The SVM Classification Leafminer-infected Leaves Based on Fractal Dimension

Da-ke Wu

School of Engineering and Technology,
Southwest University,
Chongqing, P. R.China

Chun-yan Xie

School of Engineering and Technology,
Southwest University,
Chongqing, P. R.China

Cheng-wei Ma□

Key Laboratory of Agricultural
Bio-environmental Engineering, Ministry of Agriculture,
China Agricultural University,
Beijing, P. R. China.
□ Corresponding author

***Abstract*—Leafminer is one of pest of many vegetables, and the ordinary way to control it is spraying the pesticides. Because of polluting the environment and remaining in the vegetable, the pesticides are restricted to use. In order to get the information of the pest in the vegetable before the damage was not serious, this research used the image processing technology and the fractal dimension to work out the damaged degrees of the leafminer- infected cucumber leaves. Two different kernel functions have been used to set up the classifying models, and the fractal dimension of the damaged cucumber leaf images was applied to the Threshold methods and the SVM neural network performing the recognition and classification. The numerical experimental results showed that the classification precision of the Threshold methods is the best (61%) at the (3,7) computing step, and is the worst (36%) at the (3,6) computing step, and the error number of the Threshold methods is very high (the worst is 18 times at the (3,6) computing step). The precisions of the SVM are beyond 80% by using the polynomial-based kernel function and 90% via using the RBF kernel function, and the RBF kernel based SVM excels to the polynomial–based kernel. The SVM-based fractal dimension analysis models can be used to classify the leafminer-infected cucumber leaves.**

***Keywords*—fractal dimension, SVM, image processing, leafminer**

I. INTRODUCTION

Leafminer (Liriomyza sativae Blanchard) is a pest of many kinds of vegetable and ornamental plants, such as tomatoes, beans, peas, and various cucurbits. It had been spread out of the country since the first outbreak in China occurred in 1993 (Zhao and Kang, 2000). The damage may cover so much of the leaf that the plant is unable to function, and yields are noticeably decreased. In 1995, the leafminers caused a 30%～40% reduction of output in somewhere of China (Wei, 2000). In the past, many researches were found that the spectral range from 737 to 925 nm was useful for discriminating levels of infestation, especially at wavelengths 755 and 1400 nm, where the differences in reflectance between levels were the largest. Wu et al. (2006; 2007; 2007) found that the wavelengths (771, 821, 891, 945, 1026, 1121, 1256, 1674, 1687 and 1933 nm) related with the damaged degree of the leafminer-infected leaves, and Xu et al. (2007) found that the sensitive bands of 1450 and 1900nm modeled with severity level provided the highest correlation coefficient.

The fractal dimension (FD) provides a quantitative index of the roughness of natural surfaces, and many other natural phenomena. It has been applied to many different areas in science and engineering such as graphical simulation of clouds or trees (Mandelbrot, 1982), texture analysis of some images and material surfaces (Pentland, 1984; Gagepain et al., 1985), detection the brain MRI (Uemuraa et al., 2000), et al. Automatic ways, such as machine vision (Visser, 1994; Francl, 2000; Xiong, 2002; Qiu, 2003) and artificial neural networks (Zayas, 1998; Chen, 2001) have been used to detect and recognise the parameters of agricultural pests and diseases. Support vector machine (SVM) is a powerful method for pattern recognition (Cortes and Vapnik, 1995; Vapnik, 1998). Wu et al. (2006; 2007) used the SVM to identify the damaged degree of the leafminer-infected leaves based on the NIR data, and the recognition precisions were very high (Using all 10 input vectors, the precisions were beyond 90% with the polynomial-based kernel function and beyond 97.4% with the RBF kernel function.). The SVM has also been applied widely and can be performed quite well in many other pattern recognition problems (Osuna and Girosi, 1997; Maruyama and Nakano, 2000; Qiao et al., 2005).

In the past, chemical sprays have always been used against leafminers due to their protection within the plant. However, the layman often confuses leafminer damage with other leaf diseases and also does not know how bad the damage is. Thus, our objectives were to use leaf image data to determine the leaf damaged degree and the FD in diseased plants, and use the SVM approach to classify FD data of the cucumber leaves.

