A priori information in ptychographic image reconstruction for EUV mask metrology

Ricarda Nebling, Iacopo Mochi, Hyun-su Kim, Atoosa Dejkameh, Tao Shen, and Yasin Ekinci

Paul Scherrer Institut, 5232 Villigen PSI, Switzerland

ABSTRACT

EUV photomasks define the lithographic layers of chips, which are binary structures of relatively low versatility in contrast to specimen in biology or materials science. This knowledge can be used in EUV photomask imaging and inspection methods to improve the speed or sensitivity. We present here a total variation-based phase retrieval algorithm similar to previous methods by Chang et al. and Enfedaque et al. for EUV mask imaging and metrology. Total variation (TV) regularization exploits the binary structure of the reticles, enforcing a sparse sample gradient. We compare the TV regularized algorithm, PtychoADMM, to a standard phase retrieval approach, the difference map (DM). For simulated data containing Poisson noise, we do not observe a benefit from using the TV based PtychoADMM algorithm. The reconstructed image quality is similar, while PtychoADMM being a computationally more demanding method. In future, we will investigate if TV can recover information where the standard DM approach fails, e.g. for relaxed measurement requirements like a lower signal to noise ratio or less probe overlap in the ptychography scan.

Keywords: a priori information, total variation, ADMM, phase-retrieval algorithm, ptychography, actinic patterned mask inspection, EUV metrology

1. INTRODUCTION

With extreme ultraviolet (EUV) lithography is the leading high-volume manufacturing method in semiconductor industry, and with further downscaling in future technological nodes, the demand for actinic metrology of EUV reticles will increase towards higher sensitivity and throughput. At PSI, we develop RESCAN, an actinic patterned mask inspection (APMI) microscope dedicated to EUV mask inspection and review. Our tool is based on coherent diffraction imaging (CDI): the complex image of a mask is reconstructed with a phase retrieval algorithm instead of directly collecting the image with a lens-based system. This has the advantage of avoiding expensive imaging optics, and allows to easily mount a pellicle to protect the sample. A schematic outline of our current RESCAN optics is shown in Fig. 1 (a), the concept for CDI is shown in Fig. 1 (b). To retrieve an image of the mask with a large field-of-view, we use a technique called ptychography, which was first introduced for electron microscopy and has become very popular in X-ray imaging in the last decade. The sample is scanned in overlapping positions with a well-defined coherent illumination. The information diversity in the collected diffraction patterns allows for a robust reconstruction of the complex image.

The theoretical resolution of the current RESCAN tool is 34 nm, we have demonstrated phase and amplitude defect detection down to 50 × 50 nm on mask. At the moment, we are upgrading our tool optics to a Fourier synthesis illuminator, which will enhance the resolution to 20 nm on mask. Furthermore, we are also working on an improved image reconstruction procedure from the algorithmic side, by using the a priori information available on the EUV mask. Our goal is to exploit the prior knowledge to reduce the reconstruction time, and to relax the measurement requirements.

Here, we will discuss the usage of total variation (TV) regularization for EUV mask inspection. As the mask is a (quasi) binary structure (consisting of absorber and multilayer area), the gradient of the reconstructed image of the sample should be sparse. This is enforced by using total variation regularization, which was introduced for image denoising while preserving the edges. In the following, we will introduce a ptychographic image reconstruction algorithm based on TV regularization. We will show simulation results using this algorithm and compare it to a standard algorithm without TV.

Further author information:
E-mail: ricarda.nebling@psi.ch

© 2021 SPIE

Proc. of SPIE Vol. 11875  118750M-1
Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 17 Oct 2021
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
Figure 1. (a) Schematic of the current RESCAN setup: The EUV beam from the synchrotron source is cropped by a circular aperture with variable size and focused by the condenser mirror M1 onto the EUV reticle sample. The folding mirror M2 gives an angle of incidence of 6° on the sample. The beam is reflected from the mask and a diffraction pattern for each scanned position is collected on a EUV CCD detector in the far-field. (b) Coherent diffraction imaging principle of an EUV mask. No imaging optics are used, a diffraction pattern is collected by the EUV CCD detector. To collect a full ptychography scan, the sample is moved in the two indicated lateral directions, with some illumination overlap for adjacent scan points.

