5thSASTech 2011, Khavaran Higher-education Institute, Mashhad, Iran. May 12-17.

Automatic Extraction of Buildings in a Dense Urban Area from Very High Resolution Satellite Images

M. Ghanea [a], P. Moallem [b], M. Momeni [a]
[a] Dept. of Surveying Eng., Faculty of Eng., University of Isfahan, Isfahan, Iran - (m.ghane, momeni@eng.ui.ac.ir)
[b] Dept. of Electrical Eng., Faculty of Eng., University of Isfahan, Isfahan, Iran - p_moallem@eng.ui.ac.ir
Name of the Presenter: M. Ghanea

Abstract

An inclusive accessibility of very high resolution (VHR) satellite images has generated more interest in automatic extraction of buildings for some practical applications such as updating geographic information system (GIS) database, land management, cadastre, and 3D city modeling. Extraction of buildings especially in a dense urban area containing many different and connected parts is an intricate problem due to a variety of difficulties such as shadows and similar spectra. In order to face this problem properly, we present an algorithm for extraction and improvement of buildings. The process consists of four steps: (1) separate building and non-building pixels, (2) improve the building layer, and (3) segment pixels that belong to the building layer. The proposed algorithm is evaluated for a case study in Tehran, Islamic Republic of Iran using a pan sharpened multispectral GeoEye satellite image. The experiments show that the algorithm extracts 78.3 % of buildings with a quality percentage 59.7 % in a dense urban area.

Key words: Building Extraction, Very High Resolution (VHR) Satellite Images, Dense Urban Area, K-means Clustering, Region Growing

1. Introduction

The first commercial very high-resolution satellite Ikonos moved in orbit in autumn 1999, the QuickBird satellite in October 2001, OrbView-3 was launched on 26 June 2003, and GeoEye-1 was launched on 6 September 2008. These sources of high resolution images increase the amount of useful information from local to national scales (Aplin et al., 1999). These images generate amazing detail for automatic extraction of man-made objects such as roads and buildings. Extraction of buildings has become a topic of interest due to some practical applications such as updating geographic information system (GIS) database, land management, cadastre, urban planning, and 3D city modeling.

Given the recent availability of the commercial very high resolution satellite imagery, only a few methods for building extraction have been developed. The effect of high resolution on the building extraction was reported in (Sohn and Dowman, 2001, Baltsavias et al., 2001, and Muller et al., 1997). Some challenges such as low signal to noise ratio, and interpretation restrictions are commonly arisen in extracting buildings from high resolution imagery (Jin and Davis, 2004).

The task of dividing an image into homogeneous, compact and smooth regions is referred to as image segmentation. Three levels of features can be extracted from a segmented image (Aksoy et al., 2003). Level 1 features have specifications of a single image segment, such as area, perimeter, shape index, and a range of texture measurements (Herold et al., 2003). Level 2 features explain spatial relations between two objects, such as inclusion, proximity, and

adjacency. Level 3 features are spatial patterns in which more than two objects are involved. For example, in some urban areas, buildings can be irregularly distributed, while in some suburban areas, buildings can be regularly arranged (Liu et al., 2008).

Several recent 2D building extraction methods have used clustering methods without using region-based segmentation or vice vesa. Sohn and Dowman, 2001 used a building unit shape (BUS) space generated by recursive partitioning of regions. A seeded BUS then searches for its neighbors and is grown when predefined homogeneous criterion are satisfied. Croitoru and Doytsher, 2004 explained an approach based on a modified pose clustering to combine weights, constraints, and various error sources. Müller and Zaum, 2005 presented a robust building extraction method. It started with a seed growing algorithm to segment the image, then calculated the photometric and geometric features in order to classify the segments.

In this article, a fully automatic algorithm is introduced for building extraction from high resolution multispectral GeoEye satellite imagery. For this reason, we present an automatic algorithm composed of clustering methods, morphology operators, and region-growing algorithms. The proposed algorithm is evaluated for a case study in Tehran, Islamic Republic of Iran.

The remainder of this article is organized as follows: Section 2 explains the proposed algorithm. Section 3 introduces results and analyses, and Section 4 concludes the work.

