Identification Method of Sunflower Leaf Disease Based on SIFT Point

Jun Liu, Fang Lv, and Bobo Liu
Inner Mongolia University of Technology, Hohhot, China
Email: passion_jun@sina.com, lvfang@tsinghua.org.cn, 1398155847@qq.com

Abstract—**China is the largest country in the world for planting sunflower, and sunflower is an important cash crop in Chinese agriculture. Sunflower diseases are becoming more and more serious, accurate identification and effective control of sunflower diseases are very important for local economic development in China. Traditionally, sunflower disease identification is mainly based on the eye recognition method, which has great limitations and this way is difficult to meet the development needs of modern agriculture. In recent years, with the rapid development of digital image processing technology and computer technology, it provides a new platform for development of sunflower disease identification system. This paper proposes a sunflower leaf disease identification method based on SIFT points mainly for sunflower leaf diseases. In-depth study was conducted on sunflower black spot, bacterial leaf spot, powdery mildew and downy mildew.**

Index Terms—**digital image processing, SIFT, sunflower leaf disease, identification**

I. INTRODUCTION

At present, image processing algorithm is widely used to identify plant leaf lesions. In 2014, Xu Hui and others used support vector machine to identify sunflower diseases [1]. In 2015, Di Penghui et al realized the classification and recognition of sunflower rust by extracting the color and texture features of the spots [2]. However there are deficiencies in the recognition of images by the support vector machine algorithm. This method is difficult to implement for large-scale training samples and solve multiple classification problems [3], color characteristics of lesions are vulnerable to environmental interference. SIFT algorithm is a feature detection method based on scale and angle invariant technology. It has invariance to image scaling, rotation and even radiation transform, and has higher robustness and higher computational stability [4]. Therefore, this paper selected SIFT algorithm to identify the image of sunflowers black spot, bacterial leaf spot, powdery mildew and downy mildew, which can fill the blank of domestic sunflower diseases identification method.

II. IMAGE ACQUISITION OF SUNFLOWER LEAF

About 60% of the external information that human beings obtain is from visual images, so how to obtain and process visual information is particularly important. With the rapid development of mobile communication technologies and multimedia technologies, high-resolution and high-definition camera phones have become increasingly popular in people's lives. In this paper, we use the mobile phone as the acquisition tool to take the image of the disease of sunflower leaves in the field environment.

III. SUNFLOWER LEAF DISEASE IMAGE PREPROCESSING

Since the sunflower leaf disease images acquired by mobile phones under natural lighting are inevitably affected by various factors, resulting in poor image quality and blurring. Now image preprocessing of the collected images. The aim is to improve the image quality and highlight the image features, and obtain the ideal image of easy disease identification.

The histogram equalization method has a good defogging effect, which can distribute the image histogram in the whole grayscale range to make the brightness region average distribution [5], [6]. The method of homomorphic filtering can enhance the edge of image and protect the detail, and remove the noise in the signal [7]. Therefore, this paper chooses histogram equalization and homomorphic filtering to preprocess the recognition image. Using these two methods to deal with sunflower leaf disease images (all of which uses black spot image as an example) are shown in Fig. 1-Fig. 4.

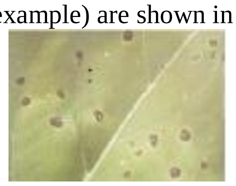

Figure 1. Foggy black spot image

Figure 2. After the histogram equalization processing

Figure 3. Black spot pathogen image

Manuscript received January 16, 2019; revised May 14, 2019.

doi: 10.18178/joig.7.2.64-67

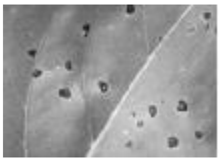

Figure 4. After homomorphic filtering

IV. Feature Extraction of Image Features of Sunflower Leaf

At present, SIFT algorithm is adopted to extract feature points of preprocessed images. The core idea of SIFT algorithm is to establish a scale space model of images and extract the key points of images. Then determine the scale, position and direction of the key points, and describe the key points to form the SIFT feature vector. Finally, matching recognition of image feature points [8].

The implementation of SIFT algorithm needs to complete two stages: The first stage is to extract the feature vector from the image to be matched, which is divided into four steps: scale space extremum extraction, accurate positioning of key points, determination of feature direction and generation of SIFT feature vector. The second stage is to complete the matching of image feature vectors [9]. The flow chart for extracting SIFT feature points and their feature vectors is shown in Fig. 5.

