An Integrated Image Analysis System for the Estimation of Soil Cover

Peter Riegler-Nurscher [a,*], Johann Prankl [a], Thomas Bauer [b], Peter Strauss [b], Heinrich Prankl [a]

[a] Josephinum Research, Rottenhauserstraße 1, Wieselburg, Austria

[b] Inst. for Land and Water Management Research, Federal Agency for Water Management, Petzenkirchen, Austria

* Corresponding author. Email: p.riegler-nurscher@josephinum.at

Abstract

The estimation of the soil cover by residues and living plant material is a fundamental issue for the sustainable cultivation of arable land. Especially, the percentage of soil covered by material is one of the main factors to protect soil from erosion. Available methods for the measurement are either subjective, depending on an educated guess or time consuming, e.g., if the image is analysed manually at grid points. First approaches using automatic segmentation of objects and classification into soil, residue and plants show promising results, but they rely on a manual adjustment of parameters. We introduce novel pixel-wise features to a special variant of machine learning technique – namely the entangled forest – in order to improve the robustness and the generalisation of the classification of individual image pixel into soil, living plants and residues.

A classical machine learning approach consists of different computer vision steps, e.g., segmentation into homogeneous patches, patch description (with colour, contour, texture ...), classification and a smoothing step. In contrast, the proposed entangled forest, a variant of random decision forest, classifies individual pixel using simple pixel-wise comparisons to neighbouring pixels with a trained offset, and in addition smoothing is achieved by maximum a posteriori features, where the decision in a specific node depends on the a posteriori label probability of a neighbouring pixel of the previous decision tree layer.

We compare our system with a data set manually annotated at grid points. The images have been taken in different lighting conditions of soil covered from 0% up to 100% with different materials, such as living plants, residues, straw material and stones. The results indicate that our method is as accurate as the manual annotation with a mean deviation of 6% to the grid method.

Keywords: image analysis, random forest, estimation of soil cover

1. Introduction

Worldwide many agricultural practices aim on soil protection strategies. National agricultural subsidy programmes tend to regulate minimum soil cover. Especially for management operations that are focusing on soil erosion mitigation the effectiveness is directly driven by the amount of soil covered. Hence there is a need for quick and reliable methods for estimating the mean soil cover rates of a field site. Figure 1 shows soil covered with living plants and residues and the classification result.

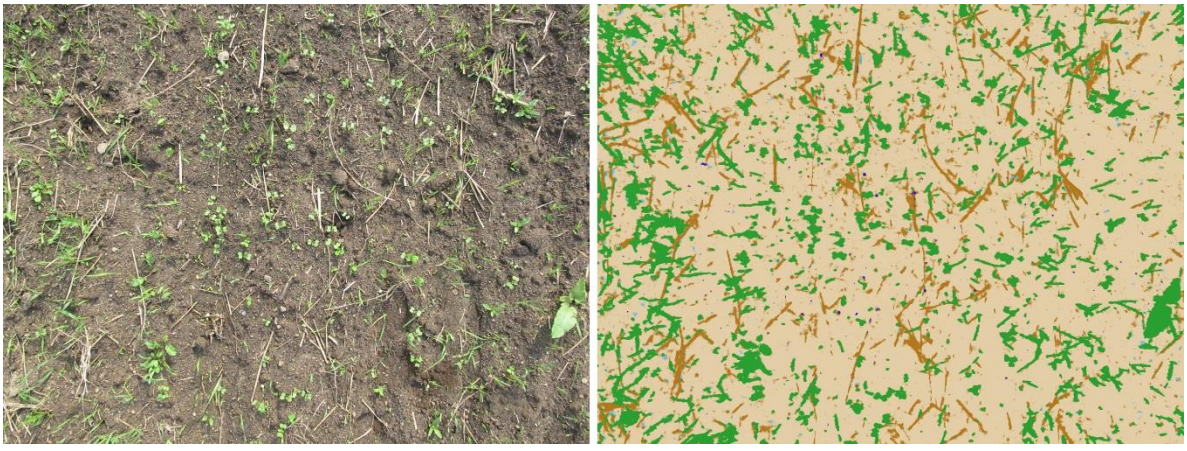

Figure 1: Soil covered with living plants and residues (left) and the classification result (right).

