Classification of Solder Joint Using Feature Selection Based on Bayes and Support Vector Machine

Hao Wu, Xianmin Zhang, Hongwei Xie, Yongcong Kuang, and Gaofei Ouyang

***Abstract*— In this paper, a feature selection and a two-stage classifier for solder joint inspection have been proposed. Using a three-color (red, green, and blue) hemispherical light-emitting diode array illumination and a charge-coupled device color digital camera, images of solder joints can be obtained. The color features, including the average gray level and the percentage of highlights and template-matching feature, are extracted. After feature selection, based on the algorithm of Bayes, each solder joint is classified by its qualification. If the solder joint fails in the qualification test, it is classified into one of the pre-defined types based on support vector machine. The choice of the second stage classifier is based on the performance evaluation of various classifiers. The proposed inspection system has been implemented and tested with various types of solder joints in surface-mounted devices. The experimental results showed that the proposed scheme is not only more efficient, but also increases the recognition rate, because it reduces the number of needed extracted features.**

***Index Terms*— Bayesian classifier, feature selection, solder joint, support vector machine (SVM).**

I. Introduction

AUTOMATED optical inspection (AOI) systems are widely used in industrial applications [1]. In surface mount technology (SMT) production, the AOI system is used to detect surface-related defects such as pseudo-solder, solder insufficient, component shifted, wrong component, and tombstone, *et al*.

Besl and Jain [2] conducted one of the earliest studies on solder joint inspection. In their work, diffusive light was used to avoid the images being saturated by a specular reflection of the solder surface. The features of solders were extracted from the low-contrast gray images [3], [4]. However, the results were discouraging due to their sensitivity to illumination conditions.

Kim *et al*. [5] used a three-layer ring-shaped light-emitting diode (LED) illumination and a charge-coupled device (CCD) camera to take gray-level images. The classification was divided into two stages. The characteristic 2-D features with simpler computation were extracted and input to a back propagation neural network for classification at the first stage. If the output value was not in the confidence, a computationally expensive algorithm based on the 3-D features with a Bayesian classifier would be used. The experimental results showed that the method had a good performance in both speed and recognition rate.

Giuseppe *et al*. [6] presented a neural network-based automatic optical inspection system for the diagnosis of solder joint defects. Five types of solder joints have been classified in respect to the amount of solder paste to perform the diagnosis with high recognition, 10 geometric features and eight wavelet features were extracted. The results have proved that the multilayer perceptron network fed with the GW-features had the best recognition rate.

Lin *et al*. [7] proposed a two-step inspection scheme. In this scheme the inspection system was divided into the screen stage and the classification stage. The screen stage was to quickly screen out most normal components to reduce overall processing time. Only one image feature was used as the screen index. At the classification stage, the neural networks were adopted to integrate all image feature information available to more precisely classify those failing to pass the screening test.

Yun *et al*. [8] proposed a method for inspecting solder joint using support vector machine (SVM). The characteristic features including the average intensity value and percentage of highlights are used to classify the solder joint. Experimental results showed the accuracy and effectiveness of the proposed method.

Although the inspection methods mentioned above with new features are efficient in the experiment, all these inspection method do not use any feature selection method to reduce feature redundancy. Also, the types of solder joints can only focus on the volume of soldering such as good solder, insufficient solder, excess solder, and no solder. Defects such as pseudo solder, tombstone, component shift, and wrong component are quite common in the industrial field. So this paper focuses on five most common defects (pseudo solder, solder insufficient, component shifted, wrong component, and

Manuscript received April 5, 2012; revised November 25, 2012; accepted November 29, 2012. Date of publication January 9, 2013; date of current version February 21, 2013. This work was supported in part by the National Science Foundation for distinguished young scholars of China under Grant 50825504, the United Fund of Natural Science Foundation of China and Guangdong province under Grant U0934004, the Fundamental Research Funds for the Central Universities under Grant 2012ZP0004, and Project GDUPS (2010). Recommended for publication by Associate Editor S. J. Mason upon evaluation of reviewers' comments.

