BIOSYSTEMS ENGINEERING 150 (2016) 89–103

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Multi-crop-row detection algorithm based on binocular vision

Zhiqiang Zhai, Zhongxiang Zhu*, Yuefeng Du, Zhenghe Song, Enrong Mao

Beijing Key Laboratory of Optimized Design for Modern Agricultural Equipment, China Agricultural University, Beijing, China

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form
3 June 2016
Accepted 14 July 2016
Published online 4 August 2016

Keywords:
Multi-crop-row detection
Binocular vision
Stereo matching
RANSAC method

Pathway determination is an important process in vision-based navigation. The pathway is very difficult to determine simply using 2D image processing, because fields are often infested with weeds, and images contain shadows, illumination variation, irregular backgrounds and other unexpected noise. Stereo vision techniques can be used to locate the spatial positions of crop rows for pathway determination. However, the stereo matching of field images is generally time-consuming and insufficiently accurate. To solve this problem, a multi-crop-row detection algorithm based on binocular vision is proposed in this paper. The algorithm is composed of the modules of image preprocessing, stereo matching and centreline detection of multiple crop rows. An accurate stereo matching method was put forward to locate the 3D position of crop rows based on the rank transformation, Harris detector and random sample consensus methods. A new method for detecting the centrelines of multiple crop rows was proposed according to their spatial distribution. The proposed algorithm was validated by comparative experiments. Regarding the proposed algorithm in situations without turnrows, the correct detection rate is greater than 92.78%; for the average deviation angle, the absolute average value is less than 1.05°, and the average standard deviation is less than 3.66°; for the processing time, the average value is less than 634 ms, and the average standard deviation is less than 101 ms. The results indicate that the proposed algorithm can satisfy the requirements of accuracy and real-time execution in field operation.

1. Introduction

Autonomous navigation for agricultural machinery is an important way to advance mechanization for modern agriculture in ways that can reduce labour intensity and improve operation efficiency and safety. Machine vision can detect a pathway in relation to crop rows or furrows and has been widely utilised as a condition awareness sensor for agricultural guidance systems (Kise & Zhang, 2008). A robust image processing algorithm for determining the pathway quickly and effectively is indispensable for achieving accurate machine-vision-based navigation. Machine-vision-based navigation can be realised by monocular vision and binocular vision using appropriate image processing. Many studies have been conducted to develop techniques for monocular and binocular vision navigation.

* Corresponding author.
E-mail addresses: srtkyxz@cau.edu.cn (Z. Zhai), zhuzhonxiang@cau.edu.cn (Z. Zhu).
http://dx.doi.org/10.1016/j.biosystemseng.2016.07.009
1537-5110/

