International Conference of Agricultural Engineering
AgEng 2014 Zurich 6–10 July

Ref: C0259

Discrimination of plants and weed by multi-sensor fusion on an agricultural robot

Christian Frese, Johannes Meyer and Christian Frey, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany

Abstract

This contribution presents an approach to automatically classify crop plants and weed based on multi-sensor information with the aim of mechanically removing the weed by an agricultural robot. First, a possible sensor setup and its calibration are outlined. As the robot moves forward, the sensor data can be aggregated into a 3D model. To increase model quality, the roll and pitch angles of the robot have to be estimated and compensated. Then, different features are computed from the sensor data and the discriminative power of the features is evaluated. The feature vector is input to a support vector machine classifier. The considered classes are crop plant, weed, and soil. The result is a 3D representation of plants and weed which can be used for automatic weed removal.
As an example application, tree nurseries are considered, especially the growth of boxwood trees. Classification and mapping results on real data acquired by a robot are reported.

Keywords: robotics, multi-sensor fusion, weed control

1 Introduction

Weed control is in many cases a cost-intensive operation, requiring either manual work to mechanically remove the weed or herbicides which may be harmful also to the crop and to the environment in general. Agricultural robotics has the potential to make mechanical weed control more cost-effective, thus allowing a more frequent application. Thereby herbicides become dispensable to a large extent, which is often beneficial for both crop and environment.

As an example application, tree nurseries are considered, especially the growth of boxwood trees (*buxus sempervirens*). Boxwood grows very slowly and requires repeated weed control during several years so that the possible savings of automation are notable. Typically, the boxwood plants are arranged in rows. Our outdoor robot IOSB.amp O1 is able to drive along such rows. It observes the plants from above and records sensor data of real boxwood trees (Figure 1).

In the literature, vision-based methods are typically used for automatic weed detection (Slaughter et al., 2008, Bossu et al., 2009). Besides segmentation approaches based on color feature thresholds, artificial neural networks are commonly used as classifiers (Tang et al., 2003). Plant classification based solely on 3D lidar data has also been studied (Weiss et al., 2010).

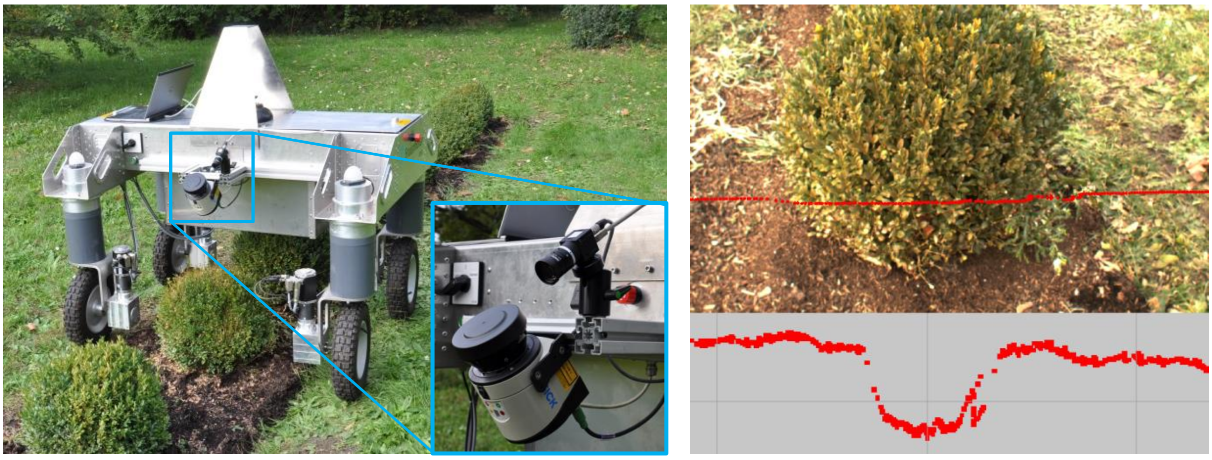

Figure 1: Robot and sensor setup. Left: Mobile robot IOSB.amp O1 driving along a row of boxwood plants, with a detail showing the sensor system consisting of a lidar sensor and a color camera. Right: Measured range data plotted in the sensor coordinate system and projected into the camera image.

