

Heikki Salo

Image Analysis with Environmental Applications

in Information Technology Master's Thesis

December 20, 2012

University of Jyväskylä

Department of Mathematical Information Technology

Author: Heikki Salo

Contact information: heikki.salo@jyu.fi

Supervisor: Tommi Kärkkäinen, Tuomo Rossi, Ville Tirronen

Title: Image Analysis with Environmental Applications

Työn nimi: Kuva-analyysi ympäristösovelluksissa

Project: Master's Thesis

Study line: Ohjelmistotekniikka

Page count: 27+31

Abstract: Remote sensing methodologies are employed in the fields of precision agriculture and forest industry. This thesis focuses on enhancing analysis process for making vegetation volume estimates from pre-processed aerial images and Digital Surface Models. An evolutionary optimisation system for learning crop field biomasses is proposed and making estimates using radiometrically corrected spectral bands with different tools is studied. This thesis also considers methods for estimating forest stem volumes by tree species.

Keywords: Remote sensing, machine learning, biomass, tree volume, k-nearest neighbours, support vector regression, optimization

Suomenkielinen tiivistelmä: Kaukokartoitusmenetelmiä käytetään tarkkuusmaataloudessa ja metsien inventoinnissa. Pro gradu -tutkielma keskittyy parantamaan analysointiprosessia kasvillisuusmäärien arvioimiseksi esikäsitellyistä ilmakuvista ja digitaalisesta korkeusmallista. Tutkielmassa esitellään evolutiivinen optimointimenetelmä viljapellon biomassojen estimoimiseksi ja tutkitaan biomassaaestimaattien tekoa eri menetelmin radiometrisesti korjatuista spektrikaistoista. Tutkielmassa pohditaan myös vaihtoehtoja puulajeittaisten tilavuuksien estimoimiseksi metsistä.

Avainsanat: kaukokartoitus, koneoppiminen, biomassa, puumäärä, k-lähintä naapuria, tukivektoriregressio, optimointi

List of publications of the author

- PI** H. Salo, V. Tirronen, F. Neri, "Evolutionary Regression Machines for Precision Agriculture", Lecture Notes in Computer Science, 2012.
- PII** I. Pölönen, I. Pellikka, H. Salo, H. Saari, J. Kaivosoja, S. Tuominen, E. Honkavaara, "Biomass estimator for CIR-image with few additional spectral band images taken from light UAS", Proceedings of SPIE 8369, Sensing for Agriculture and Food Quality and Safety IV, 2012.
- PIII** H. Salo, V. Tirronen, I. Pölönen, S. Tuominen, A. Balazs, J. Heikkilä, H. Saari, "Methods for estimating forest stem volumes by tree species using digital surface model and CIR images taken from light UAS", Proceedings of SPIE 8390, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, 2012.

Glossary

BRDF	Bidirectional reflectance distribution function
CIR	Colour-infrared image
DEM	Digital Elevation Model
DSM	Digital Surface Model
Feature	A numerical quantity that describes an object of interest
GSD	Ground Sample Distance
Hyperspectral image	An image consisting of many narrow wavelengths spaced evenly apart from each other
Multispectral image	An image consisting of multiple wavelengths
NIR	Near-infrared image
Orthoimage	Geometrically corrected aerial photograph
Precision agriculture	The study of intra-field variations to aid farming decisions
Remote sensing	non-destructive measuring of terrain properties
Texture	A feature that describes distributions of tones in an image
RMSE	Root Mean Square Error
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle

List of Figures

Figure 1. The steps of the KDD process.	4
Figure 2. Prototype of the VTT's lightweight hyperspectral imager.	10
Figure 3. Ground-truth points marked on a colour-infrared image.	13
Figure 4. Histograms of the fitness value distributions between algorithms tested in PI . ..	14
Figure 5. Gatewing X100 UAV launched with catapult to acquire airborne images.	16

Contents

1	INTRODUCTION	1
2	REMOTE SENSING AND DATA MINING	3
3	THE UASI PROJECT	9
4	APPLICATION: PRECISION AGRICULTURE	12
5	APPLICATION: FOREST INVENTORY	15
6	DISCUSSION	18
7	CONCLUSION	19
	BIBLIOGRAPHY	20
	APPENDICES	22
	A Publication PI	24
	B Publication PII	35
	C Publication PIII	46

1 Introduction

Remote sensing of vegetation is a quickly developing field. As more advanced technology becomes available new remote sensing applications can be developed. In both applications we focus on improving machine learning process to get better vegetation volume estimates. The end user of the application can then use the volume estimates to aid decision making in land-management operations.

The thesis consists of three conference papers: Publication **PI** proposes an evolutionary system for learning crop field biomasses from aerial images. The system is trained with modern meta-heuristic algorithms. In **PI** we see that a differential evolution-based optimizer outperformed two other that were based on particle swarm optimization and evolution adaptation strategy.

Publication **PII** uses infrared images with a few radiometrically corrected spectral bands and Digital Elevation Model for estimating crop field biomasses. **PII** shows that combining features from radiometrically corrected spectral bands and Digital Elevation Model gives the best results for estimating biomasses.

PIII considers methods for estimating forest stem volumes by tree species using digital surface model and colour-infrared images taken from light UAS. In **PIII** a treetop delineation method is presented and we see that photogrammetric surface model is not sufficient alone for forest applications.

My contributions in **PI** are implementation of the fitness function for the optimisation scheme and the extracted features. In **PII** I contributed the extracted features that were used in the study. I conducted the study in **PIII** apart from developing the presented tree top delineation method and collecting the ground-truth data.

The Section 2 gives a summary of remote sensing and covers the basis for the presented studies. Overview of the UASI¹ project is given in Section 3 along with the two application

1. Contact information can be found from the website <https://www.jyu.fi/it/laitokset/mit/tutkimus/uasi/>

fields presented in Sections 4 and 5. Section 6 contains discussion and Section 7 sums up the thesis.

This thesis was done as part of the Tekes² funded UASI (Unmanned Aerial System Innovations) project.

2. Tekes is Finnish Funding Agency for Technology and Innovation

2 Remote sensing and data mining

In this section we get to know basic concepts of remote sensing. Remote sensing in its entirety is a vast subject, and this section is restricted to the type of remote sensing utilized in the presented publications. The publications focus on improving a single step: analysis of data. We'll start by exploring the remote sensing in general and then concentrate on analysis and examine it with terms of data mining.

In the early days remote sensing meant merely “the observation and measurement of an object without touching it” (Jones and Vaughan 2010). While this gives a good intuition of the topic, the definition can be more formal. In their book “Remote Sensing and Image Interpretation” Lillesand, Kiefer, and Chipman (2008) define the term remote sensing as follows:

“Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation.”

The remotely collected data can be of many forms, including acoustic wave distributions, electromagnetic radiation or force distributions (Lillesand, Kiefer, and Chipman 2008). Also, sensing electromagnetic radiation is still quite a vague expression, as taking pictures with a digital camera and reading information of modern passport using RFID would still fall into same category.

The visible light, one form of electromagnetic radiation, covers the spectral range of 400 – 700 nanometers, containing bands of blue, green and red that span roughly 100 nanometers each (Lillesand, Kiefer, and Chipman 2008). That is, when we look at an object, our eyes only sense the emitted electromagnetic radiation of this range. While the common digital camera is also adjusted to cover the same spectral range, there exists a wide range of equipment for covering other ranges as well¹.

The information that can be acquired from the images depends on the sensors used but also

1. The used equipment is presented in the next Section 3

on their location relative to the target. There's challenges in aerial imaging even after succeeding in selecting the imager, optics, exposure time etc. as the atmosphere itself can have significant effect on the outcome (Lillesand, Kiefer, and Chipman 2008). The environmental conditions can also vary between the pictures or even within a single picture because of e.g. cloud shadows.

Covering ground using multiple aerial images taken from different locations presents many possibilities. If the overlap of aerial images exceed 70-80 %, it is possible to construct a *digital surface model (DSM)*² of the ground by using photogrammetry (Saari et al. 2011). DSM can be thought as a 2D digital image that contains height values as pixel values instead of colour values. A model that is created by subtracting a DSM of the ground level from a DSM of targets (e.g. vegetation) is called digital elevation model (**DEM**). We'll later cover how these models can be used in the remote sensing process.

The process of remote sensing has a goal, to extract knowledge from the remote acquired measurements. The remote sensing process consists of the following steps: *object, sensor, data, analysis* and *information*. Everything between the object and information is determined by the requirements of the application, essentially what information is to be remotely measured from where. While the process sums up the idea in short, the workflow is as abstract as possible.

To lay groundwork for the studies presented in this thesis, we'll now take another look at the analysis phase using terms of a process known as "knowledge discovery in databases" (KDD) (Fayyad, Piatetsky-Shapiro, and Smyth 1996). The steps of the KDD process are presented in Figure 1. It may be useful to think the process as a pipeline consisting of the necessary steps for turning raw data into knowledge.

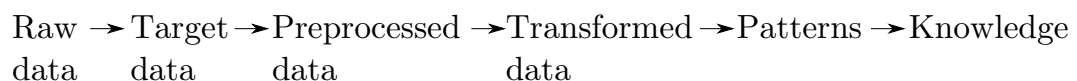


Figure 1. The steps of the KDD process.

2. The digital surface model of the ground level is also called digital terrain model (**DTM**).

Raw data as the first step in Figure 1 refers to the raw measurements, unprocessed outputs of an imaging sensor or some measurements. The images are acquired remotely whereas the measurements are *ground-truth*, they are measured on the spot.

Target data is the first refinement step that can, for example, contain removing outliers from data (Fayyad, Piatetsky-Shapiro, and Smyth 1996). The target data is a pruned version of raw data.

Preprocessing means turning the data to *samples* and can consist of numerous phases depending on the application. The samples represent the target objects as numbers or classes, that are paired by a number of *features* extracted from the remotely acquired data.

There are some requirements for preprocessing in remote sensing that should be noted. First, there should be a link between the object world and the images. When covering ground objects from airborne images this can be achieved by georeferencing single images. Aerial images can also be combined as an *orthophoto*, in which each pixel appears like it was captured from straight above, being uniform in size and clearly located in the nature.

Creating the aforementioned digital surface models are one example of the possible preprocessing steps. The last step of preprocessing is extracting features from the prepared datasets.

Feature extraction is the art of representing the target objects by descriptors extracted from the remote measurements. In hopes of the above-mentioned and other present issues have been addressed by previous phases of preprocessing, the extracted feature set aims to contain all the needed information of the objects.

Features can, for example, make use of channel intensities, texture properties or DSM properties, values of digital elevation models for example. This can be seen as compressing all the available remotely sensed information of an object only to numbers. The number of features that can be extracted is limited only by imagination, as the features can depend on parameters and can be used together to derive new features.

There exists no theoretical upper limit for the number of extracted features but as we'll soon find out, having large number of features is inconvenient. This being said, the feature extraction basically shrinks the volume of the data down by orders of magnitude.

Transformation is performed to the samples to make the learning of the data easier. The used features can have very different value ranges and are therefore not treated equally in some learners. This step could also contain encoding of the object's classes suitable to the used learners. High dimensionality (and the problems of caused by it) is also an upcoming topic. These are examples of the issues the transformation step tries to address.

The book “Dimension Reduction: A Guided Tour” Burges (2010) gives an extensive tour of dimension reduction and defines it as follows:

“Dimension means mapping of data to a lower dimension space such that uninformative variance in the data is discarded, or such that a subspace in which the data lives is detected”.

At its simplest, this can mean picking the useful features to describe the objects, which is called *feature selection*. It is somewhat inconvenient that there are no obvious implementations for the simplest concept. Feature selection can be done in several ways with or without measuring their contributions to the overall performance of the process.

The feature space can also be projected to another space by using for example Principal Component Analysis (PCA), which projects the data in terms of its *principal components*.

Patterns are used in the KDD process as a term to cover the learning of data (Fayyad, Piatetsky-Shapiro, and Smyth 1996). Here we are going to use the term **model** to represent the relation between the features and desired outcomes. This finding of a model is present in pattern recognition, machine learning, data mining and statistical literature. The book “The Elements of Statistical Learning” by Hastie, Tibshirani, and Friedman (2011) addresses the different views and gives an extensive introduction to the field.

When we predict quantitative values – like biomasses – we are performing regression (Hastie, Tibshirani, and Friedman 2011). Predicting qualitative values – tree species – is called classification (Hastie, Tibshirani, and Friedman 2011).³ In any case the goal is to find a useful

3. In statistical literature the features as defined above could be called *descriptors* or *independent variables*, where as the outcomes of the learner would be called *dependent variables* (Hastie, Tibshirani, and Friedman 2011).

approximation \hat{f} of the unknown function f that maps the extracted feature vectors \mathbf{x} as estimates y :

$$y_i = \hat{f}(\mathbf{x}_i) + \varepsilon_i,$$

where $\mathbf{x}_i \in \mathbb{R}^n$ is the feature vector of sample i , $y_i \in \mathbb{R}^1$ is the estimated value for sample i and $\varepsilon_i \in \mathbb{R}^1$ the error between the estimated and measured values. A convenient situation would be that the output of a regression or classification task is linearly dependent of the (possibly transformed) features. In this case our feature vector $\mathbf{x}_i = (x_1, x_2, \dots, x_n)$ containing n features can be projected using:

$$\hat{f}(\mathbf{x}; \beta) = \beta_0 + \sum_{j=1}^n (\mathbf{x}_i)_j \beta_j,$$

where $(\mathbf{x}_i)_j \in \mathbb{R}^1$ is the j 'th feature of i 'th feature vector, β is a coefficient vector – usually solved using *least squares* method – and β_0 is the bias of the projection. The linear model has many attractive properties and is further covered in Hastie, Tibshirani, and Friedman (2011).

