VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STAVEBNÍ
ÚSTAV GEODÉZIE

FACULTY OF CIVIL ENGINEERING
INSTITUTE OF GEODESY

EXTRACTION OF LANDSCAPE ELEMENTS FROM REMOTE SENSING DATA

EXTRAKCE KRAJINNÝCH PRVKŮ Z DAT DÁLKOVÉHO PRŮZKUMU

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE BC. JAKUB FERENCZ
AUTHOR

VEDOUCÍ PRÁCE doc. Ing. VLASTIMIL HANZL, CSc.
SUPERVISOR

BRNO 2013

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA STAVEBNÍ

Studijní program	N3646 Geodézie a kartografie
Typ studijního programu	Navazující magisterský studijní program s prezenční formou studia
Studijní obor	3646T003 Geodézie a kartografie
Pracoviště	Ústav geodézie

ZADÁNÍ DIPLOMOVÉ PRÁCE

Diplomant	Ferencz Jakub
Název	Extrakce krajinných prvků z dat dálkového průzkumu
Vedoucí diplomové práce	doc. Ing. Vlastimil Hanzl, CSc.
Datum zadání diplomové práce	30. 11. 2012
Datum odevzdání diplomové práce	24. 5. 2013

V Brně dne 30. 11. 2012

doc. Ing. Josef Weigel, CSc. Vedoucí ústavu	prof. Ing. Rostislav Drochytka, CSc. Děkan Fakulty stavební VUT

Podklady a literatura

Sborník XXII. kongresu ISPRS

Zásady pro vypracování (zadání, cíle práce, požadované výstupy)

Zabývejte se vypracováním vhodné metodiky pro extrakci krajinných prvků a budov z kombinace snímku vysokýho rozlišení a Lidaru pro účely hodnocení vývoje v záujmové oblasti. Využijte objektově orientované klasifikace (OBIA). Zhodnoťte přesnost a využitelnost navrženého postupu.

Struktura bakalářské/diplomové práce

VŠKP vypracujte a rozčleňte podle dále uvedené struktury:

1. Textová část VŠKP zpracovaná podle Směrnice rektora "Úprava, odevzdávání, zveřejňování a uchovávání vysokoškolských kvalifikačních prací" a Směrnice děkana "Úprava, odevzdávání, zveřejňování a uchovávání vysokoškolských kvalifikačních prací na FAST VUT" (povinná součást VŠKP).
2. Přílohy textové části VŠKP zpracované podle Směrnice rektora "Úprava, odevzdávání, zveřejňování a uchovávání vysokoškolských kvalifikačních prací" a Směrnice děkana "Úprava, odevzdávání, zveřejňování a uchovávání vysokoškolských kvalifikačních prací na FAST VUT" (nepovinná součást VŠKP v případě, že přílohy nejsou součástí textové části VŠKP, ale textovou část doplňují).

....

doc. Ing. Vlastimil Hanzl, CSc.
Vedoucí diplomové práce

Abstract

This master thesis deals with a classification technique for an automatic detection of different land cover types from combination of high resolution imagery and LiDAR data sets. The main aim is to introduce additional post-processing method to commonly accessible quality data sets which can replace traditional mapping techniques for certain type of applications.
Classification is the process of dividing the image into land cover categories which helps with continuous and up-to-date monitoring management. Nowadays, with all the technologies and software available, it is possible to replace traditional monitoring methods with more automated processes to generate accurate and cost-effective results.

This project uses object-oriented image analysis (OBIA) to classify available data sets into five main land cover classes. The automate classification rule set providing overall accuracy of 88% of correctly classified land cover types was developed and evaluated in this research. Further, the transferability of developed approach was tested upon the same type of data sets within different study area with similar success – overall accuracy was 87%. Also the limitations found during the investigation procedure are discussed and brief further approach in this field is outlined.

Abstrakt

Tato diplomová práce se zabývá klasifikačními metodami pro automatickou detekci různých krajinných prvků z kombinace snímků vysokého rozlišení a LiDAR dat. Hlavním cílem je představit další možnou metodu zpracování pro volně přístupná data, která může nahradit tradiční mapování pro specifické aplikace.
Klasifikace snímků je metoda, která dělí snímek do různých kategorií a zajišťuje tak stálý a aktuální monitoring. V dnešní době, s přístupem k novým technologiím a softwarům, je možné postupně nahradit tradiční monitorovací postupy plně automatizovaným procesem, jehož výstupem jsou přesná a levná data.

V tomto projektu je použita objektově orientovaná analýza snímků (OBIA) pro klasifikaci pěti terénních typů z použitých dat. Byl vyvinut a testován automatický klasifikační proces, který poskytuje 88 % úspěšnost správného přiřazení terénního typu. Následně, přenosnost tohoto postupu byla testována v jiné lokalitě s podobným úspěchem - přesnost správného přiřazení byla 87 %. Závěr práce se zabývá problémy, které se vyskytly v průběhu tvorby tohoto klasifikačního postupu a nastiňuje další varianty, které by bylo možné použít pro zlepšení celého procesu.

Keywords

classification, OBIA, eCognition, land cover, nDSM, high-resolution imagery, LiDAR

Klíčová slova

klasifikácia, OBIA, eCognition, krajinné prvky, nDSM, snímky vysokého rozlišení, LiDAR

Bibliography

FERENCZ, Jakub. Extraction of Landscape Elements from Remote Sensing Data. Brno, 2013. 73 p., 13 p. of appendices. Master's thesis. Brno University of Technology, Faculty of Civil Engineering, Institution of Geodesy. Supervisor doc. Ing. Vlastimil Hanzl, CSc..

Bibliografická citácia

FERENCZ, Jakub. Extrakce krajinných prvků z dat dálkového průzkumu. Brno, 2013. 73 s., 13 s. příl. Diplomová práce. Vysoké učení technické v Brně, Fakulta stavební, Ústav geodézie. Vedoucí práce doc. Ing. Vlastimil Hanzl, CSc..

Author’s statement:

I claim that I have completed the whole master thesis on my own with all used references named.

10/06/2013 Jakub Ferencz

Acknowledgements

The writing of this master thesis has been one of the most significant academic challenges I have ever had to face. Without the support, patience and guidance of the following people, this study would not have been completed.

Firstly, I would like to thank my supervisor at Brno University of Technology, doc. Ing. Vlastimil Hanzl, CSc. for his time, guidance and valuable notes.
I owe sincere and earnest thankfulness to, my friend and recent colleague, Rachel Alice Dalley (University of Melbourne, Australia) for her help, support and patience on both an academic and personal level. Especially for her additional time spending on correcting my 'international weird-sound' English sentences.
I would like to thank Txomin Hermosilla Gómez (University of British Columbia, Canada) for providing his valuable advices at the beginning of this research and for affording his research papers which he has recently done in the classification field.
Last, but by no means least, I thank Ing. Markéta Potůčková, Ph.D. (Charles University in Prague, Czech Republic) for introducing me image classification topic back during my studies in Denmark, and for providing access to export classification results which was necessary for the further steps of this project.

Contents

List of Figures and Tables

1 Introduction

‘Have you ever thought about how to replace a traditional mapping?’

I have, and I have realized that further use of available remote sensed data can achieve this. When I first began experimenting with extraction of landscape elements from remotely sensed data sets, I hypothesized that extraction would require constant interaction of skilled operator, which devalue whole idea behind ‘fast and effective’ mapping approach. However, I soon found that like many other applications today, developing an automated process can significantly cut the time-spending and keeps it simple for end user.

Remote sensing (RS), the acquisition of data from afar, has become a more cost-effective data acquisition technique than field surveying when conducting ‘long-term and broad area analysis’, simulating growth in a significant amount of sectors than before. (Shiba and Itaya 2006) Nowadays there is a considerable demand for fast and efficient large scale mapping, especially for planning processes and decision making in industries such as urban sprawl controlling, natural disaster management and real estate. It calls for need of remote sensing data to be converted into tangible information which can be utilised in conjunction with other data sets, often within widely used Geographic Information Systems (GIS). (Blaschke 2010)

This master thesis develops a methodology to delineate different land cover types utilising RS data processed with classification methods. The outcome provides an overview on application bridging RS, mapping and further GIS use with a focus on automated process and repeatable and transferable solutions.

Classification is the categorisation of the image elements into different land cover classes such as buildings and vegetation. The conventional pixel based classification approaches, however, have limitations when dealing with high resolution imagery, because these methods do not work with characteristics of neighbouring pixels. Over the last years, advances in computer technology lead to synchronize and harmonize approaches related to acquisition, processing, feature recognition and advanced image analysis. (Navlur 2006) For exploiting image information more ‘intelligently’, the object based image analysis (OBIA) approach can contribute to powerful automatic and semi-automatic analysis for most remote sensing applications. (Benz at al 2003) This study examines the results of applying OBIA to combination of very high resolution orthophoto and LiDAR data. There has been much research already undertaken into land cover detection with using multi-source data fusion, especially imagery and LiDAR data (Baatz and Schäpe 2000), (Rottensteiner *et al*. 2007), (Lee *et al.* 2008), (Salah *et al.* 2009).

There are five chapters in this master thesis in sum. Chapter 2 introduces methodology behind remote sensing, outline image analysis techniques and summarized recent research in the field of image classification. Chapter 3 provides results and initial steps which led to creation of classification rule set. Chapter 4 evaluates results and discusses problems during process of development. Final chapter is devoted to conclusion.

2 Pre Analysis Methodology

The aim for this thesis is to investigate the use of remotely sensed data for the different types of land cover detection, to subsequently extract them for further GIS use. In this section reader will find description of following:

- Different remote sensing data acquisition overview
- Introduction to main image classification methods
- Land cover definition
- Outline of research done in image classification field so far

The chapter will conclude with the results of the Pre Analysis as well as an explanation of why the chosen data sets, computing methods and software will be used.

There are multiple options for land cover recognition using remote sensing technology depending on amount of automation and accuracy required as well as data and computing software available. Further knowledge will be gained while exploring this area and subsequently described in this chapter. This initial investigation will determine which method will be chosen for further study and research which will then be described in the Investigation (Chapter 3).

Focusing on the detection of land cover features was chosen because when an automatic delineation of different man-made or natural features can be generated, geospatial information can be analyzed relative to the desired theme.
For example:

- Development programs – road building, zoning, urban planning, recognition of uncoordinated urban-sprawl
- Land policy - with continuous updating of remotely sensed data sets, councils could monitor illegal and unpermitted buildings or deforestation
- Agricultural program - farmers could calculate their true planted areas, influance between land cover and land use and feedbacks between these factors, including past human activity can be observed.

2.1 Introduction into the Remote Sensing Acquisition

This section will introduce Remote Sensing (RS) as the technique of obtaining information about Earth's features through instruments which are not in direct physical contact with the objects, followed by outlining the types and different approaches of data acquisition used by this technique without consideration of basic principles and theory provides a narrow. Since there are many sources available about general principles of RS, only the basics will be outlined here. The focuse of this paper will be put on post-processing of remotely sensed datasets.

There are many different definitions* which clarify RS with a common conclusion - looking but not touching. Baumann's (2008) description says

> *“ Remote Sensing is the art and science of recording, measuring and analyzing information about a phenomenon from a distance “*

or according Abdulrahman (2010)

> *“ Remote Sensing is the process of acquiring data/ information about objects/substances not in direct contact with the sensor, by gathering its inputs using electromagnetic radiation or acoustical waves that emanate from the targets of interest.“*

In this stage, the author would like to note that in this paper collocation ‘Remote Sensing’ is used as a definition involving

- Remote Sensing – way of interpreting RS data
- Photogrammetry – paradigm dealing with geometrical properties of RS data and uses mathematical principles for processing the data
- Laser scanning – methodology working with spatial information from RS data and results in 3D models from direct measuring

The history of RS can be dated from the first invention of photography in 1826, followed by the first photographs taken from balloons in the 1850's. Wider usage starts from 1909 when cameras were mounted to airplanes to provide large land observations in shorter time. The next approach in this field was made in the technological progress during WWII in radar, infrared and microwave regions developments. Besides of all these important inventions, the modern/space RS era starts in the 1960's as a result of space race between USA and USSR and when the term “remote sensing” was first time used

* see Cambell 2011, p. 6, for a summary of the main definitions of RS that have been adopted over the last few decades

by Evelyn L. Pruitt, geographer formerly with the Office of Naval Research. In 1972, Earth resource satellite Landsat – 1 was launched. From this date on rapid advances in all fields related to RS have been made – digital image processing, development of hyperspectral sensors, and so on. Nowadays, the Earth's surface can be scanned by high-tech devices mounted on platforms such as helicopters, planes and satellites in order to study various sized areas of interest.

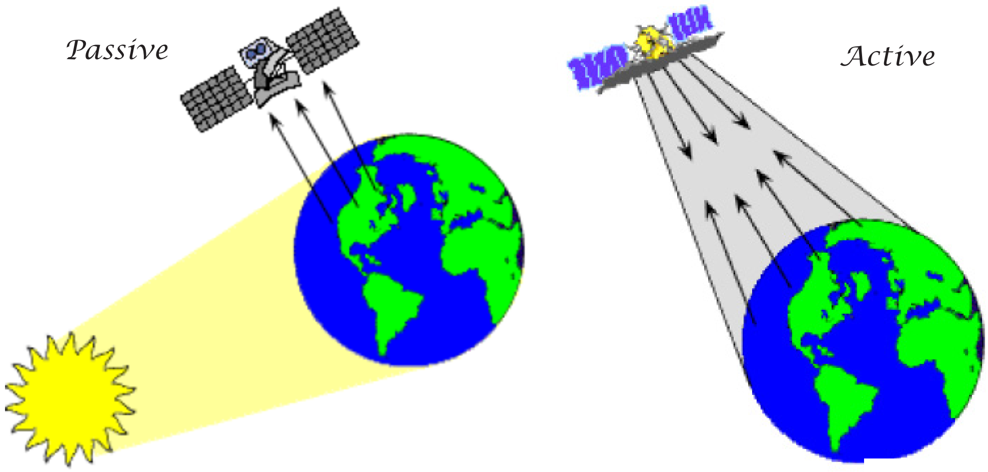


Figure 1. (Abdulrahman 2010)
Passive and Active sesning

The most important parts of devices for RS are the sensors, which can be divided into two main groups (Figure 1). The first is represented by sensors which need an external energy source – **passive** sensors. In most cases this source is the sun. This category includes film photography, infrared technology and radiometers, and works on the detection of reflective and emitted energy wave lengths from the objects. The second is an **active** sensor system which needs its own energy source to provide information about Earth's surfaces. Here, the device emits energy towards Earth and subsequently detects and measures the radiation that is reflected back or backscattered from the objects. As an example, radar sensor transmits waves and post-processes the echo coming back from the surface. (Baumann 2008)

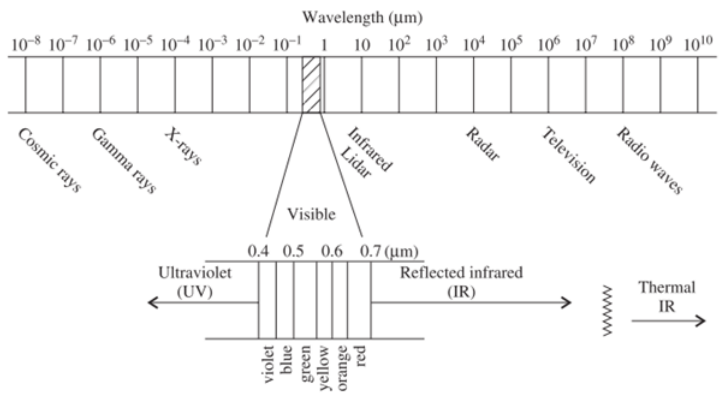


Figure 2. (Paine and Kiser, 2012)
The Electromagnetic spectrum

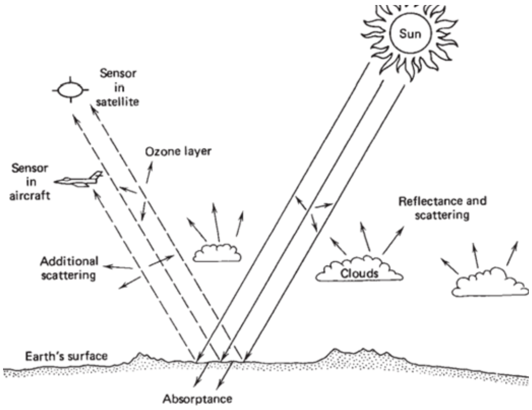


Figure 3. (Paine and Kiser, 2012)
The energy-flow profile

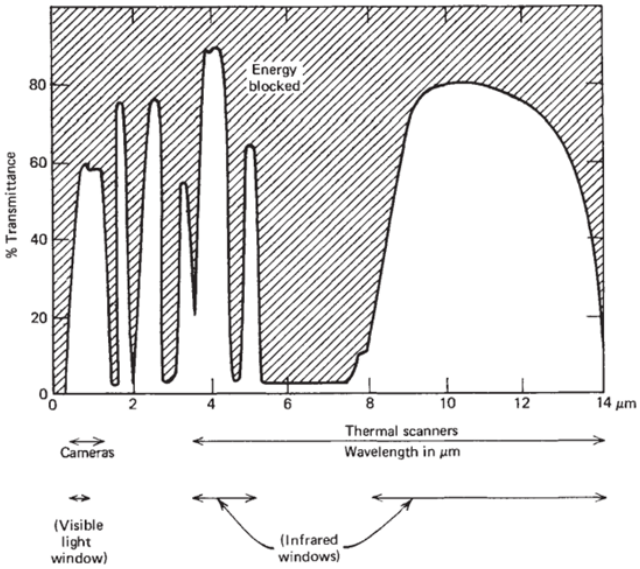


Figure 4. (Paine and Kiser, 2012)
Atmospheric windows
Figure shows wavelengths where the atmosphere is transparent. These specific EM intervals are called 'atmospheric windows'. Sensors are generally designed to record reflectance on these bands to gain more valuable information about land cover.

Electromagnetic Regions

Region Name	*Wavelength*	*Comments*
Gamma Ray	< 0.03 nm	Entirely absorbed by the Earth's atmosphere and not available for RS
X-ray	0.03-30 nm	Entirely absorbed by the Earth's atmosphere and not available for RS
Ultraviolet	0.03-0.4 μm	Wavelengths are absorbed by ozone in atmosphere
Photographic Ultraviolet	0.3-0.4 μm	Available for remote sensing the Earth. Can be imaged with cameras and sensors
Visible	0.4-0.7 μm	Available for remote sensing the Earth. Can be imaged with cameras and sensors
Near and Mid Infrared	0.7-3 μm	Available for remote sensing the Earth. Can be imaged with cameras and sensors
Thermal Infrared	3-14 μm	Available for remote sensing the Earth. This wavelength cannot be captured by film cameras
Microwave or Radar	0.1-100 cm	Longer wavelengths of this band can pass through clouds, fog and rain. Active sensors are used
Radio	> 100 cm	Not normaly used for remote sensing the Earth

Table 1. (Baumann 2008)

Whole principles of most sensors contain measuring the transmission of energy in diverse segments of the electromagnetic (EM) spectrum. For example, human eyes process this variation in the visible region of EM spectrum (390 to 750 μm for a typical eye). Sensors collect both visible and non-visible portions of the EM which can help to detect more valuable information from remotely sensed data than just the visible spectral interval. Another reason for detecting different EM spectrums is that the atmosphere is not completely transparent for every energy wave and some absorption can occur. The atmosphere transmissivity depends on the wavelength and type of radiation which pass through. All this is caused by the gases in the atmosphere which allow energy with certain wavelengths to pass through while preventing others. (Figure 4) The absorption is caused by water vapor, carbon dioxide and ozone. A sensor frequently captures information simultaneously in several regions of spectrum. These regions are called bands and are classified in nanometers or micrometers.

