Optical Digital Profilometry applications on contact holes

Jörg Bischoff, Xinhui Niu and Nickhil Jakatdar

Timbre Technologies Inc., A Tokyo Electron Company, 2953 Bunker Hill Lane, Santa Clara, CA. USA 95054

ABSTRACT

Optical scatterometry and Optical Digital Profiling (ODP) have become mainstream technology in CD and profile metrology. Without question, the extension of these techniques to measure three-dimensional patterns such as contact holes or posts is an important demand. In this paper, we demonstrate the application of ODP to contact holes and posts for both lithography and etch processes. The underlying theory based on the Rigorous Coupled Wave Approach (RCWA) is outlined and metrology results are compared with simulations.

Keywords: Scatterometry, metrology, contact holes, 3D grating patterns.

1. INTRODUCTION

Today, scatterometry is widely accepted as an enabling technology for integrated metrology and automated process control. In the beginning, the focus was on angle resolved methods such as the so-called 2θ-approach[1,2]. With the introduction of spectral methods[3] it became possible to utilize the same state-of-the-art spectral reflectometers or ellipsometers as known from thin film metrology to extract profile data. This was a breakthrough for the wide acceptance of scatterometry as a promising technology. Another benefit was the introduction of automated optimization regression which is also called Optical Digital Profiling (ODP). In this way it became possible to obtain quick results of certain amount of profiles without performing a tedious library building and consecutive searching. Moreover, ODP helps to reduce the risk of creating an ill-adapted library considerably.

Until now, almost all published applications, with the exception of a few papers[4,5], have addressed the metrology of one-dimensional (2D) gratings. However, there is a strong demand in the industry to extend scatterometry to two-dimensional (3D) patterns such as contact-hole and contact-post arrays.

The most essential part of scatterometry is the diffraction grating modeling which is called forward modeling in this context. The Rigorous Coupled Wave Approach (RCWA) has been shown its ability to ensure highly accurate results in 2D. In addition to that, it is accepted as the most versatile and robust algorithm for rigorous diffraction computations. Therefore, we developed a RCWA based algorithm for 3D structures. The most essential steps and features of this algorithm are outlined in section 2.

Although, the same basic approach (RCWA) is applied for 2D and 3D there are distinctive differences in terms of polarization decoupling, complexity, geometry and excitation conditions as well as computation time effort. This topics shall be discussed in section 3.

Section 4 presents some modeling results which demonstrate the correctness and the universality of the developed code.

This will be followed by the presentation of first metrology results, i.e., comparisons between simulated and measured signals in section 5. The extracted profiles from ODP are compared with scanning electron micrographs. Spectroscopic ellipsometry is used for the measurement data acquisition. It is shown that the

Optical Digital Profilometry provides an accurate, inexpensive, and non-destructive profile metrology for the 3D patterns.

2. THEORY

Modal methods are acknowledged to be the most effective techniques to calculate the rigorous diffraction from periodic patterns. They are assigned to the differential frequency domain methods. Particularly, the Modal Methods with Fourier Expansion (MMFE) which include the RCWA as a kind of sub class are widely preferred due to its versatility and relative ease of implementation. According to the high interest, the RCWA for 2D features has gained a high level of maturity. Early issues, particularly the failing of the algorithm for thick layers and the bad convergence in TM-polarization have been successfully removed with the use of recurrence coupling schemas such as the scattering matrix approach [6] and with the apllication of the correct Fourier factorization of products [7]. The only remaining issue is the inappropriate adaption of the boundary condition for TM-polarized light which may cause convergence problems for strong metallic materials [8]. However, it has already been shown that this can also be covered within the frame of the RCWA [9]. Since the computation effort increases, this extension should only be applied to cases where the standard RCWA fails.

After successfully solving of the convergence problems, the exp ansion of the MMFE to 3D patterns was launched [10,11]. In the following the most essential steps of the 3D RCWA algorithm will be sketched. A more detailed description can also be found [12]. An distinctive feature of the RCWA in comparison to other methods (e.g., the differential method) is the slicing of a given profile resulting in stair case approximation. Within one slice, the refraction index and therefore its permittivity is assumed to be independent on the normal direction z, i.e., $\varepsilon(x,y,z) \rightarrow \varepsilon(x,y)$.

