Classifying multisensor remote sensing data: Concepts, Algorithms and Applications

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch–Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Björn Waske

aus

Berlin

Bonn, Oktober 2007

1. Gutachter: Prof. Dr. G. Menz
2. Gutachter: Prof. Dr. S. Schmidtlein
3. Gutachter: Prof. Dr. C. Simmer
4. Gutachter: PD Dr.-Ing. K-H. Franke

Tag der mündlichen Prüfung: 14.12.2007

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert
Erscheinungsjahr: 2007

Acknowledgements

It is done – and I would like to express my gratitude to everybody who has a part in the successful completion of this work. There are undoubtedly lots of people who support me in one way or the other, and talking about everybody would be a "never ending story".

First of all I would like to thank the four members of the dissertation committee. Gunter Menz, as my main supervisor has always motivated and supported my work in a very generous way. He completely supports my ideas and the plan to staying aboard, which was certainly a fruitful experience. I'm pleased about Sebastian Schmidtlein's willingness to review my thesis.

As a team leader at the Center for Remote Sensing of Land Surfaces (ZFL), Matthias Braun has "kept things running". I am very appreciative to his support in the daily work of the thesis and the project, even though he has a tight schedule. He gives me numerous helpful comments to put the thesis in a fruitful context.

The research of the dissertation was partly inspired and based on a scientific visit to the Department of Electrical and Computer Engineering, University of Reykjavik. I'm indebted to Jón Atli Benediktsson for his great support, which made this visit possible. My visit was a valuable personal experience and a fruitful scientific cooperation.

I thank all my colleagues at the ZFL, University Bonn: Albert, Jan, Jonas, Roland, Vanessa, without all the discussions, at work and during coffee breaks the time would have been different; and not to forget Vanessa's support in fieldwork and other aspects in the Enviland project. Not to forget to thank, Ellen for catering and making things work.

I also enjoy keeping contact with Patrick Hostert (Humboldt-Universität Berlin) and look forward to collaborating with his group.

I definitely owe my thankfulness to Sebastian and Benjamin for close continuing friendship, manifold support and discussions. I am deeply indebted to Sebastian for a fruitful collaboration, for still answering my calls and e-mails, although he was more than busy; and not the least to Anna and Johann for their understanding. Special thanks also to Andreas, Calli and Tobi.

After all I express my deepest gratitude to the people, who have always supported me, and finally enabled me, to obtain my present situation: My family.

A major part of the research relevant for this dissertation was supported by the German Aerospace Center (DLR) and Federal Ministry of Economics and Technology (BMWi) under the project Enviland (FKZ 50EE0404). This work was supported in part by the German Research Foundation (DFG) under the Research Training Group 722 (Information Techniques for Precision Crop Protection) at the University of Bonn. The Envisat ASAR and ERS-2 data was provided by the European Space Agency through a CAT 1 proposal (C1.3115), the SPOT imagery through the European OASIS program (OASIS 58 - CE 6324).

Abstract

Today, a large quantity of the Earth's land surface has been affected by human induced land cover changes. Detailed knowledge of the land cover is elementary for several decision support and monitoring systems. Earth-observation (EO) systems have the potential to frequently provide information on land cover. Thus many land cover classifications are performed based on remotely sensed EO data. In this context, it has been shown that the performance of remote sensing applications is further improved by multisensor data sets, such as combinations of synthetic aperture radar (SAR) and multispectral imagery. The two systems operate in different wavelength domains and therefore provide different yet complementary information on land cover.

Considering the increase in revisit times and better spatial resolutions of recent and upcoming systems like TerraSAR-X (11 days; up to1 m), Radarsat-2 (24 days; up to 3 m), or RapidEye constellation (up to 1 day; 5 m), multisensor approaches become even more promising. However, these data sets with high spatial and temporal resolution might become very large and complex. Commonly used statistical pattern recognition methods are usually not appropriate for the classification of multisensor data sets. Hence, one of the greatest challenges in remote sensing might be the development of adequate concepts for classifying multisensor imagery.

The presented study aims at an adequate classification of multisensor data sets, including SAR data and multispectral images. Different conventional classifiers and recent developments are used, such as support vector machines (SVM) and random forests (RF), which are well known in the field of machine learning and pattern recognition. Furthermore, the impact of image segmentation on the classification accuracy is investigated and the value of a multilevel concept is discussed. To increase the performance of the algorithms in terms of classification accuracy, the concept of SVM is modified and combined with RF for optimized decision making.

The results clearly demonstrate that the use of multisensor imagery is worthwhile. Irrespective of the classification method used, classification accuracies increase by combining SAR and multispectral imagery. Nevertheless, SVM and RF are more adequate for classifying multisensor data sets and significantly outperform conventional classifier algorithms in terms of accuracy. The finally introduced multisensor-multilevel classification strategy, which is based on the sequential use of SVM and RF, outperforms all other approaches. The proposed concept achieves an accuracy of 84.9%. This is significantly higher than all single-source results and also better than those achieved on any other combination of data. Both aspects, i.e. the fusion of SAR and multispectral data as well as the integration of multiple segmentation scales, improve the results. Contrary to the high accuracy value by the proposed concept, the pixel-based classification on single-source data sets achieves a maximal accuracy of 65% (SAR) and 69.8% (multispectral) respectively. The findings and good performance of the presented strategy are underlined by the successful application of the approach to data sets from a second year. Based on the results from this work it can be concluded that the suggested strategy is particularly interesting with regard to recent and future satellite missions.

Table of contents

List of figures

List of tables

List of abbreviations

AMI	Active Microwave Instrument
ALOS	Advanced Land Observing Satellite
AP	Alternating Polarization (Envisat ASAR product)
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ASAR	Advanced Synthetic Aperture Radar
AVNIR-2	Advanced Visible and Near Infrared Radiometer type 2
ANN	Artifical Neural Network
BMWi	Bundesministeriums für Wirtschaft und Technologie (Federal Ministry of Economics and Technology)
CAP	Common Agricultural Policy
CBD	Convention on Biological Diversity
CORINE	Coordinated Information on the European Environment
CLC2000	Corine Land Cover, using year 2000 as reference
CNES	Centre National d'Etudes Spatiales
DFD	Deutsches Fernerkundungsdatenzentrum (German Remote Sensing Data Center)
DT	Decision Tree
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
EO	Eart-Observation
ERS	European Remote Sensing Satellite
ESA	European Space Agency
GEOSS	Global Earth Observation System of Systems
GMES	Global Monitoring of Environment and Security
GSE	GMES Service Element
HH	Horizontally transmitted and received
HV	Horizontally transmitted and vertically received
InVeKoS	Integriertes Verwaltungs und Kontrollsystem (Integrated Agricultural Administration and Controlling System)
IM	Image Mode (Envisat ASAR product)
JAXA	Japan Aerospace eXploration Agency
MARS	Monitoring Agriculture through Remote Sensing techniques
MLC	Maximum Likelihood Classifier
NASA	National Aeronautics and Space Administration
USGES	National Geological Surve
PALSAR	Phased Array type L-band Synthetic Aperture Radar
PRI	SAR precision image
rfs	Random Feature Selection
RF	Random Forests
SVM	Support Vector Machines
SAR	Synthetic Aperture Radar

SPOT	Systeme Pout l'Observation de la Terra
TM	Thematic Mapper
UBA	Umweltbundesamt (Federal Environmental Agency)
VV	Vertically transmitted and received

1 Introduction

Mankind is changing dramatically the global ecosystem, reasons such as the clearing of natural forests, intensifying farmland production and increasing urbanization (Foley *et al*. 2005). These impacts on the ecosystem were more dominant in the last 50 years than at any other time in human history (MA 2005). A large part of the environment is dominated by humanity and only 17% of the land surface is not directly impacted by human actions (Vitousek *et al*., 1997; Sanderson 2002). Between 1950 and 1980 more land was transformed to cropland than in the 150 years between 1700 and 1850 and today approximately 50% of Earths surface has been transformed for the cultivation of crops or to graze. More than half of the Earth's forests have disappeared by this land transformation (MA 2005).

Although modern agriculture techniques are increasing food productivity, this causes extensive environmental damage, which affects ecosystem services, including many that are important for agriculture productivity (Foley *et al*., 2005). In addition the increased demand for food has been achieved focusing efforts on planting and consuming a small variety of crops, barley, maize, rice and wheat occupied almost 40% of cropland (Tilman *et al.*, 1996). Vitousek *et al.* (1997) pointed out that land transformations are the driving force in the loss of biological diversity worldwide. Moreover these changes affect the Earth's ecosystems far beyond the boundaries of original impacted region and can influence the climate directly at local and even regional scales.

Thus, the detailed knowledge and information on land cover is an important input for decision support and environmental monitoring systems, e.g. in the area of flood forecast, subsidy control, land degradation, urban sprawl and food security. Earth-observation (EO) systems have the potential to provide spatially distributed and temporally frequent information on land cover and its environmental state over extended and remote areas. Compared to man-based field campaigns costs can be significantly reduced. Moreover satellite systems provide near-real time information, which seems particularly important for operational applications, e.g. in the context of natural hazards and disaster management.

Consequently, the manner of how our planet is observed was revolutionized during the last decades. Remote sensing became a valuable and powerful instrument to monitor the Earth and had a significant impact on the manner environmental data is acquired and analyzed (Rosenquist *et al.*, 2003). Beside scientific-driven research significance, EO-data plays a major part in supporting the operation of several multilateral environmental and political charters (Peter, 2004; Rosenquist *et al.*, 2003, Backhaus and Beule, 2005), as the Kyoto Protocol, the Convention on Biological Diversity (CBD), or the European initiatives Coordinated Information on the European Environment (CORINE Land Cover) and Global Monitoring for Environment and Security (GMES). Regarding the enormous resources of the EO-technology, Rosenquist *et al.* (2003) have mentioned that the observation of the land cover and the detection and spatial quantification of land cover change are major areas, where remote sensing imagery could be integrated to support the operation of the treaty and assist countries to meet the commitments under the Kyoto protocol. The monitoring of aforestation, deforestation and human induced land cover fragmentation, for example, is important regarding the Kyoto Protocol and CBD respectively.

The European MARS-project (Monitoring Agriculture through Remote Sensing techniques) of the Joint Research Center (JRC) integrates EO technologies to receive independent and timely information on crop areas and yields. Overall MARS provide a basis for the European capability for global food security assessment and agricultural monitoring (MARS, 2007). Another example of the relevance of remote sensing imagery is given by Europe's Common Agricultural Policy (CAP), which forces the member states to provide digital information on agricultural land use. The corresponding Integrated Agricultural Administration and Controlling System (i.e., InVeKoS, "*Intetrgriertes Verwaltungs und Kontrollsystem*") employs remotely sensed data for obligatory agricultural subsidy controls.

Driven by the Baveno manifesto in 1998, the European Commission, the European Space Agency (ESA) and national agencies have started the GMES program (Peter 2004; Backhaus and Beule 2005). GMES is aiming to provide the potential of remote sensing services operationally, to support Europe's aims in sustainable development, and environmental and security-related policies (Brachet 2004, Peter, 2004). The GMES project contains so-called service elements (GSE) as *GSE Forest Monitoring* (GSE-FM) and *GSE Land Information Services* (GSE Land). GSE Land is aimed to deliver geo-information services over wide areas and for various land applications, focusing on the GMES priorities "Land Cover Change in Europe", and "Environmental Stress in Europe". The information on land cover and vegetation is derived from remote sensing data and is harmonised and standardised enabling cross-border applications and comparisons. The information and services GSE-FM are designed to support more informed decisions, operational applications and improved policies that significantly enable sustainable forest management (GMES, 2007).

The European project CORINE Land Cover (CLC) is aimed at the provision of a unique and comparable data set of land cover for Europe. In Germany the project was led by the German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR) on behalf of the Federal Environmental Agency (UBA). The land cover mapping is performed on the basis of satellite remote sensing images on a scale of 1:100,000. CORINE Land Cover 2000 (CLC2000) is an update of an early data set from 1990 (CLC1990) and an acquisition of changes has been conduced using the year 2000 as reference. The recent update CLC2006, will be performed as a part of the GMES Fast track service on land monitoring.

Geoland, DeCover and Enviland are other selected examples of international and national projects in context of land cover monitoring. The primary objective of the German project Enviland, funded by the DLR and Federal Ministy of Economics and Technology (BMWi) is the development of novel and outstanding approaches for the generation of land cover products, utilising the synergetic usage of multitemporal SAR and optical data. The major is the development of scale-independent, stable and cost efficient methods for the generation of land cover products.

These various European and national projects underlines the relevance of land cover products in context of remote sensing, which was also confirmed by a study of German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety. The ministry has launched a study, to receive information on requirements of remote sensing, particularly in context of the GMES program (Backhaus and Beule, 2005). Regarding this study, which was completed in May 2004,

Backhaus and Beule (2005) have pointed out that the most frequently required products are focused on observation of land cover and land use.

Overall the derivation of land cover information from remote sensing imagery is an important application, regarding the support of multilateral environmental agreements, decision support and monitoring systems. The international initiative Global Earth Observation System of Systems (GEOSS) will coordinate the various requirements of different nations for temporal consistent global information as a basis for decision making. The vision is to initialize a development wherein decisions for the benefit of humankind will be supported by coordinated, comprehensive and sustainable EO-data and information, for an improved monitoring of the Earth and a better understanding of the environmental processes (GEO, 2005)

1.2 Scientific framework and scope of the dissertation

Today a variety of complementary remote sensing imagery is available, provided by different sensors that operate in different wavelengths, as synthetic aperture radar (SAR) and multispectral systems (Figure 1.2.1). The NASA's Landsat-5 TM and Landsat-7 ETM+, the ASTER sensor on the Terra platform, the French SPOT satellites or the AVNIR-2 on ALOS operated by JAXA are some examples of well-known multispectral systems. These passive optical EO systems will be supplemented by upcoming missions, as the German RapidEye. The RapidEye constellation, which will be consists of five satellites, is expected to be operationally available in early 2008.

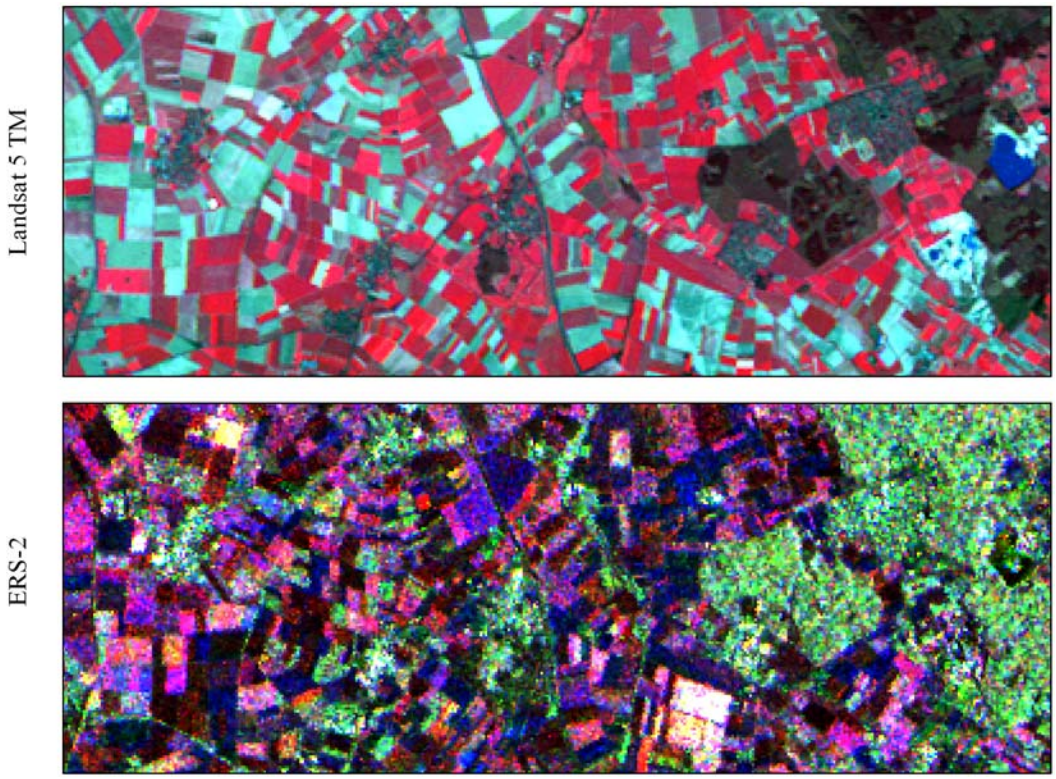

Figure 1.2.1: Landsat 5 TM image (April - bands: 4/3/2) and multitemporal SAR composite (Apr. / May / Jun.) of an agricultural region near Bonn, Germany.

The European ERS-2 and Envisat ASAR instruments as well as the Canadian Radarsat-1 are providing SAR imagery since several years. The availability of SAR data is further increased with recently launched and upcoming systems as the German TerraSAR-X, the Japanese ALOS PALSAR, the Italian Cosmo-SkyMed or the enhanced system Radarsat-2.

Multispectral and SAR systems operate in different wavelengths, ranging from visible to microwave. The interaction with the illuminated target and the amount of radiation that is reflected by the land surface varies with the wavelength. These two systems consequently provide different, but complementary information improving the monitoring of land cover. In Figure 1.2.1 a multitemporal SAR composite and a multispectral image is shown. A simple visible assessment illustrates the different nature of the data types and the dissimilar information content. Beside this, various imagery from a multitude of different SAR systems are available, which operate in different wavelength and hence provide diverse information, e.g., Envisat ASAR and ERS-2 acquired C-band data (~5.3 cm), ALOS PALSAR operates in L-Band (~23 cm) and TerraSAR-X in the short X-band (~3.1 cm) domain.

Thus, a combination of different image sources, e.g. multispectral imagery with SAR data or even different SAR imagery alone, seems worthwhile and it has been shown in several studies that multisensor analysis significantly improves the performance of land cover classifications (see Section 2). Regarding the growing data diversity, increased revisit times and enhanced spatial resolutions of upcoming and recent missions, as for example TerraSAR-X (11 days; up to1 m), Radarsat-2 (24 days; up to 3 m), or RapidEye (up to 1 day; 5 m), concepts of multisource image analysis become even more attractive. On the other hand such diverse data sets result in the demand of adequate classifier concepts, which enable the handling of diverse information, of different spectral, spatial, and temporal resolution.

Driven by the rapid development of remote sensing applications, the community has attained a step, where a multitude of adequate and widely known classifier concepts are available. Today the analysts can choose between diverse remote sensing imagery as well as a number of widely used methods. Nevertheless, the development of adequate concepts and algorithms for classifying multisource information is perhaps the most challenging research topic in the field of remote sensing (Richards, 2005).

Data sets of upcoming missions with enhanced spatial and temporal resolution might become very large and complex. Furthermore individual sources (e.g., a specific SAR acquisitions or a specific multispectral band) may not be equally reliable. One source can be more adequate to separate a specific land cover class and perhaps another source is more applicable to describe another class. Thus, it might be appropriate to weight different sources during the classification process, but conventional statistical classifiers do not allow such weighting. Moreover the derivation of spatial (e.g., texture) and temporal information might be useful to increase the classification accuracy, particularly in regard to enhanced spatial resolution and increased orbit repetition rates. On the other hand such additional derived information contributes to an increased complexity of the data set.

Another development in the classification of remote sensing imagery is that of segment-based applications. Often segment-based classifications are more accurate than conventional per-pixel approaches and reduce effects, which occur in pixel-based classifications. They seem particularly interesting for agricultural areas that are dominated by typical spatial patterns of

planted crops. Segment-based approaches, on the other hand, require the definition of an adequate segmentation scale that might be critical. In several studies this problem is solved by multilevel approaches, which employ various segmentation levels.

Thus, regarding complex multisensor and multitemporal data sets, perhaps by different segmentation scales, the imagery should be processed by adequate classifiers, which can handle such diverse data sets and makes best use of the relevant information. The increased size of the input data sets for example, demands a classifier, which is less sensitive to the curse of dimensionality (i.e., Hughes effect). On the other hand the classification of segmented imagery that is often limited to the few classifiers available in proprietary software. Moreover, widely used parametric classifiers that rely on certain conditions, e.g. a maximum likelihood classifier that often assumes a Gaussian distribution, can often not be applied to segmented data with a reduced feature space and a small number of training samples. In addition the class distributions generally cannot be modeled by an adequate multivariate statistical model, when classifying multisensor imagery.

Overall, it seems more appropriate to employ techniques, which were successfully applied to more complex data sets, as high-dimensional imagery and multisource data sets. The objective of the presented dissertation addresses the development of concepts and algorithms for an efficient image analysis of multisensor data, consisting of multitemporal SAR data and multispectral imagery:

In regard to this, the following main questions seem worth to investigate in the context of this dissertation:

- Does multisensor imagery generally improve the quality of land cover classifications?
- How do recent classifier developments perform on multitemporal and multisensor remote sensing data?
- What impact has image segmentation on classifying multisensor imagery?
- Can sophisticated classifier algorithms be further improved?

1.3 Outline of the dissertation

In Section 2 a general guide to land cover classifications is given, with a focus on multitemporal and multisource applications. In this context different classifier algorithms (e.g., maximum likelihood classifier and support vector machines) are introduced in detail. In addition, concepts for combing classifiers and segment based approaches are discussed. In Section 3 the study site and relevant data sources (i.e., remote sensing imagery and auxiliary data) are described. Afterwards the obligatory image preprocessing is briefly illustrated in Section 4. Different applications and results of multisensor image analysis are presented in Section 5. These applications are focusing on (1) the classification of multitemporal SAR data, (2) an adequate fusion of multisensor imagery, (3) the introduction of a multilevel-multisensor classification strategy, and (4) the transfer of the final classification strategy to another data set (Figure 1.3.1). Section 6 discusses results and concludes in a synopsis.

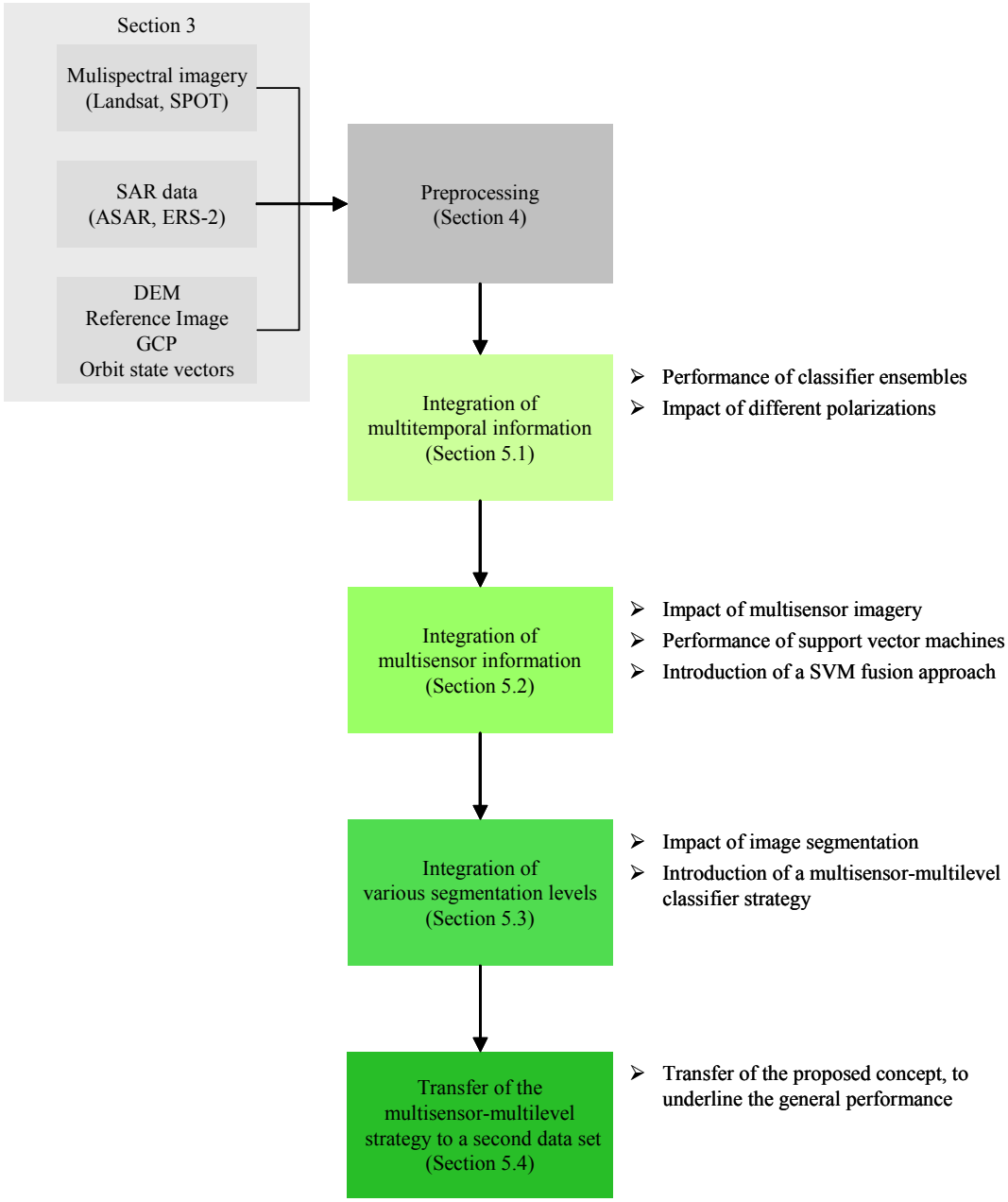

Figure 1.3.1: Overview on selected components of the dissertation. The green boxes indicate 4 different applications, with the corresponding main research questions.

2 Classifier Algorithms and Concepts

Land cover classifications are the widest used application in the field of remote sensing. Since satellite imagery is available the research community is aiming to produce precise classification maps of the data. Consequently the available literature is numerous and a significant quantity of published papers is considering the field of image classification (Wilkinson, 2005). Although the number of studies increased during the recent decades, remote sensing image analysis is not a new research subject and Richards (2005) remarked that the classification of remote sensing imagery had "its genesis in the signal processing methods of the 1950s and 1960s and their extension to handling image data".

The recent development of algorithms for land cover classification was primarily affected by three driving factors, which are closely interrelated: (1) an increased availability of diverse remote sensing imagery, (2) the enhanced computer power, and (3) a growing demand of more powerful machine learning algorithms and hence a shift from statistical methods to complex and more sophisticated methods for image classification (Richards, 2005, Jain *et al*., 2000). Beside these progresses, the use of object- or segment-based classification approaches is another development in the context of remote sensing image analysis.

Richards (2005) declares the sensor development as one driving factor for the progress in image analysis algorithms: The technical enhancement of optical EO-systems, with an increased number of bands and high spatial resolution as well as the availability of multidimensional SAR data at different wavelengths and polarizations. Beside the technical progress of the sensor systems, the rapidly increasing computer power enables faster processing of huge and diverse datasets. In addition complex data and increased performance requirements like speed and accuracy demand the development of more sophisticated classifier concepts (Jain *et al*., 2000). In contrast to this, many early classification methods were taken directly from signal processing, due to previous data limitations. Thus they were often based on simple data models and approaches (Richards, 2005).

The general aim of land cover classification is the association of each pixel within the imagery with a specific land cover class. Generally this is performed by methods of machine learning and pattern recognition. A pattern can be understood as a unique structure or attribute, which is capable to describe a specific phenomenon (e.g., a land cover type). During a classification process unknown patterns (i.e., pixels) are identified and assigned to a predefined class or they are combined to unknown clusters, depending on their similarity. In the context of remote sensing, patterns are often referred to as signatures. In the case of multispectral or hyperspectral imagery for example, the set of spectral radiances measured in the different bands can be regarded as a signature. Beside simple pixel values (i.e., spectral radiances or backscatter intensities), spatial and/or temporal image information can be used (Lillesand and Kiefer, 2000). Using texture analysis, the spatial relationships between a pixel and the neighborhood can be extracted and considered in the subsequent classification process (e.g., Haralick *et al*., 1973; Soares *et al*., 1997). The availability of image time series enables the derivation of temporal information, as multitemporal mean and variance, which can be included in the subsequent image analysis (e.g., Bruzzone *et al*., 2004).

In general the numerous classification methods can be grouped – depending on the available information – into the main categories unsupervised and supervised: In unsupervised

classifications pixels are assigned to specific classes without having a prior knowledge of the classes. During an unsupervised classification process the data is aggregated into natural groups or clusters that have similar properties. Often conventional clustering methods are used like the widely used ISODATA and the k-means algorithm (Richards and Jia, 1999; Duda *et al*., 2000.). A general problem of unsupervised algorithms is that data can be consisting of clusters with different shapes and sizes. An applicable definition of clusters and the selection of an adequate indicator for similarity are difficult (Jain *et al*., 2000). The advantage of unsupervised methods is that an image classification is possible without having a priori knowledge. On the other hand the algorithm is not providing any final membership decision and in general the analyst must assign each cluster a specific class label. Furthermore class memberships (i.e., class labels) and other knowledge about the image source do not impact the clustering algorithm, solely the interpretation of the final clusters (Jain *et al*., 2000).

Unlike unsupervised methods, supervised approaches are based on a priori knowledge of the land cover classes. Generally each class is represented by a sample set, which is used for the training of the classifier. The trained classifier is applied to the unknown pixels to determine the final class membership. A widely used classification strategy is the maximum likelihood classifier. This parametric classifier is based on the assumption that the probability density function for each class is multivariate, and often a Gaussian distribution is assumed. In contrast to parametric approaches, nonparametric methods are not constrained to any assumptions on the distribution of input data. Hence techniques as artificial neural networks, decision trees and support vector machines can be applied, even if the class conditional densities are not known or can not be estimated.

Generally all these classifier methods produce a conventional land cover map that consists of discrete categories. Each pixel is associated to a class, but only one single land cover class. In contrast to these hard classifications, a pixel can be assigned to different classes, depending on the class-specific likelihoods. The discriminant function can be assumed as imprecise or fuzzy and is leading to a soft or fuzzy classification map. Another dichotomy in context of remote sensing image classification is that of pixel-based versus segment-based methods. In segment-based approaches neighboring pixels with similar properties are combined to image segments. After the segmentation process, information on the segments' mean spectral or backscatter value, their texture and shape as well as spatial relationships can be derived and included in the classification process. In hybrid methods a pixel-based (pre-)classification is performed. Afterwards a simple majority vote is performed within each segment, to determine the final class membership of each image segment. In analogy to other studies, in this dissertation the terminology segment-based is used, contrary to the expression object-based classification. In doing so the term object is referred to real natural features, as a field plot or a settlement, which appear as a single (image)segment in the data.

Whereas dichotomies as (un)supervised and (non-)parametric define some requirements of the classifier algorithms itself, expressions as pixel- and segment-based describe the general classification concept. For example, a parametric maximum likelihood classifier, which is generally used for pixel-based hard classifications, can be applied to segmented imagery (e.g. Geneletti and Gorte, 2003) or extended to a fuzzy classification concept (e.g. Schowengerdt, 1996).

