Classification and recognition of diffraction structures using support vector machine in optical scatterometry

Jinlong Zhu[a], Shiyuan Liu[*, a, b], Chuanwei Zhang[b], Xiuguo Chen[a], and Zhengqiong Dong[a]

[a] Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China;

[b] State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan, China.

ABSTRACT

The library search is a widely used method for reconstruction of diffraction structures in optical scatterometry. In library search, an optimized set of geometrical parameters for a diffraction structure can be achieved by searching for a best match between the measured signatures and the simulated ones. The search speed and accuracy is the key to guarantee the effectiveness of this method, and some a priori geometrical model is necessary. Once the actual geometrical model of a measured signature is different from the model used in the establishment of library, the search result will be meaningless. Therefore, the classification and recognition of the geometrical profile for a measured signature is critical. In this paper, we develop two support vector machine (SVM) classifiers to deal with issue. One classifier is used to identify the geometrical profile of a diffraction structure from its measured signature, and the other one is to map the whole search range of the identified diffraction structure into a smaller one. By using some reliable and mature search algorithms, we can fast and accurately reconstruct the geometry profile of a diffraction structure in this optimized small range. Simulation and experiment have demonstrated that the SVM classifiers can identify the geometrical profile of one-dimensional trapezoidal gratings accurately, and the SVM-based library search strategy can achieve a fast and accurate extraction of parameters for diffraction structures.

Keywords: scatterometry, metrology, critical dimension (CD), structure reconstruction, library search, support vector machine (SVM), classifier, ellipsometric signature

1. INTRODUCTION

Optical scatterometry is a non-contact, non-destructive and accurate technique that is now widely used in critical dimension (CD) metrology for sub-micro and nano structures. By analyzing the scattered optical signatures, the geometrical profiles of diffraction structures can be determined. In the analysis of optical signatures, since this type of problem is ill-posed, there is no analytical solution. To solve the ill-posed inverse problem in scatterometry, some nonlinear optimization approaches have been proposed, such as Levenberg-Marquardt (LM) algorithm [1, 2], artificial neural network (ANN) [3], and ANN-LM combined method [4]. However, the nonlinear optimization approach is usually time-consuming, as the structural profile is achieved through an iterative procedure that repeatedly requires computation of forward optical modeling. This is even worse and unacceptable when dealing with two-dimensional structures or more complex structures. Therefore, the library search method is still the most attractive technique in industry [5]. By simulating the optical signatures of a predetermined set of profiles and choosing the best candidate using search techniques, the geometrical parameters can be obtained accurately. However, several challenges arise for much more accurate and faster parameter extraction when the library grows larger and larger [6]. Currently, most of the effort to solve this problem is to design new matching cost functions [7, 8] and new searching strategies [9-11]. Comparing with designing cost functions, designing search strategies is a more direct method which aims to match not only accurately, but also fast.

Another issue in library search is that the geometrical profiles of the diffraction structures are often unknown before the search itself. If the corresponding actual structures are not the same as the model in the forward optical modeling, the

* Contact author: shyliu@mail.hust.edu.cn; phone: +86 27 8755 9543; webpage: http://www2.hust.edu.cn/nom.

Metrology, Inspection, and Process Control for Microlithography XXVI, edited by Alexander Starikov,
Proc. of SPIE Vol. 8324, 83242S · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.916259

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

matching procedure will lead to be inaccurate or even erroneous. Thereofer, for a measured signature its corresponding geometrical model should be identified before the search. Recently, Gereige *et al.* proposed a method by artificial neural network (ANN) to classify geometrical profiles according to their corresponding optical signatures [12]. However, if the real geometrical parameters of a structure corresponding to the measured signature are not in the range of training data, the mapping accuracy can be guaranteed.

Support vector machine (SVM) is a kind of machine learning algorithm based on statistical learning theory (SLT) [13, 14]. By taking SLT as system information, SVM can to some extent obtain an optimal result under limited information. In this paper, we introduce the SVM to deal with those challenges in the library search. We generate two SVM classifiers. One classifier is used to identify the geometrical profile of a diffraction structure from its measured ellipsometric signature, and then other one is used to map the whole search range of the identified diffraction structure into a smaller one. In the sub-library with a smaller range, we can use some search algorithm to fast match the best signature. We have demonstrated the feasibility of the SVM-based classifiers.

2. METHODS

2.1 Principle of SVM

By using the radial basis function (RBF) as the kernel, we can nonlinearly map the input signatures to a high dimensional feature space. Then in the high dimensional feature space we can construct an optimal separating hyperplane so that we can classify those signatures. For a binary classification problem, a set of training signatures are as follows:

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_l, y_l) \tag{1}$$

$$x_i \in R^n,\ y_i \in \{-1,\ 1\},\ i = 1,\ 2,\ \ldots,\ l \tag{2}$$

where x_l is a n-dimensional vector, and l is the number of training signatures. Training signatures are used in the training of SVM network.