The authors are with School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: wuhaomoses@gmail.com; zhangxm@scut.edu.cn; xhw_cn@foxmail.com; yckuang@scut.edu.cn; scut_gaofei@yahoo.com.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCPMT.2012.2231902

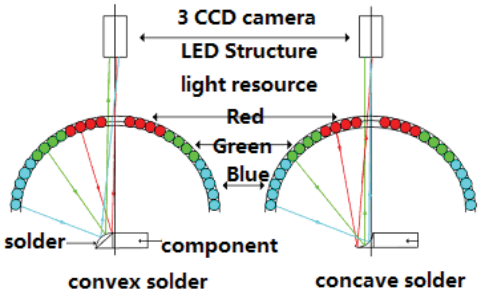

Fig. 1. Three-color LED array illumination system.

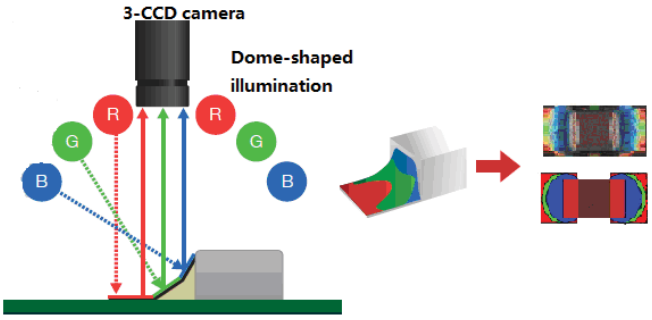

Fig. 2. Image acquisition system.

tombstone) of solder joints in industry. The classification is divided into two stages: In the first stage, the Bayesian classifier classified the solder joint into qualified or not; in the second stage, the SVM classifier classified the solder joint into multiclass in detail. Fourteen features were extracted as the input of the feature selection and classifier based on the proposed color feature and the template matching feature. This proposed scheme with feature selection is more efficient because it reduces the number of needed extracted features. The classification recognition rate also increased with this added feature selection process.

This paper is organized as follows. In Section II the illumination system and the image of solder joint with different defects are presented. The methodology of this paper and feature extraction and selection are presented in Section III. In Section IV the two stages of classification are described. Experimental results are obtained with the proposed method using a set of solder images in Section V. Finally, the conclusions are summarized at the end of this paper.

II. Image Acquirement and Solder Class Definition

The images are obtained by a CCD color digital camera and a three-color (red, green, and blue) hemispherical LED array illumination, which is shown in Fig. 1.

The surface of the solder joint holds the same property of flat mirror, which abides with the law of reflection. The red, green, and blue lights irradiate to the flat, the slow slant, and the rapid slant of the solder joint surfaces, which are reflected to the camera, respectively [9]. With such a three-color illumination and color camera system, the 3-D shape information of the solder joints can be depicted by a 2-D color image. The image acquisition process is shown in Fig. 2.

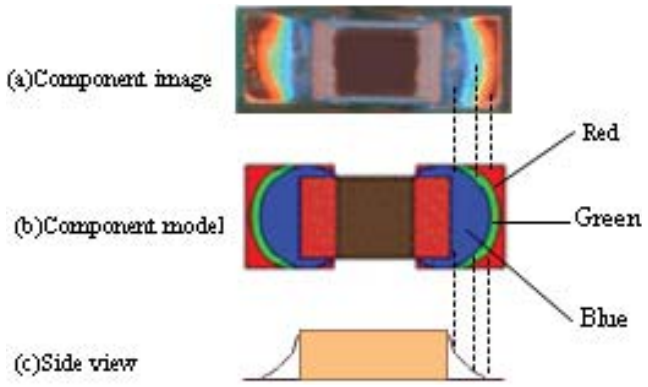

Fig. 3. Solder joint image and model.

TABLE I
Solders With Different Defects

Type	Image	Solder Model	Side View
Good solder			
Pseudo solder			
Solder insufficient			
Component shifted			
Wrong component			
Tombstone			

Fig. 3 shows the image of the solder joint and the solder model used in this paper, the types of solder joint to be inspected include good solder (Gs), pseudo solder (Cs), solder insufficient (Si), component shifted (Ct), wrong component (Wc), and tombstone (Tb). Typical images of various solder joints are shown in Table I.

