Mapping of Coca Cultivation Areas in the Meta Guaviare Region of Colombia Using an Object-based Approach of Satellite Image Analysis

Thesis · June 2007

1 author:

Leonhard Suchenwirth
TU Wien

DIPLOMARBEIT

Titel der Diplomarbeit

Mapping of Coca Cultivation Areas in the Meta-Guaviare Region of Colombia using an Object-based Approach of Satellite Image Analysis

Verfasser

Leonhard SUCHENWIRTH

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, im Juni 2007

Studienkennzahl laut Studienblatt:	A 455
Studienrichtung laut Studienblatt:	Kartographie und Geoinformation (Stzw)
Betreuer:	Univ.-Prof. Dipl.-Ing. Dr. Wolfgang Kainz

Contents

Acknowledegements

First of all I would like to thank the IVFL of the BOKU for the possibility to write my diploma thesis there. Especially to my supervisor, Thomas Bauer, I owe a great debt of gratitude for the good advice, confirming words and his almost permanent attainability; but also to Prof. Werner Schneider, as well as the whole team of the institute and my fellows writing their diploma thesis, for their warm-hearted integration. For the provision of data sets, I also want to thank the UNODC- ICMP.

I would also like to thank Prof. Kainz, the official supervisor of my thesis, as well as my other teachers at the University of Vienna, who have given me the theoretical and practical backgrounds during my studies. Thanks are also due to all my fellow students at university and all my friends that have encouraged me through all the years of studies. Especially to Marlena and Vicky, I am very grateful for their careful readings of the drafts.

Of course I also want to express my thanks to my parents, my brothers and sisters and the rest of my family, for their permanent support, encouragement and understanding.

List of Figures

List of Tables

Kurzfassung/Abstract

In der vorliegenden Arbeit geht es um die Automatisierung der bisher überwiegend visuell durchgeführten Interpretation von Satellitenbildern zur Überwachung von Kokaanbauflächen in Kolumbien. Die Arbeit steht im Zusammenhang mit dem Programm zur Überwachung von illegalen Anbauflächen (ICMP) des UNO-Büros für Drogen- und Kriminalitätsbekämpfung (UNODC). Dabei soll Expertenwissen mit einer Kombination mehrer Ansätze umgesetzt werden: Zum einen ist das die objektbasierte Bildinterpretation, die es erlaubt, das Satellitenbild aufgrund von Segmenten und nicht von einzelnen Pixeln zu analysieren. Zum anderen wird das Expertenwissen mit Hilfe eines Interpretationsschlüssels umgesetzt werden. Neben klassischen GIS-Abfragen sollen dabei auch Eigenschaften, die aufgrund menschlicher Erfahrung definiert worden sind, mit Hilfe von Fuzzy Logic beschrieben werden. Die Ergebnisse werden mit Referenzdaten verglichen, die auf einer visuellen Interpretation von Satellitendaten, aber auch aufgrund von Beobachtungen während Überflügen gemacht wurden. Aufgetretene Schwierigkeiten sowie die Anwendbarkeit der Methode in der Praxis werden diskutiert.

The thesis at hand presents the automatisation of the interpretation of satellite images to monitor coca cultivation areas in Colombia. Up to now, this task is executed visually. The thesis stands in context with the Illicit Crop Monitoring Programme (ICMP) of the United Nations Office on Drugs and Crime (UNODC). To achieve that, expert knowledge shall be applied using a combination of approaches: On the one hand, it is object-based image interpretation, allowing to analyse the satellite image on the base of segments instead of single pixels. On the other hand the expert knowledge shall be implemented with the help of an interpretation key. Besides classical GIS queries also features that are defined by human experience shall be described with fuzzy logic. The results are compared to reference data that rely on a visual interpretation of satellite images, but also on observations made during overflights. Arising difficulties and the applicability in practice are discussed.

1 Introduction

The first chapter shall introduce the reader to the problem of coca cultivation in the Andean states, especially Colombia, as well as the consumption of coca products; the reader shall learn about the United Nations Office of Drug and Crime in Vienna, and its programme to monitor the cultivation of coca plants, and the efforts so far made to monitor illicit crops.

1.1 Coca Cultivation in Colombia and other Andean States

Coca is historically spread in the Northern Andean Region of South America. This region, including the countries Bolivia, Peru and Colombia, is generally regarded as the most important area for cultivation of coca and its refinement and production of cocaine. According to the United Nations Office on Drugs and Crime (UNODC) in Vienna, there is no indication of large-scale coca cultivation outside these three countries (UNODC 2005).

Based on observations of the coca cultivation in the Andean area conducted by UNODC, the total production has slightly been increasing in the years 2004-2005 after a low value in 2003, anteceded by an absolute peak in the year 2000, but has seen different developments in the different countries.

Colombia can be regarded as the main producer of coca; in 2005, 54% of the entire coca yield was cultivated in the northernmost of the three Andean countries. The total area under cultivation increased in 2005 by 6,000 hectares to about 86,000 hectares, which is an 8% increase to the year 2004 with about 80,000 ha. The years before, cultivation was on the retreat since the year 2000, where there was an annual estimate of 163,300 ha.

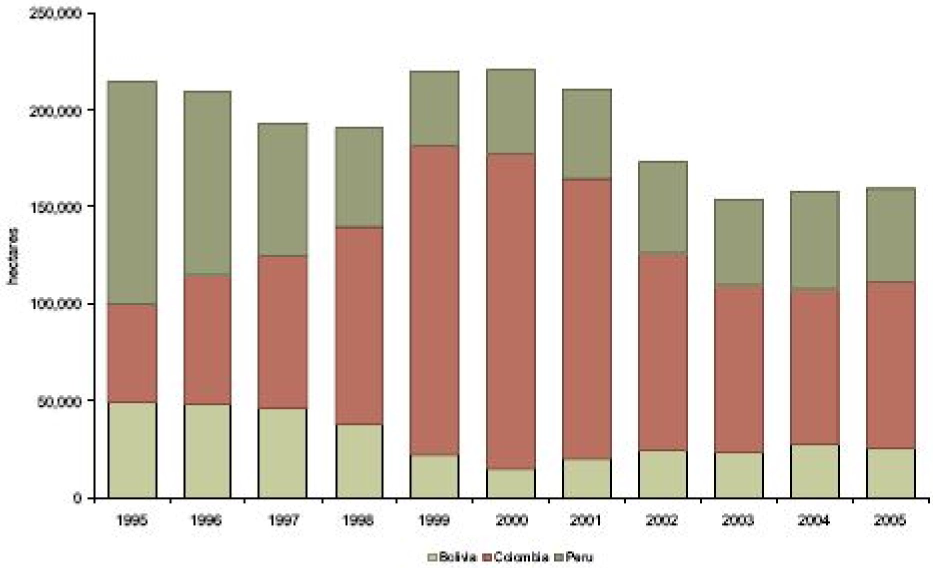

Fig. 1: Coca cultivation in the Andean region (ha), 1995-2005 (UNODC 2005)

Peru is the second largest producer of coca leaves with about 30% of the production in 2005, equalling 48,200 ha, a decrease of about 4% compared to the year 2004. The production has been remaining stable on that level for a couple of years since 1998, following a decrease from a much higher level in 1995.

Bolivia is in third place with 16% of the cultivation area, which is about 25,400 ha. In comparison to 2004, there was an 8% decrease. The lowest value could be observed in Bolivia in the year 2000, which followed a series of high-level coca production years.

While in Bolivia and Peru coca is traditionally grown for local use (e.g. for chewing and as a calcium substitute), it was introduced to Colombia in the recent decades, starting in the 1970ies. Before, it had only a marginal traditional use. Virtually the whole Colombian coca leaf production is destined for the production of cocaine; coca leaves are processed to coca paste or cocaine base in small laboratories ("kitchens") located on the farms. In Peru and Bolivia however, sun-dried leaves are traded, with regulations for traditional, commercial and industrial uses by the national governments; leaves that are traded outside this controlled market are determined for the illegal cocaine production.

In Colombia, 640 tons of cocaine were produced in 2005, about exactly as much as in 2004. During the ten years before, cocaine production has increased strongly, reaching a peak with 695 tons in 2000, having a low in 2003 with about 550 tons, and then rising up again to the last value.

In Peru and Bolivia, cocaine production has declined from more than 430 respectively 250 tons since the mid-1990ies, with the lowest value in 2000 (141 respectively 43 tons), having a smaller rise in the last years up to 190 respectively 107 tons (2004).

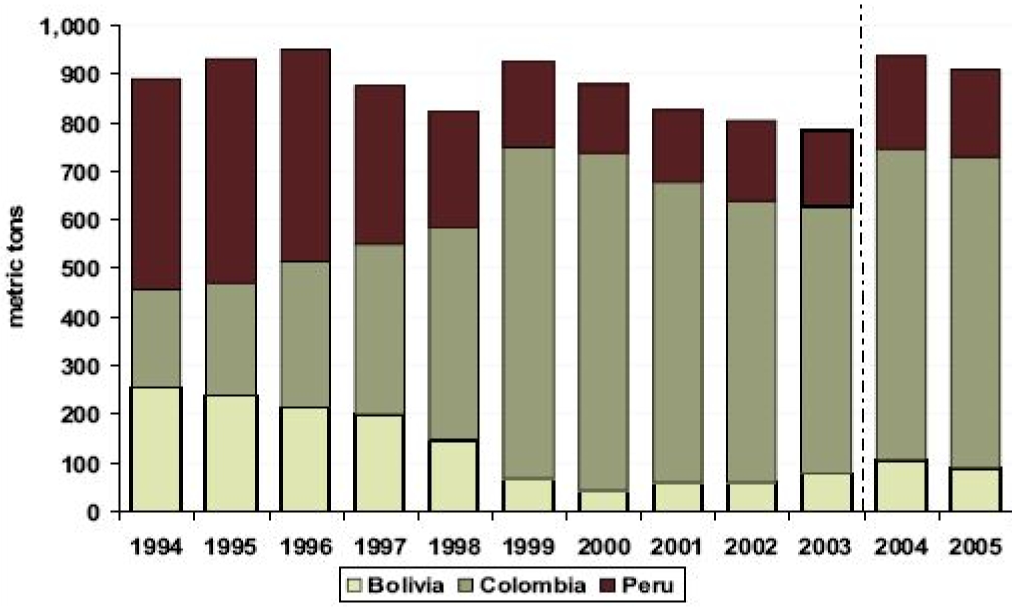

Fig. 2: Cocaine production in the Andean region (UNODC 2005)

Coca cultivation is widespread over the whole country of Colombia, with almost every department (*departamento*) being affected. Yet there are some hotspots which represent the main producing areas. Among the departments, Meta, Nariño, Putumayo, Guaviare, Vichada, Antioquia and Caquetá are most important for the cultivation. In these seven departments, about 78 percent of the cultivation in 2005 has been detected. Departments with coca cultivations are treated as regions by the UNODC surveys; actually, there are seven cultivation regions. Among these regions, Meta-Guaviare, Pacific, Central and Putumayo-Caquetá are the most important.

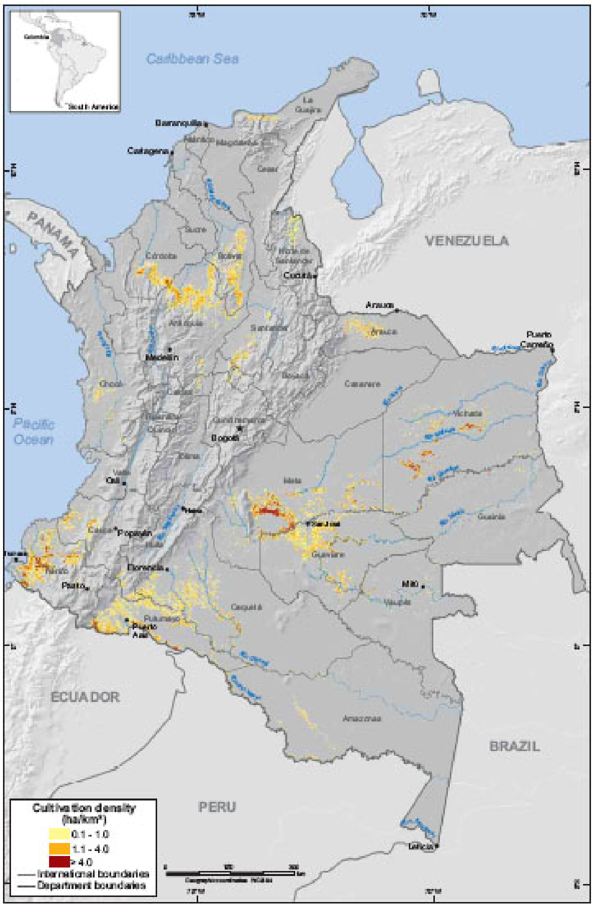

Fig. 3: Coca cultivation density in 2005, Colombia (UNODC 2005)

1.2 Consumption of Coca and Cocaine

Traditional uses of coca are chewing or eating the leaves, as well as coca-leaf tea, known as mate de coca. This tea reminds in its taste and preparation of normal black tea, and has a similar stimulant effect such as tea or coffee. These traditional products are well-known for relief and suppression of hunger, exhaustion and coldness, as well as a treatment against altitude disease. Modern products derived from coca can be chemical products, e.g. toothpaste, soap, or shampoo, as well as coca can be used for candies and cakes.

Though coca can be used for a whole range of products, it is mostly known as the basic product for the production of cocaine. Although the stimulant and hunger-suppressing characteristics of coca had been known for many centuries, the isolation of cocaine alkaloid was first done by Albert Niemann in 1860. It served as the first local anaesthetics used in medicine. Chemically, cocaine ($C_{17}H_{21}NO_4$) is a benzoylmethylgocine and a methyl-8-azabicyclo octane-2-carbonic acid methyl ester.

Its most important effect is the ability to block the start or transmission of a neuronal impulse at the spot, where it is attached to the mucous membrane (e.g. nose). After entering the bloodstream it diffuses all over the body and reaches the brain, stimulating the central nervous system. The addiction is ascribed to neurotransmtters such as norepinephrine, serotonin and dopamine.

These neurotransmitters have an effect onto the mood, and an increased excitation. After longer use of cocaine, the threshold of excitiation is increased. Thus, an exhaustion resulting in listlessness and depression leads to an increased dose to reach a level of satisfaction (after Haasen 2004).

Cocaine is consumed by sniffing, while its derivates, freebase and crack can be smoked. Freebase is produced by boiling the cocaine hydrochloride with organic solvents, such as ether or chloroform. One gram of cocaine results in 0.5 to 0.7 grams of freebase. The production of crack is safer, easier and cheaper than the production of freebase. Cocaine hydrochloride is brought up to the boil with inorganic substances like ammonia, ammonium chloride, or baking powder in the U.S., resulting in beige crack pieces. Out of one gram of cocaine about 1.3 gram of crack can be produced, that are about six to eight portions. As it decays in contact with air, it has to be stored in airproof containers of glass or plastic. Its name is derived from the crackling sound when smoked in a pipe.

In the US, there has been an enormous increase in the consumption of cocaine and its derivates in the 1980ies, with a climax in the middle of the 1980ies. In recent times it has decreased, in favour of consumption of heroine. About 0.7 percent of the population older than 12 years are consumers of cocaine or one of its derivates (Stöver, 2004).

In Europe, consumption of cocaine is still growing in the UK, and in smaller extents, also in Denmark, Germany, the Netherlands, and Spain. For Germany, Kemmesies and Werse (2004) tried to analyse trends for the consumption. It has been a society drug, with a first high in the 1920ies, when it was used by bohemians, and a second high since the 1980ies, when it became a manager drug. Whether cocaine can be seen as a rather encultured (i.e. accepted in the general society) or a decultured drug, is still a controversial issue that can be answered from different viewpoints. On the one hand, the authors do not see an enculturation process, as cocaine is still a phenomenon of fringe groups, its degree of spreading is low, there are hardly traces of a so-called consumption culture, as it is the case with legal drugs, such as nicotine, or alcohol. Another important factor is the prohibition of cocaine and its social contempt. On the other hand, the authors register an increased interest of the public in cocaine consumptions, especially in the wake of scandals of public figures and celebrities.

1.3 UNODC-ICMP

The United Nations Office on Drugs and Crime (UNODC), part of the UN secretary, is trying to combat international crime and illicit drugs. It was founded in 1997, and has about 500 staff members worldwide. The headquarters are in Vienna, Austria, a liaison office is in New York and about 20 branch offices are spread worldwide. UNODC is mainly financed by voluntary contributions by national governments, which make up about 90 percent of the budget.

The three major tasks of the UNODC can be characterised as follows (UNODC homepage 2007):

- Research and analytical activities to improve the knowledge about drug and crime issues and to expand the evidence-base for policy and operational decisions;
- Support for states to ratify international treaties, to develop national legislation on drugs, crimes and terrorism, as well as the supply with basic and supporting services for the treaty-based and governing bodies;
- Field-based technical cooperation projects to enlarge the capacity of member states in their struggle against illicit drugs, crimes and terrorism.

An important project inside the UNODC is the United Nations International Drug Control Programme (UNDCP), established in 1991. In October 1998, it was renamed the United Nations Office on Drugs and Crime (UNODC), which also administrates the Fund of UNDCP. Its objectives are to educate the world about the dangers of drug abuse and to strengthen international action against drug production, trafficking and drug-related crime. That is to be achieved by alternative development projects, anti-money laundering programmes and the illicit crop monitoring programme (ICMP), to which this thesis should be a contribution.

UNODC also provides accurate statistics through the Global Assessment Programme (GAP) and helps to draft legislation and train judicial officials as part of its legal advisory programme.

The renaming from UNDCP to UNODC in 1998 was necessary, as in 1997 the responsibility for crime prevention, criminal justice and criminal law reform was given to the UNDCP. A couple of programmes shall combat transnational crime such as corruption, organized crime, human trafficking and terrorism. The office intents to collaborate with member states to strengthen the rule of law, to promote stable and viable criminal justice systems (UNODC homepage 2007).

The objectives of UNODC’s Illicit Crop Monitoring Programme (ICMP) are to establish methodologies for the collection and analysis of data, to increase governments’ capacity to monitor illicit crops and to assist the international community in monitoring the extent and evolution of illicit crops, in the context of the elimination strategy adopted by the Member

States at the U.N. General Assembly Special Session on Drugs in June 1998 (UNODC 2005).

The information provided by the annual surveys helps guide policy interventions and constitutes a tool for planning activities to tackle the illicit crops. Surveys are available for Afghanistan, Myanmar, Laos (Opium Survey), Colombia, Peru, Bolivia (coca Survey) and Morocco (Cannabis Survey).

The monitoring systems supported by UNODC are tailored to the national contexts and include a strong capacity building element. The direct participation of UNODC in the national monitoring systems ensures the transparency of the survey activities and gives additional credibility to the published results. Through its network of monitoring experts at the headquarters and in the field, the Illicit Crop Monitoring Programme ensures the conformity of the national systems with international methodological standards and with the information requirements of the international community. The ICMP facilitates the dissemination of methodological best practices among the national systems, and it assumes a quality control function for the data produced.

In Colombia, UNODC has supported the monitoring of illicit crops since 1999, and has produced six annual surveys. In October 2003, UNODC signed a new agreement with the Colombian government to continue and expand monitoring and analytical works. In this context, the SIMCI II (*Sistema Integrado de Monitoreo de Cultivos Illícitos II*) project has established to facilitate the implementation of additional tasks in the framework of an integrated approach to the analysis of the drug problem in Colombia.

The new project foresees the creation of an Inter-Institutional Committee permanently assigned to the project in order to ensuring the transfer of know-how to the national beneficiary institutions. SIMCI II is a joint project between UNODC and the Colombian government, represented by Ministry of Interior and Justice and the International Cooperation Agency. The national counterpart and director of the project is the head of the Ministry of Interior and Justice.

The project also supports the monitoring of related problems such as fragile ecosystems, natural parks, indigenous territories, the expansion of the agricultural frontier and deforestation. It provides concrete support to the government’s alternative development projects such as the “Forest Warden Families Programme”. This campaign was initialized to motivate farmers to keep their land free of illicit crops and consists of three main components (UNODC 2006):

- Environment: Productive and sustainable projects to preserve the environment with technical support of expert entities in the training of families for the establishment
- Society: Permanent training of families in community savings, leadership, project managements among others to increase the social capital
- Economy: Financial aid to the beneficiary families (temporary)

1.4 Status Quo of Coca Monitoring in Colombia

As mentioned above, coca is monitored at present by the UNODC ICMP in cooperation with national authorities. The monitoring process is done by visual interpretation of satellite images and their comparison to additional information such as former land cover classifications, and data from former coca surveys, as well as GPS-based flight paths of airplanes used for spraying campaigns. Remote sensing is used due to comprehensible security reasons, but also for reasons of efficiency in the nation-wide survey.

For the nation-wide monitoring project, several demands onto the data in use should be taken into account:

- Data have to be available in general, i.e. there should be no restrictions for military reasons or similar
- Data have to be available for the whole national territory of 1,142,000 km² (less the islands in the Pacific and the Caribbean like San Andres or Old Providence).
- Data have to be affordable
- Data have to be up-to-date, with frequent revisits
- Data should be as free of clouds as possible.

However this is difficult to attain. It is hardly possible to meet all demands on the data; especially the frequent cloud coverage has led to certain problems, as the climate in the tropical rain forest as well as in the Andes is often humid and thus cloudy. Approximately 10-20 percent of the image is covered by clouds or cloud shadows (see figure 4).

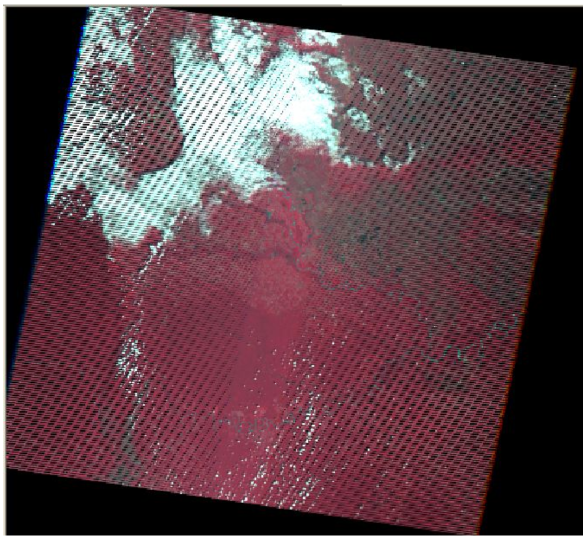

Fig. 4: Landsat 7 ETM+ image of 2004

There are actually three satellite systems and their images in use for the cocaine survey:

- Landsat7 ETM+: The Landsat 7 ETM+ (Enhanced Thematic Mapper Plus) data are collected in six spectral bands with 30 meter spatial resolution, plus an additional panchromatic band of 15 meter spatial resolution, and a 60-meter-resolution thermal infrared channel. The satellite has a 16-day repeat cycle that enhances the chance for cloud free images. The swath width is 185 km and thus appropriate for regional studies. Suitable images were identified by frequently consulting the on-line catalogue of the US Geological Survey (USGS). About 60 Landsat 7 ETM+ images would be needed to cover the whole country. Unfortunately, on the 31st of May 2003, the Scan Line Corrector (SLC) of the Landsat 7 ETM+ instrument failed. This malfunction is leading to gaps in the image, gradually diminishing towards the centre of the scene. The gaps are visible through characteristic stripes in the image. To correct them, the ICMP treated the gaps like clouds, representing an obstacle for the detection of coca fields.
- ASTER: ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) is one of the instruments of the NASA's Terra satellite. It has the highest spatial resolution and serves as a zoom lens for the other instruments aboard. The images consist of 16 spectral bands, with a spatial resolution ranging from 15 to 90 meters. Its subsystems are the Visible and Near Infrared (VNIR), the Short Wave Infrared (SWIR), and the Thermal Infrared (TIR). The vegetation cover monitoring relies mainly on the VNIR spectral bands 1,2 and 3 with a pixel size of 15 meters. The swath wide is 60 km; that requires the acquisition of substantially more images than with Landsat 7 ETM+ to cover an equivalent area. About 500 ASTER images would be necessary to cover the entire country.
- SPOT 4: The French system SPOT (*Système pour l'observation de la terre*) has a similar spatial resolution (20 meter) and swath width (60 km) as ASTER, but there are only four bands (Green, Red, NIR and MIR) as well as a panchromatic band with 10 meter resolution. The primary imaging systems on board are the *high resolution visible and infrared (HRVIR*) sensors and the *Vegetation* instrument. For covering Colombia's entire continental territory, about 500 images would be needed.

For the 2005 census, a total 68 LANDSAT images, six ASTER and 11 SPOT-4 image were analyzed, that have been captured between September 2005 and March 2006. The images cover the whole national territory of Colombia (except the islands).

