# Probabilistic Near Infrared and Depth Based Crop Line Identification

Georg Halmetschlager<sup>1</sup>, Johann Prankl<sup>1,2</sup>, and Markus Vincze<sup>1</sup>

<sup>1</sup> Automation and Control Institute, Vienna University of Technology, Austria <sup>2</sup> Josephinum Research, Wieselburg, Austria {gh,jp,mv}@acin.tuwien.ac.at

**Abstract.** We present a method for near infrared and depth data based plant soil segmentation and two alternative, probabilistic, model based methods for a subsequent real-time capable crop line identification, respectively tracking. The segmentation and both, the RANSAC based crop line identification and the cascaded particle filter crop line identification and tracking needs neither a learning phase, nor any other form of a-priori knowledge of the row structure or the observed plants. The methods are tested with different datasets of real world robotic applications, result in high detection rates, and show the adaptability of the particle filter for completely different and also changing row structures.

**Keywords:** machine vision, probabilistic crop line estimation, NIRD segmentation, RANSAC, particle filter

## 1 Introduction

Most autonomous agricultural robot systems solve the localization problem by utilizing either a high accurate and absolute, but also expensive RTK-GPS based positioning system or a local machine vision (MV) based system that determines a relative position. Besides the superior accuracy of the RTK-GPS system, the most significant disadvantages of such systems are that the negotiable track has to be defined in prior that they are expensive and that the needed external infrastructure generates often significant running costs.

In opposite, MV systems do not need any additional infrastructure or external information and by using the local available environmental structures they offer a higher flexibility. Since no external infrastructure is needed, no additional running costs are generated, what enables cheaper navigation solutions and facilitates the development of affordable autonomous agricultural machines that are also suitable for small and medium size agricultural businesses.

In the last decades a number of MV based crop line segmentation and detection algorithms have been developed. Several segmentation methods [1–4] consider spectral information (green-channel, NIR, RGB) without taking 3D information into consideration. Otherwise, pure 3D segmentation methods omit often the spectral information [5, 6]. Besides [3, 5] the crop line detection or row guidance systems are realized with single 2D image analyzing methods as the Hough transform [1], linear regression, stripe analysis, and other methods [7]. All methods need some form of prior information as the row distance, camera orientation, other algorithm specific parameters or deal with constraints concerning the line orientation.

Kise et al. [5] propose a stereo vision based row detection system, which uses a cross-correlation function in combination with an elevation map, to identify a navigation point and to calculate the desired steering angle.

In [6] Weiss and Biber develop an algorithm to segment plants, soil, and other structures in 3D point clouds based on the height differences and a RANSAC [8] ground plane estimation.

Besides the different segmentation methods, most of the algorithms need prior information or constraints concerning the crop row structure or crop line orientation and have to be reconfigured to adapt them to seasonal changes or different row structures. Our approach tries to advance the crop row detection by replacing the hard constraints with more flexible probabilistic methods that base on geometric models of the expected row structures. Our aim is to make specific initial information obsolete, independent from seasonal changes, crop line orientations, and row distances and to determine the most probable parameter configuration for the models. hh

# 2 Approach

The approach can be separated into two parts: The Near Infrared Depth (NIRD) segmentation which segments the plants, respectively the soil and prepares the data for the second part, the two alternative subsequent probabilistic crop line identifications.

#### 2.1 NIRD Plant Soil Segmentation

The NIRD segmentation utilizes the advantages of NIR imaging and improves the results with the plants' height information, extracted from 3D data. While pure height based segmentation tend to fail in early growing states, spectral information can be utilized as soon as the plants are visible. The NIRD segmentation can be divided into three steps: NIR segmentation, height segmentation, and data fusion. The realized height segmentation is similar to the method approached in [6], but adds an additional step that projects the 3D points to the NIR image. The generated height map and the NIR image are normalized and fused together by a pixel-wise multiplication. In a next step, the results are virtually projected to the ground plane. The reprojection guarantees that the line equations are rated and verified with the correct dataset by the subsequent crop line detection. Hence, it avoids wrong line ratings due to the plant's not considered height, and perspective geometry (cf. Fig. 1).

