Biosystems Engineering (2005) **90** (4), 357–367
doi:10.1016/j.biosystemseng.2004.12.008
AE—Automation and Emerging Technologies

A Stereovision-based Crop Row Detection Method for Tractor-automated Guidance

M. Kise[1]; Q. Zhang[1]; F. Rovira Más[2]

[1]Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, 1304 W. Pennsylvania Ave. Urbana, IL, USA; e-mail of corresponding author: qinzhang@uiuc.edu
[2]John Deere Technology Center, Deere and Company, Moline, IL, USA

(Received 13 May 2003; accepted in revised form 13 December 2004; published online 23 February 2005)

Steering agricultural machinery within rowed crop fields is a tedious task for producers. Automated guidance of the machinery will not only reduce operator fatigue but also increase both the productivity and safety of the operation. An essential aspect of automatic guidance is the ability to identify the pathway between the crop rows. This research was to develop an implementable row-detection algorithm for a stereovision-based agricultural machinery guidance system. The algorithm consists of functions for stereo-image processing, elevation map creation and navigation point determination. The method developed first reconstructed a three-dimensional crop elevation map from a stereovision image of crop rows and then searched for optimal navigation points from the map. The developed stereovision-based crop row detection system was tested in a soya bean field to follow both straight and curved soya bean rows at typical operating speeds. Field validation tests indicated that the stereovision-based guidance system could localise crop rows accurately and reliably in a weedy field with missing sections of soya beans. Based on crop row localisation information, an automated navigation system could guide an autonomous agricultural tractor following both straight and curved rows accurately at normal field operation speeds.

1. Introduction

Steering agricultural machinery within rowed crop fields to perform various production tasks is a tedious job for producers. To solve this problem, automated machinery guidance systems have been developed to automatically steer the machinery following crop rows to perform the required operations. The basic requirements for an automated machinery guidance system include detecting the machinery position and orientation in related to crop rows in real time; planning for an optimal path for the machinery carefully without damaging the crop; and accurately steering the machinery along the desired path. Automated guidance can greatly reduce driver fatigue and therefore increase both productivity and safety of the operation. In terms of their functionality, automated machinery guidance systems can be categorised into the two classes of operator-assisted systems and autonomous systems. An operator-assisted system only assists the human operator in guiding the machinery following the desired paths; whereas an autonomous system replaces the human driver to perform all field operations, including row following, end-of-row turning, and implement controls. Both systems require of on-board navigation sensors to provide the guidance information.

One of the most common types of navigation sensors is the global positioning system (GPS). For example, a four-antenna real-time kinematic-GPS (RTK-GPS)-based navigation system developed by a research group from Stanford University could guide an agricultural tractor following desired paths with a tracking accuracy of 0·04–0·06 m at normal field operating speeds (Bell, 2000). To compensate for GPS positioning error associated with machinery attitude, researchers at Hokkaido University integrated an inertial measurement unit (IMU) with an RTK-GPS to provide more accurate navigation information. This integrated navi-

Notation

a	proportional gain of the steering controller	x_k	deviation of a guidance path from a target path at time instant k, m
b	baseline of the stereo camera, m	x_n	longitudinal location of the navigation point in the vehicle coordinates, m
$c(\tau)$	cross-correlation between g(x) and $h(x)$ with delay τ	XYZ	vehicle axes
d	disparity	y_c	vertical location of the 3D point in the camera coordinates, m
$E\{\}$	average	y_d	horizontal offset of the camera location in the vehicle coordinates, m
f	focal length of the lens, m	y_k	observation obtained from an RTK-GPS at time instant k, m
$g(x)$	general crop model function	y_n	horizontal location of the navigation point in the vehicle coordinates, m
$\bar{g}$	mean of $g(x)$	z_c	longitudinal location of the 3D point in the camera coordinates, m
$h(x)$	data series of the crop profile obtained from an elevation map	z_d	vertical offset of the camera location in the vehicle coordinates, m
$\bar{h}$	mean of $h(x)$	α	roll angle of the camera, deg
IJ	elevation map axes	β	tilt angle of the camera, deg
m	number of data points in $g(x)$ series	δ	desired steering angle, deg
r	pixel resolution of the elevation map, m	ϕ	heading angle of a tractor with respect to target crop rows, deg
u_L	horizontal distance between a specific pixel in image L to the centre of the image, m	γ	pan angle of the camera, deg
u_R	horizontal distance between a specific pixel in image R to the centre of the image, m	v_k	measurement error of an RTK-GPS, m
v	vertical distance between a specific pixel to the centre of the image, m	ρ_x	root mean square of x_k, m
$\boldsymbol{R}_x(\theta)$	rotation matrix around X_c-axis	ρ_y	root mean square of y_k, m
$\boldsymbol{R}_y(\theta)$	rotation matrix around Y_c-axis	ρ_v	root mean square of v_k, m
$\boldsymbol{R}_z(\theta)$	rotation matrix around Z_c-axis	ρ_ω	root mean square of ω_k, m
w	inter-row spacing of the tested field, m	ω_k	RTK-GPS positioning error associated with a tractor inclination, m
x_c	horizontal location of the 3D point in the camera coordinates, m		
$X_cY_cZ_c$	camera axes		
x_d	longitudinal offset of the camera location in the vehicle coordinates, m		

gation system could guide agricultural machinery performing all field operations, including planting, cultivating, and spraying, at a travel speed of up to 3 m/s, with a tracking error of less than 0·05 m on both straight and curved paths (Kise *et al.*, 2001, 2002).

