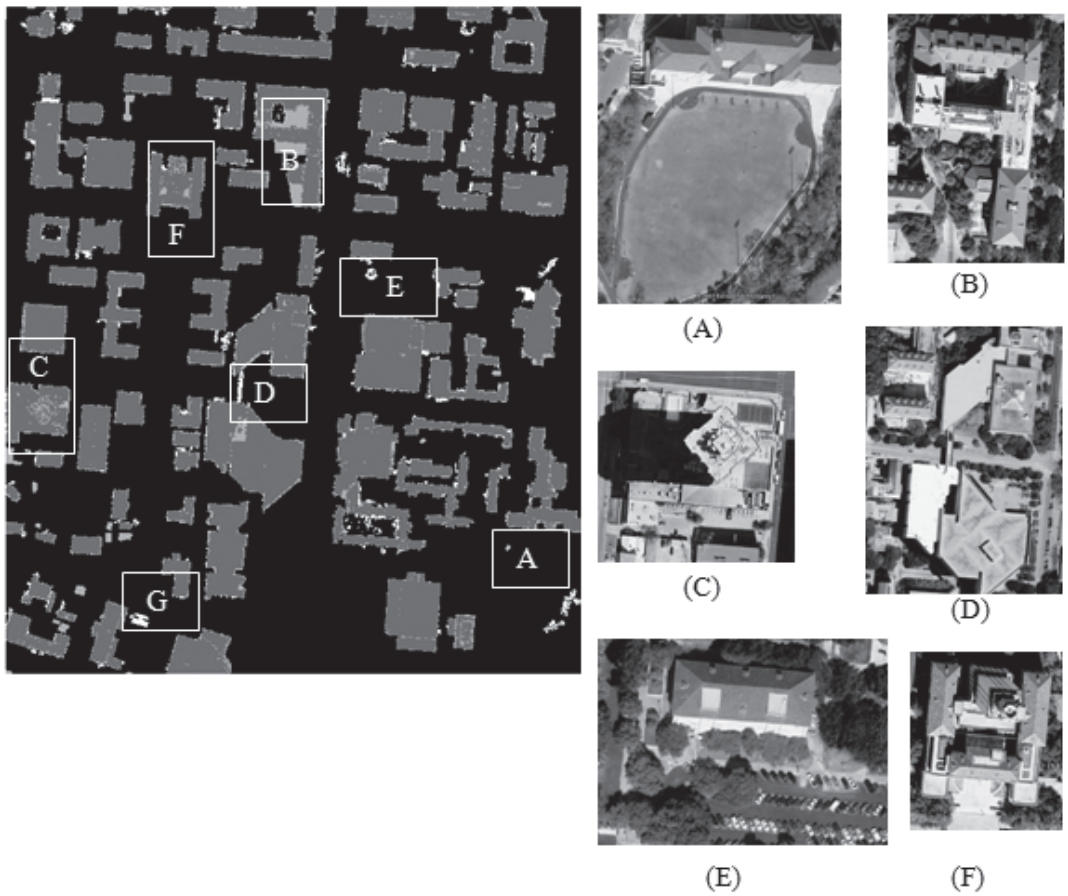

Figure 6. Building detection error analysis based on Google Earth ™ images: (a) a fitness storage room, (b) a concrete surface for an open dining area, (c) a high-rise building, (d) and (f) surfaces with the same heights as the outside ground surfaces, but are wrongly digitized as buildings, and (e) example of a dense tree area that is difficult to remove.

other multi-spectral images due to the difficulty of removing vegetation; relatively few methods only apply on lidar data. Using multi-resource images introduces errors during the fusing process due to preprocess and the differences in resolution and time. This research developed an alternative morphological building-detection method based on lidar data by gradually removing non-building pixels. The experiment's results show that this method provides a comparative performance with an overall accuracy of 95.46 percent as in the study site in Austin urban area.

Acknowledgments

This study was supported by grants to Le Wang from the National Science Foundation (BCS-0822489, DEB-0810933), and from the National Key Basic Research and Development Program, China (2006CB701304).

References

Cho, W., Y. Jwa, H. Chang, and S. Lee, 2004. Pseudo-grid based building extraction using airborne LIDAR data, *International Archives of Photogrammetry and Remote Sensing*, 35(B3):378–381.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 3735–3746.

Congalton, R., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC/Lewis Press, Boca Raton, Florida, 137 p.

Forlani, G., C. Nardinocchi, M. Scaioni, and P. Zingaretti, 2006. Complete classification of raw LIDAR data and 3D reconstruction of buildings, *Pattern Analysis & Applications*, 8:357–374.

Jensen, J.R., 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice Hall, New York, 526 p.

Jin, X., and C.H. Davis, 2005. Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information, *EURASIP Journal on Applied Signal Processing*, 2005(14):2196–2206.

Meng, X, L. Wang, J.L. Silván, and N. Currit, 2009. A multi-directional ground filtering algorithm for airborne LIDAR, *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1)117–124.

Morgan, M., and K. Tempfli, 2000. Automatic building extraction from airborne laser scanning data, *Proceedings of the 19th ISPRS Congress*, Amsterdam, The Netherlands, Book 3B, pp. 616–623.

Peng, J., and Y. Liu, 2005. Model and context-driven building extraction in dense urban aerial images, *International Journal of Remote Sensing*, 26(7):1289–1307.

Peng, J., D. Zhang, and Y. Liu, 2005. An improved snake model for building detection from urban aerial images, *Pattern Recognition Letters*, 26(5):587–595.

Rottensteiner, F., and C. Briese, 2003. Automatic generation of building models from LIDAR data and the integration of aerial images, *Proceedings of the ISPRS Working Group III/3 Workshop*, International Society for Photogrammetry and Remote Sensing, Vol. XXXIV, URL: *http://www.ipf.tuwien.ac.at/publications/2003_dresden_FR_CB.pdf* (last date accessed 18 January 2009).

Rottensteiner, F., J. Trinder, S. Clode, and K. Kubik, 2003. Building detection using LIDAR data and multi-spectral images, *Proceeding of the VIIth Digital Image Computing: Techniques and Application*s, Sydney, Australia, pp.10–12.

Sohn, G., and I. Dowman, 2003. Building extraction using LIDAR DEMs and IKONOS images, *Proceedings of the ISPRS Working Group III/3 Workshop*, International Society for Photogrammetry and Remote Sensing, XXXIV(PART 3/W13), Dresden, Germany, pp. 8–10.

Sohn, G., and I. Dowman, 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction, *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1):43–63.

Tullis, J.A., and J.R. Jensen, 2003. Expert system house detection in high spatial resolution imagery using size, shape, and context, *Geocarto International*, 18(1):5–15.

Zhang, K., J. Yan, and S. Chen, 2006. Automatic construction of building footprints from airborne LIDAR data, *IEEE Transactions on Geoscience and Remote Sensing*, 44 (9):2523–2533.