# Identifying of rice phosphorus stress based on machine vision technology

L. S. Chen[1, 2, a], S. J. Zhang[1, 2], K. Wang[1, 2], Z. Q. Shen[1, 2] J. S. Deng[1, 2]

[1]Institute of Agricultural Remote Sensing and Information Technology Application, Zhejiang University, Hangzhou, Zhejiang 310058, China
[2]Ministry of Education Key Laboratory of Environmental Remediation and Ecological Health, Zhejiang University, Hangzhou, Zhejiang 310058, China
cls512@zju.edu.cn

**Abstract:** At present, the identifying of rice nutrition stress by the chemical analysis is time-consuming and laborious. Machine vision technology can be used to non-destructively and rapidly identify rice nutrition status, but image acquisition via digital camera is vulnerable to external conditions, and the images are of poor quality. In this research static scanning technology was used to collect images of the rice's top-three leaves that were fully expand in 4 growth periods. From those images, 14 spectral and shape characteristic parameters were extracted by R,G,B mean value function and Regionprops function in MATLAB. The R, G, leaf area and Area/Perimeter ratio were chosen as the key characteristics for identifying phosphorus stress by One-Way ANOVA.The results showed that the overall recognition accuracy of phosphorus stress were 87.5 %, 100 %, 92 %, and 100% respectively. Based on the result, the methodology developed in the study is capable of identifying phosphorus stress accurately in the rice.
[L. S. Chen, S. J. Zhang, K. Wang, Z. Q. Shen, J. S. Deng. **Identifying of rice phosphorus stress based on machine vision technology.** *Life Sci J* 2013; 10(2): 2655-2663]. (ISSN: 1097-8135). http://www.lifesciencesite.com 368

**Keywords:** Rice; Scanning image; Computer vision; Phosphorus stress; Diagnosis; LOO-CV

## 1. Introduction

Phosphorus (P) is one of the indispensable minerals for plant growth. It is particularly important for promoting root development, flowering and ripening. However, most of the world's agricultural soils are deficient in P supply. For example, in China, 2/3 of the arable land suffers P deficiency. Using mineral P fertilizer can effectively overcome this problem. While the excessive use of P fertilizer may increase the risk of environmental pollution and increase production cost. Therefore, quick and accurate diagnosis of rice P nutrient status is necessary for guiding P fertilizer.

Rice with phosphorus deficiency usually has no significant symptoms before tillering stage. After the tillering stage its growth slows down significantly. And the color of new leaves is dark green, while the color of old rice leaves is grey purple, narrow and short, with small unit leaf area (Lu Jian-wei 2010).

In general, the diagnosis of phosphorus deficiency is based on chemical analysis in the laboratory (Zhao Qing-lei 2009, Li Li-mei 2007). But such methods are destructive, time-consuming, and difficult to apply in the field. However, there is a correlation between the chemical composition of leaf tissues and reflectance in the visible spectrum (Lu Jian-wei 2010), digital color image analysis might be capable of early diagnosis of phosphorus stress (L. Bacci 1998).

Previous studies had used digital camera (Jia Liang-liang 2009) and aerial Photography (Blackmer 1996) to directly acquire images for diagnosis of crop nutritient status. With simple and convenient operation the digital camera can obtain images, but the image acquisition process is affected by external conditions, and the acquired images usually have complex background, multi-redundant information and image noise, which may cause problem in image analysis. Therefore, this study used scanning as the acquisition method. Since scanning is in a closed environment, which can reduce the influence of external conditions, and ensures the reproducibility of the color and other characteristics. This method is seldom used for assessing crop nutrition.

At present, common methods for crop assessment include Bayesian classification, decision tree, neural network, etc. For example, recognition of rice disease based by Bayes method (Guan Ze-xin, 2010), study on discrimination of corn seed based on near-Infrared spectra and artificial neural network model (Chen Jian 2008), application of Clustering based decision Tree in the screening of maize germplasm (Wang Bin 2011) have been used. Those methods are suitable for the larger sample size, but have low recognition rate for the small sample size. Currently, on the basis of Vapnik-Chervonenkis dimension theory and structural risk minimization principle of statistics, an all-purpose machine learning algorithm is created by Vapnik: support vector machine (SVM) that is suitable for small sample classification (Deng Nai-yang 2004, Li Guo-zheng 2004). The advantage of SVM than earlier methods is that SVM is usually less vulnerable to over fitting problem (Wei Huang 2005), because SVM is designed to minimize structural risk, and previous techniques are usually based on minimizing empirical risk. Because of those advantages SVM has been used in studies such as land cover change detection (Nemmour, H. 2006), modeling spectral mixtures (Brown, L. 2000), Pattern recognition and data mining (Christopher and J.C. Burges 1998), soil type classification (Milos Kovacevica 2010), identification diagnosis system of rice insect pests (Shi Jing-jing 2009). However, less attention has been paid to the identification of rice nutrition stress.

The objective of this study was to identify the phosphorus stress of rice using symptomatic characteristics of leaf: R, G, leaf area and area/perimeter ratio (AP_ Ratio).

