JARQ 46 (4), 287 – 293 (2012) http://www.jircas.affrc.go.jp

REVIEW

Development of a Stereo Vision System to Assist the Operation of Agricultural Tractors

Keiji HANAWA[1]*, Takashi YAMASHITA[1], Yosuke MATSUO[1] and Yasuyuki HAMADA[2]

[1] Bio-oriented Technology Research Advancement Institution, National Agriculture and Food Research Organization (Saitama, Saitama 331–8537, Japan)
[2] National Agricultural Research Center for the Hokkaido Region, National Agriculture and Food Research Organization (Memro, Hokkaido 082–0071, Japan)

Abstract

A stereo vision system for agricultural tractors was developed. Two kinds of stereo camera unit were manufactured: one of which is mounted on the front of the tractor and the other on the cabin roof. The system can detect various formed or shaped objects, such as crop rows, ridges, uneven surfaces after tilling and marker traces in front of the tractor in agricultural fields. The system generates distant images with a 0.1-second cycle, while an internal algorithm installed in the system analyzes this distant image data three-dimensionally, and detects the positions of objects. The positional data was transferred to a steering controller of the tractor and tilling work was performed automatically by tracking these objects.

Discipline: Agricultural machinery
Additional key words: automatic steering control, image recognition, stereo camera

Introduction

In recent years, due to the aging workforce and with a lack of successors in the agricultural sector, there is a need for technical development that can ease the burden of those operating agricultural tractors, and enable highly efficient and accurate operation, even by unskilled operators. Based on these requirements, systems for operational assistance must be developed immediately.

In seeding and ridge forming work, an operator drives a tractor parallel to the traces of the previous path, while weeding or harvesting involves the operator running along crops or preformed ridges. Technology to detect crops and ridges is important to develop operational assistance system. Vision technology is considered promising because human operators recognize such objects with the naked eye and there has been considerable research into object detection. A device used to detect green crops by distinguishing them from the field surface soil with a color camera[2] has been practically applied. However, techniques capable of detecting various other crops and ridges are required. The appearances of these objects change due to sunlight, viewing perspective, soil condition and other factors in outside fields, meaning it is not easy to detect them robustly, based on their appearances and using two-dimensional images alone.

The key characteristic of crops and ridges are their three-dimensional shapes within a space. Accordingly, a technology which processes three-dimensional information, such as stereo vision, seems more effective. Research[5] exists, which detects crop rows using a stereo camera device and controls a tractor by tracking the crop rows. This features a stereo camera mounted on the front of the tractor and allows the investigation of the overall three-dimensional shape of luxuriant crop rows. However, a technique to detect a wider range of objects, such as ridges and marker traces, is required. A stereo vision system[1] was developed by a tractor maker, and has already been practically applied. Stereo vision is a useful technology to understand a three-dimensional environment. However, the design method for the stereo camera, reflecting each purpose or mounting place, and techniques exploiting three-dimensional data have not been dis-

*Corresponding author: e-mail khanawa@affrc.go.jp
Received 3 October 2011; accepted 19 March 2012.

cussed sufficiently. Moreover, it is also desirable that the stereo camera be mounted in a place where the camera itself does not disturb the practical operation.

The purpose of this research is to develop a system capable of detecting various objects such as ridges and marker traces on the field, as well as crop rows[4], and realize a system capable of controlling a tractor by tracking these objects. In this research we used stereo vision devices, which were developed for a car driving assist system[3] and incorporating functions such as adaptive cruise control and lane departure warnings. The design of the stereo camera was modified for tractor requirements and an algorithm to detect these objects were developed. In this paper we propose a method to design a stereo camera and an algorithm to analyze three-dimensional data. Moreover, one of the test results conducted in the field is mentioned.

Design of a Stereo Vision System

The stereo vision system (hereinafter referred to as “SIPS”) comprises a stereo camera unit and an image processing unit, which are based on devices produced for the car driving assist system. The electronic parts and circuits of each unit are used unchanged, and the function generating distant images from those captured by the two cameras is identical for the car driving assist system. Conversely, optical parts such as lenses and mechanical parts such as the stay joining both cameras are redesigned for the use of the tractor with the accuracy requirements and heavy vibration in mind. Two kinds of stereo cameras are manufactured, one of which is mounted on the front of an ordinary tractor (Fig. 1) and the other on the cabin roof (Fig. 2).

