Datafusion and semi-automatic map updating

Jens Popp Andersen
s001480

March 2007

Informatics and Mathematical Modelling

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

IMM-MSc:

Abstract

This thesis deals with semi-automatic map updating by the use of fusion of image data and elevation data. Three methods are developed for change detection of buildings in a digital map database. The data used for change detection consists of digitally acquired multispectral aerial imagery, digital elevation data, and the existing map database.

A part of this thesis also deals with the visual comparison of analogue scanned aerial images and digitally acquired aerial images, as well as a comparison based on information content and image noise. This comparison shows a significantly higher level of information content and much lower noise in the digitally acquired images.

The image data for change detection consists of a 4 layer image of RGB and near-infrared (NIR). The elevation data consists of a digital surface model (DSM) and a digital terrain model (DTM) used for the creation of a normalised digital surface model (nDSM). The map database is the TOP10DK from the National Survey and Cadastre - Denmark (KMS).

All three change detection methods are based on the unsupervised classification algorithm, Fuzzy Maximum Likelihood Estimation (FMLE), of either 5 layer images or 4 layer images. The algorithm both takes the spectral characteristics and the spatial characteristics into account. The 5 layer images contains the RGB, NIR, and the logarithm of the nDSM.

The first method is based on the FMLE classification of the 5 layer image. The two clusters that appears to reflect buildings are extracted, merged to one class and compared with the existing buildings in the map database. This method detects the majority (at least 80 %) of the supposed new buildings. However, several false alarms complicates the change detection.

The second method involves the FMLE classification of the 4 layer image and extraction and merging of the building clusters. After that the method is divided into two submethods 2A and 2B. In 2A an intersection between the building class and a Object Above Terrain (OAT) class is performed. The OAT class is derived from the nDSM where all elevations are above 2.5 m. The final building class is used for change detection. Method 2B also involves the class intersection with a NDVI class that describes all non-vegetated areas. Both submethods detect the majority (at least 75 %) of the supposed new buildings with a maximum of 10 false alarms.

In the third method, FMLE classification of both 4 and 5 layer images overlaid with a TOP10DK mask is performed resulting in a training sample where only existing buildings are clustered. This training sample is used in a supervised classification algorithm, Maximum Likelihood Estimation (MLE), applied with a Mahalanobis distance threshold. The method is also tested with a prior probability image defined by the existing buildings in the map database.
This third method performs worse than the two methods above. The Mahalanobis distance threshold is found to be too low. However, the inclusion of the prior probability image show to have an significant effect.

The conclusion is that the FMLE classification algorithm can be used for change detection with certain reservations. It can not be used creation of a training sample for MLE classification unless the most optimal Mahalanobis distance parameter is found.

Resumé

Dette eksamensprojekt omhandler semi-automatisk kortopdatering ved brug af fusion mellem billeddata og højdedata. Tre metoder er udviklet til ændringsudpegning af bygninger i en digital kortdatabase. Dataet der bruges til ændringsudpegning, består af digitalt optagne multispektrale luftfotografier, digitale højdedata og den eksisterende kortdatabase.

En del af dette projekt omhandler også den visuelle sammenligning af skannede analoge luftfotografier og digitalt optagne luftfotografier samt en sammenligning baseret på indhold af information of billedstøj. Denne sammenligning viser et væsentlig højere indhold af information og en meget lavere støj i det digitalt optagne fotografi.

Billeddata til ændringsudpegningen består af et 4 lags billede af RGB og nær infrarød (NIR). Højdedataene består af en digital overflademodel (DSM) og en digital terrænmodel (DTM), der bruges til dannelse af en normaliseret digital overflademodel (nDSM). Kortdatabasen er TOP10DK fra Kort & Matrikelstyrelsen (KMS).

Alle tre metoder til ændringsudpegning er baseret på en usuperviseret klassifikationsalgoritme, Fuzzy Maximum Likelihood Estimation (FMLE) af enten 5 lags billede eller 4 lags billede. Algoritmen tager både højde for de spektrale egenskaber og de spatielle egenskaber. 5 lags billedet består af RGB, NIR og logaritmen til nDSM.

Den første metode er baseret på en FMLE klassifikation af 5 lags billedet. De to klynger der bedst beskriver bygninger udtages, sammenlægges i en klasse og sammenlignes med den eksisterende kortdatabase. Denne metode detekterer størstedelen (mindst 80 %) af de forudsagte nye bygninger. Men flere falske alarmer komplicerer ændringsudpegningen.

Den anden metode involverer FMLE klassifikationen af 4 lags billedet og udtagning og sammenlægning af bygningsklyngerne. Derefter deles metoden i to delmetoder 2A og 2B. I 2A lægges fællesmængden af bygningsklassen og en *objekt over terræn* (OAT) klasse i en endelig bygningsklasse. OAT klassen er afledt af nDSM, hvor alle højder er over 2.5 m. Den endelige bygningsklasse bruges til ændringsudpegning. Metode 2B involvere også fællesmængden med en NDVI klasse der beskriver ikke beplantede områder. Begge delmetoder detekterer størstedelen (mindst 75 %) af de forudsagte nye bygninger med et maksimum på 10 falske alarmer.

I tredje metode, bestemmes en FMLE klassifikation af både 4 og 5 lags billeder overlagt med en TOP10DK maske, hvilket resulterer i et træningssæt, hvor kun eksisterende bygninger klassificeres. Dette træningssæt bruges i en superviseret klassifikationsalgoritme, Maximum Likelihood Estimation (MLE), påført en tærskelværdi for Mahalanobis afstand. Metoden testes også med et billede defineret ved en a priori sandsynlighed for de eksisterende bygninger i kortdatabasen.
Denne tredje metode fungerer dårligere end de to metoder ovenfor. Tærskelværdien for Mahalanobis afstand findes at være for lav. Men inklusionen af a priori sansynlighedsbilledet viser at have en væsentlig betydning.

Konklusionen er at FMLE klassifikationsalgoritmen kan bruges til ændringsudpegning med visse forbehold. Den kan ikke bruges til dannelse af et træningssæt til MLE klassifikation uden at den mest optimale parameter for Mahalanobis afstand findes.

Preface

This master thesis (M.Sc. thesis) was prepared at Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark (DTU) and corresponds to 30 ECTS points. The internal supervisor was Allan Aasbjerg Nielsen (IMM, DTU) and the external supervisor was Kristian Keller (COWI A/S, Kgs. Lyngby).

First of all I wish to thank Allan Aasbjerg Nielsen for help and guidance and inspiring discussions throughout this project. I also wish to thank Kristian Keller for help and for kindly providing data for this project. Finally I wish to thank Thomas Knudsen from the Danish National Space Center (DNSC) for inspiration and help.

It is assumed that the reader has some prior knowledge about image analysis and multivariate statistical analysis. Furthermore it is an advantage having some prior knowledge about map databases, image data and elevation data.

The thesis consists of a report, an appendix visualizing the data that comprises the basis of this project, and an enclosed CD-ROM that contains the project data and this thesis in PDF.

Kgs. Lyngby, March 2007

Jens Popp Andersen, s001480

Abbreviations

ALS Airborne Laser Scanning.

ASCII American Standard Code of Information Interchange. Code representing text in computers.

CCD Charge-Coupled Device. Light sensitive silicon chip used in digital cameras.

CIR Colour-InfraRed. Imagery combined by a infrared band and two visible bands - often red and green.

DDO Danmarks Digital Ortofoto, Danish for *The digital orthophoto of Denmark*.

DEM Digital Elevation Model. Other term for DTM.

DSM Digital Surface Model. Surface model including objects above surface like buildings and trees.

DTM Digital Terrain Model. Terrain model without objects above surface.

DVR90 Dansk Vertikal Reference 1990, Danish for *Danish Vertical Reference 1990*.

ECW Enhanced Compressed Wavelet. Lossy image compression format.

ED50 European Datum 1950.

ENVI ENvironment for Visualizing Images. Software for image processing in Remote Sensing.

ETRS89 European Terrestrial Reference System 1989. European datum, identical to WGS84.

FMLE Fuzzy Maximum Likelihood Estimation. Algorithm for unsupervised classification.

GIS Geographic Information System.

GPS Global Positioning System. Global position determination via satellites.

GSD Ground Sample Distance. Similar to GSI (Ground Sample Index). Indicates the ground true size of an image pixel.

GUI Graphical User Interface.

INS Inertial Navigation System. System for rotation and slope measurement of an aircraft.

KMS Kort & MatrikelStyrelsen, Danish for *National Survey and Cadastre - Denmark*.

LC Lower Center.

LIDAR LIght Detection And Ranging. Laserscanner for DEM creation.

LL Lower Left.

LR Lower Right.

MLE Maximum Likelihood Estimation. Algorithm for supervised classification.

nDSM normalized Digital Surface Model. nDSM = DSM - DTM.

NDVI Normalized Difference Vegetation Index.

NIR Near Infra-Red. Indicating near infra-red spectrum of the electromagnetic waves.

RGB Red, Green, Blue. Normal colour imagery combined by a red, green and a blue band.

RGBNIR Red, Green, Blue and Near-Infra-Red. 4 layer image of the red, green, blue and near-infrared band.

TIFF Tagged Image File Format. Lossless image compression format.

TIN Triangular Irregular Network. Interpolation method for irregular sampled points.

TOP10DK The Danish Topographic Base Map Database.

UTM Universal Transverse Mercator. Global map projection system.

WGS84 World Geodetic System 1984. Global earth-centric datum.

UL Upper Left.

UR Upper Right.

Contents

Chapter 1

Introduction

The world is constantly changing and registrations or databases of any relevance needs to be updated instantly to avoid misunderstandings. This work of maintenance is very time consuming and if just some of it can be automatized, a lot of money can be saved. As buildings are torn down and new are build, and areas of different land use are changing from one to another, maps, in both paper and digital form, needs maintenance.

1.1 Background

Since the 1960's the National Survey and Cadastre - Denmark (KMS) has managed the job of overflying Denmark with an airborne camera, taking aerial photos for photogrammetric purposes, see KMS [2007]. These traditional analogue aerial cameras have only undergone little development and since each photo have size of 23 by 23 cm, they take up a lot space for archiving. Traditionally the photos have been used for creation of orthophotos, drawing and updating topografical- and technical maps, and elevation curves by the use of photogrammetry, Mikhail et al. [2001].
During the last 10 years the scenario described above has changed a lot. Firstly,

the technology for spaceborne remote sensing has improved, increasing the resolution and therefore the ability to detect smaller objects and areas. This has improved the possibility of using satellite imagery for large scale map updating, surveying, and surveillance. Secondly the introduction of the aerial digital camera has set new standards for remote sensing. Now it is possible to detect the infra-red spectrum along with the visible light with an airborne sensing system. The developers and manufacturers of these new digital cameras promise better data quality and acquisition than for the traditional aerial film cameras.

1.1.1 The digital camera

In recent years the technology of the CCD-chip has improved, see 2.2.5, resulting in an increased image resolution and image size. Since the applied resolution for scanned analogue aerial photos are of 14 μm, the standards for the creation of a CCD-chip with sufficient resolution have been high. The possibility of using the technology for creating an aerial camera has emerged, even with a better resolution than the scanned film.
Aside from being able to detect the visual light spectrum, the digital aerial camera can also detect the near-infrared spectrum. Many spaceborne sensors, like the *Landsat Thematic Mapper* developed in the 1970's, Mikhail et al. [2001], were build as multispectral sensors detecting light in the infra-red spectrum, so the technology was at hand.

1.1.2 Change detection

Several different algorithms can be used for the detection of changes in a location. The most common method is to compare two sets of image data from the same location, but acquired at different times. This method requires that all data is observed from the same location and covering the same area and that it only involves one type of image data.
Instead of comparing two images, this project will focus on comparing an existing digital map database with a new data set, read image. The advantage of this method is that we only need newly acquired data. The existing map database is already at hand and contains all the areas of interest, objects, and other data that needs updating. This thesis will only deal with objects classified as buildings.
The map updating process can be divided into the following steps:

- Acquisition of new data.

Figure 1.1: Example of an outdated map database.

- Data classification concerning buildings.
- Detection of buildings.
- Comparison with buildings in a digital map database to detect changes.

Figure 1.1 shows an example of an outdated digital map database where two types of changes occur. A demolished building that has been replaced by a new one and a new building in the image that is not registered in the map database. A building that has changed shape in some way, either because it has been demolished and replaced by a new one or because it has been extended, will in both cases be approached as a new building.

1.1.3 Digital topographical maps

The traditional use of analogue aerial photos for map updating has always been very time consuming. Changes are detected manually by a photogrammetric operator, using stereoscopy on two overlapping aerial photos to measure and vectorize objects of interest. Along with this, field inspections are carried out to verify the changes. As more changes occur every year an automation of this process will have substantial economical benefit as much time can be saved. In this project I have chosen to focus on objects such as buildings as they are often changed, and therefore are very important for updating of maps.

1.2 Motivation

Previous work by Olsen [2004], Knudsen and Nielsen [2004], and Knudsen and Olsen [2002] have shown that it is important to include height data in methods for detection of buildings. Since buildings can have spectral similarities with roads and other asphalted areas, see figure 1.1, these areas can easily be misclassified as buildings in a data classification. In this case the height data is useful to discriminate between buildings and terrain. In the work by Jensen [2002] only digital surface models derived from laserscanned data were used to detect changes. This work involved the use of mathematical morphology and was followed up in the Ph.D. thesis by Olsen [2004] where aerial image classification were combined with digital surface models applied with an elevation threshold to extract buildings.
The combination of the image data and the height have shown to be essential for a successful detection of buildings.

1.3 Project aim

The title of this project describes two types of data processing methods. These methods will be briefly described below in the context of this project.

1.3.1 Data fusion

The word data fusion covers the method of fusing two different types of data together. In this project it means the fusion of the elevation data with the image data.
As mentioned, it is important to involve the height data in building detection to eliminate misclassified pixels in the terrain. The data fusion will primarily be done by the creation of a 5 layered image containing the visible bands of red, green, blue, the near-infrared, and a rasterized and normalized height model. Next, an unsupervised classification of this 5 layer image will be performed. Hopefully, the involvement of the height data will result in a more successful classification of the data, so that buildings can be easily detected.
A second method will be presented based on an unsupervised classification of a 4 layer image of red, green, blue, and near-infrared bands. The height data will be included by masking out low elevated pixels in the a rasterized and normalized surface model.
A third method will involve a supervised classification using the predefined

training sample derived from the unsupervised classification of the 5 and 4 layer image, applied with a mask defined by the existing map database.

1.3.2 Semi-automatic map updating

Since it is difficult to detect the precise shape of a building in an image in an automatic way, map updating can only be called semi-automatic. As mentioned much work time is used on manual map updating, especially for locating the changes. Semi-automated therefore means that if the changes can be found automatically the photogrammetric operator knows where to look for changes and much work time is saved.

1.3.3 Data

All image and height data used in this project have been provided by COWI A/S. The orthoimages are a part of the Digital Orthophoto of Denmark DDO developed by COWI A/S. All image and height data are acquired in year 2005. The map database has been provided by KMS. The data covers the town of Vejle and consists of the following:

- High resolution analogously acquired orthoimage scanned with a GSD of 0.1 m, in ECW-format.
- High resolution digitally acquired orthoimage with a GSD of 0.1 m, in ECW-format.
- Digitally acquired orthoimage mosaic with a GSD of 1.0 m, in TIFF-format.
- Digital Surface Model (DSM) as ASCII text file with a grid spacing of 1.0 m.
- Digital Terrain Model (DTM) as ASCII text file with a grid spacing of 1.0 m.
- Map database TOP10DK version 2006.

All data are geocoded in UTM zone 32 ETRS89 except for the image mosaic that is geocoded in UTM zone 32 ED50. All height data are referenced to the Danish Vertical Reference 1990 (DVR90).

1.3.4 Quality assessment

A part of this project also focuses on the range of digital aerial cameras on the market, by comparing their technical specifications mutually and with a traditional analogue camera. Furthermore, a quality comparison between a digital aerial camera, page 18, and a traditional analogue aerial camera will be carried out. This comparison will be done visually, by computing the image histograms, and by two parameters describing the information content and image noise.

1.3.5 Change detection

Figure 1.1 shows change that detection consists of both negative and positive change in the sense of demolished buildings and new-build buildings respectively. All data used in this project is the latest version from 2006. Therefore the amount of differences between the image data, height data and the digital map database are limited. Hence buildings are added to or removed from the existing map database to verify the detection of demolished buildings or new buildings.

1.4 Software

Three different software programs are used and include,

MapInfo is a powerful software for Geographic Information Systems (GIS) where different types of data are handled in different layers so they are easy to access. It can handle raster data such as images and vector data such as map databases in multiple layers. In this project it is used for selecting height and map data for further processing.

MATLAB is a highly advanced programming language with many built-in functions for implementation and development of mathematical algorithms and equations, see Mathworks [2007]. It is used for implementation of the quality assessment and the supervised classification algorithm.

ENVI/IDL is specifically designed for image analysis of remotely sensed data. ENVI has a graphic user interface where all data, functions, and algorithms are easily accessible. IDL is a highly advanced programming language in which all the image analysis algorithms are implemented. This makes it easy to change standard algorithms and to implement new algorithms. By applying some standard ENVI functions into IDL, it is possible

to access such algorithms from the ENVI user interface, see ITTVIS [2007]. In this project the unsupervised classification is done in ENVI/IDL along with the creation of the multilayered images and the rasterization of the height data.

1.5 Thesis outline

Chapter 1: Introduction. Introducing the research area, the motivation, aim and outline of this project.
Chapter 2: Data Sources. This chapter concerns the technology of data sources and how they are created.
Chapter 3: Analogue versus Digital. Chapter 3 concerns quality and information content computation of scanned aerial analogue film and digital aerial images.
Chapter 4: Classification Theory. This chapter describes the basic statistic theory of data classification and introcuces the supervised and unsupervised classification algorithms used in this project.
Chapter 5: Methods for change detection. Three different methods for change detection based on results using the unsupervised classification algorithm are presented.
Chapter 6: Evaluation of results. Evaluation of the data classifications and the change detection methods.
Chapter 7: Discussion. Data sources, data types, methods, and results are discussed along with their advantages and disadvantages. Possible sources of errors and suggestions for improvement of the change detection methods are presented.
Chapter 8: Conclusion. Conclusion of the project in general and future perspective.

1.6 Summary

The subject of this project has been introduced leading to description of related work and the motivation for this project. The project aim has been described along with the project outline.

Chapter 2

Data sources

To achieve a sufficient maintenance of a digital map database, a comprehensive data set needs to be available. As previous work has documented, see 1.2, the use of only one data type leads to an inadequate change detection. This chapter will describe the different data types and sources, and briefly describe how they are created.
In general, change detection is based on three types of data characteristics, modified from Olsen [2004]:

- Spectral characteristics is vital to distinguish between different objects. This can done at different resolutions according to the required detection accuracy. However the spectral characteristics can vary depending on the type of data acquisition.

- Spatial characteristics are important since the value of one observation or pixel is rarely independent of a neighbouring observation.

- Geometric characteristics are the desired shape of an object and are important when fusing different data types.

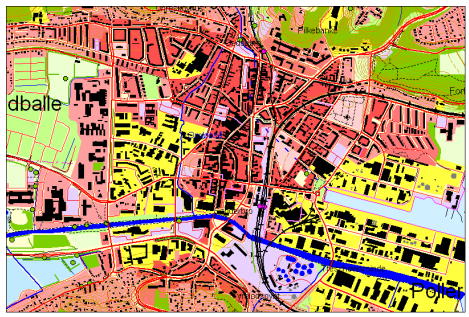


Figure 2.1: Digital map database covering Vejle.

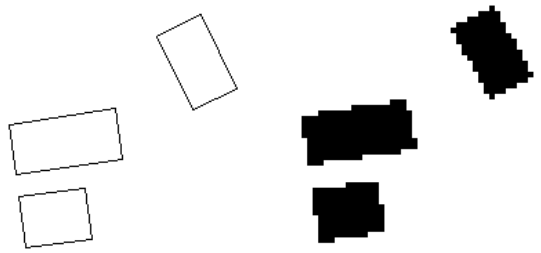
2.1 TOP10DK

The newest version of the topographical map from KMS exists in both a paper format, Kort10, and in digital format as a GIS database, TOP10DK.

Figure 2.1 shows TOP10DK covering the town of Vejle. The town consists of different areas describing the land use characteristics. Generally, densely populated areas surrounding many buildings are classified into four different areas described below, KMS [2001].

- *Lav bebyggelse.* Low buildings area mainly consists of single-family houses, wall-to-wall houses, villas and low apartment blocks. The houses contains 1 or 2 floors.
- *Høj bebyggelse.* High buildings are houses with more than 2 floors like blocks and often contains apartments, schools, service industries and other institutions.
- *Bykerne.* City centre is the most dense populated area and contains mostly the same as high building except that it is situated centrally.
- *Industri.* Areas classified as industry contains factories, shopping centres and industrial harbour areas.

The choice of which areas to acquire raw image data from, seems to be given from the listed description above. But as mentioned in the introduction, section 1.2, the use of the infra-red band does not seem to have much effect in urban areas like the city centre and industry areas where the amount of vegetation is low, see section 2.2.2 for further explanation. In suburban areas where many low



(a) (b)

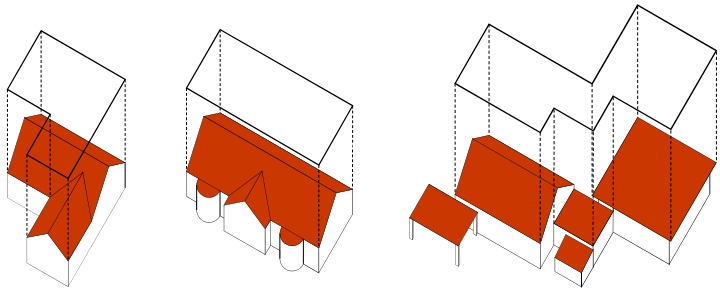
Figure 2.2: Example of building subjects on vector basis (a) and raster basis (b).

buildings and some high buildings are situated, there is much more vegetation and hence a greater contrast between buildings and surrounding areas.
The reason for not considering low-populated rural areas is that the there is a long distance between buildings, and therefore changes are expected to occur more seldom. It will therefore not be an economical advantage automatizing the digital map maintenance there. Only surface objects classified as buildings in TOP10DK will be considered in this project.

When opening TOP10DK in e.g. MapInfo all data is vectorized, appearing as line segments and polygons, Balstrøm et al. [2006]. Vector data is scale-invariant which means that zooming in on a building will not reduce the resolution. Contrarily to vector data, raster data is scale-variant and is defined by pixels having a certain resolution. Figure 2.2 shows an example of vector data and raster data of three buildings. The conversion between the two types of data can be done in ENVI.
Buildings as objects in a digital map database is registered by the eave of the building, figure 2.3a, and are shown by the ground shape. If buildings are integrated, as figure 2.3b shows, they will eventually be registered as one building. TOP10DK is georeferenced in UTM zone 32 as datum and ETRS89 as map projection.

2.2 Image data

Traditionally, aerial images have only been known in their physical form as film. In the 1960's KMS started to overfly Denmark with aerial film cameras for photogrammetric purposes. The aerial camera is mounted in an aeroplane so it is



(a) (b)

Figure 2.3: Principle of registration of building polygons (a) and registration of integrated buildings (b).

orthogonal to the ground as near as possible. The aeroplane then flies straight level in straight lines. While flying, a GPS unit logs the position of the camera every time an image is acquired. Most commonly, the photos have a mutually overlap along-track at 60 %, and 20 % across-track, Jacobi [1997]. This makes it possible to get a stereoscopic view of the images in a stereo-comparator or stereo-plotter.
Since it became possible to scan analogue films and obtain a satisfactory pixel size, this photogrammetric process has been digitized on a photogrammetric workstation, Mikhail et al. [2001].

2.2.1 Spectral properties

Electromagnetic waves from the sun are reflected from the surface of the earth to a different extend depending on the wavelength. Some intervals in the electromagnetic spectrum can hardly be transmitted through the atmosphere, and some transmits through the atmosphere with almost intact intensity especially in the visible light region and some infra-red regions, see figure 2.4. The spectral properties of different objects on the ground vary due to differences in the degree of reflectance between electromagnetic waves. Most remote sensing units are specialised to detect electromagnetic waves in the visible (VIS) and near-infrared (NIR) regions.

In Denmark, where the colours of the landscape changes significantly during the year, the choice of when to acquire aerial images is very important. It is

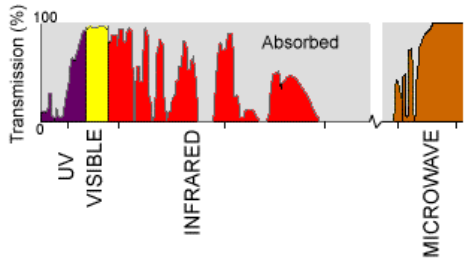


Figure 2.4: Atmospheric transmission of the electromagnetic spectrum.

Wavelength	layer
0.4 μm to 0.5 μm	blue
0.5 μm to 0.6 μm	green
0.6 μm to 0.7 μm	red

Table 2.1: Wavelengths for RGB colour image.

mainly of interest to find man-made objects such as buildings and since tall vegetation easily can cover such objects, it seems best to acquire aerial images during wintertime. However in the wintertime the angle of the sun is very low which results in objects casting long shadows that might cover the objects of interest, Olsen [2004]. By choosing springtime, just before foliation and only acquiring images when the angle of sun is highest, a satisfactory aerial image can be obtained.

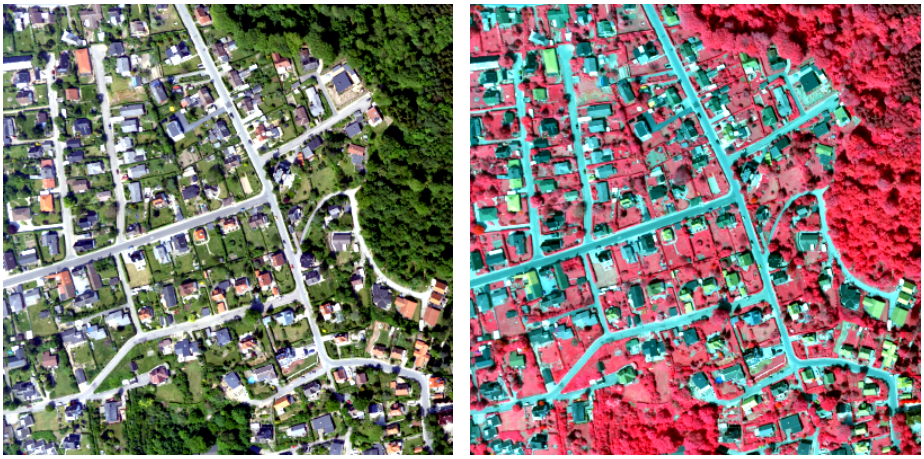
2.2.2 Visible and infra-red light

The visible light is often detected in three main bands, red, green and blue, see table 2.1. A full colour image RGB, see figure 2.5a, is then generated by assigning the three bands to three image layers so that all colours are created by a sufficient amount from each layer.

Besides having a high reflectance in the green band of the visible light, vegetation has an even higher reflectance in NIR region. Since the NIR region has a large span of wavelengths it is divided into several bands. Dependent on the aerial camera or airborne sensor the wavelength of the NIR band can vary. Since the NIR is not visible, a colour-infrared (CIR) image can be created by assigning the NIR light to the red layer, the green light to the blue layer and the red light to the green layer, see table 2.2 and figure 2.5b.

