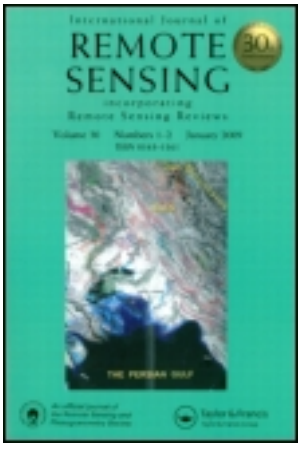

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

Quantifying the robustness of fuzzy rule sets in object-based image analysis

Peter Hofmann [a], Thomas Blaschke [b] & Josef Strobl [a]

[a] Austrian Academy of Sciences GIScience Institute, Schillerstr. 30, A-5020, Salzburg, Austria

[b] Centre for Geoinformatics, Salzburg University, Hellbrunnerstrasse 34, A-5020, Salzburg, Austria

Available online: 19 Sep 2011

To cite this article: Peter Hofmann, Thomas Blaschke & Josef Strobl (2011): Quantifying the robustness of fuzzy rule sets in object-based image analysis, International Journal of Remote Sensing, 32:22, 7359-7381

To link to this article: http://dx.doi.org/10.1080/01431161.2010.523727

International Journal of Remote Sensing
Vol. 32, No. 22, 20 November 2011, 7359–7381

Quantifying the robustness of fuzzy rule sets in object-based image analysis

PETER HOFMANN*†, THOMAS BLASCHKE‡ and JOSEF STROBL†
†Austrian Academy of Sciences GIScience Institute, Schillerstr. 30, A-5020 Salzburg, Austria
‡Centre for Geoinformatics, Salzburg University, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria

(*Received 15 May 2009; in final form 28 July 2010*)

Object-based image analysis (OBIA) has become very popular since the turn of the century. For high-resolution situations, in particular, where the objects of interest are larger than pixels, methods have been developed that build on image segmentation and on the further classification of objects rather than on pixels. Many studies have shown that OBIA methods are, in principle, more transferable and reapplicable to other images. To obtain comparable results by reapplying a given rule set on (slightly) changed conditions, the rule set must either be able to adapt to the changed conditions or it must be parameterized for manual adaptation. In this context, a rule set can be seen as the more robust the less it has to be changed, and vice versa. In this article we introduce a new method to evaluate the robustness of a rule set. The main assumption is that the amount of necessary adaptations can be measured in conjunction with the quality of classification achieved. We demonstrate that the method introduced is able to (1) evaluate the robustness of a rule set and (2) identify crucial elements of a rule set that need to be reparameterized.

1. Introduction

Classical per-pixel image analysis tends to extract the content of an image by methods of per-pixel image processing and classification. It is often hypothesized that pixel-by-pixel classifications have limitations, especially with high-resolution imagery (Blaschke and Strobl 2001). Since the turn of the twenty-first century, more than 100 satellite sensors have been launched, including an increasing number of high-resolution sensors, with the most recent WorldView-2 satellite being launched on 9 October 2009 with a spatial resolution of 0.46 m at nadir for the panchromatic band. For high-resolution information, in particular, it is widely agreed that approaches are needed that incorporate the spatial entities and relationships of the resulting pixels. In a literature study, Blaschke (2010) identified 145 journal papers that build on a recent approach known as object-based image analysis (OBIA). Sometimes the term GEOBIA (geographic OBIA) is used when distinguishing Earth Observation (EO) methods from other imaging fields. The idea of incorporating neighbourhood information is much older. Various kernel functions and moving window applications have

*Corresponding author. Email: dr.peter-hofmann@t-online.de

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2011 Taylor & Francis
http://www.tandf.co.uk/journals
http://dx.doi.org/10.1080/01431161.2010.523727

been developed over the years (Lillesand and Kiefer 2000). OBIA (for an overview see Blaschke *et al*. 2008) goes one step further; instead of processing and classifying single pixels according to their spectral values, image objects are generated. This is typically achieved through arbitrary image segmentation methods. The resulting segments are expected to be relatively homogeneous compared to the surroundings. Burnett and Blaschke (2003) called these segments 'objects candidates', which are to be recognized by further processing steps and to be transferred into meaningful objects. In a classification process these segments serve as building blocks for further analysis, readjustment of segments and finally an optimized segmentation. Meanwhile, there is also a growing scientific literature on methodological concepts of OBIA. It is well known that semantically significant regions are found in an image at different scales of analysis (Hay *et al.* 2001, 2003), and OBIA is inextricably linked to multiscale analysis concepts (Burnett and Blaschke 2003, Benz *et al.* 2004, Hay and Castilla 2008, Lang 2008), even if single levels are targeted for specific applications (Lang and Langanke 2006, Lang 2008, Weinke *et al.* 2008). Burnett and Blaschke (2003) called this OBIA concept multiscale segmentation/object relationship modelling (MSS/ORM). Lang and Langanke (2006) developed an iterative one-level representation (OLR), and Tiede *et al*. (2008) successfully applied this OLR concept to airborne light detection and ranging (LiDAR) data for tree crown segmentation. It is widely agreed that OBIA builds on older segmentation, edge detection and classification concepts that have been used in remote sensing image analysis for several decades. Nevertheless, its emergence has provided a new crucial bridge to spatial concepts applied in multiscale landscape analysis, geographic information systems (GIS) and the synergy between image objects, their radiometric characteristics and analyses of EO data (Blaschke *et al.* 2008, preface). When analysing the content of an image, OBIA tries to simulate the workflow of human visual image interpretation, first aggregating the spectral information (of the pixels) into segments representing object primitives and then applying predefined (expert) knowledge to the objects to be detected as well as to the image data given. In this way, sensor properties together with imaging conditions and the real-world objects' properties determine the characteristics of each resulting image object. In OBIA, these image object properties encompass at least geometrical, textural and colour-statistical properties. In addition, topological and scale relationships among the image objects can be used for classification purposes (see Benz *et al.* 2004, Hay and Castilla 2006, Platt and Rapoza 2008, Su *et al.* 2008). Thus, the first and most obvious criterion for success of any OBIA approach is an appropriate image segmentation method that is able to create adequate image objects. Second, for the subsequent image analysis a well-defined rule set describing the classes of concern and their respective properties is one of the key elements in OBIA.

In this article, we concentrate on an investigation of the robustness of rule sets. We hypothesize that the demand for repeatability and transferability of image analysis and image classification is increasing. The lack of transferability of per-pixel approaches is one major factor of the commercial success of OBIA methods and software in recent years. Concerning the applicability and transferability of rule sets, there are few systematic investigations beyond software-specific and software-dependent applications (Smith 2008, Walker and Blaschke 2008). The central research question in our investigation was to measure to what degree a given rule set is capable of providing comparable results on other images. 'Other images' means either spatially different scenes under similar imaging conditions or other sensor and imaging conditions but similar land use situations compared to the reference image for which the rule set

was originally developed. At first glance this constraint may seem to hamper a wider use, but it is in compliance with the vast majority of remote sensing and image analysis strategies being developed over the past decades. It makes no sense to classify completely different biogeographical or land use situations. For instance, it does not make sense to apply classification rules developed for tropical forests to a tundra ecosystem and vice versa. It may, however, be possible to transfer rule sets between images of different sensors such as QuickBird (DigitalGlobe Corporate, Longmont, CO, USA) or IKONOS and different atmospheric and seasonal conditions. As the image objects derived from the segmentation act as the building blocks of the analysis, it is important that they must be comparable. That is, the real-world objects to be detected and/or analysed must be outlined in the images comparably with respect to the images' differences. When aiming for comparable image objects, there are a variety of potentially influencing factors with regard to the quality of a segmentation. It has been successfully demonstrated that two of the most important and most obvious factors, the spatial resolution and radiometric properties of the sensor, can be determined in advance and can be compensated respectively by adapting the segmentation parameters accordingly (see Hofmann *et al*. 2008*b*). Others, such as differing illumination conditions and atmospheric conditions, are usually difficult to predict and thus a matter for appropriate methods of image preprocessing. In this article, we focus on the robustness of rule sets. For all investigations we assume that all influencing factors that are not determined by the real-world objects themselves can be compensated when obtaining image objects of comparable quality. This includes techniques that iteratively enhance an initial segmentation result by resegmenting or subsegmenting image objects according to their spatial, spectral or topological properties or class assignments. Thus, we assume that rule sets are to be applied on image segments with best possible quality concerning the underlying image data.

