Available online at www.sciencedirect.com

ScienceDirect

Microelectronic Engineering 84 (2007) 725–728

MICROELECTRONIC ENGINEERING

www.elsevier.com/locate/mee

Improvement of model kernel representation in process simulation by taking pattern correlation into account

Jianliang Li *, Qiliang Yan, Abani Biswas, Lawrence S. Melvin III

Synopsys Inc., 2025 NW Cornelius Pass Road, Hillsboro, OR 97124, USA

Available online 26 January 2007

Abstract

In the modern photolithography simulation, the computation demand on resolution enhancement techniques (RETs) and optical proximity corrections (OPCs) is proportional to the simulation runtime of the model, which is dependant on the number of the kernels retained with the constrain of the model accuracy. Thus, it is essential to retain as few kernels as possible in the model calibration. Traditionally, the kernels are retained based upon their contribution to the aerial image, which is solely determined by the magnitudes of the eigenvalues. This method works well for arbitrary photolithography masks. However, real masks are never arbitrary and random. Instead, they have regular shapes and arrangements as governed by design rules, indicating the contributions from the retained kernels are statistically correlated to each other. By taking such correlations into account, the system representation can be improved to contain fewer kernels for a constant model accuracy. In this paper, the mathematical derivation of the pattern correlation concept is discussed and the concept is applied to a contact layer illuminated by a Quasar optical system with $\lambda = 193$ nm and NA = 0.8. Significant improvement of model kernel representation is observed, four improved kernels vs 15 original kernels, and the new methodology is justified by comparing the difference of the aerial image intensities between the full kernel representation and the retained kernels representation at sampling points.

Keywords: Photolithography; Optical kernels; OPC; RET; Aerial image

1. Introduction

The Hopkin's equation underlies most of today's aerial image simulators used for photolithography and optical proximity correction (OPC). Based on the Hopkins equation, the aerial image intensity can be expressed as [1]:

$$I(x,y) = \int\int\int\int M(x_1,y_1)T(x-x_1,y-y_1;x-x_2,y-y_2)M^*(x_2,y_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2, \quad (1)$$

where $M(x,y)$ is the mask transmission function (binary mask and dark field are assumed through the study), $T(x-x_1,y-y_1;x-x_2,y-y_2)$ is so called transmission cross coefficient (Tcc) matrix. The 4-D model $T(x_1,y_1;x_2,y_2)$ can be decomposed into a series of its eigenvectors

$$T(x_1,y_1;x_2,y_2) = \sum_{i=1}^{\infty}\lambda_i K_i(x_1,y_1)K_i^*(x_2,y_2), \quad (2)$$

here λ_i is the eigenvalue corresponding to the eigenvector $K_i(x,y)$. With the expansion of a series of eigenvectors and eigenvalues, the image intensity becomes

$$I(x,y) = \sum_{i=1}^{\infty}\lambda_i\left|\int\int M(x_1,y_1)K_i(x-x_1,y-y_1)\mathrm{d}x_1\mathrm{d}y_1\right|^2. \quad (3)$$

Each eigenvector $K_i(x,y)$ (optical kernel) can be interpreted as the transfer function of a coherent imaging system. The overall aerial image intensity is the sum of the images produced by an infinite number of coherent systems.

In practice, only a certain number of optical kernels are retained based upon their importance to the aerial image, as determined by the magnitudes of the eigenvalues. This scheme works well for arbitrary photolithography masks.

* Corresponding author. Tel.: +1 503 547 6224.
E-mail address: jlli@synopsys.com (J. Li).

0167-9317/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.mee.2007.01.031

However, real masks are never arbitrary or random. Instead, they have regular shapes and arrangements as governed by design rules. For instance, the design and mask rules limit the smallest feature size and space. This implies that the kernels containing spatial frequencies higher than that determined by the design rules have limited response to the pattern, allowing new kernels to be constructed for those high frequencies. Therefore the contributions from the retained kernels are statistically correlated. Taking such correlation into account will lead to a better representation of the retained Kernels without degrading model fidelity. This correlation may also be employed to reduce the computational demands of optical proximity correction and resolution enhancement techniques (RETs) by reducing the number of retained kernels while maintaining constant model fidelity.

This study will discuss the theoretical derivation of the improved kernel representation, which is followed by the application of the concept on a real optical model and contact and polysilicon layer. The comparison of the model accuracy with original kernels and improved kernels is also presented.