II. MATERIALS AND METHODS

2.1 The support vector machine (SVM)

The statistical learning theory was developed in the late 1960s (Vapnik, 1998). In recent years, it was used in the SVM for solving real-life problems and became more and more popular (Vapnik, 1999). The principle of the SVM is briefly described below.

Starting with a set of input-output training pairs:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

$$x \in R^d, y \in \{-1, +1\} \quad (1)$$

They can be separated by a hyperplane

$$w \cdot x + b = 0 \quad (2)$$

Let us describe the hyperplanes as

$$y_i [(w \cdot x_i) + b] - 1 \geq 0 \quad (3)$$

Among the hyperplanes, the one for that the distance to the closest point is maximal is the optimal separating hyperplane. The goal of the learning processing based on SVM is to find these optimal hyperplanes to separate the training data in n dimension by using criteria describe above, and then separate the real data in the same dimension.

The SVM decision function in terms of an appropriately defined kernel function can be obtained as

$$f(x) = sign(\sum_{i,j=1}^{n} \alpha_i y_i K(x_i, x_j) + b) \quad (4)$$

Where n is the sample size and $K(x_i, x_j)$ is the kernel function and b is bias. In our paper, we used two kernels used for constructing the classifiers, a polynomial-based kernel ($K(x_i, x) = [(x_i \cdot x) + 1]^P$), a radial basis function (RBF) kernel ($K(x_i, x) = \exp(-\frac{\|x - x_i\|^2}{2p^2})$).

The coefficients α_i are obtained by solving the following quadratic optimization problem:

$$w(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (5)$$

Subject to the constrains $\alpha_i \geq 0$, $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} \alpha_i y_i = 0$.

After the construction of an SVM, a training processing is determined, and the optimal hyperplanes are formed for classifying the real data.

2.2 Calculating the FD of the cucumber leaves

Fractal feature is based on the fractional Brownian motion (FBM) model, which is a technique that combines both fractal and multi resolution image decomposition. The FBM is a part of the set of 1 / f processes, which are the generalization of the ordinary Brownian motion X (0). The FBM is non-stationary, zero-mean Gaussian random functions, which are defined as:

$$X(0) = 0 \quad (6)$$

$$P(X(t + \Delta r) - X(t) \leq x) = (2\pi)^{\frac{1}{2}} (\Delta r)^{-H} \int_{-\infty}^{x} \exp\left[\frac{-u^2}{2(\Delta r)^{2H}}\right] du \quad (7)$$

Where $0 < H < 1$ is the Hurst coefficient that characterizes the FBM. t represent different observation times of the process X.

From the equation (7), the increment $X(t + \Delta r) - X(t)$ is steady, and its variance between the Δr^{2H} is in proportion, that is:

$$E[((X(t + \Delta r) - X(t))^2] = k \Delta r^{2H} \quad (8)$$

Where the H is the Hurst coefficient. Falconer (1990) had proved that the Hurst coefficient is 2 - H when the FBM is curve and 3 – H as the FBM is the curved surface. So, in order to compute the FD successfully involves estimation of H. Wu et al. (2000) used a novel computational modeling for estimating H. This paper also used this method to estimate the H value, and hence the FD,

2.3 Materials

The larvae of the leafminers were collected at the suburb, Beijing, China. And reared on potted cucumber seedlings and honeys placed in cages (50cm length×50cm width×50cm height) fitted with 0.4×0.4mm mesh size insect-proof screens, at T = 27 ± 1°C, RH = 70% ± 5%. Cucumbers were seeded in the 72-cell plug tray, and transplanted in plastic pots of 12 cm diameter and 10 cm height with one plant per pot, when the first leaf outspreaded. The pots were placed in a greenhouse (Greenhouse conditions were kept at: 25/20℃ (winter), 30/25℃ (summer); 70% RH; a photoperiod of 14 h light: 10 h dark.) to culture, watered and fertilized regularly to guarantee the plants to growth well. After the third leaf had outspreaded, the plants were placed in the screen cages containing a colony of adult leafminers which had hatched out already one day to infect leaves. 10 seedling pots were placed in the screen cage. After 12 hours, the pots were moved in the greenhouse to culture. 10 infected leaves were collected every time.

