ISPRS Journal of Photogrammetry and Remote Sensing 126 (2017) 245–260

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Hierarchical graph-based segmentation for extracting road networks from high-resolution satellite images

Rasha Alshehhi *, Prashanth Reddy Marpu

Institute Center for Smart and Sustainable Systems, Masdar Institute of Science and Technology, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Article history:
Received 31 July 2016
Received in revised form 9 February 2017
Accepted 13 February 2017
Available online 8 March 2017

Keywords:
Road networks
Graph-based segmentation
Gabor
Morphological filtering
Hierarchical merging and splitting

ABSTRACT

Extraction of road networks in urban areas from remotely sensed imagery plays an important role in many urban applications (e.g. road navigation, geometric correction of urban remote sensing images, updating geographic information systems, etc.). It is normally difficult to accurately differentiate road from its background due to the complex geometry of the buildings and the acquisition geometry of the sensor. In this paper, we present a new method for extracting roads from high-resolution imagery based on hierarchical graph-based image segmentation. The proposed method consists of: 1. Extracting features (e.g., using Gabor and morphological filtering) to enhance the contrast between road and non-road pixels, 2. Graph-based segmentation consisting of (i) Constructing a graph representation of the image based on initial segmentation and (ii) Hierarchical merging and splitting of image segments based on color and shape features, and 3. Post-processing to remove irregularities in the extracted road segments. Experiments are conducted on three challenging datasets of high-resolution images to demonstrate the proposed method and compare with other similar approaches. The results demonstrate the validity and superior performance of the proposed method for road extraction in urban areas.

1. Introduction

Accurate and up-to-date road information is essential for many urban applications, such as automated road navigation (Li et al., 2014), geometric correction of urban remote sensing images (Auclair-Fortier et al., 2000) and updating geographic information systems (GIS) (Mena, 2003; Bonnefon et al., 2002). Rapidly changing urban environments need frequent updates of road database. Roads can be defined as long narrow regions with various orientations, lengths, and widths. Usually, the width of a road is several pixels in high-resolution satellite images. The length of a road is usually longer than buildings and longer than or equal to a street block (Liu et al., 2015; Sujatha and Selvathi, 2015). Extracting road network from up-to-date satellite images is a challenging problem due to noise and occlusions that create non-homogeneous regions leading to inaccurate classification of road segments. For instance, complex background and contextual structures (e.g., trees, shadow and vehicles on the roads) usually appear in high-resolution images. Also, some road-like segments, which have identical spectral/spatial properties such as railways and parking lots can be misclassified as roads.

Many approaches have been developed to extract roads from remotely sensed images for urban areas. Recently, there is a great interest in including spatial information e.g., morphological filtering (Liu et al., 2015; Sujatha and Selvathi, 2015; Valero et al., 2010; Maurya et al., 2011; Chaudhuri et al., 2012), shape and texture features (e.g., elongation, Gabor filtering, etc.) (Liu et al., 2015; Sujatha and Selvathi, 2015; Maurya et al., 2011; Zhou et al., 2010; Jin et al., 2012; Shi et al., 2014) along with various machine learning techniques (Maurya et al., 2011; Mokhtarzade and Zoej, 2007; Wegner et al., 2013; Mnih and Hinton, 2010).

Road extraction methodologies can be mainly classified based on two taxonomies. First, they can be divided into either road-area extraction or road-centerline extraction. Road-area extraction mainly depends on image classification and segmentation (Mnih and Hinton, 2010; Unsalan and Sirmacek, 2012; Cheng et al., 2014; Peng et al., 2011; Li et al., 2014). Road-centerline extraction methods concentrate on detecting road-skeletons (Liu et al., 2015; Sujatha and Selvathi, 2015; Shi et al., 2014; Miao et al., 2013; Miao et al., 2014; Cao and Sun, 2014; Hu et al., 2014; Shi et al., 2014; Sironi et al., 2014; Cheng et al., 2015). Second, extraction methods can be semi-automatic or automatic. In semi-automatic

* Corresponding author.
E-mail addresses: ralshehhi@masdar.ac.ae (R. Alshehhi), pmarpu@masdar.ac.ae (P.R. Marpu).

http://dx.doi.org/10.1016/j.isprsjprs.2017.02.008
0924-2716/

approaches (Liu et al., 2015; Chaudhuri et al., 2012), some prior information such as user input (e.g., seed points) or prior geographical information are required. In automatic approaches, no such prior information is required (Maurya et al., 2011; Mokhtarzade and Zoej, 2007; Wegner et al., 2013; Mnih and Hinton, 2010; Huanga and Zhanga, 2009).

