Chapter 1.1

Object-based image analysis for remote sensing applications: modeling reality – dealing with complexity

S. Lang [1]

[1] Centre for Geoinformatics (Z_GIS), Paris-Lodron University Salzburg, Schillerstr. 30, 5020 Salzburg, Austria; stefan.lang@sbg.ac.at

KEYWORDS: GMES, monitoring, geon, class modeling, cognition networks, object assessment

ABSTRACT: The advancement of feature recognition and advanced image analysis techniques facilitates the extraction of thematic information, for policy making support and informed decisions. As a strong driver, the availability of VHSR data and the ever increasing use of geo-information for all kinds of spatially relevant management issues have catalyzed the development of new methods to exploit image information more 'intelligently'. This chapter highlights some of the recent developments from both technology and policy and poses a synthetic view on an upcoming paradigm in image analysis and the extraction of geo-spatial information. It starts from a review of requirements from international initiatives like GMES (Global Monitoring of Environment and Security), followed by a discussion the possible answers from OBIA including a detailed portrait of the methodological framework of class modeling. The chapter closes with a short reflection on the required adaptation of standard methods of accuracy assessment and change detection, as well as on the evaluation of delineated and classified objects against the ultimate benchmark, our human perception.

1 Monitoring needs in a dynamic world

There is an ever increasing demand for regularly updated geo-spatial information combined with techniques for rapid extraction and targeted provision of relevant information. The need for timely and accurate geospatial information is steadily increasing, trying to keep pace with the changing requirements of the society at a global dimension. International initiatives strive for standardized solutions, as examples like cooperative effort of Global Earth Observation System of Systems (GEOSS) or the European initiative Global Monitoring for Environment and Security (GMES) impressively show. These initiatives strive to provide holistic, yet operational answers to global conventions or trans-national directives and agendas to halt uncontrollable change of physical parameters or loss of lives both potentially human-induced by unlimited growth (e.g. UN Framework Convention on Climate Change, FCCC or the UN Convention on Biological Diversity, CBD or the UN Convention to Combat Desertification, CCD[1]; EC Water Framework Directive, WFD or the EC Flora-Fauna-Habitat-Directive, FFH).

GMES: various applications – one approach

Beyond these more environmental aspects, the EC-ESA conjoint initiative GMES[2] follows the idea, that environmental integrity and societal stability both play together and may reinforce each other under certain conditions. The 'S' stands for security, and – next to environmental risks or potential hazards – the scope of observed treaties and conventions also comprises terrorism, critical infrastructure, refugees and weapons of mass destruction, to name just a few topics. Both GEOSS and GMES rely on remote sensing (RS) technology as a powerful and ubiquitous data provider, and both initiatives promote the integration of RS with in-situ data technology for the development of operational monitoring systems and integrated services, based on earth observation (EO) data. The scope is wide covering areas ranging from environmental integrity to human security, and the idea of serving this range of applications with a profound, ubiquitous set of pooled data and adaptive methods is compelling. And more than this, the approach is concept-wise sustainable, both in terms of its scientific strength and its operational capability. Operational services, delivered as

[1] http://www.unccd.int/
[2] www.gmes.info

fast-track core services (FTCS)[3] and such in preparation[4] provide status observations of highly complex systems with relevance to treaties and political agreements of different kinds. However, dealing with such advanced integrated tasks may no longer keep valid the monitoring of single compartments, but an integrated high-level approach (see 2).

Monitoring systems as required in the GMES context (Zeil et al., in press; Blaschke et al., 2007; Tiede & Lang; 2007) need to be capable of transforming complex scene content into ready-to-use information. The advancement in feature recognition and advanced image analysis techniques facilitates the extraction of thematic information, for policy making support and informed decisions, irrespective of particular application fields. The availability of such data and the increasing use of geo-information for sustainable economic development and protection of the environment have catalyzed the development of new methods to exploit image information more efficiently and target-oriented. Global commitments, directives and policies with their pronounced demand for timely, accurate and conditioned geo-spatial information, ask for an effective answer to an ever increasing load of data collected from various monitoring systems. It is obvious, yet maybe not consciously thought of, that – along with ever improved sensor technology – a technically and spatially literate user community asks for ever more advanced geo-spatial products, and expresses their needs accordingly. With an increased level of consciousness of prevailing problems the need for targeted information is rising double, it seems. The remote sensing community has to react and must deliver. When industry primarily highlights achievements in sensor developments, the efforts taken to analyze these data and to generate added value from these can hardly be underlined too much.

The upcoming paradigm of object-based image analysis (OBIA) has high potential to integrate different techniques of processing, retrieval and analyzing multi-resolution data from various sensors. By bridging technical and sector-oriented domains from remote sensing and geoinformatics we may significantly enhance the efficiency of data provision for policy making and decision support.

New data and increasing complexity: OBIA as the answer?

Recent years' advances in sensor technology and digital imaging techniques, along with ever increasing spatial detail, have challenged the re-

[3] FTCS on land, sea, and emergency
[4] Pre-operational services on security and atmosphere

mote sensing community to strive for new methods of exploiting imaged information more intelligently. The word ‘intelligence’ in this context has several facets: (1) an advanced way of supervised delineation and categorization of spatial units, (2) the way of how implicit knowledge or experience is integrated, and (3) the degree, in which the output (results) are contributing to an increase of knowledge and better understanding of complex scene contents.

New earth observation (EO) techniques and concepts from GIScience have led to the emerging field of OBIA[5]. The main purpose of OBIA in the context of remote sensing applications is to provide adequate and automated methods for the analysis of very high spatial resolution (VHSR) imagery by describing the imaged reality using spectral, textural, spatial and topological characteristics. OBIA offers a methodological framework for machine-based interpretation of complex classes, defined by spectral, spatial and structural as well as hierarchical properties (Benz et al., 2004; Burnett & Blaschke, 2003; Schöpfer et al., in press; Niemeyer & Canty, 2001; Hay et al., 2003). OBIA has been pushed by the introduction of fine resolution image data that for a broad range of application domains provides an h-res situation (Strahler et al., 1986). In h-res situations, the pixel size is significantly smaller than the average size of the object of interest. In this constellation, segmentation as a means of regionalization is an efficient means of aggregation the high level of detail and producing usable objects. Therefore, segmentation is a crucial methodological element in OBIA, but not an exclusive or isolated one (see 3).

VHRS satellite imagery is widely available now and gained popularity in research, administration and private use. If not the ‘real’ data, so at least the ‘natural color’ products can be easily accessed through web-based virtual globes like Google Earth, NASA World Wind, MS Virtual Earth and the like. Globes have penetrated daily life information exchanges, and satellite data in a ‘this-is-my-house-resolution’ have become the staple diet to feed people’s greed for immediate contextual spatial information[6].