This contribution proposes an approach to automatically classify crop plants and weed based on multi-sensor information. By multi-sensor fusion, both range and color information is available for each data point. This allows the computation of a highly distinctive feature vector which is then used to classify each point as crop, weed, or soil. The classification is performed by a support vector machine trained with hand-labeled data. Furthermore, the range data can be aggregated into a 3D model of the plantation which is an important prerequisite for successfully removing the weed without harming crop plants. The mechanical manipulation system which is developed by our project partners is outside the scope of this paper.

2 Materials and methods

2.1 Sensor setup and calibration

The current sensor setup consists of a 2D laser range finder (line scan lidar) and a color camera. The lidar sensor measures range values within an inclined plane.

The intrinsic and extrinsic camera parameters are calibrated using a standard planar calibration pattern. Then, the acquired lidar data of the calibration object is registered towards the plane estimated by the extrinsic camera calibration (Zhang & Pless, 2004).

The calibration of the sensor system allows to find for each 3D point measured by the laser range finder the corresponding pixel in the camera image (Figure 1).

2.2 Mapping

As the robot moves forward, the 2D range data can be aggregated into a 3D point cloud of the plantation. To this end, the current position and orientation of the lidar sensor has to be known. The (x, y) position within the ground plane and the yaw angle are obtained from the odometry data of the robot. The nominal height (z coordinate), roll and pitch parameters of the sensor are constant. However, due to the roughness of the terrain, the real height, roll, and pitch values of the robot will be time-variant. To increase the quality of the 3D model, these variations have to be estimated and compensated. This can be accomplished either by additional sensors such as inertial measurement units or by inferring the pose parameters from the range data using model assumptions.

For the latter approach, it is assumed that the terrain is globally planar within the observed area. The ground plane and the scan plane of the lidar intersect in a line. The position of this line in the sensor coordinate system depends on the current roll, pitch, and height values. If the roll angle is zero, the ground line appears to be perpendicular to the forward direction of the robot. If the robot is rolling, the line appears to be tilted. Depending on the pitch angle, the measured ground line is closer to the sensor or farther from the sensor. The equation of the ground line is estimated from the range data by means of the RANSAC algorithm, which discards local outliers not belonging to the ground plane (Fischler & Bolles, 1981). Then, roll and pitch angle of the sensor are reconstructed from the line equation parameters using elementary, but somewhat lengthy trigonometric equations.

The z position of the sensor can be estimated based on the assumption that roll and pitch variations even out over time. The correction of the z position is beneficial when the wheels of the robot do not have the same ground level as the observed plane, e.g., when the robot drives within the furrows of the field.

2.3 Feature extraction

As the sensor system is calibrated, both range and color information is available for each data point under consideration. Based on the sensor data, a feature vector characterizing the data point is computed. The individual features can be grouped into range, color, and texture features.

Range features such as the height (z coordinate) of the point and the standard deviation of the height capture the typical size and surface structure of different plants.

Color features are particular helpful for discriminating plants from soil. However, as the illumination is uncontrolled daylight, it is important to compute features which are invariant under intensity and illumination variations. Therefore, the color channels are preprocessed using histogram matching, and the employed color features are normalized by the intensity value of the pixel.

Texture features characterize homogeneity, contrast, and anisotropy of local structures in the gray-scale image. Size, shape, and arrangement of the leaves generate a characteristic pattern which helps to discriminate boxwood from other plants, e.g., grass having elongated structures. For each pixel corresponding to a measured 3D point, a neighborhood window in the image is considered in order to obtain information on the local image structure. The window size is a critical parameter which should be optimized for the considered task. Larger windows allow the computation of more reliable features, but impair the spatial resolution of the classification. The following texture features have been evaluated: mean and variance of gray value within the window, histograms of oriented gradients, Fourier histograms (Ursani et al., 2008), Laws' texture energy (Laws, 1980), and local binary patterns (Ojala et al., 2002).

2.4 Classification

The feature vectors of the individual data points are classified by a support vector machine (Chang & Lin, 2011) into one of the classes boxwood, weed, and soil. The weed class encompasses all vegetation other than boxwood plants.

As a basic support vector machine can handle only two classes, the one-against-one voting approach is used for this three-class problem. Radial basis functions are chosen as a kernel to transform the feature vectors.

The classifier is trained on real data in which every point has been manually annotated with the correct class label. Points which could not be assigned unambiguously to one of the three classes have been omitted in the training procedure. These points can originate from objects not belonging to any of the three classes, e.g., withered foliage lying on the ground, or from mismatches between lidar and camera data due to occlusions caused by foreground objects. The data has been recorded during several months in a plantation of spherical boxwood plants having a diameter of about 40-50cm. Training and test data sets are strictly separated: they consist of different plants and have been acquired on different days and thus under different illumination conditions.