As a machine vision system can detect a path in relation to crop rows, vision-based guidance can be used to guide machinery travelling between crop rows to perform field operations such as cultivating, chemical spraying, and harvesting (Benson *et al.*, 2003; Tillett and Hague, 1999; Tillett *et al.*, 2002). In a vision-based guidance system, crop row features are extracted from acquired field images to obtain a guidance direction. The threshold method has been applied in many vision applications to separate objects of interest from imagery (Reid & Searcy, 1987). The major challenge of the threshold method when extracting crop row features reliably from field images is the difficulty in determining an adequate threshold value under varying ambient light conditions or changing crop growth stages. The effectiveness in distinguishing crops from weeds is another challenge in determining a pathway from obtained field images. Research has been reported in attempts to improve the reliability of crop feature extraction and pathway determinations for vision-based guidance systems. Hague and Tillett (2001) exploited a method using a bandpass filter to attenuate the grey levels of weeds and shadows in field images. Pinto *et al.* (2000) attempted to apply the principal component analysis method to extract crop row features from field images. Søgaard and Olsen (2003) developed a segmentation-free method for crop row detection and localisation.

To obtain more complete field information, a stereo-vision system can provide a three-dimensional (3D) field image by combining two monocular field images taken from a binocular camera simultaneously. Such a 3D image is reconstructed based on the difference between both monocular images, and therefore is less sensitive to

ambient light changes. Owing to its information intensive feature, potential applications for stereovision have been investigated. Some applications in agriculture include estimating physical parameters of a plug tray transplant (He *et al*., 2003), approximating the mass of swimming fish (Lines *et al*., 2001), and extracting the 3D shape of live pigs (Wu *et al*., 2004). This paper will report a crop row localisation method under varying ambient lighting conditions based on stereovision images. The following sections introduce the design of system architecture for the stereovision-based crop row detector, the algorithm formulation for processing stereovision images to localise crop rows, and the validation results obtained from field tests.

2. Design of stereovision-based crop row detection system

2.1. *System architecture*

The principal objective was to create the fundamental technology for a vision-based agricultural machinery guidance system, using a stereovision-based crop row detection system. *Figure 1* shows a machinery independent signal-flow diagram of a stereovision-based agricultural machinery navigation system. To validate the developed stereovision-based crop row localisation method without loss of generality, this navigation system was installed on a John Deere 7700 tractor research platform. This research platform was modified by installing a stereovision navigation system and adding an electrohydraulic steering system in parallel to the existing tractor steering mechanism in order to implement automated steering. The core element of the stereovision navigation system was an STH-MD1 (VidereDesign, CA) stereovision camera. A personal computer was used as the navigation computer for stereovision image processing, tractor navigation parameters calculation, and steering signal generation. A microprocessor-based vehicle control unit (VCU), consisting of a single board computer and two motor driver integrated circuits (ICs), was developed to implement automatic steering. The auto-steering actuating system consisted of a solenoid-driven proportional electrohydraulic direction control valve and a potentiometer-based wheel angle sensor for implementing closed-loop steering control. During auto-steering, the VCU received steering control signals from the navigation computer *via* a RS232 serial interface and generated pulse width modulation (PWM) signals to drive the electrohydraulic steering control.

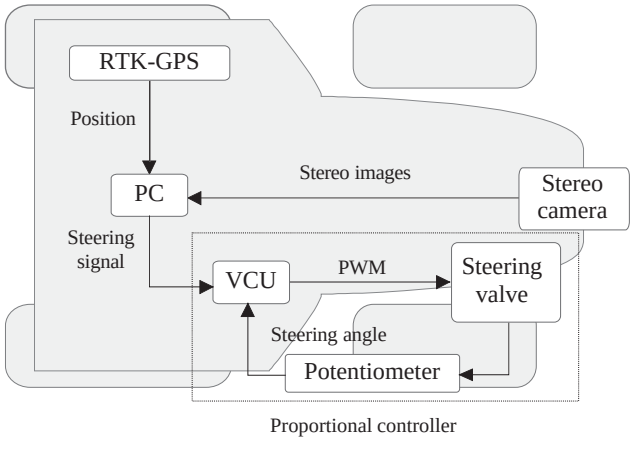

Fig. 1. Signal-flow diagram of the stereovision-based tractor navigation system: RTK-GPS, real time kinematic global positioning system; PC, personal computer; VCU, vehicle control unit; PWM, pulse width modulation

2.2. *Stereovision crop row detecting system*

The stereovision crop row detecting system was based on a stereovision camera. This stereovision system represents a 3D scene based on two plane images simultaneously taken with a binocular stereo camera. *Figure 2* shows the principle of stereovision image reconstruction based on two plane images. In this stereo geometry, the distance between the centres of two lenses is defined as the baseline b, and the focal length of the lens pair is defined as f. The horizontal distance between an identified pixel in image L to the centre of the image is defined as u_L and in image R is defined as u_R.

One of the major objectives in stereovision image processing is to calculate the depth of the object of

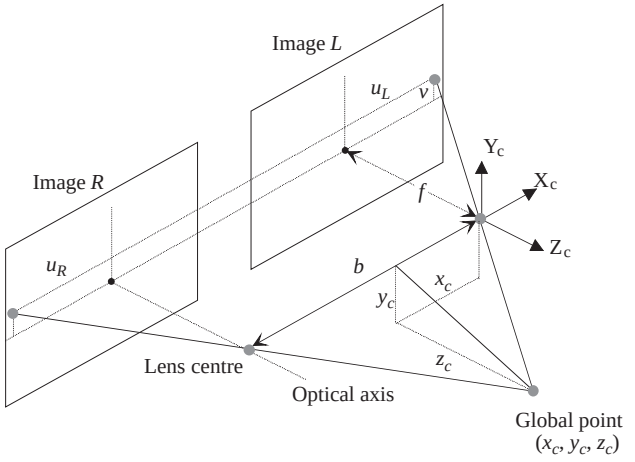

Fig. 2. Stereo geometry for stereo-image processing; global point, (x_c, y_c, z_c): u_L, horizontal distance between a specific pixel to the centre in image L; u_R, horizontal distance between a specific pixel to the centre in image R; v, vertical distance between a specific pixel to the centre of the image; b, baseline of the stereo camera; f, focal length of the lens

interest to the centre of stereo camera lens, z_c. When both image planes are laid on the same plane and their horizontal axes are aligned, the following equation can be used to calculate the object depth based on measured u_L and u_R.

