Urban slum detection using texture and spatial metrics derived from satellite imagery

Divyani Kohli, Richard Sliuzas and Alfred Stein

ITC- Faculty of Geo-Information Science & Earth Observation, University of Twente, Enschede, The Netherlands

ABSTRACT
Slum detection from satellite imagery is challenging due to the variability in slum types and definitions. This research aimed at developing a method for slum detection based on the morphology of the built environment. The method consists of segmentation followed by hierarchical classification using object-oriented image analysis and integrating expert knowledge in the form of a local slum ontology. Results show that textural feature contrast derived from a grey-level co-occurrence matrix was useful for delineating segments of slum areas or parts thereof. Spatial metrics such as the size of segments and proportions of vegetation and built-up were used for slum detection. The percentage of agreement between the reference layer and slum classification was 60 percent. This is lower than the accuracy achieved for land cover classification (80.8 percent), due to large variations. We conclude that the method produces useful results and has potential for successful application in contexts with similar morphology.

1. Introduction
An important consequence of urbanisation in developing countries is the high growth rate of urban poverty. Urban poverty is manifested most clearly in the proliferation and expansion of slums, home to a growing proportion of urban dwellers. According to UN-HABITAT (2010), the world’s slum population is expected to reach 889 million by the year 2020. Although improvements have been made in the living conditions of many slum-dwellers as a result of local and global efforts, their number is still estimated to grow by an average of 6 million every year (op. cit.). Regularly updated information on slums is therefore needed to guide policies and programmes aimed at interventions and improvements (UN-HABITAT 2014). Quantification of slum estimates by identifying and defining slums spatially in a consistent manner is required to be able to use geographical targeting for slum intervention programs (Sliuzas et al. 2008).

In remote sensing (RS), object-oriented image analysis (OOA) methods have been shown to provide improved analyses of urban areas (Cleve et al. 2008; Ebert et al. 2009; Myint et al. 2013; Taubenböck & Kraff 2014; Belgiu et al. 2014a). Such methods are also referred to as
object-based image analysis (OBIA) or geographic object-based image analysis (GEOBIA). OOA has great potential for informal settlement mapping because of its ability to include spatial, spectral and contextual characteristics similar to human cognitive image interpretation. This approach integrates expert knowledge during classification, which is a major advantage over conventional pixel-based approaches (Haala & Brenner 1999; Hofmann 2001; Hofmann et al. 2008; Kohli et al. 2013). Image segmentation in combination with a hierarchical object-based classification has been used for mapping residential land use and identifying socioeconomic status from very high-resolution (VHR) imageries (Herold et al. 2003; Stow et al. 2007, 2010). Examples include studies on determining the socioeconomic status of residential areas using spatial patterns and configuration, such as the vegetation, impervious surface, bare soil (V-I-S) model proposed by Ridd (Ridd 1995; Weeks et al. 2007; Stow et al. 2010).

In this paper, we explore the potential of using texture and spatial metrics in the Indian city of Pune, to identify and classify slum areas. The proportions of the land cover classes, i.e. vegetation and impervious, serve as proxies for the identification of slums. Previous research has shown that inclusion of contextual knowledge encoded in an ontological framework can be used for image-based classification (Kohli et al. 2013; Belgiu et al. 2014b). Use of the generic slum ontology (GSO), based upon indicators related to the morphology of the built environment, helps to address the issue of variable definitions and appearance of slums in different contexts (Kohli et al. 2012). In this paper, we adapt the GSO to slums in Pune and then follow an OOA-based classification at two spatial levels. The objectives are as follows:

- to determine the applicability of the GSO to the Indian city of Pune;
- to explore the use of texture and spatial metrics to quantify the spatial patterns at different ontological levels in an OOA environment;
- to test the methodology of slum identification based on the V-I-S model.

Previous research applied the V-I-S model to infer land use from the proportional abundance of each of the land covers (Herold et al. 2002, 2003). These studies, however, were mainly carried out for Western cities. Similar studies, dealing with slum identification as well, were done for the African city of Accra (Weeks et al. 2007; Stow et al. 2007, 2010; Stoler et al. 2012). To the best of our knowledge, such studies do not exist for Indian cities. With their complex morphologies that are a result of organic growth or their historical development, Indian cities pose a challenge for slum studies. Slum identification itself is already a complicated task considering the variability across contexts (Sliuzas et al. 2008). The aim of this paper is address these issues and detect urban slums using texture and spatial metrics derived from satellite imagery. This is done by combining the GSO, the V-I-S model and an OOA method.

2. Definition of slums

Various reasons might explain the occurrence of slums in different cities of the world. Slums tend to share some common features globally (Patino & Duque 2013). For instance, most slums are characterised by a high concentration of poor people, substandard shelter and poor physical environmental conditions. These characteristics also date back to the slums of medieval cities in the Western world (Haala & Brenner 1999; Neuwirth 2005). Over time,
different countries have developed national and, in some cases, local definitions of slums. 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 et al. (2012) developed an ontological approach to conceptualise slums using the durable housing indicator, which is most relevant for RS-based slum identification and classification. In that study, the authors conducted a survey to find the most common physical features of slums concerning the structural quality of housing, infrastructure and location. These features were compiled to form the GSO, which may be adapted to different contexts. The GSO comprises indicators related to the morphology of the built environment at three spatial levels: object, settlement and environs levels. It provides a comprehensive description of spatial characteristics and their relationships to characterise slums in a very high-resolution (VHR) satellite image. In this paper, we used the GSO as the basis for classification by adapting it to Pune, India.

3. Materials and methods

3.1. Study area

Pune is one of the fastest developing urban agglomerations in Asia and ranks eighth at national level in India (COI 2001). It covers an area of 244 km² and is the cultural and educational center of Maharashtra State. Its population has grown significantly over the past two decades. Favorable geographic location, pleasant climate, immense employment opportunities and improved educational facilities are some of the reasons attributed to high migration of people from other parts of the country. From 1981 to 1991, the population grew by 30.2 percent, and between 1991 and 2001 by 62.2 percent (Shekhar 2010). With approximately 1 million slum dwellers out of the total population of about 3.15 million (MASHAL 2011), India’s Town and Country Planning Organisation (TCPO) ranks Pune third in Indian cities with the largest number of slums.

