Automated Buildings and Road Extraction from High Resolution Panchromatic Satellite Images

[1]R. Aruna, [2]D.Sasireka
[1]PG Scholar [2]Associate Professor
Department of Computer Science and Engineering
Jayamatha Engineering College
Aralvaimozhi.

ABSTRACT
Extraction of map objects such as buildings, roads, and rivers from high resolution satellite imagery is an important task in many civilian and military applications. The study of aerial and satellite images is known as Photogrammetry; which is the science of making measurements from photographs, especially for recovering the exact positions of surface points. Moreover, it may be used to recover the location of any object (buildings), on its components and in the immediately adjacent environment. According to rapidly growing urbanization and municipal regions, it is much difficulty to detect the buildings automatically. Thus, automatic detection of buildings from remote sensing images is a hot topic and an active field of research. In the proposed system, the buildings in the urban areas can be detected by using the internal gray variance (IGV). First the satellite images are enhanced by using morphological operators, which enhances the edges of the objects. Then multiseed-based clustering technique detects the building edge using the variance in the gray levels. In order to reduce the false alarm, very small regions that are unlikely to be buildings are removed. Finally, by using the adaptive thresholding segmentation technique, the buildings are segmented.

Keywords:
Building Detection, Clustering, High Resolution, Internal Gray Variance, Morphology, Remote Sensing, Segmentation

1. INTRODUCTION

Automatic detection of geographical objects such as bridges, big buildings, or road crossings in satellite images is useful in many important applications. These applications include maintaining geographical databases, urban planning, cartography, photo interpretation, automatic land use analysis, automatic map generation, measurement of sealed areas for public authority uses, assessing the extent of damages in case of natural disasters such as floods or earthquakes and military applications including navigation or location aware systems and emergency planning systems for evacuation and fire response. Buildings form one of the most important group of manmade objects that take significant time and cost to extract, due to their variety, complexity, and abundance in urban areas and they are significant objects in remote sensing data and directly indicate inhabited areas. In most cases, buildings and roads are well recognizable by a human interpreter. An automatic system that is able to emulate a human operator is desired [11]. Automated building extraction can save both time and labor to build and update the building and road spatial database in such applications. Use of high resolution imagery facilitates greater accuracy, but in turn increases the computational and problem complexity due to noise and artifacts. An automated building-extraction strategy for high-resolution satellite imagery is proposed that utilizes structural, contextual, and spectral information. The system runs automatically without pre-classification or any training sets, although some initial algorithm parameters must be set by the user. The proposed method exploits the properties of buildings and other objects in the scene and their interactions in the resulting images. The steps in the method are: 1) local directional morphological operations to enhance the building structures; 2) multiseed-based clustering of internal gray variance (IGV) features of enhanced image; 3) identify edges of manmade structures using bimodality detection; 5) false alarm reduction using positional information of building edge; 6) adaptive threshold-based segmentation technique. The algorithm is extensively evaluated using a variety of high-resolution satellite imagery. The algorithm uses a variety of images from IKONOS and QuickBird satellites which provides global, accurate, very high-resolution multispectral and panchromatic imagery to individuals, commercial organizations, and government agencies for a variety of urban development applications. It should be noted that the resolution of panchromatic images (≤1m) from satellite imagery is greater than that for multispectral images (2.5 m for QuickBird and 4 m for IKONOS). Hence, the panchromatic

images carry more details and lend themselves to recognizing buildings more accurately, since buildings are small structures and do not carry significant color information. Therefore, only panchromatic images are used for our research. Recent availability of high-resolution satellite images provides a new data source for geospatial data acquisition.

