Convolution-variation separation method for efficient modeling of optical lithography

Shiyuan Liu,[1,2,*] Xinjiang Zhou,[1] Wen Lv,[1] Shuang Xu,[1] and Haiqing Wei[2]

[1]*Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China*

[2]*State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

*Corresponding author: shyliu@mail.hust.edu.cn

Received March 18, 2013; revised April 22, 2013; accepted May 5, 2013; posted May 8, 2013 (Doc. ID 187263); published June 18, 2013

We propose a general method called convolution-variation separation (CVS) to enable efficient optical imaging calculations without sacrificing accuracy when simulating images for a wide range of process variations. The CVS method is derived from first principles using a series expansion, which consists of a set of predetermined basis functions weighted by a set of predetermined expansion coefficients. The basis functions are independent of the process variations and thus may be computed and stored in advance, while the expansion coefficients depend only on the process variations. Optical image simulations for defocus and aberration variations with applications in robust inverse lithography technology and lens aberration metrology have demonstrated the main concept of the CVS method. © 2013 Optical Society of America

OCIS codes: (110.5220) Photolithography; (110.4980) Partial coherence in imaging; (110.1758) Computational imaging.

http://dx.doi.org/10.1364/OL.38.002168

Process variations, such as defocus and lens aberrations, always exist in real lithography processes [1]. Thus, fast and accurate lithography simulation considering process variations, especially the optical ones, is highly desirable. For optical imaging in Hopkins' theory, it is rather expensive to calculate the four-dimensional transmission cross coefficient (TCC) and integrals [2–4]. Although the sum-of-coherent-systems (SOCS) method [5] and the optimal coherent approximation [6] achieve a significant speedup by eigenanalyzing the TCC and neglecting small eigenvalues, the TCC kernels are usually obtained under the nominal (best) process condition. For a different process condition, a simulator may have to repeat the costly eigendecomposition and mask-kernel convolutions.

In this Letter, we propose a general methodology called convolution-variation separation (CVS) to speed up variational image simulations without sacrificing accuracy. The optical image $I(\mathbf{x};\mathbf{v})$ generated by a partially coherent imaging system is represented by a bilinear form [7]

$$I(\mathbf{x};\mathbf{v}) = \iint t(\mathbf{x}-\mathbf{x}_1,\mathbf{x}-\mathbf{x}_2;\mathbf{v})o(\mathbf{x}_1)o^*(\mathbf{x}_2)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2, \quad (1)$$

where $\mathbf{x}$ is the two-dimensional (2D) spatial coordinate, $\mathbf{v}$ is a vector representing a combination of multiple process parameters, $o(\mathbf{x})$ is the mask transmittance function, $*$ denotes complex conjugation, and

$$t(\mathbf{x}_1,\mathbf{x}_2;\mathbf{v}) = s(\mathbf{x}_1-\mathbf{x}_2)p(\mathbf{x}_1;\mathbf{v})p^*(\mathbf{x}_2;\mathbf{v}) \quad (2)$$

is called the double-impulse response (DIR) function [7], whose Fourier transform is the familiar TCC. Here $s(\mathbf{x}_1,\mathbf{x}_2)$ is the mutual intensity function of the source, and $p(\mathbf{x};\mathbf{v})$ is the point spread function (PSF) of the optical system with process variations.

A physical quantity dependent on both a spatial coordinate and other parameters can be represented by a sum of multiple series expansion terms, with each term consisting of one function dependent only upon the spatial coordinate and another function dependent only upon the other parameters. Such a method of separation of variables has been employed in a variety of applications [8]. In the present context, the PSF function with process variations in Eq. (2) may be expanded into a series,

$$p(\mathbf{x};\mathbf{v}) = \sum_m a_m(\mathbf{v})p_m(\mathbf{x}), \quad (3)$$

where the mth basis PSF $p_m(\mathbf{x})$ is dependent only on $\mathbf{x}$, while the mth expansion coefficient $a_m(\mathbf{v})$ is dependent only on $\mathbf{v}$. When such a variable-separated PSF is substituted into Eq. (2), we obtain a variable-separated DIR. Alternatively, the DIR itself may be expanded as