3.2. Image data

A cloud-free pan-sharpened Quickbird satellite image (0.6 m resolution) of red, green and blue bands from the year 2006 was used for this research. The reason for using a relatively old image was its availability and that it was also reasonably close to the acquisition period (2006 to 2008) of the slum atlas reference data available from MASHAL, a local NGO. The image scene covering the central part of the city was selected for analysis (Figure 1). This scene is representative of the urban form of the entire city as it contains most of the land use types found in Pune. Roads and water bodies were digitised from the image and were also used in analysis. For accuracy assessment, a polygon layer of the slum pockets, provided by MASHAL, was used as reference data (Figure 2). According to their survey, a total of 477 slum pockets were identified in Pune. Approximately 33 percent of the total population of Pune lives in only 2.34 percent of the total land of the city. Many of the slums are located on hill slopes, along rivers and in other environmentally sensitive areas (MASHAL 2011). According to Pune’s disaster management plan (PMC 2012), slums are most vulnerable to flood, earthquake and landslide disasters. Given their vulnerability, it is important to update information on slum dynamics to monitor and potentially regulate their further growth.
Updated spatial information can also help in crisis management as many slums of Pune lie in hazardous areas (Shekhar 2010).

### 3.3. Local ontology of slums in Pune

A major challenge in slum identification from satellite images is the morphological resemblance of slums to non-slum areas (Kit et al. 2012; Kuffer et al. 2014). Prior knowledge of the spatial pattern and context is thus useful in detecting slums in a particular urban environment. To do so, the GSO has been adapted to the local conditions of Pune slums. Relevant indicators were chosen and arranged to form a local slum ontology. Specific values and observations of the indicators referring to each ontological concept are given in Table 1. The table shows the specification of physical characteristics adapted from the generic ontology used for identification and delineation of slums, from environs level through settlement level to object level. The characteristics of slums in Pune were observed by visual interpretation and ground knowledge. Ground data collection included visits to some slum settlements and discussion with experts working for slum-related issues. The observations at three spatial levels of ontology are as follows:

- At environs level, slums in Pune tend to be located on steep slopes, flood-prone and marshy areas, along canals, railway tracks and major roads. These locations are hazardous and generally unattractive for planned development. Slums are also found close to areas that offer employment opportunities such as the central business district (CBD), industrial areas and middle/high-socioeconomic-status neighbourhoods.
At settlement level, most slums are highly compact, displaying a roof coverage of more than 70 percent and vegetation/open spaces comprising less than 20 percent of the settlement. The shape of the slums is often irregular, except for those along canals and roads, which tend to have a linear shape.

At object level, the small size of buildings, variable roof materials, the irregularity, narrowness and unpaved surface of roads were important observations for slum identification.

The above observations have been categorised and analysed based on the ontological framework. The relevant indicators referring to each ontological concept are further assigned specific OOA parameters (Table 1).
3.4. Conversion of local ontology to OOA parameters

OOA parameterisation was guided by the local ontology, which defines the different features for classification. Conversion of ontological indicators into object-based parameters was done qualitatively by integrating expert knowledge until object classification was achieved to a visually satisfying level. Spectral range, geometry, texture and association were mainly chosen to translate the local slum knowledge into specific OOA parameters. This study focuses at the settlement and object levels, being the most relevant with respect to the high-resolution image.

The first step in OOA is image segmentation, i.e. dividing the image into regions or objects of homogeneous pixel values within segmented objects. Multi-resolution segmentation is a bottom-up segmentation method based on a pair-wise region merging technique. The size and constituents of segments are controlled by assigning appropriate values to the key parameters, scale parameter ($SP$), shape ($w_{\text{shape}}$) and compactness ($w_{\text{compt}}$) to segment objects. For definitions of these parameters, refer to Kohli et al. (2013) or the standard reference source (Trimble 2014). A value of 0.5 was assigned to $w_{\text{shape}}$ to give equal weight to shape and spectral reflectance. Expert knowledge, driven by purposes of classification, is used to decide upon the values of $SP$, $w_{\text{shape}}$ and $w_{\text{compt}}$ (Baatz & Schäpe 2000). Two bottom-up hierarchical segmentations representing settlement level (Level 2) and object level (Level 1) were implemented (Figure 3).

At Level 1, image object primitives were generated representing the basic land cover classes such as built-up and trees. The built-up class includes building roofs and other imperious materials such as parking lots. A value of $SP = 40$ was selected for Level 1 through iterative and interactive control. Many OOA-based studies continue to use a trial and error approach to determine the appropriate $SP$ as no universally accepted tool to determine the optimal scale exists (Hu & Weng 2010; Aguilar et al. 2013). The $SP$ value is chosen based on a qualitative analysis (visual inspection) of the resulting segmentation and purpose of classification. The ESP tool proposed by Dragut et al. (2010) does provide a range of optimal

| Level                  | Indicators                  | Observation (Pune)                                                                 | OOA parameterisation                          |
|------------------------|-----------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------|
| Environ level          | Location                    | Often close to highways, major roads, railways, canals, along rivers prone to seasonal flooding, steep slopes, marshy land | Vector layers and appropriate DEM not available |
|                        | Neighbourhood characteristics| Close to employment opportunities such as CBD, industrial areas, middle/high-socioeconomic-status neighbourhoods. | Not used due to absence of spatial data on the socioeconomic status |
| Settlement level       | Shape                       | Irregular, linear; Highly dense buildings (roof area coverage > 70 percent); low vegetation and open spaces (vegetation coverage < 20 percent) | Not used; Texture – contrast; geometry – area of built-up, vegetation |
|                        | Density                     | Geometry – area, size of segments; spectral – mode, layer mean value; texture – entropy; association – distance to features | Geometry – area, size of segments; spectral – mode, layer mean value; texture – entropy; association – distance to features |
| Object level           | Building                    | Shape – variable; size – 10 to 40 m²; roof material – corrugated iron sheets, tin, concrete, plastic; colour – variable; often regular orientation but clumped together | Geometry – area, size of segments; spectral – mode, layer mean value; texture – entropy; association – distance to features |
|                        | Access network              | Irregular shape; type – paved/unpaved access streets; width – variable, often narrow | Vector layer of roads; association – relative area, distance to features |

Table 1. Adaptation of GSO concepts to an ontology for Pune with OOA parameterisation at environs, settlement and object levels.
scale parameters. The problem again is to choose the right SP according to user requirements. Our study deals with two scales of analysis, the object and settlement level. Since the ESP tool uses only one image layer as input for optimal scale determination, its implementation in the current study was not useful. After experimenting with various values (15 to 60 with an increment of 5 for Level 1 and 150 to 500 with an increment of 50 for Level 2), we chose the appropriate SP values for segmentation at the two levels. At Level 2, the segmentation was done at a coarser level to represent the settlement level of the ontology. The purpose was to delineate approximate boundaries of settlement primitives in slum areas. By settlement primitives, we mean objects representing an entire slum settlement or significant parts of a slum. Along with the three bands of the image, the GLCM contrast for the blue band (GLCM_{Con}(B)), calculated using Haralick’s method, was used as an additional layer of texture for this segmentation (Trimble 2014).

