

Row-detection on an agricultural field using omnidirectional camera.

Stefan Ericson and Björn Åstrand

Abstract—This paper describes a method of detecting parallel rows on an agricultural field using an omnidirectional camera. The method works both on cameras with a fisheye lens and cameras with a catadioptric lens. A combination of an edge based method and a Hough transform method is suggested to find the rows. The vanishing point of several parallel rows is estimated using a second Hough transform. The method is evaluated on synthetic images generated with calibration data from real lenses. Scenes with several rows are produced, where each plant is positioned with a specified error. Experiments are performed on these synthetic images and on real field images. The result shows that good accuracy is obtained on the vanishing point once it is detected correctly. Further it shows that the edge based method works best when the rows consists of solid lines, and the Hough method works best when the rows consists of individual plants. The experiments also show that the combined method provides better detection than using the methods separately.

I. INTRODUCTION

Mobile robots for use on an agricultural field require reliable sensors for both localization and perception. Today, one of the most common used sensors for agricultural machinery is the RTK-GPS. It is mainly used as position measurement for tractor autopilots, where the farmer can drive in straight rows with minimum of overlap between rows. It is also used on autonomous agricultural robots but only for research. The drawbacks of the RTK-GPS are the dropouts and the requirement of clear view of the sky. The cost has also mentioned as an issue, but promising work shows a way of building low-cost RTK-GPS using open source library [1].

The use of cameras for localization has been seen more as a complementary method to GPS. However, advantages of using camera are that it can be used for simultaneously localization and mapping (SLAM) and obstacle avoidance. It could also provide low-cost system. In the case of navigation on an agricultural structured field, one of the most important tasks is to keep track of the rows and to separate the crops from weed and soil. Several agricultural robots has been presented in the literature, some navigating using only GPS [2], others with vision [3], [4], [5], and some with sensor fusion between several sensors [6]. In [7] a mobile robot for automated weed control is presented where perspective cameras are the main sensor. In own previous work [8] a mobile robot was presented using row-following from perspective camera and visual odometry for navigation. Fig.

1 shows the mobile robot on a row-structured agricultural field.



Fig. 1: Mobile experimental robot on a row-structured agricultural field

The vision guided robots introduced so far have used perspective cameras. There are several advantages of using an omnidirectional camera instead. First more and longer rows can be captured, i.e. more robust to weed pressure and missing crops and rows. Second, it sees plants beside the robot which may give better estimate of alignment error. Finally it sees behind the robot which gives the opportunity to achieve better row-guidance at end of row, and enables monitoring of field operations. Among the omnidirectional cameras there are both cameras with catadioptric lens and with fisheye lens. The major difference is the range of azimuthal view. A fisheye lens starts from zero which means it sees straight ahead, and end somewhere above 90° . The catadioptric lens on the other side cannot see straight ahead due to its construction, but it has a wider range above 90° . The image analysis on omnidirectional images can be categorized in two groups, those who require the image to be unwrapped and those who are applied directly on the omnidirectional image. The unwrapping is a time consuming process and for real-time applications on a mobile robot the latter is to prefer. The algorithms used in this work do not need the images to be unwrapped.

Successful work on omnidirectional images has recently been presented in [9] where lines vertical to the camera is extracted. A SIFT-like descriptor is used for matching and tracking these lines between frames. This method is used for localization of a mobile robot [10], and it provides accurate heading information as well as translation. A drawback of the system is that it does not deal with tilt. In an agricultural

S. Ericson is with School of Technology and Society, University of Skövde, Skövde, Sweden stefan.ericson@his.se

B. Åstrand is with the School of Information Science, Computer and Electrical Engineering, Halmstad University, Halmstad, Sweden, bjorn.astrand@hh.se

scene the robot is moving in uneven terrain and the tilt is required to estimate the position of the row. Further more only a few lines can be expected to be radial to the camera on a field consisting of parallel rows. At the beginning of a row there will be one line ahead of the robot and while driving along a row, radial lines can be found in front and behind the robot. This is too few lines and hence this method is not suitable for this application.