Currently, digital remote sensing devices work with two types of resolution: radiometric (spectral) and geometric (spatial).
Spectral resolution involves a number of levels where the sensor can store spectral information and depends on a bit number technology. It may vary on range from 0-255 (for 8 bit technology) to 0-65,535. These values characterize the degree of reflective or emitted energy collected by a sensor.
Spatial resolution defines the smallest area on the surface for which a sensor can record spectral information. Generally for imagery it is known in terms of a pixel (picture element) which defines a finer resolution in the imagery and is influenced by several things, especially by image scale.
In the field of application, there are probably hundreds of usages typical for remotely sensed data. To mention some which are directly related to this paper – applications in cartography, land use and land cover, city planning, forestry, grassland management, soil mapping and their derivation for all other kinds of surface research like archaeology, geomorphology, disaster warning assessment.

Before continuing to next chapter the reader should understand the physical background of the RS process. To gain a deeper insight into RS fundamentals and problematics the author recommends reading "Introduction to Remote Sensing" (Cambell 2011) or/and "Physical Principles of Remote Sensing" (Rees 2001).

2.1.1 Passive Sensors

This chapter will present a snapshot of passive sensors used for acquiring data about Earth's surface. This simple overview starts with aerial cameras, which are the oldest form of the remote sensing instruments, how they have changed dramatically in recent decades. The general concept of data collection and some of the main problems with this technology and their solutions will be mentioned as well. More discussion about passive sensors will follow as well as additional RS technologies which significantly improves information gain from the surface.

Aerial Imagery

> *The ideal mapping camera would be able to capture an image with accurate, consistent geometric relationship between point on the surface and its equivalent representation on the image.*

Aerial mapping acquisition relies on the elementary camera components common to the handheld cameras we have used for everyday photography – a lens to transmit light,

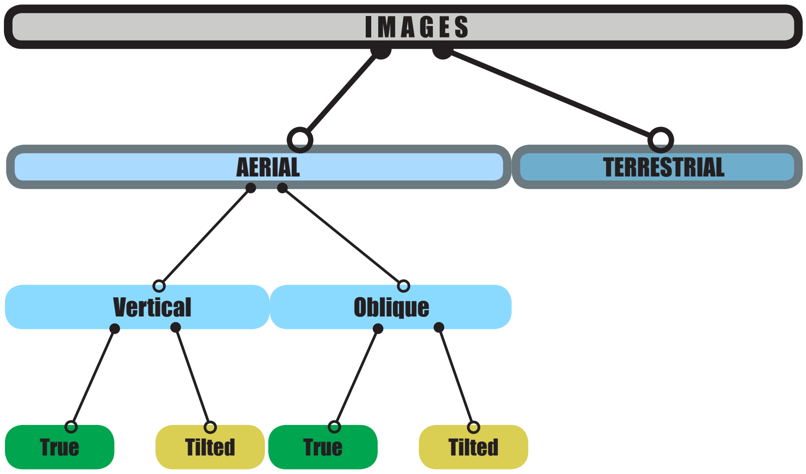


Figure 5.
Types of aerial images

light sensitive medium to record the image and a shutter for controlling entry of light and camera body. Historically, analog cameras hold problems of storage, transmission, searching and post-processing images. All this was simplified and solved by the digital development in the camera industry. Digital technologies store an image as data sequences in an individual pattern of brightness. Further, with the digital camera age comes know-hows that were not available during analog era which have radically increased aerial mapping methods. Aside from all these developments, links with positional and navigational systems and elaborate systems for annotating images need to be mentioned. For recording images, digital cameras apply either of two approaches used the most – Charged-Coupled Devices (known as CCDs chips) or Complementary Metal Oxide Semiconductors (known as CMOS) chip. Each alternative offers its own advantages and disadvantages which are more closely described in (Cambell 2011) and more advanced insights into the problematics can be found in (Coghill 2003).
Aerial images can be grouped according to their orientation to the ground in time of acquisition. (Figure 5) Vertical are called images which are in time of exposure truly vertical or they can be unintentionally tilted, where the axis of the camera is not diverse more than 3° from the vertical. Oblique images are purposely tilted in interval between 3° – 90° from vertical axes. High oblique images show the horizon, low oblique are aimed more to the ground. However oblique images are not directly used for analytic purposes, because the drastic changes in the scale that occur from foreground to background prevent convenient measurement of the distances, areas and elevations. On the other hand, their meaning raise in the field of texturing 3D buildings models with facades. Principles and further usage of this new aerial technique is described in (Stilla at al. 2009). For a comparison the scale of the vertical photo is approximately constant throughout. This makes it more useful for measurements which are more accurate.

The most valuable spectrum for aerial mapping, besides of course the one which the human eye can distinguish, is near infrared (NIR). It is mostly free of atmospheric limitations and easily distinguishes various types of vegetation covers and land-water distinctions. For simplifying the visible spectrum, panchromatic view is used as a single channel which does not distinguish between the three primary colors. Panchromatic means 'across the colors' and provides a black-and-white image representation recording brightness without separating the different colors. According to the thinking that the added detail is more valuable than a color representation, the panchromatic band was designed for detailed capturing of the scene using a capacity that might have not been adequate enough for primary colors. (Cambell 2011)

As a conclusion, these days aerial imagery offers a simple, reliable, flexible and inexpensive method for large scale mapping. Geometric errors can be well handled which allows photogrammetrists to produce accurate information from acquired images.

Satellite Imagery

In this section the basic framework for understanding the key aspects of Earth observation via satellites will be summarised to readers, with references to sources where detailed information about different satellite systems are provided.

> *The observation satellite era can be credited to Landsat satellite which has formed the model for those systems as the first successful approach in this field.*

Today, national governments and newly private corporations run multi-billion investments into the satellite systems specifically designed for Earth's surface observation to collect fast, valuable information about various topics of interest. After comparison between aerial and satellite acquisition, satellite sensors offer several advantages over aerial platforms, such as capturing a larger area in a single image, as well as systematic and repetitive coverage. In last decades the number of satellite systems increased so rapidly that it would be both unpractical and for the not advanced reader, chaotic to list them all. But for better understanding the systems can be divided into the three main categories according their usage.

The first group represents 'Landsat-like system', developed for collecting broad geographic coverage using not fine but an 'acceptable' resolution. Available data from those systems are used for a large amount of applications, mainly represented by monitoring and survey of land and water resources. In the second group of observation satellites become systems which are designed to acquire very-broad scale images with coarse resolution. The most usual sensor provides large continental or global coverage which are used to monitor environmental dynamics or common applications. The final group of satellite systems is represented by fine resolution sensors which collect information about small regions. These kinds of data sets are valuable for urban planners, large scale mapping applications and many others which require higher spatial resolution from imagery. Keeping in mind that this categorisation is imperfect because of the high number of satellite types up there. But it provides a framework to help understand the capabilities of satellite systems. A comprehensive catalog of all land observing satellites is provided in (Stoney 2008), where author describes usage, resolution, country of origin and application of most systems.

For LANDSAT info: *http://landsat.gsfc.nasa.gov/education/resources.html*

Multispectral Scanner

Another method of acquiring data about Earth's surface is using a scanning system, which includes a sensor that sweeps over the terrain to store a two dimensional image of the surface. It can be similarly used with the same operating principles in both aircraft and satellite. The most common name for a scanning system which collects data over various wavelengths is multispectral scanner (MSS). There are two main methods to acquire multispectral images. (Figure 6) One scans in a series of lines oriented perpendicularly to the direction of the flight, called an **Across-track scanner**. Lines are scanned from one side (ie. left) to other (right) using a rotating mirror for spreading arrays in predefined sequences. As the sensor platform moves forward, scans record two-dimensional images of the area of interest. In this process, reflected or emitted energy is recorded by several separated detectors which may store the UV, visible, near-infrared or thermal wavelength intervals.

The second MSS type, called an **Along-track scanner**, similarly uses the forward motion of the platform where the sensor is mounted to collect scan lines for two-dimensional images. However, they scan the whole predefined range in one step instead of using a rotated mirror. For this purpose it uses a linear array of detectors which are moved in the direction of flight – it is also called a pushbroom scanner due to reason that the sensor is 'pushed' in the flight track.

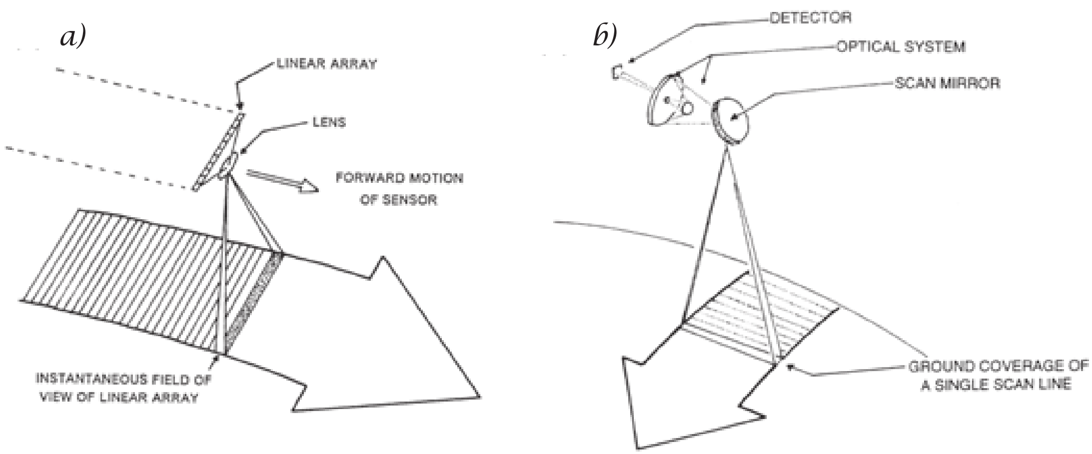


Figure 6. (Cambell 2011)
Multispectral scanner types
a) pushbroom scanner which acquires imagery line by line.
b) Along-track scanner using oscillating mirror to collect image of the scene from side to side.

There are some advantages of Along-track scanners over Across-track ones. To mention the most crucial, pushbroom motion allows detectors to measure the energy for each ground cell step for a longer time than the mirror spreading can offer. So more energy is collected that improves the final radiometric resolution. On the other hand, both scanners provide several advantages over camera systems because of the possibility to record more than visible and infrared regions. MSS is also capable of offering higher spectral resolution with comparing the camera systems.

Thermal Imagery

Thermal imagery is an important approach in the field of gaining information not easily derived from other types of imagery. Knowledge of different thermal behaviour helps to distinguish soil, construction materials and all other forms of land cover. For example, differentiations in water gained by thermal technology can uncover the presence of moisture in the soil, which is often a clue to define diverse classes of soil and rocks. Similar to previous methods, using far infrared region for thermal inspection has its own problems. Like all images, thermal imaging has geometric problems. Moreover, the analyst needs detailed quantitative interpretations of temperatures or detailed knowledge of emissivity must be available in order to reach satisfactory results. Because the thermal data sets differ much more than in the visible spectrum, it is often necessary to fuse thermal imagery with aerial to locate familiar landmarks. Far infrared wavelengths are free from the scattering but are restricted by absorption of atmospheric gases which limits its use in specific atmospheric windows.
In this category, lots of derived technologies and methods are used for highly specialized RS data acquisition such as, thermal scanners, microwave radiometers and so on.

To conclude the whole process of acquiring data using passive sensors, data quality is decreasing by limitations of both camera's parts and optical laws. Each image includes positional errors caused by the angle of the sensor optics, the motion of device, terrain relief and Earth curvature. Most of the causing optical problems are solved or suppressed to the reasonable level by using very high quality camera parts, additional capability as image motion compensation, tilt displacement and others which are precisely described at (Cambell 2011) or (Paine and Kiser 2012).

2.1.2 Active Sensors

The four previous systems outline a representative overview of passive systems. However, there are many other types of sensors which operate with different EM intervals. In this section, we briefly touch on a few of these systems to delineate the main principle behind active sensors as an alternative to the passive imagery concept.

RADAR

The name is an acronym for "Radio Detection and Ranging" (first time coined by US Navy at 1940) and represents active microwave sensors. The principle is based on instrument that transmits a microwave signal, receive its reflection from the terrain above and forms this information into image. A transmitter emits repetitive pulses of microwave energy at a known frequency. Based on this, a receiver recognizes the reflected signal and filters information from it. Regularly, the same antenna both transmits and receives the echo from the terrain. The distance is measured by the time delay between the time a signal is sent toward the surface and the time its echo is accepted. The second useful capability of the RADAR system is its ability to detect frequency and polarization shift of reflected energy. Because known wavelengths are transmitted, changes can be detected which figure as valuable information for post-processing applications.

The resolution of this technology depends on the signal length, where it can be seen that a shorter wavelength of microwave spectrum provides higher resolution. The biggest advantages of RADAR remote sensing technology are the capability of acquiring information in darkness, through cloud cover and its proficiency to observe large regions such as oceans or glaciers.

This technology would not be expected to replace imagery, aerial or satellite, but by using a combination, a more complete understanding of the land cover character is reached. Various devices sharing RADAR principles are available to collect information about Earth's surface, but most are called a "*Synthetic Aperture Radar.*" (SAR – or other meaning "*Side-looking Radar*")

LiDAR

Another member of the active sensor family is "*Light Detection and Ranging*" (LiDAR) technology. The same as RADAR, LiDAR is designed to transmit energy and receive the backscattered energy to produce an image of the acquired area. LiDAR system can measure characteristics, such as the timing of pulses and the wavelengths or divisive angles of output and input energy which allows not only brightness of the reflected energy but also changes in frequency, angular position and others differences from emitted energy to be determined. This information can be analysed and used for better description of terrain structure and vegetation features which are not expressed by optical sensors.

The whole technology works with coherent light - known as a laser; light which consists of a very narrow band of wavelengths and is monochromatic. This is used because, ordinary light, even if it is dominated by one color, still consists of many wavelengths. Laser light can be transmitted over long distances as a narrow beam which will differ very little in comparison with respect to other types of light. (Figure 7)

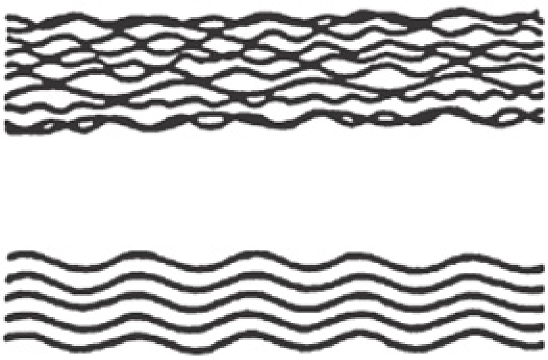


Figure 7. (Cambell 2011)
Normal (top) and **Coherent** (bottom) **light**

LiDAR as a remote sensing technology has been introduced relatively recently and it goes hand in hand with the development of technologies such as the high precision Inertial Measurement Unit (IMU) and Global Positioning System (GPS). These approaches allow precise control and recording of aircraft orientation (roll, pitch, yaw) and accurate records of geographic location of aircraft during acquiring data (GPS).

Nowadays, an airborne LiDAR scanner can transmit up to 300,000 pluses which are spread by a scanning mirror across the image swath beneath the airplane. Choice of spectrum interval to be used, depends on the purposes. For example, green in visible spectrum might be used for penetration of water bodies and near infrared for good sensitivity to vegetation.

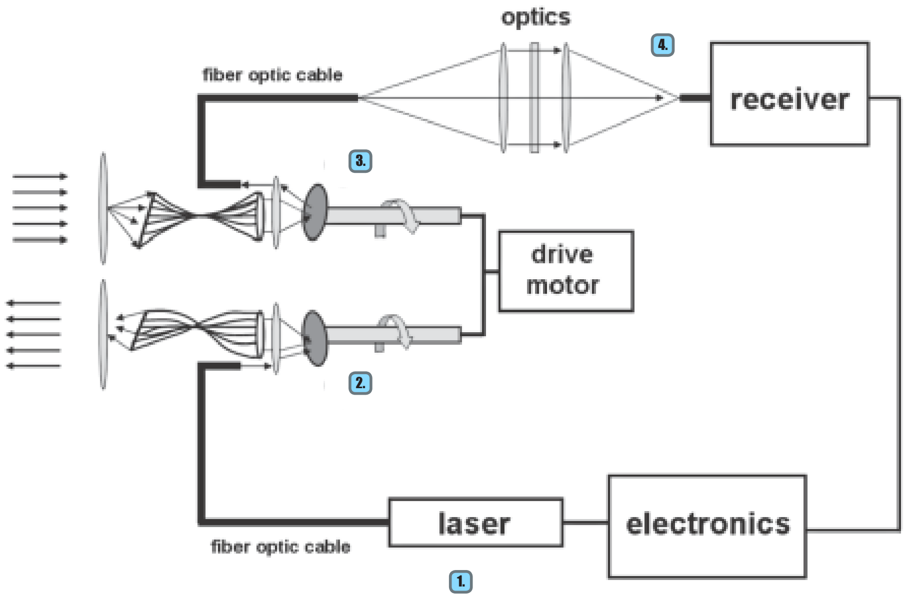


Figure 8. (Cambell 2012)
Rottating mirror laser principle

There are several alternative constructions for LiDAR technologies and one of the type consists of:

1. Laser – the electronic components generate a coherent beam which is transferred by a fiber optic cable to
2. A rotating mirror – spreads beams along the scanning line.

Generated laser light is sent through a bundle of optic cables that transmits it to the linear beam. The movement of mirror sends the beam side to side along the cross-track axis of the scanning area, recording thousands of returns each second.

3. Pulse return – goes through another optics and scanning lens
4. Receiving system – sends pulse to the electronics components

The connection with IMU and GPS allows every pulse to match to the certain point on the Earth's surface with high accuracy. The timing capability of the LiDAR system enables accurate assessment of the distance which permits to produce image with detailed and precise representation of elevation about scanning scene.

The main character of this technology stands on recording different kinds of returns from backscattered energy. (Figure 9) First (primary) return identifies upper surface of vegetation canopy. Other echoes which have passed through, represent secondary (partial) returns and carry information about lower vegetation layers and the ground surface itself.

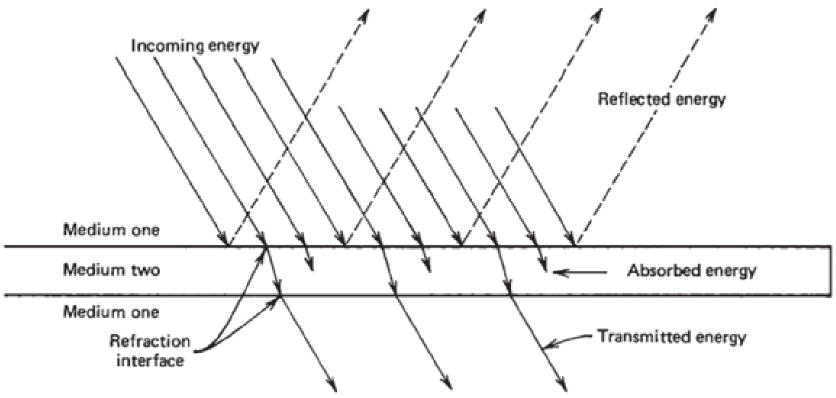


Figure 9. (Paine and Kiser, 2012)
The Interaction of EM energy
When it strikes a second medium, it may be reflected, absorbed, or reflacted and transmitted through it.

