Research of Weeds Classification System Based on Shape Feature

Liang Gao
Office of Scientific and Technological Management, Agricultural University of Hebei, Baoding 071001, P.R.China
Email: 1301683942@QQ.com

Dongming Li
College of Mechanical and Electrical Engineering, Agricultural University of Hebei, Baoding 071001, P.R.China
Email: Ldmmail@163.com

Yihua Zhang
Jibei Baoding Electric Power VOC. & TECH. College, Baoding, 071051, P.R.China
Email: Dubo80@126.com

***Abstract*—The paper studied weeds classification system using BP neural network and 6 shape parameters (the ratio of the width and the length, complete degrees, the roundness, the rectangle, ratio of the framework proportion and frame perimeter) based on such characteristic parameters as weed leaf area, perimeter, minimum bounding rectangle, circumcircle, equivalent oval as input feature vectors; and meanwhile trained and improved the system. The experiment results showed when the plant coverage was low, the classification system could identify different weeds better; otherwise, the correct recognition rate was lower.**

***Index Terms*—Weeds classification, leaf, shape parameters, neural network, coverage**

I. INTRODUCTION

The shape is a very intuitive and important feature describing the image content, so the shape of the weeds is important information source for recognizing weeds. The leaves' morphology is many and varied, and different weeds blades have specific shapes. Area, perimeter, minimum bounding rectangle, circumcircle, equivalent oval, etc, are commonly used [1-4]. As shown in Figure 1.

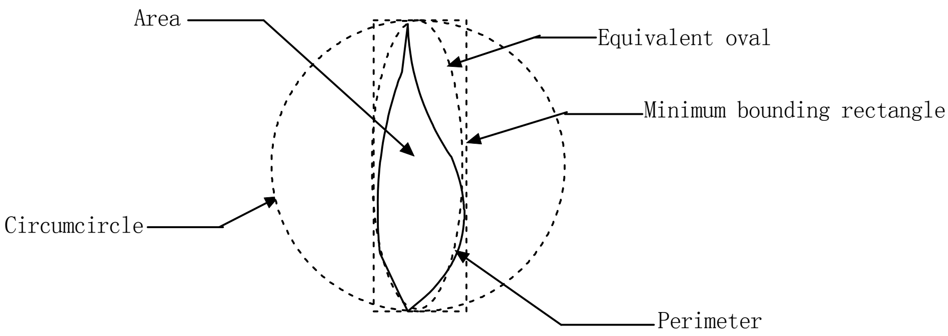


Figure 1. Leafage shape feature

Identifying the most effective characteristic parameters among many features was the key to feature selection and extraction[5]. Shape characteristic parameters had not uniform definition, so the shapes of the object, the shape differences to distinguish effectively the objects and the parameters gained conveniently, quickly could be used as the shape feature parameters. Therefore, there were many shape description methods that included not only describing on the area, edge directly but describing other lines or graphics that could represent the shape of an object, and even producing some characteristic parameters series by mathematical methods [6].

The paper studied the wheat and associated weeds and used MATLAB7.0 as analytical tools. The differences of the shape existed on the weeds and wheat was analyzed based on preliminary studies and the weeds classification system was designed, thus achieving the classification of weeds.

II. EXTRACTION METHOD OF THE SHAPE CHARACTERISTIC PARAMETERS

In the weed identification many shape characters were non-dimensional which included the ratio of the width and the length, complete degrees, the roundness, the rectangle and so on [7-9]. The paper also advised to use the ratio of the framework proportion, frame perimeter as

doi:10.4304/jsw.8.7.1736-1742

characteristics parameter to identify weeds besides using common characteristics of parameters.

Extraction of framework also known as axis transform or incinerated grass technology was a computing related with refinement. In mathematical morphology, the skeleton was an important characteristic expressing the structure and shape of the planar region as well as better describing the differences between the crop and weed plants [10-13].