eCognition developer, the software used for OOA in this study, can create a temporary layer that can be used for segmentation or classification. A temporary GLCM layer was generated where values were calculated for each object individually at Level 1, using a 256 by 256 matrix, with a direction equal to 0°. This layer and the blue, green and red bands were used as input for segmentation at Level 2. A value of $SP = 300$ was found to be visually satisfying to generate segments comprising areas with homogeneous texture. Again, a value of 0.5 was assigned to $w_{shape}$ to give equal weight to shape and spectral reflectance. The classes used for the higher-level segmentation were built-up, vegetation and shadow. Most slums in Pune comprise highly dense buildings, with a low proportion of vegetation and open space (Table 1). Thus, slum areas have a low variance within segments and a clear contrast to planned areas. The use of a GLCM contrast layer as an additional layer helped to delineate segments representing the boundaries of slums (Table 2). These were in contrast to segments in planned areas, which have a composition of different land cover features (e.g. buildings, vegetation, roads) (Figure 4).

### 3.5. Classification at Level 1

Segmentation was followed by classification at Levels 1 and 2. Segmentation results in a large number of image objects for which a variety of object characteristics or features can

![Figure 3. Method linking texture layer and spatial metrics to GSO at object and settlement levels.](image-url)
be calculated and subsequently used for classification. For each object, feature values such as spectral values, shape characteristics, texture, size and contextual relations are computed based on purpose of classification (Benz et al. 2004). Rule-based classification, also known as the membership function classifier, can be used to assign image objects to the desired

---

**Table 2.** Definitions of the parameters used in the classification. For details, refer to Trimble (2014).

| Parameter | Description | Definition |
|-----------|-------------|------------|
| GR | Green ratio | An index to measure vegetation, defined as (green / (red + green + blue)) |
| Brightness | Brightness of object | The mean intensity of all image bands for an image object |
| GLCM | Grey-level co-occurrence matrix | Proximal combinations of pixel brightness values (grey levels) within a particular band of an image |
| GLCM\(_{ent}(R)\) | GLCM entropy for red band | The measure of orderliness within the red band of the image, which relates to textural homogeneity |
| GLCM\(_{con}(B)\) | GLCM contrast for the blue band | The amount of local variation within the blue band of the image |
| Mode\(_{R}\) | Mode for red band | Returns the most frequently occurring pixel value per object in the red band |
| Area | Area of object | The number of pixels forming an image object |
| Merge | Merge region | Neighbouring image objects of the same class are merged |
| RA(S) | Relative area of neighbour object (shadow) | The area covered by image objects of a selected class, found around the selected image object, divided by the total area of image objects inside this area |
| MG | Mean green | The mean intensity of all pixels forming an image object in the green band |
| D(SL) | Distance to slums | The distance (in pixels) of an image object to the objects of selected class |
| ES | Enclosedness by slums | Image objects that are completely enclosed by objects belonging to selected class |

---

**Figure 4.** Settlement primitives at Level 2 using texture layer GLCM contrast: (a) highly dense buildings in slums outlined and displaying a clear contrast to neighbouring segments; (b) planned area with segments comprising different land cover features.
class by integrating prior knowledge (Baltsavias 2004; Belgiu et al. 2014a). In this study, knowledge of slums in the form of a local ontology was used as a basis to determine the most relevant features and corresponding thresholds to classify the image objects. A sequence of steps defining the class membership conditions forms a ruleset where each step comprises a condition/rule leading to classification of the respective class. Several steps may be required to classify a single class. All these steps were executed in eCognition developer and are explained in detail below.

Classification of the segmented image was performed in two steps. First, the non-built-up classes such as water, road, shadows and vegetation were classified. Due to the absence of the near infra-red (NIR) band, vegetation was classified using the customised arithmetic ratio, the green ratio (GR) (Shekhar 2012). GR values above 0.42 were used to classify vegetation. Shadows of trees and buildings were detected using the low values of brightness, and were classified using mean Brightness < 310. These classes were useful in defining contextual relationships for characterisation of built-up and also to avoid misclassifications. The vector layer of water bodies was used to classify water. Road classification in urban areas is quite challenging because the shadows of buildings and vegetation interfere with correct delineation. Vegetation can, in some cases, hide roads completely. Thus, the digitised road layer was used to classify roads. This road layer comprises the centerlines of the roads; the objects overlapping the layer were classified as roads.

For the classification of built-up areas, spectral measures and texture were used. The spectral feature mode for the red band, Mode(R), was first used to classify built-up. This feature calculates and returns the most frequently occurring pixel value in the red band in an object (Table 2). Mode(R) values above 360 corresponded to the built-up class, especially the bright-roofed buildings. The remaining built-up, comprising mostly dark-roofed/old buildings, was classified using texture. Previous studies have shown the usefulness of texture for built-up classification (Kohli et al. 2013; Belgiu et al. 2014a; Shekhar 2012). GLCM entropy of the red band was used to discriminate the remaining built-up from other land cover features. The entropy as a measure of disorder reflects the distribution of grey values and relates to textural homogeneity. By visual inspection, high values of entropy for the red band were found useful for classifying built-up. In this study, GLCM Ent(R) > 3.8 was used to classify the remaining built-up area. Slightly lower GLCM Ent(R) corresponds to bright roofs and bare soil. Use of Mode(R) was helpful to avoid false positives from bare soil. The results of level 1 were further used for classification of slums at Level 2.

3.6. Classification at Level 2

From the local ontology, it is evident that the slums in Pune are highly dense and often have a clear contrast to the formal areas (Figure 4). We used the level 1 (object level) classification results to calculate the spatial metrics for the V-I-S for Level 2 settlement primitives. The built-up class, in this study, is synonymous with the impervious class as it includes building roofs and other impervious materials. Considering this particular urban context, the use of soil was insignificant for classification. Area, in terms of the size of segments, is useful in identifying and characterising slums (Kohli et al. 2013; Kuffer et al. 2014). Large Level 2 segments (Area > 14,000 pixels) tend to correspond to slums and included the smallest slum in the study area. Out of these large segments those consisting of relative areas of built-up > 50 percent and vegetation < 30 percent (Table 1) were classified as slum settlements. Subsequent
steps involved bringing Level 2 classification results to Level 1 to get a final classified map together with other classes (Figure 6). Additionally a clean-up process was followed to eliminate false positives from the classified slums and reclassify any observed, misclassified slums. These steps are explained in the following section.

