University of Hohenheim

Institute for Phytomedicine

Department of Weed Science

An image analysis and classification system for automatic weed species identification in different crops for precision weed management

Dissertation

Submitted in fulfilment of the requirements for the degree „Doktor der Agrarwissenschaften“ Dr.sc.agr./Ph.D. in Agricultural Sciences

to the Faculty of Agricultural Science
University of Hohenheim

submitted: November 2010
by: Dipl.-Ing. Martin Weis
born 18. September 1972
in Detmold

Date of oral Examination:	23.6.2010
Examination Committee:	
Supervisor and Review	Prof. Dr. R. Gerhards
Co-Reviewer	Prof. Dr. W. Claupein
Additional Examiner	Prof. Dr. Dr. hc. mult. K. Köller
Vice-Dean and Head of the Committee	Prof. Dr. A. Fangmeier

I confirm that all this work is my own except where indicated and that I did not use any other than the stated resources. I have completed the dissertation independently, according to the doctoral regulations of the University of Hohenheim, § 8.2.2.

[Martin Weis] Stuttgart, November 12, 2010

Abstract A system for the automatic weed detection in arable fields was developed in this thesis. With the resulting maps, weeds in fields can be controlled on a sub-field level, according to their abundance. The system contributes to the emerging field of Precision Farming technologies.

Precision Farming technologies have been developed during the last two decades to refine the agricultural management practise. The goal of Precision Farming is to vary treatments within fields, according to the local situation. These techniques lead to an optimisation of the management practice, thereby saving resources, increasing the farmers outcome, reducing the overall management costs and the environmental impact. A successful introduction of Precision Farming involves the development of application equipment capable of varying treatments and sensor technology to measure the spatial heterogeneity of important growth factors. Such systems are able to record, store and use large amounts of data gathered by the sensors. Decision components are needed to transform the measurements into practical management decisions. Since the treatments are varied spatially, positional data, usually measured using GPS technology, has to be processed. The located measurements lead to a delineation of management zones within a field and are represented by geo-data and can be visualised in maps. The improved, detailed knowledge of the situation within the field leads to new and extended scopes of applications and allows to document the management practices more precisely.

In this work, parts of Precision Farming technology were developed for site-specific weed management. Five selected publications are presented, covering the technological prerequisites and details of the developed system.

Weed management is a necessary management practice for all cropping systems. In most crops the weeds are controlled chemically using a uniform application of herbicides within a field, although most weeds are usually not uniformly distributed. Actually the weed infestation levels vary from no, low to high densities. This property of the weed distributions makes the development of site-specific technology feasible. Large savings of herbicides can be achieved, if only the nests of high weed infestation are sprayed. A reduction of the amount of herbicides is beneficial for the economy and ecology. A prerequisite for the site-specific treatment of weeds is the knowledge about the weed infestation level and distribution within the field. From the weed infestation level, manage-

ment decisions about the type and amount of herbicide to be applied on each part of the field can be derived. The techniques for a site-specific management of weeds are reviewed in Weis et al. (2008). In addition, the paper reveals yield losses due to weed infestation and unnecessary use of herbicides as modeled based on On-Farm-Research experimental data. Sensor requirements and analysis techniques of weed detection systems are the topic of Weis and Sökefeld (2010).

A computer vision system for the site-specific measurement of the weed infestation was developed. The system uses bi-spectral camera images of red and infrared light spectra as input and processes the images to derive the local weed infestation within the field. Since there is a large number of weed and crop species to be considered and their appearance changes due to the growth stage and conditions, the system has to be flexible and must be adaptable to the occurring crop–weed communities.

The input images are processed to separate the plants from the background (soil, mulch, stones) and shape features are derived for each plant in the image. In Weis and Gerhards (2007a) the calculation of different shape features is presented. Region-based, contour-based and newly developed features, which are based on a skeletonisation operation, were computed. A subset of the features was capable to distinguish monocotyledonous, dicotyledonous weeds and *Brassica napus* L. from *Hordeum vulgare* L. with an accuracy of 98%. The shape information is used to classify the plants as weed or crop. To handle the variation of their appearance in several growth stages, a class scheme was defined and prototypes for the different species were selected and catalogued. A database was created to maintain all the necessary information about the shape parameters and weed/crop classes. The shape analysis is done with classification algorithms, which are trained with the prototype information from the database. The shape features were tested for their discriminative abilities to identify different weed and crop species, indicating that the newly developed features are well suited for a classification. The classifier is applied to classify all plant objects found in an image according to their shape. Different classification algorithms were compared in Weis et al. (2009), indicating that Support Vector Machines are suitable to handle the complexity of the stated classification problem.

The image analysis and classification result in weed maps, containing the number of weed and crop plants in each image. These weed maps are the basis for the creation of application maps. Decision rules, based

on an aggregation of classes, were applied to the weed maps to delineate management zones for each herbicide. Weis and Gerhards (2009a) focus on the creation of application maps from the classification results, deriving decision functions for the conversion of weed maps into application maps. The automatic weed sampling results showed high correlations to a manual sampling.

The application equipment, a sprayer with 3 m–wide controllable boom-sections, was capable of varying the type and amount of three herbicides simultaneously. Therefore three different application maps can be used to realise a precise mixture for each management zone within the field.

The technique was successfully applied in different crops (maize, *Zea mays* L.; spring barley, *Hordeum vulgare* L. and sugar beet, *Beta vulgaris* L.), showing the potential for an introduction into practise. The developed database and analysis techniques were the basis for the development of a commercial sensor system, which will be available soon and ready for applications on farms.

Kurzfassung Ein System zur automatischen Unkrauterkennung auf landwirtschaftlich genutzten Flächen wurde in dieser Arbeit entwickelt. Mit den erstellten Karten können Unkräuter auf Schlägen teilschlagspezifisch, nach ihrem tatsächlichen Vorkommen, bekämpft werden. Das System ist Teil der Entwicklungen im Precision Farming (Präzisionslandwirtschaft).

Precision Farming Technologien wurden in den letzten zwei Jahrzehnten entwickelt, um die landwirtschaftlichen Bearbeitungsstrategien zu verfeinern. Das Ziel von Precision Farming ist die Variation der Bearbeitung innerhalb von Schlägen anhand der lokalen Situation. Diese Techniken führen zu einer Optimierung der Bearbeitungspraxis, sparen Ressourcen, erhöhen den Ertrag der Landwirte, reduzieren die Bearbeitungskosten und schonen die Umwelt. Die erfolgreiche Einführung von Precision Farming erfordert die Entwicklung von Bearbeitungsgeräten, die die Behandlungsintensität variieren können, und Sensoren, mit denen die räumliche Heterogenität von Wachstumsfaktoren bestimmt werden kann. Solche Systeme können große Datenmengen, die von den Sensoren erfasst werden, aufnehmen, speichern und verarbeiten. Ortsgebundene Messungen führen dann zur Abgrenzung von Bearbeitungszonen innerhalb eines Schlages und werden als Geodaten erfasst und können in Karten visualisiert werden. Das verbesserte, detailliertere Wissen um die Situation innerhalb von Schlägen führt zu neuen, erweiterten Anwendungsbereichen und erlaubt eine präzisere Dokumentation der getroffenen Maßnahmen.

In dieser Arbeit wurde ein Teil von Precision Farming Technogien für die teilschlagspezifische Unkrautbekämpfung entwickelt. Fünf ausgewählte Publikationen werden präsentiert, die technologische Voraussetzungen und Details des entwickelten Systemes abdecken. Unkrautbekämpfung ist eine notwendige Maßnahme in den meisten Kulturpflanzenbeständen. In den meisten Kulturen werden Unkräuter chemisch, mit einer gleichmäßigen Ausbringung von Herbiziden, bekämpft, obwohl diese normalerweise nicht gleichmäßig über einen Schlag verteilt sind. Tatsächlich variiert die Verunkrautung von keiner, geringer bis hoher Dichte. Aufgrund dieser Eigenschaften der Unkrautverteilung können teilschlagspezifische Techniken angewendet werden. Große Mengen Herbizide können eingespart werden, wenn nur die Nester mit hoher Verunkrautung behandelt werden. Eine Reduktion der Aufwandmengen verbessert die Ökonomie und Umweltbilanz. Die Voraussetzung für eine teilschlagspezifische Bekämpfung von Unkräutern ist das Wissen um

die Stärke der Verunkrautung und ihre Verteilung im Schlag. Anhand der Verunkrautung können Entscheidungen über die Art und Menge der auszubringenden Herbizide für jeden Teil des Schlages getroffen werden. Techniken für eine teilschlagspezifische Unkrautkontrolle werden in Weis et al. (2008) vorgestellt. Zusätzlich wurden Ertragsverluste aufgrund von Verunkrautung und der unnötigen Ausbringung von Herbiziden anhand von On-Farm-Research Versuchsdaten modelliert. Anforderungen an Sensoren und Analysetechniken von Systemen zur Unkrauterkennung wurden in Weis and Sökefeld (2010) thematisiert.

Ein Computer Vision System zur Messung der Verunkrautung auf Teilflächen wurde entwickelt. Das System nutzt bispektrale Kamerabilder des roten und infraroten Lichtspektrums als Ausgangsbasis und prozessiert diese Bilder, um die lokale Verunkrautung innerhalb eines Schlages bestimmen zu können. Da es eine große Anzahl Unkraut- und Kulturpflanzenarten gibt, die berücksichtigt werden müssen und deren Erscheinungsbild sich über die Wachstumsstadien und -bedingungen hin ändern, muss das System flexibel sein und sich an die vorkommenden Kulturpflanzen-Unkraut-Gesellschaften anpassen lassen.

Die aufgenommenen Bilder werden dahingehend prozessiert, dass Pflanzen vom Hintergrund (Boden, Mulch, Steine) getrennt werden und Formparameter für jede Pflanze im Bild errechnet werden. In Weis and Gerhards (2007a) wird die Berechnung von verschiedenen Formparametern vorgestellt. Regionenbasierte, konturbasierte und neu entwickelte Parameter, die auf einer Skelettierungsoperation basieren, wurden berechnet. Mit einer Untermenge der Formparameter konnten monokotyle, dikotyle Unkräuter und *Brassica napus* L. von *Hordeum vulgare* L. mit einer Genauigkeit von 98% unterschieden werden. Die Information über die Form wird genutzt, um die Pflanzen als Unkraut oder Kulturpflanze zu klassifizieren. Damit die Variation der Erscheinungsformen in den verschiedenen Wachstumsstadien gehandhabt werden kann, wurde ein Klassenschema definiert und Prototypen für die einzelnen Arten selektiert und katalogisiert. Eine Datenbank wurde erstellt, die alle notwendige Information über die Formparameter und Unkraut-/Kulturpflanzenklassen speichert. Die Formanalyse wird mit Klassifikationsalgorithmen durchgeführt, die anhand der Prototypeninformationen aus der Datenbank trainiert werden können. Alle Formparameter wurden auf ihre Fähigkeit hin getestet, verschiedene Unkrautarten und Kulturpflanzen zu unterscheiden, dabei zeigte sich, dass die neu entwickelten Merkmale gut für eine Klassifikation geeignet sind. Die Klassifikation kann dann auf alle Pflanzenobjekte in den Bildern angewendet werden und weist diesen anhand ihrer Form eine Klasse zu. Verschiedene Klassifikationsalgorithmen wurden in Weis et al. (2009) verglichen und es zeigte sich, dass Support Vector Machines geeignet sind, die Komplexität des gestellten Klassifikationsproblemes handhaben zu können.

Das Ergebnis der Bildverarbeitung und Klassifikation sind Unkrautverteilungskarten, die die Anzahl der Unkraut- und Kulturpflanzen in jedem Bild beinhalten. Diese Unkrautkarten sind die Ausgangsbasis zur Erstellung von Applikationskarten. Entscheidungsregeln, basierend auf einer Aggregation von Klassen, wurden auf die Unkrautkarten angewendet, um Behandlungszonen für jedes Herbizid abzugrenzen. Weis and Gerhards (2009a) legen den Schwerpunkt auf die Erzeugung von Applikationskarten aus den Klassifikationsergebnissen und leiten Entscheidungsfunktionen für die Konvertierung von Unkrautkarten zu Applikationskarten ab. Die automatische Unkrautbonitur weist hohe Korrelationen mit einer manuell durgeführten Bonitur auf.

Die vorhandene Applikationstechnik, eine spezielle Pflanzenschutzspritze mit 3 m Teilbreitenschaltung, kann die Art und Menge von drei Herbiziden gleichzeitig variieren. Daher können drei verschiedene Applikationskarten genutzt werden, eine präzise Herbizidmischung für jede Behandlungszone im Feld zu realisieren. Diese Technik wurde erfolgreich in verschiedenen Kulturpflanzenbeständen (Mais, *Zea mays* L.; Sommergerste, *Hordeum vulgare* L. und Zuckerrüben, *Beta vulgaris* L.) angewendet und weist das Potenzial für eine Einführung in die Praxis auf. Die entwickelte Datenbank und Analysetechnik war die Basis für die Entwicklung eines kommerziellen Sensorsytemes, das bald verfügbar und reif für den praktischen Einsatz beim Landwirt sein wird.

Acknowledgements

I would like to thank all persons that were involved directly or indirectly in the creation of this thesis. First of all thanks go to Professor Gerhards, who supervised and reviewed the work. He used his deep understanding of the topic always in a productive and helpful manner. The contributions of the co-authors were very valuable and led to fruitful discussions, which were not limited to the topics found in the publications. Special thanks go to the colleagues at the institute and research station, namely Christoph Gutjahr, Markus Sökefeld, Helmut Kärcher and Wilhelm Hermann. Without their support the field work would not have been possible in this way. I also thank all other, not explicitly mentioned colleagues for their discussions as well as all friends for their patience.

The development is based on free software components, therefore thanks go to all contributors of free software. This work would have been much more difficult without the mature operating systems (Debian) and tools that are available to everyone.

The work was partially funded by BMELV[1] and BLE[2].

[1] Federal Ministry of Food, Agriculture and Consumer Protection (Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz)

[2] Federal Agency for Food and Forestry (Bundesanstalt für Landwirtschaft und Ernährung)

Contents

List of Figures

List of Tables

1. General introduction

1.1. Outline of the thesis

The thesis is composed of three major chapters: This chapter introduces into the research topic and localises the work in the current Precision Farming research field. Objectives of the work and considerations for the implementation are given. The developed system is presented in 1.5. The analysis steps from the input measurements to the resulting weed and application maps for site-specific herbicide applications are outlined.