2.1 Multitmeporal, Multisensor and Multisource Applications

Many remote sensing applications include data from areas, which are characterized by great temporal variability and typical spatial patterns of highly frequent land cover changes between vegetation canopies. Single-date applications are often inefficient, due to the phenology of planted crops. Whereas a specific data might not be appropriate to distinguish specific land cover classes, another acquisition date might be inefficient for other classes. In contrast to a single-date acquisition, the use of the whole multitemporal information enables a better differentiation between several classes (Figure 2.1.1)

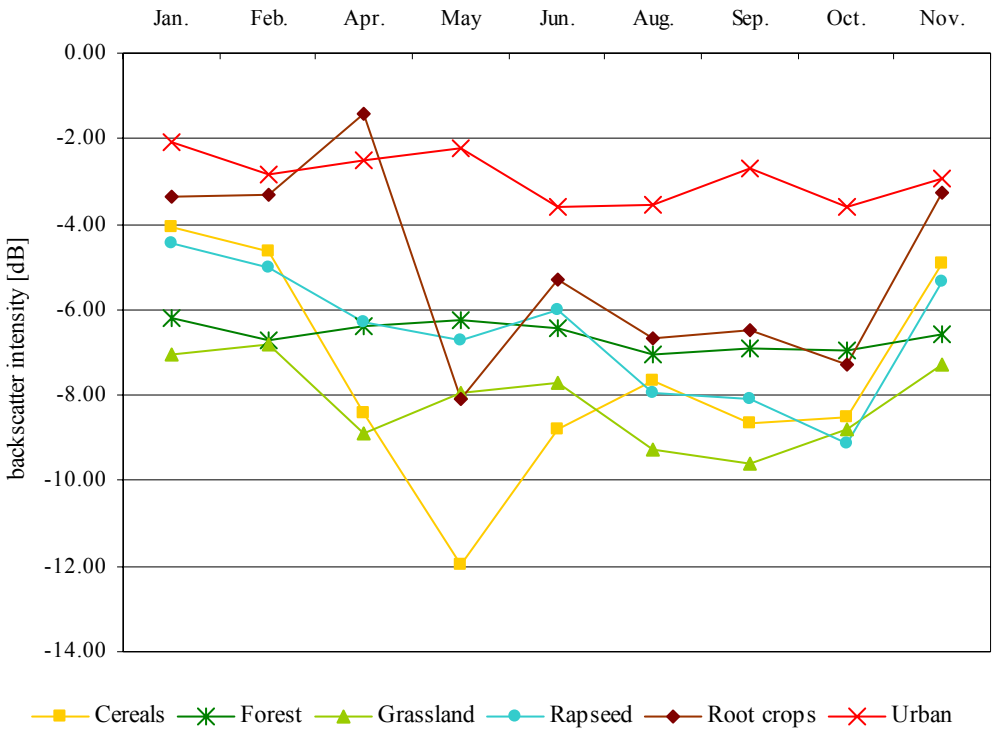

Figure 2.1.1: Comparison between multitemporal characteristics of different land cover classes, collected at a test site near Bonn in 2005 (ERS-2, C-Band, VV polarization)

This fact is confirmed by the results of several studies, where multitemporal imagery significantly improves the results achieved by monotemporal applications: In Guerschman *et al.* (2003) the impact of the number of acquisitions of a Landsat TM time series on the classification accuracy was investigated. The dimension of the time series was successively increased from 1 to 4 acquisitions, considering all potential image combinations. The results show that the accuracy is significantly improved with an increasing size of the data set. Whereas the accuracies of the 4 three-date classifications vary between 49-56%, the overall accuracy is improved up to 63%, using all images. Blaes *et al.* (2005) have used a set of 2 Spot XS and a Landsat ETM+ image for crop identifications. The classification of the full image set achieves an overall accuracy of 74%, whereas the accuracies achieved by a mono-temporal approach vary between 22% and 64%.

In another study a crop classification was increased by 26% respectively, using two Landsat TM scenes, in contrast of one acqusition (Brisco and Brown, 1995). However, these studies include multispectral remote sensing data, but the availability and efficiency of optical imagery is often limited by illumination and weather conditions, particularly in regions of Central Europe. Multitemporal datasets within one growing season can be best produced reliably, using synthetic aperture radar (SAR). This EO-system is almost independent from weather conditions, thus SAR data is particularly interesting for near-real time applications and operational monitoring systems. Considering missions with high revisit times and better spatial resolutions like TerraSAR-X and Cosmo-SkyMed, SAR-based multitemporal approaches become even more attractive. Even the classification and interpretations of SAR data seems often more difficult than those of multispectral imagery, several studies are based on SAR data and assessed the positive impact of multitemporal imagery on the classification accuracy: In Brisco and Brown (1995) the overall accuracy of a classification that is based on 4 SAR acquisitions was increased up to 24%, compared to results achieved with two images. Blaes at al. (2005) have compared the performance of various multitemporal ERS data sets. The classification accuracies were significantly increased by increasing the quantity of images. Depending on the number of scenes and the acquisition date, the overall accuracies vary between 40% and 65%. Similar results have been reported from Chust *et al*. (2004), which classified ERS data from a Mediterranean region. In addition these results show that the variance of the overall accuracy is decreased with an increasing number of images.

As mentioned the availability of multisensor imagery is increasing and in several studies the classification accuracies were increased by multisource applications (e.g., Benediktsson and Kanellopoulos , 1999; Blaes *et al*, 2005; Michelson *et al*., 2000; Le Hégarat-Mascle *et al*. 2000; Chust *et al*. 2004; Huang 2007). Brisco and Brown (1995) significantly improved the classification accuracy of two Landsat TM images by additional multi-date SAR data. In Blaes at al. (2005) Landsat imagery was combined with multitemporal ERS-2 data. Thereby total accuracy was increased at least by 5%. Huang *et al*. (2007) fused Radarsat and Landsat imagery to enhance a land cover classification. Whereas the use of two Landsat images results in a total accuracy of 74%, a combined approach increases the accuracy up to 84%.

Although some studies employ conventional methods like the well known maximum likelihood classifier (e.g., Brisco and Brown, 1995; Chust *et al*., 2004; Huang *et al*., 2007) such widely used statistical classifiers are often not optimal for classifying multisource and multitemporal data sets. In most cases the class distributions cannot be modeled by adequate multivariate statistical models (Benediktsson *et al*, 1990, Bruzzone at al., 2004). Consequently other studies use more sophisticated classification strategies and nonparametric algorithms:

Artificial neural networks (ANN) are a nonparametric method, which have been used successfully for the classification of diverse remote sensing imagery. The overall accuracies are often significantly improved compared to conventional statistical classifiers: Benediktsson *et al*. (1990) have used a backpropagation ANN for the classification of multisource data sets, containing multispectral data and topographical information. In other experiments ANN were used for the classification of time series of SAR imagery (e.g., Stakiewicz, 2006; Bruzzone at al., 2004) and multisensor data (e.g., Serpico and Roli, 1995; Bruzzone *et al*., 1999).

Besides several classifications that are based on multispectral imagery (e.g., Friedl and Brodely, 1997; Pal and Mather, 2003) and SAR data (Simard *et al*., 2000, 2002), decision tree classifiers (DT) are also used for and multisource and multitemporal applications. Fitzgerald and Lees (1994) applied DT to multisource data, including remote sensing imagery and topographical data. Another study aims on mapping multitemporal vegetation changes (Rogan *et al*. 2002). The results demonstrate that a DT is adequate for this purpose and outperforms a maximum likelihood classifier in terms of accuracy. In Hodgonson *et al*. (2003) a decision tree classifier is used for mapping impervious surfaces, using LIDAR data and air-borne imagery.

In several studies the performance of DT is further increased by classifier ensembles or multiple classifier systems: Brown de Colstoun applied a decision tree on multitemporal images from the Enhanced Thematic Mapper-Plus (ETM+) to differentiate between 11 land cover types. The overall accuracy was significantly increased by classifier ensembles techniques, as boosting. Carreiras *et al*. (2006) performed a classification of agricultural and pastures land within the Brazilian Amazon using a time series of SPOT 4 Vegetation data. In this study a DT-based classifier ensemble significantly outperforms all other approaches (i.e., maximum likelihood classifier, simple decision tree and k-nearest neighbor) in terms of overall accuracy. Briem *et al*. (2002) successfully used various classifier ensembles for classifying different multisource data sets, including SAR and multispectral imagery among other data types. Breimans' classifier system *random forests* (2001) was used in diverse remote sensing studies (Ham *et al*. 2005, Lawrence *et al*. 2006, Gislason *et al*. 2006, Pal 2004). The results in Ham *et al*. (2005) assessed a good performance of the random forests (RF) for classifying hyperspectral data with a limited sample set. In Gislason *et al*. (2006) the RF method was applied to a multisource data set, consisting of Landsat MSS data and topographical data. RF perform better than a single decision tree and comparable to other ensembles methods, whereas the computation time is much faster. Pal (2005) has used the approach for the classification of a Landsat ETM+ scene from an agricultural region. In this study the technique achieved promising results and the accuracies were comparable to computationally more complex methods like support vector machines.

In contrast to these classifier ensembles, which combine variants of the same classifier, other concepts are based on the combination of different algorithms, from now on referred to as multiple classifier systems: Steele (2000) combined a spatial classifier with a *k*-nearest neighbour and a linear discriminant algorithm, for land cover mapping of Landsat TM imagery. Liu *et al*. (2004) used a decision tree in combination with a neural network for the classification of NOAA AVHRR data. Benediktsson and Kanellopoulos (1999) combine neural networks and statistical methods for classifying multisensor data. The SAR and multispectral data undergo a separate classification. The outputs of the individual classifiers are combined by decision fusion. In Jeon and Landgrebe (1999) the concept of decision fusion was used for classifying a set of multitemporal Landsat images. Fauvel *et al*. (2006, 2007) used the strategy to increase the classification accuracy of remote sensing imagery for an urban area.

Support vector machines (SVM) are well known in the field of machine learning and pattern recognition (Vapnik, 1998; Schölkopf and Smola, 2002) and recently introduced in context of remote sensing image analysis (Huang *et al*., 2002; Foody and Mathur, 2004; Melgani and Bruzzone, 2004). The approach still exhibits further modification, e.g., in context of semi-supervised and ill-posed classification problems (Bazi and Melgani, 2006; Bruzzone *et al*.,

2006; Chi and Bruzzone, 2007). In Huang *et al*. (2002) SVM were applied successfully for classifying multispectral imagery and outperformed other methods in the very most recent cases. In Foody and Mathur (2004) a multispectral image was classified, focusing on agricultural classes, as for example wheat, sugar beet, carrots, grassland. The results clearly demonstrate that SVM perform more accurately than other classification techniques (e.g., DT and ANN). Beside a Landsat image Pal and Mather (2006) applied SVM on DAIS data from a Mediterranean region, which is predominantly used for agriculture. As in other studies the SVM produce higher accuracies than other classifiers, as a maximum likelihood classifier, an ANN and a simple decision tree. Overall it can be assessed that SVM perform well with a small number of training samples, even when applied on high-dimensional data sets (e.g., Melgani and Bruzzone, 2004, Pal and Mather , 2006). In contrast to these classifications, which are based on optical imagery, in other studies SVM were applied successfully on SAR data (Lardeux *et al*., 2006, Fukuda and Hirosawa, 2001).

Although SVM have shown promising accuracy on a multitude of different data sets, only a few studies are known which use SVM for classifying multisource or multitemporal data: In Song *et al*. (2005b), SVM were used for classifying geospatial data, consisting of multispectral images and topographical information among other data types. Halldorsson *et al*. (2003) extended a common SVM kernel function for classifying a multisource data set, containing Landsat MSS data and topographical information. In Camps-Valls *et al*. (2006) composite kernels were introduced for combining spectral and spatial information of a hyperspectral image. Fauvel *et al*. (2006) combined SVM to fuse spectral and spatial information (i.e., extended morphological profiles) of a hyperspectral ROSIS-03 data. At first the two SVM classifiers were trained separately on the spectral image on one hand and the extended morphological profiles on the other. Afterwards the outputs were fused by using different voting schemes, e.g., absolute maximum and majority voting.

Beside this shift to recent machine learning algorithms as classifier ensembles and support vector machines, segment-based approaches are another development in land cover classifications (e.g., Lobo, 1996; Tso and Mather, 1999; Marcal *et al*., 2005). In several of these studies segment-based classifications are more accurate than conventional per-pixel classifications (Tso and Mather, 1999; Smith and Fuller, 2001; Geneletti and Gorte, 2003; Lee and Warner, 2006). Whereas some studies use common statistical classifiers (e.g., Genelitti and Gorte, 2003) other segment-based classifications are based on non-parametric algorithms: Llyod *et al*. (2004) used ANN for a segment-based classification approach. Lailberte *et al*. (2007) applied a DT-classifier on segmented mulitspectral imagery. In other studies classification approaches were presented that are based on support vector machines (Marcal *et al*. 2005, Bruzzone and Carlin, 2006; van der Linden *et al*. 2007).

Contrary to several studies, which are based on a single segmentation level (e.g., Tso and Mather, 1999; Smith and Fuller, 2001; Marcal *et al*., 2005) other approaches use various levels of aggregation: Shackelford and Davis (2003) successfully combined information from pixel- and segment-level to differentiate classes that appear similar at pixel level, using a fuzzy logic classifier. Lailberte *et al*. (2007) use a fine aggregation level to classify a specific land cover class. Afterwards the class was masked out and another, coarser aggregation scale was used to differentiate other land cover classes.

However, the definition of adequate aggregation levels is critical, particularly for classifying heterogeneous imagery (e.g., van der Linden *et al*. 2007). Song *et al*. (2005a) discussed how an inaccurate segmentation reduced the classification performance. The total accuracy decreases by errors in segmentation and classification. Moreover a misclassification of a single segment results in a misclassification of all pixels included by this segment. Moreover two different natural objects (i.e., two field plots with different land cover classes) can be merged to one image segment by an imprecise segmentation. This results in a misclassification of (at least) one of these natural objects. Bruzzone and Carlin (2006) discussed the simultaneous use of different segmentation results at various scales as a feature-extraction module that adaptively models the spatial context of each pixel. They show that different levels contribute different types of information. Hence multilevel approaches (van der Linden *et al*. 2007, Bruzzone and Carlin, 2006) seem more effective.

2.2 Supervised Algorithms

The classification framework of this dissertation belongs to the group of supervised methods and various parametric and non-parametric techniques have been introduced in the literature. Supervised classifiers require a priori knowledge and each class is usually described by a training data set. In general the training data is extracted from the imagery, using expert knowledge and ground truth knowledge as reference information. This class-specific information and data structures, derived from the training samples, are used within the training process of the classifier.

A simple method is the minimum distance to mean (or minimum distance) classifier. In the initial phase of the classification procedure, the training data is used to calculate a class-specific mean vector, containing the averaged values for each image feature (e.g., spectral radiances within each band). For the classification of unknown pixels the distance between pixel and the mean vector is calculated and a pixel is assigned to the closest class (i.e., the minimum distance). Although the approach is computationally simple it has several disadvantages. The main limitation is that the inter class variances are not considered. Hence the approach is inefficient for classifying more complex class distributions (Lillesand and Kiefer, 2000).

In other concepts the data is assumed to have class-specific probability density functions. Hence a sample x belonging to class ω is assumed as a sample that is randomly selected from the class conditional probability function. A widely used decision strategy is the Bayes' decision rule, whereas the well-known maximum likelihood classifier (MLC) is a particularly case of the Bayes rules (Section 2.2.1). The MLC is derived from the Bayes rule when classes have equal priorities. It is based on the assumption that the probability density function for each class is multivariate, and often a Gaussian distribution is assumed (Lillesand and Kiefer, 2000; Richards and Jia, 1999).

Although methods like the MLC are widely used, they are often not appropriate for the classification of multisource and multitemporal data, because in most cases such data sets cannot be modeled by a convenient multivariate statistical model (Benediktsson *et al*., 1990; Bruzzone *et al*., 2004). Furthermore, individual data sources may not be equally applicable. One source – for example a specific band of a multispectral image or an specific acquisition of a multitemporal SAR data set – can be more reliable to describe a land cover class than other

imagery and perhaps another piece of information is more adequate for classifying another class type. Thus, it can be appropriate to weight the different image sources during the classification procedure. However, common statistical classification techniques do not enable such weighting processes. For these reasons, other methods are more adequate in the context of multisensor and multitemporal image analysis and several non-parametric approaches have been introduced, e.g., *k*-nearest neighbor, artificial neural networks, self-learning decision trees and support vector machines.

These approaches enable the derivation of the decision boundary without having knowledge of the class conditionals densities, by estimating the density functions or directly estimating a posteriori probabilities directly. The nearest-neighbor classifier allows the direct generation of the decision boundary on the training data, without knowing class-conditional densities or without estimating the densities functions (Jain *et al*. 2000). It's classifying an unknown pixel x to the corresponding class ω of the sample nearest to x. The *k*-nearest-neighbor rule is an extension of the single-neighbor classifier, assigning a pixel to the most frequently class among the k nearest samples. As the minimum distance classifier the method is based on metric to measure the distance between the sample patterns, and often the Euclidean metric is used. Although in some studies nearest neighbor methods are superior or comparable to other classifiers (Collins *et al*., 2004; Carreiras *et al*., 2006) more sophisticated algorithms generally achieves higher accuracies (Marcal *et al*., 2005; Debeir, 2002, Roli and Fumera, 2001). Furthermore the method needs to compute the distances between an unknown pixel and all samples in the training set. This requires that all data are stored in the memory, resulting in a computational more demanding approach.

Artificial neural networks (ANN) are a nonparametric method, which have been used successfully for the classification of diverse remote sensing imagery. The overall accuracies are often significantly compared to conventional statistical classifiers: Benediktsson *et al*. (1990) have used a backpropagation ANN for the classification of multisource data sets, containing multispectral data and topographical information. In other experiments ANN were used for the classification of multitemporal SAR data (e.g., Stakiewicz, 2006; Bruzzone *at al*., 2004) and multisensor imagery (e.g., Serpico and F. Roli, 1995). Stakiewicz (2006), for example, classified a timer series of ENVISAT ASAR and ERS-2 imagery of an agricultural region. Beside the long training time, neural networks have no consistent rules for the network design and their performance is affecting by several factors, e.g. the network architecture (Foody and Arora 1997), which is dependent on the user. A general introduction to neural networks is given by Bishop (1995). An overview in context of remote sensing is given by Kanellopoulos and Wilkinson (1997) and Benediktsson *et al*. (1990).

Decision trees (DT) are another non-parametric classifier that was applied to diverse remote sensing data sets (Friedl and Brodley, 1997; Simard *et al*., 2000; Briem *et al*., 2002; Pal and Mather, 2003; Laliberte *et al*., 2007). In contrast to other classifier algorithms, which use the whole features space at once and makes a single membership decision – a decision tree is based on a multistage or hierarchical concept. At each node the most relevant feature only is selected and used for the construction of the decision boundary. The handling of DT is rather simple and their training time is relatively low compared to computationally complex approaches as for example neural networks (Friedl and Brodley, 1997; Pal and Mather, 2003).

Beside this, their visible classification scheme allows a direct interpretation of the decision in regard to the impact of individual features.

The performance of DT is increased by classifier ensembles or multiple classifier systems (Gislason *et al*. 2004, Briem *et al*., 2002; Carreiras *et al*., 2006; Pal, 2005). By training a classifier on resampled input data (i.e., features or samples) a set of independent classifiers is generated. Afterwards the different outputs are combined to create the final result. This concept is not restricted to decision trees and classifier ensembles that are based on other algorithms have been introduced (Hansen and Salamon, 1990; Kim *et al*., 2003). Nevertheless, DT classifiers are particularly interesting due to their simple handling and fast training time. In contrast to these classifier systems, which are based on the same classifier, in other approaches different algorithms are combined (Benediktsson and Kanellopoulos, 1999; Steel, 2000; Fauvel *et al*., 2006). In doing so, the different advantages of the algorithms can be joint.

Support vector machines (SVM), which are well known in the field of machine learning and pattern recognition (Vapnik, 1998), are a recent development in the context of remote sensing image analysis (Huang *et al*., 2002, Foody and Mathur, 2004a, Pal and Mather, 2005,). The concept is based on fitting an optimal seperating hyperplane to the training samples of two classes, such that the pixels from each class are finally on the correct side of the hyperplane. The position of the hyperplane is determined only by the closest training samples of both classes. This is one reason why support vector machines perform well, even with a small number of available training samples (Huang *et al*., 2000; Melgani and Bruzzone, 2004, Pal and Mather, 2006). Foody and Mathur (2004b) pointed out that an entire description of each class is not obligatory for an accurate classification. Because only samples near the hyperplane are considered, other training data have no impact on the analysis. Nevertheless a larger number of training samples ensures the employment of adequate samples Foody and Mathur (2004a).

Even the framework of this dissertation is focused on land cover classification, a consideration of all algorithms and concepts would be beyond the scope of this work. Methods as unsupervised algorithms or fuzzy approaches are not considered in this work. The presented study is only focusing on supervised algorithms. Nevertheless, a detailed introduction of all supervised classifier concepts is also beyond the scope of the dissertation. Interested readers can be referred to several references: A general introduction to pattern recognition and classification is given in the textbooks by Duda *et al*. (2000) and Bishop (1996, 2006), and in the review paper by Jain *et al*. (2000). A detailed introduction in context of remote sensing is given by Richards and Jia (1999) and Landgrebe (2003), a general overview by Richards (2005).

The presented study employs a selection of well known methods on the one hand and recent developments in pattern recognition and remote sensing on the other: (1) Although the use of the maximum likelihood classifier might be critical in context of multisource data, a comparison is worthwhile. Regarding the numerous remote sensing applications of this well-known method, it can be assumed as reference classifier. (2) Decision trees and DT-based classifier systems are also interesting for a detailed investigation. The training time is low and the handling is rather simple; the performance can be simply increased by classifier systems. Moreover DT classifiers are often used as benchmarks in pattern recognition experiments. (3) The investigation of more sophisticated and latest developments is also interesting. In recent literature the high performance of support vector machines is reported and the method is particularly interesting

for high-dimensional datasets. Nevertheless, a few multisource and multisensor applications are known, which use support vector machines. Hence the consideration of SVM is interesting in the scientific context of the dissertation. In the following a general guide to the concepts of maximum likelihood classifier, decision trees and support vector machines is given. Afterwards the concepts of classifier systems and segment-based approaches are introduced. The section is closed with a brief overview to the methods of classifier assessment.

2.2.1 Maximum Likelihood Classification

The maximum likelihood classification (MLC) is one of the most common supervised classification techniques in the field of remote sensing. As stated before, the MLC is based on the Bayes rule when classes have equal priorities. It assumes that the probability density function for each class is multivariate, and in the context of remote sensing image analysis, often a Gaussian distribution is assumed. A pixel is finally classified to that class, for which has the highest likelihood. A detailed introduction to the MLC is given for example by Richards and Jia (1999) and Lillesand and Kiefer (2000). In the following an overview is given.

A classification problem consists of n land cover classes that are described by $\Omega = \{\omega_i\}$, with i=1,...,n. Let in a d-dimensional feature space $\Re^d$, $x = \{x_1, x_2, \ldots x_d\}$ be a sample (or pattern) belonging to a class ωi. The determination of the class membership of x is defined by the conditional probability $p(\omega i|x)$. This so-called posterior probability describes the likelihood that ω_i is the correct class for x. The maximum probability value final determines the final class membership. Consequently the decision rule can determine as follows:

$$x \in \omega_i \quad if \quad p(\omega_i \mid x) > p(\omega_j \mid x) \tag{2.2-1}$$

The decision rule in equation 2.2-1 – also called maximum a posteriori (MAP) rule – is a simplified case of the Bayes rule. In general the decisions can be weighted by a loss function that reflects the loss incurred in associating x to ω_i when the true class is ω_j. With the knowledge of all class-conditional probabilities $p(\omega_i|x)$, the Bayes decision rule can be simply applied to the imagery.

On the other hand, class-conditional densities are usually not known. In this case the training samples can be used to estimate a probability density function $p(x|\omega_i)$ for each class (Jain *et al.*, 2000). This function is given the probability to find a member from class ω_i at pixel position x. For a pixel x, a set of probabilities – with as many functions as there are classes – can be calculated that gives the relative likelihoods that a pixel belongs to each occurring class.

The relationship between the known $p(x|\omega_i)$ and the required $p(\omega_i|x)$ is described as fallows (Richard and Jia, 1999):

$$p(\omega_i \mid x) = p(x \mid \omega_i)\, p(\omega_i) / p(x) \tag{2.2-2}$$

with $p(\omega_i)$ as the probability that class ωi occurs in the imagery.

The so called a priori probability $p(\omega_i)$ is an estimation of the class membership that is based on a priori knowledge before the classification; i.e., if 30% of an image belonging to class ω_i then $p(\omega_i) = 0.3$.

The probability to find a sample of any class at location x is described by $p(x)$. Equation 2.2-1 can be modified by using 2.2-2 and removing $p(x)$ as a common factor:

$$x \in \omega_i \quad if \quad p(x \mid \omega_i)\, p(\omega_i) > p(x \mid \omega_j) p(\omega_j) \tag{2.2-3}$$

Equation 2.2.3 is based that the probability distribution $p(x|\omega_i)$ can derived from the training and that $p(\omega_i)$ is known or can be estimated. In general it is assumed that the probability density functions for the classes are of the form of multivariate normal models (i.e., Gaussian models). In doing so $p(x|\omega_i)$ is determined, solely by the class-specific mean vector and the covariance matrix. If the number of available training samples is limited, the covariance matrix can be ill-conditioned and the efficiency of the classification is reduced (Richards and Jia, 1999; Lillesand and Kiefer, 2000)

2.2.2 Decision Trees

DT classifiers successively partition the training data into an increasing number of smaller homogenous classes by producing efficient rules, estimated from the training data. Whereas a specific class can be considered in several nodes of the tree another class might be separate by only one test. The most commonly used DT classifiers are binary trees. They use only one single feature at each node, resulting in feature axis parallel decision boundaries (Figure 2.2.1). Hence, they seem less adequate for complex class distributions. On the other hand the approach has several advantages as the fast training time and simple handling. A detailed introduction to decision tree classifiers is given by Safavian and Landgrebe (1991). A brief introduction is given below.

Let, for a 3-class problem in a 2-dimensional feature space, *a*, *b* and *c* the corresponding class labels. For the example, shown in Figure 2.2.1, five tests were applied to discriminate between the three classes. A potential decision tree is shown in Figure 2.2.2

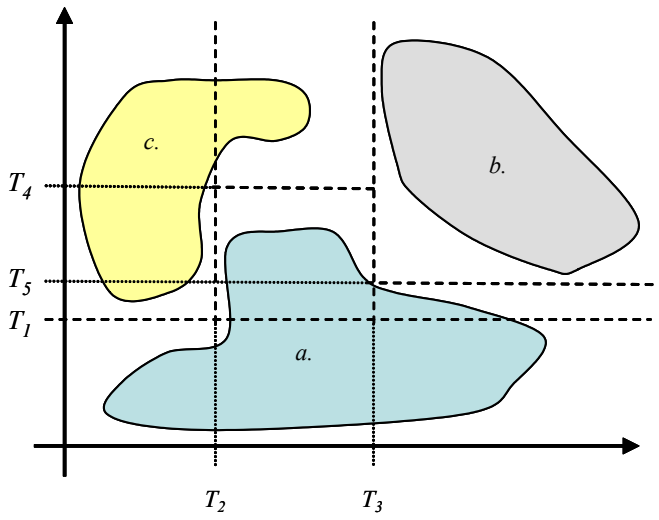

Figure 2.2.1: Distribution of the classes a, b, c in a 2-dimensional feature space. The dotted lines indicate the decision boundaries of the corresponding test T_i.

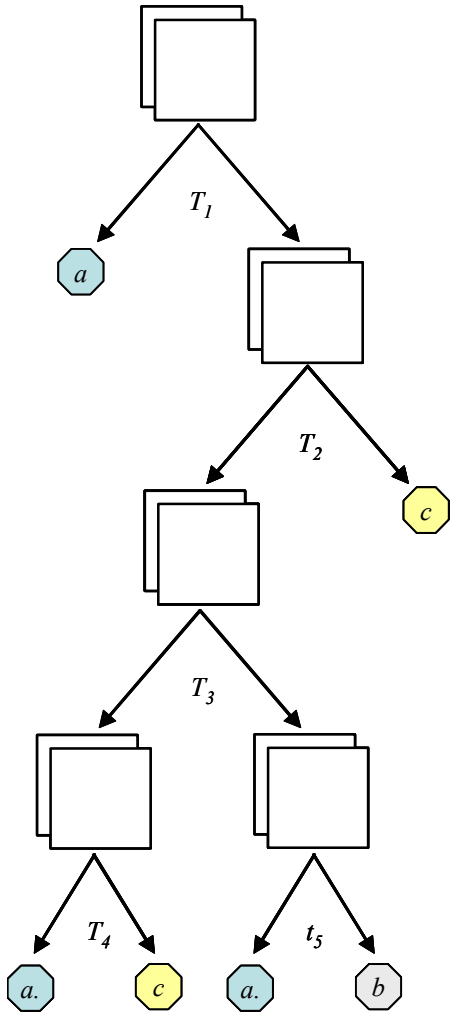

Figure 2.2.2: Schematic overview of a decision tree, separating three classes

The discrimination of the training samples is performed by the decision rule at each split, which is generally derived from the training data, using statistical methods. Consequently the method that is used for this purpose might be a crucial element of the method. The fundamental strategy underlying the creation of rule T at node N is that of purity, defining T in a way making the descendent nodes as pure as possible. The contrary formalization of the strategy is more convenient, using the impurity rather than the purity (Duda *et al.*, 2002). If all samples contained by N, belong to the same class, the node is pure and the impurity is 0, whereas it's large, if all classes are equally distributed.

Several concepts for the determination of the split rule and the measure of the impurity were introduced (Zambon *et al.*, 2006), whereas Quinlans' *information gain ratio* and Breimans' *gini index* are commonly used splitting criterions (Breiman, 1998; Quinlan, 1993).

The *information gain ratio* is based on the entropy or entropy impurity. The criterion is used to find a test that differentiate the training data set in node N, where n is the number of training samples in N with their corresponding class labels $\Omega = \{\omega_i\}$, with i=1,...,l.

The expected information provided at the node with respect of the class membership or required information to identify the class for an samples in N can be describe as:

$$E(N) = -\sum_{i=1}^{L} p(\omega_i) \cdot \log_2 p(\omega_i) \tag{2.2-4}$$

where $p(\omega_i)$ denotes the probability or the relative frequency of class ω_i:

$$p(\omega_i) = \frac{n_{\omega i}}{n} \tag{2.2-5}$$

with $n_{\omega i}$ as the number of samples belonging to class ω_i.

The test T differentiates N in z outputs $\{N_1, N_2, ..., N_z\}$. The total information content after applying the split rule T is denoted by the sum over the outputs:

$$E_T(N) = \sum_{j=1}^{z} \frac{n_j}{n} \cdot E(N_j) \tag{2.2-6}$$

The difference of equation 2.2-4 and 2.2-6 is the information gain that is achieved by separate N using rule T:

$$gain(T) = E(N) - E_T(N) \tag{2.2-7}$$

Using 2.2-7, the test T which maximizes the information gain is selected. A drawback of this approach is that the information gain prefers tests with a relatively large number of splits, i.e., an attribute with a relatively high variance or numerous values would be selected. Hence the expression is normalized by the information by the split itself:

$$E_s(T) = \sum_{j=1}^{z} \frac{n_j}{n} \cdot \log\left(\frac{n_j}{n}\right) \tag{2.2-8}$$

The final splitting criterion is defined as:

$$infogain\ ratio(T) = gain(N)/E_S(N) \tag{2.2-9}$$

The information gain ratio aims to reduce the entropy of the data that are separated by each split. The rule T at each split is selected that maximizes the reduction in the entropy (i.e, maximizes the ratio) of the descendant nodes.

The *gini index* measures the impurity at a given node and aims to separate the largest homogeneous group within the training data from the remaining samples (Zambon *et al*., 2006). The index is described as:

$$gini(t) = \sum_{i=1}^{n} p_{\omega i}(1 - p_{\omega i}) \tag{2.2-10}$$

where $p_{\omega i}$, the probability or the relative frequency of class ω_i at node T is defined as, with $n_{\omega i}$ as the number of samples belonging to class ω_i and n as the total number of samples within training set N.

$$p_{\omega i} = \frac{n_{\omega i}}{n} \tag{2.2-11}$$

with $n_{\omega i}$ as the number of samples belonging to class ω_i and n as the total number of samples within training set N. The impurity (i.e., gini index) of all sub-partitions is summed for each potential split rules and the test that causes the maximum reduction in impurity is selected (Apte and Weiss, 1997).

As other classifiers (e.g., neural networks) decision trees can be easily over trained. This is generally avoided by pruning. Pruning eliminates the inefficient, weak branches of a tree, to produce a more compact tree, which is not over trained (Duda *et al*., 2000). In general there are two concepts to generate a simplified tree: (1) the successive partitioning of the training is stopped before the whole (complex) tree is completed or (2) some parts of the DT are removed afterwards by recursive partitioning. A brief overview to pruning is given by Pal and Mather (2003).