$$t(\mathbf{x}_1,\mathbf{x}_2;\mathbf{v}) = \sum_m c_m(\mathbf{v})t_m(\mathbf{x}_1,\mathbf{x}_2), \quad (4)$$

with a suitably chosen set of linearly independent basis functions $\{c_m(\mathbf{v})\}$, which are usually orthogonal polynomials or simply powers of the variation variables, while for each m, $t_m(\mathbf{x}_1,\mathbf{x}_2)$ is the variable-separation counterpart of $c_m(\mathbf{v})$, called the mth basis DIR function. Each $t_m(\mathbf{x}_1,\mathbf{x}_2)$ may be singular-value decomposed (SVD) as

$$t_m(\mathbf{x}_1,\mathbf{x}_2) = \sum_n \mu_{mn}\phi_{mn}\varphi^*_{mn}, \quad (5)$$

where $\mu_{m1} \ge \mu_{m2} \ge \ldots > 0$ are singular values, and $\{\phi_{mn}(\mathbf{x})\}$ and $\{\varphi_{mn}(\mathbf{x})\}$ are the corresponding left and right singular vectors. Consequently, we have the following for the image intensity $I(\mathbf{x};\mathbf{v})$ under process variations:

$$I(\mathbf{x};\mathbf{v}) = \sum_m c_m(\mathbf{v})I_m(\mathbf{x}), \quad (6)$$

$$I_m(\mathbf{x}) = \sum_n \mu_{mn}[\phi_{mn}\otimes o(\mathbf{x})][\varphi_{mn}\otimes o(\mathbf{x})]^*, \quad (7)$$

 © 2013 Optical Society of America

where $I_m(\mathbf{x})$ is called the mth basis image that is independent of process variations, and $\otimes$ denotes 2D convolution. When all $\{c_m(\mathbf{v})\}$ are all real valued, the basis DIR functions $t_m(\mathbf{x}_1, \mathbf{x}_2)$ are all Hermitian, and $\{\phi_{mn}\}$ and $\{\varphi_{mn}\}$ coincide, so the number of necessary convolutions is cut in half, and the intensity reduces to the familiar form as a sum of squares [9,10]. In practical numerical calculations, the physical quantities become matrices and vectors, and the series expansions are truncated. Fortunately, their rapid convergence enables good approximation with a relatively small number of terms.

The significance of Eqs. (6) and (7) is that the fixed mask-kernel convolutions are fully separated from the variable process parameters and can be precomputed, so that optical images under a large range of process variations can be quickly calculated. This is the fundamental rationale behind the CVS methodology [11]. If L is the number of discrete sampling points of process variations, the conventional approaches need to generate and use L sets of process models. Let $T(1)$ denote the computational complexity (computing time) for simulating one sampling point of process variations, and let $T(L)$ be the total computational complexity for simulating L sampling points of process variations; then the following formula holds true for all conventional approaches:

$$T(L) = T(1) \times L. \tag{8}$$

Let T_0 denote the computational complexity of convolutions obtaining all the basis images $I_m(\mathbf{x})$ in Eq. (7), and let δT denote the extra complexity for summing up the series in Eq. (6). It is usually the case that $\delta T \ll T_0$ and $T_0 \approx T(1) \times (M+1)$, where M is the truncation order of Eq. (6). Using the method of CVS, the total complexity for simulating one sampling point of process variations is $T_{\mathrm{CVS}}(1) = T_0 + \delta T$, whereas the cost for simulating L sampling points of process variations would be

$$T_{\mathrm{CVS}}(L) = T_0 + \delta T \times L. \tag{9}$$

The superior scalability of CVS is apparent when $L > M + 1$.