The corrosion and expansion were two types of basic operation in the morphology, on which the arithmetic of skeleton extraction was based. Specific operation could be expressed as:

$$Z(X) = \bigcup_{k=0}^{K} Z_k(X) \tag{1}$$

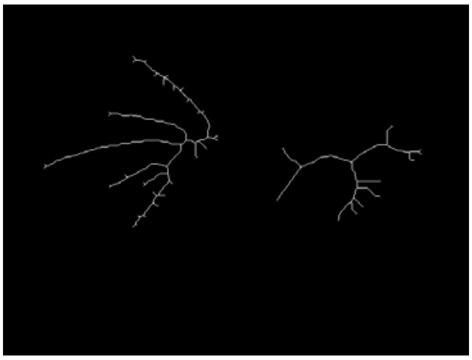
$$Z_k(X) = \bigcup_{k=0}^{K} \{(X \otimes kS) - [(x \otimes kS) \otimes S] \oplus S\} \tag{2}$$

In the formula, K referred to the number of iterative steps; beyond K times iteration, the collection X would be corroded as the empty set; $(X \otimes kS)$ represented the k times iteration for X through S, and that meant continuous corrosion. In framework extraction process, the structural element S for:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{3}$$

The binary image of monoclonal wheat and amaranthus lividus obtained by dividing was shown in Figure 2 (a), and the image after the framework extraction was shown in Figure 2 (b).

(a) Binary image



(b) Framework image

Figure 2. Binary image and framework image of Wheat and Amaranthus lividus

The above shape features extracted from the images would be translated by the actual number that each pixel represented. Calculations tended to be more complex due to the actual size differences which each pixel represented, so the study introduced a dimensionless value. Shape characteristic dimensionless was independent of scale, which provided greater flexibility and convenience for analysis.

A. Edge Detection

Edge means for collection of those pixels in which pixel gray of the image has the step change or roof change [14]. It was necessary to obtain this information before typical shape characteristics of the weed leaf area, perimeter were analyzed, so the first step was the edge detection in the plants image understanding and analysis. This paper used LOG edge detection operators; let's take binary image in Figure 2 (a) for example, and the image after edge detection was shown in Figure 3.

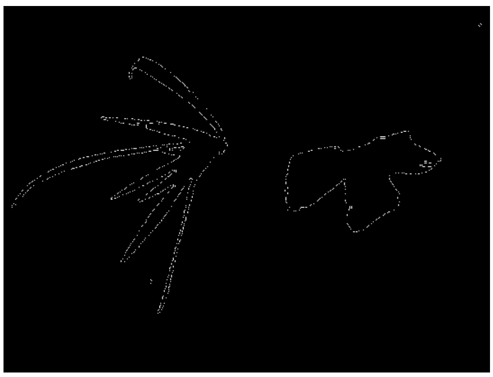


Figure 3. Image of using LOG edge arithmetic operators

B. Tag connected Domain

The feature extraction of more than one region of the image could be divided into two steps: the tagging and feature extraction. At first, different regions of the image should be tagged, and then extracted the characteristics from different regions to complete the quantitative description. This experiment used pixel tags to make a

8-connectivity regional markers on weeds and used 8-connectivity edge agreement on 2(a) the weed binary image to mark weeds with pseudo-color, after that the picture was shown as Figure 4.

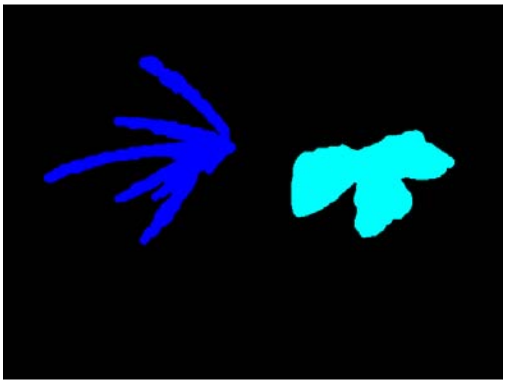


Figure 4. Area sign result

C. Shape Feature Extraction

Certain characteristics of all the connected regions could be extracted after connective regions were calibrated. To reduce the impact on which the number of pixels of different pictures made statistics, the non-dimensional characteristics were extracted besides the area and perimeter.

