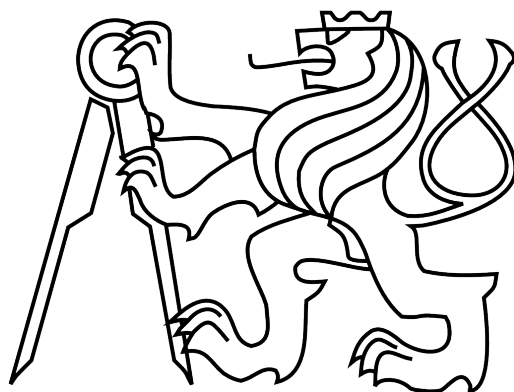


Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Cybernetics



BACHELOR THESIS

Image-based Recognition of Plants

Milan Šulc

Supervisor: Prof. Ing. Jiří Matas, Ph.D.

Study Programme: Cybernetics and Robotics

May, 2012

BACHELOR PROJECT ASSIGNMENT

Student: Milan Šulc
Study programme: Cybernetics a Robotics
Specialisation: Robotics
Title of Bachelor Project: Image-based Recognition of Plants

Guidelines:

1. The thesis builds on the results of Tomas Sixta's diploma thesis. Familiarize yourself with the thesis and the accompanying software, test it and evaluate it.
2. Familiarize yourself with the state-of-the-art in image-based recognition of plants.
3. Consider various photo types (leaf, bark, blossom, fruit/cone, needles, overall, ..) for recognition of plants, propose a visual recognition method for one of those not used by T. Sixta.
4. Implement the method and evaluate its performance on publicly available data if possible.
5. Create an Android phone application for plant recognition.
6. Implement data sharing interface.

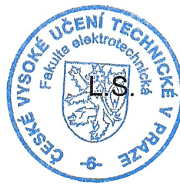
Bibliography/Sources:

- [1] Sixta, T.: Image and Video-based Recognition of Natural Objects . Prague, 2011.
- [2] Forsyth, D.A.; Ponce, J.: Computer vision: A Modern Approach. New Jersey, 2002.
- [3] Belhumeur, P.N. et al.: Searching the World's Herbaria: A System for Visual Identification of Plant Species. ECCV, 2008.

Bachelor Project Supervisor: prof. Ing. Jiří Matas, Ph.D.

Valid until: the end of the winter semester of academic year 2012/2013


prof. Ing. Vladimír Mařík, DrSc.
Head of Department




prof. Ing. Pavel Ripka, CSc.
Dean

Prague, December 1, 2011

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Milan Šulc
Studijní program: Kybernetika a robotika (bakalářský)
Obor: Robotika
Název tématu: Rozpoznávání rostlin z obrázku

Pokyny pro vypracování:

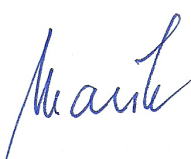
1. Práce staví na výsledcích diplomové práce Tomáše Sixty. Seznamte se s touto prací a doprovodným softwarem, otestujte a vyhodnoťte jej.
2. Seznamte se s moderními postupy obrazového rozpoznávání rostlin a současně používanými metodami.
3. Zvažte různé typy fotografií (list, kůra, květ, plod/šiška, jehličí, celek, ..) pro rozpoznávání rostlin, navrhnete vizuální metodu pro jeden, který nebyl používán T. Sixtou.
4. Metodu implementujte a, pokud možno, ohodnoťte její úspěšnost na veřejně dostupných datech.
5. Vytvořte aplikaci pro rozpoznávání rostlin na Android telefonu.
6. Implementujte rozhraní pro sdílení dat.

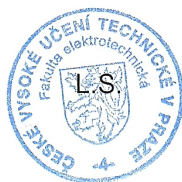
Seznam odborné literatury:


- [1] Sixta, T.: Image and Video-based Recognition of Natural Objects . Prague, 2011.
- [2] Forsyth, D.A.; Ponce, J.: Computer vision: A Modern Approach. New Jersey, 2002.
- [3] Belhumeur, P.N. et al.: Searching the World's Herbaria: A System for Visual Identification of Plant Species. ECCV, 2008.

Vedoucí bakalářské práce: prof. Ing. Jiří Matas, Ph.D.

Platnost zadání: do konce zimního semestru 2012/2013


prof. Ing. Vladimír Mařík, DrSc.
vedoucí katedry




prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 1. 12. 2011

Acknowledgement

Foremost, I would like to express my sincere gratitude to my supervisor, prof. Jiří Matas, who led me through this thesis development with great support and encouragement.

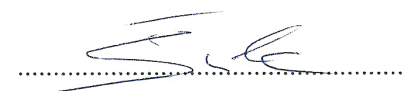
I would also like to thank my friends, colleagues and other people who helped me with interesting ideas, suggestions or support.

A special thank belongs to my family for supporting me in my studies and in my life.

Prohlášení autora práce

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne24.5.2012.....



Podpis autora práce

Abstract

This thesis has two contributions:

First, a new semi-automatic method for conifers identification based on images of needles (or more precisely images of coniferous branches) was proposed. The method's inputs are (except of the image) the coordinates of two branch-points. After finding the branch shape and the "needle area" in the branch surroundings, the texture pattern direction is described and normalised to correspond the branch direction and shape. The descriptor is then a 4-D vector (histogram) and the nearest neighbour classifier is used for identification. A dataset containing 93 images of 4 species (fir, spruce, larch, pine) was created. An experiment on this dataset shows 72.3% classification precision.

Second, a mobile application for plant identification was created, which allows users to identify plants using image recognition (from leaves and bark, as proposed by Sixta [1], and from needles, as described in this thesis), manual field guide or reading plant species characteristics. This application also allows to save the photographed images (with GPS coordinates) or upload them to a community website, which was also created within this thesis.

Abstrakt

Tato práce má dva přínosy:

Zaprvé byla navržena nová poloautomatická metoda pro identifikaci jehličnanů na základě obrázků jehličí (respektive obrázků větvíček s jehličím). Vstupem pro tuto metodu jsou (kromě obrázku) souřadnice dvou bodů větve. Po nalezení tvaru větve a "oblasti jehličí" v okolí větve je pak popsána směrovost vzorku textury, která je dále normalizována, aby odpovídala směru a tvaru větve. Deskriptorem je pak 4-D vektor (histogram) a pro identifikaci je použita metoda nejbližšího souseda. Byl také vytvořen dataset obsahující 93 fotek 4 druhů (jedle, smrk, modřín, borovice). Experiment na tomto datasetu ukazuje přesnost klasifikace 72.3%.

Zadruhé byla vytvořena mobilní aplikace pro identifikace roslin, která uživatelům umožňuje identifikovat rostliny pomocí rozpoznávání obrazu (z listů a kůry, jak navrhl Sixta [1], a z jehličí, jak je popsáno v této práci), klíče nebo čtení charakteristik druhů roslin. Tato aplikace také umožňuje ukládat pořízené fotografie (vč. GPS souřadnic) nebo je nahrána komunitní webovou stránku, která byla v rámci práce také vytvořena.

Contents

1	Introduction	7
1.1	Motivation	7
1.2	Goals	7
1.3	The structure of the thesis	9
2	Plant identification: State of the art	10
2.1	Image-based Plant Identification	12
2.1.1	Leaves	12
2.1.2	Bark	16
2.1.3	Flowers	17
2.1.4	Fruit, Cones	21
2.1.5	The overall plant image, shape, contour	21
3	Identification from needles	23
3.1	Datasets	23
3.1.1	Osterreichische Bundesforste/Archiv and Computer Vision Lab, Vienna UT	23
3.1.2	The CB Dataset	24
3.2	The texture Pattern Direction (TPD) description	27
3.2.1	The definition of TPD	27
3.2.2	Application of TPD on images of coniferous branches	28
3.3	The branch shape search	29
3.4	Detecting the width of needle texture area	30
3.5	Normalizing the TPD description of needles	31
3.6	Classification and evaluation	32
4	Implementation	33
4.1	The Android Application	33
4.1.1	Fundamental Layout	33
4.1.2	Identification using Image Recognition	34
4.1.3	Manual Field Guide	36
4.1.4	Album, Sharing, Map	38
4.2	Web	40
4.2.1	Database	40

4.2.2	Gallery	40
4.2.3	Map	41
4.3	Managing datasets	42
4.3.1	Creating descriptors	42
4.3.2	Web administration	43
5	Conclusions	44
5.1	Future work	44
6	References	46
A	Contents of the enclosed CD	49
B	Graphical resources	49

List of Figures

1	A computer vision system for identifying plants on Plummers Island, Maryland, USA. (Source: [2])	10
2	Leafsnapp app for iPhone	11
3	Encyclopedia of Life	12
4	Recognition rates of Xiao et al. (Source: [11])	14
5	Species from first "class" (Source: [8])	15
6	Species from second "class" (Source: [8])	15
7	Sixta: Confusion matrix for 11 class experiment, rows are true classes, columns estimated (Source: [1])	17
8	Saitoh: Separate HS space (Source: [15])	18
9	The overall system architecture of Kim et al. (Source: [20])	20
10	Fruit picking system of Zhao et al. (Source: [23])	21
11	The USFS-TAMU clip art (forestry.about.com): Tree silhouettes	22
12	The CB Dataset	26
13	Texture Pattern Directions - Visualisation	27
14	Neighbouring pixel colours labels for Texture Pattern Directions	27
15	TPD applied od images of coniferous branches	28
16	Finding the branch shape	30
17	Finding the area width	31
18	Left: Google Nexus One (HTC). Right: Samsung Galaxy S Plus	33
19	PlantIdent: Tab Layout - Photo Tab, Guide Tab, Album Tab	34
20	PlantIdent: Leaf recognition	35
21	PlantIdent: Bark recognition	35
22	PlantIdent: Needles recognition	36
23	A transparent identification key with images, The Natural History Museum Cromwell Road London, UK www.nhm.ac.uk/urban-tree-survey	37
24	PlantIdent: Guide	38
25	PlantIdent: Managing accounts and uploading	39
26	PlantIdent: Map - photos; photos with current location	39
27	Web: Database	40
28	Web: Gallery	41

29	Web: Maps	42
30	Creating leaf descriptors	42
31	Creating leaf descriptors	43
32	Species administration	43
33	Guide administration	43

List of Tables

1	Confusion table for conifers classification	32
2	CD contents description	49
3	List of graphical resources used in PlantIdent for Android	49

1 Introduction

1.1 Motivation

The topic of this thesis is semi-automatic image-based plant identification, its implementation and application.

