Crop Row Detection on Tiny Plants with the Pattern Hough Transform

Wera Winterhalter, Freya Fleckenstein, Christian Dornhege and Wolfram Burgard

***Abstract*—In sustainable farming, robotic solutions are in rising demand. Specifically robots for precision agriculture open up possibilities for new applications. Such applications typically require a high accuracy of the underlying navigation system. A cornerstone for reliable navigation is the robust detection of crop rows. However, detecting crops from vision or laser data is particularly challenging when the plants are either tiny or so large that individual plants cannot be distinguished easily. In this paper, we present a pipeline for reliable plant segmentation in any crop growth stage, as well as a novel algorithm for robust crop row detection that adapts the Hough transform for line detection to detect a pattern of parallel equidistant lines. Our algorithm is able to jointly estimate the angle, lateral offset and crop row spacing and is particularly suited for tiny plants. In extensive experiments using various real-world data sets from different kinds and sizes of crops we show that our algorithm provides reliable and accurate results.**

***Index Terms*—Agricultural Automation; Robotics in Agriculture and Forestry; Field Robots**

I. Introduction

In order to satisfy the rising demand for sustainable farming, autonomous robotic solutions are essential. Precision farming applications typically require a robot to accurately navigate. In this context, the robust estimation of crop rows is a crucial precondition. Applications such as weeding or fertilization require the robot to traverse the field in different growth stages of crops. Thus, a navigation system for agricultural robots needs to be robust with respect to varying plant sizes, plant types, and plants not aligned with the row structure, such as weeds.

This poses several challenges, in particular for the plant feature extraction and crop row detection that we focus on. These include changing soil conditions, tiny plants that are hard to distinguish from soil even for humans (see Fig. 1), or plants so large that rows cannot be clearly separated. In addition, weeds that do not grow in the row structure can be mistaken for crops. Such contrasting conditions call for different sensor modalities, such as high-resolution cameras for detecting small plants or laser range scanners that capture the geometry of larger plants. As we focus on cases in which a clear detection of crops is not possible, plant feature extraction algorithms provide sparse and noisy data.

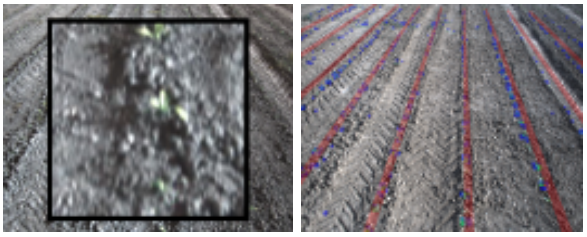

Fig. 1: Image of a sugar beet field with tiny plants (left). The enlargement shows sugar beet plants of about 1 cm. The right shows a typical pattern detected by our algorithm reprojected into the camera image.

Besides dealing with plant feature extraction, a robust crop row detection has to estimate the lateral offset of the rows as well as the angle relative to the robot and the spacing between rows. Although the latter can be assumed to be known, variations exist based on how crops were sown. Crop row detection algorithms often rely on a clear separation of plants and soil in the input data or assume a given spacing [1], [2], [3], [4], [5], [6]. Under such assumptions, it is sufficient to estimate individual crop rows. To be independent of prior knowledge of the field, it is preferable to estimate all relevant parameters and include all data.

Since we want to use the detected crop rows to localize the robot and—more importantly—its wheels with respect to the crops, the algorithm should produce crop row patterns defined in the three-dimensional local robot coordinates.

In this paper, we present a feature map extraction pipeline for different sensor modalities to capture crops of any size, especially tiny plants, that yields a two-dimensional plant map. Furthermore, we propose a novel approach for robust crop row detection that relies on minimal assumptions, most importantly that crops are sown in parallel equidistant lines. Taking only the two-dimensional map from our feature extraction pipeline as input, our approach is independent of the sensor modality. We base our crop row detection on the principle of the Hough transform [7]. In contrast to the classical Hough transform that detects individual lines, our *Pattern Hough transform* extracts a complete crop row pattern in a single step (see Fig. 1). We achieve this by transforming the two-dimensional plant feature map into the three-dimensional space of row patterns that is defined over the angle of the pattern, the lateral offset and the row spacing. We jointly estimate all parameters of a pattern, including its spacing. In contrast to incrementally fitting lines from partial information, the key idea is that all

Manuscript received: February, 24, 2018; Revised May, 25, 2018; Accepted June, 13, 2018.

This paper was recommended for publication by Editor Cyrill Stachniss upon evaluation of the Associate Editor and Reviewers' comments. This work has partially been supported by the European Commission under the grant number H2020-ICT-644227-FLOURISH.

All Authors are with the Department of Computer Science, University of Freiburg, Germany. winterhw, fleckenf, dornhege, burgard@informatik.uni-freiburg.de

Digital Object Identifier (DOI): see top of this page.

input information is used to find the best crop row pattern.

In extensive experiments we show that our algorithm works robustly on noisy and sparse data from tiny plants, as well as on dense data from larger plants. We evaluate our algorithms on different kinds of crops. This includes data from laser range sensors obtained from winter canola and sugar beet plants perceived with cameras. The results indicate that our Pattern Hough transform performs well while driving in the field and outperforms all other algorithms on noisy data in the canola data set. Our RANSAC variant that also follows our idea of using all available data to determine a crop row pattern performs better in some cases.

