Xiangyun Hu
C. Vincent Tao
Yong Hu

Geospatial Information and Communication Technology Lab (GeoICT),
Department of Earth and Space Science and Engineering,
York University, 4700 Keele Street, Toronto ON,Canada, M3J 1P3
xyhu@yorku.ca, tao@yorku.ca, yhu@yorku.ca

: Extraction, Road, Multisensor, High Resolution, LIDAR, Urban, Contextual

:

Automated and reliable 3D city model acquisition is an increasing demand. Automatic road extraction from dense urban areas is a challenging issue due to the high complex image scene. From imagery, the obstacles of the extraction stem mainly from the difficulty of finding clues of the roads and complexity of the contextual environments. One of the promising methods to deal with this is to use data sources from multi-sensors, by which the multiple clues and constraints can be obtained so that the uncertainty can be minimized significantly. This paper focuses on the integrated processing of high resolution imagery and LIDAR (LIght Detection And Ranging) data for automatic extraction of grid structured urban road network. Under the guidance of an explicit model of the urban roads in a grid structure, the method firstly detects the primitives or clues of the roads and the contextual targets (i.e., parking lots, grasslands) both from the color image and lidar data by segmentation and image analysis. Evidences of road existing are contained in the primitives. The candidate road stripes are detected by an iterative Hough transform algorithm. This is followed by an procedure of evidence finding and validation by taking advantage of high resolution imagery and direct height information of the scene derived from lidar data. Finally the road network is formed by topology analysis. In this paper, the strategy and corresponding algorithms are described. The test data set is color ortho-imagery with 0.5 m resolution and lidar data of Toronto downtown area. The experimental results in the typical dense urban scene indicate it is able to extract the roads much more reliable and accurate by the integrated processing than by using imagery or lidar separately. It saliently exhibits advantages of the integrated processing of the multiple data sources for the road extraction from the complicated scenes.

Automatic road extraction from remotely sensed imagery has attracted much attention for the last few decades. In this issue, a great number of research papers were published both in geospatial and computer vision communities. In general, road extraction consists of four steps (Gruen and Li, 1995): road sharpening, road finding, road tracking, and road linking. In the earlier research (Bajcsy and Tavakoli, 1976; Nevatia and Babu, 1980), some line detection algorithms were presented for extracting the roads from remotely sensed imagery. There is not much high-level knowledge involved in the methods for road finding. To process gaps bridging, road tracing and handle the complicated image scenes, more sophisticated strategies should be used for more reliable extraction. Knowledge or rule based methods or similar methods based on hypothesis-verification (Mckeown and Delinger, 1988; Tonjes, R., and S. Growe, 1998; Trinder and Wang, 1998) had been used for handling the issue of linear feature alignment and fragmentation. Optimal route search algorithms were frequently employed as semiautomatic road extraction. The optimization can be realized by dynamic programming (Gruen and Li, 1995; Bazohar and Cooper, 1998), snakes (Trinder and Li, 1995; Gruen and Li, 1997; Tao ., 1998; Agouris 2001) and Kalman filtering (Vosselman and de Knecht , 1995). Furthermore, contextual information supported methods (Stilla, 1995; Baumgartner 1999) were applied to extract road more reliably. Actually we can also find many strategies (Bazohar and Cooper, 1998; Couloigner and Ranchin 2000; Laptev 2000; Katartzis, , 2001, Hinz and Baumgartner; Hu and Tao, 2003; solved straightforward. Most of the methods applied to extract roads from open or rural areas were successful to some extent due to the relative simple image scene and road model. For the extraction of roads in dense urban areas, especially from high resolution imagery, there are primary obstacles which lead to unreliable extraction results: complicated image scene and road model, furthermore, occlusion caused by high buildings and their shadows. In other words, the lack of information, especially three-dimensional information is the principle difficulty in obtaining the road information with high reliability and accuracy in the urban scenes.

