ISPRS Journal of Photogrammetry and Remote Sensing 125 (2017) 225–236

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Informal settlement classification using point-cloud and image-based features from UAV data

C.M. Gevaert [a,*], C. Persello [a], R. Sliuzas [b], G. Vosselman [a]

[a] Dept. of Earth Observation Science, Faculty ITC, University of Twente, Enschede, The Netherlands
[b] Dept. of Urban and Regional Planning and Geo-Information Management, Faculty ITC, University of Twente, Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 1 July 2016
Received in revised form 6 October 2016
Accepted 24 January 2017

Keywords:
Informal settlements
Image classification
Point cloud
Unmanned Aerial Vehicles (UAV)
Feature extraction
Support vector machine

ABSTRACT

Unmanned Aerial Vehicles (UAVs) are capable of providing very high resolution and up-to-date information to support informal settlement upgrading projects. In order to provide accurate basemaps, urban scene understanding through the identification and classification of buildings and terrain is imperative. However, common characteristics of informal settlements such as small, irregular buildings with heterogeneous roof material and large presence of clutter challenge state-of-the-art algorithms. Furthermore, it is of interest to analyse which fundamental attributes are suitable for describing these objects in different geographic locations. This work investigates how 2D radiometric and textural features, 2.5D topographic features, and 3D geometric features obtained from UAV imagery can be integrated to obtain a high classification accuracy in challenging classification problems for the analysis of informal settlements. UAV datasets from informal settlements in two different countries are compared in order to identify salient features for specific objects in heterogeneous urban environments. Findings show that the integration of 2D and 3D features leads to an overall accuracy of 91.6% and 95.2% respectively for informal settlements in Kigali, Rwanda and Maldonado, Uruguay.

1. Introduction

Informal settlements are a growing phenomenon in many developing countries and the effort to promote the standard of living in these areas will be a key challenge for the urban planners of many cities in the 21st century (Barry and Rüther, 2005). These settlements refer to urban areas which lack legal tenure (Kuffer et al., 2016), and are often characterized by dense housing and sub-standard living conditions. The term is closely related to the term 'slums', referring to settlements which may lack legal tenure, lack access to water or sanitation, suffer from overcrowding and/or are characterized by non-durable housing (UN-Habitat, 2012). In the present study we utilize the term informal settlement as it is more commonly used in the remote sensing community (Kuffer et al., 2016) and due to the possible negative connotations of the term 'slum' (Gilbert, 2007). The planning and execution of informal settlement upgrading projects with the purpose of ameliorating these conditions require up-to-date base maps which accurately describe the local situation (UN-Habitat, 2012). For example, the identification of buildings gives an indication of the population in the area, classifying terrain identifies footpaths for accessibility and utility planning or free space for the location of infrastructure. However, such basic information is often lacking at the outset of upgrading projects (Pugalis et al., 2014), thus hindering the amelioration of the impoverished conditions in these areas. To create such base maps, satellite imagery is a powerful source of information regarding the physical characteristics of an informal settlement (Taubenböck and Kraff, 2013). However, as slums are often characterized by high building densities, small irregular buildings, and narrow footpaths, the spatial resolution provided by sub-meter satellite imagery is usually not sufficient (Kuffer et al., 2014). Photogrammetric workflows can extract 2D orthomosaics, 2.5D Digital Surface Models (DSMs) and 3D point clouds from overlapping aerial imagery. Although this can be done from aerial or satellite imagery, UAVs have lower operational costs and allow for flexible and fast data acquisition (Nex and Remondino, 2014). This combination of flexible data acquisition and high spatial resolution of the acquired products motivate the use of UAVs to support urban planning in dense and dynamic areas such as informal settlements. Disadvantages of the use of UAVs include

* Corresponding author.
E-mail addresses: c.m.gevaert@utwente.nl (C.M. Gevaert), c.persello@utwente.nl (C. Persello), r.sliuzas@utwente.nl (R. Sliuzas), george.vosselman@utwente.nl (G. Vosselman).

http://dx.doi.org/10.1016/j.isprsjprs.2017.01.017
0924-2716/

the limited spatial extent of UAV flights and the data processing requirements. Therefore, we consider them to be more adequate at a (settlement upgrading) project level where more detailed spatial information is required, rather than e.g. at a city level for the distinction between informal vs. formal settlements. The remaining question is then how to optimally integrate the information contained in the orthomosaic, DSM and point cloud in order to accurately classify these complex areas.

A well-known problem of classifying urban areas is the high within-class variability and low between-class variability of spectral signatures of the relevant classes. Also, when using very high-resolution (VHR) imagery, the objects to be classified are generally larger than the pixel size, which is problematic for purely pixel-based classification strategies (Blaschke, 2010). The classification of sub-decimeter orthomosaics in informal settlements can be expected to face similar problems. In the remote sensing community, a common strategy to address this issue is to include spatial-contextual features in the classification problem in addition to the spectral image attributes. Spatial-contextual information can also be incorporated through Object Based Image Analysis (OBIA), which is also currently the most common strategy for the classification of slum areas (Kuffer et al., 2016). Such approaches depend on adequate segmentation parameters, which may be difficult to transfer between study areas (Hofmann et al., 2008) or even to represent different classes within the same study area (Myint et al., 2011). Alternatively, Bruzzone and Carlin (2006) adopt a multilevel strategy to incorporate contextual features by combining the radiometric characteristics at a pixel level with attributes of larger image segments and thus avoiding the need to define one set of optimal segmentation parameters. Their approach focusses on the spectral and spatial features at the different contextual levels, but could be extended to include texture features as these have proven to be an important supplement to spectral features in urban scene classification (e.g. Puissant et al., 2005; Tong et al., 2014).

