

Improved G-K fuzzy clustering segmentation algorithm for rice damaged-spots infested by Rice Leaf Roller

Pei Wang^{1,2,3}, Guohao Jia¹, Zhiyan Zhou^{1,2,3*}

(1. College of Engineering, South China Agricultural University / Guangdong Engineering Research Center for Agricultural Aviation Application (ERCAAA), Guangzhou 510642, China;

2. National Center for International Collaboration Research on Precision Agricultural Aviation Pesticides Spraying Technology (NPAAC), Guangzhou 510642, China;

3. Key Laboratory of Key Technology on Agricultural Machine and Equipment (South China Agricultural University), Ministry of Education, Guangzhou 510642, China)

Abstract: Image segmentation of crops damaged-spots infested by insect pests under natural conditions is very important to realize the precision spray. Due to the influence of uneven lighting and random noise, the traditional method of image segmentation is difficult to achieve the ideal results. In order to overcome the complications mentioned above, an image segmentation algorithm based on OSTU binarization algorithm and improved Gustafson-Kessel (GK) Fuzzy Cluster for the rice damaged-spots infested by the Rice Leaf Roller (*Cnaphalocrocis medinalis* Guenee) was proposed in this paper. Firstly, the ultra-green equation and OSTU was utilized for image preprocessing. Secondly, take the S component of color space HSI (Hue, Saturation, Intensity) which transferred from the target image, and then the improved Gustafson-Kessel Fuzzy Cluster algorithm and Morphological Filtering were utilized to obtain the target area which the rice damaged-spots infested by Rice Leaf Roller. Experimental results showed that the accuracy rate of the proposed segmentation algorithm reached 82.4%. In order to test the effects of segmentation results in classification and recognition, three features, skewness of color feature B and R component, average of S component, were selected. The distinguished effect of each features mentioned above were showed good performance. The classification accuracy rate based on the above three features reached 94%. Efficient results were achieved by using the mentioned above method for images with the influence of uneven lighting, random noise and complex background under natural conditions.

Keywords: paddy field, precision spraying, rice leaf roller, image segmentation, Gustafson-Kessel fuzzy cluster

DOI: 10.33440/j.ijpaa.20190202.51.

Citation: Wang P, Jia G H, Zhou Z Y. Improved G-K fuzzy clustering segmentation algorithm for rice damaged-spots infested by Rice Leaf Roller. *Int J Precis Agric Aviat*, 2019; 2(2): 62–66.

1 Introduction

At present, the utilization rate of pesticides in China is only about 30%, and the improper use of pesticides has caused pollution of agricultural environments such as soil and water bodies, resulting in a vicious circle of agro-ecological chains^[1]. The rational use of pesticides is of great significance to the sustainable development of China's resource environment^[2,3]. The traditional method of pesticide spraying is to spray crops evenly in order to achieve pest control. In order to reduce the use of pesticides and improve the utilization rate of pesticides, it is necessary to carry out accurate sprays on crops. Precision spray enables on-demand drug use, thus eliminating environmental pollution caused to soil, atmosphere and water bodies due to the excessive application of pesticides. Pest identification is an important prerequisite for the precise injection of pesticides^[4].

With the rapid development of computer technology, information technology is more and more widely used in

agriculture. The damaged-spot shape of plants, the fragmentation of the edges and the holes and texture characteristics of the leaves are important source of information for pests and diseases. The same damaged-spot in a plant always has its certain shape characteristics. This allows the shape characteristics of the damaged-spot to identify the type of pest and disease and determine the degree of harm. Many scholars have studied crop pest identification by using image processing technology^[5-8].

