Understanding focus effects in submicrometer optical lithography

Chris A. Mack
National Security Agency
9800 Savage Road
Fort Meade, Maryland 20755-6000

Abstract. A new approach for characterizing resolution and depth of focus (DOF) in optical microlithography is introduced. By examination of the interaction of the aerial image with the photoresist process, a metric of image quality is defined. The variation of this metric with feature size and defocus can be used to measure the resolution and DOF. The effects of various imaging parameters on DOF can then be determined. To further study focus effects in submicrometer imaging, the lithography simulation program PROLITH (the positive resist optical lithography model) is modified to account for defocus within the photoresist film.

Subject terms: microlithography; depth of focus; resolution; lithography simulation.

Optical Engineering 27(12), 1093-1100 (December 1988).

CONTENTS
1. Introduction
2. Aerial image
3. Focus and the aerial image
4. Relationship to modulation transfer function (MTF)
5. Lithography simulation
6. Conclusions
7. References

1. INTRODUCTION
In the age of submicrometer optical lithography, focus has become a critical process parameter. Each decrease in minimum feature size is accompanied by a corresponding decrease in depth of focus (DOF). However, sources of focus errors, such as wafer warpage, topography, and the thickness of the photoresist are not being reduced in proportion to the DOF. Thus, the effects of focus on the practical resolution capabilities of a lithographic tool are becoming increasingly important.

In describing the resolution and DOF of a lithographic system, it is common to apply the Rayleigh criteria. The Rayleigh criterion for the minimum resolvable feature size is

\[ \text{resolution} = k_1 \frac{\lambda}{\text{NA}} , \]  

where \( \lambda \) is the exposure wavelength, NA is the numerical aperture of the objective lens, and \( k_1 \) is referred to as a "process-dependent constant." Typically, \( k_1 \) is in the range of 0.4 to 0.9. Similarly, the Rayleigh DOF is given by

\[ \text{DOF} = k_2 \frac{\lambda}{\text{NA}^2} , \]  

where \( k_2 \) is another process-dependent constant with values typically in the range of 0.5 to 1.0.

In the submicrometer regime, the simple Rayleigh criteria are not adequate for describing the resolution and DOF of a
microlithographic process. In fact, the common characterization of \( k_1 \) and \( k_2 \) as constants leads to many misinterpretations of these equations. A more appropriate way to view the Rayleigh criteria is as scaling equations. Resolution scales as \( \lambda / NA \), so \( k_1 \) is, in fact, the scaled resolution. Similarly, the DOF scales as \( \lambda / NA^2 \), so \( k_2 \) is the scaled DOF. The scaled quantities \( k_1 \) and \( k_2 \) are not constants and vary greatly as a function of many lithographic parameters. The Rayleigh equations give no information about the values of \( k_1 \) and \( k_2 \), their interdependence, or their dependence on other parameters.

In this paper, alternative definitions of resolution and DOF are given based on an understanding of the interactions of the aerial image with the photoresist process. Earlier studies have characterized the interaction of the aerial image with the photoresist process. This interaction points to various aspects of the aerial image that are important from a lithographic point of view. Defining a physically significant metric of aerial image quality allows one to characterize the effects of feature size and focus and leads to new definitions for resolution and DOF. The effects of numerical aperture, wavelength, feature size, and feature type also can be characterized using this technique, thereby permitting objective comparisons of different lithographic tools.

Current lithography simulation programs predict an aerial image based on the parameters of the projection tool. This image is then used to "expose" the photoresist. A tacit assumption of these models is that the photoresist thickness is less than the DOF and that, therefore, the aerial image does not change through the thickness of the resist. In the case of submicrometer imaging, however, this assumption is not adequate. Thus, Sec. 5 describes an enhancement to the lithography simulation program PROLITH (the positive resist optical lithography model), which includes the effects of defocus through the resist. Using this model, the effects of exposure and focus on linewidth control and sidewall angle can be determined and asymmetric focus-exposure diagrams can be constructed. Also, the optimum position of the focal plane can be determined.