From a scientific point of view satellite-mounted sensors and air-borne scanners have now reached the level of detail of classical aerial photography. For decades, fine-scaled (i.e. larger than 1:10,000) environmental as well as security-related applications were relying on aerial photography and visual inspection as a primary means to extract relevant information.

[5] The scientific community discusses to use the term GEOBIA to emphasize (1) the strong contribution of GIScience concepts and (2) the focus on space related applications (see Castilla & Hay in this volume).

[6] See Tiede & Lang (2007) for a discussion how the presence of familiar spatial context can be utilizing for communicating complex analytical contents.

On the other hand, medium to low resolution satellite sensors were mainly used for coarse-scaled mapping and multi-spectral classification, with a probabilistic, yet limited set of classes being targeted at (Lang, 2005). This quite dichotomous gap has been separating two application domains and the respective methods applied, e.g. in landscape-related studies which had to choose between either of them (Groom et al, 2004).

Closing this gap, but embarking on another class of problem: with the advent of digital data from airborne and satellite-borne sensors we return to the very challenge of air-photo interpretation: how do we deal with the enormous detail? Looking back to several decades of computer technology we trust in automated methods for analysis and interpretation, even of complex imaged contents. While a several problems remain a challenge, a range of tangible solutions have been developed by successfully combining GIS and remote sensing techniques for reaching closer at the photo-interpreter's capacity.

As briefly mentioned before, the need for highly accurate and regularly updated geo-spatial information cannot be met by advancements of sensor technology alone. New sensors and new kinds of data may do provide a wealth of information, but this 'information overload' needs to be conditioned, in order to fit the communication channels of the addressees. Thus, advanced methods are required to integrate single information packages. It is necessary to both synchronize technologies and harmonize approaches. The first is related to the acquisition, pre-processing, and retrieval of multi-sensor, multi-spectral, multi-resolution data from various sensors. The second deals with the integration of spatial analysis techniques into image processing procedures for addressing complex classes in a transparent, transferable and flexible manner.

The guiding principle of OBIA is likewise clear as it is ambitious: to represent complex scene content in such a way that the imaged reality is best described and a maximum of the respective content is understood, extracted and conveyed to users (including researchers). The realization, therefore, is not trivial, as the ultimate benchmark of OBIA is human perception (see 3.3). This, our visual sense of the environment is a common experience, easy to share yet difficult to express in words or even rule sets. Indeed, the challenge is to make explicit the way how we deal with imaged information in various scales, how we manage to relate recognized objects to each other with ease, how we understand complex scene contents readily. To this end, OBIA utilizes concepts from spatial thinking, which again is influenced by cognitive psychology.

2 A plurality of solutions – conditioned information and geons

An increasing detail of data and complex analysis tasks opens the door for a plurality of solutions. Often, there is no longer a single valid choice of e.g. a distinct land cover class. Rather, there is a user-driven set of classes; not necessarily restricted to extractable features, but expressed according to the very demand. Fine-scaled representations of complex real world phenomena require means for modeling the underlying complexity, for mapping the dynamics and constant changes. Automated techniques making effective use of advanced analysis methods help understanding complex scene contents and try to match the information extraction with our world view.

But OBIA is more than feature extraction (see chapter 3). It provides a unifying framework with implications for policy-oriented delivery of conditioned information. By this, it also enables monitoring of system-driven meta-indicators like vulnerability or stability.

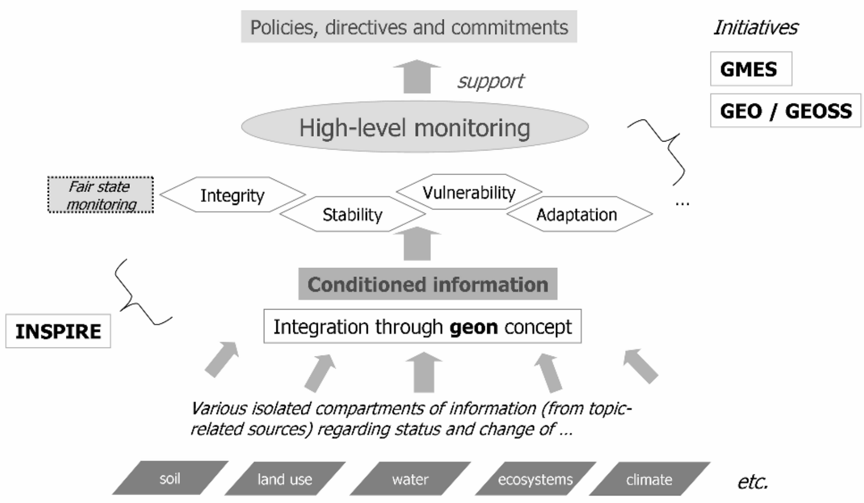

Fig. 1. High-level indicators for monitoring

From this point of view, a broad concept of manageable units makes sense. The author proposes the term *geon* (Lang & Tiede, 2007) to describe generic spatial objects that are derived by regionalization and homogenous in terms of a varying spatial phenomenon under the influence of, and partly controlled by, policy actions. A geon[7] (from Greek *gé* = Earth and *on* = part, unit) can be defined as a homogenous geo-spatial referencing unit, specifically designed for policy-related spatial decisions. The geon concept (see figure 2) can be seen as a framework for the regionalization of continuous spatial information according to defined parameters of homogeneity. It is flexible in terms of a certain perception of a problem (policy relevance/scope). It employs a comprehensive pool of techniques, tools and methods for (1) geon generation (i.e. transformation of continuous spatial information into discrete objects by algorithms for interpolation, segmentation, regionalization, generalization); (2) analyzing the spatial arrangement, which leads to emergent properties and specific spatial qualities; and (3) monitoring of modifications and changes and evaluation of state development. The latter, characterizing spatio-temporal variability require appropriate means to distinguish noise or slight modifications from real changes. In addition, there is the possibility of recovering objects in the presence of 'occlusions[8]' (i.e. data errors, measure failures, lack of data, mismatch of data due to bad referencing).

[7] The term geon (for geometric ions) was initially used by Biederman (1987), who hypothizes that cognitive objects can be decomposed into basic shapes or components. Geons in Biederman's view are basic volumetric bodies such as cubes, spheres, cylinders, and wedges. The concept used here is related, but not identical to this view.

