Target Detection and Recognition Method of Farming Machine Based on Machine Vision

Jiang Jiajun[1,a], Gui Xiangquan[1]

a.School of computer and communication, Lanzhou University of Technology, Lanzhou, 730050 China

a.Email: jjj4email@sina.com

Key words: farm; farming machine operation; background detection; machine vision

Abstract. In operation system of farming machine, the moving target detection is a key part of the monitoring system. Aimed at the shortcoming of poor detection effect of Gaussian background modeling, this paper proposes a detection method based on non-reference background model. The model utilizes the previous series of sampling values of current pixel to estimate the probability model of observed pixel; then based on the probability model to carry on binary detection of moving object. In terms of moving target recognition, this paper puts forward several characteristics, and trains identification through BP neural network. The experimental results show that the background model can effectively detect the moving target, and also achieve the satisfactory effect of recognition rate of farming machine based on machine vision.

Introduction

With the development of economy and the progress of the society, the farming machine monitoring has become a very important topic. Using camera for farm monitoring has been widely applied in social life, but in many cases, it is not appropriate for people to analyze the video sequences. Visual monitoring technology based on computer vision and image processing provides a more advanced and feasible method, its wide application prospect and potential commercial value have attracted the attention of the academic and business both at home and abroad. This paper studies the detection and recognition on moving target in farming machine operation monitoring, and puts forward some improvement methods. In view of defects of the traditional detection method like poor performance of anti-interference of external factors such as background light, etc., this paper puts forward a kind of non-reference background modeling method, this method can detect the foreground target effectively. In the aspect of farming machine target recognition, this paper puts forward several features easy to be identified, which also reached very high recognition rate through BP neural network training and test.

Moving target detection

In the operating system of farming machine, it usually uses a static camera to conduct real-time monitoring on specific locations of outdoor. Because the camera is static, so it usually uses a typical detection method of moving region, background difference method, namely the detection of moving region through comparing the current image and the background model. But the outdoor environment, due to the influence of external factors, such as leaves shaking, changes of shadows, and illumination, etc. the background region is not completely static. Therefore, how to establish the background model becomes the key of background detection method[1,2].

To extract moving target from video sequences shot by static camera is the first step in moving target recognition. The most simple background model is a pixel gray value imitated by a Gaussian distribution, this model is suitable for slow changing background, while cannot be used in handling multiple backgrounds, for example, there is shaking tree in the background. For applying in farming machine operation, the pavement, shadows, and crops can be imitated by Gaussian distribution, respectively, background pixel gray value can be imitated by these three mixed Gaussian weighted

values, the parameters of the model can be learnt and updated by increasing maximum expected algorithm. However, despite the pixel gray value is imitated by three Gaussian mixtures, the pavement is a single model assumption, etc[3,4]. In view of the complex and dynamic background introduced multiple Gaussian mixture model, however, for rapidly changing background this method cannot detect moving region accurately through a mixture of several Gaussian distributions. In addition, for background learning, multiple Gaussian mixture models have to be compromised. When the low learning rate is adopted, this model cannot respond to a sudden change of background; on the other hand, the slow foreground target will be misjudged as background. This is foreground aperture problem[5]. In order to overcome the above problem, this paper adopts a kind of non-reference background model. The experimental results show that the model can adapt to changes in the background well, and can detect the foreground target accurately.

Establishment of non-reference background model

Density estimation

The purpose of the model design is to capture the latest information in image sequence, and to capture the rapidly changing background by continuous information update. So if you want to get the exact detection, at any time, we must estimate the distribution density function through the recent historical information.

Assume $x_1, x_2, ..., x_N$ is the previous pixel value of current pixel point, x_t is the pixel value the pixel point at time t. Where, each pixel value is composed by pixel values of three color channels. Its probability density function can be non-reference estimated by the following kernel estimation function.

$$P_r(x_t) = \frac{1}{N} \sum_{i=1}^{N} K(x_t - x_i)$$

If the kernel estimation function K is Gaussian function, its probability density can be estimated by the following formula.

$$P_r(x_t) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t - x_i)^T \Sigma^{-1}(x_t - x_i)}$$

Where, Σ is covariance matrix. If σ_j^2 represents the variance of the j th color channel, and each channel is mutual independence, then

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

The density estimation can be simplified to

$$P_r(x_t) = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2} \frac{(x_{tj} - x_{ij})^2}{\sigma_j^2}}$$

It can be judged from the above formula, the pixel point is foreground or not. If $P_r(x_t) < th_1$, the pixel point is divided as foreground point; otherwise, it is divided as background point. Here th_1 is threshold value.

