An Adaptive Support Vector Machine-Based Workpiece Surface Classification System Using High-Definition Metrology

Shi-Chang Du, De-Lin Huang, and Hui Wang

***Abstract*— The shape of a machined surface significantly impacts its functional performance and exhibits different spatial variation patterns that reflect process conditions. Classification of these surface patterns into interpretable classes can greatly facilitate manufacturing process fault detection and diagnosis. High-definition metrology (HDM) can generate high density data and detect small differences of workpiece surfaces, which exhibits better performance than traditional measurement methods in process diagnosis. In this paper, a novel adaptive support vector machine (SVM)-based workpiece surface classification system is developed based on HDM. A nonsubsampled contourlet transform is used to extract features before classification with its characteristics of multiscale, multidirection, and less dimension of feature vectors. An adaptive particle swam optimization (APSO) algorithm is developed to search the optimal parameters of penalty coefficient and kernel function of SVM and is helpful to escape from the local minimum by its strong ability of global search. A varied step-length pattern search algorithm is explored to optimize the global point in every iteration of the APSO algorithm by its good performance in local search. These two algorithms are combined with their relative merits to find the optimal parameters for building an adaptive SVM classifier. The results of case studies show that the proposed adaptive SVM-based classification system can achieve a relatively high classification accuracy in the field of workpiece surface classification.**

***Index Terms*— Nonsubsampled contourlet transform (NSCT), particle swam optimization (PSO), quality control, support vector machine (SVM), workpiece surface classification.**

I. Introduction

THE classification of workpiece surface patterns is considered as an essential element in understanding the functional performance of the product and providing feedback on the manufacturing process [1]. As traditional measurement devices (such as coordinate measuring machines) measure only a few scattered points or profiles due to economic constraints, they cannot sample high density data describing 3-D surface spatial variation patterns in industrial applications. Recently, high-definition metrology (HDM) [2], [3] provides opportunities for product quality control since high density data collected by an HDM device can precisely characterize the surface and reflect the impact of manufacturing processes on the surface quality of a machined workpiece. HDM is superior to traditional measurement since it can generate a surface height map of millions of data points within seconds [4] and help us to understand and improve surface quality control strategies in high-precision manufacturing [5], process control [6], and process improvement [7], [8]. Though a lot of research has been done for image classification, research about the workpiece surface classification using high density data points collected by HDM is sparse.

The classification tasks using HDM can be divided into two main steps: 1) feature extraction and 2) classification. Since it is not appropriate to directly use the raw data collected from the workpiece surfaces to classify, features should be extracted to represent a given workpiece surface before classification. Many surface characterization approaches have been proposed to obtain abundant information about 3-D surfaces, such as a 3-D parameter set [4], [9], gray level co-occurrence matrix [10], and 2-D autocorrelation function and spectral analysis [11]. In recent years, some approaches that were first used in signal processing areas have been adopted to extract features of workpiece surfaces, such as Gabor filter banks [12], Gaussian filter banks [13], and wavelet packets [14]–[16], and after filtering, some numerical surface parameters are calculated for each sub-band to represent a given surface. In this paper, feature extraction is implemented by nonsubsampled contourlet transform (NSCT) developed in [17], which possesses the properties of full shift invariant, multiscale, and multidirection.

Automatic and accurate classifications are important tasks in industry practice and lots of methods are developed for different classification purposes [18]–[22]. Among many statistical learning-based classification methods, the support vector machine (SVM) proposed in [23] is gaining popularity due to its many attractive features and promising generalization performance. Unlike other classification methods (such as artificial neural networks), the SVM has a high capacity for generalization [24] and is used as a relatively novel statistical

Manuscript received September 10, 2014; revised January 26, 2015; accepted January 27, 2015. Date of publication May 15, 2015; date of current version September 11, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 51275558, in part by the Shanghai Rising-Star Program under Grant 13QA1402100, and in part by the U.S. National Science Foundation under Grant CMMI-1434411. The Associate Editor coordinating the review process was Dr. Ruqiang Yan. *(Corresponding author: Shi-Chang Du.)*

S.-C. Du and D.-L. Huang are with the Department of Industrial Engineering and Management, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: lovbin@sjtu.edu.cn; cjwanan@sjtu.edu.cn).

H. Wang is with the Department of Industrial and Manufacturing Engineering, Florida State University, Tallahassee, FL 32306 USA (e-mail: hwang10@fsu.edu).

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/TIM.2015.2418684

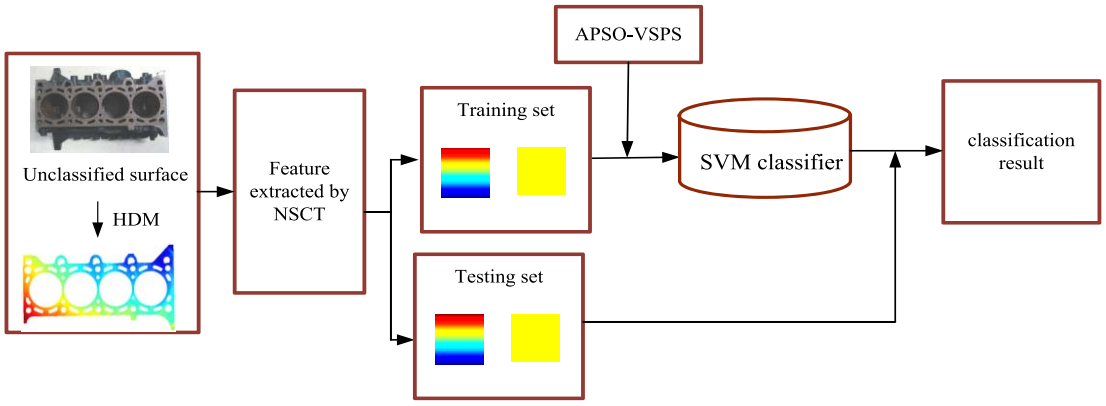


Fig. 1. Framework of the workpiece surface classification system.

learning tool [25], [26]. The SVM can effectively solve such practical classification problems of small samples, and nonlinear and high dimensions [27]–[30]. The performance of the SVM has a close relationship with the parameters of kernel function and penalty coefficient. Therefore, it is imperative to study the optimization method of SVM parameters.

Optimization of parameters (such as kernel function and penalty coefficient) in a SVM classifier face challenges. Previous parameter optimization method is adopted by experiences or experiments [31], not only leading to problems such as intensive computation and low efficiency but also limiting the application of the SVM since the selected parameters are usually not optimal. To overcome the drawback, some researchers propose many methods based on different search algorithms such as grid search [32], genetic algorithm (GA) [33], and particle swam optimization (PSO) algorithm [34], [35]. These methods have been proven effective by experiments in the corresponding articles. However, grid search has heavy burden of computation with low learning accuracy [36], and the implementation of a GA is complex and needs to design different ways in creating crossover or mutation [37].

Though PSO algorithm is a viable optimization with good robustness [34], [35], [38], it has the problem of premature convergence into local extreme points and poor local search capability [39], [40]. In a standard PSO algorithm, the inertia weight cannot well balance the ability of particles between global search and local search as it is a fixed value [41]. Inertia weight is used to characterize the extent of the impact on particle's previous velocity relative to its current speed. Therefore, inertia weight affects the balance between particles' ability of global search and local search. The adjustments methods to inertia weight currently include mainly linear change, fuzzy adaptive, and random changes. The most widely used method among them is a linearly decreasing strategy proposed in [42] and [43], but this method is likely to trap the solutions in a local optimum when solving multimodal function problems. To overcome these drawbacks, this paper proposes a novel adaptive SVM-based classification method. An adaptive PSO (APSO) algorithm is proposed to control the update of inertia weight in every iteration for each particle, which is easy to implement and helps in preventing the solution stuck at a local minimum. Then a varied step-length pattern search (VSPS) algorithm is proposed to make use of its good performance in local search, leading the entire particle swarm to search for a potential optimal solution more efficiently.

The remainder of this paper is organized as follows. Section II outlines the framework of the proposed adaptive SVM-based classification system. Section III briefly introduces NSCT theory and demonstrates how to use it to extract features. Section IV describes the theories of PSO, pattern search (PS), and the details of the proposed adaptive SVM-based classification system. In Section V, three case studies are presented to validate the proposed adaptive SVM-based classification system. The performance analysis is implemented to illustrate how the parameters affect the classification performance for high-resolution surface samples. Finally, a conclusion is given in Section VI.

II. Framework of the Proposed Classification System

This section presents an overview of the proposed classification system with the components of NSCT and a SVM classifier based on the APSO-VSPS algorithm. To be specific, in feature extraction, high-resolution images are reflected by 3-D coordinates of workpiece surfaces, which are filtered to extract details in different directions and scales by NSCT. The means and standard deviations of the coefficients of the wavelet sub-bands are calculated as feature vectors to represent a given surface. In the proposed classification method, the APSO algorithm is the main algorithm to select the optimal combination of parameters and the VSPS algorithm is nested in the main algorithm to improve the efficiency and accuracy of the search by its good performance in local search. The framework of the proposed system is shown in Fig. 1.

The procedure involves the following steps.

Step 1: HDM is employed to measure and collect 3-D coordinates from workpiece surfaces.

Step 2: NSCT is used to extract features from height maps of the 3-D coordinates. The extracted feature dataset is divided into two subsets: 1) training dataset and 2) testing dataset.

Step 3: The training dataset is input for the proposed APSO-VSPS algorithm to build a SVM classifier.