With the progress of mobile technologies and expansion of phones with advanced operating systems, new application possibilities are opening. In this thesis I focus on the possibility of using mobile phone for plant identification. Since almost all present phones have a camera, computer vision methods can provide a great help here.

Such application will be useful not only to regular users, who would like to key a plant of an unknown species. It will also find employment for other purposes, such as plant mapping, which is a common task of scientists and other professionals in fields of botany, environmental studies or gardening.

In addition, this application can also raise interest in nature among the non-expert users, who could share their photographs via Internet. From a wider point of view, a closer relation to nature can then lead to better environmental behaviour.

1.2 Goals

This project and thesis have several goals on different levels. The three main objectives are:

- improving computer vision methods used for plant identification
- creating a mobile phone application for plant identification and relevant data collection
- creating an interface for data sharing and potential community activity

For this goals, the results of T. Sixta, who proposed visual methods for broad leaf recognition and for bark recognition in his diploma thesis [1], were a very valuable basis.

One of the results of this bachelor thesis is an utilizable mobile software for plant identification/recognition. The created mobile application is called Plant Ident, which has a wide range of application, including:

Plant identification

The main function of Plant Ident is to provide a modern guide for plant identification. Both computer vision (image processing of photo taken with the phone) and user-input (answering guide questions, evaluation of the assessment) are employed. This feature would be helpful for a spectrum of users from non-experts (tourists, natural lovers, amateur gardeners) to professionals (botanists, agricultural experts, expert witnesses,..).

Providing information about particular species

The application should provide some basic information about all species in database and offer further external information sources available on the Internet. The information can also be used to assure the user of the identification results.

Plant mapping

Scientists and experts often map certain species across a given area. Current mobile phones also offer GPS positioning, which, in combination with a corresponding software, could be transformed in an effective and accessible tool for plant / species mapping.

Gathering data and creating datasets

For future development and improvements, comprehensive marked datasets would be useful. It should be possible to build new datasets easily. Also data provided from users could be utile.

Data sharing

All users will be able to upload data to the related website, where it will be accessible in an unified form.

1.3 The structure of the thesis

This thesis is structured as follows:

- First the state of the art and literature review are presented, including different approaches to image-based plant identification. In this part I mention different image-types for plant identification, such as leaves and needles, bark, blossoms, fruits or overall plant silhouette.
- Thereafter I choose needles (or more precisely images of coniferous branches) as a type of image-source for my research and propose a method to use it for plant identification. Since it is a very poorly researched area, gathering a relevant dataset was also an important step, to allow the methods evaluation.
- After describing the methodology, the implementation of Plant Ident mobile application for Android and the implementation of the related website is presented.
- In the end of this thesis I sum up the results and consider future work and potential improvements of the Plant Ident project and of the proposed needles-based plant identification method.

Contents of the enclosed CD are described in the appendix.

2 Plant identification: State of the art

The effort to identify plants and observe their distribution has accompanied humankind since time immemorial, as even our survival depended on it.

Despite all advances in computer science, most of the botanical materials and guides are still in written form.

Nevertheless, there are also attempts to integrate novel approaches to the field. Let me mention some of those projects:

Leafsnap and related projects

Leafsnap is an electronic field guide based on visual recognition of leaf shape, available for iPhone and iPad. It is a result of collaboration of Columbia University with the University of Maryland and the Smithsonian Institution.

The project is also related to the effort to digitize the complete Type Specimen Collection of the US National Herbarium (Department of Botany, Smithsonian Institution).

These activities build on the system for automated recognition of plants, which is being developed by Belhumeur et al. [2].



Figure 1: A computer vision system for identifying plants on Plummers Island, Maryland, USA. (Source: [2])

Basically, the idea of Leafsnap is very close to the project presented in this thesis. From my point of view, the main disadvantage of Leafsnap is, that the application needs a decent Internet connection to work properly. The image is sent to a server, where the recognition takes place. Without Internet connection, the recognition is not available, and with a poor signal, the upload takes too much time. That is a weak point for plant recognition software, because there are often problems with Internet connection in nature and conservation areas. In contrast, the Plant Ident software developed in this thesis implements the

image recognition in the mobile phone, which makes it serviceable even without Internet connection.

More differences:

- Leafsnap requires a fully-white background when taking photo for leaf identification, in PlantIdent leaf segmentation allows arbitrary background.
- Leafsnap uses only leaf-based recognition, PlantIdent also implements a manual guide and bark recognition.

On the other side, the Leafsnap application is well designed, user friendly and has a great database.



Figure 2: Leafsnapp app for iPhone

Encyclopedia of Life

Encyclopedia of Life (EOL.org) is an online encyclopedia trying to cover all life forms on our planet. According to the authors, the vision of this project is a global access to knowledge about life on Earth, which should increase awareness and understanding of living nature.



Figure 3: Encyclopedia of Life

2.1 Image-based Plant Identification

Up to the present, various computer-vision methods were developed or tested for plant identification, based on various photo types. For now, the leaf identification seems to be most popular, on top of that - the LeafSnap project is based on this.

Let me try to summarize possible image sources with potential existing methods and a short assessment.

2.1.1 Leaves

Although the leaf-recognition seems to be the most popular and implemented, in most cases it focuses only on broad leaves. But there is an effort to recognize needles too.

Broad leaves

As mentioned above, Sixta [1] researched broad leaves recognition. According to the requirement to work nearly realtime and precisely, he chose IDSC (Inner Distance Shape Context) proposed by Belhumeur et al. [2] as the best solution.

Belhumeur proposes an automatic segmentation with color-based EM algorithm, as white background is expected on the photos. Sixta admits different backgrounds and uses a semi-automatic segmentation method: marker-controlled watershed algorithm applied on a greyscaled image.

Other solutions for leaves broad recognition considered by Sixta were:

- using probabilistic neural network to recognize greyscale converted images according to 12 digital morphological features, which are derived from 5 basic features: diameter, physiological length, physiological width, leaf area, leaf perimeter (proposed by Wu et al. [3]).

In this work, white background is assumed. Segmentation was then done by thresholding the histogram of the greyscaled image with a fixed threshold (0.95).

- using the Kohonen SOM neural network to recognize leaves described by geometry features (ratio of the leaf length to leaf width, seven 2D movements invariants) and texture features (statistical moments from an image transformed using Discrete Wavelet Transform, at different scales). (proposed by Zhang et al. [4])

To get the contour here, image-threshold operation is applied to the grey image to obtain the binary image of leaf shape, which is further traced to produce the contour of leaf by making use of the border tracing algorithm.

- describing leaves by feature vectors containing 7 Hu moments [5] and 16 Zernike orthogonal moments [6], classifying by the moving centre hypersphere classifier (proposed by Wang et al. [7])

This work focuses on images with complicated background. To distinguish leaf from the background, an automatic marker-controlled watershed segmentation algorithm is used.

- describing leaves using SIFT (Scale Invariant Feature Transform, [9]), which provides features invariant to image translation, scaling, and rotation. (proposed by Fiel, Sablatnig [8])

Additionally, SIFT features are partially invariant to illumination changes and affine or 3D projection, and this method does not require segmentation. Multi-class SVM is used for classifying.

- fractal analysis of leaf-texture properties (proposed by Rossatto et al. [10]), which was created for identification of species from Miconieae tribe. A Bayes classifier with CDA (Canonical Discriminant Analysis - a method that removes the correlations among features and optimizes the separation between classes) is used for classification.

Xiao et al. [11] propose a new approach for plant leaf classification, which treat histogram of oriented gradients (HOG) as a new representation of shape, and use the Maximum Margin Criterion (MMC) for dimensionality reduction.

Compared to IDSC on Swedish leaf dataset and ICL dataset, this classification method seems to achieve better performance.

Algorithm\Dataset	S_ORI	S_NO	ICL_A_1	ICL_A_2	ICL_B_1	ICL_B_2	
IDSC	Average	93.73%	85.07%	95.79%	98.00%	63.99%	66.64%
	fluctuation	3.41%	4.15%	2.35%	7.02%	4.76%	11.36%
	variance	2.31E-04	5.12E-04	9.06E-05	9.98E-04	2.40E-04	2.04E-03
HOG-MMC	Average	93.17%	92.98%	96.63%	98.92%	83.35%	89.40%
	fluctuation	2.83%	3.91%	1.63%	6.92%	3.19%	11.40%
	variance	9.48E-05	1.49E-04	3.53E-05	2.22E-04	1.20E-04	1.77E-03

Figure 4: Recognition rates of Xiao et al. (Source: [11])

Needles

Fiel, Sablatnig [8] look for a solution for identification of coniferous trees - identification of needle images. It is important to highlight, that the identification photo should not cover only some needles, but the ending of a branch. Some clusters of needles should be seen there, because the arrangement of needles on the branch is pivotal. Additionally, a monochrome background is expected.

Their dataset contains 6 species of conifers common in Austria: Black pine, Fir, Larch, Scots pine, Spruce and Swiss stone pine. Those are typical Czech conifers too.

First, the species are manually divided into two "classes", which will be treated separately:

- trees on which every needle grow separately on the branch

The proposed way to differentiate the fir and the spruce is by looking at the endings of the needles and the adjustment of the needles on the branch. The spruce needles are blunt and grow only in one plane on the branch whereas the fir needles are pointed and they can grow in every direction.

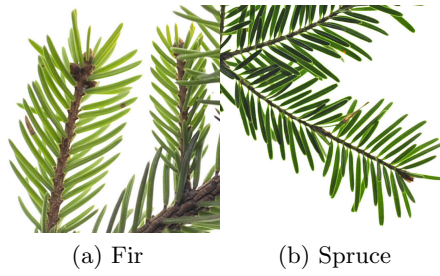


Figure 5: Species from first "class" (Source: [8])

Needle endpoints are found by calculating the skeleton of the branch and needles (separated from the background using a segmentation algorithm). Endpoints at the border of the image and endpoints lying too close to each other are skipped.

