Computer an
Computing T
in Agricultur

IFIP Advances in Information and Communication Technology 392

IFIP – The International Federation for Information Processing

IFIP was founded in 1960 under the auspices of UNESCO, following the First World Computer Congress held in Paris the previous year. An umbrella organization for societies working in information processing, IFIP's aim is two-fold: to support information processing within its member countries and to encourage technology transfer to developing nations. As its mission statement clearly states,

> *IFIP's mission is to be the leading, truly international, apolitical organization which encourages and assists in the development, exploitation and application of information technology for the benefit of all people.*

IFIP is a non-profitmaking organization, run almost solely by 2500 volunteers. It operates through a number of technical committees, which organize events and publications. IFIP's events range from an international congress to local seminars, but the most important are:

- The IFIP World Computer Congress, held every second year;
- Open conferences;
- Working conferences.

The flagship event is the IFIP World Computer Congress, at which both invited and contributed papers are presented. Contributed papers are rigorously refereed and the rejection rate is high.

As with the Congress, participation in the open conferences is open to all and papers may be invited or submitted. Again, submitted papers are stringently refereed.

The working conferences are structured differently. They are usually run by a working group and attendance is small and by invitation only. Their purpose is to create an atmosphere conducive to innovation and development. Refereeing is also rigorous and papers are subjected to extensive group discussion.

Publications arising from IFIP events vary. The papers presented at the IFIP World Computer Congress and at open conferences are published as conference proceedings, while the results of the working conferences are often published as collections of selected and edited papers.

Any national society whose primary activity is about information processing may apply to become a full member of IFIP, although full membership is restricted to one society per country. Full members are entitled to vote at the annual General Assembly, National societies preferring a less committed involvement may apply for associate or corresponding membership. Associate members enjoy the same benefits as full members, but without voting rights. Corresponding members are not represented in IFIP bodies. Affiliated membership is open to non-national societies, and individual and honorary membership schemes are also offered.

Daoliang Li Yingyi Chen (Eds.)

Computer and Computing Technologies in Agriculture VI

6th IFIP WG 5.14 International Conference, CCTA 2012
Zhangjiajie, China, October 19-21, 2012
Revised Selected Papers, Part I

Springer

Volume Editors

Daoliang Li
Yingyi Chen
China Agricultural University
China-EU Center for Information
and Communication Technologies in Agriculture (CICTA)
17 Tsinghua East Road, Beijing, 100083, P.R. China
E-mail: {dliangl, chenyingyi}@cau.edu.cn

ISSN 1868-4238 e-ISSN 1868-422X
ISBN 978-3-642-36123-4 e-ISBN 978-3-642-36124-1
DOI 10.1007/978-3-642-36124-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013930789

CR Subject Classification (1998): J.0, I.2.9-11, I.2.6, H.4.2-3, I.4.8-9, I.5.1, I.5.4, C.2.1, C.2.m, H.2.8, I.2.4, C.3

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

First of all, I must express my sincere thanks to all authors who submitted research papers to support the Sixth International Conference on Computer and Computing Technologies in Agriculture (CCTA 2012), held in Zhangjiajie, China, 19–21 October 2012.

This conference was hosted by China Agricultural University, the IFIP TC5 Working Group (WG) on Advanced Information Processing for Agriculture (AIPA), and the Agricultural Engineering Information Committee of the Chinese Society of Agricultural Engineering, and was organized by the China-EU Centre for Information and Communication Technologies (CICTA).

Proper scale management is not only the necessary approach to agro-modernization and agro-industrialization but is also required by the growth in agricultural productivity, so, the application of different technologies in agriculture is becoming especially important. 'Informatized Agriculture' and the 'Internet of Things' have been chased by many countries recently in order to scientifically manage agriculture to achieve low costs and high income. CICTA aims to promote research and development in advanced and practical technologies applied in agriculture and to encourage international communication and cooperation, and has successfully held six international conferences on Computer and Computing Technologies in Agriculture since 2007.

The topics of CCTA 2012 cover a wide range of interesting theory and applications related to kinds of technology in agriculture, including the Internet of Things and cloud computing; simulation models and decision-support systems for agricultural production; smart sensors, monitoring and control technology; traceability and e-commerce technology; computer vision, computer graphics and virtual reality; the application of information and communication technology in agriculture; and universal information service technology and service systems development in rural areas.

We selected 108 best papers among all the papers submitted to CCTA 2012 for these proceedings. The papers have divided into two thematic sections. Creative thoughts and inspirations have been discovered, discussed and disseminated. It is always exciting when creative experts, professionals and scholars get together to share inspiring ideas and hopefully accomplish great developments in the technologies in high demand.

Finally, I would like to express my sincere thanks to all authors, speakers, session chairs and attendees both coming from abroad and from mainland China for their active participation and support of this conference.

December 2012 Daoliang Li

Conference Organization

Organizer

China-EU Center for Information and Communication Technologies in Agriculture (CICTA)

Chairman

Daoliang Li

Conference Secretariat

Lihong Shen

SPONSORS

China Agricultural University

The IFIP TC5 Working Group (WG) on Advanced Information Processing for Agriculture (AIPA)

Agricultural Engineering Information Committee, Chinese Society of Agricultural Engineering

Table of Contents – Part I

Erratum

Table of Contents – Part II

Research on the "Three Networks in One" Orchard Production Information Service System

Yun Qiu, Jingchao Fan, Lin Hu, and Guomin Zhou

Agricultural Information Institute (AII) of the Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081
qiuyun@mail.caas.net.cn

Abstract. This system has, by the comprehensive use of various latest information technologies and active gateway technologies, linked the wireless communication network, fixed-line telephone network and the Internet network (Three Networks) together to realize the seamless bridging of information, so as to broaden the service scope of the application system on the internetwork. The agricultural experts and/or agricultural technology promoters could provide the information service to the grassroots orchardists by use of this system in an even more convenient and speedy manner and the grassroots orchardists could also obtain the information on the internetwork in a flexible and convenient form, so that the "problem of last kilometer" in the dissemination of orchard information has been effectively solved.

Keywords: information technology, orchardist hotline, orchard production, remote diagnosis, last kilometer.

1 Proposition of the Problems

Since our implementation of the reform and opening-up policies, China's orchard production has seen a rapid development. In 2010, China's fruit cultivation area reached 11.5440 million hectares and the output thereof achieved 214.0140 million tons [1]. China's fruit production has ranked the first place in the world [2]. However, there also exist some problems, which have restricted the healthy and sustainable development of the orchard industry. The first is the poor quality and low ratio of high-quality fruits. At present, China's ratio of high-quality fruits is less than 10% of the total output and the top-grade fruits up to the export standard is less than 5% of the total output [3], which has resulted in the low market competition and low benefit. The second is the backward sustainable development of the orchard industry and consciousness of the safety production and the weak consciousness of the non-polluted production, which has resulted in such a serious phenomenon in which the farm-oriented chemical inputs have polluted the environment and the pesticide residue remains on the fruits or the heavy metal exceeds the standard. The third is the imbalanced variety structural ratio, in which the excessive development of a single variety could hardly satisfy the market demand of the diversified consumption but has increased the pressure of marketing. The existence of these problems have been caused because of various reasons, but one of the main reasons should be the lack of

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 1–10, 2013.

the effective route to timely obtain the new technology and market information, which has resulted in the backward technology, sluggish information and blind production for the entire orchard producers.

With the continuous improvement of China's comprehensive national strength and the continuous development of the project of "Coverage of Every Village", China's vast countryside has basically realized the entrance of telephone and television into the households in the villages. At the same time, with the rapid development of China's wireless communication project, it is no longer the dream for the better-off orchardists to have their own mobile phones. Telephone and mobile phone have gradually become the tools for the broad grassroots orchardists to learn new knowledge and new technology. Then, it is a problem that should be solved in China's agricultural information technology as to how the agricultural experts or agricultural technology promoters could make full use, in an even more convenient and speedy manner, of these advanced communication tools to pass on the agricultural technology to the orchardists and solve the problems encountered by the orchardists in their production. It is, therefore, of important practical significance in respect to the promotion of China's orchard production to research and use the information technology to solve such problems as of the sluggish dissemination and low dissemination efficiency of the information on the orchard production and set up a bridge of communication between the agricultural experts and the orchardists.

2 Research on the Mode of the Orchard Production Information Service System

2.1 The Main Situation of the Telecom Facilities in Rural Area

Since our entry into the 21 century, China's government departments at different levels have attached great importance to the construction of the agricultural information network, in which China's 31 provincial-level agricultural departments (exclusive of Hong Kong, Macao and Taiwan) have established their provincial-level agricultural information network platforms and most of the provinces have established their agricultural information centers. Throughout the country, there are more than 2,000 cities and/or counties that have opened up their agricultural information service websites. Until 2010, the agricultural websites have reached 31,108 [4]. China has preliminarily formed an agricultural information network system to cover the whole country.

According to the latest statistical results released by China Internet Network Information Center (CNNIC) in January 2012, up to the end of December 2011, the scale of the Chinese netizens had broken through 500 million and reached 513 million, of which the scale of the rural netizens was 136 million, accounting for 26.5% of the total netizens [5]. As compared with that in 2006, the rural netizens had increased by 112.89 million and the proportion in the total netizens had gone up to 26.5% from the 1.6% in 2006. This could demonstrate the remarkable increase of the rural residents to approach the Internet. However, as seen from the netizens occupational structure, such netizens as engaged in the agricultural labor accounted

only for 7.2% of the total rural netizens (as shown in Figure 1) [6], which was still at a quite low level. It could be seen from it that in addition to the weak consciousness of the Chinese farmers to obtain the information by use of the Internet network, the websites are quite short of such information as being interesting, understandable and learnable for the farmers, and the farmers have quite insufficient means to obtain the website information.

As to how to make a fine job in running such websites as favored by the broad orchardists and at the same time to enable more and more orchardists to obtain the website information and to really make the orchard production technology information enter into the households in the villages, it is the first problem for the current solution to solve the "last kilometer" in the orchard informatization.

Because the per capita annual income in the rural areas is comparatively lower and the ability there to pay for the consumption is limited, the Internet-approaching equipment for the netizens in the rural areas is relatively deficient. However, with the increasing convenience of the mobile phone approaching the Internet, the mobile phone has been favored by the farmers because of its lower price. In 2009, the proportion of the rural netizens to use the desktop computer as the Internet-approaching terminals has declined to 68%, whereas the proportion of the netizens to approach the Internet by use of the mobile phone has seen a rapid increase. At present, China's mobile phone users for the Internet approach have reached 233 million, of which the rural mobile phone users for the Internet approach are about 71.89 million, accounting for 67.3% of the total rural netizens (as shown in Figure 2) [6]. Since 3G network has already covered all the county towns and most of the townships throughout the country, the mobile phone has become the mainstream of the Internet-approaching terminals for the rural netizens.

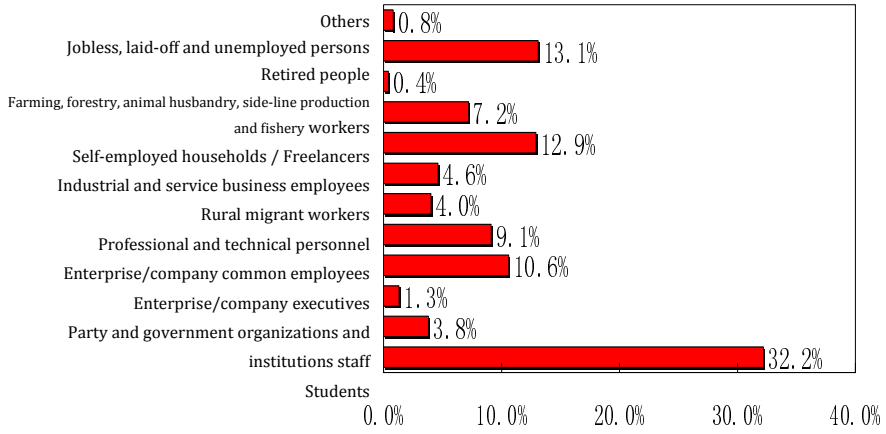


Fig. 1. Rural Netizens Occupational Structure

As shown in the statistical results released by China's Ministry of Industry and Information Technology in January 2011, up to 2010, the whole country's telephone users had totaled 1.15339 billion households, of which the rural telephone users had reached 97.76 million households, with the fixed-line telephone penetration being

22.1 set/100 people. Among them, the mobile phone users were 859 million and 3G users reached 47.05 million. The mobile phone penetration had reached 64.4 set/100 people, accounting for 74.5% of the total telephone users and about three times of the fixed-line telephone users. As the rural communication development objectives for the "11th Five-Year Plan" period, the program of "Telephone to Be Connected to Every Village and Internet to Be Approached in Every Township" has been comprehensively realized. Throughout the country, 100% of the administrative villages have been connected with the telephone, 100% of the townships have been connected with the Internet (of which 98% of the townships are connected with the broad band), 94% of such natural villages as of more than 20 households have been connected with the telephone, and nearly half of the countrywide townships have established the township information service station and county-township-village three-level information service system. China's telecommunication facilities have reached the objective of "entrance into the households in the villages" [7].

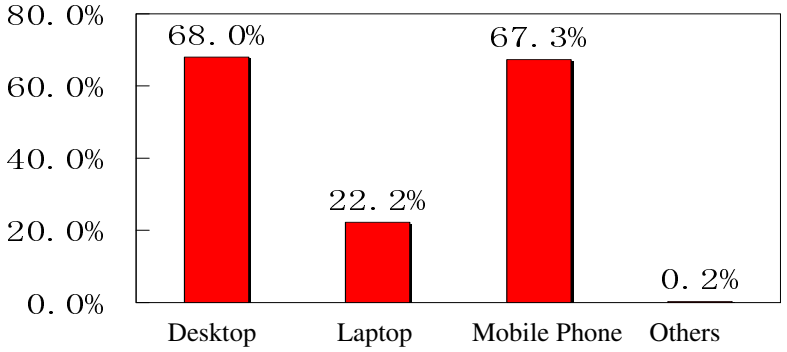


Fig. 2. Rural Netizens Internet-approaching Equipment

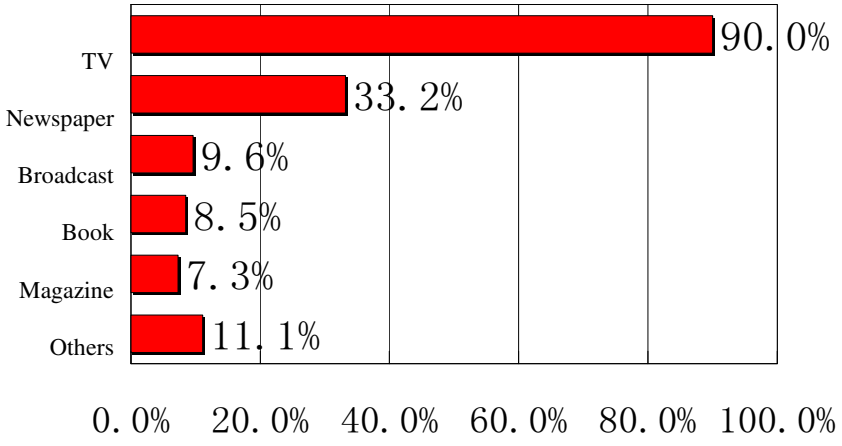


Fig. 3. Main Routes for the Non-Netizens to Obtain Information

It is discovered from the statistical findings of China Internet Network Information Center that the main route for 88.3% of the non-netizens to obtain the information is TV, 35.9% of the non-netizens to obtain the information is newspaper, and the proportion of such non-netizens whose main routes to obtain the information are broadcast, book or magazine is not more than 13% (as shown in Figure 3) [8]. It is thus clear that most of the non-netizens are to obtain the information through the TV.

2.2 The Architecture of the Fruit Production Information Service System

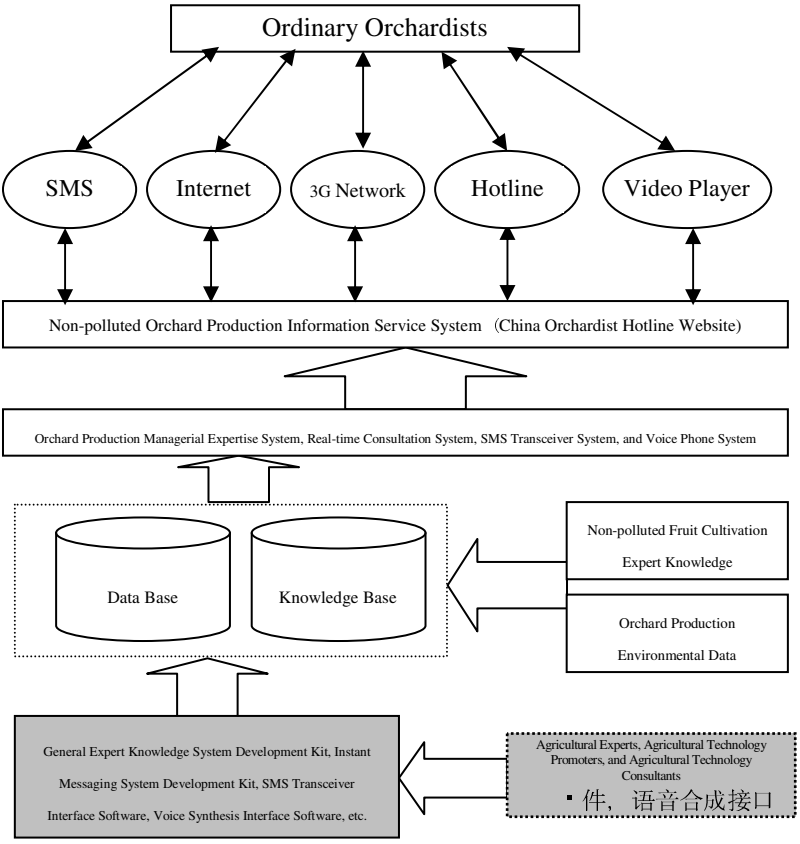


Fig. 4. Structure of the Orchard Production Information Service System

To make full use of China's existing telecommunication facilities and good use of the television and the SMS called as the "Fourth Medium" is the best route to solve the "last kilometer" in the orchard informatization. For this purpose, we have designed such a "Five-in-One" system mode as of the Internet, 3G network, hotline, SMS and video play (as shown in Figure 4). We have developed the networked "Non-polluted Orchard Production Managerial Knowledge System" by the comprehensive use of XML technology, multimedia technology, COM control technology, expert system

technology and WAP technology. By use of the streaming media processing technology of DirectShow on the WDM device driver model, we have researched and developed the "Instant Messaging Components" and established the "Non-polluted Orchard Remote Real-time Consultation System". We have provided the orchardists with the on-line consultation and the interface to conduct the "one-to-one" on-line communication with the orchardists and orchard experts. Based on the GSM text message transceiver and OCX control and by use of the API functions provided by the GSM module, we have researched and developed the "Non-polluted Orchard Production SMS Transceiver System". And furthermore, by use of CCITT No. 7 signaling, by means of the digital signal processing (DSP) technology and by use of the universal telephone voice player, we have researched and developed the "Non-polluted Orchard Production Telephone Voice Play System". By use of their mobile phones and/or telephones as the terminals, the farmers could interact with the application on the Internet network and realize the acquisition of information whenever and wherever possible.

3 Construction of the Non-polluted Orchard Production Managerial Knowledge System

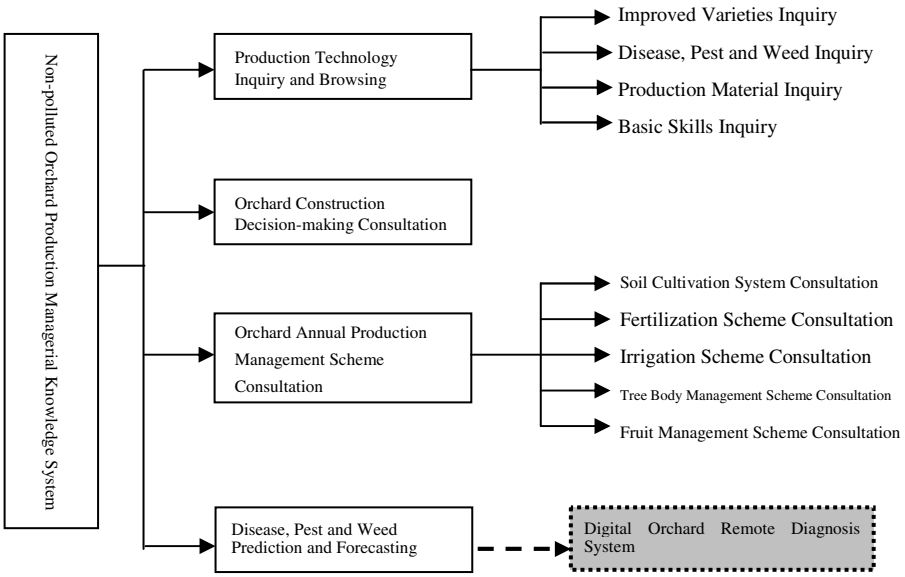


Fig. 5. Functional Diagram of the "Non-polluted Orchard Production Knowledge System"

The Non-polluted Orchard Production Managerial Knowledge System (as shown in Figure 5) includes: Orchard Production Technical Knowledge Browsing Module, Orchard Construction Consultation Module, Orchard Annual Production Management Scheme Consultation Module, and Disease, Pest and Weed Prediction and

Forecasting Module. Through the Internet or the mobile phone or telephone, the orchardists could directly communicate with the system to inquire any knowledge required to know in the production and any problems that can be solved according to certain rules.

4 Implementation of the System

Fig. 6. Interface of China Orchardist Hotline Website

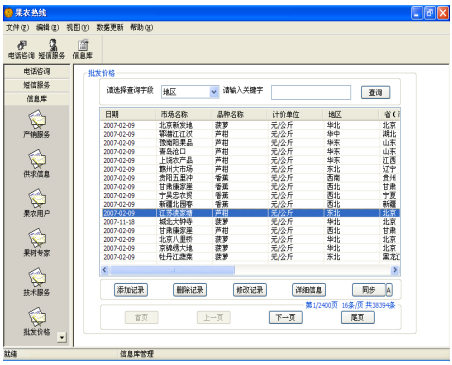


Fig. 7. Interface of the Orchardist Hotline Background Maintenance System

Fig. 8. Interface of the Digital Orchard Remote Diagnosis System

The "Three Networks in One" Orchard Production Information Service System consists of such three parts as of China Orchardist Hotline Website (as shown in Figure 6), Orchardist Hotline Background Maintenance System (as shown in Figure 7) and Digital Orchard Remote Diagnosis System (as shown in Figure 8).

4.1 China Orchardist Hotline Website

China Orchardist Hotline Website (http://www.cart.net.cn/) is the web portal of the "Orchard Production Information Service System" on the Internet. (As for the Android-edition mobile phone users, please visit http://android.cart.net.cn, but for the wap-edition mobile phone users, please visit http://wap. cart.net.cn.) The Website has integrated the "Non-polluted Orchard Production Managerial Knowledge System", "Apple Tree Plant Diseases and Insect Pests Expert System" and "Digital Orchard Remote Diagnosis System" to provide the information service of public benefit for the orchardists throughout the country. The Website includes such columns as of the recommendation of the latest fruit trees and nursery stock, recommendation of competitive orchard, practical cultivation technique, market price quotes, as well as the expert on-line consultation, expert system and training classroom.

4.2 Orchardist Hotline Background Maintenance System

By use of the C/S structure, programmed with the C++Builder language and working in the environment of Windows Operating System, the Orchardist Hotline Background Maintenance System is composed of such three parts as of telephone consultation, SMS and information base maintenance. (1) Telephone Consultation and Audio Service: It consists of such four modules as of telephone attendance, running conditions statistics, broadcasting list maintenance and voice base maintenance. (2) SMS: It includes such five parts as of text message receiving, draft mailbox, text message en route, text message sent and user address list maintenance. (3) Information Base Maintenance: This module is a background maintenance for the database of China Orchardist Hotline Website, which contains such functions as of the addition, modification, deletion and details viewing for the data set of each database, and is at the same time characterized by the data synchronization, which means to make the database of the server synchronous with the data on the customer terminals so as to realize the release on the Website of the data maintained by the agricultural experts (or agricultural technology consultant).

4.3 Digital Orchard Remote Diagnosis System

By use of the B/S structure, the Digital Orchard Remote Diagnosis System consists of such three parts as of the remote consultation, video conference and knowledge inference. (1) Remote Consultation: It contains such functions as of the image-text communication, video consultation, document transmission, electronic whiteboard, viewing and analysis of the information on the orchard cultivation environment. (2) Video Conference: It is designed to provide the real-time on-line consultation function for the experts and orchardists through the construction of the virtual video conference room. (3) Expert System: It is designed to collate the experts' experience and knowledge to form a knowledge inference machine, so as to solve, in the limited steps, the problems that the orchardists may encounter.

5 Conclusion

At present, the "Three Networks in One" Orchard Production Information Service System is mainly applied in Xingcheng and Suizhong of Liaoning Province and the outer suburbs of Beijing. This System has linked the wireless communication network, fixed-line telephone network and the Internet network together to realize the seamless bridging of information, so as to broaden the service scope of the application system on the internetwork. It has closely connected the experts of the Institute of Pomology (IP) of the Chinese Academy of Agricultural Sciences (CAAS) with the orchardists in the promotion pilot zone to realize the "all-weather" information service for the orchardists, which is greatly welcome by the orchardists in the promotion pilot zone. It has realized the rapid promotion in a large area of the orchard new technology, so as to provide the convenient means of decision-making and management for the grassroots technical personnel and broad orchardists in the non-polluted orchard production. It has realized the universality and foolproof of the non-polluted orchard production technology, so as to make it be visible, tangible and usable for the broad orchardists and become the "stationed orchard senior expert" in the orchard zone. All of these have fundamentally improved the technical level of the broad orchard producers and made up such a series of malpractice as resulted from the imperfection of the orchard technology promotion system and the shortage of the orchard technical personnel, so that the "problem of last kilometer" in the dissemination of orchard information has been effectively solved.

Acknowledgments. The Orchard Production Information Service System has been accomplished in the support of the National High-Tech (863) Program's Project of "Remote Intelligent Diagnosis System for the Field Crop Plant Diseases and Insect Pests (2007 AA10Z237)" and the Agricultural Research Findings Commercialized Project of "Remote Diagnosis System for the Orchard Plant Diseases and Insect Pests and Its Pre-production (2009 GB23260457)", for which we would like to extend our heartfelt thanks.

References

1. China Agricultural Information.Network. China's Fruit Production Situation over the Years http://www.agri.gov.cn/V20/cxl/sjfw/tjsj/sg_1/scqk/历年水果生产情况.html.
2. China Agriculture Press. China Agricultural Yearbook (2010)
3. Wang, L.: An Analysis on China's International Competitiveness in Fruits. Agricultural Economy (04), 28 (2011)
4. Ministry of Agriculture of the People's Republic of China.China's Development Report on Agricultural and Rural Informatization (2010), http://www.moa.gov.cn/ztzl/sewgh/fzbg/201112/P020111208362702958871.pdf
5. China Internet Network Information Center. China's Statistical Report on the Internet Network Development Situation (1) (2012), http://www.cnnic.net.cn/dtygg/dtgg/201201/W020120116337628870651.pdf

6. China Internet Network Information Center. China's Investigation Report on the Rural Internet Development Situation in (2009), http://www.cnnic.cn/dtygg/dtgg/201004/t20100415_13712.html
7. Ministry of Industry and Information Technology of the People's Republic of China. China's National Statistical Bulletin on the Telecommunications Industry in (2010), http://www.miit.gov.cn/n11293472/n11293832/n11294132/n12858447/13578942.html
8. China Internet Network Information Center. China's Statistical Report on the Internet Network Development Situation (1) (2007), http://www.cnnic.net.cn/uploadfiles/doc/2007/1/22/212245.doc

Rapid Identification of Waste Cooking Oil with Near Infrared Spectroscopy Based on Support Vector Machine

Xiong Shen[1], Xiao Zheng[1], Zhiqiang Song[1], Dongping He[2], and Peishi Qi[3]

[1] Institute of Mechanical Engineering, Wuhan Polytechnic University, Wuhan 430023, China
sx198711@yahoo.com.cn, zhengxiao@whpu.edu.cn, 327463922@163.com
[2] Institute of Food Science and Engineering, Wuhan Polytechnic University, Wuhan 430023, China
hedp123456@163.com
[3] Pashun Group, Wuhan 430023, China
qps@vip.sina.com

Abstract. The qualitative model for rapidly discriminating the waste oil and four normal edible vegetable oils is developed using near infrared spectroscopy combined with support vector machine (SVM). Principal component analysis (PCA) has been carried out on the base of the combination of spectral pretreatment of vector normalization, first derivation and nine point smoothing, and seven principal components are selected. The radial basis function (RBF) is used as the kernel function; the penalty parameter C and kernel function parameter γ are optimized by K-fold Cross Validation (K-CV), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), respectively. The result shows that the best classification model is developed by GA optimization when the parameters C = 911.33, γ= 2.91. The recognition rate of the model for 208 samples in training set and 85 samples in prediction set is 100% and 90.59%, respectively. By comparison with K-means and Linear Discriminant Analysis (LDA), the result indicates that the SVM recognition rate is higher, well generalization, can quickly and accurately identify the waste cooking oil and normal edible vegetable oils.

Keywords: near infrared spectroscopy, waste cooking oil, support vector machine, parameters optimization.

1 Introduction

Catering waste oils include drainage oil (in narrow sense), hogwash fat (waste cooking oil) and fried old oil. After pickling, washing, decoloration, deodorization and other processing, the catering waste oils often close to or completely achieve the national Hygienic Standard of Edible Vegetable Oil in sensory index and conventional typical properties, which consumers and government supervisors are difficult to identify by the sense of the sights and smell. At present, a complete set of testing technology standard of identification of the catering waste oil hasn't been established domestically or abroad. The Ministry of Health is requesting proposals for proposals from the public. Near Infrared Spectroscopy (NIR) technology is a

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 11–18, 2013.

nondestructive testing technique rapidly developed in recent years [1]. The domestic scholars make use of NIR qualitative analysis to research the types of edible oil [2-4], however, qualitative analysis for catering waste oil is still limited.

Support Vector Machine (SVM) is a new kind of machine learning algorithm based on the minimum principle of statistical learning theory and structural risk, which has advantages of simple structure, strong generalization ability and others. It presents many unique advantages in solving problems of pattern recognition in small sample, nonlinear, high dimension, local minimum [5]. The methods combined SVM with NIR have been applied successfully in identifying the category of tea, milk powder, apple and others [6-9]. The objective of this study is to develop a classified model for catering waste oil and four normal edible vegetable oils by combining SVM with NIR. This model provides a new approach to fast and effective identification of catering waste oil.

2 Experiments and Methods

2.1 Experimental Samples

Catering waste oils used in this experiment include drainage oil and hogwash fat obtained through different degree of refining of decoloration, deodorization, and normal edible vegetable oil which are of different brands or the same brand of different batches in major supermarkets. The samples make up of the following table 1:

Table 1. Composition of the experimental samples

	Training set	Predicting set	In total
The first category: drainage oil and hogwash fat	99	47	146
The second category: soybean oil	40	19	59
The third category: peanut oil	26	7	33
The forth category: olive oil	23	6	29
The fifth category: blend oil	20	6	26
In total	208	85	293

2.2 Experimental Methods

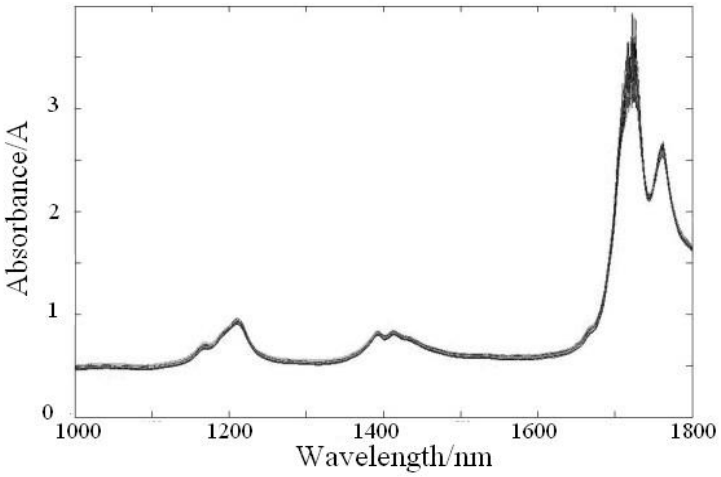
Adopt SupNIR-5700 NIRS (Focused Photonics (Hangzhou), Inc.) to collect NIR spectra of all samples. Spectral measurement of samples uses random RIMP software and its testing method is: transmission, measurement range: 1000~1800nm, scanning speed:10 times/sec, spectral resolution: 6nm, temperature of sample cell: 60°C, testing method: load the sample into the three-quarters of sample bottle, and then place the sample bottle into the sample cell. Stabilized in constant temperature for 5min, the bottle is taken out to check if there exist bubbles. It starts to collect spectrogram if there is no bubble, and each sample averages out three times.

Use NIRS random RIMP software and MATLAB7.8 to collect spectra and convert data format, use chemometrics software Unscrambler X 10.1 to pretreatment the spectral data and analyse principal component, and use SVM pattern recognition and regression software package designed by a professor Lin Zhiren from National Taiwan University to build SVM models in MATLAB7.8 and parameters optimization.

3 Results and Discussion

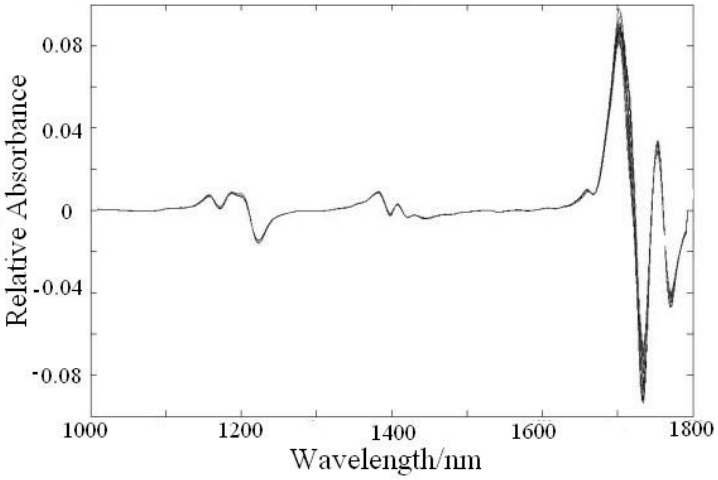
3.1 Pretreatment for Spectral Data

Besides samples' information collected through NIRS, it contains other irrelevant information and noise, therefore, it is very important and necessary to pretreatment spectra before developing model. Many kinds of methods for spectral pretreatment, including mean centralization, normalization, Savitzky-Golay smoothing, Savitzky-Golay first derivation and second derivation and so on, have been tried in this study. The attempted result indicates that NIR obtains the best pretreatment effect by combining vector normalization with Savitzky-Golay first derivation and nine-point smoothing. Fig.1 shows raw and spectra after pretreatment respectively.



(a) Raw spectra

Fig. 1. Conventional and spectra after pretreatment

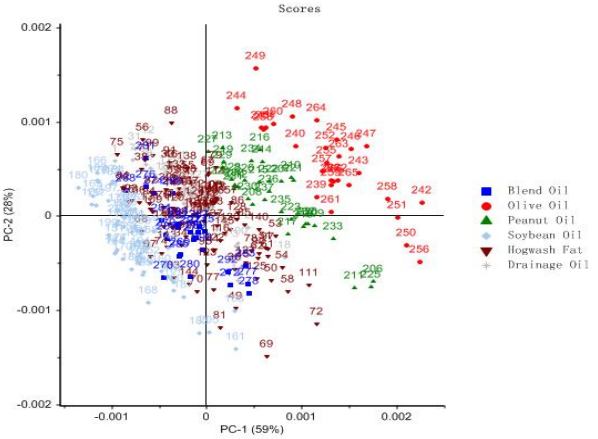


(b) Pretreatment spectra

Fig. 1. (*continued*)

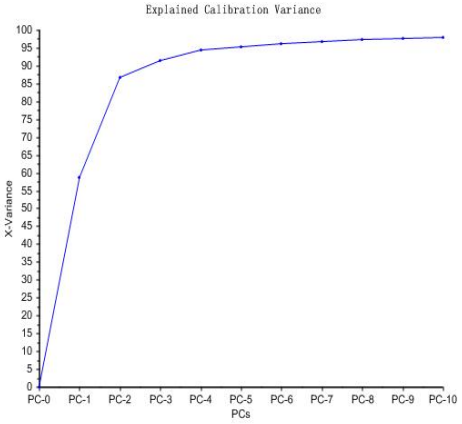
3.2 Extraction of Spectral Principal Component

Analyze the principal component of spectra after pretreatment, as shown in Fig.2-a, the X-axis stands for the first principal component (PC1), Y-axis represents the second principal component (PC2). The figure shows the good effect of sample distribution. This experiment proves that principal component can reflect most of information when principal component's accumulative contributing rate is above 95% and principal component scree plot (as shown in Fig.2-b) is quite smoothing. Therefore, this paper selects the previous seven principal components (accumulative contributing rate is 96.56%) as SVM input.



(a) PCA SCORE

Fig. 2. PCA SCORE and explained variance



(b) Explained Variance

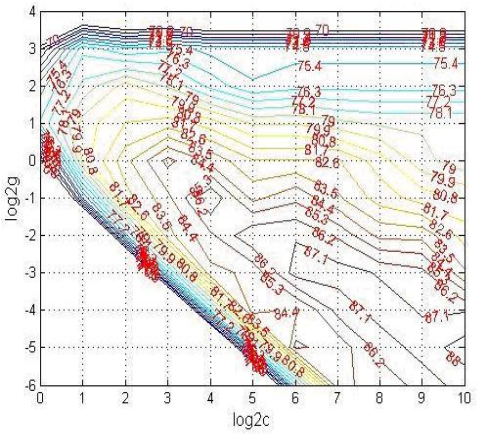
Fig. 2. (*continued*)

3.3 SVM Model Building and Parameter Optimization

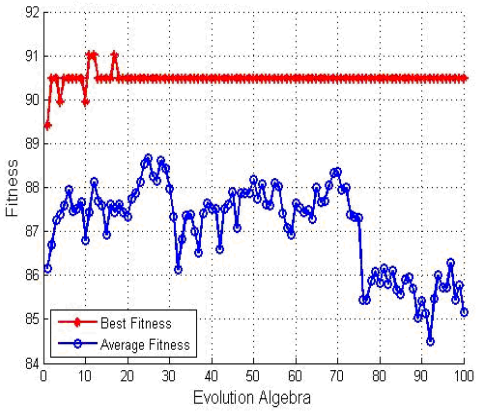
Libsvm includes two classification models: C-SVC and nu-SVC. Based on one-against-one algorithm solving multi-classes pattern recognition, this paper uses C-SVC to establish classification modeling. It needs to select kernel function and parameters when using SVM for pattern recognition. At present there is no unified international model, so we could only use experience or experimental comparison. Typically, using RBF kernel function often gets better simulation results [9], and reduces complexity of computation during the training process. Therefore, this paper makes use of RBF kernel function to establish identification model.

It is very important to select penalty parameter C and kernel function parameter γ in RBF kernel function. C is used to measure the size of the penalty, γ is used to control function regression error and directly influence the initial characteristic value and feature vector. The research respectively uses K-CV, GA and PSO algorithm to optimize the models of C and γ to reach the highest accuracy of classification of training set under the best parameters C and γ. However, it cannot guarantee the testing set to reach the highest accuracy of classification. Fig.3 shows the results of three parameters optimization. Fig.3-a gives the optimization results using K-CV parameter optimization. Fig.3-b gives the optimization results of fitness curve using GA parameter optimization, where the maximum number is 100, the population size is 20, the crossover probability is 0.8, the range of parameters C and γ are 0-1000, other parameters are by default. Fig.3-c gives the optimization results of fitness curve using PSO parameter optimization, where the maximum number of iterations is 100, the initial population size is 20, the learning factor c_1=1.5, c_2=1.7, the range of parameters C and γ are 0-1000, other parameters are by default.

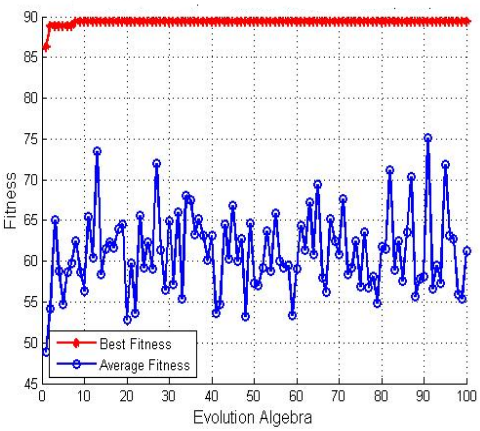
Use the default parameters (C = 1, γ= 1 / K = 0.1429) and optimal results of three different parameters to respectively establish the SVM recognition model, which are analyzed in Table 2.



(a) K-CV



(b) GA



(c) PSO

Fig. 3. The results of three parameters optimization

From the table 2, it is clear that SVM model recognition rate of the default parameters is very low, and almost four kinds of normal edible vegetable oils can be classified as catering waste oils; recognition rate of SVM model increases significantly about 90% after optimal results of different parameters of K -CV, GA and PSO. The learning ability and generalization ability of SVM classifier with optimal parameters C and γ can keep a balance and avoid the occurrence of learning state and non-learning state. Examples show that SVM classification model established when GA optimal parameters C = 911.331, γ = 2.91045, recognition rate of the 208 training sets and 85 predicting sets is 100% and 90.59% respectively, only occurs four blend oils mistaken for catering waste oil, four hogwash oils for blend oils. In the meantime, compared with methods of k-means clustering and LDA, the recognition rate of GA-SVM model is higher than those about 10%. Therefore, SVM model is superior to the methods of k-means clustering and LDA.

Table 2. Different parameters—analysis of SVM modeling results

	Default (C=1, γ=0.1429)		K-CV (C=1024, γ=0.03125)		GA (C=911.331, γ=2.91045)		PSO (C=2287.16, γ=0.01)	
	Returning error number	Predicting error number	Returning error number	Predicting error number	Returning error number	Predicting error number	Returning error number	Predicting error number
The first c ategory	0	0	2	0	0	4	2	1
The second category	40	19	1	0	0	0	1	0
The third category	26	7	0	0	0	0	0	0
The forth category	15	5	0	0	0	0	0	0
The fifth category	20	6	20	6	0	4	20	6
Recognition rate	51.44%	56.47%	88.94%	92.94%	100%	90.59%	88.94%	91.76%

4 Conclusions

The research uses GA-SVM to establish NIR classification model for catering waste oil and four normal edible vegetable oils, and determines the appropriate model parameters. The recognition rate of the established models is achieved respectively 100% for training set and 90.59% for predicting set, the recognition rate and generalization ability of GA-SVM of NIR classification model is higher than conventional analysis model, which can rapidly and accurately identifies the catering waste oil.

The sample source of catering waste oil in the research is limited and cannot completely represent diversity and complexity of catering waste oil. In addition, the law breakers usually add catering waste oil to qualified edible vegetable oil according to a certain proportion, and then sell the fake oil, therefore, it needs to further collect representative adulterated samples in the future.

It is essential to keep developing new methods of qualitative classification to research, and constantly strengthen the maintenance for the models of qualitative classification; in addition, a rapid portable detecting instrument for testing catering waste oils based on the models of NIR quantitative classification needs to be developed in order to protect the security of food production, to provide a more reliable basis for food supervisions and to prevent catering waste oils back to the table.

Acknowledgment. Funds for this research was provided by the National Science and Technology Plan Projects (2009BADB9B08), the major projects foster special of food nutrition and safety of Wuhan Polytechnic University (2011Z06), the entrust projects of Wuhan PASHUN Group green energy technology Co., LTD, and the postgraduate 2010 innovation fund of Wuhan Polytechnic University(2010cx005).

References

1. Lu, W.: Modern Near Infrared Spectroscopy Analytical Technology, 2nd edn., pp. 19–36. Chinese Oil and Chemical Press, Beijing (2006) (in Chinese)
2. Wu, J., Liu, C., Li, H., et al.: Application of NIR technology on identifying types and determining main fatty acid content of edible vegetable oil. Journal of Beijing Technology and Business University (Natural Science Edition) 28(5), 56–59 (2010)
3. Liu, F., Chen, H., Jiang, L., et al.: Rapid discrimination of edible oil by near infrared transmission spectroscopy using clustering analysis. Journal of China Jiliang University 19(3), 278–282 (2008)
4. Li, J., Fan, L., Deng, D., et al.: Principal component analysis of 6 kinds of vegetable oils and fats by near infrared spectroscopy. Journal of Henan University of Technology (Natural Science Edition) 29(5), 18–21 (2008)
5. Zhang, X.: Introduction to Statistical Learning Theory and Support Vector Machines. Acta Automatica Sinica 26(1), 32–34 (2000)
6. Chen, Q.S., Zhao, J., Zhang, H., et al.: Indentification of Authenticity of Tea with Near Infrared Spectroscopy Based on Support Voctor Machine. Acta Optica Sinica 26(6), 933–937 (2006)
7. Zhao, J., Hu, H., Zhou, X.: Application of Support Vector Machine to apple classification with near—infrared spectroscopy. Transactions of the CSAE 23(4), 149–152 (2007)
8. Wu, J., Wang, Y., Zhang, X., et al.: Applied Study on Support Vector Machines in Identifying Standard and Sub-standard Milk Powder with NIR Spectrometry. Agricultural Mechanization Sciences 1(1), 155–158 (2001)
9. Ye, M., Wang, X.: Identification of Chaotic Optical System Based on Support Vector Machine. Acta Optica Sinica 24(7), 953–956 (2004)

A Decision Support System for Fish Feeding Based on Hybrid Reasoning*

Mingfei Zhang, Huiping Xue, Lianzhi Wang, and Daoliang Li

College of Information and Electrical Engineering, China Agricultural University
Beijing, China
wangcau@163.com

Abstract. Considering the complex of decision-making process of rainbow trout feeding, a rainbow trout feeding decision support system is built by integrating case-based reasoning and rule-based reasoning with XML. This paper describes the theory of hybrid reasoning, and several experiment results with their own interpretation. The web based feeding decision support system is proved to have improved the feeding, and lessened the difficulty of obtaining expert knowledge in practice.

Keywords: Knowledge Representation, Case-based Reasoning, Decision Support System.

1 Introduction

With the Development of computer network, on-line expert decision support system plays a more and more important role [1][2]. As a key part of decision support system, the knowledge database provides important foundation for decision [3]. How to convert the expert knowledge and experiment into cases and store them in database, and work with reasoning mechanism that generate fast and correct decision information, is a key issue in practice.

The rainbow trout cultivation is complex and it is difficult to abstract the feeding knowledge. In conventional feeding decision support system, people input the quantity, growth stage, and middleweight of fish, and the information of bait feed and nutritional needs, then calculate the feeding scheme using decision model, which mainly accomplished by linear programming, RBR(Case-based Reasoning), and CBR(Case-based Reasoning). RBR is easy understanding, and can efficiently present knowledge, and the disadvantages are: difficult to get and define rules, not obvious relationship between the rules, inconvenience of process knowledge, inadvisable management and maintenance, lack flexibility in reasoning. The advantages of CBR includes: fast speed of reasoning, easy build and maintain case base, more flexible, strong self learning ability; CBR also has its limitations: not easy to express, sensitive to noise, and lack correction mechanism.

* Fund Project: Beijing Municipal Natural Science Foundation (4092024).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 19–26, 2013.

In this paper, we tried to combine RBR and CBR, and build a feeding decision support system of rainbow trout. In many cases, a simple CBR method or RBR method cannot ensure the accuracy of result when solving problems. So, the combination of RBR and CBR reasoning is a better model, in which they not only play the respective advantages, but also compensate shortcomings for each other. There are many repeats between the CBR case attributes and the RBR knowledge rules, which contribute to combining reasoning. Depending on the unified representation format (XML, Extensible Markup Language), RBR and CBR perform a seamless connection in the reasoning process. Through realizing a reasoning scheduling system, CBR can visit and call RBR auxiliary unit and RBR correction unit at each stage when they are required. Therefore, whether the RBR is used and when it is used depends on whether the CBR system request for reasoning scheduling module and when request. This ensures the parallel execution of the same knowledge in rules and cases during reasoning process, and reduces the search times, improves the efficiency, but also keeps the independence of unit itself and the entirety of system strategy.

2 Knowledge Pepresentation Based on XML

2.1 Features of XML

In decision support system, there are many traditional methods to present knowledge, such as logic, production, frame, semantic network, object oriented, agent, rough set theory, and etc. Those single methods cannot meet the needs of artificial system, and people have put forward many new ways to present domain knowledge, which combined different ways[4] . In this paper, because of the difficulty of structurization and management of mixed knowledge, we adopted the XML to present knowledge.

The advantage of XML is arithmetical, easy organized and management, therefore nowadays more and more researchers use XML to present knowledge and solve problems[5].

As an open standard, XML enables organizations and persons to build standard set that is suitable in different situations. Secondly, XML is separate, and the storage format of data does not have to follow the display format. The self-describing feature makes XML better in describing complex data relations. Those properties also make applications that based on XML more correct and effective to search data, and pay no attention to irrelevant content[6] . XML has many other advantages, such as good format, abundant display style and convenient data processing ability. All this features can help the rainbow trout feeding decision support system provide easy, capable and exact services for users.

2.2 Representation Format of Rules and Cases Based on XML

The specific feeding cases described by XML are shown below, in which the attributes of case involves caseName, caseID, caseDate, caseTime, growPeriod and their own weight. The attributes of including water temperature, DO(dissolved oxygen), the content of ammonium, pH, water quality, content of protein in feed, salinity of water, average length of fish, average weight of fish. The case results are feeding rates and feeding advices.

```
<case caseName="案例1" caseID="2" caseTime="14:00"
caseDate="2009-1-12" growPeriod="鱼种"
weight="0.287-0.287-0.162-0.109-0.073-0.049-0.033">
<caseAttribute>
        <temperature rangeLow="2" rangeHigh="20">15</temperature>
        <DO>8</DO>
        <ammonia>0.015</ammonia>
        <pH>6.8</pH>
        <quality>12</quality>
        <feed>44</feed>
        <salt>0.4</salt>
        <avgLength>15</avgLength>
        <avgWeight>0.5</avgWeight>
</caseAttribute>
<caseResult>
        <caseFeedPer>7.2</caseFeedPer>
    </caseResult>
</case>
```

We use the production rule "IF-THEN" to describe the feeding rule, whose type consists with feature attributes stored in database. In the process of reasoning, we achieved one time retrieval, and at the same match and judge case and rule, finally, generated the hybrid feeding advice. The rules are presented as follows.

```
<rule name ="规则2" type="DO">
    <premise unit="mg/L">
        <DO restrain="小于" nested="and">7</DO>
        <DO restrain="大于" nested="no">5</DO>
    </premise>
<conclusion>溶氧不足，需要及时补氧!</conclusion>
</rule>
```

3 Reasoning Based on Case

3.1 Case Reasoning Process

Case based reasoning process can be summarized by three steps[7]:

(1) Retrieve historical experiments, and find similarities with current questions.
(2) Search the most similar cases from database.
(3) Learn from retrieved cases

Meanwhile, many actions, such as retrieve, reuse, revise and save must be done in CBR:

Retrieve: using similarity match to search the most similar cases from database

Reuse: using the solutions of similar cases to solve current questions.

Revise: if the historical solution does not fit current questions, adapt it as a base solution to new questions

Save: the adapted solution become a new case, and save it to database, mainly the worthy experience and knowledge.

CBR is mainly consist of case presentation, case index, cases storage, case retrieve, case revise and inductions. Aiming at the particularities of hybrid reasoning feeding decision support system, we will discuss the case presentation and case index as below.

3.2 Case Presentation Based on XML

$CaseBase=\{case_1,case_2, \ldots,case_i\}$,

where $case_i=case(F_i,S_i)$, *present the i-th case of database.*

$F_i=(f_{i1},f_{i2},\ldots,f_{in})$ *present the feature sets of case i,*
f_{in} *present the n-th feature of case Ci*
$S_i=(s_{i1},s_{i2}, \ldots,s_{in})$ *present the solution sets of case Ci*
s_{in} *present the n-th solution of case Ci.*

In rainbow trout feeding decision support system, the case database is consist of many cases, and each case includes three child elements, which are respectively case attribute, feature attribute and solution. The case attribute classifies the case, aiming at structuring an efficient case database in indexing and retrieving. The feature attribute plays an import role in reasoning process, and the retrieval of match case mainly depends on the similarity match of feature attribute, which contains feature name and feature value. The solution is the treatment method or conclusion of corresponding cases.

3.3 Case Index

After saving cases, the system will retrieve and match the case. How to quickly and correctly obtain the case solution is not only an important aspect of evaluation, but also a core part of case based reasoning system. In the process of retrieval, efficiency is not only related to the retrieval algorithm, but also closely with the case database structure. Efficiency is even more critical considering the increasing size of the case database.

Generally, the retrieval indexes of case database are grouped by single index and multi-level index. It is comparatively simple to implement single index, and fit for the case database where not many cases are involved in early phases and the index can be set according to the feature attribute of problems. Therefore, we can have a retrieve and index in accordance with the attribute of the different stages of rainbow trout in the feeding decision system.

In the context of large scale of cases, the multi level index technology presumes there are N cases in the whole database, described as $CaseBase=\{case_1,case_2, \ldots, case_i\}$. At first, we set top level indexes of case attributes according different stages of rainbow trout, and get the first layer child case of *m-classes*:

$$IndexCaseBase= \langle Icase_1, Icase_2, \ldots, Icase_m \rangle \quad , \quad m \leqslant n$$

Then we chose the feature attribute to build the second index that weights more by using clustering method, and built child case of lower levels. As the first index, the *m-classes*

first layer case database contains many specific cases, $Icase_i = \langle case_{i1}, case_{i2}, \ldots, case_{ir} \rangle$, r is the case number in *i-class* layer cases. Finally, those specific cases cluster by the value scale of feature attribute and form the second index. The left layer index can be recursively built by choosing the minor important feature attribute according to the method described above. Fig. 1 shows the case database structured by index tree, using the methods and layer index mechanism, and case feature attribute with value ranges.

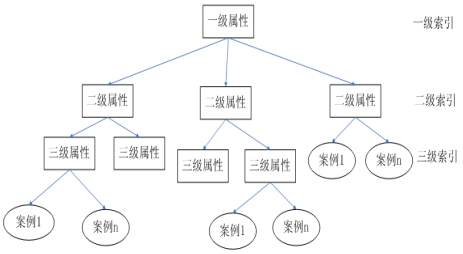


Fig. 1. Multi index of case database

4 Combination of Case Reasoning and Rule Reasoning

Considering the Complexity of decision content and the inner shortage of different decision methods, current decision support system more and more depends on multi methods in combination. Decision support system combined with different decision methods performed great effect on different decision content [8][9][10].This paper combined RBR auxiliary unit and RBR revised unit of CBR that are needed in different stages, and integrated them into a unified module. By implementing and visiting the reasoning dispatch system, we can combine CBR and RBR into the greatest extent, without breaking the dependence and completeness of every reasoning unit. In the reasoning process, whether and when to use the RBR completely depend on the request by CBR system towards the reasoning dispatch system, and once called, this not only insures the parallel implementation of same knowledge in rules and cases, decreases the retrieve count and improves the efficiency, but also keeps the integrity of system strategy and unit itself. Fig. 2 shows the hybrid reasoning frame.

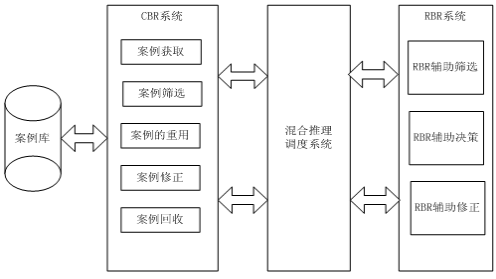


Fig. 2. Combined RBR and CBR

In a fine-feeding decision support system that combines rules and cases, the knowledge database is composed of a group of rules and cases respectively contain cultivation rule knowledge, and feeding case sample. In the reasoning process, the feeding case sample and cultivation rule knowledge support each other, and perform specific decision task together. In accordance with the hybrid reasoning frame combining rules and cases, the specific decision task will be performed as follows:

(1) According to the main impact factors that influence the growth of rainbow trout that described above, the inputted cultivation information of new case should includes:

Culture stage: adult

Water temperature:18 ℃

Dissolved oxygen: 7mg/L

PH: 7

Ammonia: 0.0015mg/L

Protein: 39.5%

Salty: 0.4mg/L

Turbidity: 10NTU

(2) According to the layered structure, we retrieve the first layer classification index in accordance with cultivation stages, and then retrieve the adult case in accordance with the cultivation stage of adult fishes. In the process of retrieval, by calling the RBR auxiliary selection module, and judging from the water temperature and dissolved oxygen in multi layer index, the retrieve will be accelerated.

(3) Calculate the similarity of retrieved case data and new case, and match the case database by KNN (K Nearest Neighbors). If there is a case that exists in case database: $CaseExist=〈e_1, e_2, \ldots, e_n〉$, and a new case :$CaseNew=〈f_1, f_2, \ldots, f_n〉$, S_i represents the similarity of feature i. If the values of feature were quantitative, then $S_i=1-abs(f_i-e_i)/e_i$; otherwise, if the values of feature were same ,then $S_i=1$; else $S_i=0$.

Hence, the similarity can be calculated as formula (1):

$$S = (\sum (w_i \times S_i) / \sum w_i) \times 100\% \tag{1}$$

When the similarity of case match is less than 0.75, after calculating by the rainbow trout feeding decision support system, all the results will be shown as Fig. 3.

Fig. 3. Results of similarity match

(4) In the reasoning process, the frame dispatch engine provides cultivation advice by calling rule reasoning module according to the cultivation information input .The advantage of hybrid feeding is that 1-time retrieval of case in decision process will serve the purpose. The RBR and CBR were performed on one round, with cross support and cross-dependence in decision reasoning process. For example, according to input cultivation information, the system gives following advices:

Feeding time: feed 2 times once a day.

Feed time: 9 :30, feed amount percentage:50%；

Feed time: 15:00, feed amount percentage: 50%.

Particles Fodders feeding advices: 5mm pellet feed

(5) Revise and reuse the case result by calling RBR auxiliary module.
If the similarity of retrieval results cannot meet the expectations, and there is corresponding warning information on feeding, we can adjust in new case in view of the warning information, and add it to the database for reversion which will finally be evaluated by experts.

For the cases with retrieval failure, we directly add it to the case database, and get it reversed to experts. For example, when current water temperature is lower than normal temperature that fits for rainbow trout, the cultivator can adjust the water temperature to maintain a healthy cultivation environment according warning information.

5 Conclusion

This paper presents a feeding decision-making method of combining the CBR and RBR, which made it more efficient in decision retrieval by using the mechanism of case index, and unified the expression forms of rules and cases by using XML to represent knowledge. The decision system combined rule reasoning and case reasoning work in practice has shown that it enhances the professionalism of feeding, and reduces the expert cost in cultivation.

References

[1] Shim, J.P., Warkentin, M., Courtney, J.F., et al.: Past, present, and future of decision support technology. Decision Support Systems 33(2), 111–126 (2002)
[2] Omar, F.E., Brian, D.F.: A web-based muti-perspective decision support system for information security planning. Decision Support Systems 50(1), 43–54 (2010)
[3] Klein, M.: Finsim expert; A KB/DSS for financial analysis and planning. Engineering Costs and Production Economics 17(1-4), 359–367 (2002)
[4] Corchado, J.M., Lees, B.: A hybrid case-based model for forecasting. Applied Artificial Intelligence 15(6), 105–127 (2001)
[5] Mehdi, B., Siavosh, K.: A new method for knowledge representation in expert system's (XMLKR). In: Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, pp. 326–331. IEEE Computer Society, Washington DC (2008)

[6] Li, G.-L., Feng, J.-H., Zhou, L.-Z.: Keyword searches in data-centric XML documents using tree partitioning. Tsinghua Science &Technology 14(1), 7–18 (2009)
[7] Marling, C., Sqalli, M., Rissland, E., et al.: Case-based reasoning integrations. AI Magazine 23(1), 69–86 (2002)
[8] Zhuang, Z.Y., Churilov, L., Burstein, F., et al.: Combining data mining and case-based reasoning for intelligent decision support for pathology ordering by general practitioners. European Journal of Operational Research 195(3), 662–675 (2009)
[9] Ting, S.L., Wang, W.M., Kwok, S.K., et al.: PACER: rule-associated case-base reasoning for supporting general practitioners in prescription making. Expert Systems with Applications 37(12), 8079–8089 (2010)
[10] Liu, K.F.R., Yu, C.W.: Integrating case-based and fuzzy reasoning to qualitatively predict risk in an environmental impact assessment review. Environmental Modelling & Software 24(10), 1241–1251 (2009)

Design and Implementation of Parent Fish Breeding Management System Based on RFID Technology

Yinchi Ma[1,*] and Wen Ding[1,2]

[1] Beijing Fisheries Research Institute, Beijing, 100068, China
[2] National Engineering Research Center of Freshwater Fisheries, Beijing, 100068, China

Abstract. In the breeding process of high quality economy parent fish, the identification and the information storage and management of the parent fish are often through artificial way. This way takes workload too much and is hard to get high reliability. It is difficult to track and manage the breeding process of the parent fish well for us. Based on the RFID (Radio Frequency Identification, RFID) technology, we developed a parent fish breeding management system which can greatly reduce the artificial workload and realize the reliable management and tracking for the breeding process of the parent fish. Test and analysis results show that the RFID Read-Write Handle-net with the antenna of 25 cm or 40 cm diameter can read and write the glass label in the muscle of the parent fish with the average distance of 16.41 cm ~ 22.82 cm. Under the support of the software system, we can realize the reliable management and tracking for the breeding process of the parent fish. The design and realization of the system have a important significance for the factory breeding of the parent fish.

Keywords: RFID, parent fish breeding, identification, storage, transmission, management, tracking.

1 Introduction

Modern aquaculture is gradually changing to facilities and factory model. Requirements for automation and information management are improving continuously now. Especially for breeding process of some high quality economic mother fish, due to the high value individual, the long breeding cycle, the recognition, storage, transmission, management and trace work of the physiology and other attributes information become very important. During the past long-term process of breeding, the recognition of the mother fish is usually through the artificial means. And the breeding worker traces and notes the key information in the notebook. Even by using the computer technology, we are also failure to form a system of management. It is difficult to trace and investigate the whole breeding process of the mother fish effectively, and often wastes effort.

---

[*] Ma Yinchi (1982 -), Beijing Fisheries Research Institute, Engineer, Master, graduated from Beijing Normal University, State Key Laboratory of Remote Sensing Science, mainly engaged in research of agriculture remote sensing and fisheries information technology.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 27–34, 2013.

RFID (Radio Frequency Identification) is a kind of non-contact automatic Identification technology began to be popular since the 1990s. The basic principle is to realize the target automatically by using radio frequency signals and space coupling (inductor or electromagnetic coupling) transmission characteristics. One identification system based RFID is usually consists of four parts including the host machine, the RFID read-write equipment, the RFID tag and the antenna. The system can work without manual intervention, and do well job in many bad environment. RFID is able to identify the fast-moving objects and multiple tags. RFID can realize the information storage and transmission of the target, and the operation is fast and simple. If RFID can work cooperate with the management information system, we can realize the automatic and intelligent information management for the target. For the mother fish, the tag will work in the body of it, the RFID technology can be qualified for the management and tracing of the breeding processing information. At present, in the logistics, tobacco management, experimental animal management and other areas, there have been some mature applications. In the factory aquaculture field, RFID technology is also used in traceability operation for the breeding objects, and has achieved good results. Combining with the actual demand of the mother fish breeding, we design and realize a mother fish breeding management system based on RFID technology. This system realizes the whole processing of identification, storage, transmission, management and tracing of the mother fish. The system will promote the automation informatization level of the mother fish breeding, and bring a important significance for the breeding of the high quality economy mother fish.

2 Materials and Methods

The glass tube RFID tag will be injected in a fixed subcutaneous muscle position of the parent fish. The RFID Read-Write Handle-net can identify the individual information of the parent fish automatically. The information will be transmitted to the PDA or portable notebook computer through the Blue-Tooth. And then the information will be transmit to the center server computer installed the breeding management system of the parent fish through the wireless network. The management system will establish a independent electronic files for every parent fish, and realize the fine management and tracing back of the whole breeding process of every parent fish. The structure of the system is shown as below (Fig.1).

Because the popular Read-Write ware usually uses TI or PHILIPS CMOS chip, such as EM4095, the emission power was limited when the CMOS chip was produced. The Read-Write ware cannot recognize the tag with a longer distance. This way has many limits in some special areas. In this article, the operation of identification and Read-Write is always done in the water, at the same time, the tag is small and the parent fish is usually big, so the Read-Write distance between the tag antenna and the CMOS chip must be limited. We design the RFID Handle-net, the communication method and the management software for the breeding demand. The major function of the system shows below.

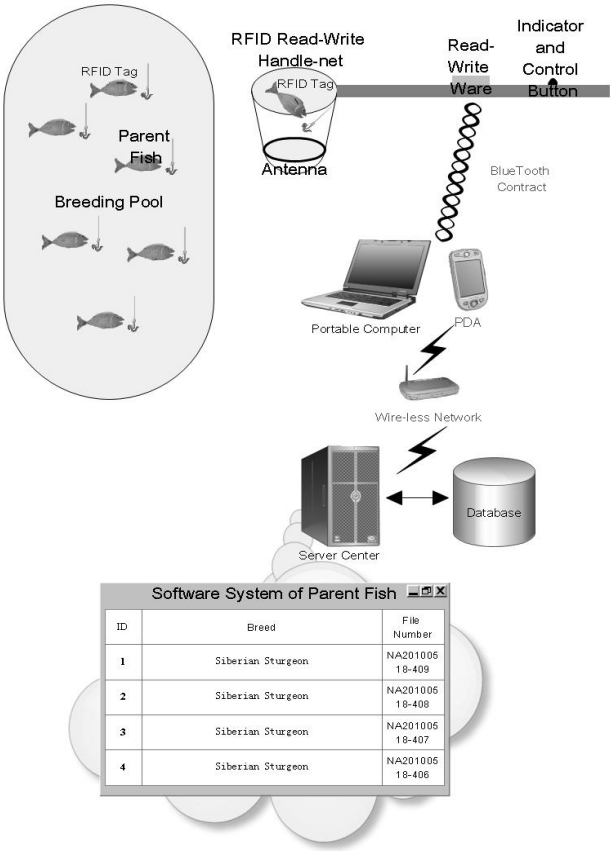


Fig. 1. System Diagram

The first, the system realizes the Handle-net identification, so the parent fish need not to be pulled out of the water, and the damage can be reduced to be the most few.

The second, the identification can be very fast, the RFID Handle-net can identify and read-write data during 3~5 minutes.

The third, the operation is simple, and one operator can finish the whole work. The RFID Handle-net works following the normal work mode, and this way has a big signification for popularization.

The forth, the system uses the wire-less communication way, and the PDA or portable computer can work in the breeding place conveniently.

The last, the system realizes the whole noting of the parent fish, and establishes the electrical file for each one. So we can manage and tracing back the whole breeding information of every parent fish.

The key technology of the system is the RFID Read-Write Handle-net. Low frequency tag can sent back the information automatically when the power big enough given by the reader. In order to read and write the information with a long distance, the

read-write equipment must give the power big enough. At the same time, the noise from the read-write equipment must be small, so that the feeble signal can be received. The system uses the glass tube tag equipped with 125KHz chip. The length of the tag is 12mm and the diameter of it is less than 2mm. This kind of tag can stay in the muscle stably and safely. The injection method is convenience and safely, and the parent fish cannot be hurt. This way has been implied in many areas such as aquiculture and circulation of high quality fish. The technology is much mature now. For the design of RFID Handle-net, we use the components separate method, so as to enlarge the emission power of the read-write equipment and improve the receiving precision. At the same time, we did the special design for the antenna. We use the way of rolling cuprum line to produce the antenna and realize the average read-write distance of 16.14cm~22.82cm by diameter 25cm circle antenna, and the average read-write distance of 16.95~22.09cm by diameter 40cm circle antenna.

The structure diagram of the RFID Read-Write Handle-net is shown as below (Fig.2).

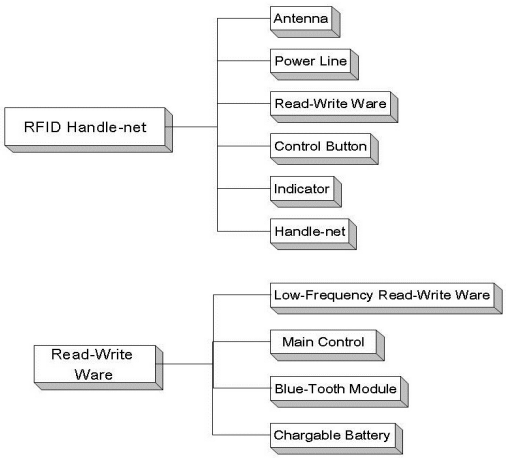


Fig. 2. Structure Diagram of RFID Read-Write Handle-net

The key technology of the RFID Read-Write Handle-net is the validity of the big power identification ware and the antenna when work in the water.

For the effect of multi-antenna identification, we use the components separate method. The power of the CMOS chip can be promoted to two Watt, so that the function of the equipment can reach the best effect (Fig.3).

The multi-route read-write ware with big power is according to the 11784/11785 protocol completely. And it gets the information from the glass tube tag by signal emission, channel selection and signal receiving.

The system software is developed based on VC++ platform and SQL server database, including three parts, such as PDA\portable computer windows client module and database management server module. The architecture scheme of the software is shown as below (Fig.4).

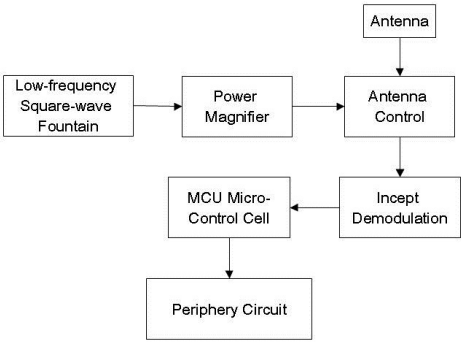


Fig. 3. High power multi-channel identifier

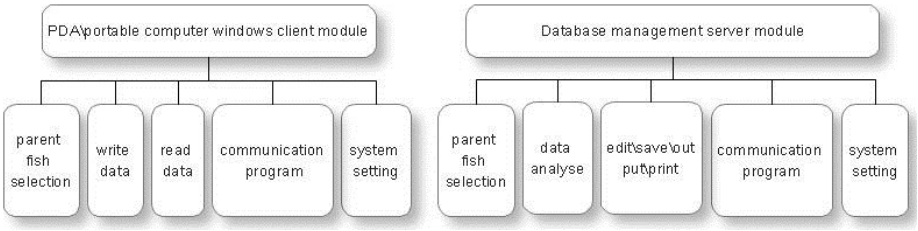


Fig. 4. Architecture scheme of the software

The whole breeding information can be found in the database by the parent fish selection module. Data can be read and written between the tag and the system by the read data and the write data module. Users can analyze the parent fish breeding data by math method through the data analyse module, and check all the life parameter of the parent fish. So users can finished the tracing job. The Bluetooth and WIFI connection and transmission can be done by the communication program. The data management server also can deal the edition, saving, outputting and printing online for the parent fish information. Users can set the system personally by the system setting module.

The interface of the system software is designed compactly, conveniently and friendly. The software function is accord with the need of aquiculture. With the data analyse module, users can check the life and breeding data of the parent fish in time, for example, year of the parent fish, data of its body, pedigree of its family, information of its breed, complexion of feeding, note of propagation, digital case history and so on. Through the analysis by math method, system can give fine data support for aquiculture management. And users can finish the tracing job during the whole breeding process for every parent fish.

The system uses Bluetooth and WIFI communication technology. Bluetooth is used for communication with a short distance. And otherwise, WIFI is used for communication with middle and long distance. Through WIFI network, uses can gather in-phase information from any breeding pool to the information center far away from locale.

The key of the system test is the effect of the RFID Read-Write ware. In the XiaoTangshan experiment base of Beijing Fisheries Research Institute, we took experiment with twelve Koi parent fish. We injected the tag into the muscle under the dorsal fin of every parent fish. Then we tested the RFID Read-Write ware with diameter 25cm an d 40cm separately, and received good results. Following the practice requirement, we design three groups of experiment.

The tag moves on the center point line of the antenna circle. The diagram shows as follow (Fig.5).

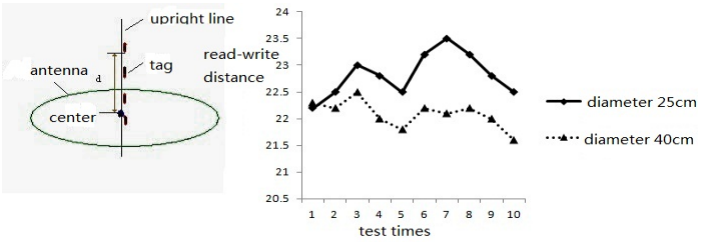


Fig. 5. Test 1 for RFID Read-Write Handle-net

The test results show that, the average read and write distance of the diameter 25cm antenna is 22.82cm, and the max distance reach to 23.5cm. the average read and write distance of the diameter 40cm antenna is 22.09cm, and the max distance reach to 22.5cm. When the tag go into the area of the antenna by the way, the effect of the diameter 25cm antenna is better than that of the diameter 40cm antenna.

The tag moves on the center line between the center of the circle and the edge of the circle. The diagram shows as follow (Fig.6).

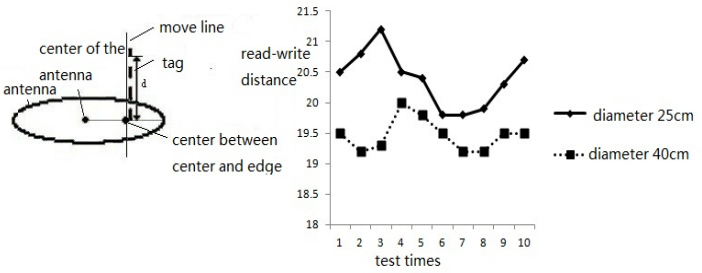


Fig. 6. Test 2 for RFID Read-Write Handle-net

The test results show that, the average read and write distance of the diameter 25cm antenna is 20.39cm, and the max distance reach to 21.2cm. the average read and write distance of the diameter 40cm antenna is 19.47cm, and the max distance reach to 20cm. When the tag go into the area of the antenna by the way, the effect of the diameter 25cm antenna is better than that of the diameter 40cm antenna.

The tag moves on the line through the point on the edge of the circle. The diagram shows as follow (Fig.7).

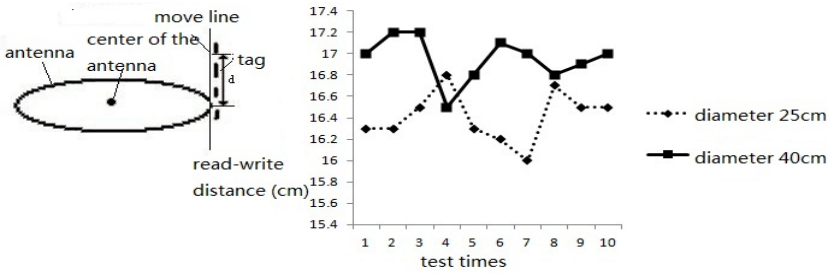


Fig. 7. Test 3 for RFID Read-Write Handle-net

The test results show that, the average read and write distance of the diameter 25cm antenna is 16.41cm, and the max distance reach to 16.8cm. the average read and write distance of the diameter 40cm antenna is 16.95cm, and the max distance reach to 17.2cm. When the tag go into the area of the antenna by the way, the effect of the diameter 25cm antenna is worse than that of the diameter 40cm antenna.

3 Conclusions

The test results from parent fish breeding site show that, RFID Read-Write Handle-net can read and write the data of the tag in the parent fish muscle steadily and veraciously when the tag move in the projection column of the antenna plane among the average distance 16.41cm and 22.82cm. In practice, users can select diameter 25cm or 40cm type antenna due to the complexion. At the same time, we have done the union test with the hardware and software of the system, the data communication state is steady and terminal test is also good for using.

In this article, we use some high value parent fish such as the Koi and the surgeon as research our object. In practice, the breeding technician can read and write the breeding information, physiological information and quality information of the parent fish automatically an d conveniently by using this system. And then they set up database to manage all the information digitally. They can form the scientific breed aquatics project by the tracing information of every parent fish. This method can give ample and credible data support for gene optimization and quality upgrading of the high value parent fish. At the same time, the design target of this system is to provide a new technology method for the high value parent fish breeding. The application environment and effect have a little limit. Design and implementation for other worse breeding environment are emphases of the next step research.

References

1. Xu, W.: Theory and Practice of Parent Fish Breed Aquatics. Aquiculture 12, 28–31 (2004); Chinese reverse to English
2. Zhao, B., Zhang, H.: Application and development of RFID technique. Electronic Design Engineering 18(10), 123–126 (2010); Chinese reverse to English
3. Shen, J.: RFID technology and Analysis of the key technology. China Computer & Communication 3, 146–147 (2010); Chinese reverse to English
4. Wang, Z., Liu, W.: The Application of electronic tag in logistics control system and the design of anti Jamming. Micro Computer 21(7-3), 154–156 (2005); Chinese converse to English
5. Jin, K.: Design and Implementation of Tobacco Input/output Warehouse Management System Based on RFID. Computer Knowledge and Technology 7(4), 813–816 (2011); Chinese converse to English
6. Wen, Z., Xia, L., Qi, C., et al.: The Application of Radio Frequency Identification (RFID) in the Management of SPF Laboratory Animals. Laboratory Animal Science 27(3), 36–38 (2010); Chinese converse to English
7. Chen, M.: Application of RFID Technology in Farm Produce Code-Chain Logistics. Journal of Yueyang Vocational Technical College 26(5), 79–82 (2011); Chinese converse to English
8. Shi, B., Zhao, D., Liu, X., et al.: Application of intelligent system based on traceability and wireless sensor network to industrial aquaculture. Fishery Modernization 38(1), 24–27 (2011); Chinese converse to English

Design and Development of Dissolved Oxygen Real-Time Prediction and Early Warning System for Brocaded Carp Aquaculture*

Huiping Xue, Lianzhi Wang, and Daoliang Li

College of Information and Electrical Engineering, China Agricultural University, Beijing, China
wangcau@163.com

Abstract. In aquaculture process, common water quality monitoring system just collect data in real-time, and data-monitoring that have been collected is time postponed. It is necessary to conduct prediction and early warning on quality in accordance with its historical state and current state. This paper aims to conduct prediction and warning in terms of the DO content in carp aquaculture using neural network and decision tree, and try to complete the dissolved oxygen Real-time prediction and early warning system through prediction and early warning model research, system design, and system development. The effect in practical application shows that the system can use the two methods to predict DO content, and conduct early warning by value prediction and rule based reasoning.

Keywords: data correct, neural network, decision tree, dissolved oxygen real-time prediction and early warning.

1 Introduction

Dissolved Oxygen (DO) is one of the key parameters in intensive aquaculture, the dissolved oxygen content plays a decisive role in affecting the aquatic feeding rate, feed utilization rate and feed conversion rate, and so, it is necessary to monitor the dissolved oxygen. At present, the DO monitor and prediction in pond breeding process cannot meet the need of real production [1], further research is still necessary. In carp breeding process, the equipment that control DO is operated by hand, which is not only waste of cost, but also badly delayed in time. Opening aeration equipment in the case of hypoxia condition has occurred, will inevitably cause losses. The traditional warning and control method just compare the current DO content and specified standards, and open aeration device if it does not meet the requirements, which means bad occurrence is not avoided by precaution.

In this paper, we use the data checking and correction method to ensure the data accuracy and effectiveness that obtained. The method of compensation calculation

* Fund Project: Beijing Municipal Natural Science Foundation (4092024).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 35–42, 2013.

was established after communication with culture experts, and it is based on the practical experience in breeding.

The system was deployed in Yixing agriculture network platform and Xiaotangshan carp farms, and the results of testing and operation achieved a desired effect comparatively. The real-time dissolved oxygen prediction and warning made breed staff can aware of the dissolved oxygen content for the moment and a period of time later (10minute, 30minute, and 60minute).

2 Theory

2.1 Data Check and Correct

Remote data checking module will check the obtained data, and call modification module and check data if the abnormal data was found, and make it into reasonable range. The specific processes is, firstly, census data and analyze the data distribution, find the maximum and minimum value, secondly, communicate with the breeding experts and after comprehensive analysis, get the range of parameters. The proportion of parameter changes is calculated by following formula:

$$\pm \max\left(abs\left(\frac{y_{i+1} - y_i}{y_i} \right) \right) \tag{1}$$

Among them: y_i represents the obtained data in time t, y_{i+1} represents the data collected in the moment $t+1$.

Calculation method of data compensation:

$$\overline{y} = \frac{1}{base} \sum_{i=1}^{base} w_i y_i \tag{2}$$

Among them: $\overline{y}$ represents the to-be-checked data, $base$ represents the number of compensation, y_i represents the historical data, w_i represents the weight. The number of compensation refers to the number of data that referenced in order to correct and compensate data when the data is missing or error.

2.2 Real-Time Prediction and Warning of DO

Natural Network

Neural network simulates human brain memory and learning activities, and can solve problems such as classification, identification and prediction [2]. BP algorithm is a kind of feedback neural network algorithm, and is realized by multiple iterations. A learning process is composed by forward propagation of data input and back-propagation of error. In this paper, we considered the actual situation of carp breeding, and take three layer network structures.

Decision Tree

Decision tree is a kind of association rules, and is based on the inductive learning method as the given sample, using up-side-down recursive manner to produce tree structure that similar to flowcharts[3]. From the root node, choose the most appropriate described attributes as branching attribute with reference of the given metric standard, and establish branches down according to different values. The advantages of decision tree classification method are: less time-consuming, simple and intuitive model, easy to understand. The C4.5(an improved algorithm of ID3) algorithm can not only deal with discrete description attribute, but also continuous description attribute. It selects the attribute that have the maximum information gain ratio and classify the training samples, aims at minimizing the system entropy when branching, and improve the computing speed and accuracy. The simulation results that using decision tree to predict dissolved oxygen content are shown as Fig. 1.

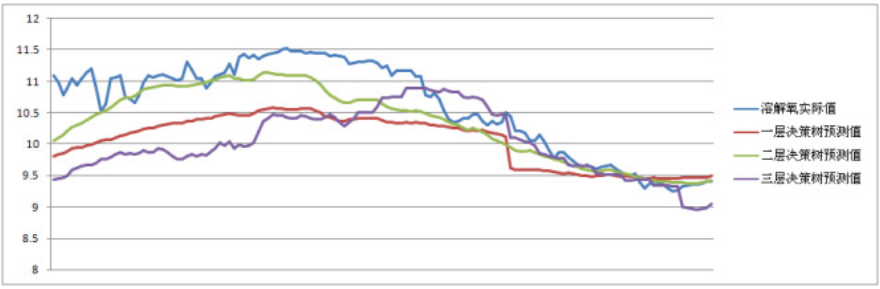


Fig. 1. DO predict results of decision tree

DO Warning

Warning refers to measure a factor in present and the future, and predict abnormal state and possible injury [4]. On the early warning process, firstly, we analyze the surrounding environment and conduct early warning on object's internal factors in quality and quantity. Secondly, identify the trends, speed and range. Finally, do the warning and remind for events that might occur and measures that can be taken. The DO warning means to analyze and evaluate the DO content of water body in a certain period of time, and determines the change trends of dissolved oxygen content, and predicts the abnormal condition that maybe happen. In the process of freshwater pond breeding, the water quality warning is sudden, lagged, complex, concentrated, and dynamic [5].

This paper do the numerical prediction according to the prediction results of neural network and decision tree, and also calculate the warning state through rule based reasoning of the various parameters of water quality and their respective effect relationship.

3 System Design

3.1 Data Collection and Database Design

The real-time database is a supporting part in development of real-time control and data acquainting system[6], which can help user collect and store data in real time and provide effective data sources for information mining in the upper layer software. Real-time database also contributes to the monitoring and optimization control, and also provides real-time data service, data management, scheduling, data analysis, decision support and remote online browsing for the enterprise production process. The database call relation of this system is shown as Fig. 2:

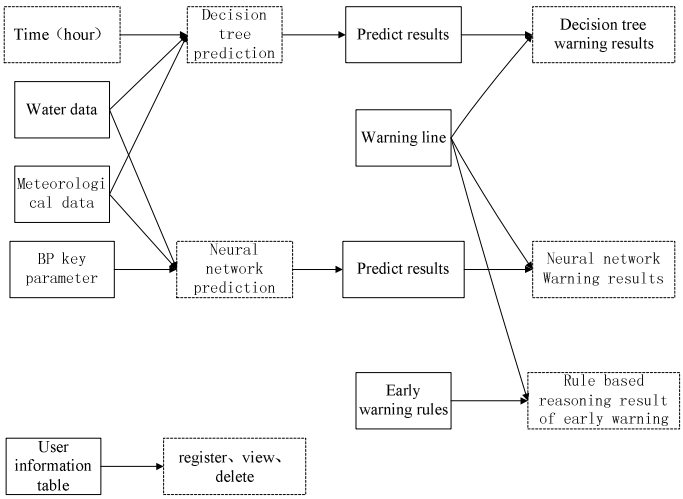


Fig. 2. Database design

3.2 The Frame of Real-Time Prediction and Early Warning System

The Real time prediction and early warning system can automatically monitoring, predict and warn on irregular data, and at the same time respond to users' requests such as viewing, modification, and so on.

The workflow of this system varies by role. First it verifies the user's jurisdiction, then responses to different service contents according to different roles. Fig. 3 shows the framework of system.

Common user: Users request for viewing the data and perform prediction, and the system responds to them. If user views the historical data, then the system calls the database of historical data directly and displays the corresponding results on the HTML page. If user views the predicted or early warning information, then the system first do the calculation process and gets the needed data, then store it into database and finally show them on a page after calling the database. The aim is: the database can save each step during handling process, which not only provides a guarantee for users' viewing, but also provides traceability data and log records when error occurs in future.

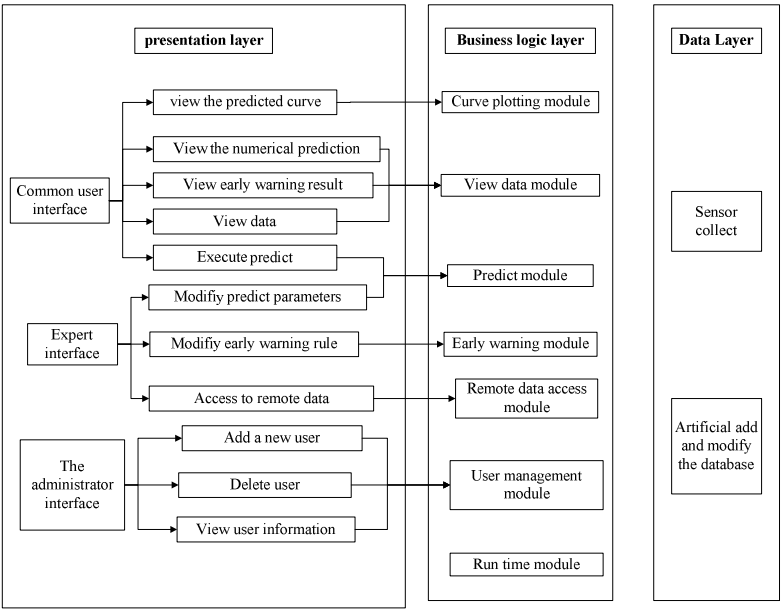


Fig. 3. System framework

Expert: Expert is responsible for updating the local database, modifying the relevant parameters of prediction and warning, such as parameters of neural network prediction, rule base of warning reasoning.

Administrator: The administrator is mainly responsible for the users data maintenance work, including viewing user information, deleting users and other functions.

In short, the system implement functions according to the request of users, while keeping a real-time prediction and warning function, and perform remote monitoring data, predicting dissolved oxygen content and executing early warning functions no matter whether there is any manual interaction.

4 System Development and Implementation

4.1 MVC Model

MVC is the abbreviation of Model, View and Controller [7]. Both The model and controller can update the view layer content, and the controller can also change the state of the model. The view layer passes the user's operation to the controller and at the same time gets a new state from the model.

The view layer is the pages that user can see, usually made of HTML, JSP, ASP, which contains form, servlet path that processing page request,i.e. the URL address of action.

Control layer processes business of view and model layer, and reads the page information, in response to user request and interactive model, and finally return information to page, which was generally handled by the configuration file web.xml.

Model layer is mainly composed of Java files, classes and servlet definition files.

According to the research object, the JSP page consists with user needs, and the model is responsible for realizing the various functional models, and servlets (control layer) response page and call for the corresponding model.

4.2 The Implementation of DO Real-Time Predict and Early Warning System

The prediction module includes two kinds of methods: neural network prediction and decision tree prediction. These serve for dissolved oxygen prediction, and belong to the same prediction module, including performing prediction, saving the results, showing the results. However, they are different in the principle of realization.

4.3 Neural Network Prediction Module

Neural network prediction module learns historical data, and input the recently acquired data to the network, then predicts the DO content in the next period. The module needs to be normalized .Neural network prediction process is divided into two main parts: preparation stage and working stage. In the preparation stage, there are two parallel activities: access to data and normalize, instantiate prediction class and acquaint parameters .In the working stage, it trains model firstly, and then uses the trained model to perform the final prediction, and save the prediction result into database, so that it contributes to the viewing and using.

4.4 Decision Tree Predict Module

This paper established the three-layer decision tree prediction module according to the prediction model previously built, and depth of 1, 2, 3 of the decision tree was set up and to predict dissolved oxygen content. Each prediction result was stored into the database, so users can view it. The implementation process of this module is divided into two parts: calculate the expected information that training set needed when doing classification, and information acquisition ratio of each attribute when the training set is partite.

4.5 DO Early Warning Module

Warning module includes two kinds of methods: according to the prediction value and rule reasoning, in which the prediction value is divided again for the predicted values of the neural network and decision tree to predict value.

According to the prediction value of early warning, each alert level is stored into database. When performing the warning, system reads warning line from database, and then the prediction value and early warning limits are compared, and finally the result of early warning is given.

When reasoning according to the rules, the system reads reasoning rules from database, calculates and infers, and gets the result of early warning. In addition, the role of experts can also modify warning rules.

5 System Testing and Running

In prediction module, the default parameter list is stored in the database, and is read from database when in use. This can not only ensure that the breeding expert modifies

the parameter into reasonable value according to their own culture experience, but also make interaction with the system possible, while the system is able to respond to the requirements of experts and calculate a more accurate prediction value.

Table 1. Default parameter list of predict module

Predict time	Default sample number when training	default learning times
10min	500	3000
30min	400	7000
60min	300	10000

After reading the relative prediction parameters, system performs the dissolved oxygen content prediction modules in different times, and gives the corresponding prediction value. The prediction results for 10 minutes are shown Fig. 4.

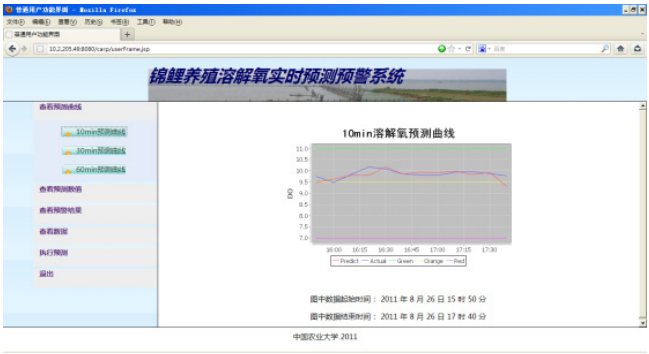


Fig. 4. Dissolved oxygen prediction curve in 10 minutes

In addition, the user can view the prediction results of dissolved oxygen content in 10 minutes, 30 minutes, and 60 minutes. Fig. 5 shows the predict value of DO content using BP algorithm.

Fig. 5. Predict results

In the parameter management module, the expert can view and modify the critical parameters, including learning rate, momentum factor, the number of hidden layer nodes, learning steps, and training times, and etc.

6 Conclusion

Taking the carp breeding as research object, this paper analyzed DO prediction and warning method in the practical breeding process, and developed a real-time prediction and warning system using Java programming technology. This paper has certain guidance significance in the practice of aquaculture production, and at the same time has reference value for other species culturing. In addition, breeding needs experts to be participating, especially in making the inference rules.

Acknowledgements. Research work of this paper was supported by The Beijing Natural Science Fund Project (No 4092024) .Thanks to the domain experts in Beijing Freshwater Fisheries Technology and Engineering Research Center for their long term support.

References

[1] von Metzen, R.P., Stieglitz, T.: A Wireless System for Monitoring Polymer Encapsulations. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2007, August 22-26, pp. 6600–6603 (2007)

[2] Caudell, T.P., Zikan, K.: Neural network architecture for linear programming. In: International Joint Conference on Neural Networks, IJCNN, June 7-11, vol. 3, pp. 91–96 (1992)

[3] Abdelhalim, A., Traore, I.: A New Method for Learning Decision Trees from Rules. In: International Conference on Machine Learning and Applications, ICMLA 2009, December 13-15, pp. 693–698 (2009)

[4] Han, L., Wang, J., Lu, C., Xie, J.: Research on early warning system of water quality safety based on RBF neural network model. In: 2010 2nd International Conference on Information Science and Engineering (ICISE), December 4-6, pp. 4667–4670 (2010)

[5] Wang, R., Chen, D., Fu, Z.: AWQEE-DSS: A Decision Support System for Aquaculture Water Quality Evaluation and Early-warning. In: 2006 International Conference on, Computational Intelligence and Security, November 3-6, vol. 2, pp. 959–962 (2006)

[6] Lu, H., Zhou, Z.: The researching and application of historical data processing in real-time database system. In: 2010 International Conference on Computer Application and System Modeling (ICCASM), October 22-24 , vol. 14, pp. V14-162–V14-166 (2010)

[7] Wojciechowski, J., Sakowicz, B., Dura, K., Napieralski, A.: MVC model, struts framework and file upload issues in web applications based on J2EE platform. In: Proceedings of the International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, February 28, pp. 342–345 (2004)

A Greenhouse Control with Sectional-Control Strategy Based on MPT Intelligent Algorithm

Fengyun Wang, Lin Mei, Wenjie Feng, Lei Wang, Limin Wang, and Huaijun Ruan*

S&T Information Engineering Research Center, Shandong Academy of Agricultural Sciences, Jinan 250100, Shandong Province, P.R. China
wfylily@163.com

Abstract. The greenhouse are classified as complex systems such as large lag, multi-input multi-output (MIMO), non-linear and difficult to create mathematic model and so on, so it is difficult to implement classical control methods for this kind of process. A new sectional-control strategy is put forward. It uses different control modes for different deviation domains. The new strategy is based on MPT intelligent control algorithm which improves on the traditional PID algorithm and adds in self-adapting, fuzzy control, expert self-tuning etc intelligent control functions. When the deviation exceeds a certain domain, fuzzy control is used to prevent the saturated integral; when the deviation reduces to within a certain domain, MPT algorithm is used to reduce the overshooting during response process and eventually eliminate the residual. The application in greenhouse control shows that the sectional-control strategy makes the output tracing the set value correctly.

Keywords: Greenhouse, MPT, Sectional-control, Fuzzy.

1 Introduction

The agricultural greenhouses were used to protect the crop against the weather changes. With technical progress, the greenhouses have become a production means used to control the crop environment in order to obtain higher quality [1]. To achieve environmental conditions favorable for plant growth, greenhouses are designed with various components, structural shapes, and numerous types of glazing materials. They are operated by hand or by control system differently according to each condition.

The control object of greenhouse is the environment system which factors such as humidity, temperature, CO_2 density and so on are affected not only by external climate, but also by the interaction between the environment and the plant. In control system, they are also controlled by the control components [2]. Therefore, the greenhouse system is a non-linear, multi-input multi-output (MIMO) complex system they present time-varying behaviors and they are subject to pertinent disturbances

* Corresponding author. Supported by: National science and technology support plan of P.R. China (2011BAD21B06), Science and Technology Development Plan of Shandong Province, China (2011GGC02035).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 43–50, 2013.

depending generally on meteorological conditions [3]. All these make it difficult to describe a greenhouse with analytic models and to control them with classical controllers.

The traditional PID algorithm was widely used because it needn't know the model of controlled object. But it is difficult to realize the high control accuracy when reduce the overshooting. In our case, we put forward a new MPT algorithm which improves the PID algorithm and adds in self-adapting, fuzzy control, expert self-tuning etc control functions. On the base, a new sectional control strategy is created. When the deviation is beyond a certain domain, fuzzy control rules are used to eliminate the integral saturation. When the deviation is within the domain, MPT is used to reduce the overshooting and eventually eliminate the residual. The feature of sectional strategy is that it needn't consider the accurate mathematical model of controlled object. Finally its application in greenhouse is given and the results show that it can obtain the ideal effects.

2 MPT Algorithm

The simplified expression of traditional PID algorithm is as follows:

$$\text{Output= Proportional action}(P)+\text{Integral action}(I)+\text{ Derivative action }(D) \tag{1}$$

In traditional PID control system, it is difficult to reduce the overshooting at the same time improving the control accuracy. The main reason is the integral action. If it reduces the integral action, the residual isn't easy to eliminate, and when there is disturbance, the speed to eliminate the error is slow. If it enhances the integral action, the oscillating process intensifies and it is hard to avoid overshooting. Therefore, PID algorithm is partly improved as follows:

$$\begin{aligned}\text{Output= Proportional action }(P)\text{ + Integral action }(I)\text{ + Derivative action}\\ (D)\text{ + Derivative-integral action }(\int I)\end{aligned} \tag{2}$$

Because derivative-integral action is added in formula (2), the integral saturation is greatly improved. But we can see from formula (2) that there is a new parameter ($\int I$), therefore these parameters must affect each other to make the new parameter more difficult to confirm. Through serious research and experimental analysis, the ratio between proportional action and derivative action is same to the ratio between the integral action and derivative action and the best ratio is related to the lag time of controlled object. If the lag time is longer, the response of proportional action decreases and the response of derivative action increases. The relationship is as follows:

$$\text{Proportional action= } K(1/t) \tag{3}$$

$$\text{Derivative action= } K(1-1/t)d \tag{4}$$

Where, K -Coefficient; t -Ratio between lag time and control period, $t \geq 1$; d -Derivative action. Then the formula of improved algorithm of PID is:

$$\text{Output}= P[1/t + (1-1/t)d] + (1/M)\int [1/t + (1-1/t)d] \tag{5}$$

In formula (5), P is used to regulate the derivation and proportion. When P increases, it increases the derivative time and decreases the proportional band at the same time. Verse On the contrary, when P decreases, it decreases the derivative time and increases the proportional band. M is similar to integral time to regulate the integral action and derivative-integral action. t is used to regulate the mutual proportion between derivative action and proportional action. If $t = 1$, the derivative action is zero. If $M = 0$, the integral action is zero.

Thus the control parameters are reduced to 3. Because the traditional PID parameters are only according to the algorithm self and its feature is that it needn't know the accurate model of controlled object. Because three improved parameters differ from the original parameters, the improved PID algorithm is defined as MPT control algorithm and the implication of three parameters as follows:

2.1 M

M50 is retaining parameter. The definition of M50 is the difference between stable measured value when the output value is 50% and stable measured value when the output value is 0%. Smaller the M value is and stronger the integral action is. Larger the M value is and weaker the integral action is (integral time adds). If $M = 0$, the integral action is cancelled.

2.2 P

P is rate parameter. P is inverse proportional to the change of measured value when the output changes 100% within one control period. The definition as follows: $P = 100/$ change of controlled parameter every second. The best value of P is usually confirmed by self-tuning method. P value affects the proportional action and derivative action. Proportional action and derivative action increase proportionally with P value increasing. Proportional action and derivative action weaken correspondingly with P value decreasing. P is unrelated to integral action.

2.3 T

T is lag time parameter. T is the time required when the output rises to 63.5% of maximum value and its unit is second (s). T is introduced and set correctly to overcome the overshooting and oscillation completely and make the response at the best speed. T affects the proportional action and derivative action. Smaller T value is, stronger the proportional action is and weaker the derivative action is. Larger T value is, the proportional action weakens and derivative action enhances. If T is less than or

equal to control period, the derivative action is cancelled completely and when the regulation becomes proportional or proportional-integral.

3 Intelligent Sectional Strategy

The intelligent control is with some intelligent characteristic as follows: (1) Integrate the human's experience with control theory. (2) With the capacity of on-line studying, modifying and creating new knowledge. (3) Deal with qualitative and quantitative information and fuzzy and accurate information. (4) With stronger logical reasoning and analytical decision-making than traditional control [4]. The intelligent control is the development of traditional control theory and combination of control theory, computer technology and artificial intelligent technology. Therefore it is with capacity of self-adapting, self-studying and self-organization[5].

3.1 Self-adapting

The greenhouse is a complex control object with severe non-linear characteristic and intercoupling between parameter and parameter which is solved by self-adapting control mode. The feature of self-adapting is that when the controlled deviation is larger than the estimated error, the self-adapting system doesn't modify MPT parameters but modify the output to lower the error. Though the modified range is limited, the phenomenon that original correct controlled parameters are modified to wrong value doesn't happen to make the system respond rapidly which greatly improve the control accuracy.

3.2 Self-tuning

In the intelligent algorithm, the traditional PID algorithm parameters have been modified to MPT parameters. In order to confirm MPT parameters, a set of self-tuning expert system is introduced. Because MPT parameters are described for controlled object, its self-adapting and self-tuning are simpler than traditional PID parameters and accurate. In general case, if the operation of self-tuning is correct, its success rate is almost 100%.

The process of self-tuning uses position control to regulate the system. After the oscillation, the lag time parameter T is confirmed by the period and the rate parameter P is confirmed by oscillation amplitude. It isn't easy to confirm the parameter M directly. For temperature, it is usually assumed that the measured value is 25℃ when the output is zero and parameter M is confirmed in accordance with the output when oscillating. For linear input, M value is its scale range. Therefore the best values of parameter M and T can be obtained by self-tuning, but parameter M is only rough. In addition, if the rate parameter or lag time is very long, self-tuning may also increase the control period to make the system conform to the actual requirements of controlled object.

3.3 Fuzzy Control

The classical fuzzy controller uses the fuzzy set theory to directly convert the language rules formed by expert knowledge or operator's experience into automatic control strategy (usually the fuzzy rules table) which doesn't rely on the accurate mathematic model of controlled object but use its language knowledge model to design and modify the control algorithm [6,7].

When the controlled deviation is very large, it hasn't too large significance for the output by formula regulation which inversely produces integral saturation. Though it uses MPT algorithm, if the controlled parameters are set incorrectly, the integral saturation or over-integral phenomena will also occur. Therefore when the controlled parameters are beyond the proportional band, the fuzzy control is used to confirm the output which can obtain ideal control effect.

3.4 Sectional Strategy

The sectional strategy adopts different control modes in the deviation domain. Its block diagram is shown in Fig. 1. When the deviation is larger than a certain domain, the fuzzy control is selected by operation mode selector to overcome integral saturation and confirm the output by fuzzy rules. When the deviation reduces within a certain domain, selector switches to MPT algorithm to decrease the overshooting during responding and eventually eliminate the residual. Therefore the sectional strategy integrates the merits of MPT algorithm and fuzzy control to improve the sensitivity and accuracy and obtain the ideal control effect.

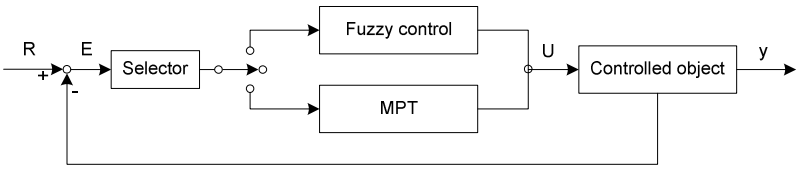


Fig. 1. Block diagram of sectional strategy

4 Application in Greenhouse

4.1 Control Parameters

The application is in a new built phalaenopsis greenhouse to effectively combine phalaenopsis planting with greenhouse control design to realize the optimal growth environment of phalaenopsis. The parameters that affect the growth of phalaenopsis are light intensity, temperature and humidity etc. The main parameters we will discuss hereinafter are light intensity and temperature.

4.1.1 Temperature

There is a critical temperature, 25℃, between the vegetative growth and reproductive growth of phalaenopsis and the critical state is during the flower spikes appearing. During the period of peduncle growing out, the temperature should be kept at 20~25℃ in daytime and 22~18℃ in nighttime. This period will last 20~30 days until the peduncle grows out. The sensitivity to temperature for various phalaenopsis is various. The minimum temperature for phalaenopsis is 2~3℃ in dry environment and 15℃ under general humidity condition. During vegetative growth phase, the temperature is usually over 25℃.

4.1.2 Light Intensity

The light needed by phalaenopsis growth is lower and continuously increases with its growth. Light intensity in various growth phases is shown in Table 1.

Table 1. Light intensity in various growth phases

Growth phase	Bottle seedling	Little seedling	Middle seedling	Big seedling	Mature plant
Light intensity (lux)	5000~7000	7000~10000	10000~15000	15000~20000	20000~40000

4.1.3 Humidity

The requirements of humidity for seedling and flowering of phalaenopsis are various. The ideal humidity during the period of seedling is from 70% in daytime to 90% in nighttime. During this period, the higher the humidity is and the quicker the seedling grows but easier the germ grows. The ideal humidity during the period of mature plant is from 60% in daytime to 80% in nighttime.

4.2 Control System

4.2.1 Hardware

The system is composed of measuring loop, controller and output loop. It may include one computer and printer depending on the requirements. In according to the production requirements, the measuring loop mainly measures the light intensity, temperature and humidity. So the sensors are light intensity sensor, indoor and outdoor air temperature sensors, soil temperature sensor, and indoor and outdoor air humidity sensor and soil humidity sensor. The controller mainly collects the signals from the sensors, then carries out A/D converting and analysis at last makes decision to the output loop, drives the relay after D/A converting so as to control the control equipments and adjust the environment conditions. The output loop mainly includes the inner and outer shade net motors, side window motor, wet curtain motor, fan, circulating motor and spray irrigation trolley.

4.2.2 Software

The system adopts MPT intelligent algorithm based sectional strategy. The develop environment for controller is Windows2000 ＋KeilC7.01＋Wave＋Labtool48. The operating enviroment is W77E58 ＋ 32KBFlash+watchdog+extending parts. The software realizes the parameters set, data and time display, alarm information hint and online state indication etc[8].

5 Application and Conclusion

The system was installed in the phalaenopsis greenhouse in Jining Agricultural Hi-tech Demonstration Zone. The ideal values of parameters are set in according to the phalaenopsis growth model [9]and expert knowledge library[10]. Fig. 2 is the comparison curve between actual measured values by the indoor and outdoor air temperature sensors and the ideal value from zero clock on 24th, Sep., 2006 to zero clock on 25th, Sep., 2006. Fig. 3 is the comparison curve between actual measured values by the indoor and outdoor light intensity sensors and the ideal value from zero clock on 24th, Sep., 2006 to zero clock on 25th, Sep., 2006.

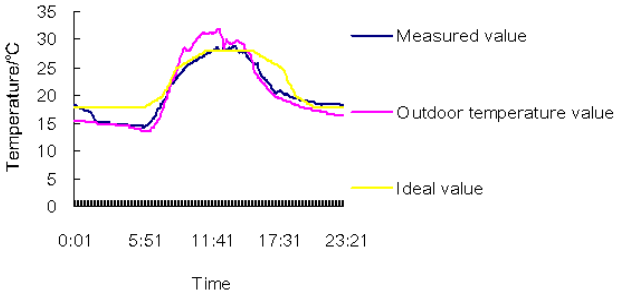


Fig. 2. Comparison curves between the measured temperature in greenhouse, outdoor temperature and ideal value

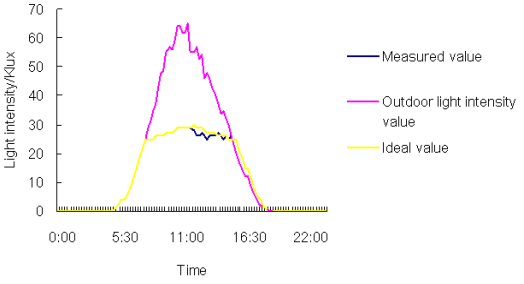


Fig. 3. Comparison curves between the measured light intensity in greenhouse, outdoor light intensity and ideal value

When it didn't take warming measures, the greenhouse only kept the inner temperature by the heat preservation of itself in night and adjusted its temperature by side window, wet curtain and fan etc. It didn't take supplementary lighting measure in night and controlled the light intensity by inner and outer shade net. We can see from the curve that the sectional strategy based system can trace the set values well.

References

1. Trabelsi, A., Lafont, F., Kamoun, M., Enea, G.: Fuzzy identification of a greenhouse. Applied Soft Computing 7, 1092–1101 (2007)
2. Wang, D.: SVM Regression Modeling for Greenhouse Environment. Transactions of the Chinese Society for Agricultural Machinery 35(5), 106–109 (2004)
3. El Ghoumari, M.Y., Tantau, H.-J., Serrano, J.: Non-linear constrained MPC: real-time implementation of greenhouse air temperature control. Computers and Electronics in Agriculture 49, 345–356 (2005)
4. Wu, H., Xie, Y., Li, Z., He, Y.: Intelligent control based on description of plant characteristic model. ACTA Automatica Sinica 25(1), 9–17 (1999)
5. Fourati, F., Chtourou, M.: A greenhouse control with feed-forward and recurrent neural networks. Simulation Modelling Practice and Theory 15, 1016–1028 (2007)
6. Zhang, H., Yang, Y., Chai, T.: Current state and development multi-variable fuzzy control (I). Control and Decision 10(3), 193–203 (1995)
7. Trabelsi, A., Lafont, F., Kamoun, M., Enea, G.: Fuzzy identification of a greenhouse. Applied Soft Computing Journal 7(3), 1092–1101 (2007)
8. Fitz-Rodríguez, E., Kubota, C., Giacomelli, G.A., Tignor, M.E., Wilson, S.B., McMahon, M.: Dynamic modeling and simulation of greenhouse environments under several scenarios: A web-based application. Computers and Electronics in Agriculture 70, 105–116 (2010)
9. Zhang, X., Liu, F., Wang, F., Liu, S., Feng, W.: Design of implementation on dynamic vegetation model of greenhouse phalaenopsis. Chinese Agricultural Science Bulletin 23(11), 398–402 (2007)
10. 2007SR15256.2007-2-1

Research on Digital Construction of Crop Plant Type Based on a Kind of Improved Functional-Structural Model and Component Technology

Zhenqi Fan*, Chunjing Si, and Quanli Yang

College of Information Engineering, Tarim University, 843300 Alaer, China
fanzhengqibtu@126.com

Abstract. Taking digital construction of crop ideotype as the target, current functional-structural model was improved, which discussed the biomass production, biomass allocation and organ reconstruction, the division of crop organ components and the way of component connections was put forward, and the internal structure of the components was explored to leaf organs as an example. At last, the process of digital construction of crop plant type was elaborated, which laid foundation for the plant type quantitative design.

Keywords: functional-structural model, componentization, crop plant type, digital construction.

1 Introduction

Botanical studies have shown that crop plant type has an important influence on crop yield[1,2]. By improving the spatial structure characteristics of crop, the access efficiency of the plant to light, water and other resources can be improved, thus, crop photosynthetic production can be distinctly increased, but traditional plant type breeding experiment is a long cycle, high cost and subject to the natural environment. With the deepening of research on virtual crop, quantitativly creating the plant type to meet the special requirements becomes a possible way to solve the above problem. To some extent, crop life activities characterize a result of the interaction between crop morphology, physiological and ecological processes and environment[3], so in order to reflect the objective reality of crop growth process, parallel simulation of crop structure and function must be carried out[4], thus one of the basic problems of using virtual crop technology to construct ideotype is the establishment of the crop functional-structural model.

In common crop functional-structural model, as shown in Figure 1, after the completion of the biomass production and allocation, only carrying on morphological reconstruction of shape factor of plant organs, such as organ length, width and area,

* Supported by the Natural Scientific Foundation of China (61062007); the Principal Fund Project of Tarim University (TDZKSS201115). Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 51–57, 2013.

but now organ not only occured morphological changes, there were the relative changes of the position and inclination, especially the leaf inclination angle has great significance on plant getting energy. Thereby, in the process of organs reconstruction, only to consider the shape factor will affect the proper biomass production and allocation of the next growth phase of organ, so that it will make greater degree of distortion compared with the reality of crop morphology.

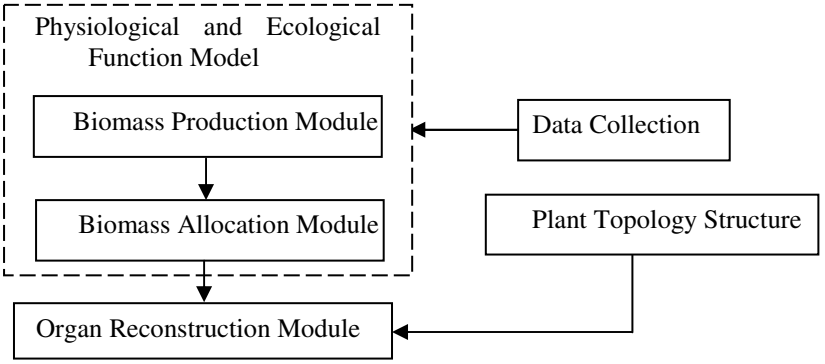


Fig. 1. Functional-Structural Model of Crop

2 Improved Functional-Structural Model

Aming at the defects of current functional-structural model, by summary of organ location factors and the horizontal, vertical variation laws, and real-time control of the dynamic changes in plant growth, morphological reconstruction of the crop plant was achieved in the way of functional-structural mutual feedback, as shown in Figure 2.

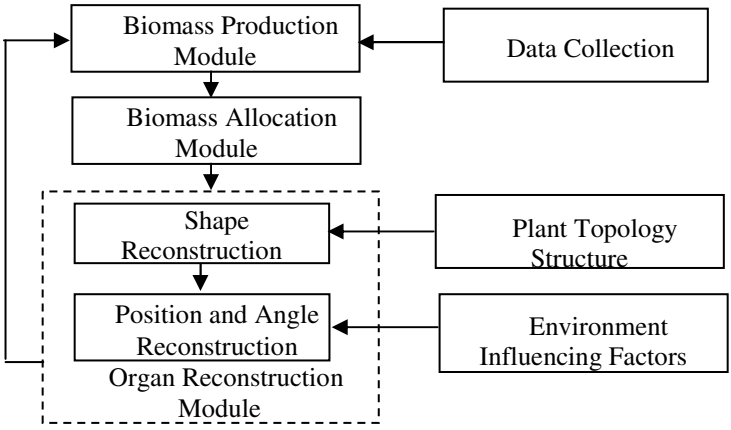


Fig. 2. Improved Functional-Structural Model

2.1 Biomass Production Module

From the perspective of biomass consumption, the produced biomass *M(t)* was studied in the No. t unit growth cycle by using the crop transpiration, structure composition and hydraulic resistance in the body [5].

$$M(t) = \sum_{i=1}^{N(t)} \frac{Energy(t)}{\frac{r_1}{Area(i,t)} + r_2} \tag{1}$$

Here, *Energy(t)* refers to the growth potential of fresh material under the control of water potential differences between inside and outside crop body and water use efficiency in the No. *t* growth unit cycle; *N(t)* refers to the number of leaves which are able to do photosynthesis in the No. *t* growth unit cycle; *Area(i, t)* refers to the area of the leaf with the serial number of *i* in the No. *t* growth unit cycle; r_1 refers to the hydraulic resistance of the leaf; r_2 refers to the total hydraulic resistance of the other parts except the leaf. However, some parameters of the model (such as *Energy(t)*, r_2) are difficult to measure.

From the energy conservation principle, combined with crop photosynthesis, the calculation method of crop biomass production was studied. [6]:

$$\int_{M_1}^{M_2} dM = \frac{1}{\sigma} \int_{t_1}^{t_2} r(1-g)(1-\lambda) I dt \tag{2}$$

Here, M_1 and M_2 represent the amount of crop dry mater at time t_1 and time t_2 respectively; σ is the energy released by burning one gram of dry matter; r is the solar energy use ratio per unit green leaf area; g is the crop respiration rate; λ is reflection leakage rate of radiation which is projected on crop plant; I is the surface radiation intensity. Based on the parameters of crop morphological characteristics (for example leaf area index, leaf azimuth), the biomass *M(t)* produced by the photosynthesis in the No. *t* growth unit cycle can be calculated.

2.2 Biomass Allocation Module

Botanical studies have shown that, in the process of transporting and distributing biomass to the organs by vascular, the transmission path has few impact on the distribution, which depends on the type of organ. Different organs (such as leaves, petioles, stems) have different competitiveness of biomass (usually represented by sink strength) and different expansion rates of themselves. The sink strength *s* of the organ can be defined as the coefficient of thermal ages [7]; expansion rate *m* is generally consistent with beta probability density function [8]. If the growth age of an organ is set as *y* (set the growth cycle as the unit), then the expansion amount of this organ in the No. *t* growth unit cycle is:

$$\Delta M(y,t) = \frac{s(t)m(t-y+1)}{\sum s(t)(\sum_{i=1}^{t_0} m(t-i+1)} M(t-1) \tag{3}$$

Here, *s(t)* refers to the sink strength of an organ in the No. *t* growth unit cycle; *t-y* refers to the actual growth age of the organ; t_0 refers to the number of organ expansion cycle.

Therefore, the total biomass *M(y, t)* of the organ with the growth age of *y* in the No. *t* growth unit cycle is:

$$M(y,t)=M(y,t-1) + \Delta M(y,t) \tag{4}$$

2.3 Organ Reconstruction Module

Based on biomass allocation, various organs, through the obtained biomass, completed the crop secondary growth or formed other new organs. Current software with powerful data analysis functions (such as Matlab) can be used to analyze the fresh weight, morphological factors, growth position and angle of the collected organs, and the data are conducted fitting by least squares or other methods, obtaining the intrinsic relationship between the fresh weight, morphology factors(such as organ length, width), location and angle to establish the function-structure coupling and morphology controllable model.

3 Components of Crop Organs

Based on object-oriented programming, further abstraction is made in the component design idea. The basic idea is to split the complex system into components according to the functions, forming component module in small scale with unitary function[9]. Component is consisted of interface and implementation body. The implementation body can be accomplished by different technologies, and the interface defined by platform-independent language is the only way for the component to interaction with the outside[10]. In the study of virtual plants, different species and different varieties of the same species may have totally different software designs. Component-based mode can unify the design mode of virtual plant software, making it consistent in structure, so it becomes a necessary condition to digitally construct crop plant type[11].

From morphologically, the crop is composed of roots, stems, leaves, fruits and other organs. Aming at corresponded organs and their growth position, morphological component modules can be built, and the components are connected according to the dependencies on the growth of crop organs, as shown in figure 3.

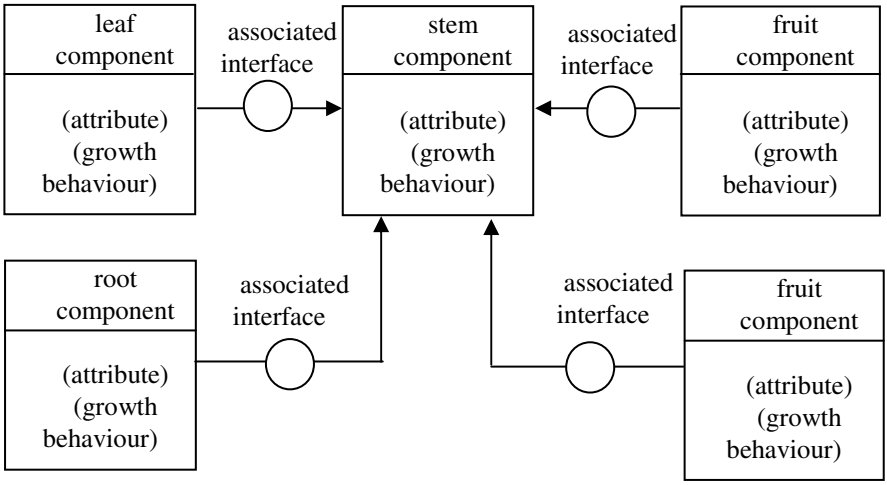


Fig. 3. Component-based organ module division and connection

The internal structures of the component modules for various organs are similar, including basic attribute structure of the organ morphology, growth behaviour (i.e., dynamic process of the growth) and associate interfaces with the other organ components or the system, here, take leaf for example, its internal structure is as shown in table 1.

Table 1. Inner Structure of Leaf Organ Component

Attribute	Behaviour	Interface
Initial growth time	Contour growth (length array, width array, time)	Location interface of leaf growth
Growth duration		Environment component interface
End of growth time	Growth duration = End of growth time − Initial growth time	Visualization component interface
Growth cycle		Expert system component
Leaf length array	Iteration number = Growth duration / Growth cycle	User interaction interface
Leaf width array		……
Current vein string	Current vein string = Vein initial string array + Iterated string array×Iteration number	
Vein initial string array		
Iterated string array		
Iteration number	……	

4 Digital Construction of Crop Plant Type

Through using Fourier equation to analyze crop canopy, achieving function description of canopy leaves quantities; Through using extinction equation to simulate light transmittance distribution in canopy, the optimal values of the parameters (such as leaf area index) was obtained and the mathematics model of the production biomass was established. And then, building the crop biomass allocation model according to the measured data of crop biomass and the organ sink strength in experiment, by analysis of the impact of environmental factors on crop morphology, crop functional-structural models of various organs were built, based on the theory of growth equation and combined with a controlled modeling method. On the basis of the above studies, crop ideotype was digitally constructed by components of the various crop organs, shown as figure 4.