For the classification of a measured signature x, the value of decision function f corresponding to x decides which class the input signature x belongs to. The decision function f can be defined by

$$f(x) = \mathrm{sign}\left[\omega \cdot \Psi(x) + b\right] \tag{3}$$

where ω is the support vector which is the linear combination of training data, b is the bias, and ψ is a RBF kernel defined by

$$\Psi(x) = \exp\left[-\frac{\left\| x - x_c \right\|^2}{2 \times \sigma^2}\right] \tag{4}$$

where x_c is the centre of kernel function, σ is the scaling factor. The ω in Eq. (3) can be expressed by

$$\omega = \sum_{i=1}^{l} \lambda_i y_i \Psi(x_i) \tag{5}$$

where λ_i is the weight coefficient of the ith input signature. By substituting Eq. (5) into Eq. (3), we can get the following equation

$$f(x) = \mathrm{sign}\left[\sum_{i=1}^{l} \lambda_i y_i \Psi(x_i) \cdot \Psi(x) + b\right] \tag{6}$$

Hence, as shown in Eq. (6), for the input signature x the decision function f can judge which class it belongs to.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

2.2 SVM-based library search strategy

In this paper, we divide the reconstruction of diffraction structures by library search into two parts. The first part is the identification for the profile of a diffraction structure by its measured ellipsometric signature (if the profile model of the structure is known, then this procedure can be skipped). Once the profile is identified, then the next step which is the search for a nearest neighbour of the measured signature in the library will be conducted. The flowchart of this method is shown in Figure 1.

For the identification stage, a SVM classifier is trained to recognize the geometrical profile of a structure by its measured ellipsometric signatures. For the binary classification task, we use numeric '1' to represent one kind of profile model, and numeric '0' to represent the other profile model. Thus, the output of classifier is '1' or '0'.

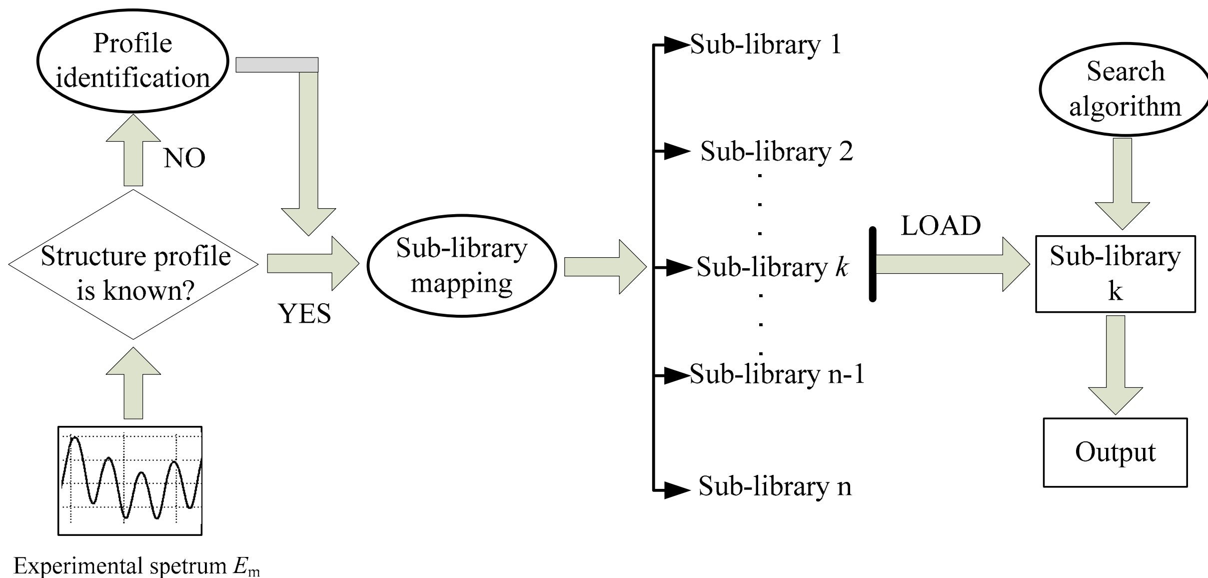


Fig. 1. Flowchart of SVM-based library search

Once the identification stage is finished, then the library search stage can be conducted. In the library search stage, instead of searching for a nearest neighbour of the measured ellipsometric signature by linear search in the whole library, we propose a SVM-based library search strategy which can reduce the searching time and at the same time keep the parameters extraction accuracy. In our SVM-based library search strategy, a multi-classification SVM classifier is trained to map the measured signature corresponding to the known structures which is identified by the first binary SVM classifier into a sub-library of the whole spectral library. In order to simplify our process, we will simulate the signatures to establish a spectral library for one-dimensional trapezoidal grating. However, the predefined model in the SVM classifier has nonnumeric information about the sub-libraries. Hence we need to translate the textual data into numerical form. Here we assume that the library is divided into M sub- libraries. The first sub-library is coded with a vector $[1_{(1)}\ldots 0_{(n)}\ldots 0_{(M)}]$ [12]. The sub-library n is coded with the vector $[0_{(1)}\ldots 1_{(n)}\ldots 0_{(M)}]$. The flowchart of the SVM-based library search method is as shown in Figure 1.

In consideration of the highly nonlinear separated ellipsometric signatures, we use RBF kernel. And in the training of SVM classifier, we use quadratic programming (QP) to search for the separating hyperplane. For the first SVM classifier which is used for identification, we will test its generalization ability. Hence a set of newly simulated testing ellipsometric signatures are used. The classification accuracy is the major target that we concern. There are many factors that affect the classification accuracy. To objectively and quantitatively estimate the relationship between those factors and classification accuracy is very difficult. Here we estimate the scaling factor in the RBF kernel and number of signatures' effect on classification accuracy.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

3. RESULTS AND DISCUSSION

3.1 Description of the Grating Models

For the purpose of geometrical profile's identification and fast extraction of geometrical parameters, simulation and experiment are conducted based on one-dimensional periodic structures. In our simulations, two grating models are used, as shown in Figure 2. The first one is a one-dimensional trapezoidal photoresist grating with a period of 400 nm deposited on a silicon substrate which is coated by an anti-reflective layer. The second one is a one-dimensional sinusoidal photoresist grating with the same materials and pitch as trapezoidal grating. In the experiment part, a one-dimensional trapezoidal photoresist grating with the same period and materials as discussed above is fabricated.

