Derivation of line edge roughness based on analytic model of stochastic exposure distribution

Rui Guo and Soo-Young Lee

Department of Electrical and Computer Engineering, Auburn University, Auburn, Alabama 36849

Jin Choi, Sang-Hee Lee, In-Kyun Shin, Chan-Uk Jeon, Byung-Gook Kim, and Hanku Cho

Samsung Electronics, Photomask Division, 16 Banwol-Dong, Hwasung, Kyunggi-Do, Korea, 445-701

(Received 30 June 2013; accepted 17 October 2013; published 5 November 2013)

As the feature size is reduced well below 100 nm, the line edge roughness (LER) becomes a critical issue since it does not scale with feature size. For minimizing the LER, it is essential to be able to accurately estimate it. A possible method for LER estimation is to rely on simulation. However, it requires time-consuming procedures, i.e., the Monte Carlo simulation for computing the exposure distribution within resist, and a resist-development simulation. In this study, an analytic method for estimating the LER, defined as the standard deviation of edge location, is developed to overcome the drawback of simulation method. This new approach first relates the stochastic exposure to the statistics of point spread functions, i.e., mean and variance, and eventually derives the variation of edge location considering critical paths in the resist development process. The analytic method achieves good accuracy compared to a simulation method and has a good potential to be employed in practice. © 2013 American Vacuum Society. [http://dx.doi.org/10.1116/1.4827816]

I. INTRODUCTION

Electron-beam (e-beam) lithography is widely employed in a variety of areas such as fabrication of photomasks, imprint lithography molds, experimental circuit patterns, etc., because of its ability to transfer ultrafine features onto the resist and eventually to the substrate material.1,2 Its main limitations are the low throughput due to the pixel-by-pixel or feature-by-feature writing and the proximity effect caused by electron scattering. The importance of developing effective and efficient schemes for correcting the proximity effect has been well recognized for a long time, and various methods were proposed and implemented by many researchers.3–7 As the feature size decreases well below microns into nanoscale (100 nm or less),8 the line edge roughness (LER) becomes an increasingly important factor that cannot be ignored since it does not scale with feature size. It can significantly affect the minimum feature size and the maximum feature density realizable in practice, and also the functionality of a device. Therefore, it is unavoidable to minimize the LER in order to maximize the feature density and enhance the yield of fabricated devices.

One important step required in developing an effective method to minimize the LER is to understand the characteristics of LER. That is, it is necessary to be able to estimate the LER accurately. A possible approach is to rely on simulation.9 Given a circuit pattern, the stochastic exposure (energy deposited) distribution is either obtained directly through the Monte Carlo simulation or computed through convolution with stochastic point spread functions (PSFs) generated via the Monte Carlo simulation. Then, the remaining resist profile (from which the LER is quantified) is derived through resist development simulation. While the simulation approach is flexible, the main drawback is that it is computationally intensive. An analytic approach to LER estimation is an alternative, which can avoid the time-consuming procedures of simulation. Some theoretical models were formulated to estimate the effects of the LER on the device behaviors,10,11 but derivation of the LER was not addressed. A comprehensive stochastic model of LER with approximate expressions of the standard deviation of the final deprotection level of polymer molecules in the resist was derived.12 An infinite contrast of development process was assumed to obtain the LER directly from the blocked polymer latent image. In another study,13 the resist development process was considered with the main emphasis on the stochastic nature of photoresist dissolution.

In this paper, a new analytic approach to analytically deriving the LER caused by the stochastic variation of exposure in the resist in e-beam lithography is described with a realistic 3-D model. The resist development process is taken into account through the conversion from exposure to developing rate. The specific objective is to derive an accurate analytic expression of LER from the statistical parameters of stochastic PSFs, i.e., the mean and variance at each location of PSF. The mean and variance of exposure at each point are computed from the mean and variance of PSFs. With exposure converted into developing rate, critical paths in the resist development process are considered to derive the variation of edge location, i.e., LER. In this study, the LER is quantified as the standard deviation of edge location. It should be pointed out that though this study focuses on the stochastic variation of exposure, the analytic method described in this paper is still applicable even when other factors contributing to the LER are taken into account.

