INTEGRATED-APPROACH-BASED AUTOMATIC BUILDING EXTRACTION

Bofang ZHAO, John C. TRINDER
School of Geomatic Engineering
The University of New South Wales
Sydney NSW2052
AUSTRALIA
bofang@gwydir.gmat.unsw.edu.au
j.trinder@unsw.edu.au

KEY WORDS: Feature Extraction, Building Detection, Building Extraction, 3D Modelling.

ABSTRACT

This paper presents an approach to automatic building extraction from large-scale aerial images. It is based on the integration of multiple information, knowledge and appropriate methods, and hierarchically consists of three major parts: building detection, building segment extraction, 3D feature matching and building modelling. Buildings are detected by segmenting DSMs using grey-scale mathematical morphological operations, and by incorporating texture and shadow information for refinement and verification. The reconstruction process is focused on the detected building regions, and is accomplished by a bottom-up process: from primitive feature derivation, building segment extraction, segment matching, plane face constructing to 3D modelling. The approach uses a generic polyhedral model for extraction, so it can deal with buildings with different shapes and structures. Test results of the approach are presented in the paper.

1. INTRODUCTION

Efficient acquisition of up-to-date 3D information of man-made objects has been increasingly in demand for various applications, as demonstrated in the literature (Grün, 1998; Weider and Förstner, 1995). Automatic building extraction from aerial imagery has thus been an important issue in digital photogrammetry, for both practical and theoretical significance. Over the last decade, through continual efforts from the photogrammetry and computer communities, a number of approaches have been proposed to cope with different situations, such as 2D feature grouping and tracking (Mohan and Nevatia, 1989; Lin et al. 1995; Kim, 1995), digital surface model (DSM) based segmentation and detection (Baltsavias et al. 1995; Weider, 1996; Berthod et al. 1995), 3D feature matching and grouping (Roux and McKeown, 1994; Henricsson, 1996; Baillard et al. 1999), and data fusion methods (Brunn et al. 1998; Haala and Hahn, 1995). In spite of the progress made, the techniques at present are still not suitable for production applications.

The causes of the difficulties in achieving automatic building extraction from aerial imagery are many. One major cause is the complexity of the images. In an aerial image, there are usually a large number of objects of different classes coexisting irregularly with buildings in the scene, such as roads, vehicles and trees. Buildings may be partially occluded by objects nearby, including neighbouring building(s). The spectral characteristics and geometric aspects of other objects may overlap with and/or be similar to buildings. Therefore, the spatial and spectral interference of these objects results in unreliable low-level segmentation and feature extraction. Also the presence of a large number of objects can lead to a huge search space for higher level processing. The great variability of buildings, in terms of either radiometric properties or geometric shapes, is another major cause of the difficulties in building extraction. As a class of cultural objects, buildings appear in various sizes, shapes, heights, structures and functions, varying from a small low simple rectangular block with a flat roof to large high complex structures with complicated multi-based roof. Buildings are also built of different materials. The radiometric and texture properties of building roofs differ from one to another and even vary with time, season and sun illumination conditions. The appearance of a building in an image is also related to the scale and orientation of the image. In addition, the perspective projection of imaging also brings in some other difficulties, such as lack of explicit 3D information of objects, shadow interference and geometric displacements. In such complex circumstances, one specific kind of cue in the vision task is obviously insufficient. It is necessary to incorporate different kinds of information and knowledge, either derived from the original aerial images or from other sources, into various stages of the process in order to reliably extract buildings.

This paper presents an automatic approach for building extraction from large-scale aerial imagery, which is based on the integration of multiple information, knowledge and appropriate methods for extracting buildings hierarchically, which is aimed at reliability and efficiency. Its brief description is presented in Section 2, and the building detection process is described in Section 3, followed by building extraction in Section 4. Results of tests are given in Section 5 and a conclusion is drawn in Section 6.

2. OVERVIEW

In this research, a building is modelled as a polyhedron, comprising planes that are connected to form a solid volume. The intersections of adjacent planes are straight lines. Furthermore, the polyhedron has a set of attributes and knowledge describing its geometry, radiometry, texture, topology, and context. From this model and for attacking the complexity of the problem, the designed system consists of three parts, namely building detection, building segment extraction and 3D segment matching and building modelling. Within this hierarchical framework, the processes of building detection and reconstruction are purposely separated to achieve greater reliability and efficiency. For the same reason 2D and 3D procedures are applied mutually cooperatively. Hence, buildings are hierarchically reconstructed from global detection to local modelling, that is, from coarse to fine details. In addition, all components of the system are modular, thus allowing flexibility and extendability of the system and components. Figure 1 shows a block diagram of the system.

