Sensors & Transducers, Vol. 157, Issue 10, October 2013, pp. 412-418

www.sensorsportal.com

Sensors & Transducers

© 2013 by IFSA

http://www.sensorsportal.com

A High-Accuracy Algorithm for Surface Defect Detection of Steel Based on DAG-SVM

[1, 2] Jili LU, [1] Mingxing LIN, [1] Yan HUANG, [1] Xiangtao KONG
[1] School of Mechanical Engineering, Shandong University,
Jinan, Shandong 250061, China
[2] School of Mechanical and Electronic Engineering, Zaozhuang University,
Zaozhuang, 277160, China
E-mail: mxlin@sdu.edu.cn

Received: 14 October 2013 /Accepted: 29 October 2013 /Published: 31 October 2013

Abstract: The quality of the steel surface is a crucial parameter. An improved method based on machine vision for steel surface defects detection is proposed. The experiment is based on 20 images for each of 6 distinct steel defects, a total of 120 defective images achieved from the detection system. 128 different features are extracted from the images and feature dimensions are reduced by the principle component analysis (PCA) based on the sample correlation coefficient matrix. Hierarchical clustering by Euclidean distance is implemented to find defect characteristics differentiation, the steel surface defects are classified based on directed acyclic graph support vector machine (DAG-SVM). The experimental results indicate that this method can recognize more than 98 % of the steel surface defects at a faster speed that can meet the demands on the steel surface quality online detection. *Copyright © 2013 IFSA.*

Keywords: Machine vision, Surface defect detection, Principal component analysis (PCA), Directed acyclic graph support vector machine (DAG-SVM), Dimension reduction.

1. Introduction

Steel is one of the most fundamental materials for automobile manufacturing machinery, ships, military defense, etc. However, there are some defects on the steel surface such as roller, bruises, edge crack, scratch, scar generated by during producing, processing, transportation, loading, unloading, storage and many other reasons. These defects can not only affect the appearance of steel, but also cause steel scrap, corrosion, stress concentration, easily to result in affecting steel performance. So, how to detect these defects of steel effectively and implement real-time detection has become a significant subject.

Surface defect detection is an important application in machine vision field. General speaking, it needs a multiple classification of the surface detects detection because there are many types of surface defects. The difficulty focuses on the design of the classifier. Currently, there are many noticeable classifier methods such as Bayesian theory [1, 2], K-means method [3], Fisher discriminant [4], K-Nearest Neighbor (KNN) [5], clustering analysis [6, 7], etc. The classification methods represented by Bayesian theory and K-means are based on traditional statistics by using empirical risk minimization in place of experiential risk minimization, however, the deviation between the empirical risk minimization and experiential risk minimization achieves the theoretical minimum only when the training sample number tends to infinity, which is hard to meet in practical applications. In addition, these kinds of method need to known the priori knowledge and model structure, which is difficult to achieve in practice. Therefore, these methods can achieve good classification in theory but not ideal in engineering application. Support vector machine (SVM) [8-12] considers both the empirical risk and the structural risk to achieve best classification with less samples, so as to improve the generalization ability of the classifier. At the same time, the method can solve the nonlinear classification by introducing kernel function, and the complexity of calculation depends on the dimension of feature space rather than the number of samples.

In this paper, a surface defect detection system for steel plate with linear CCD and monochromatic light source is proposed according to the characteristics of the steel plate surface defect and the requirements of on-line detection. The DAG-SVM [13] decision tree is employed to detect the defects of steel, which can reduce the structural risk. In order to optimize the decision tree, the hierarchical clustering method is chosen to optimize the design of the branches of tree structure according to the degree of differentiation in defect characteristics, the selection of parameters of 2-classifier in each node are discussed corresponding to SVM decision tree.