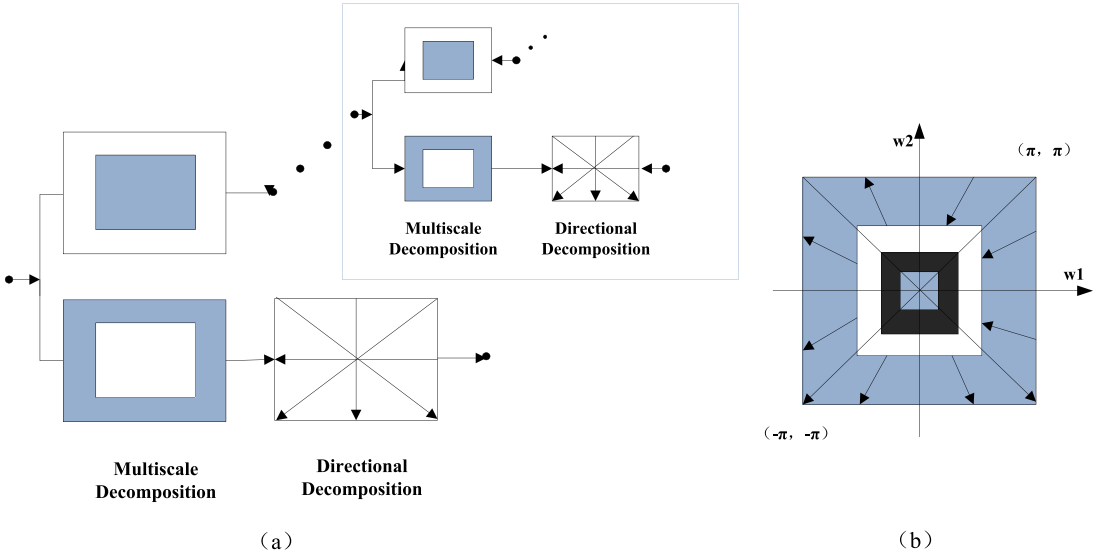


Fig. 2. NSCT. (a) Structure figure. (b) Frequency division result.

Step 4: The testing dataset is classified by the SVM classifier (built in Step 3) and the classification result is the final output.

III. Feature Extraction Using NSCT

The original data gained by HDM should not be directly used for classification since these data are just 3-D coordinates that cannot be used as features. Therefore, a feature exaction procedure is necessary. An NSCT algorithm is used to exact features in this paper because it contains many merits of few constrains, good frequency selectivity, and good decomposition of an image sub-band.

A. Nonsubsampled Contourlet Transform

NSCT combines nonsubsampled pyramids (NSPs) and nonsubsampled directional filter bank (NSDFB), and it has many properties such as multiscale, multidirection, and shift-invariance, which can effectively eliminate pseudo-Gibbs phenomenon [44] in signal processing and is sensitive to texture feature extraction [45].

The nonsubsampled pyramid filter bank (NSPFB) (or NSP) is a two-channel nonsubsampled filter bank [46], [47]. The process of NSP is similar to nonsubsampled wavelet transform [17] based on à trous algorithm [48] in achieving the multiscale decomposition of an image. The building block of an NSDFB is also a two channel nonsubsampled filter bank. The NSDFBs are iterated to obtain finer directional decomposition. The NSPFB provides multiscale decomposition and NSDFB provides directional decomposition.

Fig. 2 shows the block diagram of an NSCT that contains two parts of structure figure and frequency-division result. It can be seen that the transformation can be divided into two shift invariant parts including an NSP structure to ensure the NSCT in multiscale characteristics and a nonsubsampled directional filter under a given direction. The input is decomposed into high-pass parts and low-pass parts by the NSPFB, and then the high-pass sub-band is decomposed into several directional bands by the NSDFB. The scheme is iterated repeatedly on the low-pass sub-band.

B. Feature Extraction

Feature extraction is related to quantification of surface characteristics and its quantitative results are known as feature vectors. Optimization of these descriptive parameters is important because these parameters will influence the subsequent classification performance to some degree. A K-level NSCT decomposition is implemented to extract features. To reflect the degree of image texture, the mean and standard deviation of the coefficients matrix extracted by the NSCT for decomposition transform are used to constitute an eigenvector $f = [\mu, \delta]$. Mean μ and standard deviation δ are calculated as

$$\mu = \frac{1}{A \times B} \sum_{x=1}^{A} \sum_{y=1}^{B} |P(x, y)| \tag{1}$$

$$\delta = \sqrt{\frac{1}{A \times B - 1} \sum_{x=1}^{A} \sum_{y=1}^{B} [P(x, y) - \mu]^2}. \tag{2}$$

In (1) and (2), A denotes the number of points along x-axis, B denotes the number of points along y-axis, $P(x, y)$ denotes the coefficient matrix of wavelet sub-bands, and μ and δ denote the mean and standard deviation of the sub-bands' coefficients, respectively.

Next, by computing means and standard deviations from the coefficients of wavelet sub-bands of low-pass in K levels, $2 \times K$ characteristic values are obtained to compose a $2 \times K$-dimensional feature vector $(\mu_1, \mu_2, \ldots, \mu_K, \delta_1, \delta_2, \ldots, \delta_K)$. The decomposition by NSCT just outputs the last level's coefficients of wavelet sub-bands of low-pass. Therefore, it is necessary to find another $K-1$ levels of coefficients of wavelet sub-bands of low-pass since these coefficients affect the classification results. To wavelet sub-bands of high-pass,

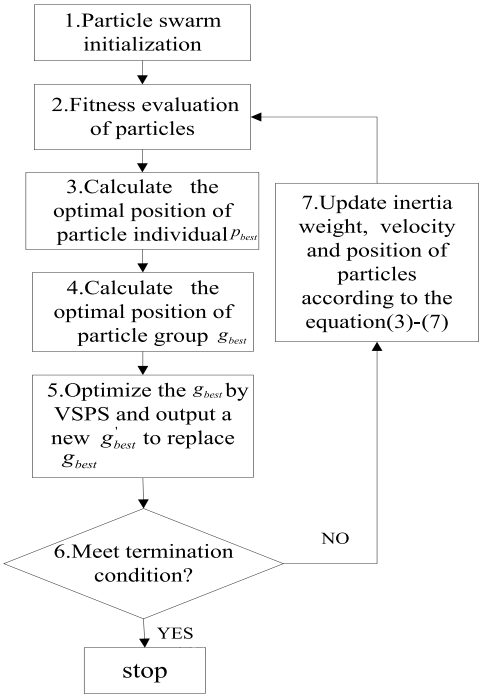


Fig. 3. Flow diagram of the adaptive parameter optimization algorithm.

means and standard deviations are calculated from the coefficient matrix in each direction ($[a_1, a_2, \ldots, a_K]$) of K levels and a $2 \times \sum_{i=1}^{K} 2^{a_i}$-dimensional feature vector is generated $(\mu_1, \mu_2, \ldots, \mu_{\sum_{i=1}^{K} 2^{a_i}}, \delta_1, \delta_2, \ldots, \delta_{\sum_{i=1}^{K} 2^{a_i}})$. By combining the $2 \times \sum_{i=1}^{K} 2^{a_i}$-dimensional high-pass feature vector and $2 \times K$-dimensional low-pass feature vectors, a $2 \times (\sum_{i=1}^{K} 2^{a_i} + K)$-dimensional feature vector is obtained to represent a given surface.

IV. Proposed Adaptive SVM Classifier

Since the SVM is currently one of the most popular classification tools, there is no need to discuss it in much detail in this paper. For more details about the SVM, please refer to [24]. Two parameters of C (controls the tradeoff between function complexity and training error of the SVM classifier) and σ [denotes the width parameter of the Gaussian radial basis function (RBF) and controls the radial scope of function] have a great impact on the classification performance of SVM classifier. An APSO algorithm and a VSPS algorithm are proposed to jointly optimize the two parameters. The flow diagram (shown in Fig. 3) is the process of the proposed APSO-VSPS algorithm, which is an APSO algorithm nested by the VSPS algorithm.

The APSO algorithm is an adjustment strategy adopting different inertia weights to update the particle swarm in the same generation, which accelerates the convergence speed and can jump out of local optimal solution. In this way, it provides a better initial value for VSPS, which is conducive to obtain the final optimal solution.

The proposed algorithm can be described as follows.

Step 1: Initialize all particles. Set initial position and velocity of particles within the allowable range randomly. The position of a particle is used to indicate a potential solution of the optimization problem in the search space. Set initial position of each particle as its local optimum p_{best} and set the best value in the set of p_{best} as g_{best}.

Step 2: Evaluate the fitness value of each particle and calculate the objective function value of each particle according to the fitness function.

Step 3: Calculate the best position of each particle p_{best}.

Step 4: Calculate the best position of all particles g_{best}.

Step 5: The g_{best} is set as the initial search point in the VSPS algorithm and the output is g'_{best} that replaces the previous g_{best}. In the vicinity of the better point for local search by VSPS, the search should continue until the result meets the accuracy requirements.

Step 6: Check the terminal condition (the optimal solution stagnates and changes within a range of $\pm 2\%$). If the termination is met, then stop. If not, go to Step 7.

Step 7: Update the inertia weight, velocity, and position of the current particle according to (3)–(7) and back to Step 2.

A. Adaptive Particle Swarm Optimization Algorithm

The PSO is an iterative optimization algorithm that is initialized to a group of random particles. The flying direction and distance of each particle may be controlled by its own velocity, while the pros and cons of each particle can be evaluated by a fitness value function. In each iteration, the algorithm is mainly to update each particle by tracking individual extreme point p_{best} and global extreme point g_{best}. When searching the two points, every particle updates its speed and position according to the following equations:

$$v_t = \omega v_{t-1} + c_1 r_1 (p_{\text{best}} - x_{t-1}) + c_2 r_2 (g_{\text{best}} - x_{t-1}) \quad (3)$$

$$v_t = \begin{cases} v_{\max}, & v_t > v_{\max} \\ -v_{\max}, & v_t \le v_{\max} \end{cases} \quad (4)$$

$$x_t = x_{t-1} + v_t \quad (5)$$

where ω is the inertia weight, c_1 and c_2 are the acceleration constraints, $r_1, r_2 \in [0, 1]$ are the random values, x_t is the current position of the particle that represents the current parameter value of C and σ in SVM, g_{best} is the best position of a group in history records, and p_{best} is the best position of the current particle in history records.

