

Design of Plant Growth Environment Monitoring System Based on Data Fusion Algorithm

Shuwei Liu,⁺ Liguang Tian, Meng Li and Hang Ding

Tianjin Key Laboratory of Information Sensing & Intelligent Control, Tianjin University of Technology and Education, Tianjin, 300222, China

Abstract. In the course of the development of facilities agriculture, plant factory is an important part. Therefore, its intelligence has become the main goal of research. This paper describes the design of a plant growth environment monitoring system, with STM32 microcontroller as the main control module which can realize the monitoring and controlling of the environmental factors in the small plant factory temperature, humidity, carbon dioxide concentration. Furthermore, it uses the data fusion algorithm to improve the detection accuracy of the sensor and PID control to improve control stability. The test results show that the system is stable and reliable, and can meet the design requirements.

Keywords: plant factory, STM32, data fusion, PID control.

1. Introduction

With the development of agricultural facilities, more and more people begin to study plant factory for the construction of a precision equipment production system with high investment and high technology, the separation between agricultural production and natural ecological chains, as well as a plant production factory farming system following the annual plan. The plant in the facility can stay in the healthiest state, for using intelligent computer and electronic sensing system to the automatic control on plant growth temperature, humidity, light intensity, illumination time, CO₂ concentration and nutrient solution conditions [1].

In the plant factory, for monitoring the three main environmental factors influencing the growth of crops including temperature, humidity and CO₂ concentration, this paper designs a system based on STM32 MCU, with multiple sensors to detect the above-mentioned factors. To ensure the most suitable conditions in the plant factory, the system adopted the BP neural network data fusion method for the detection of data fusion and PID control to adjust the environment.

2. Overall Design of the System

2.1. Overall structure of the system

In this paper, a small plant growth cabinet is used as a plant culture device. The design system includes four modules: the main control module, the human-machine interaction module, the sensor module and the execution module. The main control module uses a low power embedded system to achieve the integration of data and the control of the execution module. The human-machine interaction module uses Labview software to build the interface which can directly set the environmental parameters, display the current measured values, and preserve historical data. The sensor module uses a temperature and humidity sensor and a carbon dioxide sensor to collect the environmental information in the cabinet. The execution module regulates the

⁺ Corresponding author. Tel.: 18902002115.
E-mail address: tlg1234@163.com.

environmental parameters through the environment heater, the fan, and the carbon dioxide generator. Structural diagram of control system can be shown as Figure 1.

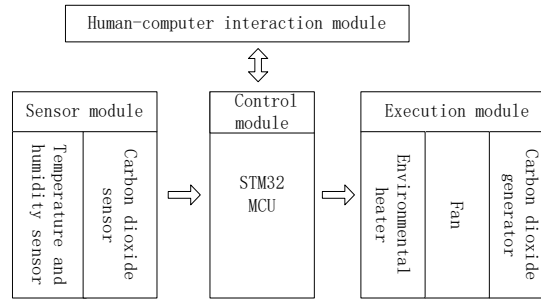


Fig. 1: Structural diagram of control system

2.2. System workflow

After the system starts working, the environmental indicators such as temperature, humidity, carbon dioxide concentration can be set through the man-machine interaction module. Then the data will be send to the main control module. The main control module will gather the information for temperature, humidity and the concentration of carbon dioxide in the cabinet through the sensor module. Then, the detection result will be reflected in the human-computer interface. Meanwhile, the result data will be compared with the setting indicators. If the index does not meet the requirements, the main control module instructs the execution module to perform the corresponding action of equipment module. The sensor module will continue the action of testing until it meets the set requirements so that the ecological environment suitable for crop growth can be simulated.

3. System hardware design

3.1. Main control module

The core of the hardware system is the main control module. In this paper, STM32F103 microcontroller is selected. It is a STM32 processor based on ARM 32 bit Cortex-M3. The highest frequency is up to 72MHz, which has 512 byte Flash capacity and 64K byte SRAM [2]. The processor has high-speed processing speed which can provide different communication ports and various interrupt response modes. It can fully fulfill the functions of the design with high precise timing counters, A/D conversion port and library functions.