$$u_L - u_R = d = \frac{bf}{z_c} \tag{1}$$

In Eqn (1), the variable d can be defined as the disparity of a stereo image. Since the disparity of a point is inversely proportional to its depth, the disparity image provides a direct (but inverse) encoding of scene depth (Scharstein, 1999). Once the disparity is determined, the 3D location of a point (x_c, y_c, z_c) in the left camera coordinates can be obtained as follows:

$$x_c = -\frac{u_L}{f} z_c = -\frac{bu_L}{d} \tag{2}$$

$$y_c = -\frac{v}{f} z_c = -\frac{bv}{d} \tag{3}$$

$$z_c = \frac{bf}{d} \tag{4}$$

This stereovision navigation used an STH-MD1 (VidereDesign, CA) stereo camera to acquire stereo images for determining the relative position between the tractor and crop rows. This stereo camera consists of two 1·3 mega pixel progressive CMOS imagers, being individually controlled *via* an IEEE 1394 interface, with 12·0 mm lenses. Depending on the resolution of the acquired image, this stereo camera can take 7 (resolution of 1280 by 1032 pixels), 25 (640 by 512 pixels), or 80 frames (320 by 256 pixels) per second. To better distinguish the vegetation from the scene, a near-infrared (NIR) filter (centre wavelength of 880 nm) was installed in front of each imager. The camera was configured a 50° horizontal field of view (HFOV) and a 38° vertical field of view (VFOV), and was installed in front of research platform 1·7 m above the ground level with a tilt angle of 31° (*Fig. 3*). With the introduced configuration and installation, this camera was capable of capturing field scenes between approximately 1 and 8 m in front of the tractor.

The original images taken by the stereo camera were transformed to the rectified images for stereo computation. This calibration process identifies the parameters related to the camera parameters and geometrical relationship between left and right imagers: the former includes the optical centre of the image, lens distortion, and the focal length; the later is the location of the right image plane with respect to left image plane. Commercial software provided by the camera manufacture was used to perform the calibration process and the stereo computation.

Fig. 3. Camera installation on research tractor platform

3. Development of stereovision-based crop row detection algorithm

3.1. *System overview and image preparation*

This research developed a stereovision-based crop row detecting algorithm for providing navigation information to guide agricultural machinery following crop rows in performing various farming operations. *Figure 4* illustrates the information flow of the algorithm: it employed a three-step process of stereo-image processing, elevation map creation and navigation point determination in searching for a row tracking solution accurately and reliably from a stereo images.

As shown in *Fig. 3*, the stereo camera was always mounted on the machinery with a tilt angle β. In many cases the camera installation could involve a pan angle γ and a roll angle α. To create a 3D field map for accurately representing relative position information between the machinery and crop rows based on the captured image, it is necessary to convert the locations of 3D points computed in the stereo process from the camera coordinates into vehicle coordinates. The axes XYZ define the vehicle coordinates with origin at a point on the ground underneath the tractor centre of the gravity (CG), and axes $X_cY_cZ_c$ the camera coordinates with origin at the centre of the left camera focal point (*Fig. 5*). Based on the camera configuration and installation described earlier, this stereovision system

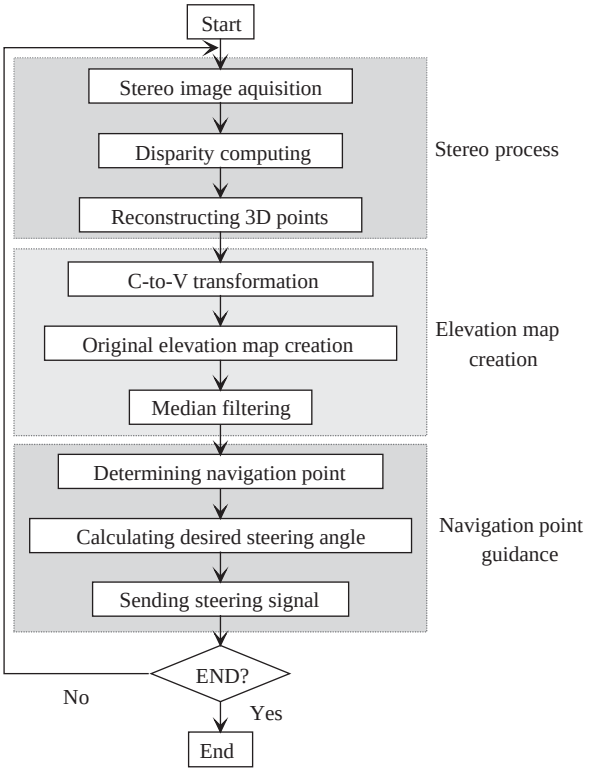

Fig. 4. Conceptual flowchart of stereovision-based crop row detecting algorithm: 3D, three dimensional; C-to-V, camera coordinates to vehicle coordinates transformation

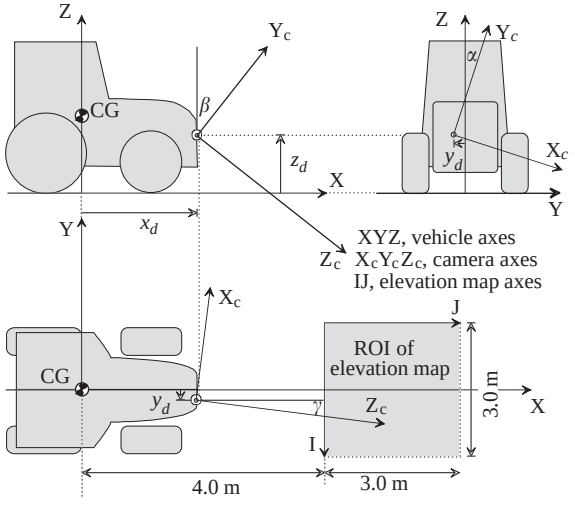

Fig. 5. Camera installation and image coordinate systems: CG, centre of gravity; ROI, region of interest; x_d, y_d and z_d, longitudinal, horizontal and vertical offsets of the camera location in vehicle coordinates; α, β and γ, roll, tilt and pan angles of camera installation

was capable of converting a field scene within a 3·0 m by 3·0 m square located in front of the machinery in to an elevation map.