Wavelength	layer
0.5 μm to 0.6 μm	blue
0.6 μm to 0.7 μm	green
0.7 μm to 0.9 μm	red

Table 2.2: Wavelengths for CIR colour image.



(a) (b)

Figure 2.5: Example of RGB image (a) and CIR image (b).

Normalized Difference Vegetation Index The availability of the NIR bands makes it possible to calculate the Normalized Difference Vegetation Index NDVI defined by equation (2.1).

$$\text{NDVI} = \frac{\text{Infrared} - \text{Red}}{\text{Infrared} + \text{Red}} \tag{2.1}$$

The NDVI value ranges from -1.0 to 1.0 where high values indicates green vegetation and low values are equivalent to non-vegetated areas like bare earth or man-made objects such as roads, houses etc. Since the observed infra-red reflection is very dependent on the sensor used, NDVI is also sensor-dependent. Non-vegetated areas range in the region from -1.0 to 0.1, see Olsen [2004]. Since this project focuses on the change detection of buildings it is very useful to be able to eliminate vegetation from image data.

2.2.2.1 Spectral properties of buildings

Since this project concerns the detection of buildings in aerial imagery it concerns the spectral properties of the roofs of buildings. Thus, it concerns the discrimination of the roofs from the surrounding terrain and objects. Figure 1.1 shows that roofs can look very different depending on the roof material. In Denmark there is a long tradition of using different colours of tiles for roofing, see figure 2.5a. The red and black tile seems to be very common along with some other colours like gray and brown. Figure 2.5a shows that within each roof colour there can be a large variation.

2.2.3 The analogue camera

The traditional analogue frame camera is a film-based camera where each photo on the focal plane has a size of 23 $\times$ 23 cm. The focal length from the lens to the focal plane is 153 mm.
Digitization of each photo is performed using a high-resolution scanner which results in a pixel size of 14 μm, Balstrøm et al. [2006].

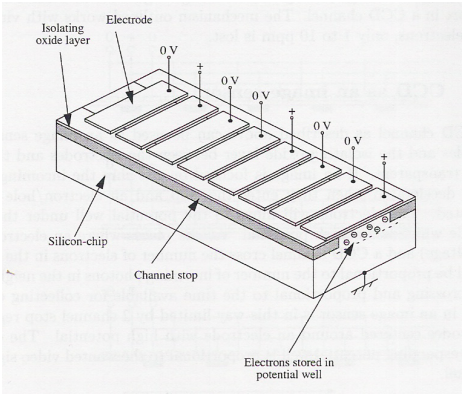


Figure 2.6: Principle of the CCD-chip, Carstensen [2002].

2.2.4 The digital camera

The basic principle of a digital camera is almost the same as an analogue camera. Here the focal plane is replaced by a silicon chip and thus the film reels become obselete as well as the procedure of digitization by scanning, Wolf and Dewitt [2000]. However digital remote sensing can work in other ways with linear CCD-arrays.

2.2.5 The CCD-chip

The most widely known technology used for image sensors is the Charge Coupled Device (CCD). The CCD-chip is a light sensitive silicon chip, consisting of an array of cells, where the light is transformed into a electric charge, see figure 2.6, Carstensen [2002]. The magnitude of the electric charge indicates the intensity of the incoming light. One cell is equivalent to one pixel, so the dimension of the chip determines the resolution of the image.

Generally speaking, there are three types of digital imaging devices used for airborne remote sensing, see Wolf and Dewitt [2000]. All of them are based on the technology of the CCD-chip:

Full-frame sensor is a two-dimensional CCD array that works in the same way as a traditional analogue camera where the sensor is mounted on the focal

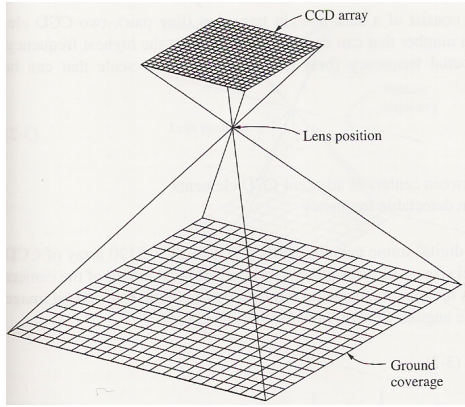
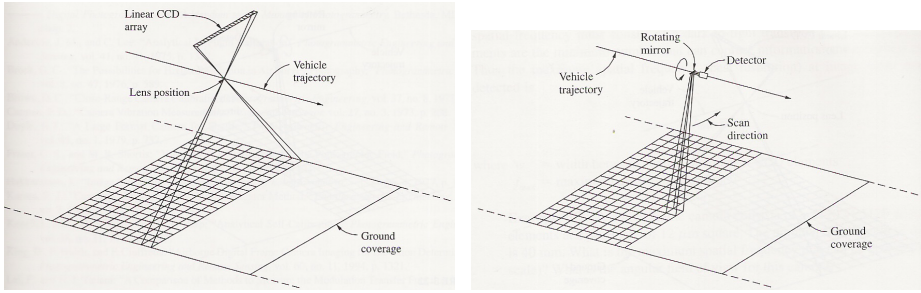


Figure 2.7: Geometry of digital frame camera, Wolf and Dewitt [2000]



(a) (b)

Figure 2.8: Geometry of a linear array sensor (a) and flying spot scanner (b), Wolf and Dewitt [2000]

plane in a single-lens camera. All CCD cells are exposed simultaneously when a image is acquired. The frame-based CCD sensor is mainly used in airborne digital cameras, but since the resolution of one CCD array is still very limited, a high resolution can be obtained by combining several CCD-array.

Linear array sensor is a linear one-dimensional CCD array that builds the image by sweeping the ground so the array is perpendicular to the flying route. This scanner is also known as a *Pushbroom* scanner and is also used in space-borne sensors.

Flying spot scanner works in almost the same way as the linear scanner except that only one cell is exposed at a time. A rotating mirror scans the ground from side to side creating a single row of the image. This principle is almost similar to those used in airborne laser scanners as described in section 2.3.1.

2.2.5.1 Panchromatic sharpening

Since it is expensive to have several high resolutional CCD arrays in one camera, the aerial camera manufacturers use an alternative technology called panchromatic sharpening.
The camera consists of only one high resolution CCD array combination used for the panchromatic band which is a wide electromagnetic band covering the visible light region. Then some CCD arrays with lower resolution are used for colour and infra-red spectrum. Now the spatial resolution of the colour images are refined using panchromatic sharpening, see Canty [2007*b*].

2.2.6 UltraCam D

The digital images used in this project are acquired with an UltraCam D which is a frame-based camera developed by the Austrian company Vexcel Corp, Vexcel [2005]. This camera consists of 13 CCD-arrays, where 9 are panchromatic and 4 are colour. A combination of the 9 panchromatic arrays results in a image of 11500 $\times$ 7500 pixels with a pixel size of 9 μm, see figure 2.9, Vexcel [2004]. The colour arrays have a reduced size and therefore result in an image of only 4008 $\times$ 2672 pixels, but with the same pixel size. By the use of panchromatic sharpening, see section 2.2.5.1 and Leberl et al. [2003], it is possible to make a colour image with the same size as the panchromatic image.
The camera records a radiometric resolution of 12 bit equivalent to 4096 colours, but the final aerial images are often quantified in to 256 colours (8 bit) as the difference can hardly be seen by the human eye. The aerial images used here are 8 bit data.
The camera detects the four multispectral bands, red, green, blue and near-infrared that results in a four layered image RGBNIR.

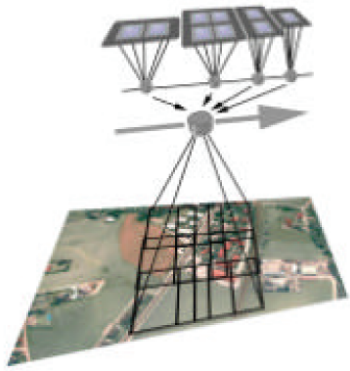


Figure 2.9: Photographic and storage unit of the UltraCam D.

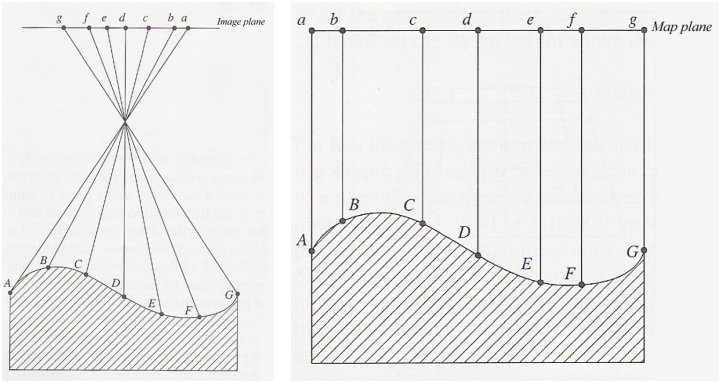
2.2.7 Other digital airborne sensors

The list below briefly describes other types of digital cameras or airborne sensors currently on the market. They all have the property of detecting the near-infrared spectrum at different wavelengths. For technical specifications of the sensor below see table 2.3.

UltraCam X Vexcel Corp. has developed an improved version of the UltraCam D called UltraCam X. The overall structure of the UltraCam X is the same as UltraCam D. The main difference is the size of the CCD-arrays and the image resolution, Vexcel [2006], see table 2.3

DMC The Digital Mapping Camera (DMC) from Intergraph is a frame-based camera similar to the UltraCam except that each frame is merged together in a different way. 4 high-resolution panchromatic camera heads and 4 low-resolution multispectral camera heads results in an output image of 13824 $\times$ 7680 pixels. Pancromatic sharpening is used here to make high-resolution multispectral images. The multispectral bands consists of red, green, blue, and NIR, where it is possible to change the desired spectrum of the NIR.

ADS 40 Is a line scanner where there are different linear CCD arrays for each of the spectral bands. The CCD colour and NIR arrays have the same resolution as the panchromatic CCD array and therefore no panchromatic sharpening is needed, Geosystems [2002].



(a) (b)

Figure 2.10: Central perspective (a) and orthographic projection (b).

JAS 150 Jena has developed a pushbroom scanner and claims that the scanner creates a near-orthogonal view of the terrain due to a stereoscopic view, GmbH [2006].

The technical specifications of the analogue camera and the digital cameras and sensors are summarized in table 2.3. The GSD and the nominal swath width are calculated for a flying altitude of 4000 m.

2.2.8 Orthoimage generation

An aerial image that comes straight from the camera is not geometrical correct. The difference between a digital map and an aerial image is that the map is an orthogonal projection of the ground objects whereas an aerial image is a central perspective projection, Mikhail et al. [2001]. Figure 2.10 shows that the distance between a set of terrain points changes their spacing in the image plane due to the variation of the terrain elevation above the reference plane (datum plane). Figure 2.11 shows the principle of this change in spacing known as the relief displacement where the perspective plane is the image plane.

In digital orthoimages the relief displacement is eliminated using a digital terrain model (DTM) of the area to reproject displaced pixels. However generation of an orthoimage by a DTM only reprojects terrain pixels so the ground plane of a building fits a digital map. The disadvantage of the orthoimage is that buildings are still in central projection. Therefore the roofs of buildings are not

	Scanned film	UltraCam D	UltraCam X	DMC	ADS40	JAS150
Camera type	Frame	Frame	Frame	Frame	Pushbroom	Pushbroom
Focal length [mm]	153	100	100	120		150
Pixel size [μm]	14	9	7.2	12	6.5	6.5
GSD [cm]	37	36	29	40	42	17
Nominal swath width [m]	6013	4140	4156	5200	4992	2066
Image size/width [mm]	230 × 230	103.5 × 67.5	103.9 × 67.8	165.9 × 92.2		78.0
Colour resolution	Full	Half	Half	Half	Full	Full
Spectral bands	3	4	4	4	4	4
Radiometric resolution [bit]	8	12	12	12	12(8)	12
Storage capacity [TB]	-	1.5	1.7	1.7	1.0 per hour	1.6

Table 2.3: A comparison of the 6 different cameras at an flying altitude of 4000 m

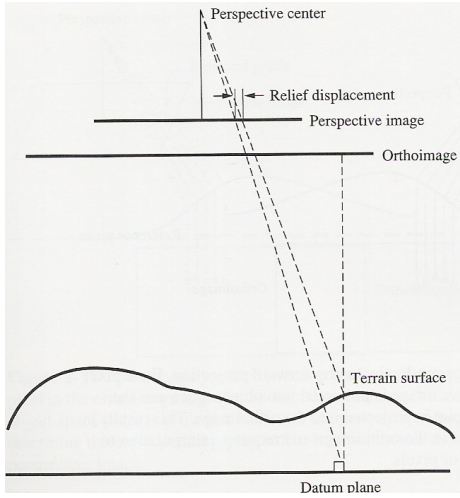


Figure 2.11: Perspective and orthographic image geometry showing the effect of relief displacement.

orthogonally correct and are consequently exposed to the relief displacement.

True orthoimage Since normal orthoimages only considers the correction of the displacement in the terrain, it can be useful to also correct objects above terrain. This is known as true orthoimages. The relief displacement of the roofs of buildings are eliminated so all buildings are in a orthographic projection and the buildings fit perfectly in a digital map.

2.2.9 Image mosaics

The image data used for the change detection is derived from an image mosaic created by georeferencing several orthoimages. Especially since the digital aerial images are of a smaller size, see table 2.3, it has become useful to georeference several orthoimages to cover a greater landscape in one image. The downside of this is that several orthoimages, even though they are acquired at the same time, can be very different spectrally because of different illumination conditions at different times as well as different locations of acquisition. This spectral difference can have a significant influence on the results of supervised and unsupervised classifications.

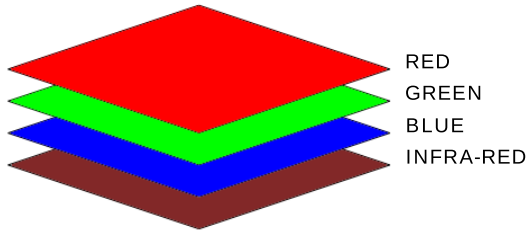


Figure 2.12: Principle of a 4 layered image RGBNIR.

Dependent on the aerial camera manufacturer or image software developer, the method for the elimination of such spectral differences can vary. The image mosaic used in this project for change detection is a 4 layer image of red, green, blue, and near-infrared (RGBNIR) see figure 2.12.

2.3 Elevation data

Traditionally, the elevation data has been created manually by a stereo operator drawing elevation curves, see Mikhail et al. [2001]. The introduction of airborne laser scanning has accelerated the process of generating elevation data.

2.3.1 Airborne laser scanning

Until recent years aerial images have been used for the production of elevation data using stereoscopy, see section 2.2.8. This method is time consuming as it is done manually by a stereo operator. However in the 1990's an Airborne Laser Scanning (ALS) for airborne collection of elevation data was introduced. Laser scanning had at that time only been known to be used in total stations for land surveying to measure distances, but now the necessary technology was available to make it airborne. ALS is also known as LIDAR (LIght Detection And Ranging) but is a more general term for laser scanning that is not necessarily airborne.
An ALS unit is an airborne active sensor that sends light pulses with the length of about 10 nanoseconds and measures the distances to the ground by recording the time for a pulse to return from the terrain surface, see figure 2.13. While scanning, a GPS unit logs the position of the aircraft. However this can not be done fast enough for every laser pulse so an Inertial Navigation System (INS)

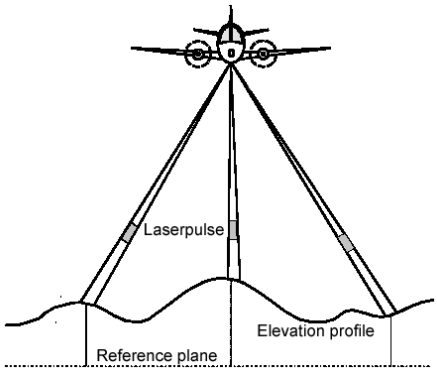


Figure 2.13: Principle of measurement by laser scanning, TopoSys [2007].

unit logs the rotation and pitch of the aircraft at every laser pulse, Balstrøm et al. [2006].

2.3.1.1 Multiple echoes

A light pulse from an ALS is infinitely narrow. It has a beam divergence that leaves a circular footprint on the ground. Dependent on the flying altitude, this footprint can have various sizes. Therefore, one laser pulse can interact at different stages through the atmosphere resulting in many echoes. The first echo comes from high objects like tree tops and roofs and the last echo comes from the ground surface, see figure 2.14. The echoes in between the first and last echo are rarely recorded.
Along with recording of the time delay of the echoes, the intensity of the received laser pulse is also recorded. The intensity can say something about the surface that the laser pulse is echoed from. A very smooth surface will reflect most of the incoming laser while on a more rough surface some of the transmitted laser will be scattered in various directions and thus be harder to detect, Elachi [1987]. Therefore this scattering mechanism reduces the intensity of the received laser pulse.

2.3.2 The TopoSys Falcon II

The laserscanner used to create the elevation data in this project is a Falcon II from TopoSys, see TopoSys [2007]. The Falcon II is a so-called glass fibre scanner emitting a laser pulse through a glass fibre, see figure 2.15a.

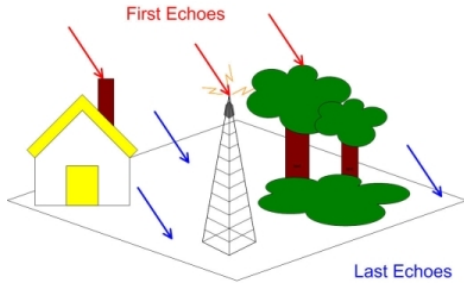
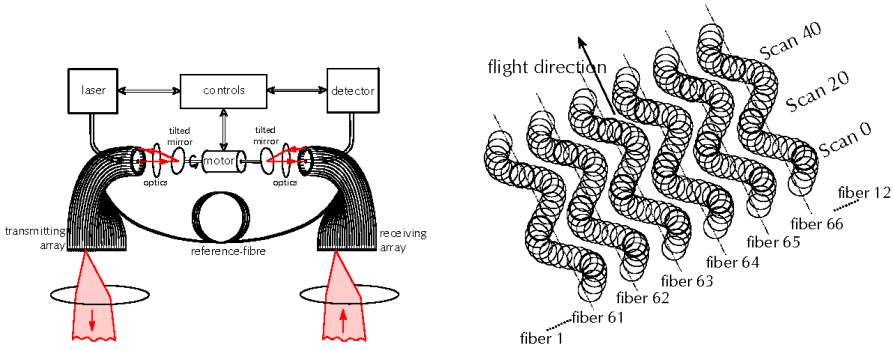


Figure 2.14: Principle of the first and last echo of a laserscanning pulse.



(a) (b)

Figure 2.15: Principle of operation of a fiber scanner (a) and resulting scan pattern (b), Katzenbeisser [2004].

Parameter	Value
Sensor type	Pulse laser
Maximum level of operation	1600 m
Scan width	14.3°
Scan rate	653 Hz
Laser pulse rate	83,000 Hz
Distance resolution	1.95 cm
Footprint size at 1000 m	0.5 m

Table 2.4: Technical specifications for TopoSys Falcon II.

From the laser generator the pulse is passed through optics and reflected by a rotating and tilted mirror into one of the glass fibres in a circular shape of glass fibres. This circular shape is transformed into a linear shape, thereby creating a linear transmitting array. The reception of the echoed laser pulses works the same way as the emission, but in the opposite direction. The resulting scan pattern can be seen in figure 2.15b. The echoed laser pulse is focused to the corresponding fibre at the receiving side and guided through the circular array and finally reflected into the center fibre leading to the detector. A central motor that controls both of tilted mirrors ensures that they are synchronized. The laser and the detector are calibrated once per revolution of the circular array through a reference-fibre. This structure ensures a stable transmission and reception of the laser pulse.
The beam divergence of the TopoSys Falcon II is 0.5 mrad which results in a footprint size of 0.5 m in diameter when the flying altitude is 1000 m.

The technical parameters of this laserscanner is summarised in table 2.4.

2.3.3 Digital representation

Digital height data are used to represent a surface determined in a three-dimensional co-ordinate system. The X and Y co-ordinate determines the point location in the reference plane and the Z co-ordinate represents the elevation. There are two main types of elevation models for both representation of the surface of a landscape:

DTM Digital Terrain Model also known as Digital Elevation Model (DEM) describes the terrain without objects such as buildings and trees.

DSM Digital Surface Model describes the surface of the landscape including buildings and trees etc.

Figure 2.16: Principle of the nDSM in profile.

These elevation models can be created in a regular grid or a triangulated irregular network (TIN), Wolf and Dewitt [2000]. A regular grid of points can then easily be rasterized so that each pixel contains the elevation of one point. The regular grid is created by linear interpolation between each point.
The elevation data, DTM and DSM, in this project are delivered as a regular grid of points in a ASCII text file, containing the X, Y, and Z co-ordinate. The first echo detected by the laserscanner described above is used to create the DSM and the last echo from the ground is used to create the DTM.

Since this project only focuses on building detection i.e. objects above the terrain the DSM can be normalized in relation to the DTM by computing the normalised digital surface model (nDSM) defined as,

$$\text{nDSM} = \text{DSM} - \text{DTM}$$

Figure 2.16 shows the principle of the nDSM that only contains the elevations from buildings and trees since the terrain in a nDSM is in reference plane at an elevation of 0 m.

2.4 Summary

Three data types have been presented, the digital map database, image data and height data. The following data will be used in this project:

TOP10DK Polygons characterised as buildings.

Image data All four multispectral bands, red, green, blue and NIR.

Elevation data DSM, DTM and nDSM.

Chapter 3

Analogue versus digital

In this chapter, a comparison will be performed of the high resolutional analogue and digital images with a GSD of 0.1 m as described in section 1.3.3. This will be carried out using visible comparison, image histograms and two parameters called entropy and SNR. This should give an indication about the quality and information content of the image data.

3.1 Entropy

Entropy is a well-known term in thermodynamics and is defined as a continuously growing parameter indicating that the disorder of the universe is increasing i.e. all differences are erased. In the same way entropy can be used in image analysis to estimate the level of information in an image, see Carstensen [2002]. Due to noise the information in an image might as well be coincidental and hence subject to uncertainty.
Since it can not be avoided that some noise is acquired in aerial images, the entropy seems to be best parameter to estimate the level of information. The entropy indicates how much data can be compressed, i.e. how many pixels there can be in 1 bit. If all pixels are very alike the entropy is low indicating that the amount of noise is low. A large entropy expresses that the pixels are very

different.
The entropy H is computed by equation (3.1),

$$H = \sum_{i=1}^{G} p_i \cdot \log_2(p_i) \quad bits/pixel \tag{3.1}$$

where p_i is the frequency of a colour in pixel i and G is the total number of colours in the image. In this case $G = 256$, see section 2.2.6.

3.2 Signal-to-Noise Ratio

Signal-to-Noise-Ratio or just SNR is a parameter estimating the relation between the clean signal S and the noise N of an observation as,

$$\text{SNR} = \frac{S}{N} \tag{3.2}$$

Since the clean signal can not be separated from the noise, equation (3.2) can not be applied. And since our test data are covering small areas with very limited variation in colours, we can use the coefficient of variation c_v,Johnson [2000], to define SNR,

$$c_v = \frac{\sigma}{\mu}$$

where μ and σ is the mean and standard deviation of the image. The coefficient of variation indicates how much data variates in relation to the position of the mean. Then equation (3.2) can be written as

$$\text{SNR} = \frac{1}{c_v} = \frac{\mu}{\sigma} \tag{3.3}$$

So a large SNR value indicates a small amount of noise and vice versa.

(a) (b)

Figure 3.1: Analogue images (a) and digital images (b) of a cross roads intersection and a shadowed garden and building.

3.3 Visible comparison

The visual comparison is simply done on two different areas, where the difference seems to be most conspicuous. Figure 3.1 shows two images from an analogue camera (left) and a digital camera (right). Figure 3.1a shows a cross roads intersection where the difference between the analogue and digital image is very clear. The white lines are much clearer and well-defined in the digital image. The pavement also seems to have a more homogeneous colour.
One of the most important issues in aerial imagery is the ability to see objects or relevant information that are covered by shadow. Figure 3.1b shows a building creating a shadow in which some information is hardly visible in the analogue part but is visible and much clearer in the digital part.

3.4 Histograms and parameters

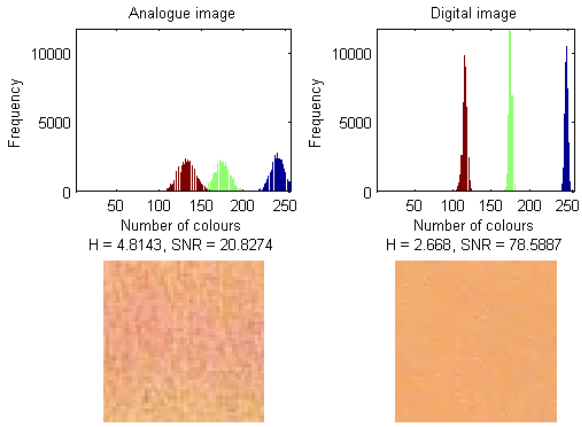
To compare imagery on the basis of histograms, entropy, and SNR, the compared images need to cover nearly the same area and be spectrally homogeneous. Since this project also concerns the location of buildings based on image classification the chosen image patches explains the two main roofing materials, see section 2.2.1. Figure 3.2 shows the comparison of a tile roof and asphalt roof as analogue

and digital image patches.

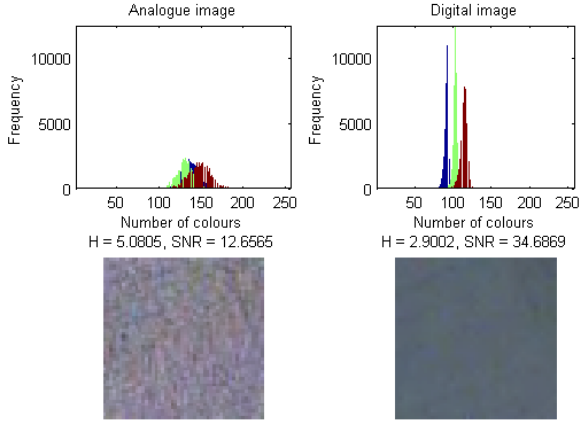
The simple comparison in figure 3.2 shows a great difference between the analogue and digital patch. The analogue image histograms have a much more flat, bell-shaped curve which is explained by a larger deviation of the image data, while the digital image histograms are more narrow explained by a smaller deviation.
The entropy H and SNR confirms the histograms. The entropy is much smaller for the digital image explaining a larger information content. The SNR is larger for the digital images because of the small standard deviation compared to the mean.
This simple comparison shows that there are some significant differencies between analogue and digital imagery which explains some important issues concerning the choice of aerial camera.

3.5 Summary

An image quality assessment has been presented as a comparison of a scanned analogue aerial image and a digital aerial image. A visual comparison of two different areas shows a greater image dynamics in the digital image. Histograms, entropy, and SNR have been computed for small image patches with homogeneous spectral properties and show a significant difference between analogue and digital images. The information content in the digital image is much larger together with a substantial lower level of noise. The higher diversity of the pixel values in the analogue image is most certainly caused by noise.



