

CLASSIFICATION OF WEED SPECIES USING ARTIFICIAL NEURAL NETWORKS BASED ON COLOR LEAF TEXTURE FEATURE

Zhichen Li, Qiu An, Changying Ji *

College of Engineering, Nanjing Agricultural University, Nanjing, Jiangsu, 210031, China

** Corresponding author, Address: College of Engineering, Nanjing Agricultural University, Nanjing, Jiangsu, 210031, China, 025-58606571 Email: chyji@sohu.com*

Abstract: The potential impact of herbicide utilization compel people to use new method of weed control. Selective herbicide application is optimal method to reduce herbicide usage while maintain weed control. The key of selective herbicide is how to discriminate weed exactly. The HIS color co-occurrence method (CCM) texture analysis techniques was used to extract four texture parameters: Angular second moment (ASM), Entropy(E), Inertia quadrature (IQ), and Inverse difference moment or local homogeneity (IDM).The weed species selected for studying were *Arthraxon hispidus*, *Digitaria sanguinalis*, *Petunia*, *Cyperus*, *Alternanthera Philoxeroides* and *Corchoropsis psilocarpa*. The software of neuroshell2 was used for designing the structure of the neural network, training and test the data. It was found that the 8-40-1 artificial neural network provided the best classification performance and was capable of classification accuracies of 78%.

Keywords: weed, texture feature, artificial neural network, neuroshell2

1. INTRODUCTION

The application of herbicides in agricultural crops has been practiced for a long time. The utilization of herbicides result in significant increasement in crop production. However the environmental pressure grows day by day with the enhancement of herbicides utilization. How to reduce the herbicides and maintain the crop production has stimulated many researchers to study.

Thompson et al. (1991) suggested that selective herbicide application (variable-rate or intermittent application technology) to weed-infested area could result in decrease of herbicide use rather than the entire field. Since herbicides are applied only to weed patches, large areas of the field remain untreated. For the technology to be effective, weeds species and spatial location must be identified from crop field. The application of machine for weed identification has increased rapidly for the recent two decades.

Burkes (2000) suggested three steps for weed discrimination using machine vision. First digital images of weed and crop must be acquired. Second the characteristic feature must be generated for discrimination between several weed species. Finally, a sufficient classification algorithm must be implemented using the feature as input and the weed class as output.

The use of color features in classical gray image texture analysis techniques was first reported by Shearer (1986). Shearer and Holmes (1990) reported accuracies of 91% for classifying different types of nursery stock by the color co-occurrence matrix (CCM) method. The traditional gray image texture features were expanded by Shearer(1986) to utilize HIS color co-occurrence method consisted of three co-occurrence matrices, one each for the hue, saturation, and intensity color features.

Meyer et al. (1998) used red, green, blue (RGB) true color to produce an excess green color feature for discriminating between four different species of weeds and soil regions with a 99% accuracy. They used the traditional gray scale co-occurrence matrix to generate four texture statistics: angular second moment, inertia, entropy, and local homogeneity. They observed classification accuracies of 93% for grasses and 85% for broadleaf categories when using the four texture feature. Tang et al. (1999) used a Gabor wavelets-based feature extraction method and neural net works to classify images into broadleaf and grass categories. They achieved 100% classification accuracies when testing 20 sample images from each of the two categories.

The resulting 11 texture feature equations are defined by Burks . The Color Co-occurrence Method (CCM) was also utilized by Burks to discriminate between six different classes of groundcover. The result showed that the CCM texture statistic procedure was able to classify five species of weed and soil with accurate of 93%.

The CCM method was utilized by Burks to evaluate three different neural network classifiers for real-time weed control systems. The result showed that the BP neural-network classifier provided the best classification performance with accuracies of 97%.

Detecting inter-row weed was researched by Mao Wenhua, the location feature of crop within field was used for discriminating weed with accuracies of 86%. The texture and position feature were used for studying by Cao Jingjing, the result of the research showed that the correct classification of weed was 93%.

Artificial Neural Network (ANN) is a data processing system based on the structure of biological neural system. Prediction with ANN is not like modelling and simulation, but by learning from the data generated experimentally or using validated models, ANN differ from conventional programs in their ability to learn about the system to be modelled without priori knowledge of the process variables relationship.