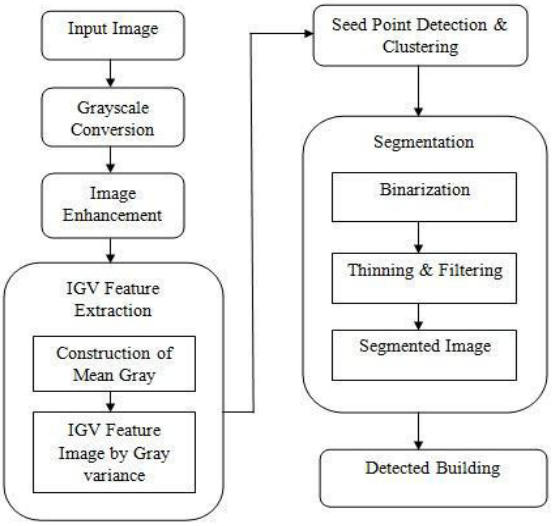
2. Literature Survey

There are many applications which require accurate information about the data contained in satellite images. Because of irregular structure and closeness of buildings in urban areas, it is difficult to extract the building automatically. To solve this problem of building detection, many strategies have been used. Here we are going to see about some of the techniques and their advantages over others.

Fraser *et al.* [12] compared buildings extracted from IKONOS imagery with those obtained using black and white aerial photographs. Thomas *et al.* [13] concluded that high-resolution imagery is a valuable tool for mapping urban areas and extracting land cover information. Sohn and Dowman [14] proposed an automatic method of extracting buildings in densely urban areas from IKONOS imagery. They detected detached buildings, however, accuracy was lacking. Others have developed algorithms utilizing feature optimization, linking edge chain, and graph matching algorithms (e.g., [15-18]) to construct objects from feature primitives. Building extraction algorithms solely based on high-resolution satellite imagery demonstrate the need for auxiliary data sources, especially those involving 3D information about the scene.

3. Building Detection Algorithm

The focus of this paper is the detection of buildings from high-resolution panchromatic images. The details of each step for detecting the buildings are presented below.



3.1 Image Enhancement

The goal of the enhancement step is to improve the visual effect in the image to facilitate geographic image interpretation and to improve the contrast between the target and non-target for high-level processing.

A key observation that is exploited here is that manmade objects tend to be more homogeneous than natural objects. Therefore, morphological operators are useful to delineate them. Since opening suppresses bright regions smaller than the structuring element, and closing suppresses dark regions, they are often used in combination as morphological filters for image smoothing and noise removal. The combination of opening and closing operations are used which uses the same structuring element along the homogeneous direction with respect to the image and a building template.

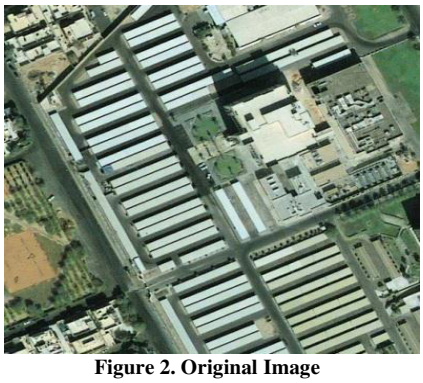


Figure 2. Original Image

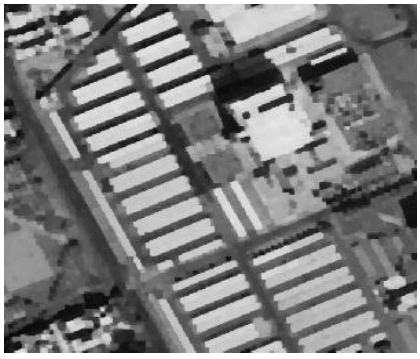


Figure 3. Enhanced Image

3.2 IGV Feature Extraction

Application of standard edge detection operator to find the boundaries of buildings is unlikely to be successful because the manmade objects are not well separated from the natural objects. After enhancement, the internal regions of the manmade structures are more homogeneous than the border region. The variation of the internal pixels of a manmade structure is low and the corresponding IGV will also small. At the same time, the IGV will be high at the border pixels of the manmade structure. Furthermore, the variations of the internal pixels and border pixels of the nonmanmade structure (natural structure) are very less. Therefore, the IGV for internal as well as border pixels of nonmanmade structures will almost the

Figure 1. System Architecture

same. So the IGV feature extraction is proposed to enhance the difference between the manmade and nonmanmade structures

which leads to the identification of boundary points of manmade structures.