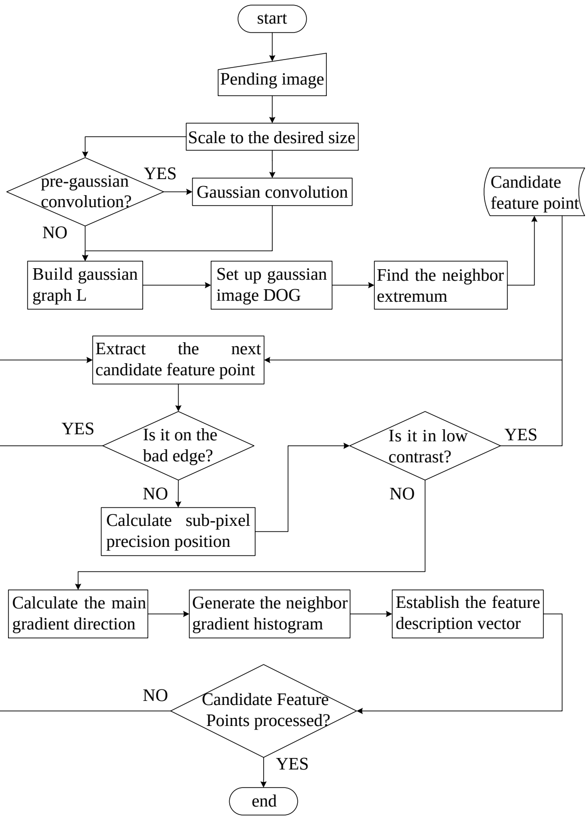

Figure 5. Flow chart for extracting SIFT feature points and their feature vectors

A. Scale Space Extreme Value Extraction

The scale space theory is the basis of detection invariance, and the Gaussian differential scale space (DoG) concept is introduced in order to obtain stable key points [10]. The local extreme value detection of DoG space needs to detect the pixel points, including the DoG values of the adjacent 8 pixels in the same layer and each of the 9 adjacent pixels in the upper and lower layers in 26 pixels. In different scales, the black spot image is shown in Fig. 6.

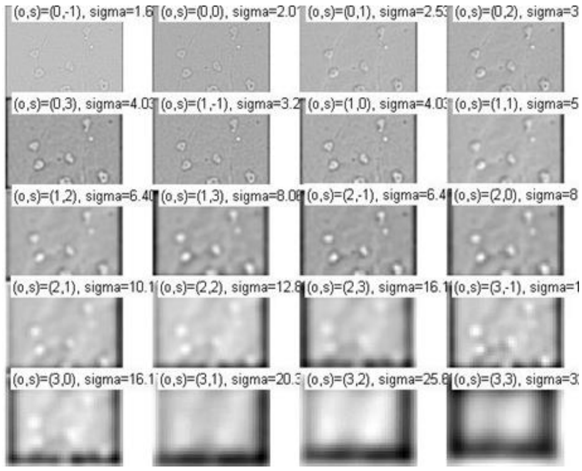

Figure 6. The black spot image at different scales

B. Accurate Positioning of Key Points

After obtaining the discrete spatial extremum point by (1), what needs to be done now is the precise positioning of the key point including its scale and position. Because the obtained discrete spatial extremum is not a real extremum, it is necessary to use the sub-pixel interpolation method to change the position of the current key point to repeatedly interpolate until it converges. At the same time, the low contrast key points and the unstable edge key points need to be removed.

C. Key Point Direction Determination

You need to specify a direction for each keypoint in order to make the feature descriptor spin-invariant. Calculate the mode and direction of the gradient in the window of the Gauss pyramid image field in which the detected key point is located, see equation (1), (2):

$$m(x,y)=\sqrt{[L(x+1,y)-L(x-1,y)]^2+[L(x,y+1)-L(x,y-1)]^2} \quad (1)$$

$$\theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \quad (2)$$

After the key point gradient is known, the histogram is used to count the pixel gradients and directions in the field. The gradient histogram divides the direction of 0~360 degrees into 36 columns, each of which is 10 degrees. The peak value of the histogram represents the direction of the neighborhood gradient in the feature point, and the maximum value in histogram is the main direction of the key point. Each key point containing the scale, position, and direction detected from the image is the SIFT feature point of the image. The extracted SIFT feature points are shown in Fig. 7, and the number of feature points is shown in Fig. 8.

D. Generate SIFT Feature Vectors.

Through the above steps, feature points have been generated. Now we need to set a descriptor for each feature point and describe it with a set of vectors to ensure rotation invariance. The description of the feature points is related to the scale of the feature points. The SIFT feature vector is generated by calculating the gradient histogram by partitioning the area around the feature points. The corresponding Gaussian image of the feature point is divided into 8 x 8 regions. Then, take the subregion of 4 by 4 as a seed point, and each point has 8 direction gradient information, and calculate its weight. Thus each feature points can produce a total of 128 gradient information, which generate 128 dimensional feature vector, then 128 dimensional vector units get SIFT descriptor, this way to enhance the stability of image matching. The feature points extracted from the black spot image and the key description of 128 dimensions are presented, as shown in Fig. 9.