for ground materials. The relative separations between ground features (i.e., asphalt road, grass, building and tree) have been compared using intensity data. It is found that the separabilities are very high for road vs. grass and road vs. tree (Song et al., 2002). In many cities, road networks are arranged in a grid structure in urban areas. These grid roads are mainly composed of parallel and orthogonal straight roads with respect to the main orientation of the network. The existence of streets can be detected much more easily from the arrangements than from imagery in which the highly complicated image content and lack of information lead to high complexity of direct extraction of the street network. It is recognized that the simple geometry and topology relations among grid streets may be used to improve the reliability of road extraction results significantly. As mentioned above, instead of using imagery, using lidar data can be easier to extract the road primitives in built-up areas, while imagery can also be used for additional information for verification and accurate extraction. Many clues of road existence can be obtained from high resolution imagery. The motivation of this paper is to explore the strategy and methodology of integrated processing of lidar data and high resolution imagery in order to obtain reliable road network information from the dense urban environment. In the followed section, the case study data is introduced and the overall strategy of the processing is given. The third section describes the road extraction methods by using of these two source of information, including road area segmentation, road clue detection and verification, fusion of the clues from the two data sources. The case study result is presented and conclusion remarks are then given.

In early 2002, Optech International, Toronto completed a flight mission of acquiring the lidar data of Toronto urban area using its ATLM 3200. The lidar dataset provided is around downtown region. The roads in the study area are coated with asphalt with pebbles or concrete. The first and last returns lidar range and intensity data were collected. The dataset contains about 10.6 million points and has a density of about 1.1 points/m^2. We generate the DTM using the last-return lidar range data, and also obtain the height data by subtracting the DTM from the range data (Hu, 2003). The height data contains height information that has removed the retain relief relative to the bare Earth, and puts all the ground features on a flat reference plane. Figure 1 (a) and (b) shows the first-return intensity data and the height data. The high resolution imagery is obtained from ortho-rectified aerial image of the same area. The image resolution is 0.5m. To do integrating processing, it is re-sampled into 1m resolution and is manually registered with the lidar data in geometry. Figure 1 illustrated the lidar data of the area. Figure 1 (c) shows an image window of the high resolution imagery. Its size is 1024 by 1024 pixel. Considering the computational cost, we carry out our extraction in this selected area, which demonstrates typical scene of dense urban area. It contains buildings with great height, roads (streets) and many kinds of typical ground objects (parking lots, grass land, trees, vehicles etc.). It is feasible to testify our method.

. data ,

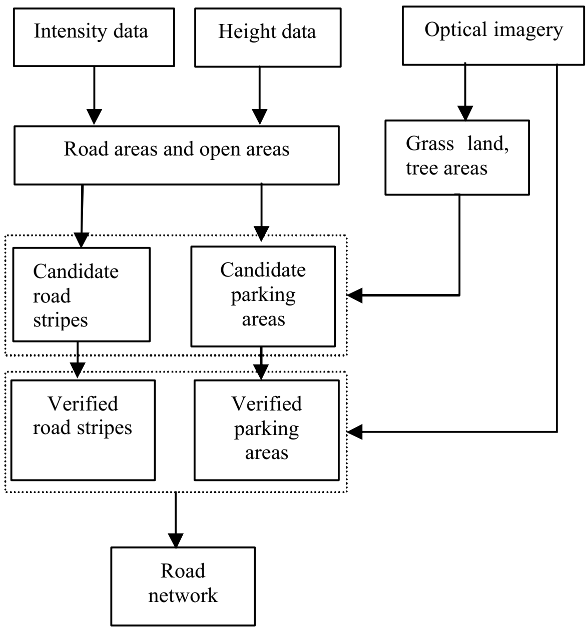


Figure 2. Workflow of integrated processing for road extraction from urban areas

We separate roads from trees, buildings and grasslands with minimum misclassification fusing the intensity and height data. In reflectivity, the spectral signature of asphalt roads significantly differs from vegetation and most construction materials. The reflectivity rate of asphalt with pebbles is 17% for the infrared laser, and no other major materials have a close reflectivity rate. In height, pavements are attached to the bare surface and appear as smooth ribbons separating the street blocks in a city.

