

Shape Feature Extraction of Wheat Leaf Disease Based on Invariant Moment Theory

Zhihua Diao*, Anping Zheng, and Yuanyuan Wu

Zhengzhou University of Light Industry, College of Electric and Information Engineering,
Dongfeng Road 5, 450002, Zhengzhou, China
{diaozhua, 13613824027}@163.com, yuanyuan539@yahoo.com.cn

Abstract. Shape feature extraction is a key research direction on wheat leaf disease recognition. In order to resolve the problem of translation, scaling and rotation transformation invariance on shape matching, the invariant moment theory was introduced to shape feature extraction and seven Hu invariant moment parameters were defined as shape features. Meanwhile the present algorithm was used and new parameters were defined for shape feature extraction research on wheat leaf disease image. The shape features suitable for two types of wheat leaf disease recognition were received and applied in wheat disease intelligent recognition system. The results show that the system recognition rate is relatively high, and can meet the practical application requirements.

Keywords: feature extraction, shape feature, invariant moment, wheat disease.

1 Introduction

With the development of technology and the improvement of computer image processing technology, image processing and pattern recognition technology can be used as an effective way for the diagnosis and identification of wheat diseases. Feature extraction is a key research topic in image processing and pattern recognition, and broadly refers to a transformation which transforms the samples in high dimensional space to low-dimensional space through the map or transformation to achieve the purpose of dimensionality reduction [1]. In wheat leaf disease aspects, shape feature is an important feature to describe the image. And as a local feature, it describes the geometric properties of the local area. As the target shape of images obtained from different vision may be involved in a large difference, translation, scale and rotation transformation invariance problems need to be solved for accurate shape matching.

* This work was financially supported by the National High Tech. R&D Program of China (2007AA10Z237), and Doctor Foundation Project of Zhengzhou University of Light Industry. Author: Diao Zhihua(1982-), male, lecturer, doctor of engineering, interesting researches are image processing and multi-class classification algorithm.

In recent years, many researchers have carried out the image shape feature extraction research and achieved certain results in agricultural field. Chen Jiajuan et al [2] extracted some shape features such as the length of the chain, chain code area and shape factor, tortuosity, irregularity and so on based on the holes and incomplete edges in cotton leaves to determine the extent of the victims of cotton pests. Xie Guojun et al [3] proposed a fruit shape feature extraction algorithm based on multi-dimensional vision system for fruit intelligent selection and grading. Cai Jianrong et al [4] obtained the best-fitting ellipse parameters to describe orange shape using genetic algorithm. Wang Xiaofeng et al [5] introduced a method to identify the leaf based on shape features of leaf images. Zhao Jinhui et al [6] extracted the lesion area width, squareness and roundness for the separation of disease lesion according to sugarcane seedling image characteristics of red rot and ring spot diseases. Yang Zengfu et al [7] detected the edge of tea buds according to the shape characteristics of tea buds, and which was used for tea buds identification. However, these studies did not consider the basic translation, scale and rotation transformation invariance problems, and had some flaws.

Invariant moment theory was introduced for shape feature extraction of wheat leaf disease image in this paper, and seven Hu Invariant moments were selected as shape parameters. Then shape features of wheat leaf disease image were extracted. A feature subset including twelve shape features was received, and was applied to wheat leaf disease identification system which acquires relatively high recognition efficiency.

2 Shape Feature Extraction

Shape feature is an important feature to describe the image. Firstly, shape often links to the target, which has a certain semantic meaning. So shape feature can be seen higher level feature than color and texture feature. In addition, the shape expression is much more complex than the expression of the color or texture in nature. Secondly, the description of the target shape is a very complex issue. In fact, we still have not found the exact mathematical definition of shapes, including geometric, statistical, or morphological definition, so that it can be consistent with people's feelings. People's feeling of shape is a general result between retina feelings and knowledge about the real world of people. In order to accurately match the shape, the problem of translation, scaling, rotation invariance needs to be solved. As the target shape of images obtained from different vision may be involved in a large difference, translation, scale and rotation transformation invariance problems need to be solved for accurate shape matching.