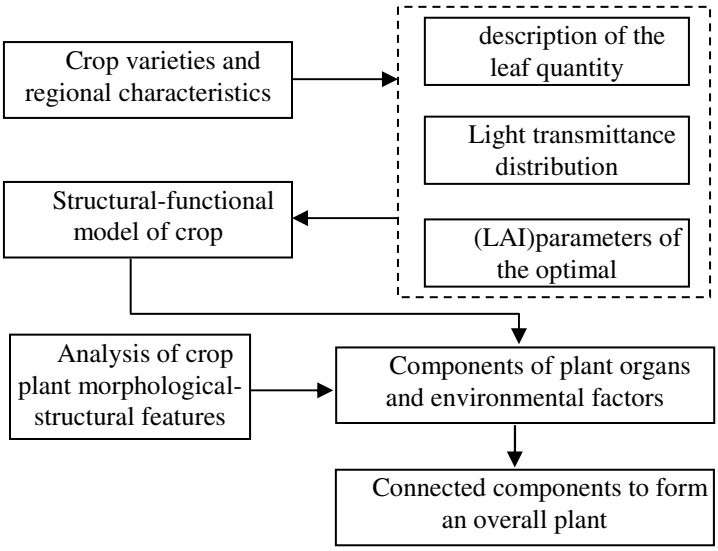


Fig. 4. Digital construction of plant type

5 Discussion

Combined with virtual crop theory and computer software development, digital design method is proposed on crop plant type, trying to achieve quantitative design of plant type to solve problems such as traditional plant type breeding inaccurate. This method is part of the process in the theoretical design stage, its effectiveness depends on the further improvement of the crop structural-functional model and organ components of the degree of precision, future research needs to be carried on to find problems and perfect the methods in the specific environment.

References

1. Yuan, L.P.: Progress of New Plant Type Breeding. Hybrid Rice 26, 72–74 (2011)
2. Sarlikioti, V., de Visser, P.H.B.: How plant architecture affects light absorption and photosynthesis in tomato:towards an ideotype for plant architecture using a functional-structural plant model. Annals of Botany 108, 1065–1073 (2011)
3. Fan, Z.Q., Si, C.J., Cao, H.W., Han, X., Li, W.H.: Virtual Plant Modeling Based on Mutual Feedback of Function-structure. Agricultural Science & Technology 12, 1972–1974 (2011)
4. Paul, C.D., Ame, L., Franc, O.H., et al.: Computing competition for light in the GREENLAB model of plant growth: A contribution to the study of the effects of density on resource acquisition and architectural development. Annals of Botany 101, 1207–1219 (2008)
5. Dong, Q.X., Wang, Y.M., Barczi, J.F., Hou, J.L.: Tomato structural-functional mode II:organ-based functional model and validation. Chinese Journal of Eco-Agriculture 15, 122–126 (2007)
6. Yang, J.A.: Calculation formula of photosynthetic yield. Chinese Journal of Agrometeorology 2, 16–20 (1981)
7. Zhang, Z.G.: Research on determination problem of plant growth's functional-structural model GreenLab. Institute of Automation of Academia Sinica, Beijing (2003)
8. Song, Y.H., Guo, Y., Li, B.G., de Reffy, P.: Virtual maize model I.biomass partitioning based on plant topological structure. Acta Ecologica Sinica 23, 2333–2341 (2003)
9. Lv, M.Q., Xue, J.Y., Hu, M.Q.: Reusable componentmodel based on software architecture. Application Research of Computer 25, 120–122, 128 (2008)
10. Chen, B., Li, Z.J., Chen, H.W.: Research on Component Models:A Survey. Computer Engineering & Science 30, 105–109 (2008)
11. Fan, Z.Q., Si, C.J., Han, X., Yang, Q.L.: Design Mode for Component-based Virtual Plant Software. Agricultural Science & Technology 13, 901–903 (2012)

Edge Geometric Measurement Based Principal Component Analysis in Strawberry Leaf Images

Jianlun Wang[1,*], Yu Han[*], Zetian Fu, Daoliang Li, Jianshu Chen, and Shuting Wang

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, P.R. China
wangjianlun@cau.edu.cn

Abstract. Edge geometric measurement analysis is an important method of image understanding and portraying the target feature. In this paper, we compress 17 interrelated shape descriptors which are based on edge geometric measure into 6 independent components, and discuss their meanings by using principal component analysis. The analyses in this article provide guidance for the shape feature optimization and accurate identification for greenhouse strawberry leaves images successfully.

Keywords: Greenhouse strawberry leaves, Shape feature description, Principal Component Analysis, Edge geometric measure.

1 Introductions

The performance comparison of several common description of shapes which are used in identify the leaves of strawberry grown in the greenhouse。 Strawberries have a short growth cycle, less disease, easy to manage, high nutritional value, economic value advantages, widely grown in the greenhouse [1]。 With the popularity of hardware and image processing technology matures, many field of agricultural engineering applications[6] have used a variety of techniques based on computer vision, such as automatic classification of the fruits and vegetables[2], quality traceability[3], the robot picking[4], growth status monitoring, and the early warning of pest and disease[5]. Among them, the target recognition are facing with many problems of different varieties of agricultural crops, complexity of the imaging background, range of issues such as characterization as a vital process in image processing, so it is difficult to form a high recognition rate of the general algorithm. Therefore, this paper according to greenhouse strawberry leaves, summarized 17 ways of shape description. On the high dimensional feature dimension reduction is compressed into 6 comprehensive index by using the principal component analysis , and has carried on the detailed explanation to provide guidance in getting more effective characteristics selection and comprehensive evaluation in practical application.

[1] Corresponding author.
[*] These authors contributed equally to the article.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 58–68, 2013.

2 The Shape Descriptors in Target Region

The main difficulties in shape quantitative description is the lack precise and uniform definition. ASM think that shape parameter which are used to describe the micro-structure have some common in no dimension, quantitative description and sensitive to shape change[7]. This paper will divide geometry characteristic which are commonly used into the following several forms：Geometric description of the target boundary（Including the ratio between perimeter and area, fineness ratio, eccentricity, etc）、Measure of the target area and its surround polygon（Including Surrounded by rectangular compact, surrounded by polygonal compact degree, etc）、Geometric description of the target area（Including Centroid offset rate, Area symmetry, Circumference symmetry, etc)、Invariant moments（Hu Invariant moments)。We will give explanation about it one by one below:（Among them, A represents the area, represents the circumference, L represents the length of bounding rectangle, W represents the width of bounding rectangle, Aconvex is reprensents target convex hull area）：

（1） Perimeter area ratio：*P/A*, Range in (0,1)。

（2） Fineness ratio：*4*Pi*A/(P*P)*, Range in (0,1)。Get the maximum values(1) while contiguous area is round。

（3） Exterior ratio：*W/L*, Used to describe the shape of the target after plastic deformation, Take the maximum value(1) of a square and circular target and it will be reduced with slender degree about slender target。

（4） Eccentricity : The eccentricity of target which is equivalent to an ellipse(E), and the equivalent ellipse is defined as a voal which has the same centroid and secondary center moment with target。

（5） Skeleton eccentricity：Eccentricity of skeleton equivalent ellipse (E), and skeleton equivalent ellipse is defined as a fitting ellipse which has the same skeleton with target。Generally, In order to avoid the influence of finely branch in skeleton, the first to do is deburring skeleton and then fitting ellipse。

（6） Compactness of bounding rectangle：*A/(L*W)*, Reflects the bump on the bounding rectangle of the target area.

（7） Compactness of bounding polygon：*A/Aconvex*, Reflects the bump on the minimum convex hull of the target area。

（8） Centroid offset rate：The ratio of the distance between centroid to boundary in the upper-left corner and lower right corner to the upper left corner, it reflects the relationship of the centroid position in the target area。

（9） Area symmetrical degrees：*min(Aleft, Aright)/max(Aleft, Aright)*, Aleft and Aright are represents the area size of left and right side of symmetry axis , regional symmetry in the area. It's the performance of regional symmetry in the area.If the region is completely symmetrical ,the maximum value will be 1.

（10） Perimeter symmetric degrees：*min(Pleft,Pright)/max(Pleft,Pright)*,Pleft and Prightare represents the perimeter size of left and right side of symmetry axis , regional symmetry in the area. It's the performance of regional symmetry in the perimeter.If the region is completely symmetrical, the maximum value will be 1.

(11) Hu Invariant moments：Ming-Kuei Hu apply algebraic invariant moment to pattern recognition and regard seven moment features about rotation and scaling translation invariant in the target area as the target shape characteristics。

3 Statistical Analysis about the Descriptor of Shape

Statistical frequency of the above shape descriptor as follows to show:

Table 1. Statistical description of shape feature

Shape descriptors	Mean	Standard Error	Standard deviation	Coefficient of variation
Perimeter area ratio	0.009293	0.000254	0.002365	0.25451
Fineness ratio	0.654973	0.005491	0.051218	0.078199
Appearance ratio	0.755404	0.010079	0.094009	0.124449
Eccentricity	0.637541	0.012837	0.119733	0.187805
Skeleton eccentricity	0.664242	0.01463	0.136462	0.20544
Bounding rectangle compactness	0.774389	0.000705	0.006572	0.008487
Bounding polygon compactness	0.957814	0.001167	0.010888	0.011367
Centroid offset rate	0.521226	0.006561	0.061194	0.117403
Area symmetric rate	0.963479	0.011393	0.106268	0.110296
Perimeter symmetric rate	0.961412	0.002981	0.027803	0.028919
Hu1 Moment	0.206821	0.001691	0.015774	0.076269
Hu2 Moment	0.003439	0.000557	0.005196	1.511056
Hu3 Moment	9.22E-06	3.03E-06	2.83E-05	3.069955
Hu4 Moment	8.95E-06	3.02E-06	2.82E-05	3.147928
Hu5 Moment	8.67E-10	5.99E-10	5.59E-09	6.445866
Hu6 Moment	5.25E-07	2.61E-07	2.43E-06	4.631451
Hu7 Moment	6.24E-09	3.06E-09	2.86E-08	4.579208

4 Principal Component Analysis of the Shape Factor

Principal component analysis is a statistical method which is first proposed by Pearson in 1901, and then developed by Hotelling (1933). Its main purpose is that to reduce the number of variables, and make change for the formation of a linear combination of a few independent variables (principal components), but by the difference of the linear combination of ingredients into the largest, makingthe original multi-dimensional characteristics of these principal components show the greatest individual differences. In short, Principal component analysis is a number of variables into a few principal components (i.e. integrated variable) dimension reduction statistical method.

The general steps are as follows:

(1) The Standardization of Original Data

First of all to the standardization of original data, each index data of the sample is mean 0, variance 1.

$$\chi_{ij} = \left(\chi_{ij} - \overline{\chi_j}\right) \Big/ \sqrt{\sigma_j} \left(i = 1,2,\cdots n,\ j = 1,2,\cdots 17\right)$$

χ_{ij} as the i sample of the j shape descriptor value, $\overline{\chi_j}$ as mean value of the j shape descriptor value, σ_j as variance of the j shape descriptor value. In the principal component analysis of SPSS statistical software, standardization is automatically executed.

(2) Calculate the Indicators of the Correlation Matrix R

$$r_{ij} = \frac{\sum_{k=1}^{n} (x_{ki} - \overline{x}_i)(x_{kj} - \overline{x}_j)}{\sqrt{\sum_{k=1}^{n} (x_{ki} - \overline{x}_i)^2 \sum_{k=1}^{n} (x_{kj} - \overline{x}_j)^2}} \left(i = 1,2,\cdots,n,\ j = 1,2,\cdots 17\right)$$

r_{ij} as matrix of I rows and j columns representing related elements, namely the index I and index J correlation coefficient.

(3) Seek the Correlation Matrix R Eigenvalue and Characteristic Vector

The characteristic equation is expressed as $|\lambda I - R| = 0$, Commonly used Jacobi method to calculate the eigenvalues and the corresponding eigenvectors, and the size of the eigenvalues in descending order. The correlation matrix are given below:

Table 2. The correlation matrix of shape characteristics about SPSS

	P/Area	Fineness ratio	Appearance ratio	Eccentricity	Skeleton eccentricity	Bounding rectangle compactness	Bounding polygon compactness	Centroid offset rate	Area symmetric rate	Perimeter symmetric rate	Hu1 invariant moment	Hu2 invariant moment	Hu3 invariant moment	Hu4 invariant moment	Hu5 invariant moment	Hu6 invariant moment	Hu7 invariant moment
P/Area	1.000	-0.390	-0.493	0.484	0.394	-0.457	-0.603	0.047	-0.246	-0.087	0.453	0.297	0.191	0.197	0.121	0.156	0.162
Fineness ratio	-0.390	1.000	0.586	-0.558	-0.339	0.377	0.685	0.084	0.095	0.347	-0.357	-0.212	-0.296	-0.294	-0.188	-0.255	-0.243
Appearance ratio	-0.493	0.586	1.000	-0.983	-0.643	0.315	0.404	0.109	0.114	0.041	-0.540	-0.394	-0.306	-0.310	-0.238	-0.291	-0.280
Eccentricity	0.484	-0.558	-0.983	1.000	0.591	-0.309	-0.400	-0.130	-0.113	-0.045	0.524	0.326	0.261	0.264	0.197	0.245	0.235
Skeleton eccentricity	0.394	-0.339	-0.643	0.591	1.000	-0.240	-0.388	-0.012	-0.154	-0.104	0.479	0.424	0.247	0.253	0.211	0.261	0.237
Bounding rectangle compactness	-0.457	0.377	0.315	-0.309	-0.240	1.000	0.536	-0.282	0.540	0.157	-0.547	-0.368	-0.604	-0.606	-0.461	-0.532	-0.540
Bounding polygon compactness	-0.603	0.685	0.404	-0.400	-0.388	0.536	1.000	-0.022	0.320	0.338	-0.438	-0.216	-0.226	-0.222	-0.106	-0.161	-0.168
Centroid offset rate	0.047	0.084	0.109	-0.130	-0.012	-0.282	-0.022	1.000	0.057	-0.011	0.084	-0.023	0.294	0.295	0.275	0.283	0.289
Area symmetric rate	-0.246	0.095	0.114	-0.113	-0.154	0.540	0.320	0.057	1.000	-0.016	-0.222	-0.282	-0.096	-0.097	-0.071	-0.096	-0.084
Perimeter symmetric rate	-0.087	0.347	0.041	-0.045	-0.104	0.157	0.338	-0.011	-0.016	1.000	-0.040	0.085	-0.291	-0.289	-0.267	-0.272	-0.282
Hu1 invariant moment	0.453	-0.357	-0.540	0.524	0.479	-0.547	-0.438	0.084	-0.222	-0.040	1.000	0.779	0.331	0.337	0.274	0.329	0.317
Hu2 invariant moment	0.297	-0.212	-0.349	0.326	0.424	-0.368	-0.216	-0.023	-0.282	0.085	0.779	1.000	0.074	0.078	0.110	0.160	0.100
Hu3 invariant moment	0.191	-0.296	-0.306	0.261	0.247	-0.604	-0.226	0.294	-0.096	-0.291	0.331	0.074	1.000	0.999	0.926	0.960	0.979
Hu4 invariant moment	0.197	-0.294	-0.310	0.264	0.253	-0.606	-0.222	0.295	-0.097	-0.289	0.337	0.078	0.999	1.000	0.927	0.961	0.980
Hu5 invariant moment	0.121	-0.188	-0.238	0.197	0.211	-0.461	-0.106	0.275	-0.071	-0.267	0.274	0.110	0.926	0.927	1.000	0.985	0.981
Hu6 invariant moment	0.156	-0.255	-0.291	0.245	0.261	-0.532	-0.161	0.283	-0.096	-0.272	0.329	0.160	0.960	0.961	0.985	1.000	0.989
Hu7 invariant moment	0.162	-0.243	-0.280	0.235	0.237	-0.540	-0.168	0.289	-0.084	-0.282	0.317	0.100	0.979	0.980	0.981	0.989	1.000

(4) Calculate the Variance Contribution Rate and Cumulative Variance Contribution Rate

Contribution rate:

$$\frac{\lambda_i}{\sum_{k=1}^{p}\lambda_k}(i=1,2,\cdots,p)$$

The cumulative variance contribution rate:

$$\frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{p}\lambda_k}(i=1,2,\cdots,p)$$

Generally it takes the cumulative contribution rate of 85% ~ 95% m eigenvalues corresponding to the former m variables as main components, number of principal components for M.

Variance and the cumulative variance contribution rate as shown in the following table:

Table 3. Variance contribution rate and cumulative variance contribution rate of shape features about SPSS

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	%of Variance	Cumulative %	Total	%of Variance	Cumulative %
1	6.783	39.901	39.901	6.783	39.901	39.901
2	3.334	19.614	59.515	3.334	19.614	59.515
3	1.504	8.847	68.362	1.504	8.847	68.362
4	1.405	8.264	76.626	1.405	8.264	76.626
5	0.933	5.489	82.115	0.933	5.489	82.115
6	0.848	4.985	87.100	0.848	4.985	87.100
7	0.608	3.578	90.679	0.608	3.578	90.679
8	0.535	3.145	93.824	0.535	3.145	93.824
9	0.409	2.408	96.232	0.409	2.408	96.232
10	0.263	1.545	97.777	0.263	1.545	97.777
11	0.191	1.123	98.901	0.191	1.123	98.901
12	0.114	0.674	99.574	0.114	0.674	99.574
13	0.052	0.308	99.882	0.052	0.308	99.882
14	0.011	0.065	99.947	0.011	0.065	99.947
15	0.007	0.042	99.989	0.007	0.042	99.989
16	0.001	0.007	99.996	0.001	0.007	99.996
17	0.001	0.004	100.00			

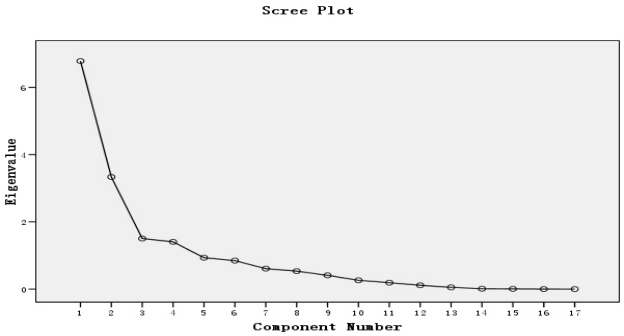


Fig. 1. The SPSS components eigenvalue scatter diagram

(5) Determine the Principal Components

From cumulative variance contribution ratio of table 3 and each component eigenvalue scatter diagram of figure 1 show that 1-6 cumulative variance contribution rate has reached 87.1 percent, After the sixth component feature value is going to be more smaller, and the transform slow. Therefore, here take first 6 comprehensive variables as main component.

(6) Determine the Principal Component Loading Matrix L

$$l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij} (i, j = 1,2,\cdots, p)$$

In this paper data loading matrix results are as follows:

Table 4. Principal component loading matrix L about SPSS

Component Matrix(a)						
Shape descriptor	Component					
	F1	F2	F3	F4	F5	F6
Perimeter area ratio	0.502	-0.492	0.070	-0.208	0.225	-0.243
Fineness ratio	-0.560	0.410	0.387	0.314	-0.026	-0.061
Appearance ratio	-0.633	0.550	0.320	-0.314	0.003	0.208
Eccentricity	0.594	-0.568	-0.322	0.295	-0.011	-0.221
Skeleton eccentricity	0.536	-0.449	-0.113	0.272	0.055	0.071
Bounding rectangle compactness	-0.745	0.059	-0.430	0.302	0.018	0.125
Bounding polygon compactness	-0.536	0.476	0.055	0.564	-0.120	0.004
Centroid offset rate	0.223	0.357	0.318	-0.070	0.772	-0.103
Area symmetric rate	-0.274	0.243	-0.569	0.334	0.481	0.257
Perimeter symmetric rate	-0.314	-0.121	0.406	0.548	0.033	-0.499
Hu1 Moment	0.643	-0.430	0.328	0.215	0.102	0.338
Hu2 Moment	0.395	-0.464	0.487	0.308	-0.041	0.475
Hu3 Moment	0.860	0.479	-0.048	0.024	-0.047	-0.049
Hu4 Moment	0.863	0.476	-0.045	0.030	-0.045	-0.050
Hu5 Moment	0.793	0.545	-0.035	0.123	-0.093	0.019
Hu6 Moment	0.846	0.495	-0.021	0.110	-0.080	0.022
Hu7 Moment	0.843	0.518	-0.039	0.082	-0.069	-0.011

(7) Calculate the Feature Vectors to Determine the Linear Expression of the Principal Component Index of the Original Data

Table 5. Characteristic vector of correlation matrix about SPSS

	Eigenvectors					
	F1	F2	F3	F4	F5	F6
P/Area	0.19	-0.27	0.06	-0.18	0.23	-0.26
Fineness ratio	-0.22	0.22	0.32	0.26	-0.03	-0.07
Min/MaxLen	-0.24	0.30	0.26	-0.26	0.00	0.23
Eccentricity	0.23	-0.31	-0.26	0.25	-0.01	-0.24
Skeleton eccentricity	0.21	-0.25	-0.09	0.23	0.06	0.08
Bounding rectangle compactness	-0.29	0.03	-0.35	0.25	0.02	0.14
Bounding polygon compactness	-0.21	0.26	0.04	0.48	-0.12	0.00
Centroid offset rate	0.09	0.20	0.26	-0.06	0.80	-0.11
Area symmetric rate	-0.11	0.13	-0.46	0.28	0.50	0.28
Perimeter symmetric rate	-0.12	-0.07	0.33	0.46	0.03	-0.54
Hu1 Moment	0.25	-0.24	0.27	0.18	0.11	0.37
Hu2 Moment	0.15	-0.25	0.40	0.26	-0.04	0.52
Hu3 Moment	0.33	0.26	-0.04	0.02	-0.05	-0.05
Hu4 Moment	0.33	0.26	-0.04	0.03	-0.05	-0.05
Hu5 Moment	0.30	0.30	-0.03	0.10	-0.10	0.02
Hu6 Moment	0.32	0.2 7	-0.02	0.09	-0.08	0.02
Hu7 Moment	0.32	0.28	-0.03	0.07	-0.07	-0.01

(8) $F_{1*6}=X_{1*17}* E_{17*6}$

E_{17*6} as eigenvector matrix in table 5, F_{1*6} as the principal component vector, X_{1*17} as the original shape descriptor vector

5 Experiments and Results

The first principal component eigenvalue is 6.783, the variance contribution rate is 39.9%, the entire data of standard variant is 39.9%, only relying on the first principal component can not reflect most information of original data. The first principal component in geometric invariant feature has high positive loads, in Appearance ratio and bounding rectangle compactness has high negative loads, the centroid offset rate has low positive loads, in the area and perimeter symmetric rate has a low negative loads, in other variables have similar secondary loads. That is to say, large (small) to the first principal components tend to be large (small) geometric moment invariants, then appearance ratio and bounding rectangle compactness tend to have smaller values (large). As a result of invariant moments are based on the regional global statistical

feature, with similar but different sets of images (like the fruit with similar to leaves) has better identification of. Therefore, the first principal component can be used to characterize the morphology of the more slender, not completely symmetrical, with concave boundary, and the global features of difference image sets with strong discernment.

The second principal component feature value is 3.334, the variance contribution rate of 19.614%, the entire data of standard variation is 19.614%.Through the observation to load factor matrix in table 4, we find that the second principal component in most variables are approximately equal to the load, so can be used as a comprehensive metric variables, reflect almost all the features of the original nature.In addition, because the main components in the bounding rectangle of compactness and area symmetrical rate have a low normal load and the perimeter symmetric rate with a low negative load. Therefore, while the second principal component analysis in comprehensive measure all the features, it slightly weak impact of foliage bending on region side.

The third principal component feature value is 1.504, showing that the entire data of standard variation is 8.847% separately. The third principal component in the area of symmetric rate ratio has high loads, in circumference symmetrical ratio, fineness rHu1 and Hu2 moments have high positive load.To the leaf image, if the edge has some curl, it may not affect the perimet circumference, but it changed the blade on the imaging surface of two-dimensional area, causing the image has high perimeter symmetric rate, and low rate of area of symmetry. Therefore, the higher of the third principal component values tend to express has an elongated oval blades, and the blade edge may appear in curling.

The fourth principal component feature value is 1.405, showing the entire data of standard variation is 8.264% separately. The fourth principal component in encircle polygon compactness and perimeter symmetric rate have high positive loads, in appearance ratio and the circumference area ratio have high negative loads, in the centroid offset ratio with low negative loads,at the moment invariant part has lower positive load. High bounding polygon compactness and perimeter symmetric ratio both show that the region is filling full, is more symmetrical convex polygon, and the appearance ratio and perimeter area ratio with higher load, the principal component has higher differentiate different width of the ability to target.To a certain extent, the fourth principal components can be associated with the first principal component is complementary, it weakens the global moment invariant statistical characteristics influence, but strengthen the shape complexity and symmetry as well as the appearance of the elongated convex characteristics.

The fifth principal component feature value is 0.933, showing the entire data of standard variation is 5.489% separately. On the center of mass migration rate of fifth principal components have significantly high positive loads the area of symmetric rate has high positive load, in addition, the other variables with have smaller load value.In general, low centroid rate figure centroid bias top right (according to the centroid offset rate calculation formula can be seen), high centroid rate figure centroid bias bottom left. Value of the larger of the fifth principal components tend to say mass at the bottom left of the symmetrical region.

The sixth principal component feature value is 0.848, showing the entire data of standard variation is 4.985% separately. In the perimeter symmetric rates on sixth principal components have higher negative loads, and in Hu1 and Hu2 with higher positive load, in other variables are not high load. Because Hu1 and Hu2 are calculated by using image moment of order two , and the image of two order moment is not included details of the general information,in the geometry mean variance, if we consider only the order for 2 sets of the moment,the original image is completely equivalent to the image centroid as the center and has the same two order moments of the ellipse. Therefore, the higher of the sixth principal component values tend to express elliptical rather than circular, may have side edges curled area,but the direction and size of the show is weak.

From the above analysis we can see that the first six principal components of the cumulative variance contribution ratio reach to 87.1%, it can be very good summary of the data group. And the six main component covers the area of concave and convex of elongated,the complexity symmetry of shape ,the centroid position and direction, as well as statistics based on geometric moment invariants and other aspects of information, a reasonable description of the region shape characteristic, the original 17 features integrated into 6, while retaining the original variables most of the information under the premise, realize dimension compression of the data.

6 Conclusion

This article from the 17 dimensions of the quantitative describe the single plane shape characteristics, including the target boundary geometric description (perimeter area ratio, fineness ratio, eccentricity and so on),the target area and its surrounding polygon metric (bounding rectangle compactness, encircle polygon compact degree and so on), the target area (centroid offset ratio, area of symmetry, perimeter symmetry degree and so on) invariant moment (Hu), covering graphics elongated, convexity, complexity, symmetry, centroid direction, ellipse, compact plumpness, number of shape, geometric invariant moment etc.. Then through by the method of principal component analysis to do variable compression on these characteristics, explaining the he first six principal components which cumulative variance contribution rate reach to 87.1%, in which elongated, convexity and geometric invariant moment contributed to the first principal component is larger; the second principal component can be used as a comprehensive measure factor, each characteristics of second main components in the contribution rate is almost equal; the third principal component more reflect the region of the ellipse and elongated; the fourth principal component in global invariant moment statistical properties of the performance is not outstanding, but focused on the expression of slim and compact plumpness; the higher of the fifth principal component values tend to areas near left lower direction of the center of mass, which leaves the image on the right is fine, but left is wide, image centroid shift; sixth main components on the ellipse has stronger descriptive power, but the direction and size of the description of is a little weak.

References

[1] Sun, S., Chen, H.: Strawberry cultivation technology in greenhourse. Rural Science and Technology (8), 74–75 (2010)
[2] Liu, Y., et al.: Fruit shape classification based on wavelet descriptors. Journal of Zhejiang University (Agriculture and Life Science Edition) (3), 322–328 (2010)
[3] Zhang, G.: Forest products traceability system research. The Modern Agricultural Science and Technology (22), 224–228 (2011)
[4] Cui, H.: Image processing technology in the application of agricultural robot. Journal of Agricultural Mechanization Research (1), 168–170 (2008)
[5] Ying, X.: Agricultural pest image remote transmission system research and design. Chinese Academy of Agricultural Sciences (2006)
[6] Jiang, S., Sun, H.: The computer image technology in Agricultural Engineering. Journal of Agricultural Mechanization Research (11), 177–178 (2006)
[7] Zhang, S.: Image engineering (Book) - image analysis, 2nd edn., pp. 311–342. Tsinghua University Press (2005)

The Research of the Strawberry Disease Identification Based on Image Processing and Pattern Recognition

Changqi Ouyang*, Daoliang Li, Jianlun Wang, Shuting Wang, and Yu Han

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, P.R. China
ouyangchangqi@126.com

Abstract. In this paper, a synthesis segmentation algorithm is designed for the real-time online diseased strawberry images in greenhouse. First, preprocess images to eliminate the impact of uneven illumination through the "top-hat" transform, and remove noise interference by median filtering. After comprehensively applying the methods of gray morphology, logical operation, OTSU and mean shift segmentation, we can obtain the complete strawberry fruit area of the image. Normalize the extracted eigenvalues, and use eigenvectors of part of the samples for training the BP neural network and support vector machine, the remaining samples were tested in two kinds of disease strawberry recognition model. Results show that support vector machines have a higher recognition rate than the BP neural network.

Keywords: strawberry fruit, gray morphology, OTSU, mean shift segmentation, Pattern recognition.

1 Introduction

Strawberry fruit image segmentation is used widely in strawberry picking and fruit grading.In strawberry picking , the strawberry fruit is mature normal strawberry, it's easy to segment the fruit image because the color of the fruit area is almost the same. In strawberry fruit grading , the strawberry fruit has already been picked, and it's not difficult for image segmentation because the background is relatively consistent. The strawberry image segmentation algorithm discussed in this paper aims to offer the objects to the real-time online identification of the three common strawberry diseases. Therefore, this algorithm is for the image segmentation of the complete strawberry fruit without picking it. As the color distribution of diseased strawberry fruit is inconsistent, and fruit background is quite complicated, it needs to combine kinds of image processing methods to obtain the complete segmentation of the strawberry fruit according to the characteristics of the diseased fruit area.

Pattern recognition is an intelligent activity, including analysis and judgment. The analysis process is to determine the division of pattern class characteristics and

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 69–77, 2013.

expression method; judgment is reflected in the characteristics of the unknown object, and its judgment belongs to a certain class. It aims to use a computer to achieve the ability to identify the class, which is the contrast of two different levels of identification ability. Pattern recognition system includes the sample acquisition, preprocessing, feature extraction, classification decisions and classifier design.

2 Material and Equipment

The objects studied in this paper are the strawberries of the three common diseases in greenhouse,as shown in Fig 1.These three diseases are powdery mildew, shrinkage and uneven ripening. The characteristic of powdery mildew is that the fruit surface is coarse and with a layer of white powder. The characteristic of shinkage is that the fruit atrophies and is deformed. The characteristic of uneven ripening is that the fruit's color is uneven, but its surface is smooth.

(a) powdery mildew (b) shrinkage (c) uneven ripening

Fig. 1. Three common diseases of strawberry

3 Image Preprocessing

As a result of noise, uneven illumination, and the transform of analog and digital signals of the CCD camera, the online images captured in the greenhouse under natural light are blurred in details ,such as its contours. Therefore, before image segmentation, the image preprocessing must be taken to eliminate the impact.

3.1 Image Denoising

To ensure that the edge of the strawberry fruit is not vague while image denoising, the median filtering is applied to remove image noise, in detail, each channel of the RGB image is removed noise through median filtering respectively, and then the denoised image can be obtained through the integration of the three denoised channels.

3.2 Uneven Illumination Effect Elimination

Strawberry images captured under natural light are greatly affected by uneven illumination. If this impact is not eliminated, the object may be segmented into disperse areas. In this paper, the "top-hat" transform is applied to eliminate the impact.

4 Image Segmentation

4.1 The Extraction and Transform of the Red and Green Area Templates

Figure 2 (a) shows the preprocessed strawberry image. The red area template can be got by 2R-G-B, as shown in Figure 2 (b), where R, G and B were the three channels of the color image in RGB space. Due to the uneven surface of the strawberry fruit, the fruit area contains many holes, which needs to be filled in order to get the closed template. Then, execute the OTSU operation, the initial fruit template can be obtained as shown in Figure 2 (c).

Conduct the logical AND operation between the template complementation and Figure 2 (a) to remove the fruit area in Figure 2 (a) while keeping the image outside of the template unchanged. The result is shown in Figure 3 (a). The green area template can be obtained by 2G-R-B, as shown in Figure 3 (b). Fig 3 (a) can be removed the green zone in similar way, and backfilled with the red template, then the result can be got as shown in Fig 3 (c). Now, the green area around the fruit is all black, and the fruit region is almost all white. The purpose of this series of operations is to separate strawberry fruit region from other objects, and create a favorable condition for the subsequent image segmentation.

(a)Preprocessed Image (b)Red region (c)Initial fruit template

Fig. 2. Initial fruit template

(a) Image erased part of the fruit (b) Green region template (c) Image backfilled the red area

Fig. 3. The separation between the fruit region and other objects

4.2 Strawberry Fruit Area Color Clustering

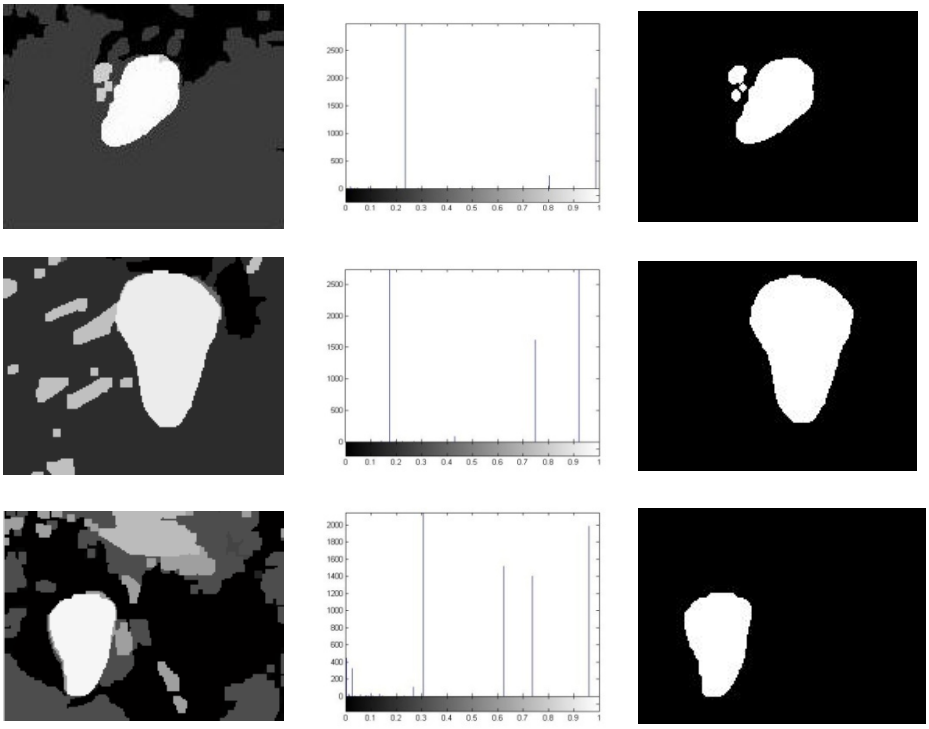
Because K-means cluster and fuzzy C-means cluster both need to set the cluster number initially, which is unknown for the image containing complicated background, it is infeasible to apply these two methods to segment the fruit image.

The mean shift algorithm is an adaptive probability density gradient ascent method of searching for the peak. It is as follows:1)Select the window size and initial position ;2)Calculate the gravity center of the window ;3)Move the center of the window to the center of gravity ;4)Repeat Step2 and Step 3 until the center of the window "convergence", that is, the moving distance of the window must be less than a certain threshold.Each pixel in the image can be set as the initial point, after performed the mean shift algorithm respectively, the initial points converged to the same point are considered as the same class. Therefore, the mean shift algorithm can be applied to image cluster segmentation.

Based on the above discussion, the mean shift segmentation algorithm is used in image segmentation, which make pixels of the strawberry fruit area cluster into the same class in color space, namely strawberry fruit regional pixels are all of the same value.

4.3 Target Area Extraction

After the operation of converting the image into grayscale and filling holses, execute the “open” operation on the image with circular structural element of a radius of 2 to remove the adhesion between the target and background. The experiments showed that: when the split threshold is 0.8, the grayscale threshold segmentation can obtain good results as shown in Figure 4; and then extract the largest connected area of the image to obtain strawberry fruit template. Combine this template and the preprocessed image by logical AND operation, then the image of the complete strawberry fruit area can be obtained.



(a) Gray image (b) Histogram of (a) (c) threshold segmentation image

Fig. 4. Threshold segmentation image

5 Strawberry Fruit Image Segmentation Algorithm

After the above discussion on the image preprocessing and image segmentation method, here are the strawberry fruit image segmentation algorithm steps, including image reading, image preprocessing, image segmentation （shown in Fig 5）.

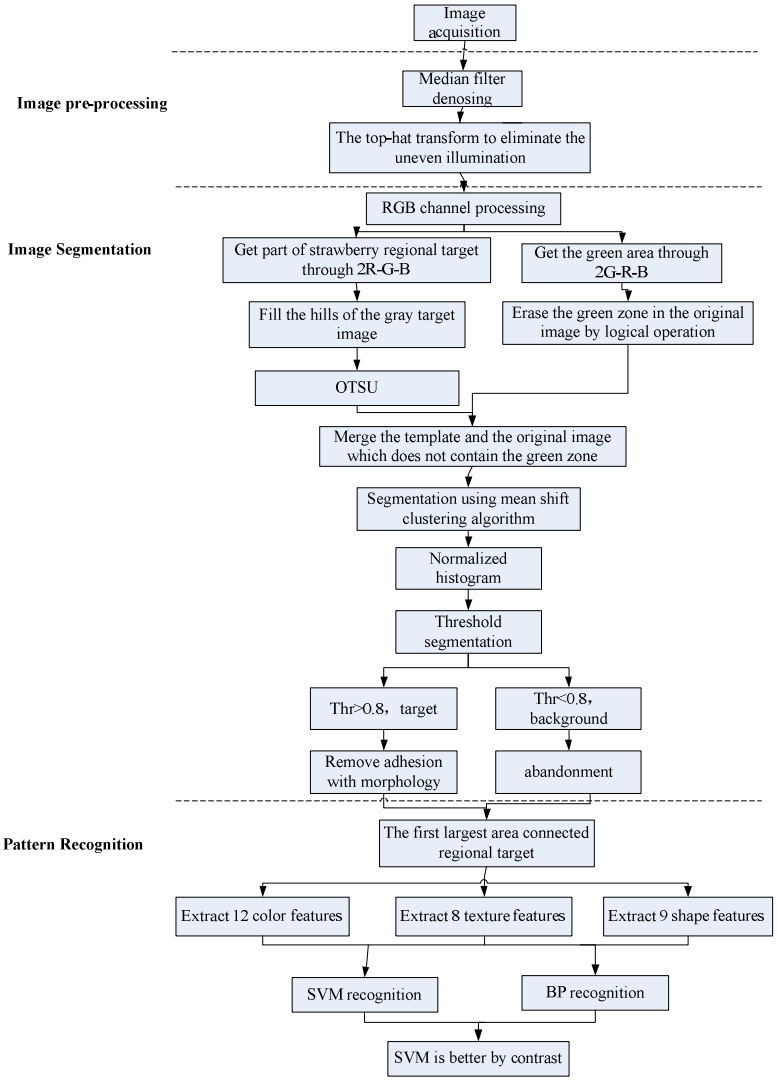


Fig. 5. The flow chart of the segmentation algorithm

6 Segmentation Results

Figure 6, Figure 7, Figure 8, and Figure 9 are the segmentation results of the above strawberry image segmentation algorithm, which corresponds to normal strawberry, powdery mildew of strawberry, shrinkage of fruit disease strawberries and uneven ripening disease strawberries respectively. Each subgraph corresponds to the result after the corresponding step of the above segmentation algorithm. The algorithm can get effective segmentation results under natural illumination.

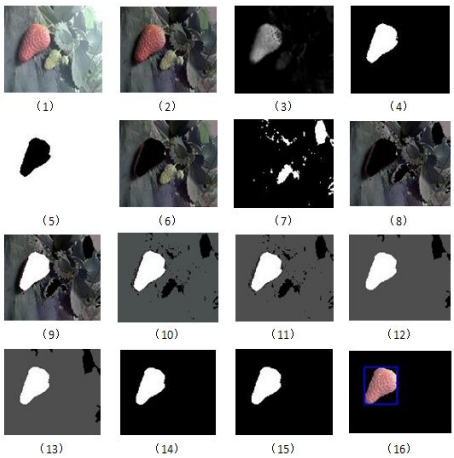


Fig. 6. Segmentation results of the normal strawberry

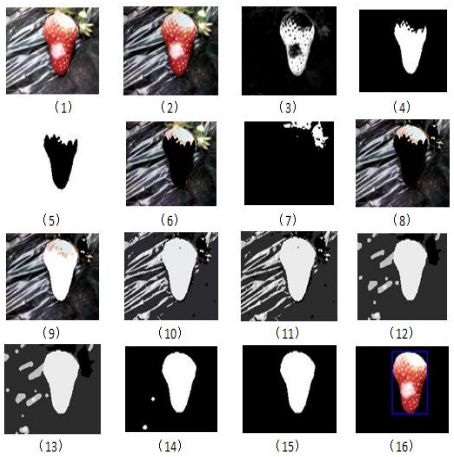


Fig. 7. Segmentation results of the powdery mildew strawberry fruit

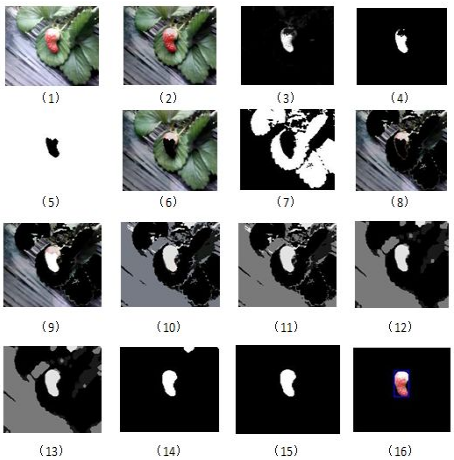


Fig. 8. Segmentation results of the strawberry fruit with the shrink disease

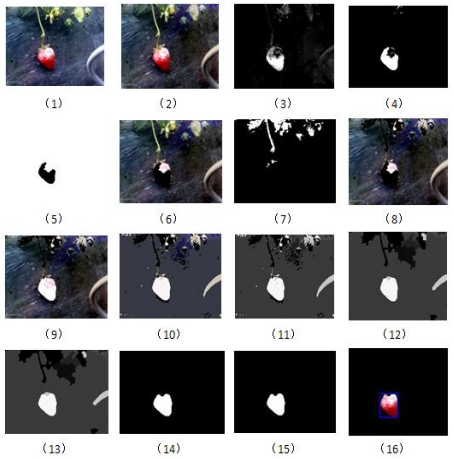


Fig. 9. Segmentation results of the strawberry fruit withthe uneven ripening disease

7 Experimental Schemes

7.1 Data Selection

a)Data access

The color, texture and shape feature were extracted from four kinds of strawberry respectively, including 12 kinds of color characteristic parameters,8 kinds of texture parameters, nine kinds of shape characteristic parameters, so each sample corresponds

to a set of characteristic data of a total of 12 +8 +9 = 29 kinds of characteristic parameters. The sample size of each type of disease strawberry and normal strawberry is 20. Set category labels are as follows: normal strawberry labels 1, powdery mildew strawberry labels 2,fruit shrink disease strawberry labels 3, and the uneven ripening disease strawberry labels 4.The two main equations are:

mean $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

standard deviation $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$

b)Division of the data

Establishing training and testing are two essential processes while using BP neural network and SVM to build the classifier. So data should be divided into two parts, one for training and one for the test. 3/4 of the data are used for training, 1/4 of the data are used for testing. Select 60 sets of data randomly as a training data set, the remaining 20 sets of data as a test data set from the data set.

c)Data preprocessing

The physical meaning of the color, texture and shape characteristic value is different, not comparable, and their range changes, so there is a big numerical difference among them. Data normalization is necessary, or it will lead to the classifier learning slows down or even does not converge. In this study, each dimensional feature data were normalized preprocessed, used in a normalized mapping shown as follows:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

In which x is the corresponding eigenvalue of a certain characteristic properties,x_{max}, x_{min}shows the largest and the smallest eigenvalue of the properties. While classifier training is completed, the input data for the test needs to be normalized in order to enter the classifier.

7.2 BP Neural Network Training and Prediction

According to Kolmogorov theorem, using a three-layer BP neural network(29 × 59 × 4)for the classifier, in which the hidden layer activation function is the non-symmetric Sigmoid function: $f(x) = \frac{1}{1+e^{-x}}$, the output layer activation function is a linear function f (x) =x, and set a target error e = 0.01. Use the training data to train the BP neural network, to make the BP neural network can predict the nonlinear function output.

Use trained BP neural network to predict the test data, and use the predicted output and actual output of the BP neural network to analyze the classification ability of BP neural network.

7.3 SVM Training and Prediction

Implementation of SVM in this experiment uses the libsvm toolbox, the kernel function uses the radial basis kernel function. The training data is used to train SVM classifiers to get the SVM classification model.

Use the trained SVM classifier to make the label prediction of test data, and analyze the Classification ability of SVM with predicted SVM output and actual output.

8 Conclusion

Image segmentation of the strawberry fruit of the three common diseases consists of two main parts: the first part is image pre-processing, including the application of median filtering to remove the noise from image acquisition, and the application of top-hat transform to eliminate the effects of uneven illumination; the second part is image segmentation, using gray morphology, OTSU algorithm, the mean shift segmentation algorithm to segment the fruit from the image. These experimental results indicate that image processing algorithm can obtain effective segmentation results to the online images of the three common kinds of diseased strawberries and nomal ones under natural illumination.

This paper uses digital image processing, pattern recognition technology to extract the integrated fruit of disease strawberries. Based on parameters of color characteristic, texture features and shape features, use BP neural network and support vector machine to make classification and recognition for disease strawberry. The experimental results show that support vector machine has a higher recognition rate than the BP neural network so that support vector machine is used as the classifier for disease strawberry.

References

1. Zhou, T., Zhang, T., Yang, L., et al.: Comparison of two algorithms based on mathematical morphology for segmentaion of touching strawberry fruits. Transactions of the CSAE 23(9), 164–168 (2007); (in Chinese with English abstract)
2. Xie, Z., Zhang, T.: A new method of segmentation of strawberry image. Journal of China Agricultural University 11(1), 84–86 (2006); (in Chinese with English abstract)
3. Zhang, T., Zhou, T.: Strawberry harvesting robot. Journal of China Agricultural University 9(4), 65–68 (2004)
4. Xie, Z., Zhang, T., Zhao, J.: Ripened Strawberry Recognition Based on Hough Transform. Transactions of the Chinese Sociey for Agricultural Machinery 38(3), 106–109 (2007)
5. Yu, S., Liu, R.: OpenCV course. Beijing University of Aeronautics and Astronautics Press, Beijing (2009)
6. Zhang, Y.: Image Processing. Tsinghua University Press, Beijing (2008)
7. Zhang, Y.: Image Analysis. Tsinghua University Press, Beijing (2008)
8. Liu, H., Shen, J., Guo, S., et al.: Digital Image Processing Using Visual C++. China Machine Press, Beijing (2010)

Selection of Leaf Orientation Insensitive Bands for Yellow Rust Detection

Lin Yuan[1,2], Jingcheng Zhang[1], Jinling Zhao[1], Shuhong Cai[3], and Jihua Wang[1,2,*]

[1] Beijing Research Center for Information Technology in Agriculture, Beijing, 100097, China
[2] Institute of Agriculture Remote Sensing and Information System Application, Zhejiang University, Hangzhou, 310029, China
[3] Hebei Agricultural Technology Extension Station, Shijiazhuang, 050000, China

Abstract. The disease detection by means of hyperspectral reflectance is influenced by the spectral differences between frontside (adaxial surface) and backside (abaxial surface) of a leaf inevitably. Taking yellow rust as an example, this study investigated the spectral differences between frontside and backside of healthy and diseased wheat leaves at grain filling stage using large size samples. We attempted to detect yellow rust with reflectance that was sensitive to the disease and insensitive to the orientation of leaves. The spectral difference between frontside and backside of leaves was analyzed by band ratioing and a pairwise t-test. The bands that were insensitive to the orientation of leaves were identified with a thresholding method. Then, with the aid of an independent t-test analysis, we recognized the bands that were sensitive to the disease. The overlapped bands were applied for developing models that quantifying disease severity by fisher linear discrimination analysis (FLDA). The results suggested that the bands within 606-697nm and 740-1000nm were suitable for disease detection yet insensitive to the orientation of leaves. Based on these bands, the model accuracies reached 71% for FLDA. These bands can be used as a basis for further selection of appropriate bands to detect yellow rust at canopy level.

Keywords: yellow rust, winter wheat, frontside and backside, fisher linear discrimination analysis (FLDA).

1 Introduction

Winter wheat (*Triticum aestivum* L.) is one of the most important crops in China. Yellow rust disease which is induced by *Puccinia striiformis* f. sp. *Tritici*, is one of the most destructive diseases that has a severe impact on both yield and grain quality of winter wheat [1-2]. Currently, scouting by human being is the widely used method for crop diseases inspection, which is not only time consuming and labor intensive, but also expensive and subjective. As the only way in obtaining the disease distribution information spatially, remote sensing is a promising alternative to traditional methods in disease monitoring [3]. A number of studies were conducted in understanding the

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 78–84, 2013.

spectral characteristics of yellow rust [4-7]. The wavelengths at 446-725 nm and 1380-1600 nm were identified as sensitive bands for disease detection [8]; Based on the spectral reflectance at 680 nm, 725 nm and 750 nm, Moshou et al. [9-10] successfully distinguished infected wheat leaves from healthy ones at an early infection stage. It should be noted that in most of the studies that were conducted at leaf level, the leaf spectra were measured only for the frontside of leaves. However, in fact, due to the curling of leaves, the frontside and backside of leaves are presenting in a mixed pattern in the nadir view at canopy level. The spectral difference between the frontside and backside of leaves may complicate the spectral response at canopy level and even lead to the failure of disease detection. Therefore, it is necessary to understand the spectral difference between frontside and backside of healthy and diseased leaves, and search for those bands that were sensitive to the disease yet insensitive to the orientation of leaves. However, the spectral difference of frontside and backside of leaves is rarely reported for healthy or diseased wheat plants. Our objectives were: (1) to analyze the spectral difference between frontside and backside of leaves at grain filling stage, and to identify bands that were insensitive to the orientation of leaves, (2) to identify those bands that were sensitive to the occurrence of disease, and (3) to develop model for disease detection based on those identified bands.

2 Materials and Methods

2.1 Study Area and Yellow Rust Inoculation

The experiment was conducted at Beijing Xiaotangshan Precision Agriculture Experimental Base, in Changping district, Beijing (40°10.6' N, 116°26.3' E) on the 2010-2011 growing season. The cultivar of winter wheat was 'Jingdong 9843', which was moderately susceptible to yellow rust. The soil at this site is a silt-clay loam. The average topsoil nutrient status (0-0.30 m depth) was as follows: organic matter 1.42-1.48%, total nitrogen 0.08-0.10%, alkali-hydrolysis nitrogen 58.6-68.0 mg kg^{-1}, available phosphorus 20.1-55.4 mg kg^{-1}, and rapidly available potassium 117.6-129.1 mg kg^{-1}. The experimental field received 200 kg ha-1 nitrogen and 450 m3 ha^{-1} water, which was a recommended rate for this cultivar. Spray method was used in inoculating yellow rust spores to wheat plants. Two different concentrations of spore solutions were applied to generate various infection levels.

2.2 Inspection of Disease Severity

The disease severity of each sample was determined by visual estimation of the cover percentage of pustules on the leaf [11-12]. To minimize the error induced by investigator, the diseased leaves were inspected by one investigator. In this study, the disease severity of leaves was determined on the basis of the proportion of infected region on the leaf (0-100%). Apart from the healthy leaves, all diseased leaves were

grouped into 3 severity classes: slight for proportion within 5-25%, moderate for proportion within 25-60%, and heavy for proportion over 60%. Since those leaves with an infected proportion less than 5% were actually difficult to be visually separated from healthy ones, they are classified as healthy leaves.

2.3 Leaf Sampling and Spectral Measurement

The leaf sampling and measurement were conducted on May 23 (grain filling stage) in the year of 2011. The spectra of frontside and backside were measured for each leaf. The leaves were cut from the plants in the fields with scissors. After that, the samples were immediately packed with ice bags and transported to a nearby indoor laboratory to be measured. A total of 91 leaf samples were measured, including 26 healthy leaves and 65 diseased leaves. All samples were separated for model calibration (60%) and validation (40%) randomly.

Leaf spectral measurements were made by a FieldSpec® UV/VNIR spectroradiometer (ASD Inc., Boulder, Colorado, USA) over the 350-2500 nm wavelengths, coupling with an ASD Leaf Clip. The spectrum of a white Spectralon reference panel (99% reflectance) was measured once for every 10 leaf measurements. Ten readings were recorded and then averaged to obtain a spectral measurement for each leaf.

2.4 Comparison between Frontside and Backside of Leaves

Apart from comparing the frontside and backside of leaves by their original spectra, the band ratioing and pairwise t-test were adopted for the spectral comparison. The band ratioing was used to emphasize the spectral difference between frontside and backside of leaves, whereas the pairwise t-test provided a more explicit way to quantify the spectral differences between two sides.

2.5 Band Selection and Disease Severity Estimation

The significance level of spectral differences between healthy and diseased leaves was quantified by an independent t-test analysis for each band, whereas the significance level of spectral differences between the frontside and backside of leaves was quantified by a pairwise t-test. To eliminate the impact from the leaf orientation, only those bands that were sensitive to disease yet insensitive to the orientation of leaves were chosen. Such band selection was facilitated by an overlapping procedure based on the results from the independent t-test and pairwise t-test as illustrated above. The efficiency of these overlapped bands was then tested by comparing the disease severity estimates generated by a Fisher linear discriminate analysis (FLDA) with the measured value [13]. Five measures, overall accuracy (OAA), average accuracy (AA), producer's accuracy, user's accuracy, and kappa coefficient were calculated from confusion matrix to evaluate the accuracies of the discriminate model.

3 Results

3.1 Spectral Differences between Frontside and Backside of Leaves

The curves of raw reflectance, band ratios between frontside and backside, and corresponding *p*-values of pairwise t-test were summarized in Fig. 1. From the ratio curves, it is obvious that the spectral differences between the frontside and backside of healthy leaves showed a similar pattern with diseased leaves. The spectral reflectance of backside was higher than the frontside in most of the bands, except for the bands at 646-690 nm. For the results of pairwise t-test, a threshold of p-value<0.05 was used to identify those insensitive bands to the orientation of leaves. The identified bands differed greatly in their positions between healthy and diseased samples. The bands within 741-923 nm were found to be orientation insensitive for healthy leaves, whereas the bands within 606-697 nm and 740-1000 nm were identified as orientation insensitive for diseased leaves. It should be noted that most of these bands were located at some specific positions that were responsible for the absorption of chlorophyll and water, and cellular structure.

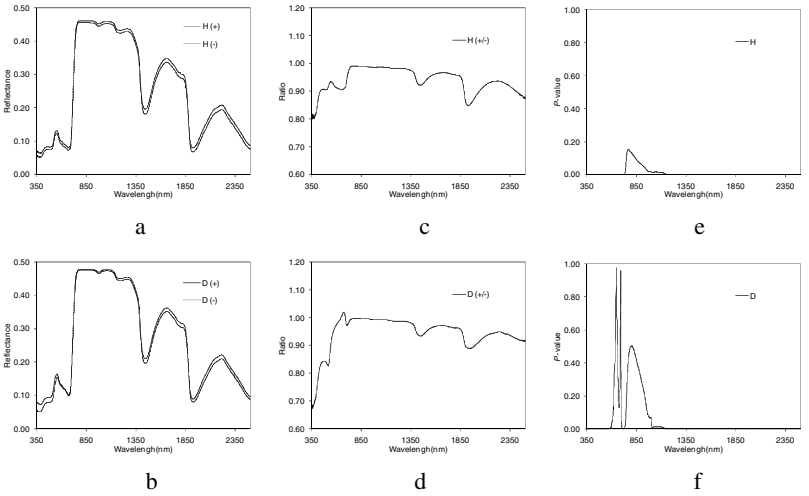


Fig. 1. Curves of original reflectance, band ratios between frontside and backside and their corresponding *p*-values based on pairwise t-test
(a, b) curves of original reflectance; (c, d) curves of band ratio; (e, f) curves of *p*-value. "H" indicates the healthy leaf samples; "D" indicates the diseased leaf samples. (+) indicates the frontside of leaves, (-) indicates the backside of leaves.

3.2 Selection of Effective Wavelengths and Disease Severity Determination

As shown in Fig. 2, the spectral difference for most bands reached the significant level except for several bands at the beginning of the spectrum. This phenomenon indicated

that the disease can induce a clear spectral response within a wide spectral region from visible to shortwave near infrared regions. Combining with the bands that were selected in section 3.1, it is possible to generate bands that were sensitive the disease yet insensitive to leaf orientation through an overlapping procedure. As shown in Fig.3, only the bands at 606-697 nm and 740-1000 nm were retained. These overlapped bands are theoretically suitable for disease detection since they are able to suppress the disturbance from the leaf orientation when monitoring yellow rust disease. Based on these bands, the FLDA model was established based on the calibration dataset (Table 1). The model accuracies were satisfactory in general, with OAA of 0.71 and kappa coefficient of 0.61 for calibration samples, and OAA of 0.70 and kappa coefficient of 0.59 for validation samples (Table 1).

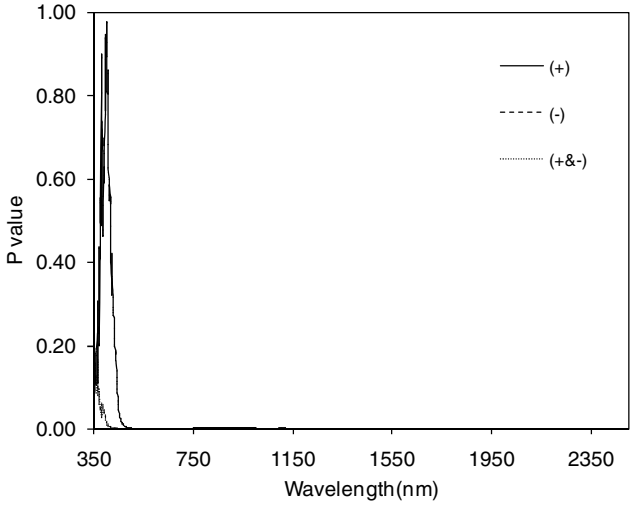


Fig. 2. Curves of the *p*-values of independent t-test between disease severity and band reflectance (+) indicates the frontside of leaves, (-) indicates the backside of leaves, (+&-) indicates the pooled of both

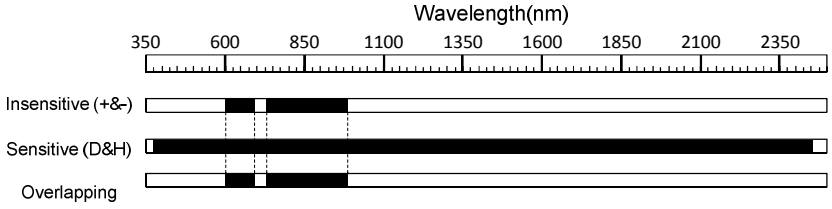


Fig. 3. Overlapping band selection
"Insensitive(+&-)" represent bands that are insensitive to the orientation; whereas "Sensitive(D&H)" represent bands that are sensitive to the yellow rust

Table 1. Confusion matrices created based on the FLDA models

		Reference								
		Healthy	Slight	Moderat	Heavy	Sum	U.'s a. (%)	OAA	AA	Kappa
Calibration	Healthy	25	4	1	0	30	83.33	0.71	0.73	0.61
	Slight	5	23	5	0	33	69.70			
	Moderate	2	5	12	1	20	60.00			
	Heavy	0	2	6	17	25	68.00			
	Sum	32	34	24	18	108				
	P.'s a. (%)	78.13	67.65	50.00	94.44					
Validation	Healthy	16	4	0	0	20	80.00	0.70	0.73	0.60
	Slight	2	16	4	0	22	72.73			
	Moderate	2	4	8	0	14	57.14			
	Heavy	0	0	6	12	18	66.67			
	Sum	20	24	18	12	74				
	P.'s a. (%)	80.00	66.67	44.44	100					

OAA is overall accuracy; AA is average accuracy; P.'s a. represents producer's accuracy; U.'s a. represents user's accuracy.

4 Discussion and Conclusion

In this study, based on the leaf spectra data with a large sample size, the reflectance of frontside and backside were found to be with significant difference at several bands. This finding was in good agreement with Zhou and Wang (2002)'s study, which compared the reflectance of frontside and backside of rice leaves under different nitrogen treatments [14]. Differences in surface reflectance were attributed to the dorsiventral morphology of wheat leaves. This might related with the larger portion of sclerenchyma, vascular bundle and bulliform cells on the upper leaf side of graminaceae [15-16]. The absolute difference of reflectance between the two surfaces was about 4% for the healthy leaves, and 3% for the diseased leaves, which means that the occurrence of disease diminished the spectral difference between frontside and backside of leaves. It was found that the reflectance of yellow rust pustules was higher than healthy positions at visible spectral region [17]. The appearance of those pustules at both surfaces for diseased leaves may whiten the spectral difference between the frontside and backside for healthy positions, which thereby lead to a reduction of spectral difference for those diseased leaves. The spectral difference between frontside and backside was relatively weak, comparing with the spectral signals that were induced by disease.

With a spectral overlapping procedure, the spectral reflectance at 606-697 nm and 740-1000 nm bands were selected for the detection of yellow rust since it can eliminate the impact from the orientation of leaves. However, it should be noted that the present study was conducted at leaf scale, which did not consider more complicated situations that might existed at canopy scale. For example some leaves may present as partially frontside and partially backside. To account for this situation, the spectral imaging techniques and some radiometric transferring models, such as PROSPECT+SAIL should be incorporated for mechanism study and model development.

Acknowledgments. This work was subsidized by National Key Technology R&D Program (2012BAH29B02), National Natural Science Foundation of China (41071276, 41101395), Beijing Natural Science Foundation (4122032). The authors are grateful to Mr. Weiguo Li, and Mrs. Hong Chang for data collection.

References

1. Li, G.B., Zeng, S.M., Li, Z.Q.: Integrated Management of Wheat Pests, pp. 185–186. Press of Agriculture Science and Technology of China, Beijing (1989) (in Chinese)
2. Strange, R.N., Scott, P.R.: Plant Disease: A threat to global food security. Annual Reviews Phytopathol 43, 83–116 (2005)
3. Muhammed, H.H.: Hyperspectral crop reflectance data for characterizing and estimating fungal disease severity in wheat. Biosystems Engineering 91, 9–20 (2005)
4. Bravo, C., Moshou, D., West, J.S., McCartney, A., Ramon, H.: Early disease detection in wheat fields using spectral reflectance. Biosystems Engineering 84, 137–145 (2003)
5. Huang, W.J., Huang, M.Y., Liu, L.Y., Wang, H., Zhao, C.J., Wang, J.D.: Inversion of the severity of winter wheat yellow rust using proper hyperspectral index. Transactions of the CSAE 21, 97–103 (2005)
6. Huang, W.J., David, W.L., Niu, Z., Zhang, Y.J., Liu, L.Y., Wang, J.H.: Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. Precision Agriculture 8, 187–197 (2007)
7. Zhang, J.C., Huang, W.J., Li, J.Y., Yang, G.J., Luo, J.H., Gu, X.H., Wang, J.H.: Development, evaluation and application of a spectral knowledge base to detect yellow rust in winter wheat. Precision Agriculture 12, 716–731 (2011)
8. Huang, M.Y., Huang, W.J., Liu, L.Y., Huang, Y.D., Wang, J.H., Zhao, C.J., Wan, A.M.: Spectral reflectance feature of winter wheat single leaf infected with stripe rust and severity level inversion. Transactions of the CSAE 20, 176–180 (2004); (in Chinese with English abstract)
9. Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., Ramon, H.: Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. Computers and Electronics in Agriculture 44, 173–188 (2004)
10. Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, A., McCartney, A., Ramon, H.: Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. Real-Time Imaging 11, 75–83 (2005)
11. Graeff, S., Link, J., Claupein, W.: Identification of powdery mildew (Erysiphe graminis sp. tritici) and take-all disease (Gaeumannomyces graminis sp. tritici) in wheat (Triticumaestivum L.) by means of leaf reflectance measurements. Central European Journal of Biology 1, 275–288 (2006)
12. Luedeling, E., Hale, A., Zhang, M.H., Bentley, W.J., Dharmasri, L.C.: Remote sensing of spider mite damage in California peach orchards. International Journal of Applied Earth Observation and Geoinformation 11, 244–255 (2009)
13. McLachlan, G.J.: Discriminant analysis and statistical pattern recognition. Wiley Interscience (2004)
14. Zhou, Q.F., Wang, J.H.: Comparison of adaxial and abaxial surface reflectance under different nitrogen level. Transactions of the CSAE 18, 34–39 (2002)
15. Takahashi, K., Mineuchi, K., Nakamura, T., Koizumi, M., Kano, H.: A system for imaging transverse distribution of scattered light and chlorophyll fluorescence in intact rice leaves. Plant Cell & Environment 17, 105–110 (1994)
16. Kirby, E.J.M.: Botany of the wheat plant. In: Bread Wheat, Improvement and Production (Curtis B C; Rajaram S; Gomez Macpherson Heds). FAO Plant Production Series No. 30, Rome (2002)
17. Devadas, R., Lamb, D.W., Simpfendorfer, S., Backhouse, D.: Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. Precision Agriculture 10, 459–470 (2009)

Forecasting the Total Power of China's Agricultural Machinery Based on BP Neural Network Combined Forecast Method

Jinyan Ju[1], Lin Zhao[1], and Jinfeng Wang[2]

[1] School of Mechanical Engineering, Heilongjiang Institute of Science and Technology, Harbin 150027, China
jujinyan2007@yahoo.cn
[2] School of Engineering, Northeast Agricultural University, Harbin 150030, China

Abstract. In view of the limitations of single forecast model, forecasted results of different models will have some differences, in order to improve the forecast precision and the forecast results reliability, on the basis of determining the single forecast model for total power of China's agricultural machinery, the nonlinear combined forecast model for total power of agricultural machinery was established based on BP neural network using MATLAB software, and then the model was trained and simulated. The simulation results show that the fitting mean absolute percentage error of nonlinear combined forecast model is 0.59%, which is lower than 2.57%,2.66% and 2.09% of exponential model, GM（1, 1) model and cubic exponential smoothing model .The established models were validated using original data of China's total power of agricultural machinery from 2009 to 2011, validation results show that the combined forecast model has the lowest forecast error 0.64%, the validation effect is the best, which can improve the forecast precision for total power of China's agricultural machinery. The total power of China's agricultural machinery was forecasted from 2012 to 2020 using the combined model, and the forecast results show that the total power of China's agricultural machinery will maintain a rapid growth trend in the next few years; it will be 1223987.1 MW by 2015 and 1603498.2 MW by 2020.

Keywords: total power of agricultural machinery, BP neural network, combined forecast, simulation[1].

1 Introduction

The total power of agricultural machinery means the sum of all mechanical power used for agriculture, forestry, animal husbandry, fishery production and transportation, which reflects the general development level of agricultural mechanization, and is the main planning index for the development of agricultural mechanization. At present, the main methods to forecast the total power of agricultural machinery are linear regression, moving average, grey GM (1, 1) model, exponential smoothing, curve

[1] Supported by Youth Foundation of Heilongjiang province Scientific Committee (Project no. QC2011C007).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 85–93, 2013.

fitting, data envelopment analysis, neural network model, combined forecast and so on [1-4]. These methods each have its advantages and disadvantages, and the forecast results are different too.

This paper chose three kinds of forecast method, which are exponential model, cubic exponential smoothing and grey GM (1, 1) respectively, to establish the single forecasting model for the total power of China's agricultural machinery. In view of the limitations of single forecast model [3-5], in order to improve the forecast precision and the forecast results reliability, the author combined forecasting results of different models together, making up for each other's deficiencies. Due to BP neural network has the ability to approximate any nonlinear function; the nonlinear combined forecast model based on BP neural network was established. Through comparing the forecast accuracy of the four developed forecast models, the nonlinear combined forecast model based on BP neural network has the best fitting effect and the forecast accuracy. Therefore, a new method for the forecast of total power of China's agricultural machinery is obtained, the forecast results can provide important theory references for China's farm machinery department to make the development planning of agricultural mechanization, publish relevant policies and apply for the financial support.

2 Forecasting Models

2.1 Exponential Model

As shown in table 1, the original statistical data of the total power of China's agricultural machinery is growing along with time, original time series from1985 to 2008 can be fitted by curve models using SPSS software. The fitting results show that the exponential model has the best fitting effect, the determination coefficient $R^2 = 0.994$, the fitting accuracy is higher, the model is significant, and the obtained forecast model is

$$Y = 199591.631 e^{0.058x} \tag{1}$$

Where Y is the total power of agricultural machinery, unit: MW; x is the time variable, corresponding value from 1985 to 2008 is respectively from 1 to 24.

The total power of China's agricultural machinery from 1985 to 2008 were fitted with Eq.1, the results are shown in table 1, and the fitting mean absolute percentage error (MAPE) is 2.57%.

2.2 Grey GM (1, 1) Forecast Model

Grey GM (1, 1) forecast model is suitable for data sequence which has small fluctuate along with time sequence, and its short-term forecast accuracy is higher[2,5]. The basic form of grey GM (1, 1) model is shown as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (k = 1, 2, \cdots, n) \tag{2}$$

If the original data sequence is given $X^{(0)}$: $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$, Where $x^{(0)}(k) \geq 0, k = 1, 2, \cdots, n$. Accumulating the original data sequence for one time to obtain the new data sequence $X^{(1)}$, $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)$, Where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, \cdots, n$ .The obtained new data sequence is a monotone growth curve, which increases the regularity of original data, and weakens the volatility. Sequence $Z^{(1)}$ is constructed by $X^{(1)}$.

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \cdots z^{(1)}(n))$$

Where $Z^{(1)}(k) = 0.5(x^{(1)}(k) + x^{(1)}(k-1)), k = 2, 3 \cdots, n$ .Provided that parameter column $\hat{a} = [a, b]^{\mathrm{T}}$,and

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

Then the least square estimation of parameter column for Eq.2 satisfies :

$$\hat{a} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Y = \begin{bmatrix} -0.059335 \\ 202182.883 \end{bmatrix} \tag{3}$$

Then the equation $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$ is called the whitening differential equation for Eq.2, and also called shadow equation, this whitening differential equation can be solved by MATLAB, and then the result is obtained:

$$x^{(1)}(k) = 3616630.945\mathrm{e}^{0.059335(k-1)} - 3407505.95 \quad (k = 1, 2, \cdots n) \tag{4}$$

Restore value: $$x^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{5}$$

The total power of agricultural machinery from 1985 to 2008 can be estimated using Eq.4 and Eq.5. The results are shown in table 1, which show that the fitting mean absolute percentage error is 2.66%.

2.3 Cubic Exponential Smoothing Model

The development of the total power of agricultural machinery not only has great relationship with that year's social development but also with the recent year's development, and cubic exponential smoothing method is suitable for this kind of

Table 1. Fitted results and errors of total power of China's agricultural machinery using different models

years	actual value /MW	exponential model		grey GM (1, 1) model		cubic exponential smoothing model		BP neural network combined forecast model	
		fitted values /MW	relative errors/%	fitted values /MW	relative errors/%	fitted values /MW	relative errors/%	fitted values /MW	relative errors/%
1985	209125	211588	1.18	209125	0				
1986	229500	224305	2.26	221085	3.67				
1987	248360	237787	4.26	234600	5.54	233575	5.95	247858	0.20
1988	265750	252079	5.14	248941	6.32	261097	1.75	266613	0.32
1989	280670	267230	4.79	264159	5.88	284861	1.49	280182	0.17
1990	287077	283292	1.32	280307	2.36	302496	5.37	287238	0.06
1991	293886	300318	2.19	297443	1.21	307194	4.53	293757	0.04
1992	303084	318369	5.04	315625	4.14	309304	2.05	305376	0.76
1993	318166	337504	6.08	334920	5.27	314824	1.05	319730	0.49
1994	338025	357790	5.85	355393	5.14	329274	2.59	338722	0.21
1995	361181	379294	5.02	377119	4.41	351863	2.58	362007	0.23
1996	385469	402092	4.31	400172	3.81	379589	1.53	388287	0.73
1997	420156	426259	1.45	424635	1.07	408479	2.78	415810	1.03
1998	452077	451879	0.04	450593	0.33	448562	0.78	449158	0.65
1999	489961	479039	2.23	478137	2.41	486245	0.76	482232	1.58
2000	525736	507831	3.41	507366	3.49	528393	0.51	518190	1.44
2001	551721	538354	2.42	538382	2.42	567439	2.85	553629	0.35
2002	579299	570711	1.48	571293	1.38	592537	2.29	583607	0.74
2003	603865	605013	0.19	606216	0.39	615896	1.99	613727	1.63
2004	640279	641377	0.17	643274	0.47	636173	0.64	643405	0.49
2005	683978	679927	0.59	682598	0.20	671016	1.90	680827	0.46
2006	725221	720793	0.61	724325	0.12	718664	0.90	725283	0.01
2007	765896	764116	0.23	768603	0.35	765317	0.08	770469	0.60
2008	821904	810042	1.44	815588	0.77	809243	1.54	815623	0.76

change trend, especially when the change trend continues for a period, this method can also be used for middle or long term forecast [4,7]. The basic form of the model is

$$Y_{t+T'} = a_t + b_t T' + \frac{1}{2} c_t T'^2 \tag{6}$$

Where $Y_{t+T'}$ is the forecasted value of the total power of agricultural machinery, t is the starting year for forecasting, a_t、b_t、c_t are smoothing coefficients , T' is the number of time cycle for forecasting.

According to the statistical data of total power of agricultural machinery from 1985 to 2008 as shown in table 1, then cubic exponential smoothing model was obtained:

$$Y_{2008+T'} = 819169.19 + 49211.63T' + 2566.02T'^2 \tag{7}$$

The total power of agricultural machinery from 1987 to 2008 can be estimated using cubic exponential smoothing equation 7, the results are shown in table 1, which show that the fitting mean absolute percentage error is 2.09%.

3 Nonlinear Combined Forecast Model Based on BP Neural Network

In view of the limitations of single forecast model, in order to improve the forecast precision and the forecast results reliability, the author combined forecast results of different models together, making up for each other's deficiencies, and then the combined forecast model was established to forecast the total power of agricultural machinery. Due to BP neural network has a series of excellent characteristics such as nonlinear, robustness, adaptive, self-organization, and especially is suitable to establish nonlinear forecast model [7-9]. Therefore, the nonlinear combined forecast model based on BP neural network was established, which can reflect the information contained in the three forecast models, and the nonlinear mapping relationship between each single forecasting results and actual values was constructed.

3.1 Basic Steps for Modeling

(1) Ascertaining the input and output vector. The forecasted values of total power of agricultural machinery from 1987 to 2008,which are obtained through the established three forecast models, are taken as the neural network input vector （P） , the actual values of the total power of agricultural machinery are taken as the neural network output vector (T).The input and output layer nodes of neural network are ascertained according to the actual problem, namely, the input layer has three nodes, and the output layer has one node, so the nonlinear mapping relationship between each single forecasting results and actual values was constructed. In this paper the basic form of nonlinear forecast model is $Y' = f(x_1, x_2, x_3)$, where $x_i (i = 1, 2, 3)$ is respectively the forecast results of exponential model, grey forecast model and cubic exponential smoothing model, Y' is the neural network output vector, f is the nonlinear function decided by neural network weights and thresholds.

(2) Ascertaining the hidden layer and hidden layer nodes. A network with a hidden layer and a linear output layer can approximate any rational function, and the forecasting accuracy can be improved by increasing the number of hidden layer nodes [7,8]. Therefore, in this paper BP neural network structure is adopted one hidden layer, namely the structure is $3-j-1$, where j is the number of hidden layer node, j can be determined according to the relationship between error and its number of nodes, and the number of node can be determined by gradually increasing or reducing method [8,9].

(3) Selecting transfer function and training function. The transfer function of hidden layer was selected Sigmoid, the output layer was selected Pureline, and the training function was selected Trainlm.

3.2 The Simulation of BP Neural Network

BP neural network is established and simulated with MATLAB. In order to prevent partial neuron nodes supersaturated, the original data are normalized with mapminmax function before the network is trained, namely [X, ps] = mapminmax (P), where X is the obtained normalized data, structural body ps records the standardization mapping relationship. The new input data also need to be normalized using mapminmax function, namely making X1 = mapminmax (' apply ', P1, ps). In order to ensure the new input data belong to [-1, 1], in the process of normalization original data, a large enough value can be input as the maximum value of original data in advance. Finally, normalized data are recovered with T = mapminmax (' reverse ',t, ps). The training parameters of the network need to be set.

Finally when the number of nodes is 2, the fitting effect and the generalization ability of network are better, the determined network structure is 3-2-1, the training results further analysis curves is shown in figure 1.

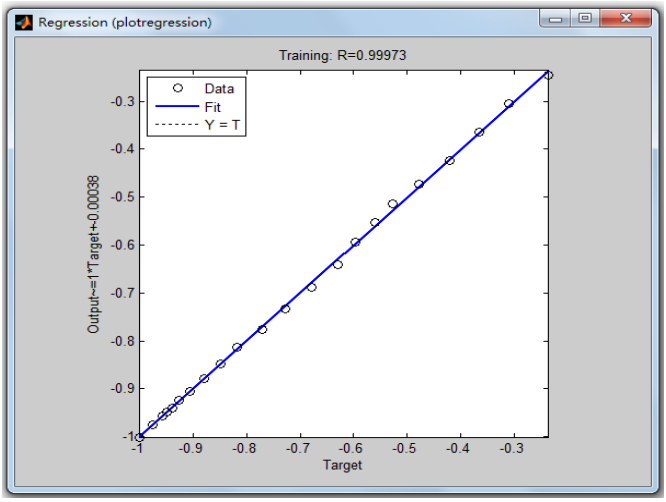


Fig. 1. Training result analysis graph of BP neural network combined model for total power of China's agricultural machinery

As shown in figure 1, the abscissa is target outputs, ordinate is network output, the ideal regression straight line is a solid line, and the optimum regression straight is a dotted line. The linear regression relationship between the network output and target output is shown in this figure, the changing rate value R is more close to 1, which means the network output and target output are more closed, the network performance is better. As shown in figure 2, the R value is 0.99973, which means the network has very good performance. At this time the fitting values of the total power of agricultural machinery are shown in table 1, and the fitting average error is 0.59%. The total power of agricultural machinery from 2009 to 2011 were used as the test samples to validate the generalization ability for the established model, forecasting values of the test samples are shown in table 2, the average forecast error is 0.64%, which has higher precision, and stronger generalization ability.

Table 2. Forecasted results and errors of test samples of total power of China's agricultural machinery

years	actual values /MW	exponential model		grey GM (1, 1) model		cubic exponential smoothing model		BP neural network combined forecast model	
		fitted values /MW	relative errors /%	fitted values /MW	relative errors /%	fitted values /MW	relative errors /%	fitted values /MW	relative errors /%
2009	874961	858729	1.86	865445	1.09	869664	0.61	869799	0.59
2010	927805	910345	1.88	918350	1.02	922724	0.55	921918	0.63
2011	970000	965058	0.51	974489	0.46	978351	0.86	976693	0.69

3.3 Forecasting Method Comparison for the Total Power of China's Agricultural Machinery

The fitting mean absolute percentage error is 0.59% for the nonlinear combined forecast model based on BP neural network, which is lower than the selected exponential model, grey GM (1, 1) model and cubic exponential smoothing model of 2.57%, 2.66% and 2.09%. The validation forecast was carried out for the total power of agricultural machinery from 2009 to 2011 using different methods. The results show that the test samples error is 0.64% for combined forecast model based on BP neural network, which has the best forecasting effect. For nonlinear combined forecast model based on BP neural network has higher fitting precision and higher forecast precision, therefore this model can effectively improve the forecast precision of total power of China's agricultural machinery. The total power of agricultural machinery from 2012 to 2020 in China was forecasted using this combined forecast method, the forecasted results are shown in table 3.

Table 3. Forecasted results of total power of China's agricultural machinery from 2012 to 2020

years	2012	2013	2014	2015	2016	2017	2018	2019	2020
forecasted value /MW	1034201	1094529	1157759	1223987	1293304	1365808	1441604	1520797	1603498

4 Conclusions

In purpose of improving forecast accuracy for the total power of China's agricultural machinery, comprehensive utilizing the effective information provided by each single forecast method, and the nonlinear mapping capability of neural network technology, the nonlinear combined forecast model based on BP neural network was put forward, and the combined forecast model for the total power of China's agricultural machinery was established.

The fitting precision and forecast accuracy of different models for the total power of China's agricultural machinery are compared, the results show that nonlinear combined forecast method based on BP neural network has significantly higher forecast precision than any other single forecast model, which effectively improves the forecast accuracy, and provides a new way to forecast the total power of agricultural machinery. The total power of China's agricultural machinery from 2012 - 2020 was forecasted with this model, and the forecasted results show that the total power of China's agricultural machinery will maintain a rapid growth trend, it will be 1223987.1 MW by 2015 and 1603498.2 MW by 2020,which is consistent with the actual development situation of China's agricultural mechanization, and can provide important theory reference for China's agricultural machinery department to make the development planning of agricultural mechanization, publish relevant policies and apply for the financial support.

References

1. Zhu, R., Wang, F.: Trend Envelope Predict and Analysis of Heilongjiang Province Agriculture Machinery Total Power. Journal of Northeast Agricultural University 38(4), 512–515 (2006)
2. Zhang, H.: Forecasting of Agricultural Mechanical Total Power and Application Comparison of Multi-mathematic Methods. Chinese Agricultural Mechanization (2), 50–52 (2006)
3. Zhu, R., Huang, Y., Yang, X.: Method for Estimating Total Power of Agricultural Machinery Based on Mixed Grey Neural Network. Transactions of the Chinese Society of Agricultural Engineering 22(2), 107–110 (2006)
4. Zhang, S., Zhao, F.: Combinatorial Forecast of Agricultural Machinery Total Power Based on Shapley Value. Transactions of the Chinese Society for Agricultural Machinery 39(5), 60–64 (2008)
5. Ju, J., Wang, J.: Prediction Method for the Operation Level of Agricultural Mechanization in Heilongjiang Province. Transactions of the Chinese Society of Agricultural Engineering 25(5), 83–88 (2009)
6. Ju, J., Wang, J., Wang, J.: Combined Prediction Method of Total Power of Agricultural Machinery Based on BP Neural Network. Transactions of the Chinese Society for Agricultural Machinery 41(6), 87–92 (2010)

7. Wang, J., Zhao, Y., Ma, L.: Research on Combination Forecasting Method and Application in Electric Power Load Forecasting. Journal of Northeast Agricultural University 39(4), 51–54 (2008)
8. Aminian, F., Dante Suarez, E., Aminian, M., Walz, D.T.: Forecasting Economic Data with Neural Networks. Computational Economics 28(1), 71–88 (2006)
9. Singh, A.K., Panda, S.S., Chakraborty, D., Pal, S.K.: Predicting Drill Wear using an Artificial Neural Network. The International Journal of Advanced Manufacturing Technology 28(5-6), 456–462 (2006)

Self-Organizing Map Analysis on Peanut Yield and Agronomy Characteristics

Yujian Yang[1] and Mingchuan Ji[2,*]

[1] S & T Information Engineering Technology Center of Shandong Academy of Agricultural Science, Information Center of Agronomy College of Shandong University
[2] The Institute of Sustainable Development of Shandong Agricultural Science Academy, Number 202 Gongye North Road, Licheng District of Jinan, 250100, Shandong Province, P.R. China
yyjtshkh@gmail.com

Abstract. The model system between peanut yield and agronomy characteristics which is nonlinear, irreversible and dissipative. The objective in the study was the peanut cultivated in the different ecological regions in Shandong province, aimed to establish the new non-nonlinear model based on Self-Organizing Maps (SOM) to improve the cultivation information of peanut growth process. In the article, applying SOM network achieved the cluster between peanut yield and agronomy characteristics about 4 variables, involved in plant height, branches, full pods and peanut yield ratio. MATLAB 7 software is used to classify 60 samplings of peanut yield and agronomy characteristics. It is concluded that the SOM network can respond the complicated information classification among each peanut yield, during the analysis, the results also showed SOM method is the most suitable for peanut yield and characteristics classification, especially analysis of clusters on basis of peanut agronomy parameters, so the study can be applied on agronomy characteristics and peanut yield of the different ecological regions in Shandong province.

Keywords: peanut, Self-Organizing Mapping, yield, agronomy characteristics.

1 Introduction

As the important oil and economic plants in China, peanut is widely planted in Shandong province, where is the core area of production and export of the superior peanut cultivars (Wan Shubo, et al., 2007, 2009)[1,2]. An important subject between peanut yield and all kinds of agronomy characteristics has been reported in recent years. The quantitative relationship focused in yield and its correlated traits (Zhang Haiyan, et al., 2006)[3]. The important study was mainly developed by the some statistical methods, such as correlation analysis and regression analysis, which carried on the variation of genetic factors and environmental conditions of peanuts

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 94–100, 2013.

(Wu Zhengfeng, 2008)[4].The capability of Self-Organizing Maps (SOM) to visualize high-dimensional data is well known, an important SOM subject between peanut yield and agronomy characteristics has been rarely reported at home and abroad. In recent years, the SOM is widely applied in many fields, such as model establishment, imaging application, text clustering. There is the rarely reports of clustering results considering peanut yield and agronomy characteristics with the support of SOM, so the main objective in the article is achievement of effective clusters about peanut yield. Correspondingly, the organization of the paper is as follows: a first section describes the basic SOM algorithm with a deeper insight of its capabilities. An approach, based on SOM visualization for peanut yield and agronomy characteristics is suggested in the next section for 60 samples. The last section discusses the obtained results, followed by the conclusion.

2 Material and Methodology

2.1 Materials and Experimental Designing

Materials and data were collected from multiple sites in 9 different ecological regions about 7 peanut varieties in Shandong province in 2007. Experimental sites with ecological regions includes Laixi city, Jiaozhou city, Changyi city, Qixia city, Haiyang city, Donggang District, Lanshan District, Dongchangfu District and Ningyang country in Shandong province. 7 peanut varieties were selected in the paper with obvious representative, they are respectively: “LuHua 11”,“LuHua 12”,“LuHua 14”,“Huayu 20”,“Weihua8”, especially,“Huayu22” ,“Huayu23”,“Huayu22” for the traditional big export-oriented peanut, “Huayu23” for the traditional small export-oriented peanut, the others is high yield peanuts. In the experimental designing, the local plant mode was selected for spring peanuts or the summer peanuts. Each site is the same planting specifications, for spring peanut, ridge is 80-85cm, wide is 50cm, the seeding approach was hill-drop, row spacing is 30cm, the hill spacing is 17cm, 9,800 plants were planted per 666.7 m2, interplanting cultivation mode of peanut was selected to be naked before 20d, 10,000 holes were dug per 666.7 m^2[5]. In the process of data measurement, 4 traits variables of peanut(plant height, branches, full pods and peanut yield ratio) was acquired involved in 60 samples. In order to achieve the SOM results, we determined the cluster numbers of peanut yield according to SOM algorithm, so we developed the preprocessing classification, there are 3 kinds of productions including 3000-4000 Kg/per ha, 4000-5000Kg/per ha and >5000 Kg/per ha. Use the SOM for clustering data correlated with the agronomy characteristics of input data (plant height, branches, full pods and peanut yield ratio), which provides a topology preserving mapping from the high dimensional space to map units. Map units, or neurons, usually form a two-dimensional lattice and thus the mapping is a mapping from high dimensional space onto a plane. In the literature [6,7], the effective characteristics of SOM method was determined by the sample distribution characteristics.

2.2 Use of the Self-Organizing Map Algorithm

2.2.1 Self-Organizing Map Algorithm

The Self-Organizing Map has been proven useful in many applications, it belongs to the category of competitive learning networks, so it is also called SOM, which is put forward by T. Kohonen(1981)[6]. It is a neural network that maps signals from a high-dimensional space to a one- or two-dimensional discrete lattice of neuron units, or is named after cluster algorithm on basis of neural network. SOM embodied the similarities in samples, achieved the transformation from high-dimensional data to low-dimensional data. In many perspectives, spatial structure and function of neurons of SOM is very important, the traditional model of neural network didn't consider the spatial structure characteristics, but SOM not only considering the structure, but also achievement of the similar input data responded the best matching units without supervision and prior knowledge[8,9,10].

SOM has the better status of non-linear clustering, it was also used to find correlations between the data by labeling the neurons of the SOM using the training set and finding the best-matching-units for every example. In theory, each neuron stores a weight. The map preserves topological relationships between inputs in a way that neighboring inputs in the input space are mapped to neighboring neurons in the map space. And by grouping units that respond to similar stimuli together. Nerve cells, neurons, in the cortex of the brain seem to be clustered by their function. During the computation, by defining the Hebbian learning rule is to determine the learning rate and update the relationship for best-matching units on the map. In essence the learning rule of the SOM defines the model as a collection of competitive units that are related through the neighborhood function. In practice, the units are placed on a regular low dimensional grid and the neighborhood is defined as a monotonically decreasing function on the distance of the units on the map lattice, thus creating a latent space, which has the dimension of the map grid and flexibility determined by the neighborhood function, and embedding when the dimension of the map grid matches the dimension of the input data manifold. In general, gaussian function is selected in the neighborhood function[11].

SOM, based on unsupervised learning, or high-dimensional observations projected to the two-dimensional coordinate system, which means that no human intervention is needed during the learning and that little needs to be known about the characteristics of the input data. In a word, Self-Organizing Maps is a categorization method, a neural network technique and the unsupervised characteristics.

The main computing steps of SOM, including construction of data sets, data preprocessing, initialization and training of input data and visualization and analysis of the correlated results. In the elaborate SOM computing process, important parameters should be also considered as the significant contents, such as Unified distance matrix (U-matrix), importance degree in the trained self-organizing map(D).

Especially, U-matrix, the parameter representation of the Self-Organizing Map visualizes the distances between the neurons, the distance between the adjacent neuons is calculated and presented with different colorings between the adjacent

nodes, for each node in the map, compute the average of the distances between its weight vector and those of its immediate neighbors. In the results, the different colors between the neurons presents the different distance significance and responds the cluster differences, average distance is a measure of a node's similarity between it and its neighbors, this can be a helpful presentation when one tries to find clusters in the input data without having any a priori information about the clusters. For the parameter D, it responds the important degree of the agronomy characteristics impact on peanut yield in the article[12,13].

2.2.2 Data Processing and Data Analysis

First, the data has to be brought into Matlab using construction of data, second, we used the function som_normalize for data preprocessing data to perform a linear scaling to unit variance. The function som_make is the basic function to use when creating and training a SOM, it is a convenient tool that combines the data of creating, initialization and training a SOM. After the data set is ready, the data set is loaded into Matlab7 and normalized, the variance normalization is used, a SOM is trained. Since the data set had labels, the map is also labeled using som_autolabel. There are a variety of methods to visualize the SOM, the basic tool is the function som_show, it can be used to show the U-matrix and the component planes of the SOM. After this, the SOM is visualized using som_show[14,15].

Here is the usage of the Matlab Toolbox to make and visualize a SOM of a data set about peanut yield and agronomy characteristics. This data set consists of four measurements from 60 samples: the first classification of peanut yield has 18 samplings, the second classification of peanut yield has 13 samplings and the third classification of peanut yield has 29 samplings. The data is in an ASCII file, the first line contains the names of the variables, the corresponding sequences, including plant height characteristics, branches characteristics, full pods characteristics, peanut yield ratio characteristics and peanut yield classification. Each of the following lines gives one data sample beginning with numerical variables and followed by labels[16].

3 Results and Analysis

We developed the cluster analysis based Matlab7 SOM Tool. the PCA-projection of both data and the map grid was required by the program computerization, in general, the projection also presented the four variables and the subspecies information from the SOM. Projection of the sample data set to the subspace spanned by its three eigenvectors with greatest eigenvalues. The SOM grid has been projected to the same subspace, and visualizes all four variables of the SOM plus the subspecies information using coordinates, neighboring map units are connected with lines, labels associated with map units are also shown.

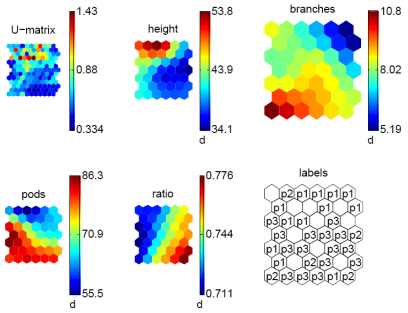


Fig. 1. Visualization of the SOM of peanut yield and agronomy parameters data

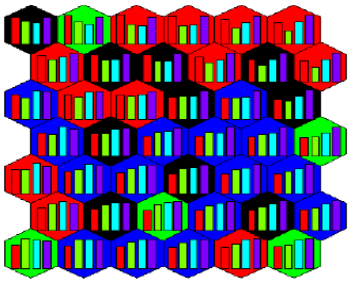


Fig. 2. SOM ordination diagram of characteristics

U-Matrix visualization provides a simple way to visualize cluster boundaries on the map, the U-matrix visualizes distances between neighboring map units, and thus shows the cluster structure of the map, high values of the U-matrix indicate a cluster border, uniform areas of low values indicate clusters themselves. For each node in the map, compute the average of the distances between its weight vector and those of its immediate neighbors, average distance is a measure of a node's similarity between it and its neighbors. Each component plane shows the values of one variable in each map unit. On top of these visualizations, additional information can be shown: labels, data histograms and trajectories.

As illustrated in Figure 1, Gradient distribution of peanut characteristics in the trained SOM ordination diagram. U-matrix on top left, then component planes(including plant height characteristics ; branches characteristics ; full pods characteristics ; peanut yield ratio characteristics), and map unit labels on bottom right. The six figures are linked by positions, in each figure, the hexagon in a certain position corresponds to the same map unit, additional hexagons exist between all pairs of neighboring map units. The map unit in top right corner has low values for plant height, branches and full pods, and relatively high value for peanut yield ratio. The label associated with the map unit is "p1" and from U-matrix it can be seen that the unit is very close to its neighbors. From the U-matrix it is easy to see that the top two rows of the SOM form a very clear cluster. By looking at the labels, it is immediately seen that this corresponds to the peanut yield (p1) subspecies, the two other subspecies peanut yield (p2) and peanut yield (p3) form the other cluster. The U-matrix shows no clear separation between them, but from the labels it seems that they correspond to two different parts of the cluster. From the component planes it can be seen that the branches characteristics and full pods characteristics are very closely related to each other in the certain degree.

Figure 2 visualizes all four variables of the SOM plus the subspecies information, 4 variables barcharts (including plant height characteristics ; branches characteristics ; full pods characteristics ; peanut yield ratio characteristics) in the topography

maps. The distribution of 3 types peanut yield in the topography maps. In every topography map, the weight of each variable is shown in the figure. The four variables shown with the different barcharts in each map unit and in the background color indicates the subspecies.

4 Conclusions and Discussion

In this paper, SOM is used to classify 60 samplings of peanut data of the different ecological regions in Shandong province, the kind of yield classification method promotes the comparative analysis for the different ecological regions in Shandong province for the yield prospective. The results showed that SOM is an excellent tool in the visualization of high dimensional data about peanut yield and agronomy characteristics and as such it is most suitable for peanut traits and yield classification, especially analysis of clusters on basis of peanut agronomy parameters.

It is concluded that the SOM network can respond the complicated information among each peanut yield. The effect of classification is good and SOM considered the complicated characteristics of peanut agronomy parameters. And it can be applied on peanut characteristics and yield. Although SOM cluster applied in peanut yield and agronomy characteristics is a preliminary trail in the study, the SOM has the obvious characteristics from the high dimensional data to low dimensional data for peanut growth process, with the development of spatial analysis technology, combining with GIS data, the spatial evaluation on peanut agronomy characteristics and yield should be explored on the precision scale to clarify the spatial variability and spatio-temporal characteristics of complicated parameters under soil-peanut system in the next step.

References

1. Wan, S.: Peanut Quality. China Agricultural Scientific Press, Beijing (2007)
2. Wan, S., et al.: The high Quality of synergism Cultivation Theory and Technology of Peanut. China Agricultural Scientific Press, Beijing (2009)
3. Zhang, H., Jiao, B., Li, G.: Analysis on the Relationship between Yield and Correlated Quantitative Character of Soybean. Journal of Shanxi ricultural Science 34(2), 27–29 (2006) (in Chinese)
4. Wu, Z., Wang, C., Zheng, Y., et al.: Analysis of characteristics and stability of peanut yield in different ecological regions of Shandong Province. Chinese Journal of Eco-Agriculture 16(6), 1439–1443 (2008) (in Chinese)
5. Yang, Y., Wan, S.: The grey correlation degree analysis on yield of peanut and traits in warm temperature regions. In: Li, D., Yang, S.X. (eds.) Computer and Computing Technologies in Agriculture, vol. 24, pp. 161–167 (2010)
6. Kohonen, T.: The self-organizing map. Proc. IEEE 78(9), 1480–1497 (1990)
7. Vesanto, J., Alhoniemi, E.: Clustering of the Self-Organizing Map. IEEE Transactions on Neural Networks 11(3), 586–600 (2000)
8. Zhang, Q., et al.: Self-organizing feature map classification and ordination of Larix principisrupprechtii forest in Pangquangou Nature Reserve

9. Zhang, Y., et al.: Soil Classification Based on Self Organizing. Feature Mapping Neural Networks
10. Moshou, D., Gravalos, I., Kateris, D., Sawalhi, N., Loutridis, S.: Condition monitoring in centrifugal irrigation pumps with self-organizing feature visualization. In: EFITA/WCCA, pp. 116-124 (2011)
11. Kohonen, T.: Self-Organizing Maps. Springer Series in Information Sciences, vol. 30. Springer, Heidelberg (1995)
12. Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J.: SOM Toolbox for Matlab5 (2000)
13. Suri, G., Zhang, J.-T., Tian, S.-G., Zhang, Q.-D., Zhang, B., Cheng, J.-J., Liu, S.-J.: Application of self-organizing map to quantitative analysis of mountain meadow in the Songshan Nature Reserve of Beijing, China. Chinese Journal of Plant Ecology 34(7), 811–818 (2010)
14. Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J.: Self-organizing map in Matlab: the SOM Toolbox
15. Vesanto, J., Alhoniemi, E., Himberg, J., Kiviluoto, K., Parviainen, J.: Self-Organizing Map for Data Mining in MATLAB: the SOM Toolbox. Simulation News Europe 25, 54 (1999)
16. Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J.: Self-organizing map in Matlab: the SOM Toolbox. In: Proceedings of the Matlab DSP Conference 1999, Espoo, Finland, pp. 35–40 (November 1999)

Modeling and Simulating of Spatial Spread of Cross-Boundary Crop Diseases

Jiaogen Zhou [2], Xu Chen[1,*], Jingyin Zhao[1,*], and Dongsheng Wang[1]

[1] Center of Information Technology in Agriculture, Shanghai Academy of Agricultural Sciences, Shanghai 201403, China
[2] Institute of Subtropical Agriculture, The Chinese Academy of Science, Changsha 410125, China
{chenxu,zjy}@saas.sh.cn

Abstract. Understanding of the timing, orientation and transmission mode of crop disease spread is very important in prevention and treatment of crop diseases, especially for cross-border crop diseases. In this paper, we introduced the complex network approach to model trans-boundary spread of diseases. In the model, the process of crop disease spread is characterized as a network, where patches of crop planting are abstracted into nodes, and the diseases spread from a node to neighboring nodes. The edge between two nodes is connected, only if the Euclidean distance between two nodes does not exceed the radius of disease spread. We obtained wheat-growing distribution data of Beijing in 2007 by image classification, and further abstracted patches of wheat growing into network nodes, and finally constructed the disease spread network to analyze its topology characteristics. The experimental results show the heavy influence of spatial constraint on disease spread.

Keywords: complex network, spatial model, crop epidemics.

1 Introduction

Occurrence and spread of crop diseases is of great harm to crop yield and quality. Forecasting of occurrence and spread of crop diseases is very significant for diseases control. Existing forecasting methods can be divided into two categories: 1) linear or nonlinear statistical models; 2) multi-spectral remote sensing prediction.

In the linear or nonlinear statistical models, there is a close association between the occurrence and spread of crop pests and diseases with environmental factors of growth and reproduction habitat (such as light, heat, water, terrain, etc.).So, based on history data extensively collected and surveyed on the disease bio-statistical information (such as spores), disease occur strength data as well as the habitat factor data during disease occurrence, a statistical model underlying relationship between disease bio-statistical data or intensity data occurs and environmental factors can be constructed to predict the range, intensity or time of future disease occurrence [1].

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 101–108, 2013.

For example, environmental data of average temperature, precipitation and relative humidity are used to determine spreading areas of wheat scab epidemic in the irrigation areas, and ground meteorological data, upper air circulation and North Pacific sea surface temperature field factor can be effectively used in the Shanghai area of wheat scab occurrence degree forecast[2]. Statistical regression model is relatively simple and accurate, and still it is time-consuming and laborious, due to the requirement of fine and enough information of disease, environmental factors and their relationships.

In environmental conditions and physiological cycle, the reflectance spectrum of the crop follows certain distribution law. However, when subjected to disease invasion, the change of crop physiology and structure will result in the change if its reflection spectrum [3].Variation of its reflectance spectra during the disease occurrence can be used as indicators to predict or identify crop diseases For example, it demonstrates that crop spectral changes before and after the study onset are used to effectively identify the corn borer pest, wheat embroidery disease and the cotton wilt disease. Overall, the multi-spectral remote sensing prediction methods can accurately monitor the scope and area of crop pests and diseases, but not be able to predict earlier occurrence of crop diseases. After the occurrence of crop diseases, the changes of its physiological and ecological structure and reflectance spectra are not synchronized, but the slow responds. In fact, the predicted results of the multi-spectral remote sensing prediction methods show the late extent of disease occurrence, but not for early prediction in advance.

In essence, statistical models or multi-spectral remote sensing prediction methods the region to be predicted as an isolated, static unit of non-interaction with neighboring areas, and cannot give a reasonable explanation of driving mechanism that the spatial spread of disease runs from infected areas spread to neighboring areas.

In this paper, we first briefly survey the advantage and disadvantage of two traditional methods of disease occurrence forecasting, and focus on the natural essence of the cross-boundary diseases, and propose a complex network-based model to simulate the spatial spread of a new disease from a early source area of its out breaking to adjacent ones, and finally throughout the whole region. Our contribution is that the experimental results show the influence of spatial constraint on disease spread using our proposed model.

The following of the paper is organized as follows. Section 2 gives an emphasis of the spatial transition of cross-boundary diseases and propose a complex network-based model to simulate its spatial spread when a new and unknown disease outbreaks. Section 3 presents experimental results on real data, and finally, Section 4 makes a conclusion of the paper.

2 Modeling of Cross-Boundary Disease Spread

In previous section, the goodness and limits of two traditional methods are briefly surveyed. This section will focus on the characteristics of cross-boundary disease transmit ion , and further give a hypothesis to model and simulate the process of its spatial spread when a new and unknown disease outbreak.

2.1 Model Hypothesis

Agricultural cross-border disease usually first occurs in a planting area (the source of the disease area), and then gradually spread to the surrounding adjacent planting areas, finally throughout the region. There is a significant difference between cross-border agricultural diseases and conventional ones, which disease transmission only from the source of the disease area spread to nearby areas that are not infected. Additionally, considering that radius of disease spread is limited, whether the disease occurrence in non-infected areas appears or not depends on its spatial distance to the source of the disease. So for any new and unknown disease, a hypothesis is proposed, that is :(1) its propagation path is only from the pathogens areas to the adjacent non-epidemic areas, and whether the outbreak of the new epidemics in crops of adjacent non-epidemic areas depends on spatial spread radium of the epidemic; (2) crops are certainly infected within a radius of the epidemic spread, regardless of environmental factors, while not outside the radius;(3)subjectively, epidemic spread is not allowed to occur between adjacent non-epidemic areas to reduce the complexity of the network.

Based on this, the spatial spread of trans-boundary diseases is considered as network evolution problem, the complex network theory method is used to solve the problem. Given study area, crop plaques are abstracted into a graph node, the onset of disease from one planting unit spread to neighboring units, even edge between two nodes, in order to build a crop epidemic communication network.

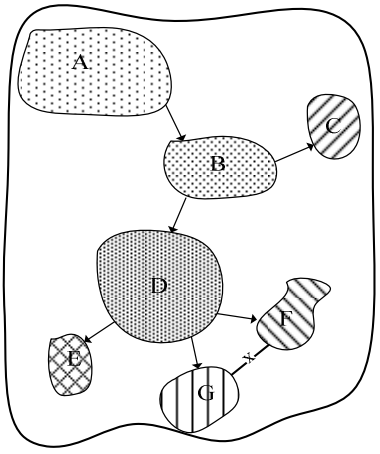


Fig. 1. The process of crop disease occurrence and spread. Polygons means different neighboring planting plaques.

To further introduce the process of construction of epidemic spread network in detail, an example of disease spread is given. Figure 1 indicates a crop planting area consisting of A~G seven polygon blocks. In the blocks, different/same crops grow, and are susceptible to the infection of a new epidemic. Supposing a new epidemic first appears in the A block, and then the epidemic spreads to the B block close to A

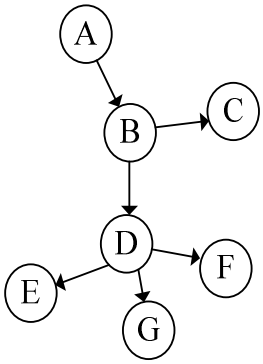


Fig. 2. Directed graph mapped into spatial spread of crop disease. the circles stand for nodes, while the arrows for the edges.

and within the spread radium, finally appear in all other blocks. By the limit of the model assumption, epidemic spread doesn't appear between the blocks of F and G, which are non-epidemic areas. The epidemic spread is abstracted to the network, which its nodes stands for the blocks and directed edge for the interaction of two nodes, and then a simple network consisting of seven nodes and six edges appears (shown in figure 2).

2.2 Evaluation Index

There is a great influence on network topology on the spread of the disease [4-6].Investigation of network topology characteristics helps to understand the disease propagation. The network structures, which meet random, uniform, small world or free scale distribution law are generated to investigate the robustness and vulnerability of different types of network on disease spread [6-8]. Network degree distribution exponent, clustering coefficient, node / edge betweenness, average path length are used to indicate the network topology characteristic [9-11]. Here two indexes of clustering coefficient and average path length are considered as the indicators of epidemic spread.

A clustering coefficient is a measure of degree to which nodes in a graph tend to cluster together. Average path length is a concept in network topology that is defined as the average number of steps along the shortest paths for all possible pairs of network nodes. It is a measure of the efficiency of information or mass transport on a network. Giving a disease spread network consisting of a set of vertices V and a set of edges E between them. The degree k_i for a vertex V_i is defined as the number of its immediately connected vertexes, E_i as the number of its immediately connected edges and local clustering coefficient (C_i) for V_i is then given by the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them. So for all vertexes, the average clustering coefficient (CC) can be obtained from the equation (2). Let d (v_i, vj) denote the shortest path between v_i and v_j. The average path length is calculated by the equation (3).

$$C_i = \frac{E_i}{k_i(k_i - 1)} \tag{1}$$

$$C_i = \frac{\sum C_i}{n} \tag{2}$$

$$APL = \frac{1}{n(n-1)} \sum_{i,j} d(v_i, v_j) \tag{3}$$

3 Experiment

3.1 Data Preparation

In this paper, all algorithms are encoded using vb.net language, and run inWin7 operation platform. The Study area of Beijing is located at latitude 39 º 28 '~ 41 º 05', longitude 115 º 24 '~ 117 º 30' with Plains of southeast, western and northern mountains, has a temperate continental climate.

In the study area, there are major land use types of grain field, woodland, orchard, and residents and facilities for agricultural land (vegetable).The land use of wheat data is abstracted from classified agricultural land in Beijing, and used to construct to the spread network of diseases, and simulate the response of diseases spread network of environmental elements, spatial elements and scale changes.

Classification of the land use of wheat in the suburb of Beijing are with the following data: the remote sensing data (4 m spatial-resolution panchromatic image of Landsat TM in March, 2007 and 100m spatial-resolution Beijing-1 satellite image in March, 2007), basic GIS data (1:10000 basic farmland of 1996, 1:100000 land use data, as well as 1:10000 Beijing administration, roads, water systems, residential and other data).To identify the land use type of wheat n the suburb of Beijing through remote image classification, data preprocessing is first performed. The process of data preprocessing includes projection transformation and convert geographic coordinates, geometric correction and image fusion. Projective transformation and conversion of geographical coordinates are done for all data, to convert to a variety of reference system to the uniform Beijing local coordinate system.

The Image geometric correction process is as follows: first crudely correct Beijing-1 satellite image and TM image geometry using1:50000 topographic map of the Beijing, and then further precisely correct them with truth ground GPS control points, in order to ensure the geometric correction accuracy of better than 1 pixel. The Beijing-1 satellite panchromatic image is high clear, but subject to the limitations of the single-band image, and so it is very difficult to distinguish its approximate surface features. TM image has multi-spectral bands characteristics of surface features, but the low resolution. Therefore, after fusion of the Beijing-1 satellite image and TM image, a high-resolution and multispectral image is obtained.

Based on the reference image (the fusion image), visual interpretation of surface features begins. In interpretation process, the following data of 1:1000000 basic farmland data of 1996, 1:100000 land use data of 2004 as well as 1:10000 Beijing administration, roads, water systems, residential and other data are taken as the priori knowledge. Clearly, the land use types of non-concern, such as cities, villages, water, roads, etc., are directly excluded, and but the type of wheat is outlined and specified attributes. The land use types not easy to determine are further confirmed through aerial photos and field investigation. Finally, the land use data of wheat (shown in figure 3) is used to map into the network which is characteristic of spatial spread of disease.

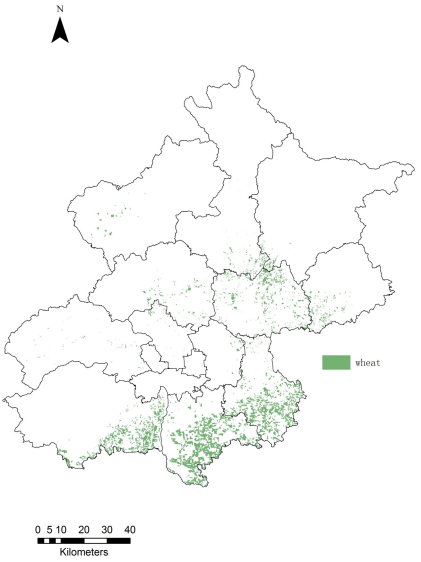
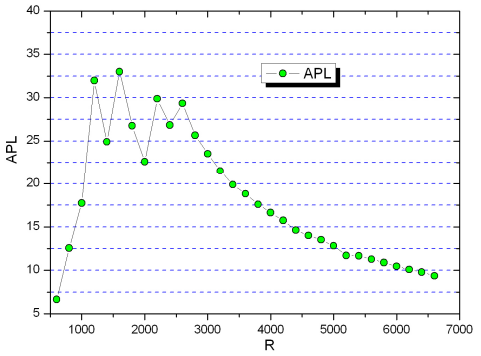


Fig. 3. The land use type of wheat

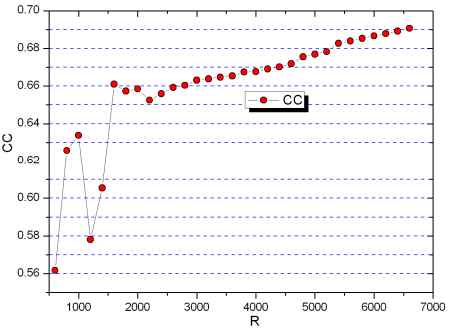
3.2 Topology Characteristics of Disease Spread

The land use of wheat extracted from remote sensing image of Beijing is mapped to an epidemic spread network. To test the influence of topology structure on epidemic spread, a new epidemic first randomly outbreak in a node, and then gradually transfers from the node to others. Considering epidemic spread radium heavily affect network structure, a variety of epidemic spread radium are set to generate the corresponding networks. For every generated network, its clustering coefficient and average path length are calculated (shown figure 4(a-b)). The results show that the network is not connected with the small radius of disease spread, and disease spread occurs only in a small amount of local connected subnets. With increasing of disease transmission radius, the network's clustering coefficient and average path length gradually show the topology features of a small world network that means when

spread radium exceeds a certain range, spatial segmentation is no longer a factor to inhibit epidemic spread. So an interesting conclusion can inferred from the above results, which spatial isolation can deter epidemic spread with small spread radium, while with the radium more than a certain range, the epidemic will quickly spread the whole region.



(a)



(b)

Fig. 4. The results of avarage path length (APL) and clustering coefficient of disease spread network, respectively

4 Conclusion

In this paper, we have introduced and discussed the problem of a crossing boundary crop disease spread, and construct a complex network-based model to simulate crop disease spread. We took wheat-growing distribution data of Beijing in 2007 by image classification, and further abstracted patches of wheat growing into network nodes, and finally constructed the disease spread network to analyze its topology characteristics. The results demonstrate that the network gradually shows the topology features of a small world network with epidemic spread radium increasing. Based on that, an interesting conclusion can be inferred that spatial segmentation may be a

reasonable treatment in control of a new epidemic spread with a small spread radium, while when the radium exceeds a certain threshold, spatial segmentation will not so work that the epidemic will quickly spread all over.

Acknowledgment. The authors are very grateful to those who collected the data of remote sensing images and GIS used in the paper.

The authors were supported financially by Shanghai Natural Science Foundation (No.11ZR1432700).

References

1. Nutter, F.W., Guan, J.: Quantifying alfalfa yield losses caused by foliar diseases in Iowa, Ohio, Wisconsin and Vermont. Plant Disease 86, 269–277 (2002)
2. Guan, J., Nutter, F.W.: Relationships between defoliation, spot pathosystem. Computers and Electronics in Agriculture 37, 97–112 (2002)
3. Crooks, W.T., Archer, D.J.: SAR observations of dry land moisture—towards monitoring outbreak areas of the brown locust in South Africa. In: Proceedings of IEEE International Symposium on Geosciences and Remote Sensing 2002, vol. 4, pp. 1994–1996 (2002)
4. Vittoria, C., Alessandro, V.: Epidemic modeling in metapopulation sysyems with hetergeneous coupling pattern: Theory and simulations. Journal of Theoretical Biology 251, 450–467 (2008)
5. Ball, F., Mollison, D., Scalia, T.G.: Epidemics with two levels of mixing. Ann. Appl. Proba. 7, 46–89 (1997)
6. Barthelemy, B., Bartar, M., Vespignani, A.: Dynamical patterns of epidemic outbreaks in complex heterogeous networks. Journal of Theoretical Biology 235, 275–288 (2005)
7. Cooper, B.S., Pitman, R.J., Edmunds, W.J., Gay, N.J.: Delaying the international spread of pandemic influenza. PLoS Med. 3, e12 (2006)
8. Colizza, V., Pastor-Satorras, R., Vespignani, A.: Reaction-diffusion process and metapopulation models in heterogeous networks. Nat. Phys. 3, 276–282 (2007b)
9. Newman, E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Physical Reviw E 64, 118 (2001)
10. Bodin, O., Norberg, J.: A network approach for analyzing spatially structured populations in fragmented landscape. Landscape Ecology 22, 31–44 (2007)
11. Albert, R., Barabasi, A.L.: Statistical mechanics of comlex network. Review of Modern Physics 74, 47–91 (2002)

Greenhouse Wireless Monitoring System Based on the ZigBee*

Minghua Shang[1], Guoying Tian[2], Leilei Qin**, Jia Zhao[1], Huaijun Ruan[1], and Fengyun Wang[1]

[1] S&T Information Engineering Research Center, Shandong Academy of Agricultural Sciences, Jinan, China
[2] Shandong Product Quality Supervision & Inspection Research Institute, Jinan, China
seqsoft@163.com

Abstract. Designed a WSN-based wireless monitoring system for greenhouse environment. The overall structure of the system is introduced. The design and realization of the monitoring node, the gateway node and the upper computer system are respectively described. Through the practical application in the greenhouse, it is proved that the system comprehensive performance is significantly better than the traditional greenhouse monitoring system.

Keywords: Greenhouse, Monitoring, WSN, ZigBee.

1 Introduction

The traditional greenhouse monitoring system is mainly based on the cable communication mode which has a series of questions such as complex wiring, difficult maintenance, inflexible deployment for sensor nodes and so on. These questions limit the popularization and application of greenhouse monitoring system to some extent. With the rapid development of modern information technology, there are many kinds of wireless communication technology such as WiFi,bluetooth,ZigBee and so on.WiFi and bluetooth etc cannot be widely used in greenhouse monitoring field because of its high cost and large power consumption etc shortcomings. As a brand new information acquisition and processing technology, the wireless sensor network based on ZigBee has a big scale, small volume, low cost, AD hoc network etc characteristics which has wide application prospects in agricultural environment monitoring field.

In view of the problems and shortcomings occurred in the current monitoring system of greenhouse environment, a wireless monitoring system for greenhouse is designed in this article which has designed low power consumption, low cost, flexible networking, friendly interface, convenient on-site and remote management. It is successfully applied.

* Supported by: Science and Technology Development Plan of Shandong Province, China (2011GGC02035).
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 109–117, 2013.

2 Overall Design

2.1 System Requirements Analysis

The environment in greenhouse has many features such as large temperature difference between daytime and nighttime, large air humidity, bad gas exchangeability, poor light intensity, strong soil acidity and so on. There are more crop species planted in greenhouse with dynamic change. The monitoring area is large and monitoring parameters are too many. In addition, the farmer is also very sensitive to whole cost and reliability of monitoring system. Through investigation and analysis, most of the current acquisition demand for greenhouse monitoring environmental parameter is mainly focused on six factors i.e. air temperature and humidity, soil temperature and humidity, light intensity, concentration of CO_2. In addition, a few greenhouses also need to collect EC value, pH value of nutrient solution and outdoor weather factors and other information. The quantity of sensor nodes can increase or decrease at random and the nodes can change its position without affecting the normal operation of the system according to crop growth, species replacement or greenhouse space structure change etc. The system is with intuitive interface, comprehensive analysis easy to use and low application cost etc.

2.2 Overall Structure

In accordance with the characteristics of greenhouse monitoring system and the above mentioned function demand, this paper integrates the wireless sensor network technology, ZigBee technology and embedded technology to design the wireless sensor network monitoring system based on ZigBee for greenhouse. The whole system hierarchy is shown in figure 1.

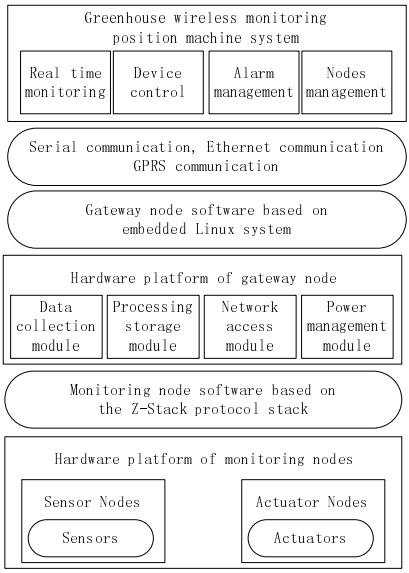


Fig. 1. Framework of the greenhouse wireless monitoring system

The system consists of monitoring node, gateway node and a host computer three layers. Monitoring node includes sensor node and actuator node deployed in greenhouse monitoring region which is automatically establish unified wireless sensor network through the ZigBee protocol. Each sensor node collects the real-time greenhouse environment data to the gateway node by multi-hop routing mode, the actuator nodes receives the real-time control orders from gateway node to control the fan etc actuating mechanism. The gateway node realizes the local communication through the serial port or remote communication through the Ethernet and GPRS etc modes which provide the sending support for monitoring data and control data. Host computer system provides a user interface to realize the interactive management operation between the user and system.

3 Monitoring Node Design

3.1 Hardware Design

Monitoring nodes are the base of greenhouse monitoring system which is the basic unit to information perception, executive control and network function of wireless sensor network. According to different division of labor, monitoring nodes are divided into the sensor node and actuator node which are introduced as follows.

1）Sensor nodes

The hardware design core of sensor node is microprocessor chip. The microprocessor of node finishes data acquisition, data processing, wireless communications etc functions in collaboration with the wireless transceiver module. The hardware structure diagram of wireless sensor node is shown in figure 2. The hardware design of wireless sensor node mainly considers the low cost, low power consumption, stability and reliability etc factors.

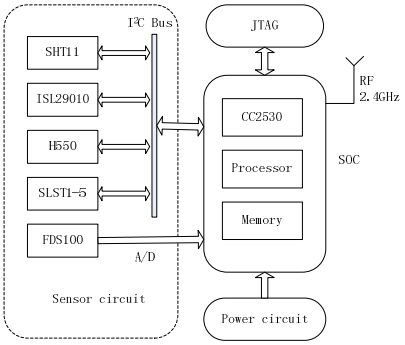


Fig. 2. The hardware structure of the sensor nodes

a）CC2530

Considering the cost and performance etc factors, CC2530 is selected. It integrates the microprocessor module and a wireless transceiver module in one single chip. CC2530 is introduced by United States TI companies for IEEE802.15.4 and ZigBee application. At present, it is also one of the most outstanding microprocessor among many ZigBee equipment product.