Chen et al.^[9] determined the damaged level of cotton pests based on the hole and edge defects of cotton leaves, the experimental results showed that the method could effectively identify the damaged level of cotton pests, and the identification error was less than 0.05. Ma et al.^[10] identified the damaged level of rice blast based on the size and shape of the disease-spot, elliptical model was used to indicate the size and direction of the disease spot characteristics, the results showed that the average accuracy rate of this method was over 90%. Qi et al.^[11] used a neural network based on color values and genetic algorithms to identify the disease spot area on the soybean leaves, accuracy of this algorithm was over 90%. Based on the characteristics of the color texture image of plant disease, Tian et al.^[12] proposed a method of applying the Support Vector Machine and Chromaticity moment analysis to identify grape disease and corn disease, and the results showed that the characteristics of the color texture image of disease were simple, fast and good classification by using

Received date: 2019-12-03 **Accepted date:** 2019-12-15

Biographies: Pei Wang, PhD, research interests: precision agriculture aviation technology and equipment, Email: wangpei@scau.edu.cn; Guohao Jia, Master, research interests: image processing, Email: wsjgha@163.com;

* **Corresponding author:** Zhiyan Zhou, PhD, Professor, research interests: precision agricultural aviation application, Email: zyzhou@scau.edu.cn.

Chromaticity moment to extract the color texture image of disease. The support vector machine classification method has better classification ability and generalization ability when there are fewer training samples for disease classification. Wang et al.^[13] first extracted the texture, color, shape and other characteristic vectors of the disease image, and then used genetic algorithm to optimize and select four independent, stable, strong classification capacity feature vectors, such as H-value, color moment, disease spot area, shape factors, etc to deal with the disease of maize leaves. The results showed that the accuracy of the recognition rate reached more than 90%. Zhao et al.^[14] first used color features to extract the non-green plant species, such as disease spots and soil, and then used area threshold segmentation method to exclude some of the soil and other non-green plant connectivity areas, and finally used chain code to calculate the shape characteristics of disease spots and soil and other non-green plant connecting areas, and separated disease spots based on the width, rectangularity and roundness of the area and effectively extracted the red rot disease and ringspot disease of sugarcane, the correct rate of image segmentation to ringspot disease reached 93%, and red rot disease reached 95%. Based on the image processing technique and the Bayes discrimination method, Guan et al.^[15] studied the recognition method of three kinds of rice diseases (blast, sheath blight, and bacterial leaf blight), the results showed that the number of texture parameters decreased to 35.2% and the highest recognition rates of four parameter sets was 97.2%, this method could be applied to recognize diseases of other crops. Sun et al.^[16] explored the detection of rice leaf roller pest by spectral technique, the analysis results showed that the canopy reflectance of severely damaged rice was higher than that of the control area, and in the near-infrared region, the red edge inflection position moved to direction of blue light with the affection severity increasing. The NIR-NDVI could be used to identify the damaged region with contrast region efficiently and the accurate rate of 25 verification samples selected randomly reached 70%. To use visible images to detect and estimate Rice Plant-hoppers (RPH) infestation areas on rice stems, Zhou et al.^[17] used a handheld camera to collect visible images of the rice stems and extracted the related image corner eigenvalues using a Smallest Univalued Segment Assimilating Nucleus (SUSAN) algorithm, and developed a linear regression model based on the corner eigenvalues, the results showed that the SUSAN corner detection algorithm could also be used to distinguish the I (infestation) and N (non-infestation) areas with high accuracy, the correlation coefficient between the number of RPHs and corners could reach 0.8277 and proposed a visible-image-based method for quantity estimation of RPH infestation.

More et al.^[18] used support vector machine technology based on a web search tool to identify pomegranate diseases and pests, with an average accuracy of 84%. Recently, deep learning technology has developed rapidly. Many researchers have begun to apply deep learning technology in agriculture. Fuentes et al.^[19] proposed a real-time algorithm based on robust deep learning target detection, which can detect a variety of tomato diseases and pests. Ding et al.^[20] developed an algorithm based on convolutional neural network to automatically identify codling moth from trap images for pest management.