The rest of the paper is organized as follows. The analytic model, the definition of terms and the assumptions are described in Sec. II. Derivation of LER from the statistical parameters of PSF is described in Sec. III. Results are presented in Sec. IV. A summary is provided in Sec. V.

---

Electronic mail: leesooy@eng.auburn.edu
II. ANALYTIC MODEL

A typical substrate system, which consists of a substrate and a certain type of resist on top of the substrate, is employed in this study as illustrated in Fig. 1, where the resist depth is along the z-dimension. The feature is W pixels wide in the x dimension, L pixels long in the y dimension, and its right edge is along the y axis, as shown in Fig. 2. Let \( \text{PSF}_{wl}(x, y, z) \) denote the PSF, which depicts the exposure distribution when point \((w, l, 0)\) is exposed, where \(-W \leq w \leq -1\) and \(1 \leq l \leq L\). Note that the peak (center) of \( \text{PSF}_{wl}(x, y, z) \) is at \((w, l, 0)\). Let \(e(x, y, z)\) and \(R(x, y, z)\) denote the exposure and resist developing rate at point \((x, y, z)\), respectively. \(R(x, y, z)\) is obtained from \(e(x, y, z)\) according to the conversion formula derived through experiment. Both \(R(x, y, z)\) and \(e(x, y, z)\) are stochastic.

The resist development process is modeled by “development paths” where a development path is a path along which resist is developed. Each development path starts from the resist surface toward the boundary of resist profile, consisting of vertical path segment (which represents the vertical development of resist) and lateral path segment (which represents the lateral development of resist). Though the resist development at each point is isotropic, considering all possible development paths starting from all points on the resist surface (and determining the boundary of resist profile by the set of square paths reaching the farthest points) tends to compensate for the potential discrepancy in the final resist profile. This has been demonstrated through simulation, i.e., the resist profiles obtained by the path-based method are very close to those by the cell-removal and fast marching methods. Given a resist layer on a cross-section perpendicular to the length dimension of a line feature, there are multiple paths on the cross-section, which start from all points with varying \(x\) on the intersection of the resist surface and the cross-section, and turn onto the resist layer. These paths move toward the edge “competing with each other.” All these paths are analyzed and the boundary (edge) of a feature (on that layer) is determined by the path reaching the farthest point, referred to as critical path. The length of the critical path is denoted by \(s\) and the developing rate along a development path is denoted by \(R(s')\), where \(0 \leq s' \leq s\) as shown in Fig. 3.

In the e-beam lithographic process, each point is exposed independent of other points. Hence, \(\text{PSF}_{wl}(x, y, z)\) and \(\text{PSF}_{wl'}(x, y, z)\) are considered to be independent of each other when \(w \neq w'\) or \(l \neq l'\). It is assumed that \(\text{PSF}_{wl}(x, y, z)\)
has a Gaussian distribution at each point with the mean \( m_{PSF}(x,y,z) \) and variance \( \sigma^2_{PSF}(x,y,z) \). Furthermore, \( PSF_{wl}(x,y,z) \) and \( PSF_{wR}(x,y,z) \) are assumed to have the same mean and variance at each point \((x,y,z)\), i.e., \( m_{PSF}(x,y,z) \) and \( \sigma^2_{PSF}(x,y,z) \). The mean and standard deviation of the PSF for one of the exposure conditions considered are plotted in Fig. 4.

Below is a summary of the notations:

1. \( PSF_{wl}(x,y,z) \): point spread function which depicts the energy distribution when point \((w,l,0)\) is exposed.
2. \( m_{PSF}(x,y,z) \): mean of the set of PSFs at point \((x,y,z)\).
3. \( \sigma_{PSF}(x,y,z) \): standard deviation of PSFs at point \((x,y,z)\).
4. \( e(x,y,z) \): exposure at the point \((x,y,z)\).
5. \( R(x,y,z) \): resist developing rate at the point \((x,y,z)\).
6. \( R(s') \): resist developing rate on the \( s' \) on the development path.
7. \( T \): developing time.
8. \( f(\cdot) \): probability density function.
9. \( f_{e(\cdot)} \): probability density function of exposure.
10. \( f_{R(\cdot)} \): probability density function of resist developing rate.