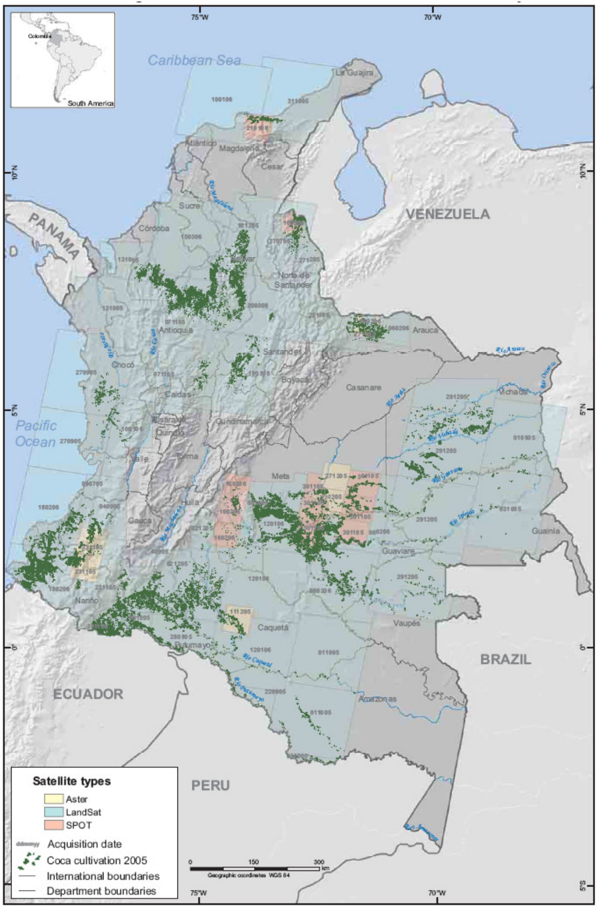

Fig. 5: Satellite images used for the Colombian coca survey 2005 (UNODC 2005)

The additional data used for the interpretation process can be found in a spatial information database (*Banco de Información Espacial*, BIE). It is to be accessed by the URL http://www.biesimci.org.

After the process of geo-referencing and coarse land cover classification of data, the visual interpretation process starts. It relies on spectral characteristics, as well as patterns and surroundings of the fields. The class coca can be described as a mixture of bare soils and rows of bushes. A phenological differentiation is not made, as there is no crop calendar for the cultivation of coca. In the interpretation process, coca crops are verified due to their spectral characteristics, their texture, shape, size of the field and contextual information (e.g.: “Has there been a coca field before?” or: “Has there been a spraying campaign and if so, when?”) At the end, a set of signatures is produced, with each signature corresponding to a class and being used with a decision rule to assign pixels to a class.

So far, coca fields are digitized on screen with the help of semi-automated procedures, such as pixel seeding. Small polygons of less than 0.25 hectares (equivalent to two or three LANDSAT 7 pixels) are deleted due to the fact that an interpretation is not reliable enough.

It was tried to remodel the entire process of interpretation with the help of an interpretation key, more precisely in form of a decision tree (see chapter 1.5).

For checking the results, verification flights by the antinarcotics police help to correct and improve the initial interpretation. This is documented by a video and a digital camera, with a mounted GPS device. These documentary photos are also an important visual aid for the land cover classification. However, these photos are no aerial photos in a strict sense (Lillesand, Kiefer, Chipman 2004). Thus they cannot serve as a substitute for satellite image. They can only serve as additional information.

As part of the programme to eradicate illicit crops, coca fields are sprayed from aircraft with toxins. These spraying lines are automatically recorded by GPS, and play an important role for the interpretation of coca fields in the following years. A buffer is calculated along the lines to model the impact of the defoliant toxin, depending on the type of the plane and the recorded spraying line. Not only the location, but also the point in time of the spraying, is important for the interpretation.

For the defoliation of coca bushes, an agent called glyphosate, formulated by the Monsanto Company as Roundup™ is used. It is described as a post-emergent, systemic and non-selective herbicide that is used widely in both agricultural and non-agricultural applications, e.g. for gardening applications (http://www.roundup.com). It is said to have been approved by the US Environmental Protection Agency for general use in 1974; in Colombia, it is mixed with water, as well as a locally produced adjuvant, Cosmo-Flux 411F; this composition is said to increase the persistence and penetration of the defoliant in to the waxy structure of the coca leafs (Messina and Delamater 2006). However, according to the UNODC (2006), there is a survival rate of ten percent of coca plants.

Due to a study on the Putumayo region done by Messina and Delamater (2006), the application of glyphosate as defoliant does not only affect coca plants and cultivation, but also food crop cultivations, such as yucca, avocados, maize, and plantains of neighbouring subsistence farmers. An unexplained difference of about 35.000 hectares between the figure calculated by that study (using fractional coverage, field data, and hybrid classification) and the official figure of reduced coca cultivation in that area by the UNODC report leads to the assumption, that besides coca also contiguous and interspersed native forest, but also food crop parcels have been collaterally damaged. Messina and Delamater (2006) come to the conclusion, that UNODC reports and data have to be read carefully and can be open to interpretation.

The SIMCI project is managed by a technical coordinator and composed of engineers and technicians: four digital image processing specialists, one field engineer, a cartographic technician, a research and analysis specialist, two assistant engineers and an administrative assistant. The team is integrated by the Inter-Institutional Committee assigned on a permanent basis to participate in the activities of SIMCI, and composed of technicians and specialists of the following government and state institutions: Ministry of Interior and Justice, its National Narcotics Bureau – DNE, Ministry of Environment and their specialized units IDEAM and Natural Parks, Ministry of Agriculture, Ministry of Social Protection (Welfare), UIAF (Ministry of Finance), Anti-Narcotics Police - DIRAN and the Geographical Institute – IGAC (UNODC 2005).

1.5 Cooperation IVFL-UNODC

Since May 2004 the Institute for Surveying, Remote Sensing and Landinformation (IVFL) of the University of Natural Resources and Applied Life Sciences of Vienna, Austria (*Universität für Bodenkultur*, BOKU) is cooperating with the UNODC ICMP. The aims of this cooperation are professional consulting and assistance as well as the research and improvement of existing methods to monitor the cultivation of illicit crops by the means of remote sensing. Its emphasis lies on the coca cultivation in the Andean states as well as the cultivation of opium poppies in the Golden triangle (Laos and Myanmar).

In September 2004, the methodology developed for the assessment of coca cultivation in Colombia was technically evaluated. The IVFL concluded that the methodology and work of the remote sensing team that performs the interpretation of the satellite images is appropriate. For quality control, the use of aerial photos was recommended for the next surveys (UNODC 2006).

To enhance and objectify the interpretation of coca fields in satellite images, an interpretation key has been developed by the project staff in cooperation with the IVFL. This task has been done in two steps: The first was to identify different factors determining the interpretation of coca fields in a selected area (departments of Meta and Guaviare).

The second step is the design of an interpretation key in form of a decision tree with the data obtained in the first step for the development of models in each region. So far, an interpretation key has been developed as an interpreter calibration technique (Bauer and Schneider 2005), as shown in figure 6. The development of the interpretation keys has been described in (Bauer and Schneider 2006;2).

First attempts have been made to formalise the Colombian interpretation key, using the software package eCognition 3.0 (Bauer and Kaiser 2006, Bauer and Schneider 2006;1).

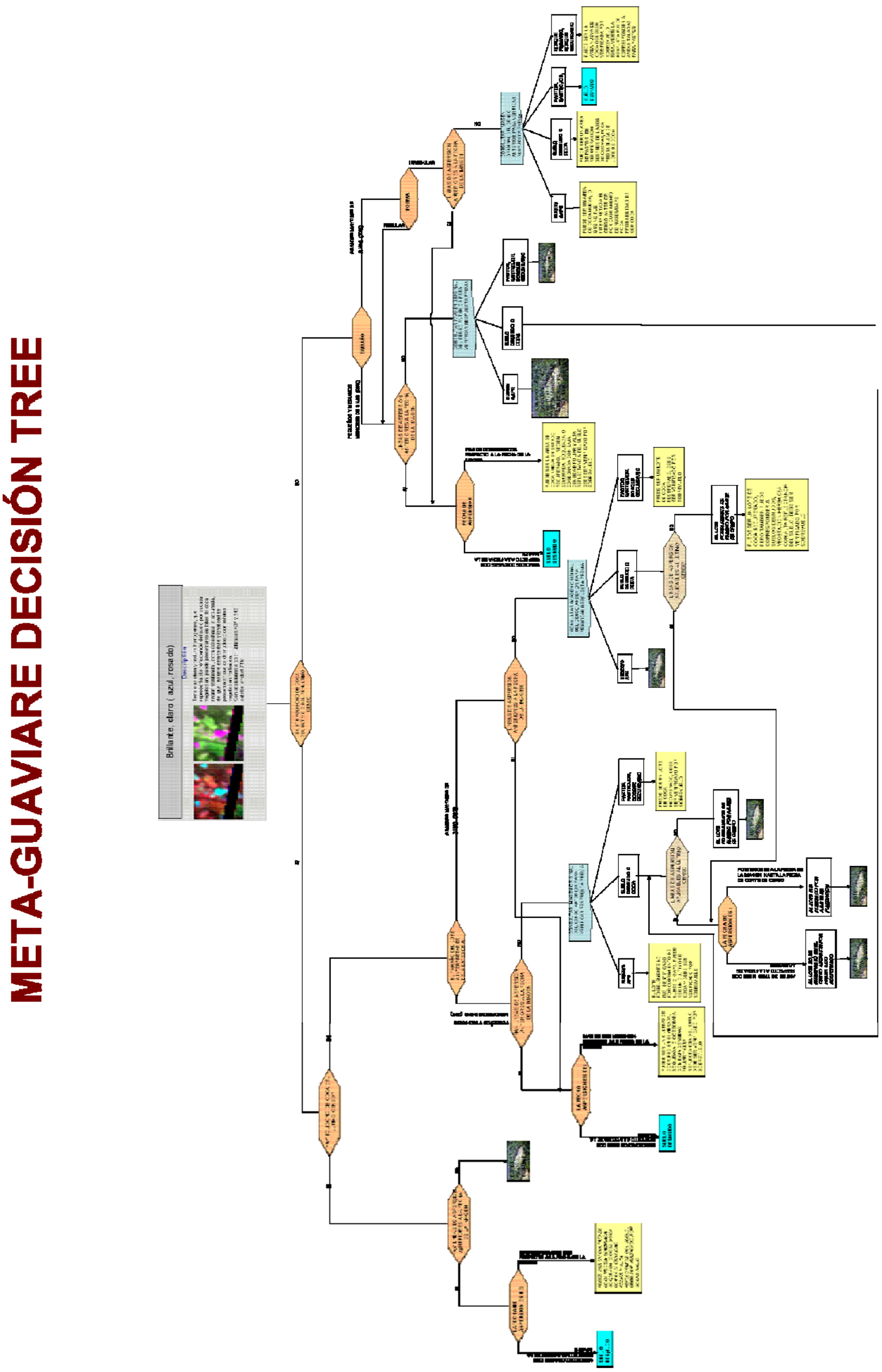

Fig. 6: Part of the Interpretation Key (Decision Tree)

1.6 Objectives of the Thesis, Presentation of Chapters

Generally this thesis should be a study of feasibility to automate the process of interpretation of coca using Landsat satellite imagery and the comparison with additional information, such as former land cover, former coca fields or spraying lines, or the query of attributes such as size, shape or texture. For this task, an object-based approach of image interpretation shall be applied, based on the software Definiens Professional 5.0 (formerly known as eCognition).

The objective of this thesis can be described as follows:

- Automating the visual interpretation of coca fields on Landsat imagery using an object-based approach, in order to objectify the results of image interpretation, i.e. the interpretation results shall no longer be dependent solely on the interpreter, but also on a traceable process in the computer.
- Finding apt segmentation parameters in order to partition the image into image objects that can represent coca fields properly concerning size, shape and texture.
- Modelling the interpretation key (decision tree) with the help of adequate software (in this case Definiens Professional 5.0), and eventually modifying it.
- An adequate validation of results.

The thesis is structured into the following chapters:

- Chapter 1 gives a general introduction into the subject matter coca cultivation, coca use and coca monitoring.
- Chapter 2 gives an introduction into object-based image interpretation.
- Chapter 3 explains the concept of image segmentation, the basic principles and technical implementation.
- Chapter 4 explains the basic principles of pixel- and object-based classification, the principles of fuzzy logic and interpretation keys.
- Chapter 5 provides the example and practical application of object-based image interpretation to detect coca fields in the Meta-Guaviare region.
- Chapter 6 shall give an overview and evaluation of the results of the interpretation conducted in chapter 5.
- Chapter 7 shall give a conclusion of the experiences and cognitions made during the interpretation process.
- Chapter 8 comprises the literature that has been used for the thesis.

2 Object based Image Analysis

In image interpretation, a new approach has been introduced during the last years. While conventional methods of image interpretation work with individual, single pixels to analyze and classify satellite images, the new approach works with image objects. What are the reasons and arguments for the new approach? Chapter 2 shall give the reader an introduction to object based image interpretation, its definition and differences to pixel-based classification.

2.1 Image Classification in General

"The overall objective of image classification procedures is to automatically categorize all pixels in an image into land cover classes or themes" (Lillesand, Kiefer, Chipman 2004). Generally, spectral, spatial and temporal image classification is distinguished (Lillesand, Kiefer, Chipman 2004). In each kind of classification, distinctive patterns are tried to be recognized. The image classifiers may also be used in combination in a hybrid mode. The recognition of patterns can be equated to the low-level, but also at least to the beginning of high-level vision by the human eye and mind (Peuquet 2002), as explained in chapter 2.1.

The automatic categorisation of all pixels in an image into land cover classes or themes is an overall objective of image classification procedures. Typically, multispectral data are applied for the classification, using the spectral pattern within the data as a numerical basis for the categorization. Therefore, different feature types show different combinations of digital numbers (DN), based on their inherent spectral reflectance and emittance properties. Therefore, spectral patterns are not geometric patterns; spectral pattern recognition refers to the family of classification procedures using the pixel-by-pixel spectral information as the basis for automated land cover classification.

The recognition of spatial patterns involves the categorization of image pixels on the basis of the spatial relationship with pixels surrounding them. When classifying spatially, parameters such as texture, pixel proximity, size, shape, directionality, repetition and context might be regarded. These classifiers try to substitute the synthesis done by the human mind during a visual interpretation process. Hence, they have the tendency to be much more complex and computationally intensive than recognition procedures of spectral patterns.

Temporal classification and pattern recognition use time as an aid for the identification of features. This is important for agricultural crop surveys, where distinct spectral and spatial changes during a growing season can permit discrimination on multidate imagery that would be impossible if only a single date was given. Lillesand, Kiefer and Chipman (2004) give an example of winter wheat, which is not distinguishable from bare soil when freshly seeded in autumn and spectrally similar to a lucernes field in springtime. An interpretation of only one date would be unsuccessful, hence, only the analysis of data from both dates leads to a successful interpretation.

The most important and 'typical' classification is the spectral pattern recognition. Two principles of spectral pattern recognition are applied, the supervised and the unsupervised classification.

In the supervised classification, spectral signatures are defined interactively in training areas, giving the user the possibility to control statistic parameters such as frequency or mean value. Subsequently, the image is classified; after classification, results are controlled and the classification is repeated, if necessary, or completed. Classes represent thematic classes; to check the reliability of classification of the entire image an accuracy assessment is performed. For the accuracy assessment, test areas that should not be identical to the training areas are used.

In the unsupervised classification, clusters are built automatically. Manipulation of classes is possible only on a low base. This classification represents spectral classes; thematic classes have to be assigned subsequently.

Thematic classes can be represented by more than spectral signature; however, this fact can be solved by the aggregation of classes, also called reclassification. On the other hand, one spectral class can represent more than one thematic class; in this case, additional information (e.g. temporal) may be needed for a separation into thematic classes.

Object classes represent clusters of pixels in a feature space, therefore a classification is a geometric division of a feature space. This multispectral feature space has as many dimensions as the image has spectral bands. In supervised classification, the following algorithms are distinguished:

- In a hyperbox or parallelepiped classification classes are described as n-dimensional cuboids. This method describes the shape of the cluster better than e.g. a minimum-distance classification and it is simple to calculate, however, not the entire feature space is assigned, as the description of clusters is restricted by the rectangular shape of the boxes, with overlapping areas leading to decision problems.
- In a minimum distance classification, for each class a mean is calculated; for each unknown pixel, the distances to all class means are calculated, and the pixel is assigned to the class with the smallest distance. Although it is mathematically simple and computationally efficient, this method has certain limitations, such a lacking sensitivity to different degrees of variance in the spectral response data.
- The maximum likelihood classification is based on the assumption that values have a Gaussian normal distribution. The algorithm is said to be the best description of the cluster, as the feature space can be classified totally, but it is numerically extensive and causes long calculating times. It is the standard method for the classification of multispectral remote sensing data. Statistically, classes are described by spectral

signatures, having a mean value, a variance of the class in the n spectral channels, and a covariance between the n spectral channels.

2.2 Pixel-based vs. object-based Image Classification

Besides the traditional pixel-based methods, as presented in chapter 2.1, object-based image classification has become more and more important in recent years. In contrast to the analysis of single pixels in the image, the object-based approach works with image objects that result from the segmentation of the image into homogenous regions, or from another perspective, the merge of those single pixels that are similar to each other. The principles of image segmentation shall be described in chapter 3.

In an article by Blaschke and Strobl (2000), it is argued that conventional pixel-based image analysis is to a fewer extent able to represent reality than by image objects. While pixel-based methods of image classification and interpretation deal with the analysis of statistical values of the satellite image, the spatial arrangements of the individual pixels are neglected. These spatial arrangements can be better recognized by image objects than by single pixels.

As Blaschke and Strobl argue, entire disciplines like geography are based rather on objects, than on pixels. In general, this difference is also known in GIS, where the difference exists between raster and vector data. In cartography, it is the idea of continuous or discrete phenomena and objects, also called continuum and discretum respectively (Hake, Grünreich, Meng 2002). These discrete objects can have a wide variety of names, such as image objects, segments, patches, and so on. Independent from their notion they share in common to be objects that have a meaning.

It is argued that object-based image interpretation is much more able to consider shape, size or colour and context information such as neighbourhood, distance and location. Especially for artificial, man-made objects, the concept of discrete objects is appropriate; this is in particular the case when working with very high spatial resolution (VHSR) data (Steinnocher et al. 2007). For natural features, object-based image classification can be used either to update existing geoinformation, or to delineate new objects that have not been mapped before.

However, object-based image information (and underlying segmentation into homogeneous objects) is facing problems, when dealing with soft boundaries that often occur in nature. In land cover there are clear, but often soft transitions, e.g. between several types of natural vegetation. An approach to handle this problem of gradual transitions is the application of fuzzy logic for classification.

Baatz and Schäpe (2000) also note a certain moment of historicity resulting in a loss of full reproducibility, as different decisions in the past may cause different decisions in the future. However, the arising differences are said to be found on edges between objects of low contrast, where the shape of an edge can be arbitrary.

In order to understand the basic principles of this new paradigm and an eventually looming paradigm shift, it is necessary to understand the basic principles of human image perception. They shall be elucidated in the following.

2.3 Human Image Perception

To understand the idea behind object-based image interpretation, it is important to know the basic concepts underlying, and the way human thinking processes (visual) information. The human mind does not only register. Processes in the mind are highly active and integrated cognitive mechanisms, which unfold in many steps between direct sensual inputs on the one hand and categories and our world knowledge on the other hand. Especially the visual sense is based on this principle.

As an illustrative example, a cartoon (figure 7) is given, where a person does not just see the apple cut into halves, but internally adds information about its consistence, similar plants, literary references, etc. The perception is therefore a construction of an object, invoked by the confrontation of sensual inputs with the knowledge base, based on cognitive processes which are activated in the mind. These inner images also include image objects with diverse relations to each other, with different types of contexts, for example spatial, temporal or functional, and to our knowledge base. Each situation has got multiple connections to events of spatial or temporal cognition. (User Guide eCognition 3, 2003)

Starting with primitive, more technical processing such as the contrast-increasing signal flow between the neurons behind the eye's retina, a process of growing abstraction and incorporation of semantic knowledge begins. It precedes from the construction of cognitive objects over primitive assignment of meaning to more and more detailed, elaborated and affective meaning, and thus also conscious functionality. However, these sequences are not rigid, but include loops of local proof and enhancement. Human perception is also influenced by emotions and affects, with each cognitive event being evaluated with affective connotation. This affect plays an important role concerning the way of further processing of information and concerning memory request. Emotions are filters which guide the focus of attention which of the many sensual, cognitive inputs is of current importance and which is not (User Guide eCognition 3, 2003).

In cognitive literature, there is a common differentiation of visual information processing into low-level and high-level vision (Ballard and Brown 1982; Kosslyn and Koenig 1992). Low-level vision is the initial process, where the raw retinal image is absorbed by the eyes; information about colour and tone, and the basic figure-ground-discrimination of the object is gained. High-level vision is based on the visual information and uses information that has been acquired by the low-level vision. Previously acquired and stored knowledge is used in order to identify the seen object, as well as to interpret and associate it with other objects. The identified object is also linked with other unseen properties. Thus, meaning is given to the scene as a whole. While low-level vision happens at a rather "hardwired" neurological level, high level vision happens at a mental level (Peuquet 2002). Neisser (1976) refers to low-level vision as the preattentive stage of cognition.

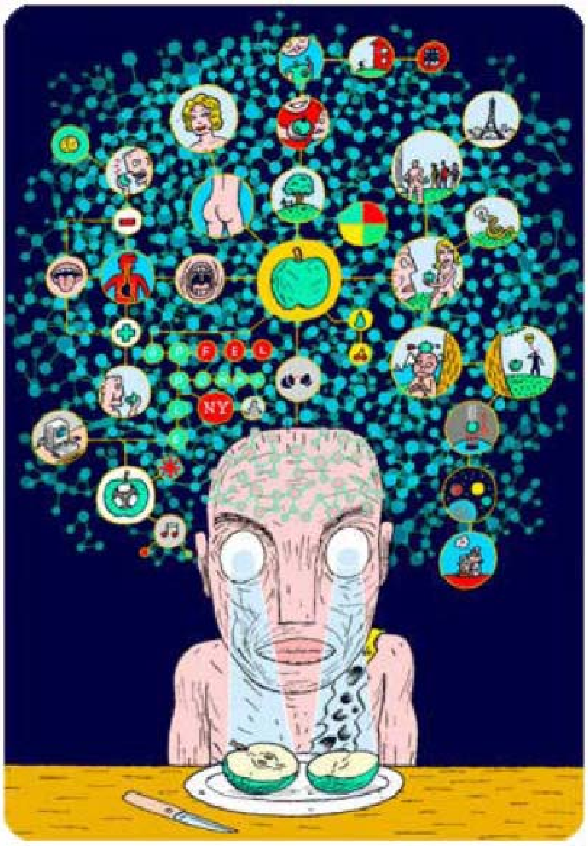

Fig. 7: Human perception of reality (User Guide eCognition 2.0)

All in all it can be concluded that humans have, contrary to computers, the ability of cognition. Most objects in nature can be observed and by comparing the newly acquired information to already existing information we can build up a vast knowledge base, as already described. Differences can be spotted easily. Furthermore, human beings are very good at spotting patterns. We can look at a picture as a whole, and see patterns, or detect differences in colours. Computers clearly do not have cognitive abilities, but they can do calculations extremely fast and precisely.

2.4 Definition and Evaluation of object-based Image Analysis

After having gained insight into the basic principles of human image perception, and the delineation to pixel-based classification, a further definition of object-based image interpretation, also known as object-based image analysis (OBIA) is necessary. As explained before, it is the process of segmenting, i.e. partitioning images into segments, and working with these segments instead of single pixels.

In geoinformation, especially in the field of remote sensing, object-based image interpretation has become an important tool, but also for computer vision and biomedical imaging. However its use has already been extended to other purposes, e.g. the geologic dating and analysis of fission-tracks (Marpu, Gloaguen, Jonckheere 2006).

It is applied to emulate or to improve human image perception and interpretation of remote sensing images by (semi-)automated means. Additionally, the process has increased repeatability as well as productivity at reduced subjectivity, labour and time.

The notion of object-based image information or analysis is still in a definition process, and is therefore a vague entity. Even the official name of this discipline is not yet readily delineated. There is still a wide range of potential alternative names. A Wiki, which was initiated by the University of Calgary after the 1st International Conference on Object-Based Image Analysis at the AGIT 2006 in Salzburg, shows some of the alternative names suggested (and their abbreviations). Each name implies a slightly different connotation (Wiki OBIA 2006):

- Geo-Object Based Image Analysis (GEOBIA, or also GOBIA)
- Object-based Analysis of Remote Sensing Images (OARS)
- Geographic Image Object Analysis (GIOA)
- Object-Based Analysis of Remotely Sensed Imagery (OBARSI)

Besides the term ‘object-based’, the term ‘object-oriented’ is used frequently. Nonetheless, ‘object-oriented’ is used more frequently for programming issues, and less frequently for the techniques of image interpretation.

So far, no consensus about the name has been reached yet. However, the term Object-Based Image Analysis is widely spread in the literature. In this thesis, this term shall be used as well.

Ditto in the Wiki, the current state of the discipline has been analyzed by the help of SWOT-analysis. It should help to detect strengths, weaknesses, opportunities and threats to this technology (Wiki OBIA 2006, Blaschke and Strobl 2001):

Strengths of object-based image analysis:

- Partition of image into objects emulates the way humans conceptually comprehend the landscape and organize it.
- Use of image objects as basic units reduces computational classifier load by orders of magnitude and enables the user to use more advanced techniques (e.g. non-parametric)
- Image objects provide features such as size, shape, texture or context relations that cannot be found at single pixels; especially texture is an important feature, also used for industrial vision.

- Image objects are less sensitive to the Modifiable Area Unit Problem (MAUP) than units which do not correspond with the structure of the phenomenon under study. The Modifiable Areal Unit Problem is described a potential source of error affecting spatial studies based on aggregated data sources (Unwin 1996).
- Image objects can easier be integrated into vector-based GIS than pixel-based classified raster images; as Blaschke and Strobl (2001) put it, object-based image interpretation is a bridge between RS and GIS.
- Several methods and packages by commercial software use an object-oriented paradigm.