To mention different designs of LiDAR, one of the most used is pulsed laser (also called waveform LiDAR) that generates timed sequences of light. This measures the time between the emitted and received pulse. Subsequently, because light velocity is known and constant (prediction), the returns are translated directly to a distance between a platform and the target.
The resolution of LiDAR systems vary on the size of the footprint that matches the target. Flying height, beam divergence and other system design parameters affect footprint size – final resolution.

The result of LiDAR surface acquisition is in almost every case a Digital Surface Model (DSM) (Chapter 3.3.1) which represents the first pulse reflection and a Digital Terrain Model (DTM) which forms the surface after removal of vegetation and manmade structures. For further applications, fusion with other remotely sensed data sets, e.g. aerial imagery, might be used for identification and extraction of land cover types. This kind of approach is useful for applications which require the separation of ground from non-ground features. And this makes LiDAR one of the few sensors that can reliably segregate into numerous covers.

To summarize LiDAR technology, because of its ability for producing highly accurate and detailed representations of terrains with the possibility of land cover separation, it can be described as a unique approach among other remote sensing technologies.

For those who find Remote Sensing fascinating and would like to gain better fundamental overview, the author refers you to an easy-to-read paper "Fundamentals of Remote Sensing" which is free accessible on the web sites of Natural Resources of Canada.

2.2 Image Analysis

> *Image analysis of remotely sensed data is the science behind extracting information from the pixels within an image.*

In this section the pixel based and object based classification techniques will be described as the main methods in the remote sensing image process. A brief overview into the problematic will be outlined along with a simple explanation of the algorithms behind them. At the end of this section, a comparison and advantages of these methods will be mentioned.

Classification of remotely sensed data sets is one of the most important procedures of extracting information about land covers related to land use. The definition may read:

> *“Classification of remotely sensed data assigns corresponding levels with respect to groups with homogeneous characteristics, with the aim of discriminating multiple objects from each other within the image.“*

(Navlur 2006) To simplify it, classification is the assignment of objects, features or areas to the classes according to their appearance in the image.

The conventional procedure of image classification consists of two stages. In the first step the recognition of real-world object categories is identified. In the context of remote sensing and focusing on the land surfaces, these categories could involve, land cover types or man-made structures with respect to the geographical scale of the study. The second stage in the classification process involves labeling of the image entities (Mather 2004). In terms of digital image processing, these labels are numerical and a pixel that is recognized as belonging to one class is given the label “1”, belonging to another class may be labeled “2”, for example. For the user, it means to determine a priori the number of categories of which the land cover should be represented, as well as giving identification labels to the pixels according to which the pixel will be subscribed to the certain class. Sometimes these steps are known as classification and identification of the image.
Another approach divides the classification process into three levels of confidence and precision.

Detection → determination of the presence or absence of features

Extraction → insight into the object’s features to assign it in to the class or category

Identification → the identity of an object that is quantified with satisfactory confidence to belong to the specific class

In contrast, a process which does not require the definition of land categories is based on clustering. The aim is to determine the number of land categories present in the image area and to allocate pixels to these categories. As a result several pixel clusters may correspondent to the same land cover category.

The process of image classification includes a broad range of decision-theoretic approaches (algorithms) for identification what is in the image and where should it belong. All algorithms work on the assumption that image objects are represented by one or more features (in our case spectral regions) and according to them it might be assigned to a specific category with other similar pixels.

2.2.1 Pixel based Methodology

Traditional image analysis techniques are grouped under the pixel based approach (per-pixel classification, pixel based analysis, pixel-by-pixel classification). Classification is realized on the basis of spectral features – on the individual vectors of pixel values. Pixel based works just with simple pixels, without any regard to surrounding areas. This approach is based upon conventional statistical techniques, such as **supervised** and **unsupervised** classification.

Unsupervised Classification

Unsupervised classification assigns image pixels on the basis of spectral properties which are allocated to the class because of their similar values. This methodology is commonly used in the case where the image should be classified into an unknown number of classes.

> *The user only needs to define basic information such as which spectral bands to use for classification and how many categories to use in classification process.*

Mather (2004) describes this process as fishing in a pond of data hoping to come up with a suitable catch. In fact, it is an automatic procedure which is highly dependent on its mathematic background. Identification of spectral classes is achieved using whatever information from the image is available, giving unsupervised classification its 'exploratory' characteristic.

The simplest unsupervised method is considered to be the **k-mean algorithm** (classifying into 'k-number' of clusters). Pixels are grouped into clusters that are nearest each other which are positioned randomly through the spectral space. After, the mean location of the cluster center is re-calculated for each cluster. The process of assigning pixels to the nearest clusters is iterated and re-calculation of cluster centers is subsequently repeated until movement of cluster centers is under the threshold (description in Figure 10). If the threshold limit is fulfilled, then the class is assigned to the cluster.

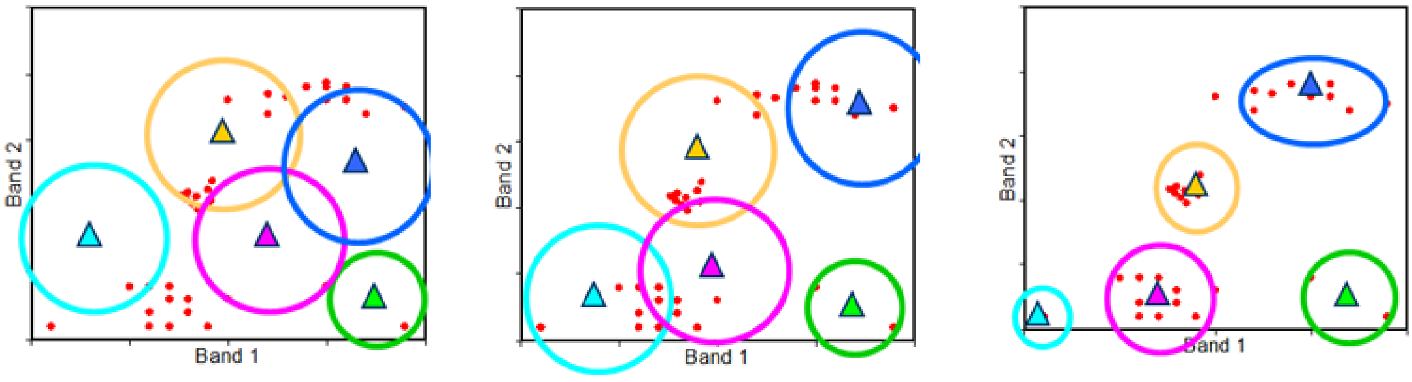


Figure 10. (Hutson 2006)
k-mean Algorithm
First iteration – Centre of clusters are set randomly and pixels will be assigned to the nearest centre.
Second iteration – The centres move to the mean-centre of all pixels in cluster
Third/n iteration – Centres have stabilised and fulfilled the threshold conditions

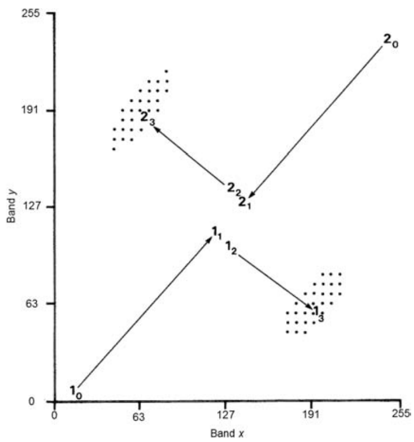


Figure 11. (Mather 2004)
k-mean Alghorithm background

The workings of this technique can be explained in the example with two well-separated groups of pixels in 2D space (each axis represents spectral band). The prediction, that there are two groups with unknown center position, will be used (Figure 11). Points 1_0 and 2_0 represent the first guess of the centers of the groups in the feature space. Next, the "shortest distance to center" decision rule is used to categorize each point (to class 1 or 2). Classifying depends on the relative Euclidean distance of the point from the initial cluster center (1_0 and 2_0). This distance is computed for each point and the point is assigned to the class 1 if d_{12} is less than d_{22}. In the case that the two squared distances are equal, then the point is arbitrarily labeled to one of these classes. After all the points are classified, the coordinates of the class centroid is calculated as a mean for each axis of all points in one class (1_1 and 2_1). With the new position of the centroids the whole process runs until there is no difference between last and previous position is detected. The final centroid position (1_3 and 2_3) is then generated for these two groups of points. Problems might appear when the points are not well separated in the feature space. Centroid boundaries become not so clear-cut which causes class membership misclassifications.

An extension to the k-mean method is the ISODATA algorithm (Iterative Self-Organising Data Analysis Techniques, with a terminal 'A' added for aesthetic reasons) which works with the assumption of an unknown number of classes. It also calculates the standard deviation for clusters and additionally provides the opportunity to merge clusters if the centers are close, split clusters with a large standard deviation into smaller ones or delete clusters which are too small. Then it reclassifies each pixel and repeats the following steps until it reaches the iteration or convergence limit. (Figure 12) After that, the class is assigned to a spectral cluster.

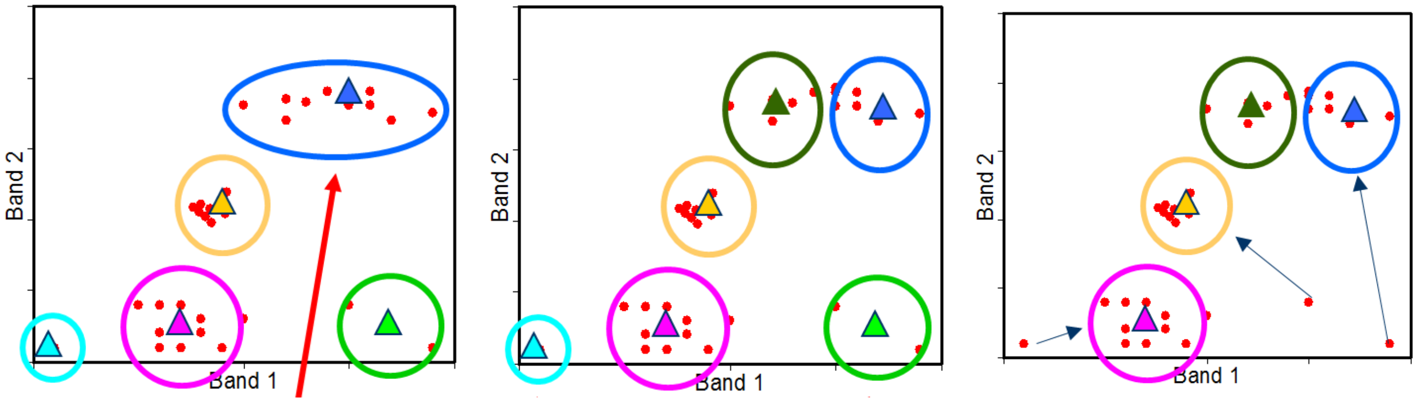


Figure 12. (Hutson 2006)
ISODATA
1. Data are clustered but blue cluster is very streched in band 1 - it will be split
2. Cyan and Green cluster include only a few pixels - they will be removed.
3. Unclassified pixels could be assigned to the nearest neighbour or stay as a unclassified.

The user specifies the threshold of the centroid scale and if a particular cluster exceeds this, it is broken into two. Once the pixels have been allocated to a certain class, the standard deviation for each axis is found and then the Euclidian distance between the cluster's centers is computed. In this step, clusters which cross the user's standard deviation limit are split in half perpendicularly to the axis where threshold was overlapped. On the other hand, close clusters which cross the lower threshold are merged. The split and merged functions are again applied until no clusters are split or merged, and none of the pixels change clusters. At the end, clusters with a low number of pixels are eliminated and these pixels are either ignored as unclassified or put back to the next iterations. Little guidance is needed to avoid endless loop when clusters are split at iteration i , merged in $i+1$, and split again at $i+2$.

Other derivations of these introduced methods can be used to fulfill the best results in pixel based image classification process, like a **modified k-mean algorithm** (described at Mather 2004).

In conclusion, unsupervised classification takes place in cases where the user does not have detailed imagery and cannot accurately specify training areas of known cover type. Generated classes may or may not correspond well with reality and the user will need to interfere more into process. For example, if the user has available small scale imagery of the city and wants to find how big area of the city district is covered by trees. Using unsupervised classification, pixels are assigned according their spectral similarity into classes which can be additionally used as an initial step prior to supervised classification – called hybrid classification. Hybrid classification is standardly used to determine the spectral classes of the imagery before conducting more detailed analyses. This allows the remote sensing program to classify the imagery based on the user-specified classes (supervised part), but will also classify other lesser known cover classes into separate groups (unsupervised part).

Supervised Classification

Supervised classification methods assign pixels to a class based on the multispectral configuration. The classes are determined on the spectral composition of training areas defined by the user, so methods are based on external knowledge of the area shown in the image. These training areas provide information about the classes which should be identified in the process of image analysis.

> *According to these training areas, users effect results by manually defining pixels into the classes.*

The supervised approach requires some input from the user before appropriate algorithms can be applied. The user usually derives this kind of information from field works or available maps.

General supervised methods are established on using statistical (parametric) or neural (non-parametric) algorithms (Mather 2004). Statistical methodology uses parameters derived from image sample data which are acquired from the users' predefined training classes. It works with features like the minimum and maximum values, the mean values or comparison of variance-covariance matrices for each of the class. Oppositely, neural methods do not rely on characteristics developed from sample data but are applied on the data directly, which might cause mis-identification in final results due to strong influence by size of the training data sets. This should be considered by users when a non-parametric approach is prioritized before parametrical, which is not so influenced by individual training pixel.

The final accuracy of supervised classification analysis will depend on:

- Quality, number and statistical nature of the users training classes
- Difference between assumption and real results derived from mathematical background

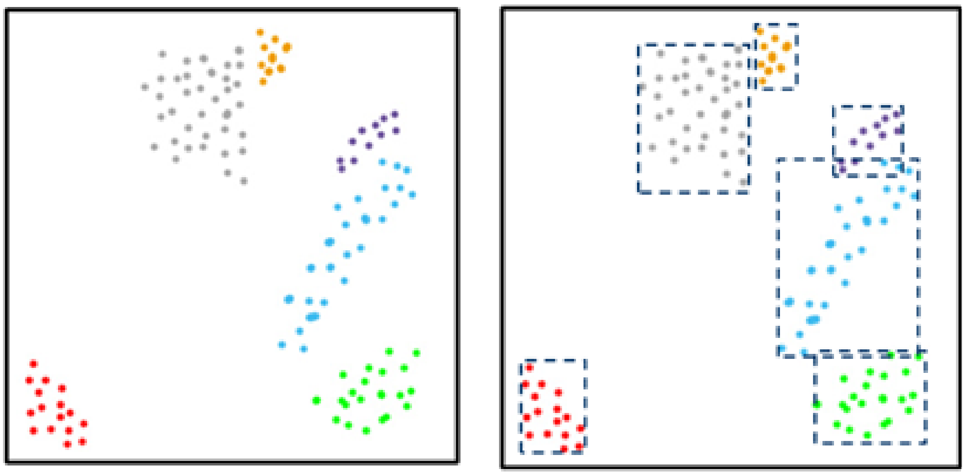
These assumptions vary on the technique used, but generally speaking, the most complex systems have more strict assumptions.

The most common algorithms in the supervised classification field will be described in following paragraphs. According to Milton and Arnold (1995) Maximum likelihood (factor of maximum probability), Parallelepiped or Nearest Neighbour principles best represent the algorithms that require the number of classes to be specified in advance.

Parallelepiped classification (box classifier) makes boundaries, which are straight and parallel, from the image data space to create a simple rule set to classify pixels. The least information from the user is needed compared to other methods described in this chapter, the user just defines n classes and estimates the minimum and maximum pixel value for n bands which are available from the image. These conditions are used as a classification threshold and estimate the position of the boundaries in n dimensional feature space (spectral bands) – pixels inside the 'box' represent one particular land cover type. Two unacceptable situations might occur when the pixel does not lie within

Figure 13. (Hutson 2006)

Parallelepiped: In figure can be seen that all class types defines parallelepiped boxes. Unfortunately, because some of parallelepipeds overlap it is possible that unknown candidate pixel might satisfy the criteria of more than one class. This can be solved by additionally customising boxes or defining new conditions e.g. it is assigned to the first class for witch it meets all criteria.



any of the box/class such as, a pixel will be assigned as unknown or when the pixel is located inside overlapping parallelepipeds. The condition for choosing the class it should become must be stipulated – the easiest way around the problem is to apply first the class inside whose boundaries the pixel falls. A more sophisticated decision rule can be based on calculations of Euclidian distance between the concerned pixel and the center of each parallelepiped and then use 'minimum distance' to find the best classification. Therefore, according to authors Lillesand *et al.* (2004) and Mather (2004), the Parallelepiped method is considered as a 'poor quality' but rapid methodology of allocating image pixels into the classes.

Maximum likelihood classification is a statistical method to help solve problems with overlapping signatures, e.g. when pixels are assigned to the class of highest probability. For describing the geometrical shape of cloud of pixels in n dimensional space an ellipsoid is used (for 3D a hyper-ellipsoid). The orientation varies on the degree of covariance between the pixels' feature. For each training class the spectral variance and covariance is calculated. The more circular shape of the ellipsis reflects a lower covariance among the bands at x- and y-axis. The shape, size and location of the ellipse are the results of the mean, variance and covariance between the two features of the pixels. Centric ellipsis (equiprobability contours) represent areas of probability of membership to the certain class. It is based on statistical knowledge that small centered ellipsis might include just a few pixels with high probability of membership to the class. Now, the distance from the center is not the only condition for deciding if a point should be assigned to one class or to another. For example, on the Figure 14 point 1 is closer to the center of the 'purple class' than to 'blue class', but according to the equiprobability contours, point 1 seems to be more likely assigned to that class. Due to the probability value, the new pixel is assigned to the class with the highest probability or unclassified if all probabilities are low. These kinds of classification results might be expected to be more accurate than a previous method (parallelepiped classification). (Milton and Arnold 1995)

The Nearest Neighbor (NN) method is perhaps the simplest of all the algorithms for predicting the class of a test example. It identifies the class of the unknown pixels on the source of its nearest neighbor whose class is known a priori. Firstly, training areas need to be picked to define the class features. Next, the algorithm searches for the closest pixels which fulfill the criteria of the training class. Distance is calculated from all training examples and the pixel with the lowest distance is called the nearest neighbor. Some approaches based on NN use weighted distances or apply additional steps for

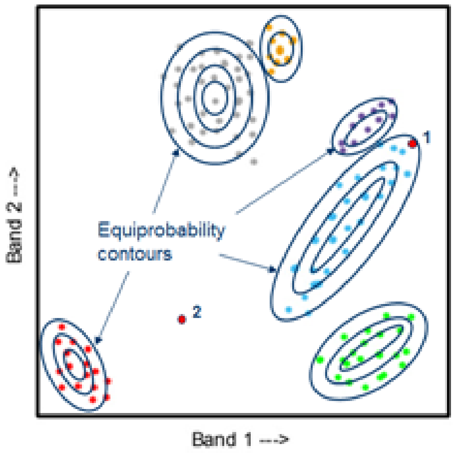


Figure 14. (Hutson 2006)

Maximum likelihood: Ellipses represent normal probability distributions of each training class. The lines, called equiprobability contours, show region of equal probability. As can be seen, point 1 would be assigned to "blue class" according its position into probability ellipses. But, point 2 would generally be unclassified due to its probability below threshold.
Drawing is in 2D spectral space, each band represents one colour.

reducing and eliminating pixels with not fitting features (Bahatia 2010). Depending on how well the pixel matches these criterion, it is assigned a 1 if it matches and 0 to remain unclassified.
In other words, it works on creating a database of criteria and training areas, for which the correct classification is known. Afterwards, when a new query is required, ie. classification of a new class, the method tries to find the nearest neighbor of the query in the training database – trying to find the classified object in database which is the most similar, which it then classifies the new object.