(a)



(b)

Figure 3.2: Analogue and digital image patch comparison of tile (a) and asphalt (b) roofing material with adjacent histograms, entropy H and SNR.

Chapter 4

Classification Theory

Explaining certain characteristics of a large data set such as an image can be difficult when no prior knowledge is available and the diversity is great. Nonetheless this can be done solely by the use of multivariate statistic analysis to classify the data into a number of classes, where each class describes data that have some similarities.
The unsupervised classification method is called the *Fuzzy Maximum Likelihood Estimation* (FMLE) and is based on the theory of the supervised classification the *Maximum Likelihood Estimation* (MLE). This chapter will first describe the theory of the supervised classification followed by the unsupervised classification theory.
The notation used is equivalent to the one used in Canty [2007*b*].

4.1 Data classification

Classification of image data is performed on the basis of the pixel characteristics. Dependent on the number of classes, pixels that have some differences might be classified as equivalent. If the number classes is small the diversity might be large in each class, and if the number of classes is large only pixels with high similarity belongs to the same class.

In supervised classification the classes are predefined as a training sample either manually by pointing out areas in the image or using unsupervised classification results.
In the unsupervised classification the classes are not given beforehand. Here the classes are determined by the spectral and spatial characteristics of the image The algorithms used here are based on multivariate statistical analysis, by assigning a membership probability to each observation explaining the probability of one observation belonging to a certain class.

4.1.1 Bayes' theorem

To assign a probability to each observations one needs to know the basic assumptions behind maximum likelihood classification. These assumptions are merely based on the theory of conditional probability and Bayes' Theorem, Johnson [2000] and Canty [2007*b*].
The conditional probability $Pr(A_k|B)$ of an event A_k occurring given that the event B occurs is given by Bayes' Theorem, equation (4.1).

$$Pr(A_k|B) = \frac{Pr(B|A_k)Pr(A_k)}{\sum_{i=1}^{m} Pr(B|A_i)Pr(A_i)} \tag{4.1}$$

The event A_k is a subset of a range of events $A_1, A_2 \ldots A_m$ for which the probability of A_i $Pr(A_i) \neq 0$ for $i = 1 \ldots m$ and $Pr(B) \neq 0$.
In the context of data classification, the event B denotes our observations (pixels) and the event A_k is a given class.
Then if we let $\boldsymbol{g}$ be our observation and $\{k|k = 1 \ldots K\}$ be a set of possible classes then equation (4.1) can be written as, Canty [2007*b*],

$$\Pr(k|\boldsymbol{g}) = \frac{\Pr(\boldsymbol{g}|k)\Pr(k)}{\Pr(\boldsymbol{g})} \tag{4.2}$$

where,

- $\Pr(k|\boldsymbol{g})$ is the posterior conditional probability for class k given the observation $\boldsymbol{g}$.
- $\Pr(k)$ is the prior probability of class k.

- $\Pr(\boldsymbol{g}|k)$ is the conditional probability of observing the value $\boldsymbol{g}$ if it belongs to the class k.
- $\Pr(\boldsymbol{g}) = \sum_{k=1}^{K} \Pr(\boldsymbol{g}|k)\Pr(k)$ is the total probability for $\boldsymbol{g}$.

4.2 Supervised classification

When we have predefined our classes we can use them to classify an image by applying the *Maximum Likelihood Estimation* (MLE).

4.2.1 Maximum likelihood classification

In MLE the posterior probability is computed for all classes and each observation is assigned to the class that has the largest posterior probability $\Pr(k|\boldsymbol{g})$. Assuming that the observations from class k are sampled from a multivariate normal distribution $N(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ our density function is given by equation (4.3),

$$p(\boldsymbol{g}|k) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Sigma}_k|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)\right) \tag{4.3}$$

and taking the logarithm of equation (4.3) equals,

$$\log(p(\boldsymbol{g}|k)) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}_k| - \frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{\mu}_k) \tag{4.4}$$

Since the first term is independent of k it can be dropped. From equation (4.4), equation (4.2) can be written as,

$$\log\Pr(k|\boldsymbol{g}) = \log(\Pr(k)) - \frac{1}{2}\log|\boldsymbol{\Sigma}_k| - \frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{\mu}_k) \tag{4.5}$$

Since $\Pr(\boldsymbol{g})$ is independent of k it has been omitted.

Equation (4.5) computes an posterior probability image for each of the classes, and the pixels with the largest posterior probability are assigned to the respective class.
To describe the relation between the posterior probabilities for each class the entropy can be computed by rewriting equation (3.1) as,

$$H = \sum_{k=1}^{K} \Pr(k|\boldsymbol{g}) \cdot \log_2(\Pr(k|\boldsymbol{g})) \tag{4.6}$$

A large entropy in this context indicates that for a given pixel the posterior probabilities $\Pr(k|\boldsymbol{g})$ are very different whereas a low entropy reflects that the posterior probabilities $\Pr(k|\boldsymbol{g})$ are very alike.
The last term in equation (4.5) is the squared Mahalanobis distance defined as,

$$d = \sqrt{(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{g}_i - \boldsymbol{\mu}_k)} \qquad \in \chi^2(p) \tag{4.7}$$

where p is the number variables e.g. the number of image layers.
Applying a classification threshold on the MLE can either be done using a minimum probability or a maximum Mahalanobis distance. This will be described further in section 5.5.3

4.2.2 Implementation in MATLAB

The supervised classification algorithm MLE has been implemented in MATLAB, see appendix C.2. The implementation concerns a MATLAB function `MLEfct.m` and a main script for execution `MLEscript.m`.

`MLEscript.m` When executing the user is asked to input the number of layers in the image, the Maximum Mahalanobis distance threshold, the image subjected to classification, and the training sample image. The training sample image has to be a binary image where each class is defined by a number from 1 to the number of classes. Unclassified areas are equal to 0. Finally the user can specify whether or not to use a prior probability image. If *no* all pixels have equal priors. If *yes* the user specifies the image defining the priors. The prior probability image also has to be a binary image where each areas with high probability have pixel values 1 and for surrounding areas the value 0.

`MLEfct.m` Uses the image, training sample, image size parameters, prior probability and distance threshold as input.
The output contains a posterior probability image for each class, the final classification and the covariance matrices of the classes.

4.3 Unsupervised classification

In unsupervised classification or clustering, classes are often denoted as clusters, since the word cluster better describes that the classification is only based on the observations.
Now if we then denote our observations $\mathcal{G}$ by the following, see Canty [2007*b*],

$$\mathcal{G} = \{\boldsymbol{g}_i | i = 1 \dots n\},$$

where $\boldsymbol{g}_i$ is one observation at a given index i and n is the number of observations. The observations are to be partitioned into a number clusters C. This can be expressed by the following,

$$C = [C_1, \dots C_k, \dots C_K],$$

where C_k is denotes the k'th cluster in a maximum number of K clusters. Then equation (4.2) can be written as,

$$Pr(C|\mathcal{G}) = \frac{p(\mathcal{G}|C)Pr(C)}{p(\mathcal{G})}, \tag{4.8}$$

where $p(\mathcal{G})$ is a normalization independent of C and $p(\mathcal{G}|C)$ is the *likelihood* of cluster C given the observations $\mathcal{G}$. The aim is to assign the observation $\boldsymbol{g}_i$ to class k so equation (4.8) is maximized. Hence the posterior probability $p(\mathcal{G}|C)$ is maximized to obtain the most optimal classification.
To describe the posterior probability of one observation $\boldsymbol{g}_i$ that belongs to a cluster k, the density function of the conditional probability $p(\mathcal{G}|C)$ is determined as the multivariate normal density function $p(\boldsymbol{g}_i|k)$ assuming that all observations are normally distributed, Canty [2007*b*].
Now $p(\mathcal{G}|C)$ equals the product of the probability density for all clusters,

$$p(\mathcal{G}|C) = \prod_{k=1}^{K} \prod_{i \in C_k} p(\boldsymbol{g}_i|k)$$
$$= \prod_{k=1}^{K} \prod_{i \in C_k} \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Sigma}_k|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)\right)$$

By taking the logarithm of the *likelihood* results in the *log-likelihood*,

$$\log p(\mathcal{G}|C) = \sum_{k=1}^{K} \sum_{i \in C_k} \left(-\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1}(\boldsymbol{g}_i - \boldsymbol{\mu}_k)\right) \tag{4.9}$$

and equation (4.8) can then be written as,

$$\log Pr(C|\mathcal{G}) = \log p(\mathcal{G}|C) + \log Pr(C) - \log p(\mathcal{G} \tag{4.10}$$

The last term is independent of C and by eliminating it the maximization of $\log Pr(C|\mathcal{G})$ is equivalent to maximizing $\log p(\mathcal{G}|C) + \log Pr(C)$.
To determine the membership of one observation to a given cluster, a membership probability u_{ki} is introduced. First a *hard* classification is defined as follows,

$$u_{ki} = \begin{cases} 1 & \text{if } i \in C_k \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

which means that one observation belongs to only one cluster, and satisfies the following conditions,

$$\sum_{k=1}^{K} u_{ki} = 1,\ i = 1 \ldots n, \qquad \sum_{i=1}^{n} u_{ki} > 0,\ k = 1 \ldots K. \tag{4.12}$$

The mean $\boldsymbol{m}_k$ and covariance $\boldsymbol{s}_k$ of the kth cluster is equivalent to,

$$\boldsymbol{m}_k = \frac{1}{n_k} \sum_{i \in C_k} \boldsymbol{g}_i = \frac{\sum_{i=1}^{n} u_{ki} \boldsymbol{g}_i}{\sum_{i=1}^{n} u_{ki}}, \; k = 1 \dots K \tag{4.13}$$

$$\boldsymbol{s}_k = \frac{\sum_{i=1}^{n} u_{ki} (\boldsymbol{g}_i - \boldsymbol{m}_k)(\boldsymbol{g}_i - \boldsymbol{m}_k)^T}{\sum_{i=1}^{n} u_{ki}}, \; k = 1 \dots K \tag{4.14}$$

When observations are classified into certain clusters there is always a change of misclassification. The cost of misclassification can be explained by the cost function $E(C)$, Canty [2007*b*].

$$E(C) = \sum_{k=1}^{K} \sum_{i=1}^{n} u_{ki} \frac{\|\boldsymbol{g}_i - \boldsymbol{m}_k\|^2}{2\sigma^2} - \log Pr(C) \tag{4.15}$$

4.3.1 Fuzzy K-means clustering

The method of using a *hard* membership probability as described above is known to be used in the K-means clustering algorithm (KM) also known as *ISODATA*, Duda and Hart [1973] or *migrating means*, Richards and Jia [1999].
In fuzzy K-means clustering the class memberships in equation (4.11) are replaced by continuous variables and clustering becomes *fuzzy*,

$$0 \leq u_{ki} \leq 1, \quad k = 1 \dots K, \; i = 1 \dots n \tag{4.16}$$

still satisfying the conditions in equation (4.12).
Then the membership probability u_{ki} is found by equation (4.17) so it minimizes equation (4.15).

$$u_{ki} = \frac{\left(\frac{1}{\|\boldsymbol{g}_i - \boldsymbol{m}_k\|^2} \right)^{1/(q-1)}}{\sum_{k'=1}^{K} \left(\frac{1}{\|\boldsymbol{g}_i - \boldsymbol{m}_{k'}\|^2} \right)^{1/(q-1)}}, \; k = 1 \dots K, i = 1 \dots n \tag{4.17}$$

The FKM algorithm iterates between equation (4.13) and (4.17) until the change of u_{ki} is significantly low, 10^{-3}.

4.3.2 Fuzzy maximum likelihood estimation

Fuzzy maximum likelihood estimation, FMLE, is an extension to the Fuzzy K-means clustering (FKM) algorithm giving it more flexibility and variability in the extent of the clusters, see Gath and Geva [1989].
In the context of the FMLE, Bayes' Theorem can be written as,

$$u_{ki} \rightarrow Pr(k|\boldsymbol{g}_i) = \frac{p(\boldsymbol{g}_i)\Pr(k)}{p(\boldsymbol{g}_i)} \tag{4.18}$$

Where $p(\boldsymbol{g}_i)$ is a normal distributed density function. The u_{ki} is the membership probability for pixel i to cluster k, and is equal to the numerator in Bayes' Theorem. The membership probability is then computed by equation (4.19).

$$u_{ki} = p(\boldsymbol{g}_i|k)\Pr(k) = \frac{1}{\sqrt{|\boldsymbol{s}_k|}} \exp\left[-\frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{m}_k)^T \boldsymbol{s}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{m}_k)\right] \frac{n_k}{n} \tag{4.19}$$

where n_k is the number of pixels in the k'th cluster. The prior probability $\Pr(k) = p_k$, equation (4.3.2), is known from the extended K-means clustering where entropy, Carstensen [2002], is used for determining the prior probability, Palubinskas [1998].

$$p_k = \frac{n_k}{n} = \frac{1}{n}\sum_{i=1}^{n} u_{ki}$$

The mean $\boldsymbol{m}_k$ and covariance $\boldsymbol{s}_k$ of the kth cluster are estimated using equation (4.13) and (4.14).

This leads to an unstable computation that is very sensitive to the initialization conditions because of the inversion of the covariance matrix $\boldsymbol{s}_k$ that can easily be singular, Eising [1999]. To avoid this the initial membership probabilities are determined by using the FKM algorithm, Canty [2007*b*], that uses a randomly generated membership probability matrix as input. To ensure that these random probabilities are the same at every classification a so-called *seed* parameter is set to a arbitrary value.

Then the FMLE algorithm can be summarised into an iteration process of 4 steps, see Canty [2007*b*].

1. First the starting values for the initial membership μ_{ki} are determined using equation (4.17).

2. Next the cluster centers $\boldsymbol{m}_k$, equation (4.13), the covariance matrices $\boldsymbol{s}_k$, equation (4.14), and the priors $\Pr(k) = n_k/n$, equation (4.3.2), are determined.

3. The new class membership probabilities μ_{ki}, equation (4.19), are determined and the columns are normalised according to equation (4.12).

4. If $\boldsymbol{u}$ has changed significantly, go to step 2, otherwise stop the process.

This iterative process, where the cost function equation (4.15) is minimized, can easily be trapped in a local optimum. Consequently, the FMLE uses simulated annealing, Canty [2007*b*], so the early membership probabilities have a high degree of random dependency and gradually the membership probabilities are allowed to influence the calculation of the mean and covariance matrices.

4.3.3 Optimal number of clusters

To determine the optimal number of clusters the data will first be clustered repeatedly into a different number of clusters starting with $K = 2$ and ending at an arbitrarily maximum number of clusters, e.g. $K = 15$. At each clustering the FMLE will determine the fuzzy hypervolume FHV as,

$$F_{HV} = \sum_{k=1}^{K} \sqrt{|\boldsymbol{s}_k} \tag{4.20}$$

and the partition density D_{PA} as

$$D_{PA} = \frac{m}{F_{HV}} \tag{4.21}$$

Where m is the membership probabilities of all observations which lie within unit Mahalanobis distance of a cluster center.

$$m = \sum_{i \in \mathcal{N}} \sum_{k=1}^{K} u_{ik},\ \mathcal{N} = \{i | i = 1 \ldots n, (\boldsymbol{g}_i - \boldsymbol{m}_k)^T s_k^{-1} (\boldsymbol{g}_i - \boldsymbol{m}_k) <\}$$

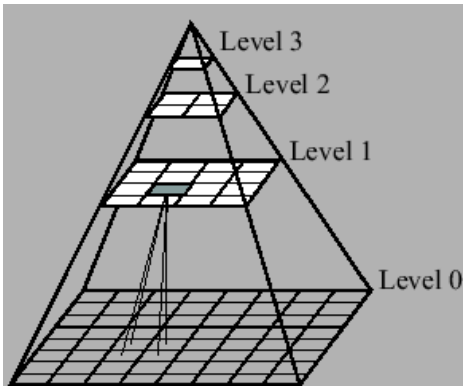


Figure 4.1: Principle of level numbering in an image pyramid.

A plot of the fuzzy hypervolume or partition density as a function of the number clusters, will show a minimum for F_{HV} and a maximum for D_{PA} for the optimal number of clusters, Gath and Geva [1989].

4.3.4 Multiresolution clustering

The FMLE is a quite slow algorithm because of the iterative processing. A method to speed it up is to perform the clustering at a coarser resolution of the image.
First a image pyramid is created, see figure 4.1 using a *discrete wavelet transform* (DWT), see Canty [2007*b*].

The original image is equivalent to pyramid depth 0 or scaling factor 1. The class membership probability determined at the coarsest resolution set by the user, and passed on to the next level as the initial membership probability, Hilger [2001] and Canty [2007*b*]. In the implementation by Canty [2007*a*] it is not possible to classify the original image at pyramid depth 0. The classification can be done at the pyramid depths of 1, 2 and 3.

4.3.5 Spatial properties

So far the FMLE has only considered the spectral information. One pixel is often dependent on the neighbouring pixels so it seems reasonable to involve the spatial information as well. This can be done using Gibbs-Markow random

fields to assign labels to each pixel in a regular lattice $\mathcal{I}$, Li [2001] and Canty [2007*b*].

$$\mathcal{I} = \{i | 1 \leq i \leq n\}$$

where n is the number of pixels and i denotes the pixel in focus. The relation between the cells of the lattice can be defined by a neighbourhood system $\mathcal{N}$,

$$\mathcal{N} = \{\mathcal{N}_i | i \in \mathcal{I}\}$$

The set of class labels can then be described as,

$$\mathcal{K} = \{k | i | 1 \leq k \leq K\}$$

where K is the number of possible classes in the image. The class label itself can be described by a discrete random variable,

$$L = \{L_1 \dots L_n\}$$

Now describing the realization of the labelling L by the configuration l,

$$l = \{l_1 \dots l_n\}, \quad l_i \in \mathcal{K}$$

The neighbourhood system can consist of either 4 or 8 neighbouring pixels. The FMLE uses a 4-neighbourhood system also known as the *first order neighbourhood system*.

Figure 4.2 shows a 4-neighbourhood with the pixel i in focus and figure the 3 possible cliques that explains the different configurations l of the neighbourhood system $\mathcal{N}$.

The probability of a given configuration l is computed by the Gibbs distribution,

$$p(l) = \frac{1}{Z} \exp\left(-\beta \sum_{c \in \mathcal{C}} V_c(l)\right) \tag{4.22}$$

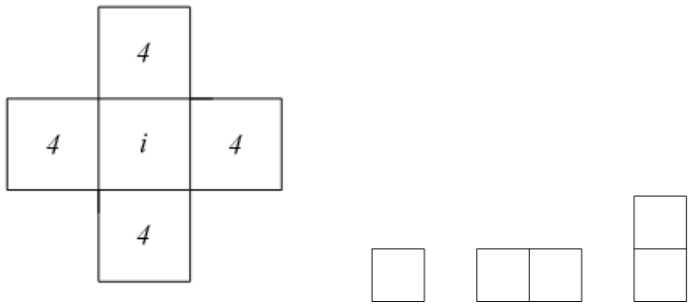


Figure 4.2: 4 pixel neighbourhood (a) and the possible cliques (b).

where Z is a normalisation factor and β is a constant. $V_c(l)$ is the clique potential for clique c and summarised in equation (4.22) for all cliques. A low clique potential indicates a more probable configuration whereas a high cliques potential describes a less probable configuration. A low β results in equal probability for all configurations and can be adjusted independently of the clique potential. In other words β adjusts the degree of spatial dependency in the final classification.
Using a 4-neighbourhood results in the following spatial membership probability for the FMLE, equation (4.23):

$$v_{ki} = A \exp(-\beta(1 - u_{k\mathcal{N}})) \tag{4.23}$$

where A is a normalisation factor which ensures that $\sum_k v_{ki} = 1$. For a more detailed description of how the v_{ki} is derived, see Canty [2007*b*].
The combined spectral membership probability u_{ki} and the spatial membership probability v_{ki} is then determined by,

$$\frac{1}{\sqrt{|\boldsymbol{s}_k|}} \cdot \frac{n_k}{n} \cdot \exp\left(-\frac{1}{2}(\boldsymbol{g}_i - \boldsymbol{m}_k)^T \boldsymbol{s}_k^{-1}(\boldsymbol{g}_i - \boldsymbol{m}_k) - \beta(1 - u_{k\mathcal{N}_i})\right)$$

where $u_{k\mathcal{N}}$ is determined for the entire image by applying the kernel,

$$\begin{matrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{matrix}$$

with the spectral membership probability image u_{ki}.

4.4 Implementation in ENVI/IDL

The FMLE algorithm has been implemented in ENVI using the IDL programming language, Canty [2007*b*] and Canty [2007*a*]. The algorithm can be added to the unsupervised classification category in ENVI. The classification result can be used like the other default algorithms to perform any kind of supervised classification and/or postprocessing steps.
Originally the implemented algorithm contains two programs, `EM.PRO` and `EM_RUN.PRO`, see Canty [2007*a*].

`EM.PRO` Takes an observation array and a initial class membership probability array u_{ki} as input and performs the iterations described in section 4.3.2. It then returns a new class membership probability array u_{ki} and repeats the iteration until $\Delta u_{ki} < 10^{-3}$.

`EM_RUN.PRO` Runs the FMLE algorithms as a GUI where all inputs are specified. The input consists of, first the multispectral image, secondly the number of clusters and thirdly the incoming and final scale factor (pyramid depth). Next, the initial annealing temperature and the spatial membership parameter β are specified. Finally the output filenames of the classified image and the membership probability image are saved.

4.4.1 Extending the algorithm

The implementation of FMLE in ENVI/IDL by Canty [2007*b*] and Canty [2007*a*], does not directly determine the optimal number of clusters. The user has to specify it without knowing the partition density or fuzzy hypervolume. These two parameters can be dragged out of the `EM.PRO` and plotted against the adjacent number of clusters.
The implemented `EM_RUN.PRO` is then changed, named `FMLE.PRO`, so the maximum number of clusters is specified instead. The part of the FMLE algorithm

where the class membership probabilities are initialised and iteratively determined, as described above, are embraced by a loop that starts at 2 clusters ($K = 2$) and stops when the K reaches the maximum number of clusters K_{max} specified by the user.
For all number of clusters $K = 2 \dots K_{max}$ the partition density and fuzzy hypervolume are stored in one array for each of the 2 parameters. The subscript in those arrays plus 2 is then equivalent to the adjacent number of clusters. The partition density is plotted against the number clusters.
Finally the FMLE is processed again using the optimal number of clusters. Both the files are contained on the enclosed CD-ROM.

4.5 Summary

The theory of the supervised classification algorithm MLE and the unsupervised classification algorithm FMLE has been presented. The MLE has been implemented in MATLAB and the FMLE is derived from Canty [2007*a*] and changed for optimal performance in this project.

Chapter 5

Methods for change detection

To detect a change between two data sets, two sets of data, acquired at different times, are needed. As mentioned in chapter 1 section 1.1.2, this project will focus on the use of newly acquired data and compare it with an existing map database, in terms of building registrations.
First this chapter describes preprocessing where test areas are chosen. Next three methods for change detection are presented.

5.1 Preprocessing

Preprocessing concerns the extraction of test areas according to the decision made in section 2.1. Firstly, two test areas are chosen from the 4 layer image mosaic covering Vejle. Figure 5.1 shows the RGB images of the two test areas overlaid the existing map database TOP10DK. The town center of Vejle is situated in the lower right corner of figure 5.1.

Next, these images are loaded into MapInfo along with the ASCII text files of the DSM and DTM as a 1×1 m grid of points according to the description in section 1.3.3. With the images overlaid with the DSM and DTM respectively, the height data that only covers the images are picked and saved as individual

Figure 5.1: The location of the two test areas in the suburban areas of Vejle.

ASCII files.
These text files are then read into ENVI using the application *Rasterize Point Data*. ENVI approaches the elevation data as being irregularly sampled and therefore uses TIN to triangulate the points. Next the points are linearly interpolated to a regular grid with a GSD of 1 m. The intensity of each pixel reflects the elevation in each point. ENVI offers another method for interpolation based on quintic polynomials which results in a more smooth surface. However linear approach was chosen, because a too smooth surface will result in closely situated buildings being highly interrelated.

To summarise the preprocessing steps are,

- Two RGBNIR images of 400 × 400 pixels (400 m × 400 m) are selected from the image mosaic.
- DSM and DTM data covering the test areas are selected.
- The DSM and DTM are rasterized using TIN and linear interpolation with a GSD of 1 m, see appendix A.

The rasterization of the elevation data is performed so that the elevations (Z co-ordinates) in each pixel are kept. Then the rasterized DSM and DTM can be used for computation of the nDSM.

The nDSM is computed in ENVI by the function `ndsm.pro`, see appendix B.1. It can be accessed in the ENVI user interface by the *band math* function in the *basic tools* menu. Possible negative elevation is equalized with 0 m to avoid spectral influence from negative elevations.

5.2 5 layered image creation

As introduced in section 1.3.1 we will try to incorporate the height data in the image data as another image layer.
Since the FMLE is based on a normal distribution, all image data has to approach a normal distribution. The histograms in appendix figure A.3 shows that the image data are normally distributed. However, figure 5.2 shows that the histograms for the nDSM's have a tendency to approach a log-normal distribution. To normalise the nDSM data, the logarithm is taken to the nDSM by the ENVI function `logndsm.pro`, see appendix B.2. It can be accessed in ENVI in the same way as `ndsm.pro`.

The pixel values in nDSM is height data so before taking the logarithm to nDSM, all pixel values are converted to byte so the pixel values ranges from 0 to 255 like the image data. The resulting log(nDSM) turns out to be very dark so the it is linearly stretched along the whole gray scale spectrum from 0 to 255, see figure A.8.
The 5 layer images are then created using the *layer stacking* option in ENVI so the 5 layers are co-registered using the correct map projection and datum, see section 2.2. Figure 5.3 shows the principle of the 5 layered image. The co-registration might influence the size of the image, and thus it is resized to the original size of 400 $\times$ 400 pixels, i.e. 400 m $\times$ 400 m.

5.3 Method 1

Method 1 is merely based on the results from the unsupervised classfication algorithm FMLE of the 5 layer image described above.
The result of the clustering is an image of a number of colours equivalent to the number of clusters. By comparing the classification with the original image the two most promising clusters, that corresponds to a building, are extracted and merged thereby creating a binary image with one class of probable buildings.
As described in section 2.1 the minimum building size registered in the digital map database is 25 m^2. Since the resulting cluster pixel size 1 m^2 we can eliminate clusters smaller than 25 pixels by using the segmentation function.