2. Research Methodology

The algorithm framework is composed of four steps: (1) separate building and non-building pixels, (2) improve the building layer, and (3) segment pixels that belong to the building layer. Fig. 1 illustrates the procedure. For separation, a k-means clustering method is used to extract building and non-building pixels. Building pixels are improved by applying morphology operators such as dilation and erosion. The segmentation step includes k-means clustering, seed generation, and region growing, respectively. The following sections describe each step in detail.

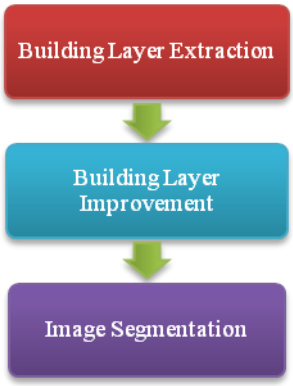


Fig 1: Proposed algorithm.

2.1 Building Layer Extraction

Separating image into building and non-building pixels is a primary step in order to prepare an approximate building layer for refinement step. K-means clustering (Shapiro and Stockman, 2001) can best be described as a partitioning method. A flow diagram of the k-means clustering algorithm is shown in Fig. 2. The function k-means partitions the pixels in the image into k mutually exclusive clusters at random, and returns a vector of indices indicating to which of the k clusters it has assigned each observation. K-means uses a two-phase iterative algorithm to minimize the sum of pixel-to-centroid distances, summed over all k clusters. The first phase, each iteration consists of reassigning pixels to their nearest centroid, all at once, followed by recalculation of centroids. This phase provides an

approximate solution as a starting point for the second phase. In the second phase, pixels are individually reassigned if the sum of distances has been reduced, and cluster centroids are recomputed after each reassignment. This two-phase procedure is repeated until it reaches the maximum iteration (m). Two parameters must be known here: 1) the number of clusters (k), and 2) the maximum iteration (m). In order to provide a binary image including building and non-building pixels, the number of clusters (k) is equivalence of 2.

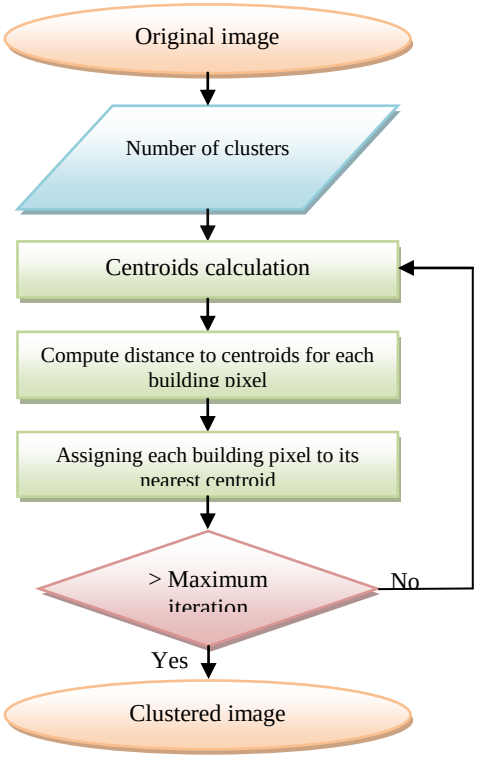
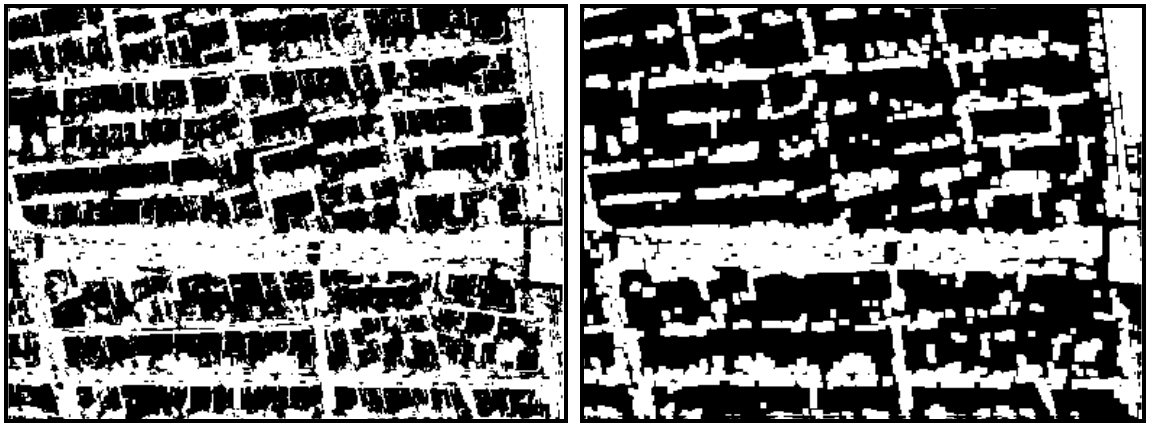