In [11] the authors present a novel edge-based method to find lines and further rectangles in catadioptric images. The lines are extracted using edge detection on a grayscale image. All connected edge points are grouped, and the endpoints are found. Lines shorter than a specified threshold are rejected. The authors use a method where the points are mapped via a sphere. In that way all straight lines in the omnidirectional image can be represented by a great circle on the sphere. Each line is then splitted so all points on the same line can be projected on one great circle. The last step in the algorithm is to merge all lines which can be represented by the same great circle. This method is applied in finding the attitude of an UAV [12] and to navigate in urban environment by finding and tracking the vanishing point [13]. However, this method requires well defined edges, consisting of connected edge points in the direction of the lines.

In an agricultural scene the rows consist of single plants placed on a line with different distance. Plants close to the robot, or in this case close to the camera, hold important information about the robot's position relative the row. In this particular area the rows may not be viewed as solid, rather as individual plants. The method presented in [11] uses edge detection such as Canny for extracting line. This method is expected to work in areas far away from the robot, where the rows will be seen as solid lines. In the areas close to the robot, the rows may only consist of the shapes of the individual plants. In this area the Hough transform is potentially better to use. A method for detecting lines in catadioptric images using Hough transform is presented in [14].

This paper contributes with a method of detecting parallel rows on an agricultural field using an omnidirectional camera. The rows can consist of either solid lines or individual plants. Both catadioptric and fisheye lens are supported, and the algorithm extracts the vanishing point which contains information about both heading and tilt. This work evaluates the two methods to find lines on agricultural scenes and suggests a method that combines the Hough transform and the edge method.

II. METHOD

This work differs from the work done in [11] in three ways. First, the calibration of the omnidirectional camera uses the Taylor model [15] which allows both catadioptric and fisheye lenses. Second, a Hough transform step is added to find rows consisting of individual plants. Finally a second Hough transform is applied to find the vanishing point of the parallel row structure.

A. Camera calibration

It is assumed that the camera uses a lens/mirror combination that provides a single effective viewpoint, which means that all incoming rays can be modeled to intersect the center point of a unit sphere. Then each ray can be represented by a unit vector on this sphere as shown in Fig. 2. Further it is assumed that the lenses and mirrors are symmetric and that the sensor plane is perpendicular to the optical axis. In [16] it is shown that both fisheye and catadioptric cameras can be projected onto the image plane using two functions g and h as shown in (1).

$$\alpha'' \mathbf{p}'' = \begin{bmatrix} h(\|\mathbf{u}''\|) \mathbf{u}'' \\ g(\|\mathbf{u}''\|) \end{bmatrix} = \mathbf{P}'' \mathbf{X} \quad (1)$$

where $\alpha'' > 0$ is factor, \mathbf{u}'' is the point on sensor plane, $\mathbf{P}'' \in \mathbb{R}^{3 \times 4}$ is projection matrix and \mathbf{X} is the scene point. In [15] a unified method to calibrate both fisheye and catadioptric camera is presented. This is done by rewriting the two functions (g and h) as one function g/h . By representing this function with a Taylor approximation, the relation between a scene point and a point in the sensor plane can be expressed as shown in (2).

$$\alpha'' \mathbf{p}'' = \begin{bmatrix} \mathbf{u}'' \\ a_0 + a_2 \|\mathbf{u}''\|^2 + \dots + a_N \|\mathbf{u}''\|^N \end{bmatrix} = \mathbf{P}'' \mathbf{X} \quad (2)$$

where a_n is coefficients of the Taylor polynomials. The advantage of using this representation is that it is valid for both catadioptric and fisheye lenses. Hence the method can be used for both types. The Matlab toolbox described in [17] is used for the calibration procedure, which provides the relation between an image point in pixels with a unit vector from the single effective view point.

B. Row model

Rows in the scene can be represented by a plane cutting the unit sphere in the camera model, i.e. each point on the row can be expressed as a vector lying both on the plane and the sphere. The plane can further be expressed by its normal, which means each row can be expressed as this normal. From here on normal is used for denoting the normal vector perpendicular to the plane. Fig. 2 shows how a row of plants is projected on the sensor plane and the corresponding image.