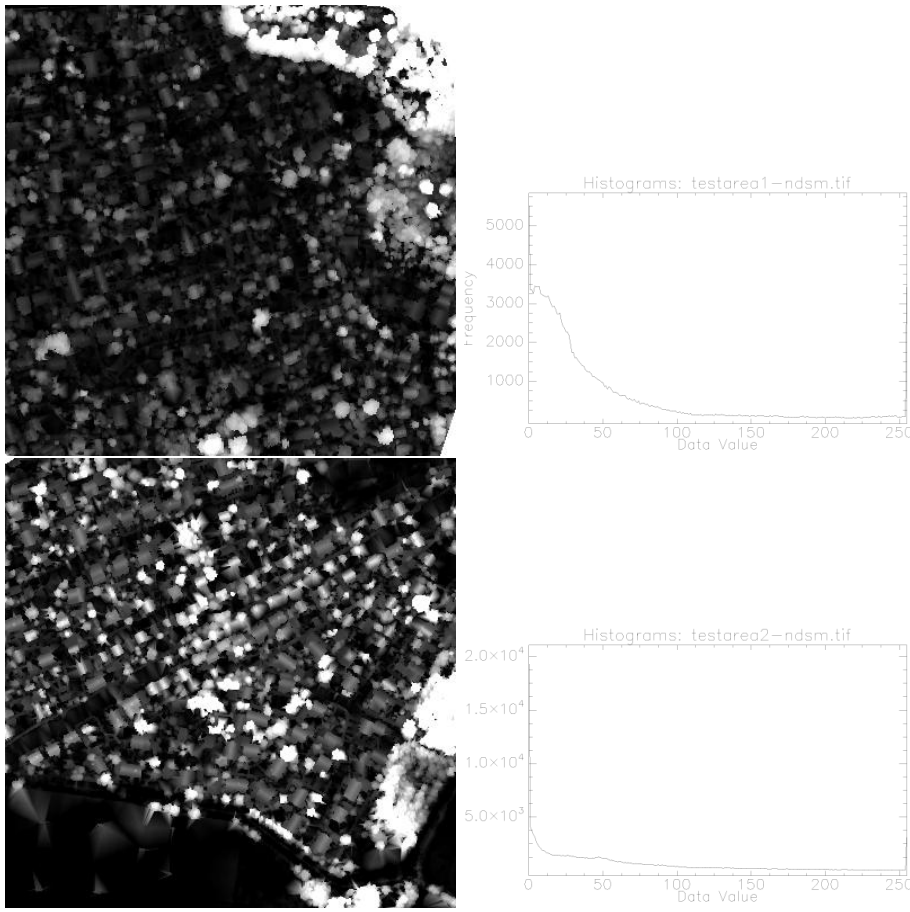


Figure 5.2: Rasterized nDSM of the two test areas and associated histograms.

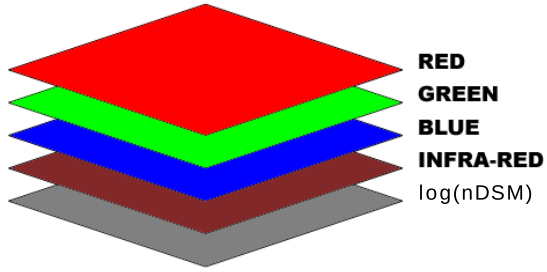


Figure 5.3: Principle of the 5 layered image.

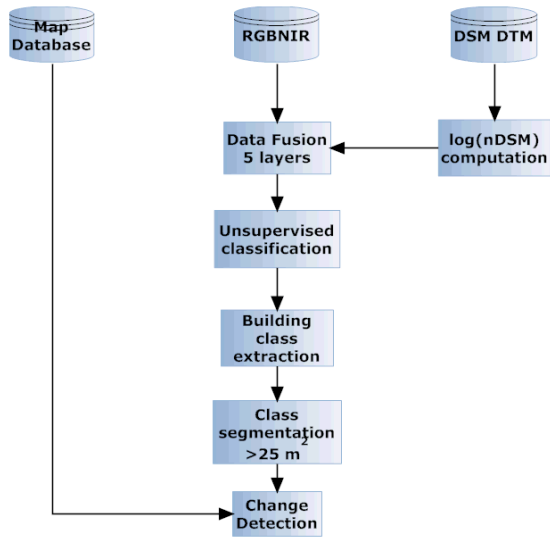


Figure 5.4: Work flow for Method 1

Now the clustered image is ready to be compared to the map database. Figure 5.4 shows the suggested work flow for method 1.

5.4 Method 2

In method 2, the raw image consists of only the 4 layered image from the mosaic. The 4 layer image is classified using FMLE at pyramid depth 2 and a spatial dependency $\beta = 1$, see the evaluation of the FMLE in section 6.1. The two most promising clusters are extracted and merged. Clusters smaller than 25 pixels or 25 m^2 are eliminated as described above.
Method 2 is then divided in to two submethods where the only difference between them is the use of the vegetation index NDVI image. Figure 5.5 shows the work flow for method 2 with a modification in method 2B encircled by a dotted line.

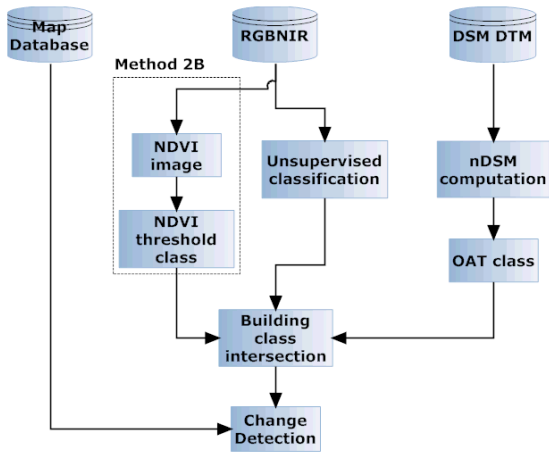


Figure 5.5: Work flow for method 2A and 2B.

5.4.1 Method 2A

Instead of integrating the height data in the unsupervised classification the nDSM is assigned a threshold value to create a classification result. As described in section 2.1, the minimum building height is $2.5m$ and by letting it act as a threshold value it is possible to create a class of height values above $2.5m$. The result is a so-called Objects Above Terrain (OAT) class where one class represents the trees and buildings containing height values above 2.5 m.
Now the intersection between the extracted building class from the FMLE and the nDSM threshhold class is computed and class regions ranging below 25 m^2 are eliminated. The final image is then ready to be compared with the existing map database.

5.4.2 Method 2B

A similar classification based on the NDVI image, see figure A.9, is applied using a maximum threshhold of 0.1 to eliminate all potential vegetation see section 2.2.2. Now the intersection between the OAT class from method 2A and the NDVI threshold class is computed thereby obtaining a new classification that shows all nDSM values above $2.5m$ without any vegetation.
Again the NDVI/nDSM class is intersected with the building cluster from the RGBNIR classified image and postprocessed as described in section 5.4.1

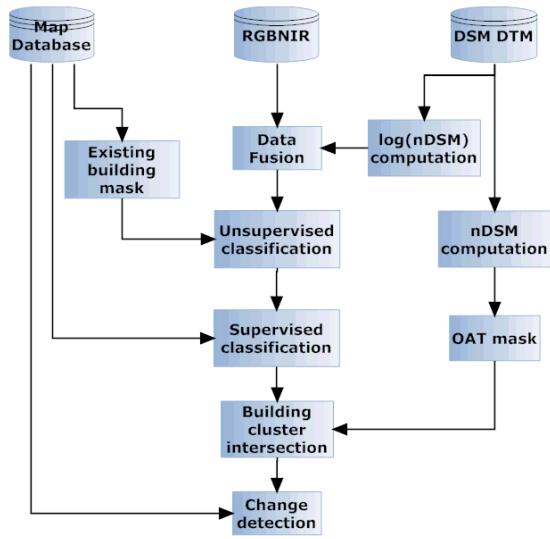


Figure 5.6: Work flow for method 3

5.5 Method 3

Method 3 is slightly more advanced than the two previous methods in the sense that it involves two data classifications and the use of a prior probability image. Method 3 contains the following steps.

- Creation of TOP10DK mask.
- Unsupervised classification (FMLE) with TOP10DK mask on 4 layer and 5 layer images.
- Extraction of training data from unsupervised classification (FMLE).
- Supervised classification (MLE) of 4 layer and 5 layer images using the training data from FMLE and possible involvement of prior probabilities.

5.5.1 TOP10DK mask

ENVI offers the possibility of masking out observations that we do not want to classify. A mask is created from an existing image that is specified by a certain

data range. To create a mask from our map database, TOP10DK, we import it to ENVI as a vector file and specify the data range from the image we want to classify, so the mask is covering our test area. In the resulting mask all buildings are *ON* and non-buildings are **OFF**, so training data from the FMLE is only derived where we know there are buildings.

5.5.2 A prior probability

We already know where there are buildings and therefore our only task is to find new buildings. This prior knowledge of the location of buildings can be used to define a prior probability image. Most often prior probability images contains the same probability in each pixel or as a randomly generated number, so all probabilities add up to 1.
Both in unsupervised and supervised classification algorithms the prior probability is used to accurate the classification.
In method 3 we will test whether or not a prior probability will have an effect on the supervised classification by applying the prior probability image described below.

Building $\Pr(k) = 0.9$

Non-building $\Pr(k) = 0.1$

As a comparison, a supervised classification will be done with equal probabilities for all observations.

5.5.3 Classification threshold

Since the Mahalanobis distance follows a χ^2 distribution, see section 4.2.1 we can apply a threshold on the MLE. We assume that most the classes have very short distances to their class centers since they are derived from very limited areas under the building mask. Thus we will use 95 % and 99 % of the χ^2 distribution. By use of a statistical table in Johnson [2000] we get table 5.1.

p	$\chi^2_{0.95}(p)$	$\chi^2_{0.99}(p)$
4	9.488	13.277
5	11.070	15.056

Table 5.1: Maximum Mahalanobis distance values.

5.6 Map updating

The classification results from the various methods are evaluated in four images.

5.6.1 Check of classification

The addition of the rasterized map database and the classification result will show an image where a successful detection will have one colour, and the lack of detection a second colour:

- Red - Lack of detection.
- Green - Succesful detection.

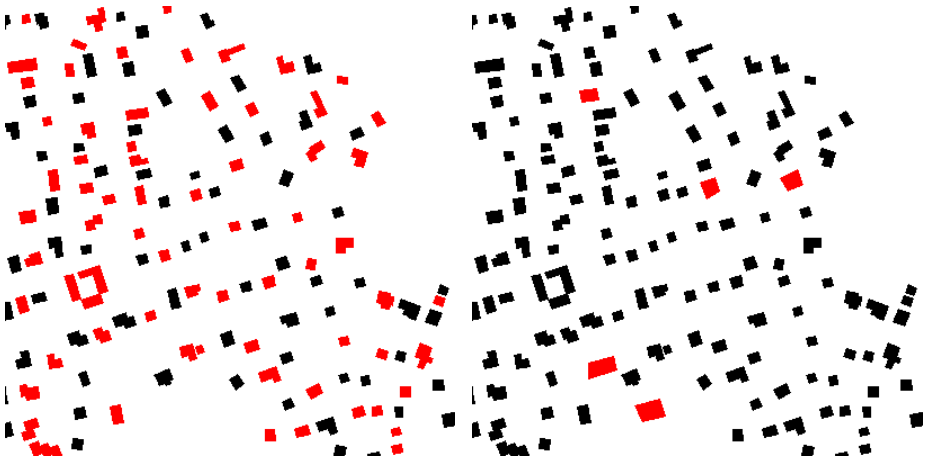
This will shows how well the change detection method detects existing buildings.

5.6.2 New buildings

To verify that the classification can find new buildings, 50 % of the buildings in the map database image are removed. The classification is then supposed to detect the removed buildings as new buildings. Figure 5.7a shows the removed buildings in red colour.
The factual and detected buildings are counted along with the amount of possible false alarms.

5.6.3 Demolished buildings

It also happens that buildings are demolished and to verify this possible buildings from the classifition result are removed. The buildings that are not detected



(a) (b)

Figure 5.7: Detection of new buildings (a) and demolished buildings (b).

by the methods are then assumed to be demolished. Figure 5.7b shows the added buildings in red.
The factual demolished buildings and non-detected buildings are counted.

5.6.4 Visual check

The building class from the different methods are vectorized in ENVI and overlaid the a RGB image of the test area to verify the change detection and to find what causes possible false alarms.

A statistical table summarizes the factual and detected new buildings and demolished buildings respectively. False alarms are counted as well.

5.7 Summary

The preprocessing and creation of the 5 layer image have been presented. Based on the available data and the description of the data sources in chapter 2 three main methods have been presented for semi-automatic map updating. Further-

more the evaluation techniques for the change detection methods have been described.

Chapter 6

Evaluation of change detection methods

This chapter concerns all unsupervised and supervised classification results and the threshold classification of the nDSM and NDVI images. The three different methods are evaluated with the original map database as described in section 5.6

6.1 Optimal settings for FMLE

As described in section 4.3.3 the optimal number of clusters can be determined by the maximum partition density. To find the right degree of influence from the spatial membership probability, three different values of β were used for a FMLE classification at the three pyramid depths of 1, 2 and 3. This test for the determination of the best composition was only performed for the 5 layered image. The resulting classification was performed in 9 different ways for each test area. The optimal number of clusters for both test areas are summarised in table 6.1. The classifications and plots of the partition density against the number of clusters can be seen in appendix D.

Based on the results in appendix D the clustering at the three different pyramid

	Test area 1			Test area 2		
	β					
Pyramid depth	0.0	0.5	1.0	0.0	0.5	1.0
1	4	4	4	6	5	5
2	6	6	**8**	7	6	**8**
3	6	5	6	6	8	6

Table 6.1: Optimal number of clusters for compressions and spatial membership β for test area 1.

depth can be summarised to the following description:

Pyramid depth 1 The number of clusters is limited by the instability of the FMLE algorithm. Classification result is not satisfactory.

Pyramid depth 2 The number of clusters seems to be adequate and so the classification result.

Pyramid depth 3 The number of clusters is not adequate and so is the classification.

The choice of the most optimal clustering composition falls on pyramid depth 2 with most spatial information taking into account with $\beta = 1$, see bolded number in table 6.1. The number of clusters is optimal and the buildings seems to belong to two clusters that separates them from surrounding areas. Figure 6.1 shows the most optimal classification of the two test areas and the plot of the partition density against the number of clusters.

6.2 Method 1

The best solution for the evaluation of the FMLE is used and the two most promising building clusters are extracted. The building clusters are verified with TOP10DK database as described in section 5.6.

6.2.1 Test area 1

The unsupervised classification of test area 1 with 8 clusters seems to give a good result in figure 6.2 UL. The buildings are generally a part of two clusters

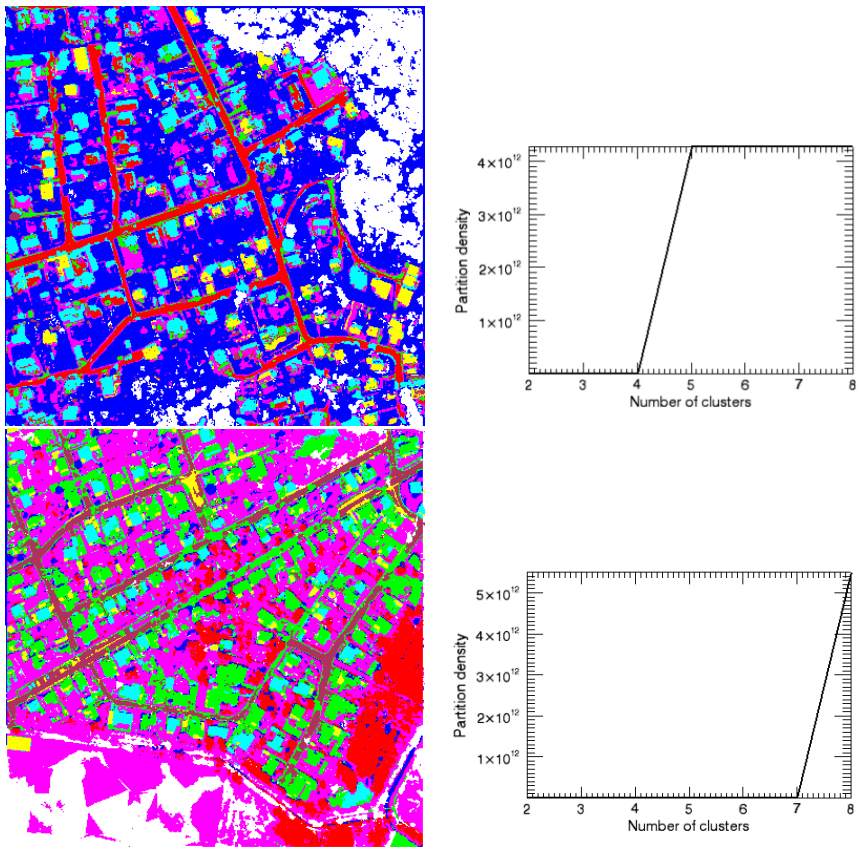


Figure 6.1: FMLE classification (left) and partition density plot (right). Upper: test area 1, lower: test area 2.

	Factual	Detected
New buildings	68	54
False alarms	40	
Demolished buildings	5	3

Table 6.2: Statistical results for: method 1, area 1

that are extracted and merged, and objects smaller than the required building size of 25 m^2 are eliminated. The classification check is fairly successful in figure 6.2 LR where most of the buildings are detected.

The supposed new buildings that are redetected are shown in figure 6.3 UL and the number is in table 6.2. It is clear that some buildings are not detected. The amount of false alarms also shows that some areas are misclassified due to spectral similarities with the buildings. Most of the false alarms can be small buildings, sheds or greenhouses that are not eliminated by the size threshold of 25 m^2 or it can be part of road segments. The buildings that are not detected seem to have a high spectral simularity with the road and hence are contained in the same class as the road..
As a coincidense two of the demolished buildings are detected, but if at least 30 % of a building have to be detected to be a new building, none of the demolished buildings are assumed to be detected.

6.2.2 Test area 2

The unsupervised classfication for test area 2 seems to work less successful than for test area 1, see figure 6.4. Again, the spectral and spatial properties of the buildings seem to join in to two clusters. The extraction and segmentation of the clusters work less well as seen for test area 1 but in the classification check in figure 6.4 all existing buildings seem to be detected.

Figure 6.5 shows that all new buildings are detected but the number of false alarms are numerous and less well-defined. Contrarily to test area 1, the false alarms can not be described as possible small greenhouses or sheds. Several road segments and bare earth areas are misclassified as a part of one of the two building clusters.

Figure 6.5 LC shows that, although some of the detected buildings are not classified as buildings according to the map database, they might still be buildings since it cannot be expected that the map database is fully updated with the image data.

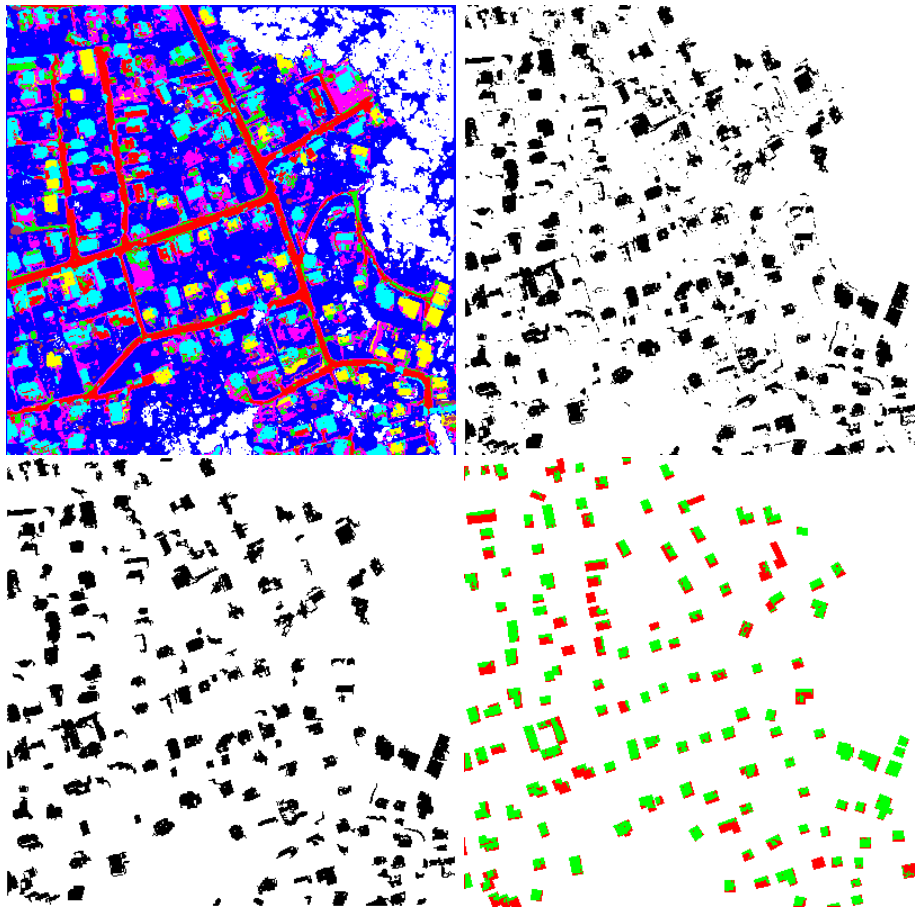


Figure 6.2: Test area 1, UL: FMLE classification of 5 layer image, UR: Extracted building clusters, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

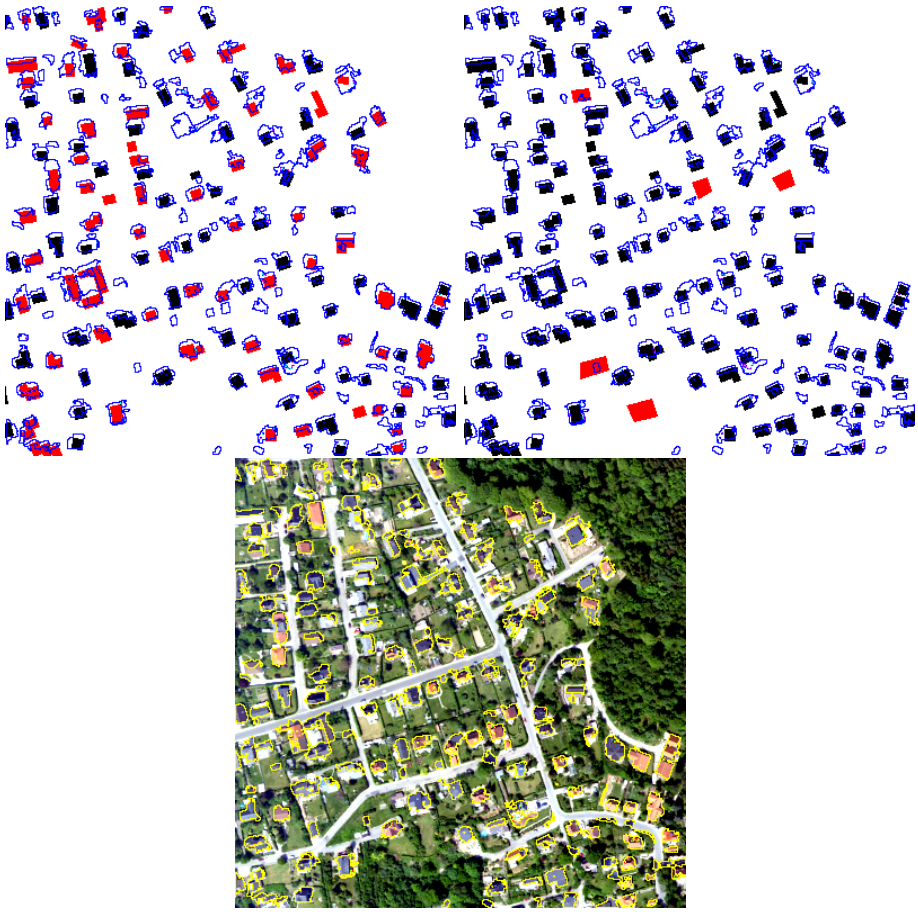


Figure 6.3: Test area 1, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

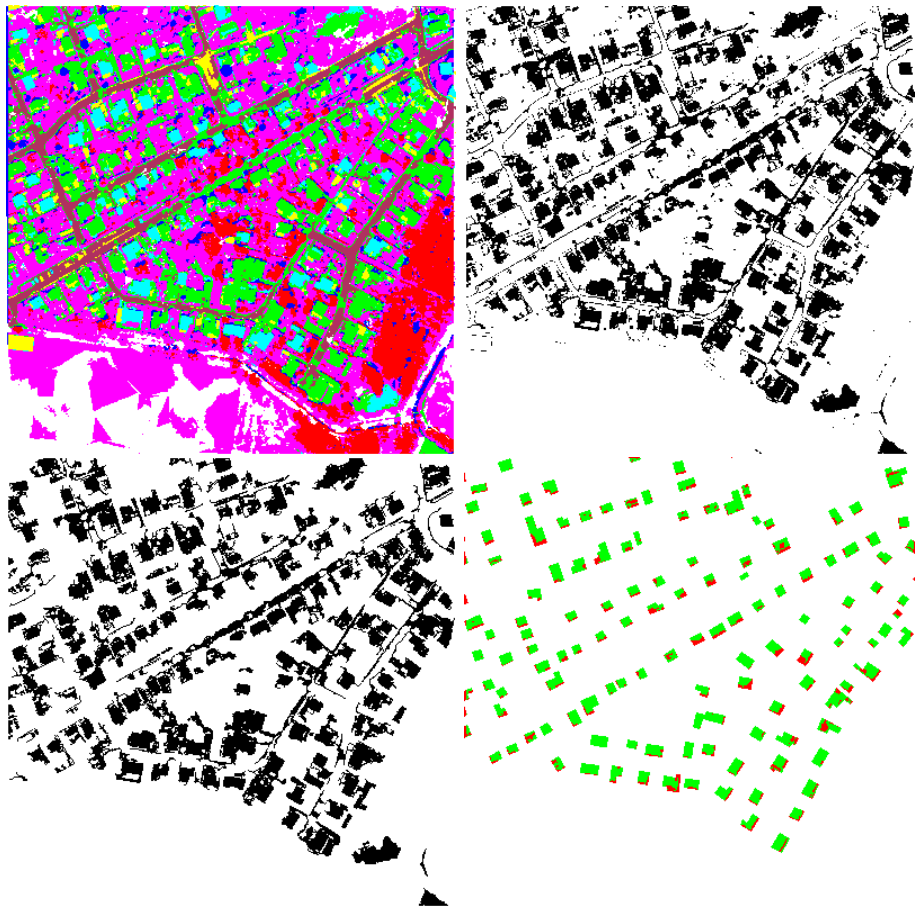


Figure 6.4: Test area 2, UL: FMLE classification of 5 layer image, UR: Extracted building clusters, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

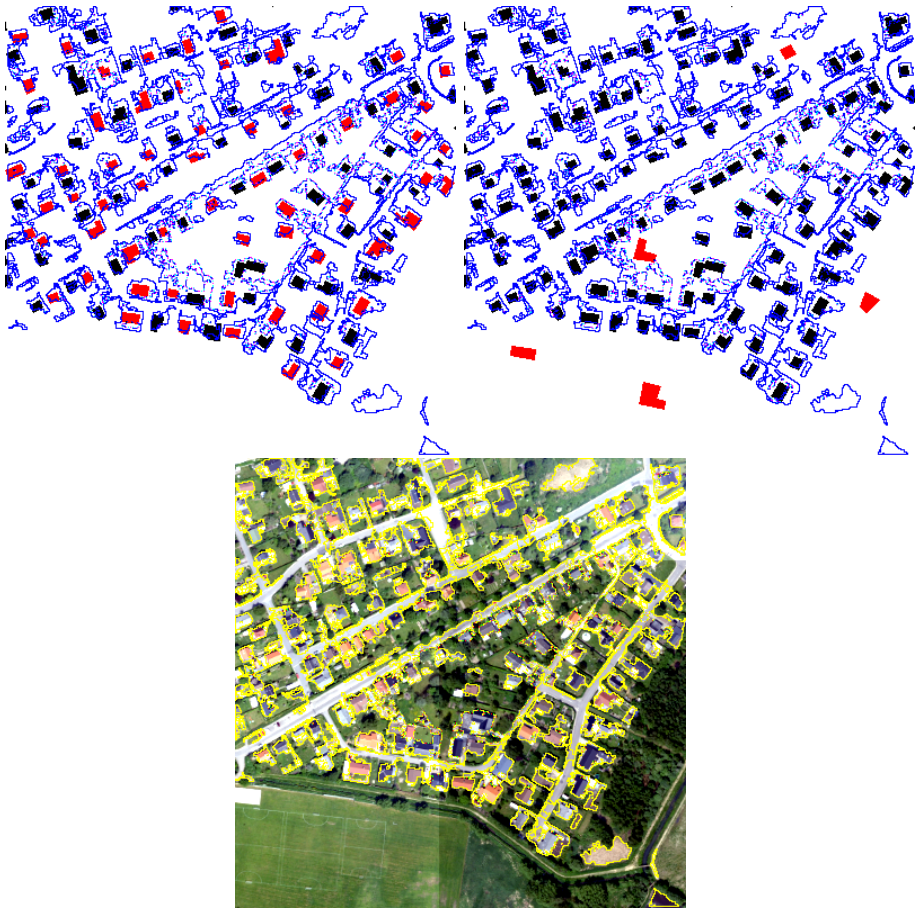


Figure 6.5: Test area 2, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

	Factual	Detected
New buildings	61	61
False alarms	≈50	
Demolished buildings	5	0

Table 6.3: Statistical results for: method 1, area 2

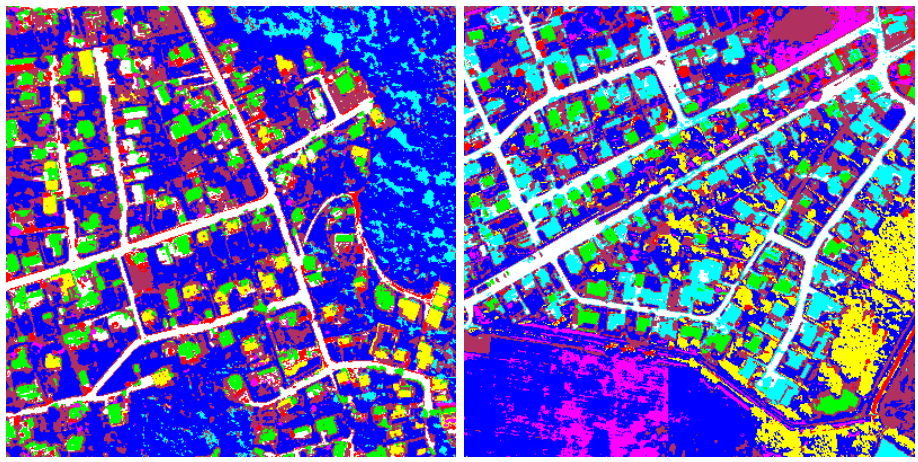


Figure 6.6: FMLE classification of RGBNIR images, Test area 1 (left) and Test area 2 (right).

