Seamless Fusion of LiDAR and Aerial Imagery for Building Extraction

Guoqing Zhou, *Senior Member, IEEE*, and Xiang Zhou

***Abstract*—Although many efforts have been made on the fusion of Light Detection and Ranging (LiDAR) and aerial imagery for the extraction of houses, little research on taking advantage of a building's geometric features, properties, and structures for assisting the further fusion of the two types of data has been made. For this reason, this paper develops a seamless fusion between LiDAR and aerial imagery on the basis of aspect graphs, which utilize the features of houses, such as geometry, structures, and shapes. First, 3-D primitives, standing for houses, are chosen, and their projections are represented by the aspects. A hierarchical aspect graph is then constructed using aerial image processing in combination with the results of LiDAR data processing. In the aspect graph, the note represents the face aspect and the arc is described by attributes obtained by the formulated coding regulations, and the coregistration between the aspect and LiDAR data is implemented. As a consequence, the aspects and/or the aspect graph are interpreted for the extraction of houses, and then the houses are fitted using a planar equation for creating a digital building model (DBM). The experimental field, which is located in Wytheville, VA, is used to evaluate the proposed method. The experimental results demonstrated that the proposed method is capable of effectively extracting houses at a successful rate of 93%, as compared with another method, which is 82% effective when LiDAR spacing is approximately 7.3 by 7.3 ft^2. The accuracy of 3-D DBM is higher than the method using only single LiDAR data.**

***Index Terms*—Aerial image, extraction, house, image processing, light detection and ranging (LiDAR), urban.**

I. Introduction

TO MEET the increasing requirement for accurate 3-D models of buildings, continuous updates for urban planning, cartographic mapping, and civilian and military emergency responses, the efforts for building extraction using a variety of data sources have been made in the past in photogrammetry, image processing, and computer vision communities. The previous methods can be grouped into three categories in terms of the employed data sources. The first category uses optical imagery, such as aerial imagery and high-resolution satellite imagery (e.g., IKONOS). The second category utilized active imagery data, such as Light Detection and Ranging (LiDAR) data and radar imagery. The third category integrates both high-resolution optical and active imagery data.

For the first category, the earliest effort can be traced back to the 1980s. For example, McKeown [26], [27] presented a knowledge-based interpretation method to recognize buildings from aerial imagery. Afterward, Matsuyama [24] developed a so-called *expert system (ES)* for building extraction. Mohan and Nevatia [28] proposed matching multiple-image technique to extract 3-D urban buildings. Lee and Lei [22] extracted house roofs through assumed gray matter on a house's surface and matching house corners and edges. Baillard and Maître [3], Fischer *et al.* [10], Kim and Muller [21], Mayer [25], and Paparoditis *et al.* [31] also presented their methods. In the 21st century, methods for the building extraction from optical imagery are still continuously explored. For example, Cord *et al.* [6], Fradkin *et al.* [11], and Jaynes *et al.* [17] presented the methods of urban buildings using multiple aerial images in dense urban areas. Jung [18], Khoshelham *et al.* [20], and Suveg and Vosselman [44] proposed detecting complex buildings from multitemporal aerial stereopairs. Peng *et al.* [32] used snake models for building detection from aerial images. Ahmadi *et al.* [1], Le Besnerais *et al.* [4], Sirmacek and Unsalan [40], and Zebedin *et al.* [49] presented their methods to extract buildings. These early methods attempted to extract the boundary of a building to identify the buildings, but separating building boundaries from non-building objects, such as parking lots, is difficult. It is especially difficult to extract the complete boundaries of a building from single optical image due to unavoidable occlusions, poor contrasts, shadows, and disadvantageous image perspectives [9].

The second category employs active imagery data, such as microwave imagery [41] and LiDAR data [13]. Particularly, the emergence of the LiDAR technique in the middle of 1990s resulted in considerable efforts for 3-D building extraction in the past decades. The methods can be grouped into two categories [51], i.e., the classification approach and the adjustment approach. The classification approach detects the ground points using certain operators designed based on mathematical morphology [29], terrain slope [2], or local elevation difference. The refined classification approach uses triangulated irregular network data structure [2] and iterative calculation [42] to consider the discontinuity of the LiDAR data or terrain surface. The purpose of these methods is to first form a grid LiDAR elevation into a depth imagery and then use image segmentation

Manuscript received October 24, 2013; accepted February 16, 2014. Date of publication April 23, 2014; date of current version May 30, 2014. This work was supported in part by the National Science Foundation (NSF) under Contract NSF 0131893, by the National Natural Science Foundation of China under Contract 41162011, by Guangxi Grand Natural Science Foundation under Contract 2011GXNSFD018001, and by the Grant of GuangXi Key Laboratory of Spatial Information and Geomatics under Contract GuiKeNeng110-31-08-01. The image segmentation via integrating LiDAR and image data processing was implemented by Ph.D. student, C. Song.

The authors are with GuangXi Key Laboratory for Geospatial Informatics and Geomatics, Guilin University of Technology, Guilin 541004, China (e-mail: gzhou@glut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2311991

0196-2892 © 2014 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

techniques to detect building footprints. The adjustment approach essentially uses a mathematical function to approximate the ground surface, which is determined in an iterative least squares process while the outliers of non-ground points are detected and eliminated. For example, Pu and Vosselman [33] proposed a knowledge-based reconstruction of building models from laser scanning data. Rutzinger *et al.* [37] made an experimental comparison and analysis for evaluating several techniques of building extraction. Sampath and Shan [38] and Zhang *et al.* [50] separated the ground and the non-ground LiDAR measurements using a progressive morphological filter. Building points were labeled via a region-growing algorithm, and the building boundaries were derived by connecting the boundary points. The major shortcoming of applying single active data is that texture and boundary information cannot be effectively reflected.

The third category uses both optical and active data since the building extraction from only monocular imagery (e.g., optical imagery, or LiDAR data) cannot reach a satisfactory result due to the strengths and weaknesses of each data source. Therefore, efficient fusion of different types of data sources can utilize the natural complementary properties of each data source and compensate for the weaknesses found from one data source to another. For instance, Hermosilla *et al.* [15], Kabolizade *et al.* [19], Zhou *et al.* [51], Yu *et al.* [47], and Zhou [52] presented the method for extraction of urban houses and road networks through the integration of LiDAR and aerial images. Habib *et al.* [14] and Schenk and Csatho [39] proposed feature-based fusion of LiDAR data and aerial imagery to obtain a better building surface description. Sohn and Dowman [43] focused on the exploitation of synergy of IKONOS multispectral imagery information combined with a hierarchical segmentation of a LiDAR digital elevation model (DEM). Rottensteiner *et al.* [35] proposed a method consisting of building detection, roof plane detection, and the determination of roof boundaries. Fujii and Arikawa [12] and Jinhui *et al.* [16] proposed integrating LiDAR, aerial image, and ground images for complete urban building modeling. O'Donohue *et al.* [30] combined thermal–LiDAR imagery for the extraction of urban man-made objects. Dong *et al.* [8], Mastin *et al.* [23], Wang and Neumann [45], and Zabuawala *et al.* [48] suggested automatic registration of LiDAR and optical imagery of urban scenes for the extraction of houses and roads.

Recently, the methods of third category have been widely used since the efficient fusion of the two types of data sources can compensate for the weaknesses of one data to another. However, most current efforts have still been staying the study of the "gray-level" fusion and few efforts on utilization of the geometric or chromatic features of buildings to support the deep data fusion for building extraction. To confront these issues, this paper develops an aspect graph-driven method to explore how the features of houses (e.g., geometry, structures, and shapes) are utilized for the data fusion between aerial imagery data and LiDAR data. The arrangement of this paper is as follows. Section II presents the aspects, aspect graphs, and corresponding data structures. Section III highlights the experiment and comparison analysis. Section IV draws up the conclusions.

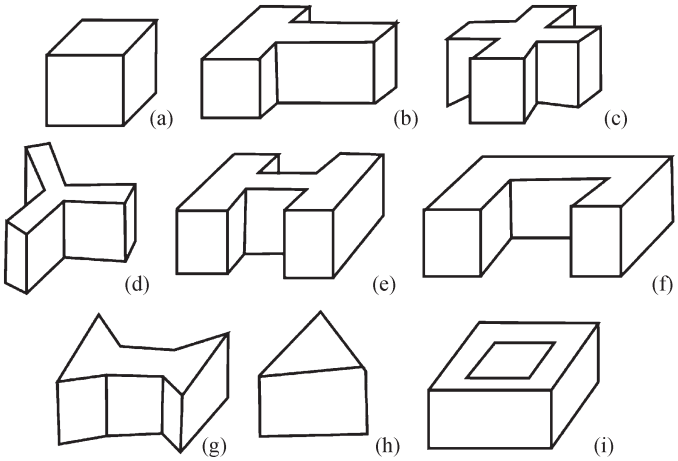

Fig. 1. Three-dimensional primitives of houses with flat roofs.

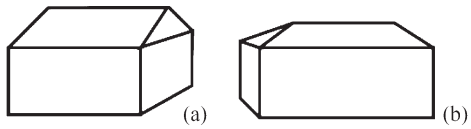

Fig. 2. Three-dimensional primitives of houses with convex roofs.

II. Principle of Aspect Code and Creation of Aspect Code Database

The concept of aspects and aspect graph techniques was first presented by Biederman [5] and then applied by Dickinson *et al.* [7] for 3-D shape recovery of industrial objects. This paper extends the early idea through developing aspects and aspect graphs of urban houses. The major work consists of the steps below.

A. Selection of the 3-D Primitives (Houses)

Since the structures of houses in urban area are intricate, the initial research in this paper only considered nine types of houses with flat roofs (see Fig. 1) and two types of houses with convex roofs (see Fig. 2) as 3-D house primitives.

B. Creation of 2-D Aspects

Dickinson *et al.* [7] employed an aspect to represent the projection of an entire 3-D industrial object, but this representation is not proper for 3-D houses due to occlusion caused by the perspective projection in aerial imagery. This paper expands this method through the use of an aspect to represent the set of topological distinct views of a 3-D house primitive. The basic idea is that an aspect represents either an entire projection or part of projection of a 3-D house primitive. Each aspect consists of a set of 2-D faces, which correspond to the surface of a primitive. An object can be modeled by a set of aspects. The relationship between 2-D aspects and 3-D primitives is depicted through an aspect hierarchical graph, which consists of three layers on the basis of the face's appearance in the aspect set (see Fig. 3).