The result of the reprojected data fusion is binarized with an adaptive thresholding method and ends in the final segmentation.



**Fig. 1.** Segmented image before (left) and after (right) it is projected to the estimated ground plane (x-y plane). Red and cyan: Estimated 2D lines which are within the ground plane (f(x, y)). Red: determined with the not-projected image data. Cyan: determined with the projected image data. If the red line is projected into the real world it will not represent the correct position of the crop line. The error is caused by the height of the observed plants and the perspective. Since the cyan line is directly estimated in the x-y plane it is unbiased (right).

#### 2.2 RANSAC Based Crop Line Detection

In a nutshell, the RANSAC algorithm fits a set of models into a set of points and finally results in a parametrization for the best fitting model [8]. Within the context of this work, it is used to search lines within the NIRD image data. Based on the assumption that the line search and the determination of their parallelism can be separated, a single line model in normal form is used that parametrizes a line with the vector  $\mathbf{r} = [r, \theta]$ . Since the 3D and color information of a point is projected onto the ground plane by the NIRD segmentation, only the points that belong to a given line within the ground plane in the 3D space, will belong to a line within the 2D image. Since the 3D information is available for each pixel a re-projection of the line to the 3D space is possible and the output of the NIRD segmentation can be used for the crop line detection.

The count of the data points significantly influences the computational costs, hence the dataset is reduced by a simple thinning method that shrinks areas in the binarized image to pixel chains (skeletons). In opposite to a center of mass data reduction method, the shrinking keeps the size and directionality information of the areas and is also suitable for elongated crop rows that appear as one connected area within the binarized segmented image.

In a first step a RANSAC line search is performed which bases on [8]. Usually the RANSAC algorithm finishes with the parametrization that results in the highest inlier ratio, but instead of rating the lines individually, the parallelism of the different lines is determined in a subsequent clustering and rating process that considers the orientation and offset values of the line in the 3D space. The clustering searches for the best lines, selects all lines within a given Euclidean distance, and ends with several bundles of parallel lines. A subsequent filter reduces lines within a bundle that offer a similar r to a single line (cf. Fig. 2). If the group consists of more than one line, the row distance can be determined.



Fig. 2. Extracted lines (left), group of parallel lines (middle), final parallel lines (right).

#### 2.3 Particle Filter Based Crop Line Detection and Tracking

In a nutshell particle filters sample a k-dimensional parameter space with M hypotheses and estimate the probability of parameter configurations for a model that is verified against a given dataset [9]. The functionality can be summarized in three steps: The prediction step that generates a new set under the consideration of the input u(t), the determination of the importance weight  $w_m(t)$ , and finally the re-sampling procedure.

For the crop line detection a geometric model of a parallel line pattern is selected that can be parametrized with the three parameters  $[r, \theta, d]$ . r and  $\theta$ represent the offset and the angle of a line's normal equation within the repetitive line pattern and d the distance between the lines. The likelihood of a line pattern to represent the best parametrization  $w_m(t)$  is directly estimated after a projection of a hypothesis from the state space to the NIRD image (cf. Fig. 3), with the plant  $(c_p)$  and soil pixel count  $(c_s)$  along the lines of the patterns.

Since it is not trivial to find a consistent probabilistic description  $w_m(t)$  that determines the probability of all three parameters at once and with the assumption that the parametrization can be separated into two independent problems, the parametrization of r and  $\theta$  and the parametrization of d can be realized with two different, cascaded particle filters. The first particle filter estimates r and  $\theta$ , while the cascaded filter estimates d based on the results of the first particle filter.

Based on an evaluation of different importance weight functions, (1) is selected for the determination of  $\theta$  and r. (2) is selected for the determination of the third parameter d.

$$w_{m,1}(t) = \frac{c_p}{c_p + c_s} \tag{1}$$

$$w_{m,2}(t) = c_p \tag{2}$$

Hence, the constraint  $d_m > \varepsilon_{max}$  has to be added to the second state space, where  $d_m$  represents the minimal value of d and  $\varepsilon_{max}$  a constant > 0. If the lower boundary is not considered, the row distance degenerates to zero.