2. AERIAL IMAGE

To simplify the analysis of a lithographic process, it is highly desirable to separate the effects of the lithographic tool from those of the photoresist process. This can be done with reasonable accuracy only if the interaction of the tool (i.e., the aerial image) with the photoresist is known. Consider an aerial image of relative intensity \( I(x) \) where \( x \) is the horizontal position (i.e., in the plane of the wafer and mask) and is zero in the center of a symmetric mask feature. The aerial image exposes the photoresist to produce some chemical distribution \( m(x) \) within the resist. This distribution is called the latent image. Many important properties of the lithographic process, such as exposure and development latitude, are a function of the gradient of the latent image \( \partial m / \partial x \). Larger gradients result in improved process latitude. It has been shown that the latent image gradient is related to the aerial image by

\[
\frac{\partial m}{\partial x} \propto \frac{\partial \ln I}{\partial x} .
\]  

(3)

The development properties of the photoresist translate the latent image gradient into a development gradient, which then allows for the generation of a photoresist image. Optimum photoresist image quality is obtained with a large development rate gradient. A lumped parameter called the photoresist contrast, \( \gamma \), can be defined that relates the aerial image and the development rate \( r \) (Ref. 1):

\[
\frac{\partial \ln r}{\partial x} = \gamma \frac{\partial \ln I}{\partial x} .
\]  

(4)

The development rate gradient is maximized by higher resist contrast and by a larger slope of the log-aerial image (the log-slope).

A second important lithographic parameter is the sidewall angle of the resist profile. There are two ways in which the aerial image affects sidewall angle. First, the latent image has a sidewall slope due to absorption. This slope is again directly proportional to the log-slope of the image. Second, the very nature of the development rate process gives rise to a sloped sidewall since the top of the resist is under attack by the developer for a longer period of time than the bottom. Neglecting absorption, the approximate slope due to development is given by

\[
\text{resist slope} \approx \frac{r(0)}{r(x)} ,
\]  

(5)

where \( r(0) \) is the development rate in the center of a space and \( r(x) \) is the development rate at the line edge (i.e., at the edge of the photoresist profile). This ratio of development rates should be maximized in order to maximize the resist slope. Further, this ratio is a function of the aerial image. A simple approximation gives

\[
\frac{r(0)}{r(x)} = f \left( \frac{I(0)}{I(x)} \right) = \left[ \frac{I(0)}{I(x)} \right]^{\gamma} .
\]  

(6)

The ratio of center to edge intensities also affects mask linearity, i.e., the ability to reproduce in photoresist the mask dimension for a variety of different feature sizes.

The above discussion gives two ways in which the aerial image and photoresist process interact. First, the slope of the log-image affects process latitude and sidewall angle. Second, the ratio \( I(0)/I(x) \) also affects sidewall angle, as well as mask linearity. Thus, there are two logical metrics by which to judge the quality of the aerial image:

\[
\frac{\partial \ln I}{\partial x} ,
\]  

(7)

\[
\frac{I(\text{center})}{I(\text{edge})} .
\]  

(8)

Which metric to choose depends on what factors limit the quality of the final photoresist image. If process latitude is considered of prime importance, then the log-image slope is the preferred metric of aerial image quality. This metric has been discussed in relation to focus effects in the excellent work of Levinson and Arnold.

The most commonly used metric of image quality is the image contrast, which is defined for a periodic pattern as

\[
\text{contrast} = \frac{I(\text{center of space}) - I(\text{center of line})}{I(\text{center of space}) + I(\text{center of line})} .
\]  

(9)
The most obvious limitation of the image contrast is that it applies only to periodic patterns (and is useful only for equal lines and spaces). Also, this definition is not directly related to the lithographic parameters of interest such as latitude and sidewall angle. For these reasons, image contrast is a poor metric by which to judge the effects of defocus.