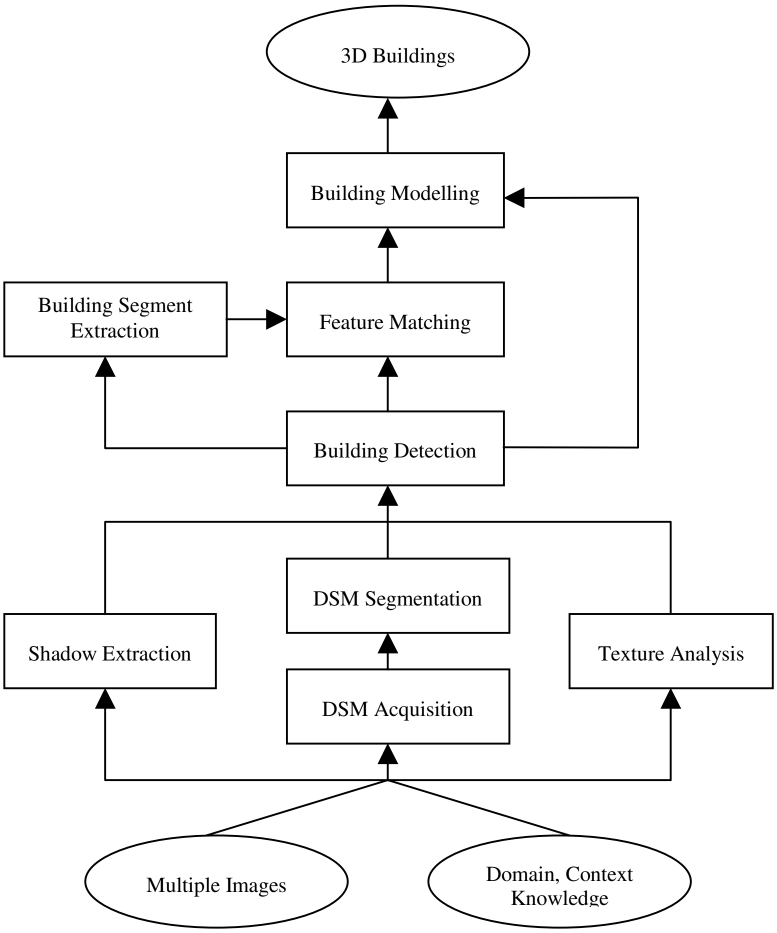


Figure 1. Block Diagram of the System

The detection process starts with segmentation of DSM (digital surface model) to derive regions of interest (ROIs) which have high expectancy of representing individual buildings. Texture and shadow information are extracted and used to refine and verify ROIs. All subsequent reconstruction procedures are focused locally on those detected building regions, thus greatly reducing the complexity of reconstruction process. Buildings are constructed in a bottom-up approach. Primitive linear features are first derived, and relevant building segments are extracted by grouping and filtering these primitive features within individual building regions. 3D lines are then generated by feature matching of these segments. Based on the matched lines, buildings are reconstructed by piecewise plane formation and plane intersection. This paper will describe the first two components of the system in Section 3 and 4 respectively.

3. BUILDING DETECTION

3.1 DSM Segmentation

Building detection is primarily based on DSM segmentation, which is motivated by the common knowledge that buildings are higher than the surrounding topographic surfaces. DSMs can be either derived by stereo matching with pairs of images, or obtained using airborne laser-scanner, as in (Hug, 1994; Hug and Wehr, 1997). Since the DSM is used only for building detection, and texture and shadow information is incorporated into the detection process, high accuracy of DSMs is not required in this step. As demonstrated in Section 5, DSMs generated from using general commercial digital photogrammetric systems, such as Socet Set and VirtuoZo, are adequate.