2. Fuzzy rule sets and robustness

Fuzzy logic is a form of multivalued logic based on fuzzy set theory to perform reasoning that is approximate rather than precise. We refer to binary sets with hard decisions as crisp logic or crisp decisions, where the members of a set have a membership value of only 0 or 1. The term 'fuzzy logic' emerged as a consequence of the development of the theory of fuzzy sets by Zadeh (1968). In fuzzy logic, the set membership values can range between 0 and 1 (inclusively). The degree of truth of a statement can range between 0 and 1 and is not constrained to the two logical values 'true' (= 1) and 'false' (= 0) as in classical predicate logic. When it is necessary to switch between the two schemata we refer to 'defuzzification', 'defuzzyfying' or 'crisping' when resulting in crisp class membership values for a final decision.

2.1 *Fuzzy rule sets in the context of OBIA*

Fuzzy rules deal with uncertain, incomplete and/or vague information in order to steer or control processes or to assign objects to fuzzy sets. Therefore, investigating the robustness of fuzzy rules or fuzzy rule sets in general means to investigate their ability to handle unpredictable situations of the rules' or rule sets' scope of application. Although a very limited number of scientists have used fuzzy sets for the delineation and optimization of objects (Gorte 1998, Hu *et al*. 2005, Wuest and Zhan 2009), we focus on the use of fuzzy logic for the classification of image objects. We hypothesize that their major advantage lies in their ability to deal with the uncertainty and

vagueness that is inherent when classifying spatial entities to distinct, user-defined classes (Bezdek and Pal 1992, Benz 1999, Hay *et al.* 2003, Benz *et al.* 2004). In OBIA this is usually done by describing the desired classes as fuzzy sets of image objects, where each class is described by at least one membership function defining the degree of membership μ for each object depending on the objects' values for a selected property. Therefore, μ lies in the range 0.0–1.0. In this way, the membership function describes the degree of fulfilment of a given condition concerning a property that an object needs for being a member of the described class. That is, $\mu = 0.0$ means no membership to the class of concern, and $\mu = 1.0$ means the conditions for membership are completely fulfilled. In practice, to crisp a classification result, a minimum threshold for μ is defined, where objects having a degree of membership below this threshold are assumed not to be a member of the respective class. The types of membership functions typically used can be categorized into the three groups, namely fuzzy-greater-than, fuzzy-lower-than and fuzzy-range, where the shape of the functions can vary (see figure 1). As indicated in figure 1, typical parameters to describe a membership function are α, β and a . While α indicates the lower border concerning a property p, β indicates the upper border, and a is the mean of the membership function's range v concerning p, that is:

$$a = \alpha + (\beta - \alpha)/2 \tag{1}$$

$$v = \beta - \alpha \tag{2}$$

Particular classes are usually defined not only by a single property but also through several fuzzy sets, where the membership functions for each property are combined by the fuzzy-logic operators 'fuzzy-and' and 'fuzzy-or'. For the 'fuzzy-and' operator the degree of membership is usually given by the property with the minimum degree of membership, and with the maximum degree for the 'fuzzy-or' operator. This way, for class description only those feature-space spanning properties are used, which are obviously necessary (see figure 2). In consequence, the resulting degree of membership (i.e. the degree of fulfilling the classification conditions) that is assigned to each

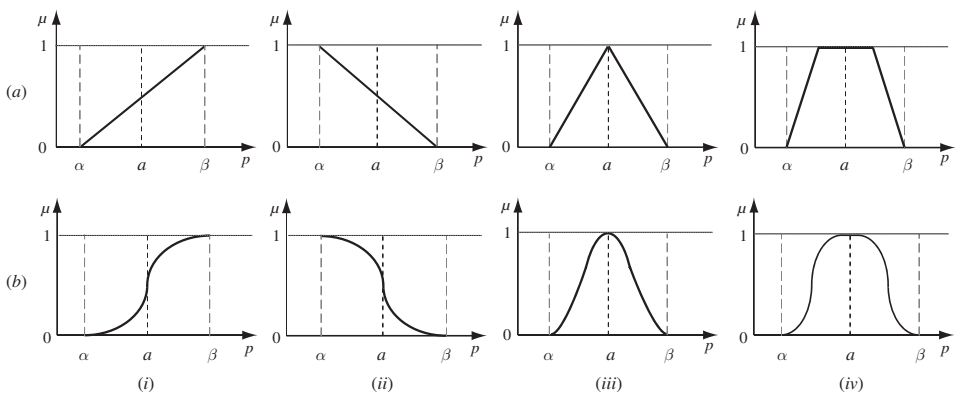

Figure 1. Typical representatives of (*a*) linear and (*b*) non-linear fuzzy membership functions with μ as the degree of membership and p as the describing property. Displayed are (*i*) fuzzy-greater-than, (*ii*) fuzzy-lower-than and (*iii*) fuzzy-range-functions with one maximum membership ($\mu(a) = 1.0$) and (*iv*) fuzzy-range-functions with a range of maximum membership.

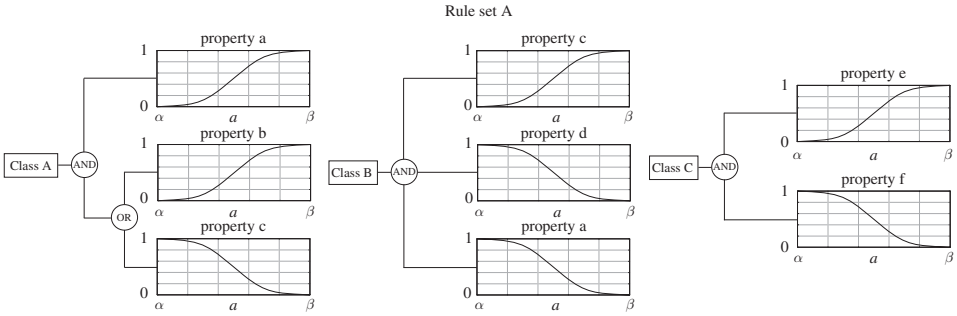

Figure 2. Example of a fuzzy rule set consisting of the classes A, B and C described by the properties a, b, c, d, e and f with different membership functions and combinations of them. For each property α, a and β are individually chosen.

object can be used to quantify the reliability of a classification (see Benz *et al.* 2004, Definiens 2004). The combination of several fuzzy sets, that is a classification scheme or a class hierarchy respectively, is called a fuzzy rule set. Applying a classification scheme described by a fuzzy rule set yields for each object a degree of membership to each class, where each object can be a fuzzy member of either no class, one class or several classes. The latter case reflects the ambiguity of the classification or the class descriptions, respectively. Although it can just reflect the reliability of a classification result, assessing the ambiguities of a fuzzy classification is nevertheless also used as a criterion of quality: the lower the ambiguity of a classification is, the higher the quality of a classification result is considered. Thus, the so-called 'classification stability' expresses for each object the difference of μ between the best and the second-best class assigned (see Benz *et al.* 2004, Definiens 2004). Nevertheless, the ambiguity of a class assignment depends on several factors; one of them is the capability of the segmentation to outline the desired objects.

2.2 *Robustness of rule sets in the context of OBIA*

The terms 'robust' and 'robustness' are used in various contexts differently but with more or less the same common understanding. For example, in software engineering, robustness expresses the ability of software to deal with errors or erroneous data. In engineering a system, a component or a process is said to be 'robust' if it functions even after faulty usage or under stressful environmental conditions (Pötzl 1996), which means beyond its specification or operating environment. In the life sciences, a system, organism or species is called 'robust' if it is tolerant against perturbation: 'robustness is a more general concept according to which a system is robust as long as it maintains functionality, even if it transits through a new steady state or if instability actually helps the system to cope with perturbations' (Kitano 2007). In communication theory, information theory and signal processing, a system or process is 'robust' if it is capable of processing perturbed information correctly. In this context, Ay and Krakauer (2007) introduced a methodology to investigate the information flow over biological networks. A comprehensive summary of the concept of 'robust' and 'robustness' is given by Jen (2003).

To define what is to be understood as 'robust' in the context of OBIA as a subdiscipline of remote sensing and image processing (Hay and Castilla 2008), we understand

the variability of the objects to be detected in different images as the perturbations for a given rule set. This means a rule set is considered to be robust if it is able to cope with the variability of the objects to be detected in varying images. For instance, a rule set designed to identify land cover objects according to the European-wide CORINE (Coordinated Information on the European Environment) land cover scheme from IKONOS images is expected to contiguously map CORINE classes from arbitrary IKONOS scenes with equal classification quality. However, the CORINE scheme is valid for the whole of Europe and for all seasons, which means the rule set must be capable of coping well with the respective regional and seasonal variations of the desired CORINE classes in the IKONOS data. We therefore understand the robustness of a rule set in OBIA for a given scope as the ability to (re)produce classification results with equal quality in different but comparable image data that have been preprocessed and segmented with comparable quality.

