

# Automated Building Extraction from High-Resolution Satellite Imagery Using Spectral and Structural Information Based on Artificial Neural Networks

Zahra Lari \*,Hamid Ebadi

Dept. of photogrammetry & remote sensing, Geodesy and Geomatics Engineering Faculty, K.N.Toosi University of Technology, No.1346, Mirdamad cross, Valiasr Avenue, Tehran, Iran  
(z\_lari@sina.kntu.ac.ir), (ebadi@kntu.ac.ir)

**KEY WORDS:** building extraction, artificial neural networks, learning, C# programming environment

## ABSTRACT:

Automatic building extraction remains an open research area in digital photogrammetry. While many algorithms have been proposed for building extraction, none of them solve the problem completely. This paper proposes a system for increasing the degree of automation in extraction of building features with different rooftops from high resolution Multispectral satellite images (e.g., IKONOS and Quickbird) in Middle East countries. Following on, the implementation and functionality of software developed on the basis of neural networks approach are also explained. As known, neural networks have capabilities as pattern recognition and object extraction for remotely sensed data. The software has been designed and developed in C# programming environment and it is rather user friendly due to the fact that little knowledge is required for the users about neural networks theory. The proposed system works in two different phases: the first phase is learning, and the second phase is application. In the first phase, the presented neural network in the system is trained with the aid of test data, and in the second phase, the system will be used for detection and extraction of buildings from satellite images. Fused images from IKONOS sensor (1m resolution) from urban area of Kashan in Iran were selected for testing this system. The building extraction results are compared with manually delineated ones. The comparison illustrates the efficiency of the proposed algorithm in which it can extract approximately 80% of buildings in the image properly.

## 1. INTRODUCTION

Automatic building extraction has been an active research topic in the field of photogrammetry and computer vision for many years. Some useful applications are automation of information extraction from images and updating of geographic information system (GIS) databases. A wide range of techniques and algorithms have been proposed for automatically constructing 2D or 3D building models from satellite and aerial imagery. Considering both radiometry and geometry, a large population of these algorithms are edge-based techniques[1,2,3] that consist of linear feature detection, grouping for parallelogram structure extraction, and building polygons verification using knowledge such as geometric structure [2,3], shadow [2], and so forth. In order to solve this complex problem, integrating the power of multiple algorithms, cues, and available data sources is needed to improve the reliability and robustness of the extraction results[4,5].

The recent availability of commercial high-resolution satellite imaging sensors such as IKONOS provides a new data source for building extraction. The high spatial resolution of the imagery specifies very fine details in urban areas and facilitates the classification and extraction of urban-related features such as roads and buildings. Since manual extraction of buildings from imagery is very slow, automated methods have been proposed to improve the speed and utility for urban map's production.

Most of the recent work on building extraction from high-resolution satellite images is based on supervised techniques. These techniques either require a classification based on initial training data to provide hypotheses for the positions and sizes of the candidate building features[6,7], or they use

training sets or model databases to classify or match the buildings[8,9].

In this paper, an automated system for extraction of buildings from high-resolution satellite imagery is proposed that utilizes structural and spectral information. Using Artificial Neural Networks algorithms, the detection percentage and quality of the building extraction are greatly improved.

## 2. SYSTEM SKETCH

The proposed automatic system consists of three main parts that each of them have specific tasks.

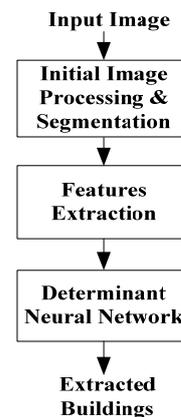


Figure1: System Sketch

In the first part, initial image processing and segmentation is done. In the second part, segment's features are extracted. And in third part, the system decides about possibility of each segment's being building based on features extracted

\* Corresponding author

using artificial neural networks. System sketch is shown in figure 1 briefly.

This system works in two phases:

1. **Learning Phase:** In this phase, the neural network presented in third part of system trains using manually saved data to reach desirable accuracy criterion.
2. **Application Phase:** In this phase, the systems tests on new dataset.

### 3. INITIAL IMAGE PROCESSING

In this part, all image processing stages including image pre-processing, image segmentation using seeded region growing algorithm and image post processing are described.

#### 3.1 Image Pre-processing

Selected images for testing this system are PAN and Multispectral images of IKONOS spacing 100km<sup>2</sup> from urban area of city Kashan in Iran. For utilization of spatial resolution and geometric accuracy of PAN and spectral capabilities of Multispectral georeferenced images, they are fused in ENVI4.0 remote sensing software package. Fused Image has capabilities of both Pan and Multispectral images.

#### 3.2 Image Segmentation Using Seeded Region Growing Algorithm

For the initial segmentation of the input image, a seeded region growing algorithm is used to find homogeneous roof regions in the image. The seed points are regularly distributed over the image with a seed point raster size set with respect to the expected roof size. For an IKONOS input image with resolution of 1m, an appropriate raster size is 10 pixel to ensure that nearly every roof region is hit. The input channel of this algorithm, is the intensity channel (I) which is calculated with following formula.