C. Edge method

Crop rows far away from the camera will be viewed as solid lines. Hence they can be modeled as lines which can be detected with an edge method. This part is similar to the method described in [11]. The first step is to apply the Canny algorithm to the image. Then a mask is applied to remove points not of interest. It can be structures from camera mounting or points close to the edge of visible field. Further each line is divided into sections where each section can be represented by one normal vector on the unit sphere. That means only points in the same direction are kept in one

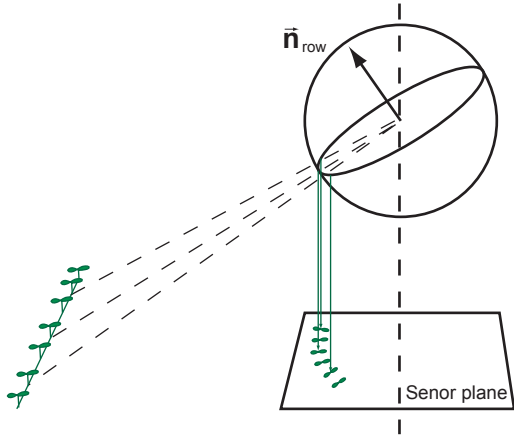


Fig. 2: Projection through the lens on to the sensor plane by a unit sphere model.

line. Minimum line length is specified and line shorter than this threshold is removed. The last step is a merging step that merges lines describing the same row. The orthogonal distance between all normals is calculated. Normals close to each other, i.e. the distance is shorter than a specified threshold, are considered to be the same normal. The new normal is estimated by using singular value decomposition of all points.

D. Hough method

The Hough method consists of a Hough transform followed by a non linear optimization step. The original image is used as input. Hence a separate mask can be used to specify the area of interest. First a Gaussian smooth filter is applied to reduce noise in the image. Then an adaptive threshold is applied to produce a binary image, where corresponding unit vectors can be calculated for each point. The threshold is tuned to allow a certain number of points to be selected. In this way the computation time of the Hough transform can be controlled. The parameters selected for the Hough transform is spherical coordinates of the normal to the plane representing the rows.

$$\vec{n}_r = [\sin(\varphi_r)\cos(\theta_r) \quad \sin(\varphi_r)\sin(\theta_r) \quad \cos(\varphi_r)]^T \quad (3)$$

where $0 < \varphi_r < \pi$ and $0 < \theta_r < 2\pi$. Let $\vec{n}_r = [n_x \quad n_y \quad n_z]^T$ denote the normal to the plane. Then the points on a row will satisfy (4).

$$n_x \cdot \sin(\varphi)\cos(\theta) + n_y \cdot \sin(\varphi)\sin(\theta) + n_z \cdot \cos(\varphi) = 0 \quad (4)$$

The size of the accumulator is kept small to increase the processing speed. Experiments in this paper uses accumulator of size 90×90 . The accuracy of the normal describing the plane is increased by applying a line fitting step as suggested in [14].

E. Finding vanishing point

The vanishing point contains information about the heading of the robot and also tilt. Similar to how points on a

row produce vectors that lies on a plane, several parallel rows will produce normals that define a plane. Hence the same method can be used to find the major direction of all rows. The normal to the multiple planes will point out the vanishing point in the image. The method used is the same as the Hough transform used for finding normals representing the rows. The number of rows is generally low which means the second Hough transform will execute fast. The range of φ and θ is selected to correspond to the range of expected vanishing points. For the case with a mobile robot in an agricultural field, the vanishing point is expected to be along the horizon. Hence points too high in the sky are rejected.

The output from the second Hough transform is used as input in a non linear optimization which gives the resulting vector describing the vanishing point. Fig. 3 shows the vector describing vanishing point, vectors describing the rows and finally each individual point that is seen in the image plane.

