Detection of Rice Exterior Quality based on Machine Vision

Xiai Chen, * Shuang Ke, ** Wenquan Chen

* School of Mechanical and Electrical Engineering, China Jiliang University, Hangzhou, China, xachen@cjlu.edu.cn
** Zhejiang Economic & Trade Polytechnic, Hangzhou, China, kkeshuang@163.com
*** China Jiliang University, Hangzhou, China, chen0596@126.com

Abstract

To investigate the detection of rice exterior quality, a machine vision system was developed. The main characteristics of rice appearance including area, perimeter, roughness and minimum enclosing rectangle were calculated by image analysis. Least Squares Support Vector Machines, Naive Bayes Classifier and Back Propagation Artificial Neural Network were applied to achieve classification of head rice and broken rice, and the classification results of three algorithms were analyzed in detail. Genetic algorithm and Particle Swarm Optimization were used to obtain the optimal values of regularization parameter and kernel radial basis function parameter, and adopt a supervised learning approach to train the Least Squares Support Vector Machines model. Meanwhile the robustness of these classification methods was tested, and the results shows that support vector machine have better classification results in this experiment. This study demonstrated the feasibility of detection rice quality using machine vision.

Keywords: Rice, Machine Vision, LS-SVM, Classification, Image Processing

1. Introduction

Commonly the quality of rice includes internal quality and exterior quality. The shape of rice is an important characteristic of rice exterior quality which directly affects the sales and prices of rice. When the rice kernels are smaller than three-fourths of whole kernel they are defined as broken rice which seriously affecting the quality of rice, so it’s very important to distinguish broken rice from head rice [1]. Currently, the appearance quality of rice is measured manually by inspector. This detection method is simple, but subjective and inefficient.

There is already a number of researching about rice quality detection [2-6]. However, most studies were about how to obtain the characteristic parameters of rice, while care less about the algorithm applied for classification of rice quality.

In this paper, machine vision was applied and the rice shape parameters were got after image processing and analysis. The classification and identification of rice grades was achieved by Least Squares Support Vector Machines (LS-SVM), Naive Bayes Classifier and Back Propagation Artificial Neural Network (BP-ANN). The Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) were used to acquire the parameters of LS-SVM. The experiment result shows that the head rice and broken rice can be effectively identified by LS-SVM using machine vision.

2. Data acquisition and analysis

The hardware of machine vision system applied in this experiment was shown in Figure 1. It’s includes image acquisition, image processing and image analysis and display.

Figure 1. Hardware Schematic Diagram of Machine Vision System
After the RGB image acquired by frame grabber card, it was transformed to a gray image and then pre-processed by image processing algorithms including Gaussian filter, erosion, dilation, binarization, edge detection, and filling. These algorithms were used to smooth and enhance the image, preparing it for further analysis.

The eight characteristics of rice image were calculated by image analysis algorithms as shown in Table 1.

<table>
<thead>
<tr>
<th>ID</th>
<th>characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>perimeter</td>
<td>perimeter of the rice image</td>
</tr>
<tr>
<td>2</td>
<td>Area</td>
<td>pixels number of the rice image</td>
</tr>
<tr>
<td>3</td>
<td>Roughness</td>
<td>Perimeter^2 / (4π × Area)</td>
</tr>
<tr>
<td>4</td>
<td>MERL</td>
<td>length of minimum enclosing rectangle</td>
</tr>
<tr>
<td>5</td>
<td>MERW</td>
<td>width of minimum enclosing rectangle</td>
</tr>
<tr>
<td>6</td>
<td>MERArea</td>
<td>MERL × MERW</td>
</tr>
<tr>
<td>7</td>
<td>MERProportion</td>
<td>MERL / MERW</td>
</tr>
<tr>
<td>8</td>
<td>solidity</td>
<td>Area / MERArea</td>
</tr>
</tbody>
</table>

Grayscale images of rice, binary image of rice after image pre-processing, and the minimum bounding rectangle of rice calculated by image analysis were shown as Figure 2, Figure 3, and Figure 4.

In order to reduce the influence of background noise, an image subtraction method was applied. First, we got the last moment pure background image, and then subtracted this background image from the rice image. The background image was updated periodically instead of being fixed, further improving the image detection and recognition accuracy.