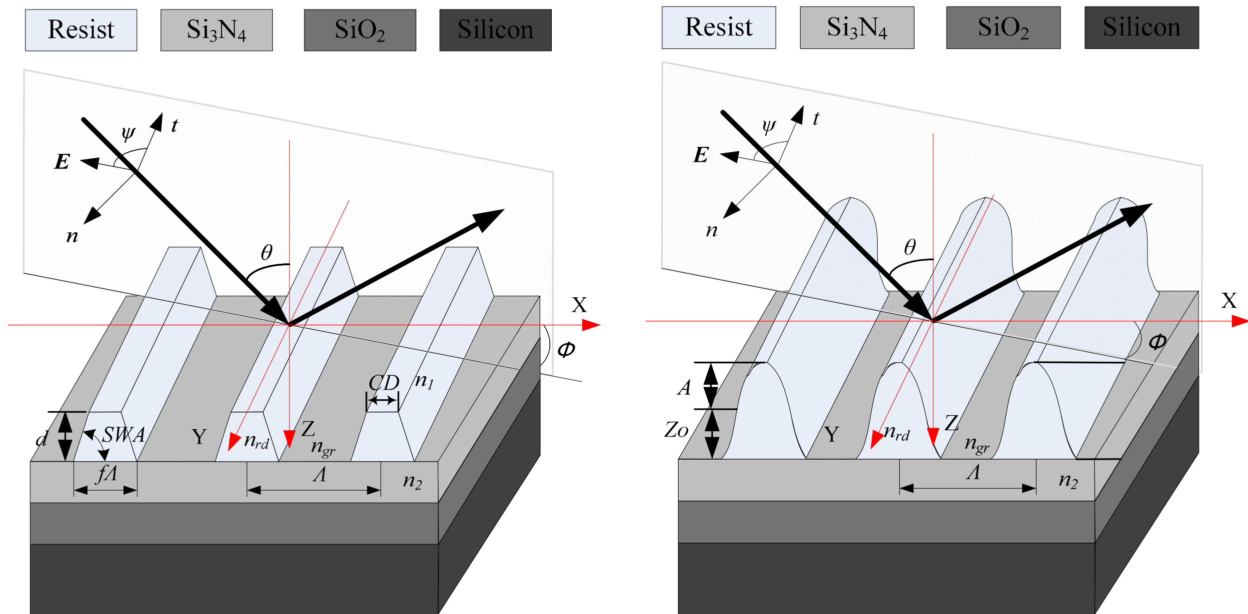


Fig. 2. Geometrical profiles of trapezoidal grating and sinusoidal grating. The left model is defined by three parameters: *d* (height), *SWA* (sidewall angle) and $f\Lambda$ (f is duty cycle, Λ is pitch). The right model is defined by two parameters: *A* (amplitude) and Z_0 (offset along Z axis).

3.2 Simulation Results

For the purpose of identifying the geometrical profiles, and in order to estimate the scaling factor in the RBF kernel and number of signatures' effect on classification accuracy, different sets of training ellipsometric signatures and different values of σ are generated, respectively. Each set is an equal mixture of signatures for trapezoidal and sinusoidal profiles. The geometrical parameters are randomly chosen from the following variation ranges: 290 < d < 320 nm, 150 < $f\Lambda$ < 190 nm, 86° < SWA < 90°, 160 < *A* < 180 nm, 160 < *Zo* < 180 nm for the trapezoidal and sinusoidal profiles. The incident wavelengths are between 380 nm and 780 nm, with a step of 10 nm and an incident angle θ=65° in the classical mounting (i.e., φ=0°).

Table 1. The corresponding ranges of the parameters for of each set

	Range of *D* (nm)	Range of $f\Lambda$ (nm)	Range of *SWA* (°)	Range of Z_o (nm)	Range of *A* (nm)	Remark
Set 1	[290 320]	[150 190]	[86 90]	[160 180]	[160 180]	No range exceeds
Set 2	[290 320]	[150 190]	[86 90]	[160 180]	[150 180]	*A* exceeds
Set 3	[290 320]	[150 190]	[86 90]	[150 180]	[160 180]	Z_o exceeds
Set 4	[280 320]	[150 190]	[86 90]	[160 180]	[160 180]	*d* exceeds
Set 5	[290 320]	[150 200]	[86 90]	[160 180]	[160 180]	$f\Lambda$ exceeds
Set 6	[290 320]	[150 190]	[80 90]	[160 180]	[160 180]	*SWA* exceeds

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

In order to test the generalization of the SVM classifier, we do not generate the testing signatures from the parameters ranges as discussed above. Here we generate six sets of testing ellipsometric signatures from the following parameters ranges, respectively (the number of signatures in each set is 400): As shown in Table 1, for the six sets of testing samples, we have five sets for each of whose one parameter range is not the same as the training samples (or exceeds). The testing results are shown in Fig. 3.