Boundary Set: A boundary set represents all subsets of lines and curves composing the faces. A complete set of nine

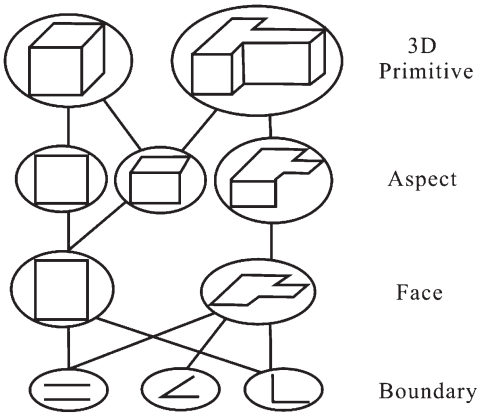

Fig. 3. Hierarchical aspect representation.

Fig. 4. Complete set of nine boundaries and their aspect coding regulation.

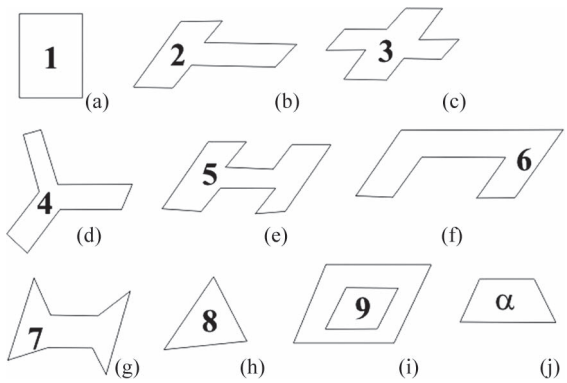

Fig. 5. Complete set of ten faces for the houses selected in Figs. 1 and 2, of which the algebraic number is the assigned code of the aspect.

boundaries is listed in Fig. 4 on the basis of the possible appearances of the selected 3-D house primitives, in which the nonlinear edges of houses are considered. The nine-boundary set represents nonaccidental contour relations such as connecting manner, parallelism, and symmetry occurring within a face.

Faces: Faces represent the set of the surface projected of houses in the aerial imagery, such as contour boundary and the relationship between the contours. Each face differs in the number of contours, the types of contours, or nonaccidental relationship between the contours. Houses are segmented into many various faces. A complete set of ten faces is identified for representing the chosen 3-D house primitives (see Fig. 5).

Aspect: An aspect is constructed by a group of faces. The aspects of houses can be formed by any surface connecting with a wall. An aspect also describes the relationship of face adjacencies and indicates the contours shared by adjacent faces.

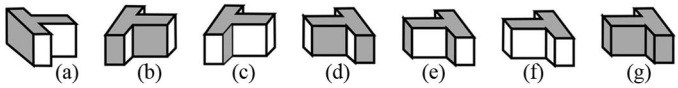

Fig. 6. Code merging regulation for houses with flat roof and wall.

C. Coding Regulation for Aspect

To extract the houses using the aspect interpretation method, aspect coding regulation has to be formulated. With the analysis of the appeared aspects in Section II-B, nine types of contour set, as shown in Fig. 4, are selected for coding. The coding regulation is formulated as follows.

1) Code of a straight line is $\mathbf{a}$ [see Fig. 4(a)].
2) Code of a curve is $\mathbf{b}$ [see Fig. 4(b)].
3) Code of the intersection of two straight lines is $\mathbf{aa}$ [see Fig. 4(c)].
4) Code of the intersection of a straight line and a curve is $\mathbf{ab}$ [see Fig. 4(d)].
5) Code of the intersection between two curves is $\mathbf{bb}$ [see Fig. 4(e)].
6) Code of two parallel straight lines is $\mathbf{a}-\mathbf{a}$ [see Fig. 4(f)].
7) Code of two parallel curves is $\mathbf{b}-\mathbf{b}$ [see Fig. 4(g)].
8) Code of a straight line opposite to another curve is $\mathbf{a}-\mathbf{b}$ [see Fig. 4(h)].
9) Code of a curve opposite to another curve is $\mathbf{b}-\mathbf{b}$ [see Fig. 4(i)].

All the basic aspects in Fig. 4 are labeled through the codes formulated above. Once these basic aspects are coded, their codes are archived in a so-called *codebase* for use of the aspect interpretation, which will be described in Section II-G. Using these basic aspects and their codes above, a complex house can be coded. For example, for the house in Fig. 1(b), its probable projections are depicted in Fig. 6 due to varying the positions of viewpoint. As observed, only the house's roof needs to be considered for aspect coding since all the walls of a house are perpendicular to the ground in the selected 3-D house primitives. Therefore, the following three additional rules for coding a house are formulated.

1) The total edge number of a face is defined as the first figure of a code. For example, if a face consists of eight edges, the first figure of a code is $\mathbf{8}$. Others are based on this analogy.
2) If all edges of a house consist of straight lines, the second figure is assigned $\mathbf{a}$ or otherwise assigned as $\mathbf{b}$.
3) The edges of a face is numbered clockwise, starting from any edge, using natural numbers, 1, 2, ..., n. If the ordered edges with even number are parallel to each other, a number $\mathbf{1}$ is assigned to the end of code with a dash line. Otherwise, $\mathbf{0}$ is assigned.

With the coding rules formulated above, a complete set of the code for 3-D house primitives selected in Fig. 2 is as follows:

$$1.\,\mathbf{4a-1};\quad 2.\,\mathbf{8a-1};\quad 3.\,\mathbf{12a-1};\quad 4.\,\mathbf{12a-0};\quad 5.\,\mathbf{12a-1};$$
$$6.\,\mathbf{7a-1};\quad 7.\,\mathbf{8a-0};\quad 8.\,\mathbf{3a-0};\quad 9.\,\mathbf{2-4a-1}.$$

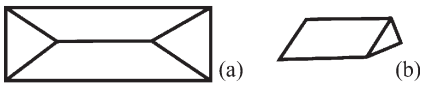

Fig. 7. Code merging regulation for houses with a convex roof.

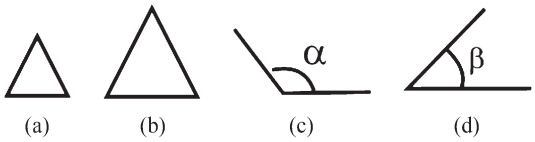

Fig. 8. Size invariance and angle invariance of the aspects.

D. Coding Regulation for Aspect Merging

The above coding regulations are for the complete projection of a house in an aerial imagery. However, a house is rarely in practice projected into a complete 2-D aspect due to unavoidable occlusion, poor contrast, shadow, and perspective projection of aerial imagery. Moreover, the houses with a convex roof consist of several surfaces. Thus, the coding regulation for merging multiple aspects is formulated. The basic idea is as follows: For a house with flat roof, its aspect is to merge a roof aspect with multiple wall aspects, whose aspects are $\mathbf{4a}$ in the selected houses. The coding regulation for merging two aspects is to merge two ordinal aspect codes into a new code with a dash line connection. For example, the emerged code of the house shown in Fig. 6(a) is $\mathbf{2-1}$, where $\mathbf{2}$ stands for the aspect $\mathbf{2}$ in Fig. 5(b), whose code of visible face is $\mathbf{8a-1}$, and $\mathbf{1}$ stands for aspect $\mathbf{1}$ in Fig. 5(a), whose code of visible face is $\mathbf{4a}$. If an entire 2-D aspect consists of several visible faces, the coding regulation is the same as that described above. For example, the house shown in Fig. 6(g) consists of a flat roof and four visible walls, thus, its code is $\mathbf{2-1-1-1-1}$.

According to the coding regulation of aspect merging, a house with convex roof, as depicted in Fig. 7(a), its merged code is $\mathbf{8-\alpha-8-\alpha}$, where $\mathbf{8}$ stands for the aspect $\mathbf{8}$ in Fig. 5(h), and $\boldsymbol{\alpha}$ stands for the aspect $\boldsymbol{\alpha}$ in Fig. 5(j). If these aspects are merged with walls, the coding regulations are the same as those formulated for flat houses.

E. Discussion of the Proposed Coding Regulation

The proposed coding regulation above has two apparent properties, namely, size invariance and angle invariance.

- *Size invariance*: The proposed coding regulation ignores the size of an aspect. For example, Fig. 8(a) and (b) have different sizes of the aspect, but they have the same code. This characteristic is called size invariance of the aspect.
- *Angle invariance*: The proposed coding regulation also ignores the size of angles. For example, Fig. 8(c) and (d) have different angle sizes of intersection, but they have the same code. This characteristic is called angle invariance of the aspect.

In addition to the above two basic characteristics, it has been observed that the proposed coding regulations have rotation invariance and scale invariance. Therefore, the characteristics of the aspect coding regulation are summarized as follows.

1) Aspect is based on a view-centered [see Fig. 9(a)–(f)],

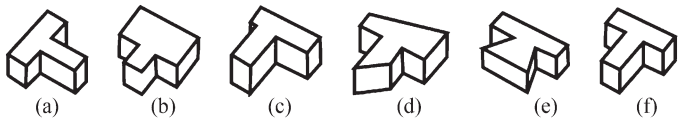

Fig. 9. Properties of aspect invariance in the coding regulation.

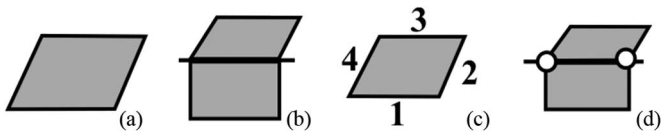

Fig. 10. Basic regulation of aspect graph construction.

2) Aspect is a rotate invariance [see Fig. 9(a) and (f)],
3) Aspect is an angle invariance [see Fig. 9(c)–(e)], and
4) Aspect is a scale invariance [see Fig. 9(a) and (f)].

With the above characteristics, it can be noted that different houses probably have the same codes such as Fig. 1(c) and (e). These characteristics do not actually affect the house extraction since our goal is to extract the houses, rather than recognize the type of a house.