Remaining endings are analysed by calculating features like eccentricity, solidity, curvature features, and the improved moment invariants by Chen [12].

- trees on which the needles grow in clusters on the branch

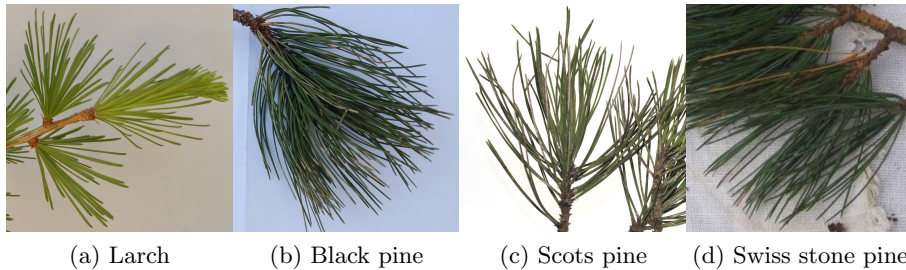


Figure 6: Species from second "class" (Source: [8])

For this "class" the number of needles in the cluster is the most important characteristic.

(Larch has a high number of needles in such a cluster, for Black Pine and Scots Pine the needles are always pairwise, Swiss Stone Pine has usually five needles in each cluster.)

The authors propose to use the same methods as described above to find the endings of the needles. Nevertheless they observe, that even though the endings have been found it is not possible to determine the number of needles of one cluster due to overlapping needles and more than one cluster at nearly the same location.

Conclusion of Fiel, Sablatnig needle-identification research:

No method has been found for the classification of the needle images. A method has been presented to distinguish between fir and spruce because the needles of these two trees differ. The images have to be in a good quality, which means that the endings of the needles must not be blurred. The difference between the needles of the other conifers are only the number of needles which grow in one cluster on the branch and due to overlapping needles these can not be counted.

2.1.2 Bark

Plants with bark include trees, woody vines, and shrubs. Bark recognition software usually focuses on trees, as their stems are thick enough.

Bark is usually described using texture features.

Sixta [1] also made a research into bark recognition.

He has considered the following methods:

- recognition based on co-occurrence matrix or long connection length emphasis (both proposed by Zheru et al. [13])
They present an average accuracy reaching 96% for 8 species (*Bombax ceiba*, *Roystonea regia*, *Lagerstoemia speciose*, *Magnolia grandiflora*, *Alhizzia Julihrrissin*, *Viburnum odoratissimum*, *Sapindus mukorossi*, *Podocarpus macrophyllus*).
- using the active appearance modelling (AAM) for bark classification (proposed by Dahl et al. [14], this work focuses on recognition by bark from images of logs)

Thereafter, he proposed a method to recognize bark using Multi-scale Local Binary Pattern (MB- LBP) on greyscaled bark images.

	Ash	Beech	Black pine	Fir	Hornbeam	Larch	Mountain oak	Scots pine	Spruce	S.s. pine	S. maple
Ash	5				1		5		1		
Beech		3									
Black pine			61			7		6		4	
Fir				47		7	1	1	3		
Hornbeam			1		12	1				2	
Larch			6			71		15	1	3	
Mountain oak					2	2	26	1	1	2	
Scots pine			6			31		49	1	4	
Spruce				2		2	3	6	82	6	
Swiss stone pine			2		2	7	5	6	2	19	
Sycamore maple									1	2	3

Figure 7: Sixta: Confusion matrix for 11 class experiment, rows are true classes, columns estimated (Source: [1])

Fiel, Sablatnig [8] also deal with bark recognition. They propose to describe bark texture using the same method as they proposed for leaf recognition: SIFT (Scale Invariant Feature Transform). In comparison to an approach of a combination of GLCM and wavelets, SIFT reaches better scores on their dataset (consisting 1183 images of 11 species).

2.1.3 Flowers

Although Sixta [1] did not use or implement any flower-based plant identification, he mentions works of Saitoh et al., Pornpanomchai et al., Cho et al., Kim et al. and Nilsback.

Saitoh et al. [15], [16] describe a flower using 8 features, which will be further described below.

First of all, such a photo of flower (blooming part) is taken, where a black sheet with a hole in the centre is placed behind the flower. This enables easier separation from the background and it also minimizes the amount of sun's ray reflection. The separation process uses the k-means algorithm for clustering the background part, which should provide better results than using single threshold. From the separated flower, following features are calculated:

- Shape features:

1. The ratio of the average width over the average height l/h
2. The number of petals, where ‘ N ’ represents ‘many’ in case of more than 7 petals
3. A central moment defined as

$$M = \frac{1}{S} \sum_m \sum_n \sqrt{(m - m_g)^2 + (n - n_g)^2} f(m, n)$$

where $f(m, n) = 1$ if the coordinate (m, n) is inside the perimeter and $f(m, n) = 0$ otherwise; S is the total number of pixels in the flower region and (m_g, n_g) is the c.

4. Roundness, which is defined as

$$R = \frac{4rS}{L^2}$$

where S is the total area of the flower region and L is the total perimeter length. We have $0 < R < 1$.

- Colour features:

At first, the image is converted from the RGB into the HSI (or HSV) coordinate representation. The HS space is then divided into 13 segments and the largest and the second largest segments are taken.

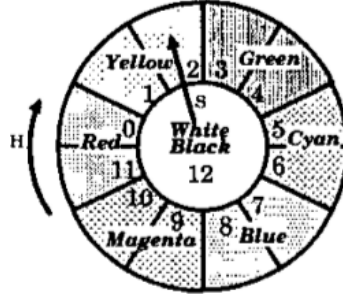


Figure 8: Saitoh: Separate HS space (Source: [15])

The features are:

5. the largest segment number,
6. its distributed value,
7. the second largest segment number,
8. its distributed value.

The recognition is done using neural networks and it is based on knowledge of both flower- and leaf-features (which were also proposed in this thesis).

After the recognition experiments, it was found that only four features out of the above eight are sufficient for flower recognition.

Pornpanomchai et al. [17] build an automatic method for recognizing a Thai herb flower based on the Minimum Distance Method.

There are five features used for description:

average red, green and blue (RGB) colours, the herb flower size and the edge of petals feature.

Experiments are done on 16 species of herb flowers. The system was tested by using 110 pictures for a training data set and 50 pictures for an un-training data set. The precision rates of the recognition system were 98.18 percent for the training dataset and 94 percent for the un-training dataset.

The presented average access time is 0.87 seconds per image, but further details about testing conditions (HW, SW) were not mentioned.

In a later work [18] Pornpanomchai et al. present Leaf and Flower Recognition System called e-Botanist.

A remarkable modification is adding a ping-pong ball aside the image to compare the size. This system also uses more features: Height ratio, area ratio, width and height ratio; roundness value; half flower area ratio, boundary feature and 6 colour subfeatures. The record with smallest Euclidean distance is returned as result.

The system achieves better results when using the "popular" features instead of the average of all features. In this case, the flowers recognition precision is around 75 percent.

The system was conducted on 25 species of flowers and the presented average access time was around 6 seconds/image. Some slowness could be attributed to the fact, that the system was implemented in Matlab.

Cho et al. [19] also propose a system for flower recognition based on content based image retrieval (CBIR). At first, a manual segmentation is done using the so-called Magnetic Lasso Tool. Thereafter twelve shape attributes (based on the shape density distribution and the edge density distribution) and six colour attributes (from HSV coordinate system) are extracted from a flower image to form a feature vector. A self-organizing algorithm

called Virus Infection Clustering is used for database clustering to improve the searching efficiency. The obtained results showed that clustering by the colour and shape features yields better retrieval results than clustering by only either colour or shape individually.

Kim et al. [20] suggest and implement a mobile-based flower recognition system using Difference Image Entropy (DIE) and contour features of the flower from the original image with multi-flower objects. An image is taken by the user using the mobile device. Then the "flower image" and "contour image" is sent to the server, where contour feature extraction is done, difference images are computed using pixel subtraction between the flower image and the normalized average flower images. Thereafter, classification is done using the minimum difference image entropy on the server side and the results are sent back to the mobile device.

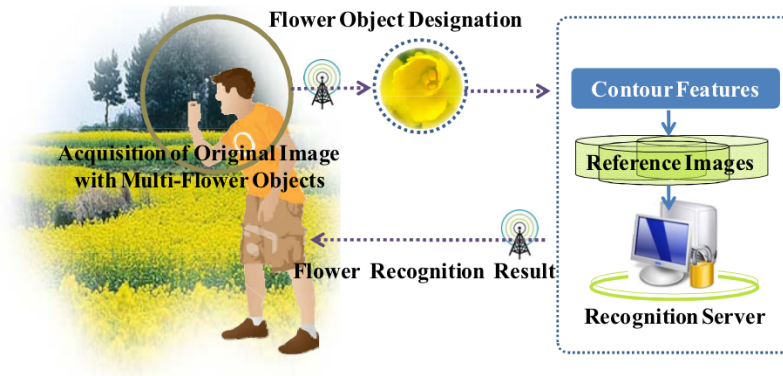


Figure 9: The overall system architecture of Kim et al. (Source: [20])

Nilsback presents an extensive work [21] on the topic of flower classification from images. An iterative segmentation scheme is developed, which is able to segment a flower given only a general colour model for flowers and no information about the flower in a particular image. This is done by iteratively fitting a geometric model and updating the colour model. The segmentation at each iteration is obtained by minimizing a Conditional Random Field using graphcuts. As the flower is extracted, it is described using features like HSV values, MR8 filters, SIFT or histogram of gradients (HOG).

A multi-class SVM is used for classification.

Nilsback also created flower databases / datasets.

2.1.4 Fruit, Cones

Next typical characteristics of plants may be fruit or cones. Although they may offer expressive optical differences, it is a seasonal matter, which is a disadvantage for plant species recognition.