III. Methodology Feature Extraction, and Selection

A. Methodology

The presented approach of this paper is shown in Fig. 4. First, the features of the input image are extracted, and the feature selection is used to select features. Then the Bayesian classifier classifies the solder joint qualified or not. In the following steps, the SVM classifier classified the solder joint into multiclass in detail.

B. Feature Extraction

In this section, two types of features were extracted from the image of solder joint: color features and template matching features. These features are described in detail below.

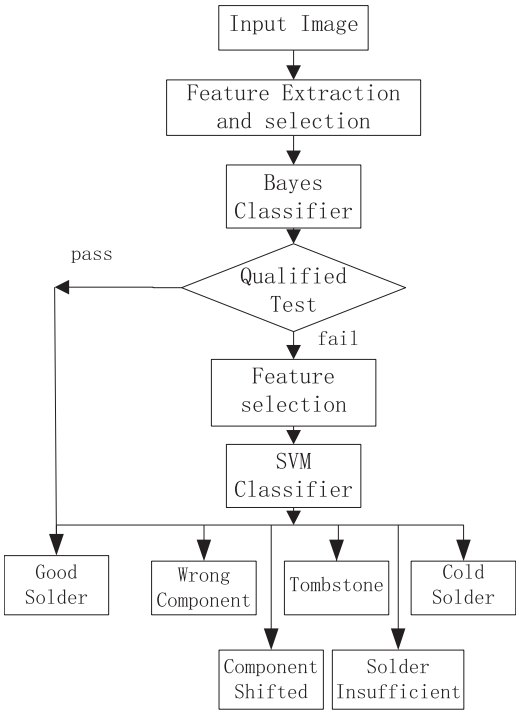

Fig. 4. Flow chart of the proposed inspection system.

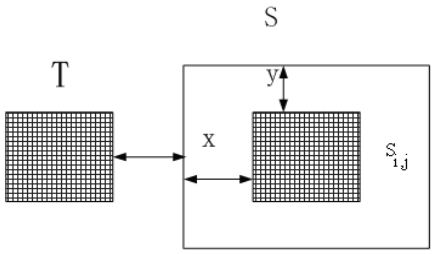

Fig. 5. Template and search range window.

1) Color Features: The color features include the average intensity value and percentage of highlights. The color features can be defined as follows:

$$\boldsymbol{X} = (X_1, X_2, X_3, X_4, X_5, X_6) \tag{1}$$

$$X_j = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} I_r(x, y) \tag{2}$$

$$X_k = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} T_r(x, y) \times 100 \tag{3}$$

where X is a six dimensional feature vector, subscripts j and k indicate the feature vector elements 1–6 respectively. $I_r(x, y)$ is the intensity of the image in the r color frame(1 = red, 2 = green, 3 = blue), $T_r(x, y)$ is the threshold value of the image in each color frame, and $M \times N$ is the number of pixels in the solder region.

2) Template Matching Feature: X_7 stands for the feature of the template match. The algorithm firstly cuts the interest area as a template and edits the template including hide the unwanted area, then uses the template to compare the image of the testing solder joint in the checking process. The comparison result is the extracted feature, which is shown in Fig. 5.

The template match method is comparing the template T and the search image S, and a similarity measure that achieves this is the normalized cross correlation (NCC) [10]

$$\begin{aligned} X_7 &= \mathrm{ncc}(x, y) \\ &= \frac{1}{n} \sum_{(u,v)\in T} \frac{t(u, v) - m_t}{\sqrt{\sigma_t^2}} \frac{s(x + u, y + v) - m_s(x, y)}{\sqrt{\sigma_s^2(x, y)}}. \end{aligned} \tag{4}$$

Here, m_t is the mean gray value of the template and σ_t^2 is the variance of the gray value, (x, y) is the current position of the template T in the search image S, (u, v) is the pixel position of the template T, $t(u, v)$ is the grey value of the template, $\mathrm{s}(x + u, y + v)$ is the grey value of the search image.