There are many types of *learners* – the different methods for acquiring the approximated \hat{f} functions. The linear model presented above is one example. The requirements of the application can narrow down the list of applicable types. Since *k-nearest neighbour* learner is referred in **PII** and **PIII** we'll introduce it quickly. K-nearest neighbours (k-NN for short) is an example of an *instance-based* learner. It is based on a simple mathematical concept yet it can be very powerful. K-NN evaluates the properties of an unknown sample by its k nearest neighbours in the training set (Hastie, Tibshirani, and Friedman 2011), selecting the most common class present or averaging some real valued attribute.

K-NN has a characteristic set of advantages and flaws. K-NN is easy to implement and its responses are easy to interpret. It doesn't do many assumptions of the data and while possibly delivering good results, it can be unstable as the available samples determine its behavior completely (Hastie, Tibshirani, and Friedman 2011).

Interpretation as the end of the KDD process is also the end product of the remote sens-

ing process. The features as well as the target attributes could have been gone through a transformation before applying a learner.

Interpretation could mean decoding the outputs of a learner from the format of the transformation process. Now the remotely sensed information can be used to aid for example a decision making process.

In summary the goal is to produce information from raw data. In restricted application areas of remote sensing – like indicating movement or sensing temperature – a small toolchain can be sufficient and no data mining concepts are required. However, there were no knowledge of such elegant solutions in the beginning of the studies of the two application fields.

It should also be noted that many of the aforementioned steps – preprocessing, transformation, patterns – can consist of several tools that require beforehand selected parameter values. For example feature extraction method can operate on the image using a fixed size window and the k nearest neighbour requires a selected k . In some cases these values can be optimized by testing the relevant value range.

3 The UASI project

This section covers the project this thesis and the presented publications is part of. We'll examine the characteristics of the UASI project using the terminology presented in the previous section. The Unmanned Aerial System Innovations (UASI) project is a collaboration between many faculties and companies in Finland and is funded by Tekes¹ (Saari et al. 2011). The UASI project's research is done in collaboration between

- **VTT** Technical Research Centre of Finland,
- **MTT** Agrifood Research Finland,
- **Metla** The Finnish Forest Research Institute,
- **JAMK** University of Applied Sciences,
- **FGI** Finnish Geodetic Institute,
- Pieneering Ltd,

and University of Jyväskylä (**JYU**), which manages the project.

The main goal of the UASI project is to evaluate how a light imaging system could be used in precision agriculture and forest inventory applications (Saari et al. 2011). The goals in the both applications are to support land-management decisions. The targets in the agriculture applications are to remotely estimate biomass and nitrogen concentrations and to identify weeds (Saari et al. 2011). The publications **PI** and **PII** study the biomass estimation. The application field and the studies are presented in more detail in Section 4.

The target in the forest application is to remotely estimate forest stand characteristics, like tree volumes by tree species (Saari et al. 2011). The publication **PIII** considers methods for estimating tree volumes. The forest application and the study is described in Section 5.

The project's data is gathered from flight campaigns performed over crop fields and forests. There have been flights using several different UAV platform and imaging equipment combinations.

1. The Finnish Funding Agency for Technology and Innovation

While the UAV's and most of the cameras have been off-the-shelf, there has also been a novel imager that has been tested: VTT has developed a lightweight hyperspectral imager that can be mounted to UAV (Saari et al. 2011). Traditionally hyperspectral imaging equipment has been so heavy that operating it had required a helicopter or a small airplane. The spectral range of the imager is 500 to 900 nanometers and it can be easily mounted to UAV platforms as it weighs less than 420 grams (Mäkynen et al. 2011).² The lightweight alternatives operate by acquiring line by line, which makes using them from a moving platform almost unthinkable as rendering the images is far from trivial.

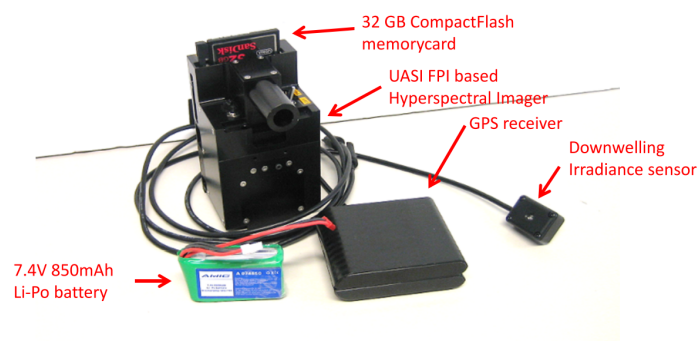


Figure 2. Prototype of the VTT's lightweight hyperspectral imager.

The new hyperspectral imager is based on Fabry-Perot Interferometer, which is used to acquire three chosen spectral bands at once with a common RGB sensor (Mäkynen et al. 2011). Very precise actuators control a very thin airgap, which is responsive for the “mapping of the spectres” (Mäkynen et al. 2011). For more detailed description of the imager see Mäkynen (Mäkynen et al. 2011; Makynen et al. 2012).

The imager captures three spectral bands at the same time, which makes it viable solution to airborne imaging. However, covering a set of spectral bands requires capturing multiple three wavelength sets that cover the needed wavelengths (Mäkynen et al. 2011). This means that while the single bands are acquired at the same time, different spectral channels aren't.

The are reasons to use UAVs instead of using larger platforms. While airplanes and satellites could withstand bigger imaging equipment and they have been in operational use for a long time, they suffer of economical and operational drawbacks. A successful image acquisition

2. The hyperspectral imager will be commercially available from Rikola Ltd. <http://www.rikola.fi>

using satellites depends heavily on the weather as clouds can block the target partially or completely. Satellite imaging has also high operational costs while having poor spatial resolution (Stafford 2000). These drawbacks of the satellite imaging are biggest benefits of using UAV based remote sensing (Berni et al. 2009). UAVs are also cheaper to operate than small airplanes (Berni et al. 2009).

Next we'll introduce some of the arrangements in terms of the remote sensing process described earlier.

The **preprocessing** of the aerial images in the UASI project has been done by Pieneering Ltd. and Finnish Geodetic Institute (FGI). Using photogrammetry in chosen application fields have required developing new methods, some of which has presented in publications (Honkavaara, Kaivosoja, et al. 2012; Honkavaara, Hakala, et al. 2012). The preprocessing for publications **PI** and **PII** covering agriculture was done by FGI and for **PIII** that covers forest inventory by Pieneering.

Not all of the acquired data was processed to be available for the presented publications.

Feature extraction and rest of the analysis process has been done by the Finnish Forest Research Institute, Agrifood Research Finland and University of Jyväskylä. For the analysis process there has been many types of images available. Features were extracted from both orthoimages created from both common and hyperspectral imager's images and digital surface models. Having both hyperspectral images and common RGB or colour-infrared images is really an advantage for many parts of the process. While the hyperspectral imager has better spectral resolution – narrow and precise wavelength ranges – it has lower spatial resolution.

The next Sections 4 and 5 covering the applications give more details of the requirements for the applications.

4 Application: Precision Agriculture

The agriculture application for a lightweight imaging system is supporting a discipline called *precision agriculture*, **PA** for short. Precision agriculture in a nutshell is using knowledge of intra-field variation to support land management decisions (Stafford 2000).

The basic concept of PA is nothing new – the treating of different parts of the field depending on their properties – has been done for centuries (Stafford 2000). For example, different surface properties – slopes and elevated areas – make the water to be unevenly available within the field, which has a direct effect on the area's growth potential.

The PA in the modern sense means using technology to aid decision making process. The beginning of the technologically intensive PA is in the mid-1980s (Robert 2002) and since then its growth has been driven by (among other factors) tightened agriculture legislation (Stafford 2000). Legislation concerning the use of fertilizers and weed management continues to tighten in Europe (Stafford 2000; Zhang, Wang, and Wang 2002). Public opinion also favors using optimized farming methods over genetically manipulated food (Stafford 2000).

The advantages of PA are measured in environmental and economical benefits (Zhang, Wang, and Wang 2002). While the needed technology is more and more available, the benefits of PA hasn't been proven except some cases (Stafford 2000). Many of the methods for estimating crop properties in the literature is only in academic use (Ehlert, Horn, and Adamek 2008).

The goal of the presented publications **PI** and **PII** is to study means of improving the analysis chain of the remote sensing process in order to support precision agriculture. Both studies make use of preprocessed orthoimages and digital surface models. Improving the preprocessing of the UAV acquired materials for precision agriculture has been studied in (Honkavaara, Kaivosoja, et al. 2012; Honkavaara, Hakala, et al. 2012).

The flight campaigns for acquiring aerial images that were used in **PI** and **PII** were conducted in July 2011 in Vihti, Finland. Commercial Panasonic Lumix NIR camera and a prototype of the VTT hyperspectral imager were flown using MD4-1000 UAV.

Figure 3 shows part of the crop field in Vihti with ground-truth points marked on a colour-



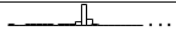
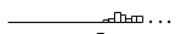
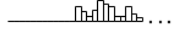
Figure 3. Ground-truth points marked on a colour-infrared image.

infrared image. All ground-truth measurement points are marked with circles, bright and dark circles indicate high and low biomass measurements, respectively. The ground-truth data was prepared by MTT and consists of 91 samples of the field that were analysed in laboratory. The ground-truth points are used in various arrangements for training and validating the methods in **PI** and **PII**.

Publications **PI** and **PII** try to enhance the remote sensing process by improving the analysis of the data. Both studies have in common that they compare performances of different transformation methods and learners (as described in Section 2).

PI uses ensemble learning for the machine learning process. The ensemble consists of transformation method and learner combinations that together produce the estimates. The learners are support vector regression machines with different kernel functions and the choices are to use principal component analysis (PCA) or simple scaling. The task of determining optimal used combinations and the parameters of the used tools is formalized as a real-valued optimization challenge, where the fitness function maps real-valued vector to average error of the system.

The optimization is done by using three modern optimization algorithms of different type,

Algorithm	Average±Std.Dev	Best	Histogram
Prox-DE	2.546±0.248	1.401	
F-PSO	3.400±0.230	2.989	
CMA-ES	3.020±0.314	2.546	

1.401 3.568

Figure 4. Histograms of the fitness value distributions between algorithms tested in **PI**.

namely Proximity based Differential Evolution (Prox-DE), Frankenstein-Particle Swarm Optimization (F-PSO) and Covariance Matrix Adaptation Evolution Strategy (CMA-ES). The arrangements of the test is explained in **PI**. In Figure 4 we can see that in this optimization task Prox-DE produced both the optimal (minimal) fitness value and produced consistently superior candidates in each run.

PII compares two learners, the k-NN that was presented in Section 2 and support vector regression. Applying *diffusion maps* as dimension reduction is also tested. Basic idea of diffusion maps is explained in **PII**. In the study it was seen that using k-NN with diffusion maps outperformed using support vector regression or k-NN alone.

The studies **PI** and **PII** make use of different data. In the time of study for **PI** (November 2011) the pre-processing of raw images was unfinished and only digital surface model and high resolution near-infrared band were available. Study for **PII** was carried out later (April 2012) when there was already preprocessed orthoimages of spectral bands available.

5 Application: Forest Inventory

This section covers the forest inventory application for a lightweight imaging system. The goal of the forest application is to offer forest stand details to aid cutting planning (Saari et al. 2011). The goals of the UASI project include developing this application and studying how UAVs and surface models generated using photogrammetry can be applied to forest applications. To succeed in developing this application would enable performing forest inventory with new kinds of platforms and devices – lighter equipment.

Some motivations for using UAV platforms were already listed in Section 3. There's been also studies that make use of satellite images, but so far the results have not been sufficient to be used to aid decision making (Tuominen and Pekkarinen 2005).

Interesting forest stand details would include tree volumes by tree-species. The total volumes are successfully acquired by using already operational methods – laser scanning – but there are only few studies that aim at species-specific volumes (Packalen and Maltamo 2006). Other forest characteristics include crown leaf area index, in which there has already been success in remotely sensing it with UAVs (Berni et al. 2009).

This thesis presents **PIII** that considers ideas for acquiring tree volumes by tree-species. In literature there are two schools of thought to achieve this. First is to study the forest as small batches (stand-level) and another is to perform individual-tree recognition (Packalen and Maltamo 2006). Both methods have delivered results but only stand-level examination is in operational use (Packalen and Maltamo 2006).

The image data that was used in **PIII** was acquired in July 2011 from Evo educational forest. In Figure 5 an unmanned aerial vehicle carrying a colour-infrared camera has just left the launchpad for acquiring aerial images. Pieneering Ltd. performed the flights and the preprocessing of the raw aerial images as orthoimages and digital surface models.

In **PIII** it was seen that DSM's that are acquired by means of photogrammetry (described briefly in Section 2) alone are not enough for estimating tree volumes, as measuring the elevation would require estimate of the ground level that couldn't be estimated from single



Figure 5. Gatewing X100 UAV launched with catapult to acquire airborne images.

surface model alone. The photogrammetry method for creating the surface model of treetops was not viable for detecting relatively small clear areas as hoped, from which the height of the vegetation could be estimated.

Lacking information of ground level is not necessarily a critical problem since public ground-level surface models are available from a public faculty at least in Finland¹. The public maps have been acquired by laser scanning. Change detection would still be valid application for photogrammetry if the data would otherwise be usable.

In **PIII** it was also find out that photogrammetry-based DSM is not suitable for tree top recognition. Detecting treetops from the surface model was demanding even for human observer. For this reason, in **PIII** we performed treetop search using colour channels and found out that performing a filtering method presented in **PIII** using green channel of the near-infrared image seems to work well compared to other strategies.

