Quality of Building Extraction from IKONOS Imagery

Jie Shan[1] and Scott D. Lee[2]

Abstract: Now to ensure the quality of automatic extraction of buildings from aerial and space imagery is not comprehensively understood. Dedicated efforts are needed in examining both the theoretical and practical aspects of the problem. This paper first presents a review of building extraction research activities, with the focus on quality-related issues. Also addressed is recent progress in the use of high-resolution satellite images. As a theoretical study, a set of five quality measures of the extracted buildings is proposed that considers all possible types of errors and describes the accuracy of building extraction relative to the ground truth. To achieve a reliable and precision result, a step-by-step quality control strategy is implemented in a classification-guided building extraction procedure. Unreliable results are identified based on the proposed quality index and excluded from further processing. Empirical results using IKONOS images are presented to demonstrate the proposed quality measures and quality control strategy.

DOI: 10.1061/(ASCE)0733-9453(2005)131:1(27)

CE Database subject headings: Imaging techniques; Satellites; Photogrammetry; Digital mapping; Quality control.

Introduction

Automatic extraction of buildings from aerial and space imagery has been a focus of research for many years. A collection of state-of-the-art articles can be found in the periodical proceedings edited by Grün et al. (1995, 1997) and Baltsavias et al. (2001a). In addition, Mayer (1999) presents a comprehensive survey of the techniques used for image-based building extraction. The data sources used for building extraction vary from images to their combinational use with additional data or information, such as topographic maps (Förstner and Pluemer 1997; Haala 1999, lidar or other elevation data (Weidner and Förstner 1995; Baillard and Maitre 1999; Hofmann 2001; Kim and Muller 2001), multispectral and/or hyperspectral images (Haala 1999; Huertas et al. 1999, 2000; McKeown et al. 1999; Lee et al. 2003), as well as various building models (Braun et al. 1995; Förstner and Pluemer 1997; Förstner 1999). A few spinoff products are available on the market, such as the semiautomatic building collection tool inJECT developed by Inpho GmbH, Feature Analyst from Visual Learning Systems, and eCognition from Definiens Imaging GmbH, all three of which are able to extract building boundaries from multispectral images based on image classification or clustering results.

These studies have recently been used for building extraction from IKONOS satellite images. Baltsavias et al. (2001b) and Fraser et al. (2001, 2002) present results on building extraction from IKONOS stereo images. A comparative study of results obtained from aerial photographs concludes that about 15% of the building areas measured in aerial images cannot be modeled using IKONOS images. An assessment based on 19 GPS surveyed check points at roof corners suggests that the IKONOS-derived building models can reach an accuracy of better than 1 m, both in planimetry and elevation (Baltsavias et al. 2001b). Sohn and Dowman (2001) use a local Fourier transformation to analyze the dominant orientation angle in a building cluster and extract rectilinear building outlines from IKONOS imagery based on a binary space partitioning tree. Dial et al. (2001) present an investigation of automated road extraction in wide suburban roads. Lee et al. (2003) report a class-guided building extraction solution and present an approach to building squaring based on Hough transformation.

Yet despite the abundant existing work on building extraction, only a few studies have addressed the quality of the extracted buildings. Baltsavias et al. (2001b) and Fraser et al. (2002) discuss the capability of IKONOS imagery and reveal its limitations for building extraction. McKeown et al. (1999) and Lee et al. (2003) provide quantitative studies for the extracted buildings from aerial and IKONOS images, respectively. These studies suggest the need for a theoretical framework that formulates the complete measures for building extraction and the related quality control issues.