2. CCD Detection System

The surface of measured steel plate is wide and continuous, tested continuously and online. The pit defect of the steel plate surface is single color and consistent with the measured steel surface but only color depth increases, these give us a hint that black and white linear CCD camera [14] can be chosen to reduce the amount of data processing and improve the detection speed. As a diffuse surface, the steel surface can scatter the light and lead to reduce illumination intensity, a monochromatic converged light source with high brightness is used to improve the lighting brightness. We design a detection system as shown in Fig. 1.

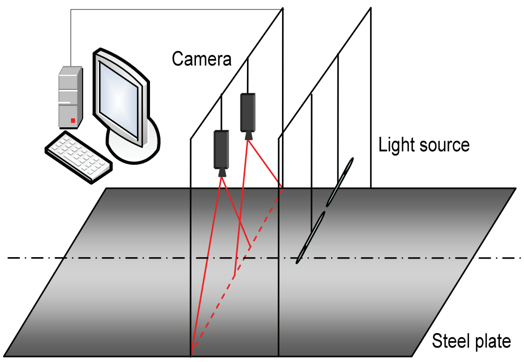

Fig. 1. CCD detection system structure chart.

Suppose the width of the measured steel surface is 1800 mm, and lateral resolution is 0.8 mm, two pixels correspond to a detection resolution, one pixel corresponds to 0.4 mm, then a steel surface with a width of 1800 mm needs 4500 pixels. As shown in Fig. 1, dual linear CCD cameras with 4096 pixels are set up on a beam, and two LED line light sources are set up on other beam. To ensure the detection area without omission, a length of 100 mm overlap detection zone is retained between the two cameras. As a result, the scanning length of each camera is 1000 mm. The cameras are installed vertically to increase the depth of field of the image system and ensure the acquisition image without distortion. Lights shine the camera scanning line to detect planar defects and pits defects simultaneously. In the case that light intensity meets the conditions, the illumination angle can be suitably reduced in order to increase the contrast of background with dimples and facilitate the detection of pit defects.

3. Defect Features Extraction

3.1. Main Surface Defect of Steel

Indentation, bubble, inclusion, crack, stains and white spot are 6 major defects of steel. The characters of single color, agreement with measured surface and only depth increased give us a hint that black and white camera, monochromatic light source can be chosen to reduce the amount of data processing and improve the detection speed. Surface defect of steel images are shown in Fig. 2.

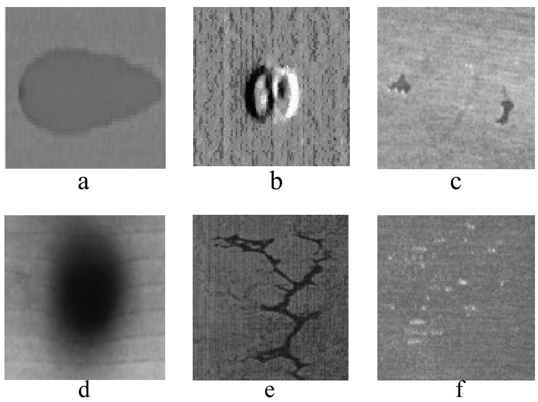

Fig. 2. Main surface defects of steel. (a) image with white spot, (b) image with bubble, (c) image with inclusions, (d) image with indentation, (e) image with crack, (f) image with stains.

3.2 Defect Segmentation

Defect segmentation is to separate the interested defect targets from the background information (such as color, outline, brightness, shape), which is the foundation of subsequent defect analysis and discrimination. It is impossible for us to accurately identify the size and quantity of the defects and make a right evaluation of surface quality for steel if a defect target does not contain the whole defects. For example, the pitting defect of steel is divided into several defect targets after edge segmentation. So, we need a merge algorithm to combine all the targets which belong to one defect into a new and complete defect. Considering the characteristics of steel surface features, we adopt the defect segmentation algorithm as shown in Fig. 3.