2.1 Solution for a slice

The first step is the solution of the diffraction problem for one slice. To this end, at first a system of differential equations is derived from the time harmonic Maxwell's equations which can be written as:

$$\begin{aligned} curl\ \vec{H}(\vec{r}) &= j\varepsilon(\vec{r}) \cdot \vec{E}(\vec{r}) \\ curl\ \vec{E}(\vec{r}) &= j \cdot \vec{H}(\vec{r}) \end{aligned} \qquad \textbf{(1)}$$

In order to make the electric and magnetic field strengths dimensionless the normalization $\omega/c = 1$ and $\mu = 1$ was utilized. In equation (1), every field vector has an additional vector character due to the modal decomposition leading to a sum over components m,n. For numerical reasons the higher orders have to be truncated. The total number of orders N which is kept in the computations is called truncation number. With:

$$\frac{\partial}{\partial \vec{r}} = \frac{\partial}{\partial z}\vec{e}_z + j(b_x \cdot \vec{e}_x + b_y \cdot \vec{e}_y) \qquad \textbf{(2)}$$

applied to (1), the normal components of the electromagnetic field can be expressed by the tangential components. Inserting this equations again into the remaining four Maxwell equations for the tangential components results finally in a first order differential equation system:

$$\frac{\partial}{\partial z}\begin{pmatrix} E_y \\ E_x \\ H_y \\ H_x \end{pmatrix} = \begin{pmatrix} 0 & D_1 \\ D_2 & 0 \end{pmatrix} \cdot \begin{pmatrix} E_y \\ E_x \\ H_y \\ H_x \end{pmatrix} \qquad \textbf{(3)}$$

The zero matrices appearing in the main diagonal reveal the transversal character of the electromagnetic field. The correct application of Li's factorization rules [11] for product ends up with the following expressions for the 2N x 2N submatrices D_1 and D_2:

$$D_1 = j \cdot \begin{pmatrix} b_y \cdot [\![e]\!]^{-1} \cdot b_x & 1 - b_y \cdot [\![e]\!]^{-1} \cdot b_y \\ b_x \cdot [\![e]\!]^{-1} \cdot b_x - 1 & -b_x \cdot [\![e]\!]^{-1} \cdot b_y \end{pmatrix} \;;\; D_2 = j \cdot \begin{pmatrix} -b_y \cdot b_x & b_y{}^2 - [\![e]\!] \\ [\![e]\!] - b_x{}^2 & b_y \cdot b_x \end{pmatrix} \qquad \textbf{(4)}$$

The computation of the particular ε-matrices is given by formulas (4) through (9) in [11].

Equation (3) could be solved by means of a diagonalization. However, a more efficient approach can be derived by making use of the ε-independence on the normal direction within a slice. Another differentiation of (3) for the normal direction z results in a second order differential equation system:

$$\frac{\partial^2}{\partial z^2}\begin{pmatrix} E_y \\ E_x \\ H_y \\ H_x \end{pmatrix} = \begin{pmatrix} D_e & 0 \\ 0 & D_h \end{pmatrix} \cdot \begin{pmatrix} E_y \\ E_x \\ H_y \\ H_x \end{pmatrix} \qquad \textbf{(5)}$$

with: $\mathbf{D_e} = \mathbf{D_1} \cdot \mathbf{D_2}$ and $\mathbf{D_h} = \mathbf{D_2} \cdot \mathbf{D_1}$.

In the next step, this system is reduced to single differential equations by means of a diagonalization of the matrix $\mathbf{D_1}$ which can be written as $\mathbf{D_1} = \mathbf{W_1} \cdot \Lambda_1 \cdot \mathbf{W_1}^{-1}$. Physically, the transition from $\mathbf{D_1}$ means the transition from the Rayleigh modes that propagate in the superstrate, substrate or in homogeneous layers to the Bragg modes that propagate in the periodic layer. The transfer matrix $\mathbf{W}$ is the matrix which comprise the eigenvectors of the characteristic matrix $\mathbf{D_1}$ (or $\mathbf{D_2}$) of the layer and the diagonal matrix Λ comprises the Eigen values λ. The fundamental solution system of the second order differential equation for the order m,n is:

$$S_{m,n}(z) = u_{m,n}(z_0) \cdot \exp\{-q_{m,n} \cdot (z - z_0)\} + d_{m,n}(z_0) \cdot \exp\{+q_{m,n} \cdot (z - z_0)\} \qquad \textbf{(6)}$$

where the $q_{m,n}$ are the complex square roots of the eigenvalues $\lambda_{m,n}$. The two basic solutions $u_{m,n}$ and $d_{m,n}$ can be considered as waves that run up and down, respectively.