2.2.3 Support Vector Machines

Support vector machines (SVM) were originally introduced as a binary classifier (Vapnik, 1998). SVM are based on an optimal linear separating hyperplane (OSH), which is fitted to the training samples of two classes within a multi-dimensional feature space. The optimization problem that must be solved is based on structural risk minimization. It aims to maximize the margins between the OSH and the closest training samples – the so-called support vectors (Vapnik, 1998). Hence, the approach only considers samples close to the class boundary and work well with small training sets, even when high dimensional data sets are classified. A detailed introduction to the general concept of SVM is given by Burges (1998) and Schölkopf and Smola (2002). An overview in the context of remote sensing is given by Huang *et al*. (2002) and Melgani and Bruzzone (2004), a brief introduction is given below:

For a binary classification problem in a *d*-dimensional feature space $\Re^d$, $x_i \in \Re^d, i = 1,2,...,M$ is a training set of M samples with their corresponding class labels $y_i \in \{1,-1\}$. The hyperplane *f(x)* is described by a normal vector $w \in \Re^d$ and the bias $b \in \Re$, where $|b|/\|w\|$ is the distance between the hyerplane and the origin, with ||w|| as the Euclidean norm from w.

$$f(x) = w \cdot x + b \tag{2.2-12}$$

The support vectors are located on two hyperplanes $w \cdot x + b = \pm 1$, which are parallel to the OSH. The margin maximization leads to the following optimization problem:

$$\min\left[\frac{w^2}{2} + C\sum_{i=1}^{L}\xi_i\right] \tag{2.2-13}$$

with ξ_i as the slack variables and C as the regularization parameter that are introduced to handle misclassified samples in non-separable cases.

The constant C is added as a penalty for cases that lie on the wrong side of the hyperplane (Figure 2.2.3). Effectively it controls the shape of the solution and thus affects the generalization capability of the SVM, e.g., a large value of C might cause an over-fitting to the training data.

In non-linear separable cases the linear approach is extended by so-called kernel methods. This concept is based on a non linear mapping of the data into a higher dimensional feature space, e.g., $\Phi : \Re^d \rightarrow H^{d'} \; (d' > d)$. In doing so, an OSH can be fit to a more complex class distribution that is generally not-separable in the original feature space. The transformed sample x is described by $\Phi(x)$ in the new high-dimensional space.

This computationally extensive transformation process of $(\Phi(x)\Phi(x_i))$ in a high dimensional space is reduced by using a positive definite kernel function k, which must meets Mercers conditions (Vapnik, 1998):

$$\Phi(x_i)\Phi(x_j) = k(x_i, x_j) \tag{2.2-14}$$

Thus the final hyperplane decision function can be described as follow:

$$f(x) = \left(\sum_{i=1}^{L} \alpha_i y_i k(x_i, x_j) + b \right) \tag{2.2-15}$$

where α_i are Lagrange multipliers.

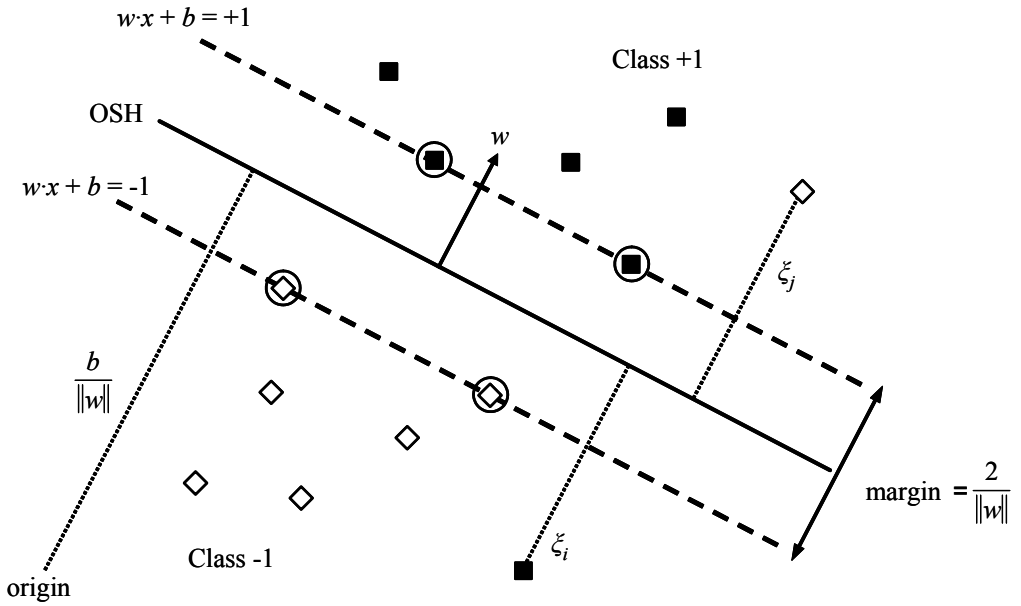

Figure 2.2.3: Concept of a SVM for a linearly non-separable case. The squares and diamonds indicate the samples of class ω_i (y_i =+1) and class ω_j (y_j =-1) respectively. The encircled samples refer to the support vectors, which lie on the two hyperplanes $w{\cdot}x$+b=-1 and $w{\cdot}x$+b=+1. The OSH lies between this two functions.

The so-called kernel-trick enables to work in the newly transformed feature space, without explicitly knowing Φ, but only the kernel function *k*. The polynomial and the Gaussian radial basis function (RBF) kernels are widely used kernel functions (Vapnik, 1998; Schölkopf and Smola, 2002):

Polynomial Kernel

$$k(x_i, x_j) = (x_i \cdot x_j + 1)^a \tag{2.2-16}$$

Gaussian kernel

$$k(x_i, x_j) = \exp\left[-\gamma \left\| x_i - x_j \right\|^2\right] \tag{2.2-17}$$

The SVM application requires the estimation of the kernel parameter, e.g., γ and the regularization parameter *C*. Different approaches for an automated model definition have been introduced (e.g., Chapelle *et al*. 2001; Chung *et al*. 2003), which based generally on a leave-one-out procedure.

In contrast to other classifier algorithms (e.g. decision tree) the original output of a SVM (2.2-15) does not contain the final class label. The imagery contains the distances of each pixel to the optimal separating hyperplane, from now on referred to as rule images. These rule images are used to determine the final class membership, depending on the multiclass strategy.

Multiclass strategies

As mentioned above, SVM have originally been introduced for binary classification problems that typically do not occur in the context of remote sensing image analysis. Hence several concepts have been developed for solving multiclass problems. In general a *n*-class problem is divided into several sub-problems and the individual binary SVM are combined in a set of classifiers. Two main strategies exist: the one-against-one strategy (OAO) and the one-against-all strategy (OAA).

Let $\Omega = \{\omega_i\}$ with $i=1,...,l$ be a set of *l* possible class labels (i.e., land cover classes). The OAO strategy is based on the training of $l(l\text{-}1)/2$ binary SVM, one for each possible pair-wise classification problem, discriminating ω_i and ω ($\omega_i \neq \omega_j$), e.g. cereals vs. forest, cereals vs. urban, forest vs. urban, etc. Consequently a set of $l(l\text{-}1)/2$ rule images is created. The sign of the distance to the hyperplane (i.e., the rule images) is used within the voting scheme. For the classification a score function S_i is computed for each class ω_i. All positive (i.e., sgn=+1) and negative (i.e., sgn=-1) votes for the specific class are summed and the final class membership of *x* is derived by a simple majority vote:

$$S_i(x) = \sum_{\substack{j=1 \\ j \neq i}}^{j} \operatorname{sgn}\left(f_{ij}(x)\right) \tag{2.2-18}$$

In contrast to the OAO approach the OAA concept is based on the training of *l* SVM. Each class ω_i is separated from the remaining $\Omega\text{-}\omega_i$, e.g., cereals vs. rest, forest vs. rest, urban vs. rest, etc. The absolute maximum distance to the hyperplane (i.e., the absolute maximum value within the

rule images) determines the final class membership, instead of using the simple sign of the decision function as done in the OAO. This approach is also known as "winner takes all", because only one value – the maximum – is used for the membership decision. In contrast to these two multicass methods, other approaches directly defined the SVM as one multiclass problem (Hsu and Lin, 2002; Sebald and Bucklew, 2001). However, a simultaneous differentiation of more than two classes leads to more complex optimization problem (Sebald and Bucklew, 2001). Consequently such approaches might be less efficient compared to conventional multiclass strategies. In Melgani and Bruzzone (2004) a computationally favorable hierarchical tree-based SVM was introduced as an alternative concept. Nevertheless compared with the original OAO concept the method achieved a slightly lower overall accuracy.

2.3 Classifier systems

Land cover classifications based on multiple classifier systems have been performed successfully during the recent years and are particularly interesting for multisource and high-dimensional data sets. Besides several applications in the context of remote sensing and pattern recognition, it has been shown theoretically that the classification accuracy can be increased by combining different independent classifiers (Schapire, 1990, Turmer and Gosh 1996, Kittler 1998). Two strategies exist to generate a multiple classifier system: (1) a combination of different classifier algorithms and (2) a combination of variants of the same algorithm.

By training the so-called base classifier on modified input data (i.e., training samples or input features) a set of independent classifiers can be produced, from now on referred to as classifier ensembles. For generating the final result these outputs are combined by a voting scheme (Figure 2.3.1). Often a simple majority vote is used, which can be effective than more complex voting strategies.

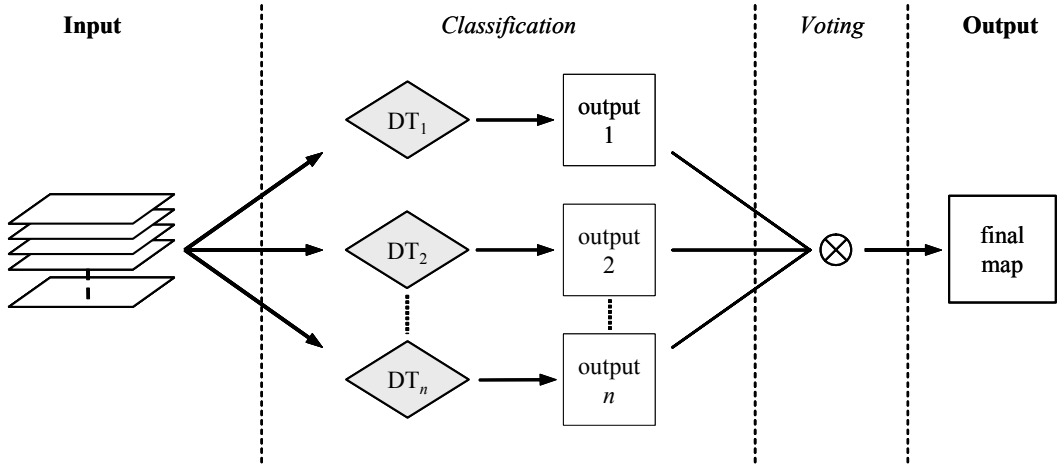

Figure 2.3.1: Schematic diagram of a classifier ensemble

The concept is based on the assumption that independent classifiers produce individual errors, which are not produced by the majority of the other classifiers. Various strategies for generating classifier ensembles have been introduced as for example a resampling of the training data (i.e., bagging or boosting) and the input features (e.g., random feature selection).

Banfield *et al*. (2007) give an introduction and assessment to various types of classifier ensembles; a brief overview to the main concepts is given below. Breimans' (1996) bootstrap aggregating (bagging) describes the random generation of training sample subsets also known as bootstrapped aggregates or bags. The approach is based on the random and uniform selection, with replacement , of *n* samples from a training set of same size *n*. i.e., a training sample can be selected several times in the same sample set and perhaps other samples are not considered in this particularly bag. Afterwards an individual decision tree is trained on each of these bags, resulting in various independent classifier outputs. The final classification map is generated by combining the individual outputs.

The concept of boosting was originally introduced by Schapire (1990) as an approach to improve the performance of a weak learning algorithm. For the iterative training process, all training samples are equally weighted in the beginning. Boosting successively changed the weights of the training samples during the training process, comparing the outputs with the known class memberships of the samples. Whereas in the initial phase all training samples are equally weighted, misclassified samples are assigned a stronger weight than those classified correctly. The next DT within the ensemble uses the newly distributed, reweighed samples. In doing so, the classifier is forced to concentrate on the misclassified samples that are more difficult to classify and can reduce the variance and bias of the classification. Unlike bagging that can be performed simultaneously, boosting generates the different classifiers in a sequentially procedure. Consequently it computation time is relatively slow. The AdaBoost.M1 approach (Freund and Schapire, 1996) is widely used in the field of pattern recognition and remote sensing.

Besides resampling of the training data, the modification of the input feature space, e.g. by a random selection of features, is another concept for generating independent classifiers (Ho, 1998; Bryll *et al*. 2003). It has been shown that this random feature selection approach can be superior to bagging and boosting (Ho, 1998; Bryll *et al*., 2003). In contrast to the two aforementioned data partitioning strategies, the training samples remain unchanged by this concept. For each DT a subset of features is created. Unlike bagging that collects *n* samples from a sample set of size *n*, the method normally selects a subset of the available input features without replacement.

Breiman's (2001) random forests technique use a set of decision trees $\{DT(x, \Theta_m), m=1,\ldots,\}$, where Θ_m denotes independent identically distributed random vectors and x an input pattern (i.e., signature). Each tree within the ensemble is trained on a subset of the original training samples (i.e., bagging); in addition the split rule at each split is determined, using only a randomly selected feature subset of the input data. A simple majority vote is used to create the final classification result (

Figure 2.3.2). The number of selected features within the subset is user-defined, and the parameter is usually set to the square root of the number of input features (Gislason *et al*. 2006). The computational complexity of the individual DT classifier is simplified, by reducing the number of features at each split.

This enables random forests to handle high-dimensional data sets. In addition the correlation between the classifiers is decreased, which generally improves the performance of a classifier

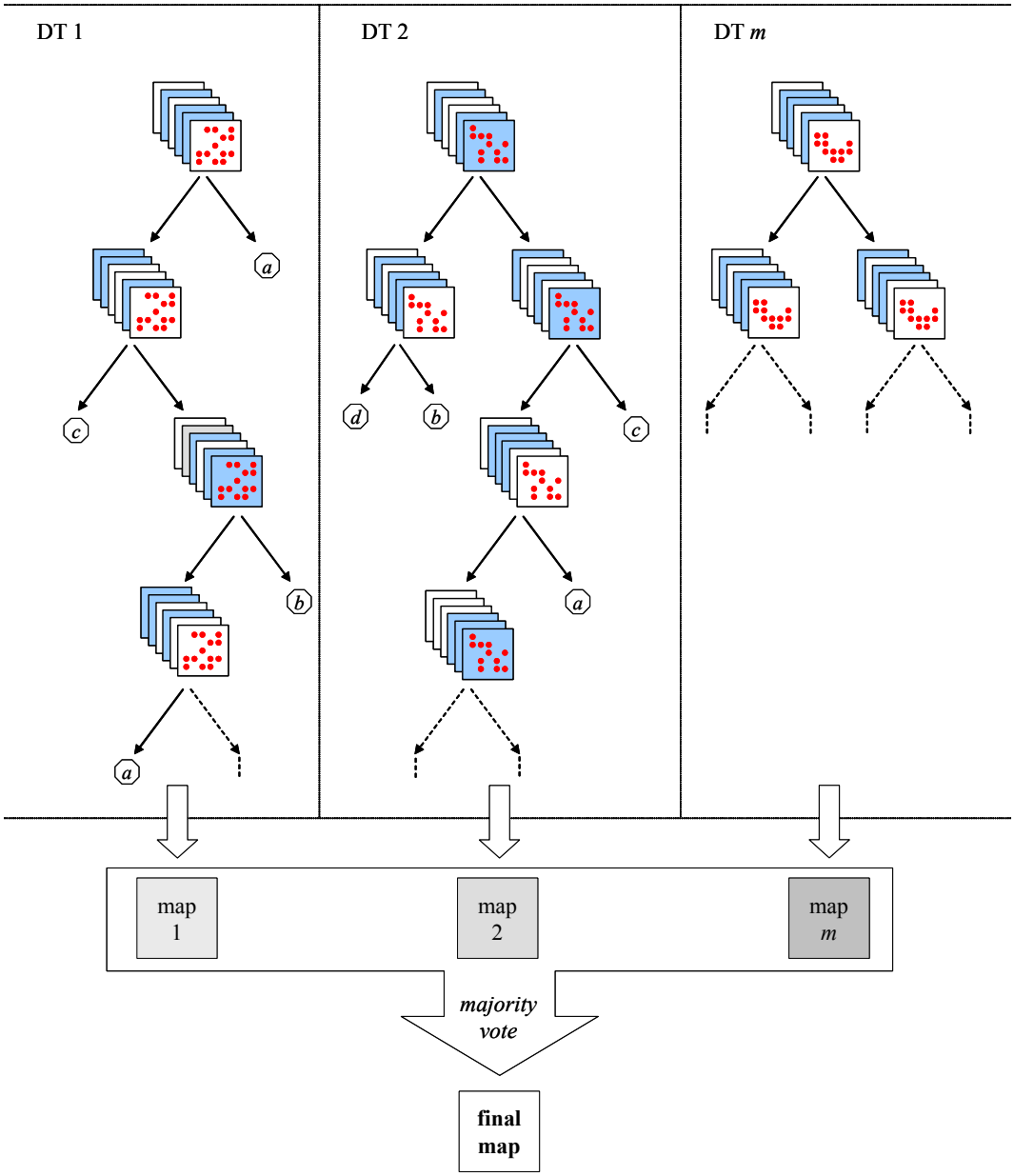

Figure 2.3.2: Schematic diagram of random forests

system. The method is from a computational view lighter than bagging and boosting concepts, because it is only based on subsets of input data (Gislason *et al*. 2006).

In contrast to the aforementioned classifier ensembles, other concepts are based on the combination of different classifier algorithms, from now on referred to as multiple classifier systems. By using different classifier algorithms, it can be assumed that different outputs are created; moreover the individual advantages of each method can be used. Perhaps the methods were developed in a diverse context and enable a different description of the same classification problem. Furthermore each method may have its own region in the feature space where it performs the best (Jain *et al*., 2000). Several of these studies are based on decision fusion, which is also used for combing different data sources (e.g., Jeon and Landgrebe, 1999;

Bachmann *et al*., 2003; Fauvel *et al*., 2006; Gamba *et al*., 2007). Decision fusion can be defined as a strategy of combining information from different data sources, after each source has been classified separately. The decision fusion concept can be based on consensus theory for example, which employs single probability functions to summarize estimates from various sources (Benediktsson and Swain, 1992), using consensus rules. The linear opinion pool and the logarithmic opinion pool are common consensus rules. Nevertheless, his technique requires the costly selection of weights, which generally reflect the reliability of the input data (Benediktsson *et al*., 1997). In contrast to this computationally complex approach, other studies uses voting concepts like majority voting and complete agreement (Benediktsson and Kanellopoulos, 1999; Fauvel *et al*. 2006a).

2.4 Segment-based classification

In segment-based approaches adjacent pixels with similar properties are aggregated into image segments. In several of these studies segment-based classifications are more accurate than conventional per-pixel classifications. By image segmentation information on the segments' mean spectral or backscatter value, their texture and shape as well as neighborhood relationships can be derived, which can be used during subsequent classification. Moreover such an approach can solve or reduce the effects of two common problems, which are often arising in pixel based classifications (Smith and Fuller, 2001; de Wit and Clevers, 2004): (1) The variability of spectral reflectance and backscatter intensities within an object as an agricultural parcel can be affected by site-internal variations, e.g., like moisture, soil heterogeneities and plant infections. Hence pixel values vary on plot level and a definite assignment to a single land cover class might be difficult. (2) The spectral properties of pixels, which lie along the boundaries of two objects might be a mixture between both classes (e.g., two different field plots). Consequently these mixed pixels do not represent either of the two land cover types. By image segmentation the pixel values can be averaged and the mixed pixels are eliminated. In regard to SAR imagery, which is affected by speckle, image segmentation has the positive effect that the noise is leveled out. In Tso and Mather (1999) the classification accuracy of multitemporal ERS-2 imagery was significantly improved by an segment based approach.

In general two main concepts of segment-based classifications can be distinguished: In the first approach the training and classification process is performed on previously segmented imagery. In contrast to this method, another approach is based a conventional pixel-based classification. After the classification is performed, the majority class within each segment is calculated and all pixels belonging to the segment are assigned to this specific class. The approaches used in this dissertation belong to the first concept.

Several techniques have been developed for image segmentation (Le Moigne and Tilton, 1995; Baatz and Schäpe, 2000; Evans *et al*., 2002). Boundary-based approaches are using the contrast in the grey-level distribution function of pixel values. Region-growing or region-based methods perform under the assumption that pixels of the same natural feature have a certain spectral homogeneity. A well-known method is based on a region-growing algorithm (Baatz and Schäpe, 2000). A description of the approach is given by Shackelford and Davis (2003), general constraints of the concept are described in Bruzzone and Carlin (2006).

In the beginning, each pixel is considered as an individual segment and is subsequently merged with adjacent segments into new larger segments. For the merge of two segments, the difference Δh between the heterogeneity of the new segment (e.g., spectral or backscatter heterogeneity) and the heterogeneity of the constituent segments is calculated as follows:

$$\Delta h = \sum_{i=1}^{b} \theta_i (\rho_1 (h_{m\,i} - h_{1i}) + \rho_2 (h_{mi} - h_{2i})) \tag{2.4-1}$$

where h_{mi} denotes the heterogeneity of the potentially merged segment for the heterogeneity measure i (e.g., the variance in a specific band). Those of the two individual segments are described by h_{1i} and h_{2i}. The number of pixels belonging to the segments is defined by ρ_1 and ρ_2. The weight θ_i controls the impact of the specific heterogeneity.

The merge is performed, when this increase is below a user defined stopping criterion, h_s. The segmentation process stops as soon as this condition can not be fulfilled by any possible merge. Consequently segment size is driven by h_s. In general, Δh is based on the spectral/back-scatter variance within user-defined bands and on segment shape (e.g., the compactness or smoothness).

2.5 Evaluation of classifiers - accuracy assessment

The evaluation of the algorithm and the corresponding classification results is an important issue in context of digital image analysis and land cover classification. Beside the performance of the algorithm, the classification result is affected by several factors: e.g., land cover types, which have to be classified, quality of the training samples and available input imagery (i.e., temporal, spatial and spectral resolution). In addition the geographical properties of a study area might have a dominant impact on classification accuracy. Smith *et al*. (2002) have shown that the classification accuracy significantly decreases with an increasing heterogeneity of the landscape and in contrast the accuracy is improved with increasing patch sizes.

A general classifier assessment can be based on a qualitative evaluation, using expert knowledge or a quantitative assessment, using statistical methods and a priori information. The main measurement of performance is the classification error or the corresponding overall accuracy. The measurement describes the total fraction to which degree the produced image classification agrees with the reality (i.e., the reference that is assumed to be truth). Thus a divergence between the map and reference information is a classification error (Foody, 2002). Contrary to the overall accuracy a classification method can be also assessed in terms of reproducibility, robustness to noise, dependency on the training patterns or computational advantages (DeFries and Chan, 2000). However, the main criterion – which is generally used in each evaluation, is the classification accuracy.

Before introducing a methodology for an accuracy assessment, it should be mentioned that – beside the classification error – several kind of errors exists (Foody, 2002). Spatial distortions, e.g., due to geometrical correction and resampling of the image data, can be a major source of misclassifications (Muller *et al*., 1998). Errors in the ground truth and reference data are another error-source that affects the classification accuracy (Foody, 2002). Man-based ground truth campaigns might be subjective and dependent on the field analysts. Moreover the field mapping

strategy is often based on natural features and objects, particularly in context of regions that are dominated by agricultural land use. Hence the sampling unit in the ground truth data might be different to the classified imagery (e.g. field plot versus pixel).

In generally the accuracy assessment is based on the accuracy or confusion matrix, which compares ground truth data with the corresponding classification for a given set of validation samples (Congaltion and Green, 1999; Foody, 2002). The accuracy matrix enables the derivation of the most common evaluation criterions: (1) overall accuracy, (2) producer accuracy, (3) user accuracy. A detailed overview is given by Foody (2002) and Congaltion and Green (1999). A brief introduction is given below.

Let, for a *l*-class classification problem, *N* be the total number of reference samples. The corresponding confusion matrix is illustrated in Figure 2.5.1. The number of samples that classified as class ω_i (i=1,2,..,l) and belong to land cover class ω_j (j=1,2,…,l) are described by n_{ij}, e.g. n_{11} denotes the number of samples that belongs to class 1 and correctly assigned to class 1, whereas n_{21} define the samples belonging to class 1, but incorrectly classified to class 2. The diagonal elements n_{cc} (i.e., the highlighted elements in Figure 2.5.1) of the error matrix contain the number of correctly classified samples. The overall accuracy can be derived by their sum divided by the total number of samples *N*:

$$overall\ \ accuracy\ = \frac{\sum_{c=1}^{L} n_{cc}}{N} \tag{2.5-1}$$

Another measurement for the assessment of the overall accuracy is Cohen's kappa coefficient, which is often used for a comparison of maps from different regions. In the presented dissertation, the overall accuracy is used.

In addition the confusion matrix can be used to derive class-specific accuracies (i.e., producer and user accuracy) or corresponding error rates. An error of omission is excluding a sample from a class in which it originally does belong. A commission error on the other hand assigned a sample to a wrong class. Consequently each error is an omission from the correct class and a commission to a wrong class. In this work the contrary (positive) formalization of the indices is used, i.e., the producer and user accuracy

The producer accuracy for class *c* is calculated by dividing the number of correct samples of *c* by the total number of reference samples of class *c*.

$$producer\ \ accuracy\ = \frac{n_{cc}}{n_{+c}} \tag{2.5-2}$$

The user accuracy describes how many samples that have been classified as c are actually belonging to class c. The measurement is derived by:

$$user\ \ accuracy\ = \frac{n_{cc}}{n_{c+}} \tag{2.5-3}$$

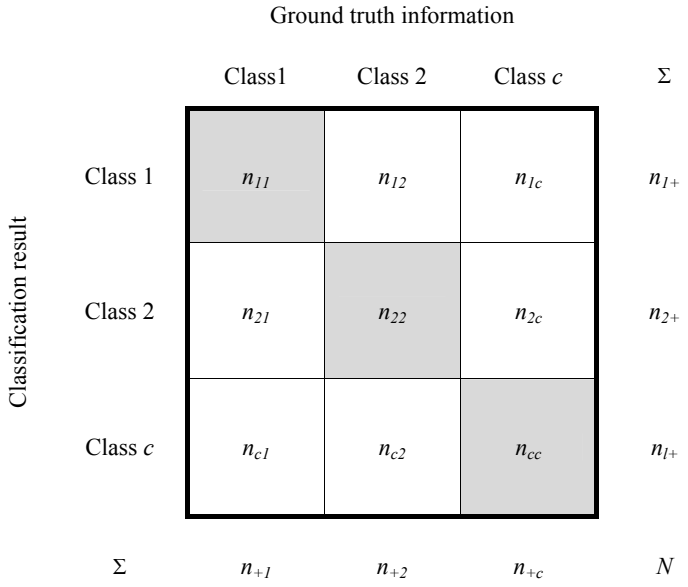

Figure 2.5.1: Schematic diagram of a confusion matrix, the highlighted elements contain the number of correctly classified samples (Congalton and Green, 1999, modified)

For classifier training and an adequate accuracy assessment independent samples sets are required. Sample sets can be generated in different ways: e.g. (1) simple random sampling, (2) systematic sampling, and (3) stratified random sampling. When using the first one each sample (e.g., pixel) has an equal chance to be selected, the systematic approach selects samples with an equal interval over the test site. The stratified random sampling combines some prior knowledge about the regions with the simple random sampling. The stratified random sampling seems advantageous, because the approach guarantees that all classes are included in the sample set (Congalton and Green, 1999). In the presented study an equalized random sampling strategy is performed, guaranteeing that all classes are uniformly included in the sample set.

3 Study site and data

3.1 Study site

The study area is located near Bonn in the German state North Rhine-Westphalia (NRW), in the Köln Aachener Bucht (Figure 3.1.1). The annual mean temperature vary approximately between 9.2°C and 9.5°C and the precipitation between 600 mm and 630 mm per year. The almost flat region is characterized by a predominant agricultural use with wheat and sugar beets as one of the main crop types. A visual assessment clearly shows that the study is characterized great variability and typical spatial patterns of high-frequent land cover changes between, relatively small, individual plots (Figure 3.1.1).

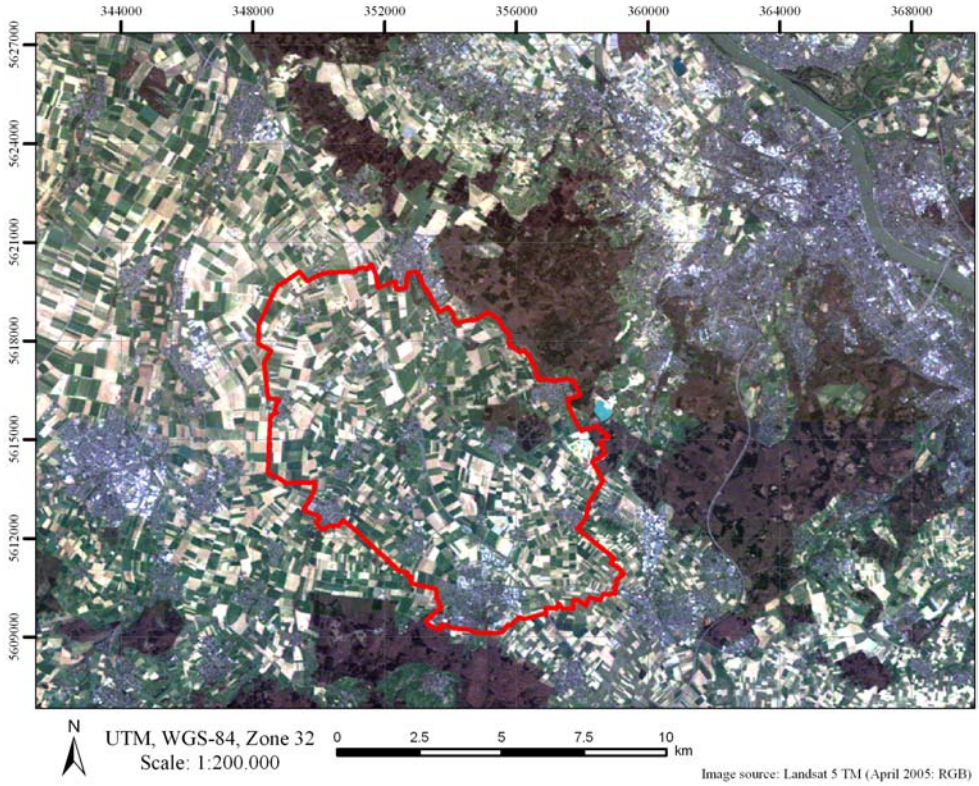

Figure 3.1.1: Location of the study site, the red polygon indicates the location of the main study site.

3.2 Remote sensing data

SAR and multispectral EO-systems operate in different wavelengths, ranging from visible to microwave domain. The interaction between the land surface and the radiation is different (e.g., Kühbauch and Hawlitschka, 2003), consequently the two systems provide different information. In the visible range of the electromagnetic spectrum the energy is highly absorbed by vital vegetation. Chlorophyll, for example, strongly absorbs the energy in the blue and range, whereas the green wavelength is highly reflected. Thus healthy vegetation is appears as green colour. The amount of reflected energy significantly increases at about 0.7 µm. The internal structures of plants that impact the electromagnetic spectrum in the range 0.7 to 1.3 µm are

clearly vary between different species. Thus, this information often enables a separation of vegetation types, which are similar in the visible domain Lillesand and Kiefer (2000).

In case of soil reflectance characteristics the signal is dependent on characteristics as soil moisture, soil texture, iron oxide and organic matter content. Beside this various properties, the reflection in the optical range is also impact by characteristics, as illumination and viewing properties and geometrical properties of target. Bauer *et al*. (1986) give an introduction and quantitative description to the complex interaction between vegetation, soil, atmosphere, and illumination and sensor geometries in context of agricultural crops (1986). A general overview can be found in several text books, e.g., Lillesand and Kiefer (2000) and Jensen (1996).

Microwaves, on the other hand, interacts with vegetation canopy primarily with their volume and their structural elements, leaves and stems. In general these interactions depend on several factors: (1) sensor parameters as wavelength, polarization and local incident angle, (2) characteristics of the illuminated target as roughness, dielectric properties and topography: The penetration depth of the electro magnetic (EM) waves generally increases with decreasing frequency. Relatively short wavelengths (i.e., high frequency) are reflected from the upper surface of a medium, e.g., from the leaves of a vegetation canopy, consequently the surface scattering from the underlying surface is minimal. Lower frequencies (i.e., higher wavelength) on the other hand, can penetrate deeper into the illuminated medium. The backscatter results from corner reflections between the stems and the underlying surface (i.e., volume scattering).