## 2. Materials and Methods

### 2.1 Experimental Design

The experiment was designed to study the rice of different P status in the greenhouse of ZiJinGang campus, Zhejiang University, Hangzhou, China in 2011. The rice variety used was ZheYou-NO.1, which was chosen as the experiment material that is susceptible to phosphorus deficiency. Pots containing clean, sieved, and thoroughly leached river sand allowing precise control of nutrient. The composition of the nutrient solution supplied was: $NH_4NO_3$ 23.57 mg/L, $K_2SO_4$ 33.6 mg/L, $CaCl_2$ 32.94 mg/L, $MgSO_4.7H_2O$ 54.32 mg/L, $Na_2SiO_3.9H_2O$ 16.5 mg/L. The experiment had five different P level treatments, 4 replications for each level, and 5 rice plants in each pot. Five P treatments (sodium dihydrogen phosphate 0 mg/L, 2.5 mg/L, 5 mg/L, 7.5 mg/L and 10 mg/L) via hydroponic solutions were applied to different pots. The nutrient solutions in pots were replaced every 14 days. Every 5 days the pH of the nutrient solution in each pot were measured and adjusted to 5.5–6.5 using 1mol/L NaOH.

### 2.2 Acquisition and analysis images of rice leaves

Leaf samples were taken on the August 4th, August 18th, August 27th, and September 8rd in 2011. The top-three leaves (Jiang Li-geng 2004, Zhu Jin-xia 2009) in 5 rice plants of 5 phosphorus levels were collected at each sampling time. After 4 dates were completed, 12 destroyed leaves were removed from 300 leaves. Thus a total of 288 rice leaves were processed with the following manner at the same day as they were collected, which were operated in the laboratory.

First, leaves were placed on a scanner (EPSON GT20000), scanner with a maximum scanning area of 11.7 × 17.0 inches, a R/G/B and BK color CCD line sensor, output image data was 16 bits per pixel per color internal and 1 to 8 bits per pixel per color extern. Resolution was set to 300dpi (dots per inch). Leaf area ($cm^2$) can be calculated by the sum of all pixels within the range of leaf multiply by $(2.54/300)^2$, length (cm) equals to the number of pixels in vein multiply by 2.54/300, and width (cm) equals to the number of pixels in the widest zone multiply by 2.54/300. Through R,G,B mean value function and Regionprops function in MATLAB (MathWorks Inc., USA), determining the leaf color characters (R,G,B) and shape characters (leaf length, leaf width, leaf area, leaf perimeter, AL_ratio, AP_ratio, eccentricity, rectangularity, area convexity, circularity, form Factor) in Table 1.

### 2.3 Research method

#### 2.3.1 The selection of optimal characteristic parameters

In this study, one-way ANOVA was used to find the optimal characteristic parameters with larger inter-group differences for modeling. One-way ANOVA is used in significance test of mean difference between two or more samples sets. A complex object usually has many factors of mutual restriction and dependence. The purpose of the Analysis of Variance is to find significant factors, the optimum level of significant factors, as well as the interactions between various significant factors by data analysis. This method divides the difference of the factor (W) between different groups into two parts, one is intra-group difference, originating from the dispersion of all samples within this group, and the other is inter-group difference, being caused by the distinction among the different groups. If the ratio (F) of intra-group and inter-group difference is near to 1, W has small effect for grouping, and compared to the randomness of data, the effect can be ignored. In other words, W is not a significant factor for grouping. Otherwise, with the bigger F, W has the more effect for grouping.

#### 2.3.2 Identification of rice phosphorus stress based on SVM

The SVM classification has two steps, training and classification. Regard the feature vectors reflecting the different categories as an input, and use an appropriate kernel function, and introduce a non-negative Relaxation term and penalty coefficient C to map the feature vectors reflecting the different categories from the low dimensional nonlinear space to high dimensional linear separable space. Then searching for optimal the separating hyperplane by solving linear equations in mapping space, to form the classifier. Above is the training process. Inputting the unknown category data which has already been preprocessed into classifier to classify, we can get the classification results. This research used SVM in Lib-SVM platform, and adopted RBF kernel function.

Table 1. The Formula and Explanation of different Characteristics

| Characteristics | Formula | Explanation |
|---|---|---|
| AL_ratio | $AL\_ratio = \frac{Area}{Length}$ | The ratio of leaf area and leaf length |
| AP_ratio | $AP\_ratio = \frac{Area}{Peremeter}$ | The ratio of leaf area and leaf perimeter |
| Eccentricity | $Eccentricity = \frac{AxisLength_{long}}{AxisLength_{short}}$ | The ratio of leaf length and leaf width |
| Rectangularity | $Rectangularity = \frac{Area_{object}}{Area_{bounding-box}}$ | The ratio of leaf area and area of the smallest encasing box area of leaf |
| Area Convexity | $Area\_convexity = \frac{Area}{Convex\ Area}$ | The ratio leaf area and area of convex Hullof leaf |
| Circularity | $Circularity = \frac{R_{inscribedcircle}}{R_{excircle}}$ | The ratio of radius of inscribed circle and circumcircle |
| Form Factor | $S = \frac{P}{4 \cdot \sqrt{A}}$ | P ,A are the perimeter and area of leaf |

$$K(X, X_i) = \exp\left\{-\frac{\|X - X_i\|^2}{2\sigma^2}\right\}$$

The classification function is

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{l} \alpha_i y_i K(X, X_i) + b\right]$$

Here σ and c are both kernel parameters which are calculated by grib.py of Lib-SVM.

For validating the model, because the small sample size, we chose Leave-One-Out Cross Validation (LOO-CV). This method supposes N samples of original data divide into N groups. Each sample as an individual validation set, while the left N-1 samples as the training set, then N models can be obtained. The average classification accuracy from those N models is the performance index of classifier (Wu Liang 2011). LOO-CV has two apparent advantages.

In each round, all the samples are used in training model, so it's closest to the primary distribution of samples. The result is quite reliable.

