Electron beam lithographic modeling assisted by artificial intelligence technology

Noriaki Nakayamada, Rieko Nishimura, Satoru Miura, Haruyuki Nomura, Takashi Kamikubo, NuFlare Technology, Inc. (Japan), 8-1 Shinsugita-cho, Isogo-ku, Yokohama 235-8522, Japan

ABSTRACT

We propose a new concept of tuning a point-spread function (a “kernel” function) in the modeling of electron beam lithography using the machine learning scheme. Normally in the work of artificial intelligence, the researchers focus on the output results from a neural network, such as success ratio in image recognition or improved production yield, etc. In this work, we put more focus on the weights connecting the nodes in a convolutional neural network, which are naturally the fractions of a point-spread function, and take out those weighted fractions after learning to be utilized as a tuned kernel. Proof-of-concept of the kernel tuning has been demonstrated using the examples of proximity effect correction with 2-layer network, and charging effect correction with 3-layer network. This type of new tuning method can be beneficial to give researchers more insights to come up with a better model, yet it might be too early to be deployed to production to give better critical dimension (CD) and positional accuracy almost instantly.

1. Introduction

The finest art of electron beam lithography is constructed on the basis of full sets of theories and formula in physics. Accelerating a charged particle and hitting a target piece requires correct application of precise knowledge of physics. The same sets of knowledge are required when we try to understand any side effects which make lithographic results imperfect. The examples of such effects are proximity effect caused by backscattered electrons\(^1\), fogging effect caused by bouncing electrons between the chamber wall and the photomask substrate\(^2\), charging effect caused by the electrostatic charge accumulated on the resist surface\(^3\), and heating effect caused by the diffusion of the heat deposited in the resist and the substrate\(^4\).

Historically, those physical effects have been modeled using a certain form of convolution of an input data and a point-spread function. An analytic function such as Gaussian has been commonly used as a point-spread function. When higher modeling accuracy is needed, researchers typically increase the number of Gaussian (multiple Gaussian) or rely on Monte-Carlo simulation in an attempt to reproduce the physical effect virtually by a numerical experiment on a computer. However, the use of Gaussian does not have any physical meaning but has only benefit in easier mathematical handling. The use of Monte-Carlo simulation does have physical meaning, but it cannot model the effect beyond the physical assumptions originally implemented.

\[ f \left[ \text{Input} \right] \times \left[ \text{kernel} \right] = \text{Output} \]

Figure 1. Object model of a convolution form.
SPIE 2017 Panel Discussion: HVM EUV Lithography: Managing without Actinic Patterned Mask Inspection

Emily Gallagher, IMEC

EUV lithography introduces 13.5nm exposure wavelength light that resolves very small photomask features on wafer. Unfortunately, the improvement in resolution also applies to unwanted defects on the photomask surface. Detecting the photomask defects that create imaging problems on wafer is an essential capability for high-yielding lithography. During the SPIE photomask conference 2017, the panel discussion addressed the gap in timing between when EUV lithography will be used in high volume manufacturing (~2019) and when actinic mask inspection will be used (~2021 or later).1 Actinic inspection, an inspection performed using light that has the same wavelength as the lithography, is viewed as the ultimate solution for predicting mask quality. Many have asked for a summary of that discussion and one is provided below.

An August group of panelists from device manufacturing companies and from inspection tool suppliers outlined options for navigating EUV lithography without actinic patterned mask inspection. Listed in their order of presentation, they were: Jeff Farnsworth, Intel; Jed Rankin, GLOBALFOUNDRIES; Byung-Cook Kim, Samsung; Weston Sousa, KLA-Tencor; Shusuke Yoshitake, NuFlare; Fei Wang, Hermes-Microvision. Jeff Farnsworth pointed out that while more defects can be tolerated during development, the number of defects and their size becomes progressively smaller as a technology moves towards HVM. While there are capability gaps now and actinic patterned photomask methods are desired, non-actinic inspection enables the detection of very small defects. Jed Rankin explained the mask manufacturing flow for EUV without actinic inspection and how that absence pushes the burden of mask qualification to wafer fab verification of mask quality through wafer printing. The cost of this shift is primarily cycle time and process complexity.

