Computers, Environment and Urban Systems 60 (2016) 37–49

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Uncertainty analysis for image interpretations of urban slums

Divyani Kohli *, Alfred Stein, Richard Sliuzas

ITC- Faculty of Geo-Information Science & Earth Observation, University of Twente, Hengelosestraat 99, P.O. Box 217, Enschede, 7500AE, The Netherlands

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 19 May 2016
Accepted 31 July 2016
Available online xxxx

Keywords:
Slums
Existential uncertainty
Extensional uncertainty
Informal settlements
Remote sensing
Random sets
Object-oriented image analysis (OOA)

ABSTRACT

Image interpretations are used to identify slums in object-oriented image analysis (OOA). Such interpretations, however, contain uncertainties which may negatively impact the accuracy of classification. In this paper, we study the spatial uncertainties related to the delineations of slums as observed from very high resolution (VHR) images in the contexts of Ahmedabad (India), Nairobi (Kenya) and Cape Town (South Africa). Nineteen image interpretations and supplementary data were acquired for each context by means of semi-structured questionnaires. Slum areas agreed upon by different experts were determined. Uncertainty was modelled using random sets, and boundary variation was quantified using the bootstrapping method. Results show a highly significant difference between slum identification and delineation for the three contexts, whereas the level of experience in slum-related studies of experts is not significant. Factors of the built environment used by experts to distinguish slums from non-slum areas or leading to deviations in slum identification are discussed. We conclude that uncertainties in slum delineations from VHR images can be quantified successfully using modern spatial statistical methods.

1. Introduction

Currently, owing to unprecedented pace of urbanization in almost all the world regions and especially in developing countries, more than half of the global population is urban (UN-HABITAT, 2010). In addition to urban population growth, governments across many countries face the challenge of high growth rates of urban poverty. Slums are an affordable housing option for the urban poor but often exhibit precarious housing conditions with poor physical and environmental characteristics (Sietchiping, 2004). Several steps are being taken at national and international levels to improve the living conditions of slum-dwellers. Many policies and programs are hence being targeted towards slum eradication by providing better infrastructures or alternative residences to the slum-dwellers. In particular, the Millennium Development Goals (MDGs) formulated by the United Nations (UN) recognized various problems related to poverty and development; and have drawn attention towards the lack of reliable data on slum areas (UN-HABITAT, 2003).

Traditional methods such as census or socioeconomic surveys are used in many countries for data collection on slums. Such methods are field-based and thus, time consuming and difficult to update regularly. A typical census survey is repeated only after ten years. Given the dynamic nature of urban areas, these data can easily become obsolete (Owen & Wong, 2013a; Taubenböck & Kraff, 2014). In addition, slums are frequently omitted from formal statistical assessments, hence updated spatial information on the concentration or location of slum dwellers is generally absent. UN-HABITAT, in a number of publications, has substantiated the need for improved information on slums (UN-HABITAT, 2010; UN-HABITAT, 2014; UN-HABITAT, 2003).

Quantification of slum estimates by identifying and defining slums spatially in a consistent manner may help in geographical targeting within slum intervention programs on global as well as local scales (UN-HABITAT, 2003). The use of remote sensing (RS) based methods can assist in making the provision of data and information on slums readily available. The availability of VHR satellite images in combination of Geographic Information System (GIS) exhibits a strong potential in identifying and characterizing slums. Using VHR images, RS-based methods can help to have a city or country wide measure of slums and facilitate comparison on global scale. The potential of RS to assist in producing statistics on slums have also been acknowledged by international experts. There is, however, a lack of robust and systematic methods (Sliuzas, Mboup, & Sherbinin, 2008). Major hurdles are different definitions across different contexts and the high variability in slum characteristics (Taubenböck & Kraff, 2014; Kohli, Warwadekar, Kerle, Sliuzas, & Stein, 2013).

