ISPRS Journal of Photogrammetry and Remote Sensing 117 (2016) 11–28

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Review Article

A survey on object detection in optical remote sensing images

Gong Cheng, Junwei Han *

School of Automation, Northwestern Polytechnical University, Xi'an 710072, China

ARTICLE INFO

Article history:
Received 10 November 2015
Received in revised form 5 March 2016
Accepted 16 March 2016
Available online 1 April 2016

Keywords:
Object detection
Optical remote sensing images
Template matching
Object-based image analysis (OBIA)
Machine learning
Deep learning
Weakly supervised learning

ABSTRACT

Object detection in optical remote sensing images, being a fundamental but challenging problem in the field of aerial and satellite image analysis, plays an important role for a wide range of applications and is receiving significant attention in recent years. While enormous methods exist, a deep review of the literature concerning generic object detection is still lacking. This paper aims to provide a review of the recent progress in this field. Different from several previously published surveys that focus on a specific object class such as building and road, we concentrate on more generic object categories including, but are not limited to, road, building, tree, vehicle, ship, airport, urban-area. Covering about 270 publications we survey (1) template matching-based object detection methods, (2) knowledge-based object detection methods, (3) object-based image analysis (OBIA)-based object detection methods, (4) machine learning-based object detection methods, and (5) five publicly available datasets and three standard evaluation metrics. We also discuss the challenges of current studies and propose two promising research directions, namely deep learning-based feature representation and weakly supervised learning-based geospatial object detection. It is our hope that this survey will be beneficial for the researchers to have better understanding of this research field.

1. Introduction

Object detection in optical remote sensing images (RSIs) is to determine if a given aerial or satellite image contains one or more objects belonging to the class of interest and locate the position of each predicted object in the image. The term 'object' used in this survey refers to its generalized form, including man-made objects (e.g. vehicles, ships, buildings, etc.) that have sharp boundaries and are independent of background environment, as well as landscape objects, such as land-use/land-cover (LULC) parcels that have vague boundaries and are parts of background environment. As a fundamental problem in the field of aerial and satellite image analysis, object detection in optical RSIs plays an important role for a wide range of applications, such as environmental monitoring, geological hazard detection, LULC mapping, geographic information system (GIS) update, precision agriculture, and urban planning.

Object detection in optical RSIs often suffers from several increasing challenges including the large variations in the visual appearance of objects caused by viewpoint variation, occlusion, background clutter, illumination, shadow, etc., the explosive growth of RSIs in quantity and quality, and the various requirements of new application areas. To address these challenges, the topic of geospatial object detection has been extensively studied since the 1980s. The low spatial resolution of earlier satellite images (such as Landsat) would not allow the detection of separate man-made or natural objects. Therefore, researchers mostly focused on extracting the region properties from these images. With the advances of remote sensing technology, the very high resolution (VHR) satellite (e.g. IKONOS, SPOT-5, and Quickbird) and aerial images have been providing us more detailed spatial and textural information. Aside from region properties, a greater range of man-made objects become recognizable and even can be separately identified than ever before because of the increased submeter resolution. This opens new prospects in the field of automatic detection of geospatial objects.

During the last decades, considerable efforts have been made to develop various methods for the detection of different types of objects in satellite and aerial images, such as roads (Barsi and Heipke, 2003; Barzohar and Coope, 1996; Chaudhuri et al., 2012; Das et al., 2011; Hu et al., 2007; Huang and Zhang, 2009; Kim et al., 2004; Laptev et al., 2000; Leninisha and Vani, 2015; Li et al., 2010; Maillard and Cavayas, 1989; Mayer et al., 2006; McKeown and Denlinger, 1988; Mena, 2003; Mokhtarzade and Zoej, 2007; Movaghati et al., 2010; Song and Civco, 2004; Trinder and Wang, 1998; Ünsalan and Sirmacek, 2012; Wang et al.,

* Corresponding author.
E-mail address: Junweihan2010@gmail.com (J. Han).

http://dx.doi.org/10.1016/j.isprsjprs.2016.03.014
0924-2716/

2015; Wang and Zhang, 2011; Zhang et al., 2011a; Zhang and Couloigner, 2006; Zhou et al., 2006; Zhu et al., 2005), buildings (Ahmadi et al., 2010; Akçay and Aksoy, 2010; Aytekın et al., 2012; Benedek et al., 2012; Durieux et al., 2008; Hofmann et al., 2002; Karantzalos and Paragios, 2009; Lefèvre et al., 2007; Lhomme et al., 2009; Mayer, 1999; Ok, 2013; Ok et al., 2013; Peng and Liu, 2005; Peng et al., 2005; Senaras et al., 2013; Shufelt, 1999; Sirmaçek and Ünsalan, 2009; Sirmacek and Ünsalan, 2011; Stankov and He, 2013, 2014; Wegner et al., 2011a, 2011b), trees (Haala and Brenner, 1999; Hung et al., 2012; Malek et al., 2014; Moustakidis et al., 2012; Yang et al., 2013), and vehicles (Eikvil et al., 2009; Grabner et al., 2008; Jin and Davis, 2007; Kembhavi et al., 2011; Leitloff et al., 2010; Moon et al., 2002; Niu, 2006; Tuermer et al., 2013; Wen et al., 2015; Yang et al., 2013, 2015; Zhao and Nevatia, 2003; Zheng et al., 2013). While enormous methods exist, a deep review of the literature concerning generic object detection is still lacking. This paper aims to provide a review of the recent progress in this field. Different from several previously published surveys that focus on a specific object class such as building (Mayer, 1999; Shufelt, 1999) and road (Mayer et al., 2006; Mena, 2003), we concentrate on more generic object categories including, but are not limited to, road, building, tree, vehicle, ship, airport, urban-area. Covering about 270 publications we survey (1) template matching-based object detection methods, (2) knowledge-based object detection methods, (3) object-based image analysis (OBIA)-based object detection methods, (4) machine learning-based object detection methods, and (5) five publicly available datasets and three standard evaluation metrics for object detection. We also discuss open problems and challenges of current studies, and propose two promising research directions in future for constructing more effective object detection framework. This survey will be especially beneficial for the researchers to have better understanding of this research field. Furthermore, to the best of our knowledge, this is the first survey paper in the literature that focuses on generic object detection in optical RSIs.