3.6.1. **Clean-up**

The Level 2 segments (classified and unclassified) were converted to objects at Level 1 (Figure 3). This step re-segmented the larger segments at Level 2 to smaller Level 1 segments. The built-up objects at Level 1 which overlapped with slum segments from Level 2 were re-classified as slum. The merge region feature was used to merge neighbouring objects of the slum class. A threshold of \( \text{Area} \leq 14,000 \) pixels was used to remove small buildings classified as slum. Relation to neighbouring objects was used to remove false positives from the Old City (Table 2). By visual inspection, the value of relative area of neighbour object shadow \( RA(S) > 0.24 \) was found appropriate to remove buildings from the Old City classified as slums. Further, mean green \( MG > 540 \) was used to remove the false positives that occurred at the bottom-right of the scene, displaying higher values as compared to slum. These were large buildings that displayed low contrast and were thus classified as slum. To include the misclassified objects on the fringes of slum areas, distance to slums \( D(\text{SL}) < 1 \) pixel was used. After merging the classified objects of the same class, the feature enclosed by slum \( ES \) was used to include any misclassified object embedded within a slum area. Table 3 shows the final ruleset summarising the sequence of steps followed for classification.

3.7. **Accuracy assessment**

Random stratified sampling was used to assess the accuracy at Level 1. A total of 250 points were generated with the aim of having a minimum of 50 points for each class (Scepan 1999; Herold et al. 2008). A standard confusion matrix was used to calculate the overall accuracy and user’s and producer’s accuracy (Congalton 1991). The points were visually interpreted on the Quickbird image for reference.

At Level 2, the slum layer provided by MASHAL was used (Figure 2). The accuracy of slum classification was assessed by using the ‘error matrix based on training and test area (TTA) mask’ in eCognition developer (Trimble 2014). The TTA mask, when used to calculate the accuracy of a single class, measures the matching of reference pixels and classified pixels inside the reference areas (Hofmann et al. 2008). The error matrix results thus give the

| Seq No. | Classification steps | Ruleset |
|---------|----------------------|---------|
| 1       | Generate object primitives at Level 1 | Multiresolution segmentation \( SP = 40, w_{\text{shape}} = 0.5, w_{\text{compt}} = 0.5 \) |
| 2       | Classify roads and water | Assign class using thematic layer |
| 2       | Classify vegetation | \( GR \geq 0.42 \) |
| 3       | Classify shadow | \( Brightness < 310 \) |
| 4       | Classify built-up | \( \text{Mode\( (R) \)} > 360, \text{GLCMEnt\( (R) \)} > 3.8 \) |
| 5       | Generate settlement primitives at Level 2 | Multiresolution segmentation \( SP = 300, w_{\text{shape}} = 0.5, w_{\text{compt}} = 0.5 \) |
| 6       | Classify slums | \( \text{Area} \geq 14,000, \text{Built-up} > 50 \% \text{and vegetation} < 30 \% \) |
| 7       | Clean-up | \( \text{Area} \leq 14,000, RA(S) \geq 0.24, MG > 540, D(\text{SL}) \leq 1, ES \) |
percentage of agreement between the reference map and an error of omission. In addition, random stratified sampling was used to assess the accuracy at Level 2. We generated a total of 150 points with a minimum of 50 points each for slums, non-slum built-up and the remaining classes. The points were visually interpreted by two slum mapping experts, referred as E1 and E2 in this paper. Accuracy assessment yielded the overall accuracies and producer’s and user’s accuracies for the two responses.

4. Results

4.1. Segmentation

For segmentation at Level 1, a value of \( SP = 40 \) worked well to segment built-up, vegetation and shadow appropriately. For the built-up class, the segments represent individual buildings in the non-slum areas and clumps of buildings in slum areas (Figure 5), as the slum buildings tend to be small and densely packed, making the individual structures indistinguishable at 60 cm resolution of the Quickbird image. For segmentation at Level 2 we used a value of \( SP = 300 \) to segment settlement primitives (Table 3). The use of a temporary layer of \( GLCM_{Con}(B) \) in addition to blue, green and red bands resulted in segments with homogeneous texture representing slum settlements (Figure 4). Segments in the planned areas thus comprise a mix of buildings, vegetation, shadows and other classes.

![Figure 5](image-url)

**Figure 5.** Classification of image at Level 1: (a) magnified view of a classified slum area; (b) segmentation result of the same area.
4.2. **Classification at Level 1**

Figure 5 shows the classified built-up, vegetation and other classes at Level 1. The segmented individual buildings in the planned areas and clumped buildings in slum areas were classified as built-up. Some confusion between the shadow and vegetation classes was observed, specifically with the shadows of tall buildings being classified as vegetation. Due to the use of GR for vegetation classification, such misclassifications occurred, though they did not directly affect the classification of slums at Level 2 as slum buildings hardly display any shadows due to their compactness. Other classes, water and roads were classified appropriately using digitised vector layers. These classification results were used as a basis for the Level 2 classification.

4.3. **Classification of slums**

The segments representing settlement primitives of slum areas were mostly composed of compact buildings (Table 1). Large segments comprising relative areas of built-up > 50 percent and vegetation < 30 percent (Table 1) were classified as slum areas. The proportion of built-up may be as high as 70 percent for entire slum settlements. A lower value was used, however, to include parts of slums with lower built-up density. Slums with high contrast to their surrounding areas were successfully delineated. False positives existed in the Old City area that also comprises areas of dense, often old and dilapidated buildings. The Old City is composed of many multi-storied buildings displaying a clear shadow. Relation with shadow RA(S) was used to remove these false positives. Misclassifications were also found in other areas where comparatively smaller segments have high built-up proportions, such as the market areas. Finally, a sequence of steps using MG, D(SL) and ES was used for clean-up (Table 3). Figure 6 shows the classification results and the ground pictures of the slums and part of the Old City.

4.4. **Accuracy assessment**

After classifying the whole scene (Figure 6), the accuracy assessment was carried out.

4.4.1. **Accuracy of basic land cover features**

The Level 1 classes were separated using spectral information and texture. The overall accuracy of the Level 1 classification was 81 percent (Table 4). The classes assessed were shadow, vegetation and built-up. The ‘others’ class (shown as unclassified on the map) comprises the remaining land cover in the scene such as bare soil and barren land. These classes were not directly relevant for this study, thus left as unclassified. The relevance, though, is in terms of misclassifications. For example, built-up was classified with maximum accuracy with high producer accuracy (88.24 percent) as well as high user accuracy (82.19 percent) though some confusion with bare soil (‘others’) is evident. Comparatively, vegetation has lower producer accuracy (80.88 percent) due to confusion with the shadow class. Vegetation has false positives from the shadow class and vice versa. Shadow is classified with the lowest accuracies (62 percent and 75.61 percent). This is due to spectral similarity to barren land, dark-coloured roofs and vegetation. The goal of Level 1 classification was to provide a land cover product
Figure 6. Final classified map of the study area with slums, ground pictures and corresponding zoomed views of the image: (a) ground photograph of buildings in the Old City; (b) ground photograph of a typical slum settlement; (c) magnified view of a part of the Old City misclassified as slum on the final map; (d) a typical slum settlement.