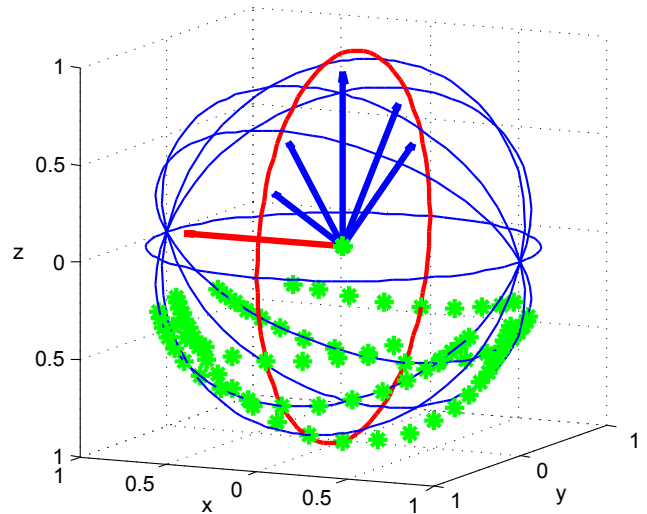


Fig. 3: Normals and planes in unit sphere describing vanishing point(red) and rows (blue). Projections of individual plants are marked in green.

F. Combined method

The combined method takes advantage of both edge method and the Hough method by feeding the second Hough transform with normals from both methods. These normals are weighted according to how many points on the row that have contributed to the normal. This leads to higher weight for rows that were built from many points. For the edge method this corresponds to longer connected edges, and for Hough method the value in the accumulator, i.e number of points contributing to the line. This step reduces the use of the merging step in the edge method since two short lines on the same great circle contributes to the same cell in the second Hough transform. To adjust the balance between the methods, a general weight w is applied, where $0 < w < 1$.

Then the normals are weighted according to (5).

$$\vec{n}_{comb} = \begin{cases} w \cdot \vec{n}_{edge} & \text{for edge normals} \\ (1 - w) \cdot \vec{n}_{hough} & \text{for Hough normals} \end{cases} \quad (5)$$

The weight w is determined by experiments and in this paper 0.7 is used.

III. EXPERIMENTS

Experiments are performed on synthetic images to evaluate algorithm performance and on real field images to validate the result.

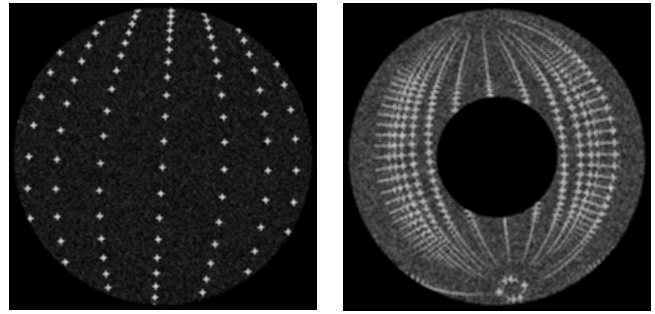
A. Hardware setup

The camera used for the experiment is a Prosilica GC2450C, which provides images with 2540x2040 pixels. It is connected using Gigabit Ethernet, and lens is of c-mount type. Two different lenses are used, one fisheye and one catadioptric. The fisheye lens is a FE185C057HA from fujinon, and the catadioptric is a PAL-S25G3817-27C from American Accurate Components, Inc. The camera is mounted on a stand with the camera pointing downwards. The camera and the two lenses are calibrated using multiple chessboard images, and the calibration data is stored for use in the generation of synthetic images.

B. Synthetic images

The synthetic images are generated by projecting rows onto the sphere. The rows are specified by the distance between rows and the distance between each individual plant. The position of a plant is then projected onto the sensor plane by using real calibration data from different lenses. A plant is build of a model with four leaves. All plants in the image are similar, i.e no scaling is applied. The camera position is defined by setting the yaw, pitch and roll relative the plane with the rows. From these two orthogonal vanishing point is calculated, one along the rows, and one perpendicular to the rows. Hence a true value for the vanishing point exists, which will be used as reference to the estimated vanishing point. Further, background noise is added to the image and a circular mask is applied. Fig. 4 shows two synthetic images with random yaw and pitch. One use calibration data from a fisheye camera and the other from a camera with catadioptric lens. The level of the background noise is 0.2 and 0.5.