### III. ESTIMATION OF LER FROM PSF

Derivation of the analytic expression of LER is carried out in three steps: (1) derivation of the exposure fluctuation from the stochastic PSFs, (2) conversion of the exposure into developing rate, and (3) derivation of the LER from the developing rate distribution.

#### A. From PSF to exposure fluctuation

The exposure at a point \((x,y,z)\) in the resist can be expressed as

\[
e(x,y,z) = \sum_{w=1}^{W} \sum_{l=1}^{L} PSF_{wl}(x-w, y-l, z),
\]

where \( w \) and \( l \) vary from \(-W\) to \(-1\) and from \(1\) to \(L\), respectively (see Fig. 2). The conciseness of expression, a vector notation of \( r_{wl} \) is adopted, which starts from point \((w,l,0)\) to point \((x,y,z)\), i.e., \( r_{wl} = \{x-w, y-l, z\} \). Then, \( e(x,y,z) = \sum_{w=1}^{W} \sum_{l=1}^{L} PSF_{wl}(r_{wl}) \). Since the relationship between PSF and exposure is linear, the mean of exposure can be expressed as

\[
m_e(x,y,z) = E\left\{ \sum_{w=1}^{W} \sum_{l=1}^{L} PSF_{wl}(r_{wl}) \right\}
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} E\{PSF_{wl}(r_{wl})\}
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} m_{PSF}(r_{wl}).
\]  

(1)

Also, the variance of exposure is

\[
\sigma^2_e(x,y,z) = \left[ E\{e(x,y,z)\} \right]^2 - \left\{ E\{e(x,y,z)\} \right\}^2
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} \sum_{w'=1}^{W} \sum_{l'=1}^{L} PSF_{wl}(r_{wl}) PSF_{wl}(r_{w'l'})
\]

\[
- \left[ \sum_{w=1}^{W} \sum_{l=1}^{L} m_{PSF}(r_{wl}) \right]^2
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} E\{PSF_{wl}(r_{wl})\}
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} m_{PSF}(r_{wl})
\]

\[
= \sum_{w=1}^{W} \sum_{l=1}^{L} \sigma^2_{PSF}(r_{wl}).
\]  

(2)

The sum of Gaussian random variables is also a Gaussian random variable. The exposure is a linear combination of PSFs which are assumed to be of Gaussian and therefore has a Gaussian distribution. Hence, it is sufficient to specify the mean and variance of exposure [Eqs. (1) and (2)] in defining the probability density function of exposure, \( f_e(\cdot) \).

#### B. From exposure fluctuation to rate fluctuation

The developing rate \( R(x,y,z) \) at each point is calculated from its exposure \( e(x,y,z) \) through a nonlinear mapping of \((e\to R)\) conversion formula, which is experimentally determined. A single line is exposed with a spatially uniform dose, and after resist development, the depth in the cross-section of remaining resist profile is measured at the center of line. This process is repeated with different dose levels. Note that the resist is developed only vertically at the center of line when the dose is spatially uniform. The resist profile can also be obtained through simulation by computing the exposure distribution in the resist and converting exposure into developing rate. The relationship between exposure and developing rate (i.e., conversion formula), which minimizes the difference between the measured and simulated depths, is obtained. The following conversion formula \( C(\cdot) \) has been

![Fig. 4.](image-url)  

Fig. 4. (a) Plot of the mean and (b) plot of the variance of the PSFs. Results are generated by Monte Carlo simulation repeating 200 times on a substrate system composed of 300 PMMA on Si. The beam energy is 50 keV and the dose is 10000 electrons per shot.


\[ R(x, y, z) = C(e(x, y, z)) = 20,000 \exp \left[ - \left( \frac{e(x, y, z) - 6.0 \times 10^{10}}{1.8 \times 10^{10}} \right)^2 \right] - 0.29, \]

\[ \text{(3)} \]

where \( R(x, y, z) \) is in nm/min and \( e(x, y, z) \) in eV/\( \mu \)m².