Its main features are as follows:

- enhanced high speed 8051 kernel;
- support the latest ZigBee 2007PRO protocol;
- support 2v-3.6v power supply range;
- 3 power management mode: the wake up mode 0.2mA, sleep mode 1uA, interrupt mode 0.4uA, with ultra low power consumption characteristics;
- high density integrated circuit.

The node design based on the CC2530 only needs few peripheral circuit to realize data acquisition and transmission which greatly improves the reliability and reduce the power consumption of the system.

b）Sensor

In the selection of sensor, it is requied that the sensor has higher precision and lower power consumption. It adopts 5 sensors in this design. Their technical parameters are as follows:

- SHT11 digital temperature and humidity sensor, detecting current 0.5mA, standby current 0.3uA, temperature accuracy ±0.5℃, humidity accuracy ±3.5%RH, I2C bus interface.
- ISL29010 digital light intensity sensor, detecting current 0.25mA, standby current 0.1uA, measurement accuracy ±50Lux, I2C bus interface.
- H550 digital CO_2 sensor, working current 15mA, accuracy ±30ppm, I2C bus interface.
- SLST1-5 digital soil temperature sensor, measuring current 1.5mA, standby current 1uA, measurement accuracy ±0.5℃, single bus interface.
- FDS100 simulation model of soil moisture sensor, working current 15mA, accuracy ≤3%, analog output signal.

2）Actuator nodes

The actuator node carries out switch control for indoor fan, sunshade etc equipment according to the control command from the host computer. The actuator node includes a drive circuit, but not includes the sensor circuit. The hardware structures of actuator node and sensor node are approximately same. The drive circuit of actuator node is mainly used to control solenoid valve associated with the actuator etc switch equipment which can output multichannel high-low level control signal. Data communication is of master-slave mode.

3.2 Node Software Design

The chip program of sensor node is based on the Z-Stack protocol. Its development environment is IAR 7.51A. Z-Stack is Zigbee protocol stack launched by TI company in April, 2007. It has been generally accepted and widely applied within the industry because of full support for Zigbee2006 and Zigbee PRO feature set and conforming to the latest intelligent energy standard. The protocol stack provides a protocol stack scheduler named operating system abstraction layer (OSAL). For developers, except being able to see the scheduling procedure, the specific implementation details of any other protocol stack operation are encapsulated in the library code. For the specific application development, the corresponding operation is finished by calling the API function interface of protocol stack, such as network device initialization, network configuration, network starting, collected data transmitting, control command receiving and so on to realize ad hoc network of wireless monitoring nodes distributed in several greenhouses. In addition, the node software designs flexible, convenient, dynamic configurable regular data collection, regular sleep and wake-up etc functions in order to further reduce the node power consumption.

4 Gateway Node Design

4.1 Gateway Node Hardware Design

The gateway node is the key device to realize the protocol conversion between wireless sensor network and external communication network. It not only has the function of data transmission, but also has equipment management function. When design the gateway node, follow the concept of modular design that the gateway system is divided into data collection module, processing/storage module, access module and power supply module.

The design is based on S3C2416 core board and establishes the hardware platform of gateway node of wireless sensor network. The hardware structure of gateway node is shown in figure 3.

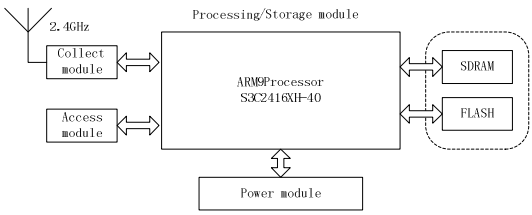


Fig. 1. The hardware structure of the gareway nodes

1）Data Collection Module

It is the coordinator node in the wireless sensor network to achieve the data collection and aggregation of greenhouse environment. In this design, the interface type between data collection module and processing/storage module is UART which communicate through serial.

2）Processing/Storage Module

It is the core module of gateway node. S3C2416 core board integrates Samsung S3C2416XH-40 processor based on ARM926EJ core which main frequency is 400MHz. In additio it also integrates 512MB DDR2 SDRAM and 128MB Nand Flash and provides abandunt peripheral equipment interfaces so as to furthest reduce the cost of the system development which is very suitable for high cost-performance and low power requirements of embedded equipment.

3）Access Module

It mainly adopts ethernet mode to connect the gateway to the external network. The core board integrates local high-speed ethernet chip LAN9220 of SMSC company. It realizes the Ethernet data transmission under the support of operating system. Network transformer uses HR601680 which main function is to match the impedance, enhance signal and realize voltage isolation etc. In addition, GPRS, as an optional way, uses MC37I module of Siemens company.

4）Power Supply Module

The module is responsible for the power supply of gateway node. The power module designed here has the function of hot plug and voltage conversion. The possible power supply mode comprises utility power, solar power and storage battery.

4.2 Software Platform Design of Gateway

Embedded Linux is a small operating system designed in accordance with the requirements for embedded operating system which consists of a kernel and some system modules customized as per requirements. Its kernel is very small and has the characteristic of multi-task and multi-process which is very suitable for transplantating to embedded systems. This paper transplants Linux2.6 kernel and associated drive on S3C2416 target platform and uses the open-source LwIP protocol stack instead of the TCP/IP protocol stack of Linux system and then designs the program on application layer of gateway node on the base of embedded Linux and LwIP. It mainly realizes two functions i.e. gateway node configuration through Web server and connecting Modbus serial communication link to ethernet through Modbus/TCP protocol.

1）Web Server Function Design

In the gateway configuration mode, the gateway node serves as the Web server, while the client is any one computer that connects to gateway RJ45 interface through the crosswire.

After gateway reset starting, operating system will start the Web service. The client sends out a request of HTTP GET method to the gateway through the browser. After receiving the request, the gateway judges the method field of message. If it is the GET method, it is the first request and returns the configuration information of Web pages and gateway embeded in the Flash out of the chip back to the client. After parameters configuration, the user clicks SUBMIT and the client sends out POST request to the gateway. The gateway erases the original configuration information in the Flash out of the chip and then writes the new information so as to ensure that the gateway configuration is not lost after resetting and the configuration information will take effect after gateway restarting.

2）Modbus/TCP Protocol Conversion Function Design

After reset starting, the gateway firstly performs a series of initialization and finally starts the Modbus server to realize the transmission between Modbus/TCP frame and Modbus RTU frame in the serial link. When a client inquires, it firstly sends out a connection request to 502 port of gateway, the gateway executes the interrupt service program to awake Modbus server in a wait state and creates the TCP connection, the client then sends a Modbus/TCP request frame and waits for a response. The gateway analyzes the frame to generate a query frame of Modbus RTU format to the serial link. If it receives RTU response frame in serial link, the frame is encapsulated into a Modbus/TCP reply frame which is sent to the Ethernet client and disconnect.

5 Design of Upper Computer System

The design is under VS.NET development environment and based on the SQL Server database and C# language to write the monitoring management software of greenhouse environmental information to complete the sensor node management and data management of greenhouse environment. Its main function is as follows:

1）Real Time Monitoring

The user can view the latest environment parameters as well as the status of fan, water pump etc control equipments in greenhouse field and can control and adjust in the current interface which is convenient for users to operate.

2）Historical Data

User can inquire monitoring data through many modes can also make a period of history data into curve to reflect the changes in greenhouse environment more intuitively.

3）Device Control

It includes two modes i.e. automatic control and manual control. In manual mode, the user can remotely control the switches of fan etc equipments. In the automatic mode, the system can automatically adjust the switches of fan etc equipments according to the environmental monitoring parameters.

4）Alarm Management

The user can define multilevel alarm conditions and can view the detailed information of set alarm. In an alarm condition, the user can specify the operation when alarming e.g. start alarm and turn on the switch of fan etc devices and send an alert notification etc.

5）Node Management

It includes the display and configuration of node ID, node location, sensor types and parameters, sampling period, running state, update time etc attributes. The user can master the working status of all monitoring node in field and discover equipment failure in time.

6 System Application

6.1 Node Deployment Solution

The system designed in this paper is applied in 1# greenhouse of Ji'nan Modern Agricultural Science and Technology Demonstration Park. It is placed 12 nodes in the vegetable cultivation area in the greenhouse where includes 10 sensor nodes and 2 actuator nodes. In addition, one gateway node is arranged in management area in the greenhouse. The air temperature and humidity sensor, the light intensity sensor, CO2 sensor and the corresponding sensor nodes are integrated into one whole body, and the soil temperature and humidity sensors are respectively connected to sensor node by a cable and the other end is inserted into the soil about 8cm, cable length 1.5～2m. Each sensor node is placed on a monitoring node position through the fixed supporting rod or rope hanging upside down mode and the height from the ground is generally about 1.2m. The sensor node uses one 1# batteries and the actuator nodes and gateway nodes use DC power supply.

6.2 System Application

The average communication distance between the nodes is about 20m, the nearest distance between monitoring node and gateway node is about 100m. Through installation and operation, after starting of gateway nodes, the average time required for node binding and self organizing network establishing is less than 1min. The sampling frequency setting scheme of sensor node is 2min for air temperature and humidity, 10min for soil temperature and humidity, 3min for light intensity and 30min for CO2 concentration. After completing data acquisition and transmission, each node will automatically enter sleep state until the next sampling cycle wakes up. Through actual test, the system can support the dynamic adjustment of sensor nodes. When add, remove the nodes or temporarily change the position of the nodes, the operation of whole wireless sensor network will not be affected. The host computer system can receive and display temperature and humidity, light intensity and CO_2 concentration etc environmental data from the sensor nodes in real time and can display real-time operating state of each node. When the collected environmental parameters exceed the alarm threshold, if the control mode is of automatic control, it will automatically start the corresponding mechanism according to the alarm processing rules to realize the automatic control of greenhouse environment.

References

1. Guo, W., Cheng, S., Li, R., et al.: Greenhouse monitoring system based on wireless sensor networks. Transactions of the Chinese Society for Agricultural Machinery 41(7), 181–185 (2010)
2. Zhang, R., Zhao, C., Chen, L., et al.: Design of wireless sensor network node for field information qcquisition. Transactions of the CSAE 25(11), 213–218 (2009)

3. Zhang, R., Gu, G., Feng, Y., et al.: Realization of communication in wireless monitoring system in greenhouse based on IEEE802.15.4. Transactions of the Chinese Society for Agricultural Machinery 39(8), 119–122 (2008)
4. Cui, X., Zhao, Z., Wang, C.: Field applications and design technologies of wireless sensor networks. National Defense Industry Press (2009)
5. Sun, H., Lu, S., Yang, S.: Research and implementation of the embedded web server based on Linux and the processor of S3C2410. Computer Applications and Software 24(2), 134–136 (2007)
6. Zhang, J., Cao, W., Wu, M., et al.: Porting Linux to S3C2410. Microcomputer Development 15(6), 142–144 (2005)
7. Wong, J., Zhang, H., Peng, D., et al.: On embedded ARM-based modbus/TCP protocal and its implementation. Computer Applications and Software 26(10), 36–38, 68 (2009)
8. Huang, Z., Deng, Y., Wang, Y.: Design course for ARM9 embedded system. Beijing University of Areonautics and Astronautics (2008)
9. Gan, Y., Wang, H., Chang, Y., et al.: Design of ZigBee gateway system based on ARM. Communications Technoloty 42(1), 199–201 (2009)
10. Li, W., Duan, C.: Experiment and Practice on ZigBee2007/PRO protocol stack. Beijing University of Areonautics and Astronautics (2009)

Application of an Artificial Neural Network for Predicting the Texture of Whey Protein Gel Induced by High Hydrostatic Pressure

Jinsong He[1] and Taihua Mu[2,*]

[1] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, 310058, China
hejinsong@mail.tsinghua.edu.cn
[2] Institute of Agro-Food Science & Technology, Chinese Academy of Agricultural Sciences, Beijing 100094, China
mutaihua@126.com

Abstract. The effects of high hydrostatic pressure (HP), protein concentration, and sugar concentration on the gelation of a whey protein isolate (WPI) were investigated. Differing concentrations of WPI solution in the presence or absence of lactose (0-20%, w/v) were pressurized at 200-1000 MPa and incubated at 30°C for 10 min. The hardness and breaking stress of the HP-induced gels increased with increasing concentration of WPI (12-20%) and pressure. Lactose decreased the hardness and breaking stress of the gel. Furthermore, these results were used to establish an artificial neural network (ANN). A multiple layer feed-forward ANN was also established to predict the physical properties of the gel based on the inputs of pressure, protein concentration, and sugar concentration. A useful prediction was possible, as measured by a low mean square error (MSE < 0.05) and a regression coefficient ($R^2 > 0.99$) between true and predicted data in all cases.

Keywords: Hydrostatic high pressure, Whey protein isolate, Gelation, Artificial neural network.

1 Introduction

Whey is a by-product of cheese manufacturing and contains about 13% protein by dry weight. Whey proteins possess outstanding physicochemical properties in gelation and binding, making them widely used as functional ingredients in many formulations of bakery, dairy, and sausage products [1]. The major constituents of milk whey protein are β-lactoglobulin (β-Lg), α-lactalbumin (α-La), bovine serum albumin (BSA), and immunoglobulins [2]. Five kinds of industrial whey proteins are currently produced, four of which have a whey protein concentration (WPC) classified into four grades according to the protein content (35%, 50%, 65%, and 70-90%). A whey protein isolate (WPI) contains >90% protein on a dry weight basis [3].

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 118–125, 2013.

One of the functional properties of whey proteins is gelation, which is induced by heat treatment [4-6], although other factors such as salt addition [7], acidification [8, 9], and enzyme treatment can cause gelation with or without heating [9-13]. Hydrostatic pressure has been shown to induce gelation in whey proteins under appropriate conditions [14-20]. The rheological properties of whey protein, i.e., storage modulus G' and loss modulus G'' [14], increase with increasing protein concentration. The gel strength also increases with increasing pressure holding time [14, 16, 18]. In addition, the contribution of intermolecular S-S bonds to the aggregation and gelation of whey proteins has been demonstrated by Tanaka *et al*. [21] and Kanno *et al*. [18].

In addition, artificial neural networks (ANNs) are among the most commonly used nonlinear techniques. An ANN is a mathematical algorithm whose structure and function is inspired by the organization and function of the human brain. The important property of a neural network is its ability to learn to improve its performance. ANN can handle nonlinear data and tend to produce lower prediction errors. ANN provides several advantages over conventional digital computations because of its faster data processing, learning ability and fault tolerance.

Controlling the texture of gel is important for the utilization of the gel in the food processing industry. In this context, new approaches are required to modulate and predict the physicochemical properties of the gel. In this present work, we studied the effect of pressure, protein concentration, and sugar concentration on the rheological properties of the pressure-induced gels formed from a WPI. The results should provide the basic information to modulate the texture of the gels by using the pressure, protein concentration, and sugar concentration as working parameters. Furthermore, a multilayer feed-forward neural network trained with an error back-propagation algorithm was employed to provide approaches to predict the physicochemical properties of the gel using these working parameters.

2 Materials and Methods

2.1 Materials

WPI powder from bovine milk was donated by the Central Research Institute of the Snow Brand Dairy Industry Co. (Saitama, Japan). This product contained 6.1% moisture, 89.8% protein, 1.8% ash, 1.3% lactose, and 0.5% lipids, and the fraction of individual whey protein was 74% β-Lg, 18% α-La, 6% BSA, and 2% immunoglobulins, according to the manufacturer. All the chemicals used were of analytical grade and were obtained from Wako Pure Chemical Industries (Osaka, Japan).

2.2 Preparation of the WPI Solution

A WPI solution (10-20%, w/v) was prepared in a 50 mM sodium phosphate buffer (pH 6.8), the lactose being added to make a final concentration of 0-20% (w/v).

Each WPI solution was poured into a Teflon tube (4 mL volume, 15 mm internal diameter and 22 mm depth) fitted with a screw cap and then subjected to 800 MPa pressure with a HR15-B2 hand-operated oil-pressure generator (Hikari Koatsu, Hiroshima, Japan). The temperature was maintained at 30°C. In each experiment, the indicated pressure was achieved within 1.5 min, held for 10 min, and then released to atmospheric pressure within 0.5 min. The pressurized sample was analyzed within 1 h of releasing the pressure.

2.3 Rheological Measurements

The hardness and breaking stress of each gel sample was measured at room temperature with a Fudoh rheometer operated under RT-2005DD software (Rheotech, Tokyo, Japan) as described elsewhere [18].

2.4 Modeling of the Neural Network

The multilayer feed-forward neural network has proven to be an excellent universal approximator of non-linear functions. In this work, a feed-forward neural network trained with an error back-propagation algorithm was employed using MATLAB (Version 2007a, Mathworks, Natick, MA). The Neural Network Toolbox was used to model the hardness and breaking stress of the gel. The input parameters chosen in this study were pressure, WPI concentration, and sugar concentration. Supervised learning was used to train this network. The predicted and desired output were compared with one another while the errors were calculated between the predicted and actual output. An error back-propagation algorithm was used to adjust the network weights. It used a Levenberg-Marquardt approach, in which the weights were changed in proportion to the value of the error gradient. The training iterations were stopped when the validation error reached a set minimum.

2.5 Statistical Analysis

Statistical analysis was performed using the package DeltaGraph, Version 5 for the Mac (SPSS Inc., USA). In order to visualize the results of multiple experiments, the average and the standard deviation were calculated.

3 Results and Discussion

3.1 Effects of Pressure and Protein Concentration

The WPI solution at a concentration of 10-16% formed a gel at 600 MPa, while the 20% WPI solution formed a gel at 400 MPa. The gels formed from the 20% WPI solution at 400 MPa and from the 16% solution at 600 MPa were soft and translucent in appearance.

The gels formed from 10% WPI solution under 600 MPa, and 12% WPI solution under 400 MPa were too soft to measure texture properties using the rheometer.

Figure 1 shows the hardness and breaking stress of those gels that were sufficiently firm formed from 12 to 20% WPI at different pressures. The hardness and breaking stress of the WPI gels increased with WPI concentration, increasing from 12 to 18% at a constant pressure and, similarly, those of each solution at a constant concentration rose as hydrostatic pressure was increased from 600 to 1000 MPa. In addition, no water was expelled from the gel matrix that had been induced from these WPI samples under pressure during the measurement of hardness and breaking stress.

The gel-forming ability of a protein depends on a critical balance between the attractive and repulsive forces of the protein molecules [22]. Hydrogen bonds, electrostatic bonds, hydrophobic interaction, and intermolecular disulfide linkages may play a role in gel formation in protein solutions [22, 23]. Under high pressure, β-Lg unfolds, resulting in the exposure of its free SH group [21, 24]. Unfolded β-Lg can interact, through SH/SS interchange reactions, with proteins containing disulfide bonds, *e.g.* α-La, BSA, and β-Lg [19, 25]. Regarding pressure-induced gelation of a WPI solution, the formation of intermolecular S-S bonds between proteins has been confirmed [18, 25], and it is likely that the gel is mainly formed by the cross-linking of intermolecular S-S bonds [25].

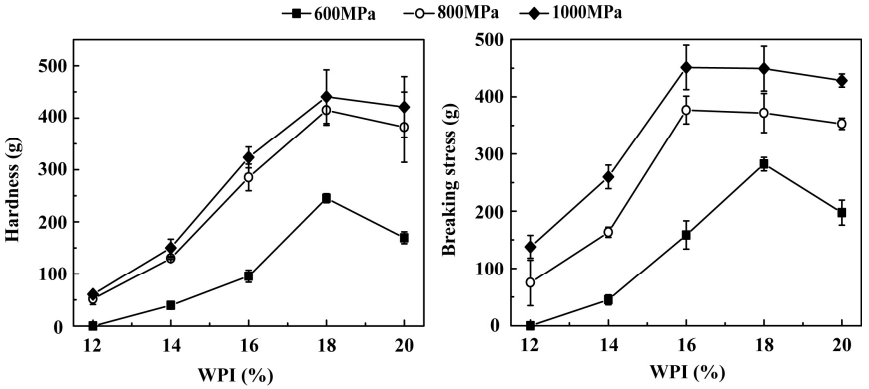


Fig. 1. Effect of pressure and WPI concentration on the hardness, and breaking stress of pressure-induced gels from WPI. WPI (12~20%, w/v) was pressurized at 600, 800, or 1000 MPa and 30°C for 10 min. Bars show the standard deviation for three measurements with different gel samples.

3.2 Effects of Sugar Concentration

To study the effect of the sugar concentration (2% to 20%) (w/v) on gels prepared from WPI in a phosphate buffer by pressurizing at 800 MPa and 30°C for 10 min, the lactose content, the main sugar contained in bovine milk, was varied. Figure 2 shows the changes in rheological properties as a function of the lactose content. The

hardness and breaking stress of the gel decreased with increasing concentration of lactose. This indicated that the sugars weakened the intermolecular interactions of the protein, seemingly preventing the cross-linking between proteins.

Sugar is preferentially excluded from the protein domain in an aqueous mixture of protein and sugar. This exclusion effect of the sugar can affect the surface tension of the water and minimize the protein-solvent interface [17, 26, 27]. As a result, the sugar protected the protein from unfolding and subsequent aggregation under pressure [28]. Sugars can also stabilize a protein against pressure-induced denaturation [17, 28]. The presence of 5% sucrose in 2.5% β-Lg decreased the protein unfolding, slightly increasing the rate of reversibility of aggregates by pressurization at 450 MPa [28]. The gelation of protein is influenced in a complex manner by the processes of protein denaturation and aggregation [22, 29]. Sugars decrease the degree of intermolecular S-S bonding of proteins [25], and restrain phase separation during the gelation of WPI under high pressure.

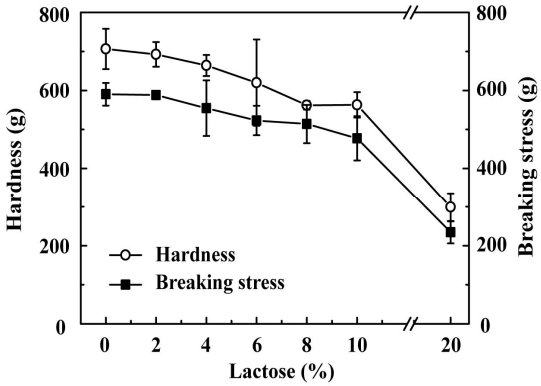


Fig. 2. Effect of lactose concentration on the hardness and breaking stress of a pressure-induced gel from WPI. WPI (20%, w/v) and the indicated concentration of lactose were dissolved in a 50 mM sodium phosphate buffer at pH 6.8, and the mixture was pressurized at 800 MPa and 30°C for 10 min. Bars show the standard deviation for three measurements with different gel samples.

3.3 Evaluation of Model Predictability

The ratio of the explained variation to the total variation, R^2, reflects the degree of fit for the mathematical model. The closer the value is to 1, the better the model fits the actual data

$$R^2 = 1 - \frac{\sum_{i-1}^{n} (y_i - y_{di})^2}{\sum_{i-1}^{n} (y_{di} - y_m)^2} \tag{1}$$

where n is the number of points, y_i is the predicted value obtained from the neural network model, y_{di} is the actual value, and y_m is the average of the actual values.

MSE is another important index to show the degree of fit of the model. It is calculated using the following equation.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_{di})^2 \tag{2}$$

In this work, the range of three independent variables (pressure, WPI concentration, and sugar concentration) for building the neural network was set. The input matrix and the properties of gel are shown in Figs. 1 and 2. To understand the generalization capacity of the network, 19 input values were divided into three sets: 13 values for the training set and three values each for the validation and testing set. Figure 3 shows the plot of the predicted values and actual values of the hardness and breaking stress of the gels as well as R^2 and MSE. The trained network gave a R^2 of 0.998 and a MSE of 0.047. The R^2 values of the training, the validation and the testing sets were 0.998, 0.999 and 0.999, while the MSE values of the three sets were 0.047, 0.042 and 0.045, respectively. The network could predict the properties of the gel within a range of ±7.8% of the actual value. The R^2, MSE and prediction range indicated a good agreement between the predicted value of the neural network model and the actual value, which also confirmed good generalization of the network.

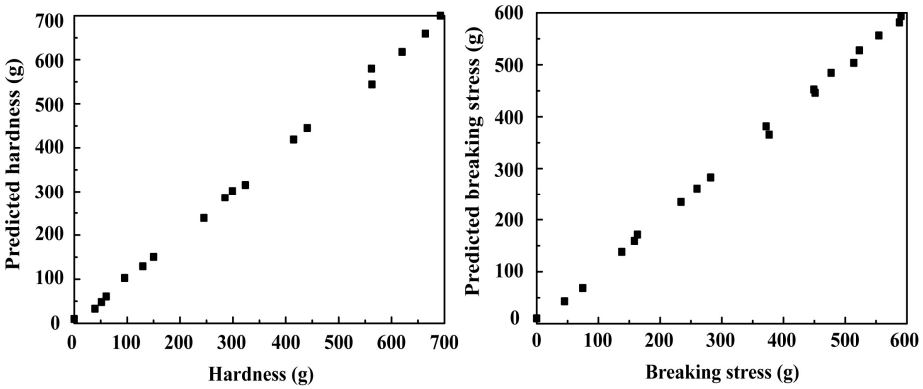


Fig. 3. Correlation between the predicted values of the neural network and the actual values for the hardness and breaking stress of a pressure-induced gel from WPI. The input matrix and the properties of gel as the output matrix are shown in Figs. 1 and 2.

4 Conclusion

The hardness and breaking stress of pressure-induced gels from WPI increased with increasing WPI concentration (12-18%) and hydrostatic pressure, and decreased with increasing lactose concentration (0-20%). The ANN provided a model to search for the non-linear nature between induced conditions and rheological properties in an efficient manner. The trained network gave an R2 value of 0.999 and an MSE value of 0.047, which implied a good agreement between the predicted values and the actual values for the hardness and breaking stress of the gels, and confirmed good generalization of the network.

Acknowledgments. The project was supported by the Zhejiang Key Scientific and Technological Innovation Team Project (2009R50001), the Department of Education of Zhejiang Province (Y201018520), and the Fundamental Research Funds for the Central Universities.

References

1. Kinsella, J.E., Whitehead, D.M.: Proteins in whey: chemical, physical, and functional properties. Adv. Food Nutri. Rese. 33, 343–438 (1989)
2. Eigel, W.N., Butler, J.E., Ernstrom, C.A., Farrell, H.M., Harwalkar, V.R., Jenness, R., Whitney, R.M.: Nomenclature of proteins of cow's milk: fifth revision. J. Dairy Sci. 67, 1599–1631 (1984)
3. Mulvihill, D.M.: Production, functional properties and utilization of milk protein products. In: Fox, P.F. (ed.) Advanced Dairy Chemistry, vol. 1, pp. 369–402. Elsevier Applied Science, London (1994)
4. Mulvihill, D.M., Donovan, M.: Whey proteins and their thermal denaturation-a review. Irish J. Food Sci. Technol. 11, 43–75 (1987)
5. Mulvihill, D.M., Kinsella, J.E.: Gelation characteristics of whey proteins and β-lactoglobulin. Food Technol. 41, 102–111 (1987)
6. Errington, A.D., Foegeding, E.A.: Factors determining fracture stress and strain of fine-stranded whey protein gels. J. Agric. Food Chem. 46, 2963–2967 (1998)
7. Barbut, S., Foegeding, E.A.: Calcium-induced gelation of preheated whey protein isolate. J. Food Sci. 58, 867–871 (1993)
8. Mleko, S., Foegeding, E.A.: pH induced aggregation and weak gel formation of whey protein polymers. J. Food Sci. 65, 139–143 (2000)
9. Britten, M., Giroux, H.J.: Acid-induced gelation of whey protein polymers: effects of pH and calcium concentration during polymerization. Food Hydrocoll 15, 609–617 (2001)
10. Faergemand, M., Qvist, K.B.: Transglutaminase: effect on rheological properties, microstructure and permeability of set style acid skim milk gel. Food Hydrocoll 11, 287–292 (1997)
11. Faergemand, M., Murray, B.S., Dickinson, E.: Cross-linking of milk proteins with transglutaminase at the oil-water interface. J. Agric. Food Chem. 45, 2514–2519 (1997)
12. Wilcox, C.P., Clare, D.A., Valentine, V.W., Swaisgood, H.E.: Immobilization and utilization of the recombinant fusion proteins trypsin-streptavidin and streptavidin-transglutaminase for modification of whey protein isolate functionality. J. Agric. Food Chem. 50, 3723–3730 (2002)
13. Wilcox, C.P., Swaisgood, H.E.: Modification of the rheological properties of whey protein isolate through the use of an immobilized microbial transglutaminase. J. Agric. Food Chem. 50, 5546–5551 (2002)
14. Van Camp, J., Huyghebaert, A.: A comparative rheological study of heat and high pressure induced whey protein gels. Food Chem. 54, 357–364 (1995)
15. Van Camp, J., Huyghebaert, A.: High pressure-induced gel formation of a whey protein and haemoglobin protein concentrate. L.-Wiss. und-Technol. 28, 111–117 (1995)
16. Van Camp, J., Feys, G., Huyghebaert, A.: High-pressure-induced gel formation of haemoglobin and whey proteins at elevated temperatures. L.-Wiss. und-Technol. 29, 49–57 (1996)

17. Dumay, E.M., Kalichevsky, M.T., Cheftel, J.C.: Characteristics of pressure-induced gels of β-lactoglobulin at various times after pressure release. L.-Wiss. und-Technol. 31, 10–19 (1998)
18. Kanno, C., Mu, T.-H., Hagiwara, T., Ametani, M., Azuma, N.: Gel formation from industrial milk whey proteins under hydrostatic pressure: effect of hydrostatic pressure and protein concentration. J. Agric. Food Chem. 46, 417–424 (1998)
19. Kanno, C., Mu, T.-H.: Gel formation of individual milk whey proteins under hydrostatic pressure. In: Hayashi, R. (ed.) Trends in High Pressure Bioscience and Biotechnology, pp. 453–460. Elsevier Science B.V. (2002)
20. Keenan, R.D., Young, D.J., Tier, C.M., Jones, A.D., Underdown, J.: Mechanism of pressure-induced gelation of milk. J. Agric. Food Chem. 49, 3394–3402 (2001)
21. Tanaka, N., Tsurui, Y., Kobayashi, I., Kunugi, S.: Modification of the single unpaired sulfhydryl group of beta-lactoglobulin under high pressure and the role of intermolecular S-S exchange in the pressure denaturation. Inter. J. Biolog. Macromol. 19, 63–68 (1996)
22. Ziegler, G.R., Foegeding, E.A.: The gelation of proteins. Adv. Food Nutr. Res. 34, 203–298 (1990)
23. Clark, A.H., Ross-Murphy, S.B.: Structural and mechanical properties of biopolymer gels. Adv. Poly. Sci. 83, 57–192 (1987)
24. Botelho, M.M., Valente-Mesquita, V.L., Oliveira, K.M., Polikarpov, I., Ferreira, S.T.: Pressure denaturation of beta-lactoglobulin: Different stabilities of isoforms A and B, and an investigation of the Tanford transition. FEBS 267, 2235–2241 (2000)
25. He, J.-S., Azuma, N., Hagiwara, T., Kanno, C.: Effects of sugars on the cross-linking formation and phase separation of high pressure induced gel of whey protein from bovine milk. Biosci. Biotechno. Biochem. 70, 615–625 (2006)
26. Timasheff, S.N.: The control of protein stability and association by weak interactions with water: how do solvents affect these processes? Annu. Rev. Biophys. Biomol. Struct. 22, 67–97 (1993)
27. Lee, J.C., Timasheff, S.N.: The stabilization of proteins by sucrose. J. Biol. Chem. 256, 7193–7201 (1981)
28. Dumay, E.M., Kalichevsky, M.T., Cheftel, J.C.: High-pressure unfolding and aggregation of β-lactoglobulin and the baroprotective effects of sucrose. J. Agric. Food Chem. 42, 1861–1868 (1994)
29. Van Camp, J., Messens, W., Clément, J., Huyghebaert, A.: Influence of pH and sodium chloride on the high pressure-induced gel formation of a whey protein concentrate. Food Chem. 60, 417–424 (1997)

The Classic Swine Fever Morbidity Forecasting Research Based on Combined Model

Yi Liang and Shihong Liu*

Key Laboratory of Digital Agricultural Early-warning Technology,
Ministry of Agriculture, Beijing, P.R. China 100081
ryuui@126.com, lius@mail.caas.net.cn

Abstract. This paper proposes the combined forecasting model which study on the classic swine fever (CSF) morbidity, using the forecasting results of ARIMA and GM (1, 1) model as the inputs of the majorizing BP neural network. Analyzing the monthly data from 2000 to 2009 and the accuracy of the forecasting results is 97.379%, more accurate and more steady than traditional methods. This research provides efficient Analytical tools for animals' diseases forecasting work, and can provide reference to other animal diseases

Keywords: Combined Model, ARIMA, GM (1, 1), GA, BP neural network.

1 Introduction

Classic Swine Fever(CSF) as one of the A level legal animal diseases of OIE, being the main control object to swine epidemics, causes a huge impact on the aquaculture industry in China[1]. Animal diseases prediction plays an important role in human health protection mechanism, and animal diseases outbreaks have complexity, therefore it has higher demands for modern society's forecasting methods.

Traditional forecasting methods have a relatively large difference in prediction accuracy, and this paper proposes a combined model which takes the optimized BP neural network by genetic algorithm (GA) as the carrier and the forecasting results of ARIMA mode and GM(1,1) model as the inputs. ARIMA model and grey forecasting model have good accuracy in data prediction, and the forecasting results after initial process by these two models is the input of optimized BP neural network, then construct the CSF morbidity forecasting model based on combined model, making empirical study aiming to the gathered data about morbidity and livestock on hand from January 2000 to June 2009(Data Source: official veterinary bulletin from MOA).

Defined the amount breeding stock at end of year as N, monthly occurrence number as M, so the formula of morbidity defined as follow:

$$Morbidity = \frac{M}{N} \times 10000\% \tag{1}$$

* Corresponding Author, Address: Agricultural Information Institute, Chinese Academy of Agricultural Sciences, 100081, Beijing, P. R. China, Tel: +86-10-82103075, Fax: +86-10-82103075.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 126–132, 2013.

10000% means the occurrence number per ten thousand, meanwhile it facilitates calculation.

2 ARIMA Model

ARIMA model that Autoregressive Integrated Moving Average, the Model regards the time series of predicted object as the variable depends on time T, and describes this time series approximately by mathematical model. We will describe the predicted object's development continuity from past value and present value after the autocorrelation of this set of random variables being identified, and then we can forecast the practical data [2].

Because the predicted model' s time series of ARIMA model is steady random series whose means is zero, so original data must be steady by first, second or natural logarithm according to their liner relation.

Pick the January 2000 to May 2008 as the model constructive data, and June 2008 to June 2009 as the verification data. Verifying and selecting the fittest model by AIC and SC criterion, and finally confirmed the model is ARIMA (1, 0, 1).

The expression of model is:

$$y(t) = 0.001022 + 0.812276\,y(t-1) + u(t) - 0.164883\,u(t) \tag{2}$$

3 Grey Model GM (1, 1)

Gray forecast model is a predicted method which builds mathematical model using little and incomplete information [3]. When we solve the practical problem using operation knowledge, formulate development strategies and policies, or make decision of vital problem, we should predict the future in a scientific way. Gray model is systemic theory based on gray information (information is incomplete, inadequacy, non-unique) in small sample, and it can provide a new solution to the problems which have complicated factors and ambiguous operative mechanism.

Give a macro forecasting data series:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(N)\} \tag{3}$$

Use AGO operator after time cumulative:

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(N)\} \tag{4}$$

If $x^{(1)}$ satisfied the first order ordinary differential equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \tag{5}$$

a is called develop gray number and it reflects the develop trend of original data and predictive data; *u* is called Endogenous control grey number, which is constant input to system and reflects the relations of data variation. This equation satisfied the initial conditions:

$$x^{(1)} = x^{(1)}(t_0) \text{ when } t = t_0 \tag{6}$$

and the answer is:

$$x^{(1)}(t) = \left[x^{(1)}(t_0) - \frac{u}{a} \right] e^{-a(t-t_0)} + \frac{u}{a} \tag{7}$$

And OLS is:

$$\hat{U} = \begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} = (B^T B)^{-1} B^T y \tag{8}$$

Put estimation value into the response time equation:

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}} \right] e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}} \tag{9}$$

And finally get the response time equation:

$$X(k+1) = 0.631815e^{0.001776k} - 0.631671 \tag{10}$$

Calculate the fitted value $\hat{x}^{(1)}(i)$, and restore the calculated result by reverse subtraction

$$x^{(0)}(i) = x(i) - \hat{x}^{(1)}(i-1)(i = 2, 3, ..., N) \tag{11}$$

4 The Optimized BP Neural Network Based on GA

Genetic Algorithm (GA) is applied to neural network widely, mainly due to the GA having a strong global search capability in a complex, polymorphism and continuous space, which can help neural network optimize its network structure and parameters. The most important operations in GA are: selection, crossover and recombination, mutation. Selection is to fix individuals for crossover and recombination and decide how many generations will be produced by these individuals, and selection always includes roulette methods, tournament methods and so on; crossover and recombination is a significant link to improve the population's qualities, which can fix new individuals combining the parents' genetic mating information, and crossover and recombination always includes real value reorganization and binary crossover; mutation is a change that individuals after crossover and recombination affects population by small probability transformation, and individuals can revolute via mutation, getting higher quality individuals, and it always includes real value mutation and binary mutation[4].

The process of GA is as follow: 1.generate the initial population and code them; 2. Analyze the fitness of individuals, and if individuals satisfies optimized principle then output the best individual with its parameters and end; else go next step; 3.select individuals by fitness and save the highest fitness one; 4.apply the cross operator on whole individuals to produce new generation; 5.apply the mutation operator on individuals of population to adjust the structure and build new individuals; 6.the individuals after selection, crossover, mutation constitute next generation and repeat 2.

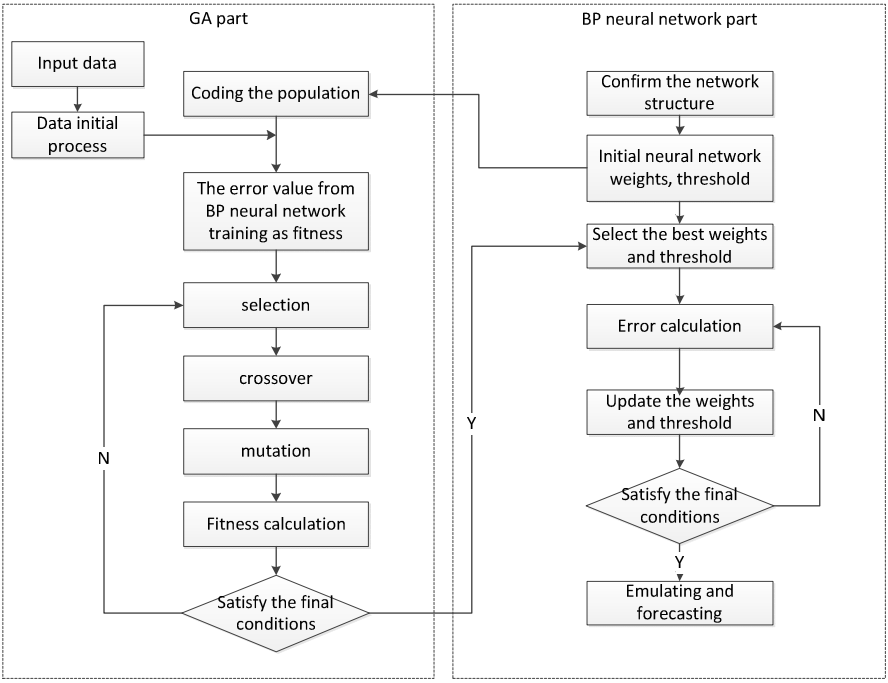


Fig. 1. Optimize the BP neural network by Genetic Algorithm

In the optimizing process of neural network, we can confirm the network structure and parameters via the global search capability of GA. We can rough adjust the network at the initial construction and get an approximate optimal solution. Meanwhile using the LM algorithm to adjust the network meticulously, in order to avoid GA algorithm sinking into a complicated iteration, which cost more time and space. GA is regarded as a decision algorithm not optimization algorithm and we can achieve the goal using rapidity and high efficiency of LM algorithm. This way balances the complexity, high efficiency and generalization of whole neural network [5].

5 Combined Model

In the practical prediction, due to the model mechanism and starting point being different from others, there are different forecasting methods upon the same problem. Different methods provide different information and forecasting accuracy. If we abandon some low accuracy methods, we will lose some useful information. Therefore, we propose a more scientific way: combining different methods in a proper way, and form the combined model methods and it is propitious to synthesize useful information from all kinds of methods and improve forecast accuracy [6].

The model in this paper contains ARIMA model, grey model and optimized BP neural network model. ARIMA model based on the time series, and analyze the data using statistic, so that we can dig internal statistic relation from morbidity data. Grey model has a good capacity to incomplete information. Because the original data has data hollow space, then grey model can simulate the whole incidence trend well and form a creditable forecast process. BP neural network is the most popular A.I. model, and has strong functions at pattern recognition and approximation of function. Take the BP neural network as the carrier of combined model can achieve a higher accuracy and save more time and space cost.

Owing to the data quality, even though the forecast result of single model has directive significance, it still can't up to the scratch [7]. Take optimized BP neural network as the carrier of combined model, and change the inputs of neural network. The forecast result of ARIMA and grey model will be the inputs and improve the inputs' quality, which improves the efficiency, quality and produce a better result.

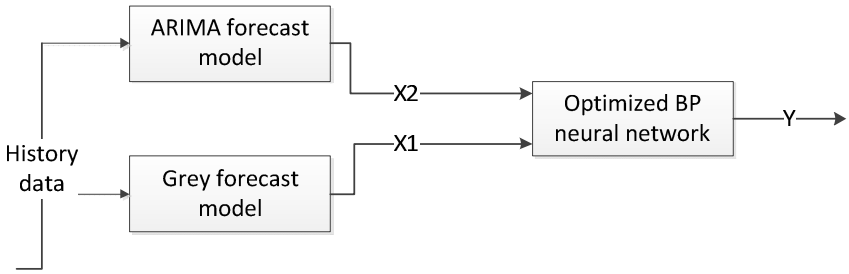


Fig. 2. Combined model structure. The results of ARIMA and GM (1, 1) as the input of BP neural network.

Combined model process is as follow:

1). Select data from January 2000 to May 2008 as the model constructive data, June 2008to June 2009 as the model verification data. Build the GM (1, 1) model and define the forecast value as X1; build the ARIMA (1, 0, 1) model and define the forecast value as X2.

2). X1 and X2 are the inputs of neural network, the real data X is the output, and construct the input-output set ((X1, X2), X). The hidden layer of neural network has 10 neurons. Inputs layer has 2 neurons includes ARIMA forecast value and grey model forecast value. Output layer has 1 neuron of real data. The configuration of GA

is as follow: the size of population is 10, evolution generation is 15, cross probability is 0.5, and the training process cost 14 iterations.

3).Optimize the neural network structure and configuration by GA, and achieve the forecast data Y, and verify the Y.

Compare the Y and real data X, broken line is the real data, and solid line is forecasting data. We can find that there is a little difference between real data and forecast data in the beginning of the graph, but at the posterior segment, the data error is almost negligible.

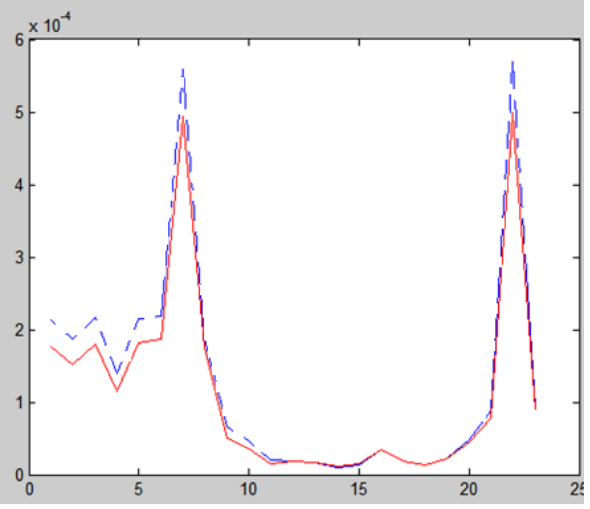


Fig. 3. Forecasting results analysis. The red solid line is the forecasting data and blue broken line is real data. The date of the result data range from June 2008to June 2009.

Table 1. Different model comparison

Model	Accuracy (%)	Mean error (%)
Optimized BP model	97.379	2.261
BP model	96.191	3.809
ARIMA	91.91	8.09
GM(1,1)	93.88	6.12

Compare the forecast value with the real value, and the forecast range is the morbidity from June 2008 to June 2009. The mean error of forecast value without optimization is 3.809%, and optimized model is 2.621%, decrease 1.188% than non-optimization model.

6 Conclusion

The combined model of this paper takes the BP neural network as the carrier, integrates the vantage of ARIMA model and GM (1, 1), and increases the accuracy of data process. Due to the optimizing by GA, strengthen the global search capability of BP neural network, avoiding the problem of traditional BP network being easily sinking into local minima, and can play the role of combined model well.

Using the combined model for animal disease forecasting application, the result proves that it is feasible. The initial data process makes use of ARIMA model's steady process capability to non-stationary, and the long term forecasting trend of grey model, and can try to apply it to the practical case. With the data being richer, there is still a large room for us to adjust the model, improve the accuracy so as to get more precise forecast value and have stronger directive significance.

References

1. Hongming, X.: Severe Animal Diseases and Their Risk Analysis. Science Press, Beijing (2005)
2. Mengquan, W., Kai, Z.: Hand-Foot-Mouth Disease Analysis and Prediction in Shandong in 2009 Based on the ARIMA Model. Ludong University Journal (Natural Science Edition), 71–75 (2011)
3. Daocai, C., Yanfang, T., Tuo, G., Miao, Y., Zheng, L., Zongzheng, M.: Predication of Irrigation Water Use Using Parallel Gray Neural Network. Transactions of the Chinese Society of Agricultural Engineering, 26–29 (2009)
4. MATLAB Chinese Forum, 30 Case Studies of MATLAB Neural Network, p. 286. BeiHang University Press, Beijing (2010)
5. Dong, Z., Kaiyuan, C.: A Genetic-Algorithm-based Two-Stage Learning Scheme for Neural Networks. Journal of System Simulation, 1088–1090 (2003)
6. Kaiyi, M.: Combination Forecast in Weight Determination and Application. Chengdu University of Technology, Chengdu (2007)
7. Ning, L.: Traffic Flow Forecasting Method Based on Combination Mode. East China University of Science and Technology, Shanghai (2010)

Application and Research of Man-Machine Interface and Communication Technique of Mobile Information Acquisition Terminal in Facility Production

Jinlei Li[1,*], Xin Zhang[1], Quanming Zhao[1], Wengang Zheng[2], Changjun Shen[2], and Zhipeng Shi[2]

[1] Beijing Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China
[2] Key Open Laboratory for Agricultural Information Technology of Ministry of Agriculture, Beijing 100097, China

Abstract. As long as the deepen application of the Object Network Technology in the facility construction, a hot topics in research is coming out, that is how to achieve mobile operation through convenient and high-efficiency multiple resources utilization, in order to increase the proportion of the efficiency and output for the high tech facility in manufacture. The mobile collection terminals could display, transfer and collect the information from wireless sensor network high efficiency. This terminal will play an important role in the Object Network Technology application system, comparing to the others, it has such advantages of the good man-machine interaction (MMI) and reliable communication technology. We develop a mobile information collection terminal basic on the wireless sensor network, the information collection technology adopts the APC240 module to realize the wireless sensor data collection by the MODBUS Agreement; adopt wireless communication module to transfer the GPRS data; MMI realize the screen touch operation, also transplant the uC/OS+GUI making the operation more simply and understandably. This system is much convenience for more rural people to operate in daily work and with more application in agricultural construction in the future.

Keywords: facility production, man-machine interface, communication technique, mobile acquisition terminal, GPRS.

1 Preface

In resent years ,the Internet of things technology and equipment to be promoted and applied in the field of facilities agriculture, including the collection of temperature, humidity, illumination and other environmental information, to predict scientifically

* Fund Projects.
The National 863 project : The Accurate Control Technology and Products of Agriculture Efficient Water Use (2011AA100509).
The science and technology program of Beijing: The Integration of Object Network Technology in Agriculture (D111100001011002).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 133–140, 2013.

though network transmission, to improve the agricultural comprehensive benefit through helping farmers to cultivate scientifically, the combination of Internet of things and facilities agriculture is a present direction of advanced facilities agriculture[1,2]. The following problems are in the specific application process: A The ways of agriculture acquirement is single and difficult to achieve the mobile operation, B The agricultural environment information transmission technology is not flexile enough, the routing is complex and maintaining is difficult, C Most man-machine interfaces (MMI) on control cabinet are not friendly that additional training is needed and unnecessary loss caused by improper operation often occurs[3]. Developing mobile information collection terminal can make man-machine interfaces and agriculture facilities combined and On-side environmental control equipment and remote monitoring center connected.

The communication technology plays an important role in the facilities agriculture, which including wired communication and wireless communication. It's known that facilities agriculture has the characteristics of object diversity, broad region, remote distribution, etc. And the communication condition is relatively backward, the last mile bottleneck phenomenon highlights in the agriculture information and communication. The wiring of wired communication is difficult in facilities agriculture, and sometimes can not be achieved, it will require a considerable investment if using the wired mode, and the distance of wired communication is relatively short, that unable to achieve remote monitoring. In resent years, the use of wireless communication has became increasing prevalent in facilities agriculture, such as distributed greenhouse monitoring controlling system based on Bluetooth technology, the monitoring system based on GSM and wireless sensor network, the above systems have good equipment interoperability, strong anti-jamming capability and good real-time performance,etc.,but some wireless communication technologies have poor stability, low reliability, high loss of power and lack transmitted distance, etc..[6].

To compare the advantage and disadvantage of WSN and GPRS communication technology, we combine with the advantages of good real-time performance low power loss from WSN technology and far transmission distance from GPRS technology, also covering shortages of each other, through designing a communication system being suitable to the Agricultural environment work. The operators of mobile information collection terminal is agricultural employee, it request a simple operation interface, also with the instruction function to make the communication between operator and equipment more smoothly which increase the efficiency. We state a MICT with low power loss, low cost, stabile communication and simple operate through the combination of Object internet communication and facilities agriculture to solve the disadvantages in daily works, such as complex control, high cost for cable arrangement and poor mobile information collection performance.

2 System Structure and Function

Regarding to daily application, the structure of this system is designed basic on agricultural wireless sensor network facilities for agricultural production mobile information collection system. Essentially, the system structure is usually shown as that in Figure 1.

The typical working way of this system is: A variety of wireless sensor nodes are laid in the collection area according to the needs of farmers, each node could collect the different agro-environmental information through the sensor probe. At the same time, a perception network (WSN) shall rapidly be constructed through self-organization between the nodes. The sensor data collected by the sensor nodes shall be transmitted to the mobile information collection terminals by multi-hop relay. Communication between the mobile information collection terminals and management and user networks need realize the WAN interconnection via GPRS. The mobile information collection terminals act as a bridge of communication between the perception network and the WAN. The perception of agricultural environmental data collected by the network shall be delivered to the WAN via GPRS and the remote monitoring shall be carried out by the end farmer.

The mobile collection terminals not only have the gateway function but also have the node management, data management and other functions in wireless sensor networks. The collection area includes the wireless sensor nodes of a variety of information, so need a unified node management, such as configuration of sensor node ID, node location, sampling period, node power and other properties.

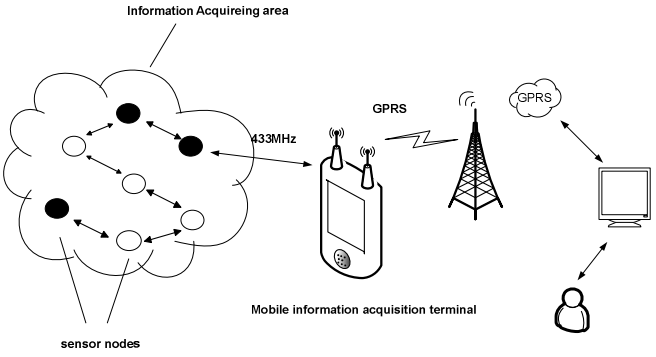


Fig. 1. Data management refers data storage display and other management to data of wireless senor network that the mobile collection terminals periodically collected

3 Hardware Design of Mobile Acquisition Terminal

The terminal includes mainly six parts: microprocessor module(MSP430F5438), GPRS module, APC240 module, man-machine interface module, memory module and power module. APC240 module is used to collect and manage the wireless sensors network data through 433MHz public frequency. Remote transmission module use SIM900A to send the information which is collected by APC240 module. The system structure is shown as that in Figure 2.

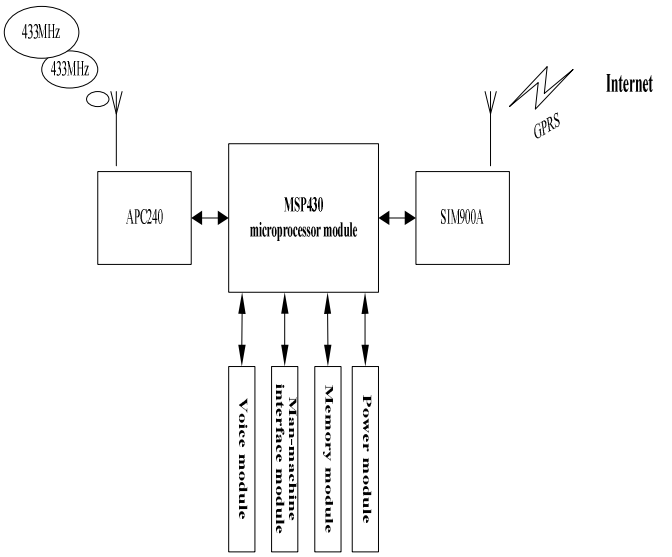


Fig. 2. Overall block diagram of system. The farmers could monitor the environmental information with the wireless temperature and humidity measurement system software.

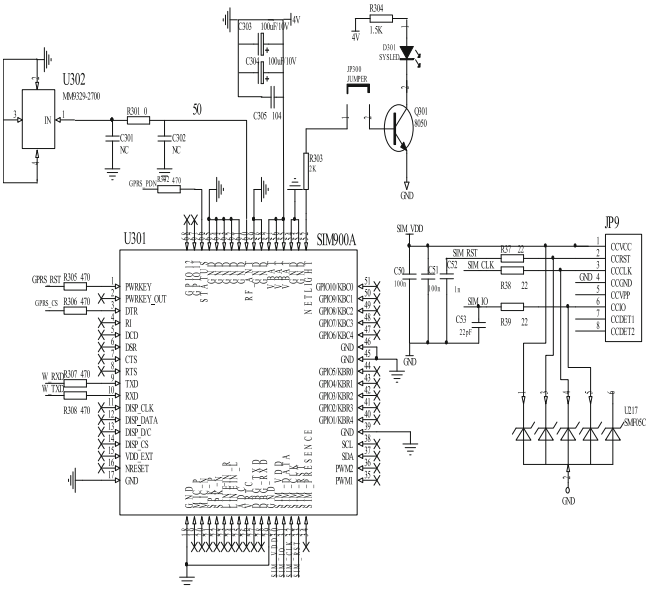


Fig. 3. U301 is the SIM900A module. The microprocessor (msp430) sends AT instructions to SIM900A module with the serial interface (W_RXD,W_TXD); JP9 is the socket of SIM card module. SIM card communicates with SIM900A through SIM_RST SIM_CLK and SIM_IO.U302 is the antenna of SIM900A module.

This system uses the SIMcom company's GPRS module SIM900A,which contains built-in TCP/IP protocol and TCP/IP AT instructions. SIM900A provides Antenna interface, Asynchronous serial interface and SIM card interface. The circuit structure of this module is shown as that in Figure 3.

4 The Software Design

4.1 The Design of Man-Machine Interface

The main interface includes mainly four parts, Time-setting sub-interface, Historical data sub-interface, Data collection sub-interface and system shutdown sub-interface. The function of time-setting sub-interface is that displaying time and setting time; in addition to display historical data, the historical data sub-interface can delete or store the sensor data; the farmers can shut down the terminal as need through the system shutdown sub-interface; the data collected by wireless sensors will be showed on data collection sub-interface. The system Man-machine Interface structure is shown as that in Figure 4.

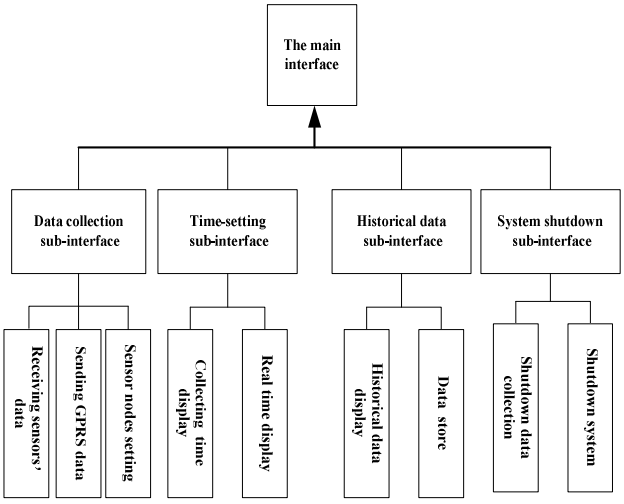


Fig. 4. Our design for MMI interface adopt the mature UC/OS+GUI technology, the operator click on the lively and visual icon but no need to input the number and characters, then it become simple, natural and friendly between man and machine.

4.2 The Design of Communication Module

The communication module includes mainly two parts: GPRS Transmission and the collection of wireless sensor network. The main work about the communication module is that setting different nodes addresses according to different sensors.

During the information transmission, the first thing is to initialize the SIM900A including the baud rate setting and mobile mode closure,etc. The key instructions are showed as follows:

```
   AT+CIPMODE=0 ;
//to chose TCP/IP mode
   AT+CIPSTART=TCP,"123.127.160.74","10005";
//to connect with Remote Monitoring System
   AT+CIPSEND="Monitoring data";
//send the monitoring data
   AT+CIPCLOSE=1 ;
//closing the  SIM900A
```

The PC software of remote monitoring center "wireless temperature and humidity collection system" could help farmers to monitor the real-time environmental information in farmland, which is developed to collect temperature and humidity data basic on the VS 2008 PLATFORM C#. Many node equipments could be connected through this software for monitoring environmental data ,such as air temperature, air humidity and soil temperature, etc. Also it could lead out the data form format through such terminal.

4.3 Implementation and Call of Main Task in System

In order to enhance the system's real-time, ucGUI multi-task system is applied and a UC/OS-II core is transplanted into MSP430 while the following content changes:

MSP430 stack member is 16, so the CPU_STK declared as unsigned integer data type. MSP430 stack grows from top to bottom. When OS_STK_GROWTH is defined as 0, it said the stack grows from bottom up, and when OS_STK_GROWTH is defined as 1, it said the stack grows from top down.

```
Typedef unsigned int            OS_STK;
#define      OS_STK_GROWTH        1
```

Transplantation of ucGUI mainly includes modification on three configuration files in the GUI/Config directory and writing of touch screen drive. LCDConf.h configuration file defines LCD control register, the display screen size and other optional features.

During normal operation of the system, Interrupt Service Routine ISR provides some related services of multi-task operating system, such as semaphore and mailbox ,etc., the flow chart of system main software is shown as that in Figure 5.

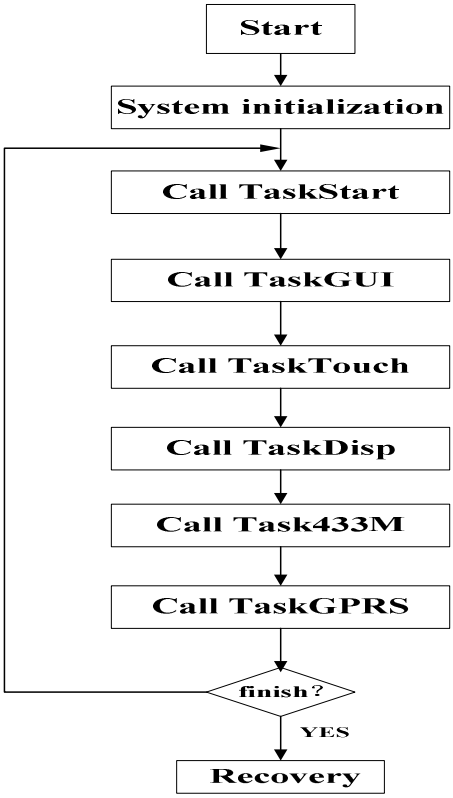


Fig. 5. The system uses six tasks, from highest to lowest priority order: TaskStart, TaskGUI, TaskTouch, TaskDisp, Task433M, TaskGPRS ,in which the TaskStart is created in main function, the other five tasks are created in the TaskStart.

5 Summarization

Through a wholly analysis to the actual demand of current facilities agriculture, then get a research of the MMI interface of mobile information collection terminals in facilities construction and communication technology. The design of mobile information collection terminals mentioned here have three advantages as follows: 1.Combining the GPRS technology and wireless sensor network makes the collection terminal does not rely on the wireless sensor and realize the real-time efficiency of remote monitoring system more stronger 2. The wireless communication module solve the complex cable arrangement limitation and realize the long-way transmission whenever and wherever 3. Successfully remove the UC/GUI format software, making the MMI more lively and simply, also solving the problems during operations to the complex control system, then low down the cost, more practically.

References

1. Pan, J.: The Object Network and Facility Agriculture. Agriculture Machinery Technology Extension (2), 45 (2010)
2. Zhou, X.: Facility Agriculture Online Monitoring System Based on Internet of Things. Journal of Taiyuan University of Science and Technology (2), 12 (2011)
3. Jia, X., Wang, C., Qiao, X.: Drive of wireless smart terminal and touch-screen human-machine interface for field irrigation 9, 8–12 (2008)
4. Ayday, C., Safak, S.: Application of Wireless Sensor Networks with GIS on the Soil Moisture Distribution Mapping, Ostrava, 25. - 28.1. (2009); Pu, Z.H. B.: Photo electricity testing technology. Machinery Industry Press, Beijing (2005)
5. Sun, Z., Zhao, W., Liang, J., Du, K., Zhang, Y.: Design and application of greenhouse remote automatic control system based on GPRS and WEB. Control & Automation (2010)
6. Texas Instruments. MSP430X5XX User's Guide (2008)

CFD Modeling and Simulation of Superheated Steam Fluidized Bed Drying Process

ZhiFeng Xiao[1], Fan Zhang[1], NanXing Wu[1], and XiangDong Liu[2]

[1] School of Mechanical and Electronic Engineering, JingDeZhen Ceramic Institute
JiangXi, 333403, P.R. China
cauxiao@126.com
[2] College of Engineering, China Agricultural University
Beijing, 100083, P.R. China
xdliu@cau.edu.cn

Abstract. An unsteady mathematical model of superheated steam fluidized bed drying process is established based on the transport process principles and computational fluid dynamics (CFD) method. The vapor-solid two-phase turbulent flow in the drying chamber is described with the Eulerian-Eulerian multiphase model. The model is solved by computer numerical simulation. The drying experiments of wet rapeseeds are conducted in a normal atmosphere. The experimental results agreeing well with the simulation results show that the mathematical model of drying process is effective.

Keywords: Mathematical model, Eulerian-Eulerian, Heat and mass transfer, Rapeseed.

1 Introduction

Using superheated steam as the drying medium and fluidizing medium, superheated steam fluidized bed drying technology combines the advantages of superheated steam drying and fluidized bed drying to improve drying quality, reduce drying time and energy consumption by high heat and mass transfer efficiency.[1,2] In recent years, some superheated steam drying models have been developed to simulate drying wet materials but are limited by their simplifications or assumptions. Taechapairoj et al.[3] developed a drying model for granular solid, which steam flow was assumed as plug flow, and interaction between particles and steam was ignored. One-dimensional single-particle models were established and integrated with a two-phase hydrodynamic model incorporating the effects of initial condensation and superficial gas flow[4], but the sample particle was fixed and immersed in inert beads and hygroscopic porous particles that were fluidized by superheated steam under reduced pressure. The initial steam condensation rate, amount of condensed water adsorbed by particles, and

[1] Author: ZhiFeng Xiao, PhD, mainly engaged in Drying Theory and Computer Simulation.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 141–149, 2013.

water evaporation rate were considered in a model used to describe the changes in temperature and moisture content of the product with time.[5-7] A coupling model was used to simulate drying of stationary porous material by superheated steam, but the initial condensation was not considered in the whole drying process.[8] The volume shrinkage of material during drying was negligible in some models.[4,9]

Superheated steam drying is a complex process accompanied by complicated heat and mass transfer, vapor–solid interaction, and particle–particle interaction in vapor–solid two-phase turbulent flow. The above simplified models are not suitable to describe the realistic drying process. Hydrodynamic behavior of solid particles is affected by interfacial force and solid stress, which are caused by vapor–solid interaction and particle–particle interaction, respectively. The interfacial force and solid stress can be described by the Gidaspow drag model[9,10] and the well-known granular kinetic theory[11,12]. Nienwland et al.[13] studied the bubble formation in fluidized bed and found that the bubble formation, size and shape are affected by particle size, density, and minimum fluidized velocity. It is helpful to establish an mathematical model for superheated steam fluidized bed drying process.

In this paper, the wet rapeseed drying experiments are conducted using superheated steam fluidized drying technology. The CFD model is established to describe the rapeseed drying process under normal pressure.

2 Experiments

2.1 Experimental Setup

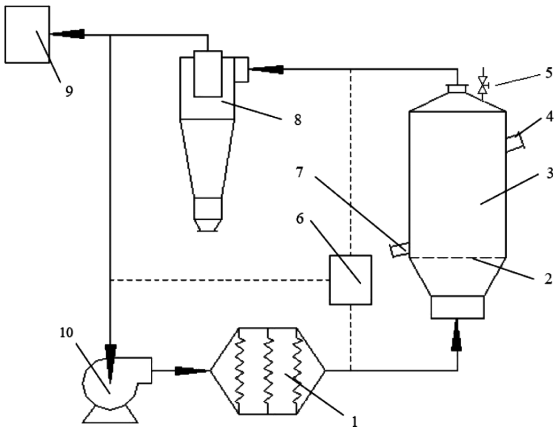


Fig. 1. Schematic diagram of the drying system

1-electric heater; 2-gas distribution orifice plate; 3-drying chamber; 4-feed inlet; 5-safety valve; 6-measurement and control system; 7-discharge outlet; 8-cyclone separator; 9-condenser; 10-blower.

Fig. 1 presents a schematic diagram of the experimental setup for the superheated steam fluidized drying. It is a circulating system consisting of a blower, electric heater, drying chamber, condenser, measurement and control system, and other devices. The cylindrical drying chamber, whose inside diameter and height are 120 and 250 mm respectively, has a conical section at the lower and upper enda. An orifice plate is placed between the cylindrical drying chamber and the lower conical section; the percentage of open area is 9.8% and the diameter of the holes is 1.5 mm. Before drying, some water is dripping into the drying chamber continuously and is heated to superheated steam by the electric heater.

2.2 Material

A small material is needed for the small drying chamber. Rapeseed has a small diameter and large specific surface area, is easily fluidized, and is simplified as a sphere in the drying model because it has an approximately spherical shape. Rapeseed was selected as the experimental material in the drying process.

2.3 Experimental Procedure

Some physical parameters of rapeseed, such as moisture content, temperature, particle density, size, and bulk porosity, are required for simulation. These parameters were measured by a moisture analyzer (MA150C-000230V1, Sartorius, Goettingen, Germany), infrared thermometer (Raynger ST6L, Raytek, Santa Cruz, CA), electronic balance (TE124S, Sartorius), oven (DHG-9140A, Jinghong, China), pycnometer, and vernier caliper. Bulk porosity can be obtained by particle density and bulk density. The normal pressure drying experiments were operated respectivly with superheated steam and hot air under the same drying conditions. The drying condition parameters (such as steam temperature, rotary velocity of blower) were set according to the experimental program, and the rapeseed was fed into the drying chamber quickly from the feed inlet. The feed inlet was closed and the rapeseeds were fluidized and dried. Sampling were performed at certain times (0, 6, 45, 90, 135, 180, 225, 270, 315, 360, 405 s) from the discharge outlet.

3 Modeling

3.1 Physical Model

The simulated object is the vapor–solid two-phase flow process and drying process in the drying chamber. To facilitate simulation, an axisymmetric cylindrical drying chamber was selected. Fig. 2 presents a schematic of the cylindrical drying chamber under a cylindrical coordinate system, in which the bottom center of the drying chamber is the origin of the coordinates. The bottom of the drying chamber is an orifice plate. Rapeseed was selected as the solid material to be dried in the small drying chamber due to its approximately spherical shape and small diameter.

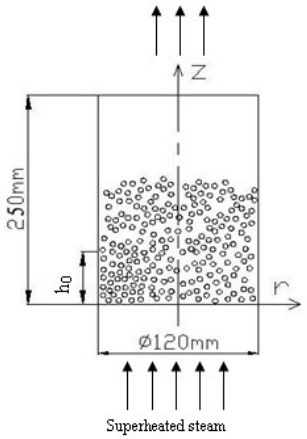


Fig. 2. Schematic of the drying chamber

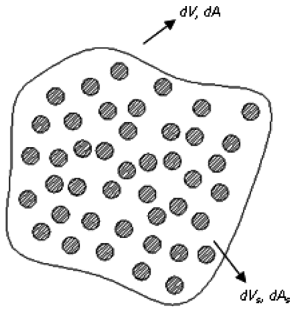


Fig. 3. Control volume of two-phase flow

3.2 Mathematical Model

The vapor–solid two-phase flow was treated as two-dimensional axisymmetric flow. The solid material was taken as isometric spherical particles, and the particle size was taken as constant during drying.

Control volume of two-phase flow system

The vapor-solid two-phase flow system in fluidized bed is generally viewed as a two-phase mixture, which vapor phase and solid phase occupy common space, interpenetrate and have respective properties such as dimension, velocity and temperature.[14] A control volume (dV) with a surface area (dA), shown in Figure 3, was chosen from the two-phase mixture in the drying chamber.

The volume fraction equation below was obeyed by at all times:

$$a_v + a_s = 1 \tag{1}$$

Continuity equation

The mass conservation equations are expressed as follows:

Vapor phase:

$$\frac{\partial}{\partial t}(a_v \rho_v) + \frac{1}{r}\frac{\partial}{\partial r}(a_v \rho_v r u_{vr}) + \frac{\partial}{\partial z}(a_v \rho_v u_{vz}) = \dot{m}_{sv} \tag{2}$$

Solid phase:

$$\frac{\partial}{\partial t}(a_s \rho_s) + \frac{1}{r}\frac{\partial}{\partial r}(a_s \rho_s r u_{sr}) + \frac{\partial}{\partial z}(a_s \rho_s u_{sz}) = \dot{m}_{vs} \tag{3}$$

The moisture mass fraction of granule X', can be predicted by Equation (4):

$$\frac{\partial}{\partial t}(a_s\rho_s X')+\frac{1}{r}\frac{\partial}{\partial r}(a_s\rho_s r u_{sr} X')+\frac{\partial}{\partial z}(a_s\rho_s u_{sz} X')=\nabla\cdot(a_s\rho_s D_{w,s}\nabla X')+\dot{m}_{vs} \quad (4)$$

Momentum conservation equations

Steam phase:

$$\frac{\partial}{\partial t}\left(\alpha_v\rho_v\vec{u}_v\right)+\nabla\cdot\left(\alpha_v\rho_v\vec{u}_v\vec{u}_v\right)=-\alpha_v\nabla P+\nabla\cdot\bar{\bar{\tau}}_v+a_v\rho_v\vec{g}+\vec{R}_{vs}+\dot{m}_{sv}\vec{u}_{sv} \quad (5)$$

Solid phase:

$$\frac{\partial}{\partial t}\left(\alpha_s\rho_s\vec{u}_s\right)+\nabla\cdot\left(\alpha_s\rho_s\vec{u}_s\vec{u}_s\right)=-\alpha_s\nabla P-\nabla P_s+\nabla\cdot\bar{\bar{\tau}}_s+a_s\rho_s\vec{g}+\vec{R}_{sv}+\dot{m}_{vs}\vec{u}_{vs} \quad (6)$$

$\vec{R}_{vs}$ and $\vec{R}_{sv}$ are drag forces between the two phases, which are defined as follows.

Energy conservation equations

The energy conservation equations of two-phase flow are formulated as follows:

Vapor phase:

$$\frac{\partial}{\partial t}\left(\alpha_v\rho_v H_v\right)+\nabla\cdot\left(\alpha_v\rho_v\vec{u}_v H_v\right)=-\alpha_v\frac{\partial P}{\partial t}+\bar{\bar{\tau}}_v:\nabla\vec{u}_v+Q_{sv}+\dot{m}_{sv}H_{evp} \quad (7)$$

Solid phase:

$$\frac{\partial}{\partial t}\left(\alpha_s\rho_s H_s\right)+\nabla\cdot\left(\alpha_s\rho_s\vec{u}_s H_s\right)=-\alpha_s\frac{\partial P}{\partial t}+\bar{\bar{\tau}}_s:\nabla\vec{u}_s+Q_{vs}+\dot{m}_{vs}H_{evp} \quad (8)$$

Heat and mass transfer model

The drying process can be divided into three distinct periods: the condensing and heating period, constant drying rate period, and falling drying rate period.

Condensing and heating period ($T_{s0}\le T_s < T_b$):

The condensation convective heat transfer coefficient is determined according to the method of McAdams[15]:

$$h_c=1.13\times\frac{6\alpha_s}{d_s}\times\left(\frac{[H_{evp}+C_{pv}(T_v-T_b)]\rho_l^{\,2}g\lambda_l^{\,3}}{\mu_l d_s(T_v-T_b)}\right)^{1/4} \quad (9)$$

The mass transfer rate between steam and particles:

$$\dot{m}_{vs}=\frac{Q_{sv}}{H_{evp}+C_{pv}(T_v-T_b)} \quad (10)$$

Where

$$Q_{sv}=h_c(T_b-T_v) \quad (11)$$

Constant drying rate period ($T_s = T_b$, $X \geq X_{cr}$):

The convective heat transfer rate:

$$h = \frac{6\lambda_v a_s Nu_s}{d_s^{\,2}} \tag{12}$$

Where

$$Nu_s = 2.0 + 0.74 Re_s^{\frac{1}{2}} Pr^{\frac{1}{3}} \tag{13}$$

The mass transfer rate:

$$\dot{m}_{sv} = \frac{Q_{vs}}{H_{evp}} \tag{14}$$

Where

$$Q_{vs} = h(T_v - T_s) \tag{15}$$

Falling drying rate period ($T_b < T_s \leq T_{v0}$, $X_{eq} \leq X < X_{cr}$):

The mass transfer rate in particles:

$$\dot{m}_{sv} = -\alpha_s \rho_s \frac{d\bar{X}}{dt} = \frac{6\alpha_s \rho_s D_{\text{eff}} (X_{cr} - X_{eq})}{d_s^2} \exp\left(-4\pi^2 \frac{D_{\text{eff}} (t - t_{\text{cr}})}{d_s^2} \right) \tag{16}$$

The heat transfer rate:

$$Q_{vs} = h(T_v - T_s) = \alpha_s \rho_s C_{ps} \frac{dT_s}{dt} + \dot{m}_{sv} H_{evp} \tag{17}$$

Numerical solution

The finite volume method was introduced to numerically solve the drying model using FLUENT. Furthermore, the heat and mass transfer model, drag force model, and physical parametric model were programmed using C and then incorporated into FLUENT through a user-defined function (UDF) to solve the drying model.

4 Results and Discussion

4.1 Drying Condition Parameters and Equipment Parameters

The inlet velocity of superheated steam was selected as 2.5 $m \cdot s^{-1}$ during the drying experiments. Some parameters used in the experiments are shown in Table 1.

Table 1. Drying condition parameters and equipment parameters

Parameters	Unit	Value
Steam pressure	Pa	1.01×10^5
Environmental temperature	K	300.2
Inlet steam temperature	K	423
Initial moisture content of rapesed (db)	$kg^{-1}kg^{-1}$	0.433
Initial particle density of rapeseed	kgm^{-3}	1184
Bone dry particle density of rapaseed	kgm^{-3}	817
Mean particle size of rapeseed	mm	1.76
Depth of static rapseed bed	mm	80
Porosity of papeseed in the static bed	—	0.43
Size of drying chamber	mm	$250\times\phi120$
Rotary velocity of fan	r min^{-1}	1840
Rated power of electric heated	kW	9

4.2 Drying Dynamic and Simulation Results

Figure 4 shows the simulation and experimental results for moisture content of rapeseed. Figure 5 shows the drying rate curves from simulation and experiment. The whole drying process can be divided into three periods: the condensing and heating period, constant drying rate period, and falling drying rate period. The simulated drying curve is in agreement with the experimental results in three periods.

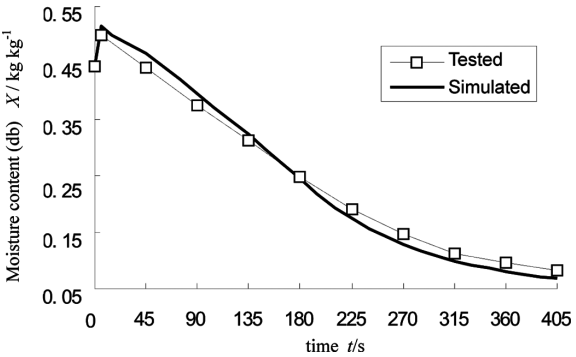


Fig. 4. Moisture content of rapeseed from simulation and experiment

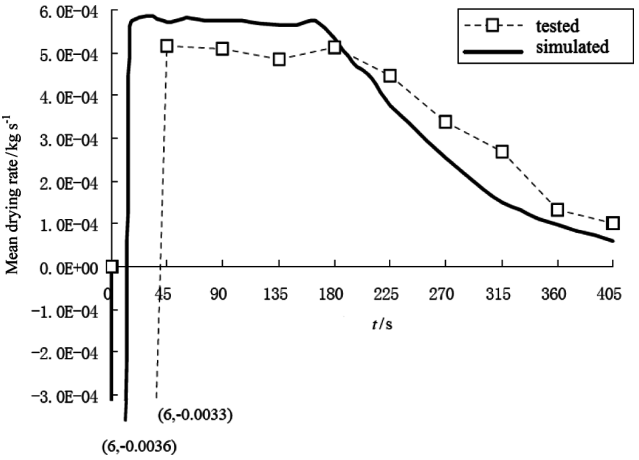


Fig. 5. Drying rate curve of rapeseed from simulation and experiment

A certain difference was found between the simulated drying curve and the experimental curve at the beginning. The simulated moisture content was greater than the experimental result in this period because the simulated curve ascends faster than latter. This discrepancy may be due to the fact that only steam condensation was taken into account and water evaporation from the material was ignored in simulation, though water evaporation from the material occurred during steam condensation. During the constant drying rate period, the simulated moisture content decreased faster; that is, the simulated drying rate was greater than that during the experiment.

5 Conclusions

The simulated drying curve obtained from the drying model was in good agreement with the experimental results. A negative drying rate occurred in the first period caused by initial steam condensation and lasted for about 6 s, and the moisture content of material increased up to a maximum. In the constant drying rate period, the simulated constant drying rate was greater than that of the experiment, which resulted in a shorter constant drying rate period. The simulated drying rate curve declined faster and a higher material temperature was obtained from simulation in the falling drying rate period.

References

1. Kudra, T., Mujumdar, A.S.: Advanced Drying Technologies. Marcel Dekker Inc., New York (2002)
2. Xiao, Z.F.: Numerical Simulation and Experimental Study on Fluidized Bed Drying Process with Superheated Steam. China Agricultural University Doctoral Thesis, Beijng (2008)

3. Taechapairoj, C., Prachayawarakorn, S., Soponronnarit, S.: Modelling of parboiled rice in superheated steam fluidized bed. Journal of Food Engineering 76(3), 411–419 (2006)
4. Tatemoto, Y., Yanoa, S., Mawatarib, Y., Nodaa, K., Komatsuc, N.: Drying characteristics of porous material immersed in a bed of glass beads fluidized by superheated steam under reduced pressure. Chemical Engineering Science 62(1-2), 471–480 (2007)
5. Soponronnarit, S., Prachayawarakorn, S., Rordprapat, W., Nathakaranakule, A., Tia, W.: A superheated steam fluidized bed dryer for parboiled rice: Testing of a pilot-scale and mathematical model development. Drying Technology 24(11), 1457–1467 (2006)
6. Poomjai, S., Thanit, S., Adisak, N., Somchart, S.: Mathematical model of pork slice drying using superheated steam. Journal of Food Engineering 104, 499–507 (2011)
7. Pakowski, Z., Adamski, R.: On prediction of the drying rate in superheated steam drying process. Drying Technology 29, 1492–1498 (2011)
8. Erriguible, A., Bernada, P., Couture, F., Roques, M.: Simulation of superheated steam drying from coupling model. Drying Technology 24(8), 941–951 (2006)
9. Wu, Z.H., Mujumdar, A.S.: Simulation of the hydrodynamics and drying in a spouted bed dryer. Drying Technology 25(1), 59–74 (2007)
10. Gidaspow, D., Bezburuah, R., Ding, J.: Hydrodynamics of circulating fluidized beds, kinetic theory approach. In: Fluidization VII, Proceedings of the 7th Engineering Foundation Conference on Fluidization, Brisbane, Australia (1992)
11. Du, W., Bao, X.J., Xu, J., Wei, W.S.: Computational fluid dynamics(CFD) modeling of spouted bed: Assessment of drag coefficient correlations. Chemical Engineering Science 61(5), 1401–1420 (2006)
12. Neri, A., Gidaspow, D.: Riser hydrodynamics: simulation using kinetic theory. AIChE Journal 46(1), 52–67 (2000)
13. Nienwland, J.J., Veenendaal, M.L., Kuipers, J.A.M., van Swaaij, W.P.M.: Bubble formation at a single orifice in gas fluidized beds. Chemical Engineering Science 51(17), 4087–4102 (1996)
14. Zhou, L.X.: Theory and Numerical Modeling of Turbulent Gas–Particle Flows and Combustion. Science Publishing Co., Beijing (1994)
15. Tao, W.Q.: Numerical Heat Transfer, 2nd edn. Xi'an Jiaotong University Press, Xi'an (2001)

Feasibility Study of Veterinary Drug Residues in Honey by NIR Detection

Hongqian Chen[1], Zhenhua Tu[2], Zhaoshen Qing[3], Xiaobin Qiu[1], and Chaoying Meng[3,*]

[1] China Agriculture University Net Center, Beijing, 100083, China
[2] China Food Industry Promotion Center, Beijing, 100062, China
[3] China Agriculture University , Beijing, 100062, China
mcy@cau.edu.cn

Abstract. In order to vertify the feasibility of the near-infrared to detect the veterinary drug in honey, studying the impact of different background correction methods for linear modeling and nonlinear modeling. Use PLSR to establish the linear model between near-infrared spectroscopy and to be measured, and LSSVR to establish the nonlinear model between near-infrared spectroscopy and to be measured. At the same time, we used Wavelet transform, Multiplicative Scatter Correction, Standard Normalized Variate, First derivative transform, Second derivative transform, Orthogonal Signal Correction and other background correction methods to improve the feasibility of the near-infrared detection.

Keywords: Near-infrared, Honey, Tetracycline.

1 Introduction

Honey is the nectar which is collected by bees from many plants, or is the sweet substance which is fully brewed in a special substance. In the process of feeding bees, beekeeper fed the bees with a large number of antibiotics such as tetracycline, chloromycetin and so on, in order to prevent the illness of the bees. This drug can be stored in honey's cells, tissues, organs, and with the food chain will threat human health. Especially in 2008, the exposure of the melamine incident highlights the problems in the field of food safety in China.

According to Hygienic Standard for Honey(GB/T5009.95): tetracycline antibiotics which are the major residues of veterinary medicine have been explicitly included in the detection range, and its limit of physicochemical index residue must be less than 0.05mg/kg. At present, as the requirements of European Union and The United States of America on pesticide and veterinary medicine

* This work is supported by National Natural Science Foundation of China(30901127) and Rapid detection of food in small and medium-sized enterprises of public service platform construction project(2011GH552096), National torch plan that environment of industrialization of construction project.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 150–156, 2013.

residue, the quality of honeys which China exports should be improved. In January 2002, The European Union banned imports of Chinese products which originated in animals, because chloramphenicol residues of shrimp exceeded the standard. It also caused a chain reaction of Japan, Canada and Hong Kong. The "Green barriers" policy which EU adopted in the honey trade has not only seriously affected the honey export of China, but also influenced the whole industry [2]. In China, the phenomenon that residues of tetracycline exceed the standard becomes more and more serious. In July 2004, Industrial and commercial bureau of Sichuan Province investigated the honey on the market: only 4 batches of honey qualified in the 30 batches, and the unqualified rate is 87%. The most prominent problem is the tetracycline antibiotic residues exceeding, and 10 batches of honey didn't qualify while the unqualified rate is 33%. In order to safeguard the interests of consumers, to ensure the quality of honey for import and export and find an effective detection method of tetracycline antibiotic residues exceeding is imminent. Efficient and practical detection technology contributes to the safety appraisal of honey, and it can ensure that China's export of honey meets all the international standards.

At present, there are several methods to detect the veterinary drug residue of honey, such as traditional chemical method, liquid chromatography, ultraviolet spectrophotometer, electrochemical analysis and so on. However, these methods are defective, as some of them are devastating while others are time-consuming and expensive. So, it needs a quick, lightweight, accurate and effective detection method in the research of tetracycline compositions in honey.

Some of previous studies have shown that Near-Infrared Spectroscopy can precisely detect the composition of honey. However, the content of tetracycline ingredients in honey is quite low, which leads that the information of near infrared region was disturbed by the background noise, so it is difficult to extract and utilize it. Meanwhile, Near infrared spectrum detection technology has been studied and applied in the detection of honey quality, identification of adulteration, plant source and place of origin. The near infrared spectrum detection technology has been used in the quality detection of honey products, and the adulteration detection, and the detection of plant source, and origin detection, but the detection of veterinary drug residues in honey has not been reported. In this study, the feasibility of detecting veterinary drug residues in honey by Near-Infrared Spectroscopy was investigated, in which variety of spectral preprocessing and non-linear quantitative model will be used.

2 Materials and Methods

2.1 Materials

This study collected the main origin of honey, include Sichuan, Qingzang, Beijing and so on. The honey’s variety include uninoral honey and multifloral honey, like Robinia pseudoacacia L, Friobotryajaponica (Thunb.) Lind, Zizyphus jujuba NM ver.

iaerirus(Bunge.)Relrl, Astragalus sivius L, Vites negundo van heterophylla (Franch.)Rehd, Tilia amurensis Rupr, Dimocarpuslongan Lour, Brassico campestris L, Litchi chine-is Soon, Eurya, Citrus reticulata Blanco, Sophoraviciifclia Hanue. There are 153 honey samples, and stored at 4℃ freezer.

2.2 Instruments and Parameters

In this study, spectral acquisition used ISF/28N Fourier near-infrared spectrometer instrument which is product in BRUKER co. This instrument's range is 3600~12500 cm^{-1}, a minimum resolution is $1cm^{-1}$. When collecting the spectrum of liquid, we can use transmission pool (optical path is 2mm or 20mm), quartz liquid transmission and reflection optical fiber (fixed optical path is 2mm) Annex.

2.3 The Acquisition of Near-Infrared Spectroscopy

The experimental spectra are collected in a temperature controlled laboratory (In this study, the temperature is 26°C). The instrument must be warm up for 30 minutes or more before test. At the same time, if honey is crystallized, the crystal honey samples must be heated to nine hours in 40 ℃ water bath and then place samples in 26°C environment temperature.

2.4 Detection of Tetracycline Content in Honey

We use high performance liquid chromatography - ultraviolet detect method to detect the content of tetracycline in honey. Honey samples dissolved in 0.1mol/L Na2EDTA-Mcllvaine solution, after centrifuged, the supernatant use Oasis HLB Solid phase extraction column and the carboxylic acid anion exchange column to purify, and use Elution to determine the volume. Finally, we use high performance liquid chromatography with UV detection at 350nm, and use external standard method (peak area) - the standard curve for quantification, and the concentration of tetracycline is 0.005mg/kg[6]. In 153 samples, we select 101 samples to detect, there are 41 samples have chemical value, and 60 samples don't have the chemical value.

3 Results and Analysis

3.1 PLSR Model of Tetracycline Content in Honey

There are 6 abnormal samples in the 41 samples, so we get 35 samples to calculate. After removing the abnormal samples, in order to get more stability of the model, the calibration set and predict set are in the ratio of 2:1, 3:1 ,7:3, 4:1, 3:2, according to K-S law to divide. And then, all set establish PLSR model respectively, all indicators shows in Table 1.

Table 1. Statistic data of honey including calibration and prediction for 5 times

Honey quality indicators	Divide number	Sample set	Number of samples	Range	Average	Standard deviation
Tetracycline(μg/kg)	1	Calibration set	27	10.2-66.8	22.1333	14.1161
		Prediction set	8	11.3-47.6	24.9125	13.2533
	2	Calibration set	24	10.2-66.8	21.7958	14.9162
		Prediction set	11	11.3-47.6	24.8909	11.2625
	3	Calibration set	21	10.2-66.8	21.8571	15.4482
		Prediction set	14	11.3-47.6	24.1357	11.2371
	4	Calibration set	25	10.2-66.8	22.1080	14.6854
		Prediction set	10	11.3-47.6	24.42	11.7571
	5	Calibration set	28	10.2-66.8	21.8143	13.9547
		Prediction set	7	11.3-47.6	26.5857	13.3714

The calibration set and the prediction set be calculate with Auto-scaling method, and then we use PLSR method for multivariate statistical analysis of the experimental data. Nonlinear iterative partial least squares (NIPALS) algorithm is used to get partial least square factors. The optimum factor number of the calibration model (# LV) is determined by two factors, LOOCV and prediction residual sum of squares (PRESS). The result of honey tetracycline Fourier transmission and reflection spectra PLSR model showed in Table 2.

Table 2. The PLSR models results of honey tetracycline content for 5 times

Honey quality indicators	Divide number	Calibration set					Prediction set		
		n^a	#LV[b]	r	SEC	RSD_c(%)	n^c	SEP	RSD_p(%)
Tetracycline (μg/kg)	1	27	4	0.6287	10.9774	49.6%	8	14.6035	58.62%
	2	**24**	**4**	**0.6170**	**11.7386**	**53.86%**	**11**	**12.9852**	**52.12%**
	3	21	4	0.5946	12.4208	56.83%	14	11.7928	48.86%
	4	25	4	0.6165	11.5629	52.3%	10	13.6492	55.89%
	5	28	4	0.6271	10.8693	49.83%	7	15.2219	57.26%

Note: a: n is number of samples in the calibration set, b: # LV is the best number of factors in calibration set, c: n is the number samples in prediction set.

The predict effect of PLSR modeling to establish honey tetracycline Fourier transmission and reflection spectra is not good, because the tetracycline content in honey is too low. The second division is much better, r is 0.6170, RSD is 52.12%. Figure 1 shows the correlation of the predictive value and the true value of the model built in the 2nd division of the tetracycline content in honey. From the figure, poor model prediction accuracy is difficult to directly meet the forecast requirements.

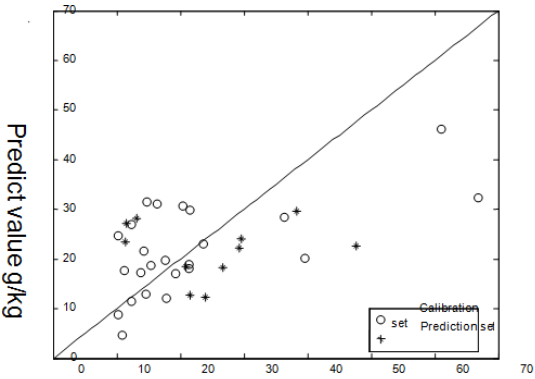


Fig. 1. Correlation between the measured and predicted values of honey tetracycline content for the second time: FT transflectance spectra (4000-12500cm^{-1})

3.2 Map of the Signal Optimization

Multiplicative scatter correction, standard variable transformation, the Fourier transform and wavelet transform are separately processed for the spectral matrix (X matrix), without taking into account the specific characteristics of the test the amount of information to be analyzed. Therefore, they are difficult to remove extraneous interference while without loss and to be measured on the spectral information. For this case, Wold, orthogonal signal correction (Orthogonal Signal Correction, OSC) method used to avoid the removal of important information on the forecast to be measured. Wold use Orthogonal Signal Correction method to resolve this problem. The OSC algorithm make the measurement information is also taken into account, and it remove the spectral information which is independent with variable matrix Y. The OSC algorithm has so many improved algorithms, one is a direct orthogonal signal correction (Direct Orthogonal the Signal Correction. DOSC), which is Westerhuis proposed, a direct orthogonal signal correction (DOSC) algorithm provides exact solution to orthogonal signal correction proposed by Wold. Just by a simple least squares method, DOSC can figure out those with the Y orthogonal, in the X matrix space and to the greatest extent describe the X matrix variation (variation) of the orthogonal factor.

We use DOSC to preprocess the near-infrared Fourier transmission and reflection spectra of honey tetracycline, and DOSC removed irrelevant information, which makes the difference among the spectra of each sample more obvious. Removing the irrelevant information with DOSC, we use PLSR model to establish the near-infrared spectrum and honey tetracycline.

Selecting parameters of DOSC and establishment the PLSR model, the number of factor selected from 1,2, and tolerance factor of DOSC from the 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, and PLSR best the main factor number is selected from 1-20. When DOSC tolerance factor is set too high, the result of model appeared over-fitting; contrary, DOSC tolerance factor is set too low, the unwanted information filter not complete, and the precision of the model are also affected. When removing a factor of DOSC, and the

tolerance factor is set to 0.35, the accuracy of the model is the highest, r increased from 0.617 to 0.6781, RSD_p (%) decreased from 52.12% to 45.65%.

3.3 Nonlinear Calibration Model of Tetracycline Content in Honey

The tetracycline content in honey is very low (10^{-9}~10^{-7}), its non-linear effects and interference may also be larger. Therefore, we use LSSVM to establish the nonlinear model of tetracycline in honey, in order to explore the feasibility of quantitative detection. RBF kernel is used in this article, and LSSVR which used RBF kernel just need confirm two parameters, namely the regularization parameter γ and kernel bandwidth parameters. The parameters γ and σ^2 used two-step grid-search and remove one across-validation to determine.

Firstly, we follow the exponential growth to primaries parameter γ and σ^2, and in the first choice, the range of parameter γ is from 1 to 109, and the range of parameter σ^2 is from 0.01 to 100. After first choice, the value of parameter γ is 1 and the value of parameter σ^2 is 81.92. Secondly, we use uniform growth to feature parameter γ and parameter σ^2, in the first choice the initial range of the parameter γ is 0.1-10, and the initial range of σ^2 is 0.01~100. The growth step of parameter γ is 0.1, and the growth step of parameter σ^2 is 0.1. After second choice, the value of parameter γ is 1.2 and the value of parameter σ^2 is 9.8. After selecting the best parameter γ and the bandwidth parameters of kernel function, we establish the LSSVR model of Tetracycline content in honey, and the result of the model shown in Table 3. From the modeling results, the prediction of using LSSVR model is better than the PLSR model, RSD_p (%) decreased from 52.12% to 48.95%.

Table 3. The PLSR and LSSVR models constructed with different spectral preprocessing methods

Honey quality indicators	Model	Parameter	r	SEP	RSD_p(%)
Tetracycline(μg/kg)	PLSR	LV=4	0.6170	12.9852	52.12%
	DOSC-PLSR	n=1, tol=0.35, LV=1	0.6781	11.3634	45.65%
	MSC-PLSR	LV=2	0.5736	12.4985	50.21%
	SNV—PLSR	LV=2	0.5734	12.4955	50.2%
	WT-PLSR	Bior3.3, J=1, LV=4	0.6169	12.9852	52.17%
	First derivative - PLSR	2-9,LV=2	0.5327	12.3201	49.50%
	Second derivative –PLSR	2-9,LV=2	0.5744	12.3930	49.79%
	LS-SVR	γ =1.2, σ^2 =9.8	0.6618	12.1838	48.95%

LV: PLSR, The number of main factors; n: DOSC, Number of factors; tol: DOS, The number of tolerance factor; J: Wavelet decomposition level; γ: LSSVR, The regularization parameter, σ^2RBF, Bandwidth of kernel function parameters.

4 Conclusion

DOSC remove the spectral information with Y orthogonal, and it make the linear relationship between spectra and to be measured is more obvious, which can improve the predictive ability of PLS Model and also simplify model. The value of RSDp from 52.12% drop to 45.65% after DOSC treatment, so DOSC is a very effective pre-processing method in PLSR regression model.

We use LSSVR to build nonlinear model of tetracycline in honey, because our data base has less number of samples. The predict precision of LSSVR is better than PLSR, the value of RSD_p(%) decrease from 52.12% to 48.95%.

However, due to complex background, peak overlap, and the strong absorb effect of water, it is hard to accurately extract the effective information of tetracycline in honey with these methods, and to achieve rapid detect requirements.

References

1. Thomas, E.V.: A primer on multivariate calibration. Analytical Chemistry 66, 795A–804A (1994)
2. Daviesa, M.C., Radovic, B., Fearn, T., et al.: A preliminary study on the characterisation of honey by near infrared spectroscopy. J. Near-Infrared Spectrosc 10(2), 121–135 (2002)
3. Despagne, F., Massart, D.L., Chabot, P.: Development of a robust calibration model for nonlinear in-line process data. Analytical Chemistry 72, 1657–1665 (2000)
4. Delphine, J.R., Leardi, R., Noord, O.D., et al.: Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration. Analytical Biochemistry 67(23), 4295–4301 (1995)
5. Serra, B., Pesudo, E.F., Pallarés, G.J.: Quantitative determination of free amino acids in honeybee collected pollen using gas chromatography and spectrophotometry. Annales des Falsifications Expertise Chimique 897, 153–166 (1991)

Study on Anti-collapse Behavior of Solar Greenhouses Covering Rigid Plate under Snowstorm

Chuanjia Hu, Yujia Dai, Jiahe Wang, Mengyan Song, Xiugen Jiang, and Min Ding*

Department of Civil Engineering, China Agricultural University, Beijing 100083, China
{chuanjiahu,xrksmy}@163.com, 530979347@qq.com, wjh_414@sohu.com,
jiangxiugeng@tsinghua.org.cn, dingmin2008@gmail.com

Abstract. To analyze the effect of stressed skin action of rigid plate covering on anti-collapse behavior of solar greenhouses under snow load, the numerical simulation on the overall collapse process of single skeleton structure and 6-skeleton overall spatial structure with rigid plate covering were conducted on ANSYS. The collapse modes of solar greenhouses and snow load-displacement curves were obtained. The effects of different parameters on the anti-collapse behavior of solar greenhouses under snow load were also analyzed. The results showed that the stressed skin action of the covering could provide lateral support for the skeleton and increase integral rigidity of the structure and the bearing capacity to resist snowstorm. The lateral support of 8mm thick PC sun board equals that of 4 purlins, 10mm thick PC sun board equals 6 purlins, and 12mm thick PC sun board equals 8 purlins. It is suggested that skeleton interval is about 1m.

Keywords: solar greenhouses, rigid plate covering, stressed skin action, anti-collapse behavior, snow load.

1 Introduction

The solar greenhouse is an agricultural building as well as a production facility. Its security under various loads is always the priority [1-4]. Much research on solar greenhouses has been focused on the lighting, insulation, heat transfer etc., whereas little was on mechanical behavior and design methods [5-7]. The lack of scientific guidance to the construction of solar greenhouses leads to many accidents, especially during extreme snowstorms in recent years [8, 9]. The collapse not only causes the loss of greenhouse facilities but also results in a pause in agricultural production. Therefore, it is essential to investigate the mechanical behavior and the collapse mechanism of solar greenhouses under various loads.

Stressed skin action refers to the strengthening effect of building's surface covering on structure's integral rigidity with its own rigidity and strength [10]. It is always treated as only a structural safety reserve in the construction. This can sometimes achieve simply safety results. But sometimes the result is opposite. The translucent covering can be used as lateral support for the greenhouse skeleton when it is rigid plate such as PC sun board or plate glass etc. Then the contribution to

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 157–165, 2013.

integral rigidity from rigid plate covering can be used. This can ensure structure security with less or no extra lateral support systems. The economy objective can thus be achieved. However, research on stressed skin action has only been emphasized in light steel frames and rigid-framed structures at present. There was little research on collapse mechanism and design theory about solar greenhouse considering stressed skin action of rigid plate. In order to make designing model of solar greenhouse better match with the actual working state, and to ensure its security and economy under extreme snowstorms, it becomes necessary to develop this research.

The numerical simulation on overall collapse process of solar greenhouses with rigid plate covering under snow load is carried out on ANSYS. The effects of parameters including component material, structure size and construction techniques on anti-collapse behavior of solar greenhouses under snowstorm are discussed.

2 Finite Element Model

2.1 Solar Greenhouse Dimension

The sectional view and dimensions of selected solar greenhouse (Liaoshen I type) are shown in Fig. 1 and table 1. Single circular steel tube is used as the skeleton [5]. The back wall is reinforced concrete structure [11].

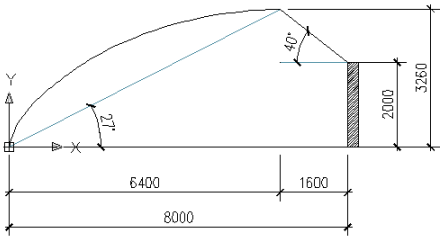


Fig. 1. Sectional view of the solar greenhouse

Table 1. The solar greenhouse dimension [8]

Parameter	Value	Parameter	Value
Span	8.00 m	Projected length of front slope	6.40 m
Ridge height	3.26 m	Projected length of back slope	1.60 m
Elevation of front roof	27°	Height of back wall	2.00 m
Elevation of back roof	40°	Reference interval of skeleton	1.00 m

In practice of Liaoshen I type, skeleton interval is generally 1m and the section size of circular steel tube is generally 30 mm × 2 mm, both of which are regarded as reference dimension. Rigid plate covering of solar greenhouse generally consists of PC sun board or plate glass etc. [11]. The thickness is generally from 6mm to 12 mm.

8 mm thick PC sun board is used as the reference. The 6-skeleton overall spatial structure with rigid plate covering and no purlins is used as the reference model (hereinafter referred to overall spatial structure) to investigate the effect of rigid plate on anti-collapse behavior of solar greenhouse structure.

2.2 Material Model

Structural steel of solar greenhouses is Q235 steel and the multi-linear kinematic hardening elastic-plastic model is used as its stress-strain relation to take material plasticity changes [12]. Mechanical properties of the steel, the PC sun board, and the plate glass are shown in Table 2. VonMises yield criterion is used as the criterion of all materials.

Table 2. Mechanical properties of materials [5, 12, 14, 15]

Materials	Thick-ness (mm)	Density ρ ($kg \cdot m^3$)	Poisson ratio μ	Elastic modulus E (GPa)	Yield stress (MPa)
Q235	2	7850	0.3	206	235
Plate glass	5	2400	0.25	70	58.8
PC sun board	8	1200	0.3	2.4	63

2.3 Element Type and Meshing

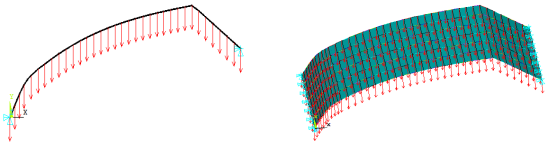
The ANSYS element BEAM188 is chosen for solar greenhouse skeleton. To PC sun board (hereinafter referred to as stressed skin), element SHELL181 is suitable for its consideration of in-plane shear deformation. Every single skeleton is divided into 40 units, and the stressed skin is divided into hexahedral elements using rules of free meshing [13].

2.4 Constraint Condition and Loading Mode

Only translational degrees of freedom of stressed skin elements are coupled with those of beam elements. Constraint condition takes the case both top and bottom hinged as the reference condition. The loading mode is to impose on the skeleton vertical down line load q to simulate snow load. The snow load can be conversed from the line load by multiplying q with skeleton length, and then dividing stressed skin area between two adjacent single skeletons. In this paper, the ultimate load refers to ultimate snow load.

2.5 Calculation Model and Analysis Method

Finite element models of solar greenhouse structure are created according to the method above, and shown in Fig. 2.



(a) Single skeleton structure (b) Overall spatial structure

Fig. 2. Calculation models of solar greenhouse

The displacement of solar greenhouse structure under actual loads is always very large due to the softness of components. Large displacement affects bearing capacity [12, 16]. Therefore, the NL GEOM command is opened to activate the large deformation effect in the analyzing process. Arc-length method based on Newton's law of Laplace is adopted. It is convenient to use static analysis of progressive loading to obtain the structure's ultimate bearing capacity by means of reasonable adjustments to overall load and loading sub-steps [17]. The descent stage of snow load-displacement curves can be received via arc-length method. And the vertex of the curve is the theoretical ultimate bearing capacity of the structure [18].

3 Results Comparison

3.1 Effect of Different Structural Calculation Model

Solar greenhouse structure is usually simplified to two-hinged arch structure or two-hinge truss arch structure to calculate the strength and deformation in plane ignoring the interaction and contribution of covering material to the greenhouse structural capacity.

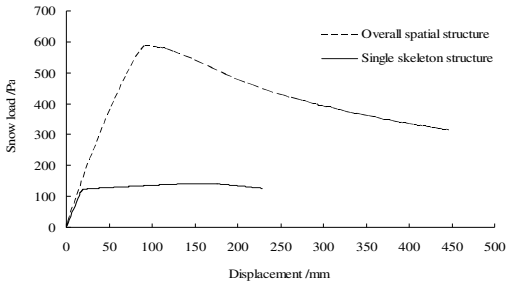


Fig. 3. Snow load-displacement curves of different structural calculation model

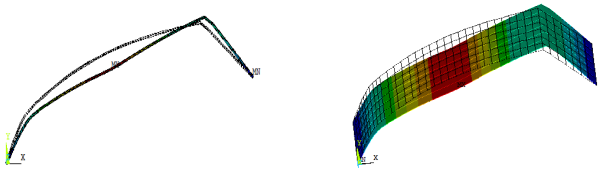
As shown in Fig.3, in the elastic stage, vertical deformation of the structure increases with the snow load increasing. After the ultimate load, the curve of overall spatial structure begins to decline, while the curve of single skeleton structure becomes horizontal. The ultimate load of overall spatial structure is 582 Pa and the maximum displacement is 448 mm. The ultimate load of single skeleton structure is only 142 Pa and the maximum displacement is 227 mm. The ultimate load of the former is 4.10 times of that of the latter, and the maximum displacement of the former

is 1.97 times of that of the latter. The reason is that the single skeleton structure has no lateral restraint support systems. Its collapse mode is out-of-plane buckling, and the steel does not fully play its role. Whereas with the lateral support provided by the stressed skin, the collapse mode of overall spatial structure is in-plane buckling, and the steel can fully play its role.

Fig.4 and Fig.5 present the comparison results of deformation before and after the ultimate load to show the buckling modes of different structural calculation models.

(a) Front view (b) Right view

Fig. 4. Deformation of single skeleton structure



(a) Front view (b) Isometric view

Fig. 5. Deformation of overall spatial structure

3.2 Effect of Different Number of Purlins

According to Part 2.1, rigid plate covering can effectively improve the bearing capacity to resist snowstorms and it can partly replace purlins to prevent out-of-plane buckling of skeletons. Here, the cross section of purlins is circular steel, and its section size is 10mm × 1mm.

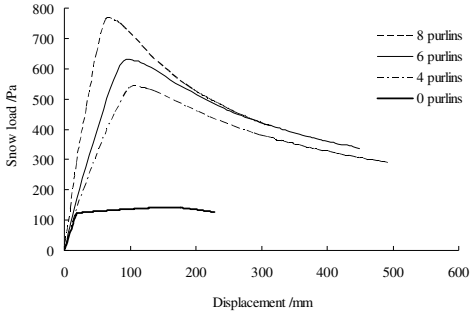
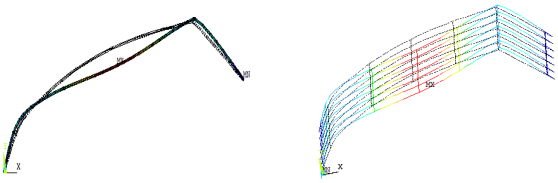


Fig. 6. Snow load-displacement curves of structures with different purlins

As shown in Fig.6, without covering materials, the more the purlins are, the greater ultimate load is, and the smaller the maximum displacement is. Curves of structures with 4, 6 and 8 purlins are similar in shape. It indicates that they have the same collapse modes and all of them are in-plane buckling. The ultimate load of the structures with 4, 6 and 8 purlins is respectively 537 Pa, 627 Pa and 761 Pa, increased by 16.8% and 21.4% in order. The structure without purlins and covering materials is equivalent to single skeleton structure. To compare with previous models, the deformation diagrams of overall spatial structure with 6 purlins are presented, as shown in Fig.7.



(a) Front view (b) Isometric view

Fig. 7. Deformation diagrams of the structure with 6 purlins

3.3 Effect of Different Skeleton Interval

As shown in Fig.8, changing skeleton interval is equivalent to change the degree of lateral support of stress skin on the structure. The greater the skeleton interval is, the smaller the ultimate load is, and the greater the maximum displacement is. The shape of curves of structures with different skeleton interval is similar, and collapse modes are all in-plane buckling. The ultimate load of structures with 1m, 1.5m and 2m skeleton interval is respectively 582 Pa, 492 Pa and 358 Pa, reduced by 15.5% and 27.2% in order. The recommended skeleton interval is 1 m considering the supporting role of stressed skin.

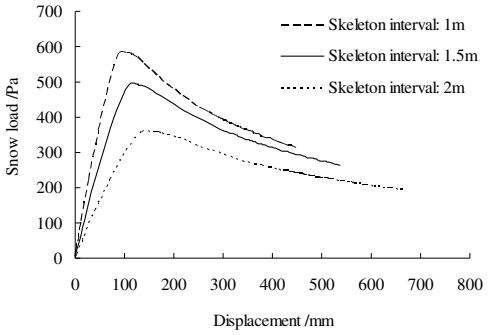


Fig. 8. Snow load-displacement curves of structures with different skeleton interval

3.4 Effect of Different PC Sun Board Thickness

As shown in Fig.9, the greater the PC sun board's thickness is, the greater the lateral support is, the greater ultimate load is, and the smaller the maximum displacement is. And the shape of curves is similar. The ultimate load of structures with 6 mm, 8 mm, 10 mm and 12 mm thick PC sun board, is respectively 448 Pa, 582 Pa, 672 Pa and 745 Pa, increased by 29.9%, 15.5% and 13.2% in order. Compared with the result of structures with different number of purlins, the ultimate load to resist snowstorms of the structure with 8 mm thick PC sun board is comparable to that of the structure with 4 purlins, and 10mm thickness comparable to 6 purlins, 12mm thickness comparable to 8 purlins.

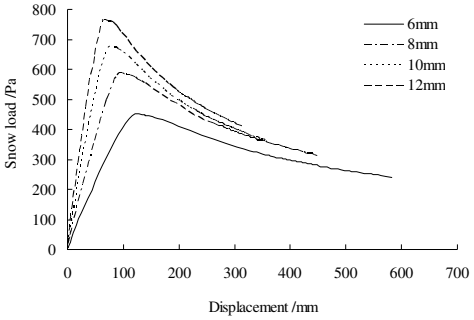


Fig. 9. Snow load-displacement curves of structures with different thickness of PC sun board

3.5 Effect of Different Covering Materials

As shown in Fig.10, the ultimate load of the structure with PC sun board is 582 Pa; while the ultimate load of the structure with plate glass is 448 Pa. The former is increased by 29.9% than the latter. The reason is that the in-plane shear stiffness of PC sun board is greater than that of plate glass. And in both cases the shape of curves is similar. It indicates that the collapse modes are both in-plane buckling.

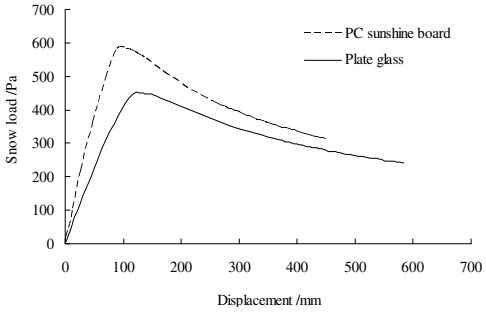


Fig. 10. Snow load-displacement curves of structures with different covering material

3.6 Effect of Different Constraint Condition

The skeleton rotation capacity at both ends is directly determined by constraint condition at each end, and the ultimate bearing capacity of overall spatial structure is also affected by constraint condition. As shown in Fig.11, the ultimate load of both top and bottom clamped case is 716Pa, while both top and bottom hinged case is 582 Pa. The ultimate loads of top hinged and bottom clamped, and top clamped and bottom hinged are respectively 627Pa and 645 Pa.

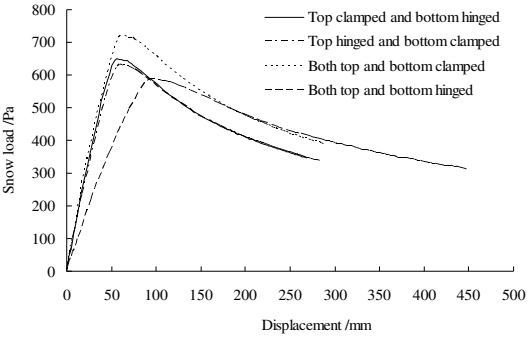


Fig. 11. Snow load-displacement curves of structures with different constraint condition

4 Conclusions

Under snow load, the collapse mode of overall spatial structure is in-plane buckling whenever the lateral support is provided by rigid plate covering or purlins. The skeleton interval determines the lateral support of rigid plate covering, thus affecting the ultimate bearing capacity to resist snowstorms. The increment of rigid plate covering's thickness improves the anti-snowstorm capacity as well. PC sun board is more effective comparing with other covering materials. Besides, increasing the skeleton steel tube section size can improve the ultimate capacity, but it will increase the amount of steel and construction cost simultaneously. Nevertheless, increasing constraint stiffness at both ends of the skeleton can increase the anti-snowstorm capacity, but causes less construction cost. Finally, rigid plate covering, such as PC sun board as translucent covering, can reduce the need of purlins. It is equivalent to purlins in some way.

However, the effect of stressed skin action in this paper is only investigated under snow load. For other loading cases, it needs further exploration. To simplify the calculation, the effect of gable at both ends of solar greenhouse is not taken into consideration. It is necessary to construct more realistic models for analysis to obtain more accurate results in future study.

Acknowledgments. Support for this research by Chinese Universities Scientific Fund No. 2011JS126, the Specialized Research Fund for the Doctoral Program of Higher Education of China No. 20110008120017, and the National Natural Science Funds No. 50979108.

References

1. Wang, C., Shi, W., Pei, X.: Comparing the Front Roof Permeated Sunlight Performances and the Arch Mechanical Performances of Four Curvilinear Roofs of Solar Greenhouse. Journal of Northwest A & F University (Natural Science Edition) 38, 143–150 (2010) (in Chinese)
2. Lu, X., Qiu, L.: Single-side Slope Sunlight Greenhouse Skeleton Finite Element Optimization. Transactions of The Chinese Society of Agricultural Machinery 36, 120–151 (2005) (in Chinese)
3. Li, C., Liang, Z., Ju, J.: Spatial Finite Element Analysis of Heteromorphous Greenhouse. Journal of China Agricultural University 12, 84–87 (2007) (in Chinese)
4. Fang, R., Lu, H.: Research on Design of Glass Greenhouse. Agricultural Engineering Technology (Greenhouse & Horticulture) 9, 32–33 (2005) (in Chinese)
5. Engel, R.D.: Using Simulation to Optimize Solar Greenhouse Design. In: 17th Annual Symposium on Simulation, NJ, USA, pp. 119–139 (1984)
6. Tiwari, G.N., Dhiman, N.K.: Design and Optimization of a Winter Greenhouse for the Len-Type Climate. Energy Conversion and Management 26, 71–78 (1986)
7. Tiwari, G.N., Dubey, A.K., Goyal, R.K.: Analytical Study of an Active Winter Greenhouse. Energy 22, 389–392 (1997)
8. Liu, J.: Structure Optimization Design for Sunlight Greenhouse Skeleton Truss-work. China Agricultural University, Beijing (2008) (in Chinese)
9. Bai, Y., Ming, Y.: Analysis of Influence Factor on Safety and Durability of Steel Skeleton of Solar Greenhouse. Housing Materials & Applications 5, 14–15 (2005) (in Chinese)
10. Yuan, X.: Skin-effect Theory and Research Profiles. Sichuan Building Materials 1(36), 73–74 (2010) (in Chinese)
11. JB/T 10286-2001, Greenhouse Load Design Specifications (in Chinese)
12. Wang, B., Jin, B., Song, J.: Finite Element Analysis of Load Bearing Capacity of Steel Skeletons for Solar Greenhouses Under Vertical Loads. Journal of Ningxia University (Natural Science Edition) 30, 337–338 (2009) (in Chinese)
13. Wang, X.: ANSYS Numerical Analysis of Engineering Structures. China Communications Press, Beijing (2007) (in Chinese)
14. Zhang, X.: Sunshine Board in Agricultural Greenhouse. Agricultural Engineering Technology: Greenhouse Gardening 5, 24–24 (2007) (in Chinese)
15. Huang, S., Pen, D.: Project Design of Greenhouse Sealed Aluminum Cover. Agricultural Engineering Technology(Greenhouse & Horticulture) 12, 10–13 (2006) (in Chinese)
16. Chen, J.: Stability of Steel Structure Theory and Design. Science Press, Beijing (2001) (in Chinese)
17. Yu, Y., Wang, J., Ying, Y.: Nonlinear Finite Element Analysis of the Bearing Capacity of the Film in Plastic Greenhouse. Transactions of the Chinese Society of Agricultural Engineering 23, 181–185 (2007) (in Chinese)
18. Yu, Y., Wang, J., Ying, Y.: Nonlinear Finite Element Analysis of the Bearing Capacity of Arch Structure in Plastic Greenhouse on Snow Load Working Condition. Transactions of the Chinese Society of Agricultural Engineering 23, 158–162 (2007) (in Chinese)

Study on Identification Method of Foreign Fibers of Seed Cotton in Hyper-spectral Images Based on Minimum Noise Fraction*

Laiqi Xu[1], Xinhua Wei[2,**], Xinyun Zhou [2], Dazhi Yu[2], and Jinmin Zhang[2]

[1] School of Electrical & Information Engineering, JiangSu University, Zhenjiang, 212013, China
[2] Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education & Jiangsu Province, Jiangsu University, Zhenjiang 212013, China
wei_xh@126.com

Abstract. In order to improve the recognition accuracy of seed cotton foreign fibers, a study on identification method in hyper-spectral images based on Minimum Noise Fraction (MNF) was proposed ,which was applied to feature extraction to reduce the dimension of hyper-spectral images. This method reduced the numbers of hyper-spectral data, lessened the images noise to the minimum , but also decreased the computational requirements for subsequent processing. The white foreign fibers and cotton which were in small discrimination were selected in this paper as the research object. The hyper-spectral images were displayed in software ENVI with 256 bands in the wavelenghth range of 871.60nm-1766.32nm. Afterwards, the images would be processed with the iteration threshold segmentation method, inflation and corrosion. Meanwhile, the correlation of template images and destination images were calculated to find the spectral peaks so that to make template matching to eliminate the images of the cotton seeds. Results of experiments show that the above methods is suitable for identifying foreign fibers of seed cotton which achieved 84.09% rate of recognition.

Keywords: seed cotton, foreign fibers, MNF, hyper-spectral data, dimensionality reduction, feature extraction, template matching.

1 Introduction

The provision of GB1103-1999 is that foreign fibers refer to the non-cotton fibers and colored fibers which have seriously affected the quality of cotton, such as chemical

* Foundation items : Supported by a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions(Jiangsu financial education (2011) No. 8)), the key laboratory of agricultural equipment intelligent high technology research in Jiangsu (BM2009703), and the Program for New Century Excellent Talents in University(NCET-09-0731).

** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 166–176, 2013.

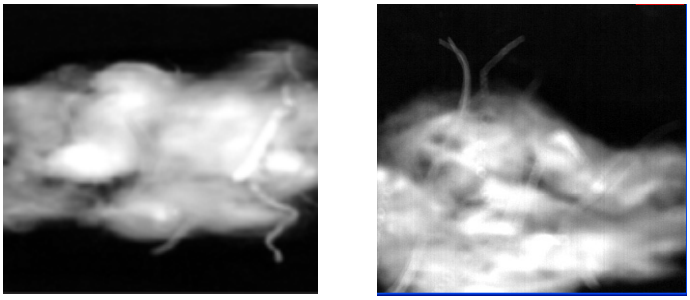
fiber, hair, silk, linen, dyed rope, plastic film, plastic rope (rope, cloth) and so on[1]. The harm caused by the foreign fibers in the textile business is huge, its performance is to reduce the production efficiency of cotton mixed with foreign fibers first. Meanwhile, it is easy to be labeled as numerous small fiber blemish, not only is difficult to remove in textile processing, but will innocently influence the quality of cloth dyeing processes that pulled off shorter and finer fiber. The formation of a large number of small fibrous blemish pinning can easily result in the efficiency weaving [2]. Therefore, in order to realize the automatic sorting of foreign fibers in cotton, a variety of techniques are tried. For example, Liu Huanjun used spectral reflectance to classify of soils[3], where machine vision technology is the primary means of on-line detection of foreign fibers in cotton.