2.1 Invariant Moment

Invariant moment is a statistical image feature which meets the translation, rotation and scale invariance, and has been widely used in image recognition. There are many invariant moments Used to describe shape which are arithmetic invariant moment, Legendre moment and Zernike moments, etc. [8].

For the integral area of S , if we have a given two-dimensional continuous function $f(x, y)$, the $(p + q)$ order moment is defined as follows,

$$m_{pq} = \iint x^p y^q f(x, y) dx dy, \quad p, q = 0, 1, 2, \dots \quad (1)$$

If the digital image $f(x, y)$ is a piecewise continuous function, and has non-zero value in a limited part of the xy plane, then we can prove the existence of various order moments. And the moment sequence m_{pq} could be uniquely determined by the image $f(x, y)$. At the same time the moment sequence m_{pq} also uniquely determine the image $f(x, y)$. The central moments can be expressed as follows,

$$\mu_{pq} = \iint (x - x_c)^p (y - y_c)^q f(x, y) dx dy \quad (2)$$

For digital images, the above integral can be replaced by the following formula,

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y) \quad (3)$$

The Standardized central moment is $\eta_{p,q}$, which can be expressed mathematically as follows,

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma} \quad \text{where } \gamma = \frac{p+q+2}{2}, p+q = 2, 3, \dots \quad (4)$$

In 1962, Hu [9] proposed the invariant moment theory of image recognition, and established a statistical feature extraction method for image recognition which has been widely used. Li [10] provided a calculation of 52 invariant moments, and seven common invariant moments are presented. From the above equation (1), (2) we could get seven H_u moment invariants, which can be expressed as follows,

$$\phi_1 = \eta_{20} + \eta_{02} \quad (5)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (6)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (7)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (8)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned} \quad (9)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) . \quad (10)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ & + (\eta_{03} - \eta_{12})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] . \end{aligned} \quad (11)$$

Although the seven invariant moments can well describe shape character, but when the image database is large, only 7 scalar features are not enough for feature extraction. Because the information carried by their own is very limited, they usually need to be combined with other conventional and global shape features in order to well describe the actual shape properties of the object.

2.2 Definition of New Shape Feature

Seven H_u invariant moments were selected as shape features in this paper, while area, perimeter, density and other parameters were selected as initial features. And two parameters were defined as shape features in intelligent identification system, which are as follows,

- 1) Number of spots (NS), which mainly used to calculate the number of spots in a disease leaf.
- 2) Ratio of area and number of spots (RANS). This parameter is an area of measurement of single spot, which mainly used to distinguish between large and small lesion spots.

3 Result and Discussion

In this study two categories of wheat rust and wheat leaf blight are as the main object, and the features are calculated according to the above definition formula. The effect drawing of shape feature extraction is shown in Figure 1.

From experimental statistics we found that the above-defined parameters of moment invariants have a good translation, scale and rotation transformation invariance. While the defined three parameters of ratio of area and number of spots, density and number of spots can be identified as core features of identification of wheat rust and wheat leaf blight. When the value of ratio of area and number of spots is more than 1200, the wheat disease can be considered as the leaf blight disease. And when the parameter is less than 500, the disease can be considered as wheat rust disease. For the parameters of density and number of spots, they need to be used in combination with the parameter of ratio of area and number of spots. When the parameter of ratio of area and number of spots is used as identification basis, the recognition rate of wheat diseases is about 95.3%. When the three parameters of ratio of area and number of spots, density and number of spots are used as identification basis, the recognition rate of wheat diseases is about 97.1%. When invariant moments

are taken into consider, and support vector machine algorithm is used as the model of disease identification, the recognition rate of wheat diseases is up to 99.4%. Using the above-defined shape features as the input of wheat disease identification system, we could obtain a higher recognition efficiency which can fully meet the needs of practical application.