3.2. Sensor module

Sensor modules include two types of sensors: One for temperature and humidity, the other for carbon dioxide. The temperature and humidity sensor adopt a digital temperature and humidity sensor chip SHT11 with self-calibration and digital signal output. It contains 14 bits of A/D converter and transmits data with quasi I₂C interface [3]. We use software to simulate I2C communication timing to receive data. Furthermore, it also has the internal function of calibrating digital signal. The range of temperature and humidity is - 40°C ~ + 110°C, 0-99.9%RH respectively. The clock pins and data pins of the temperature and humidity sensor module are connected to the PC4 and PC5 interfaces of the STM32 processor respectively. The circuit connection diagram of the temperature and humidity sensor is shown in Figure 2.

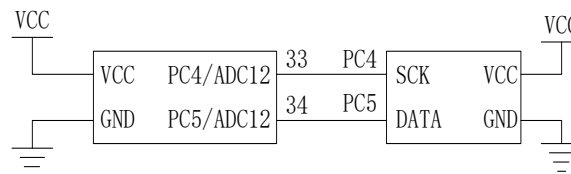


Fig. 2: Circuit diagram of temperature and humidity sensor circuit

Carbon dioxide sensor is a MG811 sensor with good stability and reproducibility. The sensor is based on solid electrolyte battery with little influence on temperature and humidity. It has good sensitivity and selectivity to carbon dioxide. The measurement range is 0-10000ppm while the analog output range is 0-4V.

4. System Software Design

4.1. Human-computer interaction module

The human-computer interaction module, as the display platform of the main control module, also plays a great role in the system. In order to implement the interface function in a more preferable way, this article uses Labview software to design. The appearance of the controls provided is similar to the actual hardware, making the interface more concise and intuitive. Block diagram programming form makes the program flow more clear and maintainable.

There are three main parts of the interface: The user login interface, the monitoring interface, and the system settings. The specific functions are shown in Figure 3.

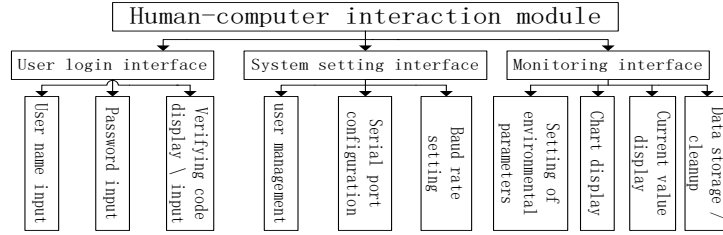


Fig. 3: Function diagram of man-machine interaction module

4.2. Data preprocessing

Because of the sudden disturbance in the growth cabinet, the performance of the sensor itself and the interference of other factors in the growth cabinet, false or even erroneous measurements will inevitably be produced during the measuring process. Then, it will lead to an error value which can be divided into systematic error, random error and rough error. The gross error needed to be eliminated as it is the major factor affecting the measurement data. There are two kinds of methods to test gross error, Pauta criterion and Grubbs criterion. Because the number of sensors in this paper is less than 10, it is more suitable for the Gra Bruce criterion based on the distribution.

Firstly, the Grubbs criterion arranged the measured values in order of size. Secondly, when the calculated Grubbs criterion number T_i is greater than the Grubbs criterion in the table number $T(n, \alpha)$ according to the sample size n and the selected judgment significance level of α , the measured value of X_i may be considered as a gross error value which should be removed. The Grubbs criterion number T_i formula is as follows.

$$T_i = \frac{|V_i|}{S} = \frac{|x_i - \bar{x}|}{S} \quad (1)$$

4.3. Data fusion algorithm

The definition of sensor data fusion can be summarized as the integration of the local distributed data resource provided by a number of similar or different sensors. Based on the analysis with computer technology, the probable redundancy and contradiction between multi-sensor information can be eliminated. Afterwards, the data can complement each other to reduce the uncertainty and improve the rapidity and correctness of the system in decision-making, planning, and responding [4]. As a result, the system can obtain more sufficient information. Because the system uses a variety of heterogeneous sensors (more than one for each type), so a hybrid fusion structure needed to be adopted. The frame diagram is shown in figure 4.

