DEM Generation and Building Detection from Lidar Data

Ruijin Ma

Abstract

Object reconstruction has attracted great attention from both computer vision and photogrammetry communities, and new technologies are being introduced into this research society. Lidar (Light Detection And Ranging) has become well recognized in the geomatics community since the late 1990s. Compared with traditional photogrammetry, lidar has advantages in measuring surface in terms of accuracy and density, automation, and fast delivery time. There is a large market in geo-data acquisition and object recognition for lidar technology (Baltsavias, 1999). In a general sense, lidar is a companion technology for traditional photogrammetry. The direct product that can be derived from lidar data is the DSM (Digital Surface Model), which depicts the topography of the earth's surface, including objects above the terrain. Further processing can be carried out to generate DEM (Digital Terrain Model) and object models like buildings, which is very useful information in telecommunication, city planning, flood control, and tourism. Morphology and classification are two commonly used methods in DEM generation and object reconstruction. However, these two methods are either sensitive to errors or of low accuracy. In this paper, a new method is proposed to extract ground points for DEM generation and to detect points belonging to buildings. A new method for boundary regularization is also proposed. The results show that buildings can be detected with high accuracy from lidar data.

Introduction

The development of lidar technology started in the 1970s and 1980s originally in the United States and Canada (Ackermann, 1999). At present, the technology is maturing and many laser-scanning systems are available on the market (Axelsson, 1999). The most attractive characteristic of lidar is that it has very high vertical accuracy which enables it to capture the Earth's surface with high accuracy. This characteristic makes lidar data very suitable for DEM generation and building reconstruction. After the pre-processing, the process of lidar data usually begins with the separation of ground points and non-ground points. Following the separation, DEM can be generated from ground points and objects like buildings can be extracted from non-ground points through further processing.

In lidar data processing for DEM generation, the critical step is to correctly separate the lidar points into ground and non-ground points. One commonly-employed procedure is to convert the lidar point data into grid formatted data; thus, some algorithms and methodologies based on image processing can be applied to the lidar grid data. To distinguish ground points from a lidar point cloud, morphology filters can be applied based on the assumption that the ground point height is lower than its neighbor object points. Another assumption is that the ground is smooth. In other words, there is no abrupt change on the ground. Some studies on this separation have been carried out and good results were produced (Weidner and Förstner, 1995; Morgan and Habib, 2002). However, morphology filters are sensitive to errors. Although a median filter can be used to decrease the effects from a single error point, the effects from errors in the form of a point patch cannot be eliminated or decreased. Kilian *et al.* (1996) used a "multi-level opening" morphology operator in order to keep small ground features while removing large non-ground objects. Because small windows have small weights, small features can still be removed. Axelsson (1999) proposed an idea to separate ground and non-ground points from lidar data which is to extract a surface from the point cloud to identify the ground surface. Controlled by defined input parameters, the moving surface can adjust itself to identify points on the ground.

"Linear prediction" is a statistical interpolation method which is employed in lidar data segmentation by researchers to generate digital surfaces (Lohmann and Koch, 1999; Lohmann, *et al.*, 2000; Kraus and Pfeifer, 1998; Lee and Younan, 2003). The interpolation is founded on the spatial correlations of neighboring points, which are expressed in a form of covariance. The covariance is calculated using a covariance function based on the distance between points. In addition, the removal of non-ground points is conducted based on a weighing scheme. Each point gains its weight based on the difference between its height and the generated DEM in a previous step. The weights are also applied into the DEM interpolation. The critical point of this method is to get a good approximation of the initial DEM. Non-ground points included in the initial DEM will be difficult to correct.

Vosselman (2000) proposed a slope-based method to filter out non-ground points. It is a modification of morphology erosion operator. A point is classified as a ground point if the maximal slope of the vectors connecting the point under test to all its defined neighbors does not exceed the maximal slope with the study area. Sithole (2001) modified this method to use different maximal slope thresholds according to local terrain characteristics. A rough slope map is needed to calculate the local slope threshold. Like

Ruijin Ma is with the CET Department, SUNY-Alfred, 10 Upper College Drive, Alfred, NY 14802. He was formerly with the Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, OH 43210.

Photogrammetric Engineering & Remote Sensing
Vol. 71, No. 7, July 2005, pp. 847–854.