A discriminating method was developed based on image features of the peanut kernels and artificial neural network by Chen Hong. The image characteristics parameters such as the color parameters HIS, and veins characteristics parameters RW,GW,BW were used as the input to the neural network set up by MATLAB. The results of the experiment show that the accuracies of the identification of the method are 95% for normal peanut kernels, 90% for slightly moldy peanut kernels and 100% for severely moldy peanut kernels, respectively.

BP neural network model was developed to map the relationship of process variables and distributing quality coefficient of Profile Modeling Spray of the Fruit Trees by Lin Huiqiang. The result shows that the correlation coefficient R between simulating outputs of the BP network and the results of experiments is 0.99, which has wide adaptability. And the BP network can be used conveniently to carryout various quantificational calculations.

The improved LVQ neural network algorithm was applied in the process to identify the grade of apples by Bao Xiaoan, The research result was that the correct identification rate of 88.9% with good stability.

A three layer feed forward neural network was established by Chen JiaJuan. The purpose of the research was to identify the corn and the background of field. The experiment showed that corn could be recognized correctly by using this automated measurement system of corn leaf color value, the judging accuracy could attain 91.6%, and corn leaf color value could also be calculated correctly.

The objectives of this paper were to develop an ANN to classify species of six weed using four texture features. The classification process using the texture feature data is very simplicity, convenience, fast and accurate. This will also assist in identifying weed from field crop real-time based on machine vision. The work of designing the structure of neural network and train or test the neural network is all done by the software neuroshell2.

2. COLOR TEXTURE FEATURE EXTRACTION METHOD

The image texture feature analysis technique selected for this study includes three processes based on the Color Co-occurrence Method (CCM). One, the image is transformed from RGB to HIS using the following equation. The purpose of this process is to reduce the storage and computing time. It can be classified the observation with the accuracy of 93% only using hue and saturation texture feature by the research of Burks^[7]. Only hue and saturation texture feature was used in this study.

$$I = (R + G + B) / 3 \quad (1)$$

$$S = 1 - \frac{3 * \text{Min}(R, G, B)}{R + G + B} \quad (2)$$

$$H = \text{ARCCOS} \left\{ \frac{(R - G) + (R - B)}{2[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \quad (3)$$

Two, one CCM matrix is generated from the HIS image. The matrix for intensity was represented by the function $P(i,j,d,\theta)$. The CCM measures the probability that a pixel at one particular gray-level will occur at a distinct distance and orientation from any pixel given that pixel has a second particular gray-level. In the function $P(i,j,d,\theta)$ i represents the gray-level of location (x,y) , and j represents the gray level of the pixel at a distance d and an orientation of θ from location (x,y) . 1 was selected for d and 0 was selected for θ in this study. For example: the image and CCM matrix are showed in figure1. The process may be understood by looking for the number of occurrence of zero adjacent to three in $I(x, y)$ in figure1 (a), we find a total of two occurrences of zero adjacent to three. This corresponds to the two value at CCM matrix location $(0,3)$.

$$I(x, y) = \begin{bmatrix} 0 & 0 & 3 & 1 \\ 2 & 1 & 0 & 2 \\ 3 & 2 & 0 & 3 \\ 1 & 2 & 1 & 3 \end{bmatrix}$$

(a)

$$P(i, j, 1, 0^\circ) = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{bmatrix}$$

(b)

Fig.1. the image and CCM matrix example: (a) gray-level image, (b) CCM matrix

Three, two CCM matrix of hue and saturation is used to generate eight texture features using the following equation.

Angular second moment (ASM):

$$H_{ASM} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} [P(i, j)]^2 \quad (4)$$

Entropy(E):

$$H_E = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \ln (P(i, j)) \quad (5)$$

Inertia quadrature (IQ):

$$H_{IQ} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j)(i - j)^2 \quad (6)$$

Inverse difference moment (IDM):

$$H_{IDM} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{P(i, j)}{1 + (i - j)^2} \quad (7)$$

Where:

$$P(i, j) = \frac{P(i, j, 1, 0)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j, 1, 0)} \quad (8)$$

And N_g = the total number of attribute levels. The calculate method for the saturation is the same as the hue equations.