Nomenclature

$\mathbf{A}$	Harris operator matrix
a_{ij}	element of $\mathbf{A}$
b	baseline distance, mm
$C(i, j, d)$	matching cost
$\mathbf{C}_D$	device coordinate system, pixel
$\mathbf{C}_I$	image coordinate system, mm
$\mathbf{C}_L$	left camera coordinate system, m
$\mathbf{C}_R$	right camera coordinate system, m
$\mathbf{C}_W$	global coordinate system, m
d	optimal disparity, pixel
$Det(\mathbf{M})$	determinant of $\mathbf{M}$
$\vec{e}$	direction vector of principal axis
f	focal length, mm
$f(i,j)$	greyscale of pixel
F_i	projection transformation function
g	number of video groups
$g(w_u, w_v)$	Gaussian filter function
h	camera height, m
$H(i, j)$	function of Harris corner point
$\mathbf{H}_i$	homography matrix
$h_0 - h_7$	elements of $\mathbf{H}_i$
(i, j)	coordinates of image point in $\mathbf{C}_D$, pixel
I_u	first gradient matrix in row direction of image
I_v	first gradient matrix in column direction of image
k	empirical value of Harris operator
$\vec{m}$	sample mean of S_p^*
$\mathbf{M}$	shape matrix of a pixel point
$\mathbf{M}_1$	transformation matrix from $\mathbf{C}_L$ to $\mathbf{C}_W$
$\mathbf{M}_2$	transformation matrix from $\mathbf{C}_D$ to $\mathbf{C}_I$
n	matching distance, pixel
n_j	number of correct processing image in each video
N_c	sample capacity of S_c^*
N_j	frame number in each video
N_i	sample capacity of S_i^*
p	confidence of model calculation
P_0	pixel point in image
P_l	projection of P_w on left image
P_r	projection of P_w on right image
P_w	point of field in $\mathbf{C}_W$
$\vec{P_i}$	coordinate matrix for line fitness
$r(i, j)$	value of rank transformation
r_L	rank transformation value of reference point
r_R	rank transformation value of matching point
R, G, B	red, green and blue chromatic values
s	size of S_i
S	sample set of RANSAC
S_c^*	final correct set of RANSAC
S_i	sample subset of RANSAC
S_i^*	consensus set of RANSAC
$\mathbf{S}_M$	scatter matrix for line fitness
S_p^*	candidate points of each crop row
$sgn(x_1, x_2)$	sign function
t	threshold of execution times
t_i	processing time of image, ms
$Tr(\mathbf{M})$	matrix trace of $\mathbf{M}$
(u_i^L, v_i^L)	coordinates of reference point of S_i in $\mathbf{C}_D$, pixel
(u_i^R, v_i^R)	coordinates of matching point of S_i in $\mathbf{C}_D$, pixel
(u_i', v_i')	estimated matching point of (u_i^L, v_i^L)
(u_i^l, v_i^l)	coordinates of P_l in $\mathbf{C}_D$, pixel
(u_i^r, v_i^r)	coordinates of P_r in $\mathbf{C}_D$, pixel
(u_o^l, v_o^l)	coordinates of origin of left image in $\mathbf{C}_D$, pixel
(u_o^r, v_o^r)	coordinates of origin of right image in $\mathbf{C}_D$, pixel
(w_u, w_v)	size of Gaussian window, pixel
(x_I, y_I)	coordinates of P_l in $\mathbf{C}_I$, mm
(x_L, y_L, z_L)	coordinates of P_w in $\mathbf{C}_L$, m
(x_W, y_W, z_W)	coordinates of P_w in $\mathbf{C}_W$, m
α_i	deviation angle in image, deg
θ	camera pitch angle, deg
μ_α	average deviation angle, deg
μ_c	average correct rate
μ_x, μ_y	pixel size of image, mm/pixel
μ_t	average processing time, ms
(ξ, η)	size of rank transformation window, pixel
(ε, λ)	size of NSAD window, pixel
σ_α	average standard deviation of deviation angle, deg
σ_t	average standard deviation of processing time, ms
σ_j^α	standard deviation of deviation angle in video, deg
σ_j^t	standard deviation of processing time in video, ms

1.1. *Monocular-vision-based navigation*

Monocular vision is more widely researched, having been utilised to guide machines in field operations, such as spraying, cultivating and harvesting. To improve the robustness of monocular image processing algorithms against environment noise and the efficiency of pathway determination, researchers have conducted extensive research and put forward several methods. For example:

(1) Methods based on expert systems. Montalvo et al. (2012) separated green crops from others with double thresholds based on Otsu's method and determined pixels to calculate the parameters of crop rows based on human knowledge on geometrical constraints of the vision system. The method can detect crop rows in maize fields with high weed pressure. Guerrero et al. (2013) further designed an automatic expert system for accurate crop row detection in maize fields based on human knowledge. The system consists mainly of an image segmentation module and a crop row detection module. Greyscale images are transformed with an excess colour index, and the positions of crop rows are initially estimated based on human input and finally modified with a Theil–Sen estimator. The system is validated to be effective to detect crop rows in maize fields with high weed pressure, but the algorithm consumes much time.

(2) Methods based on image preprocessing. Generally, images contain crops are infested with weeds, shadows, soil, gravels and so on. Appropriate image preprocessing is essential to suppress undesirable scene content. The grey levels of soil and the other non-green features can