To reconstruct a 3D field map in vehicle coordinates, it is necessary to first transform 3D points obtained from the stereovision system from camera coordinates to vehicle coordinates. This transformation uses six camera installation parameters, namely camera roll angle α, tilt angle β, pan angle γ, and position offsets (x_d, y_d, z_d), to represent the camera orientation and position in vehicle coordinates using the following equation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \boldsymbol{R}_y(\gamma)\boldsymbol{R}_x(\beta)\boldsymbol{R}_z(\alpha)\begin{bmatrix} z_c \\ x_c \\ y_c \end{bmatrix} + \begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix} \tag{5}$$

where, $\boldsymbol{R}_x$, $\boldsymbol{R}_y$, $\boldsymbol{R}_z$ are rotation matrices around X_c, Y_c, and Z_c axes as defined below:

$$\boldsymbol{R}_x(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{6}$$

$$\boldsymbol{R}_y(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

$$\boldsymbol{R}_z(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{8}$$

For the tractor platform used in this research, these six parameters were measured as 1·1°, 31·0° and −1·3° for roll, tilt, and pan angles and 2·66, 0·04 and 1·70 m for position offsets x_d, y_d, and z_d, respectively.

3.2. *Stereo-image process*

Stereo-image processing is used to determine 3D locations of the scene points of the objects of interest from the obtained stereo image. Those 3D positions, determined by means of stereo image disparity computation, provide the base information to create an elevation map which uses a 2D array with varying intensity to indicate the height of the crop. *Figure 6* shows an example of disparity computation and elevation map creation based on an actual soya bean field scene.

As shown in *Fig. 6(a)*, the original field scene image represents the soya bean rows with crop height approximately 0·4 m and row spacing of 0·75 m. The pixels in the original image represent some recognisable features of the field scene. The disparity map of an acquired stereo image [*Fig. 6(b)*] is computed by finding the corresponding points of a scene in both the left and

right images. Owing to the large number of pixels included in a stereo image, the computational load for disparity value determination is normally very large. In this research, a 320 by 240 pixel frame was chosen to collect the field scene images in order to reduce the computational load. Such an image frame will have approximately 76 800 pairs of pixels in the stereo image, with each pixel pair representing a spatial object at the corresponding location in the field. With the installation of the stereo camera on the tractor platform, this stereo image could cover a trapezoidal area of 4·4 m width on the top (the side far away from the tractor), 1·7 m width on the bottom (the side close to the tractor), and 5·5 m in length, located about 1·5 m in front of the tractor. Four crop rows can be seen within this image.

After converting the disparity map to 3D points in the camera coordinates, those 3D points were then transformed to elevation values at the corresponding locations in the vehicle coordinates as discussed previously. Such elevation values at specific locations (a small area of 0·02 m by 0·02 m in actual field) were used to form an elevation map of the field scene. Since the crop is always elevated from the soil, the pixels of an elevation map occupied by a crop must indicate larger values than the pixels occupied by the soil [*Fig. 6*(*c*)]. This conversion process (from disparity to elevation map) decreases the computational load by reducing the number of pixels from 320 by 240 points in a disparity image to 150 by 150 3D points for constructing an elevation map.

3.3. *Elevation map smoothing*

An elevation map is constructed using pixels with different intensities to represent the corresponding 3D values at the identified locations. While an elevation map can effectively represent the viewed objects, it cannot reveal the covered objects and results in an incomplete observation of the field scene. In addition, 3D points computed by a stereovision expressed by Eqn (1)–(4) get sparser as they are further away from the camera because a depth resolution of the 3D point is roughly proportional to the square of its depth. Both these limitations result in blank pixels for some locations, especially on the upper half of the image in a raw elevation map as shown in *Fig. 6*(*c*).

The missing data in an elevation map may generate unexpected noise in a crop row indicating parameters since such parameters are determined by the locations of elevated pixels. With the assumption of continuous crop plants, by which a crop plant is assumed being a continuous entity without a middle part of the plant being unexpectedly missed, a spatial filter was used to fill

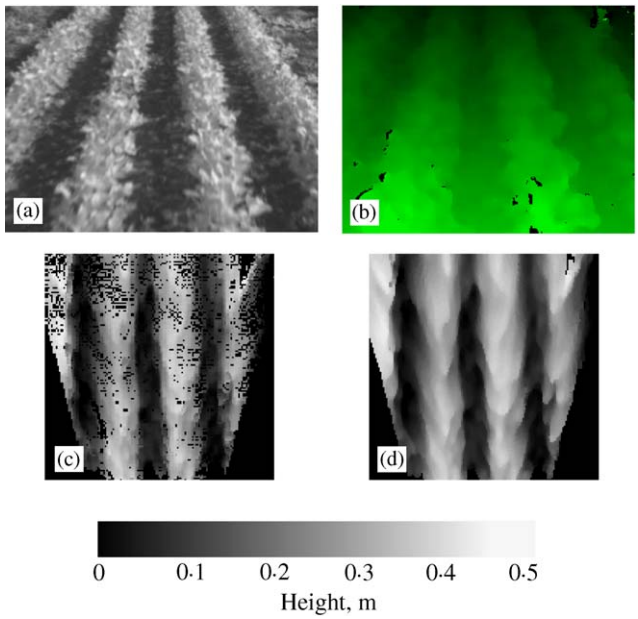

Fig. 6. Example of stereo-image processing with (a) an original image (acquired by left lens); (b) its disparity image; (c) a raw elevation map; and (d) a smoothed elevation map

the missing pixels of a field scene to provide a more accurate and stable crop row detection.