In order to solve the problems of uneven lighting and random noise under natural light conditions, and increase the correct rate of image segmentation to rice damaged-spots infested by the Rice Leaf Roller (*Cnaphalocrocis medinalis* Guenee), an image

segmentation algorithm based on OSTU binarization algorithm and improved Gustafson-Kessell (GK) Fuzzy Cluster for the rice damaged-spots infested by the Rice Leaf Roller will be proposed in this paper.

2 Materials and methods

2.1 Instrument and images acquisition method

The images acquisition instrument is a digital camera, which made in Samsung Ltd., the model is PL10. Damaged rice samples located in the campus farm of South China Agricultural University, Guangzhou City, P.R. China. The images acquired during late rice in the fall.

To reduce the influence of subjective factors, 91 of the 200 images collected were randomly selected as test images. In order to speed up image processing, the test image was converted to resolution as 320×240 by bilinear interpolation.

2.2 Image color space selection

The image of rice infested by Rice Leaf Roller under natural conditions mainly includes green and non-green, the green part mainly consists of normal leaves and weeds, and the non-green part is mainly the leaf after damaged by Rice Leaf Roller, which is yellow or white.

Usually, the color feature is selected on the image segmentation. Typically, the captured images are RGB color spaces, which are appropriately combined with different weighting factors for R, G, and B components to achieve better color characteristics. When the initial segmentation occurs, selecting the color feature 2G-R as the segmentation feature can better highlight the green part and eliminate the effect of the green part on pest damaged-spot segmentation.

RGB and HSI (Hue, Saturation, Intensity) are commonly used color spaces at present. Compare with RGB, components of HSI color space have better independence and color distribution^[21]. The S component represents the saturation of the image, proper processing of S component is a good way to address the effects of uneven image brightness.

The conversion formula for the S component from RGB is as follows:

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)] \quad (1)$$

The S component takes a value range of [0,1], adds S plus the value in (0,1), and subtracts 1 when $S > 1$, to lower the high-frequency portion of the image by moving the high grayscale range to the low grayscale range. By using this method, the effect of uneven lighting on the image can be greatly reduced.

2.3 OSTU binarization algorithm

Threshold segmentation is a common method of image segmentation, which is divided into two steps: one is to set the threshold, and the other is to compare the threshold of each pixel to classify the pixels.

The classical OSTU binarization algorithm is a very practical method in the threshold segmentation method, its basic principle is to automatically search for an optimal threshold by seeking the maximum Between-cluster Variance between target and background. Compared with other threshold segmentation methods, there are great advantages, the optimal threshold can be calculated automatically, no human intervention. It can also get better segmentation results for images with obvious gray differences.

It consists of the following steps:

L as the maximum grayscale level of image, and N as the total

number of pixels of image, n_i as the number of pixels while grayscale value is i , the probability of each grayscale value:

$$P(i) = \frac{n_i}{N} \tag{2}$$

the image is divided into two categories with an integer t ,

$$C_0 = \{0, 1, \dots, t\}$$

$$C_1 = \{t+1, t+2, \dots, L-1\}$$

Then the probability of C_0 is

$$\omega_0 = \sum_{i=0}^t P(i) \tag{3}$$

the probability of C_1 is

$$w_1 = \sum_{i=t+1}^{L-1} P(i) = 1 - w_0 \tag{4}$$

the grayscale mean of C_0 is

$$m_0 = \sum_{i=0}^t \frac{iP(i)}{w_0} \tag{5}$$

the grayscale mean of C_1 is

$$m_1 = \sum_{i=t+1}^{L-1} \frac{iP(i)}{w_1} \tag{6}$$

the grayscale average of the entire image is

$$m = w_0 m_0 + w_1 m_1 \tag{7}$$

Thus, maximum between-cluster variance is:

$$s_B^2 = w_0(m_0 - m)^2 + w_1(m_1 - m)^2 = w_0 w_1 (m_0 - m_1)^2 \tag{8}$$

When between-cluster variance reach to maximum by using ergodic method, the t is considered as the best threshold T for image segmentation.