It is inevitable that using aerial imagers of visible or near-visible spectral range some percentage of nondominant trees will be hidden from view being simply under crowns of dominant trees or in shadow. Delineation method presented in **PIII** could serve as a starting point for developing other required tools for recognizing trees.

If the performance of the presented delineation method isn't sufficient there exists other delineation methods for colour-infrared images in the literature (Brandtberg and Walter 1998).

1. For more information see the website of National Land Survey of Finland (Maanmittauslaitos): <http://www.maanmittauslaitos.fi/digituotteet/korkeusmalli-2-m>.

The method tested in **PIII** was to mix extracted stand-level features with the partially recognized and classified treetops. This way the presence of tree species wouldn't rely on stand-level features indicating presence of different tree species but on actual detections of single trees.

6 Discussion

Choosing the extracted features and performing machine learning are tasks that don't seem to have obviously well-performing implementations, which was perhaps the most striking thing for me. The target, circumstances, equipment and the analysis process are unique to each study. For example, **PI** and **PII** cover the same field but use different feature sets, thus having different starting points for performing machine learning.

Using literature one can find features that have high expectations for being useful in machine learning task. The unpleasant aspect is that introducing new features to the process the best performing choices for machine learning can alter, for example, turning a linear problem to non-linear.

In **PI** we took a stand on the issue of selecting best performing tools by means of optimization. However, for optimization to work, of the choices that must be made in the analysis process (described in Section 2) only few can be feasibly optimized at once. In other words, it requires domain knowledge to successfully select the unknown attributes worth optimizing.

Statisticians have criticized the data mining approach for the pursuing a seemingly good model with any means necessary (Fayyad, Piatetsky-Shapiro, and Smyth 1996). This means that with enough effort put into figuring a pattern in a fixed set of samples, something can always be found. In a context of an application it is essential that the found model performs sufficiently with equivalent unseen environment.

PI successfully demonstrates the formalization of an optimization task for the parameters of a machine learning problem. As seen in Section 2 there are countless variables that have an effect on the performance of the remote sensing chain and comparing numerical results between studies is therefore almost infeasible. In the context of remote sensing **PI** lacks comparison of the presented method to the currently used methods. This should be addressed in future work.

The work in **PIII** is heavily in progress. The future work shows whether or not the concept is viable to be used to enhance species-specific estimates.

7 Conclusion

In **PI** we demonstrated that selecting optimal set of tools for machine learning and optimizing the tool's parameters can be formalized as a real-valued optimization task. From the compared three modern optimization algorithms we found out that the performance of the algorithm based on differential evolution was superior in comparison to others.

Publication **PII** presents biomass estimation results of using radiometrically corrected spectral bands and digital surface model. Using diffusion maps and k-NN estimator gave best results in this test.

Forest inventory for discussed in **PIII** and ideas for learning species-specific estimates were presented. Limitations of using the photogrammetrically created digital surface model in tree volume estimation and tree top recognition were discussed.

The studies support the development of remote sensing process in the two application areas by enabling more informed decisions to be made in selecting the tools to perform the image analysis. The original papers are included in the appendix.

Bibliography

Berni, J.A. Jimenez, P.J. Zarco-Tejada, L. Suárez, V. Gozález-Dugo, and Fereres. E. 2009. “Remote sensing of vegetation from UAV platforms using lightweight multispectral and thermal imaging sensors”. In *ISPRS Hannover Workshop*.

Brandtberg, Tomas, and Fredrik Walter. 1998. “Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis”. *Machine Vision and Applications* 11, number 2 (): 64–73. ISSN: 0932-8092. doi:10.1007/s001380050091.

Burges, Christopher J. C. 2010. *Dimension Reduction: A Guided Tour*. 106. Now Publishers Inc. ISBN: 1601983786.

Ehlert, Detlef, Hans-Jürgen Horn, and Rolf Adamek. 2008. “Measuring crop biomass density by laser triangulation”. *Computers and Electronics in Agriculture* 61, number 2 (): 117–125. ISSN: 01681699. doi:10.1016/j.compag.2007.09.013.

Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. 1996. “The KDD process for extracting useful knowledge from volumes of data”. *Communications of the ACM* 39, number 11 (): 27–34. ISSN: 00010782. doi:10.1145/240455.240464.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2011. *The Elements of Statistical Learning*. 5th. Springer-Verlag.

Honkavaara, Eija, Teemu Hakala, Lauri Markelin, Tomi Rosnell, Heikki Saari, and Jussi Mäkynen. 2012. “A Process for Radiometric Correction of UAV Image Blocks” [**inlangen**]. *Photogrammetrie - Fernerkundung - Geoinformation* 2012, number 2 (): 115–127. ISSN: 14328364. doi:10.1127/1432-8364/2012/0106.

Honkavaara, Eija, Jere Kaivosoja, Jussi Mäkynen, Ismo Pellikka, Liisa Pesonen, Heikki Saari, Heikki Salo, Teemu Hakala, Lauri Markelin, and Tomi Rosnell. 2012. “Hyperspectral Reflectance Signatures and Point Clouds for Precision Agriculture by Light Weight UAV Imaging System”. In *ISPRS Annals*, 353–358.

Jones, Hamlyn G., and Robin Antony Vaughan. 2010. *Remote sensing of vegetation*. Illustrate. 353. Oxford University Press. ISBN: 0199207798, 9780199207794.

- Lillesand, Thomas M., Ralph W. Kiefer, and Jonathan W. Chipman. 2008. *Remote sensing and image interpretation*. 6th editio. 756. John Wiley & Sons.
- Mäkynen, Jussi, Christer Holmlund, Heikki Saari, Kai Ojala, and Tapani Antila. 2011. “Unmanned aerial vehicle (UAV) operated megapixel spectral camera”. In *Proceedings Vol. 8186, Electro-Optical Remote Sensing, Photonic Technologies, and Applications V*. SPIE.
- Makynen, Jussi, Heikki Saari, Christer Holmlund, Rami Mannila, and Tapani Antila. 2012. “Multi- and hyperspectral UAV imaging system for forest and agriculture applications”. In *SPIE Defense, Security, and Sensing*, pages. doi:10.1117/12.918571.
- Packalen, P, and M Maltamo. 2006. “Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs”. *Forest Science* 52 (6): 611–622. ISSN: 0015749X.
- Robert, P. C. 2002. “Precision agriculture: a challenge for crop nutrition management”. *Plant and Soil* 247, number 1 (): 143–149. ISSN: 0032-079X. doi:10.1023/A:1021171514148.
- Saari, Heikki, Ismo Pellikka, Liisa Pesonen, Sakari Tuominen, Jan Heikkilä, Christer Holmlund, Jussi Mäkynen, Kai Ojala, and Tapani Antila. 2011. “Unmanned Aerial Vehicle (UAV) operated spectral camera system for forest and agriculture applications”. In *Proceedings Vol. 8174, Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII*. SPIE.
- Stafford, John V. 2000. “Implementing Precision Agriculture in the 21st Century”. *Journal of Agricultural Engineering Research* 76, number 3 (): 267–275. ISSN: 00218634. doi:10.1006/jaer.2000.0577.
- Tuominen, S, and A Pekkarinen. 2005. “Performance of different spectral and textural aerial photograph features in multi-source forest inventory”. *Remote Sensing of Environment* 94, number 2 (): 256–268. ISSN: 00344257.
- Zhang, Naiqian, Maohua Wang, and Ning Wang. 2002. “Precision agriculture—a worldwide overview”. *Computers and Electronics in Agriculture* 36, **numbers** 2-3 (): 113–132. ISSN: 01681699. doi:10.1016/S0168-1699(02)00096-0.

Appendices

ORIGINAL PAPERS

A Publication PI

Evolutionary Regression Machines for Precision Agriculture

Heikki Salo, Ville Tirronen and Ferrante Neri

Lecture Notes in Computer Science, 2012.

© Springer-Verlag Berlin Heidelberg

Evolutionary regression machines for Precision Agriculture*

Heikki Salo, Ville Tirronen, and Ferrante Neri

Department of Mathematical Information Technology,
P.O. Box 35 (Agora), 40014 University of Jyväskylä, Finland,
Tel +358-14-260-1211, Fax +358-14-260-1021,
heikki.salo@jyu.fi, ville.tirronen@jyu.fi, ferrante.neri@jyu.fi

Abstract. This paper proposes an image processing/machine learning system for estimating the amount of biomass in a field. This piece of information is precious in agriculture as it would allow a controlled adjustment of water and fertilizer. This system consists of a flying robot device which captures multiple images of the area under interest. Subsequently, this set of images is processed by means of a combined action of digital elevation models and multispectral images in order to reconstruct a three-dimensional model of the area. This model is then processed by a machine learning device, i.e. a support vector regressor with multiple kernel functions, for estimating the biomass present in the area. The training of the system has been performed by means of three modern meta-heuristics representing the state-of-the-art in computational intelligence optimization. These three algorithms are based on differential evolution, particle swarm optimization, and evolution strategy frameworks, respectively. Numerical Results derived by empirical simulations show that the proposed approach can be of a great support in precision agriculture. In addition, the most promising results have been attained by means of an algorithm based on the differential evolution framework.

1 Introduction

Food production and agriculture have been transformed from a solar based industry into one relying on fuel, chemicals, sensors and technology. The use of chemicals and fuel increased dramatically in 60s and 70s and several concerns were stated about the effect of this increase to our health and the health of our environment. This concern and the advances in imaging technology resulted in the development of the Precision Agriculture (PA), see [10]. PA is a farming technique based on observing and responding to intra-field variations. Clearly, the observation of variations in the field is crucially important to promptly apply a countermeasure.

* This research is supported by the Academy of Finland, Akatemiattutkija 130600, Algorithmic Design Issues in Memetic Computing. A special thank to Antti-Juhani Kaijanaho for the useful discussions.

A fundamentally important entity to monitor within a field is the produced biomass since an accurate map of field biomass is necessary for crop yield estimation and optimal field management, see [13]. If an exact inventory of plant mass is known, more careful economical planning can be done. Furthermore, if some parts of the field fall behind in growth, intervention methods, such as fertilization or additional irrigation, can be used.

Field biomass mapping systems are often image based, where spectral or false color images are acquired from satellites, aeroplanes, and devices mounted on tractors and other field equipment. Field map creation is based on machine vision techniques that include a wide variety of machine learning elements where the biomass estimation is based on features and models built from the images. For example, in [14], an estimation scheme using several different vegetation indices based on relationships of multispectral images is proposed. In [6], biomass estimation is performed by means of stereoscopic vision techniques used to construct the so called Digital Elevation Models (DEM), i.e. 3-D representations of the terrain surface, from sets of ordinary aerial photographs. In [17], the combination of both multispectral images and digital elevation based measurements of the biomass is successfully proposed.

In order to estimate the biomass in a field, multiple images and measurements must be taken and the images must be processed. Thus, the problem can be presented as a non-parametric regression, which is further complicated by the large variability in images.

In this paper, we propose a chain of operations that extracts suitable information from image data and creates a non-parametric estimator for biomass using a machine learning technique for performing the non-parametric regression. This technique, is based on Support Vector Regression (SVR) (see [3]). and ensemble learning (see [11] and [1,2]).

The training of the SVM ensemble is obtained by means of three modern computational intelligence optimization algorithms, based on Evolution Strategies (ES), Differential Evolution (DE) and Particle Swarm Optimization (PSO), respectively.

The remainder of this paper is organized as follows. In Section 2 we introduce the chain of operations and the support vector regression and ensemble learning techniques. Section 3 shows the performance comparison of the three meta-heuristics considered in this study. Finally, Section 4 gives the conclusions of this work.

2 Intelligent system for biomass estimation

The proposed chain is schematically represented in Fig. 1. A set of images is taken by an Unmanned Aerial Vehicle (UAV). These images are processed into DEM and multispectral images, as shown in [17], thus producing a set of data which is processed by a machine learning technique to associate to each portion of land (patch) with a biomass value.

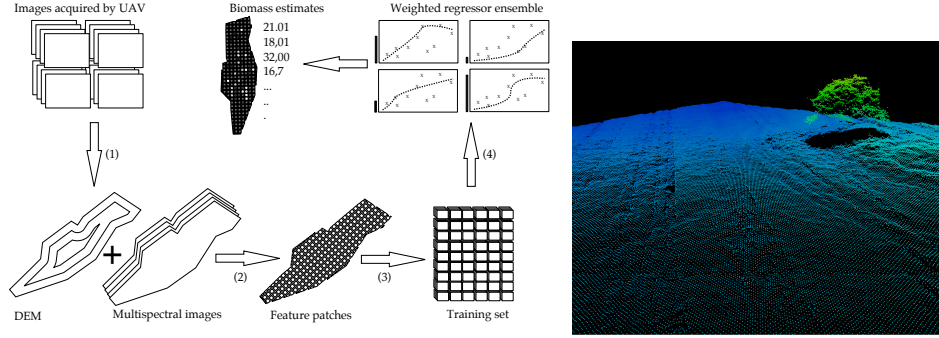


Fig. 2. A Digital Elevation Model (DEM) of a field

Fig. 1. General workflow of a biomass estimation system.

The first step in the chain (1) is collating the images acquired by an unmanned aerial vehicle into a digital elevation model (see Figure 2) and an orthographic map of the field. In our case this is done by the UAV operator using image correlation and stereoscopic vision techniques (see [12] for a survey of this topic). This phase is not parametrised in this paper.

