

Huanjun LIU

Empty glass bottle inspection method based on fuzzy support vector machine neural network and machine vision

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2010

Abstract This paper develops a computerized empty glass bottle inspection method. Wavelet transform and morphologic methods were employed to extract features of the bottle body and the finish from images. Fuzzy support vector machine neural network was adopted as classifiers for the extracted features. Experimental results indicated that the accuracy rate can reach up to 97% by using the method developed to inspect empty glass bottles.

Keywords machine vision, support vector machine (SVM), neural network (NN), morphologic method, wavelet transform

1 Introduction

Large numbers and various shapes of glass bottles are used as containers for food and drink. For example, in China over 31 million tons of beer were produced in 2008 with the majority of beer products packed in glass bottles. To ensure the quality of the final products, it is necessary to check the empty glass bottles for cleanness and breakage before the products are canned. In many cases, this kind of work is performed manually. However, manual inspection is not only expensive and time-consuming, but it is also very difficult to guarantee the quality.

Machine vision inspection system has been successfully applied in the field of integrated circuits, fruit and food quality inspection, etc. [1–4]. It also offers certain solutions for bottle inspection, but this method can only be used to inspect the bottle finish. The method described in Ref. [5] gives much attention to cracks in the upper portion of glass bottles. The captured image is corrected by

adaptive gray correction and then translated into a binary image. The binary image is judged according to conditions. However, the crack problem is only one of the defects. The inspection precision of other defects is not desirable using this method. Reference [6] proposes a method to inspect defects for empty water bottles. The defects are detected based on the variations of intensity of the images within the image segment. In this method, the decision (if the bottles are defective) is made solely dependent on pixel intensity, which is heavily influenced by noise.

Empty bottle inspection is one of the typical machine vision inspection problems. The difficulties in this inspection lie in its diversity, uncertainty and noise. The properties of body defects are diversified in position, shape and optic performance. Therefore, the features of defects in the captured image are uncertain and hardly distinguished. There are noises in bottle images. And some shadows exist in the bottle image since one cannot expect to get the same bottle optic performance. Because of the influence of environmental factors like ambient light, noises would exist in bottle images. All these difficulties make bottle inspection one of the difficult problems in intelligent inspection. For these reasons, we have developed a new machine vision method that can be used to inspect the bottle body and the finish. The system takes image photos of empty glass bottles by a digital camera, extracts and analyzes the features of the images to determine the cleanness and breakage of the empty bottles using a fuzzy support vector machine neural network.

2 Illumination system

A plate light emitting diode (LED) light is used when the camera takes a picture of the glass bottle body. The empty bottle is placed between the light and the camera, constituting the transmission-illumination relationship, shown in Fig. 1. In this case, breakages and stains on the bottle body can be clearly displayed, which is beneficial to the next step of bottle inspection. Two cameras capture

Received October 19, 2009; accepted May 21, 2010

Huanjun LIU (✉)
Faculty of Automation, Guangdong University of Technology,
Guangzhou 510006, China
E-mail: hjlju74@163.com

images of the bottle body respectively from the back and face to avoid omitting defects.

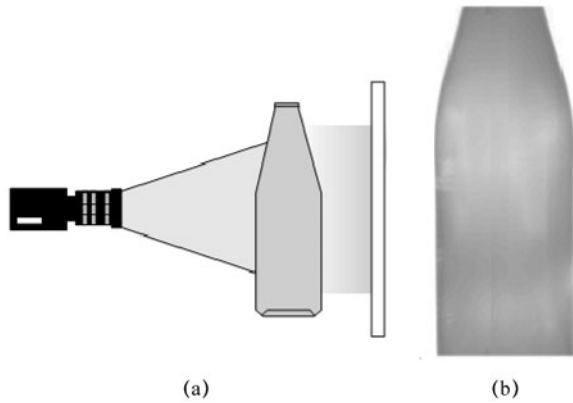


Fig. 1 Illumination of bottle body. (a) Illumination; (b) captured image

In addition, to obtain a clear image of the bottle finish, an LED light in an umbrella shape is used, shown in Fig. 2. The breakages or stains of the bottle finish can be detected by comparison of the intensity of brightness, in which the areas of breakages or stains are darker.

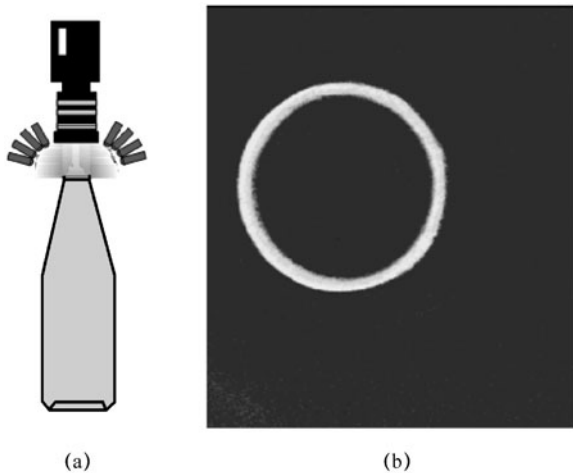


Fig. 2 Illumination of bottle finish. (a) Illumination; (b) captured image

3 Features of bottle

3.1 Marking and locating regions of interest

Since not all the information on the images of the bottles is used for further processing, it is necessary to mark the region of interest (ROI) manually in advance to reduce the processing cost. Because the defects could be everywhere, the shape of ROI is decided mainly according to the shape of the bottle. In Fig. 3, the regions of interest are marked

