CHAPTER 3
OBJECT BASED URBAN FEATURE ANALYSIS

3.1 INTRODUCTION

Important semantic information are necessary to interpret the high resolution satellite images rather than pixel information. A meaningful image object and their mutual relationships improve the image classification. A multitude of additional information that can be extracted from image data after segmentation provides the possibility to handle very high resolution data. This chapter discuss about the segmentation and its algorithm, feature description and its classification.

3.2 OBJECTS AND THEIR BEHAVIORS IN THE URBAN PLANNING CONTEXT

The definition of objects in an urban planning context may be to look at what features are currently stored and processed in existing GIS such as settlements, buildings, roads, green areas, water bodies, industrial areas, etc and these can be treated as objects. However, some of them are not likely to appear on a map at the same time.

By taking a close look at these objects, one can see that some of them are physical entities that have physical properties and physical boundaries, such as buildings, green spaces, and water bodies. And there are conceptual entities that consist of other physical entities, often with fuzzy boundaries, such as residential areas, commercial areas and industrial areas,
which directly relate to land use. A residential area may consist of buildings, gardens, footpaths, small lakes or canals, etc. One can notice that certain entities can be treated as physical entities in some cases but considered as conceptual entities in others. For instance, a lake classified as water surface in landcover classification may, in landuse classification, be identified as recreational landuse if it is located in a park, or as a fishing pool if it is located outside the built-up area. Therefore, an entity may ‘behave’ differently in different circumstances.

3.3 HIERARCHY OF PLANNING

The hierarchy of different planning is briefly discussed below

3.3.1 Regional Planning

Regional planning aims at the reasonable structure and spatial distribution of production elements at the regional scale. It deals mainly with abstract entities such as human settlements, industrial zones, and transportation networks. The central location, physical size and the spatial coverage of its influence zone on the surrounding regions are the main features to be modeled in regional planning.

3.3.2 Master Planning

Master planning aims at the sound spatial and sectional distribution of land in urban built up areas and surrounding regions.
3.3.3 Detailed Planning

Detailed planning deals mainly with organic spatial arrangement at the neighborhood level to meet the certain functions assigned to each neighborhood block.

3.4 OBJECTS AT DIFFERENT HIERARCHICAL LEVELS OF URBAN AND REGIONAL PLANNING

3.4.1 Objects at the Regional Planning Level

There are three types of objects at the regional planning level (via) point objects, line objects and area objects, corresponding to settlements, transportation networks and influence zones.

3.4.2 Objects at the Master Planning Level

There are three types of objects at the master planning level (via) point objects and line objects, which are mainly for providing spatial references, and area objects (land-use), which are the spatial partitions of major landuse classes as represented in a 2D space.

3.4.3 Objects at the Detailed Planning Level

The objects at the detailed planning level are very similar to the objects at the master planning level. The main differences are small spatial units and greater specificity in landuse functions. Therefore, landuse objects at the master planning level can be generated from landuse objects at the detailed planning level by merging objects from specific landuse classes into major landuse classes and dissolving the boundaries between neighbouring
objects within the same major landuse class. Landuse objects at the detailed planning level can be disaggregated or specified from landuse objects at the master planning level.

3.4.4 Objects at the Landcover Level

Physical entities such as buildings, roads, green spaces and water bodies are represented by objects at the landcover level. This is a fundamental base for landuse classification. Each object can be an element of a landuse class. Landcover objects can be used as an indication for determining the spatial extent of a landuse unit.

3.5 OBJECT TYPES

Objects with similar properties or similar behaviors are organized types. Similar behaviors can be identified according to various criteria or perspectives such as spatial extent and abstraction level. Types of objects will be discussed in the following subsections, according to different perspectives. The different types of object concerning different abstraction levels are briefly discussed below.

3.5.1 Elementary Objects (images or field data)

Pixels are regarded as elementary objects that have uniform geometric properties. Elementary objects share many methods or operations of pixel – or raster-based processing, such as filtering, convolution, classification.