Our first example demonstrating the CVS method is in the robust inverse lithography technology (ILT) [12]. For a clearer demonstration, we consider only the variations of defocus h. By simply choosing $c_m(\mathbf{v})$ in Eq. (6) as the mth power of h, we may approximate a defocused image as the familiar truncated Taylor series $I(\mathbf{x}; h) = \Sigma_{m=0}^{M} h^m I_m(\mathbf{x})$, and the mth basis image $I_m(\mathbf{x})$ can be efficiently calculated by Eq. (7) or by a fitting approach [13,14]. This Taylor expansion serves as a good example for the general CVS method, as it clearly demonstrates the spirit of separating the basis images from the variation coefficients. For an essentially real-valued mask, the basis images for odd defocus terms vanish [13], and it suffices to set the truncation order as $M = 4$ for image errors on the order of 10^{-4}.

We formulate the robust ILT problem as

$$\hat{o} = \arg\min_{o} \left\{ \sum_{k=1}^{L} [\zeta(h_k)(\|\mathrm{sig}[I(\mathbf{x}; h_k)] - \tilde{o}\|_2^2)] + \eta R(o) \right\}, \tag{10}$$

where $\hat{o}$ is the optimized mask, $\tilde{o}$ is the desired pattern, $\zeta(\cdot)$ is the density of a prescribed distribution of defocus, $\{I(\mathbf{x}; h_k)\}$ are the images corresponding to L sampled defocus values $\{h_k\}$, $R(o)$ is a quadratic regularization term with a weight η to promote mask manufacturability, $\mathrm{sig}[\cdot]$ is a sigmoid function modeling the resist effect, and $\|\cdot\|_2$ is the L_2 norm. Equation (10) may be solved by an optimization algorithm [15], in which the forward lithography simulation is iterated many times and each iteration requires calculating a total of L defocused images.

We performed robust ILT simulations for a desired pattern with critical dimension (CD) of 45 nm. As shown in Fig. 1(a), the desired pattern is commonly encountered in the design of static random-access memory (SRAM) circuits. The pattern had a size of 401×401 pixels with a pixel size of 2.5 nm. The partially coherent illumination source was set as annular with $\sigma_{\mathrm{out}}/\sigma_{\mathrm{in}} = 0.8/0.6$, a wavelength of 193 nm, and an NA of 1.35. The steepest-descent algorithm was used to solve Eq. (10), with the regularization weight $\eta = 0.02$. We adopted two focus blur functions to model focus variations [16]. One is a Gaussian distribution $\zeta_1(h) = \exp[-h^2/2\sigma^2]$ with $\sigma = 20$ nm to characterize the stage vibration, and the other is a modified Lorentzian function $\zeta_2(h) = \Gamma^2/(\Gamma^2 + |2h|^2)$ with $\Gamma = 30$ nm to represent the laser bandwidth. By sampling from −80 to 80 nm at each 2 nm, we obtained a set of $L = 81$ defocus points for each defocus distribution. As shown in Figs. 1(b) and 1(c), the optimized mask patterns differ for different defocus distributions, and are dramatically different from that under the nominal condition as shown in Fig. 1(d). This correlates the high sensitivity of ILT solutions to changes in parameters and process assumptions. It is particularly interesting to note that subresolution assist features (SRAFs) were generated in Figs. 1(b) and 1(c) but not in Fig. 1(d). One reason might be that a defocused image has a lower image intensity compared to an in-focus image, and thus SRAFs tend to be produced to increase the light passing through and to compensate for the loss of intensity. Figure 2 depicts the edge placement error (EPE) and the corresponding exposure-defocus (E-D) trees by setting the EPE errors to within ±10% of the CD targets. As expected, the optimized mask patterns by the robust ILT achieve smaller overall EPEs and improved E-D windows compared to that optimized for the nominal condition only. This is because the robust ILT further considers the process variations. The robust ILT benefits greatly from the CVS method for efficient simulations. With $L = 81$ defocus sampling points in each ILT iteration, the runtime was $T(L) = 415$ s for the conventional approach, while it reduced to only $T_{\mathrm{CVS}}(L) = 15.43$ s when using the CVS method. These simulations were