The polarization of microwaves is another important parameter, affecting the backscatter signal. The electro magnetic wave consists of three vector fields: (1) direction of propagation, (2) electric filed and (3) magnetic field. The active microwave consists of a polarized component, which is defined by the electro magnetic field vector of the wave (Henderson and Lewis, 1999). The (transmitted and received) electric field can be polarized horizontally (H) or vertically (V) in terms of to the direction of propagation. Whereas in co-polarized systems the transmitted and received polarization is identical, cross-polarization is referred to as different transmit-receive combinations. However, the impact of polarization and wavelength varies with varying geometrical target properties. In context of crop monitoring for example, an increasing density of a vegetation canopy (with a generally corresponding decrease in stem diameter) is leading to a decrease of the adequate wavelength. Consequently lower frequencies appear more convenient for mapping crops as maize and root crops, whereas shorter C and X-band waves seems more appropriate for monitoring cereals as wheat and barley. Vertical polarized data (VV) provide the most significant information for crops, which are dominated by vertical structures (e.g. wheat and barley). In contrast to this, the cross-polarization HV (i.e., horizontally transmitted and vertically received) is more applicable for ramified crops, as maize and rapeseed and a general differentiation between different crops types (Brisco *et al.*, 1992; Ferrazzoli, 2002). Besides the surface roughness, the backscatter mechanism of EM-waves is significantly forced by the dielectric properties of the medium. In general the dielectric constant describes the interaction of the medium with an electric field. The dielectric constant of water is significantly higher compared to those of most natural materials (Henderson and Lewis, 1999). Consequently the backscattering mechanisms are significantly influenced of the water content of the target media (e.g., soil moisture). Increasing moisture content is leading to higher backscatter intensities, whereas at the same time the penetration depth decreases.

3.2.1 Multispectral data

For this study multispectral imagery from different EO-mission were available. The Landsat 5 Thematic Mapper (TM) was launched 1984 in cooperation between the National Geological Survey (USGS) and the National Oceanic and Atmospheric Administration (NOAA). The Landsat sensor is a "whiskbroom scanner", which uses an oscillating mirror to reflect the incoming radiance to a detector system.

Furthermore data from SPOT-2 and SPOT-5 (Systeme Pout l'Observation de la Terra) were used, which were launched by the French space agency CNES in 1990 and 2002. In contrast to Landsat, the SPOT system is based on a pushbroom scanner, which uses a linear array of CCDs arranged along a line vertical to the satellite orbit track (Lillesand and Kiefer, 2000; Richards and Jia, 1999). A technical overview to the main sensor parameters of the two systems is given in Table 3.2.1 (USGS, 1984; Spot, 2005).

Table 3.2.1: Overview to sensor configurations of Landsat-5 TM, SPOT-2 and SPOT-5

Parameter	**Configuration**		
	Landsat-5 TM	*SPOT 2*	*SPOT 5*
Orbit altitude	705 km	822 km	
Orbit repitition rate	16 days	26 days	
Incident angle range	-	±31.06°	
Swath width [km]	185 km	60-80 km	
Spectral bands	6*	3**	4**
Band width [μm]	B1: 0.45 - 0.52 B2: 0.52 - 0.6 B3: 0.63 - 0.69 B4: 0.76 - 0.9 B5: 1.55 - 1.73 B7: 2.08 - 2.35	B1: 0.50 - 0.59 B2: 0.61 - 0.68 B3: 0.78 - 0.89	B1: 0.50 - 0.59 B2: 0.61 - 0.68 B3: 0.78 - 0.89 B4: 1.58 - 1.75
spatial resolution	30 m	20 m	B1-B3: 10 m B4: 20 m

* thermal band is not considered **panchromatic band is not considered

3.2.2 SAR data

In this study SAR data from the European satellites Envisat and ERS-2 are used. After the launch of the first European Remote Sensing Satellite (ERS-1) in 1991, the ERS-2 has been launched in 1995, containing the Active Microwave Instrument (AMI). The Advanced Synthetic Aperture Radar (ASAR) is the biggest instrument boarded on the ENVISAT platform, which has been launched in 2002. In comparison to the previous ERS-1/2 missions the microwave instrument was further enhanced, enabling the acquisitions of different polarizations, incident angles and image modes. An overview to the technical specifications is given by ESA (2002). Table 3.2.2 shows the main configuration parameters of the two instruments. The ASAR instrument provides 5 different *imaging modes*, with different spatial resolution and coverage. Furthermore by selecting a predefined *swath*, images with different incident angle can be acquired.

Table 3.2.2: Overview to sensor configurations of ERS-2 AMI and Envisat ASAR.

Parameter	**Configuration**	
	ERS-2	*ASAR*
Orbit altitude	785km	~799 km
Orbit repitition rate	~35 days	~35 days
Incident angle range	19-27°	14-45°*
Swath width	100 km	58-109km*
Frequenzy / wavelenght	5.3GHz / 5.66 cm	5.331GHz / 5.62cm
Polarization	VV	VV / HH / VH / HV*
Calibration accurcay	±0.5 dB	±0.5 dB
spatial resolution**	~30 m	

* dependant on selected mode, ** appr. Resolution of PRI products

The imaging modes are defined by the user before the acquisition: The *Image Mode* (IM) enables the selection between any of the 7 available swaths. Due to the dual-polarization of the system, the electromagnetic field can be polarized horizontally or vertically, resulting in co-polarized images (HH or VV). The spatial resolution of precision products is approximately 30 m, with a spatial coverage between 56 km (swath 7) and 100 km (swath 1) across track. Using the *Alternating Polarization* (AP) mode, two co-registered images are generated per image acquisition. Thus, a simultaneous acquisition of different polarized data sets is possible. In contrast to common co-polarized systems, the AP mode enables different transmit-receive combinations (i.e, cross-polarization), and operates in three potential combinations: HH/VV, HH/HV or VV/VH. Furthermore the instrument enables wide area coverage with a medium spatial resolution of approximately 150 m (*Wide Swath* mode) and 1 km (*Global Monitoring* mode). In addition a *wave mode* is available for ocean monitoring.

The selectable swaths (Table 3.2.3) enable the observation of a study area from different orbit paths, increasing the temporal resolution of the image data set. Furthermore different incident angels provide dissimilar information. Moreover for a certain application a specific incident angle is more adequate than another and perhaps for other application another incident angle is more appropriate.

Table 3.2.3: Configurations of ASAR image swaths

Image swath configurations		
Swath number	*Incident angle* [°]	*Swath width* [km]
IS 1	15.0 - 22.9	105
IS 2	19.2 - 26.7	105
IS 3	26.0 - 31.4	82
IS 4	31.0 - 36.3	88
IS 5	35.8 - 39.4	64
IS 6	39.1 - 42.8	70
IS 7	42.5 - 45.2	56

3.3 Auxiliary data

For generating training and validation data two extensive ground truth campaigns were conducted in during summer 2005 (Figure 3.3.1) and 2006 (Figure 3.3.2), with a focus on agricultural land use. Regarding the dominant agricultural land use and the typical crop penology within study area, it is assumed that no critical changes are present during the periods of image acquisition. For the field campaigns a Geographical Information System (GIS) was generated, using a priori information and high-resolution aerial photography. The polygon layers were modified, using the image information. During the field campaigns the actual boundaries of natural features were controlled by Global Positioning Systems (GPS) and the vector files within the GIS project were updated. For each object (e.g. a forest area, an agricultural parcel or an urban feature) the actual land cover was recorded on plot level. Beside the remote sensing imagery and ground truth information a Digital Elevation Model (DEM) was available.

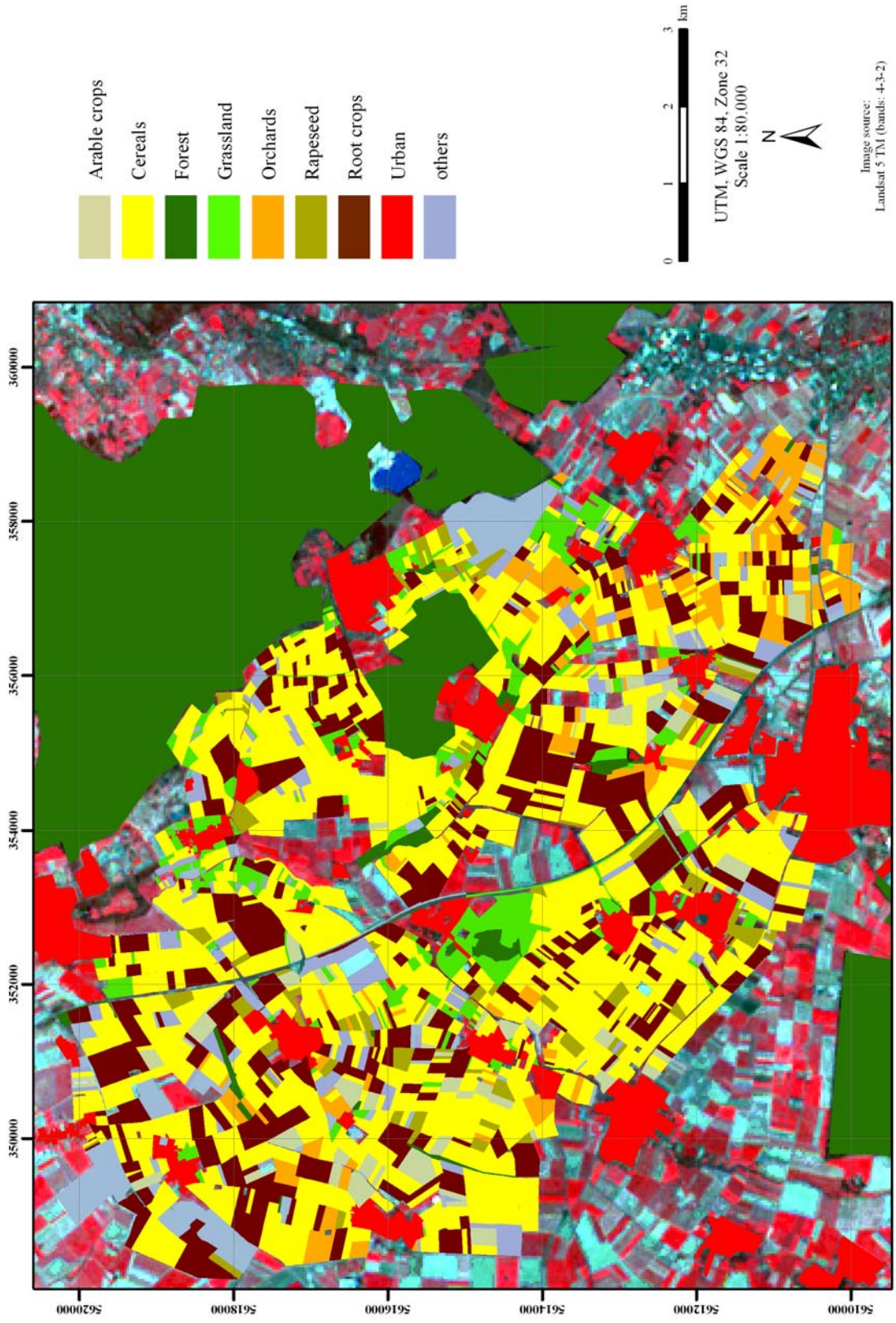

Figure 3.3.1: Study site and corresponding land use in 2005.

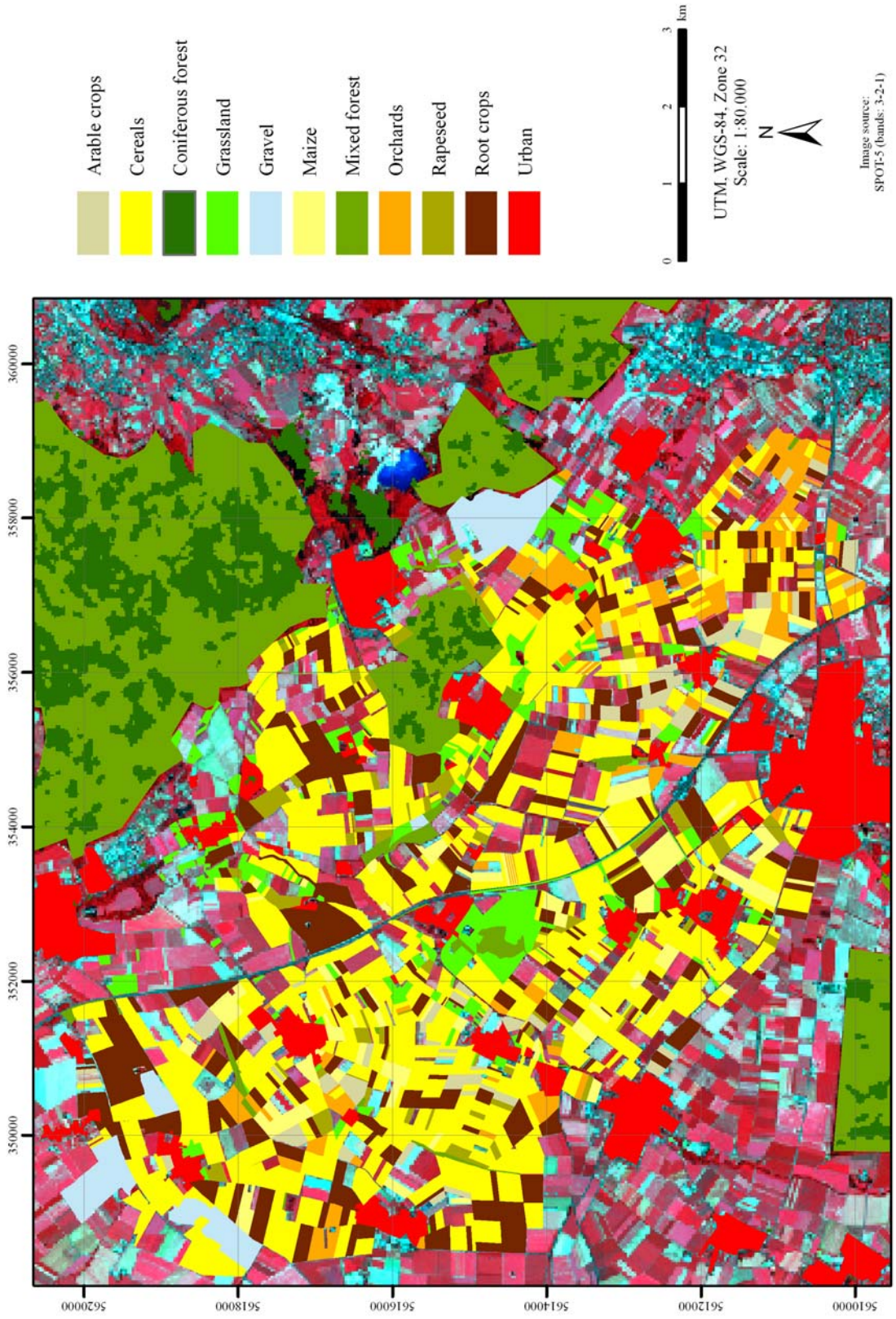

Figure 3.3.2: Study site and corresponding land use in 2006.

4 Data preprocessing

Before remote sensing data can be analyzed, a geometric and radiometric preprocessing must be performed, because the imagery can contain systematic and non-systematic distortions, as a consequence of the EO-system and the characteristics of the Earth surface. Such distortions can impact the geometry of the imagery as well as the radiometric information (Richards and Jia, 1999; Henderson and Lewis, 1999). In Figure 4.1.1 an overview to the preprocessing is given. Below a brief introduction to the different preprocessing steps is given.

4.1 Preprocessing of multispectral data

4.1.1 Radiometric correction

The signal measured at the satellite is an outcome of the interaction between the solar irradiance with the illuminated surface and its two-fold way of the radiance through the atmosphere. Due to this interaction, different radiance fluxes illuminate the target surface and can be measured at the sensor system respectively (Jensen, 1996; Mather, 1999). Hence, the at-satellite reflectance is dependent on several factors, as surface characteristics and atmospherically absorptions and scattering processes, by gaseous and aerosols components. Consequently the actual constitution of the atmosphere has a significantly impact on remote sensing imagery and a correction of these impacts seems worthwhile (Mather, 1999; Tanré 1987, Hill und Sturm, 1991). At the satellite the incoming radiances are converted to a digital signal, which can be described by a calibration function. The further radiometric pre-processing is based on actual physical measurements of spectral radiances; hence the digital numbers (DN) are converted to radiance values. The sensor calibration is described by:

$$L(\lambda) = a \cdot DN + b \tag{4.1-1}$$

with DN as digital number (i.e. pixel value), a as sensor gain and b as sensor offset.

Beside several simple approaches, as scene averaging or empirical line correction (e.g. Richards and Jia, 1999), several approaches are based on the physically modeling of the atmospheric processes (Tanré 1990, Hill and Sturm, 1991). In the study presented here, the simplified and fast correction approach *ATCOR* (Richter, 1996a; Richter, 1996b) was used. In contrast to more sophisticated radiative transfer models as *Modtran* (Berk, 1989), the approach is based on inverting previously results of Modtran calculations.

4.1.2 Geometrical correction

For matter of comparison and combination of different information sources, like remote sensing imagery from diverse sensor systems, ground truth data and topographical information (e.g., DEM), all available data have to be transferred into a reference cartographic projection system, to generate a generally valid data basis. As a consequence all geometrical errors must be corrected. Significant geometrical errors within the Landsat imagery are for example earth rotation skew, panoramic distortion, pixel size distortions, and variations in platform speed and elevation (Richards and Jia, 1999; USGS, 1984). Those distortions are considered by the system correction regularly applied to the imagery prior delivery (USGS, 1984). The system corrected

imagery need to be transferred into a reference cartographic projection system (e.g., UTM, WGS-84). The relation between the image to be corrected and the reference system is compute by identifying corresponding positions in the imagery and a reference scene that is already available in the desired cartographic projection. These positions are located by Ground Control Points (GCPs), which are used to derive a transfer function between the imagery and the reference system, i.e., the master scene. Using a DEM the image is orthorectified with spatial accuracy of approximately one pixel. The resampling is performed by *cubic convolution*. Even this technique modifies the output values; it usually generates a much smoother output image. The geometric correction of the SPOT imagery is performed similar. As before a correction of systematic distortions, including panoramic effect, earth curvature and rotation and variations in the satellite orbit, was performed prior delivery. In addition the data is corrected by normalizing the CCD response to adjust radiometric variations due to detector sensitivity (Spot, 2005). Finally, the SPOT data were resampled to the pixel size of the Landsat TM image (30 m) and orthorectified with a spatial accuracy of approximately one pixel, using a digital elevation model and the corrected Landsat scene as a reference image.

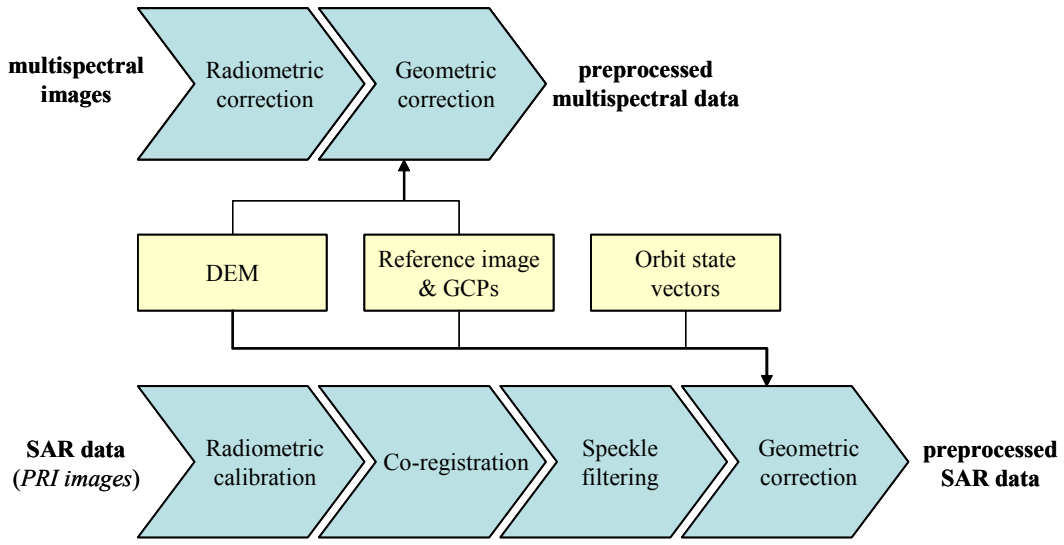

Figure 4.1.1: Schematic overview of the preprocessing

4.2 Preprocessing of SAR data

In this study SAR precision imagery (PRI) is used. Preprocessing steps as slant to ground range conversion and multilooking are already performed and the data is delivered in radar brightness, antenna pattern and range spreading loss corrected. The discussion of the generation of these data products would be beyond the scope of the dissertation. A good introduction to the topic of SAR image formation is given, e.g., by Oliver and Quegan (1998) as well as Henderson and Lewis (1999). An overview to preprocessing of ERS-2 and Envisat ASAR products is given by Lauer *et al*. and ESA (2002) respectively, whereas the preprocessing of the used data products is comparable. The preprocessing chain of the SAR data within this study is dived into four components: (1) radiometric calibration, (2) co-registration, (3) speckle filtering, and (4) geometrical correction.

4.2.1 Radiometric calibration

The radar backscatter coefficient sigma can be derived from the intensity values within the image data. For ground range detected products, the backscatter coefficient σ^0 can be derived by (ESA, 2002):

$$\sigma^0 = \frac{DN^2}{k} \cdot \sin\alpha \tag{4.2-1}$$

with DN as digital number (i.e., pixel value), α denotes the incident angle and k the calibration constant, which is provided with the image product.

Although during the normal image processing a flat earth is assumed, the radar backscatter is affected by local topography. Hence, different concepts for the correction of these effects have been introduced (Zyl *et al*., 1993, Ulander 1996). On the other hand the study area is almost flat. First test have shows no significant differences between terrain corrected and "original" image, consequently a terrain correction is not conducted. Afterwards acquisitions with same viewing properties were co-registered with a spatial accuracy of approximately one pixel, using a maximum correlation approach.

4.2.2 Speckle filtering

One drawback of SAR data is a random noise within the image, known as speckle. The speckle is caused by interferences between the radar waves reflected from multiple scattering elements within an individual resolution cell (Oliver and Quegan, 1998; Henderson and Lewis, 1999). The speckle degrades the image quality; reduces the image information content and hampers image analysis like land cover classifications and image segmentation (Touzi, 1988, Lee and Jurkevich 1989, Capstick and Harris, 2001, Herold *et al*, 2005). In the context of adaptive filter algorithms the speckle is described by a multiplicative model and assumed to be fully developed. The multiplicative speckle mode is defined as:

$$I(x,y) = I'(x,y) \cdot s(x,y) \tag{4.2-2}$$

where $I(x,y)$ is the intensity of image pixel i,j. The unbiased backscatter intensity is defined by $I'(x,y)$. The speckle noise is denoted by $s(x,y)$ and is modeled by a Gamma distribution (Lee, 1986).

A multitude of filter algorithms have been developed (Lee *et al*., 1994, for a review) and an appropriate method minimizes the speckle and preserves a maximum of image information at the same time. In several studies the impact of the filter method on the overall accuracy of a land cover classification was investigated (Capstick and Harris, 2001, Herold *et al*, 2005, Nyoungui *et al*. 2002). Although the conclusions are contrary in terms specific filter performances, the experimental results in these studies clearly demonstrate the positive effect of speckle reduction. In regard to multitemporal data sets a modification of the filter term is possible (Quegan *et al*., 2000). However, in this study a monotemporal filter approach was used, regarding the variety of the time series (i.e., different polarizations, sensors, and viewing angles). An *enhanced Frost* filter (Lopes *et al*., 1990) was applied to the imagery, using a 7×7 window. In Figure 4.2.1 a comparison between original and unfiltered images is shown.

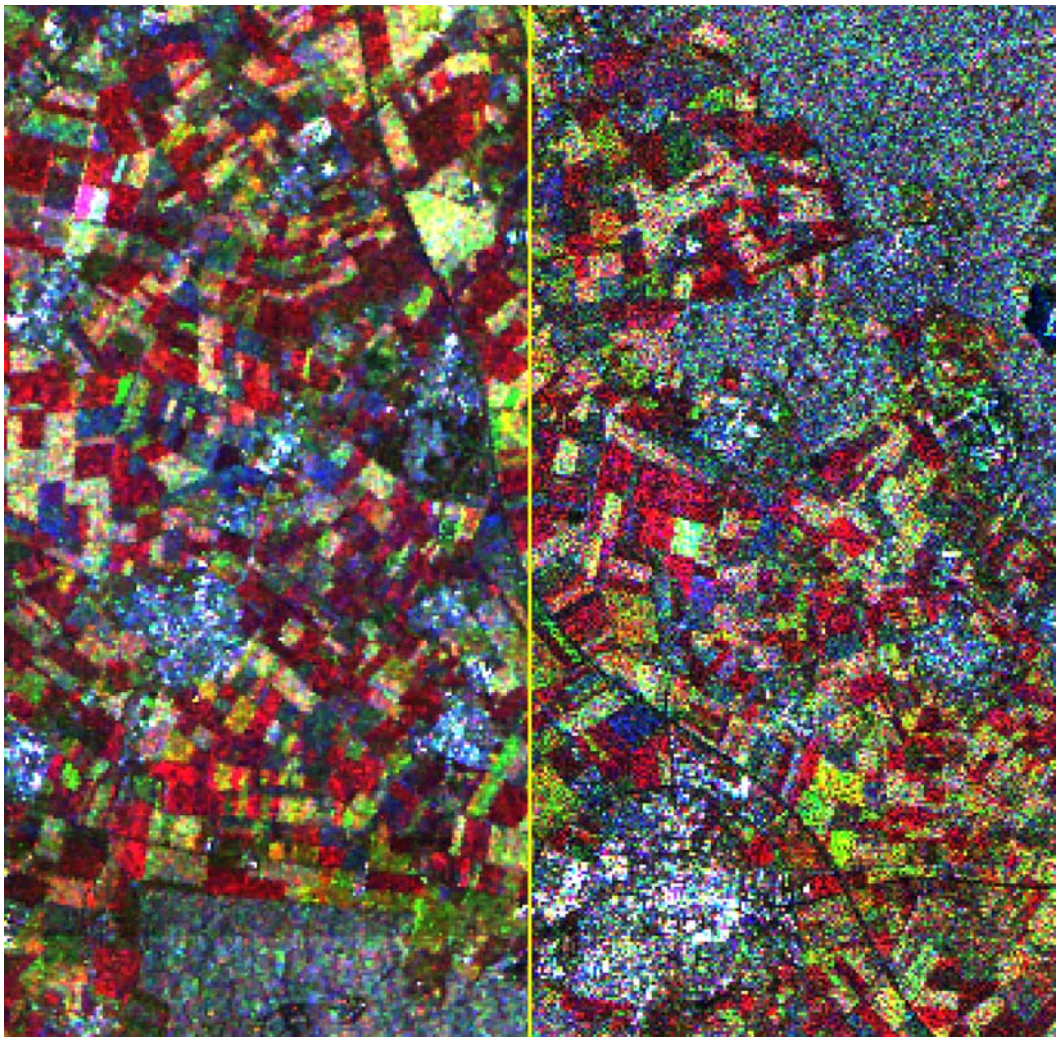

Figure 4.2.1: Multitemporal ERS-2 composite, comparison between filtered and unfiltered images.

4.2.3 Geometrical correction

For a precise geocorrection of the SAR imagery a digital elevation model and knowledge about the orbit of the sensor platform is required. Generally the orbit information is provided as orbit state vectors in the image product header. Furthermore a master scene with corresponding GCPs was used. The optical image, which was already employed for the geocoding of the multispectral imagery, was used as reference information. In doing so a more accurate spatial accuracy between the multisensor imagery is assumed. Using the (optical) master scene, GCPs were collected for the SAR imagery. Due to the prior image registration, this step must be performed only once for each set of images. Finally, the SAR images (12.5 m) were resampled to the pixel size of the Landsat TM scene (30 m) and orthorectified with a spatial accuracy of approximately one pixel, using a digital elevation model, orbit state vectors and the Landsat image as a reference information.

5 Multitemporal and multisensor land cover classification, applications and results

In the following section different multitemporal and multisensor applications are presented with their corresponding results: The first experiment is focusing on the performance of classifier ensembles, in the context of classifying multitemporal SAR data. In addition the value of multipolarized imagery is investigated. The subject of an adequate fusion of multitemporal SAR data and multispectral imagery is investigated in the second application. The application that is presented in 5.3 is dealing with the impact of image segmentation and discusses the value of a multilevel strategy. Finally a multisensor-multilevel fusion strategy is presented which is based on the combination of support vector machines and random forests. To conclude the application section, the prior introduced classification strategy is applied on another data set, to underline the aforementioned findings.

5.1 Classifying multitemporal SAR data by classifier ensembles

As stated before, monotemporal land cover classifications are often inefficient, due to great temporal differences in crop phenology, and multitemporal concepts are more appropriate in this context. Regarding upcoming SAR missions with increased revisit times and better spatial resolutions like TerraSAR-X or Radarsat-2, concepts that are based on multitemporal SAR imagery become even more promising. However, such future datasets with high temporal and spatial resolution might become very large. Although classifier ensembles used successfully for classifying multisource data and high-dimensional imagery, they have rarely been used for classifying SAR imagery and large multitemporal data sets.

In the following a random forests classifier is applied to a multitemporal SAR data set, containing Envisat ASAR and ERS-2 images. The investigation is focusing on (1) the general performance of classifier ensembles, (2) a comparison of random forests with different ensemble techniques and conventional methods, and (3) the impact of different polarizations on the total classification accuracy.

5.1.1 Dataset and Preprocessing

A dataset of 14 images from 9 acquisition dates, containing 5 Envisat ASAR alternating polarization and 4 ERS-2 precision images, was used for this experiment (Figure 5.1.1 and Table 5.1.1). Thus, the feature set, ranging from April to September, comprised information from varying phenological stages. Moreover the different polarizations provide dissimilar image information. The imagery was preprocessed as described in the aforementioned chapter (see Section 4). In this experiment the following eight classes were investigated: *Arable crops*, *Cereals*, *Forest*, *Grassland*, *Orchards*, *Rapeseed*, *Root crops*, and *Urban*. For classifier training a sample set of 300 samples per class was selected from the field survey areas using equalized random sampling as described in Section 2.5. Using the same sampling strategy an independent validation set was generated, including 500 samples per classes.