Fig 2: Flow diagram of the k-means clustering algorithm.

2.2 Building Layer Improvement

In order to refine the generated binary image, a dilation operator and an erosion operator (Najman and Talbot, 2010) with a 4 × 4 rectangular structure are applied subsequently. After refining the binary image, the product image is run out of small non-building areas such as light ducts, ventilators, aqueducts etc (see Fig. 3).



(a) (b)

Fig 3: The process of improvement the building layer: (a) Building layer with small non-building pixels, and (b) Improved building layer.

2.3 Image Segmentation

With the purpose of extracting buildings which belong to the building layer as shown in Fig. 3, we execute a segmentation method composed of k-means clustering, seed generation and region growing. First, k-means clustering partitions the building layer into predefined number of clusters. This predefined number is determined based on the variety of buildings in color, texture, and size. Each cluster includes a set of discontinuous regions (see Fig. 4b). Therefore, the histogram of the clustered image is plotted in order to compute cumulative distribution of pixels in each region (see Fig. 4c). Now, a threshold value is chosen to remove pixels which belong to small regions. Small regions include non-building regions which are similar to building regions such as roads, cars, shadows etc. The generated image is exhibited in Fig. 4d.

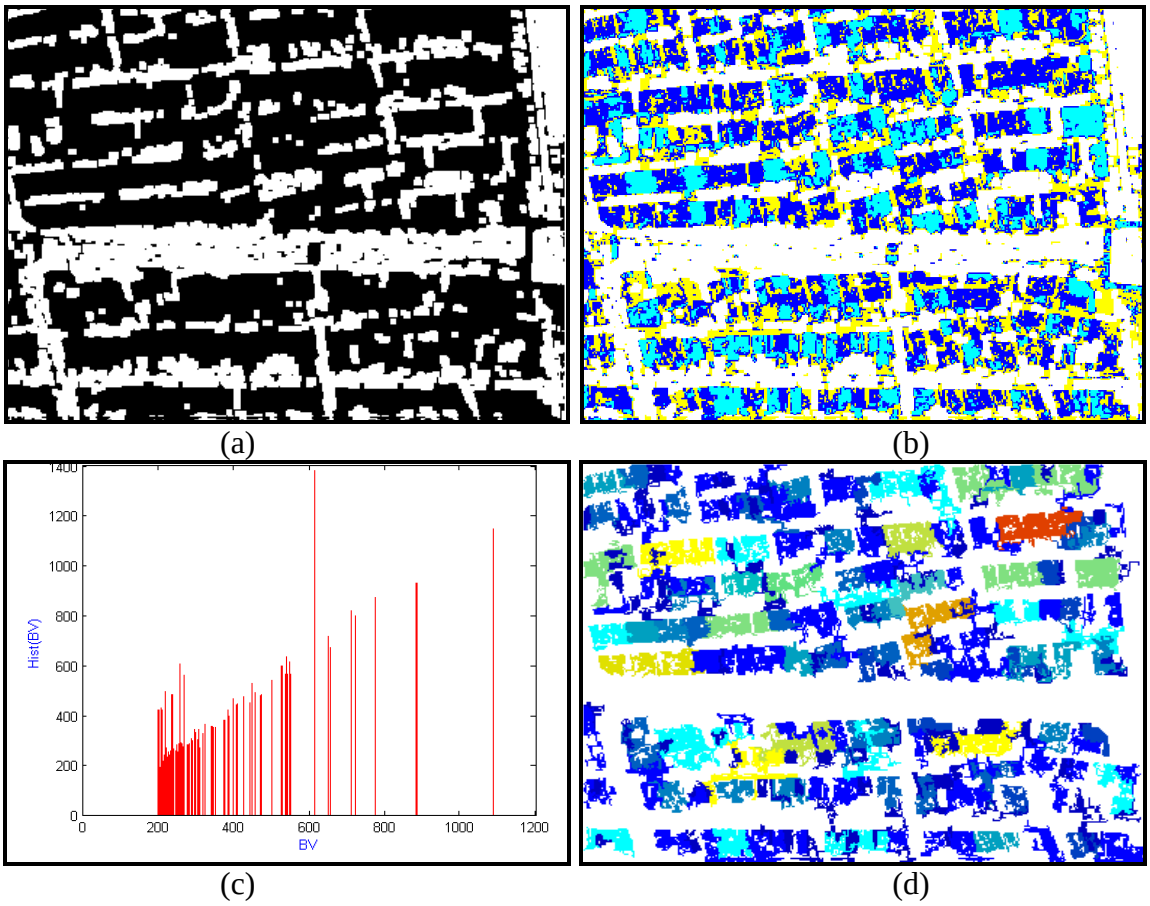


Fig 4: Applying K-means clustering and thresholding on the binary image, respectively. (a) Original image, (b) K-means clustering, (c) Histogram for the discontinuous regions, and (d) Thresholding (regions are shown with different colors).

In the second stage, seed generation is performed in order to extract seed points as the starting points for region growing stage. For extracting seed points automatically, firstly, the original image is converted into a gray scale image. Then, the position of a seed point for each region is calculated as follows:

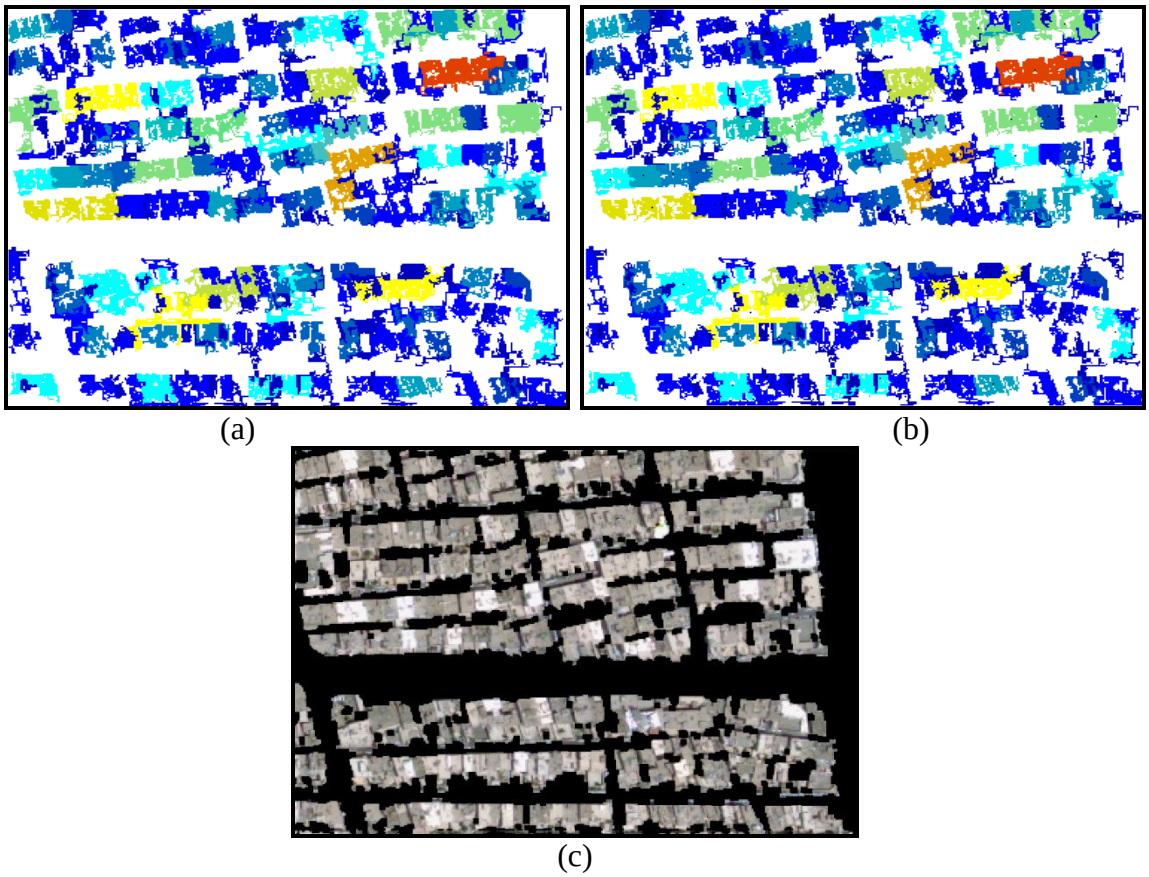
$$Seed_x = \frac{\sum_{i \in rgn} x_i . b(x_i, y_i)}{\sum_{i \in rgn} b(x_i, y_i)}$$
$$Seed_y = \frac{\sum_{i \in rgn} y_i . b(x_i, y_i)}{\sum_{i \in rgn} b(x_i, y_i)} \quad (1)$$