There are two types of elementary objects, one relating to images or field data and one relating to object fields and there is one main difference
between the two. The images or field data take Digital Number (DN) values form pixels and these DN values usually range from 0 to 255 or actual height values of laser data. The DN values of an object field are taken mainly from membership functions according to the characteristics of the object. The DN values of an object field are taken from the Boolean value 0(false) or 1(true) for crisp objects or a real value from 0 to 1 for fuzzy objects.

3.5.2 Objects at the Landcover Level

Objects at the landcover level have object ID, geometric properties such as location, size, shape, orientation, as well as class-related attributes such as class name or class ID, mean value and standard deviation of membership functions in this class.

3.5.3 Objects at the Landuse Level

Objects at the landuse level have object ID, geometric properties such as location, size, shape, orientation, as well as class-related attributes such as class name or class ID, composition and proportion of landcover types contained, number of buildings held by a landuse object.

One of the advantages of object-based image analysis is the multitude of additional information that can be derived from image objects versus the amount of information available from individual pixels. A pixel typically contains a vector of information representing each band or layer in a data set. In the case of digital imagery, the spectral response information is related as digital numbers (DN). In contrast, image objects are composed of multi-pixel groups, enabling the calculation of aggregative statistics, such as mean and standard deviation, from an object is underlying DNs. In addition to spectral-based information, information based on object size, shape, and
context can be calculated, and the information pertaining to an object sub or super-objects if a multi-level image object hierarchy has been created.

3.6 IMAGE SEGMENTATION

The basic processing units of object-oriented image analysis are segments, or image objects, and not single pixels. Advantages of object-oriented analysis are meaningful statistic and texture calculation, an increased uncorrelated feature space using shape (e.g. length, number of edges, etc.) and topological features (neighbor, super-object, etc.), and the close relation between real-world objects and image objects. This relation improves the value of the final classification and cannot be fulfilled by common, pixel-based approaches.

Directly connected to the representation of image information by means of objects is the networking of these image objects. Whereas the topological relation of single, adjacent pixels is given implicitly by the raster, the association of adjacent image objects must be explicitly worked out, in order to address neighbor objects. In consequence, the resulting topological network has a big advantage as it allows the efficient propagation of many different kinds of relational information.

Since the last two decades, processing power of affordable computers allows image processing and image segmentation. Therefore, these methods become applicable for operational remote sensing image analysis. The major advantages in this area were derived in studies for sea-ice analysis (Daida et al 1990), object-oriented image matching and certain approaches for data compaction (Ghassemian and Landgrebe 1988).
Segmentation is the subdivision of an image into separated regions. Few of the methods lead to qualitatively convincing results which are robust and applicable under operational settings. There are many segmentation methods in the research field, which are split into two regions, such as top-down or knowledge-driven and bottom-up or data-driven approach.

In top-down approach, by fixing the end product of the image, the extraction process starts and in bottom-up approach, the segments are generated based on a set of statistical methods and perform the processing. It is starting with one pixel of an object and in subsequent steps the smaller objects are merged into larger ones (more pixels). The growing decision is based on the local homogeneity criteria describing the similarity of adjacent image objects in terms of size, distance, texture, spectral similarity and form (Baatz and Schape 2000 and Schieve 2002).

As such, bottom-up methods can also be seen as a kind of data abstraction or data compression. But, as with clustering methods, in the beginning the generated image segments have no meaning, they can better be called: image object primitives. It is up to the user to determine what kind of real world objects the generated image objects represent. The basic difference between both approaches is: top-down methods usually lead to local results because they just mark pixels or regions that meet the model description, whereas bottom-up methods perform a segmentation of the complete image. They group pixels to spatial clusters which meet certain criteria of homogeneity and heterogeneity. In order to obtain image object primitives as basic processing units, the object-oriented approach to image analysis requires complete segmentation of an image. Therefore, a rough overview is given of the most common bottom-up approaches to image segmentation.
Some of the simplest approaches are all types of global thresholding. The spectral feature space is separated into subdivisions, and pixels of the same subdivision are merged when locally adjacent in the image data. Typically, this method leads to results of relatively limited quality. Oversegmentation and undersegmentation – i.e., separating into units which are too small or merging regions that do not belong to each other – take place easily without good control of meaningful thresholds. Local contrasts are not considered or not represented in a consistent way and the resulting regions can widely differ in size.