Table 6.3 shows a better detection of new buildings than test area 1. However, a much larger amount of false alarms and misclassifications are also seen. None of the demolished buildings are detected and might just be a conincidence in this context.

6.3 Method 2

Method 2, as described in section 5.4, is only based on the unsupervised classification of the 4 layer image RGBNIR. Figure 6.6 shows the FMLE classification result of the 4 layer images from the two test areas. The optimal number clusters was found to be inadequate so 8 was chosen as the number of clusters. Like the unsupervised classification of the 5 layered images the buildings in figure 6.6 are classified into two clusters. The classification seems to work quite well for both test areas without the use of height data.

	Factual	Detected
New buildings	68	51
False alarms	10	
Demolished buildings	5	0

Table 6.4: Statistical results for: method 2A, area 1

6.3.1 Method 2A

The two building clusters from the FMLE classification of the RGBNIR images are extracted and merged.

6.3.2 Test area 1

Figure 6.7 UL shows the OAT class where all low-elevated pixels are removed. Most of the vegetation is removed with the intersection of OAT class and the FMLE classification shown in figure 6.7 UR. The classification check seems to find the most of the buildings in the map database, but the building clusters are smaller than for method 1. This is because the OAT class are created from the nDSM and it is an orthogonal projection with no relief displacement of the buildings.

Figure 6.8 shows that method 2A works less well for test area 1 than method 1. Table 6.4 shows that it detects less of the new buildings, probably because the OAT class are less well-defined. But method 2A has less false alarms and it does not detect demolished buildings. The low amount of false alarms are due to the use of the OAT class.

6.3.3 Test area 2

Figure 6.9 shows the result of method 2 for test area 2. The final building class are smaller, like test area 1, due to the orthogonal projection of the nDSM. It detect almost all the existing buildings and even a new one that is not in the map database.

To detect the changes, method 2A works better for test area 2 than for test area 1. It detects even small new buildings and only 1 new building is not detected even though it is visible on the OAT class image.

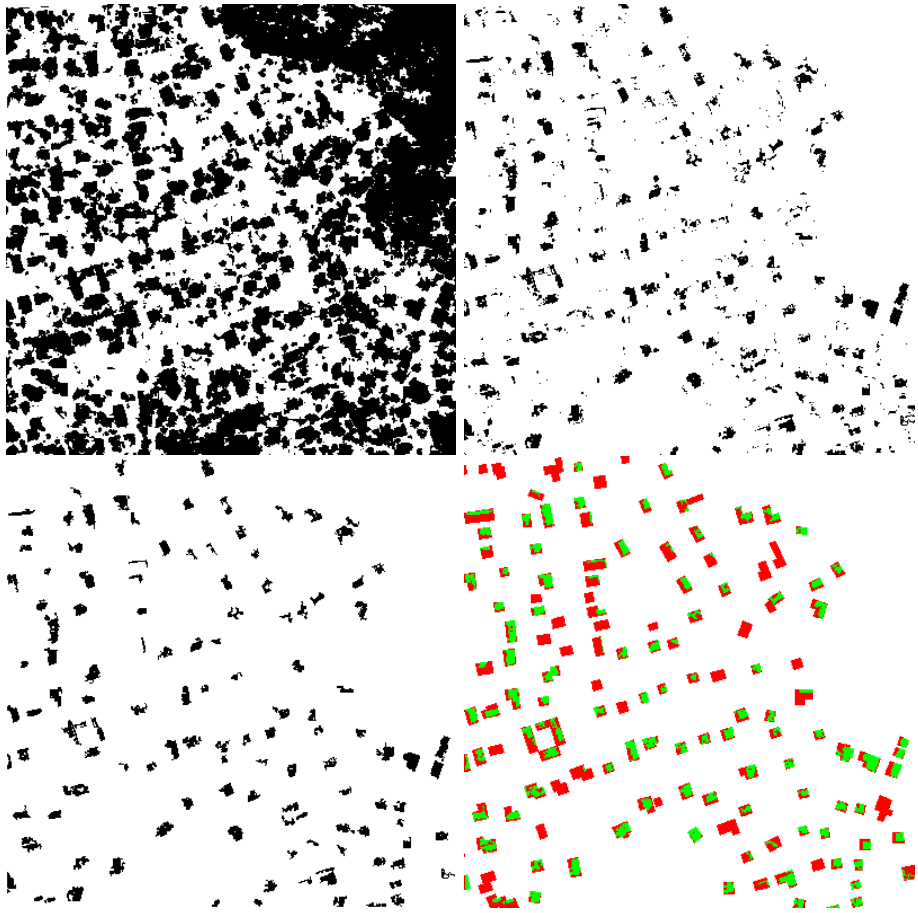


Figure 6.7: Test area 1, UL: OAT clusters, UR: Intersection of OAT clusters and FMLE classification, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

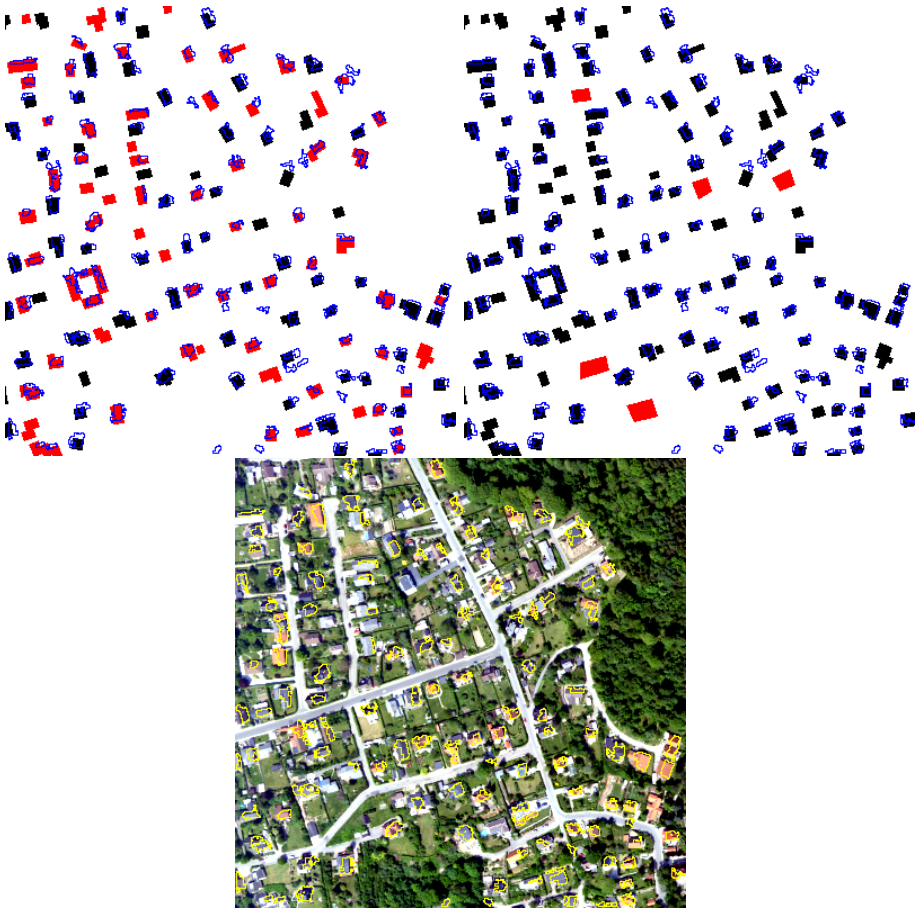


Figure 6.8: Test area 1, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

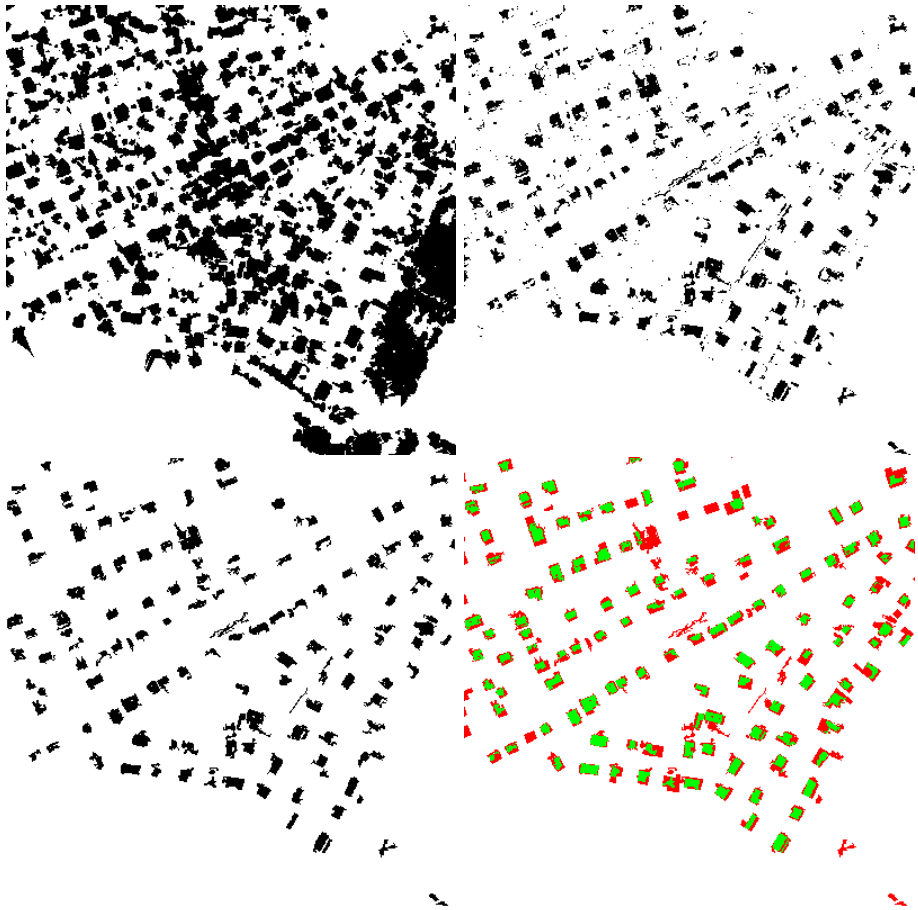


Figure 6.9: Test area 1, UL: OAT mask, UR: Intersection of OAT mask and FMLE classification, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

	Factual	Detected
New buildings	61	60
False alarms	10	
Demolished buildings	5	0

Table 6.5: Statistical results for: method 2A, area 2

Table 6.5 and figure 6.10 shows that method 2A does not detect the demolished building as just a coincidence and the amount of false alarms is small having almost the same amount as test area 1.

6.3.4 Method 2B

Now, by involving the class from the thresholded NDVI image, any false alarms in vegetated areas are expected to be eliminated.

6.3.5 Test area 1

For test area 1, figure 6.11 UL shows the thresholded NDVI image that contains one class of road, buildings, and other non-vegetated areas (all vegetation has been removed). The intersection of all the classes results in smaller building detections again due to the OAT class like described above. The classification check seems to work less well for method 2A due to the fact that less buildings are detected.

Figure 6.12 and table 6.6 clearly shows that the use of the NDVI thresholded image removes some false alarms that are probably located in vegetated areas. But it also seems to decrease the ability to detect new buildings. This is probably because some buildings are partially covered by tall trees causing that a part of the building is removed by the NDVI threshold. The final building class is then smaller than $25m^2$ and therefore removed by the size threshold.
As a coincidence method 2B does not detect any of the demolished buildings.

6.3.6 Test area 2

Figure 6.13 shows that method 2B works better for test area 2. The NDVI class are limited to the buildings and road along with some surrounding non-

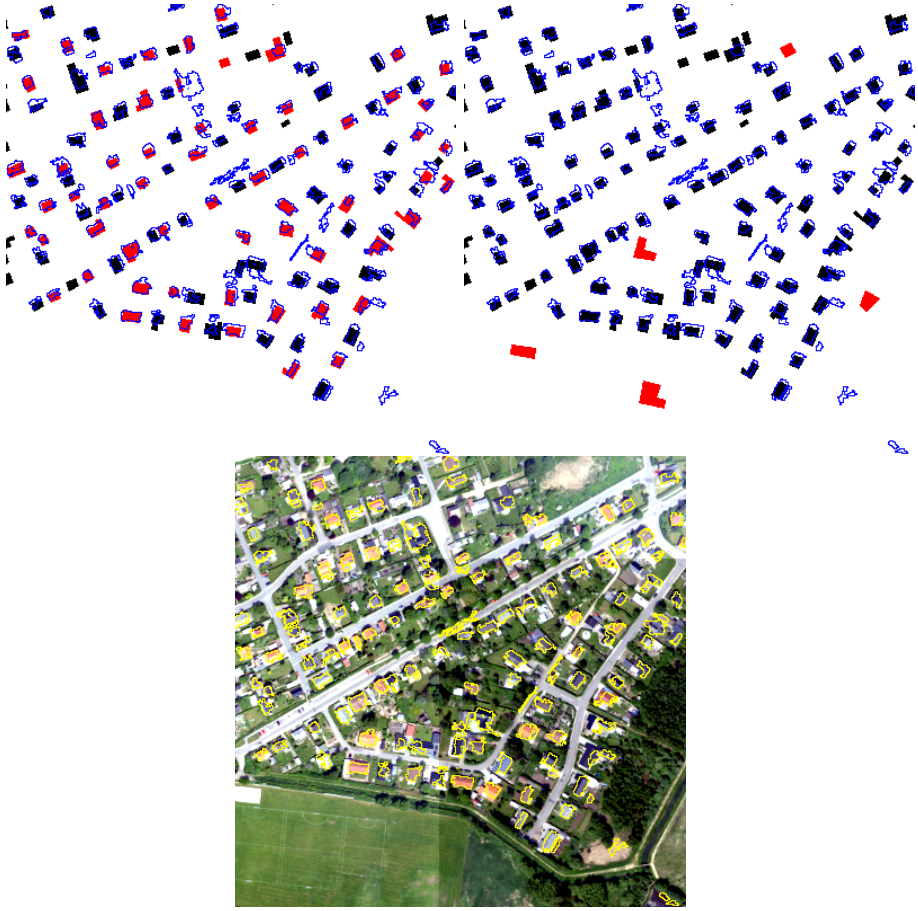


Figure 6.10: Test area 2, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

	Factual	Detected
New buildings	68	54
False alarms	6	
Demolished buildings	5	0

Table 6.6: Statistical results for: method 2B, area 1

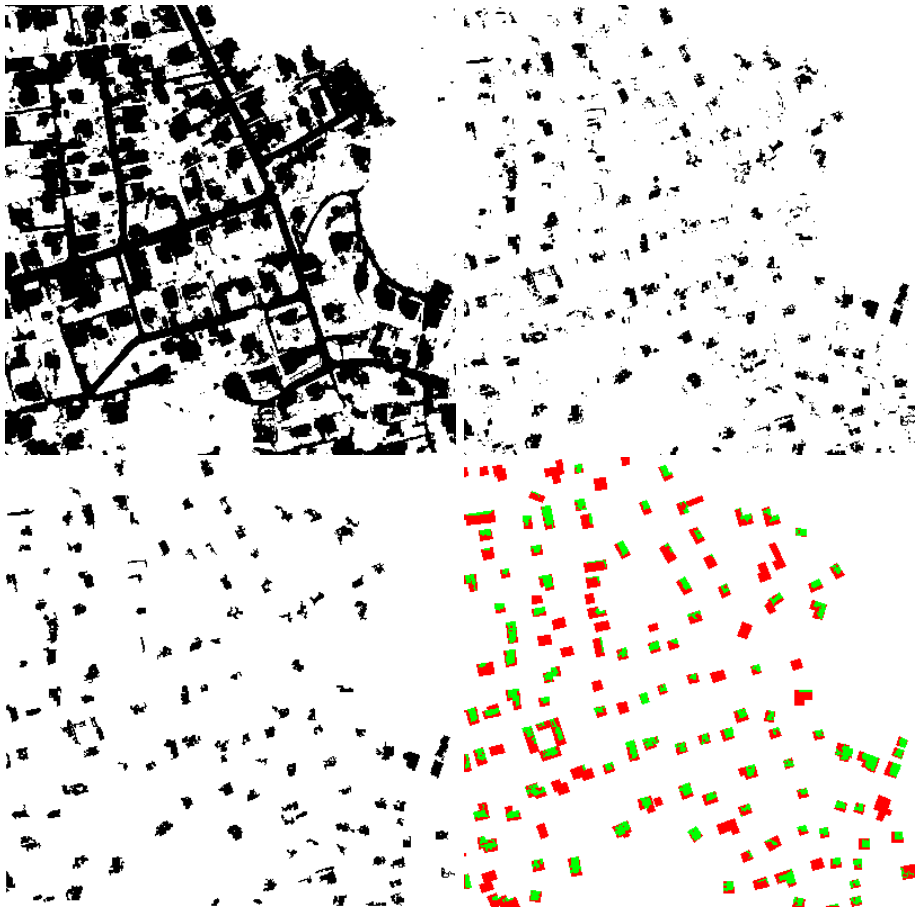


Figure 6.11: Test area 1, UL: NDVI mask, UR: Intersection of OAT cluster, NDVI cluster and FMLE classification, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

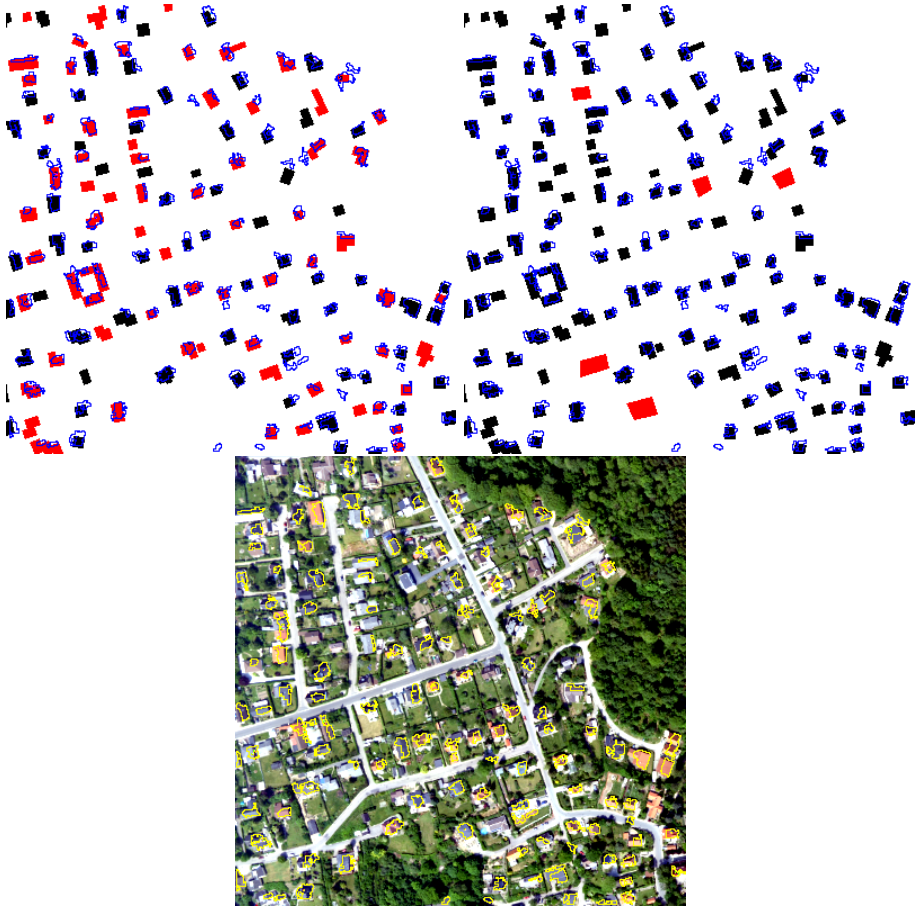


Figure 6.12: Test area 2, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

	Factual	Detected
New buildings	61	60
False alarms	8	
Demolished buildings	5	0

Table 6.7: Statistical results for: method 2B, area 2

vegetated areas. It detects more of the existing buildings but the extension of the final building class are still limited due to the OAT class.

The detection of new buildings works equally well as for method 2A. Only one building is not detected and it is the same as for method 2A, see figure 6.14. Since it is a part of the OAT class and the NDVI class it might removed by the FMLE classification. Table 6.8 shows that the amount of false alarms compared with method 2A is reduced, because of the involvement of the NDVI thresholded image. None of the demolished buildings are detected as a coincidence.

6.4 Method 3

The aim for method 3 was to derive a training sample under the TOP10DK mask of the 5 layer image and RGBNIR image. In addition, a prior probability image of known buildings would be used to see if it would have any effect. However, it turned out to be rather difficult to obtain a satisfactory result.
The use of a prior probability image did not seem to have any effect on any of the supervised classification results. However, an effect was seen on the entropy image of the class probabilities.

6.4.1 5 layer image

The supervised classification of the 5 layer images was empty for both test areas and all Mahalanobis distance thresholds. Appendix D shows the training sample and the plot of the optimal number of clusters. A possible reason for the empty classification is, that the Mahalanobis distance from a class center, e.g. class mean, to each pixel is very large, and extends the the Mahalanobis distance thresholds in table 5.1.
The entropy image computed by equation 4.6, figure 6.15 and 6.17 also verifies that only very limited areas shows large entropy values. This implies that most of the 5 layer image pixels are very far away from the class centers and have the

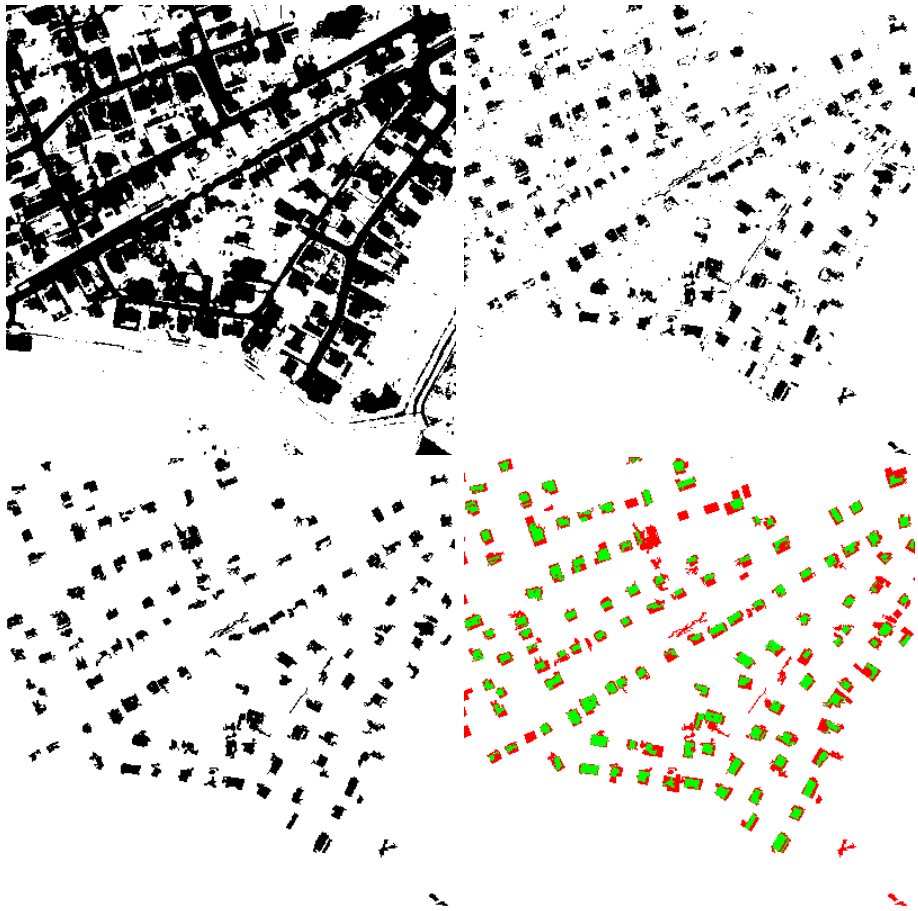


Figure 6.13: Test area 2, UL: NDVI mask, UR: Intersection of OAT mask, NDVI mask and FMLE classification, LL: Elimination of clusters $< 25m^2$, LR: Check of classification.