There is no stochastic factor to influence the experimental data, ensuring a reproducible experiment process.

#### 2.3.3 Identification of rice phosphorus stress based on Fisher discriminate analysis

Fisher discriminate analysis can judge which class the research object belongs to by observing and measuring variable value. Discriminate analysis establishes discriminate function by filtering the variables including more information from those variables that can show the classification and characteristic of observation object. With this method, misjudge rate can be the lowest. And its common formation is

$$y = a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + a_n x_n$$

Here y is discriminate value, $x_1$, $x_2$, $x_3$…, $x_n$ are variables reflecting the characters of research object, and $a_1$, $a_2$, $a_3$…, $a_n$ are discriminate coefficients.

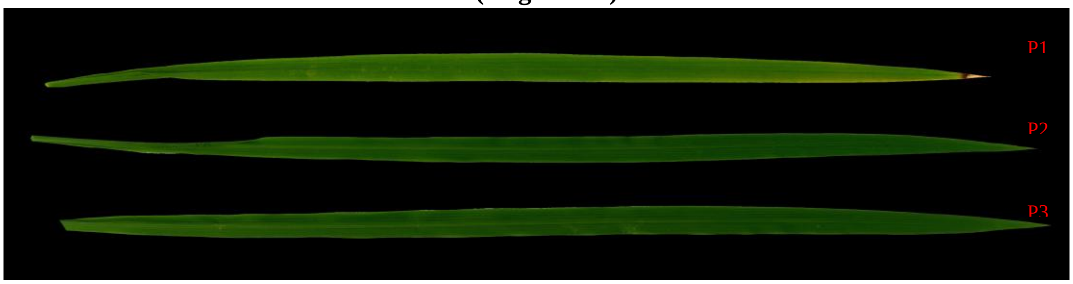
## 3. Results and Discussion

### 3.1 Optimal selections of leaves' characteristic parameters in different growth period

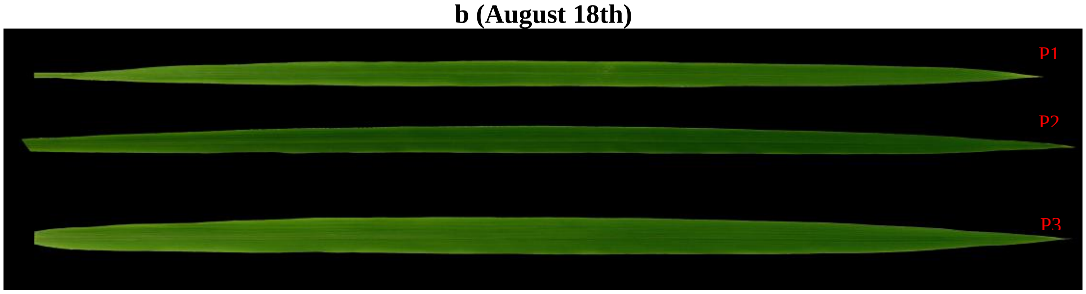
Samples collected from every growth period are divided into 3 types according to the P level: extreme deficiency (0 mg.L$^{-1}$), medium deficiency (2.5 mg.L$^{-1}$, 5 mg.L$^{-1}$, and 7.5 mg.L$^{-1}$) and normal (10 mg.L$^{-1}$), which will be represented by P1, P2 and P3, respectively. In Table 2 and 3, with the same leaf-position (leaf-position ratio) of different growth period the characteristic parameter's F value is different, which means with different P levels the inter-group difference of rice leaf's characteristic parameter are different. In different leaf-position (leaf-position ratio) of the same growth period, the inter-group differences also are different. So, in different growth period, to build models and identify different degrees of rice phosphorus stress, we should choose optimal characteristic parameters according to different inter-group difference.

Figure 1 shows that the color and shape of rice leaves in different P nutrient conditions were markedly different. In 4 growth stages, under extreme phosphorus deficiency the rice leaves appeared yellow and had smaller area. Compare with the normal phosphorus level, the rice leaves with medium phosphorus deficiency had dark green color, were narrow, and had small unit leaf area and sharply pointed.

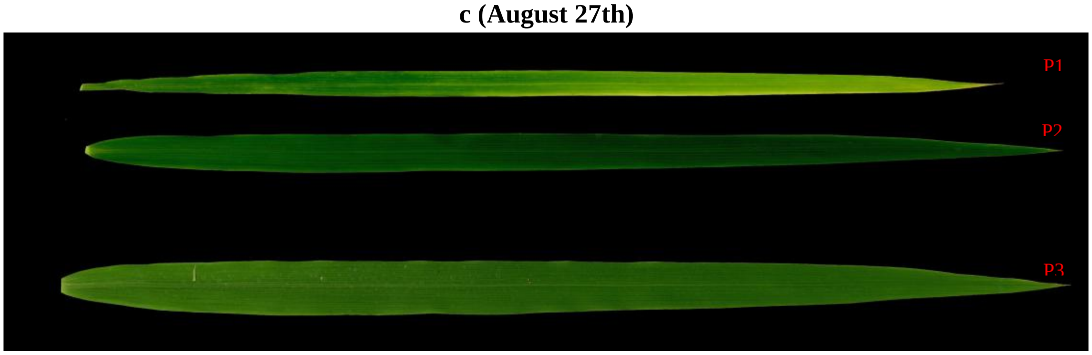

**a (August 4th)**




**b (August 18th)**




**c (August 27th)**




**d (September 8th)**

Figure 1. The difference of symptomatic characteristics in different P treatments (P1 is extreme phosphorus deficiency, P2 is medium phosphorus deficiency, and P3 is normal phosphorus level)

So the characteristic parameters which have biggest differences among 3 P nutrient levels were selected from the acquired 14 characteristic parameters to construct models for identifying rice phosphorus stress.