3. Classification method

3.1. Least squares support vector machines

Least Squares Support Vector Machines (LS-SVM) is presented by Suyken J. A. K[7]. The main differences of the SVM and the LS-SVM are as follows: 1) The SVM solution is typically found by solving a quadratic programming problem while the LS-SVM solution can be found by solving a set of linear equations. 2) The SVM has inequality constraints whereas the LS-SVM has only equality constraints. 3) LS-SVM reduces the computational cost and is easy to implement for on-line applications [8-10].
Given a training data set of \( K \) points \( \{x_i, y_i\}_{i=1}^{K} \), with input data \( x_i \in \mathbb{R}^n \) and output data \( y_i \in \mathbb{R}^n \), LS-SVM aims to build a classifier of the form as follows

\[
y(x) = \text{sign} \left( \sum_{k=1}^{N} a_k y_k \psi(x, x_k) + b \right)
\]

where \( a_k \) are positive real constants, \( b \) is a real constant and \( \psi(x, x_k) \) is a Kernel function. The following optimization problem is formulated for classification in LS-SVM.

\[
\min_{w, b, e} J_s(w, b, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{K} e_k^2
\]

Subject to constraint as follows

\[
y_k \left[ w^T \psi(x) + b \right] = 1 - e_k, \quad k = 1, \ldots, N
\]

To solve this quadratic programming, Lagrange multiplier is used as follows

\[
L_s(w, b, e, \alpha) = J_s(w, b, e) - \sum_{k=1}^{K} \alpha_k \left[ y_k \left[ w^T \psi(x) + b \right] - 1 + e_k \right]
\]

where \( \alpha_k \) are Lagrange multipliers, the conditions for optimality are given as follows

\[
\frac{\partial L_s}{\partial w} = 0 \rightarrow w = \sum_{k=1}^{K} a_k y_k \psi(x_k)
\]

\[
\frac{\partial L_s}{\partial b} = 0 \rightarrow \sum_{k=1}^{K} a_k y_k = 0
\]

\[
\frac{\partial L_s}{\partial e_k} = 0 \rightarrow a_k = \gamma e_k, \quad k = 1, \ldots, N
\]

\[
\frac{\partial L_s}{\partial a_k} = 0 \rightarrow y_k \left[ w^T \psi(x) + b \right] - 1 + e_k = 0, \quad k = 1, \ldots, N
\]

Eq. 5 can be written as the linear equations

\[
\begin{bmatrix}
I & 0 & 0 & -Z^T \\
0 & 0 & 0 & -Y^T \\
0 & 0 & \gamma I & -I \\
Z & Y & I & 0
\end{bmatrix}
\begin{bmatrix}
w^T \\
b \\
e \\
a
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0 \\
1
\end{bmatrix}
\tag{6}
\]

where \( Z = [\psi(x_1, y_1), \ldots, \psi(x_N, y_N)] \), \( Y = [y_1, \ldots, y_N] \), \( \gamma = [e_1, \ldots, e_N] \), \( a = [a_1, \ldots, a_K] \). With Mercer’s condition

\[
\Omega_k = y_k y_k^T \psi(x) = y_k y_k^T \psi(x_k), \quad k = 1, \ldots, N
\]

The equations related with \( a \) and \( b \), Eq. 6 can be transformed into

\[
\begin{bmatrix}
0 & -Y^T \\
\Omega + \gamma^{-1} I & a
\end{bmatrix}
\begin{bmatrix}
b \\
a
\end{bmatrix}
= \begin{bmatrix}
0 \\
1
\end{bmatrix}
\tag{8}
\]

Least squares method can be applied to calculate \( a \) and \( b \) from linear equations (8).

There are a lots of kernel function used in LS-SVM, and the Radial Basis Function (RBF)

\[
\psi(x, x_k) = \exp\left(-\frac{\|x - x_k\|^2}{\delta}\right)
\]

is the one most commonly used. Normally the values of the kernel
parameter $\delta$ and regularization parameter $\gamma$ were determined by a grid search and cross validation to select those values that give the smallest error on the test data set.

### 3.2 Particle swarm optimization

The particle swarm optimization is proposed by Kennedy and Eberhart in 1995 and has been applied successfully to various optimization problems [11-13]. The particle swarm optimization algorithm can complete optimization search by the collaboration and competition of the individuals in complex space. When solving optimization problem by using the PSO algorithm, each particle, that is to say the solution to the problem, is a point in the search space. Each particle has its own position vector and speed vector, which deciding its moving direction and distance. There is a fitness function which used to evaluate the superiority degree of each particle. All particles memorize and share the current optimal particle, and they will search in the search space until find the optimization solution. The iterative algorithm is used in the PSO algorithm. In each iteration cycle, every particle closes up to two points in the search space at the same time. One point is its own optimum solution in the post search, and the other is the global optimum solution for all particles in historical search. In the course of search, every particle updates its position and speed as the follow formulas.

\[
v^{(t+1)}_{id} = \omega v^{(t)}_{id} + c_1 r_1 (p_{id}^{(t)} - x^{(t)}_{id}) + c_2 r_2 (p_{gd}^{(t)} - x^{(t)}_{id})
\]