F. Creation of Aspect Graphs

The work above only created the 2-D aspects of the houses, but it is not sufficient to correctly interpret the complex houses or partially projected house caused by the occlusions if only using the above 2-D aspect codes. For this reason, creation of a hierarchical aspect graph using nodes and arcs is proposed in this paper. The hierarchical aspect graph consists of three layers, i.e., face-aspect graph, primitive-aspect graph, and aspect hypergraph. In the hierarchical aspect graph, a node represents the 2-D aspect of an entire 3-D object and an arc represents the relation between two 2-D aspects. A face is chosen as a node (element) to construct a face-aspect graph; a primitive graph is used to describe an entire projection or partial projection of a 3-D primitive; and an aspect hypergraph is constructed through the description of the connected relations between primitive graphs.

During the construction of the hierarchical aspect hypergraph, a critical step is the creation of the face-aspect graph. To this end, the following regulations are formulated.

1) A basic pattern, as shown in Fig. 10(a), is taken as a node.
2) A connected relation between two basic patterns is taken as a connected arc [see Fig. 10(b)].
3) The information such as type, edge number, and other factors in the basic pattern is taken as the attributes of the node [see Fig. 10(c)].
4) The connected information is taken as the connection attributes of the arc [see Fig. 10(d)].

The two nodes are connected using an arc. The arc is assigned with attributes. The attribute values are suggested as follows.

1) The regulations for attribute value assignment of the nodes are as follows:
   - The area of a face aspect is taken as an attribute value.

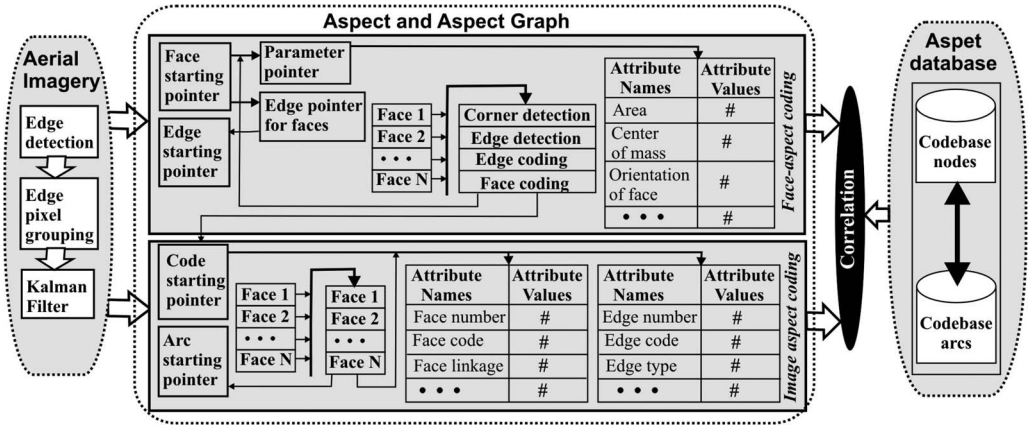

Fig. 11. Data structure of aspect database (codebase) creation.

- The perimeter of a face aspect is taken as an attribute value.
- The center coordination (X_c, Y_c) of a face aspect is taken as an attribute value.
- The orientation angle θ of a face aspect is taken as an attribute value, where

$$\theta = \tan^{-1}\left[2 \cdot M_{11}/(M_{20} - M_{02})\right]/2 \tag{1a}$$

where $M_{i,j}$ is the ith- and jth-order moment, which can be expressed by

$$M_{i,j} = \sum_{(X,Y)\in R} (X_c - X)^i (Y_c - Y)^j. \tag{1b}$$

- The recognized primitive through the above aspect interpretation method is taken as an attribute value.
- The codes describing an aspect are taken as an attribute value.

2) The regulations for attribute value assignment of the arcs are:
   - The number of the shared edge(s) is taken as an attribute value.
   - The type of the connected edge(s), straight or curve, is taken as an attribute value.
   - The slope of a straight line is taken an attribute value.
   - The curvature of a curve is taken an attribute value.
   - The information of two endpoints of the connected line is taken an attribute value.

With the regulations formulated above, the face-aspect graph and the primitive-aspect graph corresponding to describing the face aspect and the primitive aspect can be constructed. Furthermore, with the description of the connected relationship between primitive graphs, an aspect hypergraph can be constructed. The implementation of the above method and algorithms is called *aspect describer*.

G. Creation of Aspect Codes Database

The codes describing the aspects and the hierarchical aspect graphs (i.e., face-aspect graph, primitive-aspect graph, and aspect hypergraph) are taken as knowledge and archived in a database called *codebase* (see Fig. 11). The data structures for archiving the knowledge, associated with the aspect interpretation, are depicted in Fig. 11.

As shown in Fig. 11, with the established aspect codebase, house extraction using the aspect interpretation is carried out through correlated operation between the aspect codes created from aerial imagers (called *image aspect codes*) and the aspect codes archived in the codebase.The basic steps of implementing the operation are as follows:

1) The segmented aerial imagery inputs into the aspect describer.
2) The aspect describer creates the codes for houses and their connection using the aspect and the attributes in accordance with the coding regulations formulated in Sections II and III.
3) The aspect interpretation is conducted. First, the face aspect is created, and the faces describing houses are interpreted using the face-aspect correlation operation. With the implementation of face-aspect interpretation, a primitive graph, which describes an entire projection or partial projection of a 3-D primitive house, is created. The 3-D primitive houses are interpreted through the correlation operation between the image aspect graph and the codebase aspect graph in combination with LiDAR data processing. Finally, an aspect hypergraph is created, and the 3-D houses are further interpreted through the correlation operation between the image hypergraph and the codebase hypergraph with assistance of LiDAR data processing. The corresponding data structures for the implementation of the aspect interpretation proposed above are depicted in Fig. 11.

III. Experiments and Analyses

A. Data Set

Virginia Department of Transportation, contracting to Woolpert LLC at Richmond, Virginia, has established a high-accuracy test field in Wytheville, VA, for the accuracy evaluation of the LiDAR system. The field extends from the west side of Wytheville approximately 11.4 miles east, with a northern and southern extent of approximately 4.5 miles (see Fig. 12). All of the elevation data, including LiDAR and 19 ground control points (GCPs), were referenced to NAVD88

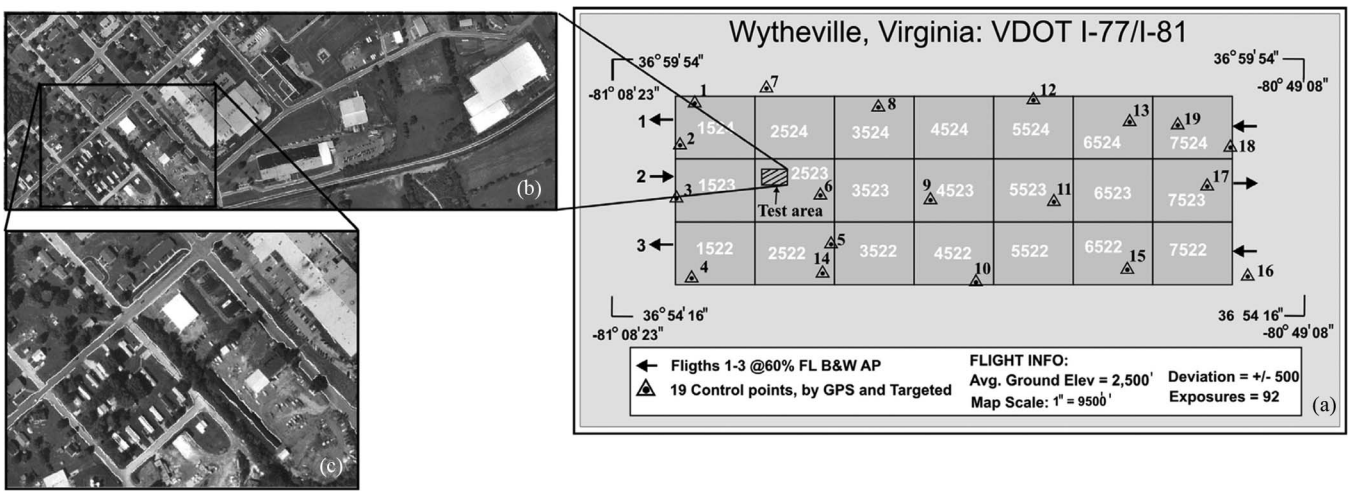

Fig. 12. Geodetic control test field in Wytheville, VA.

TABLE I
SPECIFICATION FOR LIDAR DATA COLLECTION

• **Name:** Optech 1210 LiDAR system	• **Aircraft Speed:** 202 ft/s
• **Date:** September 19, 2000	• **Flying Height:** 4500ft above ground level
• **Accuracy:** 2.0-feet	• **Scanner Field of View (half angle):** ±16 degrees
• **Density:** an average spacing of 7.3-feet	• **Scan Frequency:** 14Hz
	• **Swath Width:** 2581ft (1806 ft with a 30% sidelap)
	• **Pulse Repetition Rate:** 10KHz

TABLE II
SPECIFICATION FOR AERIAL IMAGE DATA COLLECTION

• **Camera:** Woolpert camera 5099	• **Scale:** 1:1000
• **Film:** Kodak 2405	• **Image type**: black-and-white aerial imagery
• **Focal length:** 153.087 mm	• **Pixel resolution:** 2.0 pixels
• **Total exposures:** 96	• **Date:** September 19, 2000

datum; horizontal data were referenced to NAD83/93 Virginia State Plane Coordinate system. The 19 GCP's accuracy averages 0.02, 0.02, and 0.01 m in X, Y, and Z, respectively. The density of spare LiDAR data is an average of 7.3 ft. The specification for the airborne LiDAR sensor is listed in Table I [51].

In the experimental field, aerial images using a Woolpert 5099 Camera at a focal length of 153.087 mm were acquired on September 19, 2000. The endlap of the images is about 65%, and the sidelap is approximately 30%. The specification for aerial image collection is listed in Table II [51].

In order to evaluate the proposed method, a patch is cut from the original aerial image [see Imagery ID: 2523 in Fig. 12(a)] for this experiment. The flowchart of the proposed aspect interpretation for house extraction is depicted in Fig. 13.