In the second step of the process (2), the field area is divided into test patches. Each patch contains a specific sample with a different biomass. For each patch we calculate the following features, which are used to train the estimation system:

1. A cumulative histogram of the elevation values: $\text{cdf}_i^{DEM} = \sum_{j=0}^i \frac{n_j}{n}$, where $0 < i/leg5$, and n_j is number of elevation samples where the elevation is between $\min + j(\min - \max)$ and $\min + (j + 1)(\min - \max)$, n is the number of elevation samples within the patch and \min and \max are the minimum and the maximum elevations in the patch respectively.
2. The average (μ) of the elevations measured in the patch.
3. The variance (s_{DEM}^2) of the elevations measured in the patch.
4. The variance (s_{NIR}^2) of the NIR channel responses measured in the patch.
5. A Cumulative histogram of the NIR channel responses: $\text{cdf}_i^{NIR} = \sum_{j=0}^i \frac{n_j}{n}$, where $0 < i/leg2$, and n_j is number of NIR channel responses in range $\min + j/(\min - \max)$ and $\min + (j + 1)/(\min - \max)$, n is the number of responses within the patch and \min and \max are the minimum and the maximum responses in the patch respectively.

These features must be preprocessed to equalize features with different ranges. In this paper, we consider preprocessing with both simple scaling and the principal component analysis (PCA), which can be used to reduce the feature vector dimensionality. In the case of PCA a proper ratio of dimension reduction T_{pca} must be selected properly.

These features are paired with the physically measured dry-biomass values in the step (3) to produce the training set from which we build the biomass estimator using regression analysis in the step (4). Regression analysis is the science of determining the relationship between dependent and independent variables and it is used in devising automatic prediction and forecasting tools. When the relationship between the parameters is unknown, the problem is named non-parametric regression. SVM, following the example given in [3], are used here to perform the non-parametric regression tasks. In order to construct a SVR from a set of point and value pairs, $\{(x_i, y_i)\} \subset \mathbb{R}^n \times \mathbb{R}$, we must find a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that f deviates at most an ϵ amount from the training points:

$$|y_i - f(x_i)| \leq \epsilon \quad (1)$$

$$(2)$$

while f should be as simple as possible.

In our case, the points x_i represent the features of the terrain acquired by the imaging system and the values y_i represent the biomass measures associated with the corresponding terrain features. The resulting function f will be the biomass estimator for the non-measured parts of the terrain. In order to model this problem, it is enough to consider a linear function $f(x) = w \cdot x + b$ and equate simplicity to flat slope, which can later on be generalized to a non-linear estimator by using a non-linear mapping of the data points. This results in the following optimization problem:

$$\text{minimize } \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (3)$$

$$\text{subject to } y_i - w \cdot x + b \leq \epsilon + \xi_i \quad (4)$$

$$w \cdot x + b - y_i \leq \epsilon + \xi_i^*. \quad (5)$$

Many datasets contain noise and other deviations that make it impossible to meet this constraint at all or without giving up the simplicity requirement. To properly handle these cases, the slack variables ξ_i and ξ_i^* are added to the constraint for additional flexibility and the fitness is penalized according to the parameter C . The latter parameter determines the trade-off between flatness of the function and the deviations from the estimate. By transforming inequality into equality constraints, the optimization problem is reformulated in the following way:

$$\begin{aligned}
& \text{maximize} && \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j) \\
& && - \epsilon \sum_{i=1}^l (\alpha_i - \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \\
& \text{subject to} && \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\
& && \text{and } \alpha_i, \alpha_i^* \in [0, C[,
\end{aligned}$$

where α_i and α_i^* are Lagrange multipliers. It can be observed that $w = \sum_{i=1}^l (\alpha_i - \alpha_i^*)x_i$ and $f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*)x_i \cdot x + b$. Thus, the function f is entirely characterized by the scalar product between the training points. Then, this optimization problem can be efficiently solved using quadratic programming techniques [16].

In addition, this characterization via scalar products allows an easy extension from the linear case to the non-linear one by applying a suitable non-linear mapping θ to the data prior to training the model. Although such mappings can be computationally demanding, for some θ there exist such functions k that $k(x, y) = \theta(x) \cdot \theta(y)$, which allow the efficient calculation of scalar products in the codomain of θ . These functions k are called kernel functions. Three popular kernel functions are considered in this paper: 1) Linear $k(x, y) = x \cdot y$; 2) Radial Basis $k(x, y) = e^{-\sigma \|x-y\|^2}$; 3) Sigmoid $k(x, y) = \tanh(\gamma x \cdot y + c_0)$.

In order to build up an efficient intelligent system it is fundamental to properly select the parameters ϵ , C , to design the preprocessing scheme and the related parameters as well as the kernel function k with its corresponding parameters. In this study, we propose an alternative for finding the proper estimator by combining the outputs of several, differently modelled, weaker estimators as an ensemble. Such ensemble methods have been found to be very effective tools for various machine learning tasks in a survey in [8]. In this study we model the selection of each of the regression tasks sub-components, scaling, feature reduction, and regressor kernel selection by assigning them weights. Each weight represents the selection probability of the component whose weight is associated. Thus, our problem consists of finding the optimal weights for each sub-component along with their related parameters.

Ensembles are constructed according to the optimization based scheme in Fig. 3. First, Bagging (bootstrap aggregation) is done to avoid overfitting estimators to fitness dataset and subset of 20 samples are selected from the overall training set for each regressor. Then, the components for the regressors are selected according to weights given by the optimization process and their respective parameters are picked according to Table 1. The trained regressors are then tested with the test set samples and their average error is passed to the optimizer, which then proceeds to search for better set of parameters.

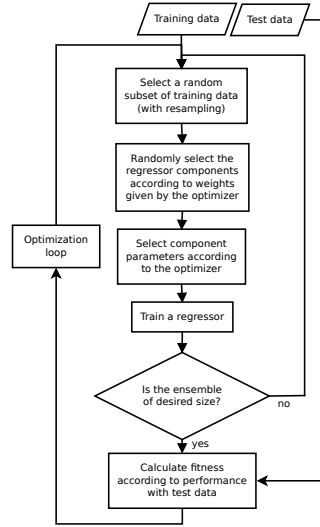


Fig. 3. General overview of the proposed system

The parameters to be selected and their respective range of variability are listed in Table 1.

Table 1. Parameters for the optimization problem

Variable	effect	range
x_1	C	$[2^{-5}, 2^{15}]$
x_2	ϵ	$[2^{-15}, 2^3]$
x_3	γ for RBF kernel	$[2^{-15}, 2^3]$
x_4	σ for Sigmoid kernel	$[2^{-15}, 2^3]$
x_5	coefficient for Sigmoid kernel	$[0, 1]$
x_6	T_{pca}	$[0.0000001, 0.5]$
x_7	Weight for Linear SVR kernel	$]0, 1[$
x_8	Weight for RBF SVR kernel	$]0, 1[$
x_9	Weight for Sigmoid SVR kernel	$]0, 1[$
x_{10}	Weight for using no reduction	$]0, 1[$
x_{11}	Weight for using PCA reduction	$]0, 1[$

The related goal is to find the set of parameters listed in Table 1 such that the median error from the actual biomass values of the samples is minimized. A training set is used to perform the machine learning while test set is used to calculate the fitness.

3 Numerical Results

This study was conducted in Agrifood Research Finland (MTT)'s experimental Hovi crop field, which is situated in Vihti, Finland. For this study, MTT arranged

a test season where growth between plots were varied using different seed and pesticide amounts during sowing.

The data consists of 91 test plots which were imaged using NIR capable UAV drone. The images were then postprocessed into an ortophotograph and a Digital Elevation Model, which describes the terrain height. DEMs are commonly used as basis for building maps and geographic information systems and can be constructed from sets of plain 2D images using image correlation and stereoscopic vision techniques (see, [12] for survey of this topic). For our application we have acquired a DEM of the target field using the stereoscopic vision techniques. The test plot locations and reference points for DEM calculation were measured using Real Time Kinematic GPS. Currently, our spectral data consists of of the DEM and near-infrared part of the spectrum, due to our UAVs occasional and catastrophic ineptitude in being aerial.

The test plots, which were randomly divided in the training and test sets, used in the training of the estimator, plus the validation set, which is used to evaluate the resulting ensembles. Each set consists of 30 samples. The target attribute in this study is the total dry biomass of the test plots and the reference values were acquired by manually collecting samples from the test plots and oven drying and weighting them.

The proposed model of the ensemble learner is trained using the three following optimization algorithms:

1. Proximity based Differential Evolution (Prox-DE) [4]
2. Frankenstein-Particle Swarm Optimization (F-PSO) [9]
3. Covariance Matrix Adaptation Evolution Strategy (CMA-ES) according to the implementation given in [7]

The Prox-DE algorithm is a Differential Evolution scheme which, instead of randomly (with uniform distribution) selecting the individuals undergoing mutation, employs a probabilistic set of rules for preferring the selection of solutions closely located to each other. The F-PSO algorithm employs a Particle Swarm Optimization structure and a set of combined modifications, previously proposed in literature in order to enhance the performance of the original paradigm. The CMA-ES is a well-known algorithm based on Evolution Strategy employing the so called maximum likelihood principle, i.e. it attempts to increase the probability of successful candidate solutions and search steps. The distribution of the solution and their potential moves tend to progressively adapt to the fitness landscape and take its shape.

For each algorithm, 75 simulation runs were run with a budget of 55 000 fitness evaluations. The parameters of the optimization algorithms are taken from the original articles in literature and are: for the Prox-DE $F=0.7$, $Cr=0.3$, $S_{pop}=60$; for F-PSO $v_{max}=1$, $w_{min}=0.4$, $w_{max}=0.9$, $wt_{max}=360$, $S_{pop}=60$, $topology_k=2000$, $topology\ update\ period=11$; for CMA-ES $\sigma=0.5$.

Table 2 shows the performance of each algorithm. The first three columns give numerical values for average, standard deviation and the best value of distribution over 75 simulations for each algorithm. The last column visualizes the

Table 2. Fitness values achieved with 75 runs.




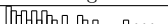

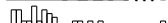
Algorithm	Average±Std.Dev	Best	Histogram
Prox-DE	2.546±0.248	1.401	
F-PSO	3.400±0.230	2.989	
CMA-ES	3.020±0.314	2.546	

Table 3. Estimation errors on the training data

Algorithm	Average±Std.Dev	Histogram
Prox-DE	5.482±4.600	
F-PSO	7.863±6.159	
CMA-ES	5.883±4.516	

distribution using sparkline histograms [15]. The distribution shows that Proximity Based Differential Evolution is both the most stable and the best performing of the algorithms in this problem. The algorithm seems to produce a good average result but is lucky in finding few extraordinarily good values during the test. The CMA-ES produces the second best values for this problem, but is less stable than the other algorithms, while the Frankenstein-PSO fails to produce a competitive result. This observed ordering of the algorithms is statistically significant (Mann-Whitney U-test, $p \leq 0.005$, see [5]). The experiment was not repeated in order to achieve the significance.

A short analysis of the convergence speed and required iterations can be made in Figure 4, which shows that CMA-ES converges slightly faster than the other algorithms.

The trained regressor ensemble is tested on the validation data not present during the training and the results are summarized in the Table 3. The table shows the average prediction error on the validation set and the sparkline histogram of the estimation error.

Discarding the failure of the F-PSO based ensemble, the result values are within $500g$, which makes the effective difference between the training methods minor. Also, the convergence behaviour suggests that if more computational resources were at hand, it could be beneficial to allow Prox-DE to run for a longer time in hope of improving the solution. When contrasting the performance to the complexity of the algorithm, it is clear that the very simple Prox-DE is the most cost efficient choice for the task. In the practical side of things, the achieved accuracy is enough to control a tractor towed fertilizer dispensers, which have relatively few dispensation settings.

4 Conclusions

A biomass estimation system based on image analysis and machine learning has been proposed in order to support precision agriculture. This system collects multiple images taken by a UAV, reconstructs a 3-D model of the field, and extracts

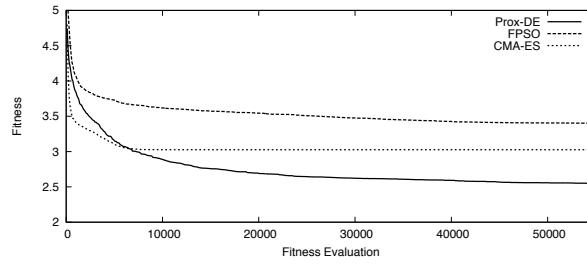


Fig. 4. Convergence of the algorithms

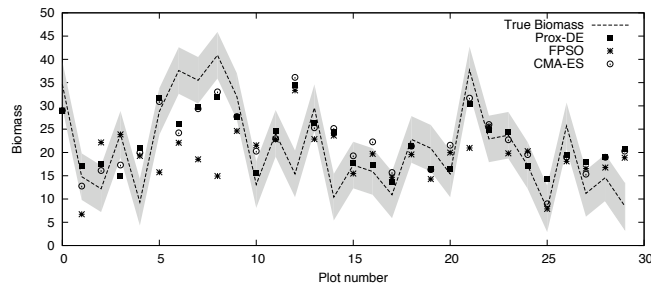


Fig. 5. Performance of the estimators with the validation dataset.

biomass information by means of a support vector regressor. The machine learning structure coordinates an ensemble of components and related parameters using the weights representing the activation probability of the corresponding components. The training has been performed by means of Prox-DE, F-PSO, and CMA-ES showed that the proposed machine learning system is reliably capable to detect the biomass present in the field by subsequent operations on the images. Amongst the three optimization algorithms considered in this study for performing the learning of the support vector regressor, Prox-DE appears to be the most reliable choice as the other two meta-heuristics seem to detect, on average, solutions characterized by a mediocre performance and then to be unable to improve upon them.

Possibilities of remote sensing applications for precision agriculture have been studied before. However, the presented estimation results in this paper are hard to compare since the premises and the target attributes in the studies differ.