The IGV is computed by first computing the gray level mean within a window w × w (w = 5). Compute the mean gray value at each pixel (x, y) in the enhanced image (*h*) within the window as,

$$\mu(x,y) = \frac{1}{w \times w} \sum_{i=-w/2}^{w/2} \sum_{i=-w/2}^{w/2} h(x+i, y+j) \quad (1)$$

Then, the candidate pixel value is generated by computing the sum of deviation from each pixel from the mean gray value within the window. Compute IGV at (x, y) within the window as,

$$IGV(x,y) = \sum_{i=-w/2}^{w/2} \sum_{i=-w/2}^{w/2} [h(x+i, y+j) - \mu(x,y)]^2 \quad (2)$$

The sum of deviations at the boundary of manmade objects is significantly more than that of points inside the manmade objects. In contrast, the sums of deviations at the boundary of nonmanmade objects are not as high as manmade objects.

3.3 Seed Point Detection

Clustering-based segmentation is used to isolate manmade structures from the background. Seed point detection is an essential part of clustering technique. Seed- based clustering methods start with some initial seed points and grow cluster around them. There are two fundamental problems with all seed-based techniques. One is the lack of a well-defined method to select initial seed points. The other problem is that these techniques are effective for spherically shaped clusters. For elongated or more complex methods do not always work well. To overcome this deficiency, multiseed based clustering techniques have been implemented.

The multiseed based clustering technique used both the information from the enhanced image and IGV feature values. First, seed points are detected using a multiseed technique of enhanced image. Then, the final seed points of IGV feature values are detected by using the seed points of enhanced image. These final seed points are called variance seeds (VS). VS of IGV feature values are used to cluster the IGV feature space.

If the multiseed technique is directly applied for seed point selection in the IGV feature space, it is computationally expensive due to the fact that it had a large data volume and the seed points are very close to each other due to small variation within the IGV feature space. As a result, the clusters are not well formed and not separated enough for accurate building extraction. Therefore, the seed points are selected from the enhanced image instead of the IGV feature image.

Let g_i, i = 0,1,……,L-1 be the gray level value of the enhanced image h and fr_i, i = 0,1,……..,L-1 be the corresponding frequency (total number of pixel) of g_i, i = 0,1,……,L-1 in the enhanced image. The seed point is defined as the mode, which is the gray value of highest frequency point, of the region in a particular iteration. Here, the measure of homogeneity is defined as the standard deviation (*Sd*) with respect to mode. The standard deviation (*Sd*) was computed as

$$Sd = \left[\frac{1}{\sum_{i=g_{min}}^{g_{max}} fr_i} \sum_{i=g_{min}}^{g_{max}} (m - g_i)^2 fr_i \right]^{1/2} \quad (3)$$

If $Sd > T_1$ (where T_1 is the predefined threshold), then the region is non-homogeneous. If the region is heterogeneous in a particular iteration, the region is split into two regions. 1) The left region whose pixel gray values are less than the mode. 2) The right region whose pixel gray values are greater than or equal to the mode. Again both the regions are recursively split if they do not satisfy the homogeneity criterion and find the modes as the seed points for the left and the right regions. Thus the seed point detection process is applied recursively until there is no region that can be split or the region is too small to split.

For each seed point ', i = 1,2,…..,k extracted from the enhanced image, find the set of all pixels of enhanced image whose gray values are ', i.e., PX_i $= \{(x_i, y_i) : g(x_i, y_i) = m_i\}$, x = 1,2,……..,M, y = 1,2,……..,N and for all i = 1,2,…..,k. where M and N are the rows and columns of the image. Then it map with the IGV feature space and find the values of IGV at those pixels, i.e., find $IGV(x_i, y_i)$, i = 1,2,……..,k. The corresponding gray variance was calculated as, $GV[i] = \sum_{\forall(x_i,y_i):g(x_i,y_i)=m_i} IGV(x_i, y_i)$, i = 1,2,……..,k. Finally, find the variance seed which are used to cluster the IGV feature space and the variance seed can be calculated as, $[i] = \frac{GV[i]}{\#PX_i}$, i = 1,2,…..,k, where "$\#PX_i$" is the number of points of the set PX_i.

3.4 Clustering

There are the *k* seed points (cluster centers) are obtained in the IGV feature space. Each sample in the IGV feature space is grouped to the nearest cluster center. This approach results in the grouping of the boundary pixels of manmade structures into a single cluster distinct from the clusters made by nonmanmade features.