All in all, supervised classification can provide highly accurate results but the whole process depends heavily on the user's skill with defining the training classes. If the chosen training areas are not representative of the range of variability found, the final classification may be much less accurate. For example if classes are very similar to each other in spectral features (e.g. roads vs. roofs) misclassification will be more common in the process of automatic classification. But still as (Ghorbani *et al.* 2006, Karl and Maurer 2009) say, supervised and unsupervised classifications are both pixel based classifications and may be less accurate than object based. The assumption comes from high-resolution images where many pixels might be classified differently but actually belong to the same class. This phenomenon is in literature called the 'salt-and-pepper-effect' and leads into the unnecessarily detailed classification.

2.2.2 Object based Methodology

> *"Object based imagery Analysis (OBIA) is a sector of GIScience, segmenting RS imagery into meaningful image-objects based upon spatial, spectral and temporal scale followed by classification."*

(Hay and Castilla 2006) It was introduced as an alternate concept in 1970s due to the problem of handling of high resolution images by the pixel based method (De kok *et al.* 1999). OBIA provides an alternate view towards image classification founded upon an object based approach which recognizes objects or features, rather than the pixel based which examines single pixels. The method is a reaction to the rapid increase of high resolution imagery where the pixel based approach is not effective anymore due to the higher number of pixels to handle.

> *OBIA brings a fresh, new perspective to the image analysis with comparison to traditional pixel based approaches.*

At its fundamental level, OBIA requires image segmentation, attribution, classification and ability to query and link individual image objects in space. (Blaschke *et al.* 2006) As it was mentioned before, the OBIA methodology works upon image objects rather than pixels. An image object is the base unit for classification, consisting of a recognized group of pixels with similar properties: spectral, textural, size and shape as well as recognizing possible spatial relations between pixels.

Segmentation is the first step which is used to 'cut' the image into objects based upon spectral and spatial criteria or thresholds for further analysis. It subdivides the image into 'pixel clusters' (objects) according to their similar characteristics. Classification is the next step which analyses these image blocks and classifies them by using different techniques. This analytical method could be seen as the product of a number of spatial and commercial drivers. OBIA is seen to be most easily integratable with GIS using RS datasets.

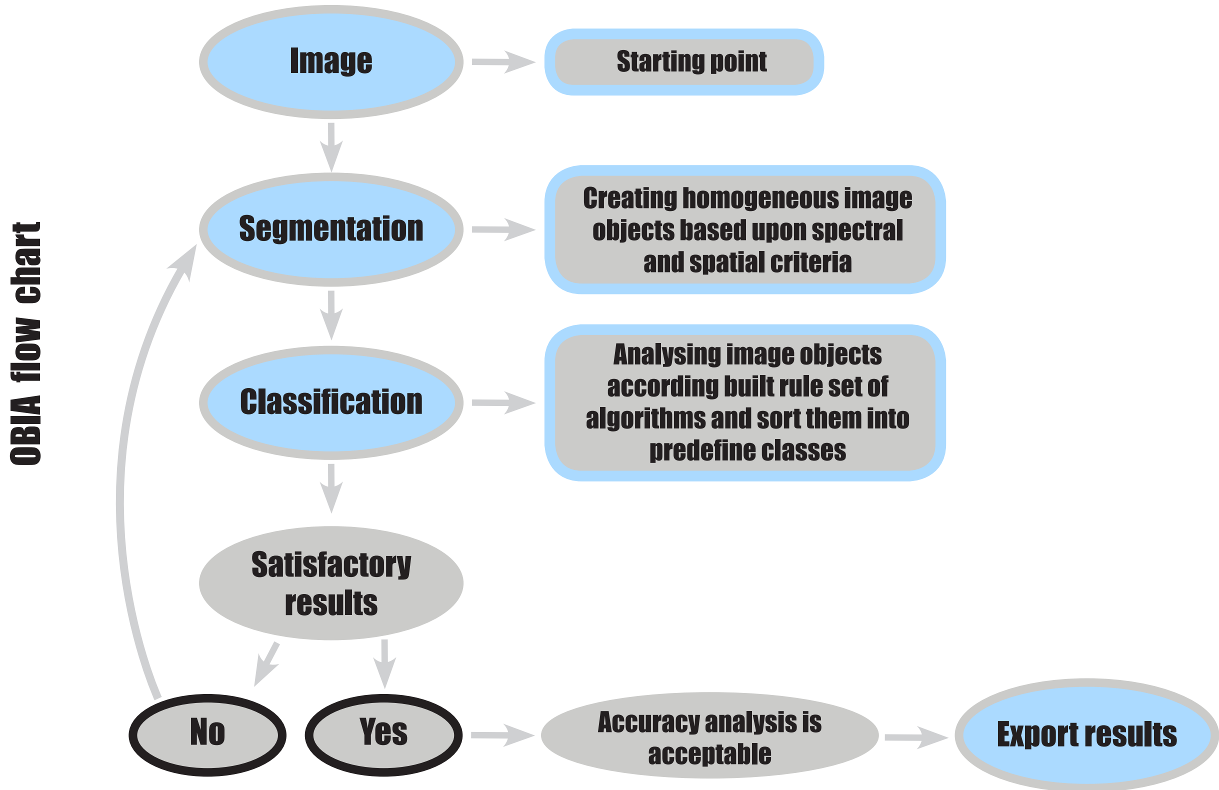


Figure 15.
OBIA flow chart

On one hand, the development of object based analysis can be seen to have been driven by new market growth in order to gain returns from the multi-billion dollar investments into high resolution remote sensing data acquisition, in which pixel based approaches are restricted. (Hay and Castilla 2006) Additionally ongoing increases in IT industry development and performance in terms of affordability and computing power, have enabled expanded availability of data and efforts in further developments. It was just an amount of time until pixel based approaches became too restrictive and object based methods were successively developed. (Hay and Castilla 2006) On the other hand, the importance of understanding images in a more meaningful way has been the most significant driver and it is now widely recognized in the spatial industry that pixel based approaches significantly neglect large areas of potential spatial analysis. Hay and Castilla (2006) also suggest that by increasing the awareness of OBIA methods, spatial information previously passed over can be utilised to provide enhanced integration with vector based GIS. Tobler's First Law of Geography states,

> " *Everything is related to everything else, but near things are more related than distant things.* "

In other words, objects are more likely to be related which are geographically closer together, so knowledge of spatial relationships and patterns are essential to an accurate understanding of the data acquired – enter OBIA. Before a more detailed process of OBIA will be described, let's look at

Why is an 'object' so important? → *Objects are the primitives that form a scene.*

The idea is based in nature, where the human brain can interpret rich information content from objects such as cars, houses, fields and other features present within the scene. (Navulur 2006)
Further emphasised by Laliberte *et al.* (2007), the analysis of objects is more propriate compared with pixels because landscapes consist of 'patches' – types of land cover. It is these 'patches' that can be made into image objects through segmentation and subsequently classified.

Figure 16. (Navulur 2006)
Object-oriented concept

1. Figure is zoomed all the way to the resolution of a few pixels. We can interpret that it is a bright feature/pixel, but we cannot assign a feature to the object.
2. Now, we have a starting point, and our brain can start retrieving various features that are elliptical in shape and have bright spectral reflectance and also use the contextual information that the bright object is surrounded by dark features.
3. This zoom level gives the first clue that the objects are probably boats in water.

4. 5. Our brain interprets to be a marina. An object-oriented approach is a first step in replicating this human interpretation process.

One of the most significant advances in OBIA is that spatial properties, such as length, width and direction can be used with spectral properties. For example, the difference between shadows and water bodies which have similar spectral parameters can be differentiated using spatial characteristics where shadows are smaller in area. Additionally, it combines current techniques for image analysis which were mentioned in the previous chapter, and some of GIS functionality.

OBIA Segmentation

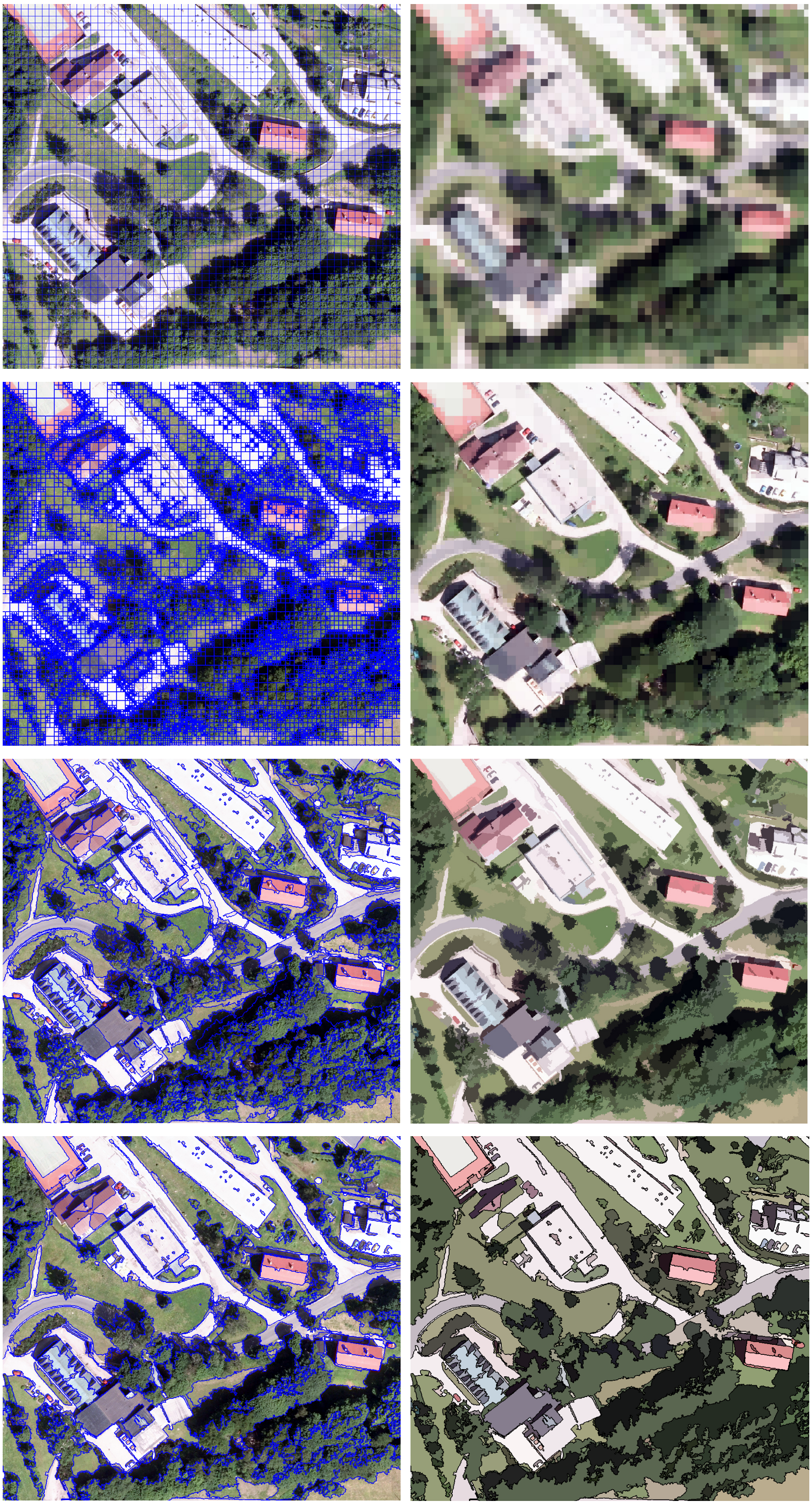
The key to land cover extraction from RS data using OBIA is considering segmentation and classification. (Jiang *et al*. 2008) Image segmentation is the first and also the most important task where the image is subdivided into smaller image objects following predefined criteria. The objects are created from pixels by using top-down (chessboard, quadtree) or bottom-up (multiresolution segmentation, spectral difference) techniques (explained by eCognition, Reference book 2012). Another author divides segmentation algorithms according to one of two processes: region merging or separation by finding edges. Navulur (2006) and Zuva *et al.* (2012) add thresholding segmentation technique as a third into the group.

Region-merging methods can be divided into two approaches: region growing and region split-and-merging. Region growing algorithms start with a single pixel and regions are grown by merging neighboring pixels that have similar properties, such as color, intensity and texture. The most generally used techniques are Thresholding method, Region Growing, Classifiers and Clustering. (Zuva *et al*. 2012)
In contrast, the region split-and-merging approach works with subdividing an image into a set of regions and then it merges or splits the regions based on the similarity rules for object creation. Thresholding based segmentation drives the comparison of pixel values with the predefined user's rules. These user's conditions may be applied globally, one rule set of conditions are applied to either the whole image area, or locally, where the image is subdivided and different threshold conditions are used for each.
Pixel relations to their local neighbour can be described by two basic properties - discontinuity and similarity. Segmentation methods based on these properties are considered as boundary/edge based techniques. (Zuva *et al.* 2012) Edge based segmentation finds the location of the pixels in the image that corresponds best to the object's boundaries seen in the image.

> *As a general rule, 'good' image objects should be as large as possible, but small enough to show contours of interest and to serve as building blocks for objects of interest not yet identified.*

The mentioned segmentation approaches are the 'building blocks' for lots of others segmentation algorithms which are derived from common mathematical background. For a deeper insight into other image segmentation techniqes which are also commonly used, beside RS, in medicine or in film industry, author refers you to many surveys available online.

Figure 17.
Different Segmentation Techniques



Chessboard:
Generates square objects of specified size, resulting in a grid pattern.

Left: object size 25
Right: segmentation result

Quadtree-based:
Generates square of different sizes based on homogeneity criteria:

- Scale - determines size of image objects
- Mode - Based on either Colour or Superobjects

Left: quadtree mode colour and scale 100
Right: segmentation result

Multiresolution:
Consecutively merges pixels or existing objects dependent upon the homogeneity criteria:

- Scale
- Shape
- Compactness

Left: Multiresolution scale 100, shape 0.1, compactness 0.5
Right: segmentation result

Spectral difference:
Merges neighbouring image objects based on mean spectral intensity difference.
Can only be used on already generated image objects.

Left: scale 25
Right: segmentation result

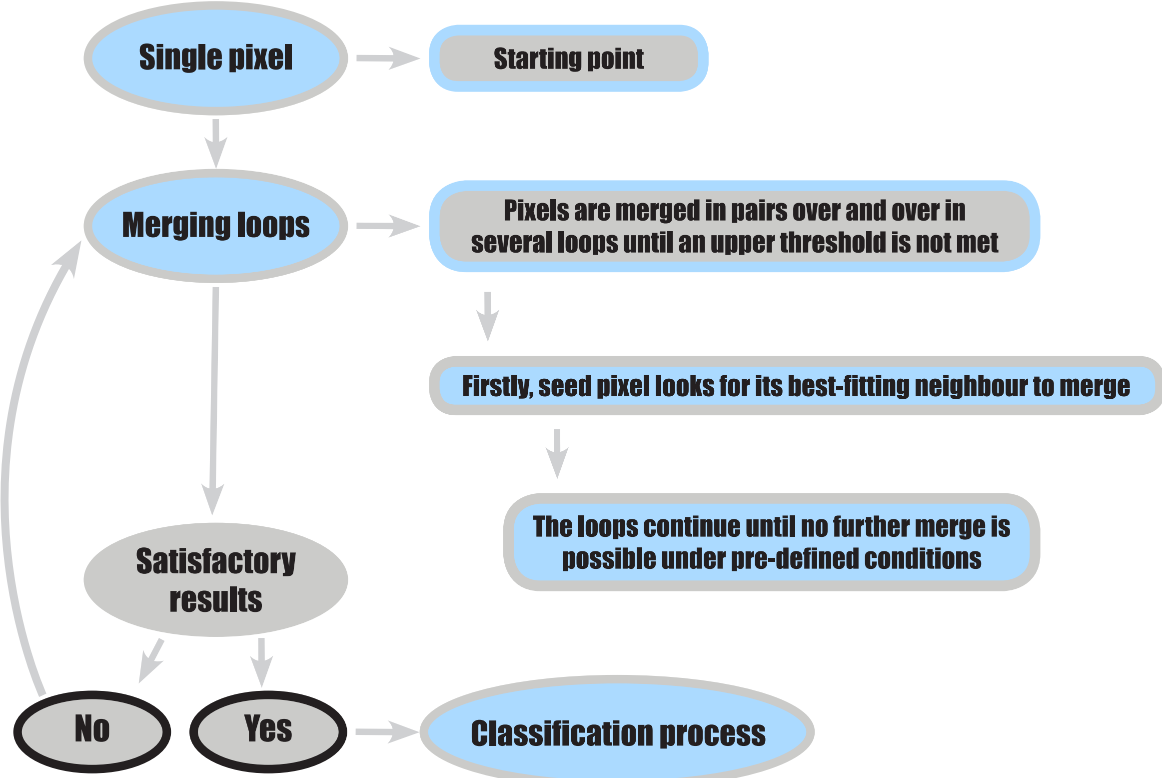


Figure 18.
Segmentation flow chart

OBIA Classification

Similar to the pixel based approach, object based classification can be supervised or unsupervised. Supervised classification involves selecting representative image objects as training samples for each different class. The objects can then be classified into a class using algorithms such as Nearest Neighbour. The unsupervised classification process requires conditions to be made, based upon image object values that differentiate one class from the other.
As mentioned before, OBIA classification works with algorithms that measure various characteristics of image objects (shape, texture, color, size, and so on). This feature is used for defining upper and lower limits of condition thresholds. Image objects within defined limits can be assigned to the common class. On the other hand, objects outside of these feature range remain unclassified (ie. they did not match any prior conditions).

The following is a brief list of commonly used features (eCognition User's guide, 2012):

- **Color:** the mean or standard deviation of each band, mean brightness, band ratios
- **Size:** length to width ratio, area, relative border length
- **Shape:** roundness, asymmetry, rectangular fit, compactness
- **Texture:** smoothness, local homogeneity
- **Class level:** relation to neighbors (class), relation to superobjects and subobjects.

As can be seen the available settings for OBIA for defining 'the best' parameters are very broad. Usually a combination of trial, error and experience is required for reaching satisfactory results from this procedure.