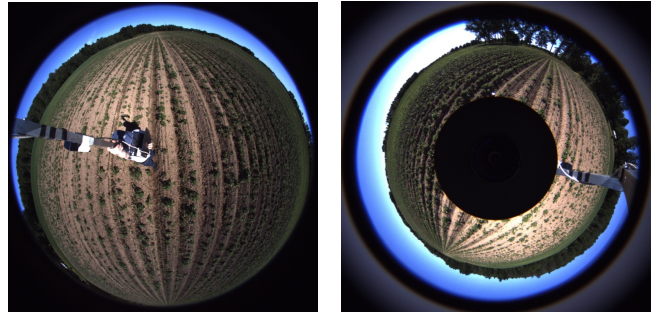
There are three types of noise that can be added to the images. First, there is background noise as mentioned above, which is created by adding salt and pepper noise followed by Gaussian filtering. Second there is a noise on the position of each individual plant. It is introduced as an error in plant distance or row distance. The latter affects how straight the rows are. The last type of noise is created by adding plants at random position. This corresponds to weed in agricultural images.



(a) Fisheye lens

(b) Catadioptric lens

Fig. 4: Synthetic images with different camera tilt and background noise level.



(a) Fisheye lens

(b) Catadioptric lens

Fig. 5: Images of row structure on real field.

C. Evaluation on real field images

The methods are validated by using images from a real field. Typical images are shown in Fig. 5. The green parts of the image is used. The main difference from the synthetic images is that these images have visible areas above the field. This area has to be removed before the analysis, which is done by tracing the boundary of the regions and then calculate a mask from the result.

IV. RESULTS

A. Evaluation on synthetic images

Rows are detected in each image using both Edge and Hough method. Fig. 6 and Fig. 7 show the result for a fisheye image and an omnidirectional image. Rows detected by the Edge method are marked in blue and by Hough method in red. It can be seen that the Hough method captures the center of the row while the edge method detects the two edge lines of the row, i.e. row boundary.

First test evaluates the accuracy of estimated vanishing point by calculating the angle between true value and the estimated. This test is done on 100 randomly generated images, where yaw, pitch and lens type are varied. Hence both on fisheye and catadioptric images are used. The only noise added is background noise with density 0.1. The plant distance is low so almost solid lines are generated. The result is shown in Table I. It shows that the mean error is very

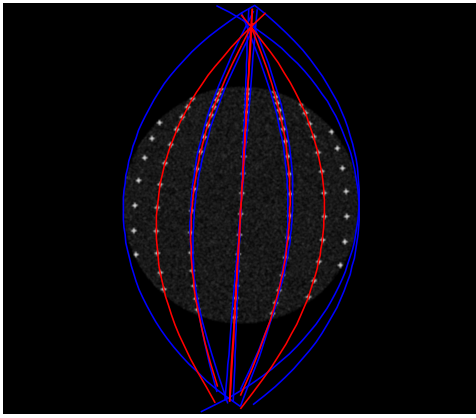


Fig. 6: Rows found in fisheye image. Edge (blue), Hough (red).

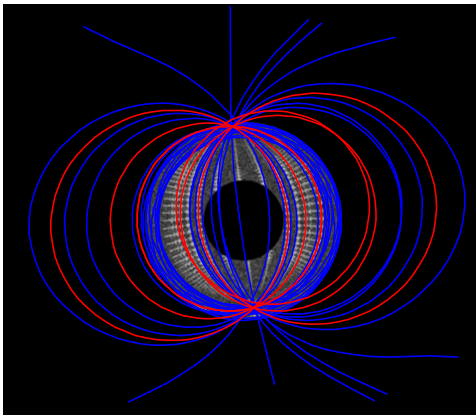


Fig. 7: Rows found in catadioptric image. Edge (blue), Hough (red).

small for correct found rows, so further experiments focus on detection rate.

The second test evaluates sensitivity to different plant distance. The detection rate is calculated as the quote between the number of correct rows found and total number of rows found. The result is presented in Fig. 8 for fisheye image and in Fig. 9 for omni image.

The notable result is that there is a major difference between the different lens types. For fisheye lens, the best result is obtained using the Hough transform, but for the catadioptric lens the Edge method works best. This is due to the shape of the rows in the images. It can be seen in the images that a catadioptric lens captures more points close to the vanishing point, and hence the rows are shown as solid lines. For the fisheye lens the rows are better described as

TABLE I: Accuracy of estimated vanishing point.