This paper addresses quality measures and quality control in building extraction with examples using IKONOS images. A set of five quantitative measures is proposed to formulate the quality of the extracted buildings. These measures describe the accuracy and possible mistakes in building extraction. The building extraction process and results undergo a quality control process through which each extracted building is associated with a quality index such that the manual interaction and assessment need only be carried out for the extracted buildings with low-quality indices. A quality control methodology that mimics the traffic signal systems is proposed and implemented based on the natural break classification method (Fisher 1958; Jenks and Caspall 1971)

To validate the proposed quality measures and control methodology, an IKONOS multispectral and panchromatic image is used for the study. Detailed quantitative results are reported to examine the quality measure and control issues in building extraction. On the one hand, the study verifies the completeness and

[1]Geomatics Engineering, School of Civil Engineering, Purdue Univ., 550 Stadium Mall Dr., West Lafayette, IN 47907-2051.

[2]Geomatics Engineering, School of Civil Engineering, Purdue Univ., 550 Stadium Mall Dr., West Lafayette, IN 47907-2051.

Note. Discussion open until July 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on January 26, 2004; approved on June 22, 2004. This paper is part of the ***Journal of Surveying Engineering***, Vol. 131, No. 1, February 1, 2005. ©ASCE, ISSN 0733-9453/2005/1-27–32/$25.00.

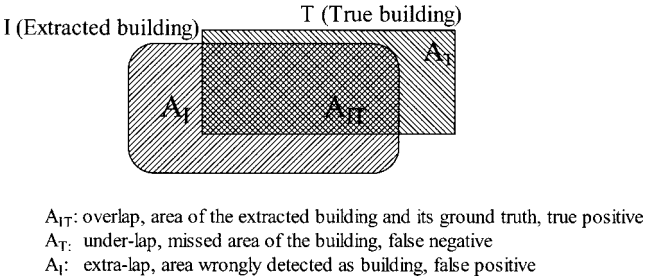

Fig. 1. Quality measures for extracted buildings

correctness of the proposed quality measures, and on the other, it also reveals in a quantitative manner the potential and limitation of building extraction using IKONOS images. The quality measures formulated here are general and can be used for building extraction using other image or data sources. They can also be easily generalized into three dimensions to assess the quality of building reconstruction.

The remainder of the paper is organized as follows. The next section briefly describes the test data and the building extraction method, illustrated with examples. A quality index is derived that examines attributes of each extracted building for quality control in the subsequent process. After that we present a quality control process that mimics the traffic signal system with which the extracted buildings are finally classified into true alarm and false alarm categories through an iterative clustering calculation. Results are presented to show the outcome of the quality control process. Next, we propose a set of theoretical measures for the quality of the extracted buildings, followed in the next section by evaluation of the quality of our tests with IKONOS images. Concluding remarks and future work are addressed at the end of the paper.

Quality Measures

This section will formulate the measures to assess the quality of the extracted buildings. Traditionally we are familiar with quality measures of point measurements, such as using the standard deviation to measure the accuracy of point coordinates. Although this approach can still be used to accurately describe the position of extracted buildings, these are complex features with a certain geometric shape of spatial extension. New measures need to be developed to describe the uncertainty or errors of the extracted buildings.

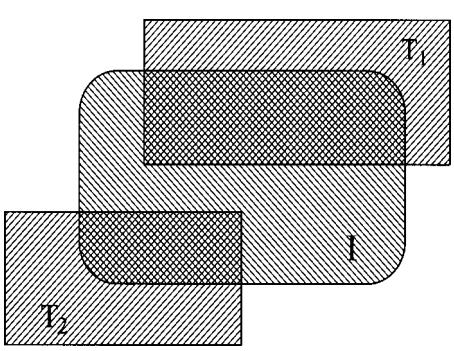

Fig. 2. Crosslap: one extracted building (I) overlaps with two true buildings (T_1,T_2)