B. Preprocessing of Aerial Imagery

Step 1—Image Enhancement and Noise Removal: Although the aerial image was captured by high quality of Woolpert aerial camera 5099, it is unavoidable for the image to involve noise. Considering the extraction of house edges in next step, a traditional 3×3 filtering algorithm is used to accentuate edges. The typical high-pass filter mask is

$$Mask\ E = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \quad (2)$$

With the proposed mask, the convolution with the original image [see Fig. 14(a)] is operated. The enhanced image is depicted in Fig. 14(b).

Step 2—Edge Detection: Yan *et al.* [46] developed detector masks for edge extraction (see Fig. 15). This paper directly employs this method for house edge extraction. The mask orientations vary at an angle of 0°, 45°, 90°, and 135°. When the masks are convolved with a moving window, the strongest response is taken as the edge, and the corresponding orientation is taken as the direction of an edge of one building boundary.

The mathematical model of the output gray value through convolution computation can be expressed by

$$g(m,n) = \max\left\{\sum_{i=0}^{2}\sum_{j=0}^{2} f(m+i,n+j) \times w_k(i,j)\right\}, \quad k = 1,2,3,4 \quad (3)$$

where $f(m,n)$ is the gray value of the original image, and $w_k(i,j)$ is the mask whose values are depicted in Fig. 15. With the convolution computation, the gray values of building edges are enhanced, and the background noise is refrained efficiently. Thus, the candidate pixels of the building's edges can be extracted. After thinning of edges, the results of the extracted edges are shown in Fig. 16.

Step 3—Edge Pixels Grouping: When detecting the building edge pixels above, a building boundary is segmented into several edge fragments. For this reason, the following task is to trace the edge fragment pixels and then group them. The method to group two neighbor edge fragments into one new edge "fragment" is proposed as follows (see Fig. 17).

1) The slopes of two neighbor edge fragments meet the condition of $|\alpha - \beta| \leq slope\_threshold$, where α and β represent the slopes of two edge fragments, respectively.
2) The distance D between the ending point of an edge fragment and the starting point of another edge fragment should meet the condition of $D \leq dist\_threshold$.

When the two conditions above are met, two neighbor edge fragments are merged by

$$AD = AO_1 + O_1O_2 + O_2D. \quad (4)$$

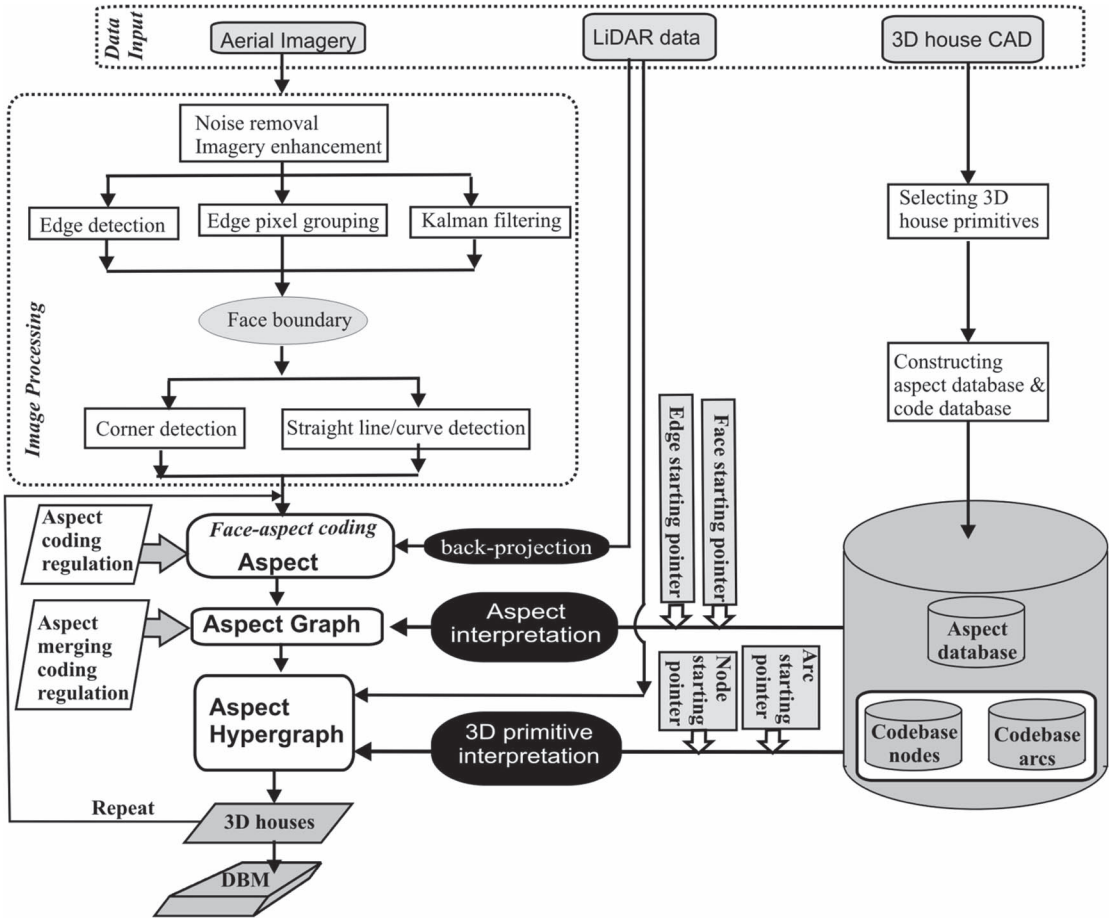

Fig. 13. Flowchart for the aspect interpretation-based house extraction.

Fig. 14. (a) Original image and (b) enhanced image using a 3 $times$ 3 filtering algorithm.

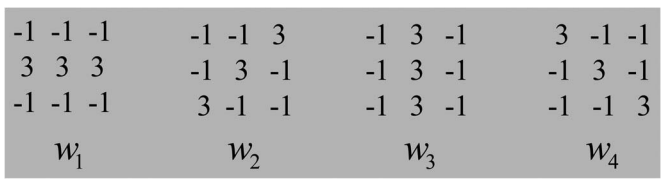

Fig. 15. Edge detection masks.

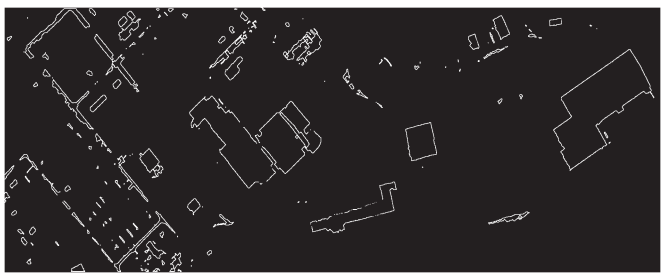

Fig. 16. Result of the extracted edges of the houses.

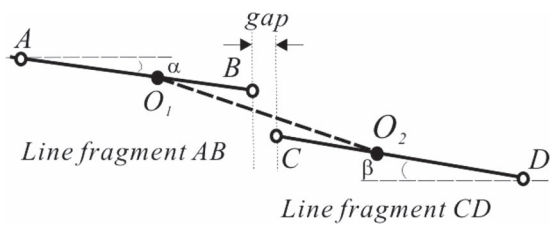

Fig. 17. Merging neighbor edge fragments.

The same operations are above repeated for each edge fragment until all edge fragments are finished. After this step, most of the edge fragments are smoothly connected, as shown in Fig. 18.

Step 3—Kalman Filter to Track Line: The above operations unavoidably cause the gaps along the edges of a house. Kalman filter algorithm is proposed to track the broken-off edges to fill the gaps. The basic idea of this method is that each edge

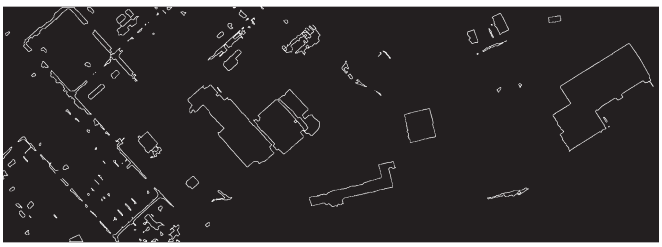

Fig. 18. Result of an extracted house boundary after implementing Step 3.

is considered as a track of a moving point on an edge. When the broken-off edge fragments are not detected, the Kalman filter keeps track of the next fragment. If there exist two edge fragments, which meet the two conditions in Step 2, they will be merged into one edge fragment. When a broken-off edge is detected, a number of the pixels are prefilled. The filled number is required to assure that the method can effectively track the next edge fragment. This computational process is mathematically described as follows: If an edge is tracked, let $X(k)$ be the state vector of movement, $x(k)$ be the moving location, $x'(k)$ be the velocity, $y(k)$ be the observation location that is the location of the detected edge point, T be the step size, and k be the tracing time; the state equation of constant velocity of point D is

$$X\left((k+1)/k\right) = AX(k) \tag{5}$$

where $X = [x \quad x']^T$, and $A = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$. The measurement model is simplified as

$$Y(k) = CX^T(k) \tag{6}$$

where $C = [1 \quad 0]^T$, and $Y(k) = [y \quad y']^T$. The predicated state equation is

$$\hat{X}\left((k+1)/k\right) = A\hat{X}(k). \tag{7}$$

The output of the filter is

$$\hat{X}(k+1) = \hat{X}\left((k+1)/k\right) + K(k+1)$$
$$\left(Y(k+1) - C\hat{X}\left((k+1)/k\right)\right) = \hat{X}\left((k+1)/k\right). \tag{8}$$

After the Kalman filter in combination with Step 2 is carried out, an entire building boundary is extracted. The result for the study area is depicted in Fig. 18.

C. Creation of Aspect Graphs and Extraction of Houses

With the detected edges above, the next task is to create aspect and aspect graphs using the proposed method in Section II and then create a digital building model (DBM) of the houses. To this end, the following steps are proposed.

Step 1—Face Coding: The first step of aspect creation is face coding to the boundary detected in Section III-B. The following steps are conducted.