It can be easily found that integrating intensity and height data may produce reliable road detection results. On the one hand, the intensity provides the spectral reflectivity, which can help identify most roads even if the objects coated by the same material are also included. On the other hand, the height data can help identify most non-building and non-forest areas even if those low open areas such as grasslands are also included. Using height information, the built-up areas with higher elevations than their surroundings will be safely removed; while using the (first-return) intensity information, the vegetated areas are easily removed. In detail, compared to roads, grasslands have different intensity although they have low elevation, trees have different values in both intensity and height, and buildings have high structures with elevation jumps although they may be coated rainproof asphalt.

After segmentation of the lidar data, the possible road areas and other areas are converted to a binary image. Figure 3 shows the segmented data. Parking lots are kept because of same reflectance and low heights as roads, and bridges and viaducts are removed because of their large heights.

From the true colour high resolution imagery, the grass lands and tree areas can be separated from the open areas. First, because the roads and parking areas are covered and coated by concrete or rainproof asphalt, the saturation of the pixels of the areas is low while in the grass lands and tree areas it is high and the hue tends to be 'green'. So using a threshold the grass lands and tree areas can be separated from the low saturation areas. Subtracting the grass lands and tree areas, we can obtain the areas containing candidate road stripes and parking areas.

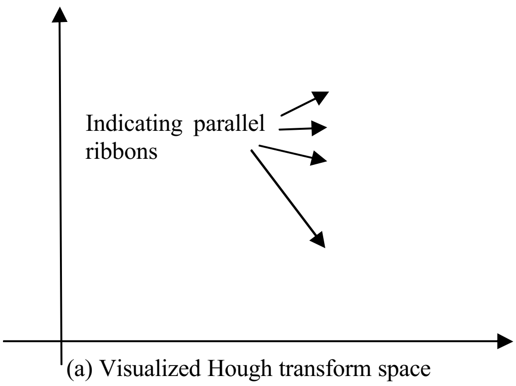
Figure 3. Extracted open areas (white) containing road stripes and other areas

The streets demonstrate ribbon features in geometry. We used a modified Hough transfer method to directly detect the candidate stripes of the streets from the segmented lidar data – the binary image. Hough transformation is frequently used for extracting straight lines. When we treat a ribbon as a straight line with the width of the street, traditional Hough transfer can be used for the detection of the streets. Figure 4 shows the Hough space after once transfer. The space is formed using the straight line as given by:

$$r = \quad \cos q + \quad \sin q \qquad (1)$$

where q is the angle of the line's normal with the x-axis; r is the algebraic distance from the origin to the line.

Instead of detecting the peak points in the transfer space, we detect the 'maximal bars' as pointed out in Figure 4. To detect all possible ribbons, first step is to determine the primary direction of the street grid. The parallel ribbons and ribbons with right angle crossing to them are also extracted. The extraction is conducted directly from the segmented binary image on contrast to extraction from 'thinned' ribbon, and the width can be estimated roughly by the bar width (the difference of). We iteratively carry out the Hough transform. In each step of transform, we only detect a maxima response in the Hough space, and then the extracted stripe pixels are removed from the binary image. This will reduce the influence of multiple peaks in the transform space. The iteration will be terminated by the trigger criteria of the maxima that indicates the length of the stripe.