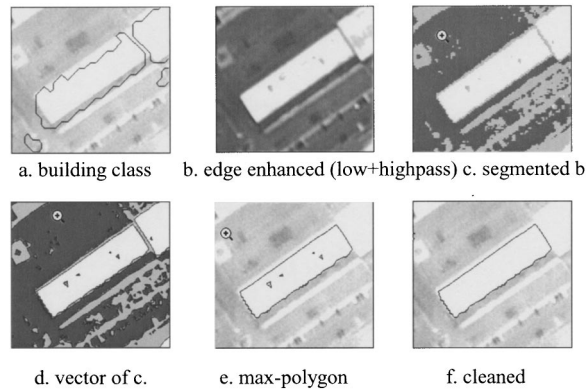

Fig. 3. Steps of class-guided building extraction: (a) building polygon from classification overlaid atop panchromatic image; (b) enhanced image with low-pass and high-pass filter; (c) segmentation of enhanced image, three classes; (d) vectorization of segmentation; (e) segmentation polygon with maximum overlap with classification polygon in (a); (f) final building polygon cleaned of speckles

Five quality measures will be defined as follows to comprehensively describe the quality of extracted buildings. As shown in Fig. 1, the true building and the corresponding extracted building are shown in rectangular and oval polygons, respectively. The two overlaid polygons are in general divided into three parts, which respectively represent the correctness and mistakes in building extraction. The first part, P_{IT}, is the common part or overlap of the true and extracted buildings; this stands for the completeness of the extracted building. The second part, P_T, is the region in the ground truth missed by the extracted building. It is therefore called underlap or misslap, which stand for the false negative error in building extraction. The third part, P_I, is extralap, namely the area that is mistakenly labeled as a building, or false positive/alarm error. To represent these quantities in a relative manner, we use the following quantitative expressions of Eqs. (1)–(3), where the operator $A(.)$ is the polygon area calculation. The common denominator in Eqs. (1)–(3) is the area of the true building polygon.

$$\text{underlap} = \frac{A(P_T)}{A(T)} \tag{1}$$

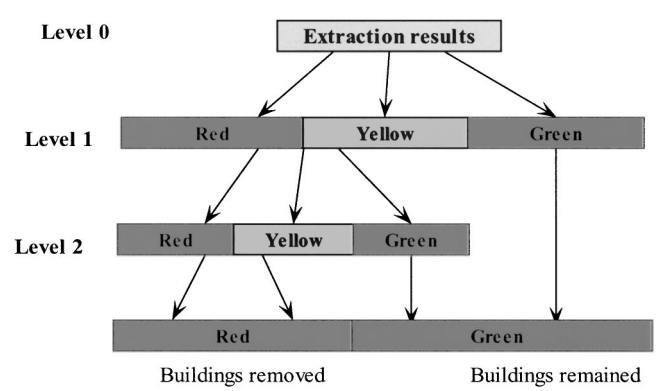

Fig. 4. Hierarchical quality control: extracted buildings are classified into red, yellow, and green classes recursively; the red class is likely a mistake, while the green class has fairly high confidence

Table 1. Results of Hierarchical Quality Control

Category	Total	Green	Yellow	Red
Level 1	459	73	206	180
Level 2	386	129	151	106

Note: Final number of building's is 202 and nonbuilding's is 257.

$$\text{extralap} = \frac{A(P_I)}{A(T)} \tag{2}$$

$$\text{overlap} = \frac{A(P_{IT})}{A(T)} \tag{3}$$

Note that another possible mistake exists besides simple underlap and extralap. As shown in Fig. 2, one extracted building may intersect with more than one true building, or, in other words, the extralap area of one extracted building relative to one true building (T1) can overlap with another true building (T2). This error is noted as crosslap and represented by the number of true buildings that an extracted building may intersect with. For example, the extracted building I in Fig. 2 has a crosslap number 2. This measure is represented in Eq. (4).

$$\text{crosslap} = \text{number of true buildings intersected} \tag{4}$$

To evaluate the overall quality of the extracted buildings, both underlap and extralap errors need to be considered. Therefore we propose the following overall quality measure:

$$\text{fitness} = \frac{A(P_{IT})}{A(P)} = \frac{A(P_{IT})}{A(P_I) + A(P_T) + A(P_{IT})} = \frac{A(I \cap T)}{A(I \cup T)} \tag{5}$$