III. COMPARISON AND ANALYSIS OF SHAPE CHARACTERISTIC PARAMETERS

The study proved that the differences of shape parameters between various weeds and the wheat were obvious, such as the ratio of width and length of large leafage category weeds, so we could use the differences of shape characteristics parameters between the wheat and weeds to distinguish them. The wheat sprouts and several kinds of common weeds with a typical representative were shown in Figure5, which was the study object of this experiment.

(a) Wheat

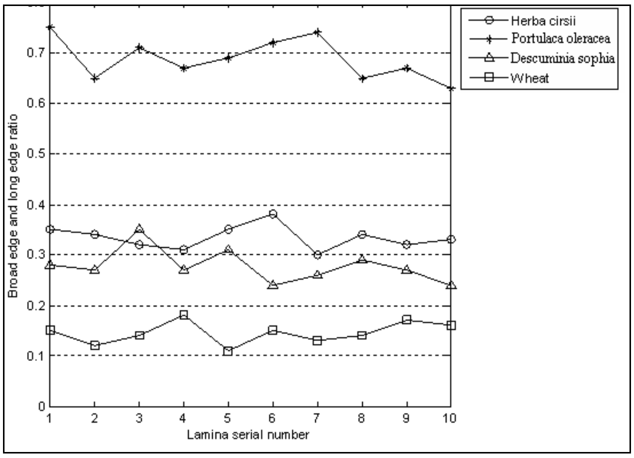
(b) Portulaca oleracea

(c) Descuminia sophia(L.)

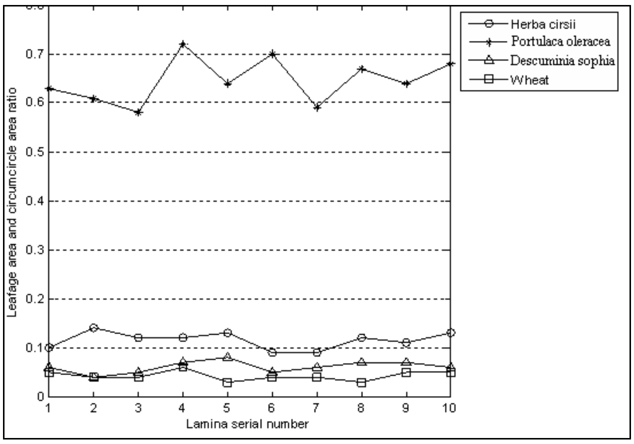
(d) Herba cirs

Figure 5. Wheat and weeds image compare

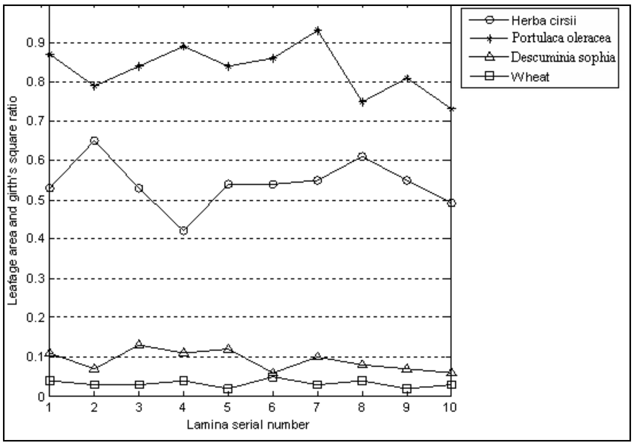
In each type of weeds, we selected 10 laminas with typical character separately, coded them and extracted the shape characteristics using above-mentioned method. The eigenvalues after extraction were expressed in line charts which were shown in Figure 6. According to the values, the ratio of broad and long edge and the ratio of framework and area of the wheat were significantly different from the other three kinds of weeds'. So the two eigenvalues could be used to identify wheat. However, in practice, there were many types of weeds, so the weeds and wheat couldn't be distinguished if we only used a single non-dimensional shape characteristics value; Moreover, this method was not only used for wheat, but the majority of crops. Therefore, we should fit these shape characteristics elements together into the shape feature set in a proper way, and then describe the shape characteristics of the plant leaves, if so, we would achieve a better recognition effect.