Fig. 1. Developed stereo camera mounted on the front of the tractor

1. Design of a Front-mount Type Stereo Camera

The first requirement for the SIPS is to cover the detection area in view. The area is set so that 1.5 square meters of ground can be covered from about 1m in front of the tractor as shown in Fig. 3. One of the SIPS designs is to be mounted on the front of the tractor (Point **A** in Fig. 3), which is near the detection area. Based on this requirement, the angle of view is set to a lateral 47° and vertical 37°, and the focal length of the lens is set to 4.4 mm.

The second SIPS requirement is to have performance capable of detecting 10 cm differences in height of the ground surface. Based on this requirement, the base line *b*, which represents the distance between the two cameras, is determined as described below. The stereo vision system can detect the disparity *d* between the images as captured by the two cameras, while the distance *D* from the camera to an object is calculated based on the triangulation principle. Here, the focal length *f*, disparity *d*, and base line *b* are used to represent the equation (1). The relation between the difference in height $\triangle H$ and that in distance $\triangle D$ is given by the equation (2).

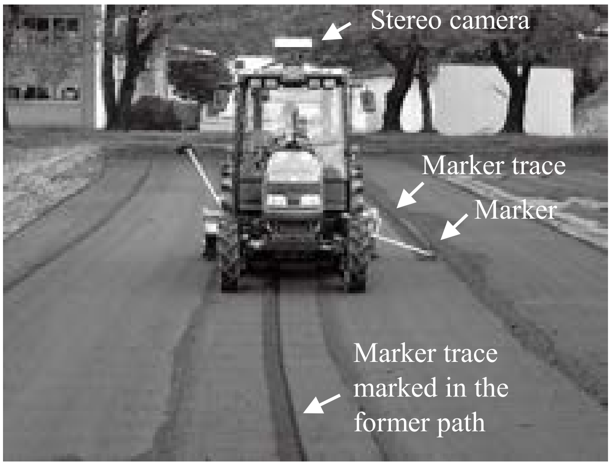

Fig. 2. Roof-mounted stereo camera, engaged in tracking a marker trace

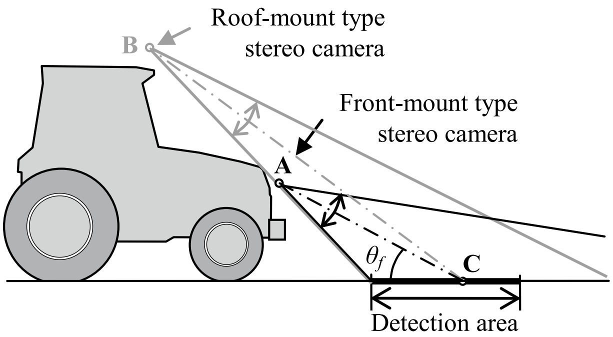

Fig. 3. Mounting location of the stereo camera and its viewing area

$$D = f \cdot b/d \qquad (1)$$

$$\triangle H = \triangle D \cdot \sin \theta_f \qquad (2)$$

θ_f is the angle between the central axis of the camera and the ground surface. To detect a 10 cm difference in height, namely 3 or more pixels in disparity difference $\triangle d$ is required on an empirical basis. By adding this condition, equation (1) is transformed into equation (3).

$$D - \triangle D = f \cdot b / (d + \triangle d) \qquad (3)$$

Since the array pitch of the CCD image sensor device of this camera is 7.4 µm, the difference of disparity $\triangle d$ of 3 pixels is 22.2 µm. In Fig. 3, point **C** represents the intersection of the camera central axis and the ground surface. The distance *D* between points **C** and **A** is about 2.5 m and the angle θ_f is 27°. Based on equation (2), $\triangle D$ is calculated at 22 cm corresponding to the difference in height $\triangle H$ of 10 cm. Substituting these values into equations (1) and (3), the base line *b* is calculated as 13 cm. The SIPS was designed with a base line of 15 cm, with some margin taken into consideration.

