The Journal of GIS Association of Korea, Vol. 16, No. 4, pp. 479-488, December 2008

Investigation of Building Extraction Methodologies within the Framework of Sensory Data

Suyoung Seo*

ABSTRACT

This paper performs investigation of the state-of-the-art approaches to building extraction in terms of their sensory input data and methodologies. For the last decades, there have been many types of sensory input data introduced into the mapping science and engineering field, which are considerably diverse in aspects of spatial resolution and data processing. With the cutting-edge technology in this field, accordingly, one of the key issues in GIS is to reconstruct three-dimensional virtual models of the real world to meet the requirements occurring in spatial applications such as urban design, disaster management, and civil works. Thus, this study investigates the strengths and weaknesses of previous approaches to automating building extraction with two categories – building detection and modeling and with sensor types categorized. The findings in this study can be utilized in enhancing automation algorithms and choosing suitable sensors, so that they can be optimized for a specific purpose.

Keywords : building extraction, sensory data, building detection, building modeling

요 약

본 논문은 건물추출 기법에 있어서 기존 연구들을 분석하기 위하여 활용센서의 관점에서 분류하고 이에 따라 주요 기법들을 분석하였다. 지난 수십 년간, 많은 유형의 센서자료가 매핑 과학기술분야에 도입되어 왔고, 그들의 특성은 공간해상력과 자료처리 측면에서 매우 다양하다. 이에 따라, GIS분야에서 이러한 자료들을 이용함으로써 실세계의 3차원 가상공간 구축이 가능해지고 있고, 이는 도시설계, 재난관리 및 토목공사 등 여러 응용분야에 활용될 수 있을 것으로 기대된다. 이와 관련하여, 본 연구는 건물추출에 대한 기존 연구들을 먼저 건물 영역탐지와 모델링의 두 가지 대주제로 분류한 후 기법들에 대한 장단점을 분석하였다. 본 연구에서 제시하는 분석결과는 목적에 따라 자동화 알고리즘을 향상시키거나 적합한 센서자료를 선택

---

* Korean Land Spatialization Group (지능형국토정보기술혁신사업단)

하는데 있어서 유용하게 활용될 수 있을 것이다.

주요어 : 건물 추출, 센서 자료, 건물 영역 탐지, 건물 모델링

1. Introduction

For the last decades, the rapid growth of sensing and computing technology in the mapping field has provided essential solutions in many spatial applications such as urban planning, telecommunication, transportation, disaster management, and other spatial analysis applications. Because of the size of spatial data and the requirement of high fidelity, many research efforts have been made in automation of data processing and quality control to meet the productivity and quality standards required in the real world.

In this context, extraction of buildings can be considered as one of the most crucial needs to be filled with to achieve virtual reality models of the real world. Accordingly, rigorous studies have been conducted to obtain building models from varying sensory input data. In this regard, in Korea, extraction of buildings has been studied using recent sensing technology such as lidar in Song et al. (2008) and BIM (Building Information Model) in Goh et al. (2008). Although previous studies provided solutions applicable, however, there certainly exist technical issues to be resolved for extracting buildings with high automation and fidelity.

Thus, this paper investigates previous methodologies of building extraction in a systematic way in order to figure out trends and key achievements of the approaches and to proceed further to advanced algorithms. For clarity, approaches to building extraction are categorized into region detection and shape modeling. Section 2 presents findings from the analysis of approaches to building region detection and Section 3 those to building model extraction. Finally, concluding remarks are stated in Section 4.

2. Studies on Building Region Extraction

This paper divided approaches to detecting building regions into the groups categorized by sensory input data as single optical image feature-driven, elevation-driven and multisensor fusion-driven approaches. Table 1 summarizes the studies which will be discussed in the following sections.

2.1 Building Region Extractions using Single Optical Image

Approaches in this group extract primitive features such as edges and homogeneous regions based on brightness contrasts and then, represent connectivities among edges and regions in a graph.