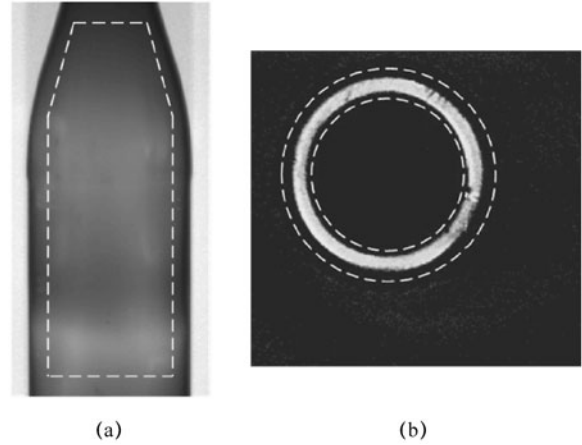


Fig. 3 Regions of interest. (a) ROI of bottle body; (b) ROI of bottle finish

with a dotted line. The computer only processes the image data within the region of interest.

3.2 Features of bottle body

3.2.1 Possible defective region identification

The body defect in the captured image is in the dark region. To label these regions, watershed transform is used. Watershed transform is a popular segmentation method coming from the field of mathematical morphology [7]. The intuitive description of this transform is quite simple: if the image is considered as a topographic relief, where the height of each point is directly related to its gray level and rain is considered gradually falling on the terrain, and then the watersheds are the lines that separate the “lakes” (actually called catchment basins) that form. Generally, watershed transform is computed on the gradient of the original image, so that the catchment basin boundaries are located at high gradient points. However, the traditional watershed transform generally leads to over-segmentation due to noise and other local irregularities of the gradient. To avoid this problem, this paper introduces some prior information to improve watershed transform.

The gradient of the image is calculated by morphology. The morphologic gradient can depend less on edge directionality [8].

The morphologic gradient of the image is computed by dilation and erosion [9]:

$$g_i = (f_i \oplus b_i) - (f_i \ominus b_i), \quad (1)$$

where $f_i \oplus b_i$ is the gray-scale dilation of f_i by b_i , and $f_i \ominus b_i$ is the erosion.

The edge is a set of points lying on the boundary between two regions. Though the edge cannot fully describe the boundary, it can show the information of the

defect region and background. Therefore, the edge is used to modify the gradient of the image. According to the characteristics of the defective region, the Sobel edge detection is selected. Modified gradient is described as follows.

$$gd(x,y) = \begin{cases} gi(x,y) + C, & \text{if there is edge point in } NB(x,y), \\ gi(x,y), & \text{else,} \end{cases} \quad (2)$$

where $NB(x, y)$ is the 3×3 neighborhood of the point (x, y) , and C is the constant.

The defective region of the bottle wall is the dark region, so the gray level of these regions is relatively low. Hence, the regional minima of the image should be in the object region. Regional minima are connected components of pixels with the same intensity value, t_m , whose external boundary pixels all have a value greater than t_m . This paper uses the regional minima as the markers.

The images of the bottle wall are segmented by the modified watershed transform, and the results are like Fig. 4.

The modified watershed transform can segment the possible defective regions, and reduce over-segmentation.

3.2.2 Features

The defective region is in dark, so the average gray level of the whole bottle body and the identified regions are

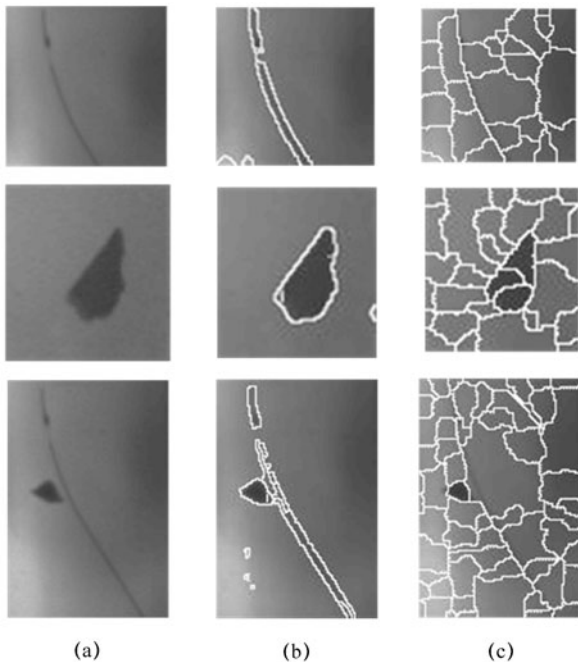


Fig. 4 Results of watershed transform. (a) Images of bottle wall; (b) results of modified watershed transform; (c) results of classic watershed transform

calculated. If the mean gray level in some regions is below the whole bottle body's, these regions may be defective.

Some features are extracted in these regions:

1) Feature 1:

The area of possible defective regions is denoted as $F_b(1)$:

$$F_b(1) = N_d, \quad (3)$$

where N_d is the number of possible defective regions.

If there are more possible defective regions with low mean gray level, the body might not be good.

2) Feature 2:

$$F_b(2) = \sum_{n=1}^{N_d} A_n, \quad (4)$$

where A_n is the area of the possible defective region.

This feature indicates the area of all possible defective regions. If this feature is big, the bottle body may be defective.

3) Feature 3:

$$F_b(3) = A_m, \quad (5)$$

where A_m is the maximum area in all possible defective regions.

4) Feature 4:

$$F_b(4) = \bar{G}_m, \quad (6)$$

where \bar{G}_m is the mean gray level in a region, of which the area is the maximum in all possible defective regions.

The gray level of the defective region is low. If the mean gray level of a region is low, this region may be a defective one.