In addition, in order to enlarge the differences of characteristic parameter between various P nutrient levels, we introduce characteristic parameter ratio between the different leaf-positions (the first leaf /the third leaf, the second leaf /the third leaf, the first leaf /the second leaf).

From Tables 2 and 3, among all characteristic parameters, the one with the largest inter-group difference in color feature are R band and G band, in shape feature are leaf length, leaf width, leaf area, leaf perimeter, area/Length, area/perimeter, which agree with physiological characteristics of rice phosphorus deficiency. After calculation it is found that F is over 3, modeling and identifying can have the highest accuracy. So we choose characteristic parameter whose F value is larger than 3 (Table 2, Table 3) for identifying utilizing SVM and Fisher discriminate analysis.

Table 2. ANOVA of rice leaves' characteristic parameters from Different growth stages

| | The first leaf | | | | The second leaf | | | | The third leaf | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 804 | 818 | 827 | 908 | 804 | 818 | 827 | 908 | 804 | 818 | 827 | 908 |
| R | 9.5 | 8.5 | 3.59 | 11.79 | 36.61 | 33.48 | 38.6 | 36.59 | 28.62 | 35.4 | 31.96 | 69.17 |
| G | 17.56 | 10.63 | 2.53 | 11.69 | 43.79 | 47.31 | 36.81 | 37.53 | 23.3 | 109.49 | 38.54 | 98.11 |
| B | 2.45 | 4.72 | 5.71 | 8.64 | 1.09 | 3.55 | 10.14 | 13.35 | 2.36 | 6.83 | 18 | 7.06 |
| Leaf Length | 26.41 | 160.65 | 8.44 | 8.34 | 19.1 | 116.84 | 61.68 | 8.46 | 11.96 | 30.87 | 11.79 | 112.49 |
| Leaf Width | 14.32 | 22.12 | 27.12 | 36.52 | 22.83 | 16.38 | 11.04 | 44.83 | 8.14 | 15.89 | 3.9 | 13.12 |
| Leaf Area | 30.59 | 51.56 | 24.22 | 13.27 | 23.19 | 27.96 | 25.23 | 28.72 | 20.68 | 31.81 | 12.49 | 54.83 |
| Leaf Perimeter | 24.94 | 136.79 | 8.88 | 8.65 | 16.71 | 129.08 | 63.53 | 8.68 | 11.35 | 27 | 13.62 | 74.87 |
| AL_ratio | 28.55 | 33.8 | 38.31 | 24.27 | 17.49 | 13.29 | 15.08 | 80.3 | 20.64 | 26.78 | 8.57 | 39.24 |
| AP_ratio | 30.79 | 36.55 | 37.84 | 24.11 | 18.56 | 13.36 | 15.16 | 82.76 | 23.73 | 26.05 | 8.39 | 43.45 |
| Eccentricity | 1.83 | 0.12 | 4.38 | 9.38 | 0.96 | 3.86 | 3.41 | 12 | 0.85 | 0.12 | 1.5 | 0.87 |
| Rectangularity | 0.92 | 2.02 | 1.04 | 0.98 | 0.66 | 1.94 | 1.44 | 1.24 | 1.67 | 6 | 0.32 | 0.83 |
| Area Convexity | 2.13 | 0.91 | 1.4 | 2.23 | 0.83 | 1.07 | 1.37 | 5.8 | 2.45 | 3.82 | 0.28 | 1.05 |
| Circularity | 1.36 | 0.6 | 5.54 | 10.82 | 0.72 | 1.71 | 1.06 | 16.35 | 1.62 | 0.16 | 0.21 | 1.79 |
| Form Factor | 0.39 | 2.55 | 6.11 | 11.66 | 0.3 | 2.61 | 3.46 | 14.12 | 2.92 | 2 | 0.38 | 6.54 |

**F value larger than 3 is in underscores; 804, 818, 827 and 908 indicate 4 growth period (August 4th, August 18th, August 27th, September 8rd), respectively**

Table 3. ANOVA of rice leaves' characteristic parameter ratio from Different growth stages