So far there are not many reliable methods available for the estimation of living and dead cover on a soil surface. Either manual analysis onsite (Marques et al., 2007; Mohammad and Adam, 2010) or manual image analysis (Hartwig and Laflen, 1978; Corak et al. 1993; Morrison et al. 1993) methods exist. Manual methods are often time consuming and very subjective and depending on the skills of the evaluating person. The measurement of living vegetation cover by using automated image analysis is becoming more common. These methods are faster in processing and can be executed

easily compared to manual cover estimation (Laliberte et al., 2010). But most of these studies focus either on dead (Obade, 2012; Pforte et al., 2012) or living (Behrens and Diepenbrock, 2006; Benett et al., 2000; Booth et al., 2005; Campillo et al., 2008; Purcell, 2000) soil cover. First attempts of using object-based image analysis for soil cover estimation were based on time intensive sampling-based analysis, where the user has to define meaningful samples for each cover type (Laliberte et al., 2010; Luscier et al., 2006; Perez-Cabello et al., 2012). Bauer and Strauss (2014) introduced a method using object-based image analysis for several soil cover types applying a rule based decision tree for classification under certain constrains. These first approaches using automatic segmentation of objects and classification into soil, residue and plants show promising results, but they rely on manual adjustment of parameters.

In this study we introduce a special variant of machine learning technique – namely the entangled forest (Montillo et al., 2011) – in order to improve the robustness and the generalisation of the classification of individual image pixel into soil, living plants and residues. The proposed entangled forest, a variant of random decision forest, classifies individual pixel using simple pixel-wise comparisons to neighbouring pixels with a trained offset. In addition smoothing is achieved by maximum a posteriori features (Wolf et al., 2016), where the decision in a specific node depends on the a posteriori label probability of a neighbouring pixel of the previous decision tree layer.

Our ambition for this study is to introduce a quick, reliable and easy to handle field method for estimation of dead and living soil cover at field scale. To benchmark our results, we compared them to the manual evaluation described in (Hartwig and Laflen, 1978). Furthermore we are focusing on implementation of the entangled forest classification algorithm into web clients and applications for mobiles to directly measure soil cover in field online.

2. Materials and Methods

The proposed algorithm consists of an offline training stage and online classification of images. During training the algorithm tries to create a soil cover model, according to the provided training data, in a supervised learning approach. The training data includes pairs of images and pixel wise masks for the corresponding classes. In the online stage each pixel of any given image is classified using the trained model. Classes are living plants, biofilm, dead residues, stones and soil. The final output of the algorithm includes classified labels for each pixel of the input image and the label distribution for the classes.

The implemented machine learning algorithm needs annotated ground truth data to learn the soil cover model. Therefore, images of different covers of residues and vegetation where taken. These images were taken from approximately 1.4m height horizontally to the ground. With state of the art handheld cameras, this results in a captured soil surface of about 1.5 m². The accuracy of the result depends on the geometric properties of image scale and viewing angle of the camera. In a next step representative areas of these images where cut out and manually pixel wise annotated. The resulting pairs of class- and image- information provide the input information for the learning algorithm.

2.1 Random Forest

For machine leaning we use a special variant of Random Forest proposed in (Montillo et al., 2011). A Random Forest consists of an ensemble of decision trees T. For the supervised training of a decision tree N labelled training images $S = \{x_i, l_i\}_1^N$ must be provided. Each data point s_i needs an image position x_i and the corresponding label l_i. During training of the tree t_i each node selects a test function, samples parameters and selects a split which maximized the information gain. According to the test a subset of the training data is either passed to the left or to the right child node. This splitting process is repeated in each node until a stopping criteria, such as a minimal information gain, a maximum tree depth or a minimum number of data points is fulfilled. During testing the data points are passed from the root node to a leaf node. In each node the test, selected in the training phase is applied and depending on the result the point is passed to the left or the right child node. Finally, the leaf node provides the empirical distribution over classes learned from the training data. The data point is classified by averaging the distributions from the leaves it reaches in each tree. Figure 2 depicts the classification procedure. The following section describes the features learned to test data points in split nodes.