Urban environments across the world are composed of a number of land cover features. Urban forms may not only vary among different countries but also within a country. Among many factors, these variations can be attributed to the topographic conditions, historical origins, urban policies and financial well-being of a country. This poses a challenge in terms of developing a standard classification technique to fit

* Corresponding author.
E-mail addresses: d.kohli@utwente.nl (D. Kohli), a.stein@utwente.nl (A. Stein), r.sliuzas@utwente.nl (R. Sliuzas).
URL: http://www.itc.nl (D. Kohli).

http://dx.doi.org/10.1016/j.compenvurbsys.2016.07.010
0198-9715/

various contexts. Slums, being a component of such urban environments and geographic locations, exhibit different appearances. There is also a wide variety of definitions for a slum. This interferes with developing a generic method for slum detection. In order to work towards monitoring slums world-wide, a global definition of a slum household was formulated as one lacking in any one of five factors: secure tenure, access to safe water, access to sanitation, sufficient living area and durability of housing (UN-HABITAT, 2006). Kohli, Sliuzas, Kerle, and Stein (2012) developed an ontological approach to conceptualize slums using the *durable housing* indicator that is most relevant for RS-based slum identification and classification. The generic slum ontology (GSO) is a framework comprising the morphological properties of the built environment which can be used to characterize slums. Its strength is context-adaptation by integrating expert knowledge. This leads to a knowledge base that can be used in turn to encode classification rules in OOA (Kohli et al., 2013). Creation of rules in OOA mimics the cognitive approach of visual image analysis. Classification using OOA is based on conceptual understanding of features of interest and may differ if performed by different operators (Belgiu, Drăguţ, & Strobl, 2014). Image interpretation, thus, contains uncertainty which may negatively impact the classification results. Further, to evaluate the results of OOA, accuracy assessment has to be carried out, e.g. by making a comparison with reference data. In many cases, a manual image interpretation is used as the reference. This approach has drawbacks in terms of the reliability of accuracy assessment results. The expertise of the interpreter in a particular field, RS skills and the time invested may lead to discrepancies in interpretation (Van Coillie et al., 2014). For slums, uncertain boundaries, poor definitions, lack of local knowledge and variability may add to the challenges and hence, uncertainties in delineations (Taubenböck & Kraff, 2014; Kit, Lüdeke, & Reckien, 2012; Sietchiping, 2004; Hofmann, Taubenbock, & Werthmann, 2015). Inaccuracies in classification may therefore not be caused by classification alone, but may occur due to the uncertainty or errors in the reference data (Albrecht, Lang, & Hölbling, 2010). If such data are used for assessing the accuracy of object based classification, it is first important to understand the attached uncertainties.

OOA based methods are lately being used for classification of slums but negligible research has been done to address the uncertainties related to slums. Several studies have shown the potential of using RS-based methods, specifically object based methods for mapping slums (Hofmann, 2001; Hofmann, Strobl, Blaschke, & Kux, 2008; Kohli et al., 2013; Kit & Lüdeke, 2013; Kit et al., 2012). Recent research in this area emphasize on identification of qualitative and quantitative indicators to detect slums in VHR data (Hofmann et al., 2015; Sliuzas et al., 2008; Kohli et al., 2012; Owen & Wong, 2013a). Some studies have particularly presented methods on structural and morphological analysis of slums vs. non-slums in spatial-quantitative manner (Shekhar, 2012; Kuffer, Barros, & Sliuzas, 2014; Taubenböck & Kraff, 2014; Stoler et al., 2012; Jain, 2007). Whereas there is literature using descriptive ways to show morphologic variability of slums (Davis, 2006; Neuwirth, 2005), there are also studies that develop classification techniques to classify or differentiate 'slum types' using RS (Owen & Wong, 2013b; Hofmann et al., 2008). While such research shows how different people 'sense' or conceptualize slums and gives an insight into the subjectivity involved, most of the RS-based studies developed a crisp classification method to map slum/non-slum areas or presented slums as an exclusive entity that can be differentiated from the surrounding formal areas (Kit & Lüdeke, 2013; Hofmann et al., 2008). Such a dichotomy may not always be applicable and requires further study in terms of uncertainty related to slum interpretation.