where *rgn* is an evaluation region and $b(x_i,y_i)$ is the brightness value of the *i*th point which belongs to the region. The product image containing regions with their seed points is shown in Figure 5.

In the third stage, region growing is applied in order to extract buildings. The procedure starts at each seed point in the image with one-pixel objects, and in numerous subsequent steps, smaller image objects are merged into bigger ones (Carleer et al., 2005). The similarity condition is defined with the help of a logical statement that is true only if pixels in the region are similar enough in terms of spectral variance property as follows:

$$\begin{cases} True & Var(b(Seed_x, Seed_y), b(x_1, y_1), .., b(x_n, y_n)) < \Delta_{rgn} \\ False & otherwise \end{cases} \quad (2)$$

where (x_1,y_1), (x_2,y_2), ..., and (x_n,y_n) are the points added to a seed point based on the similarity condition. Δ_{rgn} is a region threshold that is defined by computing the variance of all points which belong to a region. The generated image is shown in Figure 6.



(a) (b) (c)

Fig 6: Seed point generation and region growing. (a) Regions, (b) Regions with their seed points, and (c) Extracted buildings.

3. Results and Analysis

3.1. Data and Case Study

The study area is a portion of the City of Tehran, Islamic Republic of Iran. The area contains a dense urban area around the University of Tehran at the center of the city**.** We used a GeoEye satellite data which has 11 bit radiometric resolution and consists of five bands, four multi-spectral bands (i.e., Blue, Green, Red, and NIR) at a spatial resolution of 1.65 × 1.65 m, and one panchromatic band at a spatial resolution of 0.41 × 0.41 m. Before building extraction, the multi-spectral bands and the panchromatic band are pan sharpened. The size of sub-image is 2500 × 2500 pixels (i.e., 1 × 1 km). The result of building extraction for a subsection of the study area is shown in Figure 7.

(a) (b)

Fig 7: Building extraction, before and after: (a) Before (original image), (b) After (extracted buildings).