Since the conversion formula is a monotonic function, from the probability density function of exposure \( f_e \), the probability density function of developing rate \( f_R \) can be derived as

\[
f_R(R(x, y, z)) = f_e(C^{-1}(R(x, y, z))) \frac{dC^{-1}(R(x, y, z))}{dR(x, y, z)}. \tag{4}
\]

From the distribution of resist developing rate \( f_R \), the mean \( (m_s(x, y, z)) \) and variance \( (\sigma_s^2(x, y, z)) \) of \( \frac{1}{R} \) at point \( (x, y, z) \) can be derived as

\[
m_s(x, y, z) = \int_{0}^{\infty} f_R(R(x, y, z)) \frac{dR(x, y, z)}{R(x, y, z)},
\]

\[
\sigma_s^2(x, y, z) = \int_{0}^{\infty} f_R(R(x, y, z)) \left( \frac{dR(x, y, z)}{R^2(x, y, z)} \right) dR(x, y, z) - m_s^2(x, y, z).
\]

\[ \text{C. From rate fluctuation to LER} \]

The LER can be considered as variation of the edge location on a layer of the remaining resist profile. More specifically in our model (see Sec. II), it is the variation in the length of the longest development path (critical path), which determines the edges (boundaries) of a feature, given a developing time \( T \). Note that the edge location, equivalently the length of the critical path, is denoted by \( s \). For a sufficiently long line, \( \sigma(s|T) \), the standard deviation of \( s \) given \( T \), may be used to quantify the LER which is the variation of edge location along the length dimension of the line, i.e.,

\[ \text{LER} = \sigma(s|T). \]

In order to compute \( \sigma(s|T) \), it is necessary to derive the probability density function \( f(s|T) \), which describes the variation of edge location given a developing time. According to the Bayes’ theorem, it can be expressed as

\[ f(s|T) = \frac{f(T|s)f(s)}{f(T)}. \tag{5} \]

The main task is to derive \( f(T|s) \), \( f(T) \) and \( f(s) \). Note that \( f(T) \) describes the possible unintended variation of developing time, \( f(s) \) depicts the variation of edge location considering all possible developing time \( T \) defined by \( f(T) \), and \( f(T|s) \) describes the variation of developing time given an edge location.

\[ \text{1. } f(T|s) \]

Given the length of the critical path \( (s) \), the developing time \( (T) \) can be expressed as a line integral of \( \frac{1}{R} \) along the critical path as illustrated in Fig. 3, i.e.,

\[ T = \int_{0}^{l_s} \frac{ds'}{R(x, y, z)} = \int_{0}^{l_s} \frac{ds'}{R(s')}. \tag{6} \]

Assume a Gaussian distribution of \( \frac{1}{R} \) with the mean \( (m_s(x, y, z)) \) and variance \( (\sigma_s^2(x, y, z)) \) and its independence between \((x, y, z)\) and \((x', y', z')\). Therefore, the mean \( (m(T|s)) \) and variance \( (\sigma^2(T|s)) \) of the developing time given the length of the critical path can be expressed as a line integral of the mean \( (m_s(x, y, z)) \) and variance \( (\sigma_s^2(x, y, z)) \) of \( \frac{1}{R} \), respectively, i.e.,

\[ m(T|s) = \int_{0}^{l_s} m_s(x, y, z) ds' = \int_{0}^{l_s} m_s(s') ds', \tag{7} \]

\[ \sigma^2(T|s) = \int_{0}^{l_s} \sigma^2_s(x, y, z) ds' = \int_{0}^{l_s} \sigma^2_s(s') ds', \tag{8} \]

where \( m_s(s') \) and \( \sigma^2_s(s') \) denote the mean and variance of \( \frac{1}{R} \) along the path.

Given the length of the critical path \( (s) \), the developing time \( (T) \) is a linear combination [Eq. (6)] of \( \frac{1}{R} \) which is assumed to have a Gaussian distribution, and therefore the distribution of the developing time given the length of the critical path is a Gaussian distribution. Hence, it is sufficient to specify the mean \( (m(T|s)) \) and variance \( (\sigma^2(T|s)) \) of developing time \( (T) \) given the length of the critical path \( (s) \) in defining the \( f(T|s) \).