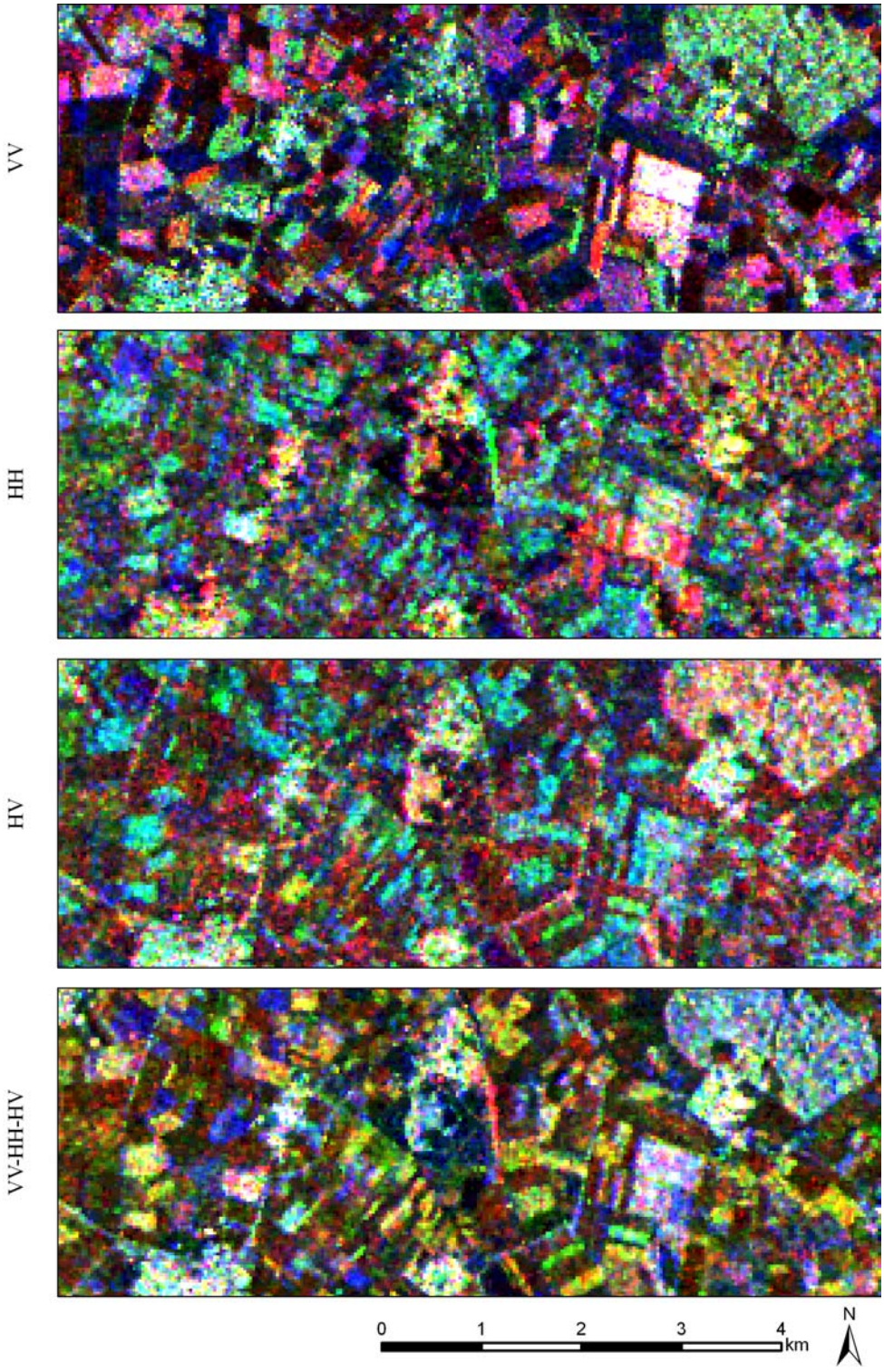

Figure 5.1.1: Subset of the multitemporal SAR data set, VV Datum (Apr. 21 / May 26 / Jun 30), HH (Apr.12/ Jul. 22 / Sep. 18), HV (Apr. 12 / Jul. 22 / Sep. 18), and HH-HV-VV (May 26 / Jul. 22 / Jul. 22).

Table 5.1.1: Image characteristics multitemporal SAR data set, 2005.

Sensor	Image characteristics SAR data			
	Date	*Track/Swath*	*Polarization*	*Orbit*
ASAR	12-Apr-05	6208	HH / HV	Asc
ERS-2	21-Apr-05	337	VV	Des
ERS-2	26-May-05	337	VV	Des
ERS-2	30-Jun-05	337	VV	Des
ASAR	13-Jul-05	3029	HH / HV	Asc
ASAR	22-Jul-05	7158	HH / HV	Asc
ERS-2	4-Aug-05	337	VV	Des
ASAR	14-Aug-05	2487	HH / HV	Asc
ASAR	18-Sep-05	2487	HH / HV	Asc

5.1.2 Methods

As stated in Section 2, classifier ensembles can be constructed using different approaches. In this study the performances of four different ensemble approaches was investigated: (1) bagging, (2) boosting, (3) random feature selection (rfs), and Breimans' random forests (RF). The random forests application is based on a FORTRAN code by Breiman and Cutler (2003). For the other ensembles the DT algorithm *C5.0* (Quinlan 1993) was used, which was used successfully in several remote sensing studies (e.g., Friedl *et al*. 1999; Simard *et al*, 2002). For matters of comparison, two common classification methods were applied to the data set, a maximum likelihood classifier (MLC) and a conventional decision tree (DT).

Beside the number of trees within the ensemble the feature partitioning methods (i.e., rfs and RF) requires the definition of the number of randomly selected features that are used for the entire tree and at each split respectively. In the beginning the feature subset size of the rfs and RF was increased stepwise and individual classifier ensembles were generated. In doing so the ideal feature subset size (in terms of classification accuracy) was verified. A subset size of 10 was used for the rfs approach. For the RF classifier a feature subset size of 3 was selected. This value is approximately equivalent to the standard value, which is often denoted as square root of the number of features (Gislason *et al*., 2006) In order to investigate the impact of the number of trees on the total accuracy, the size of each ensemble was increased stepwise (from know on referred to as *ensemble size*). To underline the value of multipolarization imagery the RF classifier was applied to individual image subsets, containing single polarizations (i.e., HH, HV, and VV) and combinations of these data sets (i.e., HH+VV).

5.1.3 Experimental results

The accuracy assessment shows that a classifier ensemble significantly improves the results of a conventional decision tree and performs better than a maximum likelihood classifier (49.7%). Whereas the overall accuracy of a simple decision tree is 51.8%, the classifiers ensemble already performs better, irrespectively of the generation strategy (Figure 5.1.2). Even ensembles of a small size (e.g., 10) outperform a simple DT in terms of total accuracy and achieve accuracies between 59.7% (RF) and 62.4% (boosting). Regarding the differences between the 4 ensemble methods, it can be assessed that in case of very small ensembles boosting performs slightly better than the other approaches. On the other hand the different concepts perform comparably, with a larger ensemble size, whereas random forests method outperforms the other approaches.

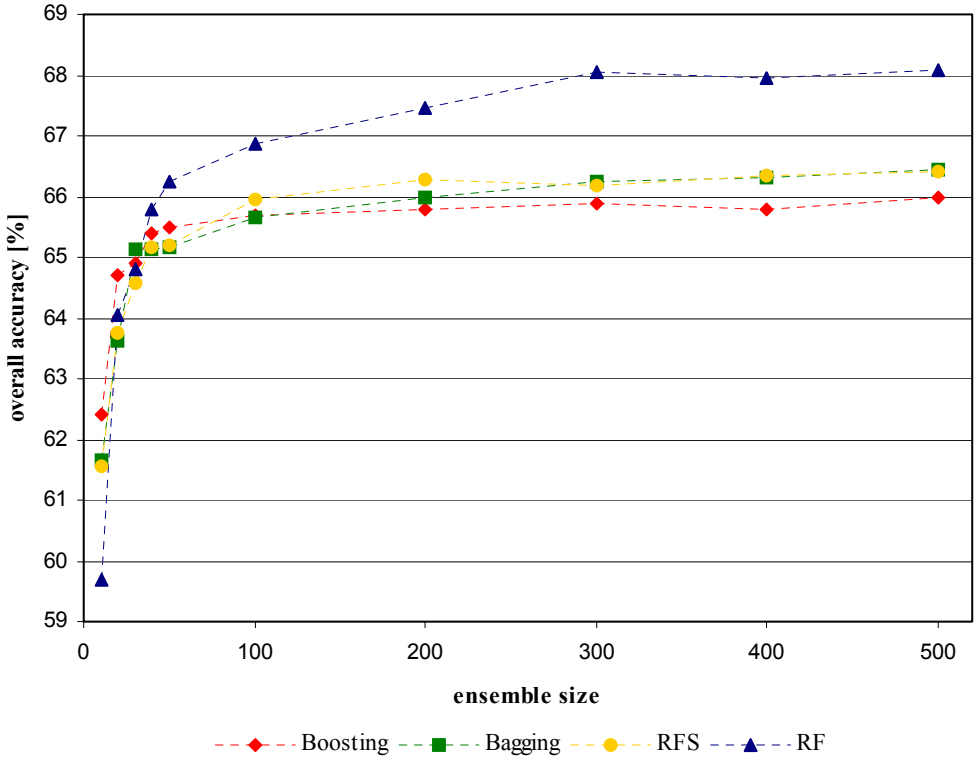

Figure 5.1.2: Overall accuracies [%], applying different classifier ensembles on the multitemporal SAR data set.

Regarding the class-specific accuracies, i.e., user and producer accuracy, the random forests achieve the highest accuracies in the most cases (Table 5.1.2). In general the classification accuracy for the land cover classes *Arable crops* and *Orchards* is lower compared to other classes. A reason for the lower accuracy could be the variability within the two classes. Both classes are characterized by spatial varieties, e.g. the surface beneath *Orchards* is often covered by meadow. Hence this class can appear as a mixture between the two classes *Grassland* and *Forest*. In contrast to this, classes like *Cereals* and *Root crops* are less variable and consist mainly of winter wheat and sugar beets, respectively.

These classes consequently were classified with relatively high accuracies. A detailed investigation of the impact of different polarizations shows the worth of the multisource data set, consisting of ERS-2 and Envisat ASAR data of various polarizations (Table 5.1.3). The random forests classification results, which are based on a single-polarization, vary between 43% and 49%. In contrast to that, the accuracies achieved with two polarizations range between 59% and 62%. However employing all data (i.e, three different polarizations) the overall accuracy is further increased by 6%. The positive impact of different polarizations was also confirmed by other studies (e.g., Del Frate *et al*., 2003; Stankiewicz, 2006). Beside the advantage of multipolarized imagery, the combination of the ASAR and ERS-2 products also increases the multitemporal information content. Hence, one might argue that the improved classification accuracy is a result of the additional multitemporal information. However, the ASAR alternating polarization imagery that provide two polarizations with a single acquisition, clearly demonstrate the positive effect of two polarizations. Whereas the individual classification results of the HH and HV data are 46% and 49% respectively, the total accuracy is increased by up to 10% by a combination of both polarizations. An assessment of the individual class-specific accuracies confirms the advantage of multipolarized data sources (Table 5.1.2): *Cereals* are best discriminated by the VV polarization, whereas the highest accuracies for *Grassland* and *Urban* are based on the HH data. The 5 other land cover classes are best classified by the HV imagery.

Table 5.1.2: Class-specific accuracies [%], using different classifier ensembles, with a size of 500.

Land cover class	**Boosting**		**Bagging**		**RFS**		**random forests**	
	prod. acc.	*user acc.*	*prod. acc.*	*user acc.*	*prod. acc.*	*user acc.*	*Prod. acc.*	*user acc.*
Arable crops	56.0	55.6	**63.0**	54.4	57.8	**57.9**	62.8	54.7
Cereals	73.0	66.0	73.4	**66.6**	72.8	65.2	**74.6**	65.3
Forest	76.2	70.9	80.0	69.0	78.4	69.0	**82.0**	**71.1**
Grassland	72.0	71.3	71.4	73.0	73.0	70.3	**73.2**	**74.2**
Orchards	49.6	54.3	50.4	56.5	47.8	55.8	**51.2**	**59.8**
Rapeseed	**68.4**	71.5	66.2	70.6	67.4	70.8	68.0	**73.4**
Root crops	67.8	60.2	65.2	64.9	**68.0**	64.9	66.8	**66.5**
Urban	64.8	80.4	61.8	80.5	65.8	77.2	**66.0**	**84.2**

Table 5.1.3: Overall accuracy [%], applying RF on different single- and multi-polarization datasets.

Polarization	**Overall accuracy**
ERS-2 VV	43.4
ASAR HH	45.9
ASAR HV	49.0
ASAR HH-HV	58.6
ASAR / ERS-2 HH-VV	60.2
ASAR / ERS-2 HV-VV	61.7
all data	68.1

An assessment of the individual class-specific accuracies confirms the advantage of multipolarized data sources (Table 5.1.4): *Cereals* are best discriminated by the VV polarization, whereas the highest accuracies for *grassland* and *urban* are based on the HH data. The 5 other land cover classes are best classified by the HV imagery. These findings correspond to other studies: Ferrazzoli (2001) mentioned that VV polarization is more adequate for monitoring of crops that are dominated by vertical structures, as wheat and barley, whereas HV is more applicable for ramified crops, as root crops and rapeseed. Brisco *et al.* (1992) have underlined the general value of cross-polarization for discriminating crop types.

Table 5.1.4: Class-specific accuracies, using different polarizations.

Land cover class	**HH**		**HV**		**VV**	
	prod. acc.	*user acc.*	*prod. acc.*	*user acc.*	*prod. acc.*	*user acc.*
Arable crops	35.4	30.4	**38.6**	**34.6**	31.4	32.0
Cereals	31.4	31.7	39.4	38.7	**67.4**	**58.9**
Forest	51.6	40.7	**69.6**	**57.3**	38.2	32.9
Grassland	**69.2**	**63.3**	47.6	48.9	60.8	52.6
Orchard	28.8	36.1	**46.2**	**42.1**	25.6	31.9
Rapeseed	44.0	47.8	**60.4**	**62.0**	34.0	38.1
Root crops	42.4	49.1	**52.2**	**52.2**	50.6	45.3
Urban	**64.2**	**71.3**	38.2	63.0	38.8	51.9

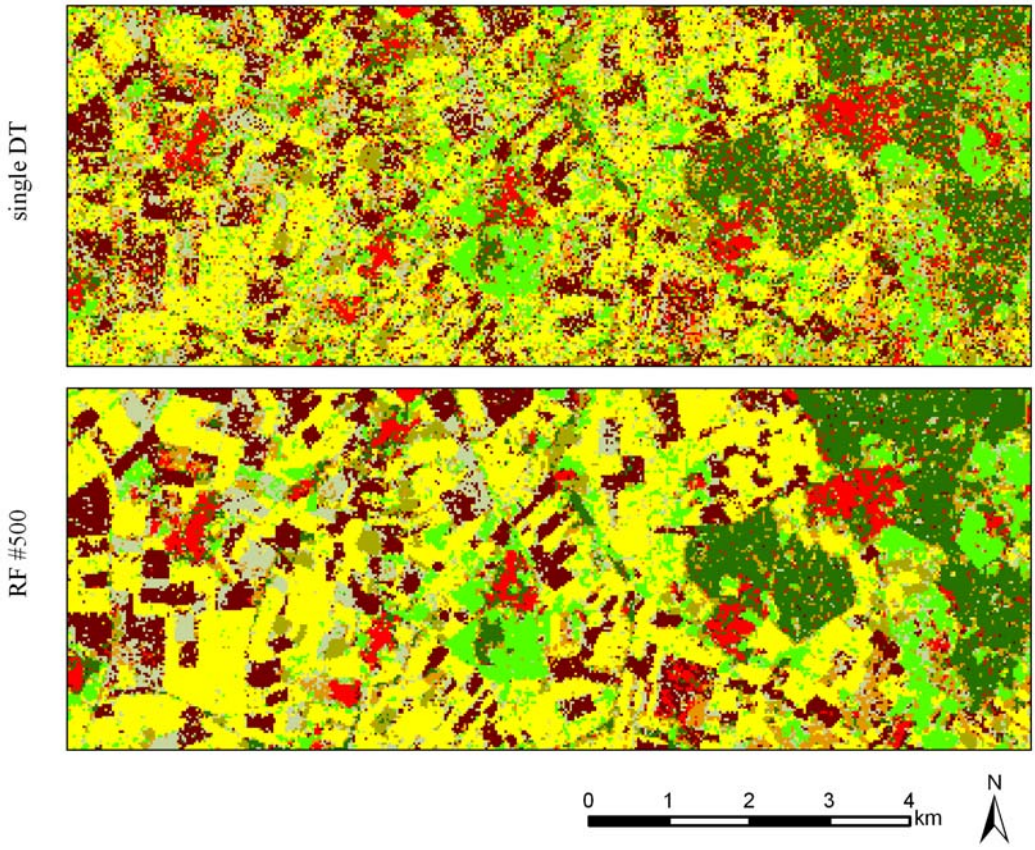

Figure 5.1.3: Classification result, using a single DT and random forests with an ensembles size of 500.

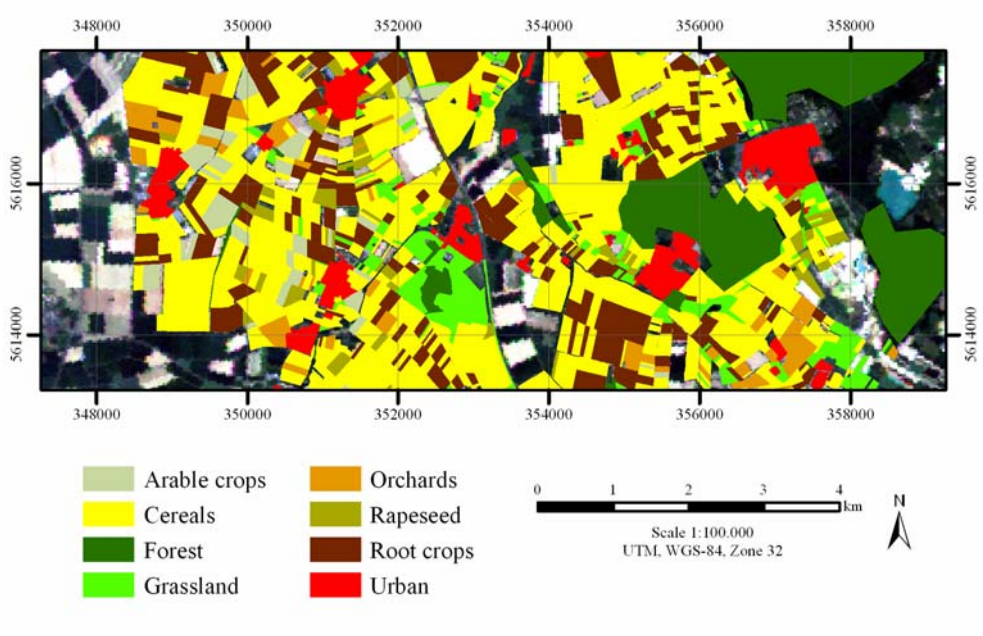

Figure 5.1.4: Corresponding subset of the land cover map 2005.

A visible comparison (Figure 5.1.3) of the classification maps underlines the improvement of accuracy: The map produced by a simple decision tree shows the general structures of the study site. Nevertheless parcels that are homogeneous in reality appear very noisy, sometimes to a degree that the true land cover class can not be assigned to the entire area. Sharp boundaries along individual fields appear blurred; consequently they are hard to be to identify. The confusion and noise is significantly reduced by random forests. Almost all features can be assigned to a specific land cover and edges appear clear.

5.1.4 Discussion and Conclusion

The investigation clearly shows that results from classifier ensembles are superior to those from a simple decision tree. Regardless of ensemble size and the method used for the generation of classifier ensembles a higher accuracy is reached. Comparing the results achieved by different classifier ensembles, it can be assessed that random forests are superior to other concepts in terms of overall accuracy. The class-specific accuracy assessment shows that the approach achieves the highest user and producer accuracies for almost each class.

The main reason for the success can be the underlying assumption of classifier ensembles. As mentioned above (see Section 2), the strategy requires the generation of independent classifiers and the performance is directly influenced by this independency. Given the relative independence of images from different acquisition times, a concept which is based on a random selection of the input features (i.e., rfs and RF) seems well suited for multi-temporal approaches. The independence of the individual classifiers and thus the performance of the ensemble are further increased by combining this feature selection with a random selection of training samples, as done in RF. Boosting performs slightly better with a very low number of decision trees. On the other hand boosting must be computed sequentially. Hence it is computationally much more costly than random forests.

In the visual interpretation of the produced maps the positive effect of the statistical assessment is more than confirmed: the degree of noise is significantly decreased and the image is classified into more homogeneous areas. The differences between the maps from the simple decision trees and classifier ensembles are comparable to the differences between pixel-based and segment-based approaches. Classifier ensembles – particularly random forests – thus appear very well suited for noise inherent SAR data. The impact of different polarizations on the classification accuracies demonstrates that the various data sets contribute unequally to the classification of different land cover classes. Hence, the different sources appear not equally reliable. The good results for separating agricultural classes, which are hard to be distinguished in monotemporal imagery, generally underline the importance of multitemporal analyses as well as the positive impact of different polarizations.

5.2 Classifying multisensor data by support vector machines

As stated in the prior sections, SVM often outperform other algorithms in terms of classification accuracies, even when classifying high-dimensional data sets with small training sets. Nevertheless only a few approaches are known, which use SVM for classifying multisource datasets: Whereas Fukuda and Hirosawa (2001) and Song *et al*. (2005b) used "conventional" SVM for classifying multisource data, in other applications the SVM approach was modified for classifying diverse datasets: Halldorsson *et al*. (2003) extended a radial basis kernel function for a combined classification of Landsat MSS imagery and topographical data. Camps-Valls *et al*. (2006) proposed so-called composite kernels (i.e., a combination of different kernel functions) for the classification of spectral and spatial information of a hyperspectral imagery. Fauvel *et al*. (2006) applied two individual SVM classifiers to the original spectral data set and the corresponding spatial information, which was derived by using extended morphological profiles. The SVM were trained separately on the two sources, the spectral image and the spatial information. Afterwards the different outputs were fused by various voting schemes, e.g., absolute maximum and majority voting. Overall the results of these studies have shown that the performances of a SVM classifier can be improved by modifying the kernel functions or using separate SVM when classifying diverse datasets. Thus, in case of multisensor imagery, it could be more applicable to handle various sources separately.

In the following a multisensor classification strategy is presented, which is based on support vector machines. The individual image sources, i.e., the SAR data and the multispectral image were preliminary classified, using conventional SVM. For the subsequent fusion process the outputs of the SVM classifiers are combined by voting strategies, i.e., majority voting and the absolute maximum rule. In a more sophisticated decision fusion approach the outputs of the SVM are combined by an additional SVM classifier to generate a multisensor classification. The investigation is focused on (1) the general impact of multisensor imagery on the classification results, (2) the performance of support vector machines for classifying multitemporal/multisensor imagery, and (3) the potential of an alternative modified SVM-based classification strategy.

5.2.1 Data set and preprocessing

For the application a multisensor data set from 2005 was used, consisting of multitemporal SAR data and a Landsat image from May-26 (Figure 5.2.1). The SAR imagery contains the same Envisat ASAR and ERS-2 data that were already used in Section 5.1 (see Table 5.1.1). The data was preprocessed, following the methods introduced in Section 4. The multisensor data set was coregistered with a spatial accuracy of approximately one pixel. In addition the images were normalized before the SVM training, to simplify the parameter selection for the kernel function.

For the presented investigation, two different training sets for 8 land cover classes, i.e., *Arable crops*, *Cereals*, *Forest*, *Grassland*, *Orchards*, *Rapeseed*, *Root crops*, and *Urban*, were generated, using equalized random sampling. Each class has the same sample size, containing 50 and 300 samples per class, respectively (from now on referred to as training set #50 and training set #300). In doing so, the impact of the samples size on the classifier performance can be investigated. Using the same sampling strategy as before, an independent validation sets was generated, containing 4000 samples, 500 of each class.

The SVM was applied on three different data sets: (1) the multitemporal SAR data set, (2) the Landsat TM image, and (3) the multisensor data set, containing both sources. In addition different fusing schemes are presented.

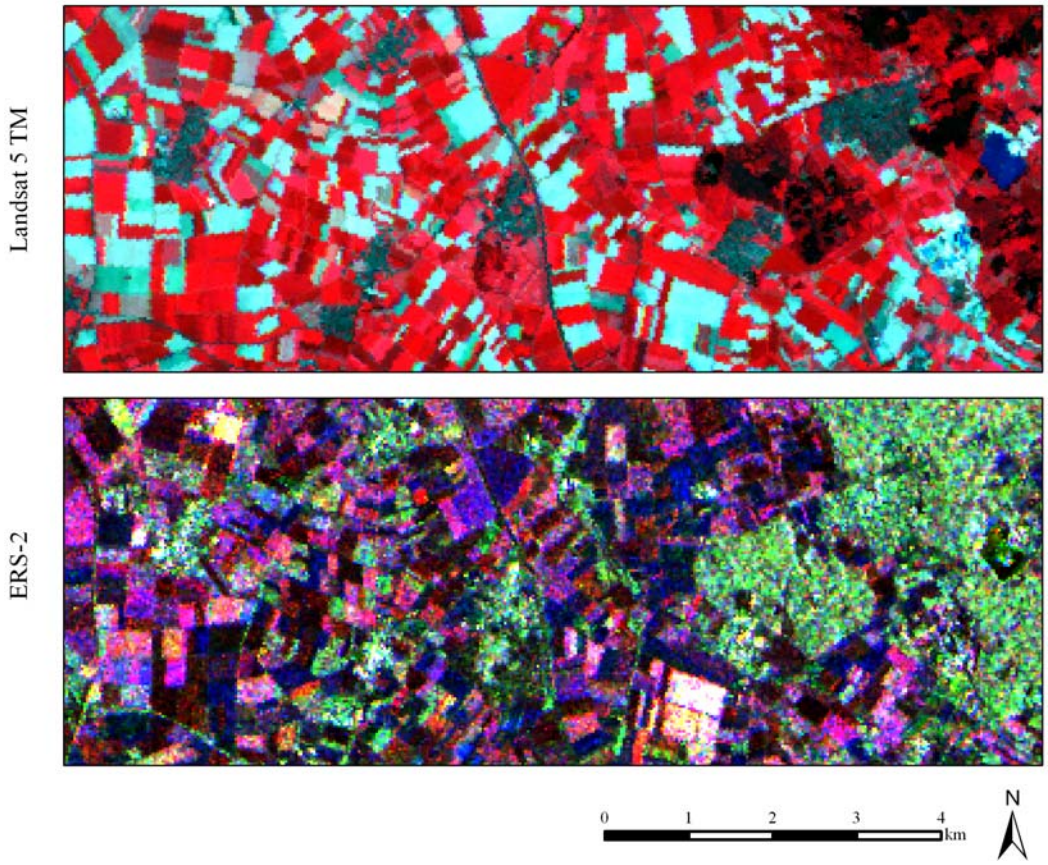

Figure 5.2.1: Subset of the multisource data set 2005. Multispectral false-colour composite (bands: 4/3/2) and multitemporal ERS-2 composite (Apr-21 / May-26 / Jun-30).

In the experiments a Gaussian kernel was used, which was employed in several remote sensing studies (e.g., Huang *et al*., 2002; Melgani and Bruzzone, van der Linden; 2007). In comparison to a polynomial kernel, which requires more parameters than a Gaussian kernel, the computational complexity of model is decreased. In contrast to a simple linear kernel, which is a special case of the Gaussian kernel, a Gaussian RBF function can handle more complex, non-linear class distributions (Keehrti and Lin, 2003). Thus the use of the Gaussian kernel seems adequate in context of this study. As state before the SVM training requires the estimation of the kernel parameter γ and C, which is normally solved by approaches, which based generally on a leave-one-out procedure.

In the presented application, the training of the SVM with Gaussian kernel and the generation of the rule images were performed using *imageSVM* (Janz *et al*., 2007). *imageSVM* is a freely available IDL/ENVI plug-in that is based on the LIBSVM approach by Chen and Lee (2001) for the training of the SVM. The strategy determines the best values for γ and C within a user defined range of possible parameters using a 10-fold cross validation. In doing so, the outputs of the final discriminant function (i.e. rule images) were calculated for each binary classification

problem. Following the common OAO procedure (i.e., majority vote) the final classification results were generated for the three data sets (single SVM). Although the OAO approach results in a larger number of binary SVM classifiers, the approach was selected, because the classification problem as a whole is separated into many simpler ones. The OAA strategy on the other hand, which is based on a separation between a specific class and the rest, can involve the estimation of a more complex hyperplanes. Regarding the 8 land cover classes, the application of the SVM with the OAO strategy generates 28 rule images per data set (see Section 2.2.3).

Beside the application of one single SVM on the whole multisource data set, the pre-classifications of the two single-source SVM were combined to create the final result. For this purpose, three different decision fusion strategies were applied to the rule images: (1) a simple majority vote, (2) an absolute maximum rule and (3) an additional SVM.

Similar as in the conventional majority vote, which is used for the OAO multiclass strategy, all positive and negative votes for a specific class are summed (see Section 2.2.3). Contrary to the single-source classification, the rule images of both sources (i.e., SAR and Landsat) are used for the multi-source majority voting. The second fusion approach – the absolute maximum rule – is based on the concept which is used in the conventional OAA strategy. The final class is the one with the highest absolute value (i.e., distance to the hyperplane). The original OAA decision rule was modified for the multisensor classification problem, similar to what was done in Fauvel *et al*. (2006): For each binary two-source classifier, separating the classes ω_i and ω_j ($\omega_i \neq \omega_j$), the rule images of the two SVM (i.e., f_{SAR} and f_{TM}) were compared. The absolute maximum value within the rule images (i.e., maximum distance) determines the decision for this two-source classification problem (5.2-1). Afterwards, a simple majority vote is applied to derive the final class membership.

$$F_{\max}(x) = AbsMax\left|f_{SAR}(x), f_{opt}(x)\right| \tag{5.2-1}$$

For the third fusion concept all information from the two rule image sets (i.e., 2×28) are simply combined in to a single data set, containing 56 rule images. In doing so a feature vector is generated that includes the outputs $f_{sar}(x)$ and $f_{opt}(x)$ of the two individual SVM classifiers. To determine the final class membership an additional SVM (SVM fusion) was trained on this data set (Figure 5.2.2). Besides the SVM, three different classifier algorithms were applied on the data sets: (1) MLC, (2) DT, and (3) RF. Contrary to the SVM fusion, which performs the data fusion at decision level, the other approaches perform the fusion at the data level (maximum likelihood classifier) and at the feature level respectively (decision tree and random forests). A comparison with these algorithms is worthwhile, because of this dissimilarity and also because of the numerous studies that are uses the well-known methods.

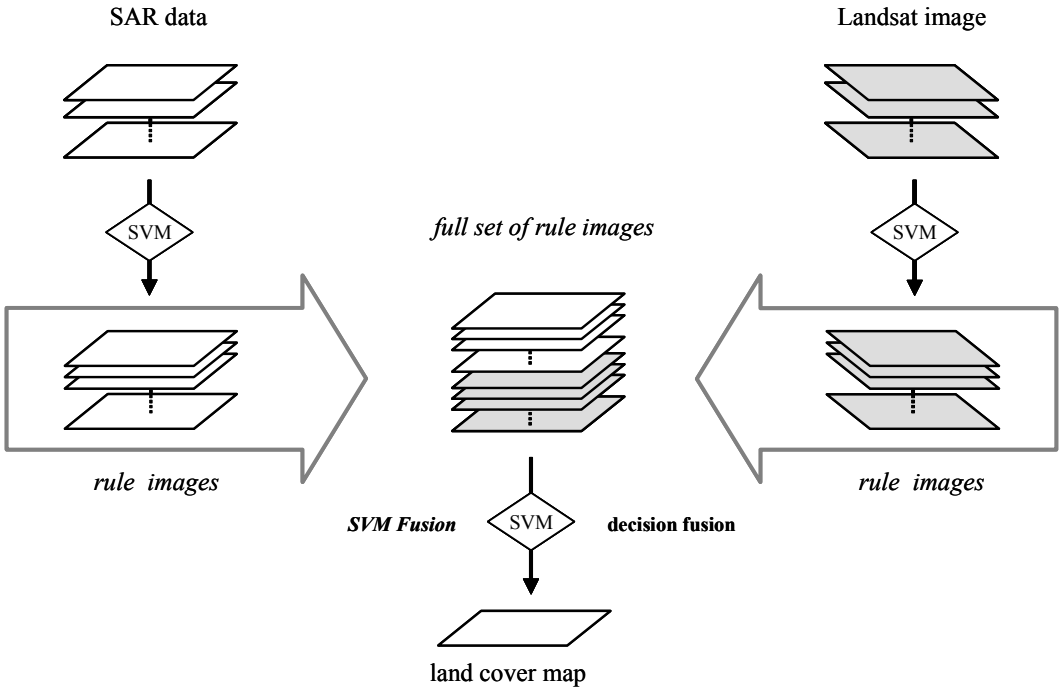

Figure 5.2.2: Schematic diagram of the SVM-based decision fusion

5.2.2 Experimental results

The experimental results clearly show the positive impact of a multisensor data set. Independently of the training sample set, the overall accuracy is significantly increased by using multisensor imagery (Table 5.2.1). Comparing the multisource-based results of the single SVM, the total accuracy increased by 10% compared to the classification results achieved on the optical images (training set #300) and by 14% when compared to the accuracy achieved on the SAR data (training set #50).

Table 5.2.1: Overall accuracies [%], achieved by SVM, using different data types and training sets.

Data Set	Training set	
	#50	*#300*
SAR	58.5	64.7
TM	63.8	68.0
SAR + TM	72.6	78.3

Comparing the different methods for classifying multisensor imagery, the results show that the single DT performed worst in terms of accuracy. The SVM outperform the other algorithms in terms of accuracy, except from the RF, which generally achieved the highest overall accuracies. Hence random forests could seem more adequate for classifying the multisensor imagery. On the other hand the performance of multisource SVM can further increased when the SVM training process is conducted separately for the two image sources and an adequate fusion technique is used (see Table 5.2.2).