Extensive grid search optimizations have been performed in order to select suitable parameters for classification and feature extraction. The classifier has been trained using different parameter settings, and the classification success has been evaluated on the labeled test set each time. This procedure has been applied to optimize the parameters of the support vector machine, of the radial basis function kernel, and of different features, especially the window sizes for the texture features.

Additionally, the classification performance when using feature vectors composed of varying feature sets has been evaluated in order to assess the distinctive power of certain groups of features.

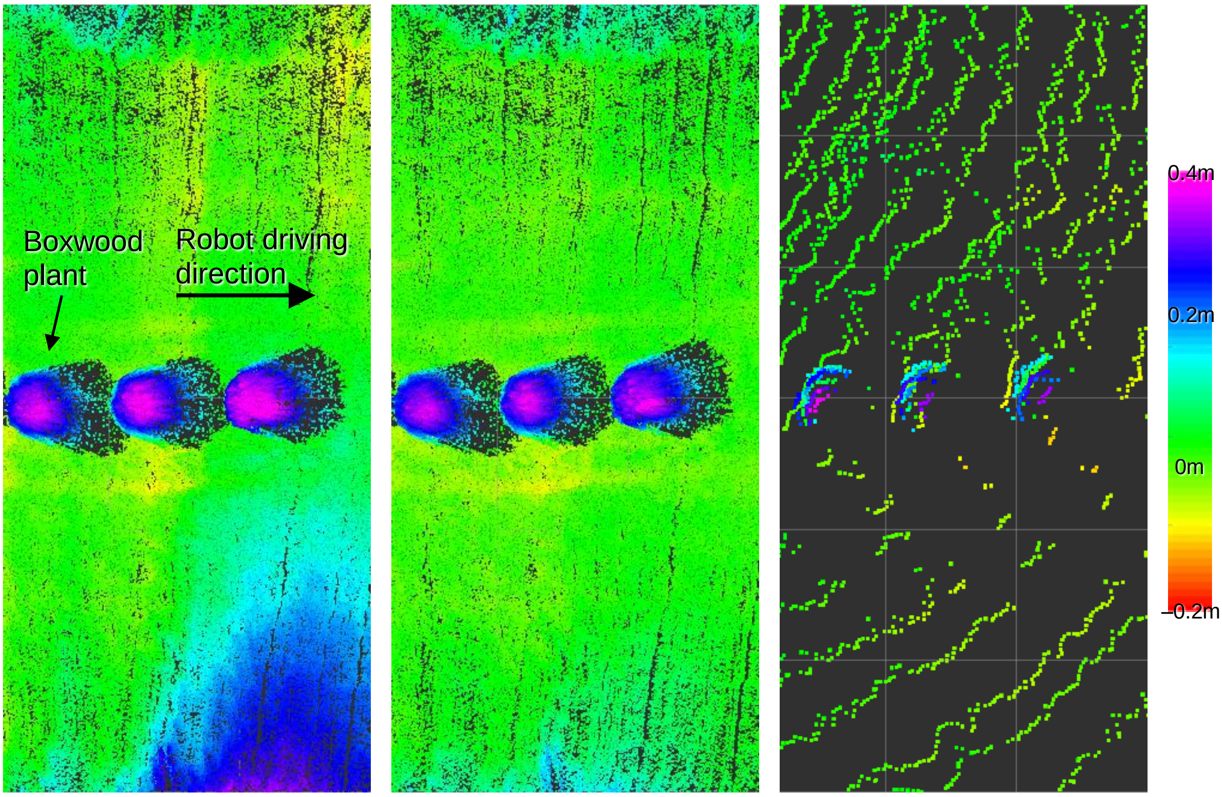

Figure 2: 3D mapping of the plant cultivation. Left: Top view on the uncorrected point cloud computed from the linescan lidar data. Center: Corrected point cloud aggregated using the estimated roll, pitch and z values. Right: Single scan of a Velodyne 3D lidar. The height (z value) is color-encoded as indicated.