| | first leaf /third leaf | | | | second leaf /third leaf | | | | first leaf /second leaf | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 804 | 818 | 827 | 908 | 804 | 818 | 827 | 908 | 804 | 818 | 827 | 908 |
| R | 4.14 | 16.52 | 9.99 | 10.9 | 5.4 | 28.32 | 11.3 | 28.79 | 1.21 | 4.14 | 4.6 | 0.17 |
| G | 1.23 | 9.88 | 5.48 | 4.84 | 1.98 | 16.85 | 6.79 | 6.43 | 1.09 | 2.32 | 2.75 | 0.43 |
| B | 0.6 | 3.14 | 3.24 | 8.01 | 0.24 | 7.47 | 4.14 | 4.4 | 1.31 | 4.56 | 0.16 | 3.81 |
| Leaf Length | 7.2 | 21.62 | 0.81 | 0.47 | 1.59 | 13.99 | 3.61 | 4.19 | 7.42 | 2.06 | 2.08 | 2.26 |
| Leaf Width | 1.64 | 8.09 | 3.97 | 2.79 | 0.54 | 4.22 | 0.19 | 2.5 | 2.2 | 5.85 | 4.97 | 3.88 |
| Leaf Area | 7.63 | 31.11 | 3.44 | 0.76 | 0.11 | 5.01 | 4.53 | 0.8 | 13.5 | 42.47 | 1.64 | 2 |
| Leaf Perimeter | 7.37 | 15.69 | 0.77 | 0.23 | 1.41 | 9.16 | 3.77 | 3.09 | 6.85 | 3.42 | 1.91 | 2.36 |
| AL_ratio | 1.88 | 8.47 | 6.4 | 3.02 | 0.25 | 2.16 | 0.39 | 14.22 | 6.65 | 52.22 | 15.54 | 2.65 |
| AP_ratio | 1.95 | 7.29 | 6.68 | 2.44 | 0.35 | 1.76 | 0.59 | 13.21 | 6.66 | 45.76 | 14.91 | 2.64 |
| Eccentricity | 2.06 | 0.23 | 2.45 | 6.7 | 0.53 | 1.22 | 0.07 | 14.25 | 1.69 | 4.16 | 7.76 | 4.12 |
| Rectangularity | 0.96 | 4.99 | 0.15 | 0.71 | 0.8 | 10.7 | 0.11 | 0.67 | 0.69 | 1.2 | 0.95 | 1.57 |
| Area Convexity | 1 | 3.17 | 0.15 | 1 | 1.6 | 3.26 | 0.27 | 1.43 | 1.41 | 0.6 | 0.52 | 3.42 |
| Circularity | 1.06 | 0.93 | 1.48 | 2.52 | 0.47 | 1.6 | 0.04 | 7.8 | 1.38 | 3.85 | 5.35 | 2.44 |
| Form Factor | 2.36 | 1.77 | 1.3 | 2.02 | 2.1 | 4.8 | 0.79 | 6.49 | 0.75 | 21.42 | 8.21 | 2.91 |

**F value larger than 3 is in underscores; 804, 818, 827 and 908 indicate 4 growth period (August 4th, August 18th, August 27th, September 8rd), respectively**

### 3.2 SVM modeling and validation

In order to know the identification accuracy of the different degrees of phosphorus stress in rice's different growth period, as well as to choose the optimal leaf-position. In every growth period we chose characteristic parameters of different leaf-position (leaf-position ratio) with a larger-than-3 F value for modeling and classification, and then validating through LOO-CV.

From Figure 2, we can obtain the training accuracy of identifying different degrees of phosphorus stress with SVM, and the validation accuracy of identifying with LOO-CV. In Figure 2, on August 4th, choosing 3rd leaf's characteristic parameters to identify, the maximum identifying accuracy (87.5%) was obtained and the validation accuracy with LOO-CV was 79.19%, this validation accuracy corresponded closely to the maximum validation accuracy (1st leaf, 83.33%). So, on August 4th, the identifying accuracy depending on 3rd leaf's characteristic parameters was maximum, and also stable. On August 18th, with 2nd /3rd and 1st/2nd leaf the identifying accuracy all was maximizing (100%),and using1st/2nd leaf had higher validation accuracy(95.65%) than 2nd/3rd leaf , so this study thinks using 1st/2nd leaf's characteristic parameters for identifying rice phosphorus stress on August 18th can get the best recognition accuracy and reliable result. On August 27th, when using the 1st leaf's characteristic parameters to classify, the training and validation accuracy get to the highest, respectively 92% and 88%. On September 8th, using 1st, 2nd, 3rd leaf to classify, the training and validation accuracy both reach 100%. As a result, any one of them can be used to diagnosis the phosphorus stress of rice.

