LANDTECHNIK 73(1), 2018, 1–9
DOI:10.15150/lt.2018.3176

LANDTECHNIK
AGRICULTURAL ENGINEERING

Object recognition and position determination in agriculture using the example of a coupling assistant

Tobias Blume, Ilja Stasewitsch, Jan Schattenberg, Ludger Frerichs

The increasing performance of modern tractors increases the demands on the operator. In order to further support the operator, the development of new assistance systems is necessary. For this reason, the Institute of Mobile Machines and Commercial Vehicles (IMN) at the Technische Universität Braunschweig is conducting research into assistance systems based on optical sensors that increase ease of use, quality of work and safety. These assistance systems are being investigated in a publicly funded research project together with AGCO Fendt as an associated partner.

Keywords
Image recognition, pose estimation, assistance system, sensor system

The rapid development in the field of optical 3D sensors has caused that three-dimensional data can be acquired and calculated cost-effectively and in real time (Aldoma et al. 2012). With the help of this 3D data, subsystems can be automated that previously could only be executed by humans themselves. Within the scope of a research project at the Technische Universität Braunschweig, five different assistance functions for mobile machines were investigated with a 3D sensor. For this purpose, a stereo camera was mounted on the rear part of the cab roof of a tractor in order to obtain a three-dimensional image of the rear space. With the help of this data the following functions were realized:

- a detection of attachments for automatic parameterization of the machine
- a collision assistant to increase safety for man and machine
- an assistant for automatic coupling to attachments
- a passive lane guidance assistant for mounted implements
- a manoeuvering assistant for trailers with single and multiple set of axles

These assistance systems have in common that an object has to be classified and/or its position in the room has to be determined. With these two pieces of information, the above-mentioned applications can be implemented. In the following, the coupling assistant is used as an example to describe how the detection of the attachments and path control work.

received 18 October 2017 | accepted 12 February 2018| published 21 February 2018

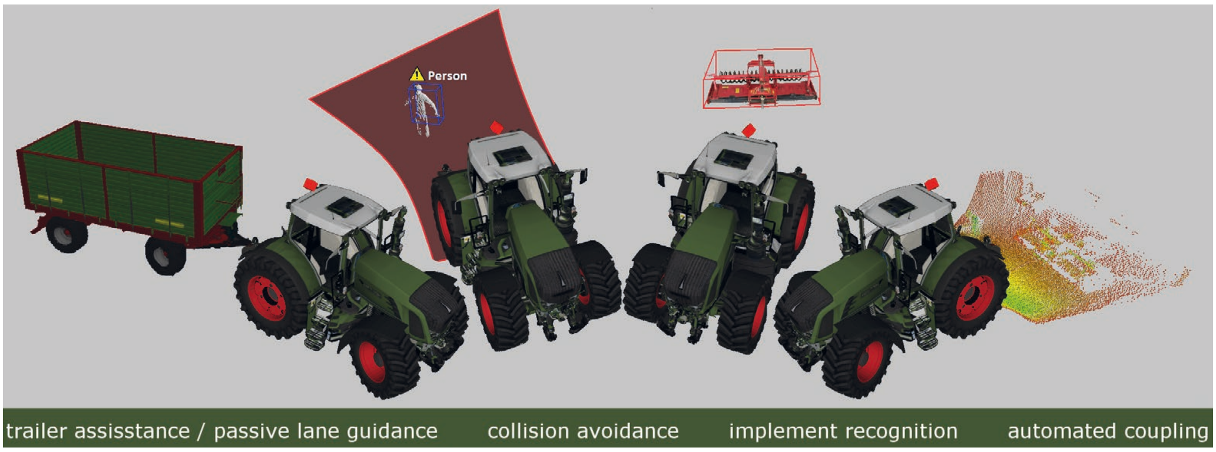

Figure 1: Schematic representation of the assistance functions

System description

For the development of the assistance systems, a FENDT 724 was equipped with a stereo camera (Multisense S21). This sensor is directly connected to a computing unit (Lenovo Thinkpad W540) for testing purposes. The motion commands (steering angle and speed) are generated from the optical data and transmitted to the tractor by an Ethernet CAN adapter. For steering, the signals of the VarioGuide tractor tracking system are imitated, for speed control the signals of the multifunction joystick. The power supply of the systems is decoupled from the tractor and located in a separate control box on the tractor. Figure 2 shows the test setup for automated coupling to a rotary harrow. The implement shown in the picture must be recognized (classified) in the first step. In the second step, the position in the three-dimensional space is calculated. With the help of these information, a path to the implement can be calculated and the tractor can be guided via the path to the implement by a path controller.