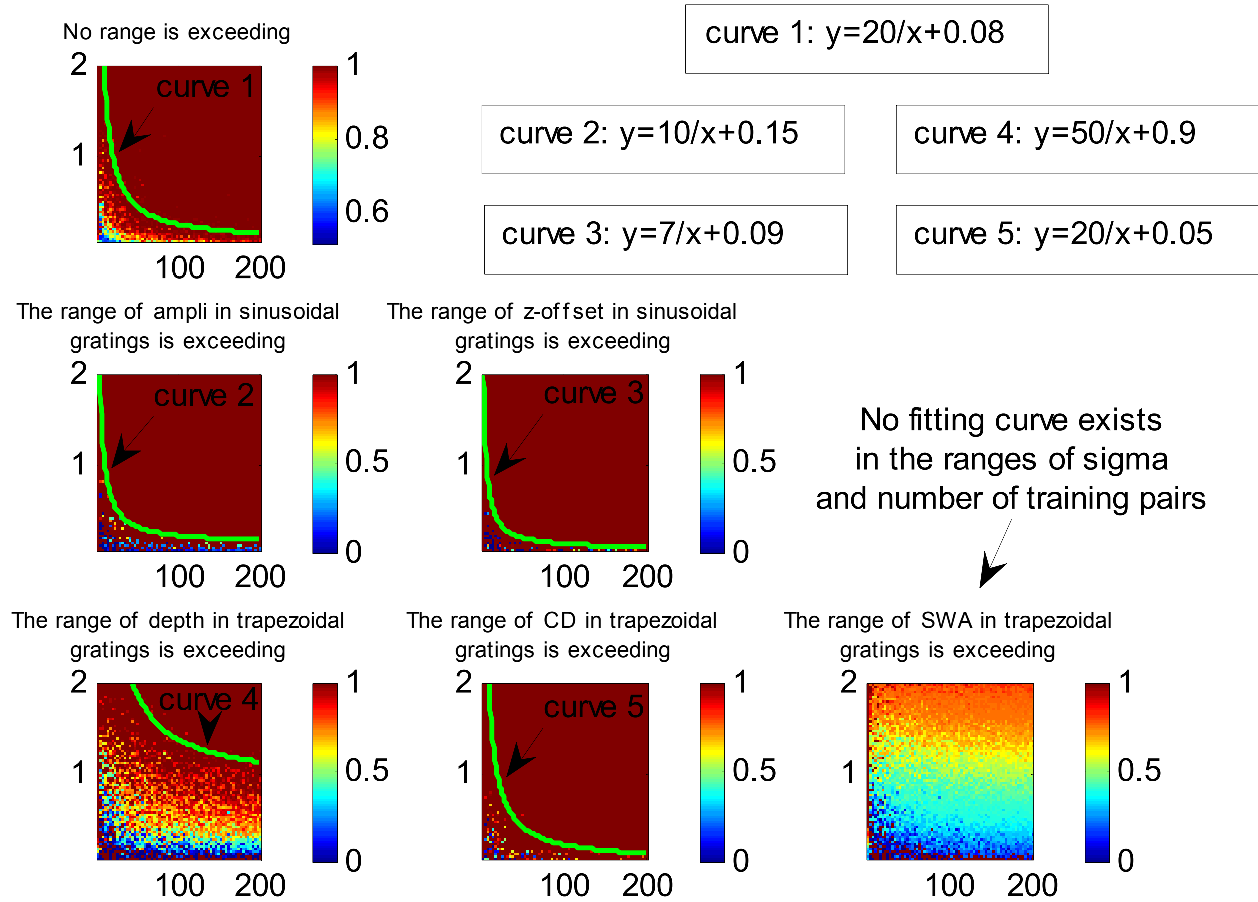


Fig. 3. Classification accuracy for trapezoidal grating and sinusoidal grating under different parameter ranges. The parameter range of the first sub-image in the top row is the same as training pairs.

In Figure 3, the X axis of the six sub-images corresponds to number of training signatures, and the Y axis corresponds to the value of sigma in RBF kernel. The ranges of the sigma and number of training signatures for the six sub-images are the same, which are: 0.1 < sigma < 2, 1 < number of training signatures < 200. The increment in X axis and Y axis is 1 and 0.02, respectively. We can find that for the five sub-images on the left in Figure 3 there exists a curve which separates the image according to classification accuracy. Those curves can be described as hyperbolas with different biases. If a point is above those curves, the classification is nearly100%. For a point under those curves, the classification accuracy is below 100% and unstable. Hence we indicate that the ability of generation of SVM classifier is excellent which can help to classify those geometrical profiles even their parameter values are not in the ranges of training pairs. Also, we can find that for the classification of trapezoidal gratings with the exceeding range of *d* (depth), more training pairs and bigger sigma are necessary in specified ranges when compared to any other aspect. However, there exists no hyperbola in the last sub-image. In the sixth sub-image with an exceeding range of *SWA* we can find that from the bottom left corner to top right corner the classification accuracy is arising. Hence we suppose that a bigger sigma and bigger number of training signatures can make the classification accuracy close to 100%.