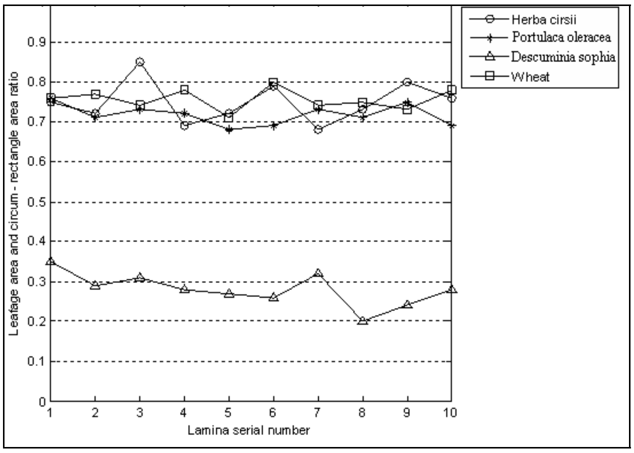
(a) Broad edge and long edge ratio



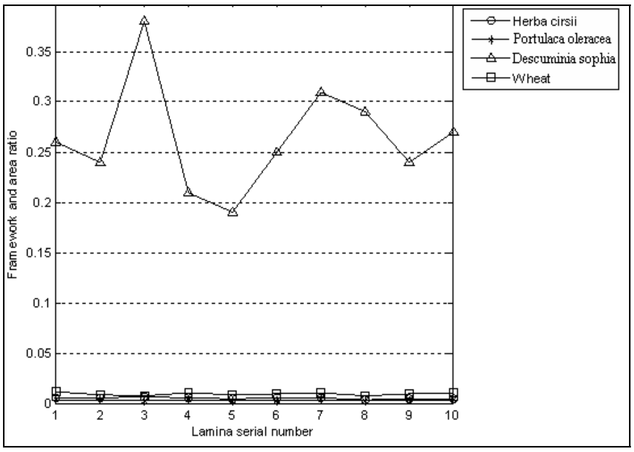
(b) Leafage area and circumcircle area ratio



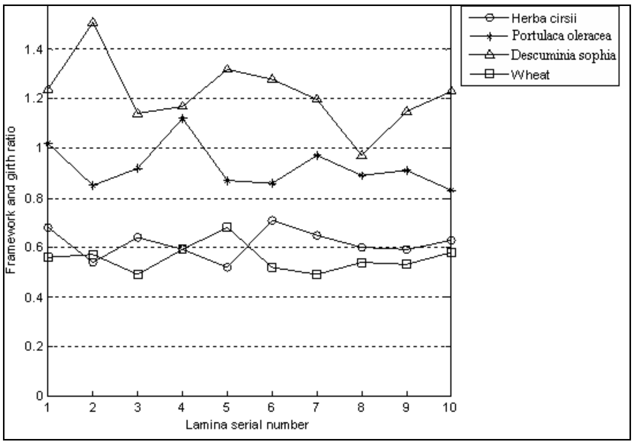
(c) Leafage area and girth's square ratio



(d) Leafage area and circum- rectangle area ratio



(e) Framework and area ratio



(f) Framework and girth ratio

Figure 6. Feature parameter graph

IV. ESTABLISHMENT AND TEST OF THE WEED CLASSIFICATION SYSTEM

A. Classification system based on BP Neural Network Structure

BP neural network structure was shown in Figure 7: the network model that was divided into three layers achieved the idea of multi-layer network learning. When an input pattern given to network was sent from the input layer unit to the hidden layer unit where it was conducted successively, and then was sent to the output layer unit. Through the output layer unit's processing, an output mode called forward propagation was generated [15-17]. Compared with the desired output mode, if the output response had error which did not meet the requirements, it would be sent back-propagation. It meant error values were sent back layer-by-layer along the connection path and the layers of connection weights fixed [18][19].