[8] This term, again, is borrowed from Biederman (1987)

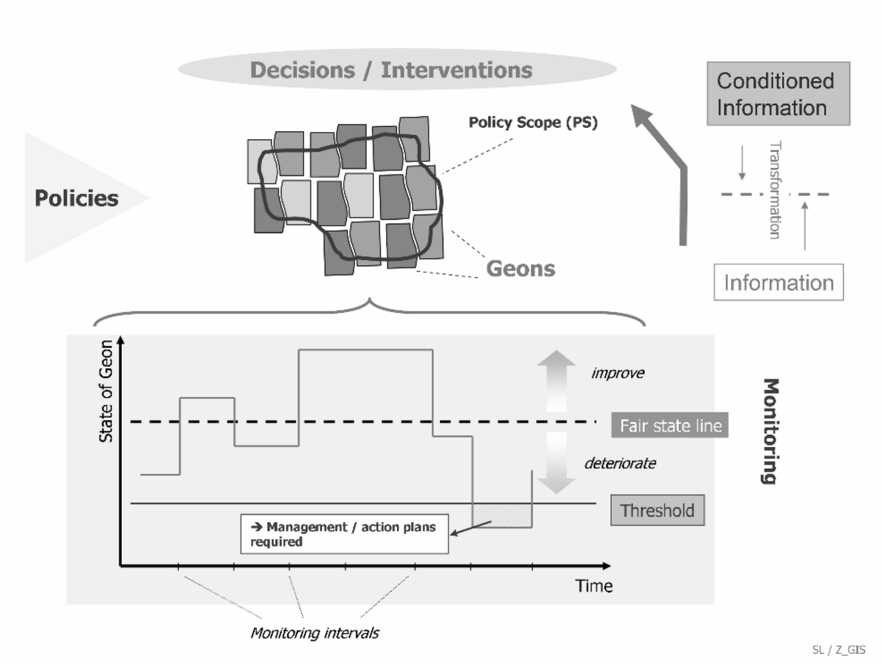

Fig. 2. The Geon Concept

Within the spatial extent in which a certain policy applies (policy scope, PS), a group of geons makes up a spatially exhaustive set (geon set). PS comprises the spatio-temporal extent in which a certain policy is valid. This extent usually coincides with administrative units, but not necessarily does: in the context of the EC Water Framework Directive, catchments function as reference units. In cases when PS is defined by legal boundaries, the spatial limit of the geon set, as derived functionally, may not fully coincide with PS. As policies address various scale domains and their implications apply to particular domains, a geon set is scale-specific and adapted to the respective policy level. Several geon sets may exist, altogether forming in a spatial hierarchy. Using geons, we are capable of transforming singular pieces of information on specific systemic components to policy-relevant, conditioned information. Geons are of dynamic nature. Monitoring the spatio-temporal development of geons is critical for assessing the impact of policies and the compliance with obligations or commitments attached to those. The ‘fate’ of a geon is monitored using the fair state concept, which takes into account the natural dynamics (improvement or deterioration in regard to an optimum or average state). Management actions have to be taken, when a certain threshold is reached.

Irrespective of the very concept applied for naming delineated units, and irrespective of the different fields of use, OBIA aims at the delineation and the classification of relevant spatial units. The way to perform this task is an integrated, cyclic one, and in the following section this will be discussed, under the heading ‘class modeling’.

3 Class modeling

From a methodological point of view, one may observe a convergence of various techniques from formerly distinct GIS and remote sensing embankments; aiming at the aforementioned purpose, OBIA is trying to bridge these. OBIA rests upon two interrelated methodological pillars, i.e. (1) segmentation / regionalization for nested, scaled representations; (2) rule-based classifiers for making explicit the required spectral and geometrical properties as well as spatial relationships for advanced class modeling. We speak of ‘image analysis’ and not merely of ‘image classification’, since the process of OBIA is iterative rather than a linear and strict subsequent one. The process of OBIA is a cyclic one. It is usually not enough to think of (a) delineation and (b) labeling[9]. By its iterative nature, the process is highly adaptive and open for accommodating different categories of target classes, from specific domains, with different semantics, etc. Class modeling (Tiede et al., 2006; Tiede & Hoffmann, 2006) enables operators to tailoring transformation of scene contents into ready-to-use information according to user requirements. It supports data integration and the utilization of geo-spatial data other than images (e.g. continuous data like altitude or data representing administrative units).

[9] To underline this, Baatz et al. (this volume) propose the term “object-oriented” to be used instead of “object-based”, because the former is more target-oriented, teleological, whereas the latter may be misleading and implying a more static concept. However, since the term “object-oriented” is strongly occupied by computer scientists, the author stays with “object-based”.

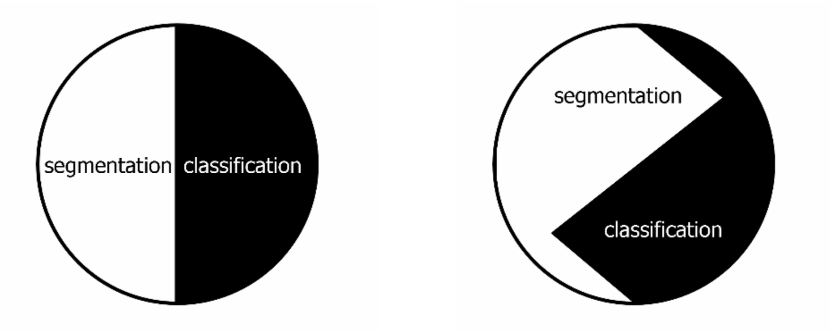

Fig. 3. Class modeling: the symbology marks the move from distinct realms of segmentation and classification towards an interlinked concept

Class modeling (as for example realized by the modular programming language CNL, cognition network language), provides flexibility in providing problem-oriented solutions for advanced analysis tasks. Examples are scene-specific high-level segmentation and region-specific multi-scale modeling (Tiede et al., this volume) or the composition of structurally defined classes as proposed by Lang & Langanke (2006). The latter was successfully realized in a study on semi-automated delineating biotope complexes (Schumacher et al., 2007, see figure 9a).