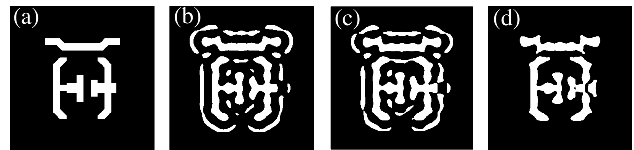

Fig. 1. (a) Desired SRAM pattern, and the optimized mask patterns obtained after 100 iterations under (b) defocus variation with a Gaussian distribution, (c) defocus variation with a modified Lorentzian distribution, and (d) nominal condition.

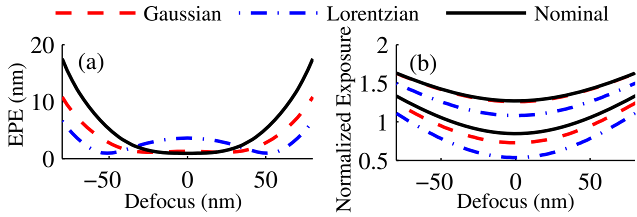

Fig. 2. (a) EPEs and (b) E-D trees of the optimized mask patterns obtained, respectively, under the nominal condition and defocus variations with the Gaussian and the modified Lorentzian distributions.

conducted with running MATLAB codes on an HPZ800 Workstation with 3.46 GHz Opteron processors.

Our second example demonstrating the CVS method is an efficient simulation of imaging under lens aberrations. We consider Zernike coefficients up to the Kth-order, so $\mathbf{v} = \mathbf{Z} = (Z_1, Z_2, \ldots, Z_K)^T$ is a $K \times 1$ vector representing process variations. Under small aberrations, the aberration-induced intensity may be approximated by a $\mathbf{Z}$-power series truncated to the first-order,

$$I(\mathbf{x}; \mathbf{Z}) = I_0(\mathbf{x}) + I_1(\mathbf{x}) = I_0(\mathbf{x}) + \sum_{k=1}^{K} Z_k I_{\text{lin}}^{(k)}(\mathbf{x}), \quad (11)$$

where $I_0(\mathbf{x})$ is called the aberration-free basis image, and $I_{\text{lin}}^{(k)}(\mathbf{x})$ represents the linearly aberrated basis images. Equation (11) is in CVS form, as the Zernike coefficients are separated from their corresponding basis images. For a given mask at a given defocus, all these basis images can be precomputed and stored in advance, according to Eq. (7). This has motivated us to develop a method for small aberration metrology in partially coherent systems by using a single defocused intensity measurement $I(\mathbf{x}; \mathbf{Z})$, in which the basis image $I_{\text{lin}}^{(k)}(\mathbf{x})$ is defined as the aberration sensitivity for the kth Zernike order, and is used for reconstructing the corresponding Zernike coefficient.

Figure 3 depicts one of the image intensity simulation results for a binary mask pattern as shown in Fig. 3(a). The induced aberration is shown in Fig. 4(a) by choosing Zernike coefficients Z_2 to Z_{37} randomly from the range [−15 mλ, 15 mλ], which leads to a total aberration ranging from −60 to 60 mλ. The partially coherent illumination source was set as quadrupole with $\sigma_{\text{out}}/\sigma_{\text{in}} = 0.8/0.4/45°$, a wavelength of 193 nm, an NA of 0.8, and a defocus of 180 nm. As shown in Fig. 3, the normalized intensity error is on the order of 10^{-4}, which indicates a good accuracy of image calculations for small aberrations. Figure 4 depicts the wavefront reconstructed from the measured Zernike coefficients up to the 37th order, showing a satisfactory accuracy on the order of 2 mλ for wavefront metrology. However, this accuracy degrades when the aberration becomes larger; i.e., the reconstructed aberration error is more than 2 mλ when the aberration surpasses 60 mλ. To improve the accuracy, we need to consider further the higher Zernike cross terms, such as the quadratic terms [17,18], by truncating the $\mathbf{Z}$-power series to a higher order.