3. FOCUS AND THE AERIAL IMAGE

Figure 1 shows the well-known effect of defocus on the aerial image. Both the edge slope of the image and the center intensity decrease with defocus, and the intensity at the mask edge remains nearly constant. To examine the behavior of the log-slope, the aerial images of Fig. 1 were used to calculate the log-slope, which is plotted in Fig. 2. Clearly, the log-slope varies considerably with horizontal position x. To compare aerial images using the log-slope, one must pick an x value to use. An obvious choice is the mask edge. Thus, all subsequent reference to the slope of the log-aerial image will be at the mask edge. Now the effect of defocus on the aerial image can be expressed by plotting log-slope as a function of defocus, as shown in Fig. 3. Superimposed on this curve is a graph of the ratio of center to edge intensities. Note that these two metrics of image quality give nearly identical variation with defocus to within a scale factor. The agreement between these two metrics is very good for images near the resolution limit but worsens for larger features. Since this paper deals with high resolution imaging, the use of the log-slope is sufficient to characterize the degradation of the aerial image with defocus.

Some useful information can be obtained from a plot of log-slope versus defocus. As was previously discussed, both process latitude and sidewall slope vary directly with the log-slope of the image. Thus, minimum acceptable process latitude and sidewall slope specifications translate directly into a minimum acceptable value of the log-slope. The log-slope versus defocus curve then can be used to give a maximum defocus to keep the process within this specification. If, for example, the minimum acceptable log-slope of a given process was determined to be 3.5 µm⁻¹, the maximum defocus of 0.6 µm lines and spaces on a 0.35 NA g-line printer would be, from Fig. 3, about ±0.5 µm. This gives a practical definition of the DOF that separates the effects of the aerial image and the photoresist process. The printer determines the shape of the log-slope defocus curve, and the process determines the range of operation (i.e., the minimum log-slope value). If the minimum log-slope needed was 4.5 µm⁻¹, one would conclude from Fig. 3 that this printer could not adequately resolve 0.6 µm lines and spaces. Thus, resolution can also be determined from a log-slope defocus curve.
To define resolution, consider Fig. 4, which shows the effect of feature size on the log-slope defocus curve. If, for example, a particular photoresist process requires a log-slope of 3.5 \( \mu \)m\(^{-1}\), one can see that the 0.6 \( \mu \)m features will not be resolved, the 0.7 \( \mu \)m features will be resolved only when in perfect focus, the 0.8 \( \mu \)m features will have a DOF of ±1 \( \mu \)m, and the 0.9 \( \mu \)m features will have a DOF of ±2 \( \mu \)m. Obviously, the DOF is extremely sensitive to feature size, a fact that is not evident in the common Rayleigh definition. Since DOF is a strong function of feature size, it is logical that resolution is a function of DOF. Thus, in the situation shown in Fig. 4, if the minimum acceptable DOF is ±1 \( \mu \)m, the practical resolution is 0.8 \( \mu \)m lines and spaces. Resolution and DOF cannot be independently defined but rather are inter-dependent.

The log-slope defocus curve can now be used to explore the effects of various parameters on the resolution and DOF. The numerical aperture is one of the most important parameters defining lithographic performance, and yet it is the most misunderstood. The Rayleigh DOF seems to predict a dramatic decrease of DOF with increasing numerical aperture. Figure 5 shows the effect of numerical aperture on the log-slope defocus curve of 0.6 \( \mu \)m lines and spaces. The effect is to improve the aerial image log-slope for all values of defocus. Thus, the DOF increases with increasing numerical aperture for a given feature size near the resolution limit, over the range of numerical apertures shown.