Using the spectral analysis technology distinction sample type is a fast and accurate method. For example, Chenyang created that the characteristic spectra of 3 kinds of gases and the weights was classified by similarity analysis [4]. These methods all directly realized the variety distinction using the test specimen spectrum characteristic peak. However, machine-based visual of the cotton foreign fibers used in indirect detection is the spectral characteristics of the material. Thus, Li Bidan in Tsinghai University examined the colorless plastic, the jute, the white hair silk by using the infrared absorption characteristic of foreign fiber [5]. Jia Dongyao and Ding Tianhuai produced the differences of foreign fibers of composition, physical and chemical parameters [6]. Church J.S. et al. used the near-infrared wave Bands within 1000nm-l700nm to realize in the wool polymer impurity examination [8]. Hyper-spectral image technologies for the detection of foreign fibers in cotton are not much visible in the literature at home and abroad. By analyzing the foreign fibers in the near-infrared hyper-spectral transformation algorithm to choose the best band of foreign fibers in cotton, has not been found in Chinese literature.

This paper takes advantage of the high-resolution digital camera system to get the high spectral images between seed cotton and 5 kinds of white foreign fibers in the wavelenghth range of 871.60nm-1766.32nm. Hyper-spectral imaging overcomes the limitation of the traditional single-band, as well as the band range limitations which have a relatively narrow band interval. And more number of characteristics of the band are able to get the spectral space features continuous and fine spectral characteristics. Because of the large amount of letter of hyper-spectral, multiband and high redundancy of the feature makes its information processing difficult, and data dimensionality reduction problem has been the application of hyper-spectral information processing problems [9,11]. Yang Wenzhu et al.in China Agricultural University proposed a method for selecting optimal detection bands, based on the extreme distribution of reflectance differences between cotton fiber and foreign fibers[12]. On one hand, data dimension reduction can make the images away from the noise and improve the data quality .On the other hand, removing the worthless band image can reduce the amount of computation and the band number of purposes, as well as the image processing efficiency. To effectively use of hyper-spectral data, dimension reduction is an integral part. The MNF transformation would solve the problem which mentioned.

2 Research Data

In this study, seed cotton and five typical white fibers including white confetti, white hair, polypropylene yarn, and white chemical silk, white plastic were selected. This paper takes advantage of the high-resolution digital camera system to get the high spectral images at the range of 871.60nm-1766.32nm. Each of the band is 4.3nm with256 bands.



a) sample1 b) sample 2

Fig. 1. Foreign fiber samples in seed cotton

3 Research Methods

3.1 The Minimum Noise Fraction (MNF) Transformation

MNF (Minimum Noise Fraction) is one of the common methods of feature extraction in hyper-spectral data [13]. The MNF transformation is essentially two cascaded principal components(PC) transformations[14]. First, the high-pass filter is used in processing the whole image data to get the noise covariance matrix C_N and digitalization.

$$D_N = U^T C_N U \tag{1}$$

In the formula, D_N is the diagonal matrix of the C_N characteristic value which is according to the size descending sequence arrangement. U is the orthogonal matrix which is constituted by the C_N characteristic vector, N is the wave band number. The formula (1) can be further transformed into

$$I = P^T C_N P \tag{2}$$

Where I is the unit matrix, P is the transformation matrix,which $P=UD_N^{-1/2}$. The original image X can be transformed to the new space Y. The data after transforming contained in the noise is with unit variance and without correlation between bands. Standard PC transformation is carried in images. Transformation formula is

$$\hat{C} = P^T C_D P \tag{3}$$

C_D ***is*** the covariance matrix of the original image, $\hat{C}$ is the transformed matrix. The above equation can be further transformed into a diagonal matrix:

$$\hat{D} = J^T \hat{C} J \tag{4}$$

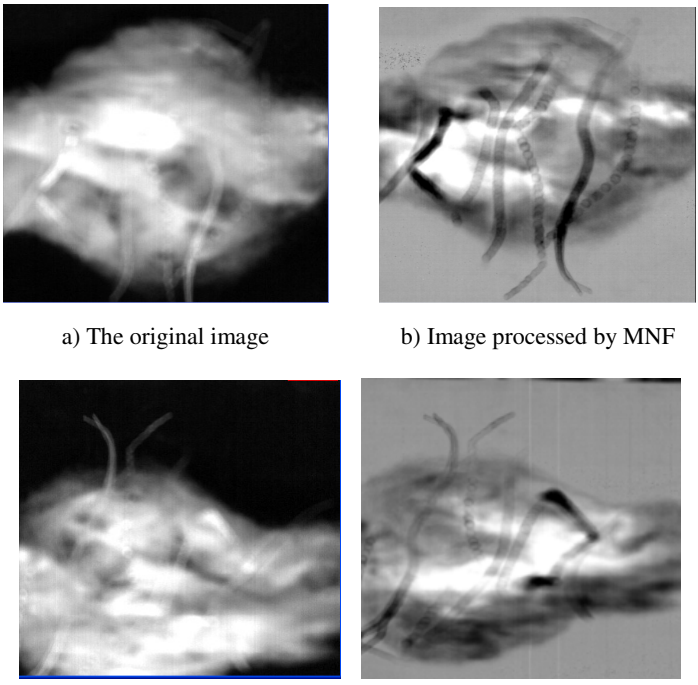
According to the size descending sequence arrangement constitution, $\hat{D}$ is the diagonal matrix of $\hat{C}$ characteristic value , J is the orthogonal matrix which constitutes by the $\hat{D}$ characteristic vector.

Through the above steps, MNF is got to the transformation matrix:

$$T_{MNF} = PJ \tag{5}$$

Using the MNF Transformation can remove noise from data by performing a forward transformation and determine which bands contain the coherent images by examining the images and eigenvalues.

As shown below (Fig.2) is the images of foreign fiber before and after by MNF transformation.



a) The original image b) Image processed by MNF

c) The original image d) Image processed by MNF

Fig. 2. Images processed before and after by MNF

3.2 Iteration Threshold Segmentation

For the online visual examination system, it is the key aspects with algorithm rapidity and accuracy. Therefore, choosing a quick division method is necessary under the guarantee division precision. It is more effective for border-based segmentation method in dealing with some relatively high contrast edges images. However, the content of foreign fibers in cotton is extremely small, and became finer after opening. With the low contrast between the target and background, segmentation method based on boundary does not apply to processing the cotton fibers images. Due to the gray of the cotton fiber image varies from the type of foreign fibers, the fixed threshold method can not achieve precise segmentation. In this case, dynamic threshold method is needed. Iterative threshold segmentation method is statistically significant for the best segmentation threshold, which is more suitable for cotton fiber.

The principle of the iterative method is that the gray distribution between foreground and background in the images do not overlap with each other. With this premise, the threshold segmentation is realized in two types of objects based on the idea of approximation. In General, the gray images within the change between 0-255 gray-scale changes is 256 levels. Some steps are given below:

Step1: Find out the maximum and minimum gray value of images to record as Z_{max} and Z_{min} respectively. Thus, make the initial threshold $T_0 = (Z_{max} + Z_{min}) / 2$;

Step2: According to the threshold T_K, to do image segmentation for the foreground and background to get the average gray values Z_o and Z_B•

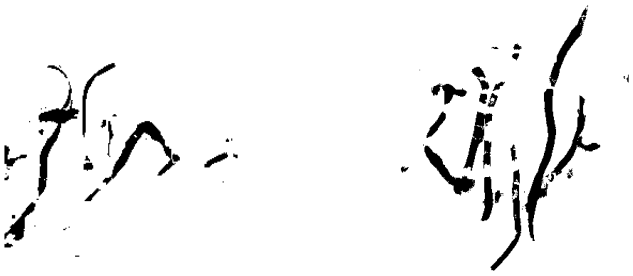
Step3: Get the new threshold for $T_{K+1} = (Z_o + Z_B) / 2$;

Step4: If $T_K = T_{K+1}$, the income is for threshold; Otherwise, turn to the Step2 to do the iterative calculation.

3.3 Template Matching, Corrosion and Dilation

Template matching is an important part of digital image processing. After the image processing, cotton seed in-depth can have the remarkable images characteristic, in this case, a method that eliminate cotton seed image is needed. Establishing a template of the average size model and carrying on the match with the image is named template match. In this article, the average value of cotton seed template was first found out, and then pattern recognition method was used. Meanwhile, two dimensional Fourier were transformed between the template image and matching image, between which the correlation was calculated .The highest position of the spectrum that match template objects.

After the MNF processing, iterative processing, corrosion and dilation, the foreign fibers were divided with the division effect achieves 84.09%. Now enumerates two pictures to take the demonstration, the demonstration is as follows:



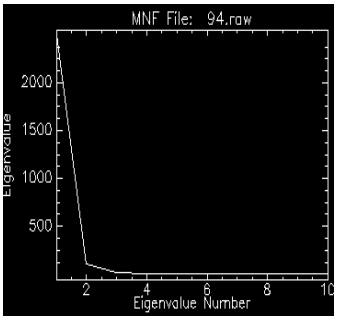
a) Image 1 b) Image 2

Fig. 3. Final segmentation images

4 Results and Analysis

4.1 The Results of Minimum Noise Fraction (MNF) Transformation

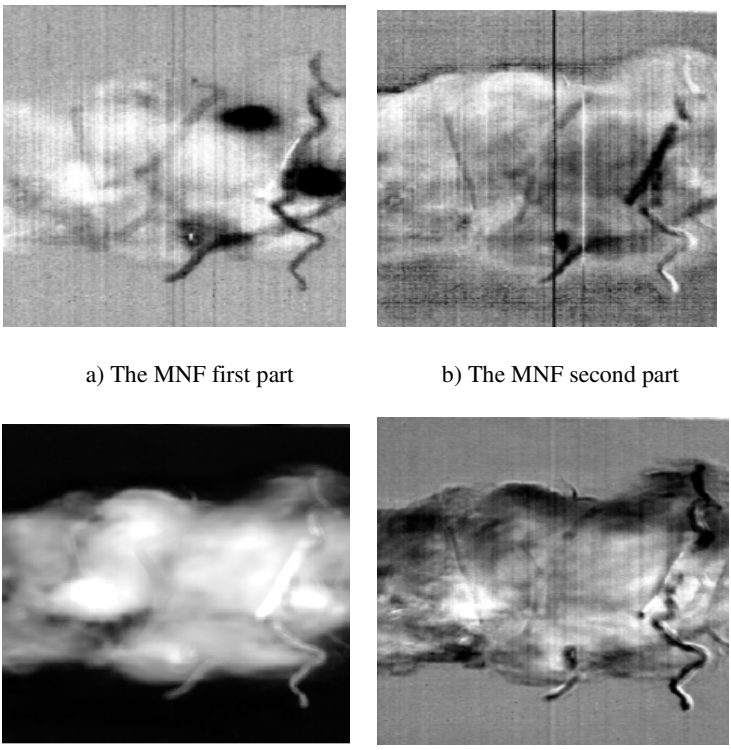
In the figure 4, there are the eigenvalue of the MNF and 10 top features images of the cumulative variance, which is generated by a MNF transformation.



MNF	Eigenvalue	Percent
1	2987.5520	85.91%
2	155.8276	90.39%
3	13.6063	90.78%
4	6.0610	90.95%
5	5.4138	91.11%
6	4.9964	91.25%
7	3.9639	91.37%
8	3.7245	91.47%
9	3.1142	91.56%
10	2.3820	91.63%

Fig. 4. Eigenvalue of the MNF and 10 top features images of the cumulative variance

It is known that the transformed image of cumulative variance reaches top 10 bands of 87.92%. And most amounts of images are contained in top 10 bands. Typically, after the MNF transformation, the band that contains the data Eigenvalues is greater than 1. When the Eigenvalue is close to 1, it indicates that the images have only noise. As we can see in figure 4, the fifth band characteristics are 1, which contains part of the noise. The first 4 after MNF processing components is shown in Figure 5.



a) The MNF first part b) The MNF second part

c) The MNF third part d) The MNF fourth part

Fig. 5. The Results of MNF

4.2 Results of Iteration Threshold Segmentation

In order to deal with further better in the image processing, appropriate image segmentation is needed. First of all, the iterative method is imported into the MATLAB to figure an iterative process. Combined with histograms, the threshold values used for the calculation of the iterative method are proposed. The method is based on 256 gray-scales of foreign fiber images, and the initial value of the last iteration is selected.

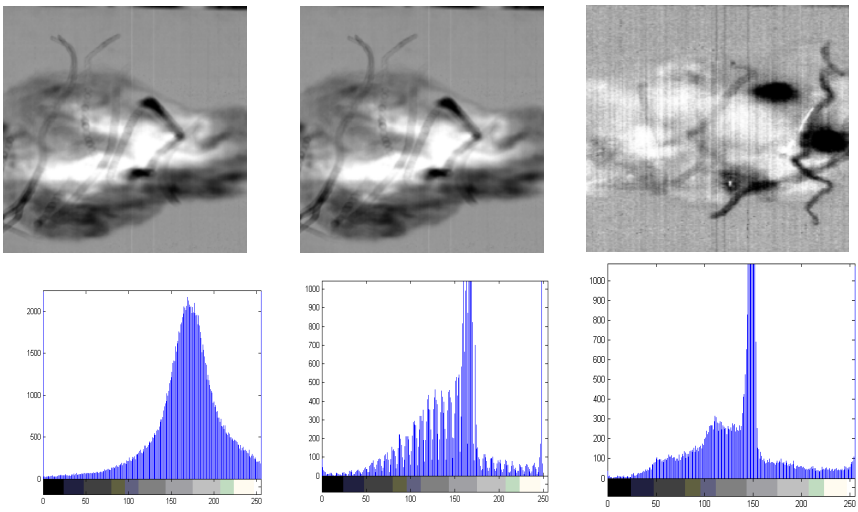


Fig. 6. Gray images 1,2,3 and its histograms

Table 1. The results of optimal threshold

Image number	Image 1	Image 2	Image 3
Optimal threshold	154	139	112

According to the principles of iterative methods, iterative calculation should be done. Finally, the thresholds are as shown in table 1.

4.3 Results of Template Matching

To calculate correlation of template image and destination image, the template matching should be selected . They are listed as follows.

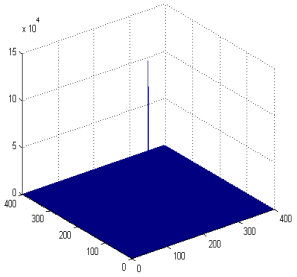
Step1: Convert into a template (396,398) binary image, and then do a matching image to convert a (396,398) binary image.

Step2: To do two-dimensional Fourier transformation of the template image and match the image.

Step3: Calculate correlation of template image and destination image. The method is to match the image rotated 180 degrees, then convolution calculation techniques based on fast Fourier transform evaluated. If the convolution rotates 180 degrees, the convolution and related calculations are equivalent.

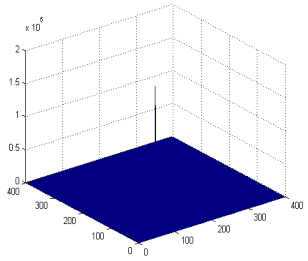
Step4: Find one of the three highest spectral peaks from the generated spectrum image spectral peaks. The three highest positions of the spectral peaks are matched with the template object.

a) Iteration and corrosion image

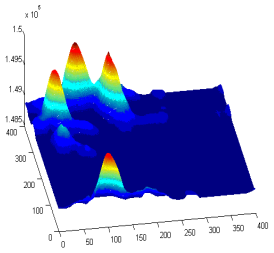


b) Fourier transformation

c) Template



d)Fourier transformation



e) Spectrum image

f) Match point

g)Final segmentation image

Fig. 7. The result of segmentation process

4.4 Data Statistical Results

With the proposed method, segmentation and dimensionality treatment has been used in the hyper-spectral images of the cotton foreign fiber, in order to show the method, there are 176 sets of data tests, and the experimental results are expressed as follows.

Table 2. Results of testing sets of samples

Name	Sample number	Correct identification number	Correct identification (%)
white confetti	32	30	93.75
white hair	39	29	74.35
polypropylene silk	36	32	88.88
white chemical silk	39	35	89.75
white plastic	30	22	73.33
Sum	176	148	84.09

5 Conclusions

Through theoretical analysis and experimental analysis of this article, we find that Green and other people's classic MNF transformation methods have the good effects in dimensionality reduction, separation of the of noise and feature extraction for similar surface features gathered data. Currently result fully illustrates the advantages of MNF transformation method. However, for the time complexity of the algorithm of hyper-spectral images, it is necessary for us to do further simplifying and further mining algorithm.

References

[1] Fang, H.: Analysis of application for the MQZY seed cotton foreign fibers. Cotton Processing in China (4), 12–14 (2010)
[2] Li, M.: Features and application of online foreign fiber detecting and clearing device. Shanghai Textile Science&Technology 34(1), 15–18 (2006)
[3] Liu, H., Zhang, B., Zhang, Y., et al.: Soil taxonomy on the basis of reflectance spectral characteristics. Spectroscopy and Spectral Analysis 28(3), 624–628 (2008)
[4] Chen, Y., Jia, L., Zhang, T., Guo, P., Wang, X., Chang, S.: A classification method for nonlinear fluorescent spectra based on edges matching. Acta Physica Sinica 59(1), 271–279 (2010)
[5] Li, B., Ding, T., Jia, D.: Design of a sophisticated foreign fiber separator. Transactions of the Chinese Society for Agricultural Machinery 37(1), 107–110 (2006)
[6] Jia, D., Ding, T.: The use of fiber infrared absorption characteristics of lint impurities detection method. Infrared and Millimeter Waves 24(2), 147–150 (2005)
[7] Xia, X., Ji, J., Chen, J., Liao, Q., Yang, Z.: Analysis of soil physical and chemical properties by reflectance spectroscopy. GEO-edge (4), 354–362 (2009)
[8] Church, J.S., O'Neil, J.A.: The detection of polymeric contaminants in loose scoured wool. Vibrational Spectroscopy 19(2), 285–293 (1999)
[9] Tong, Q., Zhang, B., Zheng, L.: Hyper-spectral remote sensing principles techniques and applications. Higher Education Press, Beijing (2006)

[10] Hughes, G.F.: On the Mean Accuracy of Statistical Pattern Recognition. IEEE Trans. Info. Theory 14(1), 55–63 (2008)
[11] Du, P., Chen, Y.: Advances in remote sensing science, pp. 369–370. Xuzhou China mine University Press (2007)
[12] Yang, W., Li, D., Wei, X., et al.: Selection of optimal band for detecting foreign fibers in lint cotton using spectroscopic analysis. Transactions of the CSAE 25(10), 186–192 (2009)
[13] Liu, P., Lin, H., Sun, H., Yan, E.: Dimensionality reduction method of Hyperion Eo-1 data. Journal of Central South University of Forestry &Technology 31(11), 34–38 (2011)
[14] Green, A., Berman, M., Switzer, P., Craig, M.D.: A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing 26(1), 65–74 (1988)

Study on Agricultural Information Push Technology Based on User Interest Model

Xiaorong Yang[1,*], Qingtian Zeng[2], Nengfu Xie[1,*], Lihua Jiang[1,*]

[1] Agriculture Information Institute, Chinese Academy of Agriculture sciences, Beijing, P.R. China

[2] Shandong Science and Technology University, Shandong Province, P.R. China

Abstract. Traditional agricultural information systems lack availability because of ignoring user preferences. This paper proposes a user interest modeling method which combines auto-modeling with artificial modeling. The value of user interest measure can be obtained by counting the frequency of information accesses. The statistics method based on the heat of information accesses is adopted to perform cluster agricultural information push. The result of experiment proved the user interest modeling method can express a user's interest in agricultural information better. The intelligently and personalization of the system are improved.

Keywords: User interest model, Information push, Personalized service.

1 Introduction

Traditional agricultural information systems focus on the integration and query optimization of information resources. The systems lack availability because of ignoring user preferences. Personalized services are the inevitable tendency of Internet. The personalized system can push satisfactory information resources to users actively by analyzing their personality and habits. Pablo Castells from Spanish Madrid university applied semantic-based personalized techniques to develop an extensible retrieval system which could automatically estimate the value of user personalization. Alexander Pretschner from Germany and Susan Gauch from America cooperatively studied an ontology-based personalized retrieval system. The system could reorder and filter search results by matching search results and user character so that the search performance of the system was improved. Automatic Department of Tsinghua University presented a maximum distance-based Ranking algorithm to match information and developed the Bookmark system. The system adopted expanding the Bookmark function of the browser to record a user's access behavior and obtained a user's demand by analyzing his commentary. Professor Qingtian Zeng from Shandong Science and Technology University studied the obtaining technique of users' explicit

* Key Laboratory of Agricultural Information Service Technology（2006-2010）, Ministry of Agriculture, The People's Republic of China.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 177–182, 2013.

and implicit knowledge needs. And he applied the technique to develop the E-learning system and multi-user interactive Question-Answering system.

Usual user interest model can be based on neural network, evaluation matrix, vector space model or ontology. However the neural network-based model cannot be readily intelligible and is now rarely applied. The poor adaptability of evaluation matrix-based model results in the difficulty in renewing a user's interest. The ontology-based model adopts multi-level domain knowledge to denote users' interest. At present building ontology is only manual or semiautomatical so that it takes too much time. The vector space-based model often results in deviation. So this paper integrates auto-modeling and artificial modeling to build user interest model. The value of user interest measure can be obtained by counting the frequency of information accesses. The statistics method based on the heat of information accesses is adopted to perform cluster agricultural information push.

2 User Interest Model

User interest model is a kind of user description which is algorithms-oriented and has a specific data structure. It can store and manage a use's background information and history behavior of accesses. It can record a user's interest points and describe his relatively inflexible information needs at a certain time. This paper integrates auto-modeling and artificial modeling to build user interest model. On one hand

In this paper, a combination of automatic modeling and artificial modeling construct user interest model, on the one hand, to record the user's behavior in the data access and mining user access logs to complete the modeling of the user, while allowing users to manually customize their own user model, which can be as effective as possible, complete access to users' needs and to meet agriculture personalized information retrieval service requirements（Barrett,R. et .al，2007）.

2.1 Automatic Modeling

Automatic modeling by mining the log information can analyze the behavior of different users, and dynamically learn and improve in order to obtain the explicit needs and implicit needs of the user. For example, whenever the user to enter the agricultural science and technology information sharing system, user interest in a variety of behaviors will be recorded, including his clicking on the information classification of agricultural information resources, entering your search keywords, and browsing information. The order of these information classification accessed in one operation will be recorded such as A→D→E→C→D→B is a resource access sequence after a user login system which means that the user has accessed the system resources in the class A, D, E, C, D, B. There are many ways which can automatically create a user model. This paper uses a word frequency mining algorithms accessed high-frequency access in the recent time（Zeze Wu et .al, 2009）to establish the user interest model. The algorithm filters out the retrieval keywords which were most frequently used recently, and represents the user model with a weighted directed graph or a weighted vector

model. The user interest model is shown in Figure 1. In order to distinguish different resources junction point between a user's interest points set and interest vector set, a set of rectangular nodes with weights is used to represent a user's interest point set and a directed graph of circular nodes with weights is used to represent a user's interest vector set.

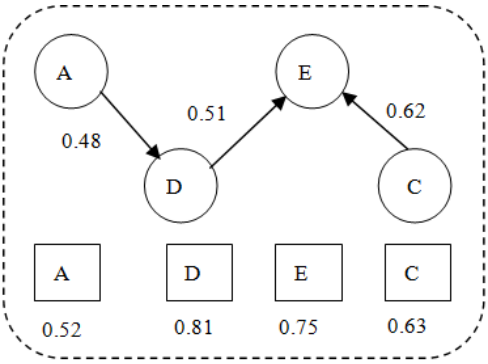


Fig. 1. User Interest Model Map

2.2 Artificial Modeling

On the basis of the automatic modeling, by personalized mapping the fields of data sheet in an agriculture database, each user can customize his favorite search criteria of agricultural information resources so as to meet the requirements of personalization. Besides, a user can improve the user model by filling in his personal information such as gender, age, occupation, education, personal interests and so on. Thus the system can cluster users so as to improve the user model.

3 Push Technology of Agricultural Information Resources Based on User Demand-Driven

3.1 User Interest Statistics Based on the Number of Access

As a traditional user interest statistical method, the principle of user interest statistics based on the number of access is that the more a user is interested in some information resource, the more frequently he access the resource. The resource tree is displayed to users at agriculture database granularity. When a user accesses the resource R of a database D in a time TS, Interest (D) of the database D increases. Combined with the user's access frequency Freq (d), the interest of the user on the database D can be gotten according to the following method（Xiaomin Ying,2003）:

$$I=Interest(D)*Freq(D) \quad (1)$$

Interest (D) represents a user's degree of interest for database D. And Freq (D) represents the frequency of a user access to the database D.

But this method simply focuses on the number and frequency of a user's access to resources without the attenuation of the user's interest with time. On this basis, this paper adds time decay for the user interest. Time-attenuation coefficient α can be gotten according to the following formula:

$$\alpha=10^{-\text{passedTime/secsInMonth}} \tag{2}$$

PassedTime represents elapsed time. And secsInMonth is a constant which is 30 days a month, that is 2592000 seconds.

User's actual interest can be derived according to the following formula（Yong Li,2002）:

$$I'=I*\alpha \tag{3}$$

User interest calculated according to the above algorithm corresponds to the resources tree which is displayed after a user logins system. Database accessed more frequently will be enhanced its sort. A user's interest in agricultural information resources can be counted step by step and information push can be achieved implicitly.

3.2 Agricultural Information Resources Push Based on the Heat of Access

For groups of users, sometimes a certain category of information resources is generally concerned. For example "the price of wheat today " is closely related to life. During individual users are searching, this kind of information is easily ignored due to different interest areas of users or different concern in current time. But the information is actually helpful for the user. By mapping multiple users' interest for an information resource, the weight of the information resources can be calculated:

$$I(r)=\sum I(i)*\alpha,\quad i=1,2,3\ldots n \tag{4}$$

According to the weight, the sort of information resources produces the hottest resources sequences. When he retrieves in the agricultural science and technology information database, a user can get not only matching results but also pushed information which is the most accessed by other users. Pushed information may be useful for him. Thus clustering information resources push is achieved.

In view of a larger amount of user group's access, a certain threshold should be set to ensure the effectiveness and timeliness of agricultural information resources recommended. When the access weight of certain type of information resources within a certain time limit T can not reach the specified experience value e, the information resources will be removed from the recommended resource queue to ensure the efficiency and timeliness of information resources recommended queue.

4 The Experimental Results

According to the above techniques and methods, the intelligent retrieval platform of agricultural science and technology information based on user interest model was designed and developed. The platform adopts the user interest model to obtain users' need preferences and pushes the interested information to them. Not only the information meeting the search keywords is retrieved, but also the information which contains the synonyms and related information of the search keywords is also retrieved. Retrieval results are displayed to the user in accordance with his interest priority. By sorting resources dynamically based on user interest, customizing personalized fields and recording accessed information, personalized service can be achieved. As the log of user behavior, user history records are very important for building user model, speculating user behavior and mining user interest. Because user behavior is difficult to speculate and users ' history data are more complicated, a user can not get effectively feedback when he checks his history data. This platform uses a uniform approach- "unified recording, classify browsing" to integrated into users' access records to the user model. The platform collects a user interest in the classification of resources based on each user's behavior. Whenever a user accesses a resource, the system will record his behavior. When he login the system next time, the system will automatically display the frequently accessed database resources to the top row.Besides, he can define the fields in the database as own favorite language or vocabulary in order to make retrieval conditions more in line with his habits. He can also operate his own accessed records to determine whether they are useful. By analyzing the accessed information resources, the system can obtain user interest for a certain type of information resource and improve user interest model. The example showed that the proposed user interest model is better able to push useful resources for users.

5 Conclusion

This paper used an agricultural information push method based on user interest model and presents the combination of automatic modeling and artificial modeling to build user interest model. Further the study proposed the user interest statistics techniques based on the number of access and agricultural information push based on the heat of access. According to the above techniques and methods, the intelligent retrieval platform of agricultural science and technology information based on user interest model was designed and developed to verify the effectiveness of the proposed method.

Acknowledgements. The work is supported by the special fund project for Basic Science Research Business Fee "The demonstration of Tibet agricultural information personalized service system", AII (No. 2012-J-08).

References

1. Yang, X.: Study and Application of Key Technologies for Distributed Agricultural Science and Technology Information Sharing. [Doctoral Dissertation]. Graduate School of Chinese Academy of Agricultural Sciences, Bejing (2011)
2. Barrett, R., Maglio, P.P., Kellen, D.C.: How to Personalize the Web. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI 1997). ACM Press (2007)
3. Wu, Z., Zeng, Q., Hu, X.: Mining Personalized User Profile Based on Interesting Points and Interesting Vectors. Information Technology Journal 6(8), 830–838 (2009)
4. Ying, X.: Study on User Modeling Techniques for Internet personalized service. [Doctoral Dissertation]. National University of Defense Technology, Changsha (2003)
5. Li, Y.: Study and Application of Ontology-based Personalized User Modeling Techniques in Intelligent Retrieval. [Master Dissertation]. National University of Defense Technology, Changsha (2002)

Research on Computer Vision-Based Object Detection and Classification

Juan Wu[1], Bo Peng[1], Zhenxiang Huang[1], and Jietao Xie[2]

[1] College of Information and Electrical Engineering, China Agricultural University, 100083, Beijing, China
[2] Suzhou Institute for Advanced Study, University of Science & Technology of China, 215100, Suzhou, China
{juanwu_cau,pengbo_cau,zhenxianghuang,xjtuxjt}@126.com

Abstract. Computer vision techniques become particularly important in agriculture applications due to their fast response, high accuracy and strong adaptability. Two of the most demanding and widely studied applications relate to object detection and classification. The task is challenging due to variations in product quality differences under certain complicate circumstances influenced by nature and human. Research in these fields has resulted in a wealth of processing and analysis methods. In this paper, we explicitly explore current advances in the field of object detecting and categorizing based on computer vision, and a comparison of these methods is given.

Keywords: Computer Vision, Detection, Classification.

1 Introduction

The interest of this topic in agriculture is motivated by the promise of many applications, such as agricultural and food product inspection [1]. Compared with traditional mechanical technologies, computer vision methods enable more efficient detecting and classifying. For example, categorizing fruit for uniformity can benefit from the advances. Based on computer vision methods can rapid grading fruit quality with respect to color [2], shape, crack and defect, and also have sufficient adaptability to variations, while traditional methods cannot.

Impacted by nature or other complicated factors, agriculture products are distinct with each other greatly. Some issues are often unavoidable, for example, lighting conditions vary over time, image blurs or tilts, or object is partly occluded. These are the reason why the topic is challenging, interesting and worth researching. In this paper, a detailed overview of current main advances in the field is provided. Object detection and classification are discussed independently. Moreover, we focus on color and shape these two salient features for detection.

2 Detection

The literature related to this topic is vast, and we attempt to present two main corresponding groups in the following sections: color-based and shape-based

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 183–188, 2013.

approaches. Shape and color features are two key indicators for fruit and vegetables when they are to be inspected or classified. If agriculture products were improperly fertilized or sprayed with insecticide, deformation or abnormal color will be induced.

2.1 Color-Based Approaches

Color information is the most significant visual feature in agriculture products and it has good performance on invariance of size and view point change. Color based approaches usually aim at segmenting out the desired colors of interest region for further processing. The main limitation with these methods is that they are very vulnerable to nature illumination change.

Color Thresholding. Color thresholding is particularly useful and simple for segmenting image with strong contrast between object and background. By setting the color value range and classifying pixels, thresholding can be done and object can be segmented. When a pixel matching, if its color value is among the reference range, the pixel can be called a required object pixel, if not it will be considered as a background pixel. RGB (Red-Green-Blue) is the most intuitionistic color space. Its advantage is fast enough for real-time performance, due to no need for color space transformation.

HIS or Other Color Space Transformation. HIS (Hue-Intensity-Saturation) color space has good consistency with human vision. Significantly eliminating the influence of brightness changes is its most attractive characteristic, by dividing an overall intensity value from saturation and hue two values. But it will take some time and hardware cost to transform the captured image represented by RGB value into HIS value. CIE model takes human visual characteristics and environmental conditions into consideration, and gets more immune to lighting. However, setting parameters is complicated when applied to specific environment. There are other more color spaces like HSV and L*a*b* [3].

Color Indexing. Color indexing is a fast, straightforward and effective detecting method. Its fundamental principle is comparing the colored objects of two images with their color histograms, as color histogram is used to index the object and store into the template database. The potential problem is its computation will grow enormously in complex outdoor scenes.

2.2 Shape-Based Approaches

We divide shaped-based approaches into two categories: boundary-based approaches and region-based approaches [4]. The former encodes the object's boundary, including Chain code, Fourier descriptor [5], Hough transform [6], spline estimation and bending energy. Many statistic features can be used to describe simple shape region, such as area, projection, Euler number, centrifugal rate, rectangular degrees

and direction. As for complicate region, we could divide the observation into cells, each of which encodes part of the observation locally.

Object shape detection techniques are usually used for further segmentation. Most of them require big computation and attention about position, viewpoint and partial occlusion.

Fourier Descriptor. Fourier Descriptor is the Fourier transform coefficient of the shape boundary curve, which is a very classical method in transform domain. Its main advantages are based on mature theory of Fourier analysis, and invariant to position, rotation and size under certain circumstances. Also, high accuracy can be acquired with only a few low order coefficients. However, it is very sensitive to noise, and influenced by the curve shape and initial point's position.

Hough Transform. Hough transform is originally used for detecting lines and circles, but gradually its related transforms have been generalized to detect arbitrary shapes. The essence of these approaches is the transformation from image space to the defined parameter space. When the results accumulated in parameter space, it can be used to detect or describe the straight lines or curves in the image. This method is inherently robust, as noise, gap, partial deformation and occlusion can be handled [7], but appear a decreased strength. The potential issue is extensive computation and memory hungry.

Moment. Assuming grayscale image normalized as a probability density of bivariate random variable, the random variable can be described by moment. Moment is a kind of linear characteristic. It can be used to describe local features of binary or gray images. This approach in a manner is invariant to image translation, rotation and scale. Usually, moments include invariant moment, Legendre moment Zernike moment and others. Their shortcoming is difficult to present the object's local information and hard for high order moment to directly associate with object's visual features.

3 Classification

Relying on the same features extracted from the object, good classification approaches can get better results. Three commonly used methods exist are described below.

3.1 Template Matching

This method is mainly for searching maximum matching [8]. All required object features to be classified are stored in a database, including considerable variation in performance, viewpoint, and illumination. Each potential object is normalized in size and compared with every template of the same feature. Template matching is fast and easily to modify for new classes. Its potential issue is time-consuming when geometry

transform is complex. The issue can be compensated with genetic algorithm due to its powerful searching ability, quick response and less templates for matching.

3.2 Neural Network

This method is currently one of most successful and widely used learning classification algorithms [9]. Its advantage is that it is tolerant to fault and is powerful to classify. Through training the neural networks with specific color, shape pictogram, features of interest region can be recognition. Its commonly models contain Back Propagation Network, Multilayer Perception Network, Hopfield Network, Radial Basis Network, etc.

3.3 Support Vector Machine

Support Vector Machine extended statistic learning theory to classification, employing Structural Risk Minimization principal to improve generalization capability. It is able to solve problems with respect to nonlinear [10], high dimension, local minimum and also invariable with viewpoint.

4 Discussion

Various commonly detecting and classifying techniques based on computer vision have been presented. A comparison of performances of these techniques is given in Table 1.

Table 1. Comparison of detection and classification approaches

Name	Advantage	Limitation
Color Thresholding	Simple, fast.	Sensitive to illumination.
HIS color Transform	Immune to lighting changes.	The transform will cost computation.
Color Index	Straightforward, quick.	Great computation time in complex scenes.
Fourier Descriptor	Invariant to viewpoint.	Accuracy degrades if curve shape or initial point is not desired.
Hough Transform	Allows slightly shape imperfect or occlusion.	Computationally costly and memory hungry.
Moment	Invariant to viewpoint.	Hard to present local information.
Template Matching	Quick and easily to extent for new classes.	Imperfect shape may pose a problem.
Neural Network	Dynamic handle changes, fast matching.	Demand large training data for good network.
Support Vector Machine	Invariant to viewpoint, allow partial occlusion.	Difficult for multi-category classification and massive samples.

Color and shape are two essential features for detecting. Color-based approaches are fast and straightforward. The shortcoming is vulnerable with illumination changes, not suitable for weak or reflective conditions. Compared to color-based approaches, shape-based methods face more limitations. When detecting object's boundary, they should be immune to changes, shape imperfect, or partial occlusion. But the issue can be compensated with color-based approaches.

Neural network is very popular for classification. Its performance depends on which what network architecture is applied and how well the network is trained. Support Vector Machine is invariant to rotation, size, slightly tilted and even partial occlusion. Template matching is fast and easily to extend with new classes, but computationally costly when geometry transform is complex. However, if it combined with genetic algorithm, its response rate can be improved. Moreover, the hybrid method wouldn't require abundant templates to be prepared carefully.

Some of these approaches are robust but computation hungry, while others are simple but unable to tackle changes. Research in this field is significant and useful, deserving wider attention.

5 Conclusion

A research of the popular object detection and classification based on computer vision has been present in this paper. The research includes description of: (1) existing object detection methods associated with color or shape, (2) important methods developed to tackle the object classification problem, and (3) comparison of performances of these techniques.

References

1. Brosnan, T., Sun, D.W.: Improving Quality Inspection of Food Products by Computer Vision-—a Review. Journal of Food Engineering 61, 3–16 (2004)
2. Lee, D.J., Archibald, J.K., Xiong, G.M.: Rapid Color Grading for Fruit Quality Evaluation Using Direct Color Mapping. IEEE Transactions on Automation Science and Engineering 8, 292–302 (2011)
3. Mendoza, F., Dejmek, P., Aguilera, J.M.: CalibSrated Color Measurements of Agricultural Foods Using Image Analysis. Postharvest Biology and Technology 41, 285–295 (2006)
4. Moreda, G.P., Muñoz, M.A., Ruiz-Altisent, M., Perdigones, A.: Shape Determination of Horticultural Produce Using Two-Dimensional Computer Vision – a Review. Journal of Food Engineering 108, 245–261 (2012)
5. Guo, C.L., Ma, Q., Zhang, L.M.: Spatio-temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform. In: CVPR, Anchorage, Alaska, USA, pp. 1–8 (2008)
6. Maji, S., Malik, J.: Object Detection using a Max-Margin Hough Transform. In: CVPR, Miami Beach, Florida, pp. 1038–1045 (2009)
7. Murillo-Bracamontesa, E.A., Martinez-Rosasb, M.E., et al.: Implementation of Hough Transform for Fruit Image Segmentation. Procedia Engineering 35, 230–239 (2012)

8. Ren, F.X., Huang, J.S., Jiang, R.Y., Klette, R.: General Traffic Sign Recognition by Feature Matching. In: IVCNZ, pp. 409–414 (2009)
9. Torres, M., Hervás, C., García, C.: Multinomial Logistic Regression and Product Unit Neural Network Models: Application of a New Hybrid Methodology for Solving a Classification Problem in The Livestock Sector. Expert Systems with Applications 36, 12225–12235 (2009)
10. Li, Z.F., Tang, X.O.: Bayesian Face Recognition Using Support Vector Machine and Face Clustering. In: CVPR, Washington, DC, vol. 2, pp. 374–380 (2004)

Automatic Detection of Kiwifruit Defects Based on Near-Infrared Light Source*

Pingping Li, Yongjie Cui**, Yufeng Tian, Fanian Zhang, Xiaxia Wang, and Shuai Su

College of Mechanical and Electronic Engineering, Northwest A&F University,
Yangling Shaanxi, 712100, P.R. China
{lipingping,cuiyongjie,tianyufeng,zhangfanian,
wangxiaxia,sushuai}@nwsuaf.edu.cn

Abstract. A mathematical model that expresses the relationship between Near-infrared light intensity and automatic threshold for automatic kiwifruit surface defect detection was established. By applying different levels of Near-infrared light intensity to machine vision system, 268 images were collected. Then the images were processed with MATLAB using the method to detect kiwifruit defects based on Near-infrared light source .The obtained 268 sets of data on Automatic Threshold T_0 and Manual Threshold T_1 were divided into 19 groups according to different aperture and light intensity. After processing data, a series of linear equations about the relationship between Near-infrared light intensity and Automatic Threshold T_0, with function fitting coefficient of $R^2 > 95\%$ was obtained. Finally, relationship between T_0 and T_1 was analyzed according to the effectiveness of image processing results and constant P was introduced to revise Automatic Threshold T_0. Thus, a mathematical model needed to gain kiwifruit defects detection threshold, namely Model Threshold T, was established.

Keywords: Image processing, Near-infrared light intensity, Automatic threshold, Manual threshold, Model threshold, Linear relationship.

1 Introduction

The kiwifruit, or often shortened to kiwi in many parts of the world, is the edible berry of a cultivar group of the woody vine Actinidia deliciosa and hybrids between this and other species in the genus Actinidia. It's nutritious and has high medicinal value [1]. China is one of the major areas in producing kiwi. Currently, kiwifruit postharvest sorting processing is still performed manually and its surface defects judgment depends on human completely. Standard Sphere Method [2-5] which uses machine vision technology to separate the fruit surface defects has achieved good results. However, with the traditional RGB [6][7], CCD imaging systems [8-11], the angle between the camera and edge light reflection direction of sphere and ellipsoid fruits remains very large. According

* Foundation item: The Project-sponsored by SRF for ROCS, SEM (KS08021101), Project supported by the National Natural Science Foundation of China (61175099), Northwest Agriculture and Forestry University Talent Fund (Z111020902).

** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 189–198, 2013.

to Lambertian light laws of reflection, the fruit edge and surface defects both have lower gray level. Therefore it is difficult to detect and distinguish the defect. In addition, there is a challenge in uniform illumination with the RGB and CCD systems. To overcome these difficulties, near-infrared hyper spectral is used by many research institutes to test external qualities of agriculture products such as maturity of strawberry [12], bruises of strawberry [13] [14], apple [15] and chestnut [16] defects etc.

There are two assumptions in Standard Sphere Method. First assumption is that the shape of fruit is either standard sphere or ellipsoid. Second assumption is that light illumination in machine vision field is uniform. Under these conditions, optical system consisting of lens and camera can be a linear system [17]. In the case of kiwifruit, its shape is quasi-ellipsoid. With the Near-infrared light source, the issue of fruit surface reflective area can be overcome as it plays the role of uniform illumination according to preliminary studies [18][19]. This paper aims at collecting images of defective kiwifruit with different light intensity by adjusting the Near-infrared light intensity, then processing images by detection methods for kiwifruit surface defects grading, and finally analyzing the relationship between binarization threshold of kiwifruit defects and Near-infrared light intensity, to establish a mathematical model.

2 Material and Method

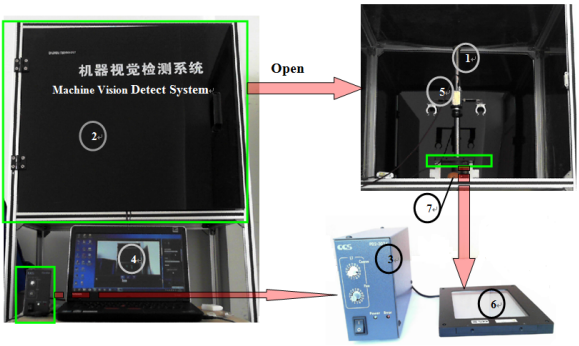
2.1 Test Materials and Equipment

Test kiwifruit samples were “Qin Mei” cultivars. They were bought from Xizhai Village, Qinghua Country, Mei County, Shaanxi Province. All of the fruits (total 182 samples) were picked up on site and kept in cartons. Weight of a single fruit ranged from 69.6g to 224.0g. Tests were carried out at College of Mechanical and Electronic Engineering, Northwest A&F University.

Firstly the fruits were classified; according to the type of defects such as sunburn, parasitic spot. Then intact fruits were scratched to make them defective.

In terms of equipment, lamp house was made of black organic plastic material to avoid external light interference. The camera used was DALSA CCD camera (matching dedicated image acquisition and debugging software Sapera Cam Expert) and the lens was Camera FUJINON HF16HA-1B (aperture range 1.4~16.0). PC used to save data after image processing was Lenovo ThinkPad E420. LFX2-100IR850 Near-infrared light source and dedicated power supply PD-3024-K1320-014 with Light source extension line (1m, 24V) also consisted the test equipment. As a background a white soleplate was provided to enhance the contrast of the background and kiwifruit. The distance between camera FUJINON HF16HA-1B and kiwifruit was 370mm and distance between Near-infrared light source and kiwifruit was 110mm, as shown in Fig. 1.

LFX2-100IR850 Near-infrared light source needed to be attached to the light source dedicated power supply CCS PD-3024- K1320-014 with a light source extension line (1m, 24V).There are 16 brightness coarse adjustment gears and 16 fine adjustment gears respectively. Fig.1 shows pictures of Near-infrared plane illuminator LFX2-100IR850 and light source dedicated power supply CCS PD-3024- K1320-014. The light source needed to be warmed up for 30 minutes before usage.



1. Lifting frame 2. Lamp house 3. Light intensity regulator
4. PC 5.CCD Camera 6.Near-infrared light source
7. Kiwifruit

Fig. 1. Kiwifruit Machine Vision Detect System

2.2 Research Program

As shown in Fig.2, the left dotted line box shows the image processing of kiwifruit defects detection [18]. Although this method is able to extract the fruit surface defects, there are two problems exist:

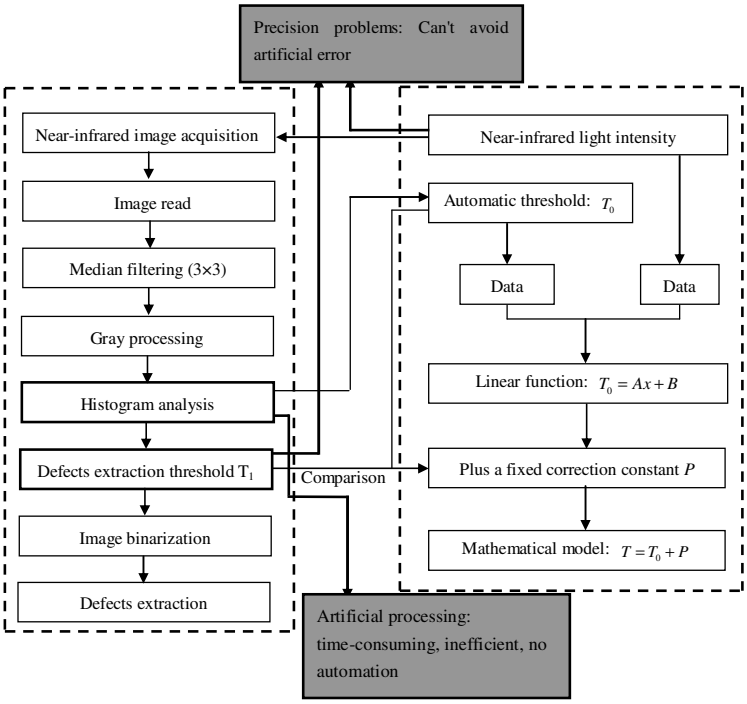


Fig. 2. The Block Diagram of Research

Problem 1: It's time-consuming to extract the image parameters of fruit surface defects manually and the automatic detection of fruit defects is a mainstream in the development of nondestructive testing of the fruits.

Problem 2: As shown in Fig.2, the automatic defection threshold T_0 will produce error and it will influence the accuracy of fruit defection.

To extract the fruit defects threshold automatically, this paper proposed a method based on Near-infrared light intensity with Automatic Threshold T_0 and thus established a mathematical model. The dotted line box on the right is the solution diagram. As T_1 obtained by manual analysis will produce error. Therefore, Automatic Threshold T_0 was introduced to solve the problem, where T_0 is automatically generated by the image processing program and it reduces the systematic errors effectively. T_0 cannot be used to detect defect directly. After comparing T_1 with T_0, a correction constant P was introduced to revise T, as $T= T_0+P$.

2.3 Test Methods and Procedures

Scratch wound is one of kiwifruit surface defects which is the most difficult defect to be extracted. So we took kiwifruit with scratch wound as objects to establish the relationship between Near-infrared light intensity and scratch defects and to determine the automatic threshold of surface defects.

Table 1. Test conditions and image collection

Aperture (1.4~16.0)	Coarse adjustment (1~16)	Fine adjustment (1~16)	Image quantity
1.4	1	3~11	9
3.7	2、3	1~16	32
4.0	3、4	1~16	32
6.0	5~10	1~16	96
8.0	5	1~16	16
8.0	6~16	1、9、16	33
12.0	6~16	1、9、16	32
16.0	8~16	1、16	18
Total			268

Test Procedures

(1) Put the scratch kiwifruits into lamp house; adjust the camera aperture according to Tab.1. Adjust the brightness coarse adjustment gears and fine adjustment gears in accordance with Tab.1. Finally collect images using supporting software Sapera Cam Expert and adjust R, G, B to 0 dB.

(2) Number the collected images and light intensity record mode. E.g. coarse adjustment gear 1 and fine adjustment 1~16 are recorded as 101,102,103, 104…116, while coarse adjustment gear 1~16 and fine adjustment 1 are recorded as 1601, 1602, 1603, 1604…1616.

(3) Process the numbered images using the method described in [18], and record T_0, T_1 and image processing results.Tab.2 is part of the test data records.

(4) Change the Near-infrared light intensity from 103 to 111, gain T_0 using the method as mentioned in Step **(3)**, and adjust the threshold interval to gain T_1. It is obtained by adding a constant P to T_0, as shown in Tab.4.

(5) Finally, analyze and process the data with Microsoft Excel. Although the final threshold used in image processing is defect detection threshold, the mathematic model expresses the relationship between T_0 and Near-infrared intensity. The reason is that T_1 is a manual selected threshold according to the effectiveness of image processing results. If T_1 is selected to analyze data, it will increase the error. Thus T_0 obtained by MATLAB is used to establish the exact relationship between threshold and light intensity. Then analyze the relationship between T_1 and T_0. Finally introduce a correction constant P to revise T_0.

Table 2. Part of the test data

No.	Aperture	Light intensity	Automatic Threshold T_0	Manual Threshold T_1
001	1.4	103	0.0941	0.1400
002	1.4	104	0.1333	0.2400
003	1.4	105	0.1686	0.2590
004	1.4	106	0.2039	0.3590
005	1.4	107	0.2353	0.4000
006	1.4	108	0.2706	0.4400
007	1.4	109	0.3020	0.5300
008	1.4	110	0.3333	0.5533
009	1.4	111	0.3686	0.6700

3 Results and Discussion

3.1 Mathematical Model

The data were analyzed using Microsoft Excel, and the data in Tab.2 were used to produce the function graph in Fig.3. The intensity of Near-infrared was from 103 to 111, and the automatic threshold interval was from 0.0900 to 0.4000. Fig.3-a is the linear diagram of T_0 and light intensity, in Tab.3, we established the equation T_0=0.0339x-3.3884, R^2=0.9992. Fig.3-b is the linear diagram of T_1 and the light intensity and we can find out the equation T_1=0.0614x-6.1688, R^2=0.9844.

The function fitting degree in Fig.3-a is higher than that in Fig.3-b. This is because the T_0 is free from the human error of T_1.

Similarly, we processed the 19 groups' data and the results were recorded inTab.3.

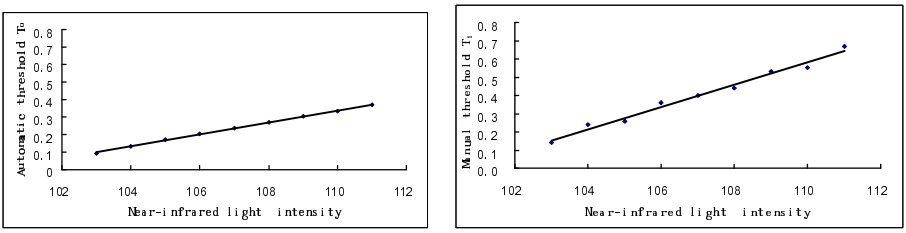
As a result, the function relation graphs between Near-infrared intensity and T_0 was gained. It was linear. All of the 268 groups of data were processed by this method, the results were shown in Tab.3, with $R^2 > 95\%$. Thus the relation between Near-infrared intensity and T_0 can be defined as follows:

$$T_0 = Ax + B\,. \tag{1}$$

Where: T_0, x, A, B stand for automatic threshold, Near-infrared intensity, light intensity coefficient, Near- infrared fine adjustment influence coefficient, respectively.

（1） In Tab.3, the functions from 01 to 11 were gained under the same aperture and light intensity with both brightness coarse and variable fine adjustment. They showed linear relationship. In addition, value A and B are variables. Value A increased as the light intensity enhanced, while value B had the opposite trend.

（1） The functions from 12 to 19 were gained under the same aperture and light intensity with brightness fine adjustment and variable coarse adjustment. They also showed linear relationship. In addition, value A didn't change as the light intensity enhanced, while value B increased with light intensity enhancing.



a **b**

Fig. 3. The relation between light intensity and T_0、T_1(Data in Tab. 2)

（2） Considering that T_0 was not used to detect the fruit surface defects directly, it was necessary to define the correction constant P value. The relationship between Near-infrared intensity and defection threshold was defined as follows:

$$T = T_0 + P \,. \tag{2}$$

This is because the T_0 is free from the human error of T_1.

Where: T, T_0, P stand for defects threshold, automatic threshold, threshold correct constant, respectively.

Table 3. The relation between light intensity and automatic threshold

No.	Aperture	Coarse adjustment	Fine adjustment	Function (T_0=Ax+B)	R^2
01	1.4	1	3~11	T_0=0.0339x - 3.3884	0.9992
02	3.7	2	1~16	T_0= 0.0057x - 1.0295	0.9975
03	3.7	3	1~16	T_0= 0.0063x - 1.7055	0.9985
03	4.0	3	1~16	T_0=0.0050 x - 1.3321	0.9976
04	4.0	4	1~16	T_0=0.0052 x - 1.8516	0.9935
05	6.0	5	1~16	T_0= 0.0027x - 1.1500	0.9882
06	6.0	6	1~16	T_0= 0.0026x - 1.3268	0.9893
07	6.0	7	1~16	T_0= 0.0027x - 1.6486	0.9865
08	6.0	8	1~16	T_0= 0.0028x - 1.9274	0.9927
09	6.0	9	1~16	T_0= 0.0029x - 2.2458	0.9930
10	6.0	10	1~16	T_0= 0.0027x - 2.2822	0.9808
11	8.0	5	1~16	T_0= 0.0011x - 0.4516	0.9545
12	8.0	6~16	1	T_0= 0.0002x - 0.0209	0.9995
13	8.0	6~16	9	T_0= 0.0002x - 0.0109	0.9987
14	8.0	6~16	16	T_0= 0.0002x - 0.0029	0.9993
15	12.0	6~16	1	T_0= 0.0001x +0.0099	0.9980
16	12.0	6~16	9	T_0=0.0001x + 0.0129	0.9981
17	12.0	6~16	16	T_0=0.0001x + 0.0077	0.9929
18	16.0	8~16	1	T_0=0.00003x+0.0210	0.9634
19	16.0	8~16	16	T_0=0.00003x+0.0291	0.9867

Table 4. The correction constant P for surface defects extraction threshold

T_0	$T_0<0.1$	$0.1\le T_0<0.2$	$0.2\le T_0<0.3$	$0.3\le T_0<0.4$	$T_0\ge 0.4$
P	0.0400	0.1000	0.1600	0.2200	0.2800

3.2 Results and Discussion

T_0 was obtained by processing the 268 images and divided into five intervals. One image was selected from each interval for the manual threshold, and the results are shown in Fig.4. Fig.4-a is a Near-infrared original image, Fig.4-b is a threshold binary image of T_0 and Fig.4-c is the binary image of T_1, Fig.4-d is the binary image of Model Threshold T.

(1) The image processing results are shown in Fig.4 which mainly include flaws and scratch. This processing the reflection of light effectively and simplify the image processing based on Near-infrared light to detect surface defects of kiwifruit. Using P and interval of T_0 could detect the defects of fruit. There was some difference in the extent of noise influence. Results showed that interval of T_0 value from 0.1 to 0.2 had better effects than those of other intervals. Thus mathematical morphological operation was used to remove the noises.

(2) Using T_0 alone does not detect the surface defects for kiwifruit from Near-infrared image; however, it provided the precondition to quantify the standard of image segmentation reasonably, which meant that we could process different brightness images under the same standard. In Fig.5, for example, the binary images of T_0 were almost the same though different intervals were selected.

(3) Compared with the results of image processing in [18], the results in this paper were much better as it overcame the influence of fruit calyx.

(4) Fig.4-a is Near-infrared original image. Fig.4-d is binary image of T. Fig.4-c is binary image of T_1. When two images were compared, both detected the fruit surface defects. Therefore to make further comparison, the pixel values of the two binary images were calculated. As for Tab.5, Where: $p_{i1}$is pixel value of Fig.4-d and p_{i2} is pixel value of Fig.4-c.

Table 5. The pixel values p_{ij} for surface defects extract threshold

T_0	$T\ (T=T_0+P)$	T_1	p_{i1}	p_{i2}
0.0784	0.1184	0.1180	2789	2789
0.1490	0.2490	0.2450	1458	2214
0.2745	0.4345	0.4305	7275	5521
0.3765	0.5965	0.6050	7022	8459
0.4039	0.6839	0.7000	3716	5450

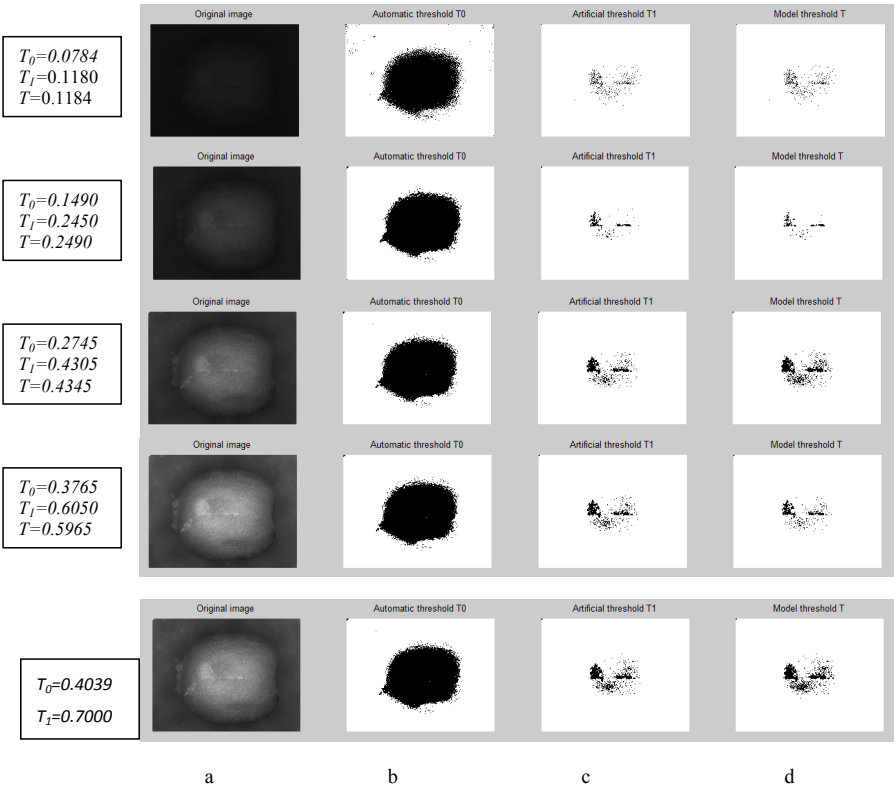


Fig. 4. Five image processing results

4 Conclusion

The function of Near-infrared light intensity and automatic threshold showed obvious linear relationship, with function fitting coefficient of $R^2 > 95\%$. A linear mathematical model was developed to detect threshold of kiwifruit surface defects automatically.

In the five intervals of T_0, the Near-infrared image processing excluded the interference of the fruit calyx and therefore this method was more effective to detect fruit defects of surface flaws and scratches, especially with T_0 range of 0.1 to 0.2.

Overall our data provide a promising tool for automatic defects detection in the fruit grading system.

Acknowledgments. This research was funded by The Project-sponsored by SRF for ROCS, SEM (KS08021101), Project supported by the National Natural Science Foundation of China (61175099) and Northwest Agriculture and Forestry University Talent Fund (Z111020902).

References

1. Rashidi, M., Seyfi, K.: Classification of Fruit Shape in Kiwifruit Applying the Analysis of Outer Dimensions. Int. J. Agric.Biol. 5, 759–762 (2007)
2. Tao, Y., Wen, Z.: An Adaptive Image Transform for High-speed Fruit Defect Detection. Transactions of the ASAE 42(1), 241–246 (1999)
3. Feng, B., Wang, M.: Study on Identifying Measurement about Default of Fruit in Computer Vision. Journal of China Agricultural University 7(4), 73–76 (2002)
4. Fu, F., Ying, Y.: Gray Level Transform Model of Ball Image and Its Application in Citrus Image Correction. Transactions of the Chinese Society of Agricultural Engineering 20(4), 117–120 (2004)
5. Ying, Y., Fu, F.: Color Transformation Model of Fruit Image in Process of Non-destructive Quality Inspection Based on Machine Vision. Transactions of the Chinese Society for Agricultural Machinery 35(1), 85–89 (2004)
6. Cheng, F., Ying, Y.: Inspection of Mildewed Rice Seeds Based on Color Feature. Transactions of the CSAE 35(4), 102–105 (2002)
7. Zhu, W., Cao, Q.: Defect Segmentation of Tomatoes Using Fuzzy Color Clustering Method. Transactions of the Chinese Society of Agricultural Engineering 19(3), 133–136 (2003)
8. Pang, J.: Study on External Defects Classification of Navel Orange Based on Machine Vision. Zhenjiang University, Hangzhou (2006)
9. Yang, F., Zhu, S., Qiu, Q.: Prickly Ash Appearance Quality Detection Based on Computer Vision and its Implementation in MATLAB. Transactions of the CSAE 24(1), 198–202 (2008)
10. Diaz, R., Gil, L., Serrano, C., Blasco, M., Moltó, E., Blasco, J.: Comparison of Three Algorithms in the Classification of Table Olives by Means of Computer Vision. Journal of Food Engineering 61(1), 101–107 (2004)
11. Blasco, J., Aleixos, N., Moltó, E.: Computer Vision Detection of Peel Defects in Citrus by Means of a Region Oriented Segmentation Algorithm. Journal of Food Engineering 81(3), 384–393 (2007)
12. Nagata, M., Jasper, G.T., Taiichi, K., Cui, Y., Yoshinori, G.: Predicting Maturity Quality Parameters of Strawberries Using Hyperspectral Imaging. In: ASAE/CSAE Annual International Meeting, 2004 (2004)
13. Nagata, M., Jasper, G.T., Taiichi, K., Hiroshi, T.: NIR Hyperspectral Imaging for Measurement of Internal Quality in Strawberries. In: ASABE Annual International Meeting 2005 (2005)
14. Jasper, G.T., Nagata, M., Taiichi, K.: Detection of Bruises in Strawberies by Hyperspectral Imaging. In: ASABE Annual International Meeting 2006 (2006)
15. Zhan, H., Li, X., Wang, W., Wang, C., Zhou, Z., Huang, Y.: Determination of Chestnuts Grading Based on Machine Vision. Transactions of the CSAE 26(4), 327–331 (2010)
16. Wen, Z., Tao, Y.: Dual-camera NIR/MIR Imaging for Stem-end/Calyx Identification in Apple Detect Sorting. Transactions of the ASAE 43(2), 449–452 (2000)
17. Li, J., Xue, L.: A Study on Navel Orange Grading System Based on Computer Vision. Jiangxi Agricultural University 28(2), 304–307 (2006)
18. Li, P., Ding, X., Su, S., Cui, Y.: A Method for Surface Defects Detection in Kiwi Fruit Classification. CSAE (2011)
19. Li, Q., Wang, M.: A Fast Identification Method for Fruit Surface Defect Based on Fractal Characters. Journal of Image and Graphics 5(2), 144–148 (2000)

The Fractal Dimension Research of Chinese and American Beef Marbling Standards Images

Jianwen Chen, Meiying Liu*, and Li Zong

College of Engineering, Huazhong Agricultural University, Wuhan, Hubei Province 430070

Abstract. Beef marbling level is the most important indicators in the evaluation of beef quality. The fractal dimension is closely related to marbling level. In this paper, the theory of fractal dimension is used to analyze beef marbling standards images in China and USA. After comparing several different dimension calculation method, the final method is improved box-counting dimension. Linear regression model of dimension calculation value with this method and marbling level is built. The model results are satisfactory through examination. For further use of samples, this model settles the foundation to establish the grade evaluation methods.

Keywords: marbling, longissimus dorsi, fractal dimension, beef quality, standards images.

Introduction

Assessment of beef quality, is highly valued in foreign countries. Developed countries proposed grading standards earlier, and generated significant economic benefits. China had a late start in this area, and began research in 2000. The assessment of beef quality, marbling quality is the most important indicator. Marbling usually refers to the section pattern of cattle at the 11-13 sternocostal or 5-7 sternocostal fat department [1] of the longissimus dorsi. Marbling is divided into six levels in USA, from 3 to 8, the quality is getting better and better. Marbling level of 1 and 2 beefs are chopped to use in tin system, do not directly sell. In China, marbling is divided into four levels. From 1 to 4, the quality is getting worse. Beef quality grading has a significant economic effect. Different quality grades of beefs have the difference prices. However, in China and abroad, the main method of beef grade assessment is manual measurement and artificial sensory evaluation, which have low efficiency and error shortcomings.

Fractal theory was proposed by Mandelbrot. The fractal dimension is important object in fractal theory research [2]. It is a number which describes the complexity of fractal collection. Recently the fractal theory is applied to the assessment of marbling in China and abroad. Some scholars expressed their views, and had mixed results.

* Supported by the fundamental research funds for the central universities, program No.52902-090020/199.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 199–209, 2013.

This paper carries out the corresponding analysis of BMS (beef marbling standards) images in China and America. Specific analysis indicators are the box-counting dimension[3], differential box-counting dimension, information dimension. Firstly, the background[4]of beef eye muscle image is removed through using the knowledge of machine vision, then extracting the longissimus dorsi. The fractal dimension of longissimus muscle is increased with the improvement of the marbling quality. After acquiring dimension of different grades marbling, regression mathematical model is established to distinguish the grade.

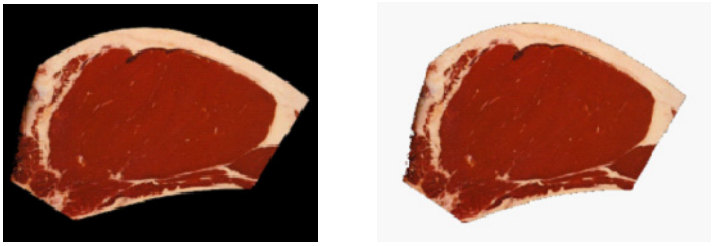
1 Image Processing

To get rid of some interference information , background of the beef image should be removed .Then, by the method of binarization, mathematical morphology, sent shadow method, etc, the fat and proud flesh which surround beef eye muscle[5] can be removed. Sometimes part of the proud flesh and longissimus muscle contact too close to be divided , because artificial boundaries can open them, so achieving the ultimate goal would be no problem.

1.1 Background Removed

Background removal is benefit to extract the longissimus muscle[6]. Threshold method is used. The background is black, while the beef is red and white color, so the difference is significant. Standards images are only two major categories of the target and background. So only to select a threshold, it is called single threshold [7] split. This approach let the gray value of each pixel in the image to compare with threshold. The gray value of pixels which are greater than the threshold is for a class, and the gray value of pixels which are less than the threshold value is to the other class. Significant color is different [8] at the background and beef, so selecting a appropriate threshold[9], the background can be removed.

Fig.1 are the third grade image of USA, and the effective image of background removal. To make the effect more obvious, the background is set to white.



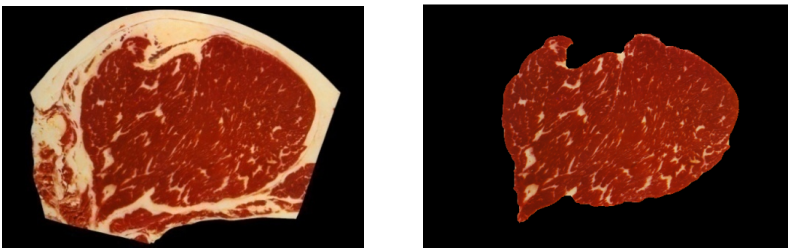
(a). The level 3 figure (b). Background eliminate

Fig. 1. Third grade image of America removes background

1.2 The Extraction of Longest Back Muscle

The peripheral of beef eye muscle is the fat and proud flesh[10].To extract the longissimus dorsi, they should be get rid. The image type conversion, mathematical morphology[11], the mark function, properties function, sent shadow method, image restoration, etc ,are used [12] to get the longissimus muscle.

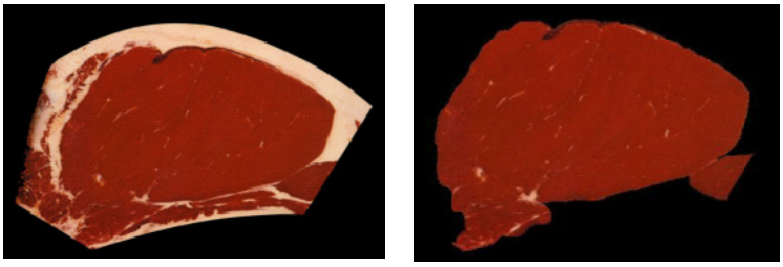
Fig.2 show the American seventh level image and processed result.



(c). American 7th level image (d). Longissimus muscle

Fig. 2. The American 7th level image and processed result

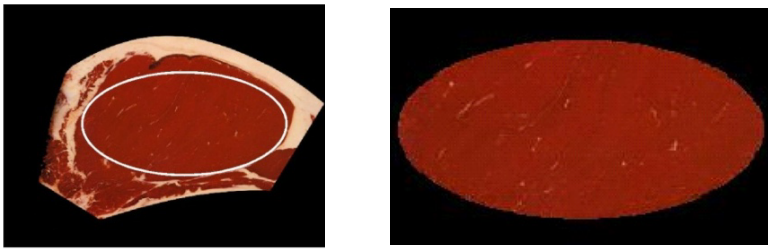
For some images, some proud flesh and longissimus muscle is too tight[13] to be separated. For example, Fig.3 show third level image and processed result. This requires other methods ,so several methods are used to compare for the best.



(e). Level 3 original (f). Longissimus muscle diagram with proud flesh

Fig. 3. The third level image and processed result

The shape of longissimus muscle likes an oval. Tring to draw a white oval with right size to be the artificial boundaries, the effect is well as shown in Fig.4-(h).



(g).To add a ellipse (h). Processed results

Fig. 4. Adding ellipse and then extracting the longissimus dorsi muscle

Fig.4- (h) has no proud flesh, but longissimus muscle area decreases.

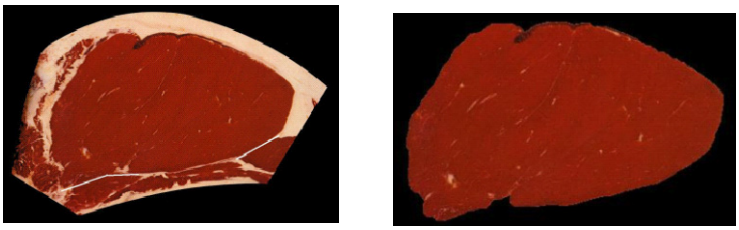
The iterative morphology method is also used. Firstly, let Fig.3-(f) be binary. In addition , structural element of "disk" [14] is used. Figure erodes and dilates four times, and then using a “for” loop program[15] which makes the image to recruit into RGB diagram.

Fig.5 shows the effect. However, Fig.5 still has a few proud flesh which is not removed, and a few part of longissimus muscle is split off. It is better than Fig.4-(h).Although, the best method is needed.

Fig. 5. Iterative corrosion and expansion to image

Fig.4-(g) paints oval to remove proud flesh, but longissimus muscle is lost a few. To further improve the method, if at the close connection of proud flesh and longissimus muscle, drawing a curve to be artificial boundaries, the effect may be well.

On the third level image of the United States, at the closely connection of proud flesh and longissimus dorsi, two white curves are drawn. Then dealing with separately [16], the proud flesh is removed, as shown in Fig.6.



(i) .Add curves (j). Processing results

Fig. 6. Drawing curves and then extracting the longissimus muscle

Some standards images, at the departments of proud flesh and longissimus muscle connecting closely, also do this processing, the result is quite good too. Fig.7 are longissimus muscle diagrams of Chinese second and third grade images,for example.

(k). Longissimus muscle of Level 2 (l). Longissimus muscle of Level 3

Fig. 7. Longissimus muscle diagrams of Chinese standards images.

2 The Application of Fractal Dimension

Four kinds of fractal dimension calculation methods and MATLAB software[17] are used. Method 1 is traditional box-counting dimension algorithm. Method 2 is the differential box-counting algorithm. Method 3 is information dimensional algorithm. Method 4 is improved box-counting dimension algorithm. These four methods are with a number of step r_i to divide the grids that cover the longissimus muscle image. If fat particles are in the mesh[18], counting them and summing to get Nr, then marking point ($\log(1/r_i)$, $\log(Nr_i)$) on a double logarithmic coordinates, fitting them with a straight line, finally absolute value of the line's slope is dimension value.

2.1 Compare Fractal Dimension Calculation Method

2.1.1 Method One Is Traditional Box-Counting Dimension Algorithm

The calculation of the box-counting dimension, intuitive understanding, is counting the number of lattice [19]. Binary image is covered with small square box of different side length, and different side length of the small square box to cover it, the box number is also different. Side length r and a total of not empty small box N(r), meet

the relationship test: D_b =-lg N (r) / lg (1/ r). D_b is the box dimension. Scale size r is usually 2^n. A series of non-empty box number are acquired in the different proportion sizes of r. The reciprocal of r, and the number of these non-empty boxes, are set in double logarithmic coordinate[20], by least squares linear to fit them, then the absolute value of slope is the box dimension . Table 1 is results of this method to obtain.

Table 1. Results of traditional box counting dimension method for calculating values

American level	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
Fractal dimension	1.2222	1.2324	1.2343	1.2185	1.2314	1.2204
Chinese level	Level 1	Level 2	Level 3	Level 4		
Fractal dimension	1.2719	1.2763	1.2714	1.2751		

This method is time-consuming[21] a bit long, about three seconds. The law of increasing dimension value with the high quality of marbling is less obvious.

2.1.2 Method Two Is Differential Box-Counting Dimension Algorithm

Differential box-counting algorithm[22] has three dimensions. The third dimension is the image gray. Let M × M size image divided into S × S sub-block (M / 2 ≥ S > 1, S for integer), and r = S/M. To imagine the images into surfaces[23] of three-dimensional space, x, y represents plane position, and z-axis represents the gray value. X-Y plane is divided into a lot of s × s grid. In each grid is an s × s × s box. For different r and the calculation of non-empty box N_r, using the least squares linear to fit, the fractal dimension D could be obtained. The calculation results are shown in Table 2.

Table 2. Results of differential box-counting dimension method to calculate values

American level	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
Fractal dimension	2.1263	2.1253	2.1450	2.1250	2.1352	2.1322
Chinese level	Level 1	Level 2	Level 3	Level 4		
Fractal dimension	2.2983	2.2485	2.1709	2.1139		

The differential box-counting algorithm, to the Chinese images, the values are well; but the images to the United States, less than ideal, the law that the fractal dimension value increments with high quality of marbling is not very obvious.

2.1.3 The Third Method Is the Information Dimension

This method is similar to traditional box-counting dimension algorithm[24]. Let N be the total fat information elements, N_i is fat information on the number of elements contained in each cover, the probability that fat distribution of information in each

coverage is: $P_i = N_i / N$. Amount of fat information is $I_i = -P_i \ln P_i$. The amount of fat information is I (r). Changing the scale r, I (r) with the 1 / r meet the relationship test: $I(r) \propto (1 / r)^D$. D is its information dimension. Results are shown in Table 3.

Table 3. Results of information dimension method to calculate values

American level	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
Fractal dimension	1.7516	1.7479	1.7492	1.7500	1.7521	1.7624
Chinese level	Level 1	Level 2	Level 3	Level 4		
Fractal dimension	1.6978	1.6834	1.6667	1.6942		

Information dimension method, its effect is not very satisfactory. The law[25] that incremental dimension value with the high quality of marbling is also not very obvious.

2.1.4 Method Four Is an Improved Box-Counting Dimension

For the above three methods, the values which are obtained are not very satisfactory. Method four is proposed: an improved box-counting dimension. It is used in the binary image of the longissimus dorsi. Method four has improvements and advantages as follows.

2.1.4.1 Square (box) side lengths are with a lot of data, meaning that the mesh number of times is a lot, not just the 2 ^ n. It can help improve data accuracy.

2.1.4.2 When the image size could not be a square side length divisible, then rounding the excess with only small "margins" section, thus reducing outside interference.

2.1.4.3 It gives value "1" pixel of binary chart, accounting for the entire map of the area proportion. This reflects objectively the number of marble pattern.

2.1.4.4 D numerical is strong regularity that they show an increasing trend with marbling levels increase.

2.1.4.5 Calculation of the dimension values is short time-consuming. Evaluating the code, the results show out immediately.

The calculation results of method four are shown in Table 4.

Table 4. Results of improved box-counting dimension method to calculate values

American level	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
Fractal dimension	1.8159	1.8242	1.8575	1.8646	1.8754	1.8960
Chinese level	Level 1	Level 2	Level 3	Level 4		
Fractal dimension	1.9630	1.9521	1.9168	1.8819		

This method is time-consuming short. When the program is evaluated, the results show out immediately. In the images of China and the United States, the values obtained increase with the quality of marbling became high.

After comprehensive comparison, method 4, an improved box-counting dimension method is the best, so use it to judge the grade.

2.2 Calculating and Testing of Improved Box-Counting Dimension Method

Detailed results of calculation and test are in Table 5 and Table 6. Exel is used to test significance of calculated data through method 4. That is to view R^2 value of fitting equation.

Table 5. Fitting results of American images' fractal dimension value

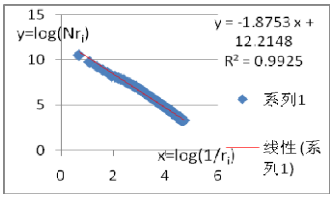
American level	Fitting equation	Dimension	R^2
3	y = -1.8159 x + 11.9629	1.8159	0.9894
4	y = -1.8242 x + 11.9983	1.8242	0.9899
5	y = -1.8578 x + 12.1418	1.8578	0.9914
6	y = -1.8646 x + 12.1702	1.8646	0.9919
7	y = -1.8753 x + 12.2148	1.8753	0.9925
8	y = -1.8960 x + 12.3010	1.8960	0.9936

Table 6. Fitting results of Chinese images' fractal dimension value

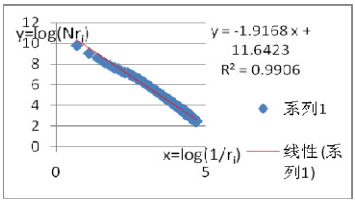
Chinese level	Fitting equation	Dimension	R^2
1	y = -1.9630 x + 11.8332	1.9630	0.9936
2	y = -1.9521 x + 11.7887	1.9521	0.9927
3	y = -1.9168 x + 11.6423	1.9168	0.9906
4	y = -1.8947 x + 11.5479	1.8947	0.9899

Seeing from the tables, the fitting equations are obvious in significance level.

Two images'fitting maps are giving out. Fig.8 are the fitting diagram of the seventh grade of USA, and third grade of China.



(m). Fitting figure of American seventh grade image



(n). Fitting figure of Chinese third grade image

Fig. 8. Two standards images fit in diagram

3 Marbling Grade of Mathematical Model

3.1 Mathematical Model of American Images

Assuming that the regression equation is: T = a + b * D .Level value T with the box-counting dimension value D, as well as the T^2, D^2, D * T, etc[26], are used. Through regression analysis, a regression equation is established about the American images:

$$T=-105.9740+60.0746D \tag{1}$$

T is level value, and D is fractal dimension value in equation (1).Through F-test, the equation is significant at level of α=0.05.

3.2 Test of Equation (1)

Results of the equation (1) are examined, only validation of level 5 is mistaken. The accuracy rate is 83.33%, as shown in Table 7.

Table 7. Level verification of equation (1)

The actual level	Calculated value	Round
3	3.1155	3
4	3.6141	4
5	5.6146	6
6	6.0411	6
7	6.6899	7
8	7.9274	8

3.3 Mathematical Model of Chinese Images

Similarly, through regression analysis, combined with the corresponding value, regression equation is established:

$$T=68.7368-34.3471D \tag{2}$$

T is level value, and D is fractal dimension value in equation (2).Through F-test, the equation is significant at level of α=0.05.

3.4 Test of Equation (2)

Results of the equation (2) are examined.The accuracy rate is 100%, as shown in Table 8.

Table 8. Level verification of equation (2)

The actual level	Calculated value	Round
1	1.3134	1
2	1.6876	2
3	2.9003	3
4	4.0990	4

4 Conclusion and Suggestion

This study, first of all is the standards images' processing that to extraction longissimus muscle. To figures of some proud flesh and longissimus dorsi closely connected, three image segmentation methods are compared. The optimal method is to add curve manually.

For Chinese and American images, four fractal dimension methods are used. After comparison, the final choice is improved box-counting dimension algorithm. Dimensions obtained through this method are regularity, that the D values become larger with increased beef marbling levels.

Using regression equations to establish a linear mathematical model, the effect is well. Mathematical model of American images, the accuracy rate is 83.33%; to Chinese images, the accuracy rate is 100%.

BP network, and support vector machine modeling methods can be tried to compare the classification results. Some rounded values, although are correct, but they are close to the median. Trying other mathematical models, the accuracy may be higher.

References

1. National beef grading methods and standards-NY / T. 676-2010
2. Chu, D., Wang, X.: Fractal geometry in applied geophysics. Foreign Oil and Gas Technology 4(9), 44–52 (1995)
3. Chen, K.: Beef marbling box-counting dimension and information dimension determination. Agricultural Engineering 23(7), 145–149 (2007)
4. Brosnan, T., Sun, D.: Inspection and grading of agricultural and food products by computer vision systems to review. Computers and Electronics in Agriculture 36, 193–213 (2002)
5. Chen, K., Qin, C., Ji, C.: The research of cattle carcass eye flesh image segmentation method. Agricultural Machinery 37(6), 155–158 (2006)
6. Gerrard, D.E., Gao, X., Tan, J.: Beef marbling and color score determination by processing. Journal of Food Science 61(1), 145–148 (1996)
7. Chen, K., Qin, C., Mcdonald.: Segmentation of Beef Marbling based on Vision Threshold. Computers and Electronics in Agriculture 62(2), 223–230 (2008)
8. Jackman, P., Cheng, G., Jindu.: Predition of beef eating quality from colour, marbling and wavelet texture features. Meat Science 80(4), 1273–1281 (2008)

9. Yang, H.: Image segmentation research through threshold value method. Natural Science Newspaper 33(2), 135–137 (2006)
10. McDonaId, T.P., Chen.: Separating connected muscle issue on images of beef carcass eyes meat. Trans. of ASAE 55(6), 2059–2065 (1990)
11. Ren, F., Tu, K., et al.: Application of image processing technology evaluation of beef marbling. Meat Research (4), 14–15 (2002)
12. Shiraniata, K., Hayashi, K., et al.: Grading meat quality by image processing. Pattern Recognition 33, 97–104 (2000)
13. Zhao, J., Liu, M., Zhang, H.: Research of beef longissimus dorsi image split and marbling extracting technique based on mathematical morphology. Agricultural Engineering 20(1), 144–146 (2004)
14. Hu, X., Dong, C.: MATLAB from entry to master. People's Posts and Telecommunications Press (2010)
15. Shen, Z., Gao, F., Li, C., et al.: Beef grading technology based on computer vision research progress. Food Science and Technology 29(6), 304–306 (2008)
16. Quevedo, R., Calos, L.G., Aguilera, J.M., Cadoche, L.: Description of food surfaces and microstructural changes using fractal image texture analysis. Journal of Food Engineering 53(4), 361–371 (2002)
17. Tan, J.L.: Meat quality evaluation by computer vision. Journal of Food Engineering 61, 27–35 (2004)
18. Sun, Y., Xianyu, J., Shi, J.: The computer vision-based analysis method of cooling beef tenderness. Agricultural Machinery 34(5), 102–105 (2003)
19. Li, P., Xing, L., Pan, J., Gu, X.: Calculation of the fractal dimension and the image edge extraction. Journal of Jilin University (Information Science) 29(2), 152–156 (2011)
20. Pentland, P.: Fractal-based description of natural scenes. IEEE Transactions on Pattern analysis and Machine Intelligence 6(6), 661–674 (1984)
21. Zhang, Q., Wang, Z.: Proficient in MATLAB image processing. Electronics Industry Publishing House (2009)
22. Zhang, J.: Fractal, 2nd edn. singhua University Press (2011)
23. Liu, M., Zhang, X.: Differential box dimension method on the weld edge detection. Hebei Industrial Technology 26(5), 300–302 (2009)
24. Yin, Y., Li, P., Kang, Y., et al.: A Survey and Research of fractal theory. Science and Technology (15), 152–170 (2007)
25. Zhao, H., Yang, G., Xu, Z.: The comparing of Image fractal dimension calculation methods. Computer System Apply 20(3), 238–241, 246 (2010)
26. Li, Y., Hu, C.: Experimental design and data processing, 2nd edn., pp. 82–93. Beijing Chemical Industry Press (2011)

Research and Application of Human-Computer Interaction System Based on Gesture Recognition Technology

Zhenxiang Huang, Bo Peng, and Juan Wu

College of Information and Electrical Engineering,
China Agricultural University, Beijing 100083, China
{zhenxianghuang,pengbo_cau,juanwu_cau}@126.com

Abstract. With the gradual improvement of the computer performance and the use of computer more deeply in many fields , mouse and keyboard, the traditional human-computer interactive way, show more and more limitations. In recent years,gesture recognition interaction based on machine vision is used more and more widely due to its simple,nature,intuitive and non-contact advantages, which is becoming the research hotspot in the world. This paper mainly studies the main idea about DTW,HMM,ANN and SVM methods used in gesture recognition.This paper also expounds the basic model and future applications of gesture interaction system based on machine vision and research meaning of its application in agriculture.

Keywords: gesture recognition, DTW, HMM, ANN, SVM, agriculture.

1 Introduction

With the rapid development of information technology and gradual increase of computer's performance, human-computer interaction has become an important part in people's daily life when the use of computer is deepening in various fields.

The traditional human-computer interactive way,such as mouse and keyboard,shows more and more limitations,especially in emerging application fields of virtual reality, augmented reality and pervasive computing.The fact promotes the research on human-computer interaction technology develop towards the direction of people-oriented, free and direct manipulation.Gestures, playing an important role in the expression of a specific intent in people's daily life, are adopted as a new human-computer interaction way in recent years. The gesture interaction based on machine vision has become the research focus of human-computer interaction due to its simple, natural, intuitive and non-invasive advantages.Gesture recognition is a technology which identifies a variety of gestures according to certain rules by computer , instructs the computer to translate into corresponding control commands or semanteme with the goal of computer operating or information exchange[1].

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 210–215, 2013.

2 Gesture Interaction System Model Based on Machine Vision

The gesture interaction system based on machine vision input gestures through the acquisition module, in multocular case,cameras are distributed in accordance with a certain relationship in front of users[2],in monocular case,it require the plane where the camera is and the plane of the user's hand movements should be in the same level[3].

After reading the gesture images, interaction system detects gestures from the data stream on the basis of complexion and motion information[4],then separates the gesture signal from the video signal and launch trajectory tracking.The CamShift algorithm[5] and Kalman filter algorithm are commonly used.In the gesture analysis stage, the analysis process includes feature extraction and model parameter estimation.In the identification stage, gestures are classified by the model parameters and generated description as required.At last,system drive the specific applications according to generated gesture description.

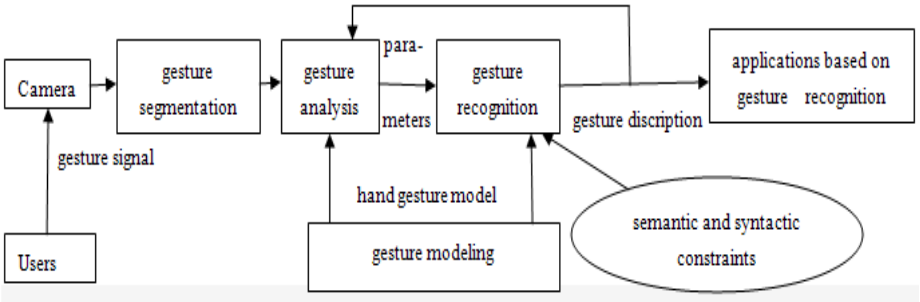


Fig. 1. Basic gesture interaction system model based on machine vision

3 Main Algorithms on Dynamic Gesture Recognition

Trajectories matching method and state space modeling method are the main dynamic gesture recognition algorithms.The most representative trajectories matching algorithm is Dynamic Time Warping(DTW) algorithm, and the typical state space modeling algorithms include Hidden Markov Model(HMM), Artificial Neural Network(ANN) and Support Vector Machine(SVM).

3.1 DTW Algorithm

Human activities can be modeled by structures of high-dimensional temporal trajectories,gesture recognition can then be performed by measuring the similarity or the distance between the input gesture trajectories and existing gesture trajectories template from the training sample set[6]. The Dynamic Time Warping(DTW) algorithm is designed to exploit some observations about the likely solution to make

the comparison between sequences more efficient. DTW algorithm assumes that the endpoints of two modes are aligned accurately,then it translates the matching problem into the problem how to use dynamic programming techniques to find the optimal path through the limited grid efficiently.The sequences are warped non-linearly in the time dimension to determine measure of their similarity independent of certain non-linear variations in the time dimension. The advantages of DTW method are simple in concept, efficient and allow sufficient flexibility between the conceptual model and reference model[7].

3.2 HMM Algorithm

The state space modeling method is a algorithm that models for spatiotemporal characteristics of the gesture trajectories,it treats the trajectories as a series of transfer between the states. The specific approach is to learn the state transition parameters from training samples at first,then appraise the trajectories with different gesture models to achieve results. The typical algorithms include HMM,ANN and SVM.

A Hidden Markov Model(HMM) is a double stochastic process model which is represented with parameters and used to describe probability statistical properties of stochastic processes.It includes the random process of the state transition and the random process of observation symbols output. It evolves from the Markov Chain.Markov Chain is a mathematical system that undergoes transitions from one state to another, between a finite or countable number of possible states. It is a random process characterized as memoryless: the next state depends only on the current state and not on the sequence of events that preceded it. HMM can be regarded as a general form without the constraints of the Markov Chain[8]. As a widely used statistical method,HMM that general topology structure has good ability to describe the spatiotemporal variation of the hand signals, but the result is uncertain due to HMM has more than one migration curve for a given output result,so it can't determine the state matrix of input set by output results. In addition, HMM can't defined migration rules accurately, for example, the number of states and the number and types of migration.It depends on priori knowledge and try to determine these rules. In fact, the migration mapping result of Markov Process isn't satisfactory and the topology structure of HMM is general. These all cause the model too complicated in analysis of hand signals and large calculation of training and recognition.

3.3 ANN Algorithm

An Artificial Neural Network(ANN) is a technology that imitates the structure and/or functional aspects of biological neural networks by advanced engineering technologies. Its purpose is to enable the robot to perceive,learn and reasoning like human brain. ANN consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. BP Neural Network Model is used most widely in recent years. It is a multilayer feedforward neural networks of one-way transmission, have unilaminar or multilayer hidden layer

nodes except input nodes and output nodes, no coupling among nodes of the same layer, input signal pass through all layer nodes in turn from input layer nodes to output layer nodes. Output results of each layer nodes only affect output results of the next layer nodes.BP Neural Network Algorithm has characteristics of self-organizing and self-learning for information processing, has strong ability of anti-noise, model spread and processing incomplete models.Its main defect is low convergence rate and unavoidable overfitting phenomenon.

3.4 SVM Algorithm

A Support Vector Machine(SVM) is a machine learning method based on Statistical Learning Theory(SLT), primarily research how to get the best results from small sample.In other word, it seeks a optimal balanced scheme between complexity of model and learning ability on the basis of limited sample information[9]. It solves practical problems well, such as small sample, nonlinearity, over learning, high-dimension pattern recognition and local minima, by the method of structural risk minimization principle and the kernel function. Due to its success in dealing with the relationship between the dimensions and computed strength and enhancing the computational efficiency greatly,SVM plays an important part in real-time dynamic gesture recognition.

4 Gesture Recognition Interaction Application in Agriculture

As a new human-computer interactive way that different from traditional interface, the characteristics of gesture recognition interaction like intuitiveness and naturalness, detemines its good application prospects in many fields,for example,in agriculture.

In the virtual agricultural field,operating objects in virtual environment by gestures can simulate the plant shape that crops may reach maximum yields accurately, imitate crop row spacing of intercropping and companion planting intuitively. These can provide useful reference for planning cultivated land reasonably. By the simulation of virtual agriculture,learning of agricultural technology will be visualization, interaction will be realistic,then the agricultural practitioners will understand and master agricultural knowledge better.

In the field of agricultural production, gesture recognition interaction will lower the use barrier of computer,which will help farmers to use intelligent systems for agricultural production conveniently.It also can control agricultural robot in long distance for collection or fishing in extreme environments,such as poor weather conditions[10].

Gesture recognition also can be used for interaction of information display system in public places including agricultural production site where mouse and keyboard aren't suitable to be used[11]. In addition, gesture recognition interaction can be used for distance education system and video conference system to facilitate agricultural production and spread last agricultural information and technologys[12].

5 Gesture Recognition Interaction Application Prospects

With the help of gesture recognition technology, remote control of household appliances such as TV, DVD, stereo can be achieved and provide convenience for people's life[13][14],game operation will also be more natural and game playing experience will be enhanced.

Gesture recognition technology can be used for sign language translation system[15],then deaf-mute people can communicate with computer and normal human by it,which will improve their education level and provide convenience for two-side communicate.

With the development of gesture recognition technology, mouse and keyboard being replaced in the near future will not be a dream only, which will reduce the cost of human-computer interaction and e-waste, be consistent with the historical trend of green low-carbon.

6 Conclusion

As an important development direction of new generation human-computer interaction system, hand gesture recognition has raised widespread concern of researchers and have achieved certain results. Gesture recognition technology combined with agriculture informationization is good for the promotion of intelligent agricultural production equipment, rapid popularization of agricultural technology and fine control of agricultural production. These will significantly reduce production costs, increase productivity and make an important contribution to world food security.

References

1. Shneiderman, B.: Direct manipulation:A Step Beyond Programming Languages. IEEE Computer 16, 57–69 (1983)
2. Olsen Jr., D.R.: Where Will We Be Ten Years from Now. In: 10th annual ACM Symposium on User Interface Software and Technology, pp. 115–118. ACM, New York (2007)
3. Pavlovic, V.L., Sharma, R., Huang, T.S.: Visual Interpretation of Hand Gestures for Human-Computer Interaction:A Review. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 677–695 (1997)
4. Howe, L.W., Farrah, W., Ali, C.: Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition. In: International Symposium on Information Technology, Kuala Lumpur, Malaysia, vol. 2, pp. 1–7 (2008)
5. Mahmoudi, F., Parviz, M.: Visual Hand Tracking Algorithms. In: International Conference on Geometric Modeling and Imaging, London, England, pp. 228–232 (2006)
6. Psarrou, A., Gong, S., Walter, M.: Recognition of human gestures and behaviour based on motion trajectories. Image Vision Comput. 20, 349–358 (2002)
7. Ren, H.B., Zhu, Y.X., Xu, G.: Vision Based Recognition of Hand Gestures:A Survey. Acta Electronica Sinica 2, 118–121 (2000)

8. Charniak, E.: Statistical Language Learning. MIT Press, Cambridge (1993)
9. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, England (2000)
10. Wang, W.T., Li, S.Q.: Research and Implementation of the Object-Oriented Virtual-Hand Technology. Computer Engineering & Science 2, 45–47 (2005)
11. Chen, Q., Malric, F., Zhang, Y., Abid, M., Cordeiro, A., Petriu, E.M., Georganas, N.D.: Interacting with Digital Signage Using Hand Gestures. In: Kamel, M., Campilho, A. (eds.) ICIAR 2009. LNCS, vol. 5627, pp. 347–358. Springer, Heidelberg (2009)
12. Mitra, S., Acharya, T.: Gesture Recognition:A Survey. IEEE Transactions on Systems, Man and Cybernetics-part C: Applications and Reviews 37, 311–324 (2007)
13. William, T.F., Craig, D.W.: Television Control by Hand Gestures. In: IEEE International Workshop on Automation Face and Gesture Recognition, Zurich (1995)
14. Chang, J.S., Kim, S.H., Kim, H.J.: Vision-Based Interface for Integrated Home Entertainment System. In: Marques, J.S., Pérez de la Blanca, N., Pina, P. (eds.) IbPRIA 2005. LNCS, vol. 3522, pp. 176–183. Springer, Heidelberg (2005)
15. Yang, M.H., Ahuja, N., Tabb, M.: Extraction of 2D motion trajectories and its application to hand gesture recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 1061–1074 (2002)

Crop Model-Based Greenhouse Optimal Control System: Survey and Perspectives

Qiaoxue Dong, Weizhong Yang, Lili Yang*, Yifei Chen, Shangfeng Du, Li Feng, Qinglan Shi, and Yun Xu

Department of Electronic Information, College of Information and Electrical Engineering, China Agricultural University, Box 63, 100083, Beijing
llyang@cau.edu.cn, dongqiaoxue@163.com

Abstract. A survey is presented of the developments in scientific literature on the greenhouse optimal control. The related problems to optimal control were discussed, and the schematic diagram of optimal control system, combined crop models and optimization algorithm were covered focusing on issues such as: model simplification and validation, definition of objective function, input or output constraints and dynamic optimization, and especially structure-function Greenlab model was made s a first exploration for the optimal control application. Not only the above research issues were listed in the paper, the perspectives and the solutions are presented as well. It is pointed out that only the crop growth model is integrated into the greenhouse climate optimal control, the best economical result and energy-saving can be warranted.

Keywords: Greenhouse optimal control, Crop growth model, Greenlab model, Dynamic optimization, Energy saving.

1 Introduction

The operation cost of current modern greenhouse is usually high due to the complexity of greenhouse-crop production system, so many efforts have been made to develop optimal management and control strategy which aims to reduce the energy consumption, thus to improve the efficiency of greenhouse production[1,2,3,4,7,10]. Simulation and experimental work were also carried out to support the effectives of optimal control system. For instance, Van Ooteghem found that through the simulation of a optimal closed loop controls, the boiler use was reduced, thereby reducing fossil energy use, and gas use was decreased by 77% compared to a conventional greenhouse[5]; Van Henten made a experimental comparison of the performance between optimal control approach with the traditional control system, and the results showed that with optimal control , energy and carbon dioxide are used more efficiently[3].

However, it is still not an easy task to apply the optimal control concepts into greenhouse production practice , because the current models for greenhouse and crop

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 216–224, 2013.

are usually too complex to be incorporated into the optimal control. But crop models are essential, and without the support of crop models, it will be difficult for optimal control system to achieve the appropriate results, because the dynamic greenhouse-crop models output the state variables required by objective function[5,13]. The quality of crop models will affect the preformance of the optimal control , so for application of optimal control, accurate and simplified models are required , i.e. Only if to build an optimal control system combined appropriate crop growth model, energy efficient operation can be implemented[1]. In this paper, we will give a brief description about such optimal control system, and its recent advances about related optimal control problem will be presented in detail in order to exploit the possibilities for practical application.

2 The System Structure of Crop Model-Based Optimal Control

In order to understand the system structure of optimal control, in this section we will illustrate it by first comparing with the conventional climate control system in current practice (Fig.1). In conventional control of the greenhouse climate, the grower defines the set-points of the greenhouse climate variables such as temperature, humidity and carbon dioxide concentration , and the climate control computer works to achieve the desired climate using measurements and feed-back control techniques. During the growth season, the grower may modify the setting trajectory based on the observation of crop growth state and their experience.

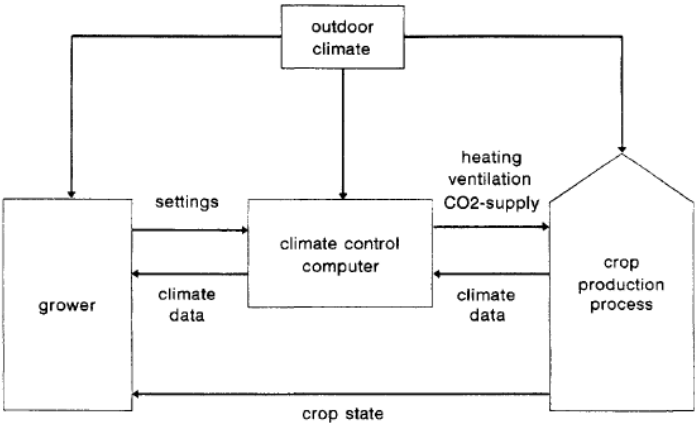


Fig. 1. A schematic diagram of the conventional climate control procedure [3]

In the above traditional system, trajectory settings were determined by the grower according to heuristic rules, which will have a definite effect upon the consumption of energy and other resources, as well as on growth and development of the crop, but the exact effect is unknown to the grower, and can at present only be inferred from experience [1].

We can see that the traditional control system of greenhouse climate don't take into account the costs in an explicit manner, and provide the control actions by focusing almost always on the performance. So the costs for keeping that performance might be unacceptable[10].

The optimal control strategy provide an alternative method to consider energy consumption [7,10]. It consists of models for the greenhouse as well as for the crop, a suitable objective function and an optimization algorithm(Fig.2), which will be introduced in the following.

In Fig.2 system, the control inputs are not the settings defined by the grower like above conventional system, because the incorporation of the heuristic control rules in the optimal control decreases the optimal control freedom. Therefore the greenhouse –crop model used in this research is based on actuator values as control inputs. But the input and output constraints are still need to be given by the grower. Other inputs include external weather data, and because the results of the optimal control strongly depend on the weather conditions, reliable forecasts are also needed in the optimal control system.

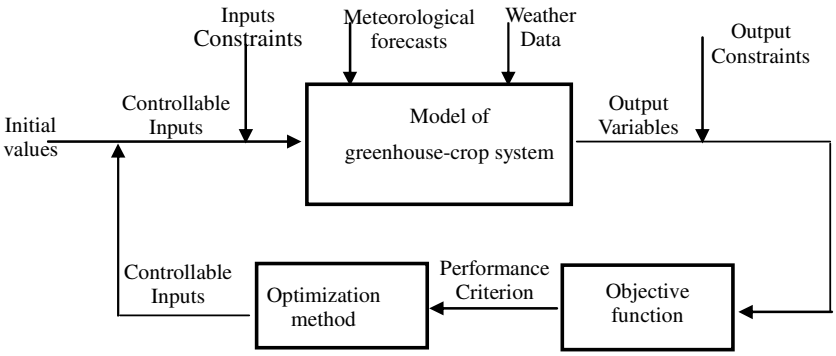


Fig. 2. The system structure of model-based optimal climate control

Optimal control approach is based on a mathematical model for calculating greenhouse energy consumption and on a mathematical method for minimizing total energy consumption, so the models for greenhouse climate as well as the crop are required. In a formal way, the model is usually represented by state equation[3],i.e.

$$\frac{dx}{dt} = f(x,u,v,t), \qquad x(t_b) = x_b \tag{1}$$

in which x are the state variables, u are the control inputs, v are the external inputs, t denotes time and dx/dt represents the rate of change of the state in time. The initial state of crop production process is denoted by $x(t_b)$ in which t_b represents the planting date. The state of production process is represented by variables relating to crop such as fresh weight, as well as to variables describing the indoor climate such as air temperature, humidity and carbon dioxide concentration.

The function $f(.)$ describes the dynamic behavior of greenhouse climate and crop, i.e. modeling methodology. In most literatures about optimal control strategy, the greenhouse climate model used is mainly based on energy and mass balance equations[4,5] , the details of which will not be explained in this paper, and the description about the energy balance can be found elsewhere[11,16,17]. While the crop growth model used in present optimal greenhouse control will be described in detail in section 3.

The objective function in the optimal system Fig.2 summarizes the output variables to one single performance criterion, which depends on the requirement of user, and it will affect the performance the optimal control. In general, the performance criterion can be summarized into two category: (1) economic performance criterion; (2) cost function. According, the definition of objective function is

$$J(u) = \Phi(x(t_f), t_f) - \int_{t_b}^{t_f} L(x,u,v,t)dt \tag{2}$$

Or

$$J(u) = -\Phi(x(t_f), t_f) + \int_{t_b}^{t_f} L(x,u,v,t)dt \tag{3}$$

Where $\Phi(.)$ is the term related with the expected crop economic value and $L(.)$ term related with the operation costs of the climate conditioning equipment. The optimal control problem defined by Eq. 2 is to determine the trajectory of the control inputs $u^*(t)$ over $[t_b, t_f]$ which maximize the economic revenue or production, while the performance criteria defined by Eq.3 aims to minimize the total energy consumption subjected to input and output constraints.

The optimization method is to find the solution of above optimal control problem defined by the objective function. In section 4, several commonly used optimization algorithm will be discussed and reviewed briefly.

3 Greenhouse Crop Model Oriented for Optimal Control

The Model quality is a fundamental aspect to achieve adequate control performances. Here some successful crop models in horticultural practice will be discussed in order to exploit their possibilities applied in the optimal control. At present, the main obstacles are that the current greenhouse crop models are too complex, or require too many data or too much calculation, which is not suitable for optimal control task, so simple or very compact models are required.

Current horticultural crop models are mostly based on a photosynthesis model and an evaporation model, and the typical instance is TOMGRO[12,18]. Although TOMGRO model has been validated and proved enough accurate , it has too many parameters to be directly involved into the optimal control method . So its simplified

version is proposed in order to be suitable for the optimal application . For instance , Pucheta restricted TOMGRO model to two state variables : dry weight and number of leaves which are modeled by Eq.4. to implement the optimal greenhouse control of tomato-seedling crops[15].

$$\dot{W} = E[P_g(T, S_{PAR}, CO_2) - R_m(T)W], \qquad \dot{N} = r_m\, r(T) \tag{4}$$

Where W and N are the time-dependent state variables representing the total dry weight and the number of leaves, respectively. $P_g(.)$ is the photosynthesis rate which depends on the temperature , solar radiation and the concentration of CO_2; $R_m(.)$ is respiration rate of the leaves, which only depends on temperature.

van Straten also described a simplified model which has similar modeling mechanism with the description in Eq.5 , but it gives a fairly general representation of crop production[1].

$$\begin{aligned} \dot{x}_{\mathrm{n}} &= q_{\mathrm{P}}\{C, I, \cdot\} h\{x_{\mathrm{n}}\} - (1+\theta) s\{x_{\mathrm{n}}\} \sum_i q_{\mathrm{G}i}\{T, x_{\mathrm{s}i}, \cdot\} - \sum_i q_{\mathrm{M}i}\{T, x_{\mathrm{s}i}, \cdot\} \\ \dot{x}_{\mathrm{s}i} &= s\{x_{\mathrm{n}}\} q_{\mathrm{G}i}\{T, x_{\mathrm{s}i}, \cdot\} \quad i = 1, \cdots, N \end{aligned} \tag{5}$$

Where $\dot{x}_n$ and $\dot{x}_{st}$ are two state variable representing not-structural biomass and structural biomass. C , I and T are the temperature , solar radiation and carbon dioxide concentration, respectively.

Above models is only a general description for optimal control-oriented application. Ioslovich fruther proposed a three-stage growth model for tomato named MBM-A[2], which is modelled in a more adequate and accurate way . In MBM-A model, two state variables namely accumulated vegetative dry mass x and the harvestable red fruits y are defined and they have different differential equations as described by Eq.6.

$$\begin{cases} \dfrac{dx}{dt} = M(t,U) f(x), \quad \dfrac{dy}{dt} = 0 & vegetative\ \ stage \\ \dfrac{dx}{dt} = M(t,U)(1-\alpha) f(x), \quad \dfrac{dy}{dt} = M(t,U) g(y)/K_c & vegetative\text{–}reproductive\ \ sta \\ \dfrac{dx}{dt} = 0, \quad \dfrac{dy}{dt} = M(t,U) f(x)\eta & reproductive\ \ stage \end{cases} \tag{6}$$

From Equation 6, the time-dependent state variables x and y relies on control variables U , including greenhouse heating, ventilation and CO2 enrichment, which will influence the greenhouse climate such as temperature and CO2 concentration. Further assumption is that the mean daily temperature is strongly correlated with the mean daily light, and light is strongly correlated with photosynthesis. In Eq.6, unknown coefficients are need to be calibrated, and the model was validated against TOMGRO.

The recently emerged plant structure-function model Greenlab also provides the possibilities to be introduced into the optimal control strategy [6]. Because it is an efficient dynamic process of balancing the simplicity and complexity. It consists of biomass production and allocation equations, and the plant produces biomass by leaf photosynthesis,i.e.

$$Q(i) = F(S(i), E(i), r1, g(r2, S(i), N(i)) = \frac{E(i)S(i)}{r1}[1 - \exp(-r2\frac{\sum_{j=1}^{N(i)} S_j(i)}{S(i)})] \tag{7}$$

where $Q(i)$ represents the dry biomass created at the growth cycle i and we define the growth cycle (GC) as the thermal time necessary for each plant axis to develop a new growth unit (GU); S_j is the blade area of the j th leaf, and $S(n)$ is a kind of ground projection area of the leaf surface; Parameter $r1$ sets the leaf-size effect on transpiration per unit leaf area, while $r2$ accounts for the effect of mutual shading of leaves according to the Lambert–Beer's law.; $N(i)$ is the number of leaves; $E(i)$ is the average biomass production potential depending on environmental factors, such as light, temperature and soil water content.

The biomass produced by photosynthesis is redistributed among all the organs according to their demands and sink strength:

$$\Delta q_o(i, j) = \frac{P_o \times f_o(j)}{D(i)} Q(i-1) \tag{8}$$

Where Δq_o is the biomass increment of o-type organ; P_o are the organ sink strengths and are model hidden parameters. f_o are normalized distribution functions characterizing the evolution of the sink strengths; $D(i)$ is the total biomass demand.

The Greenlab model provides an interaction interface with climate and an explicit definition about growth cycle(the duration of which can vary from several days(tomato) to one year), so it describes the evolution of the plant structure periodically. Furthermore, the validation study of this Greenlab model has been well made[8,22] , and the research work shows possible applications of GreenLab in optimization and control for agronomy.

4 Optimization Algorithm

The Effective optimization methods play an key role on finding the solution of optimal control problem. The above-stated control problems in section 2 belong to dynamic optimization problems which can be solved by choosing a dynamic programming technique, such as Hamilton-Jacobi-Bellman equation[2], Lagrange Multiplier technique [19], Genetic algorithm[21], Model-based Predictive control[10]and Receding Horizon Optimal Controller[5], and so on.

However, optimal control concepts and nonlinear dynamic programming (NDP) in particular have almost not been used in practice, due to the implementation complexity. Therefore the researchers try to improve the optimization techniques

mentioned above to be capable of using crop growth models. For example, Luus proposed iterative dynamic programming(IDP)procedure to solve the optimal control problem[14]. IDP—a modified dynamic programming is a computational procedure that allows one to obtain the optimal control trajectories for time varying nonlinear processes, with any type of performance indices and with restrictions on both state and control variables. The application of this algorithm to a continuous system requires discretization of the differential equations that model the process, and quantification of the variables of state, decision, control and time. Ioslovich used the Hamilton-Jacobi-Bellman formalism in the form of Krotov-Bellman sufficient conditions and the optimal value of the constant seasonal control intensity can be approximately obtained from the MBM-A model by Eq.3 [2]. Because during each sampling period, the simulation of the associated optimal closed loop control system is very time consuming in general. to drastically limit the simulation time, the optimal control problems can be solved by the receding horizon optimal controller[5,21] used genetic algorithm combined with Greenlab model to achieve the simulation for greenhouse water supply optimization problem, which open the possibilities for Greenlab model to be applied to optimal greenhouse climate control .The advantage of genetic algorithm is to avoid local minima.

Optimal control problem based on non-linear and non-quadratic performance criteria, such as those discussed in this paper , are very difficult to solve analytically. Iterative schemes need to be used to achieve a numerical solution of the mathematical problem. Dynamic optimization problems can be numerically solved by direct or indirect methods. Over the last years, a number of numerical search methods have been developed. For instance, Van Henten proposed a modified steepest ascent algorithm and used a forth order Runge-Kutta integration algorithm to obtain the numerical solution for optimal control trajectories[3]. An indirect gradient method was proposed which has proven its effectiveness in optimal greenhouse control and many other fields[5,20]. The search procedure to find optimal value is described as the following:

- start the simulation model with initial values for a set of input variables;
- calculates the output variables by the simulation model and an objective function value;
- derives an improved set of controls the optimization procedure and starts a new simulation procedure

The above feedback loop is repeated until the objective function value converges on the optimum .Since the simulation model runs independently from the optimization algorithm, there are in principle no restrictions with respect to its structure. The drawback is ,there is no certainty finding the absolute minimum or maximum[13]. Besides this, the selection of initial values also need to paid more attention because it will affect the performance of optimal algorithms.

5 Discussion and Perspectives

This paper presented a framework of crop model-based optimal greenhouse control system and key factors, which shows optimal control of the greenhouse is feasible.

However, even if some valuable research work has been made during the last decades, we can see that it is still a long way for the optimal control strategy to be applied into greenhouse practice, especially for China horticulture production. To solve above problem, current elaborate and complex horticultural crop models are needed to simplified to be suitable for optimal control, and only if are the simplified models calibrated and experimentally proven accuracy, they can be incorporated into the optimal algorithms; Efficient optimal algorithms are required to solve the time-consuming and stability problem, and time-scales problem induced by crop growth response and control actions are also the obstacles to hamper the application of optimal control. Despite the challenging list of issues that need further investigation, we are convinced that with the improvement of modeling technology and computer power, the optimal control methodology based on crop growth model must find its way for on-line production practice, and provide an effective tool to achieve the energy saving.

Acknowledgments. This work is supported in part by Natural Science Foundation of China (#61174088) and Fundamental Research Funds for CAU(#2012QJ078).

References

1. Van Straten, G., Challa, H., Buwalda, F.: Towards user accepted optimal control of greenhouse climate. Computers and Electronics in Agriculture 26, 221–238 (2006)
2. Ioslovich, I., Gutman, P.O., Linker, R.: Hamilton-Jacobi-Bellman formalism for optimal climate control of greenhouse crop. Automatica 45, 1227–1231 (2009)
3. Van Henten, E.J., Bomtsema, J., van Straten, G.: Improving the efficiency of greenhouse climate control: an optimal control approach. Netherlands Journal of Agricultural Science 45, 109–125 (1997)
4. Chalabi, Z.S., Bailey, B.J., Wilkinson, D.J.: A real-time optimal control algorithm for greenhouse heating. Computers and Electronics in Agriculture 15, 1–13 (1996)
5. Van Ooteghem, R.J.C., van Willigenburg, L.G., van Straten, G.: Receding horizon optimal control of a solar greenhouse. Acta Hort. (ISHS) 691, 797–806 (2005)
6. De Reffye, P., Hu, B.G.: Relevant qualitative and quantitative choices for building an efficient dynamic plant growth model: GreenLab Case. In: Hu, B.-G., Jaeger, M. (eds.) Plant Growth Modeling and Applications: Proceedings-PMA 2003, pp. 87–107. Tsinghua University Press and Springer, Beijing (2003)
7. Dieleman, J.A., Marcelis, L.F.M., Elings, A., Dueck, T.A., Meinen, E.: Energy saving in greenhousees: optimal use of climate conditions and crop management. Acta Hort. (ISHS) 718, 203–210 (2006)
8. Dong, Q.X., Wang, Y.M., Barczi, J.F., De Reffye, P.: Tomato growth modeling based on interaction of its structure-function. In: Hu, B.-G., Jaeger, M. (eds.) Plant Growth Modeling and Applications: Proceedings - PMA 2003, Tsinghua, pp. 250–262. University Press and Springer, Beijing, China (2003)
9. Farquhar, G.D., von Caemmerer, S., Berry, J.A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta 149, 78–90 (1980)
10. Blasco, X., Martınez, M., Herrero, J.M., Ramos, C., Sanchis, J.: Model-based predictive control of greenhouse climate for reducing energy and water consumption. Computers and Electronics in Agriculture 55, 49–70 (2007)

11. Jollie, O., Danloy, L., Gay, J.B., Munday, G.L., Reist, A.: HORTICERN: An improved static model for predicting the energy consumption of a greenhouse. Agric. Forest Meteorol. 55, 265–294 (1991)
12. Jones, J.W., Dayan, E., Allen, L.H., van Keulen, H., Challa, H.: A dynamic tomato growth and yield model (TOMGRO). Trans. ASAE 34, 663–672 (1991)
13. Lentz, W.: Model applications in horticulture: a review. Scientia Horticulturae 74, 151–174 (1998)
14. Luus, R.: Luus-Jaakola Optimization Procedure. Recent Res. Devel. Chem. Eng. 4, 45–64 (2000)
15. Pucheta, J.A., Schugurensky, C., Fullana, R., Patino, H., Kuchen, B.: Optimal greenhouse control of tomato-seedling crops. Computers and Electronics in Agriculture 50, 70–82 (2006)
16. Van Bave, C.H.M., Takakura, T., Bot, G.P.A.: Global comparison of three greenhouse climate models. Acta Hortic 174, 21–33 (1985)
17. Bot, G.P.A.: Greenhouse climate: from physical processes to a dynamic model. Ph.D. Thesis, Agricultural University, Wageningen (1983)
18. Dayan, E., van Keulen, H., Jones, J.W., Zipori, I., Shmuel, D., Challa, H.: Development, calibration and validation of a greenhouse tomato growth model:I Description of the model. Agricultural Systems 43, 145–163 (1993)
19. Bertsekas, D.P.: Constrained optimization and lagrange multiplier methods. Academic Press, London (1982)
20. Van Willigenburg, L.G., van Henten, E.J., van Meurs, W.T.M.: Three time-scale receding horizon optimal control in a greenhouse with a heat storage tank. In: Proceedings of the Agricontrol 2000 Conference, Wageningen, The Netherlands, pp. 12–14 (2000)
21. Wu, L., Le Dimet, F.X., Hu, B.G., Cournede, P.H., De Reffye, P.: A water supply optimization problem for plant growth based on GreenLab model. Journal of ARIMA 3, 194–207 (2005)
22. Zhan, Z.G., Wang, Y.M., De Reffye, P., Wang, B., Xiong, F.: Architectural modeling of wheat growth and validation study. In: Proceedings of ASAE Annual International Meeting, Milwaukee, USA, pp. 1–14 (2000)

Cold Chain Logistics Monitoring System with Temperature Modeling

Shaoxin Guo[1,2], Fan Zhang[1,2], and Jianqin Wang[1,2,*]

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing, 100083, China
[2] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
wjqcau@126.com

Abstract. Cold chain logistics guarantees the quality and freshness of some especial agriculture products. At present cold chain logistics monitoring system can just monitor single point of the temperature in the cold chain vehicle body, it cannot estimate the space temperature distribution of cold chain vehicle body. This paper introduces a cold chain logistics monitoring system based on ARM Linux. The data of temperature value in the cold chain vehicle body and the geographic location of cold chain vehicle will be monitored in real time. This paper presents a space temperature modeling method based on mutli-point temperature value to estimate the space temperature distribution of cold chain vehicle body.

Keywords: cold chain logistics, ARM, temperature modeling.

1 Introduction

The specific characteristics of some special agriculture products like fish, egg, meet and vegetables determined the logistics of those agriculture products must be in low temperature. So the definition of cold chain logistics is based on freezing technology, by the method of refrigeration technique to guarantee the freshness and quality of agriculture products [1].

In the developed country such as USA and Japan, cold chain logistics system was built with the core of information technology, by the use of storage technology, transportation technology, distribution technology, loading and unloading technology, and stock control technology [2].

In recent years by the establishment of the cold chain logistics grid, cold chain logistics have been covered in everywhere of China. But currently there are still some problems of cold chain logistics in China. The problem is the facilities of cold chain logistics is not perfect, the low level of Market-oriented and the competitive environment of cold chain logistics companies is very bad due to the low price competition [3].

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 225–233, 2013.

The most important key to solve the problem is to build the facilities of cold chain logistics, to enhance the management of cold chain logistics companies, to build the standard of cold chain logistics. Additionally the management of the cold chain logistics should be within information technology [4].

So the technologies and facilities for cold chain logistics are necessary for the development of the modern cold chain logistics [5]. Therefore, to build a cold chain logistics monitoring system is the way to enhance the efficiency of management in cold chain logistics.

Li and You proposed a temperature tracking system based on iButton-DS1923 [6], but this system cannot estimate the temperature distribution of the cold chain vehicle body and cannot locate the cold chain vehicle. Wang proposed a temperature monitoring system based on RFID [7]. But this system cannot monitor the temperature in real-time.

This paper introduces a cold chain logistics monitoring system with temperature modeling. There are two sub-systems of this monitoring system, one is embedded system and the other is control center. Embedded system collects the data about temperature value of cold chain vehicle body and the geographic location of the vehicle, and then sends the data to control center in real-time. The program run in control center receives the data from embedded system, draws the graph of temperature modeling and also shows the geographic location of the cold chain vehicle.

The definition of temperature modeling in this paper is based on the method of interpolation to estimate the temperature distribution in the cold chain vehicle body. There are some interpolation method is widely used. Those methods are distance weighing, interpolating polynomials, Kriging and spline method [8]. In the group of the interpolation method, the distance weighing method is most convenient, but huge error will occur. Spline interpolation method is based on some extreme point, by the way of control estimate variance to get the smooth spline by polynomial. But it is not fit to temperature modeling, because the calculation is huge. The Kriging interpolation can get the optimum linear unbiased estimate, but Kriging interpolation needs many interpolation points. In the temperature modeling system the number of interpolation point is limited, so Kriging interpolation is not fit to temperature modeling. The space interpolation method for temperature modeling should not only more convenient and precise, but also less calculation.