A DSM not only models all 3D objects in the scene, but also models the terrain surface, on which the objects are located. Hence, the first step is to separate "high objects" from the terrain surface. Several techniques have been presented for this step, including edge detection in DSMs, grey-scale mathematical morphology method (Weidner and Förstner, 1995) and "multiple height bins (MHB)" method (Baltsavias et al., 1995). In our approach, grey-scale mathematical morphology is applied. Specifically, it involves four steps: (i) "erosion" operation followed by "dilation" operation on DSM with a structural element *S* to derive an approximate DEM; (ii) subtraction of DEM from DSM to yield "height peaks"; (iii) thresholding the "height peaks" to extract regions of interest (ROIs); (iv) filtering ROIs. During the above process, domain knowledge of minimum size, maximum size and minimum heights of buildings in the area, together with the context knowledge of the image scale are employed.

3.2 ROI Refinement

It is clear that besides buildings, the initial ROIs derived from DSM may also include other "high objects" like trees. If DSM is accurate enough, information of surface normals of the DSM can be used to distinguish buildings from other objects (Brunn and Weidner, 1998). In our approach, texture information has been utilised for this purpose. While existing texture measures are usually unable to distinguish all classes of objects present in aerial images with desirable accuracy, texture characteristics within the constrained areas, i.e. the detected ROIs, are discernible. Texture filters proposed by Laws (1980) are used for texture calculation on ROIs in our approach. Classification of the texture images can be accomplished by Bayesian methods. For practical purposes it seems necessary to establish a texture database in the system and to develop different approaches for the various types of buildings. Since the proposed system is modular, modifying this component will not affect other modules of the system. After classification, shape attributes are calculated for each refined ROI as auxiliary information for subsequent processing. Domain knowledge of minimum size of buildings is again applied to reject small regions. Since the building region on texture image is usually smaller than its true size, due to the effect of widow size of the texture filters, the refined ROIs are expanded up to half of the window size through morphological operation "dilation".

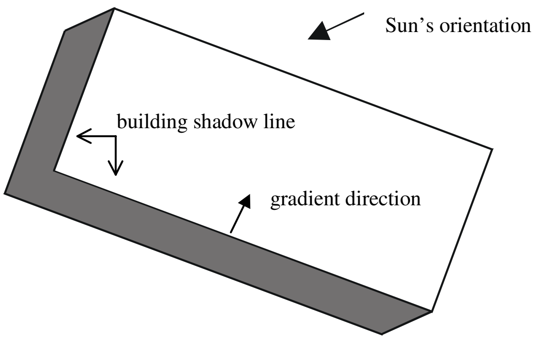


Figure 2. Building Shadow Lines

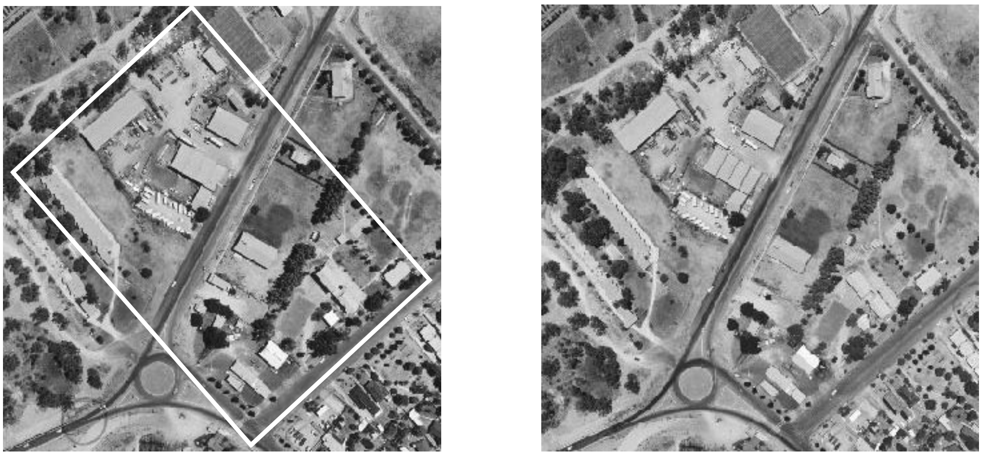
3.3 ROI Verification

To improve reliability in the reconstruction process, the hypothesised ROIs require further verification. In our system this evidence is provided by shadows cast by the buildings. In addition, building shadow lines will also be used in reconstruction process. Extracting shadows in aerial images may be done simply by thresholding in simple situations (Huertas and Nevatia, 1988; Irvin and McKeown, 1989). In order to handle complex aerial images, the method proposed in (Liow, 1990) has been extended by incorporating shadow intensity information. As illustrated in Figure 2, under non-vertical sun illumination a building casts a dark shadow strip. Since in many cases the shadow line opposite the building does not always appear, due to the interference of other objects nearby, only the building shadow line is of interest in this study. The building shadow line is extracted using three criteria: average gradient magnitude along the shadow line, relative orientation between sunlight and gradient direction, and average intensity of the shadow line's neighbouring region in the sunlight direction. The Sun's orientation and elevation are automatically derived from longitude, latitude and imaging time.