The mean of exposure along the path is approximately linear to edge location when it is close to the target width \((X = 0)\) as illustrated in Fig. 5(a). Since the conversion formula [Eq. (3)] is an exponential function, \( m_r(s') \) should also be an exponential function of \( s' \) as shown in Fig. 5(b). Therefore, \( m(T|s) \), which is a line integral of \( m_r(s') \) [Eq. (7)], can be approximated as an exponential function. A similar relationship can be applied to \( \sigma^2(T|s) \). Hence, exponential functions are used to approximate the mean \( (m(T|s)) \) and variance \( (\sigma^2(T|s)) \) of developing time given the length of the critical path, i.e.,

\[ m(T|s) \approx a_1 \exp(b_1 s), \]

\[ \sigma^2(T|s) \approx a_2 \exp(b_2 s), \]

where \( a_1, b_1, a_2, \text{ and } b_2 \) are constants dependent on the circuit pattern and PSF. They are derived by curve-fitting onto the \( m(T|s) \) and \( \sigma^2(T|s) \) which are generated from Eqs. (7) and (8), as illustrated in Figs. 5(c) and 5(d).

\[ \text{2. } f(T) \text{ and } f(s) \]

The probability density function of developing time \( f(T) \) can be assumed or derived through experiment. In this study, the developing time \( T \) is assumed to have a Gaussian distribution with the mean \( \mu_T \) and variance \( \sigma_T^2 \). The probability density function of edge location \( f(s) \) is derived from \( f(T) \) and \( m(T|s) \),
where \( m(T|s) \) is used as the mapping from \( T \) to \( s \). Since \( m(T|s) \) is a monotonic function, \( f(s) \) can be derived as

\[
f(s) = f(T) \frac{d(m(T|s))}{ds},
\]

as illustrated in Fig. 6.

3. \( f(s|T) \), \( m(s|T) \), and \( \sigma^2(s|T) \)

From Eq. (5), \( f(s|T) \) is obtained, and the mean and variance of edge location given a developing time can be computed as

\[
m(s|T) = \int_0^\infty s f(s|T) ds,
\]

\[
\sigma^2(s|T) = \int_0^\infty s^2 f(s|T) ds - E^2\{s|T\}.
\]

Let \( A_T \) and \( B_T \) denote the following expressions:

\[
A_T = \frac{a_1 b_1}{\sqrt{2\pi} \sigma_T} \exp \left( \frac{T^2 - 2\mu_T}{2\sigma_T^2} \right) - \frac{a_1}{2\sigma_T^2} \frac{\mu_T a_1}{\sigma_T^2},
\]

\[
B_T = \frac{a_1 b_1}{\sqrt{2\pi} \sigma_T^2} \exp \left( \frac{T^2 - 2\mu_T}{2\sigma_T^2} \right) \frac{\mu_T a_1}{\sigma_T^2} \frac{\sigma_T^2}{\sigma_T^2} \left( \frac{b_1 - b_2}{2} \right).
\]

Then, \( E\{s|T\} \) and \( \sigma^2\{s|T\} \) can be derived as

\[
m(s|T) = \frac{A_T}{B_T},
\]

\[
\sigma^2(s|T) = \frac{2A_T A_T^2}{B_T^2}.
\]

It needs to be pointed out that the interaction between adjacent development paths is not taken into account in the above derivation. The interaction tends to decrease the variation (of path length) among the paths and reduces the actual LER. Therefore, an adjustment needs to be made to \( \sigma^2(s|T) \) in Eq. (9) in order to obtain an accurate estimate of LER.

D. Adjustment

The difference in developing time between adjacent development paths, given an edge location, may be considered as a way to quantify the level of interaction between them. In this study, the correlation coefficient of developing time (\( \rho_T \)) between adjacent development paths is employed in adjusting \( \sigma(s|T) \) in Eq. (9). Noting that developing time can be computed by integrating \( \frac{1}{\sigma} \) along a developing path, \( \rho_T \) is computed as the integral of the covariance of \( \frac{1}{\sigma} \) along the development path, normalized by the integral of the variance of \( \frac{1}{\sigma} \).

