Proceedings of the 2010 International Conference on Wavelet Analysis and Pattern Recognition, Qingdao, 11-14 July 2010

BASELINE DETECTION AND MATCHING TO VISION-BASED NAVIGATION OF AGRICULTURAL ROBOT

CUI-JUN ZHAO[1], GUO-QUAN JIANG[2]

[1] School of Resources and Environment Engineering, Henan Polytechnic University, Jiaozuo 454000, China
[2]School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China
E-MAIL: zhaocuijun@hpu.edu.cn, jiangguoquan@hpu.edu.cn

Abstract:

An automatic guidance model based on machine vision for detection and localization of crops rows is presented. The machine vision system consists of a color video camera and a computer. The camera is mounted on the head directly above the robot as a navigation sensor. When the agricultural mobile robot goes forward, the camera captures images continuously and transferred to the computer. First, pattern recognition and image processing were used to obtain quasi navigation baseline. Second, the real navigation line was extracted from quasi navigation baseline via Hough Transform. Then the mobile robot can be guided by the navigation line matching dynamically by itself. Test results indicate that the model has simple and robust algorithm, low-level requirements for software and hardware and was capable of navigating an agricultural autonomous robots traveling between crops rows.

Keywords:

Hough transform; Baseline matching; Path planning

1. Introduction

1.1. Background

A number of different sensor methodologies have been proposed or developed for agricultural machine guidance [e.g. global positioning system (GPS), inertial, machine vision, buried cable]. Today the guidance researches of agricultural mobile robot mainly concentrate on both machine vision and GPS which are of the most promising methods [1, 2]. In contrast to a real-time differential global positioning system (RTD-GPS), Machine vision is cheaper and has a higher precision. In addition to the virtues mentioned above, machine vision can provide local or relative information. So machine vision guidance was used in this project.

1.2. Main features of vision's information on a farm

There are two main features of vision's information on a farm:

Features one: On a farm, the physiognomy is often uniform and the species of crops are little and easy to differentiate. So, the images based machine vision will have a more clear contour and regular geometry graphics.

Features two: In general, each crop row (ribbing or furrow, etc.) on a farm is an approximate straight line and between crop rows (ribbings, or furrows, etc.) are parallel to each other. Thus, the crops rows (ribbings, furrows, etc.) can be used for mobile robot baselines [3]. These baselines can be described by y=kx+b. In fact, they are often called quasi-lines. That's to say, they don't mean real straight lines, because these baselines are approximate straight lines which formed by lots of dots. What's more, some segments of them maybe blurry or disconnected.

2. Model of machine vision processing

Agricultural mobile robot experimental equipment is primarily composed of four parts: computer, micro mobile vehicle and Charge Couple Device (CCD). When the micro vehicle is moving, CCD captures the images of crop rows continuously, navigation baselines can be obtained after image processing and Hough transform in a computer. Then according to the guidance information, the computer sends the anticipant oriented-wheel control angle to mobile vehicle. At the same time, system accepts the feedback information in order to adjust the mobile vehicle more effectively, as is shown in Figure 1.

Generally, transverse distance and navigation angle can statically describe the relative position of mobile robot. But the oriented-wheel control angle need the control information in time. Therefore, this model adopts linear feedback to count control variable by three state variables which are transverse distance, navigation angle and current fore-wheel control angle. The expression is as follows [4]:

$$a_d = K_1 e_l + K_2 e_h + K_3 a \tag{1}$$

Where e_l, e_h, a refer to transverse distance, navigation angle and current fore-wheel sheer angle respectively. K1, K2,

978-1-4244-6531-6/10/$26.00 ©2010 IEEE

K3 are the feedback parameters of three state variables correspondingly. a_d is the anticipant oriented-wheel control angle.

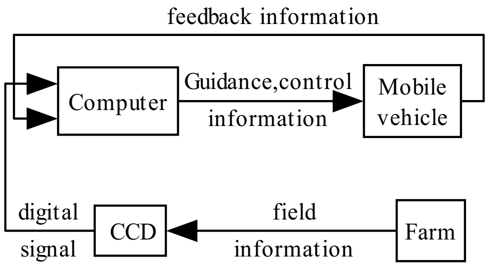


Figure 1. Machine vision navigation system model

3. Camera calibration, image collection and image processing

3.1. Camera calibration and image collection

When mobile robot goes forward, on one hand, CCD camera captures the images from a farm continuously, on the other hand, the coordinate changes from world coordinate (x_w, y_w, z_w) of scene to image coordinate (x', y'). The transform relations are as follows:

$$f:(x_w, y_w, z_w) \to (x', y') \quad (2)$$

$$(\hat{x}', \hat{y}') = (x', y') + w_r \quad (3)$$

Where expression (2) is the linear model and expression (3) is its nonlinear one. Thereinto, w_r is the white noise which maybe come from the uneven road, vibration of system or tiny curvature change of navigation baseline, etc.