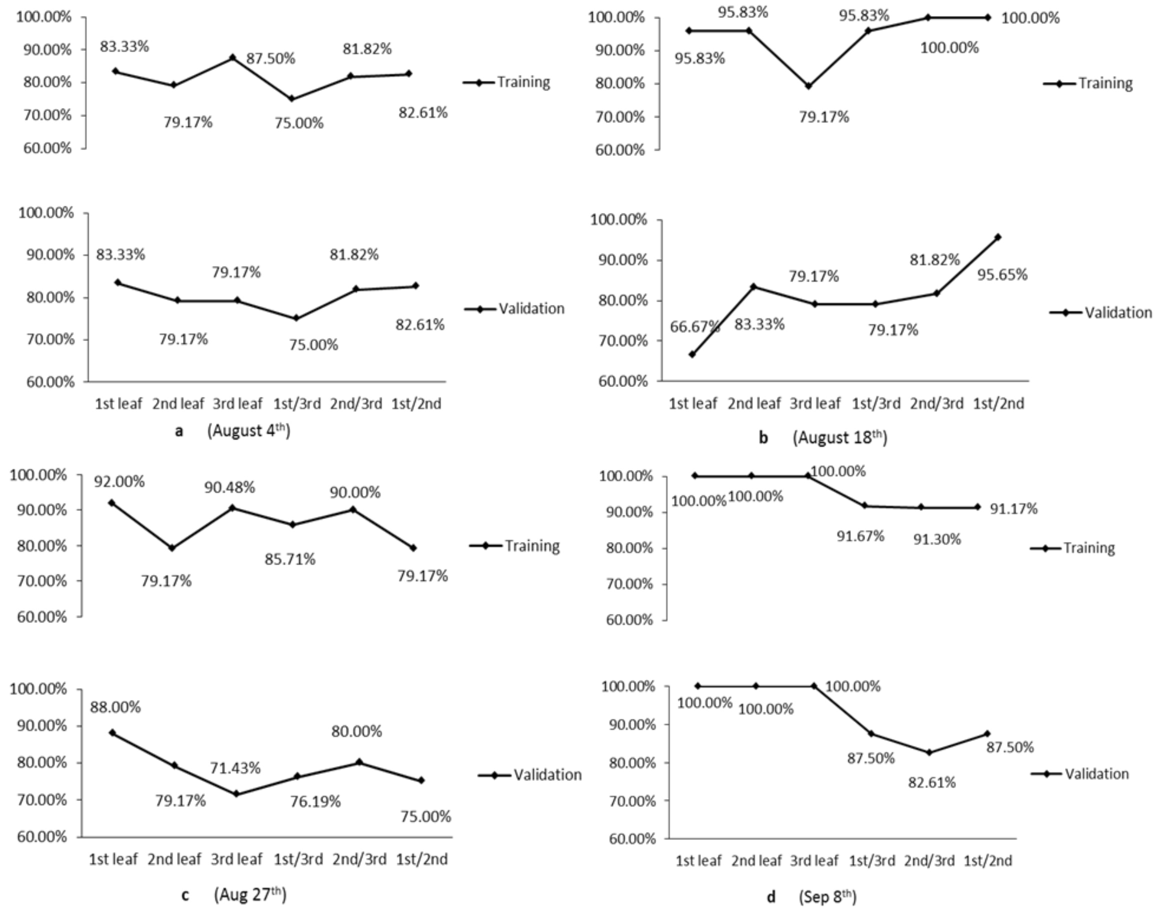



Figure 2. The SVM training accuracy and the LOO−CV validation accuracy of the different leaf-position in different growth stages

Figure 2 indicates the optimized leaf-position (leaf-position ratio) whose characteristic parameters can be used to identification rice phosphorus stress with the maximum accuracy. But in order to know the identifying accuracy of different degrees of phosphorus stress, we had done further research. On August 4th, the research finds the identification accuracy was up to 100% under P1 and P2. But, the identification misdiagnosed to be medium phosphorus deficiency. On August 18th, recognition rate all achieved 100% under P1, P2 and P3. On August 27th, under P1 the recognition accuracy was 100%, under P2 the recognition accuracy was only 92.58% with the left misdiagnosed to be P3, and under P3 it was 80% with the left misdiagnosed to be P2.

**3.3 Identification results with Fisher Discriminate Analysis Fisher**

After optimized choice of leaf samples' characteristic parameters by one-way ANOVA, chose F value being larger than 3 from different leaf-position to diagnose rice's phosphorus stress by Fisher discriminate analysis. On August 4th, 18th ,27th and September 8th, for different degree of phosphorus stress the maximum training accuracy respectively were 91.70% (2nd leaf), 95.80% (1st/2nd), 91.70% (2nd leaf), and 100% (1st, 2nd, 3rd), and the corresponding validation accuracy respectively were 79.20%, 83.30%, 58.30%, 100%. On August 27th, with the 2nd leaf's characteristic parameters the training accuracy was maximum for identifying, but the model was not stable as the low validation accuracy (58.30%). At the same growth period, to identify with the 1st leaf's characteristic parameters resulted in a low training accuracy (90.50%) which was nearly with 2nd leaf (91.70%), but its validation accuracy increased to 81.00%. So we chose the 1st leaf on August 27th.