There has been some discussion recently about the role of the numerical aperture in the DOF. In particular, many authors have used image simulators such as PROLITH or SAMPLE\(^7\) to look at the effects of very high numerical apertures on DOF. The conclusion drawn based on these studies is that DOF increases with numerical aperture to a point and then decreases for very high numerical apertures.\(^6,8\) However, these results should be viewed with some skepticism. The aerial image models used by SAMPLE and PROLITH are different, but both are based on the assumptions of scalar wave theory. The models are known to break down for high numerical apertures, where electromagnetic diffraction theory is required.\(^9\) Thus, the effects of very high numerical apertures on DOF cannot be well understood using the scalar models in SAMPLE or PROLITH.

The role of wavelength in DOF is also misunderstood. Although Eq. (2) seems to indicate worse DOF with shorter wavelength, Fig. 6 shows that DOF improves as wavelength decreases. Note that the effect of wavelength is different from that of the numerical aperture in that the curves do not converge at zero slope as they do for increasing numerical aperture. Figures 5 and 6 show clearly the danger of using the
Rayleigh criterion for comparing the DOF of different printers (i.e., different values of wavelength and numerical aperture). The log-slope defocus curve can be used to compare different printers objectively. For example, there has been much discussion about the advantages of lower wavelength versus higher numerical aperture. It is common to compare a g-line, 0.42 NA system with an i-line, 0.35 NA system. Both have the same value of λ/NA (almost) and thus, according to the Rayleigh criterion, the same resolution. In terms of the log-slope curve, the same value of λ/NA corresponds to the same value of the log-slope of the image with no defocus (Fig. 7). The practical resolution is defined as the smallest feature meeting a given log-slope specification over a given focus range. If a process requires a log-slope of 4 μm⁻¹ and a focus budget of ±1 μm, Fig. 7 shows that the i-line system will resolve a 0.6 μm feature, but the g-line system will not. Thus, the lower wavelength system has better resolution even though λ/NA is the same.

It is important to note that all of the aerial image calculations presented in this paper assume diffraction-limited lens performance, i.e., ideal lenses. Obviously, the ideal lens does not exist, and thus real lenses have log-slope versus defocus curves that are degraded to some extent from the ideal curves shown here. To a first approximation, the aberrations in an optical system can be thought of as a “built-in defocus,” where the degradation of the image is roughly equivalent to defocusing by a certain amount. An effective defocus of 0.5 μm is probably typical. This translates into a shift to the left of the log-slope curve. When comparing different lenses, as was done above, one must keep in mind that one lens may be farther from the ideal than the other.

Figure 8 shows the differences between an isolated line, an isolated space, and equal lines and spaces. The differences are quite remarkable. For very small features [Fig. 8(a)], packed lines and spaces have the worst resolution/DOF. Interestingly, for the 0.6 μm features the curves for isolated features cross, in this case at about 1 μm defocus. For defocus less than 1 μm the isolated line has better image quality than the space. However, for greater defocus the isolated line falls off quite rapidly, and the isolated space has the better quality image. This variation with feature type is strongly dependent on the feature size. For 0.8 μm features [Fig. 8(b)] there is little difference between the various feature types for the conditions given. For 1.0 μm features the situation is completely reversed, with equal lines and spaces having the best log-slope over the entire focal range.

As a final example, the coherence of the illumination affects image quality, as shown in Fig. 9. For the conditions given, a partial coherence of 0.5 is optimum. As expected, however, this result is very feature-size dependent.
4. RELATIONSHIP TO MTF

The log-slope of the aerial image has a strong relationship with the MTF of the optical system. This relationship can be elucidated for the simple case of incoherent illumination. For a periodic mask pattern, the aerial image takes the form of a Fourier series,

$$I(x) = a_0 + 2 \sum_{n=1}^{\infty} a_n \cos \left( \frac{2 \pi n x}{p} \right) \text{MTF} \left( \frac{n}{\nu_0} \right), \quad (10)$$

where

$$a_0 = \frac{w}{p}, \quad a_n = -\frac{\sin(n\pi w/p)}{n\pi}, \quad \nu_0 = \frac{2 \text{NA}}{\lambda},$$

which is the cut-off frequency, and p is the pitch, w is the linewidth, and n is the diffraction order. For the case of equal lines and spaces, Eq. (10) simplifies to

$$I(x) = \frac{1}{2} + 2 \sum_{n=1}^{\infty} \frac{\sin(n\pi/2)}{n\pi} \cos \left( \frac{\pi n x}{w} \right) \text{MTF} \left( \frac{n}{\nu_0} \right). \quad (11)$$