2.4 Gustafson-Kessel Fuzzy Clustering Algorithm

The GK algorithm^[22] is a powerful clustering technique that is widely used in areas such as image processing, classification and system recognition. It uses membership to determine the extent to which each pixel belongs to a category^[23].

For the image pixel sample set, $X=(x_1, x_2, \dots, x_N)$, the GK algorithm uses iterative optimization target function to obtain the target classification, equivalent to seek peak value of objective function, that is, the minimum value of the objective function:

$$J(U, V, \{A_i\}) = \sum_{i=1}^K \sum_{k=1}^N (\mu_{ik})^m (x_k - v_i)^T A_i (x_k - v_i) \tag{9}$$

where, $A_i = [\rho_i \det(F_i)]^{1/n} F_i^{-1}$, $\rho_i = |A_i|$; μ_{ik} is the membership of the pixel x_k that belongs to i category; v_i is the center of category i ; m represents the constant of weight, usually $m=2$; k is the number of clusters; N is the total number of samples, and at the same time, For F_i , there are:

$$F_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (x_k - v_i)(x_k - v_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \tag{10}$$

2.5 Improved Gustafson-Kessel Fuzzy Clustering Algorithm

The flow diagram of improved Gustafson-Kessel fuzzy clustering algorithm was shown in Figure 1. Firstly, the super-green equation and OSTU binarization algorithm was used to remove the green part when images inputted, and then created color-coded images, converted the color space to HSI, selected S component for segmentation feature extraction and classification. GK Fuzzy Clustering Algorithm was used for damaged-spots segmentation after clustering. What follows was process the segmentation result by Morphological filter and then extracted the target damaged-spots in the image.

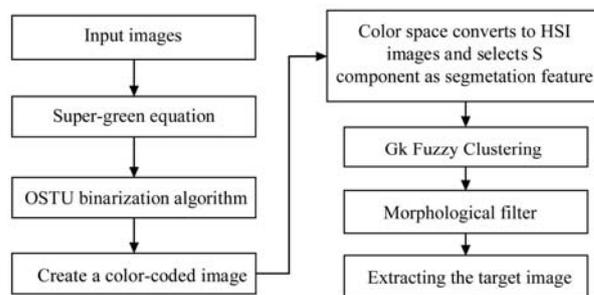


Figure 1 Flow diagram of improved Gustafson-Kessel fuzzy clustering algorithm

3 Results and discussion

3.1 Optimization for the number of clusters

As shown in formula 9, k is the number of clusters, the selection of clustering number k has an important impact on the segmentation results. Unsuitable m value will seriously affect the segmentation performance of the algorithm. In order to select the best k value of clustering number, 50 images were segmented. According to the comparison of the experimental results, it was found that the segmentation effect was better when clustering number $k \geq 4$, and iteration stop condition $\epsilon=10^{-3}$. As shown in Figure 2, the results were $k=2$, $k=3$ and $k=4$ respectively. Figure 3 showed the time required for taking different k values, in which, the time used was shorter and the segmentation effect was better when $k=4$.