$$m_t = \frac{1}{n} \sum_{(u,v)\in T} t(u, v) \tag{5}$$

$$\sigma_t^2 = \frac{1}{n} \sum_{(u,v)\in T} (t(u, v) - m_t)^2. \tag{6}$$

Analogously, m_s (x, y) and (x, y) are the mean value and variance in the search image at a shifted position of the template

$$m_s(x, y) = \frac{1}{n} \sum_{(u,v)\in T} s(x + u, y + v) \tag{7}$$

$$\sigma_s^2(x, y) = \frac{1}{n} \sum_{(u,v)\in T} (s(x + u, y + v) - m_s(x, y))^2. \tag{8}$$

While the template matches the image perfectly if $X_7 =$ NCC $(x, y) = \pm 1$, large absolute value of the X_7 generally indicates that the template closely corresponds to the image part under examination, while value close to zero indicates that the template and image do not correspond well. The feature extraction result X_7 is got by calculating the NCC, which is shown in (4).

These seven features given above are extracted in the component body and problem solder pad separately, so the total feature number is 14.

C. Feature Selection

The goal of the feature selection phase is to know what the influence of individual features is, as well as how certain combinations of features influence the classifier performance. There are 14 different features. To do this, a feature selection method based on information gain is used.

If the target attribute can take on c different values, then the entropy of S relative to this c-wise classification is defined as

$$\mathrm{Entropy(S)} = \sum_{i=1}^{c} -\mathrm{p}_i \log_2 \mathrm{p}_i \tag{9}$$

where p_i is the proportion of S belonging to class i. Note that the logarithm has the base 2 because entropy is a measure of the expected encoding length measured in bits. Note that if

the target attribute can take on c possible values, the entropy can be as large as $\log_2 c$.

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (10)$$

where values (A) is the set of all possible values for attribute A, and S is the subset of S for whose attribute A has value v (i.e., S,= {s ∈S|(s)=v)). The first term in (10) is just the entropy of the original collection S, and the second term is the expected value of the entropy after S is partitioned using attribute A. The expected entropy described by this second term is simply the sum of the entropies of each subset S_v, weighted by the fraction of examples S_v/S that belong to S_v. Gain (S, A) is therefore the expected reduction in entropy caused by knowing the value of attribute A.

This method calculates the probability of an instance being a segment border (prior probability) and compares this to the probability of a segment border given that a feature has a certain value. The higher the change in probability, the more useful this feature is.

IV. Classification

In this paper, a two-stage classifier for the inspection of solder joint using extracted features is proposed. In the first stage, the Bayesian classifier classified the solder joint qualified or not; in the second stage, the SVM classifier classified the solder joint into multiclass in detail.

A. Bayesian Classifier

The first-stage classifier is the Bayesian classifier which minimizes the total expected loss.

To minimize the average probability of error, we should select the class that maximizes the posterior probability $p(w_i|x)$, where w_i is the ith class and x is the input vector. It can be represented as follows.

Determine w_i if $p(w_i|x) > p(w_j|x)$ for all $j \neq i$.

Now, the classifier can be expressed by the following decision function $d_i(x)$:

$$d_i(x) = p(w_i|x). \quad (11)$$

According to the Bayesian formula

$$p(w_i|x) = \frac{p(w_i)\, p(x|w_i)}{p(x)} \quad (12)$$

where $p(x|w_i)$ is the probability density function of class w_i and $p(w_i)$ is the probability of occurrence of class w_i. Because $p(x)$ is positive and common to all $d_i(x)$, it can be dropped from the decision function. Then the decision function becomes

$$d_i(x) = p(x|w_i)\, p(w_i). \quad (13)$$

From the output of decision functions, input vector can be classified into the corresponding class which has the maximum value of decision function.

$$c_k = \left\{ x | x \in C,\, d_x = \max_{i \in C}(d_i) \right\} \quad (14)$$

where c_k represents the class of the kth input vector and C is the set of classes.