A significant feature of OBIA is allowing analysis to be undertaken on different levels. Spatial information allows more than one level of analysis which is ideally beneficial for landscape analysis that usually requires multiple, related levels of segmentation. For example, if vegetation was the top level, trees, shrubs and low vegetation would be sub levels (Figure 19). More meaningful results can be reached by OBIA in comparison with pixel based image classification in terms of defining edges or boundaries between different classes.

After finishing the OBIA classification procedure, the output is a classified image which becomes part of the further investigation (ie. exported to a shape file for use in GIS). Accuracy assessment can be evaluated through tools such as a Confusion Matrix which represents the amount of correct classification associated with each class.

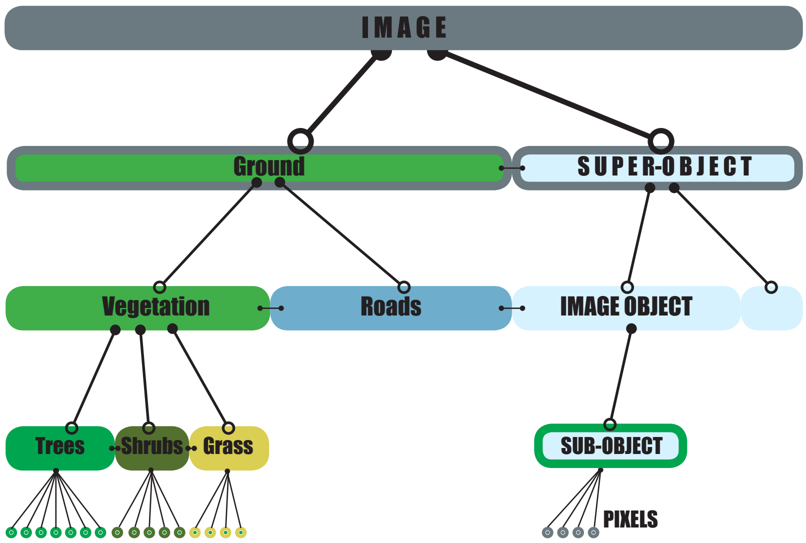


Figure 19. Class Hierarchy

All created objects in OBIA are part of the image object hierarchy which has many different levels. Each level works as a virtual copy of the image, represents information about particular parts of the image. Objects are linked to neighbouring objects on the same level.

The biggest difference between approaches in classifying is that OBIA segments the image into image objects which are simple and easy to handle in the further classification process. This methodology appeared when development in IT advanced, and faster processing of high resolution imagery was required. The advantages of this method are the significant improvement of data processing times and particularly the ability to quickly delineate edges within a large amount of data. OBIA produces classification results that are extremely close to that of what a human would interpret from the image, subsequently automating the procedure and reducing the manual efforts by technicians. In other words the OBIA provides classification that is closer to what the real world looks like.

2.3 Land Cover and Land Use

Before defining land cover and land use, this paper agrees with the opinion that for practicality, land cover and land use must be considered together, while also knowing the distinction between the two. (Cambell 2011)
Availability of digital image processing is the main objective of this thesis, and deeper knowledge about ‘what needs to be found in the image’ is the potential to automate land cover/use mapping for big areas.
European Communities (2001) official paper defines land covers as “a physical description of space, (bio)physical cover of the Earth’ surface.” For the author this definition is pretty rough and sees the term land cover as

> *Kinds of vegetation that cover the Earth’s surface with materials that form the surface in the places where vegetation is lacking.*

Now, the categories of various (bio)physical elements should be distinguished, eg. vegetation (trees, lawns, …), bare soil, hard surfaces (rocks, man structures, …), water bodies and so on. All this is needed for development of classification systems which varies on the purpose of research. A unique classification system can be developed to best fit needs of a particular project or for example, the most widely used U.S. Geological Survey’s Land Use and Land Cover classification system, developed during the 1970s, can be followed.(Cambell 2011) See U.S.Geological classification system in Appendix A.

Land use, by contrast, describes a usage of the land surface by humans. The majority of the literature perceives it just in a socio-economic context (eg. Industrial, residential or commercial purposes), but European Communities (2001) study adds also agricultural purposes.
Land use might be defined as

> *a series of operations on the land, carried out by humans, with the intention to obtain products and/or benefits through using land resources.*

In comparison to land cover, land use is difficult to observe. It is often hard to decide for example, what grasslands are used for, just by looking at the image. It requires accurate understanding of the relation between land cover and land use, to find best solution for planning purposes or collecting data.
Cambell (2011) considers RS as an accurate and cheap acquisition technique for land cover and land use mapping in part because objects can be seen in the context of the neighbouring features.

Land Cover → surface features

Land Use → economic function

2.4 Literature Summary

This section will outline the general opinion of how the researcher's results will profit the field of image processing, a suggestion of methods on how to handle a fully automatic approach, the type of data for the best land cover classification and a brief overview into several studies which were completed in recent years.

An obvious tendency of RS imagery extraction is to lead in various planning and modeling applications, followed by high demand for quick, cost efficient and precise information. On the other hand, there are still difficulties with manual interaction from skilled modelers, which slow down the whole process of valuable information extraction. The alternate semi-automatic approach, where efficient human interaction occurs, can meet the high precision needs but according to overall scientists' opinion **the automatic processes seem to be the only way to satisfy the growing trend for the future.**
In the past, most the image classification have been undertaken by using pixel-based method. Results of fusion with available high resolution imagery have found a "salt and pepper" effect that contributes to the inaccuracy. Technological advances in recent years ask for more development in methods for handling the growing availability of high resolution imagery. The answer for them is the object-based classification approach, which can deal with a wide range of different data storage within advanced database structures.
Authors prove that the most recent achievements in the automate image processing are centered to the integration of data combining two or more sources. Their research shows high potential of fusion from different spatial sources – mostly multi-sensor imagery or a combination of imagery and LiDAR data.
In the following paragraphs, a brief overview of recent research in this field will be provided.

Rottensteiner *et al.* focuses on two topics. The paper deals with a technique for automated generation of 3D building models from a combination of aerial imagary and LiDAR. Secondly, it describes an object-oriented paradigm for handling these kinds of topographic data sets. The authors' test the OBIA technique for storing large sets of RS data,concluding that the OBIA method has a great potential for combining different spatial sources for automated and optionally semi-automated processes. (Rottensteiner *et al*. 2002)

Jiang *et al.* promotes object-based methods to solve building footprint extraction problems that occur with pixel-based techniques. The paper describes advantages of dealing not only with the spectral information but also with the shape, contextual and semantic information which rapidly improves results. The authors use eCognition software package to efficiently handle high resolution imagery with high-reaching precision. The crucial step for the entire classification process is considered to be the segmentation component. In their research multi-resolution segmentation with the right parameter scale is believed as an optimal solution for reaching sufficient accuracy. (Jiang *et al*. 2008)

Yunhao *et al.* deals with land cover extraction in urban area. This paper presents a hierarchical object-oriented classification method based on QuickBird imagery combined with LiDAR. Authors focus on the distinguishing problem between water bodies and shadows, due to spectral similarities. Additionally, the paper provides discussion about improving total accuracy between pixel-based and object-based methodologies. (Yunhao *et al*. 2008)

Weih *et al.* provides a comparison of object-based classification with supervised and unsupervised pixel-based analysis using multi-temporal imagery, SPOT-5 imagery and high-spatial resolution orthophoto for analysis. On these three sets of data, authors show the importance of multi-temporal and multi-spatial imagery for total classification accuracy. Their objective determined that object-based analysis produces statistically more accurate land cover classification than the pixel-based approach when applied to the same imagery. The accuracy assessment used was an error matrix to determine differences. (Weih *et al*. 2009)

Chmiel and Fijałkowska examine different approaches applied to thematic accuracy assessment for object-based image classification. For their research, the authors used examples of images with different resolutions, including a very high resolution one. Several accuracy assessments methodologies were put into process which confirmed that every evaluation process is very sensitive and can be dependent on different factors – proper selection of accuracy indicators and a more objectively applied procedure can allow for achieving a useful accuracy figures. (Chmiel and Fijałkowska, 2012)

Walter uses a quite different approach from most of other researchers using object-based classification method for detecting a land cover. In the first step, supervised maximum likelihood classification is utilized over RS data to classify it into different land cover classes. The training areas are obtained from an already existing GIS database which avoids having to manually pick examples. Afterwards, results are compared with a GIS object database to detect any changes that occurred or which were classified incorrectly. (Walter 2004)

Yan *et al.* tested land cover mapping using pixel-based and object-based methods and compare them over the same data sets. Pixel-based classification and supervised maximum likelihood classification algorithm were utilised and compared with the object-based method run on a region-growing multi-resolution segmentation and a soft nearest neighbor algorithm. Classification accuracy was evaluated using ground referenced data and error matrices were formed and compared. This led to the conclusion that classification accuracy between these two methods was distinctively different with significantly better results for object-based approach. (Yan *et al*. 2006)

Zabuawala *et al.* also uses a fusion of aerial imagery with LiDAR as the best solution for an automated delineation of land cover features. Firstly, LiDAR data is used for classifying ground and non-ground characters following by solution of detecting trees and wires. Secondly, imagery with watershed segmentation technique is applied for precise delineation of building footprints. Lastly, whole process was used to another area to proof efficiency. (Zabuawala *et al*. 2009)

Hermosilla *et al.* describes a methodology for mapping urban land-use types working with multiple data sources. A combination of high spatial resolution imagery, LiDAR and cadastral plots is presented. Cadastral plots help reduce the volume of information that needs to be manually interpreted and different classification algorithms are applied according to the cadastral area (for example, residential, industrial, greenlands). For classification, a supervised approach with a hierarchical tree is used based on heritages, where a branch is added below with a new condition, but keeping those conditions above. The paper concludes with promising results provided by external and internal additional features into the classification process. (Hermosila *et al*. 2012)

For sum of the available literature author refers reader to Blashke (2010) who summarised in his research paper overview of available literature focused on OBIA paradigm available to date of writing.

2.5 Pre Analysis Conclusion

To conclude the analysis investigation, the final data types, software and methods chosen for further exploration in this thesis will be described along with the reasons why. All decisions were made according to recent research and availability for the author.

Method: The object-based approach, OBIA, has been chosen for image classification according to the conclusions of many publications that focus on automated image processing from RS data sets who find significantly improved results using this method over the pixel-based approach.

Data Types: To follow on from the methodology described by these publications, a combination of a high resolution RGB+NIR orthophoto along with height information from LiDAR data will be used, improving the entire process as well as the final accuracy. The land cover detection problems found when using only imagery, such as occlusion and tree foliage interference can be solved when combining with these two data types. To combine high resolution imagery with any height data for the land cover, correct geo-referencing for alignment of both data sets is required.
Another reason for this decision, is that these data types are the most easy-accessible geo data which fit the requirements for classification research.

Software: After choosing to explore the object based approach, eCognition Developer was the obvious choice to investigate building boundaries because it is based upon the OBIA principles and has free available trial version. Additionally it has the capacity to combine and analyses the two data sources that were decided to work with.

3 Investigation

This chapter will outline the procedures undertaken during the investigation for this master thesis. Using the orthophoto and LiDAR data available, classification will be executed using OBIA principles with the aim to delineate land cover types from their surroundings. The investigation was undertaken using the free available version of eCognition Developer Trial 8.

The chapter will begin with a summary of the main questions which will be answered, description and explanation of the data processing required before analysis, followed by how the workspace in eCognition is set up. The object based approach requires initial segmentation to create appropriately sized image objects and some classification algorithms customized by autor will be outlined. Following this, the process of how the objects will be classified according to rules or threshold conditions will be described. The image classification has great potential for flexibility in defining rules which will make up a large part of the investigation.

3.1 Problem Statement

After exploration into image classification techniques and data sources; decisions were made to explore how image classification techniques can be utilised with a combination of a high resolution orthophoto and LiDAR data to delineate different land cover types with as much automated development as possible. The idea is to improve a process to classify land cover, to subsequently generate suitable data for further GIS use.
The following question will be answered:

> *How well the OBIA approach can be used to detect and accurately delineate different land cover types using an orthophoto combined with height information from LiDAR?*

This will be assessed against the following criteria:

1. Highest degree of automation possible
2. Resulting classification accuracy
3. Transferability of established strategy

The Transferability process will be assessed to see if it can also be used in other geographical areas with the same success as the study area.

According to the author's beliefs, the outcomes of the project will be seen to play a significantly increasing role in the future fields of spatial visualisation in map based information and change the way of re-using available data sets for simple extraction of certain information.

3.2 Available Data

When considering the choice of the study area for this research, it was determined there should be some criteria to base it upon, which will be described as follows.

The whole idea of creating automatic recognition rule set stands upon its usage for commercial purposes. The author wants to prove that further post-processing of already available data will play a significant role in the field of land cover mapping, which is still mostly held by field workers. An area of dense vegetation with built-up areas and water bodies would help to increase the ability of simulation various land cover types for detection. It was found that a part of Krkonošský Národní Park in Czech Republic (KRNAP – The Krkonoše Mountains National Park) fulfilled these criterions by having diverse vegetation covers, a variety of buildings, paved and unpaved roads, etc. Figure 20 shows relative position of chosen study area.

To gain the best possible results, a high resolution orthophoto with RGB+NIR bands was chosen. This is a digital aerial photo image in its raw form that has been orthorectified (geometrically corrected) to a suitable DTM. This process is required to be able to measure true distances from the imagery with minimised distortion. For a detailed description upon the steps needed to create an orthophoto the author refers the reader to works such as Aerial Mapping Methods and Applications (Falkner and Morgan 2001) as well as the many publications available online.

When searching for LiDAR data, the DTM and DSM data were available from KRNAP already generated from the raw LiDAR form. This meant that the usual processes of creating the DTM and DSM were bypassed. However, this meant that there were no intensity values to use and that LiDAR data would be used for the height information only.

The LiDAR data were acquired by company Geodis, via laser scanner Riegl LMS-Q680i in dates between 24/07/2012 – 18/08/2012. Equipment was mounted on the airplane Zlín Z-Z37. The geo-referencing of the point clouds was calculated from GPS and IMU vectors, exporting to .las file (v1.2) through software RiProcess (v1.5.7). DSM and DTM, converted into 1 m grid, were generated using software Microstation V8, modul TerraScan 012.020 and TerraModeler 012.008.

The DTM and DSM of the area were available in ASCII format with three columns, YXZ with a regular grid of 1 m.

Data Information	
Ortophoto imagery	
Horizontal accuracy	0.300 m
Pixel Resolution	0.125 m
Date of Acquisition	18-29/06/2012
DTM/DSM	
Resolution	grid 1 x 1 m
Hight accuracy	0.087 m (accuracy of point clouds)
Horizontal accuracy	0.145 m (accuracy of point clouds)
Date of Acquisition	24/07/2012 – 18/08/2012

Table 2.

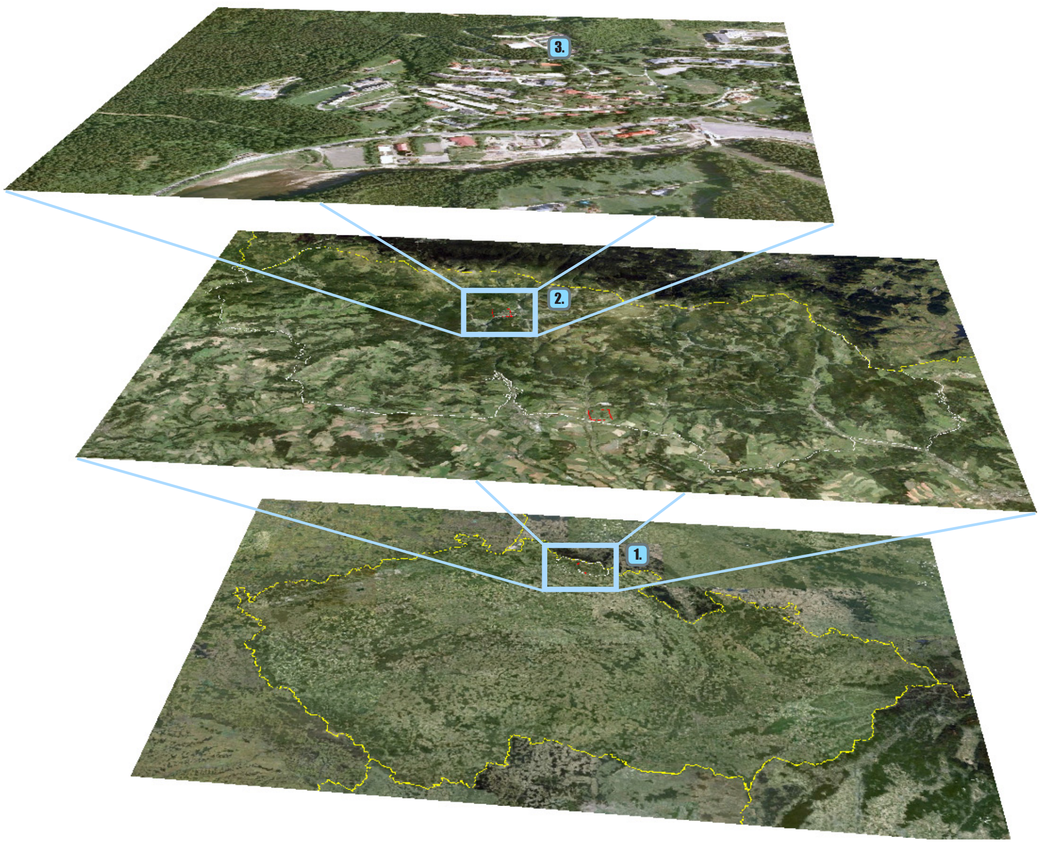


Figure 20.
Location

1. Krkonošský Národní Park, Czech Republic
2. Špindlerův Mlýn / Bedřichov (name of the town)
3. Study area

3.3 Initial Steps

3.3.1 Creation of Normalised Digital Surface Model

The DTM, DSM and orthophoto were already processed and geo-referenced which meant the only data processing required is creating an appropriate height model to combine with the orthophoto. The height model to be used for this process is the normalised DSM (nDSM). This is a subtraction between the two terrain models available (Figure 21) and is used during the image classification process to distinguish above ground features. This is ideal for the process of isolating buildings and trees from surroundings. The DTM and DSM were converted into rasters prior to the subtraction and the result is a .tif file containing only height values above the ground. This was all done within ArcGIS software.

DSM - DTM = nDSM

Please see Appendix A. for further details on how it was generated.