II. Related Work

Crop row detection has received considerable interest in the past decades. Most approaches use vision data. On RGB data, often pixel-wise color indices are applied to segment vegetation. Various different indices have been proposed, for example *excess green* that emphasizes green values [8] or the *triangular greenness* index that is based on a color spectrum around green [9]. If specialized near infrared cameras are available, often the normalized difference vegetation index is used [10]. Based on such a segmentation, crop row detection can be performed. Kise et al. introduced an algorithm based on stereo vision to create an elevation map [1]. Under the assumption that the row spacing is known and the heading of the sensor corresponds roughly to the heading of the crop rows, four rows are extracted from the elevation map. Other vision-based row detection methods divide the image into a low number of predefined horizontal strips. In each strip they determine center points of rows and connect these points to lines [2], [3]. Another method uses a bandpass filter computed from a known row spacing to create a crop row template that can be matched against the image [11]. All these approaches assume a given row spacing and that the heading of the sensor is approximately aligned with the crop rows. Template matching can also be used to estimate row spacing and offset of a curved pattern, when relying on the assumption that the sensor heading is aligned with the crop rows. Here the template is shifted along the image to compute a minimal score [12]. Explicitly considering curved patterns is not critical, as curvature found in most fields only leads to a lateral deviation of a few centimeters from a straight line in the usable sensor range. The approach by Montalvo et al. assumes that the number and the position of the rows in the image is known [13]. They then perform a linear regression in the region of interest.

A method that estimates the angle and spacing of the rows is presented by English et al. [14]. They convert the image into a top-down view, sum the green values of each column in the image and compute the variance of the resulting sum. The image is skewed, and the procedure is repeated to find the angle with the highest variance in the image columns to estimate the heading. Peaks in the sums of the image columns correspond to crop rows. Further work integrated this with a learning based approach that also takes depth data into account [15]. For every new field, a manually labeled image and training run over 20 s is required. They achieve a lateral error of between 1.6 cm and 3.1 cm depending on the field.

Several approaches are based on the Hough transform. They estimate single lines or a pattern under the assumption of a given row spacing [4], [5], [6], [16]. These differ in how the Hough transform is applied. To get more stability against noise, giving the line a predefined width can be achieved by summing neighboring pixels [4]. By separating an image into three parts of width equal to the given spacing it is possible to incorporate information from parallel lines. The individual parts are summed before applying the Hough transform [5]. Without the given number of crop rows, but a known spacing one can also sum all pixels in parallel lines in the Hough space [6], [16]. While all these approaches aim to tailor the classical Hough transform [7] towards crop row pattern detection, all make assumptions on known values such as a given spacing or the number of crop rows. In contrast, our approach is related to the generalized Hough transform [17], as we adapt the model to directly fit the desired shape of parallel equidistant lines.

All in all, most existing approaches work on vision data, while our approach is capable of extracting and using plant features from vision and laser data and is thus independent from the sensor modality. Furthermore, many approaches assume that the crop row spacing is given a priori or that the heading of the sensor corresponds to the heading of the crop rows. Our approach is able to estimate both the heading and the spacing of the crop rows in addition to the lateral offset and thus does not make the assumption of pre-defined values. In particular, our Pattern Hough transform extracts a crop row pattern in a single step instead of detecting individual lines.

III. Feature Map Extraction

For robust crop row detection it is necessary that the algorithm works with various sensor modalities to be applicable in different environments. We achieve this by computing feature maps from arbitrary sensor data. These feature maps provide a generic input for the crop row detection algorithms. A feature map is a two-dimensional grid map, where each cell contains a weight that describes the likelihood of vegetation being present. Such feature maps are defined in the coordinate frame of the robot and located on the ground plane. This ensures that real-world geometric relations between plants are recovered in the feature map (see Fig. 2). In particular crop rows sown in parallel equidistant lines appear as such in the feature map. This is especially important for vision data, where dependent on the perspective crop rows do not appear parallel. We give two procedures for feature map extraction for either vision or laser data.

a) Vision-based Segmentation: Given RGB images of plants, we first compute the *triangular greenness* index for each pixel. This index forms a triangle in RGB space around the spectral values of leaf chlorophyll [9]. Thus larger index values imply that this pixel belongs to a plant. We use a threshold of 0.06 to segment vegetation based on this index.

To rectify the image and project image pixels into the feature map in the ground plane the camera must be intrinsically and

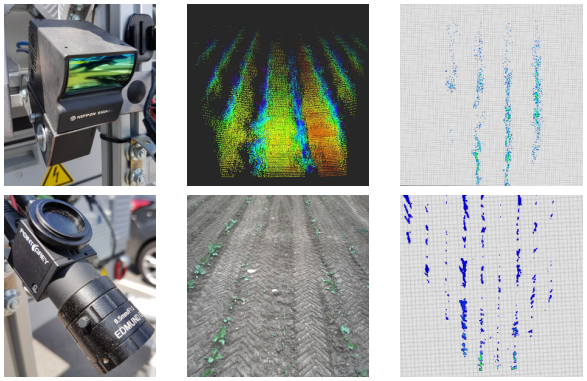

Fig. 2: This figure shows the 3d laser and vision camera mounted on our robot; their raw data, i.e., 3d point clouds or RGB images; and the resulting feature maps. Note that in both cases plant features form equidistant and parallel lines.

extrinsically calibrated to the robot. We intersect the 3d ray from the camera origin through each pixel with the ground plane to determine the corresponding cell in the feature map and increase its weight by the index value. The projection into the coordinate frame of the robot makes the feature maps independent of the sensor placement.