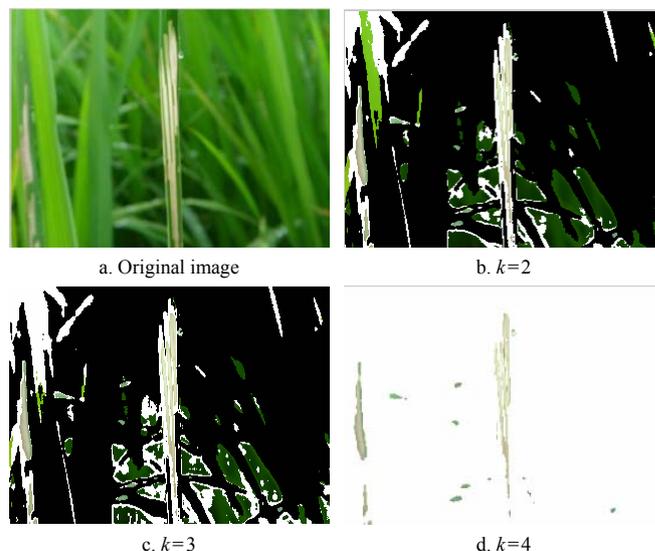


Figure 2 Segmentation performance of different k values

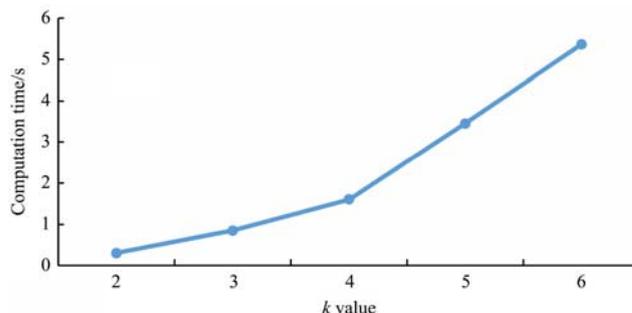


Figure 3 Computation time of different k values

3.2 Comparison of image segmentation results of different algorithm

To distinguish the more obvious color characteristics in the image, first remove the green information, transform the image by

using ultra-green equation, and then segment by using OSTU binarization algorithm, as shown in Figures 4b and Figure 4c. Mark the original image to get the damaged-spot in the image, as shown in Figure 4d. Enhance the S component of the image to improve clarity and reduce the effects of uneven brightness, as shown in Figure 4e and 4f. Finally, segment the image with the proposed GK fuzzy clustering algorithm, as shown in Figure 4g. Segmentation results were processed by morphological filtering. In the end, the rice damaged-spots infested by Rice Leaf Roller were accurately extracted from the original image, shown in Figure 4h.

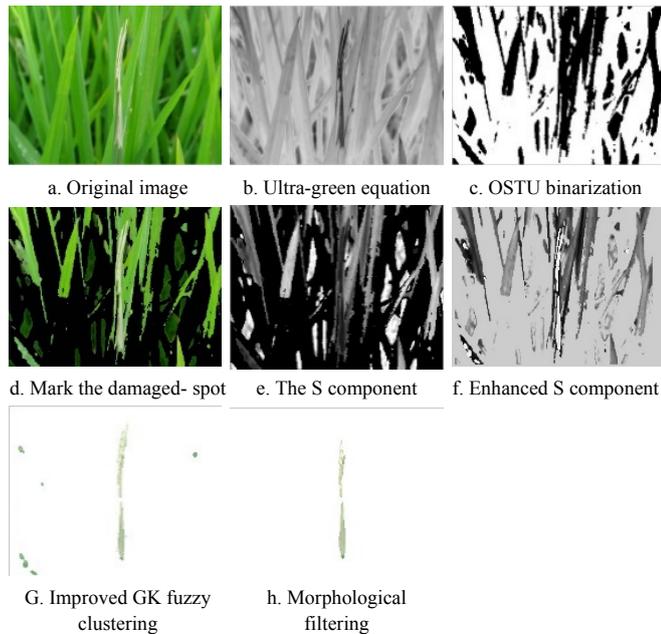


Figure 4 Comparison of image segmentation results of proposed algorithm in different step

In order to verify the accuracy of the proposed algorithm, 91 sample images were selected randomly, and segment by the proposed algorithm of this paper. The results showed that there were 75 images, which the damaged-spots were segmented completely. The accuracy rate of the proposed segmentation algorithm reached 82.4%.

The random sampling test images were also segmented by using the standard Fuzzy C-Means (FCM) algorithm and OSTU binarization algorithm, the results were shown in Figure 5. There was noise in the segmentation result when using FCM, which cannot give the outline of the damaged-spot. There was also random noise when using OSTU.

3.3 Feature extraction and classification

In classification, the information of segmented original image was too large and disorderly, which seriously affected the speed and efficiency of classification. Compared with the original image, the information of the feature component is being greatly reduced, which is more conducive to classification.

