International Journal of Applied Earth Observation and Geoinformation 15 (2012) 57–69

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images

Juan P. Ardila*, Wietske Bijker, Valentyn A. Tolpekin, Alfred Stein

Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, 7500 AA Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 30 September 2010
Accepted 23 June 2011

Keywords:
GEOBIA
Object based image analysis
Tree mapping
Tree crown identification
Context classification
Object classification

ABSTRACT

Municipalities need accurate and updated inventories of urban vegetation in order to manage green resources and estimate their return on investment in urban forestry activities. Earlier studies have shown that semi-automatic tree detection using remote sensing is a challenging task. This study aims to develop a reproducible geographic object-based image analysis (GEOBIA) methodology to locate and delineate tree crowns in urban areas using high resolution imagery. We propose a GEOBIA approach that considers the spectral, spatial and contextual characteristics of tree objects in the urban space. The study presents classification rules that exploit object features at multiple segmentation scales modifying the labeling and shape of image-objects. The GEOBIA methodology was implemented on QuickBird images acquired over the cities of Enschede and Delft (The Netherlands), resulting in an identification rate of 70% and 82% respectively. False negative errors concentrated on small trees and false positive errors in private gardens. The quality of crown boundaries was acceptable, with an overall delineation error <0.24 outside of gardens and backyards.

1. Introduction

Multiple benefits from the presence of urban trees have been described extensively, mainly by forestry institutes and local authorities that are aware of the balance that should exist between urban infrastructure and green spaces (McPherson et al., 1994; Konijnendijk, 2005; McHale et al., 2007). These studies have stressed the relevance of monitoring the state of urban trees to quantify economic benefits and facilitate forestry interventions. Since detailed inventories of the constantly changing urban green ecosystem are costly and difficult to update with traditional field survey methods, alternative solutions have been sought (Ward and Johnson, 2007). Experience with remote sensing imagery in forest plantations, however, has indicated a number of factors constraining the semi-automatic identification of tree crowns. Such factors include: (1) the limited spatial resolution of satellite images with respect to the size of tree crowns, (2) the increase of within-crown spectral variance in very high resolution (VHR) imagery, and (3) the low spectral separability between tree crowns and other vegetated surfaces (Pouliot et al., 2002; Gougeon and Leckie, 2006; Hirschmugl et al., 2007). Specific characteristics of urban areas also hinder the semi-automatic image identification of trees. (1) In cities, trees spatially coexist with urban elements like buildings, roads, sidewalks and canals, which results in a complex arrangement of the image space. (2) There is a large variation in tree structural characteristics, such as height, crown shape, crown diameter and canopy cover. (3) Depending on planting practices, trees may be isolated, evenly spaced, in irregular spatial patterns or in groups of interlocked trees.

As these factors limit the applicability of spectral classifiers for tree identification, a promising solution is to address the complexity of the urban space by using image context. Context, defined as any information that can be used to characterize the situation of an entity, is an essential element for feature recognition (Abowd et al., 1999; Oliva and Torralba, 2007). In the identification of trees from digital images, context can foster better classification results by modeling conditions of the spatial distribution of trees with respect to other urban elements. As such, contextual rules can model the occurrence of trees along roads, in private gardens or in green areas, and depending on sun illumination, image shadows may be used to further improve tree identification.

Geographic object-based image analysis (GEOBIA) has been proposed as a method to bridge the gap between the increasing amount of detailed geospatial data and complex feature recognition problems (Blaschke, 2010). GEOBIA formulates the processing and analysis of homogeneous regions, referred to as *image-objects*, which interact and evolve during the classification process. Within a GEOBIA approach, context is modeled through the topologic relations of neighboring image-objects which are generated with a

* Corresponding author. Tel.: +31 53 487 4444.
E-mail addresses: ardila14710@itc.nl (J.P. Ardila), bijker@itc.nl (W. Bijker), tolpekin@itc.nl (V.A. Tolpekin), stein@itc.nl (A. Stein).

0303-2434/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jag.2011.06.005

segmentation technique. This is an advantage over pixel-based analysis, where context is limited to the local interaction of individual pixels within a window of an specific size (Blaschke, 2003).

In this work we consider VHR images those providing a spatial resolution better than one meter in the panchromatic mode (e.g., Geoeye, Worldview-2, and Quickbird). Determining aspects that make GEOBIA an attractive approach for the semiautomatic identification of urban tree crowns in VHR satellite images are: (1) There is not a unique scale for the analysis of geographic elements in remote sensing (Hay et al., 1997). In fact, a multiscale approach is needed for detection and analysis of vegetation (Marceau et al., 1994). In the context of this research, trees can stand alone, forming regular patterns, or aggregate in groups of interlocked trees. GEOBIA enables multiscale image analysis (Hay and Castilla, 2008). (2) With increasing resolution so does the within-class spectral variance, which ultimately results in a low classification accuracy of pixel based-classifiers (Woodcock and Strahler, 1987). An object-based approach deals more efficiently with the *high resolution problem* (Hay and Castilla, 2008). (3) Image-objects offer a wide range of features for image analysis not available when considering individual pixels (Blaschke, 2003). Texture, shape and contextual features are key to the identification of trees in urban scenes. (4) End-users of tree inventories require products that quantify the state of tree resources in terms of discrete units, e.g., individual trees or tree groups. The GEOBIA approach provides a direct link from image-objects to meaningful geographical objects (Hay and Castilla, 2008).