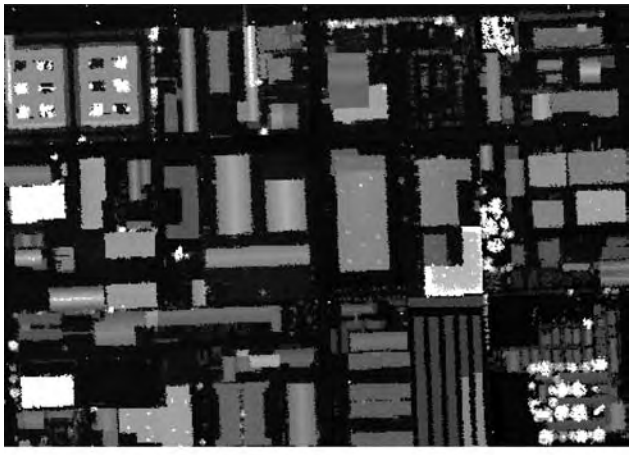
0099-1112/05/7107–0847/$3.00/0

morphology operators, the problem with these kinds of methods is how to determine the optimal neighbor of a point, i.e., the determination of the structure elements. A large neighbor relationship will eliminate small features, while a small neighbor relationship will include non-ground points, such as points on a large building.

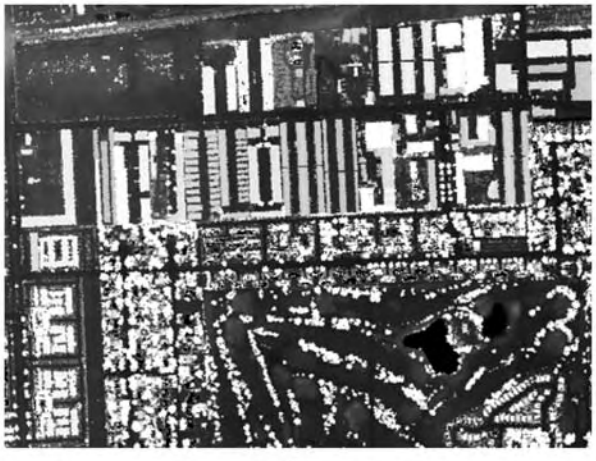
Several studies have been conducted by researches to perform building detection from lidar data. After performing a "bottom up" region growing segmentation, Matikainen, *et al.* (2003) used the height information to separate trees and buildings from the ground. Then, a fuzzy classification method was applied to detect buildings based on three attributes: the Gray Level Co-occurrence Matrix (GLCM) homogeneity of height, the GLCM homogeneity of lidar intensity, and the average length of edges from a "shape polygon" derived from the segment under test. Rottensteiner and Briese (2002) first used the height difference between the DSM and the DEM, a morphological opening operator, and the size measurement to detect buildings and tree groups. A polymorphic feature extraction method based on the Förstner operator was then used to detect "point-like" pixels. Based on the analysis of the number of "point-like" pixels of each segment, tree groups were eliminated. This whole process was repeated before the final building regions were generated. Lohmann (2001) investigated the Gaussian Laplace (GL) filter to detect break lines like dike edges by proposing to use the mean curvature to overcome the difficulty in determining thresholds for the GL filter result. This method could be used to detect building outlines. Brunn and Weidner (1997) used a Bayesian Network classification to detect buildings based on the height difference between the DSM and the DEM, detected step edges, and the surface normal variation. Axelsson (1999) used a classification method based on the minimum description length criteria to separate buildings and trees. The cost function was calculated from the second derivatives of the surface. Some research on building detection was also accomplished based on the integration of multi-source data (Haala and Brenner, 1999; Schiewe, 2003).

The information delivered directly by lidar data is the height information of the Earth's surface. Other information, such as texture can be derived based on the point neighbors, and this kind of information can be used to perform lidar data segmentation (Elberink and Maas, 2000; Lohmann, 2001). Some instances of texture that can be derived from the height information are slope, variance, and aspect. In the study described in this paper, two textures were derived and used to separate ground and non-ground lidar points. Ground points were used to generate DEM, while non-ground points were further processed to extract building regions, and regular geometric building boundaries were obtained.

Two data sets were used in this research experiment which were provided by Harris County, Texas. The horizontal point spacing of Harris data set is approximately one meter with a horizontal accuracy of approximately 0.5 meter. The vertical accuracy RMSE is reported by the data provider to be approximately 15 centimeters. Figure 1 shows these two data sets in grid format. The first data set has an approximately 2 meters difference in terrain topography, while the second has a difference of around 4.5 meters including a park area within the study area.