Because of the continuity characteristic of temperature distribution, this paper introduces a temperature modeling method based on three dimensions Lagrange interpolation [9]. Within the temperature modeling method, the temperature in the cold chain vehicle body will be estimated. So that it is very convenient to monitor the temperature distribution in the cold chain vehicle body.

2 Design of Embedded System

This embedded system was installed in the cold chain vehicle, the core of the embedded system is a S3C6410 MCU, and it is a 32-bit ARM11 RISC microprocessor. The memory in this embedded system is 128MB DDR RAM and 1GB NAND FLASH. The MCU controlled 8 temperature sensors, which installed in

the 8 vertex of the cold chain vehicle body. GPS module is used to get the cold chain vehicle geographic location. GPRS modem in embedded system is used to communicate with control center. Figure 1 shows the structure of embedded system. S1 to S8 means temperature sensor No1 to No8.

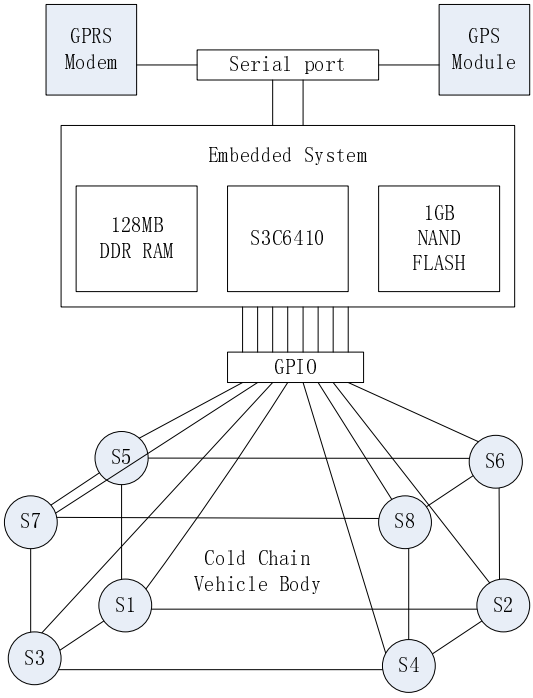


Fig. 1. The 8 temperature sensors, GPS module and GPRS modem are connected to the embedded system

The operating system of embedded system is Linux. Linux is famous for its efficiency flexibility portability and stability. The file system in the embedded system is yaffs2. The application run in the embedded system is to get the temperature value from temperature sensor and cold chain vehicle geographic location from GPS module, and then transfer the data to control center by GPRS modem.

2.1 The Method of Temperature Measurement

The temperature sensor DS18B20 is used to measuring the temperature, it can just use one wire bus to communicate with the MCU [10]. The 8 DS18B20 connected to the S3C6410 GPIO port. The Linux driver is used to receive the data from DS18B20. The Linux drivers are like an interface of hardware and Linux kernel. There are two sort of device controlled by Linux driver, char device and block device. The DS18B20 is controlled as char device by the Linux driver. The process of read temperature value is shown in table 1.

Table 1. The process of receive temperature value from DS18B20

Step	Function
1	Open device
2	Select which ds18b20
3	Reset the ds18b20
4	Send skip check rom command [0xCC]
5	Send convert temperature command[0x44]
6	Reset the ds18b20
7	Send skip check rom command [0xCC]
8	Send read data register command[0xBE]
9	Read temperature data

2.2 The Method of Vehicle Location

GPS module is used to locate the cold chain vehicle. GPS (Global Positioning System) is a space satellite navigation system which can provides location almost anywhere in the earth and time information. The GPS module connected to the embedded system through serial port. GPS module send the data include UTC time, longitude, latitude and other information to the MCU through serial port. The program just analyze the frame first with $GPGGA. This frame is most important in GPS, it include most geographic position information, and this frame used very widely. Based on the NMEA-0183 protocol the format of $GPGGA frame is shown in table 2[11].

Table 2. The format about GPGGA frame

$GPGGA,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,<9>,M,<10>,M,<11>,<12>*xx<CR><LF>	
<1>	UTC time (hhmmss.sss)
<2>	Latitude (ddmm.mmmm)
<3>	Hemisphere of latitude (N or S)
<4>	Longitude (dddmm.mmmm)
<5>	Hemisphere of Longitude (E or W)
<6>	State of GPS
<7>	Number of Satellite (00 to 12)
<8>	horizontal dilution of precision (0.5 to 99.99)
<9>	Altitude (-9999.9to9999.9meters)
<10>	Height of geoid above WGS84 ellipsoid
<11>	Time since last DGPS update
<12>	DGPS reference station id

The process of get data about latitude and longitude is first open the device of serial port (/tty/SAC1), and then set baud rate 9600, using read() function to read the string of GPS module send, and then using the sscanf() function to match the GPGGA frame and get the data about latitude and longitude.

2.3 The Method of Real-Time Data Transmission

GPRS modem is used to transfer the data between embedded system and control center. GPRS (General packet radio service) is a packet oriented mobile data service on GSM mobile phone. Because of the packet communication technology is used in GPRS system, the usages of network resources will be greatly optimized. GPRS network supports the TCP/IP protocol, and it can connect to the X.25 network [12]. The data transport rate of GPRS is ranged from 40Kbps to 100Kbps [13]. In China GPRS network has almost covered everywhere.

GPRS modem connects to Internet base on PPP protocol by AT command. PPP (point-to-point protocol) is the protocol about data transmission between two nodes of network. The link that the protocol creates is Full-duplex. The PPP protocol transport data through the point to point link. PPP protocol packaging the IP datagram and combine the datagram in GPRS packet, finally GPRS modem send the packet off. The merit of PPP is used easily and widely also it supports user verification and IP distribution.

PPP protocol is consists of three parts [14]:

1. High-Level Data Link Control (HDLC) protocol used to package IP packet to link for transmission.
2. Link Control Protocol (LCP) used to set up PPP communication.
3. Network Control Protocol (NCP) that is run atop of PPP and used to negotiate options for a network layer protocol running atop PPP.

The process of the establishment about PPP link is shown in figure 2.

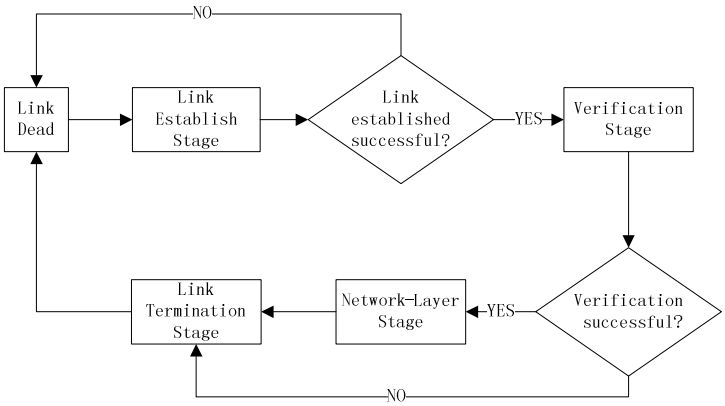


Fig. 2. There are 4 steps to establish the PPP link. Step 1 is link establish stage. Step 2 is user verification stage. Step 3 is network-layer protocol stage. And step 4 is link termination stage.

When the PPP link is established, socket API is used to send and receive data in software layer. Socket API is provided both in Linux and Windows. Control center bind the process of monitoring program to the port, and listening socket request and accept the request. The program in embedded system call the send function to send the data to the control center and the program in control center call the recvfrom function to receive the data.

3 Space Temperature Modeling

3.1 Three Dimensions Lagrange Interpolation

In order to estimate the temperature in the space of the cold chain vehicle body by the temperature sensor installed in the 8 vertex of the cold chain vehicle body. The method of three dimensions Lagrange interpolation is used. The three dimensions Lagrange interpolation is based on one dimension Lagrange interpolation.

suppose that in three-dimensional space {(x,y,z)|x ϵ [a,b],y ϵ [c,d],z ϵ [e,f]},and it was divided in :

$$a = x_0 < x_1 < .... < x_m = b \tag{1}$$

$$c = y_0 < y_1 < .... < y_n = d \tag{2}$$

$$e = z_0 < z_1 < .... < z_l = f \tag{3}$$

So the m-th power Lagrange basis function about x is:

$$l_i(x) = \prod_{\substack{t=0 \\ t \neq i}}^{m} \frac{x - x_t}{x_i - x_t} \tag{4}$$

The n-th power Lagrange basis function about y is:

$$l_j(y) = \prod_{\substack{t=0 \\ t \neq j}}^{n} \frac{y - y_t}{y_j - y_t} \tag{5}$$

The l-th power Lagrange basis function about z is:

$$l_k(z) = \prod_{\substack{t=0 \\ t \neq k}}^{l} \frac{z - z_t}{z_k - z_t} \tag{6}$$

So the three-dimensional Lagrange basis function is the product of each basis function:

$$l_{ijk}(x, y, z) = l_i(x) l_j(y) l_k(z) \tag{7}$$

Finally the Lagrange interpolation polynomial is:

$$P(x,y,z)=\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{l}f(x_i,y_j,z_k)l_{ijk}(x,y,z)=\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{l}f(x_i,y_j,z_k)\prod_{\substack{t=0\\t\neq i}}^{m}\frac{x-x_t}{x_i-x_t}\prod_{\substack{t=0\\t\neq j}}^{n}\frac{y-y_t}{y_j-y_t}\prod_{\substack{t=0\\t\neq k}}^{l}\frac{z-z_t}{z_k-z_t} \tag{8}$$

Can prove that the equation (8) is unique[15].In this polynomial $f(x_i, y_j, z_k)$ is the temperature sensor value in the direction vector(x_i, y_j, z_k) of the cold chain vehicle body. The direction vector (x,y,z) is the point that the temperature to estimate. And the P(x,y,z) is the point(x,y,z) estimated temperature.

3.2 Experiment of Space Temperature Modeling

Suppose that the length of the cold chain vehicle body is L, the width of the cold chain vehicle body is W, the height of the cold chain vehicle body is H. In this experiment L is 5.13 meters, W is 2.91 meters and H is 3.86 meters. The temperature sensor parameter about location and temperature value is shown in table 3.

Table 3. The location and temperature value about the 8 temperature sensors

ID	Location	Value
1	(0,0,0)	4.11
2	(0,0,L)	3.16
3	(0,W,0)	4.67
4	(0,W,L)	4.12
5	(H,0,0)	3.10
6	(H,0,L)	4.95
7	(H,W,0)	3.34
8	(H,W,L)	4.23

The 8 temperature sensors was installed in the 8 vertex of the cold chain vehicle body. The temperature sensors in each vertex of cold chain vehicle body and its value are shown in figure 3.

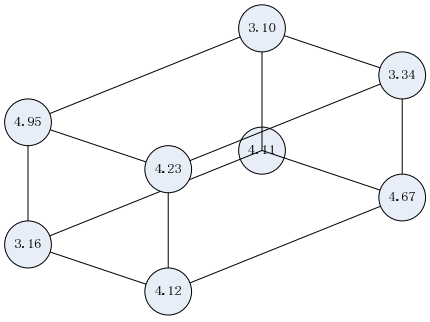


Fig. 3. The position of the 8 temperature sensors and the value of each temperature sensor are displayed

By the use of three-dimensions Lagrange interpolation, the temperature value in anywhere of cold chain vehicle body can be estimated. Figure 4 shows the graph of both temperature ruler and temperature distribution.

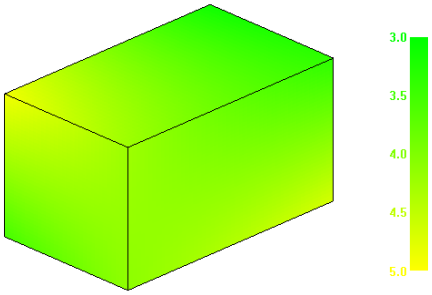


Fig. 4. The temperature ruler and temperature distribution are displayed. If the temperature is near 5 Celsius degrees the color turn to yellow, and if the temperature is near 3 Celsius degrees the color turn to green.

4 Conclusions

This paper proposed a cold chain logistics monitoring system with temperature modeling, and the experiment of this system indicates that:

1. By method of three dimensions Lagrange interpolation, the temperature in anywhere of the cold chain vehicle body can be estimated accurately.
2. The information includes latitude and longitude of cold chain vehicle is also obtained. It reflects the geographic location of the cold chain vehicle.
3. The transmission of all the data is in real-time.

So that the manager in the cold chain logistics company just sit in front of the control center computer he can view the information about the temperature distribution and the geographic location of each cold chain vehicle in the company. And the efficiency of the management in cold chain logistics company will be enhanced by this cold chain logistics monitoring system.

Acknowledgments. This paper is supported by Chinese Universities Scientific Fund (2012QT003).

References

1. Dong, H.: China's agricultural products Cold-chain logistics status, Problems and Solutions. Ecological Economy (10), 255–257 (2009)
2. Liu, L., Li, J.: Summary of Development Pattern and Government Behavior of Cold Chain Logistics of Agricultural Products. Food Science 29(9), 680–683 (2008)

3. Wu, Q.: The Current Situation and the Countermeasures of China's Cold Chain Logistics Development. China Business and Market 25(2), 24–28 (2011)
4. Wu, J., Jin, T.: Comparative Study on the Development Levels of Domestic and International Cold Chain Logistics for Agricultural Products Based on AHP. Journal of Anhui Agricultural Sciences (6), 3663–3665 (2012)
5. Zhang, X., Xing, S., Fu, Z., Tian, D.: Current situation, development trend and countermeasures of aquatic products cold-chain logistics technology. Fishery Modernization 38(3), 45–49 (2011)
6. Li, P., You, T.: Cold chain recording instrument of temperature and humidity tracking based on iButton-DS1923. Mechanical & Electrical Engineering Magazine 26(10), 64–68 (2009)
7. Wang, T., Zhang, X., Chen, W., Fu, Z., Peng, Z.: RFID-based temperature monitoring system of frozen and chilled tilapia in cold chain logistics. Transactions of the Chinese Society of Agricultural Engineering 27(9), 141–145 (2011)
8. Lin, Z., Mo, X., Li, H., Li, H.: Comparison of Three Spatial Interpolation Methods for Climate Variables in China. Acta Geographica Sinica 57(1), 47–56 (2002)
9. Feng, T.: Multivariate Polynomial Interpolation. Journal of Chongqing Three Gorges University 25(3), 129–132 (2009)
10. Maxim Integrated Products, http://datasheets.maxim-ic.com/en/ds/DS18B20.pdf
11. Yoon, B.J., Park, M.W., Kim, J.H.: UGV (Unmanned Ground Vehicle) Navigation Method using GPS and Compass. In: SICE-ICASE International Joint Conference (2006)
12. Zhai, Z., Cai, S.: GPRS/GPS/GIS Based Vehicle Navigation and Monitoring System. Bulletin of Surveying and Mapping (2), 34–36 (2004)
13. Xia, X., Xu, H., Lai, J., Wen, A.: Wireless Meter Reading Based on GPRS. Techniques of Automation and Applications 28(3), 62–64 (2009)
14. Liu, G.: Design and Implementation of PPP in uC/OS-II Operating System. Degree thesis, Northeastern University (2005)
15. Wei, L.: Polynomial Interpolation Problem in R~S Space. Degree thesis, Dalian Jiaotong University (2009)

Wheat Three-Dimensional Reconstruction and Visualization System*

Hao Zhang[1], Qiang Wang[1], Hui Zhang[1], Yali Ji[1], Xinming Ma[1,2], and Lei Xi[1,**]

[1] Information and Management Science College, Henan Agricultural University, Henan, P.R. China
{zhanghaohnnd,hnaustu}@126.com
[2] Agronomy College Henan Agricultural University, Henan Agricultural University, Henan, P.R. China
xinmingma@126.com

Abstract. The study object is wheat. Through the observation and analysis on the morphological structure of wheat plant, the concepts of macrostate, substate and microstate are put forward, the wheat growth mechanism model is established, and thus the description and control of wheat morphogenesis process could be realized. Based on the branch structure characteristics of plant, the wheat morphological data model based on axis structure is proposed to realize the structural storage of plant topological structure and organ morphological feature data. On this basis, based on the thought of "growth model - morphological feature model - geometric modeling - display model", the technical framework of wheat form visualization is established and on the VC++ platform, the virtual wheat growth system is built with OpenGL. The running result of the system shows that wheat morphological feature can be well simulated with the system to realize the virtual display of the growth process in the individual growth period of wheat.

Keywords: Wheat, Plant morphological, Virtual, Visualization.

1 Introduction

Virtual plant has been the research hotspot in digital agriculture, virtual reality and other interdisciplinary fields. It is aimed at revealing the complex relationship among plant growth, physiological process and environmental factors with a dynamic and visualized virtual scene, therefore, it has bright prospect in agriculture, forestry, ecology and remote sensing. A lot of researches have been carried out on virtual plants at home and abroad. Among them, in terms of the reconstruction of plant morphological structure, L system established by the University of Calgary, Canada[1], the AMAP system by CIRAD in France[2] and the dual-scale automata

* This work is partially supported by China National 863 Plans Projects #2008AA10Z220 and key scientific and technological project of Henan Province #082102140004.
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 234–243, 2013.

proposed by Sino-French Laboratory of Institute of Automation, Chinese Academy of Sciences[3] and so on are included. L system focuses on the expression of plant topological structure, and describes plant morphology and growth pattern with abstract rules, but its defect lies in that it is too complex, and it is very difficult to extract and define the growing rule of a specific type of plant. AMAP and dual-scale automata can well describe the plant physiological feature and growth process, however, AMAP is suitable for modeling tall plants, and dual-scale automata cannot describe the organ growth state, thus, it is not fully applicable to the reconstruction of wheat plant morphology. In terms of the organ morphological modeling based on process, Mengjun, et al. described leaf vein curve with leaf specific weight, leaf inclination angle and leaf length, they could accurately describe the curl and distortion of wheat leaves in combination with the leaf shape and the geometric modeling of leaf edge; Wu Yanlian constructed the geometric model based on the characteristic parameters of organ morphology, including leaf, stalk and spike[5-6], which could dynamically describe the formation process of organ. In terms of the research on wheat three-dimensional visualization system, Chen Guoqing et al. proposed wheat three-dimensional visualization technical framework based on growth model[7], but the structural coupling of morphological model, growth model and visualization model was not taken into consideration. In addition, the structural storage of wheat plant topological structure and organ morphology characteristic data has not been reported.

On the basis of referring to the existing research findings and the observation and analysis on morphological structures of wheat plant, the concepts of macrostate, substate and microstate are put forward, the wheat growth mechanism model is established, and thus the formalized description of wheat morphogenesis process is realized. Through the analysis on the branch structure of wheat plant, the wheat morphological data model based on axis structure is proposed to realize the structural storage of plant topological structure and organ morphological feature data. On this basis, according to the thought of "growth model - morphological feature model - geometric modeling - display model", the visualization framework of wheat morphology is designed. On the VC++ platform, the virtual wheat growth system is built with OpenGL to realize the three-dimensional visualization of the morphological changes of organ and individual plant of wheat in the whole growth period.

2 The Technical Framework of Wheat Form Visualization

To accurately reconstruct wheat three-dimensional morphology, the technical framework of morphological visualization of wheat is designed by following the thought of "growth model - morphological feature model - geometric modeling - display model". The framework consists of virtual environment, growth model base, the model base of organ morphological characteristics, growth engine, visual engine, and wheat morphological data model and other components, as shown in Fig.1. 1) The virtual environment mainly includes weather data, soil information, variety data and cultivation management data. 2) Wheat-Grow is adopted for wheat growth model,

and the system mainly uses the output growing degree days, leaf area index, nutrients, water and other influencing factors[8-11]. 3) Organ morphology characteristic parameter model is mainly composed of 4 parts, morphology characteristic parameter models of leaf, stalk, sheath and wheat head[6,12-13]. 4) The growth engine is made up of macrostate trigger, blade element trigger, growth engine drive, and branching (tillering) trigger and organ extension trigger. 5) Visualization engine is composed of engine drive, organ geometric modeling model and visualization model. 6) Wheat morphological data object is used to store plant topological structure and organ morphological characteristic data.

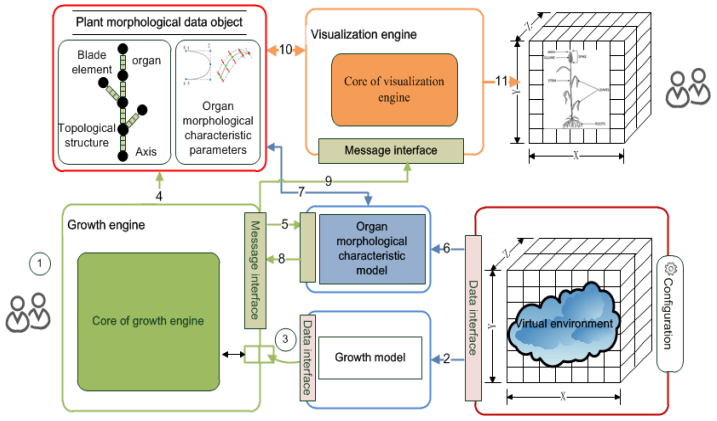


Fig. 1. Technical framework of morphological visualization of wheat

Driven by the growth engine, the system invokes the growth model and morphological characteristic model of organ, calculates the physiological age of the wheat according to the virtual environment data, and then saves the plant topological structure at the physiological age and the related morphological characteristic parameters of leaves, stalks, nodes and spikes into the wheat morphological data model. The visualization engine receives the visual message sent by the growth engine, conducts precise geometric modeling of the organ according to the wheat morphological data model and splices the organs according to the topological structure of the organ to form an individual so as to realize the demonstration of three-dimensional configuration and visual computing function of wheat organs and individuals.

3 Wheat Growth Mechanism Model

Wheat growth mechanism model is composed of macro automata, child automata and micro automata.

The growth process of wheat is a unidirectional and irreversible evolutionary process, which passes through a number of distinct physiological stages, namely

growth elements. In the model, macro-state represents the growth unit. Blade element, the meristematic unit of plant, is represented by sub-state. Since the growth unit is composed of blade elements, macro-state is constituted by sub-states, moreover, the blade elements in the same growth unit are in the same physiological stage. Changes of blade elements are embodied by the meristematic organs forming blade elements, and macro-state represents the growth state of the meristematic organs making up of blade elements.

According to the analysis mentioned above, the topological structure of wheat is generated by the growth mechanism model of wheat through the combination of macro-state, sub-state and micro-state and cycle simulation of growth process of wheat. In the system, the growth mechanism model is realized by the growth engine.

A. Macro Automata

Macro automata could represent the state transition among the growth units of wheat. It is defined as a hexahydric group: AP ::= $<Q_p,\pi_p , S_{GU} ,C_p,\delta_p ,F_p>$. Where, Q_p is the set of the finite state of macro automata, πp is the initial probability vector$(\pi_{pj})_j = (p (Q_{pi} = j))_j$; S_{GU} is the clock cycle that is needed for the growth of a growth unit; C_p is the time of macro-state circulation; $\delta_p \subseteq Q_p \times Q_p$ is the transition condition between states; Fp is the termination condition. The parameters in the automata are determined with the aforementioned wheat growth model in the system.

B. Child Automata

Child automata could express the state transition relationship among blade elements within a growth unit. Blade elements of wheat are classified as ml type and mt type; the former is composed of leaf, sheath and internode and the latter is constituted by spike and internodes under spike.

Child automata are defined as a hexahydric group, including AC::= $< Q_c, \pi_c , S_{ME} , C_c ,\delta_c$ and $F_c>$. Q_c indicates the set of finite state of child automata and πc is defined as a vector of initial probability $(\pi_{cj})_j = (p (Q_{ci} = j))_j$. S_{ME} is the clock cycle during which a blade element outgrows. Cc expresses the cycling times of child automata. $\delta_c \subseteq Q_c \times Q_c$ indicates transition conditions between states and F_c is the termination condition. Parameters of child automata could be determined by phyllochron and growth model.

As wheat grows in cluster with a main stem and several lateral stems (tillering), 2 types of mechanisms need to be introduced in order to reflect wheat morphogenesis truthfully. 1) One is branching (tillering) mechanism which could decide time, quantity and disappearance of tillering and effects of external environment on tillering. A great number of studies have been carried out on this aspect[14-16]. 2) The other is synchronous mechanism. Lateral axis of wheat (tillering) has synchronous growth relation with its parent axis and its physiological age is the same to that of parent axis. However, the number of its meristematic units can not be greater than that of its parent axis. Therefore, when parent axis reaches a certain physiological stage, its lateral axis will also reach synchronously.

C. Micro Automata

Micro automata show transition relationship between specific states of organs within the blade element and are defined as a hexahydric group, including AM::= < Q_m, π_m , S_{OG} , C_m , δ_m and F_m>. Q_m indicates the set of finite state of micro automata and πm is defined as a vector of initial probability (π_{mj}) $_j$ = (p (Q_{mi} = j))$_j$. S_{OG} is a clock cycle in which the micro automata could change. $\delta_m \subseteq Q_m \times Q_m$ indicates transition conditions between states and F_m is the termination condition. Micro automata could be decided by rule of synchronous growth of wheat organs.

4 Plant Morphological Data Model of Wheat Based on Axis Structure

Plant morphology of wheat is manifested as branch structure composed of axes with the main axis and lateral axes of all levels included (tillering). Essence of plant morphological description of wheat is to reveal the following aspects: 1) Branch network, namely axis information, such as branch numbers of axis (tillering), blade element number, organ set, arraying order, etc.. Logical relation between axes and morphological description are also included, such as mapping relationship between branching (tillering) and parent axis, angles between axes, etc.. 2) Morphological characteristics of organs, such as length and width of leaf organ, angle between stem and leaf, length and thickness of internode organ, etc.. On this basis, plant morphological data model of wheat is constituted by set of axis objects, axis objects and organ objects, which describes relationship among them and morphological characteristics of organs from three aspects of axis, blade element and organ.

A. Set of Axis Objects

Set of axis objects is showed in Fig.2. As elements of a set, axis objects of all levels are ordered according to their occurrence time.

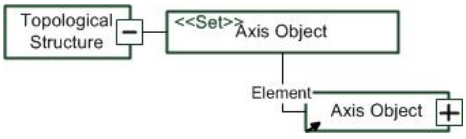


Fig. 2. Set of axis objects

B. Axis Object

Structure of axis object is shown in Fig.3. 1) Identity of axis object is its key which reflects logical relations between axes. 2) Branching (tillering) number indicates number of branches in an axis. 3) Blade element number indicates number of blade elements in an axis. 4) Angle of parent axis shows the angle between an axis and its parent axis. 5) Leaf object set is the set of leaf organs constituting the current axis and its elements are leaf objects. 6) Sheath set is the set of sheath organs constituting the current axis and its elements are sheath objects. 7) Internode object

set is the set of internode organs constituting the current axis and its elements are internode objects. 8) As to spike object, axis which could grow into a spike only has one spike organ.

In actual application of this model, organs constituting the blade elements could be added according to practical requirements. Morphological characteristic data of every organ object are recorded, which is not described in detail here.

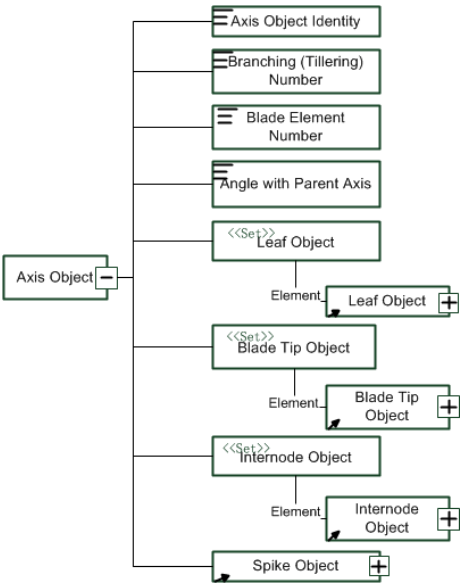


Fig. 3. Axis object

5 Geometric Modeling of Wheat Organ

A. Geometric Modeling of leaf

Wheat leaf is constituted by blade and sheath. Sheath is under leaf and it is in an open cylindrical shape, surrounding the internode completely. NURBS surface modeling is applied to both sheath and blade in this study. Based on length and width of leaf, angle between stem and leaf, sheath length and other parameters output by organ morphological characteristic model mentioned above, control points of NURBS surface are determined. Control points of blade are included in 5 rows and 7 lines. Control points of the 4th column of every row are decided by leaf curve, other control points are on two sides of control points in 4th row when leaf is unfolded and the distance between two points is decided by leaf width. Geometric modeling of untwisted blade is shown in Fig.4.

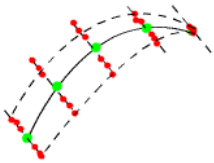


Fig. 4. Geometric schematic diagram of untwisted blade

Twisted modeling of blade could be regarded as rotating the center-top part of blade around curve of leaf vein under the condition of unchanged vein curve. In terms of NURBS surface, it is manifested as rotating control points in 2nd and 3rd rows on blade around curve of leaf vein. As shown in Fig.5, l1, l2, l3 and l4 are lines connected by the control points in last 4 rows (control points in 5th row are all at the highest point of blade) and p1, p2, p3, p4 and p5 are 5 control points in curve of leaf vein. Through rotating 14 and 13 around tangent line of curve of leaf vein, twisted modeling could be achieved, namely, rotating 14 around the vector which passes point p4 and is parallel to p3p5 with an angle of ro and rotating 13 around vector which passes p3 and is parallel to p2p4 with an angle of ro. Twisting angle of r is within 0-180°.

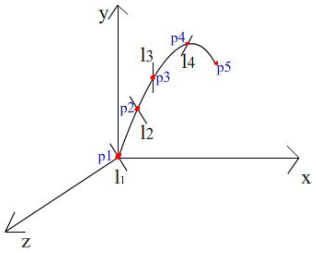


Fig. 5. Geometric schematic diagram of twisted blade

Sheath of wheat is in an open cylindrical shape. Its modeling is mainly determined by radius R and height H. Control points are set to be included in 5 rows and each row has 7 control points. The first row is set to be a square with a side length of 2 x R which is defined as a circle with a radius of R. The x and z coordinates of control points in row 2, 3 and 4 are equal to those of row 1 and y coordinate has an equidistance rising. The y coordinate of control points in row 5 increases with the same space and x and z coordinates form into an open square, thus defining an open circle, as shown in Fig.6.

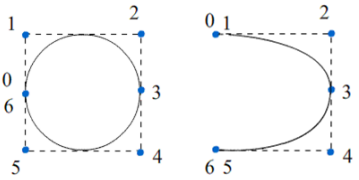


Fig. 6. Geometric schematic diagram of sheath

In order to ensure smooth connection between blade and sheath, y coordinate of control points of blade in first row could be set to be equal to that of control points in the highest part of sheath. Thus, control points in the 10 rows and 7 columns could form a NURBS surface to simulate the modeling of whole leaf, as shown in Fig.7.

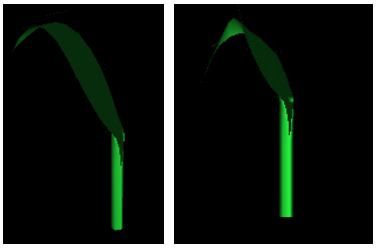


Fig. 7. Morphological visualization of blade and sheath of wheat

B. Geometric Modeling of Internode

Wheat internode is manifested as a cylinder which grows to be thick and long gradually during the growth process and this is simulated by cylinder in quadric surface. Based on internode length and thickness output by morphological characteristic model of aforementioned organ, length and diameter of cylinder are determined.

C. Geometric Modeling of Wheat Spike

Structure of spike is relatively complex and could be divided into grain, axis, peduncle and awn. Therefore, method of combining several basic graphics was adopted to construct the three-dimensional geometric model of spike. Axis and pedunde of spike were modeled by cylinder in quadric surface. For species which had awn, model construction of awn adopted cylinder and ellipsoid was used to construct grain model. Visualization effects of spike are shown in Fig.8.

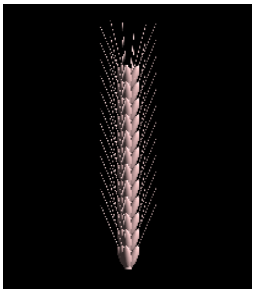


Fig. 8. Morphological visualization of spike of wheat

6 Example Analysis and System Implementation

Based on the technical frame of wheat morphological visualization technology, virtual growth system of wheat was developed using VC + + and OpenGL. With the variety of "Yumai 34" used as the test materials, morphological characteristic parameters of "Yumai 34" were extracted based on meteorological data, soil characteristics, variety

parameters, cultivation measures, etc. during wheat growth in Science and Educational Garden of Henan Agricultural University in 2008 and 2009. With variety parameters and model parameters extracted, plant topological structure and morphological characteristic parameters of all organs were generated using virtual wheat growth system. Three-dimensional morphology of all organs was drawn and three-dimensional morphological reconstruction of wheat individual was achieved, as shown in Fig.9.

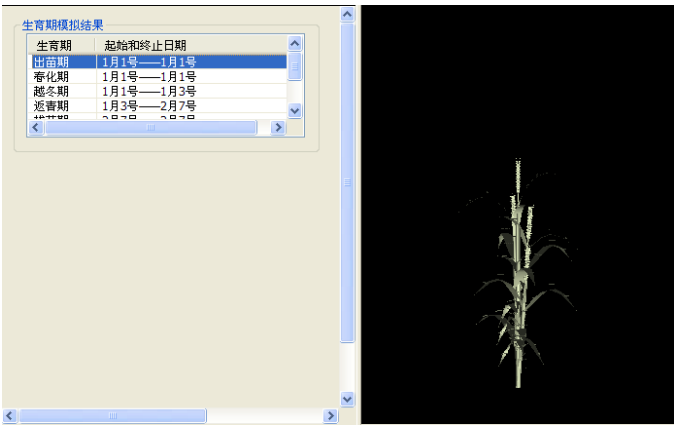


Fig. 9. Interface of virtual wheat growth system

7 Conclusions

1) Following the thought of "growth model-morphological characteristic model-geometric modeling-visualization model", technical frame of morphological visualization of wheat was put forward in this study based on growth automata model of wheat, plant morphological data model and geometric modeling of organ. This frame combined the morphological model, growth model and visualization model organically. Based on the technical frame, virtual wheat growth system was established which could achieve three-dimensional visualization of wheat plant organ and morphological change in the full growth period of an individual.

2) Concepts of macro state, child state and micro state were put forward in this study from the perspective of botany and wheat growth automata model was constructed. Through introducing organ growth state, this method could reflect changes of blade element during growth precisely, which overcame the deficiency of dual-scale automata in this aspect. Besides, growth mechanism of wheat was also considered in this model and it combined with the agricultural knowledge effectively, facilitating simulation of wheat morphological process. This model had certain universality and could offer ideas and methods for studies on morphology and structure of other cereal crops.

3) Plant morphological data model of wheat based on the axis structure was put forward in this study and this achieved structural storage of topological structure of wheat plant and morphological characteristic data of organ. Combination of all organs could be controlled conveniently and effectively using this model and virtual display

of the wheat individual was achieved. This method was also applicable to structured storage of morphological data of other cereal crops.

4) This study did not involve the crop root. Therefore, further studies should be carried out on how to construct the morphological model of crop root so as to achieve the complete visualization of crop organ, individual and population.

References

1. Prusinkiewicz, P., Lindenmayer, A.: The Algorithmic Beauty of Plants. Springer, Heidelberg (1996)
2. De Reffye, P., Fourcaud, T., Blaise, F., Barthelemy, D., Houllier, F.: A Functional Model of Tree Growth and Tree Architecture. Silva Fennica 31(3), 297–311 (1997)
3. Xing, Z., de Reffye, P., Fan-Lun, X., Bao-Gang, H., Zhi-Gang, Z.: Dual-scale Automaton Model for Virtual Plant Development. Chinese Journal of Computers 24(6), 608–615 (2001)
4. Jun, M., Xinyu, G., Chunjiang, Z.: Geometry Modeling and Visualization of Above-ground Organs of Wheat. Journal of Triticeae Crops 29(1), 106–109 (2009) (in Chinese)
5. Yanlian, W., Weixing, C., Liang, T., Yan, Z., Hui, L.: OpenGL-based Visual Technology for Wheat Morphology. Transactions of the CSAE 25(1), 121–126 (2009) (in Chinese)
6. Guoqing, C., Yan, Z., Weixing, C.: Modeling Leaf Sheath and Internode Growth Gynamics in Wheat. Journal of Triticeae Crops 25(1), 71–74 (2005) (in Chinese)
7. Guoqing, C., Yan, Z., Hui, L., Weixing, C.: Morphogenesis Model-based Virtual Growth System for Organs and Plant of Wheat. Transactions of the CSAE 23(3), 126–130 (2007) (in Chinese)
8. Meichun, Y., Weixing, C., Weihong, L., Haidong, J.: A Mechanistic Model of Phasic and Phenological Development of Wheat I. Assumption and Description of the Model 11(3), 355–359 (2000) (in Chinese)
9. Meichun, Y., Weixing, C., Weihong, L., Shaohua, W.: A Simulation Model of Above-ground Organ Formation in Wheat. Acta Agronomica Sinica 27(2), 222–229 (2001) (in Chinese)
10. Tiemei, L., Weixing, C., Weihong, L.: A Simulation Model of Photosynthetic Production and Dry Matter Accumulation in Wheat. Journal of Triticeae Crops 21(3), 26–30 (2001) (in Chinese)
11. Weixing, C., Tiemei, L., Weihong, L.: Simulating Organgrowth in Wheat Based on the Organ-weight Fraction Concept. Plant Production Science 5(3), 248–256 (2002) (in Chinese)
12. Zihui, T., Yan, Z., Xia, Y., Yongchao, T., Xiaojun, L., Weixing, C.: Modeling Spike Growth Dynamics in Winter Wheat. Journal of Triticeae Crops 26(4), 93–97 (2006) (in Chinese)
13. Guoqing, C., Yan, Z., Weixing, C.: Modeling Leaf Growth Dynamics in Winter Wheat. Acta Agronomica Sinica 31(11), 1524–1527 (2005) (in Chinese)
14. McMaster, G.S., Klepper, B.: Simulation of Shoot Vegetative Development and Growth of Unstressed Winter Wheat. Ecological Modeling (53), 71–74 (1991)
15. Yushan, S.: A Staistic Analysis of the Development of Tillerings in Winter and Its Use in Production. Acta Agronomica Sinica 8(11), 49–56 (1982) (in Chinese)
16. Tiemei, L., Weixing, C., Weihong, L., Jie, P., Meichun, Y., Wenshan, G.: Simulation on Wheat Tillering Dynamic. Journal of Huazhong Agricultural University 20(5), 416–421 (2001) (in Chinese)

Study on Agricultural Condition Monitoring and Diagnosing of Integrated Platform Based on the Internet of Things

Jinying Yu and Wei Zhang

Information Institute of Science and Technology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

Abstract. In order to meet the needs of modern agriculture, with the internet of things and related technologies, the author has developed a set of agricultural condition monitoring and diagnosing of integrated platform for experiment. The platform used the client and server mode, include environmental factor of monitoring platform, intelligent decision sub platform and expert remote science and technology advisory services subsidiary platform. Environmental factors collected by data acquisition module, and transmit through network. Audio and video data used the XMPP protocol extension Jingle transmission; Use the knowledge base, fuzzy theory, self learning algorithm of intelligent to achieve decision-making and control. The platform has been experimented in the vegetable greenhouses, which can carry out remote supervisory and automated control and expert remote diagnosis and consultation, improving the facilities agriculture management level.

Keywords: Internet of Things, Agricultural condition monitoring, Automatic regulation, Distance-vision consulting and diagnosing.

Introduction

With the development of information technology, more and more intelligent systems have been applied to agricultural production. The strawberry grower Norcal Harvesting living in Oxnard of California the United States, installed a set of networking system developed by Climate Minder. According to the air and the soil condition, the system can real-time track plant condition. It can automatically trigger the related behaviors, such as water or adjust the temperature [1]. The system based on GPRS wireless communication and the internet technology has been developed in the Jiangsu University, which can remotely collect and monitor the environment factor data in the aquatic products breeding[2]. Zhengyi Ren and Juanjuan Pan constructed a system that can remote diagnostics and monitoring the facilities gardening pests [3]. The Information Institute of Science and Technology Beijing Academy of agriculture and forestry sciences developed a remote video consultation and diagnosis system [4]. But all of the above systems are not forming an organic whole. The agricultural condition monitoring and diagnosing of integrated platform

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 244–250, 2013.

based on the internet of things can organically integrate the remote monitoring the produce environmental factor, the video surveillance of production site, automatic control and remote video consultation for the diagnosis all together, which can improve the agricultural fine management level.

1 Platform Construction

The agricultural condition monitoring and diagnosing of integrated platform based on the internet of things includes environmental factor of monitoring platform, intelligent decision sub platform and expert remote science and technology advisory services subsidiary platform. It is mainly composed of a sensor, camera, internet, intelligent decision and other network technology and device.

1.1 The Environmental Factor of Monitoring Platform

The environmental factor of monitoring platform is mainly composed of a sensor, camera, switches and other components. Various sensors are responsible for collecting each environmental factor data, transmission to the server by the network. The data collected by the camera transmit to the concentrator. All the real time environment factor data and the video data concentrate through the switch, then transmit the data to the server by the network. The platform used the cable and the switch completes the production monitoring of the temperature and humidity, illumination, carbon dioxide and other environmental factors and the video monitoring. The installation is simple, and has the strong stability and the strong data exchange [5-7].

1.2 The Intelligent Decision Sub Platform

The intelligent decision sub platform comprises input module, output module, threshold module, decision module, alarm control module. All kinds of monitoring data input through the input module. When the data reaches the threshold, the platform will generate an alarm signal, sound, Email, SMS and other forms of real-time alarm, for reminding the user to accurately grasp the watering, shading and other in the appropriate time; According to the abnormal situation, decision module can also independently find solutions from the knowledge base, then the platform transfers the control information through the output module to each actuator, automatically control temperature, humidity etc.with right measures.

1.3 The Expert Remote Science and Technology Advisory Services Subsidiary Platform

The expert remote science and technology advisory services subsidiary platform is mainly composed of the video module, audio module, text communication module, application sharing module, file transmission module, communicated in XMPP

protocol. The video module carries out video compression, transmission and decompression. The audio module provides the voice compression, transmission and decompression, and the video module combines with the audio module, both of the modules can realize the remote "face to face" real time communication between the experts and the users for remote diagnosis. When the users and experts need to communicate by text, the text communication module will come into play. Through the application sharing module the experts can reveal some applications to the user, such as PPT, word, furtherly they can transmit information to the users through the file transmission module. According to the need, the platform also can carry out remote training.

2 The Structure and Principle of Platform

2.1 The structure of platform

The structure of platform is divided into two parts, server and client. The server locates in the center, which is responsible for the exchange of information and information management. The clients are divided into two genres: one genre is the information collection client, including the sensors of temperature, humidity, light and carbon dioxide and other environmental factors, camera as well as intelligent decision platform. The other genre is the user client, including the internet monitoring advisory terminal and the mobile phone monitoring advisory terminal [8, 9]. In the course of practical application, the environmental factor of monitoring platform continuously collects data, when the data is abnormal, the intelligent decision sub platform will automatically alarm and control; and the users can view the environment factor of real-time data, history data and video scene through the client; when the users meet with the complex problems, they can remotely contact with expert "face to face".

2.2 The Principle of Platform

The server-side realize environmental factor data acquisition and management, video data collection and management, expert "face to face" information communication and management with the XMPP protocol, Socket technology, combination authentication. The environmental factor data can be operated with the smart sensor, RJ45 modules, cable, Socket technology, network camera, etc. With H.264 media encoding technology, G.723.1, AMR coding technology, XMPP protocol the platform can realize remotely expert "face to face" consulting diagnosis. The intelligent decision sub platform using the knowledge base, fuzzy theory, self learning algorithm to decision control. The multi point control unit through the XML language realizes each module message control. The user client using H.264 media encoding technology, G.723.1, Socket technology, XMPP protocol, combination authentication technology, implements consulting diagnosis, video monitoring, environment monitoring and other functions. Technical diagram as shown below.

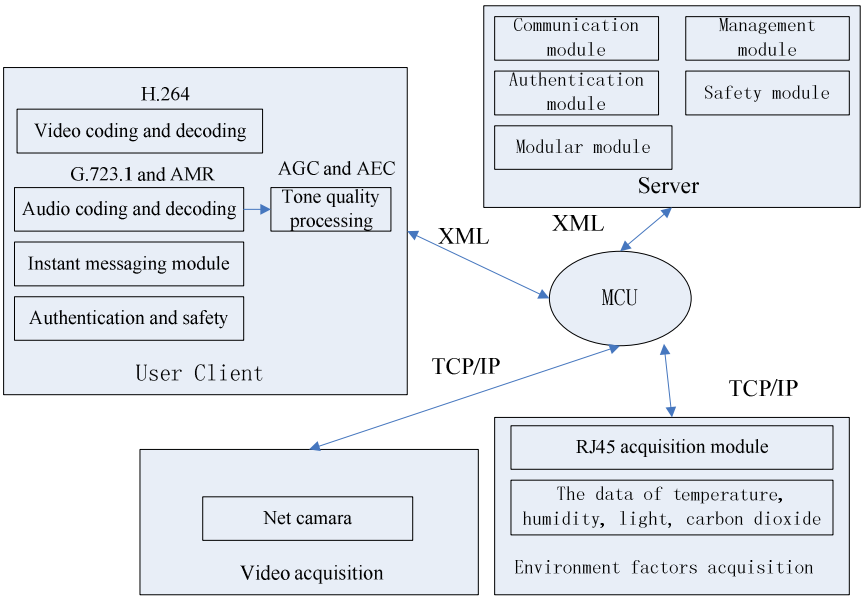


Fig. 1. System structure diagram

3 Platform Test and Analysis

3.1 Experimental Design

The vegetable greenhouse cultivation of cucumber has been designed as an experimental environment and the author has completed the crop condition monitoring diagnosis experiment [10-12], in the experiment, the author has designed four sensors such as Temperature, humidity, light, carbon dioxide ,etc. All the data are collected through the RJ45 collection module and they are transmitted to the server through the network .The users can monitor the data through the Web; the author has designed the threshold for each factor, when the monitoring value exceeds the threshold ,it will alarm and tell us which environmental factor is unnormal in the greenhouse. The platform has been placed on the internet; The experts can face to face communicate with the users and guide the users in actual production through the XMPP server.

3.2 Experimental Materials

Four sensors: the temperature sensor, the humidity sensor, the light sensor, the carbon dioxide sensor (supply voltage of 12V DC, output 4-20mA current signal); RJ45 acquisition module (8 inputs, the network output); the Server; Framework 3.5; the PC computer terminal with USB camera; the heating equipment: hot stove.

3.3 Experimental Method

The experimental utilize the field tests. In the 70 m×10 m vegetable greenhouses, the experimenters have installed one temperature sensor, one humidity sensor, one light sensor, one carbon dioxide sensor in the center. The experimenters have erected one network camera at the end of the greenhouse. In order to complete the experiment, the experimenters in the greenhouse will man-made change some environmental factor specifically, such as use the hot stove to enhance the greenhouse temperature to check whether there is an alarm. Experts can remote access check local actual condition by the PC.The local experimenters also through the local computer terminals check the experts' calls.

3.4 Experimental Results

Fig. 2. The experimental results

In the Experiment the various functions of the system platform have been verified. The expert can real-time monitor the vegetable greenhouse through the browser, he can read each environmental factor, such as data, video etc., check the local alarm information of the high temperature, remotely "face to face" exchange successfully with local experimenters. The experimental result show as above Fig.2, the expert is viewing the environmental factors data.

4 Conclusion and Discussion

The agricultural condition monitoring and diagnosis of integrated platform based on the internet of things creatively put environmental factor remote monitoring, remote monitoring of production site, production automation control and remote video

consultation for the diagnosis organically together. All the data of environmental factors and the live video are collected through the network cable. They can transfer to the server side by the switch. The intelligent decision sub platform on the server can compare the environmental factors data with the threshold all the time. When the monitoring value exceeds the threshold, the system will alarm, and the sub platform will make an order to the actuator to open or close, to realize automatic regulation; the users who have known the alarm can manual control by themselves. The users can remotely view site status at any time. When the plant diseases and there are insect pests, after logining the experts can not only see crop diseases, they also can view real-time and historical data, which is significant to improve the accuracy of diagnosis to plant diseases and insect pests, timely to resolve the disease, to reduce the loss .

Along with the internet technology development, the platform function need to be extended to further standardize; how to further standardize networking standards, how to transplant 3G mobile network platform to realize mobile monitoring and consulting diagnosis etc, all these need to continue to study and explore [13-16].

References

1. Jin, P.: With the internet of things to promote the modern facility agriculture. Agricultural Machinery Market 3, 27–28 (2010)
2. Fan, Q., Lai, R., Qiu, Q., Chen, Y., Lin, T., Zhang, X.: Development and application of online distant diagnosis of tobacco diseases and insects. Chinese Tobacco Science 29(1), 60–61 (2008)
3. Ren, Z., Pan, J., Gu, P., Yang, J., Su, N., Ren, A., et al.: Establishment of diagnosis and remote monitoring system for diseases and pests in protected cultivation. Acta Agriculturae Boreali-Occidentalis Sinica 19(3), 62–65 (2010)
4. Zhang, W., Yu, J., Yu, F., Luan, R.: Study on agricultural distance monitoring and diagnosing integration platform based on XMPP. Chinese Agricultural Science Bulletin 27(11), 151–154 (2011)
5. Wang, Y., Bai, Z.: Design of system auto controlling and measurement of parameters of environment formodern greenhouse. Microcomputer Information 24(20), 140–141 (2008)
6. Ding, X., Jin, H., Zhang, H., Yu, J.: Changes of illuminations and light transmissivities in different greenhouses. China Cucurbits and Vegetables 24(1), 1–4 (2011)
7. Zhang, F., Bian, X.: Middle ware technology of internet is the important links in internet of things chain. Sci-tech Innovation and Productivity 3, 41–43 (2011)
8. Wu, H.: The enter of 3G applications into the era of the internet of things. Journal of Changchun Normal University 30(1), 66–68 (2011)
9. Wang, D., Nie, Y., Wang, L., Li, C.: Research on strategic conception of green agricultural products application platform of internet ofthings in Heilongjiang reclamation Area. Ecological Economy 233(12), 145–147 (2010)
10. Jin, P.: The internet of things application in the modern facility agriculture. Agriculture Engineering Technology (Agricultural Product Processing Industry) 5, 40–41 (2010)
11. Jin, P.: The internet of things and the modern facility agriculture. Agriculture Machinery Technology Extension 2, 45 (2010)

12. Lai, W.: Agricultural internet of things application in the modern flower industry. China Flowers & Horticulture 20, 28–29 (2010)
13. Sun, Z., Du, K., Yin, S.: Development trend of internet of things and perspective of its application in agriculture. Agriculture Network Information 5, 5–8 (2010)
14. Cao, J., Fan, Y., Zhu, K., Wang, Z., Dai, Q.: Analysis of internet of things in Jiangsu agricultural development. Jiangsu Journal of Agricultural Sciences 26(6), 1402–1405 (2010)
15. Zhao, J., Wang, Y., Yang, M., Tan, X.: IOT application in agricultural pest disasters. Communications Technology 43(11), 49–51 (2010)
16. Ma, X., Wang, W., Han, J., Yuan, S., Li, P., Pang, C., et al.: To accelerate the realization of agricultural modernization with internet of things technology. Journal of Shanxi Agricultural Sciences 39(4), 376–378 (2011)

The System of Anti-bud Injury in Seedcane Cutting Based on Computer Vision

Yiqi Huang, Xi Qiao, and Jian Yang

College of Mechanical Engineering, Guangxi University, 530004, Nanning, China

Abstract. This article establishes a system of anti-bud injury in sugarcane cutting based on computer vision using MATLAB software as a development platform. The seedcane cutter box cuts twice each turn. Match the accurate shutter trigger time interval according to the rotation speed of the cutter , the transportation speed of sugarcane, in order to make the position of collected image of the target happen to be the next cutting site. Image processing is based on digital image processing tools of Matlab, and the image is processed from two aspects of the cane edge of the curve smoothness and internode surface color using Matlab by which sort out the image suitable for the identification of sugarcane internode, and then deal with the collected images by filtering denoising and binarization to determine whether the cutter cuts sugarcane internode. When the recognition threshold is 3500, the recognition accuracy of this system is 99%. After modifying the recognition threshold to 4700, its recognition accuracy is 100%.

Keywords: MATLAB, Injury bud, Cut sugarcane, Iterative threshold, Internode.

Introduction

China is one of the largest sugarcane-producing countries, as the third largest sugar-producing country in the world. Development of sugarcane industry directly affects our country's tens of millions of sugarcane farmers' livelihood and the development of sugar industry [1]. The majority of the sugarcane-producing states in the world have realized the mechanization of sugarcane cultivation to a certain extent, but there are deficiencies about it. The foreign planter has a good performance and the function tends to perfect, but it not yet equipped with professional device of anti-injured buds, and it's too expensive, with too complex institutions, so the foreign planter is difficult to popularize in sugarcane-producing areas of China. However, the domestic planter is more difficult to achieve the purpose of automatic anti-injury buds in the process of sugarcane species cut. At present, the field research at home and abroad still stays at an early exploratory stage. There are some similar researches, such as Lu Shang-ping extracted and recognized sugarcane internodes' characteristics based on machine vision [2]. Abroad, Moshashai K of Iran used grayscale image threshold segmentation method to do a preliminary study on sugarcane internodes' recognition [3].

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 251–259, 2013.

Our country has the large sugarcane cultivation area. However, in the sugarcane planting process, two main problems exist, one is the low degree of automation planting, and another is the high of bud injury rate in the automation planting process. With the continuous development of the machinery industry and the improvement of computer technology [4], this paper uses the technology of computer image processing to effectively process the cane image and analysis the characteristic parameters. This device consists of computer and camera. The study in the paper is to use MATLAB software to process the collected images and judgment the cutter cut the sugarcane internode. The biggest feature of image processing technology is to use a machine or computer instead of the human eye and brain to directly get the target image, and then process the target image and extract the effective and specific information to achieve the perception of object. Especially with the continuous development of image processing computer equipment, the high frequency of CPU, the high-capacity physical memory and image digitizing equipment are constantly updated to make the collection of the fast moving and multi-purpose image become a reality[5~7].

1 Device Structure

The cane was transported to the cutter box by the clamping mechanism of the roller, with the synchronization between Conveyor speed and cutting hob speed. A pair of hob in cutter box cutting twice each turn. According to the rotation speed of the cutter , the accurate shutter trigger time interval has been designed to match with it so that the collected image of target happens to be the next cutting site. Collected target image transmitted through the data cable to the computer, the images were processed and identified by MATLAB software, then to estimate whether they are sugarcanes. At last, computer feedback the information to the motor controller, the controller drive motor to move the clamping mechanism 15mm, to avoid cutter cutting sugarcane in order to reduce the rate of injury bud . Structure of the device shown in Figure 1:

1、Cutter box 2、Video camera 3、Clamping mechanism 4、Computer 5、Cane

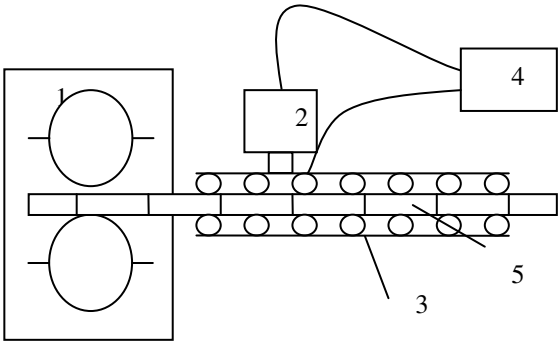


Fig. 1. Structure diagram

2 Image Acquisition

Use MVC360MF digital camera for image acquisition of the target .The speed of the sugar cane transported by clamping mechanism is the same with the line speed of hob which inside the cutter box ,They are 0.1m/s-0.3m/s in proper function. The camera image acquisition rate is 110 fps , so the cane has moved about 1mm - 3mm in the time of acquisition of an image. According to the agronomic requirements， the sugarcane stem segments for two or three buds. For example, Guangxi generally grown black fruit sugarcane (black sugar cane) , randomly selected ***q*** internodes of the ***P*** black canes , and measure the length of each internode, then calculate the average length of the sugarcane internode:

$$\overline{X} = \frac{\sum_{P=1}\sum_{q=1} X_{Pq}}{Pq} \qquad (2.1)$$

The X_{Pq} is the length of the n-th internode of the N-th cane

In this paper, taking the sugarcane stem segments for 150mm meets the agricultural demands. In order to make the collected sugarcane image consistent with the sugarcane cutting parts, the time of collection interval(fps) must be designed to match with the conveying speed of the cane. So that it correctly position the cutting parts and avoid redundant image information.The formula of calculate the frame interval as follows[9]:

$$F=（150-1000Vl）/1000Vl \qquad (2.2)$$

Of which: ***F*** - frame interval; ***V*** - cane conveyor speed (m / s); ***l*** = 0.009s for the time required to capture single image;

3 Image Process

3.1 Image Pretreatment

The study object in image processing is the internodes and stem of sugarcane. For example black cane, compare with the stem,there are significant changes in the shape and color of the sugarcane internode. Overall, its morphology is characterized by the different of the edge curve smoothness of the seedcane, the sugarcane internode color and the cicatricle shape in the internode.

In case to accurately extract the useful information of the image, it is necessary to pretreat the image. For a clear distinction between the image of sugarcane internode and sugarcane stem, and show obvious characteristics in each image, use the method of currently used to process the image,for example: Gray-scale adjustment、Filtering noise and Morphological processing. Gray-scale adjustment is the use of contrast enhancement method to enhance the contrast of the various parts of the image. Filtering noise is that use the method of neighborhood average to reduce the noise, which can effectively eliminate the Gaussian noise. Expression for:

$$g(x,y)=\frac{1}{H}\sum_{(i,j)\in S} f(x-i,y-j) \tag{3.1}$$

Of which: ***H*** is the total number of pixels contained within the neighborhood ***S***; ***S*** is a predetermined neighborhood (the neighborhood does not include point of (x,y)).

Neighborhood average ,cause image blur. So in this study, take the threshold neighborhood average. A (3×3) window move along the image (line by line, row by row), Find the average of all pixel gray values Unless pending pixel in the window. If the absolute value of pending pixel gray value minus the average value beyond a predetermined threshold, the pixel gray-scale use of average value instead of; otherwise, keep the pixel gray-scale invariant. Formula of the threshold neighborhood average:

$$g(x,y)=\begin{cases}\dfrac{1}{H}\displaystyle\sum_{(i,j)\in S} f(x-i,y-j) \\ \qquad f(x,y)\end{cases} \tag{3.2}$$

$$\left|f(x,y)-\frac{1}{H}\sum_{(i,j)\in S} f(x-i,y-j)\right|>Z$$

$$\left|f(x,y)-\frac{1}{H}\sum_{(i,j)\in S} f(x-i,y-j)\right|\le Z$$

Of which: ***Z*** is predetermined threshold value.

Image pretreatment steps are as follows:First, Gray-scale conversion of the original image.Then add Gaussian noise which mean is zero and variance is 0.2 into the grayscale image. Pretreatment of sugarcane internode image and sugarcane stem image show in the Figure 2 and the Figure 3. Internode and stems distinguish significantly, the gray-scale of Image changes tend to be gentler in the sugarcane stem pretreatment image.

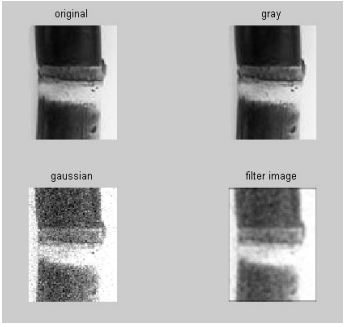


Fig. 2. Pretreatment of sugarcane internode image

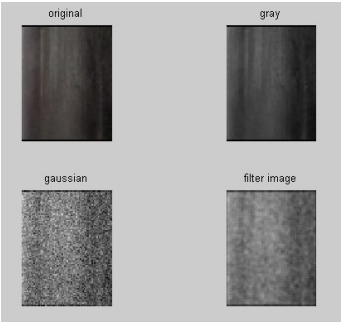


Fig. 3. Pretreatment of sugarcane stem image

3.2 Image Binarization

Binarization of the image is that set the grayscale of the point in the image into 0 or 1, in other words ,the whole image appear black and white[10-12]. Select the appropriate threshold through 256 brightness levels in the image into two kinds,the image was binarized still reflect the global and local features of image. The two-dimensional matrix elements of the binary image only by 0 and 1.The binary image can be seen as one special image which includes only black and white. Suppose the original image pixel size is $M \times N$, and $0 \le I(i,j) \le (L-1)$, formula of the threshold:

$$T = \left(\sum_{K=1}^{L-1} N_K \times K \right) \Big/ (M \times N) \tag{3.3}$$

Of which: N_k is the numbers of the pixel grayscale values is ***K***.

From that can get the binary image, and $B(i,j) = \begin{cases} 1 & I(i,j) > T \\ 0 & I(i,j) \le T \end{cases}$.

Black sugarcane internode and stem color have clear difference between black and white. Binarization is conducive to the computer identify the target. The key of Image binarization processing is to select the threshold, that direct impact on the final result.

In threshold processing , generally the most commonly and simple way is artificial selection. Adjust the ***T*** of the binarization threshold, by a large number tests of threshold adjustment, find that the target of image binarization is very clear when ***T*** is between 0.5 and 0.7.The internode and the stem can be very obvious distinction after binary the images of the internode and the stem.This way is based on subjective thinking of people, it usually chooses out a satisfactory threshold. But the system requires no man-made intervention to automatically select the threshold. So this system selects an iterative method to automatically select the appropriate threshold for each image.

According to the results of the image pretreatment,sum the value of the matrix elements of the gray image,and then averaging would calculate the iterative initial threshold:

$$T_0 = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} K_{ij}}{M \times N} \tag{3.4}$$

K_{ij} :The value of point (i,j) which is the point of the image pixel size $M \times N$ is ***K***.

Assume that the gray value of gray-scale image with ***L*** levels（0~L-1）, According to the Threshold ***T*** divided the image into two regions—R_1 and R_2, formulas of the average gray value(μ_1 and μ_2) of the regions R_1 and R_2:

$$\mu_1 = \frac{\sum_{i=0}^{T_i-1} in_i}{\sum_{i=0}^{T_i-1} n_i} \qquad \mu_2 = \frac{\sum_{i=T_i}^{L-1} in_i}{\sum_{i=T_i}^{L-1} n_i} \tag{3.5}$$

n_i :The number of the level i of the gray level appeared

After calculated μ_1 and μ_2, with the following formula to calculate the new threshold T_{i+1}:

$$T_{i+1} = \frac{1}{2}(u_1 + u_2) \tag{3.6}$$

Then put the results of the 3.6 -style into the 3.5 and 3.6 –cycle, if $|T_{i+1} - T_i| \le 0.1$, select T_{i+1} as the binarization threshold.

Selected results of the iteration as the binarization threshold, the result of the pretreatment image binaried as Figure 4 shown in. After processed, outline of black sugarcane internode has become more clearer, showing more prominent details of the image, the position of the sugarcane internode is also available to identify; The white part of the black sugarcane stem image is not existed, more to improve the accuracy of the stem and the internode can be correctly identified.

Fig. 4. Image of black sugarcane nodes and stem processed by binarization

3.3 Process Image Data and Judge the Result

For the computer to accurately determine the sugarcane internode and stem , need to inverted image and flip the gray values of the binary image. In a word, it transfers black to white and white to black. And then sum the values of all pixels, the formula as follow:

$$A = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} B[i, j] \tag{3.7}$$

Of which: $B[i, j]$ is the tag values of the point, ***m*** is total number of marker pixel ***I,*** ***n*** is total number of marker pixel ***j***.

Binary image corresponds to a value, experiment showed that the value of sugarcane internode is far less than the sugarcane stem. By compare the values, Computer can accurately identify the sugarcane internode and stem.Then it issues commands to the motor controller.

3.4 Experimental Results and Analysis

Take 100 samples tested,including 84 pictures of sugarcane stems ,16 pictures of sugarcane internode.Each picture sum A value in the image processing,statistical results as shown in figure 5

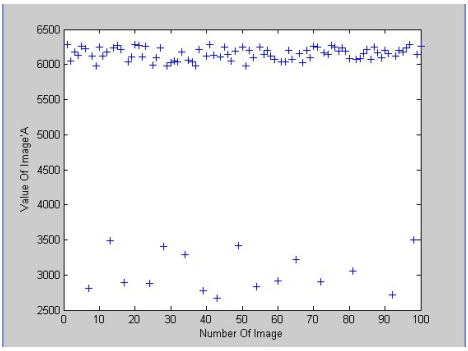


Fig. 5. A value of sugarcane internode charts

Seen from chart 5, a value of sugarcane stem is generally between 5900-6300,a value of sugarcane internode is generally between 3500-2600, specific statistical data as shown in Tabel 1.In experiment,select 3500 to identify the critical value,resulting in a miscarriage of justice.After change the critical to 4700,whether the node or the stem have large buffer,all samples of the recognition accuracy rate is 100%.

Table 1.

Statistical	Number	$\overline{A}$	A_{min}	A_{max}
sugarcane stem	84	6153.1	5978	6300
sugarcane internode	16	3154.3	2664	3505
the recognition accuracy rate		99%		

The system studied in this article was based on the develop platform of MATLAB, compared with the traditional image processing methods, MATLAB seldom need own programming ,the program directly through the function call,this save a lot of time.In addition,this related to all data and mathematical formulas by using MATLAB to call and calculate the results ,fast calculation ,which is the other image processing software don't have[13].

4 Conclusions

This study selected sugarcane node and stem as the research object, and using machine vision to track parts of cutter cutting sugarcane preliminary build a sugarcane cutting test rig based on computer vision.

Using the neighborhood average to make the target of image prominent, internodes and stems obviously distinguished, and the image gray-scale changes tend to be more gentle which is good for the next processing.

Using an iterative approach to obtain the threshold of the binarization procedure, which can eliminate the influence of human factors.

The recognition system was tested with a result of 100% recognition accuracy, and the algorithm execution time of a single image is 0.634s which provide a reference for domestic and foreign sugarcane machinery not equipped with anti-hurt bud cut-off device, and fill the blank of the image processing technology at home and abroad in the application of sugarcane stem cutting.

Using computer image processing technology to achieve fast and accurate judgment of the cutter when cutting sugarcane parts and reduce the rate of sugarcane internodes of injured buds, sugarcane production costs, and save the sugarcane speices, and improve labor productivity. The visual device solves the technical bottleneck of the mechanization of sugarcane injury prevention bud, besides, this device can also solve the problems of high costs of field test, and improve the sugarcane cutting automaticity.

There are definitely some issues needed to be studied further in this article: to develop the appropriate algorithms for different types of sugarcane; to improve the developed algorithm, such as improving the pretreatment algorithm and the threshold value selection; how can the existing cutter better cope with this system.

Acknowledgements. This work was financially supported by Guangxi Key Laboratory of Manufacturing System &Advanced Manufacturing Technology（Grant No. 10-046-07S04）.

References

[1] (April 21, 2011), http://zhidao.baidu.com/question/68598663.html?fr=ala0

[2] Lu, S., Wen, Y., Gewei, P.: Recognition and Features Extraction of Sugarcane Nodes Based on Machine Vision. Journal of Agricultural Machinery (10), 190–194 (2010)

[3] Moshashai, K., Almasi, M., Minaei, S., Borghei, A.M.: Identification of sugarcane nodes using image processing and machine vision technology. International Journal of Agricultural Research 3(5), 357–364 (2008)

[4] Chen, C., Xu, J.: Sugarcane Cultivation Science. Guangxi Science and Tecenology Press, Guangxi (2009)

[5] Ma, X.: Precision Metering Device performance image processing techniques to detect. Journal of Agricultural Machinery 7, 34–37 (2001)

[6] Hu, S.: Metering device performance testing technology based on computer vision. Jilin University, Changchu (2001)

[7] Hu, S.: Filing performance image processing detection metering device. Journal of Agricul Engineering 29(5), 56–59 (2002)

[8] (August 15, 2011), http://wenku.baidu.com/view/51ffae7202768e9951e738f1.html

[9] Li, W.: Seeding precision detection technology based on computer vision research. Doctor Dissertation of China Agricultural University 6 (2004)
[10] Qing, X.: MATLAB image processing and interface programming, pp. 227–254. Publishing House of electronic industry, Beijing (2009)
[11] Dong, Q.Y., Song, C.C., Ben, C.S., Chun, L.Q.: On-line measurement of deposit dimension in spray forming using image processing technology. Mater. Process.Technol. 172(2), 195–201 (2006)
[12] Dworkin, S.B., Nye, T.J.: Image Processing for machine vision measurement of hot formed parts. Mater. Process.Technol. 174(1), 1–6 (2006)
[13] Qiao, X., Huang, Y.: MATLAB in precision seeding machine monitoring device in Application Research. Chinese Agricultural Mechanization (2), 108–110 (2011)

A Multi-parameter Integrated Water Quality Sensors System

Mingli Li[1], Daoliang Li[1,*], Qisheng Ding[1,2], Ya Chen[1], and Chengfei Ge[2]

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, P.R. China
dliangl@cau.edu.cn
[2] Jiangsu Normal University, Xuzhou, Jiangsu, P.R. China

Abstract. Because of the rapid development of aquaculture in China at present, it is more and more urgent to apply hi-tech devices in aquaculture field to guarantee its efficiency. There are already many devices used for monitoring water quality have been developed. However, many of them are only in a step of academic research or too expensive to apply to practices. In this paper, a multi-parameter integrated water quality sensors system which has a practical value based on self-contained design is introduced and the hardware and software of this system are researched and discussed. A test conducted at Taihu Lake in Jiangsu province, China shows that this system can perform well. And this system has many features such as low cost, low consumption, multi-parameter and real-time data upload. The features and good performance above suggest the practical and potential application of the system in water monitoring. And this system still has much space to improve. It may become lighter, more integrated, and more portable in the future.

Keywords: water quality monitor, multi-parameter, aquaculture, low consumption.

1 Introduction

China has a flourishing aquaculture and the production of it ranks first in the world for many years which occupies above 70% of the world. And output of aquaculture in China is still increasing increasingly [1-2]. Thus aquaculture has made much contribution to development of Chinese agriculture. However, compared to other developed country, aquaculture in China has many weak points such as low efficiency and high consumption due to the low level of technology and management [3]. We still monitor water quality through experience and visual observation. Even though we also sample water for experimental analysis, lacking real-time monitoring and adjustment cause its low accuracy. In addition, experimental test costs much, has a long circle and collects limited data [4]. Water quality is a vital factor in the aquaculture. Short of monitoring of water quality parameters such as pH, Dissolved

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 260–270, 2013.

Oxygen (DO) and temperature can cause the low quality of water. What's more, the problems above may lead to waste of forage, residue of medicine and bacterial reproduction which have big terrible impact to aquaculture of our country. For example, the EU restricted the import of shrimps from China due to the medical residue a few years ago [5]. Therefore the high technology and smart management are important to raise production output and quality, improve productive efficiency, guarantee production safety and achieve sustainable development of aquaculture [6].

There are some researches in this field in China. A plan about a real-time multi-parameter test system is put forward by Zhang Libao from Qingdao University which can monitor four parameters of water quality as pH, temperature, DO(dissolved oxygen), and conductivity continuously in real time [7]. A real-time smart water quality monitoring system is researched by Ma, congguo from Jiangsu University, achieving smart controlling and information sharing in aquaculture. There are also some researches abroad [8]. A Multi-Sensor System is developed by O. Postolache from Portugal which can test turbidity, pH and temperature of water quality [9]. Losordo, Piedrahita and Ebeling from California researched an automated water quality data acquisition system based on self-contained microprocessor. This system can monitor and record weather data and pond environmental data [10]. However, researches and plans inland are often far from perfect and stay in a developing step. They cannot be applied in practice. On the other hand, productions abroad are often too expensive to afford by common people. Price of a set of foreign monitoring system can be higher than ￥100,000, thus they are not economical in large scale aquaculture.

Aiming at problems above, this paper develops a multi-parameter integrated water quality sensor system which can achieve multiple parameters collection, data storage and upload. 6 parameters of water quality are considered: pH, dissolved oxygen(DO), temperature, conductivity, NH3+ and water level which play important roles in water quality. For example, pH can influence the solubility and biological availability of water. Short of DO can generate toxic substances in water. Moreover, suitable temperature and other parameters also contribute to the good growth of aquatic life.

Each parameter of water quality can be monitored through this system of aquaculture to make it convenient for people to observe the states of water and adjust water quality in time to fit the requirements of aquaculture. And this system can achieve self-identification, self-correction and self-complement. It also has characteristics as low cost, low consumption, anti-interruption, multi-channel collection and easy operation.

2 Principle and Structure of System

2.1 Principle of System

Principle of multi-parameter integrated water quality sensors system is shown in Fig.1. The whole system can be divided into three layers: Application layer•process layer and perception layer. Perception layer face the monitoring objects. Application layer faces users. And they are connected by process layer.

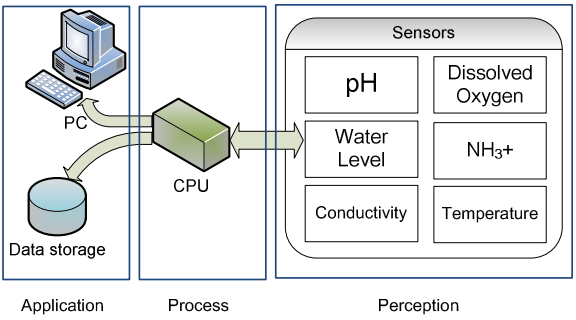


Fig. 1. Overall Principal of System

Perception layer is composed of 6 channels of water quality sensors which can collect 6 parameters of water quality. After the layer acquires these parameters and change them into electric signal through transducer, signals are sent to process layer.

Process layer can receive the analog and digital signals and analyze them through MCU msp430. And then MCU has AD conversion on the signals of pH, Dissolved Oxygen (DO), conductivity, NH3+ and use RS-485 to collect signals of temperature and water level. Finally, CPU revises these data by software and transmits them to application layer.

Application layer receives data from process layer. And then it will store these data in the flash and upload them to PC. In the application layer, users can configure and read parameters of sensors and read real-time data of every channel through software Unilog.

2.2 Structure of System

Multi-parameter integrated water quality sensors system is shown in Fig.2. The left part shows the appearance of the system and the right part shows the section of system. From these figures we can see that this system is a cylinder, with a hook on its top. PCB is placed at the middle of system. Batteries are put near the top of the system. Sensors and water pump connected with PCB are fixed at the lower place. Signals are collected through sensors from the bottom of the system.

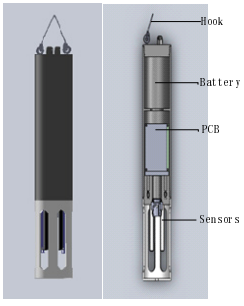


Fig. 2. Multi-parameter integrated water quality sensors system

3 Design of System Hardware

3.1 Hardware Structure

Fig.3 shows the hardware diagram of the multi-parameter integrated water quality sensors system. The hardware of system is mainly consist of CPU, sensors, PC, clock module, storage module, power module and water pump.

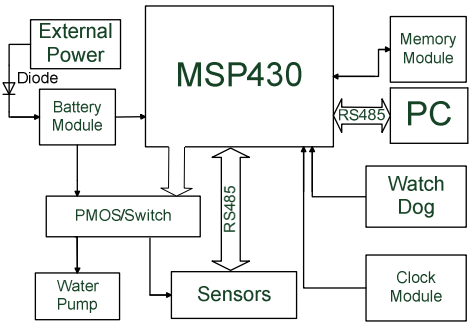


Fig. 3. Hardware diagram of system

3.2 Design of Circuit

CPU and Storage

Multi-parameter integrated water quality sensors system use MSP430F1612 made by TI as CPU. This is a mixed signal micro-controller which has 5 low-power modes. Proper clock mode can be chosen according to our real need in order to decrease the energy consumption. It also has rich resources inside, such as timer A, timer B, 12-bit ADC module, Watch dog, 8-bit general I/O. CPU connects the data storage part and parameter storage part through SPI and I2C respectively. Parameter storage part uses chip FM24CL64 with 64k ROM to store parameters of channels, serial ports, communication, dormancy and ADC. Data storage part use chip M23PE80 to store data which are processed by CPU. This chip is a page-erasable serial flash memory with byte-alterability.

Water Pump

Water pump can wash sensors every day under the control of software, thus to make sure the continuing of service and get rid of trivial human washing work. Circuit of water pump module is shown in Fig.4. It mainly consists of a PMOS and a BJT. Port SP7 connected with I/O of CPU is the enable port. When water pump is working, CPU will give a high level to this port and port POWER+ will get 9.6V to provide water pump with proper working voltage.

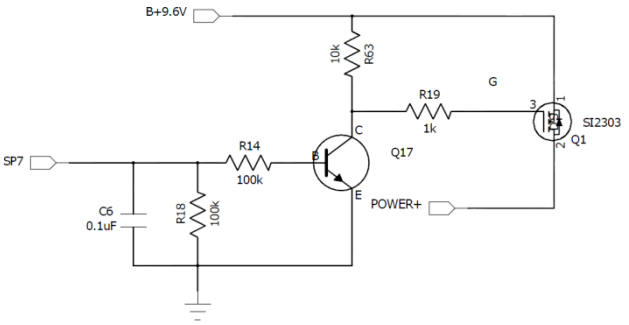
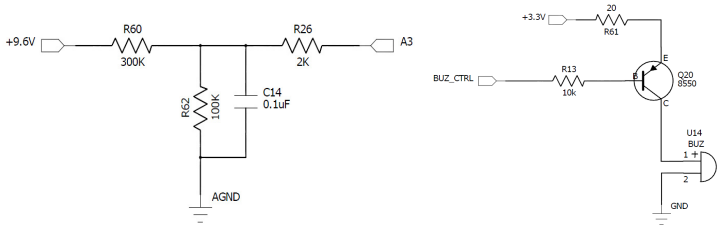


Fig. 4. Wash pump control module

Power Module

Nickel-Hydrogen Battery is used in this system. It has features of long service life, no pollution and rechargeable capability to meet the requirements of self-contained sensors. The rated voltage of this battery is 1.2V. 8 batteries are used in series, so the output of power module is 9.6V. In order to supply voltage for every module they need, this system uses an ultra-low dropout regulator LP2981. It has an output tolerance of 0.75% and is capable of delivering 100-mA continuous load current. The output of this chip is 3.3V which supply voltage to CPU, communication module, and storage module. Power module can also charge the battery automatically when the battery voltage is low to guarantee continuing work. In addition, power module provides protect measures to prevent accidents caused by miscellaneous interferences such as excessive voltage and current. In order to protect system from under voltage, 1/4 of battery voltage is input into CPU through port A3. And the voltage of this Pin is monitored every regular interval. Circuit of this voltage monitor module is shown in fig.5 (a). When the voltage is too low, buzzer will ring to give a warning signal. Structure of buzzer module is shown in figure 5(b). BUZ_CTRL is an enable port. CPU controls this module through a BJT.



(a) (b)

Fig. 5. Protection module

Communication Module

This module can realize the exchange of data among PC, CPU and sensors. Sensors and PC are connected to CPU with two serial ports. Through this module, PC can

transmit commands to CPU or sensors in order to read and change parameters of every channel and check real-time data. The acknowledge signals are also transmitted to CPU or PC via communication module.

Serial communication is easier and simpler than parallel communication in the microprocessor circuit which has little data throughput. Therefore communication in this system is achieved by chip MAX3485 which is a RS-485/RS-422 serial transceiver. Its transmission rate is high to 10Mbps. The interface between CPU and RS-485 is shown in figure 6. RS485_B- and RS485_A+ are connected with sensors and 430_URXD1 and 430_URXD1 are connected with CPU.

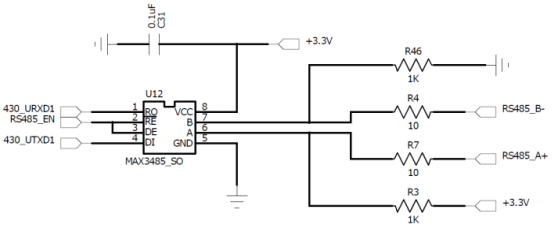


Fig. 6. Communication module

Moreover, there is also communication between CPU and storage module. The I²C bus protocol is used in communication between FM24CL64 and CPU. This is a serial extended bus, using two-wire system. Every node is linked to clock line SCL and data line SDA and every device has a unique address and independent electrical characteristic which can simplify the structure of circuit. It can realize the modularization and standardization design of circuit system [11]. SPI bus protocol is used in communication between chip M23PE80 and CPU. SPI is serial peripheral interface which is a four-wire system and full duplex. SPI provides programmable clock and have write conflict protection and bus contention protection [12-13].

4 Design of System Software

4.1 Program Flow

Software of multi-parameter integrated water quality sensors system combining with its hardware realizes the monitoring of 6 parameters of water quality. The main part of software includes initialization module, serial port communication module, ADC module, data storage module and dormancy module as shown in fig.7. Timer interruption and serial port interruption are used to control data collection, monitor and communication task. When there is no task on the go, software can change MCU into dormancy mode. At this time, CPU works in LPM (low-power mode) which is the embodiment of low-consumption.

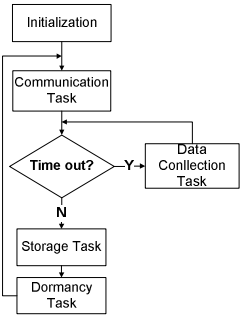


Fig. 7. Main flow of software

4.2 ADC Module

ADC module can process the analogy signals from sensors, transferring them into digital value. Timers are used to trigger the ADC task. There are 2 channels of signal need to be AD converted. Before the start of ADC operation of one channel, software will read the information of the channel such as power style, number of channel, address, state of channel and collection style. CPU repeats collection and processing of these 2 channels one by one. Data which have been processed are stored in a TEDS (Transducer Electronic Data Sheet). When all the data of sensors have been stored in the TEDS, the TEDS will be stored in a FLASH. And these data can be read by PC and analysis of water quality can be done according to these data. The flow of ADC module is shown in fig.8:

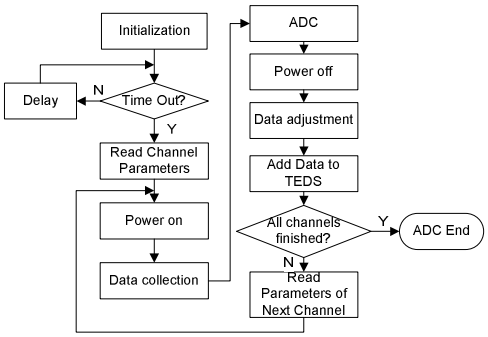


Fig. 8. Flow of ADC module

The step Data adjustment shown in the flow above is used to revise data to more accurate ones through a mathematical calculation. And the results are given by:

$$y = A(x + B) \tag{1}$$

Where A and B are coefficients. The default values of A and B are 1.0 and 0.0 respectively. And the real values of A and B vary according to specific applications.

4.3 Communication Module

Software of this module can code command frame according to target device address and purpose of instructions. And CRC is used as check code [14]. Later these commands will be sent to PC or sensors by serial ports.

When serial port receives command frames, software can calculate the checksum to judge whether the frame is correct. And then software will analyze command frame to extract address and purpose of the command frame. If the address is not the CPU itself, command frame will be sent to the sensor which has the corresponding address. Finally, command will be implemented in the correct device.

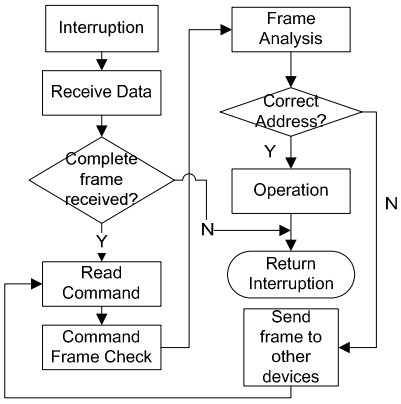


Fig. 9. Flow of communication module

4.4 Dormancy Module

In order to save energy, multi-parameter integrated water quality sensors system is in dormancy mode when there is no active task. After a circle of operation, CPU mode is changed into LPM. The consumption of CPU in LPM is a thousandth of normal mode. When timer is timed out, CPU will quit LPM and a new circle of operation starts.

5 Results and Discussion

5.1 Configuration of System Parameters

A multi-parameter integrated water quality sensors system was installed in Yixing, China to test the water quality of Taihu Lake. Sensors connected to main part of

system have been put in the water to start to acquire data of 6 water quality parameters. First of all, in order to ensure that data can be collected correctly, every module of software has many parameters to configure. Each device should have a unique address which guarantees that the communication information can be transmitted to right places. Second, time parameters should be configured to ensure that every operation should last proper minutes. Among all the parameters, important ones have been set as Tab.1 below before the beginning of monitoring. They are set by the software Unilog.

Tab.1 shows the basic time relative parameters in each module.

Table 1. Configuration of time parameters

Storage Parameters	
Data storage time interval	600000ms
Record start time	2011-10-10 0:00
Record end time	2011-10-17 0:00
Communication Parameters	
Baud rate of serial ports	9600B
Dormancy Application	
Dormancy Circle Time	60000ms
ADC Parameters	
ADC measurement time interval	600000ms
Battery monitoring time interval	600000ms
Frequency for ADC for each channel	50

5.2 Responses of Water Quality Sensors

After a one-week monitor, the multi-parameter integrated water quality sensors system has collected lots of information of water quality. Data is collected every 10 minutes, so there are 1008 records have been acquired for each water quality index in a week. Rely on these records, charts which show trends of 4 indexes have been plotted. And they are shown in Fig.10. From curves in these charts, we can analyze the states of water quality parameters as DO, pH, conductivity and NH3-H. Parameter as pH is more stable than the rest. And others as DO and conductivity are changing obviously based on the circle of the day. We can notice that DO concentration is lower than the normal standard from the curve of DO of the fifth day. Thus on these days we should take steps to enhance the content of DO. For example turning on the oxygen enhance machine.

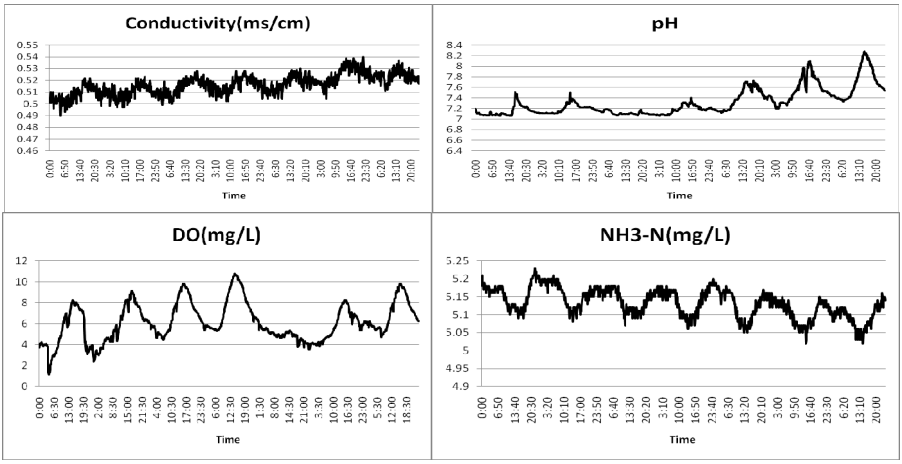


Fig. 10. Curves of water quality parameters over time

6 Conclusion

This paper introduced a procedure of development for multi-parameter integrated water quality sensors system. This system can monitor water quality to meet the demand of large scale aquaculture and it can be used in practical applications. Because this system can make the monitor operations easier, more effective, and more accurate, it saves much time and money of human in a long term. This method may replace the traditional monitor way which depends on human's observation and takes people much more time for monitoring water quality and cannot obtain precise data. As shown above, this system is able to obtain states of 6 parameters for water quality continuously in humans' absence. And the data of water quality is available on PC in real time. We can get much useful information for estimating water quality based on these responses of sensors. And the result proves that the system works well in practice.

Acknowledgements. This work was supported by 948 Project of China Agriculture Ministry (2010-Z13). And the programs "Development and Applications of sensor network applied to monitor bloom of blue-green algae in Taihu Lake" (2010ZX03006-006).

References

1. National Bureau of Statistics of China. China Statistical Yearbook (2011), http://www.stats.gov.cn/tjsj/ndsj/2011/indexch.htm
2. Yang, J., He, Z.: Discussion about current situation of aquiculture and its prospect in China. China Science & Technology Overview 5, 248 (2012)

3. Research group of the Ministry of Agriculture Bureau of fisheries aquaculture. Research of main aquaculture in China. China Fisheries 2,11–13 (2006)
4. Wang, P., Miao, L., Zhang, Y., Tang, T., Huang, J.: On-line water quality monitoring system integration technology in aquaculture. Fishery Modernization 6, 18–22 (2008)
5. Yu, Z., Yang, F.: Aquatic disease and pollution-free cultivation. Jiangxi Agriculture Technology 4, 13–15 (2003)
6. Li, D., Wang, J., Duan, Q., et al.: Research of integrated digital aquaculture system. China Science and Technology Achievements 2, 8–11 (2008)
7. Zhang, L.: Online multi-parameter water quality monitoring system. Qingdao University, Qingdao (2007)
8. Ma, C., Zhao, D., Qin, Y., Chen, Q., Liu, Z.: Intelligent Monitoring and control for aquiculture process based on field bus. Transactions of Chinese society in agricultural machinery 38(8), 113–115 (2007)
9. Postolache, O., Pereira, M., Ramos, H.: An internet and microcontroller-based remote operation multi-sensor system for water quality monitoring. In: IEEE Conference Publications, vol. 2, pp. 1532–1536 (2002)
10. Losordo, T.M., Piedrahita, R.H., Ebeling, J.M.: Automated water quality data acquisition system for use in aquaculture ponds. Aquaculture Engineering 7, 265–278 (1988)
11. Wang, C., Di, Q.: Control method of I^2C bus memory. Changchun Normal Collage Journal (Nature) 27(4), 30–32 (2008)
12. Long, A.: Design of FRAM based on SPI bus interface. Jishou University Journal (Nature) 29(6), 62–64 (2008)
13. Zhang, B., Liu, Y., Rong, J.: IP design and achievement of general SPI bus. China Integrated Circuit 20(7), 43–47 (2011)
14. Luo, Y.: Frame error correction in WLAN system using CRC. High Technology Letters 14(4), 12–15 (2004)

Research and Development of Decision Support System for Regional Agricultural Development Programming*

Jiangang Liu[1], Yongchang Wu[2], Tingting Tao[1], and Qingquan Chu[1,**]

[1] College of Agronomy and Biotechnology, China Agricultural University, 100193 Beijing, China
ljgwr0619@sina.com
[2] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, 100000 Beijing, China
cauchu@cau.edu.cn

Abstract Under the guidance of the agricultural system theory, operational research theory and decision-making support system theory, the regional agricultural development decision support system (RADDSS) was developed in this study ,in which different analysis method and models was integrated. By providing data, right models and analysis methods, RADDSS can assist decision-makers and administrators to solve half-structured and unstructured problems, improving level of management on agriculture. The agriculture in Xuchang has been analyzed using the system constructed in this study. The distribution of local agricultural production elements was rather reasonable, indicating that the district was suitable for agriculture development. The insufficient agricultural investment is the main factor that limits yield, and the more investment, the better yield. The level of agricultural modernization in Xuchang for nearly 20 years has been rising gradually, but still at the initial stage. The agricultural sustainable development potential index of Xuchang was 40.2, which was still at a low level. The gross output value of agriculture in Xuchang was predicted to undergo year-on-year growth. Fourteen multiple cropping patterns were designed and chose by RADDSS, and then the Effiency-oriented and grain-oriented program were proposed. The above-mentioned results indicate that the system can be applied to the analysis and decision-making of regional agriculture.

Keywords: region, agricultural development, evaluation on resources, programming, decision support system.

1 Introduction

The agriculture and rural economics of China have achieved outstanding success since the reform and opening up. The agriculture of China has entered upon a new historical stage. Its are main the feature emphasizes expression to be in the following

* Founded by National main basic research and development plan （2009CB118608）.
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 271–281, 2013.

respects[1] .firstly, Amid a lot of new changes taking place in agricultures, many new problems and contradictions arise, which are mainly reflected in the following aspects[2]. The alternating buyers' market and sellers' market that the primary products have been experiencing are gradually transitioning into the buyers' market. Also, the structural and regional oversupply of a few products occurs. Secondly, the sustainable development of agriculture in the new stage requires adjusting the agricultural structure. The utilization of agricultural resources is unreasonable. Some areas suffer severe shortage of agricultural resources while some other areas' resources are idle or wasted. Farmers' income does not increase along with the increase of grain yields, resulting in the low-speed rural economic development. The gap between the east and the west district is expanding, so that the imbalance in regional economy is aggravated [3] .Thirdly, the similarity in industrial structure has led to the situation of the blind competition for some regions. Such situation weakens the organic connections and the complementary advantages among those regions. In some regions, the phenomenon of "blind follower" and "herd mentality" occurs during the adjustment of agricultural structure.

Due to the disparity on natural and social conditions of different regions, There was significant distinct variation in the agriculture of China. Therefore, bases on the analysis of agricultural resources and regional variations, the studies manage to find out the advantages and disadvantages of the regions and to fully explore the regional characteristics, at last , a specialized and high-efficient regional layout should be put forward. Since the agricultural production system is a complex system based on the regional natural resources, the decision-making and administration on agricultural problems needs to take many factors into account [4]. By applying information technology for managing and analyzing information, decision support system can be helpful for administrators in different levels to make the right administrative decisions [5]. However, research and development in the agricultural information technology and the decision-making system in china is relatively late compared with the developed country. Because of lacking information and agricultural expert system support, mistakes in agricultural decision-making can hardly be avoided, which leads to a great loss of agriculture. Therefore, it is very urgent to build up all sorts of agricultural decision support systems. Based on the system theory, optimal theory and decision-making support system theory [6-9], by utilizing the information technology, the regional agricultural development decision support system (hereinafter referred as RADDSS) was researched and developed in this study.

2 The Main Functions of RADDSS

RADDSS is an intelligent decision support system [10] designed for agricultural administrative and companies at medium or small size, by providing users information and data required for decision-making, predicting related projects, offering a variety of alternative proposals, and making a selection and evaluation. By using RADDSS, users can evaluate resources, predict market and make a plan for production in a region. Users could also compare different plan, and make a judgment

over and again through the human - computer interface, so as to support and assist decision makers to solve decision-making problems or prepare plans for regional agricultural development. Specifically, RADDSS encompasses evaluation on resources and modernization level, and assessment in agricultural sustainable development, multi-cropping design, planting structural optimization and analysis in agricultural product market.

2.1 Evaluation on Regional Resources

The agricultural production potential of a region or a production unit depends on the natural resources, while the fulfillment degree of potential is partly determined by social and economic conditions. Therefore, the evaluation on resources is vital for decision-making. By using RADDSS, users can make a comprehensive evaluation and analysis on the natural resources, the social and economic conditions and the agricultural sustainable development level in a region.

2.1.1 Evaluation on Natural Agricultural Resources

By using RADDSS, users could analyze and evaluate natural resources of land, water, climate; base on the fundamental data on regional agricultural population, per capita farmland area, per capita forest area, per capita grassland area, per capita water area, per capita water resources; and calculate the index of regional natural resource in a region. RADDSS can reflect the trend of regional natural resources through dynamic analysis on changes among different regions as well.

2.1.2 The Evaluation and Analysis on Agricultural Modernization Level

There are some big differences between different regions on agriculture develops level in China. Then evaluating a region's agricultural development level correctly is vital. The evaluation on a region's agricultural modernization level is an important prerequisite for inspecting the development conditions of the agriculture and rural areas [11]. The established model for evaluating and analyzing the agricultural modernization level in RADDSS is used to measure a region's agricultural modernization level. The weight analysis method is largely adopted by RADDSS evaluation. In the process, indicator system, which consists of 1 overall index, 9 subject indices and 26 group indices, is divide into three levels, i.e. the overall indicator, the subject indicator and the group indicator. The analytic method is to evaluate the weight of subject index and each group index, while the weighted quadrate method and weighted summation method are adopted to calculate the comprehensive index of agricultural modernization..

2.1.3 The Analysis on Agricultural Sustainable Development Potential

The sustainable development of agriculture refers to the sustainable development of a compound system, which consists of such interacting sub-systems of nature, society and economy [12]. Therefore, a plenty of indices, such as population, resources, environment, economy and society in a certain region will be involved in the analysis of agricultural sustainable development. The indicator system for sustainable

development could be divided into descriptive indices and evaluative indices. Descriptive indices refer to those are difficult to be defined in quality and quantity and reflect the compound relation between society and nature as well as economy. While the evaluative indices refer to those are quantitative. In the model, the overall index is decomposed into three subject indices, i.e. the agricultural production sustainability, the ecological environmental sustainability and the agricultural economic sustainability. According to the accounts of the group indices and the subject indices determined by the expert review and analytic method, the weighted summation method could be adopted to figure out the overall index, i.e. the agricultural sustainable development potential index which could evaluate regional sustainable development level. In addition, the model can make a historic analysis on regional sustainable development level and a comparison analysis among different regions.

2.1.4 Comprehensive Evaluation on Resources

In this module, RADDSS build up a combination-barrel model to evaluate the resources of nature, society and economy in a region comprehensively. combination-barrel model is based the quantitative functional relationship on the amount of resources and crop productivity, By using the "barrel model" user can find restricted factors for agricultural development of a region. The factors combination barrel model can be used to evaluate the natural resources and crop demand quantitatively, to assess the restriction of limiting factors quantitatively and to determine the first or the second limiting factor, then making man-made investment targeting limiting factors, making full use of all kinds of natural resources and avoiding blind investment.

2.2 Prediction and Decision

Agricultural production system is a complex system based on the regional natural resources and economic characteristics, it consists of such subsystems of grain, economic crop, vegetable production and horticulture, fisheries, animal husbandry and other subsystems. Furthermore, it not only embodies biological and economic characteristics, but also involves a multitude of certainty and uncertainty factors that are either linear or nonlinear with interwoven effects and influences. Because of the complexity of influencing factors and the low accuracy in predication, the predicted result for macro agricultural production is imprecise. In this study, artificial neural network based on error back propagation was used, which can enhance the reliability of predicting effects. RADDSS attempts to predict the agricultural macro data [13]. BP (Back Propagation) neural network model is mainly adopted in the model. BP algorithm is widely populated and important for training neural network forward, which is a kind of teaching network performing a very powerful self-organizing and self-adaptive ability. Once the model is finished, we can input parameters, and select input and output factors, and predict and make decision after model training, and then user can obtain an effective result. In return, the result of prediction and decision-making can be used as an instruction for agricultural production and agricultural development programming.

2.3 Farming System Optimization and Design

Farming systems of crop production in China are quite complicated. The traditional multi-cropping design is a manual process based on the agricultural natural resources and economic conditions in a region, which will be then validated in production. In this study a new method of optimization and design was brought out based on crop ecological adaptability. The model can analyze regional climate and economic conditions, then find out the mainly suitable crops for combinations, and list the anniversary of the multi-cropping patterns and various parameters, so as to offer choices for decision makers. Besides, the model can analyzes and calculates the production potential of each crop, for example, photosynthetic production potential, light temperature productivity potential, climate production potential and comprehensive production potential.

Moreover, RADDSS can provide a comprehensive programming for the planting production in a region [14]. A multi-objective method is mainly adopted for programming and design. However, in the previous study, the linear programming method was applied. However, the practical production is multiple objectives, in which not only the production efficiency is considered, but also the ecological, economic and social benefits. As a matter of fact, the optimization results of a single objective functional linear programming tend to fail the requirements of practical problems. In this view, RADDSS uses the multiple objectives programming to solve regional agricultural programming problems which not only provides optimal solutions, but also offers alternative schemes for decision makers to choose.

2.4 The Analysis on Agricultural Product Market

Agricultural product market and price is always changeable because it is constrained by the law of supply and demand and the government policies. Besides, as one of the main factors that affect decision makers, agricultural product market and price affect agricultural products in the coming year. This study uses GM (1, n) method and time series model to predict the market and price of main agricultural products, the predicted results are presented in graphics or other visual styles, which directly reflect the fluctuation of price in agricultural products and provide evidence for programming and decision making.

3 The Structure of RADDSS

RADDSS focuses on supporting or assisting users to make a decision for half-structural and nonstructural problems rather than replacing users to realize the automatic decision-making. For this purpose, the basic functions of RADDSS are as follows: the first function is to collect, store and provide internal and external data related to agricultural management decision, the second function is to offer common system models which are equipped with the internal and external data for prediction, decision, evaluation and Analysis, the third function is to provide expert advice for the macro management, the fourth is to provide user-friendly interface that is easy for

users to operate and master. To support the operation of the above functional modules, the organizational structure of RADDSS consists of man-machine interface, database, model bank and method bank. In the process of system development and design, RADDSS adopts the most popular programming tool (Visual C++) that is object-oriented. Moreover, the entire RADDSS system is running on Windows95/98/2000/XP.

3.1 Man-Machine Interface

Being a link between the system and users, the man-machine interface exchanges and communicates information between RADDSS and users. It acts as a component of RADDSS logic structure and weaves together with other components in the program. The powerful interface function and data processing function of Visual C++ are given full play in the man-machine interface. Windows for explanation and description are set to make a brief instruction about its preselected functions. The input and output information is presented in such directly visual forms of windows, charts, words and others, and all inputs with are designed with powerful checking function and fault tolerance.

3.2 Database System

Database system is made of database and its management system. The database constructed by two main databases, several auxiliary databases and database management modules, that stores data and data management operations, includes a RADDSS system which involves basic data, an important intermediate data and the results. While the external information base is to stock the market price of crop seeds, chemical fertilizers and pesticides and main crops, agricultural meteorological database, as well as the data related to the future decision-making regions and its subordinate administrative region, such as population, labor force, land resources, basic statistics in national economy, areas and scales of agricultural production industries, costs, material consumptions and other data about agricultural management and decision-making. The internal one deposits labor index, farmland area, average yield per area, capital and other data related to internal users. Apart from the basic functions of data input, modification, deletion, searching, database management system also can extract data from main database to generate new database, realizing the data transferring between varieties of languages through the interface program and providing basic data for other subsystems according to the information demands of decision makers.

3.3 The System of Model Library and Method Library

The system of model library and method library is an important part of RADDSS system and a crucial tool to support decision-making. All models are included in model library, such as the models for analysis, evaluation, prediction and optimization. The main function of model library management system is to provide

model advice, generate models, modify models and solve problems using models. While method base management system is made up by method base and its management module. The traditional DSS method base is actually a model that refers to a solution seeking to a model. This study defines the method bank as one for decision-making and analysis, which breaks through the traditional concept about DSS method base. The system base is a kind of algorithm that collects goal programming, GM (1, n), multiple linear regression, and principal component for analysis and so forth [15].

The models and the methods used in the system are as follows: a) the model for agricultural factors evaluation, including barrel model, model for principal components analysis; b)the model for agricultural sustainable development evaluation; c) the model for agricultural modernization level analysis; d) the model for prediction, including neural network model, grey model, model for time sequences analysis ; e) the model for agricultural planning, including linear planning model and objective programming model; f) the model for resources evaluation, including weight analysis method, weighted integration method, and weighted sum method; g) the model for factors analysis, including regression model and model for difference analysis.

3.4 Knowledge-Base System

Knowledge-base system is consisted by knowledge-base, inference engine, self-learning machine and knowledge-base management module. The knowledge-base is to store some relevant knowledge, such as knowledge about professional agriculture and decision-making, expert's experience in making decision, scientific data. Actuality, the expression of agricultural knowledge is the key component.This study combines frame method and production rule as well as the relational database together, i.e. using frame method and production rule to form rule database, while using relational database to store knowledge elements and establish logic relations between them.

4 The Analysis of Agricultural Development in Xuchang Using RADDSS

In this paper, we analyze, evaluate and plan the agricultural production in Xuchang, Henan province with the established RADDSS.

Xuchang city is located in the central part of Henan. There is three counties, two cities, and a district with a total area of 4061.49 square kilometers, 304,600 hectares of cultivated land, and 0.084 hectares of arable land per capita area. In 2000, total income in agricultural production hits 3,737,360,000 yuan with 575,200 hectares sowing areas in whole year. Xuchang has a long history with an earlier agricultural development. The climate are suitable for agricultural production, include adequate light, a long frost-free period and four distinct seasons. Moreover, there are 314 farming days with daily average temperature above 0 ℃, 217 frost-free period days

and 631.1-736.0mm rainfall. Therefore, the favorable climate and the suitable conditions for crops make Henan become a major producing region for rice and economic crops in China. For example, it's rich in wheat, tobacco, cotton, corn, soybean and sweet potato.

Making an analysis for Xuchang with RADDSS. Firstly, agricultural resources in Xuchang were evaluated with barrel model- analysis model for resources combination, the graphic of "barrel model" is as follows:

According to the graphic of "barrel model" of Xuchang (Figure 1), we can see that the gap among "missing" height is small, which indicates the matching for agricultural production factors of Xuchang is reasonable and suitable for agricultural development. The values of elements of equivalent are shown in the table 1.With investment of only 2952 yuan/hm^2 that is below the average level of Henan province, capital is the first limiting factor in Xuchang. Since agricultural investment is the primary factor that restricts yields, we must raise agricultural investment to increase crop yields.

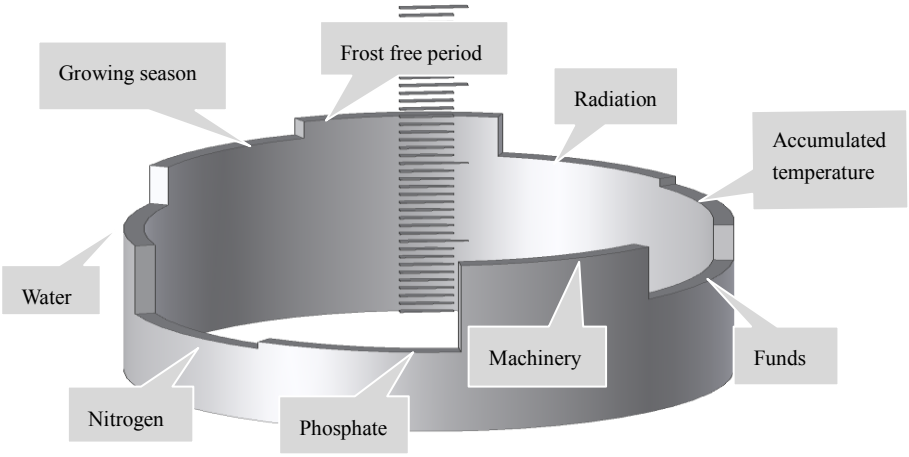


Fig. 1. Resource Evaluation in Xuchang- Cask model diagram

In addition, the rainfall in Xuchang is 700.7 mm/ year that can't meet the water requirement for the two or three crops a year; most rains fall in July, August and September; the effective irrigation can only cover approx. 70.50% of all farmlands in Xuchang; and the average irrigation power per farmland is only 810 watts/hm^2 , that is far below the demand. In this regard, water, which restricts agricultural production in Xuchang, has become the primary limiting one of natural factors. Only can we harvest in crops by increasing the coverage of farmland irrigation and the number of irrigational machines.

Table 1. The values of elements of equivalent in Xuchang

Elements	Radiation	Frost free period	Growing season	Nitrogen fertilizer	Phosphate fertilizer	Accumulated temperature	Water	Machinery	Funds
Value-s (KG)	2285.71	2500	2000	1453.33	2125	1800	1900	1013.64	925

Secondly, the agricultural modernization level was analyzed in Xuchang. This study analyzes the data in 1980, 1985, 1990 and 2009 in Xuchang, and figures out the agricultural modernization comprehensive index of the 4 historical stages. The results are in the following table.

The Agricultural modernization comprehensive index from 1980 to 2009 was calculated by the system, we know that the comprehensive index of agricultural modernization, which is still in its initial stage, is 45.9 % until 2009 in Xuchang. However, when we analyze the development process, the index has increased 29.6% in the past 29 years from 1980 to 2009, and the fastest development happened from 1990 to 1998 and reaches 13.2%.

To further analyze the 25 group indices, we can find that the progress in agricultural modernization is rapid in the 1990s in Xuchang, Henan province, which attributes to two reasons, one reason is the bigger rise in agricultural production level and a more reasonable structure in agricultural production. For instance, the production proportion of farming land is 1.7 times, agricultural labor productivity 4.14 times, meat of pigs and sheep 2.8 times in 1998 than in 1990, which changes the traditional agricultural mode. The second reason is the increase of farmers' income and the enhancement of consumption ability. From 1990 to 1998, the net income of farmers raised 4.18 times which promoted the growth of consumption and the improvement of consumption structure.

Thirdly, the level of agricultural sustainable development in Xuchang was evaluated. It's turned out the potential index of agricultural sustainable development in Xuchang is in low level with 40.2.The agricultural production sustainability, economic sustainability and ecological sustainability all stay below the middle level of China. This shows that in order to catch up with the developed district, Xuchang needs to accelerate the development of agriculture and rural economy while paying attention to the protection of both resources and environmen.

Fourthly, the agricultural production in Xuchang was predicted by using artificial neural network model. In this study, we uses *gross agricultural income* (Y) as output neurons, and planting areas of economic crops and food crops, rural electricity, chemical fertilizer dosages and several other factors as input neurons for network learning and prediction from 1990 to 2009 in Xuchang. The application of the system has been validated, which performed well in predicting gross agricultural income. Specifically, the learning rate is defined as 0.85, ε= 0.0001 and the data from 1990 to 2009 is taken as a set of learning samples. We predict the gross agricultural income with the trained neural network in 2013, 2014 and 2015,and the results indicate that the gross agricultural income will increase year by year.

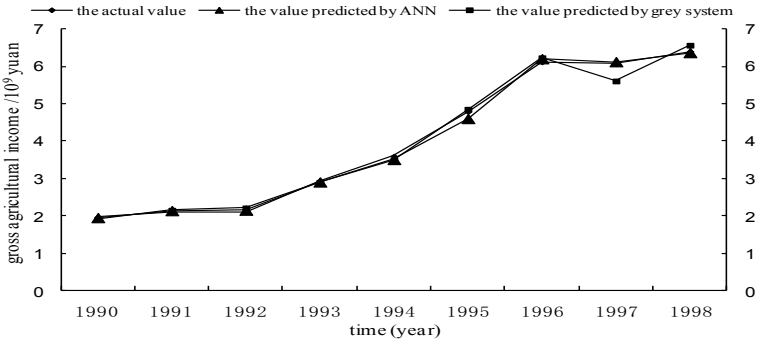


Fig. 2. Comparison of the actual value , the value predicted by ANN and the value predicted by grey system

Finally, the multi-cropping methods [16] were applied to design farming system in Xuchang, 14 kinds of suitable planting modes were screened out and their potential were analyzed respectively. Meanwhile, after optimizing the crops layout in Xuchang with the multi-objective method, two proposals based on analysis and calculation are put forward,which include grain-oriented project and benefits-oriented project. The grain-oriented project increases grain by 174,009.2 tons while the benefits-oriented project increases income by 439,762,000 yuan. The comparison analysis for the two projects obviously shows that, the grain-oriented project makes a weaker economic performance than benefits-oriented project as a result of ensuring grain output in Xuchang. Therefore, the low efficiency of agriculture in Xuchang cannot be solved by the planting sector alone, and it is unrealistic to increase farmer's income and accelerate the development of rural economy in Xuchang by relying on the nation to mark up the price of national grains. Seeking high-efficient sectors rather than planting sector is an imperative way out.

5 Conclusion

The body structure of RADDSS was build, and different analysis methods, models and theories were integrated into the system. The developed system can provide the users with all kinds of data information, model and analytical method. By providing a macroscopic insight on the local resources, society and economic development, as well as the stability and development direction of regional agricultural system, RADDSS offers proposals for the decision-making of administrators, operators, managers and production departments, and offers decision-making support for the administration and programming of agricultural administration, which improing the administrative and decision-making level on agricultural production, optimizing the allocation of resources, reducing the failures of agricultural macroscopic decision, promoting the agricultural development and raise the farmers' income.

References

[1] Zhao, J., Luo, Q., Deng, H., Yan, Y.: Opportunities and challenges of sustainable agricultural development in China. Philosophical Transactions of the Royal Society 363, 893–904 (2008)

[2] Wang, Y., Zhang, X., Tan, J., Han, Y.: Problem and Solutions of Soil and Fertilizers in Agricultural Sustainable Development. Chinese Agricultural Science Bulletin 24, 278–281 (2008)

[3] Zhou, X.-P., Gu, X.-K., Ding, N., Fan, P.: Change of Cultivated Land Protection Concept in Developed Region in China and Discussion on its Mechanism. China Land Science 23, 43–47 (2009)

[4] DeVries, F.P.: Systems Approaches For Agricultural Development. Springer, New York (1993)

[5] Power, D.J., Sharda, R.: Decision Support Systems. Springer, New York (2009)

[6] van Ittersuma, M.K., Ewerta, F., Heckeleib, T., Weryc, J., Alkan Olssond, J., Andersen, E., et al.: Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150–165 (2008)

[7] Fischer, G., Ermolieva, T., Ermoliev, Y., Sun, L.: Risk-adjusted approaches for planning sustainable agricultural development. Stochastic Environmental Research and Risk Assessment 23, 441–450 (2009)

[8] Roetter, R.P., Hoanh, C.T., Laborte, A.G., Van Keulen, H., Van Ittersum, M.K., Dreiser, C., Van Diepen, C.A., De Ridder, N., Van Laar, H.H.: Integration of Systems Network (SysNet) tools for regional land use scenario analysis in Asia. Environmental Modelling & Software 20, 291–307 (2005)

[9] Chen, W.: Decision Support System and Development. Tsinghua University Press, Beijing (1994)

[10] Zhu, X.: Intelligent and Object-oriented Decision-making Support System. Zhejiang University Press, Hangzhou (1992)

[11] Yue, W., Zhu, C.: World comparison of appraisement of Comprehensive Targets of the level of Agricultural Modernization in China. Research of Agricultural Modernization 12, 5–7 (1991)

[12] Zhu, X.: Sustainable Development of Agriculture and Rural Modernization. Agricultural Modernization Research 3, 4–7 (1998)

[13] Zhang, Z.: Studies on agriculture production using artificial neural networks based on crops culture complete collected information complex concert control theory. Ph.D. Thesis, China Agricultural University (1999)

[14] Li, B.: Structural Model for Optimizing in Agricultural Production System and Application. Agricultural Systematic Science and Comprehensive Research 4, 25–27 (1999)

[15] Pei, X.: Multivariate Statistical Analysis and Application. Beijing Agricultural University Press, Beijing (1990)

[16] Wang, H.: Structural Conception and Design of Agricultural Resources Management and Optimized Expert System for Cropping Regulation. The Nation's Thesis Collection of Youth Symposium of Farming, pp. 9—14. Nanjing University Press, Nanjing (1992)

Designation of R&D on Pig Production Intelligent Monitoring and Early Warning*

Fantao Kong**, Liyuan Xin, Wen Yu, Jianzhai Wu, and Yongen Zhang

Agricultural Information Institute of Chinese Academy of Agricultural Sciences/Key Laboratory of Agri-information Service Technology, Ministry of Agriculture, P.R. China, Beijing 100081
kft606@163.com

Abstract. This paper analyzed live pig production, consumption and price, and proposed the monitor system and the early warning mechanism. It is indicated that the study has a theoretical and practical significance in this area in China. In the view of the unusual fluctuation in the current live pig and its product prices, we carried out more works in live pig risk factors, fluctuation, detection methods, model building, as well as R & D relevant circumstances. In order to facilitate future studies in this field and to provide a useful exploration, we additionally proposed the implementation of the program of research and design at the last part of the paper.

Keywords: live pig prosperity, price fluctuation, forecasting, price movement.

1 Background

Pork is a main source of animal-derived food for Chinese urban and rural residents. In 2010, pigs in the stock totaled 46.46 million in China, compared with the years 2000 and 1990, increased 11.59% and 28.20% respectively, and 1.52 times that of 1980. Slaughter capacity was 66.69 millions, rose 28.58% and 115.18% respectively compared with 2000 and 1990, and was 3.36 times that of 1980. Pork production increased to 5.07 million tones, 27.87% and 122.31% up from 2000 and 1990 respectively, and 4.47 times that of the year 1980. Though the livestock produce structure had been adjusted, and the rate of pork production in the meat decreased step by step, it still accounts for 63.98%. The pork possession per capita is 37.82 kg in

* This paper is supported by (1) the 12th 5-year plan National Key Technology R&D Program " Agricultural Production and Market Circulation Matching Management and Research and Demonstration of Information Service Key Technology"（2012BAH20B04）; and (2) Basic R&D Research Fund for National Non-profit Research Institutes "Construction of Agricultural Information Monitoring and Early-warning Technology and System"（2012ZL017）.

** KONG Fan-tao, born in 1968, Male, PhD, his interests research include Agricultural Information Analysis. XIN Li-yuan, Master; YU Wen, PhD s; WU Jian-zhai, PhD; ZHANG Yong-en, PhD.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 282–289, 2013.

2010, which is 3.29 times that of the year 1980, and 20.86% and 89.56% more than that in 2000 and 1990 respectively.

However, from January to November 2000, pork price in China tended to fluctuate continuously and significantly (Figure 1). The price was stable in the period of January 2000 to May 2003. But after then, the price went higher and fluctuated largely, especially after June 2006, the pork price increased from 10.58 yuan/kg to 26.08 yuan/kg in February 2008. During the 20 months, the price fluctuation range got 146.50%. After that, it quickly declined by 39.88% to 15.68 yuan/kg, and followed by 16.04 yuan/kg. In September 2011 the price reached 30.35 yuan/kg, continuously went up 86.61% in 15 months, The fluctuated pork price seemed like a "roller coaster", which made "live pig production stable development" difficult to be realized. So, it is in urgent need to make a live pig price control plan.

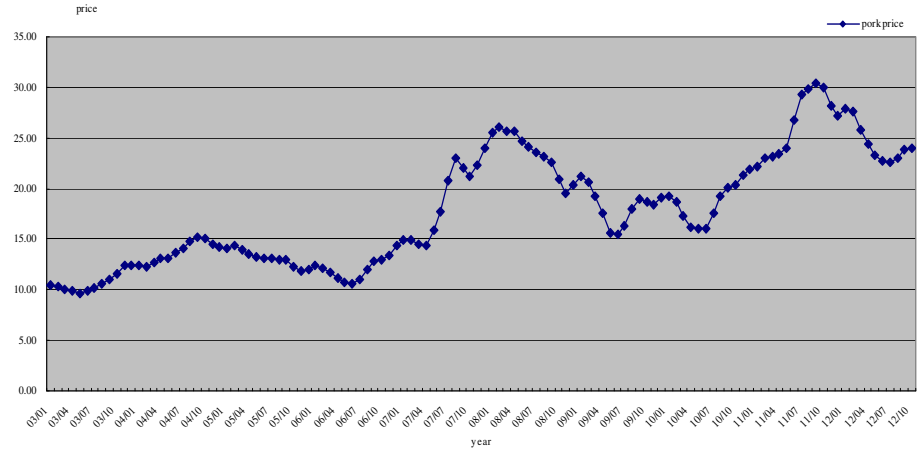


Fig. 1. Trend of Pork Price(2003-2012)yuan/kg

The price of pork has been an important livelihood issue, and concerned by the society. For example, in March 2011, during NPC (the National People's Congress) and CPPCC (the Chinese People's Political Consultative Conference), "live pig" became a hot topic of representatives and committee members. Continuous fluctuations of pork price had been a very important factor to push CPI (the consumer price index) fluctuations1. In July 2011, Premier Wen Jiabao asked about the price 3 times during 10 days, he said that " we should study the fundamental policy, avoid cyclical fluctuations out of pig production, and try to achieve the sustainable development of stable and healthy". At the same time, the Ministry of Agriculture unveiled a number of initiatives to encourage pig production, to guarantee market supply and stabilize market prices. Therefore, facing the abnormal fluctuations of live pigs and their products prices, we should do more research on monitoring of risk warning mechanism, which is important theoretical and practical significance.

In recent 10 yr, the scholars at home and abroad have achieved an important progress in pig production and consumption, price volatility, monitoring and early warning, and etc. From The Big Crisis in the 1930s', the economists have been

seeking theoretical explanations for the agricultural price fluctuations. For pig production and the pork price volatility, Kaldor proposed the cobweb model. Arthur A. Harlow determined the cobweb model as the applicability of the theoretical framework to explain the pig cycle when he analyzed the relationship between live pig production and price. Bollerslev proposed GARCH model. Anthony N. Rezitis applied the standard GARCH model, the asymmetric GARCH model and nonlinear GARCH models to study the pork expect and price volatility. Li Bing-long explained the reasons for pork price volatility in China. Xin Yin, Tan Xiang-yong built a causing model of pig price fluctuations. Peng Tao studied the "convergence-type spider web" trend of pig production and prices. Chen Yongfu et al. analyzed pig production and price cycles by continuous - short-term (PT) model and the information share (IS) model. Zhang Cun-gen, Wang Ji-min, Liu Gui-zhen also studied more on pig price volatility from different angles and by using different methods.