The proposed concept – which combines the individual rule images of the pre-classification by an additional SVM – outperforms the other algorithms and fusion strategies in terms of accuracies. Compared to the single SVM that was trained directly on the whole multisource dataset the overall accuracies are increased by up to 3%. Compared to the RF, the differences between the overall accuracies are less dominant, but the fusion approach performs slightly better. In contrast a fusion by the simple majority seems inefficient and the overall accuracy even drop below the accuracy achieved by a single SVM, whereas the application of the absolute maximum rule is more efficient and increases the total accuracies.

Table 5.2.2: Overall accuracies [%] for multisensor data, using different classifiers and fusing schemes.

Classifier / fusion strategy	**Training set**	
	#50	*#300*
MLC	64.8	74.0
DT	61.2	69.1
RF	74.4	79.6
Single SVM	72.6	78.0
Majority vote	72.5	75.6
Absolute maximum	74.3	78.6
SVM fusion	75.6	80.4

In general the experimental results clearly demonstrate that the overall accuracy is improved by increasing the number of training samples. Whereas the SVM fusion is increased by approximately 5%, the MLC is improved by 9%, due to a larger sample set.

Regarding the class-specific accuracies (i.e., producer and user accuracy), the experimental results show the positive effect of the proposed fusion technique. This impact is particularly obvious for classes, which are classified with relatively low accuracies by both single-source classifications. Compared to the SAR data and the Landsat image the producer accuracy of *Arable crops* is increased by 15% and 12% respectively, that of *Orchards* is increased at least by 18%. The producer accuracy of these classes is increased by 9.8% and 6.2% using the SVM fusion concept, in comparison to the application of a single SVM on the full multisensor dataset. Contrary, there is a reduction in accuracies (1.2% - 2.4%) for a few classes (*Cereals*, *Forest*, *Grassland*) comparing the proposed strategy with a single (multisource) SVM. But those classes were still classified with relatively high accuracies. Another advantage of the multisensor fusion concept is that the achieved producer and user accuracies are less variable (Table 5.2.3 and Table 5.2.4): Whereas the application of the single SVM results in standard deviations of 9.8 and 11.7 respectively, the SVM fusion reduces the values to 6.7 and 8.5. Thus the results are more homogeneous over the different land cover classes. Moreover compared to the results, which are based on a single data source, the producer and user accuracies of the multisensor classification are more balanced.

Table 5.2.3: Class-specific accuracies [%] of single-source results, using SVM and training set #300.

Land cover class	Producer accuracies		User accuracies	
	SAR	*TM*	*SAR*	*TM*
Arable crops	59.6	53.4	50.5	59.2
Cereals	73.4	80.2	69.0	59.9
Forest	78.6	92.2	69.2	95.3
Grassland	67.8	41.8	71.8	54.0
Orchards	55.2	45.4	50.5	44.9
Rapeseed	59.4	74.2	65.9	79.4
Root crops	62.2	77.4	67.9	66.4
Urban	61.8	79.2	80.9	87.4

Table 5.2.4: Class-specific accuracies [%] of multisource results (SAR+TM), using single SVM applied on the whole data set and SVM fusion, using training set #300.

Land cover class	Producer accuracies		User accuracies	
	Single SVM	*SVM fusion*	*Single SVM*	*SVM fusion*
Arable crops	65.4	75.2	65.1	73.5
Cereals	82.4	80.8	76.3	80.4
Forest	94.6	92.2	94.4	93.7
Grassland	73.6	72.8	77.6	81.9
Orchards	67.2	73.4	60.7	67.4
Rapeseed	82.4	84.2	88.2	88.0
Root crops	73.4	79.4	75.8	74.9
Urban	85	85.4	88.7	85.5
standard deviation	*9.9*	*6.7*	*11.7*	*8.5*

The visual assessment interpretation (Figure 5.2.3) clearly shows a high degree of noise in the SAR-based classification map. In contrast to this, the classification that is based on the Landsat scene appears more homogenous. The comparison of the single-source maps shows some divergences within the results, underlining the fact that the two sources provide different information. The fused map on the other hand, appears as a mixture of both single-source maps. Contrary to the visibly good performance of the multisensor classification, the elimination of the highway in the center of the image, which was classified correctly in the multispectral data, might be considered as a possible disadvantage of a multisensor approach. In Figure 5.2.4 the final classification result is shown.

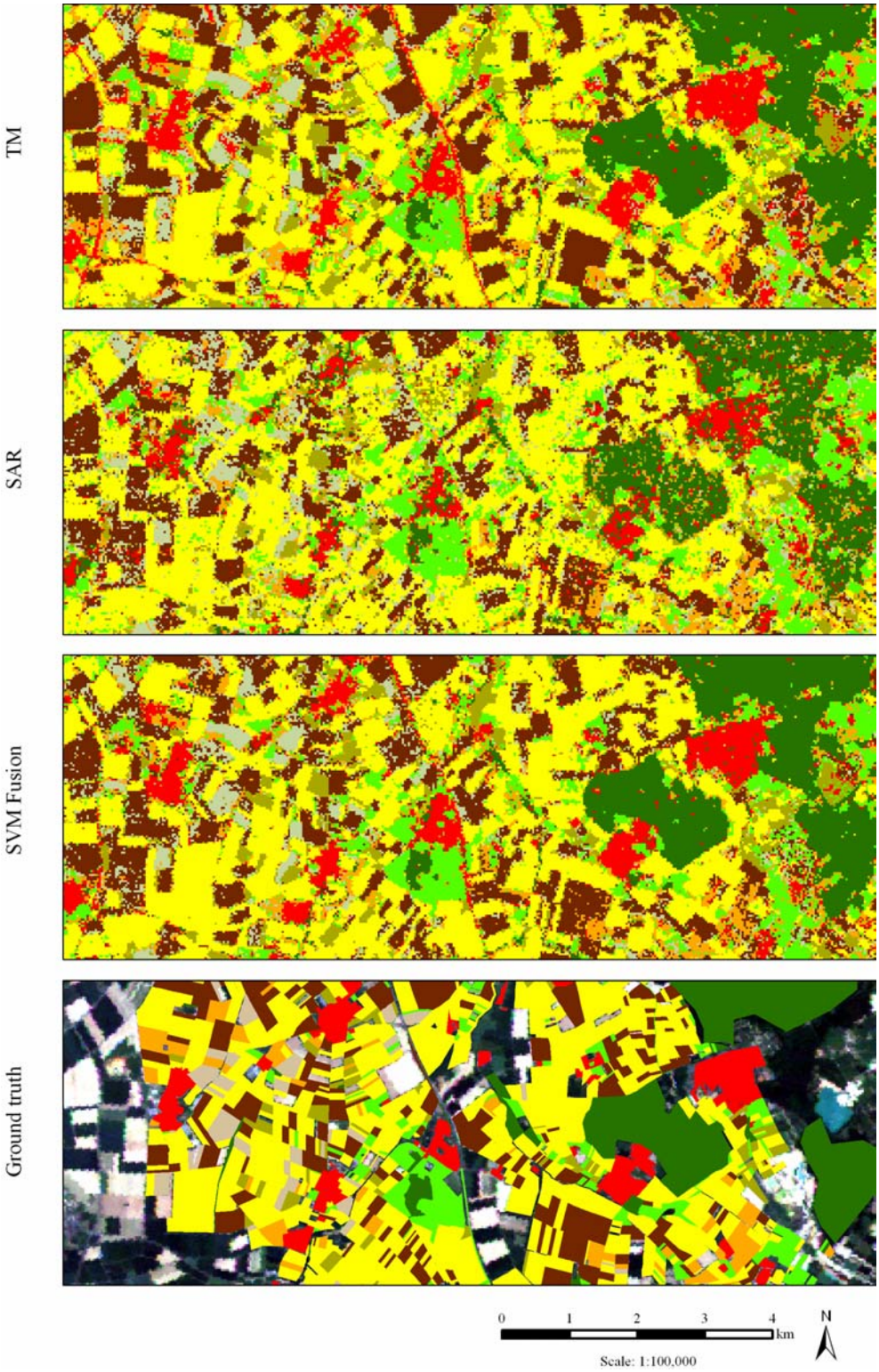

Figure 5.2.3: Ground truth information (legend see Fig. 5.2.4) and classification maps of the two single-source SVM classifications and the SVM fusion result

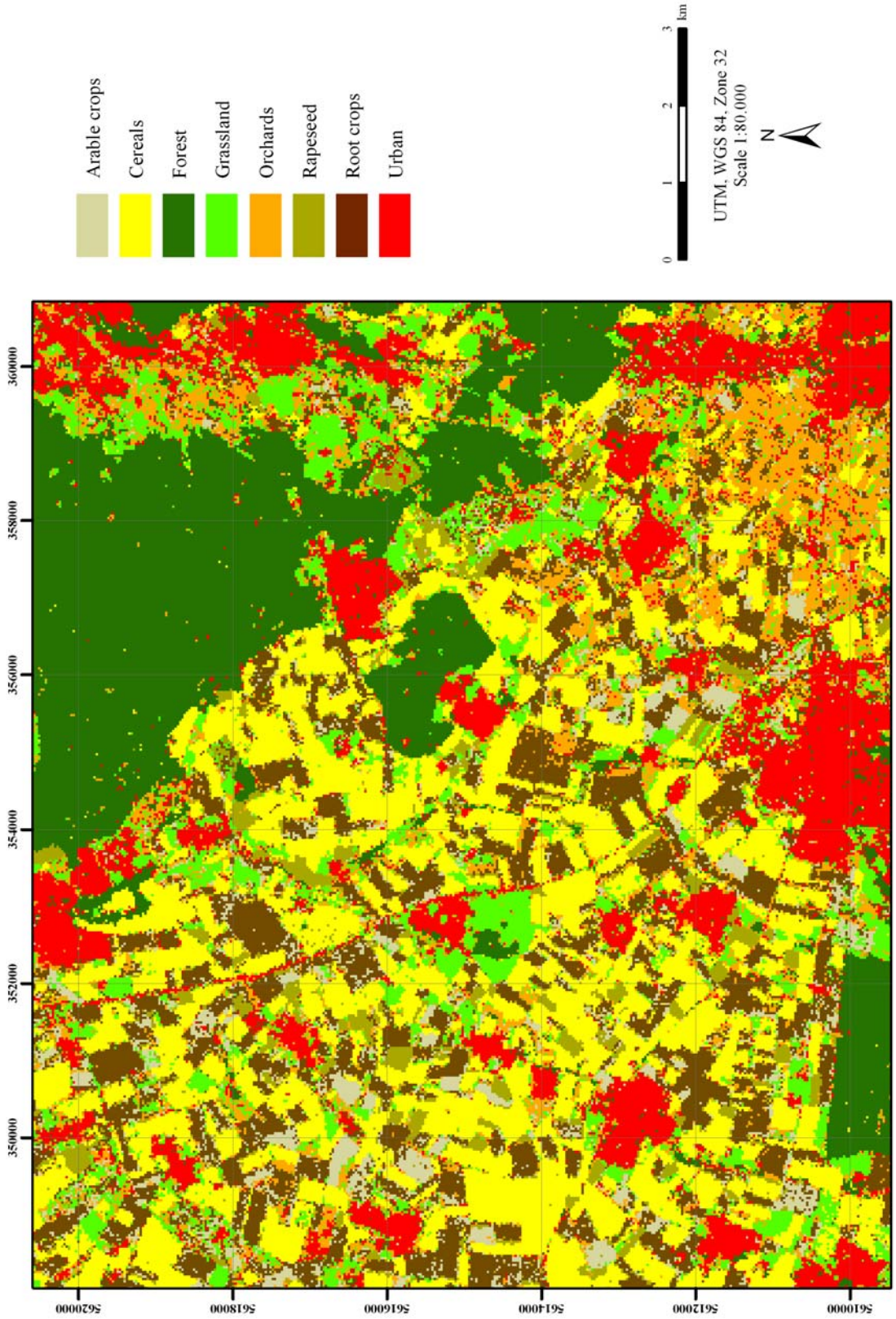

Figure 5.2.4: Classification result, using SVM fusion.

A detail visual interpretation of the classification results that are based on different data sources, clearly illustrates the different and complementary information of the imagery (Figure 5.2.5 and Figure 5.2.6). Using the TM image, the *Arable crops* parcels are incorrectly classified as *Root crops* and some fractions as *Urban* areas. In the SAR data, these regions are almost correctly classified as *Arable crops*, and the data fusion eliminates most of the errors achieved in the multispectral result. Within the largest *Arable crops* field, some pixels in the TM image are misclassified as *Urban* land cover, whereas these pixels are correctly classified by the SAR data. Contrary to this, pixels that are accurately recognized in the TM image are classified as *Orchards* in the SAR data. This confusion is eliminated by the proposed fusion concept. The classification results within *Grassland* and *Orchards* regions are other selected examples to underline the value of multisensor image analysis (Figure 5.2.5).

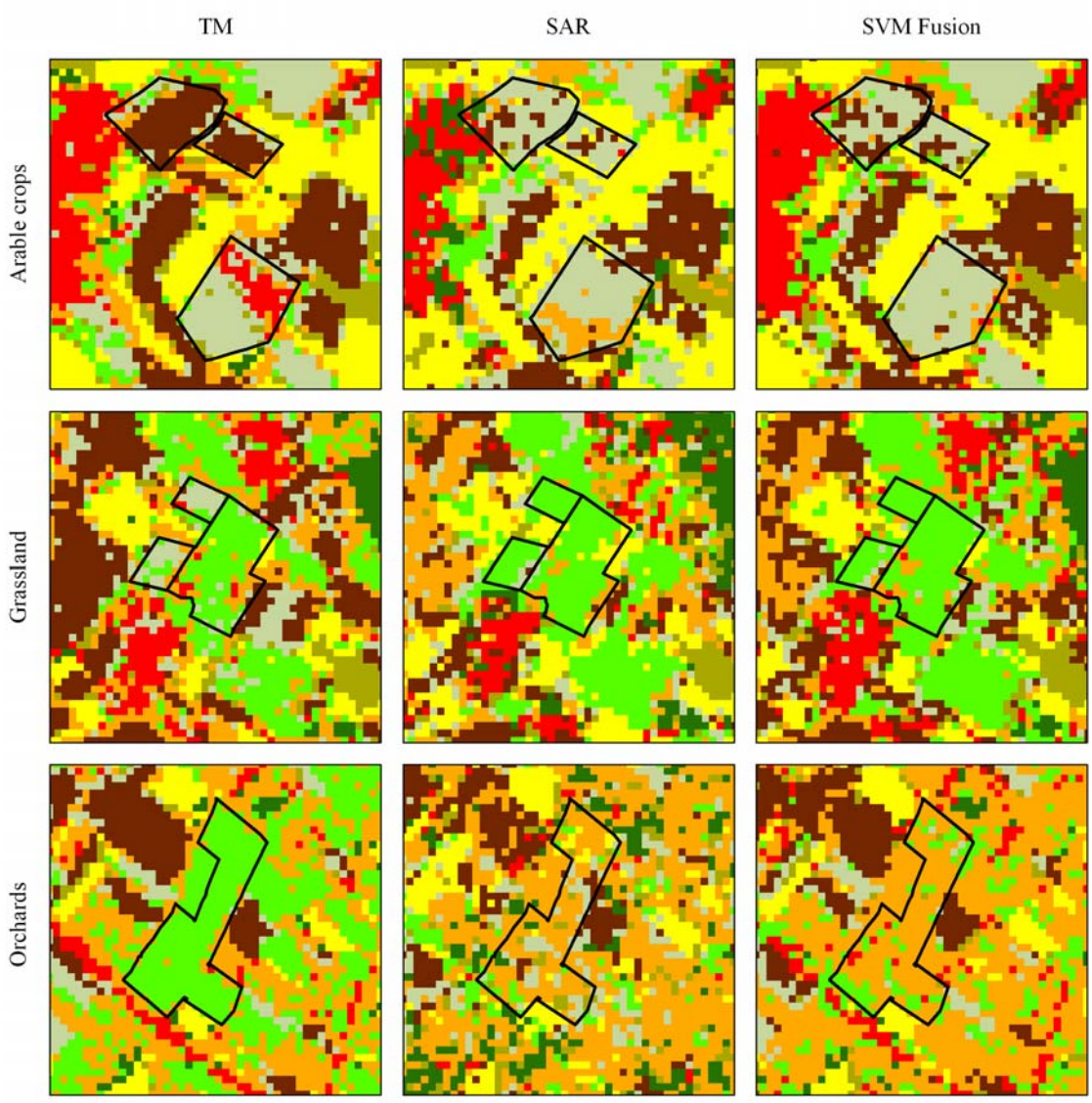

Figure 5.2.5: Comparison between single-source and fused classification results. The polygons indicate parcels of *Arable crops*, *Grassland*, and *Orchards*.

The TM-based classification results in a misclassification of *Grassland* pixels and classifies the two smaller patches as *Arable crops*. The pixels within the large *Orchards* region are erroneously assigned to *Grassland*. Contrary to this, the SAR data recognized most pixels within these regions accurately and thus improves the final classification results.

The classification outputs within *Forest* and *Urban* regions visualize the positive impact of the TM data, underlining the inherent noise in the SAR data on the one hand and the homogenous outputs of the Landsat image on the other (Figure 5.2.6). The noise within the SAR-based classification, which is mainly confusion between *Forest* and *Urban*, is significantly reduced by the multisensor fusion. Moreover the regions appear more homogenous than the TM-based classification.

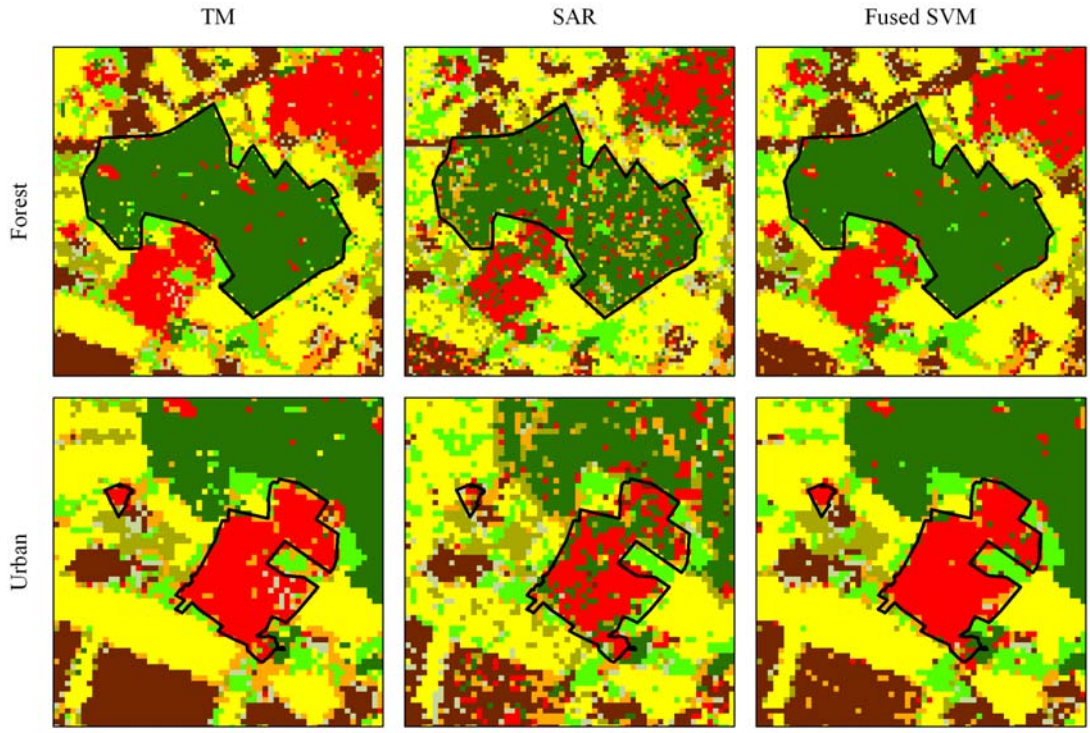

Figure 5.2.6: Comparison between single-source and fused classification results. The polygons indicate parcels *Forest* and *Urban*.

5.2.3 Discussion and conclusion

In this chapter, a strategy for an improved classification of multisensor imagery, consisting of multitemporal SAR data and a multispectral image, was introduced. The SVM fusion is based on the decision fusion of support vector machines that were individually trained on different image sources, i.e., the SAR data and Landsat image. In general, SVM might be more accurate in classification of complex data sets when compared to conventional multivariate classifiers, since a convenient statistical model is often not known for such imagery. Furthermore the performance of a statistical classifier as the MLC is more dependent on the size of the training sample set. These assumptions are conformed by the experimental results.

The accuracy assessment shows that the separate training of the SVM and a subsequent fusion of the pre-classified outputs by another SVM outperforms the other parametric and non-parametric classification methods including random forests and a single SVM, which was

trained on the full multisensor dataset. Particularly the accuracies of land cover classes, which are critical in terms of accuracy, were significantly increased.

The main reason, that the proposed classification strategy outperforms the single SVM in terms of accuracy, could be the heterogeneous multisensor imagery. That multitemporal SAR data and optical imagery provide different information is clearly shown by the class-specific results of the single-source classifications. In addition the imagery may not be equally reliable and one image-source (e.g. a specific SAR acquisition or band of the Landsat image) might be more adequate to describe a class, and perhaps another source seems more adequate for another class. The application of a single SVM for the full multisensor dataset requires the definition of one specific kernel function. Regarding the heterogeneous imagery, it seems more adequate to define the kernel functions for each data source individually and combine the derived outputs in a subsequent process.

In general it can be assessed that SVM and RF classifiers perform well on the multisensor imagery. The “simple” RF can be pointed out as a simple yet accurate approach. The results are further improved by the proposed decision fusion strategy. It is assumed that a modification of the presented conceptual framework would further increase the classifier performances.

In other studies the total accuracy was imporved by using segment-based approaches, which seems also interesting for the study site in this dissertation. Hence in the following step the fusion strategy is further modified and transferred to a diverse dataset, which consist of multisensor imagery at different levels of segmentation.

5.3 Classifying multisensor data by a multilevel decision fusion concept

As introduced before, one development in the classification of remote sensing imagery is that of segment-based approaches. On the one hand segment-based approaches can overcome problems of pixel-based classifications. On the other hand the definition of a single segmentation level might be critical, and the application of multilevel concepts seems more adequate.

In the following, the fusion strategy introduced in Section 5.2 is modified and extended to a multilevel component. The approach is applied to a multisource dataset, consisting of multitemporal SAR data and a Landsat image at various segmentation scales. Image segmentation is performed separately for each source and each segmentation level is separately classified with support vector machines. Afterwards the decision fusion is performed. Beside an additional SVM, the fusion is also performed by random forests. Both methods perform well in the prior presented applications. Hence the concept is assumed to make an efficient utilization of both, the ability of SVM to model complex class distribution in a high dimensional feature space and the strength of RF to utilize diverse combinations of input features for optimized decision making. The results are compared to the results achieved by conventional SVM and RF, which were applied on the whole multisource-multilevel dataset. In the prior section it was clearly demonstrated that the maximum likelihood classifier or a simple decision tree are achieving lower results than the other approaches, consequently a comparison with theses algorithms is neglected. The investigations in this section are focusing on (1) the general impact of image segmentation on classifying multisensor imagery, (2) the performance of a multilevel classification approach, and (3) the combination of different classifier algorithms, i.e., SVM and RF.

5.3.1 Data set and preprocessing

The same multisource data set as in Section 5.2 was used for this approach. As a last pre-processing step, image segmentation was performed. Various algorithms for image segmentation have been introduced (Baatz and Schäpe, 2000; Evans *et al.*, 2002). In this study a frequently applied method was used (Shackelford and Davis, 2003; Song *et al.*, 2005a, van der Linden *et al.*, 2007). This method is based on a region-growing algorithm (Baatz and Schäpe, 2000). General constrains of the segmentation concept are introduced in Bruzzone and Carlin (2006). A brief description is given in Section 2.4.

The different levels of aggregation are generated for each data set using only the spectral or backscatter intensity information. For both data sources, the multitemporal SAR data and the Landsat image, the average segment sizes are between approximately 10 and 55 pixels (Figure 5.3.1). The segments of the smallest aggregation level *#1* include usually only fractions of natural features and a few pixels. In contrast to this, outlines in the two larger segmentation levels #2 and #3 rather correspond to natural objects. Further aggregation merges several natural features into single image segments and therefore do not provide any useful additional information. Consequently larger levels were not considered in the subsequent classification process.

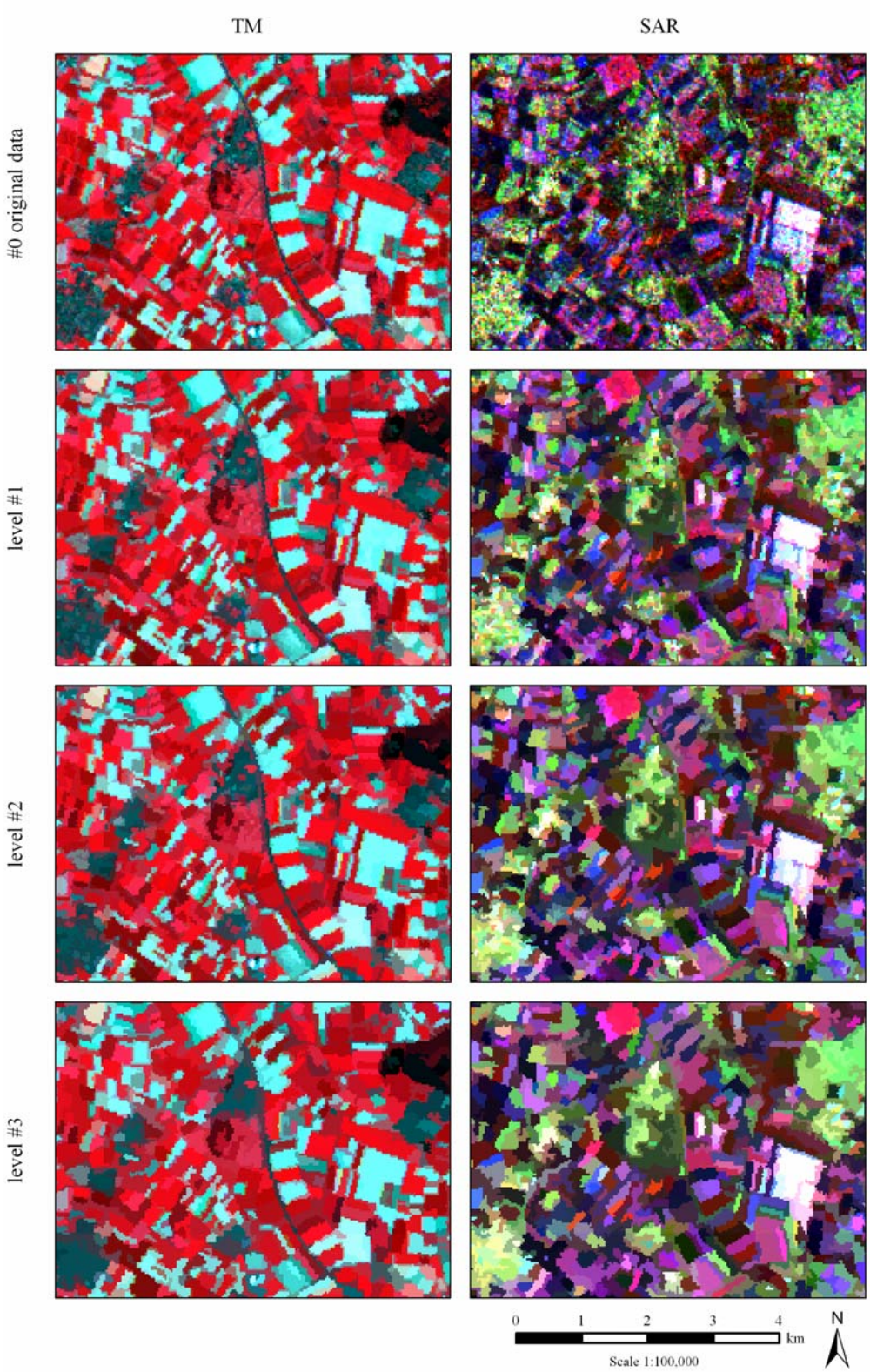

Figure 5.3.1: Subsets of the original Landsat 5 TM image and the SAR data, with the corresponding segmentation levels.

5.3.2 Methods

The proposed concept for the classification of multilevel data from two different sources is illustrated in Figure 5.3.2. As a last preprocessing step, the SAR data and Landsat image are independently segmented. Afterwards a *pre-classification* is performed: individual SVM are trained on the original pixel-based data and each aggregation level of the two data sets, generating the corresponding rule images. The application of the SVM with the OAO multi-class strategy provides 28 rule images per source. These rule images are combined to one data set, which consequently includes 8×28 rule images. Finally the information of these rule images is combined by performing an additional SVM classification on the outputs as already done in Section 5.3 (i.e., *SVM fusion*). Furthermore random forests are used for fusing the rule images (from now on referred to as *RF fusion*).

Analogous to the other applications the same 8 land cover classes were considered for the classification: (1) *Arable crops*, (2) *Cereals*, (3) *Forest*, (4) *Grassland*, (5) *Orchards*, (6) *Rapeseed*, (7) *Root crops*, and (8) *Urban*. Again, 150 samples per class were selected from the ground mapping by equalized random sampling, guaranteeing that all 8 classes are included in the sample set. These samples were used during all stages of the classification approach. Once an independent validation set with 500 samples per class was employed for the accuracy assessment.

The SVM was performed as done before in Section 5.3., employing *imageSVM* (Janz *et al.*, 2007). Using a Gaussian kernel and the OAO multiclass-strategy, individual rule images were generated for each aggregation level and the original pixel level of the SAR data and the multispectral image. The rule images were fused employing random forests (*RF fusion*) and an additional SVM (*SVM fusion*), which were trained on all rule images (Figure 5.3.2). To assess the performance of the presented multilevel-multisource fusion concepts, a conventional SVM and RF were applied directly to (1) each individual aggregation level, (2) individual multilevel stacks from both source, (3) a single-level multisource dataset, and (4) a stack of all aggregation levels from both sources.

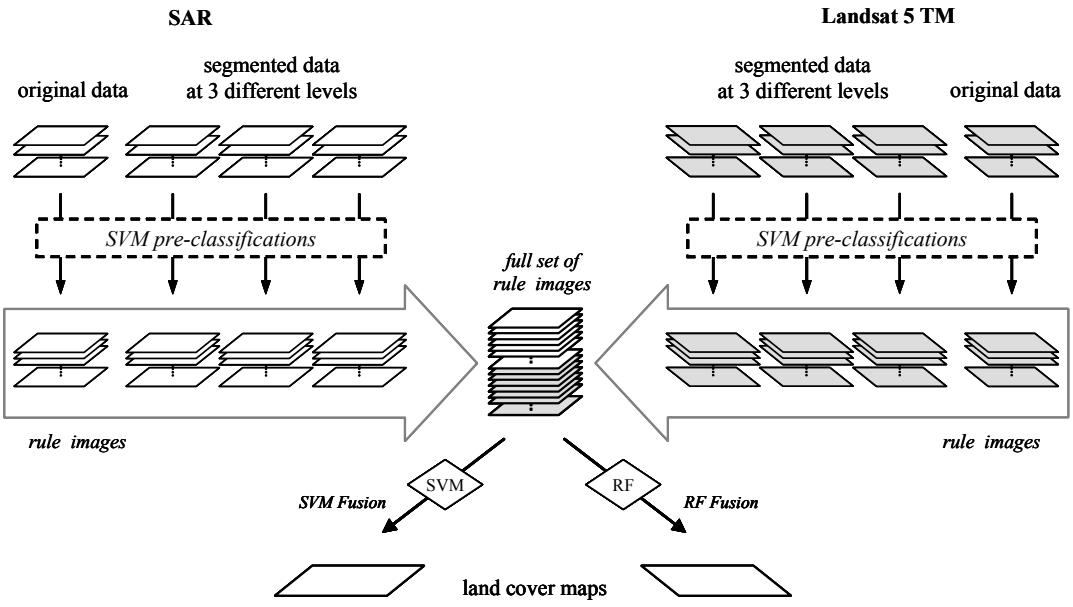

Figure 5.3.2: Schematic overview on the multisensor-multilevel fusion concept.

