



Crop edge detection based on stereo vision

Johannes Kneip, Patrick Fleischmann*, Karsten Berns

Robotics Research Lab, Department of Computer Science, Technische Universität Kaiserslautern, Gottlieb-Daimler-Str., 67663 Kaiserslautern, Germany



ARTICLE INFO

Article history:

Available online 18 October 2019

Keywords:

Agricultural automation
Computer vision for automation
Visual-based navigation
Advanced driver-assistance systems (ADAS)

ABSTRACT

This paper focuses on the development of a crop edge detection algorithm based on the point cloud produced by a stereo camera system using the GPU for fast matching of the camera images. The approach utilizes the 3D characteristics of the transition between the crop and the stubbles or the ground. Therefore, the point cloud is sorted into a grid of cells to create an elevation map. A segmentation in crop and ground is obtained using the Expectation-Maximization algorithm with a Gaussian Mixture Model to represent the distribution of the cell's heights. This segmentation is Bayesian filtered over a short time frame to create a more robust segmentation result. Afterward, the resulting potential crop edge locations are processed using robust linear regression to come up with an overall linear crop edge model. The implemented system has been tested in a series of experiments with detailed results stated at the end of this work.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

In 2009, *The Food and Agriculture Organization (FAO)* of the *United Nations* predicted the world's population to reach 9.1 billion by 2050 [1]. With a global human population growth of around 75 million annually, the challenge for the future will be to provide enough food. To give an example, the world's demand for cereal in 2010 was 2.1 billion tons, whereas the estimated need in 2050 is 3.0 billion tons, which is an increase of 43%.

In the current era of globalization and digital revolution, the agricultural market segment reacts accordingly contemporary. The need to feed the ever-increasing number of people from agricultural land will also be of major importance in the future. The area of land used will at best remain constant but will most likely further increase over the years. The trend towards a considerably greater demand and use of all sorts of agricultural machinery is set and recognizable. The scope of this problem has broadened significantly in recent years and the market evolves solutions to address and counter this topic. A rethinking of how crop production is done has taken place or is ongoing to identify and solve pending or emerging issues.

Driven by economic factors (fuel, product selling prices), political factors (restrictions, subventions), ecological factors (weather conditions, soil quality), and social factors (security of land tenure, skills, and roles) manufacturers rethought existing technologies and strive for innovative solutions for intelligent and sustainable farming methods. Research on automation of agricultural

vehicles has accelerated recently, which of course is partially also a consequence of the current era of technology [2,3]. Some demands for new autonomous machinery is to improve their efficiency while for example minimizing their soil compaction [4], which leads to improved yield and profitability. With regard to increasing vehicle performance, the operator is one of the most considerable limitations [5]. The most concerning factor due to his significant impact on quality is the tedious and monotonous work for long hours. The operator has to supervise the field operation while steering, which requires a high level of concentration and is quite demanding over time. The market came up, among others, with assistance systems for example for tractors or combine harvesters. With the aid of such systems, the operator can, therefore, pay more attention to other important processes like the implement performance [6]. The system running in the background is capable of steering the vehicle along a predefined path, e.g., via GNSS, LIDAR, or other sensors like machine vision for the environment perception.

The product solutions of such already existing systems significantly support the farmers and help to cut costs. In order to reach a broad range of (potential) customers another wish of manufacturers is to offer affordable and reliable systems. While GNSS guidance systems ensure a certain precision, usually within an error of less than 10 cm, the costs are rather high [7]. A low-cost and more adaptive solution to this could be utilizing camera sensors to allow the local environment perception.

In this paper, we present a stereo vision setup for online and adaptive detection of crop and cutting edges (see Fig. 1). It was chosen to build the application based on a stereo camera, as the sensor is – compared to LIDAR – a rather inexpensive solution for local navigation and environment perception. Although LIDAR

* Corresponding author.

E-mail addresses: j.kneip@cs.uni-kl.de (J. Kneip), [\(P. Fleischmann\)](mailto:fleischmann@cs.uni-kl.de), [\(K. Berns\)](mailto:berns@cs.uni-kl.de).



Fig. 1. Visualization of the threaded application: crop edge detection – in this case, in a wheat field.

measurements are known to be more precise, the further advantages like robustness, a higher density in the VFOV of the point cloud, and the availability of color data were the basis for this decision. Besides, preliminary tests have shown, that the laser measurements are heavily affected by dust which is caused by the harvesting process. Unlike monocular vision, stereo allows to work with geometric constraints and overcomes the problems of a fixed camera to ground plane calibrations, e.g., in sloped terrain. To some extent, stereo vision is also able to deal with shadows as the texture changes in both images simultaneously and therefore does not affect the matching algorithm. Besides, the calculated 3D information, from which the height and volume of the crop can be estimated, is valuable information for further improvements in the harvesting process.

Additionally, as for local guidance, no a priori map of the area is needed to operate based on this device, and an additional safety-relevant obstacle detection system may be easily integrated. Thus, the advantages of stereo vision, the benefits of positioning the camera on top of the cabin [8], together with the knowledge gained through previous investigations, and the need for robust crop edge detection algorithms, motivated this research. The proposed system consists of three decisive modules: *Stereo Vision System* (see Section 3), *Elevation Map* (see Section 5) and *Robust Linear Regression* (see Section 7) as presented in the schematic overview (see Fig. 2).

2. Related work

Even though there has been related research in this area, published papers are not too many and sometimes kept partly confidential. Prior used methods can be summarized into three groups:

1. Image segmentation and edge detection

- Robert, Sobel, Prewitt, Kirsch, and Laplacian of Gaussian (LoG) filter
- Gray level calculation, Haralick features
- Morphologic processing, e.g., erosion, dilation
- K-means clustering

2. Scan line, Hough transform, and Least-squares method

3. Wavelet analysis and fuzzy logic recognition

Table 1 gives an overview of publications since 1996. It is sorted by year in descending order and provides a summary of the systems main approach, used sensors, root-mean-square deviation of the detection, and type of crop the system can operate with.

When designing a system for an agricultural environment the circumstances that are naturally given and that are likely to occur must be considered. Depending on which characteristics the system should be able to cope with and prior knowledge about the vehicle, implement and field conditions, the sensors and system architecture is build upon. In the following, typical characteristics in the context of crop edge detection field scenarios are listed:

- Ground surfaces are mostly even, but may also be uneven
- Operating areas are large
- Weather conditions (fog, rain, sun, dust, etc.) affect sensors and influence soil/ground conditions
- Large cultivated areas which may contain weeds and barren areas (holes)
- Obstacles and landmarks
- Visual appearance (e.g., color, texture) of the crop may change during different growth stages
- Crop being collected might not be uniform (e.g., flattened due to a storm, varying crop heights)
- Vibrations generated from the operating machine

A first step in the detection of a crop edge is to segment the observations into two classes, namely *crop*, and already harvested areas from now on referred to as *ground*.

In the following, prominent approaches on how to detect the actual crop edge and how to perform the segmentation task are presented.

2.1. Height-based segmentation

A couple of researchers such as [14,22,23] used the geometrical characteristics of a field's crop edge. While [22] used a 1D scanning laser rangefinder for automated guidance of agricultural vehicles such as a windrow or a combine harvester, [14] conducted tests with a LIDAR for crop row localization in a row-planted soybean field with three-dimensional field information obtained using a pan-tilt unit.

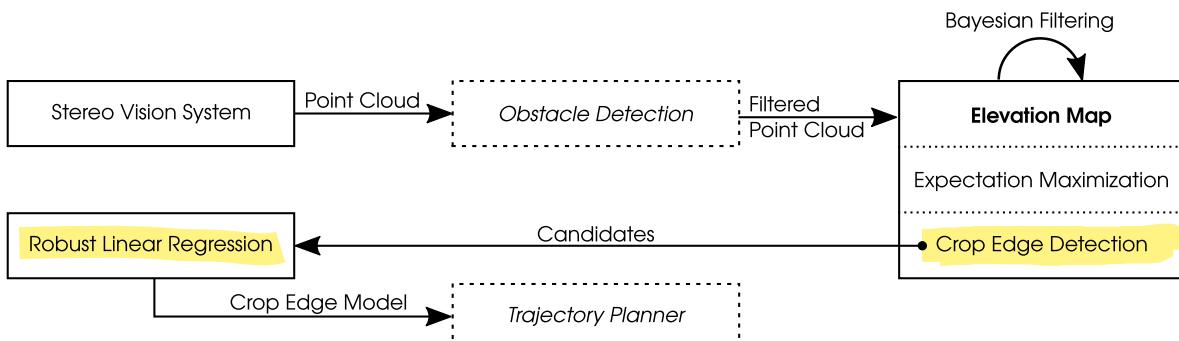


Fig. 2. Schematic overview of the proposed system.

Table 1

Overview of related work since 1996.