In summary, among the above measures, underlap and extralap respectively represent the two common possible mistakes in building extraction. Underlap means missing a true building, namely a false negative error, whereas extralap stands for the wrong claim of buildings, namely a false positive error. Crosslap means the number of true buildings an extracted building may intersect with. Ideally, the cross-lap number should be 1 while cross-lap 0 corresponds to a completely false alarm. The overlap and fitness represent the completeness of the extracted building. Overlap stands for the percentage of the correctly detected building, while fitness considers both extralap and underlap errors and is the most critical quality measure. All measures are given in a relative manner, defined as a ratio of polygon areas. Finally, it should be noted that overlap and underlap measures are not independent; their sum is equal to 1, but they reflect error and accuracy, two aspects of the extracted buildings.

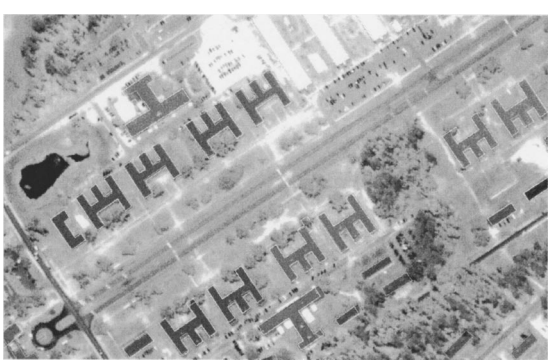

Fig. 5. Extracted and squared buildings

Building Extraction

Building extraction is conducted by using IKONOS satellite images. The study area covers $5.8 \cdot 3.6\ \text{km}^2$ surrounding Camp Lejeune, North Carolina. The images are georectified with a nominal horizontal accuracy of 25 m by the image vendor Space Imaging, Inc. (Dial 2000).

A class-guided approach (Lee et al. 2003) is used to extract the buildings. In general, the approach consists of two major sequential steps: classification and refinement. The classification uses multispectral images, with the objective of obtaining approximate locations and shapes of potential buildings. Supervised classification classifies the image into seven classes: building, road, tree, grass, marsh, sand, and water. As the primary interest, the majority of building pixels ($>90\%$) are correctly identified and correspond to the general building shapes on the ground. However, a visual check suggests that a large number of nonbuilding pixels are incorrectly labeled as buildings (false positive errors). Most of these pixels are road pixels and form a lot of small, irregular speckles scattered over the entire image. Some of them appear in the vicinity of buildings and severely change the building shapes. All this makes it impossible for the building class to be directly used without further refinement.

For refinement, building classes are first vectorized to form polygons of the potential buildings. The refinement is conducted through image segmentation and building squaring based on the panchromatic image guided by the building class. Each potential building polygon obtained in the classification will define an extent window for the refinement. The corresponding panchromatic image within the extent window is segmented (usually into three classes) using the ISODATA (iterative self-organizing data analysis techniques algorithm) (Mather 1999) clustering approach. The segmentation results are vectorized into polygons, which are then compared with the potential building polygon obtained from the classification. The segmented polygon that has a large overlap and small nonoverlap with the classification polygon will be chosen as the candidate building polygon. A final squaring or delineation process is then applied to regularize the building boarder. This process is illustrated step by step in Fig. 3.

The candidate building is selected based on a quality index modified by using the quality measures described above. This selection is made possible by the use of both multispectral and panchromatic images. Referring to Fig. 1, the classification and segmentation polygons are denoted as P_X and P_S, respectively, while their overlap part is denoted as P_C. In the ideal situation, the overlap should be the same as both the classification result and the segmentation result. Therefore, we can use the following equation to attribute the selected candidate building polygon:

$$Q = \frac{2A(P_c)}{A(P_x) + A(P_s)} \tag{6}$$

The value of the quality index varies from 0 to 1. If there is no overlap between the classification and segmentation results, the quality index will be zero, but if the two polygons exactly overlap each other, then the quality index will be 1, which means a perfect match between the segmentation and classification results. Segmentation polygons with the largest quality index will be selected as the candidate building. Every such selected building polygon will have the quality index associated with it as an attribute that will be used for quality control and evaluation in the subsequent process.

