Research Article

Surface Defect Target Identification on Copper Strip Based on Adaptive Genetic Algorithm and Feature Saliency

Xuewu Zhang, Wei Li, Ji Xi, Zhuo Zhang, and Xinnan Fan

Computer and Information College, Hohai University, Changzhou 213022, China

Correspondence should be addressed to Xuewu Zhang; lab_112@126.com

Received 22 February 2013; Accepted 21 June 2013

Academic Editor: Yudong Zhang

Copyright © 2013 Xuewu Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To enhance the stability and robustness of visual inspection system (VIS), a new surface defect target identification method for copper strip based on adaptive genetic algorithm (AGA) and feature saliency is proposed. First, the study uses gray level cooccurrence matrix (GLCM) and HU invariant moments for feature extraction. Then, adaptive genetic algorithm, which is used for feature selection, is evaluated and discussed. In AGA, total error rates and false alarm rates are integrated to calculate the fitness value, and the probability of crossover and mutation is adjusted dynamically according to the fitness value. At last, the selected features are optimized in accordance with feature saliency and are inputted into a support vector machine (SVM). Furthermore, for comparison, we conduct experiments using the selected optimal feature subsequence (OFS) and the total feature sequence (TFS) separately. The experimental results demonstrate that the proposed method can guarantee the correct rates of classification and can lower the false alarm rates.

1. Introduction

With the development of production and processing, the quality requirements of products appearance have become increasingly higher, which caused the manufacturers to ensure nondestructive inspection for providing products. However, due to the variety of conditions in VIS, the size and the shape of the defect targets in detected object are usually indefinite, especially the false targets. For example, the falling iron scraps, the flying moth, the oil droplet, and so forth in copper strip production lines might be classified as true targets. Therefore, an automatic target identification system is necessary to be designed to eliminate the false targets. In this paper, the target identification system has the following stages: feature extraction, feature selection, and feature optimization and classification.

Feature extraction is a key factor for defect inspection. Actually, metal surface defects are mostly similar to texture patterns. Numerous methods have been proposed to extract textural features, and cooccurrence matrix method is one of statistical methods. Huang et al. [1] proposed an inspecting technology using gray level cooccurrence matrix to extract colorific and structural textures for solders of a flexible

printed circuit (FPC). The experiments indicated that the defective solder is obviously different from the nondefective solder in several kinds of quantified characters. Zhang et al. [2] developed a vision inspection system for the surface defects of strongly reflected metal. In this system, spectral measure approach based on Fourier spectral is used to compute textures which are inputted into a SVM to detect defects. Moment is a linear characteristic which has translation invariance, scaling invariance, and rotation invariance. Ping et al. [3] used the moment invariants to pick up the characters of typical copper surface defects. Tolba et al. [4] used GLCM and HU invariant moments based on Learning Vector Quantization (LVQ) classifiers to detect defects of textiles, and the correct defect detection rate is 98.64% with an average false acceptance rate of 0.0012.

Feature selection is an important preprocessing step in target identification system (TDS). Selecting an OFS that preserves classification accuracy is a growing important problem because of the increasing size and dimensionality of real data sets. Achieving reduction of the relevant features number without negative effect on classification accuracy is a goal which greatly improves the overall effectiveness of the target identification system. After analyzing the advantages

FIGURE 1: Block diagram of target identification.

and disadvantages of filtering and wrapping approaches for feature selection, Wang and Zheng proposed a hybrid approach named filtering-wrapping feature selection (FWFS) [5]. This approach uses information gain to evaluate feature's relevance to the target class and mutual information to evaluate redundancies among features. Then, the actual classifier is used as a "black box" to evaluate the fitness value. Rodriguez et al. [6] proposed a feature selection method, named Quadratic Programming Feature Selection (QPFS), which can limit the computational complexity for large data sets. Comparatively speaking, genetic algorithm has strong ability in local search, and the local optimal solution can be avoided in searching process [7], it solves the shortcomings of slow convergence speed and high time complexity. Yang et al. [8] presented an improved genetic algorithm (IGA) to select the optimal feature subset effectively and efficiently from a multicharacter feature set (MCFS). The IGA adopts segmented chromosome management scheme to implement local management of chromosome. A segmented crossover operator and a segmented mutation operator are employed to operate on these segments to avoid invalid chromosomes. The feature selection step of Zhang's method is based on adaptive simulated annealing genetic algorithm, which can guarantee the correct rate of classification and can improve the efficiency [9].