Note that, because of the sine term, only odd diffraction orders contribute to the aerial image. If the feature size is small enough, only the first diffraction order is used to generate the image. This occurs when

$$w < 0.75 \frac{\lambda}{\text{NA}}. \quad (12)$$

For these small features, the aerial image simplifies further to

$$I(x) = \frac{1}{2} + \frac{2}{\pi} \cos \left( \frac{\pi x}{w} \right) \text{MTF} \left( \frac{1}{\nu_0} \right). \quad (13)$$

One can see that the aerial image has become a sinusoid.

The above expression can now be used to compute the slope of the image at the mask edge:

$$\frac{\partial \ln I}{\partial x} = \frac{4}{w} \frac{1}{\nu_0} \text{MTF} \left( \frac{1}{\nu_0} \right). \quad (14)$$

Thus, the edge slope of the log-aerial image is directly related to the value of the MTF at the first diffraction order. The degradation of the MTF with defocus results directly in a degradation of the log-slope. For the case of partially coherent illumination, an MTF cannot be defined. However, the concept of MTF remains important for understanding the image formation process. The above analysis suggests that the log-slope can be used as a measure of MTF and thus as a measure of the imaging capabilities of an optical system.

Note that from Eq. (11) the intensity at the mask edge is always 0.5 for incoherent illumination. The center intensity can be obtained from Eq. (13) and the ratio of center to edge intensities can be calculated:

$$\frac{I(\text{center})}{I(\text{edge})} = 1 + \frac{4}{\pi} \frac{1}{\nu_0} \text{MTF} \left( \frac{1}{\nu_0} \right). \quad (15)$$

Thus, the ratio of center to edge intensities is related to the log-slope by a scale factor for small features, as was shown in Fig. 3.

5. LITHOGRAPHY SIMULATION

Primary parameter lithography models such as PROLITH\textsuperscript{3,4} can be used to study in detail the effects of focus on the lithographic process. Previously, these models have assumed that the aerial image is a constant throughout the thickness of the resist film. This is equivalent to saying that the resist thickness is less than the DOF of the process, so the image does not defocus as it propagates through the resist. In submicrometer imaging, however, this approximation is no longer valid. Thus, to accurately describe the effects of focus one must take into account the defocusing with the resist. The most rigorous solution to this problem is to calculate the image in two or three dimensions using, for example, the extended source method\textsuperscript{10} or electromagnetic diffraction theory.\textsuperscript{9} Such rigorous approaches are quite complicated, however, and a simple extension of current modeling techniques will now be proposed.

An integral part of current lithography models such as PROLITH is the assumption that the image is a plane wave traveling normal to the resist surface. In such a case the aerial image $I(x)$ and the standing wave intensity $I_s(z)$ may be calculated independently. The total intensity is

$$I(x,z) = I(x)I_s(z), \quad (16)$$

where $z$ is the depth into the resist and is zero at the top. Yeung's rigorous solutions\textsuperscript{9,10} do not make these assumptions, and the total intensity is not separable. To avoid the complications of Yeung's calculations, we will assume that the aerial image and standing wave intensity are still separable, but make the aerial image a function of depth into the photoresist:

\[I(x,z) = I(x)I_s(z), \quad (16)\]
Fig. 10. Focus latitude as a function of exposure as predicted by PROLITH. A 1 µm resist film with typical g-line parameters was assumed.