Closed polygons are generated by joining primitive features with consideration of their topological relationship. In the approaches, lines are aggregated into closed polygons by assembling line segments which are likely to compose parallelograms while considering the viewing geometry of camera models. To evaluate the likelihood of line combinations, Lin and Nevatia (1998) used certainty factors to reduce search space. Kim and Nevatia (1999) generated closed polygons based on neural networks, whereby Baysian networks were exploited to express evidences derived from natural dependencies.

Katartzis et al. (2001) classified relationships between features using a Markov Random Field (MRF), whereby relationships are classified into supporting or competing ones, describing geometric and photometric attributes of the features. For verification of buildings hypothesized, shadows and walls were exploited in Lin and Nevatia (1998). In Collins et al. (1998), through camera models, building boundary hypotheses were projected into other images and verified with features collected along epipolar zone.

Table 1. Studies on Building Region Extraction

Sensory Data	Studies
Single Optical Image	Levitt (1995), Collins et al. (1998) Lin and Nevatia (1998) Kim and Nevatia (1999) Katartzis et al. (2001)
Elevation Data	Baltsavias et al. (1995) Eckstein and Munkelt (1995) Haala and Hahn (1995) Weidner and Föstner (1995) Berthod et al. (1995) Bailard et al. (1998) Brunn and Weidner (1998) Wang (1998) Cord and Declerq (2001) Seo (2003)
Data Fusion	Csatho et al. (1999) McKeown et al. (1999) Haala and Brenner (1999) Tao and Yasuoka (2002)

Levitt (1995) exploited both edges and homogeneous regions together to find building regions and presented a snake algorithm for verification and refinement. Their snake algorithm takes boundaries of the homogeneous regions as initial outlines of buildings and integrates the geometry of edges and regions into the energy function. The results show, however, that some building regions were missed due to low contrasts along building outlines and lack of evidences.

To summarize, although approaches in this group have presented rigorous perceptual organization of primitive features, their performance, however, has been shown to have limitations to extract complex buildings due to the geometric capacity of a single image.

2.2 Building Region Extractions using Elevation Data

The approaches in this group can be subdivided into the following categories: morphological operation, comparison of neighboring heights, edge detection, optimal configuration and contour based methods.

In approaches using morphological operations, non-ground objects whose size is larger than a specific window size are suppressed by morphological opening (Baltsavias et al., 1995; Eckstein and Munkelt, 1995; Haala and Hahn, 1995; Weidner and Föstner, 1995; Brunn and Weidner, 1998). Then, building regions are obtained by subtracting the derived surface from the original DSM (Digital Surface Model), producing a surface model referred to as a normalized DSM. To perform the process properly, however, a suitable window size should be selected to detect the largest building while preserving terrain features.

In Berthod et al. (1995) and Cord and Declerq (2001), homogeneous regions were extracted and classified into ground and nonground regions by comparison of elevations of their neighboring regions.

Baltsavias et al. (1995) and Wang (1998) extracted edges and classified their shapes to distinguish building regions from other features. It is an advantage of this type of approaches that there is no need to set the structural element size as compared with approaches using morphological operations.

Brunn and Weidner (1998) and Baillard et al. (1998) extracted building regions through optimal configuration of region classifications. Brunn and Weidner (1998) classified regions by Bayesian networks where elevation data and step and crease edges were considered together. Baillard et al. (1998) extracted building blobs by MRF optimization of ground and nonground segments.

Utilization of contours has been attempted in Baltsavias et al. (1995) and Seo (2003). In Baltsavias et al. (1995), discrete elevation edges were utilized from coarse to fine levels. Seo (2003) established contour graphs to represent hierarchical adjacency among contours and analyzed slopes between neighboring contours to detect building regions.

In general, approaches in this group have shown good approximations of building regions with less suffering from invalid hypotheses.

2.3 Building Region Extractions using Data Fusion

Because of availability of input data from a variety of sensors and their strengths, data fusion-based approaches have been performed rigorously as follows.