Five selected publications are reproduced in chapter 2. They cover the prerequisites and applicable techniques for site-specific weed control and models to assess the effect of weeds and herbicide on the yield (section 2.1). An overview about sensor technology and analysis techniques for weed detection, as they were published in literature, is given in section 2.2. The developed system and results achieved with it are the focus of the publications in sections 2.3, 2.4 and 2.5.

A general conclusion and discussion of the achieved results are given in chapter 3. Additional resources and technical details about the developed software can be found in the appendix. There is one bibliography at the end of this document, which contains also the referenced literature of the appendix.

1.2. The need for weed management

Weed management is an important practise in crop production, since weed infestation in agricultural fields cause high yield losses. Oerke (2006) estimates the worldwide potential loss due to all pests at 40%–80%, varying for different crops, and identifies the potential of yield losses for weeds with 34% to be the highest of all pests.

Weed control strategies include preventive, biological, physical and chemical methods. In Europe chemical weed control plays the most important role in weed management. These methods are very effective and cause lower costs compared to other management practises, such as mechanical weeding. However, herbicide applications impose environmental risks. In Europe, therefore strict regulations exist for the application of herbicides, e.g. the German plant protection law (PflSchG (1986)) requires the farmers to use economic weed thresholds for the management. Integrated pest management (IPM) techniques also try to minimise the undesirable side effects and use of crop protection products by combining different strategies to optimise the results (Radcliffe et al. (2009)). Key factors for the selection of an appropriate strategy are the knowledge about the pests, a thorough monitoring and economic as well as biologic thresholds, which are used for the decision about the type of treatment.

These regulations and recommendations target uniform treatments of a field. However, weeds are not uniformly distributed in agricultural fields, but tend to occur in patches (Dieleman and Mortensen (1999, Gerhards and Christensen (2003)). If the treatments are varied within a field according to the particular local situation, large savings can be accomplished. The so-called site-specific treatments improve, both, the economic and ecological output. Different management zones can be drawn within a field, if the variations are known. Site-specific treatments are the topic of Precision Farming research, of which an overview is given in section 1.3.

One of the concerns regarding site-specific weed management was the question if the partially low or no herbicide input could increase the level of weed infestation and/or their distribution in the subsequent years. Ritter and Gerhards (2007) studied the long term effects of site-specific weed management over a period of eight years, no significant increase in weed intensity due to a site-specific management practise were found. Thus, this management strategy has the potential to be introduced into practise.

In this work a weed detection system was developed to support the site-specific treatment of weeds. One of the basic requirements is a detailed estimation of the varying weed infestation within a field (Wiles (2005)). This can be done manually, which is too time consuming and expensive for application on large areas. Therefore a computer vision system for the automation of the weed sampling was developed, which measures the weed infestation using images taken in the field. An automation increases the spatial sampling ratio and can be applied to different crops due to its configurable analysis. The spatial information on weed infestation allows their monitoring in a map, which is an efficient decision tool for treatment optimisation.

A large number of different species are to be considered. Thus a challenge for the system development was the precise species discrimination. In an arable field, there is usually a high number of weed species, each having a characteristic appearance. But the appearance changes with phenological development and can be affected by growth conditions.

This problem can only be tackled by a large database comprising all these variations of the appearances by finding stable parameters that can be used to differentiate between the species. For the development of such a database, images were taken in the green house and in field experiments, covering the most important weed and crop species at different growth stages. The system is described in more detail in section 1.5 and the publications (chapter 2).

1.3. Precision Farming

The term Precision Farming designates an emerging technical development for farmers and has been developed during the last two decades (Mondal and Tewari (2007)). Precision Farming introduces new aspects to classical management strategies, refining decision and application boundaries spatially. Three major components can be identified in Precision Farming systems: first, sensor technology to measure the variability within a field (Godwin and Miller (2003)), second, an analysis and decision component to derive a management strategy (Kudsk (2008)) and third, variable rate application technology (VRT) to vary the treatment according to the chosen decision. Precision Farming technologies are developed for several application domains in agriculture (Auernhammer (2001)) and combine information technology with variable rate application technology and decision systems.

Precision Farming applications were primarily developed for the following application domains:

Guidance systems: allow the precise steering within the field, avoiding overlapping application areas (Zier et al. (2008)).

Precise sowing: a homogeneous number of seeds, precise aligned seeds (with equal spacing) or the variation of sowing density can be achieved.

Fertilisation: the amount of fertilisers is adapted to the actual nutrition status within the field (Biermacher et al. (2009)).

Plant protection: variation of pesticides (herbicide, fungicide and insecticide) within a field (Miller (2003)).

Soil management: tillage (e.g. ploughing intensity/depth) according to the soil properties (Adamchuk et al. (2004)).

Irrigation: precise irrigation according to the soil water status (Al-Kufaishi et al. (2006)).

Yield mapping: for quality control of the management decisions and yield (Arslan and Colvin (2002)).

Documentation: all taken actions can be documented precisely for each management zone, including the information about the total amount of material and working hours.

This list is not complete, further uses are being researched and developed (Godwin and Miller (2003)). The guidance, fertilisation and yield mapping systems were the first to be available and adopted by farmers (Reichardt et al. (2009)). The work of this thesis refers to Precision Farming for plant protection.

1.3.1. GPS and GIS for Precision Farming

A crucial prerequisite for the Precision Farming is the use of positioning services and the processing of location based data.

GPS Most systems nowadays use GPS (Global Positioning System) receivers to measure the position within the field. Some GPS receivers can be operated in different modes: either the position solution is derived only from satellite signals (GPS mode) or additionally a reference signal of one or more base stations can be used (RTK–real time kinematic modes). The latter system configuration allows to minimise measurement errors and therefore provides a more accurate position (standard deviation of centimetres instead of decimetres/meters). In the conducted experiments an RTK-capable receiver was used. A reference signal, received by a radio modem (450 MHz UHF), was available. All GPS coordinates are geographic (latitude, longitude) and measured in the WGS 84 reference system, which defines an ellipsoid used for measurements around the earth. To be able to work in a metric system, these coordinates (representing angles) have to be projected, e.g. with a transverse mercator projection. The projection used in Germany and mostly throughout this work is called Gauß-Krüger[1], which is an adapted transverse mercator projection, but using a different ellipsoid (Bessel), so that a transformation between the ellipsoids is necessary first. This transformation can be a source of inaccuracies, since several (Helmert) parameter sets exist for the transformation (for global/local levels). These parameters are known with a precision of not more than several decimetres (in projected system). Using different Geographic Information Systems (GIS) software can lead to shifted results, if their parameter sets are not the same. This problem was avoided by keeping the original measurements (in WGS 84) throughout the analysis.

GIS For the processing of spatial data (geodata) geographical information systems are used. With the help of such systems the data can be stored, edited and analysed. Simulations and geostatistical computations can be accomplished within these systems. The data sets and results of computations can be visualised using these systems, which leads to the creation of thematic maps. Frequently used computations are combinations of different data sets by intersection operations and spatial interpolation of variables.

In Precision Farming, GIS is also used to create application maps, which are used by application equipment to vary the treatment, e.g. the herbicide dose and mixture (Fig. 1.9). Precision Farming terminals with the possibility to use GPS and application maps command the equipment in the field. Farm management information systems (FMIS) usually have basic GIS functions included and combine economic, inventory and man-

[1] Gauß-Krüger Projection will successively be replaced by ETRS89/UTM (European Terrestrial Reference System 1989/Universal Transverse Mercator) in the future (Ellipsoid: Geodetic Reference Systems 1980, GRS80).

agement data with the real estates and economic data of a farm. Additionally some vendors of Precision Farming systems provide software with basic GIS functionality, they are at least capable to create maps of the measurements and recommendations. For more advanced GIS analyses (geostatistics, intersections, data combination) additional desktop GIS software has to be used, which requires expertise and training.

1.4. Objectives of this work

A site-specific management of weeds requires the acquisition of the weed infestation on a sub-field level. Several weed species have to be considered for each crop, especially the ones causing high yield losses. Experts sample weeds in the field by looking at a predefined area (e.g. $0.25\,\mathrm{m}^2$) and counting the individual plants or estimating the percentage of weed/crop cover. This is performed on regular grids or with irregular sampling strategies in the field. The experts make use of the morphological properties and the general appearance and shape of the individual plants for the recognition.

A computer vision system for automated weed detection therefore can be oriented at this technique. The distinction of crop and weed species on the plant level requires the measurement of high resolution images with clearly visible plants. With image processing techniques shape features can be derived for each plant in the image to describe its appearance. It is assumed, that a species have a characteristic appearance, which can be measured by such shape features.

The objectives of the work can be summarised as follows:

- build an operational image processing system for the analysis of images taken in weed infested fields, which identifies different species according to their shape, section 1.5
- create an image database for the most abundant weed species in winter cereals (*Hordeum vulgare L., winter barley*, *Triticum aestivum* L.), maize (*Zea mays* L.) and sugar beet (*Beta vulgaris* L.), section 1.5.3
- calculate different shape features and investigate the appropriateness of them for the purpose of crop–weed discrimination in different growth stages, section 1.5.2

- select suitable features as well as supervised and unsupervised classification algorithms for automatic plant species identification, section 1.5.4
- create weed application maps in combination with decision support systems and GIS, section 1.5.4

A weed detection system was developed to address these objectives. The next section gives an overview about it and illustrates the necessary analysis steps to create weed and application maps.

1.5. Weed detection system

The work in this thesis describes a system for the automatic detection of weeds using image processing. A weed detection system requires a certain generality and flexibility for applications under various field conditions. This can only be achieved if the system design allows maintenance, modifications and extensions. Thus the system was designed according to the following criteria:

- Modularity: the system consists of separate modules, which can easily be changed independently from each other. Interfaces for the data storage/exchange between the components have been defined.
- Traceability of the results: all parameters and data are stored, so that the results are transparently connected to the input data.
- Extensibility: new analysis methods can be integrated, extending the system.

To fulfil these criteria, the system was divided into several modules, and interfaces were defined between them:

- Image processing software, computes the shape features from the input images (see Appendix A).
- Classification software, classifies the objects in the images.
- Data storage, a database to hold all processed data.

- GUI (graphical user interface) for the analysis software (see Appendix A.1).
- GIS software to use and modify the resulting map data. OpenJump[2], a free GIS software, with PIROL[3] additions for Precision Farming was used in this thesis.

All developed software was written based on free software components (FOSS[4]). GPL (Gnu Public License, FSF (1991)) and LGPL (Lesser Gnu Public License, FSF (1999)) licensed software and libraries were used. Thus, the components allowed to understand and review its functionality and adapt it to the needs of this work.

Fig. 1.1 illustrates the general layout of the system for site-specific weed management highlighting the internal and external interfaces. Each arrow represents either a software or hardware interface for the data flow. Input are images and their GPS position, measured in the field, as well as decision rules for the weed management with herbicides, output is the herbicide application decision (white boxes). The system implements image processing and classification (upper left) using a knowledge base, which stores the information about plant species and their shape information. The result of the analysis and the position information can be used to derive a decision on a sub-field level and control the herbicide application in the field.

1.5.1. Image acquisition

The automatic sampling system is applied in the field and has to operate under adverse and changing weather conditions and shaking of the equipment. Special bi-spectral cameras take the input images, one image of the red spectrum (R) and the other of the infrared spectrum (IR) of the light. Changes of the illumination, for instance clouds, can cause variations of the image intensities. Robust equipment was chosen and methods were developed, that can adjust for the changes (Sökefeld et al. (2007)). Both images are aligned to each other and show the same scene of the ground. The distance of the cameras to the ground is approximately one meter and they are directed vertically downwards, leading

[2] http://openjump.org/
[3] http://www.pirol.fh-osnabrueck.de/pirol-openjump.html
[4] Free, Open Source Software.

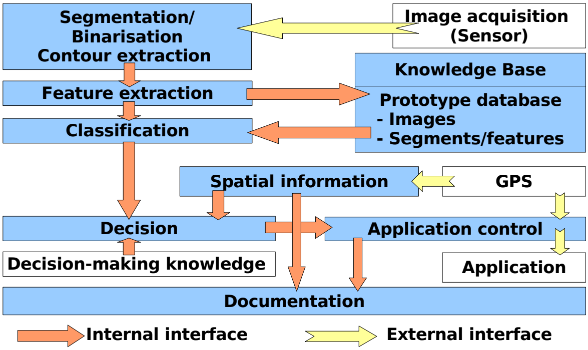


Figure 1.1.: Site-specific system scheme: The system for the detection of weeds and application of herbicides consists of several components that interact with each other. The data flow is denoted by arrows.

to an image of 40×30 cm of the field. All images and the corresponding GPS position are gathered with a field measurement software and stored on a harddisk. The measurement software controls the shutter times of the cameras, adjusting for changing illumination conditions. A difference image (DIFF) is computed from the infrared (IR) and red (R) images (Fig. 1.2). Even in cluttered environments this technique leads to a good separation of the plants: in Fig 1.2 crop residues of maize are visible in the red and infrared images, but disappear in the difference image.

These difference images can be thresholded with a grey level value, separating the foreground (plants) from the background (soil, mulch, stones). In the resulting binary images the foreground pixels are coded with white colour and the background with black colour (as in Fig. 1.7). The binary images can be further preprocessed with morphological operators (Soille (2003)). The basic operations are dilation and erosion with a round element, which enlarge respectively shrink the object at the border. Closing operators (dilation followed by erosion) can be used to connect nearby regions, which were separated by the thresholding, the edges can be smoothed. A closing operator was used in Fig. 1.2 for the segmentation (SEG). Opening operators (erosion followed by dilation) do the opposite: they separate objects at thin connections and very small

Figure 1.2.: Red (R), infrared (IR) and difference (DIFF) images (histogram stretched for print), (SEG) is the segmented image. The difference images (DIFF) is used for the analysis. Plants appear bright in the infrared and dark in the red image, the background components (soil and crop residues) are dark in the difference image. This can be segmented (SEG) into foreground objects (white) and background.

objects (regarded as noise) are suppressed. A segmentation algorithm then groups all foreground pixels, which are surrounded by background pixels (connected components) to segments. These segments are the objects of interest, as they correspond to whole plants or parts thereof.

1.5.2. Shape feature computation

The next step of the analysis is the computation of shape features for each segment to analyse the appearance and compare them with prototypes. Shape features can either be derived from the complete set of object pixels (region-based) or the outer border (contour-based). A variety of features were developed by image processing researchers. For a shape based comparison of different segments these should have the following relevant characteristics:

- independence of the position in the image: the position of a plant within the image is of no importance,
- independence of rotation: a rotated plant should have the same shape features,

- independence of size: congruent shapes should have the same feature representation.