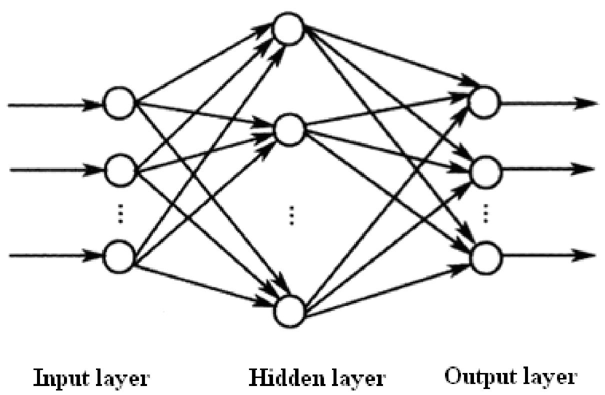


Figure 7. Three layers BP neural network structure

B. Design and Training of Neural Network

The experiment considered synthetically implicit network layers, hidden layer nodes, activation function, error function forms as well as the learning rate and designed a weed classifier to detect weeds or crops.

① First, the network input and the target sample should be required. According to above procedure, 40 groups shape characteristics samples (see Figure 5 line map specific data) were obtained, which respectively contained 6 non-dimensional numerical value like the ratio of the width and the length, complete degrees, the roundness, the rectangle , ratio of the framework proportion, frame perimeter. Different data sets corresponded to different plants, such as Herba cirsii, Galium tricorne, artemisiaapiacea and wheat. The details were presented in Table1.

② P was for sample input vector of the network and T was objective vector of the network. The data gained from Table I was as follows:

$$
\begin{aligned}
P = [&0.35,0.10,0.53,0.75,0.0057,0.68;\\
&0.75,0.63,0.87,0.87,0.0031,1.02;\\
&0.28,0.06,0.11,0.35,0.26,1.24;\\
&0.15,0.05,0.04,0.76,0.012,0.56;\\
&0.34,0.14,0.65,0.72,0.0051,0.54;\\
&0.65,0.61,0.79,0.71,0.0040,0.85;\\
&0.27,0.04,0.07,0.29,0.24,1.51\\
&\qquad\qquad\ldots\ldots\\
&0.16,0.05,0.03,0.78,0.011,0.58]\\
T = [&0,0;0,1;1,0;1,1;0,0;0,1;1,0;\ldots\ldots;1,1]
\end{aligned}
$$

③Network training. After the network was created and the network error was assumed, network was trained and met the requirement, and the weeds classification system was finally achieved.

C. Weeds Classify System Test

Neural network classify system was used for weed identification experiment. In accordance with previous experimental results, neural network of 0.002 error-learning was selected. The purpose of network testing was to determine whether the network met the requirements of practical applications. It should be noted that testing data should be different from training sample data. Otherwise, the results of the test were always satisfied.

Weed identifiable experiment had two cases: one was high leaves coverage (more than 30%), the other was relatively low leaves coverage (less than 10%). Through the experiment of 40 plants for each type, the results showed in Table II and III.

The experimental results showed: when the coverage was high, the overall average correct recognition rate was 73.75%, including 77.5% correct crop recognition rate and 72.5% correct weed recognition rate; when the coverage was low, the overall average correct recognition rate was 91.25%, including 90% crop recognition rate and 91.7% weed recognition rate.