In forecasting and early-warning, Sims proposed VAR model to study the agricultural price transferring and fluctuations. B. Petersen raised the food security by the computerized monitor methods, so as to reveal the risk factors for each stage of production. Shao Renyuan used "Monte Carlo simulation model" to simulate pork price and forecast net income for the pig producers. M.L. Shiha used CBR-based analysis method to construct a price forecasting model. Boccaletti.S. used causal analysis method to analyze the Italian wholesale pork prices and retail prices. Ma Xiaobin built VAR model to forecast pig market price. Zhao Ruiying established a pig price risk warning model based on BP artificial neural network. Wang Mingli et al. studied the persistent effects of the pork market based on the variance ratio measure of the random shocks. The Commerce Department of China established a pig monitoring and early warning system in Sichuan Province, who therefore was regarded as the first one to develop this system. As for economic prosperity and its monitoring methodology, W.C. Mitchell discussed the possibility of the economic prosperity indicator for monitoring the macroeconomic theory. Economic prosperity monitoring method was first applied in the U.S. macroeconomic analysis, subsequently, the scope of application expanded to dozens of countries and regions. Its applications also expand to more related fields, i.e., the international economic prosperity, the regional economic prosperity, industry prosperity, economy and market prosperity. Abdollahzadeh adopted the synthetic index and proposed agricultural development index. In recent years, China has made a good progress in prosperity monitoring research and application. The study was launched in a social risk, the SME economy, business prosperity monitoring, the macroeconomic climate, regional economic prosperity. Up to now, the prosperity monitoring has been widely used and accepted. Around the prosperity research goals, we have engaged in a large number of studies on the Composite Index, the AC models, data base, neural network-based monitoring, and early warning, and etc.

Even though the important progress has been made on the research of pig production monitoring and early warning, many problems still remain to be resolved in this field. Firstly, integrative study is limited in pig production, pork prices and pig prosperity rules, but the laws of pig prosperity have not been developed. Secondly,

the pig prosperity risk factors are not formed. Thirdly, the research on price fluctuation is so less that the monitoring data have seldom been used and the monitoring methods have not been established. Actually the monitor research has not been carried out in China. Fourthly, early warning is mentioned more, while the research on establishing early warning indicators, dividing into early-warning warning-limit, building the model and its mechanism are limited. The early warning mechanism of the theoretical framework is not yet paid special attention.

2 Main Content of R&D

Focusing on exploration and analysis of the pig prosperity fluctuation rules, the optimization of the monitoring pig prosperity and the solution of the model of pig prosperity warning three key issues referred above, the goals of the R & D are designed as follows: (ⅰ) verify pigs risk and its risk factors; (ⅱ) propose the rule of the pig prosperity fluctuations; (ⅲ) construct the pig prosperity monitoring methodology system; (ⅳ) build the pig prosperity warning model; (ⅴ) form a theoretical framework of risk early warning mechanism for the pig prosperity monitoring. To achieve these goals, the main five aspects of R & D were studied by us as follows.

Pig Risk and Its Factors. We carried out the research on pig natural risks and economic risks, i.e., the basic properties, types, the degree of risk influencing, the probability of risk occurrence, risk loss, and conducted the pig prosperity risk policy scenario simulation. The pig risk factors were analyzed by using the approach of risk management. Based on a number of risk factors, the risk factors known as the "candidate risk factors", which may affect the consumption of pig production, were initially selected. Comparing with the pig fluctuation, assimilating to historical data of candidate risk factor, and scenario simulating to the impact of the candidate risk factor, analyzing the sensitivity and relevance of the simulation results, and we finally confirmed the risk factors of the live pig.

Pig Prosperity Fluctuating Rules. According to the pig industry development situation, combining the sectors of pig production, circulation and consumption, we built the pig sentiment indicator system using quantitative and qualitative analysis, and integrating other relevant indicators. By comparing and selecting of quantitative and qualitative methods, such as Analytic Hierarchy Process (AHP), theory of Gray Model GM (1, 1), Support Vector Machines (SVM), BP artificial neural network, Delphi Method, the link between the characteristics of pig production and consumption in different periods and backgrounds was formed. In addition, in order to determine the indicators in the basic period; we therefore put forward the scientific "pig prosperity index algorithm". In the future, we are going to analyze the historical data of live pig production to get live pig prosperity index data; and to apply live pig prosperity index algorithm to work out the live pig prosperity index in the past 30 years. Sequentially, we will explore the character, circle, account of increasing and summarize the law of live pig prosperity fluctuation.

Pig Prosperity Monitoring Method. By applying monitoring methods in actual warning and emergency, and the risk information on pig prosperity, the research was carried out on pig prosperity monitoring methods involving the aspects of survey range, place, quantity, methods and data transformation.

Pig Prosperity Warning Model. The safety interval and the risk interval of the pig industry were obtained using Value at Risk (VaR) to measure and assess the risk of the pig industry. Then by comparing the Pig prosperity index with the safety interval and the risk interval of pig risk, we determined the degree of pig prosperity. The warning results showed with the lights. The releasing mechanism was first established in China. The mechanism consists of the main body, subject, publishing content, distribution channels, means and frequency of publication.

Pigs Prosperity Monitoring and Risk Warning Mechanism Theory. We formed the A • D • Hall three-dimensional structure for pig prosperity monitoring and early warning with the methodology of systems engineering. Studying on the pig risks and risk factors, the pig prosperity fluctuation, the pig prosperity monitoring method, and the pig prosperity warning model, we applied the basic conceptions of the system science, risk management and early warning analysis, and finally formed the theoretical framework of the pigs prosperity monitoring and risk early warning mechanism.

3 The Program Implementation of R&D Designation

To collect and analyze the historical data of live pig industry in the past 30 year, we made the literature analysis and spot investigation. Especially, for the missing key data, in order to get the exact information, we also took the spot investigation. Our works supported the project strongly in data.

The Value method (Value at Risk of VaR) was adopted to calculate the risk in risk management. The method was based on extreme value theory and the quartile theory of VaR calculation. Under a certain confidence level (c) and within the time interval, the means of maximum loss expected in a normal market environment was defined as:

$$P_r\left\{r_i \le F_{i/i-1}^{-1}(P) / I_{i-1}\right\} = 1 - C$$

In order to test the VaR model, Kupiec statistical tests were applied to verify its validity. Through mathematical statistics and econometric methods, Analytic Hierarchy Process (AHP), Theory of gray system Model--- GM (1, 1), Support vector machine (SVM), BP artificial neural network, Delphi Method and other quantitative and qualitative methods were compared to determine the base period index, and therefore the scientific pig prosperity index algorithm was finally proposed. The prosperity indicators, the warning status were denoted with Liket's five-point scale method. The safety interval and risk interval of the swine industry are corresponded to no warning, and warning in the early warning system. As the warning system included light warning, middle warning, heavy warning and great warning, we adjusted and calibrated the related parameters to optimize early warning model.

Theoretical analysis methods were also used in our study, i.e., the methodology of system engineering was applied to establish A • D • Hall three-dimensional structure for pig prosperity monitoring and early warning. We proposed the comprehensive risk analysis and assessment of system to pigs, pig prosperity fluctuation, monitoring method, early warning techniques and methods, pig risk prevention and control, and the theory framework and risk early warning mechanism. The specific works involved were as follows. Firstly, we confirmed the issues to be researched. Based on China pig production and consumption practices, we carried out literature analysis, more spot investigations, theory analysis, and expert consult. Secondly, we collected and analyzed data and information of live pig in the past 30 years using risk management so as to find live pig risk and its risk factors. Thirdly, we built the pig prosperity index system to propose the pig prosperity index algorithm, and explore the law of live pig prosperity fluctuation. Fourthly, we established pig prosperity monitoring system including the contents of flowing such as the monitoring area, sites, indictors, methods, information transmission and data processing, and we also founded seven pig bases to monitor pig prosperity. In the end, we utilized the monitor data and calculated the live pig prosperity. With live pig risk consults, we constructed the live pig prosperity early-warning system as well as the pig warning releasing mechanisms and risk warning based on the three-dimensional structure of AD Hall. We put forward a theoretical framework for the pig prosperity monitoring. The project adopted the technology roadmap shown in Figure 2.

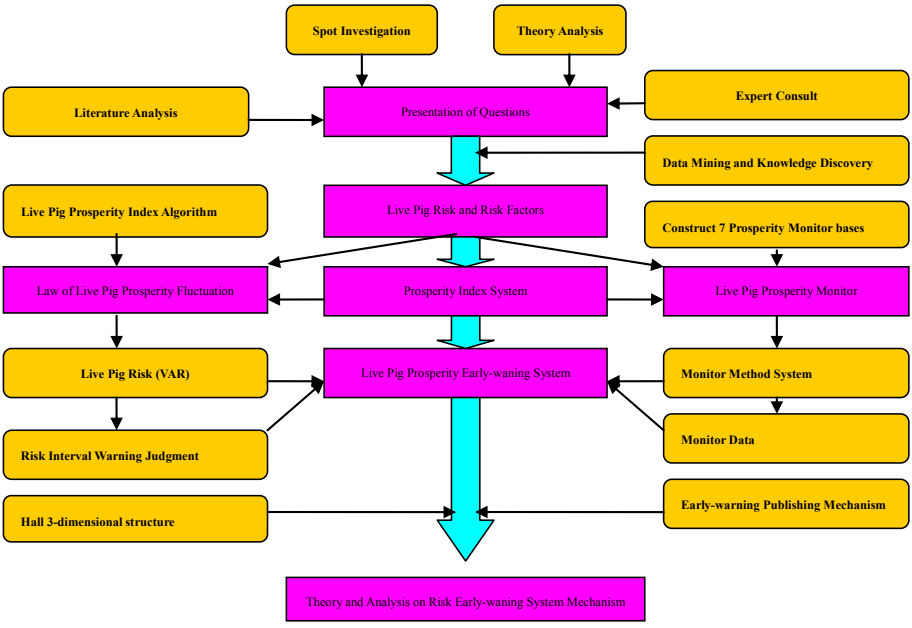


Fig. 2. Technology roadmap adopted in this program

To date, we have established 7 demonstration bases to monitor pig prosperity, such as the Fengtai District of Beijing, Huailai County and Dingxing County? in Hebei Province, and the Jia Taifeng Agricultural Science and Technology Demonstration Park in Suzhou of Jiangsu Province, Beijing, Tianjin, Jilin, Shandong, Guangdong, Hunan, Sichuan and other provinces and cities. Therefore, it is possible for us to take advantages of the demonstration bases on the agricultural market price for short-term forecasting. In times to com, we will set up bases to monitor pig prosperity, so as to provide data support and demonstration pilot and to collect and transfer monitor with the PDA data. Agricultural Information Institute of the Chinese Academy of Agricultural Sciences (AII-CAAS) and Tsinghua University have jointly developed a portable agricultural market information collection (handheld PDAs), and now it is at the stage of pilot consulting, planning and promoting. The project intends to use a handheld PDA to collect monitoring data in order to improve the intelligent, precision and timeliness of it. Integrating the related system and demonstrating the warning results, Prof. Xu Shiwei, the chief scientist of AII-CAAS, has established China's agricultural monitoring and early warning model system (the CAMES model), and achieved initial results, thereby provided a good platform for cooperative research and development on this study.

References

1. National Bureau of Statistics of China. China Statistical Yearbook. China Statistics Press, Beijing (2011)
2. Yu, H., Xu, L., Wang, H.: Monitor and Information Integration Technology. Tsinghua University Press, Beijing (2011)
3. PIG becomes one of hot topics (March 11, 2011), http://news.gdswine.com/yangzhuxinwen/zhuanjiapinglun/20110311/114372.html
4. Han, Y., Liu, X.: Caculations of China Pork Price's Fluctuation Affecting others price and CPI. Chinese Rural Economy (5), 12–20 (2011)
5. Xinhuanet. Premier Wen Jiabao asked pork price (July 14, 2011), http://finance.people.com.cn/nc/GB/15157579.html
6. Harlow, A.A.: The Hog Cycle and Cobweb Theorem. Journal of Farm Economics 42(4), 842–853 (1960)
7. Bollersl, E.V.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometircs (31), 307–327 (1986)
8. Li, B., He, Q.: Analysis on the Short-Term Fluctuations of Pork Prices and Its Reasons in China. Issues in Agricultural Economy (10), 18–22 (2007)
9. Chen, Y., Ma, G., Wu, B., Qian, X.: Study on China's Swine Price Discovery Forming Mechanism ——Empirical Analysis on the Relationship Between Interregional Price. Scientia Agricultura Sinica 44(15), 3279–3288 (2011)
10. Liu, C., Wang, J.: Outlook of Pig price Increasing and its Devlopment. Agricultural Outlook (9), 43–47 (2011)
11. Sims, C.A.: Macroeconomics and Reality. Economitrica 48(1), 1–48 (1980)
12. Petersen, B., Knura-Deszczka, S.: Computerised food safety monitoring in animal production. Livestock Production Science 76(3), 207–213 (2002)
13. Shao, R.: The Design and Evaluation of Price Risk Management Strategies in the U.S. Hog Industry. Graduate School of The Ohio State University, Columbus (2003)

14. Shiha, M.L., Huang, B.W.: Farm price prediction using case-based reasoning approach—A case of broiler industry in Taiwan. Computers and Electronics in Agriculture 66(1), 70–75 (2009)
15. Wang, M., Li, W.: Dissolution of Pig Price Circle and Measurment of Random Shock Effect. Journal of Agrotechnical Economics (12), 68–87 (2010)
16. Abdollahzadeh, G., Kalantari, K., et al.: Spatial Patterns of Agricultural Development: Application of the Composite Index Approach (A Case Study of Fars Province). Journal of Agricultural Science and Technology 14(1), 51–64 (2012)
17. Zhang, J., Zhang, X., Wang, J.: Business Cycle Tracer:A New Approach to Business Cycle Analysis. Journal of Systems Science and Mathematical Sciences 31(2), 241–250 (2011)
18. He, Y., Zhang, F., Huo, Y.-Q.: The Combining Early-warning of Booming Signals and Indicators Based on AC Model. Soft Science 25(3), 130–134 (2011)

Research on the Inconsistency Checking in Agricultural Knowledge Base

Nengfu Xie

Agricultural Information Institute, The Chinese Academy of Agricultural Sciences, Beijing, P.R. China
Key Laboratory of Digital Agricultural Early-Warning Technology, Ministry of Agriculture, P.R. China
nfxie@caas.net.cn

Abstract. The paper will propose a method to check the knowledge inconsistency in the agricultural knowledge Base, which is one the main measures to evaluate the agricultural knowledge base. This paper will pay attention to analyzing main types of knowledge inconsistencies and factors arousing it, and then discuss possible solving strategies which reduce agricultural knowledge inconsistency, so that agricultural knowledge is utilized correctly. In our practical application, it is very effective to find the inconsistency.

Keywords: Agricultural ontology, Inconsistency, Agricultural knowledge.

1 Introduction

People have realized that ontology is an important for knowledge reuse, knowledge share and modeling. In philosophy, ontology is a systematic explanation of existence, and is about the essence of description [1]. In cyberspace, lowercase letter "o" represents ontology, which means an entity, the consequence of analyzing and modeling by ontology. That is, to abstract a group of concepts and relationship between concepts from one field in the objective world [2]. Nowadays ontology is widely used in information systems, natural language understanding and knowledge systems. In agriculture, knowledge will be made use of effectively if we use ontology to organize agricultural knowledge. Agricultural knowledge is a production factor of high quality, which can improve the labor force and capital production in agriculture and accelerate agricultural informatization. What's more, agricultural knowledge is a kind of special knowledge, which lays the foundation of agricultural information application. Systems in agriculture education, language processing and expert systems all depends on agricultural knowledge.

On the other hand, in the process of organizing agricultural ontology based knowledge, differences will arise on concepts when people cooperate to construct knowledge base. And in the process of agricultural knowledge formalization, individuality error and editing error exist. Because of the large quantity of agricultural

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 290–296, 2013.

knowledge, text knowledge itself has inaccuracy, which leads to knowledge inconsistency. In philosophy, as time goes on, correctness of knowledge may be not right. It is easy to know the consequences of knowledge inconsistency. In military, knowledge inconsistency may cause wrong missiles launch and disability of aircrafts. In agriculture, farmers cannot grasp the right time of fertilization and reliable market information.

The paper will discuss the problems in the process of agricultural knowledge construction, and concludes the corresponding strategies of checking.

2 Definition of Inconsistency

In general, the knowledge consistency means some judgment accords with both history's judgments and the current facts. On the other hand the inconsistency means the contradiction between history's judgments and the current fact. From the aspects of ontology, the consistency means logic relations of terminology are consistent, while inconsistency means conflicts existing between some parts of ontology. For example, we define grain crops and cash crops as disjoint classes that have not the same instances. If the class wheat belongs to both grain crops and cash crops, the inconsistency will occur.

In the paper, agricultural ontology consistency includes the consistency of the definition of ontology and knowledge based on ontology, which means we can not get the conflict knowledge from knowledge base. Generally, whether knowledge base exists conflict knowledge depends on the following conditions:

1) The consistency of concept defining. That is to say, the formal definition contains the same means with informal one. Take the concept dogs as an example. If the formal definition of dogs goes with that of the concept cat, it brings inconsistency.

2) The consistency of concept extension. In terms of formal or non-formal concept definition, it can bring out conflict knowledge by concept explanation (include reasoning). For example, cats can catch mice, but we cannot say that mice can catch cats.

3) The consistency of axiom. The axiom system will not reason the conflict knowledge.

In the view of knowledge application, knowledge base can guide users to make the right decisions and ensure no confusion conclusion arising. In brief, the consistency is an important criterion to evaluate an ontology-based knowledge base. Knowledge inconsistency will lead to unreliable service, which threatens the knowledge correctness [4]. The paper will propose a method of ontology consistency checking.

Definition 1: Given knowledge base K, knowledge inconsistency problem is a 3 triple KI=(K,Y,Q), which satisfies that

- $Y=\{y_1,y_2,\ldots, y_n\}$ is Knowledge operation set.
- Q is a given knowledge query.

Definition 2: knowledge inconsistency problem KI=(K,Y,Q). If there exist knowledge conflict in K, it satisfies the following conditions:

- $\exists$ k, k11, k22, …, k1j $\in$ K, y11, y12, …, y1j $\in$ Y, $\sum_{1}^{j} y_{1l}(k_{1l}) \models k \wedge$ k→Q. the Symbol |= indicates "reason out" and → represent " can satisfy" .
- $\exists$ k, k11, k22, …, k1m $\in$ K, y21, y22, …, y2j $\in$ Y, $\sum_{1}^{j} y_{2l}(k_{2l}) \models \neg k \wedge \neg$k→Q. And then we judge out that the knowledge base has inconsistent knowledge.

From the above definition, the knowledge base has inconsistency if there are two pieces of contradictory knowledge. It is very import to find a mechanism or method to checking the inconsistency of knowledge base.

3 Agricultural Knowledge Inconsistency Problem Analysis

To explain inconsistency problems, we assume that agricultural knowledge adopts frame-based representation and organized by agricultural ontology. Firstly, some definitions is written as the following:

Definition 3: individual-of(i,C) means that i is a unit of category C. instance-of (i,C) means i is an example of category C.

Definition 4: IsRParent(A,B) means A is the direct father category of category B.

Definition 5: IsRSubclass(A,B) means A is the direct sub category of category B.

Definition 6: IsParent(A,B) means A is the father category of category B.

Definition 7: IsSubclass(A,B) means A is the sub category of category B.

Definition 8 : if C_1 is called the sub-class of C_2, we write if as subcategory(C_2, C_1). We can get the fact that instance-of(i,C_2)→instance-of(i,C_1) and for each i individual-of(i,C_2) → individual-of(i, C_1) is true for each i.

Definition 9: IsHParent(A,B)means that A is the sub category of category B, or B is the sub category of category A.

Definition 10: IsHasBaseParent(A,B) means that A and B is the sub category of a category, and IsHParent(A,B) is true.

3.1 Category Error

In agricultural ontology, the most important and basic semantic relationship is inheritance. Ontology constructed by this relationship can be taken as a description of a hierarchical model. It can also be taken as an information category system. Therefore, category error checking is the most basic requirement.

3.1.1 Loop Error

When classifying a concept, we usually take sub class as a partition of the concept. Though this method satisfies the integrity of concept definition, sub category error arises. We call it ownership contradiction of categories, shown as Fig. 1. We assume $A=\sum P_i$ represents category B has sets as IsRParen(Pi, B). To any Pk、Pj∈A, if IsHasBaseParent(Pk, P_j) is true, ownership contradiction of categories arises. For example, to IsRParent(grain crops, rear crops) and IsParent(cash crops, rear crops), if IsHasBaseParent(grain crops, cash crops) is true, contradiction arises.

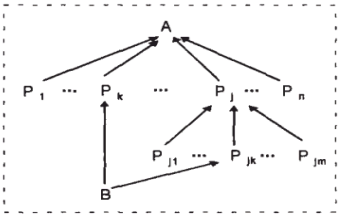


Fig. 1. The Class ownership contradiction

The checking algorithm is described that for each class B, finding $A=\sum Pi$ with IsRParen(Pi, B). If there exists Pk, Pj∈A, IsHasBaseParent(Pk, P_j) will be true. The method will find out conflict knowledge and mark it so that knowledge developers can delete error relations.

3.1.2 Classifying Category

When classifying a concept, we usually take sub class as a partition of the concept. Though this method satisfies the integrity of concept definition, sub category error arises. We call it ownership contradiction of categories, shown as Fig.2. We assume $A=\sum Pi$ represents category B has sets as IsRParen(Pi, B). To any Pk、Pj∈A, if IsHasBaseParent(Pk, P_j) is true, ownership contradiction of categories arises. For example, to IsRParent(grain crops, rear crops) and IsParent(cash crops, rear crops), if IsHasBaseParent(grain crops, cash crops) is true, contradiction arises.

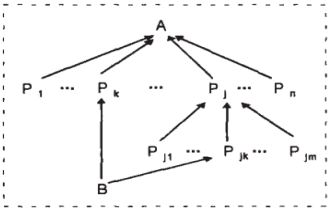


Fig. 2. The Class ownership contradiction

The checking algorithm is described that for each class B, finding $A=\sum Pi$ with IsRParen(Pi, B). If there exists Pk, Pj∈A, IsHasBaseParent(Pk, P_j) will be true.

The method will find out conflict knowledge and mark it so that knowledge developers can delete error relations.

3.2 Class Definition Error

Class definition of ontology should be correct and have clear semantic. Naturally, inconsistency of class definition may cause ambiguity and knowledge inconsistency of instances. Therefore before acquiring knowledge, we must ensure the consistency and integrity of class definition. Particularly, under cooperation of several engineers, different agricultural terminology of the same concept may arise, which causes redundancy of category definition. So we must define clearly to ensure one concept with one terminology in ontology, and defining synonyms is allowed. In another situation, the same terminology may represent several concepts. It is also caused by unclear class definition, and it is reflected in the relation, attribute and facet definition.

3.3 Axiom Inconsistency

Axiom is used to limit contacts between categories, attributes and relations, to ensure consistency of frame-based knowledge. Besides, axiom is used in reasoning, to supply the system with intellectualized judges. Axiom Inconsistency means axiom stands reasoning in different situations, while inconsistency is the opposite. Any axiom system should satisfy the consistency, otherwise geometric systems built by it will be conflict and such system is Worthless. Formally, we can take axiom system as a logic system with first-order logic. In that sense, axiom inconsistency means the axiom "yes=no" arises in the logic system. For example, such axiom P and Q are conflict, if

Axiom P: $\forall X \in$plantation crops), GreaterThan (MaturityDate(X), SeedingDate(X))).

Axiom Q: $\forall X \in$plantation crops), LessThan (MaturityDate (X), SeedingDate(X))).

Formally, we can take the axiom defined by ontology as first-order logic. Then axiom consistency checking can be translated to first-order logic operations.

3.4 Definition and Description Error

3.4.1 Contradiction of Case Ownership

If there exists an instance i and category A and B, which satisfies instance-of (i, A)$\wedge$instance-of(i,B) $\rightarrow$IsHparent(A,B), then instance inconsistency arises, shown as Fig. 3. For example, for IsParent(grain crops, rear crops) and IsParent(cash crops, rear crops), if instance-of(Ararat wheat, grain crops) and instance-of(Ararat wheat, cash crops) are true at the same time, then contradiction arises. Because IsHparent (grain crops, cash crops) is false.

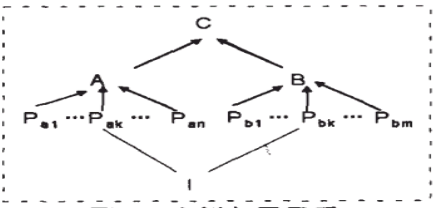


Fig. 3. The instance ownership contradiction

Checking Operations: for each case i, if there exists category A ≠B, which satisfies instance-of(i,A) ∧ instance-of(i,B), then check IsHparent(A,B). Assume IsHparent(A,B) is false, and knowledge developers should delete relation error. Besides, if there is IsParent(A,B), delete instance-of(i,B) too.

3.4.2 Instance Level Contradiction

We can use a tree to represent ontology category structure. If there is IsRParent(A,B), node B serves as son node of A. Thus all leaf nodes consist the biggest partition of the tree root. We call the leaf set as the ideal category of the tree root, and nodes as the leaf category, shown in Fig.4. Atypical cases arise when it is not the instance of ideal category, but of non-leaf nodes. For example, divide horticultural crops into three categories as fruit trees, vegetables and flowers. If we define lilies as an instance of the concept crop, then the relation as IsInstance(lily, P_i) in the categories.(P_i is a leaf category).

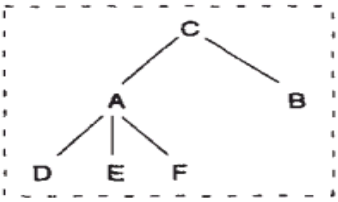


Fig. 4. The concept tree

The checking algorithm is described that this operation is relatively easy. For each case i, find category A which satisfies instance-of(i,A), and judge whether A belongs to the non-leaf category. If it is yes, then we mark the error.

3.4.3 Description Contradiction of Instances

Instance description should be clear. Xiong et al. [4] define a language in which it is possible to specify the inconsistency rule and the possibilities to resolve the inconsistencies. Egyed et al. [5] define such an approach for resolving inconsistencies in UML models. They take into account the syntactical constraints of the modeling language and they only consider the impact of one consistency rule at a time. Xie et al.[3] design a fused method to solve the inconsistency. In the fact, But in the process

of formalization, sometimes different names represent the same instance. Therefore we need to describe and analyze synonyms to judge whether they are the same instances. This method reduces redundancy and error caused by one name with different instances. Besides, in the process of realizing cases, we need to check attribute values and relation constraint error. For example, the instance wheat describe that it has main producing countries: China and Russia and the number of its main producing countries is less than 2". We can use axiom system to check such error.

4 Conclusions

From the point of inconsistency, this paper discusses the possible situations and causes of the agricultural knowledge inconsistency in the process of building agricultural knowledge. It also discusses the corresponding processing strategies to reduce knowledge inconsistency as much as possible. It is very important to check the knowledge inconsistency for knowledge base-based application and make full use of agricultural knowledge. In the future, more studies should be focused on the effective and quick algorithms to check inconsistency of agricultural knowledge bases and the method should be put into the practical application for find new method to solve the inconsistency in the knowledge base.

Acknowledgment. The work is supported by the special fund project for basic science research business Fee, AIIS "Tibet agricultural information personalized service system and demonstration" (No.2012-J-08) and The CAAS scientific and technological fund project "Research on 3G information terminal-based rural multimedia information service" (No.201219).

References

1. Gill, Y., Blythe, J.: PLANET: a shareable and reusable ontology for representing plans. In: Proceedings 17th International Conference on AI (2000)
2. Tamma, V.A.M., Bench-Capon, T.J.M.: A knowledge model to support inconsistency management when reasoning with shared knowledge. In: Proceedings of the IJCAI 2001 Workshop on Ontologies and Information Sharing (2001)
3. Xie, N.F., Wang, W.S., Yang, X.R., Jiang, L.H.: Rule-Based Agricultural Knowledge Fusion in Web Information Integration. Sensor Letters 10(2), 635–638 (2012)
4. Xiong, Y., Hu, Z., Zhao, H., Song, H., Takeichi, M., Mei, H.: Supporting automatic model inconsistency fixing. In: Van Vliet, H., Issarny, V. (eds.) Proc. ESCE/FSE 2009, pp. 315–324. ACM (2009)
5. Egyed, A., Letier, E., Finkelstein, A.: Generating and evaluating choices for fixing inconsistencies in UML design models. In: Proc. Int'l Conf. Automated Software Engineering, pp. 99–108 (2008)

Applications of Internet of Things in the Facility Agriculture

Linli Zhou, Liangtu Song, Chengjun Xie, and Jie Zhang

Institute of Intelligent Machines, Chinese Academy of Sciences
230031, Hefei China
Linlizhou@iim.ac.cn

Abstract. It is a trend to use information technology to lead the development of modern agriculture. The IntelliSense Internet of Things will be an important support for intensive, high-yield, high-quality, efficient, ecological security agriculture. In this paper, we give solutions and key technologies of facilities agriculture based on the Internet of Things technology. On this basis it designs and implements facility cultivation greenhouses. Practice has proved that the Internet of Things is the development of modern agriculture productivity. It has an important significance in raising the level of agricultural development, improving the overall efficiency of agriculture, promoting the upgrade of modern agricultural transformation.

Keywords: Internet of Things, agricultural facilities, wireless sensing, intelligent control.

1 Introduction

Facilities agriculture is a new agricultural industry which has a high degree of intensification. It is an important part of modern agriculture. In recent years, the technology of Internet of Things which continues to evolve and mature has injected new vitality into the development of agricultural facilities. IntelliSense chips, mobile embedded systems such as the Internet of Things technology in modern agriculture are gradually widened. Using wireless sensor networks can reduce the impact of human consumption and the farmland environment. Extensive use of automation, intelligent remote-controlled production equipment can obtain accurate crop and crop information. Through these, people who stay at home can monitor a variety of field information. This can achieve the scientific cultivation, scientific monitoring and production management and promote modern agriculture development pattern.

2 The Design of facility Agricultural Based on Internet of Things

The facility agricultural system based on the Internet of Things technology is divided into three levels: the perception layer, transport layer and application layer. Perception

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 297–303, 2013.

layer is mainly responsible for data-aware acquisition; the transport layer is mainly responsible for the perception of data transmission; application layer is mainly responsible for sensing data analysis, statistics, and early warning, automatic control and scientific decision-making. Fig.1 shows the overall architecture diagram of the facility agriculture.

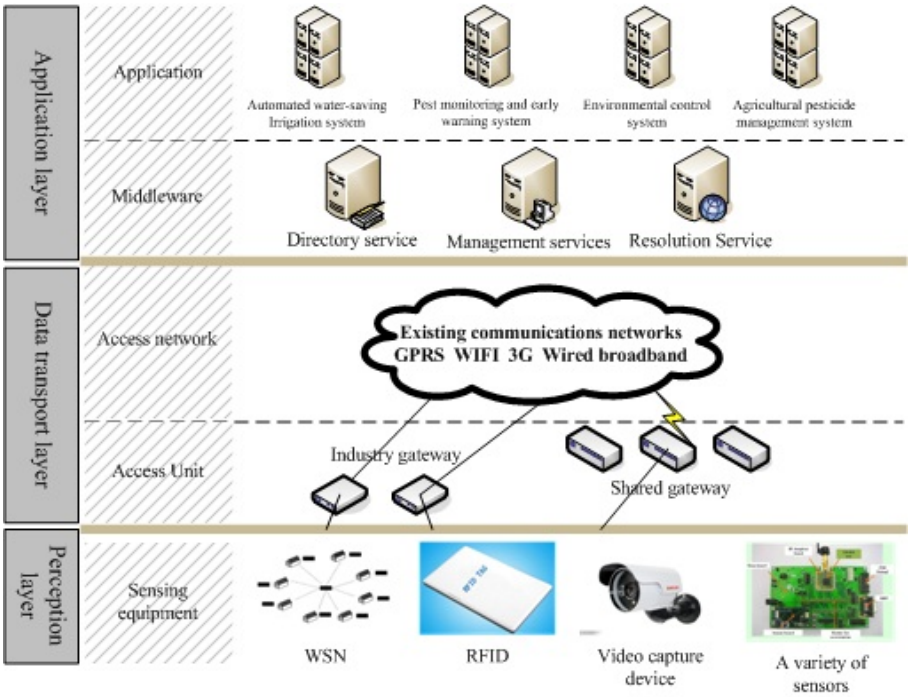


Fig. 1. The architecture diagram of facility agriculture

2.1 Data Perception and Acquisition

Information sensing technology of facility agricultural is stressed that the perception of intelligence and automation of information. Perception information should be processed intelligently and transported wirelessly. The Facility agricultural information has high demand to the perception and perception equipment. They must enable quick determination of the object, but also need to have a dynamic, continuous determination of the characteristics of the wireless sensing and wireless transmission. The sensing devices need to have the characteristics of micro, reliability, energy-saving, environmental adaptability, low cost and intelligent. Information-aware technology can be used for soil parameters, the continuous monitoring of plant nutrients, and the fast dynamic of the environmental parameters.

2.2 Wireless Sensor Networks

Wireless sensor networks include wireless sensor nodes, sink nodes, the routing node, the central base station, network data server and remote access node. Sensor node is responsible for the collection, a variety of soil and environmental parameters of the storage location. These parameters include air temperature, humidity, sunshine intensity, soil temperature, moisture, PH value. Sensor node sends data to the sink node through a variety of means of communication. Aggregation node is responsible for data collection, filtering and storage in wireless sensor clusters, and to communicate with the wireless routing nodes forward data timely.

2.3 Application of Decision-Making Platform

Application of decision-making platform includes the intelligent processing of information, cloud storage and application. Facility agriculture in the prenatal, delivery, postpartum industrial chain, relying on the cloud application service platform for the base production, the platform needs to be developed greenhouse environment control systems, fertilizer drug control management system, pest and disease monitoring and early warning systems, agricultural quality of the distribution process safety monitoring system. For government decision-making, technical guidance to farmers, public consumption and other aspects of the platform required the development and application of centralized display system, pests and diseases of the joint prevention and control command and decision system, remote expert guidance systems, facilities, vegetables, green resume system, etc..

3 The Key Technologies of Facility Agriculture

3.1 RFID Technology

Facility agricultural system has a large number of device management and remote control, you need to effectively identify and distinguish the equipment. Uniquely identify and read by the RFID technology, together with the GPS positioning method can control and manage the equipment location of Things, so as to achieve the purpose of the overall perception and intelligent control. There are more RFID products, but only fewer products for facility agricultural use.

3.2 Sensor Technology

The sensors are the eyes and ears of the information age; the sensor has a wide application in facility agriculture. The sensors can monitor the environmental conditions of the crop and the dynamic data which got through the facilities intervention. Commonly used sensors of agricultural facility include light sensors, temperature and humidity sensors, pressure-sensitive (fluid) sensor, CO2 sensor, the value of the sensor, as well as plant growth characteristics of sensors and other sensors.

3.3 Automation Technology

Facility agricultural needs automation technology for environmental regulation. According to the set of environmental conditions the system get control of greenhouse fertilization, irrigation, opening and closing doors and windows, lifting moderate light adjustments and others through the controlling system and execution system. The nature of the automation technology is mechatronics, controller and executive body is the core of the automation technology. Automation technology is the traditional industries; the development is already quite mature. Equipment intelligent control has been achieved. Development of automation technology which greatly improves labor productivity, and contributed to improve the system closed-loop control of the Internet of Things has been developing rapidly in a wide range of needs.

3.4 Wireless Data Transmission Technology

With the development of information technology, the wireless transmission technology has been developing rapidly. ZigBee technology occupies a dominant position. ZigBee technology is a unified standard, short-range wireless communications technology, with the characteristics of low power, low cost and versatility. WIFI has a high transmission rate, transmission distance and cover a wide range of features; it is widely used in data transmission in computer networks and mobile communications. GPRS and 3G digital communications technology has begun to be applied in the field of agricultural facility. It has a stable network; the terminal price is low and the maturity of the technology and other advantages.

3.5 Intelligent Information Processing Technology

Intelligent information processing system has been used in the field of agricultural facilities, real-time control of the parameters of the environmental temperature and humidity, light, fertilizer and liquid level. The development of cloud computing has injected new elements for facilities management in agriculture.

4 Implementation of the Facility Agriculture Project

On the basis of the research of facility agricultural structure and key technologies, sensing technology, infinite communication technology, computer network technology, agricultural resources, database technology, Internet of Things technology were used to build management technology platform for facility agriculture. It develops intelligent, accurate scientific production management for the majority of cooperatives, grower, facilities and agricultural enterprises and other users. Developed science intelligent precision operation scheme can make reasonable use of resources; improve product yield and efficiency; and improve the quality of agricultural products so as to enhance the market competitiveness of industry standards and product. Fig.2 shows the function and structure of facility agriculture.

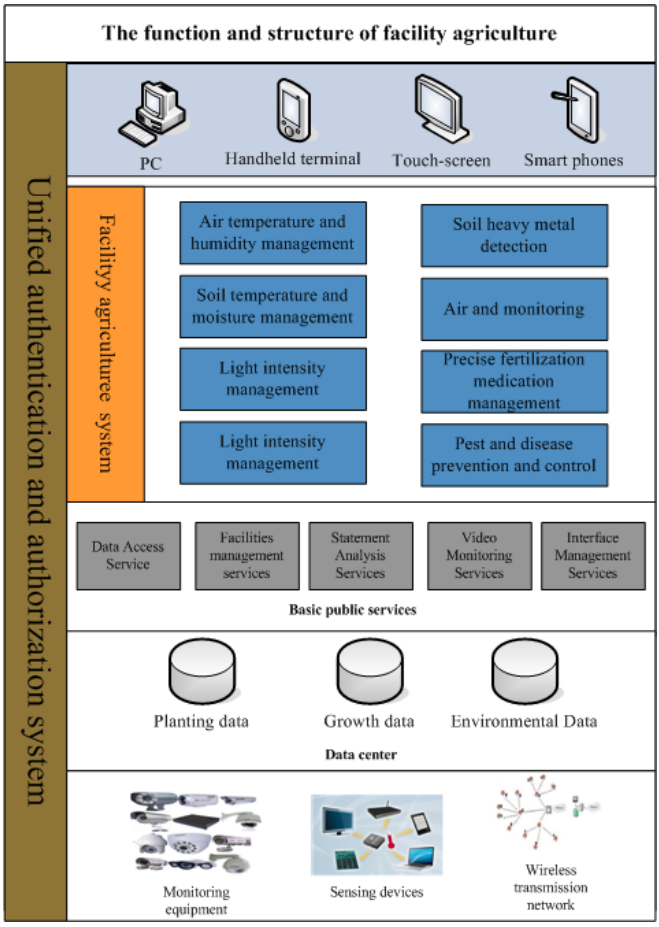


Fig. 2. The function and structure of facility agriculture

4.1 Greenhouse Environment Monitoring System

According to the facilities greenhouse cultivation of environmental conditions, we selected major environmental impact factor parameter detection. These parameters include meteorological parameters (light, temperature, humidity), environmental parameters (CO 2 and harmful gases such as ammonia, nitrous acid gas, CO, etc.), soil parameters (temperature, humidity and pH, EC, soil composition).The environmental information within the greenhouse which gets by the integration of multiple sensors is sent to the smart processing platform by wireless sensor networks. It provides data support for agricultural decision-making. Greenhouse intelligent monitoring system is composed by the host computer, acquisition controller, network adapter, sensors, outdoor weather stations, and video capture equipment, and other components.

A host computer can connect multiple Acquisition Controllers and Acquisition Controllers connect the various sensors and control equipment to form a separate

acquisition and control unit. Acquisition and control unit connected to the bus structure on the network communicate with the host computer independently. It controls the system remotely through the cable and the Internet way; regulates the temperature, light, ventilation, carbon dioxide supply, the supply and pH of the nutrient solution value, EC values.

Wireless sensor networks use hierarchical network structure, and a single greenhouse is a measurement of the wireless sensor network control area. Sensor nodes and control nodes in the network are self-organization; information is passed between nodes using multiple routes agreement. Data from all sensor nodes via a gateway to reach the control center, the feedback for greenhouse environment control information reaches the control node through the gateway. Control information controls valves, fans, temperature and other equipments to achieve the wireless monitoring and control of greenhouse environment. Fig.3 shows The greenhouses intelligent planting schematic.

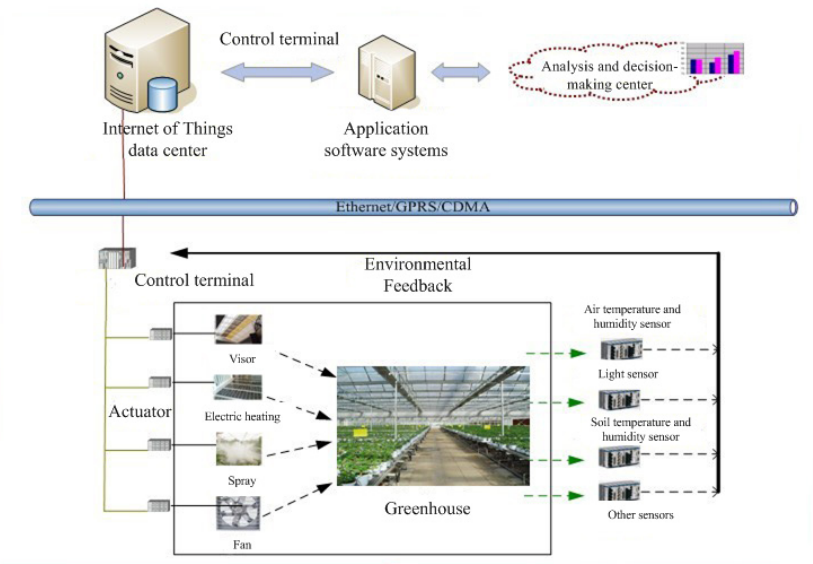


Fig. 3. The greenhouses intelligent planting schematic

4.2 Digital Management System for Facility Agriculture

Digital management technology analyzes intelligently the collected information; grasps all aspects of animal and plant growth status and makes decisions. The main contents are: monitoring, collection, processing and wireless transmission important environmental factors about vegetable growth; collecting data about facility vegetable growth; Modeling the growth simulation and optimization.

This system enables the query and distribution of agricultural information, the expert decision the Knowledge optimizes decision-making and analysis. It can achieve the efficient integration of information technology and network, and ultimate realize the agricultural precise digital control management.

5 Summary

It is a trend to use information technology to lead the development of modern agriculture. The IntelliSense Internet of Things will be an important support for intensive, high-yield, high-quality, efficient, ecological security agriculture. Agricultural facilities can be time-out order-season production, improve land productivity, resource utilization and labor productivity. It has an important meaning for overcoming the bottleneck of resources and the environment of the agricultural development, transforming agricultural development, and enhancing the competitiveness of agriculture.

As a comprehensive application of technology in different disciplines, facility agricultural based on Internet of Things technology integrates a variety of technologies such as sensor, automation, communications, computer and animal plant sciences. It can be predicted that facility agricultural will have a rapid development in the promotion of agricultural machinery, sensors, information and communications and cloud computing technologies. It will play a major role to improve the overall efficiency of agriculture, promote the upgrade of modern agricultural transformation.

Acknowledgments. This study is financially supported by the National Key Technology R&D Program of China (NO: 2012BAH20B02).

References

1. Collier, T.C., Kirschel, A., Taylor, C.E.: Acoustic localization of ant birds in a Mexican rainforest using a wireless sensor network. Journal of the Acoustical Society of America 128(1), 182–189 (2010)
2. Ignacio Huircan, J., Munoz, C., Young, H., et al.: ZigBee-based wireless sensor network localization for cattle monitoring in grazing fields. Computers and Electronics in Agriculture 74(2), 258–264 (2010)
3. Nadimi, E.S., Sogaard, H.T.: Observer Kalman filter identification and multiple-model adaptive estimation technique for classifying animal behaviour using wireless sensor networks. Computers and Electronics in Agriculture 68(1), 9–17 (2009)
4. Fukatsu, T., Hirafuji, M.: Field Monitoring Using Sensor Nodes with a Web Server. Journal of Robotics and Mechatronics 17(2), 164–172 (2005)
5. Nadimi, E.S., Sogaard, H.T., Bak, T., et al.: ZigBee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass. Computers and Electronics in Agriculture 61(2), 79–87 (2008)
6. Baggio, A.: Wireless Sensor Networks in Precision Agriculture. In: Proc. of the ACM Workshop on Real-World Wireless Sensor Networks, REAL WSN 2005 (2005)
7. Pierce, F.J., Elliott, T.V.: Regional and on-farm wireless sensor networks for agricultural systems in Eastern Washington. Computers and Electronics in Agriculture 61(1), 32–43 (2008)
8. Lea-Cox, J.D., Ristvey, A.G., Arguedas, F.R., et al.: Wireless sensor networks for real-time management of irrigation and nutrient applications in the greenhouse and nursery industry. Hortscience 43(4), 1103 (2008)
9. Mizunuma, M., Katoh, T., Hata, S.: Applying IT to Farm Fields A Wireless LAN. NTT Techinical Review 1(2), 6–60 (2003)
10. Green, O., Nadimi, E.S., Blanes-Vidal, V., et al.: Monitoring and modeling temperature variations inside silage stacks using novel wireless sensor networks. Computers and Electronics in Agriculture 69(2), 149–157 (2009)

Automatic Navigation Based on Navigation Map of Agricultural Machine*

Jianjun Zhou**, Xiu Wang***, Rui Zhang, Qingchun Feng, and Wei Ma

National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China

Abstract. This article mainly studies automatic navigation control methods for agricultural machinery. The navigation system takes RTK-DGPS and inertial measurement unit as positioning sensors. The preview control mode was researched for curve path tracking. The preview control model is to determine the aiming point and calculate the lateral deviation. Automatic turning was realized by fuzzy control method based on genetic algorithm. This control method can online adjust the parameters of fuzzy control using the genetic algorithm, in order to realize adaptive control. Thus the navigation control precision can be improved by the methods studied. The navigation control methods can meet the demand of agricultural machinery operating.

Keywords: automatic navigation, path tracking, genetic algorithm, self-adaptive fuzzy control.

1 Introduction

Automatic navigation of farm machinery is an essential technique to realize variable rate applications and collecting field information [1]. The potential benefits of automated agricultural tractors include increased productivity, increased application accuracy, and enhanced operation safety [2, 3].

Automatic guidance of agricultural tractors has been studied over the past several decades. Various guidance technologies, including mechanical guidance, machine-vision guidance, radio navigation, and ultrasonic guidance, have been investigated [4]. In recent years, high-accuracy Global Positioning System (GPS) receivers are widely used as guidance sensors [5]. Using an absolute positioning system, GPS-based guidance technology has the potential to achieve completely autonomous navigation [6,7]. Numerous fuzzy control applications on vehicles have been reported. Todo et al. developed a fuzzy controller that utilized the offset and the orientation errors to control the steering of a mobile robot and resulted in satisfactory trajectory tracking performance [8, 9].

* Supported by national natural science funds (31101088).

** JianJun Zhou, doctor, assistant researcher, mainly research direction is agricultural intelligent equipment.

*** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 304–311, 2013.

The basic idea of this article conveyed was researching fuzzy control method for automatic steering and methods for curve path tracking. In order to validate these methods, an agricultural robot for precision farming was developed, which are refitted from a tractor. Navigation control system was developed based on the robot using RTK-DGPS, digital compass and other sensors.

2 Materials and Methods

Field operation maps for autonomous navigation tractor are often created off-line using a GIS, and then be loaded into the navigation computer before the operation starting. Navigation map is a shape format point map. An autonomous field operation of a tractor was a sequence of instructions that both guide the tractor movement in the field and control synchronously the concurrent tractor operations. Each instruction contains a location in the field (x_i, y_i) and desirable control values, ($z1_i$, $z2_i$, … , zk_i), at that location. Control values include steering wheel angle, action of machine, and so on.

2.1 Preview Point Searching Arithmetic

An important content in the path tracking is searching dynamically the preview point in field navigation map. Fig. 1 shows the work flow of the preview point searching. Firstly the preview points are searched within the circle of radius equal to 1.8 meters, with the current navigation position as center on the navigation map. If the point's amount of searching result is zero, then the searching extent is extended. The preview points are searched one time again. If the search result is still zero, then the preview point is the one which id is equal to ID1 plus two. If the searching results points are equal to or greater than three, the preview point is the biggest of ID number within the searching range (ID number of point is increased in navigation map).

2.2 Computing Cross-Tracking Error Sign

In curve tracking the relationship of the current position of vehicle and path is complicated, computing the current vehicle lateral deviation error is difficult. This paper provides a vector multiplication method to determine the lateral error. As shown in Fig.2, the relationship between point p' and line P_0P can be denoted by the location vector P_0P' and vector P_0P. The relationship between vector P_0P' and vector P_0P can be described by these two vector cross multiplication. Vector cross product accords with the right hand regulation. Formula (1) and (2) showed the two vector cross multiplication process, the position relationship between points P 'and line P_0P can be easily obtained from the result of formula (2).

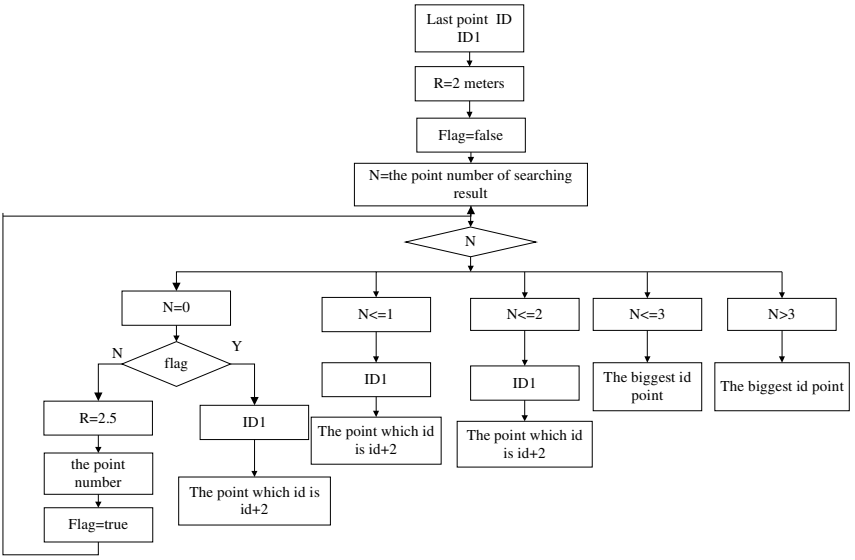


Fig. 1. Work flow of the preview point searching

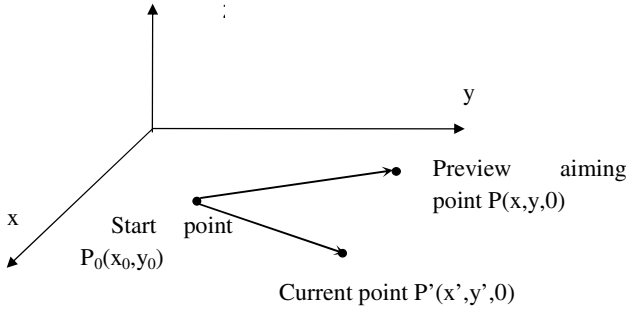


Fig. 2. XTE flag computing methods

$$\vec{u} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x'-x_0 & y'-y_0 & 0 \\ x-x_0 & y-y_0 & 0 \end{vmatrix} = \vec{k}\begin{vmatrix} x'-x_0 & y'-y_0 \\ x-x_0 & y-y_0 \end{vmatrix} \quad (1)$$

$$u=(x'-x_0)(y-y_0)-(y'-y_0)(x-x_0) \quad (2)$$

2.3 Dynamic Path Search Algorithm

In Fig.3, predefined route is $\{\ldots,p_{n-2},p_{n-1},p_n,p_{n+1},p_{n+2},\ldots\}$, and the current tractor position is point P_c. The predefined circle route points are searched within the circle. The circle's center is Pc. as center, radius is R. P_{n-1},P_n, P_{n+1} are the points in the search circle. P_n is the nearest point in predefined path from P_c recent points. EF direction is the direction of the tractor body, which can be got through the electronic compass acquisition; NM is the front wheel turning direction that can be obtained by angle sensor.

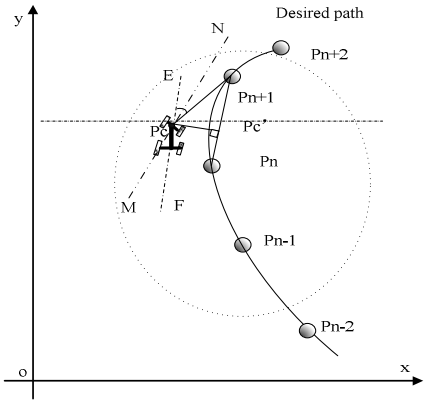


Fig. 3. Curve path tracking map

The start point coordinates of P_nP_{n+1} line is (x_1, y_1). The end point coordinates of P_nP_{n+1} line is (x_2, y_2). The equation of AB line is:

$$Ax + By + C = 0$$

in the equation:

$$A = y_2 - y_1$$

$$B = x_1 - x_2$$

$$C = x_2 y_1 - x_1 y_2$$

The current position of tractor is P (x_3, y_3).

The distance between point p and the AB line is:

$d = \dfrac{|Ax_3 + By_3 + C|}{\sqrt{A^2 + B^2}}$, d is the cross track error. $\angle EP_cP_{n+1}$ is the heading error.

Thus XTE and heading error can be as input for expected steering angle decision algorithm, according to the fuzzy control output can control the tractor automatic tracking curve path.

2.4 Self-adaptive Fuzzy Controller Based on Genetic Algorithm

The steering control of a vehicle is realized by using a step motor. Steering control is based on the error between the desired and the actual wheel angles. In this system turning control is realized by fuzzy control method. The design of the fuzzy steering controller was based on typical vehicle steering responses. Two inputs of fuzzy control are XTE and heading error. The output from the fuzzy controller was a real–valued steering rate command to drive the step motor.

System adopts fuzzy control arithmetic to control the turning of the front wheel. The cross track error E and heading error H was the input of fuzzy controller. The output was the expected angle U. The cross track error basic domain is [-24cm, 24cm]. Quantification factors are 0.5. The heading error basic domain was [-15, 15]. Quantification factor of this variable is also 0.5.

$$U = -<\alpha E + (1-\alpha)H>, \alpha \in (0,1) \tag{3}$$

By adjusting the correction factor α, the weighted degree of the lateral deviation error *E* and heading error can be changed, thus the fuzzy rule could be changed. The correct factor of the fuzzy control was online adjusted by the genetic algorithm, in order to reach the adaptive control. Control algorithms principle was as Fig. 4.

Genetic algorithm is robust, parallel search and group optimization. It is a kind of effective parameters optimization method. Its basic operation mainly includes encoding, decoding, selection, crossover and mutation, the design of genetic optimization algorithm is as follows.

Encoding and Decoding. Coding method using binary code which search ability is strong. Chromosome is length for 8 binary strings, α accounted for eight. The corresponding decoding way is as follows.

$$\text{a=binrep(a)} \tag{4}$$

binrep(a)---- integer expressed by a eight bit string of binary

Structure of the Fitness Function. The selection of fitness function directly influenced the convergence rate of the GA and where to find the optimal solution. Because of the system designed is small overshoot according to ITAE index department. In the agricultural vehicles to be automatic navigation, lateral deviations are the primary indicators to weight control effect. So the reciprocal of ITAE index of the lateral error are taken as the fitness function.

Genetic Operator. Genetic operator mainly includes two crossover and mutation genetic operator. Crossover operator adopts single point crossover modes. Mutation operator is logical opposition to every bit of chromosomes according to mutation probability. Float chart of genetic algorithms are as follows.

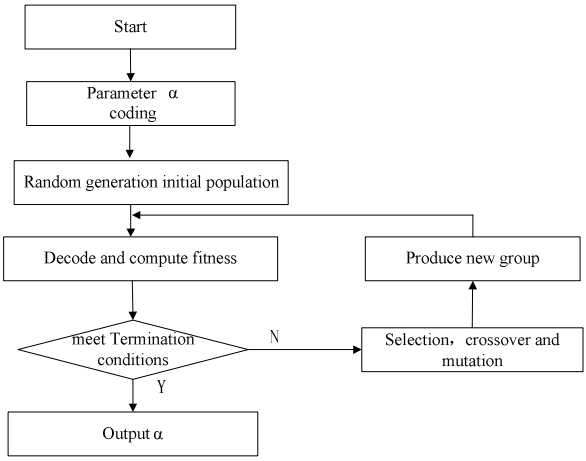


Fig. 4. Genetic algorithms flow chart

3 Experiments on the Refitted Tractor

The experimental vehicle is a refitted tractor. The tractor includes steering controller, posture information acquiring sensors and GPS receiver. The posture information acquiring sensors includes angle transducer and digital compass. The angle transducer transforms the angle mechanical rotation to the electrical signal. It can survey the change of the angle displacement. The angle transducer is equipped in the axis of front wheel of the tractor and used for measuring turning angle of the front wheel. The type of the angle transducer is WYT-AT-1.

The type of the digital compass is LP3300. There are three orthogonal magnetic field sensors and two axle inclination angle sensors in this digital compass. LP3300 outputs data through the RS-232 format and the configuration of RS232 port is "9600, n, 8, 1". Each frame outputs 20 bytes hexadecimal number.

3.1 Vehicle Navigation Control Methods System

In this research, RTK-GPS, digital compass and angle sensor are used for automatic navigation. GPS can provide the absolute position of the vehicle. At first, the path of tractor is preformed through GIS. Tractor can compare the appointed path with the current position to control the steering. Thus automatic navigation can be realized.

The type of RTK-GPS is Trimble 4700. The GPS data is received from serial port. The serial port can be easily set using MSComm active X control. After setting the parameter of serial port and opening the serial port, the GPS data can be received. According to the data format of NMEA 0183, useful GPS data can be extracted. The system uses GGA format.

3.2 Experiment

The experiment location is in the Xiaotangshan precision agricultural demonstration station, and the experiment was preceded in March 2012.

Firstly artificial driving tractors walked on a curve path. The tractor's path could be recorded through the data acquisition program, and was saved as a text file. File data included GPS data of tractor x, y, heading and id numbers. This text files could be input by Arcview software and generated a GIS format layer shape. This layer was the predefined curve path for tractor tracking.

It can be analyzed that is XTE is less than 12cm, the biggest error is 0.12cm, and the average error is 5cm .The lateral error is shown in Fig. 5.

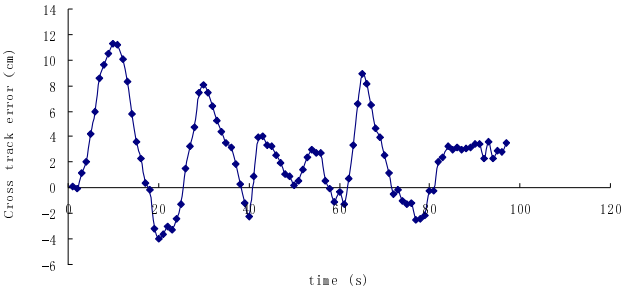


Fig. 5. The lateral deviation of curve path tracking

4 Conclusion

The main goal of this study was to develop a dynamic path search algorithm and fuzzy control method based on genetic algorithm capable of automatic navigation along the curved path on a field. The dynamic path search algorithm of this article is to determine the aiming point and calculate the lateral deviation. Self-adaptive fuzzy controller based on genetic algorithm could .The cross track error was less than 12cm at a travelling speed of 1m/s. The developed dynamic path search algorithm and Self-adaptive fuzzy controller based on genetic algorithm could achieve the design goal of guiding an autonomous agricultural tractor to track curve path.

References

1. Han, S., Zhang, Q., Noh, H., Shin, B.: A dynamic performance evaluation method for DGPS receivers under linear parallel-tracking applications. Transactions of the ASAE 47(1), 321–329 (2004)
2. Luo, X., Ou, Y., Zhao, Z., Zhao, X.: Research and development of intelligent flexible chassis for precision farming. Transaction of the CSAE 21, 83–85 (2005)

3. Zhou, J.-J., Wang, M.-H., Zhang, M.: Automatic navigation for agricultural vehicle based on fuzzy control. Transactions of the Chinese Society for Agricultural Machinery 40(4), 151–155 (2009)
4. Reid, J.F., Zhang, Q., Noguchi, N., Dickson, M.: Agricultural automatic guidance research in North America. Computers and Electronics in Agriculture 25, 155–167 (2000)
5. Yukumoto, O., Matsuo, Y., Noguch, N.: Robotization of agricultural vehicles (part 2)–description of the tilling robot. Journal of Agricultural Engineering Research 34, 107–114 (2000)
6. He, Q.: Automatic guidance system for tractor inter-row operations based on multiple sensors. China Agricultural University, Beijing (2007)
7. Zhang, Q., Qiu, H.: A Dynamic Path Search Algorithm for Tractor Automatic Navigation. Transactions of the ASAE 47(2), 639–646 (2003)
8. Reid, J.F., Zhang, Q., Noguchi, N., Dickson, M.: Agricultural automatic guidance research in North America. Computers and Electronics in Agriculture 25, 155–167 (2000)
9. Toda, M., Kitani, O., Okamoto, T., Torii, T.: Navigation method for a mobile robot via sonar–based crop row mapping and fuzzy logic control. Agri. Eng. Res. 72, 299–309 (1999)

Mathematical Study of the Effects of Temperature and Humidity on the Morphological Development of *Pleurotus Eryngii* Fruit Body

Juan Yang[1], Jingyin Zhao[1], Hailong Yu[2], Yunsheng Wang[1], Ruijuan Wang[2], and Lihua Tang[2]

[1] Technology & Engineering Research Center for Digital Agriculture, Shanghai Academy of Agricultural Sciences Shanghai 201423, P.R. China
yangjuan@saas.sh.cn
[2] Institute of Edible Fungi, Shanghai Academy of Agricultural Sciences Shanghai 201423, P.R. China

Abstract. The king oyster mushroom (*Pleurotus eryngii*) is one of the most popular mushrooms in Asia, Europe and North America. Its growth is influenced by environmental conditions, particularly air temperature and humidity. Using traditional methods of cultivation in plastic greenhouses, P. eryngii can be grown only in spring and autumn. However, mass cultivation using modern commercial methods is unrestricted by season, but precise environmental control is crucial, as the process is costly and a high energy consumption. Through experimentation and analysis of the effects of variations in air temperature and relative air humidity on morphological development, a mathematical model for these effects was developed. This model can be used to predict the optimal ranges of air temperature and humidity for various indices of morphological development of the fruiting body, and to guide commercial production. Using the model, standard values of morphological development, optimal environmental control ranges for each day of the fruiting body growth period were calculated. These values provide a reference suitable for application to actual production.

Keywords: Fruit body, morphology, air temperature, relative humidity, mathematical model, Pleurotus eryngii.

1 Introduction

Mushrooms are known for their nutritional and medicinal value, and the diversity of their bioactive components[1]. These organisms have long been valued as highly tasty and nutritious foods by many societies throughout the world. Among the edible mushrooms, the king oyster mushroom (*Pleurotus eryngii*) is one of the most popular in Asia, Europe and North America[2]. P. eryngii has several bioactive components, such as β-glucan[3] and ribonuclease[4] , which have been isolated from the fruiting bodies of edible mushrooms. However, the fresh fruiting bodies are mainly used as a food sources. In Japan, production increased from 60 t in 1995 to over 29,000 t in

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 312–323, 2013.

2003[5]. Even greater increases in production have occurred in China, where commercial production began in the late 1990s. Chinese growers produced an estimated 7300 t in 2001, and 114,100 t in 2003[6-7]. In the US, commercial production began in 2000 and has reached 85 t by 2004[8].

As a commodity, the ideal morphology of the *P. eryngii* fruit body is generally columniform. The cap is smaller than that of the natural state, and can even be smaller than the stipe. Air environment seems to be very important for induction of fruiting body formation. The appropriate air environment is not only crucial for production, but also influences the ideal morphology. When air temperature and humidity are lower than optimal, the fruiting body does not growth, and can even wilt and die. Conversely, when air temperature and humidity are higher than optimal, the fruiting body grows quickly, but a higher proportion of anamorphic fruiting body.

Most studies focusing on the relationship between fungal growth and air environment have been qualitative, because of the large range of the environmental control variables in conditions of traditional cultivation in plastic greenhouses. In modern methods of mass commercial cultivation of *P. eryngii*, air environment is monitored through air temperature and relative humidity probes placed in the cultivation room. Air temperature and humidity are then adjusted in real time through air conditioning, humidifiers and fans controlled by a dedicated computer. Environmental control is thus very precise. Accurate environmental control can reduce the overall costs through increased yield.

The objective of this research, therefore, was to analyze the effect of different air temperatures and relative humidity on the morphological growth of the *P. eryngii* fruiting body, through experimental methods. A mathematical model for the effects of air temperature and humidity on growth was then established. Using the model, the appropriate environmental ranges for morphological growth of *P. eryngii* in commercial production were obtained. Finally, the model could make useful predictions about air environment variables to guide environmental control in the commercial production of *P. eryngii*.

2 Materials and Methods

2.1 Experiments

In this study, the air climate control in the growing rooms was designed and manufactured by Patron AEM; temperature, humidity and CO_2 concentrations in the growing rooms could be controlled very precisely and efficiently. There were two treatments. The first was with air temperatures of 14°C, 15°C, 16°C, 17°C and 18°C. The second was with relative humidity of 89%, 91%, 93%, 95% and 97%. The intensity of illumination in the growing room was 500-1000Lx. The experiment was a 2 (supplement) × 5 (treatment) design with three replicates per treatment. Morphological growth indicators were measured daily (every 24 h), with 8 bottles for each experiment. Each value is the mean of 24 measured results (3 replicates × 8 bottles).

The morphological growth is characterized by three indicators: first, cap diameter, which reflects the size of the circular cap of fruiting body; second, stipe diameter, which is measured at the base and reflects the approximate size of the columniform stipe of the fruiting body; and third, stipe height, which is measured from the base of the stipe to the cap. Each day, two or three of the bigger fruit bodies in each cultivated bottle of eight examples were measured by vernier caliper.

2.2 Model Validation

To validate the model the three morphological growth indicators described above were computed with the model and then compared with field data. The following three indices were computed under two situations to assess the closeness of the estimated data to the measured values: correlation coefficient (*R*), bias (*Bias*)[9], root mean square error (*RMSE*)[10], and the percent root mean square error (*%RMSE*)[11].

$$Bias = \frac{1}{n}\sum \left| OBS_i - SIM_i \right|$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (OBS_i - SIM_i)^2}{n}}$$

$$\%RMSE = RMSE \times 100 / (\frac{1}{n}\sum OBS_i)$$

OBS_i and SIM_i are the measured and simulated value, respectively, for the *i*th data point of *n* observations.

3 Results

During the growth period of the fruiting body, the three morphological growth indices of *P. eryngii* showed significant changes over different days (Table 1). We found that the optimal air temperature and humidity facilitated cap expansion, stipe heightening and thickening. For any single growth indicator, the optimal air temperature and humidity was different on each day (Table 2). Multiple statistical comparisons (Table 2) showed that over five days of growth, the optimal air temperature was 16°C, 17°C and 18°C in the first three days, and 17°C or 18°C, and 16°C for the last two days. Variation in air relative humidity over a range of 89–97% did not significantly affect morphological growth of the fruiting body in these experiments (Table 2), but comparatively, the optimum relative humidity was 97%, followed by 95% and 93%.

Table 1. The values of morphological growth indices of *P. eryngii* on each day.

Indices \ Time/d	1	2	3	4	5
Cap diameter	0.729 [a]	1.164 [b]	1.890 [c]	2.922 [d]	4.041 [e]
Stipe height	1.540 [a]	2.415 [b]	3.608 [c]	4.774 [d]	5.825 [e]
Stipe diameter	1.170 [a]	1.433 [b]	1.881 [c]	2.347 [d]	2.639 [e]

*The value of each index was each day's mean growth for both treatments. Different letters indicate statistically different values (ANOVA/LSD) ($P<0.05$).

Table 2. The multiple comparisons result showing the influence of different air temperatures and relative humidity levels on morphological growth of the fruiting body of *P. eryngii* on each day.

	Time/d	Cap diameter	Stipe height	Stipe diameter
Air temperature treatment	1	18°C,17°C,16°C	16°C,17°C,18°C	16°C,18°C,17°C
	2	17°C,16°C,18°C	16°C,17°C,18°C	17°C,18°C,16°C
	3	17°C,18°C	18°C,16°C	16°C,18°C,17°C
	4	17°C	18°C	16°C,18°C
	5	17°C	18°C	16°C
Air relative humidity treatment	1	97%,95%,91%	93%,89%,95%	97%,95%,89%
	2	97%,95%,91%	93%,97%,91%	93%,91%,97%
	3	97%,89%,93%	93%,97%,89%	93%,89%,91%
	4	93%,97%,89%	97%,93%,89%	97%,93%,89%
	5	97%,91%	97%,93%,91%	97%,93%

*The relative humidity/temperature was not considered in the air temperature/relative humidity treatment, and the effect of air temperature/relative humidity was analyzed on morphological growth.

*The former three air temperature and relative humidity conditions that showed no differences between them were listed in the table.

There is complexity and uncertainty in biological growth and *P. eryngii* is no exception. The optimal air temperature is independent of optimal relative humidity. For example, the optimum air temperature for stipe heightening is 18°C, but at that temperature, 93% is not the only optimum relative humidity. Therefore, it is necessary to further analyze and confirm the optimal air environment.

The optimal air temperature for cap expansion is 17°C (Table 3) under relative humidity conditions of 93–97% analyzed above. The optimal relative humidity for cap expansion is dependent on temperature (Table 3): 97% at 16°C and 18°C, and 91–97% at 17°C. So the optimal air temperature and relative humidity conditions for cap expansion are 17°C and 97% respectively.

Under better relative humidity conditions of 93–97%, the optimal air temperature for stipe heightening is 18°C at 93% and 97% relative humidity, and is 17°C and 18°C at 95% relative humidity. The optimal relative humidity for stipe heightening is

also dependent on temperature (Table 3): 97% at 16°C; 91%, 95% and 97% at 17°C; 93% and 97% at 18°C. So the optimal air temperature and relative humidity conditions for stipe heightening is 18°C and 97% relative humidity.

Under better relative humidity conditions of 93–97%, the optimal air temperature for stipe thickening are 16°C and 17°C at 95% and 97%, 16°C and 18°C at 93% relative humidity. The optimal relative humidity for stipe thickening is also dependent on temperature (Table 3): 93% and 97% at 16°C; 91%–95% at 17°C; 97% at 18°C. So the optimal air temperature and relative humidity conditions for stipe thickening is 16°C and 93% relative humidity.

Table 3. The results of multiple comparisons of influence of different air relative humidity/temperature treatment on fruiting body growth under specific air temperature/relative humidity conditions

	#1	16°C	17°C	18°C	#2	91%	93%	95%	97%
Cap diameter	93%		√		16°C				√
	95%		√		17°C	√	√	√	√
	97%		√		18°C				√
Stipe height	93%			√	16°C				√
	95%		√	√	17°C	√		√	√
	97%			√	18°C		√		√
Stipe diameter	93%	√		√	16°C		√		√
	95%	√	√		17°C	√	√	√	
	97%	√	√		18°C				√

#1 the optimum air temperature in temperature treatment under specific air relative humidity condition; #2 the optimum air relative humidity in humidity treatment under specific air temperature condition. The air relative humidity/air temperature conditions that no difference for the morphology growth were marked by the symbol "√".

Through comprehensive analysis of the data, we believe that the optimal air temperature and relative humidity conditions for morphological development of the fruiting body in the growing room is 16–18°C at a relative humidity of 93% and above. On different days, the optimal air temperature and relative humidity conditions for morphological development of the fruiting body are slightly different. In the appropriate range, lower air temperature and relative humidity are better for stipe thickening, higher air temperature and relative humidity are better for stipe heightening.

4 Model Description

4.1 The Morphological Development Simulation Model

4.1.1 Effects of Air Temperature on Morphology Development (ETMD) Sub-model

From the experimental results, the relationships between air temperature and cap diameter, or stipe diameter were single-peak curves; the relationship between air

temperature and height of the stipe was a single-peak curve on the first day, and linear on the other days. Mathematically, the effects of air temperature on morphological development of the fruiting body can be presented as follows:

$$D_{JG} = a_{11} \cdot T^2 + b_{11} \cdot T + c_{11} \tag{1}$$

$$L_{JB} = \begin{cases} a_{21} \cdot T^2 + b_{21} \cdot T + c_{21} & t = 1 \\ u_1 \cdot T + k_1 & t > 1 \end{cases} \tag{2}$$

$$D_{JB} = a_{31} \cdot T^2 + b_{31} \cdot T + c_{31} \tag{3}$$

where D_{JG} is cap diameter, L_{JB} is stipe height, D_{JB} is stipe diameter, T is air temperature (14 –18°C), a_{11}, b_{11}, c_{11}, a_{21}, b_{21}, c_{21}, a_{31}, b_{31}, c_{31}, u_1 and k_1 are parameters, t is time (1–5 days).

With values of T, D_{JG}, L_{JB} and D_{JB} measured experimentally, a_{11}, b_{11}, c_{11}, a_{21}, b_{21}, c_{21}, a_{31}, b_{31}, c_{31}, u_1 and k_1 were estimated through the nonlinear least squares method using *Statistica* 6.0. Because of the small effect of humidity observed experimentally on morphological development of the fruiting body, humidity was not considered in the parameters estimated. As there were substantial differences in the three growth indicators and temperature on different days of the fruiting body growth period, the parameters were estimated separately for each day (see Table 4) over five days.

Table 4. Estimated parameters in Eqs. (1) – (3)

Time/d	a_{11}	b_{11}	c_{11}	a_{21}	b_{21}/u_1	c_{21}/k_1	a_{31}	b_{31}	c_{31}
1	-0.02	0.76	-5.84	-0.03	0.87	-5.69	0.02	-0.56	4.69
2	-0.03	1.10	-8.74		0.27	-1.93	-0.02	0.80	-6.02
3	-0.03	1.25	-9.81		0.42	-3.02	0.02	-0.40	3.68
4	-0.04	1.49	-10.73		0.47	-2.64	0.02	-0.50	4.78
5	-0.22	7.11	-52.11		0.46	-1.57	0.06	-1.78	16.08

4.1.2 Effects of Relative Humidity on Morphological Development (EHMD) Sub-model

Although there was not significant observed effect of varying relative humidity on morphological development, air humidity is almost certainly an important environmental factor for the production of *P. eryngii*. Therefore the mathematical relationship between air humidity and morphological development was also simulated. From the results, the optimal air temperature for *P. eryngii* is 16°C and above. At temperatures of 16°C and higher, the morphological development of the fruiting body has the following mathematical relationship with relative humidity:

$$D_{JG} = a_{12} \cdot H^2 + b_{12} \cdot H + c_{12} \tag{4}$$

$$L_{JB} = u_2 \cdot H + k_2 \tag{5}$$

$$D_{JB} = a_{32} \cdot H^2 + b_{32} \cdot H + c_{32} \tag{6}$$

in which H is relative humidity (89–97%), a_{12}, b_{12}, c_{12}, a_{32}, b_{32}, c_{32}, u_2 and k_2 are parameters.

With the experimentally measured H, D_{JG}, L_{JB} and D_{JB}, a_{12}, b_{12}, c_{12}, a_{32}, b_{32}, c_{32}, u_2 and k_2 were estimated using the same method as for the parameters in Eqs. (1) – (3). As there were substantial differences of morphological development indices on different days of fruiting body growth period, the parameters were estimated separately for each day (see Table 5) over five days.

Table 5. Estimated parameters in Eqs. (4) – (6)

Time/d	a_{12}	b_{12}	c_{12}	a_{32}	b_{32}	c_{32}	u_2	k_2
1	-0.004	0.66	-29.74	-0.001	0.10	-3.26	0.01	0.44
2	-0.01	1.78	-81.63	-0.002	0.40	-17.12	0.09	-5.65
3	-0.0003	0.05	-0.69	-0.002	0.43	-17.81	0.10	-4.74
4	-0.005	0.94	-40.78	-0.001	0.10	-2.39	0.12	-6.23
5	-0.02	3.68	-166.35	0.004	-0.69	34.34	0.09	-2.41

4.2 Validation of the Morphological Development Simulation Model

The morphological development simulation model was validated by assessing the correlation coefficient (*R*), bias (*Bias*), root mean square error (*RMSE*) and the percent root mean square error (*%RMSE*) for the measured and estimated values. The detailed results are shown in Table 6.

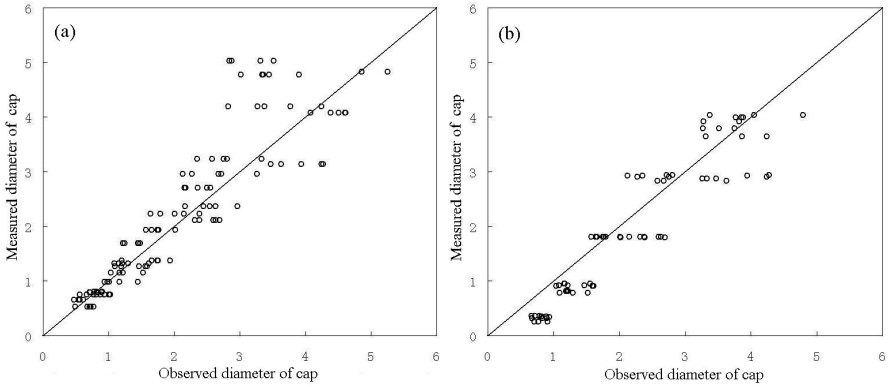
From the mathematical model, the effect of air temperature gave the following results: *R, Bias*, *RMSE* and *%RMSE* were 88.6%, 0.325 cm, 0.335 and 13.42%, respectively, for cap diameter (Eq.1); 93.7%, 0.226 cm, 0.305 and 8.40%, respectively, for stipe height (Eq.2); and 92.6%, 0.173 cm, 0.213 and 10.09%, respectively, for stipe diameter (Eq.3).

The effect of relative humidity resulted in the following:R, Bias, RMSE and %RMSE were 85.4%, 0.350 cm, 0.396 and 17.42%, respectively, for cap diameter (Eq.4); 91.7%, 0.297 cm, 0.392 and 9.84%, respectively, for stipe height (Eq.5); 93.4%, 0.237 cm, 0.276 and 19.97%, respectively, for stipe diameter (Eq.6).

The estimated morphological development indices are correlated well with the measured values. This is also shown from Figures 1-3.

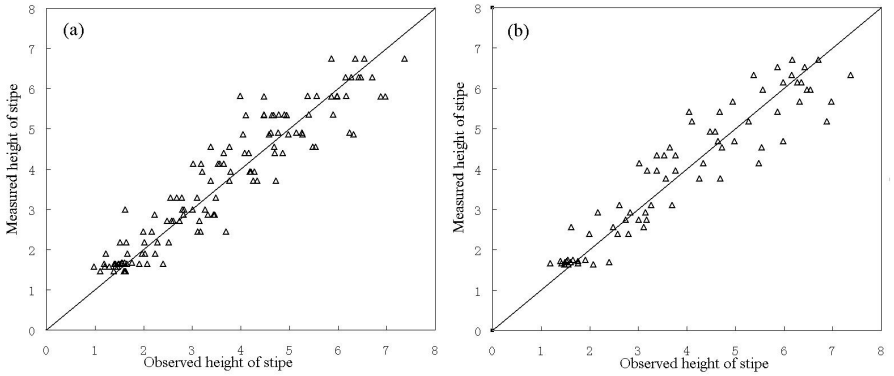
Table 6. The values of validation indices for the morphological development simulation model

Model	Index	Model validation indexes			
		R^2	*Bias /cm*	*RMSE*	*%RMSE*
ETMD sub-model	Cap diameter	88.6%	0.325	0.335	13.42%
	Stipe height	93.7%	0.226	0.305	8.40%
	Stipe diameter	92.6%	0.173	0.213	10.09%
EHMD sub-model	Cap diameter	85.4%	0.350	0.396	17.42%
	Stipe height	91.7%	0.297	0.392	9.84%
	Stipe diameter	93.4%	0.237	0.276	19.97%



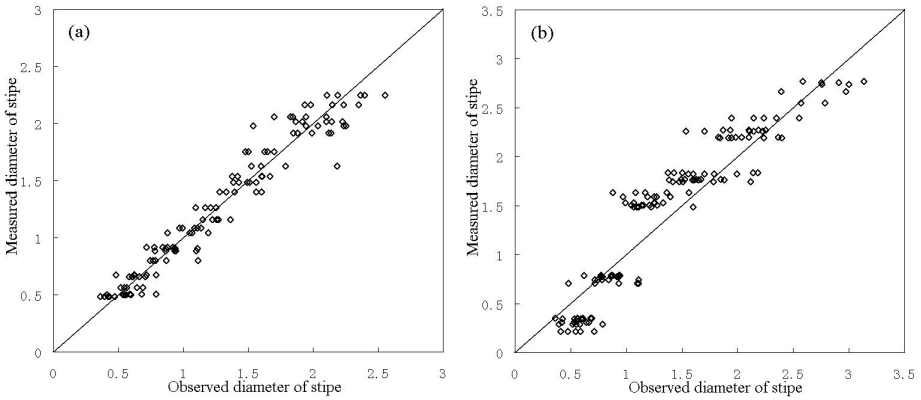
(a) Simulations using the effects of air temperature on the morphological development sub-model (Eq. 1); (b) simulations using the effects of relative air humidity on the morphological development sub-model (Eq. 4). The dashed line (—) is the 1:1 line.

Fig. 1. Measured vs. simulated diameter of the cap



(a) Simulations using the effects of air temperature on morphological development sub-model (Eq. 2); (b) simulations using the effects of relative air relative humidity on the morphological development sub-model (Eq. 5). The dashed line (—) is the 1:1 line.

Fig. 2. Measured vs. simulated height of the stipe



(a) Simulations using the effects of air temperature on the morphological development sub-model (Eq. 3); (b) simulations using the effects of relative air humidity on the morphological development sub-model (Eq. 6). The dashed line (—) is the 1:1 line.

Fig. 3. Measured vs. simulated diameter of the stipe

4.3 The Appropriate Environment Range Model for Morphological Development

4.3.1 The Appropriate Air Temperature Range Sub-model for Morphological Development

From equations (1) – (3), we obtained the following formulas:

$$T_{opt-JG} = [\frac{-b_{11} - \sqrt{b_{11}{}^2 - 4a_{11}(c_{11} - SD_{JG})}}{2a_{11}}, \frac{-b_{11} + \sqrt{b_{11}{}^2 - 4a_{11}(c_{11} - SD_{JG})}}{2a_{11}}] \tag{7}$$

$$T_{opt-LJB} = \begin{cases} [\frac{-b_{21} - \sqrt{b_{21}{}^2 - 4a_{21}(c_{21} - SL_{JB})}}{2a_{21}}, \frac{-b_{21} + \sqrt{b_{21}{}^2 - 4a_{21}(c_{21} - SL_{JB})}}{2a_{21}}] & t = 1 \\ (SL_{JB} - k_1) \Big/ u_1 & t > 1 \end{cases} \tag{8}$$

$$T_{opt-DJB} = [\frac{-b_{31} - \sqrt{b_{31}{}^2 - 4a_{31}(c_{31} - SD_{JB})}}{2a_{31}}, \frac{-b_{31} + \sqrt{b_{31}{}^2 - 4a_{31}(c_{31} - SD_{JB})}}{2a_{31}}] \tag{9}$$

$$T_{opt} = T_{opt-JG} \cap T_{opt-LJB} \cap T_{opt-DJB} \tag{10}$$

where $T_{opt\text{-}DJG}$ is the optimal air temperature for cap diameter, $T_{opt\text{-}LJB}$ is the optimal air temperature for stipe height, $T_{opt\text{-}DJB}$ is the optimal air temperature for stipe diameter. T_{opt} is the optimal air temperature for the fruiting body, which is the intersection of $T_{opt\text{-}DJG}$, $T_{opt\text{-}LJB}$ and $T_{opt\text{-}DJB}$. SD_{JG}, SL_{JB} and SD_{JB} are standard values of cap diameter,

stipe height and stipe diameter on different days, and is taken from the average value of the indices in this study (see Table 7).

4.3.2 The Appropriate Relative Humidity Range Sub-model for Morphological Development

From equations (4) – (6), we obtained the following formulas:

$$H_{opt-JG} = [\frac{-b_{12} - \sqrt{b_{12}{}^2 - 4a_{12}(c_{12} - SD_{JG})}}{2a_{12}}, \frac{-b_{12} + \sqrt{b_{12}{}^2 - 4a_{12}(c_{12} - SD_{JG})}}{2a_{12}}] \quad (11)$$

$$H_{opt-LJB} = (SL_{JB} - k_2) \Big/ u_2 \quad (12)$$

$$H_{opt-DJB} = [\frac{-b_{32} - \sqrt{b_{32}{}^2 - 4a_{32}(c_{32} - SD_{JB})}}{2a_{32}}, \frac{-b_{32} + \sqrt{b_{32}{}^2 - 4a_{32}(c_{32} - SD_{JB})}}{2a_{32}}] \quad (13)$$

$$H_{opt} = H_{opt-JG} \cap H_{opt-LJB} \cap H_{opt-DJB} \quad (14)$$

where $H_{opt\text{-}DJG}$ is the optimal relative humidity for cap diameter, $H_{opt\text{-}LJB}$ is the optimal relative humidity for stipe height, $H_{opt\text{-}DJB}$ is the optimal relative humidity for stipe diameter. H_{opt} is the optimal relative humidity for the fruiting body, which is the intersection of $H_{opt\text{-}DJG}$, $H_{opt\text{-}LJB}$ and $H_{opt\text{-}DJB}$.

4.3.3 Simulated Optimal Environmental Conditions for the Fruiting Body Growth Period

Using equations (7) – (14), the optimal environmental control standards can be computed for each day of the fruiting body growth period (Table 7).

Table 7. The standard values of morphological development and the simulated optimal environment conditions in the fruiting body growth period

Time /d	SD_{JG} /cm	SL_{JB} /cm	SD_{JB} /cm	Optimal temperature range /°C	Optimal humidity range /%
1	0.729	1.540	1.170	[15.4,16.9]	[87,96]
2	1.164	2.415	1.433	[15.9,19.1]	[89,95.4]
3	1.890	3.608	1.881	[15.8,16.4]	[87,96]
4	2.922	4.774	2.347	[15.8,16.4]	[91,96]
5	4.041	5.825	2.639	[14.5,17.2]	[91,94]

5 Conclusion

When compared with traditional cultivation in plastic greenhouses, modern methods of commercial mass cultivation of fungi has the advantage of uniform growth, high efficiency and is unrestricted by season. However, production is costly and has high energy consumption. Therefore, optimal environmental control is crucial. Higher fungi are considered difficult to cultivate in the laboratory without complex growth medium[12] because of the limited growing space in a laboratory. Environmental control needs to be very precise in modern commercial production of fungi, and the application of the optimal climate control parameters is crucial for ideal fruiting body development, and therefore these study results have great practical significance. Using the simulation models morphological development rate and optimal climate control values can be obtained to guide the actual environmental control settings in the commercial production of *P. eryngii*.

In general, fungi are very difficult to study through experimental means alone because of the complexity of their natural growth habitat (e.g., soils) and the microscopic scale of growth (e.g., tip vesicle translocation and hyphal tip extension)[13]. Some research focus on shape is on the cell[14].Mathematical modeling provides a complementary, powerful and efficient method of investigation. The aim of mathematical modeling is to reduce a complex (biological) system into a simpler (mathematical) system that can be analyzed in far more detail and from which key properties can be identified[13].

Temporal effects were considered in the model. The parameters in the model varied for each day. The relationship between the indices of morphological development and air environment in this paper was represented by a quadratic equation with one variable, reflecting the experimental results which showed that the estimated optimal air temperature and relative humidity are not the highest values of appropriate air temperature and relative humidity. The calculated optimal air temperature and relative humidity values predicted by the model are different for each day. For this reason, environmental parameters should be varied daily in the commercial production of *P. eryngii*, and not set at temperatures of 16–18°C or 93% above relative humidity, as analyzed to be optimal by the experiments. This also illustrates the practical value of the model.

We conclude that our model described in this article provides a powerful tool to predict the morphological development rate according to air environment, and to guide real-time adjustment of air environment for commercial production of *P. eryngii*.

Acknowledgements. This work was supported by grants from the National Natural Science Foundation of China (No. 30800765) and the Project of the Science and Technology Commission of Shanghai Municipality, China (Nos. 08DZ2210600 and 08QA14058). Carbohydr. Polym.

References

1. Ng, T.B.: A review of research on the protein-bound polysaccharide (polysaccharopeptide, PSP) from the mushroom Coriolus versicolor (Basidiomycetes: Polyporaceae). Gen. Pharmacol. 30, 1–4 (1998)
2. Rodriguez Estrada, A.E., Royse, D.J.: Yield, size and bacterial blotch resistance of Pleurotus eryngii grown on cottonseed hulls/oak sawdust supplemented with manganese, copper and whole ground soybean. Bioresour. Technol. 98, 1898–1906 (2007)
3. Carbonero, E.R., Gracher, A.H.P., Smiderle, F.R., et al.: A β-glucan from the fruit bodies of edible mushrooms Pleurotus eryngii and Pleurotus ostreatoroseus. Carbohydr. Polym. 66, 252–257 (2006)
4. Ng, T.B., Wang, H.X.: A novel ribonuclease from fruiting bodies of the common edible mushroom Pleurotus eryngii. Peptides 25, 1365–1368 (2004)
5. Yamanaka, K.: Cultivation of new mushroom species in East Asia. In: Tan, et al. (eds.) Proceedings of the Fifth International Conference on Mushroom Biology and Mushroom Products, Shanghai, China, April 8-12; Acta Edulis Fungi 12 (suppl.), 343–349 (2005)
6. Chang, S.T.: Witnessing the development of the mushroom industry in China. In: Tan, et al. (eds.) Proceedings of the Fifth International Conference on Mushroom Biology and Mushroom Products, Shanghai, China, April 8-12; Acta Edulis Fungi 12 (suppl.), 3–19 (2005)
7. Tan, Q., Wang, Z., Cheng, J., Guo, Q., Guo, L.: Cultivation of Pleurotus spp. in China. In: Tan, et al. (eds.) Proceedings of the Fifth International Conference on Mushroom Biology and Mushroom Products, Shanghai, China, April 8-12; Acta Edulis Fungi 12 (suppl.), 338–349 (2005)
8. Royse, D.J., Shen, Q., McGarvey, C.: Consumption and production of recently domesticated edible fungi in the United States with a projection of their potential. In: Tan, et al. (eds.) Proceedings of the Fifth International Conference on Mushroom Biology and Mushroom Products, Shanghai, China, April 8-12; Acta Edulis Fungi 12 (suppl.), 331–337 (2005)
9. Kobayashi, K., Salam, M.U.: Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92, 345–352 (2000)
10. Moreno-Sotomayora, A., Weiss, A.: Improvements in the simulation of kernel number and grain yield in CERES-Wheat. Field Crops Res. 88, 157–169 (2004)
11. Jansen, P.H.M., Heuberger, P.S.C.: Calibration of process-oriented models. Ecol. Model 83, 55–66 (1993)
12. Bazała, M.A.: Transport of radiocesium in mycelium and its translocation to fruitbodies of a saprophytic macromycete. Environ. Radioact., 1–3 (2008)
13. Boswell, G.P., Jacobs, H., Gadd, G.M., et al.: A mathematical approach to studying fungal mycelia. Mycologist 17, 165–171 (2003)
14. Braaksma, A., van Doorn, A.A., Kieft, H., et al.: Morphometric analysis of ageing mushrooms (Agaricus bisporus) during postharvest development. Postharvest Bio. Tech. 13, 71–79 (1998)

Development of a Web-Based Prediction System for Wheat Stripe Rust

Weigang Kuang[1], Wancai Liu[2], Zhanhong Ma[1], and Haiguang Wang[1,*]

[1] Department of Plant Pathology, China Agricultural University, Beijing 100193, China
[2] National Agro-Tech Extension and Service Center, Ministry of Agriculture, Beijing 100125, China
wanghaiguang@cau.edu.cn

Abstract. A web-based prediction system for wheat stripe rust was developed based on B/S (Browser/Server) mode in this study. Some existing prediction models of wheat stripe rust were collected, analyzed and then stored in SQL Server 2005 database according to certain rules. All these models could be called through this web-based system and used to predict wheat stripe rust. Meanwhile, Using multiple regression analysis principle, prediction regression model could be built based on the input historical data of wheat stripe rust through the network programming via this system, and significance tests of prediction factors could be conducted to obtain optimal prediction model and the built model could be stored into the model database for further prediction of this disease. Using WebGIS technologies, the prediction results of wheat stripe rust could be displayed in different colors in the web map according to the prediction values of disease prevalence. The web-based prediction system for wheat stripe rust developed in this study provided a convenient and fast way for the prediction of wheat stripe rust.

Keywords: wheat stripe rust, web-based system, prediction, regression model, WebGIS.

1 Introduction

Wheat (*Triticum aestivum*) with large plating area, is one of the major food crops in China. Pest damage to wheat yield and quality is very serious. Wheat stripe rust (or wheat yellow rust), caused by *Puccinia striiiformis* f. sp. *tritici*, is the most important wheat disease in China [1]. It is an important airborne plant disease. Severe epidemics of wheat stripe rust occurred in 1950, 1964, 1990 and 2002 in China and caused server yield losses resulting in significant economic losses [1], [2]. Therefore, timely and accurate prediction of wheat stripe rust has an important significance for taking timely and effective disease control measures and safe production of wheat.

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 324–335, 2013.

Many methods used to predict wheat stripe rust have been reported, such as regression analysis [3], [4], [5], [6], [7], discrimination analysis [8], [9], [10], Markov forecast method [11], principal component analysis [12], grey model forecast method [13], [14], neural networks [15], [16], [17], [18] and support vector machine [19]. Generally, prediction functions of wheat stripe rust by using these methods were realized on single PC. A few of single computer versions of wheat stripe rust prediction systems have been developed [20], [21], such as PANCRIN that could be used to make simulation experiments on the pandemic of wheat stripe rust in China [21].

WebGIS technology is an emerging technology used in plant protection. It is convenient to process geographic information based on Internet by using WebGIS technology. It plays an important role in pest monitoring and forecasting by making full use of the processing capabilities of geographic information system (GIS) for spatial data and combining GIS with network technologies effectively. Based on long-term study of Sudden Oak Death in California coastal region, Kelly and Tuxen constructed a warning and monitoring website for Sudden Oak Death in California called the "OakMapper" by using WebGIS technology [22]. The information of the website is quarried and browsed in the web map. Users can get the latest information of disease occurrence and learn some knowledge related disease prevention by visiting the website. PhytoPRE+2000, an Internet based version of decision support system for potato late blight was developed for more efficient services in monitoring and management of this disease in Swiss [23]. A potato late blight monitoring and warning system "china-blight" (www.china-blight.net) was constructed based on B/S (Browser/Server) internet structure by combining information technology with the principles of plant disease epidemiology [24]. The system could provide the information on the infection risk of late blight pathogen in different districts of China in the coming 48 hours and also could provide decision support for the chemical control of this disease. With the development of information technologies, the Internet based or WebGIS based plant disease prediction systems play a more and more important role in monitoring and warning of plant diseases.

To the authors' knowledge, there is not any report about web-based wheat stripe rust prediction system. In this study, a prediction system for wheat stripe rust in China was developed using WebGIS technologies. Prediction of wheat stripe rust could be conducted using the existing prediction models or prediction regression models built by using this system based on the input historical data by users. The prediction of this disease by using the information and network technologies was achieved. A convenient and fast way was provided for the prediction of wheat stripe rust and the web-based services also could be provided for integrated control of this disease.

2 Development Platform of This System

The web-based prediction system for wheat stripe rust was developed based on the .NET platform. C# and JavaScript network programming language were used as the system development languages. Internet Information Server (IIS) was used as the web server and ArcGIS Server 9.3 of ESRI Company was used as the map server. The data

information was stored in SQL Server 2005 database. The Microsoft Visual Studio 2008 was used as the development tool to develop the web-based prediction system for wheat stripe rust.

3 Overall Structure of This System

The web-based prediction system for wheat stripe rust was developed based on B/S mode which was a three-layer model structure including web application service layer, logic layer and database layer. The system was composed of the client side and the background (as shown in Fig. 1). Using the historical disease prediction models and the models built by using the system based on the input historical data, the prediction of wheat stripe rust could be implemented and the prediction results could be displayed in different colors in the web map at the client side according to disease prevalence. The background included user management module and model management module that could be used to manage the users and to manage the database, respectively.

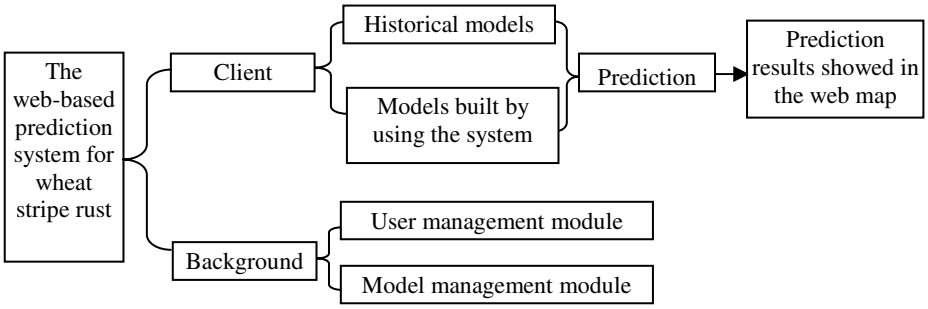


Fig. 1. Overall structure of the web-based prediction system for wheat stripe rust

4 Function Realization of the Web-Based Prediction System

The disease prediction function of this web-based system was realized using two methods. One was to realize the function by using the historical models. The existing prediction models were collected and stored in the background database. The models could be queried from the database and then the suitable model could be selected to predict wheat stripe rust. Another method was to realize the prediction function by using the models built based on the input historical data of wheat stripe rust. The function structure chart of the web-based prediction system for wheat stripe rust was shown in Fig. 2.

4.1 Building Model Database of Wheat Stripe Rust

Historical regression prediction models of wheat stripe rust were collected through reviewing relevant literature. The models were stored into model database mainly according to provinces where the disease occurred, counties where the disease occurred, model types, coefficient value of impact factor, the values of constant term of the models, use instructions of the models, and so on. The model database was composed of prediction models and original data table. The models were stored in the model table including the fields, i.e. fid, fdisesse, fprov, fcity, fmodel, fx1, fx2, fx3, fx4, fx5, fx6, fchang and fexplain (as shown in Table 1). The input historical data of wheat stripe rust used to build regression model was stored in the original data table.

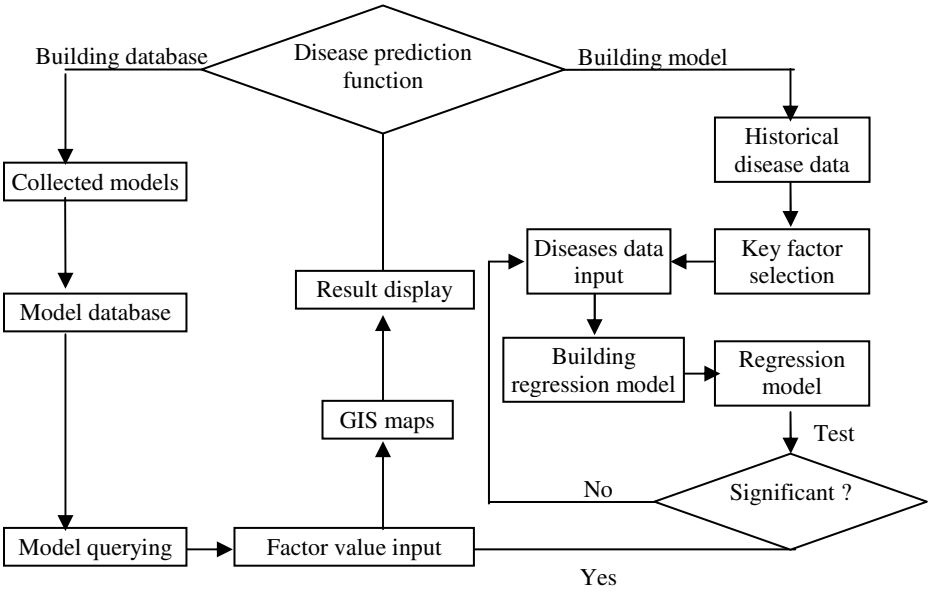


Fig. 2. Function structure chart of the web-based prediction system for wheat stripe rust

Table 1. Disease prediction model table

Field name	Data type	Comments
fid	int	
fdisesse	nvarchar(50)	Disease name
fprov	nvarchar(50)	Provinces where the disease occurred
fcity	nvarchar(50)	counties where the disease occurred
fmodel	nvarchar(50)	Model type
Fx1, Fx2, Fx3, Fx4, Fx5 and Fx6	float	The coefficient value of impact factor 1, 2, 3, 4, 5 and 6
fchang	float	The values of constant term
fexplain	text	Use instructions of the models

The regression prediction models of wheat stripe rust that were applied in Ganggu County in Gansu Province, Longnan in Gansu Province, Hanzhong in Shaanxi Province, Shanxi Province and so on, respectively, were collected for the web-based prediction system (as shown in Table 2). Users can query the models of corresponding regions by inputting information at the client side, and then use the corresponding model to carry out the prediction of wheat stripe rust.

Table 2. Some historical prediction model stored in the model database of the web-based prediction system for wheat stripe rust

The suitable region for using prediction model	Prediction model	Variable	Reference
Hanzhong in Shaanxi Province	$y = -4.699158 + 0.199445x_1 + 0.052963x_2 + 0.314299x_3 + 0.014257x_4 + 0.008677x_5$	x_1 is the amount of pathogen in autumn, x_2 is the amount of pathogen in spring, x_3 is the average temperature in April, x_4 is the precipitation in April, x_5 is the proportion of susceptible variety area, and y is disease prevalence.	[5]
Shanxi Province	$y = -0.2766 + 0.0249x_1 + 0.0104x_2 + 0.0132x_3$	x_1 is the precipitation in the middle ten days of April to the first ten days of May, x_2 is the precipitation in the middle ten days of May and the last tens of May, x_3 is the average relative humidity in May, and y is disease prevalence.	[6]
Gangu County in Gansu Province	$y=-2.07472 + 0.242533x_1 + 0.06744x_2 - 0.0091x_3 + 0.011865x_4 + 0.090152x_5$	x_1 is the average temperature in March, x_2 is the precipitation in March, x_3 is the precipitation in April, x_4 is the precipitation in May, x_5 is the susceptible variety area, and y is disease prevalence.	[7]
Nonglan in Gansu Province	$y = -0.547 + 0.025\ x_1 + 0.03\ x_2 + 0.29\ x_3 + 0.081\ x_4$	x_1 is the percentage of diseased fields in autumn in the preceding year, x_2 is the proportion of susceptible variety area, x_3 is the average temperature in January in Huixian, x_4 is the average maximum temperature in March in Wudu, and y is disease prevalence in Spring in Nonglan.	[25]
Jincheng in Shanxi Province	$y = 3.7644 + 0.0070x_1 - 0.0097x_2 + 0.0273x_3 + 0.1079x_4 - 0.2145x_5$	x_1 is the precipitation in the last ten-day period of July in the preceding year, x_2 is sunlight in the first ten days of November in the preceding year, x_3 is sunlight in the first ten days of March in the preceding year, x_4 is the average temperature in the middle ten days of April, x_5 is the precipitation in the last ten-day period of April, and y is disease prevalence.	[26]
Pingliang in Gansu Province	$y = -0.1736 + 0.2051x_1 + 0.4345x_2 + 0.0089x_3$	x_1 is disease prevalence in spring, x_2 is the amount of oversummering pathogen, x_3 is the total precipitation in September and October, and y is disease prevalence on autumn wheat seedlings.	[27]

4.2 Realization of Automatically Modeling Using the System

The quantity of collected regression prediction model of wheat stripe rust is limited, and most of the models could only be used locally. With the changes in the preconditions, the accuracy of the models should be further improved. Therefore, the model construction module was designed in the web-based prediction system for wheat stripe rust so that the prediction models could be built by using the system. Users can establish regression prediction models using the principle of multiple regression with the help of network programming based on the input historical data of wheat stripe rust. F test on the independent variables at the level of 0.05 or 0.01 could be conducted by using the system to help users to get optimization model.

Regression analysis is a statistical method used to determine the quantitative relationship between two or more variables. Monadic linear regression is to use one of the major factors as independent variable to explain the change of the dependent variable. Multiple regression is to use two or more factors as independent variables to explain the change of the dependent variable. Least square method is used to calculate parameters during regression analysis.

K linear regression using *N* groups of observations can be solved through the following matrix (as shown in Equation (1)).

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \dots \\ B_K \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{bmatrix} \tag{1}$$

in which, vector *y* is observation variable vector, matrix *x* is the matrix of observation variables, vector *B* is regression parameter vector, vector *u* is random error vector. The simplified form of Equation (1) could be expressed as Equation (2).

$$Y = XB + U \tag{2}$$

According to the principle of least square method, the estimate value of the minimum parameter vector *B* could be expressed as Equation (3).

$$B = (X'X)^{-1}(X'Y) \tag{3}$$

Automatically modeling using the system could be realized in accordance with the following steps. Prediction models of wheat stripe rust could only be built by using regression analysis method via this system. Firstly, the targeted region where disease prediction result would be used should be determined. Secondly, model type should be chosen according to the related data. Five kinds of regression models could be built via this system including binary linear equation, ternary linear equation, first-order linear equation with four independent variables, first-order linear equation with five independent variables, and first-order linear equation with six independent variables. After choosing the proper model and clicking the OK button, a system interface as

shown in Fig. 3 would be displayed in the client browser. Then the prediction factor values could be input into the system. After inputting all the historical data into the system and clicking the building model button. The system would deal with the data by using regression analysis method. Finally, the prediction model would be built just as shown in Fig. 4.

After building the prediction model by using regression analysis method via the web-based prediction system for wheat stripe rust, F tests on the independent variables at the level of 0.05 or 0.01 could be conducted to select the model with more accurate by using the system and the optimization model could be obtained. Then the relevant information about the model including the applicable area and the instructions of variables should be input into the system through system interface so that the model could be used in correct way later. The model could be stored into the model database and be used as the same as the historical regression prediction models of wheat stripe rust.

Fig. 3. The system interface for inputting data

4.3 Realization of Prediction Function Based on WebGIS

Using the web-based prediction system for wheat stripe rust, the relevant regional prediction model can be queried through the user interface of the system. According to the instruction of the selected model, after the corresponding information being input into the system, the model could be used to predict wheat stripe rust. In order to

Fig. 4. The prediction model built by using the web-based prediction system for wheat stripe rust

make the prediction results more intuitive, WebGIS technologies were combined with the network technologies in this system and the results could be displayed in different colors in the web map at the client side according to disease prevalence.

Spatial database was established to store national geographic information. The maps used in this system were administrative electronic maps of China which were acquired from the National Geographic Information System in China. The map scale was 1:400000000. The e00 format of the maps was converted to the Shape format by using ArcCatalog software firstly. The ArcGIS 9.3 was used to deal with the vector maps. And the map information was released by using ArcCatalog software.

In the web-based prediction system, the maps were divided into three layers; the first was the county and city boundary layer, the second was the province boundary layer and the third was Chinese boundary layer. Disease prevalence of wheat stripe rust was divided into five classes represented by 1, 2, 3, 4 and 5, respectively. The greater the grade of disease is, the more severe the disease prevalence is. Class 1, 2, 3, 4 and 5 are displayed in green, purple, blue, yellow and red, respectively. This function of the system could be demonstrated by the prediction model of wheat stripe rust in Gangu County that was built by Fan et al. [7]. Firstly the background model

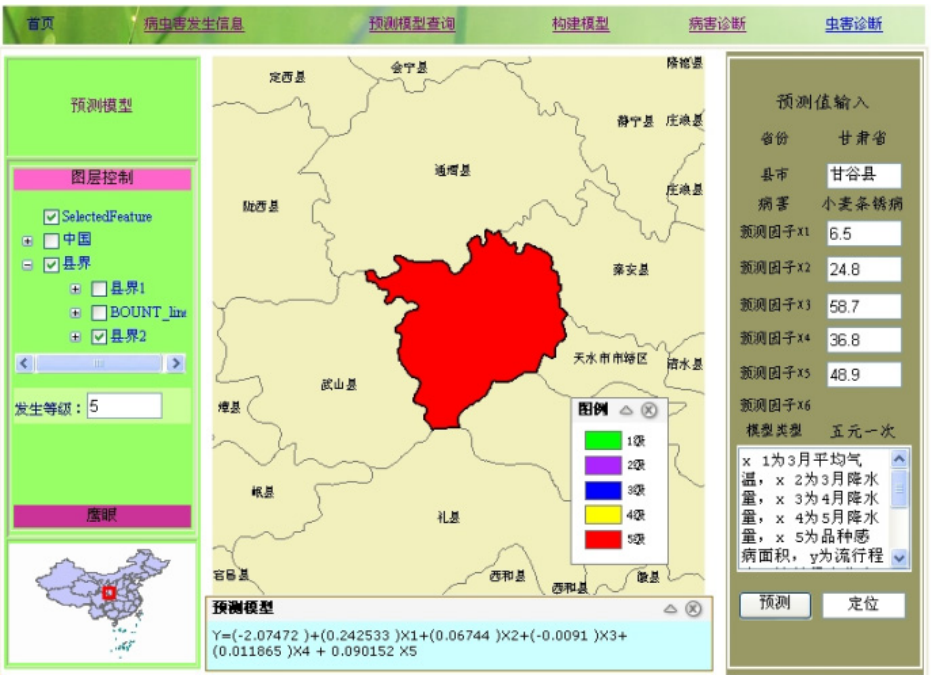


Fig. 5. The prediction results of wheat stripe rust for Gangu County using the web-based prediction system

that was suitable for Gansu County was called in accordance with the rules of the system. Then the model could be used to predict the disease prevalence of wheat stripe rust in Gangu County. The prediction result was shown in Fig. 5.

5 Maintenance of Model Database

Due to the changes of environments conditions, tillage practices and so on, the accuracies of the prediction models would decrease. In order to improve the model database continually, the prediction model can be added or modified via system background. Thus the validation and verification of the prediction models could be ensured. After inputting the correct user's name and password, the user can enter the system background. The kinds of regression prediction models stored in background were divided into 5 types just like described above. When the user selects the model type that he or she would like to add, the number of variables is determined. Then the user can fill in the form in the system interface with relevant information according to the instruction of the system. In order to ensure that the model could be correctly used later, the variables and using method of the model should be described in the system. After all information is input, the user could click the "Add" button and the model would be stored into the model database. By using the query function, the added models could be selected to predict wheat stripe rust.

6 Conclusion and Discussion

A web-based prediction system for wheat stripe rust was developed using network information technologies in this study. The existing prediction models stored in the model database of the system could be used to predict wheat stripe rust. Moreover, regression models could be built via the system by using C# network programming technologies based on the principle of multiple regression analysis. The optimized models could be stored into the model database and users could select the suitable model among the database to predict wheat stripe rust in certain region. Thus the waste of information resources and the repeat development of prediction models could be avoided. This system built in this study provided a convenient and fast way for the prediction of wheat stripe rust.

In general, the prediction of wheat stripe rust was conducted on single PC. The web-based prediction of wheat stripe rust could not be carried out before in China. However, the web-based prediction of other plant diseases has been implemented and plays an import role in practical application in China [24]. Based on B/S structure, the system built in this study could display the prediction results in different colors by means of the powerful spatial information processing ability and graphical display function of WebGIS technologies. By combining GIS maps with the prediction models, the system makes the prediction results of wheat stripe rust more intuitive. The prediction of wheat stripe rust could be conducted by using the models built based on the related data via the system and the historical regression prediction models stored in the model database. This web-based prediction could make the prediction of wheat stripe rust more efficient and more convenient and then could provide better services for control decision making.

The development of the web-based prediction system for wheat stripe rust was a meaningful attempt in informationization and automatization of monitoring and early warning on wheat stripe rust. But the functions of spatial analysis and prediction of this system still need further improvement. There are many kinds of methods to construct the prediction models of wheat stripe rust. Each kind of method has its advantages and disadvantages. Only the historical regression prediction models were collected and stored in the model database of this system. And the models that could be built by using this system were also regression models. To make this system more useful, other types of prediction models and other kinds of modeling methods should be added to this system. The prediction results would become more accurate on the basis of complementary advantages of different prediction methods. In order to improve the reliability of results, multiple prediction results using different methods should be showed in overlay format. Thus a scientific basis would be provided for the integrated management of wheat stripe rust.

Acknowledgments. This work was supported in part by National Key Technology R&D Program (2012BAD19B04) and Special Fund for Agro-scientific Research in the Public Interest (200903004 and 200903035).

References

1. Li, Z.Q., Zeng, S.M.: Wheat Rusts in China. China Agriculture Press, Beijing (2002) (in Chinese)
2. Wan, A., Zhao, Z., Chen, X., He, Z., Jin, S., Jia, Q., Yao, G., Yang, J., Wang, B., Li, G., Bi, Y., Yuan, Z.: Wheat Stripe Rust Epidemic and Virulence of *Puccinia striiformis* f. sp. *tritici* in China in 2002. Plant Disease 88, 896–904 (2004)
3. Zeng, S.M.: Interregional Spread of Wheat Yellow Rust in China. Acta Phytopathologica Sinica 18, 119–223 (1988) (in Chinese)
4. Yang, Z.W., Shang, H.S., Pei, H.Z., Xie, Y.L.: Dinamic Forecasting of Stripe Rust of Winter Wheat. Scientia Agricultura Sinaca 24, 45–50 (1991) (in Chinese)
5. Hu, X.P., Yang, Z.W., Li, Z.Q., Deng, Z.Y., Ke, C.H.: Studies on the Prediction of Wheat Stripe Rust Epidemics in Hanzhong District of Shaanxi Province. Acta Univ. 28, 18–21 (2000) (in Chinese)
6. Fan, S.Q., Xie, X.S., Li, F., Yin, Q.Y., Zheng, W.Y.: Forecast Model for Prevalent Stripe Rust in Winter Wheat in Shanxi Province. Chinese Journal of Eco-Agriculture 15, 113–115 (2007) (in Chinese)
7. Fan, Z.Y., Cao, S.Q., Luo, H.S., Jin, S.L.: Prediction Model on Wheat Stripe Rust in Gangu Country. Gansu Agriculture Science and Technology 10, 9–21 (2008) (in Chinese)
8. Chen, G., Wang, H.G., Ma, Z.H.: Forecasting Wheat Stripe Rust by Discrimination Analysis. Plant Protection 32, 24–27 (2006) (in Chinese)
9. Chen, G., Wang, H.G., Zhang, L.D., Wang, T., Ma, Z.H.: Preliminary Research on the Regional Relationship of Epidemic of *Puccinia striiformis* in China. Chinese Agricultural Science Bulletin 22, 415–420 (2006) (in Chinese)
10. Yun, X.W., Wang, H.G., Ma, Z.H.: Forecast of Wheat Stripe Rust by Upper-air Wind. Chinese Agricultural Science Bulletin 23, 358–363 (2007) (in Chinese)
11. Qiang, Z.F.: Markov Forecast of Wheat Stripe Rust in Qinghai in 1998. Plant Protection 25, 19–22 (1999) (in Chinese)
12. Yuan, L., Li, S.Q.: Application of Principal Component Analysis of Wheat Stripe Rust. Computer Engineering and Design 31, 459–461 (2010) (in Chinese)
13. Pu, C.J.: On Periodic Epidemic Regular Pattern and Prediction of Wheat Stripe Rust in Gansu Province. Acta Phytopathologica Sinica 28, 299–302 (1998) (in Chinese)
14. Liu, R.Y., Ma, Z.H.: The Prediction Methodology of Wheat Stripe Rust Using Combination Model Based on GM (1, l). Journal of Biomathematics 22, 343–347 (2007) (in Chinese)
15. Hu, X.P., Yang, Z.W., Li, Z.Q., Deng, Z.Y., Ke, C.H.: Prediction of Wheat Stripe Rust in Hanzhong Area by BP Neural Network. Acta Agriculturae Boreali-occidentalis Sinica 9, 28–31 (2000) (in Chinese)
16. Zhang, J.: Research on Chaotic Characteristics of the Disaster Rate of Crops and Its GA-BPNN Forecasting Model. Jour. of Northwest Sci-Tech Univ. of Agri. and For(Nat. Sci. Ed.). 34, 63–66 (2006) (in Chinese)
17. Jin, N., Huang, W.J., Jing, Y.S., Wang, D.C., Luo, J.H.: Long-term Meteorological Prediction of Country Wide Wheat Stripe Rust by Genetic Neural Network. Chinese Journal of Agrometeorology 30, 243–247 (2009) (in Chinese)
18. Wang, H.G., Ma, Z.: Prediction of Wheat Stripe Rust Based on Neural Networks. In: Li, D., Chen, Y. (eds.) CCTA 2011, Part II. IFIP AICT, vol. 369, pp. 504–515. Springer, Heidelberg (2012)

19. Wang, H.G., Ma, Z.: Prediction of Wheat Stripe Rust Based on Support Vector Machine. In: Proceedings of 2011 Seventh International Conference on Natural Computation (ICNC 2011), pp. 389–393 (2011)
20. Xiao, C.L., Zeng, S.M.: A Prototype Expert System for Long-term Prediction of Stripe Rust Epidemics. Acta Agriculturae Universitatis Pekinensis 16(suppl.), 126–132 (1990) (in Chinese)
21. Zeng, S.M., Zhang, M.R.: Simulation Experiments on the Pandemic Dynamic of Wheat Stripe Rust in China by Means of a Simulation Model PANCTRIN. Acta Agriculturae Universitatis Pekinensis 16(suppl.), 151–162 (1990) (in Chinese)
22. Kelly, N.M., Tuxen, K.: WebGIS for Monitoring "Sudden Oak Death" in Coastal California. Computers, Environment and Urban Systems 27, 527–547 (2003)
23. Forrer, H.R., Steenblock, T., Fried, P.M.: Monitoring of Potato Late Blight in Switzerland and Development of PhytoPRE+ 2000, an Internet Based Decision Support System. Journal of Agricultural University of Heibei 24, 38–43 (2001)
24. Hu, T.L., Zhang, Y.X., Wang, S.T., Yang, J.Y., Zhang, Y., Cao, K.Q.: Construction and Implementation of China-light: a Monitoring and Warning System on Potato Late Blight in China. Plant Protection 36, 106–111 (2010) (in Chinese)
25. Xiao, Z.Q., Li, Z.M., Fan, M., Zhang, Y., Ma, S.J.: Prediction Model on Stripe Rust Influence Extent of Winter Wheat in Longnan Mountain Area. Chinese Journal of Agrometeorology 28, 350–353 (2007) (in Chinese)
26. Cheng, H.X., Wang, C.M., Shuai, K.J., Li, Y.F., Song, J.F., Wang, J.M.: Meteorological Forecast Models of Wheat Pests in Jincheng City, Shanxi Province. Jiangsu Agricultural Sciences 6, 159–163 (2010) (in Chinese)
27. Liu, X.L.: Study on Prediction Model of Stripe Rust on Autumn Wheat Seedlings in Pingliang, Gansu Province. Gansu Agriculture Science and Technology 10, 37–38 (2010) (in Chinese)

Research on Semantic Text Mining Based on Domain Ontology

Lihua Jiang[1,2], Hong-bin Zhang[3], Xiaorong Yang[1,2], and Nengfu Xie[1,2]

[1] Agricultural Information Institute of Chinese Academy of Agricultural Sciences, Beijing, 100081
[2] Key Laboratory of Digital Agricultural Early-Warning Technology, Ministry of Agriculture, P.R. China, Beijing, 100081
[3] Institute of Agriculture Resources and Regional Planning of Chinese Academy of Agricultural Sciences, Beijing, 100081
Jianglh@caas.net.cn

Abstract. Text mining is an effective means of detecting potentially useful knowledge from large text documents. However conventional text mining technology cannot achieve high accuracy, because it cannot effectively make use of the semantic information of the text. Ontology provides theoretical basis and technical support for semantic information representation and organization. This paper improves the traditional text mining technology which cannot understand the text semantics. The author discusses the text mining methods based on domain ontology, and sets up domain ontology and database at first, then introduces the "concept-concept" correlation matrix and identifies the relationships of conceptions, and puts forward the text mining model based on domain ontology at last. Based on the semantic text mining model, the depth and accuracy of text mining is improved.

Keywords: ontology, domain ontology, text mining, semantic text mining.

1 Introduction

Natural language is the main means to exchange and express thoughts and ideas in social-economic living for people. Though study of natural language for a long time, the ability of interpretation is still limited. The existed technology has solved single sentence analysis, but it is hard to cover all language phenomenon, especially for the whole paragraph and chapter.

In the early nineteenth century, the data mining technology based on statistical technology had developed maturely, and application in the large-scale structural relational database achieved success. So people want to apply data mining technology to analyze natural language text and the method is called text mining or knowledge discovery in text. Difference with the traditional natural language processing words and sentences, the main goal of text mining is concentrated on founding hidden meaningful knowledge in mass text set, that is the understanding of text set and the relationship among texts. Now most of the text mining applications are lack of semantic level consideration, only in grammar level processing, so the obtained result is not pretty.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 336–343, 2013.

2 Text Mining Based on Domain Ontology

Text mining or knowledge discovery in text is the process of extracting unknown, useful and understandable pattern or knowledge in mass text data（Hearstma, 1997; Feldman R, 1995）. The research object is semi-structural or unstructured and natural language text contains multi-level ambiguities. So text mining has brought a lot of difficulties.

The traditional text mining method based on vector space model represented the text to lexical frequency vectors and mined the vectors (Mothe, 2001; Ghanem, 2005). The defect of this method is only dealing with document forms and ignoring the semantic role. This is the primary cause that traditional text mining can not live up to expectations. So text mining technology is needed to combine with semantic analysis to realize text mining in semantic level (Xin Xu, 2004; Chih-Ping Wei, 2008; Hui-Chuan Chu, 2009). Appling domain ontology to text mining provides theoretical support for semantic text mining and also provides a feasible technology way.