(a)



(b)

Figure 1. Two data sets in grid format used in this study: the elevation varies (a) from 16 to 59 meters, and (b) from 18 to 63 meters.

Lidar Data Segmentation

The purpose of lidar data segmentation is to separate the lidar point data into different classes depending on specific application requirements. To generate the DEM of the area under study, the ground lidar points should be distinguished. In urban and suburban areas, the ground is relatively smooth without abrupt changes in topography. Under these conditions of lidar data resolution and accuracy, smooth ground without abrupt change can be treated locally as a planar surface. That means a group neighboring ground lidar points, for example, points within a 3 pixel by 3 pixel window, can form a local planar surface with regression accuracy no less than the accuracy of the original data points. This characteristic of forming a local planar surface is used together with size measurement in this paper to distinguish points on planar surfaces, such as ground and building roofs from points falling onto trees or shrubs.

For each lidar point, its neighboring points are used to calculate a regression planar surface, and how well the tested point matches the regression planar surface indicates whether it is on a planar surface. The neighbor is defined as a square window of 3 pixels by 3 pixels in this experiment. Two measurements for this tested point can be derived using this regression planar surface. One is the Root Mean Square (RMS) calculated from the neighboring points used to derive the regression planar surface, and the other is the difference between the actual height value of the test point and its height value calculated from the regression planar surface. Both measurements can be used to test whether or

not a point falls onto a planar surface formed by its neighboring lidar points. These two measurements correlate with each other and thus produce similar results. Points with a small RMS or such a difference are classified as points on ground or building roofs, and points with a large RMS or difference are classified as points falling onto objects with non-planar surfaces, such as, trees and shrubs.

Based on the vertical accuracy of original lidar data, a threshold of the difference or RMS is determined. In this example, 30 centimeters is used as the threshold of the fitting difference for classification which is double the stated vertical accuracy of original lidar data. This threshold can correctly classify most planar points, approximately 96 percent assuming a normal error distribution. Further processing will test and classify more points as ground points. Points with a difference smaller than the threshold are classified as points falling onto a planar surface. The classified planar points include points on the ground, points on building roofs, and some scattered points on other objects like trees and cars.

For easier processing, the lidar point data is converted into a regular grid data format using a binning process (Hu and Tao, 2002) with a grid size of one meter. The planar-surface fitting algorithm was implemented on the grid data. Figure 2 shows the calculated planar-surface fitting differences of the two used data sets with large-difference points represented as light pixels. Detected points or pixels falling onto planar surfaces were extracted for further processing. Connected regions in the grid were detected and labeled, and these areas of these connected regions were calculated by counting the pixel numbers within these regions.

To differentiate ground points from building roof points, it is assumed that the ground points form one or more connected planar surfaces that have larger areas than the largest building within the study area. This assumption is true in urban and suburban areas. Regions with areas larger than the largest building size within the study area were extracted as ground regions. The ground region can also be simply selected manually. Lidar points falling within the ground regions were extracted and used to generate an initial DEM. Figure 3 shows the work flow for this process, and Figure 4 shows the detected points on planar surfaces and points on the ground. It should be noted that the manually-picked ground region is the one connected by, or including, the road network. Not every ground region must be selected at this point, such as the inner court ground areas. These areas will be included in the ground region during the DEM refinement procedure which is presented in the following section.