Csatho et al. (1999) presented a general strategy to recognize buildings using multisensor data sets that include aerial images, LIDAR data and multispectral images. McKeown et al. (1999) proposed two methods to detect building regions: stereo-based and classification-based methods. They utilized regions from classification of hyperspectral images and DSM to prune line segments which are not likely related to building boundaries. It is shown that both methods could reduce the number of false hypotheses remarkably.

Tao and Yasuoka (2002) exploited heuristic knowledge about the response of object surfaces to sensors, whereby. NDVIs (Normalized difference Vegetation Indexes) from IKONOS images and elevations from DSM data were analyzed together for the knowledge-based classification. Haala and Brenner (1999) merged DSM and multispectral data at pixel level and performed ISODATA classification of the merged data so that objects could be classified into buildings, trees, and grass-covered areas.

To summarize, fusion-based approaches to building detection have shown to be effective to search for building regions and further to verify and characterize them because of data redundancy and compensation.

3. Studies on Building Model Extraction

In building model extraction, the procedure needs to assemble primitive features into building models and then can verify them in aspects of their topology and geometry if other evidences are available. Although rigorous research efforts have been made to automate the procedure, however, it is not trivial due to the complexity of building shapes and incomplete collection of primitive features. This section investigates approaches to building model extraction in terms of data type of input data and involvement of human interaction (Table 2).

Table 2. Studies on Building Model Extraction

Sensory Data	Studies
Single Optical Image	Lin and Nevatia (1998) McKeown et al. (1999)
Stereo-images	Roux and McKeown (1994) Collins et al. (1995) Willuhn (1996) Fischer et al. (1998) Henricssion (1998) Baillard et al. (1999)
Elevation Data	Weidner and Föstner (1995) Brunn and Weidner (1998) Cord and Declerq (2001)
Data Fusion	Brenner and Haala (1998) Haala and Brenner (1999) McKeown et al. (1999) Ameri (2000) Vosselman and Dijkman (2001) Seo (2003)
Semi-automatic Way	Gruen (1998) Gruen and Wang (1998) Läbe and Gülch (1998) Gülch et al. (1999) Rottensteiner (2000)

3.1 Building Model Extractions using Single Optical Images

Lin and Nevatia (1998) and McKeown et al. (1999) presented building modeling methods which utilized walls and shadow casts in an image whereby the illumination direction and camera geometry were considered in geometric modeling of building heights. Then, for verification of building models, building edges on ground and shadows were exploited.

Because available features are limited to one image, the approaches exploit the domain knowledge about camera model, building shapes and other object space conditions to predict three-dimensional shapes of buildings. Although the approaches have great merits in that three dimensional buildings can be automatically generated from a single two dimensional image, however, they are shown to be limited to simple shapes building models and quality of resulting models seems to depend on camera geometry and illumination condition.

3.2 Building Model Extractions using Stereo-images

Conventionally, aerial images with overlaps have been the typical data type for building model extraction. Their epipolar geometry inherent in overlapped images facilitates generation and verification of building models. Grouping mechanism, however, needs to overcome incompleteness of features in image domain and the complexity of building models in object space.

To generate building models, Roux and McKeown (1994) presented an approach utilizing the connectivity among corners and combining line segments into 3D planes. They analyzed radiometric consistency between primitive features to discard invalid surfaces and avoided combinatorial explosion of matching pairs of corners based on epipolar constraints, height ranges, corner types and corner directions. Collins et al. (1995) determined the most probable height of a building using a voting scheme using the distribution of heights of 3D line segments and models building shapes through estimation of parameters while considering coplanarity and rectangularity. Bailllard et al. (1999) extracted 3D lines through matching and grouped and refined them by a half plane whereby its rotation parameter about the lines is determined with features in other overlapped images. Then, they aggregated the half planes with other coplanar and collinear planes into a building model.