4. BUILDING EXTRACTION

The extraction process is carried out in a data-driven, bottom-up approach: from the derivation of primitive features to the formation of building structures and building models. The complete extraction and reconstruction process involves these steps: primitive feature derivation, building segment extraction, segment matching and plane construction, and 3D building modelling, of which the first two steps are to be discussed in this section. As the detected building regions, i.e. ROIs, provide focus of attention for the extraction process, all these steps, except the first one, are executed locally on each individual detected building regions.

Corresponding to the employed generic polyhedral model of buildings, Burns's "straight line algorithm" (Burns et al., 1986) is applied to generate the primitive linear features. The algorithm extracts linear features by finding regions of similar gradient direction. For each extracted feature, a set of attributes is defined and calculated for subsequent processing. Due to the complexity in the images stated in Section 1, the primitive features extracted from aerial images are usually highly fragmented, especially for those derived from large-scale aerial images in urban areas. Hence they need to be grouped. The grouping is based on the perceptual grouping method (Lowe, 1985; Mohan and Nevatia, 1992), and is achieved through a series of operations (Trinder and Zhao, 1998). After grouping, those features located within individual detected building regions are selected as building segments for the 3D matching and reconstruction process. All extracted building segments are stored in a graph, in which segments and their geometric relationship are represented as nodes and arcs. Accurate individual building boundaries are constructed by traversing the graph using constrained depth-first search to find closed loops in the graph.



(a) (b)

Figure 3. Original Images (a) Left Image (b) Right Image

5. TEST AND RESULTS

A pair of aerial images of Mildura, New South Wales was used for the tests (Figure 3), each with about 1600 x 1650 pixels. The image scale is 1:4,000 and the pixel size is 24 x 24 cm^2 on ground. The area covered by the images is a mixture of industrial and residential areas, with many trees, trucks and roads around some of the buildings requiring extraction. The sizes of buildings vary greatly. Tests have been carried out on a rectangular portion, marked on Figure 3(a). The length and width of the rectangle also represent the X and Y directions of DSM to be generated.

A DSM was generated on Socet Set digital photogrammetric workstation, with a ground sample distance (GSD) of 0.5 meters in both X and Y directions, as is shown in Figure 4. The generated DSM was segmented to derive initial regions of interest (ROIs), using the method described in Section 3.1. The height threshold parameter used for segmentation is 3 metres. The initial ROIs were projected back onto the images via collinear equations. Subsequent processing has been performed on both left and right images, but only those results yielded from the left image are presented below. Figure 5 presents the results of initial DSM segmentation in image space.