2.4 Computing the damaged degrees and the H values

The leaf image was taken by the Casio EX-Z3 camera in the natural light, and then processed to grey image, segmented and calculated the total leaf areas and the damaged leaf areas. At last, the damaged degree (DD) of the leave was worked out by using the followed equation:

DD= (The total leaf areas - The damaged leaf areas)/ The total leaf areas (8)

The software Matlab 6.5 was used to compute H values with self-written programs. Then the H values were classified by the SVM models.

a **b**

Figure 1. Cucumber leaf image infected by the leafminer

(a. Original image; b. Grey image)

2.5 Training and testing datasets

Total 95 cucumber leaves were used to analysis in the SVM models, and 31 used as training datasets and 64 used as testing samples (see Table 1). 21 computing steps and the kernels (the polynomial-based and the RBF kernels) were used in the SVM models. And 28 cucumber leaves (6 ordinary and 22 damaged) were used in the Threshold method classification models.

TABLE 1 SAMPLES FOR SVM CLASSIFICATION

The DD datasets %	Total training number	Total sample number
0	6	13
0-10	6	12
10-30	7	15
30-50	7	14
50-100	5	10
Total	31	64

III. RESULTS AND DISCUSSIONS

3.1 H values and the FD of the leafminer -infected cucumber leaves

When the leaves were infected by the leafminers, the leaf surface was damaged because they ate leaf mesophyll tissues (the leaf lost its chlorophyll, water, and other materials). From the figure 2, it is found that the *H* values are different with different computing steps. At some computing steps, the *H* values of the leaves are significantly different, so we can use the *H* values to classify the damaged leaves.

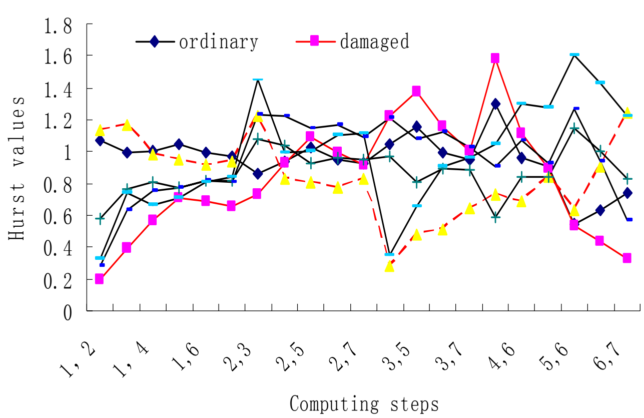

Fig.2 Hurst values of the cucumber leaves with different computing steps

It is showed that the FD of the leafminer-infected leaves is different from the figure 3. The FD between the ordinary and the damaged leaf images has many intersections, so it is hard to classify by using theTM, while the TM can classify the leaves at some computing steps.

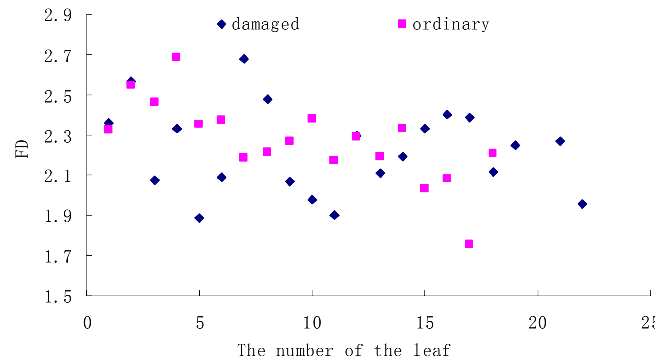

Fig.3 Fractal dimensions of cucumber leaves

3.2 Thresholds classification (TM) based on the FD

The TM is good at the linear problem. From the table 2, the classification precision of the TM is the best (61%) at the (3,7) computing steps, and is the worst (36%) at the (3,6) computing steps. In a word, the classification precision is very low. From the figure 4, the error number of the TM was very high, and the worst is 18 times at the (3,6) computing steps. So, at that step, the leaves can not be classified. Therefore, the nonlinear methods must be used to classify.