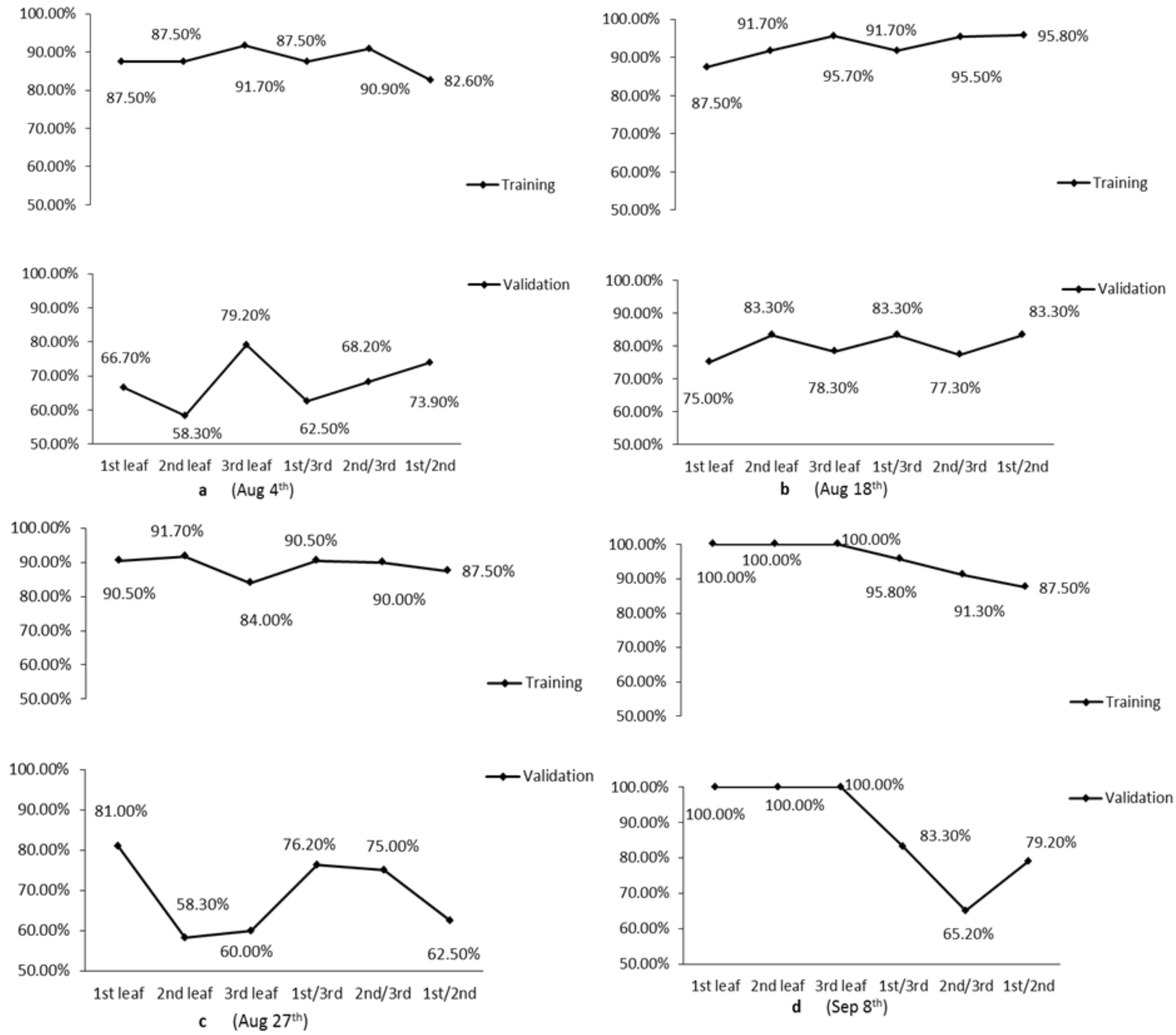



Figure 3. The Fisher discriminate analysis training accuracy and the validation accuracy of the different leaf-positions at the different growth stages

After the identifying of rice phosphorus stress of every growth period by Fisher discriminate analysis, the study found the optimal leaf-position which had the maximum identifying accuracy of phosphorus stress. Table 4 shows the optimal leaf-position with the maximum identifying accuracy of phosphorus stress in every growth period by Fisher discriminate analysis. It was coincident with what tested by SVM. In this case, it means when diagnosing rice phosphorus stress by SVM the leaf-position with the maximum identifying accuracy is also the one which has the most obvious difference of physiological characteristic from different degree of phosphorus stress.

Table 4. ANOVA of rice leaves' characteristic parameter ratio from Different growth stages

| Leaf-position (Growth period) | 3rd leaf (804) | 1st/2nd leaf (818) | 1st (827) | 1st,2nd,3rd leaf (908) |
|---|---|---|---|---|
| SVM | 87.50% | 100.00% | 92.00% | 100% |
| LOO-CV[a] | 79.17% | 95.65% | 88.00% | 100% |
| Discriminate Analysis | 91.70% | 95.80% | 90.50% | 100% |
| LOO-CV[a] | 79.20% | 83.30% | 81.00% | 100% |

**a. LOO-CV only conducted in the cases in this analysis. In validation, every case is sorted by the function derived by all cases but this one**

When identifying different degree of rice phosphorus stress, it was found that in different growth period the identification accuracy were different by using different leaf-position's characteristic parameters. With the 3rd, 1st/2nd, 1st, and 1st-2nd-3rd leaf's characteristic parameters from 4 growth period, respectively, Table 4 showed the maximum identifying accuracy by SVM and Fisher discriminant analysis. It agrees with the rice physiological characters under phosphorus stress in all growth period. On August 4th, rice was in the early part of tillering stage. Because of the mobility of phosphorus in rice plant, phosphorus deficiency symptom is mainly appeared in the 3rd leaf. And, at this period, rice requires little phosphorus for growth, so there's no significant difference between P2 and P3. On August 18th, most of phosphorus in the 3rd leaf had moved to the 1st leaf, and phosphorus in the 2nd leaf also gradually moved to the 1st leaf. So the 2st leaf had the most significant difference of phosphorus deficiency symptom. As the ratio of different leaf-position's characteristic parameters can enlarge this difference, the 1st/2nd leaves' characteristic parameter had the largest difference, so it also enlarged the identifying accuracy of different degree of phosphorus stress. The phosphorus content of the 2nd leaf gradually stopped changing with time going, on August 27th the 2nd leaf almost stopped moving phosphorus to the 1st leaf. So the ratio of different leaf-position's characteristic parameters no longer enlarged the difference, at the same time the 1st leaf was affected by phosphorus deficiency, rice leaves showed the obvious phosphorus deficiency symptom. On September 8th, as a result of long-time phosphorus stress, every leaf-position of rice showed extreme phosphorus deficiency symptom. At this period, by any leaf-position the identifying accuracy can get the maximum. For the same reason, the rice was under phosphorus stress for long term, the ratio of different leaf-position's characteristic parameters cannot enlarge the difference of different degree of phosphorus deficiency symptom.

**4. Conclusions**

This paper takes the three most anterior leaves of rice under different phosphorus nutrition condition as the object of research. In laboratory condition, spectral and morphological characteristic parameters were acquired from the scanning image of rice leaves to diagnose the phosphorus status. Also this study introduce characteristic parameter ratio to increase identification accuracy. Finally, SVM is used to identify the rice's phosphorus nutrient level.

During four growth periods, the overall recognition rates of different phosphorus nutrition level are 87.5%, 100%, 92% and 100%, respectively. In detail, P1's recognition rates are the highest, which all reach 100%. And P2's recognition rates are 100%, 100%, 92.85%, 100%, respectively, and P3's recognition rates are 40%, 100%, 80%, 100%, respectively. On August 4th, 60% of P3 samples are misclassified as P2, so it may be difficult to identify P deficiency in very early growth period. And in other growth periods, the recognition rates are high for different P nutrient status.