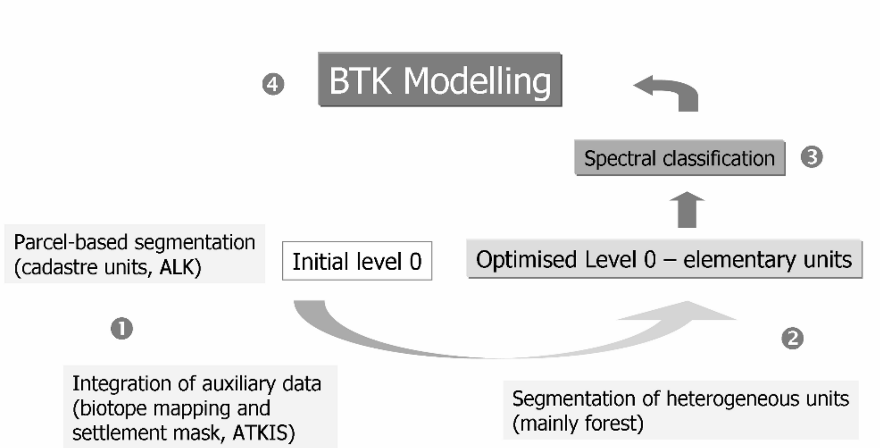

Fig. 4. Class modeling – a cyclic process. The example of modeling biotope complexes

From pixels (via image regions) to image objects

In object-based image analysis, the 'image object' is the central methodological element and as an object of investigation, it resides somewhere between application-driven plausibility and technology-driven detectability. To this end, we conjoin image segmentation with knowledge-based classification. Image segmentation decreases the level of detail, reduces image complexity, and makes image content graspable. Segmentation produces image regions, and these regions, once they are considered 'meaningful', become image objects; in other words an image object is a 'peer reviewed' image region; refereed by a human expert. A pixel as a technically defined unit can be interpreted in terms of its spectral behavior, in terms of the aggregation of spectral end-members, or in terms of its neighborhood. A pixel cannot be assigned a valid corresponding real-world object, but an image object can. Overcoming the pixel view and providing image objects that 'make sense' opens a new dimension in rule-based automated image analysis; image objects can be labeled directly using a range of characteristics, including spatial ones, or they can be used for modeling complex classes based on their spatial relationships. Coupled with e.g. a rule-based production system we can make expert knowledge explicit by the use of rules (see below).

Hierarchical, multi-scale segmentation

Multi-scale segmentation has often been linked with hierarchy theory (Lang, 2002). This is an appealing concept, and the comparison seems obvious as both hierarchy theory and multi-scale segmentation deal with hierarchical organization. Still we need to be careful: hierarchy theory proposes the decomposability of complex systems (Simon, 1973), but imagery is just *a* representation of such systems. An imaged representation is in several aspects reductionism: it is a plane picture only revealing reflection values. So hierarchy theory forms a strong conceptual framework, rather than a methodological term for multiple segmentation cycles. What we delineate, needs to be assigned a function first (image regions – image objects – (geons) – classes, see above). We should be aware that hierarchical segmentation at first hand produces regions of increasing average size (or number of pixels, respectively). But hierarchy theory is not about size, it deals with increasing degree of organization (Laszlo, 1972; Szaramovicz, 2004). What makes a strong link to hierarchy theory is not multiple segmentation alone, but the way we approach complexity, how we model and decompose it, and how we integrate our knowledge about it.

When fitting image complexity into hierarchical levels, it does not happen independently from human perception (Lang 2005). Allen & Starr (1982) point out that "discrete levels need to be recognized as convenience, not truth" and levels would emerge "as device of the observer" (ibid.). While drastically expressed, we should be aware that instead of questioning the ontological truth of scaled representations, we should rather focus on their epistemological relevance for target objects. Human perception is a complex matter of filtering relevant signals from noise, a selective processing of detailed information and, of course, experience. To improve automated object extraction we therefore seek for mimicking the way how human perception works (see 3.3 and Corcoran & Winstanley, this volume).

Is there one single set of multiple segmentations applicable 'horizontally' over the entire scene? The multi-scale option does not always lead to satisfying results. This applies in scenes, where multiple scales occur in different domains. Tiede et al. (this volume) discuss an application of regionalized segmentation (Lang, 2002).

Nested scaled representations need to consider scale. While this statement is somewhat tautologically, there is no single solution to this and different approaches exist to address this. One, aiming at a strict hierarchical representation, performs multi-scale segmentation with coinciding boundaries. In other words: a super-object gets assigned exactly n sub-objects (Baatz & Schäpe, 2000). The advantage is a clear 1:n relationship between super- and sub-object. On the other hand, since boundaries are 'cloned' up the scaling ladder (Wu, 1999), boundaries will not be generalized. It is scale-adapted, but not scale-specific. On the other hand, scaled representations are scale-specific, if there is – as in visual interpretation – a match between average size and generalization of boundaries. This is for example realized in the software SCRM (Castilla, 2004). The drawback is, however, boundaries do not coincide and cannot be used for 'direct' modeling (but see Schöpfer et al., Weinke et al., this volume).

Knowledge representation and cognition networks

Knowledge plays a key role in the interpretation-focused, value-adding part of the remote sensing process chain (Campbell, 2002). We have at our disposal a huge store of implicit knowledge and a substantial part of it is used in image interpretation (ibid.). By training we subsequently complement implicit knowledge with a more formalized knowledge obtained through formal learning situations (e.g. the specific spectral behavior of stressed vegetation) and experience. From an artificial intelligence (AI) perspective two components of knowledge can be distinguished, procedural and structural knowledge. Procedural knowledge is concerned with the specific computational functions and is therefore represented by a set of rules. Structural knowledge implies the way of how concepts of e.g. a certain application domain are interrelated: in our case that means, in how far links between image objects and 'real world' geographical features is established. Structural knowledge is characterized by high semantic contents and it is difficult to tackle with. A way to organize structural knowledge is the use of knowledge organizing systems (KOS), either realized by graphic notations such as semantic networks (Ibrahim, 2000; Pinz, 1994; Liedtke et al., 1997; Sowa, 1999) or and more mathematical theories like formal concept analysis (FCA, Ganter & Wille, 1996). Within image analysis semantic nets and frames (Pinz, 1994) provide a formal framework for semantic knowledge representation using an inheritance concept (*is part of, is more specific than, is instance of*). As semantic nets need to be built manually, they allow for controlling each and every existing connection once being established. With increasing complexity the transparency and operability will reach a limit. Bayesian networks are manually built, but the weighting of the connections can be trained, though it has to be trained for every connection.

