Plant Stem Detection and Position Estimation using Machine Vision

Sebastian Haug[1], Peter Biber[1], Andreas Michaels[1], and Jörn Ostermann[2]

[1] Robert Bosch GmbH, Corporate Research, Germany,
sebastian.haug@de.bosch.com
[2] Leibniz Universität Hannover, Germany,
ostermann@tnt.uni-hannover.de

Abstract. In this paper we propose a machine vision based approach to plant stem detection. A sliding window classifier is applied to predict whether each local window in the image displays a stem region or not. The resulting plant stem certainty map is post-processed and non-maxima suppression is used to generate the stem position candidates. Our approach is evaluated using real world image data from a commercial organic carrot farm that was captured with the autonomous field robot Bonirob.

1 Introduction

Robotic precision agriculture activities require detailed information about individual plants. The plant stem position is such an important plant property as it describes the location of a plant in the field. Precise detection of the plant position is a prerequisite for many plant selective tasks like e. g. single plant weed control.

A plant stem detection system is required in the RemoteFarming.1 project [2] where an autonomous field robot Bonirob [11] is built for weed control in organic farming (see Figure 1). The robot is equipped with cameras, a plant classification system [4], this plant stem detection system as well as a manipulator and a mechanical weeding tool [6]. The weeding tool treats single weed plants individually, only when it is applied as close to the stem as possible it is able to treat weed efficiently. The diameter of the weeding tool is 11 mm, i. e. for optimal regulation the position error should be less than 5.5 mm.

We present a machine vision approach to determine the stem of plants in field images. A sliding window is applied and each image patch representing the local neighborhood in the image is classified whether it displays a stem or non-stem region.

Our method addresses the general task where the stem positions of all plants (crop and weed) in the image must be detected. Only multi-spectral images without any additional scene information are used as input data. Prior segmentation into individual plants or leaves which is required by several related approaches is not necessary. This allows easier handling of real word field situations with many plant species and overlap between plants.

Fig. 1: Bonirob field robot during RemoteFarming.1 field tests. The image acquisition and weed control systems are mounted in the covered area below the robot.

2 Related Work

The detection of plants and the estimation of their precise location has been studied from different perspectives.

First, on a coarse level, row detection methods have been applied to detect the positions of row crops in fields [1]. Some work has focused on even more constrained scenarios where the underlying spatial arrangement of the plants is to a large extent known in advance (e. g. when crops grow in a regular grid). These methods are not applicable in our use-case: In the field the spatial distribution of all plants is not known, especially we are also interested in the position of weed plants.

Some work relies on external sensing to (re-)detect the plant position in the field. Nørremark et al. [9] and Sun et al. [12] study the use of RTK GPS to map the position of seeds during sowing to locate plants later when weeding or harvesting is done. Additionally, 3D sensing and processing has been applied to the problem: Nakarmi & Tang [8] use side-view depth images to measure inter plant spacing (stem to stem) in the field and achieve 1.7 cm root mean squared error. However, these methods rely on more complex (3D) or external, expensive sensing (RTK GPS) and are not applicable when only mono camera images are available.

Kiani & Jafari [5] segment field images into individual plants and derive the centroid position of each plant. Their experiments with corn show that the centroid is a good estimate of the plant stem location. Midtiby et al. [7] perform leaf segmentation and then determine plant stem candidates by searching from the leaf tip. The information of multiple leaves is fused to determine an estimate for the plant stem emerging point. These methods rely on a segmentation of the image into plants or leaves which can be very challenging in the field.

3 System Description

The image based plant stem detection system comprises four steps, see Figure 2.

[Data Acquisition & Background Removal] → [Sliding Window Feature Extraction] → [Classification] → [Stem Detection & Position Estimation]

Fig. 2: Processing steps of plant stem detection and position estimation system.

3.1 Data Acquisition and Background Removal

The input consists of multi-spectral images (visible and near-infrared spectrum) captured by a top-down looking camera mounted in the utility bay of the field robot Bonirob. The multi-spectral images are used to determine a binary mask by thresholding in Normalized Differential Vegetation Index (NDVI) image space (see [4] for more details). The threshold is calculated with Otsu's method [10] using the test images, then the determined threshold is fixed and used for all test and train images. The resulting biomass mask which removes soil pixels is applied to the near-infrared image and all stem detection calculations are performed on these masked near-infrared images. Figure 3 shows an example input image (left) and the resulting image after background removal (right).