The proposed method optimizes the OFS judging by the feature saliency (i.e., the contribution degree of feature), which can increase the accuracy rate and robustness of the TDS. Saliency is the human perception of certain quality and quantity which can representatively reflect the difference between targets and other ones. For performance comparison, we do experiments on defects detection using the OFS and the TFS.

The organization of this paper is as follows: in Section 2, the proposed method is explained by three steps, feature extraction, feature selection, and feature optimization; Section 3 shows the experimental results and discusses the performance of the proposed method; conclusions are given in Section 4.

2. Target Identification System Design

Figure 1 shows the target identification process that contains two steps, training step and testing step. In the training step, the identification system firstly extracts the initial features from each image in training set to form the initial feature sequence set, which is then inputted into feature selection to obtain the optimal feature sequence and the weights. In the testing step, the weighted OFS is inputted into SVM classifier to implement the target identification.

- (1) *Feature extraction.* This study uses invariant moments and textural features to ensure that the extracted feature sequence can satisfy the requirements of target identification with validity, less computation quantity, and good robustness.
- (2) *Feature selection.* Genetic algorithm is a faster global optimizing algorithm with constant feedback correction. Consequently, the paper adopts GA to select the optimal feature sequence.
- (3) *Feature optimization.* Generally, each feature in OFS has different contribution to target identification. Therefore, this paper measures the feature saliency as the weights to update the existing identification model, which can guarantee the robustness and can further improve the performance of the identification model.
- (4) *Classification and identification.* SVM has been proposed as popular tools for classification problems [10]. After eliminating the features with less contribution in the training step, SVM can identify the target more accurately.

3. Methodology

3.1. Feature Extraction

3.1.1. HU Invariant Moments. Moments can be used for representation of a two-dimensional image on the basis of

the Papoulis uniqueness theorem [11]. Invariant moment theory based on region shape recognition was proposed by Hu [12] at first.

The $(p + q)$ th-order central moment of digital images is defined as

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q \rho(x, y). \quad (1)$$

The normalized central moment is

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}}, \quad p, q = 1, 2, 3, \dots \quad (2)$$

Hu proposed seven invariant moments which satisfied the conditions of translation invariance, scaling invariance, and rotation invariance. These moments can be written as follows:

$$\begin{aligned} \phi_1 &= \mu_{20} + \mu_{02}, \\ \phi_2 &= (\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2, \\ \phi_3 &= (\mu_{30} - 3\mu_{11})^2 + (3\mu_{21} - \mu_{03})^2, \\ \phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \\ \phi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \\ &\quad \times [(\mu_{30} + \mu_{12})^2 - 3(\mu_{30} + \mu_{12})^2] \\ &\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \\ &\quad \times [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], \\ \phi_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ &\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), \\ \phi_7 &= (3\mu_{21} - \mu_{03})(\mu_{03} + \mu_{21})[3(\mu_{30} + \mu_{21})^2 - 3(\mu_{21} + \mu_{03})^2] \\ &\quad - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) \\ &\quad \times [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]. \end{aligned} \quad (3)$$

The aforementioned seven parameters $\phi_1 \sim \phi_7$ are used as identification feature.1 to feature.7 in this paper.

3.1.2. Texture Features. Texture features can be used to measure the characteristics of smoothness, roughness, and regularity, and so forth. There are primarily three description methods for texture feature: statistical method, structured method, and spectrum method [13]. GLCM is an excellent statistical method in texture analysis, which can commendably convert gray value into texture information. GLCM is the statistics of probability $p(x, y)$ of a pair of image elements gray value with certain position relationship. The 14 feature parameters inferred from GLCM have all contained texture information. Nevertheless, they could bring about information redundancy when describing texture of surface defect

Feature sequence	1	2	3	4	5	6	·	·	·	l
Chromosome	1	0	0	1	0	0	·	·	·	1

FIGURE 2: Chromosome coding of feature sequence.

image. This paper chooses the following four parameters with preferable descriptiveness and independence.