Table 4. The producer’s and user’s accuracies of the land cover classes at Level 1.

|          | Shadow | Vegetation | Built-up | Others | Total | PA (percent) | UA (percent) |
|----------|--------|------------|----------|--------|-------|--------------|--------------|
| Shadow   | 31     | 10         | 0        | 0      | 41    | 62.00        | 75.61        |
| Vegetation| 10    | 55         | 0        | 0      | 65    | 80.88        | 84.62        |
| Built-up | 4      | 1          | 60       | 8      | 73    | 88.24        | 82.19        |
| Others   | 5      | 2          | 8        | 56     | 71    | 87.50        | 78.87        |
| Total    | 50     | 68         | 68       | 64     |       | OA 80.80 percent |              |

PA = producer accuracy, UA = user accuracy, OA = overall accuracy.
which could form the basis for further analysis at Level 2. An accuracy of 81 percent can be considered good and encouraging for Level 2 analysis (Herold et al. 2003).

4.4.2. Accuracy of classified slums

The accuracy results demonstrate that 60 percent of the pixels of the reference map were also detected as slums in classification at Level 2; i.e. the error of omission was 40 percent. Errors in the classified map can be attributed to a number of reasons. First, the amount of generalisation in the outlines of the reference layer could have led to errors (Figure 7a). Secondly, the process of defining slums in India can be rather complicated due to political and administrative interests. The definition used by MASHAL for slum delineations, which incorporates these political and administrative perspectives, also impacts our results. Thus, many areas that are morphologically slum-like are classified as slums but are absent in the reference layer (Figure 7b). However, in a data-poor environment like India, this is the best reference that could be procured. Thirdly, some of the polygons of the reference layer show...
vegetation on the ground instead of slum areas (Figure 7c). These could be small slum settlements hidden under trees and hence missed out in the classification. During field work, we observed that many buildings in the Old City and market areas were as dilapidated as slum buildings, but these areas are, by definition, not considered slums by MASHAL but some of these areas are classified as slums. High-density built-up and low vegetation are used as a proxy for slum presence. Though this approach has worked for classifying slums in the scene, there are misclassifications from high-density non-slum areas (such as the Old City and market areas). These misclassifications could also be attributed to the errors stemming from basic land cover classification (Figure 7d). There are some parts of the Old City where use of shadow could not remove misclassified slums. Though some of the misclassifications were removed by using MG, false positives still exist. Finally, the use of a single class for error matrix calculation results in the percentage of agreement between the reference map and the slum classification. This is because of the absence of non-desired classes in the reference layer (non-slum in this case). The overall accuracy thus ignores the false positives outside the reference layer and also the unclassified pixels inside the reference layer.

### 4.4.3. Accuracy comparison with visual interpretations

The overall accuracy resulting from the visual interpretation by E1 and E2 was 68 percent and 71.33 percent. The classes assessed were slums and non-slum. The non-slum class consisted of other built-up. Since the accuracy calculation was at settlement level, the experts were instructed to allocate shadows, vegetation and roads within a slum settlement to the slum class. The ‘rest’ class comprised the remaining land cover in the scene such as water bodies, shadow, bare soil and barren land (Table 5). Clearly there are higher accuracies for slum classification using visual interpretation than using the vector layer of the Pune slum atlas. The producer’s accuracies for both E1 and E2 are much higher than the respective user’s accuracies. Apparently, both experts identified some classified slums as non-slum built-up. These interpretations were from the Old City and other informal-looking formal areas. There was also some confusion on including or excluding the fringes of slum settlements within the slum class.

**Table 5. The producer’s and user’s accuracies of the slum classification resulting from visual interpretation by E1 and E2.**

|          | Slums | Non-slum | Rest | Total | PA (percent) | UA (percent) |
|----------|-------|----------|------|-------|--------------|--------------|
| **E1**   |       |          |      |       |              |              |
| Slums    | 30    | 4        | 2    | 36    | 83.33        | 60.00        |
| Non-slum | 12    | 33       | 9    | 54    | 61.11        | 66.00        |
| Rest     | 8     | 13       | 39   | 60    | 65.00        | 78.00        |
| Total    | 50    | 50       | 50   | 50    |              |              |
| **E2**   |       |          |      |       |              |              |
| Slums    | 38    | 5        | 2    | 45    | 84.44        | 76.00        |
| Non-slum | 10    | 42       | 21   | 73    | 57.53        | 84.00        |
| Rest     | 2     | 3        | 27   | 32    | 84.38        | 54.00        |
| Total    | 50    | 50       | 50   | 50    |              |              |

PA = producer accuracy, UA = user accuracy, OA = overall accuracy.

The overall accuracy thus ignores the false positives outside the reference layer and also the unclassified pixels inside the reference layer.

**4.4.3. Accuracy comparison with visual interpretations**

The overall accuracy resulting from the visual interpretation by E1 and E2 was 68 percent and 71.33 percent. The classes assessed were slums and non-slum. The non-slum class consisted of other built-up. Since the accuracy calculation was at settlement level, the experts were instructed to allocate shadows, vegetation and roads within a slum settlement to the slum class. The ‘rest’ class comprised the remaining land cover in the scene such as water bodies, shadow, bare soil and barren land (Table 5). Clearly there are higher accuracies for slum classification using visual interpretation than using the vector layer of the Pune slum atlas. The producer’s accuracies for both E1 and E2 are much higher than the respective user’s accuracies. Apparently, both experts identified some classified slums as non-slum built-up. These interpretations were from the Old City and other informal-looking formal areas. There was also some confusion on including or excluding the fringes of slum settlements within the slum class.
5. Discussion

Unprecedented urbanisation in many countries of the developing world has led to growth of slums. Slums provide an affordable housing option to those who cannot find a place in the formal housing market. Slums form an integral part of city landscapes in many cities in India. With the growing economy, more and more people opt to leave villages for cities in search of better opportunities. Pune city in India is one of the fastest growing metropolitan areas with almost 40 percent of its people living in slums (MASHAL 2011). Considering the scale of the slum population, detecting and mapping existing slums may facilitate better policy formulations for improvement and to provide infrastructure or alternatives. It could also be important to monitor such developments to get updated information for crisis management (PMC 2012). The fact that many slums in Pune are located in hazardous areas, considered unattractive for any planned development, substantiates this need. Mapped slum settlements may be used as an important input for climate change mitigation activities (Kit et al. 2012). Considering the location of a settlement (environ) in addition to other indicators, in a previous work a GSO was developed to be used as a basis for systematic slum identification using VHR images (Kohli et al. 2012). Subsequently, the GSO was adapted for slum detection and classification in the Indian city of Ahmedabad (Kohli et al. 2013).