2.2 Matrix coupling

For the remainder of the solution we will stick closely to the algorithms which are described in detail in Li's paper [6]. Once, the solutions, i.e., the eigen decompositions, for the various slices are obtained, the total solution can be assembled by making use of the boundary conditions of the electromagnetic field. The tangential field components are equivalent across a flat interface (flatness was a precondition for the slicing). This leads to formula (2a) in [6] which couples the up and down waves across the interface. The Bragg waves propagate independly from each other within a slice:

$$\begin{pmatrix} u_{m,n}(z_0 + t) \\ d_{m,n}(z_0 + t) \end{pmatrix} = \begin{pmatrix} \exp(-q_{m,n} \cdot t) & 0 \\ 0 & \exp(+q_{m,n} \cdot t) \end{pmatrix} \cdot \begin{pmatrix} u_{m,n}(z_0) \\ d_{m,n}(z_0) \end{pmatrix} \qquad \textbf{(7)}$$

Herein, t is the thickness of slice i and z_0 is the absolute coordinate of the interface in normal direction. After reformulation and involvement of the propagation step, one obtains the explicit formulas (5-7) in [6] for the coupling between the up and down waves across one full slice, comprising one interface crossing and one propagation step. Herein the matrices $\mathbf{t}$ and $\bar{t}$ are called interface and layer matrix, respectively.

Now, it would be straightforward to couple all these solutions together resulting eventually in an expression that relates the fields, or up and down waves, respectively, at the front side $u^{(f)}$ and $d^{(f)}$ of the grating stack with those on the back side $u^{(b)}$ and $d^{(b)}$ connected by a transfer matrix. The fields are shown schematically in figure 1. However, what we are really aiming at is a relation between the cause and response fields, i.e., between $d^{(f)}$ and (often dropped) $u^{(b)}$ as excitation fields on the one hand, and $d^{(b)}$ and $u^{(f)}$ as detection fields on the other hand. This means one has to reformulate the transfer matrix which may lead to an ill-posed matrix inversion problem. However, this should be avoided because of the numerical instabilities and bad convergence. Instead of, the so-called scattering matrix (S-matrix) approach was introduced to overcome this convergence and stability issue. Unlike the transfer matrix approach, the S-matrix approach relates the cause and response waves from the beginning:

$$\begin{pmatrix} u^{(f)} \\ d^{(b)} \end{pmatrix} = (S) \cdot \begin{pmatrix} d^{(f)} \\ u^{(b)} \end{pmatrix} = \begin{pmatrix} r^f & t^b \\ t^f & r^b \end{pmatrix} \cdot \begin{pmatrix} d^{(f)} \\ u^{(b)} \end{pmatrix} \qquad \textbf{(8)}$$

The S-matrix consists of four usually densely populated submatrices. If one drops the excitation from the back side the submatrix r^f gives the reflection at the front side and the matrix t^f gives the transmission at the back side for front side excitation.

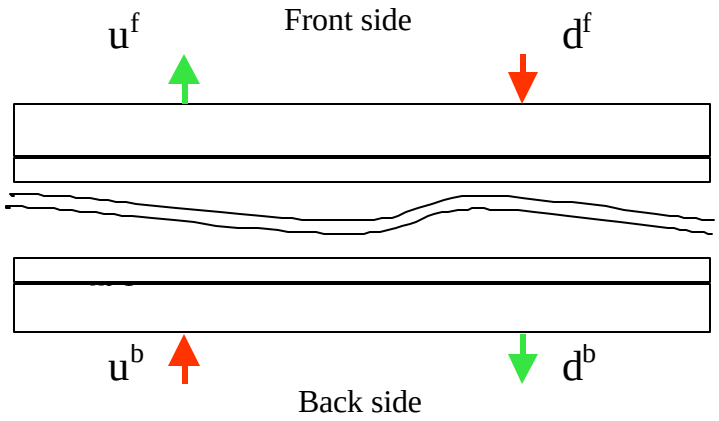


Figure 1: Fields at the front and back side of a grating stack (d^f and u^b are cause waves whereas d^b and u^f are response waves)

In order to avoid the numerical problems connected with the reformulation of the transfer matrix, it is reasonable to start directly with the scattering matrix and couple layer by layer to the existing S-matrix. To this end, the scattering matrix has to be initialized appropriately, i.e., $\mathbf{r}^{f,b} = \mathbf{0}$ and $\mathbf{t}^{f,b} = \mathbf{1}$. The elementary scattering matrices for an interface and a layer, $\mathbf{s}$ and $\overline{\mathbf{S}}$ are given by the formulas (10a) through (12a) in Li Feng Li's paper [6].