Region growing algorithms cluster pixels starting from a limited number of single seed points. These algorithms basically depend on the set of given seed points and often suffer from a lack of control in the break-off criterion for the growth of a region.

Often used in operational applications are different types of texture segmentation algorithms. They typically obey a two-stage scheme (Mao and Jain 1992, Hofmann et al 1998). In the modeling stage characteristic features are extracted from the textured input image and range from spatial frequencies (Hofmann et al 1998), MRF models (Mao and Jain 1992) and co-ocurrence matrices (Haralick et al 1973) to wavelet coefficients (Salari and Ling 1995), wave packets (Laine and Fan 1996) and fractal indices (Chaudhuri and Sarkar 1995). In the optimization stage features are grouped into homogeneous segments by minimizing an appropriate quality measure. This is most often achieved by a few types of clustering cost functions (Mao and Jain 1992, Hofmann and et al 1998 and Manjunath and Chellappa 1991). Although texture segmentation leads to reproducible, for specific applications excellent and sometimes high speed results, they are mostly applicable to a limited number of types of image data, texture types and problems. Texture often
must be very regular to be recognized and the results cannot be achieved on any chosen scale.

Common alternatives are knowledge-based approaches and they try to incorporate knowledge derived from training areas or other sources into the segmentation process (Gorte 1998). These approaches mostly perform a pixel-based classification, based on clustering in a global feature space. Segments are produced implicitly after classification and simply by merging all adjacent pixels of the same class. In doing so, these approaches are typically not able to separate different units or objects of interest of the same classification. Furthermore, the information on which classification can act typically is limited to spectral and filter derivates.

A further relatively common procedure is watershed segmentation (Wegner et al 1997). It got its name from the manner in which the algorithm segments regions into catchments basins. Typically, the procedure first transforms the original data into a gradient image. The resulting gray tone image can be considered as a topographic surface. If flood this surface from its minima and if prevent the merging of the waters coming from different sources, partition the image into two different sets: the catchments basins and the watershed lines. The catchments basins should theoretically correspond to the homogeneous gray level regions of this image. This method works for separating essentially convex and relatively smooth objects of interest that even may touch slightly in relatively homogeneous image data. When it works, it is convenient, fast and powerful. However, for remote sensing data, which typically contain a certain noise and not always strong contrasts, this method is typically not able to achieve appropriate results.
3.6.1 Creation of an Object

In subsequent steps, smaller image objects are merged into bigger ones. Throughout this pair wise clustering process, the underlying optimization procedure minimizes the weighted heterogeneity of \( n \cdot h \) of resulting image objects, where \( n \) is the size of a segment and \( h \) is a parameter of heterogeneity. In each step, that pair of adjacent image objects is merged which results in the smallest growth of the defined heterogeneity. If the smallest growth exceeds the threshold defined by the scale parameter, the process stops. The above steps are followed to optimize the segmentation parameters.

3.6.2 Segmentation Parameters

There are five parameters which are having functions for defining the heterogeneity while doing the segmentation in the object oriented image processing. Such parameters are scale, color, shape, compactness and smoothness (Baatz and Schaâpe, 1999). A brief description of each parameter is given below.