When the discussion shifted to the tool suppliers of inspection systems, the inspection points during blank fabrication, mask build and at the wafer fab were matched to existing inspection tools and technologies. The blank manufacturer must identify phase defects. This is possible today, but only at low volumes. Phase defects must be detected before the absorber is deposited and with accurate location so that blanks and patterns can be matched during mask build in a process known as ‘pattern shift’. With effective pattern shift in place, phase defects remain covered with absorber on the final photomask so they have no imaging impact on wafer. During mask build, optical systems are available that meet throughput and initial 7nm node size requirements, but could miss EUV-specific imaging defects. Once pellicles are introduced, the pellicle must be transparent to the inspection wavelength for through-pellicle inspection, or waived in favor of inspecting wafers imaged with the pellicle in place. Electron beam mask and wafer inspection methods were mentioned by all tool suppliers as a viable option for detecting small defects, though this technique cannot be deployed through the pellicle. Fei Wang stated that resolution below 2nm has been demonstrated but that a very large gap in throughput requirements must be addressed with multi-beam architecture, higher data rates, lower noise and more advanced algorithms for detection.

The panel session was constructed to extract opinions from the audience as well as the panelists. The group consensus was that no new inspection tool development was needed for the initial introduction of EUV. To ensure high yields, actinic blank inspection is considered mandatory, even if actinic patterned mask inspection is not available. Finally, the existing tooling can find the required defects with a combination of blank, mask and wafer inspections, but the complex web of inspections and logistics is a strong motivator for the actinic inspection that could restore manufacturing flows to 193nm lithography baseline of complexity, accountability and cost.

Panelists and the audience responding to questions during the SPIE Photomask conference panel discussion. (photos courtesy of Bernd Geh, Carl Zeiss SMT)

electrons is often denoted as $\eta$. The profile of the backscattered exposure $g(x)$ is usually obtained by Monte-Carlo simulation and later approximated by Gaussian expression, with its 1 sigma of about 10um. Fig.5(a) shows the test pattern to demonstrate the proximity effect modeling. The test pattern is composed of three blocks; 55um wide 250nm L/S block, 60um wide block with big 4um patterns on both sides of each 250nm line, then again 55um wide 250nm L/S block. The test pattern was written on positive chemically amplified resist whose sensitivity was around 20uC/cm$^2$, coated on a normal chrome-on-glass substrate, and exposed by electron beam mask writer EBM-8000. Correction of proximity effect was not applied.

Fig.5(b) shows the measured CD results of 250nm lines. The CD becomes wider in the center block because of the bigger backscattered exposure from the 4um pattern. The measured CD can be expressed in a convolutional form as below:

$$CD(x) = CD_0 + K \int D(x') \cdot g(x - x')dx', \quad (2)$$

where $CD_0$ is the baseline CD of 250nm line, $K$ is a gain factor and $D(x') = 1$ on the exposed lines and patterns and $D(x') = 0$ in the unexposed area. By fitting the CD data of Fig.5(b) with the expression (2) with $g(x)$ as Gaussian profile, the best fitting radius was obtained as 9.1um.

To convert the convolutional expression of equation (2) to the neural network scheme, we use the simple 2-layer network described in Fig.2 and by equation (1). However as discussed in section 2, one must put a restriction on the weights to make them...
work as a point-spread function. Here, we put a constraint that for any two weights which have the same absolute distances between the input nodes and the output nodes, those two weights must have identical value.

\[ \omega_{ji} = \omega_{ji}' = g_k, \quad \text{if } |x_j - x_i| = |x_{ji} - x_i|, \quad \text{where } k = |i - j| \]  

(3)

There is another constraint to make offset \( b_j \) as constant number \( (\beta = b) \), otherwise these bias factors will absorb every variation in measured CD so that the errors will never propagate to the weights \( w_{ji} \). The conversion function was chosen to be linear \( f(x) = x \) in this example. We used standard back propagation technique for the learning algorithm.

Fig. 5(c) shows the fitting result after learning was performed. The residual 3 sigma was reduced to 2.9nm from original 4.0nm with Gaussian fitting in Fig. 5(b).

Fig. 6 compares the Gaussian profile of 9.1um radius and the tuned point-spread function after machine learning. This experiment was performed on the normal chrome-on-glass substrate, but the machine learning clearly predicted the existence of shorter range effect and multiple-Gaussian-like signature. Note that the machine learning extracted the backscattering energy profile only from the experimental data, without any help of physical knowledge or Monte-Carlo simulation.