	Mean error (10^{-3} rad)	Max error (10^{-3} rad)	Variance 10^{-8}
Edge method	0.12	1.8	7.9
Hough method	0.036	0.17	0.13
Combined method	0.029	0.35	0.23

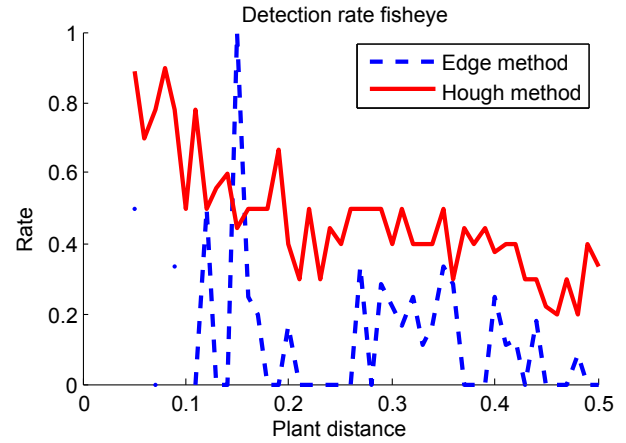


Fig. 8: Detection rate for fisheye image

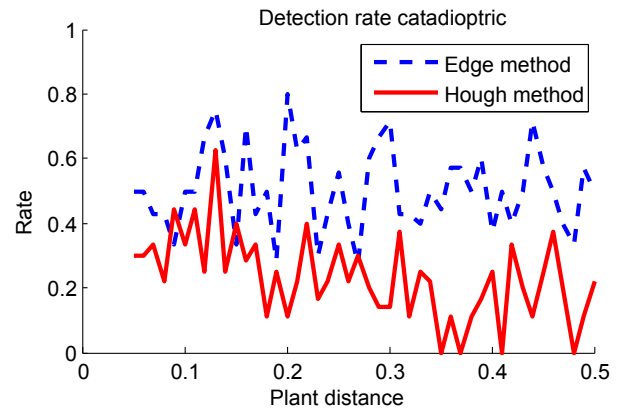


Fig. 9: Detection rate for omni image

individual plants, and hence Hough transform is better.

The last test on the synthetic images compares the two individual methods to the combined method. Images are generated randomly with different noise levels, lens types and plant distance. Each image is evaluated using all three methods. The result is presented as pass or fail, where it is classified as pass if the estimated vanishing point is within a certain distance from the reference. The amount of noise added is high which gives a lot of failures. Table II shows the result.

TABLE II: Comparison between methods.

	Edge method (failures)	Hough method (failures)	Combined method (failures)
Fisheye lens	30	18	15
Omni lens	27	48	23
Sum	57	66	38

B. Evaluation on real field images

The result from the test on real field images is shown in Fig. 10 and Fig. 11. Rows are found with both edge and Hough method in the fisheye image. In the catadioptric image no edge lines are found. The Hough transform provides two candidates of vanishing point, where one is aligned with the row structure, and the other is rejected. These images also show that there is a small error in the calibration since the curvature of the lines is too small.

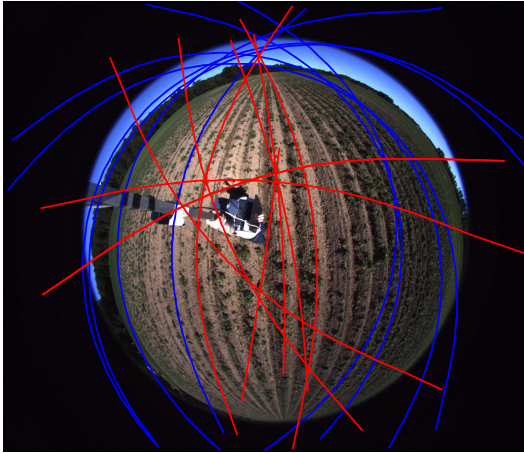


Fig. 10: Lines found in fisheye image. Edge (blue), Hough (red).

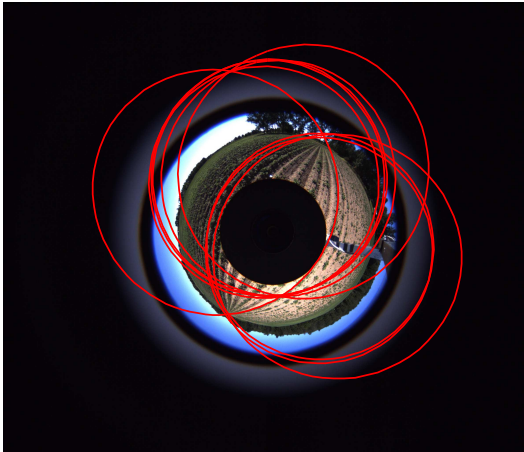


Fig. 11: Line found in catadioptric image. Edge (blue), Hough (red).