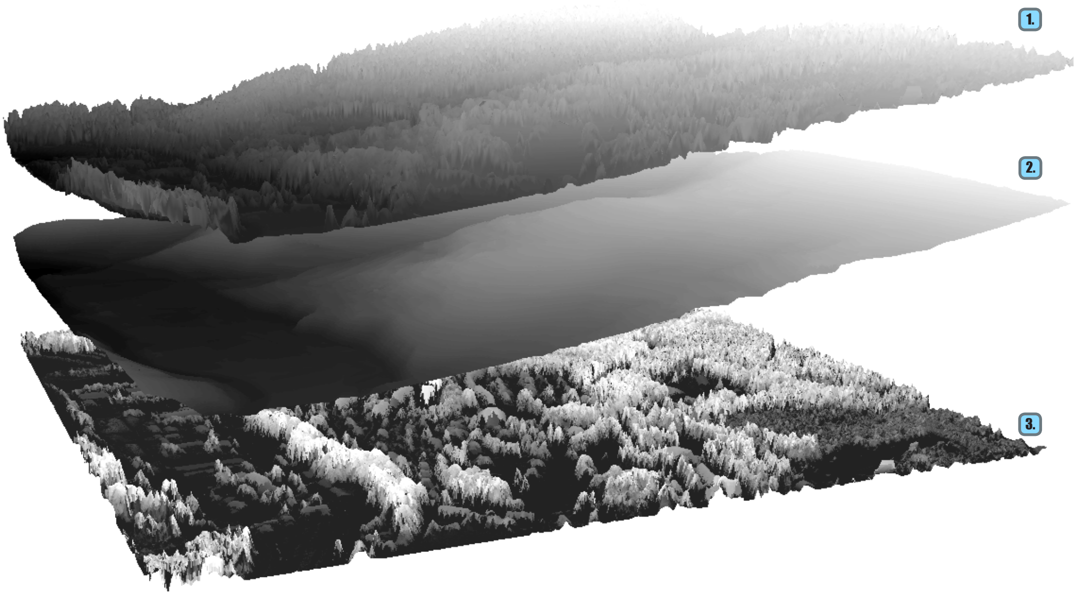


Figure 21.
nDSM Creation
Layers of 1. DSM, 2. DTM, 3. nDSM

3.3.2 Importing Layers

In this part, the explanation how available data sets were imported to make a 'starting point' for further classification is provided.

Firstly, it is important to ensure that the data sets to be imported as layers are geo-referenced correctly so that they overlay each other precisely. Secondly, due to the complexity of the algorithms used for segmentation, a smaller representative subset of the investigation area was selected to decrease the time spent (Table 3). The subset was chosen (Figure 22) because initial exploring found this area to be representative of problems that needed to be solved in this research. These include the variety of different land cover types in reasonable large area.

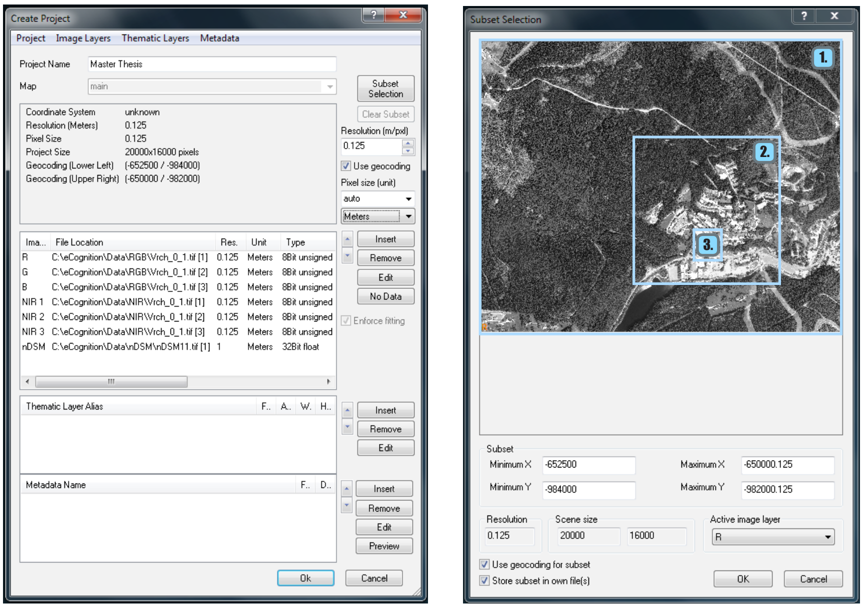


Figure 22.
eCognition interface for data set up

1. Available orthophoto image
2. Available nDSM area
3. Subset area chosen by author to find suitable rule set and minimise amount of data for caluculation

Calculation spend time depended on area size

	size [pxl]	*time [min]*	*number of objects*
Area 1.	20 000 x 16 000	40:13	2 840 806
Area 2.	10 000 x 10 000	7:20	132 592
Area 3.	1 511 x 1 600	0:21	5 110

Table 3.

3.3.3 Rule set Creation

Segmentation

After importing the layers the next step is to create the rule set by firstly identifying the correct parameters for creating appropriate image objects in a segmentation algorithm. This step is considered as a crucial 'building block' for the whole automatic rule set establishment due to all the thresholds and conditions that will be applied on image objects created by segmentation algorithm.

As it was outlined in the Pre Analysis section (Chapter 2), there are different segmentation techniques that enable similar pixels to be clustered together. Each has different parameters to consider and the possibility to weight layers depending on their priority in the segmentation. These have been previously defined, such as the bottom up and top down approaches of Multiresolution and Quad Tree, respectively.
In this case, Multiresolution was initially chosen as the optimal segmentation algorithm for OBIA procedures. Developed by Baatz and Schäper (2000), this algorithm finds small objects (the pixel clusters) and then merges them if they have similar characteristics as defined by user - scale, shape and compactness (Figure 23). Multiresolution Segmentation is defined by the eCognition Userguide (2012) as

> *'a bottom-up segmentation algorithm based on a pairwise region merging technique'.*

In other words, it is an optimization procedure which, for a given number of image objects, 'minimizes the average heterogeneity and maximizes their respective homogeneity.' The figure below displays how the algorithm utilises these parameters.

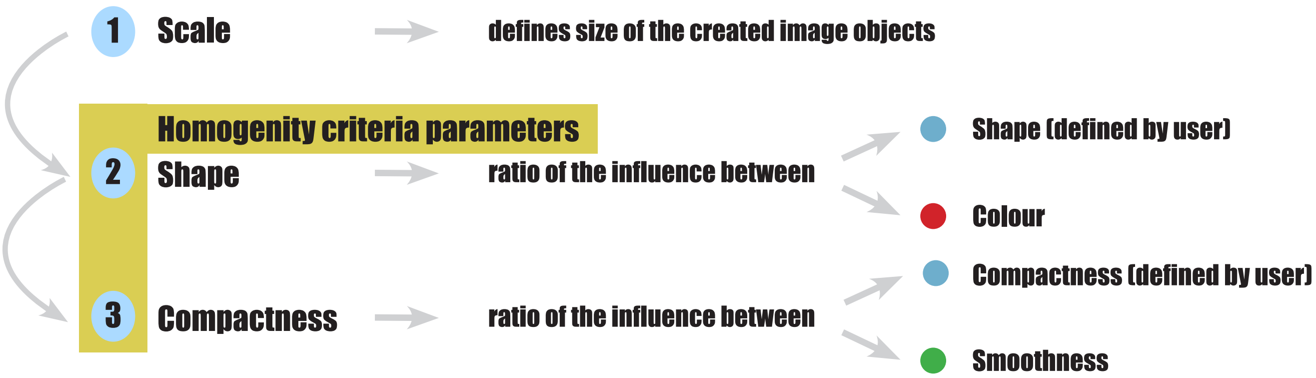


Figure 23.
How Multiresolution segmentation works

The Multiresolution segmentation algorithm optimises the ratio between spectral and spatial heterogeneities via formula (eCognition Reference Book 2012)

$$h = w * h_{colour} + (1 - w) * h_{shape}$$

Where h characterises a heterogeneity value calculated with a pair of image objects, h_{colour} represents spectral heterogeneity, and h_{shape} indicates spatial heterogeneity. w indicates weight of ration between colour and shape.

Spectral heterogeneity is defined as

$$h_{colour} = \sum w*\sigma$$

where w represents weight value of each image layer and σ indicates the standard deviation of pixels within a pair of image objects.

The **scale** parameter controls the size of the objects created, therefore it is dependent upon the scale of the image. The larger scale the image, the higher values needed, the smaller scale the image the smaller value required. By increasing the scale, the size of the objects becomes larger. The scale needs to be adjusted according to the size of the object that is needed to be classified. As seen in Figure 24, making the scale too small increases the number of objects which makes further classifications more difficult in terms of time taken and final object identification of various land cover types. However, too large a scale will find e.g. the buildings footprint but the object will also 'bleed' onto the surrounding surfaces of similar characteristics.

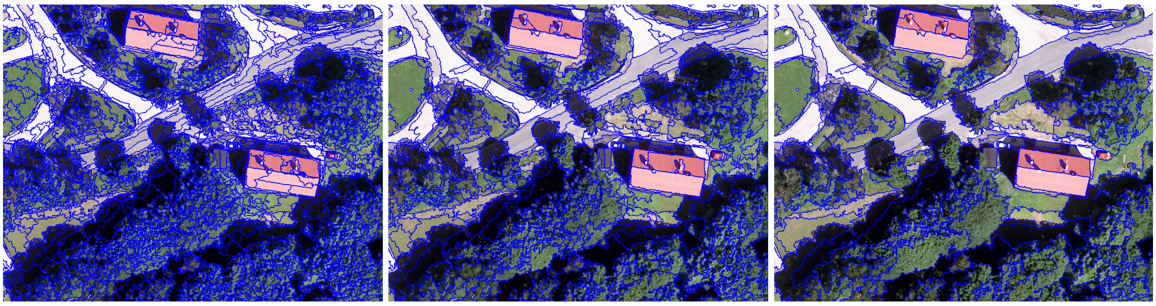


Figure 24.
Example of different scale value
Left: Scale = 25 ; Shape = 0.1 ; Compactness = 0.5
Middle: Scale = 50 ; Shape = 0.1 ; Compactness = 0.5
Right: Scale =100 ; Shape = 0.1 ; Compactness = 0.5

Homogeneity criteria are calculated from three basic elements - colour, compactness and smoothness. In most cases related to image post-processing, the colour criterion is the most important for creating meaningful image objects. However, for improving the quality of segmentation, a certain degree of shape homogeneity is required. The **shape** principles help us avoid highly fractured image objects. It is a ratio of the influence that shape has over colour in the segmentation, ie 'a weighting of 0.9 for shape infers a 0.1 weight for colour.'
By decreasing the shape value, the resulting objects are then based upon different

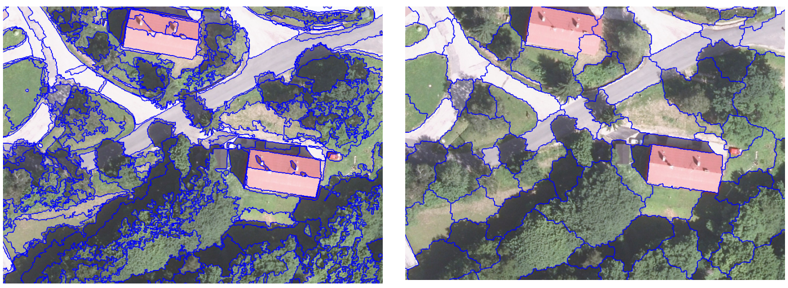


Figure 25.
Example of different shape value
Left: Scale = 100 ; Shape = 0.1 ; Compactness = 0.5
Right: Scale = 100 ; Shape = 0.9 ; Compactness = 0.5

colours which create difficulties when there are different shades of one colour in the same material. When you increase the shape value, the objects try to preserve the block structure, not the actual outlines of the buildings or roads.

The eCognition Reference Book (2012) gives the formula, where the shape is a function of compactness and smoothness, to create the value as below:

$$h_{shape} = w_{cpt} * h_{cpt} + (1 - w_{cpt}) * h_{smooth}$$

Where h_{cpt} and h_{smooth} are compactness and smoothness values respectively and w_{cpt} is the compactness weight value defined by the user.

Similarly for **compactness**, it is a ratio between compactness and relative smoothness. Compactness is defined as a deviation from a compact shape. It can be calculated by a ratio between the pixel perimeter length and the square root of the number of pixels forming the object. It might also be said that the compactness optimises the degree of smoothness based on pixels. Figures 25 and 26 display comparisons of different shape and compactness values. By decreasing the value, the actual shape is endeavoured to be preserved in the resulting object. Inversely, by increasing the compactness value, the smoothness quantity decreases and the object becomes more compacted and condensed.

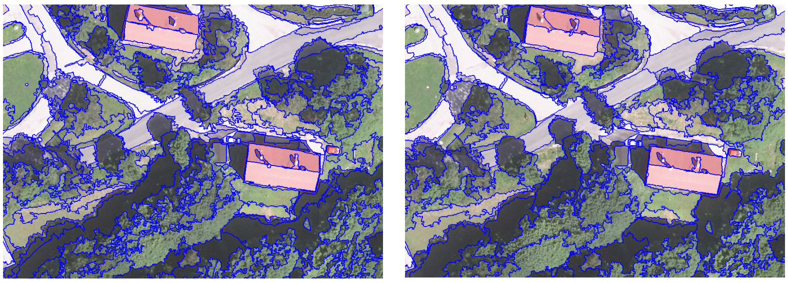
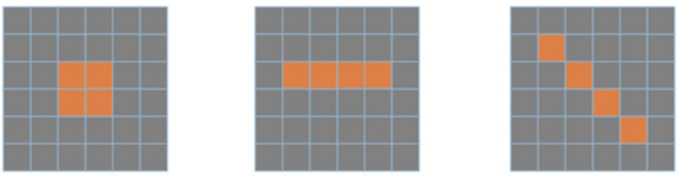


Figure 26.
Example of different compactness value
Left: Scale = 100 ; Shape = 0.1 ; Compactness = 0.1
Right: Scale = 100 ; Shape = 0.1 ; Compactness = 0.9

The figures below shows the effect of changing the compactness parameter. It is important to know that it is still related to the shape value. eCognition Reference book (2012) gives examples that show how compactness and inversely smoothness is computed.



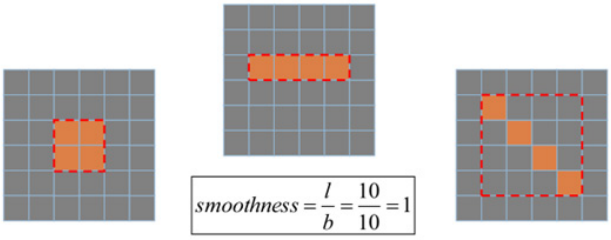
$$cpt = \frac{l}{\sqrt{n}} = \frac{8}{\sqrt{4}} = \frac{8}{2} = 4$$

$$cpt = \frac{l}{\sqrt{n}} = \frac{10}{\sqrt{4}} = \frac{10}{2} = 5$$

$$cpt = \frac{l}{\sqrt{n}} = \frac{16}{\sqrt{4}} = \frac{16}{2} = 8$$

Figure 27.
Formula for calculating compactness
l - size of the object boundry (perimeter)
n - number of pixels

Smoothness defines the homogeneity of a shape. It is a ratio of the pixel perimeter length and shortest possible border length of a bounding box of an object (parallel to the raster). It can be seen the same value can come from different pixels arrangements.



$$smoothness = \frac{l}{b} = \frac{10}{10} = 1$$

$$smoothness = \frac{l}{b} = \frac{8}{8} = 1$$

$$smoothness = \frac{l}{b} = \frac{16}{16} = 1$$

Figure 28.
Formula for calculating smoothness
l - size of the object boundry (perimeter)
b - size of the rectangular boundry

One of the most valuable features in segmentation process is the weighting of the available band layers. The layers can be weighted according to the influence they will have upon the segmentation. It order to retain the linear edges from the orthophoto, the nDSM layer was weighted as zero. This means that the nDSM layer is subdivided according to the same results of the segmentation of the RGB layers. It results in the linear built-up edges of the imagery being kept. Segmentation results using the same settings can be seen when the nDSM is weighted as 1 and the other layers as zero for segmentation. (Figure 29) As can be seen segmention using the nDSM results in jagged building edges.

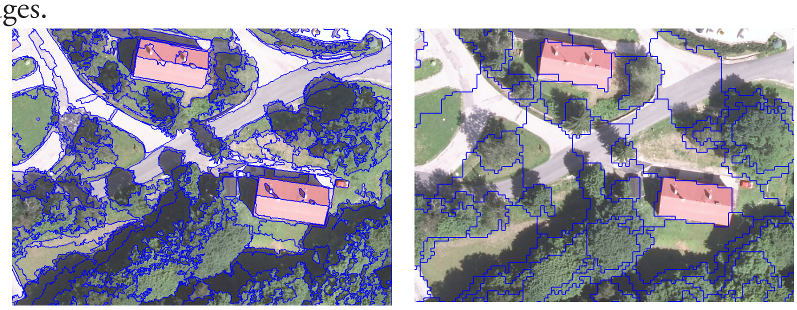


Figure 29.
Example of weighting different layers
Left: weight 1 is given to RGB layers
Right: weight 1 is given to nDSM layer

After becoming familiar with the Multiresolution image segmentation technique which required extensive time, experience and knowledge, the segmentation values were set for segmenting data into image objects, upon which the whole classification process will run. Due to the aspiration of delineating different land cover types, the size and shape for each segmented image object should fit various constraints. The best image object for building extraction would be rectangular with a high value defined for the shape. On the other hand, the roads have greater linear shape which is better described by a higher smoothness value, and so on. That is the reason the author looked for the 'optimal' size, shape and compactness settings for Multiresolution segmentation.

Minho (2012) divides segmentation qualities into three main categories:

1. over-segmentation – the created objects are too small relative to features of interest
2. optimal segmentation
3. under-segmentation – inadequate low numbers of objects which lead to merging different features into one object

The final decision was made after trial-and-error research undertaken by the author and the scale was set to 23. This scale delineates also smaller details quite well for the resolution of available data; shape value is given 0.7, which works well for more rectangular fit objects and also for linear ones; compactness is represented by value 0.3, which still merges similar objects into bigger objects and does not make unreasonably small objects. Segmentation runs on RGB layers only. (Figure 30)

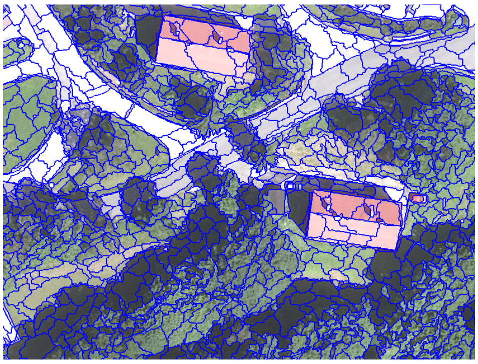


Figure 30.
Final Segmentation Parameters
Scale = 23 ; Shape = 0.7 ; Compactness = 0.3

Initially, the author wanted to work in different segmentation levels for each land cover class, but further detailed research showed no significant improvement for this kind of data sets after applying diverse segmentation values during the classification process.

Class Hierarchy

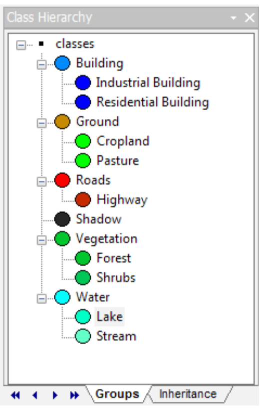


Figure 31.
Land cover classes
Possible ways how 'main' land cover types can be divided

Before the rule set creation will be described, the land cover classes which are to be separated from the data set need to be defined. For the purpose of this paper, the author focuses on the delineation of main land cover types. All other land types can be considered as subclasses and can be additionally classified from the parent class (roads – highways). Successful automatic classification of buildings, roads, high vegetation, water bodies and ground is the fundamental aspect which is to be examined in this study. (Figure 31)

Classification

This stage will discuss the possibilities and steps used to classify objects created in the segmentation process to the right class. This part can be considered the most interesting due to the variability of options provided. There are already many algorithms defined within the software to classify the objects according to spectral, spatial or contextual values, but it requires experience and initiative to combine these in a way that will reach satisfactory results in the field of classification.
As outlined in Problem statement (Chapter 3.1), the highest degree of automation is the main aim to be accomplished and all particular research steps will lead to provide satisfactory automatised classification of chosen land types.