To accelerate the convergence speed while not being trapped in the local optima, an adaptive ω-strategy and the concrete equations are adopted as follows:

$$\omega(i, x) = u(d_i, x) * \left[\omega_1 + (\omega_2 - \omega_1) * \frac{\text{CurIter}}{\text{MaxIter}} \right] \quad (6)$$

where MaxIter is the maximum number of iteration, CurIter is the current number of iteration, ω_1, ω_2 are the initial value and final value of ω, respectively

$$u(d_i, x) = \begin{cases} 1 + \theta, & d_i \le S_1 M \\ 1, & S_1 M < d_i < S_2 M \\ 1 - \varsigma, & d_i \ge S_2 M \end{cases} \quad (7)$$

where M is the population size, S_1, S_2 are the control parameters, and θ, ς are the adjustment parameters, and satisfy the following conditions: $S_1 < S_2 < 1$, $\theta > 0, \varsigma > 0$.

B. Varied Step-Length Pattern Search Algorithm

In this section, the VSPS algorithm is embedded within the APSO algorithm due to its good performance in local search. Hooke and Jeeves [49] first proposed PS in 1961, and then Torczon [50] completed the convergence proof of the PS algorithm for unconstrained problems and gave a framework for a class of PS methods, which contain coordinate search method, Hooke–Jeeves (H–J) method, and Powell conjugate direction method [51]. The H-J PS algorithm [49] alternates the implementation of two searches from the initial point: axial search and PS. The first type of search is an exploratory move (also called the axial search explained in the subsequent Step 2) designed to determine the new base point and the direction conducive to decrease the values of fitness function, and move in the direction of m-axis. The second type of search is a pattern move designed to utilize the information acquired in the exploratory moves, and accomplishes the actual minimization of the function by moving in the conjunction direction of two adjacent base points.

The H–J PS algorithm is a constant nonlinear function optimization method that is widely used in many fields. It is a simple iteration as it does not require the objective function to be derivable. Besides, the H–J PS algorithm has good performance in local convergence and is sensitive to initial value in consideration of local convergence. However, in the process of the H–J PS algorithm, the same step length for search in all directions restricts the convergence speed and accuracy and if the search point is a local maximum point of the function, such an exploratory move may lead to a direction of selecting an inferior point rather than the better one. Therefore, a VSPS algorithm is proposed to avoid these drawbacks, which consists of a variable step change strategy, accelerating factors, and a new way for exploratory moving. The flow diagram of the proposed VSPS algorithm is shown in Fig. 4.

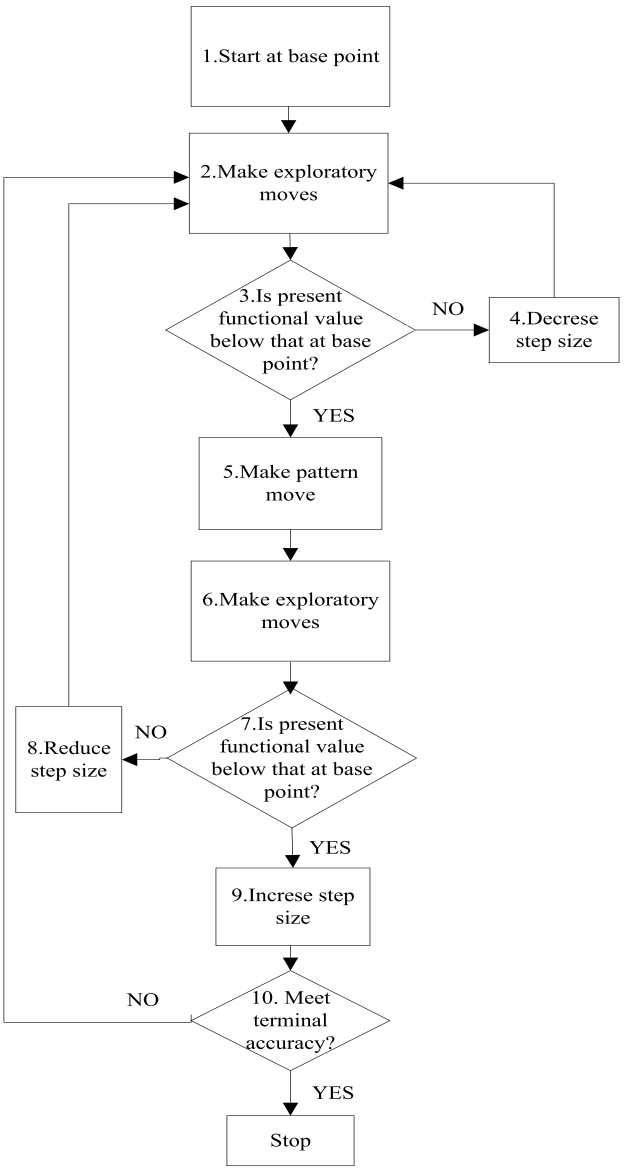


Fig. 4. Flow diagram of the proposed VSPS algorithm.

To describe the two modes of axial search and PS, the procedure is presented as follows. With respect to $\min f(x)$ (where min is to minimize and $f(x)$ is the objective function), let $e_i = (0, \ldots, 0, 1, \ldots, 0)^T, i = 1, 2, \ldots, m$ and e_i denote unit vector and m denote m-direction of the axis. $y^i (i = 1, 2, \ldots, m)$ is used to indicate a starting point search in the direction of the ith axis e_i.

Step 1: Select the initial point x^1, initial step length $\chi > 0$ (here, χ is a m-dimensional vector rather than a numerical value), acceleration factor $\eta \geq 1$, increase rate of step length $\lambda > 1$, decrease rate of step length $\beta \in (0, 1)$. ε is the accuracy requirement. Let $C_{\text{pso}}, \sigma_{\text{pso}}$, the search loop step number $k = 1 \cdot x^k$ is used to denote the kth base point.

Step 2: Start the axial search mode. In every axial search, there are two search directions: one is along the positive direction $(+e_i)$ of the axis and the other is the opposite $(-e_i)$, $i = 1, 2, \ldots, m$. Compare the value of $f(y^i + \chi(i)e_i) = f_{i1}$ (positive direction), $f(y^i - \chi(i)e_i) = f_{i2}$ (negative direction) and $f(y^i) = f_{i3}$. If $f_{i1} < f_{i3} \leq f_{i2}$ or $f_{i1} < f_{i2} < f_{i3}$, let $y^{i+1} = y^i + \chi(i)e_i$ and $\chi = \chi \times \lambda$ (the search along the positive direction of the axis is effective); if $f_{i2} < f_{i3} \leq f_{i1}$ or $f_{i2} < f_{i1} < f_{i3}$, let$y^{i+1} = y^i - \chi(i)e_i$ and $\chi = \chi \times \lambda$ (the search along the negative direction of the axis is effective); else let $y^{i+1} = y^i$ and search along another axis. The search should be along m axes and a value of $f(y^{m+1})$ is obtained. An example is shown in Fig. 5(a) in a 2-D coordinate system. x^1 and x^2 are base points, y^1,y^2,y^3 are points found by axial search and pattern search, e_1 and e_2 are unit vectors, and O is the origin of the coordinate.

Step 3: If $f(y^{m+1}) \geq f(x^k)$ (it means axial search mode fails and that process should be repeated until it succeeds), go to Step 4. If $f(y^{m+1}) < f(x^k)$, go to Step 5.

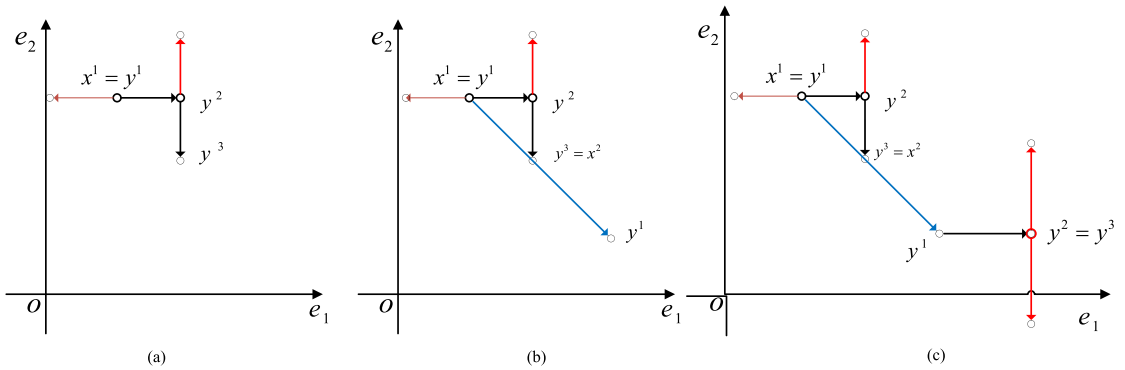


Fig. 5. Axial search and PS in a 2-D coordinate system.

Step 4: Decrease the step size $\chi = \chi \times \beta$ and go back to Step 2.

Step 5: Start the PS mode. Let $x^{k+1} = y^{m+1}$ and the direction of $(x^{k+1} - x^k)$ is probable to decrease the function value. Therefore, PS will be along that direction, which is $y^1 = x^{k+1} + \eta \times (x^{k+1} - x^k)$ [shown in Fig. 5(b) in 2-D coordinate system].

Step 6: Do the same as Step 2 [shown in Fig. 5(c) in a 2-D coordinate system].

Step 7: Let $k = k + 1$. If $f(y^{m+1}) < f(x^k)$ (it means the PS mode is successful and another PS mode can begin with an acceleration at the point of y^{m+1} obtained in Step 6), go to Step 9. If $f(y^{m+1}) \geq f(x^k)$ (it means PS mode is failed and another iteration of VSPS can begin at the point of y^{m+1} obtained in Step 2), go to Step 8.

Step 8: Reduce the step size $\chi(\chi = \chi \times \beta, \eta = \eta \times \beta)$, and go back to Step 2.

Step 9: Increase the step size $\chi = \chi \times \lambda$, and go to Step 10.

Step 10: If $\chi < \varepsilon, i \in [1, m]$, then stop. If not, go back to Step 2.