Interesting areas for future studies would be to compare results with differently produced orthophotographs and Digital Elevation Model. Also, with a larger set of extracted features it would be interesting to shift the focus particularly to optimizing methods for dimension reduction.

References

1. Breiman, L.: Bagging predictors. *Machine learning* 24(2), 123–140 (1996)
2. Breiman, L.: Random Forests. *Machine Learning* 45(1), 5–32 (2001)
3. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V.: Support vector regression machines. *Advances in neural information processing systems* pp. 155–161 (1997)
4. Epitropakis, M., Tasoulis, D.K., Pavlidis, N.G., Plagianakos, V.P., Vrahatis, M.N.: Enhancing differential evolution utilizing proximity-based mutation operators. *IEEE Transactions on Evolutionary Computation* 15(1), 99–119 (2011)
5. Fay, M.P., Proschan, M.A., Others: Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys* 4, 1–39 (2010)
6. Gimelfarb, G.L., Haralick, R.: Terrain reconstruction from multiple views. *Computer Analysis of Images and Patterns* (1997)
7. Hansen, N., Müller, S.D., Koumoutsakos, P.: Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation* 11(1), 1–18 (2003)
8. Maclin, R., Opitz, D.: Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research* 1(11), 169–198 (1999)
9. Montes De Oca, M.A., Stutzle, T., Birattari, M., Dorigo, M.: Frankenstein’s PSO: A Composite Particle Swarm Optimization Algorithm. *IEEE Transactions on Evolutionary Computation* 13(5), 1120–1132 (2009)
10. Moran, M.S., Inoue, Y., Barnes, E.M.: Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing of Environment* 61(3), 319–346 (1997)
11. Schapire, R.: The strength of weak learnability. *Machine learning* 5(2), 197–227 (1990)
12. Scharstein, D.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision* 47(1), 131–140 (2002)
13. Serrano, L., Filella, I.: Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Science* 40(3), 723–731 (2000)
14. Thenkabail, P., Smith, R., De Pauw, E.: Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment* 71(2), 158–182 (2000)
15. Tirronen, V., Weber, M.: Sparkline Histograms for Comparing Evolutionary Optimization Methods. In: *Proceedings of 2nd International Joint Conference on Computational Intelligence*. pp. 269–274 (2010)
16. Vapnik, V.N.: *The Nature of Statistical Learning Theory, Statistics for Engineering and Information Science*, vol. 8. Springer (1995)
17. Zebedin, L., Klaus, a., Grubergermayer, B., Karner, K.: Towards 3D map generation from digital aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing* 60(6), 413–427 (Sep 2006)

B Publication PII

Biomass estimator for CIR-image with few additional spectral band images
taken from light UAS

Ilkka Pölönen, Ismo Pellikka, Heikki Salo,
Heikki Saari, Jere Kaivosoja, Sakari Tuominen, Eija Honkavaara

Proceedings of SPIE 8369, Sensing for Agriculture and Food Quality and Safety IV, 2012.

© SPIE

Biomass estimator for NIR image with a few additional spectral band images taken from light UAS

Ilkka Pölonen^a, Heikki Salo^a, Heikki Saari^b, Jere Kaivosoja^c, Liisa Pesonen^c, Eija Honkavaara^e

^aDepartment of Mathematical Information Tech., University of Jyväskylä, P.O.Box 35, 40014

^bVTT Photonic Devices and Meas. Sol., Espoo, Finland

^cMTT – Agrifood Research Institute Finland, Helsinki, Finland

^eFinnish Geodetical Institute, Masala, Finland

ABSTRACT

A novel way to produce biomass estimation will offer possibilities for precision farming. Fertilizer prediction maps can be made based on accurate biomass estimation generated by a novel biomass estimator. By using this knowledge, a variable rate amount of fertilizers can be applied during the growing season. The innovation consists of light UAS, a high spatial resolution camera, and VTT's novel spectral camera. A few properly selected spectral wavelengths with NIR images and point clouds extracted by automatic image matching have been used in the estimation. The spectral wavelengths were chosen from green, red, and NIR channels.

Keywords: NIR, spectral imager, estimator, biomass estimation, Fabry-Perot, precision farming, fertilizer

1. INTRODUCTION

From an ecological and economical point of view, fertilization is a notable feature of farming. The right timing of fertilization and the right amount of fertilizer are essential in maximization of grain harvest. Also the surrounding nature is connected to fertilization. If a farmer over-fertilizes his fields, then more fertilizers end up in the surrounding nature and cause eutrophication. To predict how to fertilize, the farmer has to have knowledge about the current situation of his field. We present an innovative way to produce biomass estimation for precision farming. Fertilizer prediction maps can be made based on accurate biomass estimation generated by our novel biomass estimator. Using this knowledge, a variable rate amount of fertilizers can be applied during the growing season. The system consists of a light UAS, an NIR camera, and VTT's Fabry-Perot tunable filter based spectral camera. A few properly selected spectral wavelengths with NIR images and point clouds extracted by automatic image matching will be used in the estimation. The spectral wavelengths are chosen from green, red, and NIR channels. The research project concept has been presented at SPIE 8174 [1, 2].

In the proposed concept we used a Fabry-Perot Interferometer based hyperspectral imager compatible with the lightweight UAS platforms. The concept of the hyperspectral imager has been developed by VTT [1, 2, 3]. Spectral camera images, NIR images, and point clouds extracted by automatic image matching were used for biomass estimation. Images at a few properly selected spectral bands made it possible to achieve more accurate biomass estimation than what can be achieved with pure NIR images. The University of Jyväskylä has developed a way to produce biomass estimation for wheat and barley, which can be generalized to many applications. MTT's test field at Vihti, Finland, has been used as a training database for this estimator. Green, red, and some NIR channels from hyperspectral images will be added to the NIR data. These spectral bands are centered at 570, 660, 740, 800, and 855 nm. Fertilizer prediction maps will be made based on the estimator results. In addition, variation of soils will be taken care of in the estimation. Fields need to be imaged before the growing season in order to get the soil information for the estimation.

The objective of this investigation is to evaluate different biomass estimators that utilize the novel imaging technology. The imaging system was presented by Mäkynen et al. [1] and Saari et al. [2] and the data processing and the first assessments of the biomass estimator were presented by [4,5].

1.1 Test field

The test field for the training database is located in Southern Finland. For the variance test, the field was seeded with wheat and barley, both with two varieties (Anniina, Kruunu, Voitto, and Saana, respectively). Also different amounts of fertilizer were used as we can see in Figure 1.

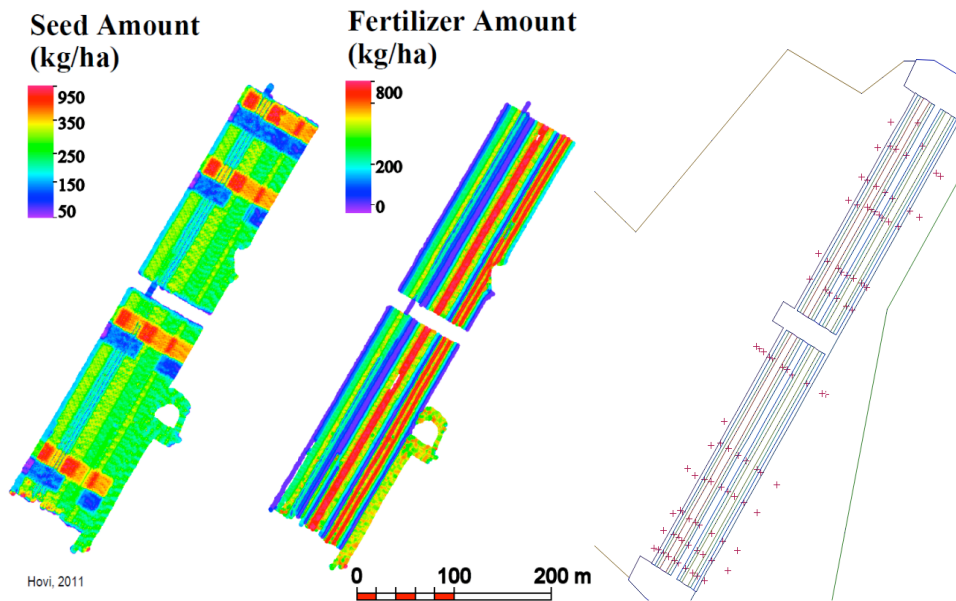


Figure 1. The applied amounts of seeds and fertilizers and the locations of 91 sample points.

The area consisted of wheat and barley test plots in which the amount of seeds and nitrogen fertilizer was varied in the range of 0 – 150 % for the seeds and 0 – 180 % for the nitrogen fertilizer (Figure 1). The type of soil and its moisture varied from north to south. The test field was planned to widely cover the realistic vegetation differences and its reflectations. 11 accurate ground control points and four reflectance reference panels were installed in the test area to allow geometric and radiometric quality assessment.

For estimator building and testing we collected 91 test samples from the field. The test samples were about one square feet large (33 cm x 33 cm) and were located with centimeter-level accuracy with RTK-GPS. All the plants were taken from the sample area. The number of main and axillary shoots was calculated, and their masses were weighted. Also the nitrogen content for each sample was determined. The samples were taken immediately after the flights. This produced labels for the training and test data. Beside these samples, also 91 samples were taken during the harvesting time. Yield and nitrogen contents were determined from those samples.

1.2 Cameras and light UAV

The novel concept was to integrate stereoscopic, lightweight, high-spectral, and high-spatial resolution imaging in the estimation process. The Microdrones MD4-1000 UAV quadcopter with a 1000 g payload capacity operated by the Finnish Geodetic Institute was used as the platform.

The hyperspectral imaging was carried out using the UASI FPI based Hyperspectral Imager developed by the VTT Technical Research Centre of Finland [1,2]. Four flights in total were performed with different sets of spectral steps and resolutions. The imaging resolution was set to VGA (640x480 pixels) to enable the recording of a continuous

spectrum for each ground pixel. The average flying speed of the MD4-1000 was 3 m/s. The whole spectral data cube could be recorded in 1.5 s, thus the last raw image was shifted by around 32 pixels related to the first raw image of a spectral data cube. The data of the UASI hyperspectral system consists of overlapping images recorded at around 50 different spectral bands. The flights were carried out from the flight height of 140 m which provided a 13 cm ground sample distance (GSD); the forward overlaps were about 80%. One of the flights was selected for the further analysis.

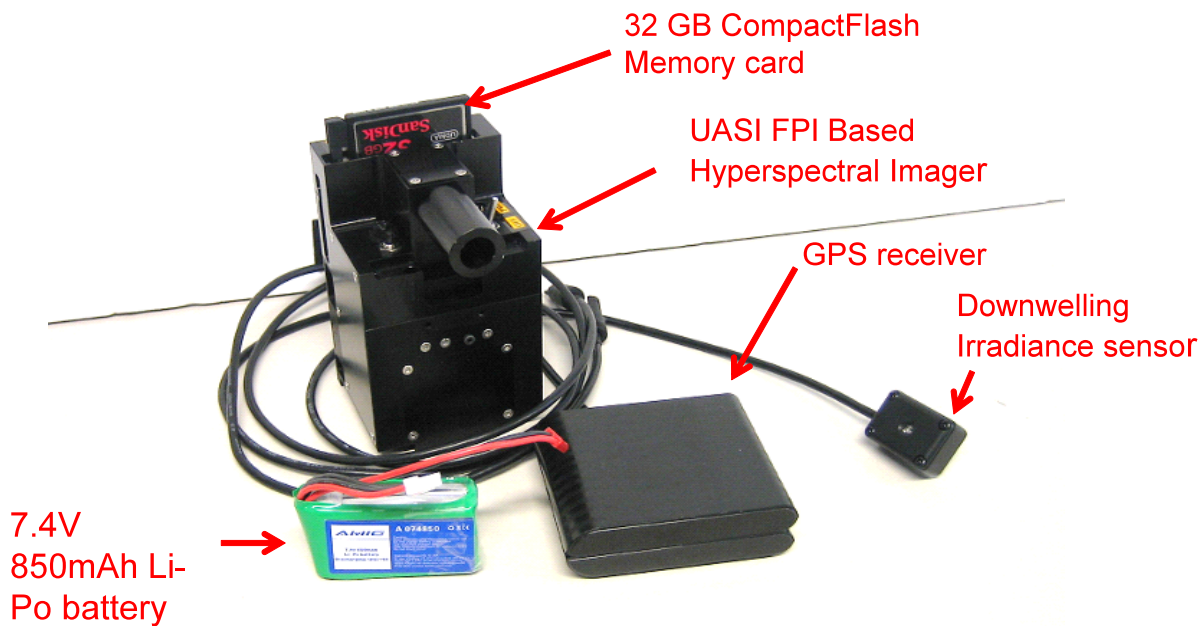


Figure 2. Microdrones MD4-1000 UAV helicopter and VTT's hyperspectral camera.

A customized Panasonic Lumix GF 1 camera (focal length 2 cm, image size 4016 x 3016 pixels, pixel size 4.5 micrometers) was used to collect high spatial resolution images. Images were collected in sunny weather from a flying height of approximately 140 m from the ground level (3 cm GSD) and forward overlaps were approximately 90%.