2. Theoretical background

To simplify the analysis, the aerial image intensity in Eq. (3) can be expressed in a vector and matrix format since all of the masks and kernels are decomposed into the Fourier–Bessel functions [2]. Let the mask be expressed as a vector X, and the eigenvectors (or eigenfunctions) of Tcc as u_i, then the intensity can be written as

$$I(x,y) = \sum_{i=1}^{\infty}(X^{\mathrm{T}}u_i)^2. \tag{4}$$

It should be noted that each element in X and u_i is a function of reticle position (x,y). In the above equation, the summation is usually dominated by the first several terms, if the eigenvectors are sorted in the way that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_\infty$. Therefore, in practice, the summation can be truncated and only first n terms are retained as,

$$I(x,y) \approx \sum_{i=1}^{n}(X^{\mathrm{T}}u_i)^2. \tag{5}$$

In the above scheme, the truncation to n terms is solely based on the magnitudes of eigenvalues of Tcc matrix. This is the best solution for *arbitrary* masks. In reality, however, masks often have some certain patterns as dictated by the design rules. For example, the mask usually has a minimum feature size and minimum space. This implies that the components of the mask vector X are statistically correlated to each other. Taking such correlation into account might lead to further reduction of computational demands without loosing the mode accuracy.

To find a better representation of the model kernels, consider a new set of functions w_i, which is generated by rotating the original eigenfunctions u_i:

$$\begin{aligned}(w_1, w_2, \ldots, w_n) &= \mathbf{W} = \mathbf{UR} \\ &= (u_1, u_2, \ldots, u_n)(r_1, r_2, \ldots, r_n)\end{aligned} \tag{6}$$

where $\mathbf{R}$ is an orthogonal rotation matrix [3], $\mathbf{RR}^{\mathrm{T}} = \mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$ (identity matrix).

After the matrix rotation, the aerial image intensity becomes,

$$\begin{aligned}I(x,y) &= X^{\mathrm{T}}\mathbf{UU}^{\mathrm{T}}X = X^{\mathrm{T}}\mathbf{URR}^{\mathrm{T}}X^{\mathrm{T}} \\ &= X^{\mathrm{T}}\mathbf{WW}^{\mathrm{T}}X = \sum_{i=1}^{n}(X^{\mathrm{T}}w_i)^2 \\ &= \sum_{i=1}^{n}(X\mathbf{U}^{\mathrm{T}}r_i)^2\end{aligned} \tag{7}$$

The above equation indicates that after an arbitrary rotation of the eigenfunctions, the original signal values can still be evaluated from the contribution of the rotated eigenfunctions, in which the relative contribution of each rotated eigenvector in the summation is different than the contribution of the original eigenvector. In order to squeeze the contribution of the signal value as much as possible into the first few model terms, thus employing smallest number of rotated eigenvectors to represent for a constant model accuracy, the optimum rotation matrix $\mathbf{R}$ needs to be identified.

One method to squeeze the signal value into the first few terms is to maximize each term in the summation, $X^{\mathrm{T}}\mathbf{U}r_i$, one by one with the best vector r_i, which is a unit vector and represents the ith column of the matrix $\mathbf{R}$. For instance, consider the first term, $X^{\mathrm{T}}\mathbf{U}r_1$. Suppose there is a mask with s sampling points. A new matrix $\mathbf{X}$ can be constructed with the mask vectors X_i at all of the sampling points as,

$$X = \begin{pmatrix} X_1^{\mathrm{T}} \\ X_2^{\mathrm{T}} \\ \vdots \\ X_s^{\mathrm{T}} \end{pmatrix}. \tag{8}$$

Then the first term contribution of optical signal at all sampling points can be described by a vector e_1:

$$e_1 = \begin{pmatrix} X_1^{\mathrm{T}}Ur_1 \\ X_2^{\mathrm{T}}Ur_1 \\ \vdots \\ X_s^{\mathrm{T}}Ur_1 \end{pmatrix} = XUr_1. \tag{9}$$

Note, the norm of the above vector, $e_1^{\mathrm{T}}e_1$, represents the overall contribution of optical signal from the first time in the summation (7), and hence, it needs to be maximized. Since r_1 is the first column of the orthogonal matrix $\mathbf{R}$, and

$$e_1^{\mathrm{T}}e_1 = r_1^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{XU}r_1, \tag{10}$$

the best r_1 would be the unit vector that maximizes the quadratic form (10). Following the same argument, the ith column of the rotation matrix $\mathbf{R}$ is the unit vector that maximizes the quadratic form (10), subject to the con-