3 Results

3.1 Mapping

Figure 2 shows a detail of the 3D point cloud acquired by the robot while driving along the row of boxwood plants. The spherical shape of the boxwood trees is clearly visible in the color-encoded view from above onto the point cloud. In the uncorrected point cloud shown on the left, an apparent “hill” can be seen in blue color. This is however a false estimation caused by the robot heavily rolling and pitching due to a pot-hole in the ground which can be seen in yellow next to the second boxwood plant. Such artifacts can be avoided when estimating the sensor pose from the lidar data as proposed in Section 2.2. Then, the respective area appears to be more or less flat and the map is generally smoother. Ground truth data of the scene has been acquired by a Velodyne 3D lidar sensor in a single scan. By comparing the point clouds in Figure 2, it can be verified that the “hill” occurring in the uncorrected data is indeed an artifact.

Similar results are obtained by a quantitative analysis. The deviation of the points outside the row of plants from an idealized plane model is more than halved when using the corrected pose during the aggregation of the point cloud. It reaches about the same value as on the Velodyne point cloud. Additionally, the height mismatch of the computed point cloud to the Velodyne data is significantly reduced compared to the uncorrected point cloud.

3.2 Classification

Feature vectors on the margin of the classes in the feature space are selected as support vectors by the training algorithm. Figure 3 shows the corresponding image pixels for support vectors of the two-class decision problems boxwood against weed and soil against weed, respectively. These examples indicate that the support vectors indeed represent the limiting cases in the training data. In total, about 6.4% of the training data points have been selected as support vectors. This percentage represents an estimate for the expected classification error (Cristianini & Shawe-Taylor, 2006).

On the different test sets, an average classification success of 95.41% has been obtained. Furthermore, only 0.39% of the boxwood points have been incorrectly classified to be weed. Hence the classification of the boxwood plants is very reliable, whereas weed and soil are mixed up sometimes, especially if they are in close vicinity. But this is uncritical as the mechanical manipulator will always treat a region a little larger than the detected weed area. Figure 4 illustrates typical classification results obtained by the proposed method.

An iterative feature selection procedure has been applied in order to assess the performance of different groups of features with regard to the classification task. The procedure can be illustrated by means of the excerpt of the results shown in Figure 5. First, the classification

Figure 3: The data points marked in red have been selected as support vectors for the discrimination of boxwood against weed (top row) and soil against weed (bottom row).

success rate for each group of features is evaluated individually. The most successful group is selected, in this case the range features. Then, the union of the selected feature set with each of the remaining groups is evaluated on the same test set. In Figure 5, the addition of the color features showed the largest increase in classification performance. In the third step, the statistical texture features (mean and variance of gray value within a local window) have been chosen in addition to the previous set of range and color features. Among the advanced texture features, local binary patterns performed best, followed by Fourier histograms, Laws' texture energy, and histograms of oriented gradients.

Figure 6 shows the benefits of the multi-sensor approach using both range and image features by plotting the classification success for different groups of features.

4 Conclusions

This contribution has presented a multi-sensor approach to distinguish crop plants from weed. The acquired range and image data allows both weed detection and 3D mapping of the plantation. Weed detection has been performed by classifying each data point based on range, color, and texture features. The classifier has achieved good success rates on real test data.

The resulting data can be used for mechanical weed control as well as for other applications such as cutting the boxwood trees into the desired (spherical) shape or rating the health of the plants.

Further work will address the optimization of the computational performance and the evaluation on different data sets encompassing, e.g., boxwood plants of different sizes. Additionally, the integration of range data and classification results into a model based on geometric primitives will be investigated.

Figure 4: Examples of classified data points projected into the camera image.