Developing rulesets in OOA requires considerable image analysis skills, expertise in the application area and time (Belgiu et al. 2014a). Thus, there have been recent efforts to make rulesets/methods more transferable for wider application (Hofmann et al. 2011; Kohli et al. 2013; Belgiu et al. 2014a). In our previous work, we identified a set of stable parameters, in an OOA environment, that worked on subsets of varying characteristics for classifying built-up and slums. In this paper, we take that work further by developing a method that combines the GSO, parameterisation in OOA (Table 1) and a V-I-S model (Ridd 1995) to detect slums in an Indian city. This approach not only considerably reduces the number of parameters used in the ruleset, it also presents the possibility of successful application in similar morphological contexts. Our study also builds upon similar work where the authors use ontology-driven class definitions for classification of informal settlements in Quickbird images of Rio de Janiero (Hofmann et al. 2008). The complexity of ruleset in that study, however, may restrict the transferability of the method.

Abundance of built-up/impervious and low vegetation served as a proxy for the presence of slums at the settlement level. A hierarchical approach comprising the two levels of ontology, the object level and the settlement level, was used in an OOA environment. The object level represented the basic land cover features of the V-I-S model. The parameters GR, Brightness, Mode(R) and GLCM_{Ent}(R) were useful for basic land cover classification of vegetation, shadow and built-up, respectively. Roads and water were classified by using vector layers. Road classification is one of the trickiest in RS as vegetation and shadow can hide them completely. With RS offering a bird’s-eye view, we are left with conceptual problems where an object virtually has two classes, vegetation and road. The choice, then, could be driven by purpose of classification. To overcome this problem, we digitised the major/main roads in the study area. The minor/secondary roads were purposely ignored. The classification of roads in planned areas could negatively impact quantification of vegetation class. This could further influence our results as vegetation and built-up are used as a proxy for slum presence.
The overall accuracy achieved for land cover classification is 81 percent with a kappa agreement of 0.74. Whereas there have been discussions in previous research on the suitability of error matrices and kappa for measuring accuracy (Pontius & Millones 2011), there are many OOA studies that continue to use them (Herold et al. 2003; Myint et al. 2013; Belgiu et al. 2014a; Belgiu & Drăguț 2014). In the absence of a detailed and updated reference map, a set of randomly selected pixels was considered to be representative of the ground condition (Congalton 1991; Myint et al. 2013). There are mis-classifications where built-up mixes with bare soil; specifically dark-coloured roofs made of rusted corrugated iron sheets display low GLCM Ent(R) similar to rough and porous bare soil. There is noticeable mixing between the shadow and vegetation class, with shadow being classified with minimum accuracy. Due to the sun angle at the time of acquisition of the image, some smaller trees seem to hide under the shadow of higher trees with large canopies, resulting in low brightness values, and hence being classified as shadows. The use of a near-infrared (NIR) band would have allowed a better discrimination of vegetation from other classes. The purpose of our study using only RGB data is to show what can be achieved when NIR data are not available. This is a common situation in many cities that, if using satellite images at all, restrict themselves to purchasing RGB data because they are primarily interested in the image as a backdrop or for data extraction via interpretation. In addition many cities do not have sufficient in-house remote sensing knowledge to fully understand the analytical implications of having access to all bands. As RGB data are commonly available within city authorities it is relevant to examine whether OOA approaches in combination with a reasonably simple V-I-S model can also generate acceptable results (slum maps in this case) as this would mean that the approach is also widely applicable.

The settlement level provided boundaries for the quantification of land cover features. Automatic delineation of quantifiable segments for appropriate spatial metrics analyses is challenging considering the heterogeneity of urban areas. Previous studies (Herold et al. 2002, 2003) used manually delineated homogeneous urban regions, also referred to as homogeneous urban patches (HUP) or enumeration units (Weeks et al. 2007), for determining land use using spatial metrics analyses. A recent publication automated generation of HUPs at settlement level by using a large SP value for segmentation (Kuffer et al. 2014). We used the same approach by using a \( SP = 300 \) value to generate settlement primitives representing parts of slum settlements and mix of different land cover features in planned areas. We refrain from using the term HUP for settlement primitives as they do not entirely comply with the ‘rules for HUPs’ presented in Herold et al. (2002). In addition to the three bands of the Quickbird image, we used the texture \( GLCM_{con}(B) \) to segment settlement primitives. \( GLCM_{con}(B) \) has been identified as a stable parameter for slum classification (Kohli et al. 2013). A large value of \( SP \) helps to generate large segments representing slum settlements, but it also poses a challenge as under-segmentation may occur. This is likely to be the case if non-slum buildings on the fringes are included in a segmented slum due to spectral similarity. Use of texture largely solved this problem by restricting the segmentation, as most slums in Pune comprise highly compact buildings and hence low variance within segments (Figure 4). After experimenting with different scale parameters visually, a value of \( SP = 300 \) was found appropriate for spatial metrics analyses. Subsequently, Area, expressed in terms of number of pixels and spatial metrics, was used to classify slum settlements. The parameters Area, \( RA(S) \) and \( MG \) were suitable for eliminating false positives. \( D(SL) \) and \( ES \) were uniformly useful for reclassifying any misclassified slum objects embedded within the slums.
The percentage of agreement, at 60 percent, is rather low but encouraging considering the complexities involved in image-based slum identification (Kohli et al. 2013; Kuffer et al. 2014). One reason for such a low accuracy value is that remote sensing patterns and criteria for mapping used by MASHAL are different. We carried out alternative accuracy assessments using visual interpretations by two experts (E1 and E2) to account for such distortions. The resulting overall accuracies for E1 and E2 were 68 percent and 71 percent respectively. The producer’s accuracy of slum classification for both E1 and E2 was high (83 percent and 84 percent), but the user’s accuracy was much lower (60 percent and 76 percent respectively). These figures show the existence of considerable discrepancies and errors of omission. To better account for such discrepancies input from more experts could be used in order to systematically quantify the uncertainties related to visual interpretation of slums.