2. Design of a Stereo Camera of the Roof-mounted Type

A stereo camera which is mounted on the cabin roof, at point **B** in Fig. 3, was designed. By mounting it on the cabin roof, the negative influences of dust and vibration, operational disturbances and maintenance work will all be reduced.

The same requirements as mentioned above are applied. Since the camera was mounted at a height of 2.7 m, the distance *D* from camera to ground is extended to 4.6 m. To cover the 1.5 square meters of the ground surface in view, the focal length of the lens is extended to 8.2 mm. Next, the base line *b* is determined, in order to fulfill the requirement of detecting a 10 cm difference in height of the ground. Since the angle θ_f is 36°, based on equation (2), $\triangle D$ is calculated as 17 cm, corresponding to the difference in height $\triangle H$ of 10 cm. Since the image sensor of this camera is the same as that of the above-mentioned camera, the disparity $\triangle d$ with 3 pixels is 22.2 µm, hence the base line *b* is calculated at 32 cm by substituting these values into equations (1) and (3). The SIPS was designed with a base line of 35 cm, taking some margin into consideration.

3. Image Processing Unit

Video signals from the stereo camera unit are input into the image processing unit and the following process is executed at a 0.1-second cycle. Fig. 4 shows a ridge formed in a field, an example image captured by a stereo camera of the front-mounted type. Fig. 4-a shows an image obtained by the right camera (512 pixels wide ×200 pixels tall, 256 gray scale). Here, since the image sensor operates with an interlaced scan, each figure is displayed at a size equivalent to 400 pixels in the vertical direction. The right camera image is divided into blocks 128 wide × 50 tall, each of which is 4 × 4 pixels in size. The right camera image is used as a reference, and the matching point is sought for each block along the epipolar line on the left camera image. A gray scale, area-based matching process is applied to detect the matching point, and the sum of the absolute difference (hereinafter referred to as "SAD") is used as the evaluation function.

The disparity of the two images is searched until 128 pixels, and output per 0.25 pixels by the sub-pixel calculation. This processing is performed for all blocks on the right camera image, and the disparity value of each of the 6,400 blocks is generated. The aggregate of these disparity values, as in Fig. 4-b, is called a distant image. Here, when the brightness difference between pixels next to the right and left is less than the threshold value, the disparity value is nullified. Most of the black area on the distant image is one in which the disparity value was deleted due to this condition. This image processing is also the same as that used for cars and performed with a 0.1-second cycle by the dedicated gate array.

Fig. 4-c shows the height of the ground surface, as calculated from each piece of distant image data, while Fig. 4-d shows the color codes changed for each 33 mm

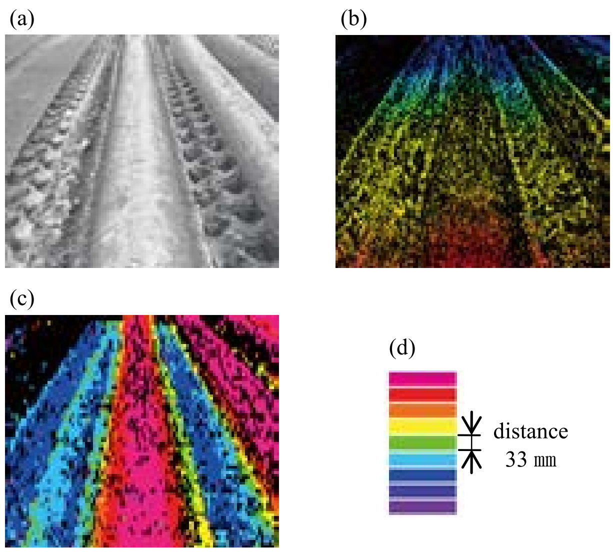

Fig. 4. Images from a front-mounted stereo camera
(a) Right camera image. (b) Distant image. (Disparity: large = red, small = blue) (c) Detected height. (d) Legend of the height and painted color.

height. The variations in height of the detected ground surface are found to be within 33 mm. Fig. 5 shows an example of the images captured by the stereo camera of the roof-mounted type, while Fig. 5-c shows the detected height of the ground surface. The height differences of the ground surface are similar to those in Fig. 4-c. When comparing the two stereo cameras: front- and roof-mounted, as far as the height of the ground surface is concerned, their detention accuracy is almost the same.