Weaknesses of object-based image analysis:

- Software products are overly complicated, due to their flexibility adjustment tools.
- Large datasets are difficult to process, e.g. images with a two-digital number of mega-pixels; enhanced tiling or multiprocessing solutions are needed.
- No unique solution of segmentation; e.g. a changed bit-depth of heterogeneity measure leads to different segmentations; this can be compared to the fact that also human interpreters will not delineate the same objects.
- Lack of paradigmatic concepts about the relationship between image objects and landscape objects
- Poor understanding of scale and hierarchical relations among objects derived at different resolutions, e.g. the ratio between segments at coarse resolutions to those at fine resolutions

Opportunities of object-based image analysis:

- Object-oriented methods have been successfully applied to a lot of different problems, thus also to object-based image analysis; methods from biomedical applications and computer-vision can also be applied.
- Ontological foundations are already offered by integrative object-based proposals.
- New knowledge and consensus may be achieved and spread faster using new IT tools, such as wikis for example.
- A continuously growing community using the new technology for different purposes promotes the development and the scientific differentiation of special methods and solutions for specific fields, such as forestry, agriculture, urban mapping and planning, resources exploration, etc.
- New technologies such as symmetric multiprocessing, parallel processing or grid computing can be used for the processing of large datasets.

Threats/Risks to object-based image analysis:

- Object-based image analysis is not by far an operationally established scientific paradigm.

- There is a risk of isolation of object-based image analysis from object-oriented concepts and methods.
- There is a high risk of commercialisation driven by major software providers resulting in a less scientific consensus and definition of object-based image analysis; e.g. different packages, software formats, options, syntax, algorithms, mode of operation.

2.5 Software Issues

For object-based image analysis, the company Definiens AG (formerly Definiens GmbH, and even before Delphi2 Creative Technologies GmbH) in Munich offers substantial software. First, this software was named eCognition, later it was renamed into Definiens Professional. The company was founded in 1994 by Nobel Laureate Gerd Binnig.

The software is used in two fields of applications: geosciences and medical applications. Medical applications include the analysis of cell-based assays, histopathology (i.e. automated digital pathology) and non-invasive imaging. Besides these, also other applications for the software have been found, e.g. the analysis of minerals for the geological dating using fission tracks (Marpu, Gloaguen, Jonckheere 2006).

In geosciences, besides classical uses for remote sensing, it is used for defence and security, infrastructure planning, and natural resources management. Special applications include assisted target recognition for maritime surveillance, as well as automated terrain mapping for unmanned vehicle navigation.

Inside the software, a broader range of sub-packages is offered: Definiens eCognition Server, Definiens Developer, Definiens Analyst, Definiens Architect, Definiens Viewer and Definiens Professional.

For this project, Definiens Professional 5.0 is used. It offers comprehensive functionality with the following highlights (Definiens Website):

- Data and information fusion options for many types of images and vector data
- Multi-resolution image segmentation providing hierarchical networks of image objects
- Outstanding performance for panchromatic, multi-hyperspectral imagery, lidar, infrared and polarimetric SAR data from space-borne and airborne imaging platforms
- Rule-based image analysis to incorporate expert knowledge
- Semantic class hierarchies and relationships
- Fuzzy classification functionality to emulate inherent uncertainties, thereby increasing accuracy assessment
- Standard input and output formats ensuring seamless workflows when extracting information from remotely sensed data, or even GIS-ready results

3 Image segmentation

The concept of object-based image analysis needs a technique to extract image objects from an image. These image objects are attained by image segmentation. In this chapter, theoretical concepts are presented, as well as algorithms that are available in the software.

3.1 Theoretical Approaches

Image segmentation is a term that has been used with various meanings and contexts by different authors. Although not using the term ‘segmentation’, Haralick, Shanmugam and Dinstein (1973) use the concept for the definition of “a set of meaningful features to describe the pictorial information from a block of resolution cells”, with the objective to create areas with similar texture. Pavlidis (1977) uses it for the subdivision of a picture into regions which have a certain uniformity. In order to keep an analysis as general as possible it is assumed that there is a given uniformity predicate. He gives an idea of what a formal description of segmentation may look like (Pavlidis 1977).

Pinz (1994) describes image segmentation as a process to divide an image into non-overlapping, expedient and homogeneous parts, whose unification results in the former image. Criteria for the homogeneity of resulting image objects can be colour, gray value, texture etc. Image objects are called ’image events’, the associated data structure is called ‘token’. A ‘tokenset’ represents all tokens produced by a certain segmentation algorithm and is equivalent to a data file. Each token has a couple of features that give a description of its attributes, including its location. Baatz and Schäpe have published a well-known article about multiresolution segmentation (2000). They also define the results of segmentation as meaningful homogeneous areas, whose scale is to be significantly larger than the scale of image noise or the texture.

3.2 Subtypes of Segmentation

Generally, most segmentation approaches can be divided into two groups: area-based algorithms and edge-based algorithms (Pinz 1994, Blaschke and Strobl 2001). In the following, some of the established methods shall be explained in brief.

3.2.1 Area-based Segmentation

3.2.1.1 Histogram Thresholding:

The setting of global thresholds is probably the easiest form of area-based segmentation. If there is a strong contrast between fore- and background, results are said to be very promising. When working with a number of objects with several gray values, several thresholds have to be set. So, this concept based on crisp sets soon reaches its limits, when handling a lot of different values (Pinz 1994).

3.2.1.2 Region Growing

This algorithm is used very often, even for pixel-based image analysis. Its algorithm can be described as follows: First, so-called ‘seed cells’ are spread over the image. In a bottom-up approach, they can be chosen at random, while in a top-down approach assumptions about the expected image content are used for selection. Continuing, in each parallel step all seed cells’ neighbours are regarded, and attached, as long as they fit into the homogeneity criterion and do not belong to a different region. As soon as two regions contact each other, the constraints of homogeneity are checked. If they are similar, they amalgamate. The last steps are repeated until no changes are made in the structure. If there are still pixels that have not been assigned to a class, they are either marked as background, or coherent pixels constitute a new, additional region, or new seed cells have to be set and the second step is repeated (Pinz 1994).

3.2.1.3 Split and Merge:

While Region Growing is using a ‘merge’ strategy (single pixels are added to a region), a `split’ strategy starts with a rough segmentation of the image and improves this segmentation by dividing the objects into more homogeneous parts. Thus ‘Split and Merge’ is a combination of both strategies, with a couple of possible algorithms that can be used. Horowitz and Pavlidis (in Pinz 1994) have proposed an algorithm using a quadtree structure for image segmentation, with a merge of the square objects.

3.2.1.4 Blobs and Scale Space

Compact objects are also called blobs, in contrast to long, linear objects. When viewing the blob at different scales, one can say that the surviving of the blob in several scales is a measure for its significance. Lindeberg and Eklundh (in Pinz 1994) described such ‘scale space blobs’. With decreasing or increasing scale, blobs can disappear, fuse with other blobs; respectively they can be divided into several blobs or can be created completely new.

3.2.2 Edge-based Segmentation

An edge can be defined as the frontier between two homogeneous areas in an image. In contrast to area-based segmentation, image objects are to be segmented by detecting their edges. In a lot of algorithms, a bottom-up approach is basic for the detection of edges and lines.

At first, images have to be smoothed, as they suffer from noise; otherwise a plethora of artefacts would be constructed. After the smoothing, the concrete algorithm to detect the lines and edges starts. Small remaining gaps are to be closed, as the results are not perfect yet. Also, short edges are removed, as they represent artefacts as well. Edges that have survived the last steps have to be grouped, continued und linked to each other, so that complete networks can be obtained. These networks can now be assigned to image object borders.

3.2.2.1 Representation of Edges by Chain Code

Edges, lines and curves can be represented by a couple of chain codes, e.g. the Freeman or RULI chain code. The Freeman chain code starts with an edge point, and codes the next pixel by its direction from the staring pixel. Either four (North, East, South, West) or eight directions (N, NE, E, SE, S, SW, W, NW) are used. Its result is a chain of compass directions.

The RULI chain code looks at a curve entering a pixel. It can leave the pixel either to the right (R) or to the left (L), go straight through it, i.e. intersect it (I), or make a U-turn (U). Its result is a chain of codes for each curve. For the exact localisation of the curves, a curve pyramid is used, representing the curves in different spatial resolutions.

3.2.2.2 Gradients and Zero Crossings

An edge is representing a place, where two regions with different grey values are neighbours, or, from a mathematical perspective, they represent a gradient between the two different grey values. Zero-crossings are turning points of the second differential derivation. Pinz (1994) describes Sobel and DoG (Difference of Gaussians) operators.

3.3 Segmentation Algorithms in Definiens Professional / eCognition

Definiens Professional offers a couple of segmentation algorithms: multiresolution segmentation, spectral difference segmentation, chessboard segmentation and quad tree based segmentation; probably the most frequently used is the multiresolution segmentation algorithm. From a theoretical background, it is a fractal-based approach.

Baatz and Schäpe (2000) have set up the following objectives for (multiresolution) segmentation:

- High Quality Image Object Primitives: Primitives should be a universal high-quality solution that is applicable and adaptable to many problems and also textured image data of arbitrary type
- Multiresolution: Objects of interest typically appear simultaneously on different scales in an image. The extraction of meaningful image objects should consider the scale of the problem to be solved. Therefore the scale of resulting image objects should be unrestrictedly adaptable to fit the scale of task.
- Similar Resolution: Almost all attributes of image objects – colour, texture or shape – depend on the scale to a fewer or larger extent. Only structures of similar size or scale are of comparable quality. Therefore, the size of all resulting image objects should be of comparable scale.
- Reproducibility: Segmentation results should be reproducible.
- Universality: Results should be applicable to arbitrary types of data and problems.
- Speed: Reasonable performance should be reached even on large image data sets.

As criteria for the evaluation of the segmentation results, Baatz and Schäpe (2000) defined quantitative and qualitative criteria. Quantitative criteria can be - based on a certain definition of heterogeneity for image objects and a certain average size of image objects: Average heterogeneity of image objects should be minimized; each pixel is weighted with the heterogeneity of the image objects to which it belongs.

Qualitative criteria, considering the human eye as an evaluation resource, can be the consistent handling of local contrasts, as well as the segmentation of image regions of a similar dimension. Another qualitative criterion can be the information that can be extracted from image objects for further successful processing.

3.3.1 Multiresolution Segmentation

This is the most important algorithm, based on a heuristic optimization procedure that locally minimizes the average heterogeneity of image objects for a given resolution. It can be applied on a pixel level or on an image object domain. If a pixel level is selected, a new image object level is created above the pixel level moving all existing levels one level up.

Baatz and Schäpe (2000) describe it as a region merging technique. At the beginning, each pixel forms one image object or region. At each step, a couple of image objects are merged into one larger object, with the merging decision based on local homogeneity criteria, describing the similarity to adjacent image objects. The homogeneity criterion is not solely a ‘fit’ or ‘fit not’ criterion; a merging cost is assigned to each possible merge. These merging costs represent the degree of fitting. A possible merge is executed, if its degree of fitting is smaller than a least degree of fitting. The whole procedure stops when there is no possible merge left. The smaller the least degree of fitting, the fewer the number of merges permitted. Thus the size of resulting image objects will grow with the least degree of fitting value. Due to this property, this parameter is referred to as ‘Scale Parameter’. A merge with a smaller degree of fitting (i.e. a smaller value) than the scale parameter is said to fulfil the homogeneity criterion.

The two main components of multiresolution segmentation are decision heuristics and the definition of a homogeneity criterion to compute the degree of fitting for a pair of image objects (Baatz and Schäpe 2000).

Decision heuristics can be described as the method to determine the image objects that will merge at each step. Beginning at a random object A, an adjacent object B, that will be merged, can be found by a couple of different heuristics. Baatz and Schäpe (2000) distinguish between four constraints in terms of choice: fitting, best fitting, local mutual best fitting and global mutual best fitting.

The degree of fitting, the second main component of multiresolution segmentation is described by Baatz and Schäpe (2000) as the similarity of objects that are near to each other in a d-dimensional feature space. It can be defined as the Euclidian distance of single pixel values; optionally, the value can be standardized by the standard deviation over all segments of the feature in each dimension. After the merging of segments, their heterogeneity apparently increases. Although the heterogeneity is to be minimized, a rule needs to be found in order to adapt the degree of fitting.

Also, the form heterogeneity has to be maintained. Two possible definitions are offered. One is the deviation from an ideal compact form given by the relation between factual edge length l and the square root of the object size n (in pixels), or in other words, the edge length of square of n pixels. The second definition of form heterogeneity is the deviation from the shortest possible edge length given by the bounding box b of the segment. It is the relation between the factual length l and the edge length of the bounding box. In a raster the edge length of the bounding box is also the shortest possible edge length for an arbitrary segment (Baatz and Schäpe 2000).

Figure 8 shows an example for a multiresolution segmentation with different scale parameters. Segments (image objects) are delineated with white lines.

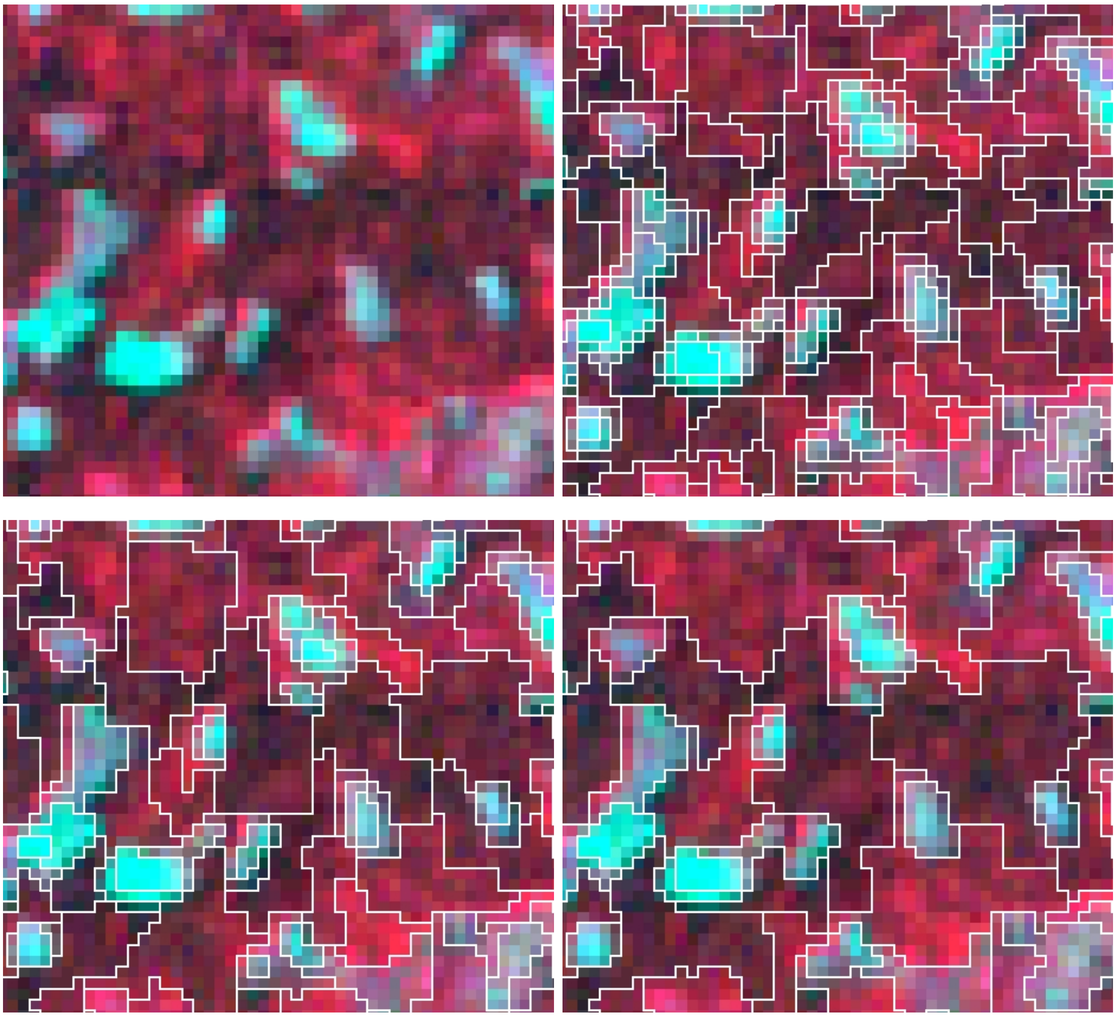

Fig. 8: Landsat image segmented with multiresolution segmentation with scale parameters 3, 5 and 8

3.3.2 *Spectral Difference Segmentation*

This algorithm is designed to refine existing segmentation results by merging neighbouring image objects that have similar spectral values (Definiens Professional 5 User Guide 2006). Neighbouring objects are merged, if their mean layer intensity values are below the maximum spectral difference. It cannot be used to create new image object levels based on the pixel level domain.

Besides the level name, the user has to enter:

- the maximum spectral difference defines the amount of spectral difference in the new segmentation for the image objects. If the difference is below this value, neighbouring objects are merged.

- the image layer weights offer, as in the other segmentation algorithms, the possibility to assess their importance for the segmentation, with values between 0 and 1. The Definiens Professional 5 Reference Book recommends assessing the spatially coarser thermal layer 6 of Landsat7 images the value 0 to avoid deterioration of the segmentation result due to the blurred transient between image objects in this layer.
- thematic layers: also here a thematic layer can be added, which will lead to additional splitting of the objects.

This algorithm can be used additionally to refine the results of the multiresolution segmentation, as it helps merging objects with similar values; thus, large homogeneous areas can be created regarding spectral difference. In an example given in figure 9, forest areas are merged, as their spectral difference is quite low.

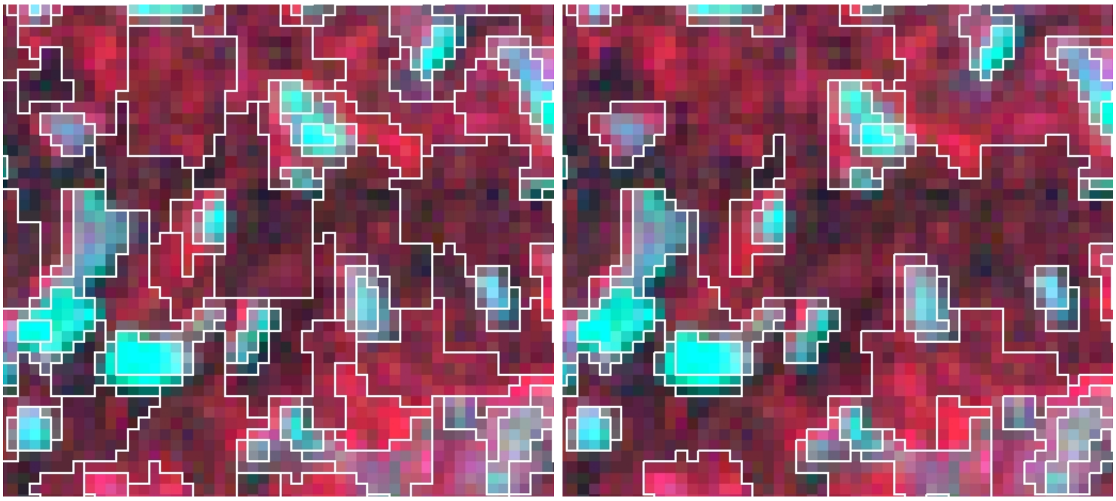

Fig. 9: Satellite image with multiresolution segmentation (scale parameter 5), segments merged by spectral difference segmentation 5

3.3.3 *Chessboard Segmentation*

This segmentation algorithm splits the pixel domain or an image object domain into square image objects. The most important segmentation parameter is the object size, i.e. the pixel size or length of edge of the chessboard. For creating objects of the size of one pixel, the object size is one, for objects with an edge length of five pixels (square of 25 pixels) the object size is five, and so on.

In the project, the chessboard segmentation does not really seem to be helpful and applicable as it does not differentiate between different colour tones or different shapes. So it has remained out of consideration for the further proceeding. Figure 10 shows a chessboard segmentation.

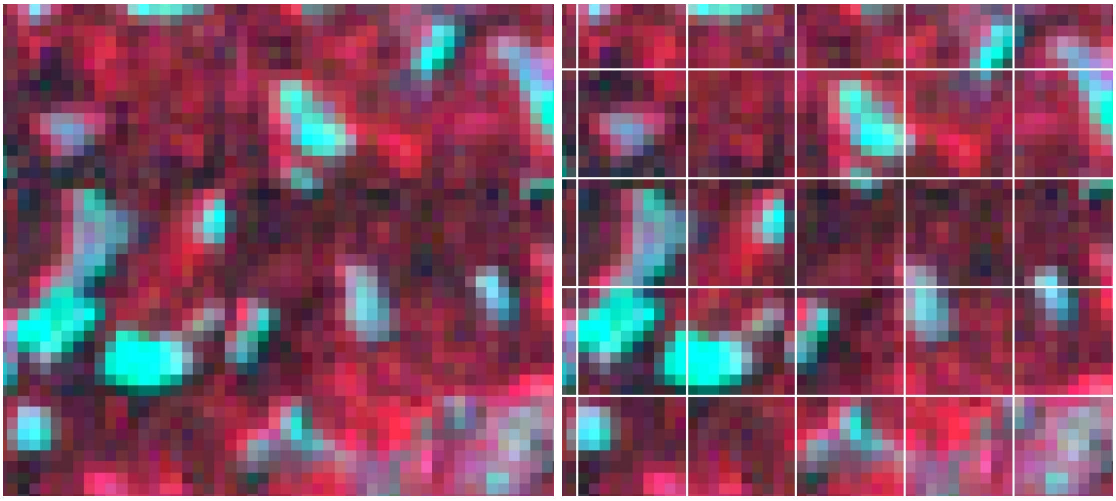

Fig. 10: Satellite image segmented with chessboard segmentation (chessboard size 10)

3.3.4 Quad tree based Segmentation

This segmentation splits a pixel image into square image objects of maximal size that are then split into four equal square image objects in a recursive procedure until a level of homogeneity inside a square is reached. The quad tree (or quaternary tree), used for two-dimensional data, is besides the binary tree and the octree (for one- respectively three-dimensional data) an example for a tree structure; it serves the purpose to describe elementary spatial data structures Quad trees are used for a couple of purposes, e.g. there are point quad trees, region quad trees or point-region (PR) quad trees (Kainz 2007).

Its segmentation parameters in the Definiens software are the mode, which can be colour (the maximal colour difference within each square image object is less than the scale value) or super object form (each square image object must completely fit into the superobject, so there must be an additional upper image level), and the scale that defines the maximum colour difference of the segments. The smaller the scale value, the finer the square segments are. It can only be used in conjunction with the colour mode. An additional function offers the possibility to weight the different colour bands of the images as well as the thematic layers included in the project. However, the image layer weight is only binary, offering a yes/no-weighting instead of a numerical weighting between 0 and 1.

A comparison between segmentations with differently weighted layers and equally weighted layers showed that the segmentation with equally weighted layers had smaller segments than the segmentation with differently weighted layers.

In the project, also the quad tree segmentation (though offering a more refined instrument for segmentation than the chessboard segmentation) was not applied because of the square image objects, which were not useful for the coca fields that mostly have an irregular shape. Figure 11 shows a quad tree based segmentation.

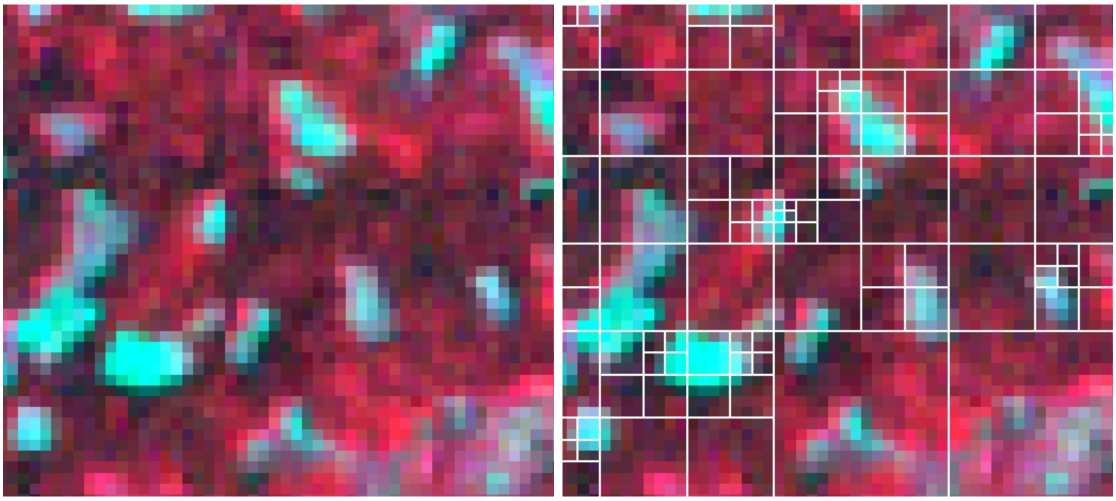

Fig. 11: Satellite image segmented with quad tree based segmentation (scale 70)

4 Knowledge-based Image Analysis

In an example given by Richards (1993), a skilled photo interpreter analysing spatial data will not analyse single pixels, but will focus the concentration on regions. In a combined use of Landsat imagery, radar data and a soil map, the interpreter can apply his expert knowledge by subsequent decisions, coming to the final conclusion that a certain spot on Landsat image has to be a certain type of land cover. Richards deduces that the key to the interpreter's success lies in his or her knowledge and asks if that knowledge can be emulated by an automated process. The solution is suggested to be a knowledge-based image analysis system. This chapter will be engaged with methods and systems to transform human expert knowledge into technical systems. The technique of fuzzy logic and its application on image classification shall be explained. Another sub-chapter informs about image interpretation keys in general.