Figure 2: Test setup for coupling to implements (blue: implement, green: test vehicle, yellow: sensor, red: power supply)

The Robot Operating System (ROS) forms the basis for the research of these applications. This open source framework provides functions for hardware abstraction, various device drivers and standardized message formats. Furthermore, ROS regulates the exchange of messages between individual programs, which are also called „nodes". Each node has the possibility to provide or receive information by means of subscribers and publishers based on the TCP/IP protocol. This offers a high degree of modularity during development because individual functions can easily be replaced, extended or transferred to other computers.

In addition to ROS, the open source software GAZEBO as an independent simulation environment is another important component within the project. The created virtual vehicle is integrated into GAZEBO in a XML file. In this file the masses, moments of inertia and friction coefficients are defined. The vehicles can also be equipped with virtual sensors and actuators using plug-ins programmed in Python or C/C++. The vehicles are then loaded into the simulation environment and controlled by the aforementioned publishers and subscribers.

Classification and positioning

Classifying and detecting the position of implements is the biggest challenge in the research project described here, but also necessary for the development of the assistance systems. In the following, three different approaches for classifying objects and pose estimation are described. The first two approaches are based on the processing of 3D point clouds with the software library Point Cloud Library (PCL). The third approach uses a neural network to process the stereoscopic camera's 2D images.

For classification with PCL, descriptors are used to extract characteristics from the sensor data, which can also be uniquely identified at different viewing angles, sensor noise and resolutions. In the PCL there is a wide range of different descriptors that can be divided into two groups: global and local descriptors. Global descriptors describe an object as a whole, which makes it necessary to segment the objects from the sensor data beforehand. Local descriptors, on the other hand, extract characteristics that only describe a small part of an object. Segmentation is not necessary here, but several local characteristics must be in a certain geometric relation to each other in order to correspond to an object. As a result of these differences, there are two ways for the program sequence (Figure 3).

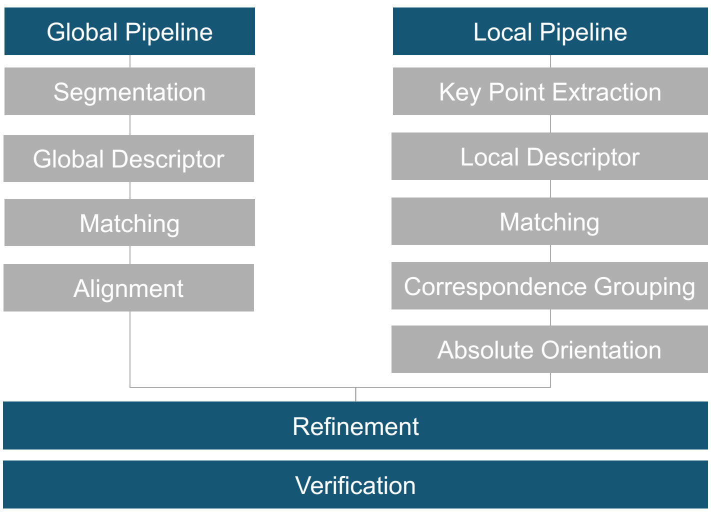

Figure 3: Program sequence for classification and positioning of objects (Aldoma et al. 2012)

Both approaches were implemented for the coupling assistant. The Viewpoint Feature Histogram (VFH) was used as global descriptor (Blume et al. 2015). This descriptor uses the sensor data to calculate geometric information and the relative position between sensor and implement, which are stored in a histogram. For classification purposes, the histogram is compared with histograms that have already been classified in a database. If there is a high degree of agreement, it can be concluded that it is the same object. The alignment of the implement is then also carried out with the help of this histogram. However, a direct reading of the position is not possible. Instead, the histogram is compared with other histograms in which the position is known. The disadvantage of this method is the large amount of training data needed to generate the classified histograms.