3. MATERIAL AND METHOD

3.1 Weed Species

Six weed species common to south china row crops were selected for this study. The weed species were *Arthraxon hispidus*, *Digitaria sanguinalis*, *Petunia*, *Cyperus*, *Alternanthera Philoxeroides* and *Corchoropsis psilocarpa*. Every of the six species were grown from one experimental field of Nanjing Agricultural University. The weed species have been grown under normal ambient conditions until they reached an appropriate maturity level. The number of different size leaves for studying were 14, 11, 8, 9, 8, 10 respectively for the above species.

3.2 Image Acquisition and processing System

A three CCD camera of Olympus N438 was used for collecting the digital RGB images. The weed leaves were put on a white paper immediately after getting from the field. The image was taken outdoors under natural sunlight.

The image was acquired and digitized into 24-bit (RGB) images with a resolution 640*480. All of the leaves images were recorded on an U-disc of the camera and then downloaded into the personal computer of PentiumIII 700 MHz CPU .

Six species of weed were captured in 640*480 pixel digital images which included both white paper and weed leaf. The images for every species of weed are presented in figure 2. A computer software program using VC++ was developed to generate RGB image files, create a HSI color model of the image, generate color co-occurrence matrices, and calculate CCM statistics for the specified matrices.

3.3 NeuroShell 2

NeuroShell 2 is a software program by Ward System Group.Inc. The program can mimic the human brain's ability to classify patterns or to make predictions or decisions based upon past experience.

NeuroShell 2 combines powerful neural network architectures, a Microsoft® Windows icon driven user interface, sophisticated utilities, and popular options to give users the ultimate neural network experimental environment. It is recommended for academic users only, or those users who are concerned with classic neural network paradigms like backpropagation. Users interested in solving real problems should consider the NeuroShell Predictor, NeuroShell Classifier, or the NeuroShell Trader. The central interface of NeuroShell 2 is as example of figure 3.

NeuroShell 2 enables you to build sophisticated custom problem solving applications without programming. You tell the network what you are trying to predict or classify, and NeuroShell 2 will be able to "learn" patterns from training data and be able to make its own classifications, predictions, or decisions when presented with new data.



Arthraxon hispidus



Digitaria sanguinalis



Petunia



Fig.2 Six species of weed

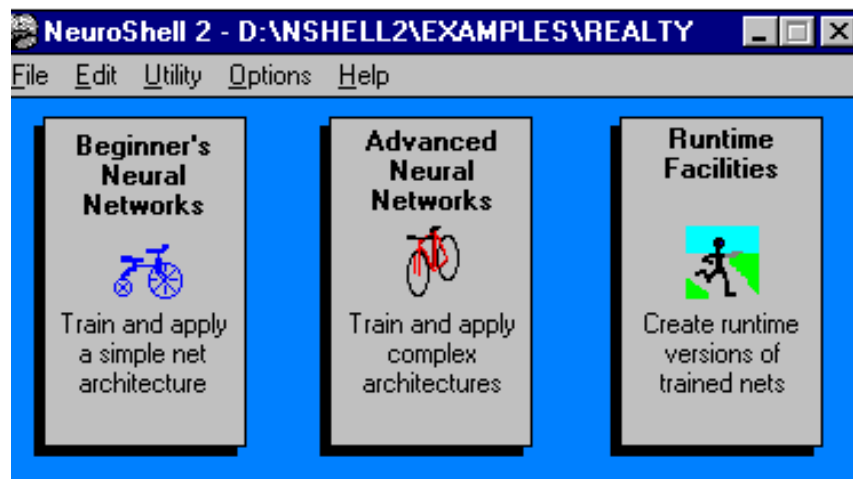


Fig.3 The central interface of NeuroShell 2

4. RESULT AND DISCUSSION

4.1 Neural network model

Three layered ANN was developed to classify leaf species of six weed. Eight in the input layer represented eight input parameters and one nodes in output layer to classify *Arthraxon hispidus*, *Digitaria sanguinalis*, *Petunia*, *Cyperus*, *Alternanthera Philoxeroides* and *Corchoropsis psilocarpa*. The number of hidden layer nodes was chosen according to ANN performance.

The backpropagation algorithm was used for ANN training. The best testing error was taken as criterion to stop training. The errors were

examined by the testing sets, which were not used for ANN training, were applied to evaluate ANN performance. The learning rate is set to value of 0.1.

4.2 Network training and testing

From the generated 60 data, 45 and 15 data sets were randomly selected as Training and testing respectively. After every data set training, ANN weights were adjusted. In the beginning of training, testing error decreased with training process. Training was continued until testing error did not decrease.