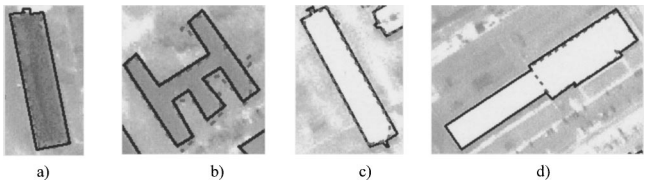

Fig. 6. Quality statistics of extracted buildings (solid: ground truth; dashed extracted results): (a) overlap: 89.9%; underlap: 10.1%; extralap: 3.2%; fitness: 87.0%; (b) overlap: 91.1%; underlap: 8.9%; extralap: 11.4%; fitness: 81.8%; (c) overlap: 91.8%; underlap: 8.2%; extralap: 0.8%; fitness: 91.0%; and (d) overlap: 56.4%; underlap: 43.6%; extralap: 0.1%; fitness: 56.4%

Quality Control

The extracted buildings undergo a quality control process whose objective is to separate the extracted buildings with a high level of confidence from those that are possible false alarms. Through this process, it is expected that the manual and human interaction for quality control can be significantly reduced or even minimized. To reach this objective, a hierarchical quality control strategy is proposed that mimics the traffic signal system, which classifies the extracted buildings into three classes (Förstner 1999): the green class has fairly good quality; the yellow class has only a certain level of confidence and needs an additional check; and the red class has a high likelihood of being a mistake. The difficulty is to determine the break points of the three classes so that each class can have a minimum classification error.