When \( \rho_T \) is closer to \(-1\) as illustrated in Fig. 7(a), the developing time for the same edge location is very different between adjacent development paths. This leads to a high level of interaction between the adjacent paths, and therefore, a larger adjustment (reduction) is to be made. On the other hand, when \( \rho_T \) is closer to 1 as illustrated in Fig. 7(c), the developing time for the same edge location is almost the same between adjacent development paths, i.e., a low level of interaction, and thus the adjustment (reduction) needs to be small. A reasonable multiplicative adjustment factor may be formulated as \( 1 - \alpha \frac{1 - \rho_T}{2} \), where \( \alpha \) determines the maximum adjustment and \( 0 < \alpha < 1 \). Note that the adjustment factor is 1 (i.e., no adjustment) when \( \rho_T \) is 1 and \( 1 - \alpha \) when \( \rho_T \) is -1

\[
LER = \sigma(s|T) \left( 1 - \alpha \frac{1 - \rho_T}{2} \right).
\]

The correlation coefficient \( \rho_T \) is computed between adjacent development paths and therefore is affected by the distance between the adjacent paths, denoted by \( \Delta y \). For example, a smaller \( \Delta y \) makes \( \rho_T \) larger (closer to 1). Hence,
The final estimate of LER is given as

\[
LER = \sigma(s|T) \left( 1 - \min \left( 1, \frac{\sigma(s|T)}{\Delta y} \right) \right) \left( 1 - \rho_T \right).
\]

(11)

IV. RESULTS AND DISCUSSION

The proposed analytic method for derivation of LER has been compared with the simulation method, which employs the Monte Carlo simulation for generating the stochastic exposure distribution and a path-based method for resist development simulation. In the model employed for this comparison, the substrate system for this comparison is composed of 300 nm PMMA on Si, and the beam energy is 50 keV. In this study, the LER is estimated with edge location varied by changing the developing time.

In Fig. 8, the results for a single feature of size 80 nm × 256 nm (W × L in Fig. 2) are provided. The dose is 1000 electrons per shot. It can be seen that the LER estimated by the analytic method is well matched with that by the simulation method on all three layers of resist. In Fig. 9, the results for a lower dose of 40 electrons per shot are given with a single feature of size 60 nm × 256 nm (W × L). A similar observation, i.e., a close match between the analytic and simulation results, can be made. In Fig. 10, the LER by the analytic method for different feature sizes (widths) is shown. The width of feature (W) is varied from 30 nm to 90 nm. This result verifies that the LER is independent of feature size.
A multiline pattern is also considered in order to examine the general applicability of the analytic method. The pattern consists of 41 lines where each line is of size \(25 \text{ nm}/\sqrt{2}\times4000 \text{ nm}\) and the space between lines is 25 nm. The dose is 1000 electrons per shot. The LER is estimated in the corner and center regions of the pattern and the results are provided in Fig. 11. A good match between the analytic and simulation results at both regions can be observed. It is also seen that the LER in the center region is smaller. This is mainly because more exposed points contribute to the center region, compared to the corner region, leading to a lower stochastic fluctuation of exposure.

V. SUMMARY

As the size of feature continues to shrink, the LER which does not scale with feature size has become a critical issue to be addressed in pattern transfer. For an effort to reduce the LER, it is essential to be able to estimate it accurately. The main drawback of a simulation approach to estimating the LER is the intensive computation required. In this study, an analytic method for estimating the LER has been developed to avoid the time-consuming simulation procedure. The method derives the mathematical expression of LER given the statistical parameters of stochastic PSF and a pattern specification based on the concept of development path. It has been shown that the results obtained by this analytic method are closely matched with the simulation results for both a single feature and a pattern of multiple lines. Though the LER caused by the stochastic fluctuation of exposure only is considered in the model adopted in this study, the method must be still applicable to the cases where other sources of LER are taken into account. Therefore, it can be said that the analytic method has a good potential to be used in practice.

ACKNOWLEDGMENT

This work was supported by a research grant from Samsung Electronics, Co., Ltd.