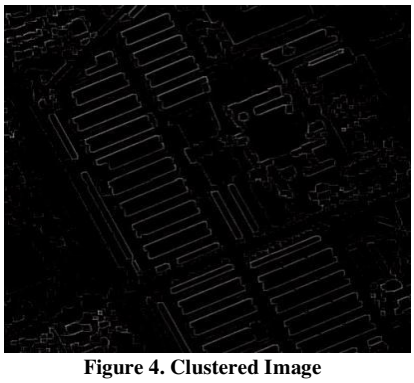


Figure 4. Clustered Image

3.5 Binarization

The clustered image from the previous steps is still a gray level image. The manmade objects have natural brightness

properties. In addition, the border regions of the manmade structures have higher values than its internal values. Still, it is difficult to extract the border regions of the manmade object accurately from the clustered image. To overcome this problem, an automatic threshold-based binarization technique is used. Although threshold-based binarization is simple, automatic detection of the threshold value is difficult. Let P be the population of points in a cluster in the IGV space. P is divided into two components subpopulations, say $P_{s(u)}$, $P_{g(u)}$ such that

1) $P_{s(u)}$ contains all the pixels with cluster value ≤ some u and $P_{g(u)}$ contains all the pixels with cluster value > u.
2) The variances of $P_{s(u)}$ and $P_{g(u)}$ are small relative to the variance of P.

Let us assume that n and σ^2 are the total frequency and variance of P, $n_{s(u)}$ and $\sigma^2_{s(u)}$ be the total frequency and variance of $P_{s(u)}$ and similarly let $n_{g(u)}$ and $\sigma^2_{g(u)}$ be the total frequency and variance of $P_{g(u)}$. Now, the gray level value μ' can be determined as,

$$\mu' = \frac{n_{s(u)}\sigma^2_{s(u)} + n_{g(u)}\sigma^2_{g(u)}}{n\sigma^2}, \qquad cl_{min} \le u \le cl_{max} \tag{4}$$

where cl_{min} and cl_{max} are the minimum and maximum gray value values of P. The gray level value μ' is called bimodality parameter. Now, the cluster image is binarized by using this bimodality parameter μ'

$$V(x,y) = \begin{cases} 0, & if\ cl_f(x,y) < \mu' \\ 1, & if\ cl_f(x,y) \ge \mu' \end{cases} \tag{5}$$

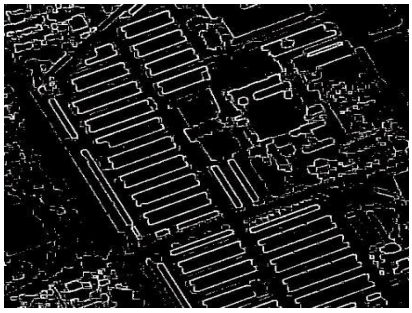


Figure 5. Binary Image

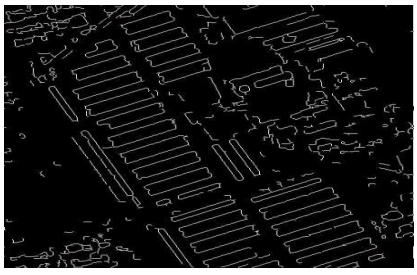


Figure 6. Small Component Filtered Image

3.6 False Alarm Reduction

Many false alarms present in the image after the binarization step can be removed by using some domain knowledge. For reducing the false alarms, the thinning and component filtering operations are used. The binary image formed after the steps described before highlights the boundaries of the manmade structures. But these boundaries are thicker. To improve the results, we use a thinning step to derive the linear/curvilinear representation of the boundaries. After thinning, many small regions remain due to noise. Very small regions that are unlikely to be buildings can be removed. These small regions are removed using a simple filtering step based on the length. Thus, all thinned regions whose lengths are less than a predefined threshold are eliminated for further processing.