Feature 4 and the two following features show the characteristics of a possible defective region, of which the area is the largest. This region is one of the most possible defective regions. The size of the defective region is not too small; thus, if Feature 4 of a region is large, this region may be a defective one.

5) Feature 5:

$$F_b(5) = \sum_{j=\bar{G}_m-1}^{\bar{G}_m+1} P_m(j), \quad (7)$$

where $P_m(j)$ is the probability density function of the gray level j in a region, of which the area is the maximum.

The gray level of the pixel in defective region is not different. Therefore, if this feature of a region is bigger, this region may be the defective one.

6) Feature 6:

$$F_b(6) = A_g, \quad (8)$$

where A_g is the area of a region, in which the mean gray level is the maximum in all possible defective regions.

When the mean gray level of a region is the maximum, it is another possible defective region. This feature and the two following features show the characteristics of this region. Also, they have the same meaning as Features 3–5.

7) Feature 7:

$$F_b(7) = \overline{G}_g, \tag{9}$$

where \overline{G}_g is the mean gray level in a region, in which the mean gray level is the maximum.

8) Feature 8:

$$F_b(8) = \sum_{j=\overline{G}_g-1}^{\overline{G}_g+1} P_g(j), \tag{10}$$

where $P_g(j)$ is the probability density function of gray level j in a region, where the mean gray level is the maximum.

3.3 Features of bottle finish

3.3.1 Scanning methods

While extracting features of the bottle finish, in view of the bottle finish shape, the circular law is used to carry out scanning. In scanning, the bottle finish center is taken as the center of a circle; each point is scanned by changing the radius and central angle. Because the round size of a real glass bottle finish ring varies in reality, the round width of the ring in the finish image can be also varied. Hence, the scope of the ring's radius is given in advance. The ring's radius of the normal bottle finish lies within this range. The scanning point is obtained by

$$\begin{cases} x = x_C + r\cos\theta, \\ y = y_C - r\sin\theta, \end{cases} \tag{11}$$

where (x_C, y_C) is the center of the bottle finish, the scope of the r is (r_1, r_2) , and θ ranges from 0° to 359° .

In scanning, the average gray level of the central angle θ is calculated by

$$L(\theta) = \frac{\sum_{r_1}^{r_2} g(x,y)}{r_2 - r_1}, \tag{12}$$

where $g(x, y)$ is the gray level of (x, y) , and (x, y) is decided by Eq. (11).

$L(\theta)$ of the bottle finish is shown in Fig. 5.

3.3.2 Features

When the quality of the bottle finish meets the standard, its image should be a ring with a consistent width and a smooth gray level. $L(\theta)$ will thus make little changes in different central angles. Otherwise, $L(\theta)$ would have significant changes. There are jagged noises and other information in the curve of $L(\theta)$. The multilevel approximation coefficients of one-dimensional (1D) wavelet transform [10] are used to reduce the noise. The Daubechies wavelet is chosen. The multilevel approximations of $L(\theta)$ are shown in Fig. 6.

Level 3 approximation coefficients not only keep the essential information but also reduce the noise. Therefore, they are chosen as the base of the features. These features are expressed as follows:

$$F_f(i) = \begin{cases} C_A(i), & \text{if } C_A(i) \text{ is an extreme,} \\ 0, & \text{others,} \end{cases} \tag{13}$$

where $C_A(i)$ are the level 3 approximation coefficients.