4.3 Hidden layers and nodes

In the study, different hidden layer nodes were chosen to select the best production results (Table 1). Learning rate was set at 0.1. Nodes in hidden layer were varied from 20 to 60.

Table 1. Effect of nodes in hidden layer

Nodes of hidden layer	20	30	40	50	60
Accuracy (%)	72	76	78	78	78

4.4 Input variables importance

Every input node was connected with 40 hidden nodes, then to output nodes. The importance of an input variable to ANN predictions was compared by comparing sum of its connecting weights (Table 2).

Table 2. Attribution of the input invariable

Variable	ASM	E	IQ	IDM
Weight	0.312	0.308	0.25	0.13

5. CONCLUSION

Classification species of weed by ANN using simulation data is a simple and convenient method. Classification accuracy could be increased by careful selection of hidden nodes and appropriate of learning rate. In this study, the appreciate structure of artificial neural network is 8-40-1 with optimal learning rate is 0.1. The Angular second moment and entropy of weed texture feature were important to all of the input variables. The inverse difference moment of weed texture feature is the minimum affective to the outputs.

ACKNOWLEDGEMENTS

Funding for this research was provided by Hubei Provincial Department of Education (P. R. China). The first author is grateful to the Wuhan Polytechnic University for providing him with pursuing a PhD degree at the Wuhan University of Technology.

REFERENCES

- Bao Xiaon, Zhang Ruilin , Zhong Lehai. Apple grade identification method based on artificial neural network and image processing. *Transaction of the CSAE* 2004,20(3),109-111
- Cao Jingjing, Wang Yiming, Mao Wenhua et al. Weed detection method in wheat field based on texture and. Position features. *Transaction of the CSAM* 2007.38(4),107-110.
- Chen Jiajuan, Ji Shouwen, Ma Chenglin. Investigation on automated color measurement of corn leaves based on genetic neural network. *Transaction of the CSAE* 2000,16(3),115-117
- Cheng Hong, xiong Lirong, Hu Xiaobo, et al. Identification method for moldy peanut kernels based on neural network and image processing. *Transaction of the CSAE* 2007,23(4),158-161.
- G.E.Meyer, T.Mehta, M.F. Kocher,etl.1998. Textural image and discriminant analysis for distinguishing weeds for spot spraying.*Transactions of ASAE*,41(4),1189-1197
- Lin Huiqiang, Xiao Lei, Liu Caixing, et al. Neural network model for profile modeling spray of chemical to fruit trees and its applications. *Transaction of the CSAE* 2005,21(10),95-99
- Mao Wenhua ,Wang Yiming, Zhang Xiaochao et al. Machine vision system used for real-time detection inter-row weed. *Transaction of the CSAE* 2003,19(5),114-117.
- Ramesh, M. N., Kumar, M. A. & Rao, P. N. S. (1996). Application of artificial neural networks to investigate the drying of cooked rice. *Journal of Food Process Engineering*, 19, 321-329.
- Shearer,S.A., and R.G.Holmes.1990. Plant identification using color co-occurrence matrices. *Transactions of the ASAE* 33(6):2037-2044.
- Shearer.S.A.1986. Plant identification using color co-occurrence matrices derived from digitized images. PH.D. thesis. Ohio university, Agricultural Engineering.
- T.F.Burks, S.A.Shearer, F.A.Payne.2000. Classification of weed species using color texture features and discriminant analysis. *Transactions of ASAE*,43(2),441-448
- T.F.Burks, S.A.Shearer, J.R.Heath,etl.2005. Evaluation of neural-network classifiers for weed species discrimination. *Biosystems Engineering*(2005) 91(3),293-304
- T.F.Burks, S.A.Shearer, R.S. Gates, K.D. Dono.2000. Backpropagation neural network design and evaluation for classifying weed species using color image texture feature.*Transactions of ASAE*,43(4),1029-1037
- Tang L. , L.F. Tian, B.L.Steward, et al. Texture-based weed classification using Gabor wavelets and neural network for real-time selective herbicide applications. *ASAE* ,1999 ,Paper No.993036
- Thompson,J.F.,J.V.Stafford, and P.C.Miller. 1991. Potential for automatic weed detection and selective herbicide application. *Crop Prod.* 10(4):254-259.