The use of this reasonably simple approach shows the potential for transferability of our method. Care must be taken, however, when this method is applied in a context with completely different characteristics. A suitable approach would be adapting the GSO to that particular context and modifying thresholds or parameters based on its characteristics. Addition of new parameters may also be required for improved classification. It is interesting that some of the parameters such as $GLCM_{Ent}(R)$, $GLCM_{Con}(B)$, Area and ES identified as most stable for built-up and slum classification in a previous study in another Indian city, Ahmedabad (Kohli et al. 2013), could be used in this study as well.

Previous work has also been done on two small extracts of the city of Sao Paulo where the knowledge-based system Inter-Image was applied (Novack & Kux 2010). For such small areas, higher accuracies were obtained, as upon a careful selection these have a higher uniformity: the dwellings were raised in approximately the same period and have developed along similar lines. In terms of the GSO, the Sao Paulo study is at the object level (Level 1). Our study builds upon this work by providing a methodology to consider a larger area that contains more variation in history and ground conditions. This unavoidably leads to a decrease in accuracy. A possible way to overcome this reduction would be to use a hierarchical approach: splitting up the larger area into a mosaic of smaller sub-areas, classifying these and then integrating towards the settlement level. We leave this for future work.

Slums are not just physical entities; they also have administrative and political definitions and local perceptions attached to them. Thus, it can be difficult to identify them solely by visual interpretation, particularly if one lacks local knowledge. In the process of developing a semi-automated slum identification process, we found that ontologies can help to bridge the gap between technology and the real world by integrating that expert (local) knowledge about the context. It also helped to systematise the process by considering different spatial levels where a set of indicators may be chosen depending on context for an OOA. In Pune, slums generally have a distinct appearance compared to planned residential complexes but there were a few exceptions, e.g. the Old City core, which exhibited similar characteristics to slums. Pune and many other cities in India have a historic old centre or core. If these city cores have dilapidated over time, they are likely to exhibit similar physical characteristics to many slums and can therefore be readily misclassified as such (Figure 6). During our field-visit to Pune, we observed that some of these buildings had already collapsed and many need upgrading in order to be considered as durable houses: in terms of the UN-HABITAT definition these are slum buildings but they are not so numerous and concentrated as to form a slum settlement. Though shadow was useful in removing false positives from the Old City, the inclusion of some buildings negatively impacts our accuracy statistics, as these are
excluded from the reference layer, which is based upon a different political-administrative definition of slum (Figure 7). This problem exists in Pune and is known to also exist in other Indian cities. It is also quite conceivable that this issue will arise in other contexts as well.

Additional information in terms of surface and terrain elevation conditions may help in improving the results: many slums are located on steep slopes and surface height information can be used to differentiate slum buildings from formal, high-rise buildings. Pune's slum buildings generally tend to be made of less-durable materials and thus have lower heights. Other than the false positives in the Old City, the methodology works well in the other parts of the image, especially for highly dense slums showing a clear contrast to other built-up areas, though there are false positives from market areas because of their relatively large and highly dense roofs contributing to the high roof coverage. The larger slums with a relatively uniform texture are generally identified correctly, while some small slums are often missed because at settlement level they merge with other land cover features.

Finally, the low agreement between the reference and slum classification can also be attributed to the uncertainties in the reference data themselves. The reference data are prepared by manual image interpretation, which may contain uncertainty and hence negatively impact the accuracy results (Figure 7). The expertise of the interpreter in this field, RS skills and the time invested may have led to discrepancies in interpretation (Van Coillie et al. 2014). In the case of slums, different conceptualisation, fuzzy boundaries, poor definitions and variability could add to uncertainties in delineations. Inaccuracies in classification can thus be due to errors in the reference data (Albrecht et al. 2010) and should be understood. This forms the topic for our ongoing investigation where we are studying the uncertainties attached to the existence/presence and spatial extent of slums.

6. Conclusions

This study presented a method for slum detection from VHR images based on the morphology of the built environment as expressed in generic slum ontology (GSO). It consists of segmentation followed by hierarchical classification using an object-oriented image analysis (OOA). It used texture in combination with spatial metrics based on the V-I-S model and was applied to a Quickbird image of the city of Pune, India. It turned out to be a promising method for slum identification. Contextual knowledge in the form of a local slum ontology guided the classification process. Our analysis identified that textural features derived from a grey-level co-occurrence matrix (GLCM), contrast (GLCM$_{con}$) in addition to three bands of an image were appropriate for segmentation at settlement level. A clear difference in classification accuracy was revealed between the two ontological levels. The overall accuracy at Level 1 of 80.8 percent was remarkably high and adequate as an input to the GSO considering the limited spectral resolution of the Quickbird image and also the absence of the near-infrared (NIR) band. The percentage of agreement between the reference layer and the slum classification at the settlement level was 60 percent. This value is promising considering that a rather simple method was applied for slum identification, consisting of the size of image segments (Area) and the proportions of built-up and vegetation. The use of soil was not useful for the classification in this particular context, but it may be relevant in other contexts. The parameters relative area of shadow (RA(S)), mean green (MG), distance to classified slums (D(SL)) and enclosed by slums (ES) were useful for the final classification. Our results are potentially useful for city managers especially in cities where no alternative data are available.
to get updates regarding slums on the ground. Further study of the characteristics of slums in terms of patterns, size, density and their manifestations in the images may be useful. In such future research, relationships with other classes derived at the environs level of ontology are likely to be used to refine results.

Acknowledgements

We would like to thank MASHAL, Pune for providing data for this research. We would also like to thank Dr. Sulochana Sekhar for her valuable advice and assistance in field-work.

Notes on contributors

Shekhar is currently an associate professor and head of the Department of Geography, School of Earth Sciences at the Central University of Karnataka, Gulbarga, India. She is an urban geographer concerned with urban problems in India. She is keen to apply remote sensing and GIS technologies to understanding urban issues in India and to finding practical solutions to them.

References

Aguilar, M., Saldaña, M. and Aguilar, F., 2013. GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments. *International Journal of Remote Sensing*, 34 (7), 2583–2606.

Albrecht, F., Lang, S. and Hölbling, D., 2010. Spatial accuracy assessment of object boundaries for object-based image analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38 (4), C7.

Baatz, M. and Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: J. Strobl, ed. *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium*. Heidelberg: Herbert Wichmann Verlag, 12–23.

Baltzavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58 (3–4), 129–151.

Belgiu, M. and Drăgut, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 96, 67–75.

Belgiu, M., Drăgut, L. and Strobl, J., 2014a. Quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using WorldView-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 205–215.

Belgiu, M., Hofer, B. and Hofmann, P., 2014b. Coupling formalized knowledge bases with object-based image analysis. *Remote Sensing Letters*, 5 (6), 530–538.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58 (3–4), 239–258.