Arivazhagan et al. [22] summarize: *The computer vision strategies used to recognize a fruit rely on four basic features which characterize the object: intensity, colour, shape and texture.*

They propose an efficient fusion of colour and texture features for fruit recognition based upon the statistical and co-occurrence features derived from the Wavelet transformed sub-bands and use the minimum distance classifier for recognition.

Image processing methods can be also applied for the fruit quality control in the food industry or for developing automatic fruit picking systems.

Zhao et al. [23] propose On-tree Fruit Recognition for automatic fruit (apples) picking.



Figure 10: Fruit picking system of Zhao et al. (Source: [23])

2.1.5 The overall plant image, shape, contour

The overall image could also be a source for plant identification. Sometimes people try to identify trees by their silhouette.

Still, trees and plants generally can have various body structure and silhouette strongly depending on the local conditions. That is why contour- or silhouette- based recognition would be quite a complicated task.

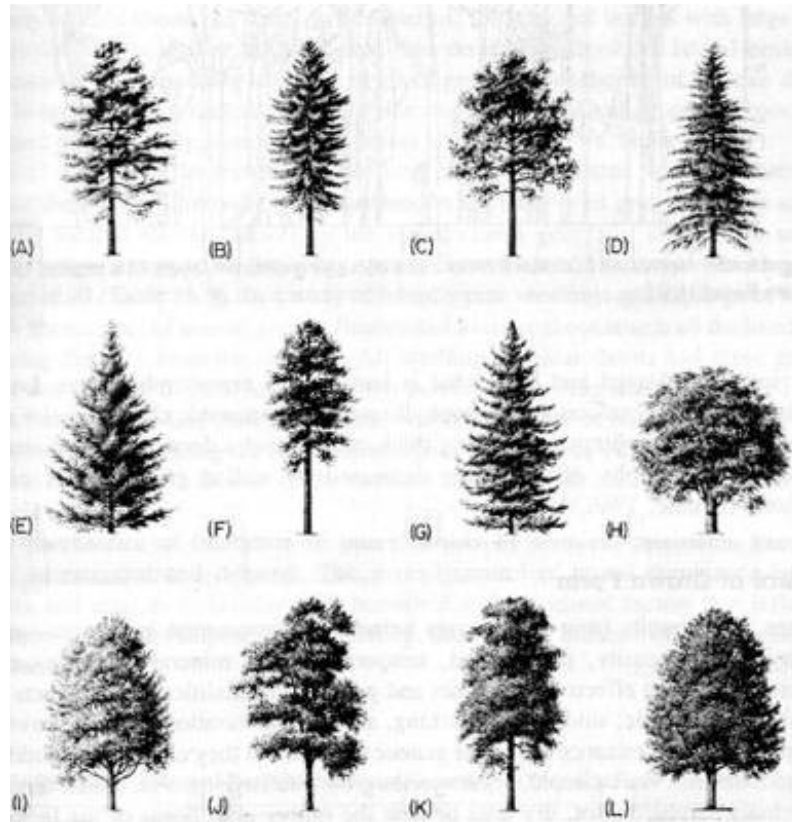


Figure 11: The USFS-TAMU clip art (forestry.about.com): Tree silhouettes

3 Identification from needles

As we can see from the prior exploration, even though conifers (lat. Pinopsida) are a widespread class of plants, there has not been found a usable method for conifers-recognition based on needles. (Of course, bark recognition can be used instead for coniferous trees.)

Nevertheless, according to the Report on Forestry of the Czech Republic [24], conifers currently constitute over 73% of the forest land area in the Czech Republic.

Therefore I decided to make some research and experiments about needles-based identification. From the very beginning I have to mention, that "Identification from needles" is a simplistic title, because the recognition algorithm proposed in this thesis does not work with images of single needles. As mentioned by Fiel, Sablatnig [8], the needle arrangement on the branch is a very important information. Therefore, an image of branch with needles is taken for recognition.

In the above mentioned work, it has been observed, that the needles segmentation or counting is a very complicated issue. The arrangement of needles is usually not optimal and they are often hiding or overlapping eachother.

That is why I will not focus on finding individual needles or needle endpoints, but on convenient description of needles as a texture. The texture pattern direction on both sides of the branch should provide be a valuable description.

For that, the knowledge of branch position in the image will be a relevant information.

3.1 Datasets

As far as I know, there is no public available dataset for needles-based (or coniferous-branches-based) image recognition.

3.1.1 Österreichische Bundesforste/Archiv and Computer Vision Lab, Vienna UT

Nonetheless, Fiel, Sablatnig in the work [8] used such a dataset, but it has not been published yet. Luckily it has been made accessible to me by Mr. Fiel, so that - with

courtesy by Österreichische Bundesforste/Archiv and Computer Vision Lab, Vienna UT
- I can use it for academic purposes.

This dataset consists of diverse photos of coniferous branch-endings on white or natural (inhomogeneous) background, and it covers the following species:

- Black pine
- Fir
- Larch
- Scots pine
- Spruce
- Swiss stone pine

3.1.2 The CB Dataset

For the purposes of development and testing, I also photographed, gathered and labelled images of coniferous branches and created the Coniferous Branches Dataset (CB Dataset), which is also an attachment to this thesis.

This dataset focuses primarily on trees typically found in the Czech Republic. The Report on Forestry of the Czech Republic [24] says the most common conifers (according to the % of forest land area they constitute) are spruce (51.92%), pine (16.81%), larch (3.88%) and fir (1%).

Considering this, I decided to make a dataset containing the following species:

- Norway Spruce (*Picea abies*)
- Scots Pine (*Pinus sylvestris*)
- European Larch (*Larix decidua*)
- Grand Fir (*Abies grandis*)

Although the Grand Fir (*Abies grandis*) is not an originally European tree, as it was imported from North America, today it is commonly found in Czech region. Since it is less sensitive to the air pollution than the European Silver Fir (*Abies alba*), it is easier found near urban areas or in city parks.

Unfortunately, as I am not a dendrologist, I can not confidently label a tree species, since many related species have a very similar characteristics. Therefore, I photographed only labelled trees.

Photos of those trees were taken in the areal of the Czech University of Life Sciences in Prague and in the Botanical garden of the Faculty of Science, Charles University in Prague. In both cases, trees were labelled by dendrologists from the corresponding university.

All images represent a coniferous branch on an (almost) white background.

Spruce (20):



Pine (25):



Larch (31):



Fir (17):



Figure 12: The CB Dataset

3.2 The texture Pattern Direction (TPD) description

As mentioned before, a description of "texture direction" would be very helpful in the intention of needles-based plant recognition. Therefore I created an operator to describe the texture pattern direction.

The introduced Texture Pattern Direction operator gives a very simple, but applicable description of an image. For every foreground pixel it either returns one of four directions or decides, that the direction is not definite.

In this case, white background is expected behind the coniferous branch. That is why all "non-white" (with reserve) pixels can be taken as foreground.

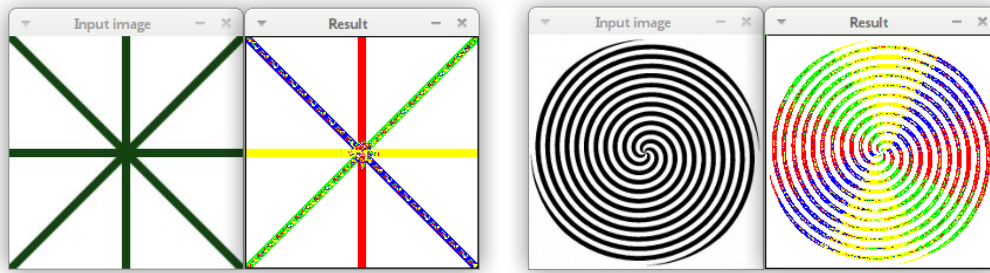


Figure 13: Texture Pattern Directions - Visualisation

3.2.1 The definition of TPD

Lets choose a pixel on position (x, y) with colour C in the image. Neighbouring pixels with colours C_1, \dots, C_8 are used for texture direction determination, labelled as follows:

C_8	C_1	C_3
C_6	C	C_5
C_4	C_2	C_7

Figure 14: Neighbouring pixel colours labels for Texture Pattern Directions

From those, the total of colour differences from the centre is computed (as the distance in 3-dimensional RGB space) for every direction.

$$\delta_1 = |C - C_1| + |C - C_2|$$

$$\delta_2 = |C - C_3| + |C - C_4|$$

$$\delta_3 = |C - C_5| + |C - C_6|$$

$$\delta_4 = |C - C_7| + |C - C_8|$$

Then the maximum differences total δ_m is found and directions m is returned as the result of TPD operator.

If there is not an unequivocal maximum, zero is returned.

3.2.2 Application of TPD on images of coniferous branches

As said before, I defined the Texture Pattern Direction to use it for needles-based recognition. Lets see the result of its application on images of coniferous branches from my dataset:

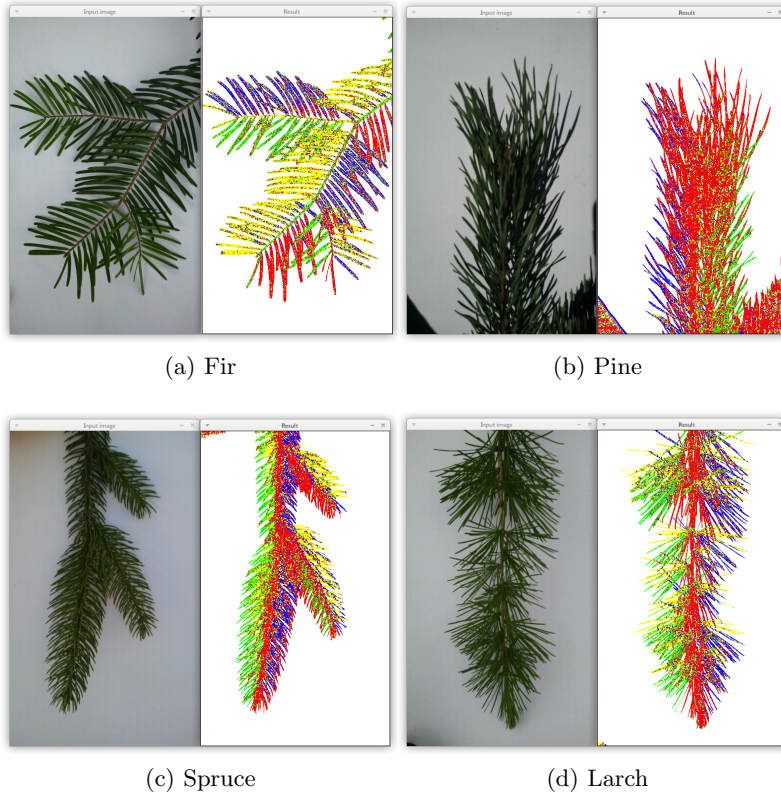


Figure 15: TPD applied od images of coniferous branches

The experiments have shown, that TPD description denotes the direction of needles on the branch.