This paper proposes: (1) to develop a generic and reproducible set of contextual GEOBIA methods for the identification of urban tree crown objects in The Netherlands using VHR imagery and (2) to assess the accuracy and suitability of this approach for tree identification. We implement the GEOBIA methods on a pair of QuickBird (QB) scenes captured over the cities of Enschede and Delft and address the quality of identification of individual trees and tree groups with object-based accuracy indicators.

2. Test sites and data

The Netherlands show significant urban forestry activities in their more than 300 urban settlements. As horticulture and arboriculture are important economic activities, there is a great variety of tree species, including many ornamental trees. Planting and maintenance of trees is controlled by municipalities and land-owners in public and private areas respectively. In this research we selected two test areas, namely, the Bothoven district in Enschede, and the downtown area in Delft (Fig. 1). These sites contain several tree species with a variation of distance between trees, crown size and surrounding elements, resulting in challenging and realistic benchmarking cases for tree detection. In the study areas, trees are mainly deciduous (>80%), with trees planted along the roads, water channels, in grassland areas, public spaces and in private gardens. Depending on landscape design and urban space, trees exist in isolation, in pairs, forming linear patterns or groups of interlocked trees. Crown diameter ranges from 1 to 25 m.

For the study panchromatic and multispectral QB images were acquired in leaf-on conditions with acquisition dates and image characteristics as specified in Table 1. The images were obtained in standard processing level, where a radiometric and geometric correction is applied by the image provider. GEOBIA methodology was implemented on the digital values (DN) of the pan-sharpened images obtained after enhancing the spatial resolution of the original QB images using a high-pass filter (HPF) algorithm (Chavez et al., 1991). The HPF algorithm used the panchromatic and multispectral information to produce a four band image with 0.6 m resolution and similar spectral characteristics. For information on urban infrastructure and landuse the digital topographic map of the Netherlands Top10NL (scale 1:10,000) was used. A polygon reference layer of tree crowns was constructed as described in Section 4.2. Fig. 1 presents the pan-sharpened image windows of the test areas together with the reference tree crown layer.

GEOBIA was implemented in eCognition Developer 8 software (Definiens, 2010), using the multiresolution segmentation algorithm to generate image-objects (Baatz and Schäpe, 2000) with modifications to address several object scales as explained below. Object features used in this paper are defined in the application user guide (Definiens, 2009). Training areas were selected in subsets of the QB images with no overlap in respect to test areas and selected on the basis of specific image context for the developed rules. The following section presents the general method formulated for identification of tree crowns and Section 4 presents its specific implementation in the test sites.

3. GEOBIA approach for tree identification

Due to the complexity of the urban scenes and of several factors limiting digital tree detection, we divide the tree crown GEOBIA methodology into specific classification strategies depending on image characteristics and the local context of the trees on the ground. Based on a visual analysis of the scenes we propose a sequence of steps to: (1) identify grassland areas; (2) generate tree crown *object primitives* using normalized difference vegetation index (NDVI) local contrast measurements; (3) detect isolated trees which present a high contrast to the background; (4) identify trees and tree groups in a multiscale approach; (5) identify very small trees alongside the roads using local maxima and shadow information; (6) separate adjoining trees using morphological operations and (7) identify remaining trees using a local maxima and region growing method. Lastly, we integrate the results of each GEOBIA step and report on the individual trees and groups of tree-objects. Fig. 2 presents a diagram of the proposed workflow described in this section.

3.1. Masking of grassland areas

Tree crown identification is largely affected by the low spectral separability of tree crown pixels with respect to other elements of the image. In particular, vegetated surfaces such as grasslands and shrubs hinder the correct identification and delineation of tree crowns (Bunting and Lucas, 2006). Hence, a first step is to identify grassland areas and exclude them prior to tree detection. We propose to base this identification on the spectral properties and texture values of image-objects. Image texture, considered as a measure of the spatial variation of image tone or intensity (Haralick, 1979), is useful to identify grassland areas since the spectral variance of grassland objects is smaller than that of tree crowns in VHR imagery. Moreover, since grassland areas occur at different scales, we formulate a multiscale segmentation approach where objects are iteratively generated with varying segmentation scale parameter p_{seg} which decreases during each iteration step by multiplication with an update factor $p_{upd} \in (0, 1)$. Thus, the size of image-objects, depending on p_{seg}, decreases with each segmentation step as long as p_{seg} is higher than a threshold p_{min}. This procedure is presented in Algorithm 1 in Appendix A.

In this procedure, for each set of objects generated with p_{seg}, unclassified objects are assigned to *grass* if their near-infrared (NIR) standard deviation is below a threshold (std_{thr}), or remain unclassified otherwise. In addition, we define a class (*grass_neigh*) to account for the edge texture effect of grassland regions that occurs when there is a strong contrast between the image texture of adjacent landcovers. This happens for instance, in transition areas between grasslands and tree crowns (Ferro and Warner, 2002).