Cleve, C., Kelly, M., Kearns, F.R. and Moritz, M., 2008. Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. *Computers, Environment and Urban Systems*, 32 (4), 317–326.

COI, 2001. *Census of India* [online]. Office of the Registrar General and Census Commissioner, Government of India. Available from: [http://www.censusindia.gov.in/](http://www.censusindia.gov.in/) [Accessed 24 June 2011].

Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37 (1), 35–46.
Drăguț, L., Tiede, D. and Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24 (6), 859–871.

Ebert, A., Kerle, N. and Stein, A., 2009. Urban social vulnerability assessment with physical proxies and spatial metrics derived from air- and spaceborne imagery and GIS data. *Natural Hazards*, 48 (2), 275–294.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2–3), 130–137.

Herold, M., Scepan, J. and Clarke, K.C., 2002. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. *Environment and Planning A*, 34 (8), 1443–1458.

Herold, M., Liu, X. and Clarke, K.C., 2003. Spatial Metrics and Image Texture for Mapping Urban Land Use. *Photogrammetric Engineering & Remote Sensing*, 69 (9), 991–1001.

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A. and Schmullius, C., 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, 112 (5), 2538–2556.

Hofmann, P., 2001. Detecting informal settlements from Ikonos image data using methods of object oriented image analysis-an example from Cape Town (South Africa) In: Jürgens, C. ed. *Remote sensing of urban areas*. 41–42.

Hofmann, P., Strobl, J., Blaschke, T. and Kux, H., 2008. Detecting informal settlements from Quickbird data in Rio De Janeiro using an object based approach. In: T. Blaschke, S. Lang and G.J. Hay, eds. *Object-Based Image Analysis*. Berlin: Springer, 531–553.

Hofmann, P., Blaschke, T. and Strobl, J., 2011. Quantifying the robustness of fuzzy rule sets in object-based image analysis. *International Journal of Remote Sensing*, 32 (22), 7359–7381.

Hu, X. and Weng, Q., 2010. Impervious surface area extraction from IKONOS imagery using an object-based fuzzy method. *Geocarto International*, 26 (1), 3–20.

Kit, O., Lüdeke, M. and Reckien, D., 2012. Texture-based identification of urban slums in Hyderabad, India using remote sensing data. *Applied Geography*, 32 (2), 660–667.

Kohli, D., Sliuzas, R., Kerle, N. and Stein, A., 2012. An ontology of slums for image-based classification. *Computers, Environment and Urban Systems*, 36 (2), 154–163.

Kohli, D., Warwadekar, P., Kerle, N., Sliuzas, R. and Stein, A., 2013. Transferability of Object-Oriented Image Analysis Methods for Slum Identification. *Remote Sensing*, 5 (9), 4209–4228.

Kuffer, M., Barros, J. and Sliuzas, R.V., 2014. The development of a morphological unplanned settlement index using very-high-resolution (VHR) imagery. *Computers, Environment and Urban Systems*, 48, 138–152.

MASHAL, 2011. *Pune city slum atlas*. Pune: Maharashtra Social Housing and Action League.

Myint, S.W., Galletti, C.S., Kaplan, S. and Kim, W.K., 2013. Object vs. pixel: a systematic evaluation in urban environments. *Geocarto International*, 28 (7), 657–678.

Neuwirth, R., 2005. *Shadow cities: a billion squatters, a new urban world*. New York, NY: Routledge.

Novack, T. and Kux, H.J.H., 2010. Urban land cover and land use classification of an informal settlement area using the open-source knowledge-based system InterIMAGE. *Journal of Spatial Science*, 55 (1), 23–41.

Patino, J.E. and Duque, J.C., 2013. A review of regional science applications of satellite remote sensing in urban settings. *Computers, Environment and Urban Systems*, 37, 1–17.

PMC, 2012. *Integrated ward level disaster management plan*. Pune: Pune Municipal corporation.

Pontius, R.G. and Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32 (15), 4407–4429.

Ridd, M.K., 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing*, 16 (12), 2165–2185.

Scepan, J., 1999. Thematic validation of high-resolution Global Land-Cover Data Sets. *Photogrammetric engineering and remote sensing*, 65 (9), 1051–1060.

Shekhar, S., 2010. *Applying GIS and RS for modeling the growth of slums in Pune City, Maharashtra, India*. Thesis (M.Sc.). University of Salzburg.
Shekhar, S., 2012. Detecting slums from QuickBird data in Pune using an object-oriented approach. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39 (8), 519–524.

Sliuzas, R., Mboup, G. and Sherbinin, A., 2008. Expert group meeting on slum identification and mapping. Enschede, The Netherlands: ITC.

Stoler, J., Daniels, D., Weeks, J.R., Stow, D.A., Coulter, L.L. and Finch, B.K., 2012. Assessing the utility of satellite imagery with differing spatial resolutions for deriving proxy measures of slum presence in Accra, Ghana. *GIScience & remote sensing*, 49 (1), 31–52.

Stow, D., Lopez, A., Lippitt, C., Hinton, S. and Weeks, J., 2007. Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data. *International Journal of Remote Sensing*, 28 (22), 5167–5173.

Stow, D.A., Lippitt, C.D. and Weeks, J.R., 2010. Geographic object-based delineation of neighborhoods of Accra, Ghana using QuickBird satellite imagery. *Photogrammetric Engineering & Remote Sensing*, 76 (8), 907–914.

Taubenböck, H. and Kraff, N.J., 2014. The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data. *Journal of Housing and the Built Environment*, 29 (1), 15–38.

Trimble, 2014. Reference Book. *eCognition Developer 9.0*. München, Germany.

UN-HABITAT, 2006. *Analytical perspective of pro-poor slum upgrading frameworks*. Nairobi, Kenya: United Nations Human Settlements Programme.

UN-HABITAT, 2010. *State of World’s Cities 2010/2011*. UK and USA: United Nations Human Settlements Programme.

UN-HABITAT, 2014. *Streets as tools for urban transformation in slums: A street-led approach to citywide slum upgrading*. Nairobi, Kenya: United Nations Human Settlements Programme.

Van Coillie, F.M.B., Gardin, S., Anseel, F., Duyck, W., Verbeke, L.P.C. and De Wulf, R.R., 2014. Variability of operator performance in remote-sensing image interpretation: the importance of human and external factors. *International Journal of Remote Sensing*, 35 (2), 754–778.

Weeks, J., Hill, A., Stow, D., Getis, A. and Fugate, D., 2007. Can we spot a neighborhood from the air? Defining neighborhood structure in Accra, Ghana. *GeoJournal*, 69 (1-2), 9–22.