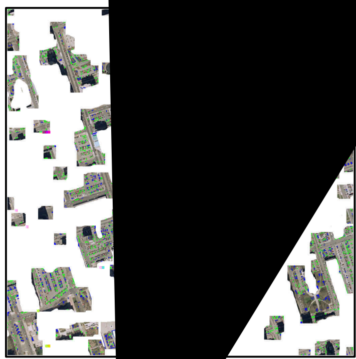
(a) Visualized Hough transform space

(b) Detected road stripes displayed on the image

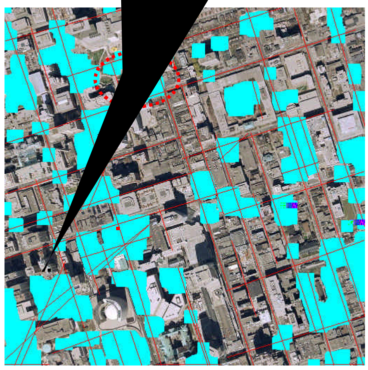
Figure 4. Road stripe detection by Hough transform from segmented lidar data

The detected primary streets by Hough transform are possible streets and just straight line equations (parameters). To form a real street 'grid', we should identify the candidates and remove some wrong segments. The first step is to overlay the straight lines onto the binary image. For each line, break it to be segments where it transverses building areas. It can be fulfilled simply by the binary image. Thereafter, each verified line segment is adjusted by geometric correction — to move it to be in the street centre where the dual distance between it and the building edge is equivalent.

We judge that the short segments going through the big open areas are with low possibility of being a part of the street and high possibility of being a parking area. To verify a parking area, we employ the vehicle clue to confirm the area. The vehicles are extracted by a pixel based classification method. Some samples of vehicles are provided by manual digitization, and they are used for learning the pixel intensity value of the vehicles. The possible pixels of the vehicles in the road and parking areas are shown in green and blue colours in the Figure 5 (a). In the study, the open areas contain roads and parking lots. We assume a region with nearly squared shape and big area has high possibility of being parking lots. A morphologic operation is applied to the binary image to detect the big open areas. In Figure 5 (b), the highlighted areas are possible open areas rather than roads, but the roads could go though the area. Combining the analysis result of shape and vehicle clue from lidar data and the optical imagery, we compute the 'score' of an open area of being a parking lot. The high score indicates the high possibility of being parking lot. By computing the length of the segment which goes through the parking area, the segments mostly lie in the parking areas a [redacted] the short segments cir [redacted] most likely gong throug [redacted]



(a) Vehicles [redacted] of parking areas



(b) Verified road stripes displayed on the image

Figure 5. Road stripes and parking areas verification

The road topology is formed by intersecting the road segments extracted from the previous steps, as shown in Figure 6.

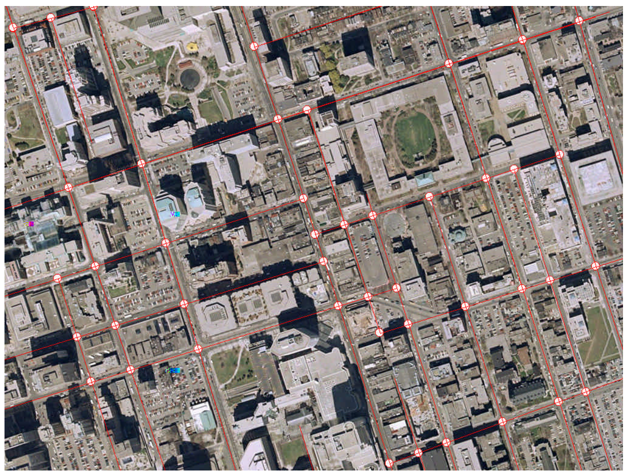


Figure 6. Road grid formation

We develop a road extraction method using lidar data and high resolution optical images. The method tackles the problem of extracting grid roads in urban areas with dense buildings. Using lidar data, the difficulty of resolving the occlusion of roads in optical images is eliminated. It demonstrates the potential and power of using lidar data to extract information from complicated image scenes. To obtain more reliable results, image analysis (to detect contextual objects: grasslands, parking lots, vehicles etc.) for contextual information extraction is integrated into the whole procedure. It greatly improves the final results in correctness and accuracy. The work described in this paper clearly indicates that involving multiple source of information will definitely improve the extraction results in the complicated scene. Future work will include testing the method using more datasets and developing algorithms of adaptive threshold determination in the multi-step processing, which will be a challenging work.