Detection Algorithm

The generated distant image is an aggregate of 6,400 sets of disparity data. First, the value of disparity *d* for each block of the distant image is substituted into equation (1), and the distance *D* between the camera and each object is calculated. Next, using the coordinate values of each block on the image, and known parameters such as the mounted position and the camera viewing angle, the coordinate values within the three-dimensional space of each object are calculated. These 6,400 three-dimensional coordinate values are then analyzed by the algorithm described below, and various objects such as crop rows, ridges, uneven differences, plow ditches and marker traces can be detected. In this paper, an algorithm to detect marker traces is described.

Fig. 2 shows a situation where the marker trace on the ground in front of the tractor is detected by a stereo camera of the roof-mounted type. The tractor travels automatically by tracking the marker trace, while tilling is handled by the working machinery. The marker used here is commercially available, and has a structure in which streaky ditches and ridges are formed by using a metal disc to scrape out soil. The width of the marker trace was about 25 cm, and the difference in height between the ditch and ridge was about 10 cm.

Fig. 6 shows an image of the marker trace captured by the stereo camera of the roof-mounted type, while Fig. 7 shows the bird's-eye view generated with the three-dimensional coordinate values of each block. The vertical axis in the figure indicates the forward frontal direction, while the scale indicates distances from the front of the tractor. The color display for each data indicates heights from the average ground level. In Fig. 7, a marker trace located slightly to the left, a streaky series of the ditch in blue, and the ridge in yellow are all observed. This result indicates that the three-dimensional coordinate values of each block are almost accurate.

The white frame lines in Fig. 7 shows a range in which a marker trace has been detected, which is zoned into 7 segments, each of which is 1.66 m crosswise and 256 mm lengthwise. First, the data within the range is projected onto the vertical plane for each segment. Fig. 8 shows the results of the projected segment located at the

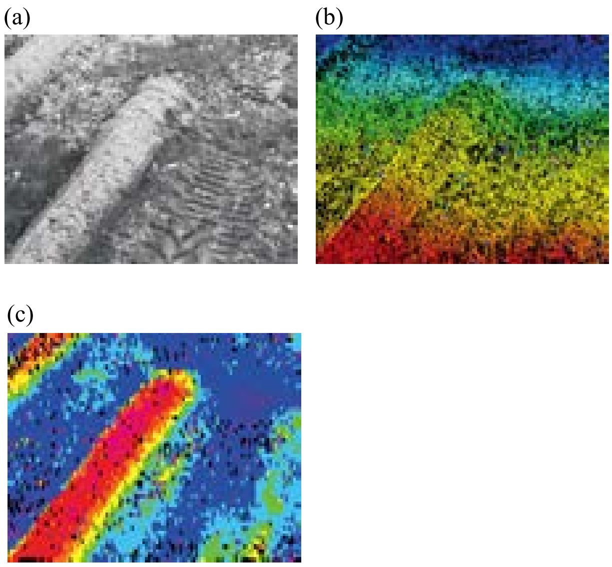

Fig. 5. Images from a roof-mounted stereo camera
(a) Right camera image. (b) Distant image. (Disparity: large = red, small = blue) (c) Detected height.

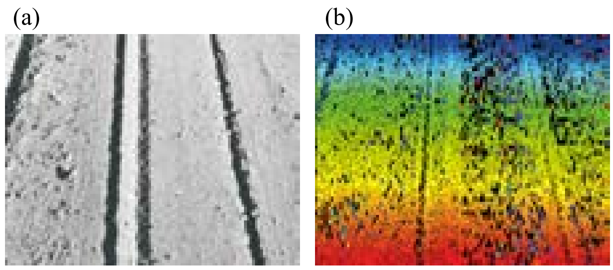

Fig. 6. Images of the stereo camera
(a) Right camera image. (b) Distant image.