Furthermore, the availability of 3D data are an important supplement to the orthomosaic as the inclusion of height information has been shown to greatly increase classification accuracy of urban scenes (Hartfield et al., 2011; Longbotham et al., 2012; Priestnall et al., 2000). Especially the extraction of a normalized DSM (nDSM), which gives the elevation of pixels above the terrain, is useful for identifying elevated objects in urban scenes (Weidner and Förstner, 1995) and distinguishing between low vegetation and high vegetation (Huang et al., 2008). A recent overview of building detection methods based on aerial imagery and LiDAR data indicates that state-of-the-art techniques which have access to both imagery and height information can identify large buildings with a very high correctness and completeness (Rottensteiner et al., 2014). However, these building detection algorithms face difficulties when the buildings are relatively small (i.e. less than 50 m^2), or when the height of the terrain is not uniform on all sides of the building due to sloped terrain. Unfortunately, informal settlements are often characterized by these challenging conditions, which emphasizes the need to investigate the synergies between 2D and 3D features to fully exploit the available UAV data and obtain a high classification accuracy.

Existing strategies regarding the combination of 2D and 3D features are often based on the integration of LiDAR with multispectral aerial imagery. Yan et al. (2015) cite a number of studies where nDSM data derived from LiDAR was combined with vegetation indices from multispectral imagery to classify urban scenes (e.g. Hartfield et al., 2011). Other methods make use of elevation images which directly project the 3D points onto a horizontal plane without taking into account interpolation techniques which are typically applied for DSM extraction. Processing this summarized information in 2D space rather than the original 3D space can decrease computing costs (Serna and Marcotegui, 2014). In another example, Weinmann et al. (2015) describe a generic framework for 3D point cloud analysis which includes spatial binning features or accumulation maps, which are similar to elevation images. They define a horizontal 2D grid and calculate: the number of points within each bin, maximum height difference and standard deviation of height difference within each cell. Serna and Marcotegui (2014) use elevation maps to define the: minimum elevation, maximum elevation, elevation difference, and number of points per bin as a basis for detecting, segmenting and classifying urban objects. However, this method assumes the ground is planar. Guo et al. (2011) combined geometrical LiDAR features and multispectral features from imagery to analyse which features were most relevant to classify an urban scene into: building, vegetation, artificial ground, and natural ground. They use elevation images to include the inclination angle and residuals of a local plane, but found that the maximum height difference between a LiDAR point and all other points within a specified radius was the most relevant feature.

There are two main limitations of the previous methods. Firstly, most methods explicitly or inherently assume the terrain to be planar. Attributes such as the maximum absolute elevation or height above the minimum point within a horizontal radius, which are often considered to be the most relevant features (Guo et al., 2011; Yan et al., 2015), will not serve to distinguish between buildings and terrain in a settlement located on a steep slope. Secondly, the methods generally focus on pixel-based features, or local neighbourhood features. However, Vosselman (2013) and Xu et al. (2014) indicate that segment-based point cloud features provide important supplementary information to pixel-based attributes. Similarly, Myint et al. (2011) found that 2D object-based attributes significantly improve the classification of urban scenes from VHR satellite imagery. Studies investigating the importance of features for urban scene classification should therefore consider segment-based features as well as point-based features.

The objective of this paper is to integrate the different information sources (i.e. UAV point cloud, DSM, and orthomosaic) and to analyse which 2D, 2.5D, and 3D feature sets are most useful for classifying informal settlements, a setting which challenges the boundaries of existing building detection algorithms. In an effort address the challenge of identifying salient features in various conditions, UAV datasets over informal settlements in two different countries are compared. Feature sets describing 2D radiometrical and textural features from the orthomosaic, 2.5D topographical features from the DSM, and 3D features from the point cloud are selected from literature. Both pixel- or point-based features and segment-based features are included. The suitability of the feature sets for classifying informal settlements are tested through their application to two classification problems. The classification is performed using Support Vector Machines (SVMs), which have been shown to be very effective in solving nonlinear classification problems using multiple heterogeneous features. The first classification problem identifies major objects in the scene (i.e. buildings, vegetation, terrain, structures and clutter), whereas the second attempts to describe semantic attributes of these objects such as roof material, types of terrain, and specific structures such as lamp posts and walls. The results presented here are an extension of previous research regarding the suitability of various features sets for the classification of an informal settlement in Kigali, Rwanda (Gevaert et al., 2016) in two significant ways. Firstly, the suitability of the feature sets in a different setting is analysed through the application of the same framework to an informal settlement in Maldonado, Uruguay. Secondly, we provide an extensive analysis of the most suitable features per class, which supports other researchers in identifying which features could be most relevant for their specific classification problem.