As plants further away from the camera appear smaller, i.e., occupy a smaller area in the image, we re-weight the feature points based on the squared distance from the camera. Finally we retain the best 25% of all feature points to remove noise.

b) Laser-based Segmentation: We use extrinsically calibrated 3d laser range scanners that provide 3d point clouds relative to the robot. As these are less dense than vision data, we integrate the point clouds over 2.5 seconds using the odometry of the robot (see Fig. 2).

Since the ground is usually the lowest in the local robot frame, we consider all points that are below a certain height threshold as ground points. The remaining non-ground points are then treated as vegetation. Instead of a fixed threshold, we dynamically determine this at the upper 10% of all height values, i.e., we retain the 10% highest points above the ground. The weight of the feature map cell corresponding to the 2d projection of the point is increased by its height above the ground.

IV. Crop Row Detection

We aim for a crop row detection that determines a set of rows localized in metric space relative to the robot and is therefore suited for applications like localization or mapping. The input is a two-dimensional feature map in the robot coordinate frame described in the previous section. In most fields crops form lines that—at least locally—appear parallel and equidistant. Therefore our goal is to extract a crop row pattern $p_{o,\theta,s}$ of parallel equidistant lines from a feature map. Here, o is the offset of the line closest to the origin of the robot frame, θ is the angle of the pattern lines, and s is the spacing between lines in the pattern (see Fig. 4).

We cannot expect the vegetation segmentation used to compute the feature maps to be perfect. This holds especially under harsh conditions, when large plants overlap crop rows or plants are extremely small and thus feature points are sparse. Moreover false positives during segmentation or weeds produce noise that is not aligned with the crop row pattern. To be robust in such environments we propose to extract all rows of a pattern at once, thus considering all available data in contrast to only detecting individual lines.

In this section we present multiple approaches for crop row detection. The first two algorithms use the Hough transform [7]. This is a robust algorithm developed to detect individual lines in images. We describe its application as a naive approach to detect a single crop row, which—with a pre-given spacing—defines a pattern. This approach is widely used for crop row detection [4], [5], [6], [16]. We denote this as the *Line Hough* transform. We then describe our *Dual Line Hough* transform that uses the results of the Line Hough transform to also estimate the row spacing by extracting a second parallel line.

In contrast to extracting individual lines, our *Pattern Hough* transform follows our idea to extract all lines of a pattern at once. Finally as a comparison, we present a variant of the random sample consensus (RANSAC) algorithm [18], the *Pattern RANSAC*, that also determines all rows jointly.

In general all these algorithms aim to find a model, e.g., a line, that best fits the data, in our case the feature map. Algorithms based on the Hough transform compute the model with maximum support by transforming the 2d data from the feature map into a histogram h over the space of possible models $\mathcal{H}$, e.g., the space of lines. For each vegetation cell c in the feature map the subset of models $H(c) \subset \mathcal{H}$ that contain c is determined. For each element in $H(c)$ the corresponding histogram value is increased by the weight of the cell. Intuitively, each bin in the histogram represents a model and the value of that bin corresponds to the weight of all cells supporting this model. The maximum value of h then represents the result of the algorithm, i.e., the best fitting model. In contrast, RANSAC samples a minimal number of feature map cells that define a candidate model. Through iterative sampling a large number of candidate models is obtained. For each model the support is computed by accumulating the weights of the feature map cells that lie on this model. The model with the maximum support is then the best fitting model.

We use the following parametrizations for lines and crop row patterns. We represent crop rows as lines in 2d space. All lines $l_{r,\theta}$ are represented in the Hesse normal form

$$l_{r,\theta} := \left\{(x,y) \in \mathbb{R}^2 \mid r = x \cdot \cos(\theta) + y \cdot \sin(\theta)\right\}, \quad (1)$$

where θ defines the normal vector $\vec{n}_\theta = (\cos(\theta), \sin(\theta))^t$ to the line $l_{r,\theta}$ and r is the signed distance of the line to the origin. An illustration of the Hesse normal form is shown in the left image of Fig. 3. A crop row pattern $p_{o,\theta,s}$ is a set of parallel equidistant lines, i.e.,

$$p_{o,\theta,s} := \{l_{r,\theta} \mid r = n \cdot s + o \text{ for any } n \in \mathbb{Z}\}, \quad (2)$$

where θ is the angle of the normal vector of all lines, s is

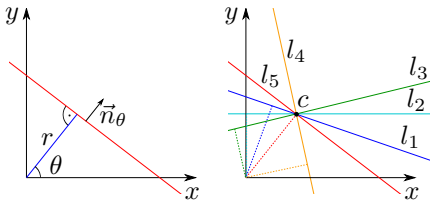

Fig. 3: The left side illustrates the Hesse normal form of a two-dimensional line (red), where θ defines the angle of the normal to the line and r the distance to the origin. The right side visualizes lines $l_1, \ldots, l_5$ of the set $H_L(c)$ for a cell c.

the spacing or distance between consecutive lines and o is the offset of the pattern from the origin. An example of such a pattern is shown on the left side of Fig. 4.

A. Line Hough Transform

The Line Hough transform is often used for crop row detection as it determines the best fitting line from a set of points. The Hough space $\mathcal{H}_L$ of this transform is the space of lines (see Equation 1) and defined as

$$\mathcal{H}_L := \{l_{r,\theta} \mid r \in \mathbb{R}, \theta \in [0, \pi)\}. \tag{3}$$

Given the coordinates of a cell c in the grid, the Hough transform maps c onto the set of lines $H_L(c)$ that pass through c. For each possible value of θ, Equation 1 defines the matching value of r (see Fig. 3). The corresponding value for each (θ, r) in the Hough space is incremented to account for each possible line. As this leaves θ free and r bound, for each cell a one-dimensional subset of $\mathcal{H}_L$ is examined.