The VSPS algorithm improves the optimization process of detecting move by enabling the detection of step length in each direction to change differently according to the variations of objective function value in all directions, which makes the search direction closer to the optimal pattern in the descent direction. The acceleration factor is shrinking with the search conducted, which is helpful to search more in detail and is not easy to skip the optimal point.

V. Case Studies

A. Case 1: Engine Block Top Surface

This section validates the proposed classification system based on measurement data from engine block top surfaces. The measurement is conducted by a laser holographic interferometer (ShaPix Detective [52]), which is capable of gathering up to one million data points over a surface area of 300 × 300 mm^2 within 1 min. The basic height (Z) accuracy of it is 1 μm, while the lateral (X, Y) resolution is around 0.3 mm.

Six surface samples are selected from different engine blocks (shown in Fig. 6) manufactured in different conditions. It is apparent that surface samples 1–3 (called set A) are machined in one condition, while surface samples 4–6 (called set B) are machined in another condition. ShaPix is used to scan these six surfaces and collect 3-D coordinate data from each surface. Two samples of the color-coded measurement results from each data set are shown in Fig. 7.

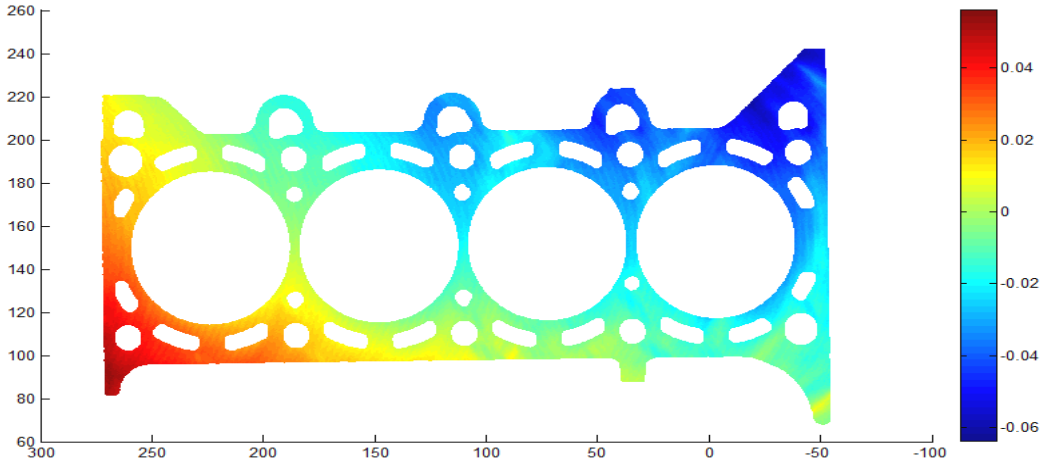
Fig. 6. Engine cylinder block.

It is obvious that the height of both sets decreases from left to right while the former set has a larger fluctuation range in the surface than the latter. As the car engine block is a critical functional part, low precision will lead to poor engine performance such as leakage or bore distortion. Therefore, it is imperative to identify the defective blocks from the normal ones through classification.

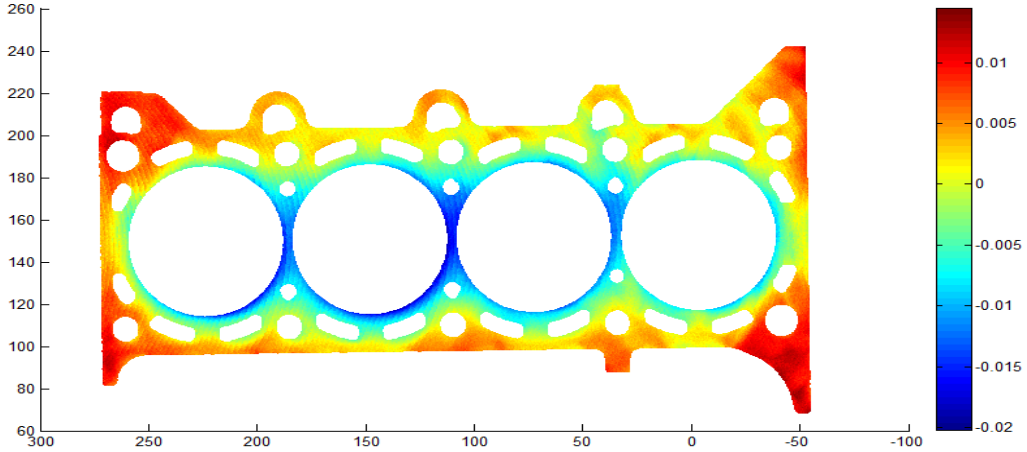
B. Feature Extraction

Data could be obtained by ShaPix over an entire surface containing about 790 000 points and it is necessary to extract representative features first. Several square areas are selected, for example, 20 (shown in Fig. 8), with the size of 128 by 128 over an area of 16 384 mm^2 to replace the whole surface. In addition, the selected areas should be the places with no holes on the surface and should be of the same location of each surface. These locations cover most critical areas that impact the surface functional performance during assembly.

To simplify the illustration, six patches of the same location from the six initial surfaces are selected, which are used as the input for NSCT. Fig. 9 shows the six surfaces chosen for classification. Surfaces 1–3 exhibit a similar pattern in spatial data and surfaces 4–6 exhibit a different spatial distribution in height data. This is evident as the former three surfaces are



Surface 1 from set A



Surface 4 from set B

Fig. 7. Two samples of the color-coded measurement results from data sets A and B.

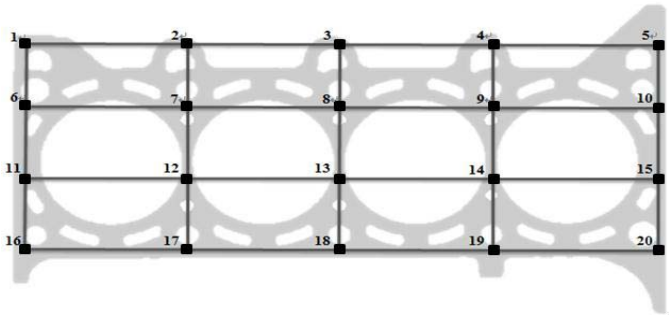
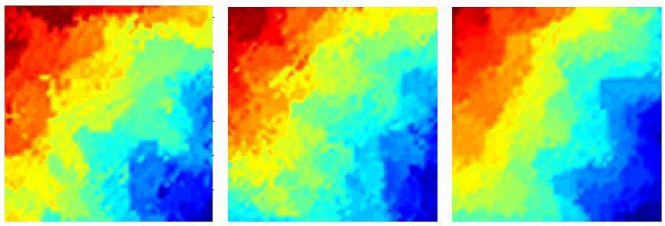
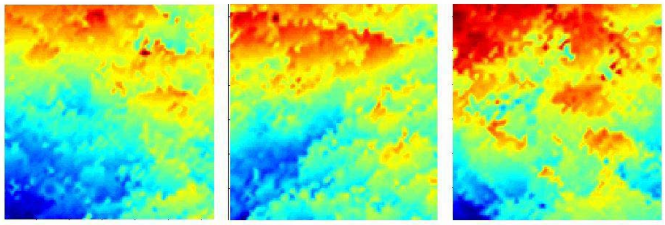


Fig. 8. Optimization of small surface samples through a grid chart.



Surface 1 to 3(from left to right)



Surface 4 to 6(from left to right)

Fig. 9. Six surfaces selected from sets A and B.

from one set, while the latter three are from another. And it is also apparent that surfaces in different sets look different because they are machined by tools in different conditions.

To extract features, we first partition each surface into $4 \times 4 = 16$ equal small surfaces with a size of 32 by 32 over an area of 1024 mm^2. As there are six original surfaces, we can get $16 \times 6 = 96$ samples. Then three-level NSCT decomposition is applied to extract feature vectors, which contains the means and standard deviations of every coefficient matrix of the wavelet sub-bands. Then, the final output of feature extraction is a matrix containing 96 56-D feature vectors.

TABLE I
CCP OF FOUR TYPES OF KERNEL FUNCTIONS

Kernel function	Linear kernel	Polynomial kernel	RBF	Sigmoid kernel
CCP (%)	87.5	91.67	93.75	87.5

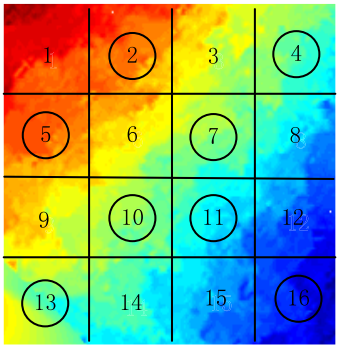


Fig. 10. Generation of training samples and testing samples.

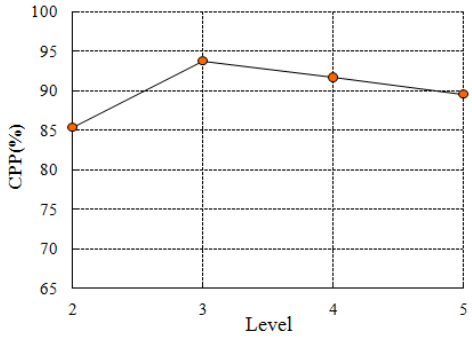


Fig. 11. CCPs using NSCT with different levels.

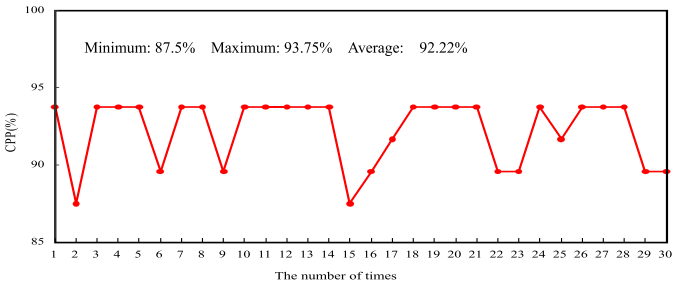


Fig. 12. Classification results of the proposed adaptive SVM-based system.