Authors	Approach keywords	Sensors	RMSE (cm)	Crop
[9]	Otsu threshold, Least-squares fit	LIDAR, IMU	10.15	Wheat
[10]	Entropy segmentation, Hough transform	RGB-camera	-	Wheat
[11]	HSV color-space segmentation	RGB-camera, GPS	-	Wheat
[12]	Inverse-perspective mapping, Color segmentation	RGB-camera	-	Rice
[13]	Color segmentation, Least-squares fit	RGB-camera	-	Wheat, rice, rape
[14]	Model correlation	LIDAR, IMU, GPS	7	Soybean
[15]	Wavelet transformation, transformation	RGB-camera	-	Wheat
[16]	Gray-level distribution	Monochrome-camera	11.4	Rice
[17]	Density grid, 3D edge detection	Monochrome stereo camera	-	Corn
[18]	Pixel-distribution, Point hough transform	RGB-camera	-	Wheat, corn
[8,19,20]	Adaptive gray level segmentation, Blob analysis, Fuzzy logic	Monochrome-camera	-	Wheat, corn
[21]	Color segmentation (RGB), Shadow compensation, End-of-row detection	RGB-cameras, GPS	5–30	Alfalfa hay, sudangrass
[22]	Model correlation	LIDAR	<10 (max. 100)	Wheat
[23]	Markov-segmentation, Image statistics	Monochrome-camera	<10	Grass
[24,25]	Color segmentation (RGB), Shadow compensation	RGB-camera	5–30	Alfalfa hay, sudangrass

In all these three publications, the crop and ground classes are modeled by a function (a unit step or a Heaviside) which follows the shape profile of a crop edge. After measuring the crop profile (e.g., via LIDAR or stereo camera vision), a point cloud containing height values is obtained. The measured data set is then correlated with the function model using the Pearson correlation coefficient formula at some given delay.

For this methodology to work correctly, the sample size should not be too small and should only contain a few to no outliers. In the case of a bivariate normal distribution of the sample data, each correlation coefficient is a maximum likelihood estimate of the population correlation coefficient and it is impossible to construct a more accurate estimate. However, this correlation method is neither distributionally robust nor robust to outliers, but according to researchers results in sufficient estimates. Nevertheless, areas of flattened or other types of unusual crop occurrences are not handled.

2.2. Color-based segmentation

The second group of works which could be found is based on color segmentation. For all reviewed algorithms to work properly and deliver persuasive performance, the areas of crop and ground in the image have to be descriptively different in color. The harvested ground area usually occurs in soil-related brownish colors, whereas the crop area is differently pronounced by the presence of leaf canopy with its respective color. Robustness to color-based segmentations is increased if the color of the two classes across the harvest area is relatively consistent. Researchers from the Carnegie Mellon University used a New Holland 2550 Speedrower to build the Demeter system [21] for automated navigation in harvesting tasks in sorghum and alfalfa fields. Details on this often-cited system are furthermore published in [24,25]. Here, an RGB and a monochrome camera with 6 band-pass filters in the range of near-ultraviolet to near-infrared was used for an initial examination of various color discriminants. As an example, the best discriminant for alfalfa hay was selected as the ratio R/G, since the within-class standard deviation of crop and ground was sufficiently low and their means revealed a good separation. Finding the best-fit step function of a scan line of the image row with the lowest least-squared error results in a cut-edge candidate at a particular column. This column index is found by means of an exhaustive search (brute force) across all possible value defined by the image width. The error of the fitted step function is computed by calculating the mean of the discriminant values in each line.

2.3. Texture segmentation

A further group of approaches is based on the idea to segment the different textures of crop and ground. Debain et al. [23] have chosen a set of Haralick texture features [26] to distinguish between these two classes to detect the crop edge in grass fields. These features are computed in a grid with predefined cell size and forms the input for an unsupervised segmentation algorithm using Markov fields. The actual crop edge is found by correlating a model function to the rows in a feature image. The results of the individual correlated feature responses are merged with the evidence theory (probability theory).

Another approach presented in [16] is only using the gray-level (GL) values of an image captured with a cabin-mounted monochrome camera. The idea is to utilize the GL values to split the image into two regions of crop and ground. The underlying assumption of the methodology is that brighter pixels (high GL value) represent the crop and darker pixels (low GL value) ground regions. At this point, it should be noted that we found in our research, with respect to having different data sets, that this important assumption does not hold and is not generally valid. In two predefined region-of-interests (ROI), in the lower and upper part of the image, the sum of GL values is computed for each column. Computing the largest gradient in the upper and lower region defines the pixel positions of the cut-edge. The evaluation using 33 images showed that the performance of this algorithm is very poor as only 27% (9 of 33) of the cut-edges have been correctly detected. The researchers stated the failures might be due to shadow near the cut-edge and mechanical vibration of the combine.

Based on the results of these previous publications, we decided to start with a height-based approach for our crop edge guidance system for two important reasons. First, detection based on the geometrical profile is more independent from the crop type as its color and texture do not play a role. Second, methodologies based on 3D data are also less sensitive to illumination difficulties like shadows or daylight changes.

3. Stereo vision system

The stereo system used on our test platform (see Section 8) consists of two stereo cameras. As a result of the investigations of Benson et al. [8], both sensors are mounted to the cabin [8] – one in each frontal outer corner – and provide a broad overview of the harvesting area in front of the combine (see Fig. 3). In their research, the authors considered different mounting positions of cameras for perception during combine guidance, as each position allows for different approaches to capturing crop edge characteristics. Furthermore, they concluded that the best position is above

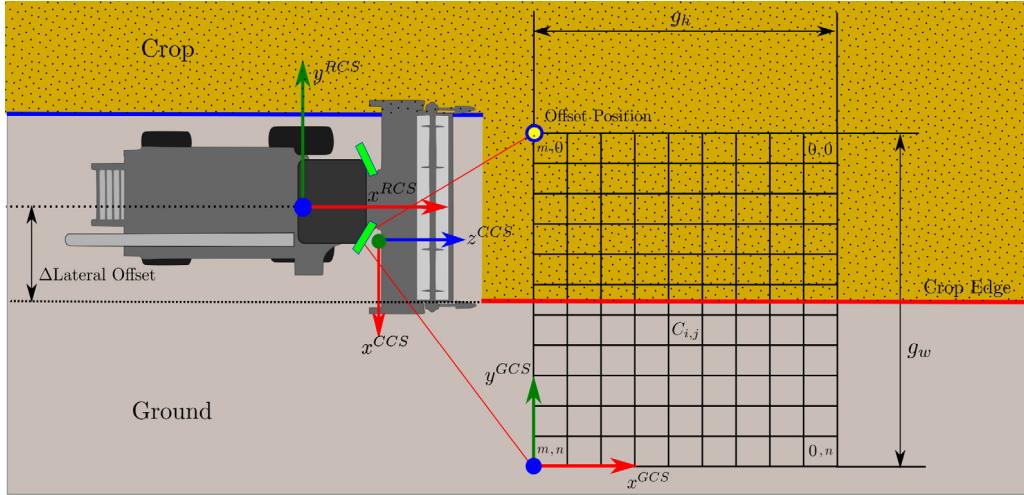


Fig. 3. Sketch of a harvesting scenario in bird's-eye view with the used elevation grid map. The axes of the Robot- (RCS), Camera- (CCS), and the Grid-Coordinate-System (GCS) are color-coded in red (x), green (y) and blue (z). The two slightly outwards rotated and downwards tilted stereo cameras (green) are mounted on the roof of the cabin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the cabin which, for example, is less influenced by vibrations generated from the operating machine.

To reduce the load of the CPU and overcome the difficulty to achieve real-time performance, the disparity map of the stereo system is generated by the *libSGM*¹ library, a CUDA implementation of the Semi-Global Matching (SGM) algorithm [27]. In comparison to the CPU-implementation, a considerably increased frame rate from 10 Hz to 40 Hz with images of size (width \times height) = (640×512) at 64 disparity levels is achieved. We extended the library to allow for a disparity refinement step using quadratic interpolation of the calculated matching costs in the disparity space image (DSI). Using the camera's intrinsic and extrinsic parameters, two point clouds (left and right stereo system) are obtained which are projected to the vehicle-fixed Robot Coordinate System (RCS) ($(x^{RCS}, y^{RCS}, z^{RCS})$; see Fig. 3) using the known position of the cameras on the combine harvester.

4. Height segmentation

Depending on the season, each crop type near maturity stipulates a different plant height. At this time, wheat has an approximate height of 0.5–1 m, rapeseed of 0.3–1.5 m and grass ranges between 0.1 m and 1 m. A generally valid height segmentation is desired to handle different types of crop. Thus, the segmentation algorithm has to be adaptive and to some extent universal/unsupervised. The problem can be formulated as follows:

$$f(h) = \begin{cases} \text{crop}, & \text{if } h \geq T \\ \text{ground}, & \text{otherwise,} \end{cases} \quad (1)$$

where the function f evaluates a height value h and assigns it to a class according to the threshold T . Reviewing the height distribution of multiple point clouds of harvesting scenarios revealed a bi-modal distribution. This property makes it possible to use automatic clustering techniques like Expectation–Maximization (EM) with two Gaussian Models to come up with a threshold T .

A paired *t*-test with samples of each class was used to evaluate the class separation. The difference between the two classes *crop* and *ground* found to be statistically significant ($\alpha = 0.001$) as the evaluation of height estimates of several wheat (see Table 2)

and rapeseed fields revealed. These results indicate that accurate and reliable segmentation is possible.