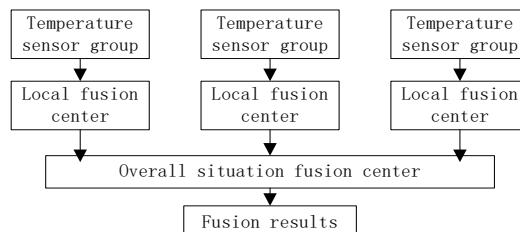


Fig. 4: System chart of fusion structure

For the fusion of the same type sensor, the original data collected by the sensor can be fused directly in the data layer without processing. For different types of multi-sensor information, it need to fuse in the feature layer or the decision layer. During the feature level fusion, we should extract features from sensor data to comprise feature vectors which should be input to pattern recognition processing module in the next step. Finally, we use neural network, clustering algorithm or template matching to recognize. The system uses Bayesian estimation method for the fusion of individual environmental factors. Meanwhile, the global fusion center adopts BP neural network method.

4.3.1. Bayesian estimation

Bias estimation is based on statistical data fusion algorithm. Bias statistics theory makes full use of prior knowledge which regards every inspection process as a continuous revision process of prior knowledge^[5]. If the data collected by two similar sensors at the same time are $X=\{X_1, X_2\}$, the fusion data is used as the Bias formula.

$$f(\mu | X_1, X_2) = \frac{f(\mu, X_1, X_2)}{f(X_1, X_2)} \quad (2)$$

If the parameter $\mu \sim N(\mu_0, \sigma_0^2)$, and $X_k \sim N(\mu, \sigma_k^2)$, So you can get it.

$$\begin{aligned} f(\mu | X_1, X_2) &= \frac{1}{f(X_1, X_2)} \prod_{k=1}^2 \frac{1}{\sqrt{2\pi\sigma_k}} \exp\left\{-\frac{1}{2}\left(\frac{X_k - \mu}{\sigma_k}\right)^2\right\} \times \frac{1}{\sqrt{2\pi\sigma_0}} \exp\left\{-\frac{1}{2}\left(\frac{\mu - \mu_0}{\sigma_0}\right)^2\right\} \\ &= \alpha \exp\left\{-\frac{1}{2} \sum_{k=1}^2 \left(\frac{X_k - \mu}{\sigma_k}\right)^2 - \frac{1}{2} \left(\frac{\mu - \mu_0}{\sigma_0}\right)^2\right\} \end{aligned} \quad (3)$$

It is known from the upper formula that the exponential distribution is the quadratic function of μ under the condition of known measured data. So (4) also conforms to the normal distribution, if $\mu \sim N(\mu_N, \sigma_N^2)$, then

$$f(\mu | X_1, X_2) = \frac{1}{\sqrt{2\pi\sigma_N}} \exp\left\{-\frac{1}{2}\left(\frac{\mu - \mu_N}{\sigma_N}\right)^2\right\} \quad (4)$$

The formula can be obtained by (2) and (3)

$$\mu_N = \left(\sum_{k=1}^2 \frac{X_k}{\sigma_k^2} + \frac{\mu_0}{\sigma_0^2} \right) / \left(\sum_{k=1}^2 \frac{1}{\sigma_k^2} + \frac{1}{\sigma_0^2} \right) \quad (5)$$

So the bias estimate of μ is

$$\hat{\mu} = \int_{\Omega} \mu \frac{1}{\sqrt{2\pi\sigma_N}} \exp\left\{-\frac{1}{2}\left(\frac{\mu - \mu_N}{\sigma_N}\right)^2\right\} d\mu = \mu_N \quad (6)$$