3. Quantifying the robustness of fuzzy rule sets

To quantify the robustness of a given rule set, it is indispensable to define what is to be understood as more and as less robust. As stated in §2, a robust rule set must be able to achieve classification results of comparable quality for image data with comparable characteristics. The quality of classification results in remote sensing is typically estimated through a (random) comparison of the results with a reliable mapping of the classified site or an ad hoc manual interpretation of random points selected. For per-pixel classifications, the vast majority of methods can be characterized as the random point method described by Story and Congalton (1986); for an overview see Congalton and Green (1999). Only more recently have a significant number of scientists encountered problems when applying the same methods to the resulting objects (Zhang *et al.* 2005, Albrecht 2008, Grenier *et al.* 2008, Platt and Rapoza 2008). Objects are technically polygons or, when addressed in a raster domain, regions. At present, a variety of classification quality measures exist that indicate the conformance and non-conformance of the classification results with the reference mapping (ground truth), but this is a developing field (Blaschke 2010). In principle, this conformance can be expressed either class-wise or globally. It is usually expressed within an interval of 0.0 (non-conformance) to 1.0 (absolute conformance) (see Congalton and Green 1999, Lillesand and Kiefer 2000, ISO/TC 211 2003, Benz *et al.* 2004, Definiens 2004). Schiewe and Gähler (2008) proposed the so-called Fuzzy Certainty Measure (FCM), which takes uncertainties in the reference data and the classification results into account respectively.

In the following section we assume that the quality of a classification result to be defined globally is given by an arbitrary quality criterion q with a value of $0.0 \leq q \leq 1.0$. To quantify the robustness r of a rule set R we consider the following: let $\mathrm{R_r}$ be the initial rule set that has been developed on the reference image $\mathrm{I_r}$ and applied on $\mathrm{I_r}$ with quality q_r. When reapplying $\mathrm{R_r}$ to a number of comparable images I_i (with $1 \leq i \leq \mathrm{n}$), we obtain for each image a classification result with quality q_i. Each of them is either different from or equal to the quality q_r achieved in $\mathrm{I_r}$. We can then consider three general scenarios for evaluation:

Scenario (*a*): The initial rule set $\mathrm{R_r}$ is not adapted because the quality q_i of the classification result in I_i is at least as good as in the reference image; that is, $\mathrm{R_r} = \mathrm{R}_i$ and $q_i \geq q_\mathrm{r}$.

Scenario (*b*): The initial rule set $\mathrm{R_r}$ is going to be adapted until the quality of the classification result in I_i is at least as good as in the reference image; that is,

before adaptation we have $R_r = R_i$ and $q_i < q_r$ and after adaptation we have $R_r \neq R_i$ but $q_i \geq q_r$.

Scenario (*c*): The initial rule set R_r has been changed to R_i but the quality of the reference classification could not be achieved; that is, $R_r \neq R_i$ and $q_i < q_r$.

3.1 *Quantifying deviations in quality of classification results*

Evaluating the robustness r_i of an unchanged rule set that has been applied on image I_i (i.e. $R = R_r = R_i$), the robustness concerning this image can be expressed by the ratio of the quality values:

$$r_i = \frac{q_i}{q_r}, \tag{3}$$

for $q_r > 0.0$, where the greater r_i is the more robust the rule set is concerning image I_i. This is equivalent to Kitano's evaluation function for biological robustness. Moreover, for $r_i > 1.0$, a better result in I_i can be achieved and vice versa for $r_i < 1.0$. For all the images I_n under investigation, the mean robustness can be expressed by:

$$r = \frac{1}{n}\sum_{i=1}^{n} r_i, \tag{4}$$

which means the greater r the more robust R is for $q_r > 0$. Furthermore, this method can be applied to measure r if the inequality for q in scenario (*a*) is not fulfilled.

3.2 *Quantifying the deviation of adapted rule sets*

Concerning scenarios (*b*) and (*c*), the original rule set R_r has been changed for image I_i either until $q_i \geq q_r$ could be achieved or not. Therefore, a new rule set R_i was created with $R_i \neq R_r$. By detecting all differences between R_r and R_i, that is by summing all adaptations performed, we obtain a measure for the deviation d of R_i from R_r. As we are focusing on fuzzy rule sets we have to consider the following potential types of adaptation:

Type C. Adding or removing (or deactivating) a class.

Type O. Change of the fuzzy-logic connection of membership functions; that is, switch a fuzzy-and operator to a fuzzy-or operator and vice versa.

Type F. Adding, removing (or deactivating) or changing an already existing fuzzy membership function.

Consequently, we can write for d:

$$d = \sum_{i=1}^{c} C_i + \sum_{i=1}^{o} O_i + \sum_{i=1}^{f} F_i, \tag{5}$$

where c indicates the number of all adaptations of type C, o indicates those of type O and f those of type F; C_i is the ith adaptation of type C, and similarly for the other types. Adaptations of type F can be differentiated further for quantification purposes:

to reduce the potential complexity we assume that membership functions describing $\mu(p)$ are only used once per class. We can then subdivide the changes of type F into two further types:

Type F^a: Adding, removing (or deactivating) a membership function or changing the type of a membership function from:

fuzzy-greater to fuzzy-lower and vice versa
fuzzy-range to fuzzy-greater or fuzzy-lower and vice versa
linear to non-linear shape and vice versa

Type F^b: Changing the range v of a membership function; that is the type and shape of the function remain unchanged but the function is shifted and/or stretched or compressed along the p-axis (see figure 3).

For the investigation of the robustness of a rule set, the quality of F^a and F^b is different: changes of type F^a involve changing the semantics of a membership function with respect to the class description by either changing the feature space (adding and/or removing) or changing the type of the membership function. Changes of type F^b are just changes of the position and/or extent of a class in the feature space. They can therefore be understood as value adjustments but the principal semantics of the membership function remains as it was. Therefore, to quantify the deviation of R_i from R_r for changes of type F^a, it is enough to summarize them. However, for changes of type F^b we have to measure the amount of change concerning the parameters a and v of each membership function, that is we have to measure shifts, stretches and/or compressions of the membership functions. This can be done by either measuring their absolute change Δa and Δv or their relative change δa and δv. The latter has the advantage of being independent of the properties' differing value ranges. Hence, for each membership function we have to consider shifting-changes δa and stretching-changes δv by taking the parameters a_r and v_r in the reference rule set R_r and the adapted parameters a_i and v_i of the corresponding membership function in rule set R_i into account:

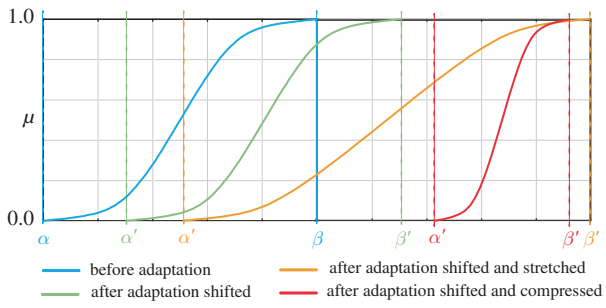

Figure 3. Some examples of changes of type F^b for a given membership function (blue). The colours indicate different kinds of change of type F^b; α', β' indicate the respective changed lower and upper border of the function after adaptation.

$$\delta a = \begin{cases} \left(1 - \dfrac{a_i}{a_\mathrm{r}}\right) & \text{for } a_i > a_\mathrm{r} \text{ (positive shift)} \\ 0 & \text{for } a_i = a_\mathrm{r} \text{ (no shift)} \\ \left(1 - \dfrac{a_\mathrm{r}}{a_i}\right) & \text{for } a_i < a_\mathrm{r} \text{ (negative shift)} \end{cases} \quad \text{and} \quad \delta v = \begin{cases} \left(1 - \dfrac{v_i}{v_\mathrm{r}}\right) & \text{for } v_i > v_\mathrm{r} \text{ (stretch)} \\ 0 & \text{for } v_i = v_\mathrm{r} \text{ (linear shift)} \\ \left(1 - \dfrac{v_\mathrm{r}}{v_i}\right) & \text{for } v_i < v_\mathrm{r} \text{ (compression)} \end{cases} \tag{6}$$

with $a_\mathrm{r} \neq 0$ for positive shifts and $a_i \neq 0$ for negative shifts. For the case $v_i = 0$ and/or $v_\mathrm{r} = 0$ a change of type F^a (changing the type of a membership function) is given because there was or is no membership value range of $0.0 \leq \mu(p) \leq 1.0$ given for p; that is, the membership function of the reference rule set has been changed to a crisp single value rule ($v_\mathrm{r} \neq 0$ and $v_i = 0$) or a crisp single value has been changed to a membership function ($v_\mathrm{r} = 0$ and $v_i \neq 0$). The relative change δF of a fuzzy membership function F is then given by:

$$\delta F = \delta a + \delta v \tag{7}$$

Finally, we can measure all deviations of rule set R_i from R_r by summing them:

$$d = \sum_{i=1}^{c} C_i + \sum_{i=1}^{o} \mathrm{O}_i + \sum_{i=1}^{f_\mathrm{a}} \mathrm{F}^a{}_i + \sum_{i=1}^{f_\mathrm{b}} \delta F_i \tag{8}$$

where f_a indicates the number of changes of type F^a (adding or deleting a membership function) and f_b indicates the number of relative changes of type F^b (shifting and/or stretching along the p-axis) and δF_i is the ith relative change of type F^b, that is of a fuzzy membership function as defined in equation (7). For practical applications, if $a_i < 0$ and $a_\mathrm{r} < 0$ is given, δa should be calculated inversely. Furthermore, for simple changes of signs, that is $a_i = -a_\mathrm{r}$, δa should be set to 1. Additionally, to avoid eliminations when calculating d, each δF should be calculated by adding δv and the absolute value of δa.