3.6.3 Algorithms for defining the parameters

3.6.3.1 Definition of heterogeneity

Heterogeneity in eCognition considers as primary object features color and shape. The increase of heterogeneity ‘\( f \)’ has to be less than a certain threshold.

\[
f = W_{color} \cdot \Delta h_{color} + W_{shape} \cdot \Delta h_{shape}
\]

\( W_{color} \in [0,1], W_{shape} \in [0,1] \)

\( W_{color} + W_{shape} = 1 \)
where

- \( w_{\text{color}} \) - Weightage for color factor
- \( w_{\text{shape}} \) - Weightage for shape factor
- \( \Delta h_{\text{color}} \) - Color heterogeneity
- \( \Delta h_{\text{shape}} \) - Shape heterogeneity

The weight parameters \((w_{\text{color}}, w_{\text{shape}})\) allow adapting the heterogeneity definition to the application. The spectral heterogeneity allows multi-variant segmentation by adding a weight \( w_c \) to the image channels \( c \). Difference in spectral heterogeneity \( \Delta h_{\text{color}} \) is defined as following:

\[
\Delta h_{\text{color}} = \sum_c W_c \left( \eta_{\text{merge}} \sigma^c_{\text{merge} \cdot} - \left( \eta_{\text{obj}_1} \sigma^c_{\text{obj}_1} + \eta_{\text{obj}_2} \sigma^c_{\text{obj}_2} \right) \right) \quad (3.2)
\]

where

- \( \eta_{\text{merge}} \) - number of pixels within merged object
- \( \eta_{\text{obj}_1} \) - number of pixels in object 1
- \( \eta_{\text{obj}_2} \) - number of pixels in object 2
- \( \sigma^c \) - standard deviation within object of channel

Subscripts \( \text{merge} \) refer to the merged object, object 1 and object 2 are prior to merge, respectively. The shape heterogeneity \( \Delta h_{\text{shape}} \) is a value that describes the improvement of the shape with regard to smoothness and compactness of an object’s shape.

\[
\Delta h_{\text{shape}} = w_{\text{compt}} \cdot \Delta h_{\text{compt}} + w_{\text{smooth}} \cdot \Delta h_{\text{smooth}} \quad (3.3)
\]

\[
\Delta h_{\text{smooth}} = \eta_{\text{merge}} \frac{l_{\text{merge}}}{b_{\text{merge}}} - \left( \eta_{\text{obj}_1} \frac{l_{\text{obj}_1}}{b_{\text{obj}_1}} + \eta_{\text{obj}_2} \frac{l_{\text{obj}_2}}{b_{\text{obj}_2}} \right)
\]
\[ \Delta h_{\text{compt}} = \eta_{\text{merge}} \frac{l_{\text{merge}}}{\sqrt{\eta_{\text{merge}}}} - (\eta_{\text{obj}_1} \frac{l_{\text{obj}_1}}{\sqrt{\eta_{\text{obj}_1}}} + \eta_{\text{obj}_2} \frac{l_{\text{obj}_2}}{\sqrt{\eta_{\text{obj}_2}}}) \]

where

- \( l_{\text{merge}} \) - Length of merging distance
- \( b_{\text{merge}} \) - Perimeter of object’s bounding box.

The scale parameter is the stop criterion for optimization process. Prior to the fusion of two adjacent objects, the resulting increase of heterogeneity ‘f’ is calculated. The scale parameter determines the threshold ‘t’ and it checks the result, if \( t = W \) (scale parameter), then no further fusion takes place and the segmentation stops. The segmented image is shown in fig 3.1. This produces extensive semantic information for an object in terms of spectral, shape and neighborhood which is shown in fig 3.2. The larger the scale parameter, the more objects will be fused and the larger the objects grow.