Image data processing included determination of exterior orientations, generation of 3D point clouds and digital surface models (DSMs), and calculation of radiometrically corrected orthophoto mosaics. The geometric processing was carried out using the Bae Systems SOCET SET v. 5.5 commercial photogrammetric workstation. A self-calibrating bundle block adjustment method was applied to determine the exterior orientations; in the case of UASI each individual channel was processed separately. The 3D point clouds and surface models were extracted with a 10

cm point interval from high spatial resolution GF1 images by novel dense image matching techniques; the Next Generation Automated Terrain Extraction software (NGATE) was used. For the radiometric processing and image mosaic generation a novel processing software that is under development at the FGI was used. A radiometric block adjustment approach was developed to eliminate the intensity differences between different images caused by the illumination changes and sensor effects, and the object reflectance anisotropy effects caused by varying view/illumination geometry in central perspective images (the so-called bi-directional reflectance distribution function, BRDF, correction). Finally, orthophoto mosaics were calculated with the help of the exterior orientation information, the DSM, and radiometric correction parameters; GSD was 3 cm with GF1 images and 20 cm with hyperspectral images. It was estimated that the final positional accuracy of the GF1 mosaic was about 20 cm and the hyperspectral mosaic about 50 cm. Examples of point clouds and image mosaics are shown in Figures 3 and 4. Details of the image processing are given in [4,5].

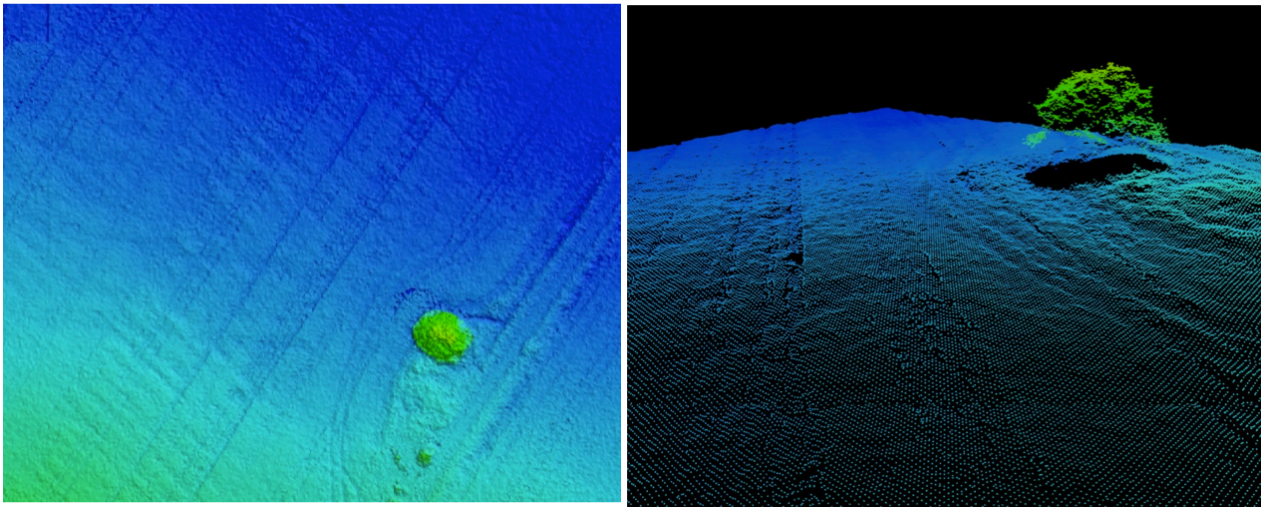


Figure 3: A digital surface model showing a tree in the test area.

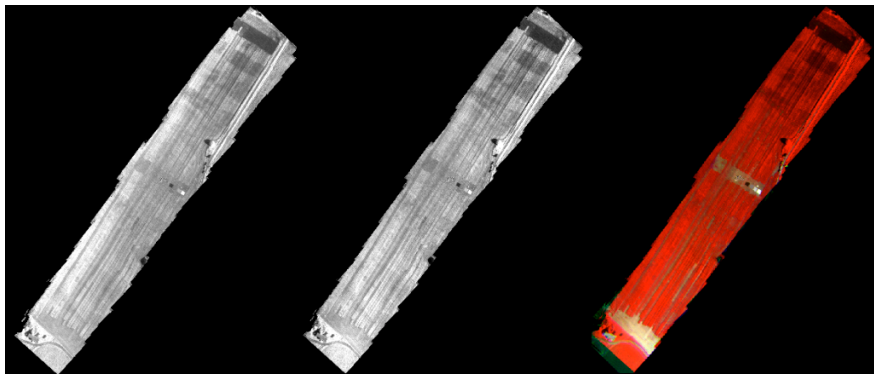


Figure 4: Image mosaics. From left: the original DN mosaic and the BRDF-corrected mosaic of the red edge channel and the BRDF-corrected color-infrared mosaic of spectral images (north is up and east is to the right, sun illumination comes from south-east). The BRDF correction compensated efficiently the brightening of the mosaic due to the view-illumination geometry in the north-west part.

2. FEATURE SPACE

In this section we describe the features extracted from images and the DSM.

2.1 Features from Hyperspectral Images

Hyperspectral imaging provides a great set of different features of the target. One hyperspectral image data cube contained around 50 different spectral bands to be used in the analysis. This enables calculating several different results from the data. In this study we used five interesting spectral bands at 570, 660, 740, 800, and 855 nm. From these bands we calculated the following features for the sample areas:

- Green channel: channel 10 (green) (5x5 pixels (1 m²) average).
- Red channel: channel 25 (red) (5x5 pixels (1 m²) average).
- Red edge channel: channel 36 (red edge) (5x5 pixels (1 m²) average).
- NIR1 channel; channel 43 (nir1) (5x5 pixels (1 m²) average).
- NIR2 channel: channel 49 (nir2) (5x5 pixels (1 m²) average).
- NDVI: NDVI using NIR2 and Red channels (5x5 pixels (1 m²) average).
- SR: Simple ratio using NIR2 and Red channels (5x5 pixels (1 m²) average).

The feature space based on hyperspectral images is in R^7 .

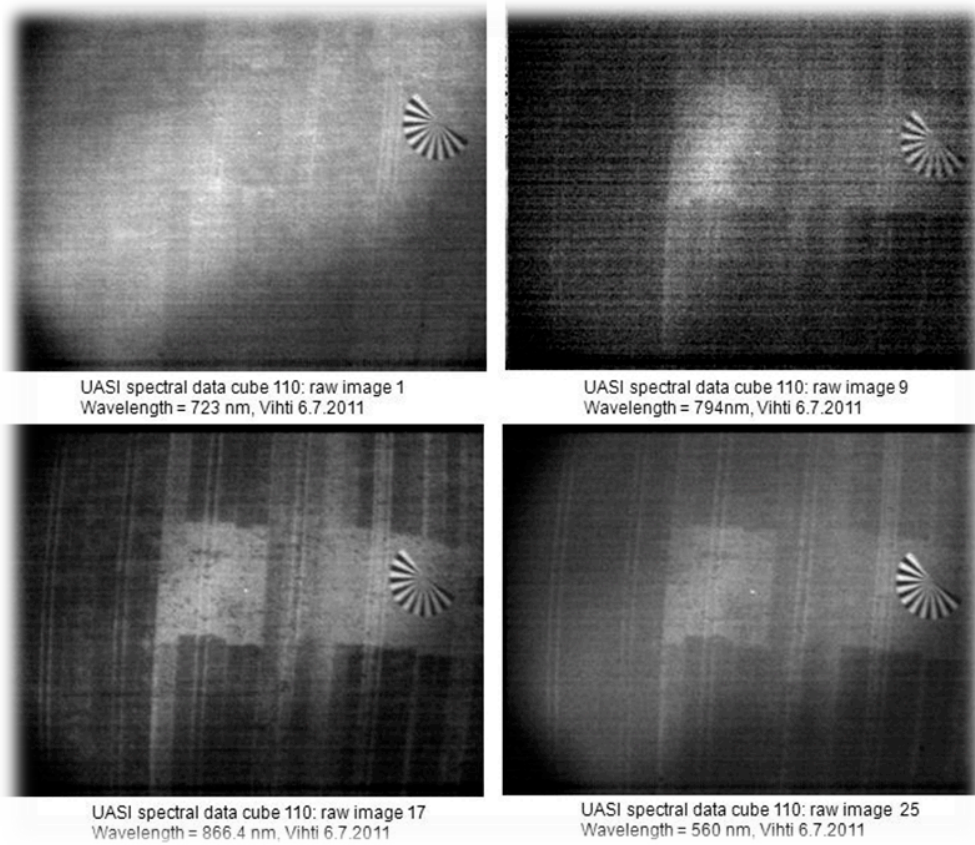


Figure 5: Spectral images with different wavelengths.

2.2 Digital Elevation Model for the Field

Based on the digital surface model it was possible to form a digital elevation model. Our DSM accuracy was 0.1x0.1 meter. We had also a digital terrain model (DTM), which was from earlier LIDAR research. The DTM's accuracy was 2x2 m. We calculated neighbors for the points in the DSM and subtracted the DTM's height from that of the DSM.

From the DSM and DEM we calculated the following different feature sets for the sample areas:

- DSM1: mean value of heights of 25 nearest neighbors
- DSM2: subtraction of 25 neighbors' minimum value from maximum
- DEM1: height of 25 nearest neighbors
- DEM2: mean value of heights of 25 nearest neighbors
- DEM3: subtraction of 25 neighbors' minimum value from maximum

The feature space based on the DEM and DSM is in R^{29} .

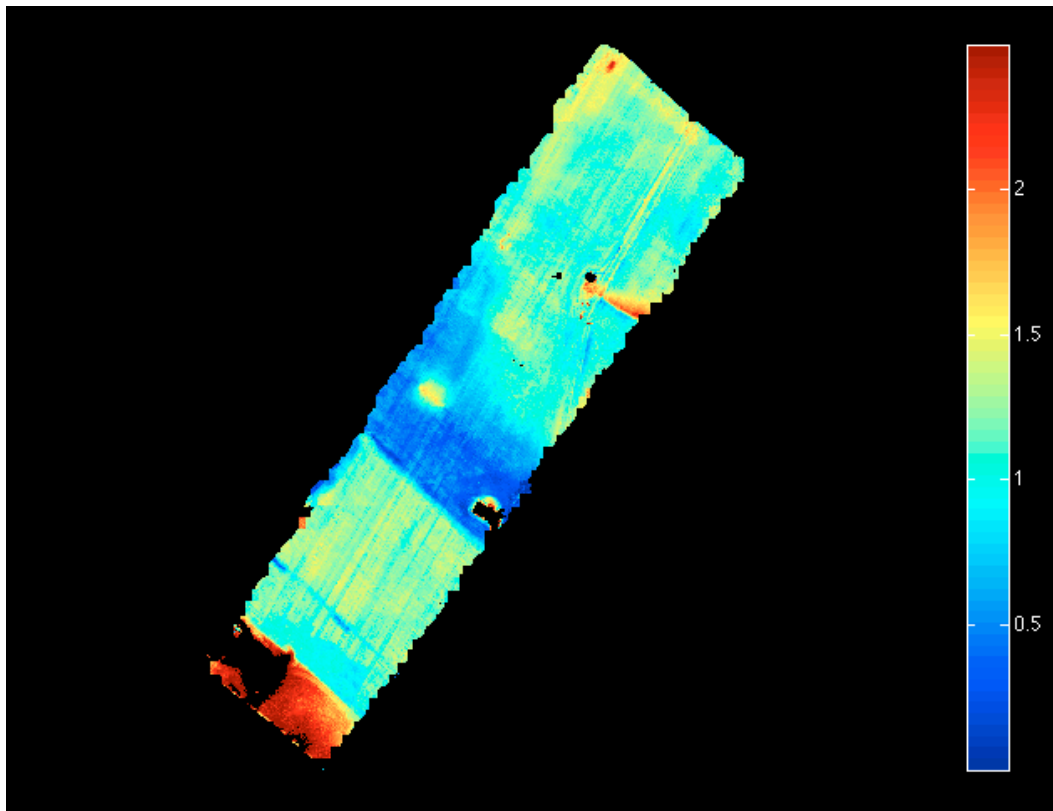


Figure 6: The digital elevation model in meters.

3. BIOMASS ESTIMATOR

After feature extraction we have 36 dimensional feature spaces in our hands. Our novel approach to the biomass estimation is based on diffusion maps and the k nearest neighbors classification. The basic idea in the estimation is to reduce the feature space of the labeled data set (includes N data points) to the lower dimension with diffusion maps. After dimension reduction we embedded new points to the embedded space and for these data points we calculated five nearest neighbors. From these neighbors we can approximate an estimation for the embedded new data points. The estimator is based on supervised learning. It needs data points for training and testing.

Diffusion maps are a relatively new dimensional reduction method presented by Coifman and Lafon [6]. The method is based on harmonic geometric analysis and spectral graph theory. Diffusion maps generate an efficient representation of complex geometric structures in a high-dimensional space. The basic idea is to construct random walks between data points x_i to achieve the Markov transition matrix. For this we used a normalized graph Laplacian. Using the Gaussian heat kernel we can calculate the matrix

$$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}, \text{ where } \epsilon > 0 \text{ and } i, j = 1, \dots, N.$$

After this we normalize W to be row stochastic and form a graph Laplacian. From the eigenvalues and eigenvectors of the graph Laplacian we can form diffusion distances. A diffusion map embeds the data into the Euclidian space. The Euclidean distance is equivalent to the diffusion distance. Therefore, it is possible to visualize data in the Euclidian space.

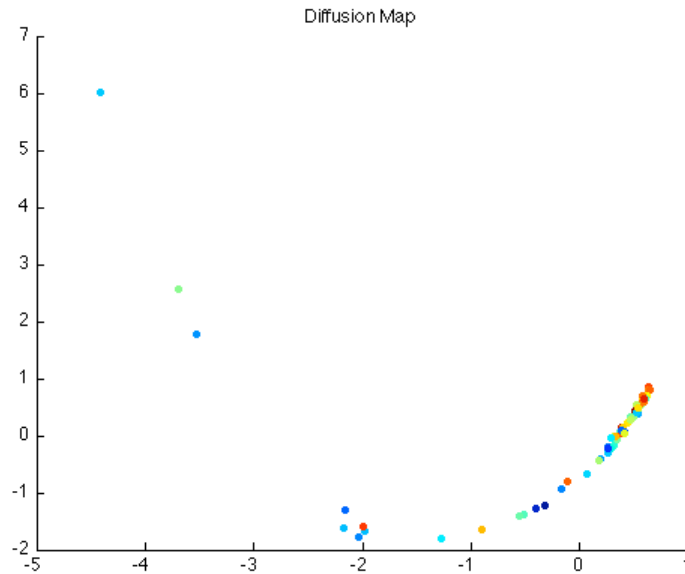


Figure 7: Diffusion mapping, 2nd and 3rd eigenvalues.