- *Face filling:* Using a four-neighbor seed point growing algorithm, each face is filled with gray value, 255. An example filled by the shadow is shown in Fig. 19. With face filling, the area of each face is obtained by accounting the number of the filled pixels. After filled, a binary imagery is created.
- *Boundary detection:* With the binary imagery created above, a gray threshold of 100 is set up to detect the boundary. After this, thinning edges and deleting edge burrs are implemented consequently. Finally, the boundaries with a pixel width are obtained and displayed in Fig. 18.
- *Boundary vectorization:* With the result obtained above, the boundary of each face is vectorized using an eight-neighbor tracking operator. With the operation, the perimeter of each boundary can be obtained and is used as an attribute value of the face aspect.
- *Corner detection:* With the vectorized boundary of each face above, the corner detection is conducted using a method called *the maximum curvature difference* between *forward K steps* and *backward K steps*. The basic steps of this method are (see Fig. 19):

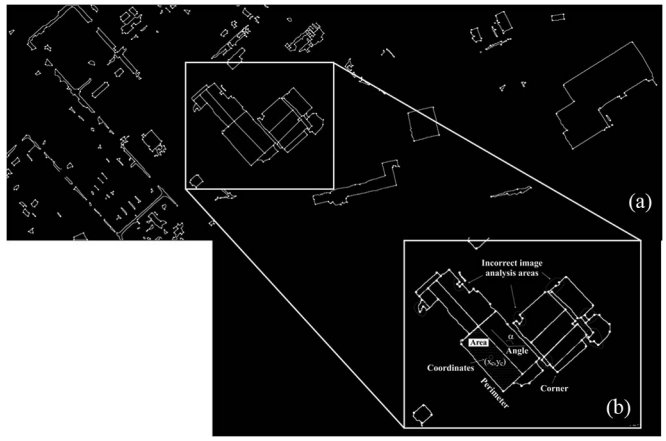

Fig. 19. Procedures of coding a house using the formulated coding regulation, in which the small circles indicate an endpoint or the corner of a line or a curve. The orientation and length of a straight-line segment are calculated. The area, perimeter, and the central coordinates of a face are calculated as well.

TABLE III
ATTRIBUTE VALUES OF NOTES AND ARCS

Attribute Values For Note	Attribute Values For Arc
• Area • Perimeter • Central coordinates • Orientation angle θ of a face-aspect • Codes describing face-aspect	• The coordinates of a shared edge • The types connected edges • The slope of straight line • The curvature of a curve • The information of two end-points

Selecting one pixel along the boundary, we called *central pixel*, and calculating the curvatures of the line located in the selected pixel (noted k_0), the forward 3 pixels' (noted k_{+3}), and the backward 3 pixels' (noted k_{-3}), respectively. The curvature differences between the selected and the forward pixels (noted k_{0+3}) and between the selected and the backward pixels (noted k_{0-3}) are calculated, respectively. The corner is detected by using the second curvature difference, i.e.,

$$Curvature_0 = |k_{0-3} - k_{0+3}|$$
$$\times \begin{cases} if\ Curvature > threshold & \text{noted as corner} \\ if\ Curvature \le threshold & \text{noted as no corner.} \end{cases} \tag{9}$$

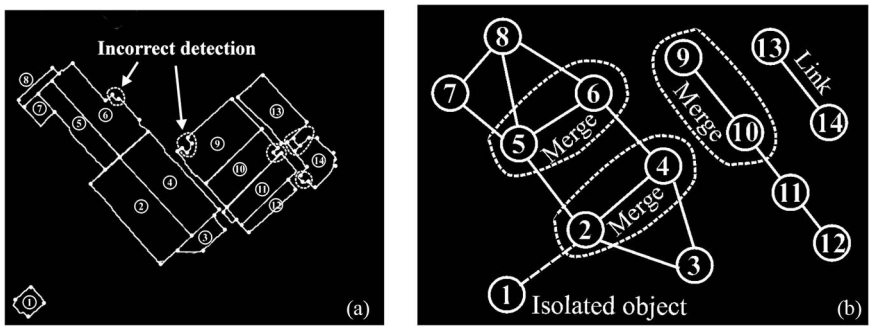

Fig. 20. Illustration of aspect and aspect graph construction, (a) the segmented houses, and (b) and their aspect graphs.

If $Curvature_0$ is greater than the given threshold, the pixel is determined as a corner; otherwise, it is not a corner pixel. Sometimes, it is necessary to suppress the local nonmaximum since the multiple $Curvature_0$'s surrounding the central pixel simultaneously meet the condition of (9). With the proposed method here, the result of the detected corners is displayed in Fig. 19.

- *Determination of straight line or curve:* When using the aspect or aspect graph for house interpretation, the property of a line segmentation, either a straight line or a curve, has to be determined. To this end, the curvature of the line segmentation between two corners is calculated. If the curvature is close to zero, the line segment is considered as a straight line, otherwise, as a curve.
- *Face coding*: With the operations above, face coding is conducted in terms of the coding regulation formulated in Sections II-C and II-D. When the above operations are finished, the face codes are completed. An example is depicted in Fig. 19.

Step 2—Creation of Aspect Graphs: With the faces coded above, the aspect graph is constructed. The basic steps include: First, each face is given an attribute name using natural numbers 1, 2, 3, ... Meanwhile, the attribute values are calculated using the parameters listed in Table III. Second, the aspect graph is constructed using nodes and arcs. Each face is taken as a node of the aspect graph, and each of the shared edge is determined by the same imagery coordinates and taken as an arc of the aspect graph [see Fig. 20(b)]. Third, a primitive aspect graph is constructed using the aspect merging regulation described in Section II-D. For example, face aspects 5 and 6, face aspects 9 and 10, and face aspects 2 and 4 are merged in Fig. 20(b). In addition, face aspect 1 is an isolated aspect, which represents an independent house. Fourth, the aspect graphs between the created in this section and the archived in the database are correlated. With this operation, the aspect graph using the proposed face-coding regulation can effectively remove the mis-segmented area [see Fig. 20(a) and (b)].

Step 3—Coregistration Between Aspect and LiDAR Data: The created aspects above unavoidably contain the objects that do not belong to houses, such as vehicles or parking lots. For this reason, the application of elevation information driven by LiDAR data will help separate ground objects and non-ground objects. A common method is to coregister the LiDAR data and the created aspects, with which the house's boundaries and 3-D coordinates can be obtained and are used to further create a DBM. This paper develops a method called *aspect-based coregistration*. The basic idea is: 1) The aspect is selected as coregistration elements since the aspect consists of the house boundary, which is important invariant geometric elements in house; 2) LiDAR data are back-projected onto aspect imagery plan to determine which LiDAR footprints are located within an aspect; 3) the algorithm for coregistration between the LiDAR data and the aspects is developed to create the DBM (see Fig. 21). The detailed steps are presented below.

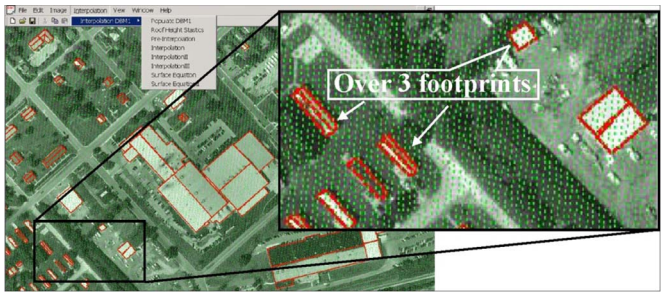

Fig. 21. Coregistration of aerial imagery and LiDAR footprints.

- **Determination of reference system:** The LiDAR data with X-, Y-, and Z-coordinates have been referenced to NAVD88 datum at height; and the NAD83/93 Virginia State Plane Coordinate system in planimetric coordinates. Thus, the datum of the LiDAR data set is used as the common reference framework.
- **Back Projection of LiDAR onto Aspect** Imagery **Plane:** The LiDAR data are back projected onto the created aspect plane, also called "*aspect imagery*" for removal of the ground objects and addition of the objects missed by operation of aspect creation in Step 2. The basic steps consist of:
  1) Back projection of LiDAR data onto the aspect plane (called aspect imagery) using the collinearity equations accompanying with the known interior and exterior orientation parameters of the aerial images, which are provided by a vendor.
  2) Correlation of the created aspects and LiDAR data and identification of the following four types of correlation cases (see Fig. 20):
     - Case 1: LiDAR data are correctly matching with the created house aspects.
     - Case 2: LiDAR data are partially matching with the created house aspects.

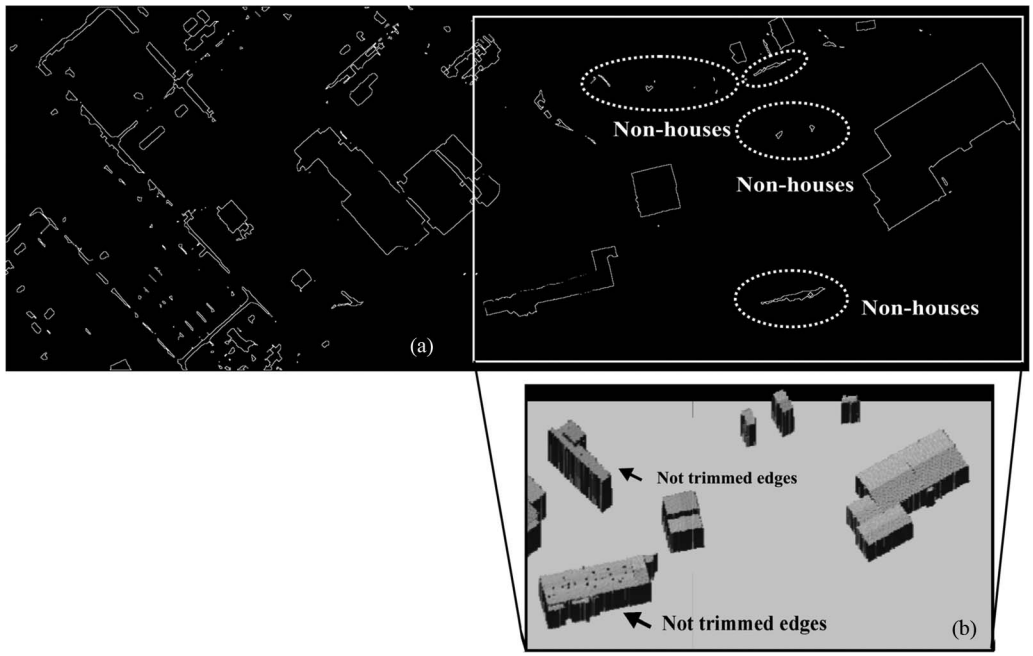

Fig. 22. Coarse 3-D models of the houses through coregistration between the aspect and the LiDAR data.