(9)

\[
x^{(t+1)}_{id} = x^{(t)}_{id} + v^{(t+1)}_{id}
\]

(10)

where $d = 1, 2, \cdots n$, $i = 1, 2, \cdots m$, $n$ is the dimension of the search space, and $m$ is population size. $x^{(t)}_{id}$ is the current position of particle $i$ at a certain iteration $t$. $v^{(t)}_{id}$ is the velocity of particle $i$ at a certain iteration $t$, deciding the direction of $x^{(t+1)}_{id}$. $p_{id}$ is the optimum solution of particle $i$ in the post search, $p_{gd}$ is the global optimum solution for all particles in historical search, $r_1$ and $r_2$ are random values between 0 and 1, $c_1$ and $c_2$ are learning parameters, usually they are the constants in the range $[0, 2]$. $\omega$ is the inertia weight, which can affect the searching ability. Usually, the inertia weight $\omega$ can be described as follow.

\[
\omega(t) = \omega_{in} - \omega_{in} \frac{T_{end}}{T_{max}} t
\]

(11)

where $\omega_{in}$ is the original inertia weight value, $\omega_{end}$ is the inertia weight value at the most iterative time, $T_{max}$ is the whole iterative time, $t$ is the current iterative time.

### 3.3. Genetic Algorithm

Genetic algorithm is a stochastic global searching and optimizing algorithm which is used to solve complicated problems. GA including five steps of initialization, selection, recombination, mutation, and termination [14]. GA simulates natural evolution process, it start with an initial set of random solutions called population and use some measures of fitness to evaluate each chromosome. Through the recombination process, chromosomes with the higher probability are preserved. Then mutation was used to produce new variable values and prevent optimization process falling into local optimal. After several generations, the algorithms converge to the best chromosome, which hopefully represents the optimum [15, 16].

Genetic algorithm was applied in calculating the kernel parameter $\delta$ and of the regularization parameter $\gamma$ of LS-SVM in this paper. Compared to the grid search method, genetic algorithm greatly improves the problem solving speed.
3.4. Naive Bayes classifier

The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Naive Bayes classifier has gained popularity in solving various classification problems including microarray data analysis [17-19].

Naive Bayes classifier assuming strong independence within attributes of an instance.

\[
p(C|A_1, \ldots, A_n) = \frac{1}{Z} \prod_{i=1}^{n} p(A_i|C)
\]

where \( Z \) is a scaling factor depends only on \( A_1 \) through \( A_n \), \( C \) is the class variable, \( A_1 \) through \( A_n \) are independent attributes variables, and \( p(C) \) is called class prior probability.

3.5. BP Artificial neural networks

Back Propagation Artificial Neural Network is one of the most widely used neural networks. It is a single-direction multilayer neural network that contained of input layer nodes, output layer nodes and one or more layers of hidden nodes. The information transfer from nodes to nodes of different layers, and the degrees of the connections are controlled by the connection weights. The connection weights are adjusted on the basis of data by training. BP Artificial Neural Network input information from the input layer and output the decisions to the output layer. Actually, the BP Artificial Neural Network is a kind of highly nonlinear mapping from input to output [20-22].

3.6. Model validation

Model validation method was applied in order to evaluate the performance of classification model. The goal of cross-validation is to estimate the expected level of fit of a model to a data set that is independent of the data that were used to train the model.

Commonly used model validation methods include Holdout validation, K-fold validation and leave-one-validation. Leave-one-out cross-validation (LOOCV) is one of the most widely used cross-validation method, it involves using a single observation from the original sample as the validation data, and the remaining observations as the training data. This is repeated such that each observation in the sample is used once as the validation data.

4. Classification result

The Radial Basis Function was selected as the kernel function of LS-SVM in this experiment. Gaussian distribution and kernel density estimation was used in Naive Bayes Classifier respectively.

In order to realize the classification of rice, two different four-layer neural networks was constructed in this paper, which with 6 hidden layer neurons and 8 hidden layer neurons respectively. \textit{Trainlm} was selected as the BP Artificial Neural Network training function, \textit{purelin} was selected as the output layer as transfer function and \textit{tansig} was selected as the hidden layer transfer function. The training goal was set to 0.001 and the maximum number of training was set to 10,000.

The eight parameters calculated by image analysis of each 200 head rice and 100 broken rice were used as the training and testing data by cross validation.