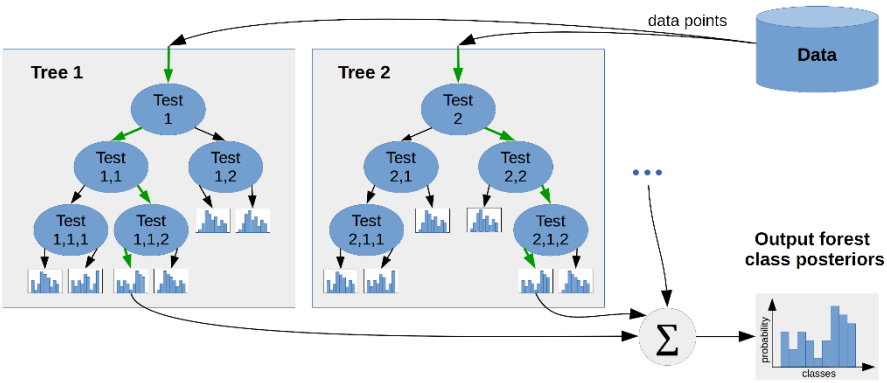

Figure 2: Classical Random Forest, consisting of a set of decision trees which provide a leaf node distribution for the learned classes, e.g. soil, organic and inorganic material.

2.2 Features

Classical classification systems use complex image features, such SIFT (Lowe, 2004) or HOG (Dalal and Triggs, 2005). Instead, we implement simple features which are fast to compute. Our features are designed, either to capture the local neighbourhood with the colour mean, colour variance and simple texture coefficients or to describe the semantic and the appearance context at a larger distance. All features are computed in LAB CIE 94 colour space, where the colour channels are separated from the brightness. Hence, brightness can be weighted differently or for individual features it can be skipped completely to achieve robustness against brightness changes.

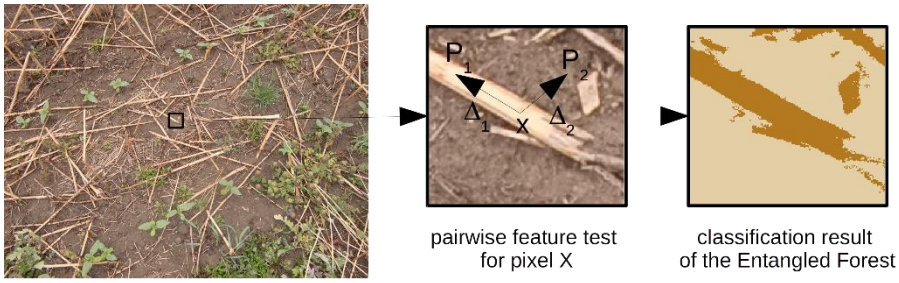

Figure 3: Pairwise color comparison (pairwise feature) for a pixel X

Absolute colour features

$$f_{color,\{A,B\}}(\mathbf{x};\boldsymbol{\Delta}_1,\mathbf{R}_1) = \bar{I}_{\{A,B\}}\big(\mathbf{R}_1(\mathbf{x}+\boldsymbol{\Delta}_1)\big)$$

are computed from the colour channels I_A and I_B. $\boldsymbol{\Delta}$ defines an offset and $\mathbf{R}$ is a region from which the mean colour value is computed. The coordinate $\mathbf{x}$ specifies the pixel to classify.

Pairwise features

$$f_{colorDiff\{L,A,B\}}(\mathbf{x};\boldsymbol{\Delta}_1,\mathbf{R}_1,\boldsymbol{\Delta}_2,\mathbf{R}_2) = \bar{I}_{L,A,B}\big(\mathbf{R}_1(\mathbf{x}+\boldsymbol{\Delta}_1)\big) - \bar{I}_{L,A,B}\big(\mathbf{R}_2(\mathbf{x}+\boldsymbol{\Delta}_2)\big)$$

formalize image gradients from offsets $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$. Image gradients, i.e. the difference of colour values is inherently robust against brightness changes. Thus pairwise features are computed from all three colour channels L, A and B. We compute the difference of mean colour values from the regions R_1 and R_2. These regions can have arbitrary shape (elongated or rectangular, parallel or normal), hence this feature is able to reflect geometric properties in a supporting area. Figure 3 shows an example of a pairwise feature comparison and the final classification result after applying the complete Entangled Forest.