Finally, the diffraction efficiencies or ellipsometric parameters (such as tan ψ and cos δ) can be claculated from the corresponding elements of the submatrices $\mathbf{r}^{f,b}$ or $\mathbf{t}^{f,b}$.

3. EFFECTIVITY CONSIDERATIONS

The computation complexity in modal methods is directly associated with the matrix size via the truncation number N. Because the Eigen decomposition and various matrix muliplications which are the most time consuming steps scale with N^3 the number of retained orders has a paramount impact on the computation speed. Unlike in 2D where we have only diffraction orders in one frequency dimension, there are diffraction order in two dimensions for crossed gratings. Based on this, the expected computation complexity increase for 3D related to 2D is 2 to 3 magnitude. In this study, some measures in algorithm and computation optimization are used to reduce the computation time in 3D. All in all, reasonable computation times can be achieved by making use of the opportunities given by physics and modern computing sciences.

4. MODELING

In this section, some simulations will be discussed that show both the correctness of the code as well as its universality. Actually, the accuracy of the modeling strongly impacts the accuracy of the whole ODP method. Therefore, it is essential to ensure the accuracy of the code. Moreover, simulations help to gain more insight into the applicability of a code. This shall be discussed with the two most important metrology related process situations namely the development inspection (DI) and the final inspection (FI). The target of the development inspection is to measure the developed photoresist profile whereas for the final inspection the etched profile is of interest. Moreover, the following simulations comprise both metrology hardware namely reflectometry with normal incidence and ellipsometry with oblique incidence.

4.1 Slicing

The slicing criterion is one way for the validation of a RCWA implementation. Basically, the simulated diffraction efficiency should converge to a fixed value while increasing the number of slices that approximate a given geometry. Because of the electromagnetic boundary conditions, the slice thickness should be adapted to the slope of the profile, i.e., steep slopes need less slices and shallow slopes need more slices. An equal thickness for all slices results in the case of a constant slope such as a trapezoid or a cone. The code was tested with several related cases. One example was an array (period = 300 nm) of circular holes in PR being 150 nm in diameter (top) and having 80 degrees slope angle on a silicon substrate. The number of slices was gradually increased from 3 through 10 while the total depth of the holes was kep constant (300 nm). Figure 3a shows the diffraction efficiency versus the number of slices for the DI-case. Obviously, the signal converges towards 30.8% with increasing number of slices. Having a more close look at the relative change it can be found that the change is about 0.3% while going from 5 to 10 slices. Therefore, 5 slices might be enough to ensure an accurate result. This corresponds to a stepwidth of about 10.5 nm in lateral direction from one slice to the next one. A slicing criterion of about $\lambda/30$ follows relative to the wavelength within the resist (500 nm/1.57) .

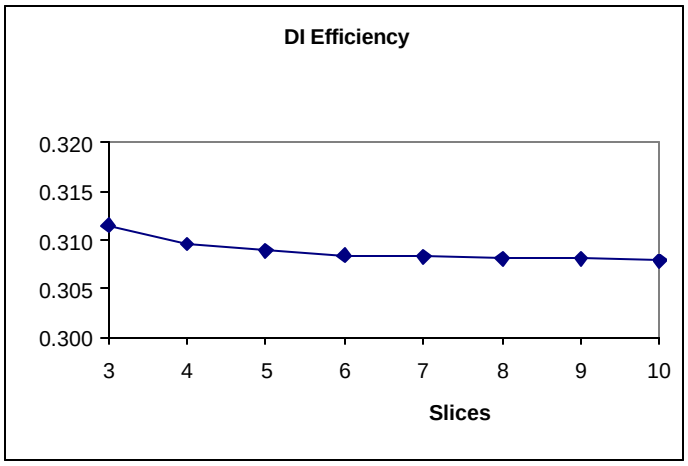


Fig. 3a: Zeroth order efficiency versus the number of slices for the DI reflectometry of a contact hole array

In addition to that, the ellipsometric DI situation with 65 degrees angle of incidence (AOI) was modeled. The resulting tan ψ and cos δ are plotted in figure 3b. Unfortunately, the convergence with increasing slice number is worse compared to the reflectometer case. The reasons might be the oblique incidence and the increased sensitivity particular of the phase.