Variance estimation

There are two major causes for the changes of the brightness of the pixel value. One cause is the great leap of the brightness of this point created by different targets mapping on the same pixel point at different moments. Another cause is for the same target mapping on the same pixel point in a short period, the change of local brightness generated by image fuzzy. Covariance matrix Σ shall reflect the change of local brightness caused by image fuzzy, rather than the leap change of

brightness value. For different color channel, its variance is different due to local change varies.

In order to estimate the variance σ_j^2 of the j th color channel, it should calculate the difference $\left| x_{i,j} - x_{i+1,j} \right|$ of each pair of adjacent sampling values in samples first, then get the mid-value m_j of this series of difference value. Due to $x_{i,j}, x_{i+1,j}$ are two continuous brightness values，So usually they follow the same distribution, only some follow cross-distribution. Here, we assume the brightness follows normal distribution $N\left(\mu, \sigma^2\right)$, then，$\left(x_{i,j} - x_{i+1,j}\right)$ also follows normal distribution $N\left(0, 2\sigma^2\right)$. Therefore, the variance σ_j^2 of the j th color channel can be estimated by the following formula:

$$\sigma_j^2 = \frac{m_j}{0.68\sqrt{2}}$$

Because the differences are all integer values, so we can obtain a more accurate mid-value by linear fitting.

3.3 Restrain error detection

For the outdoor environment with changing background, there are two main sources of error detection. First, the uniform random noise exists in whole image; second, there is no background movement shown in the background model. In addition, the offset of the camera can also cause many error detections. If some pixel points in the background would be detected as foreground points due to movement occupying new pixel points, and not be included in model. But this moving part of background has a high probability at the original pixel point, so it belongs to background distribution. In general, only a small offset between two adjacent frames, so we can decide whether this point is an error detection due to background changes by the background distribution of adjacent region of pixel point detected in previous stage.

Assume x is the foreground point obtained by previous background detection, $x_{t,j}$ is the value of the j th color channel of this pixel point at t moment. Then the probability of pixel offset can be defined as：

$$P_N\left(x_t\right) = \max_{y \in N(x)} \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2}\frac{\left(x_{t,j} - y_{b,j}\right)^2}{\sigma_j^2}}$$

Where, $N(x)$ is the background point adjacent to pixel point x, $y_{b,j}$ is the value of the j th color channel of background point y. If $P_N > th_2$, this foreground point is divided into background, so the error detection caused by small background movement can be eliminated in this way.

Experimental result

Target detection

We shoot a series image sequence as experiment sample at the fixed angle of intersection of farm road, respectively. Take 15 frames image per second, image size is 240×320.Experiment system is established on 1 G internal storage of Intel PIV-3GCPU PC, and make experimental simulation with Matlab 7.1.For outdoor scene, background pixel value changing with time, the changing condition of grey level of some pixel at 900s as shown in figure 1. Figure 2 is the performance comparison of the Gaussian model and the background model, it can be seen that with the increase of the contrast ratio of target and background, the error rate of non-reference background model is significantly lower than that of Gaussian model. The experimental results of target detection

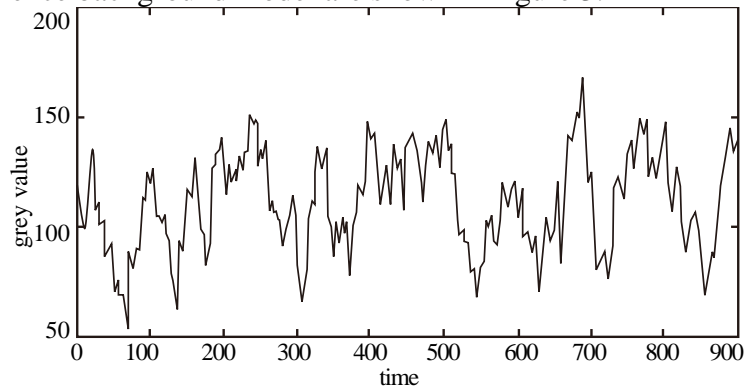obtained by non-reference background model are shown in figure 3.