After the determination of the likelihood for each hypothesis, a low variance sampling redraws a set of particles (cf. [9]). The influence of the robot's movement (u(t)) on the hypotheses in the parameter space has to be considered for

the prediction step and is modeled with,

$$\begin{bmatrix} r_{bl,i}(t)\\ \theta_{bl,i}(t) \end{bmatrix} = \begin{bmatrix} -x_{bl}(t)\cos\left(\theta_{w,i}\right) - y_{bl}(t)\sin\left(\theta_{w,i}\right) + r_{w,i}\sin\left(\theta_{w,i}\right)\\ \theta_{w,i} - \alpha_{z,bl}(t) \end{bmatrix} \quad , \quad (3)$$

where  $[x_{bl}(t), y_{bl}(t), \alpha_{bl}(t)]$  describe the robot's pose and  $[r_{w,i}, \theta_{w,i}]$  a static line in the world coordinate system.  $[r_{bl,i}, \theta_{bl,i}]$  represent the polar coordinates of a line in the robot coordinate system. With the time derivative of (3), the velocity vectors which correspond to the different hypotheses for a movement of the robot along its x-direction can be determined (cf. Fig. 3).



**Fig. 3.** Projection of a hypothesis into the camera image (left) and movement of the hypotheses depicted with a vectorfield, if the robot moves along its x-axes (right).  $r_h$ : line equation vector of a horizontal line,  $r_p$ : line equation vector of a parallel line,  $v_r$ : velocity of the robot, vectorfield: velocity vectors of the hypotheses (points) in the state space for the given  $v_r$ .

Since the robot will move most of the time parallel to the detected rows and the parallel line expressions stay static within the state space (cf. Fig. 3), it is assumed that the prediction can be modeled with Gaussian noise.

## 3 Experiments and Results

First, a pure NIR segmentation is compared with the final NIRD segmentation to highlight the improvements that can be achieved if the height information is added to the segmentation step. Second, the applicability of the two approached crop detection algorithms and their performances for different plant and row structures are evaluated. All experiments were performed with seven real in-field datasets which offer different row organizations, plant structures, plant sizes, and row irregularities.

The datasets were recorded with a camera carrying vehicle (CCV) that moved parallel to the rows. For the 3D data gathering the stereo vision system Bumblebee2 was used. The NIR images were captured with an industrial camera (DBK 31AF03) equipped with an additional optical 850nm bandpass filter. The algorithms were tested on a LENOVO W530 notebook equipped with an Intel Core i7-3630QM CPU @ 2.40GHz  $\times$  8, 7.4 GiB RAM and a NVIDIA Quadro K1000M graphic card. Each experiment consists of ten repeated runs for each dataset and the accumulated results are shown in Fig. 5 and Fig. 6.

## 3.1 NIR vs. NIRD Segmentation

Figure 4 gives a comparison between a NIR and a NIRD segmentation. In opposite to the NIR segmentation, the NIRD approach filters out plants that are close to the estimated ground plane such as dead plants or smaller weeds that would influence the crop row detection and result in an overall better row segmentation. The NIRD segmentation requires for the given setup 15ms, excluding the 3D data preparation which consists of the point cloud generation and a subsequent estimation of the ground plane.



Fig. 4. Comparison of a pure NIR segmentation and a NIRD segmentation for radish (a), onions (b), and carrots (c). The first row shows an RGB image of the field structure, the middle row contains the results from the NIR segmentation, and the lower row displays the corresponding NIRD segmentation.

## 3.2 Evaluation of the RANSAC Algorithm

As expected the thinning based data downsampling keeps, compared to a  $0^{th}$  and  $1^{st}$  spatial moments based approach, more pixels, retains the orientation of closed areas, and offers nevertheless a significant reduction of the initial dataset.

The RANSAC crop line identification offers high parameter detection rates for all three parameters up to 94% for densely seeded plants which appear along a row as elongated areas (cf. Fig. 5). Since the row space parameter is a constant, the row space estimations are accumulated and depicted within a histogram including the best fitting normal distribution. If the row spaces are tighter and the plants are clearly separated from each other, diagonal line bundles offer a similar high inlier rating as vertical lines and the detection ratio decreases to 63%. Very unstructured rows result in a detection ratio of 75%. The RANSAC algorithm requires for an estimation in average 83ms.