Hierarchical strategies have been presented to improve feature aggregation and to come up with building complexity. Willuhn (1996) proposed a blackboard reasoning system, whereby knowledge about images, features and building models are taken into account for at feature, structural and conceptual levels. At feature level, primitive features are extracted and their relationships are utilized to produce 3D line segments. Next, at structural level, 3D lines are grouped into 3D surface patches based on the knowledge-based rules about chromatic and photometric attributes. At conceptual level, natural characteristics of the buildings such as roof ridge lines and symmetric relationships were proposed to be exploited.

Fischer et al. (1998) proposed to use corner models as bases to aggregate primitive features and to generate higher level of abstractions. They hypothesized building part models from 3D corner models and assembled them into a complete building model by merging and gluing processes. Henricsson (1998) stated that color attributes can be strong cues for grouping beside geometric attributes. They extracted and grouped regions based on the similarity of color attributes, whereas they generated 3D planes from 3D lines using geometric attributes. Then, building models were generated by joining 3D patches with consideration of topological relationships between surface patches.

To summarize, approaches using stereo-images have been shown to have abilities to extract complex buildings. Nevertheless, however, matching ambiguities caused a non-trivial number of invalid hypotheses to be verified that those approaches suffered from.

3.3 Building Model Extractions using Elevation Data

To extract building models, Brunn and Weidner (1998) extracted crease and jump edges, and grouped homogeneous regions into planar surface patches. Then, they modeled buildings with regularities such as parallelism and symmetry constituted among surface patches. Weidner and Föstner (1995) exploited building boundaries obtained from building detection to generate parametric and prismatic building models, whereby the boundaries from images were refined based on a MDL (Minimum Description Length). Cord and Declerq (2001) performed a Monte Carlo type of simulation to group 3D points directly to surface patches. In their approach, the number of patches to be extracted was predefined and optimal configuration between points and planes which can maximize probabilities predefined was determined.

From the observations above, elevation data can be stated to be advantageous over optical intensity imagery such that approach using them could generate three-dimensional components of a building model without suffering from matching ambiguities. One crucial drawback of elevation data, however, such as lidar data and DSM from aerial photographs is that localization of building boundaries in the horizontal plane cannot be performed precisely as compared with aerial photographs.

3.4 Building Model Extractions using Data fusion

Modeling buildings with features from multisensory input data has certainly advantages because of their synergy effect when combined, producing redundancy and enhancing geometric accuracy. Data to be integrated are obtained from various sources such as aerial images, LIDAR data, ground plan, multispectral images and hyperspectral images. Usage of ground plans from topographic maps and GIS data has been studied in Brenner and Haala (1998), Haala and Brenner (1999) and Vosselman and Dijkman (2001). In the approaches, roof surface patches were initiated from ground plans and decomposed into rectangular regions. Then, they segmented further lidar points in each rectangular region by a Hough transform and least squares fitting and refined planar patches by splitting and merging processes. Utilizing hyperspectral imagery was presented to verify building model hypotheses generated from panchromatic imagery and to identify materials of roof surfaces by McKeown et al. (1999).

Building extraction which exploits elevation data and aerial images has been studied in Ameri (2000) and Seo (2003). Ameri (2000) extracted planar surface patches from elevation data and generated building model hypotheses by combining them. Then, building hypotheses were verified and refined through being compared with edge pixels in aerial images. Seo (2003) proposed a grouping method, utilizing pair of symmetric planar patches called wing model, to aggregate 3D planar patches into a complete building model and verified a building model based on edge vectors in aerial images.

In summary, the previous approaches utilizing multisensor data have shown efficiency in finding valid building models and localizing their entities.

3.5 Building Model Extractions using Semi-automatic Way

Because of the variety and complexity of building structures in the real world, human interpretation of the building models has been considered highly reliable as human operator can perceive and classify building shapes quickly with high fidelity. Nonetheless, however, automation of feature collection and accurate adjustment of them to a certain building model are needed to be combined for better productivity, requiring development of semi-automatic building extraction systems.