Chaudhuri et al. (2012) proposed a semi-automatic method to extract road-centerline by applying directional Mathematical Morphology (MM), which is one of the commonly used methods to extract oriented, thin and line-like features. In Liu et al. (2015), the authors integrated MM (e.g., path opening/closing) approach by adapting volunteered geographic information (VGI), captured in the OpenStreetMap (OSM) database along with shape features (e.g., elongation and compactness) as prior knowledge to extract main road-networks from satellite images. Unsalan and Sirmacek (2012) suggested a method, that utilizes edge pixels to detect road centers by kernel-based density estimation. This process is followed by extracting road-shape segments based on a binary balloon algorithm and tracking road-segments based on a graph theoretic approach. Miao et al. (2013) proposed an automatic method, in which potential road segments are first obtained based on shape and spectral features, followed by multivariate adaptive regression splines modeling to extract road centerline. Hu et al. (2014) proposed a method where adaptive mean shift algorithm is first used to define road-center points. It is followed by voting to enhance salient linear features and then Hough transform to extract arc regions of the road-centerline. Most of these mentioned approaches integrated spectral and spatial features. However, selection of the most appropriate spatial features that can be used to extract different topologies (e.g., shape, orientation and scale) of roads in different urban images is a challenging problem. Some methods also consider contextual feature descriptors based on neighborhood, but the window-size of feature descriptors is defined subjectively.

Mokhtarzade and Zoej (2007) extracted roads from satellite images using artificial Neural Network (NN) classification based on texture features derived using Gray Level Co-occurrence Matrix (GLCM)(e.g., contrast, energy, entropy, and homogeneity). Huanga and Zhanga (2009) developed a methodology to extract road centerlines based on geometrical features of road-spectral variations by using Support Vector Machine (SVM) classifier. These methods are pixel-based methods where the complete contextual structure is not considered. Recently, probabilistic graphical models such as Markov Random Fields (MRF) (Perciano et al., 2011) and Conditional Random Fields (CRFs) (Wegner et al., 2013) are widely used for extracting roads based on contextual information. Furthermore, deep Convolution Neural Networks (CNNs) (Mnih, 2013; Shu, 2014; Saito and Aoki, 2015; Saito et al., 2016) have been achieving impressive state-of-the-art performance for extracting roads. In Mnih (2013), a supervised deep Neural Network (NN) architecture was proposed to extract roads more accurately from noisy data. Saito and Aoki (2015) and Saito et al. (2016) proposed a model averaging methods with various CNN parameters to produce more accurate roads. Although, the accuracy of these approaches is high, they need a large training database to have good road/non-road classification. Moreover, these approaches produce results where they have many discontinuous road-regions affecting the completeness of the extracted road network.

Urban classes (e.g., roads, building, parking lots, trees, etc.) typically appear at different scales in an image. The extraction method needs to take into account the scale to detect road and eliminate non-road pixels. Therefore, it is important to work with a multi-scale approach.

In this paper, we propose a graph-based segmentation to extract roads in urban areas. An input image is first segmented into superpixels to represent homogeneous regions. In order to eliminate the background, multi-channel Gabor and morphological filters are used to derive features to produce the initial road-network. However, road-superpixels of the initial road-network maybe be discontinuous and can be affected by over and under-segmentation. Therefore, hierarchical merging and then splitting are carried out to improve the results. Merging is based on grouping smaller segments to produce an approximate road-network. Splitting is the process to eliminate non-road superpixels from the extracted road-network. The criteria for merging and splitting are based on similarity of spectral and spatial features of superpixels. A graph-theoretic approach is then used to model contextual structure between adjacent segments to produce the complete road-network. The main contributions of the proposed method are summarized as follows:

- Many road extraction methods are based on the pixel-based classification which produces heterogeneous results. Pixel-based approaches cannot discriminate between contextual structures. Therefore, we suggest using superpixels to incorporate spectral and spatial information based on contextual structure.
- We introduce a new hierarchical graph-based approach to extract road-network more efficiently and accurately based on two main steps: merging and splitting. Hierarchical merging is used to group possible road superpixels to produce an intermediate result based on contextual, spectral and spatial features. Splitting is used to eliminate artifacts from the intermediate output.
- Many recent methods fail to differentiate road from other road-like regions (e.g., parking lot, railway, narrow shadow). We integrate texture features to discriminate road-regions from other regions which have identical shape and spectral properties.
- To overcome the weakness of the existing methods while dealing with intersections and other discontinuous regions due to complex backgrounds, the shortest path algorithm is proposed to complete these unconnected road-regions.
- A simple regularization method is suggested to smooth the segmentation results.

The remaining of this paper is organized as follows: we first describe the proposed method in Section 2, followed by a description of the experiments and the discussion on the final road-network including a comparison with other methods, Section 3. Finally, conclusions and future works are drawn in Section 4.

2. Methodology

The flowchart of the proposed method is depicted in Fig. 1. There are three main stages of the proposed method: (A) pre-processing, (B) graph-based segmentation and (C) post-processing. Pre-processing consists of two major steps to extract the most informative features based on: (i) Gabor filtering and (ii) morphological filtering. Graph-based segmentation is based on hierarchical merging or splitting. Post-processing consists of two steps: linking segments and removing small artifacts to obtain a complete road-network map.

2.1. Pre-processing

Pre-processing is used to filter the effect of background variations and produce more effective feature image which shows a high contrast between road and non-road pixels to facilitate the segmentation process. Here, we consider two well known filtering methods: (i) Gabor filtering to discriminate road texture from non-road texture and (ii) Morphological filtering to eliminate the background. Fig. 2 summarizes the pre-processing stage.