- Case 3: LiDAR data have no matching with the created house aspect.

For Case 2 and Case 3 above, there are three possibilities: 1) Edges of the houses in aerial imagery possibly have not been completely detected due to the traditional optical imagery problems, such as low gray contrast; 2) partial or complete occlusion due to the property of the perspective projection; and 3) the created aspect is part of an entire house. For Case 3 above, there are two possibilities: First, the size of the house is less than 14×14 ft^2 in length and width, i.e., less than two LiDAR footprints were shot on the house roof. Second, the created aspects are for ground objects such as parking lots, which have no elevation information.

This paper only considers the cases that over three LiDAR footprints are shot on the aspects, i.e., Case 1 and Case 2.

With the above coregistration between aspects and LiDAR data, the roof equation can be established using a local plane fitting equation, i.e.,

$$aX + bY + cZ = 1 \tag{10}$$

where a, b, and c are unknown coefficients, and X, Y, and Z are coordinates of LiDAR footprints shot within the house's roof. As shown in (10), at least three LiDAR footprints are required to determine the equation coefficients a, b, and c by least squares estimates. With this algorithm, a coarse 3-D model of houses can be generated. As observed in Fig. 22, the boundaries of the 3-D model of each house are very coarse since it is difficult to extract the boundaries of each house from single LiDAR data.

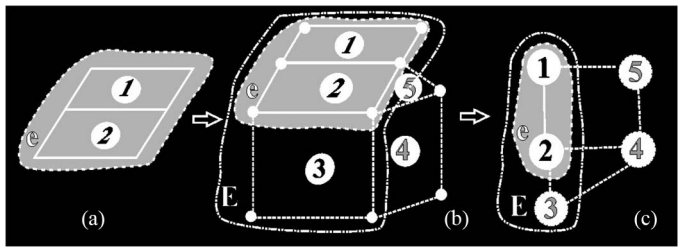

Fig. 23. Illustration of 3-D primitive graph creation procedure.

In addition, some objects are wrongly detected as aspect (i.e., non-ground objects) in the operation of the edge detection and the aspect creation using the method described above [see Fig. 22(a)]. However, they are in fact ground objects since they have no elevation when coregistered with the LiDAR data [see Fig. 22(b)].

Step 4—Construction of Aspect Hypergraphs: Step 3 unavoidably missed the detection of a few houses and/or wrongly detected the houses. For this reason, this paper uses the aspect hypergraph in combination with LiDAR data to overcome the wrong determination. The basic procedures are: For a given face aspect [see Fig. 23(a)], we check whether the LiDAR is shot within the face aspect. If only one LiDAR footprint is shot, it is supposed that the face aspect may represent a house. We further examine the connected face aspect manually. If the connected face aspect has no LiDAR shot, it is supposed that this face aspect is probably a wall [see Fig. 23(b)]. Consequently, this wall face aspect will be merged with the roof face aspect to construct a 3-D volume primitive graph [see Fig. 23(c)]. If two neighbor face aspects have no LiDAR shot, it is supposed that the face aspects represent other objects such as parking lots. Repeat the procedures above and manually check whether

all houses are detected one by one. If not, the house has to be manually added.

Step 5—House Extraction Using Aspect Interpretation: The primary design for the extraction of houses using aspect interpretation is a correlation operation between the aspect codes created from aerial imagery (called *image code*) and the aspect codes archived in the codebase (called *codebase*). The correlation operation is carried out by the attribute values assigned in both the nodes and the arcs of the aspect graphs. Each node and each arc are composed of many attribute values; a linear combination of these attributes is constructed by

$$node_{\text{attribute}} = Area + Perimeter + CentralCoord + Orientation + Code_{\text{aspect}} \quad (11)$$

$$arc_{\text{attribute}} = Coord_{\text{Two-ends}} + Type_{\text{edge}} + Slope_{\text{line}} + Curvature_{\text{curve}}. \quad (12)$$

The correlation coefficient maximum of the note/arc attributes between the obtained from aerial imageries and the archived in the database is taken as criterion. With the new combined variables in (11) and (12), the two criteria are employed to determine their correlation, respectively. The first criterion is the correlation coefficient maximum of the node, which is employed for determining the candidate house under a given threshold of correlation coefficient. The mathematical model is expressed by

$$r_i = \begin{cases} r_i & \text{if } r_i \le R_\theta \\ 100 & \text{if } r_i > R_\theta \end{cases} \quad r_i \subseteq R \quad (13)$$

where R_θ is the threshold, and R is the set of r_i. R_θ is set 0.95 in this paper. The successful operation of this step will be able to interpret and find a significant amount of geometric information of a house, such as aspect property, geometry, gray area, relationship, and separate the ground objects (e.g., parking lots) and non-ground objects (e.g., houses). The second criterion is a correlation coefficient maximum of the arcs between the obtained from the aerial imagery and the archived in the database. The same operation as the node correlation is carried out. The successful operation of this step will be able to interpret and find a primitive graph, which describes an entire projection or partial projection of a 3-D primitive (house). If an aspect is identified as a house, we will further determine how many LiDAR footprints hit its roof and then calculate the house's model, which will be described in Step 7.

Step 6—Repeat the Operations Above: Repetition of the operations in Step 4 and Step 5 is required in order to avoid missing the extraction of houses.

Step 7—Creation of Accurate 3-D Model of Houses: An aspect with its correct interpretation is capable of accurately depicting the boundary of a house. With the accurate boundary of house, a precise 3-D model can be created using the method proposed in [51]. The major difference between this step and Step 3 is that boundary information of the aspect interpreted is used to fit the planar equation. Thus, the boundary of the 3-D model appears to be trimmed, i.e., an accurate DBM describing houses can be established (see Fig. 24). As compared with Fig. 22, it is found that the aspect-based 3-D model is much more accurate than other methods.

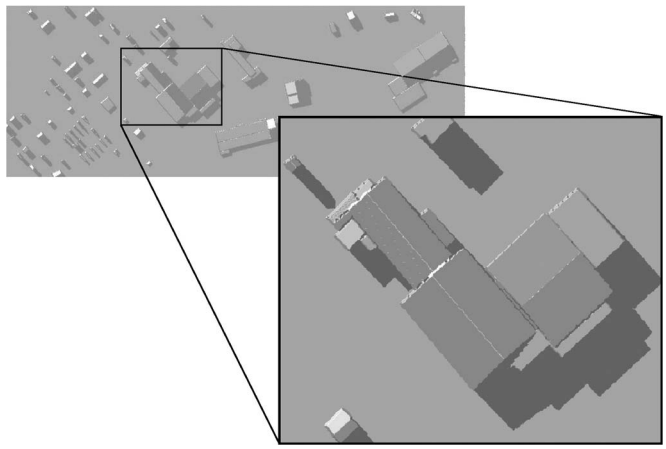

Fig. 24. Houses extracted by the proposed method.

D. Occlusion Analysis

Height trees unavoidably occlude the houses, resulting in missing the house detection. In order to evaluate the impact of the proposed method to the tree occlusion, this paper extends the experiment for another aerial image (see ID 3523 in Fig. 12). The experimental results are depicted in Fig. 25. As checking the extracted houses [see Fig. 25(b)] and the corresponding original image [see Fig. 25(a)] manually, it is found that a few houses cannot be detected when using the proposed method. Majority of the undetected houses are from either tree occlusion or the house roof's gray values close to the trees [see Fig. 25(b)]. The rate of successfully extracted houses is 93%, i.e., 113 out of 122 houses are detected using the proposed method. The major reasons for missing the detection of house using our method are probably caused by (a) when a high tree occludes most part of a house, the house cannot be detected since the house aspect cannot be formed after the aerial image is processing using the method proposed in this paper; (b) if the less than three LiDAR footprints shot a house, the house cannot be detected since the height information cannot be obtained. We manually checked the detected and undetected houses and found that eight undetected houses are due to the tree occlusions or the house roof's brightness close to trees, and only one house is due to the less than three LiDAR footprints onto the house. This fact demonstrated that the tree occlusion to houses and house roof's brightness close to the surrendering objects significantly impact the house detection when using the proposed method.

E. Comparison Analysis

The comparison between our method and the hierarchical morphological filtering (*called Method* 1 hereafter), which was proposed by Rottensteiner [36] is conducted. The basic idea of Method 1 is to first remove ground footprints using a ground-filtering algorithm and then further remove the remaining non-building pixels using the size, shape, height, and height

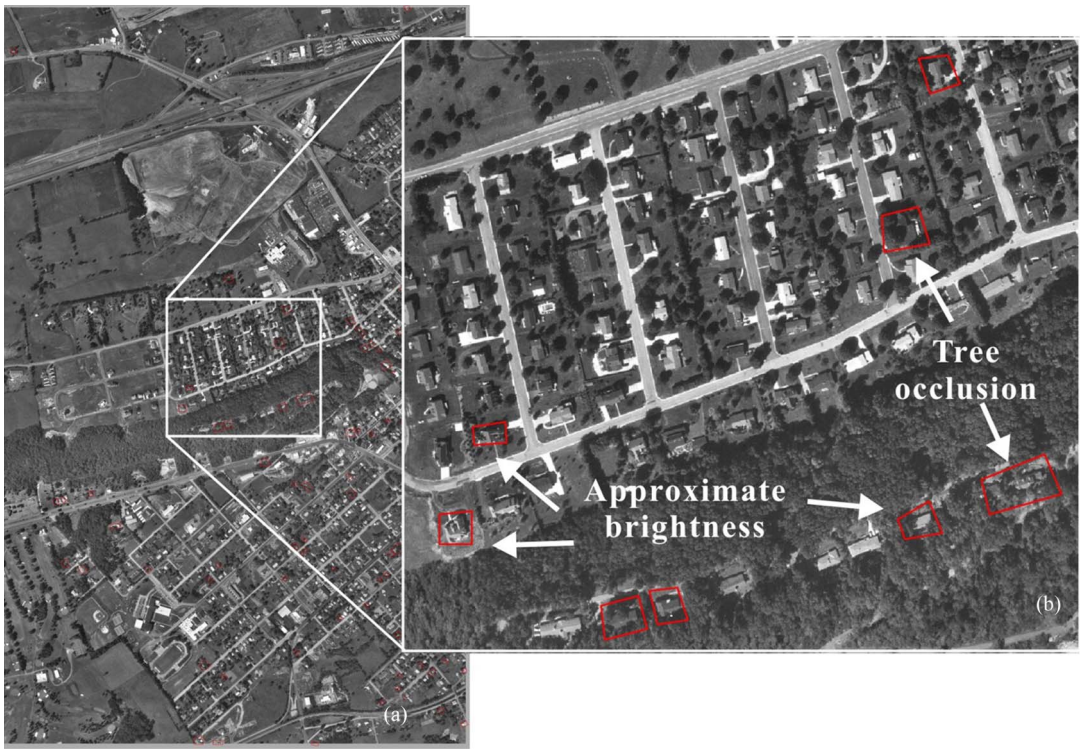

Fig. 25. Tree occlusion impact analysis to the proposed method and the undetected houses.