Variance feature

$$f_{var,\{L,A,B\}} = VAR_{\{L,A,B\}}(\mathbf{R}_1(\mathbf{x}+\boldsymbol{\Delta}_1))$$

of a supporting region R_1 is also computed from all three colour channels L, A and B. Optionally, the region can be computed at an offset $\boldsymbol{\Delta}_1$ to the target pixel $\mathbf{x}$ to classify. This feature is beside the Linearness-/Pointness-Feature a simple and fast to compute value which represents the smoothness of a region.

Linearness-/Pointness-Feature

$$f_{lt\{L\}} = \underset{R}{\mathrm{VAR}}\left(\arctan\left(\frac{g_y}{g_x}\right)\right)$$

and

$$f_{g\{L\}} = \underset{R}{\mathrm{MEAN}}\left(\|g_x\| + \|g_y\|\right)$$

describe the variance of the orientation and the mean amplitude of image gradients **g** in a region R centred around a pixel **x**. The idea is that small values of f_{lt} indicate line structures, e.g. coming from a stem of a plant, whereas f_g indicates the reliability of the value.

Entangled Features enable modelling of the context of individual pixel. The idea is that during testing on novel images intermediate label predictions of pixel aids the classification of nearby pixel. Here, we extend the MAPClass entangled features of (Montillo, et al. 2011) to colour constrained maximum a posteriori (Col-MAPClass) features

$$f_{MAPClass\{A,B\}}(\mathbf{x}; \boldsymbol{\Delta}_1, \mathbf{P}_1, C, T_c) = \begin{cases} \underset{c}{\mathrm{argmax}}\, p\left(c; n_{\mathrm{P}_1}\right) = C \wedge \left\|\bar{I}_{\{A,B\}}\left(\mathbf{R}_1(\mathbf{x}+\boldsymbol{\Delta}_1)\right) - T_c\right\| < T & 1 \\ otherwise & 0 \end{cases}$$

Col-MAPClass tests if the maximum a posteriori (MAP) class of a probed pixel $\mathbf{P_1}=\mathbf{x}+\boldsymbol{\Delta_1}$ is equal to a particular class C. $p(c;n_{P1})$ is the posterior class distribution of the node of $\mathbf{P}_1$. The posterior probability can be accessed because the trees are trained in a breadth first manner and an association between the pixel and the tree node is maintained. The colour term in the equation above compares the mean colour in a region R_1 with an offset $\boldsymbol{\Delta_1}$ to a learned colour value T_c. This feature enables smoothing of classification probabilities within one closed framework, which is typically done in a second processing step with a MRF/ CRF (Shotton et al. 2010). In addition the colour term prevents smoothing across object boundaries, e.g. from a particle on the ground to the soil.

2.3 Training and Inference

In principle we apply the standard training procedure described in (Criminisi and Shotton, 2013). The first step before training a decision tree or using it for inference is the data pre-processing. We convert RGB images to the LAB CIE 94 colour space, compute the corresponding grey scale image gradient angle and amplitude maps and create integral images of all these channels. Integral images enable a fast estimation of the features independent of the supporting region size R. To achieve randomized trees without being correlated, we apply bagging of the data and select 50% randomly sampled images. Because the number training labels per class is highly unequally distributed, the random selection of images is guided towards an equal levelling of the labels. In case of small bags, where equal levels of the label distribution is not achieved, we additionally weight each data point with a class weight inversely proportional to the total number of points for the corresponding class in the bag.