(1) Angular second-order moment (energy)

$$A_1 = \sum_x \sum_y p(x, y)^2. \quad (4)$$

(2) Entropy

$$A_2 = -\sum_x \sum_y p(x, y) \log p(x, y). \quad (5)$$

(3) Contrast (inertia moment)

$$A_3 = \sum_x \sum_y m^2 [p(x, y)], \quad m = |x - y|. \quad (6)$$

(4) Relevance

$$A_4 = \frac{\sum_x \sum_y xy p(x, y) - \mu_1 \mu_2}{\sigma_1^2 \sigma_2^2}, \quad (7)$$

where, μ_1 , μ_2 , σ_1^2 , and σ_2^2 are, respectively, defined as

$$\begin{aligned} \mu_1 &= \sum_x x \sum_y p(x, y), \\ \mu_2 &= \sum_y y \sum_x p(x, y), \\ \sigma_1^2 &= \sum_x (x - \mu_1)^2 \sum_y p(x, y), \\ \sigma_2^2 &= \sum_y (y - \mu_2)^2 \sum_x p(x, y). \end{aligned} \quad (8)$$

This paper uses the mean value and the standard deviation of the aforementioned four parameters (i.e., energy, entropy, contrast, and relevance), total 8 subfeatures as identification feature.8 to feature.15 in this paper.

3.2. Adaptive Genetic Algorithm. Genetic algorithm is an adaptive optimization algorithm simulating biological evolution mechanism [7], which has strong searching capabilities in parallel pattern space and can quickly approach to the global optimal solutions.

(1) Chromosome Coding and Initial Population Setting. As shown in Figure 2, if the primary feature sequence extracted from training set contains l features, define a 0-1 binary code with chromosome length of l , where $l = 15$ corresponding to the preceding extracted features. If the i th chromosome is 1,

the corresponding feature is selected, otherwise not selected. Each chromosome corresponds to a feature subsequence.

Setting N randomly generates chromosomes as initial population. This paper take the $N = 1000$ in order to ensure the individual diversity of the population.

(2) *Fitness Function Designing.* On one hand, the larger the feature numbers in the feature subsequence, the more complicated the identification model, which could decrease the last identification performance with greater computational cost and declining antinoise capability.

On the other hand, the identification accuracy can be determined by the total error number To_e and the false alarm number Mi_e when the true/false target numbers are already known.

Therefore, the abovementioned three factors must be considered simultaneously when evaluating the feature subsequence from the practical application. The fitness function is defined as

$$F(f_i) = -(n \lg(L) + To_e \lg(To) + Mi_e \lg(Mi)), \quad (9)$$

where f_i is feature vector corresponding to feature subsequence, L is the total features number (i.e., the chromosomes length), To is the total training images number, and Mi is the true defects number.

From (9), we can see that the fewer the selected features, the total errors, and the false alarm numbers, the larger the received fitness value.

(3) *Genetic Operation.* The individual evolution is finished by genetic operators in GA. In this paper, the initial crossover operator $op_c = 0.8$, and the initial mutation operator $op_v = 0.01$. This paper adopts two-point crossover method to adaptively adjust the probability of crossover and mutation in the searching process.

The adaptive crossover operator is defined as

$$op_c = \begin{cases} op_{c1} - \frac{(op_{c1} - op_{c0})(F' - F_{av})}{F_{mx} - F_{av}}, & F' \geq F_{av}, \\ op_{c0} - (op_{c1} - op_{c0}) \frac{F'}{F_{av}}, & F' < F_{av}. \end{cases} \quad (10)$$

The adaptive mutation operator is defined as

$$op_v = \begin{cases} op_{v1} - \frac{(op_{v1} - op_{v0})(F_{mx} - F)}{F_{mx} - F_{av}}, & F \geq F_{av}, \\ op_{v0} - (op_{v1} - op_{v0}) \frac{F}{F_{av}}, & F < F_{av}, \end{cases} \quad (11)$$

where F' is the larger fitness value of the two crossover individuals and F_{mx} and F_{av} are the largest fitness value and the average fitness value separately.

(4) *Termination Conditions of AGA.* The AGA is terminated when the iteration times of AGA are equal to the maximum iteration times $MxGe$. In this paper, we set $MxGe$ to 400.