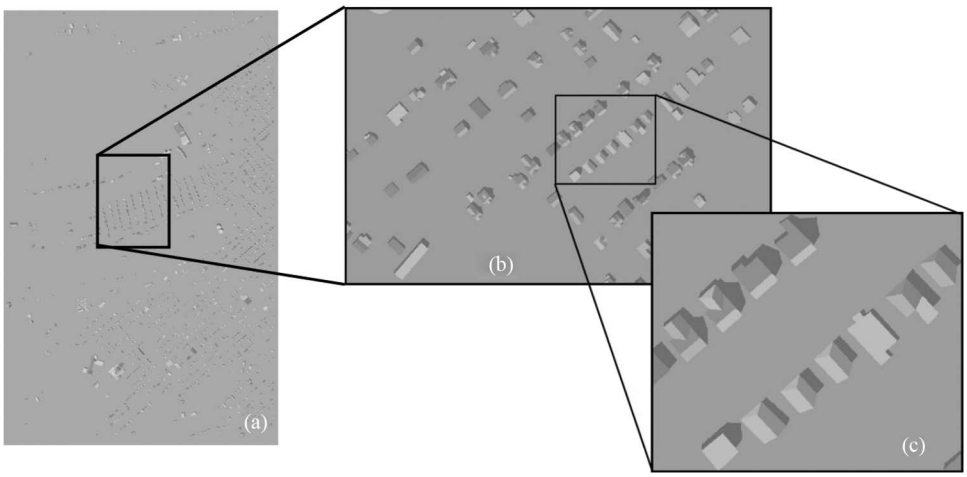

Fig. 26. Houses extracted in the aerial image (ID 3523) by the proposed method.

difference between the first and last returns. This section gives a brief description below for Method 1, and its details can be referenced to [36].

- **Ground Filtering:** The primary purpose of the method is to apply the multi-directional ground filtering algorithm to identify LiDAR ground footprints and generates a Digital Elevation Model (DEM). A 2-D mesh is first overlaid on the LiDAR point clouds, and the elevation of the nearest point is assigned to the cell. The ground filter algorithm is used to separate ground and non-ground footprints using the given thresholds of slopes, the elevation difference between a point and its local minimum elevation, and the elevation difference between a point and the nearest ground point. With successful implementation of this operation, DEM, only composed of non-ground objects, is created at the same resolution [see Fig. 27(a) and (b)].
- **Removing Non-building Pixels:** With the created coarse DEM above, the coarse DEM are mainly composed of vegetation, buildings, and other objects, such as vehicles. Accordingly, three pixel-based operations are used to break non-building blocks into smaller fragments on the basis of the fact that houses differ from other objects in height, size, and shape, and laser lights possess a relatively strong ability to penetrate vegetation and hardly pass solid house roofs. This operation can remove objects that are shorter and lower than the minimum house in length and in height from the house candidates within coarse DEM, such as shrubs and vehicles. Meanwhile, vegetation is also filtered out from coarse DEM. Finally, an accurate DEM

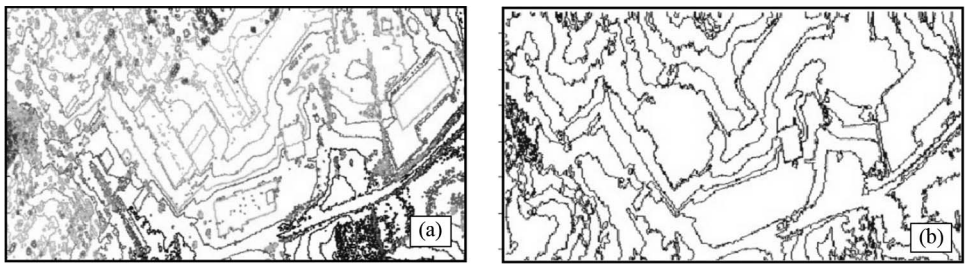

Fig. 27. (a) Original DSM and (b) DEM created by the ground filtering algorithm (courtesy of Ren *et al.* [34]).

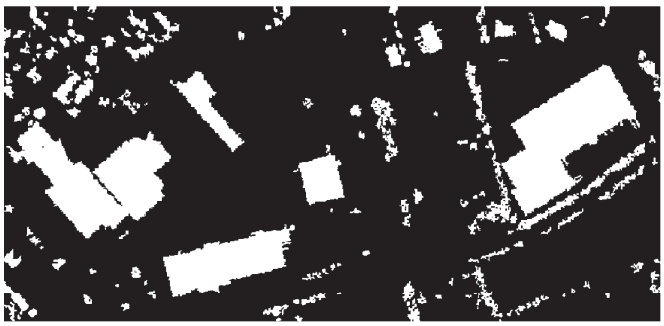

Fig. 28. Non-ground objects extracted using single LiDAR data.

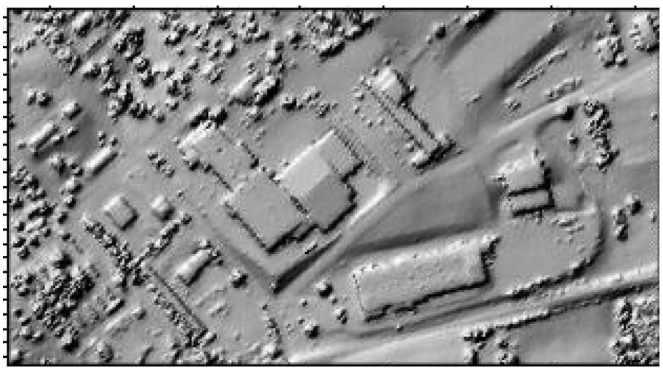

Fig. 29. Visualization of the houses extracted by using single LiDAR data.

only composed of houses is created, and the boundary of each house is vectorized into vectors (Fig. 28).

In order to compare the accuracy and advantages of our method compared with *Method 1*, we visualize the houses extracted by *Method 1*, as depicted in Fig. 29. As observed and compared by Figs. 24–29, it is found that [34].

1) The rate of extracting houses is 82% (i.e., 100 out of 122 houses) using Method 1. The major reasons for the undetected house using Method 1are probably caused by the fact that Method 1 only used height information, i.e., LiDAR data did not use the aerial image information. When the trees surround the house, Method 1 will be blind to either trees or a house.
2) Method 1 only creates coarse boundary (see Fig. 29), whereas our method is able to create a higher accuracy value of DBM than Method 1 does (see Figs. 24 and 26) since our method proposed aspect interpretation operation, which accurately fused the LiDAR data and the imaged houses.

IV. Conclusion

The main contribution of this paper lies in the development of a seamless fusion between LiDAR data and aerial imagery. This fusion differs from the traditional methods, which simply merged aerial imagery with LiDAR data. The proposed method is: 1) to create the aspects and the aspect graphs from an aerial imagery; and then 2) merge the created aspects with LiDAR data. The major characteristics of the proposed method lie in: 1) The aspect and the aspect graph contain significant amount of knowledge of a building, such as geometric features, shapes, and structures; therefore, the proposed method is a type of geometric information-driven method for the extraction of houses; and 2) data fusion is carried out at each process of the house detection and the house description; hence, it is thus called a seamless fusion in this paper.

The main work of the proposed method includes the following. 1) Aerial image preprocessing, including edge detection using a detector mask, edge pixel grouping, and gap filling, is developed. 2) Eleven types of 3-D house primitives are selected, and their projections are represented by the aspects. The coding regulations are formulated for creation of the (hierarchical) aspect graphs on the basis of the results of aerial image processing and the results of LiDAR data processing. In the aspect graphs, the notes are represented by the face aspect, and the arcs are described by the attributes using the formulated coding regulations. 3) The coregistration between the aspects and the LiDAR data is implemented using the correlation operation. As a consequence, the aspects and/or the aspect graphs are interpreted for the extraction of houses, and the extracted houses are fitted using a planar equation for the creation of a DBM.

The experimental field, which was established by Virginia Department of Transportation, contracting to Woolpert LLC at Wytheville, VA, is used to evaluate the proposed method. The experimental results have demonstrated that the proposed method is capable of effectively extracting houses at a rate of 93%, compared with another method, which is 82% effective with LiDAR spacing of approximately 7.3 by 7.3 ft^2 in the experiment area. The accuracy of the 3-D DBM is higher than the method using only single LiDAR data.

Acknowledgment

The authors would like to thank Virginia Department of Transportation and Woolpert LLC at Richmond, VA, for

providing the experimental data. They would like to especially thank F. Sokoloski and Q. Xiao for their discussion of the technology of LiDAR data processing, as well as for their kind assistance in LiDAR data check and delivery.

REFERENCES

[1] A. Ahmadi, A. Salman, M. J. Valadan Zoej, H. Ebadi, H. Abrishami Moghaddam, and A. Mohammadzadeh, "Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 12, no. 3, pp. 150–157, Jun. 2010.

[2] P. Axelsson, "Processing of laser scanner data—Algorithms and applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 2/3, pp. 138–147, Jul. 1999.

[3] C. Baillard and H. Maître, "3-D reconstruction of urban scenes from aerial stereo imagery: A focusing strategy," *Comput. Vis. Image Understanding*, vol. 76, no. 3, pp. 244–258, Dec. 1999.