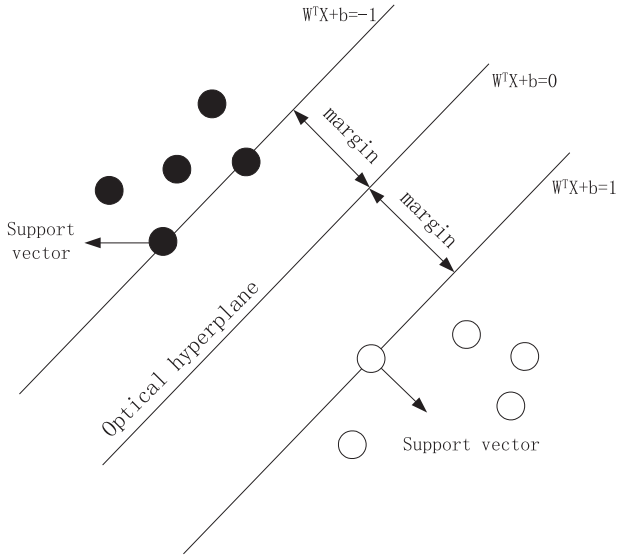

Fig. 6. Summary of the basic idea for SVM.

B. SVM

The second-stage classifier is the classifier. classifier is gradually developed base on statistical learning theory. First, the algorithm transforms input space into a high-dimensional space through nonlinear transform which is generated by the kernel function. Then, it seeks optimized linear classification, which requires not only correctly separating the two types, but also ensuring that the classification interval is the largest (see Fig. 6) [11], [12].

Set linear separable dataset (x_i, y_i), $i = 1, 2, \ldots, N \times \in R_d \times$ is eigenvector, N is the number of samples, $y \in \{+1 - 1\}$ is category label. In d-dimensional space, linear discriminant function is

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i (x, x_i) + b \right). \quad (15)$$

So the optimal separating hyper-plane equation is $w^T x + b = 0$ (see Fig. 6). To find the hyper-plane, it needs to solve the following quadratic programming problem:

$$\min \frac{1}{2} \| w \|^2 \quad (16)$$

$$\text{s.t.} y_i (w x_i + b) \geq 1. \quad (17)$$

For nonlinear problems, the SVM model adopts the feature mapping by introducing the kernel function K

$$K(x_i, x_j) = \Phi(x_i)\Phi(x_j). \quad (18)$$

There are four main kernel functions: linear, polynomial, radial basis function (RBF), and sigmoid kernel function.

In many practical situations, a separating hyper-plane does not always exist. To allow the possibilities of hyper-plane exist, slack variables $\xi_i \geq 0, i = 1, 2, \ldots, l$, are introduced which modify the cost function to penalize any failure to meet the original inequalities. The problems therefore become

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C^T \xi \quad (19)$$

$$\text{s.t. } y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (20)$$

where, the constraint parameter C controls the tradeoff between the dual objectives of maximizing the margin of separation and minimizing the misclassification error. Parameter w is weight vector, b is a scalar threshold and $\Phi(x_i)$ is the high-dimensional feature space which is nonlinearly mapped from the input space x.

For its dual problem, the Lagrange multiplier function is employed. The original function is thus modified as

$$\max w(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^{l} \alpha_i Q(i, j) \alpha_j \qquad (21)$$

$$\text{s.t.} \sum_{i=1}^{l} y_i \alpha_i = 0, 0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, l \qquad (22)$$

where, $Q(i, j) = y_i y_j \mathrm{K}(\mathrm{x}_i, \mathrm{x}_j)$, C is the penalty parameter and α_i is the Lagrange multiplier. After calculating, the original decision function is modified as follows:

$$f(x) = \mathrm{sgn}((\mathrm{w}^*, \mathrm{x}) + \mathrm{b}^*) = \mathrm{sgn}\left\{ \sum_{\mathrm{i}=1}^{\mathrm{n}} \alpha_{\mathrm{i}}^* \mathrm{y_i K(x, x_i)} + \mathrm{b}^* \right\}. \qquad (23)$$

The SVM is originally designed for binary classification, but it can be conveniently used in multi-class classification problems by combining a series of binary classifiers. The main combination methods are one-against-one, one-against-rest, and the directed acyclic graph SVM (DAG-SVM) [13].