4.1 The principle of knowledge-based Image Analysis

Although there is more than one possibility to model expert knowledge, the most common representation of knowledge is to apply rules. Richards (1993) puts it like the following: if there is a certain condition, then an inference has to occur. The condition can be either true or false; and it can be a simple logical expression or a compound logical statement, with several components are linked through the logical 'or' and 'and' operations. Kainz (2006) differentiates between deduction (*modus ponens*) and induction (*modus tollens*) for the implementation of rules. In the *modus ponens*, a premise determines a conclusion. So if the premise is true, the conclusion is also true. In the *modus tollens*, a premise determines a conclusion as well. If the conclusion is not true, then the premise is not true as well.

In an analysis system, knowledge bases can contain a plethora of rules of different types, which are created by experts in particular fields of work. Some of theses rules can be strong, while others offer only weak support. In the case of several candidate classes finding support among the rules, an inference mechanism has to lead to a decision. In the case of multiple data sources or sensors, it is recommended to decompose the problem into a set of individual analyses with an adjacent combination of results with the help of a separate expert system which is able to perform a joint analysis of sources (Richards 1993).

In order to have an efficient analysis system, Benz et al. (2003) have set up some requirements. These requirements include the sensor characteristics, as well as appropriate analysis scales and their combination; also the identification of typical context hierarchical dependencies are required, and the consideration of inherent uncertainties of the entire information extraction system, beginning with the sensor and going up to the fuzzy concepts for the requested information.

Although Richards (1993) does not use the term fuzzy logic yet when describing knowledge-based image analysis, it is obvious that he refers to it indirectly. The concept of fuzzy logic shall be explained in the following.

4.2 Fuzzy Classification

To understand the concept of fuzzy classification, knowledge about fuzzy logic is needed. In general, fuzzy logic defines fuzzy sets in contrast to classical (crisp) sets. A crisp set is defined by a sharp delineation at the ‘border’ of the set. An element either belongs to a set, or it does not. In a fuzzy set, however, the delineation of the set is not sharp, but fuzzy, i.e. blurred. The delineation is performed by a membership function, with values ranging between 0 (no membership of the set) and 1 (definite membership of the set). A membership value of 0.5 (crossover point) represents the boundary of the comparable crisp set. As a mathematical formula, a membership function is defined as

$\mu_A : U \rightarrow [0,1]$ with A standing for the fuzzy set of the universe U.

An element can be completely contained in the set, it can be not contained in the set, or it can be contained just a bit. There are two types with a number of subtypes of membership functions known: Linear and sinusoidal functions. Operations such as unions, intersections or complements on fuzzy sets can be compared to operations on crisp sets.

In a classical example to compare crisp and fuzzy sets (given by Kainz 2006), persons can be assigned to body height classes, such as short, average, and tall. Three persons A, B and C have a body height of 185 cm, 165 cm and 186 cm. In a classical crisp set, class boundaries can be set to (-, 165] for short, (165, 185] for average and (185,-) for tall people. So, person A belongs to the average class, person B to the short class and person C to the tall class. Although A is almost as tall as C, they fall into different classes, when classified in the crisp set.

In the fuzzy set approach, membership functions need to be defined for the three classes. For the short class, a linear membership function can be chosen, ranging from one (shorter than 150 cm) to zero (taller than 180 cm). For the average class, the membership function goes up from zero at 150 cm up to one at 175 cm, and then decreases to zero again at a value of 180 cm. The tall class is defined by a membership function starting with value zero at 170 cm, increasing linearly up to value one at 200 cm.

Due to the fuzzy set approach, it is much easier to show that A and C have about the same height and belong to a higher degree of membership to the average class than to the tall class.

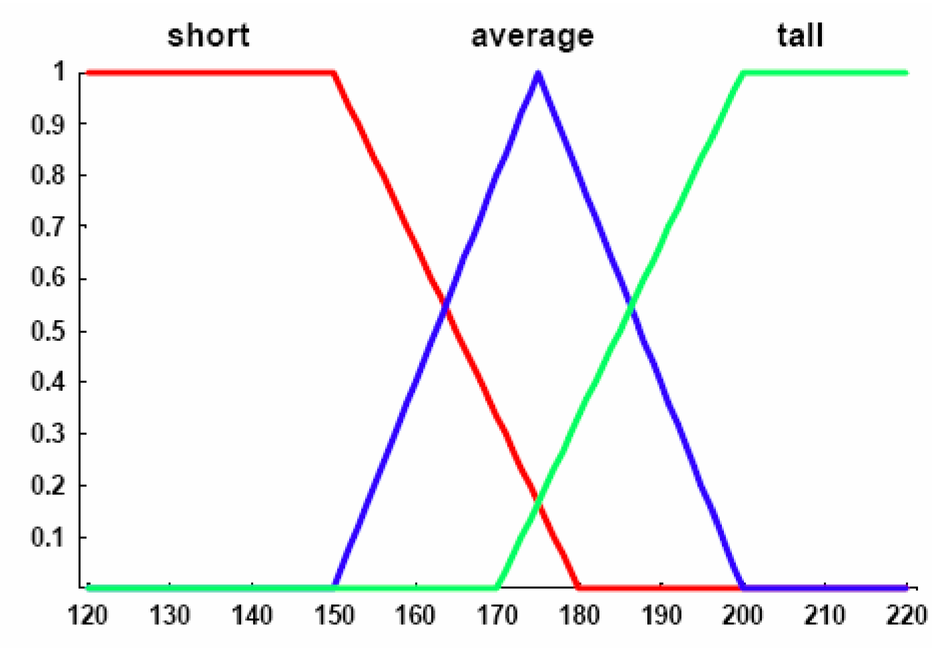

Fig. 12: Membership functions for short, average and tall (Kainz 2006)

With the help of fuzzy logic, linguistic variables (e.g. height) and their values (e.g. short, average or tall) can be translated into mathematical terms, as well as modifiers such as ‘very’, ‘somewhat’, slightly’ etc. These modifiers are also called hedges; they can be expressed by operators such as normalization, concentration, dilation, negation and contrast intensification. Kainz (2006) gives examples of how these hedges can be transferred into membership functions.

The process of turning a crisp system into a fuzzy system is called fuzzification, its reverse process defuzzification. Defuzzification is important for the production of maps of land cover, or land use applications (Benz et al. 2003). Therefore, fuzzy classification results have to be translated back into crisp values, by either assigning an object to a certain class or not. If an object is shared by several fuzzy sets, the class with the highest membership value is chosen. If the membership is below a certain value, e.g. 0.5, it will not be assigned to a crisp set class, in order to ensure a minimum reliability.

In the field of geoinformation, fuzzy logic can be applied in spatial analysis, fuzzy reasoning, and the representation of fuzzy boundaries, as many spatial phenomena have an inherent fuzziness. Also in remote sensing, fuzzy logic can be applied (Lillesand, Kiefer and Chipman 2004). One approach is fuzzy clustering that is described to be similar to a K-means unsupervised classification, however with fuzzy boundaries instead of hard boundaries; therefore, the membership value describes the distance of a pixel to the means of all classes. A second approach is a fuzzy supervised classification, similar to a maximum likelihood classification. But in contrast to the maximum likelihood classification, training areas are not purely homogeneous, but can be a combination of pure and mixed sites. In an example by Lillesand, Kiefer and Chipman (2004), a vegetation classification can include a pixel with grades of 0.68 for the class forest, 0.29 for the class street, and 0.03 for the class grass. Nevertheless, all grades have to summarize at the end to the value 1.

In object-based image analysis, fuzzy logic can be applied as it standardizes features and allows the combination of features, even if they have a different range and dimension; it also provides a transparent and adaptable feature description, especially if compared to neural networks; and it allows the formulations of complex feature descriptions with the help of logical operators and hierarchical class descriptions (User Guide eCognition 2003). Benz et al. (2003) estimate the fuzzy classification systems as very powerful soft classifiers, that are able to incorporate inaccurate sensor measurements, vague class descriptions and imprecise modelling.

The concept of fuzzy logic can be an appropriate tool for the differentiation between similar or blending classes. To define classes when analyzing images, an interpretation key can be a good instrument.

4.3 Image Interpretation Keys

Image interpretation keys are very powerful tools to model human knowledge. The visual interpretation of images or photos has a very long tradition; this is of course also the case with aerial photography and satellite images. Although computational image processing and analysis have spread out almost everywhere, human interpreter knowledge is still crucial for the interpretation and assessment of remotely sensed images.

Therefore, the knowledge, experience, training, as well as visual analytical abilities have a high influence on how an image is interpreted, and on how the objects are detected and assigned to a land cover class. The image interpretation process thus involves a lot of subjective judgement. The interpreter bases his or her decisions on features like tone, colour, shape, texture, and additionally to that, context information and knowledge about the area being interpreted. This information is however subjective and thus different interpreters in one project do not have the same experience and the same training programme, which makes it harder to compare or combine the results of two or more interpreters.

To minimize subjectivity of the visual interpretation process, measures like a calibration of the interpreters can be employed, with the aim of a comparable knowledge level of all interpreters. Another measurement could be the automation of the interpretation process. However, even in an automated image analysis the results are influenced by the choice of algorithm and the selection of training areas. Anyhow, the results will be homogeneous and duplicable.

The crucial tool for bringing interpreters to a comparable level of knowledge, experience and notion of the topic is the interpretation key. This can be described as a guide for the object categories to be interpreted and their characteristic features on the image. It standardizes and documents the knowledge and experience of the interpreter involved in the project, and it makes interpreters organize their findings in a logical system. Especially in large projects, such a guideline is of crucial importance for coordinating the work of several interpreters to obtain homogeneous results. It serves as a reference tool for novice interpreters and as decision finding tool for experienced interpreters in case of doubts. Moreover it can be useful to explain the complex work to persons not being familiar with image interpretation. Additionally, it should serve as the basis for a knowledge based, automatic classification (Bauer and Schneider 2006;2).

Due to Mintzer (1975), several types of interpretation keys can be distinguished. The choice of the type depends on the scope, the technical level of the interpreter, and the intrinsic character of the interpretation. Normally, an interpretation key should include pictorial as well as descriptive documentation and reference data, for example photographs of the area. The documentation has to provide a general description of the land cover or land use types to be mapped.

Generally speaking, there are two ways of how an interpretation key can be organised:

- A selective or example key is a series of examples of image or image clippings of the different object (terrain) classes to be interpreted.
- A decision tree or elimination key assists the interpreter to proceed step-by-step from the general to the specific.

The selective or example key contains numerous typical example pictures and textual descriptions of objects in a given category. They are organised for comparative use, where the interpreter selects the example that nearly coincides in its image feature with the image to interpret. The interpreter has to select the example that resembles most closely the feature or condition found on the image under consideration.

A critical aspect concerns the number of representative examples to include in the key as it is difficult to describe all forms of appearance. If the interpreter cannot find an example that matches the feature on the image, a wrong interpretation might be the result (Mintzer 1975). The selective key is useful where the categories to be identified are highly variable in thematic and spatial composition and, as a consequence, in their appearance on the images.

Subtypes of the selective key are the following:

- Essay key: objects are described in textual form using images for illustrations only.
- File key: composed of one or more selected images, with notes concerning their interpretation. This type of key is generally assembled for use by an individual interpreter.
- Photo index key: composed of one or more selected images, together with notes concerning their interpretation, assembled for rapid reproduction and distribution to other interpreters.

- Integrated-selective key: images and recognition features for any individual object, within a subject or regional key, are associated in such a manner that by reference to the appropriate part of the key the object can be identified.

The decision or elimination key guides the interpreter in a step-by-step process from the general to the specific through the elimination of all categories except the correct one. They often have the form of a dichotomous key which offers the choice between two alternatives; based on the features used for the description of an object, the interpreter has to decide if a statement is true or false and then follows the scheme to the next decision.

A decision tree bears the risk that just one single incorrect decision (e.g. if the interpreter is forced to make an uncertain choice between two (or more) unfamiliar image characteristics) will lead to an incorrect branch of the tree and thus finally to an incorrect identification. That is particularly a problem if the wrong decision occurs at an early stage in the sequence of queries. In comparison, an example or selection key provides more redundant information descriptions, with a reduction of the risk of an incorrect assignment.

However, it is stated by Bauer and Schneider, as experience shows, that a wrong decision occurs usually at the end of the process at an advanced stage, when differentiating between two very similar characteristics. If designed properly, a decision tree is said to give more reproducible results than a selection key. The decision tree resembles a fixed cooking recipe and is thus easier to execute.

A couple of different features are used for the image interpretation, according to the Manual of Photographic Interpretation (1997), there are ten different basic elements: tone and colour, size, shape, height, shadow, texture, pattern, site, association and time. Depending on the type of images (aerial photographs or satellite images), different features can be used for the interpretation. The lower the spatial resolution, the more important are colour or context, the higher the resolution the more shape or shadow can be detected. The following features can be taken into consideration; their importance for the project shall be assessed as well.

- Tone and colour: Tone describes the relative brightness of image objects; on black and white photographs, the tone varies only from black to white, with various shades of gray in between. On colour images, hue, saturation and brightness make up the colour and thus contribute to the interpreter's ability to differentiate between objects. The ASPRS manual (1997) sees photographs as a collection of discrete elements or as an imperfect representation of a continuum. Tone and colour, especially brightness are relative and can vary within a scene. They also vary with the season of the year, and the position of the sun in relation to the camera position. These features therefore are not very robust: they can be used as absolute features (as constant characteristics) for identifying objects to a very limited degree only. Haralick, Shanmugam and Dinstein (1973) point out a strong, inextricable relationship between the concepts of tone and texture. For the project, this is one of the most basic and important features.
- Size: The size of an object is one of the most important as well as most robust features of the image. In general, the scale has to be known for exploiting the size or height

feature; in other words, measuring information has to be available in image information. Another possibility is the definition of relative sizes among objects in the image. The size of an unknown object can be set in relation to the size of a known object. For the project, this feature is definitely important.

- Shape: Shape refers to the general form of depicted objects as depicted in the image. As a significant and robust feature it may serve to identify some objects without taking other characteristics into account. Besides the identity, the shape can also give a clue for its significance and function. The plan view of an object can be so different from its familiar side view that its shape is surprisingly difficult to interpret; examples are highway interchanges or buildings, like the Pentagon. For the mapping of coca, the shape has been assessed as a feature, but it has turned out to be problematic to define it.

- Height: The height information can be detected only in stereoscopic aerial images, for the interpretation of vegetation in satellite images, this feature cannot be taken into consideration for satellite image analysis. For the project, this feature has not made sense so far.

- Shadow: This feature, mainly usable for very high resolution images like aerial photographs for direct interpretation, can also have an influence on tone and colour on lower resolution satellite images, if different image acquisition directions are involved. In aerial images, it provides additional and partly redundant information. On Landsat images, the shadow normally has no importance for the detection of objects; however, cloud shadows may impair the analysis of the image, additionally to the loss of information that is covered by the clouds. This feature plays a certain role for the interpretation, as the shadows impaired the detection of fields.

- Texture: The texture describes the fine-structured nearly periodic spatial tonal or colour variations inside an image. Texture is produced by the spatial arrangement of elements making up the objects which are too small to be discernible. Depending on the scale, illumination and view direction in conjunction with shadow effects may be of decisive influence on the texture. Texture usually is a robust feature for lower resolution images, and can be unreliable for high resolution images. It is difficult to describe a texture verbally. It involves aspects of grain size, grain shape, and their influence on texture, tone and colour contrasts, homogeneity etc. Haralick, Shanmugam and Dinstein (1973) see a strong relationship between tone and texture: whilst tone is based on the varying shades of gray of resolution cells in an image, texture is the spatial, statistical distribution of gray tones. Although this feature has been implied into the interpretation key, it has not been determined clearly.

- Pattern: Pattern defines the spatial arrangement of objects; compared to the feature texture, it is coarser and not periodic. Patterns can be natural or cultural, or result from their interaction. Man-made patterns normally tend to be more rectilinear or otherwise regular, they can persist a long time after human activities have ceased. This feature has not been taken into account for the detection of coca fields.

- Context: This feature includes aspects of site, association and time. Individual objects on the ground usually cannot be identified when viewed isolated. Their identification is possible only when they are seen in association with each other. This involves a "convergence of evidence" in the progression of the interpretation and represents an important skill to be developed by the interpreters. Interpretation based on context involves a reasoning process that uses all the principles of interpretation to relate an object to its surroundings. Expert knowledge in the special field of the interpretation topic usually is required. For the detection of coca fields, this feature has been rated as extremely important.
- Site: Site refers to the relationship of objects and phenomena in images to their geographic locations or terrain conditions. By knowing the geographic location of objects, the variety of possible interpretations can be greatly reduced, and a lot of detailed information, not necessarily visible in the image, can be concluded. Aspect, topography, geology, soil, vegetation and the varied imprints of man's culture are distinctive factors when examining a site, e.g. some crops grow only up to distinctive height or slope. Although this feature has not been taken into account in the interpretation key, it is an important feature for the detection of fields.
- Association: Association means the spatial relationship of objects and phenomena; it is an abstraction of the proximity and connectedness of objects in the environment; the identification of one object would indicate the likely presence of other commonly associated objects. It can be a useful clue in identifying cultural features. Therefore this feature may have a certain importance, e.g. coca fields are situated near rivers or settlements.
- Time: Time is an element that refers to the temporal relationship of objects and phenomena, e.g. a crop calendar. Certain objects or phenomena can be detected more easily at particular times, requiring multi-temporal or multi-seasonal images. For the interpretation of coca, time is not a useful feature, as there is no well defined crop calendar for Colombia.

In general, man-made features are easier to detect on an image than natural features (Blaschke and Strobl 2000; Lillesand, Kiefer and Chipman 2004). For a reliable interpretation of these, training and field experience are often essential to ensure consistent results. However, as Lillesand, Kiefer and Chipman (2004) state, a number of keys have been successfully employed for the identification of agricultural crops or species of trees. But such keys are normally developed and used on a region-by-region basis in that the appearance of vegetation can vary widely with location and season.

Bauer and Schneider (2006;2) come to the conclusion that an interpretation key serves as documentation of the entire interpretation process. It is a difficult task to transform the knowledge of human interpreters into a logical system. The key should be developed in conjunction with all persons involved in the interpretation process and should be based on a broad agreement of all interpreters.

The development of an interpretation key is an important task in order to conduct a proper survey. It is a prerequisite for a reproducible interpretation of all kind of remote sensing imagery. A major benefit of the development of an interpretation key is that the interpreters engaged in the development of the key deepen their insight into the interpretation process and become aware of the visual and mental processes and the expert knowledge taking effect in the interpretation. They thereby improve their interpretation abilities.

An interpretation key has to be adapted regularly. Its development is a dynamic process, with a regular verification and adjustment necessary. Furthermore, an interpretation key is only valid for a specific geographic area, and for a specific season. Temporal aspects may be included in an interpretation key. In case of large projects, as for the monitoring of illicit crops within a country, parts of the key can be transferred to other regions.

To enhance the work flow of image interpretation and to make it more objective, an interpretation key can be useful.

5 Case Study – Implementation of a Decision Tree

This chapter represents the core of the thesis, the practical application of object-based image analysis to detect coca fields in the Meta-Guaviare region. The area under investigation shall be introduced to the reader, as well as the underlying data sources, and the preparation of data. Then, referring to chapter 3, the apt segmentation adjustments for the process shall be explained. In a very important section of this chapter (5.5), the implementation process of the decision tree will be explained, first on a small subset, then on the entire Landsat image.

5.1 Area under Investigation

The area under investigation is the area of a Landsat image that is located in the Colombian departments (*departamentos*) of Meta and Guaviare (Figure 14). These two departments are situated in the landlocked center of Colombia, in the East of the Cordillera Oriental. Neighbouring departments are Vichada and Guainia in the east, Vaupes in the southeast, Caquetá in the southwest, Huila in the west, and in the north there are the federal district of the capital Bogotá and the departments of Cundinamarca and Casanaré.

The department of Meta has about 770.000 inhabitants in 2005 (DANE 2005), and a size of 85.216 km², about as much as the republic of Austria. Its capital Villavicencio (about 385.000 inhabitants, DANE 2005) lies in the Northwest, next to the border of the Cundinamarca department with the national capital Bogotá. Historically, it served as a station for cattle on its way to Bogotá. Extensive, partly semi- intensive ranching represents the agricultural basis of the department (Pombo 1990). In total there are 29 denominated municipalities; larger ones are (besides the capital) Acacias, Granada, Puerto López, Puerto Gaitán, Cumaral and San Martin (DANE 2005).

The department consists of three physiogeographic regions: in the west, the edge of the Cordillera Oriental, with heights more than 4000 meters, the Serrania de la Macarena and the plain in between. The second zone is a plain in the center of the department which is transected by rivers into smaller hills. The third zone, the biggest one, consists of hot and humid lowlands, with the lowest height of about 125 meters. The mean temperature lies around 27° C. There are a couple of national parks such as Páramo de Sumapaz, Cordillera de los Picachos, Tinigua and the Serrania de la Macarena NP.

The department is not one of the poorest departments though in 2000 about 57 percent of people living below the poverty line.

The department of Guaviare is smaller than Meta, not only in terms of size (55.391 km²), but also in terms of population (about 81.000 inhabitants in 2005), resulting in a much lower population density. There are just four official municipalities, of which the capital, San José de Guaviare, represents the largest settlement with about 45.000 inhabitants (DANE 2005). It is situated right at the Guaviare river, the northern border to the department of Meta.

The department is characterized by plain lowland basins, with rivers rich in minerals ("*rio blanco*"), coming from the Cordillera Oriental and heading towards the Orinoco in the north, and in the south the "*rios negros*", heading towards the Amazonas, As they have their origins in the primeval forests these rivers are poor in minerals, resulting in a dark tone of the water. The climate is in general very humid and hot with day temperatures between 25 and 30° C.

With about 80 percent of people living below the poverty line in 2000, Guaviare is one of the poorest departments. In this department, still a lot of indigenous tribes can be found, such as the Guahibo and Guayabero, which belong to different linguistic groups. In the late nineteenth century, colonists arrived in order to exploit caoutchouc (rubber) and later to grow coca, the most important crop of the department at the moment. The agriculture is characterised by shifting cultivation (Pombo 1990).

Meta and Guaviare represent one of the most important and notorious producing regions for coca. Its level of coca cultivation is the highest of Colombia, with 25970 hectares in 2005 (Meta: 17305 ha; Guaviare: 8658ha), nine percent less than in 2004 (28509 hectares totally; Meta: 18740 ha; Guaviare: 9769 ha). Nonetheless, the level was even higher in 1999, when 39819 hectares (Meta: 11834 ha; Guaviare: 28435) were cultivated. Guaviare was the department where coca cultivation first appeared in Colombia in the 1970ies. Since then, coca cultivation has remained important for the department (UNODC 2006). Cultivation is mainly concentrated along the southern bank of the Guejar river (going from the Serrania de la Macarena mountains and National Park in the west to the Ariari river in the east), as well as along the other rivers in that region. Especially around Guaviare's capital San José de Guaviare there is a high concentration of coca plants (figure 13). The town is described even in tourist literature as a coca capital (Braune and Semper 2006).

Generally, the region, especially the department of Guaviare is described as a place difficult to access, it is hardly developed by tourist infrastructure, because of the low population, and the dominance of the narcoguerrilla. Tourism is described as very risky, especially in Guaviare (Braune and Semper 2006, García Marín 2000).

However, Meta's national park Sierra La Macarena, representing the oldest national park in Colombia founded in 1948, is promoted as a major tourist attraction. The mountain range, about 120 km long and 30 km wide, is famous for its wide variety of species, a lot of them endemic. There are more than 500 bird species, pre-Columbian rock paintings and the Caño Cristales river, described as one of the most colourful and beautiful rivers in the world. However, this national park is also known as the protected area that is most affected by illicit crops that were covering 3,354 ha in 2005.

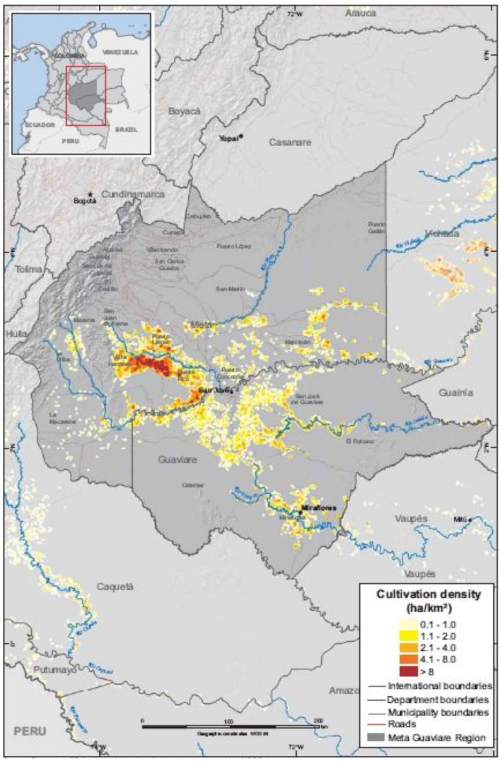

Fig. 13: Coca Cultivation in Meta and Guaviare (UNODC 2005)

Both departments suffer from displacement of people due to violence, armed conflicts, drug trafficking and the search for better living conditions. There have been about 52,000 people displaced in Meta and about a fourth of that number in Guaviare. Though these figures are relatively low in comparison to numbers in northern Colombia, especially in Antioquia. But, in relation to population size, these values rank quite high for the year 2004, especially in the sparsely populated Guaviare.