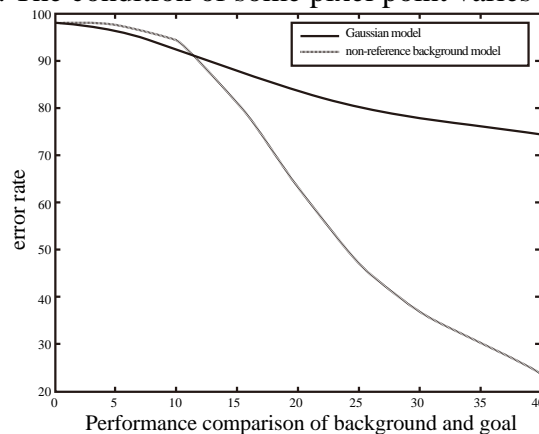Figure 1. The condition of some pixel point varies over time

Figure2. Performance comparison of non-reference background model and Gaussian model

Target recognition

This paper selects square stool to simulate agricultural robot, the installation height of cameras of both side is about 47.5 cm, the horizontal Angle of the left and right camera is about 35.7°，baseline distance is 18 cm, IPC (industrial personal computer) CPU is 2.5 Ghz, with 8G internal storage, at the same time, make a plumb line centered on the left camera, the intersection point of plumb line with the ground determined as the origin O of world coordinate system, X axis along the body level of agricultural robot to the left, Y axis along the optical axis direction of the camera perpendicular to X, Z axis is upward and vertical with the ground, as shown in figure 3
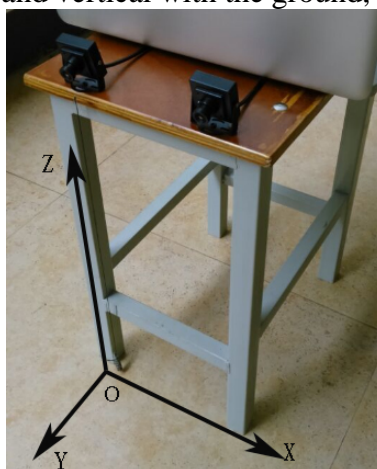

Figure 3. Binocular system of agricultural robot set up in this paper

Due to all selected navigation reference points are on the field ground, so let Z plane always equals to 0, i.e. no matter where agricultural robot to march on, the world coordinate system is established by this method, thus the obtained navigation reference point relative to the world coordinates of agricultural robot body.

This paper collects crops as sample, 48 samples of each crop, among which 36 samples of each

category used for training study, the other 12 samples for test. Set target error of network training is 0.05, the initial weight randomly selects between（-1，1）, coefficient of inertia item is 0.95, the initial value of learning rate is 0.1, the increased learning rate is 1.05, the decreased learning rate is 0.7, and maximum error ratio is 1.04.

To conduct training study for sample set, learning time is 130 s, part of sample data of the training set as shown in table 1. Test the testing set samples with trained network, the mean square error is 0.0247, the maximum error component is 0.2164, the recognition rate is up to 92%. Sample data of specific test set and test results are shown in table 2.Convergence curve as shown in figure 4.
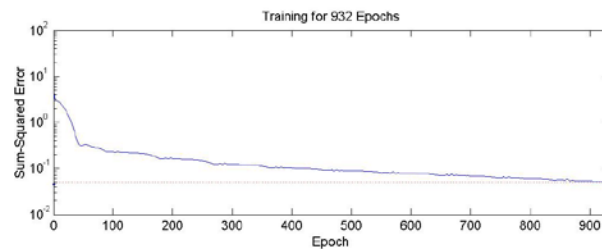

Figure 4. BP network training convergence curve

Conclusion

In the agricultural monitoring system, moving target detection and recognition is the key technique in the operating system of farming machine. Based on the summary of previous background modeling, this paper adopts a kind of detection method based on non-reference background model. This model utilizes a series of latest sampling values to estimate the probability model of observation pixel point; then through the binarization detection of moving target on the probability model, finally it obtains good detection effect via noise suppression processing. In terms of moving target recognition, this paper puts forward several characteristics, and conducts training study by BP neural network. The experimental results show that a high recognition rate can be achieved through the classification of crops with this characteristic.

Reference

[1] Jinyu Hu, Zhiwei Gao and Weisen Pan. Multiangle Social Network Recommendation Algorithms and Similarity Network Evaluation[J]. Journal of Applied Mathematics, 2013 (2013).

[2] Shuang Zhou, Liang Mi, Hao Chen, Yishuang Geng, Building detection in Digital surface model, 2013 IEEE International Conference on Imaging Systems and Techniques (IST), Oct. 2012

[3] Jie He, Yishuang Geng, Kaveh Pahlavan, Toward Accurate Human Tracking: Modeling Time-of-Arrival for Wireless Wearable Sensors in Multipath Environment, IEEE Sensor Journal, 14(11), 3996-4006, Nov. 2014

[4] Jinyu Hu and Zhiwei Gao. Distinction immune genes of hepatitis-induced heptatocellular carcinoma[J]. Bioinformatics, 2012, 28(24): 3191-3194.

[5] Na Lu, Caiwu Lu, Zhen Yang, Yishuang Geng, Modeling Framework for Mining Lifecycle Management, Journal of Networks, 9(3), 719-725, Jan. 2014