Due to the viewing direction of the stereo camera, the pixels of crop rows are often spatially located in the longitudinal direction. A rectangle (1 by 3 pixels) spatial median filter with longer side in the longitudinal direction was selected to enhance the quality of the raw elevation map. Using this approach, the filter is formed to replace the value of the centre pixel using the median value of all neighbouring pixels. In the case when the centre pixel is located on the border of the elevation map, the median value is computed based on remaining neighbouring pixels. *Figure 6*(*d*) shows a filtered elevation map processed using a 1 by 3 median filter. The resulting map indicates that the developed spatial median filter could effectively fill the blank pixels with reasonable estimates and result in a smooth elevation map for further processing.

3.4. *Guidance parameter extraction*

After being smoothed, the elevation map serves as a navigation map in this stereovision-based navigation system. To fulfil this function by extracting guidance parameters from this map, a cross-correlation function was developed to identify the navigation point and to calculate the desired steering angle.

Figure 7 illustrates the principle of the navigation point extraction. Using this approach, the navigation

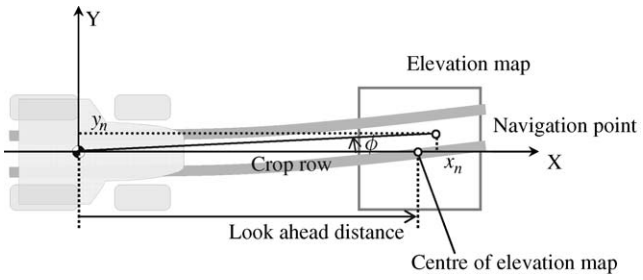

Fig. 7. The algorithm of navigation point guidance: x_n, longitudinal location of the navigation point in the vehicle coordinates; y_n, horizontal location of the navigation point in the vehicle coordinates; ϕ, heading angle of a tractor with respect to target crop rows

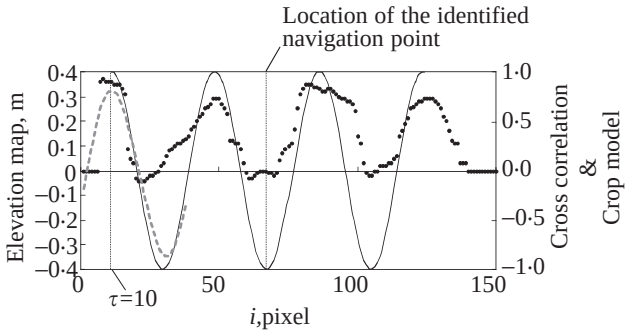

Fig. 8. Illustration of navigation point extraction by means of cross-correlation analysis: ……, elevation map; ——, crop model ($\tau = 10$); - - - -, cross-correlation

point extracted from an elevation map is similar to the way a human determines the target while driving a tractor. In this research, the centre of the elevation map was defined as the viewpoint for navigation, and the middle point between two crop rows being followed was defined as the navigation point. The longitudinal distance from tractor CG to the viewpoint was defined as the 'look-ahead distance'. When there is a lateral offset detected between the viewpoint and the navigation point, it indicates that a steering correction should be made.

After the navigation point is identified, the steering angle for correcting the offset can be determined using the following equations.

$$\phi = \tan^{-1}\left(\frac{y_n}{x_n}\right) \tag{9}$$

$$\delta = a\phi \tag{10}$$

where: (x_n, y_n) is an identified navigation point with respect to vehicle coordinates; ϕ is a heading angle of a tractor with respect to target crop rows; and a is a proportional steering control gain. The look-ahead distance and the proportional steering control gain are empirically identified as 4·5 m and 1·0, respectively.

To effectively identify the navigation point from the elevation map, a general crop model has been created to represent crop rows in the map. This crop model is designed to meet the requirement that the model can depict the inter-row spacing of the crop rows regardless of the height of the crop. This research found that a three-cycle cosine function was able to depict four rows of crop in an elevation map, and thus defined the general crop model:

$$g(i) = \cos(2\pi ri/w) \quad \left(0 \leqslant i \leqslant 3\frac{w}{r}\right) \tag{11}$$

where: $g(i)$ is crop model function; r is the pixel resolution of the crop elevation map and w is the inter-row spacing in m.

Based on this general model, the period of the cosine function was determined according to the inter-row spacing w of the crop (the value for w was 0·75 m in an example case), and the resolution r of elevation map (the value for r was 0·02 m). After the general crop model is defined, the next step is to extract the navigation point from the stereo field scene image by means of aligning the crop row profiles represented in a cross-section of obtained crop elevation map with the general crop model. *Figure 8* shows an example case of navigation point extraction when the tractor was travelling in a soya bean field with average plant height of approximately 0·4 m.

A cross-correlation matching method (Tsai *et al.*, 2003) was applied to determine the navigation point. Based on this method, the profile line representing the crop in the crop elevation map was selected as one data series and the general crop model function as the other series. During the analysis, a match algorithm will search for the optimal navigation point by adjusting the delay τ to maximise the correction index $c(\tau)$ between the crop profile and the general crop model using the following equation:

$$c(\tau) = \frac{\sum_{i=0}^{m-1}(h(i+\tau) - \bar{h})(g(i) - \bar{g})}{\sqrt{\sum_{i=0}^{m-1}(h(i+\tau) - \bar{h})\sum_{i=0}^{m-1}(g(i) - \bar{g})}} \tag{12}$$

where: $h(i)$ is the data series of crop profile obtained from the crop elevation map; $g(i)$ is the data series provided by the general crop model; $\bar{h}$ and $\bar{g}$ are the means of corresponding series; and m is the number of data points in g series.

During the search, Eqn (12) is repeatedly computed by changing the value of delays τ from zero to one complete cycle. The τ resulting in the maximum value

for $c(\tau)$ is selected as best matching point for the crop profile line and the general crop model line. As illustrated in *Fig. 8*, one complete cycle of the cosine curve model used for the example presented here consists of 38 discrete τ's. By sliding the general crop model function from a value for τ of 0 to 37, a maxim um correlation value (0·81) was reached at a value of 10 for τ.