3.2. Accuracy Assessment

To evaluate the accuracy of our algorithm, we calculated factors which are introduced by Shufelt, 1999 and Lee et al., 2003. The extracted and the manually buildings are compared pixel-by-pixel. The accuracy assessment involves computation of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values by the comparison of the manual and the automatic method. TP refers to the pixels determined as building both in the manual and the automatic method. TN refers to the pixels that are not determined as building in the manual and the automatic method, respectively. FP refers to the buildings that cannot be determined and FN refers to the pixels which are determined as buildings although they are not in the manual method. The following factors are computed with the help of these values:

$$
\begin{aligned}
&Branching\ Factor = \frac{FP}{TP}, \\
&Miss\ Factor = \frac{FN}{TP}, \\
&Detection\ Percentage = \frac{100 \cdot TP}{TP + FN}, \\
&Quality\ Percentage = \frac{100 \cdot TP}{TP + FN + FP}
\end{aligned}
\tag{3}
$$

The results of the accuracy assessment show that the algorithm provides 78.3 % of DP which shows the appropriate generality of the algorithm. The QP is calculated as 59.7 % that is suitable for a dense urban area (see Table 1).

Number of pixels				Percentage			
TP	TN	FP	FN	BF	MF	DP	QP
2662354	1789432	1059291	738923	0.40	0.28	78.3	59.7

Table 1. Accuracy assessment of the proposed algorithm.

4. Conclusions

In this article, we presented a fully automatic algorithm for building extraction from a 0.41 m multispectral pan sharpened GeoEye satellite image. Our building extraction detected 78.3 % of the building areas in a dense urban area with a quality percentage of 59.7 %. However, there are still some buildings which are not extracted due to the similarities between buildings and adjacent properties. We intend to extend our algorithm in two ways. We plan to use boundary optimization methods in order to improve building boundaries. We plan to determine the initial values automatically in order to achieve the best results.

References

Aksoy, S., Tusk, C., Koperski, K., & Marchisio, G. (2003). Scene modeling and image mining with a visual grammar. *Frontiers of Remote Sensing Information Processing*. World Scientific: Singapore, 35-62.

Aplin, P., Atkinson, P.M., & Curran, P. J. (1997). Fine spatial resolution satellite sensors for the next decade. *International Journal of Remote Sensing*, 18(18), 3873-3881.

Baltsavias, E., Pateraki, M., & Zhang, L. (2001). Radiometric and geometric evaluation of Ikonos geo images and their use for 3D building modeling. *Joint ISPRS Workshop on High Resolution Mapping from Space*.

Carleer, A.P., Debeir, O., & Wolff, E. (2005). Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering & Remote Sensing*, 71(11), 1285-1294.

Croitoru, A., Doytsher, Y. (2004). Right-angle rooftop polygon extraction in regularised urban areas: cutting the corners. *The Photogrammetric Record*, 19, 311-341.

Herold, M., Liu, X., & Clarke, K. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering & Remote Sensing*, 69 (9), 991-1002.

Jin, X., & Davis, C.H. (2005). Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. *Journal on Applied Signal Processing*, 14, 2196-2206.

Lee, D. S., Shan, J., & Bethel, J. S. (2003). Class-guided building extraction from Ikonos imagery. *Photogrammetric Engineering and Remote Sensing*, 69(2), 143-150.

Liu, Z., Cui, S., & Yan, Q. (2008). Building extraction from high resolution satellite imagery based on multi-scale image segmentation and model matching. *International Workshop on Earth Observation and Remote Sensing Applications, IEEE.*

Muller, J.P., Ourzik, C., Kim, T., & Dowman, I.J. (1997). Assessment of the effects of resolution on automated DEM and building extraction. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, II, 245-256.

Müller, S., Zaum, D. W. (2005). Robust building detection in aerial images. *CMRT05, IAPRS*, Vienna, Austria, Vol. XXXVI, Part 3/W24.

Najman, L. & Talbot, H. (2010). *Mathematical morphology: from theory to applications, (Eds)*. ISTE-Wiley, 520.

Shapiro, L. G., & Stockman, G. C. (2001). *Computer Vision*. New Jersey, Prentice-Hall, 279-325.

Shufelt, J. A., (1999). Performance evaluation and analysis of monocular building extraction from aerial imagery. *IEEE Transactions on Pattern Analysis, Machine Intell*, 21(4), 311-326.

Sohn, G., & Dowman, I. J. (2001). Extraction of buildings from high-resolution satellite data. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, III, 345-355.