Features are the characteristics of one kind of target different from other targets. Good features can eliminate the interference of unimportant variables, and make image classification achieve high accuracy. In order to test the effect of segmentation results in classification and recognition between infested leaves and healthy leaves, three features, skewness of color feature B and R component, average of S component, were selected. The distinguish effect between infested leaves and healthy leaves of each features mentioned above were shown in Figure 6. The

classification accuracy rate based on the above three features reached 94%.

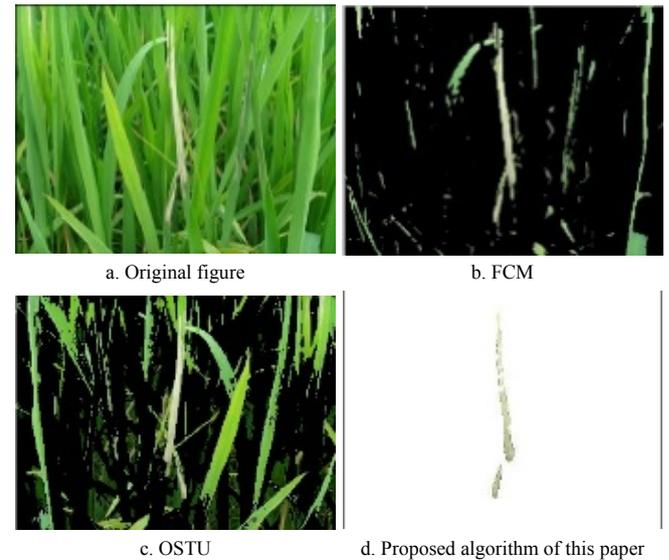


Figure 5 Comparison of image segmentation results of different algorithm

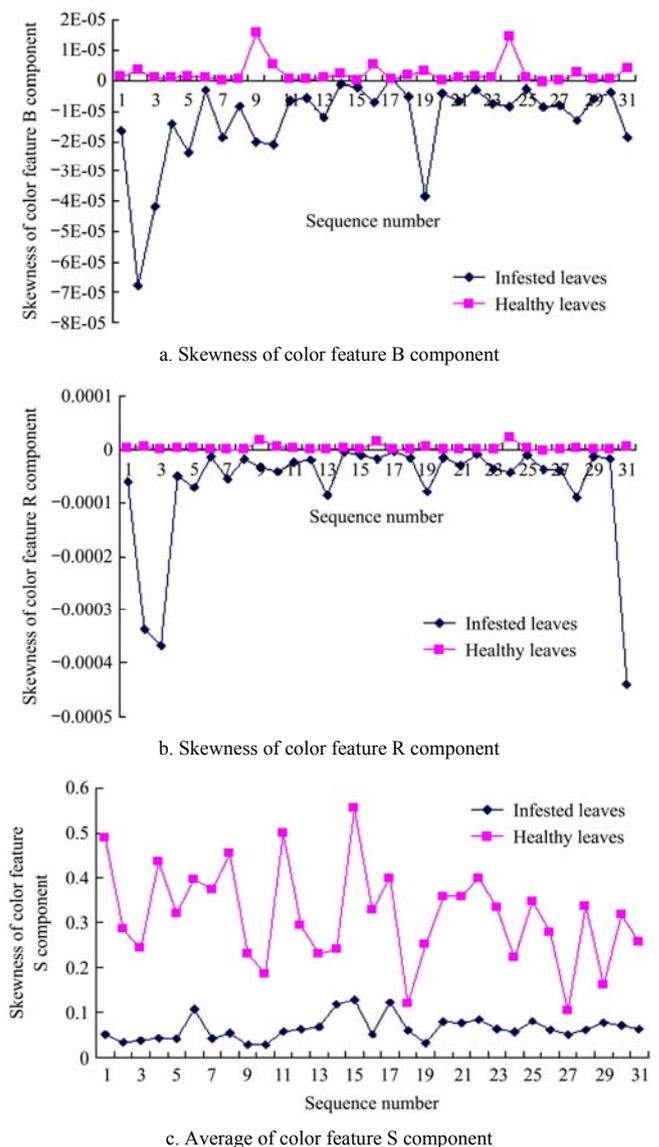


Figure 6 Distinguish effect between infested leaves and healthy leaves based on given features