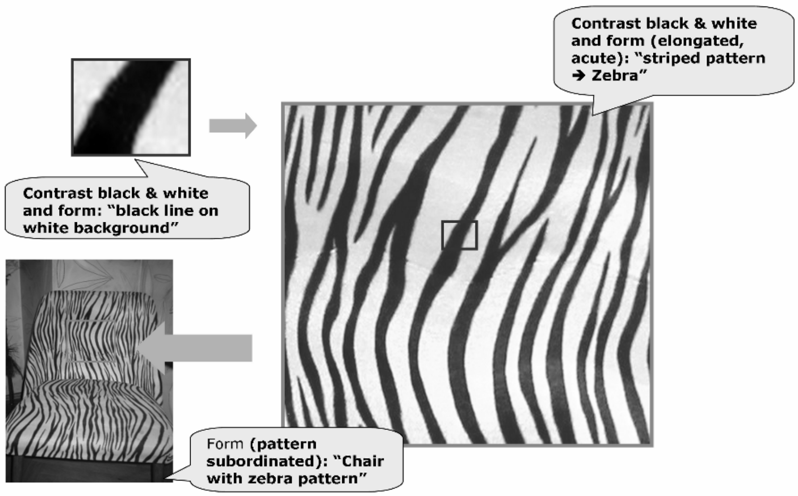

Fig. 5. Colour, form and arrangement evoke certain parts of our knowledge and experience (from Lang, 2005)

Over several decades techniques were developed to group pixels into spectrally similar areas and link spectral classes – if possible – to information classes. The pixel based approach is considered intrinsically limited (Blaschke & Strobl, 2001; Benz et al., 2004; Schöpfer et al., in press), since only spectral properties of (geographic) features are taken into account. The 'picture element' as a technical driven smallest unit integrates signals but does not reflect spatial behavior in a sufficient manner. Even if direct neighborhood is considered by applying kernel-based techniques, the 'environment' remains a square or any other predefined regular geometric shape. Modeling in this case is confined to spectral characteristics and related statistical behavior (texture, etc.). Spatial explicit characteristics are left aside.

The process of OBIA is supported by the use of so-called cognition networks (Binnig et al., 2002) or related concepts of KOS that provides the framework for modeling user-oriented target classes and their composition by spatial components. A cognition network controls the system of target classes and the class definitions as well as the mode of representation (e.g. regionalized segmentation or one-level vs. multi-level representation, see Tiede et al., this volume; Weinke et al., this volume). It provides the basis for a rule-based production system, which is controllable and transferable, as well transparent to both operators and users. Even if there are means and techniques to formalize knowledge and to encapsulate it into rule bases, the vast intuitive knowledge of a skilled interpreter operative for many years is hard to encapsulate in a rule system. Transferring existing experience effectively into procedural and structural knowledge remains a challenge of AI systems, especially taking into consideration the user-oriented plurality of solution, as discussed above. Cognition Network Language (CNL, Baatz et al., this volume), the meta-language to set up rule sets in Definiens software[10] offers a range of tools and algorithms to even address complex target classes. Establishing a rule set is often time-, labor- and therefore cost-intensive. But once a system is set up, and proved to be transferable, the effort pays off. The application of it to like scenes does not imply linear effort, as visual interpretation does. Therefore, in terms of operability one needs to distinguish between establishing a cognition network, and its, mostly scene-depending, parameterization.

OBIA can play a key role for image understanding (Lang & Blaschke, 2006). The entire process of image analysis is characterized by the transformation of knowledge. Finally, a scene description representing the image content should meet the conceptual reality of an interpreter or user. By establishing a body plan for the classes targeted at, knowledge is stepwise

[10] See www.definiens-imaging.com

adapted through progressive interpretation and class modeling. By this, knowledge is enriched through analyzing unknown scenes and transferring knowledge will incorporate or stimulate new rules. A particularity with spatial characteristics is related to the notion that parameters for spatial characteristics are difficult to determine. Whereas spectral reflectance can be measured directly by spectrometers on the ground, spatial or hierarchical properties are often ill-defined, less explicit and therefore difficult to be used as distinctive rules. Fuzzification of ranges (e.g. “rather small in size” or “rather compact”) is one way to tackle this problem, but it is often not the range as such that is ambiguous, but the very spatial property itself.

Pro-active classification

Operators applying methods of object-based image analysis must be ready for taking over responsibilities. Isn’t that contradictory? In most publications we read about ‘automated’ or at least ‘semi-automated’ analysis. Automation, of course, is the overall aim of using this approach – like with any other computer-based technique. However, with increasing complexity of the scenes and the classes or features to extract, we need to carefully feed the system with our experience, in a usable form. The approach towards this goal must be interdisciplinary: modeling complex target classes using spatial and structural characteristics not only requires computational skills, but also a wealth of knowledge about the area and the composition of the imaged setting.

Maybe it sounds provokingly, but it may be stated that a standard supervised multi-spectral classification is mechanistic. The machine is fed with samples, which we assume to be correct, and then provides a corresponding result. The process involves a certain level of probability but highly controllable. Class modeling, in contrast, is not mechanistic, but systemic. It not only requires ‘supervision’, but pro-active engagement from the operator. The combination of existing techniques incorporate know-how and existing experience in the different areas: (1) the modeling stage of the classes relies on expert knowledge that can build upon manual interpretation skills; (2) users that practice pixel-based statistical approaches can utilize their understanding in machine based classifications; (3) experiences in semi-automatically detecting and delineating features on high resolution data can be used for the classification process as such. Therefore object-based methods will integrate the existing remote sensing know how rather than replacing prevailing classification practices.

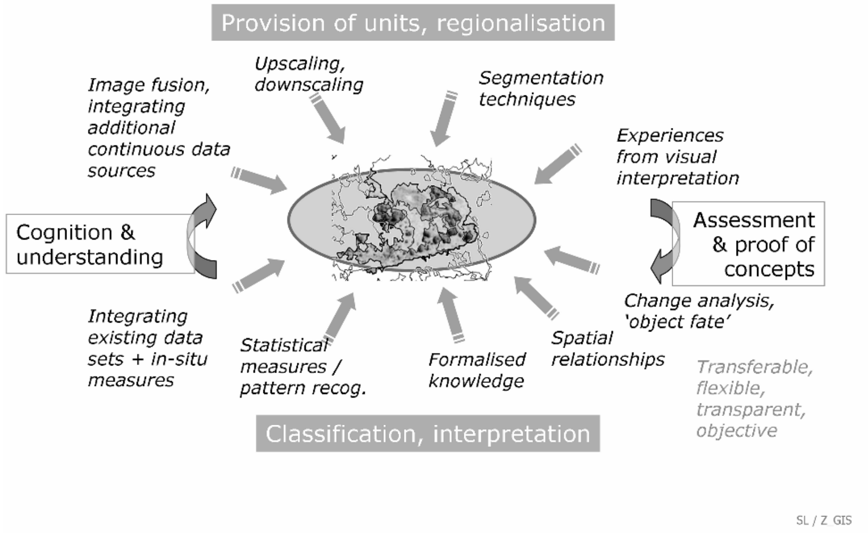

Fig. 6. OBIA as an integrated approach (from Lang, 2005, modified)

4 Object assessment and evaluation

Introducing new ways of image analysis is not free of challenges in terms of adapting both core and complementary methods (Lang & Blaschke, 2006). Among others, this applies for object-based accuracy assessment and object-based change detection. 'Object-based' in this case means that accuracy or changes are assessed in such a way that spatial properties are unambiguously reflected. Both spatial-explicit assessment techniques are facing similar challenge of comparing data sets according to criteria, which are hardly reducible to binary decisions of true or false. Possible solutions are discussed in this volume and preceding studies under different aspects (Schöpfer et al, this volume; Schöpfer & Lang, 2006; Lang et al., in press; Castilla & Hay, 2006; Weinke et al; this volume; Weinke & Lang, 2006). Thus, in the following two sections only key aspects are summarized.