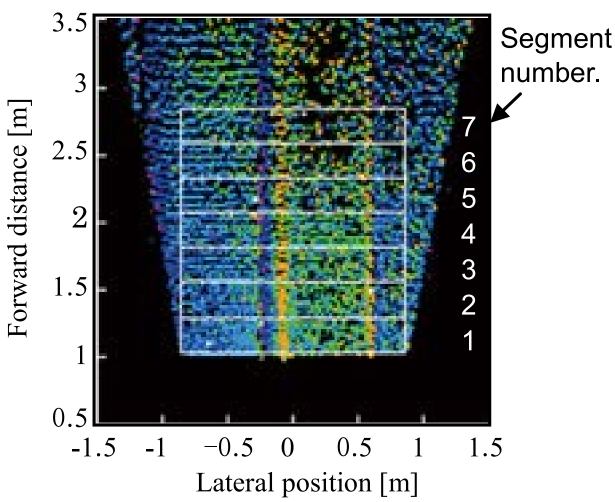

Fig. 7. Projected bird's-eye image

nearest distance, with the cross-section of the ground surface clearly visible. Against these projected data, thinning and interpolation are carried out, and the contour of the ground surface is generated.

The size of the marker trace fluctuates depending on the stiffness of the ground soil. Moreover, depending on the angle from which the marker trace is perceived, the cross-section is an N- or reverse N-shape. Subsequently, as shown in the upper part of Fig. 8, 4 types of shape patterns: small/large and right/left reverse are prospectively memorized. The shape pattern which gives the optimal match is selected at the start of the detection process. In the example shown in Fig. 8, the second shape pattern from the left is selected, which is a reasonable result.

Next, pattern matching calculation of the selected shape pattern is performed against the contour of the ground surface for each segment. The evaluation function is composed of SAD. The degree of matching is calculated while moving the shape pattern bit by bit from the left end to the right end of the contour. When such calculation of the degree of matching is performed against the contour of 7 segments, its value is generated as a two-dimensional distribution within the detection range of the white frame lines in Fig. 7, as shown in Fig. 9. By applying the Hough transfer method, a series of areas showing a higher degree of matching can be detected as an approximated straight line. Fig. 10 shows the detected straight line.

For the front-mount type stereo camera, only certain parameters used in the equation calculating three-dimensional coordinate values differ, and the algorithm for analyzing the values of three-dimensional coordinates is completely the same as above. One feature of the SIPS is that despite any differences in the design specification of the stereo camera, the same algorithm can be applied if the objects are the same.

The above-mentioned is the algorithm for detecting marker traces, but in the process of pattern matching against the ground surface contour, other objects can also be detected by changing the shape pattern to be used. The shape pattern in Fig. 11-a enables the detection of uneven ground surfaces, field boundary, plow ditches, etc., while a pattern like that in Fig. 11-b enables the detection of ridges like the shapes shown in Figs. 4 and 5.