C. Classification Results

In Fig. 10, 1–16 represents 16 surface samples. In terms of each original surface sample, eight circled surface samples are randomly selected as training samples and the other eight ones are testing samples.

Since there are six surfaces in this case, the total number of training samples and testing samples are both $6\times 8 = 48$. As to the selection of the kernel function, there are four common kernel functions: 1) linear kernel function; 2) polynomial kernel function; 3) RBF; and 4) sigmoid kernel function. In fact, the RBF is the most widely used kernel function and the reasons are as follows: 1) the RBF can realize the nonlinear mapping; 2) the number of parameters that affects the model complexity in the RBF is few; and 3) RBF kernel has less numerical difficulties.

Table I shows the correct classification percentage (CCP) for the four types of kernel functions. The results represent that the CCP of the RBF is the highest and it is reasonable to be chosen as the kernel function of the SVM. It is, however, worth noting that although the kernels presented in Table I are the most popular, multiple other functions (such as Laplace and Fourier) can also be used for the classification task as well. This paper focuses on the optimization of parameters of the SVM rather than the selection of kernel functions, and therefore, other types of functions are not compared here.

The level of NSCT has an effect on the classification accuracy. To study the influence of different decomposition levels on the classification accuracy, levels from 2 to 5 are adopted, respectively. For the purpose of comparison, each level of NSCT implemented to extract features should use the same data set. The final classification results of each level are shown in Fig. 11.

Fig. 12 shows the classification results for the six engine cylinder block surfaces (surface areas are of the same location from each surface).

To obtain a more accurate result, the experiment is repeated 30 times. The CCP is used as an evaluation index. The proposed system finds the CCP of 93.75% repeatedly (19/30 in the experiments), which shows good performance of stability. The average CCP is 92.22%.

Table II presents the final results of 30 experiments by the proposed adaptive SVM-based classification system, which contains the classification accuracy and corresponding parameter combination. It can be seen that although the accuracies are the same, the parameter combinations are quite different. The detailed analysis will be discussed in the following section.

To analyze the time duration of the proposed algorithm, it was implemented in the MATLAB R2012b programming environment on a PC with Intel Core2 Duo T6600 CPU running Windows 7. The typical time complexity of the SVM is between $O(\mathrm{Nsv}^3 + \mathrm{LNsv}^2 + \mathrm{dLNsv})$ and $O(dL^2)$, in which Nsv is the number of support vectors, L is the number of training samples, and d is the dimensional number (the original dimensional number that has not been mapped to high-dimensional space) of each sample. The time complexity of APSO-VSPS is $\max\{O_{\mathrm{APSO}}, O_{\mathrm{VSPS}}\} = O_{\mathrm{APSO}} = N^* m^* T_T$, in which N is the number of particles, m is the maximum

TABLE II
RESULTS OF THE PARAMETERS AND CCP BY THE PROPOSED ADAPTIVE SVM-BASED SYSTEM

Index	1	2	3	4	5	6	7	8	9	10
C	104.03	244.06	216.68	292.57	153.53	105.09	153.53	93.23	108.42	45.64
σ	4.57	9.49	1.15	1.19	1.94	8.96	1.94	4.4	7.87	5.52
CCP	93.75	87.5	93.75	93.75	93.75	89.58	93.75	93.75	89.58	93.75
Index	11	12	13	14	15	16	17	18	19	20
C	75.07	242.3	174.78	257.03	277.36	225.89	120.45	163.66	137.57	28.02
σ	4.66	1.53	1.81	1.85	8.91	6.28	2.01	2.25	2.06	9.71
CCP	93.75	93.75	93.75	93.75	87.5	89.58	91.67	93.75	93.75	93.75
Index	21	22	23	24	25	26	27	28	29	30
C	57.45	150.42	233.79	46.28	40.15	249.95	50.64	263.05	269.68	204.88
σ	5.43	9.78	6.05	6.14	6.1	1.03	5.91	1.41	3.06	3.76
CCP	93.75	89.58	89.58	93.75	91.67	93.75	93.75	93.75	89.58	89.58

TABLE III
COMPUTATIONAL TIME OF THE PROPOSED METHOD

	Time (unit: second)	Standard deviation
Feature Extraction[1]	0.0459	0.0009
Parameters optimization[2]	7.0856	0.7663
Training[3]	0.0349	0.0060
Classification (testing)[4]	0.0081	0.0012

1: Time spent on extracting features of a surface sample.

2: Time spent on optimizing parameters of penalty coefficient and kernel function.

3: Time spent on training classifiers with the parameters optimized by APSO-VSPS (the training set includes 48 surface samples).

4: Time spent on classifying an unlabeled surface sample (input is extracted feature vector) with the selected classifiers obtained by training.

TABLE IV
CCP BY DIFFERENT PARAMETER COMBINATIONS OF C AND σ

σ	C					
	10	20	50	100	200	300
1	87.5	87.5	87.5	87.5	89.58	93.75
2	87.5	87.5	89.58	89.58	93.75	89.58
5	87.5	87.5	93.75	91.67	89.58	89.58
10	87.5	89.58	91.67	89.58	87.5	87.5
50	93.75	89.58	87.5	87.5	85.42	85.42
100	89.58	87.5	87.5	85.42	91.67	89.58
200	87.5	87.5	85.42	91.67	89.8	87.5
300	87.5	87.5	85.42	89.58	87.5	87.5

number of Iterations, and $T_T T_T$ is the computational time of each particle in every iteration.

To evaluate the actual duration of the implemented algorithm, statistical methods including mean and standard deviation values are used. The former determines the most probable execution time, while the latter gives the information about the size of outlayers caused by the operating system's work regime [53]. The training time, testing time, and parameters optimization time of the presented method in this case study are given as follows, both of which are the average of 30 times. The time spent on feature extraction is also provided in Table III. The total implement time cost is about 7.1745 s.

D. Sensitivity Analysis

1) Sensitivity Analysis of the Range of Parameters in APSO-VSPS: Table IV presents that the optimization of parameters has a great impact on the classification performance.

The parameter of the penalty coefficient (C) controls the extent of the punishment for classifying wrong samples and the value of C denotes the degree of punishment for misclassified samples. If value of C tends to infinity, all constraints must be satisfied, which means that all the training samples should to be correctly classified and would result in the complexity of the classification of hyperplane and intensive computation. From Table IV, it can be seen that the classification accuracy shows a trend that it first increases and then decreases on

TABLE V
FREQUENCY OF THE CCP BY APSO-VSPS ALGORITHM

Times of value number		Combination of initial step-length in VSPS					
		(0.1,0.1)	(1,1)	(1,9)	(10,2)	(10,10)	(20,3)
CCP	87.5	14	8	5	6	4	2
	89.58	4	10	3	8	8	7
	91.67	4	4	3	5	1	2
	93.75	8	8	19	11	17	19

the whole with the increasing value of the penalty coefficient (C). The selection of C is a compromise since the optimal number of support vectors and CCP can be obtained when C reaches a certain value; if value of C continues to increase, it will only add the training time and have no improvement for the performance of SVM. The experiments with the σ parameter show that as long as it tends to be zero and the vector of Lagrange multiplier is nonnegative, all training samples are support vectors and the SVM classifier classifies the whole training set correctly. If σ tends to be infinite, which indicates that more training samples will be in the interval (between two different classes of data) or be misclassified, the classification accuracy of training samples by SVM is zero, and SVM classification performance will drop substantially since all the samples are classified as the same class. If value of σ is appropriate, the number of support vectors can decrease and the ability to classify test samples will be greatly enhanced. Therefore, choosing the appropriate parameter combination together can construct a classifier with high performance. Besides, for datasets obtained from different kinds of workpiece surfaces, the parameter combination should be reoptimized because the ranges are not the same.

2) Sensitivity Analysis of the Step Length of Pattern Search in APSO-VSPS: If the length of the step size during the normal PS is the same in all directions, the deviation between the direction of pattern movement and optimal descent direction is large, limiting the convergence speed and accuracy. Thus, a nonmonotonic and VSPS method is adopted to change the step length and make the direction of the pattern move closer to the optimal descent direction. However, the initial step length should be selected appropriately as large step length may lead to a skip out of the optimal point and small step length may lead to a stuck in local convergence that is not globally optimal. Table V represents the frequency of the CCP by APSO-VSPS with different lengths of step size (operation for 30 times totally).

3) Comparison Analysis for APSO, APSO-PS, and APSO-VSPS: Table VI presents the comparison of APSO, APSO-PS, and APSO-VSPS in the aspects of maximum, minimum, and average and standard deviation of the CCP.

From Table VI, it can be seen that the CCP of APSO-VSPS is much higher than that of APSO since PS performs well in local search, which does good to the performance of global search by APSO. Besides, it also can be seen that although the standard deviation, and maximum and minimum CCPs of APSO-PS are equal to those

TABLE VI
COMPARISION OF APSO, APSO-PS, AND APSO-VSPS

CCP (%)	Maximum	Minimum	Average	Standard deviation
APSO	91.67%	85.42%	87.43%	0.027
APSO-PS	93.75%	87.5%	89.65%	0.021
APSO-VSPS	93.75%	87.5%	92.22%	0.021

TABLE VII
CCP OF THE STANDARD SVM

	Classification result						CCP
Pattern	1	2	3	4	5	6	
1	**8**	0	0	0	0	0	100
2	0	**8**	0	0	0	0	100
3	0	**2**	**6**	0	0	0	75.0
4	0	0	0	**5**	**3**	0	62.5
5	0	0	0	0	**7**	**1**	87.5
6	0	0	0	0	0	**8**	100
average							87.5

TABLE VIII
CCP OF THE PROPOSED ADAPTIVE SVM METHOD

	Classification result						CCP
Pattern	1	2	3	4	5	6	
1	**8**	0	0	0	0	0	100
2	0	**8**	0	0	0	0	100
3	0	**1**	**7**	0	0	0	87.5
4	0	0	0	**7**	**1**	0	87.5
5	0	0	0	0	**7**	**1**	87.5
6	0	0	0	0	0	**8**	100
average							93.75

of APSO-VSPS, the average CCP of APSO-VSPS is higher than that of APSO-PS because APSO-VSPS adopts VSPS, which is uneasy to step out of the primal point. Therefore, APSO-VSPS can perform better. According to the comparisons, APSO-VSPS is superior to APSO and APSO-PS.

