International Journal of Applied Earth Observation and Geoinformation 30 (2014) 158–166

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Hierarchical segmentation of urban satellite imagery

Bardia Yousefi [a,*], Seyed Mostafa Mirhassani [a], Alireza AhmadiFard [a], MohammadMehdi Hosseini [b]

[a] Department of Electrical & Robotic Engineering, Shahrood University of Technology, Shahrood, Iran
[b] Electrical Engineering Department, Islamic Azad University, Shahrood Branch, Shahrood, Iran

ARTICLE INFO

Article history:
Received 4 September 2013
Accepted 22 January 2014
Available online 7 March 2014

Keywords:
Very high resolution satellite imagery
Gabor wavelet
Bayesian classifier
Relaxation labeling

ABSTRACT

This paper proposes a method to combine contextual, structural, and spectral information for classification. The method is an integrated method for automatically classifying urban-area objects in very high-resolution satellite imagery. The approach addresses three aspects. First, the Gabor wavelet is applied to the image along with morphological operations, with the sparsity of the outcome considered. A Bayesian classifier then categorizes the different classes, such as buildings, roads, open areas, and shadows. There are some false positives (wrong classification), and false negatives (non-classification) in the initial results. These results can be corrected by the relaxation labeling categorization of the unknown regions. The novelty of the proposed approach lies in the extensive use of spatiotemporal features considering the sparsity of urban objects. The results indicate improvement in classification through relaxation labeling compared with existing methods.

1. Introduction

Remote sensing imagery allows the monitoring of the surface and atmosphere of the earth on different scales. The development of imagery sensors has increased the availability of high-quality remote sensing images. Such developments in remote sensing imagery and the increased availability of commercial and high-resolution imagery have made the classification of urban areas a well-researched topic in the field of computer vision and image processing. Classification is a requirement of various applications, such as urban development planning and growth (see Kaya and Curran, 2006, the monitoring of urban subsidence (Zhang et al., 2011)), emergency response(see Van der Sande et al., 2003), and earth survey (e.g. Canadian Urban Land-Use Survey in Guindon and Zhang, 2007). The classification of remote sensing images of urban areas provides information and means useful for surveillance, earth survey, map updating and GIS (Ahmadi and Ebadi, 2009), urban planning, emergency response, management, and security applications. Therefore, automatic and semiautomatic methods to classify roads, buildings, and other land cover types in urban areas are of significant interest. This research examines the high-level classification and accurate segmentation of urban satellite images. Several studies have focused on the classification (Jin and Davis, 2005) and segmentation of urban areas and map generation. However, the misclassification problem remains. Misclassification includes false positives and false negatives. False positives occur when regions are incorrectly classified, whereas false negatives represent unclassified regions. Both problems are considered misclassifications and are thus the focal points of the map generation in this study. These problems can be solved by increasing the rate of classification and reducing the rate of misclassification.

1.1. Information combination

This study aims to consider a method to classify urban area images with the Gabor wavelet to extract features. A Bayesian rule classifier and relaxation labeling are used as post-processing techniques. Relaxation labeling is also used to reduce false positive and false negative results. Moreover, this study aims to increase accuracy by increasing the extraction rate of the components of urban area images. The proposed approach is a technique to classify very high-resolution satellite images of urban regions. Different aspects such as structural and contextual techniques and spectral methods are considered in classification. Structural techniques classify images by considering the structure of image components, that is, by using several special shapes or a special morphology of the components. Pesaresi and Benediktsson (2001) were the first to investigate morphological profiles and differential morphological profiles (Pesaresi and Benediktsson, 2001). Contextual methods follow the combinations and symmetric and systematic arrangements of components. Two main perpendicular directions of the road can be detected with a spatial signature-weighted Hough transform (SSWHT) (Jin and Davis, 2005). To complete classification, the image is rotated at an angle of less than 45.

* Corresponding author.
E-mail address: bardia.yousefi@ieee.org (B. Yousefi).

0303-2434/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2014.01.014

Consequently, the most important directions of roads and buildings are the horizontal and the vertical in the image space. Spectral techniques to classify satellite images depend on spectral consideration and an improved probabilistic neural network (Zhang et al., 2009). The accessibility of commercial high-resolution satellite imaging sensors, such as IKONOS and Quick-Bird, provides data sources for extracting building images. High spatial resolution images display significant details on urban areas and significantly facilitate the classification of urban-related features, such as roads (Benediktsson et al., 2003; Guindon, 2000; Araya and Cabral, 2010; Jin and Davis, 2005; Zhang et al., 2011), and buildings (Benediktsson et al., 2003; Hartfield et al., 2011; Yousefi et al., 2012; Matikainen and Karila, 2011; Melgani et al., 2000; Segl and Kaufmann, 2001; Dell'Acqua and Gamba, 2001; Shackelford and Davis, 2003; Sohn and Dowman, 2001; Fazan and Poz, 2013). Launched in September 1999, IKONOS is a major commercial high-resolution satellite. IKONOS obtains 1 m panchromatic and 4 m multispectral images. Important features, such as individual roads and buildings in the urban environment, can be classified because of the high geometric accuracy and spatial resolution of IKONOS imagery. IKONOS also supplies accurate geodetic coordinates. Van der Sande et al. (2003) presented an approach to assessing flood damage and risk through the segmentation of IKONOS-2 imagery to map land cover. Zhang et al. (2011) proposed segmentation using a smoothness constraint. This method develops progressive TIN densification (PTD) using point cloud segmentation.