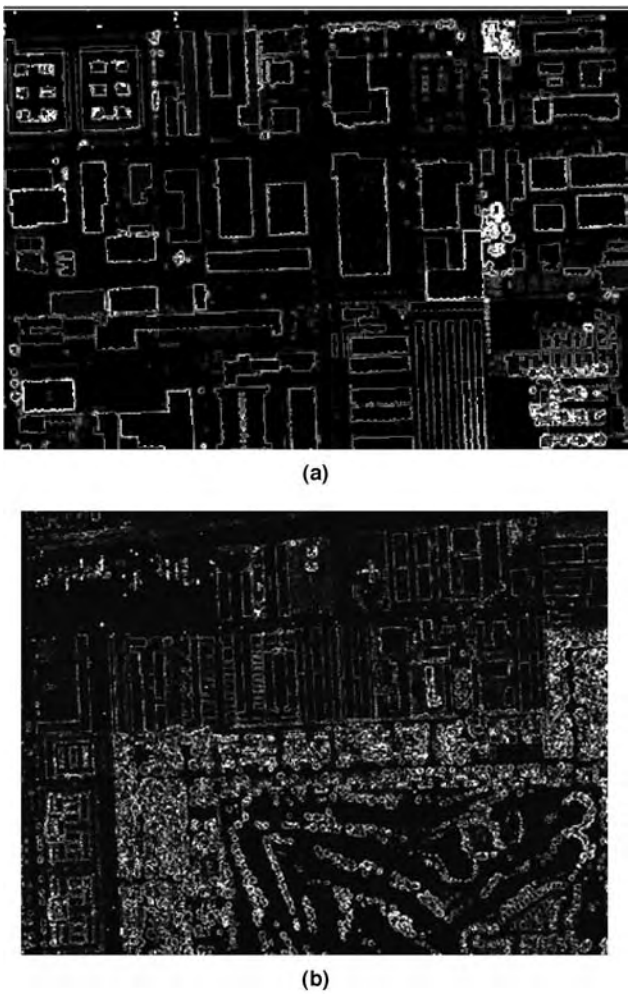


Figure 2. The difference texture of each point calculated from planar surface fitting for both data sets: the differences vary (a) from 0 to 3 meters, and (b) from 0 to 2 meters.

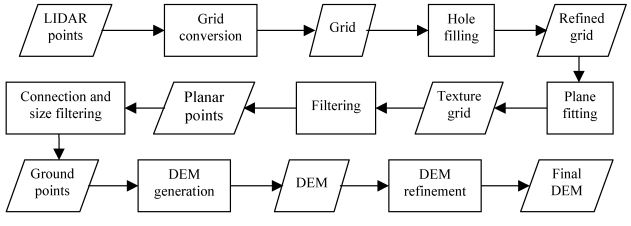


Figure 3. The flow of DEM generation from lidar points.

DEM Generation and Building Detection

After the ground points were detected, a DEM was generated using an interpolation method. In this experiment, a Triangulated Irregular Network (TIN) was created first, and then a DEM grid was generated from the TIN model. However, this DEM is a rough DEM whose accuracy can be improved. A refinement was conducted on the initial DEM to get better accuracy by comparing the lidar point heights with the DEM elevation. In the planar-surface fitting procedure, points falling on the boundary regions of a planar surface also had larger fitting differences due to the height jumps occurring in the boundary regions. These boundary points, that are not initially detected as ground points, should be included in the final generated DEM. After the initial DEM was generated from the classified ground points, all lidar points classified as non-ground points were compared with the DEM. Those points with differences smaller than the threshold were reclassified as ground points, and these ground points were used to refine the DEM. Again, the difference threshold in the comparison is the double that of the stated vertical accuracy of the original lidar point data, i.e., 30 centimeters. This comparison also correctly classifies ground points below trees, which were excluded at the beginning.

The ground points were updated to include new detected ground points from the comparison of the previously generated DEM and the actual point heights. The new generated DEM had better accuracy since more ground points were

Figure 4. Points detected falling onto planar surfaces (a) and points detected falling onto the ground (b).

included. The procedure was repeated until no significant number of new ground points was detected or a fixed iteration number was reached. Figure 5 shows the DEMs generated from the lidar data. The procedure above was repeated four times for both data sets before the final DEMs were generated. The terrain differences within this study area were about 2 meters and 4.5 meters, respectively. The DEM in Figure 5 was stretched to demonstrate detailed information. It can be seen from Figure 5 that some artificial features exist in the DEM, such as the triangles in building footprints. These features are caused by the interpolation of grid DEM from TIN models. Further study will be conducted on interpolation methods for DEM generation, such as spline approximation, linear prediction, and *k*-order Voronoi approach.