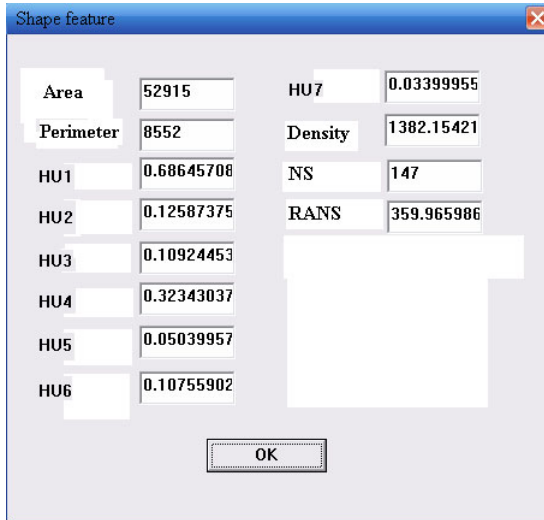


Fig. 1. Interface of shape feature extraction

4 Conclusion

Feature extraction process is a process to reduce the entropy of image information, and how to define and use features to express the image is the basis of image recognition [10]. In this study, combined with the current research of shape feature extraction, the invariant moment theory was introduced to shape feature extraction and seven Hu invariant moment parameters were defined as shape features. Meanwhile new parameters were defined for shape feature extraction research on wheat leaf disease image. The results showed that the system recognition rate is relatively high, and can meet the practical application requirements.

The shape features in this study is relatively easy to extract, and could reflect the morphological characteristics of two kinds of wheat disease images. But how to extract more features that can directly reflect the characteristics of different wheat disease images is the direction of future efforts.

Acknowledgments. This work was financially supported by the National High Tech. R&D Program of China (2007AA10Z237), and Doctor Foundation Project of Zhengzhou University of Light Industry.

References

1. Wang, J., Ci, L., Yao, K.Z.: A Survey of Feature Selection. *Computer Engineering and Science* 27, 68–71 (2005) (in Chinese)
2. Chen, J.J., Ji, S.W., Li, J., Zhao, X.D.: Automatic Measurement of Danger Degree of Cotton Insect Pests Using Computer Vision. *Transactions of the CEAE* 12, 157–160 (2001) (in Chinese)
3. Xie, G.J., Cao, Q.X., Liu, J.Z., Guo, F., Zhou, J.L.: Method for fruit shape feature acquisition based on multidirectional vision. *Transactions of the CSAE* 23, 127–132 (2007) (in Chinese)
4. Cai, J.R., Fan, J., Li, Y.L., Zhao, J.W., et al.: Shape feature extraction of on-tree citrus based on genetic algorithms. *Journal of JiangSu University(Natural Science Edition)* 28, 469–472 (2007) (in Chinese)
5. Wang, X.F., Huang, D.S., Du, J.X., Zhang, G.J.: Feature Extraction and Recognition for Leaf Images. *Computer Engineering and Applications* 3, 190–193 (2006) (in Chinese)
6. 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* 39, 100–103 (2008) (in Chinese)
7. Yang, F.Z., Yang, L.L., Tian, Y.N., Yang, Q.: Recognition of the Tea Sprout Based on Color and Shape Features. *Transactions of the Chinese Society for Agricultural Machinery* 40, 119–123 (2009) (in Chinese)
8. Ding, M.Y., Chang, J.L., Peng, J.X.: Research on moment invariants algorithm. *Journal of Data Acquisition and Processing* 7, 1–9 (1992) (in Chinese)
9. Hu, M.K.: Visual pattern recognition by moment invariant. *IEEE Trans. Information Theory* 8, 179–187 (1962)
10. Li, Y.J.: Reforming the theory of invariant moment for pattern recognition. *Pattern Recognition* 25, 723–730 (1992)
11. Yu, X.W., Shen, Z.R., Gao, L.W., Li, Z.H.: Feature Measuring and Extraction for Digital Image of Insects. *Journal of China Agricultural University* 8, 47–50 (2003) (in Chinese)