A drawback of this method is of course that if no observable height difference is found this method fails. Such a case could, for example, be cutting grass in an early growth state where the height difference would only range within a few centimeters. In such a case the classes variance will overlap and make a proper distinction challenging.

5. Elevation grid map

Given the stereo point cloud $P = \{p_1, \dots, p_k\}$ of the scene at time t in-front of the harvester projected to the RCS, the points are sorted into a 2D grid map of size $(g_h \times g_w)$ consisting of cells $C_{i,j}$ as shown in Fig. 4c. Thus, the perceived environment is rasterized into small areas (cells) to be able to approximate the real world and allow for parallel computations. A cell's size is initialized on startup and is parameterized by $(c_h \times c_w)$. For most of the tests, a dimension of $c_w = 10\text{--}15\text{ cm}$ and $c_h = 20\text{--}25\text{ cm}$ was chosen depending on the crop type and the field conditions. This rectangular cell shape was selected because a higher resolution in width is desirable to be more precise on the crop edge position. The grid map frame is positioned in the horizontal (x^{RCS}, y^{RCS}) plane with a predefined offset position $o \in \mathbb{R}^2$ to the grid coordinate system origin. The offset o_x allows to position the grid map in front of combine header and thus skip invalid and unusable measurements in this area. Additionally, the offset o_y is used to limit the area in which the crop edge can occur, independent of the header width. As the output of the stereo matching algorithm depends on the camera's baseline only disparities up to a specific range are useful. For this reason, the grid is limited in the x -direction to be able to cut off the point cloud at a predefined depth x_{\max} . As the crop edge only appears in a short section of the point cloud, likewise the y -direction is limited to a certain length y_{\max} .

In each cell, a height representative $h_{i,j}^{\text{median}}$ of the 3D points that fall into this specific cell is calculated as the median of the points' height. The median has been chosen in preference to the mean since its a more robust measure of the central tendency. As the cells are independent of each other, this step is parallelized. This statistical measure is then stored in a circular buffer of capacity three so that at most the last three height representatives are kept. As agricultural vehicles in harvesting scenarios usually drive

¹ <https://github.com/fiixstars/libSGM>.

Table 2

Characterization of the *crop* and *ground* classes in four sample fields with statistical significance of the separability by height (in the RCS).

Field	Samples	Crop		Ground		t-value	p-value
		Mean height [m]	Deviation	Mean height [m]	Deviation		
A	205	0.58	0.033	0.043	0.054	84.88	<0.001
B	340	0.62	0.027	0.037	0.028	76.74	<0.001
C	427	0.42	0.015	0.046	0.014	70.22	<0.001
D	191	0.67	0.029	0.042	0.025	72.67	<0.001

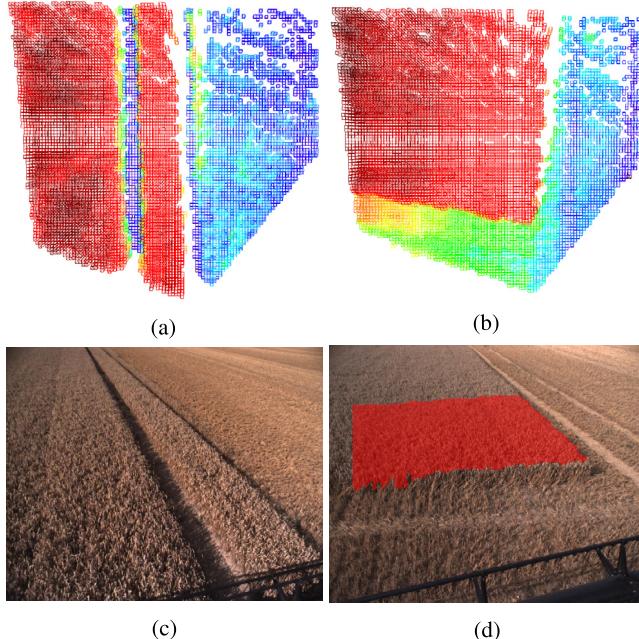


Fig. 4. (a), (b) Bird's-eye view of an elevation map with heights color-coded from low (blue) to high (red). (c), (d) corresponding camera images. (d) Based on (b) the portion of the point cloud with *crop*-membership visualized as an overlay on the stereo camera input image. Due to the baseline of the stereo camera the predefined depth cut off is set to $x_{\max} = 12.5$ m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

at low speed, this is a feasible approach. If at some point in time it is not possible to compute a height representative and a cell's buffer consists of more than two representatives, one is dropped to avoid lag and distortions of the map and the real world. Empty cells (holes) in the grid, mostly caused by occlusions or by the stereo vision system itself as the depth output is only valid up to a specific range, are efficiently closed using the mean height of all heights stored in the buffer. As a result, the grid map models the current elevation of the real world in front of the combine.

5.1. Expectation–maximization with Gaussian mixture model

Based on our observations, heights of the elevation map can be modeled as a mixture of k Gaussians, where $k = 2$. Here, it has to be mentioned, that a preceding filter removes large obstacles like trees, power poles, and other agricultural machines (see [28]) as these objects would falsify this assumption. An algorithm to estimate the parameters of a probability density function of a Gaussian Mixture Model (GMM) is the Expectation–Maximization (EM) algorithm [29]. The assumption is that the set of computed height representatives \mathcal{H} of all cells holding a representative are generated from a set of Gaussians of unknown parameters. EM is utilized to estimate these parameters to determine the mixture model a single height representative $h_i \in \mathcal{H}$

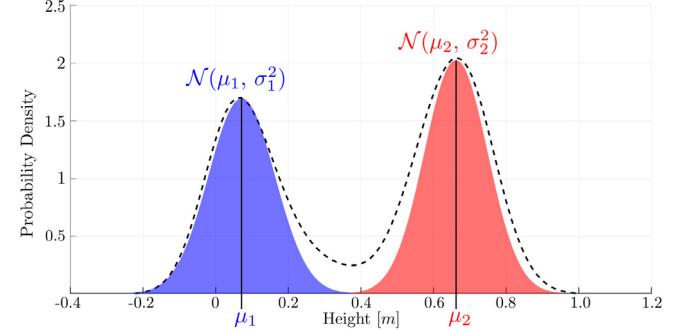


Fig. 5. Example of a probability distribution (black dotted line) of the heights in the elevation map estimated by a Gaussian for each class of *ground* height μ_1 (blue) and *crop* height μ_2 (red). In this scenario, stubbles left from wheat harvest caused a slight shift of the *ground* height from zero to $\mu_1 = 0.082$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

originates from, by maximizing a likelihood function of the variables in the model which are unobserved. These parameters $\theta = (\pi_1, \pi_2, \mu_1, \mu_2, \Sigma_1, \Sigma_2)$ are the weight π_i of the mixture, the mean μ_i and the covariance Σ_i which are estimated iteratively.

Previous tests showed that the Otsu Threshold method [30] could be utilized for this segmentation task, too, but threshold results in ambiguous cases, e.g. with a skewed bi-modal distribution or slightly tri-modal distribution, in contrast to EM was unconvincing. Nevertheless, the system setup allows to change the used segmentation method easily and for the future other segmentation algorithms could be run, too.

The results of EM, apart from the estimated parameters θ , are for each cell with h_i a most probable explicitly assigned Gaussian component label and log-likelihood estimate $\log p(h_i | \theta)$. To accelerate the parameter estimation process, prior parameters θ of the last frame are used to provide an initial estimate for the subsequent calculation. This approach is especially permitted as the environment between two point clouds only changes insignificantly as the combine drives at low speed.

Considering the original problem of segmenting the harvest area into two regions, the resulting μ_1 and μ_2 can be interpreted as approximates of *crop* and *ground* heights, see Fig. 5. Also taking a look at the variance of each component reveals how scattered the height measurements are.

5.2. Bayesian filtering

Similar to occupancy grid maps, each cell in the elevation map stores a binary crop value, which specifies whether it contains *crop* (1) or can be considered *ground* (0). The EM log-likelihood estimates are used in a Bayes filter to continuously update each cell's belief in currently holding points belonging to the *crop* class. We use the sigmoid function (2) to obtain a probabilistic crop map with scaled probabilities in a range between [0, 1]

$$p(C_{i,j} | h_i^{(1:t)}) = 1 - \frac{1}{1 + e^{\log p(h_i^{(1:t)} | \theta)}} , \quad (2)$$

where $h_i^{(1:t)}$ is the set of all height measurements up to time t . Therefore, the segmentation result label $L_{i,j}(b)$ for a cell $C_{i,j}$ can be determined based on a specified threshold level b of belief in crop (3)

$$L_{i,j}(b) = \begin{cases} 1, & \text{if } p(C_{i,j} | h_i^{(1:t)}) \geq b \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

which was set to $b = 0.8$ in our experiments.