2. TOTAL VARIATION REGULARIZED PTYCHOGRAPHIC IMAGING

In ptychography, the object \( O \) (here the EUV mask sample), is scanned by the illumination function or probe \( P \) in overlapping positions. The exit wave \( \psi \) just behind the object is assumed to be the object function multiplied with the illumination function,\(^7\)

\[
\psi_j = P \circ S_j O, \tag{1}
\]

with \( j \) the current scan position for a total number of \( J \) scanned points, and \( \circ \) the pointwise multiplication. \( S_j \) describes a matrix that selects the \( j^{th} \) illuminated object region. In Fourier space, this is formulated as:

\[
A_j = \mathcal{F}(P \circ S_j O) \tag{2}
\]

with \( \mathcal{F} \) being the Fourier transform. To retrieve the object and probe from intensity measurements, one tries to solve the phase problem, or respectively to minimize the Fourier constraint,\(^7,13\)

\[
G(\Psi_j) = \frac{1}{2} \| \sqrt{|\Psi_j|^2 - \sqrt{I_j}} \|^2, \tag{3}
\]

that requires the reconstructed intensity, \( |\Psi_j|^2 \), to match the one measured on the detector, \( I_j \), while additionally enforcing the overlap constraint given in eq. 1.

To exploit the binary structure of the EUV masks, we also use the TV functional,

\[
\lambda \| \nabla O \|_1, \tag{4}
\]

in the problem formulation given below, enforcing a sparse object gradient.

All above constraints together give the phase-retrieval problem with total variation regularization that can be formulated as the following constraint optimization problem:

\[
\min_{O,P,\Psi} G(\Psi_j) + \lambda \| q \|_1 \tag{5}
\]

s.t. \( \Psi_j - A_j(P,O) = 0 \),

\[
q - \nabla O = 0,
\]
PtychoADMM

Initialization

\( O \in \mathbb{C}^{m \times m}, \quad S_j \in \mathbb{C}^{n \times n} \)
\( P \in \mathbb{C}^{n \times n} \)
\( \psi \in \mathbb{C}^{J \times n \times n} \)
\( q \in \mathbb{C}^{2 \times n \times n} \)
\( \Lambda_1 \in \mathbb{C}^{2 \times n \times n} \)
\( \Lambda_2 \in \mathbb{C}^{J \times n \times n} \)

for \( k = 0 \) to \( \text{maxit} \)

1. Fourier constraint
   \( \Psi_j^{(k+1)} = \text{FFT}(\psi_j^{(k)}) \)
   \( \Psi_j^{(k+1)} = \frac{T_1 + T_2}{1 + T_2} \circ \text{sign}(\Psi_j^{(k+1)}) \)
   \( \psi_j^{(k+1)} = i\text{FFT}(\Psi_j^{(k+1)}) + \Lambda_j^{(k)} \)

2. Probe update
   \( P_j^{(k+1)} = \sum_j (S_j \Omega^{(k)} + \psi_j^{(k+1)}) \)

3. Object update using the conjugate gradient method
   \( -r_1 \nabla O_j^{(k+1)} + r_2 \sum_j S_j^T ((P_j^{(k+1)})^2 \circ S_j O_j^{(k+1)}) + r_1 \nabla \cdot (\Lambda_j^{(k)} + q^{(k)}) + r_2 \sum_j S_j^T (P_j^{(k+1)} \circ \psi_j^{(k+1)}) \)

4. Update \( q \) and multipliers
   \( q_j^{(k+1)} = \max\{0, |\nabla O_j^{(k+1)}| - \Lambda_j^{(k)} - \frac{\Lambda_j^{(k)}}{r_1} \circ \text{sign}(\nabla O_j^{(k+1)} - \Lambda_j^{(k)}) \}
   \( \Lambda_j^{(k+1)} = \Lambda_j^{(k)} + \mu_j^{(k+1)} - \nabla O_j^{(k+1)} \)
   \( \Lambda_j^{(k+1)} = \psi_j^{(k+1)} - (P_j^{(k+1)} \circ S_j O_j^{(k+1)}) \)
   \( \psi_j^{(k+1)} = P_j^{(k+1)} \circ S_j O_j^{(k+1)} - \Lambda_j^{(k+1)} \)

\( P_j^{(\text{maxit})}, O_j^{(\text{maxit})} \)

Table 1. Pseudocode of the PtychoADMM algorithm.

with \( q \) an auxiliary variable for the TV constraint.