The basic idea of one-against-one method is to choose different two classes and construct all possible binary sub-classifiers. On the other hand, the one-against-rest method compares a given class with all the others put together. In DAG-SVM, it divides all samples into two broad categories classes and goes on division in each category until there are only two classes left.

In the study by Hsu and Lin [14], all these methods have been compared. The one-against-one and DAG methods perform well. Due to their higher classification accuracy and faster classification speed, these methods are more suitable for practical use. However, the construction of DAG consumes more time and the structure of DAG is precarious. In our study, there are only six classes needed to be classified. So, it is quite reasonable to use one-against-one method, without a lot of classifiers.

V. Experimental Results

To evaluate the performance of the proposed method, experiments are performed with printed circuit boards by an AOI system developed by the authors [15]. First, a number of representative solder joint images, which are divided into training samples and testing samples, are selected as experimental samples. Second, the inspection performance was shown detailed with the proposed method. Finally, the performance of SVM classifier in the second stage was evaluated by comparing it with the other methods.

A. Sample Image Data

In this paper total number of 280 chip components (type 0402) with good and various defect types of solder joint are selected for experiment. The matrix associated with each image of the chip component has as dimensions 100 × 80 pixels in the component body area and 60 × 80 pixels in the solder pad area in the AOI systems developed by the authors. Each solder joint had been pre-inspected by an expert inspector and categorized into different classes, as shown in Table II.

TABLE II
Number of Samples

Type	Number
Tombstone	19
Wrong component	30
Solder insufficient	30
Component shifted	44
Pseudo solder	48
Good solder	109
Total	280

TABLE III
Result of the Information Gain Ranking Algorithm on All Features

Rank	Feature	Information Gain
1	4	1.4746
2	5	1.3452
3	6	1.2918
4	1	1.2415
5	7	1.2221
6	3	1.1165
7	14	1.0772
8	2	0.8715
9	11	0.502
10	8	0.3863
11	10	0.3814
12	12	0.32
13	13	0.1308
14	9	0

B. Feature Selection

The results of the information gain ranking in order can be seen in Table III. The 4th feature has the largest information gain value of 1.4746, and the 9th feature has the lowest information gain value of 0. All these features are selected according to the rankings in the table. When the selected number of features is 1, then only the 4th feature are selected, when the selected number of features is 2, the 4th and 5th feature are selected and so on.

C. Bayesian Classifier

In the first stage of the Bayesian classifier, the classification performance is shown below. The total 14 features [8] are extracted, the correct rate rise as the increase of the number of features, the detail data and the increasing curve was shown in Table IV and Fig. 7 separately. When the number of features

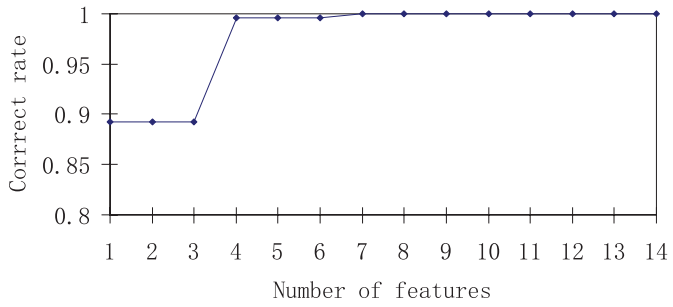

Fig. 7. Correct rate curve with Bayes.

TABLE IV
CORRECT ACCURACY OF BAYES WITH DIFFERENT NUMBER OF FEATURES

Number of Features	Correct Rate	Number of Features	Correct Rate
1	0.8929	8	1.00
2	0.8929	9	1.00
3	0.8929	10	1.00
4	0.9964	11	1.00
5	0.9964	12	1.00
6	0.9964	13	1.00
7	1.00	14	1.00