The obtained optimal τ value from the cross-correlation analysis is then used to realign the general crop model. After the general crop model is realigned at the obtained optimal τ point (a value for τ of 10 in this example), the middle point of the general crop model is defined as the navigation point for the cross-section of the crop elevation map being investigated. The search starts at the centre cross-section of the elevation map (in this example the search was started at a value of 75 for *j*). By repeating the same process on cross-sections shifting away from the starting point to both sides (such as a value for *j* of 74, 76, 73, 77, …,), the navigation points for those cross-sections can also be identified. This search process is complete when the navigation point becomes greater than a threshold correlation value (in this research the threshold was set at 0·8). *Figure 9*(*a*) shows the navigation points identified on all cross-sections based on the elevation map presented in *Fig. 6*(*d*). *Figure 9*(*b*) reveals the variation of the obtained maximum cross-correlation values in the longitudinal direction. The variation range is between 0·5 and 0·9. As the distance from the camera increases, the cross-correlation value of the identified navigation points are stable at a level of 0·8 or greater and the identified navigation points are located near the centre of the inter-row space. In comparison, the cross-correlation values decline to 0·5 as the distance from the camera decreases. This is due to the blank pixels caused by the crop located behind the scene of the stereo camera. To obtain a consistent performance in identifying navigation points, a threshold value of 0·8 was selected because a cross-correlation value greater than this threshold representing a reliable match between the crop model and the crop elevation map. *Figure 9*(*c*) shows that the maximum cross-correlation value at the centre cross-section is reached at the centre of the inter-row space. It should be noted that the search process in the actual program was completed at the centre cross-section because the program obtained a cross-correlation (0·81) greater than the threshold (0·8) in this particular case. Both results indicate that this approach can accurately identify a navigation point based on the crop elevation map, and it is better to select the navigation point as close as the longitudinal centre of the elevation map for achieving more stable and accurate guidance from the perspective of navigation point guidance. In this research, the cosine function was used to represent soya bean crop at heights between 0·3 and 0·5 m. When the crop is in early stage or the crop size is very thin and the rows are sparse, other crop models, such as the rectangular wave function used for sugar beets, may be more suitable.

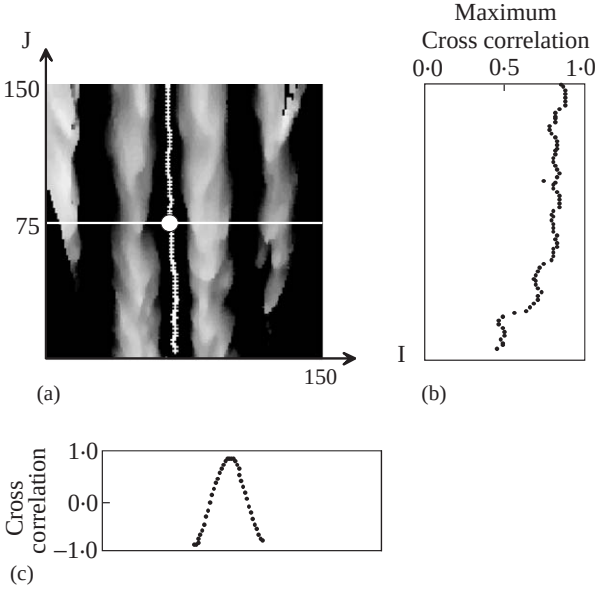

Fig. 9. Identified navigation points: (a) navigation points on the map; (b) maximum cross-correlation variation on longitudinal direction; (c) cross-correlation variation at the centre cross-section (white horizontal line in the map represents the centre cross-section); IJ, elevation map axes

4. Results and discussion

A series of automated tractor guidance tests in rowed soya bean fields were conducted to validate the performance of the developed stereovision-based navigation system. Sixteen rows of soya beans were planted using a four-row planter in four different patterns as listed in Table 1.

To provide the base information for evaluating the navigation system performance, the positions of those crop rows were measured using an RTK-GPS (horizontal accuracy 0·02 m, 1 sigma level) mounted on the cabin directly over the CG while a human driver navigated the path (called the reference path). The same RTK-GPS setup was used to record tractor trajectories during the validation tests. In this research, the guidance accuracy of the navigation system is quantitatively evaluated by comparing the tractor path measured using the RTK-GPS under automatic control (called the guidance path) to the reference path. It should be noted that this evaluation method had some inevitable error

Table 1
The crop row patterns (amplitudes and periods) for validating the stereovision-based crop row following agricultural machinery navigation system

Path	*Amplitude, m*	*Period, m*
A	0·80	58·3
B	0·50	52·0
C	0·25	58·2
D	0·20	80·5

factors, such as RTK-GPS measurement error and the GPS antenna positioning error associated with the tractor inclination. The mathematical model of this RTK-GPS based measurement system is expressed by

$$y_k = x_k + \omega_k + v_k \tag{13}$$

where: k is the time in discrete-time system, x_k is the lateral deviation of the guidance path from the reference path; ω_k is the RTK-GPS measurement error; v_k is the positioning error associated with the tractor inclination; and y_k is the observation obtained from the RTK-GPS.

The root mean square (RMS) of x_k is used to quantify guidance accuracy. Based on Eqn (13), the RMS of y_k can be determined using the following equation:

$$\begin{aligned}\rho_y^2 &= E\{y_k^2\}\\ &= E\{(x_k+\omega_k+v_k)^2\}\\ &= E\{x_k^2+\omega_k^2+v_k^2+2x_k\omega_k+2x_kv_k+2\omega_kv_k\}\\ &= \rho_x^2+\rho_\omega^2+\rho_v^2+2E\{x_k\omega_k+x_kv_k+\omega_kv_k\}\end{aligned} \tag{14}$$

where: ρ_y is the RMS of y_k; $E\{\}$ represents average; and ρ_x, ρ_ω, and ρ_v are RMS of corresponding parameters.