The camera is mounted on the head directly above the robot. In the experiment period, w_r is not taken into account in order to simplify the research. According to relation of perspective projection, transform between world coordinate (x_w, y_w, z_w) of scene and image coordinate (x', y') can be expressed as [5]:

$$\frac{x'}{x_w} = \frac{y'}{y_w} = \frac{o}{z_w} \quad (4)$$

Where o refers to the focus of CCD lens. This equation expresses that the position of navigation baselines such as ribbings, furrows, stubbles or harvested crops can be obtained if CCD is fixed.

This system is used Microsoft Visual C++6.0 as developing language. In general, we select the image digitizer board and CCD camera according to actual circumstance. There is Software Development Kit (SDK) with image digitizer boards which can be used to capture the images. Of course, we can also use Video for Windows (VFW) which provided by Microsoft to capture images.

3.2. Image processing

It sounds a powerful tool for graphic element extraction from images for Hough transform in theory. However, the application of HT has been limited to small-size images for a long time. Besides the well-known heavy computation in the accumulation, the peak detection and the line verification become much more time-consuming for large-size images. So the images must be pre-processed in order to improve the real-time computing.

According to the images captured by CCD, in this study, the image processing adopts four steps: median filtering, binary-conversion of image with Otsu, dilation and edge detection. Experiment shows that this method of image processing is feasible and efficient.

- Median filter

As a nonlinear filter, median filter can suppress impulsive noise while preserving the edges of images. In this paper, we used median filter to enhance the image. The main image enhancement techniques are intensity adjustment and noise removal. The median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.).

The concrete steps are:

Giving each array pixel [i, j] a advisable neighborhood area, for example 3×3. We count respectively the average value of left-top corner neighborhood, left-bottom corner neighborhood, right-top corner neighborhood and right-bottom corner neighborhood. The lowest average value of them is used as the new value of the array pixel[i, j].

Average value expression:

$$v = \sum f^2(i,j) - (\sum f(i,j))^2 / N \quad (5)$$

- Binary-conversion of image with Otsu

Before binary-conversion of image, we should make gray-level transform with image, mostly, RGB linear transform, 2G–R–B linear transform (D.M. Woebbecke et al., 1995) are used widely.

Among the global thresholding techniques, the Sahoo (1988) study concluded that the Otsu method (Otsu, 1979) was one of the better threshold selection methods for general real world images with respect to uniformity and shape measures. According to the image features of a farm, we can regard ribbings, furrows or crop rows as a foreground, others as a background. The key is how to select appropriate threshold value to distinguish them. The binary-conversion of image with Otsu algorithm is competence for this difficulty problem.

The basic principle of Otsu is looking for an optimal threshold value to divide gray-level histogram of an image into two parts on the condition that between-cluster variance is maximal. Suppose there are 1~M gray levels in a gray image. Level i has n_i pixels. The number of total pixels is:

$$N = \sum_{i=1}^{M} n_i \tag{6}$$

The probability of occurrence of level gray i is defined as:

$$P_i = n_i / N \tag{7}$$

The average gray-level of the entire image is computed as:

$$\mu = \sum_{i=1}^{M} i \cdot P_i \tag{8}$$

In the case of single thresholding, the pixels of an image are divided into two classes $C_0 = \{1, 2, \cdots, k\}$ and $C_1 = \{k+1, k+2, \cdots, M\}$, where k is the threshold value. C_0 and C_1 are normally corresponding to the foreground and the background.