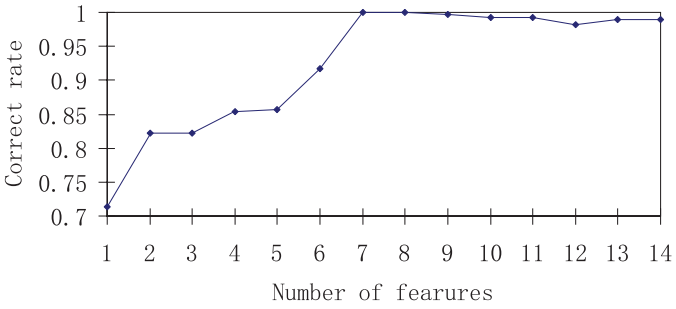

Fig. 8. Correct rate curve with SVM.

increases to seven, the correct rate rises to 100%. So in the first stage of inspection process, we need only seven features to get the highest inspection accuracy. The reason is seven features have enough information to classify the solder joint data set. Additional features cannot add more information.

D. SVM

In the second stage of the SVM classifier, the classification performance also was shown below. The total 14 features are extracted, the kernel function used here is polynomial (cost = 1, gamma = 2, degree = 3), the correct rate rise as the increase of the number of features, the detail data and the increasing curve are shown in Table V and Fig. 8 separately. When the number of features increases to seven, the correct rate rises to 100%. Then the correct rate begin to decrease as the number of features increases to 9. From the description above, we can know that the inspection accuracy rises as the number of features arises because the total information gain increases. But as the number of features increase to nine however, accuracy begins to suffer because the features are redundant.

TABLE V
ACCURACY OF SVM WITH DIFFERENT NUMBER OF FEATURES

Number of Features	Correct Rate	Number of Features	Correct Rate
1	0.7143	8	1.00
2	0.8214	9	0.9964
3	0.8214	10	0.9929
4	0.8536	11	0.9929
5	0.8571	12	0.9821
6	0.9167	13	0.9893
7	1.00	14	0.9893

TABLE VI
SVM CLASSIFICATION ACCURACY BY DIFFERENT KERNELS

Kernel (parameters)	Classification Accuracy
Linear (c=1)	99.29%
Polynomial (c=1,g=2,d=3)	100%
RBF (c=1,g=2)	86.43%
Sigmoid (c=1,g=2)	38.93%

TABLE VII
PERFORMANCE OF VARIOUS CLASSIFIERS

Classifier	Feature Number	Modeling Time(s)	Correct Rate
K-NN (K=1)	8	0.01	99.64%
Tree	7	0.04	99.28%
BP	7	0.66	100%
SVM	7	0.13	100%

Different kernels such as linear, polynomial, RBF, and sigmoid are used to SVM. Their highest classification effects are shown in Table VI. The polynomial kernel SVM has the highest accuracy, so we think that the features in our paper have a polynomial relationship.

To compare the classification performance in the second stage, various statistical classifiers and back propagation (BP) neural networks have been implemented, the highest correct rate of the classifier and the corresponding feature number have been compared, as summarized in Table VII. The K-NN and Tree method have a less modeling time and a lower correct rate at the same time. The BP and SVM classifier get similar highest correct rate when the number of feature is seven, but the BP classifier takes 0.66 s to build the model, which is longer than 0.13 s for SVM. The SVM with seven features has the best inspection performance. The inspection accuracy rises as the number of features arises because the total information gain increases, but when the number of features increase to 8, accuracy begins to decrease since the features are redundant.

The Bayesian classifier gets the highest inspection accuracy when the selected feature number is seven from the Part C, so we use seven features in this two-stage classifier.

VI. Conclusion

Two-stage approaches were evaluated in this paper, one is the binary classification approach, and the other one is multiclass classification approach.

After the total 14 features were extracted in the feature extraction step, a feature selection method based on information gain was used, then the Bayesian classifier determined if the solder joint qualified or not, and the SVM classifier classified the solder joint into multiclass.

The experimental results have shown that the proposed two-stage classifier with feature selection method achieved a high performance in recognition rate.

Acknowledgment

The authors would like to thank Corporation Folungwin Dongguan, Dongguan City, China, for its experimental facilities and testing area.