3.3 *Quantifying the robustness by rule set deviations and quality deviations*

To evaluate situations as described in scenario (c), that is rule set R_r has been changed to rule set R_i with rule set deviation d_i and the quality q_i of the classification of image I_i could not achieve the respective quality of image I_r, the robustness r_i concerning image I_i can be described by:

$$r_i = \frac{q_i / q_\mathrm{r}}{d_i + 1} \tag{9}$$

and therefore the mean robustness r for all images investigated can be determined as described in equation (4). It is clear that, when comparing equation (9) with equation (3), if rule set R_i was not adapted, that is $\mathrm{R}_i = \mathrm{R}_\mathrm{r}$, d_i equals 0.0 and hence equation (9) is identical to equation (3), which means the robustness is only expressed by the ratio

of the classification qualities achieved. In addition, for the situation where the classification result in I_i was better than in I_r and the rule set was not changed, r_i is greater than 1.0. Vice versa, the closer r_i is to 0.0, the less robust R_i could have been applied on I_i because either the deviation d_i between the reference rule set and the adapted rule set was very high, or the quality of the reference image could hardly be achieved ($q_i << q_r$). Further information that is derived automatically with equation (9) is the relationship between effort and benefit, that is between the amount of adaptation of R and the respective raising of quality: assuming the case $q_i >> q_r$. That is, in the reference image I_r, only a poor quality compared to I_i could have been achieved but therefore the rule set R_i is very different to R_r. The amount of deviations d_i between R_i and R_r is then very high and therefore r_i is still low. On the contrary, for the case when the quality could have been raised dramatically with little effort, d_i remains very low and r_i rises.

3.4 *Extensions by weighting rule set deviations*

For some investigations the types of necessary adaptations might be of different relevance in order to evaluate the robustness of a rule set; for example, removals or additions of classes (type C) change the semantics of a rule set more than changing the shapes of membership functions (type F^b). The types of adaptation can be weighted respectively by individual weights w_c, w_o, wf_a and wf_b, that is equation (8) can be extended to:

$$d = w_c \sum_{i=1}^{c} C_i + w_o \sum_{i=1}^{o} O_i + wf_a \sum_{i=1}^{f_a} F^a{}_i + wf_b \sum_{i=1}^{f_b} \delta F_i \tag{10}$$

In some cases individual adaptations are more important than others; for example, the addition or removal of class C_1 changes the semantics of a rule set more than that of another class C_2 or C_3. Respectively, weighting single adaptations by individual weights w_i extends equation (10) to:

$$d = w_c \sum_{i=1}^{c} w_i C_i + w_o \sum_{i=1}^{o} w_i O_i + wf_a \sum_{i=1}^{f_a} w_i F^a{}_i + wf_b \sum_{i=1}^{f_b} w_i \delta F_i \tag{11}$$

4. An approach to evaluating the robustness of fuzzy rule sets

Evaluating the robustness of a fuzzy rule set means at the first stage identifying and, if possible, quantifying the deviation between a reference rule set R_r and an adapted rule set R_i with respect to the deviation in classification quality. By doing so, for all investigated images I_n, we obtain a mean robustness as described in equations (9) and (4). In the second stage, it is interesting to know the reasons for the observed deviations; for example, why a certain class had to be added or why certain membership functions had to be added, removed or changed. Evaluating these deviations yields two different results: (*a*) the variability of the scope, for example, the variability of land cover classes over time and region, or (*b*) revealing principal mistakes in the design of the rule set in the assumptions made for the desired classes. For the latter case, a tabular analysis of all deviations, especially those of type F^a and F^b, can subsequently lead to a specific enhancement of the rule set design.

To prove this approach, a rule set was designed to identify and differentiate the three types of form 'triangle', 'circle' and 'square' in appropriately segmented artificial

images, as illustrated in figure 4. Then, a rule set R_r is developed that assumes that the three classes can be differentiated by their colour. This rule set uses the reference image I_r as shown in figure 4, consisting of three classes, namely 'triangle', 'circle' and 'square'. Each class is described using the colour fraction of each RGB band. The colour fraction indicates the relative brightness of a respective band against the mean brightness of an object. In some software packages, such as Definiens eCognition (Definiens AG, Munich, Germany), the colour fraction is expressed by the so-called ratio of a band. The respective fuzzy membership functions for the properties 'ratio red', 'ratio green' and 'ratio blue' are described in figure 5.

When classifying I_r with R_r we obtain a classification result of quality $q_r = 1.0$. To investigate the robustness of R we now apply R_r to another comparably segmented image as outlined in figure 6. We apply $R_r = R_1$ on I_1 and observe a classification quality of $q_1 = 0.33$. Moreover, we observe that only the class 'circle' was correctly classified. Hence, we adjust R_r to create R_2 by adjusting the colour parameters until $q_2 = q_r = 1.0$. The necessary adjustments are listed in table 1.

Figure 4. Reference image used to create a rule set to identify objects of type 'triangle', 'circle' and 'square'.

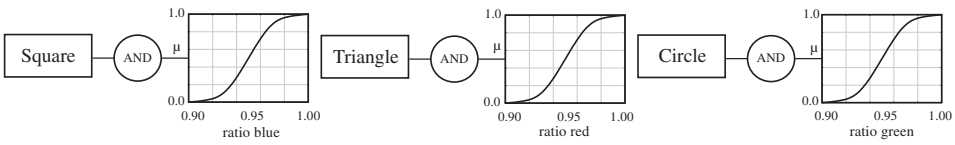

Figure 5. Reference rule set to identify object classes as shown in figure 4.

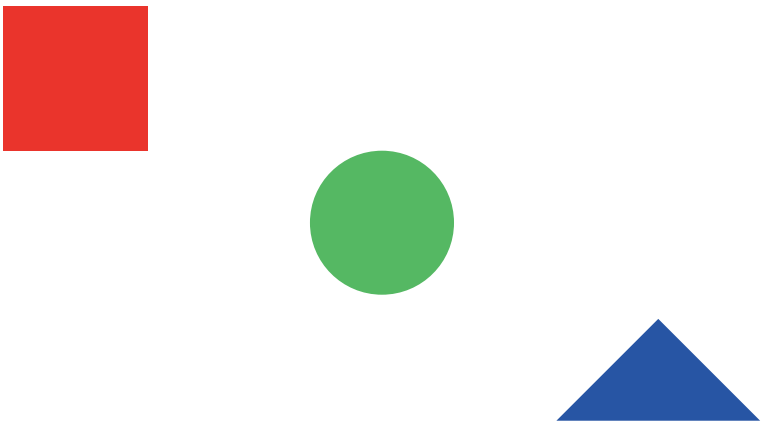

Figure 6. Comparable image I_i with comparable segmentation to figure 4.

Table 1. Listing of amount and types of adjustments to create rule set R_2.