6. Crop edge detection

The crop edge is the transition point of standing crop and ground. Therefore, we assume that the *crop* and *ground* classes can be modeled by a model function $g(x)$ in the $(y^{\text{GCS}}, z^{\text{GCS}})$ plane. Such a function follows the profile shape of a perfect crop edge, which is why we have used a unit step function:

$$g(x) = \begin{cases} \text{crop} \Leftrightarrow 1, & x < 0 \\ \text{ground} \Leftrightarrow 0, & x \geq 0 \end{cases} \quad (4)$$

The exact point of transition (edge) between 0 and 1 (at $x = 0$) is known a priori, so that we defined the crop edge to live on a 2D linear subspace of the elevation map at ground level height.

The proposed grid setup allows for a parallelized row-wise processing to detect the crop edge in each row, from now on referred to as scan line, separately. A scan line \vec{S}_m is a $(n+1)$ -dimensional vector consisting of the height segmentation results $L_{m,j}(b)$ of cells in grid row m . Scan lines only containing *crop* or *ground* are excluded from further processing steps since no edge is present and are marked as invalid ($\vec{S}_m = \text{invalid}$). Now, each scan line is correlated with g to locate a possible 2D location of the crop edge, in the future called *Crop Edge Candidate* (CEC).

Following the work of [14,22,23] the correlation coefficient at some given delay d is calculated using (5)

$$r(d) = \frac{\sum_{i=0}^n (g(i-d) - \bar{g})(\vec{S}_{m;i} - \bar{S}_m)}{\sqrt{\sum_{i=0}^n (g(i) - \bar{g})^2} \sqrt{\sum_{i=0}^n (\vec{S}_{m;i} - \bar{S}_m)^2}}, \quad (5)$$

where \bar{g} and \bar{S}_m are the respective means. A series of correlation coefficients is generated by calculating all possible delays $d \in [0, \dots, n]$. All sample points outside the range of the model function are ignored, and their computation is skipped. The argument at which (5) reaches its maximum r_{\max} corresponds to a cell position which possibly contains a crop edge point, see Fig. 6. As the above approach is performed for each scan line, a set E of cells evolves (6)

$$E = \left\{ C_{i,j} \mid i = 0, \dots, m; j = \arg \max_{k \in [0, n]} r(k) \right\}. \quad (6)$$

Up to this point, the computed crop edge location is without regard to the cell resolution inside the grid. Therefore, additionally half of a cell's width needs to be subtracted for a *crop-ground-edge* and added for a *ground-crop-edge*. The set of real world crop edge points is given by (7)

$$\mathcal{P} = \left\{ (x, y) \mid x = E_{\text{center},x}, y = E_{\text{center},y} \pm \frac{c_w}{2} \right\}, \quad (7)$$

where E_{center} is a cell's center of a cell of set E .

7. Linear crop edge model

In the previous step, the set \mathcal{P} containing 2D points in the $(x^{\text{GCS}}, y^{\text{GCS}})$ plane located at the very crop edge position are identified. We made the initial assumption of the crop edge most

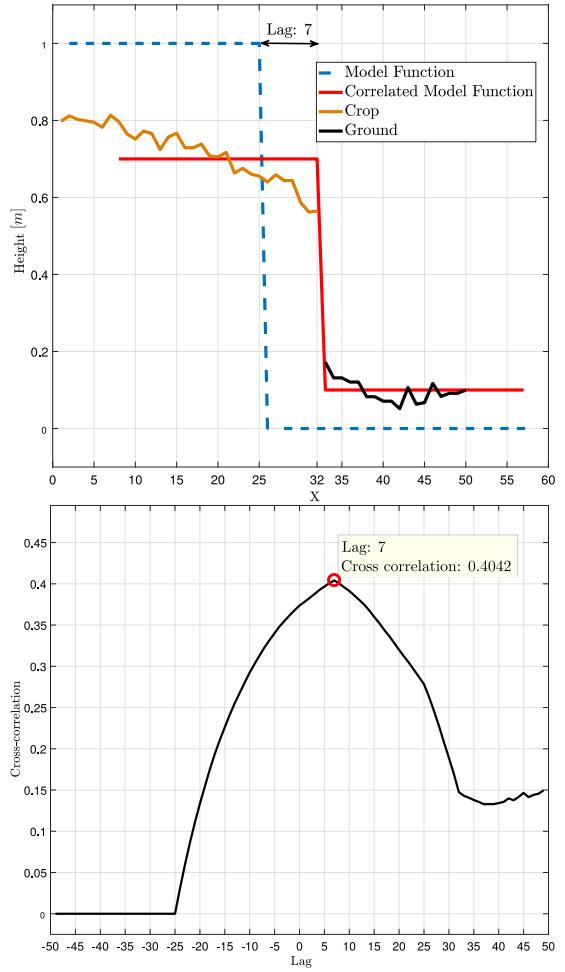


Fig. 6. Exemplary computation of the crop edge position in a segmented height scan line S_0 via cross-correlation. Cell heights of *crop* (orange) and *ground* (black) are correlated with the function model g (dotted blue) which has its known edge position at $x = 25$. The maximum correlation $r_{\max} = 0.4042$ is at a lag of 7, hence cell $C_{0,32}$ contains a CEC. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

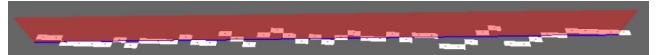


Fig. 7. An exemplary 3D linear model of the crop edge. The crop edge candidates (white rectangles) in each grid row contribute to the model (blue line) fitted via robust linear regression. The 3D edge is formed by adjusting the model with the estimated crop and ground height from the GMM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

likely being of linear nature (see [8,14,23]) and use robust linear regression to model their relationship of spatial distribution (see Fig. 7). Partially flattened crop is not uncommon and can be caused for example by bad weather. With respect to the elevation map, the result can be that an entire scan line only contains *ground*-samples, in which case no CEC would be found, or the CEC may be strongly shifted. Ordinary least-squares (OLS) estimators are too sensitive to such outliers as the example in Fig. 8b shows. Hence, the estimated model would be unstable (low distributional robustness).

Some quantitative results of both fits together with their parameters estimated are given in Table 3. By Robust Fitting the standard error (SE) decreased by -0.0452 and the square of the correlation coefficient (R^2) increased by 39.6%. Very noticeable is

the p -value (probability) of the slope of the robust fit is much greater than $\alpha = 0.01$ for the F-statistic of the hypotheses test equal to zero or not. Since the p -value equals 1.0 this means the slope is even less significant than the 1% significance level given the other intercept term in the model. Not surprisingly, this allows the model to predict only based on the interception term which makes total sense when the robust residual plot is considered.

7.1. Model fitting

To reliably estimate a model and suppress outliers we use the Huber M-Estimator [31]. The set \mathcal{P} is directly used for model fitting with all points $d_i \in \mathcal{P}$, $d_i = (x_i, y_i)$. Based on the predictor values x_i and the measured responses in each grid row, namely the crop edge positions y_i , a predicted response \hat{y}_i of the model is obtained. The robust linear regression model is then given by (8)

$$\hat{y}_i = b + mx_i + \epsilon_i, \quad (8)$$

where parameter b is the line intercept with the y^{GCS} -axis, m the slope of the line and the disturbance term $\epsilon \sim \mathcal{N}(0, \sigma^2)$. The M-estimation is solved by using the iteratively re-weighted least squares (IRLS) algorithm. To compensate the impact of a single estimated model on the combine guidance trajectory and maintain to some extent the knowledge of the crop edge gained in previous models, the exponentially weighted moving average (EWMA) (9) is used

$$\hat{Y}^t = \beta\hat{y}^t + (1 - \beta)\hat{Y}^{t-1}, \quad (9)$$

where β is the degree of weighting decrease.

7.2. Model crop edge filtering

The smoothed model \hat{Y}^t allows for the prediction of the CEC in each grid row at time $t + 1$. One could filter each new CEC based on the model and therefore use the 95% prediction interval (PI) which gives an estimate of where one can expect the next crop edge candidate estimates in 95% of the cases, see Fig. 8a. But our experiments showed that this assumption forces the model to be too strict and does not properly account for unpredictable crop edge occurrences (see Section 8). The scan line of grid row m is therefore only marked as an outlier ($S_m = \text{invalid}$) if its corresponding CEC is not included in the interval (10)

$$[\hat{y}_m \pm \gamma], \quad (10)$$

where γ is a predefined parameter; in our experiments set to $\gamma = 5 \cdot c_w$.