With the Line Hough transform one can easily extract a best-matching single crop row line l_{r^*,θ^*}. Many approaches consider this sufficient and employ the Hough transform or other line detection algorithms with a pre-given row spacing s^* to define a pattern p_{r^*,θ^*,s^*} [4], [5], [6], [16]. Extracting additional lines for other crop rows is not trivial as, for example, the second best line in the Hough space is often similar to the best line, i.e., only using a slightly different θ or r. For crop row detection, however, one is usually interested in lines that represent a different crop row.

B. Dual Line Hough Transform

Given the information already computed by the Line Hough transform, i.e., $\mathcal{H}_L$, it is possible to retrieve a second crop row and thus the pattern spacing under the reasonable assumption that this line is parallel to the best matching line and one has a rough idea of the spacing. We thus only assume that instead of a pre-given spacing we have a minimum spacing s^- and maximum spacing s^+ that in practice are set to be within 15 cm around the expected spacing.

Given the best matching line l_{r^*,θ^*} from the Line Hough transform, we now define a new search pattern that does not consider the full Hough space, but only

$$\mathcal{H}'_L := \{l_{r',\theta^*} \in \mathcal{H} \mid s^- \leq \|r^* - r'\| \leq s^+\}. \tag{4}$$

The best value in $\mathcal{H}'_L$ represents the parallel line to l_{r^*,θ^*} with the highest support within the minimum and maximum spacing. The difference between the offsets of these two lines thus determines the row spacing. As this variant of the Hough transform uses two lines to define a crop row pattern, we denote it as the *Dual Line Hough* transform.

C. Pattern Hough Transform

Although the approach described in the previous section estimates a crop row pattern, it only uses the information from two lines to estimate the spacing, while the offset and normal angle are determined by a single line. Especially in data that is not perfectly segmented, we aim for an approach that considers all available data to determine the best pattern. To this end we follow the principle of the Line Hough transform and create the Pattern Hough transform that estimates a complete crop row pattern in a single step.

The key idea is to directly transform into the space of row patterns $\mathcal{H}_P$ instead of the space of lines $\mathcal{H}_L$. The set of all such patterns $p_{o,\theta,s}$ as defined in Equation 2 is given by

$$\mathcal{H}_P := \{p_{o,\theta,s} \mid s \in \mathbb{R}_+, o \in [0, s), \theta \in [0, \pi)\}. \tag{5}$$

An example of how changing the angular parameter θ or the spacing parameter s affects the pattern is shown in Fig. 4.

Similar to the Line Hough transform, we map a cell c onto the set of all patterns $H_P(c) \subset \mathcal{H}_P$ that matches this cell and increase the corresponding entry in the histogram. Here, a pattern matches a cell if it contains a line going through c. Such a pattern $p_{o,\theta,s} \in H_P(c)$ is defined by values of θ and s, as the offset o is then bound by combining the constraints from Equation 1 and Equation 2, yielding $c_x \cdot \cos(\theta) + c_y \cdot \sin(\theta) = n \cdot s + o$ for some $n \in \mathbb{Z}$ such that $o \in [0, s)$. To determine $H_P(c)$ we thus iterate over all possible values of θ, and all values of s in $[s^-, s^+]$. Therefore a two-dimensional subset of $\mathcal{H}_P$ is examined. Note that shifting the offset of a pattern by the spacing s results in an equivalent pattern according to $r = n \cdot s + o$ for some $n \in \mathbb{Z}$ (see Equation 2). We therefore constrain the offset to be in the range $[0, s)$ to gain a unique pattern $p_{o,\theta,s}$ Equation 5.

Although in comparison to the Line Hough transform our Pattern Hough transform adds one dimension to the histogram and the search space, the additional computational requirements are limited since—by our design—the range of the offset parameter o is bounded by the spacing parameter s. Finally, the result of the Pattern Hough transform is the maximum value in $\mathcal{H}_P$ and directly represents the best matching pattern over all parallel equidistant lines.

D. Pattern RANSAC

The Pattern RANSAC is a RANSAC variant that directly computes crop row patterns. To find candidate patterns, we first sample three points from the feature map. The first two points define a line $l_{r,\theta}$. The distance of the third point to this line determines the spacing s. If s is not within $[s^-, s^+]$ we reject the candidate pattern and continue sampling. To evaluate a candidate pattern $p_{r,\theta,s}$ we accumulate the weight of all cells in the feature map that lie on this pattern. We retain the candidate pattern with the highest accumulated weight as the result. Similar to the Pattern Hough transform the Pattern

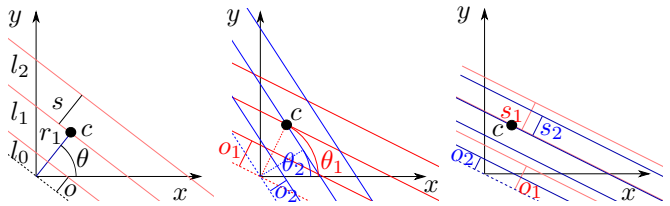

Fig. 4: Left: The lines l_0, l_1, l_2 (red) of a pattern with parameters o (offset), θ (angle), s (spacing). The blue line depicts the distance r_1 of line l_1 from the origin, equal to $1 \cdot s + o$. Middle: Two patterns from the set $H_P(c)$ for a cell c with angles θ_1 (red) and θ_2 (blue) and the same spacing. Right: Two patterns with the same θ for spacings s_1 (red) and s_2 (blue). Note that a given spacing determines the offset of the pattern through c.