For all of our experiments, we limit the number of trees to 10 and stop the expansion at a depth of 25 levels. In addition we stop the growing of the tree in case the information gain of the best split in the particular node drops below 0.2 or less than a minimum number of 5 points are left. In contrast to (Wolf et al., 2016), we do not individually specify the number of randomly selected features, thresholds and parameters. Instead, we perform training in a two-step procedure. First, we select features and parameter equally distributed and in a second learning step we use the prior distribution of the first run to guide the feature selection. This method more often selects complex features with a higher number of parameters and thus achieves a higher learning accuracy. In both runs the number of random sample for each split is set to 4000.

2.4 Dataset and Validation of the Method

In general a well-balanced training set that includes all expected test images provides the best training results. To consider this, our training dataset consists of more than 200 different images and their annotated class information. We selected representative parts with a size of approximately 640x480 pixels from high resolution images (12M pixels). The pixel-wise class information was done by humans using a semi-automatic pre-segmentation method.

For validation of the proposed technique, the photogrammetric grid method, similar to (Hartwig and Laflen, 1978), was applied to the full resolution images. A grid of 160 x 160 pixels is overlaid on a test image, resulting in 432 crossing points for a 4000 x 3000 pixel image. Each of this grid crossing point is manually assigned to a class. The classification results of the algorithm are compared to the ground truth classifications of the human.

Obtaining the ground truth data from images is a very difficult task and often results in strong variations between different evaluators (Hartwig and Laflen, 1978).

3. Results and Discussion

In figure 4 different images of various plants and coverages taken under different lighting conditions are shown. The results of the pixel-wise classification are displayed below each image. In a further step the relative coverage in percent is estimated from those resulting class labels. The visual inspection of this images and the classification indicates that living plant material is classified in an accurate way. The cover classes biofilm and stones are not as accurate as the results of the other classes. This is due to an underrepresentation of stones and biofilm in the training dataset and the ambiguity especially for stones covered with soil.