All three classification algorithm used the same training samples data and testing samples data, which shown as Table 2. To test the robustness of these classification methods, 10% and 20% noise was added respectively, which showed as Table 3 and Table 4.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Recognition rate</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Head rice</td>
<td>Broken rice</td>
</tr>
<tr>
<td>LS-SVM</td>
<td>99.67%</td>
<td>100%</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>99.67%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 3. Classification result of rice detection with 10% Noise

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Recognition rate</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-SVM</td>
<td>99.67% 100% 99%</td>
<td>RBF Kernel</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>99.33% 99.5% 99%</td>
<td>Gaussian distribution</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>99.33% 99.5% 99%</td>
<td>kernel density estimation</td>
</tr>
<tr>
<td>BP ANN</td>
<td>98.33% 99.5% 96%</td>
<td>8 hidden layer neurons</td>
</tr>
<tr>
<td>BP ANN</td>
<td>99% 99.5% 98%</td>
<td>6 hidden layer neurons</td>
</tr>
</tbody>
</table>

Table 4. Classification result of rice detection with 20% Noise

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Recognition rate</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-SVM</td>
<td>99.67% 100% 99%</td>
<td>RBF Kernel</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>99% 99.5% 98%</td>
<td>Gaussian distribution</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>99% 99.5% 98%</td>
<td>kernel density estimation</td>
</tr>
<tr>
<td>BP ANN</td>
<td>98.67% 99% 98%</td>
<td>8 hidden layer neurons</td>
</tr>
<tr>
<td>BP ANN</td>
<td>97% 97% 97%</td>
<td>6 hidden layer neurons</td>
</tr>
</tbody>
</table>

The experiment results show that appearance quality of rice can be effectively classified. Dealing with the data of noise pollution shows that least support vector machines have very strong generalization ability. The identification result of BP Artificial neural network is affected by training sample size, training errors, and the structural complexity of network. For the number of hidden number of BP Artificial neural network, too small networks are unable to adequately learn the problem while excessively large networks tend to overfit the training data and consequently result in weak performance. The LS-SVM is an algorithm based on the statistical learning theory for small samples, which transforms the problem of searching for the optimal hyperplane and it’s well suited for solving data as type of this experimental. The classification result of this experiment indicate that the LS-SVM outperforms the Naive Bayes Classifier and BP Artificial neural network significantly, when the training samples are limited.

Grid search is one of the conventional approaches to determine hyper-parameters. However, it needs an exhaustive search over the space of hyper-parameters, which must be time consuming[23]. The regularization parameter was set to be 500 and the RBF kernel parameter was set to be 100 in LS-SVM classification Algorithm after optimization.

Genetic Algorithms and Particle Swarm Optimization were applied in getting the regularization parameter \( \gamma \) and the RBF kernel parameter \( \delta \) of LS-SVM. Compared to grid search and cross validation, the speed of obtaining optimal parameters has been greatly improved by Genetic Algorithms and Particle Swarm Optimization, which shown as Table 5. The Operating System of tested computer was windows XP, which had Intel Core 2 Duo P8600 processors and 2GB DDR2 memory.

Table 5. Optimization of regularization parameter and RBF kernel parameter of LS-SVM

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>cross-validation</th>
<th>Time-consuming (s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GD-LS-SVM (Grid Research)</td>
<td>Leave-one-out cross-validation</td>
<td>1.6805e+004</td>
<td>( \gamma [0, 2000] ) ( \delta [0, 2000] ) Search step; 10</td>
</tr>
<tr>
<td>GA-LS-SVM (Genetic Algorithm)</td>
<td>Leave-one-out cross-validation</td>
<td>735.1400</td>
<td>( \gamma [0, 2000] ) ( \delta [0, 2000] ) Number of the initial population; 30</td>
</tr>
<tr>
<td>PSO-LS-SVM (Particle Swarm Optimization)</td>
<td>Leave-one-out cross-validation</td>
<td>2.7828e+003</td>
<td>( \gamma [0, 2000] ) ( \delta [0, 2000] ) Number of the initial population; 20 Maximum search speed; 100</td>
</tr>
</tbody>
</table>
5. Summary

This paper developed a machine vision system for rice appearance quality detection. After image acquiring, image pre-processing algorithm was employed to enhance the image quality. The characteristics of rice shape including area, perimeter, roughness and minimum enclosing rectangle were calculated by image analysis. Least Squares Support Vector Machines, Naive Bayes Classifier and Back Propagation Artificial Neural Network were used to realize classification. Meanwhile the robust of the three algorithms was tested by adding 10% and 20% noise respectively. The results show that, LS-SVM optimized by GA and LS-SVM optimized by PSO in this experiment have the best performance. The experiment showed that based on machine, the appearance quality of rice can be identified effectively by LS-SVM, and the recognition accuracy of 99.67% was achieved.

6. Acknowledgment

This research work is supported by Scientific Research Fund of Zhejiang Provincial Education Department (Grant No. Y201121959).

7. References