The study provides evidence for quick diagnosis of P nutrient status, which makes it possible to accurately identify rice phosphorus stress with scanning technology and SVM.

With a portable scanner, the images of rice leaves can be collected quickly by the static scanning technology under field conditions, and the rice phosphorus stress might be identified for guiding fertilization in time. Other crops such as maize, wheat with phosphorus stress, they usually have some symptom by the shape and color of leaf, so this method can be used to diagnose the phosphorus situation for them. So, the technology introduced in the study has the application value and development potential.

**Acknowledgements:**

The author thanks Mr. Cai guangzhe for helping with scanning rice leaves. The author also thanks the Professor Wang Ke for his constant encouragement for the work. Financial assistance is from the National Natural Science Foundation of China (Grant No.31172023).

**Corresponding Author:**
Prof. Shen Zhangquan,
E-mail: zhqshen@zju.edu.cn
Prof. Wang Ke,
E-mail: kwang@zju.edu.cn
Institute of Agricultural Remote Sensing and Information Technology Application, Zhejiang University, Hangzhou, Zhejiang 310058, China.

**References**

[1] Brown, L., Jin, M. C., Leblanc, S. G., et al. 2000. A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: an image and model analysis. REMOTE SENS ENVIRON. 71:16-25.
[2] Chen Jian, Chen Xiao, Li Wei, et al. 2008. Near-Infrared Spectra and Artificial Neural Network Model. SPECTROSC SPECT ANAL. 28(8):1806-1809.
[3] Christopher J.C. Burges. 1998. A Tutorial on Support Vector Machines for Pattern Recognition. DATA MIN KNOWL DISC. 2:121-167.

[4] Deng Nai-yang, Tian Ying-jie. 2004. New Method of Data Minning-Support Vector Machine. Science Press.
[5] Guan Ze-xin, Tang Jian, Yang Bao-jun, et al. 2010. Study on Recognition Method of Rice Disease Based on Image. Chin J Rice Sci. 24(5): 497-502.
[6] Jia Liang- liang, Fan Ming-sheng, Zhang Fu-suo, et al. 2009. Nitrogen Status Diagnosis of Rice by Using a Digital Camera. SPECTROSC SPECT ANAL. 29(8): 2176-2179.
[7] Jiang Li-geng, Cao Wei-xing, Jiang Dong, et al. 2004. Distribution of leaf nitrogen, amino acids and chlorophyll in leaves of different positions and relationship with nitrogen nutrition diagnosis in rice. Acta Agronomica Sinica. 30(8): 739-744.
[8] Li Li-mei, Guo Rong-fa, Guo zhi-juan, et al. 2007. Studies on differences in phosphorus nutrition of rice genotypes. Journal of Northwest A & F University (Nat. Sci. Ed.). 35(11):75-79.
[9] L. Bacci, M. De Vincenzi, B. Rapi, et al. 1998. Two methods for the analysis of colorimetric components applied to plant stress monitoring. COMPUT ELECTRON AGR. 19:167-186.
[10] Li Guo-zheng, Wang Meng, Zeng Hua-jun. 2004. An Introduction to Support Vector Machines and other Kernel Based Learning Methods. Bejing: Publishing House of Electronics Industry Press.
[11] Lu Jian-wei, Li rong, et al. 2010. Common rice nutrient deficiency symptom and correction technology. Chinese Agricultural Press.
[12] Milos Kovacevica, Branislav Bajata, Bosko Gajic. 2010. Soil type classification and estimation of soil properties using support vector machines. GEODERMA. 154: 340-347.
[13] Nemmour, H., Chibani, Y. 2006. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. ISPRS J PHOTOGRAMM. 61: 125.
[14] Shi Jing-jing, Liu Zhan-yu, Zhang Li-li. 2009. Hyperspectral Recognition of Rice Damaged by Rice Leaf Roller Based on Support Vector Machine. Chin J Rice Sci. 23(3):331-334.
[15] Tracy M. Blackmer, James S. Schepers. 1996. Aerial Photography to Detect Nitrogen Stress in Corn. J PLANT PHYSIOL (148): 440-444.
[16] Wang Bin. 2011. Application of Clustering-based Decision Tree in the Screening of Maize Germplasm. Agricultural Science & Technology. 12(10): 1449-1452.
[17] Wei Huang, Yoshiteru Nakamori, Shou-Yang Wang. 2005. Forecasting stock market movement direction with support vector machine. COMPUT OPER RES. 32: 2513-2522.
[18] Wu Liang, Luo Hang-zai, Zhu Ming-hua, et al. 2011. Efficient Leave-One-Out Evaluation of SVM Models. Proceedings of 2011 IEEE International Conference on Intelligent Computing and Intelligent Systems Vol. 3.
[19] Zhao Qing-lei, Wang Kai-rong, MA Jia-qing, et al. 2009. Effects of Long-Term Application of Different Fertilizer Patterns on Rice Paddy Soil Phosphorus and Rice Phosphorus Nutrition. ACTA AGRONOMICA SINIC. 35(8): 1539–1545.
[20] Zhu Jin-xia, Deng Jin-song, Shi Yuan-yuan, et al. 2009. Diagnosis of rice nitrogen status based on characteristics of scanning leaf. SPECTROSC SPECT ANAL. 29(8):2171-2175.

6/20/2013