The rest of the paper is organized as follows. Section 2 briefly introduces the taxonomy of methods for object detection. In Sections 3–6, we exhaustively review template matching-based object detection methods, knowledge-based object detection methods, OBIA-based object detection methods, and machine learning-based object detection methods, respectively. In Section 7, we review five publicly available datasets and three standard evaluation metrics. Section 8 discusses the challenges of current studies and proposes two promising research directions to advance the field. Finally, conclusions are drawn in Section 9.

2. Taxonomy of methods for object detection

In the last decades, a large number of methods have been developed for object detection from aerial and satellite images. We can generally divide them into four main categories: template matching-based methods, knowledge-based methods, OBIA-based methods, and machine learning-based methods. These four categories are not necessarily independent and sometimes the same method exists with different categories. Fig. 1 shows a taxonomy of geospatial object detection studies, in which rounded rectangles with solid borders illustrate our scope in this paper.

According to the template type selected by a user, template matching-based methods are further divided into two classes as rigid template matching and deformable template matching. As for knowledge-based object detection methods, we mainly review two kinds of most widely used prior knowledge, namely geometric information and context information. Generally, OBIA-based object detection methods involve two steps: image segmentation and object classification. With regards to machine learning-based methods, we mainly focus on reviewing three crucial steps that play important roles in the performance of object detection. They are feature extraction, optional feature fusion and dimension reduction, and classifier training. In the step of feature extraction, we introduce five types of typical features including Histogram

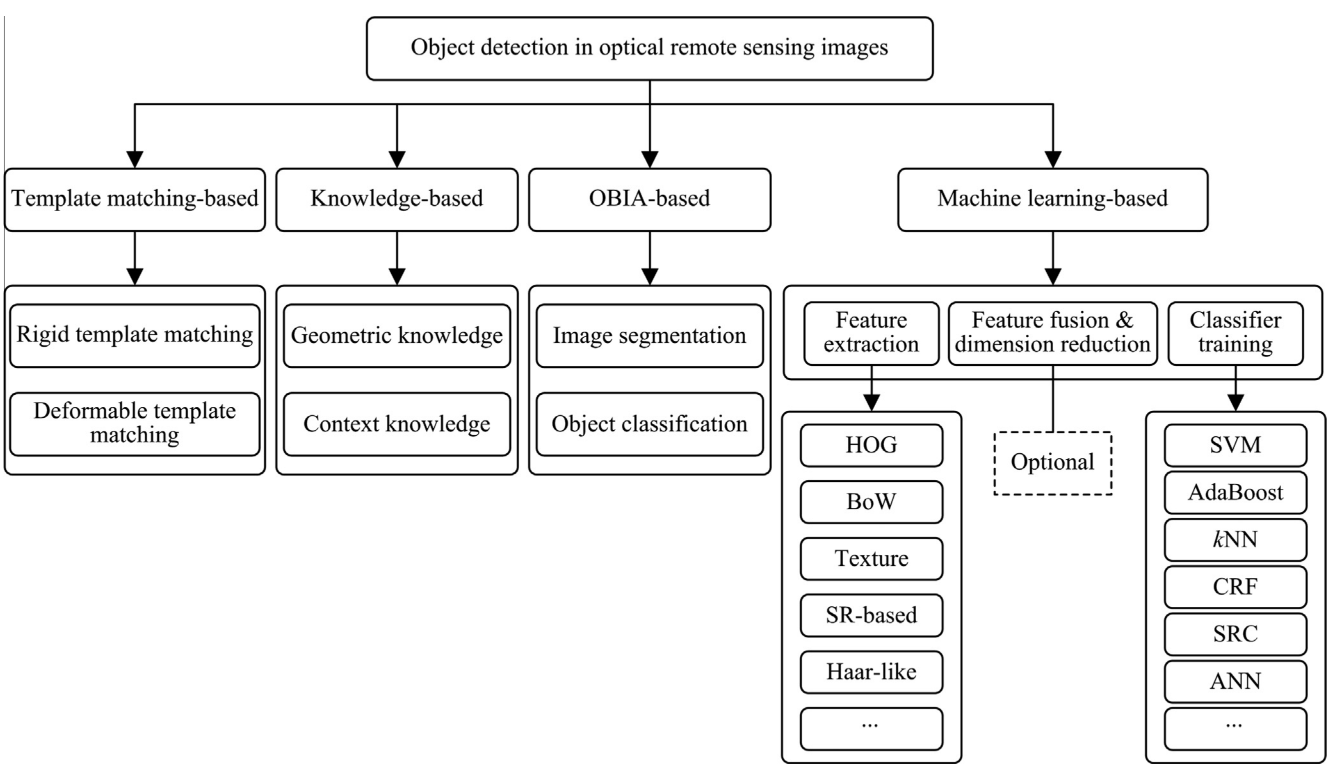

Fig. 1. Taxonomy of methods for object detection in optical RSIs. Rounded rectangles with solid borders illustrate our scope in this paper.