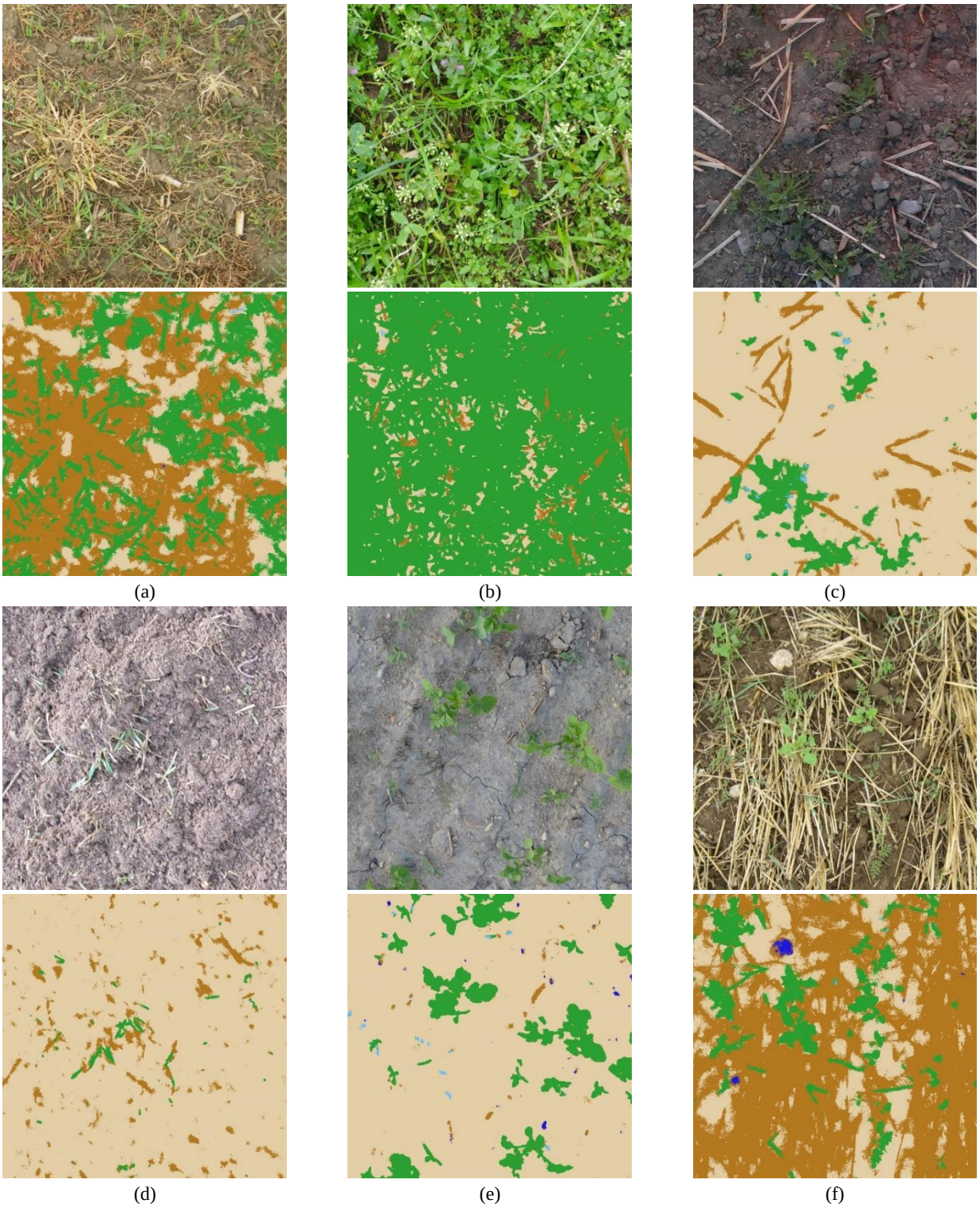

Figure 4: Examples of different soils surface covers (top) and the classification results (bottom).

The validation dataset consists of 180 images and their classifications provided by the photogrammetric grid method, described in Section 2.4. Each of these images has a coverage value between 0 (no occurrence) and 1 (fully covered) for each class based on the grid points. Figure 5 (a-c) show the comparisons between the coverage values from human classification and our Entangled Random Forest (ERF) classification results. In the figures the linear regression over all test samples is visualised. An ideal algorithm would deliver an identity function as regression function of y = x, each human annotation x corresponds to the exact same ERF result y. For the classification of soil we get a regression equation of y = 0.964x + 0.036 with a coefficient of determination R^2 = 0.92, corresponding to an average error of about 4%. Living plant classification results in an error of about 3% and dead residue in 8% error.

The presented method is implemented in C++ using OpenCV and the Eigen library. For simple on field usage of the method, we integrate the algorithm into a web based service including an Android application.