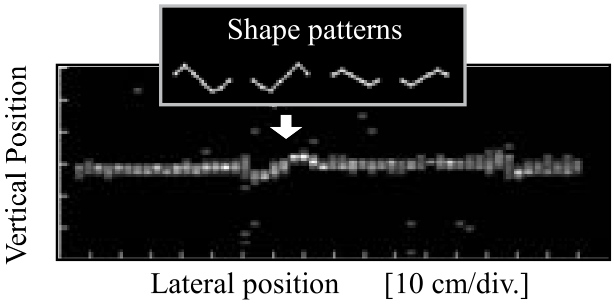

Fig. 8. Profile of the ground surface

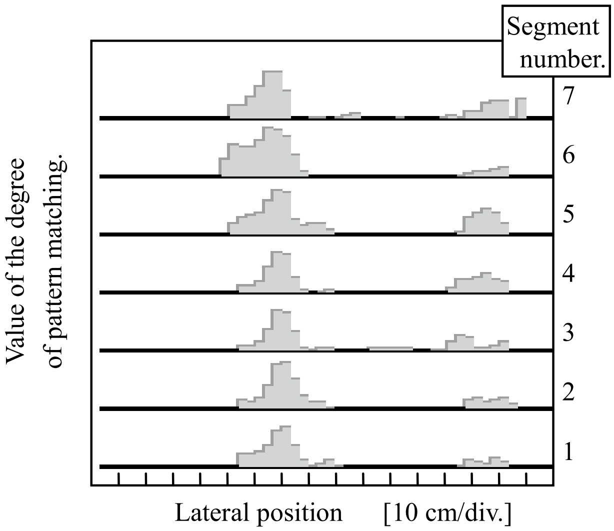

Fig. 9. Profile of the degree of pattern matching

Test Results

A detection algorithm was installed to the SIPS, and

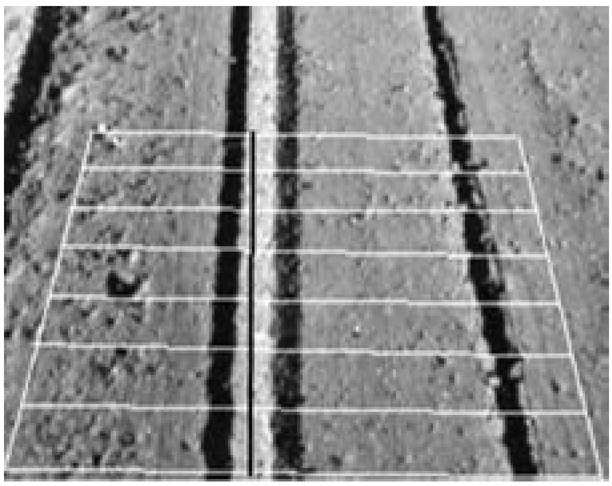

Fig. 10. Detected marker trace (black line)

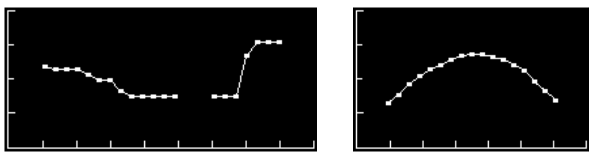

Fig. 11. Examples of the shape pattern

mounted on a tractor, the steering angle of which can be automatically controlled. The SIPS succeeded in detecting various objects, including marker traces, with a 0.1-second cycle, while the positional data of the detected objects was transferred to the tractor controller.

A tilling work test was performed by making the tractor travel tracking marker traces, as shown in Fig. 2. The operational procedures for this test initially include tilling work performed on the first path, followed by the marker trace formed at the working position on the second path. The operator manually steers the tractor to the start position on the second path and starts the detection of marker traces by the SIPS, whereupon the automatic steering control is done. The relative position of the marker trace with the tractor is detected at the start position, and the data thus acquired are memorized by the system. The system also controls the steering angle automatically and allows the tractor to track the marker trace, ensuring the relative position with the trace can be maintained. Tilling work and the formation of a marker trace are performed simultaneously. When the tractor reaches the edge of the field, the operator turns off the automatic steering control, manually turns the tractor, drives to the start position of the next path, and reactivates the SIPS and the automatic steering control. Repeating this operational sequence facilitates tilling work on a continuous basis.

The field used for this test was 19 × 55 m and a rotary tilling device 2.1 m wide was attached to the tractor. The height difference of the maker trace between the ridge and ditch was about 10 cm. The roof-mounted stereo camera succeeded in detecting this maker trace stably. The tractor was controlled and the tilling work was performed by 8 paths (4 returns). Fig. 12 shows the trajectories of the tilling device edge having remained on the field. An auto-tracking type surveying device (nominal accuracy: ±1 cm) was used to measure the trajectory. One of the requirements of this system is that the tractor can travel maintaining a given off-set from the operational trace remaining on the previous path. The distances between trajectories as shown in the lower right of Fig. 12 are measured at 20 cm intervals and the calculated dispersion values are shown in Table 1.

We set a target performance for the accuracy of automatic tracking to within 5 cm by referencing skilled operators. The average standard deviation was 3.4 cm and all values of each path satisfied the target performance. However, the standard deviation of each path showed a tendency to increase. Some error of direction of the tractor existed when the operator manually drove the tractor to the next start position. This error distorted the trajectory of the following automatic steering control, which also had some effect on the successive path. Another operation assist function, which guides the tractor to the next start position, would also be desirable.