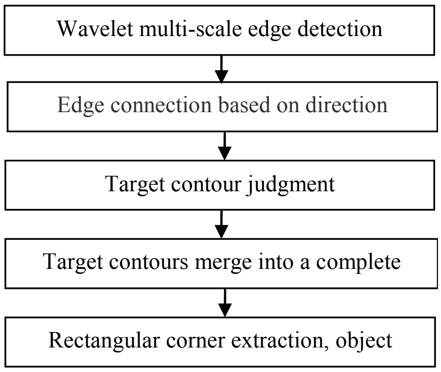

Fig. 3. Defect segmentation flowchart.

For the defects shown in Fig. 2, we obtain the corresponding segmentation results (as shown in Fig. 4) using the method as shown in Fig. 3.

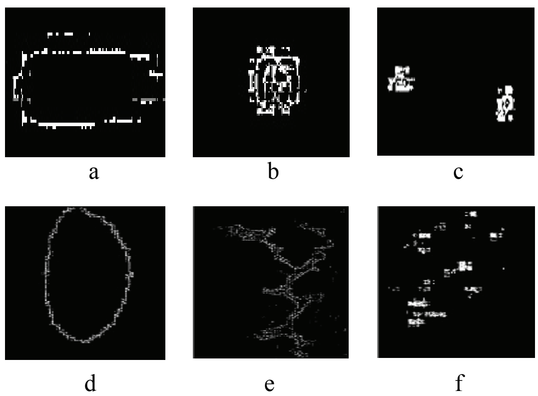

Fig. 4. Results of defect segmentation.

3.3. Dimension Reduction and Features Selection

For steel surface defect detection, it needs to further determine the type of defect after segmentation in order to meet the requirements of practical application.

The experiment studied the following surface defect characteristics of steel: geometric features, gray level, projection features and texture features, a total of 128 kinds. Including 18 kinds of geometric features which describe the size and shape of the defect, such as rectangularity, compactness, duty cycle, ovality, eccentricity, seven invariant moments, number of boundary fitting straight line and minimum slope, maximum slope, average slope and slope variance, etc. 27 kinds of gray level features which describe the gray distribution, such as the maximum, minimum, average, variance, slanting degrees, kurtosis, energy, entropy of target gray level, gray scale gradient, edge gradation; 40 kinds of projection features including slanting, kurtosis, energy, entropy, waveform, pulse characteristics, peak, margin characteristics, average, variance of x-projection , y-projection, 45° projection and 135° projection; 43 kinds of texture features mainly including the duty ratio, waveform, pulse characteristics, peak, margin characteristic, average, variance, slanting degrees, kurtosis, energy and entropy of the horizontal projection and vertical projection of HH1 and LH1 respectively.

If these 128 kinds of features were put into pattern classifier directly without any process, they will reduce the efficiency of pattern classifier due to the too large dimension of features. Therefore, we need to further choose the effective characteristics to reduce the dimensions.

The feature dimensions are reduced by the principle component analysis (PCA) based on the sample correlation coefficient matrix. According to testing results we know that geometric features are associated with edge length, area and shape, but not with image inside information; while all other 110 types of features are related to image content information. So these 128 types of features are divided into two groups for principal component analysis, 18 types of geometric features and other 110 types of features.

Suppose $X_i=(X_{i1},X_{i2},\cdots,X_{ip})',i=1,2,\cdots,n$ is a random sample with capacity of n taken from $X=(X_1,X_2,\cdots,X_p)'$, thus we have the sample covariance matrix S and the sample correlation coefficient matrix R:

$$S=(s_{ij})_{p\times p}=\frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})(X_k-\bar{X})' \quad (1)$$

$$R=(r_{ij})_{p\times p}=\left(\frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}\right) \quad (2)$$

Main features are selected by sample covariance matrix.