As the departments belong to the main coca cultivation areas, there is a high presence of illegal armed groups in the area especially in the Meta department, but also, to a smaller extent in the Guaviare department. In 2004, there are some larger groups of the left-wing FARC (*Fuerzas Armadas Revolucionarias*) guerrillas, and smaller groups of the paramilitary right-wing AUC (*Autodefensas Unidas de Colombia*), but no groups of the Marxist, left-wing ELN (*Ejército de Liberación Nacional*) guerrillas. For obvious reasons, there is a certain connection between the average number of people enrolled in illegal armed groups in a municipality and the presence of coca cultivation in the municipality (Ministry of Defence, in UNODC 2005).

To motivate farmers to keep their land free of illicit crops, the “Forrest Warden Programme” has been created nationwide by the UNODC, consisting out of several components regarding environmental, social and economic needs, as described in chapter 1.3 and by UNODC (2006).

5.2 Data Issues

The data used for this project were offered by the SIMCI II project in Colombia. Besides satellite image data, additional context information such as spray lines data as well as data of previous coca fields have been used.. External data are not necessary, but can be consulted for orientation purposes, such as information about villages and settlements.

5.2.1 Image Data

The basic raster data available for the project are a temporal series of Landsat images ranging from 2001 to 2006. They are located on the 7th path in the 58th row. The images cover the central part of the department of Meta and the northwestern part of the province of Guaviare (Figure 14). These Landsat ETM image data were provided by the ICMP of UNODC and are available from USGS/EROS, Sioux Falls, SD.

The characteristics of Landsat 7 –ETM+ are, as already presented in chapter 1.4, its 30 meters spatial resolution (with additional 15-meter panchromatic and 60 meter thermal band), its 16-day repeat cycle, and its swath width of 185 km.

Although about half a dozen images were available (table 1), mainly the images of 2005 and 2004 were used. The remaining satellite images served for reference purposes. Even though there are also SPOT and ASTER images in use for the SIMCI project, this thesis project focuses only on the use of Landsat imagery; very high resolution images or aerial photographs would be even more desirable for a reliable detection of potential cultivation areas, but they are not available due to economic restrictions.

Date	Comments
9th September 2006	High cloud coverage, about 30 percent of the image
12th January 2006	Scattered clouds; about 10 percent cloud coverage; thermal channel 6 missing
08th October 2005	Used for automating of interpretation
24th December 2004	Used for previous land cover classification
03rd October 2003	Very high cloud coverage, about 50 percent image
03rd March 2001	Free of clouds and of SLC-stripes, but out-of-date

Tab. 1: Satellite data available for the thesis

Though aerial photos are not available, there is a series of photos that have been taken during various overflights over the region. The images have been shot with a digital camera which is combined with a GPS unit to reconstruct the position of the plane. These images can serve as a kind of reference images and can give a better impression of what the coca cultivation areas look like. However, these photos have not been of practical use for the automatisation of the interpretation. This is due to lacking information about the direction of the camera while photographing, as well as there are no details about aperture angle and so on.

Fig. 14: Position of the satellite image of 2005 inside Colombia and the study area

5.2.2 Additional Context Data

There are two kinds of vector-based context information resources that are crucial for the interpretation process. On the one hand, there is information about coca fields from recent surveys: Polygons from the years 2003 and 2004 are available and can be used for the interpretation process. These data derive from the visual interpretation of Landsat images as well as from observations made during overflights done by Colombian staff (figure 15, left).

Another important source of information are the spraying lines: Here, the GPS-tracks of the plane spraying the toxin onto the plants are recorded and buffered. The date of the spraying set in relation to the date of acquisition of the satellite image is an important piece of information (figure 15, right).

5.2.3 Reference Data

Additionally, polygons from 2005 can be used; similar to the coca fields of the years before, these polygons indicate the position of coca fields that have been detected by visual interpretation. The data of 2005 serve as reference data for the accuracy assessment, in contrast to the data of 2004 and 2003, which serve as context information for the interpretation process.

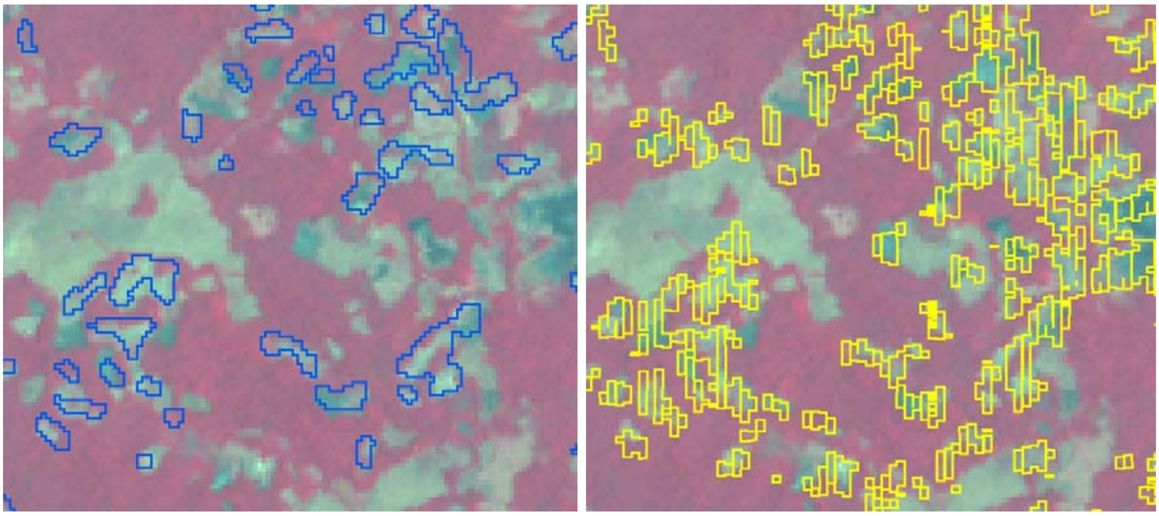

Fig. 15: Coca fields (blue, on the left) and spraying lines (yellow, on the right)

5.2.4 Data-related Problems

There are several data-related problems that have been encountered to aggravate the usability for interpretation. One of the basic problems when handling the data is their geometric mis-registration. So, errors can occur as the data sets do not match exactly. This is due to the use of different coordinate systems and coarsely rectified satellite images. Consequently, before the interpretation process, even more when it is automated, all relevant data have to be rectified (see chapter 5.3.1).

The Landsat images in use are subject to the failed scan line corrector (compare chapter 1.3), so nearly the entire image has thin black stripes running (with a width of at least one pixel, i.e. 30 meters or more) from southeast to northwest. Solely a stripe in the middle of the image is affected to a fewer extent by the SLC-off phenomenon.

As the area is part of the tropic equator zone, rich in rivers and forests evaporating water, there is always a cloud cover; thus about ten to 20 percent of the image cannot be interpreted as the land cover is not visible; additionally, cloud shadows affect the interpretability, as spectral values are modified.

The vector data containing the polygons of former coca censuses do not only refer to interpretation of satellite images, but do also refer to observations made from planes. Thus, when trying to collect sample areas for classification in accordance with the vector data, one has to keep in mind that vector data do not match perfectly with satellite data regarding the content.

The vector data containing the spraying lines information seem to be incomplete, as some areas strongly affected by coca cultivation have not been sprayed at all. It is not comprehensible, if the information has been masked out, or if the mentioned zones simply have not been sprayed at all.

As already mentioned, the spatial resolution (30 meters) of Landsat images is not optimal for interpretation, as smaller structures may remain undetected, and misconceptions can occur. This is especially problematic, when the small average size of coca fields, lying below 1.3 hectares, is considered.

5.3 Data Pre-processing

5.3.1 Image Rectification

As the data provided by the Colombian SIMCI project are geometrically miss-registered (compare chapter 5.2.4), the different data sets have to match in order to be combined. For this project, all relevant raster and vector data have been registered to the satellite image of 2005. All data were also registered to the Universal Transversal Mercator (UTM) zone 18 N, based on the Geodetic Reference System (GRS) 80. This UTM zone covers the investigation area.

The rectification is done with ground control points (GCPs). These can be highway intersections or distinct shoreline features (Lillesand, Kiefer and Chipman 2004); in the investigation area, features such as corners of fields, river mouths or road crossings can serve as GCPs. To increase the stability of the georeferencing, these control points should be distributed over the whole image.

5.3.2 Land Cover Classification

As the interpretation key also uses land cover classification of the previous year as reference information, the satellite image of September 24th 2004 has to be classified. There is one classification already executed by the Colombian project staff; however, this classification covers just a part of the satellite image, while a great part of the image is clipped out; therefore an independent land cover classification has been made, based on a pixel-wise classification. The existing classification done by Colombian staff was used for reference purposes. The Colombian classification consists of about 15 classes, however serving for the classification of the whole country. For the project, 13 classes have been created, based only on information available in the Landsat image of 2004 (Tab. 2).

Classes	Class Number	Classes (used in interpretation key; spanish)	Classes (Colombian Classification)
SLC off	0	*Gaps*	-
Bare Soil	1	*Suelo Desnudo*	Bare Soils
Stagnant water	2	-	Water Bodies
River	3	-	Water Bodies
Crop	5	*Coca*	Crops
Grassland	6	*Pasto, Vegetacion herbacea*	Grassland and Shrubs
Inundated Areas	12	-	Inundated areas
Cloud Shadow	13	-	Shadows
Secondary Forest	14	*Bosque secundario, rastrojo*	Secondary Forest and Shrubs
Primary Forest	15	*Bosque primario*	Primary Forest and Rainforest
Cloud	19	*Nubes*	Clouds
Slash and Burn	20	-	-
Sand Bank	33	-	Sand Banks
-		-	Roads
-		-	Urban and populated areas
-		-	Rock outcrops
-		-	Other

Tab. 2: Classes created for the land cover classification of 2004

The classification has been performed as a supervised classification, with training areas representing the features to be mapped and class signatures to be calculated. In order to increase the reliability of the classification about 5 training areas for each class have been chosen, for some cases more, in some classes less. As an algorithm, the maximum likelihood classification has been chosen (compare chapter 2.1.)

When preparing the classification and defining the classes a couple of problems have been encountered. One of the main difficulties for an automatic or supervised classification of vegetation and land cover in Colombia is the missing of a well defined crop calendar. Due to the situation in the tropics, there is no division of the year into phenological stages, as would be the case for moderate climates. Hence, most crops, including coca are cultivated and harvested throughout the whole year, making it difficult to separate coca plants from other crops based on temporal pattern analysis. Thus an automated land cover classification cannot be used to detect coca cultivation, but rather to study and put down the various land cover classes present on the image.

Another main problem is the similarity and confusion of various classes crop, grassland, bare soil and settlement as well as rock outcrop that have very similar spectral patterns. In order to reduce annoying misclassifications, the classes settlements and rock outcrops were left out; they were only of minor importance for the project. Also the land cover class road was not classified, as roads are too similar to other land cover classes and also were of minor importance for the project.

Ambiguities in the translation: As the decision tree has been published in Spanish, however most of the literature and this thesis are written in English, some ambiguities have to be dispelled. The terms *pasto*, *vegetacion herbacea* and *rastrojo*, used in the decision tree, can lead to definition problems. *Pasto*, meaning meadow or grass, can be assigned relatively easy to the land cover class “grassland and shrubs”; while *rastrojo*, meaning a stubble field can be associated with the class *bosque secundario y rastrojos altos* (“secondary forest and shrubs”). The term *vegetacion herbacea*, meaning herbaceous vegetation seems to have so far no counterpart in the land cover classification; however, it can be associated with the class “grassland and shrubs” (see table 2).

5.4 Segmentation

For the appropriate segmentation of the Landsat image, a broad range of segmentation parameters had to be tested and the correct parameters had to be chosen. In order to find the right parameters, a small test area of about 18 x 20 km has been clipped as a subset, where the range of different parameters has been tested. The test area should contain coca fields, in order to control the size of the image objects generated by the segmentation. The vector layer containing coca fields in 2004 has been added, visible through blue polygons in the images.

The first decisive question is the choice of the correct segmentation algorithm. As already described in chapter 3.3 there are three segmentation algorithms available in the Definiens Professional 5.0 software kit, plus the spectral difference segmentation algorithm which can be applied to refine an existing segmentation. As already indicated, the chessboard as well as the quadtree-based segmentation algorithms can be seen as inept for the purpose of this work, as the generated segments are square; however, the coca fields to be detected are rarely square. Hence, the parameters apply onto the multiresolution segmentation; additionally, parameters for the spectral difference segmentation have been tested.

The most important parameter in a multiresolution segmentation is, as already described, the scale parameter. Values between 1 and 10 have been tested, each one with the standard ratio between colour and shape of 0.9 to 0.1 and smoothness to compactness 0.5 to 0.5. A scale parameter of 1 produces very small image objects, often with a size of only one pixel. Scale parameter 3 results in a less detailed segmentation, but it is still too fragmented. Scale parameter 5 has provided good results, as the image is not too fragmented, but single cultivation areas can still be distinguished. Scale parameter 8 produces image objects that have too heterogeneous spectral values; this is even more the case with scale parameter 10. Image objects are too large and are not congruent with the coca cultivation areas.

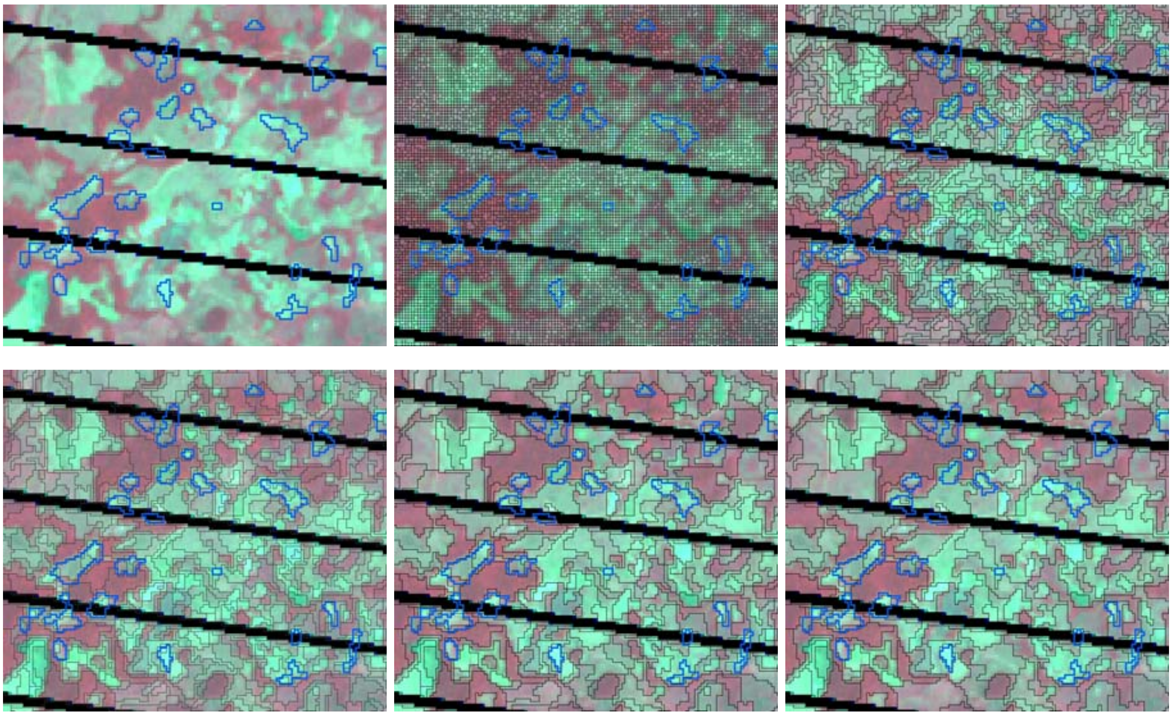

Fig. 16: Multiresolution segmentation of satellite image (left above) with scale parameters 1, 3, 5, 8 and 10 (right below)

Besides the scale parameter, another important parameter is the homogeneity criterion, consisting of colour, smoothness and compactness. Leukert (2002) has shown that results can vary considerably when modifying this tool.

The first parameter to be tested was the shape/colour ratio. The default value is 0.1 for shape and 0.9 for colour, as colour is much more important for defining spectral homogeneity than shape. In figure 17 there are segmentations with a medium scale parameter value of 5, and differing values for colour and shape. The first example shows the default settings, the following examples point out settings from 1 for colour and 0 for shape up to 0.1 for colour and 0.9 for shape. It becomes obvious that a segmentation with a stronger accent on the shape criterion does not come up to the landscape consisting of the scattered fields. Resulting polygons are too heterogeneous.

Another parameter, though of only secondary importance, is the relation between smoothness and compactness that constitutes the shape value inside the homogeneity criterion. The default ratio for these values is 0.5 to 0.5. To get a better impression of how these parameters work, each of them has been set to the value 1 and the other one to the value 0, with a standard colour / shape ratio of 0.9 to 0.1. Apparently, a smoothness value of 0 to compactness value of 1 creates image objects that are quite compact but are heterogeneous, while a smoothness value of 0 and a compactness value of 1 result in rather lengthy image objects, which give a hulking impression.

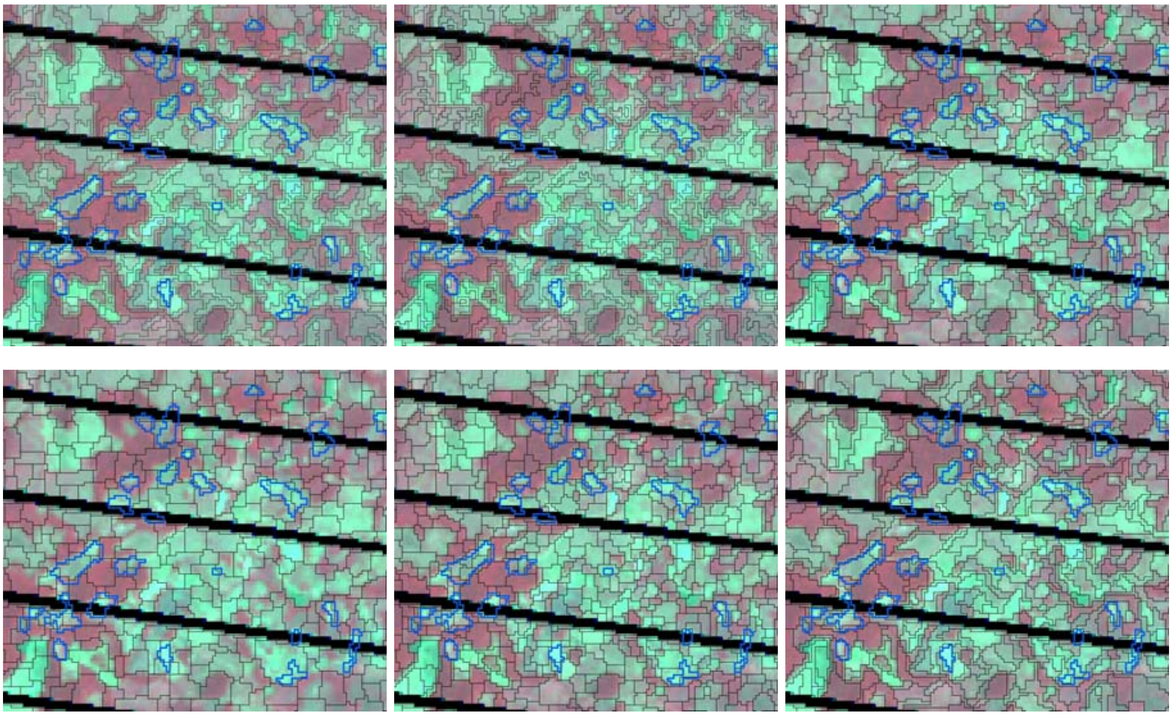

Fig. 17: Multiresolution segmentation with different homogeneity criterion settings (from left above to right below: colour 0.9 shape 0.1; colour 1 shape 0; colour 0.5 shape 0.5; colour 0.1 shape0.9; colour 0.5 compactness 1 smoothness 0; colour 0.5 compactness 0 smoothness 1)

After having segmented with the multiresolution algorithm, it is reasonable to merge neighbouring image objects with equal or very similar spectral values. In the following, a spectral difference segmentation has been carried out comparing different values. The most promising results have been reached with spectral difference values of 3 and 5.

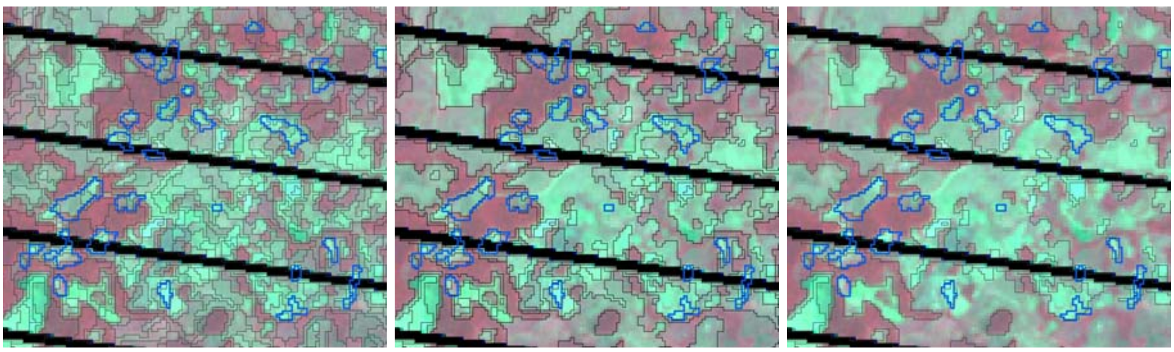

Fig. 18: Image with segments created by multiresolution segmentation (left) merged with spectral difference segmentation (middle, left)

As there is the possibility to weight the spectral bands of an image differently, also this parameter has been investigated. For that purpose, a correlation matrix of the spectral bands, comparing the correlation of the seven channels in Landsat 7 ETM+, has been made (Table 3). It can give useful information, also for an optimal band combination when classifying the images. For the subset that is under investigation a correlation matrix has been calculated. Apparently, bands 1 and 2 have a high correlation, also band 3, while bands 4 and 5 have a lower correlation with the first three bands as well as with each other. Thermal band 6 has correlation values between 0.68 and 0.81. The medium infrared band 7 has a high correlation value with medium infrared band 5, but a comparatively low value with the near infrared band 4.

Band	1	2	3	4	5	6	7
1	1,000						
2	0,969	1,000					
3	0,870	0,933	1,000				
4	0,864	0,845	0,686	1,000			
5	0,803	0,836	0,847	0,703	1,000		
6	0,814	0,792	0,687	0,799	0,719	1,000	
7	0,724	0,777	0,858	0,556	0,957	0,614	1,000

Tab. 3: Correlation matrix for spectral bands of Landsat image

A couple of different layer weight combinations have been tested. Table 4 gives a short overview of the different settings that have been tested. They have all been tested with 4 different spectral differences.

Layer weighting	Comment	Suitability
1-1-1-1-1-0-1	Weighting all layers equally except the thermal layer	good
1-0.8-1-1-0.8-0-1	Weighting layer 2 and 5 less due to high correlation with layer 1 respectively 7	good
0.5-0.5-1-1-1-0-0.5	Weighting layers that are important for visual interpretation in band combination 5-4-3	good
0.5-0.5-1-1-0.5-0-1	Weighting layers that are important for visual interpretation in band combination 4-3-7	good
0.5-0.5-1-1-1-0-1	Weighting all layers that are important for visual interpretation for 4-3-7 and 5-4-3	good
1-1-0-0-0-1-0	Weighting only layers 1, 2 and thermal layer 6	poor

Tab. 4: Tested Weighting settings

In theory, a plethora of combinations are possible. However, this tool should not be overrated, as the sum of all chosen weights for image layers is internally normalized to 1 (User Guide Definiens Professional 5, 2006). Finally, it can be said that the image layer bands do not have a very high influence on the results of the segmentations.

The additional context information (previous coca fields, spraying lines, previous land cover classification) has been weighted during the segmentation process. The weighting of the context based information is necessary to create segments, which can be queried by additional information. The weighting of these vector layers results in smaller segments, as every border of an object (e.g. a coca field) is considered for the segmentation.

On the basis of the observations described above, it can be stated that the majority of settings for segmentation that can be modified do not need to be changed, default settings will be applied. Neither a change of the homogeneity criterion, i.e. colour, shape, smoothness and compactness, nor refined choices of layer weights were able to improve the results of segmentation significantly. Rather, worse results were attained when changing the homogeneity criterion. Only a variation of the scale parameter of the multiresolution segmentation or changes in the spectral difference modified the results. Also the weighting of thematic layers was important.

Consequently, three levels of segmentation have been applied:

- a multiresolution segmentation with scale parameter 5
- a spectral difference segmentation of 3 (based on the multiresolution segmentation)
- a spectral difference segmentation of 5 (based on the multiresolution segmentation)

5.5 Implementing the Interpretation Key

When implementing the interpretation key into the software, all the different parameters have to be considered and paid attention to. This is one of the most challenging tasks of this work.

One of the most important points that has to be borne in mind is that the coca fields that have been detected visually are very heterogeneous. Thus it is almost impossible to classify all coca fields detected by the human eye correctly with the automated procedure. However, the important question arises, if the process should result into a higher or a lower extent of the real existing areas, i.e. should the classification result into a higher number of possible coca fields, or into a lower number of coca fields, so the interpreter has to take an additional view onto the areas.

The first alternative, classifying a higher number of areas, could be described as a more generous one, giving the interpreter the task to have a look with human intelligence onto the results and have a final decision. It can be called an overassessment. However, if the human-eye-screening is performed inattentively, it might result into an unjustified classification of non-coca fields as coca.

The second alternative, classifying a lower number, i.e. a more cautious approach, will have the consequence that coca areas remain undetected, so an underassessment takes place. It would be in the field of responsibility of the interpreter to find more fields than the fields already indicated.