The constrained optimization problem is solved similarly to Chang et al.\(^{14}\) and Enfedaque et al.\(^{15}\) using a Lagrange multiplier technique called alternating direction method of multipliers (ADMM). As the problem is non-smooth, one cannot directly apply a gradient based optimization technique.\(^{16}\) ADMM finds the solution by using partial derivatives of the augmented Lagrangian.\(^{17}\) Minimizing the derivatives for all variables directly gives update functions, that can be summarized to the algorithm presented in Tab. 1. The algorithm is called PtychoADMM, after Chang et al.\(^{14}\) and Enfedaque et al.\(^{15}\) To the best of our knowledge, solving the blind ptychography problem (retrieve object and probe) with total variation regularization has not been demonstrated before.

In the pseudocode, \( \Lambda_1 \) and \( \Lambda_2 \) are the Lagrangian multipliers, \( r_1, r_2 \), and \( \lambda \) are scalar parameters. \( \text{sign} \) stands for the complex sign function, that keeps the phase but sets the amplitude to unity with respective sign. FFT denotes the fast Fourier transform algorithm, iFFT its inverse.

3. SIMULATION RESULTS

We compare the PtychoADMM algorithm to a standard phase-retrieval algorithm, the difference map (DM),\(^{13}\) for simulated data. The object is generated from a mask layout file and contains a Siemens star pattern. We assume a binary absorber and add a phase factor corresponding to a 70 nm thick TaBN layer under 6° angle of incidence. The illumination function is assumed to be the aerial image on the object of the circular aperture with aperture diameter of 0.5 mm. The illumination wavelength is 13.5 nm, and the sample is assumed to be placed 62 mm from the detector. We further assume a 1800 × 1800 pixel detector with a pixel size of 13.5 µm. To generate the diffraction patterns, the illumination function is multiplied with the object. The Fourier transformation squared gives the intensity on the detector plane. The diffraction patterns are scaled to a maximum of 65'000.
counts per pixel (as for a uint16 CCD camera), and we add Poisson noise. We choose a circular scan pattern of three concentric rings with a step size of 1.2 µm.

In the following, we will compare four reconstructions. First, using the standard DM algorithm as described by Thibault et al.,\textsuperscript{13} Second, from a DM version with TV regularization, inspired by Tanksalvala et al.,\textsuperscript{18} that combined a RAAR approach with TV. Then we discuss reconstructions of the simulated dataset using the PtychoADMM algorithm with and without TV regularization. The case without TV for PtychoADMM is obtained by setting parameters \( r_1 \) and \( \lambda \) to zero, and removing the update for \( q \) and the Lagrangian multiplier \( \Lambda_1 \) from the iteration. In all reconstructions, we use the following parameters: \( r_1 = 0.5 \), \( r_2 = 1 \), and \( \lambda = 0.05 \). These are chosen for DM and PtychoADMM to be comparable. The object is initialized as unit amplitude with a small random phase, the initial probe is the same one as used to generate the dataset. The initial exit wave \( \psi \) is set equal to the probe, all other variables are initialized as zero.

The algorithms are run for 500 iterations each, starting to update the probe from the 50\textsuperscript{th} iteration on. We use a binning of 2 \( \times \) 2 pixels in all reconstructions to shorten the reconstruction time. For the DM with TV run, the TV regularization is started from the 80\textsuperscript{th} iteration, and is applied to the central 800 \( \times \) 800 object pixels only.

\[ E_{\text{obj}} = \frac{\sum_j \sum_{u,v} |O_{uv}^{\text{ref}} - \hat{O}_{uv}|^2}{\sum_j \sum_{u,v} 1}. \]  

(6)

It measures the difference of the reconstructed object magnitude to a reference, here the ideal object magnitude filtered by a lens with NA according to the simulated experimental setup. All object error curves are plotted in Fig. 4 (a). All four algorithms reach a low object error after a few iterations, only PtychoADMM (with TV) seems to diverge slightly with more iterations. The lowest object error for each curve is plotted in Fig. 4 (b) (the filled dots), with the respective iteration number where the error is reached written on top. All lowest errors are reached before or at 215 iterations, indicating that 500 iterations would not be necessary to reconstruct the object. Also, they all are about 0.02 a.u., with a difference from the largest to the lowest object error of 0.004 a.u.. As observed by eye from the reconstructions, all four algorithms seem to reconstruct the object well with little noise and artifacts remaining.