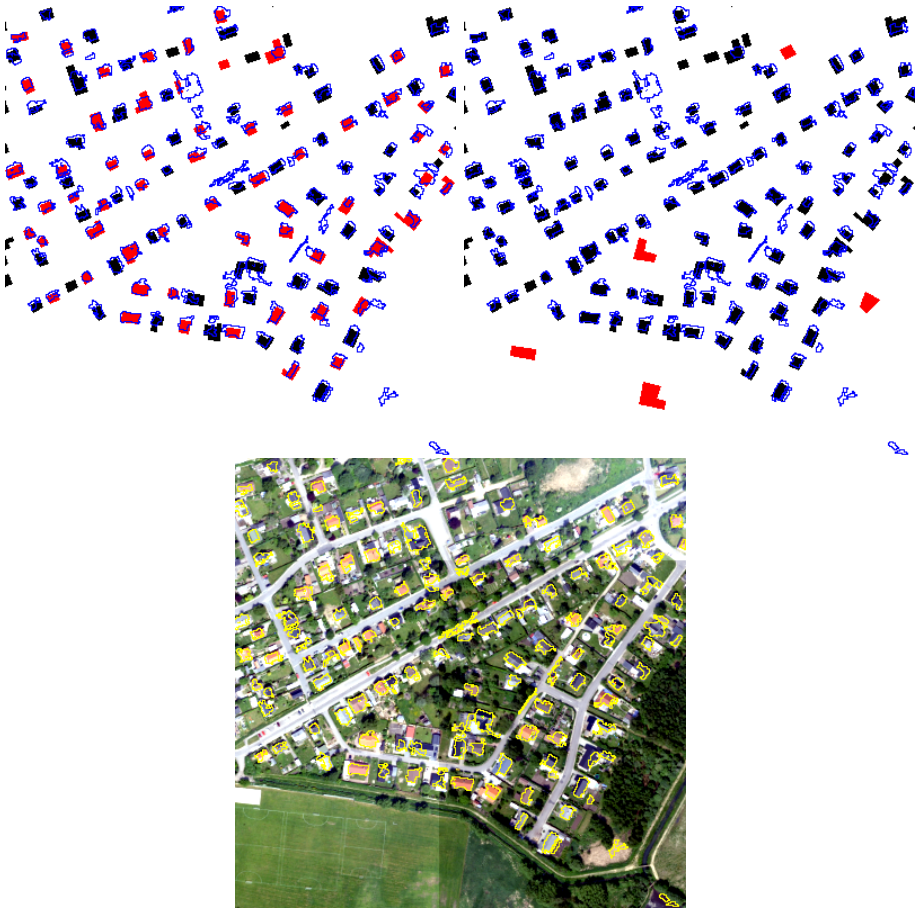


Figure 6.14: Test area 2, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.

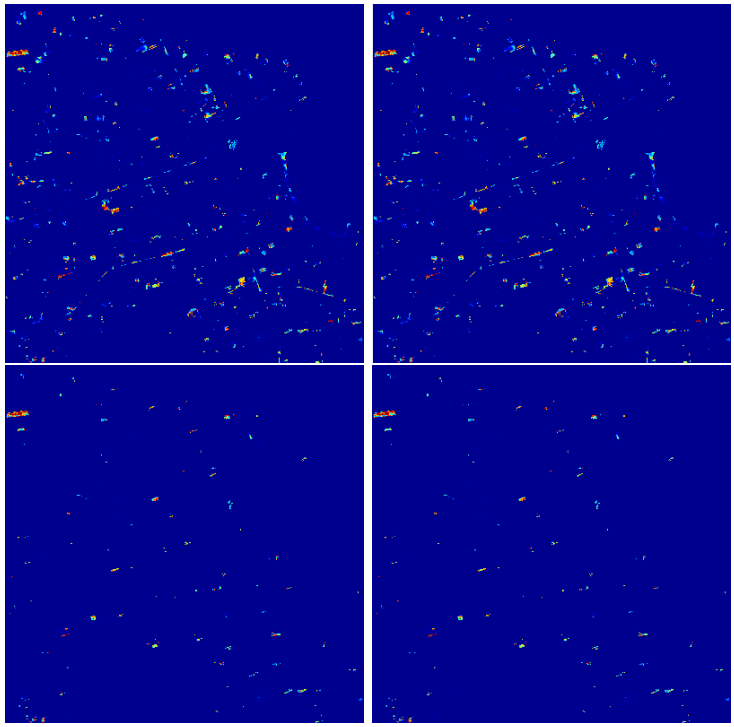


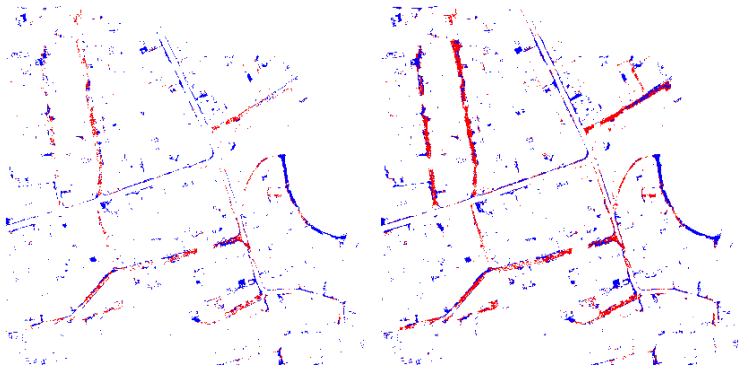
Figure 6.15: Entropy images test area 1, UL: No prior 95 %, UR: No prior 99 %, LL: With prior 95 %, LR: With prior 99 %.

same Mahalanobis distance to each class.

6.4.1.1 Test area 1

The entropy image of the 5 layer image for test area 1, figure 6.15, shows a significant effect using a prior probability image defined by the existing map database. Several possible false alarms seems to be eliminated by using a prior probability image. Furthermore the Mahalanabis distance threshold seems to work best for 99 % using a prior probability image.

To understand how the classification works, experiments on Mahalanobis distance parameter of 30 and 40 were carried out see figure 6.16. It does not positively detect what seems to be buildings. However, increasing Mahalanobis distance threshold implicates more pixels. Mostly what seems to be road segments are detected, with several spurious pixels.



(a) (b)

Figure 6.16: Increase of Mahalanobis distance, (a) 30, (b) 40.

6.4.1.2 Test area 2

The same tendency for test area 2 was seen as for test area 1. The entropy image, figure 6.17, shows very limited areas with high entropy, indicating that only few pixels have a high degree of membership to just one class. For test area 2 the use of a prior probability image also have an effect in eliminating areas that are not buildings.

Increasing the Mahalanobis distance parameter to 30 and 40 was also tested for test area 2. However, the supervised classification was empty in both cases.

6.4.2 4 layer image

The supervised classification of the RGBNIR image works better than for 5 layer image however only for test area 1.

6.4.2.1 Test area 1

Figure 6.18 shows the entropy images for the MLE of the RGBNIR image applied with a TOP10DK mask. Visually there is a clear difference compared with the MLE classification of the 5 layer image. An increasing amount of pixels seems to have a larger membership to just one of the classes. However, the implication of the prior probability image defined by the existing map database eliminates

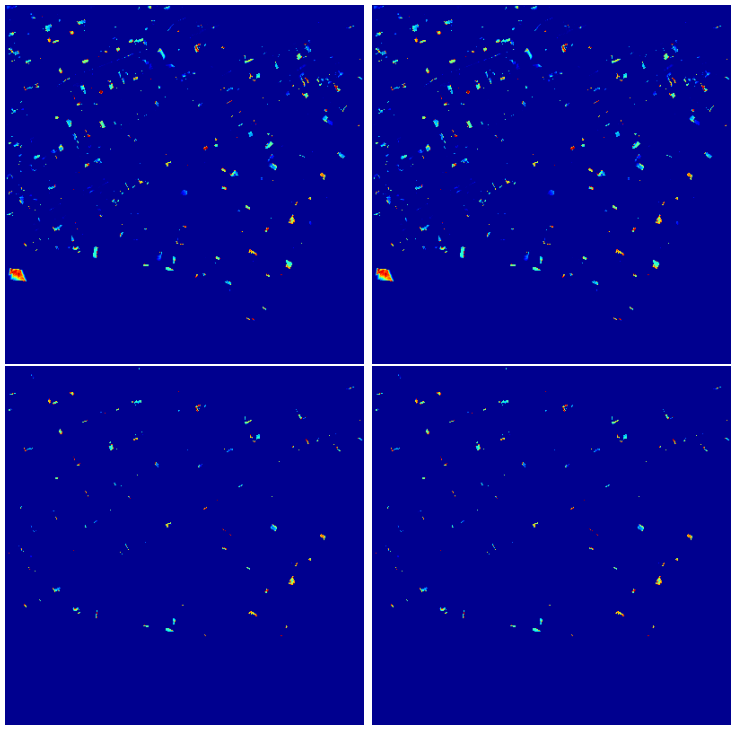


Figure 6.17: Entropy images test area 2, UL: No prior 95 %, UR: No prior 99 %, LL: With prior 95 %, LR: With prior 99 %.

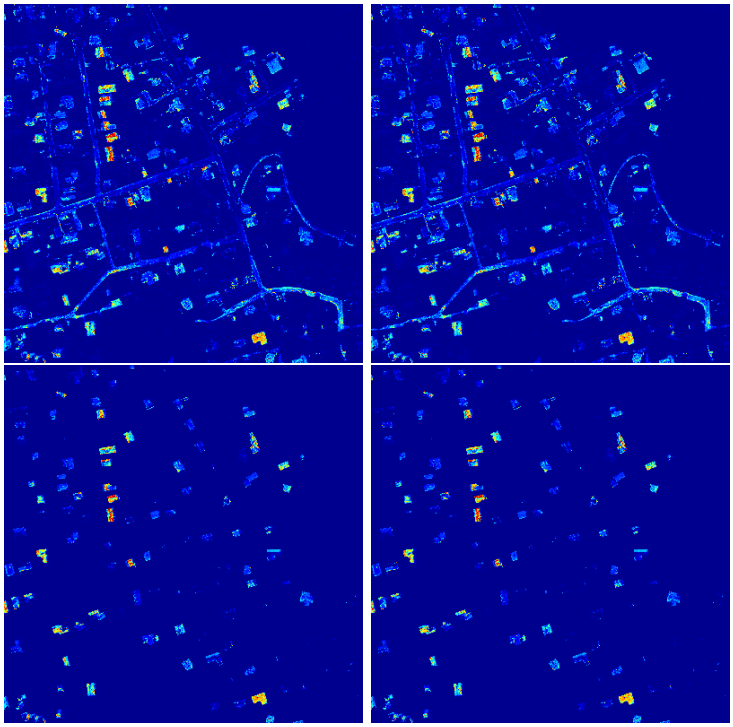
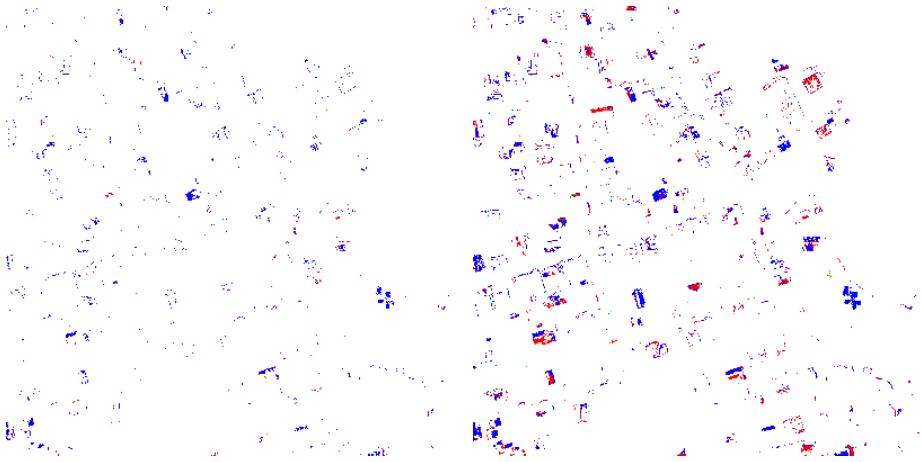


Figure 6.18: Entropy images test area 1, UL: No prior 95 %, UR: No prior 99 %, LL: With prior 95 %, LR: With prior 99 %.

the misclassified areas.

Figure 6.19 shows the supervised classification of test area 1 with no prior probability image, where it seems to succesfully detect some buildings. Even if the entropy image shows a difference by using a prior probability image defined by the map database, the classification does not show this difference. This might be expected since the computation of the Mahalanobis distance is not dependent on the prior probability, see section 4.2.1. There is however, a significant difference between using 95 % or 99 % of the χ^2 distribution. Applying a Mahalanobis distance threshold of 13.277 seems to result in the largest classes.

Intersecting this classification with the nDSM class creates figure 6.20 where some buildings seems to be detected. But as might be expected, taking the previous results into account, it does only detect few new building and no demolished buildings. Clearly, if the Mahalanobis distance parameter is increased the classification would work better if it is intersected with the OAT class to



(a) (b)

Figure 6.19: Classified image for test area 1, (a): No prior 95 %, (b): No prior 99 %

	Factual	Detected
New buildings	61	10
False alarms	0	
Demolished buildings	5	0

Table 6.8: Statistical results for: method 3, area 1

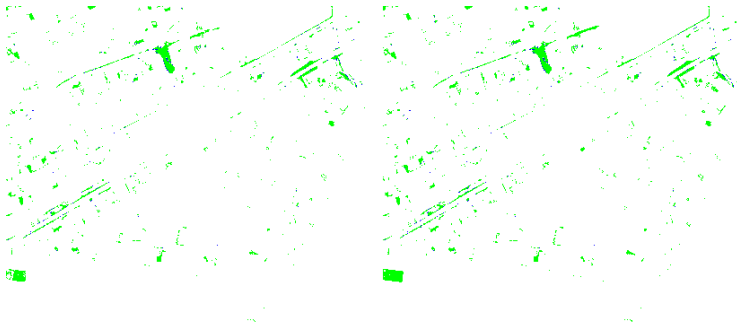
eliminate road segments and other spurious pixels. Summarising the detected buildings in table 6.8 shows a poor result as only 10 of the so-called new buildings are detected. However none of the demolished buildings are detected and there seems to be no false alarms.

An increase of the Mahalanobis distance shown in figure 6.21 will probably result in a more successful detection if it is intersected with the OAT class since it seems to detect most of the buildings in both cases (30 and 40). However, to find the right Mahalanobis distance will be a subject for future work.

6.4.2.2 Test area 2

Test area 2 worked out like the classification of the 5 layer image, section 6.4.1. Even though the entropy images at figure 6.22 shows a larger diversity in the

Figure 6.20: Test area 2 MLE, UL: Detection of new buildings, UR: Detection of demolished buildings, LC: Vectorized building clusters on RGB image.



(a) (b)

Figure 6.21: Increase of Mahalanobis distance, (a) 30, (b) 40.

entropy values, a successful classification could not be obtained. As seen for test area 1, it seems that most of the road segments could be eliminated from the classification by applying a prior probability image.

Testing test area 2 with larger Mahalanobis distances showed the same tendency as for the 5 layer image. Areas along roads and spurious pixels seem to have the shortest distance to the class centers as figure 6.23 shows. Increasing the distance does not have a positive effect on building detection.

6.5 Summary

The FMLE for building detection and three methods presented in chapter 5 have been evaluated. Method 1 and 2 have been evaluted for the detection of new buildings and the detection of demolished buildings and both methods seems to work reasonable well. Method 3 however, had difficulties in detecting buildings in both the 5 layer and RGBNIR images by the MLE and does only seem to work for test area 1 of the RGBNIR image.

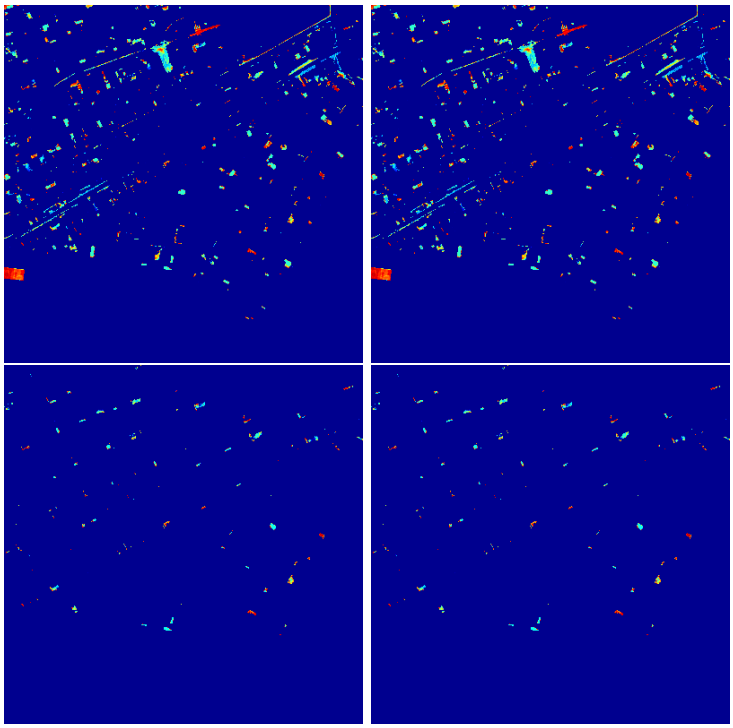
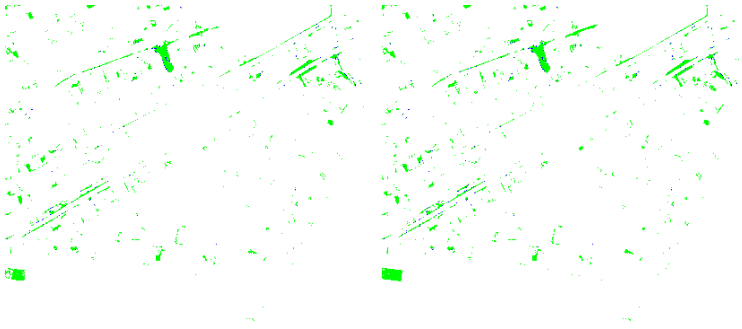


Figure 6.22: Entropy images test area 2, UL: No prior 95 %, UR: No prior 99 %, LL: With prior 95 %, LR: With prior 99 %.



(a) (b)

Figure 6.23: Increase of Mahalanobis distance, (a) 30, (b) 40.

Chapter 7

Discussion

In this chapter, the methods and all the results obtained in the project will be discussed. Furthermore possible sources of errors will be discussed leading to suggestions on improvements and modifications of the change detection methods. Finally, suggestions on future work will be considered.

7.1 Data

This section will discuss the two types of data, image data and digital elevation data.

7.1.1 Image data

The image data are as mentioned acquired with a digital aerial camera resulting in 4 layer aerial images containing the VIS and the NIR spectrum. Even though it is possible to acquire the NIR spectrum with a traditional film camera the main advantage of using a digital camera is that the whole visible spectrum can be acquired along with the NIR.
A disadvantage of the UltraCam D by Vexcel is that the final image size is

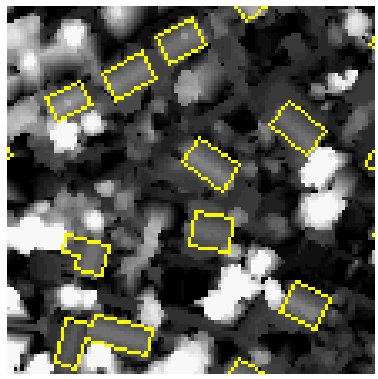


Figure 7.1: Zoom of the log(nDSM) overlayed with TOP10DK.

smaller than an analogue film of 23 × 23 cm, and therefore more images must be acquired to cover the same area. However, the storage issue and scanning of the analogue films must be the main disadvantage of the aerial film cameras.

7.1.2 Digital elevation data

The rasterization of the elevation data and the following nDSM computation shows that the buildings can be seen in contrast to the surrounding terrain. However, since it is only used for the creation of the OAT class it is only the elevation in each pixel that matters.
Concerning the computation of the log(nDSM), the spectral contrast of the buildings are much more important since it is involved in the FMLE classification. Figure 7.1 shows the log(nDSM) overlayed with the TOP10DK where it is difficult to see the spectral difference between the buildings and surrounding terrain. Since the trees are taller than the buildings, they can easily be discriminated.

The rasterization of the DTM and DSM in ENVI is dubious since it, in the nDSM and log(nDSM), can be difficult to distinguish between the spectral properties of the building roofs and the surrounding areas.

As described in section 2.3.1.1 the ALS also detects the intensity of the reflected laser pulse. This might be a useful feature for describing the surface from where the laser pulse is reflected. In connection with the detection of buildings with

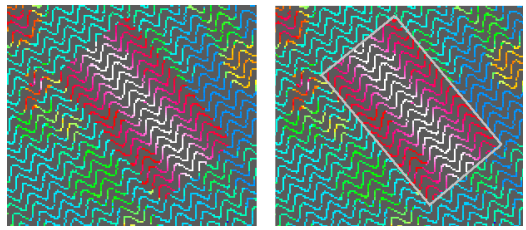


Figure 7.2: High density colour-coded DSM with an easily detectable building, Katzenbeisser [2004].

different roofing as described in section 2.2.2.1, the intensity measure from the laserscanner might be useful to improve the building detection since it describes the texture of the roof.

Figure 7.2 shows a colour-coded elevation image of the raw elevation data from the laserscanner, TopoSys Falcon II, Katzenbeisser [2004]. The dense cloud of points clearly shows the scan pattern and the very precise limits of a building. Instead of using the DTM and DSM with point spacings of 1 m, this raw elevation data gives a much more precise building detection. However, the amount of data is also much greater which requires a huge amount of storage capacity for even a small area.

7.2 Quality assessment

The quality assessment clearly shows that the level of information is greater in digital images than scanned analogue films. The visual comparison shows that the digital images are much more detailed and objects more well-defined. The level of noise is also much lower in the digital images confirming that much less information is lost. However, the presented method for noise estimation and information content is only representative for homogeneous image patches. Larger images covering larger areas with different information content have to be evaluated both visually and by comparing the accuracy of measured distances in the orthoimages.

7.3 Unsupervised classification

The evaluation of the unsupervised classification method shows that the amount of classes has to be at least 7 or 8 to obtain a satisfactory classification in order for the buildings to be discriminated from surrounding areas. The involvement of the spatial dependency is also shown to have a significant effect on the classification since many spurious pixels are eliminated.
As mentioned, one of the advantages of the FMLE is that opposed to many other unsupervised classification algorithms it does not attempt to equalise the sizes of all the clusters. Since buildings have various sizes and the test areas have different amounts of buildings it seems important to take this feature into account. However, the FMLE algorithm is slow and unstable compared to many other and more simple unsupervised classification algorithms. The initial classification at a coarser scale of the image pyramid accelerates the process and does not even reduce the properties of the final classification.
The use of the optimal number of classes can be questioned since it does not give a satisfactory result in every case. The FMLE classification of the RGBNIR images compared with the FMLE classification of the 5 layer image does not seem to have significant differences. As mentioned above in section 7.1.2 the spectral properties of the buildings in the log(nDSM) are not very clear resulting in a low influence on the unsupervised classification. Hence the advantage of using the log(nDSM) as an extra image layer is not very distinct.

7.4 Method 1

Method 1 seems to be successful since it detects almost all buildings. However, the evaluation of the FMLE classification shows that the required number of clusters has to be at least 7 or 8. Since the maximum partition density determines the number of clusters, it can vary from test area to test area. The optimal number of clusters can easily be lower than 7 or 8 and thus method 1 will not work as described in this project. In this case it would be necessary to intersect building class with OAT class and/or NDVI class to eliminate misclassified pixels.

Figure 7.3 shows that method 1 successfully detects a possible new building in test area 1 that is not registered in the existing map database TOP10DK, although it seems that some narrow road segment is misclassified as a building. The number of false alarms is largest for method 1 where some of them might be small sheds, greenhouses or outhouses that are not registered in the map database. However, only a field inspection would confirm that.

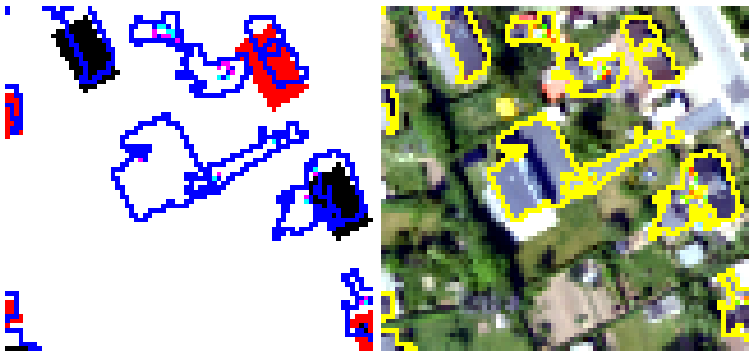


Figure 7.3: Possible new building not registered in the existing map database.

For test area 2 method 1 works less well. Many pixels along road segments are mis-classified and the building detections are hard to discriminate from one another as well as the false alarms. Here it seems obvious that a intersection with the OAT class from method 2 would eliminate several false alarms.
The extraction of the most promising building classes are performed manually in this project. A suggestion on how to automatize this could be to choose the two classes that have the largest amount of pixels that intersects with the existing map database.

7.5 Method 2

Method 2 works as well as method 1. When including the NDVI class in method 2B some false alarms located in vegetated areas seems to be eliminated. The extent of the building detections are more limited for method 2. This is due to the fact that the OAT class, derived from the orthogonal projection of the nDSM, eliminates the visibility of relief displacement in the classified RGBNIR orthoimage.
As opposed to method 1, the use of OAT class in method 2 also eliminates false alarms around existing buildings, that as mentioned are possible outhouses, sheds, or greenhouses, due to the elevation threshold of 2.5 m.
However, method 2 does not detect the non-registered new building that is found in figure 7.3 of method 1, since it is not found in the nDSM image in figure 7.4. This is most likely because the DSM and DTM of the laserscanned data are not fully updated with the image data or because this new building is lower than 2.5 m.

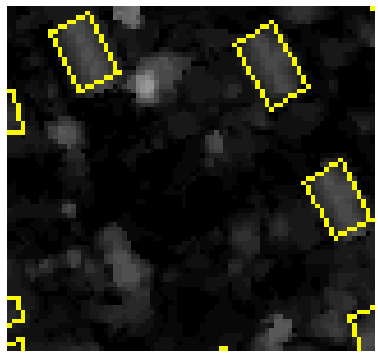


Figure 7.4: New house from method 1 is not present in the nDSM.

7.6 Method 3

The first line of approach to method 3 was to extract a training sample from areas where there are buildings. The expected result was a supervised classification with a Mahalanobis distance threshold classifying only the roofs of buildings. However, this approach was not shown to be successful. Even if the parameter is increased significantly areas with spectral similarities with the roofs, are misclassified as being buildings. As mentioned, there is a low diversity in the classes of the training samples for both test areas. A possible reason for this is that the training sample is derived from a very limited area and thus have a low deviation from the class mean. This results in very large Mahalanobis distances to unclassified pixels. For test area 1 the training sample is slightly larger than test area 2 and thus the deviation in each of class is larger. Since the buildings in the image data are exposed to the relief displacement they do not fit the TOP10DK mask perfectly. Pixels that might belong to a surrounding area are involved in the derivation of the training sample whereas other pixels that belong to a building roof are masked out. A way to increase the area, from where the training sample is derived, is to increase the size of the buildings in the TOP10DK mask. This could be done by applying mathematical morphology, Carstensen [2002]. Furthermore, by using true orthoimages, where the roofs fits the TOP10DK mask, all pixels involved in the creation of the training sample reflects the roof of a building.
The use of the prior probability image defined by the existing buildings have a significant effect in the entropy images as might be expected.
Method 3 was also tested with no Mahalanobis distance parameter. In the resulting classification buildings could hardly be discriminated from the roads and would have little value on their own.