Gruen (1998) and Gruen and Wang (1998) presented methods to generate complex building models using critical points in their systems, called TOBAGO and CC modeler. Visual inspection observed critical points and computerized procedure analyzed their distribution and generated building models based on predefined rules, called parser. In Läbe and Gülch (1998), Gülch et al. (1999), and Rottensteiner (2000), initial building model and points were given by human operator and building models were generated and refined by the collection of linear features within the proximity of the model.

To summarize, semi-automatic approaches can be considered to have productivity to meet the current needs of real world mapping projects.

4. Conclusion

This paper manifested the trend and issues of building extraction methodologies through analytical investigation, stating their strengths and weaknesses in association with sensory input data. The findings indicate that the quality of buildings extracted depends largely on the type of sensory input data and that utilization of domain knowledge and combination of other available data be important not only to circumvent limitations of a certain type of sensory data but also to expedite the processing time.

One aspect still needed to establish a high performance building extraction system, however, is to assess the influence of input data quality on output building models so that research efforts can be focused more on bottlenecks in the processing flow, which may be found in any factors among primitive features, matching, error propagation, data models, aggregation, data availability, processing time, and other factors.

Acknowledgement

This research was partly supported by a grant (08KLSGB01) from Cutting-edge Urban Development – Korean Land Spatialization Research Project funded by Ministry of Land, Transport and Maritime Affairs of Korean government.

References

Ameri, B., 2000. Automatic recognition and 3D reconstruction of buildings through computer vision and digital photogrammetry. Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.

Baillard, C., Dissard, O., Jamet, O., Maitre, H., 1998. Extraction and textural characterization of above-ground areas from aerial stereo pairs: a quality assessment. ISPRS Journal of Photogrammetry and Remote Sensing 53 (2):130-141.

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

Berthod, M., Gabet, L., Giraudon, G., Lotti, J., 1995. High-resolution stereo for the detection of buildings. In: Gruen et al. (1995), pp. 135-144.

Brenner, C., Haala, N., 1998. Fast production of virtual reality city models. In: Fritsch, C., M., E., M., S. (Eds.), IARPS Commission IV Symposium on GIS - Between Vision and Applications. Stuttgart, Germany, pp. 77-84.

Brunn, A., Weidner, U., 1998. Hierarchical Bayesian nets for building extraction using dense digital surface models. ISPRS Journal of Photogrammetry and Remote Sensing 53 (5), 296-307.

Collins, R. T., Jaynes, C. O., Cheng, Y.-Q., Wang, X., Stolle, F., Riseman, E. M., Hanson, A., 1998. The ascender system: Automated site modeling from multiple aerial images. Computer Vision and Image Understanding 72 (2), 143-162.

Cord, M., Declercq, D., 2001. Three-dimensional building detection and modeling using a statistical approach. IEEE Transactions on Image Processing 10 (5), 715-723.

Csatho, B., Schenk, T., Lee, D.-C., Filin, S., 1999. Inclusion of multispectral data into object recognition. In: IAPRS Vol.32, Part 7-4-3 W6. Valladolid, Spain, pp.53-61.

Eckstein, W., Munkelt, O., 1995. Extracting objects from digital terrain models. In: Schenk, T. (Ed.), Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes. Proc. SPIE 2572. pp. 222-231.

Fischer, A., Kolbe, T. H., Lang, F., Cremers, A. B., F¨orstner, W., Pl¨umer, L., Steinhage, V., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. Computer Vision and Image Understanding: CVIU 72 (2), 185-203.

Goh, I.-D., Choi, J.-H., Kim, E.-D., Jeong, Y.-S., Lee, J., 2008, "Extracting building geometry from BIM for 3-D city model," The Journal of GIS Association of Korea 16(2):249-261.

Gruen, A., Kuebler, O., Agouris, P. (Eds.), 1995. Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkh¨auser Verlag, Ascona.

Gruen, A., 1998. TOBAGO - a semi-automated approach for the generation of 3-D building models. ISPRS Journal of Photogrammetry and Remote Sensing 53 (2):108-118.

Gruen, A., Wang, X., 1998. CC-Modeler: a topology generator for 3-D city models. ISPRS Journal of Photogrammetry and Remote Sensing 53: 286-295.