3.3 The branch shape search

The branch position and shape will be a useful information for detecting further description features.

The start- and end-point of the usable part of branch can be selected by user, as - from experience - its very comfortable for users of touchscreen phones to mark a few points in the image.

With the knowledge of those points, the branch shape can be found using dynamic programming. First, a grid of points will be computed between the start- and end-point, since it would be needless to go through every pixel.

Then every point p_i of this grid will be evaluated with a price $P(p_i)$, which will consist of unary term $f(p_i)$ and the minimal binary term $g(p_i, p_j)$ over all points p_j from the previous level (points row).

$$P(p_i) = f(p_i) + \min_j g_j(p_i)$$

The unary term $f(p_i)$ is a sum of texture incongruence with the angle to the start-point, texture incongruence with the angle to the end-point, and the Euclidean distance to the end-point.

The binary $g(p_i, p_j)$ term is a sum of the price $P(p_j)$, Euclidean distance between p_i and p_j , and the squared difference between current and previous branch-segment angle.

Texture incongruence with an angle

To get the texture incongruence with an angle, as used in the unary term, a histogram of TPD descriptors in a circle around the given point is used.

As the histogram consist of four orientations, where each can be understood as two angles, a difference Δ from given angle α is the smaller angle difference.

$$\Delta_d = \min_k \left| \frac{(d-1)\pi}{4} + k \cdot \pi - \alpha \right|$$

Now, if the histogram frequency in direction d is h_d , the texture incogurence with an angle α can be expressed as

$$T(\alpha) = \sum_d \Delta_d \cdot h_d$$

To accentuate the negative impact of bigger incogurence, $T^2(\alpha)$ is used for this feature.

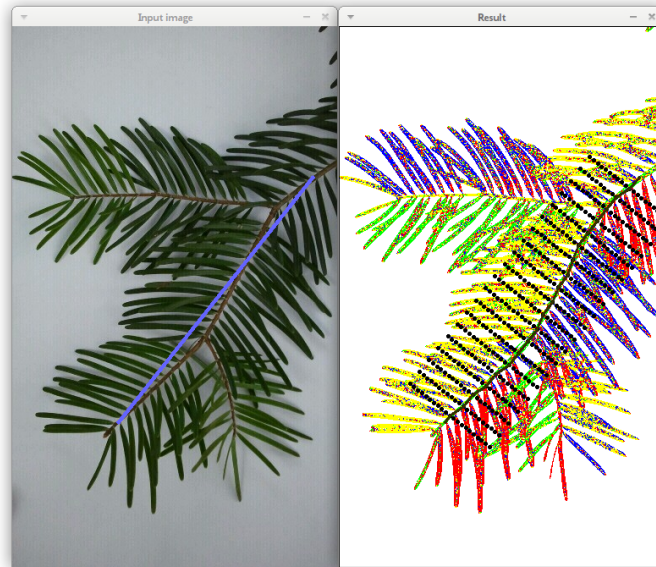


Figure 16: Finding the branch shape

3.4 Detecting the width of needle texture area

The next step after finding the branch shape should be finding the useful area (in the branch surroundings) for needles description. Such area is given by shifting the local branch part on its normal. The branch does not have to be divided into as many parts as the number of grid levels (from shape finding), a lower number will be sufficient. Since we can easily tell the foreground from the background (as almost white colour is expected for the background), the background-occurrence can be used to find the area width W .

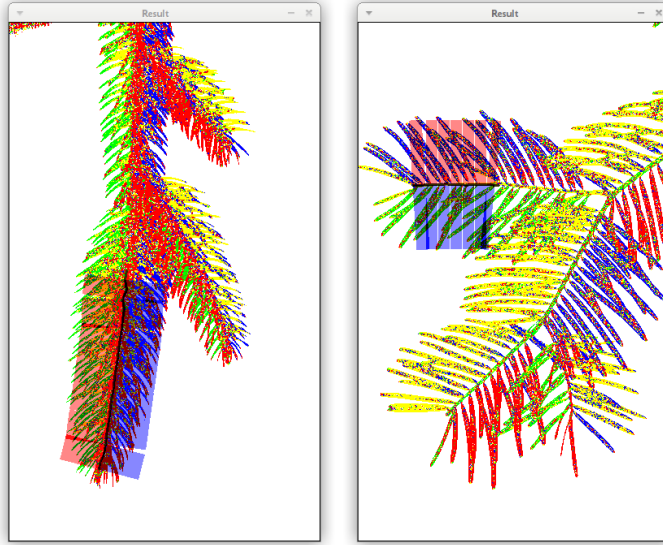


Figure 17: Finding the area width

Through experiments I found the following method to be working fairly well:

The minimum width W_0 is given. Thereafter, rectangles with width W_i are being added, since one of the following conditions is true:

- the background percentage b_i in rectangle i is higher than 95%
- the background percentage b_i in rectangle i is higher than 70% and higher than 125% of the "expected value" e_i , where

$$e_i = \frac{e_{i-1} + b_{i-1}}{2}$$

The final area width is then

$$W = W_0 + \sum_i W_i$$

3.5 Normalizing the TPD description of needles

Now, for every area described before, a TPD histogram can be made, where for every orientation the number of corresponding TPD values is calculated. But because of different branch-part-angles, the histograms don't overlap, as their frequencies (angle differences from the branch-part-angle) differ.

The histogram is shifted to normalized frequencies (angles) $0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$ using polynomial interpolation. Specifically, Lagrange polynomials of 3rd degree are used for interpolation:

$$L(x) = h_1 \cdot l_1(x) + h_2 \cdot l_2(x) + h_3 \cdot l_3(x) + h_4 \cdot l_4(x)$$

$$l_1(x) = \frac{(x - a_2)(x - a_3)(x - a_4)}{(a_1 - a_2)(a_1 - a_3)(a_1 - a_4)} \quad l_2(x) = \frac{(x - a_1)(x - a_3)(x - a_4)}{(a_2 - a_1)(a_2 - a_3)(a_2 - a_4)}$$

$$l_3(x) = \frac{(x - a_1)(x - a_2)(x - a_4)}{(a_3 - a_1)(a_3 - a_2)(a_3 - a_4)} \quad l_4(x) = \frac{(x - a_1)(x - a_2)(x - a_3)}{(a_4 - a_1)(a_4 - a_2)(a_4 - a_3)}$$

After the shift, for each side of the branch a histogram can be counted as the sum of the shifted partial histograms. Because the direction of branch is unequivocal, as its end and beginning may be mistaken, next normalisation takes place. The second and fourth value of the histogram (both in the distance of $\frac{\pi}{4}$ from direction of the normal) are, if need be, replaced, so that the higher value is on the second place. This normalisation also solves the opposite histograms on different sides of the branch, so only one histogram can be made as their total.

Finally, the histogram values are divided by the highest one, so that, after this adjustment, the highest value will be 1.

That is how we will get a descriptor (a histogram that can be represented as a 4-D vector), which will be used for classification.

3.6 Classification and evaluation

For the purposes of testing, the CB Dataset was split in half, into a training and testing dataset. The nearest neighbour algorithm (NN) is used for classification.

The results of classification between 4 species (classes) are:

34 conifers were classified right, 13 were classified wrong, which means 72.3% precision.

	Fir	Spruce	Larch	Pine
Fir	7	0	0	2
Spruce	1	6	3	0
Larch	0	2	12	1
Pine	1	3	0	9

Table 1: Confusion table for conifers classification

4 Implementation

4.1 The Android Application

The mobile PlantIdent application is being developed for the Android operating system. For development and testing purposes, (HTC) Google Nexus One and Samsung Galaxy S Plus were used.



Figure 18: Left: Google Nexus One (HTC). Right: Samsung Galaxy S Plus

4.1.1 Fundamental Layout

For the draft of the Graphical User Interface for PlantIdent, three main operation groups were considered:

- Photo taking and editing
- Guide, going through database
- Album of taken photos

For simple and transparent switching between those groups, Android Tab Layout was chosen to be the fundamental layout of this application.

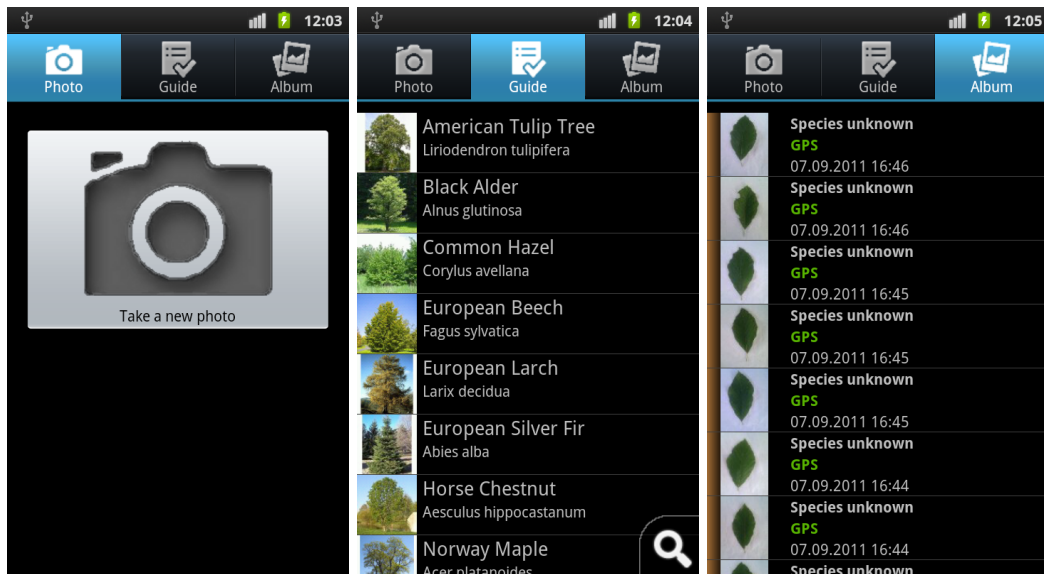


Figure 19: PlantIdent: Tab Layout - Photo Tab, Guide Tab, Album Tab

4.1.2 Identification using Image Recognition

The application allows plant identification using leaf-, bark- or needles-based image recognition, where leaf- and bark-based computer vision methods are based on the diploma thesis of Sixta [1].