Since we chose a conversion function to be linear \( f(x) = x \), the learning result is all the same as that obtained by traditional method such as inverse fourier transform or inverse matrix solution. This is because we adopted only the linear exposure model in equation (2). In future, if we adopt non-linear resist development model and if a more appropriate non-linear function is chosen for the conversion function \( f(x) \), the machine learning may have a chance to provide even better result than traditional methods.

### 3.2 Charging effect correction with 3-layer network model

Resist surface charging has been one of the top critical error sources which degrade the image placement accuracy of the advanced photomasks written by electron beam lithography. There are two major components in the modeling of charging mechanism, one is direct charging and the other is fogging charging. Fig. 7 shows the object model of charging effect. Input object is the exposure intensity and output object is the positional error. In between the input and output objects, there is a middle object to express the charging magnitude. Connecting the middle (charging) object and output (positional error) object is what we call response function \( R_x(x, y) \) and \( R_y(x, y) \) which describe the beam deflection by a unit...
charge, so it is a kind of point-spread function. Connecting the input (exposure) object and the middle (charging) object is described separately for direct charging and fogging charging. Direct charging is obtained from exposure intensity $D$ via a non-linear conversion function $Cd(D)$, whereas fogging charging is obtained via convolution of exposure intensity $D$ with a fogging kernel $g(x,y)$ to first obtain the fogging intensity $F$, then $F$ is converted by a linear function $Cf(F)$.

This object model is converted to a neural network model as shown in Fig.8. Fogging kernel is now expressed as Fig.5(c). Measured CD of the test pattern and machine learning result is Fig.6. Comparison of backscattering energy profile between Gaussian (blue) and machine learning result (red).

Again, we put similar restrictions as equation (3) to make the weights work as a point-spread function.

$$\begin{align*}
g_{ji} &= g_{pfw}, & \text{if } (x_j - x_i) = (x_f - x_f) \text{ and } (y_j - y_i) = (y_f - y_f), \\
R_{xji} &= R_{xpf}, \text{ and } R_{yji} = R_{ypf}, & \text{if } (x_k - x_i) = (x_f - x_f) \text{ and } (y_k - y_i) = (y_f - y_f),
\end{align*}$$

(4)

where $x = (x, y)$ is the coordinates of input (dose) layer nodes, $x = (x, y)$ is the coordinates of middle (charging) layer nodes, and $x = (x, y)$ is the coordinates of output (positional error) layer nodes. The difference between the equation (3) and (4) is that the constraint is made on the absolute distance between the nodes in the equation (3) so that the kernel becomes symmetric, whereas the constraint is made on the vector between the nodes in the equation (4) so it allows the kernel to be asymmetric.

We also note here that $Cd(D)$ to be non-linear function is important in this neural network model. If both $Cd(D)$ and $Cf(F)$ are the same linear function, two convolution can turn into one convolution then this 3-layer network can be converted to 2-layer network. As
described in section 3.1, 2-layer network with linear conversion function can be solved by any traditional methods. Since Cd(D) is non-linear function in this charging effect case, the use of this machine learning framework makes some sense.

Fig.9 shows the experimental result of positional error caused by charging effect. We wrote 20mm x 20mm exposure pad first then wrote 40mm x 40mm metrology matrix (Fig.10(a)) and measured the positional errors (Fig.10(b)). After applying conventional CEC, the residual error can be observed mainly on the left edge of the exposure pad (Fig.10(c)). Conventional CEC uses symmetric Gaussian fogging kernel as shown in Fig.10(a). After machine learning step was applied for a set of positional error maps with various pattern densities, the fogging kernel after learning showed obviously shifted and elliptically distorted shape as seen in Fig.10(b). The shift of fogging kernel to the left side is the effort of the machine to explain the remaining error on the left edge of the pad, and it was successful. By applying this learned fogging kernel together with the learned response function (not shown in this paper), the residual 3 sigma of X positional error was reduced from 2.1 to 1.5nm.