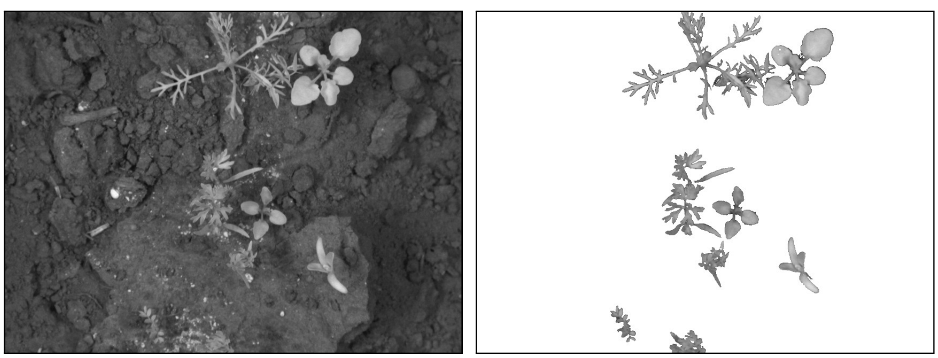


Fig. 3: Background removal: Near-infrared channel of input image (left) and the masked near-infrared image (right).

3.2 Sliding Window Based Feature Extraction

Stem detection and position estimation calculations are performed by applying a sliding window to the image. At every window position where biomass is located at the window center, an image patch is extracted. From each image patch a feature vector is derived which is then passed to a classifier.

First, data must be generated for the classifier training step. This is achieved using images from the field, where plant stem positions were labeled by a human expert. Figure 4 summarizes the training data extraction step: A regular grid is applied to the image, but only where biomass is located and no stem label is in proximity. These positions are used to extract patches that do not display stem regions and are the negative training instances (positions marked with red boxes in Figure 4). At and in the close neighborhood of every manually labeled stem position patches that display the stem region are extracted (positions marked with stars in Figure 4) and these are positive training instances.

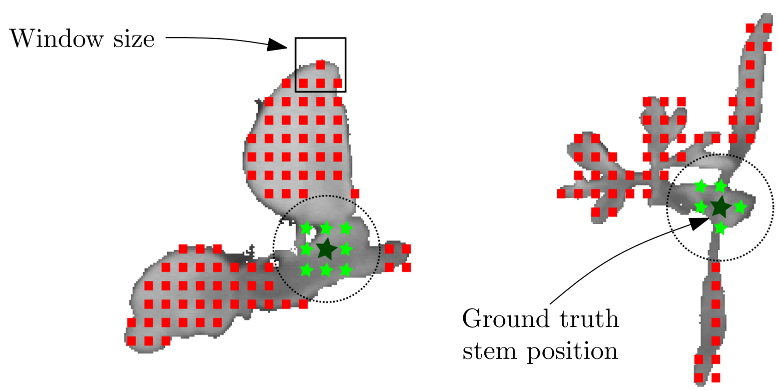


Fig. 4: Sliding window based patch extraction during training phase: Negative (non-stem) patches are extracted at the grid locations marked by the red box. At and around the manually labeled ground truth stem position (dark green star) positive patches are sampled. These positions are marked by green stars.

From these image patches features are extracted. A set of 8 statistical and 4 geometrical features are extracted from 4 scales of each image patch (in total 48 features per patch):

Statistical features Minimum, maximum, range, median, mean, standard deviation, skewness and kurtosis of biomass pixels in image patch.

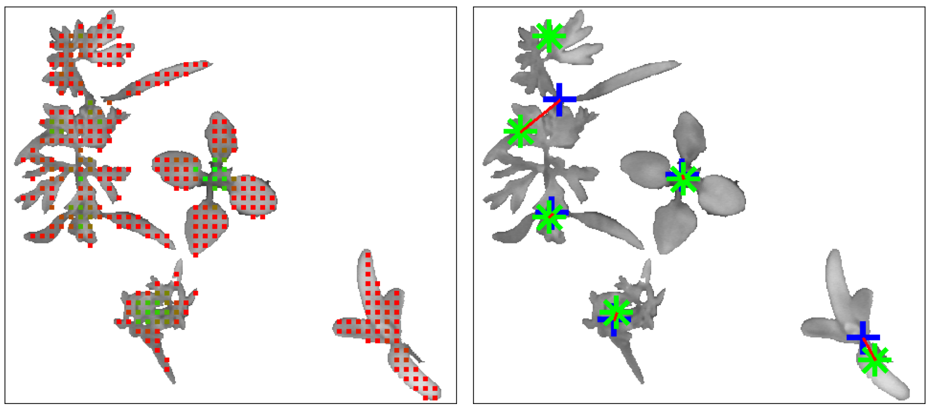
Geometrical features Distance of center of gravity of biomass in image patch from patch center, Mean and Standard Deviation of distance of every biomass pixel from center of gravity in patch and area of biomass in patch.