Some of the computed features are not independent of the size, nevertheless they are useful for the discrimination: since the plants are growing, the areasize for example (number of pixels a segment consists of) facilitates to distinguish later growth stages from earlier ones.

Some well known shape features like compactness, central moments, Hu moments (Hu (1962), region-based) and Fourier features (contour-based) were computed. Additionally new features were developed, which are based on a skeletonisation of the segments. The skeleton, which is the central line of a region, is combined with a distance function (Soille (2003, pp. 64 ff)) of the region. The distance transform assigns each pixel of a region a value for its distance to the border of the region. The combination leads to a vector of distance values, which can be interpreted as leaf thickness features. These features are especially useful to discriminate broad-leaved weed from grass leaves, which are narrow and elongated (Weis and Gerhards (2007a), section 2.3).

The discriminant abilities of the newly computed features were compared with ones that were implemented in software of previous research (IMPPAS, Oebel (2006)). In Fig. 1.3 the first two discriminant functions of each software are opposed. The newly developed features lead to a better discrimination between the three classes *Hordeum vulgare* L. (crop), monocotyledonous weeds and dicotyledonous weeds. The clusters for the classes in the feature space are less overlapped (Fig. 1.3 right) and thus provide a better distinction. Table 1.1 contains a confusion matrix for the classification of test data with the discriminant functions of the new features, derived from an independent training set. Monocotyledonous and dicotyledonous species can be separated very well, wrongly classified plants are mostly monocotyledonous weeds classified as *Triticum aestivum* L.. SPSS (SPSS Inc (2004)) was used for the calculation.

Overlaps of objects are a problem especially for monocotyledonous plant species with elongated leaves. These situations can be distinguished by a subset of features, like presented in Fig. 1.4. The data set shown contains non-overlapped and overlapped samples of *Triticum aestivum* L., *Chenopodium album* L. and *Viola arvensis*.

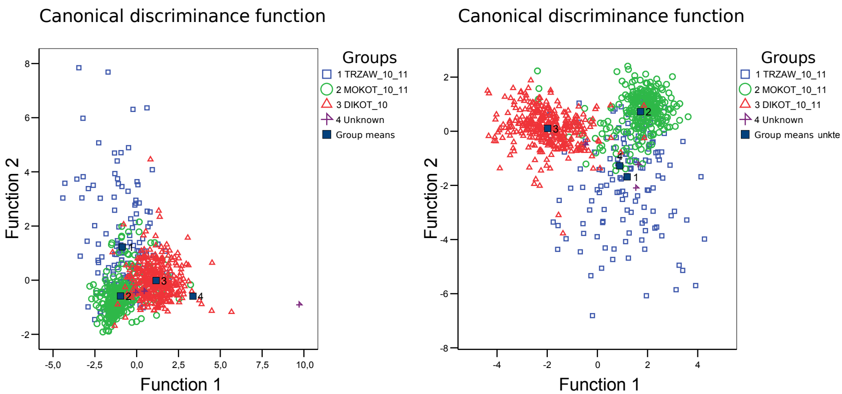


Figure 1.3.: Discriminant functions (first two) for two feature sets. Left: features of previously used software (IMPPAS), right: features of the developed software. TRZAW: *Triticum aestivum* L., MOKOT: monocotyledonous weeds, DIKOT: dicotyledonous weeds. 'Unknown' are samples without class assignment (to be disregarded). The growth stages were BBCH 10&11.

predicted\true	TRZAW_10_11	MOKOT_10_11	DIKOT_10_11
TRZAW_10_11	63.7	27.4	6.2
MOKOT_10_11	7.1	88.9	3.7
DIKOT_10_11	2.3	1.7	95.7

87.2% samples of the test data set were correctly classified

Table 1.1.: Confusion matrix for crop–weed discrimination. The data was split into 50% training data and 50% test data and the discriminance functions of Fig. 1.3 (right) used. TRZAW: *Triticum aestivum* L., MOKOT: monocotyledoneous weeds, DIKOT: dicotyledoneus weeds. The growth stage was BBCH 10&11.

[Erratum: figure was not complete]

Figure 1.4.: Overlap detection using a selection of features: Overlapped and non-overlapped objects in the feature space. The overlapped objects are bigger and thicker due to their complex structure. The features are areasize – size of segment; drear – rearmost distance to centre of gravity, along main axis; skelmax – maximum distance of skeleton to border, representing the maximum thickness of the object.

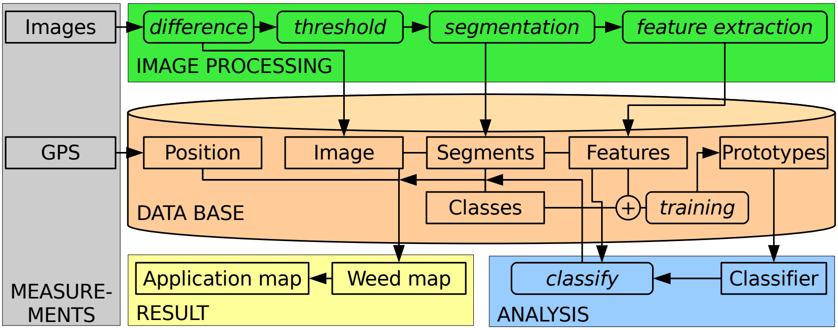


Figure 1.5.: System scheme with data flow in the database: The input data (left) is processed (top) and all data is stored in a database. Training data can be generated from the image segments and is used for a classification (lower right). The resulting weed and application maps can be assembled from the classification data and positions of the images (lower left).

1.5.3. Database

A database is the central storage place for the data analysis. Fig. 1.5 illustrates the data flow of the system. The data generated by the analysis components (image processing and classification) are stored in the database. The resulting map data with the aggregated weed densities can be retrieved from the database and further processed in GIS software to generate application maps. Figure 1.6 shows parts of the relational database model.

The shown relations are implemented by foreign key constraints between tables in the database. a shows the class model, a class is defined by the three parts EPPO, BBCH and segspecial. EPPO denotes the species as defined in the EPPO-Code system (EPPO (2007)), BBCH denotes the phenological growth stage (Hess et al. (1997)) and segspecial is an additional modifier for situations stemming from image processing (e.g. single leaf segmentation or overlapping). More than 250 classes were actually used and created according to the scheme in the database. A subset of species can be grouped to a 'crop–weed community' (Ta-

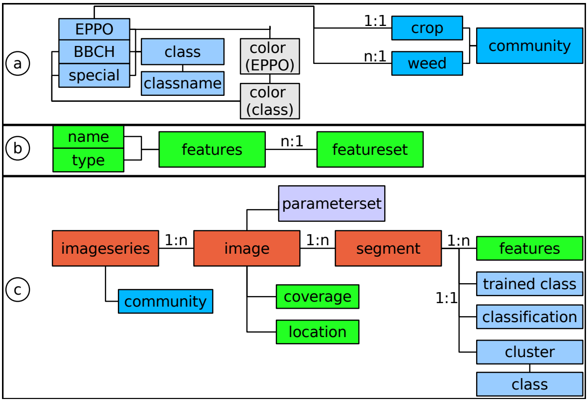


Figure 1.6.: Database layout: The lines represent relations. The colours denote the type of information. Blue: class information, green: feature information, red: image information, grey: colour information. (a): class relations, (b): feature relations, (c): image series relations. A parameter set consists of image processing parameters.

ble 1.2). The concept of communities can be used to define common (problematic) weed species for different crops. Table 1.2 lists several examples of the defined communities in the database. b shows the feature relations: a subset of features can be grouped to feature sets. c are the relations within one of the image series: the file locations of the images belonging to the series are stored. For each image taken in the field the image processing parameters and GPS position are added. The results of the image processing are segments and their shape features, the relations to the source images are preserved. Class information for these segments, together with feature vectors, define the prototypes and provide the training data for the classification. Fig. 1.7 shows examples for prototypes, which were used in Weis et al. (2009) (section 2.5). Training data sets can be composed from the data base, thus the data can be reused for different classifications, e.g. training data obtained from greenhouse images can be used for the analysis of field data.

1.5.4. Classification and weed map creation

A classification assigns classes to all objects in an image series. Classification algorithms can be grouped into two types: supervised and unsupervised classifiers. Supervised classifiers are trained with a training data set that provides samples for the assignment of classes and feature vectors. This type of classifiers was used to create weed maps from the defined prototypes in the database.

Unsupervised classifiers, aka. clustering algorithms, create a separation of the data set solely based on feature vectors, without class information. The resulting clusters group feature vectors according to their distance in the feature space. This technique was used to group similar shaped plants with the computed shape features (Weis and Gerhards (2009b)). If the clusters represent the desired classes, these can be mapped to each other. A fully automated classification was performed this way on a field data set of a winter wheat crop (*Triticum aestivum* L.). The resulting clusters separated dicotyledonous and monocotyledonous weeds and could be mapped to species (*Alopecurus myosuroides* Huds. and *Veronica persica* Poir.). The system provides clustering algorithms for the selection of prototypes. A clustering of a data set can be used as starting point to assign classes to objects of similar shape (Appendix, Fig. A.6). The predefined classes can be assigned to the clusters, and

Table 1.2.: Crop–weed communities, as they are defined in the database (see also Fig. 1.6). These are used for the selection of prototypes. The five letter acronyms are the EPPO-codes, their Latin names are given below the table.

Crop	Monocotyledonous weeds	Dicotyledonous weeds
HORVS	ALOMY APESV AVEFA BROST POAAN	ATXPA CENCY GAETE GALAP CAPBP LAMAM LAMPU MATIN MATCH PAPRH POLAV POLCO POLLA SENVU SINAR SONAR STEME VERPE VIOAR THLAR CIRAR CONAR
ZEAMX	AGRRE HORVS POAAN TRZAW DIGSA ECHCG SETVI	AMARE CHEAL CIRAR CONAR GALAP LAMPU POLAV POLCO SOLNI SONAR STEME VERHE VERPE
BEAVP	AGRRE POAAN	ATXPA CHEAL CIRAR CONAR GALAP GERDI LAMPU POLAV POLCO SONAR STEME VERHE VERPE

AGRRE: *Agropyron repens) (L.) P.Beauv. (Elymus repens (L.) Gould)*, ALOMY: *Alopecurus myosuroides* Huds., AMARE: *Amaranthus retroflexus* L., APESV: *Apera spica venti* L. Beauv., AVEFA: *Avena fatua* L., ATXPA: *Atriplex patula* L., BEAVP: *Beta vulgaris* L., BROST: *Bromus sterilis* L., CAPBP: *Capsella bursa-pastoris* (L.) Medik., CENCY: *Centaurea cyanus* L., CHEAL: *Chenopodium album* L., CIRAR: *Cirsium arvense* (L.) Scop., CONAR: *Convolvulus arvensis* L., DIGSA: *Digitaria sanguinales* (L.) Scop., ECHCG: *Echinochloa crus-galli* (L.) P.Beauv., GAETE: *Galeopsis tetrahit* L., GALAP: *Galium aparine* L., GERDI: *Geranium dissectum* L., HORVS: *Hordeum vulgare* L., LAMAM: *Lamium amplexicaule* L., LAMPU: *Lamium purpureum* L., MATCH: *Matricaria chamomilla* L., MATIN: *Matricaria inodora* L., PAPRH: *Papaver rhoeas* L., POAAN: *Poa annua* L., POLAV: *Polygonum aviculare* L., POLCO: *Polygonum convolvulus* L., POLLA: *Polygonum lapathifolium* L., SENVU: *Senecio vulgaris* L., SETVI: *Setaria viridis* (L.) P.Beauv., SINAR: *Sinapis arvensis* L., SOLNI: *Solanum nigrum*, SONAR: *Sonchus arvensis* L., STEME: *Stellaria media* (L.) Vill./Cyr., THLAR: *Thlaspi arvense* L., TRZAW: *Triticum aestivum* L., VERHE: *Veronica hederifolia* L., VERPE: *Veronica persica* Poir., VIOAR: *Viola arvensis*, ZEAMX: *Zea mays* L.

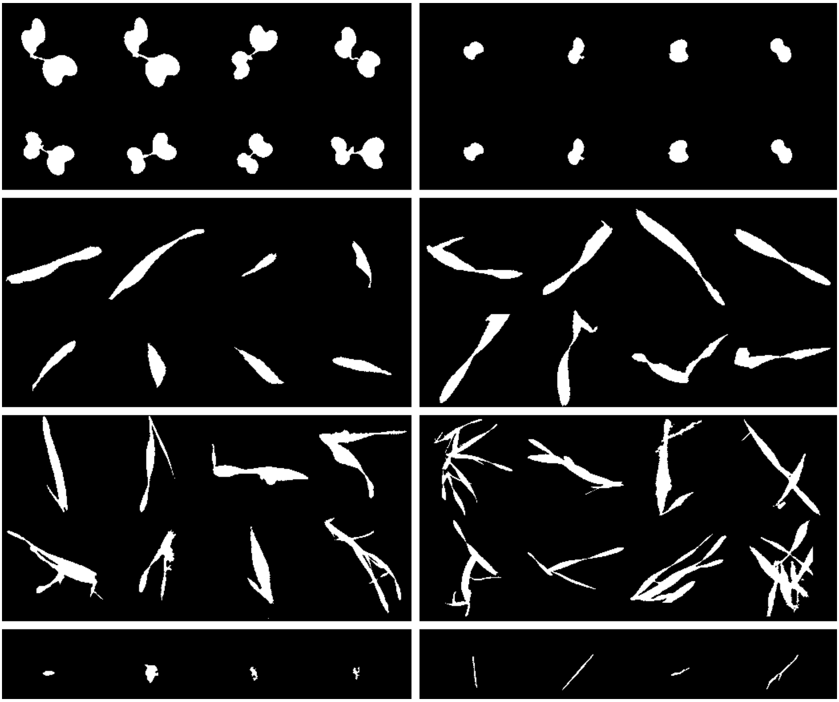


Figure 1.7.: Training data examples from the database (from top to bottom), according to the developed class naming scheme:
BRSNN10N, BRSNN10L;
HORVS10N, HORVS12L;
HORVS12N, HORVS12O;
NOISE00X, NOISE00L.
BRSNN: *Brassica napus* L., HORVS: *Hordeum vulgare* L., NOISE: compact (X) and elongated (L) noise objects; the numbers denote the phenological stage on BBCH scale, N denotes whole plants, L denotes single leaves, O overlapped objects. These examples belong to the paper in section 2.5.

suitable prototypes are then selectable by a 'belongs/does not belong to the class' decision. Thus the creation of prototypes is simplified and difficult-to-separate classes can be identified, if they are grouped by the clustering.