To quantitatively compare the reconstructed object quality, we also use the object error:\textsuperscript{19}

To show that a difference of 0.004 a.u. is not significant, we additionally run 200 iterations using the rPIE algorithm,\textsuperscript{20} starting from the lowest object error reconstruction as initial guess. It has been shown that the rPIE algorithm robustly converges to the optimal solution if the starting guess is close enough to avoid local minima.\textsuperscript{21} The object error for the rPIE reconstructions is shown in Fig. 4 (a) as the dashed lines, with the lowest object
Figure 3. (a) - (d) Reconstructed objects cropped to the central Siemens star pattern at the respective lowest object error iteration. For (a) DM, (b) DM with TV regularization, (c) PtychoADMM without TV, and (d) PtychoADMM. In (e) - (h), the respective reconstructions after additional 200 iterations with the rPIE algorithm, starting from the lowest object error iteration. The same scale as shown in Fig. 2 (b) applies here.

The error reached in (b) as the hollow dots. Using rPIE, the object error is going down by more than 0.01 a.u. for all studied cases. The corresponding reconstructions are shown in Fig. 3 (e) to (h). For all algorithms, rPIE improves the reconstructed image quality visibly, reducing the noise and showing sharper features.

Figure 4. (a) Object error for the different compared algorithms. The dashed lines correspond to additional 200 iterations with the rPIE algorithm, using the lowest object error reconstruction as initial guess. (b) Plot of the lowest object error for each algorithm. The filled dot is the lowest error after 500 iterations of the respective algorithm, with the iteration number where the error is reached written on top. The hollow dot corresponds to the reconstruction after additional 200 iterations of rPIE.

We furthermore reconstruct a second dataset, assuming an identical experimental setup, to calculate the Fourier ring correlation (FRC). All correlations are shown in Fig. 5, we use the 1-bit threshold criterion to calculate the reproducible resolution for each algorithm.

DM, DM with TV, and PtychoADMM all show a resolution around 230 nm, only PtychoADMM without TV has a smaller reproducible resolution of 168 nm. Nevertheless, the PtychoADMM curve without TV (red) shows a small dump around 5 $\mu$m$^{-1}$, almost crossing the threshold. The smaller resolution could be well within the variance of the correlation. To complete the picture, more independent reconstructions and respective FRC curves would be necessary.

4. CONCLUSIONS AND OUTLOOK

We present here a total variation regularized ptychographic imaging algorithm, that exploits the fact that EUV mask are binary structures. The PtychoADMM algorithm is similar to the method described by Chang et al. and Enfedaque et al., and combines blind ptychography (retrieval of probe and object) with TV regularization. For the simulated Siemens star dataset, we cannot observe a benefit from using TV regularization during the reconstruction procedure. A significant improvement in retrieved image quality is achieved by additionally using the rPIE algorithm. As PtychoADMM has a large memory footprint and is a computationally demanding method due to the additional auxiliary variables and multipliers used, we currently do not see an advantage of using PtychoADMM over a standard approach like the DM. In future, we will study whether TV regularized image reconstruction is advantageous for relaxed measurement conditions, like a lower signal to noise ratio or an increased step size with less overlap. The additional TV constraint might help to recover information where DM fails. Furthermore, we will investigate the possibility to use the mask design file as prior knowledge, similarly to Ansuinelli et al. The main challenge for a layout based prior is the alignment of the reconstructed image to the mask reference during the iteration procedure.

ACKNOWLEDGMENTS

We thank Manuel Guizar-Sicairos for his valuable insights and discussions. Part of this research was performed at the Swiss Light Source, Paul Scherrer Institut. Co-authors (R.N. and A.D.) thank the Swiss National Science Foundation for financial support (SNF Grant Number: 200021_172768).

REFERENCES