Suppose $S=(s_{ij})_{p\times p}$ is the sample covariance matrix, the eigenvaluess are $\hat{\lambda}_1\ge\hat{\lambda}_2\ge\cdots\ge\hat{\lambda}_p\ge 0$, and the corresponding unit orthogonalization eigenvalues are $\hat{\partial}_1,\hat{\partial}_2,\cdots,\hat{\partial}_p$, then we get the principle component of i^{th} sample:

$$Y_i=\hat{\partial}_i' X=\hat{a}_{i1}X_1+\hat{a}_{i2}X_2+\cdots+\hat{a}_{ip}X_p, i=1,2\cdots,p \quad (3)$$

Then n observations $X_k=(X_{k1},X_{k2},\cdots X_{kp})'$, $(k=1,2,\cdots,n)$ of X are plugged into Y_i one by one, n observations ($Y_{ki}(k=1,2,\cdots,n)$) of i th principal component Y_i are obtained,

$$\begin{cases} Var(Y_i)=\hat{\alpha}_i' S\hat{\alpha}_i'=\lambda_i, i=1,2,\cdots,p \\ Cov(Y_i,Y_j)=\hat{\alpha}_i' S\hat{\alpha}_j'=0, i\neq j \\ \sum_{i=1}^{p} Var(Y_i)=\sum_{i=1}^{p} s_{ii}=\sum_{i=1}^{p}\hat{\lambda}_i \end{cases} \quad (4)$$

So the contribution rate of i^{th} sample's principal component is gotten:

$$Cov\hat{\lambda}_i=\frac{\hat{\lambda}_i}{\sum_{i=1}^{p}\hat{\lambda}_i} \quad (5)$$

And the accumulating contribution rate of top m samples' principal components:

$$Cov\hat{\lambda}_{1\sim m}=\frac{\sum_{i=1}^{m}\hat{\lambda}_i}{\sum_{i=1}^{p}\hat{\lambda}_i} \quad (6)$$

Different dimensions may increase the dispersion degree difference of values, so we adopt standardized variables (sample correlation coefficient matrix) to eliminate the effects of different dimensions.

$$X_i^*=\left(\frac{X_{i1}-\overline{X}_1}{\sqrt{s_{11}}},\frac{X_{i2}-\overline{X}_2}{\sqrt{s_{22}}},\cdots\frac{X_{iP}-\overline{X}_P}{\sqrt{s_{PP}}}\right)', i=1,2,\cdots,n \quad (7)$$

The sample covariance matrix after standardization is the sample correlation coefficient matrix R of original data. We use the eigenvalues of R and the corresponding unit orthogonal eigenvectors and the principal component analysis process above-mentioned to obtain the sample principal component after standardization.

We choose 6 kinds of defects, 20 images of each defect, that total to 120 images. Here, geometric features, for example, 18 kinds of geometric features can be given in the form of equations:

$$B_1=\begin{pmatrix} v_{1-1} & v_{1-2} & \cdots & v_{1-k} & \cdots & v_{1-18} \\ v_{2-1} & v_{2-2} & \cdots & v_{2-k} & \cdots & v_{2-18} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{i-1} & v_{i-2} & \cdots & v_{i-k} & \cdots & v_{i-18} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{120-1} & v_{120-2} & \cdots & v_{120-k} & \cdots & v_{120-18} \end{pmatrix} \quad (8)$$

Here, element v_{i-k} of B_1 represents eigenvalue of the first k geometric features in image i. Using the principal component analysis based on sample correlation coefficient matrix above, we get eigenvalues and accumulative contribution rate of geometric features, as shown in Fig. 5.

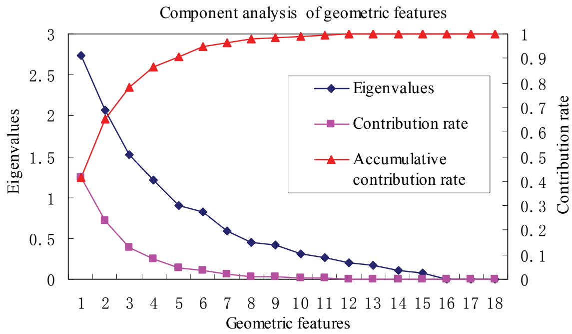

Fig. 5. Component analysis of geometric features.