4.3.2. BP neural network method

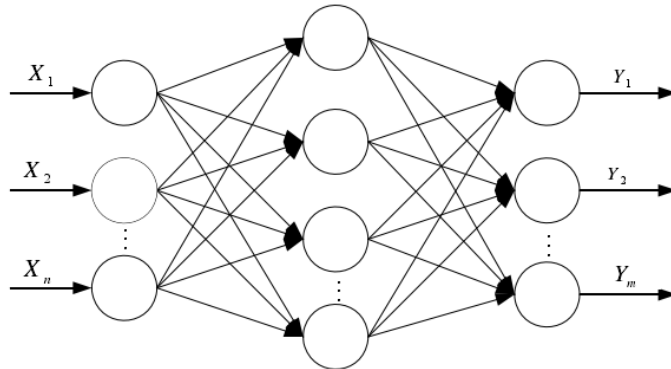


Fig. 5: Structural model of BP neural network

Neural networks can be classified according to the topology of the network and the flow of information. When taking topological structure as basis, neural network can be divided into hierarchical structure, interconnected structure and sparse structure according to the way of connection between neurons. And it can be divided into feedforward and feedback neural network models [6] according to the flow of internal information. The BP neural network used in this system belongs to the hierarchical and feedforward neural network model. The hierarchical structure model includes the input layer, the output layer, and the hidden layer. The network structure diagram is shown in Figure 5. In the feedforward neural network model, the information transmission route is from the input layer to the hidden layer, then to the output layer, and there is no feedback path between each layer.

In most cases, the nonlinear Sigmoid function, one of the excitation functions, is always used in the BP neural network:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

Only when the characteristics of the system change in the positive and negative interval, the symmetric Sigmoid function, also called the hyperbolic function, is applied as the excitation function.

$$f(x) = \text{than}(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (8)$$

The learning process of BP neural network can be divided into two aspects. One is the forward propagating of signal. During this process, if the output cannot reach the desired value, the error of the output signal would propagate along the original connection path in an opposite direction, which is called back propagation. During the back propagation, the weights of each neuron would be repeatedly adjusted until the minimum error value of the mean square value is reached. The basic learning ideas are as follows [7]:

(1) The weights of the neural network are W_{ij} , V_{jt} , and the threshold values are respectively θ_j and γ_t . The initialization of BP neural network is completed by assigning the initial value to the four values mentioned above. Then, input the first couple of BP neural network learning samples and use the following formulas to calculate the value of input u_j and output h_j of each neuron in the middle layer.

$$u_j = \sum_{i=1}^n w_{ij} x_i - \theta_j \quad (9)$$

$$h_j = f(u_j) = \frac{1}{1 + e^{-u_j}} \quad (10)$$

(2) Calculate the input l_t and output y_t in the output layer of the neuron, as well as the weight error δ_t of nerve unit t .

$$l_t = \sum v_{jt} h_j - \gamma_t \quad (11)$$

$$y_t = \frac{1}{1 + e^{-l_t}} \quad (12)$$

$$\delta_t = (c_t - y_t) y_t (1 - y_t) \quad (13)$$

(3) Update the connection weight W_{jt} and V_{jt} as well as threshold value θ_j and γ_t of neural network.

$$v_{jt}(N+1) = v_{jt}(N) + \alpha \delta_t h_j \quad (14)$$

$$w_{jt}(N+1) = w_{jt}(N) + \alpha \delta_j x_i \quad (15)$$

$$\theta_j(N+1) = \theta_j(N) + \beta \delta_j \quad (16)$$

$$\gamma_t(N+1) = \gamma_t(N) + \beta \delta \quad (17)$$

(4) Train all samples through the steps above. If the following conditions are satisfied, finish the training, or carry out a new round of learning.

$$\left| \sum_{k=1}^z E_k \right| \leq \varepsilon \quad (18)$$

In the formula above, ε is a presupposition precision and E_k stands for square error.

$$E_k = \frac{1}{2} \sum_{t=1}^q (c_t - y_t)^2 \quad (19)$$

4.4. PID control algorithm

The system involves complex environmental system and PID control system is more applicable to single environmental factor controlling, so an optimized PID control system combining BP neuron network was applied to the design of the system. This PID controller consists of two parts: the classic PID controller and the neural network. PID formula is as follows [8]:

$$\Delta u(k) = k_p [e(k) - e(k-1)] + k_i e(k) + k_d [e(k-1) - e(k-2)] \quad (20)$$

In the formula above, k_p is a proportionality constant; $k_i = T/T_i$ is an integral constant; $k_d = T_d/T$ is a differential constant; T is the sampling time; T_i is an integral time; and T_d is a differential time.

The system does not need a specific mathematical model. It can use the self-learning and adaptability of BP neural network to approximate any nonlinear function, so as to generate a specific model to control and optimize the system. Three parameters, K_p , T_i and T_d , can be adjusted by using self-learning of BP neural network to find out the optimal control rate. The adaptive PID controller realized by neurons has short transient time and small overshoot, which improves the stability and anti-interference of the system so as to maintain the system at its set value.