7.3. Model confidence

In order to assess the quality of the currently detected crop edge, we introduced a confidence value $v_{\text{confidence}}$ (11)

$$v_{\text{confidence}} = w_0 v_{\text{model_compliance}} + w_1 v_{\text{correlation}} + w_2 v_{\text{valid}} \quad (11)$$

which is a weighted linear combination of three separate confidence values (12)–(14) within the range $[0, 1]$. The three associated parameters w_0, w_1, w_2 allow to adjust the confidence preferences. In our experiments, the weights were set to $w_0 = 0.4, w_1 = 0.4, w_2 = 0.2$. The first confidence value $v_{\text{model_compliance}}$ (12) gives an indication of how well the newly found CECs correspond with the latest established crop edge model

$$v_{\text{model_compliance}} = 1 - \frac{\sum_{i=0}^{s-1} |\hat{Y}_{i,y}^{t-1} - \mathcal{P}_{i,y}|}{\gamma s}, \quad (12)$$

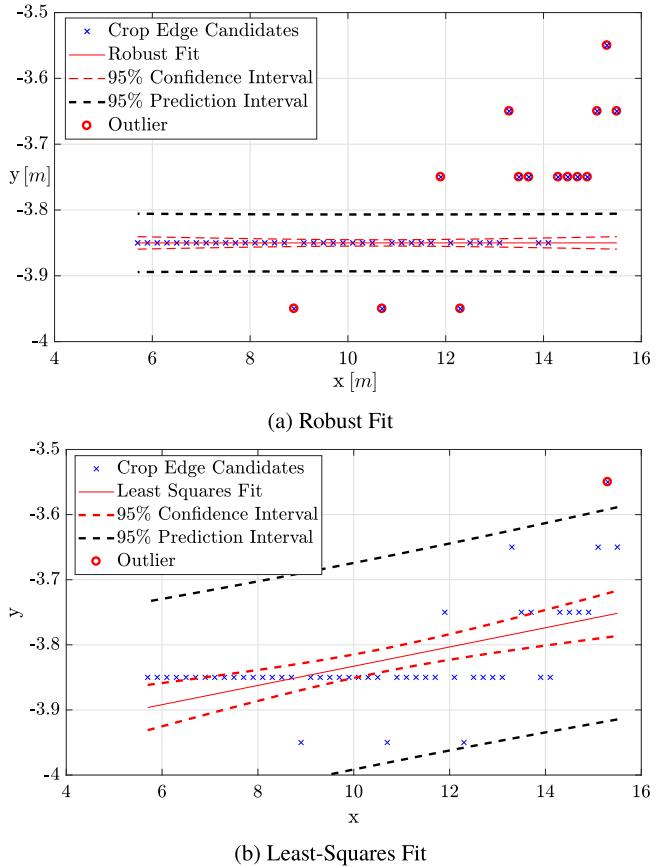


Fig. 8. An exemplary robust linear crop edge model with a length of 10 m. Out of the 50 CECs (plotted in RCS) found, the subset of 36 CECs (72%) contribute to the new crop edge model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where $s = |\{x \in \{0, \dots, m\} | S_x = \text{valid}\}|$ is the number of valid scan lines and $\mathcal{P}_{i,y}$ the y -value of a respective crop edge point in a scan line i . The second value $v_{\text{correlation}}$ (13) is an overall measure of how well the model function was correlated in each scan line.

$$v_{\text{correlation}} = \frac{\sum_{i=0}^{s-1} r_{\max,i}}{s} \quad (13)$$

The last indicator v_{valid} (14) is a simple ratio of the number of valid scan lines and the total number of scan lines

$$v_{\text{valid}} = \frac{s}{m+1} \quad (14)$$

The combined confidence value could be used as feedback for the combine operator or further influence the subsequent guidance system.

7.4. Start and end of crop detection

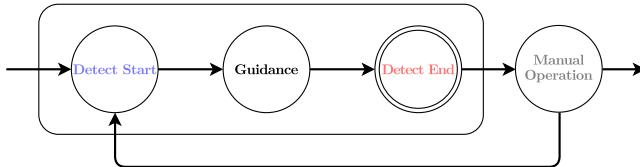
For the guidance system, it is a pleasant feature to actually know where the crop edge starts and ends. The start detection allows the guidance system to steer the harvester towards the crop edge and begin the harvest process from thereon. On the other hand, the detection of the end of crop allows for a smooth transition to subsequent processes such as performing a turning maneuver or lowering the speed and just handing the control back to the harvester operator.

On system start-up, the initial state is set to *Detect Start*. In this state, the starting point of the crop edge is searched for.

Table 3

Quantitative results of the two fitting methods shown in Fig. 8—the **Difference** row shows the actual benefits of the Robust Fitting method (**SE** = standard error, **R**² = coefficient of determination).

Fitting method	Intercept (b)	p-value	Slope (m)	p-value	SE	R ²
Least-squares	-3.9805	5.62e-61	0.0148	1.34e-05	0.0621	0.315
Robust	-3.85	2.24e-87	0.0	1.0	0.0169	0.711
Difference	+0.1305		-0.0148		-0.0452	+0.396

**Fig. 9.** State machine of the start and end of crop detection process.

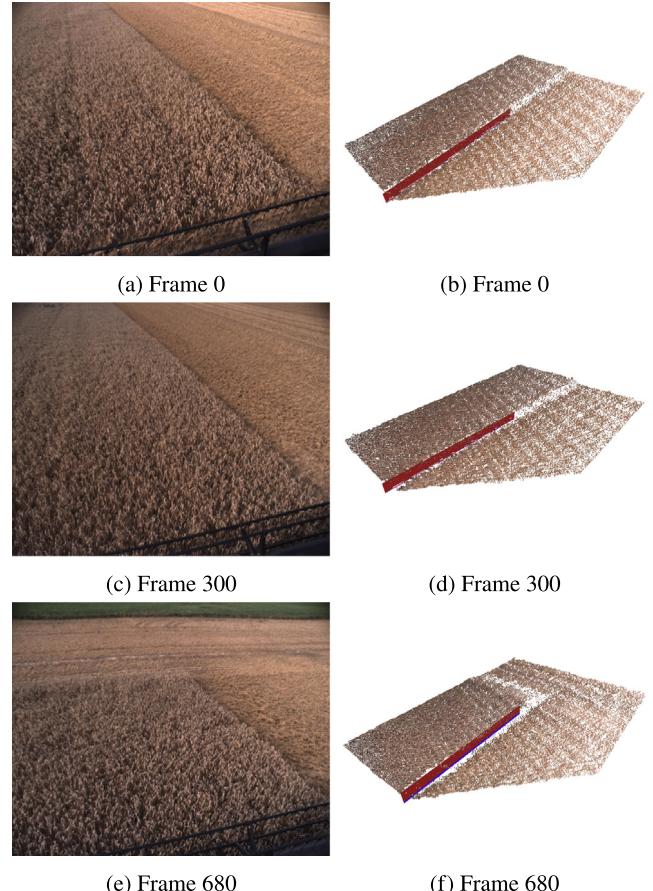
For this reason, in each row of the elevation map, beginning from the harvester, the ratio of ground classified cells and the maximum number of cells in this row is computed. If the amount of ground labeled cells exceeds a predefined threshold (e.g., 90%), the row is considered *empty*. The row with the crop edge starting point is eventually found when a subsequent *non-empty*-row of *empty*-rows is detected. At this point the system state changes to *Guidance* which executes the actual crop edge guidance. At some point in time, the harvester will reach the end of the field or therefore also the end of the crop. Since the crop end is exactly the opposite direction of detection of the start, the same technique is used. In this case, cells considered empty after non-empty-rows are taken into account so that if no subsequent rows filled with crop are found, the guidance system is disabled and control is handed over to the operator. A state machine of this process is shown in Fig. 9.

8. Experiments and results

A John Deere Combine equipped with an Intel Core i7-4790K@4.0 GHz CPU was used as a test platform. Two stereo camera systems mounted to the sides of the cabin – slightly shifted and angled towards each side of the head – are used to capture the images of 640×512 pixels. With the proposed system implemented in the C++ Real-time Robotic Framework FINROC [32], multiple experiments have been conducted. In preparatory tests, logged data of actual harvesting in wheat and rapeseed fields in summer has been used. For the conclusive experiments, the combine harvester was driven with an RTK-DGPS guidance system on parallel trajectories defined by the header width. To evaluate the performance and quality of our system, the crop edge models together with the point cloud were logged. Later on, they have been first visually and then geometrically examined by comparing them in three-dimensional space.

During our tests, the elevation map was parameterized as follows: $g_h = 50$, $g_w = 75$, $c_h = 0.25$ m and $c_w = 0.1$ m. The cell width c_w needed to be chosen with care because the disparities at the very crop edge at plant height are blurred due to stalks and ears sticking out. The chosen value was sufficient enough to capture the crop edge properties and still allowed for precise detection of the edge location. In the following, 3 experiments of 3 different fields are presented which demonstrate the performance of the system and show its applicability for real-world scenarios.

The first data set shows the results of a 452.5 m run of harvesting wheat at an average speed of 7 km/h with a mounted header of 8 m in width, containing a series of 700 stereo image pairs. The crop was threshed on a sunny day, with dry soil but

**Fig. 10.** (a), (c), (e) Selected frames of a recording while harvesting wheat showing the start, center, and end of the field together with the detected crop edges (b), (d), (f).

no dust turbulences. Furthermore, no obstacles like little bushes or irregularities such as flattened crop are contained. At the beginning of the series, the combine's header is located right at the crop edge (Fig. 10a), ready for harvesting. After manually driving the harvester along the crop edge (Fig. 10c), the standing crop itself ends approx. around frame 680 (Fig. 10e).