E. Comparisons of the Proposed Adaptive SVM Method and Other Methods

1) Comparison of the Standard SVM and Proposed Adaptive SVM Methods: Weka [54] with Libsvm toolbox [55] is chosen as a software tool for computation. The standard SVM is used to classify the six patterns that represent for six different surfaces (1–3 belong to set A while 4–6 belong to set B) in detail. Here, the default parameters of $C = 1, \sigma = 0$ are selected for the standard SVM. Tables VII and VIII show

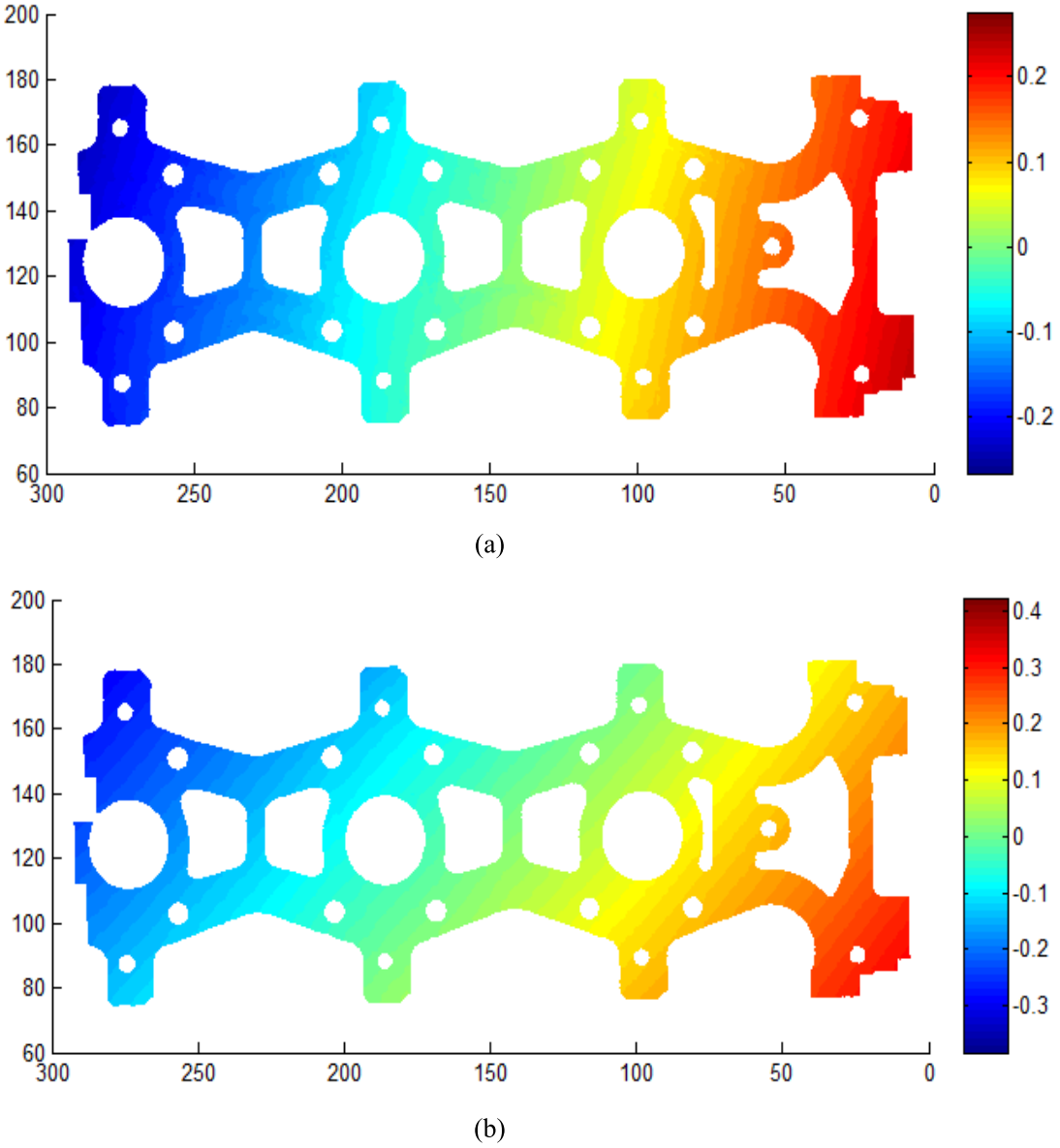


Fig. 13. Height map of two different cylinder heads. (a) New cylinder head. (b) Used cylinder head.

the results of CCPs of the standard SVM and the proposed adaptive SVM, respectively.

According to Table VII, classifications of several patterns are entirely correct and the final classification result is the average CCP of six patterns. There also exists a phenomenon that misclassification is more likely to occur in the same set as they are too similar to differentiate.

Comparing Table VII to Table VIII, it can be observed that the classification result of the proposed adaptive SVM method is 6.25% higher than that of the standard SVM and proposed adaptive SVM methods can increase the classification accuracy of misclassified patterns in Table VII to some extent. It is also apparent that the proposed adaptive SVM method can improve the CCP of patterns with much misclassification (pattern 3 with two misclassifications and pattern 4 with four misclassifications), but it cannot improve the CCP of misclassified patterns to 100%.

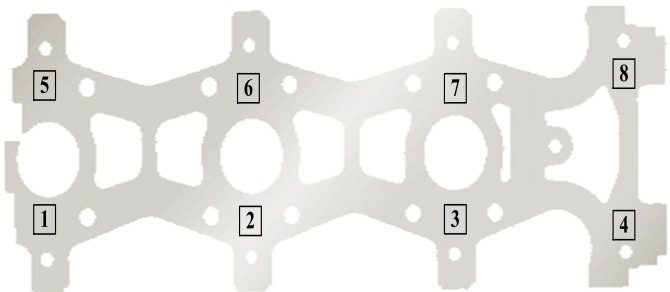


Fig. 14. Eight locations selected from the surface.

2) Comparison of the Proposed Adaptive SVM Method and Other Classification Methods: There are several other classic classification methods such as bagging-SVM [56], random subspace-SVM [57], naive Bayesian [58], Adaboost.M2-SVM [59], logistic regression [60],

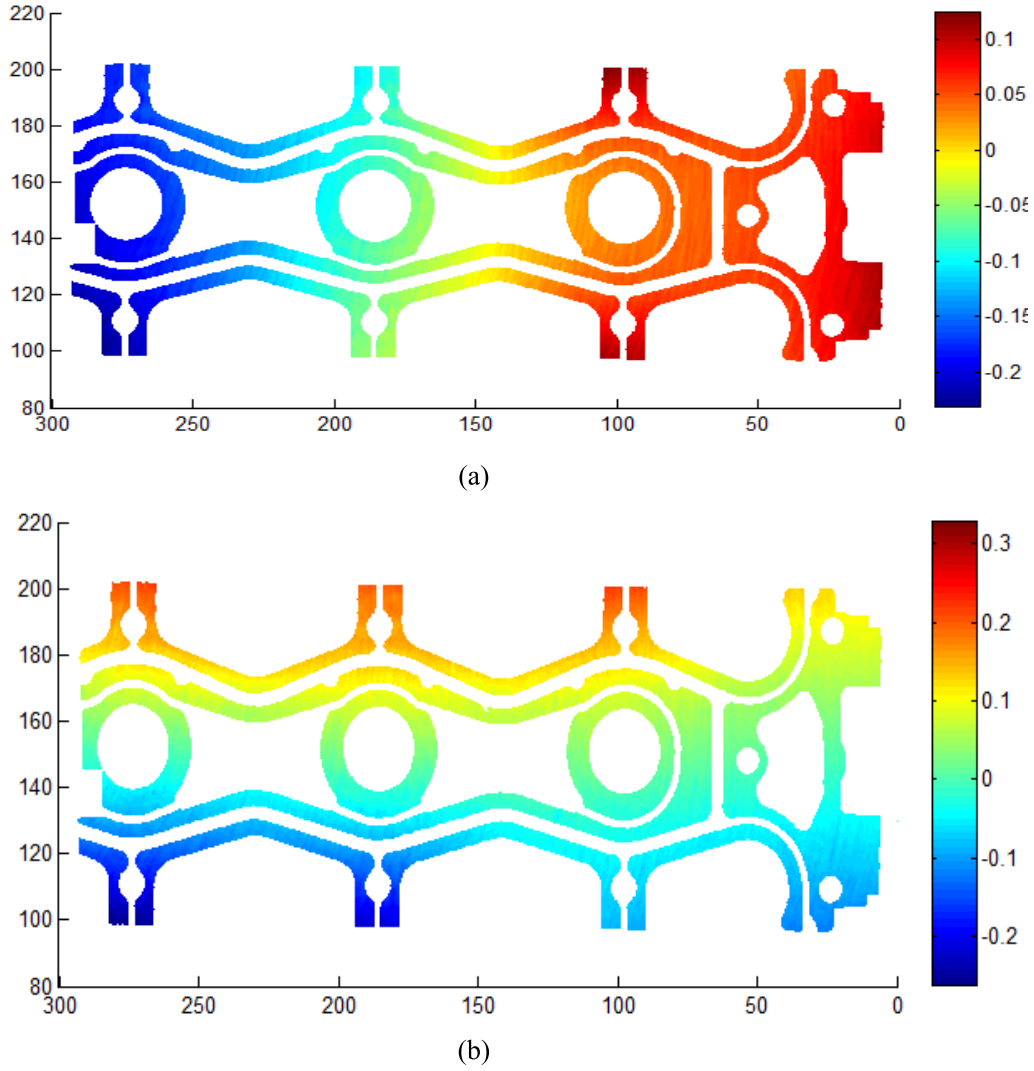


Fig. 15. Height map of two different cam carriers. (a) New cam carrier. (b) Used cam carrier.

neural network [61], fuzzy logic [62], and rough sets [63]. The performances of some methods are greatly affected by their parameters and architectures. For instance, rough sets require setting discretization and reduction methods, and neural network has several specific architectures (multilayered perceptron, RBF network, etc.) to choose. Due to limited space, the comparisons with the rough sets and neural network methods are not provided in this paper. The classification results of the other strategies are shown in Table IX.