Both alternatives are not completely satisfying, but rely on the given circumstances (image data, vector data); it seems to be more appropriate for the task to choose the overassessment approach, i.e. rather classifying too many fields as coca rather than classifying too less.

5.5.1 The Interpretation Key

The interpretation key has been designed by the Colombian project staff members in form of a decision tree (chapter 4.3). Two decision trees have been designed, one for bright areas, one for dark areas.

In the decision tree for dark areas the first query is whether the texture is fine or coarse; this query is not done in the decision tree for bright areas, as they represent bare soil areas, without plant cover. The next query deals with the question if there has been a coca field before (either in one of the last two years, or only in the year before) or not. In case of a positive response, the presence of spraying lines is queried. If there have been spraying lines, additionally the date of the spraying is requested. If the area has not been sprayed or there has been no coca before, the size, and the shape (for areas with more than three hectares) is queried. In either case, a query of the previous (i.e. in the year before) land cover is queried. Depending on the type of land cover, and of other query results, a certain class (either coca, a status to be checked) is assigned to the area. In case that an area is free from coca, it is assigned either to the class grassland or bare soil, depending on the colour (bright or dark).

5.5.2 Colour query

This is the first step when implementing the decision tree and it is crucial for the results. Two types of colour brightness were taken into account; they have just been called “bright” and “medium and dark” (these are treated jointly). It refers to the spectral band combinations 4-3-7, or 5-4-3. In the band combination 4-3-7 it will be a blue tone, while in 5-4-3, it will be a rather pink colour tone, with variations of hue.

The definition of the colours “bright”, “medium” or “dark”, and where the borders between these values lie, have not been explicitly mentioned; their definition depends on the experience of the interpreter. Even in the description of the features of the interpretation key (Bauer and Schneider 2005) there are just visual examples for the types of brightness, but no mathematical clues, such as RGB or IHS values or similar.

As there was no mathematical information given by the interpretation key, this query is hard to automatise, and should be done by an expert. However, as no such expert knowledge was available directly for this thesis, the expert knowledge was tried to remodel by a manual selection of sample areas from the segments in the image. The sample editor offered by Definiens Professional offers a good chance to differentiate between “bright” and “medium/dark” areas; as the difference between bright and medium or dark areas is vague, a fuzzy logic operator is used to describe the classes and to select samples.

Nevertheless, the act of selecting of samples and assigning them to either the class “bright” or “medium/dark” depends on subjective and individual decisions of the person that is in charge of preparing the decision tree; that can lead to a certain insecurity.

To increase the reliability of the samples, samples for “bright” and “medium/dark” areas have been chosen in accordance with the vector layer containing the coca fields.

It is important to keep in mind that this parameter relies a lot on subjective decision; hence it is one of the fine-tuning adjustments that only an interpreter with high experience for that region and task can determine.

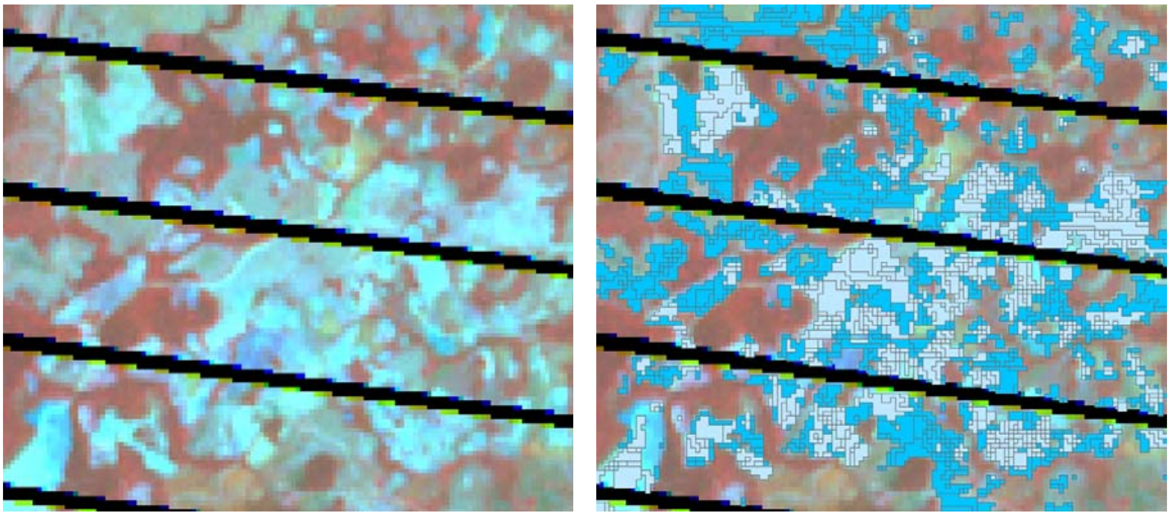

Fig. 19: Results of colour query (with dark objects in dark blue, bright objects in light blue)

5.5.3 *Texture Query*

The texture query applies only to dark/medium areas. This query is one of the more complicated tasks to realise in Definiens Professional, as the texture is not an easy-to-define attribute. Definiens Professional offers a concept of texture with two types of calculations: those based on sub-objects, as well as calculations based on a grey-level co-occurrence matrix (GLCM) or grey-level difference vector (GLDV).

Calculations based on sub-objects require a level that is below the object level to be handled. Features of the layer value texture, such as mean of standard deviation, average mean difference to neighbours can be calculated as well as form texture parameters such as mean and standard deviation of the area, mean and standard deviation of density, mean and standard deviation of asymmetry and mean and standard deviation of direction.

In a nutshell, the grey-level co-occurrence matrix (GLCM), based on Haralick, Shanmugal and Dinstein (1973), is a tabulation of how often different combinations of pixel grey level occur in an image. Depending on the direction (0°, 45°, 90°, 135°, with 0° standing for vertical, 90° standing for horizontal direction) a different matrix can be calculated, or a sum of all directions. Every GLCM is normalized through the partition of every cell value by the sum of all cell values. The resultant matrix is symmetrical. Its diagonal elements represent pixel pairs with no grey level difference; cells that are one cell away from the diagonal, represent pixel pairs with only one grey level and so forth.

A related approach to measure texture is the grey-level difference vector (GLDV) that represents the sum of the diagonals of the GLCM. It counts the occurrence of references to the neighbouring pixels' absolute differences. Both GLCM and GLDV methods are independent of the image's bit depth, with 8-bit-data considered to be most reliable.

Definiens Professional offers a wide variety of GLCM/GLDV features, such as homogeneity, contrast, dissimilarity, entropy, angular second moment, mean and correlation for GLCM, and angular second moment, entropy, mean and contrast for GLDV. Calculations of texture features based GLCM and GLDV are said to be very CPU demanding because of the calculation of the GLCM (Definiens Professional Reference Book 5, 2006), in practical work this did prove true.

When determining the appropriate features and parameters, layer value texture features based on sub-objects were checked; in combination with GLCM homogeneity values have been tested. It turned out though, that the application of mere layer values, based on sub-objects, yields more reliable results than the application of a mixture of sub-object-based and GLCM-based parameters.

For the project, an additional segmentation was carried out, regarding all objects classified as possible dark or medium colour fields. This segmentation has a scale parameter of one, to have the highest number of segments that are as small and thus homogeneous as possible. The smaller the sub-objects, the more the feature value approaches the standard deviation calculated from single pixels. Based on that additional layer, the calculation of standard deviations was carried out, for the layers 3, 4, 5 and 7 which are important for visual image interpretation. The standard deviation has been assessed with a sinusoidal curve going up from zero standard deviations for membership value zero up to seven standard deviations with membership value one for coarse texture, and vice versa for fine texture.

However, results did not turn out to be very convincing (Figure 20). A difference between fine and coarse texture was not apparent. All in all, it can be said that texture is a feature which is hard to grasp. For the final execution of the decision tree, this feature has been left out (see chapter 5.5.10).

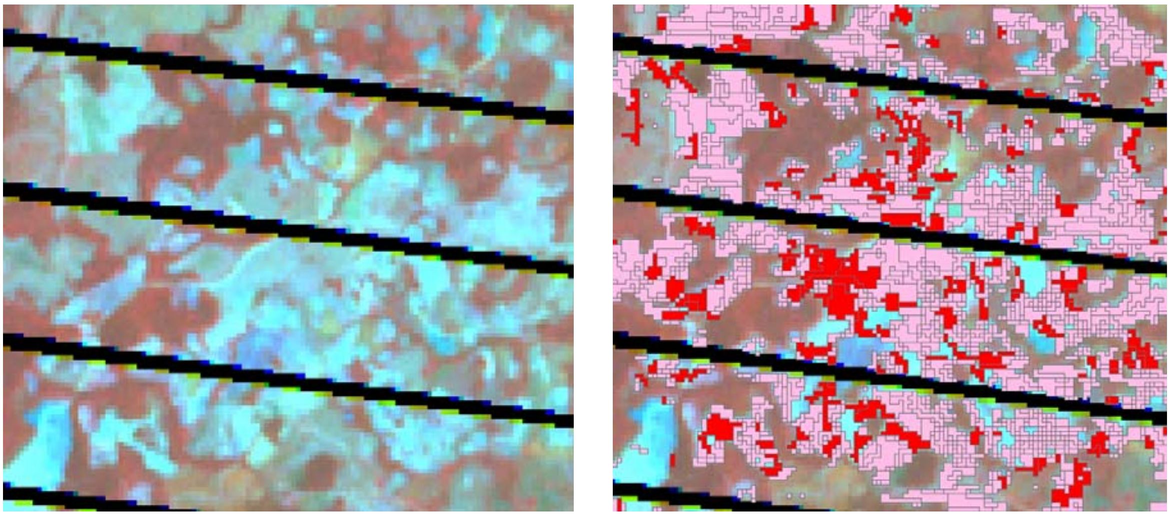

Fig. 20: Results of texture query (with objects having a fine texture in pink, and objects having a coarse texture in red)

5.5.4 Size Query

The query of size is one of the tasks of the implementation of the interpretation key that is more standardized and causes few problems to handle. As Definiens Professional offers the feature area size as one of the shape parameters, the differentiation between areas with more than or with less than three hectares is dispatched quickly.

When executing this query, one central problem arises: As every image object is treated independently, also here only the size of the particular image object is regarded. However, if a potential coca field with a size larger than three hectares is segmented into a couple of image objects with a particular size smaller than three hectares, the result will be arguable.

To represent large areas realistically through image segments, small neighbouring areas can be merged. But, due to the merge image objects can be bulky and unrealistic. Additionally, context information (such as previous land cover class) can be blended and thus be useless through the merge of image objects.

Finally, it has been decided to put the size query at the very end of the process tree, in order to be able to merge objects having the same context information. Figure 21 clarifies the query of size, with objects smaller than three hectares in bright red, objects larger than three hectares in dark red.

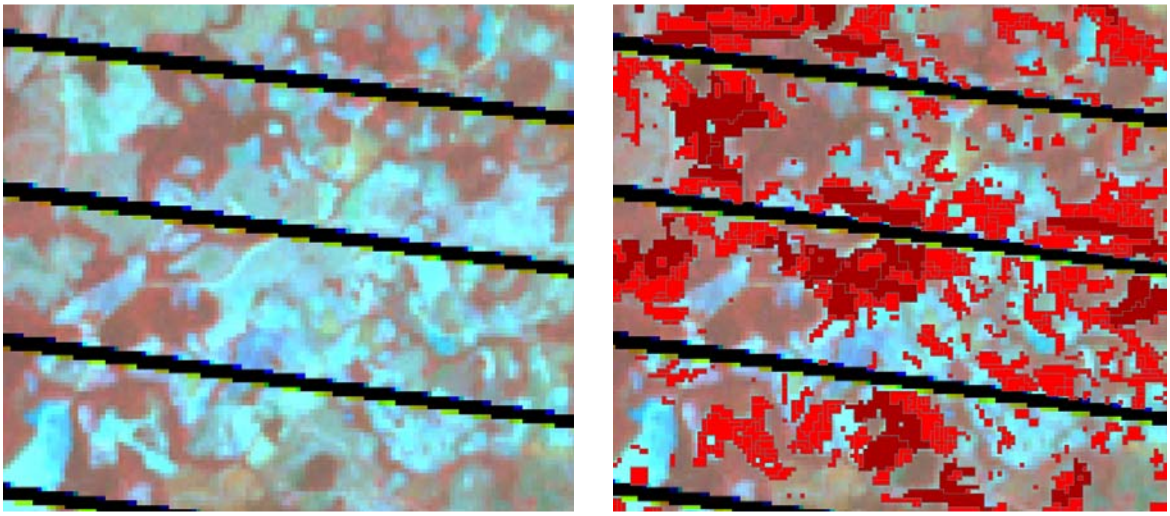

Fig. 21: Results of the size query (with objects smaller three hectares in bright red, objects larger in dark red)

5.5.5 *Shape Query*

Affiliated to the query of size is the query of shape, which is applied in the case of cultivation areas that have a size of more than three hectares. In the interpretation key, regular shapes are distinguished from irregular shapes; according to Colombian interpreters (in Bauer and Schneider 2005) most of the coca fields in the Meta-Guaviare region are regularly shaped.

Defining the shape parameters is not an easy task, as defining irregularity of shape is quite a fuzzy concept. Thus, fuzzy logic has been a very appropriate means when defining irregular and regular shapes. In the first approach to automatise the interpretation key, Bauer and Schneider (2006;2) did not apply this criterion at all.

Definiens Professional offers a couple of parameters defining generic shape features. However, when examining these parameters only a few of the offered parameters were found to be appropriate for the definition of a fuzzy logic membership function. In order to have reliable parameters, they should not depend on absolute values, but on relative values. Thus, absolute parameters such as area, length, width, border length, or the elliptic or rectangular fit are not appropriate for the description of an irregular or regular-shaped object. Two relative parameters, the compactness and the shape index, have been chosen and tested. Results were not too convincing.

Although a slightly satisfying solution has been found to query shape, this feature is also hard to grasp and thus to define; so, this feature will be left out, due to the simplifications made in 5.5.10. Additionally, results of the classification hardly vary, when removing the shape query. Figure 22 show objects larger than three hectares that are assigned to be either regular or irregular in shape.

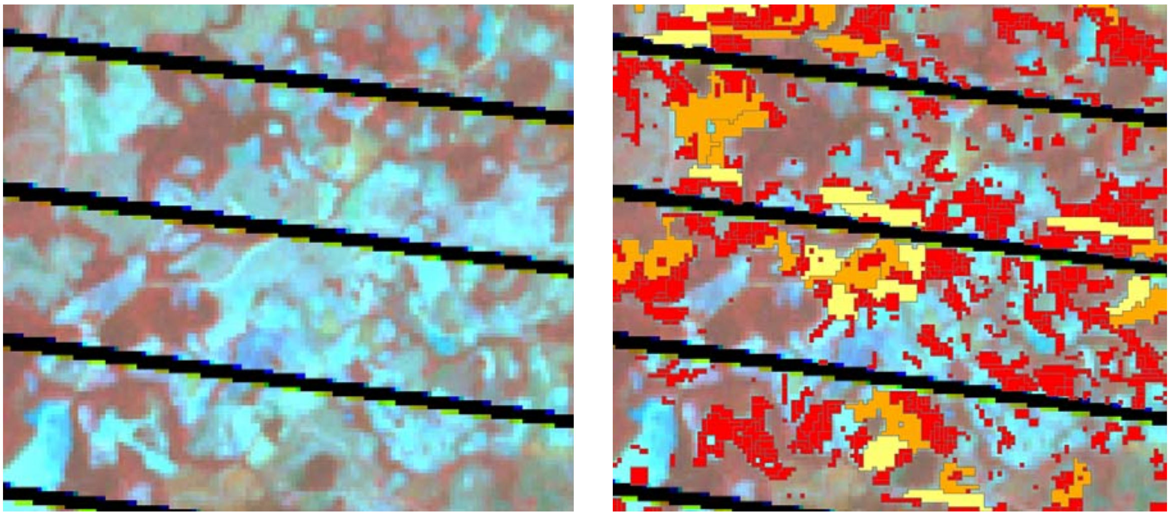

Fig. 22: Results of the shape query (with irregularly shaped objects in orange, and regularly shaped objects in yellow)

5.5.6 Query of previous Coca Fields

This query has to check, if there has been a coca field in the last year or the year before. In the interpretation key, there are two queries: One querying if there has been coca in either the last year or the year before, and a second one querying if there has been a coca field in only the last year. These queries are performed with the help of the additional information containing coca fields of the year 2004 and 2003, as described in chapter 5.2.2. They are added as a thematic layer into the project. The query has to capture all coca fields of the year 2004 and 2003 in the region. This is done by querying all object IDs. Results are shown in figure 23.

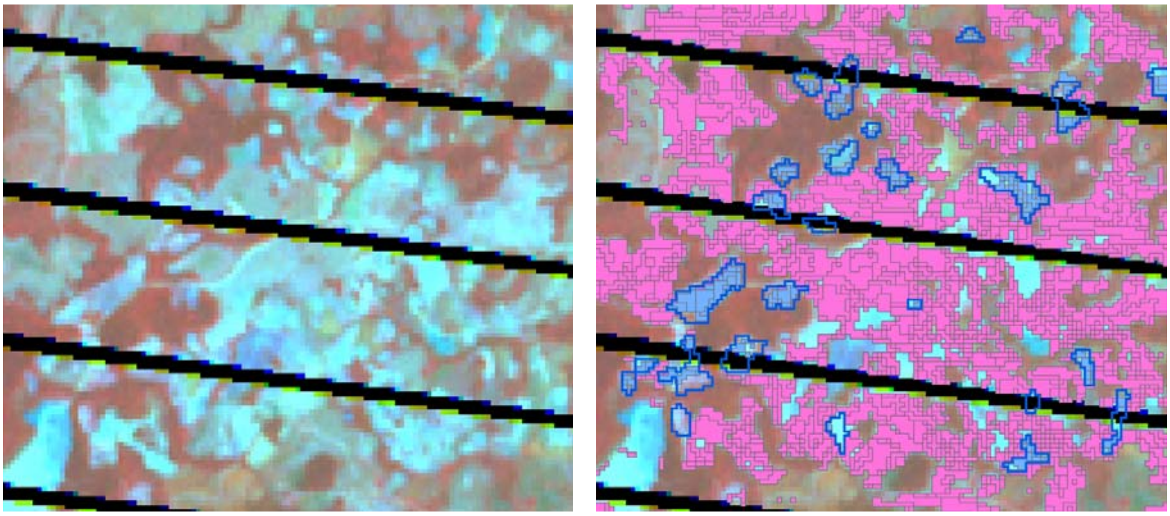

Fig. 23: Query of previous coca fields (with blue indicating that here has been a coca field before and pink that there has not been a field)

5.5.7 Queries of Spraying Lines

Spraying lines of the same year or of the former year are important features for the determination of coca fields. This query is also performed by the help of the vector layers containing spraying lines of the years 2004 and 2005. Spraying lines of 2005 are important for determining if a particular area has been sprayed before. This query is essential, as every branch of the decision tree runs through it. When implementing the query, all spraying lines have to be given an additional attribute, which is requested.

Adjacent to the query of the spraying lines is the query of the point of time of the spraying in relation to the date of the satellite image. If the spraying took place more than two months earlier, an area is treated differently from an area that has been sprayed in the last two months before the recording of the satellite image. In the vector layer an additional attribute field called “Date” is created from the information of the vector layer month. As the satellite image was recorded in October of 2005, all spraying lines before August 2005 are more than two months old, thus they are assigned value 1. All other spraying lines are assigned value 2.

Spraying lines of the previous year (2004) are requested only in a couple of queries, but are not as essential as the spraying lines of the same year.

Figure 24 shows a query of spraying lines. In case there has been a spraying, the area is coloured yellow; in case that there has not been a spraying the area is painted in blue.

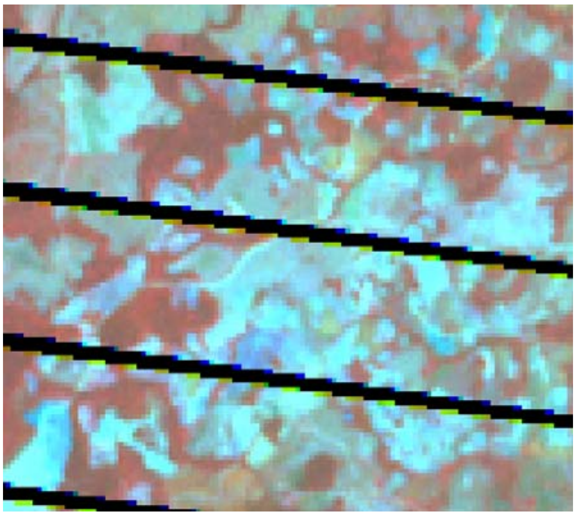 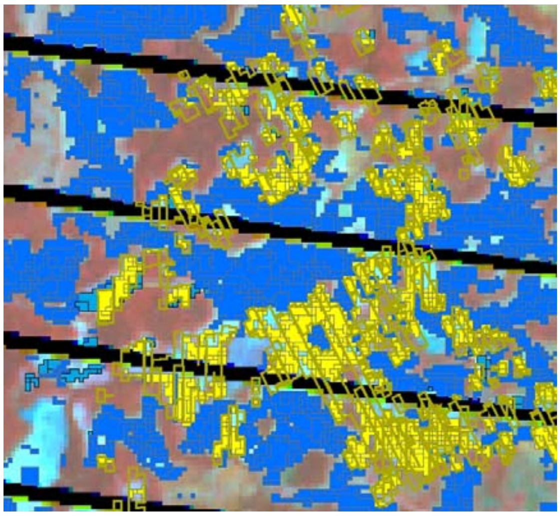

Fig. 24: Query of spraying lines (with areas that have been sprayed in yellow, and areas that have not been sprayed in blue)

5.5.8 Query of previous Land Cover

The final instance of the decision process whether a designated area is a coca cultivation area or not is the query of the land cover of the previous year. Depending on that land cover class, a cultivation area can be assigned either the final classes coca field (with the slightly variations of stable, new, recovered), bare soil, grasslands, or, as in the majority of cases, areas that need to be checked by overflight because of some vagueness. The land cover query is always executed if a field has not been a coca field in the previous year (i.e. after a size and/or shape query, thus on all areas that are either smaller than three hectares, or that are larger than three hectares, with either a regular or irregular form).

Depending on the branch of the decision tree, three or four land cover classes are taken into account. For example, in the case of bright colour coca fields that have an irregular shape, four possibilities of land cover classes are taken into account: cloud/SLC-off, bare soil, grassland, or primary/secondary forest are asserted the possible outcome classes. Figure 25 shall give an idea of how the query of previous land cover works.

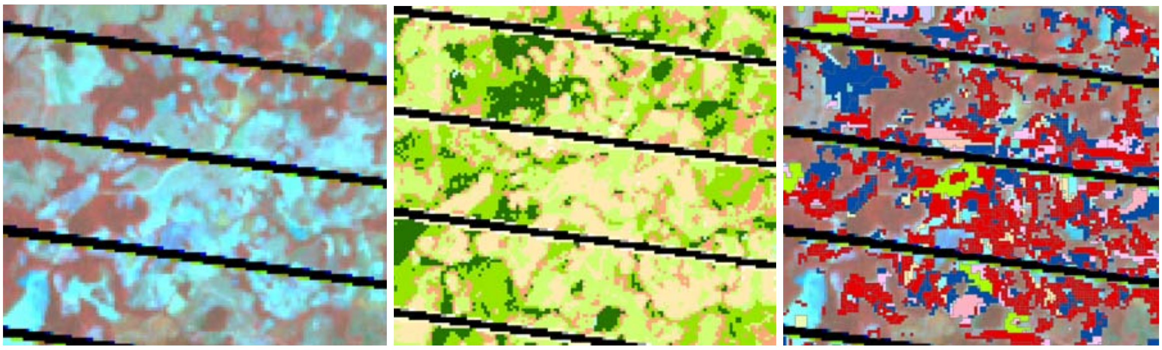

Fig. 25: Query of previous land cover

5.5.9 A System for the Decision Tree

As the decision tree shows a complex interlaced structure, with a couple of interconnections (which reduce its volume, but increase the complexity), finding a way through the structure, and remodelling it with the computer programme, is a difficult task. Thus an orientation or navigation system is required, which has to be comprehensible, reliable and traceable and its terms should be distinctive.

Most of the decisions are either binary (e.g. has there been a coca field in the recent survey- yes or no; size smaller or bigger than three hectares, shape regular or irregular, etc.), or refer to three or at most four land cover classes. Following that structure, a code system for the different outcomes has been developed, with two values, 1 and 2, or four land cover classes.

Value 1 stands for affirmation (“Yes”) in a Yes-/No-query (e.g.: Have there been spraying lines before?), but also for a field size smaller than three hectares, a regular shape, or a difference in time of less than two months. In the visualisation of the decision tree, all these variables stand on the left side of the query.

Value 2 stands for negation (“No”) in Yes-/No-query (e.g.: Have there been spraying lines before?), but also for a field larger than three hectares, an irregular shape, or a difference in time of more than two months. In the visualisation of the decision tree, all these variables stand on the right side of the query.

The land cover classes (cloud/SLC-off, bare soil, grassland, or primary/secondary forest) are labelled with an abbreviation; CL stands for cloud/SLC-off, BS stands for bare soil, GR stands for grassland, SF stands for Primary/Secondary Forest.

The values A and B stand for the bright respectively medium/dark colour tone.