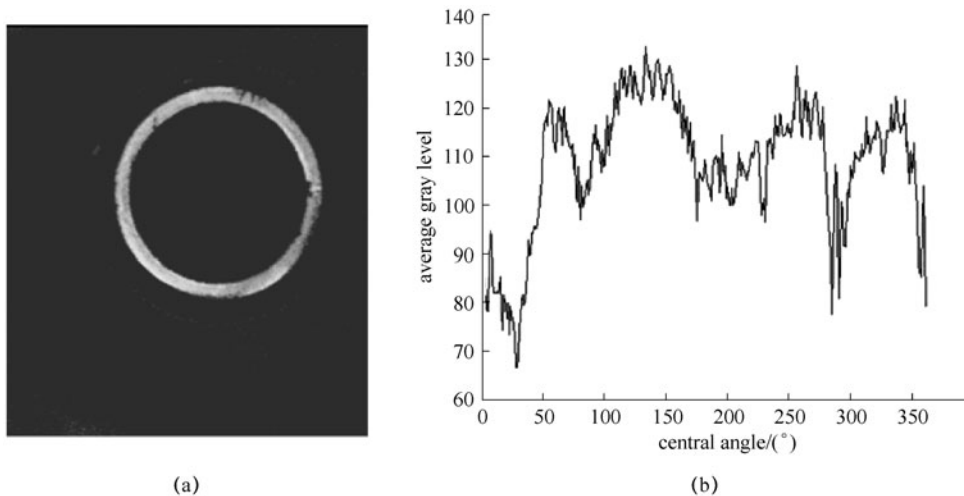


Fig. 5 $L(\theta)$ of bottle finish. (a) Bottle finish; (b) $L(\theta)$

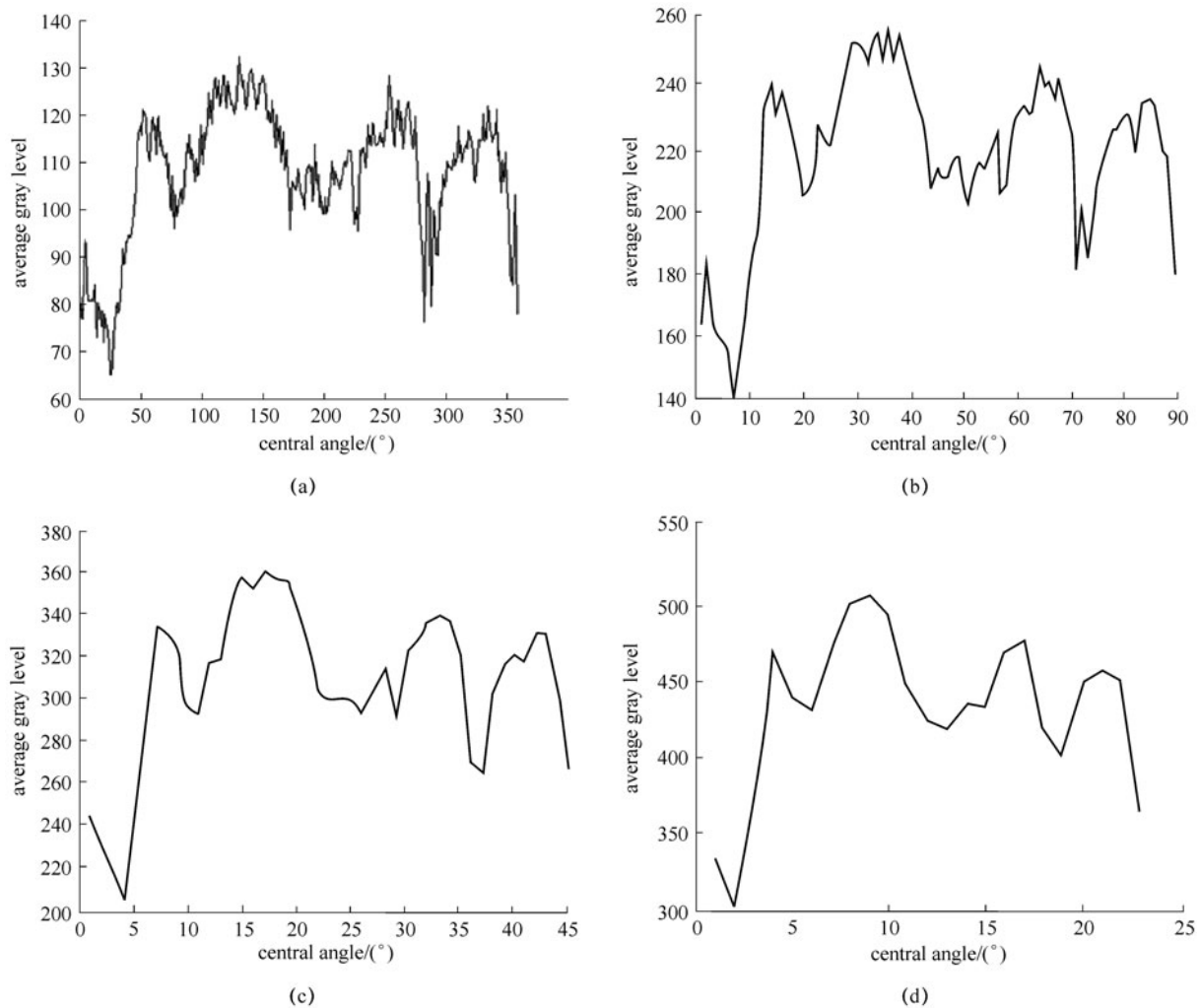


Fig. 6 Multilevel approximations. (a) Originality; (b) level 2 approximation; (c) level 3 approximation; (d) level 4 approximation

4 Classifier based on fuzzy support vector machine neural network

Bottle defects are diversified, so the bottle inspection is a small-sized sample problem. Also, the defects in the image are hardly described. To solve these problems, we propose a fuzzy support vector machine neural network.

Support vector machines (SVMs) are proposed initially in the field of machine learning, to classify problems on (typically large) sets of data having an unknown dependency on (possibly many) variables. SVMs are based on structural risk minimization methods, and produce a decision surface as the optimal hyperplane that separates the two classes with maximal margin [7,11]. SVMs have been used as one of the highest performance classifying systems because of their ability to generalize well. For SVM, some parameters, for example, kernel functions, need to be chosen. It is a difficult task to choose the parameters related to the model of the object. In

practice, the SVMs obtained from the learning are insufficient to completely classify all unknown samples. Also, SVMs cannot ensure to provide the global optimal classification performance over all samples.

Neural networks (NNs) have been exploited in many applications and a few learning algorithms have been developed. One of the main applications employed is data classification. Studies have reported that SVMs are generally able to deliver higher classification accuracy than other data classification algorithms [12,13]. Nevertheless, NNs can adjust the parameters more easily than SVMs.

Fuzzy theory uses fuzzy sets instead of normal sets. It can process the fuzzy information. Fuzzy theory simulates the way of human thinking. Its fault tolerance is good. A fuzzy support vector machine neural network is used in this study as the classifier. It combines fuzzy theory with SVMs. We use an optimization method adopted from a genetic crossbred algorithm to select the

primary parameters. Since SVM resembles NN in structure, back propagation (BP) learning method is adopted to optimize the parameters of the fuzzy support vector machine neural networks.

4.1 Fuzzy support vector machine (FSVM)

FSVM consists of fuzzy layer and SVMs. The structure is shown in Fig. 7.

Fuzzification is the function of the fuzzy layer. The features are inputted into the fuzzy layer, and translated into fuzzy outputs. This layer uses Gaussian function as the membership function. The function is as follows:

$$\mu_i(x_i) = e^{-\left(\frac{x_i - a_i}{b_i}\right)^2} \quad (14)$$

Then, SVMs are used as the classifier for fuzzy outputs.