According to Table IX, it can be observed that the proposed adaptive SVM method outperforms other methods in most of the 30 cases and the proposed adaptive SVM method achieves the highest average CCPs in all the methods. This proves that the proposed adaptive SVM method has a better performance comparing with those commonly used methods. Due to a large production volume in engine production, an improvement of 1% in CCP could lead to a significant savings of cost incurred by potential scraps.

TABLE IX
CCP OF DIFFERENT STRATEGIES (AVERAGE OF 30 TIMES)

Strategy		CCP (%)
Other methods	Bagging-SVM	90.69
	Random Subspace-SVM	90.97
	AdaBoost.M2-SVM	85.56
	Naive Bayesian	87.50
	Logistic regression	87.50
	Fuzzy logic	89.50
Proposed method	the proposed adaptive SVM method	93.75

F. Case 2: Cylinder Head Cover Surface

The second workpiece surface is the surface of a cylinder head made of aluminum alloy. The height map of the two different (new and used) surfaces is shown in Fig. 13 and the units of the x-axis, y-axis, and height are millimeters. Eight locations are selected (locations 1–8, shown in Fig. 14) from the surface with the same size and then eight small surfaces are obtained.

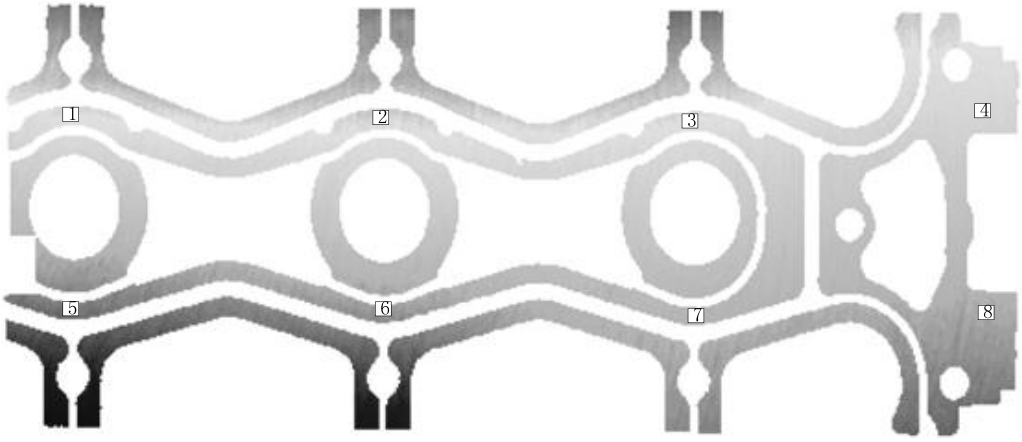


Fig. 16. Eight locations selected from this surface.

TABLE X
CCP OF THE PROPOSED ADAPTIVE SVM METHOD

	Classification result						CCP
Pattern	1	2	3	4	5	6	
1	**8**	0	0	0	0	0	100
2	0	**8**	0	0	0	0	100
3	0	**1**	**7**	0	0	0	87.5
4	0	0	0	**7**	**1**	0	87.5
5	0	0	0	0	**8**	0	100
6	0	0	0	0	0	**8**	100
average							95.83

TABLE XI
CCP OF THE PROPOSED ADAPTIVE SVM METHOD

	Classification result						CCP
Pattern	1	2	3	4	5	6	
1	**8**	0	0	0	0	0	100
2	0	**8**	0	0	0	0	100
3	0	**2**	**6**	0	0	0	75
4	0	0	0	**7**	**1**	0	87.5
5	0	0	0	0	**8**	0	100
6	0	0	0	0	0	**8**	100
average							93.75

To classify the cylinder head cover surfaces, the proposed adaptive SVM-based classification method is adopted and the same process should be done as shown in case 1. The classification accuracy is 95.83%, which is higher than that of case 1 since the samples in this case are more apparent in difference. An optimal set of parameters is $C = 214.28$, $\sigma = 6.91$. Table X presents the classification results of the proposed adaptive SVM-based classification system, where patterns 1–6 represent six different surfaces (1–3 from one set and 4–6 from another).

G. Case 3: Pump Valve Plate Top Surface

The third workpiece surface is the surface of a cam carrier. The height map of the two different (new and used) surfaces is shown in Fig. 15 and the units of the x-axis, y-axis, and height are millimeters. Eight locations are selected (locations 1–8, shown in Fig. 16) from the surface with the same size and then eight small surfaces are obtained.

To classify the cam carrier cover surfaces, the proposed adaptive SVM-based classification system is adopted and the same process should be done as shown in Case 1. The classification accuracy is 93.75% and an optimal set of parameters is $C = 91.54$, $\sigma = 4.05$. Table XI presents the classification results of the proposed adaptive SVM-based classification system, where patterns 1–6 represent six different surfaces (1–3 from one set and 4–6 from another).

VI. CONCLUSION

A novel adaptive system based on SVM is proposed to classify workpiece surfaces using HDM. In this classification system, an NSCT is first used to extract features and an adaptive SVM-based classification method is proposed, which consists of the selection of appropriate parameters C and σ to build a SVM classifier for the purpose of classification. Three case studies are conducted to evaluate the classification performance. The results demonstrate that the proposed adaptive SVM-based classification system can not only enhance the capability of APSO algorithm in global search but also reduce the likelihood that the identified solution being trapped in local optimum. The proposed adaptive SVM-based classification system is effective and efficient in the classification of workpiece surfaces.

The proposed adaptive SVM-based classification system can be applied in the following three aspects.

1) *Process Monitoring:* The proposed system shows a good classification performance in workpiece surface. If the

categorized defective workpieces account for a high proportion of the total amount, it is likely that problems exist in the manufacturing process.

2) *Fault Diagnosis:* The proposed system can be applied to classification and recognition of the fault types of detected faults by taking advantage of clear physical explanation of the workpiece surface classes one practitioner has created.
3) *Machine Tool Condition Prediction:* Classification of workpiece surfaces provides an important indicator of abnormal machine tool conditions, such as chatter, wear, and breakage. A classifier built by the proposed system can classify the workpiece surfaces machined by machine tool under abnormal conditions into different classes and a certain class is corresponded to a certain condition of machine tool. Once a workpiece surface machined by a certain tool is classified as one class with known conditions, the machining tool condition can be predicted for diagnosis.

There is a limitation in the proposed system. The parameters of SVM classifier need to be optimized again with the proposed system when classifying another new kind of workpiece surfaces, which means that the proposed system could not provide a constant SVM classifier that is optimal for different kinds of workpiece surfaces.

REFERENCES

[1] D. J. Whitehouse, "Surface metrology," *Meas. Sci. Technol.*, vol. 8, no. 9, pp. 955–972, 1997.

[2] Z. Huang, A. J. Shih, and J. Ni, "Laser interferometry hologram registration for three-dimensional precision measurements," *J. Manuf. Sci. Eng.*, vol. 128, no. 4, pp. 1006–1013, 2006.

[3] M. Wang, L. Xi, and S. Du, "3D surface form error evaluation using high definition metrology," *Precis. Eng.*, vol. 38, no. 1, pp. 230–236, 2014.

[4] J. Leopold, H. Günther, and R. Leopold, "New developments in fast 3D-surface quality control," *Measurement*, vol. 33, no. 2, pp. 179–187, 2003.

[5] H. Liu, Y. Shi, L. Yin, W. Jiang, and B. Lu, "Roll-to-roll imprint for high precision grating manufacturing," *Eng. Sci.*, vol. 11, no. 1, pp. 39–43, 2013.

[6] L. Zhou, H. Wang, C. Berry, X. Weng, and S. J. Hu, "Functional morphing in multistage manufacturing and its applications in high-definition metrology-based process control," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 1, pp. 124–136, Jan. 2012.

[7] H. T. Nguyen, H. Wang, and S. J. Hu, "Characterization of cutting force induced surface shape variation in face milling using high-definition metrology," *J. Manuf. Sci. Eng.*, vol. 135, no. 4, pp. 041014-1–041014-12, 2013.

[8] H. T. Nguyen, H. Wang, and S. J. Hu, "Modeling cutter tilt and cutter-spindle stiffness for machine condition monitoring in face milling using high-definition surface metrology," *Int. J. Adv. Manuf. Technol.*, vol. 70, nos. 5–8, pp. 1323–1335, 2014.

[9] L. Blunt and X. Jiang, *Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards*. London, U.K.: Kogan Page Science, 2003.

[10] K. V. Ramana and B. Ramamoorthy, "Statistical methods to compare the texture features of machined surfaces," *Pattern Recognit.*, vol. 29, no. 9, pp. 1447–1459, 1996.

[11] W. P. Dong, P. J. Sullivan, and K. J. Stout, "Comprehensive study of parameters for characterising three-dimensional surface topography: IV: Parameters for characterising spatial and hybrid properties," *Wear*, vol. 178, nos. 1–2, pp. 45–60, 1994.

[12] D.-M. Tsa and S.-K. Wu, "Automated surface inspection using Gabor filters," *Int. J. Adv. Manuf. Technol.*, vol. 16, no. 7, pp. 474–482, 2000.

[13] M. Zhang, E. Levina, D. Djurdjanovic, and J. Ni, "Estimating distributions of surface parameters for classification purposes," *J. Manuf. Sci. Eng.*, vol. 130, no. 3, pp. 031010-1–031010-9, 2008.

[14] S. Fu, B. Muralikrishnan, and J. Raja, "Engineering surface analysis with different wavelet bases," *J. Manuf. Sci. Eng.*, vol. 125, no. 4, pp. 844–852, 2003.