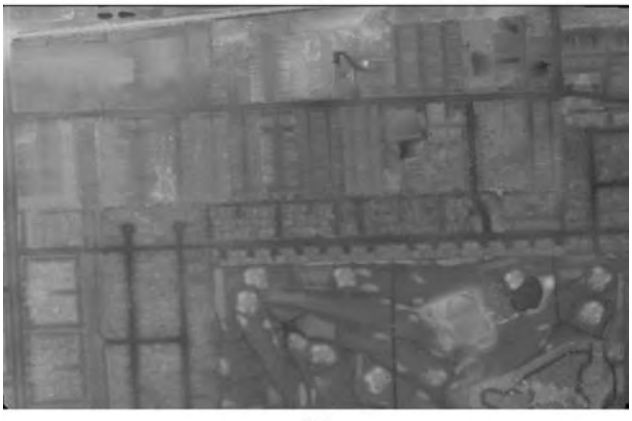
After a DEM was generated, objects on the ground like buildings, trees and cars were distinguished by analyzing the object heights with the DEM. A Normalized Digital Surface Model (NDSM) was generated by subtracting the DEM from the original lidar point grid. In the NDSM, objects such as buildings could be viewed as sitting on a level planar surface; thus, these objects could be detected by simply checking their height values in NDSM. Points with heights larger than three meters in NDSM were extracted as building and tree points, and points lower than three meters were classified as objects like cars and shrubs, which were not processed further in this study. Figure 6 shows the regions detected as buildings and trees. The detected regions on the image margin were caused by the different sizes of the original lidar and interpolated DEM grids. The DEM grids were smaller than the original lidar grids because some points on the original grid margins were not detected as ground points. Thus, the ground point coverage was smaller, and consequently the interpolated DEMs were smaller than the original lidar grid data. These regions were eliminated during the building detection process described in the following sections.

Buildings are assumed to have planar roof surfaces, while trees do not. During the DEM generation, planar surfaces have been extracted by planar-surface fitting method proposed in this study. Obviously, the conclusion can be drawn that the intersection of these two data sets, the planar surfaces and the high objects in the NDSM, are building regions. Figure 7 shows the detected building regions. Scattered regions with small areas were eliminated with a threshold of building size, which is the smallest building size known as a prior knowledge in the area under study. For the detected building regions, a morphology dilation operator was used to compensate for the lost points on building boundaries during the planar surface detection. The window size is the same as the window used to calculate the planar fitting difference. Regions connected to the image boundary were eliminated because these buildings were cut by the image boundary, and they were not complete within the data sets.

Building Boundary Regularization

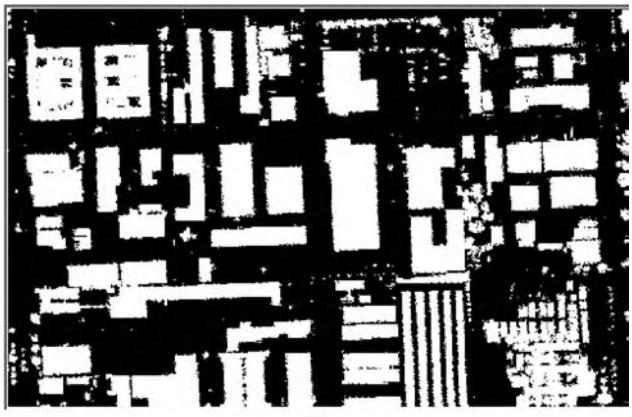
After the building regions were detected, the building boundary of each region was extracted. Due to noise, building boundaries were broken into small, ragged segments. The buildings were assumed to have rectangular-shape boundaries, which means the boundary segments are either parallel or perpendicular to each other. The traditional polygon simplification method cannot guarantee regular building shapes. To eliminate the effects of noise and to get parallel

(a)



(b)

Figure 5. DEM generated from lidar data (artificial features are caused by the interpolation from TIN models to grid models).



(a)

(b)

Figure 6. Points detected as falling onto buildings and trees from NDSM.

or perpendicular boundary segments, some researchers have used the Minimum Description Length (MDL) method to regularize the ragged building boundaries (Weidner and Förstner, 1995). MDL is a statistical method, and requires intensive computation. Mayer (2001) used a constrained active contour method in optimizing building boundaries, however, this method requires a good approximation. In addition, it cannot merge or eliminate small line segments. In this study, a new regularization method was developed to determine rectangular-shape building boundaries.

To use the regularization method, the extracted building boundaries were generalized first to avoid fragmented line segments and redundant points. Then, the regularization was conducted on the generalized building boundaries. The algorithm can be described as follows, with a flow chart presented in Figure 8:

- The azimuth of each line segment of a building boundary was calculated, and all segments were clustered into two classes according to their azimuths. A segment was classified into class **A** if the difference between its azimuth and the averaged azimuth (the cluster center) of class **A** was smaller than the difference between its azimuth and the averaged azimuth of class **B**. The result of this step is two groups of line segments, which are supposed to be perpendicular to each other. At the beginning, the two cluster centers were chosen as the azimuths of two segments, which has a difference of larger than 60 degrees;
- For each segment class, the weighted average of the azimuths was calculated. The weight used was the length of each line segment: $azimuth = \Sigma^{l_i * azimuth_i} \big/ \Sigma^{l_i}$, l_i is the length of i^{th} segment in one class. This matches the observation that a longer line segment has higher azimuth accuracy than a shorter line segment, provided that the end point position accuracy is the same. The output of this step is two azimuths that are supposed to be perpendicular to each other;
- A weighted adjustment using a Gauss-Markov model was carried out to make the azimuths of these two classes perpendicular. Again, the weight was calculated as the total length of all segments in each class. After the adjustment, these two azimuths assigned to two classes are now perpendicular;
- All segments were adjusted so that each was assigned the azimuth of the class it belongs to, and each passed the central point of the original segment. Up to this point, all the line segments of the building under investigation are parallel or perpendicular;
- Adjacent parallel segments were grouped to form one line segment. Such a line had the same azimuth and it passed the calculated central point, which is a weighted average of the central points on the grouped adjacent segments. For segments parallel to each other but not adjacent: if the distance between them was smaller than a pre-defined threshold, they were adjusted to pass through the same line. The intermediate line segments, usually forming a "U" shaped polyline, are kept untouched. The threshold is an experimental value. A value of two meters was employed in this study, which corresponds to two pixels in the DEM grid;
- Regularized building boundaries were calculated by intersecting adjacent line segments.

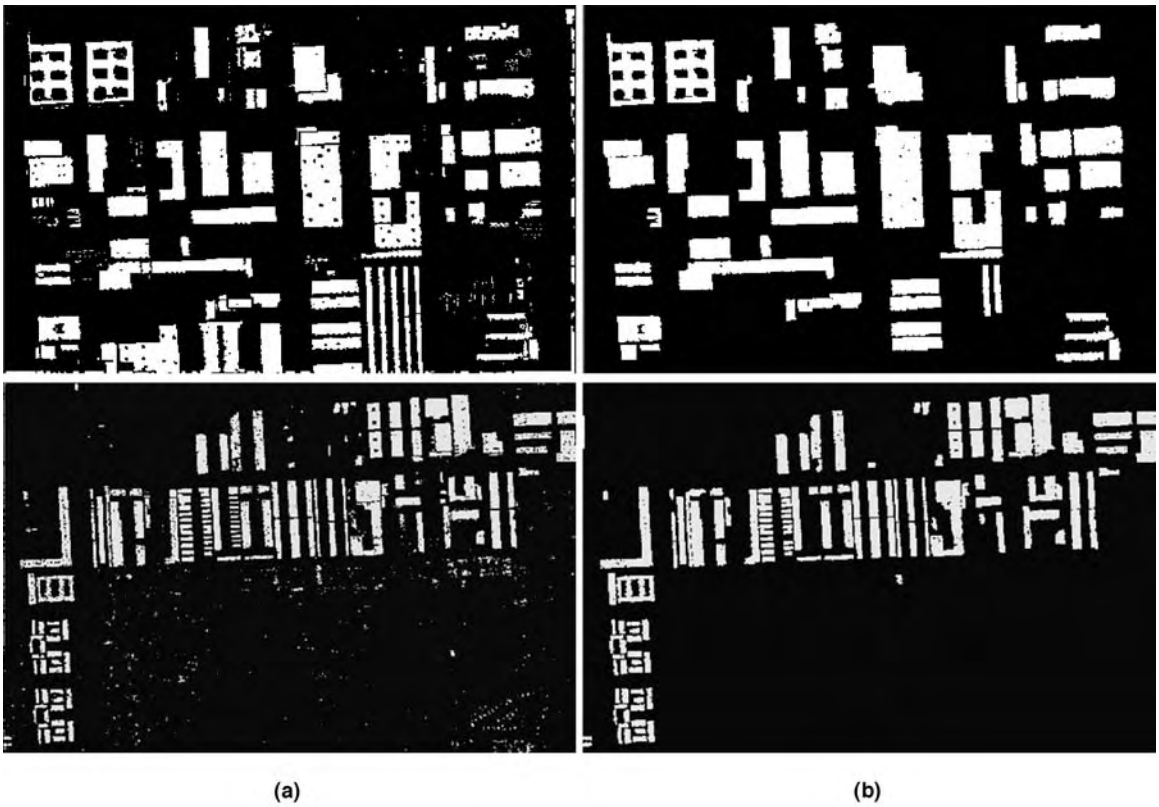


Figure 7. Detected building regions after a morphology open operator (a) and refined building regions using the size constraint, with a morphology dilation operator to compensate the boundary points lost during planar surface fitting calculation (b).