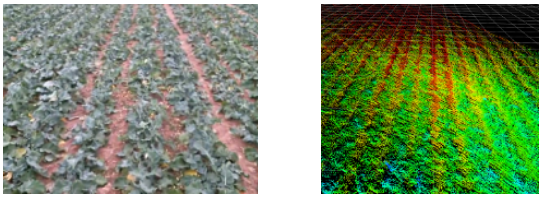

Fig. 5: This figure shows a web cam image of the canola field and the corresponding laser range data.

RANSAC therefore finds the pattern that is supported by most data across the feature map.

V. Evaluation

We evaluate the performance of our approach regarding robustness on various real-world data sets featuring different crops and sensor modalities. Robustness in our case means that a crop row detection algorithm is reliably able to produce crop row patterns that are close to the actual real-world situation. This is crucial especially as an input for autonomous navigation.

We ran all four crop row detection algorithms presented in Section IV on every data set, which we processed either with the vision or laser feature map extraction respectively. In all cases the same parameter sets have been used independent of the data set and algorithm. The histograms for the Hough based algorithms had an angular resolution of $0.57°$ and an offset and spacing resolution of 1 cm. As the Pattern RANSAC algorithm relies on an incremental improvement, we evaluated this with 2 500, 5 000, and 25 000 iterations and repeated each run five times to account for its probabilistic nature. We chose the number of iterations so that one variant is faster than the Pattern Hough transform, one takes a similar amount of time and one gets close to the best possible result when time is not an issue.

In the following, we first describe the hardware used to record the data sets and to perform the evaluation, then we introduce our data sets before investigating the robustness of crop row detection in general. Afterwards we discuss data sets that are especially challenging in detail and close with considerations on the applicability for real-world scenarios.

Fig. 6: Our agricultural robot, the BoniRob, on a leek field (left). The right side shows a georeferenced image of a sugar beet field overlaid with a crop row pattern (magenta) extracted from the image data.

A. Hardware Description

All data has been recorded with our agricultural robot BoniRob (see Fig. 6). It was driven at different speeds with a maximum of 1 m/s manually and 0.2 m/s autonomously. Vision data was recorded with a PointGrey Blackfly color camera with five megapixels and laser data was recorded with a Nippon-Signal FX-8 3d laser range scanner (see Fig. 2). Both sensors are mounted centered in the front of the robot at a height of about one meter above the ground and are tilted downwards by about $25°$. Crop row detection on the robot during autonomous navigation was performed on a Pokini i2 computer with an Intel Core i7-4600U CPU with 16 GB RAM. Evaluations in this section were performed on an Intel Core i7-4770 CPU with 16 GB RAM.

B. Data Sets

We collected data from various different plants. We recorded vision data from sugar beet fields with medium sized plants of about 5 cm (see Fig. 2 (bottom)) and tiny plants that had just recently emerged. With a size of about 1 cm these plants are hard to perceive even for a human (see Fig. 1). Furthermore, we recorded 3d laser data from a canola field (see Fig. 5), a corn stubble field, and a leek field (see Fig. 2 (top)). In the canola and leek data sets, the plants are in a later growth stage and thus can be easily seen in the laser range data. The corn stubbles are harder to segment as they are small and therefore hardly recognizable in the laser data. They also do not present a distinctive color difference to the soil in images.

All data sets were split into *in row* motions, where the robot drives in the field aligned with crop rows and *transition* motions, where the robot either leaves or enters the field. The covered distance for driving *in row* and the covered angle for *transitions* are listed in Table I. When leaving or entering the field crop rows are usually only partially visible and the robot is not necessarily aligned with the field, e.g., when turning towards it as in the leek and sugar beet transition data sets (see Table I). For the canola and corn transition data sets, the robot was mostly leaving the field and thus not changing its orientation notably. Such scenarios present hard challenges to crop row detection algorithms as less data is available to detect crop row patterns, while at the same time the sensors see areas not part of the field that also contain vegetation. Nevertheless these situations are especially

	in row					transition				
	Canola	Corn	Leek	Medium Sugar Beet	Tiny Sugar Beet	Canola	Corn	Leek	Medium Sugar Beet	Tiny Sugar Beet
number of feature maps	47	24	79	70	63	59	61	87	55	88
distance / angle covered	43.94m	35.69m	383.36m	70.81m	109m	1.98°	8.00°	45.84°	49.46°	62.72°
mean vegetation density	0.7%	1.6%	0.5%	0.6%	0.5%	0.6%	0.6%	0.7%	0.3%	0.3%

TABLE I: Data set properties. When driving *in row*, the angle does not change a lot, and during *transitions*, the robot does not cover much distance. The vegetation density denotes the percentage of cells that contain plant features.