[4] G. Le Besnerais, M. Sanfourche, and F. Champagnat, "Dense height map estimation from oblique aerial image sequences," *Comput. Vis. Image Understanding*, vol. 109, no. 2, pp. 204–225, Feb. 2008.

[5] I. Biederman, "Human image understanding: Recent research and a theory," *Comput. Vis., Graph Image*, vol. 32, no. 1, pp. 29–73, Oct. 1985.

[6] M. Cord, M. Jordan, and J.-P. Cocquerez, "Accurate building structure recovery from high resolution aerial imagery," *Comput. Vis. Image Understanding*, vol. 82, no. 2, pp. 138–173, May 2001.

[7] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld, "3-D shape recovery using distributed aspect matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 174–197, Feb. 1992.

[8] H. L. Dong, M. L. Kyoung, and U. L. Sang, "Fusion of lidar and imagery for reliable building extraction," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 215–225, Feb. 2008.

[9] E. Nima, V. Zoej, M. Javad, M. Z. Sahebi, and A. Mohammadzadeh, "Automatic building extraction from LiDAR digital elevation models and WorldView imagery," *J. Appl. Remote Sens.*, vol. 3, no. 1, pp. 033571-1–033571-12, Jan. 2009.

[10] A. Fischer, T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, and V. Steinhage, "Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D," *Comput. Vis. Image Understanding*, vol. 72, no. 2, pp. 185–203, Nov. 1998.

[11] M. Fradkin, H. Maître, and M. Roux, "Building detection from multiple aerial images in dense urban areas," *Comput. Vis. Image Understanding*, vol. 82, no. 3, pp. 181–207, Jun. 2001.

[12] K. Fujii and T. Arikawa, "Urban object reconstruction using airborne laser elevation image and aerial image," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2234–2240, Oct. 2002.

[13] P. Gamba and B. Houshmand, "Joint analysis of SAR, LiDAR and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings," *Int. J. Remote Sens.*, vol. 23, no. 20, pp. 4439–4450, Jan. 2002.

[14] A. Habib, M. Ghanma, M. Morgan, and R. Al-Ruzouq, "Photogrammetric and lidar data registration using linear features," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 6, pp. 699–707, Jun. 2005.

[15] T. Hermosilla, L. A. Ruiz, J. A. Recio, and J. Estornell, "Evaluation of automatic building detection approaches combining high resolution images and LiDAR data," *Remote Sens.*, vol. 3, no. 6, pp. 1188–1210, Jun. 2011.

[16] H. Jinhui, S. You, and U. Neumann, "Integrating LiDAR, aerial image and ground images for complete urban building modeling," in *Proc. 3rd Int. Symp. 3DPVT*, Jun. 2006, pp. 184–191.

[17] C. Jaynes, E. Riseman, and A. Hanson, "Recognition and reconstruction of buildings from multiple aerial images," *Comput. Vis. Image Understanding*, vol. 90, no. 1, pp. 68–98, Apr. 2003.

[18] F. Jung, "Detecting building changes from multitemporal aerial stereopairs," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3/4, pp. 187–201, Jan. 2004.

[19] M. Kabolizade, H. Ebadi, and S. Ahmadi, "An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data," *Comput., Environ. Urban Syst.*, vol. 34, no. 5, pp. 435–441, Aug. 2010.

[20] K. Khoshelham, C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti, "Performance evaluation of automated approaches to building detection in multi-source aerial data," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 1, pp. 123–133, Jan. 2010.

[21] T. Kim and J.-P. Muller, "Automated urban area building extraction from high resolution stereo imagery," *Image Vis. Comput.*, vol. 14, no. 2, pp. 115–130, Mar. 1996.

[22] Y. C. Lee and K. S. Lei, "Machine understanding: Extraction and unification of manufacturing features," *IEEE Comput. Graphics Appl.*, vol. 7, no. 1, pp. 20–32, Jan. 1987.

[23] A. Mastin, J. Kepner, and J. Fisher, III, "Automatic registration of LiDAR and optical images of urban scenes," in *Proc. IEEE CVPR Workshops*, 2009, pp. 2639–2646.

[24] T. Matsuyama, "Knowledge-based aerial image understanding system and expert systems for image processing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 3, pp. 305–316, Mar. 1987.

[25] H. Mayer, "Automatic object extraction from aerial imagery—A survey focusing on buildings," *Comput. Vis. Image Understanding*, vol. 74, no. 2, pp. 138–149, May 1999.

[26] D. M. McKeown, "Automating knowledge acquisition for aerial image interpretation," *Comput. Vis., Graph, Image*, vol. 46, no. 1, pp. 37–81, Apr. 1989.

[27] D. M. McKeown, "Knowledge-based aerial photo interpretation," *Photogrammetria*, vol. 39, no. 3, pp. 91–123, Sep. 1984.

[28] R. Mohan and R. Nevatia, "Using perceptual organization to extract 3-D structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 11, pp. 1121–1139, Nov. 1989.

[29] M. Morgan and K. Tempfli, "Automatic building extraction from airborne laser scanning data," *Int. Arch. Photogramm. Remote Sens.*, vol. 33, no. B3, pp. 616–623, 2000.

[30] D. O'Donohue, S. Mills, S. Bartie, P. Park, and A. David, "Combined thermal–LiDAR imagery for urban mapping," in *Proc. 23rd IVCNZ*, 2008, pp. 1–6.

[31] N. Paparoditis, M. Cord, M. Jordan, and J.-P. Cocquerez, "Building detection and reconstruction from mid- and high-resolution aerial imagery," *Comput. Vis. Image Understanding*, vol. 72, no. 2, pp. 122–142, Nov. 1998.

[32] J. Peng, D. Zhang, and Y. Liu, "An improved snake model for building detection from urban aerial images," *Pattern Recog. Lett.*, vol. 26, no. 5, pp. 587–595, Apr. 2005.

[33] S. Pu and G. Vosselman, "Knowledge based reconstruction of building models from terrestrial laser scanning data," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 575–584, Nov. 2009.

[34] Z. Ren, M. Cen, and G. Zhou, "Filtering method of LiDAR data based on contours," *J. Remote Sens.*, vol. 13, no. 1, pp. 55–62, 2009, (Chinese).

[35] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 2, pp. 135–149, Jun. 2007.

[36] F. Rottensteiner, "Automatic generation of high-quality building models from LiDAR data," *IEEE Comput. Graphics Appl.*, vol. 23, no. 6, pp. 42–50, Nov./Dec. 2003.

[37] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 11–20, Mar. 2009.

[38] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne lidar point clouds," *Photogramme. Eng. Remote Sens.*, vol. 73, no. 7, pp. 805–812, Jul. 2007.

[39] T. Schenk and B. Csatho, "Fusion of lidar data and aerial imagery for a more complete surface description," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 34, no. 3A, pp. 301–317, 2002.

[40] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, Apr. 2009.

[41] E. Simonetto, H. Oriot, and R. Garello, "Rectangular building extraction from stereoscopic airborne radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2386–2395, Oct. 2005.

[42] G. Sithole and G. Vosselman, "Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 59, no. 1/2, pp. 85–101, Aug. 2004.

[43] G. Sohn and I. Dowman, "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 1, pp. 43–63, May 2007.

[44] I. Suveg and G. Vosselman, "Reconstruction of 3D building models from aerial images and maps," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3/4, pp. 202–224, Jan. 2004.

[45] L. Wang and U. Neumann, "A robust approach for automatic registration of aerial images with untextured aerial LiDAR data," in *Proc. IEEE CVPR Workshops*, 2009, pp. 2623–2630.

[46] G. Yan, G. Zhou, and C. Li, "Automatic extraction of power lines from large-scale aerial images," *IEEE Trans. Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 387–391, Jul. 2007.
[47] Y. Yu, B. P. Buckles, and X Liu, "Residential building reconstruction based on the data fusion of sparse LiDAR data and satellite imagery," in *Proc. 5th ISVC, II*, vol. 5876 LNCS, *Lecture Notes in Computer Science*, 2009, pp. 240–251.
[48] S. Zabuawala, H. Nguyen, H. Wei, and J. Yadegar, "Fusion of LiDAR and aerial imagery for accurate building footprint extraction," in *Proc. SPIE, IS T Electron. Imaging, Image Process., Mach. Vis. Appl.*, Feb. 2009, vol. 7251, pp. 72510Z-1–72510Z-11.
[49] L. Zebedin, A. Klaus, B. Gruber-Geymayer, and K. Karner, "Towards 3D map generation from digital aerial images," *ISPRS J. Photogramm. Remote Sens.*, vol. 60, no. 6, pp. 413–427, Sep. 2006.
[50] K. Zhang, J. Yan, and S. Chen, "Automatic construction of building footprints from airborne LiDAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2523–2532, Sep. 2006.
[51] G. Zhou, C. Song, J. Simmers, and P. Cheng, "Urban 3D GIS from LiDAR and digital aerial images," *Comput. Geosci.*, vol. 30, no. 4, pp. 345–353, May 2004.
[52] G. Zhou, "Geo-referencing of video flow from small low-cost civilian UAV," *IEEE Trans. Autom. Eng. Sci.*, vol. 7, no. 1, pp. 156–166, Jan. 2010.

Guoqing Zhou (M'02–SM'05) received the Ph.D. degree from Wuhan University, Wuhan, China, in 1994.

He was a Visiting Scholar with the Department of Computer Science and Technology, Tsinghua University, Beijing, China, and a Postdoctoral Researcher with the Institute of Information Science, Beijing Jiaotong University, Beijing. He continued his research as an Alexander von Humboldt Fellow with the Technical University of Berlin, Berlin, Germany, and afterward became a Postdoctoral Researcher with The Ohio State University, Columbus, OH, USA, from 1998 to 2000. He later had been with Old Dominion University, Norfolk, VA, USA, as an Assistant Professor, Associate Professor, and Full Professor in 2000, 2005, and 2010, respectively. He is currently with Guilin University of Technology, Guilin, China, as a "1000 Plan Talent Scholarship Professor" honored by the Chinese Government. He has published 2 books and more than 200 referred publications. He has worked on 48 projects as a Principal Investigator or a Co-Principal Investigator and participant as a Postdoctoral Researcher.

Xiang Zhou, photograph and biography not available at the time of publication.