Remote Sensing of Environment 209 (2018) 793–807

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements

Patrick Aravena Pelizari*, Kristin Spröhnle, Christian Geiß, Elisabeth Schoepfer, Simon Plank, Hannes Taubenböck

German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Wessling, Germany

ARTICLE INFO

Keywords:
Very high spatial resolution imagery
Data fusion
Spectral-spatial features
Feature selection
Object-based image analysis (OBIA)
Classification
Built-up area
Refugee camp mapping

ABSTRACT

Detailed and up-to-date knowledge on the situation in temporary settlements of forced migrants plays an important role for effective humanitarian assistance. These settlements emerge as planned or spontaneous camps or camp-like structures, characterized by a small-scale physical morphology and high dynamics. Information on the built-up area (BUA; i.e. areas occupied by buildings) in these settlements provides important evidence on the local situation. The objective of this work is to present a generic procedure for the detailed extraction of BUA in complex temporary settlements from very high spatial resolution satellite data collected by different sensor types. The proposed approach is embedded in the methodological framework of object-based image analysis and is compound of i) the computation of an exhaustive set of spectral-spatial features aggregated on multiple hierarchic segmentation scales, ii) filter based feature subset selection and iii) supervised classification using a Random Forest classifier. Experimental results are obtained based on Pléiades multispectral optical and TerraSAR-X Staring Spotlight Synthetic Aperture Radar satellite imagery for six distinct but representative test areas within the refugee camp Al Zaatari in Jordan. The experiments include a detailed assessment of classification accuracy for varying configurations of considered feature types and training data set sizes as well as an analysis of the feature selection (FS) outcomes. We observe that the classification accuracy can be improved by the use of multiple segmentation levels as well as the integration of multi-sensor information and different feature types. In addition, the results show the potential of the applied FS approach for the identification of most relevant features. Accuracy values beyond 80% in terms of κ statistic and True Skill Statistic based on significantly reduced feature sets compared to the input underline the potential of the proposed method.

1. Introduction

1.1. Temporary settlement analyses – the benefit of Earth observation

As a consequence of the numerous ongoing crises, large scale displacement of people has reached an unprecedented level in recent history. According to the United Nations High Commissioner for Refugees (UNHCR, 2016), there were about 65.3 million forcibly displaced people worldwide in 2015, including 21.3 million refugees, 40.8 million internally displaced people (IDPs) and 3.2 million asylum seekers. Being forced to flee their homes due to conflict situations, manmade or natural disasters, these people belong to the most vulnerable in the world. Most of them seek protection and shelter in urban environments (Taubenböck et al., 2018), but there are still a large number of refugees living in self-settled or planned camps (UNHCR, 2016). Having arrived in a camp, refugees are generally exposed to poor living conditions with limited access to water, nutrition, medical care and sanitary facilities. Although they are supposed to be temporary, most of these camps are maintained for years or even decades. Thereby complex settlement structures emerge and originally extemporary buildings are solidified and extended (Herz, 2006; Dalal, 2014). At the end of 2015, about 4 million of the world's refugees gather in camps (UNHCR, 2016). Most of these camps are managed or supported by national or international relief organizations, which supply the camps with essential facilities for survival. For effective camp management and decision-making, humanitarian organizations require reliable and up-to date information about the situation on the ground (Bjørgo, 2000; UNHCR, 2000, 2005). In this context, population distribution and numbers are crucial information for relief operations, e.g. to enhance the logistical support of aid agencies (UNHCR, 2007; Ehrlich et al., 2009).

Remote sensing data provide independent, area-wide and up-to-date information on the camp situation and thus can complement

* Corresponding author.
E-mail address: patrick.aravenapelizari@dlr.de (P. Aravena Pelizari).

https://doi.org/10.1016/j.rse.2018.02.025
Received 3 July 2017; Received in revised form 13 December 2017; Accepted 16 February 2018

information usually collected by field observations (e.g. Schöpfer et al., 2015). In some situations, Earth observation (EO) is the only reliable or independent source of information. This accounts particularly in situations where the on-ground situation is unclear due to uncontrolled growth and arrival of new migrants as well as for cases in which field assessments are either unsafe due to conflict situations, or provide false information due to stakeholder bias and/or politics (Bjørgo, 2000; UNHCR, 2000).

Especially with the continuous advent of satellite sensors providing data of increasing spatial and temporal resolution, the role of EO-based applications for support to humanitarian relief becomes more and more important (Kranz et al., 2010; Kuffer et al., 2016a). In case of population estimation, the use of very high spatial resolution (VHR) satellite data allows for the detection of single dwelling units or the built-up area (BUA) in a camp. An approximate figure of the number of people living in a camp can then be derived based on estimated occupancy rates (i.e. average number of people per shelter or per sqm). Furthermore, if total counts of population are available (e.g. registration figures), such information can form the spatial base entity for population disaggregation. In this way information on the spatial pattern of the population (e.g. population density) and respective changes over time can be given (e.g. Lang et al., 2010). The present study focuses on the BUA rather than on single dwelling units. BUA is defined here as the detailed spatial delineation of areas occupied by buildings. Thereby, the term building refers to any type of man-made temporary settlement structure with a roof.