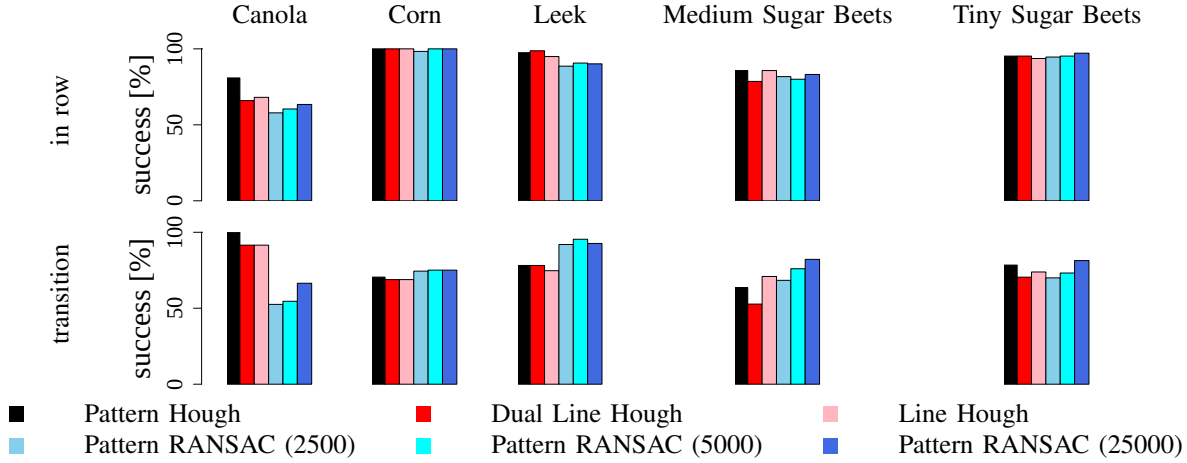

Fig. 7: This figure shows the success rates for all algorithms on each data set.

important for autonomous navigation as they allow the robot to accurately leave and enter a field.

To evaluate the success rate we manually labeled crop row patterns for integrated point clouds in the laser data sets and images for the vision data sets. As feature maps extracted during transitions potentially change in appearance quickly, we evaluated a feature map every 0.1 s. When driving in a row, we only evaluated a feature map every 5 s since the surroundings stay similar for a long time due to low speed. The number of feature maps is shown in Table I.

C. Robustness

Robustness is crucial for guiding an autonomous vehicle. We define robustness by the percentage of successfully extracted patterns, i.e., that the angular and lateral error are within reasonable thresholds for navigation. Here, the error of the spacing is not essential, as navigation relies on the angular and lateral errors. We quantify the error between the computed pattern and the labeled ground truth pattern by three measures: the angular error (theta error) between the normal angles of the patterns, the spacing error as the difference of the computed spacing to the ground truth, and the lateral error. The lateral error is defines as the lateral distance of the computed pattern in comparison to the ground truth pattern at a point one meter in front of the robot.

We say that a pattern extraction was successful, if the angular error is smaller than $10°$ and the lateral offset does not exceed $0.10\,\mathrm{m}$. Both measures are crucial to control the angle and sideways tracking of a robot following a crop row [4]. The thresholds were determined in real-world experiments on our BoniRob (see Fig. 6).

Fig. 7 shows the success rates for all algorithms and data sets. For all approaches, the success rate depends on the data set. For small plants, when feature points are extracted accurately while driving *in* the sugar beet field, the linear structure is clearly visible and therefore all algorithms perform well. In the medium sugar beet data set, the Pattern Hough transform performs best with 85%, while other algorithms show similar performance. This is also true for most other data sets, while driving *in row*, besides the canola, where crops were larger and thus a clear line is not easily visible (see Fig. 11).

Most data sets pose harder challenges when the robot is *transitioning* as less data is visible overall and the sensors not only view the field, but also possible vegetation off the field. This can be seen clearly in the corn data set, where the stubbles present quite sparse data and thus there is not much noise driving in the field. This results in almost 100% success rates for all algorithms here. In contrast, when not all data is visible while entering or leaving the field, success rates drop to around 75% for all algorithms.

For larger plants, as in the canola and leek fields or the medium sugar beets, plant feature points are distributed across the larger plant surface, so that the linear structure is harder to recognize. Therefore, in these data sets we see notable differences in the success rates also while driving *in row*. Here, the Pattern Hough transform still performs well in comparison, where especially in canola with large plants all other approaches perform worse.

When driving in row the Pattern Hough transform usually performs similar or better than others. This is different during transitions. For canola all Hough-based algorithms have higher

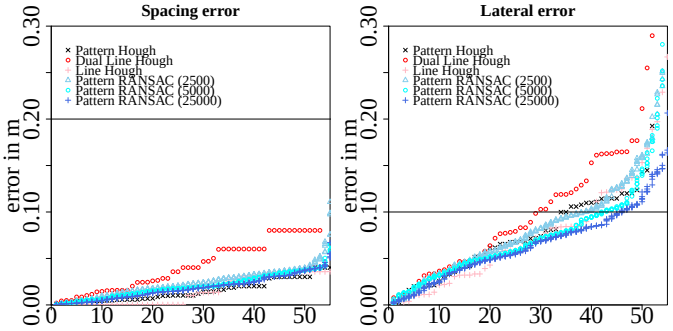

Fig. 8: Evaluation of spacing error (left) and lateral error(right) in the *medium sugar beet transition* data set. Shown are individual errors for each measurement for a data set. The error values have been sorted by size. The black horizontal line shows our threshold for a successful pattern extraction.