The trajectory generated by the direction vectors of the detected crop edge is visualized in Fig. 11a. As shown in the enlarged area from frame 190 to 200, the trajectory itself is shown as a blue line with the corresponding direction vector as a red arrow. The trajectory is almost a straight line since the crop edge was more or less straight, too. Noticeable, at the field end in frame 680 (see Fig. 10e) the trajectory starts curving, which is due to missing crop edge candidates behind standing crop at the end of the field. The procedure described in Section 7.4 detected the field end in frame 685, which would then allow to stop the guidance system and hand the control back over to the combine operator. Based on the direction vectors drawn in Fig. 11a, the plot depicted in Fig. 11b shows the heading angle with reference to a GPS guidance system driving along the crop edge. Furthermore, the

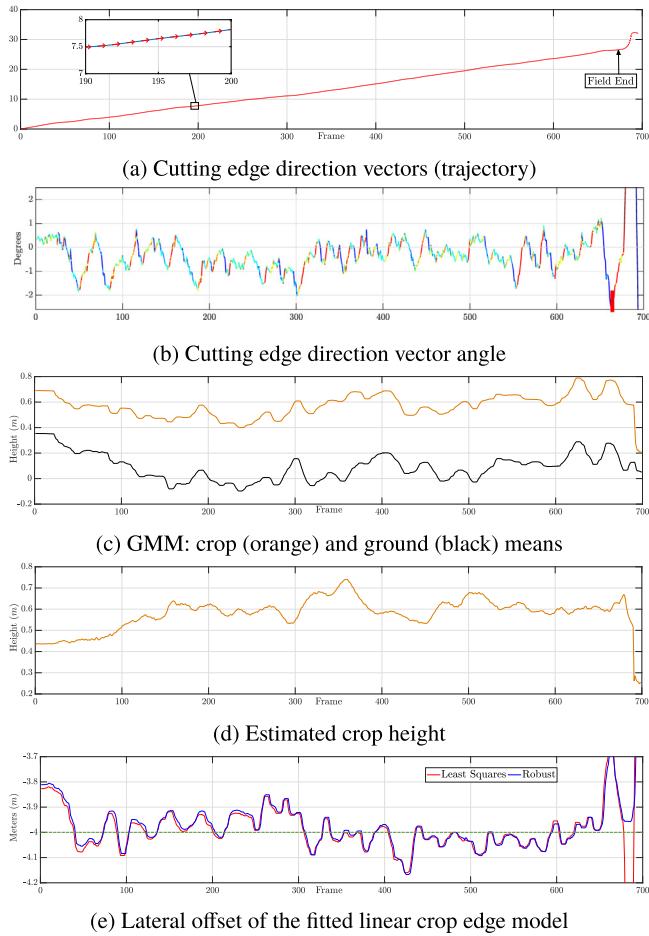


Fig. 11. Evaluation of the data set shown in Fig. 10. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

color-coding from green (low) to red (high) shows the gradient between two consecutive direction vectors. The height of the crop and ground class is estimated using EM with a Gaussian Mixture Model. The class means which have been determined by the algorithm for the recording are plotted in Fig. 11c, where the orange line shows the crop and the black line the ground heights. It should be noted, that this figure visualizes the estimated class height in the area covered by the elevation map and any unevenness on the ground and on a softer or harder surface directly influences these estimates. Just as in Fig. 11c, the actual overall crop height can be estimated as the difference of crop and ground height which is demonstrated in Fig. 11d. In order to have a more realistic cell-wise estimate of the crop height, the ground height can be subtracted from the cell's median height representative. Around frame 680 the sudden drop of the crop height indicates the missing crop since the field end is reached.

As the last analysis, the lateral offset (red, blue) of kinematic center (RCS) and crop edge model located at the header divider is shown in Fig. 11e for both fitting methods. Here, it has to be taken into consideration, that the header dividers have an offset of approx. ± 4 m which is the reference (green) in this diagram. In summary, the RMSE of the heading angle amounts to 0.94° and for the lateral offset 0.09 m in comparison to track driven by the DGPS-based guidance system (with RTK) used to record the data for evaluation.

The second experiment has also been conducted on a wheat field. The data set which has been recorded contains a series

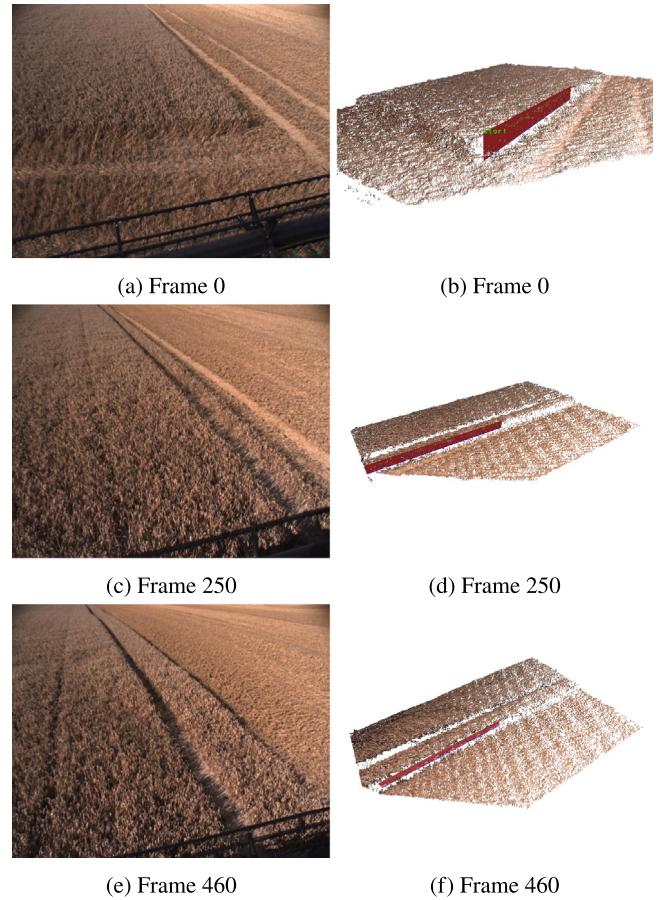


Fig. 12. (a), (c), (e) Selected frames of the 2nd field evaluation showing the start, center, and end of the field together with the detected crop edges (b), (d), (f).

of 480 stereo image pairs. Just as in first recording, the wheat was harvested during a sunny day, with dry soil and almost no dust. As irregularities, small bushes and tramlines are included. These tramlines add additional crop edges, in this case very close to the edge of interest, which makes the detection extremely challenging. Another special feature of this data is that it starts with an offset from the actual crop edge. As shown in Fig. 12a, the combine is initially located a couple of meters in front of the crop. Therefore, the combine first needs to be steered towards the crop edge to start harvesting. This process is visible in the trajectory plot in Fig. 13a within the first 50 frames. As already mentioned, the occurrence of the tramline, which introduces a new crop edge in the middle of the recording, is also visible on the graph. The quantitative results of this experiment using the proposed height-based crop edge detection can be found in Fig. 13. Furthermore, 3 examples of detected crop edges are drawn in the perceived point cloud together with their respective input frames in Fig. 12. The Start and End of Crop Detection introduced in Section 7.4 has been tested, and the start of the crop was successfully determined (see Fig. 12b – marked with the green start tag).

At the beginning of the test drive, the resulting trajectory depicted in Fig. 13a is slightly curved as the combine first has to drive towards the crop edge. Around frame 230, a small crop stripe next to a tramline starts to broaden, which introduces a potentially new crop edge. In frame 260 the algorithm adjusts itself to follow the new crop edge. Fig. 13b presents the crop edge model heading angle in degrees with reference with a DGPS-guidance system (with RTK) driving along the crop edge. The

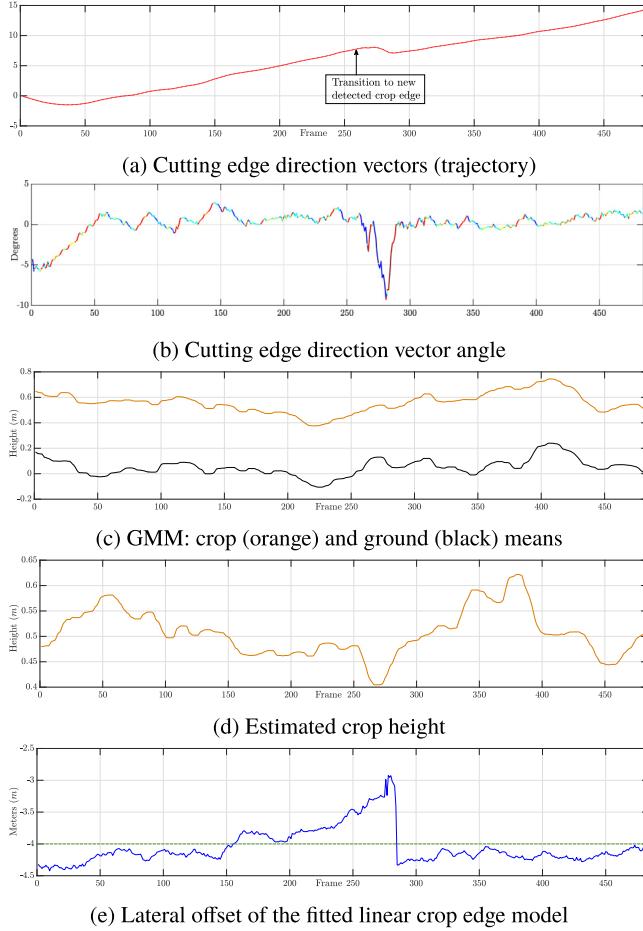


Fig. 13. Experimental results of the data set shown in Fig. 12. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

color-coding follows the same scheme that was introduced for the first experiment. In the first 50 frames, as already mentioned, one can observe the steering towards the crop edge, which is fully covered starting at frame 50 where the angle is close to 0. Also quite noticeable is the amplitude between frame 260 and 290, where the algorithm executes the transition towards the new crop edge. The detected crop edge introduces a negative direction angle, thus, steering towards the outer right side is performed. Analogous to the first evaluation, the class means μ_1 and μ_2 estimated for ground and crop using EM with a Gaussian Mixture Model is shown for this field in Fig. 13c.