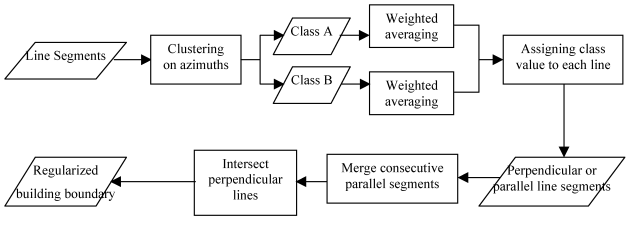


Figure 8. Flow chart for boundary regularization.

Figure 9 shows an instance of building boundary regularization using the method proposed above. The advantage of this method is that it takes global information into account by calculating and adjusting the azimuths using the line segment lengths as weights. This method agrees with the fact that a longer segment has better azimuth accuracy with the same end point position accuracy.

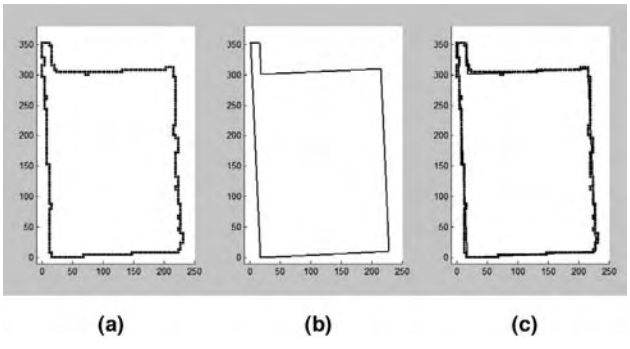
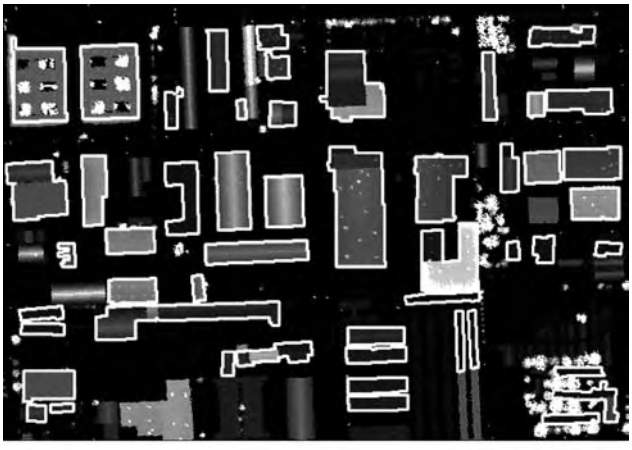


Figure 9. Results of building boundary regularization: the boundary before regularization (a); the boundary after regularization (b); and the comparison of the regularized boundary with the original boundary (c).

Results and Analysis

Figure 10 shows the regularized building boundaries with the background of the DSM generated from the lidar data. It should be noted that there are some buildings missing their boundaries. The reasons for this are:

- Buildings connected to the data boundary were not processed because they were not complete;
- Some buildings have "arch" shaped roofs which can be verified by checking the elevation profile. Thus, they do not match the assumption that a building roof is a planar surface. Consequently, they cannot be detected by the planar-surface fitting method proposed in this paper;
- Some buildings are concealed by trees. As a result, the detected planar surfaces were smaller than the region size threshold employed during building detection. Subsequently, these buildings were eliminated.



(a)

(b)

Figure 10. Regularized building boundaries overlaid on the DSM generated from the lidar data.

The accuracy assessment was conducted by comparing the results with visual interpretation of lidar DSM due to the lack of ground truth. Excluding buildings covered or obscured by trees, around 80 percent buildings were correctly detected in data set one, while around 93 percent in the second data set. One reason for the low accuracy in data set one is that there are several buildings with "arch" shaped roofs. To solve this problem, the threshold of planar facet determination can be relaxed. But certainly, this will introduce occlusion error. Another potential approach is to use complex surface models to detect non-planar roofs.