In this embedded space we estimated data points by calculating the nearest neighbors for new data points. To embed the new data points from the feature space to the embedded Euclidean space we used Geometric Harmonics [7]. Geometric Harmonics is based on the Nyström method [8]. Basically, the idea in Geometric Harmonics is to extend empirical functions f defined on set A to another set B . In our case set A is a training set and B is a test set. Function f is composed during the diffusion map process.

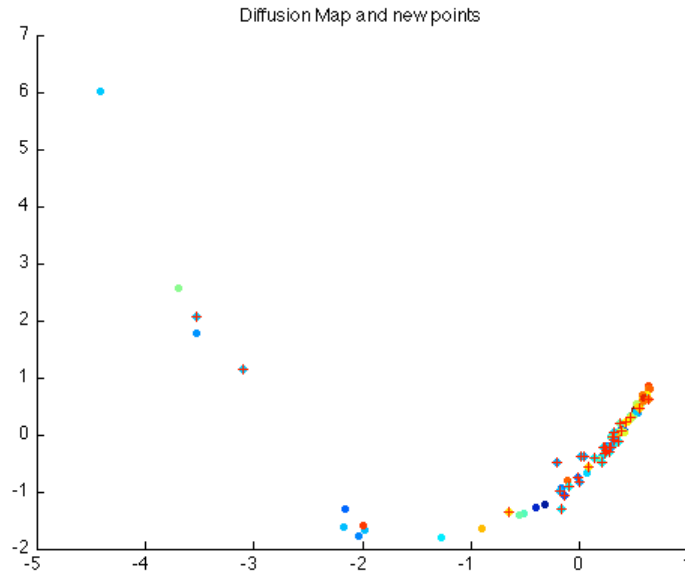


Figure 8: Diffusion mapping with embedded new data points, 2nd and 3rd eigenvalues.

For the embedded new points we calculated five nearest neighbors from the training set. Based on these neighbors we calculate mean values for these points from a label set.

We also implemented the nearest neighbors and support vector regression based estimators, giving us alternative results that can be compared to the diffusion map based method.

4. ESTIMATION RESULTS

Unnatural variance in the test field was large due to different amounts of fertilizer and seeds per square meter. To reduce the effect of this unnatural variance we clustered data based on the knowledge on how much seeds and fertilizer had been used, how deep the sowing had been done, and what kind of strain were present in that test sample. We reduced dimension with diffusion maps and then used a kmeans algorithm for clusterization. Based on this clusterization the data was divided into the training and testing sets in the ratio 2:1. In the training set there were 60 samples and in the testing set 31 samples. First we tested only features from hyperspectral images. We ran the estimations 50 times. The results are presented in Table 1.

Table 1. Estimation results only from hyperspectral images.

ESTIMATION RESULTS ONLY FROM HYPERSPECTRAL IMAGES (mean values of 50 runs)		
	R ²	RMSE (g/ft ²)
KNN.....	0.02	9.24
SVR.....	-0.27	11.12
DM + KNN.....	0.12	9.78

In addition, we increased the amount of features with DEM and DSM based features that are described in Section 2. The increased feature space made estimators more precise as we can see in Table 2.

Table 2. Estimation results from hyperspectral images and height features.

ESTIMATION RESULTS WITH HEIGHT FEATURES (mean values of 50 runs)		
	R^2	RMSE (g/ft ²)
KNN.....	0.16	9.14
SVR.....	-0.07	10.19
DM + KNN.....	0.13	9.14

After this we decreased the amount of hyperspectral image features and used only the near infrared and red edge channels and height features. It seems to be that the near infrared features and the red edge channel characterize features. After this estimation the results improved significantly.

Table 3: Estimation results from NIR images, red edge channel, and height features.

ESTIMATION RESULTS WITH HEIGHT AND NIR FEATURES (mean values of 50 runs)		
	R^2	RMSE (g/ft ²)
KNN.....	0.39	7.83
SVR.....	0,29	8,43
DM + KNN.....	0.50	7.05

Lastly, we tried only DEM and DSM based features. However, it did not improve the results as we can see in Figure 4.

Table 4: Estimation results from height features.

ESTIMATION RESULTS ONLY FROM HEIGHT FEATURES (mean values of 50 runs)		
	R^2	RMSE (g/ft ²)
KNN.....	0.17	9.09
SVR.....	0.06	9,69

As we can see, the best result was achieved with combined diffusion maps and a knn estimator. The used feature space was a combination of hyperspectral images, which are near infrared, and height features. It seems that our DEM adds value to the estimation. Also our approach in high-dimensional data analysis is productive.

5. CONCLUSIONS AND FUTURE WORK

The novelty of this work comes from using BRDF-corrected images and the DEM in a few different estimators. Within the presented results it seems to be possible to estimate the amount of biomass in the fields from light UAV with hyperspectral and NIR cameras. Based on these estimates it is possible to create more precise fertilization plans.

Our research project continues until the end of 2012. Our ambition is to create better results for the estimation of the biomass and nitrogen in the field. We are also interested in predicting the total amount of the yield. In order to have better hyperspectral imaging quality in the UAV flight campaigns in summer 2012, VTT will develop an updated UAV-compatible imaging system for forest and agriculture applications [9].

For the campaigns of summer 2012, the test field will be more homogeneous with less variance in the amount of seeds, fertilization, and variety. Based on these campaigns' imaging results, we are going to include texture features to our feature space. With improved hyperspectral imaging quality we are able to use more spectral channels in the estimator and to also add spectral features to the feature space.

ACKNOWLEDGEMENT

The research has been funded by Tekes (Finnish Funding Agency for Technology and Innovation), VTT Strategic Research, the University of Jyväskylä, MTT Agrifood Research Finland, and Finnish Forest Research Institute. We would also like to thank Mr. Teemu Hakala and Mr. Tomi Rosnell from Finnish Geodetic Institute (FGI) for operating the MD4-1000 UAV helicopter.

- [1] Mäkynen, J., Holmlund, C., Saari, H., Ojala, K., Antila, T., "Unmanned aerial vehicle (UAV) operated megapixel spectral camera", Proc. SPIE 8186B (2011).
- [2] Saari, H., Pellikka, I., Pesonen, L., Tuominen, S., Heikkilä, J., Holmlund, C., Mäkynen, J., Ojala, K. and Antila, T., "Unmanned Aerial Vehicle (UAV) operated spectral camera system for forest and agriculture applications", Proc. SPIE vol. 8174 (2011).
- [3] Saari H., "Spectrometer and interferometric method", US Patent US 8,130,380, Mar. 6, 2012.
- [4] Honkavaara, E., Hakala, T., Saari, H., Markelin, L., Mäkynen, J., Rosnell, T., "A process for radiometric correction of UAV image blocks." Photogrammetrie, Fernerkundung, Geoinformation (PFG) 2/2012.
- [5] Honkavaara, E., Kaivosoja, J., Mäkynen, J., Pellikka, I., Pesonen, L., Saari, H., Salo, J., Hakala, T., Markelin, L., Rosnell, T., "Hyperspectral reflectance signatures and point clouds for precision agriculture by light weight UAV imaging system" Accepted for publication in ISPRS Melbourne Congress, Australia, 2012.
- [6] Coifman, R., Lafon, S., "Diffusion maps". Applied and Computational Harmonic Analysis 21(1), 5–30 (2006)
- [7] Coifman, R., Lafon, S., "Geometric harmonics: A novel tool for multiscale out-of-sample extension of empirical functions", Applied and Computational Harmonic Analysis 21(1), 31-52 (2006)
- [8] Nyström, E.: "Über Die Praktische Auflösung von Integralgleichungen mit Anwendungen auf Randwertaufgaben", Acta Mathematica, 54(1), 185-204 (1930)
- [9] Mäkynen, J., Holmlund, C., Saari, H., Mannila, R., Antila, T., "Multi- and hyperspectral UAV imaging system for forest and agriculture applications", To be published in SPIE . Vol. 8359-25 (2012).

C Publication PIII

Methods for estimating forest stem volumes by tree species using digital surface model and CIR images taken from light UAS

Heikki Salo, Ville Tirronen, Ilkka Pölönen,
Sakari Tuominen, Andras Balazs, Jan Heikkilä, Heikki Saari

Proceedings of SPIE 8390, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, 2012.

© SPIE

Methods for estimating forest stem volumes by tree species using digital surface model and CIR images taken from light UAS

Heikki Salo^a, Ville Tirronen^a, Ilkka Pölönen^a, Sakari Tuominen^b, Andras Balazs^b, Jan Heikkilä^c and Heikki Saari^d

^a Department of Mathematical Information Tech., University of Jyväskylä, P.O.Box 35, 40014

^b Metla – Finnish Forest Research Institute, Vantaa, Finland ^c Pieneering Ltd., Helsinki, Finland ^d VTT – Photonic Devices and Meas. Sol., Espoo, Finland

ABSTRACT

In this paper we consider methods for estimating forest tree stem volumes by species using images taken from light unmanned aircraft systems (UAS). Instead of using LiDAR and additional multiband imagery a color infrared camera mounted to a light UAS is used to acquire both imagery and the DSM of target area. The goal of this study is to accurately estimate tree stem volumes in three classes. The status of the ongoing work is described and an initial method for delineating and classifying treetops is presented.

Keywords: Remote sensing, forest, tree stem volume, species-specific, UAV, UAS, CIR images, individual tree recognition

1. INTRODUCTION

Forest inventing using remote sensing has been actively under study in last decades. In this paper, we consider methods for remote sensing tree volumes by species from small (20 m by 20 m) forest stands.

The starting point of remote sensing forests has traditionally been to study forest with features that describe a small neighbourhood of trees. While producing good results in total tree volumes, volumes by tree species are more challenging to estimate accurately. There are studies^{1,2} on estimating species-specific tree stem volumes using airborne laser scanning and aerial photographs.

Recently, forest inventing based on recognition of single trees has also been under many studies.³ Single tree recognition consists of recognizing tree tops and classifying trees to species. Main drawbacks of inventing using single recognized trees are computational complexity and the fact that recognizing non-dominant trees is significantly harder.⁴ When studied using orthoimages and digital elevation models the smaller trees under dominating trees are totally left without sight.

In last two decades airborne laser scanning (ALS) has been used successfully for forest inventory in Finland, Norway and Sweden^{4,5} ALS data is often used as point height distributions that describe the ground in plot level. However, many LiDAR equipments are so heavy that they cannot be flown using UAV's that can only carry payloads in range of 0.5 kg.

If methods based on UAV equipment can be used they offer a lightweight solution for getting information from small or medium-sized areas. The orthoimage and Digital Surface Model (DSM) are processed using photogrammetry from acquired digital images from a regular CIR camera flown by an UAV system.

In this paper, we consider additional features derived from recognized single trees to be used along with stand-level features. The added features could add value as they provide describe structure of the stand, while not being sufficient to perform well alone.

The remainder of this paper is organized as follows. In Section 2 we describe the study material. Section 3 shows an overview and details of the used methods. Finally, Section 4 contains conclusions and ideas for future work.

For further author information: (Send correspondence to Ilkka Pölönen)
Ilkka Pölönen: E-mail: ilkka.polonen@jyu.fi, Telephone +358 400 248 140
Heikki Salo: E-mail: heikki.salo@jyu.fi, Telephone +358 50 33 97 894

2. MATERIALS

The study area is located in the municipality of Lammi, in Southern Finland (approximately 61° 19' N and 25° 11' E). The study area covers a part of state-owned Evo educational forest area covering approximately 2000 ha. The forests of the study area are dominated by coniferous tree species, mainly Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). Of the deciduous species birches (*Betula pendula* and *B. pubescens*) are most common. Other, mainly non-dominant tree species present in the study area are aspen (*Populus tremula*), grey alder (*Alnus incana*), rowan (*Sorbus aucuparia*), contorta pine (*Pinus contorta*), larches (*Larix sp.*) and firs (*Abies sp.*).

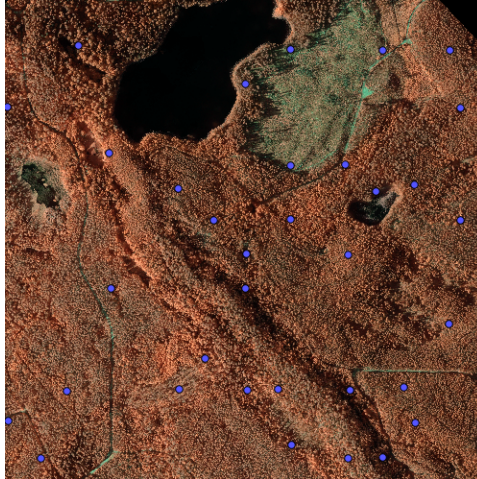


Figure 1. Field plots on CIR orthoimage with GSD of 10 cm.