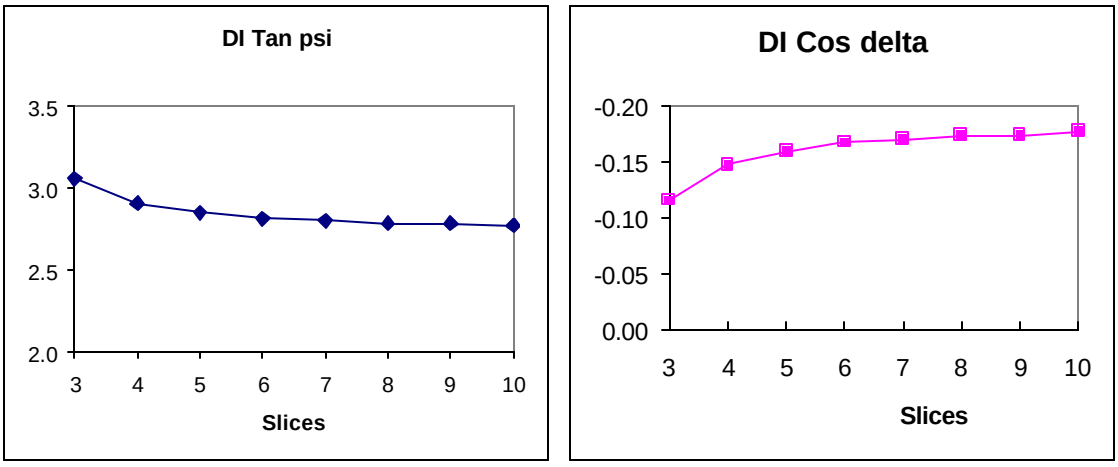
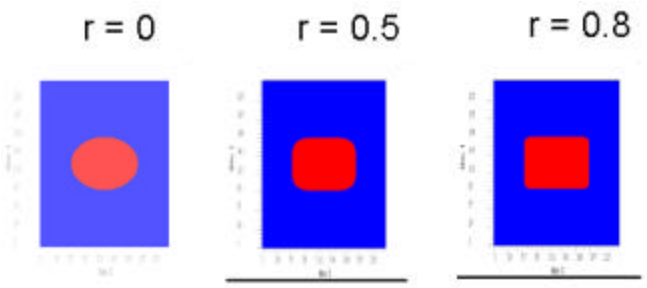


Fig. 3b: Tan ψ (LHS) and cos δ (RHS) versus the number of slices for the DI ellipsometry of a contact hole array

Further work should be devoted to the estimation of the slcing criterion for the FI case. In any case, more slices will be necessary for the same geometry as in DI due to the much higher real part of the etched material (e.g. 4.3 for Si at 500 nm).

4.2 Feature shape

A very important feature in photolithography is the patterning of contact holes. Rectangular shapes on the mask become rounded due to the transfer function of the stepper optics. In our algorithm, this was taken into account by means of a so-called rectangularity measure r that defines the deviation from a rectangle. The limits are r = 1 (rectangle) and r = 0 (ellipse). A visual impression is given in figure 4 that shows top down views for 3 different r-values.



A gradual change of the diffraction signal is expected while changing the shape of the features. It is of big interest of course, how sensitive the signal is with changing hole shape. To this end, a few simulations were performed. For the DI case, a contact hole array equal to that of subsection 4.1 was considered. The top view shape of the holes is changed from circular to quadratic. Figure 5 shows the diffraction efficiency (AOI = 0°) and the ellipsometric signals (AOI = 65°), respectively versus the rectangularity r. The conclusion from this plots is that first, the signals are sensitive to the shape and second, that the tan ψ gives the strongest relative change (about 20%).

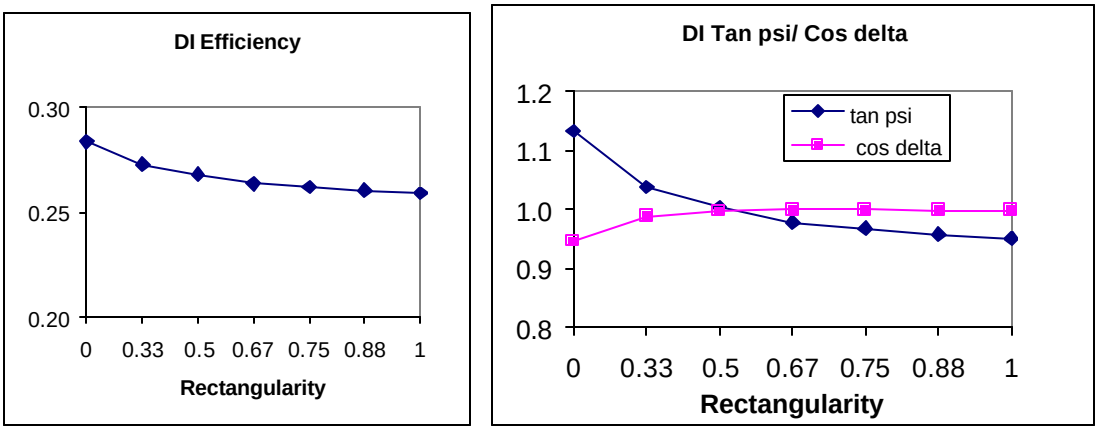


Fig. 5: Zero order efficiency (LHS) and tan ψ and cos δ (RHS) versus the "rectangularity" of contact holes in PR (DI)