From the data in Fig. 5, we can see that the accumulative contribution rate of the top 5 eigenvalues of sample covariance matrix has reached more than 90 %. With the accumulative contribution rate of 90 % as a standard, accordingly, we can use the principal components associated with the top 5 eigenvalues instead of the original 18 variables. In this way, the dimension reduction rate can reach to 18/5=3.6, while the information is kept at a rate of 90 %.

Therefore, the same method of the principal component analysis based on correlation coefficient can be used for the subsequent 110 eigenvalues. We can see that the accumulative contribution rate of the top 11 eigenvalues of sample covariance matrix has reached 91.5 %. We can use the principal components associated with the top 11 eigenvalues instead of the original 110 variables. In this way, the dimension reduction rate can reach to 110/11=10, while the information is kept at a rate of 90 %, and greatly improve the speed of the operation.

4. Defect Detection

4.1. DAG-SVM Decision Tree

Directed acyclic graph support vector machine (DAG-SVM) [13] is a new method of multiple classifiers combination of several binary classifiers. This method combines the SVM and decision tree, constructs $n(n-1)/2$ binary SVM classifications to realize the multi-classification.

16 types of main defect features are chosen according to the principal component analysis based on the sample correlation coefficient. Because of the obvious difference between geometric features and gray level features, the features are divided into two

parts (5 kinds of geometric features and 11 kinds of gray level features) that respectively is calculated. The calculation has two steps, the first step is to detect the presence of defects; the next is to classify the surface defect based on DAG-SVM. In the classification process, using the DAG-SVM decision tree, samples are judged by the principal component features. Those with obvious differentiation in geometry are categorized by geometry features while others without obvious geometry features are put into gray classifier for classification. In general, defects d, e, f that have obvious difference in geometry features on the principal component are classified by their geometric features; defects a, b, c that have not obvious differences in geometry features on the principal component are classified further by their gray features. The DAG-SVM decision tree of steel surface defects is shown in Fig. 6.

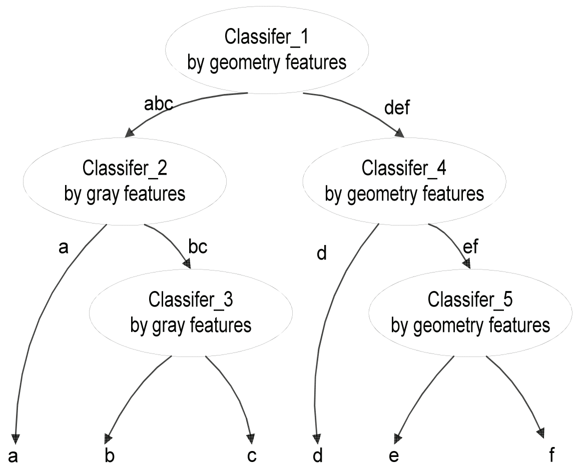

Fig. 6. DAG-SVM decision tree of surface defect detection.

4.2. Defect Features Clustering

In order to reduce the error rate of classifier and optimize design of classifier, we need to find the distinguishable differences of all the defect features, therefore, all defect samples cluster analysis in the initial stage. In this paper, the hierarchical clustering method is adopted, using the Euclidean distance to measure the similarity degree.

Suppose there are n samples with p indictors, the Euclidean distance between the sample X_i and sample X_j is:

$$d(X_i, X_j) = \left[\sum_{k=1}^{p} (X_{ik} - X_{jk})^2 \right]^{1/2} \qquad (9)$$

The average distance $D(G_p, G_q)$ between class G_p and class G_q is:

$$D(G_p, G_q) = \frac{1}{n_p n_q} \sum_{i \in G_p} \sum_{i \in G_q} d(X_i, X_j) \qquad (10)$$

The defect feature with larger distance is designed as the upper nodes of DAG-SVM while the defect feature with smaller distance as the lower nodes, to optimize the design of the DAG-SVM and reduce the accumulated errors.