The classification results (class resp. cluster assignments) are also stored for each segment and can then be assembled to result data sets and maps. Coverage maps contain the overall plant coverage (in percent and pixel) as attribute for each image. The resulting weed maps contain the total number of objects assigned to each class for every image, resulting in a point layer with class attributes. The overall coverage level and measured weed densities for two species (*Cirsium arvense* (L.) Scop. and *Convolvulus arvensis* L.) are shown in Fig 1.8 for a field with a maize (*Zea mays* L.) crop. Additionally manual measurements were made and the nests with high infestation levels for these two species were mapped. From these weed maps an application map was created (Fig. 1.9). A square grid is the base geometry for the application maps, covering an area of 3.3 ha. In Fig. 1.9 a grid of 6×6 m was used, the data of the point layer was intersected with the grid cell layer. Each grid cell represents a small management zone, for which application decisions are derived.

The interpolation was done using a GIS, mean values of the weed densities of all points located in a grid cell were computed. According to the scheme in Weis and Gerhards (2009a) (section 2.4), the weed densities of the classes were aggregated to new attributes containing all weed species, that are targeted by one of the herbicides. For the application map in Fig. 1.9 three classes of one species, as defined in the training and classification step, were combined: *Cirsium arvense* (L.) Scop. classes were CIRAR14N, CIRAR15N and CIRAR00L for whole plants in growth stages BBCH 14&15 and any single leaves. The densities of each class are termed ω and represent the number of class assignments per image. The leaf class (CIRAR00L) density $\omega_{CIRAR00L}$ was weighted with a factor 0.3 in a linear equation, leading to a new attribute ω_{CIRAR}:

$$\omega_{CIRAR} = \omega_{CIRAR15N} + \omega_{CIRAR14N} + 0.3\omega_{CIRAR00L}$$

On these aggregated values decision rules can be applied to transform the weed map into an application map. The decision rules contain threshold values for no, low, medium and high weed infestation. Since the control of *Cirsium arvense* (L.) Scop. is done using a herbicide which not varied in the amount, only one threshold was used to create the map in Fig. 1.9. The application map can be used in the field to control the herbicide application. 70% of the herbicides can be saved using this approach, since only 30% of the field are to be sprayed (green areas), although some of the uninfested grid cells are sprayed due to false alarms (false positives). An area of 0.8 ha was selected to be sprayed outside of the nests (green areas not covered by manual measurements). The comparison with the manually measured infestation shows, that nearly all nests with high infestation are covered by the application, only some smaller areas were not covered (false negatives). These make up 0.3% of the application map area and 8% of the manually measured nests. These parts were missed, since there were no images taken in that area respectively there were no weed plants visible in the images of these field parts.

Fig. 1.10 shows the application software for the sprayer. The amount of three herbicides can be controlled on each of the seven 3 m wide boom-sections by a variation of the water pressure. The software uses GPS positions to derive the amount of herbicide for each boom-section from the map and controls the sprayer via a serial interface connection.

The CERBERUS® sprayer (Fig. 1.11) is the result of a prototype development and can vary the amount of three herbicides individually. Three independent water circuits exist. The seven boom sections can be controlled individually for each of the circuits, resulting in a precise application for each of the herbicides.

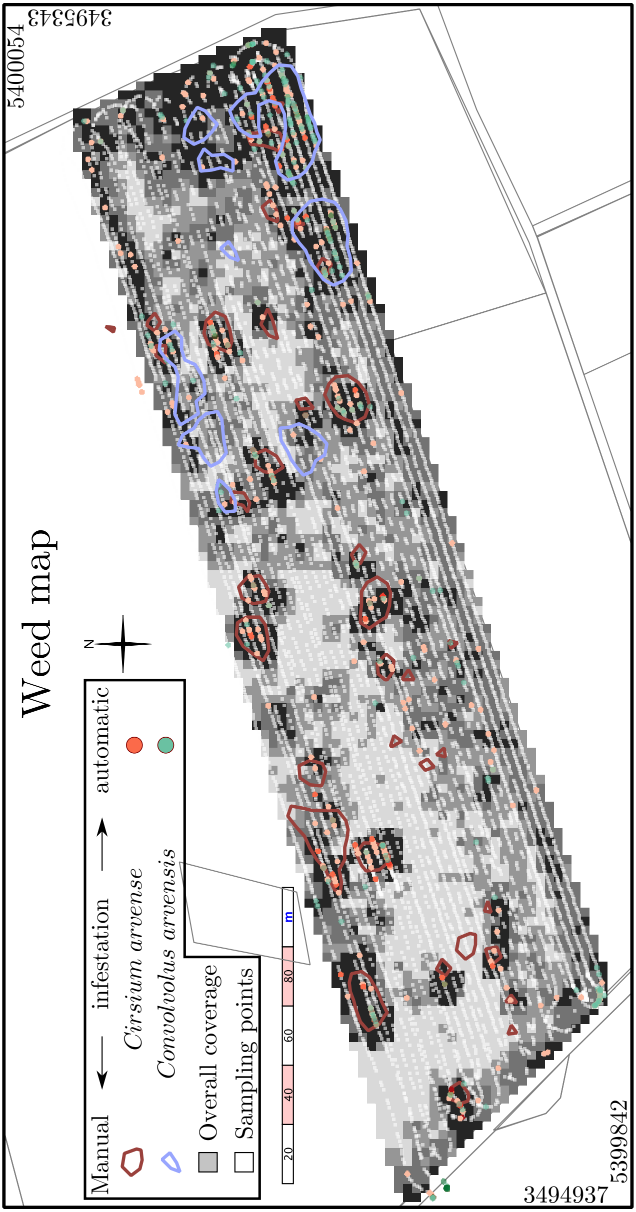


Figure 1.8.: Weed map, the dots are the results of the automatic weed sampling, the polygons were measured manually. Additionally the overall coverage and sampling points are shown. The coordinates are given in Gauß-Krüger projection (DHDN Zone 3).

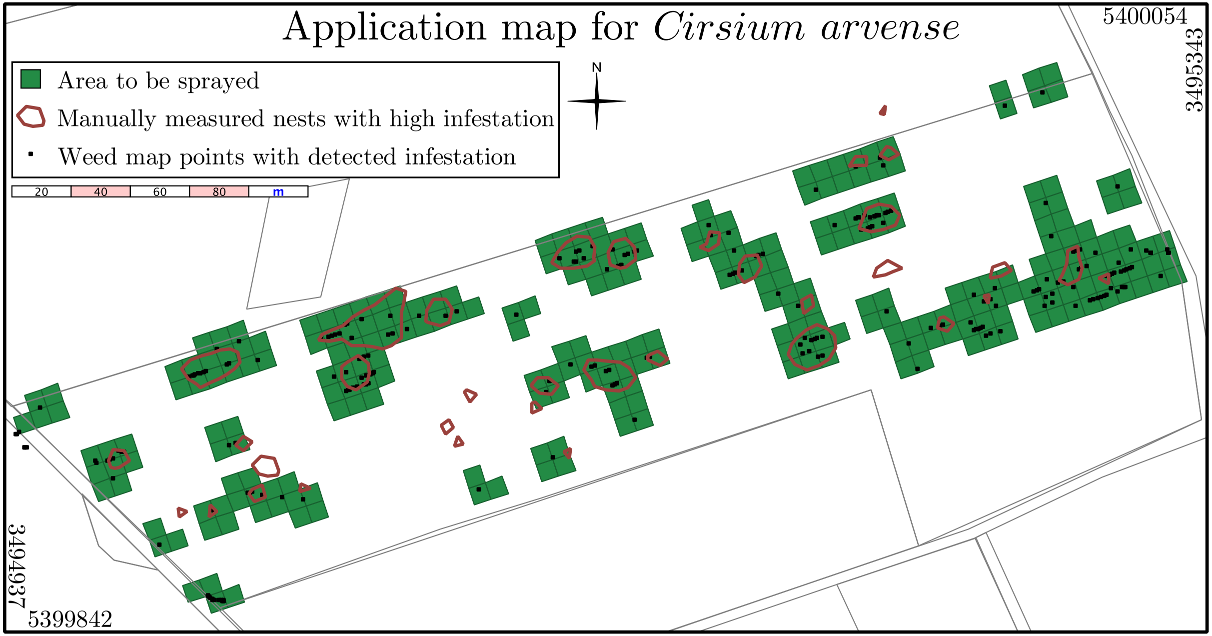


Figure 1.9.: Application map for *Cirsium arvense* (L.) Scop., generated from the weed map in Fig. 1.8. The weed map values were interpolated to a grid of 6×6 meters with a 3 m buffer. The coordinates are given in Gauß-Krüger projection (DHDN Zone 3).

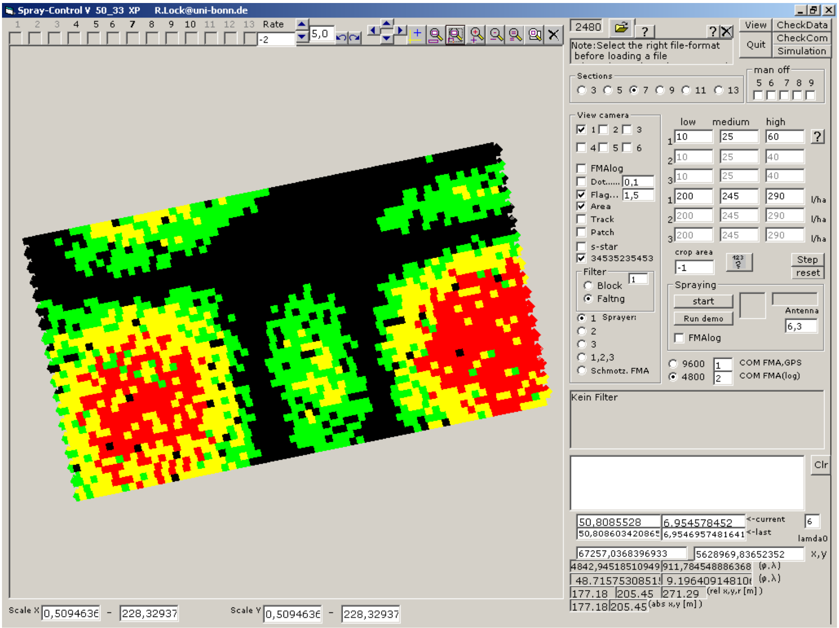


Figure 1.10.: Spray control software: this screenshot shows the application control software for the sprayer. The application map is read according to the GPS-position and controls the amount of herbicide. Three herbicides can be used simultaneously with the sprayer (Fig. 1.11). Red/yellow/green/black areas: high/medium/low/no amount of herbicide.

Figure 1.11.: Sprayer with seven boom sections and three separate herbicide tanks (CERBERUS®).

2. Publications

The publications related to the work are listed as follows:

Reviewed publications: Ritter et al. (2008, Weis et al. (2008, Weis and Sökefeld (2010)

Conference proceedings: Sökefeld et al. (2006, Weis and Gerhards (2007c, Weis and Gerhards (2007b, Weis (2007, Ritter et al. (2007, Weis and Gerhards (2007a, Gutjahr et al. (2008, Gutjahr et al. (2008, Weis and Sökefeld (2010, Weis and Gerhards (2009b, Weis et al. (2009, Weis and Gerhards (2009a) Sökefeld et al. (2006, Weis and Gerhards (2007c, Weis and Gerhards (2007b, Weis (2007, Weis and Gerhards (2007a, Gutjahr et al. (2008)

Other: Weis (2009)

In this chapter a selection of the papers are reproduced, as summarised in table 2.1. The first section 2.1 introduces into the techniques, which can be used for precision weed management (Weis et al. (2008)). The prerequisites for site-specific weed management are reviewed and the system for automated weed detection is outlined. Decision rules for this kind of weed management were developed and applied to field data.

A general overview of the related literature for automated weed detection is given in section 2.2 (Weis and Sökefeld (2010)). This publication reviews the approaches found in the literature and introduces into the sensor technology which can be used to derive weed infestation measures.

Section 2.3 contains a publication (Weis and Gerhards (2007a)) introducing the developed system and the image processing in more detail. A general outline of the developed system is given and results for a field experiment in *Hordeum vulgare* are presented.

Section 2.4 focuses on the decision component (Weis and Gerhards (2009a)). This publication details the creation of application maps from the weed maps, which are generated by the classification.

Finally advances in the classification step are the topic of section 2.5 (Weis et al. (2009)). Since the classification is one of the crucial steps for the results, this part of the system was looked into on more detail. Different classification algorithms were applied to a dataset and compared to each other. A varying complexity of the class definition in the input data lead to the conclusion, that Support Vector Machines (SVMs) are a suitable algorithmic solution to the weed classification problem.

Please note that only the abstracts of the publications are included in the electronic version due to the coprights of the publishers.

Table 2.1.: Overview of chapter publications: bibliographic entries

2.1—Weis et al. (2008) Weis, M., C. Gutjahr, V. Rueda Ayala, R. Gerhards, C. Ritter, and F. Schölderle (2008, Dec). Precision farming for weed management: techniques. *Gesunde Pflanzen 60*, 171–181.

2.2—Weis and Sökefeld (2010) Weis, M. and M. Sökefeld (2010, August). *Precision Crop Protection* (1 ed.)., Chapter Detection and identification of weeds. In: Oerke, E.-C., R. Gerhards, G. Menz, and R. A. Sikora (Eds.) Heidelberg, Germany: Springer Verlag, in press.

2.3—Weis and Gerhards (2007a) Weis, M. and R. Gerhards (2007a, June). Feature extraction for the identification of weed species in digital images for the purpose of site-specific weed control. Stafford, J. (Ed.) (2007, June). *Precision agriculture '07*, Volume 6, The Netherlands. 6th European Conference on Precision Agriculture (ECPA): Wageningen Academic Publishers, pp. 537–545.

2.4—Weis and Gerhards (2009a) Weis, M. and R. Gerhards (2009a, July). Automatic derivation of weed densities from images for site-specific weed management. In C. Lokhorst and R. d. L. J.F.M. Huijsmans (Eds.), *JIAC2009 Book of abstracts*, Wageningen, Netherlands, pp. 349–354. ECPA (European Conference on Precision Agriculture): Wageningen Academic Publishers. Only available on CD-rom (PDF-file) - Proceedings of the Joint International Agricultural Conference, 6.-8. July.