Statistical methods have been used in classification and described through mathematical presentations. These methods have been combined with other techniques (Zhang et al., 2009). Several classification methods use statistical techniques. Pattern recognition methods are extensively used. Jin and Davis (2005) presented an integrated strategy to identify buildings in 1 m resolution satellite imagery of urban areas through structural, contextual, and spectral information. In this study, we use the multi-direction Gabor wavelet filter to preprocess data, the Bayesian classifier to classify the components of urban areas, and relaxation labeling to increase classification accuracy. The remainder of this article is organized as follows. Section 2 presents the proposed method, Section 3 discusses the results, and Section 4 concludes the paper.

2. Methodology

Urban satellite objects are identified and categorized as buildings, streets, roads, shadows, and so on. In this section, we present a technique to segment urban satellite images. The proposed approach has three parts: preprocessing, the core approach, and post-processing. Preprocessing is used to enhance image quality and adopt frequency components to improve image contrast and the visibility of urban area objects. A Gabor filter bank is used to amplify the amplitude and spatiotemporal features of images. Thus, for 2D remote sensing images, we develop 3D comprehension using the Gabor wavelet.

2.1. Gabor wavelet

The Gabor wavelet is used as a feature extractor. The technique highlights the components that require enhancement in spatiotemporal features. Different directions apply this filter to increase the quality of image components.

A Gabor wavelet dictionary comprising n directions and m scales is expressed as $GW_j(\theta, \omega)$, $j = 1, \ldots, m \times n$. Where, $\theta \in \{(k\pi/n), k = 0, \ldots, n-1\}$ and $\omega = \{(\sqrt{2}/i), i = 1, \ldots, m\}$. Gabor wavelet features capture the object form with small variance in size, location, and posture. The overall shape structure is maintained throughout the recognition process. The response or convolution of each element offers form information with θ and ω.

$$I_{gM}O = \langle GW, I \rangle = \sum \sum GW(x_0 - x, y_0 - y : \omega_0, \theta_0) I(x, y). \quad (1)$$

Let GW_j is a $[x_g, y_g]$, I is a $[x_i, y_i]$ matrix. The response of I to GW is a $[x_i + x_g, y_i + y_g]$. Therefore, previous convolutions in both matrices must be padded with sufficient zeros. Afterward, morphological operation is used on the target image. Different shapes of structural elements are used to extract distinctive image classes e.g. A stretched rectangular shape is used for roads, whereas a square rectangular shape is used for buildings.

2.2. Directed graph and factor graph

Using the Gabor wavelet as a phase-dependent feature extractor is a technique to extract edges. This technique is considered a novelty for the proposed approach. Fig. 1 shows the directed and factor graphs, which reveal the topology of the proposed technique. Using the multi-directional Gabor filter bank produces different effects and results to be fused by summation. Applying the Gabor wavelet to different directions and fusing various outcomes result in sparseness of output. A directed graph is used to apply the Gabor wavelet and fusion. This technique reveals the angle of rotation and the Gaussian envelope on the Gabor wavelet. The first layer of the directed graph represents the input layer. Using the Gabor wavelet at different angles at stable frequency and a constant location of the peak of the Gaussian envelope reduces the variety of parameters in the Gabor wavelet and prevents computational load in the system.

The various ways of applying the Gabor wavelet can be chosen arbitrarily, and the techniques can be modified depending on the application. The second layer is a Bayesian classifier that uses the output of Gabor wavelet outcomes from the first layer. Every Gaussian envelope has one outcome. These outputs are aggregated to create a final outcome and send the information for decision making. Fusion is done by summation of the different direction outcomes with respect to sparsity. In the morphological operation, structural elements of various sizes are used to determine the appropriate size for extracting building images from urban area images. Each urban area component in an image should have a structural element with a suitable size. The size of an urban area object is relevant to the size of the structural element. To locate a building in urban areas, images should have a structural element that is rectangular or near-quadrangular. The following steps are considered:

1. A structural element is used to enhance the image.
2. The structural element is rotated and resized.
3. Morphological operation is performed on the original image to visualize the target elements.

Finally, the Bayesian discrimination function is used to segment urban objects.

2.3. Bayesian rule and segmentation

The features extracted by the Gabor wavelet are fused by summation. The fused images are then classified into two categories: C_l, C_{nl}. These classes depend on the selected criterion. We consider our method output, which is represented by y and has a value of 0 or 1. A value of 0 indicates that the initial image is of a non-building, and 1 indicates buildings in the satellite image. The posterior probability function with regard to several events in one large set is

$$P(x_1, x_2, x_3, \ldots, x_n | y) = P(x_1|y)P(x_2|y, x_1)P(x_3|y, x_1, x_2)\ldots P(x_n|y, x_1, x_2, x_3, \ldots, x_{n-1}) \quad (2)$$