Fig. 5. RANSAC results for elongated rows, row space (right):  $\mu = 0.735$ .

### 3.3 Evaluation of the Particle Filter

All in all, the experiments result in high detection ratios up to 98% for r and  $\theta$ , and demonstrate that the detection is able to adapt within seconds to changing row structures. Figure 6 shows that the filter automatically adapts the row space parameter to the double row distance if whole crop lines are missing, but keeps the correct estimation if the crop lines offer scattered gaps. The precise offset plot additionally shows the manual correction of the CCV's orientation during the data capturing. Table 1 shows a comparison of the results for different plant types and row structures. The particle filter algorithms requires in average 145ms for one iteration.

## 4 Discussion

Two probabilistic methods using geometric models for a crop row determination were tested with different in-field datasets and show good results for the determination of the orientation and offset of parallel lines for elongated row structures. The RANSAC algorithm results in lower detection ratios for the estimation of



Fig. 6. Particle filter results for separated plants, blue curve: first run, other colors: repeated runs, reorientation after a drift of the CCV at t=15s, missing row from t=16s to t=25s.

plant type	row structure	RANSAC	particle filter
carrot	elongated	94.40%	97.71%
onion	separated	75.47%	80.53%
radish	separated	62.94%	95.62%

Table 1. Crop line detection results, RANSAC vs. particle filter.

the correct parameters for separated plants while the introduced cascaded particle filter delivers also sufficiently precise results for such row structures and  $\varepsilon_{max} = 0.15...0.25$ . Both algorithms result in noise afflicted parameter values for the row distance. Since the row space is the only parameter that is constant during the robot's motion, the results can be improved with additional filters. Depending on the variance of the prediction step and the activation status of the particle deprivation avoidance of the second particle filter, the results can be optimized either for a crop line structure identification, or a crop line tracking and in-line navigation. Additional the evaluated computational times for the segmentation and the crop line detection indicate a real-time capability of the algorithms. Further the real movement of the robot will be considered in future works for the crop row estimation, to get more stable results for the orientation and offset with a Kalman-filter based fusion of the crop line determination and the (visual) odometry data.

### Acknowledgment

This work was funded by Sparkling Science – a programme of the Federal Ministry of Science and Research of Austria (SPA 04/84 - FRANC).

## Bibliography

- Åstrand, B., Baerveldt, A.J.: A vision based row-following system for agricultural field machinery. Mechatronics 15(2) (2005) 251 – 269
- [2] Jiang, G.Q., Zhao, C.J., Si, Y.S.: A machine vision based crop rows detection for agricultural robots. In: Wavelet Analysis and Pattern Recognition (ICWAPR), 2010 International Conference on. (2010) 114–118
- [3] Ruiz, G.R., Gracia, L.M.N., Gil, J.G., Junior, M.C.S.: Sequence correlation for row tracking and weed detection using computer vision. In: Workshop on Image Analysis in Agriculture, 26-27. (2010)
- [4] Romeo, J., Pajares, G., Montalvo, M., Guerrero, J., Guijarro, M., Ribeiro, A.: Crop row detection in maize fields inspired on the human visual perception. The Scientific World Journal 2012 (2012)
- [5] Kise, M., Zhang, Q., Rovira Más, F.: A stereovision-based crop row detection method for tractor-automated guidance. Biosystems Engineering 90(4) (2005) 357–367
- [6] Weiss, U., Biber, P.: Plant detection and mapping for agricultural robots using a 3d lidar sensor. Robotics and Autonomous Systems 59(5) (2011) 265 273 Special Issue ECMR 2009.
- [7] Fontaine, V., Crowe, T.: Development of line-detection algorithms for local positioning in densely seeded crops. Canadian biosystems engineering 48 (2006) 7
- [8] Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6) (1981) 381–395
- [9] Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics (Intelligent Robotics and Autonomous Agents). The MIT Press (2005)