Then, touching briefly at the strands of cognitive psychology, the chapter concludes with a note on object evaluation, for our results need to be proven at the ultimate benchmark, our human perception.

Object-based accuracy assessment

Quantitative site-specific accuracy assessment (Congalton & Green, 1999) using error matrices and specific assessment values (such as error of commission, error of omission, kappa ^) is widely used for evaluating the probability of correct class assignments and the overall quality or reliability of a classification result. Within OBIA, point-based accuracy assessment only gives indication on thematic aspects (labeling). Thematic assessment can be checked by generating random points within objects and comparing the labels against a ground truth layer. Alternatively, a set of objects can be selected in advance and be used as reference information. The decision of being thematically correct may not be a binary one: in fuzzy-based systems, the assessment of class membership is rather qualified by a confidence interval. Still, thematic accuracy assessment may be dubbed '1-dimensional': there is one specific label, i.e. the most likely classification result, to be checked on one specific site. Note that also checking an entire object in terms of its label is a point-based approach[11]. Spatial accuracy instead requires different ways of assessment. There are at least two aspects to be considered: (1) the appropriateness of an object's delineation (match of extend and shape with real situation) and (2) the precision of boundary delineation (match with scale applied).

In smaller test areas with a limited number of larger objects, every single object may be assessed individually: classified image objects can be visually checked against manual delineation (e.g. Koch et al., 2003). Still, a quantitative assessment requires at least some basic GIS overlay techniques. But performing hard intersections implies operational problems of producing sliver polygons and the like. A solution of 'virtual overlay' has been proposed by Schöpfer & Lang (2006) (see also Schöpfer et al., this volume), looking at specific object fate. This comprises object transition (fate in time) and object representation (fate in terms of different ways of delineation). Generally speaking, we encounter difficulties in performing object-based accuracy assessment, which satisfies the needs as being discussed by Congalton & Green (1999): (1) a 100% geometrical fit between reference and evaluation data is usually not given due to the issue of scale and the different ways of delineation; (2) the thematic classes are not mutually exclusive when using fuzzified rule bases. In other words, the accuracy is also a matter of geometrical and semantic agreement (Lang, 2005).

[11] This can be considered a special case of ecological fallacy (Openshaw, 1984), or better: individualistic fallacy, as we assume correctness for an entire area based on point-based observation. See also the discussion about polygon heterogeneity Castilla & Hay, 2006.

Object-based change detection and analysis

Monitoring is about detecting, quantifying and evaluating changes. Without proper change assessment, any decision making process is vague, and any management measure to be taken is ill-posed. Changing objects (or geons, see the discussion about fair state above) do not only change in terms of label, but also – and usually more often – in terms of their spatial properties. A habitat under observation may not have changed its class over time, given that is was measured on a specific point located e.g. in the centre of the habitat area, where there have no changes occurred. On the other hand, it may have been substantially shrinked through activities around it. So its function of providing living space for certain organisms may not be further fulfilled. In terms of its spatial component, object-based change detection faces a similar class of problems as object-based accuracy assessment. Common image-to-image or map-to-map comparisons (Singh, 1989) are site-, but not object-specific, i.e. they refer to pixels. Any aggregated measure based on this, becomes spatially implicit. Object-based change analysis needs to specifically compare corresponding objects. Methodological frameworks have been discussed by Blaschke, 2005; Niemeyer & Canty, 2001; Straub & Heipke, 2004; Schöpfer & Lang, 2006; Schöpfer et al., this volume. Like with object-based accuracy assessment, vector overlays (intersections) produce very complex geometry, which is later on difficult to handle and critical in terms of post-processing. Visual inspection is subjective and time-intensive and therefore of limited operational use.

In GIScience there are generic concepts available to describe the spatial relationships among spatial objects (e.g. Mark, 1999 or Hornsby and Egenhofer, 2000) which are built upon sets of spatial relationships. The relationships among spatial objects are built on basic topological principles like containment, overlap, and proximity. In addition to topology, the increase or decrease in size are two further qualities describing temporal dynamics, with presence or absence of objects form the extremes. While these basic categories for describing mutual spatial relationships are conceptually clearly distinguishable, in reality a combination or any kind of transition may occur. The problem is usually reinforced by the effects of generalization. Spatial changes may get completely averaged by applying smaller scales of representations (figure 7).

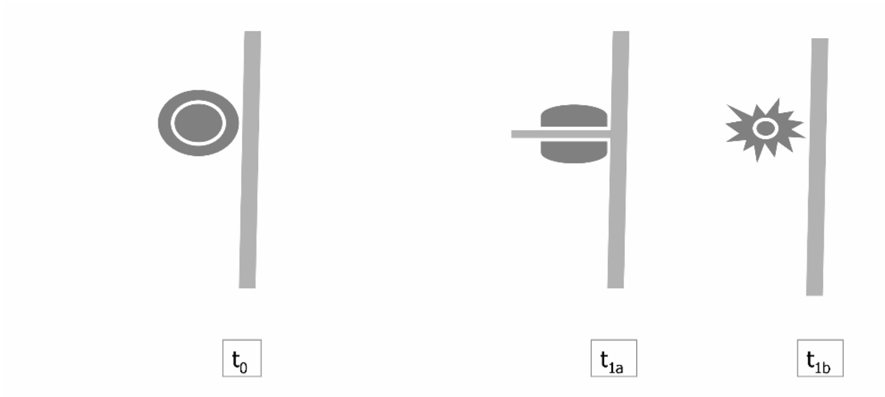

Fig. 7. Substantial changes that may only occur in fine-scaled representation. Left: Habitat next to a road in state t_0. Right: state t_{1b} - the habitat is split by a small road that (1) is not reflected in coarser scale and (2) only slightly decreases habitat area; state t_{1b}: the habitat has been influenced in terms of its boundaries; functional properties have changed substantially (the core area, indicated in white, has shrinked), but area has remained the same.

Object evaluation – the ultimate benchmark

It is recommendable, in the entire process of OBIA, not to forget the ultimate benchmark, i.e. our (visual) perception. The machine is supportive – it reacts on parameters, though the expert has to decide (figure 8).