Research shows that the use of the hybrid kernel yields a better performance than those with a single common kernel [14]. Hence, the hybrid kernel is applied in this study. The kernel function adopted in this paper is as follows:

$$K(x, x_i^*) = k_1(x \cdot x_i^*)^d + k_2 e^{-r|x - x_i^*|^2} \quad (15)$$

Genetic algorithms (GAs) constitute the global optimization techniques known to be successful in many domains. Thus, a GA based selection of components for FSVM is proposed in this study. This method is employed to first optimize FSVM. The accuracy of classification and the risk of classifier are often used to evaluate the performance of classification. But the traditional GA only performs optimization process according to one goal. This paper adopts a crossbred genetic algorithm, which simulates crossbreeding in biology. There are two different fitness functions, with which the individuals are selected for breeding. So this algorithm suits the optimization of FSVMs.

The flow chart of this algorithm is shown in Fig. 8.

By this method, chromosomes are encoded as the real

number. The structure of chromosomes is as follows:

$$(GF_{a1}, GF_{b1}, \dots, GF_{am}, GF_{bm}, KP_1, KP_2, \dots, KP_n), \quad (16)$$

where GF_{ak} and GF_{bk} , $k = 1, 2, \dots, m$, are the parameters of Gaussian functions in the fuzzy layer, KP_k , $k = 1, 2, \dots, n$, are parameters of the kernel function.

The initial population is randomly crafted in different regions. There are two fitness functions, with which the individuals are selected for breeding. The two fitness functions are produced according to the accuracy of classification and the risk of the classifier. One fitness function is given by

$$Fit_{k1} = \frac{1}{1 - C_k}, \quad (17)$$

where C_k , $k = 1, 2, \dots, n$, is the accuracy of classification. This accuracy is obtained by u -fold cross-validation. In u -fold cross-validation, the training sets are first divided into u subsets of equal size. Sequentially, one subset is tested using the classifier trained on the remaining $u-1$ subsets.

The risk can be estimated by Vapnik and Chervonenkis (VC) dimension, which is hard to calculate directly. Therefore, RT is employed to denote the leave-one-out bound [11].

$$RT = \frac{N_{sv}}{l}, \quad (18)$$

where l is the number of training data set, and N_{sv} denotes the number of support vectors.

Another fitness function is given by

$$Fit_{k2} = \frac{5}{RT_k}. \quad (19)$$

In reproduction, one half is selected according to fitness function Fit_{k1} , and the other half is selected according to fitness function Fit_{k2} . The elitist selection (10%) and roulette wheel selection operators are employed for reproduction.

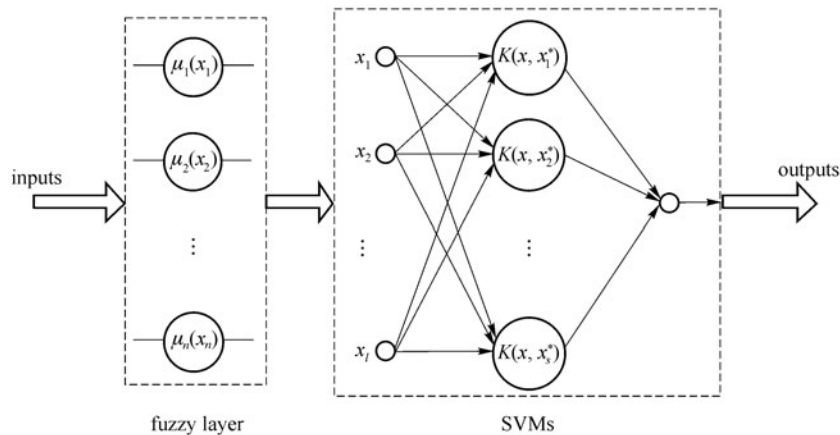


Fig. 7 Structure of FSVM

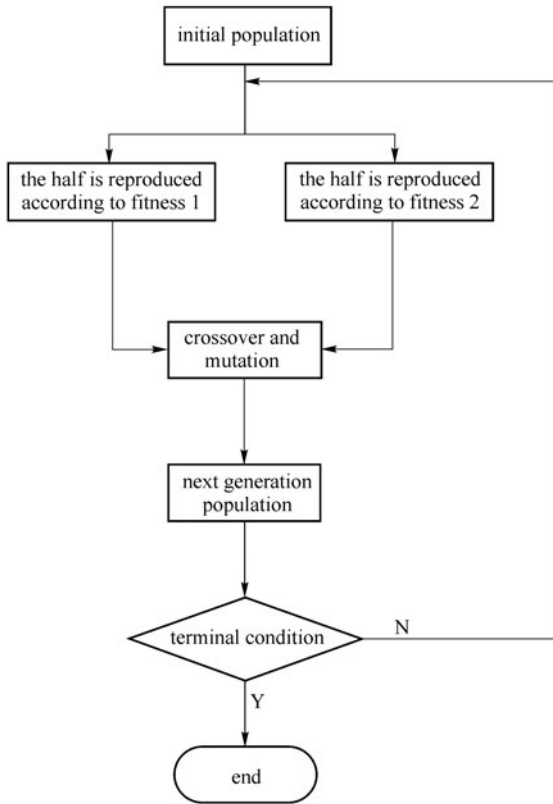


Fig. 8 Flow chart of proposed algorithm

The crossover operator is as follows. For two chromosomes $A_i = (a_1, a_2, \dots, a_n)$ and $B_i = (b_1, b_2, \dots, b_n)$, the chromosomes after crossover are $A'_i = (a'_1, a'_2, \dots, a'_n)$ and $B'_i = (b'_1, b'_2, \dots, b'_n)$, where

$$a'_i = \beta_i a_i + (1 - \beta_i) b_i, \quad (20)$$

$$b'_i = \beta_i b_i + (1 - \beta_i) a_i, \quad (21)$$

where β_i is a random number in $[0, 1]$.