Type C	Type O	Type F[a]	Type F[b]
–	–	Delete 'ratio red' in class 'triangle' Add 'ratio blue' in class 'triangle' Delete 'ratio blue' in class 'square' Add 'ratio red' in class 'square'	–
$\Sigma = 0$	$\Sigma = 0$	$\Sigma = 4$	$\Sigma = 0$

Before adapting the rule set, its robustness concerning image I_1 was at $r_1 = q_1 = 0.33$. After adaptation the quality was at $q_2 = 1.0$. This was obtained by the adaptations depicted in table 1. We can now write for $d_1 = 4$ and for $r_2 = 1/5 = 0.2$. However, it is easy to comprehend that for situations as outlined in figure 7, R_r and R_2 will fail and a new rule set R_3 has to be developed, taking shape criteria rather than colour criteria into account to distinguish the desired classes. In the present case

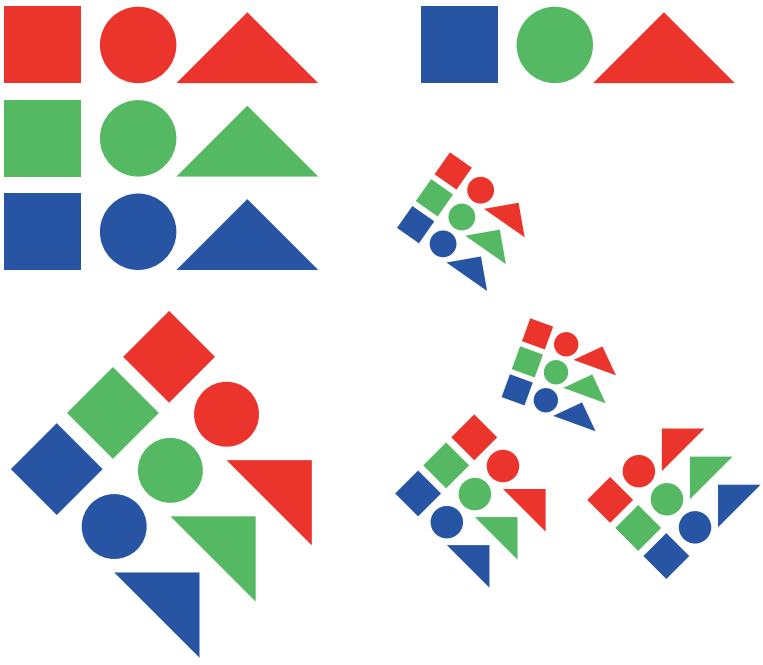

Figure 7. Comparable image I_2 with comparable segmentation to I_r.

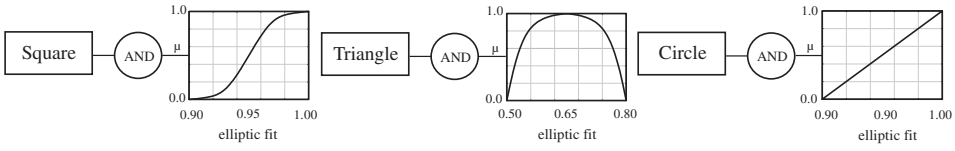

Figure 8. Adjusted rule set R_3 describing the desired classes just by shape criteria.

Table 2. Listing of amount and types of adjustments to create rule set R_3.

Type C	Type O	Type F[a]	Type F[b]
–	–	Delete 'ratio red' in class 'triangle' Delete 'ratio blue' in class 'square' Delete 'ratio green' in class 'circle' Add 'elliptic fit' in class 'triangle' Add 'elliptic' fit in class 'circle' Add 'elliptic fit' in class 'square'	–
$\Sigma = 0$	$\Sigma = 0$	$\Sigma = 6$	$\Sigma = 0$

we have deleted all ratio properties and added for each class an appropriate membership function for the shape property 'elliptic fit' (see Definiens 2004), which leads to a deviation $d_2 = 6$ but to a classification quality of $q_3 = 1.0$ (figure 8 and table 2).

If we now apply R_3 on the images I_1 and I_r we observe for all images a classification result with quality $q_3 = q_4 = q_5 = 1.0$ without any adjustments concerning I_2 and I_r; that is $d_3 = d_4 = 0.0$ and therefore $r_3 = r_4 = r = 1.0$. That is, rule set R_3 seems to be most robust for the scope of detecting triangles, squares and circles in the artificial images used here because for all images investigated until now the mean robustness r was at $r = 1.0$. This value is greater than the mean robustness that could be achieved with R_1 and R_2: $1/2(r_1+r_2) = 1/2(0.33 + 0.2) = 0.265$.

5. Some empirical test-bed scenarios evaluating the robustness of rule sets

As indicated earlier, when developing rule sets for image analysis, there is usually little known *a priori* about the whole variability of the rule set's scope. Thus, assumptions about the target classes are made and formalized accordingly during the development stage (here through fuzzy rule sets). In this way the validity of these assumptions can be verified in every case when applying the rule set to one or more reference image(s). This is an iterative process because the rule set is permanently changed until an obvious maximum of classification quality is achieved (see Hofmann 2005, Leukert 2005). 'Transferability' and 'robustness' are then evaluated by applying the developed rule set on the image data of the rule set's scope on which it has not yet been applied.

5.1 *Detecting informal settlements from very high resolution satellite images*

In the following example a rule set is evaluated that was originally designed to identify and extract informal settlements from high resolution multispectral satellite images such as IKONOS and QuickBird. Thus, it was basically designed according to the radiometric characteristics of these sensors. The reference rule set was developed using an IKONOS scene from Cape Town, South Africa, acquired in March 2000. With respect to the findings described in Hofmann *et al*. (2008*b*), critical object properties for the class descriptions were avoided. The development strategy, design and structure of the rule set followed the approach as depicted in Hofmann *et al*. (2008*a*). Applying the reference rule set as described in table 3 on the reference image led to the results shown in figure 9, with the quality described in table 4.

The rule set was then applied without any adaptations on a QuickBird scene from Rio de Janeiro acquired in May 2002, which was segmented with respectively adapted parameters, as described in Hofmann *et al*. (2008*b*). As expected, the quality of the classification result was far below that achieved in the IKONOS scene (see table 4).

Table 3. Reference rule set to detect informal settlements developed on IKONOS scene from Cape Town.

Class	Property	Membership function	Parameters for reference rule set IKONOS Cape Town			
			α_r	β_r	v_r	a_r
Top level (Level 2)						
Settlement	Area of sub-objects (1)		42.00	45.00	3.00	43.50
	Asymmetry		0.96	0.97	0.01	0.96
	Average mean difference to neighbours of sub-objects (NIR channel) (1)		93.70	326.70	233.00	210.20
	Relative area of small shadows/dark objects and sub-objects (1)		0.01	0.01	0.00	0.01
	Relative area of vegetation sub-objects (1)		0.40	0.50	0.10	0.45
Formal settlement	Area		1800.00	1900.00	100.00	1850.00
	Not informal settlement	–	–	–	–	–
Informal settlement	Area of sub-objects (1)		39.00	41.00	2.00	40.00
	Asymmetry of sub-objects: mean (1)		0.55	0.57	0.02	0.56
	Relative area of bright small roofs/objects and sub-objects (1)		0.03	0.04	0.01	0.04
	Relative area of red roofs sub-objects (1)		0.01	0.01	0.00	0.01
	Relative area of small shadows/dark objects and sub-objects (1)		0.03	0.04	0.02	0.03
	Relative area of vegetation sub-objects (1)		–	–	–	–
Base level (Level 1)						
Small shadows/dark objects	Area		35.00	40.00	5.00	37.50
	Ratio blue		0.20	0.33	0.13	0.26
Vegetation	Ratio NIR		0.29	0.30	0.01	0.30
Red roofs	Ratio red/ratio green		1.09	1.10	0.01	1.10
Bright small roofs/objects	Area		40.00	60.00	20.00	50.00
	Brightness		750.00	775.00	25.00	762.50
	Shape index		–	–	–	–

Note: Ratio NIR describes the object-related relative brightness of the near infrared (NIR) channel against the mean brightness.

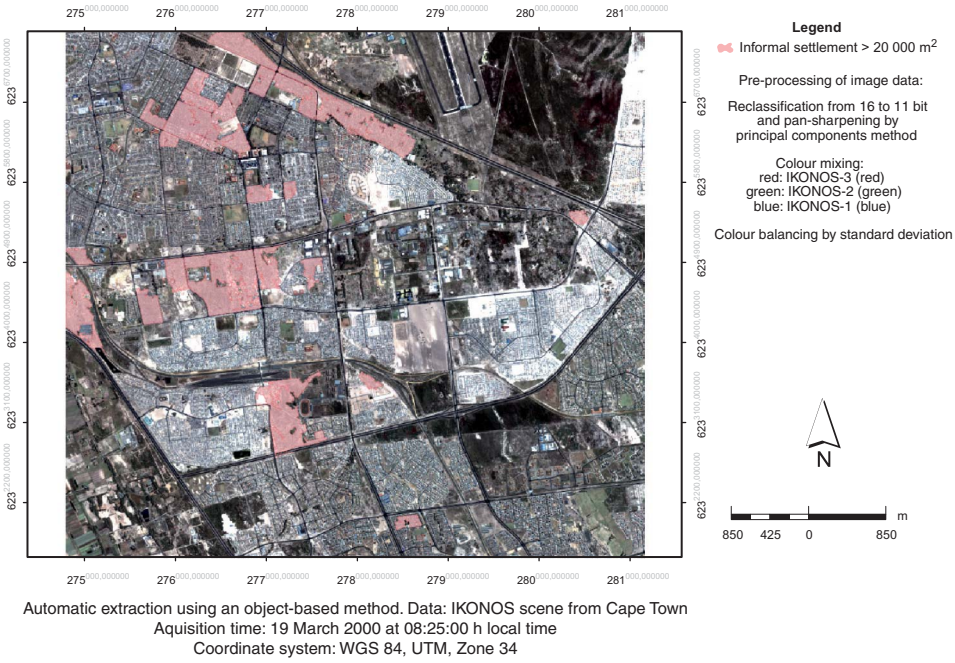

Figure 9. Classification result for reference rule set applied on reference image: informal settlements in Nyanga/Crossroads (Cape Flats, Cape Town, South Africa).