Similarly, the overall crop height determined in the elevation map is given in Fig. 13d and shows a small variance of the plants' height. Finally, the estimated offset regarding the ground truth (generated using market-leading GNSS guidance system) is given in Fig. 13e. The transition to the new crop edge is quite visible around frame 280 as a sudden offset drop by 1.3 m.

The 3rd example covers another crop type, namely rapeseed. Compared to wheat or grain fields, the crop edge was much more challenging to detect as the rapeseed plant has a more bushy shape. Additionally, the field did not only contain the crop itself but bigger green bushes and due to dry soil also dust turbulences.

Fig. 14a shows a representative intermediate stage of the rapeseed field in the late hours of the day. Based on this input, the approach was also able to accurately detect the crop edge as shown in Fig. 14b. This detection capability is also proven by the

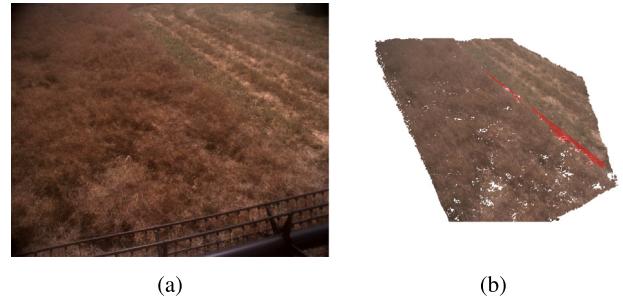


Fig. 14. (a) Example frame captured while harvesting a rapeseed field, (b) detected crop edge for this image.

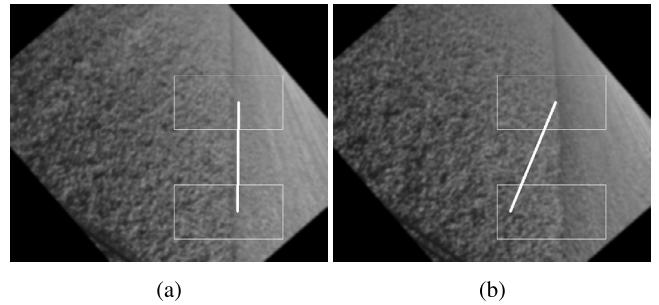


Fig. 15. Adaptation of the approach detailed in [16] to the introduced camera setup: (a) best detection according to the evaluation set (frame 190), (b) worst estimation of the crop edge (frame 410).

video stills which are summarized in Fig. 16 where the detected transition between the standing plants and the harvested ground is highlighted as a red line which is a back-projection of the 2D crop edge model to image coordinates.

In the last evaluation, the quality of the proposed detection is compared to hand-labeled ground truth and the approach presented in [16]. For that purpose, the proceeding given in [16] had to be adapted to get meaningful results for the introduced camera setup and type. First, the technique assumes that the sensor is mounted close to the crop edge which appears vertically in the image. To achieve this, the image was rotated in a first step as shown in Fig. 15. Furthermore, the camera used for our investigations has color capabilities and exhibits much more details compared to the device used by Iida et al. Thus, the images have been converted to grayscale and blurred with a Gaussian blur filter to remove local minima. Besides, the ROIs had to be adapted to the camera resolution and the previous rotation as shown in Fig. 15. Finally, also the selection of the gradient which is supposed to model the transition from crop to ground had to be modified as the proposed calculation was often running into local minima at the standing crop. Here, the rightmost of the 5 largest gradients was selected instead. The extension described by Iida et al. could unfortunately not be realized as no individual stubble lines are visible in the recorded data (compare Fig. 12).

In total, 61 frames of a data recording out of a 417 m run have been labeled in a way that the cutting edge was approximated as a line by a human. Examples of such ground truth lines are shown in Fig. 20 where they are visualized in green. For comparison of both the algorithm proposed in this paper and the in [16], the outputs were compared to the hand-labeled data in the following way: First, the direction deviation of the crop edge lines was determined by calculating the enclosed angle. The result of this evaluation is shown in Fig. 17 where our method is shown in



Fig. 16. Detection results in a rapeseed field where the red line depicts the crop edge detected based on the stereo data – every 30th frame is shown.

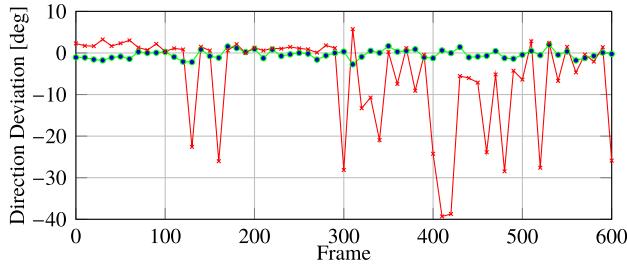


Fig. 17. Deviation of the crop edge direction compared to the ground truth labels: our method (green) and [16] (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

green and direction deviation computed based on the output of lida's algorithm is denoted in red. It can be seen from the graph that the proposed approach is very close to the ground truth with an average error of 0.87 deg and a maximum deviation of -2.74 deg. It is furthermore visible, that this deviation is almost equally distributed with a slight tendency towards the harvested area. In contrast, the procedure published in [16] shows huge outliers with a large average deviation of 7.34 deg for the crop edge direction.

As a second metric, the distance of the cutting edge determined by both methods was compared to the ground truth by computing the distances between both lines at 3 different positions. These different distances are visualized by blue lines in Fig. 20 and are in average 6, 8, and 11.5 m in front of the camera (varies with ground profile and pitch of the machine).

A detailed analysis of this investigation is given in Figs. 18 and 19. In both Figures, the blue plots are representing the estimated offsets at a distance of 6 m, the red ones at 8 m, and green color is used for the distance farthest away from the camera (11.5 m). Looking at the diagram, it can be seen that the introduced method delivers results which are very close to the human-labeled data. The average absolute errors are only 0.070 m, 0.067 m, and 0.077 m at the 3 considered positions. Contrary to this finding, the deviations considering the spacial shift of lida's technique are considerably higher. Here, the average values are 0.478 m, 0.382 m, and 0.347 m with peaks of up to 1.479 m. Moreover, a general tendency to place the crop edge to the already harvested area is visible. To get a more intuitive impression of these results, the best and worst camera frame of

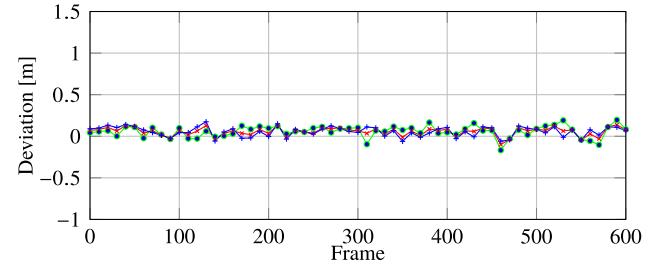


Fig. 18. Spatial deviation of the detected crop edge compared to human-labeled data of the proposed approach at 3 different positions: 6 (blue), 8 (red), and 11.5 m (green) in front of the camera. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

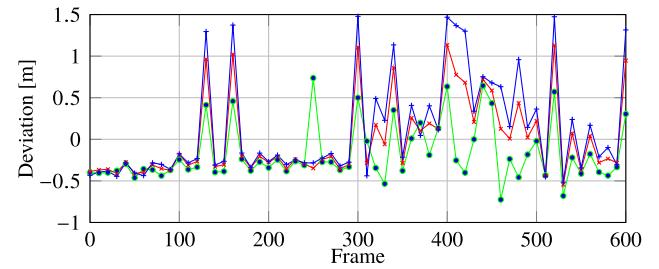


Fig. 19. Spatial deviation of the detected crop edge determined for the procedure presented in [16] (see Fig. 18 for details on colors used).

each approach according to both evaluation criteria is printed in Fig. 20.