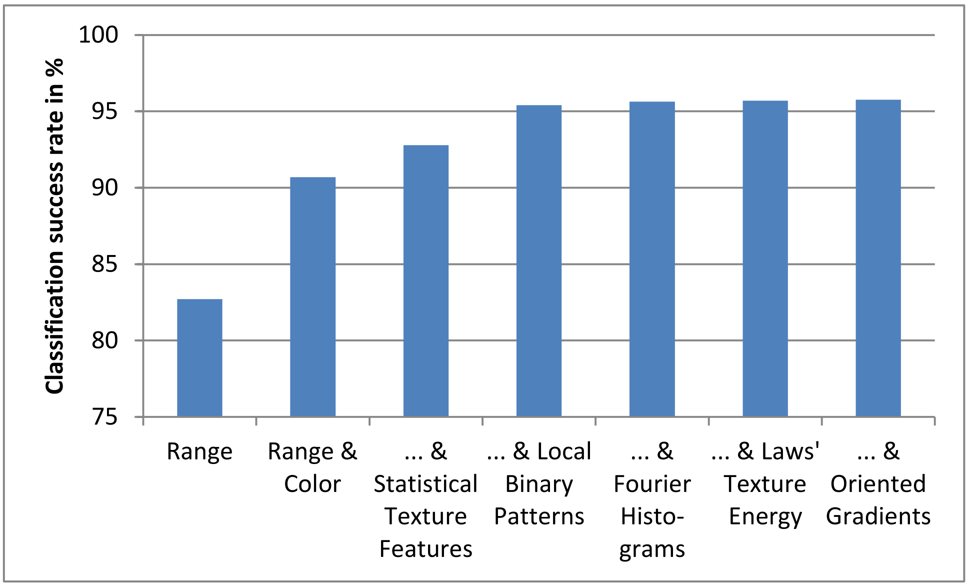

Figure 5: Results of the iterative feature selection procedure. Each feature set encompasses all features mentioned to the left of it, as suggested by the notation "… &".

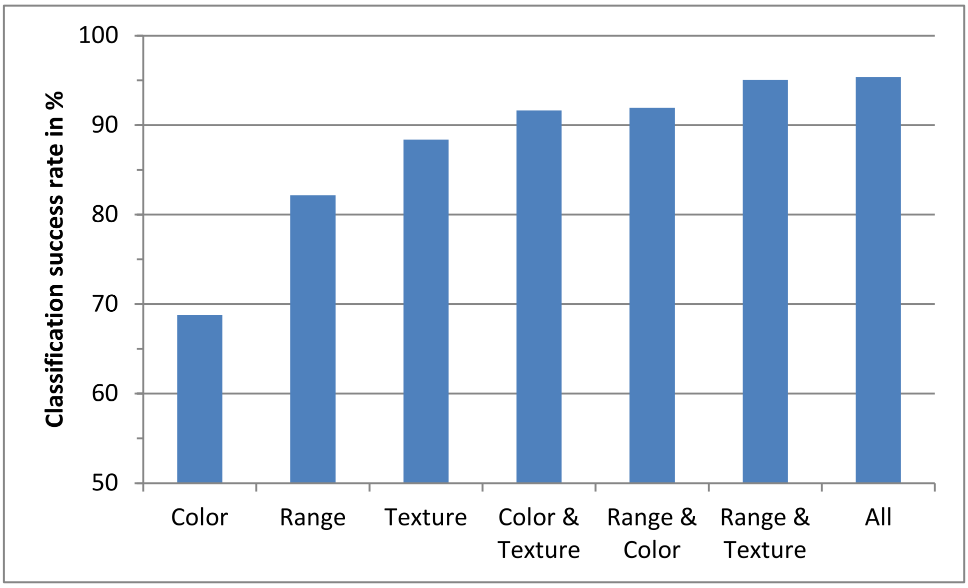

Figure 6: Classification success when using different groups of features. While a good performance can be achieved using only image features (color & texture), the addition of the range features further improves the classification success.

5 Acknowledgements

This work has been supported by the German Federal Ministry of Education and Research (BMBF) within the project "AgriApps" under grant No. 01IM12002C.

6 References

Bossu, J., Gée, Ch., Jones, G., & Truchetet, F. (2009). Wavelet transform to discriminate between crop and weed in perspective agronomic images. *Computers and Electronics in Agriculture, 65*, 133-143.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2 (3)*.

Cristianini, N., & Shawe-Taylor, J. (2006). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.

Fischler, M., & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24 (6)*, 381-395.

Laws, K. (1980). Rapid texture identification. *Proceedings of the Seminar on Image Processing for Missile Guidance,* 376-380.

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24 (7)*, 971-987.

Slaughter, D. C., Giles, D. K., & Downey, D. (2008). Autonomous robotic weed control systems: a review. *Computers and Electronics in Agriculture*, *61 (1)*, 63-78.

Tang, L., Tian, L. F., & Steward, B. L. (2003). Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network. *Transactions of the ASAE*, *46 (4)*, 1247-1254.

Ursani, A. A., Kpalma, K., & Ronsin, J. (2008): Texture features based on local Fourier histogram: self-compensation against rotation. *Journal of Electronic Imaging*, *17 (3)*.

Weiss, U., Biber P., Laible, S., Bohlmann, K., & Zell, A. (2010): Plant species classification using a 3D LIDAR sensor and machine learning. *9th International Conference on Machine Learning and Applications*, 339-345.

Zhang, Q., & Pless, R. (2004). Extrinsic calibration of a camera and laser range finder (improves camera calibration). *International Conference on Intelligent Robots and Systems (IROS)*, 3:2301-2306.