5. EXPERIMENTAL

The first experiment we consider is a lithography patterning process. We have a resist/ARC/SiO_2/Si stack. The patterning mask has 0.2μm/0.3μm space/hole gratings on a 50μmx50μm testing area. The grating is measured by the Sopra SE 300 spectroscopic ellipsometer. Fig 6 shows the fit between the silmulated and

measured signal. Fig 7 shows the corresponding contact hole profile and it comparison with the top-down CD-SEM image.

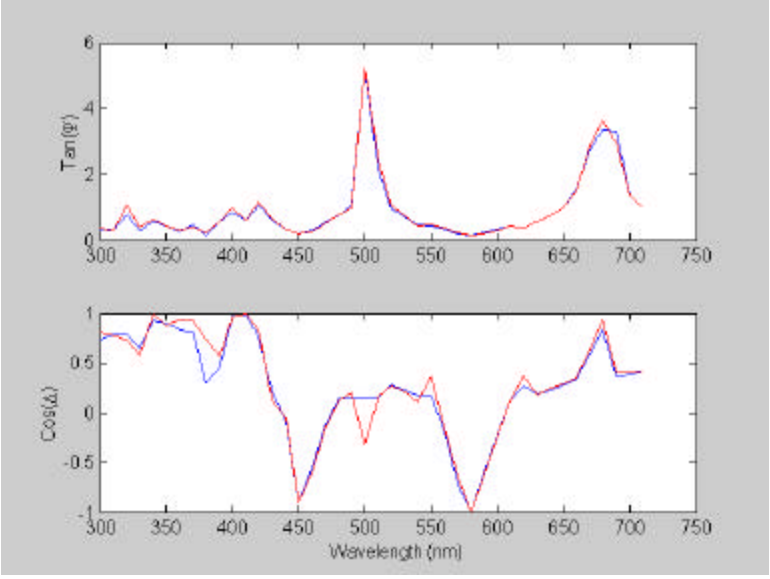


Fig. 6: Matching on the simulated and measured tanψ and cosδ for the lithography process.