4.3. Kernel Function of SVM

Due to steel has good discrimination in surface defect features, we choose Gaussian radial basis function (RBF) with few parameters as kernel function to reduce the complexity of algorithm, just determine the punish parameter c and the kernel function parameter g. In order to obtain higher classification accuracy, we use the K-fold cross validation (K-CV) method to adjust parameters of classifier.

Specific implementation: the data set is divided into K subsets, one as a test set while the rest of k-1 as training sets. Generally the data set is divided randomly or equally. Parameters c and g take a variety of possible values in a certain range randomly or according to prior knowledge to constitute a multiple classifier. Then the training data set is put into the multiple classifiers, take the parameters with highest classification accuracy is taken as the best parameters of c and g. However, there might be more groups of c and g corresponding to the highest classification accuracy. The parameters of c and g with minimum value of c are taken as optimal parameters. If the minimum value c corresponds to multiple values g, it takes the first retrieved value g as the optimal parameter. Because parameter c is corresponding to the confidence range of the classifier, high confidence range can lead to expected risks and over learning, the training set has high classification accuracy while the test set has very low classification accuracy. For this reason, the smallest possible parameters should be chosen on the premise of sufficient classification accuracy.

5. Results and Discussion

5.1. Classifier Parameters

For the parameters of classifiers are determined by geometric features and gray features. We choose 10 samples for each defect. There are 6 types of defects, a total of 10×6=60 training samples. Each has 5 geometric characteristic values and 11 gray characteristic values of principal component analysis, all of which forms a 60×16 data array as train_data. Each sample has a category number, all of which forms a 60×1 vector as train_class. We can get the parameters in the subsequent classifiers in the same way as the optimal parameters of classfier_1 can be obtained by the design ideas of K-fold cross validation method. The optimal parameters of 5 classifiers we get are shown in Table 1.

Table1. Parameters of classifiers.

Classifier	Parameter c	Parameter g
Classifier_1	0.082	0.758
Classifier_2	0.250	1.320
Classifier_3	0.00001	1.321
Classifier_4	0.144	2.297
Classifier_5	0.00097	2.296

5.2. Results of Classification

The abscissa represents 60 test samples and the ordinate represents 6 types of defects. Blue diamonds in figure represent real classes of each test sample and red squares mean the results of classification of the test samples. From Fig. 7 we can see that there is just 1 fault occurring amongst 60 samples by using our classifier, the classification accuracy is above 98 % which can meet the application's demand.

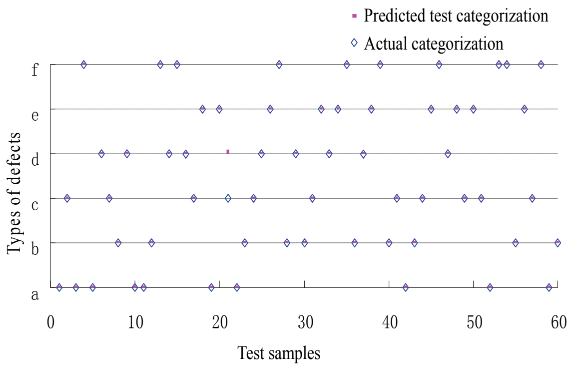

Fig. 7. Defects classification results.

6. Conclusions

1) A new steel surface defect detection method based on DAG-SVM is proposed. This method studies 6 types of steel defects, 20 images for each defect, a total of 120 defective images, and extracts 128 types of features. Experimental results indicate that this method can recognize above 98 % surface defect of steel, which can meet the demands of the steel surface quality online detection.

2) Feature dimensions are reduced by PCA based on the sample correlation coefficient matrix. To get a high computational efficiency and keep target information as much as possible, the geometrical features and gray features are separated to reduce dimensions. For these defect data samples of steel, 16 principal component parameters are kept after application of principal component analysis with the dimension reduction rate of 8:1.