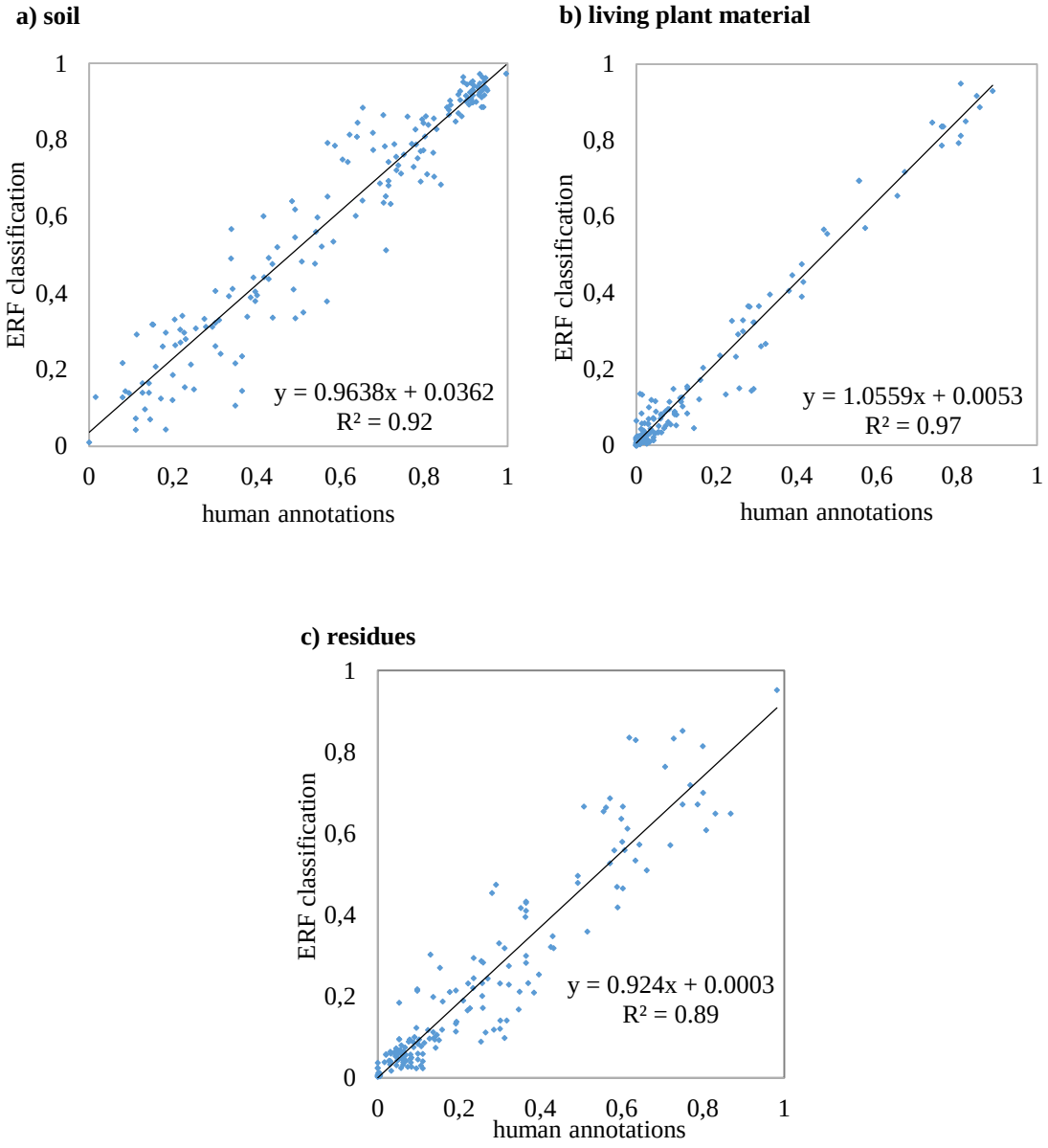

Figure 5: Comparison between reference annotations and entangled random forest results, for (a) soil, (b) living plant and (c) residue cover; dots indicate observation points; continuous lines = regression.

4. Conclusions

The purpose of this publication was to introduce a new analysis method for the estimation of various soil cover types in one single step and to compare this new method to a well-established and widely used method of manual image classification. The proposed entangled forest method was a successful method of image classification. The used training data set seemed to be chosen in a proper way to get best fitting classification algorithms for analysing a broad range of different soil surface conditions. Using the proposed methodology, it is possible to collect soil cover information directly at the field site. Compared to other methods (Bauer and Strauss, 2014) the degree of objectivity is higher and the uncertainties compared to a method of manual evaluation are similar. Due to the learning procedure an objective way of finding good classification is given. A further step forward would be to increase the number of images and cover the whole range of different types of soil surface cover for arable land on different soils.

Acknowledgements

The research leading to this work has received funding from the Lower Austrian government and the Austrian Research Promotion Agency. Furthermore this contribution is partly funded by BiodivERsA/FACCE-JPI – VineDivers Project.

We would like to thank Aigner Franz, Fila Konstanze, Jungwirth Silvia, Karlinger Johann, Karner Matthias, Klipa Vladimir, Koch Adelheid, Leitzinger Maria, Naderi Aziza, Pirker Philipp, Renz Petra and Sonnleitner Claudia for the painstaking work of manual image classification of the training and test data set.