The Fast Point Feature Histogram (FPFH) has been implemented as a local descriptor. In order to reduce computing time, this descriptor is not applied to all sensor data points, but only to certain key points. These key points can be selected via an edge detector, for example. The local descriptor then calculates, similar to the VFH, a histogram describing the geometric properties around the sensor data point. This feature, described by the histogram, is significantly reduced in complexity compared to the VFH, since it only describes a small part of the object. For classification, the feature is compared with the feature stored in a database. If the match is high, a pair of points is formed. If there are enough point pairs fixed in a geometric relationship to each other, the implement was successfully detected. The position can be calculated by transforming the point pairs. The advantage of this approach is that only one 3D model of each implement is needed. A time-consuming training phase, in contrast to the global descriptors, is no longer necessary. The disadvantage is the high number of parameters that influences the result.

Newer approaches for classification and positioning are mainly based on neural networks. The reason for this is the increasing performance of modern graphics cards over the last 10 years, which enable an efficient training. Convolutional neural network is mostly used for 2D images. This type of neural network usually consists of so-called convolutional layers in combination with pooling layers for feature extraction. Fully-connected layers can be connected downstream for classification and position determination (Redmon et al. 2015).

The neural network YOLOv2 used in this project consists exclusively of standard convolution layers and max-pooling layers. As input variable, an image is scaled to the size of 416 x 416 pixels. This picture is divided into 169 equal sized parts (13 in x- and y-axis each) and five boxes are calculated for each of these parts. The result of the neural network is a vector with five delimitation boxes for each of the 169 cells, which contains the position data (X-position, Y-position, height, width) and a probability digit that reflects whether the delimitation box completely encloses an object. At the same time, the probability distribution for each trained class is calculated for each cell. These are calculated with the limitation boxes and filtered with a limit value. The result is the 2D position and the class of the implement located behind the vehicle (Figure 4).