4. Summary

We presented a proof of concept of kernel tuning scheme by machine learning applied to the electron beam lithographic modeling, using two examples of PEC and CEC. Same scheme can be applied to the other modeling which uses any kind of kernel expression with convolution form. The tuned kernel can give researchers next suggestions to come up with better physical modeling or to understand the limitation of a too much simplified physical model which does not take into account the actual mechanical structure or electrical configuration. In future, it may be more common to use this type of practice in the electron beam lithographic modeling, replacing some portion of more dedicated work which could have been achieved only by Monte-Carlo simulation with highest level of physical knowledge in the past.

5. References


Figure 9. Experimental result of positional error caused by charging effect, (a) test layout, (b) result without correction, (c) residual error after conventional CEC is applied.

Figure 10. Fogging kernel profile, (a) before learning (symmetric) and (b) after learning (shifted and elliptic).
Industry Briefs

■ Expected EUV Insertion into Late 10nm and 7nm Technology Offerings

Recent EE Times articles indicate that insertion of extreme ultraviolet (EUV) lithography is certain in late 10 nm and 7 nm node technology offerings. It is expected that TSMC, Samsung, and Global Foundries will use EUV lithography to reduce the number of mask layers required in the back end of line (BEOL) and middle of line (MOL) modules. In addition, Intel is expected to make use of EUV in a late 10 nm-plus offering expected in 2019. EUV promises clear improvements in edge placement accuracy and cycle time by reducing the number of mask layers required for key modules. To reduce possible EUV contamination, ASML is expected to introduce pellicles in time to enable second generation 7nm technology offerings, but there is still work to be done to reach transmission values at targeted source power.

While EUV introduction into manufacturing is almost certain, important challenges exist for resist materials to enable 5 nm node offerings. In particular, resist related defectivity and required dose ranges will require improvements to ensure smooth EUV insertion into 5 nm. Another important area of development for EUV insertion is EUV mask process technology. Industry executives expect that EUV mask technology flows will likely see the insertion of multi-beam mask writers to enable quicker mask fabrication turnaround times. In addition, increased use of mask process correction is expected to become an increasing part of typical EUV mask process flows to enable more stringent mask shape fidelity requirements expected in late 7nm and 5 nm technology nodes. Inspection of EUV masks will most likely be done with existing inspection technology until actinic tools become available.

■ Record Year for DRAM

2017 was an extremely good year for DRAM which culminated in a record 4th quarter revenue of $21 billion dollars according to IC insights. Demand for DRAM was driven primarily by shortages in fab capacity and yield challenges for sub-20nm nodes. In addition, increased demand from server based systems common in cloud data centers coupled with increases in game console memory demands as well as mobile memory demand contributed to record sales. While DRAM has historically operated in a boom-bust cycle, it will be interesting to see if continued industry demand will soften downturns as smart phones and compute paradigms shift toward heavier use of artificial intelligence and cognitive compute applications.

■ Is an Industry Slowdown Expected in 2019 and Beyond?

2017 was a year of significant growth for the semiconductor industry and 2018 is expected to see solid growth. A worrying sign observed by industry analysts is the growth in inventory for key consumer electronics firms. Since a significant component of semiconductor sales is driven by consumer demand, analysts are concerned that inventory buildups at consumer firms is an early signal that 2019 may see a slowdown.
Join the premier professional organization for mask makers and mask users!

About the BACUS Group

Founded in 1980 by a group of chrome blank users wanting a single voice to interact with suppliers, BACUS has grown to become the largest and most widely known forum for the exchange of technical information of interest to photomask and reticle makers. BACUS joined SPIE in January of 1991 to expand the exchange of information with mask makers around the world.

The group sponsors an informative monthly meeting and newsletter, BACUS News. The BACUS annual Photomask Technology Symposium covers photomask technology, photomask processes, lithography, materials and resists, phase shift masks, inspection and repair, metrology, and quality and manufacturing management.

Individual Membership Benefits include:
- Subscription to BACUS News (monthly)
- Eligibility to hold office on BACUS Steering Committee

Corporate Membership Benefits include:
- 3-10 Voting Members in the SPIE General Membership, depending on tier level
- Subscription to BACUS News (monthly)
- One online SPIE Journal Subscription
- Listed as a Corporate Member in the BACUS Monthly Newsletter

You are invited to submit events of interest for this calendar. Please send to lindad@spie.org; alternatively, email or fax to SPIE.