In this article we propose a hierarchical quality control method

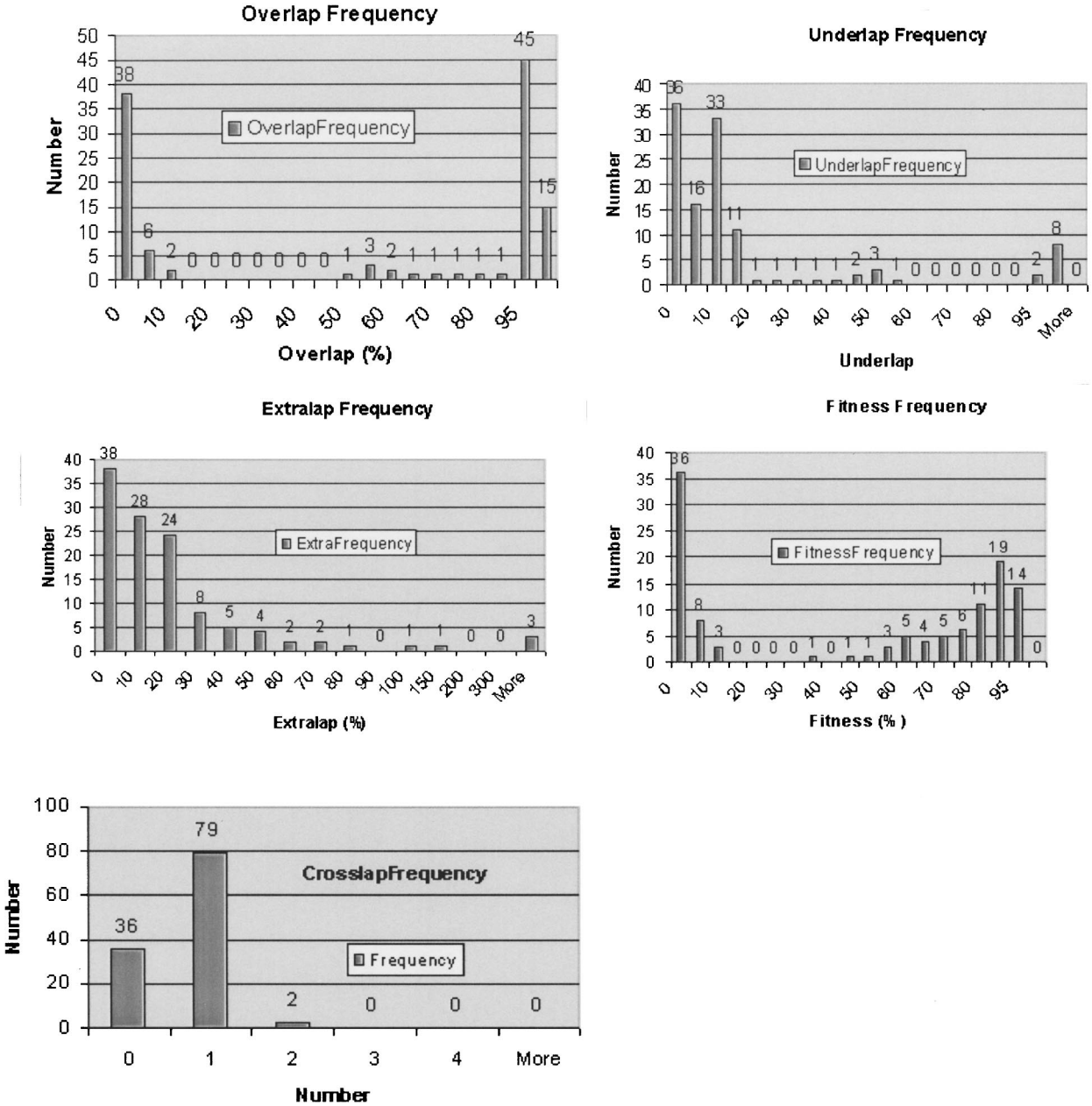

Fig. 7. Quality statistics for final extracted buildings

based on the natural break classification approach (Fisher 1958; Dent 1996). This classification approach divides the data into classes such that the variation within classes is as small as possible, while the variation among classes is as large as possible. As is shown in Fig. 4, level 0 contains all the extracted possible buildings, which are classified into three classes, shown in green, yellow and red in level 1 according to their quality indices calculated using Eq. (5). This process is continued for the combined red and yellow classes in level 1 to further screen out possible good buildings in level 2.

The reason we include both red and yellow classes for further review is to reduce the false negative mistake, namely not to wrongly exclude true buildings. Such a mistake is considered to be more severe than false positive. At the last level, the yellow and red classes will undertake a further inspection or be treated as mistakes, while all the green classes through this process will be taken as buildings of high confidence. This process continues for two times as shown in Fig. 4. The final building category contains buildings labeled as green in each level. Through this process, we expect to separate the good buildings from the suspicious ones so that the manual check can be carried out only for those red buildings. Our study shows that this quality process is necessary due to the large percentage of false alarms in the original classification results. Table 1 presents the number of buildings in each class at different levels and the final number of buildings removed as possible false alarms and retained as correct buildings with high confidence.

Final removed and retained buildings through the quality control process are shown in Fig. 5. The dark lines are manually delineated buildings that will be used as a ground truth for quality evaluation later. After removing the red and yellow polygons, which are likely wrongly extracted buildings and will receive a manual confirmation, the final selected buildings are squared and shown in Fig. 5.

Quality Assessment

A 1 km·1 km area is chosen to assess the quality of the extracted buildings, and 112 buildings in this area are visually identified and manually delineated and then used as the ground truth to assess the quality of the automatically extracted buildings. Fig. 5 shows a subset of the extracted buildings followed by a squaring process in (Lee et al. 2003).