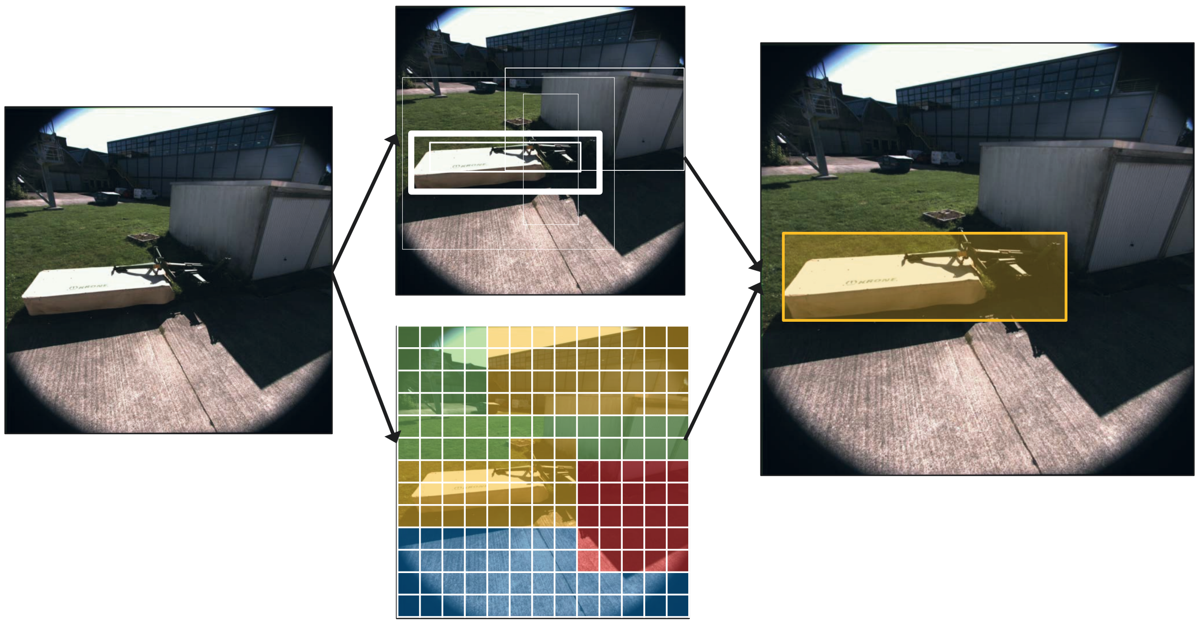

Figure 4: Construction of the neural network YOLO using the example of mower identification according to Redmon et al. (2016)

In contrast to the program sequences of the global and local descriptors, the neural network does not calculate a position in 3D space. For this reason, the six degrees of freedom of the implement are calculated with the aid of a downstream algorithm, the iterative closest point algorithm (ICP). This algorithm makes it possible to adapt 3D point clouds to each other. The neural network delimitation box is transformed into the depth image of the stereo camera to estimate the position of the implement. This position serves as the starting point for the ICP, to which a 3D model of the implement is transformed. This 3D model is iteratively adapted to the sensor data to reduce the point distances between the two point clouds. The result is the transformation between implement and sensor and corresponds to the relative position between tractor and implement.

Path Tracking Control

Beside the position of the implement and the path planning, the assistance system for automatic coupling to implements requires a path tracking control. The path planning algorithm generates a path from the tractor to the implement. The function of the track control is to keep the tractor on this path in such a way that the error is minimal. In particular, the lateral error and the tractor's orientation to the implement at the end of the path are relevant for a successful coupling. Depending on the coupling design, the permissible errors are only a few centimeters and angular degrees.

A model predictive control has been designed for this automated coupling system. In this control method, a model is used to determine the future behavior in such a way that an optimal sequence of the control variables minimizes the errors to future reference values. The sequence of the control variables is optimal with regard to the selected weighting matrices, since these are solved as a mathematical optimization problem. In real-time applications, this must be solved in a fixed frequency. This is necessary because in each iteration only the first set of the calculated sequence is used for the control, so that model errors and disturbances can be compensated.

This increases the calculation time exponentially with the prediction horizon and the model complexity (number of states in the state-space model). In addition to predicting future behavior, the use of optimization has the advantage that restrictions of states and manipulated variables can be taken into account. Furthermore, this method can be used to control all non-linear systems, and because of the state-space representation all states can be controlled. The controller was implemented in C++ with the open source library ACADO Toolkit (Quirynen et al. 2015). The ACADO Toolkit allows the implementation of model predictive controllers in real time, because the computing time is shortened by efficient optimization procedures and by exporting the integrations into efficient C-code.

For the model predictive control the tractor is modeled as a kinematic state-space representation according to equation 1.1 to 1.5:

$$\dot{x}_K = v_T \cos(\theta) + v_T \frac{l_K}{l_T} \tan(\varphi_L) \sin(\theta) \qquad \text{(Eq. 1.1)}$$

$$\dot{y}_K = v_T sin(\theta) - v_T \frac{l_K}{l_T} \tan(\varphi_L) \cos(\theta) \qquad \text{(Eq. 1.2)}$$

$$\dot{\theta} = \frac{v_T}{l_T} \tan(\varphi_L) \qquad \text{(Eq. 1.3)}$$

$$\dot{\varphi}_L = \frac{1}{T_L} \left(\varphi_L^d - \varphi_L\right) \qquad \text{(Eq. 1.4)}$$

$$\dot{\varphi}_L^d = u. \qquad \text{(Eq. 1.5)}$$

Figure 5 shows the kinematic relations of the state-space model (Eq. 1.1 to 1.5), where point K describes the positions x_K and y_K, and θ the orientation of the tractor's coupling point in a global coordinate system. The constant v_T describes the vehicle velocity at the rear axle. The system input u is the derivation of the desired steering angle φ_L^d whereby the desired steering angle is used as the input into a PT1 element. This lag element takes the time delay of the tractor's steering control loop into account.