[15] Y. Liao, D. A. Stephenson, and J. Ni, "Multiple-scale wavelet decomposition, 3D surface feature exaction and applications," *J. Manuf. Sci. Eng.*, vol. 134, no. 1, pp. 011005-1–011005-13, 2012.

[16] Y. Li and J. Ni, "B-spline wavelet-based multiresolution analysis of surface texture in end-milling of aluminum," *J. Manuf. Sci. Eng.*, vol. 133, no. 1, pp. 011014-1–011014-11, 2011.

[17] A. L. de Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[18] S. He, K. Li, and M. Zhang, "A real-time power quality disturbances classification using hybrid method based on S-transform and dynamics," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2465–2475, Sep. 2013.

[19] G. Panahandeh, N. Mohammadiha, A. Leijon, and P. Handel, "Continuous hidden Markov model for pedestrian activity classification and gait analysis," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 1073–1083, May 2013.

[20] Z. He, Q. Wang, Y. Shen, J. Jin, and Y. Wang, "Multivariate gray model-based BEMD for hyperspectral image classification," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 889–904, May 2013.

[21] W. Li, S. Zhang, and G. He, "Semisupervised distance-preserving self-organizing map for machine-defect detection and classification," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 869–879, May 2013.

[22] Z. He, Y. Shen, Q. Wang, and Y. Wang, "Optimized ensemble EMD-based spectral features for hyperspectral image classification," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 5, pp. 1041–1056, May 2014.

[23] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1999.

[24] S. R. Gunn, "Support vector machines for classification and regression," Faculty Eng., Sci. Math., School Electron. Comput. Sci., Univ. Southampton, Southampton, U.K., ISIS Tech. Rep., 1998.

[25] D. R. Salgado and F. J. Alonso, "Tool wear detection in turning operations using singular spectrum analysis," *J. Mater. Process. Technol.*, vol. 171, no. 3, pp. 451–458, 2006.

[26] S. Du, D. Huang, and J. Lv, "Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines," *Comput. Ind. Eng.*, vol. 66, no. 4, pp. 683–695, 2013.

[27] S. Du and J. Lv, "Minimal Euclidean distance chart based on support vector regression for monitoring mean shifts of auto-correlated processes," *Int. J. Prod. Econ.*, vol. 141, no. 1, pp. 377–387, 2013.

[28] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[29] R. J. Malak and C. J. J. Paredis, "Using support vector machines to formalize the valid input domain of predictive models in systems design problems," *J. Mech. Design*, vol. 132, no. 10, pp. 101001-1–101001-14, 2010.

[30] S. Du, J. Lv, and L. Xi, "On-line classifying process mean shifts in multivariate control charts based on multiclass support vector machines," *Int. J. Prod. Res.*, vol. 50, no. 22, pp. 6288–6310, 2012.

[31] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Netw.*, vol. 17, no. 1, pp. 113–126, 2004.

[32] J. Lin, J. Zhang, and J. Lin, "A fast parameters selection method of support vector machine based on coarse grid search and pattern search," in *Proc. 4th Global Congr. Intell. Syst. (GCIS)*, 2013, vol. 3, no. 4, pp. 77–81.

[33] P.-W. Chen, J.-Y. Wang, and H.-M. Lee, "Model selection of SVMs using GA approach," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2004, pp. 2035–2040.

[34] S.-W. Lin, K.-C. Ying, S.-C. Chen, and Z.-J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1817–1824, 2008.

[35] S. Du and L. Xi, "Fault diagnosis in assembly processes based on engineering-driven rules and PSOSAEN algorithm," *Comput. Ind. Eng.*, vol. 60, no. 1, pp. 77–88, 2011.

[36] Q. Huang, J. Mao, and Y. Liu, "An improved grid search algorithm of SVR parameters optimization," in *Proc. IEEE 14th Int. Conf. Commun. Technol. (ICCT)*, Nov. 2012, pp. 1022–1026.

[37] C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Syst. Appl.*, vol. 31, no. 2, pp. 231–240, 2006.

[38] S. L. Ho, S. Y. Yang, G. Z. Ni, and K. F. Wong, "An improved PSO method with application to multimodal functions of inverse problems," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1597–1600, Apr. 2007.
[39] Y. Li, Y. Peng, and S. Zhou, "Improved PSO algorithm for shape and sizing optimization of truss structure," *J. Civil Eng. Manage.*, vol. 19, no. 4, pp. 542–549, 2013.
[40] L. Xiao, W. Zhang, and Y. Zhang, "Particle swarm optimization with adaptive local search," *Comput. Sci.*, vol. 34, no. 8, pp. 199–201, 2007.
[41] W. Guo and G. Chen, "Fuzzy self-adapted particle swarm optimization algorithm for traveling salesman problems," *Comput. Sci.*, vol. 33, no. 6, pp. 161–185, Jun. 2006.
[42] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. IEEE Int. Conf. Evol. Comput.*, May 1998, pp. 69–73.
[43] Y. Shi and R. Eberhart, "Parameter selection in particle swarm optimization," in *Evolutionary Programming VII*. Berlin, Germany: Springer-Verlag, 1998, pp. 591–600.
[44] R. R. Coifman and D. L. Donoho, "Translation-invariant de-noising," in *Wavelets and Statistics*. Berlin, Germany: Springer-Verlag, 1995, pp. 125–150.
[45] S. Li, X. Fu, and B. Yang, "Nonsubsampled contourlet transform for texture classifications using support vector machines," in *Proc. IEEE Int. Conf. Netw., Sensing Control*, Apr. 2008, pp. 1654–1657.
[46] B. Yang, S. Li, and F. Sun, "Image fusion using nonsubsampled contourlet transform," in *Proc. 4th Int. Conf. Image Graph.*, Aug. 2007, pp. 719–724.
[47] J. Zhou, A. L. Cunha, and M. N. Do, "Nonsubsampled contourlet transform: Construction and application in enhancement," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2005, pp. 469–472.
[48] M. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, Oct. 1992.
[49] R. Hooke and T. A. Jeeves, "'Direct search' solution of numerical and statistical problems," *J. Assoc. Comput. Machin.*, vol. 8, no. 2, pp. 212–229, 1961.
[50] V. Torczon, "On the convergence of pattern search algorithms," *SIAM J. Optim.*, vol. 7, no. 1, pp. 1–25, 1997.
[51] C. Audet and J. E. Dennis, Jr., "Pattern search algorithms for mixed variable programming," *SIAM J. Optim.*, vol. 11, no. 3, pp. 573–594, 2001.
[52] *Introducing the Latest in High-Definition, Non-Contact Metrology Shapix 1500 Series*. [Online]. Available: http://www.coherix.com
[53] P. Bilski and W. Winiecki, "Methods of assessing the time efficiency in the virtual measurement systems," *Comput. Standards Interf.*, vol. 34, no. 6, pp. 485–492, 2012.
[54] *Waikato Environment for Knowledge Analysis*. [Online]. Available: http://weka.wikispaces.com
[55] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. ID 27.
[56] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.
[57] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.
[58] P. Domingos and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," *Mach. Learn.*, vol. 29, nos. 2–3, pp. 103–130, 1997.
[59] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Proc. 2nd Eur. Conf. Comput. Syst. Sci.*, 1995, pp. 23–37.
[60] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: A methodology review," *J. Biomed. Inform.*, vol. 35, nos. 5–6, pp. 352–359, 2002.
[61] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3642–3649.
[62] M. Thakkar, M. Bhatt, and C. K. Bhensdadia, "Fuzzy logic based computer vision system for classification of whole cashew kernel," *Comput. Netw. Inf. Technol.*, vol. 142, pp. 415–420, Jun. 2011.
[63] Y.-C. Hu, "Rough sets for pattern classification using pairwise-comparison-based tables," *Appl. Math. Modell.*, vol. 37, nos. 12–13, pp. 7330–7337, 2013.

Shi-Chang Du received the B.S. and M.S.E. degrees in mechanical engineering from the Hefei University of Technology, Hefei, China, in 2000 and 2003, respectively, and the Ph.D. degree in industrial engineering and management from Shanghai Jiao Tong University, Shanghai, China, in 2008.

He was with the University of Michigan, Ann Arbor, MI, USA, from 2006 to 2007, as a Visiting Scholar. He is currently an Associate Professor with the Department of Industrial Engineering and Management, Shanghai Jiao Tong University. He has authored 30 papers in journals, including the *ASME Journal of Manufacturing Science* and *Precision Engineering*. His current research interests include quality engineering, including quality control with analysis of variation flow, monitoring, and diagnosis of manufacturing process by support vector machine and Bayesian.

Dr. Du's research has been supported by the National Natural Science Foundation of China, the Ministry of Science and Technology of China, and the Shanghai Association for Science and Technology.

De-Lin Huang received the B.S. degree in control engineering from Northeast University, Qinhuangdao, China, in 2013. He is currently pursuing the Ph.D. degree with the Department of Industrial Engineering and Management, Shanghai Jiao Tong University, Shanghai, China.

His current research interests include the monitoring of product quality based on adaptive support vector machine.

Hui Wang received the B.S. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001, the M.S.E. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2003, and the Ph.D. degree in industrial engineering from the University of South Florida, Tampa, FL, USA, in 2007.

He is currently an Assistant Professor with Florida State University, Tallahassee, FL, USA, and an Adjunct Faculty with the University of Michigan. His research in recent years has been conducted in close collaboration with a metrology company and U.S. manufacturing industries, including Ford Motors, Dearborn, MI, USA, and General Motors, Detroit, MI, USA. His current research interests include manufacturing system design, automation, and process control by integrating applied statistics, image processing, optimization, and control theory with engineering knowledge with broad applications, including automotive, energy system, semiconductor, and nonmanufacturing.

Dr. Wang is a member of the IIE Transactions, the Institute of International Education, and the Institute for Operations Research and the Management Sciences. His research has received a number of best paper awards from the ASME Manufacturing Science and Engineering Conference and the International Conference on Frontiers of Design and Manufacturing, and featured article recognition of the IIE Transactions. His research has been sponsored by the U.S. National Institute of Standards and Technology and the National Science Foundation.