The brief discussion about various image analysis techniques commonly used in the remote sensing industry is provided in Image Analysis section (Chapter 2.2). Knowledge of these techniques is valuable in developing rule bases to extract required features from the image. For this research, the rule based classification approach was chosen, due to no further human interaction needed after development. It consists of predefined conditions and thresholds according to which classification process is executed. Determined features can be explained by a set of rules which follow a series of logical steps. This is built on an existing set of variables available from used data sets – image object feature values (Figure 32). Mostly, the user should be familiar with spectral, spatial and contextual features of the objects and the associated phenomenology. (Table 4) To develop this comprehension there are several data mining techniques to better understand the relationship between a specific thematic class and variables, such as spectral bands, customized features, height information layers.

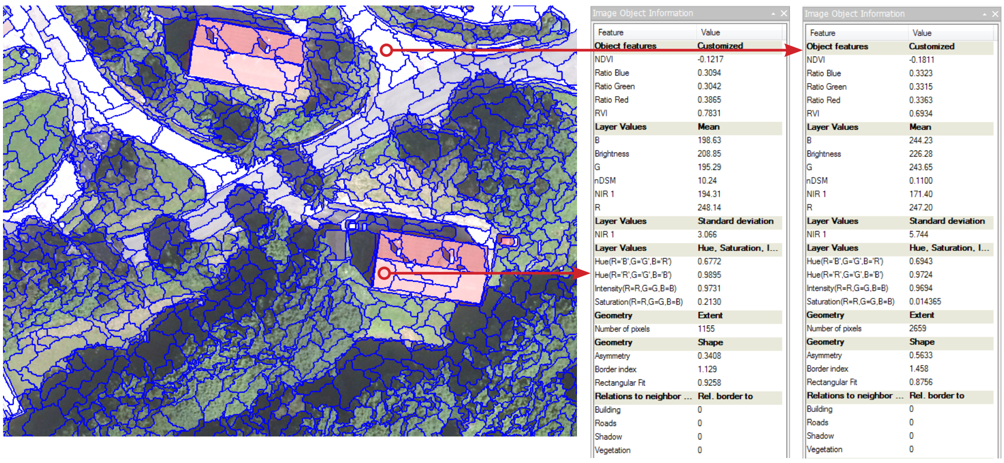


Figure 32.
Identifying image object information

Before the classification rule set creation can start, it is important to make up a strategy (Figure 33) based on the available data sets (Chapter 3.2). To apply simpler and more efficient constraints when classifying particular land cover types, the classification process is done in two levels. Firstly, a coarse classification built on the height information is utilised to separate 'ground' - roads, water and soil, and 'above' objects – buildings, high vegetation. All other classification steps are applied under the coarse classification, which leads into better final results due to dealing with lower number of objects when

Spectral, Spatial and Contextual Parameters

Object Spectral Feature	*Application tips*
Mean spectral values	Vegetation has high NIR values, and water has low spectral values in all bands.
Brightness	Bright objects such as metal rooftops, snow, and others have high brightness values. Water bodies, shadows, asphalt, and other dark objects have low brightness values.
Ratios	Blue ratio is an important tool that can be used to identify water and shadows Green and Red ratios are useful for bare soil identification.
MaxDiff	Useful for impervious surface extraction.
Standard deviation	Large bare soil patches tend to have low red standard deviation.
Object Spatial Feature	
Area	Classify water bodies into lakes or ponds, based on area.
Width	Useful for road extraction.
Length/Width	Classify large narrow features such as streams, rivers, and man-made features like interstates.
Asymmetry	Man-made objects are symmetrical.
Density	Industrial and commercial areas tend to have higher density
Compactness	Man-made objects tend to be more compact than bare soil.
Object Contextual Feature	
Mean difference neighbor	Identify features that have contrast to neighbors such as to roads.
Std to neighbor	Residential features have large standard deviation values with neighbors.
Distance to a class	This is analogous to buffering in GIS and can be used for several applications, such as identifying features within several meters of an oil pipeline, for monitoring and encroachment.
Ratio to scene	Cloud detection where the cloud DN ratio has high values within the scene
Relative border to neighbors	Can be used to resolve water bodies that are misclassified as shadows. Classify an island if a land object is 100% surrounded by water.

Table 4. Navlur (2006)

conditions are applied. Secondly, more comprehensive thresholds using spectral, spatial or contextual information are utilised to classify, or misclassify wrongly assigned object, to the specific class.

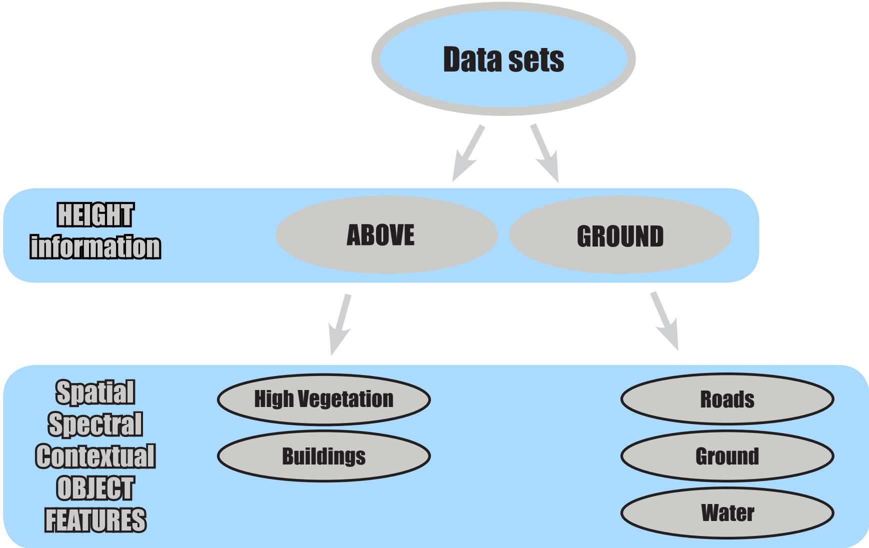


Figure 33.
Classification strategy

During the rule set developing process, there were some specific steps used which might be interesting to mention here. Besides using spectral values available from data sets bands, customised features were defined by the author to investigate whether experimenting might lead to better classification results.
Firstly, the author tries to use the rationing image manipulation technique. It transforms DN value of pixels (in case of OBIA, image objects) in any one band by value of another band. One of the commonly used applications of rationing is to get rid of the dark shadows – they have values near to zero within each spectral band. For the purpose of this study, the author investigates Red, Green and Blue ratio via formula

$$\text{Ratio} = \text{Band}_{R,G,B}/(R+G+B)$$

Secondly, author focuses on applying various types of vegetation indices (VIs) for better delineation of land cover classes. VIs are defined as a combinations of surface reflectance of different wavelengths designed to highlight a particular property of vegetation according plant foliage.
One of the most well known and most frequently used is Normalized Difference Vegetation Index (NDVI). It uses highest absorption and reflectance region of chlorophyll what makes it robust over a wide range of analysing vegetation. NDVI is defined by formula

$$\text{NDVI} = (\text{NIR}-R)/(\text{NIR}+R)$$

Due to an area within the image with high percentage of vegetation available, a Ratio Vegetation Index (RVI) was also tested. It is built upon a knowledge that with increasing canopy the NIR increases too, whereases red will decrease. It is defined by formula

$$\text{RVI} = \text{NIR}/R$$

Further, Soil-Adjusted Vegetation Index (SAVI) was investigated for it's resemblance to the NDVI. It works with some added terms to adjust for different brightness values of background soil. It is defined by formula

$$SAVI = ((NIR-R)/(NIR+R+L))*(1+L)$$

where L represents amount of visible soil and it varies from 0 to 1.

Another commonly used approach in the image analysis field is transforming RGB colour space into three bands in HSI colour space. The result of the transformation represents bands which are in correlation to a specific feature. The equations for transformation are as follows

$$I = (R+G+B)/3$$
$$S = (1-3[\min\{R,G,B\}]/(R+G+B)$$
$$H = \cos^{-1}\{[(R-G)+(R-B)]/2*[(R-G)^2+(R-B)*(G-B)]^{1/2}\}$$

According to ENVI's research there is more than 150 VIs published in scientific literature, but only a small subset are systematically used for image analysis. The author investigates just a few of them with the most promising reviews for this kind of image analysis.

For the final rule set of this study, NDVI, RVI and HIS transformation were successfully applied to improve classification conditions. Here, the author would like to highlight that the final decision was based upon trial-and-error investigation and for different data or purposes other customized features may work better.

Table 5 summarises the features which are used to complement the rule set for automatic extraction of different land cover types developed and investigated in this paper. For finding the right interval for determining feature to create classification condition, image object information can be plotted into a 2D diagram, where axes are representing the feature value of interest. In this case, manually picking examples of examined land cover type objects will represent the 'true' value of searching objects. This approach can be used as a leading tool to clarify object feature values set as axes for creating threshold conditions. (Figure 34) A similar approach is applied during the whole process of classification rule set establishment.

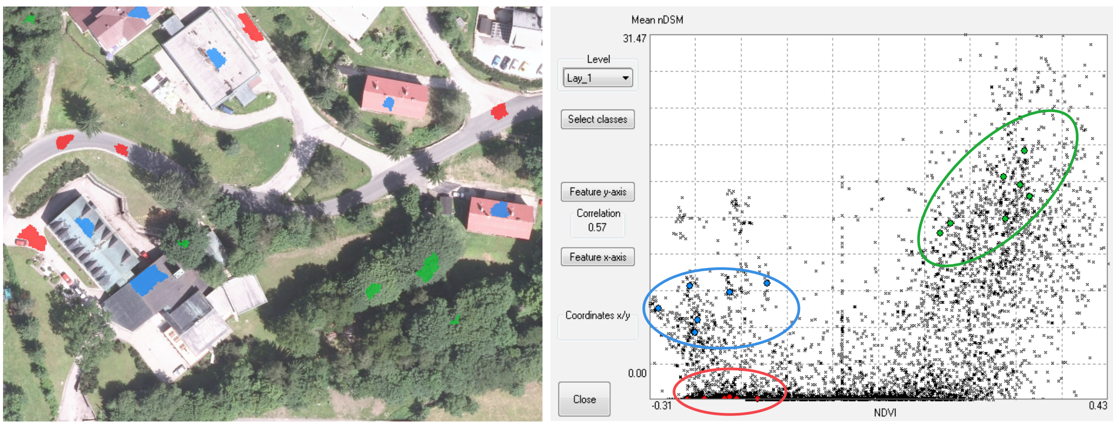


Figure 34.
Plot of image object values for different classes

Left: User's randomly chosen image object samples of different classes

Right: Distribution of image object information in 2D diagram; x- axis represents NDVI; y- axis represents nDSM (height)

* one black cross represent one image object from the image

Used Features for Developing Classification Rule Set

Object Feature	*Application*
nDSM	Determines ground and above object extracted from height information.
NDVI	Distinguishes vegetation from man-built structures.
RVI	Separates vegetation and man-built structures.
Relative border	Helps assign object to the class according its neighbours.
Merge region	Joins image object assigned to the same class into one.
Border index	Describes how jagged the image object is in comparison to rectangular approximation.
Rectangular fit	Describes how well an image object fits into a rectangular with the same parameters.
HSI Transformation Hue	Helps assign shadows to roads or ground.
HSI Transformation Saturation	Helps distinguishes roads from ground.
Mean NIR	Uses mean intensity value from an image object.
Brightness	Calculates mean value of intensity from available spectral bands.

Table 5.

Possible ways of how to obtain interval values for image object features have been showed so far. But how to know which image object feature should be used for the best classification of individual class? Different land cover types contain various characteristics through different features. To find the right one, the author uses the image object transformation according the feature value of interest. (Figure 35) Image objects are displayed in shades of grey within the feature interval. In this way, it is possible to visually investigate whether this specific feature can be used for distinguishing features e.g. roads from ground.

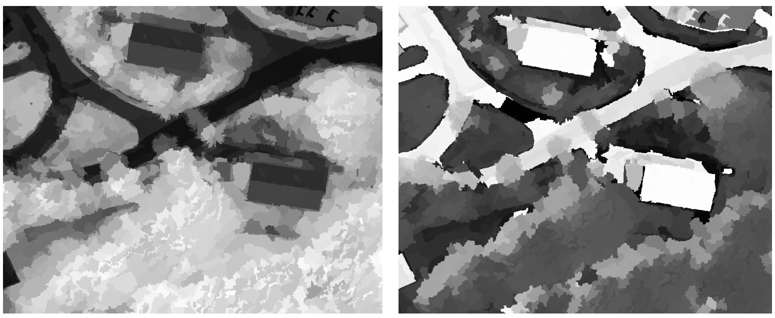


Figure 35.
Image object color transformation according the specific feature
Left: NDVI values
Right: HSI Hue values

The main approaching steps for rule set creation were mentioned above. For a closer understanding of mathematics behind algorithms used in this study, the author refers you to eCognition Reference Book (2012), where basic principles available in eCognition software are given or to Computer Processing Of Remotely-Sensed Images (Mather 2004) for a more advanced description of the image processing problematics. The developed rule set with a step-by-step description of utilizing it in the investigation area can be found in Appendix C, with figures showing classification changes after running single rule set conditions.

The classification results based on a 'one click' rule set can be seen in Figure 36. Evaluation, transferability and overall limitations of this process are given in the following Discussion (Chapter 4).

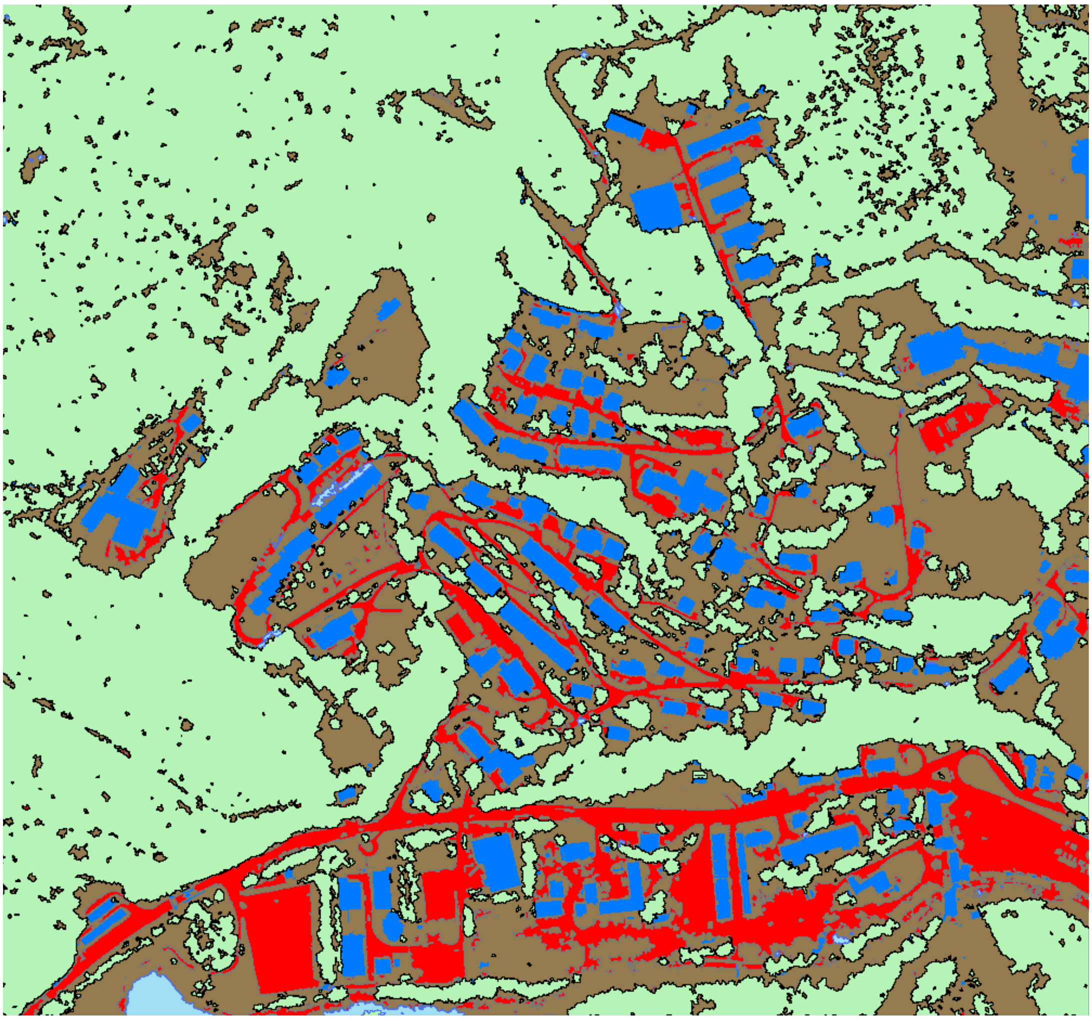


Figure 36.
Classification results
Exported classification results in a shape file via full license eCognition software.

4 Discussion

This chapter will provide a discussion summary based upon the Investigation in the previous chapter. To begin with, two quality analysis procedures will be undertaken to verify classification process. Next the limitations and problems that have been encountered during the process and how they have been solved, or why they have not, will be discussed.

4.1 Quality Analysis

This section will undertake an analysis of the final results. Quality analysis can be evaluated through numerical comparison and transferability of developed approach to another area. There will be two measures of quality to assess the classification procedure.

> 1. Confusion matrix – will reflect a final accuracy of the classification process
> 2. Transferability of the rule set – will describe how well can be developed procedure applied to a different area

The Confusion matrix assessment technique will be applied to evaluate the quality of the classification by calculating an overall of correctly classified accuracy percentage.
The next assessment of accuracy is based on analysis between the reference and classified polygons. The final evaluation of accuracy will be to test how well the rule set has been constructed by applying it to another test area to assess how well it can work there.

Confusion Matrix

A confusion matrix is a way of comparing two maps through quantitative sampling by cross validation. One map is the original image or a reference map and the other is the resultant map from the image classification process. A variety of measures can be computed to describe the accuracy of the classified map with respect to the reference map.
A sample area (m^2) is chosen on the reference map and compared with the corresponding area in the classified map. Rows of the matrix (Table 6) represent the reference map, and the columns represent the resultant classes of the classified image. If the 'reference' sample is correctly classified, this is recorded in the matrix as so. If not, it is recorded in the appropriate class column.
For example, if a 'building' sample in the reference map is correctly classified as 'building', then a tally starts in the row and column corresponding to the 'building' class ie. 'A'. However, if 'building' is incorrectly classified as a road then a tally is started in the row 'buidling' and column 'road', ie. 'B'.

Calculation of Confusion Matrix

		Classification Map: Building	Classification Map: Road	Row Total	Producer's Accuracy	Errors of Ommission
REFERENCE MAP	Building	**A**	**B**	**E**	**A/E**	**B/E**
REFERENCE MAP	Road	**C**	**D**	**F**	**D/F**	**C/F**
REFERENCE MAP	Column Total	**G**	**H**			
	User's Accuracy	**A/G**	**D/H**			
	Errors of Commission	**C/G**	**B/H**			

Overall Accuracy **(A+D)/(E+F)**

Mean Accuracy **(A/E+D/F)/n** n - no. of classes

Table 6.