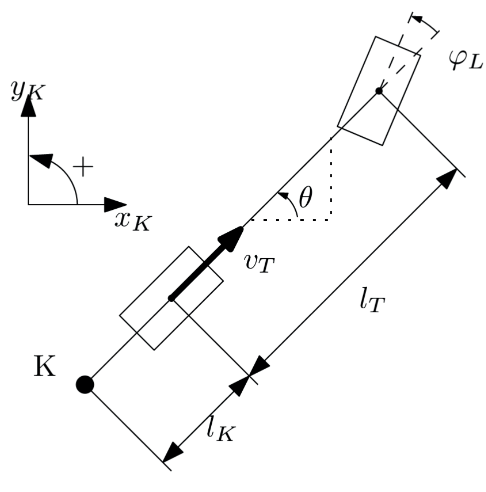

Figure 5: Sketch of the kinematic model

The equations 1.1 to 1.5 show that the tractor's coupling point is controlled. Usually, the rear axle is used in a path tracking control as the point of interest. The control of the rear axle has turned out to be inappropriate due to excessive errors in lateral displacement and orientation at the end of the path. The system input as the derivation of the desired steering angle φ_L^d was chosen, so that the change of the desired steering angle can be considered for comfort reasons as a restriction -0,1 rad/s ≤ u ≤0,1 rad/s in the optimization problem. The tractor's steering dynamic is modeled as a first-order lag element, whereby the time constant T_L = 0.375 s was determined by step responses. The physical restrictions for the steering angle were determined with -0,5 rad ≤ φ_L ≤ 0,5 rad. The sampling frequency was set to 10 Hz. This is larger than the steering dynamics and is sufficient for this application. In each iteration, the steering angle sequence is calculated for a path track of 1 m or 10 s at a velocity of -0.1 m/s, but only the first steering angle is used to control the tractor. Only the system states x_K, y_K and θ_K are weighted in the optimization problem. With an approach to the implement the orientation is becoming more and more important, since orientation at the end of the path is more decisive for successful coupling than the position.

Other simpler path tracking controls (Hellstrom and Ringdahl 2006, Kanayama et al. 1990) than the described model predictive one, with using non-linear feedforward control, have proved by our tests to be unsuitable. These controls generally have a lower control quality, no prediction and no consideration of restrictions as well as steering dynamics. Furthermore, a path tracking control (Müller and Deutscher 2007) was investigated, which is used in the automotive sector for automated parking systems. In this control, the nonlinearities are compensated by a feedback linearization. The steering dynamics is also modeled as a first-order lag element. However, in this application the rear axle is controlled, since the authors didn't find a so-called flat output for the differential equations 1.1 to 1.5. This led to the fact that in this application the coupling was not successful under certain position errors and localization disturbances. The model predictive controller used in this project has the disadvantage that it requires a high computing effort. A separate computing unit has been provided for this purpose. For product development the tractor could be modeled as a linear parametric variable system, i. e. the tractor model is linearized in many different operating points. Subsequently, an explicit model predictive controller could be used, which is pre-calculated offline and exists as a lookup table. This eliminates the calculation of a non-linear optimization problem in real time. This approach has been implemented for a non-linear single-track model to control the vehicle's rear axle on a path in Besselmann and Moraril (2009).

Results

Decisive for a successful coupling to an implement are the lateral offset, which must be less than 3 cm, and an angular accuracy of less than 2.5° between implement and vehicle. Table 1 shows the accuracy for reaching the coupling point for six different test runs from different positions. On the one hand, the deviations measured by the sensor at the end of each trip are shown, which reflect the inaccuracies of the model predictive controller. On the other hand, the absolute error is shown. This results from the optical position determination.