In order to validate our guess, we simulated another image with a bigger range of sigma and bigger range of number of training signatures, as shown in Figure 4.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

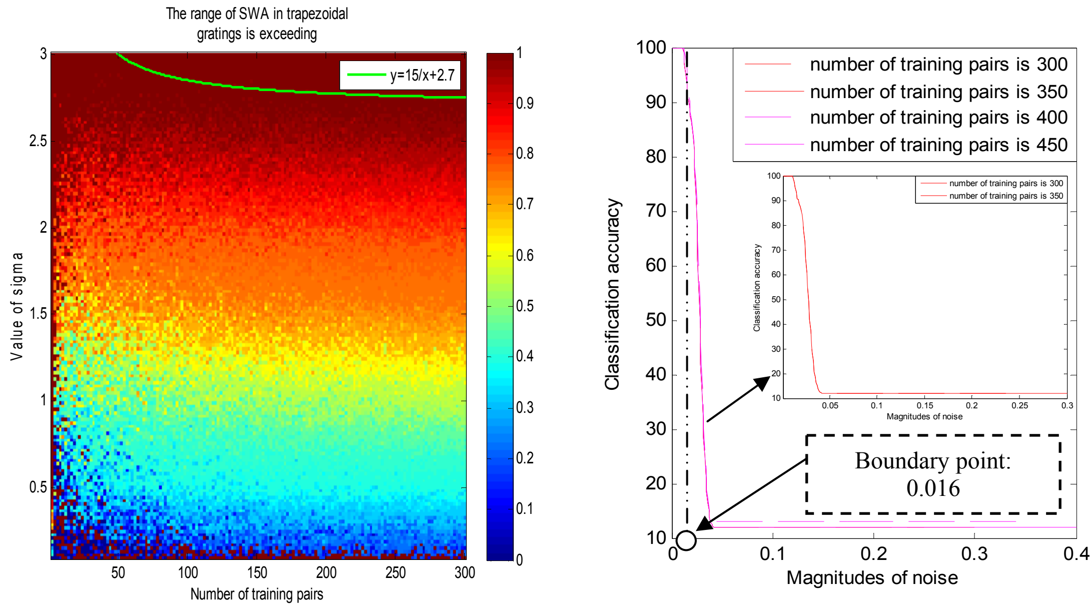


Fig. 4. Classification accuracy under different conditions. The left is a classification accuracy image for trapezoidal grating and sinusoidal grating under the exceeding range of *SWA*. The exceeding *SWA* range of the right sub-image is: 80° < *SWA* < 90°. The right is four classification accuracy curves with different number of training pairs under different magnitudes of noise.

The left image in Figure 4 indicates that if we want to classify a trapezoidal grating whose range of SWA is out of the range for training signatures, more training signatures and bigger sigma are necessary for the correct classification. Different magnitudes of noise may lead to different classification accuracy. Hence we add Gaussian noise into the testing signatures. The standard deviation of the Gaussian (magnitude) model is set as a percentage of the mean of the simulated signatures [15]. As shown in the right of Figure 4, each curve has a boundary point which divides the classification accuracy into two parts: one part indicates the classification accuracy is 100%, the other one indicates the classification accuracy is below 100%. Also, we can hardly find the two red curves, because the two red curves are overlaying (As shown in the sub-image on the right side). This indicates that the two curves have the same classification effect. In Table 2, we show the boundary point (maximal level of noise) for each set of training signatures in detail.

Table 2. Boundary point of each set of training signatures

	Number of training pairs	The maximal level of noise (Boundary point)	Remark
Set 1	300	0.0147	If the level of noise is bigger than the maximal level of noise (boundary point), the classification accuracy is below 100%.
Set 2	350	0.0149	
Set 3	400	0.0155	
Set 4	450	0.016	

Once the relationship exists in sigma, number of training signatures and classification accuracy is confirmed, we next can use the method discussed above to continue our SVM-based library search strategy. For the discussion above, we chose the magnitude of noise as 0.001, and the number of training signatures was chosen as 350. We then continue our improved library search strategy by using a 25-classification classifier. The extracted parameters errors and searching time by SVM-based library search strategy is compared with the results by linear search method of the whole library. The results are as shown in Figs. 5 to 7.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

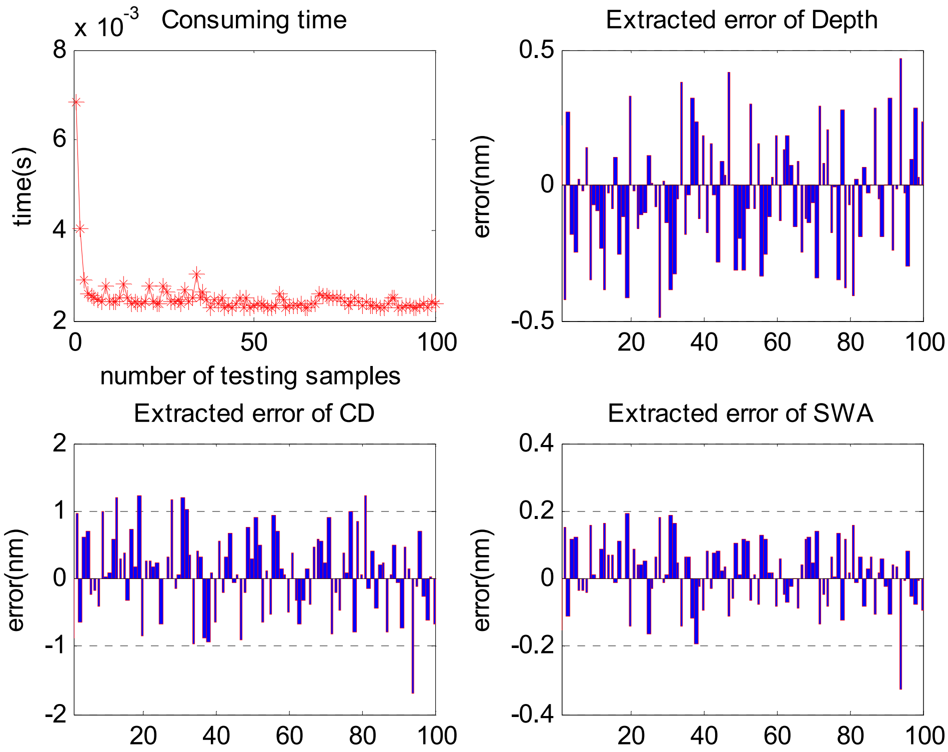


Fig. 5. The searching time and extracted errors of depth, CD and SWA by linear library search method.