Gülch, E., M¨uller, H., L¨abe, T., 1999. Integration of automatic processes into semiautomatic building extraction. In: IAPRS Vol.32, Part3-2W5, Proceedings of ISPRS Conference on Automatic Extraction Of GIS Objects From Digital Imagery. München.

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

Haala, N., Hahn, M., 1995. Data fusion for the detection and reconstruction of buildings. In: Gruen et al. (1995), pp. 211-220.

Henricsson, O., 1998. The role of color attributes and similarity grouping in 3-D building reconstruction. Computer Vision and Image Understanding 72 (2): 163-184.

Katartzis, A., Sahli, H., Nyssen, E., Cornelis, J., 2001. Detection of buildings from a single airborne image using a Markov random field model. In: Proc. of IEEE International Geoscience and Remote Sensing Symposium (IGRASS 2001). Sydney, Australia.

Kim, Z., Nevatia, R., 1999. Uncertain reasoning and learning for feature grouping. Computer Vision and Image Understanding 76 (3): 278-288.

Läbe, T., Gülch, E., 1998. Robust techniques for estimating parameters of 3D building primitives. In: Proceedings of ISPRS Commission II Symposium. Cambridge, UK, pp. 169-176.

Levitt, S., 1995. Building recognition in aerial images. URL citeseer.nj.nec.com/levitt95building.html Lin, C., Nevatia, R., 1998. Building detection and description from a single intensity image. Computer Vision and Image Understanding 72 (2): 101-121.

Lin, C., Nevatia, R., 1998, "Building detection and description from a single intensity image," Computer Vision and Image Understanding 72 (2): 101-121.

McKeown, D. M., Cochran, S. D., Ford, S. J., McGlone, J. C., Shufelt, J. A., Yocum, D. A., 1999. Fusion of HYDICE hyperspectral data with panchromatic imagery for cartographic feature extraction. IEEE Transactions on Geoscience and Remote Sensing 37 (3):1261-1277.

Rottensteiner, F., 2000. Semi-automatic building reconstruction integrated in strict bundle block adjustment. In: Proceedings of the XIXth ISPRS Congress at Amsterdam, International Archives of Photogrammetry and Remote Sensing 33(B3), Amsterdam, The Netherlands, pp. 461-468.

Roux, M., McKeown, D. M., 1994. Feature matching for building extraction from multiple views. In: Proceedings of the 1994 IEEE Conference on Computer Vision and Pattern Recognition. Seattle, WA.

Seo, S., 2003. Model-Based Building Extraction from LIDAR and Aerial Imagery, Ph.D. Dissertation, The Ohio State University, Columbus, Ohio, 138 p.

Song, C.-C., Lee, W.-K., Jeong, H.-S., Lee, K.-K, 2008, 3D surface model reconstruction of aerial LIDAR (Light Detection And Ranging) data considering land-cover type and topographic characteristic, The Journal of GIS Association of Korea 16(1):19-32.

Tao, G., Yasuoka, Y., 2002. Combining high resolution satellite imagery and airborne laser scanning data for generating bareland DEM in urban areas. In: International Workshop on Visualization and Animation of Landscape, IAPRS Vol. 34, Part 5/W3. Kunming, China.

Vosselman, G., Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. In: Workshop on Land Surface Mapping Characterization using Laser Altimetry. Annapolis, MD.

Wang, Z., 1998. Extracting building information from LIDAR data. In: Schenk, T., Habib, A. (Eds.), IAPRS Object Recognition and Scene Classification from Multispectral and Multisensor Pixels. Vol. 32(3/1). Columbus, OH, pp. 279-284.

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

Willuhn, W., Ade, F., 1996. A rule-based system for house reconstruction from aerial images. In: International Conference on Pattern Recognition. International Association on Pattern Recognition, IEEE Computer Society Press, pp. 885-889.

Received (September 14, 2008)
Revised (December 17, 2008)
Accepted (December 17, 2008)