9. Conclusions and outlook

In this paper, we have presented a robust crop edge detection approach to allow for automated guidance of combine harvesters equipped with stereo camera(s) as perception sensor(s). The approach utilizes the 3D characteristics of the transition between the crop and the stubbles/ground. With the help of an elevation map, the 3D point cloud is partitioned into scan lines allowing for parallelized processing. This continuously updated grid is used to calculate the median height in each cell where each row in this map represents a height scan line with some offset to the

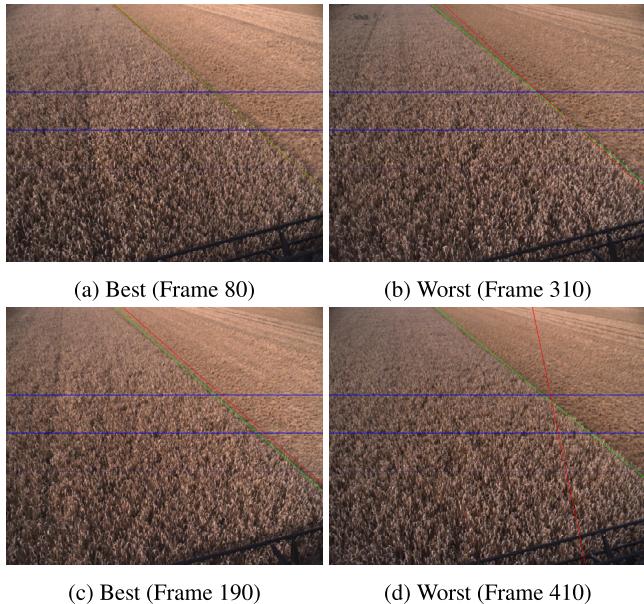


Fig. 20. Best and worst detection result of the proposed methodology (a, b) and the adapted approach described in [16] (c, d) for the hand-labeled data set: ground truth (green), detection (red), positions used to evaluate the offsets (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

harvester. A fast height segmentation was achieved by a frame-wise iterative Expectation–Maximization of a Gaussian Mixture Model. To come up with an overall crop edge model for the current input point cloud, a linear model is fit to the crop edge positions via robust linear regression.

Multiple experiments in different fields have shown, that the proposed system is capable to detect the crop edge in wheat and rapeseed fields accurately. Furthermore, these tests also included severe environmental conditions like dust, herbs, and challenging lighting conditions.

In the future, we plan to fuse our height-based approach with a texture – and color-based crop edge detection to further improve the guidance precision – for example in cases of flattened crop or grass fields where the height difference between ground and crop can be minimal.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] N. Alexandratos, J. Bruinsma, et al., World agriculture towards 2030/2050: the 2012 revision, Tech. rep, ESA Working paper FAO, Rome, 2012.
- [2] Q. Zhang, F.J. Pierce, Agricultural Automation: Fundamentals and Practices, CRC Press, 2016.
- [3] M. Kassler, Agricultural automation in the new millennium, *Comput. Electron. Agric.* 30 (1–3) (2001) 237–240.
- [4] R. Fernández, H. Montes, C. Salinas, Vis-nir, swir and lwr imagery for estimation of ground bearing capacity, *Sensors* 15 (6) (2015) 13994–14015.
- [5] K. Fitzpatrick, D. Pahnos, W.V. Pype, Robot windrower is first unmanned harvester, *Ind. Robot.: Int. J.* 24 (5) (1997) 342–348.
- [6] M. Pérez Ruiz, S. Upadhyaya, GNSS in Precision Agricultural Operations, Intech, 2012.
- [7] A. English, P. Ross, D. Ball, P. Corke, Vision based guidance for robot navigation in agriculture, in: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE, 2014, pp. 1693–1698.
- [8] E. Benson, J. Reid, Q. Zhang, Machine vision-based guidance system for agricultural grain harvesters using cut-edge detection, *Biosyst. Eng.* 86 (4) (2003) 389–398.
- [9] Z. Teng, N. Noguchi, Y. Liangliang, K. Ishii, C. Jun, Development of uncut crop edge detection system based on laser rangefinder for combine harvesters, *Int. J. Agric. Biol. Eng.* 9 (2) (2016) 21–28.
- [10] G. Wu, Y. Tan, Y. Zheng, S. Wang, Walking goal line detection based on dm6437 on harvesting robot, in: *International Conference on Computer and Computing Technologies in Agriculture*, Springer, 2011, pp. 351–361.
- [11] M.Z. Ahmad, A. Akhtar, A.Q. Khan, A.A. Khan, Simplified vision based automatic navigation for wheat harvesting in low income economies, 2015, CoRR abs/1501.02376.
- [12] W. Cho, M. Iida, M. Suguri, R. Masuda, H. Kurita, Using multiple sensors to detect uncut crop edges for autonomous guidance systems of head-feeding combine harvesters, *Eng. Agric. Environ. Food* 7 (3) (2014) 115–121.
- [13] Z. Tian, X. Junfang, W. Gang, Z. Jianbo, Automatic navigation path detection method for tillage machines working on high crop stubble fields based on machine vision, *Int. J. Agric. Biol. Eng.* 7 (4) (2014) 29.
- [14] J. Choi, X. Yin, L. Yang, N. Noguchi, Development of a laser scanner-based navigation system for a combine harvester, *Eng. Agric. Environ. Food* 7 (1) (2014) 7–13.
- [15] Y. Ding, D. Chen, S. Wang, The mature wheat cut and uncut edge detection method based on wavelet image rotation and projection, *Afr. J. Agric. Res.* 6 (11) (2011) 2609–2616.
- [16] M. Iida, Y. Ikemura, M. Suguri, R. Masuda, Cut-edge and stubble detection for auto-steering system of combine harvester using machine vision, *IFAC Proc. Vol.* 43 (26) (2010) 145–150.
- [17] F. Rovira-Más, S. Han, J. Wei, J.F. Reid, Autonomous guidance of a corn harvester using stereo vision, *Agric. Eng. Int.: CIGR J.* (2007).
- [18] Z. Lei, W.S. Mao, C.B. Qi, Z.H. Xia, Crop-edge detection based on machine vision, *N. Z. J. Agric. Res.* 50 (5) (2007) 1367–1374.
- [19] E. Benson, J. Reid, Q. Zhang, Machine vision-based guidance system for an agricultural small-grain harvester, *Trans. ASAE* 46 (4) (2003) 1255.
- [20] E. Benson, J. Reid, Q. Zhang, et al., Development of an automated combine guidance system, *Dev. Automat. Comb. Guid. Syst.* (2000) 1–11.
- [21] T. Pilarski, M. Happold, H. Pangels, M. Ollis, K. Fitzpatrick, A. Stentz, The demeter system for automated harvesting, *Auton. Robots* 13 (1) (2002) 9–20.
- [22] T. Chateau, C. Debain, F. Collange, L. Trassoudaine, J. Alizon, Automatic guidance of agricultural vehicles using a laser sensor, *Comput. Electron. Agric.* 28 (3) (2000) 243–257.
- [23] C. Debain, T. Chateau, M. Berducat, P. Martinet, P. Bonton, A guidance-assistance system for agricultural vehicles, *Comput. Electron. Agric.* 25 (1–2) (2000) 29–51.
- [24] M. Ollis, A. Stentz, Vision-based perception for an automated harvester, in: *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications*, in: *IROS'97*, vol. 3, IEEE, 1997, pp. 1838–1844.
- [25] M. Ollis, A. Stentz, First results in vision-based crop line tracking, in: *Proceedings of IEEE International Conference on Robotics and Automation*, Vol. 1, IEEE, 1996, pp. 951–956.
- [26] R.M. Haralick, K. Shanmugam, et al., Textural features for image classification, *IEEE Trans. Syst. Man Cybern.* 3 (6) (1973) 610–621.
- [27] H. Hirschmüller, Stereo processing by semiglobal matching and mutual information, *IEEE Trans. Pattern Anal. Mach. Intell. (ISSN: 0162-8828)* 30 (2) (2008) 328–341.
- [28] P. Fleischmann, K. Berns, A stereo vision based obstacle detection system for agricultural applications, in: D.S. Wettergreen, T.D. Barfoot (Eds.), *Field and Service Robotics: Results of the 10th International Conference*, Springer International Publishing, Cham, 2016, pp. 217–231.
- [29] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the em algorithm, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 39 (1) (1977) 1–38.
- [30] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man Cybern.* 9 (1) (1979) 62–66.
- [31] P.J. Huber, Robust Statistics, Wiley Series in Probability and Statistics, Wiley, 2005.
- [32] M. Reichardt, T. Föhst, K. Berns, Introducing FINROC: A Convenient Real-time Framework for Robotics based on a Systematic Design Approach, Technical report, Robotics Research Lab, Department of Computer Science, University of Kaiserslautern, Kaiserslautern, Germany, 2012.



Johannes Kneip received his M.Sc. in computer science from the Technische Universität Kaiserslautern in 2017. Since then, he works as a developer at the Continental AG, A.D.C GmbH in Lindau. During his studies, his field of interest consisted of autonomous mobile robotics with respect to ADAS, image processing, mapping, and navigation.



Patrick Fleischmann received his Dipl.-Technoinform. degree from the Technische Universität Kaiserslautern in 2010. Since the end of 2012, he is employed as a research assistant at the Robotics Research Lab in Kaiserslautern. His research interests lie in the area of robust sensor data and image processing for environmental modeling and recognition. Furthermore, he is developing ADAS for agricultural applications.



Prof. Dr. Karsten Berns received his Ph.D. from the University of Karlsruhe in 1994. Since 2003 he is a full professor for robotic systems at the Technische Universität Kaiserslautern. Present research activities are in the area of autonomous mobile robots and humanoid robots with a strong focus on control system architectures and behavior-based control.