All the extracted buildings are assessed against the above proposed quality measures in Eqs. (1)–(5). Fig. 6 presents the values of quality measures for some delineated buildings, where solid and dash lines are for ground truth and extracted results, respectively. As shown in this figure, the extracted buildings (a)–(c) have rather good overall quality. The calculated quality measures reflect the consistency of the extracted buildings relative to their ground truth. The overlap rate can be as large as 91.8%, while the extralap or false alarm rate may reach 11.4% for some building due to misinterpretation and misalignment at the building borders. Fig. 6(d) is a special case, where the manual and automatic delineation disagrees on whether the object consist of one or two buildings. Under such circumstance, the quality measures reflect this disagreement of a large underlap and small fitness rate, though the extralap rate or false alarm rate is still very small.

Fig. 7 presents the statistics for 117 final extracted buildings: 36 of the building polygons have zero crosslap, which means they are completely false alarm, and 60 (53.7%) building polygons have an overlap larger than 95%, which indicates a very good coverage of the ground truth. However, as shown in the extralap chart, about 80 well-extracted buildings (good overlap rate and crosslap number) have above 10% false alarm errors. A few of them can even be as high as 300%, caused by trees and occlusion in the vicinity of the buildings. As an overall quality measure, 47 (42.0%) of the total true buildings have an overall fitness rate better than 85%.

Conclusions

Quality control and assessment are crucial for a successful implementation of an automatic process such as building extraction. The combined use of multispectral and panchromatic images makes it possible to obtain a quality index to attribute to each extracted potential building for quality control and assessment purposes. The hierarchical quality control methodology can distinguish fairly well between the true buildings and false alarms in the building extraction results. On the one hand, this will greatly increase the reliability of the final outcomes from an automatic extraction process, and on the other, significantly reduce human interaction with the computer. The proposed five quality measures provide a comprehensive definition and description about the accuracy and uncertainty of the extracted buildings. Based on these quality measures, it is shown that 53.7% of the extracted buildings have an overlap rate of over 95% while 42.0% have an overall fitness better than 85%.

Several factors undermine the quality of the class-guided building extraction approach. Our experience shows that misclassification of spectrally similar features (for example, buildings and roads) remains a major source of error. Limitations on image spatial resolution cause a certain ambiguity in separating adjacent buildings. Low image contrast makes it difficult to separate a building from its vicinity features. To further raise the level of automation for building extraction, images with higher spectral, spatial, and radiometric resolution can be of great help. The combinational use of heterogeneous data such as image and lidar (light detection and ranging) data may compensate for the limitation of current image data sources.

References

Baillard, C., and Maitre, H. (1999). "3D reconstruction of urban scenes from aerial stereo imagery: A focusing strategy." *Comput. Vis. Image Underst.*, 76(3), 244–258.

Baltsavias, E., Grün, A., and van Gool, L., (eds.). (2001a). *Automated extraction of man-made objects from aerial and space images (III), Ascona, Switzerland*, Balkema, Lisse, The Netherlands.

Baltsavias, E., Pateraki, M., and Zhang, L. (2001b). "Radiometric and geometric evaluation of Ikonos GEO images and their use for 3D building modeling." *Proc., Joint ISPRS Workshop, High Resolution Mapping from Space 2001* (CD-ROM), Hannover, Int. Society for Photogrammetry and Remote Sensing, London.

Braun, C., et al. (1995). "Models for photogrammetric building reconstruction." *Comput. Graphics*, 19(1), 109–118.

Dent, B. D. (1996). *Cartography—Thematic map design*, Wm. C. Brown Publishers, 132–137.

Dial, G. (2000). "IKONOS satellite mapping accuracy." *Proc., ASPRS Annual Conf.* (CD-ROM), American Society for Photogrammetry and Remote Sensing, Bethesda, Md.