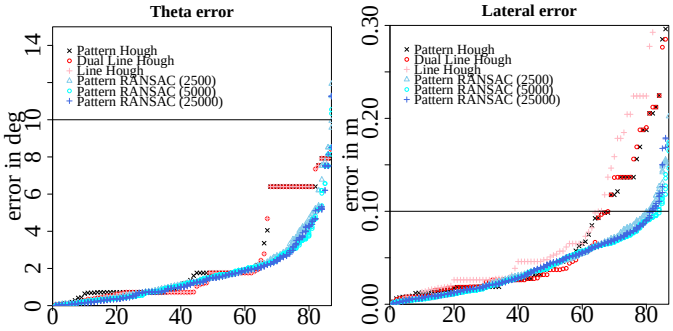

Fig. 9: Evaluation of theta error (left) and lateral error (right) in the *leek transition* data set.

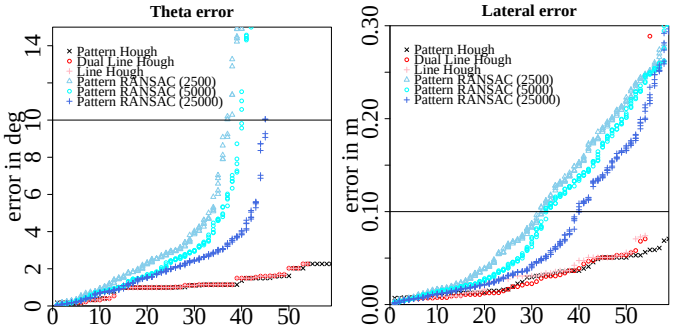

Fig. 10: Evaluation of theta error (left) and lateral error (right) in the *canola transition* data set.

success rates than RANSAC, while the opposite is true in leek. For the medium sugar beets the Dual Line Hough shows worse success rates than the Pattern or Line Hough. We consider these cases in detail in the next section.

D. Discussion

In this section we show how the different algorithms behave in special cases by investigating the distributions of individual errors for a specific data set. In particular, we are interested in why RANSAC appears to be more robust than the Hough variants in the leek and medium sugar beet transition data sets, as can be observed in Fig. 7.

For the medium sugar beets, we take a closer look at the spacing and offset errors shown in Fig. 8. Most prominent in both plots is the considerable error when applying the Dual Line Hough transform. Here, estimating the spacing only using the closest parallel line is not always correct, as the sugar beets were not sown evenly on the field. This shows that, in realistic scenarios, an estimation of the crop row spacing across all data is necessary.

While the Pattern Hough transform only finds a pattern with a suitable lateral error for about 35 of the 60 images, for most of the other images the lateral error is only about $2-3\,cm$ above the threshold. As the choice of $10\,cm$ as a threshold for the lateral error is arbitrary, we conclude that, even though by our definition the Pattern Hough transform has a lower success rate than RANSAC, in this case it performs reasonably well although at slightly less precision than the Pattern RANSAC.

For the leek transition data set, a detailed evaluation plot is shown in Fig. 9. We observe that about ten images (70-80) show a larger theta error for the Pattern Hough transform. In this case, the data was quite sparse, which caused the increase in theta error by about $4°$ in comparison to RANSAC. Although the theta error in these cases is still below the threshold, the slightly off-angle pattern causes lateral errors exceeding the success threshold.

In contrast, RANSAC performs a lot worse than the Pattern Hough transform in the canola transition data set (see Fig. 7). It is clear from Fig. 10 that the theta error diverges from image 35 to 40, which subsequently causes large lateral errors. In this case however, the failure of the Pattern RANSAC to correctly determine the angle is too large to be recovered by adjusting the threshold. Note that even though the Pattern Hough transform has a success rate of 100%, it does not yield perfect results: The lateral error still goes up to $7\,cm$. This is low enough for navigation applicability and thus we consider it a success.

Overall, the Pattern Hough performs well in row, and especially outperforms other algorithms in the noisy canola data set. In contrast the Pattern RANSAC has higher success rates while transitioning in the leek data set and slightly higher success rates while transitioning in corn and medium sugar beets, where more iterations show improvements. Most importantly, in the cases where the Pattern Hough transform had lower success rates, we observed from the individual error plots that its error values did not diverge. The lower success rates were explained by lower precision in comparison with the Pattern RANSAC. These results show that using the Pattern Hough transform as correction input of a localization filter is viable, since a filter does not require perfect measurements to enable reliable navigation.

E. Applications for Navigation

The main use case for crop row detection is as an input for localization during autonomous navigation. Here, sufficiently fast computation is essential. We measured the computation time of each algorithm and report the mean time over each data set in Table II. Here, the Line Hough and Dual Line Hough are an order of magnitude faster in comparison to the Pattern Hough as the Hough space is of a lower dimensionality.

	Canola	Corn	Leek	Medium Sugar Beet	Tiny Sugar Beet
Line Hough	8	13	8	9	8
Dual Line Hough	8	12	8	9	8
Pattern Hough	89	119	71	53	41
RANSAC 2500	34	40	33	27	27
RANSAC 5000	69	79	64	53	52
RANSAC 25000	340	391	319	259	255

TABLE II: Mean computation times in ms.

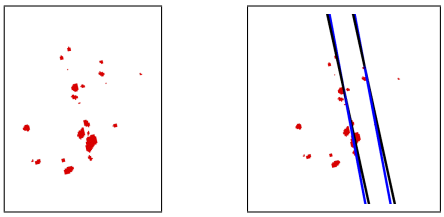

Fig. 11: A feature map from the canola transition data set (left), overlaid with correct row patterns from the Pattern Hough (black) and RANSAC 25000 (blue) (right). Both algorithms are able to extract row patterns in difficult scenarios.