Leaf-based recognition

- The centre of the image is assumed to be foreground (leaf), the borders are assumed to be the background (surface)
- Watershed segmentation (followed by hole-filling algorithm) is applied on a greyscaled copy of the image.
- User is asked to visually check the segmentation results and alternatively correct the foreground- and background- markers to get a new segmentation.
- Egg-shaped pattern CCD (Centroid Contour Distance) curve is used to determine leaf orientation
- 64 IDSC (Inner Distance Shape Context) descriptors are created
- Nearest neighbour classifier is applied for comparison with dataset descriptors to get scores used for sorting results

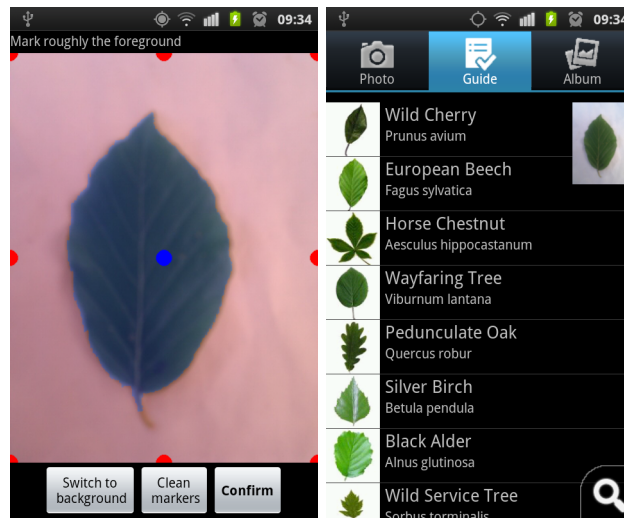


Figure 20: PlantIdent: Leaf recognition

Bark-based recognition

- User is asked to select the applicable area of the photo
- LBP (Local binary patterns) algorithm is used to create pattern descriptors for bark
- Nearest neighbour classifier is applied for comparison with dataset descriptors to get scores used for sorting results

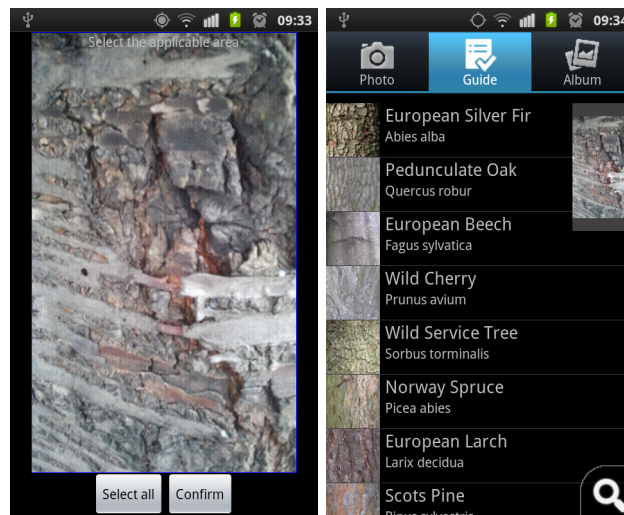


Figure 21: PlantIdent: Bark recognition

Needles-based recognition

- User is asked to select the start- and end-point of the branch
- The branch shape and applicable surroundings are found
- The 4-D vector describing the Texture Pattern Direction towards the local branch direction is created
- Nearest neighbour classifier is applied for comparison with dataset descriptors to get scores used for sorting results

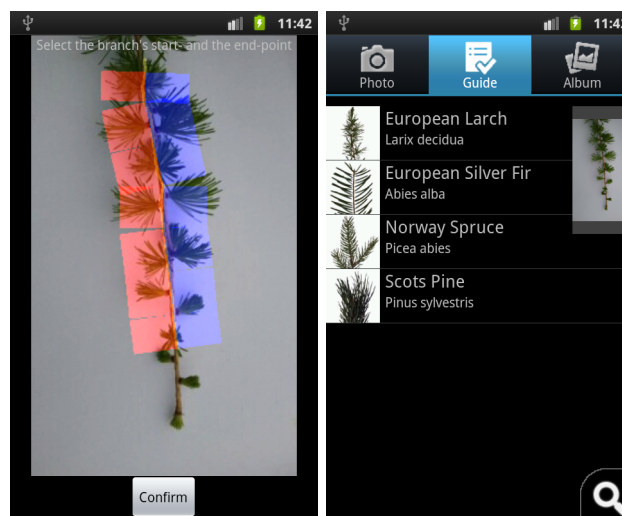


Figure 22: PlantIdent: Needles recognition

After the recognition process, application switches to the Guide Tab and shows sorted results with thumbnail. A preview of identified image is also shown for comparison.

Unlike the rest of the application, which is written only in Java using Android SDK, the descriptor creation and comparison is implemented in libraries developed in C++ using Android NDK (Native Development Kit) and the OpenCV library.

4.1.3 Manual Field Guide

Botanists, as well as non-expert users, usually work with "paper" identification keys or field guides. Such tools are usually in form of a dichotomous key, in better case it has images to make users find easier.

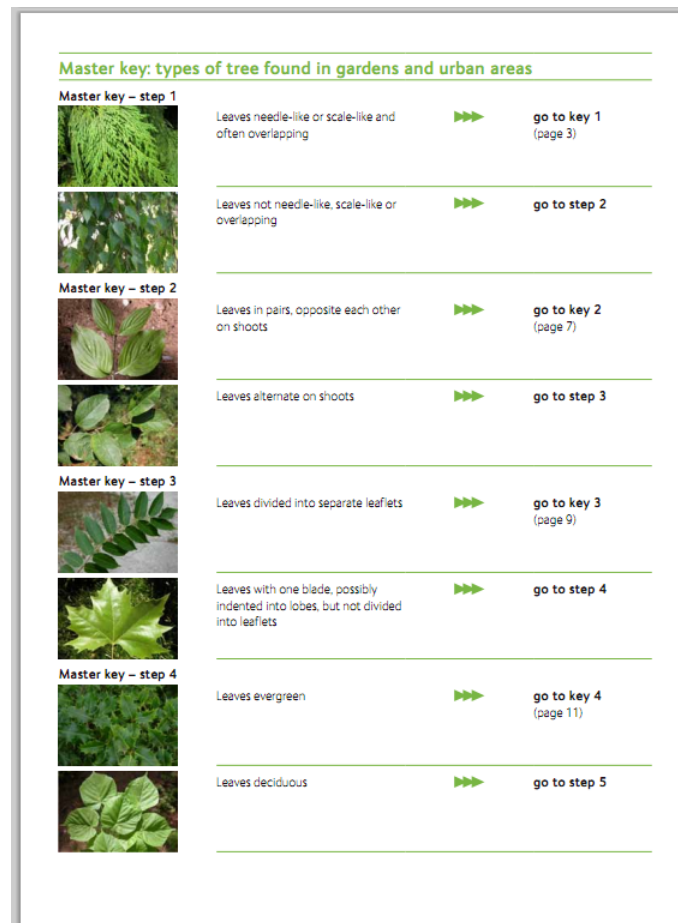


Figure 23: A transparent identification key with images,
 The Natural History Museum Cromwell Road London, UK
www.nhm.ac.uk/urban-tree-survey

In contrast with such "paper" keys and guides, an electronic field guide can provide more comfort and usability, because it can be designed to:

- choose multiple answers for one question
- change previous answers while leaving the other as they are
- automatically change shown plants after new answer (saves time)
- easily get more information about any of shown plants

The field guide in PlantIdent was implemented with focus on these improvements.

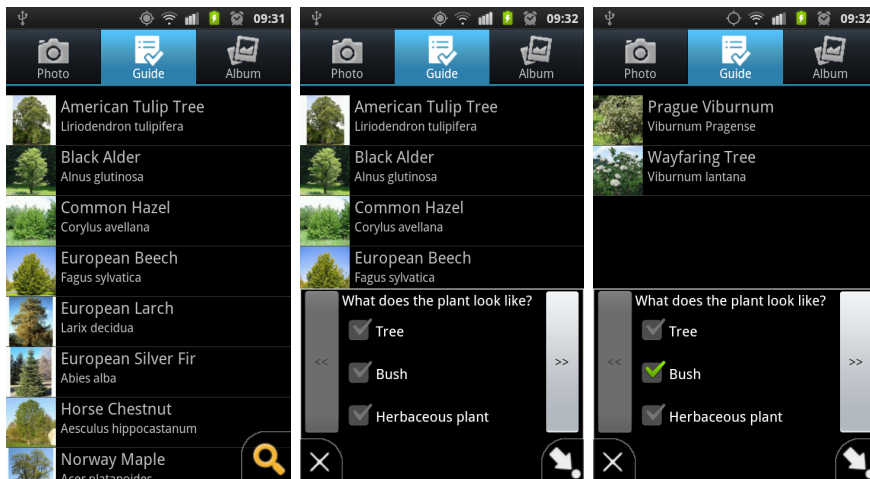


Figure 24: PlantIdent: Guide

4.1.4 Album, Sharing, Map

In the Album Tab all saved plant photos are listed. For every listed plant, user can:

- change identification
- upload it to the PlantIdent web gallery
- show it on Map (if photo GPS is available)
- remove it from the Album

Uploading / sharing

To upload photos, an user account is needed. The registration is very easy and can be done both from the mobile phone and from the web site.