: The authors are grateful for the support of a GeoIDE project, Canada: 'Automating 3D Feature Extraction and Change Detection from High Resolution Satellite Imagery with Lidar and Maps'

Agouris, P., A. Stefanidis, and S. Gyftakis, 2001. Differential snakes for change detection in road segments, , 67(12):1391-1399.

Alharthy, A., Bethel, J., 2003. Automated road extraction from lidar data, , May 5-9, Anchorage, Alaska (American Society for Photogrammetry and Remote Sensing, Anchorage, Alaska), unpaginated CD-ROM.

Barzohar, M., and D.B. Cooper, 1996. Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation, , 18(7):707-721.

Baumgartner, A., C. Steger, H. Mayer, and W. Eckstein, 1999. Automatic road extraction based on multi-scale, grouping, and context. , 65(7):777-785.

Couloigner, I., and T. Ranchin, 2000. Mapping of urban areas: a multiresolution modeling approach for semi-automatic extraction of streets, , 66(7):867-874.

Fiset, R., F. Cavayas., M. Mouchot., B. Solaiman and R. Desjardins, 1998. Map-image matching using a multi-layer perceptron: the case of the road network, 53(2):76-84.

Gruen, A., and H. Li, 1997. Semiautomatic linear feature extraction by dynamic programming and LSB-snakes, , 63(8):985-995.

Hinz S., A. Baumgartner, 2003. Automatic extraction of urban road networks from multi-view aerial imagery, , 58: 83-98.

Hofmann, P., 2001. Detecting buildings and roads from IKONOS data using additional elevation information, , 6: 28-33.

Hu X., C.V. Tao, 2003. Automatic extraction of main-road centerlines from high resolution satellite imagery based on perceptual grouping. , May 5-9, Anchorage, Alaska (American Society for Photogrammetry and Remote Sensing, Anchorage, Alaska), unpaginated CD-ROM.

Hu X., C.V. Tao, 2004. Automatic Highway Extraction from High Resolution Imagery by an Energy Minimizing Based Perceptual Grouping Method, , 2004, Vol 58 (1), 41-50.

Hu, Y., 2003. Automated extraction of digital terrain models, roads and buildings using airborne lidar data, PhD dissertation, Department of Geomatics Engineering, University of Calgary, 206 p, URL: http://www.geomatics.ucalgary.ca/links/GradTheses.html.

Katartzis A., H. Sahli, V. Pizurica, and J. Cornelis, 2001. A Model-Based Approach to the Automatic Extraction of Linear Features from Airborne Images, , 39(9): 2073-2079.

Laptev I, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, 2000. Automatic extraction of roads from aerial images based on scale space and snakes , 12(1):23-31.

Song, J.H., Han, S.H., Yu, K., Kim, Y., 2002. Assessing the possibility of land-cover classification using lidar intensity data, , 9-13 September, Graz, vol. 34, 4 p.

Steger, C., 1998. An unbiased detector of curvilinear structures, , 20(2):113-125.

Tonjes, R., and S. Growe, 1998. Knowledge based road extraction from multisensor imagery, , 32(Part3/2):387-393.

Trinder, J.C., and Y. Wang, 1998. Automatic road extraction from aerial images , 8 (4):215-224.

Vosselman, G., and J. de Knecht, 1995. Road tracing by profile matching and Kalman filtering, (A. Gruen, O. Kuebler, and P. Agouris, editors), Birkhaeuser Verlag, Basel, pp. 265-274.

Zhang, C., Baltsavias, E., Gruen, A., 2001. Knowledge-based image analysis for 3D road reconstruction, , 1(4): 3-14.