In class C_0, average gray-level is:

$$\mu(k) = \sum_{i=1}^{k} i \cdot P_i \tag{9}$$

The total pixels:

$$N_0 = \sum_{i=1}^{k} n_i \tag{10}$$

The probabilities is :

$$w_0 = \sum_{i=1}^{k} P_i = w(k) \tag{11}$$

In class C_1, average gray level is: $\mu - \mu(k)$, the number of total pixels is: $N - N_0$ and the probabilities is:

$$w_1 = 1 - w(k) \tag{12}$$

The mean gray-level values of the two classes can be computed as:

$$\mu_0 = \mu(k)/w(k) \text{ and } \mu_1 = [\mu - \mu(k)]/[1 - w(k)] \tag{13}$$

The average gray-level of the whole image is:

$$\mu = w_0\mu_0 + w_1\mu_1 \tag{14}$$

The between-cluster variance:

$$\begin{aligned} \sigma^2(k) &= w_0(\mu - \mu_0)^2 + w_1(\mu - \mu_1)^2 \\ &= w_0 w_1 (\mu_0 - \mu_1)^2 \end{aligned} \tag{15}$$

Equation (15) can be predigested as:

$$\sigma^2(k) = [\mu \cdot w(k) - \mu(k)]^2 / \{w(k) \cdot [1 - w(k)]\} \tag{16}$$

Select optimal k from 1~M on the condition that $\sigma^2(k)$ is the maximal. $\sigma^2(k)$ is the goal select function[6,7].

After k was selected, binary-conversion of image can use k as a threshold value.

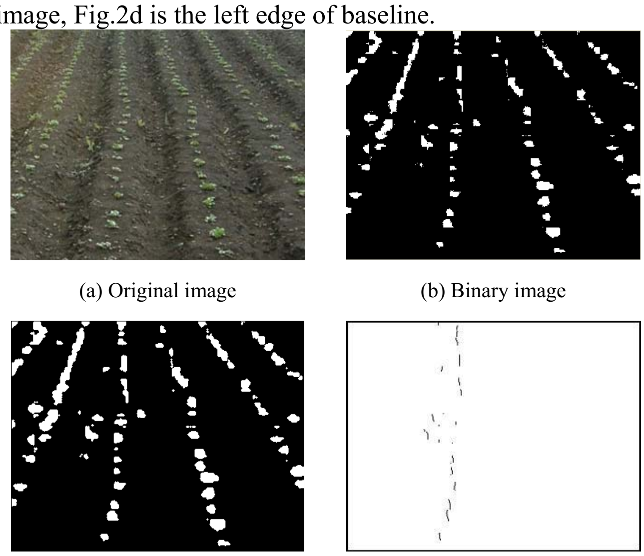
- Dilation of binary image

In order to reduce the noise in binary images, dilation of image is added. This paper uses 4-neighbors Dilation. That is, as long as one pixel of current pixel's 4-neighbors is white, the current pixel will be filled with white[8].

- Edge detection

After conversion of a binary, image of ribbings, furrows, stubbles or harvested crops has a certain width. We can use the left edge or right edge to obtain navigation baseline based on Hough transform. The approach for edge detection is: we can regard the pixels which gray change from 0 to 1 or 1 to 0 as the left edge or right edge.

The following are experiment images. Fig.2a is original image, Fig.2b is binary image, Fig.2c is the dilation of binary image, Fig.2d is the left edge of baseline.



(a) Original image (b) Binary image

(c) Dilation image (d) Edge detection image

Figure 2. Result of the image processing experiment

4. Hough transform and path designing

4.1. Hough transform obtain parameters of navigation baselines

After image processing, now the parameters of navigation baselines can be obtained by Hough transform. The Hough transform can be used to determine the parametrisation of straight lines and curves in an image[9~11].

The standard HT for straight line is depicted by expression 17, where (x, y) denotes a point in the Cartesian coordinates, i.e., the image space, and (ρ,θ) represents its parameter in the polar coordinates, also called the HT parameter space. All points on the same line in the image space will intersect at one point in the HT parameter space.

$$\rho = x\cos\theta + y\sin\theta \tag{17}$$

The concrete algorithm of HT is as follows:

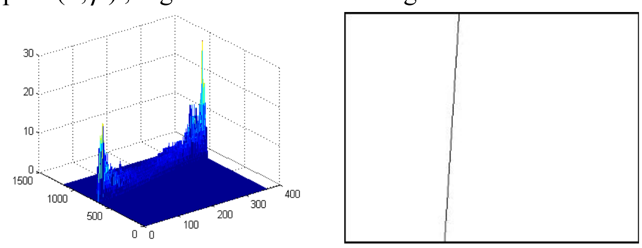
Step one. Construct a accumulator array. For straight line detection, the HT maps each edge pixel (x,y) from the image space into a parameter space of (θ,ρ), where contributions from each feature point to each possible set of (θ,ρ) are accrued. For this purpose the parameter space is divided into cells with each cell corresponding to a pair of quantized (θ,ρ). A accumulator array is often used to represent the quantized space.