Assuming that ω_k and v_k are white noises uncorrelated with x_k:

$$E\{x_k\omega_k\}=0, \quad E\{x_kv_k\}=0, \quad E\{\omega_kv_k\}=0 \tag{15}$$

Substituting Eqn (15) into Eqn (14), the RMS of the guidance path deviation from the target path ρ_x can be determined by

$$\rho_x = \sqrt{\rho_y^2-\rho_\omega^2-\rho_v^2} \tag{16}$$

In Eqn (16), ρ_y is measured using the RTK-GPS and ρ_ω is the RTK-GPS measurement accuracy (0·02 m for the system being used in this research). To estimate ρ_v, a survey on the test field was performed and the result revealed that the field was nearly flat. The roll angle variation measured using a fibre optical gyro attitude sensor mounted on a tractor was within $\pm 0{\cdot}4^\circ$, and repeat trials on same path showed that the standard deviation in roll angle measurement was 0·05°. At this condition, the estimated ρ_v value was less than 0·002 m on the tractor platform (on which the GPS antenna was installed 2·2 m above the ground level) and was small enough to be ignored in Eqn (16).

To validate the effectiveness and the accuracy of the stereovision crop-row detecting algorithm in providing real-time navigation information, the tractor platform was guided automatically using the stereovision-based navigation system to track crop rows of different curvatures as defined in Table 1 at various travelling speeds between 1·0 and 3·5 m/s. In these field tests, the stereovision system updated the navigation information at 5 Hz, and the RTK-GPS recorded tractor CG trajectory at 10 Hz. Results obtained from validation tests verified that the developed stereovision-based navigation system could reliably and accurately detect navigation points for guiding the tractor following crop rows without overrunning the crop at normal field operation speeds.

Figure 10 shows the crop row tracking result obtained from one of the auto-guidance tests conducted on a near straight row (Path C defined in Table 1) at 2·5 m/s. A total of 333 stereovision images were collected and processed during this particular test, and all computed navigation points were located within the inter-row space. Field observation on this test indicated that the automatically guided tractor did not run over any soya

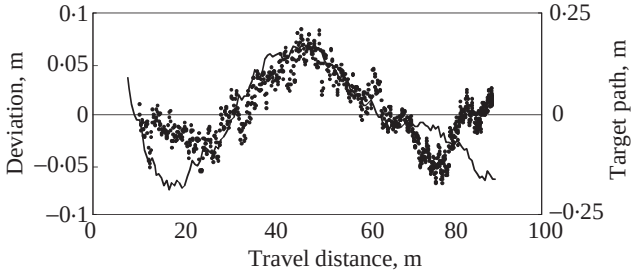

Fig. 10. Row tracking result of an auto-guided tractor following a straight path at 2·5 m/s: ……, lateral deviation of the guidance path to the target path; ——, reference path

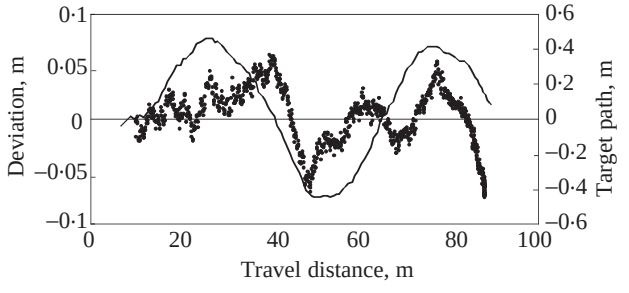

Fig. 11. Row tracking result of an auto-guided tractor following a curved path at 2·0 m/s; ……, lateral deviation of the guidance path to the target path; ——, reference path

bean plants. Quantitative analysis on the result indicated that the RMS of the lateral deviation between the guidance path and the reference path was 0·03 m. From *Fig. 10*, it can be seen that the lateral deviation to the reference path was periodically changed and shows a strong correlation to the reference path. This phenomenon was likely caused by the non-linearity of the steering system. Although a 0·03 m RMS error is favourable, the guidance accuracy could be improved by adopting enhanced steering control methods, such as feedforward plus proportional integral differential (PID) control and fuzzy control (Qiu *et al.*, 2001; Qiu and Zhang, 2003).

Similar results were obtained from field tests when the automatically guided tractor was following the path having a larger curvature (path B defined in Table 1). *Figure 11* presents the lateral deviation of the guidance path to the reference path while following path B at a travel speed of 2·0 m/s. All the computed navigation points during this curved row following test were located between the inter-row space which indicated a 100% accuracy in determining navigation points. A strong correlation between the guidance path and the reference path also existed due to poor performance of the steering controller. Very consistent results were obtained from other paths while the auto-guided tractor was travelling at speeds between 1·0 and 3·0 m/s, as shown in Table 2. In all field tests, the auto-guided tractor did not damage any crop. The path following tolerant was 0·10 m in tests having an inter-row spacing width between adjacent rows of 0·55 m and a tire width of 0·45 m.

Validation tests were also conducted on fields with significant amounts of weeds growing in between the crop rows. *Figure 12* illustrates the auto-guidance result obtained from one example of this case. The weeds had filled in most of the inter-row space, and were reflected in the elevation map. From *Figure 12(c)*, it is clear that weeds-free areas can be easily identified in an elevation map, even those areas are blocked or isolated by weeds. Identified weed-free areas can provide sufficient information to support the stereovision-based system to detect the navigation points [the white spot plotted in *Fig. 12(c)*]. At the cross-sections with weeds in between rows, the cross-correlation evaluation function always gives a correlation value smaller than the threshold value. By eliminating those points, the stereovision-based row detect algorithm can identify a correct navigation path and guide the tractor accurately. The validation tests conducted on weedy rows resulted in comparable navigation accuracy to those conducted on weed-free rows.