The forest area has a dense grid of sample plots measured for research purposes by HAMK University of applied sciences. The field data employed in this study consists of hundred fixed-radius (9.77 m) circular plots that were measured in 2007-2010 (updated to 2010). From each plot, all living trees with a breast-height diameter at least 50 mm have been measured as tally trees. From each tally tree the following variables have been measured: location (compass bearing and distance from plot centre), tree species, crown layer, diameter at breast height, height and height of living crown. The plots were located with Trimble's GEOXM 2005 Global Positioning System (GPS) device, and the locations were processed with local base station data, resulting in an average error of approximately 0.6 m.



Figure 2. A Gatewing UAV system

Aerial images were acquired from flights performed in August 2011. The flights were performed with Gatewing that had an autopilot and Ricoh GR Digital III NIR camera mounted in it. Aereas covered in this study were

covered with four separate flights performed between 12am and 6pm. The total coverage of the images were first seen as success, but the areas far from nadir were skewed and made processing images challenging.

We also have areas covered using VTT's lightweight hyperspectral imager. The hyperspectral images taken 2011 aren't yet processed as orthoimages as the Fabry-Pérot interferometer prototype⁶ with the selected optics didn't provide images with sufficient amount of luminosity. The same prototype performed well in the field of agriculture, and in summer 2012 a new prototype that addresses this luminosity issue will be tested.

The digital surface models and CIR orthoimages used were processed by PIEngineering Ltd.*. The ground sample distance (GSD) of orthoimage is 10 cm and the GSD of current DSM is 20 cm.

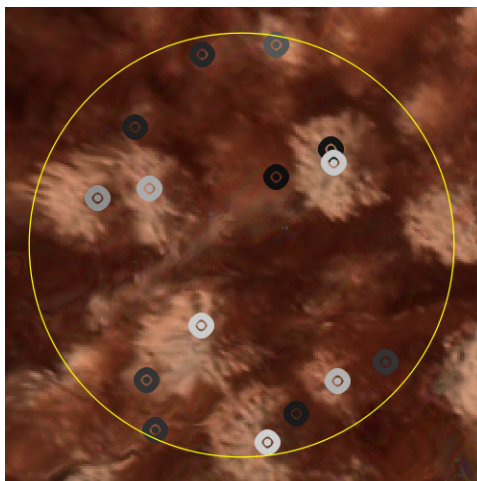


Figure 3. A view to a single field plot projected on a CIR image. Dots represent indexed trees, the brighter the spot the taller the tree.

For calculating an digital elevation model (DEM) out of this digital surface model, additional information was required as the DSM didn't provide enough ground points to evaluate the height of the vegetation. We used additional ground level dataset acruired earlier by laser scanning. So far the imagery acquired by photogrammetry this way can therefore be used only for updating forest information using lightweight setup, since it cannot be used alone.

As there were twists in original images, the orthoimage and digital surface model have silhouettes of treetops between trees instead of ground. This leads to difficulties in recognizing individual treetops and in estimating the ground level.

Currently, the number of sample plots available to study is relatively small, since the image preprocessing is at the moment demanding and isn't fully automated. We're investigating methods for making it possible to automate the processing of the orthoimage and digital surface model.

3. METHODS

A full estimation chain from pre-processed DEM and orthoimages to estimation results goes as follows:

$$\mathbf{G} \rightarrow \mathbf{S} \rightarrow \mathbf{F} \rightarrow \mathbf{D} \rightarrow \mathbf{R}$$

where \mathbf{G} stands for reading ground-truth from plot data, \mathbf{S} for separating data to train and test sets, \mathbf{F} for extracting all feature sets, \mathbf{D} for dimension reduction and finally \mathbf{R} for regression. We're studying the effects

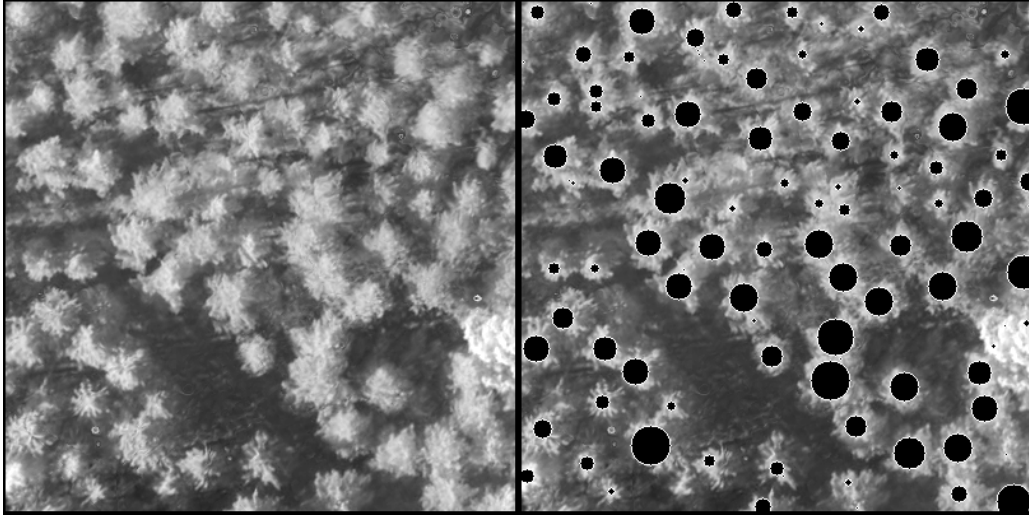


Figure 4. Simple treetop delineation by a scale space method.

of different choices in all steps. Most of all we're interested in finding well-performing features and feature combinations. The remains of this Section outlines some thoughts and details of the above mentioned steps.

The number of ground plots available to study is currently relatively low. For this reason, the selection of train and test sets must be considered carefully. For providing a fair separation of plots for train and test data we're using a k-means clustering algorithm. The plots are clustered by their tree volumes by species (pine, spruce and deciduous) to clusters, which are then unpacked to split the plots to train sets and test sets.

Plot-level features currently under study consist of height level distribution features from DEM and image features from the CIR image, much like in studies performed with ALS data combined with aerial images.¹ There are several studies that focus on plot level tree volumes and species-specific tree volumes. While the CIR images provide many possibilities from learning plot-level features, they also make possible detecting treetops.

On the assumption, that treetops would be efficiently characterized by DEM, we've tested several local mode detectors such as hill climbers and the mean shift⁷ mode detectors. However, on some plots we found that the DEM was too inaccurate in delineating trees even for a human observer. Thus, we reverted to a simple scale-space investigation of the green channel of the CIR images. First, the image is decomposed into a gaussian pyramid, from which we can obtain multiple scale difference-of-gaussians (DoG) filters. In our case, we can detect treetops with a sufficient accuracy by selecting areas where both the first vs. third and the second vs. fourth octaves of the DoG filters have high response (see Figure 4 for an example result). In other words, the pixelwise indicator image S for a treetop is as follows:

$$S(x, y) = \begin{cases} 1, & \text{when } \text{DoG}_{4,2} > t_1 \text{ and } \text{DoG}_{0,3} > t_2 \\ 0, & \text{otherwise,} \end{cases}$$

where the $\text{DoG}_{w,h}$ is defined as

$$[\text{DoG}_{j,k}(f)](x, y) = [G(x, y, k\sigma) - G(x, y, j\sigma)] * f(x, y),$$

and the thresholds t_1 and t_2 are selected according to luminance of the dataset (here they are set to 0.02 and 0.06 respectively), and the scale constants are set according the image resolution. To filter out the inevitable false detections by the coarse texture of the images, we cull the detections using a height limit on the DEM and

*Additional information can be found from PIEngeering website at <http://www.pieneering.fi>

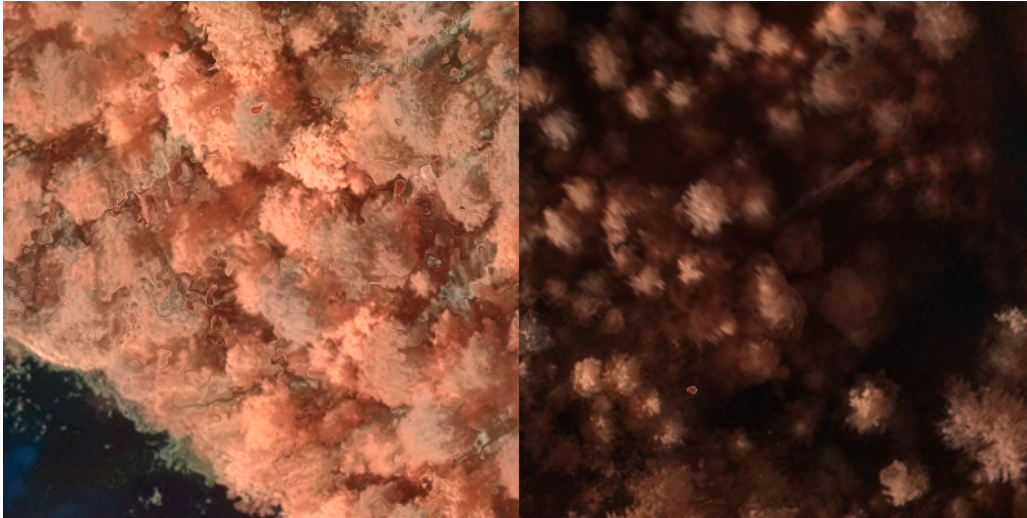


Figure 5. A pick of challenging parts of orthoimages displaying challenges like twirls, blur and varying lightning conditions.

restricting ourselves to blobs, which cover, at minimum, a circle of 50 cm on the terrain. At the moment this seems to match well to the average true positive detections.

For each blob, we extract a rectangular subimage of twice the radius of the detected blob. This cropped image is then used for feature extraction and subsequently for training a Support Vector Machine classifier to recognize the tree species. Initial results on classifying treetops to three classes using a small subset of the data set seem promising, yet the difficulty of the task varies greatly throughout the orthoimage. The classifying is at the moment performed using only the histograms of the available color channels.

It's possible that further studying the derived features from recognized treetop would enhance the performance. The tree height and stem diameter are the main contributors to tree stem volume, while the crown diameter only contributes to that information. There are studies⁸ of measuring crown diameter and it's influence to tree volumes.

As regressor there are currently two different types of regressor under study. K nearest neighbours (k-NN) and support vector regressors (SVR) represent machine learning with instance based learning and model-fitting learning. Both of these algorithms have been widely used and studied in the field of remote sensing.

4. CONCLUSIONS AND FUTURE WORK

Main challenges regarding this study have been related to the quality of the raw CIR images. The photogrammetric composition of DSM and orthoimage will get easier and produce better results when repeated using a better CIR camera with more suitable focal lengths to avoid having to use skewed part of the images far from the nadir.

Using a better camera will also provide better spectral resolution, which will make the tree recognition easier. For the flight campaigns in summer 2012 a new version of VTT's lightweight hyperspectral imager will also be used. The advantages of using hyperspectral images will be studied.

We will also study effects on bidirectional reflectance distribution function (BRDF) based correction to orthoimages. BRDF corrections would likely give benefits to the descriptiveness of the extracted features.

Plot-level features should also be selected carefully in order to maximize the performance. Genetic algorithms could be used in feature selection.⁹ Genetic algorithms have already shown good performance in forest invention.¹⁰ Used parameters could also be optimized together with different estimator and dimension reduction combinations.¹¹

The digital surface model acquired by means of photogrammetry doesn't seem to be sufficient to stand-alone usage as it doesn't contain enough ground points for calculating elevation model. Likely application could therefore be for getting updates to forest stand details using UAS's for evaluating storm damages.

ACKNOWLEDGEMENTS

This research has been funded by Tekes (Finnish Funding Agency for Technology and Innovation), VTT Strategic Research, the University of Jyväskylä, Metla Finnish Forest Research Institute, Pioneering Ltd. and MTT Agrifood Research Finland.

REFERENCES

- [1] Niska, H., Skon, J.-P., Packalen, P., Tokola, T., Maltamo, M., and Kolehmainen, M., "Neural Networks for the Prediction of Species-Specific Plot Volumes Using Airborne Laser Scanning and Aerial Photographs," *IEEE Transactions on Geoscience and Remote Sensing* **48**, 1076–1085 (Mar. 2010).
- [2] Packalen, P. and Maltamo, M., "Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs," *Forest Science* **52**(6), 611–622 (2006).
- [3] Korpela, I., *Individual tree measurements by means of digital aerial photogrammetry*, PhD thesis (2004).
- [4] Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H. k., Persson, A. s., and Söderman, U., "Laser scanning of forest resources: the nordic experience," *Scandinavian Journal of Forest Research* **19**, 482–499 (Dec. 2004).
- [5] Næsset, E., "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sensing of Environment* **80**, 88–99 (Apr. 2002).
- [6] Mäkynen, J., Holmlund, C., Saari, H., Ojala, K., and Antila, T., "Unmanned aerial vehicle (UAV) operated megapixel spectral camera," in [*Proceedings of SPIE*], **8186**, 81860Y–81860Y–9 (Oct. 2011).
- [7] Comaniciu, D. and Meer, P., "Mean shift analysis and applications," in [*Proceedings of the Seventh IEEE International Conference on Computer Vision*], 1197–1203 vol.2, IEEE (1999).
- [8] C Popescu, S., H Wynne, R., and F Nelson, R., "Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass," *Canadian Journal of Remote Sensing* **29**(5), 564–577 (2003).
- [9] Chaikla, N., "Genetic algorithms in feature selection," in [*IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028)*], **5**, 538–540, IEEE (1999).
- [10] Holopainen, M., Haapanen, R., Tuominen, S., and Viitala, R., "Performance of airborne laser scanning- and aerial photograph-based statistical and textural features in forest variable estimation," in [*SilviLaser 2008*], 105–112 (2008).
- [11] Salo, H., Tirronen, V., and Neri, F., "Evolutionary Regression Machines for Precision Agriculture," *Lecture Notes in Computer Science* **7248**, Springer Berlin Heidelberg, Berlin, Heidelberg (2012).