The seeded region growing algorithm starts at the pixel position of each seed point and compares this pixel's value with the neighbouring pixel values. If the neighbour pixel values lie inside a given tolerance  $t_T$ , the neighbouring pixels belong to the same region as the seed point. The region growing goes on recursively with the newly added pixels and ends, when no new neighbouring pixels which fulfil the condition can be found. An example of an image segmentation made with the described procedure is shown in Figure. 2.B.

belong to the same region as the seed point. The region growing goes on recursively with the newly added pixels and ends, when no new neighbouring pixels which fulfil the condition can be found. An example of an image segmentation made with the described procedure is shown in channel of this algorithm, is the intensity channel (I) which is calculated with following formula.

The seeded region growing algorithm starts at the pixel position of each seed point and compares this pixel's value with the neighbouring pixel values. If the neighbour pixel values lie inside a given tolerance  $t_T$ , the neighbouring pixels belong to the same region as the seed point. The region growing goes on recursively with the newly added pixels and

ends, when no new neighbouring pixels which fulfil the condition can be found. An example of an image segmentation made with the described procedure is shown in Figure. 2.B.

#### 3.3 Image Post Processing

The goal of the post processing stage is to improve the image segmentation procedure for the following feature extraction and classification. The post processing consists of functioning opening and closing operators on the segmented image. By use of opening and closing operators, the impure edges of the regions are smoothed and small holes in the regions are closed. In Figure 2.C, the result of opening and closing operations is shown.

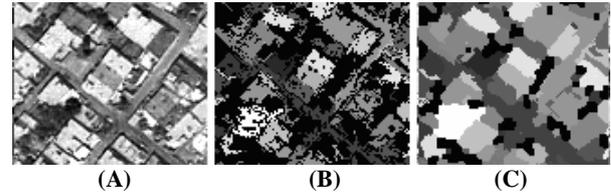


Figure 2.:results of initial processing of image,(A)input image,(B)segmented image, (C)result of opening and closing operators functioning on segmented image

### 4. FEATURES EXTRACTION

In this section, we describe extraction procedure of features which are basis of image classification and building detection. These features consist of: area, perimeter, mean colour and intensity, roundness, compactness and structural features of each region.

#### 4.1 Geometric Features

This subsection describes used geometric features and their calculation method. These features include area and perimeter.

- **Area:** the area of each region is calculated by counting the number of its pixels. If the calculated value is more than 10000 pixels the hypothesis of being building for that region will be rejected.
- **Perimeter:** the perimeter of each region is calculated by counting its boundary pixels.

#### 4.2 Structural Features

Structural features of each region are very important because the structure can differentiate buildings from other objects. These features consist of: roundness, compactness, lengthiness and specific angles of each region.

- **Roundness:** This feature is independent of region's size and calculated as ratio of area to square of perimeter. It varies from 0 to 1.

$$Roundness = \frac{4\pi \cdot Area}{Perimeter^2} \quad (1)$$

- **Compactness:** The compactness of each region is defined as number of opening and closing functions repetitions to remove each region completely.
- **Specific angles:** In [10] different types of region axes are introduced. The main axis is defined as line between two region contour points that have maximum distance among each other. The cross axis is defined as vertical line to the main axis that connects two contour points with the maximum distance to each other. The two ancillary axes are defined as vertical lines to the main axis with the maximum distance from the contour to the main axis. For each side of the main axis one ancillary axis is defined. The axes calculation results in six points which lie on the contour of the investigated region. The corresponding hexagon approximates the region's shape. An example of such a hexagon is shown in Figure 3.

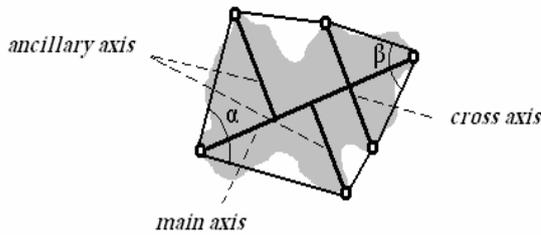


Figure 3: Approximating hexagon model

#### 4.3 Photometric features

- **Mean color and intensity:** For considering spectral features of objects, Mean intensity and color (R,G,B) of all pixels in each region are calculated of photometric features of that region.

### 5. NEURAL NETWORK AND LEARNING ALGORITHM

The neural network used in this system is a tree-layer perceptron. A specific weight is determined for each of input values. This network will function well if all weight coefficients select truly. In training process of neural network, this coefficient will be modified.

In training process of multi-layer neural network, at first weight coefficients of networks are selected randomly. Then a pattern is presented to network and its output will be calculated. Comparison of real outputs and desired outputs makes weight coefficients to be modified until we reach valid outputs. This process stops when network's output for specific number of inputs is correct. This kind of learning is called as supervised learning.

To acquire input data for training neural network, input image is divided to smaller parts. Then, some parts are selected randomly and initial image processing algorithms will be applied to segment them. After image segmentations, class type of each region will be determined manually and saved in a database.