The mutation operator is given by

$$a'_i = \begin{cases} a_i + f(t, a_{i\max} - a_i), & \text{rad} = 0, \\ a_i - f(t, a_i - a_{i\min}), & \text{rad} = 1, \end{cases} \quad (22)$$

where rad is a random number, and

$$f(t, y) = y \left(1 - r \left(1 - \frac{t}{T} \right)^2 \right), \quad (23)$$

where t is the number of current generation, T is the maximum generation, r is a random number in $[0, 1]$.

The probabilities of crossover and mutation are decided adaptively, that is to say, these probabilities relate to the situation of evolution.

The probability of crossover is calculated by

$$P_C = \begin{cases} \frac{f_{\max} - f_b}{f_{\max} - \bar{f}}, & f_b > \bar{f}, \\ 0.8, & f_b \leq \bar{f}, \end{cases} \quad (24)$$

where f is the hybrid fitness, which is given by

$$f = 0.6\text{Fit}_{k1} + 0.4\text{Fit}_{k2}, \quad (25)$$

and f_b is the bigger f of the two chromosomes, f_{\max} is the maximal f in the population, \bar{f} is the mean f of the population.

The probability of mutation is

$$P_M = \begin{cases} \frac{0.5(f_{\max} - f)}{f_{\max} - \bar{f}}, & f > \bar{f}, \\ 0.5, & f \leq \bar{f}, \end{cases} \quad (26)$$

where f is the hybrid fitness of the chromosomes to be mutated, and is calculated by Eq. (25), f_{\max} and \bar{f} are the same as Eq. (24).

The procedure would not be terminated until both the changes of average fitness 1 and average fitness 2 between two neighboring populations are less than the thresholds.

4.2 Fuzzy neural networks

The best parameters cannot be obtained only by GA. Therefore, the NNs learning methods were used in our study to optimize the parameters on the bases of the result of crossbred genetic algorithm. Since SVM is reported to be similar to NN [15,16], it is logical to consider FSVM as the fuzzy neural network. The structure of fuzzy neural networks is shown in Fig. 9.

Layer 1 is the fuzzy layer, of which the function is to transform input vectors to fuzzy variables. The relations of the input with the output are as follows:

$$\mu_i(x_i) = e^{-\left(\frac{x_i - a_i}{b_i}\right)^2}, \quad (27)$$

$$I_i^{(1)} = x_i, \quad (28)$$

$$O_i^{(1)} = \mu_i(I_i^{(1)}), \quad (29)$$

where $i = 1, 2, \dots, n$, n is the number of input vectors.

Layer 2 is the middle layer.

$$I_i^{(2)} = O_i^{(1)}, \quad (30)$$

$$O_i^{(2)} = I_i^{(2)}. \quad (31)$$

Layer 3 is the kernel function layer; the relations between the input and the output are as follows:

$$I^{(3)} = O^{(2)}, \quad (32)$$

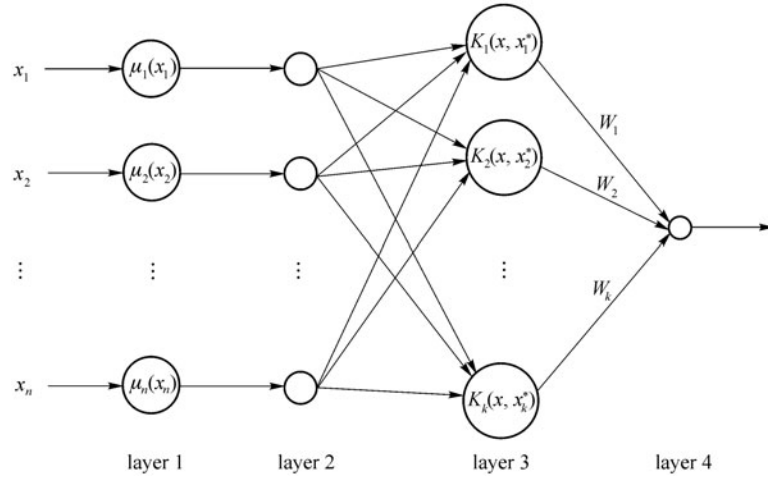


Fig. 9 Fuzzy neural network

$$O_j^{(3)} = K_j(I^{(3)}, x_j^*), \quad (33)$$

$$K_j(x, x_j^*) = k_{1j}(x \cdot x_j^*)^{d_j} + k_{2j}e^{-r_j|x-x_j^*|^2}, \quad (34)$$

where $j = 1, 2, \dots, k$, k is the number of the support vectors in the FSVM, $K(x, x_j^*)$ is the kernel function.

Layer 4 is the output layer; the relations between the input and the output are as follows:

$$I^{(4)} = \sum_{j=1}^k O_j^{(3)} \cdot W_j, \quad (35)$$

$$Y = O^{(4)} = I^{(4)}. \quad (36)$$

The learning algorithm of networks is adopted in a gradient descent way. The initial parameters are given according to the FSVM trained. Therefore, x_j^* are the support vectors, W_j are a_j which are the Lagrange multipliers corresponding to support vectors in FSVM, and a_i and b_i are the same as the fuzzy layer in FSVM. The detailed learning algorithm is introduced in the Appendix.

5 Procedures of inspection methods

5.1 Procedure of bottle body inspection

The procedure of bottle body inspection methods in this paper are summarized as follows.

Step 1 The regions of interest are marked as Fig. 3(a).

Step 2 The possible defective regions in the bottle body images are segmented by the modified watershed transform.