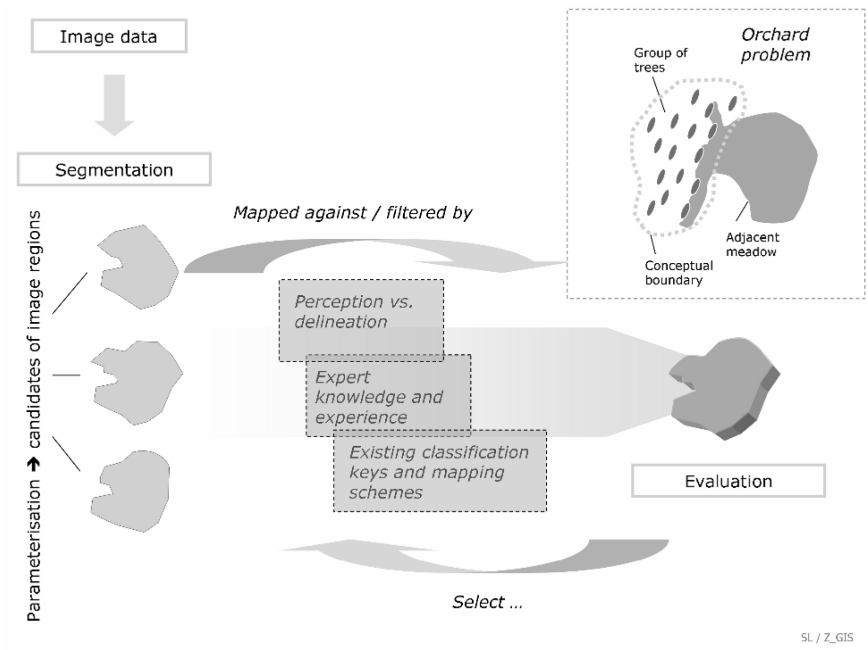

Fig. 8. Object delineation: the expert filter and the orchard problem (from Lang, 2005; Lang & Langanke, 2006, modified)

Region-based segmentation algorithms, like the name indicates, produce regions according to a certain criterion of homogeneity (spectral similarity, compactness, etc.). Due to their bottom-up nature, they are limited in providing delineations of aggregates that consists of high contrast, but regularly appearing (sub-)objects. These kinds of structural arrangements, such as an orchard (Lang & Langanke, 2006) or a mire complex with pools and hummocks (Burnett et al., 2003), are readily delineated by humans, though hard to grasp by a machine. This is a different kind of homogeneity: regularity in structure (repetitive patterns) or conformity (i.e. constancy) in change.

The *orchard problem* and related problems (Lang & Langanke, 2006) arises when addressing geographical features that exhibit conceptual boundaries rather than 'real' ones. Consider an orchard, which is delineated on an aerial-photograph with ease, because of the specific arrangement of fruit trees in a matrix of grass. Even, if the orchard is situated next to a meadow with the same spectral behavior as the matrix, humans would draw the boundary line somewhere in between. Reasons for that can be found in the human way to deal with heterogeneity according to the gestalt laws (Wertheimer, 1925) and other principles of human vision. Both the factor of good gestalt and the factor of proximity and continuation explain why humans would delineate an object on a higher level. As found out by Navon (1977), a scene is rather decomposed than built-up: if segmentation routine start from pixels, it can hardly mimic the way of visual processing, namely to start from a global analysis of the overall pattern and then to proceed subsequently to finer structures.

The limitations as pointed out above may require coping strategies like the introduction of hybrid techniques. This means in this context the complementary usage of machine-based automated delineation of basic units and high-level aggregation by a human interpreter (see figure 9).

Fig. 9. Delineation of habitat complexes: full automated class modeling vs. hybrid approach.

Recent advances in visual intelligence research have found further explanatory rules for the interpretation of geometrical (or spatial) structures (Hofman, 1998), and some of these provide valuable hints for the shortcomings of the difficulties we are facing when trying to automate the interpretation of complex scenes. There are rules concerning the way how something is interpreted in a constructive way, e.g. how lines in 2D are interpreted in 3D (straight lines, coinciding lines, collinear lines). But when dealing with satellite image data or air-photos these rules are less important, since the perspective is always vertical and requires abstraction, anyway. Others make us favor constellations which likely exist, and neglect theoretical, unlikely ones. This implies utilizing certain ‘world views’, e.g. the rule of regular views, which excludes some constellations to be very unlikely, is based on the concept of regularity; and this is a matter of experience. The Gestalt law of symmetry, though being powerful in explaining this phenomenon in part, is not capable to cover all cases (ibid.). But subjectively perceived structures based on constructed boundaries can be very complex.

6 Conclusion

This chapter has highlighted 'drivers' for object-based image analysis as well as some of the 'responses' as they became key characteristics of OBIA. The aim of the opening chapter for the first book section asking "Why object-based image analysis?" was to put forward profound ideas and basic concepts of this new approach, as it was to discuss the tasks challenging it. Both motivation and requirements for OBIA were presented in the light of a world of increasing complexity to be addressed by multiple solutions for global monitoring issues. Conceptual elements of this new approach were discussed considering spatial as well as perceptual aspects. Drawing from those, methodological implications have been pointed out in terms of adaptations and further development of traditional methods, empowered by a successful integration of GIS and remote sensing techniques. The subsequent chapters of this section will complement these views with many additional aspects presented from various angles and backgrounds.

Acknowledgements

Foundation of the conceptual work presented here was laid in various research projects carried out at the Centre for Geoinformatics (Z_GIS). The author highly appreciates the many valuable inputs from members of the OBIA research group at Z_GIS. The author also expresses gratitude and thanks to the reviewers for valuable comments and corrections.

References

Allen TFH & TB Starr (1982) Hierarchy. University of Chicago Press, Chicago.
Baatz M & A Schaepe (2000) Multi-resolution segmentation. An optimization approach for high-quality multi-scale image segmentation. In: J. Strobl, T. Blaschke & G. Griesebner (eds.), Angewandte Geographische Informationsverarbeitung XII, Wichmann Verlag, Heidelberg, pp. 12-23.
Benz U, P Hofmann, G Willhauck, I Lingenfelder & M Heynen (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-

ready information. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, pp. 239-258.

Biederman I (1987) Recognition-by-Components: A Theory of Human Image Understanding. Psychological Review 94(2):115-147.

Binnig G, M Baatz J Klenk & G Schmidt (2002) Will Machines start to think like humans? Europhysics News, vol. 33 (2)

Blaschke T (2005) A framework for change detection based on image objects. In: Erasmi, S. et al. (eds.), Göttinger Geographische Abhandlungen 113, pp. 1-9.

Blaschke T, P Zeil, J Strobl, S Lang, D Tiede, M Möller, G Triebnig, C Schiller, M Mittlböck, B Resch (2007) GMES: From research projects to operational environmental monitoring services. ISPRS Workshop High-Resolution Earth Imaging for Geospatial Information, May 29 – June 1, 2007, Hannover.