With the help of ontology to mine text sets, it amounts to a "domain expert' is equipped to the process of text mining to guide the whole process. According to features of text mining applications, ontology is divided into common ontology and domain ontology. Common ontology usually starts with the epistemology of the philosophy and extracts the relationship of general objects. The typical semantic dictionary established by common ontology is English WordNet and Chinese HowNet. At present, there are many text mining methods based on WorNet and HowNet (Rosso, 2004；Sedding J，2004；Raymond, 2000；Y. Ino, 2005；Shehata S, 2009). But the text mining methods based on common ontology are very difficult to achieve good results in some specific domains. Therefore, some researchers began to develop text mining study based on domain ontology. Bloehdom etc. put forward the OTTO frame（OnTology Based Text mining frame wOrk）(S. Bloehdorn, 2005).OTTO used text mining to learn the target ontology from text documents and used then the same target ontology in order to improve the effectiveness of both supervised and unsupervised text categorization approaches. The bag of words representation used for these clustering methods is often unsatisfied as it ignores relationships between important terms that do not co-occur literally. In order to deal with the problem, Hotho etc. integrate core ontology as background knowledge into the process of clustering text documents（Hotho A，2003）. Song etc. suggested an automated method for document classification using an ontology, which expressed terminology information and vocabulary contained in web documents by way of a hierarchical structure (Song Mh, 2005).

In our country, knowledge engineering lab of computer science department in Tsinghua university developed text mining platform based semantic web. And also there are some researchers who discussed applications of semantic processing technology in text mining. Xuling Zheng etc.proposed a corpus based method to automatically acquire semantic collocation rules from a Chinese phrase corpus, which was annotated with semantic knowledge according to HowNet（Zheng Xuling etc.，2007）. By establishing domain ontology as the way of knowledge organization, Guobing Zhou etc. introduced a novel information search model based on domain ontology in semantic context（Zou Guobing etc., 2009）. An Intelligent search method

based on domain ontology for the global web information was proposed to solve the problem of low efficiency typical in traditional search engineers based on word matched technology by Hengmin Zhu（Zhu Hengmin etc., 2010）. In order to improve the depth and accuracy of text mining, a semantic text mining model based on domain ontology was proposed by Yufeng zhang etc.. And in this model, semantic role labeling was applied to semantic analysis so that the semantic relations can be extracted accurately (Zhang Yufeng etc., 2011).

Taken together, the research of semantic text mining based on domain ontology is still in the domestic research theory spread stage, and relative actively in foreign countries. But there is few whole text mining based on domain ontology solutions results. And the research scope is only in foundation of shallow knowledge such as classification and clustering of text (Bingham,2001; Montes-y-Gómez, 2001)but rarely in rich useful deep semantic knowledge such as semantic association foundation(Zelikovitz,2004)、topic tracking(Aurora, 2007) and trend analysis (Pui Cheong Fung, 2003)and so on.

3 Key Technology of Text Mining Based on Domain Ontology

At present, Most of the ontology systems are almost the same in basic structure. Most of ontology is described for the entity, conception, generally properties and relationship. That is, by the way of some rules, the characteristic features and the corresponding parameters of the entities or conceptions are studied. At the same time, the relationship of entity and conception is described.

3.1 Knowledge Presentation Based on Domain Ontology

Knowledge presentation is the foundation of text mining. The quality of the knowledge representation is related to the efficiency of text mining. This paper uses WordNet and OWL to organize and present knowledge. This paper studies the following ideas: after data preprocessing, universal transformation format and conception extraction, domain ontology and ontology database are built by WordNet and OWL. Knowledge construction is defined and knowledge index is built in conceptual level.

In this paper, we take agriculture ontology as an example. Agriculture ontology is the system including agriculture terms, definition and standard relationship of terms. It is also formalize expression of conceptions and the relationship among conceptions. At the same time, it can not only deal with the inner relations among agriculture subject tale, but also more formal special relations. Formalize agriculture ontology is defined:

$$\text{Agri\_Onto}=(\text{Onto\_Info}, \text{Agri\_Concept}, \text{AgriCon\_Relation}, \text{Axion})$$

Therein, Onto_Info is the basic information of ontology including name, creator, design time, modification time, aim and knowledge resource and so on; Agri_Concept is the set of agriculture conceptions; AgriCon_Relation is the conceptual relation set including hierarchical relationship and the hierarchical relationships; Axion includes existing axiomatic set in ontology.

3.2 Construction Conceptual Semantic Correlation Matrix

Combined with the features of domain ontology and text mining, this paper proposes the text mining model based on domain ontology, and the main idea is as following: at first, "conception-conception" correlative matrix of domain ontology is constructed, and then key words in the collected documents are extracted in the light of domain ontology and the documents are represented to vector space model in order to calculate similarity between documents by reference to the "conception-conception" correlation matrix to realize document clustering. If new conception is found in the process of document clustering, the ontology database should be enlarged.

In domain ontology, there are vocabulary to represent classes and conceptions which are not only the bridge to communicate with classes and conceptions but also basic elements to represent classes. The relationships of domain ontology depend on the words to connect, so vocabulary is the key of construction domain ontology. In this study, the class vocabulary and conception vocabulary are extracted from text documents to make up vocabulary set $C=\{C_1,C_2,...,C_T\}$, in which T is the sum of conceptions in ontology. This paper uses matrix to structure representation method of conception relatedness, as follows:

$$R=\begin{bmatrix} R(C_1,C1) & R(C_1,C_2) & ... & R(C_1,C_T) \\ R(C_2,C_1) & R(C_2,C_2) & ... & R(C_2,C_T) \\ .. & ... & ... & ... \\ R(C_T,C_1) & R(C_T,C_2) & ... & R(C_T,C_T) \end{bmatrix}$$

In the matrix, $R(C_1,C_T)$ represents semantic relativity of C_1 and C_T. The result of $R(C_1,C_T)$ is semantic relativity of C_1 and C_T. A lot of scholars have discussed the conception relativity calculation. Generally speaking, there are two kinds of methods, information capacity method and the concept distance method. This paper adopts concept distance method, because in the ontology conceptions are arranged in tree. The method of concept distance can not only reduce complexity of algorithm but also easy to calculate. Before calculation the relativity, first of all to do the following description: ① if C_i and C_j is similar, the $R(C_i,C_j)=1$; ② if C_i and C_j is not similar, according document [25] this paper adopts the following formula to calculate semantic relativity:

$$R(C_i,C_j)=\frac{(Dist(C_i,C_j)+\alpha)*\alpha*(d(C_i)+d(C_j))}{CE(C_i,C_j)*2*Dep*\max(|d(C_i)-d(C_j)|,1)}$$

In the formula, $d(C_i)$ and $d(C_j)$ represent the layers that C_i and C_j located. $Dist(C_i,C_j)$ is the total weights in the shortest root from C_i to C_j in domain ontology tree. Dep is the max depth of ontology tree. α is a controllable parameter, generally more than equal to zero.

3.3 Identification of Relationship of Conceptions

The aim of text mining is to find the inherent, useful knowledge and implicit relationship. This paper introduces the identification mode to judge weather there is

relation $(C_i \cup C_j)$ between C_i and C_j.The method of this paper is to set a threshold to judge the relation between C_i and C_j. If $C_i \cup C_j$ accounts for high proportion that is C_i always accompany C_j, C_i and C_j inevitably has the relationship. If the value of $(C_i \cup C_j)$ is higher than threshold, there is correlation between C_i and C_j. At the same time, the implicit correlation is found.

4 Text Mining Model Based on Domain Ontology

4.1 System Frame

Design ideas in the following figure：

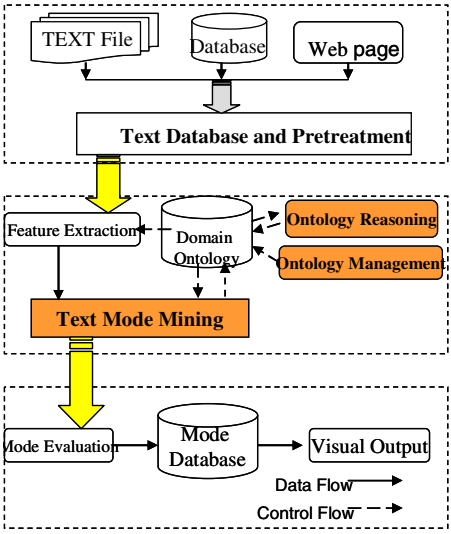


Fig. 1. Text Mining Model Based on Agriculture Ontology

4.2 Workflow of Text Mining Based on Domain Ontology

The text mining model based on domain ontology includes six parts.

(1) Ontology management

Ontology management is the core of the whole text mining model and provides semantic support. It mainly takes charge of building, storage, maintenance and optimization domain ontology. After comparison several ontology building methods, this paper adopts Protégé developed by USA Stanford University to set up domain ontology. The building of ontology is a process of accumulation and updating domain knowledge.

⑵ Text database and text data pretreatment

The text data source of text mining is unorganized text documents, including Web pages, files, Words and Excels, PDF documents, E-mails and so on. Before acquisition text information, the text data must be pretreated including data cleaning such as noise reduction and duplication removal, data selection, text segmentation such as Chinese Word Segmentation and Paragraphs segmentation.

⑶ Text information extraction

After pretreatment, the text data must be clean and then feature information must be extracted including word segmentation, feature representation. After feature extraction, the text data will be changed to text information. And the text information can be stored in formal of structured or semi-structured.

⑷ Text mode mining

Text model mining is based on text information and needs the support of system resources including domain ontology database, ontology reasoning and ontology management. The module of text model mining use semantic model mining arithmetic to mine deep semantic knowledge.

⑸ Ontology reasoning

Ant colony optimization is introduced, combined with searching algorithm to realize intelligent semantic reasoning and provides technological support for mining deep semantic mode. The effect of ontology reasoning is to reason mining mode and obtain deeper level mode to avoid common sense knowledge.

⑹ Evaluation and output

The knowledge obtained from mining module may be inconsistent, non-intuitive and difficult to understand. So it is necessary to post process text knowledge including knowledge valuation and accept or reject, elimination knowledge inconsistency.

5 Conclusion

In order to improve the depth and accuracy of the text mining, a semantic text mining model based on domain ontology is proposed. In this model, conceptual semantic correlation matrix is applied to semantic analysis so that the semantic relations can be extracted accurately. The text mining model based on domain ontology in this paper can mine deep semantic knowledge from text documents. The pattern got has great potential applications.

Acknowledgment. The work is supported by the special fund project for basic science research business Fee, AIIS “Tibet agricultural information personalized service system and demonstration” (No.2012-J-08) and The CAAS scientific and technological fund project “Research on 3G information terminal-based rural multimedia information service” (No.201219).

References

1. Hearstma: Text datamining Issues, techniques, and the relationship to information access. Presentation Notes for UW/MS Workshop on Data Mining (1997)
2. Feldman, R., Dagan, I.: Knowledge discovery in textual databases (KDT). In: Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD 1995), pp. 112–117. AAAI Press, Montreal (1995)
3. Mothe, J., Chrisment, C., Dkaki, T.: Information mining – use of the document dimensions to analyse interactively a document set. In: European Colloquium on Information Retrieval Research, pp. 6–20 (2001)
4. Ghanem, M., Chortaras, A., Guo, Y., Rowe, A., Ratcliffe, J.: A grid infrastructure for mixed bioinformatics data and text mining. Computer Systems and Applications 34(1), 116–130 (2005)
5. Xu, X., Cong, G., Ooi, B.C., Tan, K.-L., Tung, A.K.H.: Semantic Mining and Analysis of Gene Expression Data. In: Proceedings 2004 VLDB Conference[C], Morgan Kaufmann, St Louis, pp. 1261–1264 (2004)
6. Wei, C.-P., Yang, C.C., Lin, C.-M.: A Latent Semantic Indexing-based approach to multilingual document clustering. Decision Support Systems 45(3), 606–620 (2008)
7. Chu, H.-C., Chen, M.-Y., Chen, Y.-M.: A semantic-based approach to content abstraction and annotation for content management. Expert Systems with Applicaions 36(2), 2360–2376 (2009)
8. Rosso, P., Ferrettle, E., Jmenez, D., et al.: Text Categorization and Information Retrieval Using WordNet Senses. In: The Second Global Wordnet Conference, GWC 2004, Czech Republic (2004)
9. Sedding, J., Kazakov, D.: Wordnet-based text document clustering. In: Proceedings of the Third Workshop on Robust Methods in Analysis on Natural Language Data (ROMAND), Geneva, pp. 104–113 (2004)
10. Kosala, R., Blockeel, H.: Web Mining Research: A Survey. ACM SIGKDD (2), 1–15 (2000)
11. Ino, Y., Matsui, T., Ohwada, H.: Extracting Common Concepts from WordNet to Classify Documents. Artificial Intelligence and Applications, 656–661 (2005)
12. Shehata, S.: A Wordnet-based Semantic Model for Enhancing Text Clustering. In: 2009 IEEE International Conference on Data Mining Workshop, pp. 477–482 (2009)
13. Bloehdorn, S., Cimiano, P., Hotho, A., Staab, S.: An Ontology-based Framework for Text Mining. LDV-Forum 20(1) (2005)
14. Hotho, A., Staab, S., Stumme, C.: Wordnet improves text clustering. In: Proceedings of the Semantic Web Workshop at SIGIR-2003, 26th Annual International ACM SIGIR Conference (2003)
15. Song, M.H., Lm, S.Y., Park, S.B.: Ontology-based automatic Classification of Web Pages. International Journal of Lateral Computing 1(1) (2005)
16. Zheng, X., Zhou, C., Li, T.: Automatic Acquisition of Chinese Semantic Collocation Rules Based on Association Rule Mining Technique. Journal of Xiamen University (Natural Science) 46(3), 331–336 (2007)
17. Zou, G., Xiang, Y.: Information Search Model Based on Domain Ontology. Journal of Tongji University (Natural Science) 37(4), 545–549 (2009)
18. Zhu, H., Ma, J., Huang, W., Fan, H.: Study on Method of the Global Web Intelligent Search Based on Domain Ontology. Journal of the China Society for Scientific and Technical Information 29(1), 9–15 (2010)

19. Zhang, Y., He, C.: Research on Semantic Text Mining Based on Domain Ontology. Journal of the China Society for Scientific and Technical Information 30(8), 832–839 (2011)
20. Bingham, E.: Topic identification in dynamical text by extracting minimum complexity time components. In: Proceedings of ICA 2001, pp. 546–551 (2001)
21. Montes-y-Gómez, M., Gelbukh, A., López-López, A.: Discovering ephemeral associations among news topics. In: Proceedings of IJCAI 2001 Workshop on Adaptive Text Extraction and Mining, pp. 216–230 (2001)
22. Zelikovitz, S.: Transductive LSI for Short Text Classification Problems. In: Proceedings of the 17th International FLAIRS Conference. AAAI Press, Miami (2004)
23. Pons-Porrata, A., Berlanga-Llavori, R., Ruiz-Shulcloper, J.: Topic discovery based on text mining techniques. Information Proceeding and Management (43), 752–768 (2007)
24. Pui Cheong Fung, G., Xu Yu, J., Wai, L.: Stock prediction: Integrating text mining approach using real-time news. In: 2003 IEEE International Conference on Computational Intelligence for Financial Engineering, pp. 395–402 (2003)
25. Wu, J., Wu, Z., Li, Y.: Web Service Discovery Based on Ontology and Similarity of Words. Chinese Journal of Computer 28(4), 595–602 (2005)

The Research and Design of the Android-Based Facilities Environment Multifunction Remote Monitoring System*

Lutao Gao, Linnan Yang, Lin Peng, Yingjie Chen, and Yongzhou Yu

College of Basic Science & Information Engineering, Yunnan Agricultural University, 650201, Kunming, China
{Gaolutao_11,penglin2286351,yyzclc}@163.com, lny5400@sina.com, 24424448@qq.com

Abstract. For the actual demand of national facilities environment remote monitoring system combined with mobile Internet technology is cheaper, simpler to operate and better performance of mobility management. This article describes the research of the Android-based facilities environment multifunction remote monitoring system that based on the Android Smartphone as the terminal, was combined with wireless camera, relay group, wireless AP + temperature and humidity (light intensity) sensor and so on. The system implements remote facilities environment factors monitoring, real-time video monitoring, maintenance and management for remote server. The test and application shows that is stable, cheap, good mobility and easy to operate, it is a strong practicality and application prospects.

Keywords: Mobile Internet, Android Smartphone, Facilities Environment, Remote Monitoring.

1 Introduction

With the rapid development of Chinese agricultural facilities environment remote monitoring system is an important factor to improve facilities for agricultural production automation and efficient[1]. In recent years, the national agricultural engineering researchers used computer-controlled technology, web technology, GPRS and GSM technology[2,3] based on the PC (Personal Computer), PDA (Personal Digital Assistant) application terminal, develop and design a range of facilities environment remote monitoring and control systems[4,5], has played a positive role in promoting the development of Chinese agricultural facilities.

This year, with the rapid development of mobile Internet and networking technology continues to mature, especially the world's rapidly growing popularity of smart phones, the mobile network terminals application is more widely. In the present facilities of agricultural environmental monitoring system, there are lots of PDA

* Agriculture Science Technology Achievement Transformation Fund (2011GB2F300013).

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 344–351, 2013.

application terminal for greenhouse control system, but because the PDA is a single function, and development environment in the human-computer interaction, cross-platform performance is poor, limiting the effect of promotion and application of its facilities in agricultural environments. In recent years, the Android operating system and Android Smartphone has been rapid development, especially the launch of the Android 3G Smartphone fewer than 1000 Yuan (RMB), the Android Smartphone market share in Chinese rising [6]. In the one hand, the Android smart phone is the set of calls, multimedia, Internet smart terminal; it is not only inexpensive, but also multifunction[7]. The other hand, the Android operating system is open source and free, it not only reduces the system development costs, but also has a better human-computer interaction technology because of object-oriented Java language supporting. Therefore, Android Smartphone application terminal with mobile Internet technology has a positive meaning to build multi-purpose environmental monitoring system for Chinese agricultural facilities to further improve our facilities for agricultural automation and intelligent.

The article describes the system is based on Android technology, Socket technology and Java technology, group of wireless relay, wireless sensors and wireless cameras, over Android application terminals, research and development facility environment based on the Android multi-function remote monitoring system.

2 System Architecture Design

The article based on Android Smartphone application terminal for the real facilities environment monitoring system status, combines the working principle of wireless cameras and wireless sensors, and gives the design of system architecture. The figure of system architecture shows as Figure 1:

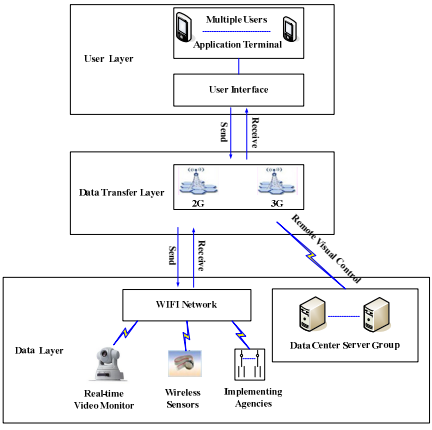


Fig. 1. System architecture

As Figure 1 shows, the system architecture contains user, data transfer and data layers. User layer is based on the Android Smartphone application terminal for sending and checking relevant instruction and information by human-computer interaction. Data transfer layer refers to the entire system operation that depends on network environment. The system is connected to the Internet by WIFI or 3G network for data transmission. Data layer contains video, environmental factors (temperature and humidity, light intensity) collection of information, such as, sun visor, water screen, heating and other facilities environmental site implementing agencies, as well as for data storage, centralized management of servers in the data center, and so on.

3 System Design and Implementation

The system based on the Android Smartphone application terminal to achieve the following three main features:

(1) Implementation remote real-time video monitoring of the facility agriculture;
(2) Combination of wireless relay group and wireless sensors such as temperature and humidity, light intensity, remote real-time monitoring to agricultural environment;
(3) Remote control of facility agriculture data center servers.
Core system class diagram is Figure 2:

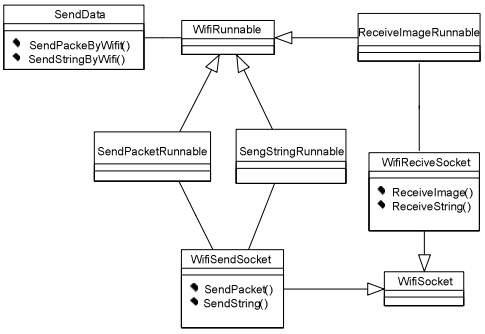


Fig. 2. System class diagram

As Figure 2 shows, This article achieves the SendData, SendPacketRunnable, SengStringRunnable, WifiSendSocket, WifiReciveSocket, ReceiveImageRunnable, and other core classes by overriding methods of the parent class and inheritance and overrides the appropriate methods. The core class diagram design for code optimization and integration of the system of development laid the Foundation important technologies.

3.1 System Design and Realization of Real-Time Video Monitoring

Aiming at the wireless camera command protocol, used Android technology and Socket Technology, The article Completed remote real-time video monitoring and control function based on Android Smartphone. The technical route process as shown in Figure 3:

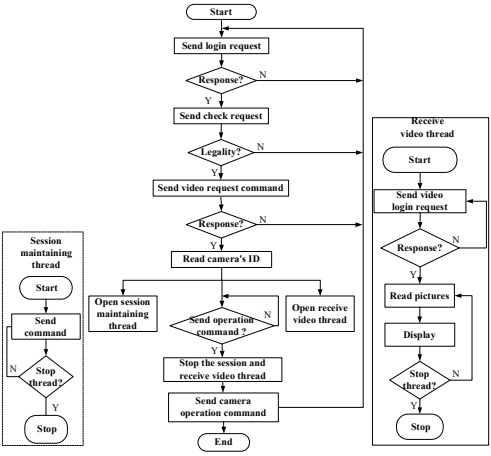


Fig. 3. Technical route flowchart

You can see from Figure 3, this opening session and video get the thread method, to achieve the functionality of video transmission and session continued. At the same time, this article also uses memory auto clean mechanism, thereby ensuring that the video transmission in the process, due to lack of memory of the Terminal and is causing the lockup issues.

3.2 Design and Implementation of System Facilities Environment Monitoring Module

The module to temperature humidity and light strength sensor for information collection Terminal, to wireless AP for frontal network access device, to wireless following electrical group for implementation body control device, and to the Android smart phone for application terminal, achieved remote real-time monitoring facilities environment within temperature humidity, illumination strength. It is able to through remote control unlimited following electrical group to achieve remote real-time control facilities environment system, its system implementation flowchart as Figure 4:

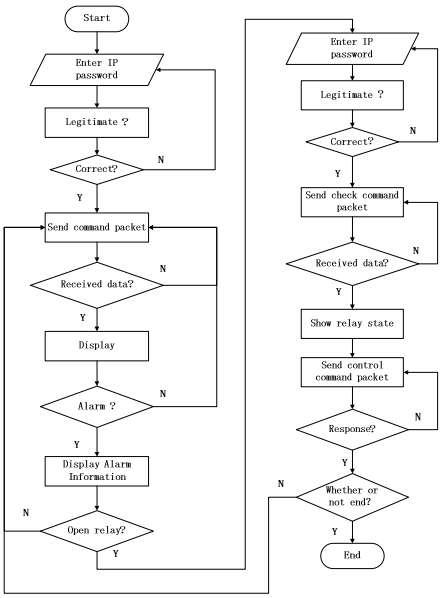


Fig. 4. Facilities environment monitoring system flowchart

As shown in Figure 3, the article point to the transfer protocol for wireless sensor to achieve the real-time data acquisition terminal for data transmitted to mobile terminals, automatic alarm when facilities environment exceeds a predetermined value. At the same time, using wireless relay group for weak electricity to force electricity controling. It achieves that mobile terminal directly control facilities environmental site visor, water screen, heater and functions of system.

3.3 Design and Implementation of Remote Server Control Module

This module is primarily in Smartphone application for terminal objects server-control. It achieves remote management features and maintains a data center server by Android mobile phone with coordinate positioning technologies and Socket like technologies. Its server control activity figure as shown in Figure 5:

The article point to the achievement of control server from Android mobile phone by Socket technology and Android technology. As the remote control of server activity figure shows, the module achieves following functions: When the user clicks the phone screen, a click command sent to the server. When the user double-clicks on a mobile phone screen, a double-click command sent to the server. When a user under the mobile phone screen, pop up the dialog box that you are able to choose right click, delete, copy, cut, paste operations. At the same time it is able to do function over PC such as IE browsing, file management, system management, and so on.

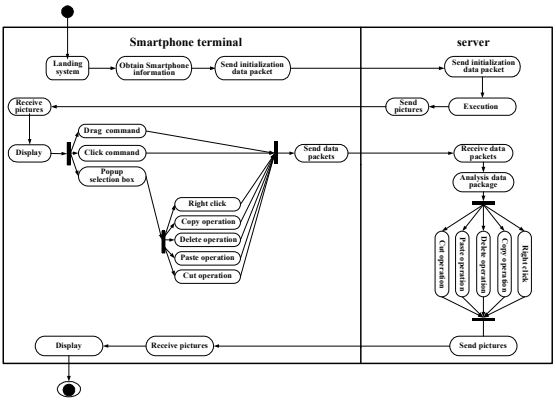


Fig. 5. Remote control of server activity

4 System Integration and Implementation

The system application program development environment is built over JDK +Eclipse +Android SDK+ADT. It support for the JAVA programming language[8]. It is better than QT application development platform based on C/C++ on security, human-computer interaction, and cross-platform[8,9].

In this development environment, using Android, Java and Socket technology, combine human-computer interaction technology with Android, to integrate and apply for facility environment multifunction remote monitoring system based on Android. The user operations flowchart as shown in Figure 6:

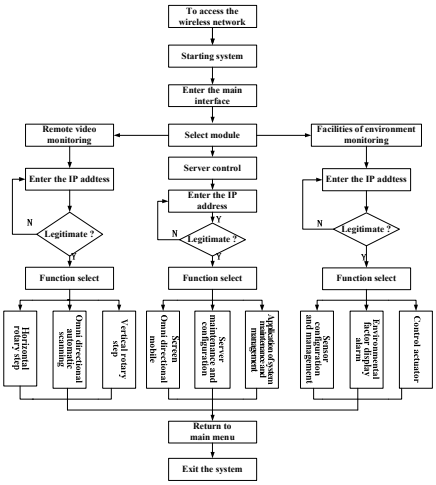


Fig. 6. System user operations flowchart

This system is applied by the agricultural science and technology achievements transformation Fund "Test and demonstration of Yunnan information service platform for Standardization of characteristic fruit and vegetable production ". The part of system interface as shown in Figure 7:

Fig. 7. Part of system interface

5 Discussion and Conclusions

The system uses the mobile Internet technology, to Android Smartphone application for the Terminal, combined with wireless data collection Terminal, using Android technology research and development "facility environment multifunction remote monitoring system based on Android", achieved through the Android Smartphone to remote facilities environment monitoring of video information, environmental factors, as well as the ability to remotely controlling, managing, and maintaining a data center server.

Display of system features and application effect:

(1) Using Android technology, based on the major mobile Internet devices – smart phones replace traditional PC, and applied to the facilities environment monitoring systems, reducing the cost of system development and investment;

(2) Compared with the single network applications such as PDA Terminal, Android smart phone applications more powerful, more large and universal application, have better market infrastructure and its application development environment based on JAVA language to make them better human-computer interaction and cross-platform performance;

(3) The system is feature-rich, secure and reliable, reducing labor costs, and good mobile performance by actual application in project. It gets very good application value.

References

1. Sun, Z., Cao, H., Li, H., et al.: GPRS and WEB based data acquisition system for greenhouse environment. Transactions of the CSAE 22(6), 131–134 (2006)
2. Zhou, G., Zhou, J., Miao, Y., et al.: Development and application on GSM based monitoring system for digital agriculture. Transactions of the CSAE 21(6), 87–91 (2005)
3. Yao, L.-B.: Design of Supervisory Control System for Greenhouse Based on Industrial Ethernet and Web Browse. Journal of Changzhou Vocational College of Information Technology 10(3), 23–25 (2011) (in Chinese with English abstract)
4. Hu, J., Yu, Y., Diang, M., Wang, W.: Technology of a fertilizing expert system PDA for crop growing. Transactions of the CSAE 22(8), 149–152 (2006) (in Chinese with English abstract)
5. Li, W., Lu, D.-X., Liu, C.-A.: Research on automation testing on Windows mobile-based device. Computer Engineering and Design 27(21), 4055–4057 (2006) (in Chinese with English abstract)
6. Hall, S.P., Anderson, E.: Operating Systems For Mobile Computing. Journal of Computing Sciences in Colleges Archive 25(2), 64–71 (2009) (in Chinese with English abstract)
7. Han, Y., Shan, X., Zhu, D., Zhang, X., Zhang, N., Zhan, Y., Li, L.: Design and Implementation of Locust Data Collecting System Based on Android. In: Zhou, M., Tan, H. (eds.) CSE 2011, Part II. CCIS, vol. 202, pp. 328–337. Springer, Heidelberg (2011)
8. Shang, M., Qin, L., Wang, F., et al.: Information collection system of wheat production risk based on Android smartphone. Transactions of the CSAE 27(5), 178–182 (2011) (in Chinese with English abstract)
9. Song, X., Zhou, D.: Development and Research of Application Based on Android Platform. Software Tribune 10(2), 104–106 (2011) (in Chinese with English abstract)

Application of the ARIMA Models in Drought Forecasting Using the Standardized Precipitation Index

Ping Han[1], Pengxin Wang[2,*], Miao Tian [2], Shuyu Zhang[3], Junming Liu[2], and Dehai Zhu[2]

[1] Department of Physics, College of Science, China Agricultural University, West Campus, Beijing 100193, P.R. China
hanrose@cau.edu.cn

[2] Department of Geographic Information Engineering, College of Information and Electrical Engineering, China Agricultural University, East Campus, Beijing 100083, P.R. China
wangpx@cau.edu.cn

[3] Remote Sensing Information Center for Agriculture of Shaanxi Province, Xi'an 710015, P.R. China
shuyuzhang88@sina.com

Abstract. The standardized precipitation index (SPI) was used to quantify the classification of drought in the Guanzhong Plain, China. The autoregressive integrated moving average (ARIMA) models were developed to fit and forecast the SPI series. Most of the selected ARIMA models are seasonal models (SARIMA). The forecast results show that the forecasting power of the ARIMA models increases with the increase of the time scales, and the ARIMA models are more powerful in short-term forecasting. Further study was made on the correlation coefficient between the actual SPIs and the predicted ones for the forecasting. It is shown that the ARIMA models can be used to forecast 1-month leading values of all SPI series, and 6-month leading values for SPI with time scales of 9, 12 and 24 months. Our study shows that the ARIMA models developed in the Guanzhong Plain can be effectively used in drought forecasting.

Keywords: drought forecasting, standardized precipitation index, autoregressive integrated moving average model.

1 Introduction

Drought is a slow-onset and creeping natural hazard that occurs in all regions of the world. Prolonged multiyear drought has caused significant damages in natural environment as well as in human lives. The Estimation for the cost of drought in the United States ranges from $6 to $8 billion annually [1]. In China, the amount of loss caused by drought ranks the first in all natural hazards. With the increase of the population and the severity of drought, an effective mitigation of the impacts caused by drought is imperative.

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 352–358, 2013.

In 1965, Palmer presented a drought index that incorporated antecedent precipitation, moisture supply, and the pioneering evapotranspiration [2]. McKee et al. developed the SPI as an alternative to the Palmer Index for Colorado [3]. The SPI is considered to have several advantages over the PDSI [3,4,5]. The first advantage is that the SPI is based only on precipitation [3]. Due to this reason, the SPI is also not adversely affected by topography. Second, the SPI is calculated on various timescales, which allows it to describe the various types of droughts: the shorter time scales for meteorological and agricultural droughts, and the longer ones for hydrological drought. Third, because the SPI is normally distributed, the frequencies of drought events at any location for any time scale are consistent. Hayes et al. argued that the SPI detects moisture deficits more rapidly than the PDSI, which has a response time scale of approximately 8 to 12 months [5]. Paulo et al. used several drought indices in Portugal, and found that the SPI showed its superiority for the purpose of drought monitoring [6,7]. Labedzki used SPI to analyse the local meteorological drought and evaluate the drought risk in Bydgoszcz, Poland [8].

The time series forecasting has been widely applied and become an important approach of drought forecasting. One of the most widely used time series model is the autoregressive integrated moving average (ARIMA) model [9]. The wide application of the ARIMA model in many areas is due to its flexibility and systematic search (identification, estimation and diagnostic check) in each stage for an appropriate model [10]. The ARIMA model has several advantages over other approaches, such as moving average, exponential smoothing, neural network, and in particular, its forecasting capability and its richer information on time-related changes [11, 12]. The ARIMA models have also been used to analyze and model hydrologic time series [11,13,14]. Fernandz et al. used SARIMA model to forecast stream-flow in a small watershed in Galicia [15]. Durdu developed linear stochastic models for forecasting droughts in Turkey using SPI series as drought index [16].

The Guanzhong Plain is located in the northwest of China. This area is subjected to water stress and drought conditions. The frequency of drought is on average about once in 7 years [17]. The Plain has flat terra and fertile soil, and is the political and economical center of Shaanxi Province. Drought forecasting for this area can help to mitigate the effects of drought, and is in favor of effective water resource management. In this paper, The SPI is used as a drought index to describe the drought condition of the Guanzhong Plain. The SPI time series of multiple time scales in the Guanzhong Plain are calculated. The ARIMA models are applied to simulate and forecast the SPI series.

2 SPI Time Series Forecasting Models

The SPI time series of multiple time scales can be computed (typically 3, 6, 9, 12 and 24 months) according to the McKee's method[4]. The classification of dry and wet spells resulting from the values of the SPI is shown in table 1.

Table 1. Drought classification of SPI

SPI value	Class
>=2	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1 to -1.49	Moderately dry
-1.5 to -1.99	Severely dry
<=-2	Extremely dry

The SPI data set from 1966 to 2003 is used for model development. The data set of 2004 is used for model validation. Based on the steps for developing the ARIMA models [18,19], the ARIMA models fit for SPI3, SPI9, SPI12 and SPI24 are identified respectively. They are SPI3, MA(2); SPI6, ARIMA(1,0,0)$(0,0,1)_6$; SPI9, ARIMA(1,0,0)$(0,0,1)_9$; SPI12, ARIMA(1,0,0)$(0,0,1)_{12}$; and SPI24, ARIMA(0,1,0)$(2,1,1)_{24}$. It is observed that most of the SPI time series have the seasonal features. As the time scale increases, the seasonal feature is more and more distinct, and the series need to be seasonally differenced.

The forecast is done for 12 leading months using the selected models, i.e., forecast the SPI values in 2004. By comparing the predicted data with the original data, the forecasting capability of the models is discussed.

3 Results and Discussion

SPI3, SPI6, SPI9, SPI12 and SPI24 are respectively fitted by the selected best models from historical data. The forecast is done with 12-month lead-time. The values of SPI in 2004 are predicted. Fig. 1 shows the plots of fitting (forecasts for 1966-2003) and forecasting (forecasts for 2004) using the chosen ARIMA models, in which only the values of SPI in 1990-2004 are displayed for clarity.

Observing the fitting figure, how well the chosen model fits the data series is indirectly shown. In Fig. 1, it is obvious that the fitted data follow the original data very closely, especially for higher SPI series (SPI12 and SPI24). This indicates that the chosen best ARIMA models capture the patterns of the SPI series. However, the models that fit the data well do not all have good forecasting capability. The predicting power of the ARIMA models should be valuated. The forecast with 12-month lead-time are analyzed. It is found that the predicted SPI values fit the original ones better and better with the increase of the time scale in Fig. 1. The values of SPI3 predicted for 2004 tend to be constant. However, this situation is changed in SPI6, SPI9, SPI12 and SPI24. The predicted data gradually tends to present the stochastic change of the SPI series. It indicates that the forecasting power of the ARIMA models is improved as the time scale increases.

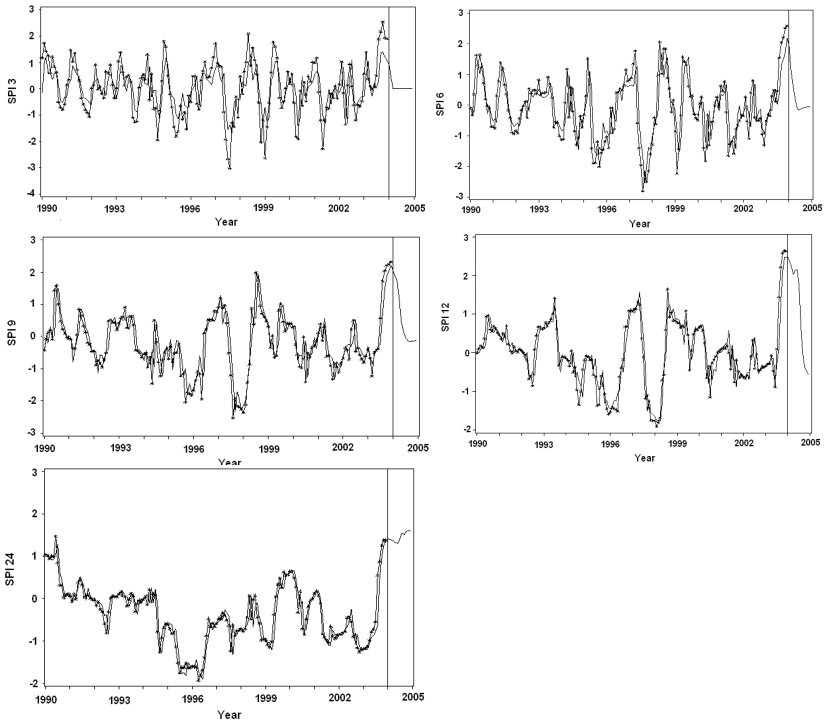


Fig. 1. The SPI series, the fitted series and the forecasts for 12 months ahead (1990-2004). The curve with stars is for the calculated SPI. The one without stars is for fitting and forecasting. The predicted values of SPI for 2004 are on the right side of the reference line.

By comparing the calculated values with the predicted ones, the absolute percentage errors (APE, $\left|(X_i - \hat{X}_i)/X_i\right| \times 100$) are calculated as an analysis on the forecasting power, as shown in Table 2. It is obvious that all SPI series have less APEs on 1-month ahead forecasting. The mean absolute percentage error (MAPE) is 5.8. This demonstrates that the ARIMA models have better result on short-term forecasting. The APEs seem to be less for higher SPI series (SPI12 and SPI24). For 1-month ahead, the APEs of all SPI series decrease with the increase of the time scale. For 1-9 month ahead, the APEs of SPI12 are all less than 6, and the APEs of SPI24 are generally less than 10, with only two exceptions (13.3 and 17.4). These results indicate that if the SPI series have a longer time scale, the ARIMA models will have a better forecasting power.

In order to exactly determine how good the forecasting power of the ARIMA models is, the forecasts are done with 1-month to 8-month lead-time. For example, 2-month lead-time forecast means that during March 2004 the forecast for May 2004 is

Table 2. Absolute percentage errors of SPI3, SPI6, SPI9, SPI12 and SPI24 for 12-month lead-time. Only values less than 20 are shown.

SPI series	One month	Two months	Three months	Four months	Five months	Six months	Seven months	Eight months	Nine months	Ten months	Eleven months	Twelve months
SPI3	8.1											
SPI6	8.4	15.1										
SPI9	8.9	14.9										
SPI12	2.2	0.9	5.5	0.9	0.4	5.7	5.4	1.8	5.0	13.2		
SPI24	1.4	4.1	13.3	8.7	3.9	17.4	2.9	7.6	2.5	9.7	8.0	5.3

performed. The correlation coefficients between the actual SPIs and the predicted ones from the ARIMA models are used as the criteria to evaluate the fit. Table 3 provides the coefficients for 1-8 month lead-time. It is observed that with more and more longer lead-time the coefficient decreases. To a given lead-time, the coefficient gradually increases with the increasing of the time scale, which shows that the ARIMA model is an appropriate model for the SPI series, and its power of forecasting is better for higher SPI series (SPI9, SPI12 and SPI24). Evidently, the ARIMA models give good forecast results for 1-month lead-time with the correlation coefficients ranging from 0.873 to 0.980. For high SPI series (SPI9, SPI12 and SPI24), the forecasts have good results up to 6-month lead-time with the correlation coefficients larger than 0.7. Therefore, the selected ARIMA models have a better power of forecasting for the SPI series. The ARIMA models can be used to forecast the change of the series of SPI3 and SPI6 one to two months ahead, and the future change of the series of SPI9, SPI12 and SPI24, six months ahead. The ARIMA model therefore exhibits a strong forecasting capability.

Table 3. Correlation coefficients of different lead-time. Blank cells are for the non-existing correlation coefficients.

SPI series	One-month lead-time	Two-month lead-time	Three-month lead-time	Four-month lead-time	Five-month lead-time	Six-month lead-time	Seven-month lead-time	Eight-month lead-time
SPI3	0.873	0.417	0.360					
SPI6	0.888	0.762	0.673	0.580	0.514	0.248		
SPI9	0.955	0.889	0.830	0.786	0.772	0.803	0.695	0.523
SPI12	0.989	0.985	0.985	0.985	0.935	0.742	0.538	0.388
SPI24	0.980	0.885	0.876	0.852	0.821	0.775	0.523	0.475

4 Conclusions

The Guanzhong Plain is one of the areas subjected to drought. In this study, the ARIMA models are developed to forecast drought using the SPI as the drought index. The results show that the selected ARIMA models are appropriate for the SPI series. Exactly the same as in many other applications[10,14], the ARIMA models demonstrate better capability in short-term forecasting. The absolute percentage errors (APE) between the actual SPIs and the predicted ones for 12-month lead-time are all less on 1-month ahead forecasting. The correlation coefficients between the actual data and the predicted ones are all larger on 1-2 month lead-time forecasting. Moreover, in this study, the ARIMA models also demonstrate a better forecasting power even on 6 month leading values for SPI9, SPI12 and SPI24. The study also shows that the ARIMA models have a good forecasting capability for the SPI series with longer time scales. The APEs are less for higher SPI series (SPI12 and SPI24). To a given lead-time, the correlation coefficient gradually increases with the time scale increasing. This may be because the increase in the length of the time scale effectively reduces the noise of the SPI series. Therefore, the selected best ARIMA models developed from the SPI time series can be used for drought forecasting in the Guanzhong Plain.

This study contributes to the exploration of a feasible way on drought forecasting in the Plain. The results demonstrate that the developed ARIMA models have a good forecasting power and can be used for drought forecasting. Further study should be focused on improving the precision of the model forecasting, and on studying the types of droughts described by the SPI series with different time scales, which will effectively assist the local authorities to mitigate the impacts caused by drought, and to reasonably use the water resources.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (Grant No. 41071235 and 40871159), the China National Support Program of Science and Technology (Grant No. 2012BAH29B03), and China Agricultural University's Grant, 2012JC004.

References

1. Schubert, S., Koster, R., Hoerling, M.: Predicting Drought on Seasonal-to-decadal Time Scales. Bulletin of the American Meteorological Society 88, 1625–1630 (2007)
2. Palmer, W.C.: Meteorological Drought. Research Paper No. 45, 58 p. US Department of Commerce Weather Bureau, Washington, DC (1965)
3. McKee, T.B., Doesken, N.J., Kleist, J.: The Relationship of Drought Frequency and Duration to Time Scales. In: Eighth Conference on Applied Climatology, pp. 179–184. American Meteorological Society, Anaheim (1993)
4. Guttman, N.B.: Comparing the Palmer Drought Severity Index and the Standardized Precipitation Index. Journal of the American Water Resources Association 34, 113–121 (1998)

5. Hayes, M.J., Svoboda, M., Wilhite, D.A., Vanyarkho, O.: Monitoring the 1996 Drought Using the SPI. Bulletin of the American Meteorological Society 80, 429–438 (1999)
6. Paulo, A.A., Pereira, L.S., Matias, P.G.: Analysis of Local and Regional Droughts in Southern Portugal Using the Theory of Runs and the Standardized Precipitation Index. In: Rossi, G., Cancelliere, A., Pereira, L.S., Oweis, T., Shatanawi, M., Zairi, A. (eds.) Tools for Drought Mitigation in Mediterranean Regions, pp. 55–78. Kluwer, Dordrecht (2003)
7. Paulo, A.A., Pereira, L.S.: Drought Concepts and Characterization, Comparing Drought Indices Applied at Local and Regional Scales. Water International 31, 37–49 (2006)
8. Labedzki, L.: Estimation of Local Drought Frequency in Central Poland Using the Standardized Precipitation Index. Irrigation Drainage 56, 67–77 (2007)
9. Box, G.E.P., Jenkins, G.M.: Time Series Analysis Forecasting and Control, 525 p. Holden-Day, San Francisco (1976)
10. Zhang, G.P.: Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Neurocomputing 50, 159–175 (2003)
11. Yurekli, K., Kurunc, K., Ozturk, F.: Application of Linear Stochastic Models to Monthly Flow Data of Kelkit Stream. Ecological Modeling 183, 67–75 (2005)
12. Hu, W.B., Tong, S.L., Mengersen, K., Connell, D.: Weather Variability and the Incidence of Cryptosporidiosis: Comparison of Time Series Poisson Regression and SARIMA Models. Annals Epidemiology 17, 679–688 (2007)
13. Fernando, D.A.K., Jayawardena, A.W.: Generation and Forecasting of Monsoon Rainfall Data. In: Proceedings of the 20th WEDC Conference, Colombo, Sri Lanka, pp. 310–313 (1994)
14. Modarres, R.: Streamflow Drought Time Series Forecasting. Stochastic Environmental Research and Risk Assessment 21, 223–233 (2006)
15. Fernandez, C., Vega, J.A., Fonturbel, T., Jimenez, E.: Streamflow Drought Time Series Forecasting: a Case Study in a Small Watershed in North West Spain. Stochastic Environmental Research and Risk Assessment 23, 1063–1070 (2008)
16. Durdu, O.F.: Application of Linear Stochastic Models for Drought Forecasting in the Buyuk Menderes River Basin, Western Turkey. Stochastic Environmental Research and Risk Assessment 24, 1145–1162 (2010)
17. Editing Committee on Brief Actual Records of Historical Natural Disasters for Shaanxi: Brief Actual Records of Historical Natural Disasters for Shaanxi, 271 p. China Meteorological Press, Beijing (2002) (in Chinese)
18. Bowerman, B.L., O'Connell, R.T.: Forecasting and Time Series: an Applied Approach, 521 p. China Machine Press, Beijing (2003)
19. Box, E.P., Jenkins, G.M., Reinsel, G.C.: Time Series Analysis: Forecasting and Control, 3rd edn., 327 p. Post & Telecommuncations Press, Beijing (2005)

Study of Cluster Formation Algorithm for Aquaculture WSN Based on Cross-Layer Design

Xufeng Hua[1,2], Chengxun Chen[2], Yunchen Tian[1,2], Yongjun Guo[2], and Kezhi Xing[2]

[1] Department of Computer Sciences and Information Engineering, Tianjin Agricultural University, 300384 Tianjin, China
[2] Tianjin Key Laboratory of Aquatic Ecology and Aquaculture, Tianjin Agricultural University, 300384 Tianjin, China
huaxufeng@163.com

Abstract. Aquaculture WSN is composed of a large number of nodes with different monitoring function. In this paper, we present a cluster formation algorithm in Sensor MAC for event-driven Aquaculture WSN. The technique proposed is based on cross-layer design which is adopted to reduce the energy waste and guarantee the data transmissions. The event-driven Aquaculture WSN is simulated by OPNET and MATLAB. Simulation results show that the proposed protocol saves node energy, shortens average packet latency, and improves event detection reliability.

Keywords: Cluster Formation Algorithm, Cross-Layer Design, Event-Driven, Aquaculture WSN.

1 Introduction

A large number of water quality monitoring nodes with low-power radio compose Wireless Sensor Network (WSN) for Aquaculture. In event-driven Aquaculture WSN, such as water quality changing, aquatic organisms tracking and water pollution detection, monitoring nodes only send little condition information to sink during usual time and send emergent information when event occurs.

Plenty of redundant information is sent by monitoring nodes deployed in Aquaculture WSN when specific event occurs. Topology features and data fusion must be considered in protocol design. Routing protocol is responsible for forwarding messages from source nodes to sink, which greatly influences the performance of data fusion, packets delay and node energy consumption. Meanwhile, Medium Access Control (MAC) layer directly controls the actions of wireless communication module which are regarded as the main part of energy dissipation [1]. Therefore, it is important that design of Sensor MAC (SMAC) protocol to the lifetime of network.

Low Energy Adaptive Clustering Hierarchy (LEACH) [2] applies TDMA within a cluster. There is a cluster head among each cluster. Instead of transmitting the data to the sink directly, the sensors send their data to the cluster-head. The cluster head relays the data to the sink node after data aggregation. LEACH assumes all nodes

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 359–365, 2013.

have data to transmit at all times, which did not consider the unique features of event, such as event severity level and occurrence position. EDC algorithm [3] and event driven data reporting algorithm [4] have been proposed for event-driven WSN. However, little consideration has been given to the features of regular status when sensor nodes only send little condition information to sink.

In schedule-based MAC protocol, TDMA is an important approach that is inherently collision free and avoids unnecessary idle listening. An intra-cluster communication bitmap-assisted (BMA) MAC protocol [5] is proposed for large-scale cluster-based WSN. Energy efficient TDMA [6] allows nodes with no data to transmit keep their radios off during their allocated time slots to reduce energy consumption. But when nodes have little information to transmit, it will achieve lower channel utility.

SMAC [7] is a contention-based MAC protocol especially designed for WSN. It forces sensor nodes to operate at low duty cycle by putting them into periodic sleep. It uses RTS/CTS/DATA/ACK to reduce collision and overhearing. However, packet collision problem will become severe when data traffic is heavy.

2 Details of Cluster Formation Algorithm

A cluster formation algorithm in Sensor MAC is presented in this paper for event-driven Aquaculture WSN, by which two network statuses: regular status and event status are defined. On regular status, no specific event happens and monitoring nodes only send normal detection information to report the usual state of Aquaculture WSN. Monitoring nodes energy consumption on regular status should be as low as possible. On event status, several monitoring nodes detect a specific event in the Aquaculture WSN, and the event must be transmitted to the sink node promptly. In order to reduce the energy waste, the techniques are adopted on regular status, in proposed event-driven cluster formation algorithm. In order to guarantee the event is efficiently transmitted to sink, on event status, an event-driven cluster formation algorithm is proposed according to the cross-layer theory.

When no water quality variation event happens, the approaches in SMAC protocol are adopted to reduce energy consumption. To save energy, monitor nodes go to sleep periodically with a low duty cycle. Using the mechanism of RTS/CTS/DATA/ACK, water quality monitor information of the normal network state is directly transmitted to sink, to reduce collision and overhearing.

When water quality variation event happens, all monitor nodes which have sensed the event are required to form into a cluster. Only one cluster head is elected according to the remaining energy level of the nodes in the cluster. All non-cluster head nodes transmit their water quality monitoring data to the cluster head using energy efficient TDMA approach. Using RTS/CTS/DATA/ACK technique, the cluster head node performs data fusion and transmits data to the remote sink. In the process of cluster formation, the proposed event-based cluster formation algorithm can guarantee that the event is efficiently transmitted to sink with low energy and delay.

Description of the event-driven cluster formation algorithm is:

In route layer, water quality monitor nodes calculate data arrival rate. If the data arrival rate is higher than threshold T_a, which indicates that a water quality variation event happens, monitor nodes will send cluster formation indication message to MAC layer to change the mechanism at MAC layer and perform cluster formation at route layer.

The operation of the event-driven cluster formation algorithm is divided into three phases: cluster set-up, data-transmission and cluster-dissolution. As shown in Fig.1., the data-transmission phase consists of several TDMA frames which include intra-cluster transmission and cluster head to sink transmission.

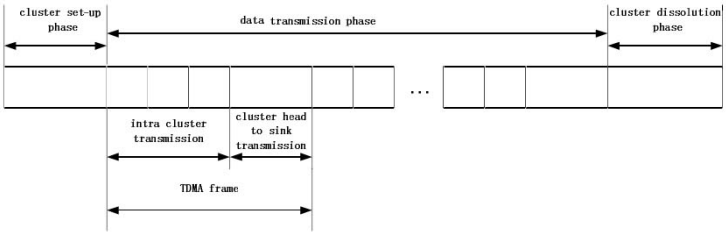


Fig. 1. Operation of event-driven cluster formation algorithm

Steps of cluster set-up phase are:

- Step 1: Monitor nodes that have sensed the water quality variation event broadcast a HELLO message when the next listen period comes. This message contains the node's ID and remaining energy. From now on, these nodes will not follow the sleep/wake-up schedule.
- Step 2: By comparing the remaining energy, the node with the minimum remaining energy will elect itself to be the cluster head. It must be sure that there is only one cluster head in this cluster.
- Step 3: The cluster head node broadcasts an advertisement message (ADV_CH) to let all the other nodes know that it has chosen this role.
- Step 4: Each non-cluster node informs the cluster head node with a join-request message (JOIN_REQ).
- Step 5: The cluster head node sets up a TDMA schedule and transmits it to the nodes in the cluster.
- Step 6: After the TDMA schedule is known by all nodes in the cluster, the cluster set-up phase ends.

When receiving any overhead packets in cluster set-up phase, the nodes that didn't detect the water quality variation event will set their NAV and go to sleep for an estimation time.

The data-transmission phase consists of several frames, where nodes send their data to the cluster head at most once per frame during their allocated transmission slot, and also the radio components of each node are allowed to be turned off at all

times except their own transmission duration to reduce the energy waste. Once the cluster head receives all the data, it performs data aggregation. In our analysis, we assume that all individual signals can be combined into a single representative signal. The resultant data are sent from the cluster head to the sink using RTS/CTS/DATA/ACK approach which will guarantee the transmission. As mentioned above, only one cluster was formed where the event happened in event-based cluster formation algorithm. Therefore, the approach of direct sequence spread spectrum (DSSS) in LEACH is not necessary to be used in our proposed protocol, which reduces the complexity of protocol.

We assume the event ends and the cluster should be dissolved, when the cluster head node receives no data in one frame. The cluster head broadcasts a cluster over message to dissolve the cluster. When the non-cluster nodes receive the OVER message, they return to their periodic sleep/wake-up schedule as on regular status.

3 Simulation and Analysis

The simulations are performed on OPNET and MATLAB, and 50 nodes are placed in a 100m×100m square region with no mobility. Each monitor node can communicate with sink directly. Three power consumption levels are set: 0.03 mW for sleep, 30 mW for receiving and idle listening, and 81 mW for transmitting. On regular status the packets interval time of CBR traffics is 10s. When the packets interval time is smaller than 1.432s, which is the time length of one sleep/wake-up period in SMAC protocol with 10% as its duty cycle, a water quality variation event is assumed to occur and the Aquaculture WSN turns into event status. In simulation, a water quality variation event is triggered in the middle of the topology at simulation time of 100s.

The simulation parameters of event-driven Aquaculture WAN are:

- Band-Width: 20Kbps
- Interval for Regular: 10s
- Interval for Event: 0.1s
- Tx Range: 250m
- Carrier Sensing Range: 550m
- Event Packets Number: 20
- Initial Energy: 100J

The metrics to evaluate the MAC protocols are: event get ratio, packets end to end delay, network energy consumption and network remaining energy. In event-driven cluster formation algorithm, Event get ratio (EGR) is defined as $(P_r \times D)/P_t$. In which P_r is the total received packets, P_t is the total transmitted packets when a water quality variation event occurs, D is the data aggregation degree, N_n is the number of nodes which sensed the event. We set D decreases with the increase of N_n. EGR is defined to indicate the delivery ratio of a water quality variation event in the Aquatic WSN. Packets end to end delay is the average delay experienced by a message from source node to sink, which indicates the real-time performance of the protocol. Network energy consumption is the total energy consumption among sensor nodes, which

represents the network lifetime. Network remaining energy is the remaining energy of the network at different simulation time, which indicates the energy consumption rate.

The nodes number (N_n) represents the influential range of the specific event and it also indicates the sensing range of an individual sensor node. In our simulation, N_n is changed from 3 to 30.

As illustrated in Fig.2., a large number of packets collision and lost in SMAC for its contention mechanism, which means the event can not be successfully delivered to sink. By using TDMA approach, event-driven cluster formation algorithm gains a significant improvement of EGR. The results show that while the number of nodes, N_n, increases from 3 to 30, packets delay increases for both protocols in Fig.3. The event packets can be sent to sink promptly for event-driven cluster formation algorithm, which has resolved the long delay problem in SMAC.

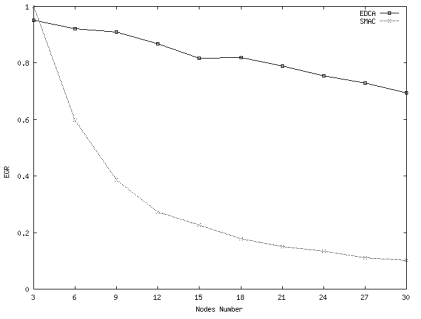


Fig. 2. EGR

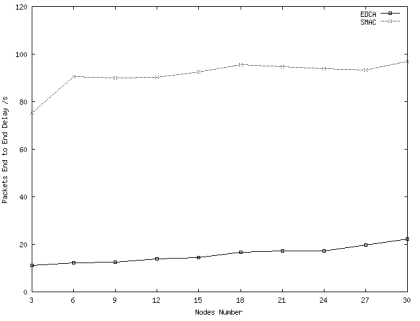


Fig. 3. Packets end to end delay

As illustrated in Fig.4., network energy consumption in SMAC protocol is the same with different number of nodes. It is mainly due to the fixed duty cycle in SMAC. It also can be seen that the larger of N_n, the more energy will be cost in event-driven cluster formation algorithm for more nodes will join in the cluster formation. Although there is cluster formation overhead in event-driven cluster formation algorithm, energy-TDMA approach is adopted which will reduce packet collisions and the energy consumption effectively.

Network remaining energy with simulation time after the event happens is shown in Fig.5. When an event occurs, event-driven cluster formation algorithm allows the nodes that have sensed the event to form into a cluster at once, therefore the energy is consumed faster. With the increasing of the nodes, the trend of rapid energy consumption becomes longer. When the cluster set-up phase ends, energy consumption becomes slower for the energy-efficient TDMA approach. Nodes always have data to transmit in each wake-up period in SMAC protocol after an event happened, therefore, the energy is consumed rapidly and the situation can not turn better.

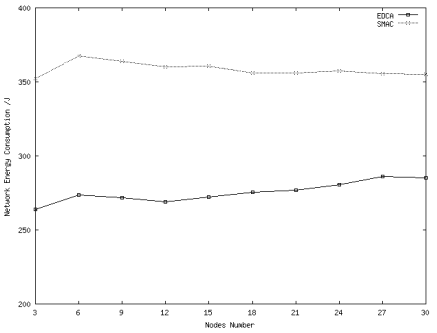


Fig. 4. Aquaculture WSN energy consumption

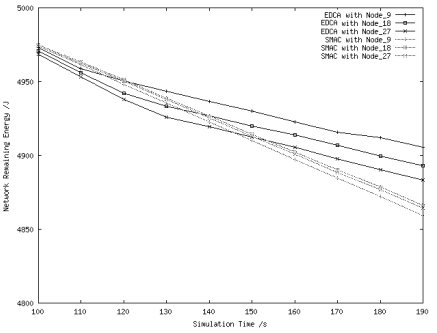


Fig. 5. Aquaculture WSN remaining energy

4 Conclusion

Event-driven Aquaculture WSN is an importance application of Aquaculture of Internet of things. Based on the features of event-driven Aquaculture WSN, a cross-layer designed cluster formation algorithm is proposed. The basic techniques in SMAC protocol are adopted for regular status, while new cluster formation algorithm is proposed for event status. The simulation experiments demonstrate event-driven static distributed Aquaculture WSN. The results show that the algorithm proposed achieves good performance on the water quality variation event detection reliability, average packets latency and monitor node energy consumption.

Acknowledgments. This research is jointly sponsored by Special Fund for Agro-Scientific Research in the Public Interest (No. 201203017), Key Projects in Tianjin Science and Technology Pillar Program (No. 11ZCKFNC00400), National Spark Planning Project (No. 2011GA610009).

References

1. Georgios, Y.L., Jing, L., Joseph, P.: A Cluster-Based Power-Efficient MAC Scheme for Event-Driven Sensing Applications. Ad Hoc Networks 5, 1017–1030 (2010)
2. Jamieson, K., Balakrishnan, H., Tay, Y.: A MAC Protocol for Event-Driver Wireless Sensor Networks. J. Wireless Sensor Networks 3868, 260–275 (2009)
3. Ilker, D., Cem, E.: MAC Protocols for Wireless Sensor Networks. IEEE Communications Magazine 44, 115–121 (2008)
4. Tashtarian, F., Tolou-honary, M.: An Energy Efficient Data Reporting Scheme for Wireless Sensor Networks. In: IEEE/ACS International Conference on Computer Systems and Applications, pp. 223–228. IEEE Press, New York (2010)

5. Zhiwei, Z., Xinming, Z., Peng, S.: A Transmission Power Control MAC Protocol for Wireless Sensor Networks. In: Proceedings of the Sixth International Conference on Networking, pp. 5–15. Scientific Research Publishing, New York (2011)
6. Venkatesh, R., Katia, O., Garcia, J.: Energy-Efficient Collision-Free Medium Access Control for Wireless Sensor Networks. Wireless Networks 12, 63–78 (2009)
7. Tuirkmen, C., Farid, N.: A Cross-Layer Optimization Approach for Efficient Data Gathering in Wireless Sensor Networks. In: IEEE International Networking and Communications Conference, pp. 101–106. IEEE Press, New York (2010)

The Survey of Fishery Resources and Spatial Distribution Using DIDSON Imaging Sonar Data*

Wei Shen[1], Long Yang[2], Jin Zhang[1,**], and Guangxiong Peng[3]

[1] College of Marine Sciences, Shanghai Ocean University, Shanghai, 201306, China
j_zhang@shou.edu.cn
[2] The First Institute of Oceanography, SOA, Qingdao, 266061, China
[3] School of Geosciences and Info-Physics, Central South University, Changsha, 410083, China

Abstract. In recent years, the DIDSON (Dual-frequency IDentification SONar),which can provide almost-video-quality images to identify objects even in turbid water has been used in enumerating fish populations, underwater structures inspection, oil/gas leakage detection and identification, underwater security, evidence searching, ship's hulls and ports safety inspection, and underwater navigation and so on.

In this paper, a designed vessel-mounted observing systems collected DIDSON data and GPS data according to designed lines in the Dishui Lake, shanghai city. Every line area was calculated according to the depth and sonar open-angle. Trained observers counted manually each fish image in every DIDSON file. Then based on the area and number of fish observed, the average density and zoning density of the Dishui lake was calculated. At last the whole fish number and distribution was calculated based on the ARCGIS software.

The practice proved the method based on DIDSON Data is very feasible, effective and accurate for fishery resources estimating.

Keywords: DIDSON, Survey, Fishery Resources, Spatial Distribution.

1 Introduction

The Dishui Lake of Shanghai Lingang Town was completed on 2003. The area of whole lake is 5.56 square kilometers, the average depth is 3.7 m, and the most deep depth is 6.2 m. Since 2007, the management of the lake has put about 315 ton fry into the lake so far, the fry were mainly silver carp and bighead carp. At present, how much fish in the lake is a question. So the lakes management decides to detect the fishery resources for the further catching and management.

In the water , sonar equipment is the main observation and investigation equipment[1]. Acoustic fisheries resources survey is fast and efficient investigation of biological resources, and no damage for fisheries[1-3].

* Supported by the Innovation Program of Shanghai Municipal Education Commission(12ZZ159), and China Postdoctoral Science Foundation (2011M501296, 2012T50832).
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 366–375, 2013.

Currently, the main advanced acoustic instruments for fishery resources survey are Simrad EY-60[4], Bisonics DT-X[5], and etc.. These instruments generally utilize echo integration method[6-8] to assess the total amount of resources. When using these acoustic instruments, firstly the target strength of fish should be acquired, and then the standardized calibration should be implemented. Usually the acoustic image displayed on the screen are not dynamic, the accuracy of the fish counting is not very high. In addition, these methods also need to set the various parameters of fish, which will lead to significant differences on the identifying result[4].

This project utilized DIDSON (Dual-Frequency Identification Sonar) to detect the Dishui lake fishery resources, and acquired three days first-hand data. Based on echo counting method[9], the number of fish was counted from DIDSON image. Then the average density and zoning density of fish were calculated. At last, the total amount was calculated and the fish spatial distribution was plotted on a map based on ARCGIS software.

2 DIDSON

DIDSON (Dual-frequency IDentification SONar), which is invented by Washington University, can provide almost-video-quality images to identify objects even in turbid water. DIDSON has been used in fisheries management, underwater structures inspection, oil/gas leakage detection and identification, underwater security, evidence searching, ship's hulls and ports safety inspection, and underwater navigation and so on.

No other sonar in its class is able to deliver the image quality and frame rate of the DIDSON. The resulting real-time, near video-quality data is clear enough to study the behavior of darting fish, even in opaque waters.

A defacto standard for the exacting marine biology audience, this dual frequency sonar is also ideal for construction inspection, pipe laying observation, hull and berth sweeps and other mission critical jobs down to 300 meters.

The DIDSON is versatile, too, with a 1.1 MHz detection frequency capable of ima ging targets up to 30 meters away.Stored images can be converted to AVI files and JP EGs, and a size and distance tool is included, making analysis accurate and presentati on easy. [9]

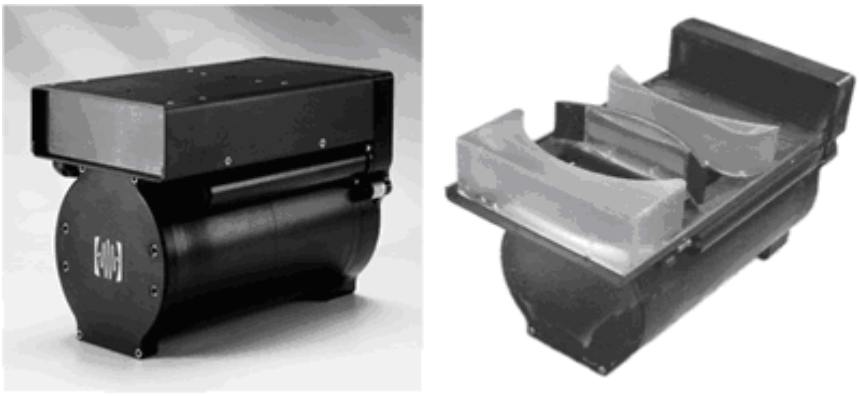


Fig. 1. Standard DIDSON (Left: main body, Right: acoustic lens)

3 Fishery Resources Detection and Assessment

3.1 Area and Methods

The three-day detection almost covered whole Dishui Lakes, each designed survey line spacing was 50 m. Due to driving, weather conditions, underwater fishing nets and barriers, the actual survey routes spacing was about 100 m as shown in Fig. 2.

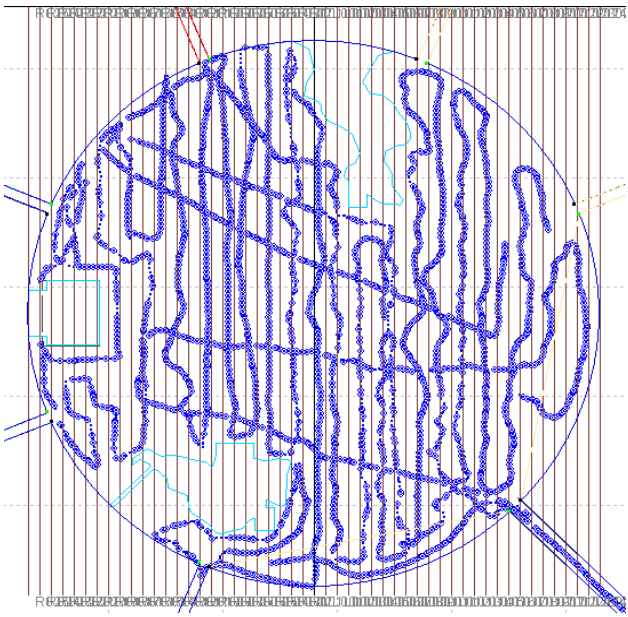


Fig. 2. The actual detecting route in Dishui lake, shanghai, China

The detection system(Fig.3) comprises the DIDSON transducer, a set-top control box, a data cable, control software, GPS and an associated laptop computer. The DIDSON transducer is directly connected to the set-top box, which is linked to the laptop via an ethernet connection. The acoustics image is transferred from the unit to the laptop via the control software, which displays the data as a streaming image. The DIDSON Image can then be directly viewed or saved on computer. The GPS positioning information is transferred into the navigation software and the actual survey routes are recorded (Fig.2). The practices improved the DIDSON transducer should be installed under water 0.5m and the lens angle 45°is best.

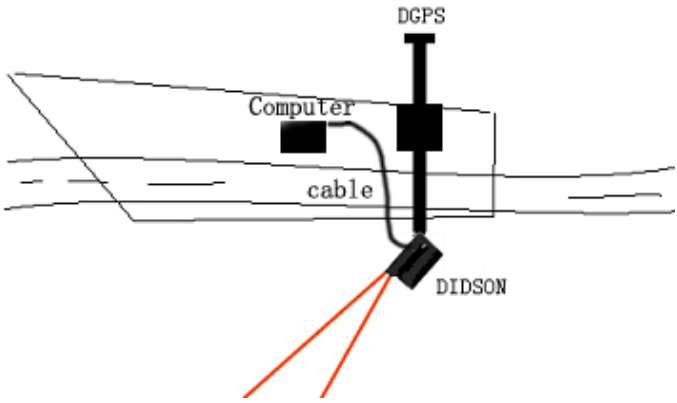


Fig. 3. Measurement of vessel and equipment installation

3.2 Fish Counting Automatically Based on DIDSON Image

Using DIDSON image (Fig.4) and semi-automated counting software[10,11] developed ourself, the fish number of every line was counted and the location was recorded[12,13]. The fish length was calculated based on the pixel value of identified fish in the image. At last, The number checked via manual review would be saved to table (Table 1) for late statistical analysis.

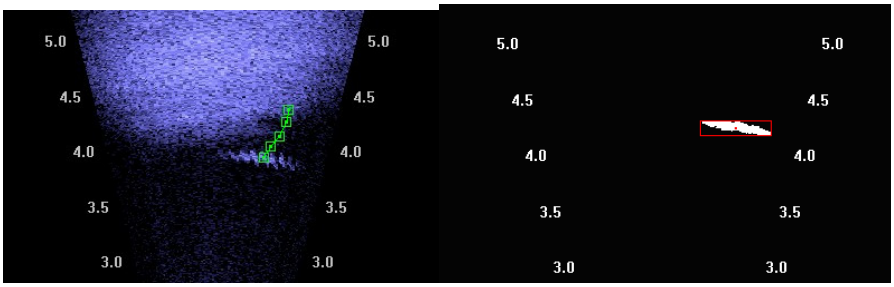


Fig. 4. Fish Identification and Tracking

Table 1. The fish count table from DIDSON image

N	E	Local X	Local Y	Num	Len(cm)
30.90368	121.9308	-36690.4516	44345.14614	1	96
30.90346	121.9308	-36714.39585	44347.71281	2	60
30.90338	121.9308	-36723.94869	44348.47933	2	72

3.3 Catch Statistics

By the actual measurement data form the Dishui Lake management Company, the fish species composition, body length, weight, and other relevant data were acquired. Based on caught 41 silver carp, the relationship between weight and length (Fig.5) was established using the power function regression method.

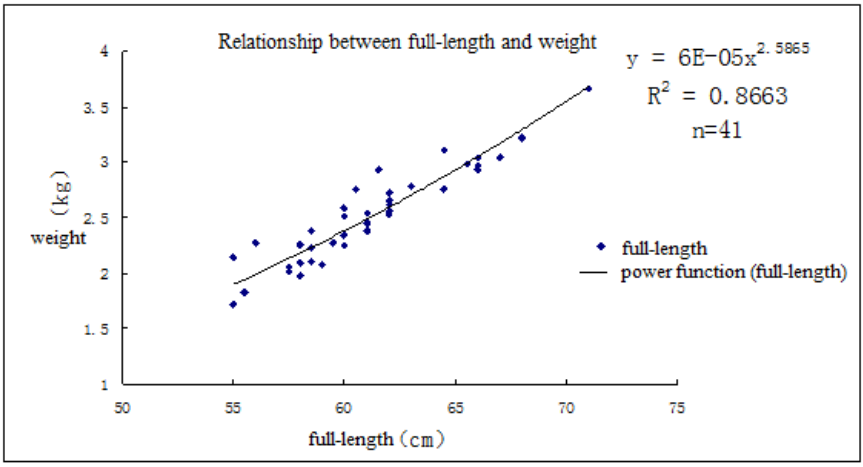


Fig. 5. Relationship between full-length and weight of silver carp (R2 is fitting degree, n is the number of samples)

The power function is

$$W = 0.00006L^{2.5865} \tag{1}$$

3.4 Fishery Resources Quantitative Assessment Methods

The average density method and zoning density method are used in fish resources quantitative assessment in this paper.

3.4.1 The Average Density Method

The whole lake fish quantity value N is defined as

$$N = \rho.S \tag{2}$$

where ρ is fish average density, S is the whole lake area. The value ρ is calculated by every line fish density ρ_i .

$$\rho = \frac{1}{n}\sum_{i=1}^{n}\rho_i \tag{3}$$

The ρ_i is calculated by the fish number and the area of line i.

$$\rho_{\mathrm{i}} = \frac{\mathrm{N_i}}{\mathrm{S_i}} \tag{4}$$

Where the value $\mathrm{S_i}$ is defined as

$$\mathrm{S_i} = \mathrm{L_i} \times \frac{\mathrm{R_i}}{2} = \frac{1}{2}\sum_{n=1}^{n} r_n l_n \tag{5}$$

Where the $\mathrm{L_i}$ is length of line i and the $\mathrm{R_i}$ is the slant-range from the transducer to the water bottom. The r_n is changed by the water bottom terrain from time 1 to n seconds. The l_n is computed by GPS positioning data.

Based on these methods above, the length, area, fish number, fish density of each survey line are calculated.

Table 2. The detection data of each survey line

Line	Length(m)	Area(m2)	fish number	fish density (num/m2)
No1	3109	15206.1	485	0.031895095
No2	4561			
No3	1871	5697.7	385	0.067571125
No4	3740	13552.6	1038	0.07659047
No5	3828	12001.1	1152	0.095991201
No6	3907	17333.9	1198	0.069113125
No7	4605			
No8	6703	21239.2	482	0.022693887
No9	1827	3451.6	269	0.077934871
No10	962	3167.7	190	0.059980427
No11	7312	22139.6	865	0.039070263
No12	3901	12206.9	823	0.067420885

The whole lake fish density ρ is 0.0608261349 via formula 3, and the Dishui lake area is 5.56 sq.km so the whole lake fish number N is 338193.

By the site statistics data, the mean size of Dishui lake fish is 64.26cm. So the fish average weight is 2.847kg, and the whole lake fish gross is 962835kg.

3.4.2 The Zoning Density Method

The fish average density method ignored the number difference of different regional, and it could not express the actual distribution of the fish, and it's accuracy was affected by the survey lines distribution. So the zoning density method was designed to improve the count accuracy.

The value $K1$ (num/m^2) was designed to express the fish density of every survey point.

$$K1 = \frac{n}{l \cdot d} \tag{6}$$

Where value n is the detected fish number, value l is the distance between two survey points, value d is the DIDSON scan width.

The value $K2$ (kg/m^2) was designed to express the fish grass of every survey point.

$$K2 = \frac{nW}{l \cdot d} = \frac{0.00006n}{l \cdot d} \cdot L^{2.5865} \tag{7}$$

Where value L (cm) is length of fish, W (kg) is the fish weight.

Based on formula 6 & 7, the value $K1$ and $K2$ was calculated.

Then using Inverse Distance Weighted (IDW) interpolation within ARCGIS, the whole lake fish density were processed as shown Fig.6.

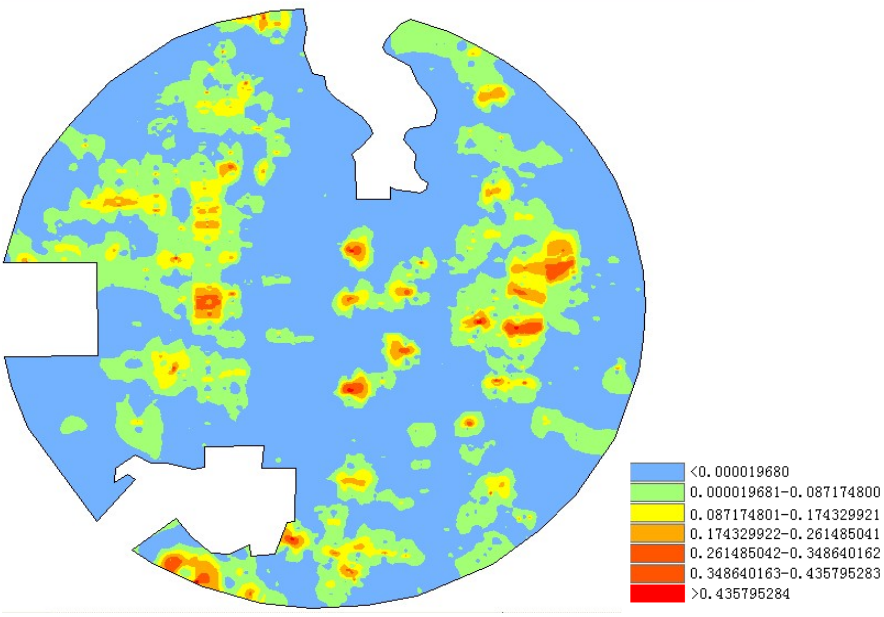


Fig. 6. Map of K1 zoning density (unit: num/m2, grid resolution 10m)

Based on the K1 Grid data, the whole Dishui lake fish number was calculated via ARCGIS Surface Volume tool. And the whole number is 345253.

Based on the K2 Grid data, the whole Dishui lake fish gross was calculated via ARCGIS Surface Volume tool. And the whole fish gross is 939885kg.

3.5 Fishery Resources Spatial Distribution

In order to understand the fish spatial distribution of Dishui lake, the pie charts of detected fish number were plotted on a map based on the counted fish number of every points. From the map, the Dishui lake fish spatial location are expressed visually (Fig. 7). Actually the Figure 6 expresse the fish spatial distribution also.

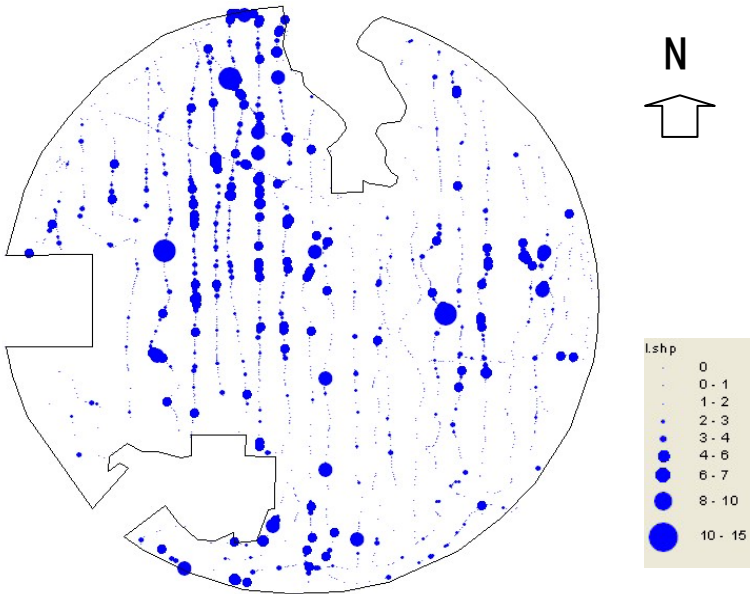


Fig. 7. Map of Dishui lake fish distribution

4 Analysis and Discussion

4.1 Summary and Analysis

In this paper, two kinds of assessment methods were used to calculate the total fish amount in Dishui lake. Whith average density method, the Dishui lake fish number is 338193 and the fish gross is 962835kg. And the number is 345253 via zoning density method, and the fish gross is 939885kg.

The assessment value by two methods is close, and the fish fry data supplied by Dishui lake management department was analyzed to prove that the two methods are

effective. Overall, the average density method is simple, easy to operate. The zoning density method needs a large amount of calculation and requires the assist by GIS software.

Practice proved that the average method is simple and effective, can be directly applied to the assessment of fishery resources. The zoning density method witch can visually analyse fish distribution and the amount of the different regions needs lots of equally distributed detection data to interpolate the whole area data.

In comparison with the lake depth map, the fish were mainly concentrated in two deep-water areas and several estuaries. Because the main fish silver carp like warm area in winter, so they lied in deep water and lakebed. In addition, the bait was rich in estuary, so they foraged there.

4.2 Question and Discussion

Due to technical and capacity constraints, there are several deficiencies need for improvement.

In this detection, the fish species relatively simple and mainly concentrated in the bottom of the lake, so we used a flat density. But in many cases, different fish distributed in different water depth. So the 3D fish density in order to more accurately measure the 3D distribution will be survey in the next study.

Sonar instruments have certain blind spots, for example DIDSON has a big noise away from the lens 1 meter, where the fish can not be quantified. So how to solve this question is a key point for the next research.

When using the acoustic instruments to survey fishery resources underwater, due to the movement of the boat and the noise of boat motor, part of the fish will be driven. So how to decide the drive coefficient is very difficult but necessary, and the current research is still very scarce.

Applying DIDSON sonar to detect and assess fishery resources in China has just begun, but the practices proved that it could be used effectively in large-waters fishery resources detection and assessment.

References

1. Gunderson, D.R.: Surveys of Fisheries Resources, p. 69. John Wiley & Sons, New York (1993)
2. Elliott, J.M., et al.: A comparison of three methods for assessing the abundance of Arctic charr, Salvelinus alpinus, in Windermere (northwest England). Fisheries Research 53(1), 39–46 (2001)
3. Tan, X., Kang, M., Tao, J., et al.: Hydroacoustic survey of fish density, spatial distribution, and behavior upstream and downstream of the Changzhou Dam on the Pearl River. China. Fish Sci., 891–901 (2011)
4. Tan, X., Shi, J., Zhang, H., et al.: Hydroacoustic assessment of fish resources in the Lake Qinghai with EY60 echosounder. Lake Sci. 21(6), 865–872 (2009)

5. Wang, C.-R., Zhang, H., Du, H., et al.: Hydroacoustic assesment of abundance and spatial distribution of Gymnocypris przewalskii in Qinghai Lake with BioSonics DT-X echosounder. Freshwater Fisheries 41(3), 15–21 (2011)
6. Diner, N.: Correction of school geometry and density: an approach based on acoustic image simulation. ICES Coop. Res. Rep. 238, 27–51 (1999)
7. Misund, O.A., Aglen, A., Fronaes, E.: Mapping the shape, size, and density of fish schools by echo integration and a high-resolution sonar. ICES J. Sci. 52, 11–20 (1995)
8. Gauthier, S., Rose, G.A.: Acoustic observation of diel vertical migration and shoaling behaviour in Atlantic redfishes. Fish Biol. 61, 1135–1153 (2002)
9. Wang, J., Zhang, C., Wang, D.: Acoustic assessment of silver carp and bighead carp in Qinghe Reservoir. South China Fisheries Science 6(5), 50–55 (2010)
10. Han, J., Asada, A., Mizoguchi, M.: DIDSON-based Acoustic Counting Method for Juvenile Ayu Plecoglossus altivelis Migrating Upstream. Journal of the Marine Acoustics Society of Japan 36(4), 250–257 (2009)
11. Han, J., Asada, A., Honda, N., et al.: Automated Acoustic Method for Counting and Sizing Farmed Fish during Transfer using DIDSON. Fisheries Science 75, 1359–1367 (2009)
12. Tong, J., Han, J., Shen, W., Xu, P.: Mosaicing of Acoustic Video Images for Underwater Structure Inspection. In: 29th International Conference on Ocean, Offshore and Arctic Engineering, Shanghai, China (2010)
13. Tong, J.-F., Han, J., Shen, W.: Preliminary Research on the Image Processing of Acoustic Camera and Its Application in Fishery. Hunan Agricultural Sciences 17, 149–152 (2010)

Study on Cultivated Land Concentrated Areas Delineation Based on GIS and Mathematical Morphology: A Case of Miyun County and Pinggu District in Beijing

Yanmin Ren[1], Yongxia Yang[2], Yuchun Pan[1], Yu Liu[1], Yunbing Gao[1], Xiumei Tang[1], and Zhixuan Zeng[1]

[1] National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
[2] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

Abstract. The basic farmland within the protection areas was required to be of high-quality and connective in the general plans for land use. A new method was developed to delineate the boundaries of high-quality, concentrated and connective cultivated land. Based on mathematical morphology principles and GIS methods, the high quality of cultivated land blocks could be identified through dilation and erosion operations according to this rule that which inside distance was less than threshold d_1 and the number of blocks was larger than 3. This method was validated with the farmland classification data of Miyun County and Pinggu District in Beijing City, and 91% of high-quality, concentrated and the connective cultivated land could be identified by this method, so it would provide the reference method for delineating the basic farmland areas scientifically and reasonably in the general plans for land use.

Keywords: mathematical morphology, concentrated and connective, GIS, cultivated land.

1 Introduction

Cultivated land resources were scarce and it was a basic state policy that everyone should cherish and utilize rationally every inch of the farmland in China [1]. Basic farmland protection system was set up and basic farmland regions were delineated to execute strict protection rules [2]. Basic farmland was the arable land which was of high-quality, good land geographic conditions, complete water conservancy facilities, concentrated and connective etc [3]. A lot of researches had been done on how to delimit basic farmland, and much emphasis had been put on the principles of quality, the distance from the farmland to the settlement. However, the distribution and concentration of the farmlands were usually ignored in these principles. These principles made the cultivated land blocks scattered, and it was difficult to realize scale management of agriculture and furthermore, it resulted in that a large number of

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 376–383, 2013.

scattered basic farmland had been transferred. Basic farmland protection areas planning introduced the principle of " high-quality and concentrated ", and it would focus on protecting the blocks of high-quality and concentration, reducing flow, realizing the scale of agriculture as well as controlling the non-point pollution. As a result, the concentrated and connective concept in the space planning of basic farmland was helpful to realize the management of "high-quality, concentration and connective ".

At present, the researches about connectivity at home and abroad mostly focused on Landscape ecology and ecological protection and other fields, and rarely on cultivated land protection [4]. In recent years, more and more researches were concerned about connectivity in China, such as: Yun Wenju, Zhou Shangyi[5] proposed the methods for measuring space connectivity by calculating the link index of blocks with grids; Duan Gang[6] raised a method based on the buffer analysis of vector; Guo Zihan[7] determined whether the cultivated land was connected by setting up distance threshold for the grid length. These methods were helpful to select the cultivated land areas which were of high-quality, concentrated and connective. However, too long calculating time and the distance threshold which was difficult to ascertain had a significant impact on delimiting concentrated areas in the study area. As a new subject of image processing and analysis, the principal theory and methods of Mathematical Morphology had been successful applied in the fields such as the biomedical, remote telemetry, highway and transportation, mechanical engineering and weather etc. For example, Li Deren [8] analyzed Chinese characters of scanned binary images; Feng Qingzhi [9] used these theories and methods in image enhancement and detection for license plates and complicated fingerprints; Wu Dan [10] had studied application status in GIS and recognition and the position of image mark; Zhang Qingnian [11] proposed the method of identification and generalization of clusters by mathematical morphology and proved to be quite effective.

Considering that the cultivated land concentrated area could be regarded as area clusters, this research referenced Zhang Qingnian's methods on identifying area clusters. Firstly, the basic operations of mathematical morphology were introduced. Secondly, the identifying methods were probed to delineate concentrated area of cultivated land based on mathematical morphology. Thirdly, take the farmland classification data of Miyun County and Pinggu District in Beijing City for example, this method's feasibility was verified and it would provide the reference method for delineating basic farmland protection areas scientifically and reasonably in the general plans for land use.

2 Theory and Methodology

2.1 Basic Operations of Mathematical Morphology

The areal geographical phenomenon of discontinuous distribution was suitable to be expressed by raster[11]. Mathematical morphology could analyze the structure of data and extract features through transforming the raster data structure [12]. Grid-based areal features were expressed in polygon, called blocks. The kind of map data was easily processed using the mathematical morphological methods.

Mathematical morphology was a nonlinear image processing and analysis theory, which discarded the traditional view of value modeling and analysis. It described and analyzed the image from the perspective of set operations, focusing on the geometric structure of image, in order to extract the targets' size, shape, connectivity, convexity, smooth , directional characteristics and so on[11]. In some sense, mathematical morphology opened the new theory and the new method of image processing, analysis and identification. The operations allowed for efficient handling of grid data in morphological transform and pattern recognition, as well as other types of data.

Basic operations of mathematical morphology included dilation erosion and opening, closing, thinning, thickening, top-hat transform and low-hat transform derived from dilation and erosion. Among them, dilation and erosion were the most basic operators of mathematical morphology. Because the other operators could be derived by dilation and erosion, this paper just introduced the two operators.

X was input image, B was structural element,

(1) **dilation** : $X \oplus B = \{x+b : x \in X, b \in B\} = \bigcup_{b \in B} X[b]$

Geometrically, dilation was the union of all calculation results with input image X translation of b.

(2) **erosion** : $X \Theta B = \{x : x+b \in X, b \in B\} = \bigcap_{b \in B} X[b]$

Geometrically, erosion was the union of all calculation results that input image X was translated –b, also was composed of all points within input image X after structural element B was translated x.

Operation results of dilation and erosion were shown in fig.1 below.

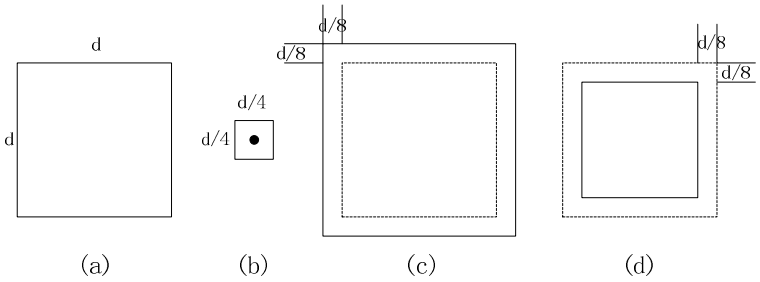


Fig. 1. dilation and erosion

(a) Input Image X (b) Structural element B (c) dilation, X $\oplus$ B (d) erosion, X Θ B

2.2 Identifying Cultivated Land Concentrated Areas by Mathematical Morphology

"Connectivity" is an indicator which showed the spatial distance and distribution of cultivated land blocks with the same property. When the distances among cultivated land blocks were smaller than the threshold value, they could be considered to be

connective[5]. "Clusters" was a group of features whose distances were less than the threshold value L, and the clusters were usually defined by human eyes. In this paper, a new method of identifying cultivated land block clusters, based on mathematical morphology, was developed to identify and delineate connective cultivated land regions.

Cultivated land clusters were identified based on two criterion, (1) the distances among features inside clusters; (2) the ratio between distances inside and outside cluster. Cultivated land blocks were considered as connective when their maximum distance d1 inside clusters was less than the distances mean value d2 outside clusters, while the block number of cluster was not less than 3.

Four steps were included to implement this method, as were shown as follows:

(1) Determinating the distance threshold d1 of concentrated areas. Firstly, cultivated land blocks were selected randomly as a cluster in the image. Secondly, the distances from each block to other blocks inside clusters were calculated. Thirdly, the maximum distance value was considered as the threshold of concentrated areas. The implement methods were as: There were n blocks Pk(k∈(1,n)) in the image X, the formula was needed to be looped until. B was the structural element with radius of 1, w was the number of loops, dk was the minimum distance between the block Pk and the others(k∈(1,n)) which equaled with w, the maximum of dk was considered as the threshold of concentrated areas d1. Image GH was assigned by X.

(2) Identifying the blocks aggregation which internal distance was less than threshold d1 and the number of blocks was not less than 3. Firstly, Image PR was dilated by structural element Br with radius of d1/2, and the arithmetic expression was . Image PR was composed of blocks Ck (whose number was h)(k∈(1,h)). Secondly, the aggregation inside Ck could be got through (k∈(1,h)). The number of blocks n inside was counted, and judged whether n was less than 3 or not. If n<3, the blocks were considered scattered and they would be deleted from GH, then they could be expressed as. Otherwise, the blocks were considered as connective, and the aggregation was initialized.

(3) Judging whether the blocks aggregation met with the criteria d1/d2≤1/2. The formula was needed to iterate by structural element B with radius of 1 until it met with this criteria. The variables were as follows: w was the iteration number, d2 was the minimum distance between the initialized aggregation and the other surrounding blocks, which was equal to w. If d1/d2≤1/2, PCk was seen as clusters.

(4) Identifying the aggregation for which inside distance was smaller than before by reducing d1 in the rest of the blocks. The identified clusters were deleted from image GH and it could be expressed by the formula GH=GH-PC, after uniting with all recognized clusters through (k∈(1,w)).

If, d1 was assigned with the value of d1-d*2% . Go back to Step 2 when d1≥Ds*St/Ss to identify the clusters with smaller inside distance in the rest of the aggregation until all clusters inside image X were iterated through. Among them, Ds was the minimum distance discriminated by human's eyes in the target image. Ss and St were the scale denominators of the original image and the target image, respectively.

The operation flow to identify the cultivated land concentrated areas was shown in Fig.2 below.

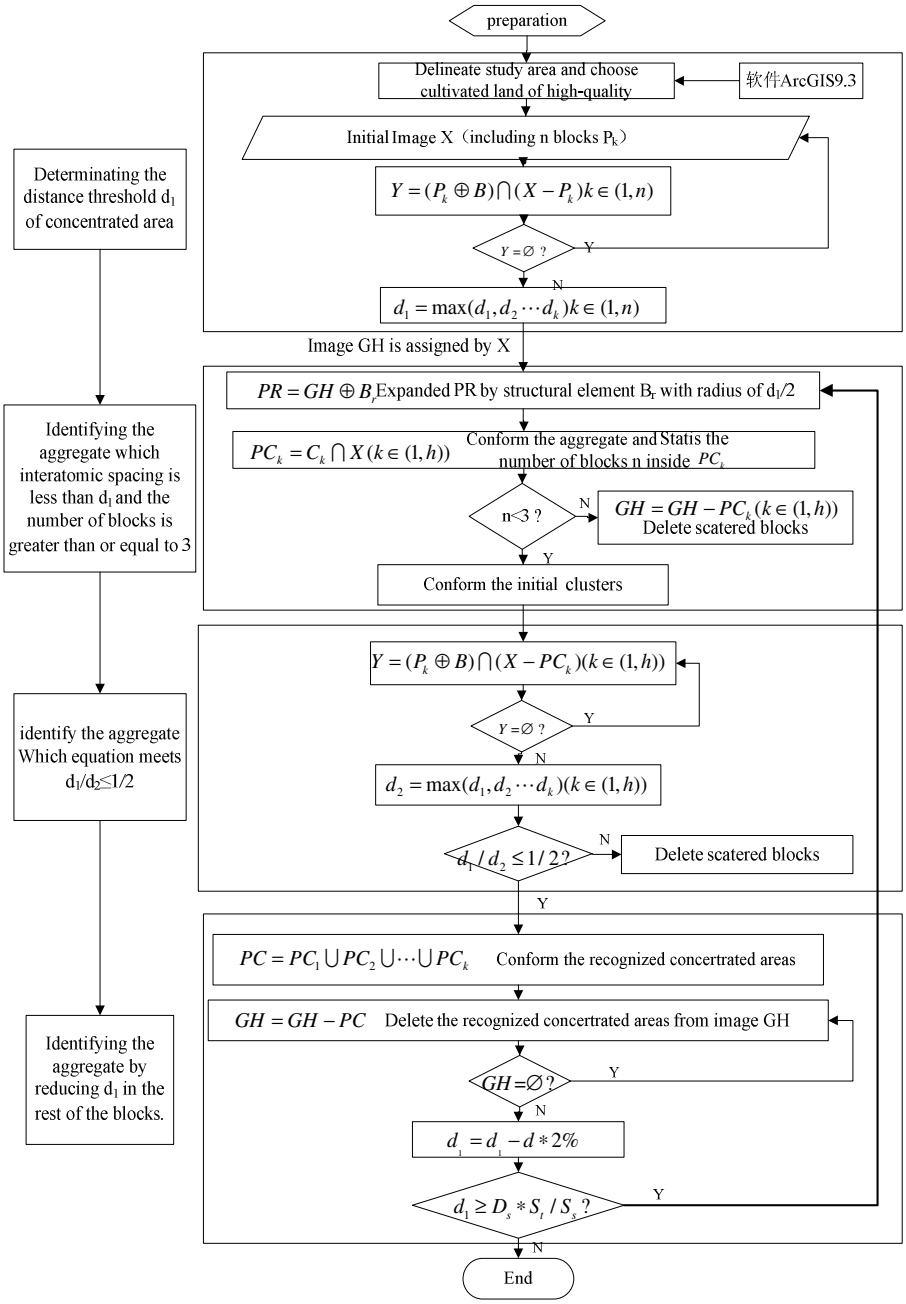


Fig. 2. Operation flow to identify the cultivated land concentrated areas

3 Application and Demonstration

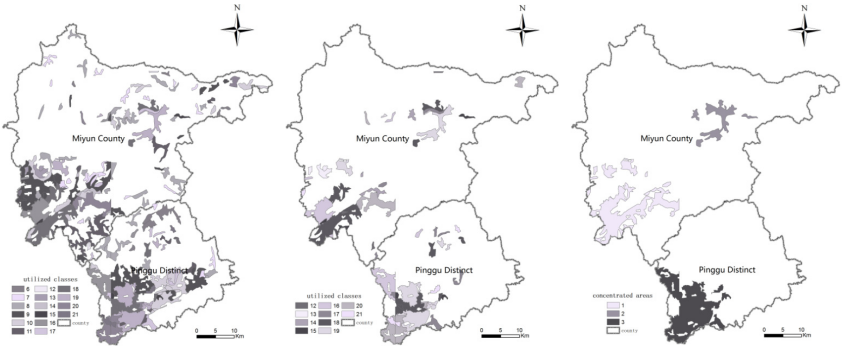
The method based on dilation and erosion operators of mathematical morphology, was applied and validated by identifying the concentrated areas of cultivated land in Miyun County and Pinggu District with the help of the Classification of Agricultural Land Map.

Classification of Agricultural Land Map in was Miyun County and Pinggu District in Beijing City was shown in Fig.3 (a) and its coordinate projecting system was Beijing 54. The agricultural land classes were determined by the project team according to the technical line of "Regulation for Classification of Agricultural Land" [15]. The regulation was the result of comprehensive evaluation in accordance with local natural factors including climate, soil, topographty as well as the utilization level and so on. The utilized agriculture land was ranged from 6 to 21grades and total area was 52548 hectares. According to the standard of protecting basic farmland, it was defined that the farmland above 11 grades was the high-quality cultivated land (In fact, the high-quality grade could be defined by local land administration departments). The high –quality agriculture land was ranged from 12 to 21 grades and total area was 42799 hectares, accounting for 81.45% of the total utilized agriculture land area.

The steps to delineate the concentrated area of high-quality cultivated land were described as follows.

Step 1: The farmland blocks above 11 grades (high quality) were extracted based on their property value, and they were shown in Fig.3 (b).

Step 2: Identifying the concentrated areas of cultivated land based on mathematical morphology. After rasterizing the resulting image in step 1, the operation flow in Fig.2 were to run to identify the cultivated land concentrated areas. Finally, 3 concentrated areas were identified, and they were expressed by different colors in Fig.3 (c). The 3 concentrated areas were 13647 hectares, 4008 hectares, 14953 hectares, respectively.



(a) (b) (c)

Fig. 3. Delineate concentrated areas of high-quality cultivated land

(a)the map of utilized classes ,(b)the map of high-quality,(c) the map of concentrated areas of high-quality cultivated land

The result showed that the delineated concentrated areas of high-quality cultivated land was 32608 hectares which was accounting for 91% of the total high-quality cultivated farm land area (35862 hectares) in Miyun County and Pinggu District in Beijing City. It showed that this method was feasible and precise result was attained. Therefore, it was helpful for delineating basic farmland protection areas.

4 Conclusion and Discussion

Protecting basic farmland was an elementary state policy in china. A new method was developed in this research to delineate the connective concentrated basic farmland based on mathematical morphology principles and GIS technology. Furthermore, this method was validated with the farmland classification data of Miyun County and Pinggu District in Beijing City and 91% of the connective and concentrated farmland could be identified by this method. The achievements of this research were as follows: (1) the method was able to locate the position of the connective cultivated land and it could be identified satisfactorily with great computational efficiency and high computational accuracy; (2) delineating concentrated and connective cultivated land was helpful to consolidate basic farmland as well as to control basic farmland transformation. Therefore, it was of significant meaning to protect and supervise basic farmland.

At present, there had been few studies on the connectivity of cultivated land, and the attempt was made based on mathematical morphology in this paper, it showed that the method was effective to delineate connective and concentrated cultivated land. However, further efforts should be made on these aspects: (1) the maximum threshold value of cultivated land block area should be defined, for the reason that when the area of cultivated land were large enough to management connectively, it was not meaningful to delineate any more;(2) the number of the scattered cultivated land should be further considered according to different situations. In this paper, the blocks number was set 3, when the scattered cultivated land were less than 3, they were not taken into the concentrated and connective the aggregates. When the area of cultivated land blocks was large, the error would be amplified, so improvement should be made for this method in the further researches latterly.

Acknowledgments. This work was supported by the National Science-technology Support Projects for the 12th five-year plan under Grant No.2011BAD04B0302.

References

1. General land use planning regulations for examination and approval. No.7 of China Land Administration Bureau (1998)
2. Regulations on the Protection of Basic Farmland. Order No.257 of the State Council (released on December 27, 1998)
3. Zhou, S., Zhu, A., Qiu, W.: GIS based connectivity analysis and its application in prime farmland protection planning. Transaction of CSAE 24(07), 72–77 (2008)

4. Marullia, J., Mallarach, J.M.A.: GIS methodology for assessing ecological connectivity: Application to Barcelona Metropolitan Area. Landscape and Urban Planning (2005)
5. Yun, W., Zhou, S., Zhu, A.: Protect high-quality cultivated land Concentrated and connectively. China Land and Resources News 3-21(005) (2008)
6. Gang, D.: Study on the method of basis farmland Protection spatial Planning in basis of agricultural land classification. Chang'an University, Xi'an City (2009)
7. Guo, Z., Yang, Y.: GIS-based farmland connectivity analysis methods research and system implementation. Geography and Geo-Information Science 26(3) (2010)
8. Li, D., Chen, X.: Mathematical Morphology and its application binary image analysis. Journal of Wuhan University of Surveying and Mapping 14(3) (1989)
9. Feng, Q.: The method of image processing and application based on mathematical morphology.Journal of China, Criminal Police University (2011)
10. Wu, D., Liu, X., Shang, J.: The application and prospect of Mathematical Morphology.Journal of Engineering Graphics (2003)
11. Zhang, Q., Qin, J.: Identification and generalization of area clusters by mathematical morphology. Geographical Reseach 19(1) (2000)
12. Ma, F., Li, D.: Application of mathematical Morphlogy in GIS spatial analysis. Journal of Wuhan Technical University of Surveying and Mapping 21(1) (1996)
13. Chen, S., Li, Y.: A new algorithm of minimum convex hull and its application. Geography and Geo-Information Science 25(5), 43–45 (2009)
14. Liu, R., Yang, D., Li, Y., Chen, K.: A improved algorithm of minimum convex hull. Journal of Geodesy and Geodenamics 31(3), 30–33 (2011)
15. Ministry of Land and Resources. Regulation for Classification of Agricultural Land. China Standard Press, Beijing (2003)

Design and Implementation of Rapid Grading Platform for Shape and Diameter of Oranges Based on Visual C#.NET*

Wenshen Jia[1], Wenfu Wu[1], Fang Li[1], Ligang Pan[2,3], Zhihong Ma[2,3], Miao Gao[2,3], and Jihua Wang[2,3,**]

[1] School of Biological and Agricultural Engineering, Jilin University, Changchun 130022
[2] Beijing Research Center for Agrifood Testing and Farmland monitoring, Beijing 100097, China
[3] National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
wangjh@nercita.org.cn

Abstract. Research on digital image processing technology, which began in 1960s, stepped into an active research stage in late 1970s and early 1980s. It was firstly used in industrial and biomedical fields. Although it put into use in agricultural research very late, it has a broad prospect. In order to realize navel orange grading, this paper used Visual C # .NET program to develop navel orange shape and diameter rapid grading based on machine vision image feature. Southern Jiangxi navel orange was used as the research object. The color and shape feature of the navel orange was extracted. The image data was processed through Sobel Operator algorithm and standard median filter. Results show that the digital image processing technology based on C # program is feasible for shape grading of navel orange. It also provides a new method for navel orange grading detection.

Keywords: Visual C#, navel orange, agricultural products, image processing, shape.

Introduction

Machine vision technology is a computer image processing technology. It is simulate the human visual system to extract information from the image of objective things. Processing of such information achieves detection, measurement and control for the target. The use of machine vision technology for agricultural grading and detection

* This work is supported by Research and Demonstration of Authenticity Identification and Quality Safety Traceability Technology of Agricultural Products(201203046) funded by Special Fund for Agro-scientific Research in the Public Interest as well as the Restoration Technology Integration and Demonstration on Disaster destroyed and Wastewater Irrigated farmland（2011BAD04B04） funded by the National Key Technology R&D Program.
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 384–389, 2013.

means to use custom-built computer software computing processing of agricultural products image and effective decomposition of the image information. Through the machine vision, the agricultural products of pixel, boundary, shape and so on, has obtained its own characteristics, with a certain pattern to match the image. Finally get to determine the agricultural products quality[1].

Navel orange is one of the largest fruits of our production. At the same time, it is also an important trade fruit. Because of the detection and the grading technology is backward, causing the listed level of navel orange mixed, which affect its commodity value and in particularly cause the lack of competitiveness in the international market. At present, the navel orange grading relies on manual completed in Chinese. The grading results are less consistently and less efficient due to individual diversity of workers. Using machine vision to classify navel orange, which is objectivity strong, stable standard, consistency, high efficiency, and lossless. It is one of the effective ways to solve the artificial grading problem, and machine vision technology is also a hot research topic in agricultural applications[2].

Dimension is one of the main quality parameters in navel orange, also one of the important bases for grading. The sales clerks usually sell navel orange with the dimension grading. Depending on machine vision technology we can obtain the navel orange image, and select the relevant parameters. We can obtain data like volume, weight, diameter of the navel orange and so on.

We develop a grading of software for the navel orange based on machine vision technology. And in accordance with the CAC standards, using diameter of navel orange, we can grade the navel orange.

1 Test Materials

Test specimen: Navel orange

Producing area: Southern Jiangxi, purchased yuanfang supermarket in Changchun

2 System Design

2.1 Hardware Architecture

Computer：Hasee Core i3 380M, Memory 6G, Operating System Win7 Ultimate

Light box size: 500 mm×500 mm×400 mm, white background

Light: PAK T5, tricolor circular fluorescent lamp, Power 40W, Color Temperature 6500K, placed on the light box at the top of the central

Camera: Logitech C270, 300 million pixels, the camera installed in the center of the light, Mounting Height: 380 mm from the lens to light box bottom.

2.2 Software Design

The system use C # language development, a new programming language, which originate from .NET[3]. C # language evolve from C and C++, that is a Simple,

modern, type-safe and object-oriented language. C # is designed for a wide range of enterprise applications running on the .NET platform. .NET Framework includes the common language runtime (CLR) and .NET Framework Class Library (FCL). FCL provides trusteeship applications, and writes object-oriented API[4]. When we write .NET Framework applications, we do not consider to the Windows API, MFC, ATL, COM or other tools and technologies. .NET provides great convenience to programmer, so they only focus on the algorithm in the program is enough. Relative to C and C++ program development, writing the same piece of code, C # is not only the development cycle is short, a small amount of code, and readability. C # requires a lower level of the programmer, and the entry time is shorter. So that the programmer only focus on the algorithm of the digital image, that is greatly improving the efficiency of software development, is very conducive to the application of image processing.

3 Image Acquisition and Pretreatment

3.1 Image Acquisition

Image acquisition for navel orange use light box which have circular fluorescent lamp and white background. The circular fluorescent lamp first preheats 5 minutes, to ensure the stability of light. Each navel orange collects one picture, and collected a total of 10 pictures.

3.2 Image Pretreatment

In order to facilitate detection and extraction of image features for navel oranges, the original image using Photoshop CS5 software batch program for unified clipping zoom 512 x 512 pixels image.

4 Image Processing

4.1 Navel Orange Shape Extraction

The detection of edge is often uses the airspace differential operator. It is completed through templates and image convolution. There is always a gray edge between two adjacent areas of different gray values. Gray edge is caused by mutation of gray values. The discontinuous of gray values can be easily detected by seeking first-order and second-order derivative of gray values. Existing local edge detection methods include first-order derivative(Sobel Operator, Roberts Operator), second-order derivative(Laplace Operator), template operations(Prewitt Operator, Kirsch Operator, Robinson Operator) and so on[5]. This program use Sobel Operator algorithm.

Sobel Operator weight discrete data and it is use small convolution template (Fig 1). It use convolution of template and corresponding image data as the similar results. The two direction templates are horizontal direction edge detection and vertical edge detection, respectively.

1	2	1
0	0	0
-1	-2	-1

（a） Vertical gradient direction to detect the horizontal edge

1	2	1
0	0	0
-1	-2	-1

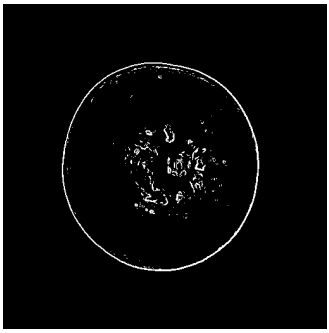
（b） Horizontal gradient direction to detect the vertical edge

Fig. 1. Sobel operator convolution template

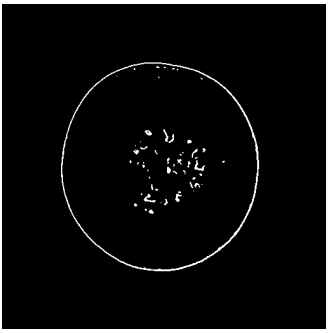
In the image processing the steps of using Sobel operator to detect image edges are as follows[6]:

1） Respectively, both directions template along the image move from one pixel to another pixel, and the center pixel of the template coincide with a location in the image.
2） The coefficients within the template and its corresponding image pixel by multiplying.
3） Get the sum of all the multiplied values.
4） Assign the maximum of the two convolutions to the pixel value of the corresponding template center location, as the pixel of gray value.

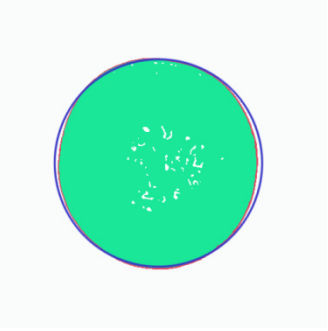
（a） Original image



（b） Edge detection image



（c） Median filtered image



（d） Circular contrast

Fig. 2. Median filtered image

Fig 2 (b) by the navel orange original image in Fig 2 (a), calculated after Sobel Operator method.

Fig 2 (c) median filtering images by edge detection image in Fig 2 (b), calculated after median filtering algorithm.

Standard median filter[7] (SMF), which mainly depend on the quick sort algorithm and is a nonlinear filtering method with less blurred edges, not only remove or reduce the random noise and pulse interference, but also retain edge information. After median filtering navel orange shape is clearly visible, it provides a basis for the further identification of the navel orange appearance grade.

4.2 Navel Orange Diameter Converted

Any color can be expressed by the tristimulus values. The three kinds of stimuli are called as three kinds of primary colors, usually refers to red, green and blue[8][9]. In 1931 the CIE-RGB color system selects the red (R) for the 700 nm, green (G) for the 546.1 nm, blue (B) for the 435.8 nm. Navel orange generally was orange so the blue component sensitive, which blue pixel value, is less than 220 after the test images judged to be the navel orange part. The effect is very good.

As shown in Fig 2 (d), the difference between navel orange and circular is only 1120 pixels, the different area of navel orange and circular is only 1.4% of the total area. Therefore, it is feasible to derive the diameter of navel orange from circle formula.

The camera lens is 380 mm away from the bottom of the light box. As the lens away from the sample is fixed height, therefore, as long as the statistical proportion of navel orange pixels share of the whole image can be converted navel orange cross-sectional area. In accordance with the largest cross-sectional area of the navel orange, its diameter could be calculated through area formula of a circle. Table 1 shows the level of navel orange graded by diameter in accordance with the CAC standard[10].

Table 1. CAC Navel orange grading standards

Level	Diameter (mm)
0	92 – 110
1	87 – 100
2	84 – 96
3	81 – 92
4	77 – 88
5	73 – 84
6	70 – 80
7	67 – 76
8	64 – 73
9	62 – 70
10	60 – 68
11	58 – 66
12	56 – 63
13	53 – 60

5 Conclusions

This paper studied on navel orange grading detection based on machine vision technologies, and established a navel orange image acquisition system. Through the navel image analysis and processing, this paper realizes the accurately distinguish between navel orange and the background based on Sobel Operator algorithm and standard median filter. Through the detection of the blue band of the image, this method can quickly distinguish between navel orange and the background, and convert the diameter of the navel orange. This method improves the efficiency of image processing, and is of great significance for online rapid detection. The results show that the detection using machine vision technology based on C # program for navel oranges' shape grading is feasible and provide a new idea for the grading detection of navel orange.

Acknowledgment. This work is supported by Research and Demonstration of Authenticity Identification and Quality Safety Traceability Technology of Agricultural Products(201203046) funded by Special Fund for Agro-scientific Research in the Public Interest as well as the Restoration Technology Integration and Demonstration on Disaster destroyed and Wastewater Irrigated farmland （2011BAD04B04） funded by the National Key Technology R&D Program.

References

[1] Li, J., Xue, L.: A study on navel orange grading system based on computer vision. Acta Agriculturae Universitatis Jiangxiensis 28(2), 304–307 (2006)
[2] Brosnan, T., Sun, D.W.: Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering 61(1), 3–16 (2004)
[3] Anderson, T.: Back in the studio-Visual Studio 2008. Personal Computer World 30(11), 148–149 (2007)
[4] Christian, N., Bill, E.: Professional C# 4.0 and.NET 4. Wrox (2010)
[5] He, C.H., Zhang, X.F., Hu, Y.C.: A study on the improved algorithm for Sobel on image edge detection. Optical Technique 38(3), 323–327 (2012)
[6] Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Electronic Industry Press (2011)
[7] Gou, X.M., Jia, X.H.: Digital Image Processing, Edge Detection Technique. Zhongyuan Institute of Technology (6), 64–70 (2007)
[8] Cao, L.P.: Machine recognition of citrus variety based on the fractal dimensions of perimeter-area. Transactions of the CSAE 26(2), 351–355 (2010)
[9] Cao, L.P., Wen, Z.Y., Shen, L.M.: Sugar Content and the Valid Acidity Test of the Citrus Based on the Fractal Dimensions of Hue. Transactions of the Chinese Society for Agricultural Machinery 41(3), 143–148 (2010)
[10] Codex Standard For Oranges. Codex Stan 245 (2004)

A Fast Processing Method of Foreign Fiber Images Based on HSV Color Space

Qinxiang Wang[1], Zhenbo Li[3,4,5], Jinxing Wang[2], Shuangxi Liu[2], and Daoliang Li[3,4,5,*]

[1] Shandong Provincial Key Laboratory of Horticultural Machineries and Equipments, Shandong Agricultural University, Taian 271018, China
[2] College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China
[3] China-EU Center for Information and Communication Technologies in Agriculture of China Agriculture University, Beijing 100083, P.R China
[4] Beijing Engineering & Technology Research Center for Intern.et of Things in Agriculture, Beijing 100083, P.R. China
[5] Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, P.R. China
dliangl@cau.edu.cn

Abstract. Traditionally, it was hard for image segmentation to suit the cotton image segmentation of foreign fibers. To solve this problem, this paper proposed an image segmentation method of foreign fibers based on HSV color space. The value of foreign fibers images' S channel was enhanced in this method to improve the contrast of foreign fiber and its background which help the subsequent image segmentation. The result of experiment shows that the method could highlight the images of foreign fibers, speed up subsequent image segmentation and realize fast image segmentation.

Keywords: Cotton, Foreign fibers, S channel, Image segmentation.

1 Introduction

Foreign fibers in cotton refer to the fibers mix in the raw cotton production, process and transportation. These fibers are non-cotton fibers and colored fibers which terribly affected the cotton and its production (Li et al., 2006). Most of the cotton enterprises pick out foreign fibers artificially, which waste both time and energy (Li, 2006). Nowadays, Machine vision technology is the main method to process online automatic detection of foreign fibers.

The aim of image segmentation is to partition the image into meaningful connected components to extract the features of the objects (Zhang et al.,2011).The segmentation results are the foundation of all subsequent image analysis and understanding, such as object representation and description, feature measurement, object classification, and scene interpretation, etc. Thus, throughout the image processing, image segmentation is

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 390–397, 2013.

an extremely important aspect. Various image segmentation methods are reported in the literature (Bakker et al., 2008; Kim et al., 2003;Pichel et al., 2006). In recent years, researchers have developed more efficient but also more complicated methods for segmentation. The methods mentioned above may work well in their specific context, but they are not capable of segmenting images of cotton foreign fibers because of the low contrast (Zhang et al., 2011).

As a result, a non-linear enhancement method based on S channel of cotton images is proposed in this paper. First, the RGB color space of the original images is transformed into HSV, and the values of S channel are extracted. Second, the non-linear enhancement model is constructed via saturation histogram analysis. Finally, the best thresholding is got through the iterative method.

2 Samples and Methods

2.1 Sample Preparation

The foreign fibers used in this research, including feather, hair, hemp rope, plastic film, polypropylene twine, colored thread, pieces of cloth, etc., as shown in Fig. 1, were collected from cotton mills. Adequate pure lint without foreign fibers was also prepared for making the lint layer.

Fig. 1. Foreign fiber samples

2.2 Image Acquisition

Put the prepared foreign fibers into sufficient lint without any of them before and opened the lint completely with opening machine. In the meantime, negative pressure was produced in the pipe through the fan, which leads the lint moving in it. Finally, an online scan industrial camera was used to scan these layer samples and acquired 24 bit true color images.

2.3 Choose of Color Space

Color spaces consist of many types such as RGB, HSV, YCbCr, CMY, etc. Acquired foreign fiber images are based on the RGB color space, a common color presentation method which is not suitable for human visual system. While HSV, describes the color with hue, saturation and value from the aspect of human visual system, is advantageous to the process and recognition of images. Hue refers to different colors, saturation refers to the various shades and value refers to the degree of lightness and darkness.

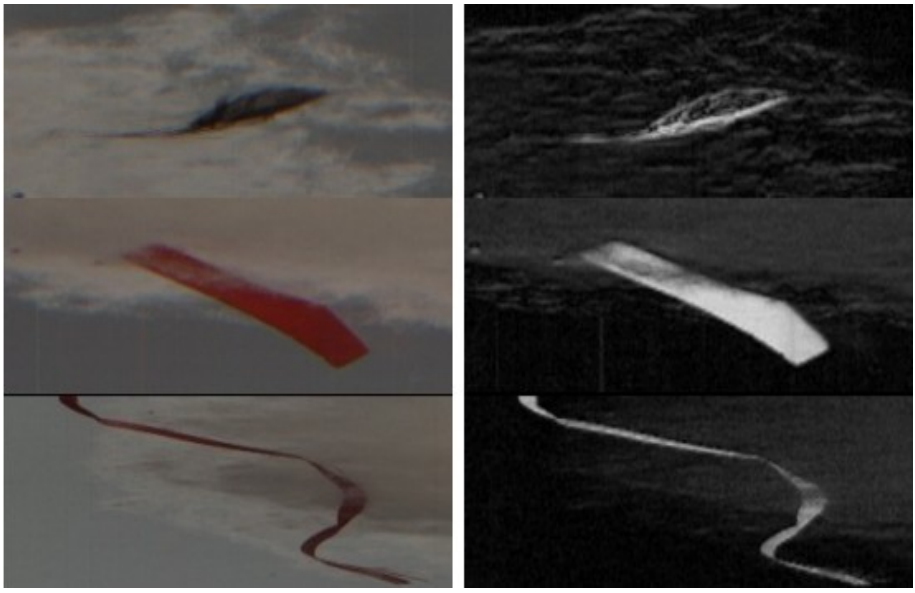
The conversion from RGB to HSV can be completed with the following formula

$$\begin{cases} H=\begin{cases}\theta & ;G\geq B\\ 2\pi-\theta & ;G\leq B\end{cases}\\ S=1-\dfrac{3}{(R+G+B)}\left[\min(R+G+B)\right]\\ V=\max(R,G,B)\end{cases} \tag{1}$$

In (1), $\theta=\arccos\left\{\dfrac{\left[(R-G)+(R+B)\right]\times 0.5}{\left[(R-G)^2+(R-B)(G-B)\right]^{\frac{1}{2}}}\right\}$, R, G, B respectively response to different color component in collected images of foreign fibers in cotton; H, S, V, respectively response to saturation and brightness component of images after converted (Wei et al.,2008).

2.4 Saturation Component Enhancing Method

The three typical types of the foreign fiber images and their saturation images are shown as Fig. 2.



(a) (b)

Fig. 2. (a) Three typical types of the foreign fiber images, (b) Foreign fiber saturation images

Image enhancement can be divided to two kinds, space domain enhancement and transform domain enhancement. Machine vision system always chooses space domain enhancement due to the requirement of speed.