5. System Test

In order to verify the feasibility and reliability of the system, the system was applied to a self-developed intelligent growth cabinet. The parameters about environmental factors were set as the following: temperature 23°C, humidity 60%, carbon dioxide 1000ppm. The computational results of the system are shown in Fig. 6.

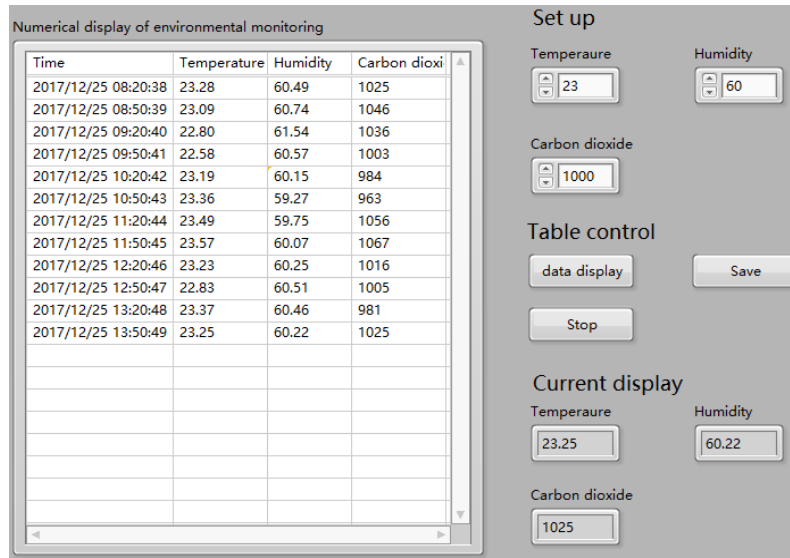


Fig. 6: The monitoring interface of the system

From the numerical display of environment monitoring in the table above, we can see that the system can successfully control the temperature within the range of $\pm 0.5^\circ\text{C}$, the humidity within the range of $\pm 2.5\%$, and the carbon dioxide concentration within $\pm 5\%$. So, the whole monitoring system has reached the design goal, which can meet the needs of monitoring system in plant factory.

6. Conclusion

Aiming at meeting the requirement of developing intelligent plant factory, A plant growth environment monitoring system is designed. The hardware and software design in this system is based on embedded

microcontroller, and the accurate monitoring of temperature, humidity and carbon dioxide concentration is realized by the integration of multi-sensor information. In addition, the system can run with a set value by using PID controller, and it has a good operation and display interface. In a word, the system meets the design requirements.

7. References

- [1] Qichang Yang. The development strategy of plant plant [J]. Science & Technology Review, 2014,(10):20-24.
- [2] Bowen Liu. Yan Sun. A classic example of embedded real time operating system μ C/OS- II-Based on STM32 processor [M]. Beijing: Beijing University of aerospace and Aeronautics and Astronautics Press, 2012.
- [3] Fan wu. Study on cucumber biomass accumulation model under controlled environment [D]. Jiangsu, Jiangsu University, 2009:9-21.
- [4] Lin Shao. Application of multi sensor fusion technology in environmental monitoring system in animal house [D]. Hebei: Agricultural University of Hebei, 2013.
- [5] Jiang He, Youchao Sun. Application of multi source reliability data fusion method based on Bayes theory [J]. Aircraft Design, 2012, 6(32):52-55.
- [6] Dachun Yang. Behavior recognition based on genetic algorithm to optimize BP neural network [D]. Liaoning, University of Science and Technology Liaoning, 2012.
- [7] Xue Li, Hanping Mao, Zhiyu Zuo, etc. Study of MATLAB neural network prediction based on turf evapotranspiration in Greenhouse [J]. Journal of Anhui Agricultural Sciences, 2008, 36(16):6609-6610.
- [8] Binfeng Gu, Yunshi Xiao, Yongqing Su. Research on fuzzy PID control method for agricultural greenhouse temperature based on data fusion [J]. System simulation technology, 2006, 10(4):198-201.