As desired, RANSAC 2 500 is faster than our Pattern Hough, RANSAC 5 000 is comparable, and RANSAC 25 000 is five times slower, but can get better results than with 5 000 iterations. The larger times for all algorithms on the Corn data set can be explained by the higher vegetation density in the feature maps while driving in row (see Table I), as all vegetation features need to be evaluated (RANSAC) or transformed (Hough). Overall, all computation times are reasonable for online operation and in fact we use the Pattern Hough as part of our autonomous navigation system successfully [1].

Besides autonomous navigation crop row detection can also be used for mapping, either based on the sensors of the robot, or as shown in an example in Fig. 6 (right) to extract semantic information of crop rows from a georeferenced overhead image acquired with a UAV [2]. To extract such a map, we used our vision-based segmentation on the overhead image to produce a feature map and applied the Pattern Hough transform. This information can then in turn be used by a ground vehicle to localize itself—using crop row detection—in the field.

VI. Conclusion

In this paper we presented a novel approach that robustly creates a sensor-independent feature representation from different sensor modalities and detects crop rows from noisy or sparse data. A key idea is to determine a pattern that is best supported by all data in contrast to (incrementally) extracting single lines. Our Pattern Hough transform estimates the spacing between equidistant rows directly. While other approaches rely on strong assumptions and prior information limiting the search space, the Pattern Hough performs a global optimization over all necessary parameters with all data provided. Estimating the complete pattern instead of reconstructing the spacing in post-processing steps or assuming it given constitutes a sound method for crop row detection. Our Pattern RANSAC follows the same philosophy.

Our approaches work particularly well on tiny plants, a use case that is rarely tackled by other systems. Our results show that we achieve high robustness especially when driving in the field, which makes it well suited for the intended application of precision farming. We furthermore showed that in some hard cases the Pattern RANSAC performs better than the Pattern Hough based on our requirements for precision, although even in these cases the Pattern Hough transform did not fully diverge from the ground truth. We are already using the Pattern Hough transform as part of our system for autonomous navigation. In future work we will focus on increasing the robustness of the complete navigation system leveraging our crop row detection methods.

[1] A video demonstrating our approach on different fields can be found at https://www.youtube.com/watch?v=0VIwuCaTHPM.

[2] Thanks to Raghav Khanna from ETH Zurich for providing the image.

References

[1] M. Kise, Q. Zhang, and F. R. Más, "A stereovision-based crop row detection method for tractor-automated guidance," *Biosystems Engineering*, vol. 90, no. 4, pp. 357–367, 2005.

[2] H. Søgaard and H. Olsen, "Determination of crop rows by image analysis without segmentation," *Computers and Electronics in Agriculture*, vol. 38, no. 2, pp. 141–158, 2003.

[3] N. Tillett and T. Hague, "Computer-vision-based hoe guidance for cereals—an initial trial," *Journal of Agricultural Engineering Research*, vol. 74, no. 3, pp. 225–236, 1999.

[4] B. Åstrand and A.-J. Baerveldt, "A vision based row-following system for agricultural field machinery," *Mechatronics*, vol. 15, no. 2, pp. 251–269, 2005.

[5] T. Bakker, H. Wouters, K. van Asselt, J. Bontsema, L. Tang, J. Müller, and G. van Straten, "A vision based row detection system for sugar beet," *Computers and Electronics in Agriculture*, vol. 60, no. 1, pp. 87–95, 2008.

[6] V. Leemans and M.-F. Destain, "Line cluster detection using a variant of the hough transform for culture row localisation," *Image and Vision Computing*, vol. 24, no. 5, pp. 541–550, 2006.

[7] P. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3.069.654, 1962.

[8] G. E. Meyer and J. C. Neto, "Verification of color vegetation indices for automated crop imaging applications," *Computers and Electronics in Agriculture*, vol. 63, no. 2, pp. 282–293, 2008.

[9] D. Constantin, M. Rehak, Y. Akhtman, and F. Liebisch, "Detection of crop properties by means of hyperspectral remote sensing from a micro UAV," *Bornimer Agrartechnische Berichte*, vol. 88, pp. 129–137, 2015.

[10] A. Pérez, F. López, J. Benlloch, and S. Christensen, "Colour and shape analysis techniques for weed detection in cereal fields," *Computers and Electronics in Agriculture*, vol. 25, no. 3, pp. 197–212, 2000.

[11] T. Hague and N. Tillett, "A bandpass filter-based approach to crop row location and tracking," *Mechatronics*, vol. 11, no. 1, pp. 1–12, 2001.

[12] I. Vidović, R. Cupec, and Ž. Hocenski, "Crop row detection by global energy minimization," *Pattern Recognition*, vol. 55, pp. 68–86, 2016.

[13] M. Montalvo, G. Pajares, J. Guerrero, J. Romeo, M. Guijarro, A. Ribeiro, J. Ruz, and J. Cruz, "Automatic detection of crop rows in maize fields with high weeds pressure," *Expert Systems with Applications*, vol. 39, no. 15, pp. 11 889–11 897, 2012.

[14] A. English, P. Ross, D. Ball, and P. Corke, "Vision based guidance for robot navigation in agriculture," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[15] A. English, P. Ross, D. Ball, B. Upcroft, and P. Corke, "Learning crop models for vision-based guidance of agricultural robots," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys. (IROS)*, 2015.

[16] J. Marchant, "Tracking of row structure in three crops using image analysis," *Computers and Electronics in Agriculture*, vol. 15, no. 2, pp. 161–179, 1996.

[17] D. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.