Conclusions

From the study presented in this paper, the new method for DEM generation and building detection is very efficient and effective, especially in urban and suburban areas. The new boundary regularization method takes global information into account, and thus produces very promising results; the computation expense is much lower than the MDL method. Due to the low point density of the data sets used in this study, one point per square meter, small features like dormers on buildings cannot be recognized. Further study will be conducted on 3D building model reconstruction, detecting buildings with non-planar surfaces, and building detection and reconstruction from the fusion of multiple data resources like lidar, color, and hyper spectral images.

Acknowledgments

The authors would like to express their sincere gratitude to Harris County Flood Control District, Texas for providing the experimental data. The authors are also grateful to Mr. Qian Xiao at Woolpert, Inc. for his support during the study.

References

Ackermann, F., 1999. Airborne laser scanning—present status and future expectations, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (1999):64–67.

Axelsson, P., 1999. Processing of laser scanner data—algorithms and applications, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (1999):138–147.

Baltsavias, E., 1999. A comparison between photogrammetry and laser scanning, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (1999):83–94.

Brunn A., and U. Weidner, 1997. Extracting buildings from digital surface models, *International Archives of Photogrammetry and Remote Sensing*, 32, Stuttgart.

Elberink, S.O., and H.G. Maas, 2000. The use of anisotropic height texture measurements for the segmentation of airborne laser scanner data, *International Archives of Photogrammetry and Remote Sensing*, 33, Amsterdam.

Fua, P., and C. Brechbüler, 1997. Imposing hard constraints on deformable models through optimization in orthogonal subspaces, *Computer Vision and Image Understanding*, 65(2):148–162.

Haala, N., and C. Brenner, 1999. Extraction of buildings and trees in urban environments, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (1999):130–137.

Hu, Y., and C.V. Tao, 2002. Bald DEM generation and building extraction using range and reflectance lidar data, *Proceeding of ACSM-ASPRS 2002 Annual Conference*, Washington, D.C. unpaginated CD-ROM.

Kilian, J., N. Haala, and M. Englich, 1996. Capture and evaluation of airborne laser scanner data, *International Archives of Photogrammetry and Remote Sensing*, 31, B3.

Kraus, K., and N. Pfeifer, 1998. Determination of terrain models in wooded areas with airborne laser scanner data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53 (1998):193–203.

Lee, H., and N.H., Younan, 2003. DEM extraction of lidar returns via adaptive processing, *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):2063–2069.

Lohmann, P., and A. Koch, 1999. Quality assessment of laser-scanner-data, *http://www.ipi.uni-hannover.de/html/publikationen/1999/koch/isprs99%20koch%20lohmann.pdf*, (last date accessed: 13 April 2005).

Lohmann, P., 2001. Segmentation and filtering of laser scanner digital surface models, *International Archives of Photogrammetry and Remote Sensing*, 34, 2.

Lohmann, P., A. Kock, and M. Schaeffer, 2000. Approaches to the filtering of laser scanner data, *International Archives of Photogrammetry and Remote Sensing*, 33, Amsterdam.

Matikainen, L., J. Hyyppä, and H. Hyyppä, 2003. Automatic detection of buildings from laser scanner data for map updating, *International Archives of Photogrammetry and Remote Sensing*, 34, Dresden.

Mayer, S., 2001. Constrained optimization of building contours from high-resolution ortho-images, *IEEE International Conference on Image Processing*, Thessaloniki, Greece.

Morgan, M., and A. Habib, 2002. Interpolation of lidar data and automatic building extraction, *Proceeding of ACSM-ASPRS 2002 Annual Conference*, Washington, D.C., unpaginated CD-ROM.

Rottensteiner, F., and C. Briese, 2002. A new method for building extraction in urban areas from high-resolution lidar data, *International Archives of Photogrammetry and Remote Sensing*, 34, Graz.

Schiewe J., 2003. Integration of multi-sensor data for landscape modeling using a region based approach, *ISPRS Journal of Photogrammetry and Remote Sensing*, 57 (2003):371–379.

Sithole, G., 2001. Filtering of laser altimetry data using a slope adaptive filter, *International Archives of Photogrammetry and Remote Sensing*, 34, 3W4.

Vosselman, G., 2000. Slope based filtering of laser altimetry data, *International Archives of Photogrammetry and Remote Sensing*, 33, Amsterdam.

Weidner, U., and W. Förstner, 1995. Towards automatic building extraction from high-resolution digital elevation models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 50 (1995):38–49.

(Received 16 December 2003; accepted 30 March 2004; revised 21 April 2004)