Table 2
The root mean square (RMS) error of lateral deviation of the guidance path from the reference path under various test conditions

Path	*Speed, m/s*	*RMS of the deviation, ρ_x, m*
A	1·0	0·04
A	2·0	0·05
B	1·0	0·03
B	2·0	0·03
C	1·0	0·03
C	2·0	0·03
C	3·0	0·03
D	2·0	0·04
D	3·0	0·04

5. Conclusions

This research investigated the fundamental technology for a stereovision-based agricultural machinery crop-row tracking navigation system. A crop-row detecting method, consisting of a stereo-image, processing module, a crop elevation map creating module and a navigation point searching module, has been developed and validated. Field test verified that a stereovision-based navigation system could reliably capture two field scene images simultaneously and extract the feature objects from both images to form a disparity image. Based on the obtained disparity image, a three-dimensional (3D) crop elevation map could be created to

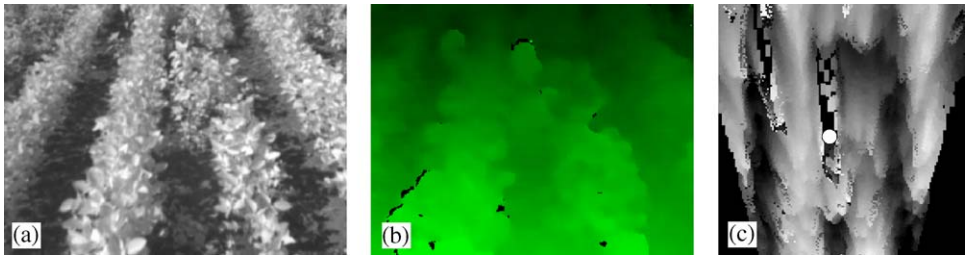

Fig. 12. Stereovision process for a weedy field with curved crop rows; (a) original image; (b) disparity image; and (c) elevation map with identified navigation point (white spot)

provide reliable and accurate tractor navigation information for crop-row tracking. This navigation system can effectively navigate a tractor following crop rows in a weedy field. Quantitative analyses on the field test results indicated that the root mean square (RMS) error of lateral deviation was less than 0·05 m when a tractor was guided automatically using the stereovision navigation system following on both straight and curved rows at a speed up to 3·0 m/s. Although this stereovision-based navigation system should be capable of detecting different types of row crops, this research has validated the system only with soya beans. Validating the applicability of the developed system in detecting various types of crops and/or at different growing stages is an interesting challenge for further investigation.

Acknowledgements

The material presented in this paper is based upon work supported by the USDA HATCH Funds (ILLU-10-352 AE). Japan Society for the Promotion of Science provided funds for Dr. Michio Kise performing post-doctoral research at the University of Illinois at Urbana-Champaign. Deere and Company provided a research tractor platform to support the development of the stereovision navigation system. All support mentioned is gratefully acknowledged. Any opinions, findings, and conclusions expressed in this publication are those of the authors and do not necessarily reflect the views of the University of Illinois at Urbana-Champaign or the USDA.

References

Bell T (2000). Automatic tractor guidance using carrier-phase differential GPS. Computers and Electronics in Agriculture, **25**(1–2), 53–66

Benson E R; Reid J F; Zhang Q (2003). Machine vision-based guidance system for agricultural small-grain harvester using cut-edge detection. Biosystems Engineering, **86**(4), 389–398

Hague T; Tillett N D (2001). A bandpass filter-based approach to crop row location and tracking. Mechatronics, **11**(1), 1–12

He D X; Matsuura Y; Kozai T; Ting K C (2003). A binocular stereovision system for transplant growth variables analysis. Applied Engineering in Agriculture, **19**(5), 611–617

Kise M; Noguchi N; Ishii K; Terao H (2001). Development of agricultural autonomous tractor with an RTK-GPS and a FOG. Proceedings of The Fourth IFAC Symposium on Intelligent Autonomous Vehicles, Sapporo, Japan, pp 99–104

Kise M; Noguchi N; Ishii K; Terao H (2002). The development of the autonomous tractor with steering controller applied by optimal control. Proceedings of Automation Technology for Off-Road Equipment, Chicago, USA, pp 367–373

Lines J A; Tillett R D; Ross L G; Chan D; Hockaday S; McFarlane N J B (2001). An automatic image-based system for estimating the mass of free-swimming fish. Computers and Electronics in Agriculture, **31**(2), 151–168

Pinto F A C; Reid J F; Zhang Q; Noguchi N (2000). Vehicle guidance parameter determination from crop row images using principal component analysis. Journal of Agricultural Engineering Research, **75**(3), 257–264

Qiu H; Zhang Q (2003). Feedforward-plus-PID controller for an off-road vehicle electrohydraulic steering systems. Journal of Automobile Engineering, **217**(5), 375–382

Qiu H; Zhang Q; Reid J F (2001). Fuzzy control of electrohydraulic steering systems for agricultural vehicles. Transactions of the ASAE, **44**(6), 1397–1402

Reid J F; Searcy S W (1987). Vision-based guidance of an agricultural tractor. IEEE Control Systems Magazine, **7**(12), 39–43

Scharstein D (1999). View Synthesis using Stereo Vision. Lecture Note in Computer Science. Springer-Verlag, Berlin

Søgaard H T; Olsen H J (2003). Determination of crop rows by image analysis without segmentation. Computers and Electronics in Agriculture, **38**(2), 141–158

Tillett N D; Hague T (1999). Computer-vision-based hoe guidance for cereals—an initial trial. Journal of Agricultural Engineering Research, **74**(3), 225–236

Tillett N D; Hague T; Miles S J (2002). Inter-row vision guidance for mechanical weed control in sugar beet. Computers and Electronics in Agriculture, **33**(3), 163–177

Tsai D; Lin C; Chen J (2003). The evaluation of normalized cross correlations for defect detection. Pattern Recognition Letters, **24**(15), 2525–2535

Wu J; Tillett R; McFarlane N; Ju X; Siebert J P; Schofield P (2004). Extracting the three-dimensional shape of live pigs using stereo photogrammetry. Computers and Electronics in Agriculture, **44**(3), 203–222