Figure 7. Extract of black spot image SIFT points

```
computing octave 3
initial keypoints # 0.
                Time (0.000 s)
total keypoints: 43
```

Figure 8. Number of SIFT points extracted from black spot image

Variable Editor - descr1

descr1 <43x128 double>

		2	3	4	5	6	7	8	9	10
1	30	0.2690	0.1146	0.0254	0.0487	0.2690	0.1806	0.0249	0.0414	0.1167
2	33	0.0172	0.0120	0.0231	0.0255	0.1631	0.0856	0.0746	0.0430	0.2630
3	81	0.0207	0.0630	0.0188	0.1252	0.2330	0.2330	0.0190	0.0647	0.1920
4	76	0.0449	0.1018	0.0302	0.2487	0.2487	0.2487	0.0481	0.2136	0.0965
5	19	0.0865	0.0520	0.0099	0.0205	0.2375	0.2250	0.1726	0.0098	0.0805
6	64	0.1697	0.0364	0.0187	0.0391	0.2655	0.2309	0.2403	0.0200	0.1526
7	60	0.0299	0.0647	0.0259	0.1053	0.1366	0.2461	0.0504	0.2461	0.2461
8	48	0.0074	0.0114	0.0099	0.1324	0.1010	0.2625	0.0646	0.2625	0.2625
9	89	0.0192	0.0280	0.0093	0.2434	0.0798	0.2149	0.0528	0.0679	0.0719
10	19	0.2549	0.2549	0.0134	0.0944	0.2549	0.2549	0.0637	0.0506	0.2432
11	13	0.2577	0.1779	0.0195	0.1100	0.2577	0.2577	0.0746	0.0499	0.2577
12	85	0.0822	0.1481	0.0174	0.2408	0.2568	0.2568	0.0372	0.0721	0.2515
13	32	0.0877	0.1218	0.0151	0.2580	0.2580	0.2580	0.0451	0.0884	0.2064

Figure 9. 128 key point descriptors generated by SIFT algorithm of black spot image

V. THE SIFT ALGORITHM IDENTIFIES THE IMAGE OF SUNFLOWER DISEASE

Identification of sunflower disease is the matching of image feature vector. Feature point matching is achieved by comparing the Euclidean distances of two images as matching similarity [11]. When the SIFT feature vectors of the two images are generated, then taking a SIFT key point in the unrecognized image and the two SIFT key points of the Euclidean nearest neighbor distance and the second nearest neighbor distance of the image in the sample library and finally determine whether the matching point is accepted by the ratio (closest distance to closest distance).

The threshold of ratio was 0.75 by conducting a lot of experiments. The smaller the Euclidean distance means the higher the similarity, so accept the matching point when the ratio is less than 0.75, otherwise it will be discarded. Through analysis of the data of the research object, it is shown that the matching degree is higher when the number of image matching points is greater than or equal to 5. Therefore, this article selects the number of matching points to be 5, and it can be determined that the image matching is successful. The SIFT points feature matching is performed between the image to be identified and the image in the sample library. The matched image is shown in Fig. 10, and the number of matching points of the two image features is shown in Fig. 11.

The Fig. 11 shows that match the image with the sunflower leaves get black spot image feature matching by eight feature matching points, matching points greater than 5, so think of unknown image for sunflower leaf black rot.

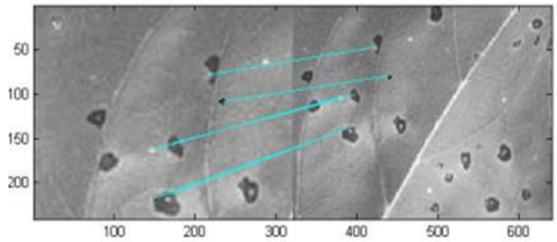

Figure 10. Feature Matching Diagram

```
Computing matches.
Found 8 matches.
Matched in 0.097 s
```

Figure 11. Number of feature matching points

VI. ANALYSIS OF RESULTS

This paper selects phone collected images: 120 samples of bacterial leaf spot, powdery mildew, black spot, and downy mildew, of which 60 are used as training samples and 60 are used as test samples. Four different diseases were identified by SIFT matching algorithm, and the recognition situation was shown in Fig. 12.