TABLE 2 CLASSIFICATION RESULTS OF THRESHOLD METHODS BASED ON THE CUCUMBER LEAF FRACTAL DIMENSIONS

Steps	Thresholds H values	Correct number	Error number	Classification Precision %
1,6	2.922	14	14	50
1,7	2.928	15	12	54
2,3	2.844	10	18	36
2,4	2.589	12	16	43
2,5	2.325	11	17	39
2,6	2.298	13	15	46
2,7	2.249	11	17	39
3,4	2.207	11	17	39
3,5	2.187	15	13	54
3,6	1.873	15	13	54
3,7	2.007	17	11	61

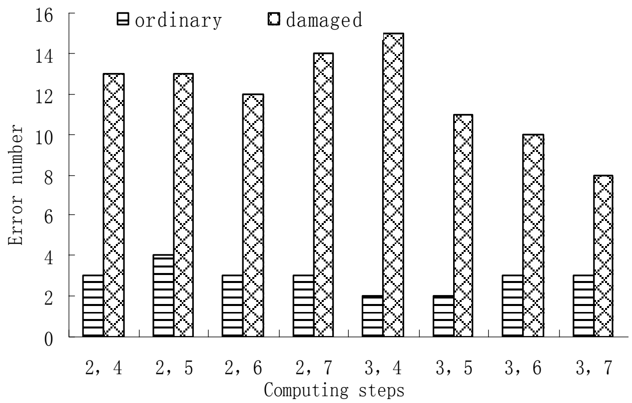

Fig.4 Distribution of error samples using the threshold classifier

3.3 SVM classification based on the FD

The SVM is a powerful method for pattern recognition and is good at the nonlinear problems. The SVM neural networks of polynomial-based and radial basis kernels have been applied as the recognising and classifying system. Note that the SVM network has only one output neuron, enabling the two-class recognition. To solve the five-class problem we have applied its extension in the form of "one-against-one" approach. For the five-class recognition problem we had to train ten neural two-class recognising networks. Table 3 was showed the results of the performance of the classifier with different kernel functions. It was found that the recognition provisions were higher (the precisions were beyond 80% by using the polynomial-based kernel function and 90% via using the RBF kernel function, and the RBF kernel based SVM excels to the polynomial–based kernel.) than the TM's. The classification performances of the group 10-30% damaged degrees were not good as the other groups. It may be caused by the overlapping of the FD of the cucumber leaves, which was needed further study.

TABLE 3 CLASSIFICATION RESULTS OF SVMS BASED ON THE FD

The damaged degree %	The polynomial –based kernel function		The RBF kernel function	
	Classification Precision %	Error data	Classification Precision %	Error data
0	77	3	85	2
0-10	75	3	83	2
10-30	73	4	80	3
30-50	79	3	86	2
50-100	80	2	90	1

IV. CONCLUSIONS

The paper has presented the FD of the leafminer-infected leaves recognising applications based on the TM and the SVM.

The proposed method has been tested on the samples. Two different kernel functions have been considered to set up the classifying models, and the FD of the damaged cucumber leaf image was applied to the TM and the SVM neural network performing of the recognition and classification.The classification precision of the TM is the best (61%) at the (3,7) computing steps, and is the worst (36%) at the (3,6) computing steps. The error number of the TM was very high, and the worst is 18 times at the (3,6) computing steps. The precisions of the SVM were beyond 80% by using the polynomial-based kernel function and 90% with the RBF kernel function, and the RBF kernel based SVM excels to the polynomial–based kernel. When the input vectors were fewer, the recognition precisions were lower. From the experiment, we can use the SVM-based FD models to classify the cucumber leaves infected by the leafminers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the National Key Technology R and D Program (Grant No. 2006BAD28B07) and the Chongqing Natural Science Foundation of China (Grant No. CSTC, 2007BB1365).