V. CONCLUSIONS AND FUTURE WORKS

In this paper we have evaluated two methods of extracting rows from an omnidirectional camera, and we have suggested an own combined method. The result from tests on synthetic images shows that the vanishing point can be estimated with high accuracy. It also shows that the detection rate varies between the methods and the lens used. The edge based method works best when the rows consist of solid lines, and the Hough method works best when the rows consist of individual plants. Hence, the two methods will work differently

on images from fisheye respectively catadioptric lenses. The result from the combined method shows that better detection is obtained than using the methods separately.

Further work is to evaluate different resolutions of images, and to transform the calibration data to the new resolution. The algorithm will be used for positioning a mobile robot on an agricultural field. That would require a real time implementation of the code and a control system of the steering.

REFERENCES

- [1] T. Takasu and A. Yasuda, "Development of the low-cost rtk-gps receiver with an open source program package rtklib," in *International Symposium on GPS/GNSS, International Convention Center Jeju, Korea*, November 4-6 2009.
- [2] A. Stoll and H. Dieter Kutzbach, "Guidance of a forage harvester with gps," *Precision Agriculture*, vol. 2, no. 3, pp. 281–291, 2000.
- [3] J. Billingsley and M. Schoenfisch, "The successful development of a vision guidance system for agriculture," *Computers and Electronics in Agriculture*, vol. 16, no. 2, pp. 147 – 163, 1997, robotics in Agriculture.
- [4] J. Reid and S. Searcy, "Vision-based guidance of an agriculture tractor," *Control Systems Magazine, IEEE*, vol. 7, no. 2, pp. 39 – 43, apr 1987.
- [5] N. D. Tillett, T. Hague, and S. J. Miles, "Inter-row vision guidance for mechanical weed control in sugar beet," *Computers and Electronics in Agriculture*, vol. 33, no. 3, pp. 163 – 177, 2002.
- [6] 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.
- [7] 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.
- [8] S. Ericson and B. Åstrand, "A vision-guided mobile robot for precision agriculture," in *Proceedings of 7th European Conference on Precision Agriculture*, Wageningen, Netherland, July 7-9 2009, pp. 623–630.
- [9] D. Scaramuzza, N. Criblez, A. Martinelli, and R. Siegwart, "Robust feature extraction and matching for omnidirectional images," in *Field and Service Robotics: Results of the 6th International Conference (STAR: Springer Tracts in Advanced Robotics Series Volume 42)*, vol. 42. Springer, 2008, pp. 71–81.
- [10] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *IEEE Transactions on Robotics, Special Issue on Visual SLAM*, vol. 24, no. 5, 2008.
- [11] J. Bazin, I. Kweon, C. Démonceaux, and P. Vasseur, "Rectangle extraction in catadioptric images," in *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, 2007, pp. 1–7.
- [12] C. Démonceaux, P. Vasseur, and C. Pegard, "Uav attitude computation by omnidirectional vision in urban environment," in *2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 2017–2022.
- [13] J. Bazin, I. Kweon, C. Démonceaux, and P. Vasseur, "A robust top-down approach for rotation estimation and vanishing points extraction by catadioptric vision in urban environment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008. IROS 2008*, 2008, pp. 346–353.
- [14] X. Ying and Z. Hu, "Catadioptric line features detection using hough transform," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4, 2004.
- [15] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A flexible technique for accurate omnidirectional camera calibration and structure from motion," in *Proceedings of IEEE International Conference on Computer Vision Systems*, vol. 1. Citeseer, 2006, pp. 45–52.
- [16] B. Micușik, "Two-view geometry of omnidirectional cameras," Ph.D. dissertation, Czech Technical University, 2004.
- [17] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Proc. of the IEEE International Conference on Intelligent Systems, IROS06, Beijing, China*. Citeseer, 2006.