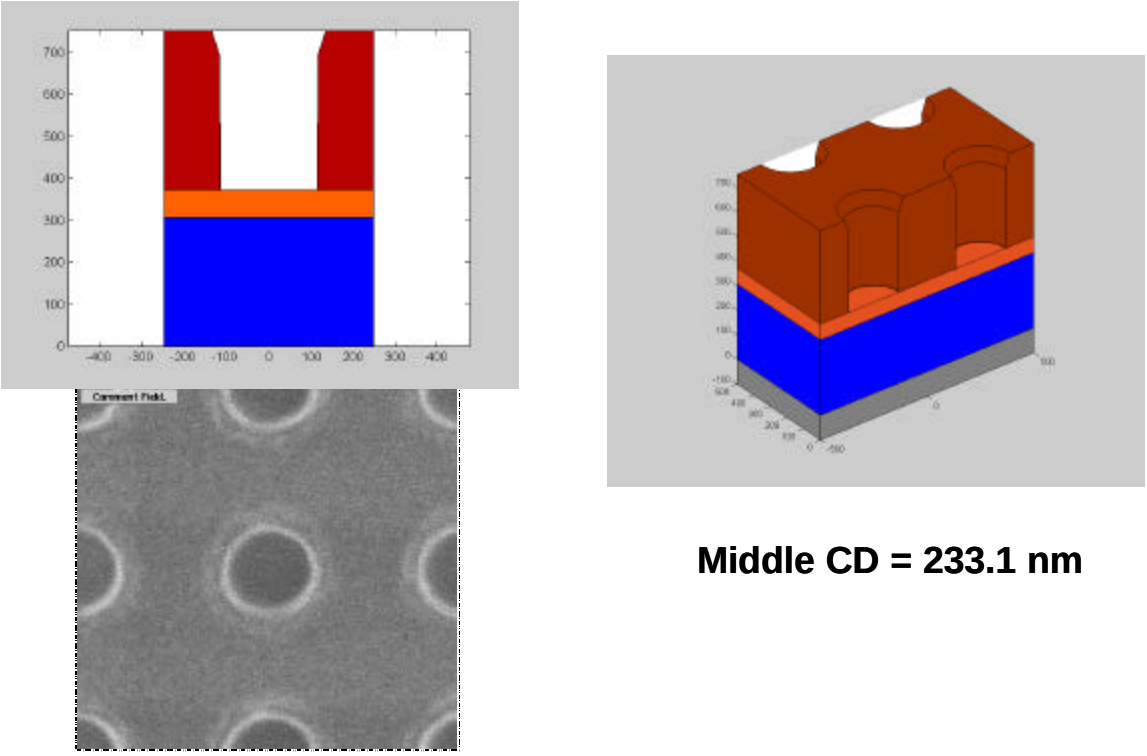


Fig. 7: Extracted DI contact hole profile and its comparison with top-down CD-SEM image.

For the verification purpose, we also measured the has 0.2µm/0.3µm Line/Space patterns. Fig 8 shows the fit between the measured and simulated signals. The extracted profile matches the image from top-down CD-SEM.

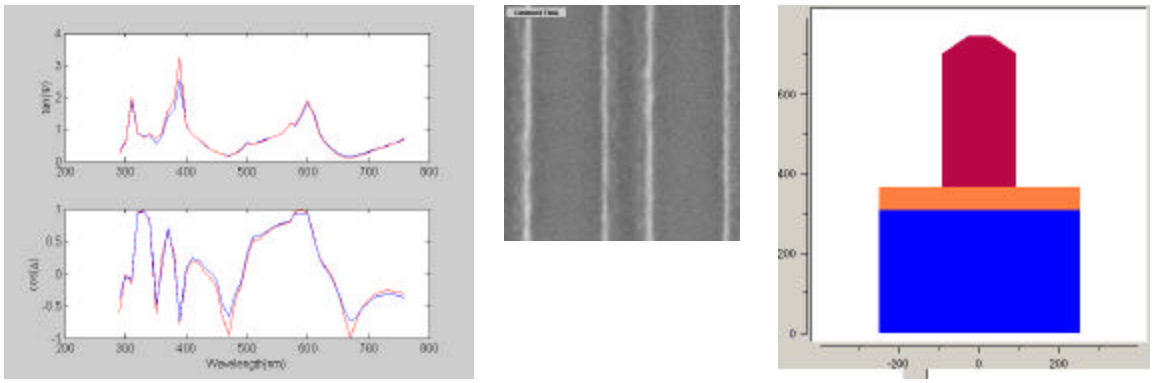


Fig. 8: Matching on the simulated and measured tanψ and cosδ for the 2D structure in lithography process and the corresponding profile.

The second experiment we consider is an etch process. We have a SiO_2/Si stack. The patterning mask has 0.2µm/0.3µm space/hole gratings on a 50µmx50µm test area. Fig 9 shows the fit between the silmulated and measured signals. Fig 10 shows the corresponding contact hole profile and it comparison with the cross sectional SEM image.

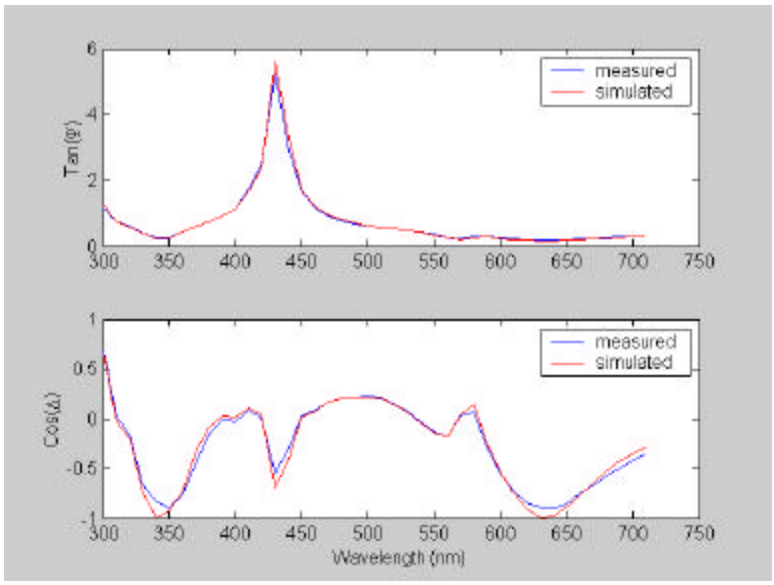


Fig. 9: Matching on the simulated and measured tanψ and cosδ for the etch process.