Dial, G., Gibson, L., and Poulsen, R. (2001). "IKONOS satellite imagery and its use in automated road extraction." E. Baltsavias et al., *Automated extraction of man-made objects from aerial and space images,*

Balkema, Lisse, The Netherlands, 349–358.
Fisher, W. D. (1958). "On grouping for maximum homogeneity." *Am. Stat. Assoc. J.*, 53, 789–798.
Förstner, W. (1999). "3D-city models: Automatic and semiautomatic acquisition methods." D. Fritsch and R. Spiller, *Photogrammetric week*, Wichmann, Heidelberg, Germany, 291–303.
Förstner, W., and Pluemer, L. (1997). *Semantic modeling for the acquisition of topographic information from images and maps*, Birkhaeuser Verlag, Basel, Switzerland.
Fraser, C., Baltsavias, E., and Grün, A. (2002). "Processing of Ikonos imagery for submetre 3D positioning and building extraction." *ISPRS J. Photogr. Remote Sens.*, 56(3), 177–194.
Fraser, C., Hanley, H. B., and Yamakawa, T. (2001). "Sub-metre geopositioning with Ikonos GEO imagery." *Proc., Joint ISPRS Workshop, High Resolution Mapping from Space* (CD-ROM), Int. Society for Photogrammetry and Remote Sensing, London.
Grün, A., Baltsavias, E., and Henricsson, O., eds. (1997). *Automatic extraction of man-made objects from aerial and space images*, Birkhaeuser Verlag, Basel, Switzerland.
Grün, A., Kuebler, O., and Agouris, P., eds. (1995). *Automatic extraction of man-made objects from aerial and space images*, Birkhaeuser Verlag, Basel, Switzerland.
Haala, R. (1999). "Combining multiple data sources for urban data acquisition." D. Fritsch and R. Spiller, *Photogrammetric week*, Wichmann Verlag, Heidelberg, Germany, 329–340.
Hofmann, P. (2001). "Detecting buildings and roads from IKONOS data using additional elevation information." *GeoBIT/GIS*, 6, 28–33.
Huertas, A., Kim, Z., and Nevatia, R. (2000). "Multisensor integration for building modeling." *Proc., Computer Vision and Pattern Recognition* (CD-ROM). IEEE, New York.
Huertas, A., Nevatia, R., and Landgrebe, D. (1999). "Use of hyperspectral data with intensity images for automatic building modeling." *Proc., 2nd Int. Conf. on Information Fusion* (CD-ROM).
Jenks, G. F., and Caspall, F. C. (1971). "Errors on choropleth maps." *Ann. Assoc. Am. Geogr.* , 61, 217–244.
Kim, J.-R., and Muller, J.-P. (2001). "Assessment of automated techniques for extracting vegetation and buildings from 1 m stereo multispectral IKONOS and 1 m pan-sharpened IKONOS coupled with scanning laser altimetry." *Proc., Joint ISPRS Workshop, High Resolution Mapping from Space 2001* (CD-ROM), Int. Society for Photogrammetry and Remote Sensing, London.
Lee, D. S., Shan, J., and Bethel, J. (2003). "Class-guided building extraction from IKONOS imagery." *Photogramm. Eng. Remote Sens.*, 69(2), 143–150.
Mather, P. (1999). *Computer processing of remotely-sensed images: An introduction*, Chapter 8, 2nd Ed., Wiley, New York.
Mayer, H. (1999). "Automatic object extraction from aerial imagery—A survey focusing on buildings." *Comput. Vis. Image Underst.*, 74(2), 138–149.
McKeown Jr., D., Cochran, S., Ford, S., McGlone, C., Shufelt, J., and Yocum, D. (1999). "Fusion of HYDICE hyperspectral data with panchromatic feature extraction." *IEEE Trans. Geosci. Remote Sens.*, 37(3), 1261–1277.
Sohn, G., and Dowman, I. J. (2001). "Extraction of buildings from high resolution satellite data." E. Baltsavias et al., *Automated extraction of man-made objects from aerial and space images*, Balkema, Lisse, The Netherlands, 345–355.
Weidner, U., and Förstner, W. (1995). "Towards automatic building reconstruction from high resolution digital elevation models." *ISPRS J. Photogr. Remote Sens.*, 50(4), 38–49.