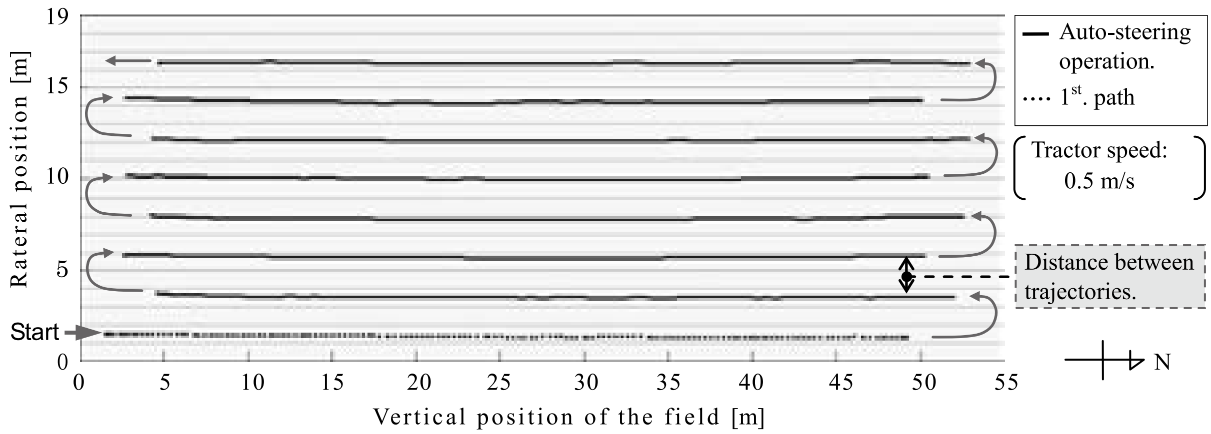

Fig. 12. Test result: trajectories of the tilling device edge

Table 1. Standard deviation of the distance between the trajectories

Path number	Between 1-2 path	Between 2-3 path	Between 3-4 path	Between 4-5 path	Between 5-6 path	Between 6-7 path	Between 7-8 path	Total average
Standard deviation	2.8 cm	3.3 cm	2.6 cm	2.5 cm	3.9 cm	4.1 cm	4.5 cm	3.4 cm

Conclusion

A stereo vision system for agriculture tractors was developed and two kinds of stereo camera were manufactured: one of which the front-mount type and the other, the roof-mount type. Despite differing distances between the cameras and detection areas, we succeeded in obtaining almost identical detection capacity for both stereo cameras.

An algorithm capable of detecting ridges, marker traces and other objects was developed. The distant image data, which is generated by the stereo vision system, is first transformed into three-dimensional coordinate values, from which cross-sectional profiles of the ground surface are then extracted. The positions of the target objects are detected from these profiles using the pattern matching method. This algorithm can be applied for both types of stereo camera when some camera parameters are changed during the process of calculation of the three-dimensional coordinate values from the distant image data.

A field test involving the detection and tracking of marker traces was conducted. The target performances, namely the ability to detect 10 cm height differences and tracking to within 5 cm were both achieved. This result verified the validity of the methods used to design the stereo camera and devise the detection algorithm.

In some cases, the shadow of the tractor covered its detection area. When the sunlight was very strong, blocked up shadows affected the camera image. The performance of the SIPS was restricted in these cases because no distant image data was generated in the shadows. Further improvements in the camera performance are desirable.

Acknowledgements

We express our thanks to those concerned of Fuji Heavy Industries LTD, Subaru Research Center for helping us with our research project with the provision of stereo vision devices for the cars used in this research.

References

1. Claas KGaA mbH (2010) CAM PILOT System. Brochure from “http://www.claas.com”.
2. Garford Farm Machinery (2009) Robocrop system. Brochure from “http://www.garford.com”.
3. Hanawa, K. et al. (2001) Development of stereo image recognition system for ADA. *Proc. IEEE Intelligent Vehicle Symposium*, 177–182.
4. Hanawa, K. et al. (2007) Study of stereo vision processing methods for operating assist of agricultural vehicles. *IEEJ Trans. EIS-C*, **127**, 513–520 [In Japanese].
5. Kise, M. et al. (2005) A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering*, **90**, 357–367.