### 3.7 CLASSIFICATION

Classification means assigning a number of objects to a certain class according to the classes description. Thereby, a class description is a description of the typical properties or conditions the desired classes have. The objects then become assigned (classified) according to whether they have or have not met these properties/conditions. In terms of database language one can say the feature space is segmented into distinct regions which leads to a many-to-one relationship between the objects and the classes. As a result each object belongs to one definite class or to no class. Classic classifiers in remote sensing (e.g., maximum-likelihood, minimum-distance or parallelepiped) thereby assign a membership of 1 or 0 to the objects, expressing whether an object belongs to a certain class or not.
Such classifiers are also called hard classifiers since they express the objects’ membership to a class only in a binary manner. In contrast, soft classifiers (mainly fuzzy systems and/or Bayes classifiers) use a degree of membership/a probability to express an object’s assignment to a class. The membership value usually lies between 1.0 and 0, where 1.0 expresses full membership/probability (a complete assignment) to a class and 0 expresses absolutely non-membership/improbability. Thereby the degree of membership/probability depends on the degree to which the objects fulfill the class-describing properties/conditions.

The main advantage of these soft methods lies in their possibility to express uncertainties about the class’ descriptions. It makes it also possible to express each object’s membership in more than just one class or the probability of belonging to other classes, but with different degrees of membership or probabilities. With respect to image understanding these soft classification results are more capable of expressing uncertain human knowledge about the world and thus lead to classification results which are closer to human language, thinking and mind. In other words: soft classifiers are more honest than their hard counterparts. But many applications using land use or land cover information are unable to handle soft classification results. Thus, soft classification results must be hardened, which can lead to doctored classification truths and accuracies.

In the classification methods, they can basically be separated into supervised and unsupervised methods. While supervised methods ask the user how the desired classes look, unsupervised methods are almost user independent. They rather can be seen as statistical grouping methods, sorting objects by their properties into clusters with similar properties. Since unsupervised methods work almost automatically, supervised methods have to be trained by the user – usually either by taking samples or by describing
the classes’ properties. Therefore, the class-describing information must be as accurate, representative and complete as possible, which is in most cases effectively impossible.

Hence, a class description can only be a general estimation of the desired classes properties. Estimating the properties also means assuming a more or less known uncertainty about the class description or a known vagueness about the properties’ measured values. Formulating these uncertainties can only be achieved using soft classifiers. Comparing unsupervised and supervised classification methods, each has its advantages and disadvantages.

Unsupervised methods are noticeably faster than supervised ones, but since they are just a special way of sorting algorithms, their results have to be interpreted by the user – which can be tough in some cases and lead to numerous repetitions of the classification with slightly adjusted parameters. Another advantage of unsupervised classifiers is their ability to analyze the objects’ statistics completely and systematically. Thus, the results of an unsupervised classification can give useful indications of detectable classes, but, in general, formulating uncertainty is only possible if related to the classification parameters, not to the classes and their properties themselves. In contrast, supervised classification methods can be more labor intensive since the user has to describe the classes’ properties either explicitly or by taking samples as typical representatives.

3.7.1 Classification Systems

The most powerful soft classifiers are classifiers based on fuzzy systems. Fuzzy logic is a mathematical approach to quantifying uncertain statements. The basic idea is to replace the two strictly logical statements
“yes” and “no” by the continuous range of \([0…1]\), where 0 means “exactly no” and 1 means “exactly yes.” All values between 0 and 1 represent a more or less certain state of “yes” and “no.” Thus, fuzzy logic is able to emulate human thinking and take into account even linguistic rules. Fuzzy classification systems are well suited to handling most vagueness in remote sensing information extraction. Parameter and model uncertainties are considered as using fuzzy sets defined by membership functions. Instead of the binary “true” and “false” multivalued fuzzy logic allows transitions between “true” and “false.” Additionally, there are more or less strict realizations of the logical operations “and” or “or.”