There are various measures available to be read which are usually expressed as a percentage including:

Measures of correctly classified areas

- The "Overall Accuracy" for the classification, can be calculated by the diagonal sum divided by the total number of square meters
- The proportion of correctly classified areas is calculated for each row from the ratio of the number of correctly classified areas and the total number of areas in that row. This is called " Producer's Accuracy" because it is a measure of how well the analyst did when generating the classification map.
- "User's Accuracy" can be similarly calculated in the same way as "Producer's accuracy" using the proportion of correctly classified areas and total number of areas in each column. This gives a measure of the probability of correctly labelled areas during the classification process.
- "Mean Accuracy" can be computed from the sum of the producers accuracy divided by the number of classes.

Measures of incorrectly classified areas

- The number of square meters that have been incorrectly attributed to a class, can be computed by dividing the total number of the non-diagonal cell values in each row by the sum of the row total giving the "Error of Omission".
- Similarly, the same method can be used upon the columns, computing the "Error of Commission".
- Checks: The "Error of Omission" is 100% minus the "Producer's Accuracy" and "Error of Commission" is 100% minus the "User's Accuracy".

The total amount of correctly classified areas can be read from the sum of the cell values along the diagonal. Checks can be done by the sum of the row totals and sum of the column totals gives the total area number of the entire matrix.

To create the reference map for the confusion matrix, the orthophoto was manually vectorised in ArcMAP. It is important to note that the accuracy of the reference map corresponds to that of manual vectorisation, and as such cannot be measured.
However, this is deemed a sufficient technique for the purpose of this procedure. The vectorising of the orthophoto makes this process semi-automatic, otherwise you would have to manually assess the classification of each image object, increasing time spent and decreasing the accuracy, especially if there is a large investigation area.

To obtain the data for final confusion matrix a manually vectoring reference map was undertaken to evaluate classification developed in this study. This involved the intersection between reference and classification map, calculate the overlap area across different classes and the resulting numbers were then input into the Table 7 to achieve an overall accuracy.

Confusion Matrix (Area 1)

		CLASSIFICATION MAP				*Row*	*Producer's*	*Errors of*
		Building	*Road*	*Vegetation*	*Ground*	*Total*	*Accuracy*	*Ommission*
REFERENCE MAP	*Building*	7542	28	301	280	8152	93%	7%
	Road	349	11436	229	2674	14687	78%	22%
	Vegetation	22	33	25852	1611	27519	94%	6%
	Ground	725	1150	2814	29361	34249	86%	14%
	Column Total	8638	12647	29196	34126			
	User's Accuracy	87%	90%	89%	87%			
	Errors of Commission	13%	10%	11%	13%			
	Overall Accuracy	88%						
	Mean Accuracy	88%						

Table 7.
Values represent area in square meters

The confusion matrix shows overall accuracy of developed classification rule set to be 88%. Here, the author would like to point out that manual vectorisation may carry some difficulties, especially with delineation of non-manmade land covers. In some point, it might affect final accuracy. Due to mentioned problem, this paper provides detailed analysis of buildings, which is believed, can be manually delineated with satisfactory accuracy.
The 44 buildings were chosen from study area 1, where the intersection of reference map and classification map was investigated. The accuracy of correctly classified buildings area as a building is 93%. The table and figure of this process is in Appendix D, where the table involves area number (m^2) of true value (from reference map), classified area extracted by developed rule set and overhanging value (%) which shows if classified building 'bleeds' onto the surroundings (is over-classified - possitive value) or the building is classified incompletely (negative value).

Transferability of Classification Rule Set

The second method of evaluation that will be tested in this research is how well the classification rule set would work upon another area (Area 2). If the rule set is able to classify an acceptable amount of classes then it can be deemed transferable.
The test Area 2 (Svoboda nad Úpou) chosen for this analysis is an area ~22 km east of the Area 1 (Špindlerův Mlýn). As can be seen in Figure 37, the area is a combination of infrastructure and residential area with small amount of vegetation cover. The same procedures were undertaken to create the nDSM and then in eCognition the same rule set was applied to classify the identical land cover classes.

Figure 37.
Area 2 - Svoboda nad Úpou
Subset of training area where transferability of classification rule set was tested

To evaluate quality analysis for this area, the confusion matrix was created following the same principles described before (Table 8).

Confusion Matrix (Area 2)

		CLASSIFICATION MAP				Row	Producer's	Errors of
		Building	*Road*	*Vegetation*	*Ground*	*Total*	*Accuracy*	*Ommission*
REFERENCE MAP	*Building*	4658	15	80	252	5005	93%	7%
	Road	25	2103	31	895	3054	69%	31%
	Vegetation	12	3	4092	656	4762	86%	14%
	Ground	308	193	792	10058	11352	89%	11%
	Column Total	5003	2314	4995	11861			
	User's Accuracy	93%	91%	82%	85%			
	Errors of Commission	7%	9%	18%	15%			
	Overall Accuracy	87%						
	Mean Accuracy	84%						

Table 8.
Values represent area in square meters

It can be seen that the overall accuracy of developed classification in different area is 87%. The number is very similar to area 1 where the rule set was originally created what makes developed rule set transferable and utilisable as a decent 'building block' for similar purposes based on this kind of the data sets.
However, small differences in classification accuracy of single classes through these two areas can be seen. Most likely, it is caused by different time of image acquisition which leads to slightly different spectral behaviour upon which is the rule set based on.

4.2 Limitations

This section will outline some problems which were encountered during the classification process with explanation of the causes.

Firstly, author would like to explain why the class 'water' was not a part of classification evaluation but was defined in the final rule set. The reason is that area which was used for rule set creation contains just a small bit of water body what limited whole process of defining proper classification condition. Due to this, author considers inappropriate to evaluate small sample and includes it to the overall accuracy analysis.

Further, with combination of different data sources, data based problems appear. In some areas with the nDSM covering a building would overhang the imagery building boundary. During the classification process, this meant that some areas adjacent to the buildings would be classified building when using height thresholds. (Figure 38) This issue is not a result of incorrect geo-referencing but problem might be caused in process of orthorectifying aerial image. Most likely, it is result of applying different interpolation techniques in the place of building footprints in DTM upon which the imagery was orthorectified. But since there is no available information how was this certain orthophoto created, the 100% confidence cannot be stated here. In this stage of investigation, this problem cannot be avoided, just taken into consideration when making the rule set and implement other than height information for correct delineation of buildings.

Figure 38.
Shifted data sets

Due to similar features and the height to the other classified buildings, the bridge was incorrectly identified as a building as seen in Figure 39. With the current rule set, this issue cannot be solved automatically and further rules would need to be derived to get rid of this misclassification.

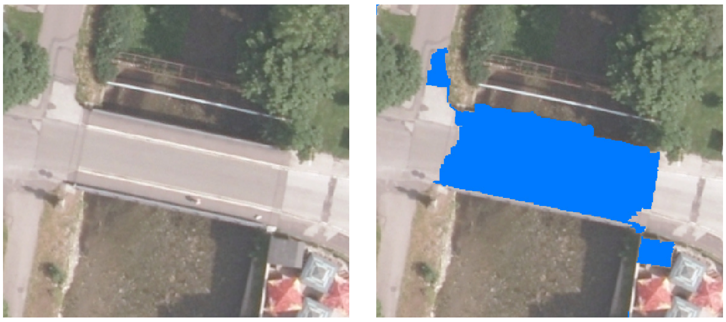


Figure 39.

Misclassification

Bridge misclassified as a building due to similar height and other values

Another unsolved but challenging problems appeared with classifying the shadows to the proper land cover type. Especially, successful classification of shadows covering roads was considered as a crucial and an additional time was given to solve it. Using HSI colour transformation (Figure 35) partly helps to distinguish roads covered by shadow, but it does not work throughout whole area of investigation (Figure 39).

Figure 40.

Classification of shadows

Top: Shows example of successful classification of shadow to the class road

Bottom: Shows a part of road which was not successfully classified from shadows

After concluding the evaluation phase there are some general aspects that have emerged from the results discussion that can be noted here. When searching for the optimal rule set, an appropriate balance of classification needs to be taken into account. This means making rules that will achieve the maximum classification, and bearing in mind that some will remain misclassified. With the inherent data characteristics and limitations as mentioned previously, 100% classification is not reasonable to expect. In this stage, readers must emphasise the need for a balance between what is statistically sound and what is practically attainable.

In terms of the quality assessment testing, in overall the results were positive. The confusion matrix gave an overall accuracy of 88 % in Area 1, where the rule set was originally developed, and 87% in Area 2, where the transferability of the developed approach was tested. These numbers indicate a good rule sets performance with reference to similar studies in recent years.

5 Conclusion

The aims of this master thesis were to perform automated process of different land cover types classification from remotely sensed data sets, which represent fast and efficient large scale mapping tool. Main focus was on development of the classification rule set of the fusion between imagery and LiDAR data sets to delineate the land cover classes with acceptable success.

In relation to the problem definition, the choice to combine high resolution orthophoto with nDSM generated from LiDAR data using OBIA methods turned out to be a successful selection. The data fusion worked satisfactory and each data type complemented the other's limitations relatively well, particularly the height value information made a significant improvement when distinguishing above features from the ground. On the other hand, it caused some subsequent issues that needed to be solved during rule set development which was held within eCognition software.

The final results were assessed against the initial criteria in the problem statement (Chapter 3.1) as are as follows:

> 1. Once the rule set is developed the entire process becomes highly automated because there is no more manual processing required after its creation, it can be simple executed upon the data.
> 2. The overall resulting accuracy is 88% of correctly classified examined land covers areas upon the confusion matrix, which is the main accuracy assessment tool for this field. The additional evaluation run over the comparison measures between the reference map and classification results of buildings footprints only. It bypassed the problems with manually vectorising land cover types which have unclear boundaries and supported the higher overall percentage of correctly classified buildings – 93% without any particular blunders.
> 3. The transferability of the rule set into a different geographical area was held as a final evaluation test of the robustness and how well the rule set could be used for different areas. The results from applying the same rule set on the same data sets in different area showed high potential. Overall accuracy reached 87% and for buildings only, the accuracy was 93%.

From the experience acquired so far, it is author's opinion that the classification technique based on orthophoto and LiDAR data fusion has proven recent research in the image classification field and is a viable option to increase final accuracy. Additionally, there is scope for further exploration into using more data types to improve classification process, such as information from geospatial database or national cadastral layer, which can be used as an identification tool for applying right classification conditions.

"I hope that posterity will judge me kindly, not only as to the things which I have explained but also as to those which I have intentionally omitted so as to leave to others the pleasure of discovery."

Ren´e Descartes

References

Abdulrahman, A., *"Remote Sensing."* First edition, University of Technology, Republic of Iraq, 2010.

Anderson, J. R., Hardy, E. E., Roach, J. T., Witmer, R. E., *"A Land Use and Land Cover Classification System for Use with Remote Sensor Data."* U.S. Geological Survey, Professional Paper 964, p. 28, Reston, VA, 1976.

Aplin, P., Smith, G.M., *"Advances in Object-based image classification."* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing, 2008.

Baatz, M., Schäpe, A., *"Multiresolution Segmentation–An Optimization Approach For High Quality Multi-Scale Image Segmentation."* Angewandte geographische informationsverarbeitung, Edited by: Strobl, J., Blaschke, T. and Griesebner, G. 12–23. Heidelberg: Wichmann-Verlag, 2000.

Baumann, P., *"Introduction to Remote Sensing."*, State University of New York, USA, 2008.

Bhatia, N., *"Survey of Nearest Neighbor Techniques."*, International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010.

Blaschke, T., Lang, S., *"Object based image analysis for automated information extraction - a synthesis."* MAPPS/ASPRS 2006 Fall Conference November 6 – 10, 2006, San Antonio, Texas.

Blaschke, T., *"Object based image analysis for remote sensing"*, Z_GIS Centre for Geoinformatics and Department for Geography and Geology, University of Salzburg, ISPRS Journal of Photogrammetry and Remote Sensing 65, 2010, 2–16.

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

Cambell, J. B., Wynne, R. H., *"Introduction to Remote Sensing."*, Fifth edition, ISBN 0-470-84918-5, 2011.

Chmiel, J., Fijałkowska, A., *"Thematic accuracy assessment for object based classification in agriculture areas: Comaparative analysis of selected approaches."* Warsaw University of Technology, Faculty of Geodesy and Cartography, Division of Photogrammetry, Remote Sensing and Spatial Information Systems, Pl. Politechniki 1, 00-661 Warsaw, Poland, 2012.

Coghill, J., *"Digital Imaging Technology."* Based on material developed by Albert Theuwissen, 2003.

De kok, R., T. Schneider and U. Ammer., *"Object Based Classification And Applications In The Alpine Forest Environment."* In: *"Fusion of Sensor Data, Knowledge Sources and Algorithms."* Proceedings of the Joint ISPRS/Earsel Workshop, 3–4 June 1999, Valladolid, Spain. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 32(7-4-3W6), 1999.

eCognition® Developer 8.7 Reference book. 2012 Trimble Germany GmbH, München

eCognition® Developer 8.7 User Guide. 2012 Trimble Germany GmbH, München

Escalante-Ramírez, B., *"Remote Sensing – Advanced Techniques and Platforms."* Copyright © 2012 InTech, 2012.

European Communities, *"Manual of concepts on land cover and land use information systems."* Luxemburg: Office for Official Publications of the European Communities, 2001, ISBN 92-894-0432-9

Falkner, E., Morgan, D., *"Aerial Mapping Methods and Applications."* Second Edition, CRC Press, 2001, Print ISBN: 978-1-56670-557-8 eBook ISBN: 978-1-4200-3244-4

Ferencz, J., Dalley, R. A., *"Classifying Remotely Sensed Data for 3D Building Visualisation."* Semester project, supervisors Peter Cederholm and Jens Juhl, L8SM2012_02, Aalborg University, Denmark, 2012.

Ghorbani, A., Bruce, D., Tiver, F., *"Specification: A problem in rangeland monitoring."* In: Proceedings of the 1st International Conference on Object-based Image Analysis (OBIA), 4th-5th July 2006, Salzburg, Austria, 2006.

Hay, G.J., Castilla, G., "Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT)." International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences 35 (4/C42), 2006.

Hermosilla, T., Ruiz, L.A., Recio, J.A., Cambra-López, M., *"Assessing contextual descriptive features for plot-based classification of urban areas."* Landscape and Urban Planning 106 (2012) 124-137, 2012.

Horňáká, M., *"Automatická extrakce budov a zpevněných ploch z dat velmi vysokého rozlišení v suburbánní zóně Prahy."* Dipomová práce, Univerzita Karlova v Praze, Fakulta Přírodovědecká, Katedra Aplikované geoinformatiky a kartografie, 2011.

Hutson, R., *"Remote Sensing and Land Classification."* Remote Sensing Group - Plymouth Marine Laboratory, 2006.

Jiang, N., Zhang, J.X., Li H.T., Lin, X.G., *"Object-Oriented Building Extraction By Dsm And Very High-Resolution Orthoimages."* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37(3b): 441-446, 2008.

Karl, J. W., Maurer, B. A., *"Multivariate correlations between imagery and field measurements across scales: comparing pixel aggregation and image segmentation."* Landscape Ecology. 4: 591-605, 2009. DOI: 10.1007/s10980-009-9439-4.

Laliberte, A.S., Fredrickson, E.L., Rango, A., *"Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands."* Photogrammetric Engineering and Remote Sensing 73(2): 197-207, 2007.

Lee, D., Lee, K., Lee, S. *"Fusion of LIDAR and imagery for reliable building extraction."* Photogrammetric Engineering & Remote Sensing 74 (2): 215–226, 2008.

Mather, P. M., *"Computer Processing Of Remotely-Sensed Images: An Introduction."* 3rd edition, John Wiley & Sons, 2004.

Milton, J.S., Arnold, J.C., *"Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences."* McGraw Hill, 1995.

Minho, K., *"Object-based Spatial Classification of Forest Vegetation with IKONOS Imagery."* A Dissertation Submitted to the Graduate Faculty of The University of Georgia, 2009.

Natural Resources Canada, *"Fundamentals of Remote Sensing."*

Navulur, K., *"Multispectral image analysis using the object-oriented paradigm."* CRC Press/Taylor & Francis, ISBN-13: 978-1-4200-4306-8, 2006.

Paine, D. P., Kiser, J. D., *"Aerial Photography and Image Interpretation."* Third Edition, Copyright © 2012 John Wiley & Sons, Inc. 2012.

Rees, W.G., *"Physical Principles of Remote Sensing."* Second edition, Scot Polar Research Institute, University of Cambridge, 2001.

Rottensteiner, F., Jansa, J., *"Automatic Extraction of Buildings From Lidar Data and Aerial Images."* Symposium on Geospatial Theory, Processing and Application, Ottawa, 2002.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K. *"Building detection by fusion of airborne laserscanner data and multi-spectral images: Performance evaluation and sensitivity analysis."* ISPRS Journal for Photogrammetry and Remote Sensing 62(2): 135-149, 2007.

Salah, M., Trinder, J.C., Shaker, A. *"Evaluation Of The Self-Organizing Map Classifier For Building Detection From Lidar Data And Multispectral Aerial Images."* Journal of Spatial Science 54(2): 15-34, 2009.

Shiba, M., Itaya, A. *"Using eCognition for Improved forest Management and monitoring Systems in Precision Forestry."* IUFRO Precision Forestry Symposium, 2006.

Stilla U., Rottensteiner F., Paparoditis N., *"Dense Matching in High Resolution Obliqe Airborne Imagery."* France, (Eds) CMRT09. IAPRS, Vol. XXXVIII, Part 3/W4, 2009.

Stoney, W. E., *"ASPRS Guide to Land Imaging Satellites. Bethesda."* MD: American Society for Photogrammetry and Remote Sensing, www.asprs.org/news/satellites/ASPRS_DATABASE_021208.pdf. 2008.

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

Walter, V., *"Object-based classification of remote sensing data for change detection."* ISPRS Journal of Photogrammetry & Remote Sensing 58 (2004) p. 225–238, 2004.

Weih, R., Riggan, N., *"Object-based classification vs. Pixel-based classification: Comparitive importance of multi-resolution imagery."* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C7, 2009.

Yan, G., Mas, J., Maathuis, B., Xiangmin, Z., Van Dijk, P., *"Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China."* International Journal of Remote Sensing Vol. 27, No. 18, 4039–4055, 2006.

Yunhao Ch., Wei S., Jing L., Zhongping S., *"Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas."* Advances in Space Research 43 (2009) 1101–1110. 2008 COSPAR. Published by Elsevier Ltd, 2008.

Zabuawala, S., Nguyen, H., Wei, H., Yadegar, J., *"Fusion of LiDAR and Aerial Imagery for Accurate Building Footprint Extraction."* Image Processing: Machine Vision Applications II, edited by Kurt S. Niel, David Fofi, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7251, 2009.

Zuva, T., Olugbara, O. O., Ojo, S. O., Ngwira, S. M., *"Image Segmentation, Available Techniques, Developments and Open Issues."* Canadian Journal on Image Processing and Computer Vision Vol. 2, No. 3, March 2011.

Appendix