3.7 Segmentation

When the targets and the background in the image are well contrasted, thresholding with a suitable value can result in effective segmentation. The approach proposed by Otsu [7] is used that algorithm chooses the intensity value that maximizes a measure of class separability computed as the ratio between class variance to the local variance. Enhanced image is used for final building detection step by local adaptive thresholding-based segmentation procedure. Any local adaptive thresholding-based operation depends on the region size, i.e., the area of the image in which the operation will be implemented for segmentation. Each connected edge region of the final building edge image is considered as candidate for region thresholding. The adaptive threshold-based segmentation procedure using the final building edge and enhanced images is described as below.

1) First, find four points, i.e., left-most, right-most, topmost, and bottom-most points of a particular connected edge region from the final building edge image. These four points give the bounding rectangle region and this rectangle is the region of interest for segmentation.
2) Find the corresponding rectangular region of the enhanced image.
3) Apply Otsu's algorithm within the rectangular region of the enhanced image and find the threshold value (O_T).
4) Segment the rectangular region of the enhanced image as

$$g(x,y) = \begin{cases} 0, & if\ h(x,y) < O_T \\ 1, & if\ h(x,y) \ge O_T \end{cases} \tag{6}$$

where the pixel (x, y) belongs to the rectangular region
5) Repeat the above steps for all connected edge regions of the final building edge image.

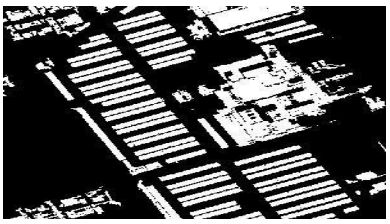


Figure 7. Final Detected Building Structures

Finally, after segmentation the set of building structures are extracted from the image. This procedure is less immune to noise but time consuming when it processed for large images.

4. RESULT

In this paper, we have evaluated the performance of our approach with scenes obtained from IKONOS and QuickBird panchromatic imagery. We have also compared our results with other algorithms in literature that have used IKONOS and QuickBird images. Here, we compare the results with the actual ground truth derived manually. From the table, we can see that the detection percentage for the proposed algorithm is higher than the other methods. Simultaneously, the branch factor is smaller than or comparable to the other methods and our proposed algorithm is significantly faster than other proposed methods in literature. It should be noted that our segmentation approach is, at the core, threshold-based and relies on the homogeneity of the building tops. If the top of the building is partially bright and partially dark, then this technique will detect only a part of the building. In general, the algorithm fails to accurately detect buildings with mixed (shade or texture) rooftops.

TABLE I
Evaluation Results of Four Images

Image	Proposed Method				
	TP	FP	TN	DP (%)	BF (%)
Image 1	35	2	4	89.7	5.4
Image 2	36	1	5	87.8	2.7
Image 3	4	1	2	66.7	20
Image 4	13	1	1	92.8	7.1
Mean	DP = 84.25% BF = 8.8% (Overall)				

The two metrics used in our research are

$$\text{Detection Percentage (DP)} = \frac{100 \times TP}{TP + TN} \quad (7)$$

$$\text{Branch Factor (BF)} = \frac{100 \times FP}{TP + FP} \quad (8)$$

DP describes how many of the existing buildings in the scene are found by the automatic approach. BF indicates how many buildings are found erroneously. TP (True Positive) is a building detected by both a person and the algorithm, FP (False Positive) is a building detected by the automatic approach but not a person, and TN (True Negative) is a building detected by a person but not by the algorithm.

5. CONCLUSION

Satellite imagery has significant information and form an important source of knowledge in many applications. Therefore, extraction of salient features from satellite imagery is a critical task. Buildings form an important class of objects in satellite imagery that carry important information about human activities. State-of-the art in automated extraction of buildings from satellite imagery is in its early stages. Two of the most difficult problems in developing accurate algorithms for this task include low signal-to-noise ratio and weak object signal in the images. Most of the algorithms in literature are not fully automated. In this work, a fully automated method to accurately extract buildings from high-resolution panchromatic remotely sensed imagery is presented. The approach exploits both the spectral and spatial properties of buildings using a multistep approach. The proposed algorithm is extensively evaluated using a variety of images from IKONOS and QuickBird satellites and compared with existing algorithms. The results demonstrate that the algorithm is both more accurate and efficient in comparison to other algorithms in literature.