The algorithm flow chart is shown in Figure 3.

FIGURE 3: The feature selection algorithm flow chart based on GA.

The specific algorithm steps are as follows (where Ge is the iterations):

Population initialization: $Co \leftarrow$ randomly generating N chromosomes with length of k ;

while ($Ge \leq MxGe$ or $op_v \leq 0.09$), do

(1) fitness valuation: figure out $F(x)$ for each individual x in Co ;

(2) adaptive adjustment of op_c and op_v : using two-point crossover method to any two individuals;

(3) selection, crossover, mutation;

(4) population updating: $Co \leftarrow Co_n$;

end

output the OFS from Co .

3.3. *Feature Saliency Measuring.* This paper optimizes the selected features by measuring feature saliency in accordance with probabilities of each feature in features space. Suppose that the OFS is $f_o = [f_{o1}, f_{o2}, \dots, f_{om}]$ after using AGA; then make a certain feature $f_{oi} \in f_o$ input into SVM for classification. The total accuracy rate is

$$P_{ri} = \frac{To - To_{ei}}{To}. \quad (12)$$

FIGURE 4: Defect samples.

The target accuracy rate is

$$P_{tri} = \frac{Mi - Mi_{ei}}{Mi}. \quad (13)$$

The contribution degree D_i of feature f_{oi} is defined as

$$D_i = \frac{1}{2} (P_{ri} + P_{tri}). \quad (14)$$

Then, the normalization is applied to the contribution degree D_i , and the weight of feature f_{oi} is defined as

$$W_i = \frac{D_i}{\sum_{i=1}^m D_i}. \quad (15)$$

3.4. Basic Theory of SVM. Support vector machine (SVM) [10] is a learning method, which is based on structural risk minimization (SRM) rules. It means that the indicator function set S is decomposed into several subsets function sequence $S_1 \subset S_2 \subset \dots \subset S_n \subset S$, arranging the subsets according to the VC dimension size. After weighing the subset of the empirical risk and confidence range, we will obtain the smallest actual risk. It combines the theory of maximum interface classifier with the method based on kernel, and it can achieve the globally unique optimum solution. Besides, it has unique advantages in small samples, nonlinear problem, and high-dimensional pattern recognition problem. The basic idea of support vector machine is transforming the input space to a high-dimensional space by nonlinear conversion defined by inner product function firstly. In SVM, the choice of the appropriate kernel functions is an important factor, and different kernel function can form different algorithms, which will directly influence the generalization ability and error

control of SVM. Gaussian radial basis kernel, polynomial kernel, B-spline kernel, Fourier kernel, and sigmoid kernel are frequently used kernel functions.

In this paper, the one-versus-one method is chosen to build the SVM classifier, and the polynomial kernel is used as the kernel function.

4. Experimental Results and Discussion

In this section, the authors present the experiment results followed by previous approaches on copper strips that are provided by XINGRONG manufacture corporation in Changzhou, Jiangsu province, China to evaluate the performance of the proposed identification method. As shown in Figure 4, we choose eight typical defect target images that include true defect samples (burr, loophole, perforation, and pit) and false defect samples (moth, inclusion, iron scrap, and paper scrip). To extend the image database, we rotate the image sample and transform it in different resolutions. After that, we randomly choose 214 true defect images and 362 false defect images as the training set; the remaining 153 true defect images and 423 false defect images are used as testing set.

4.1. Feature Extraction. Figure 5 shows the spatial distribution map of each extracted feature of the defect images. From the maps, we can see that each feature can reflect the differences between true defects and false defects in a certain degree.

4.2. Feature Selection

(1) Single Feature. Table 1 shows the identification results using single feature. Figure 6 corresponds to Table 1. Table 1

FIGURE 5: Continued.

FIGURE 5: Spatial distribution maps of features.

TABLE 1: Identification accuracy rates using single feature.