2.5—Weis et al. (2009) Weis, M., T. Rumpf, R. Gerhards, and L. Plümer (2009, August). Comparison of different classification algorithms for weed detection from images based on shape parameters. Zude, M. (Ed.) (2009, August). *Image analysis for agricultural products and processes*, Volume 69 of *Bornimer Agrartechnische Berichte*, Potsdam-Bornim. CIGR, ATB - Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. 15. Workshop Computer-Bildanalyse in der Landwirtschaft, 27–28 August 2009, pp. 53–64.

2.1. Precicion farming for weed management: techniques

Abstract (german) Teilschlagspezifische Unkrautbekämpfung hat in den letzten Jahren zunehmendes Interesse im Bereich der Präzisionslandwirtschaft gefunden. Die Bekämpfung von Unkräutern auf Teilflächen innerhalb eines Schlages erfordert die Messung der unterschiedlichen Unkrautdichten. Entscheidungsmodelle helfen bei der Auswahl und der Steuerung der Maßnahmen abhängig von der tatsächlichen Unkrautsituation. Die Unkrautbekämpfung kann entweder mittels Herbiziden oder mechanisch erfolgen. Eine teilschlagspezifische Herbizidapplikation kann einen Großteil der Herbize einsparen. Mechanische Unkrautbekämpfungstechnik, die auf die Verunkrautungssituation abgestellt ist, kann in einem weiten Spektrum an Kulturen angewendet werden.

Teilschlagspezifische Techniken für die Identifizierung und Bekämpfung von Unkräutern werden vorgestellt. Ein System für die Differenzierung von Unkräutern und Kulturpflanzen mittels Bildanalyse kann Unkrautkarten automatisch erstellen. Modelle zur Beschreibung der Auswirkungen der Unkräuter auf den Ertrag werden entwickelt und in On-Farm-Research Versuchen angewendet. Ökonomische Schadschwellen werden abgeleitet und können für eine Herbizidapplikation mit einer auf Teilflächen steuerbaren Spritze umgesetzt werden.

Keywords: teilschlagspezifische Unkrautkontrolle, Unkrautkartierung, chemische Unkrautbekämpfung, mechanische Unkrautbekämpfung, Expertensysteme

Abstract (english) Site-specific weed control techniques have gained interest in the precision farming community over the last years. Managing weeds on a subfield level requires to measure the varying density of weeds within a field. Decision models aid in the selection and adjustment of the treatments depending on the weed infestation. The weed control

Originally published as Weis et al. (2008)

can be done either with herbicides or mechanically. A site-specific herbicide application technology can save large amounts of herbicides used. Mechanical weed control techniques adapting to the weed situation in the field are applicable to a wide spectrum of crops.

Site-specific techniques for the detection and management of weeds are presented. A system for the discrimination of different weed species and crops from images is described, which generates weed maps automatically. Models for the yield effect of weeds were developed and applied in On-Farm-Research experimental setups. Economic weed thresholds are derived and used for a herbicide application with a patch sprayer.

Keywords: site-specific weed control, weed mapping, chemical control, mechanical control, expert systems for weed control

2.2. Detection and identification of weeds

Abstract This section reviews the approaches for the automation of weed detection. Site-specific plant protection needs to address the varying weed infestation, but the automation is only partially solved and research is still ongoing. The properties for plant species distinction as well as approaches that use them are presented. The focus is on image based methods, of which an example is given.

Originally published as Weis and Sökefeld (2010)

2.3. Feature extraction for the identification of weed species in digital images for the purpose of site-specific weed control

Abstract Automated weed detection and classification allow a high spatial density of measurements and can therefore be used for site-specific application of herbicides in variable rates.

A system for the detection and classification of different crops and weed species is presented. Near range images were taken with a bispectral camera (IR+VIS) mounted on a vehicle at a speed of about 8 km/h. The techniques used analyse the images including preprocessing steps to reduce noise and to obtain comparable results, even under the influence of different image qualities. A segmentation of green plants and background is achieved by binarisation.

The shapes of all plants were extracted and shape parameters, contour and skeleton features were calculated. The features were used to classify different weed and crop species. Their discriminant abilities were tested using data mining and classification algorithms, including discriminant analysis. Different feature sets were compared to each other and the most promising were selected for classification. The classification of an image series taken in a field with *Hordeum vulgare* in 2006 resulted in a correct classification of 98%.

Additionally an image database with weed and crop samples was created, which can be used as prototypes to set up and test different evaluation approaches. This database helps to develop new approaches and makes them comparable to each other.

Keywords: precision weed management, digital image analysis, shape analysis, patch spraying, weed mapping, discriminant analysis

---

Originally published as Weis and Gerhards (2007a)

2.4. Automatic derivation of weed densities from images for site-specific weed management

Abstract Site-specific herbicide applications can save large amounts of herbicides and improve management practices. One crucial part of a system for site-specific weed management is the measurement of the spatial variability of weed densities. A system was developed to identify different weed species from images taken in the field. The automation has the potential to increase the spatial density of weed sampling points. A manual sampling with high density of points is unfeasible due to the costs. Image processing algorithms are used to generate a shape description for each plant in the image. A classifier can be constructed that assigns weed and crop classes to the plants based on the shape features. Weed density maps are generated using the results of the classification. The weed maps are transformed to application maps, which are used for the site-specific herbicide application.

The shape of the plants vary with their growth stage and may be segmented into parts, e.g. single leaves, in the image processing. Therefore different classes for each species need to be introduced into the process. To avoid over- or underestimation of the actual number of weeds some of the classes are aggregated using weight factors. To derive transformation functions the results of the automatically derived weed counts are compared to manual measurements of weed densities, which were derived from the images and in the field. The results show that the raw classification results are linearly related to the actual number of manually counted weeds and the results can be used as input for a site-specific decision component.

Keywords: precision weed management, digital image analysis, shape analysis, patch spraying, weed mapping

Originally published as Weis and Gerhards (2009a)

2.5. Comparison of different classification algorithms for weed detection from images based on shape parameters

Abstract Variability of weed infestation needs to be assessed for site-specific weed management. Since manual weed sampling is too time consuming for prctical applications, a system for automatic weed sampling was developed. The system uses bi-spectral images, which are processed to derive shape features of the plants. The shape features are used for the discrimination of weed and crop species by using a classification step.

In this paper we evaluate different classification algorithms with main focus on k-nearest neighbours, decision tree learning and Support Vector Machine classifiers. Data mining techniques were applied to select an optimal subset of the shape features, which then were used for the classification. Since the classification is a crucial step for the weed detection, three different classification algorithms are tested and their influence on the results is assessed. The plant shape varies between different species and also within one species at different growth stages. The training of the classifiers is run by using prototype information which is selected manually from the images.

Performance measures for classification accuracy are evaluated by using cross validation techniques and by comparing the results with manually assessed weed infestation.

Keywords: weed mapping, weed detection, digital image analysis, shape analysis, feature selection, classification, supervised learning

Originally published as Weis et al. (2009)

3. General discussion and conclusion

In this work, a system was developed and implemented for the automated weed detection in agricultural fields. The developed system uses red and infrared camera images and GPS positions as input and determines the weed infestation level by image processing and shape based classification. The infestation is measured on the plant level by class assignments for each visible plant in the image. Previous research found the shape based approach to be applicable in greenhouse experiments under controlled lighting conditions (Woebbecke et al. (1995)) and in the field (Pérez et al. (2000)). If RGB (red, green, blue) images were used, additional colour features could be computed for the discrimination. Woebbecke et al. (1995) did not find improvements of the classification with colour, but Åstrand and Baerveldt (2002) included colour features successfully in sugar beet (*Beta vulgaris* L.). This improved the classification result from 86% (shape features only) over 92% (color features only) to an overall classification rate of 97%. Burks et al. (2000) additionally proposed colour texture features to discriminate weed species. The analysis was successfully applied to field images with artificial lighting. In cluttered situations, as they can be found in grassland images, RGB colour features Gebhardt and Kühbauch (2007) were found to be useful for the segmentation of *Rumex obtusifolius* L.. The segmentation and shape based approach of this work does not lead to results in these cases, since the plants in grassland are overlapping. Since changes in the illumination lead to large variations of the colour features, the colour image approach should be applied in controlled illumination conditions. The contrast of the red and infrared reflection is higher than between red and green spectra, as measured by standard camera equipment (Sökefeld et al. (2007, Weis and Sökefeld (2010)). Therefore a high image quality was achieved in the presented work, leading to stable segmentation results. This is a necessary prerequisite for the shape based discrimination, since variations of the objects due to differences of the illumination are avoided.

Difference images of the aligned infrared and red images were the input for the image analysis component of the system. The modularity of the developed system allows the exchange of software components, thus providing a flexible framework for the research. It can be extended to use other input data, e.g. the image processing software already is capable of colour image segmentation based on EGI (excess green index, Rasmussen et al. (2007)) or a HSV (hue, saturation, value) colour space transformation (Appendix A). A database was created to store the information about the images, the corresponding processing and shape parameters in a unified way, providing interfaces for all software components. Images were taken in the greenhouse and within different crops in the field. At the time of writing the database contained 360.000 images, providing a thorough basis for the research. The different plant species found in the images have to be trained by prototypes, which were defined according to a class scheme that includes the species and phenological stage information. The database allows the flexible selection and reuse of training data for the analysis with classification algorithms.

The shape based approach has proven its applicability for the weed detection, but also has limitations. Due to the inherent properties of photo-optical measurements, the 3-dimensional space is measured by a 2-dimensional image by central projection along the optical path. This leads to variations of the appearance of objects, if they have a different orientation towards the lens of the camera. If the leaves of plants are twisted, rolled or bended, their appearance in the image and also the shape features are changed. Since the objects of study are biological entities and the growth can vary due to the outer conditions, plants of the same species can have a variety of appearances, even if they are in the same growth stage. Especially in late growth stages plants have a complex structure, which leads to a different appearance for each individual. The complexity of the weed detection, due to the large number of species and their varying appearance, was dealt with by the database approach.

There are species with very similar appearance: some of the grass weeds for example cannot be distinguished only according to their shape, even experts have to use additional properties to identify them, which are not visible in the image (e.g. a reddish colour of the stem, form of auricles and ligule, hairs, . . .). Dicotyledonous weeds in an early (two-leaf, germination) growth stage (BBCH 10–11) also have similarities, which makes it difficult to exactly identify them without additional information. The weed species can be grouped into classes of similar appearance for the automated detection. The basic classes, monocotyledonous and dicotyledonous weeds, can be separated very well using the shape based approach, as the publications indicate (Weis and Gerhards (2007a), section 2.3 and Weis et al. (2009), section 2.5). This can even be achieved by unsupervised classification (clustering) without any training information, if the number of different species is low (Weis and Gerhards (2009b)).

Each herbicide targets several weed species, which are sensitive to it. Usually one herbicide to control monocotyledonous weeds is combined with a second one to manage dicotyledonous weeds. Often one special weed species has to be controlled by a third selective herbicide. Thus the grouping of the weed species should be done according to the herbicides in use, leading to a correct decision even in cases of misclassification between weeds, which are controlled by the same herbicide. The proposed classification approach can be refined with this prior knowledge and might improve the classification especially for difficult-to-separate groups.

The image processing approach assumes, that the plants can be separated in the image, but overlapping plants and parts thereof result in complex, combined objects. These cannot be distinguished using the shape information, although it is possible to analyse, that overlapping occurred. Overlapping often occurs with monocotyledonous plants due to their elongated leaves, leading to large, complex objects. Some of the computed features, among others the newly developed skeleton features, are suitable to identify them as being overlapped. The overlap problem is worse in later growth stages. Within a completely 'closed' canopy this approach cannot be used. Since the application of herbicides takes place, while the plants are in an early development stage, the overlapping problem is regarded to be of minor importance in practise.

Other image processing techniques and features might solve the overlapping problem, but they require a more complex model and increase the computation time, which is a critical factor for online systems. In Weis and Gerhards (2007b) a structural approach was proposed, that tries to solve the problem by high-level vision algorithms in a top-down and bottom-up approach. It was recommended to segment an object into parts and model the plant as complex, hierarchical object consisting of these parts (top-down). Plant parts, which were already split in the thresholding step, can be aggregated with a bottom-up approach in a similar way.

A somewhat simpler approach can be used to improve the results in row crops (Burgos-Artizzu et al. (2009)). Crop row detection can be used to weight the type of plant species according to their position: since no crop is expected between the rows, any plant growing there must be a weed (Åstrand and Baerveldt (2004)). This approach requires, that the crop rows can be identified within the image. Nevertheless a species discrimination is necessary for a precise management with specialised herbicides. On the one hand this is difficult in highly infested areas, since a high coverage can obfuscate the direction of the crop-rows, on the other hand the field of view of the camera should fit to more than one row. If the latter requirement leads to a lower resolution of the image, the possibilities for the shape based detection can be affected. The development of in-row mechanical weeding equipment needs the precise position of plants as input to be able to only affect weeds and spare the treatment of the crop plants. The sensor and analysis technique used in this work can be adapted to identify single crop plants, the additional use of location constraints can also lead to significant higher detection rates.

There are other benefits and application areas of the sensor and analysis technology developed for these studies. Several spectral indices were proposed to be computed from the infrared and red spectrum of the light, which correlate with plant health, nutrition level and leaf area index (e.g. NDVI, Shafri et al. (2006)). As the camera images have a very high spatial ground resolution compared to remote sensing systems, plant health might be assessed on a higher scale, even for single plants. Such systems might require a more thorough observation of the environmental atmospheric and lighting conditions for a normalisation of the measured intensities.

The cameras and image analysis were successfully applied in the greenhouse for dose response experiments. A determination of the biomass can be done from the difference images, since the amount of phytoactive, living plant parts can be assessed. Preliminary results correlated very well with the dry matter, which is tedious and time consuming to measure by hand. Since the measurements are non-destructive, time-series for the response of plants to herbicide applications can easily be measured.

More experiments are to be conducted to prove the general applicability of the system for this purpose.

The ongoing development of sensors and processing power leads to more diverted, specialised systems and applications. Sensor fusion approaches combine different sensor technology and data types (e.g. 2D/3D), thus improving the foundation for new analysis techniques (Klose et al. (2008, Piron et al. (2009)). The fusion of different data sets can improve the results, if the weaknesses of one sensor can be overcome with additional information of another. Especially the robotics research targets the sensor fusion, and the developed technology and analysis approach can be one part of the data acquisition for a precise guidance of the equipment. An integration with other sensor data can improve the robustness of the analysis and optionally widens the scope of the application to other Precision Farming fields. Sensor technology, which has been used by researchers for weed detection, were reviewed in Weis and Sökefeld (2010) (section 2.2), but further research is necessary to elaborate the best combinations.