REFERENCES

1. Ahmady .S, Ebadi .H, Zouj .M .J .V, and Moghaddam .H .A, "Automatic building extraction from high-resolution aerial images using active contour model," in *Proc. Int. Archi. Photogramm. Remote Sens. Spatial Inf. Sci.*, Beijing, China, 2008, vol. XXXVII, pp. 453–456, Part-B3b..
2. Chaudhuri .D, Kushwaha .N .K, and Samal .A, "Semi-automated road detection from high-resolution satellite images by directional morphological enhancement and segmentation techniques," in *Proc. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1538–1544, Oct. 2012.
3. Cui .S, Yan .Q and Reinartz .P, 2012, "Complex building description and extraction based on Hough transformation and cycle detection," *Remote Sensing Letters,* vol. 3, pp. 151-159, 2012.
4. Khoshelham .K, Li .Z and King .B, 2005, "A split-and-merge technique for automated reconstruction of roof planes," *Photogrammetric engineering and remote sensing,* vol. 71, pp. 855-862, 2005.
5. Mayunga .S .D, Zhang .Y, and Coleman .D .J, "Semi-automatic building extraction utilizing QuickBird imagery," *IAPRS*, vol. XXXVI, pp. 29–30, 2005, Part-3/W24.
6. Muller .S and Zaum .D .W, "Robust building detection in aerial images," in *Proc. Joint Workshop ISPRS Ger. Assoc. Pattern Recog. Object Extr. 3-D City Models Road Databases Traffic Monit. Concepts Algorithms Eval.*, Vienna, Austria, 2005, vol. 36 (3/W 24), pp. 143–148.

7. Otsu .N, "A threshold selection method from gray level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
8. Peng .J, Zhang .D, and Liu .Y, "An improved snake model for building detection from urban aerial images," *Pattern Recog. Lett.*, vol. 26, no. 5, pp. 587–595, 2005.

9. Segl .K and Kaufmann .H, "Detection of small objects from high-resolution panchromatic satellite images based on supervised image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 9, pp. 2080–2083, Sep. 2001.
10. Theng .L .B, "Semi-automatic building extraction from satellite imagery," *J. Eng. Lett.*, vol. 13, no. 3, pp. EL_13_3_5, 2006.
11. Wang .J, Yang .X, Qin .X, Ye .X, and Qin .Q, "An efficient approach for automatic rectangular building extraction from very high resolution optical satellite imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 388–392, Mar. 2015
12. Fraser, C.; Baltsavias, E.; Gruen, A. Processing of IKONOS imagery for sub-meter 3D positioning and building extraction. *ISPRS J. Photogramm.* 2002, *56*, 177-194.
13. Thomas, N.; Hendrix, C.; Congalton, R.G. A comparison of urban mapping methods using high-resolution digital imagery. *Photogramm. Eng. Remote Sensing* 2003, *69*, 963-972.
14. Sohn, G.; Dowman, I. Building Extraction Using LiDAR DEMs and IKONOS Images. In *Proceedings of the ISPRS Working Group III/3 Workshop*, Dresden, Germany, 8–10 October 2003; Volume 34, Part 3/W13, pp. 8-10.
15. Wei, Y.; Zhao, Z.; Song, J. Urban Building Extraction from High-resolution Satellite Panchromatic Image Using Clustering and Edge Detection. In *Proceedings of 2004 IEEE International Symposium on Geoscience and Remote Sensing*, Anchorage, AK, USA, 20–24 September 2004; Volume 7, pp. 2008-2010.
16. Liu, Z.; Wang, J.; Liu, W. Building Extraction from High Resolution Imagery Based on Multi-Scale Object Oriented Classification and Probabilistic Hough Transform. In *Proceedings of 2005 IEEE International Symposium on Geoscience and Remote Sensing*, Seoul, Korea, 25–29 July 2005; pp. 2250-2253.
17. Haverkamp, D. Automatic Building Extraction from IKONOS Imagery. In *Proceedings of the ASPRS Annual Conference*, Denver, CO, USA, 23–28 May 2004.
18. Ortner, M.; Descombes, X.; Zerubia, J. Building Extraction from Digital Elevation Model. In *Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, 6–10 April 2003; Volume 3, pp. 337-340.