7.7 Change detection

The methods for change detection can also be discussed. The detection of so-called new buildings seems to be successful, at least for method 1 and 2. However, in reality we do not have any prior knowledge about the new buildings and the change detection have to be confirmed by a field inspection.
The detection of the so-called demolished buildings can also be questioned. In most cases it does not work as intended. The location of the 5 demolished buildings are completely arbitrary since we do not know where there have been buildings. In reality it will give a more reliable result if the change detection methods were compared with an older version of the map database TOP10DK.

Chapter 8

Conclusion

In this first part of this project a simple comparative quality assessment was performed for digital and analogue images. This was done both visually and concerning the deviation of the image data. This comparison showed that the level of information content in the digitally acquired image is much larger, which is further supported by a significantly lower level of noise. It can therefore be concluded that there are significant advantages in using digitally acquired images instead of digitized analogue imagery.
These evident benefits of the digital imagery is very useful in defining the spectral properties of different roof materials. However, it can be difficult to standardize since it is very dependent on the type of data source and time of acquisition.

The unsupervised classification algorithm FMLE has been used for classification of RGBNIR and 5 layer images. A successful classification has shown to be very dependent on the optimal number of clusters. The strength of the FMLE classification is that it successfully can detect buildings at a lower resolution (pyramid depth 2) than the original image. This accelerates the algorithm. However, the instability of the algorithm can become a problem, since a semi-automatized procedure requires a stable algorithm in order to work sufficiently well. It is therefore unclear if it is a optimal algorithm for change detection.

Three change detection methods have been presented and evaluated, where method 1 and 2 successfully detects almost all existing buildings. These two methods can then be used for semi-automatic map updating.
It can be concluded that method 1 works if the number of clusters are adequate for discrimination of the buildings from surrounding areas. However, for test area 2 the amount of false alarms makes it hard to discriminate between the different buildings. For test area 1 method 1 detects what seems to be a new builing that is not registered in the map database TOP10DK. The involvement of the log(nDSM) does only seem to have little influence on the final classification.
Method 2 works better for both test areas and it can be concluded that including both the NDVI class and OAT class gives the most satisfactory result with the fewest false alarms. Therefore, it seems important to use the NDVI image to eliminate false alarms in vegetated areas.
Method 3 works inadequately. The method only detects 10 buildings by supervised classification by using the training sample derived from the RGBNIR image with the TOP10DK mask. However, the use of a prior probability image defined by the existing buildings in the map database, seems to have some effect on the supervised classification.

8.1 Future perspectives

The quality assessment shows that new digitally acquired images contain much more information which is and can become very useful. Different objects are much more well-defined in the digital acquired images which makes it possible to measure distances in the image more precisely.

The rasterization of the DSM and DTM needs to be more precise, since they are used to create the nDSM and log(nDSM). Even though the log(nDSM) only has little effect on the unsupervised classification it might be improved if the rasterization of the DSM and DTM is improved and the spectral properties of the buildings in the log(nDSM) are more clear. This would also improve the nDSM, resulting in more well-defined buildings in the OAT class.
An improvement of method 1 can be done by intersection of the building class with the OAT class to eliminate several false alarms.

Method 3 needs to be improved by extending the buildings in the TOP10DK mask to extract a larger training sample. In combination with this the inclusion of the prior probability image defined by the existing buildings extended by

mathematical morphology would probably give a more successful classification. However, more tests are needed to find the most optimal Mahalanobis distance parameter.

Appendix A

Data set

Figure A.1: TOP10DK overlayed with test areas in RGB.

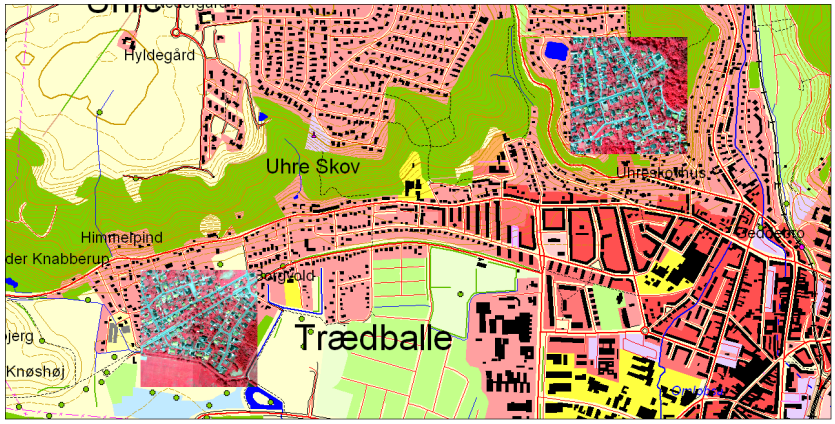


Figure A.2: TOP10DK overlayed with test areas in CIR.

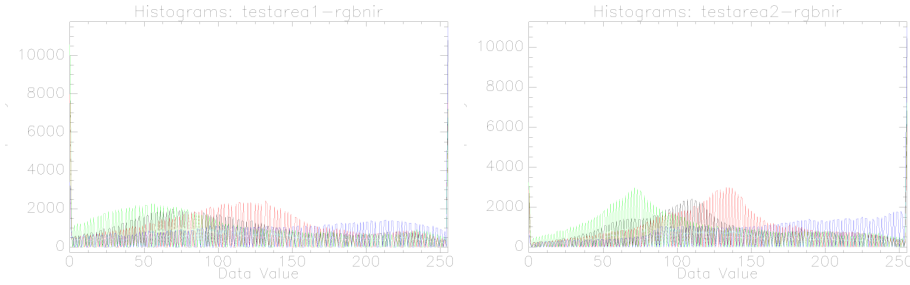


Figure A.3: Image histograms for test area 1 (left) and test area 2 (right).

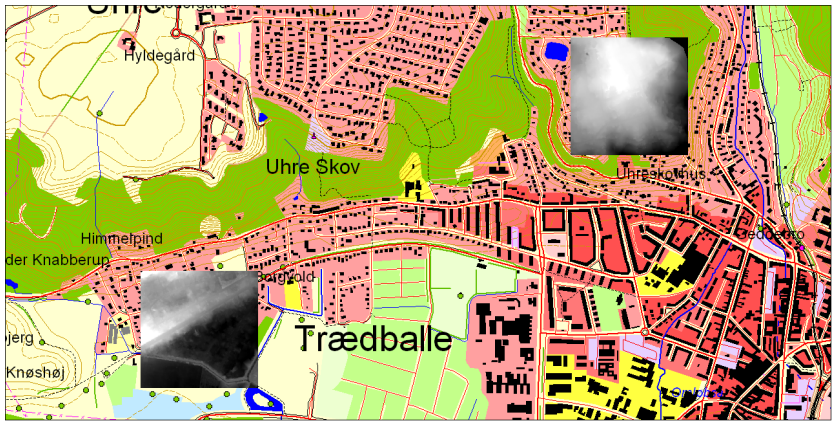


Figure A.4: TOP10DK overlayed with test areas in DTM.

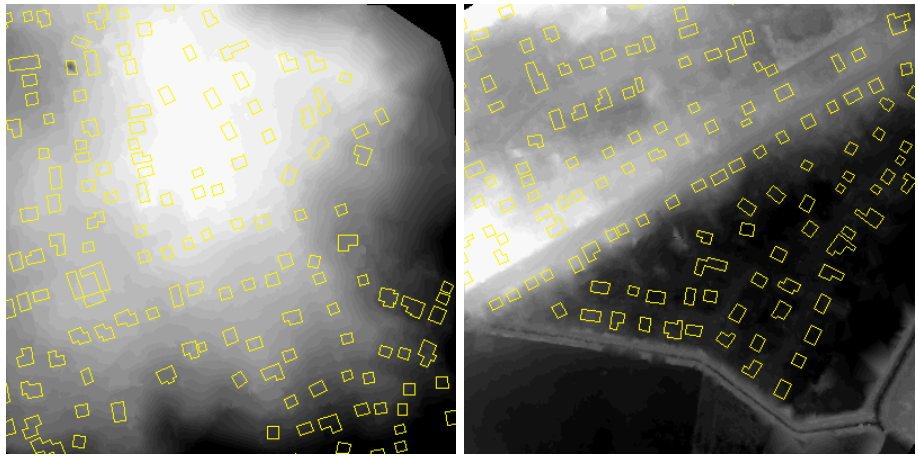


Figure A.5: DTM for test area 1 (left) and test area 2 (right)

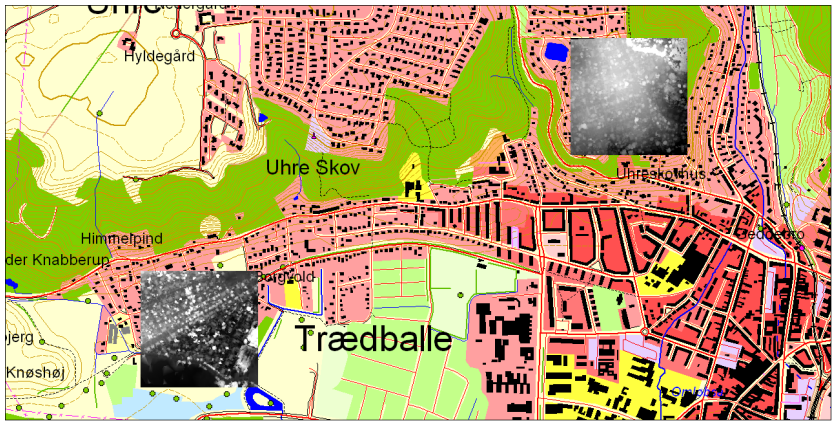


Figure A.6: TOP10DK overlayed with test areas in DSM.

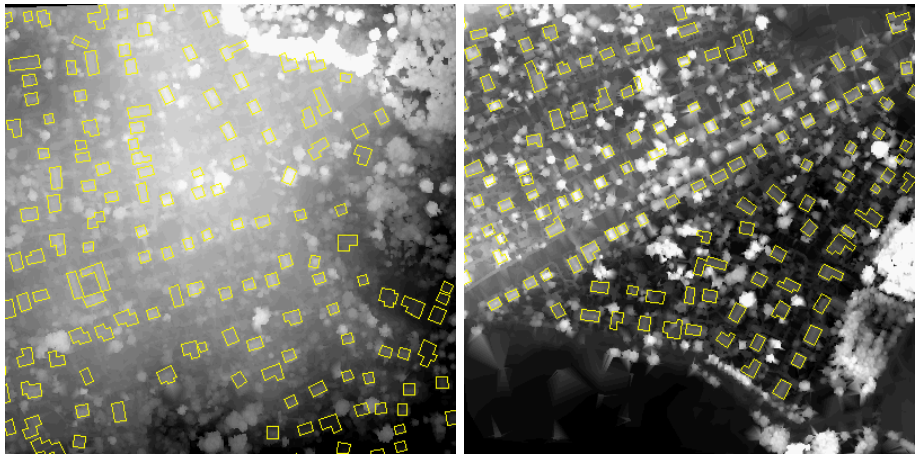


Figure A.7: DSM for test area 1 (left) and test area 2 (right)

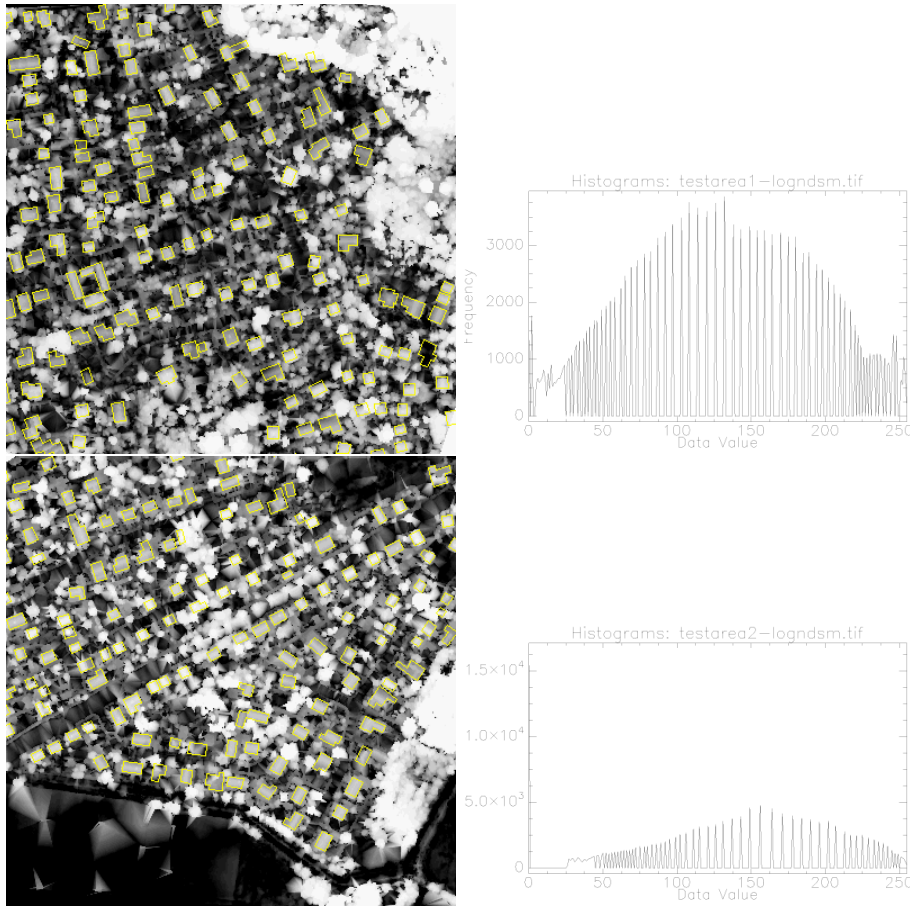


Figure A.8: log(nDSM) and stretched histogram for test area 1 (upper) and test area 2 (lower).

Figure A.9: NDVI images of test area 1 (left) and test area 2 (right).

Appendix B

ENVI Code

B.1 nDSM function

```
; Function for determination of a nDSM.
; Run by band math under basic tools menu and restore ndsm.exp.
; Assign DTM to b1 and DSM to b2.

; Author: Jens Popp Andersen
; Master thesis: Data fusion and semi-automatic map updating.

function ndsm, b1, b2
    result = float(b2)-b1
    ;Possible negative values are 0.0
    ptr = where(result lt 0., count)
    result[ptr] = 0.0
    return, result
end
```

B.2 log(nDSM) function

```
; Function for determination of a logarithmic nDSM.
; Run by band math under basic tools menu and restore logndsm.exp.
; Assign DTM to b1 and DSM to b2.
; Possible negative values are 0.0

; Author: Jens Popp Andersen
; Master thesis: Data fusion and semi-automatic map updating.

function logndsm, b1, b2
    result = float(b2)-b1
    ; all values below 0.0 equals 0.0
    ptr = where(result lt 0., count)
    result[ptr] = 0.0
    ; logarithm to all values greater than 0.0
    logptr = where(result gt 0., count)
    result[logptr] = alog(result[logptr])
    ; all 0.0 values are assigned lowest logarithmic value
    zeroptr = where(result eq 0., count)
    result[zeroptr] = min(result)
    return, result
end
```

Appendix C

MATLAB Code

C.1 Quality assessment

C.1.1 Main script and SNR computation

```
% Script for computation of histogram, entropy and SNR
% for RGB colour images.
% Images contains of 256 colours.
close all
clear all
format compact

nbands = 3;
ncolour = 256;

% Input analogue image.
fid=0;
while fid < 1
   filename=input('Open analogue image: ', 's');
   [fid,message] = fopen(filename, 'r');
   if fid == -1
```

```
     disp(message)
   end
end
analogue=fread(fid,'uint8');
fclose(fid);

% Input digital image.
fid=0;
while fid < 1
   filename=input('Open digital image: ', 's');
   [fid,message] = fopen(filename, 'r');
   if fid == -1
     disp(message)
   end
end
digital=fread(fid,'uint8');
fclose(fid);

% Image size and number of variables.
alength = length(analogue);
dlength = length(digital);

if alength ~= dlength
    error('Analogue and digital image are not same size');
end

nrows = sqrt(alength/nbands);
ncols = nrows;
nobs = ncols*nrows;

% Images are reshaped to obs.*band -array.
ana=reshape(analogue,nobs,nbands);
dig=reshape(digital,nobs,nbands);

% Calculation of the entropy and SNR.
for i = 1:nbands
    HA(:,i) = entropy(ana(:,i),nobs,ncolour);
    HD(:,i) = entropy(dig(:,i),nobs,ncolour);
    SNRA(:,i) = mean(ana(:,i))/(std(ana(:,i)));
    SNRD(:,i) = mean(dig(:,i))/(std(dig(:,i)));
end

% Mean values of all bands.
HAm = mean(HA,2);
```

```
HDm = mean(HD,2);
SNRAm = mean(SNRA,2);
SNRDm = mean(SNRD,2);

maxy = max([max(max(hist(ana,ncolour))); ...
max(max(hist(dig,ncolour)))]);

% Visualizing histograms, resulting values and images
% with zero mean and unit variance.
figure,
subplot(221), hist(ana,ncolour), title('Analogue image')
xlabel('Number of colours'), ylabel('Frequency'),
axis([1 ncolour 0 maxy])
subplot(222), hist(dig,ncolour), title('Digital image')
xlabel('Number of colours'), ylabel('Frequency'),
axis([1 ncolour 0 maxy])
subplot(223), imshow(reshape(ana,nrows,ncols,nbands)./ncolour)
title(['H = ',num2str(HAm), ', SNR = ',num2str(SNRAm)])
subplot(224), imshow(reshape(dig,nrows,ncols,nbands)./ncolour)
title(['H = ',num2str(HDm), ', SNR = ',num2str(SNRDm)])
```

C.1.2 Entropy function

```
% Function for entropy computation.
%
% Input:    im (band as column vectors),
%           nobs (number of observations (rows*columns)).
%           ncolor (number of colours in image).
% Output    H (entropy value).

function H = entropy(im,nobs,ncolor)

% Number of colours
h = hist(im(:),ncolor)';

% Probability of each colour value.
p = h./nobs;

% Only probabilities larger than 0 are used.
P = p(p~=0);

% The entropy H.
```

```
H = -sum(P.*log2(P));
```

C.2 Supervised classification

C.2.1 Main script

```
% Main script for supervised classification MLE.

% User input:   Number of image bands (layers).
%               Maximum Mahalanobis distance threshold.
%               Multispectral image.
%               Training sample image (binary image file).
%               Prior probability image (optional).
% User output:  Classified image (.bin file).

clear all
close all
format compact

% User input.
nbands = input('Number of bands in multispectral image: ');
thresh = input('Maximum Mahalanobis distance: ');

% Reading multispectral image into X.
fid=0;
while fid < 1
   filename=input('Open multispectral image: ', 's');
   [fid,message] = fopen(filename, 'r');
   if fid == -1
     disp(message)
   end
end
X = fread(fid,'uint8');
fclose(fid);

Xlength = length(X);

nrows = sqrt(Xlength/nbands);
ncols = nrows;
nobs = nrows*ncols;
```

```
if ncols ~= nrows
    error('Not a quadratic image!')
end

% Reading training sample into T.
fid=0;
while fid < 1
   filename=input('Open training sample image: ', 's');
   [fid,message] = fopen(filename, 'r');
   if fid == -1
     disp(message)
   end
end
T=fread(fid,'uint8');
fclose(fid);

stats = regionprops(T,'area');
nclass = length(stats);

% Requesting use of map database as prior probability image.
prquest = input...
('Use existing map database as a prior probability? Y/N [Y]: ','s');
if isempty(prquest)
    prquest = 'Y';
end

if prquest == 'Y'
    fid=0;
    while fid < 1
        filename = input('Open prior probability image: ', 's');
        [fid,message] = fopen(filename, 'r');
        if fid == -1
            disp(message)
        end
    end
    prior = fread(fid,'uint8');
    fclose(fid);
    % prior probability from map database.
    idhigh = find(prior==1);
    idhighl = length(idhigh);
    prior(idhigh) = 0.9/idhighl;
    idlow = find(prior==0);
    idlowl = length(idlow);
```

```
    prior(idlow) = 0.1/idlowl;
elseif prquest == 'N'
% static prior probability of 1/nclass.
    prior = ones(nobs,1)/nclass;
end

classimage = input('Save classified image: ','s');

[P, I, S] = classfct(X, nbands, nclass, nrows, ncols, ...
nobs, T, prior, thresh);

figure,
colormap jet
subplot(221), imagesc(reshape(P(:,1),nrows,ncols)'), axis image
title('Class 1')
subplot(222), imagesc(reshape(P(:,2),nrows,ncols)'), axis image
title('Class 2')
subplot(223), imagesc(reshape(P(:,3),nrows,ncols)'), axis image
title('Class 3')
subplot(224), imagesc(reshape(P(:,4),nrows,ncols)'), axis image
title('Class 4')

% Computation of the entropy.
H = -sum(P.*log2(P),2);
idnan = find(H==NaN);
H(idnan) = 0;

% Entropy image
figure, colormap jet
imagesc(reshape(H,nrows,ncols)'), truesize,
axis off, title('Entropy image')

L = reshape(I,nrows,ncols,1)';

cmap = [1 1 1; 1 0 0; 0 1 0; 0 0 1];

figure
colormap(cmap)
imagesc(L), truesize, axis off, title('Classified image')

fwriteid = fopen(classimage, 'w');
count = fwrite(fwriteid,(I),'uint8');
status = fclose(fwriteid);
```

C.2.2 Maximum Likelihood function

```
% Function for Maximum likelihood classification.
%
% Input:    X (raw image),
%           nbands (number of band (variables) in input image).
%           nclass (number classes in training sample).
%           thresh (maximum Mahalanobis distance).
%           T (training sample).
%           prior (prior probability image).
% Output:    I (classified image, row vector).

function [Pim I S] = classfct(X, nbands, nclass, nrows, ncols, ...
nobs, T, prior, thresh)

X=reshape(X,nobs,nbands);

var = [];
noc = [];
my = [];
S2 = [];

% Computation of mean and covariance for all class values
% different from 0.
for i = 1:nclass
    aux = X(find(T==i),:);
    var = [var; aux];
    noc = [noc size(aux,1)];
    my = [my; mean(aux,1)];
    S2 = [S2 cov(aux,1)];
end

% Covariance matrix reformed to 3D matrix.
S = reshape(S2,nbands,nbands,nclass);

% Log-likelihood computation.
for i = 1:nclass
    Xl = X-repmat(my(i,:),nobs,1);
    D(i,:) = sum((Xl'.*(inv(S(:,:,i))*Xl')));
    P(:,i) = log(prior)- ...
     repmat(0.5*log(det(S(:,:,i))),nobs,1)-(0.5*D(i,:)');
    Pim(:,i) = exp(P(:,i));
    idx = find(D(i,:)>thresh);
```

```
    P(idx,:) = 0;
end

% Largest membership and subscript for each class are extracted.
[C I] = max(P,[],2);

% All zero probabilities are unclassified (class 0).
idnull = find(C==0);
I(idnull) = 0;
```

Appendix D

FMLE classification

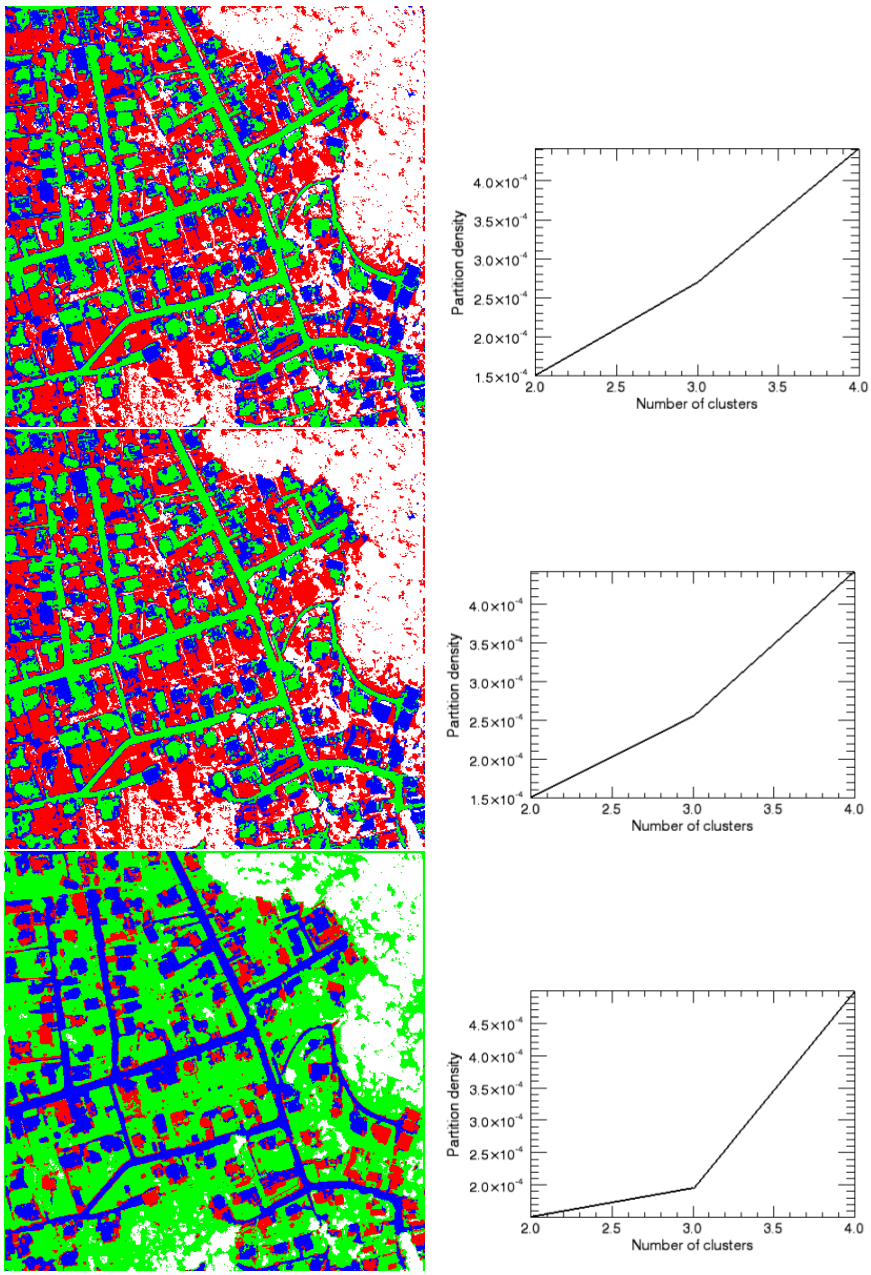


Figure D.1: FMLE classification of test area 1 at pyramid depth of 1 with plot of partition density, Upper: $\beta = 0$, Center: $\beta = 0.5$, Lower: $\beta = 1$.