5.3.3 Experimental Results

Single-source results

The different segmentation levels of the two image sources were separately classified using both a common SVM and RF. Experimental results clearly show the positive effect of image segmentation. Independent from the classifier and data source, the total accuracy is significantly improved by prior image segmentation. Using the coarse segmentation level #3, the total accuracies increased by 14% and 9% compared to the classification results achieved on the original pixel-based SAR and Landsat imagery, respectively. Even the use of the fine segmentation level #1, increases the overall accuracy of the SAR data by 10% and of the optical image by 6% (Table 5.3.1). The results clearly demonstrate that image segmentation is particularly worthwhile for the noise inherent SAR imagery: The accuracies of the original SAR data are below that of the multispectral imagery, whereas the results are comparable at level #1 and #2. The accuracies of the SAR data are further increased by employing level #3 and outperform that of the multispectral Landsat image. The application of the classifiers to stacked data sets including the pixel and all aggregation levels of an individual image source do not further improve the results (Table 5.3.1).

Table 5.3.1: Overall accuracy [%], using SVM and RF on single-source imagery at different segmentation levels.

Segmentation level	**SAR**		**TM**	
	SVM	*RF*	*SVM*	*RF*
#0 (original data)	63.3	65.0	69.8	68.8
#1 (~10 pixel)	73.3	75.6	73.9	74.8
#2 (~23 pixel)	75.0	78.1	75.5	77.2
#3 (~55 pixel)	77.8	79.7	75.3	77.6
all levels	75.9	79.9	74.5	77.7

Regarding the class-specific accuracies the results illustrate the dissimilarity of the different image sources and aggregation levels. Neither an individual image source (i.e., SAR and multispectral) nor a specific level of segmentation is ideal to differentiate between all classes (Table 5.3.2 and Table 5.3.3). Whereas in the SVM-based classifications *Grassland* and *Orchards* better differentiated in the SAR data, irrespective of the segmentation level, the TM scene is more appropriate for the other classes. Regarding the producer accuracies achieved by the SVM on the TM image for example, it can be assessed that the segmentation level #2 is adequate to describe *Root crops* and *Cereals*, whereas *Arable crops* and *Grassland* are better classified by segmentation level #3. Consequently a definition of an adequate segmentation level for all class is difficult.

Table 5.3.2: Producer and user accuracies [%], using SVM on single-source two data sets at different segmentation levels.

Land cover class	SAR				TM			
	Producer Accuracy		User Accuracy		Producer Accuracy		User Accuracy	
	#2	*#3*	*#2*	*#3*	*#2*	*#3*	*#2*	*#3*
Arable crops	72.4	72.8	71.7	83.1	71.2	81.6	79.1	70.2
Cereals	79.8	80.2	63.1	60.1	81.2	74.4	67.6	65.6
Forest	88.2	91.0	87.9	96.2	93.4	92.2	96.5	96.9
Grassland	65.8	65.4	77.4	70.8	55.8	62.0	71.7	70.3
Orchard	75.4	77.8	66.4	72.6	62.6	63.0	58.0	67.2
Rapeseed	70.8	75.6	85.7	88.1	73.8	70.6	82.4	77.6
Root crops	71.6	76.4	71.5	71.7	80.2	74.8	66.5	70.7
Urban	76.2	83.2	83.9	90.0	86.2	84.0	88.9	87.1
standard deviation	*6.8*	*7.5*	*9.2*	*12.2*	*12.3*	*10.4*	*12.8*	*11.0*

Table 5.3.3: Producer and user accuracies [%], using RF on single-source data sets at different segmentation levels.

Land cover class	SAR				TM			
	Producer Accuracy		User Accuracy		Producer Accuracy		User Accuracy	
	#2	*#3*	*#2*	*#3*	*#2*	*#3*	*#2*	*#3*
Arable crops	77.4	81.0	75.7	77.1	77.0	83.6	71.6	70.4
Cereals	81.6	76.4	66.9	64.4	76.2	72.6	74.4	73.5
Forest	90.4	92.8	86.3	95.5	92.8	92.6	96.1	97.9
Grassland	70.8	69.6	79.0	75.0	61.0	70.4	72.8	67.3
Orchard	80.2	81.8	70.5	73.2	71.6	66.2	62.9	70.9
Rapeseed	77.4	80.6	82.2	83.6	75.4	76.0	83.2	82.3
Root crops	72.0	71.8	81.3	83.3	77.2	74.0	70.8	74.0
Urban	75.2	83.6	88.7	90.9	86.4	85.6	89.8	87.9
Standard deviation	*6.2*	*7.3*	*7.5*	*10.0*	*9.5*	*8.8*	*11.0*	*10.5*

Multisensor-Multilevel Results

To generate the final *multilevel-multisensor* classification, the different segmentation levels of both data sources were combined by the presented strategy of *SVM fusion* and *RF fusion*. Moreover usual SVM and RF classifier were trained directly on a stacked dataset, containing the imagery and not the SVM rule images from the pre-classifications of each level and image source. The accuracy assessment clearly shows the positive influence of a synergetic use of multisensor imagery at various segmentation levels: The highest accuracies achieved on single-source imagery are 79.9% on the SAR data and 77.7% on the multispectral image (Table 5.3.1), whereas the application of SVM and RF directly on of the stacked data set increases the accuracy at least up to 81.1% and 83.6% respectively (Table 5.3.4). The performance is further increased to 84.9% by the proposed fusion strategies (*RF fusion*).

Table 5.3.4: Overall accuracy [%] of the multilevel-multisensor classification, using different classifiers and fusion strategies.

Classifier / fusion strategy	Overall accuracy
single SVM	81.1
SVM fusion	83.5
single RF	83.6
RF fusion	84.9

The advantage of using diverse image sets is also underlined by the class-specific accuracies, shown in Table 5.3.5. In general fusing different image sources increase the class specific accuracies. Moreover the accuracies are more balanced. Producer and user accuracies of the single-source results vary strongly (Table 5.3.2 and Table 5.3.4) whereas the classifications that are based on the fusion concept are less variable, resulting in a reduced standard deviation (Table 5.3.5). This underlines the worth of combining different sensor sources at different segmentation levels.

Table 5.3.5: Class-specific accuracies [%], using different fusion strategies.

Land cover class	SVM fusion		RF fusion	
	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
Arable crops	79.6	84.8	78.2	95.1
Cereals	81.4	78.8	85.2	76.3
Forest	94.8	96.9	94.6	97.7
Grassland	76.6	76.4	82.2	76.7
Orchard	81.6	78.7	83	81.2
Rapeseed	84.8	84.8	89	79.6
Root crops	80.8	78.6	78.8	85.3
Urban	88.2	89.4	88.4	92.3
standard deviation	*5.7*	*6.9*	*5.6*	*8.5*

One might argue that a multisensor fusion of individual segmentation level could be more appropriate than the use of various levels. To assess the impact of the multilevel approach and to underline the value of the proposed mutisensor-multilevel strategy, the RF fusion was applied to a multisensor data set, including only single segmentation scales. The achieved classification results demonstrate the worth of a multilevel approach (Table 5.3.6): Whereas the use of a single segmentation level increases the multisensor pixel-based result by 3.7% and 6.4%, using level #1 and #2 respectively, the multisensor-multilevel classification outperforms all single-level classifications in terms of accuracy, resulting in an overall accuracy of 84.9% (Table 5.3.4).

Table 5.3.6: Overall accuracy [%], fusing different individual segmentation levels, using RF fusion.

Segmentation level	Overall accuracy [%]
#0 (pixel)	77.4
#1	81.1
#2	83.8
#3	83.5

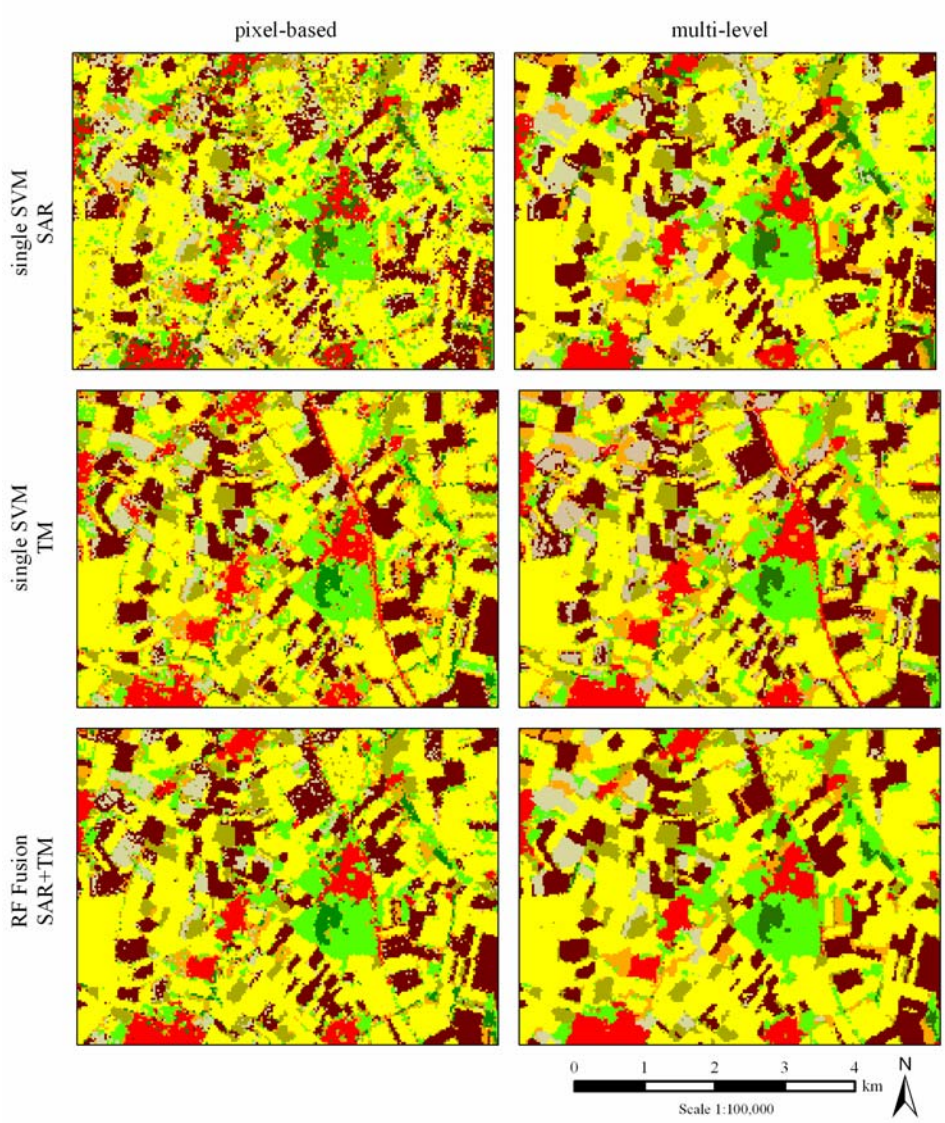

Figure 5.3.3: Classification results, using a conventional SVM and RF fusion on pixel data and multilevel data

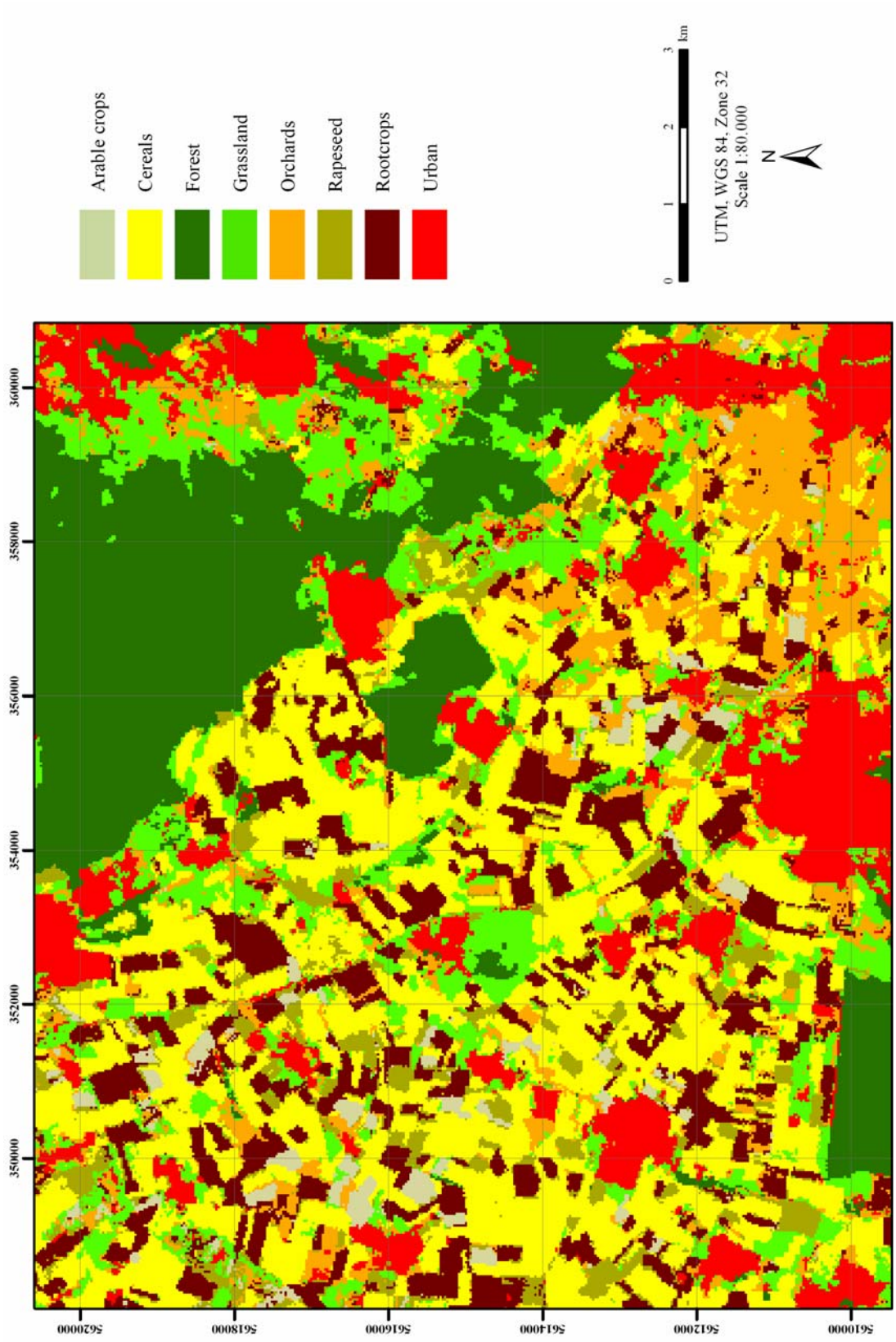

Figure 5.3.4: Multisensor-multilevel classification result, using RF fusion

A visual assessment of the classification results confirms the positive impact of the classification strategy. Concerning the image segmentation, the noise within the pixel-based results is clearly reduced, in particular the speckle in the SAR imagery. Natural objects, e.g. field parcels, appear more homogeneous. Moreover some typical errors, e.g. a confusion of forest and urban in the SAR data, are eliminated. Furthermore the classification results illustrate the complementary character of the multitemporal SAR data and the multispectral image. Contrary to the visibly good performance of the RF fusion, and as in the approach introduced in Section 5.2, the disappearance of the highway that was recognized in the TM data (located in the in the middle of the image subset), shows possible drawbacks of the multisensor classification.

5.3.4 Discussion and Conclusion

In this section the issue of classifying multisensor data at different segmentation levels was addressed. The use of information from different sensors as well as and the integration of the spatial context at various segmentation level with different scales performs well. Both, the sensor fusion as well as the image segmentation contributes to an increased overall accuracy.

The RF fusion achieves an accuracy of 77.4% when applied to the two data sets (i.e., SAR and TM) at pixel level. This is significantly higher than results achieved by single-source pixel-based classification. The accuracy is improved by 3.7% and more by prior image segmentation and the multisource results that are based on a single segmentation level are generally higher than results achieved on segmented single-source imagery. Finally results achieved by the proposed multilevel-multisource approach are always better than those achieved on any other combinations of data. The assessment of the producer and user accuracies underlines the general good performance of the proposed RF fusion concept. Results from SVM fusion on the other hand, show a very positive balance between producer and user accuracies of the classes.

The experimental results underline the different nature of the image sources, which provide diverse information. Moreover different aggregation levels contribute unequally to the classification of the various land cover classes and the definition of an ideal segmentation level is difficult. The multilevel approach that is based on the RF fusion strategy appears appropriate for decision making in this context. As in the previous sections both classification techniques, the SVM and RF, perform well. The main reason for the success of the proposed RF fusion is the sequential use of both algorithms. The complex class distributions that are modelled by the SVM and the RF classifier, which is based various combinations of input features for optimized decision making. In this context the pre-classification by SVM can be regarded as a class-specific data transformation. By applying a SVM, the image data is transformed into a new feature space that is made-up of the distance values of the individual SVM rule images. This transformation is performed separately for each image source and segmentation level and the individually generated rule images are better comparable than the original data sets. This could be simplifying the definition of the split rules during training of the RF and thus, results in the highest accuracies. The experimental results also underline the findings in 5.2, that it is more than adequate to define the kernel functions for each data source and level separately, instead of using a single kernel function for the whole data set.

5.4 Transfer of the multisensor-multilevel strategy

The presented applications have shown that the two methods, random forests and support vector machines are adequate for classifying multisource dataset. Moreover a combination of the two classifiers in a multisensor-multilevel fusion strategy achieves the highest classification accuracies. Thus in the following the proposed concept is applied to a data set from another year, to assess the reproducibility and transferability of the approach. The classification complexity is increased by additional land cover classes. Furthermore the multispectral time series is extended.

5.4.1 Dataset and preprocessing

As before, a multisensor dataset consisting of SAR and optical imagery was available for the known study site. The SAR dataset includes 4 Envisat ASAR AP and 6 ERS-2 acquisitions from a period between April and September 2006 (Table 5.4.1). Contrary to the other applications, the multispectral dataset includes three SPOT acquisitions (Table 5.4.2). The preprocessing is performed following the common procedures that are described in Section 4. For matter of comparison all data are resampled to 30 m spatial resolution. Again as a last preprocessing step, multilevel image segmentation was performed, generating different segmentation levels for each data set, i.e., the multitemporal SAR and the multitemporal Spot imagery (Figure 5.4.1).

Table 5.4.1: Image characteristics multitemporal SAR data set, 2006.

Sensor	**Image characteristics SAR data**			
	Date	*Track/Swath*	*Polarization*	*Orbit*
ERS-2	06-Apr-06	337	VV	Des
ASAR	16-Apr-06	2487	HH / HV	Des
ERS-2	11-May-06	337	VV	Des
ASAR	21-May-06	2487	HH / HV	Des
ERS-2	09-Jun-06	258	VV	Asc
ERS-2	15-Jun-06	337	VV	Des
ASAR	25-Jun-06	2487	HH / HV	Des
ERS-2	20-Jul-06	337	VV	Des
ASAR	30-Jul-06	2487	HH / HV	Des
ERS-2	24-Aug-06	337	VV	Des

Table 5.4.2: Image characteristics multispectral imagery, 2006.

Sensor	**Image characteristics, multispectral imagery**			
	Date	*# band*	*Band width [µm]*	*Spatial resolution [m]*
SPOT 2	11-May-06	3	0.5-0.89	20
SPOT 5	24-Jun-06	4	0.5 - 1.75	10/20
SPOT 2	17-Jul-06	3	0.5-0.89	20

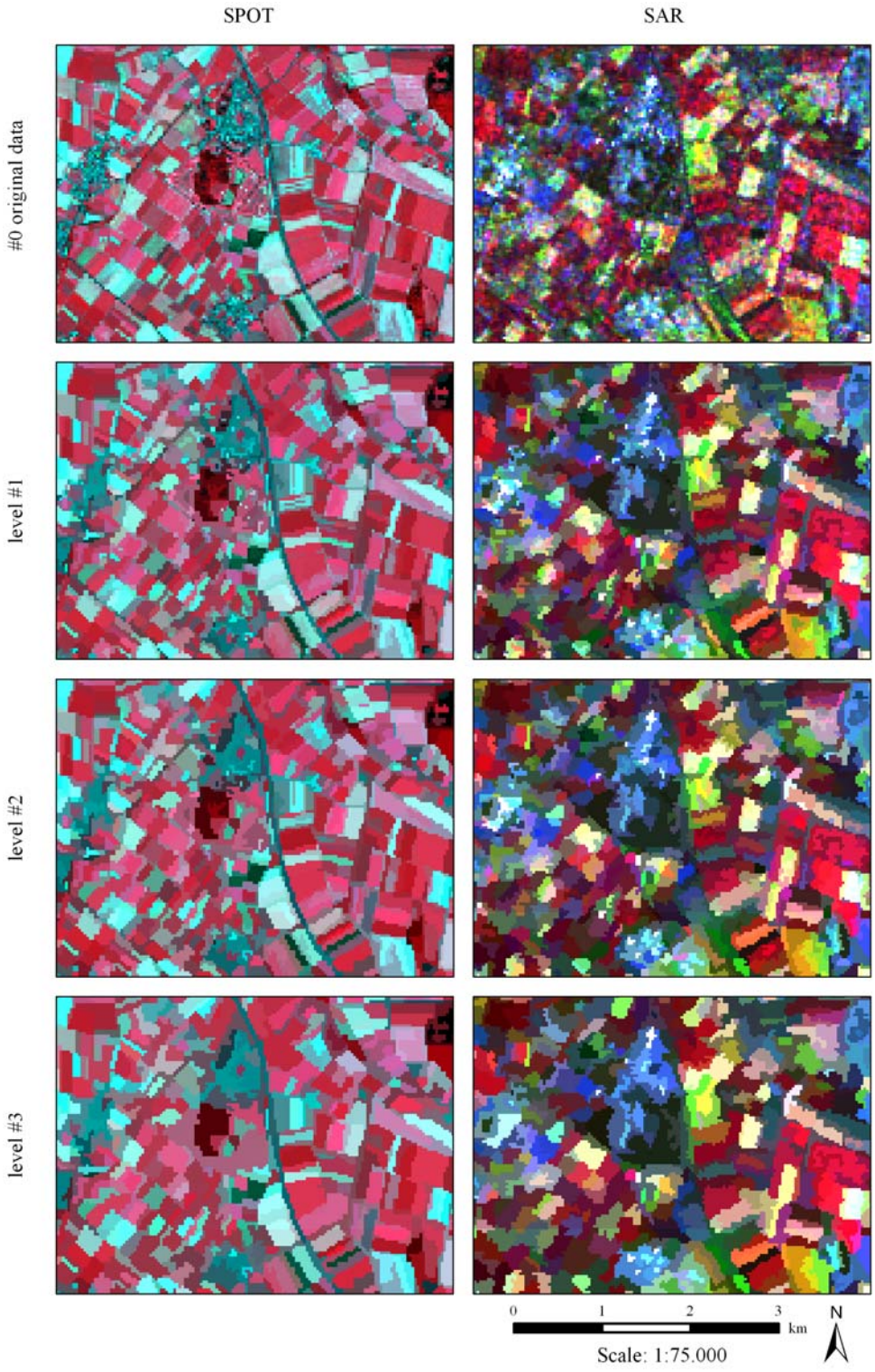

Figure 5.4.1: Subsets of the original SPOT and SAR data, and corresponding segmentation levels.

5.4.2 Methods

The approach is following the methods described in Section 5.3. As a last preprocessing step, the SAR data and Spot images are independently segmented, using only the spectral or backscatter intensity information. For both data sets three segmentation scales are generated (#1, #2, #3), with an average segment sizes are approximately 11 and 50 pixels. Separate SVM are trained on the original pixel-based data and each segmentation level of the two data sets, generating the corresponding rule images. The information is combined by applying random forests on the stacked rule images (i.e., RF fusion).

For matters of comparison conventional SVM with OAO strategy and conventional RF were applied to various data sets. In contrast to the SVM fusion and RF fusion these classifiers were directly trained on the original data and not on the corresponding rule images (i.e., single SVM and single RF). The methods were applied to (1) each individual aggregation level, (2) multilevel stacks from single sources, and (3) a stack of all aggregation levels from both data sources. In addition to this, the RF fusion was applied on a multisensor data set, which includes only single segmentation levels.

In contrast to the other applications 11 land cover classes were classified: *Arable crops*, *Cereals*, *Coniferous forest, Mixed forest, Grassland*, *Gravel pits, Maize*, *Orchards*, *Rapeseed, Root crops,* and *Urban*. Again, 150 samples per class were selected from the ground mapping, using by equalized random sampling strategy.

5.4.3 Experimental results

Single Source results

As in Section 5.3, the classification results clearly show the positive effect of image segmentation. Independent from the algorithm and data source, the total accuracy is significantly improved. The positive impact of image segmentation is more dominant for the speckle inherent SAR data. Whereas the image segmentation of the SAR data improves the overall accuracy by up to 21% that of the multispectral imagery is increased by up to 4.5%. A "simple" combination of all segmentation levels do not improves the results, expect for the RF-based classification of the Spot images (Table 5.4.3). Moreover the application of a SVM on a multilevel data set reduces the overall accuracies, achieved on a single segmentation level. In contrast to the results in 5.3, the classification results achieved on the multispectral imagery are always more accurate compared to the results achieved on the SAR data.

Table 5.4.3: Overall accuracy [%], using individual SVM and RF on SAR and SPOT data from 2006, at different segmentation levels.

Segmentation level	SAR		SPOT	
	SVM	*RF*	*SVM*	*RF*
#0 (pixel)	45.9	48.7	75.2	73.8
#1	60.1	63.8	76.9	77.3
#2	63.0	66.6	75.7	77.9
#3	67.2	68.7	78.1	78.3
all levels	60.1	69.8	78.4	79.2

The difficulty to define an ideal segmentation level is confirmed by the class-specific accuracies (Table 5.4.4 and Tabel 5.4.5). The results demonstrate the differences of the various image sources and segmentation levels; no single level seems to be ideal to differentiate between all classes Using a SVM for example, *Cereals* and *Root crops* are better classified by segmentation level #2, but *Rapeseed* is more accurately classified better by level #3. Scale #3 achieves the highest overall accuracies from all single-level results, whereas the results of level #2 are less variable, resulting in a reduced standard deviation of the class-specific accuracies.

Table 5.4.5: Producer and user accuracies [%] and corresponding standard deviation, using individual SVM on SAR and multispectral data sets at different segmentation levels.

Land cover class	SAR				SPOT			
	Producer Accuracy		User Accuracy		Producer Accuracy		User Accuracy	
	#2	*#3*	*#2*	*#3*	*#2*	*#3*	*#2*	*#3*
Arable crops	57	74	57.3	61.1	74.2	71	72.6	72.3
Cereals	73.4	67.4	52.2	52.1	80.6	80.6	66.9	67.5
Coniferous	48.4	68.2	50.2	58.8	84.4	92.2	66.7	66.5
Mixed forest	51	48.4	48.7	57.8	56	54	75.3	82.3
Grassland	60.2	58.2	64.3	73.3	65.4	67.2	69.7	74.2
Gravel	72	89.6	90	85.2	78.2	92	86.9	90.2
Maize	67.6	61.2	67.7	68.3	76.2	83	91.5	92.8
Orchard	64.6	68.6	63.5	65.2	78.6	78.6	66.0	68.71
Rapeseed	67.2	68.2	86.82	90.7	81.2	79.6	80.6	88.4
Root crops	64.6	61.8	62.9	62.5	76.6	78	75.9	75.6
Urban	67.4	74.2	65.1	76.8	81.4	82.8	91.5	93.4
standard deviation	*8.1*	*10.5*	*13.5*	*12.0*	*8.2*	*10.9*	*9.7*	*10.5*

Table 5.4.4: Producer and user accuracies [%] and corresponding standard deviation, using individual RF on SAR and multispectral data sets at different segmentation levels.

Land cover class	SAR				SPOT			
	Producer Accuracy		User Accuracy		Producer Accuracy		User Accuracy	
	#2	*#3*	*#2*	*#3*	*#2*	*#3*	*#2*	*#3*
Arable crops	65.6	71.8	60.4	62.0	73.2	75.6	74.1	72.8
Cereals	73.4	71.0	56.5	54.9	81.2	77.6	72.8	68.8
Coniferous	64.2	69.2	52.2	58.0	81.4	75.6	67.1	62.9
Mixed forest	38.0	44.0	48.0	55.8	59.6	54.8	74.5	68.2
Grassland	63.6	62.0	68.8	74.7	67.0	68	71.9	75.6
Gravel	78.2	88.0	89.1	89.8	85.8	95	87.2	91.3
Maize	71.4	69.2	75.2	67.7	82.2	85.2	91.9	92.8
Orchard	69.0	69.8	66.0	70.2	82.0	83	67.5	75.0
Rapeseed	75.0	74.0	88.7	84.3	85.8	84.4	89.2	87.0
Root crops	67.8	61.4	68.5	68.1	77.6	77	78.7	79.2
Urban	66.4	75.4	69.3	78.5	81.6	85.2	89.7	92.4
standard deviation	*10.6*	*13.3*	*10.8*	*11.6*	*8.2*	*10.5*	*9.3*	*10.6*

Multisensor-Multilevel Results

For a multisensor-multilevel fusion, the different source, which were separately pre-classified by a SVM, were combined by random forests (RF fusion). For matter of comparison a single SVM and a single RF were trained directly on a stacked data set, consisting of the imagery and not the SVM rule images (single RF and single SVM).

As the results achieved in 5.3, the positive impact of a synergetic use of multisensor imagery at different segmentation levels is demonstrated. Where as the highest single-source result achieved on a single segmentation level is 68.7% and 78.3% respectively (Table 5.4.3), the classification of the mulitsensor-mulitlevel data set achieves accuracies of 77.5% and 80.1%. The performance of a multilevel-multisensor classification is further increased by up to 82.7% (RF fusion) by the proposed fusion strategy (Table 5.4.6). Most of the class-specific accuracies are increased by the RF fusion. The standard deviations of the class-specific accuracies are reduced by the multilevel-multisensor approach in some cases, compared to the results achieved with single-source data at individual segmentation levels (Table 5.4.7). The final classification result of the multilevel-multisensor approach for 2006 is shown in Figure 5.4.2.

Table 5.4.6: Overall accuracy [%], using different classifiers on the multilevel-multisensor data set.

Classifier algorithm	Overall accuracy [%]
Single SVM	77.5
Single RF	80.1
RF fusion	82.7

Table 5.4.7: Producer and User accuracies [%], using RF fusion.

Land cover class	Producer accuracy	User accuracy
Arable crops	80.6	76.0
Cereals	86.0	72.9
Coniferous	89.8	67.9
Forest	55.8	81.8
Grassland	72.2	77.3
Gravel	95.0	96.9
Maize	84.0	94.0
Orchard	86.4	82.0
Rapeseed	88.0	94.4
Root crops	82.4	80.5
Urban	89.0	94.9
standard deviation	*10.6*	*10.0*

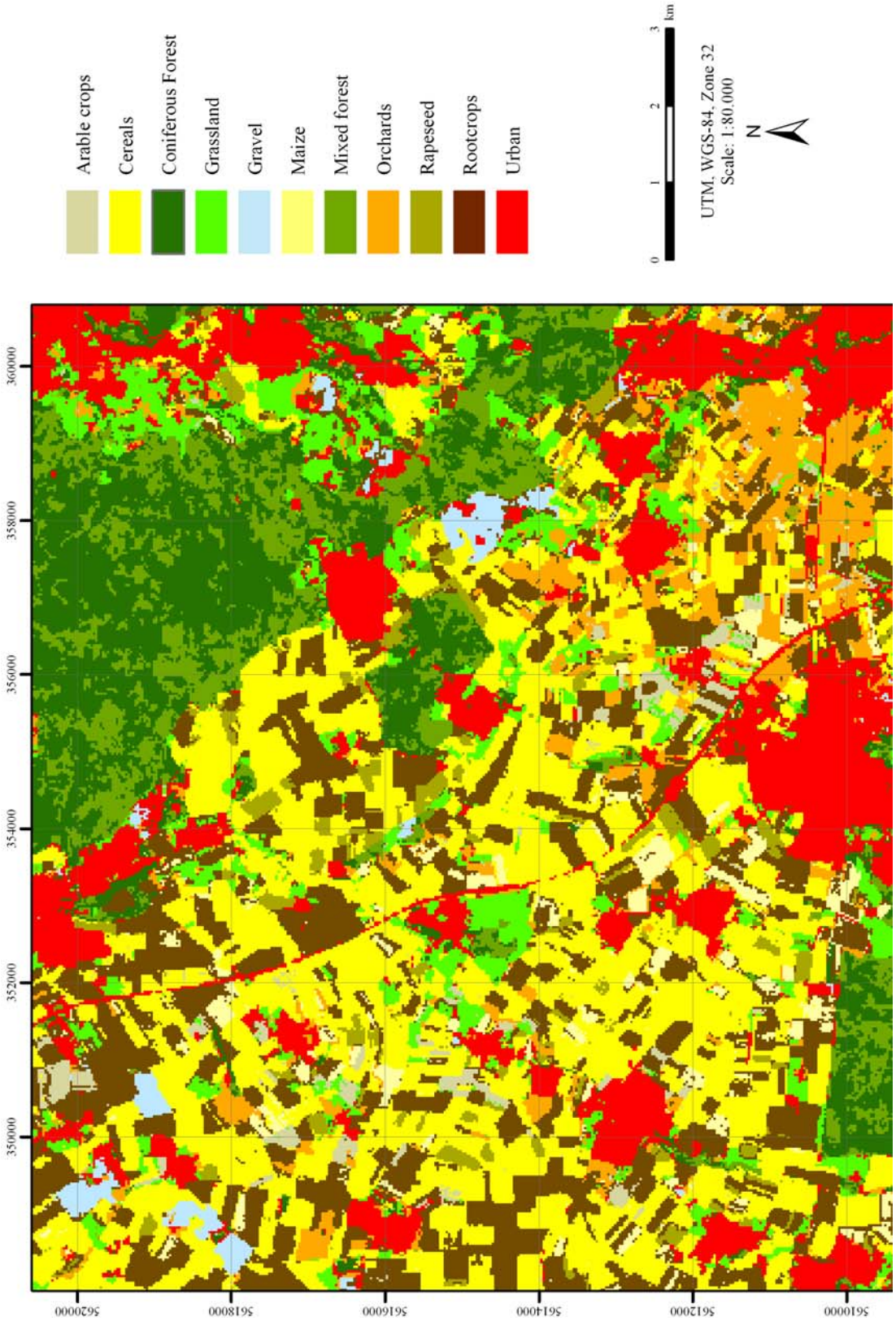

Figure 5.4.2: Multisensor-multilevel classification result 2006, using RF fusion.