Table 1: Lateral misalignment and angular deviation to the coupling point in six attempts to couple to a mower unit

Approach	Lateral offset		Angular deviation	
	sensor	absolute	sensor	absolute
1.	1.1 mm	4 mm	1.99°	0.44°
2.	3.9 mm	8 mm	1.06°	0.95°
3.	1.5 mm	8.5 mm	1.59°	2.41°
4.	8.4 mm	24 mm	0.77°	0.45°
5.	0.9 mm	2 mm	0.87°	1.34°
6.	3.3 mm	9.5 mm	0.91°	1.62°

The required tolerances were maintained in all six tests. The position deviations due to the controller are negligible with less than 1 cm. The angular deviations are, with sometimes up to 2°, the greater challenge. It can be clearly seen that the biggest errors occur in the optical position determination of the implement. This is caused by noise in the depth due to sensor image mainly.

Conclusions

The five presented assistance systems were successfully researched. The automated coupling to implements was demonstrated at various events using the algorithms presented here. In the field of object recognition, the neural networks in particular have proven their worth. These allow high detection rates with a large number of different implements. The approaches presented with the help of 3D descriptors are especially also of importance for alternative sensor concepts such as 3D laser scanners due to their transferability.

References

Aldoma, A.; Marton, Z.-C.; Tombari, F.; Wohlkinger, W.; Potthast, C.; Zeisl, B.; Rusu, R. B. ; Gedikli, S.; Vincze, M. (2012): Tutorial: Point Cloud Library: Three-Dimensional Object Recognition and 6 DOF Pose Estimation. IEEE Robot. Automat. Mag. 19(3), pp. 80–91

Besselmann, T.; Morari, M. (2009): Autonomous vehicle steering using explicit LPV-MPC. In: Control Conference (ECC), 2009 European, IEEE, pp. 2628–2633

Blume, T.; Schattenberg, J.; Frerichs, L. (2015): Innovative Assistance Systems based on a Backward-Looking 3D-Time of Flight Camera. In: 73rd International Conference on Agricultural Engineering LAND.TECHNIK AgEng 2015, Innovations in Agricultural Engineering for Efficient Farming, Hannover, 6–7 November 2015, VDI-Bericht 2251, pp. 505–510

Hellstrom, T.; Ringdahl, O. (2006): Follow the Past: a path-tracking algorithm for autonomous vehicles. International journal of vehicle autonomous systems 4(2–4), pp. 216–224

Kanayama, Y.; Kimura, Y.; Miyazaki, F.; Noguchi, T. (1990): A stable tracking control method for an autonomous mobile robot. In: IEEE International Conference on Robotics and Automation, Cincinnati, USA, 13–18 May 1990, Proceedings, pp. 384–389

Müller, B.; Deutscher, J. (2007): Orbital tracking control for car parking via control of the clock using a nonlinear reduced order steering-angle observer. In: European Control Conference (ECC), Kos, Greece, 2–5 July 2007, pp. 1917–1924

Quirynen, R.; Vukov, M.; Zanon, M.; Diehl, M. (2015): Autogenerating microsecond solvers for nonlinear MPC: a tutorial using ACADO integrators. Optimal Control Applications and Methods 36(5), pp. 685–704

Redmon, J.; Divvala, S.K.; Girshick, R.B.; Farhadi, A. (2016): You Only Look Once: Unified, Real-Time Object Detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 27–30 June 2016, https://arxiv.org/abs/1506.02640, http://dx.doi.org/10.1109/CVPR.2016.91

Redmon, J.; Farhadi, A. (2017): YOLO9000: Better, Faster, Stronger. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 21–26 July 2017, https://arxiv.org/abs/1612.08242, http://dx.doi.org/10.1109/CVPR.2017.690

Authors

Dipl.-Ing. Tobias Blume and **M. Sc. Ilja Stasewitsch** are scientific assistants, **Dipl.-Ing. Jan Schattenberg** is Deputy Director and **Prof. Dr. Ludger Frerichs** is Director of the Institute of Mobile Machines and Commercial Vehicles at TU Braunschweig, Langer Kamp 19a, 38106 Braunschweig, Germany, e-mail: t.blume@tu-bs.de.

Acknowledgments

The project is supported by funds of the German Government's Special Purpose Fund held at Landwirtschaftliche Rentenbank.