The results of the presented analysis were targeted at mechanical as well as chemical weed control. Most application maps were created for herbicide applications due to the potential of high savings (Biller (1998, Gerhards et al. (2002, Oebel and Gerhards (2005)). In wheat (*Hordeum vulgare*), grass-weeds were often controlled by ACCase-inhibitors (e.g. Fenoxapropethyl), applied post-emergent. Annual broad-leaved species were controlled by a different herbicide (e.g. ALS-inhibitor or Auxin-like herbicide) and special weeds like *Galium aparine* L. or *Agropyron repens) (L.) P.Beauv. (Elymus repens (L.) Gould)* required the use of a third herbicide, such as Fluroxypyr (*Galium aparine* L.) of Meso-+Iodo-Sulfuron (*Agropyron repens) (L.) P.Beauv. (Elymus repens (L.) Gould)*). All three herbicides need to be applied site-specifically according to the spatial variability of weed-species. The three-tank GPS-controlled sprayer Cerberus®, which was used for site-specific weed control (Gutjahr and Gerhards (2010)), simultaneously realised three spray maps loaded on board of the computer of the sprayer.

Tank-mix applications—as they are often used for uniform weed control strategies—could be avoided with the presented site-specific approach of weed control and therefore, herbicide savings were higher than for a simple site-specific application with only one spray tank. In addition to that, negative impacts of herbicides on the crop were lower with this approach (Gutjahr et al. (2008)). A similar strategy was realised in maize (*Zea mays* L.), sugar beet (*Beta vulgaris* L.) and summer barley (*Hordeum vulgare* L.).

Precision weed management strategies—as they are described in this thesis—also help to maintain the efficacy of chemical weed control methods. Since new herbicide targets have not been discovered during the past 20 years (Zwerger and Ammon (2002)), many weed populations were selected, which express a resistance against herbicide. The number of resistant populations increase with higher selection pressure due to herbicide use. Therefore, precision weed management techniques can be expected to lower the selection pressure and slow down the spread of herbicide-resistant weed populations. High density weed patches can also be controlled with a combination of chemical, preventive and physical practises, which again would decrease weed competition and the risk for herbicide resistance.

For now the system uses an 'offline' approach: the field measurements are analysed in the lab and application maps are generated from the result. These maps are used in a second step to apply herbicides site-specifically using a suitable sprayer, which can vary the amount of three different herbicides. This setup allows to precisely estimate the total amount of needed herbicides before the application takes place, thereby reducing the amount of remaining herbicides in the sprayer to a minimum, but requiring to go two times over the field.

A combination of offline and online approach can be used to aid the classification and decision process. Historical maps provide data about the previous weed infestation and can be considered for the decision about the necessity of herbicide application. Since weed patches are spatially stable, a value for the possible weed infestation can be estimated from weed maps of previous years. Researchers have used such historical weed maps successfully to apply herbicides site-specifically (Christensen and Heisel (1998, Mortensen (2002)). The developed system creates the data for these approaches and can be a valuable tool for population dynamics research.

A product for the introduction to agricultural practise would have to use an 'online' approach, analysing the images in the field, deriving a decision and immediately controlling the application equipment. A stand-alone camera and processing unit is under development, inspired by this work. The weed sensor system fulfils the prerequisites to be operated in the field. That system will solve some of the technical aspects like robustness of the equipment and necessary computing power for the real time image processing and classification. Interfaces for the seamless integration with existing farm equipment ensure the usability and flexibility of the system. Precision Farming equipment nowadays can be connected via ISOBUS interfaces, allowing the control of the equipment and data transfer between the components. Such standardised components can be used interchangeably and provide the foundation for different Precision Farming applications. A documentation of the decisions and differently treated management zones thus can easily be achieved with standardised equipment.

Since such a system should operate in an online approach, it cannot be seen independently from the herbicide application equipment. The decision component has to be adapted to the possibilities and layout of the sprayer: for now there were not many sprayers developed for site-specific precise management of weeds. The requirements for the application technology are high: an optimal sprayer can vary the dose and mixture of different herbicides in very short time intervals (less than one second), using only plain water and the given herbicide formulations. This could be achieved using direct-injection systems (Vondřička (2007)) near the nozzles of the sprayer or premixed herbicides, which are circulating in smaller parts the sprayer system. The trade-off between the amount of technical parts needed and the costs has to be considered besides security and cleaning considerations in the handling of unmixed herbicides.

The development of a commercially available sensor hopefully boosts the research and development in this area, making the goal of site-specific herbicide application feasible for a broad implementation in practise.

References

Adamchuk, V. I., J. W. Hummel, M. T. Morgan, and S. K. Upadhyaya (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture 44*(1), 71 – 91.

Al-Kufaishi, S. A., B. S. Blackmore, and H. Sourell (2006, August). The feasibility of using variable rate water application under a central pivot irrigation system. *Irrigation and Drainage Systems 20*(2-3), 317–327.

Arslan, S. and T. S. Colvin (2002). Grain yield mapping: Yield sensing, yield reconstruction, and errors. *Precision Agriculture 3*(2), 135–154.

Åstrand, B. and A.-J. Baerveldt (2002). An agricultural mobile robot with vision-based perception for mechanical weed control. *Autonomous Robots 13*(1), 21–35.

Åstrand, B. and A.-J. Baerveldt (2004). Plant recognition and localization using context information. In *Proceedings of Mechatronics and Robotics 2004 (MechRob2004)*, Aachen, Germany, pp. 1191–1196. Sascha Eysoldt Verlag. special session Autonomous Machines in Agriculture.

Auernhammer, H. (2001, Feb). Precision farming – the environmental challenge. *Computers and Electronics in Agriculture 30*, 31–43.

Biermacher, J. T., F. M. Epplin, B. W. Brorsen, J. B. Solie, and W. R. Raun (2009). Economic feasibility of site-specific optical sensing for managing nitrogen fertilizer for growing wheat. *Precision Agriculture 10*(3), 213–230.

Biller, R. H. (1998, December). Reduced input of herbicides by use of optoelectronic sensors. *Journal of Agricultural Engineering Research 71*, 357–362.

Burgos-Artizzu, X. P., A. Ribeiro, A. Tellaeche, G. Pajares, and C. Fernández-Quintanilla (2009, March). Improving weed pressure assessment using digital images from an experience-based reasoning approach. *Computers and Electronics in Agriculture 65*(2), 176–185.

Burks, T. F., S. A. Shearer, and F. A. Payne (2000). Classification of weed species using color texture features and discriminant analysis. In *Transactions of the ASAE*, Volume 43, pp. 441–448. American Society of Agricultural Engineers.

Christensen, S. and T. Heisel (1998). Patch spraying using historical, manual and real-time monitoring of weeds in cereals. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz Sonderheft XVI*, 257–263.

Dieleman, J. and D. Mortensen (1999). Characterizing the spatial pattern of *Abutilon theophrasti* seedling patches. *Weed Research 39*, 455–467.

EPPO (2007). Harmonized classification and coding of the uses of plant protection products. *EPPO Bulletin 37*(1), 25–28.

FSF (1989, 1991). Gnu general public license. FSF: Free Software Foundation, Inc.

FSF (1999). Gnu lesser general public license. FSF: Free Software Foundation, Inc.

Gebhardt, S. and W. Kühbauch (2007, Apr). A new algorithm for automatic rumex obtusifolius detection in digital images using colour and texture features and the influence of image resolution. *Precision Agriculture 8*, 1–13.

Gerhards, R. and R. Belz (Eds.) (2008, March). *Special Issue XXI Weed Biology and Weed Control*, Volume Special Issue XXI of *Journal of Plant Diseases and Protection*, Stuttgart-Hohenheim, Germany. German Phytomedical Society (DPG): Eugen Ulmer KG. 24th German Conference on Weed Biology and Weed Control, March 4-6.

Gerhards, R. and S. Christensen (2003). Real-time weed detection, decision making and patch spraying in maize, sugar beet, winter wheat and winter barley. *Weed Research 43*(6), 385–392.

Gerhards, R., C. T. M. Sökefeld, W. Kühbauch, and M. Williams (2002, March). Site specific weed control in maize, sugar beet, winter wheat and winter barley. *Precision Agriculture 3*(1), 25–35.

Godwin, R. and P. Miller (2003). A review of the technologies for mapping within-field variability. *Biosystems Engineering 84*(4), 393–407.

Gutjahr, C. and R. Gerhards (2010, August). *Precision Crop Protection* (1 ed.)., Chapter Decision rules for site-specific weed management, pp. 223–239. In Oerke et al. Oerke, Gerhards, Menz, and Sikora (2010).

Gutjahr, C., B. Huber, M. Weis, and R. Gerhards (2008, May). Untersuchung eines automatisierten bildanalytischen Verfahrens zur Erfassung der Unkrautverteilung in landwirtschaftlichen Kulturflächen. See Zude and Ruckelshausen (2008), pp. 76–83.

Gutjahr, C., M. Weis, M. Sökefeld, C. Ritter, J. Möhring, A. Büchse, H. Piepho, and R. Gerhards (2008, March). Erarbeitung von Entscheidungsalgorithmen für die teilflächenspezifische Unkrautbekämpfung. See Gerhards and Belz (2008), pp. 143–148. 24th German Conference on Weed Biology and Weed Control, March 4-6.

Herold, B. (Ed.) (2006). *Computer-Bildanalyse in der Landwirtschaft - Workshop 2006*, Volume 56 of *Bornimer Agrartechnische Berichte*, Potsdam-Bornim. ATB - Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. 12. Workshop Computer-Bildanalyse in der Landwirtschaft.

Hess, M., G. Barralis, H. Bleiholder, L. Buhr, T. Eggers, H. Hack, and R. Stauss (1997). Use of the extended BBCH scale - general for the descriptions of the growth stages of mono- and dicotyledonous weed species. *Weed Research 37*(6), 433–441.

Hu, M. K. (1962, Feb.). Visual pattern recognition by moment invariants. *IRE Transactions Information Theory 8*(2), 179–187.

Klose, R., M. Thiel, A. Ruckelshausen, and J. Marquering (2008, Sept. 2008). Weedy – a sensor fusion based autonomous field robot for selective weed control. In VDI (Ed.), *LAND.TECHNIK 2008*, Volume 2045 of *VDI-Tagungsbände*, Stuttgart, pp. 167–172. VDI: VDI Verlag. Agricultural Engingeering.

Kudsk, P. (2008, Mar). Optimising herbicide dose: a straightforward approach to reduce the risk of side effects of herbicides. *The Environmentalist 28*, 49–55.

Mierswa, I., M. Wurst, R. Klinkenberg, M. Scholz, and T. Euler (2006, August). Yale: Rapid prototyping for complex data mining tasks. In L. Ungar, M. Craven, D. Gunopulos, and T. Eliassi-Rad (Eds.), *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining*, New York, NY, USA, pp. 935–940. ACM.

Miller, P. (2003, May). Patch spraying: future role of electronics in limiting pesticide use. *Pest Management Science 59*, 566–574.

Mondal, P. and V. Tewari (2007). Present status of precision farming: A review. *International Journal of Agricultural Research 2*(1), 1–10.

Mortensen, D. A. (2002). Crop/weed outcomes from site-specific and uniform soil-applied herbicide applications. *Precision Agriculture 3*(1), 95.

Oebel, H. (2006, July). *Teilschlagspezifische Unkrautbekämpfung durch raumbezogene Bildverarbeitung im Offline- und (Online-) Verfahren (TURBO)*. Ph. D. thesis, Universität Hohenheim, Fakultät Agrarwissenschaften.

Oebel, H. and R. Gerhards (2005). Site-specific weed control using digital image analysis and georeferenced application maps – first on-farm experiences. In J. Stafford (Ed.), *Proceedings 5th European Conference on Precision Agriculture (ECPA)*, Uppsala, Sweden, pp. 131–138. Wageningen Academic Publishers.

Oerke, E. (2006). Crop losses to pests. *Journal of Agricultural Science 144*(1), 31–43.

Oerke, E.-C., R. Gerhards, G. Menz, and R. A. Sikora (Eds.) (2010, August). *Precision Crop Protection* (1 ed.). Dordrecht, Heidelberg, London, New York: Springer Verlag.

PflSchG (1986). Gesetz zum Schutz der Kulturpflanzen (Pflanzenschutzgesetz). In der Fassung vom 14.5.1998, zuletzt geändert durch Zweites Gesetz zur Änderung des Pflanzenschutzgesetzes vom 22.6.2006.

Piron, A., V. Leemans, F. Lebeau, and M. Destain (2009). Improving in-row weed detection in multispectral stereoscopic images. *Computers and Electronics in Agriculture 69*(1), 73–79.

Pérez, A., F. López, J. Benlloch, and S. Christensen (2000, Feb). Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture 25*, 197–212.

Radcliffe, E. B., W. D. Hutchison, and R. E. Cancelado (Eds.) (2009). *Integrated Pest Management: Concepts, Tactics, Strategies and Case Studies*. Cambridge University Press.

Rasmussen, J., M. Nørremark, and B. Bibby (2007). Assessment of leaf cover and crop soil cover in weed harrowing research using digital images. *Weed Research 47*, 299–310.

Reichardt, M., C. Jürgens, U. Klöble, J. Hüter, and K. Moser (2009). Dissemination of precision farming in germany: acceptance, adoption, obstacles, knowledge transfer and training activities. *Precision Agriculture*.

Ritter, C., D. Dicke, M. Weis, H. Oebel, H. Piepho, A. Büchse, and R. Gerhards (2008, Jun). An on-farm approach to quantify yield variation and to derive decision rules for site-specific weed management. *Precision Agriculture 9*(3), 133–146.

Ritter, C., D. Dicke, M. Weis, M. Sökefeld, H. Oebel, H. Piepho, A. Büchse, and R.Gerhards (2007, June). Analysing decision rules for site-specific weed control subject to variable weed distribution and soil quality in winter wheat. In *Proceedings of the 14th European Weed Research Society (EWRS) Symposium*, Hamar, Norway, pp. 131. EWRS.

Ritter, C. and R. Gerhards (2007). Can short-term gains in site-specific weed management be sustained over multiple years? In *Proceedings of the 2nd Conference on Precision Crop Protection*, Bonn, Germany.

Shafri, H. Z. M., M. A. M. Salleh, and A. Ghiyamat (2006, June). Hyperspectral remote sensing of vegetation using red edge position techniques. *American Journal of Applied Sciences 3*(6), 1864–1871.

Soille, P. (2003). *Morphological Image Analysis* (2nd ed.). Heidelberg: Springer-Verlag.

Sonka, M., V. Hlavac, and R. Boyle (1999, September). *Image Processing: Analysis and Machine Vision* (2 ed.). Computing Series. Chapmann & Hall.