In case of several possibilities for the encoding of one outcome class, because of an interconnection between nodes of the decision tree, the variant with the lowest number (i.e. coming from the left) is used; however, the suffix –multi shall indicate that there are more than one possible ways through the decision tree in order to get to that point. The first information gives information about the class, to which the area has to be assigned, due to the interpretation key.

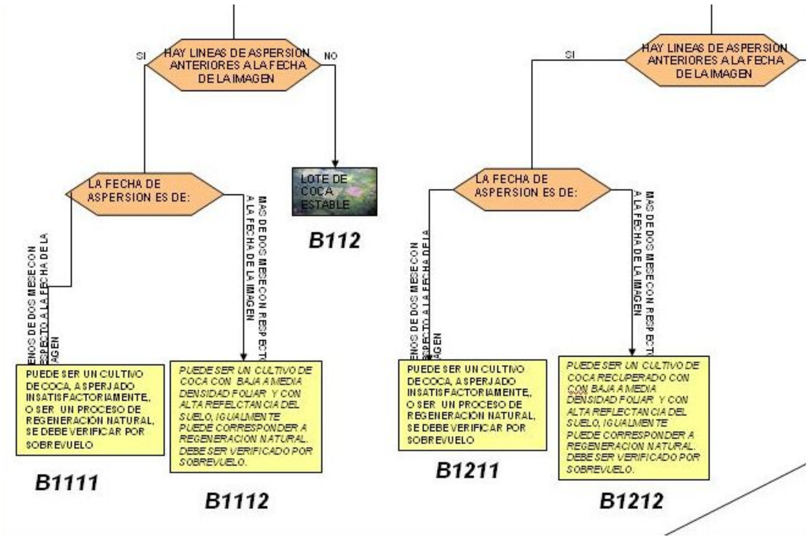

Fig. 26: Decision tree with coded branches and classes

Examples for deciphering the code:

- 2005baresoilB1111: This code stands for a bare soil area in 2005, based on a dark color area (A), where there have been coca fields from previous (1) and the last survey (1), where there have been spraying lines (1) and the spraying has taken place less then two months ago (1)
- 2005tobecheckedA2222BS: This code stands for a area that needs to be checked during an overflight by plane in 2005, based on a bright colour area (A), where there have not been coca fields from previous surveys (2), with an area size larger than three hectares (2) and irregular shape (2), where there have been no spraying lines (2), and the land cover classes from the previous classification has been bare soil (BS)

In some cases, the classification process stops at one of the classes in-between, and does not continue. It is not yet clear, why the process stops in these cases; however, to have a reasonable final classification result, these areas have to be assigned to a possible outcome class. For that reason, the additional outcome class '2005tobechecked_incomplete' has been created. All classes, which at the end of the classifications do not belong to on of the subclasses of '2005tobechecked_total', '2005grassland_total', '2005baresoil_total' or '2005coca_total', are assigned to the class '2005tobechecked_incomplete'.

5.5.10 Modelling of Queries with Definiens Professional's Process Tree

For implementing the complicated structure of the decision tree, the Process Tree of Definiens Professional 5 offers a clear visual aid to control and document the process. Also the interlaced structure of the decision tree can be reproduced, with a hierarchy of processes.

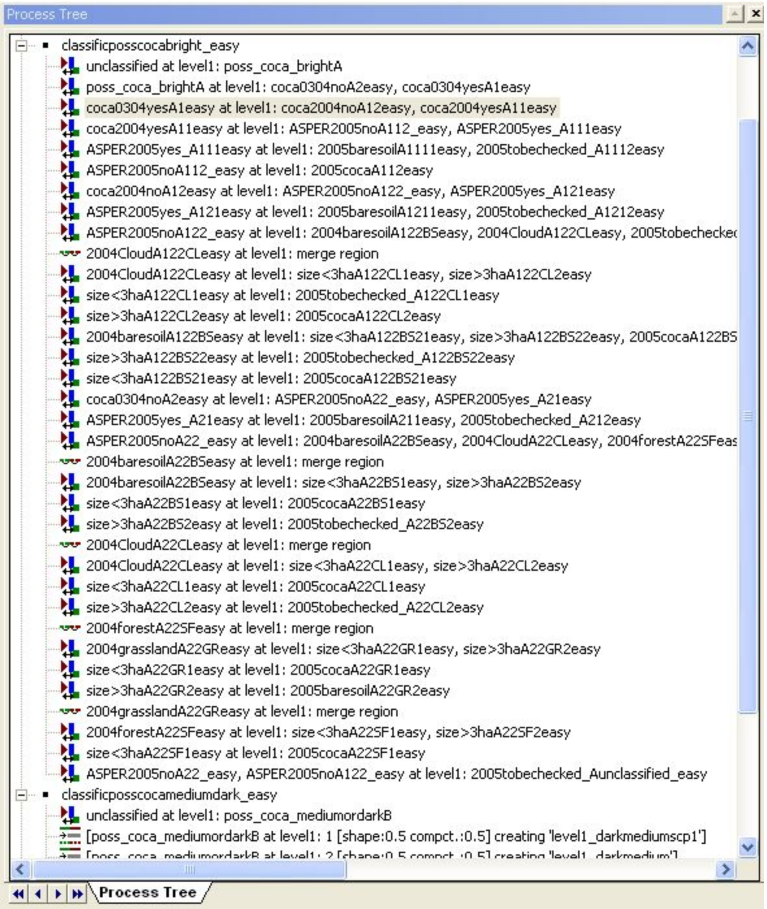

Fig. 27: Linear structure of queries

However, for the practical use and application, it has become obvious that executing a hierarchical structure of queries takes much more time than executing a linear structure that generates the same results. Although the clearness and ability to duplicate the process are better visualised through an interlaced, hierarchical structure, the linear structure is to be preferred for reasons of procession speed.

5.5.11 A simplified Approach – Modifications of the Decision Tree

As the structure of the decision tree has a couple of complicated features, there is an approach to simplify it. Both decision trees have been modified, considering these simplifications. The simplifications apply to the features texture, shape, size, and the interlaced structure of the decision tree.

The first modification applies to the feature texture: As the texture feature is quite difficult to define, it has been omitted. Thus, the classification of dark and medium coloured coca uses a decision tree reduced by half of its volume, covering only one instead of two pages (if printed out).

Although a reasonable approach has been found to query the feature shape, an exact definition of regularity/irregularity has not been found; additionally, resulting classes are in most of the cases very similar. Thus, this query has been omitted.

The size feature has not been omitted; however, it has been set to the end of the queries, following the previous land cover queries. By that, there are no ambiguities when querying the previous land cover in areas that have more than three hectares. In the simplified version, only segments with the same previous land cover are merged.

As the interlaced structure complicates the understanding and readability of the decision tree, each branch of the decision tree has been depicted. Therefore, in combination with the omissions of features, the whole structure is clearer. While in the original version of the interpretation key, outcome classes represented a couple of different possibilities of how a possible coca field could be classified, in the simplified interpretation key each outcome class represents just one possibility to reach it.

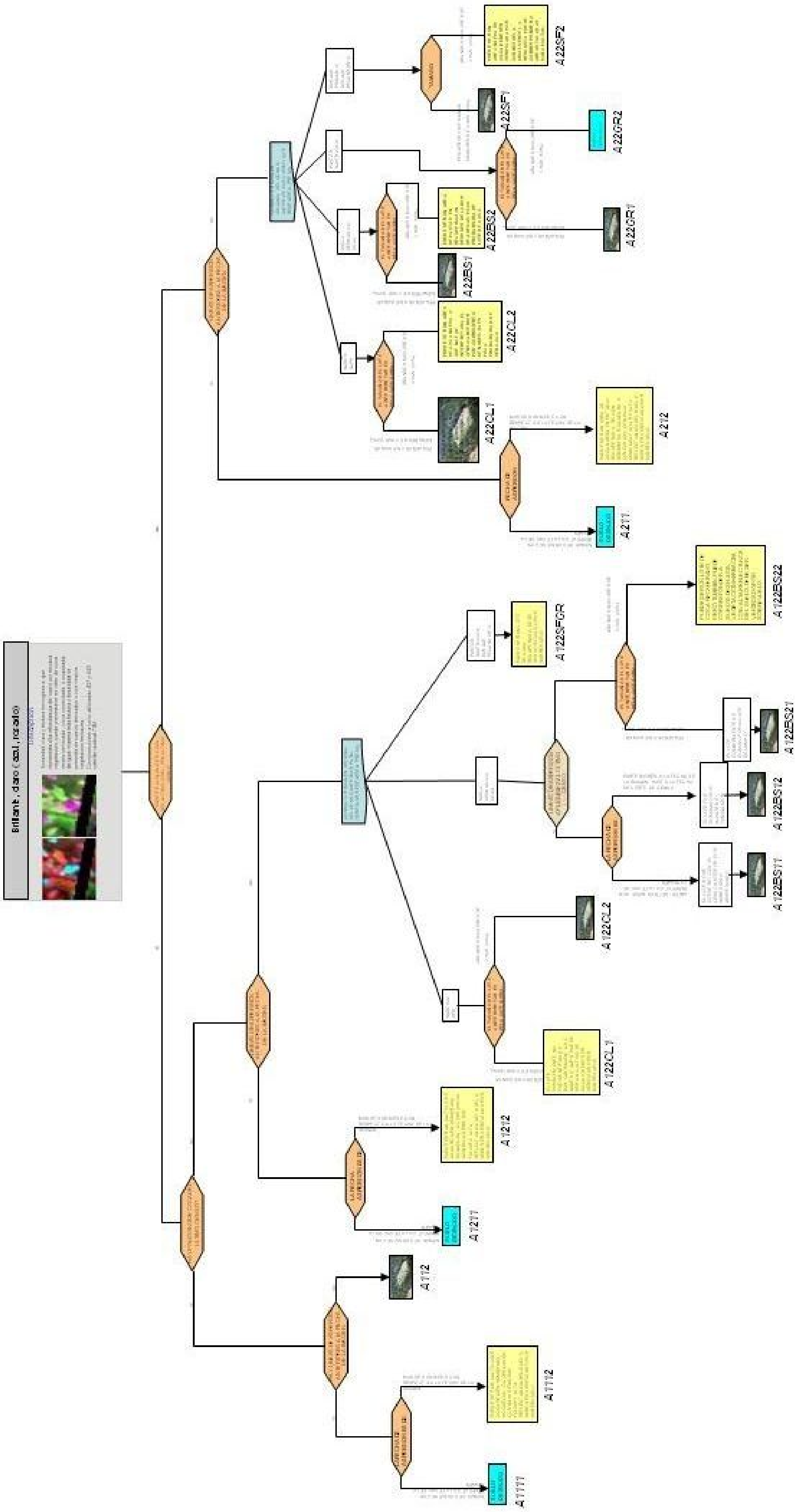

Fig. 28: Remodelled decision tree for bright tones

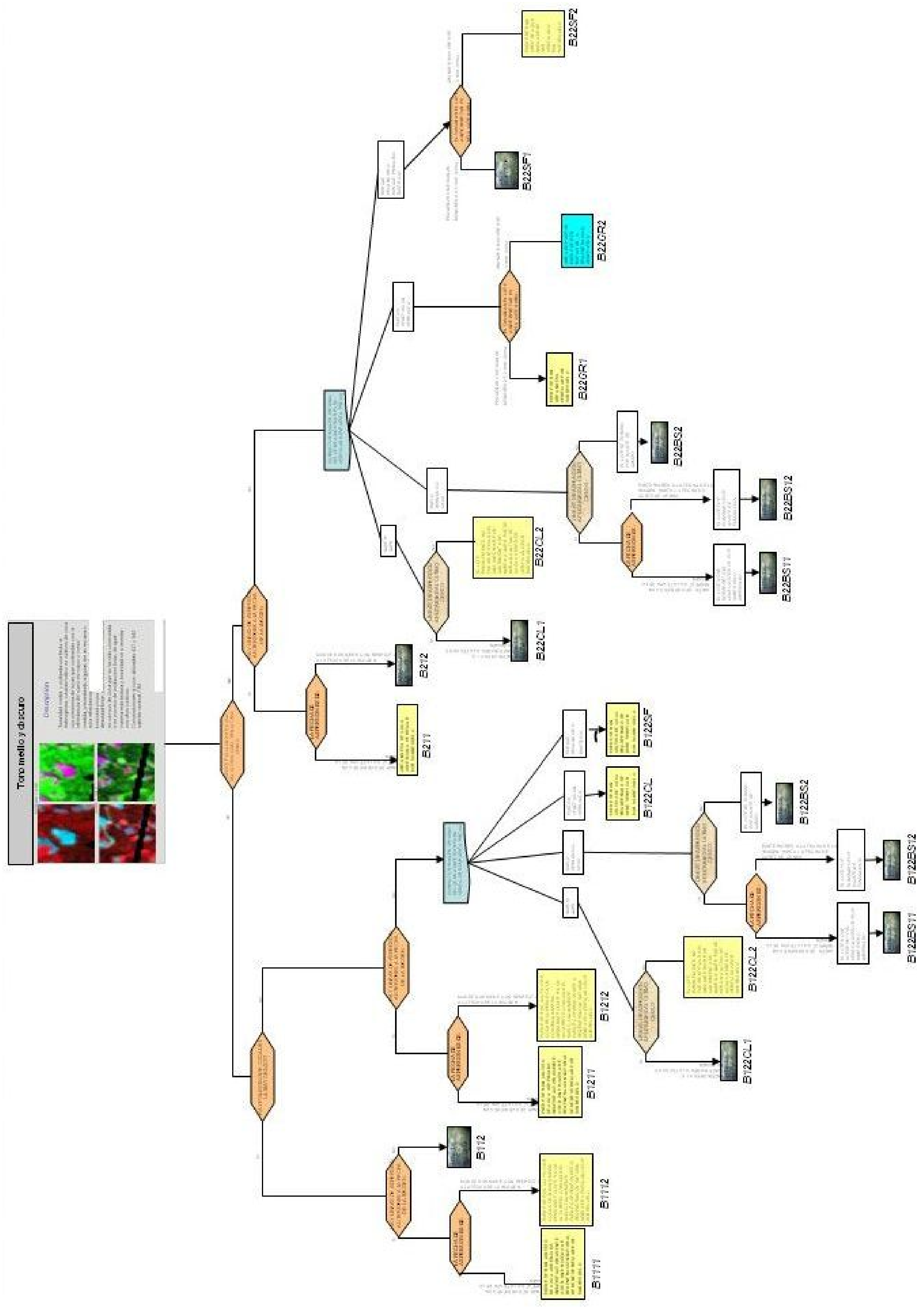

Fig. 29: Remodelled decision tree for medium and dark tones

5.6 Implementing the Process to the whole Landsat Image

To develop the process alls queries and analyses have been performed on the test area, i.e. a small subset of the entire Landsat image. However, as the process has to be applied to the whole satellite image, tests have to be made with segmentation of the entire image. As the file is of course more extensive, the computer capability has to be a lot higher. The random access memory has to be larger, as well as an increased processor power; a large hard disk capability with a lot of empty space (at least 20 Gigabytes) is also very important, where temporary files (.tmp) can be saved.

The process of segmenting the Landsat image takes very long, as the size of the subset is only 2.8 megabytes, while the entire Landsat scene has more than 386 megabytes. Every single process, whether it is loading or saving a project, loading data or processing data, is very time and CPU demanding. For instance, the segmentation of the whole Landsat scene at scale parameter 5, not respecting the thematic layers, will take about two hours on a system with a 2.8 Ghz processor and 500 MB RAM.

One problem for the classification of bright respectively medium or dark possible coca fields is the fact that samples have been selected only on the test area. In order to increase the reliability of the classification of possible coca fields, samples from all over the image should be chosen, in accordance with existing coca fields. However, navigating along the image and choosing samples is a time-demanding task itself, when executed with Definiens Professional.

Concluding, the central points that have to be kept in mind when transferring the interpretation process from the small clipped test area to the entire Landsat image are:

- Samples for the colour definition should be selected from affected regions on the entire Landsat image
- An adequate hardware for the processing of images is needed, with sufficient free space on the hard disk, as well as a processor and random access memory with sufficient capacities.

6 Accuracy Assessment

This chapter represents the results of the research to automate the interpretation of coca, with the aid of the interpretation key. It shall introduce the reader to the established parameters to measure the accuracy of the classification.

At a first glance, results seem to be sobering. Despite the remodelling, implementing and executing the process tree, a lot of areas are classified as coca that turn out to be no coca after verification (by the coca field vector layer of 2005 used for ground truthing) and vice versa. Considering the outcomes more thoroughly, one can see that the result is overestimated; more coca fields have been detected by the automated approach than by the visual interpretation. The manifold reasons will be tried to be given below.

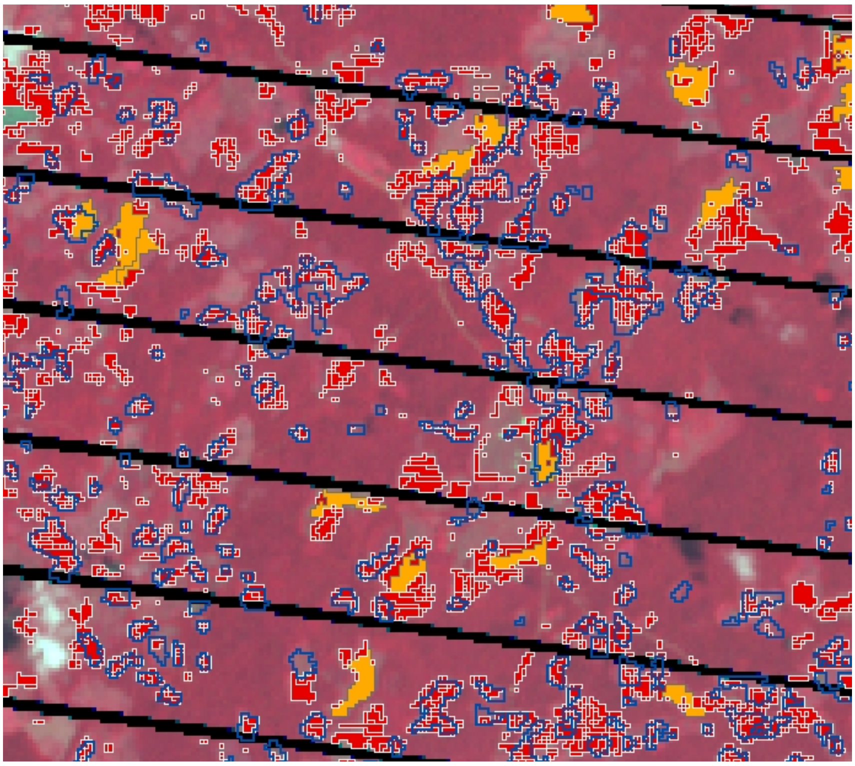

Fig. 30: Results of the classification for the entire Landsat image (red: coca; orange: area to be checked; blue polygons: visually interpreted coca fields) Multiresolution segmentation, Scale Parameter 5

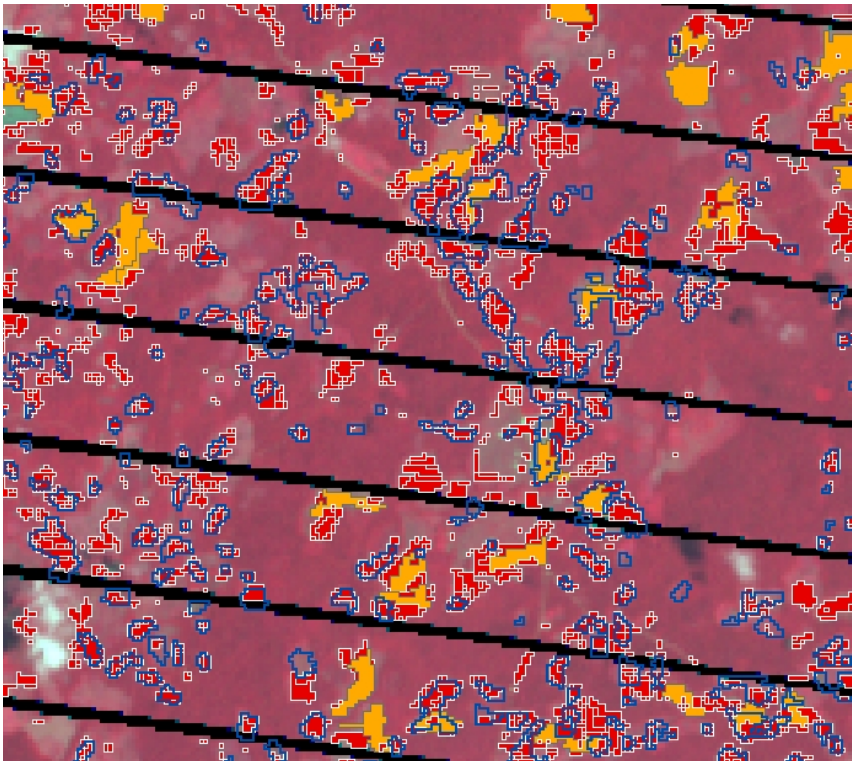

Fig. 31: Results of the classification for the entire Landsat image (red: coca; orange: area to be checked; blue polygons: visually interpreted coca fields) Segmentation with spectral difference 3

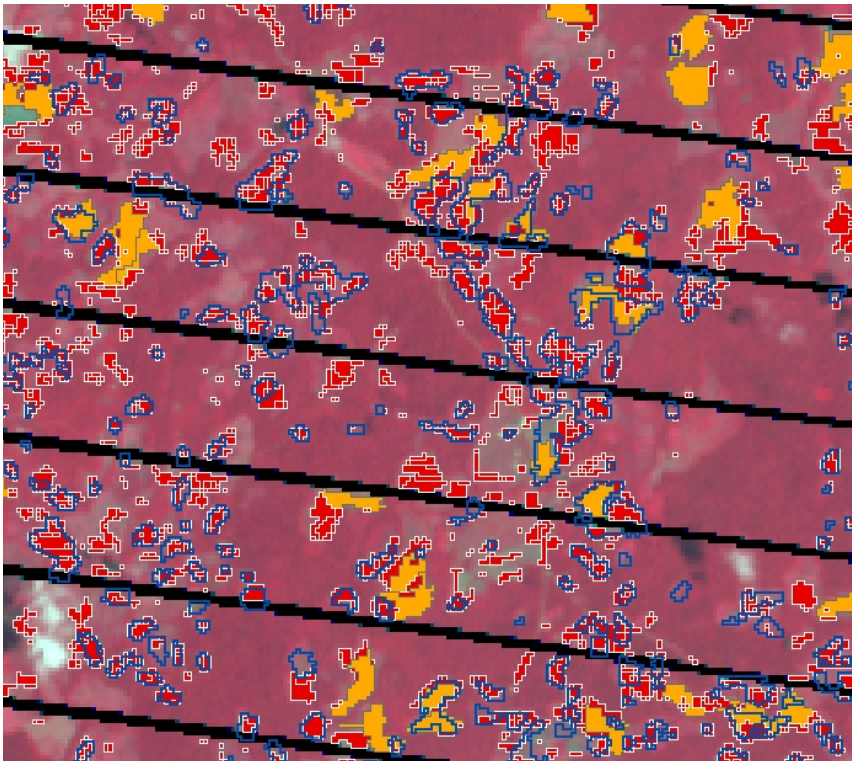

Fig. 32: Results of the classification for the entire Landsat image (red: coca; blue: area to be checked; green polygons: visually interpreted coca fields) Segmentation with spectral difference 5

To have a scientifically reliable evaluation of results, a classification accuracy assessment needs to be done (Lillesand, Kiefer and Chipman 2004). For that, one of the most common means is to prepare a classification error matrix, also known as confusion matrix or contingency table. These matrices compare the relation between the known reference data, i.e. coca fields in 2005, and the corresponding results of the automated classification on a category-by-category basis. As the number of rows and columns is equal to the number of categories whose classification accuracy will be assessed, the matrix is square. In the row, the classification data is represented, while in the column, the training set data (or ground truth data) is given. Every element in the matrix that is not diagonal (running from upper left to lower right) can be seen as errors of omission (exclusion of elements) or commission (inclusion of elements). Omission errors correspond to non-diagonal column elements, while commission elements are represented by non-diagonal row elements.

Several kinds of accuracy are distinguished:

- Overall accuracy: It represents the division of the total number of correctly classified pixels by the total number of pixels considered in the matrix.
- Producer's accuracy: It represents the division of the number of correctly classified pixels in each category by the number of reference pixels used for that category (the column total). This figure indicates how well reference pixels of the given cover type

are classified. To explain it more pictorially, it represents the probability that a certain object of reality is classified correctly.

- User's accuracy: It represents the division of the number of correctly classified pixels by the total number of pixels that were classified in that category (the row total). This figure serves as a measure to indicate the probability that a pixel classified into a certain category actually represents that category on the ground. To explain it more pictorially, it is the probability to find a certain object of a classification map in reality on the ground.
- The Kappa coefficient, also named "KHAT" (in Lillesand, Kiefer and Chipman 2004) is a measure of the difference between the actual agreement between reference data and an automated classifier, and the chance agreement between the reference data and a random classifier. It serves as an indicator of the extent to which the percentage correct values of an error matrix are due to "true" agreement versus "chance"agreement. In general parlance, it expresses how much the classification is better than a classification by chance.

In the ideal case, the Kappa coefficient approaches 1, as the true agreement approaches 1 and chance agreement approaches 0. E.g., a Kappa coefficient of 0.2 indicates that a given classification is by 20 percent better than a random assignment of pixels.