V. CONCLUSION

TABLE I.
FOUR PLANTS STYLEBOOK DATA

Wwatch number	Swatch input features data						Regimentation
	Broad edge and long edge ratio	Leafage area and circumcircle area ratio	Leafage area and girth's square ratio	Leafage area and circum-rectangle area ratio	Framework and area ratio	Framework and girth ratio	
1	0.35	0.10	0.53	0.75	0.0057	0.68	Herba cirsii
2	0.75	0.63	0.87	0.87	0.0031	1.02	Herba Portulacae
3	0.28	0.06	0.11	0.35	0.2600	1.24	Descuminia Sophia
4	0.15	0.05	0.04	0.76	0.0120	0.56	Wheat
5	0.34	0.14	0.65	0.72	0.0051	0.54	Herba cirsii
6	0.65	0.61	0.79	0.71	0.0040	0.85	Herba Portulacae
7	0.27	0.04	0.07	0.29	0.2400	1.51	Descuminia Sophia
……	……	……	……	……	……	……	……
40	0.16	0.05	0.03	0.78	0.0110	0.58	Wheat

TABLE II.
ENTIFY RESULT OF BESTROW RATIO WAS HIGH

Plant name	number	Identify right	Identify error	Right ratio (%)	Error ratio (%)
Wheat	40	31	9	77.5	22.5
Herba Portulacae	40	29	11	72.5	27.5
Descuminia sophia	40	28	12	70.0	30.0
Herba cirsii	40	30	10	75.0	25.0

TABLE III.
DENTIFY RESULT OF BESTROW RATIO WAS LOW

Plant name	number	Identify right	Identify error	Right ratio (%)	Error ratio (%)
Wheat	40	36	4	90.0	10.0
Herba ortulacae	40	35	5	87.5	12.5
Descuminia sophia	40	38	2	95.0	5.0
Herba cirsii	40	37	3	92.5	7.5

The paper analyzed how to identify crops and weeds with different shapes, studied the establishment and training of the neural network which had relatively better recognition rate for the average image, and achieved more satisfactory results. However, there were deficiencies that should be further researched. Recommendations must continue to be considered:

(1) At different stages, the leaf shape of crops was significantly different and the object of study of this subject was limited. Aimed at the features of different crops at different stages, a database of the shape features should be set up and the professional weed identifier should be developed.

(2) The recognition algorithm should be further improved and simplified, the system running speed and the recognition accuracy should be improved and meanwhile the recognition rate should be further increased when blade coverage was high.

(3) The paper studied static images. Dynamic processing of the image should be further studied, the images should be disposed as collected and the impact factors in motion should be studied.

ACKNOWLEDGEMENTS

This work was supported by Department of Science and Technology of Hebei Province (No. 10220925) and Hebei Engineer Research Center of Information Technology for Rural Area.

REFERENCES

[1] Lamm R D, D C Slaughter, D K Giles, “Precision weed control system for cotton”, *Transactions of the ASAE,* vol. 1, pp. 231-238, 2002.

[2] Udomsit Sukakanya, Kriengkrai Porkaew, “Modeling A Generic Web Classification System Using Design

Patterns", *Journal of Computers,* Vol 6, No 10, pp. 2212-2220, 2011.

[3] Cheng H, Shi Z X, Yao W, et al, "Corn breed recognition based on support vector machine", *Transactions of the CSAM*, vol. 40, no. 4, pp. 180-183, 2009.

[4] DongMing Li, Bo Du, and Li Zhang, "The Design of the Weeds Classification System Based on BP Neural Network", *Advances in Computer Science and Information Engineering 2012(Conference*), vol. 2, pp. 625-630, 2012.

[5] Liu L B, Zhou G M, "Identification method of rice leaf blast using multilayer perception neural network", *Transactions of the CSAE*, vol. 25, no. 2, pp.213-217, 2009.

[6] Gai S P,GaiI W L, Wang R X, "Identification Seed Puritv of Six Maize Hybrids by SSR Markers", *Seed*, vol. 29, no. 7, pp.44-47, 2010.

[7] Paliwal J, Visen N S, Jayas D S, et al, "Cereal grain and dockage identification using machine vision", *Biosystems Engineering*, vol. 85, no. 1, pp.51-57, 2003.