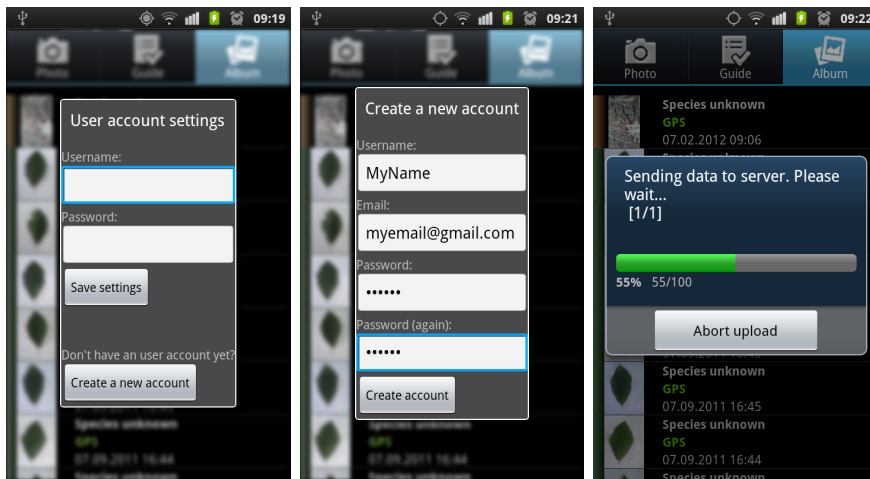


Figure 25: PlantIdent: Managing accounts and uploading

From the Album Tab, users can upload either the chosen photo or all new (=yet not uploaded) photos.

Map

When taking photos with PlantIdent, GPS coordinates are saved if possible. Those photos can be then shown on the Map.

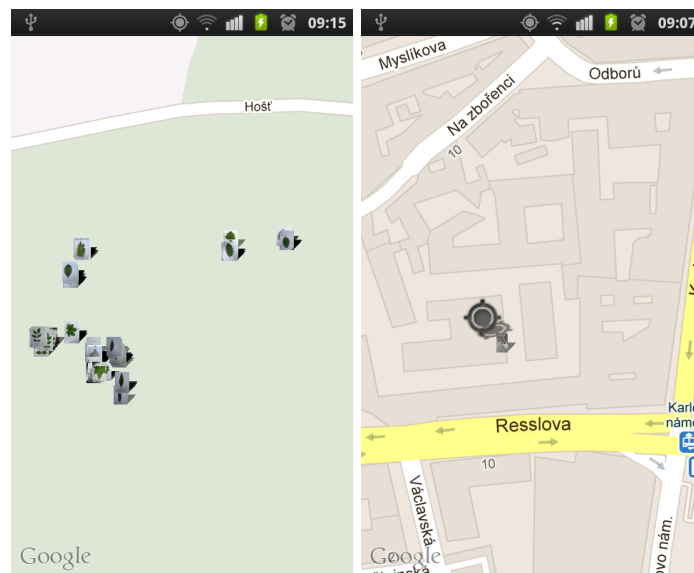


Figure 26: PlantIdent: Map - photos; photos with current location

Using Android Google Maps API, Plant Ident allows user to see a Google Map with photos from his Album.

Still, an application for plant identification can be often used far from the roads, which may also complicate orientation from a Google Map, when searching a plant. Therefore, if GPS signal is available, a pointer with users current position is shown.

4.2 Web

The Web gallery and presentation is another part of the Plant Ident project.

The main 3 fuctions, which are connected together, are:

- Database
- Gallery
- Map

The implementation uses PHP and MySQL on the server side, JavaScript and XHTML with CSS on the client side.

4.2.1 Database

The Database allows users and visitors to go trough the current PlantIdnet database and display available info about each plant.

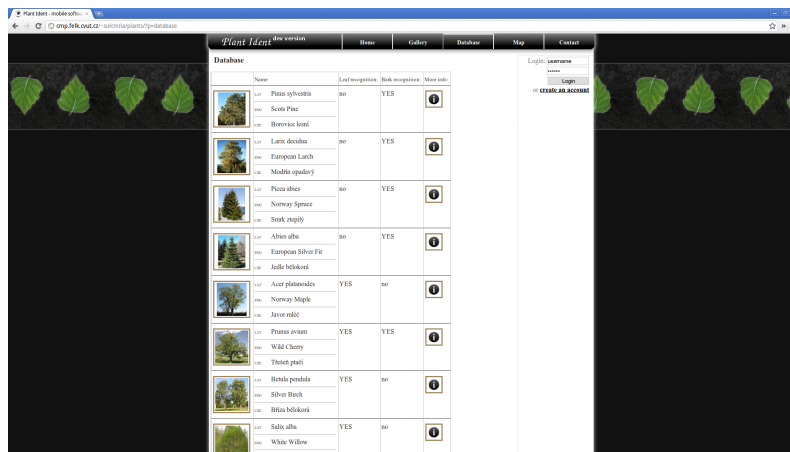


Figure 27: Web: Database

4.2.2 Gallery

In Gallery, all uploaded photos are listed. By choosing one, you open a detail. In this view, users can see also other photos of the same plant (marked by author) and comment photos.

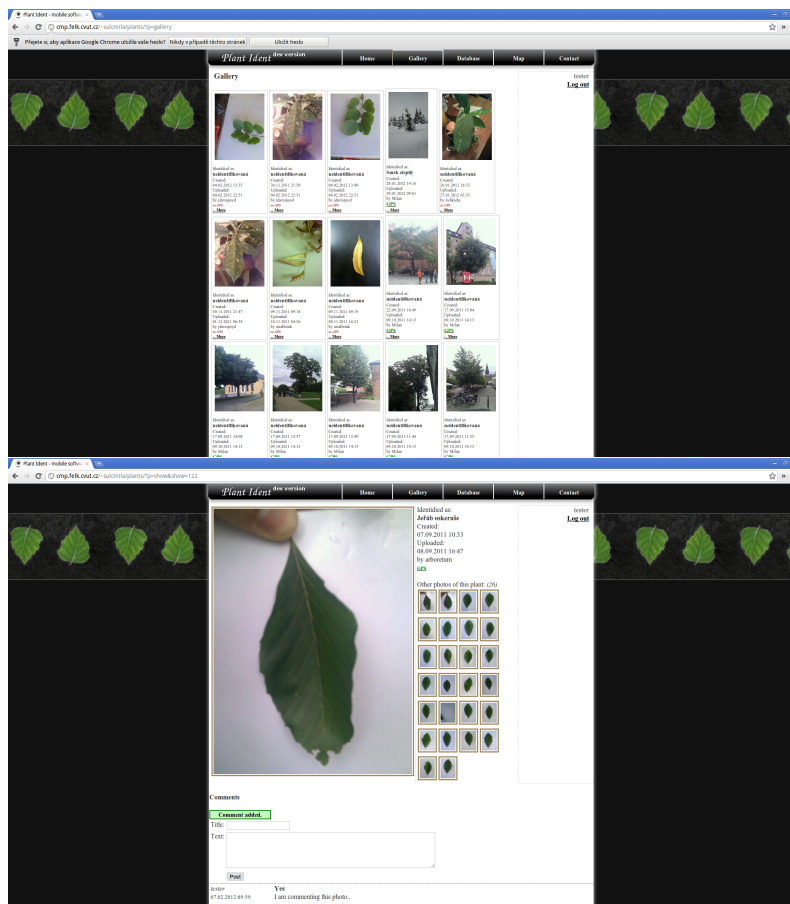


Figure 28: Web: Gallery

4.2.3 Map

All photos with GPS coordinates are marked on the Map, which was designed using Google Maps JavaScript API V3. This API should be faster than the previous one, more applicable to mobile devices allows, and among others also supports StreetView.

By clicking on a photo icon, an information window appears.

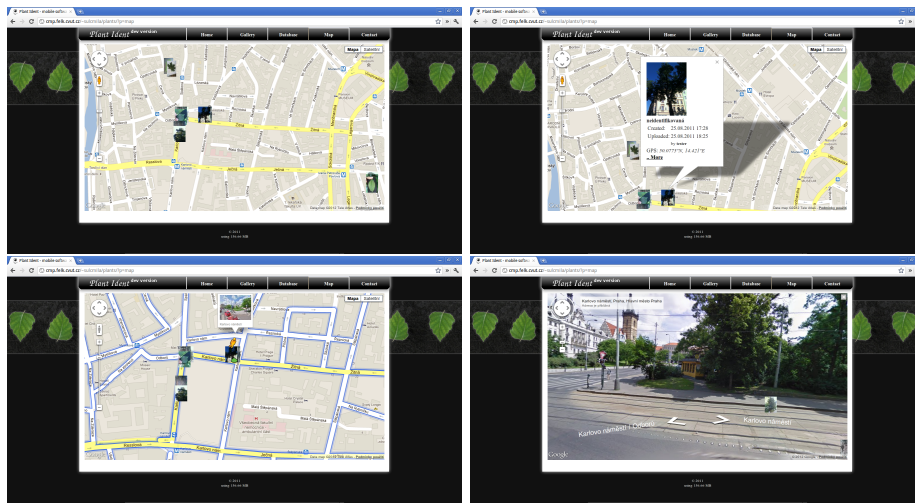


Figure 29: Web: Maps

4.3 Managing datasets

To manage datasets, a web administration a few helpful tools and scripts were developed.

4.3.1 Creating descriptors

Besides the program for needles-descriptor creation described above, two more programs have been written in C++ with OpenCV for leaf- and bark- descriptor creation.

Leaf descriptors

To create leaf descriptors from a photo, foreground and background has to be marked. After segmentation, the result is shown and the descriptor is saved into a txt file opened in a text editor.