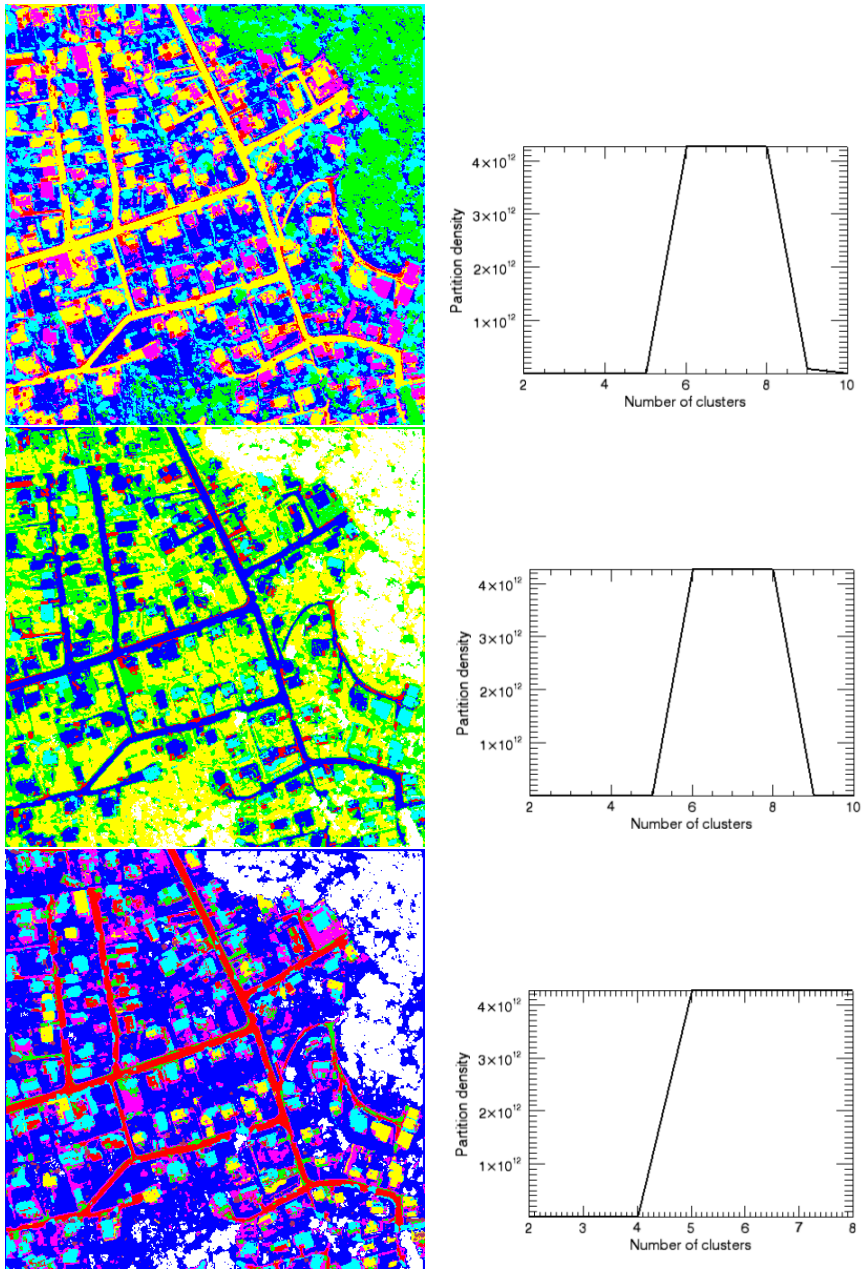


Figure D.2: FMLE classification of test area 1 at pyramid depth of 2 with plot of partition density, Upper: $\beta = 0$, Center: $\beta = 0.5$, Lower: $\beta = 1$.

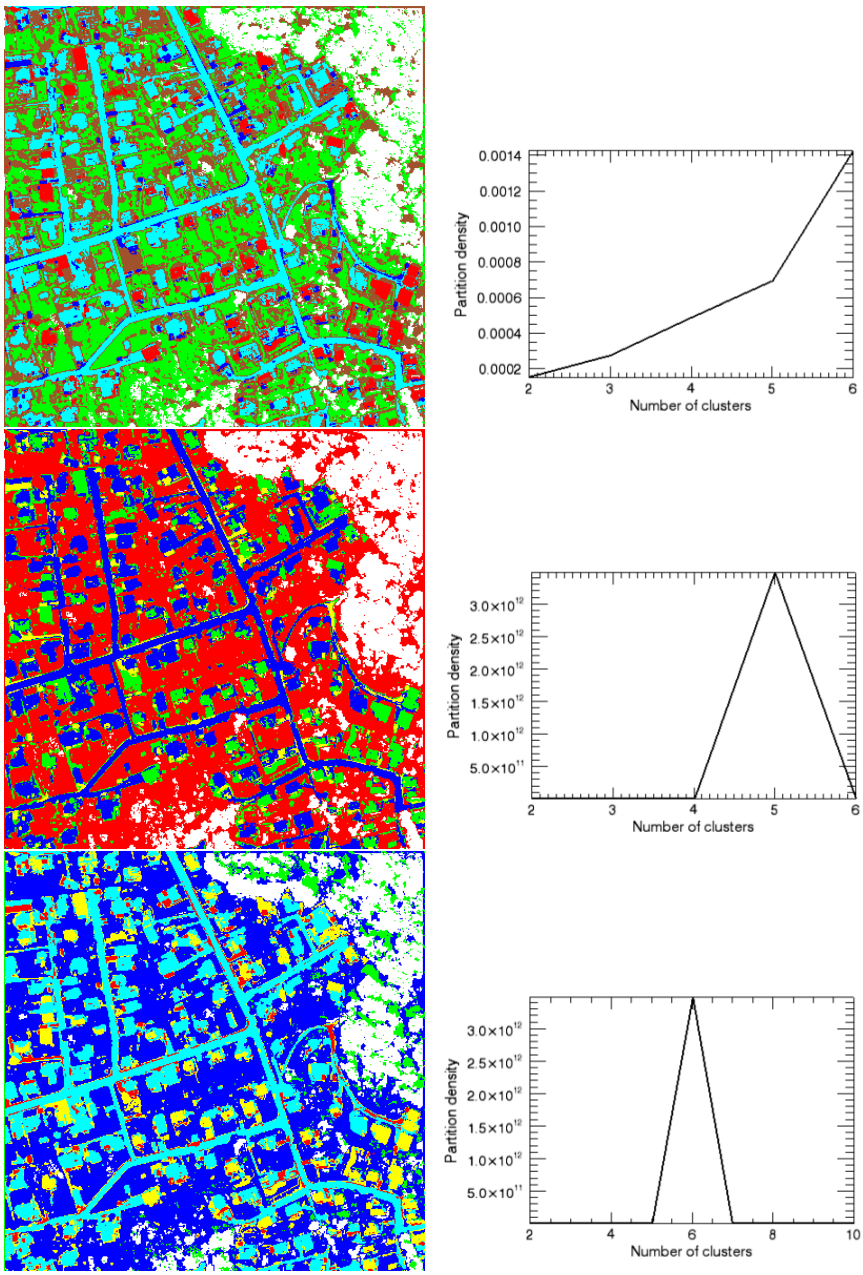


Figure D.3: FMLE classification of test area 1 at pyramid depth of 3 with plot of partition density, Upper: $\beta = 0$, Center: $\beta = 0.5$, Lower: $\beta = 1$.

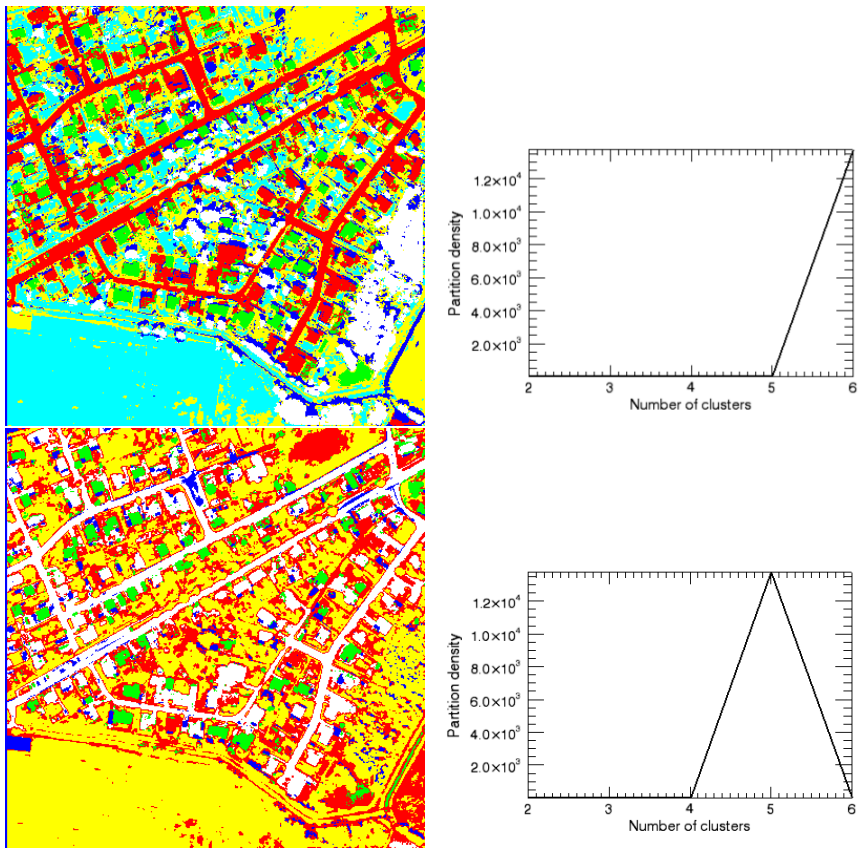


Figure D.4: FMLE classification of test area 2 at pyramid depth of 1 with plot of partition density, Upper: $\beta = 0$, Lower: $\beta = 0.5$.

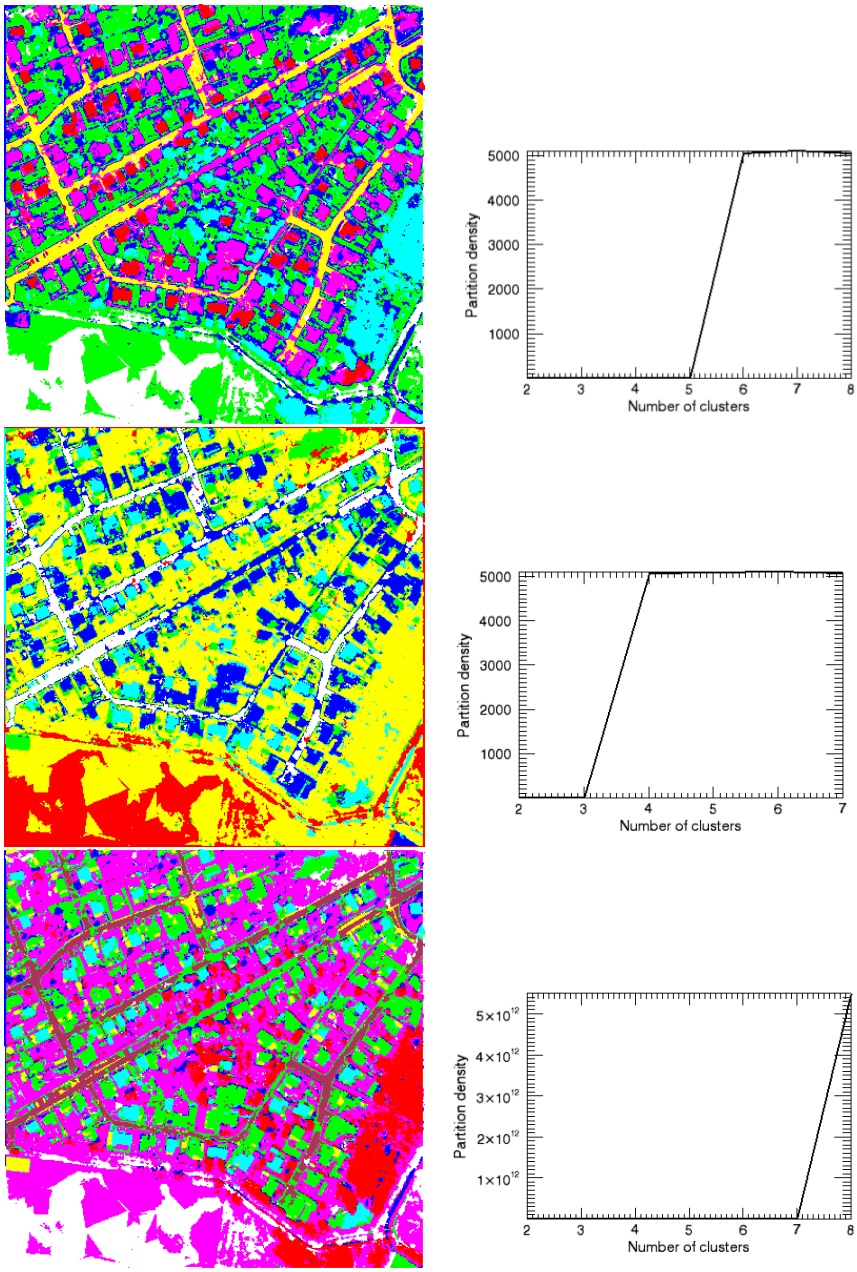


Figure D.5: FMLE classification of test area 2 at pyramid depth of 2 with plot of partition density, Upper: $\beta = 0$, Center: $\beta = 0.5$, Lower: $\beta = 1$.

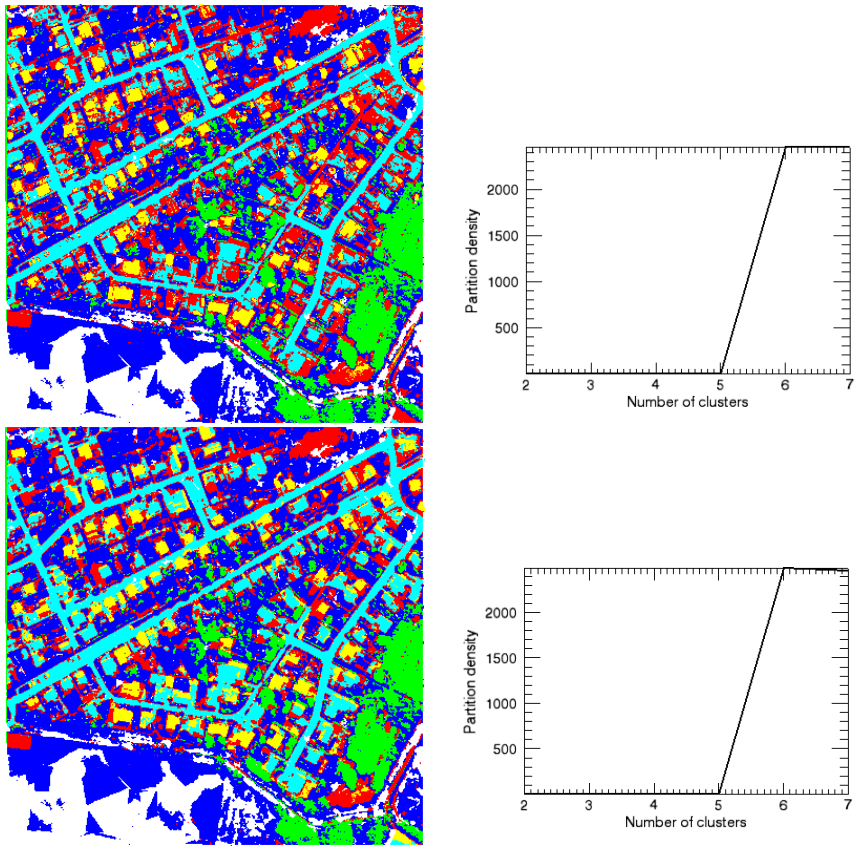


Figure D.6: FMLE classification of test area 2 at pyramid depth of 3 with plot of partition density, Upper: $\beta = 0$, Lower: $\beta = 1$.

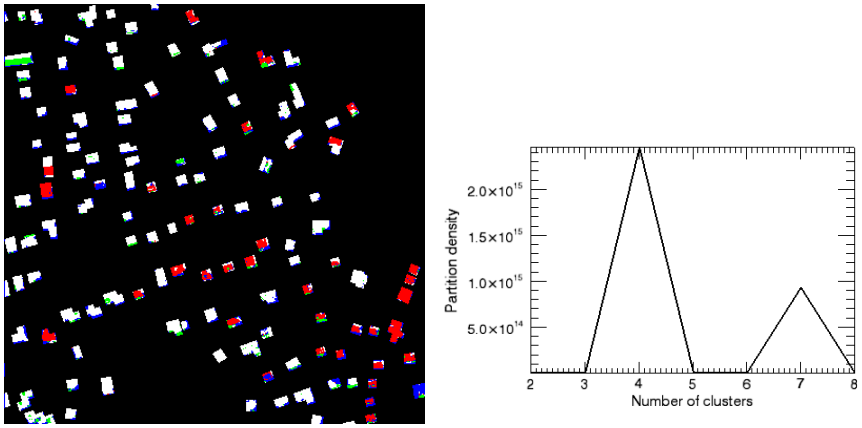


Figure D.7: FMLE classification of test area 1 with mask at pyramid depth of 2 and $\beta = 1$ and plot of partition density.

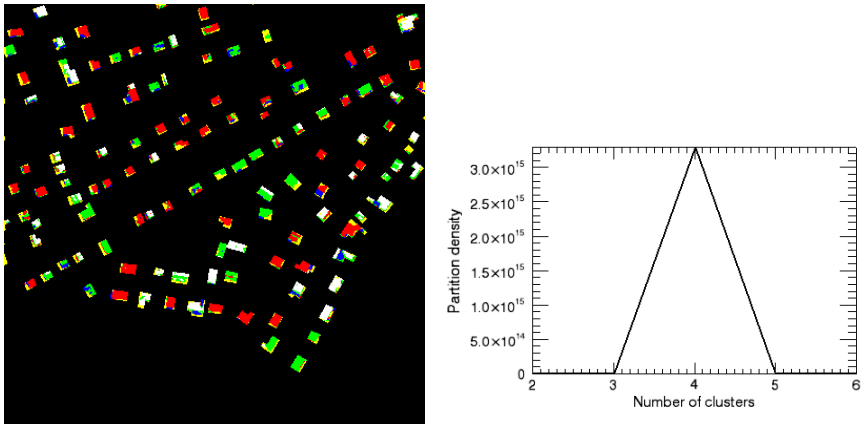


Figure D.8: FMLE classification of test area 2 with mask at pyramid depth of 2 and $\beta = 1$ and plot of partition density.

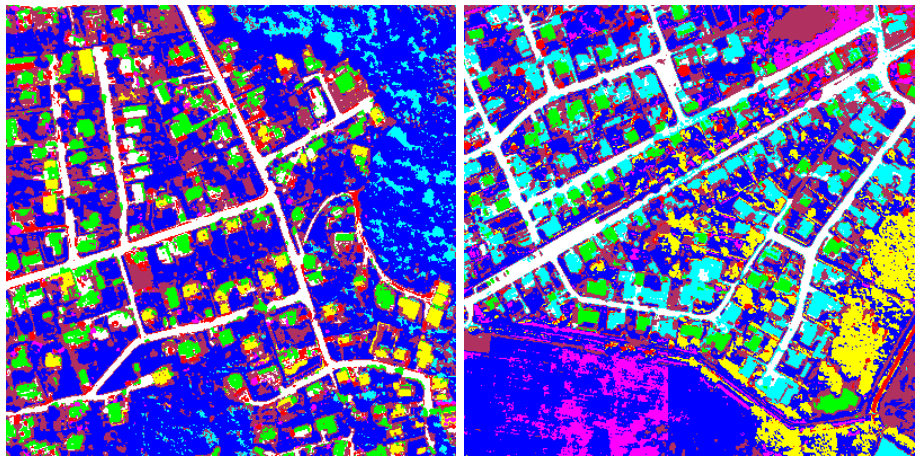


Figure D.9: FMLE classification of RGBNIR image with 8 clusters of test area 1 (left) and test area 2 (right).

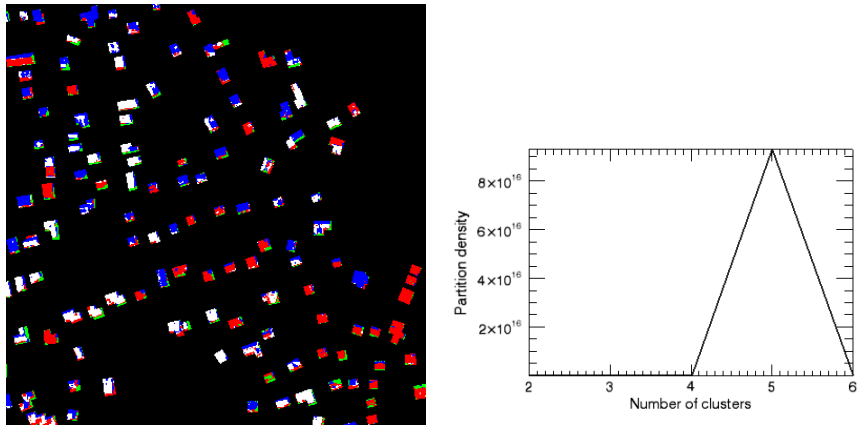


Figure D.10: FMLE classification of test area 1 with mask at pyramid depth of 2 and $\beta = 1$ and plot of partition density.

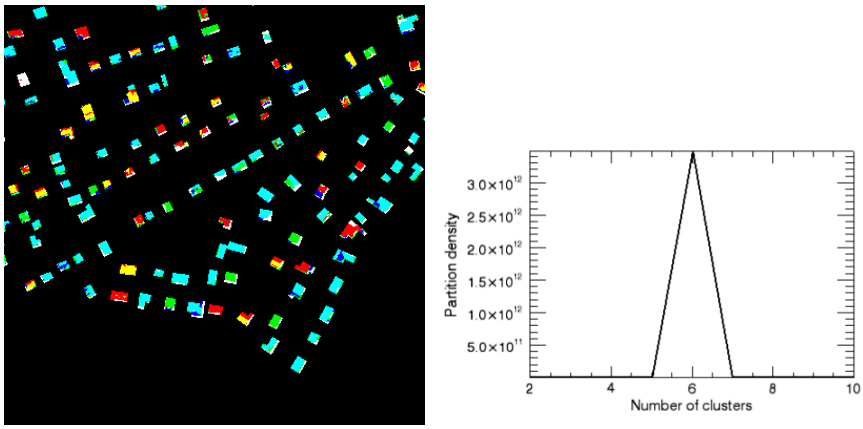


Figure D.11: FMLE classification of test area 2 with mask at pyramid depth of 2 and $\beta = 1$ and plot of partition density.

Bibliography

Balstrøm, T., Jacobi, O. and Bodum, L. [2006], *Bogen om GIS og geodata*, Forlaget GIS & geodata, København, Denmark.

Canty, M. J. [2007*a*], 'Dr. Mort J. Canty's homepage at Jülich Research Centre'. **URL:** *http://www.fz-juelich.de/ste/remote_sensing*

Canty, M. J. [2007*b*], *Image Analysis, Classification and Change Detection in Remote Sensing - with algorithms for ENVI/IDL*, CRC Press, Taylor & Francis Group, LLC, Boca Raton, Florida, USA.

Carstensen, J. M. [2002], *Image Analysis, Vision and Computer Graphics*, IMM, Technical University of Denmark, Kgs. Lyngby.

Duda, R. O. and Hart, P. E. [1973], *Pattern Classification and Scene Analysis*, John Wiley & Sons, Inc., New York, USA.

Eising, J. [1999], *Lineær algebra*, Institut for matematik, Technical University of Denmark, Kgs. Lyngby, Denmark.

Elachi, C. [1987], *Introduction to the Physics and Techniques of Remotes Sensing*, John Wiley & Sons, Inc.

Gath, I. and Geva, A. B. [1989], 'Unsupervised optimal fuzzy clustering', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7), 773–781.

Geosystems, L. [2002], 'ADS40 airborne digital sensor'.

GmbH, J.-O. [2006], 'Jena airborne scanner JAS 150'. **URL:** *http://www.jena-optronik.com/cps/rde/xchg/SID-26EE34DB-C7968293/optronik/hs.xsl/3919.htm*

Hilger, K. B. [2001], Exploratory analysis of multivariate data, PhD thesis, Technical University of Denmark. IMM-PHD-2001-89.

ITTVIS [2007], 'Web page of ITT Industries, Inc.'.
URL: *http://www.ittvis.com*

Jacobi, O. [1997], *Digital Kortlægning*, Institut for planlægning, DTU, Kgs. Lyngby, Denmark.

Jensen, T. B. S. [2002], Opdatering af topografiske kort - ændringsudpegning af bygninger ved brug af digitale overflademodeller fra laserscanning, Master's thesis, Technical University of Denmark.

Johnson, R. A. [2000], *Probability and statistics for engineers*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA.

Katzenbeisser, R. [2004], 'Calibration and data validation of lidar fiber scanner', *Presented at ASPRS Annual Conference 2004* .

KMS [2001], TOP10DK geometrisk registrering, specifikation udgave 3.2.0, Technical report, National Survey and Cadastre - Denmark.

KMS [2007], 'Homepage of the National Survey and Cadastre - Denmark'.
URL: *http://www.kms.dk*

Knudsen, T. and Nielsen, A. A. [2004], 'Detection of buildings through multivariate analysis of spectral, textural, and shape based features', *Proceedings on IGARSS-2004* .
URL: *http://geodesy.spacecenter.dk/∼thk/pubs/*

Knudsen, T. and Olsen, B. P. [2002], 'Detection of buildings in colour-infrared aerial photos for semi-automated revision of topographic databases', *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (3B), 120–125.
URL: *http://geodesy.spacecenter.dk/ thk/pubs/*

Leberl, F., Gruber, M., Ponticelli, M., Bernoegger, S. and Perko, R. [2003], The UltraCam large format aerial digital camera system, The American Society for Photogrammetry and Remote Sensing.
URL: *http://www.vexcel.com/products/photogram/ultracam/*

Li, S. Z. [2001], *Markov Random Field Modeling in Image Analysis*, 2 edn, Science Workbench, Springer.

Mathworks [2007], 'Homepage of The Mathworks, Inc.'.
URL: *http://www.mathworks.com*

Mikhail, E. M., Bethel, J. S. and McGlone, J. C. [2001], *Introduction to Modern Photogrammetry*, John Wiley & Sons, Inc., New York, USA.

Olsen, B. P. [2004], Maintenance of Digital Topographical Map Databases - Change Detection, PhD thesis, National Survey and Cadastre - Denmark.
URL: *http://geodesy.spacecenter.dk/~bpo/index.html*

Palubinskas, G. [1998], 'K-means clustering algorithm using the entropy', *SPIE (European Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing), September, Barcelona* **3500**, 63–71.

Richards, J. A. and Jia, X. [1999], *Remote Sensing Digital Image Analysis*, Springer.

TopoSys [2007], 'Web page of TopoSys, Topographische Systemdaten GmbH'.
URL: *http://www.toposys.com*

Vexcel [2004], 'UltraCam D, technical specifications'.
URL: *http://www.vexcel.com/products/photogram/ultracam/*

Vexcel [2005], 'UltraCam D, large format digital aerial camera'.
URL: *http://www.vexcel.com/products/photogram/ultracam/*

Vexcel [2006], 'UltraCam X, large format digital aerial camera'.
URL: *http://www.vexcel.com/products/photogram/ultracamx/*

Wolf, P. R. and Dewitt, B. A. [2000], *Elements of Photogrammetry with Applications in GIS*, The McGraw-Hill Companies, Inc., USA.