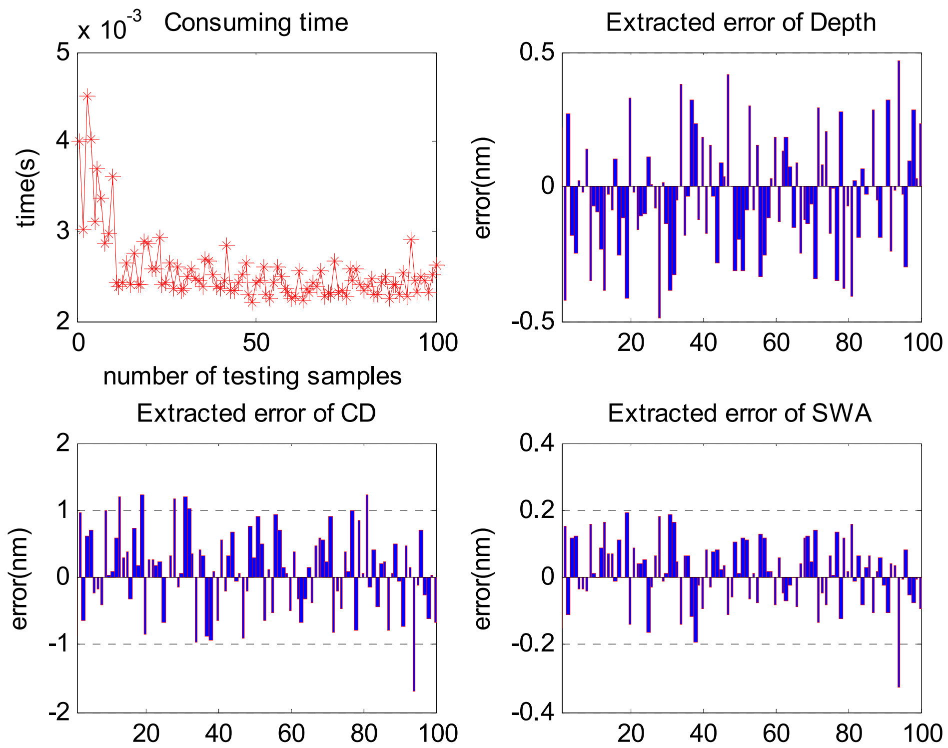


Fig. 6. The searching time and extracted errors of depth, CD and SWA by SVM-based library search method.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

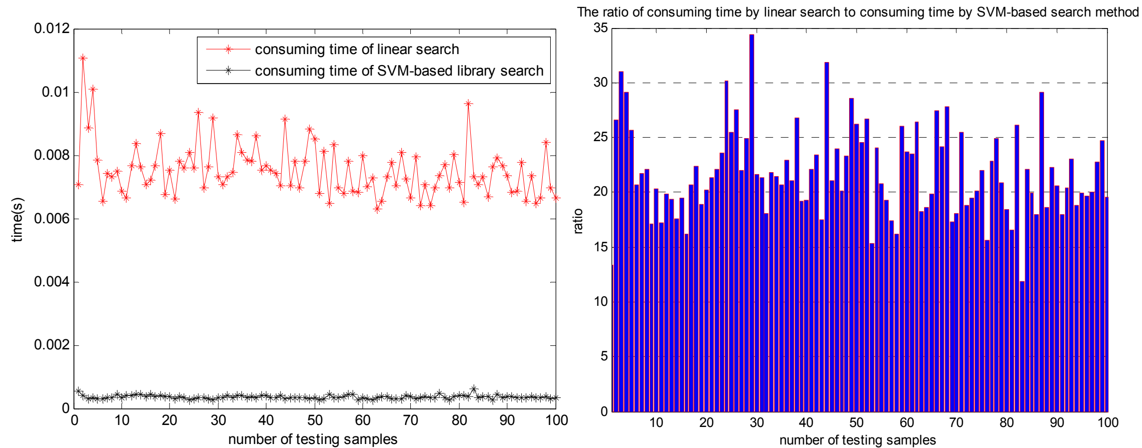


Fig. 7. The searching time of linear search method and SVM-based search method, respectively.

As shown in Figure 7, the searching time by linear search is nearly 20 times than the SVM-based method. Also, as shown in Figure 5 and Figure 6, the extracted parameters errors by the two different methods are nearly the same when set the initialized condition properly, which means the classification accuracy ratio is nearly 100% for each of the testing samples. Hence, we indicate that the SVM-based library search method is a speed controllable method which only adds a time-consuming but off-line procedure.

3.3 Experimental Results

In our experiments, a trapezoidal grating was fabricated in order to validate experimentally the SVM-based library search strategy. Figure 8 depicts the geometrical model and top-down scanning electron microscopy view of one-dimensional trapezoidal grating. The measured value of CD is 172 nm by scanning electron microscopy.