[8] Astrand B, Baerveldt A J, "A vision based row-following system for agricultural field machinery", *Mechatronics*, vol. 15, no. 2, pp.251-269, 2005.

[9] Granitto P M, Verdes P F, Ceccatto H A, "Large-scale investigation of weed seed identification by machine vision", *Computers and Electronics in Agriculture*, vol. 47, no. 1, pp.15-24, 2005.

[10] Meyer G E, Neto J C, Jones D D, et al, "Intensified fuzzy clusters for classifying plant, soil and residue regions of interest from color images", *Computers and Electronics in Agriculture*, vol. 42, no. 3, pp. 161-180, 2004.

[11] Victor. Alchanatis, Leonid. Ridel, Amots. Hetzroni, et al, "Weed detection in multi-spectral images of cotton fields", *Computers and Electronics in Agriculturem*, vol. 47, no. 3, pp. 243-260, 2005.

[12] Persson Maria, Åstrand Björn, "Classification of crops and weeds extracted by active shape models", vol. 100, no. 4, pp.484-497, 2008.

[13] Zhang Yun, Slaughter David C., Staab, Erik S., "Robust hyperspectral vision-based classification for multi-season weed mapping", ISPRS *Journal of Photogrammetry and Remote Sensing*, vol. 69, pp. 65-73, 2012.

[14] Pereira Luís A.M., Nakamura Rodrigo Y.M., de Souza Guilherme F.S., Martins Dagoberto, Papa João P., "Aquatic weed automatic classification using machine learning techniques", *Computers and Electronics in Agriculture* , vol. 87, pp. 56-63, 2012.

[15] Cruz-Ramírez M., Hervás-Martínez C., Jurado-Expósito M., López-Granados F., "A multi-objective neural network based method for cover crop identification from remote sensed data", *Expert Systems with Applications*, vol. 39, no. 11, pp. 10038-10048, 2012.

[16] Li Hui, Qi Lijun, Zhang Jianhua, Ji Ronghua, "Recognition of weed during cotton emergence based on principal component analysis and support vector machine", *Transactions of the Chinese Society of Agricultural Machinery*, vol. 43, no. 9, pp. 184-189+196, 2012.

[17] Md. Khalilur Rhaman, Tsutomu Endo, "Recurrent Neural Network Classifier for Three Layer Conceptual Network and Performance Evaluation", *Journal of Computers,* Vol. 5, No. 1, pp. 40-48, 2010.

[18] Lihua Fu, Meng Zhang, Zhihui Liu, Hongwei Li, Tingting He, "Adaptive Hybrid Feature Extraction for Leaf Image Classification by Support Vector Machine", *Journal of Convergence Information Technology*, Vol. 7, No. 2, pp. 160-166, 2012.

[19] Fan Lin, Wenhua Zeng, Jianbing Xiahou, Yi Jiang, "Optimizing for Large Time Delay Systems by BP Neural Network and Evolutionary Algorithm Improving", *Journal of Software*, Vol 6, No 10, pp. 2050-2055, 2011.

Liang Gao received his B.S. degree from College of Mechanical and Electrical Engineering, Agricultural University of Hebei in 1995. He received his M.S. degree from Agricultural University of China in 1998. He is an associate professor in Agricultural University of Hebei. His research interests include Image Processing and agricultural informatization.

Dongming Li, Correspondence author. He received his B.S. degree from College of Agricultural Engineering, Shenyang Agricultural University in 2002. He received his M.S. degree from College of Mechanical and Electrical Engineering, Agricultural University of Hebei in 2008. He is an associate professor in Agricultural University of Hebei. His research interests include computer application in power system, and agricultural informatization. He has authored and coauthored more than ten journal and conference papers in these areas.

Yihua Zhang received her M.S. degree from College of Mechanical and Electrical Engineering, Agricultural University of Hebei in 2008. Her current research interests include computer application, application of artificial intelligence in power system, and control of power distribution grids.