The visual interpretation of the final classification result, confirms the aforementioned findings. Natural features, e.g. field parcels, appear almost more homogeneous and the typical noise is eliminated, due to image segmentation. Even smaller natural features are correctly classified.

The RF fusion was also applied to multisensor imagery, using only individual segmentation levels (Table 5.4.8). As in the section before, results clearly demonstrate the positive impact of image segmentation and the value of the multilevel fusion strategy: The RF fusion, which is applied on pixel data, results in a total accuracy of 75.6%, whereas a "simple" segmentation improves the results by 4% and more. The use of all segmentation levels and the application of the introduced multisensor-multilevel approach results in the highest overall accuracy.

Table 5.4.8: Overall accuracy [%], applying RF fusion on individual segmentation levels.

Classifier algorithm	Overall accuracy [%]
#0 (pixel)	75.6
#1	79.7
#2	80.5
#3	80.5
all levels	82.7

5.4.4 Conclusion and Discussion

The results clearly confirm the findings from Section 5.3. The combination of different data sources as well as the integration of spatial information at different scales (i.e., image segmentation) contributes significantly to improved classification accuracy. The definition of an ideal single segmentation level is critical and even if a "simple" single-level approach improves the results of a pixel-based classification, a multilevel strategy achieves the highest accuracies.

In contrast to the single-sensor results in 5.3, the extended multispectral time series outperforms the SAR data in terms of accuracy, irrespective of the segmentation level. This fact underlines the worth of multitemporal optical imagery. Nevertheless, the integration of SAR data still improves the classification accuracies. Thus, the value of multisensor image analysis is more than confirmed. It can be expected that the positive impact is further strengthen by recent and upcoming multitemporal SAR data with enhanced spatial resolution and increased revisit times.

6 Synopsis

6.1 Main Findings

It can be demonstrated that the different information provided by synthetic aperture radar and multispectral sensor systems are complementary for an improved land cover classification. Although multisensor approaches can improve the results irrespectively from the classifier algorithm, the application of recent non-parametric concepts is useful. Support vector machines and random forests significantly outperform the classification results achieved by conventional algorithms. As shown in other studies, image segmentation usually improves the classification accuracy compared to common pixel-based approaches. Nevertheless, the definition of a single segmentation level is difficult, and a multilevel approach is worthwhile. Finally it can be assessed that the results achieved by recent classifier developments (i.e., support vector machines and random forests) are further improved by modifying and combining these concepts. In the following the findings of the different application are summarized and discussed.

The results clearly demonstrated that the RF approach is well suited for classifying multitemporal SAR data. The achieved overall accuracies are higher than that of traditional classifier and other classifier ensembles. Compared to other recent developments as SVM, RF performs at least comparable. The SAR inherent noise in significantly reduced and the image is classified into more homogeneous areas. The differences between the maps from the simple decision trees and classifier ensembles are comparable to the differences between pixel-based and segment-based approaches. In addition the result achieved on the SAR data underlines the assumption and findings in other studies that multitemporal and multipolarized imagery is worthwhile for an improved crop differentiation.

The classification accuracies are further improved by combining multitemporal SAR data and multispectral imagery. Even when a set of multispectral imagery is available, which can achieve higher overall accuracies compared to SAR data, a combination is useful. In general this improvement is achieved for all considered land cover class. Although the results of conventional classifiers can be improved by such heterogeneous multisource data sets, the use of complex techniques is more adequate in this context. Especially random forests can be pointed out as an accurate as well as simple approach. Nevertheless the performance of these two recent classifier algorithms can be further improved in context of classification accuracy, as demonstrated by the experimental results.

The training of SVM for the full multisensor data sets demands the definition of a specific kernel function. In context of heterogeneous imagery, it is more adequate to define individual kernel functions for each data source and fuse the outputs in a subsequent process.

The sequential use of SVM and RF combine the strengths of both algorithms: SVM can model complex class distributions and RF performs an optimized decision making on a basis of various combinations of input features. By generating SVM-based pre-classifications, the imagery is transformed into a new feature space that is composed of the distance values of the SVM outputs. This pre-classification, which is based on a one-against-one multiclass-strategy, can be regarded as class-specific data transformation. The values of the generated rule images

are better comparable than the original data sets. This simplifies the determination of the split criterion during the training phase of the RF. Consequently the application of the RF fusion leads to the highest accuracies, as shown in the results.

Segment-based approaches are well suited for increasing the classification accuracy. Especially the segmentation inherent speckle reduction significantly improves the accuracy for the SAR data and emphasizes the quality of multitemporal data for the classification of agricultural areas. On the other hand the result underlines the assumption and findings from other studies, that different segmentation levels provide different information and that a definition of a single segmentation size is critical. The different image sources provide diverse information and various aggregation levels contribute unequally to the classification of the various land cover classes. Thus the different information sources are not necessarily equally reliable. One specific source (i.e. image type or specific aggregation level) is more applicable to describe a specific class and not appropriate for another class, whereas another one seems more adequate to classify another land cover class. Consequently the use of information from different sensors as well as the integration of spatial context from various segmentation levels with different scales is valuable. This statement was clearly confirmed by the results of the presented applications.

6.2 Summary

Overall the results underline that multitemporal and multisensor imagery are worthwhile for an improved land cover classification. The use of image segmentation further improves the results and is particularly worthwhile for the noise inherent SAR data. The adequate utilization of the wide information content of such complex data sets requires recent classifier developments. The synergetic use of different classifier algorithms seems useful, to combine the strengths of different techniques. Regarding the initial research questions, the findings can be summarized:

- Multisensor imagery improves the accuracy of land cover classifications, normally independent form the land cover class and the classifier algorithm.
- Recent classifier developments as classifier ensembles and support vector machines are adequate in this context and significantly outperform conventional methods.
- A “simple” image segmentation can increase the classification accuracy. Nevertheless the definition of a single segmentation level for all land cover classes and the whole multisensor data set is critical. Thus, a multilevel approach is to prefer.
- The performance of recent classifier developments can be further improved by a combination of different concepts.

6.3 Prospect

Regarding the continuously increasing availability of various remote sensing data, multisensor applications become even more attractive. Moreover, recent and upcoming satellite missions, as TerraSAR-X, Radarsat-2, provide spatially enhanced and temporally more detailed information: The improved spatial resolution might result in more valuable textural information content for example, whereas increased revisit times results in more current information on the environmental state. SAR missions as ALOS PALSAR and TerraSAR-X are the first space-borne systems, which provide temporally frequent polarimetric imagery. In addition several recent SAR systems operate in different wavelength (e.g., X-band, C-band, L-band). Overall a multitude of different but complementary information on land cover is provided. On the other hand the dimensionality and complexity of such data sets is further increased and thus it might demand more sophisticated classifier concepts. Future research within this context should concentrate on the integration of these new data sets to include additional information provided by recent Earth Observation systems. Regarding space-borne polarimetric SAR data the integration of methods, which were originally introduced for airborne data seems feasible.

In addition it is expected that the positive impact of segmentation and the presented multilevel strategy becomes even more dominant in the context of high-resolution imagery. Thus future studies may concentrate on the integration of additional segment features as shape on the one hand and alternative segmentation algorithms on the other hand.

The pre-processing of remote sensing data is an important step for image analysis, particularly in multisensor image analysis. One research focus could be set on an adequate multisensor preprocessing, for example an improved co-registration between SAR systems and multispectral imagery or a temporal extension of the segment-based speckle reduction, as introduced in Waske *et al*. (2007).

Overall the universality of the proposed concepts and findings can be underlined by transferring the methods to other study sites. Thus, beside an application on the well-known agricultural test site near Bonn, the investigations should be performed on urban regions as well as rural test sites, with different environmental settings. In context of operational applications and products, the integration of enhanced land cover products into monitoring and decision support systems seems interesting to underline the general worth of the products and upcoming missions.

In summary, the following main subjects are worth to investigate in detail: (1) the derivation of additional temporal information, using temporal high resolution imagery, as RapidEye and TerraSAR-X data, (2) the integration of textural information, e.g. derived from spatial high resolution SAR data, as TerraSAR-X or Cosmo-SkyMed, (3) the utilization of polarimetric satellite imagery, as provided by ALOS PALSAR and TerraSAR-X, and (4) the use of multifrequenzy approaches, using for example, ALOS Palsar, TerraSAR-X or Cosmo-SkyMed, and Radarsat-2 data.

References

Baatz M and Schäpe A (2000): Multiresolution Segmentation: An optimization approach for high quality multi-scale segmentation. *Angewandte Geographische Informationsverarbeitung XII*, Salzburg, 2000.

Backhaus R and Beule B (2005): Efficiency evaluation of satellite data products in environmental policy. *Space Policy*, vol. 21(3), pp. 173-183.

Banfield RE, Hall LO, Bowyer KW, and Kegelmeyer WP (2007): A comparison of decision tree ensemble creation techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29(1), pp. 173-180.

Bauer ME, Daughtery CST, Biehl LL, Kanemasu ET, and Hall FG (1986): Field spectroscopy of agricultural crops. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24(1), pp. 65-75.

Bazi Y and Melgani F (2006): Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44(11), pp. 3374-3385.

Benediktsson JA, Swain PH, and Ersoy OK (1990): Neural network approaches versus statistical-methods in classification of multisource remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28(4), pp. 540-552.

Benediktsson JA and Swain PH (1992): Consensus theoretic classification methods. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22(4), pp. 688-704.

Benediktsson JA, Sveinsson JR, and Swain PH (1997): Hybrid consensus theoretic classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35(4), pp. 833-843.

Benediktsson JA and Kanellopoulos I (1999): Classification of multisource and hyperspectral data based on decision fusion. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37(3), pp. 1367-1377.

Benediktsson JA, Pesaresi M, and Arnason K (2003): Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41(9), pp. 1940-1949.

Benediktsson JA, Palmason JA, and Sveinsson JR (2005): Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(3), pp. 480-491.

Berk A, Bernstein LS, and Robertson DC (1989): *MODTRAN: A Moderate Resolution Model for LOWTRAN 7*. GL-TR-89-. 0122, Geophysics Laboratory, Bedford MA.

Bishop CM (1995): *Neural Networks for Pattern Recognition*, Oxford University Press.

Bishop CM (2006): *Pattern Recognition and Machine Learning*, Springer.

Blaes X, Vanhalle L, and Defourny P (2005): Efficiency of crop identification based on optical and SAR image time series. *Remote Sensing of Environment*, vol. 96(3-4), pp. 352-365.

Brachet G (2004): From initial ideas to a European plan: GMES as an exemplar of European space strategy. *Space Policy*, vol. 20(1), pp. 7-15.

Breiman L and Culter A (2002): *Random Forests*. Software available at (last access: Sep. 2007): http://www.stat.berkeley.edu/~breiman/RandomForests.

Breiman L (2001): Random forests. *Machine Learning*, vol. 45(1), pp. 5-32.

Briem GJ, Benediktsson JA, and Sveinsson JR (2002): Multiple classifiers applied to multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40(10), pp. 2291-2299.

Brisco B, Brown RJ, Gairns JG., and Snider B (1992): Temporal ground-based scatterometer observations of crops in western Canada. *Canadian Journal of Remote Sensing*, vol. 18(1), pp. 14-22.

Brisco B and Brown RJ (1995): Multidate SAR/TM Synergism for crop classification in Western Canada. *Photogrammetric Engineering & Remote Sensing*, vol. 61(8), pp. 1009-1014.

Bruzzone L, Marconcini M, Wegmuller U, and Wiesmann A (2004): An advanced system for the automatic classification of multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42(6), pp. 1321-1334.

Bruzzone L, Chi MM, and Marconcini M (2006): A novel transductive SVM for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44(11), pp. 3363-3373.

Bruzzone L and Carlin L (2006): A multilevel context-based system for classification of very high spatial resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44(9), pp. 2587-2600.

Burges CJC (1998): A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery*, vol. 2(2), pp. 121-167.

Camps-Valls G, Gomez-Chova L, Munoz-Mari J, Vila-Frances J, and Calpe-Maravilla J (2006): Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, vol. 3(1), pp. 93-97.

Capstick D and Harris R (2001): The effects of speckle reduction on classification of ERS SAR data. *International Journal of Remote Sensing*, vol. 22(18), pp. 3627-3641.

Carreiras JMB, Pereira JMC, Campagnolo ML, and Shimabukuro YE (2006): Assessing the extent of agriculture/pasture and secondary succession forest in the Brazilian Legal Amazon using SPOT VEGETATION data. *Remote Sensing of Environment*, vol. 101(3), pp. 283-298.

Chang C-C and Lin C-J (2001): LIBSVM: a library for support vector machines. Software available at (last access: Sep. 2007): http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chapelle O, Vapnik V, Bousquet O, and Mukherjee S (2002): Choosing multiple parameters for support vector machines. *Machine Learning*, vol. 46(1-3), pp. 131-159.

Chi MM and Bruzzone L (2007): Semisupervised classification of hyperspectral images by SVMs optimized in the primal. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45(6), pp. 1870-1880.

Chung KM, Kao WC, Sun CL, Wang LL, and Lin CJ (2003): Radius margin bounds for support vector machines with the RBF kernel. *Neural Computation*, vol. 15(11), pp. 2643-2681.

Chust G, Ducrot D, and Pretus JLL (2004): Land cover discrimination potential of radar multitemporal series and optical multispectral images in a Mediterranean cultural landscape. *International Journal of Remote Sensing*, vol. 25(17), pp. 3513-3528.

Collins MJ, Dymond C, and Johnson EA (2004): Mapping subalpine forest types using networks of nearest neighbour classifies. *International Journal of Remote Sensing*, vol. 25 pp. 1701-1721.

Congalton RG and Green K (1998): *Assessing the Accuracy of Remote Sensed Data: Principles and Practices*. Lewis Publisher, London.

Davidson MWJ, Le Toan T, Mattia F, Satalino G, Manninen T, and Borgeaud M (2000): On the characterization of agricultural soil roughness for radar remote sensing studies. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38(2), pp. 630-640.

De Colstoun ECB, Story MH, Thompson C, Commisso K, Smith TG, and Irons JR (2003): National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sensing of Environment*, vol. 85(3), pp. 316-327.

Del Frate F, Schiavon G, Solimini D, Borgeaud M, Hoekman DH, and Vissers MAM (2003): On the potential of multi-polarization and multi-temporal C-band SAR data in classifying crops. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41(7), pp. 1611-1619

Duda RO, Hart PE, and Stork DG (2000): *Pattern Classification*. 2nd Ed., Chichester, New York, John Wiley & Sons.

ESA (2002): *Envisat ASAR product handbook*, Issue 1.1, December, 2002.

Evans C, Jones R, Svalbe I, and Berman M (2002): Segmenting multispectral Landsat TM images into field units. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40(5), pp. 1054-1064

Fauvel M, Chanussot J, and Benediktsson JA (2006): Decision fusion for the classification of urban remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44(10), pp. 2828-2838.

Fauvel M, Chanussot J, and Benediktsson JA (2006): A Combined Support Vector Machines Classification Based on Decision Fusion. *Proc. Int. Geoscience and Remote Sensing Symp.*, Denver, USA, 2006.

Ferrazzoli P (2001): SAR For Agriculture: Advances, Problems and Prospects. pp. 47-56. *Int. Symp. on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, Sheffield, 2001.

Fitzgerald RW and Lees BG (1994): Assessing the classification accuracy of multisource remote-sensing data. *Remote Sensing of Environment*, vol. 47(3), pp. 362-368.

Foley JA *et al*. (2005): Global consequences of land use. *Science*, vol. 309(5734), pp. 570-574.

Foody GM and Arora MK (1997): An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing*, vol. 18(4), pp. 799-810.

Foody GM (2002): Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, vol. 80(1), pp. 185-201.

Foody GM and Mathur A (2004a): A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42(6), pp. 1335-1343.

Foody GM and Mathur A (2004b): Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, vol. 93(1), pp. 107-117.

Foody GM and Mathur A (2006): The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, vol. 103(2), pp. 179-189.

Freund Y and Schapire RE (1996): Experiments with a new boosting algorithm. *Proc*. 13th Int. *Conference of Machine Learning*, Bari, Italy.

Friedl MA and Brodley CE (1997): Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, vol. 61(3), pp. 399-409.

Friedl MA, Brodley CE, and Strahler AH (1999): Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37(2), pp. 969-977.

Fukuda S and Hirosawa H (2001): Support vector machine classification of land cover: application to polarimetric SAR data. *Proc. Int. Geoscience and Remote Sensing Symp.,* Sydney*,* Australia, 2001

Geneletti D and Gorte BGH (2003): A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing*, vol. 24(6), pp. 1273-1286.

GEO (2005): *Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan. Reference Document*, available at: (last access Sep. 2007): http://www.earthobservations.org.

Gislason PO, Benediktsson JA, and Sveinsson JR (2006): Random Forests for land cover classification. *Pattern Recognition Letters*, vol. 27(4), pp. 294-300.

GMES (2007): *Global Monitoring for Environment and Security*. http://www.gmes.info (last access Sep. 2007)

Guerschman JP, Paruelo JM, Di Bella C, Giallorenzi MC, and Pacin F (2003): Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing*, vol. 24(17), pp. 3381-3402.

Halldorsson GH, Benediktsson JA, and Sveinsson JR (2003): Support vector machines in multisource classification. *Proc. Int. Geoscience and Remote Sensing Symp.*, Toulouse, France, 2003.

Ham J, Chen YC, Crawford MM, and Ghosh J (2005): Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(3), pp. 492-501.

Haralick RM, Shanmugam K, and Dinstein I (1973): Textural features for image classification. *IEEE Transactions on Systems, Man, Cybernetics*, vol. 3, pp. 610-621

Hegarat-Mascle S, Quesney A, Vidal-Madjar D, Taconet O, Normand M, and Loumagne C (2000): Land cover discrimination from multitemporal ERS images and multispectral Landsat images: a study case in an agricultural area in France. *International Journal of Remote Sensing*, vol. 21(3), pp. 435-456.

Henderson FM and Lewis AJ (1999): *Principles and Applications of Radar Imaging*. 3rd Ed, John Wiley & Sons, New York.

Herold ND, Haack BN, and Solomon E (2005): Radar spatial considerations for land cover extraction. *International Journal of Remote Sensing*, vol. 26(7), pp. 1383-1401.

Hill J and Sturm B (1991): Radiometric correction of multitemporal thematic mapper data for use in agricultural land-cover classification and vegetation monitoring. *International Journal of Remote Sensing*, vol. 12(7), pp. 1471-1491.

Hill MJ, Ticehurst CJ, LEE JS, Grunes MR, Donald GE, and Henry D (2005): Integration of optical and radar classifications for mapping pasture type in western Australia. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(7), pp. 1665-1681.

Hodgson ME, Jensen JR, Tullis JA, Riordan KD, and Archer CM (2003): Synergistic use of lidar and color aerial photography for mapping urban parcel imperviousness. *Photogrammetric Engineering & Remote Sensing*, vol. 69(9), pp. 973-980.

Hsu CW and Lin CJ (2002): A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, vol. 13 pp. 415-425.

Huang C, Davis LS, and Townshend JRG (2002): An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, vol. 23(4), pp. 725-749.

Huang H, Legarsky J, and Othman M (2007): Land-cover classification using Radarsat and Landsat imagery for St. Louis, Missouri. *Photogrammetric Engineering & Remote Sensing*, vol. 73(1), pp. 37-43.

Jain AK, Duin RPW, and Mao JC (2000): Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(1), pp. 4-37.

Janz A, Schiefer S, Waske B, and Hostert P (2007): imageSVM - A user-oriented tool for advanced classification of hyperspectral data using support vector machines. *Proc. of the 5th workshop of the EARSeL SIG on Imaging Spectroscopy,* Bruges, Belgium, 2007.

Jensen JR (1996): *Introductory digital image processing: A remote sensing perspective.* 2nd Ed., Prentice Hall, Upper Saddle River.

Jeon B and Landgrebe DA (1999): Decision fusion approach for multitemporal classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37(3), pp. 1227-1233.

Kareiva P, Watts S, McDonald R, and Boucher T (2007): Domesticated nature: Shaping landscapes and ecosystems for human welfare. *Science*, vol. 316(5833), pp. 1866-1869.

Keehrti SS and Lin CJ (2003): Asymptotic behavior of support vector machines with Gaussian kernel. *Neural Computation*, vol. 15 pp. 1667-1689.

Kittler J (1998): Combining classifiers: A theoretical framework. *Pattern Analysis and Applications*, vol. 1(1), pp. 18-27.

Kuehbauch W. and S. Hawlitschka (2003): Remote Sensing - A Future Technology in Precision Farming. *Proc. of POLinSAR*, Frascati, Italy, 2003.

Laliberte AS, Fredrickson EL, and Rango A (2007): Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. *Photogrammetric Engineering & Remote Sensing*, vol. 73(2), pp. 197-207.

Lardeux C, Frison PL, Rudant JP, Souyris JC, Tison C, and Stoll B (2006): Use of the SVM classification with polarimetric SAR data for Land use cartography. *Proc. Int. Geoscience and Remote Sensing Symp.*, Denver, USA, 2006.

Laur H *et al*. (2002*):* Derivation of the backscattering coefficient σ0 in ESA ERS SAR PRI products*. ESA Document ES-TN-RE-PM-HL09*, Issue 2, Rev. 5d.

Lawrence RL, Wood SD, and Sheley RL (2006): Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, vol. 100(3), pp. 356-362.

Lee JS (1986): Speckle suppression and analysis for synthetic aperture radar images. *Optical Engineering*, vol. 25(5), pp. 636-643.

Lee JS and Jurkevich I (1989): Segmentation of SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27(6), pp. 674-680.

Lee JS, Jurkevich I, Dewaele P, Wambacq P, and Oosterlinck A (1994): Speckle Filtering of Synthetic Aperature Radar Images: A Review. *Remote Sensing Review*, vol. 8, pp. 313-340.

Lee and Warner (2006): Segment based image classification. *International Journal of Remote Sensing*, vol. 27(16), pp.3403-3412.

Lillesand TM, Kiefer RW, and Chipman JW (2000): *Remote sensing and image interpretation*. 4th Ed., John Wiley & Sons, Chichester, New York.

Liu WG, Gopal S, and Woodcock CE (2004): Uncertainty and confidence in land cover classification using a hybird classifier approach. *Photogrammetric Engineering & Remote Sensing*, vol. 70(8), pp. 963-971.

Lloyd CD, Berberoglu S, Curran PJ, and Atkinson PM (2004): A comparison of texture measures for the per-field classification of Mediterranean land cover. *International Journal of Remote Sensing*, vol. 25(19), pp. 3943-3965.

Lobo A, Chic O, and Casterad A (1996): Classification of Mediterranean crops with multisensor data: Per-pixel versus per-object statistics and image segmentation. *International Journal of Remote Sensing*, vol. 17(12), pp. 2385-2400.

Lopes A, Touzi R, Nezry E (1990): Adaptive speckle filters and scene heterogeneity, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28(6), pp. 992-1000.

MARS (2007): *Monitoring Agriculture through Remote Sensing techniques*. http://www.marsop.info/ (last access Sep. 2007).

Marcal ARS, Borges JS, Gomes JA, and Da Costa JFP (2005): Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing*, vol. 26(7), pp. 1347-1362.

Mather PM (1999): *Computer processing of remotely sensed images: An introduction*. 2nd Ed., John Wiley & Sons, Chichester, New York.

Melgani F and Bruzzone L (2004): Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42(8), pp. 1778-1790.

Michelson DB, Liljeberg BM, and Pilesjo P (2000): Comparison of algorithms for classifying Swedish landcover using Landsat TM and ERS-1 SAR data. *Remote Sensing of Environment*, vol. 71(1), pp. 1-15.

Millennium Ecosystem Assessment (2005): *Ecosystems and Human Well-being: Current State and Trends*. Island Press, Washington, DC.

Muller SV *et al*. (1998): Accuracy assessment of a land-cover map of the Kuparuk river basin, Alaska: considerations for remote regions. *Photogrammetric Engineering & Remote Sensing*, vol. 64(6), pp. 619– 628.

Nyoungui AN, Tonye E, and Akono A (2002): Evaluation of speckle filtering and texture analysis methods for land cover classification from SAR images. *International Journal of Remote Sensing*, vol. 23(9), pp. 1895-1925.

Oliver CJ and Quegan S (2004): *Understanding SAR images*. SciTech Pub., Raleigh, NC.

Pal M and Mather PM (2003): An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, vol. 86(4), pp. 554-565.

Pal M and Mather PM (2005): Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, vol. 26(5), pp. 1007-1011.

Pal M (2005): Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, vol. 26(1), pp. 217-222.

Pal M and Mather PM (2006): Some issues in the classification of DAIS hyperspectral data. *International Journal of Remote Sensing*, vol. 27(14), pp. 2895-2916.

Peter N (2004): The use of remote sensing to support the application of multilateral environmental agreements. *Space Policy*, vol. 20(3), pp. 189-195.

Quegan S, Toan TL, Yu JJ, Ribbes F, and Floury N (2000): Multitemporal ERS SAR analysis applied to forest mapping. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38(2), pp. 741-753.

Quinlan JR (1993): C4.5: Programs For *Machine Learning*. Morgan Kaufmann, Los Altos.

Richards JA and Jia X (2003): *Remote Sensing Digital Image Analysis: An Introduction*. 3rd ed., Springer, New York.

Richards JA (2005): Analysis of remotely sensed data: The formative decades and the future. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(3), pp. 422-432.

Richter R (1996): Atmospheric correction of satellite data with haze removal including a haze/clear transition region. *Computers & Geosciences*, vol. 22(6), pp. 675-681.

Richter R (1996): A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing*, vol. 17(6), pp. 1201-1214.

Rogan J, Franklin J, and Roberts DA (2002): A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sensing of Environment*, vol. 80(1), pp. 143-156.

Rosenqvist A, Milne A, Lucas R, Imhoff M, and Dobson C (2003): A review of remote sensing technology in support of the Kyoto Protocol. *Environmental Science & Policy*, vol. 6(5), pp. 441-455.

Sanderson EW, Jaiteh M, Levy MA, Redford KH, Wannebo AV, and Woolmer G (2002): The human footprint and the last of the wild. *Bioscience*, vol. 52(10), pp. 891-904.

Schowengerdt RA (1996): On the estimation of spatial-spectral mixing with classifier likelihood functions. *Pattern Recognition Letters*, vol. 17(13), pp. 1379-1387.

Schölkopf B and Smola A (2002): *Learning with Kernels*. MIT Press, Cambridge, MA.

Sebald DJ and Bucklew JA (2001): Support vector machines and the multiple hypothesis test problem. *IEEE Transactions on Signal Processing*, vol. 49(11), pp. 2865-2872.

Serpico SB and Roli F (1995): Classification of multisensor remote-sensing images by structured neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33(3), pp. 562-578.

Shackelford AK and Davis CH (2003): A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41(10), pp. 2354-2363.

Simard M, Saatchi SS, and De Grandi G (2000): The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38(5), pp. 2310-2321.

Smith JH, Wickham JD, Stehman SV, and Yang L (2002): Impacts of Patch Size and Land-Covr Heterogeneity on Thematic Image Classification Accuracy. *Photogrammetric Engineering & Remote Sensing*, vol. 68, pp. 65-70.

Smith GM and Fuller RM (2001): An integrated approach to land cover classification: an example in the Island of Jersey. *International Journal of Remote Sensing*, vol. 22(16) pp. 3123-3142.

Soares JJ, Rennó CD, Formaggio AR, da Costa Freitas Y. C, and Frery AC (1997): An Investigation of the Selection of Texture Features for Crop Discrimination Using SAR Imagery. *Remote Sensing of Environment*, vol. 59(2), pp. 234-247.

Song M, Civco DL, and Hurd JD (2005): A competitive pixel-object approach for land cover classification. *International Journal of Remote Sensing*, vol. 26(22), pp. 4981-4997.

Song XM, Fan GL, and Rao M (2005): Automatic CRP mapping using nonparametric *Machine Learning* approaches. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(4), pp. 888-897.

Stankiewicz KA (2006): The efficiency of crop recognition on ENVISAT ASAR images in two growing seasons. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44(4), pp. 806-814.

Steele BM (2000): Combining multiple classifiers: An application using spatial and remotely sensed information for land cover type mapping. *Remote Sensing of Environment*, vol. 74(3), pp. 545-556.

Tanre D, Deschamps PY, Duhaut P, and Herman M (1987): Adjacency effect produced by the atmospheric scattering in thematic mapper data. *Journal of Geophysical Research-Atmospheres*, vol. 92(10), pp. 12000-12006.

Tanré D *et al*. (1990): Description of a computer code to simulate the satellite signal in the solar spectrum: the 5S code. *International Journal of Remote Sensing*, vol. 11, pp. 659-668.

Tilman D (1999): Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proc. of the National Academy of Sciences,* vol. 96(11), pp. 5995-6000.

Tso B and Mather PM (1999): Crop discrimination using multi-temporal SAR imagery. *International Journal of Remote Sensing*, vol. 20(12), pp. 2443-2460.

Turmer K and Gosh J (1996): Classifier combining: analytical results and implications, *Proc. of Conference on Artificial Intelligence*, Portland, USA.

Ulander LMH (1996): Radiometric slope correction of synthetic-aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34(5), pp. 1115-1122.

USGS - United States Geological Service (1984): *Landsat User Handbook*.

van der Linden S, Janz A, Waske B, Eiden M, and Hostert P (2007): Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. *Journal of Applied Remote Sensing*, submitted.

van der Linden S, Waske B, and Hostert P (2007): Towards an optimized use of the spectral angle space. *Proc. of the 5th workshop of the EARSeL SIG on Imaging Spectroscopy*, Bruges, Belgium, 2007.

Van Zyl, Chapman BD, Dubois P, and Shi JC (1993): The effect of topography on SAR calibration. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31(5), pp. 1036-1043.

Vapnik VN (1998): *Statistical Learning Theory*. John Wiley & Sons, Chichester, New York.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997): Human domination of Earth's ecosystems. *Science*, vol. 277(5325), pp. 494-499.

Waske B, Braun M, and Menz G (2007): A segment-based speckle filter using multi-sensoral remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, 4 (2), pp. 231- 235.

Wilkinson GG (2005): Results and implications of a study of fifteen years of satellite image classification experiments, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(3), pp. 433-440.

Zambon M, Lawrence R, Bunn A, and Powell S (2006): Effect of alternative splitting rules on image processing using classification tree analysis. *Photogrammetric Engineering & Remote Sensing*, vol. 72(1), pp. 25-30.