SPSS Inc (2004). SPSS for Windows. Rel. 12.0.2 (Standard Version).

Stafford, J. (Ed.) (2007, June). *Precision agriculture '07*, Volume 6, The Netherlands. 6th European Conference on Precision Agriculture (ECPA): Wageningen Academic Publishers.

Sökefeld, M., R. Gerhards, H. Oebel, and R.-D. Therburg (2007, June). Image acquisition for weed detection and identification by digital image analysis. See Stafford (2007), pp. 523–529.

Sökefeld, M., R. Gerhards, H. Oebel, M. Weis, P. Risser, and M. Pflugfelder (2006). Wissensbasierte Artenerkennung anhand von Konturmerkmalen der Pflanzen und verschiedenen Klassifikationsalgorithmen. See Herold (2006), pp. 51–59. 12. Workshop Computer-Bildanalyse in der Landwirtschaft.

Vondřička, J. (2007, Nov). *Study on the process of direct nozzle injection for real-time site-specific pesticide application.* Ph. D. thesis, Rheinische Friedrich-Wilhelms-Universität, Bonn. Forschungsbericht Agrartechnik des Arbeitskreises Forschung und Lehre der Max-Eyth-Gesellschaft Agrartechnik im VDI (VDI-MEG).

Weis, M. (2007, October). Spatial data infrastructures for precision farming data - standards and system design criteria. In H. Bleiholder and H.-P. Piepho (Eds.), *Agricultural Field Trials - Today and Tomorrow*, Beuren, Stuttgart, pp. 254–257. Universität Hohenheim, Institut für Pflanzenbau und Grünland: Verlag Grauer.

Weis, M. (2009, Jan). Intelligente Sensorik in der Unkrautbekämpfung. Agrarstellenmarkt, Beruf und Bildung 02, Proplanta GmbH & Co. KG, 70599 Stuttgart-Hohenheim.

Weis, M. and R. Gerhards (2007a, June). Feature extraction for the identification of weed species in digital images for the purpose of site-specific weed control. See Stafford (2007), pp. 537–545.

Weis, M. and R. Gerhards (2007b, October). Identification of weeds from digital images. In H. Bleiholder and H.-P. Piepho (Eds.), *Agricultural Field Trials - Today and Tomorrow*, Beuren, Stuttgart, pp. 258–263. Universität Hohenheim, Institut für Pflanzenbau und Grünland: Verlag Grauer.

Weis, M. and R. Gerhards (2007c, May). Qualitative und quantitative Messung der Verunkrautung in Kulturpflanzenbeständen mittels

Bildanalyse. See Zude (2007), pp. 67–74.

Weis, M. and R. Gerhards (2009a, July). Automatic derivation of weed densities from images for site-specific weed management. In C. Lokhorst and R. d. L. J.F.M. Huijsmans (Eds.), *JIAC2009 Book of abstracts*, Wageningen, Netherlands, pp. 349–354. ECPA (European Conference on Precision Agriculture): Wageningen Academic Publishers. Only available on CD-rom (PDF-file) - Proceedings of the Joint International Agricultural Conference, 6.-8. July.

Weis, M. and R. Gerhards (2009b, August). Detection of weeds using image processing and clustering. See Zude (2009), pp. 138–144. 15. Workshop Computer-Bildanalyse in der Landwirtschaft, 27–28 August 2009.

Weis, M., C. Gutjahr, V. Rueda Ayala, R. Gerhards, C. Ritter, and F. Schölderle (2008, Dec). Precision farming for weed management: techniques. *Gesunde Pflanzen 60*, 171–181.

Weis, M., T. Rumpf, R. Gerhards, and L. Plümer (2009, August). Comparison of different classification algorithms for weed detection from images based on shape parameters. See Zude (2009), pp. 53–64. 15. Workshop Computer-Bildanalyse in der Landwirtschaft, 27–28 August 2009.

Weis, M. and M. Sökefeld (2010, August). *Precision Crop Protection* (1 ed.)., Chapter Detection and identification of weeds, pp. 119–134. In Oerke et al. Oerke, Gerhards, Menz, and Sikora (2010).

Wiles, L. J. (2005, Mar). Sampling to make maps for site-specific weed management. *Weed Science 53*, 228–235.

Woebbecke, D., G. Meyer, and K. von Bargen (1995). Shape features for identifying young weeds using image analysis. *Transactions of the American Society of Agricultural Engineers 38*, 271–281.

Woebbecke, D., G. Meyer, K. von Bargen, and D. Mortensen (1995). Color indices for weed identification under various soil, residue and lighting conditions. *American Society of Agricultural Engineers* (259-269).

Zier, P., K. Hank, and P. Wagner (2008). Economic potential of autoguidance systems. *Berichte über Landwirtschaft: Zeitschrift für Agrarpolitik und Landwirtschaft 86*(3), 410–432.

Zude, M. (Ed.) (2007, May). *13. Workshop Computer-Bildanalyse in der Landwirtschaft & 4. Workshop Precision Farming*, Volume 60 of *Bornimer Agrartechnische Berichte*, Potsdam-Bornim/Stuttgart. ATB - Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V.

Zude, M. (Ed.) (2009, August). *Image analysis for agricultural products and processes*, Volume 69 of *Bornimer Agrartechnische Berichte*, Potsdam-Bornim. CIGR, ATB - Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V.: Leibnitz Institute for Agricultural Enineering (ATB). 15. Workshop Computer-Bildanalyse in der Landwirtschaft, 27–28 August 2009.

Zude, M. and A. Ruckelshausen (Eds.) (2008, May). *14. Workshop Computer-Bildanalyse in der Landwirtschaft*, Volume 62 of *Bornimer Agrartechnische Berichte*, Potsdam-Bornim/Osnabrück. ATB - Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V.

Zwerger, P. and H. U. Ammon (2002). *Unkraut: Ökologie und Bekämpfung*. Stuttgart (Hohenheim): Ulmer (Eugen).

A. Image processing software

Software was written for the image processing and classification. To handle the amount of data generated during the image processing and classification a database layout was developed. All information about the images, the processing parameters, features and results are stored in the database. 156 series consisting of 360,000 images and 7,300,000 segments were stored in the database at the time of writing. This way the data handling is unified and it was possible to develop modular software for the creation and usage of the data. The central component for the image processing is a program with the name SEGMENTATION, that does the image preprocessing, segmentation, binarisation and the feature extraction. Difference images or color images can be used as input, the latter can be segmented with two different algorithms (options `-y, -Y, -z, -Z`). A threshold can be given (option `-t`) or automatically derived from the image (option `-T`, Sonka et al. (1999, p. 129)). Preprocessing operation include morphological operators (options `-e, -o`), with `-B` regions can be dicarded, that are cut by the borders of the image. The region for the computation can be restricted to a bounding box (`-U`) or a predefined region of a binary image (`-u`). Holes in segments can be filled with `-f` and a criterion can select segments according to their size (options `-m, -M`). The computation of the features is controlled via options `-n, -N -l, -L`, to only compute the binary image `-C` can be used. A row histogram can be received with option `-H`, summing up the values along the rows. If the camera is directed in crop row direction, the maxima of the histogram can be used to identify these rows. The results are stored either in XML[1]-files (for 'standalone' operation, option `-f, -a`) or in the database (option `-q`). The storage of intermediate results (binary image, segments, ...) is controlled by option `-s`. The program can be run with or without display of the intermediate processing results (options `-D` and `-d`). A classification can be performed and the classification results be

[1]XML: eXtensible Markup Language

displayed ('online mode', options `-I, -i, -w, -W`).

A separate program named CLASSIFICATION uses the feature data in the database to classify (supervised or unsupervised) all segments of a series, the options are similar to the ones of `segmentation` (without image processing related). The result of the classification is written back into the database. Another program can be used for this task: RAPIDMINER (Mierswa et al. (2006)). A flexible setup for the parametrisation of the classification was created, which allows the usage of different algorithms and data sets (Fig. A.7).

Segmentation parameters These are the options for the software SEGMENTATION:

```
Usage: segmentation [OPTIONS]... [FILES]...

  -h, --help                   Print help and exit
  -V, --version                Print version and exit
  -t, --threshold=INT          set theshold value  (default='50')
  -T, --autothreshold          use automatically derived threshold
                                 (default=off)  (default=off)
  -m, --minimumsize=INT        set minimum size for regions  (default='1')
  -M, --maximumsize=INT        set maximum size for regions (0=no limit)
                                 (default='0')
  -F, --fillholes              fill holes in regions  (default=off)
  -y, --iscolorimage=STRING    assume the image is a color image, use named
                                 color segmentation approach 'egi|hsv'
                                 (default 'egi'), for hsv see also -Y,-z,-Z
                                 (default='egi')
  -Y, --minsaturation=FLOAT    set the minimum saturation threshold for color
                                 segmentation  (default='0.05')
  -z, --mincolor=FLOAT         set the minimum color threshold for the
                                 segmentation of color images (range=0-1)
                                 (default='0.25')
  -Z, --maxcolor=FLOAT         set the maximum color threshold for the
                                 segmentation of color images (range=0-1)
                                 (default='0.4')
  -e, --erosion=INT            erode the regions with a circle of specified
                                 size (default=1), negative values: dilate
                                 (default='0')
  -o, --open=INT               open/close the regions with a circle of
                                 specified size (default=1), negative values:
                                 close  (default='0')
  -B, --noborderregions        disregard regions which are cut by the border
                                 of the image (default=off)  (default=off)
  -v, --verboseness=INT        set verbose level, 0=nothing  (default='0')
```

```
-d, --debugview=INT          enable debug viewing of image processing steps,
                               higher values show more  (default=‘-1’)
-D, --nokey                  Do not wait for a key pressed, if debugview is
                               enabled (off=wait)  (default=off)
-s, --save-interim=INT       save interim results to current directory,
                               higher number saves more: (1) binary image,
                               (2) segments, (3) clipped original files, (4)
                               skeleton of segment (5) distance image;
                               default is 0  (default=‘0’)
-f, --fileout=STRING         Outputfile (features in xml)
                               (default=‘out.xml’)
-a, --appendxml              Append XML output to Outputfile (-f, fileout)
                               (default=off)
-c, --classtoset=STRING      Set the class in the XML (for all regions)
                               (default=‘’)
-X, --numberoffffeatures=INT  Number of Fourierfeatures in output
                               (default=‘20’)
-x, --numberofcssfeatures=INT Number of CSSfeatures in output  (default=‘5’)
-n, --nocss                  Do not compute CSS (Curvature Scale Space)
                               (off=compute)  (default=off)
-N, --noff                   Do not compute Fourierfeatures (off=compute)
                               (default=off)
-l, --nofeat                 Do not compute Areafeatures (off=compute)
                               (default=off)
-L, --noskeleton             Do not compute Skeleton (off=compute)
                               (default=off)
-C, --coverageonly           Only compute coverage, nothing else
                               (default=off)
-u, --usemaskimage=STRING    use mask from image (min/max coordinates are
                               computed)  (default=‘mask.png’)
-U, --usemaskcoordinates=STRING
                             use mask min/max coordinates (format:
                               'minx:miny:maxx:maxy')
                               (default=‘0:0:1024:766’)
-q, --sqlparameters=STRING   store features to a Mysql database with the
                               parameters
                               'host:db:table:user:password:parametersetname'
-I, --inmodel=STRING         Input file with trained model to use
                               (default=‘model’)
-i, --minquality=FLOAT       Minimum quality (of classifier output)
                               (default=‘0’)
-w, --classify=STRING        classify data using one of the following:
                               rbf|lvq  (default=‘rbf’)
-W, --featuresetname=STRING  featuresetname  (default=‘default’)
-H, --columnhistogram        compute column sums/histogram  (default=off)
```

A.1. Graphical user interface

A graphical user interface (GUI) was created for the image processing and classification. It was developed as web application, which has the benefit to be available via Intranet and Internet to the interested researchers. The web interface covers most of the functionality of the software components. The components can be parametrised and run consecutively. The results of the computations can be visualised and the raw data as well as resulting maps can be downloaded for further usage. The application provides several interactive web pages:

- *Edit classes*: Classes can be defined here (Fig. A.1).
- *Select parameters/set classes*: Image processing parameters can be tested and stored as sets for further usage (Fig. A.2 and Fig. A.3)
- *Classification*: Several classification algorithms are available to classify all objects of a series according to the training data (Fig. A.4).
- *Classification (RM)*: Same as Classification, but using different software (Rapidminer), which provides a more flexible classification framework (Fig. A.5).
- *Clustering*: Clustering algorithms (unsupervised classifiers) can be used to group objects according to their shape features (Fig. A.6).
- *Clustering (RM)*: Clustering algorithms implemented with Rapidminer, more different algorithms available (Fig. A.7).
- *Data Administration*: Image series can be defined, information about existing data sets (images, training data, parameter sets, classification statistics) can be reviewed and batch processing of image series is possible on this page (Fig. A.8).
- *Edit Traindata*: Training data sets can be reviewed and changed on this page (Fig. A.9).
- *Train from image*: Training data sets can be generated from the images by selecting objects and assigning classes to them. Results of the image processing and classification can be visualised (Fig. A.10, A.11 and A.12).

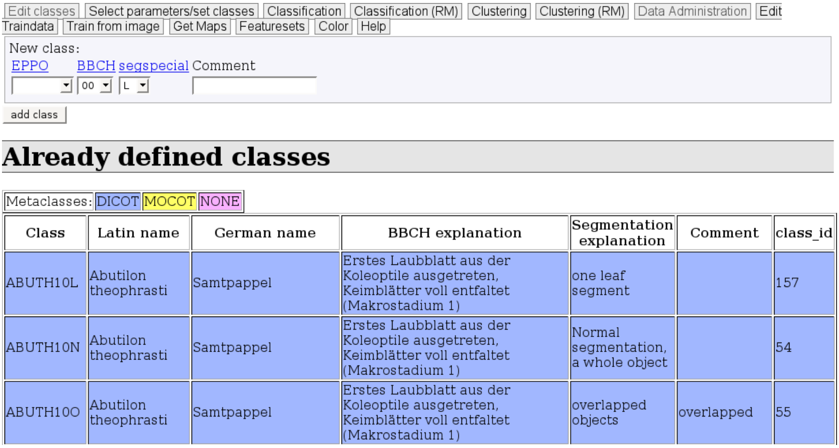


Figure A.1.: Class definition page. A new class can be generated using the three drop down boxes, already defined classes are listed.