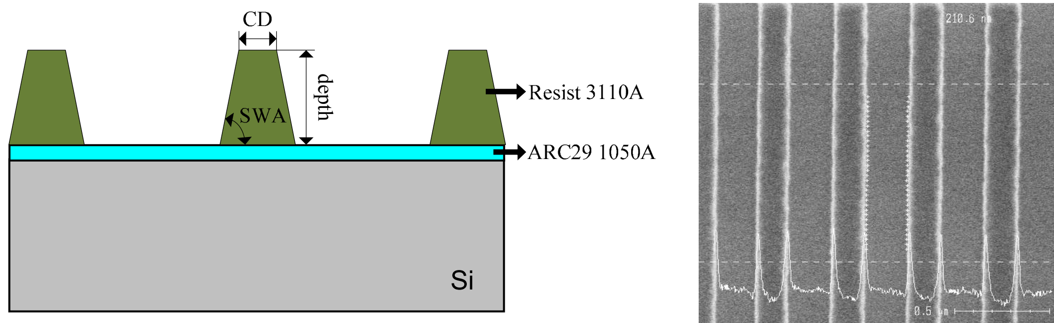
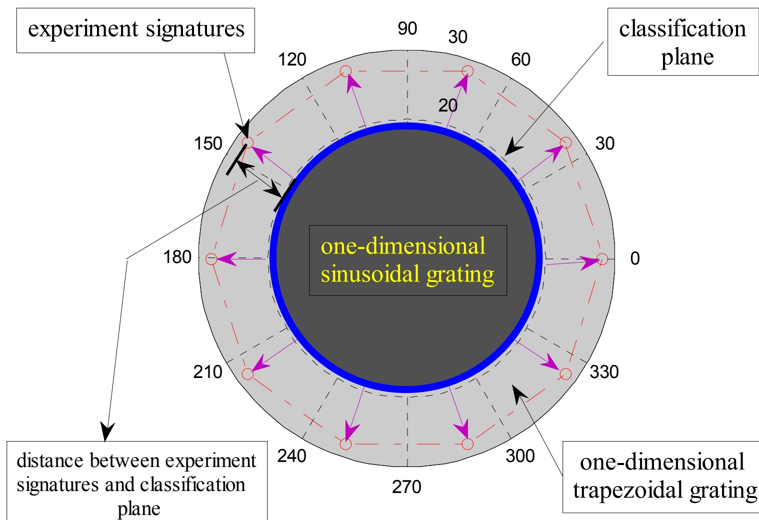


Fig. 8. Geometrical model and top-down scanning electron microscopy view of the one-dimensional trapezoidal photoresist grating with a pitch of 400 nm.

Here we extract the depth, CD and SWA (the nominal sidewall angle is 90°, but the actual value of SWA may have a little bias compared to the nominal value. Hence, the SWA also needs to be extracted). We established a simulated signature library. We set the CD from 150 to 190 nm, the depth from 290 to 320 nm, both with an increment of 1 nm, and the SWA from 86° to 90°with an increment of 0.1°. We repeatedly measured the patterned region for 10 times (different measurements may contain different noise levels). A ten-classification classifier was trained in the SVM-based method, and the experimental results are shown in Fig. 9, Fig. 10, and Table.3.

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms



Measurement Order	Euclidean distance
1	28.2190
2	28.2212
3	28.2221
4	28.2207
5	28.2182
6	28.2148
7	28.2136
8	28.2141
9	28.2143
10	28.2140

Fig. 9. Classification for the experiment signatures of one-dimensional trapezoidal grating. The data in the right table represents the Euclidean distance between experiment signatures and classification plane as shown in the left image.

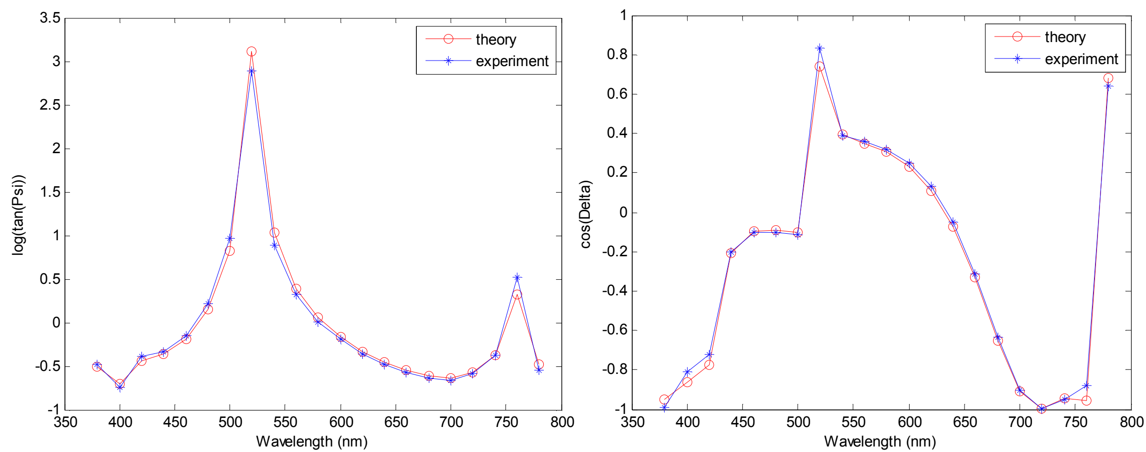


Fig. 10. Comparison of theoretical and experiment signatures

Table 3. Comparison of linear search and SVM-based library search

Measurement order	Nominal CD 185 nm, SEM CD 172 nm		Nominal depth 310 nm		Nominal SWA 90°		Ratio of time
	Linear	SVM-based	Linear	SVM-based	Linear	SVM-based	
1	164	164	297	297	88.4	88.4	11.0
2	163	163	299	299	88.4	88.4	8.9
3	162	162	301	301	88.0	88.0	12.8
4	164	164	298	298	88.4	88.4	13.1
5	162	162	303	303	88.0	88.0	10.9
6	162	162	303	303	87.6	87.6	11.1
7	161	161	304	304	87.6	87.6	10.6
8	162	162	303	303	87.6	87.6	11.1
9	162	162	303	303	87.6	87.6	10.5
10	162	162	303	303	87.6	87.6	11.1

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

Figure 9 indicates that for the 10 measurements, although different measurements may contain different noise levels, the classification is correct (experiment signatures are represented by small red circles). Also, as shown in the right table, the Euclidean distances between experiment signatures and classification plane for all the measurements are nearly the same, which means the noise level for every measurement has limited influence on the classification accuracy.