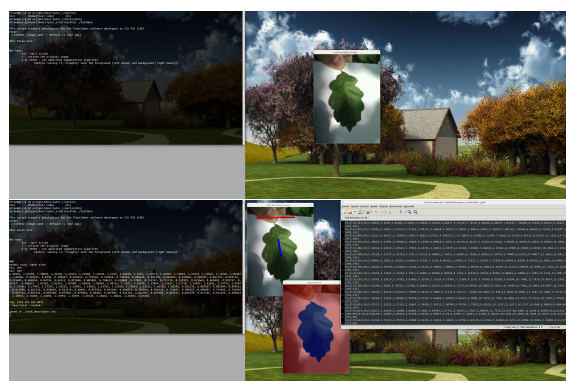


Figure 30: Creating leaf descriptors

Bark descriptors

To create bark descriptors from a photo, only the input photo is shown for check, after confirm the descriptors are saved into a txt file opened in a text editor. 12 descriptors for different scales are created for every bark photo.



Figure 31: Creating leaf descriptors

4.3.2 Web administration

For managing the database (plants, descriptors, characteristics, guide questions and answers) a web administration was created, which is accessible only for user with higher permissions. This database can then be used to generate datasets using PHP scripts.



Figure 32: Species administration

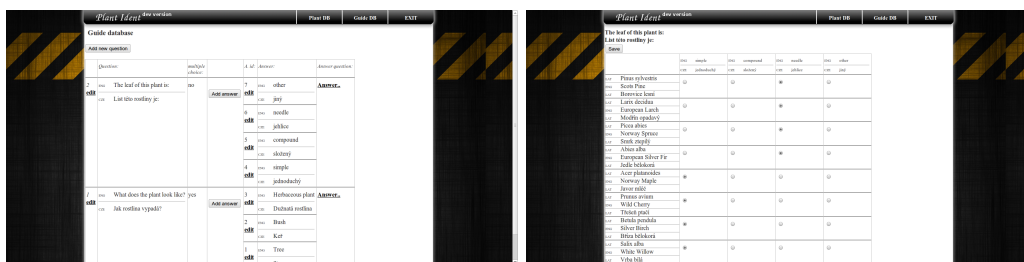


Figure 33: Guide administration

5 Conclusions

In this thesis I dealt with the computer vision methods for plant identification. For the purposes of conifer identification, I gathered a small dataset (CB Dataset) containing photographs of 4 conifer species (Norway Spruce, Scots Pine, European Larch, Grand Fir) and offered a method for their recognition based on images of needles (or more precisely images of coniferous branches). This method builds on the connection between local branch orientation and its surroundings texture pattern direction. The branch shape is found using dynamic programming, as the user marks the start- and end- point. The texture pattern direction is described through an operator I proposed in this thesis. The final descriptor is a 4-D vector, which allows a fast and simple classification executed by the nearest neighbour algorithm. A test on the CB Dataset shows 72.3% precision between 4 classes. This score could be further improved by using a wider dataset and by adding more characteristics features to the descriptor.

For the purposes of plant species recognition, species mapping and datasets gathering, an Android application was created. It builds on the computer vision methods for leaves and bark recognition proposed by T. Sixta [1] and the method for needles-recognition proposed in this thesis. It also offers a Manual Field Guide, a gallery with photographed data (with their GPS coordinates), a map with photo position illustration, an interface for uploading taken photos, etc.

The related website is also a part of the project. It allows users to see the uploaded photos, display them on a map, comment them or display the mobile applications dataset.

Recently, a researcher from the Czech University of Life Sciences Prague showed potential interest in our application, where both the use as a teaching aid and the use for their arboretum mapping are discussed.

We are also negotiating with a potential foreign investor.

5.1 Future work

In the future development, some other recognition methods may be implemented. These could focus on other image sources (blossoms, fruits, overall look, needles,..) or try to find more precise algorithms to process currently used image sources (leaves, bark or needles). For example, Xiao et al. [11] suggest leaf recognition based on histogram

of oriented gradients (HOG) representation with Maximum Margin Criterion (MMC) for dimensionality reduction. The needles-based recognition should be further researched, as certainly more describing features can be found. Also an improvement, which would allow recognition on non-homogeneous background, would be very practical. For further efforts, a wider datasets would be helpful, preferably labeled by a botanist or dendrologist.

As follows from discussion with the potential investor, an iPhone version of the application would be also appreciated.

With the expansion of tablets, a tablet version of the application could be also useful, as the larger screen can display further information. We have tested the iPad version of Leafsnap and, in my opinion, it offers a considerable comfort.

Cooperation with more institutions interested in plant identification or related studies will be a very helpful advance.

6 References

- [1] Sixta T.. *Image and Video-based Recognition of Natural Objects*. Czech Technical University in Prague, 2011.
- [2] Belhumeur P.N., Chen D., Feiner S., Jacobs D.W., Kress W.J., Ling H., Lopez I., Ramamoorthi R., Sheorey S., White S., Zhang L.. *Searching the World's Herbaria: A System for Visual Identification of Plant Species*. In European Conf. on Computer Vision, LNCS vol. 5305: 116–129. Springer, 2008.
- [3] Wu G.S., Bao F.S., Xu E.Y., Wang Y.X., Yi-Fan Chang Y.F., Xiang Q.L.. *A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network*. Signal Processing and Information Technology, 2007 IEEE International Symposium on.
- [4] Zhang L., Kong J., Zeng X., Ren J.. *Plant Species Identification Based on Neural Network*. Natural Computation, 2008. ICNC '08. Fourth International Conference on.
- [5] Hu M.K.. *Visual pattern recognition by moment invariants*. IRE Trans. Inform. Theory 8: 179–187, 1962.
- [6] Teague M.. *Image analysis via the general theory of moments*. JOSA, Vol. 70, Issue 8, pp. 920-930, 1980.
- [7] Wang X.F., Huang D.S., Du J.X., Xu H., Heutte L.. *Classification of plant leaf images with complicated background*. Appl. Math. Comput. 205(2): 916–926, 2008.
- [8] Fiel S., Sablatnig R.. *Automated identification of tree species from images of the bark, leaves or needles*. Technical Report, CVL-TR003, Computer Vision Lab, Institute of Computer Aided Automation, Vienna University of Technology, 2010.
- [9] Lowe D.G.. *Object recognition from local scale-invariant features*. In Proc. Seventh IEEE International Conference on Computer Vision The, 2: 1150–1157, 1999.
- [10] Rossatto D., Casanova D., Kolb R., Bruno O.. *Fractal analysis of leaf-texture properties as a tool for taxonomic and identification purposes: a case study with species*

- from *Neotropical Melastomataceae (Miconieae tribe)*. *Plant Systematics and Evolution* 291(1): 103-116, 2011.
- [11] Xue-Yang Xiao, Rongxiang Hu, Shan-Wen Zhang, Xiao-Feng Wang. *HOG-Based Approach for Leaf Classification*. Hefei, China, 2010.
- [12] Chau-Chin Chen. *Improved moment invariants for shape discrimination*. *Pattern Recognition*, 26(5):683–686, 1993.
- [13] Zheru C., Li H., Wang C.. *Plant species recognition based on bark patterns using novel Gabor filter banks*. *Neural Networks and Signal Processing*, 2003. Proceedings of the 2003 International Conference on.
- [14] Dahl A.B., Aanæs H., Larsen R., Ersbøll B.K.. *Classification of Biological Objects using Active Appearance Modelling and Color Cooccurrence Matrices*. SCIA'07 Proceedings of the 15th Scandinavian conference on Image analysis, 2007.
- [15] Saitoh T., Kaneko T.. *Automatic Recognition of Wild Flowers icpr*, vol. 2, pp.2507, 15th International Conference on Pattern Recognition (ICPR'00) - Volume 2, 2000.
- [16] Saitoh T., Aoki K., Kaneko T.. *Automatic Recognition of Blooming Flowers*. ICPR '04 Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04), 2004.
- [17] Pornpanomchai Ch., Sakunreraratsame P., Wongsasirinart R., Youngtavichavhart N.. *Herb flower recognition system (HFRS)*. *Electronics and Information Engineering (ICEIE)*, 2010 International Conference On, 2010.
- [18] Pornpanomchai Ch., Member, IACSIT, Kuakiatngam Ch., Supapattranon P., Siriwis-esokul N.. *Leaf and Flower Recognition System (e-Botanist)*. *IACSIT International Journal of Engineering and Technology*, Vol.3, No.4, August 2011.
- [19] Cho S.Y., Lim P.T.. *A novel Virus Infection Clustering for Flower Images Identification*. *Pattern Recognition*, 2006. ICPR 2006. 18th International Conference on.
- [20] Kim J.H., Huang R.G., Jin S.H., Hong K.S.. *Mobile-Based Flower Recognition System*. *Intelligent Information Technology Application*, 2009. IITA 2009. Third International Symposium on, 2009.

- [21] Nilsback M.E.. *An automatic visual Flora – segmentation and classification of flower images* (PhD thesis). University of Oxford, United Kingdom, 2009.
- [22] Arivazhagan S., Newlin Shebiah R., Selva Nidhyanandhan S., Ganesan L.. *Fruit Recognition using Color and Texture Features*. Journal of Emerging Trends in Computing and Information Sciences, VOL. 1, NO. 2, Oct 2010. Tamilnadu, India, 2009-2010.
- [23] Zhao J., Tow J., Katupitiya J.. *On-tree Fruit Recognition Using Texture Properties and Color Data*. 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. Sydney, 2005.
- [24] The Ministry of Agriculture of the Czech Republic. *The Report on Forestry of the Czech Republic in 2010*. Prague, 2011.

Appendices

A Contents of the enclosed CD

Directory	Content description
Branch script	The script for needles-images processing for PC
Dataset	The CB dataset
PlantIdent	The Android project folder, including native codes
PlantIdent web	The website sources and database design
Thesis	This thesis in PDF format.

Table 2: CD contents description

B Graphical resources

Resource	Info
Android Mega Pack	http://champi-design.deviantart.com/
Lovely Bones Icons by Archigraphs	http://www.archigraphs.com/

Table 3: List of graphical resources used in PlantIdent for Android