Feature	Training set		Testing set	
	Total correct rate	Target correct rate	Total correct rate	Target correct rate
f1	0.8056	0.7533	0.8554	0.7727
f2	0.5637	0.4801	0.5373	0.3359
f3	0.6119	0.5003	0.6883	0.4988
f4	0.6715	0.7092	0.7004	0.5013
f5	0.5508	0.3342	0.6017	0.3081
f6	0.7702	0.6547	0.8009	0.6833
f7	0.8113	0.7087	0.8351	0.7104
f8	0.7518	0.6476	0.7991	0.6506
f9	0.8683	0.7812	0.8790	0.7941
f10	0.6013	0.4747	0.5706	0.3987
f11	0.8896	0.8900	0.8711	0.9014
f12	0.9004	0.9151	0.9107	0.8992
f13	0.9417	0.9450	0.9625	0.9633
f14	0.8006	0.7229	0.8513	0.7094
f15	0.5725	0.3918	0.5631	0.3614

TABLE 2: Results for AGA in 10 times.

Times	OFS coding	Feature numbers	Fitness value	Consuming time (s)
1	[10000111011110]	9	-9.0103	64.03
2	[10000111011110]	9	-9.0103	74.67
3	[10000111011110]	9	-9.0103	70.79
4	[10000111011110]	9	-9.0103	72.88
5	[10000111011110]	9	-9.0103	69.53
6	[10000111011110]	9	-9.0103	68.96
7	[10000111011110]	9	-9.0103	79.25
8	[10000111011110]	9	-9.0103	61.94
9	[10000111011110]	9	-9.0103	56.78
10	[10000111011110]	9	-9.0103	67.41
Average	[10000111011110]	9	-9.0103	68.62

FIGURE 6: Identification accuracy rates using single feature.

and Figure 6 further illustrate that each of the 15 features extracted in this paper has a certain identification ability for the true and false defects.

(2) *Multifeature*. As shown in Table 2, the selected OFS using AGA in 10 times is the same, that is $[f_1, f_6, f_7, f_8, f_9, f_{11}, f_{12}, f_{13}, f_{14}]$. The corresponding fitness value is -9.0103 , and the average consuming time is 68.62 s, which interprets that the processing speed of the proposed method is very fast.

Table 3 shows the experimental results using the selected OFS inputted into SVM. The total error using the OFS for training set is 9, and the false alarm number is 3; thus, the total is 0.9817, and the target correct rate is 0.9939. The total error using the OFS for testing set is merely 5, and the false alarm number is just 1; thus, the total correct rate is 0.9886, and the target correct rate is 0.9903. Compared with the results in Table 1, we can see that the identification accuracy using OFS is higher than using single feature independently.

Combined with Table 1, the weight of each feature in OFS calculated by (12) to (15) is shown in Table 4.

TABLE 3: Identification accuracy of optimal feature subsequence.

Training set				Testing set			
Total errors	False alarm	Total correct rate	Target correct rate	Total errors	False alarm	Total correct rate	Target correct rate
9	3	0.9817	0.9939	5	1	0.9886	0.9903

TABLE 4: Weight of each feature in optimal feature subsequence.

	f1	f6	f7	f8	f9	f11	f12	f13	f14
D_i	0.7795	0.7125	0.7600	0.6997	0.8248	0.8898	0.9078	0.9434	0.7618
Weight	0.1071	0.0979	0.1044	0.0961	0.1133	0.1222	0.1247	0.1296	0.1047

TABLE 5: Identification accuracy of total feature sequence.

Feature coding: [1111111111111], Fitness value: -19.1774							
Training set				Testing set			
Total errors	False alarm	Total correct rate	Target correct rate	Total errors	False alarm	Total correct rate	Target correct rate
10	7	0.9796	0.9624	14	13	0.9682	0.8738

In order to compare the performance of the proposed method, we use TFS inputted into SVM classifier under the same condition. And the identification results are shown in Table 5. From Table 5 we can see that the total errors for training set is 10, and the false alarm numbers is 7, thus the total is 0.9796, and the target correct rate is 0.9624. The total error for the testing set is 14, and the false alarm number is 13; thus, the total correct rate is 0.9682, and the target correct rate is 0.8738. By comparing Table 5 with Tables 1 and 3 it is observed that the identification accuracy using TFS for training set is higher than either single feature, but lower than that using OFS. Furthermore, the identification accuracy using TFS for the testing set is far below than using OFS, which further interpreted that the proposed feature selection method increases the identification accuracy rate greatly.