4 Conclusions

(1) Under the condition of natural light, due to the influence of uneven lighting and random noise, the traditional method of image segmentation is difficult to achieve the ideal result. The feasibility of proposed image segmentation algorithm, which based on OSTU binarization algorithm and improved GK Fuzzy Cluster for the rice damaged-spots infested by the Rice Leaf Roller, was verified. Experimental results showed that the accuracy rate of the proposed segmentation algorithm reached 82.4%.

(2) The skewness of color feature B and R component, average of S component were extracted as features of segmentation images, the classification accuracy rate based on the above three features reached 94%. Efficient results were achieved by using the mentioned above method for images with the influence of uneven lighting, random noise and complex background under natural conditions.

(3) The proposed algorithm in this paper has achieved good results in the segmentation of damaged-spots infested by the Rice Leaf Roller. In the future, the next research will focus on how to carry out batch classification in-situ, and give the results of damage level evaluation.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (31871520 and 31371539), the National Key R&D Program of China (2018YFD0200301), Natural Science Foundation of Jiangsu Province of China (BK20160510), Science and Technology Plan of Guangdong Province of China (2017B090903007), Science and Technology Plan of Guangzhou City of China (201807010111) and Innovative Research Team of Agricultural and Rural Big Data in Guangdong Province of China (2019KJ138). We also thank the anonymous reviewers for their critical comments and suggestions to improve the manuscript.

[References]