Blaschke T & J Strobl (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS, vol. 6, pp. 12-17.

Burnett C & T Blaschke (2003) A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling, pp. 233-249.

Burnett C, K Aavikso, S Lang, T Langanke & T Blaschke (2003) Mapping mire forests as hierarchy of objects. In: IUFRO conference "decision support for multiple purpose forestry, 23.-25.4. 2003. Vienna. CD-Rom

Campbell JB (2002) Introduction to Remote Sensing. The Guilford Press, New York.

Castilla, G & G Hay (2006) Uncertainties in land use data. Hydrology and Earth System Sciences, 3, 3439-3472.

Castilla GA Lobo & J Solana (2004) Size-constrained region merging (SCRM): a new segmentation method to derive a baseline partition for object-oriented classification. Proceedings of SPIE, 5239, pp. 472-482.

Congalton RG & K Green (1999) Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publishers, Boco Raton.

Egenhofer MJ (1994) Pre-Processing queries with spatial constraints. Photogrammetric Engineering & Remote Sensing, 60 (6), pp. 783-790.

Ganter B & R Wille (1996) Die Formale Begriffsanalyse. Springer, Berlin.

Groom G, CA Mücher, M Ihse, T Wrbka (2004) Remote sensing in landscape ecology. Landscape Ecology, pp. 1-18.

Hay G, T Blaschke, DJ Marceau & A Bouchard (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. Photogrammetry and Remote Sensing, vol. 1253, pp. 1-19.

Hofmann DD (1998) Visual Intelligence. How we create what we see. W. W. Norton & Company, Inc.

Hornsby K & MJ Egenhofer (2000) Identity-based change: a foundation for spatio-temporal knowledge representation. International Journal of Geographical Information Science, 14 (3), pp. 207-224.

Ibrahim AE (2000) An intelligent framework for image understanding. Available at: http://www.engr.uconn.edu/~ibrahim/publications/image.html (May, 2004).

Koch B, M Jochum, E Ivits & M Dees (2003) Pixelbasierte Klassifizierung im Vergleich und zur Ergänzung von pixelbasierten Verfahren. Photogrammetrie Fernerkundung Geoinformatik, vol. 3, pp. 195-204.

Lang S & D Tiede (2007) Conditioned Information For GMES Applications. Presentation at the Definiens GMES Research Award 2007, March 29.

Lang S & T Blaschke (2006) Bridging remote sensing and GIS - what are the most supportive pillars? In: S Lang, T Blaschke & E Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.

Lang S & T Langanke (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. In: Photogrammetrie, Fernerkundung, Geoinformation, 1/2006, 5-18

Lang S (2005) Image objects vs. landscape objects. Interpretation, hierarchical representation and significance, Salzburg, (unpublished PhD thesis).

Lang S (2002) Zur Anwendung des Holarchiekonzeptes bei der Generierung regionalisierter Segmentierungsebenen in höchstauflösenden Bilddaten. In: T. Blaschke (ed.), Fernerkundung und GIS: Neue Sensoren – innovative Methoden, Wichmann Verlag, Heidelberg, pp. 24-33.

Laszlo E. (1972) The Systems View of the World. George Braziller, New York.

Liedtke CE, J Bückner, O Grau, S Growe & R Tönjes (1997). AIDA: a system for the knowledge based interpretation of remote sensing data. Third International Airborne Remote Sensing Conference, 7-10 July 1997, Copenhagen, Denmark, Copenhagen.

Navon EJ (1977) Forest before trees: the precendence of global features in visual perception. Cognitive Psychology, vol. 9, pp. 353-383.

Niemeyer I & MJ Canty (2001) Object-oriented post-classification of change images. Proceedings of SPIE's International Symposium on Remote Sensing, Toulouse, 17-21 September 2001, vol. 4545, pp. 100-108.

Openshaw S (1984) The modifiable areal unit problem, Concepts and Techniques in Modern Geography 38: 41.

Pinz A. (1994) Bildverstehen. Springer, Wien.

Schöpfer E. & S Lang (2006) Object fate analysis - a virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: S Lang, T Blaschke & E Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.

Schöpfer E., S Lang & J Strobl (in press) Segmentation and object-based image analysis. In: C. Juergens & T. Rashed (eds.), Remote sensing of urban and suburban areas, Kluwer, Amsterdam. pages pending.

Schumacher J, S Lang, D Tiede, J Rietzke, J Trautner & D. Hölbling (2007) Einsatz von GIS und objekt-basierter Analyse von Fernerkundungsdaten in der regionalen Planung: Methoden und erste Erfahrungen aus dem Biotopinformations-und Managementsystem (BIMS) Region Stuttgarts. In: Angewandte Geografische Informationsverarbeitung XIX, pages pending.

Simon HA (1973) The Organization of Complex Systems. In: H.H. Pattee (ed.), Hierarchy Theory. The Challenge of Complex Systems, George Braziller, New York, pp. 1-28.
Singh A (1989) Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10, pp 989-1003.
Sowa JF (1999) Knowledge representation: logical, philosophical, and computational foundations. Brooks Cole Publishing Co, Pacific Grove.
Strahler AH, CE Woodcock & J Smith (1986) Integrating per-pixel classification for improved land cover classification. Remote Sensing of the Environment, vol. 71, pp. 282-296.
Straub BM & C Heipke (2004) Concepts for internal and external evaluation of automatically delineated tree tops. International Archives of Photogrammetry and Remote Sensing, vol. 26 (8/W2), pp. 62-65.
Szaramovicz M (2004) "Hierarchy" und "Scale" - Interessante Denkmuster für die Landschaftsplanung?. Naturschutz und Landschaftsplanung, 36 (1), 21-26.
Tiede D & S Lang (2007) Analytical 3D views and virtual globes - putting analytical results into spatial context. ISPRS Joint Workshop Visualization and Exploration of Geospatial Data, June 27-29, Stuttgart.
Tiede D & Ch. Hoffmann (2006) Process oriented object-based algorithms for single tree detection using laser scanning data. EARSeL-Proceedings of the Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb 2006, Vienna.
Tiede D, S Lang & C.Hoffmann (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: S. Lang, T. Blaschke and E. Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.
Wertheimer M (1925) Drei Abhandlungen zur Gestaltpsychologie. Verlag der Philosophischen Akademie, Erlangen.
Wu J (1999) Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing, vol. 25 (4), pp. 367-380.
Zeil P, S Lang, E Schöpfer, D Tiede, T Blaschke (in press) Monitoring for Human Security – reflecting the ‚S' in GMES, EARSeL Workshop GMES, Bolzano, 7-8 June, 2007.