In this system, for each edge pixel (x,y), vary θ from -90° to 90° and calculate $\rho = x\times\cos\theta + y\times\sin\theta$. The total quantizing dots are $\theta\times\rho$. Coordinate of each cell array is (θ,ρ). At the beginning, all cells are initialized by 0.

Step two. Scanning the whole binary image, if the current pixel is black, the corresponding elements in accumulator array add 1.

Step three. Scan the accumulator array and search the accumulator array for peaks. The peaks correspond to the parameters of the navigation baseline.

The following are the results of Hough transform experiment, where Figure 3a is the peaks of parameter space (θ,ρ), Figure 3b is the real navigation baseline.



(a) Peaks of parameter space (b) Navigation baseline

Figure 3 Result of Hough transform experiment

4.2. Path designing

When the mobile robot moves, on the one hand, the controller can obtain real time navigation information based on machine vision via image processing and Hough transform. On the other hand, the controller may get position information from the pose sensor on the wheel [12, 13]. So, the system regards the vision information as a norm, matching with pose information. When they both match well, the controller make no adjust. As long as both of them do not match with each other any more, according to the departure of parameter θ and ρ, the system quickly adjusts the pose of robot and guarantee it move forward along the right way.

5. Conclusions

In this paper we have presented an automatic guidance model based on machine vision. Firstly according to the farm's feature such as the uniform physiognomy, little crops species and nearly parallel crop rows, we used image processing method aim to distinguish between crop and background, further obtain image edge. Then we adopted Hough transform to exact navigation baseline and in the end we developed path designing to adjust the position of agricultural mobile robot. Experiment on a farm shows that this model can attain navigation parameters in real time and guide the mobile robot effectively.

Acknowledgements

This research was supported by the doctorial fund (Grant No. B2010-27).

REFERENCES

[1] T. Hague, J.A. Marchant, N.D. Tillett, "Ground based sensing systems for autonomous agricultural vehicles", Computers and Electronics in Agriculture, , Vol 25, No. 2 , pp. 11-28, January 2000.

[2] J.N. Wilson, "Guidance of agricultural vehicles a historical perspective", Computers and Electronics in Agriculture, Vol 25, No. 2, pp.3-9, January 2000.

[3] Yang WeiMin, Li TianShi, Jia HongShe, "Simulation and experiment of machine vision guidance of agriculture vehicles", Transactions of the Chinese Society of Agricultural Engineering, Vol 20, No. 1, pp. 160-165. February 2004.

[4] Zhou Jun, Liu Chengliang, Ji Changying, "Predictive tracking and control method for vision-guided navigation of agricultural robot", Transactions of the

Chinese Society of Agricultural Engineering, Vol 20, No.6, pp.106-110, December 2004.

[5] Zhu Rujun, "Study of Navigation and Positioning Technique Based on Computer Vision", Master Thesis in BeiHang University, 2004.

[6] N. Otsu., "A threshold selection method from gray-level histogram", IEEE Trans. SMC, Vol 9, No. 1, pp. 62-66, 1979.

[7] Zhang XiaoPing, FanJiuLun, Pei JiHong, "Two Improved Forms to Otsu's Histogram Thresholding Method", Journal of Xi'An University of Post and Telecommunications. Vol 8, No. 4, pp. 62-66, April 2003.

[8] Xu Hui. Visual C++ Digital Image Practical Projects Choiceness. Posts & Telecom Press, Beijing 2004.

[9] B.Chen, S.Tojo, K.Watanabe. "Machine Vision Based Guidance System for Automatic Rice Transplanters", Applied Engineering in Agriculture, Vol 19, No. 1, pp. 91-97, 2003.

[10] Rongben Wang. "A Vision Navigation Algorithm Based In Linear Lane Model", Proceedings of the IEEE International Conference on Intelligent Vehicles Symposium, Vol 4, pp. 240-245, 2000.

[11] Hough, P.V.C. Method and Means for Recognizing Complex Patterns. U.S. Patent No. 3069654, December, 1962.

[12] E.R.Benson, J.F.Reid, O.Zhang.Ma. "Machine Vision-Based Guidance System For An Agricultural Small-Grain Harvester", Transactions of the ASAE, Vol. 46, No. 2, pp. 551-558, 2003.

[13] T.Torii, T.Teshima, T.Okamoto, K.Imou, K.Taniwaki. "Autonomous Navigation Of Rice Husbandry Vehicle Using Machine Vision", ASAE/CSAE-SCGR Annual International Meeting, Paper No. 993005, 1999.