1.2. Image features for Earth observation-based built-up area detection

The development of methods for the derivation of thematic land use/land cover (LULC) information such as BUA from remote sensing imagery has been a major research subject of the remote sensing community in the past decades. EO data from latest VHR sensors such as the multispectral (MS) sensors WorldView (WV) 1–4 or Pléiades allow for LULC mapping at an unprecedented level of spatial detail. Simultaneously, the benefits of increasing spatial resolution are associated with an increasing mismatch between single pixels and the real world objects they depict. Those real world objects are typically represented by numerous pixels. Particularly in heterogeneous environments such as settlement areas this induces high intra-class and low inter-class variability of the different classes in the spectral domain. This poses challenges for an accurate classification (Blaschke, 2010). In addition to MS data, VHR Synthetic Aperture Radar (SAR) data such as imagery from the sensor systems TerraSAR-X (TS-X) or COSMO-Skymed additionally exhibit new potential for accurate settlement mapping (Taubenböck et al., 2012; Chini et al., 2009). The SAR backscatter signal represents a complex combination of various sources, which can provide additional information about objects on the ground due to their distinctive backscatter signature as induced by characteristic geometric as well as dielectric properties and surface roughness. However, being subject to geometric perturbations due to the side looking geometry of SAR sensors (i.e. double bounce, layover, foreshortening and shadowing effects) as well as the speckle effect, the resulting imagery can be difficult to interpret (Brunner et al., 2008; Gamba, 2013).

Recent studies on remote sensing based thematic mapping propose several image processing concepts in order to cope with these challenges. Most of them are based on spectral-spatial homogenization through the extraction of meaningful image features incorporating information beyond the margins of single pixels.

Aiming to capture distinct spatial grey tone dependency patterns, a variety of studies integrate texture features into the classification procedure. A popular and approved approach for computing such measures is the grey-level co-occurrence matrix (GLCM; Haralick et al., 1973). GLCM texture features have demonstrated to beneficially complement MS information in optical data (e.g. Carleer and Wolff, 2006; Pacifici et al., 2009; Geiß et al., 2015; Kuffer et al., 2016b) as well as in single- (e.g. Gamba et al., 2011; Ban et al., 2015; Uhlmann and Kiranyaz, 2014) and multi-polarized (e.g. Du et al., 2015; Masjedi et al., 2016; Wurm et al., 2017) SAR backscatter information.

An additional group of image features being able to complementary encode spectral and spatial information constitutes morphological profiles (MPs). These kinds of features are built upon mathematical morphology (MM) operations (Soille, 2004). MPs rely on the idea of comprehensively describing image structures by their morphological intrinsic characteristics exploited through sequential morphological transformations of the image data applying a structuring element (SE) of increasing size (Pesaresi and Benediktsson, 2001). MPs and its derivatives have proven particularly effective for classification of urban land cover in VHR MS and hyperspectral imagery (Benediktsson et al., 2003; Fauvel et al., 2008; Tuia et al., 2009; Dalla Mura et al., 2010; Ghamisi et al., 2015; Geiß et al., 2016b). None the less, the utility of MPs with regard to the classification of SAR data is sparsely documented in literature. Among the few available studies, Chini et al. (2009) assess the potential of anisotropic MPs for the classification of urban land cover based on single polarized VHR TS-X Stripmap data (6 m spatial resolution) achieving promising accuracy levels. Du et al. (2015) integrate MPs into an array of polarimetric image descriptors for LULC classification based on Radarsat-2 fine quad-pol data (8 m spatial resolution) and report a significant boost in overall accuracy. Wurm et al. (2017) deploy MPs supplementary to texture information for the mapping of inner urban structures, specifically informal settlements using dual-pol TS-X Stripmap mode imagery. They conclude that MPs computed from 6 m spatial resolution imagery do not allow for a meaningful representation of individual objects such as slum dwellings. This indicates that a detailed extraction of temporary settlement BUA under the consideration of MPs might require higher resolved SAR data.

A prominent methodological concept to face the challenges of VHR remote sensing image classification is object-based image analysis (OBIA). OBIA relies on the aggregation of pixel values to meaningful image objects using a segmentation procedure (Benz et al., 2004; Blaschke, 2010). An object-based representation of the imagery allows for a straight forward regularization of the data based on common measures of central tendency or spread (e.g. *mean*, *median* or *standard deviation*). Additionally, it facilitates the spectral-spatial integration of the pixel information (e.g. spectral values, indices and texture) with geometric characteristics (i.e. object shape and extent) as well as object-based contextual measures (e.g. topological relationships) from single or multi source data into the classification procedure (Stumpf and Kerle, 2011; Geiß et al., 2015). The integration of object-based features calculated from a sequence of multiple hierarchical segmentation-levels for classification has been shown to be superior to single-level approaches (Bruzzone and Carlin, 2006; Taubenböck et al., 2010). Such a multi-level strategy on the one hand allows for more adequate object-based feature representations of all different classification targets of interest (e.g. building rooftops of different types and sizes) and on the other hand for the consideration of their spatial context, as represented by affiliated super-object information.

1.3. Studies on building extraction in temporary settlements

Emerging planned or spontaneous temporary settlements such as refugee camps adapt to different natural, social and political conditions. Their physical morphology differs from settlements intended to be permanent. They are composed of small-scale ground-level dwellings of different types (e.g. tents, containers, huts) and materials (e.g. plastic or metal sheet, loam or wood). High temporal dynamics (e.g. due to variations in population pressure), spatial limitations and haphazardly building (e.g. due to uncontrolled population influx) typically induce heterogeneous patterns of BUA often exhibiting high densities. These characteristics impose specific challenges for satellite-based building extraction with respect to temporary settlements.

A variety of methods focusing on the extraction of buildings in