3.3 Classification

Using the features extracted from the positive and negative image patches a classifier is trained. Here a Random Forest [3] is used because of its state of the art performance and the additional output of certainty scores during prediction.

3.4 Stem Detection and Position Estimation

During application of the system on a new image the sliding window based feature extraction is applied to the whole image. Using the trained classifier stem probability is predicted at every window position. Figure 5a displays the output of the classification step which can be plotted as a plant stem certainty map.



(a) Stem certainty map in color code from red (no stem) to green (stem).

(b) Stem predictions (green *) and ground truth (blue +) stem positions.

Fig. 5: Stem certainty map and step position estimates. Best viewed in color.

The position of plant stems is predicted using the plant stem certainty map. First, the certainty map is slightly filtered with a Gaussian filter to smooth the individual predictions. Then, non-maxima suppression is used to generate discrete stem position estimates. Figure 5b displays the resulting stem predictions (green *) and for visual comparison ground truth stem positions (blue +).

4 Experiments and Results

The stem detection accuracy is evaluated using an image dataset captured in an organic carrot farm using the autonomous field robot Bonirob. The images display carrot and weed plants and here we study the detection of the stem of all plants in the images. The dataset is split and 45 images are used to extract training data, 25 images are used to evaluate the performance. The ground resolution of the images is 7 px per mm. The window size is set to 30 px, during detection the window stride was set to 10 px. The scales of the window size were set to 1.0, 0.75, 0.5 and 0.25 rounded to full pixels. The Random Forest was trained with 25 trees, for all other parameters the default is chosen.

Visual analysis of the stem prediction indicates good performance. Figure 6 shows four output images with annotated ground truth for comparison.

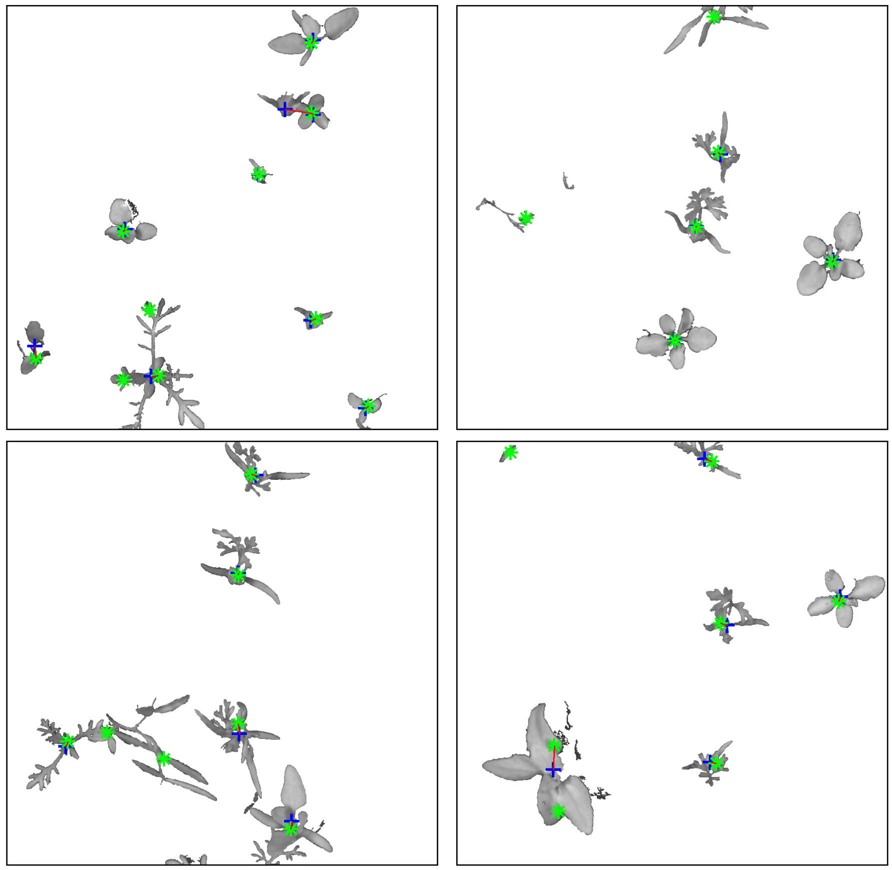


Fig. 6: Image with plant stem detections (green *) and ground truth stem positions (blue +). The deviation from the closest predicted stem center is marked by a red line. Best viewed in color.