Many studies have evaluated the correspondence between image interpretations by different interpreters (Congalton & Mead, 1983; Lunetta, Liames, Knight, Congalton, & Mace, 2001; Foody & Boyd, 2013; Edwards & Lowell, 1996). These studies focused on evaluating agreements of interpreters on different classifications and focused on rural applications. In particular, interpreters asked to distinguish between different forest land-cover types, reported a wide variation. Studies of the classification of urban areas have been given less attention. Urban areas are generally considered as ensemble of objects with determined boundaries (Campari, 1996). They comprise objects like buildings, roads and pavements that may be easily distinguishable from one another. Alternately, there are uncertain objects or concepts in an urban area, the accurate delineation of which can be complicated. An example is the delineation of the boundary of a city or classifying urban and suburban areas (Stein, Hamm, & Qinghua, 2009). Other examples may be distinctions between private and public open space, vegetated/landscaped areas in cities.

As stated by Couclelis (1996), *"Boundaries constitute the outer limits of individual entities, but also the locus where two or more different entities meet; they enclose and separate, divide and join, distinguish and juxtapose, contain and include, create interiors and exteriors, help tell same from other."* Taking this definition, the current study considers uncertainty as the inability to draw a clear boundary about an object's existence, and where it ceases to exist. In a remote sensing environment, this links image features with objects on ground. In an urban context, detection of boundaries between different homogeneous morphologic zones can help to understand constituent patterns and spatial dynamics (Taubenböck, Esch and Roth, 2006). Several studies have attempted to develop methods to derive homogeneous urban zones using RS or ancillary spatial data. Taubenböck, Habermeyer, et al. (2006) used a moving-window approach and Savitzky-Golay filtering in combination with built-up densities to segment homogeneous zones and infer socio-economic characteristics directly from RS data. Other studies have used textural analysis (Pesaresi & Bianchin, 2000) and geostatistical analysis for urban pattern characterization (Brivio & Zilioli, 2001). While RS-based border identification for urban zoning deals with intrinsic difficulties such as edge effect, pre-defined borders such as road layers have sometimes been used as ancillary data (Bauer & Steinnocher, 2001; Batty & Longley, 1987; Taubenböck, Habermeyer, et al., 2006).

A slum area can be considered a type of homogeneous zone with specific spatial characteristics. There is, however, a great deal of complexity in case of slums as they can exhibit different appearances and definitions depending on context, making it challenging to detect with RS techniques. An additional challenge is posed where there is uncertainty in the exact location of boundaries on ground also, as for example, to find out where exactly a slum transitions into a non-slum. In other words, there could be vague boundaries leading to uncertainty. This considers first the question whether a slum exists, and if so: how its extent can be determined. Related uncertainty is termed existential and extensional uncertainty, respectively (Molenaar, 2000). Existential uncertainty expresses the uncertainty about the existence of a slum in reality. It thus refers to the possibility that a slum is delineated by experts on an image while in reality it does exist; this may depend upon experts' inexperience and conceptual differences in interpretation. Extensional uncertainty implies that the area covered by a slum can be determined with limited certainty, i.e. with boundaries that reflect different perceptions of slums by experts. Extensional uncertainty in slum identification includes differences in expertise among the experts, the applied slum definitions, time invested and the degree of generalization.

The aim of this paper is to study deviations in slum identification and their delineations as observed from VHR images. Existential and extensional uncertainties of slums are estimated using images from three contexts interpreted by 19 experts. Images from Ahmedabad (India), Nairobi (Kenya) and Cape Town (South Africa) were used. First, the percentage of slum area agreed by most experts and the uncertain area is calculated. Second, we model uncertainty in terms of random sets and apply bootstrapping method to show confidence in various delineations and boundaries. Finally, we identify the factors of the built environment that experts use to distinguish slums from non-slum areas and also study the factors leading to deviations in slum identification in the three contexts.