In conclusion, a general CVS formulism is proposed for efficient simulations of imaging under process variations. The CVS advantage is clearly demonstrated in examples

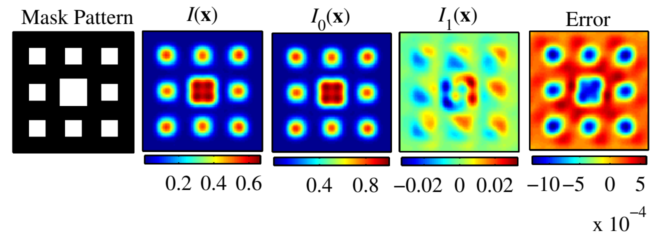

Fig. 3. Binary mask pattern and image intensity results under the aberration shown in Fig. 4(a).

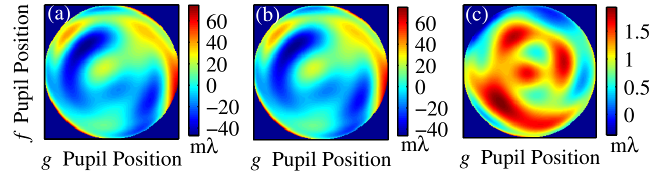

Fig. 4. Example of aberration reconstruction. (a) Unknown aberration and (b) reconstructed aberration, and (c) wavefront error.

of robust ILT and aberration metrology, and is expected to strengthen in applications with more process variables.

This work was funded by the National Natural Science Foundation of China (grants 91023032, 51005091, and 51121002), the Specialized Research Fund for the Doctoral Program of Higher Education of China (20120142110019), and the National Science and Technology Major Project of China (2012ZX02701001).

References

1. A. K. Wong, *Optical Imaging in Projection Microlithography* (SPIE, 2005).
2. H. H. Hopkins, Proc. R. Soc. A **217**, 408 (1953).
3. E. Kintner, Appl. Opt. **17**, 2747 (1978).
4. K. Yamazoe, J. Opt. Soc. Am. A **25**, 3111 (2008).
5. N. B. Cobb, "Fast optical and process proximity correction algorithms for integrated circuit manufacturing," Ph.D. dissertation (University of California at Berkeley, 1998).
6. Y. C. Pati and T. Kailath, J. Opt. Soc. Am. A **11**, 2438 (1994).
7. B. E. A. Saleh, Opt. Acta **26**, 777 (1979).
8. W. Miller, Jr., *Symmetry and Separation of Variables* (Cambridge University, 1984).
9. A. Starikov and E. Wolf, J. Opt. Soc. Am. **72**, 923 (1982).
10. B. J. Davis and R. W. Schoonover, Opt. Lett. **34**, 923 (2009).
11. H. Wei, "Computational efficiency in photolithographic process simulation," U.S. patent 7,788,628 (August 31, 2010).
12. Y. Shen, N. Jia, N. Wong, and E. Y. Lam, Opt. Express **19**, 5511 (2011).
13. P. Yu, D. Z. Pan, and C. A. Mack, Proc. SPIE **6156**, 397 (2006).
14. K. Yamazoe and A. R. Neureuther, Appl. Opt. **50**, 3570 (2011).
15. A. Poonawala and P. Milanfar, IEEE Trans. Image Process. **16**, 774 (2007).
16. T. Brunner, D. Corliss, S. Butt, T. Wiltshire, C. P. Ausschnitt, and M. Smith, J. Microlith. Microfab. Microsyst. **5**, 043003 (2006).
17. S. Y. Liu, W. Liu, and T. T. Zhou, J. Micro/Nanolith. MEMS MOEMS **10**, 023007 (2011).
18. S. Y. Liu, S. Xu, X. F. Wu, and W. Liu, Opt. Express **20**, 14272 (2012).