Table 4. Classification qualities achieved with reference rule set applied on reference image (IKONOS) and QuickBird scene with adapted segmentation and adapted rule set.

			Mean classification stability for 'informal settlement'		
		Overall accuracy	True positives	False positives	True positives minus false positives
IKONOS	Reference rule set	0.80	0.92	0.84	0.08
QuickBird with adapted segmentation	Without adapted rule set	0.00	–	–	–
	With adapted rule set	0.68	0.94	0.89	0.05

Consequently, the rule set was adapted to the situation in the QuickBird scene as outlined in table 5 until a maximum quality for the classification was accomplished, leading to the results depicted in figure 10 and table 4. The adaptations performed are documented in table 5. According to equation (8), a deviation of $d = 0 + 0 + 2 + 56.05 = 58.05$ was measured. In conjunction with table 4 this leads to a robustness of $r = 0.014$ for the quality criterion overall accuracy (OA) derived from an error matrix as described in Lillesand and Kiefer (2000) and $r = 0.011$ for the quality criterion 'difference of classification stability between true positives and false positives' (see table 4). Concerning the 'classification stability' of the true positives only, $r = 0.017$.

Table 5. Adaptations of reference rule set and deviations between reference rule set and adapted rule set to detect informal settlements.

Class	Property	Membership function	Parameters for QuickBird Rio after adaptation				Deviations		
			α_1	β_1	v_1	a_1	δv	δa	δF
Top level (Level 2)									
Settlement	Area of sub-objects (1)		50.00	53.00	3.00	51.50	0.00	0.18	0.18
	Asymmetry		0.96	0.97	0.01	0.96	0.00	0.00	0.00
	Average mean difference to neighbours of sub-objects (NIR channel) (1)		120.00	350.00	230.00	235.00	0.01	0.12	0.13
	Relative area of small shadows/dark objects and sub-objects (1)		0.01	0.01	0.00	0.01	0.00	0.00	0.00
	Relative area of vegetation sub-objects (1)		0.40	0.50	0.10	0.45	0.00	0.00	0.00
Formal settlement	Area		1800.00	1900.00	100.00	1850.00	0.00	0.00	0.00
	Not informal settlement	–	–	–	–	–	–	–	–
Informal settlement	Area of sub-objects (1)		36.00	38.00	2.00	37.00	0.00	0.08	0.08
	Asymmetry of sub-objects: mean (1)		0.61	0.63	0.02	0.62	0.00	0.10	0.10
	Relative area of bright small roofs/objects and sub-objects (1)		0.06	0.07	0.01	0.06	0.00	0.71	0.71
	Relative area of red roofs sub-objects (1)		0.01	0.02	0.01	0.02	49.00	1.79	50.79
	Relative area of small shadows/dark objects and sub-objects (1)		0.03	0.04	0.02	0.03	0.00	0.00	0.00
	Relative area of vegetation sub-objects (1)		0.09	0.10	0.01	–	–	–	–
Base level (Level 1)									
Small shadows/dark objects	Area		35.00	40.00	5.00	37.50	0.00	0.00	0.00
	Ratio blue		0.25	0.38	0.13	0.32	0.00	0.20	0.20
Vegetation	Ratio NIR		0.29	0.30	0.01	0.30	0.00	0.00	0.00
Red roofs	Ratio red/ratio green		1.25	1.30	0.05	1.28	4.00	0.16	4.16
Bright small roofs/objects	Area		40.00	60.00	20.00	50.00	0.00	0.00	0.00
	Brightness		750.00	775.00	25.00	762.50	0.00	0.00	0.00
	Shape index		1.20	1.30	0.10	–	–	–	–

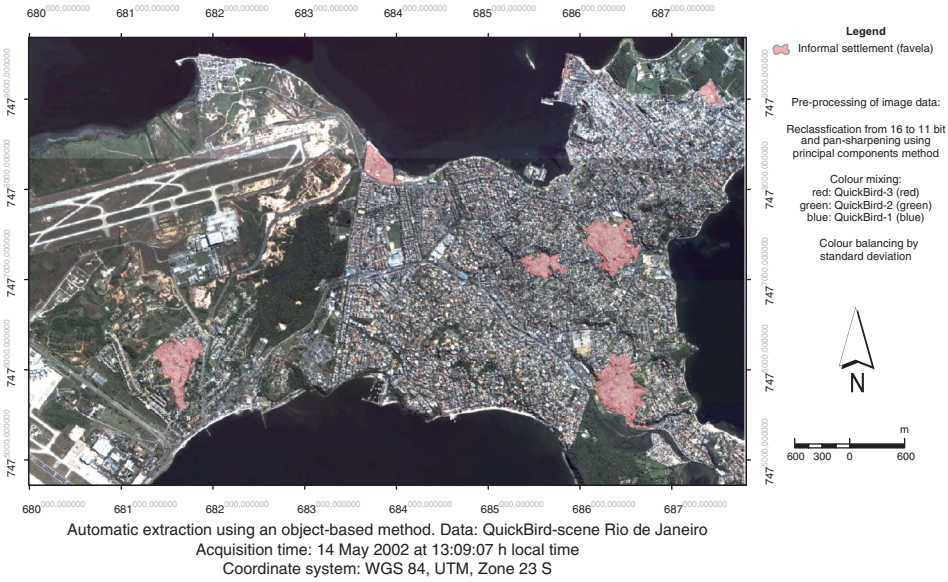

Figure 10. Classification result with adapted rule set on QuickBird scene from Rio de Janeiro: informal settlements (favellas) at the Ilha do Governador, Rio de Janeiro (Brazil).

5.2 *Detecting urban green from very high resolution satellite images*

In this second test-bed a rule set was developed to classify and differentiate urban green areas in the city of Salzburg, Austria, based on a QuickBird scene from June 2005 (see Hölbling 2006). For our research the rule set was reapplied to a subset of the QuickBird scene and an appropriate subset of an IKONOS scene from November 2008, covering the same area (see figure 11). For the QuickBird subset an OA (see §5.1) of 0.77 could be achieved. To apply the rule set to the IKONOS scene, the initial segmentation was adapted as described earlier. First, the original rule set was applied at the segmented IKONOS scene, which led to a complete unclassified scene. Second, the original rule set was adapted in the same way as described previously until a visually convincing classification result was achieved. After applying the adapted rule set, its deviations and robustness were measured. As depicted in table 6, from the 32 properties used for the IKONOS subset, 14 properties (deviations of type F^a) were added or the principal type of the membership functions used for these properties was changed. Additionally, 18 membership functions were shifted, compressed or stretched (type F^b), which led to a deviation between the rule sets of $d = 60.61$. With this adapted rule set, the classification had an OA of 0.53, which led to a final robustness of the rule set of $r = 0.011$.