Denote the saturation value of original image in the position of (x, y) to be S(x, y), the enhanced saturation value to be I(x, y)

$$I(x,y)=\sqrt[2]{\left|S(x,y)-\overline{S}\right|^3} \tag{2}$$

In the formula, $\overline{S}=\frac{1}{ab}\sum_{x,y\in(a,b)} S(x,y)$, parameter S(x, y) is known, a, b successively represent the length and width of image, as the size of collected images is 4096×128, so the formula of enhanced model can be defined as

$$I(x,y)=\sqrt[2]{\left|S(x,y)-\frac{1}{ab}\sum_{x,y\in(a,b)} S(x,y)\right|^3} \tag{3}$$

2.5 Method of Image Segmentation

The identification of cotton foreign fibers is carried out through real-time monitoring. However, image segmentation threshold value will change when image collection environment changes. Therefore, this paper chooses iteration method to calculate threshold value of foreign fibers automatically on the premise of accuracy and speed.

Image segmentation procedures of iteration method are described as follows (Deng et al., 2010).

(1) To calculate the maximum saturation value Zmax and minimum saturation value Zmin of the saturation image, we define the original thresholding $T_0=\{Tk|k=0\}$ to be

$$T_0 = \frac{Z_{\max}+Z_{\min}}{2} \tag{4}$$

(2) Segment the image into target and background according to thresholding Tk, and calculate their average saturation value (Zo and ZB).

$$Z_o = \frac{\sum_{Z(i,j)<T_k} Z(i,j)\times N(i.j)}{\sum_{Z(i,j)<T_k} N(i.j)}$$
$$Z_B = \frac{\sum_{Z(i,j)>T_k} Z(i,j)\times N(i.j)}{\sum_{Z(i,j)>T_k} N(i.j)} \tag{5}$$

Z(i , j) is the image saturation value of the position (i, j), N(i, j) is the weight coefficient of position (i, j), generally take N(i, j)=1.0.

(3) New thresholding Tk+1 is Calculated.

$$T_{k+1} = \frac{Z_O+Z_B}{2} \tag{6}$$

(4) Repeat steps (2) and (3), until the difference between Tk and Tk+1 is smaller than defined parameter or reach certain iteration times.

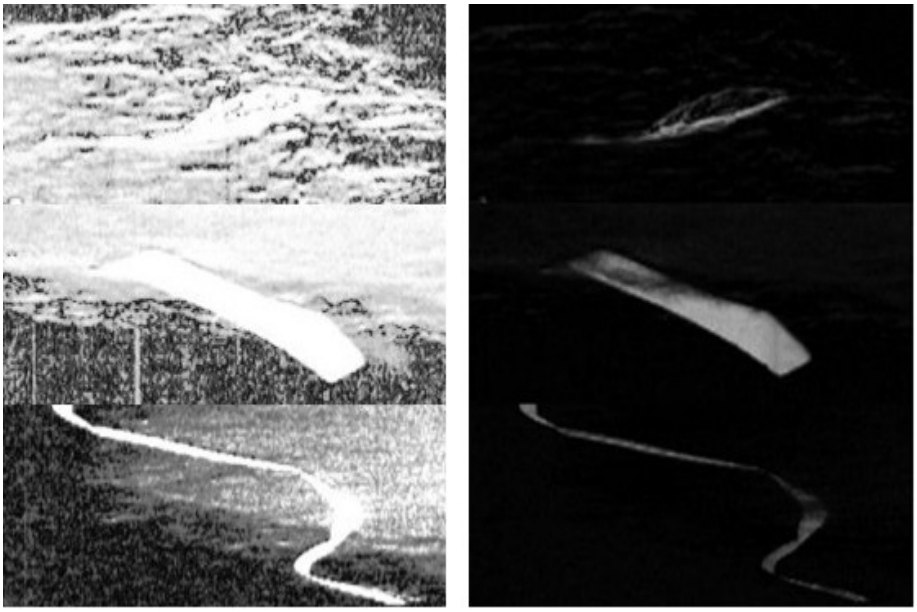
(5) Compare the thresholding Tk+1 with the saturation value of each pixel, if the saturation value is bigger than Tk+1, the pixels will be set as target, otherwise, the pixels can be considered as the background. The formula is as following

$$g(i,j) = \begin{cases} 0, f(i,j) < T_{k+1} \\ 1, f(i,j) \geq T_{k+1} \end{cases} \tag{7}$$

3 Results and analysis

3.1 Analysis of Enhanced Result of Foreign Fiber in Cotton

The images of foreign fiber shown in Fig. 2 are chosen as the input images. Fig. 3(a) is the enhanced images by histogram equalization, and Fig. 3(b) is the enhanced images based on the method of this paper.



(a) (b)

Fig. 3. (a) Enhanced images by histogram equalization, (b) Enhanced images based on the method of this paper

The results show that histogram equalization method can enhance not only the foreign fiber targets but also the background. Therefore, it cannot increase the contrast between targets and background. The difficulty of segmentation hasn't been reduced. On the contrast, conclusion can be drawn that the segmentation method which is proposed in this paper can remain targets' details, in the meantime, inhibit the background effectively.

3.2 Analysis of Image Segmentation Results

The contrastive result of 1000 images between direct iteration method and iteration based on enhancing method is shown in Fig. 4. The direct iteration method can be

success to handle the sliced foreign fibers such as paper and plastic pieces but not to the linear such as polypropylene yarn. The main reason is that the size of linear fibers in images is so small that the background is missed to recognize it as an object easily. Consequently, the approach proposed in this paper is more effective for segmentation of foreign fiber.

(a) (b)

Fig. 4. Processed images using the iteration method: (a) segmentation results of the original saturation images, (b) segmentation results of the enhanced saturation images

4 Conclusions

(1) To speed up enhancement, according to the features of foreign fiber images, we adopt air space method to enhance foreign fiber images.
(2) This research adopts iteration method to segment images and obtains accuracy segment results which avoid the inaccuracy results created by direct iteration method.

Acknowledgement. This research was supported by Agricultural Science and Technology Achievements Transformation Project Funds (2012GB23600629) and Science and Technology Planning Project of Shandong Province, China (2012GNC11202).

References

1. Li, B., Ding, T., Jia, D.: Design of a sophisticated foreign fiber separator. Agricultural Machinery Journal 37(1), 107–110 (2006)
2. Li, M.: Features and application of online foreign fiber detecting and clearing device. Shanghai Textile Science & Technology 34(1), 15–18 (2006)
3. Zhang, X., Li, D., Yang, W., Wang, J., Liu, S.: A fast segmentation method for high-resolution color images of foreign fibers in cotton. Computers and Electronics in Agriculture 78, 71–79 (2011)
4. Bakker, T., Wouters, H., van Asselt, K., Bontsema, J., Tang, L., Muller, J., van Straten, G.: A vision based row detection system for sugar beet. Comput. Electron. Agric. 60(1), 87–95 (2008)
5. Kim, B.G., Shim, J.I., Park, D.J.: Fast image segmentation based on multiresolution analysis andwavelets. Pattern Recognit. Lett. 24(6), 2995–3006 (2003)
6. Wei, J., Fei, S., Wang, M., Yuan, J.: Research on the Segmentation Strategy of the Cotton Images on the Natural Condition Based upon the HSV Color-Space Model. Cotton Science 20(1), 34–38 (2008)
7. Pichel, J.C., Singh, D.E., Rivera, F.F.: Image segmentation based on merging of sub-optimal segmentations. Pattern Recognit. Lett. 27(10), 1105–1116 (2006)
8. Deng, L., Xu, J., Cheng, X.: Application Research on Segmentation Algorithm for Sun-image Based on Iteration Threshold. Computer and Modernization 10, 72–74 (2010)

An Intelligent Four-Electrode Conductivity Sensor for Aquaculture

Jiaran Zhang[1,2,3], Daoliang Li[1,2,3,*], Cong Wang[1,2,3], and Qisheng Ding[1,2,3,4]

[1] China-EU Center for Information and Communication Technologies in Agriculture of China Agriculture University, Beijing 100083, P.R. China
[2] Beijing Engineering & Technology Research Center for Internet of Things in Agriculture, Beijing 100083, P.R. China
[3] Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, P.R. China
[4] Jiangsu Normal University, Xuzhou, Jiangsu, P.R. China
dliangl@cau.edu.cn, li_daoliang@yahoo.com

Abstract. Conductivity is regard as a key technical parameter in modern intensive fish farming management. The water conductivity sensors are sophisticated devices used in the aquaculture monitoring field to understand the effects of climate changes on fish ponds. In this paper a new four-electrode smart sensor is proposed for water conductivity measurements of aquaculture monitoring.The main advantages of these sensors include a high precision, a good stability and an intrinsic capability to minimize errors caused by polarization. A temperature sensor is also included in the system to measure the water temperature and, thus, compensate the water-conductivity temperature dependence. The prototype developed is appropriate for conductivity measurement in the range of 0-50mS/cm, 0-40 ℃. A cure relationship was found between the out-put value of each standard solution measured by the sensor and the electric conductivity concentration.

Keywords: four-electrode, water-conductivity, intelligent sensor.

1 Introduction

As we all known, fish require levels of salinity (salt), hardness (Calcium and Magnesium) and low levels of various nutrients like Phosphorus, Ammonium and nitrate to maintain their daily lives. Therefore, understanding the concentrating of these ionized chemicals in water of fishponds is critical to successful aquaculture.

Electrical conductivity is commonly used in hydroponics, aquaculture systems to monitor the amount of salts, nutrients or impurities in the water[1]. The nominal values of water-conductivity are generally used to measure the concentration of ionized chemicals in water, though it does not distinguish individual concentrations of

* Corresponding author. Postal address: P.O. Box 121, China Agricultural University, 17 Tsinghua East Road, Beijing 100083, PRC.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 398–407, 2013.

different ionic chemicals mixed in water. Water-conductivity measurements are of paramount importance in water-quality monitor of aquaculture since high or low conductivity levels, relative to their nominal values, can be used to detect the environmental changes of aquaculture fishpond.

Temperature is also an important variable when conductivity measurements are concerned. Temperature is itself a water quality parameter and an influence variable that affects conductivity measurements[2]. For raw water the temperature coefficient is about 2% per ℃. This means that acceptable conductivity measurement accuracy implies temperature measurement in order to obtain a temperature compensated conductivity measurement for a given reference temperature, typically 25 ℃ or 20 ℃.[3]

There are two main types of conductivity sensors: electrode (or contacting sensors) and toroidal or inductive sensors[4][5][6]. Electrode sensors contain two, three, or four electrodes. The conductivity (σ) is directly proportional to the conductance (1/R). The proportionality coefficient (KC) depends on the geometry of the sensor that must be designed according to the target conductivity range[7][8]. Toroidal or inductive sensors usually contain two coils, sealed within a nonconductive housing. The first coil induces an electrical current in the water, while the second coil detects the magnitude of the induced current, which is proportional to the conductivity of the solution.[9]

As far as water conductivity measurement by electrodes is concerned, several solutions have been proposed[10], and many commercial types of equipment are available from many manufacturers[11][12]. Well known limitations of a two or three electrodes sensor have been identified in literature[13]——The main problems associated with water-conductivity measurements are sensitivity to polarization effects in the situation of long time power, measurement selectivity, and too expensive price to application of aquaculture in china. Also this measurement solution implies a large number of repeated calibration procedures are required as the fouling drawback.

In this paper, the attention is focused on the smart four-electrode conductivity sensor for water-quality monitoring in aquaculture. The main advantages associated with the designed sensing unit are its high precision, an intrinsic capability to minimize errors caused by polarization. Another important advantage of the proposed solution for conductivity measurements are its simple, accurate method of temperature compensation, good linear behavior that enables the calibration of conductivity sensor to have the better accuracy by using several standard solutions. This paper includes the design of water-conductivity measurement hardware, description of temperature compensation, calibration of conductivity sensor and capability test experiments of the developed prototype.

2 Design of the Four-Electrode Conductivity Sensor

The intention of this part is to present and characterize a prototype for water-conductivity measurements based on a four-electrodes sensing unit which using in the

aquaculture. The smart sensor also includes a temperature sense part to provide compensation of conductivity measurements caused by temperature variation.

2.1 Principle of Measurement

For sensing water conductance, metal-solution interface modeling should be taken into account as shown in Fig. 1[14][15]. Assuming the behavior of electrolyte resembles a pure resistor and neglecting the polarization, the charge transfer back and forth the interface results in the reduction-oxidation reaction, and the double layer capacitive effects caused by charging of the electrode-solution at the interface. The two processes can be modeled as a resistive (R_{CT} and Z_W) and a capacitive (C_{DL}) component in parallel.

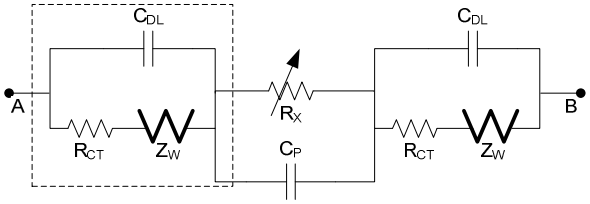


Fig. 1. Equivalent circuit for an electrochemical cell

The charge transfer resistance R_{CT} is highly nonlinear and depends on both the concentration of ions in the solution and the applied potential. The Warburg impedance Z_W is due to the ion diffusion process in proximity to the interface. The double layer capacitance C_{DL} depends on the material of the electrodes and on the ion concentration, with typical values in the range 10–40μF/cm^2.

To minimize these factors, a four-electrode cell was projected and implemented. This type of cell is like a four-terminal precision resistor: Two electrodes are used to force a uniform time varying electric field and the other two measure the voltage.

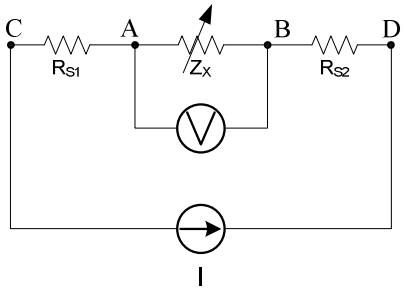


Fig. 2. Impedance sensing by four terminal sensing technique

Four-electrode water-conductivity sensors are widely used to reduce contact and interface interference in conductance measurements[16]. The primary advantage is that the current and voltage electrodes of sensors are separated, eliminating the contribution of wiring impedance and contact resistances R_S as illustrated in Fig. 2. While the alternating current running through two connections, C and D, a voltage drop can be measured across the impedance Z_x by voltage sensing in A and B connections. Disregarding wiring resistances, sensed impedance can be determined by Ohm's law as the current is known. The approach, based on current-forcing and voltage-sensing, is particularly suited for the water-conductivity sensors which are used to eliminate polarization and fouling effects in aquaculture.

2.2 Design of the Hardware

The structure of intelligent water-conductivity sensor mainly contained excitation current source model, constant-current source, conductivity measurement system, temperature sensor, signal conditioning model, power source model, RS485 field bus, microprocessor and Sensors Electronic Data Sheets (TEDS), hardware structure is shown in Fig. 3.

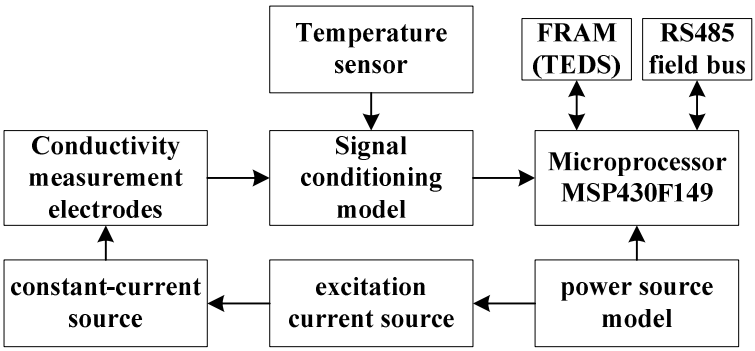


Fig. 3. Hardware structure of the sensor

Excitation current source generates equal and opposite current by means of a dual, bidirectional and single-pole/double-throw (SPDT) CMOS analog switches. It is of paramount importance in water-conductivity sensing unit. Constant current source is used to avoid the fluctuations of output current due to the change of load variations. The system is proposed in this paper adopts the following circuit as shown in Fig. 4.

The alternating current via two excitation electrodes EI1 and EI2,hence magnetic field is generated in solution, a voltage drop can be measured across the measure electrodes EV1 and EV2. Alternating voltage transform to direct current by signal conditioning model. Standardized RS-485 interface is integrated with and possesses self-recognition capability provided by its TEDS. Power source model ensure each model fully functional at supply voltages.

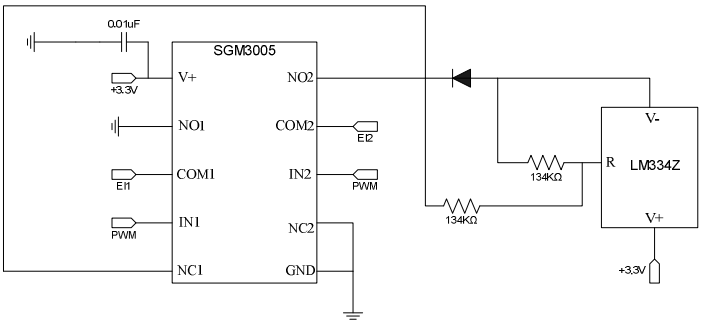


Fig. 4. Circuit diagram for the sensor

A negative temperature coefficient thermistor precision achieved 0.1℃ is adopted to temperature sensing unit.

2.3 Temperature Compensation

Temperature compensation is of important to water-conductivity measurement as well. Generally, the conductivity of common electrolytes increases with increasing temperature. In this paper, a valid temperature-compensated method by software is given out.

The measurements for different concentrations of KCl confirm the linear variation of the conductivity with temperature:

$$\sigma_t = \sigma_{ref} \times \left[1 + \alpha \cdot \left(t - t_{ref}\right)\right] \tag{1}$$

where σ_t is the conductivity at any temperature t (in degrees celcius), σ_{ref} is the conductivity at the reference temperature t_{ref} (in degree celcius), and α is the temperature coefficient of the solution at t_{ref} .α can be calculated of every solution at the reference temperature t_{ref} by measured the values of σ_t, σ_{ref} and t. A temperature value of 20 ℃ is chosen at the temperature reference in this paper.

3 Experiment

The basic experiments were conducted to evaluate the characteristics of the water-conductivity sensor. All solutions were prepared with distilled water and analytical grade chemicals. The KCl solutions of different electric conductivity were prepared. All experiments were carried out in the lab. In addition to distilled water, five KCl solutions of different concentrations are presented as standard solutions to calibrate the water-conductivity of the sensor. They are 4.6mS/cm, 10.4mS/cm, 20.7mS/cm, 32.1mS/cm and 52.6mS/cm.

After the calibration was completed, measurements are repeated 10 times in standard solutions of 4.10mS/cm and 9.20mS/cm every ten minutes to test the reproducibility of sensor. Besides, the experiments of accuracy analysis, stability test and effect of temperature compensation have be done by measure different conductivity KCl solutions.

4 Result and Discussion

The results of experiments were listed to evaluate the characteristics of the water-conductivity sensor at this part. By these experimental data, the characteristics of conductivity curve, reproducibility, accuracy, stability and effect of temperature compensation are verified.

4.1 Water-Conductivity Sensor Calibration

The curvilinear relation of out-put voltage values changing with different of conductivity at the reference temperature(20℃) is presented respectively in Fig 5. The conclusion that a better measurement effect the sensor has when the conductivity is lower than 32.1mS/cm, a low resolution at 32.1mS/cm to 52.6mS/cm the sensor has can be draw from the Fig.5. It indicates that the low sensitivity the sensor has when the water-conductivity is higher than 32.1mS/cm.

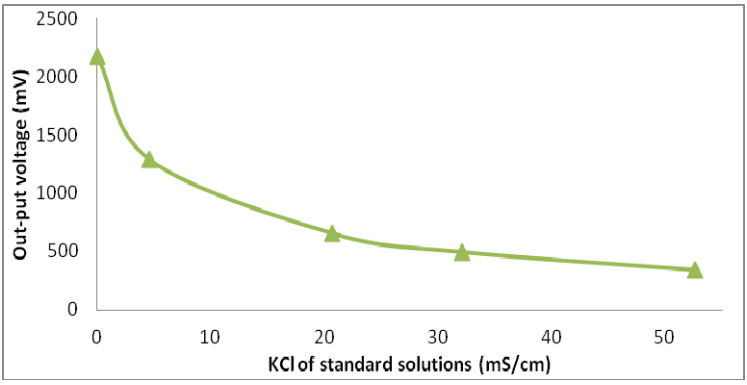


Fig. 5. Curve between out-put voltage and KCl standard solutions

4.2 Reproducibility of the Sensor

Fig. 6 shows the results of repeatedly measuring the water-conductivity of the sensor. As the Fig. shows, the measurement reproducibility is very high and the absolute measurement error is between -0.04~ +0.05 for these samples.

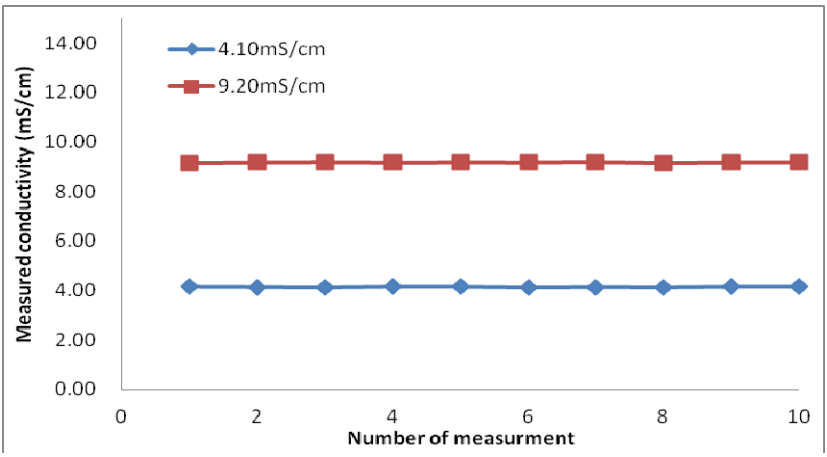


Fig. 6. Reproducibility of measured water-conductivity

4.3 Accuracy Analysis

Measurement accuracy reflects the closeness between the measurement result and the true value of the measure. Table 1 shows the results of continuous monitoring of KCl standard solutions (0.5, 10.0, 20.7, 32.1 and 50.0mS/cm) with the sensor. As the table shows, the measurement accuracy is very high and the relative measurement error is within ±1.5% for all these samples.

Table 1. The results of accuracy testing

Water sample	Measurements								Average value	Absolute measurement error	Relative measurement error
	1	2	3	4	5	6	7	8			
0.50	0.50	0.49	0.51	0.47	0.47	0.52	0.48	0.5	0.49	0.01	1.50%
10.00	9.91	9.93	9.90	9.92	10.00	10.06	10.08	9.96	9.97	0.03	0.30%
20.70	20.57	20.77	20.89	21.00	20.38	20.62	21.04	20.71	20.75	0.05	0.23%
32.10	31.77	31.87	32.12	32.29	32.01	32.05	32.31	32.01	32.05	0.05	0.15%
50.00	50.04	50.44	50.04	49.34	50.46	50.02	50.18	50.2	50.09	0.09	0.18%

4.4 Stability Test

Generally stability refers to the ability of metrological characteristics of the measuring instrument does not change with time. Fig. 7 shows the results of measuring the water-conductivity of KCl standard solution (4.20mS/cm, 9.20mS/cm) using the sensor. Measurements of one half day at 5 min intervals were recorded. Results show

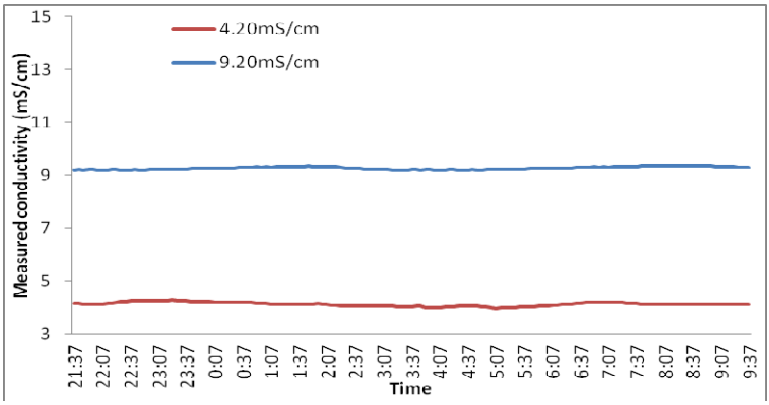


Fig. 7. Experimental results of Stability

no significant change and the variation is within ± 0.2mS/cm for each samples. The sensor gives a stable output value.

4.5 Effect of Temperature Compensation

The effect of temperature compensation in accordance with the method mentioned in this paper is shown in Fig. 8. Measurements of one half day at 5 min intervals were recorded for KCl standard solution of 4.20 and 9.20mS/cm. As the Fig. shows, the effect of temperature compensation is relative effective at 0~30℃, however, the values of measured conductivity are absolutely lower than the actual value at 40 ℃. The data indicates that the method of temperature compensation have a certain error when the temperature above 30℃.

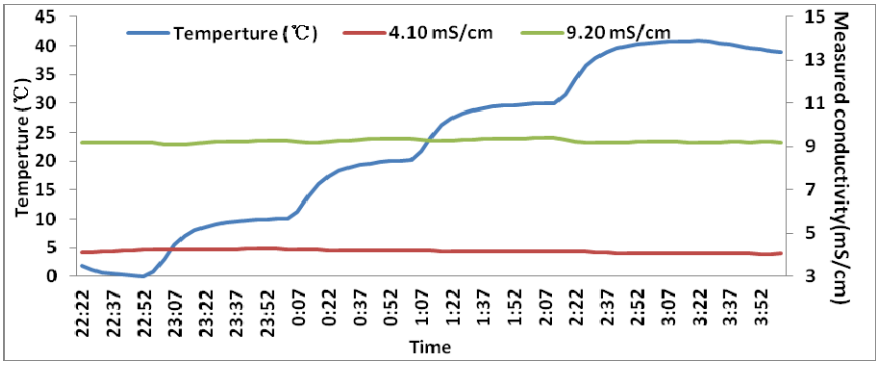


Fig. 8. Effect of temperature compensation

5 Conclusion

The proposed prototype is an attractive solution for water quality measurements systems in fishery. The sensor is based on a novel four-electrode conductivity cell that eliminates errors caused by fouling and polarization. Main characteristics of the proposed prototype include an automatic temperature compensation of conductivity measurements, a low sensitivity to disturbances caused by electrolytic polarization, double layer. Measurement system's conditioning signal circuitry and digital signal processing assure an appropriate conductivity measuring range, good measurement reproducibility, accuracy and stability.

Future works will focus on improving the resolution at the solution of high conductivity so that expanding measurement range, in addition, a more simple and effective method of temperature compensation needs to be presented.

Acknowledgments. The authors would like to thank Development and Applications of sensor network applied to monitor bloom of blue-green algae in Taihu lake (2010ZX03006-006), and Beijing Natural Science Foundation "Integrations methods of digitalization technologies in intensive fish farming" (4092024) for their financial support.

References

[1] Wei, Y., Ding, Q., Li, D., Tai, H., Wang, J.: Design of an Intelligent Electrical Conductivity Sensor for Aquaculture, pp. 1044–1048 (June 2011)
[2] Standard Test Methods for Electrical Conductivity and Resistivity of Water, D1125, American Society for Testing and Materials, Conshohocken, PA
[3] Komarek, M., et al.: A DSP Based Prototype for Water Conductivity Measurements. In: Proceedings of the IEEE Instrumentation and Measurement Technology Conference on IMTC 2006 (2006)
[4] Fougere, A.J.: New non-external field inductive conductivity sensor (NXIC) for long term deployments in biologically active regions. In: OCEANS 2000 MTS/IEEE Conference and Exhibition, vol. 1, pp. 623–630 (September 2000)
[5] Karbeyaz, B., Genr, N.: Electrical Conductivity Imaging via Contacless Measurements: An Experimental Study. IEEE Transactions on Medical Imaging 22(5), 627–635 (2003)
[6] Ripka, P.: Advances in Fluxgate Sensors. Sensors and Actuators A 106, 8–14 (2003)
[7] Orion, T., Webster, J.: The Measurement, Instrumentation, and Sensors Handbook. CRC Press (1999)
[8] Keitlley.: Four-Probe Resistivity and Hall Voltage Measurements with the Model 4200-SCS. App. Note 2475 (2002)
[9] Ramos, P.M., et al.: A Four-Terminal Water-Quality-Monitoring Conductivity Sensor. IEEE Transactions on Instrumentation and Measurement 2008 57(3), 577–583 (2008)
[10] Stogryn, A.: Equations for calculating the dielectric constant of saline water. IEEE Trans. Microw. Theory Tech. MTT-19(8), 733–736 (1971)
[11] TBI-Bailey Controls.: Process monitoring instruments, Product Specification—E67-23-1
[12] http://www.coleparmer.com

[13] Kisza, A.: The capacitance of the electric double layer of electrodes in molten salts. Electroanal. Chem. 534(2), 99–106 (2002)
[14] Hyldgard, A., Mortensen, D., Birkelund, K., Hansen, O., Thomsen, E.V.: Autonomous multi-sensor micro-system for measurement of ocean water salinity. Sensors and Actuators A 147, 474–484 (2008)
[15] Bard, A.J., Faulkner, L.R.: Electrochemical methods, 2nd edn. John Wiley and Sons (2001)
[16] Crescentini, M., Bennati, M., Tartagni, M.: Design of integrated and autonomous conductivity–temperature–depth (CTD) sensors. International Journal of Electronics and Communications (2012)

Integration and Development of On-Site Grain Yield Monitoring System Based on IPC

Xiang Guo, Lihua Zheng*, Xiaofei An, Jia Wu, and Minzan Li

Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agriculture University, Beijing 100083, China
zhenglh@cau.edu.cn

Abstract. Variable Rate Technology (VRT) is the goal of precision agriculture. In ordet to realize VRT, it is necessary to measure yield accurately and build yield map in real time. A grain yield monitoring system based on IPC (Industrial Personal Computer) was integrated and developed to fit for China actual conditions and combines. The system consisted of a hardware part and a software part. The hardware part included IPC, CAN-bus module, GPS receiver, GPRS transmission module, and yield monitoring sensors. The CAN-bus module, GPS receiver and GPRS transmission module were integrated with IPC. The yield monitoring sensors included a dual-plate differential flow transducer, a grain temperature sensor, a grain humidity sensor, a ground speed transducer, a grain elevator speed sensor and a header height sensor, which were connected to the CAN-bus module. The software part included five main modules, device settings module, data receiving module, data delivering module, data processing module, and data analyzing module. Firstly, the system used CAN-bus technology to construct a sensors network so that the signals of the grain flow, grain temperature and humidity, ground speed, elevator speed and the header height could be sent to the CAN-bus via CAN converting module. Then, the IPC collected data packets on the CAN-bus and recorded and displayed the yield data after analyzing and calculating through yield models. At the same time, IPC collected geographic information from GPS receiver and stored it with the yield data from CAN-bus on this location together as one record. Finally, the system uploaded the collected data to the host server in real-time or packaged it via GPRS. The experiments showed satisfactory results.

Keywords: IPC, grain yield monitoring system, CAN-bus, GPS, GPRS.

1 Introduction

With the rapid growth of world population, the demand for food increases day by day. Precision agriculture technology has been showing powerful superiority[1]. It is necessary to develop the precision agriculture mechanized equipment supporting for the implementation of precision agriculture. The harvester with the grain yield monitoring system can collect geographic information via GPS receiver, provides real-time production data intuitively and accurately by using yield map. It provides the necessary information supporting for decision-making of irrigation, fertilization,

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 408–417, 2013.

planting and pesticides spraying according to spatio-temporal variety, hence farming inputs will be much saved, the environment pollution will be reduced, the costs will be reduced and the land yields will be increased[2,3].

Precision agriculture has already made a considerable development since the late 1980s. In recent years, thousands of professional academic researches on Precision Agriculture have been carried out. Leading by the United States, Britain, Canada, Australia, Japan and other developed countries, the developing countries such as Brazil, Malaysia, and China have launched many research projects on precision agriculture technology, sucn as development of PA equipment and demonstration projects. On the other hand, up to now no commercial grain yield monitoring system was used to monitor grain on-site harvest in China [4].

The combine with GPS positioning system and yield sensors has been produced by a couple of large international agricultural machinery manufacturing enterprises, and the information processing system has also been developed[5]. These systems were generally compatible with the GPS receiver and memory card to enable real-time yield monitor and generate yield map automatically[6]. However, importing overseas system is higher cost, and meanwhile the yield monitoring software is just appropriate to the corresponding foreign harvesters so that it is difficult to match with the Chinese domestic harvester models. And it is so difficult for ordinary Chinese users to operate the yield monitoring system because the overseas system interface is all operating in English. There is an urgent need to develop a lower-cost, on-site monitoring system fitting for the conditions of domestic combine and needs of local farmers in China.

2 Integration and Development of the Grain Yield Monitoring System

2.1 Hardware Structure

The hardware based on IPC integrated three modules, CAN-bus module, GPS module and GPRS module. The CAN-bus module connected with all sensors needed by the yield monitoring system, GPS module received position information in real time, and GPRS module communicated with the host server[7,8]. The system structure is shown in Fig.1.

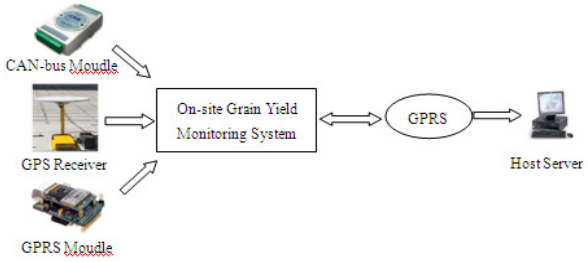


Fig. 1. System Structure

2.2 Hardware Integration and Development

The grain yield monitoring system was compose of the IPC, USB-CAN interface card, analog signal-CAN converter module, digital signal-CAN converter module, conditioning circuit, power conversion module, GPS receiver, GPRS module and yield monitoring sensors[9]. The hardware integration scheme is shown in Fig.2.

Fig.3 shows the modules need to be integrated together, including IPC mainboard, GPS receiver, GPRS module and CAN-bus module. Fig.4 shows the yield monitoring sensor group including the dual-plate differential flow transducer, the grain temperature sensor, grain humidity sensor, ground speed transducer, grain elevator speed sensor and header height sensor[10]. The system integration schematic is shown in Fig.5.

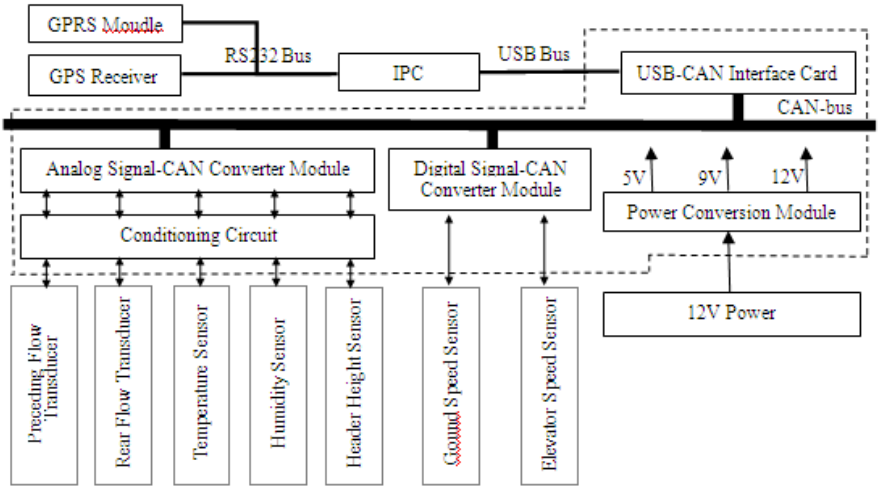


Fig. 2. Hardware integration scheme

(a) ARM9 IPC mainboard (b) GPS Receiver (c) CAN-bus Module (d) GPRS Module

Fig. 3. The grain yield monitoring system integration modules

(a) Flow Transducer (b) Header Height Sensor (c) Temperature/Humidity Sensor (d) Hall Sensor

Fig. 4. Yield monitoring sensor group

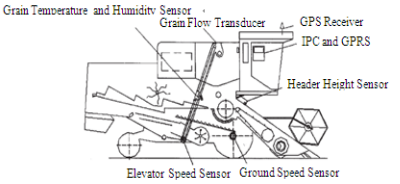


Fig. 5. System integration schematic[11]

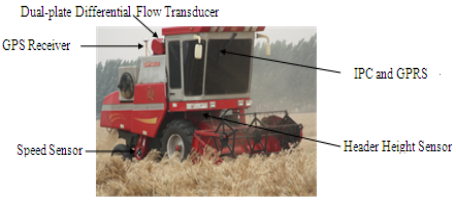


Fig. 6. On-site winter wheat harvesting[12]

Firstly, via CAN-bus the system got the signal from the header height sensor to determine whether the system started to work. Secondly, yield relevant parameters were collected through CAN-bus including the grain flow coming from the dual-plate differential flow transducer, the grain temperature and humidity from temperature and humidity sensor, harvester speed from the ground speed transducer and rotate speed of elevator from the elevator speed sensor. And then the grain yield was calculated using the data collected from those sensors. Finally, on-site grain yield map was drawn using the yield and the real-time geographic location data getting from GPS receiver. Fig.6 shows on-site winter wheat harvesting in Huantai County, Shandong Province in North China on 13th of June, 2012.

3 Software Development and Implementation

The software system was designed running on the IPC. It was embedded stand-alone structure. The data outflowing from the CAN-bus and GPS flowed into the on-site grain yield monitoring application to be processed, analyzed, stored and uploaded to the yield server through GPRS module. Yield data were designed to be sent to the server every 30s by default or at anytime by clicking the corresponding function button. They were packaged as a file to be sent after harvesting finished. The data flow diagram of the system is shown in Fig.7.

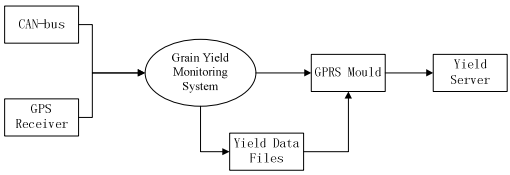


Fig. 7. Data flow diagram of system software

3.1 System Design

The system included five main functions: device settings, data receiving, data processing, data delivering and data analyzing. The usecase diagram of the system is shown in Fig.8.

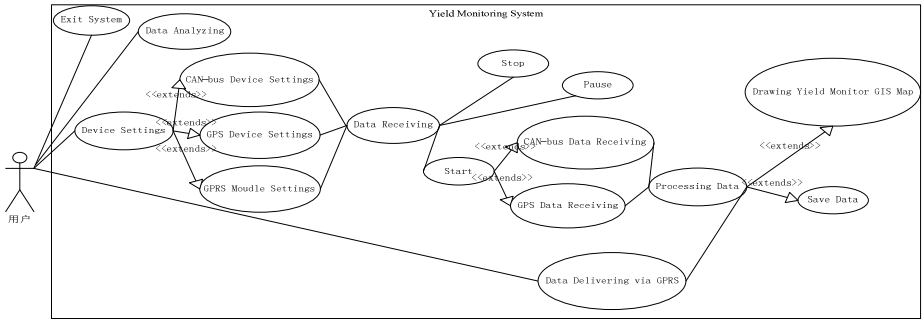


Fig. 8. Use case diagram of the system

Device settings function set three devices (CAN-bus module, GPS receiver, GPRS module) up; Data receiving function collected and analyzed data from two devices (CAN-bus module, GPS receiver); Data processing function calculated and saved yield data and drew yield GIS map; Data delivering function sent yield data to the yield server; Data analyzing function translated binary yield data to text data. Accordingly five modules were designed to fulfill four functions, as shown in Fig.9.

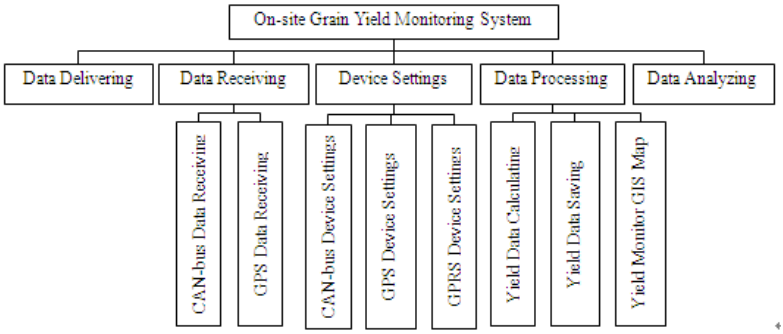


Fig. 9. Software system modules

3.2 CAN-Bus Data Designing and Analyzing

CAN-bus is a kind of bus with multi-master mode serial data communication. The short frame structure of CAN-bus does not occupy the bus for too long and the interference keeps lower. Thus it ensures real-time communication and fits for field work[13]. CAN-bus is easy to construct an underlying control network of open, digital, multi-point communications. Compared with traditional distributed control system, it has more advantages of all digital, distributed control, two-way transmission, openness, etc.[14]. The data frame issued by CAN converter module is shown in the Tab.1.

Table 1. Data frame resolving

CAN Data	Analytical Content
0010	Command Frame,“00”is a flag, “10” means 16B。
408096BF96B914A6	Response Frame,“40” is a flag, “8096” is data from channel 0. “BF96” is data from channel 1. “B914” is data from channel 2. “A6” is upper byte of channel 3.
819680A787A28000	Response Frame, “81” is a flag, “96” is lower byte of channel 3. “80A7” is data from channel 4. “87A2” is data from channel 5. “8000” is data from channel 6.
C1800D	Response Frame, “C1” is a flag, “800D” is data from channel 7.

Assuming that the actual voltage was V_m, the measured voltage was A_{data} and K_r was the coefficient in corresponding measurement range. The baseline value was set as 0x8000 according to the CAN module provider’s specification[15]. When A_{data}>0x8000, the actual voltage was positive and calculated using formula (1). When A_{data}<0x8000, the actual voltage was negative and calculated using formula (2). The coefficient of K_r was defined as shown in Tab.2.

$$V_m = \frac{(A_{data} - 0x8000)}{0x8000} \times K_r \tag{1}$$

$$V_m = (-1) \times \frac{(0x8000 - A_{data})}{0x8000} \times K_r \tag{2}$$

Table 2. Coefficient of K_r

Measuring Rage	Coefficient of K_r	Measuring Rage	Coefficient of K_r
±150mV	0.15625	±2.5V	2.5
±500mV	0.625	±5.0V	5
±1.0V	1.25	±10.0V	10

3.3 System Development and Implementation

The main flow chart of the system is shown in Fig.10. The main interface is shown in Fig.11.

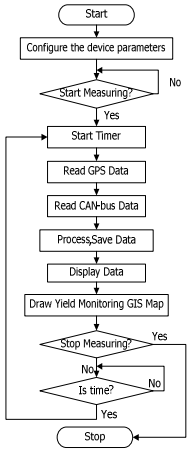


Fig. 10. Main flow chart of the system

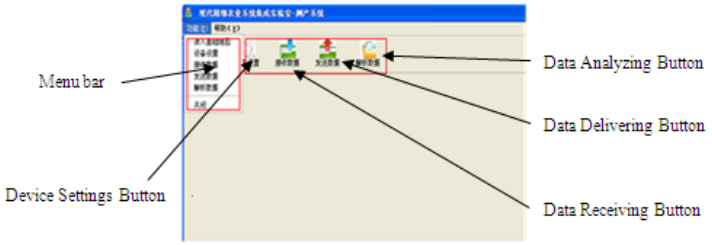


Fig. 11. The main interface

3.3.1 Device Settings Module

The Device settings module was composed of CAN-bus device setting, GPS device setting and GPRS device setting. The interface of the device settings module is shown in Fig.12.

Fig. 12. Interface of device settings

(1) CAN-bus device Setting

Configuring the parameters of CAN-bus made CAN-bus device ready to receive the data from each sensor. It was implemented by calling the corresponding API functions provided by the hardware device. The flow chart of CAN-bus configuration is shown in Fig.13(a).

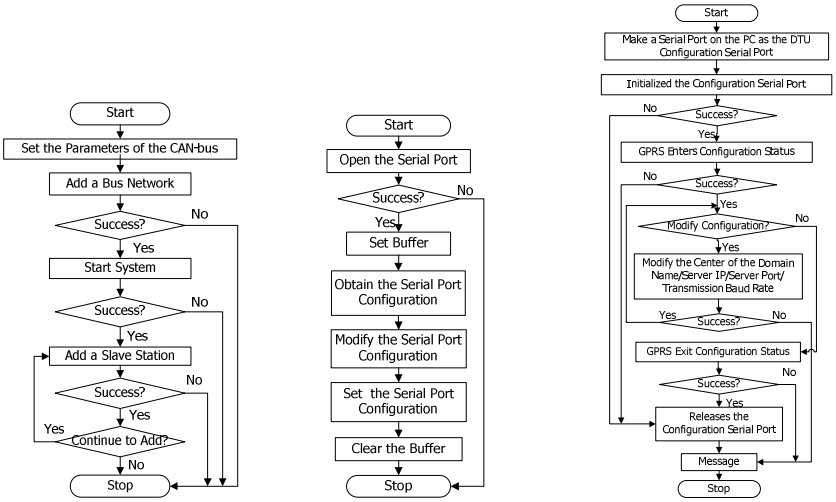
(2) GPS device setting

A class (CserialPort) was developed to operate GPS device. The GPS device configuring made the GPS receiver ready to receive the GPS data. The flow chart is shown in Fig.13(b).

(3) GPRS device setting

Configuring the GPRS device made the IPC connect with the server. DTUSet.dll, DTUSet.lib and dtucfgheader.h header files were added into the project and the API functions provided by them were used to make a serial port as the DTU configuration

serial port in the IPC, and then the center of the domain name, server IP, server port, transmission Baud rate were set to configure the GPRS device. The flow chart of GPRS device configuration is shown in Fig.13(c).



(a) CAN-bus configuration (b) GPS device configuration (c) GPRS device configuration

Fig. 13. Flow charts of device settings module

3.3.2 Data Receiving Module

Firstly, the system used CAN-bus technology to compose a sensors network and the signals were sent to the CAN-bus via CAN converting module. Then the IPC collectd data packets from the bus. The on-site yield monitoring system received data at the same frequency as CAN-bus's sampling. The Flow chart of data receiving module is shown in Fig.14, and the interface is shown in Fig.15.

3.3.3 Data Analyzing Module

The received data were stored in .dat files in binary. The system translated the .dat file into a readable file and saved in .txt format so that it was convenient for farmers to read the data. The flow chart of data analyzing module is shown in Fig.16.

3.3.4 Data Delivering Module

By default, a set of yield data was sent to the host computer per 30s via GPRS. It also could be sent by clicking button at any time, or sent in data package. The result on server side is shown in Fig.17.

3.3.5 Yield GIS Map Drawing

In the data receiving interface (Fig.15), the on-site yield GIS map was drawn in real time. In the map, GPS coordinates were transformed into the system interface coordinates and the depth of red color was used to represent the measured yield. The Flow chart of yield GIS map drawing is shown in Fig.18.

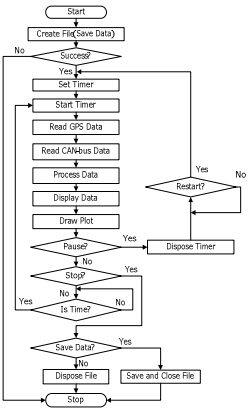


Fig. 14. Flow chart of data receiving module **Fig. 15.** Interface of Data Receiving module

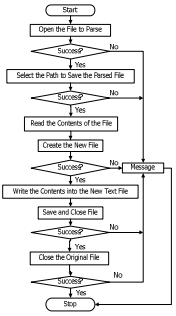


Fig. 16. Data analyzing module **Fig. 17.** Server side test result

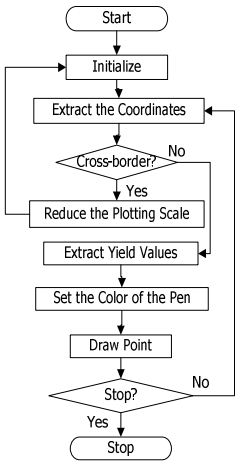


Fig. 18. Yield GIS map drawing

4 Conclusions

According to the conditions and actual needs of the domestic harvesting machinery in China, an on-site grain yield monitoring system was developed based on the vehicle IPC. It could collect yield data from CAN-bus sensor network in real-time, acquire GPS data from GPS receiver and delivery the yielding data to the server via GPRS. The winter wheat yield model was embedded in the system. Moreover, it could provide on-site yield GIS map plotting on the IPC.The on-site grain yield monitoring system was running in the winter wheat field in north China. And the result showed that the system could meet on-site yield monitoring and yield visualization needs.

Acknowledgements. This research was supported by National High-tech R&D Program of China (Contact Number: 2012AA101901), National Science and Technology Support Program (2012BAH29B04, 2011BAD21B01).

References

[1] Li, B.-Q., Tian, H., Fan, L.-H., et al.: Research advancement of precision agriculture. Joural of HeNan Agricultural University 4(38), 1 (2004)
[2] Wang, F.-H., Zhang, S.-Z.: Reserch Progress of the Farming Informatiion Collections Key Technologies on Precision Agriculture. Transactions of the Chinese Society for Agricultural Machinery 39(5), 1 (2008)
[3] Ma, L.-J.: Research on an Intelligent Mointor System of Grain Combine Harvester. China Agricultural University, Beijing (2006)
[4] Zhao, W.-Y., Yang, S.-M., Yang, Q., Yang, S.-C.: The Development and Thinking of the Precision Agriculture Technology. Journal of Agricultural Mechanization Research (4), 3 (2007)
[5] Jie, Z., Liu, H.-J., Hou, F.-Y.: Research advances and prospects of combine on precision agriculture in China. Transactions of the CSAE 21(2), 1–3 (2005)
[6] Chen, S.-R.: Research and Development of the Grain Combin Harvester yield Monitor System. In: The 14th National Combine Harvester Technology Development and Market Dynamics Seminar, p. 3, 4 (2007)
[7] Wang, Q.: Research on an Intelligent Monitor for Yield Mapping System of Grain Combine Harvester. China Agricultural University, Beijing (2005)
[8] Zhang, Z.-Y., Yu, L.-P., Li, L.-G.: GPRS Communication Technology. Digital Technology and Application (6), 1 (2011)
[9] Wang, M.-H.: Precision Agriculture, pp. 132–134. China Agricultural University Press (2011)
[10] Wu, G.: Development of Grain Yield Monitor System with CAN-bus. China Agricultural University, Beijing (2011)
[11] Wang, B.: Development of Advanced Sensing Technology of Grain Yield Monitor System. China Agricultural University, Beijing (2009)
[12] Zhuang, W.-D., Wang, X.: Studies on the GPS Data Processing and Absolute Positional Accuracy. Jounrnal of Heilongjiang August First Land Reclamation University 15(2), 1–2 (2003)
[13] Wang, J.-B., Xu, B.-G.: Analysis and Online Evaluation of CAN Message Real-time Performance. Control and Decision 22(4), 1–4 (2007)
[14] Tuo, B., Ju, H.: Design of General Data Acquisition Equipment Software System Based on Filed-bus. Journal of ChengDu University of Information Technology (21), 1 (2006)

Development and Performance Test for a New Type of Portable Soil EC Detector

Xiaoshuai Pei, Lihua Zheng, Yong Zhao, Menglong Zhang, and Minzan Li*

Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China
limz@cau.edu.cn

Abstract. The soil electrical conductivity (EC) refers to the capability for soil to conduct current. It is a comprehensive reflection of soil salinity and moisture. Therefore, acquiring soil EC rapidly and accurately can provide better guidance for farming production. Based on improving four-electrode method, a new portable soil EC detector with six electrodes was developed and its performance was tested. Inside two electrodes and outside two electrodes were used to measure soil EC near the surface and in deeper soil, respectively. And middle two electrodes were used to input a constant current to soil. The stability tests of the current source showed that the amplitude fluctuation was less than 3%.

Keywords: Soil EC detector, Four-electrode method, Precision agriculture.

Introduction

China is the largest agricultural producer and consumer in the world. With the rapid development of agriculture, large amounts of chemical fertilizers and pesticides were used to increase the crop yield. It has resulted in soil erosion, agricultural pollution and groundwater pollution and decreased soil productivity consequently. Precision agriculture, as a new farming approach, aimed at avoiding waste and excessive fertilizer spraying caused by blind input in farmland[1,2].

Precision agriculture technology is an information-based agricultural management system. Implementing spatial-temporal variable management could achieve true sense of the "intensive farming" based on information and advanced technology[3]. Soil researches has shown that soil electrical conductivity (EC) contains a wealth of information for analyzing soil nutrients and physicochemical properties[4] and can reflect the soil salinity, moisture, organic matter content, soil texture, structure and porosity, etc. Acquiring soil EC effectively is of great significance to determine the spatial-temporal distribution of soil parameters[5].

In this paper, we tried to improve the traditional current-voltage four-electrode method to better fit for in-situ measuring, and aimed at developing a in-situ soil EC

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 418–424, 2013.

detector with low price, easy operation, high measurement precision, integral control procedures and data processing procedures.

1 Materials and Methods

1.1 Principle of Soil EC Detector Development

Figure 1 is a typical structure of the current-voltage four-electrode approach, including two current electrodes (J and K) and two voltage electrodes (M and N). By loading constant amplitude current as excitation signal through two current electrodes, the soil EC can be calculated according to the voltage collected from two voltage electrodes.

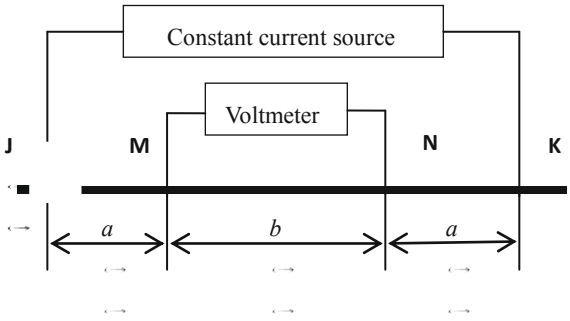


Fig. 1. Principal diagram of current-voltage four-electrode method

The earth is complex object with uncertain cross-sectional area and length. Researches showed that the earth EC could be measured according to formula (1)[6]

$$\sigma = \frac{\left(\frac{1}{d_{JM}} - \frac{1}{d_{JN}}\right) - \left(\frac{1}{d_{KM}} - \frac{1}{d_{KN}}\right)}{2\pi} \frac{I}{V_{MN}} = k(a, b)V_{MN}^{-1} \quad (1)$$

where, σ (with unit of S/m) is the value of EC ; I (with unit of A) is the current provided by the constant current source. V_{MN} (with unit of V) is the voltage measured between the M and N electrode; $k(a,b)$ (with unit of m) is the function with variables of d_{JM}、d_{JN}、d_{KM} and d_{KN}; Besides, $a=d_{JM}=d_{KN}$, $b=d_{MN}$.

To measure the EC in deep soil layer, two electrodes were added on the basic four electrode theory, which is shown in Figure 2. Moreover, accurate circuits were designed to ensure constant alternating current could be loaded between two current electrodes, and the voltage of inside and outside electrodes could be collected accurately.

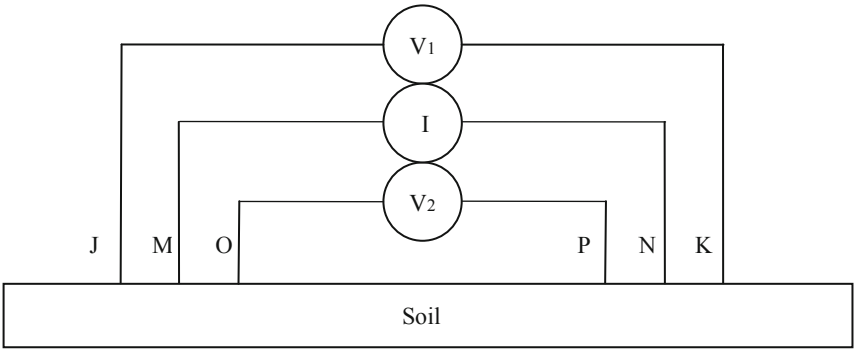


Fig. 2. Schematic of improved current-voltage four-electrode method

2 Results and Discussions

2.1 Soil EC In-Situ Detecting System Design

In-situ measurement system of soil EC was designed as shown in Figure 3. It mainly includes signal generating circuit, output signal conditioning circuits and data acquisition circuit.

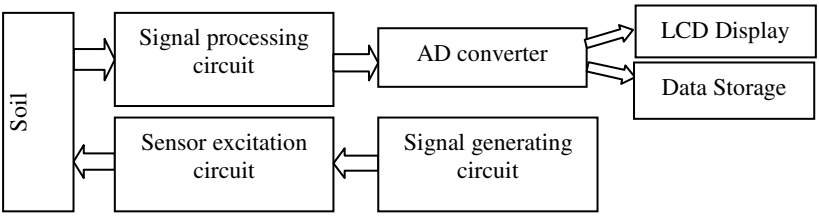


Fig. 3. System diagram

2.2 Circuits Development

2.2.1 Sinusoidal Signal Generating Circuit

ICL8038 was used as waveform generator to output sinusoidal signals. After processed by LM324, the circuit produced stable alternating current source avoiding distortion caused by oversized signal.

Figure 4 is 300Hz sinusoidal signal generating circuit. C_{15} and R_2 were used to implement low-pass filter, playing the role of cutting off direct current and conducting alternating current. Resistors of (R_{14}+R_x) (R_{15}+R_{13}-R_x) were used to adjust the amplitude of current source of I_1, I_2 and duty cycle square wave. Tests indicated that in

order to minimize the total harmonic distortion of signal (TDH), it was necessary to keep (R_{14}+Rx) and (R_{15}+R_{13}-R_x) equal. R_{16} was used to adjust the shape of the sinusoidal signal. Tests showed that when R_{16}=82 kΩ, TDH kept the minimum. C_3 and (R_{14}+R_x) were used to adjust the frequency of output sinusoidal signal calculated by formula (2).

$$f=0.3/[(R_{14}+R_x)]/C_3 \tag{2}$$

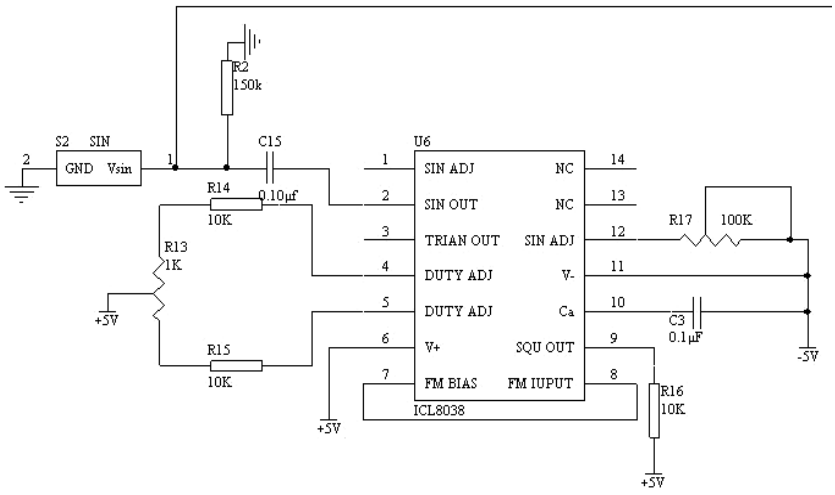


Fig. 4. Sinusoidal signal generating circuit

2.2.2 Constant Current Source Circuit

Experiments and literatures suggested that the soil EC measuring accuracy mainly depended on the constant amplitude alternating current source when using current-voltage four-electrode method[7,8]. Figure 5 is the circuit diagram of stable alternating current source controlled by LM324. LM324 series are devices of the four op amp with differential input and have some advantages compared with standard op amp. With 3~32V supply, the quiescent current is only one fifth of the MC1741's. In digital systems, it can easily provide the necessary interface circuitry without extra supply. In Figure 5, the benchmark (pin 3 in LM324, the input signal) of the AC source is Sinusoidal signal, which is the output signal from the sinusoidal signal generator ICL8038.

This circuit was with high output impedance and the output current could be controlled by programming digital potentiometer. It could be automatically adjusted with different impedance conditions and designed to fit for the soil EC with low regularity and wide changes, so that the system could reach high accuracy without changing the system parameters artificially.

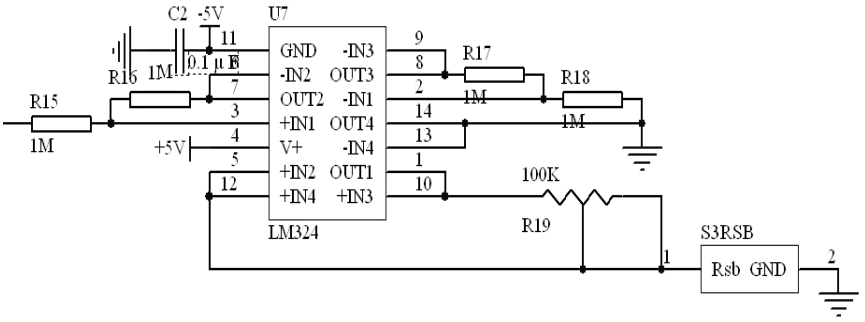


Fig. 5. Constant current source circuit

2.2.3 Differential Circuit

Unbalanced texture and structure of soil are likely to cause the instability of the signal. A subtraction circuit was designed to reduce signal instability, remove unreasonable signal, enhance useful signal and improve the accuracy. The differential circuit is shown as Figure 6. It was the combination amplifying circuit of inverting input and non-inverting input, and implemented the subtraction between Vin+ and Vin-. In the ideal conditions, two input voltages of the op amp were equal. It meant that there were virtual short and common-mode voltages between both ports of the op amp. When R_5=R_6 and R_7=R_8=R_9=R_{10}, *V*in can be calculated by formula (3).

$$V\text{in}=1+2R_{17}/R_{16}[(V\text{in}+)-(V\text{in}-)] \tag{3}$$

Because of common-mode voltage, the op amp with higher common-mode rejection ratio should be chosen and the bias voltage of the op amp should be taken into consideration. In the case of the above two factors and cost reason, LM358 and OP07 were used in this paper. This circuit was with high input impedance and low output impedance, which is suitable for the signal conditioning circuit.

2.3 Design and Development of In-Situ Soil EC Detector

The electrode and structure and structure of the detector, signal generating circuit, signal conditioning circuit, LCD module and storage module were designed and developed. In addition, the system was set aside many peripherals used for future extensions.

2.3.1 Hardware Integration and Development

The structure of the in-situ soil EC detector is shown in Figure 7. The instrument consisted of handle, straight pipe, electrodes, data acquisition controller and other components. After measuring the voltage of the electrodes, the detector calculated EC, and then displayed and stored data.

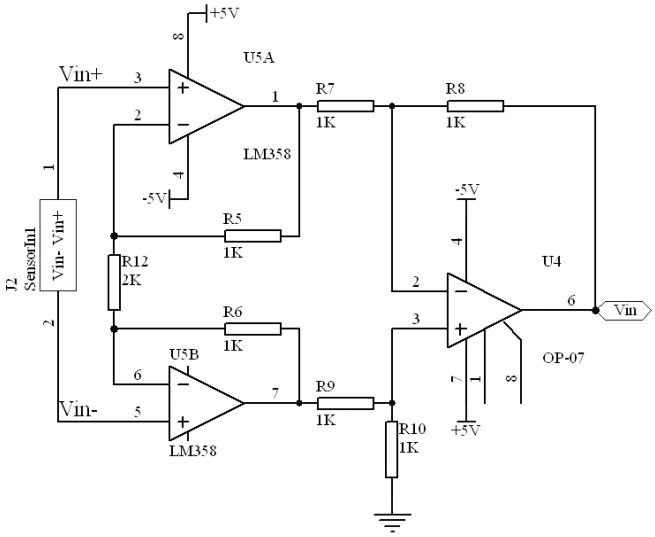


Fig. 6. Differential circuit

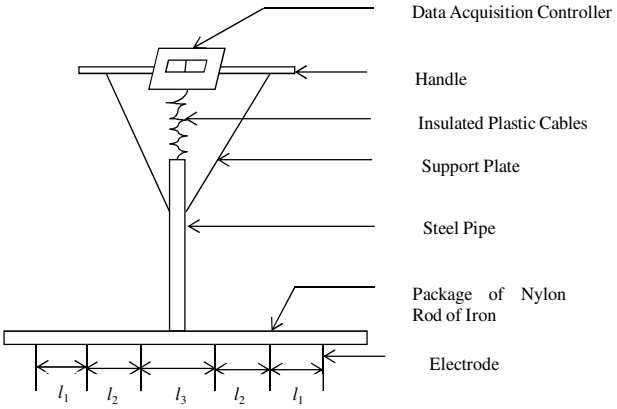


Fig. 7. Schematic of the in-situ soil EC detector

2.3.2 Software Design and Development

The overall software flow chart is shown in Figure 8.

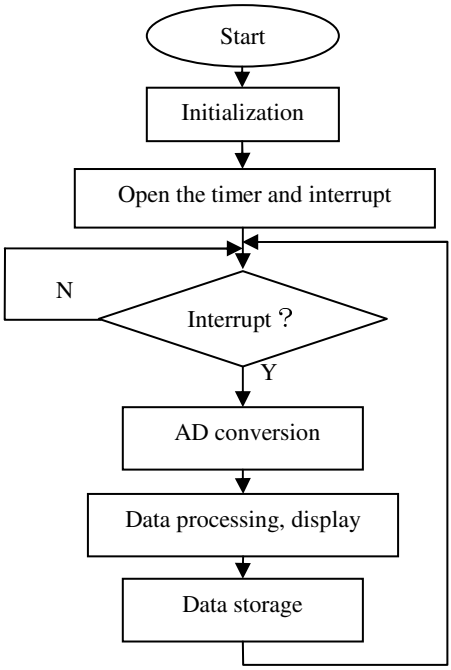


Fig. 8. Overall flow chart of the system software

Acknowledgements. This research was financially supported by National Science and Technology Support Program (2011BAD21B01) and NSFC program (61134011).

References

1. Wang, M.: Development of precision agriculture and innovation of engineering technologies. Transactions of the Chinese Society of Agricultural Engineering 15(1), 1–8 (1999)
2. Zhao, C.: Research and Practice of Precision Agriculture. Science Press, Beijing (2009)
3. Li, M.: Precision agriculture characteristics, current status and development. Agricultural Machinery Market 3, 38–41 (2008)
4. Kitchen, N.R., Sudduth, K.A., Drummond, S.T.: Soil electrical conductivity as a crop productivity measure for claypan soils. Prod. Agric. 12, 607–617 (1999)
5. Sudduth, K.A., Drummond, S.T., Kitchen, N.R.: Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. Computers and Electronics in Agriculture 31, 239–264 (2001)
6. Sun, Y., Wang, M.: A mathematical model and its experimental study for a kind of measurement method of soil electric conductivity. Transactions of the Chinese Society of Agricultural Engineering 17(2), 20–23 (2001)
7. Zhang, M., Li, M., Zhao, Y.: The research and development of vehicle-mounted soil EC monitoring system. Proceeding of CSAE, 10–22 (2011)
8. Zhang, J., Li, M., Kong, D., Zou, Q.: Artificial Intelligence in Real-Time Evaluating Electrical Conductivity of Greenhouse Substrate. In: Li, D., Wang, B. (eds.) Artificial Intelligence Applications and Innovations. IFIP, vol. 187, pp. 609–615. Springer, Boston (2005)

Retracted: Water Temperature Forecasting in Sea Cucumber Aquaculture Ponds by RBF Neural Network Model

Shuangyin Liu[1,2,3], Longqin Xu[1], Ji Chen[2,3], Daoliang Li[2,3,*], Haijiang Tai[2,3], and Lihua Zeng[2,3,4]

[1] College of Information, Guangdong Ocean University, Zhanjiang Guangdong 524025, China
[2] China-EU Center for ICT in Agriculture, China Agricultural University, Beijing 100083, China
[3] Beijing Engineering Research Center for Agricultural Internet of Things, China Agricultural University, Beijing 100083, China
dliangl@cau.edu.cn
[4] College of Mechanical and Electrical Engineering, Agricultural University of Hebei, Banding 071001, China
hdlsyxlq@126.com

Abstract. Water temperature is considered to be the most important parameter which can largely determine the aquaculture production of sea cucumbers, so it is extremely important to monitor and forecast the water temperature at different water depths. As the change of water temperature is a complex process which can not be exactly described with a certain formula, the artificial neural network characterized by non-linearity, adaptivity, generalization, and model independence is a proper choice. This paper presents a RBF neural network model based on nearest neighbor clustering algorithm and puts forward four improved methods, then integrates them into an optimization model and verifies it on matlab platform. Finally, a comparison between the optimized RBF model and the original RBF model is made to confirm the excellent forecasting performance of the optimized RBF neural network model. This paper provides a relatively impeccable learning algorithm to complete the choice of radial basis clustering center in the process of RBF network design, and obtains a high forecasting precision so that the demand of water temperature forecasting in sea cucumber aquaculture ponds can be satisfied.

Keywords: RBF neural network, nearest neighbor clustering algorithm, sea cucumber, water temperature.

Introductions

Sea cucumber is a kind of traditional Chinese seafood with high edible and medicinal value. In the process of rearing sea cucumber, water temperature is the most important parameter of the water quality. On the one hand, water temperature influences the growth rate and development of sea cucumbers as well as their distribution within the

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 425–436, 2013.

pond environment. Sea cucumber has a specific range of water temperature that it can tolerate and high water temperatures can adversely affect it by limiting its habitat or can even result in sea cucumber mortality[1],[2]. On the other hand, water temperature can directly or indirectly influence the other water quality factors, so that it can largely determine the productive capacity of aquatic ecosystems. For example, when water temperature rises, dissolved oxygen will reduce, salinity will rise, PH value will reduce and ammonia nitrogen content will rise, the balance among the water quality factors will be seriously disrupted when the water temperature is excessively high, which can result in the death and rot of sea cucumbers. Therefore, it is extremely important to monitor and forecast the water temperature at different depths in order to control the water temperature stratification during the process of sea cucumber breeding[3].

So far, the major models of forecasting the water temperature generally include stochastic models, regression models, deterministic model and Empirical models. Because water temperature is influenced by water depth, weather, aquatic activities and many other factors, the change process of water temperature can not be exactly described with a certain formula[4]. In a specific application, each type of model has advantages and drawbacks. Stochastic and regression models use statistical techniques where water temperature is related to relevant input parameters (e.g., Solar radiation, air temperature, water depth)[5],[6]. Benyahya et al. provided a review of statistical models[7]. One limitation of this modeling approach is that it explains very little of the underling physical processes. On the contrary, deterministic models are used to predict river water temperature using a mathematical representation of the underlying physics of heat exchange between the river and the surrounding environment[8]. However, a significant drawback of these deterministic models is the amount of data required (information on hydrology and meteorology) and these data is difficult to acquisition.

Empirical models such as artificial neural networks (ANNs) have been used as a viable alternative approach to physical models[9]. It has specific features such as non-linearity, self-adaptivity (i.e., learning from inputs parameters), self-generalization, associated memory, and model independence (no a priori model needed)[10]. ANNs can learn from patterns and capture hidden functional relationships in a given data even if the functional relationships are not known or difficult to identify [10]. Using the training methods, an ANN can be trained to identify the underlying correlation between the inputs and outputs[11].

Radial basis function neural networks (RBF-NNs), in particular, have been extensively studied by researchers in nonlinear identification and water quality forecasting areas such as Dissolved Oxygen Content forecasting [12],[13] and water temperature forecasting also in financial applications such as option pricing and exchange rates forecasting [13]. In this particular type of NN, a RBF takes the role of the activation functions of the network. It has been proved that a RBF-NN can approximate arbitrarily well any multivariate continuous function on a compact domain if a sufficient number of radial basis function units are given. Guo et al. built

two artificial neural network (ANN) models, a feed-forward back-propagation (BP) model and a radial basis function (RBF) model, to simulate the water quality of the Yangtze and Jialing Rivers in reaches crossing the city of Chongqing, P. R. China, the experimental results showed that RBF neural network has characteristics such as small network size, fast learning speed, high precision and so on, which is particularly well suited for problems in which datasets contain complicated nonlinear relations among different inputs, and thus more suitable for exactly describing the forecasting process of water temperature [10],[14],[15],[16],[17].

Although it has excellent features, RBF neural network is limited in academic research and industrial applications, for the difficult points of RBF neural network design include the determination of hidden nodes number and radial basis clustering center[18],[19] .At present, the biggest deficiency of the RBF network is that there is no relatively impeccable learning algorithm to complete the choice of radial basis clustering center, which leads to a deviation on the forecasting results. To overcome the defect, we put forward an improved nearest neighbor clustering learning algorithm which four optimization schemes are joined in the existing algorithm, and a well test result is obtained.

In conclusion, the neural network model based on RBF algorithm is an advisable choice in the field of water temperature dealing with various complex physical processes. This paper first develops an adaptive RBF neural network model, then quests the optimization method and compares the results obtained by the optimized RBF model and the original RBF model, finally provides guidelines of the use of optimized method for future water temperature forecasting purposes. Wherein, the test results show that the optimized RBF algorithm can obtain the smallest forecast error with a shorter training time.

1 Materials and Methods

1.1 Study Area and Data Source

The data used in this study were produced by a Digital Wireless Monitoring System for Aquaculture Water Quality. The system has been installed at China Agricultural University-Dongying Aquaculture Digital System Research Center in Shandong province, where is near the Yellow Rive estuary. In the study, three sea cucumber ponds and a settling pond are selected as monitor objects, wherein the area of each experimental pond is about 1300m^2, and the water level is about 1-2m.

Because the water temperature is mainly affected by the weather factors, we develop the system to monitor water temperature, dissolved oxygen, salinity, PH value and water level of the sea cucumber ponds as well as air temperature, solar radiation, wind speed and rainfall capacity around the sea cucumber ponds. The structure diagram of the Digital Wireless Monitoring System is showed as Fig. 1.

The system comprises four parts: data collection nodes, routing nodes, on-site monitoring center and remote monitoring center. Each data collection node assembles

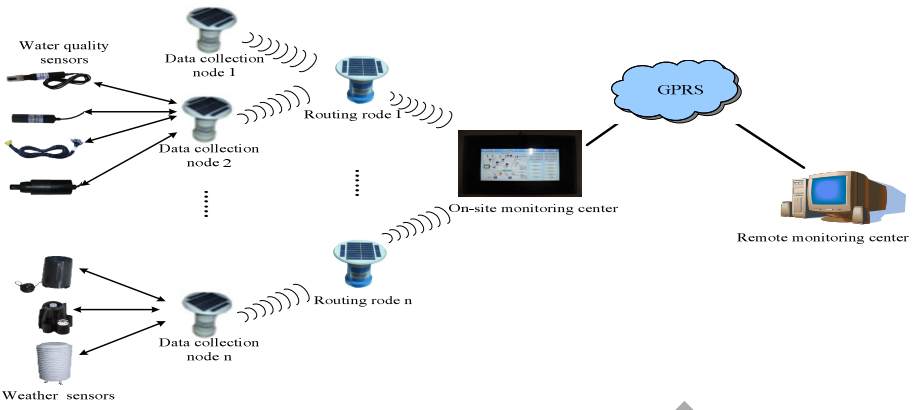


Fig. 1. Structure diagram of Digital Wireless Monitoring System

a number of sensors and a RF modem, and fixed in sea cucumber aquaculture ponds by buoys. On-site monitoring center includes on-site TPC, WSN unit and GPRS unit, remote monitoring center is a computer with fixed IP address.

The working process of the system is as follows, data collection nodes send data to the on-site monitoring center through routing nodes. On-site monitoring center collects all the data and then communicates with remote monitoring center through GPRS. From both on-site monitoring center and remote monitoring center, users can get the water quality data and weather data collected by the data collection nodes.

The data used in the paper include water temperature data at 4 different depths and real-time meteorological data in Dongying from August 7 to August 13, 2009. Data collection time interval is 10 minutes, and we use data of first 6 days for neural network training and the data of latter 1 day for test.

1.2 RBF Neural Network

ANN uses a multilayered approach that approximates complex mathematical functions to process data. An ANN is arranged into discrete layers each layer consisting of at least one neuron. Each node of a layer is connected to nodes of preceding and/or succeeding layers but not to nodes of the same layer with a connection weight. Thus, as the number of layers and nodes in each layer increases, the process becomes more complex demanding more computational effort. In general hydrologic and environmental problems are complex and require a complex ANN structure for prediction purposes. The number of layers and nodes in each layer is problem specific and needs to be optimized [20],[21].

Nearest neighbor clustering algorithm is an online adaptive dynamic clustering algorithm without confirming the number of hidden units in advance, which can obtain a optimal RBF network with a short learning time and a small amout of computation.

1.3 Network Optimization

By studying and analyzing the nearest neighbor clustering learning algorithm, we have found 4 deficiencies:

(1) It is not reasonable in some cases that only input information is used to determine whether a sample belongs to a clustering .

(2) It is not appropriate with a fixed clustering radius when there is a big difference in sample distribution density.

(3) In most cases, taking $C_j = \frac{1}{s}\sum_{j=1}^{k} x_j^k$ as a clustering is more reasonable, in which s is the sample number of subset k.

(4) It doesn't take the learning errors as performance index to do the iterate in learning process, which will be limited on the occasions with high precision requirement.

Aiming at the defects of original algorithm, we have made corresponding improvements, the specific steps are as follows:

Step (1) Compute the distance d_{ij} (i , j =1, 2, 3, …, N, N is the sample total) between each sample and their average $\overline{d}$:

$$d_{ij} = \sqrt{\left\|x^i - x^j\right\|^2 + \left\|y^i - y^j\right\|^2} \tag{1}$$

$$\overline{d} = \frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} d_{ij} \tag{2}$$

Find out the nearest t samples to sample i, then set the distance as d_j respectively and work out their average distances.

$$\overline{d}_i = \frac{1}{t}\sum_{p=1}^{t} d_i^p \tag{3}$$

Step (2) Beginning from the first sample data set (x^1, y^1), create a clustering center on x^1 expressed as $\overline{x}^1$, then set A(1)=y^1, B(1)=1 and select a initial clustering radius r.

Step (3) Suppose the clustering center number of (x^k, y^k) is m and the center is $\overline{x}^1, \overline{x}^2, ..., \overline{x}^m$ respectively, which means that there are m hidden units in the network as above-mentioned. Then find out the distance to these clustering centers respectively which expressed as $|x^k - \overline{x}^i|$,i=1,2,…,m.

Set $|x^k - \overline{x}^j|$ as the smallest distance, the clustering j is the nearest neighbor clustering of sample x^k. Then make

$$r_k^* = \frac{r \bullet \bar{d}_k}{\bar{d}} \tag{4}$$

If $$\sqrt{\left\|x^k - \bar{x}^j\right\|^2 + \left\|y^k - \bar{y}^j\right\|^2} \le r_k^* \tag{5}$$

take the sample (x^k, y^k) into clustering j, and

$$A(j) = A(j) + y^k, B(j) = B(j) + 1, \bar{x}_{tmp}^j = \bar{x}_{tmp}^j + x^k \tag{6}$$

$\bar{x}_{tmp}^j$ is clustering j and the sum of inputs belong to it. Otherwise, take the sample (x^k, y^k) as a new clustering center , and set

$$\bar{x}^{m+1} = x^k, m = m + 1, A(m) = y^k, B(m) = 1 \tag{7}$$

Step (4) After determining the clustering, set $\bar{x}^i = \bar{x}_{tmp}^i / B(i)$ and put learning samples into network, then get the fitting error sum of squares E,

$$E = \sum_{k=1}^{N} e_k^2, e_k = f^k - y^k \tag{8}$$

f^k is the output when input is x^k, at this time we can adjust the network parameters according to general gradient algorithm,

$$\sigma = \sigma - \eta \frac{\partial E}{\partial \sigma} \tag{9}$$

η is learning rate.

Step (5) The network output after learning is

$$f(x^k) = \frac{\sum_{j=1}^{m} A(j) \prod_{i=1}^{n} \exp\left(-((x_i - \bar{x}_i^j)/\sigma)^2\right)}{\sum_{j=1}^{m} B(j) \prod_{i=1}^{n} \exp\left(-((x_i - \bar{x}_i^j)/\sigma)^2\right)} \tag{10}$$

2 Implementation, Results and Discussion

This paper constructs a RBF neural network using five factors: air temperature, wind speed, solar radiation, water level and the previous water temperature value as network inputs to forecast the water temperature in sea cucumber aquaculture ponds. The Gaussian function and the pure linear purelin type transformation function are separately used as hidden layer neuron and output layer neuron. And the result of this model is the water temperature value 0.5 hour ahead, the forecasting value should be in high coordination with the observed value.

In order to evaluate the performance of proposed algorithm effectively, we developed the original algorithm and the optimized algorithm on matlab platform respectively, and the forecasting results comparison diagrams are showed as follows:

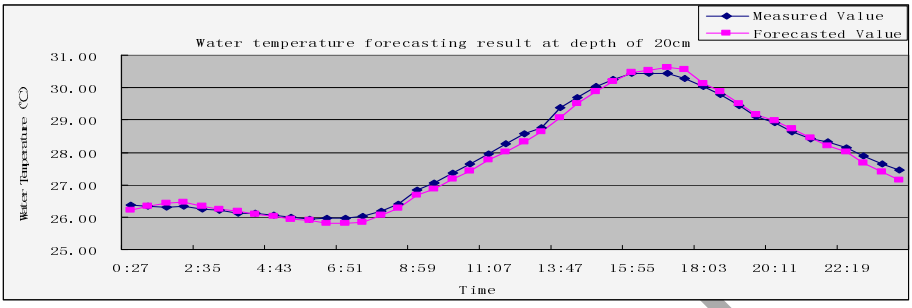


Fig. 2. Water temperature forecasting result diagram at depth of 20 cm

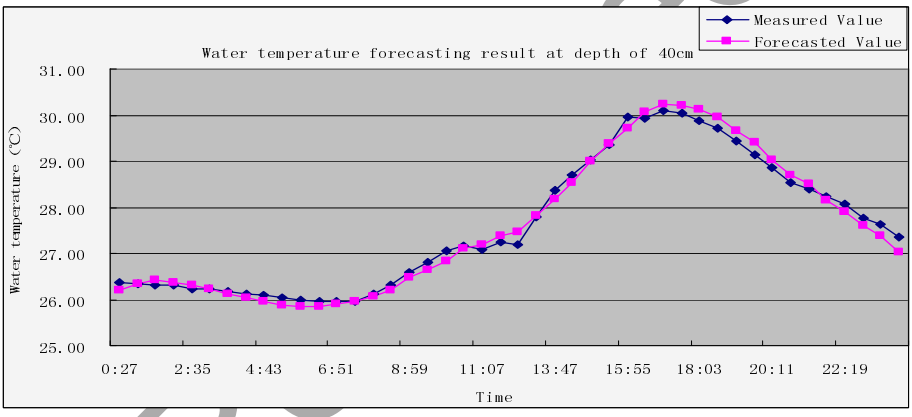


Fig. 3. Water temperature forecasting result diagram at depth of 40 cm

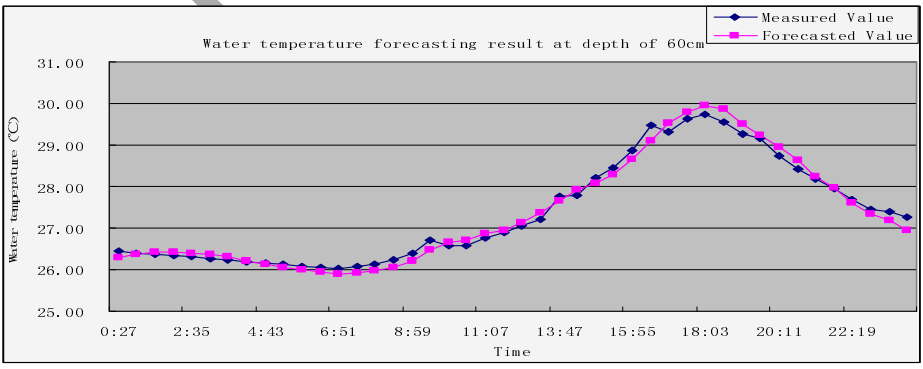


Fig. 4. Water temperature forecasting result diagram at depth of 60 cm

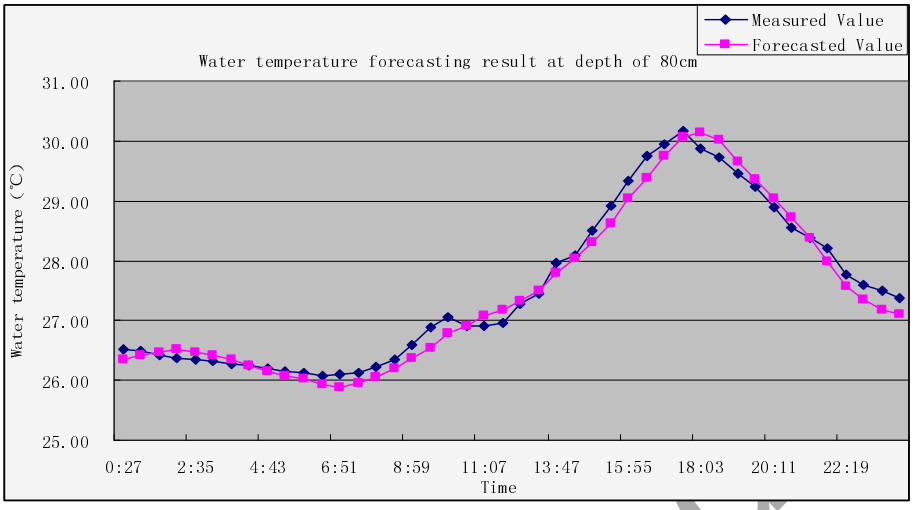


Fig. 5. Water temperature forecasting result diagram at depth of 80 cm

The forecasting performances of the optimized method is measured by using four different statistical efficiency criteria to evaluate the relative strength and weakness of the various models developed. These are mean relative error (MRE), minimal relative error, maximal relative error and root of mean square error (RMSE). Each term is estimated from the predicted and observed water temperature (targets). All of these efficiency terms are unbiased as they use error statistics relative to the observed values.

Table 1. Forecasting performance comparison of the optimized RBF network at different depths

Depth of water	*MRE*	*Minimal relative error*	*Maximal relative error*	*RMSE*
20cm	0.457%	0.029%	1.131%	0.154
40cm	0.464%	0.003%	1.168%	0.157
60cm	0.477%	0.061%	1.257%	0.160
80cm	0.580%	0.016%	1.241%	0.191

The four forecasting results comparison diagrams and the forecasting performance comparison table show that the RBF neural network can effectively forecast the change of water temperature 0.5 hour ahead, the mean square error of the forecasting

results can be controlled in less than 0.2, and the maximal relative error can be controlled within 2%. Wherein the forecasting precision of RBF neural network has a slight decline with the increase of depth, this is because the input factors of neural network we used (weather factors mainly) have a gradually diminished influence on water temperature with the increase of depth. But fortunately, the tiny change does not have a big impact on forecasting result.

In addition, in order to verify the effectiveness of the optimization method, we compare the forecasting results between the original model and the optimized model at the depth of 20 cm, at the same time, the results obtained by each optimization method are also compared.

Table 2. Comparative analysis of forecasting performance efficiency of each optimization method at depth of 20 cm

Method	*MRE*	*Training time*
original RBF network	0.638%	7.6s
optimized RBF network by adding output information	0.597%	7.8s
optimized RBF network by adding adaptive adjustments	0.481%	6.2s
optimized RBF network by replacing the final clustering	0.574%	7.4s
optimized RBF network by gradient optimization	0.498%	7.1s
final optimized RBF network	0.457%	6.9s

Table 2 shows that the optimized network has an obviously better result than original network, these results are displayed not only in the forecasting precision, but also in training time. The final optimized RBF network makes the MRE reach a minimum with the shortening of the training time. And the optimization performances of each method are as follows:

(1) the optimized algorithm adds influence of output information in the process of judging whether a sample can form a new clustering (formula 5) , which makes the determination of clustering more reasonable, and the result shows that it can improve the forecasting precision a little, but the training time is prolonged at the same time;

(2) adaptive adjustments of new clustering radius based on sample density (formula 4) can not only improve forecasting precision, but also shorten the training time greatly;

(3) we replace the final clustering input by clustering input and the input centers of samples belonging to it (formula 6 and step 4) , so the clustering can reflect and replace the samples belong to it more accurately, and the forecasting precision and the training time can be simultaneously improved a little;

(4) we only adjust one-dimensional parameters in process of gradient optimization (formula 9), as a result, there is little effect on learning speed with a big forecasting increase.

(5) at last, this paper develops a forecasting network with optimal performance by integrating all the optimization modes (formula 10), and the final optimized RBF network has the minimum forecasting precision and a shorter training time.

3 Conclusion

Based on nearest neighbor clustering learning algorithm, a RBF neural network model of forecasting the water temperature in sea cucumber aquaculture ponds has been established. In order to maximise the forecasting precision and shorten learning time of network, four improved methods aiming at the defects of nearest neighbor clustering learning algorithm have been put forward. Then we integrated these methods reasonably after comparing the effects of each improved method and achieved an optimal optimization performance.

The results show that RBF neural network is a proper method for modeling a heat transfer problem due to the lack of information about internal process and boundary conditions. The optimized network has a better forecasting performance than the original network, wherein the mean square error of the forecasting results can be controlled in less than 0.2, and the maximal relative error can be controlled within 2%, thereby satisfying the demand of water temperature forecasting in sea cucumber aquaculture farms.

However, the forecasting precision of RBF neural network has a slight decline with the increase of depths, which owing to the input factors of neural network we used (weather factors mainly) having a gradually diminished influence on water temperature with the increase of depths, Fortunately, the tiny change does not have a big impact on forecasting results.

This paper proposed four improved methods, then contrasted and integrated them by experiments respectively, finally obtained an optimization algorithm with 0.181% decrease in MRE and 0.7s decrease in training time compared with the original algorithm. But restricted by the monitoring data, we can only estimate the water temperature stratification conditions by forecasted value of water temperature at the four different depths, the water temperature forecasting performance on the bottom water can not be verified, which needs to be improved in the future. Moreover, because the nearest neighbor clustering can not reflect the average value of samples, it is a feasible way to create a new kind of clustering algorithm aiming at this defect.

Acknowledgements. This research is financially supported by the National Key Technology R&D Program in the 12th Five year Plan of china (2011BAD21B01), Guangdong Science and Technology Program Project (2012A020200008) and the Zhanjiang Science and Technology Program Project (2010C3113011).

References

1. Ji, T., Dong, Y., Dong, S.: Growth and physiological responses in the sea cucumber, Apostichopus japonicus. Aquaculture 283, 180–187 (2008)
2. Xu, L., Liu, S.: Water quality prediction model based on APSO-WLSSVR. Journal of Shangdong University (Engineering Science) 42(5), 80–86 (2012)
3. Wang, F., Yang, H., Gao, F., et al.: Effects of acute temperature or salinity stress on the immune response in sea cucumber Apostichopus japonicus. Comparative Biochemistry and Physiology Part A 151, 491–498 (2008)
4. Hernández, J.M., León-Santana, M., et al.: The role of the water temperature in the optimal management of marine aquaculture. European Journal of Operational Research 181, 872–886 (2007)
5. Ahmadi-Nedushan, B., St-Hilaire, A., Ouarda, T.B.M.J., et al.: Predicting river water temperatures using stochastic models: case study of the Moisie River (Québec, Canada). Hydrological Processes 21(1), 21–34 (2007)
6. Chenard, J.F., Caissie, D.: Stream temperature modeling using artificial neural networks: application on Catamaran Brook, New-Brunsick. Canad Hydrological Processe 22(17), 3361–3372 (2008)
7. Benyahya, L., St-Hilaire, A., Ouarda, T.B.M.J., et al.: Modeling of water temperatures based on stochastic approaches: case study of Deschutes River (Oregon, USA). Journal of Environmental Engineering and Science 6, 437–448 (2007)
8. Caissie, D., Satish, M.G., El-Jabi, N.: Predicting water temperatures using a deterministic model: application on Miramichi River catchments (New Brunswick, Canada). Journal of Hydrology 336(3-4), 303–315 (2007)
9. Falca, A.O., Langlois, T., Wichert, A.: Flexible kernels for RBF networks. Neurocomputing 69, 2356–2359 (2006)
10. Sahoo, G.B., Schladow, S.G., Reuter, J.E.: Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models. Journal of Hydrology 378, 325–342 (2009)
11. Palit, A.K., Popovic, D.: Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications, 1st edn. Springer (2005)
12. Zemouri, R., Racoceanu, D., Zerhouni, N.: Recurrent radial basis function network for time-series prediction. Engineering Applications of Artificial Intelligence 16, 453–463 (2003)
13. Han, H., Chen, Q., Qiao, J.: An efficient self-organizing RBF neural network for water quality prediction. Neural Networks 24, 717–725 (2011)
14. Guo, J., Li, Z.: Artificial neural network modeling of water quality of the Yangtze River system: a case study in reaches crossing the city of Chongqing. Chongqing Univ: Eng. Ed. 8(1), 1–9 (2009)
15. Romero, C.E., Shan, J.: Development of an artificial neural network-based software for prediction of power plant canal water discharge temperature. Expert Systems with Applications 29, 831–838 (2005)
16. Malinowski, P., Sułowicz, M., Bujak, J.: Neural model for forecasting temperature in a distribution network of cooling water supplied to systems producing petroleum products. International Journal of Refrigeration 34, 968–979 (2011)
17. Huang, B., Langpap, C., Adams, R.M.: Using Instream Water Temperature Forecasts for Fisheries Management: an Application in the Pacific Northwest. Journal of the American Water Resources Association 47, 861–876 (2011)

18. Wu, A., Hsieh, W.W., Tang, B.: Neural network forecasts of the tropical Pacific sea surface temperatures. Neural Networks 19, 145–154 (2006)
19. Awad, M., Pomares, H., Rojas, I.: Prediction of Time Series Using RBF Neural Networks: A New Approach of Clustering. International ARAB Journal of Information Technology 6, 138–143 (2009)
20. Zhang, J., Hu, S.: Chaotic time series Prediction based on RBF neural networks with a new clustering algorithm. Acta Physica Sinica 56, 713–719 (2007)
21. Lee, C.M., Ko, C.-N.: Time series prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm. Neurocomputing 73, 449–460 (2009)
22. Ignacio, J., Mulero-Martnez.: Analysis of the errors in the modelling of manipulators with Gaussian RBF neural networks. Neurocomputing 72, 1969–1978 (2009)
23. Hsu, C., Chiu, C., Tsai, J.: Indirect adaptive self-organizing RBF neural controller design with a dynamical training approach. Expert Systems with Applications 39, 564–573 (2012)

Discussion on Calculation Method of Social Stability Price of Farmland Requisition Price: Taking Bazhou City as an Example*

Yapeng Zhou[1], Lin Liu[2,3], Yang Yang[1], Li Zhang[1], Hao Xu[1], Yigong Zhang[1,**], and Zhiwei Li [1]

[1] College of Natural Resources & Environment Sciences, Agricultural University of Hebei, Baoding 071001, China
[2] School of Land Science and Technology, China University of Geosciences, Beijing 100083, China
[3] Shijiazhuang Engineering Technology School, Shijiazhuang 050061, China
{zhouyp,zhangyg63}@hebau.edu.cn

Abstract. Land is a kind of basic resources and of great importance to the survival and development of human beings. Due to the conflict between the limited amount of land and the unlimited development of human beings, land surely plays a significant role in national security and social development. The academic circle has introduced the social stability price to the pricing system of farmland, which reflects the social stability value of farmland. This paper analyses the composition of social stability function of farmland and takes Bazhou city as an example to discuss two calculation methods of social stability price of farmland, a direct method and an indirect method.

Keywords: Farmland, Social Stability Price, Calculation.

The farmland price we refer to at present usually only involves its productivity price, which is the price of farmland as productive power or a means of production, the opposite side of value in use. And the social value of farmland, which lies in the fact that farmland provides farmers with survival guarantee and social welfare, and farmland is also of the function of national security and social stability as it provides food for the society, is ignored. At present, knowledge of its social stability function and social stability price is just limited to the aspect of food security and the elaboration on its definition and functions is not clear and complete[1].

1 Definition of the Social Stability Function and Price of Farmland

The social stability function of farmland means the function that farmland performs so that the state and the social development can enjoy a safe and stable state which is

* The paper is supported by the Hebei Provincial Department of education funding for a project, (project number is 2009451).
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 437–443, 2013.

healthy, high-efficiency and not threatening and accepted by human beings. The risk of this kind of stable state is predictable and accepted by human beings.

The social stability price of farmland means that the social stability function of farmland is monetized or realized in the market, which becomes the social stability price of farmland.

Price is the embodiment of value and comes from the realization of function. In order to figure out the social stability price of farmland, the social stability function of farmland should be analyzed first, so that its social stability price can be obtained [2].

2 The Composition, Characteristics and Influencing Factors of Social Stability Function of Farmland

2.1 The Composition of Social Stability Function of Farmland

Generally speaking, the function of land lies in the four aspects, namely production function, environment function, bearing function and space function [3]. As far as farmland is concerned, its social stability function also lies in the four aspects. First, due to its productivity, farmland can yield various kinds of crops which are necessary for the survival of human being, meet the requirement of human beings for food, and ensure the food security. Second, farmland is a constituent of the whole ecological environment and farmland is indispensable to ensure the structural integrity of the ecological environment and realization of its functions. Third, human beings carry out their activities on land and farmland as a constituent of land also serves the bearing function, which ensures the economic development and social stability and progress. Finally, as the survival and development of humans need some space and before they find a usable space in addition to the earth, land is the only choice of human beings (Figure 1).

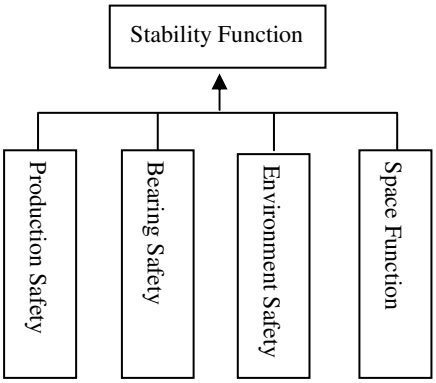


Fig. 1. Constitutional Diagram of Social Stability Function of Farmland

2.2 The Characteristics of Social Stability Function of Farmland

Due to the characteristics of farmland and influence of the external environment on farmland, the social stability function of farmland is characterized by the following aspects: eternity, spatial differentiation, time variability, limited function, complicated influencing factors and various function ways.

2.3 The Influencing Factors of Social Stability Function of Farmland

Farmland is an open complicated system, with many influencing factors of complicated functions. Our analysis reveals that the influencing factors that affect social stability function of farmland cover the following aspects (Table 1):

Table 1. Table of Influencing Factors of Social Stability Function of Farmland

	Aspects Influenced	Influencing Factors
Social Stability Function of Farmland	Number	taken up by construction, industrial restructuring of agriculture, taken up by people and lying waste, destroyed by natural disasters, etc.
	Quality	water resources, lighting, temperature, nutrients, environment pollution, etc.
	Social Economy	safety quantity demanded within planning period, investment in agriculture, local living standards, construction of infrastructure, relevant policies, etc.
	Science and Technology	irrigation techniques and facilities, seed resources, cultivation techniques, etc.

3 Calculation of Price of Social Stability Function of Farmland

The influence of various influencing factors on the function of farmland is reflected directly by price. There are two ways of price realization; one is direct calculation, which is rather difficult; the other one is indirect method, which evaluates the price in accordance with the substitution principle. The concrete ideas are as follows:

3.1 Direct Calculation Method

This calculation method begins with the social stability function of farmland, evaluates the importance of social stability function of farmland to the whole society, works out the weight of the function and then finds out the influencing factors that influence the security function of farmland. Taking the economic conditions and scientific levels into account, quantify these influencing factors and the social stability price of farmland is obtained. The calculation equation of social stability price in a unit area is as follows:

$$S=S1+S2+S3+S4 \quad (1)$$

In this equation, S is the social stability price of farmland in a unit area.

S1 is the production safety price in a unit area.
S2 is the bearing safety price in a unit area.
S3 is the environment safety price in a unit area.
S4 is the space safety price in a unit area.

The price of each function is the function of the influencing factors. The influencing factors are already listed in the table above and they can be used in calculation. The equation for the function price in a unit area of each part is

$$SI=f(X1,X2,X3,\ldots XI) \tag{2}$$

In this equation, SI is the function price in a unit area of the four parts in the social stability function of farmland

XI is the various factors of stability function price in each part

The present conditions show that the influencing factors of the last three kinds of stability function price are of complicated function and it is hard to calculate the price. Therefore, production safety price is used to replace the stability price of the whole society temporarily.

3.2 Indirect Calculation Method

The calculation of the social stability price can be based on the substitution principle and use land reclamation fees as the quantized value of the social stability function, because the land reclamation fee is the compensation of value for land taken up. The charging of land reclamation fee is a economic means to ensure the dynamic balance of the total cultivated land in China and is the basis for food safety in the society. The determination of standards of land reclamation fees should first consider the land capital input of land with average quality levels, including tangible farmland water conservancy facilities and intangible soil economic fertility. For the tangible farmland water conservancy facilities, the quantity of value can be determined based on the replacement cost; for the intangible soil economic fertility, the formation cycle of soil fertility as well as the capital input each year within the cycle should be considered. According to the sum of capital input each year within the formation cycle of soil economic fertility, the quantity of value is determined by gains and losses of investment within the formation cycle of economic fertility. As far as soil fertility is concerned, due to difference in starting reclamation, hysteresis quality of reclamation benefit and gradualness of improving soil fertility, land reclamation fees are different. At present, for the wasteland which is easy to reclaim, it is predicted that in the first four years, there is profit, but not much. Generally, from the fifth year on, the profit begins to increase. Therefore, taking the input factors to form enough soil fertility and the losses of profit in the first four years, suppose the average input level of land reclamation is T yuan/mu, the yield level is P yuan/mu, the capitalization rate is r, the average interest rate of recent years, so the present value of input for reclaiming wasteland for five years is

$$V1=T+T/(1+r)+T/(1+r)2+T/(1+r)3+T/(1+r)4 \tag{3}$$

The present value of loss in revenue in four years is

$$V2=P/(1+r)+P/(1+r)2+P/(1+r)3+P/(1+r)4 \tag{4}$$

The social stability price of farmland V=V1+V2. With the land development and management gathering momentum and increase of demand for land from the society,

land reserve resources are decreasing day by day, land reclamation becomes more and more difficult and the land reclamation fees will increase day by day.

4 An Attempt to Calculate the Social Price of Farmland in Bazhou City

4.1 A Brief Introduction to Bazhou City

Bazhou city is located in the south-central part of Hebei province and the center of the triangle area formed by Beijing, Tianjin and Baoding. With the north latitude of 38°58′57″--39°13′10″ and east longitude of 116°15′13″--116°55′09″, Bazhou city is situated at the semiarid continental monsoon climate zone in temperate zone. Adjacent to Tianjin and Wuqing on the east, bordering on Xiongxian on the west, adjacent to Wenan on the south and bordering on Guan, Yongqing and Anci district of Langfang City on the north, it covers an area of 80027.25 square kilometers, among which 57348.41hectares of farmland, accounting for 71.7% of the total area.

4.2 An Attempt to Calculate the Social Stability Price

4.2.1 Direct Method

Here the production safety price of the first part is calculated. The concrete process is as follows: taking the practical situation of Hebei province into consideration, evaluate the function of the production capacity of farmland of Bazhou city in Hebei province and get the weight. Then consider the food safety in Hebei province and arable land reduction of Bazhou city within a certain planning period, get the food safety value that farmland in Bazhou city should have in the planning period, quantify it and obtain the production safety price of farmland in Bazhou city.

According to the statistics of 2003, the grain out of Hebei province was 23.878 million ton and that of Bazhou city was 177.198 thousand ton.

Calculate the production safety weight of Bazhou city:

PS= grain output of the small area/grain output of the large area
=grain output of Bazhou city/grain output of Hebei province
=177198/2387.8
=0.0074

Based on the population base and the natural growth rate of Hebei Province in 2003, it is predicated that the population of Hebei will reach 70.2318 million in 2010. If 350 kg of grain per capita is determined to be the survival safety standards, 24581130647 kg will be needed. If 420 kg of grain per capita is the nutrition safety standards, 29497356776 kg will be needed. Based on the grain output of Hebei province, the food gap is 703130646.8 kg and 5619356776 kg respectively. Taking PS, the production safety weight of farmland of Bazhou city, into account, the selling price of wheat is converted into the production safety price of farmland and the total amount of production safety function value of farmland in Bazhou city is obtained.

SZ1=quantity demanded for survival safety in Hebei province×production safety weight of Bazhou city×selling price of wheat

=703130646.8×0.0074×0.7
=3642216.75（yuan）

SZ2= quantity demanded for nutrition safety in Hebei province×production safety weight of Bazhou city×selling price of wheat

=5619356776×0.0074×0.7
=29108268.1（yuan）

The area under cultivation in Bazhou city at the beginning of 2002 was 48830 hectares and the area at the end of 2003 was 46148 hectares, reducing by 1341 hectares annually. Based on this reduction amplitude, by the year of 2010, the area under cultivation in Bazhou city will have reduced to 36761 hectares. Let's assume that the production level remains the same as the present level and the influence by scientific progress and social economy are not taken into consideration. That is to say, the various influencing factors except the amount and the safety need amount all remain the same. So the production safety price in a unit area changes into the function of the amount and the safety need amount. Then the amount of the production safety value of farmland in a unit area is obtained by calculation with the equation above. Namely,

Survival safety price in a unit area:
S=*f*(total amount of safety need, area under cultivation at the end of the period)
=SZ1/ area under cultivation at the end of the period=3642216.75/36761
=99.08（yuan/hectare）
nutrition safety price in a unit area:
S=*f*(total amount of safety need, area under cultivation at the end of the period)
=SZ2/ area under cultivation at the end of the period=29108268.1/36761
=791.82（yuan/hectare）

4.2.2 Indirect Method

The equation above is adopted.

V1＝T＋T/(1＋r)＋T/(1＋r)2＋T/(1＋r)3＋T/(1＋r)4
V2＝P/(1＋r)＋P/(1＋r)2＋P/(1＋r)3＋P/(1＋r)4
social stability price： V=V1＋V2

According to a survey, the average input level each year into agricultural production in Bazhou city is 5250—9900 yuan per hectare and the output level is 15000 yuan per hectare. The annual input into land reclamation is 9900 yuan per hectare and the interest rate is 3%. Therefore, the present value of input into land reclamation is

V1=9900＋9900/(1+3%)＋9900/(1+3%)2＋9900/(1+3%)3＋9900/(1+3%)4
=46699.2（yuan/hectare）
The present value of loss in revenue for 4 years is
V2=15000/(1+3%)＋15000/(1+3%)2＋15000/(1+3%)3＋15000/(1+3%)4
=55756.5（yuan/hectare）
Therefore, V=V1+V2=102455.7（yuan/hectare））

5 Discussion

The result of the direct calculation has already deducted the amount of grain based on the production capacity of farmland and the production safety price of it is calculated, with the reduction of arable land taken into consideration. As the safety function of grain production of farmland is calculated, arable land is considered rather than other kinds of farmland. When the social stability price of farmland is calculated, the price of only one kind of crop－wheat, is calculated and the increase in grain yield due to the scientific progress and the increase in the input into social economy is not considered. In practice, the price of many kinds of crops should be used in the calculation and the influence of social progress and increase in input into social economy on the increase in grain yield should also be considered when the production safety price of farmland is determined. In addition, the calculation only covers the production safety price and does not involve the safety function price of the other three kinds due to differnt calculation method and data source.

Compare the result of the direct calculation and that of the indirect calculation, we can see that there is great difference in fees used for cultivating land and the safety price, as the direct calculation does not take the other three kinds of price into consideration.

References

1. Liu, X.: On Choice of Strategic Objectives and Models of Food Security in China. Journal of Shanxi Finance and Economics University 26(3) (2004)
2. Liu, L.: Land Resource Science, pp. 6–7. China Agricultural University Press, Beijing (2001)
3. Liu, H.: On Composition and Quantification of Farmland Price in China. Land Science of China (2000)

Quantitative Analysis of and Discussion on Social Security Price of Farmland in Land Requisition Price*

Yapeng Zhou[1], Lin Liu[2,3], Yang Yang[1], Ying Chen[1], Hao Xu[1], Wenting Zhao[1,**], and Na Hao[1]

[1] College of Natural Resources & Environment Sciences, Agricultural University of Hebei, Baoding 071001, China
[2] School of Land Science and Technology, China University of Geosciences, Beijing 100083, China
[3] Shijiazhuang Engineering Technology School, Shijiazhuang 050061, China
zhouyp@hebau.edu.cn, zwt1964b@sina.com

Abstract. This paper points out that the irrational calculation of the social security price of farmland leads to the relatively low price of the land requisition price. This paper begins with the analysis of the social security function of farmland to farmers and puts forward the idea that the social security value of farmland should include the following four aspects, namely, the value of the endowment insurance, the medical insurance, the job security and the inheritance of property. Based on the principle that the urban residents and the rural residents are equal and taking the pricing level of the social security system of residents into account, this paper quantifies the value of the four aspects of the social security value of farmland, adds up the four aspects, analyzes them and gets the social security price of farmland.

Keywords: Farmland, Social Security, Quantitative Analysis.

With the unceasing development of the market economy in China, radical changes have occurred to the employment system of enterprises and the traditional compensation and placement system cannot function well. Therefore, how to calculate the land requisition price of farmland in the new era rationally has become a major problem that needs resolving urgently. The social security price of farmland is an important part of its price structure, but the research into its calculation method is still limited to the theoretical research. Therefore, the research into the rational calculation method of social security price of farmland will provide a basis for building a sound land requisition pricing system of farmland [1,2].

1 Definition of Social Security Price of Farmland

The social security system is the social service and measures provided by a state or a community so as to improve and enhance the material standards as well as the living standards of its residents, and it is also a sign of social civilization and social progress.

* The paper is supported by the Hebei Provincial Department of education funding for a project, (project number is 2009451)
** Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 444–450, 2013.

On March 22, 2005, the No. 233 file, Technical Guide to Setting Comprehensive Land Price in Requisition Area (Exposure Draft) issued by the General Office of Ministry of Land and Resources defined the social security price of farmland as "the due compensation accepted by farmers of whom land is requisitioned and does not get the basic living allowances, to receive education and reemployment training" [2,3]. Therefore, the social security price of farmland is monetized functions of the four aspects, namely, the endowment insurance, medical insurance, job security and inheritance of property for farmers.

2 The Price Structure of Social Security Price of Farmland

2.1 The Function of Endowment Insurance

Land provides not only a means of production for farmers in the contemporary age, but also a basis for their children to live by cultivating the land and support their elders.

2.2 The Function of Medical Insurance

Family has become the basic unit for farmers to get medical insurance. For most rural families, land is the most basic and most important means of production and land revenue is the most important or even the sole source of family security. Therefore, land is of important medical security function for farmers.

2.3 The Function of Job Security

Land is the important place for farmers to work on and live on and also basis for their subsistence. By means of working on land, farmers can obtain the basic security for their food and closing. They all amply show that land is of important job security function or basic subsistence guarantee function for farmers.

2.4 The Function of Property Inheritance

In accordance with relevant laws, land is of property inheritance function for farmers, it not only guarantees the basic subsistence of farmers, but also provides their offspring with the basic survival conditions [3,4].

3 The Calculation of Social Security Price of Farmland

The new edition of The Law of Land Administration adopts the "method of multiplicity of production value" to calculate the placement and compensation fee (i.e. the social security price). This kind of calculation is very subjective and lacks a reasonable basis, and thus the result fails to reveal the social security value of land. Therefore, this paper will calculate the endowment insurance price of land, medical insurance price of land,

job security price of land and property inheritance price of land respectively, then add them up, modify them reasonably and finally obtain the social security price of land [4].

At present, more and more farmers will leave their land, and thus the social security value of land will gradually be replaced by the social security system of the state, and farmers will enjoy the same social security as that of the urban residents. Therefore, when the social security price of land is calculated, the choice of the parameters as well as the value can consult the social security standards of the urban residents, i.e. reach uniformity between urban residents and rural residents.

3.1 Calculation of Endowment Insurance Price

The endowment insurance price of the requisitioned land should be calculated based on the amount of money paid for the individual endowment insurance. By reference to the endowment insurance system of urban residents and taking the individual endowment insurance, scale of premium rate of Taipingshengshi Changshou endowment insurance A (see Table 1) of China Pacific Insurance Co., Ltd. as an example, the per capita single premium of endowment insurance can be calculated based on the following equation.

$$Ye = (Yma \times Bm + Ywa \times Bw) \times Mi / Mo$$

In this equation, Ye—the per capita single premium of endowment insurance

a—average age

Yma—base number of the single premium of endowment insurance of male citizens at the age of a

Ywa—base number of the single premium of endowment insurance of female citizens at the age of a

Bm—the proportion of male population to the total population

Bw— the proportion of female population to the total population

Mi—basic allowances of farmers (insurance benefit standard per month)

Mo—base number of the premium per month

Table 1. Scale of Premium Rate of Taipingshengshi Changshou Endowment Insurance A of China Pacific Insurance Co., Ltd.(Male citizens get insurance benefit when they reach 60 years old, while female citizens reach 55 and both get 100 yuan each month) unit：yuan

Purchasing Insurance at the Age of	Single Premium for Males	Single Premium for Females	Purchasing Insurance at the Age of	Single Premium for Males	Single Premium for Females
0	3642	5278	30	7956	11283
1	3719	5412	31	8170	11573
2	3816	5550	32	8390	11871
3	3916	5692	33	8616	12177
…	…	…	…	…	…
27	7347	10454	57	16173	
28	7545	10723	58	16575	
29	7747	11000	59	16981	

3.2 Calculation of Medical Insurance Price

The medical insurance price of the requisitioned land should be calculated based on the amount of money paid for the individual medical insurance. Therefore, by reference to the individual medical insurance, scale of premium rate of Taipingshengshi Changjian Medical Insurance (see Table 2) of China Pacific Insurance Co., Ltd., the per capita single premium of medical insurance can be calculated based on the following equation.

$$Ym=Pma\times Bm+Pwa\times Bw$$

In this equation, Ym—the per capita single premium of medical insurance

a—average age

Pma—base number of the single premium of medical insurance of male citizens at the age of a

Pwa—base number of the single premium of medical insurance of female citizens at the age of a

Bm—the proportion of male population to the total population

Bw—the proportion of female population to the total population

Table 2. Scale of Premium Rate of Taipingshengshi Changjian Medical Insurance A of China Pacific Insurance Co., Ltd.（insured till 70 years old） unit：yuan

Purchasing Insurance at the Age of	Single Premium for Males	Single Premium for Females	Purchasing Insurance at the Age of	Single Premium for Males	Single Premium for Females
16	865	743	39	1329	1128
17	884	759	40	1347	1140
18	903	775	41	1364	1151
…	…	…	…	…	…
35	1250	1071	58	1335	1008
36	1270	1086	59	1295	971
37	1290	1101	60	1246	928
38	1310	1115			

3.3 Calculation of Job Security Price

Land is the place of employment for most farmers and is of job security function for them. Therefore, the loss of land for farmers is equal to the loss of jobs for urban residents, so the job security price of land is the price of unemployment insurance. The unemployment insurance is a government action and now is not taken over by insurance companies by means of charging a single premium. When the unemployment insurance is calculated, the years before the retirement of farmers should be worked out based on the average age of farmers and during these years, farmers should be treated as unemployed, and thus the government should pay them one-off unemployment compensation. As the number of years to receive unemployment compensation is determined by the number of years residents paying unemployment insurance before

unemployment, the maximum number of years is only 24 months. If the resident fails to find a job in 24 months, he should receive the minimum living allowances instead of unemployment compensation. As farmers have never bought any unemployment insurance, the unemployment compensation of 2 years at most can be ignored and they can receive the minimum living allowance of 2 years. By reference to the minimum living allowances from the time of unemployment to the retirement, the equation is as follows:

$$Yuc=M\times(dm\times Bm+dw\times Bw-a)$$

In this equation, Yuc—per capita job security price
M— minimum living allowance per capita
a— average age of citizens
dm—age of retirement of male citizens
Bm—the proportion of male population to the total population
dw—age of retirement of female citizens
Bw—the proportion of female population to the total population

In this equation, the value of M, the minimum living allowance per capita, is the sum of annual basic living expenses per capita, annual basic medical care expenditure per capita (referring to the treatment of common diseases, not including treatment for major diseases) and annual educational expense per capita under the current education level (including compulsory education and job training).

3.4 Calculation of Property Inheritance Price

The Law of the People's Republic of China on Land Contract in Rural Areas currently in effect has given farmers the long-term right of use of the collective land and their right of use of land is the survival guarantee that farmers leave with their offspring. The property inheritance value of the original farmland can be calculated by calculating the education expenses that ensure the offspring of farmers receive the same education as that of the local urban residents and enable them to have the same career competence as the offspring of the urban residents. The education expenses refer to the expenses that should be used to pay for the job training which enables the offspring of farmers to receive the same education as that of the urban residents other than the nine-year compulsory education. The equation is as follows:

$$Yi=A\times(D1-D2)\times T$$

In this equation, Yi— per capita property inheritance price
A—the annual education expense for citizens per capita between the age of 16 and 25
D1—the average years of education of urban residents between the age of 16 and 25
D2—the average years of education of rural residents between the age of 16 and 25
T—the proportion of rural population between the age of 16 and 25 to the total population

Due to the great difference in the development of different areas and the ideology of people in China, the years of education of the offspring of the farmers in some areas may reach or exceed the average years of education of the offspring of the urban residents. In that case, as the equation of job security price above has included the

education expense into the minimum living allowance per capita, M, double counting here should be avoided. That is, when D2≥D1, Yi=0.

3.5 Determination of Social Security Price of Farmland

In those areas where farmers have other noneconomic income, land provides part of security function for farmers and the land requisition price should also include this part of social security price. Hence, the social security price per capita is related to the proportion of land revenue of farmers to their total revenue. The equation is as follows:

$$Ysp = (Ye + Ym + Yuc + Yi) \times CN / C2$$

In this equation, Ysp—social security price of land per capita
Ye—endowment insurance price per capita
Ym—medical insurance price per capita
Yuc—job security price per capita
Yi—property inheritance price per capita
CN—annual net income of land of farmers
C2—annual total net income of farmers

The social security price of land per capita is obtained by calculation in the equation above, while the social security price of land in a unit area is related to land area each farmer owns. The equation is as follows:

$$V = Ysp/Q$$

In this equation, V—average social security price of a unit area
Ysp—average social security price of land per capita
Q—land area a farmer owns per capita

At present, in some areas where fine agriculture or characteristic agriculture is developed, the income of farmers from farming is far more than that of the ordinary land cultivation. However, fine agriculture and characteristic agriculture are the key supported industries of the local government, the production and management land of which will not be requisitioned. Therefore, this kind of sampling point should be eliminated and not be calculated as an exception.

In the process of farmland requisition, the insurance function of farmland to farmers is often ignored or in the actual operation, the calculation method of retirement pension of insurance companies is adopted, which result in some problems such as the extremely low compensation price. In the process of calculating endowment insurance price, medical insurance price, job security price and property inheritance price, some factors such as the inflation of prices and inflation are ignored. With the development of economy, the inflation rate and inflation of prices might change dramatically and modification should be made according to the practical situation when the social security price is calculated. Whether the structure of the social security function of farmland is complete is determined by whether the social security system of China is complete. With the perfection of social security system of China and the disappearance of limit between urban areas and rural areas, the structure of social security price of farmland will change accordingly.

References

1. Dai, S.: Constructing Two-way Land Requisition Compensation System: Requisition and Purchase. China Real Estate Information 6, 61–63 (2004)
2. Li, L.: Predicament and Way Out of Land-lost Farmers in Guangxi Province. South Land Resources 5, 8–10 (2004)
3. Liu, Q., Li, Y., Liu, J.: Discussion on Rational Compensation Standard of Requisition of Suburb Land in China. Economic Forum 14, 10–11 (2004)
4. Guo, L., Chen, L.: Constructing Compensation and Placement Mechanism with Social Security Function for Land Requisition. Real Estate in China and Abroad Herald 18, 20–22 (2003)

Erratum: Water Temperature Forecasting in Sea Cucumber Aquaculture Ponds by RBF Neural Network Model

Shuangyin Liu[1,2,3], Longqin Xu[1], Ji Chen[2,3], Daoliang Li[2,3], Haijiang Tai[2,3], and Lihua Zeng[2,3,4]

[1] College of Information, Guangdong Ocean University, Zhanjiang Guangdong 524025, China
[2] China-EU Center for ICT in Agriculture, China Agricultural University, Beijing 100083, China
[3] Beijing Engineering Research Center for Agricultural Internet of Things, China Agricultural University, Beijing 100083, China
dliangl@cau.edu.cn
[4] College of Mechanical and Electrical Engineering, Agricultural University of Hebei, Banding 071001, China
hdlsyxlq@126.com

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 425–436, 2013.

DOI 10.1007/978-3-642-36124-1_54

The paper "Water Temperature Forecasting in Sea Cucumber Aquaculture Ponds by RBF Neural Network Model" authored by Shuangyin Liu, Longqin Xu, Ji Chen, Daoliang Li, Haijiang Tai and Lihua Zeng, DOI 10.1007/978-3-642-36124-1_51, appearing on pages 425-436 of this publication has been retracted due to plagiarism. It is a plagiarized version of the paper "Water Temperature Prediction in Sea Cucumber Aquaculture Ponds by RBF Neural Network Model", authored by Min Sun, Ji Chen, and Daoliang Li, published in the proceedings of the IEEE International Conference on Systems and Informatics (ICSAI 2012) held in Yantai, China on May 19-20, 2012; DOI 10.1109/ICSAI.2012.6223239.

The original online version for this chapter can be found at http://dx.doi.org/10.1007/978-3-642-36124-1_51

Author Index