Step 3 The eight features of bottle body are extracted by Eqs. (3)–(10) in these possible defective regions.

Step 4 The FSVM is constructed, and the parameters are optimized by the crossbred genetic algorithm.

Step 5 This FSVM is taken as the NN, and the initial parameters of the NNs are decided by this FSVM.

Step 6 This NN is learned by gradient descent methods.

Step 7 Finally, this NN is used to classify the features, and the inspection results of the bottle body are obtained.

5.2 Procedure of bottle finish inspection

Step 1 The regions of interest are marked as Fig. 3(b).

Step 2 The circular law is performed to scan, the curve $L(\theta)$ is obtained by Eq. (12).

Step 3 The features of the bottle finish are calculated according to Eq. (13).

Steps 4–6 They are the same as the bottle body inspection.

Step 7 The features of the bottle finish are classified by the fuzzy support vector machine neural network.

6 Experiments

Experiments were conducted to examine this method. A prototype equipped with an annular conveyor was developed according to Sect. 2. 500 images of each bottle finish and bottle body have been captured separately using this prototype. The bottle inspection method developed in this study was programmed by C language on a P4 2.6.

First, the images of the bottle finish and bottle body were compared to the real empty bottles to identify the empty bottles. Then the features from the 300 images were extracted according to the rules described in Sect. 3, which would be used as a base for learning the fuzzy support vector machine neural networks. After the fuzzy support vector machine neural networks were trained, they started

Table 1 Comparison of accuracy of detecting glass bottle body using three methods*

samples	accuracy/%		
	SVM	NN	fuzzy support vector machine neural network
good (112)	95.0	94.8	97.0
bad (88)	96.8	95.0	97.7

* The bottle body inspection speed is about 674 bottles per minute. The defect larger than 1.5 mm² can be detected by the method proposed in this paper.

Table 2 Accuracy comparison of detected glass bottle finish using three methods**

samples	accuracy/%		
	SVM	NN	fuzzy support vector machine neural network
good (121)	96.5	96.3	97.7
bad (79)	97.4	97.3	98.5

** The inspection speed is about 645 bottles per minute. The inspection method proposed in this paper can detect the defect that is larger than 0.5 mm² on the bottle finish.

to judge the features extracted from the other images. In crossbred genetic algorithm, the initial population is randomly crafted in ten regions. The number of the population is 20. The terminated threshold is 0.5. In training fuzzy neural networks, the learning rate is 0.01. In comparison with the fuzzy support vector machine neural network, the SVM and NNs were also applied. The RBF kernel function was selected for the SVM. The function is as follows:

$$K_{\text{RBF}}(x, x_i^*) = e^{-r|x-x_i^*|^2}. \quad (37)$$

According to experience, r is set as 0.5. The number of neurons in the hidden layer of neural networks is the number of neurons in the input layer adding one. The BP method is applied to train neural networks, and the learning rate is 0.4. The means were calculated based on the experiments repeated ten times and shown in Tables 1 and 2.

Experiment results indicated that the fuzzy support vector machine neural network is superior to the support vector machine and the neural network.

7 Conclusion

This study offers an intelligent inspection method to examine empty glass bottles. Clear images of the bottle finish and bottle body were obtained through a specially-designed machine vision system. The regions of interest in the image were marked and located. The possible defective regions of the bottle body were labeled by morphologic methods, and the features were summarized after comparing with the real glass bottle defects. The bottle finish features were extracted by the method based on 1D wavelet transform, which can reduce the noise and features number. Then these features were classified by fuzzy support vector machine neural networks. The fuzzy support vector machine neural networks combined the fuzzy theory with the SVMs, and utilized an optimization

method based on a genetic crossbred algorithm to choose the parameters primarily. Then the BP learning method is adopted to optimize the parameters of the fuzzy support vector machine neural networks. As a result, the classifier possesses not only good generalization ability but also strong anti-noise ability. The performance of the fuzzy support vector machine neural networks is better than SVM and BP NN, and it is adaptable to the bottle inspection, often with a problem of uncertainty and noise. Experiments have been performed on the prototype to prove whether the methods can work effectively. The results showed that the method developed in this study was able to detect defective glass bottles with an accuracy rate up to 97%.

Appendix

The learning algorithm of fuzzy neural networks is as follows.

The learning errors of networks are defined as follows:

$$J = \frac{1}{2}(D - Y)^T(D - Y), \quad (A.1)$$

where D is the anticipant output. If the empty bottle is good, D is defined as 0.8; otherwise, D is defined as -0.8 .

The revised formula of network's weight W_j is

$$\Delta W_j = -\frac{\partial J}{\partial W_j} = (D - Y)O_j^{(3)}, \quad (A.2)$$

where $j = 1, 2, \dots, k$, and k is the number of support vectors.

$$W_j(t + 1) = W_j(t) + \eta \Delta W_j, \quad (A.3)$$

where η is the learning rate.