3) The DAG-SVM method is implemented to classify the surface defects of steel. A 5-classifier decision tree is designed to classify 6 types of steel defects. The K-fold cross validation (K-CV) method is employed to optimize the classifier parameters and obtain higher classification accuracy.

Acknowledgements

The authors gratefully acknowledge the support of the natural science foundation (61301242) and the natural science foundation of Shandong Province (ZR2010EM037), the scientific research foundation for the returned overseas Chinese scholars, Sate Education Ministry.

References

[1]. J. M. Bernardo, A. F. M. Smith, Bayesian theory, *Wiley*, 2009.

[2]. F. Pernkopf, Detection of surface defects on raw steel blocks using Bayesian network classifiers, *Pattern Analysis and Applications*, Vol. 7, Issue 3, 2004, pp. 333-342.

[3]. T. Kanungo, D. M. Mount, N. S. Netanyahu, et al., An efficient k-means clustering algorithm: analysis and implementation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, Issue 7, 2002, pp. 881-892.

[4]. M. Sugiyama, Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis, *The Journal of Machine Learning Research*, Vol. 8, 2007, pp. 1027-1061.

[5]. L. Tomczak, V. Mosorov, D. Sankowski, J. Nowakowski, Image defect detection methods for visual inspection systems, in *Proceedings of the 9th International Conference 'The Experience of Designing and Application of CAD Systems in Microelectronics' (CADSM'07)*, art. No. 4297617, 2007, pp. 454-456.

[6]. A. Y. Ng, M. I. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, *Advances in Neural Information Processing Systems*, Vol. 2, 2002, pp. 849-856.

[7]. Fei He, Jin-Wu Xu, Zhi-Guo Liang, et al., Hot rolled strip state clustering based on kernel entropy component analysis, *Journal of Central South University: Science and Technology,* Vol. 43, No. 5, 2012, pp. 1732-1738 (in Chinese).

[8]. Zong-Hua Zhang, Hong Shen, Application of online-training SVMs for real-time intrusion detection with different considerations, *Computer Communications*, Vol. 28, No. 12, 2005, pp. 1428-1442.

[9]. S. Nashat, A. Abdullah, S. Aramvith, et al., Support vector machine approach to real-time inspection of biscuits on moving conveyor belt, *Computers and Electronics in Agriculture*, Vol. 75, Issue 1, 2011, pp. 147-158.

[10]. S. Ghorai, A. Mukherjee, M. Gangadaran, et al., Automatic defect detection on hot-rolled flat steel products, *IEEE Transactions on Instrumentation and Measurement*, Vol. 62, Issue 3, 2013, pp. 612-621.

[11]. Jing Bai, Li-Hong Yang, Xue-Ying Zhang, An anti-noise SVM parameter optimization method for

speech recognition, *Journal of Central South University: Science and Technology*, Vol. 44, No. 2, 2013, pp. 604-611 (in Chinese).

[12]. Hai-Peng Ren, Zhan-Feng Ma, Strip steel surface defect recognition based on complex network characteristics, *Acta Automatica Sinica*, Vol. 37, No. 11, 2011, pp. 1407-1412. (in Chinese)

[13]. J. C. Platt, N. Cristianini, J. Shawe-Taylor, Large margin DAGs for multiclass classification, *Advances in Neural Information Processing Systems, MIT Press, Cambridge*, Vol. 12, 1999, pp. 547–553.

[14]. Wu-bin Li, Chang-Hou Lu, Jian-Chuan Zhang, A lower envelope Weber contrast detection algorithm for steel bar surface pit defects, *Optics & Laser Technology*, Vol. 45, Issue 1, 2013, pp. 654–659.

2013 Copyright ©, International Frequency Sensor Association (IFSA). All rights reserved.
(http://www.sensorsportal.com)