6. Discussing and interpreting rule set deviations

Besides the measurable deviation and robustness presented in tables 5 and 6, it is also possible to analyse the deviation for each property, which helps to identify crucial aspects of the rule set and to interpret possible reasons for these deviations. In our first test-bed scenario (informal settlements), changes of type F^a (adding or removing a membership function or changing the type of a membership function) took place

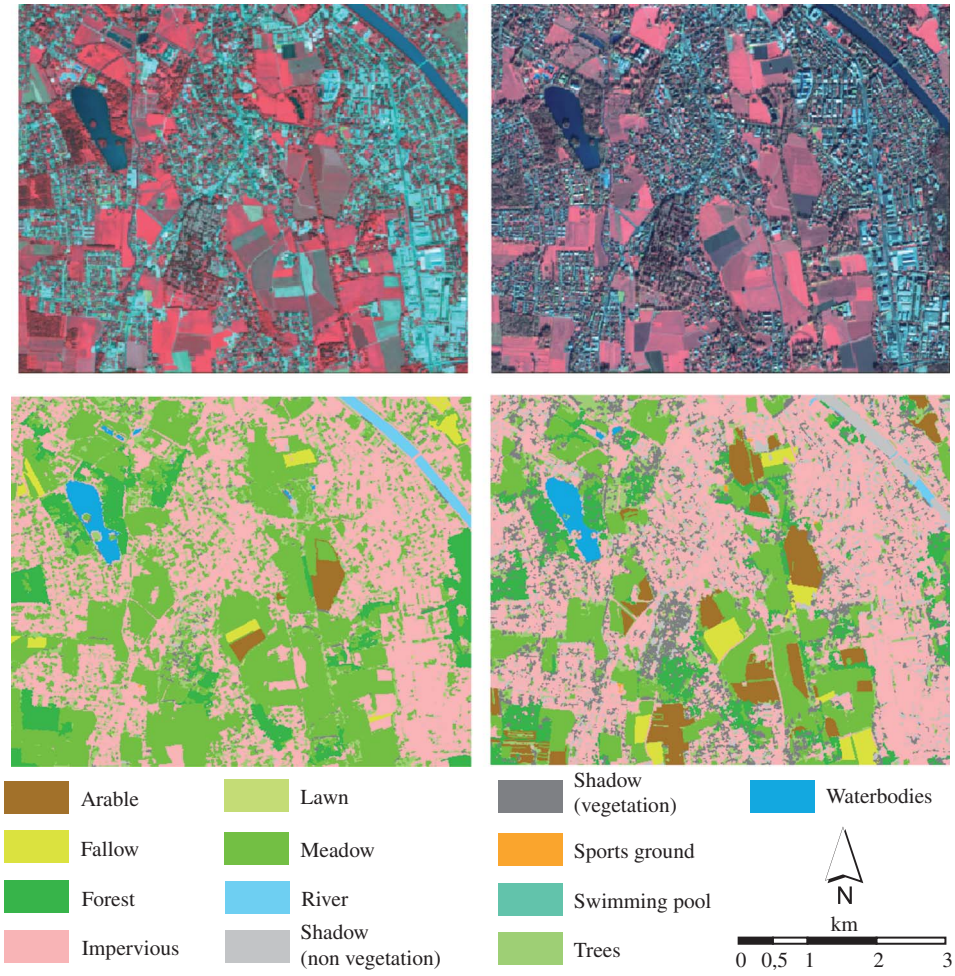

Figure 11. Land cover classification result for QuickBird (left) and IKONOS (right) subsets from Salzburg (Austria) for urban green mapping.

for the properties 'relative area of vegetation sub-objects (1)' (adding), describing the class 'informal settlement' at the upper segmentation level, and 'shape index' (adding), describing the class 'bright small roofs/objects' at the lower segmentation level. The addition of 'relative area of vegetation sub-objects (1)' can be interpreted as a consequence of differences in the general settlement structures between Rio de Janeiro and Cape Town; although the percentage of vegetation in informal settlements in South Africa is usually relatively low, it is in general higher in the favellas of Rio de Janeiro. Nevertheless, the vegetation fraction is relatively lower in Rio de Janeiro's favellas than in other (formal) settlement areas in Rio. Thus, it became necessary to explicitly formulate this difference of settlement structure accordingly in the rule set for the Rio de Janeiro scene. A 'shape index' was added that describes the smoothness or roughness of an object's border within a value range of 1 (maximum smoothness) and ∞ (maximum roughness). This might instead be due to differences in the sensors; the class 'bright small roofs/objects' acts as an indicating class for the objects at the upper

Table 6. Deviations between rule sets to detect urban green from a QuickBird scene from May 2005 and an IKONOS scene from November 2008.

		Deviations			
		Type F^b			Type F^a
Class	Property	δv	δa	δF	Add/remove/change
Forest	GLCM contrast (all directions)	–	–	–	Add
	Mean NIR	19.000	2.492	21.492	–
	Standard deviation NIR	–	–	–	Change
Arable	Compactness	–	0.500	0.500	–
	Mean red	0.829	0.167	0.995	–
	Standard deviation NIR	0.937	0.118	1.054	–
Fallow	Mean red	4.000	0.769	4.769	–
Waterbodies	Mean green	0.000	0.406	0.406	–
River	Mean green	1.500	1.191	2.691	–
	Standard deviation blue	–	–	–	Add
Swimming pool	Area (m^2)	–	–	–	Change
	Brightness	4.000	1.243	5.243	–
	Compactness	0.000	0.077	0.077	–
	Mean blue	0.968	0.465	1.433	–
	Ratio red	1.000	0.156	1.156	–
	NDVI	–	–	–	Add
	Ratio NIR	–	–	–	Add
	Standard deviation blue	–	–	–	Add
	Standard deviation NIR	–	–	–	Add
Shadow	Brightness	0.922	0.283	1.205	–
Shadow (vegetation)	NDVI	0.000	0.909	0.909	–
Shadow (not vegetation)	NDVI	0.000	0.909	0.909	–
Vegetation	NDVI	–	–	–	Add
	Ratio NIR	–	–	–	Remove
Non-vegetation	NDVI	–	–	–	Add
	Ratio NIR	–	–	–	Remove
Trees	Existence of 'forest' (0)	–	–	–	Remove
	Standard deviation NIR	1.000	0.556	1.556	–
Meadows	Standard deviation NIR	0.333	0.217	0.551	–
Sports ground	Ratio blue	0.934	0.391	1.325	–
	Standard deviation blue	0.000	0.333	0.333	–
Meadow	Standard deviation NIR	–	–	–	Add

segmentation level through the property 'relative area of bright small roofs/objects'. This property indicates the density of small buildings with a bright roof (or other small and bright objects), which can be assumed not to be a typical dwelling of informal settlements, such as a toolsheds or garage. Because of the higher spatial resolution of the QuickBird sensor, the outlines of each object are depicted more precisely. Hence, adding 'shape index' to the class description of 'bright small roofs/objects' enhances the detection of objects of this class. Regarding the deviations of type F^b (shift and/or stretch or compression of a membership function along the p-axis), the relatively high

value of δF for the property 'relative area of red roofs sub-objects (1)' is notable. Its contribution to the relative low values of d and r is explained by its high amount towards d. That is, the high value of $\delta F = 50.79$ for the relative deviation of 'relative area of red roofs sub-objects (1)' is equivalent to a contribution of 90.62% to the sum of all deviations expressed by δF and to a contribution of 87.50% to the overall deviation ($d = 58.05$) of the adapted rule set. This relatively high ratio of δF for the property 'relative area of red roofs sub-objects (1)' indicates that this property must be crucial in terms of the rule set's robustness. Thus, if it were possible to avoid using this property or to exchange it with a more stable property, a higher robustness might be possible. A possible reason for this outstanding deviation is that different materials were used for building the dwellings; in South Africa materials such as metal or wooden sheets combined with plastic materials are very common whereas in Brazil and especially in Rio, clay and brick stones are more common. Therefore, revising the rule set by taking into account different local contexts is a reasonable strategy. However, those properties whose δF was at zero seem to be stable and therefore reusable for applications within the scope given; that is these properties with the selected values seem to be most suitable for reapplying this rule set to detect informal settlements in IKONOS and QuickBird images.

As the rule set of the second scenario (urban green) is relatively complex, we focus our discussion only on changes. Most notable is the high number of deviations of type F^a (table 6): for the classes 'vegetation' and 'not vegetation' the feature 'ratio NIR', which describes the object-related relative brightness of the NIR channel against the mean brightness, was replaced by 'NDVI', describing the mean value per-object of the normalized difference vegetation index as described in Lillesand and Kiefer (2000). As these exchanges do not affect the semantics of the rule set but enhance the separation of 'vegetation' and 'not vegetation', these changes could also be weighted by zero. This would reduce the deviation to $d = 56.61$ and increase the robustness to $r = 0.012$. Furthermore, it is noteworthy that the classes 'forest', 'river', 'swimming pool' and 'meadow' were extended by the textural features 'GLCM (contrast)', which indicates the contrast of an object based on the grey level co-occurrence matrix after Haralick *et al.* (1973), and 'standard deviation in the blue and/or NIR channel'. This indicates that at least these classes are better detected (in the IKONOS image) by adding texture features to the feature space. The feature 'existence of 'forest' (0)', which determines whether or not there are objects of 'forest' in the direct neighbourhood of an object, was removed for the class 'trees' because it turned out to be obsolete. Finally, the description of 'area (m^2)' of the class 'swimming pool' has been changed from a 'fuzzy-greater-than' to a 'fuzzy-lower-than' function with appropriately changed values for α and β. These two changes can also be interpreted as changes that do not affect the principal semantics of the rule set and can therefore each be weighted by zero. We would then obtain a deviation of $d = 54.61$ and a robustness of $r = 0.013$. Regarding the deviations of type F^b, most of them can be interpreted as changes due to seasonal effects. In particular, the relatively high measured deviation of 'mean NIR' in the class 'forest' is noted. This 'fuzzy-greater-than' function has been compressed and shifted by changing its range from 500–600 to 155–160. For the case this seasonal effect could be eliminated by the rule set; for example, in terms of an automatic adjustment of the value range with respect to the season, the deviation was at $d = 33.12$ and the robustness at $r = 0.02$.