The precision of the stem predictions can also be measured by evaluating the distance between predictions and the annotated ground truth stem positions.

In the 25 test images our system estimates 178 stem positions while the human annotator marked 138 stem positions. When setting the threshold for a correct stem detection to 5.5 mm (half the size of the diameter of the weeding tool) we detect 80.4 % of the ground truth plant stems. The mean position error of the detected stems is 1.88 mm $\pm$ 1.26 mm standard deviation.

The system achieves a stem detection precision that is well suited for weed control with the current mechanical tool. However, some of the labeled stems are missed. Misses happens e. g. when plants grow close together and the non-maxima suppression rejects one of the stems. Additional detections happen e.g. for chamomile and carrot plants. Both species have pinnate leaves and some leaf areas look like plant stem regions. Also, overlap between plants can create patches that look similar to stems in the top-down view. To overcome these issues and further improve performance a larger dataset or plant species specific stem classifiers might help. This needs more research and is future work.

5 Conclusions

The objective of this paper is the detection and position estimation of plants in images. A sliding window based image patch extraction and classification approach is combined with non-maximum suppression to determine plant stem positions. The system is evaluated using field data that was captured with the autonomous field robot Bonirob and results indicate good performance. It achieves a detection rate of 80.4 % of the ground truth stem positions with a mean position error of 1.88 mm.

In the future a larger training and test dataset will be collected with the field robot Bonirob. Furthermore, the system could make use of plant classification information to further improve results by training species specific stem detection classifiers.

Acknowledgments

The project RemoteFarming.1 is partially funded by the German Federal Ministry of Food, Agriculture and Consumer Protection (BMELV).

References

1. Åstrand, B., Baerveldt, A.J.: An agricultural mobile robot with vision-based perception for mechanical weed control. Autonomous robots 13(1), 21–35 (2002)
2. Bangert, W., Kielhorn, A., Rahe, F., Albert, A., Biber, P., Grzonka, S., Haug, S., Michaels, A., Mentrup, D., Hänsel, M., Kinski, D., Möller, K., Ruckelshausen, A., Scholz, C., Sellmann, F., et al.: Field-robot-based agriculture: RemoteFarming.1 and BoniRob-Apps. VDI Agricultural Engineering 2013 pp. 439–446 (2013)
3. Breiman, L.: Random forests. Machine learning 45(1), 5–32 (2001)
4. Haug, S., Michaels, A., Biber, P., Ostermann, J.: Plant classification system for crop / weed discrimination without segmentation. In: Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on. IEEE (2014)
5. Kiani, S., Jafari, A.: Crop detection and positioning in the field using discriminant analysis and neural networks based on shape features. Journal of Agricultural Science and Technology 14(4), 755–765 (2012)
6. Michaels, A., Albert, A., Baumann, M., Weiss, U., Biber, P., Kielhorn, A., Trautz, D.: Approach towards robotic mechanical weed regulation in organic farming. Autonomous Mobile Systems 2012 pp. 173–181 (2012)
7. Midtiby, H.S., Giselsson, T.M., Jørgensen, R.N.: Estimating the plant stem emerging points (pseps) of sugar beets at early growth stages. Biosystems Engineering 111(1), 83–90 (2012)

8. Nakarmi, A., Tang, L.: Automatic inter-plant spacing sensing at early growth stages using a 3d vision sensor. Computers and Electronics in Agriculture 82, 23–31 (2012)
9. Nørremark, M., Griepentrog, H.W., Nielsen, J., Søgaard, H.T.: The development and assessment of the accuracy of an autonomous gps-based system for intra-row mechanical weed control in row crops. Biosystems Engineering 101(4), 396–410 (2008)
10. Otsu, N.: A threshold selection method from gray-level histograms. Automatica 11(285-296), 23–27 (1975)
11. Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., et al.: Bonirob – an autonomous field robot platform for individual plant phenotyping. Precision agriculture 9, 841 (2009)
12. Sun, H., Slaughter, D., Ruiz, M.P., Gliever, C., Upadhyaya, S., Smith, R.: Rtk gps mapping of transplanted row crops. Computers and electronics in agriculture 71(1), 32–37 (2010)