To train the neural network used in this system, different saved regions are chosen randomly and numeric features of them which calculated in section 4 are entered in networks as inputs, network's output is compared with desired output and network's weight coefficients are modified using back-

propagation learning algorithm. This process repeats until network's output could recognize the class type of each region with desirable accuracy. After fulfilling this process, this neural network is trained and ready to work.

### 6. RESULTS EVALUATION

After training the neural network used in this system, we must evaluate the degree of its success in building recognition. So parts of image which haven't used in training phase are presented to system. Then, the output of system will be compared with real results and the accuracy of system will be calculated.

Each part of image after initial processing stage is segmented and features of each region (segment) will be calculated distinctly and set as neural network's input. If numeric value of neural network's output for a region is more than a threshold, it will be recorded as a building by system and its location will be shown by blue color on image. Figure 4 shows a sample of system output.

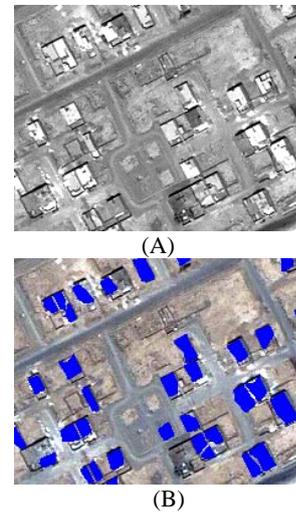


Figure 4. a sample of system output, (A) a small part of satellite image, (B) system output

For evaluation of this system, we use parts of image with size of 1000\*1000 pixels which cover an area about 1km<sup>2</sup>. for comparison of output result and the results which saved manually and evaluation of system accuracy, two measurements which are calculated so are used:

$$DP = \frac{100.TP}{TP + TN} \quad (2)$$

$$BF = \frac{100.FP}{TP + FP} \quad (3)$$

Where TP (true positive) is the number of buildings detected by both a person and the automatic approach, FP (false positive) is the number of buildings detected by the automatic approach but not a person, and TN (true negative) is the number of buildings detected by a person but not by the automatic approach. The detection percentage (DP) determines how many of the existing buildings in the image are found by the automatic approach and the branch factor (BF) shows how many buildings are detected wrongly. For

comparison, the results which acquired from four distinct parts of image are shown in table 1.

Image segment NO	Area type	DP	BF
1	Urban-old	81.6%	19.1%
2	Urban-new	84%	17.8%
3	Suburb	91.7%	20.6%
4	Industrial	80.4%	21.3%

Table 1. Acquired results from 4 distinct part of image

## 7. CONCLUSION

In this paper, a system is proposed for automatic building detection from satellite images based on neural networks algorithms. In this system initial image processing stage is implemented at first and image segmentation procedure is done. Then suggested features are calculated for each region and a three layer perceptron neural network is trained for detection of buildings in satellite images. After these processes the system is evaluated using DP and BF parameters. The mean value of these calculated parameters shows that has the capability of building detection in desirable level.

A. K. Shackelford and C. H. Davis, 2003, A combined fuzzy pixelbased and object-based approach for classification of high resolution multispectral data over urban areas, IEEE Trans.Geosci. Remote Sensing, vol. 41, no. 10, pp. 2354–2363.

Gamba, O. Hellwich, and P. Lombardo, 2003, Theme issue on algorithms and techniques for multi-source data fusion in urban areas, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, no. 1-2, pp. 1–3.

H.Mayer, 1999, Automatic object extraction from aerial imagery—a survey focusing on buildings, Computer Vision and Image Understanding, vol. 74, no. 2, pp. 138–149.

D. S. Lee, J. Shan, and J. S. Bethel, 2003, Class-guided building extraction from Ikonos imagery, Photogrammetric Engineering and Remote Sensing, vol. 69, no. 2, pp. 143–150.

C. Jaynes, F. Stolle, and R. Collins, 1994, Task driven perceptual organization for extraction of rooftop polygons, in Proc. IEEE Workshop on Applications of Computer Vision (WACV '94), pp.152–159, Sarasota, Fla, USA.

J. A. Benediktsson, M. Pesaresi, and K. Arnason, 2003, Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, IEEE Trans. Geosci. Remote Sensing, vol. 41, no. 9, pp. 1940–1949.

C. Lin and R. Nevatia, 1998, Building detection and description from a single intensity image, Computer Vision and Image Understanding, vol. 72, no. 2, pp. 101–121.

## REFERENCES

Z. Kim and R. Nevatia, 1999, Uncertain reasoning and learning for feature grouping, Computer Vision and Image Understanding, vol. 76, no. 3, pp. 278–288.

K. Segl and H. Kaufmann, 2001, “Detection of small objects from high-resolution panchromatic satellite imagery based on supervised image segmentation,” IEEE Trans. Geosci. Remote Sensing, vol. 39, no. 9, pp. 2080–2083.

Baxes, G. A., 1994, Digital image processing: principles and applications. John Wiley & Sons, Inc., New York, NY, USA.

S. M`uller and D. W. Zaum, 2005, Robust building detection from aerial images, IAPRS, Vol. XXXVI, Part 3/W24, Vienna, Austria.