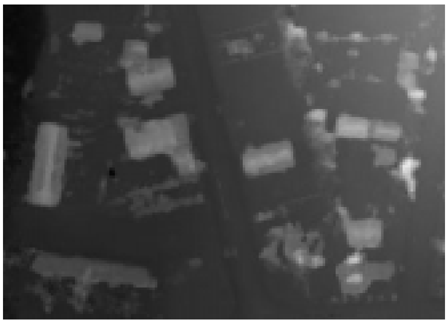


Figure 4. DSM

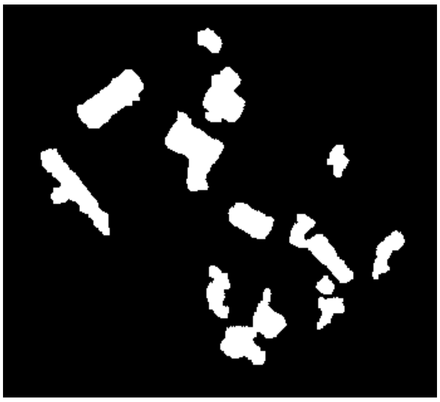


Figure 5. Initial ROIs

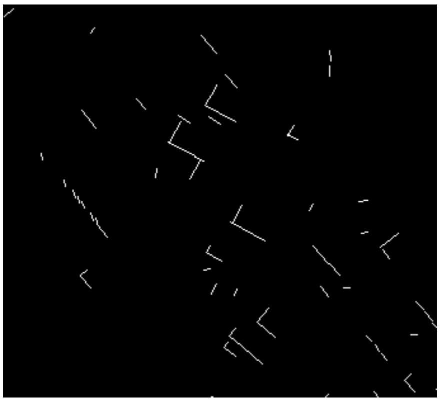


Figure 6. Building Shadow Lines

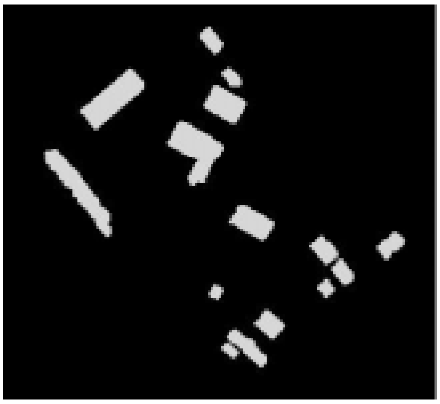


Figure 7. Detected Building Regions

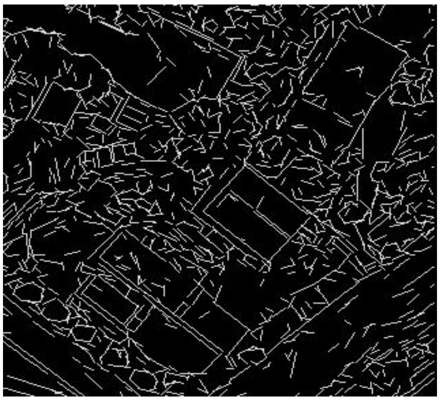
The initial ROIs were refined with texture information and verified by building shadow lines, which were extracted using approaches presented in Section 3.2 and 3.3 respectively. Figure 6 illustrates the extracted building shadow lines. Figure 7 shows the final ROIs, from which it can be seen that all buildings in the test area have been successfully detected and separated from other objects. Based on the detected building regions, buildings were extracted and reconstructed locally through all steps described in Section 4. Figure 8 shows a set of processing results carried out on a window of 500 x 500, presenting the original image (a), extracted linear features (b), detected building regions (c) and the extracted building boundaries (d) respectively.

6. CONCLUSIONS AND FUTURE WORK

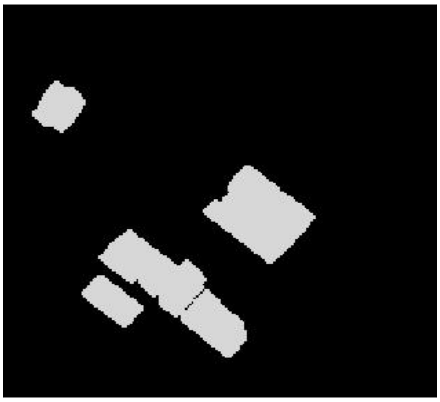
In this paper, we have presented a hierarchical framework for automatic building extraction from aerial images. It is based on the integration of multiple information, knowledge and algorithms, with close cooperation between 2D and 3D information. The whole task is accomplished through global detection to local extraction and modelling. The system applies a generic polyhedral model, and is hence potentially able to handle buildings of different shapes and structures.

The test results show that its two major components: building detection and segment extraction, operate with satisfactory performance. At present, feature matching and 3D modelling components are under development. Future work will be focused on these areas.

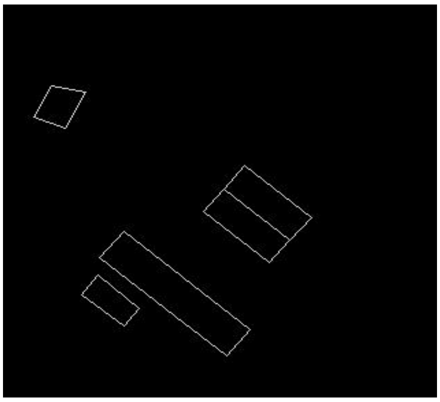
(a) Original Image



(b) Linear Features



(c) Detected building regions



(d) Extracted building boundaries

Figure 8. Processing Results

REFERENCES

Baillard, C., Schmid, C., Zisserman, A. and Fitzgibbon, A., 1999. Automatic line matching and 3D reconstruction of buildings from multiple views. *International Archives of Photogrammetry and Remote Sensing,* Vol.32, Part 3-2W5, pp. 69-80.