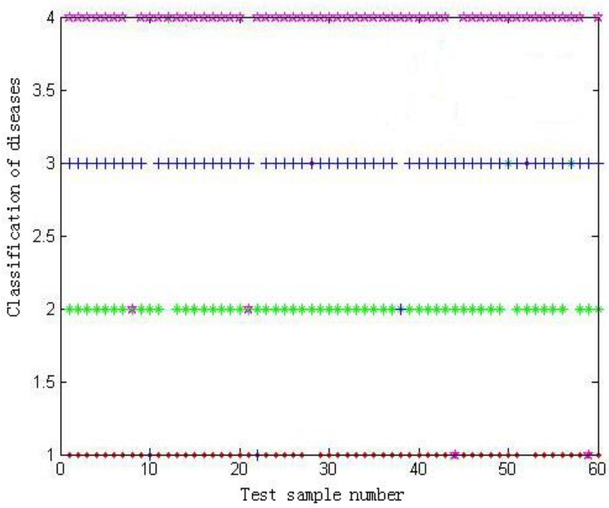

Figure 12. Four diseases identification in sunflower leaf

Among them, the vertical coordinate 1 indicates bacterial leaf spot disease, 2 represents powdery mildew, 3 represents black spot disease, 4 represents downy mildew disease. The x-coordinate represents the test sample number. The results of statistical analysis and analysis are shown in Table I.

TABLE I. IDENTIFYING DISEASES OF SUNFLOWER LEAF DISEASES BASED ON SIFT POINTS

Name of sunflower disease	Training sample size	Test sample size	Proper identification number	Correct recognition rate
Black spot disease	60	60	57	95.00%
Downy mildew	60	60	56	93.33%
Powdery mildew	60	60	57	95.00%
Bacterial leaf spot disease	60	60	58	96.67%

As can be seen from Table I, the recognition rate can reach 93.33% when using SIFT matching algorithm to identify four diseases of sunflower leaf. And the recognition effect is better.

VII. CONCLUSION

This paper mainly describes the use of SIFT algorithm to identify different diseases of sunflower leaves. Firstly, the collection of image of sunflower disease is described. Secondly, the image is preprocessed before image recognition. In this paper, histogram equalization is used to defog the image and the image is denoised by homomorphic filtering. Thirdly, the feature vector of image is extracted by SIFT algorithm. Finally, the characteristic vector was matched and the identification of the disease image was completed.

REFERENCES

[1] H. Xu, F. Lv, and X. M. Yue, "Discussion on sunflower leaf disease diagnosis based on imaging identification," in *Applied Mechanics and Materials*, Trans Tech Publications, 2014, vol. 577, pp. 1202-1206.

[2] P. Di, F. Lv, and X. Wang, "The research on the feature extraction of sunflower leaf rust characteristics based on color and texture feature," in *Proc. IEEE International Conference on Computational Intelligence and Communication Networks*, 2015.

[3] Y. Mu, *et al.*, "Hash-SVM: Scalable kernel machines for large-scale visual classification," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 979-986.

[4] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 17th IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, pp. 1150-1157.

[5] C. Y. Wong, *et al.*, "Histogram equalization and optimal profile compression based approach for colour image enhancement," *Journal of Visual Communication and Image Representation*, vol. 38, 2016.

[6] M. Abdullah-Al-Wadud, *et al.*, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, 2007.

[7] K. Delac, M. Grgic, and T. Kos, "Sub-image homomorphic filtering technique for improving facial identification under difficult illumination conditions," in *Proc. International Conference on Systems, Signals and Image Processing*, 2006, pp. 21-23.

[8] Y. Wang, *et al.*, "License plate recognition based on SIFT feature," *Optik-International Journal for Light and Electron Optics*, vol. 126, no. 21, pp. 2895-2901, 2015.

[9] S. Zhang and H. P. Zuo, "Research on SIFT algorithm of feature matching," *Modern Computer*, vol. 7, p. 12, 2010.

[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[11] D. Cheng, Y. Li, and R. Yu, "Image matching method based on improved SIFT algorithm," *Computer Simulation*, vol. 7, pp. 285-289, 2011.

Jun Liu, born in March 1993 in Luan City, Anhui Province, China, received his bachelor's degree from the College of Electronic and Information Engineering, Anhui Jianzhu University in 2016. Since 2017, he has been studying for a master's degree in Information Engineering College of Inner Mongolia University of Technology, and is currently preparing his master's degree thesis. His research interests include machine vision, machine learning, digital image processing, pattern recognition and other fields.

Fang Lv was born in Baotou City, Inner Mongolia Autonomous Region, China in March 1964. She received his Bachelor of Science degree and Master of Engineering degree from the Department of Physics, Inner Mongolia University and the Department of Automation, Tsinghua University in 1987 and 2003, respectively. She is currently a professor at the Information College of Inner Mongolia University of Technology. She teaches courses in electromagnetic fields and electromagnetic waves, microwave technology and digital image processing. Her research interests include machine learning, digital image processing, pattern recognition and other fields.

Bobo Liu, born in Shanxi Province in July 1990, received her bachelor's degree from Ren Ai College of Tianjin University in June 2015. In August 2018, she received her master's degree from the Information Engineering College of Inner Mongolia University of Technology, and is currently engaged in image processing related fields. Her research interests include machine learning, machine vision, digital image processing and other fields.