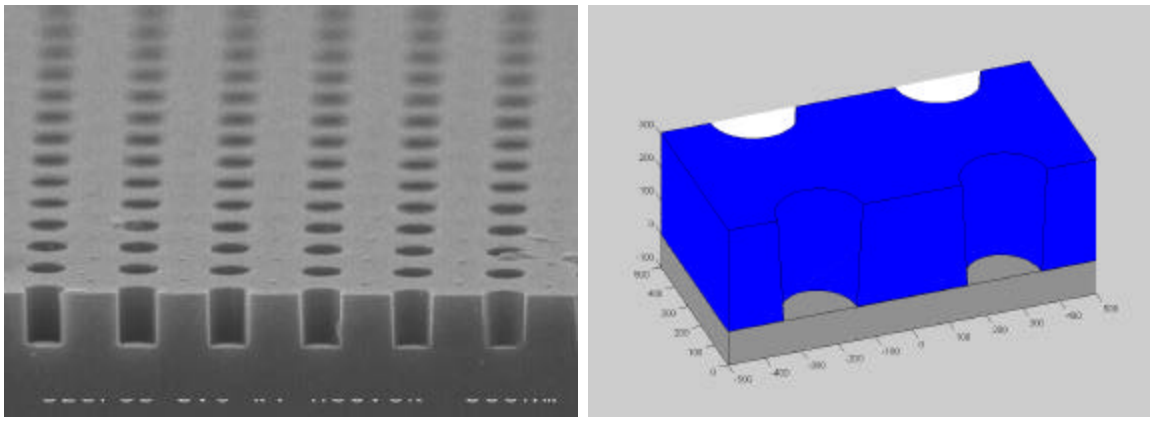


Fig. 10: Cross-sectional view of the etched contact hole.

6. CONCLUSIONS

In this paper, it is shown that the contact/via grating physics has been developed and verified. Measurements from both DI and FI processes are conducted and good results are achieved as well. It is demonstrated that optical scatterometry provides an accurate, in expensive, and non-destructive profile metrology for the 3D test gratings.

REFERENCES

1. C.J. Raymond, M.R. Murnane, S.L. Prins, S.S.H. Naqvi, J.R. McNeil and J.W. Hosch, "Multi-parameter CD measurement using scatterometry", *Proceedings SPIE* **Vol. 2725**, pp.698-709, 1996.
2. J. Bischoff, J.W. Baumgart, und H. Truckenbrodt, "Photoresist metrology based on light scattering", *Proceedings SPIE* **Vol. 2725**, pp.678-689, 1996.
3. X. Niu, N. Jakatdar, J. Bao and C.J. Spanos, "Specular Spectroscopic Scatterometry", IEEE Transactions on Semiconductor Manufacturing, **Vol. 14**, pp. 97-111, 2001.
4. M. Yeung and E. Barouch, "Electromagnetic scatterometry applied to in situ metrology", *Proceedings SPIE* **Vol. 4344**, pp.484-495, 2001.
5. J. Bischoff, L. Hutschenreuther, H. Truckenbrodt, J. Bauer, U. Haak and T. Skaloud: "Characterization of 3D resist patterns by means of optical scatterometry", *Proceedings SPIE* **Vol. 3743**, pp. 49-60, 1999.
6. Lifeng Li, "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings," J. Opt. Soc. Am. **A 13**, pp. 1024-1035, 1996.
7. Lifeng Li, "Use of Fourier series in the analysis of discontinuous periodic gratings," J. Opt. Soc. Am. **A 13**, pp. 1870-1876, 1996.
8. E. Popov, M. Neviere, B. Gralak and G. Tayeb, "Staircase approximation validity for arbitrary-shaped gratings," J. Opt. Soc. Am. **A 19**, pp. 33-42, 2002.
9. E. Popov and M. Neviere, "Maxwell equations in Fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media," J. Opt. Soc. Am. **A 18**, pp. 2886-2894, 2001.
10. E. Noponen and J. Turunen, "Eigenmode method for electromagnetic synthesis of diffractive elements with three-dimensional profiles," J. Opt. Soc. Am. **A 11**, pp. 2494-2502, 1994.
11. Lifeng Li, "New formulation of the Fourier modal method for crossed surface-relief gratings," J. Opt. Soc. Am. **A 14**, pp. 2758-2767, 1997.
12. J. Bischoff, "Theoretical and experimental investigations of light diffraction from micro-patterned multilayer stacks," Thesis Dr.-Habil., TU Ilmenau/ Germany 2001.