Baltsavias, E, Mason, S. and Stallmann, D., 1995. Use of DTMs/DSMs and orthoimages to support building extraction. In Grün, A., Kübler, O. and Agouris, P. (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images,* Birkhäuser Verlag, Basel, pp. 199-210.

Berthod, M., Gabet, L., Giraudon, G. and Lotti, J. L., 1995. High resolution stereo for the detection of buildings. . In Grün, A., Kübler, O. and Agouris, P. (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images,* Birkhäuser Verlag, Basel, pp. 135-144.

Bruun, A., Gülch, E., Lang, F., Förstner, 1998. A hybrid concept for 3D building acquisition. *ISPRS Jouranl of Photogrammetry and Remote Sending,* Vol.53:119-129.

Brunn, A. and Weidner, U., 1998. Hierarchical Bayesian nets for building extraction using dense digital surface models. *ISPRS Jouranl of Photogrammetry and Remote Sending,* Vol.53:296-307.

Burns, J. B., Hanson, A. R., and Riseman, E. M, 1986. Extracting straight lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol.8:425-455.

Grün, A., 1998. TOBAGO – a simi-automated approach for the generation of 3-D buildings models. *ISPRS Jouranl of Photogrammetry and Remote Sending,* Vol.53:108-118.

Haala, N. and Hahn, M., 1995. Data fusion for the detection and reconstruction of buildings. In Grün, A., Kübler, O. and Agouris, P. (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images,* Birkhäuser Verlag, Basel, pp. 211-220.

Henricsson, O, 1996. *Analysis of Image Structures Using Color Attributes and Similarity Relations*. Ph.D. thesis, Institute of Geodesy and Photogrammetry, Zürich.

Huertas, A. and Nevatia, R., 1988. Detecting buildings in aerial images. *Computer Vision, Graphics, and Image Processing,* Vol.41:131-152.

Hug, C., 1994. The scanning laser altitude and reflectance sensor – an instrument for efficient 3D terrain survey. *International Archives of Photogrammetry and Remote Sensing,* Vol.30, Part1, pp.100-107.

Hug, C., Wehr, A., 1997. Detecting and identifying topographic objects in imaging laser altimetry data. *International Archives of Photogrammetry and Remote Sensing.* Vol.32, Part 3-4W2, pp.19-26.

Irvin, R. B. and McKeown, D. M., 1989. Methods for exploiting the relationship between buildings and their shadow in aerial imagery. *IEEE Transactions on Systems, Man, and Cybernetics,* Vol.19(6):1564-1675.

Kim, T., 1995. *Automated 3D Modelling of Buildings from Aerial and Space Imagery Using Image Understanding Techniques*. Ph.D. thesis, University College London.

Laws, K., 1980. *Texture Image Segmentation.* Ph.D. thesis, University of Southern California.

Lin, C., Huertas, A. and Nevatia, R., 1995. Detection of buildings from monocular images. In Grün, A., Kübler, O. and Agouris, P. (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images,* Birkhäuser Verlag, Basel, pp.125-134.

Liow, Y. T., 1990. Use of shadows for extracting buildings in aerial images. *Computer Vision, Graphics, and Image Processing,* Vol.49:242-277.

Lowe, D. G., 1985. *Perceptual Organisation and Visual Recognition*. Kluwer Academic Publisher, Boston.

Mohan, R. and Nevatia, R., 1989. Using perceptual organization to extract 3-D structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol.11(11):1121-1139.

Roux, M. and McKeown, D. M., 1994. Feature matching for building extraction from multiple views. In *Proceedings of DARPA Image Understanding Workshop*, Monterey, CA, pp. 331-349.

Trinder, J. C. and Zhao, B., 1998. Heuristic feature derivation for automatic building extraction. In Li, D., Gong, J. and Chen, X. (eds), *Spatial Information Science, Technology and Its Applications*, WTUSM Press, Wuhan, pp.237-242.

Weidner, U., 1996. An approach to building extraction from digital surface models. *International Archives of Photogrammetry and Remote Sensing,* Vol.31(B3):924-929.

Weidner, U., and Förstner, W., 1995. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS Jouranl of Photogrammetry and Remote Sending,* Vol.50(4):38-49.