7. Conclusions

The method introduced enables the evaluation of the robustness of (fuzzy) rule sets for a defined scope through comprehensive and objectively measurable values. Thereby, the two criteria achievable classification quality and amount of deviation from a reference rule set are the determining factors for the evaluation. An approach to quantify the robustness against quality and deviation is introduced (§3) and applied to an artificial example (§4) and to two real-world test-beds (§5). We have demonstrated how crucial elements of a fuzzy rule set can be identified and how the observable deviations of a rule set can be evaluated and interpreted (§6). These capabilities open new avenues for future rule-base designs aimed at transferable classification applications in large production environments with several or even hundreds of images.

Acknowledgements
We thank Heinz Rüther, University of Cape Town, for his support, Space Imaging for the provision of the IKONOS scene and INTERSAT for providing the QuickBird scene. We also thank Daniel Hölbling for providing the results of his research and the reviewers for their constructive comments.

References

Albrecht, F., 2008, Assessing the spatial accuracy of object-based image classifications. In *Geospatial Crossroads @ GI_Forum '08. Proceedings of the Geoinformatics Forum Salzburg*, A. Car, G. Griesebner and J. Strobl (Eds.), pp. 11–20 (Heidelberg: Wichmann).

Ay, N. and Krakauer, D.C., 2007, Geometric robustness theory and biological networks. *Theory in Biosciences*, **125**, pp. 93–121.

Benz, U., 1999, Supervised fuzzy analysis of single and multi-channel SAR data. *Transactions on Geoscience and Remote Sensing*, **37**, pp. 1023–1037.

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

Bezdek, J. and Pal, S., 1992, *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data*, p. xi (New York: IEEE Press).

Blaschke, T., 2010, Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, **65**, pp. 2–16.

Blaschke, T., Lang, S. and Hay, G.J. (Eds.), 2008, *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Lecture Notes in Geoinformation and Cartography (Berlin: Springer).

Blaschke, T. and Strobl, J., 2001, What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS – Zeitschrift für Geoinformationssysteme*, **14**, pp. 12–17.

Burnett, C. and Blaschke, T., 2003, A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling*, **168**, pp. 233–249.

Congalton, R.G. and Green, K., 1999, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices* (Boca Raton, FL: Lewis Publishers).

Definiens, 2004, *eCognition 4.0 User Guide*, 4th edn. (Munich: Definiens Imaging).

Gorte, B., 1998, *Probabilistic Segmentation of Remotely Sensed Images*. ITC Publication Series No. 63 (Enschede, NL: ITC Publication).

Grenier, M., Labrecque, S., Benoit, M. and Allard, M., 2008, Accuracy assessment method for wetland object-based classification. In *GEOBIA 2008 – Pixels, Objects, Intelligence. GEOgraphic Object Based Image Analysis for the 21st Century. International*

Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C1, G.J. Hay, T. Blaschke and D. Marceau (Eds.), pp. 285–289 (Calgary, AL: University of Calgary).

Haralick, R.M., Shanmugam, K. and Dinstein, I., 1973, Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, **3**, pp. 610–621.

Hay, G.J., Blaschke, T., Marceau, D.J. and Bouchard, A., 2003, A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry and Remote Sensing*, **57**, pp. 327–345.

Hay, G.J. and Castilla, G., 2006, Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **36**, CD-ROM.

Hay, G.J. and Castilla, G., 2008, Geographic Object-Based Image Analysis (GEOBIA): a new name for a new discipline. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 93–112 (Berlin: Springer).

Hay, G.J., Marceau, D.J., Dube, P. and Bouchard, A., 2001, A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology*, **16**, pp. 471–490.

Hofmann, P., 2005, Übertragbarkeit von Methoden und Verfahren in der objektorientierten Bildanalyse – das Beispiel informelle Siedlungen [Transferability of Methods and Procedures in Object-Oriented Image Analysis – the Example of Informal Settlements]. PhD thesis, University Salzburg, Austria.

Hofmann, P., Strobl, J. and Blaschke, T., 2008*a*, A method for adapting global image segmentation methods to images of different resolutions. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **38**, CD-ROM.

Hofmann, P., Strobl, J., Blaschke, T. and Klux, H., 2008*b*, Detecting informal settlements from QuickBird data in Rio de Janeiro. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 531–533 (Berlin: Springer).

Hölbling, D., 2006, Objekt-basierte Klassifikation relevanter urbaner Grünstrukturtypen auf höchstauflösenden Fernerkundungsdaten – Automatisierung und Übertragung [Object-based classification of relevant urban types of estate with remote sensing data of highest resolution – automation and transferability]. MSc thesis, University Salzburg, Austria.

Hu, X., Tao, C.V. and Prenzel, B., 2005, Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity, and color features. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 1399–1406.

ISO/TC 211, 2003, ISO 19114 Geographic information – Quality evaluation procedures. Draft International Standards. International Organization for Standardization.

Jen, E., 2003, Essays & commentaries: stable or robust? What's the difference? *Complexity*, **8**, pp. 12–18.

Kitano, H., 2007, Towards a theory of biological robustness. *Molecular Systems Biology*, **3**, pp. 1–7.

Lang, S., 2008, Object-based image analysis for remote sensing applications: modelling reality – dealing with complexity. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 1–25 (Berlin: Springer).

Lang, S. and Langanke, T., 2006, Object-based mapping and object-relationship modelling for land use classes and habitats. *Photogrammetrie, Fernerkundung, Geoinformation*, **1**, pp. 5–18.

Leukert, K., 2005, Übertragbarkeit der objektbasierten Analyse bei der Gewinnung von GIS-Daten aus Satellitenbildern mittlerer Auflösung [Transferability of the object-based analysis at the extraction of GIS-data from satellite images of medium resolution]. PhD thesis, University of the German Federal Armed Forces, Munich, Germany.

Lillesand, T.M. and Kiefer, R.W., 2000, *Remote Sensing and Image Interpretation*, 4th edn., pp. 568–575 (New York: John Wiley & Sons).

Platt, R.V. and Rapoza, L., 2008, An evaluation of an object-oriented paradigm for land use/land cover classification. *The Professional Geographer*, **60**, pp. 87–100.

Pötzl, M., 1996, Robuste Tragwerke [Robust Frames]. *Bauingenieur*, **71**, pp. 481–488.

Schiewe, J. and Gähler, M., 2008, Modelling uncertainty in high resolution remotely sensed scenes using a fuzzy logic approach. In *Object-Based Image Analysis*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 755–768 (Berlin: Springer).

Smith, G.M., 2008, The development of integrated object-based analysis of EO data within UK national land cover products. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 513–528 (Berlin: Springer).

Story, M. and Congalton, G.C., 1986, Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, **52**, pp. 397–399.

Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T.M., Suppiah, I. and Hashim, S.A.M., 2008, Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing*, **29**, pp. 3105–3117.

Tiede, D., Lang, S. and Hoffmann, C., 2008, Domain-specific class modelling for one-level representation of single trees. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 133–151 (Berlin: Springer).

Walker, J. and Blaschke, T., 2008, Object-based landcover classification for the Phoenix metropolitan area: optimization vs. transportability. *International Journal of Remote Sensing*, **29**, pp. 2021–2040.

Weinke, E., Lang, S. and Preiner, M., 2008, Strategies for semi-automated habitat delineation and spatial change assessment in an Alpine environment. In *Object-Based Image Analysis. Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G.J. Hay (Eds.), pp. 711–732 (Berlin: Springer).

Wuest, B. and Zhang, Y., 2009, Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**, pp. 55–64.

Zadeh, L., 1968, Fuzzy algorithms. *Information and Control*, **12**, pp. 94–102.

Zhang, Q.F., Molenaar, M., Tempfli, K. and Shi, W., 2005, Quality assessment for geospatial objects derived from remotely sensed data. *International Journal of Remote Sensing*, **26**, pp. 2953–2974.