Figure 10 is a comparison of theoretical and experiment signatures in one measurement, the extracted values of parameters corresponding to the theory signatures are 303 nm, 162 nm, and 87.6°, respectively. It demonstrates an excellent agreement between the three parameter theory and experiment.

As shown in Table 3, the parameter extracted values by linear search method and SVM-based method are all the same, but the searching time of SVM-based method is only about ten percent of the searching time by linear search method. Hence, we indicate the SVM-based method is a time-controllable and accurate method which can be applied in OCD metrology.

4. CONCLUSIONS

In this paper, we have introduced an SVM-based classification technique that can be applied in the identification of geometrical profiles of diffraction structures. Our simulation and experiment have shown that the SVM classifiers can accurately identify the geometrical profile of one-dimensional trapezoidal grating even though the different levels of noise exist in the signatures. We also have obtained several curves that represent the relationship among the classification, the number of training signatures, and the value of sigma. For the case of a point that is above those curves, the classification accuracy is nearly 100%. By further extending the application of SVM in scatterometry-based structure reconstruction, an SVM-based library search strategy is proposed. Before searching for the nearest neighbor of an experimental signature in a library, a multi-classification SVM classifier is trained. Once a measured signature is obtained, the SVM classifier can map the signature into its corresponding sub-library. By searching in the sub-library, the searching time can be reduced significantly compared to the linear search in the whole library. The simulation and experiment have demonstrated that the SVM-based library search strategy can achieve a robust and fast extraction of structural parameters.

ACKNOWLEDGMENT

This work was funded by National Natural Science Foundation of China (Grant No. 91023032, 51005091) and National Instrument Development Specific Project of China (Grant No. 2011YQ160002).

REFERENCES

[1] H. T. Huang and F. L. Terry, "Spectroscopic ellipsometry and reflectometry from gratings (scatterometry) for critical dimension control and in situ, real-time monitoring," Thin Solid Films **468**, 339-346 (2004).

[2] J. M. Holden, T. Gubiotti, W. A. McGaham, M. Dusa, and T. Kiers, "Normal-incidence spectroscopic ellipsometry and polarized reflectometry for measurement and control of photoresist critical dimension," Proc. SPIE **4689**, 1110-1121 (2002).

[3] S. Robert, A. M. Ravaud, S. Reynaud, S. Fourment, F. Carcenac, and P. Arguel, "Experimental characterization of subwavelength diffraction gratings by an inverse scattering neural method," J. Opt. Soc. Am. A **19**, 2394-2402 (2002).

[4] C. W. Zhang, S. Y. Liu, T. L. Shi, and Z. R. Tang, "Improved model-based infrared reflectrometry for measuring deep trench structures," J. Opt. Soc. Am. A **26**, 2327-2335 (2009).

[5] X. Niu, N. Jakatdar, J. Bao, and C. J. Spanos, "Specular spectroscopic scatterometry," IEEE Trans. Semicond. Manufact. **14**, 97-111 (2001).

[6] X. Niu, "An integrated system of optical metrology for deep sub-micron lithography," Ph.D. Dissertation, University of California at Berkeley (1999).

[7] J. B. Loudermilk, D. S. Himmelsbach, F. E. Barton, and J. A. Haseth, "Novel search algorithms for a mid-infrared spectral library of cotton contaminants," Appl. Spectrosc. **62**, 661-670 (2008).

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms

[8] R. R. Nidamanuri, B. Zbell, "A method for selecting optimal spectral resolution and comparison metric for material mapping by spectral library search," Prog. Phys. Geog. **34**, 47-58 (2010).
[9] S. Ramaswamy, K. Rose, "Adaptive cluster-distance bounding for nearest neighbor search in image databases," Proc. IEEE **6**, 381-384 (2007).
[10] C. M. Chen, Y. B. Ling, "A sampling-based estimator for top-k query," Proc. ICDE **18**, 617-627 (2002).
[11] R. Weber, H. J. Schek, and S. Blott, "A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces," Proc. VLDB **24**, 194-205 (1998).
[12] I. Gereige, S. Robert, S. Thiria, F. Badran, G. Granet, and J. J. Rousseau, "Recognition of diffraction-grating profile using a neural network classifier in optical scatterometry," J. Opt. Soc. Am. A **25**, 1661-1667 (2008).
[13] C. Cortes, V. Vapnik, "Support-vector networks," Mach. Learn. **20**, 273-297 (1995).
[14] H. Drucker, D. Wu, V. Vapnik, "Support vector machine for spamcategorization," IEEE Trans. Neural Networks **10**, 1048-1054 (1999).
[15] R. M. Al-Assaad, D. M. Byrne, "Error analysis in inverse scatterometry. I. Modeling," J. Opt. Soc. Am. A **26**, 326-328 (2007).

Downloaded from SPIE Digital Library on 05 Apr 2012 to 115.156.244.32. Terms of Use: http://spiedl.org/terms