- *Get Maps*: Training data, feature data and results of the classification are provided on this page. The data can be downloaded and be used in desktop software (Fig. A.13).
- *Featuresets*: Subsets of features are defined and reviewed on this page (Fig. A.14).
- *Colour*: A unique colour for each species is defined on this page, leading to unique class colours. These are used for the resulting label images and allow the integration of the classification result into images (Fig. A.15).
- *Help*: Explains the functionality of the provided pages.

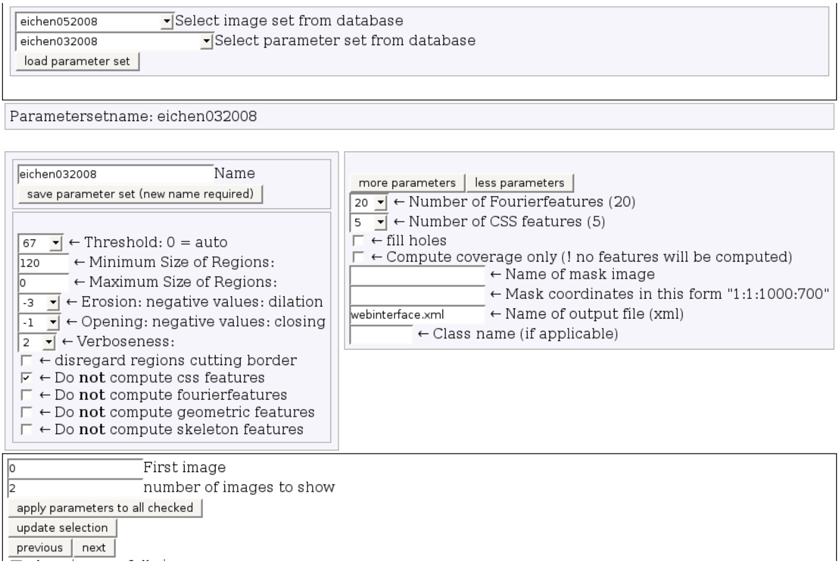


Figure A.2.: Definition of image processing parameters. The parameter sets can be tested (Fig. A.3) and stored into the database.

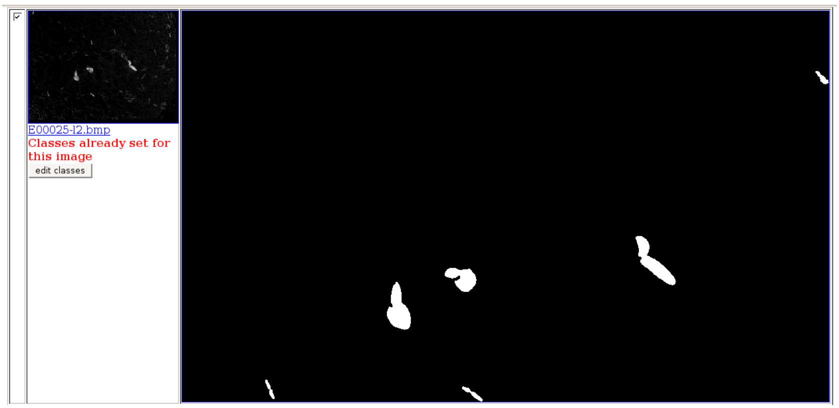


Figure A.3.: Result of applied image processing parameters.

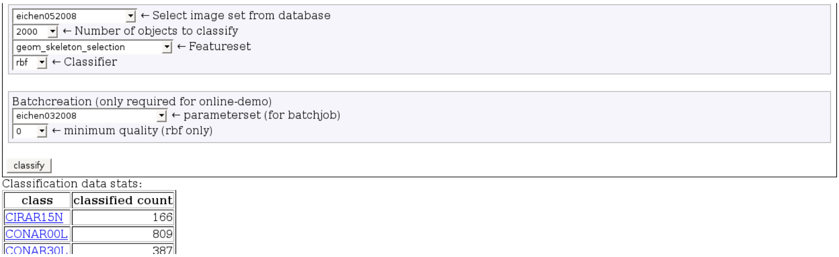


Figure A.4.: The parameters for the classification of a data set are defined on top, below classification statistics are shown.

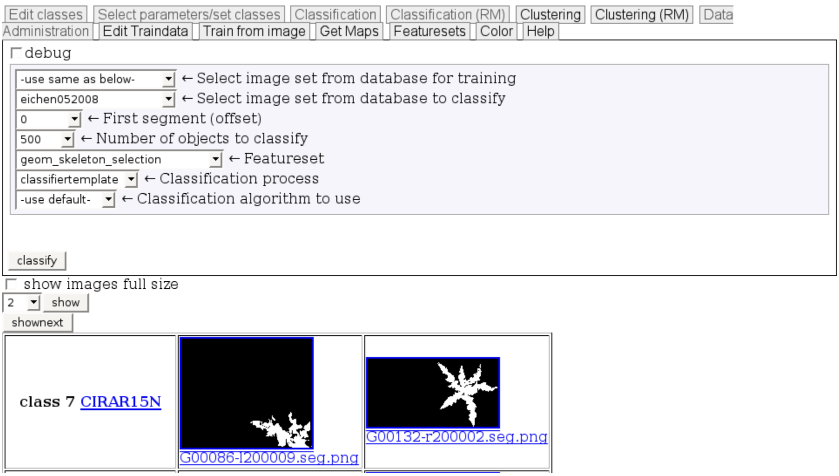


Figure A.5.: The parameters for the classification of a data set with RapidMiner are defined on top, below classification results (segments ordered by class) are shown.

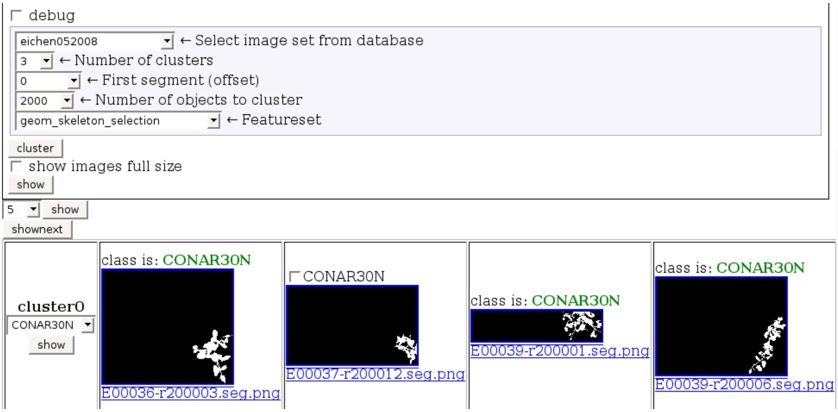


Figure A.6.: Clustering of a data set, below clustering results (segments ordered by cluster) are shown. The chosen class for each cluster (left) can be set using the check box of each segment. This way training data can be generated efficiently.

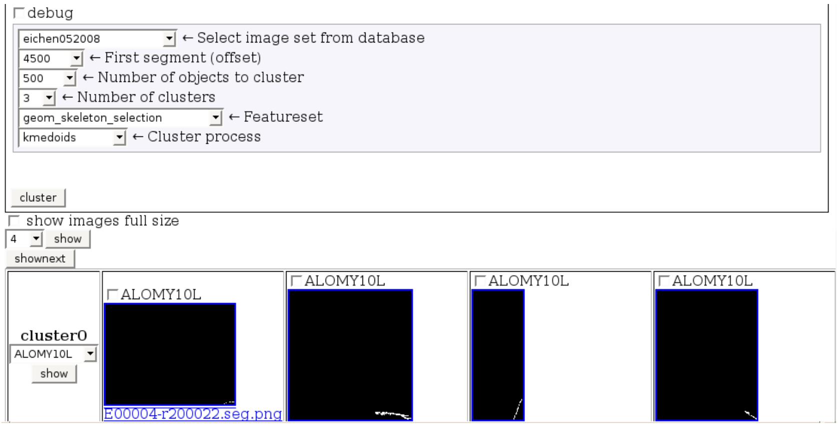


Figure A.7.: Clustering of a data set with RAPIDMINER, below clustering results (segments ordered by cluster) are shown. The chosen class for each cluster (left) can be set using the check box of each segment. This way training data can be generated efficiently.

Figure A.8.: Data administration interface. Statistics of existing data sets can be reviewed (with image and training data lists), data sets (series) can be generated or deleted and batch processed on this page.

Figure A.9.: Training data review page. The defined training data can be reviewed and changed here.

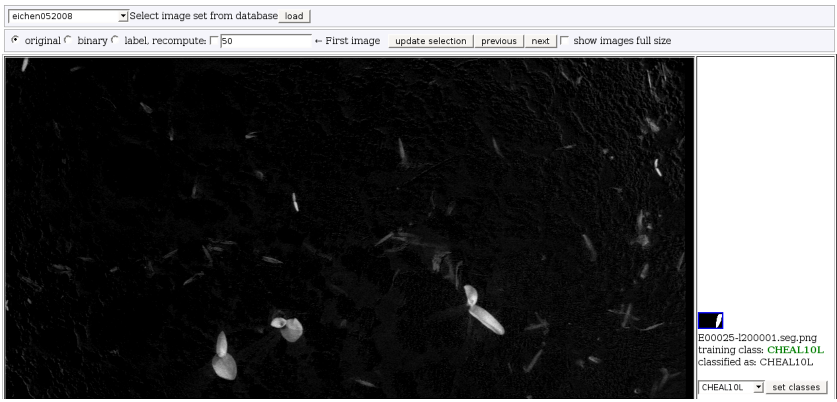


Figure A.10.: Training data creation from difference image. Objects can be selected in the image and a training class set (on the right)

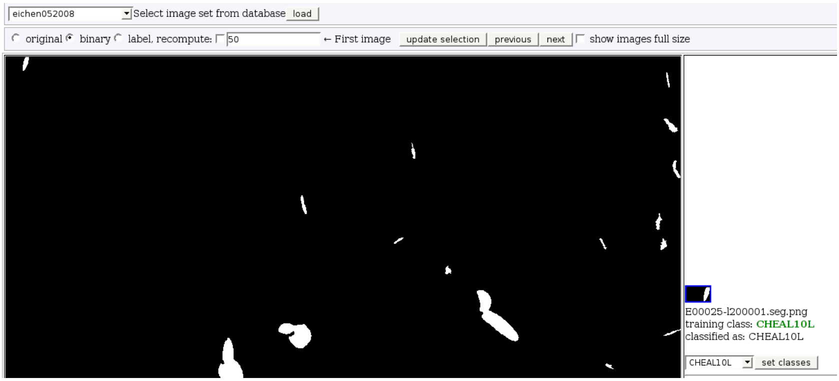


Figure A.11.: Training data creation from binary image. Objects can be selected in the image and a training class set (on the right)

Figure A.12.: Training data creation from label image. Objects can be selected in the image and a training class set (on the right)

Figure A.13.: Download raw data and maps. On this page the training data, features and maps can be downloaded.

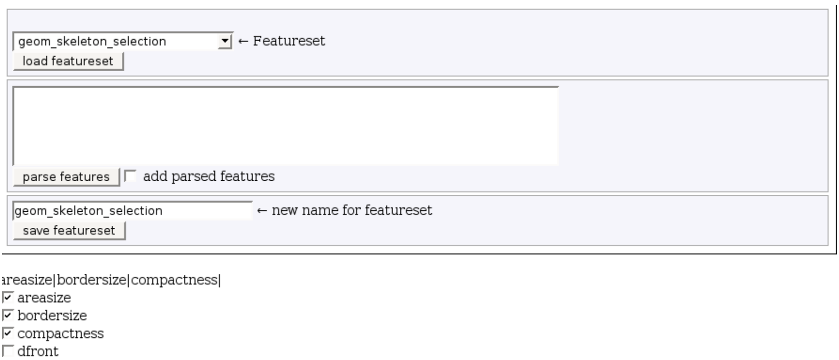
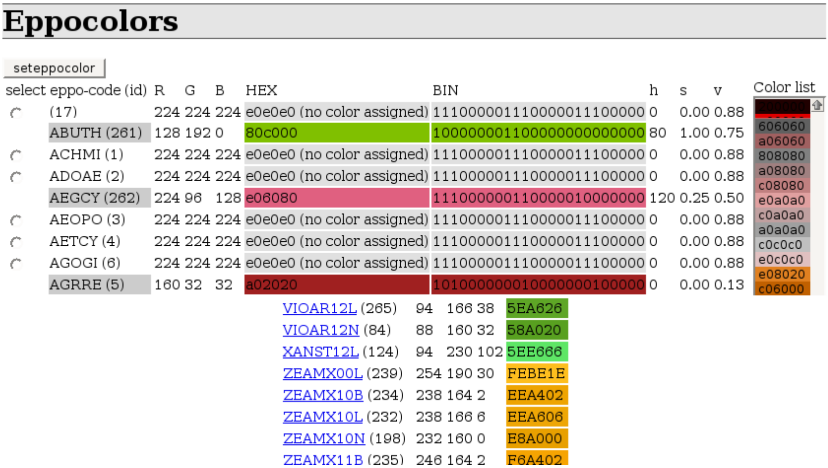


Figure A.14.: Feature set creation page. A subset of features can be selected on this page for further usage (classification and clustering).



Eppocolors

seteppocolor

select	eppo-code (id)	R	G	B	HEX	BIN	h	s	v
◌	(17)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
	ABUTH (261)	128	192	0	80c000	100000001100000000000000	80	1.00	0.75
◌	ACHMI (1)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
◌	ADOAE (2)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
	AEGCY (262)	224	96	128	e06080	111000000110000010000000	120	0.25	0.50
◌	AEOPO (3)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
◌	AETCY (4)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
◌	AGOGI (6)	224	224	224	e0e0e0 (no color assigned)	111000001110000011100000	0	0.00	0.88
	AGRRE (5)	160	32	32	a02020	101000000010000000100000	0	0.00	0.13

Color list: 606060, a06060, 808080, a08080, c08080, e0a0a0, c0a0a0, a0a0a0, c0c0c0, e0c0c0, e08020, c06000

VIOAR12L (265)	94	166	38	5EA626
VIOAR12N (84)	88	160	32	58A020
XANST12L (124)	94	230	102	5EE666
ZEAMX00L (239)	254	190	30	FEBE1E
ZEAMX10B (234)	238	164	2	EEA402
ZEAMX10L (232)	238	166	6	EEA606
ZEAMX10N (198)	232	160	0	E8A000
ZEAMX11B (235)	246	164	2	F6A402

Figure A.15.: Class colour definition. On this page a colour can be chosen and reviewed for each species (The numbers denote the RGB/HSV values). The resulting colours for each class are displayed below. They are used in the label images (see Fig. A.12).