The output of a fuzzy classification system, where the membership degree to each land cover or land use class is given for each object. This enables detailed performance analysis and gives insight into the class mixture for each image object. This is a major advantage of soft classification. The maximum membership degree determines the final classification to build an interface to crisp (boolean) systems. Fuzzy systems consist of the following steps, such as fuzzification and fuzzy rule base.

a) **Fuzzification**

Fuzzification describes the transition from a crisp system to a fuzzy system. It assigns a membership degree (membership value) between 0 and 1 to each feature value. The membership value is defined by membership function. Depending on the shape of the function, the transition between “yes” and “no” can be crisp (for a rectangular function) or fuzzy (Figure 3.1, set M). The set of feature values which produce a membership value higher than 0 can be called a fuzzy set. In general, the broader the membership function, the vaguer the underlying concept; the lower the membership values, the more uncertain is the assignment of the set.
Figure 3.1 Crisp and fuzzy set of different features

The combining of different features within a fuzzy system is always done after the feature is fuzzified. Therefore, all input values for fuzzy combinations are in the range between 0 and 1, independent of the dynamic of the originally crisp features. This simplifies working in a high-dimensional feature space with different dynamics and features of various types, e.g., backscatter from different sensors, geographic information, texture information and hierarchical relations. For successful classification a deliberate choice of membership function is crucial. This allows the introduction of expert knowledge into the system. The better the knowledge about the real system is modeled by the membership functions, the better the final classification result.
Figure 3.2 Fuzzy set of different features

It is possible to define more than one fuzzy set on one feature, e.g., to define the fuzzy sets low, medium and high for one object feature. The more the memberships overlap, the more objects are common in the fuzzy sets and the vaguer the final classification. Figure 3.2 shows three fuzzy sets defined for the feature $x$: low, medium and high. They are characterized by overlapping triangular membership functions. For an image object with a feature value of $x = 70$, the membership to class low is 0.4, to class medium is 0.2 and to class high is 0. If the feature value $x$ equals to 200, the membership to the classes is 0, 0, 0.8, respectively.

b) Fuzzy rule base

Fuzzy rule base is a combination of fuzzy rules, which combine different fuzzy sets. The simplest fuzzy rules are dependent on only one fuzzy set. Fuzzy rules are “if – then” rules and if a condition is fulfilled, an action takes place. Referring to fig. 2, the following rule could be defined: “If feature $x$ is low, “then” the image object should be assigned to land cover $W$. 

![Diagram showing fuzzy sets for different feature values with memberships to low, medium, and high sets at feature values 70 and 200.](image-url)
In fuzzy terminology this would be written: If feature x is a member of fuzzy set *low*, then the image object is a member of land cover *W*. According to the definition in fig. 2, in case feature value x = 70, the membership to land cover *W* would be 0.4, in case x = 200, the membership to land cover *W* would be 0.

To create advanced fuzzy rules, fuzzy sets can be combined. An operator returns a fuzzy value that is derived from the combined fuzzy sets and this value is derived depends on the operator. The basic operators are “and” and “or.” “and” represents the minimum, meaning that the minimum value of all sets defines the return value. “or” represents the maximum value, meaning that the maximum value of all sets defines the return value. The results are very transparent and ensure independence of the sequence of logic combinations within the rule base (A “and” B gives the same result as B “and” A). In addition a hierarchic structure following common logic (e.g., A “or” (B “and” C) equals (A “or” B) “and” (A “or” C)) can be created easily. A fuzzy rule base delivers a fuzzy classification, which consists of discrete return values for each of the considered output classes (figure 3.3). These values represent the degree of class assignment.

![Figure 3.3](image-url)  
*Figure 3.3  Fuzzy classification for the urban, water and vegetation classes*
\( \left( \mu_{\text{urban}} \text{ (object)} = 0.6, \mu_{\text{water}} \text{ (object)} = 0.8, \mu_{\text{vegetation}} \text{ (object)} = 0.3 \right) \)

It is considered that the fuzzy classification gives a possibility for an object to belong to a class. While classification is based on probability, it gives a probability that belong to a class. A possibility gives information on a distinct object and a probability relies on statistics and gives information on many objects. Whereas the probability of all possible events adds up to one, this is not necessarily true for possibilities.

3.8 CONCLUSION

From the above concept of object oriented approach in very high resolution imagery will be segmented using the color, shape, compactness, smoothness and scale parameter. By optimizing those objects the semantic information will be created. The fuzzy rule based classification will solve the fuzzification on the decision making process.