The neural network has been used in pattern recognition widely. Hundreds of defect images are randomly chosen from the database as the training sample; then we choose the other hundreds of images as the test sample. Table 6 shows the identification results using RBF neural network.

From Tables 3, 5, and 6, we can see that the identification accuracy of SVM can be as high as 99.03%, and in most instances this number is more than 95%. However, the identification accuracy of RBF neural network is no more than 95%. It is obvious that the performance of the proposed method is better than the RBF neural network.

5. Conclusions

In this study, a new surface defect target identification system for copper strip based on adaptive genetic algorithm and feature saliency was developed. Genetic algorithm has the advantage of fast convergence and can avoid involving into local optimal solution. In the proposed method, the probability of crossover and mutation was adjusted dynamically according to the fitness value, which had been calculated by integrating total error rate and false alarm rate. Furthermore,

TABLE 6: Identification accuracy of RBF neural network.

Defect type	Total	Correct	Accuracy
Pits	126	112	88.9%
Loophole	342	323	94.4%
Burr	215	203	94.4%
Perforation	106	96	90.6%

to evaluate the performance of the proposed method, comparison between the selected feature subsequence and total feature sequence has been implemented. The experimental results demonstrate that the proposed approach increases the correct rate and lowers the false alarm rate. In the proposed method, feature extraction can decrease the dimension of features and can increase the speed of processing. However, crossover and mutation operators in genetic algorithms always work under the condition of a certain probability; it will lead to “degradation” phenomenon inevitably, such as prematurity and species diversity decreasing. Some potential priori knowledge of the actual problem itself cannot be applied in the genetic algorithms. The future work should attempt to verify the robustness and effectiveness of our method in practical application.

Acknowledgment

This paper is supported by the National Natural Sciences Foundation of China (Grant no. 61273170).

References

- [1] J.-X. Huang, D. Li, F. Ye, and W. Zhang, “Detection of surface defect of solder on flexible printed circuit,” *Optics and Precision Engineering*, vol. 18, no. 11, pp. 2443–2453, 2010.
- [2] X.-W. Zhang, Y.-Q. Ding, Y. Lv, A. Shi, and R. Liang, “A vision inspection system for the surface defects of strongly

- reflected metal based on multi-class SVM,” *Expert Systems with Applications*, vol. 38, no. 5, pp. 5930–5939, 2011.
- [3] W. Ping, Z. Xuwu, M. Yan, and W. Zhihui, “The copper surface defects inspection system based on computer vision,” in *Proceedings of the 4th International Conference on Natural Computation (ICNC '08)*, pp. 535–539, October 2008.
 - [4] A. S. Tolba, H. A. Khan, A. M. Mutawa, and S. M. Alsaleem, “Decision fusion for visual inspection of textiles,” *Textile Research Journal*, vol. 80, no. 19, pp. 2094–2106, 2010.
 - [5] W. Wang and Q.-H. Zheng, “Feature selection for text categorization using filtering and wrapping,” *Journal of Computational Information Systems*, vol. 2, no. 4, pp. 1333–1342, 2006.
 - [6] I. Rodriguez-Lujan, R. Huerta, C. Elkan, and C. S. Cruz, “Quadratic programming feature selection,” *Journal of Machine Learning Research*, vol. 11, pp. 1491–1516, 2010.
 - [7] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Boston, Mass, USA, 1989.
 - [8] W.-Z. Yang, D.-I. Li, and L. Zhu, “An improved genetic algorithm for optimal feature subset selection from multi-character feature set,” *Expert Systems with Applications*, vol. 38, no. 3, pp. 2733–2740, 2011.
 - [9] H. Zhang, R. Tao, Z. Li, and H. Du, “A feature selection method based on adaptive simulated annealing genetic algorithm,” *Binggong Xuebao/Acta Armamentarii*, vol. 30, no. 1, pp. 81–85, 2009.
 - [10] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
 - [11] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, Mc Graw-Hill Kogakusha, Akita, Japan, 1965.
 - [12] M. Hu, “Visual pattern recognition by moment invariants,” *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 2002.
 - [13] Y. Chiou, C. Lin, and B. Chiou, “The feature extraction and analysis of flaw detection and classification in BGA gold-plating areas,” *Expert Systems with Applications*, vol. 35, no. 4, pp. 1771–1779, 2008.