References

Bauer, T., Strauss, P., 2014. *A rule-based image analysis approach for calculating residues and vegetation cover under field conditions*; CATENA, 113, 363-369

Benett, L.T., Judd, T.S., Adams, M.A., 2000. *Close-range vertical photography for measuring changes in perennial grasslands.* J. Range Manage., 53, 634-641

Behrens, T., Diepenbrock, W., 2006. *Using digital image analysis to describe canopies of winter oilseed rape (Brassica napus L.) during vegetative development stages.* J. Agron. Crop Sci., 192, 295-302

Booth, D.T., Cox, S.E., Fifield, C., Phillips, M., Williamson, N., 2005. *Image analysis compared with other methods for measuring ground cover.* Arid Land Res. Manage., 19, 91-100

Campillo, C., Prieto, M.H., Daza, C., Monino, M.J., Garcia, M.I., 2008. *Using digital images to characterize canopy coverage and light interception in a processing tomato crop.* HortScience, 43, 1780-1786

Corak, S.J., Kaspar, T.C., Meek, D.W., 1993. *Evaluating methods for measuring residue cover.* J. Soil Water Conserv., 48, 70-74

Criminisi, A., Shotton J., 2013 (eds). Decision Forests for Computer Vision and Medical Image Analysis. Springer London.

Dalal, N., Triggs, B., 2005. *Histograms of oriented gradients for human detection.* In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE.

Hartwig, R.O., Laflen, J.M., 1978. *A meterstick method for measuring crop residue cover.* J. Soil Water Conserv, 33, 90-91

Laliberte, A.S., Browning, D.M., Herrick, J.E., Gronemeyer, P., 2010, *Hierachival object-based classification of ultra-high-resolution digital mapping camera (DCM) imagery for rangeland mapping and assessment*, J. Spat. Sci. 55, 101-115

Lowe, D. G., 2004. *Distinctive image features from scale-invariant keypoints*. International journal of computer vision, 60(2), 91-110.

Luscier, J.D., Thompson, W.L., Wilson, J.M., Gorham, B.E., Dragut, L.D., 2006. *Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots.* Front. Ecol. Environ., 4, 408-413

Marques, M.J., Bienes, R., Jimenez, L., Perez-Rodriguez, R., 2007. *Effect of vegetal cover on runoff and soil erosion under light intensity events.* Rainfall simulation over USLE plots. Sci. tot. Environ. 378, 161-165

Mohammad, A.G., Adam, M.A., 2010. *The impact of vegetative cover type on runoff and soil erosion under different land uses.* Catena 81, 97-103

Montillo, A., Shotton, J., Winn, J., Iglesias, J.E., Metaxas, D., Criminisi, A., 2011. *Entangled Decision Forests and Their Application for Semantic Segmentation of CT Images.* Information Processing in Medical Imaging, Volume 6801 of the series Lecture Notes in Computer Science, 184-196

Morrison, J.E. Jr., Huang, C., Lightle, D.T., Daughtry, C.S.T., 1993. *Residue Cover measurement techniques.* J. Soil Water Conserv, 48, 479-483

Obade, V. de P., 2012. *Review Article: Remote sensing, surface residue cover and tillage practice.* J. Environ. Prot., 3, 211-217

Perez-Cabello, F., Cerda, A., de la Riva, J., Echeverria, M.T., Garcia-Martin, A., Ibarra, P., Lasanta, T., Montorio, R., Palacios V., 2012. *Micro-scale post-fire surface changes monitored using high spatial resolution photography in a semiarid environment: A useful tool in the study of post fire soil erosion processes.* J. Arid Environ, 7, 88-96

Pforte, F., Wilhelm, B., Hensel, O., 2012. *Evaluation of an online approach for determination of percentage residue cover.* Biosyst. Eng., 112, 121-129

Purcell, L.C., 2000. *Soybean canopy coverage and light interception measurements using digital imagery*. Crop Sci., 40, 834-837

Shotton, J., Winn, J., Rother, C., Criminisi, A., 2010. *TextonBoost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture*, Layout, and Context. In IJCV 2010

Wolf, D., Prankl, J., Vincze, M., 2016. *Enhancing Semantic Segmentation for Robotics: The Power of 3D Entangled Forests.* IEEE Robotics and Automation Letters (Volume:1, Issue: 1), 49-56