References

[1] M. Moganti, F. Ercal, C. H. Dagli, and S. Tsunekawa, "Automatic PCB inspection algorithms: A survey," *Comput. Vis. Image Understand.*, vol. 63, no. 2, pp. 287–313, 1996.

[2] P. J. Besl, E. J. Delp, and R. Jain, "Automatic visual solder joint inspection," *IEEE J. Rob Autom.*, vol. 1, no. 1, pp. 42–56, Jan. 1985.

[3] S. L. Bartlett, P. J. Besl, and R. Jain, "Automatic solder joint inspection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 1, pp. 31–43, Jan. 1988.

[4] M. R. Driels and D. J. Nolan, "Automatic defect classification of printed wiring board solder joints," *IEEE Trans. Compon. Hybrids Manuf. Technol.*, vol. 13, no. 2, pp. 331–340, Jun. 1990.

[5] T. H. Kim, T. H. Cho, Y. S. Moon, and S. H. Park, "Visual inspection system for the classification of solder joints," *Pattern Recognit.*, vol. 32, no. 4, pp. 565–575, Apr. 1999.

[6] G. Acciani, G. Brunetti, and G. Fornarelli, "Application of neural networks in optical inspection and classification of solder joints in surface mount technology," *IEEE Trans. Indus. Inform.*, vol. 2, no. 3, pp. 200–209, Mar. 2006.

[7] S.-C. Lin, C. H. Chou, and C.-H. Su, "A development of visual inspection system for surface mounted devices on printed circuit board," in *Proc. 33rd Annu. Conf. Indus. Electron.*, 2007, pp. 2440–2445.

[8] T. S. Yun, K. J. Sim, and H. J. Kim, "Support vector machine-based inspection of solder joints using circular illumination," *Electron. Lett.*, vol. 36, no. 11, pp. 949–951, 2000.

[9] F. Wu, X. Zhang, and Y. Kuang, "An AOI algorithm for PCB based on feature extraction inspection," in *Proc. 7th World Congr. Intell. Control Autom.*, 2008, pp. 240–247.

[10] H. Wu, X. M. Zhang, and Y. C. Kuang, "The research of the PCB location based on three layers of MARK point," in *Proc. Int. Conf. Adv. Design Manuf. Eng.*, 2011, pp. 869–876.

[11] M. V. M. Yeo, X. P. Li, and K. Shen, "A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system," *Safety Sci.*, vol. 47, no. 1, pp. 115–124, 2009.

[12] L. C. Jiao, F. Liu, J. Liu, *Intelligent Data Mining and Knowledge Discovery*. Xi'an: Xidian University, China, 2006.

[13] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, " Large margin DAGs for multiclass classification," in *Proc. Neural Inform. Process. Syst.*, 1999, pp. 547–553.

[14] C. W. Hsu and C. J. Lin, "A comparison of methods for multi-class support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Feb. 2002.

[15] F. Wu and X. Zhang, "Feature-extraction-based inspection algorithm for IC solder joints," *IEEE Trans. Compon., Packag. Manuf. Technol.*, vol. 1, no. 5, pp. 689–694, May 2011.

Hao Wu was born in 1986. He is currently pursing the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China.

His current research interests concern pattern recognition, particularly in automated optical inspection and pattern classification.

Xianmin Zhang was born in 1964. He received the Ph.D. degree in mechanical engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1993.

He is currently a Professor with the South China University of Technology, Guangzhou, China. He has authored or co-authored over 150 technical papers. His current research interests include AOI system, elastic and compliant mechanism analysis and design, dynamics and vibration control of precision instruments.

Hongwei Xie was born in 1981. He received the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China, in 2011.

He is currently a Post-Doctoral Research Fellow with the South China University of technology. His current research interests concern machine vision, particularly in AOI and pattern recognition.

Yongcong Kuang was born in 1970. He received the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China, in 2003.

He is currently an Assistant Professor with the South China University of Technology. His current research interests include machine vision and its applications in inspection.

Gaofei Ouyang was born in 1978. He received the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China, in 2007.

He is currently an Assistant Professor with the South China University of Technology. His current research interests include machine vision and its applications in inspection.