Table 5 shows the overall accuracy, as well as the producer's and the user's accuracy as well as the Kappa coefficient, for a couple of classifications with different settings. In many confusion matrices there is a large number of classes, which is controlled by random samples representing the ground truth (e.g. Lillesand, Kiefer, Chipman 2004, Richards 1993). In the project described here, it is an area-wide comparison between ground truth (the coca fields detected by visually interpretation) and the results of the automatic classification. Only two classes have been taken into account: Areas with coca, and areas without coca. This matrix is based on a pixel-per-pixel accuracy assessment.

The table is to be read as follows: While the first two rows represent values for a classification executed on just a small subset, the other rows represent classifications of the entire Landsat image. The segmentations on the entire satellite image are based on a multiresolution segmentation with scale parameter 5 (abbreviated as MultRes 5) and subsequent segmentations based on spectral difference segmentations with scale parameters 3 or 5 (abbreviated as SpDiff3 or SpDiff5). According to the interpretation key, classification results can either be coca or areas that will need to be checked during an overflight. The term 'total' marks that both types of fields have been assessed, while the term 'pure' means that only areas that have been classified as coca without any doubt have been assessed.

	Overall Accuracy	Kappa coefficient	Producer's Accuracy		User's Accuracy	
			Non-Coca	Coca	Non-Coca	Coca
SpDiff 5 (Subset,pure)	90,77%	0,4253	92,31%	67,13%	97,73%	36,27
SpDiff 5 (Subset, total)	90,75%	0,4257	92,28%	67,38%	97,75%	36,25%
MultRes 5 (LS, pure)	98%	0,1352	98,12%	57,98%	99,88%	7,92%
MultRes 5 (LS, total)	96,91%	0,0968	97%	62,64%	99,89%	5,52%
SpDiff 3 (LS, pure)	97,56%	0,1204	97,66%	62,44%	99,89%	6,93%
SpDiff 3 (LS, total)	98,58%	0,1738	98,7%	55,05%	99,87%	10,59%
SpDiff 5 (LS, pure)	97,9%	0,1362	98%	61,51%	99,89%	7,93%
SpDiff 5 (LS, (total)	97,42%	0,136	97,54%	61,88%	99,87%	7,97%

Tab. 5: Overall accuracy, Kappa coefficient, producer's and user's accuracy for classifications

On the first glimpse, results of the accuracy assessment (except for the small subset) seem to be confusing for the experienced human interpreter: While the total accuracy reaches almost 100 percent, the Kappa coefficient is quite low, not even reaching 0,20; while producer's and user's accuracy for non-coca reaches almost 100 percent, the corresponding values of coca lie around 60 percent at the producer's accuracy, and between five and eleven percent for the user's accuracy.

To explain this difference, one has to keep in mind that only two classes have been taken into consideration, of which the first, the non-coca class covers the largest part of the classified image. Thus it should not be too astonishing that this class has been classified correctly in wide parts.

The decisive coca class, however, covering just an inferior part of the territory was not always classified correctly; many more areas have been classified as coca that according to the ground truth are not coca. That is why the user's accuracies of the coca class, representing the ratio of the number of pixels correctly classified as coca by the total number of pixels classified as coca is so much smaller than the corresponding producer's accuracies, representing the ratio of the number of pixels correctly classified as coca by the total number of pixels representing the ground truth of coca.

For this project, the producer's accuracy is more important than the user's accuracy. Existing coca fields (the ground truth or training set data) should be classified as good as possible. If even more areas are classified as coca, also known as overestimation, the effects are less problematic than in case of an underestimation, when fewer fields are detected.

As table 5 presented above is based on pixel-per-pixel accuracy assessment, one might argue that the approach has been object-based, so the classification should be object-based as well. For example, if a coca field of the ground truth has been detected only partially, it still can be seen as fully detected. Table 6 shows the number of coca fields that have been detected, in relation to the total number of coca fields to be detected. Thus, these ratios can be compared to the producer's accuracy.

	Number of detected coca fields	Relation number of detected coca fields / total number of coca fields
SpDiff 5 (Subset, pure)	292	292/328 = 89,02%
SpDiff 5 (Subset, total)	295	295/328 = 89,94%
MultRes 5 (LS, pure)	6845	6845/10755 = 63,64%
MultRes 5 (LS, total)	6933	6933/10755 = 64,46%
SpDiff 3 (LS, total)	6924	6924/10755 = 64,38%
SpDiff 3 (LS, pure)	6808	6808/10755 = 63,28%
SpDiff 5 (LS, (pure)	6727	6727/10755 = 62,55%
SpDiff 5 (LS, (total)	6903	6903/10755 = 64,18%

Tab. 6: Ratio of detected coca fields / total number of coca fields

Here, the values are more promising. While the accuracy reaches more than 89 percent for the small subset, about 64 percent are reached for the entire Landsat image, depending on the settings of the classification. The classification based on a multiresolution segmentation with a scale parameter of five reaches the best value with 64,46 percent, when including areas that need to be checked.

As the scan line corrector (SLC) failed, the data covered below the black stripes cannot be classified by automatic means. However, during a visual interpretation, areas can be detected by deduction, or by observation from overflights. When executing the accuracy assessment, ground truth data overlapping with the lines of the failed SLC will be seen as a misclassification by the automated classification; hence, it is advisable to mask out the lines and erase them from the ground truth vector objects. This problem has been solved with a mask covering the SLC stripes (figure 33).

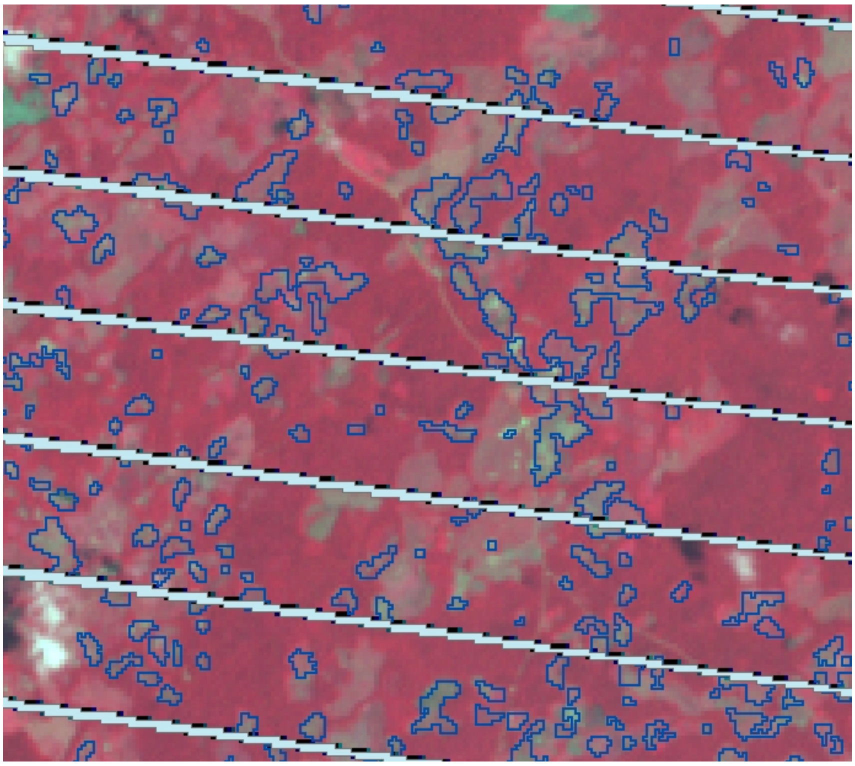

Fig. 33: Mask covering stripes produced by the failed scan line corrector

In case of fields that have not been detected by the automated classification, but have been detected by visual interpretation (i.e. ground truth), the main reason to be given is the feature colour. The colours of these fields often differ from the description stated in the interpretation keys. While coca fields are descried as bright, medium or dark blue, these fields often appear as green, or rather pink, others have a very mixed structure of pixels. One might assume that these pixels have been in a different phenological stadium then the rest of pixels. Also, when the satellite image was taken, some coca fields might not even have existed, but were planted later, and have been discovered during an overflight.

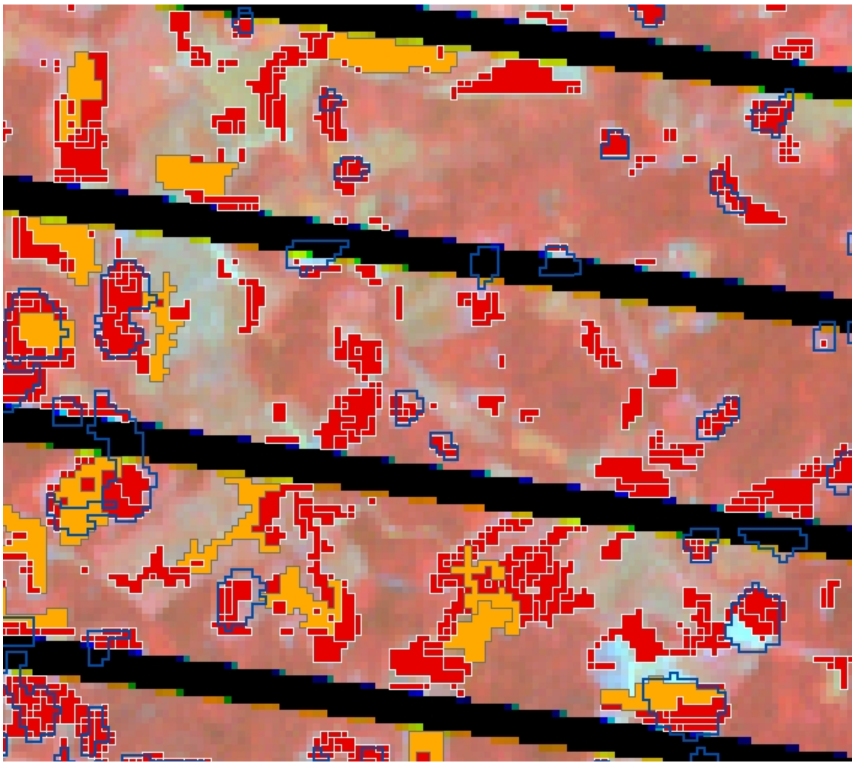

Fig. 34: Comparison visual interpretation (blue polygons) – automated interpretation (red polygons: coca – orange polygons: area to be checked)

In case of overestimation, i.e. fields that have been classified as coca by the automated process but have not been classified as coca by visual interpretation, it is the spectral similarity to other types of land cover. Several different types of land cover have very similar spectral values, for example the land cover types of bare soil and rock outcrop. Hence, the rock outcrops in the Serrania de la Macarena National Park are mistaken for bare soil, and are classified as possible coca fields. Also some pasture regions, especially in the northern part of the image, are completely free of coca cultivations; however, during the automated classification they are classified as potential coca fields, and therefore object to all queries of the interpretation key. During the accuracy assessment, there is a huge overassessment of potential coca fields, leading to a very low accuracy. In order to anticipate this problem, one could mask out certain regions of the image. This could also be helpful for areas like the Serrania de la Macarena mountains in the homonymous National Park, where rock outcrops are mistaken for bare soil due to spectral similarity.

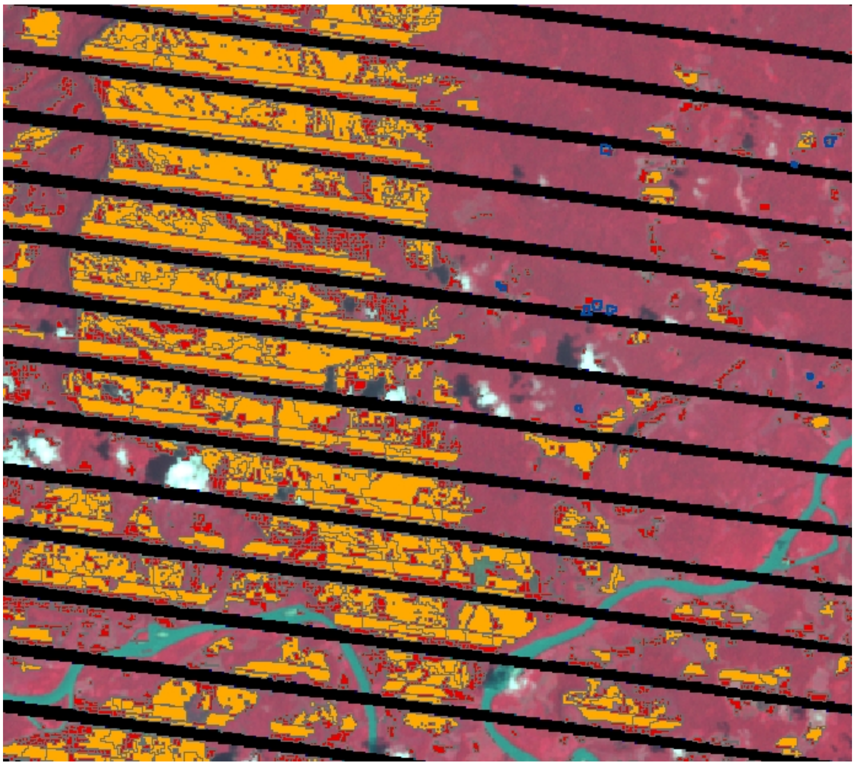

Fig. 35: Detected coca fields (red) and fields to be checked (orange) in comparison to visually interpreted fields (green) in the Serrania de la Macarena National Park Area

As already mentioned before, the data of the coca fields in 2005 which are used for ground truth, i.e. verification, were gained not exclusively by visual interpretation of the satellite images, but also during overflights by plane. Hence, areas that are no possible coca cultivation areas according to the satellite image can be marked as coca fields in the ground truth data, due to observations made during the overflight. A more detailed analysis of the results of the accuracy assessment is difficult, as the original visually interpreted reference data have been corrected after the overflights; uncorrected data, representing exclusively the results of the visual interpretation, have not been available. Also no differentiation has been made between the two main output classes coca and areas that might be coca areas and have to be checked. In the automated classification, a differentiation has been made, due to the rules stated by the interpretation keys. Thus, a comparison between automatically and visually interpreted data tends to be difficult.

7 Conclusions and Outlook

In the following, conclusions of the thesis shall be drawn. New insights shall be explained, the main problems that have been encountered shall be described, and finally an outlook shall be given. The objective of this thesis is to investigate a possible automatisation of the visual interpretation of coca fields in Colombia, using knowledge-based methods of classification, such as interpretation keys, and object-based image analysis.

It has turned out that one of the first points to mention is the quality and condition of data. Besides the lacking rectification, the vector data containing the visual interpretation of coca fields did not contain exact information whether polygons stem from interpretation of the satellite image or from observations made during overflights. It would be more adequate to compare results of the visual interpretation without the observations to the automated interpretation. Additionally, the spatial resolution of the Landsat 7 ETM+ images causes problems, as the average size of coca areas is quite small with 1.3 hectares, and the fact that smaller structures remain undetected.

In order to classify coca fields, appropriate image objects have to be created by segmentation that should be congruent to the visually interpreted fields. It has emerged that in many cases it is not possible to delineate image objects which match with objects that would have been demarcated by a human interpreter. Although a wide variety of segmentation parameters are provided by the software, only few of them can help to emulate the human delineation of objects. In many cases, image objects are problematic regarding the size, the heterogeneity or the shape.

During the implementation of the interpretation procedure as described in the interpretation key, several difficulties have occurred. A couple of features, such as colour, texture or shape are difficult to describe by parameters, and thus also in the interpretation key these features have not been stated sufficiently. Other features represent basic GIS queries, such as former land cover, or previous coca occurrence, which are definitely facile to implement.

When comparing results of the automated interpretation with the visual interpretation, it becomes evident that sufficient accuracy cannot be reached by the automated process at the moment. To improve results for future interpretations, several postulations can be set up.

On the one hand, the visually interpreted coca fields should be distinguished into data that have been interpreted on the satellite image, and data that have been added later due to observations. Additionally, satellite data with a higher spatial resolution should be used.

On the other hand, features such as colour, texture or shape, which rely on human perception and bias, shall be described in a more formal way, in order to emulate them better with the help of the image objects. Additionally, the structure of the interpretation key is strongly interlaced; so, as to improve the readability, the structure should be remodelled to be more linear. Here, new approaches of pruning (Rokach and Maimon 2005) should be taken into consideration.

As an outlook, one can say that a complete automatisation of interpretation is still a long way off in the future. Even if parameters will be refined and enhanced, there are still too many parameters that are human based and depend on human experience.

However, automated interpretation can be a valuable pre-screening tool for the visual interpretation, and can thus help to accelerate the process of interpretation. To improve this tool, further research on the transfer of human knowledge and experience into digital systems will be necessary.

8 Literature

-ASPRS – American Society for Photogrammetry and Remote Sensing (1997): Manual of Photographic Interpretation. Second Edition. Chief Editor: Philipson W.R., Bethesda, MD

-Baatz, M, Schäpe, A (2000): Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl/Blaschke/Griesebner (eds.): Angewandte geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000. Wichmann Verlag, Heidelberg; p. 12-23

-Ballard, D, Brown, C (1982): Computer Vision. Prentice-Hall, Englewood Cliffs, NJ

-Bauer, T, Kaiser, G (2006): Knowledge Transfer – Formalising an Interpretation Key. In: Lang S, Blaschke T, Schöpfer E (eds.): 1st International Conference on OBject-based Image Analysis (OBIA 2006), July 4-5, 2006, Salzburg, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVI (4/C42)

-Bauer, T, Schneider W (2006;1): Automatisation of the Interpretation Process for Coca Monitoring in Colombia. Unpublished project report on behalf of ICMP/UNODC

-Bauer, T, Schneider W (2006;2): Toolkit for the development of interpretation keys for illicit crop monitoring. Unpublished project report on behalf of ICMP/UNODC

-Bauer, T Schneider W (2005): Development of an Interpretation Key and an “Interpreter Calibration Technique” for the Monitoring Project in Colombia. Unpublished project report on behalf of ICMP/UNODC

-Benz, U, Hofmann, P, Willhauck, G, Lingenfelder, I, Heynen, M (2003): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. In: ISPRS Journal of Photogrammetry and Remote Sensing 58, p. 239-259

-Blaschke, T, Strobl, J (2001): What’s wrong with pixels? Some recent developments interfacing remote sensing and GIS In: Geo-Informations-Systeme. Zeitschrift für raumbezogene Information und Entscheidungen; GeoBIT 6/2001, p.12-17

-Braune, H, Semper, F (2006): Nah dran Kolumbien. Sebra, Hamburg

-Definiens Professional 5 User Guide (2006) Website: www.definiens.com (last query 13th September 2006)

-Definiens Professional 5 Reference Book (2006) Website: www.definiens.com (Last query 27th February 2007)

-Departamento Administrativo Nacional de Estadistica (DANE) (2005) Website www.dane.gov.co (Last query 14th May 2007)

-García Marín, J (2000): Colombia. Guía Azul. Gaesa, Madrid

-Haasen, C (2004): Die Wirkung von Crack. In: Stöver H, Prinzleve (eds.): Kokain und Crack – Pharmakodynamiken, Verbreitung und Hilfeangebote. Lambertus, Freiburg i. Breisgau

-Hake, G, Grünreich, D, Meng, L (2002): Kartographie. Eighth Edition. De Gruyter, Berlin, New York

-Haralick,R M , Shanmugam, K and Dinstein, I (1973): Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics; Vol. SMC-3, No. 3, p. 610-621, November 1973

-Kainz, W (2006): Fuzzy Logic and GIS. Handout for the lecture "Fuzzy Logic und GIS", Winter Semester 2006. University of Vienna

-Kainz, W (2007): Geographic Information Science, Technology, and Infrastucture. Handout for the lecture "Geographische Informationssysteme", Summer Semester 2007. University of Vienna

-Kemmesies, U and Werse, B (2004): "Das Gute" (kontrollierter) Kokainkonsum zu Beginn des 21. Jahrhunderts: Anzeichen eines Enkulturationsprozesses? In: Stöver H, Prinzleve (eds.): Kokain und Crack – Pharmakodynamiken, Verbreitung und Hilfeangebote. Lambertus, Freiburg i. Breisgau

-Kosslyn, S M , Koenig, O (1992): Wet mind.: The new cognitive neuroscience. Free Pree, New York

-Leukert, K (2002): Untersuchungen zur Segmentierung von Satellitenbilddaten für die Extraktion von GIS-Objekten; www.definiens.com/pdf/publications/Leukert_DGPF2002.pdf (last query 24th November 2006)

-Lillesand, T.M., Kiefer, R.W. and Chipman, J.W. (2004): Remote Sensing and Image Interpretation. Fifth Edition. Wiley, New York

-Marpu, P R, Gloaguen, R, Jonckheere, R (2006): Automatic identification of fission tracks using object-based image analysis, proceedings of the European Conference on Thermochronology, 30 July–04 August 2006, Ventura, B. & Lisker, F. (Editors). - Hannover: Dt. Ges. für Geowiss., H. 49, p.101-113

-Messina, J P, and Delamater, P L (2006): Defoliation and the war on drugs in Putumayo, Colombia. In: International Journal of Remote Sensing, Vol. 27, No. 1-2, January 2006, p. 121-128

-Mintzer, O (1975): Engineering Applications. In: Manual of Remote Sensing, Volume II – Interpretation and Applications. p. 1955-2109. ASP - American Society of Photogrammetry. Chief Editor: Colwell, R.N.; Falls Church (Virginia)

-Neisser, U (1976): Cognition and reality. Freeman, New York

-Pavlidis, T (1977): Structural Pattern Recognition. Springer-Verlag. Berlin, Heidelberg, New York

-Peuquet, D J (2004): Representations of Space and Time. Guilford, New York

-Pinz, A (1994): Bildverstehen. Springer Verlag Wien, New York

-Pombo, D (1990): Perfil ambiental de Colombia. Escalá, Bogotá

-Rokach, L and Maimon, O (2005): Top-Down Induction of Decision Trees Classifiers - A Survey. In: IEEE Transactions on Systems Man and Cybernetics Part C – Applications and Reviews 35 (4): p. 476-487

-Steinnocher, K, Kressler, F, Franzen, M, Ries, C (2007): Objekt-orientierte Klassifikation von Orthophotos zur Aktualisierung der digitalen Katastralmappe. In: Österreichische Zeitschrift für Vermessung und Geoinformation 1/2007: p. 3-9

-Stöver, H (2004): Bestandsaufnahme "Crack-Konsum" in Deutschland: Verbreitung, Konsummuster, Risiken und Hilfeangebote. In: Stöver H, Prinzleve (eds.): Kokain und Crack – Pharmakodynamiken, Verbreitung und Hilfeangebote. Lambertus, Freiburg i. Breisgau

-Unwin, D (1996): GIS, spatial analysis and spatial statistics. Progress in Human Geography. 20, p. 540-551

-UNODC Official Homepage (2007) http://www.unodc.org (last query 21st January 2007)

-United Nations Office on Drugs and Crime (UNODC) (2005): Colombia Cultivation Survey for 2004, June 2005

-United Nations Office on Drugs and Crime (UNODC) (2006): Colombia Cultivation Survey for 2005, June 2006

-User Guide eCognition 2.0 (2001) Website www.rstc.msstate.edu/wfd/ecog/User_Guide/index.htm (last query: 25th April 2007)

-User Guide eCognition 3.0 (2003) Website: www.definiens.com (last query 10th December 2006)

-User Guide eCognition 4.0 (2005) Website: www.definiens.com (last query 5th December 2006)

-Article OBIA. In: University of Calgary; Wiki: http://wiki.ucalgary.ca/page/OBIA (last query 25th April 2007)

Curriculum Vitae

Leonhard Suchenwirth

Born: 15/06/1981 in Kassel, Germany

Nationality: German

Student adress: Piaristengasse 33/6, 1080 Wien, Austria

Home adress: Seestraße 12, 82211 Herrsching-Breitbrunn, Germany

leonhard.suchenwirth@univie.ac.at

Education:-1987-1991 Grundschule Herrsching am Ammersee

-1991-2000 Secondary school: Max-Born-Gymnasium, Germering

-since Oct 2001 Student at the University of Vienna: Geography

-Sept 2003 - Feb 2004 Exchange study at the University of Barcelona

-since Mar 2004 Specialisation on Cartography and GIS

Languages: -English: fluent -Catalan: fluent -Latin

-Spanish: fluent -French: advanced

Computer:-Microsoft Office: advanced

-ArcGIS, Arc View: advanced

-ERDAS Imagine 8.7: basic

-Definiens Professional / eCognition Professional 4.0 and 5.0: advanced

-Macromedia Dreamweaver, Director, Freehand, Flash: basic

-Adobe Photoshop, Illustrator: advanced

Internships and work experience:

-Dec 2002 - Jan 2003: Österreichisches Ost- und Südosteuropa Institut, Vienna; Data research for internet portal on Cooperation in the Danube Region

-Aug - Sep 2004: Niedersächsiches Landesamt für Ökologie (NLÖ), Hildesheim, Germany; Assistance in the GIS department; participation in GIS-trainings and UMN-Mapserver User Conference

-Sept 2005: Freytag & Berndt und Artaria Verlag, Vienna; Cartography, hiking maps; working with MicroStation and I/RAS

-since Oct 2005: University of Vienna: Tutor for cartographic course: “Spatial Reference Systems” by Prof. Wolfgang Kainz

-Feb – Apr 2006: German Aerospace Center (DLR) – Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany: Disaster Observation Center (ZKI): Working with high resolution images, GoogleEarth-KML, Disaster monitoring

-Jul 2006: Assistance and participation at the international conference “Geoinformation Connecting Societies (GICON)” in Vienna

Vienna, June 2007

Leonhard Suchenwirth

Ich versichere:

- dass ich die Diplomarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.
- dass ich dieses Diplomarbeitsthema bisher weder im In- noch im Ausland (einer Beurteilerin/ einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.
- dass diese Arbeit mit der vom Begutachter beurteilten Arbeit übereinstimmt.

Wien, Juni 2007 Unterschrift