REFERENCES

[1] D. W. Wei. "Actualities of the Liriomyza sativae Blanchard in China". Guang xi nong ye ke xue, 2000, 6: 320-324 (in Chinese)
[2] Y.X. Zhao and L. Kang. "Cold tolerance of the leafminer Liriomyza sativae (Dipt., Agromyzidae)". J. Appl. Ent. 2000, 124: 185-189
[3] D. K. Wu and C. W. Ma. The support vector machine (SVM) based near-infrared spectrum recognition of leaves infected by leafminers. In: Proceedings of 2006 International Conference on Innovative Computing, Information and Control. VOL 3: 448-451
[4] D. K. Wu, C. W. Ma, S. F. Du. Influences of different damaged degrees of leafminer-infected leaves on the near-infrared spectral reflectance. Transactions of the Chinese Society of Agricultural Engineering, 2007, 23 (2): 156~159 (in Chinese)
[5] D. K. Wu, C. W. Ma, S. F. Du. Classification Visible Spectra of Leafminer-Infected Leaves By Support Vector Machine. Transactions of the Chinese Society of Agricultural Machinery,2007, 38(10): 87~90 (in Chinese)
[6] B.B. Mandelbrot. The fractal geometry of nature. San Francisco, CA: Freeman, 1982.
[7] A. Pentland. Fractal-based description of natural scenes. IEEE Trans Pattern Anal Machine Intell 1984; PAMI-6: 666–74.
[8] J. J. Gagepain, J Groslambert, R Brendel. The fractal dimension of phase and frequency noise: another approach to oscillator characterization. Proceedings of the 39th

Annual Frequency Contr. Symposium Philadelphia, PA, May 1985, p. 31–9.
[9] K. Uemuraa, H. Toyamab, and S. Baba1, et al. Generation of fractal dimension images and its application to automatic edge detection in brain MRI. Computerized Medical Imaging and Graphics,2000,24:73-85
[10] U. Visser, U. Voges, U. Streit. "International of AI-, database- and telecommunication-techniques for the plant protection expert system PRO-PLANT". In: Proceedings of the International Conference of Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE), San Antonio, 1994
[11] L. J. Francl, C. Larson, and E. D. De Wolf. "Description and evaluation of the NDSC regional wheat disease forecasting system". In: Proceedings of 2000 National Fusarium Head Blight Forum, Erlanger, Kentucky, 2000.
[12] X. M. Xiong and Ch. Y. Ji. "Application of the Combination of Genetic Algorithm and Artificial Neural Network to Prediction of Cucumber Downy Mildew", Transactions of the Chinese Society of Agricultural Machinery, 2002, 33(4): 69-71(in Chinese)
[13] D. Y. Qiu, Ch. H. Zhang, H. T. Zhang, et al.. "Application of neural networks in the recognition of stored-grain pests". Transactions of the CSAE, 2003, 19 (1): 142-144 (in Chinese)
[14] I. Y. Zayas and P. W. Flinn. "Detection of insects in bulk wheat samples with machine vision". Trans of the ASAE,1998,(3):88-888
[15] J. J. Chen, Sh. W. Ji, J. Li, et al.. "Automatic measurement of danger degree of cotton insect pests using computer vision" Transactions of the CSAE, 2001, 17 (2): 157-160 (in Chinese)
[16] C. Cortes and V. Vapnik, "Support vector networks, Mach. Learn". 1995, 20 (3) : 273–297.
[17] V. N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[18] V. N. Vapnik. "An overview of statistical learning theory. IEEE Transactions on Neural Networks", 1999, 10 (5), 988– 999.
[19] K. Maruyama and Y. Nakano, "Recognition method for cursive Japanese wordwritten in Latin characters", In: Proceedings of the Seventh IWFHR, 2000, pp. 133–142.
[20] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection", In: Proceedings of the CVPR'97, 1997, pp. 130–136.
[21] Y. H. Qiao., J. L. Liu, et al. "SVM classification of human intergenic and gene sequences". Mathematical Biosciences, 2005, 195:168–178
[22] K. Falconer.Fractal Geometry:Mathematical Foundations and Applications[M].New York:Willey,1990
[23] GH Wu, YJ Zhang and XG Lin. Fractal-based auto correlation description and classification of natural texture images. Journal of Tsinghua University (Science and Technology), 2000, 40(3): 90~93 (in Chinese)