The revised equations of the factors in wavelet basis function are

$$\Delta k_{1j} = -\frac{\partial J}{\partial k_{1j}} = (D - Y)W_j \left(I^{(3)} x_j^* \right)^{d_j}, \quad (A.4)$$

$$\Delta k_{2j} = -\frac{\partial J}{\partial k_{2j}} = (D-Y)W_j e^{-r_j |I^{(3)} - x_j^*|^2}, \quad (\text{A.5})$$

$$k_{1j}(t+1) = k_{1j}(t) + \eta \Delta k_{1j}, \quad (\text{A.6})$$

$$k_{2j}(t+1) = k_{2j}(t) + \eta \Delta k_{2j}, \quad (\text{A.7})$$

$$\begin{aligned} \Delta d_j &= -\frac{\partial J}{\partial d_j} \\ &= (D-Y)W_j \left[k_{1j} \left(I^{(3)} x_j^* \right)^{d_j} \ln \left(I^{(3)} x_j^* \right) \right], \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned} \Delta r_j &= -\frac{\partial J}{\partial r_j} \\ &= -(D-Y)W_j \left(k_{2j} e^{-r_j |I^{(3)} - x_j^*|^2} \left| I^{(2)} - x_j^* \right|^2 \right), \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} \Delta x_j^* &= -\frac{\partial J}{\partial x_j^*} \\ &= (D-Y)W_j \left[k_{1j} d_j \left(I^{(3)} x_j^* \right)^{d_j-1} I^{(3)} \right. \\ &\quad \left. + k_{2j} e^{-r_j |I^{(3)} - x_j^*|^2} \left(2r_j \left| I^{(3)} - x_j^* \right| \right) \right], \end{aligned} \quad (\text{A.10})$$

$$d_j(t+1) = d_j(t) + \eta \Delta d_j, \quad (\text{A.11})$$

$$r_j(t+1) = r_j(t) + \eta \Delta r_j, \quad (\text{A.12})$$

$$x_j^*(t+1) = x_j^*(t) + \eta \Delta x_j^*, \quad (\text{A.13})$$

$$\begin{aligned} \Delta a_i &= -\frac{\partial J}{\partial a_i} \\ &= \sum_{j=1}^k (D-Y)W_j \left[k_{1j} d_j \left(I^{(3)} x_j^* \right)^{d_j-1} x_j^* \right. \\ &\quad \left. + k_{2j} e^{-r_j |I^{(3)} - x_j^*|^2} \left(-2r_j \left| I^{(3)} - x_j^* \right| \right) \right] \\ &\quad \cdot 2e^{-\left(\frac{I_i^{(1)} - a_i}{b_i} \right)^2} \frac{I_i^{(1)} - a_i}{b_i^2}, \end{aligned} \quad (\text{A.14})$$

$$\begin{aligned} \Delta b_i &= -\frac{\partial J}{\partial b_i} \\ &= \sum_{j=1}^k (D-Y)W_j \left[k_{1j} d_j \left(I^{(3)} x_j^* \right)^{d_j-1} x_j^* \right. \\ &\quad \left. + k_{2j} e^{-r_j |I^{(3)} - x_j^*|^2} \left(-2r_j \left| I^{(3)} - x_j^* \right| \right) \right] \end{aligned}$$

$$\cdot 2e^{-\left(\frac{I_i^{(1)} - a_i}{b_i} \right)^2} \frac{\left(I_i^{(1)} - a_i \right)^2}{b_i^3}, \quad (\text{A.15})$$

where $i = 1, 2, \dots, n$, n is the number of input vector.

$$a_i(t+1) = a_i(t) + \eta \Delta a_i, \quad (\text{A.16})$$

$$b_i(t+1) = b_i(t) + \eta \Delta b_i. \quad (\text{A.17})$$

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. U0735003, 60604006), the Ph.D. Programs Foundation of Ministry of Education of China (20070562005), and the National Natural Science Foundation of Guangdong province of China (Nos. 07117423, 6021452).

References

1. Aleixos N, Blasco J, Molto E, Navarron F. Assessment of citrus fruit quality using a real-time machine vision system. In: Proceedings of the 15th International Conference on Pattern Recognition. 2000, 1: 482–485
2. Du C J, Sun D W. Comparison of three methods for classification of pizza topping using different colour space transformations. Journal of Food Engineering, 2005, 68(3): 277–287
3. Shankar N G, Zhong Z W. Defect detection on semiconductor wafer surfaces. Microelectronic Engineering, 2005, 77(3–4): 337–346
4. Jiang B C, Tasi S L, Wang C C. Machine vision-based gray relational theory applied to IC marking inspection. IEEE Transactions on Semiconductor Manufacturing, 2002, 15(4): 531–539
5. Ma H M, Su G D, Wang J Y, Ni Z. A glass bottle defect detection system without touching. In: Proceedings of the First International Conference on Machine Learning and Cybernetics. 2002, 2: 628–632
6. Shafait F, Imran S M, Klette-Matzat S. Fault detection and localization in empty water bottles through machine vision. In: Proceedings of E-Tech 2004. 2004, 30–34
7. Vapnik V N. Statistical Learning Theory. New York: John Wiley & Sons Inc, 1998
8. Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62–66
9. Gonzalez R C, Woods R E. Digital Image Processing. 2nd ed. Upper Saddle River: Prentice Hall, 2002
10. Mallet S. A Wavelet Tour of Signal Processing. 2nd ed. New York: Academic Press, 1999
11. Vapnik V, Levin E, Le Cun Y. Measuring the VC-dimension of a learning machine. Neural Computation, 1994, 6(5): 851–876
12. Dumais S, Platt J, Heckerman D, Sahami M. Inductive learning algorithms and representations for text categorization. In: Proceedings of the Seventh International Conference on Information and Knowledge Management. 1998, 148–155
13. Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks, 2002, 13(2): 415–425
14. Tan Y, Wang J. A support vector machine with a hybrid kernel and

- minimal Vapnik-Chervonenkis dimension. IEEE Transactions on Knowledge and Data Engineering, 2004, 16(4): 385–395
15. Smola A J, Schölkopf B, Müller K R. The connection between regularization operators and support vector kernels. Neural Networks, 1998, 11(4): 637–649
16. András P. The equivalence of support vector machine and regularization neural networks. Neural Processing Letters, 2002, 15(2): 97–104