- [1] Qiu B J, Li H F, Wu C W, et al. On variable-rate spraying equipment and its key technology. *Journal of Jiangsu University*, 2004, 25(2): 97–101. doi: 10.3969/j.issn.1671-7775.2004.02.002. (in Chinese)
- [2] Huang Y X, Yang Q. Impact of precision agriculture on environment. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, 10(25): 250-254. doi: 10.3969/j.issn.1002-6819.2009.z2.048. (in Chinese)
- [3] Zhao G Q, Wang J, Mu X T. Research Advancement of Precision Agriculture. *Meteorological and Environmental Sciences*, 2007, 30(1): 84-88. doi: 10.3969/j.issn.1673-7148.2007.01.018. (in Chinese)
- [4] Lan Y B, Chen S D, Fritz B K. Current status and future trends of precision agricultural aviation technologies. *International Journal of Agricultural & Biological Engineering*, 2017, 10(3): 1–17. doi: 10.3965/j.ijabe.20171003.3088.
- [5] Li B, Liu Z Y, Huang J F, Zhang L L, et al. Hyperspectral identification of rice diseases and pests based on principal component analysis and probabilistic neural network. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, 9(9): 143–147. doi: 10.3969/j.issn.1002-6819.2009.09.026. (in Chinese)
- [6] Ma X D, Guan H O, Huang Y. Investigation on the Infectious Degree of Maize Leaf Based on Image Processing. *Journal of Agricultural Mechanization Research*, 2009, 11(11): 102–103. doi: 10.13427/j.cnki.njyi.2009.11.032. (in Chinese)
- [7] Zhou Z Y, Luo X W, Zhang Y, et al. Machine-based technologies for detecting and monitoring insect pests of crops: A review. *Acta Entomologica Sinica*, 2010, 53(1): 98–109. doi: 10.16380/j.kcxb.2010.01.006. (in Chinese)
- [8] Luo X W, Zang Y, Zhou Z Y. Research progress in farming information acquisition technique for precision agriculture. *Transactions of the CSAE*, 2006, 22(1): 167–173. doi: 10.3321/j.issn:1002-6819.2006.01.037. (in Chinese)
- [9] Chen J J, Ji S W, Li J, et al. Automatic measurement of danger degree of cotton insect pests using computer vision. *Transactions of the CSAE*, 2001, 17(2): 157–160. doi: 10.3321/j.issn:1002-6819.2001.02.040. (in Chinese)
- [10] Ma D G, Shao L S, Ge J, et al. Detection of the harm degree of rice blast and rice sheath blight. *Chinese Agricultural Science Bulletin*, 2008-9: 485–489. (in Chinese)
- [11] Qi G Y, Ma X D, Guan H O. Extraction of the image of soybean target leaf spot based on improved genetic algorithm. *Transactions of the CSAE*, 2009, 25(5): 142–145. doi: 10.3969/j.issn.1002-6819.2009.05.027. (in Chinese)
- [12] Tian Y W, Li C H. Research on recognition of cucumber disease based on image processing in sunlight greenhouse. *Journal of Agricultural Mechanization Research*, 2006, 0(2): 151–153. doi: 10.3969/j.issn.1003-188X.2006.02.055. (in Chinese)
- [13] Wang N, Wang K R, Xie R Z, et al. Maize leaf disease identification based on fisher discrimination analysis. *Scientia Agricultura Sinica* 2009, 42(11): 3836–3842. doi: 10.3864/j.issn.0578-1752.2009.11.010. (in Chinese)
- [14] Zhao J H, Luo X W, Zhou Z Y. Image Segmentation Method for Sugarcane Diseases Based on Color and Shape Features. *Transactions of the Chinese Society for Agricultural Machinery*, 2009, 39(9): 100-103. (in Chinese)
- [15] Guan Z X, Tang J, Yang B J, et al. Study on recognition method of rice disease based on image. *Chinese Journal of Rice Science*, 2010, 24(5): 497-502. doi: 10.3969/j.issn.1001-7216.2010.05.009. (in Chinese)
- [16] Sun H, Li M Z, Zhou Z Y, et al. Monitoring of cnaphalocrocis medinalis guenee based on canopy reflectance. *Spectroscopy and Spectral Analysis*. 2010, 30(4): 1080–1083. doi: 10.3964/j.issn.1000-0593(2010)04-1080-04. (in Chinese)
- [17] Zhou Z Y, Zang Y, Yan M L, et al. Quantity estimation modeling of the Rice Plant-hopper infestation area on rice stems based on a 2-Dimensional Wavelet Packet Transform and corner detection algorithm. *Computers and electronics in agriculture*, 2013, 101: 102–109. doi: 10.1016/j.compag.2013.12.013.
- [18] More S, Nighot M. AgroSearch: A Web Based Search Tool for Pomegranate Diseases and Pests Detection Using Image Processing. *International Conference on Information & Communication Technology for Competitive Strategies*, 2016, pp 1–6. doi: 10.1145/2905055.2905102
- [19] Fuentes A, Yoon S, Kim S C, et al. A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors*, 2017, 17(9): 2022–2043. doi: 10.3390/s17092022
- [20] Ding W G, Taylor G. Automatic moth detection from trap images for pest management. *Computers & Electronics in Agriculture*, 2016, 123(C), 17–28. doi: 10.1016/j.compag.2016.02.003
- [21] Bu J, Wang X Y, Sun Y F. A FCM Based Image Segmentation Algorithm using Multi-color Components. *Journal of Image and Graphics*, 2008, 13(10): 1837–1840. doi: 10.11834/jig.20081005. (in Chinese)
- [22] Babu'ska R, Van der Veen P J, Kaymak U. Improved Covariance Estimationfor Gustafson-Kessel Clustering. *Proceedings of the 2002 IEEE International Conference*, 2002. doi: 10.1109/FUZZ.2002.1006654.
- [23] Peng D Q, Li J Q, Lin Y Q. FCM Image Segmentation Based on the Spatial Restrained Fuzzy Membership. *Computer Science*, 2010, 37(10): 257–259. doi: 10.3969/j.issn.1002-137X.2010.10.062. (in Chinese)