\[ I(x,z) = I_i(x,z)I_s(z) \]  

The image \( I_i(x,z) \) is calculated using the standard aerial image model by defocusing different amounts for different values of \( z \). The depth into the resist is not the defocus distance, however. Rather, the defocus distance \( \delta \) is given by Ref. 6 as

\[ \delta(z) = \delta_0 + \frac{z}{n} \]  

where \( \delta_0 \) is the defocus distance at the top of the resist and \( n \) is the index of refraction of the resist. Although the aerial image is an even function of defocus (a defocus distance of \( -\delta \) produces the same image as a defocus of \( \delta \)), it is not an even function of \( \delta_0 \). Since \( \delta_0 \) can be thought of as the vertical position of the wafer plane in a stepper, focus effects in submicrometer lithography are not symmetric about the optimum focal position. In fact, the position of the optimum focal plane is not obvious when the resist thickness is not negligible compared to the DOF.

PROLITH v1.4 includes the defocus model described above. Using this program, the effects of defocus and exposure on submicrometer features can be simulated. Figure 10 shows the common plot of linewidth versus focus for different values of exposure. The defocus distance \( \delta_0 \) is zero when the focal plane is at the top of the resist. If the wafer is moved upward 1 µm from this position, the value of \( \delta_0 \) is -1 µm (i.e., the focal plane is 1 µm below the resist surface). One can see that these curves are not symmetric. The optimum exposure appears to be about 120 to 125 mJ/cm² with mask bias or about 100 mJ/cm² without. The optimum focal position is about -0.3 to -0.5 µm. In other words, the focal plane should be about ½ to ½ of the way down into the resist.

Another way to represent the data of Fig. 10 is the focus-exposure process volume. In Fig. 11, the values of focus and exposure that result in a ±10% variation in linewidth from the nominal are plotted. The result is a process window, within which the linewidth specification is met. Other specifications, such as sidewall angle, can also be plotted,6 as shown in Fig. 11.

In addition to an asymmetric behavior of linewidth versus focus, the above defocus model also predicts that the shape of the resist profile behaves differently with defocus. Figure 12 shows the variation of the resist profile through focus. When the focal plane moves from below the resist surface to above, the profile changes from convex to concave. This variation of resist profile with focus and the asymmetric focus-exposure curves have been confirmed by the more rigorous model and by experiment.9

6. CONCLUSIONS

The Rayleigh criteria for resolution and DOF are not adequate in describing submicrometer optical lithography. In fact, it is quite easy to misinterpret the Rayleigh criteria and draw completely inaccurate conclusions. Thus, a new approach to characterizing resolution and DOF has been introduced. By examining the interaction of the lithographic tool (via the aerial image) with the photoresist process, a metric for judging aerial image quality has been established. By examining the effects of feature size and defocus on this metric, accurate and meaningful definitions of resolution and DOF can be made. This technique also leads to an understanding of the influence of various parameters on the resolution/DOF and the ability to compare the theoretical performance of different lithographic tools.

An enhanced version of PROLITH has been introduced to account for defocus within the photoresist layer. With this model, various focus effects can be characterized in great detail.

7. REFERENCES

Fig. 12. Variation of the resist profile shape with defocus (as predicted by PROLITH).

2. R. Hershel and C. A. Mack, “Lumped parameter model for optical lithography,” in 
3. C. A. Mack, “PROLITH: a comprehensive optical lithography model,” 
in Optical Microlithography IV, H. L. Stover, ed., Proc. SPIE 538, 
4. C. A. Mack, “Advanced topics in lithography modeling,” in Advances in 
submicron optical lithography: part 2,” in Optical Microlithography VI, 
7. W. G. Oldham, S. N. Nandgaonkar, A. R. Neureuther, and M. O’Toole, 
“A general simulator for VLSI lithography and etching processes: Part 
I—application to projection lithography,” IEEE Trans. Electron Devices 
8. A. Suzuki, S. Yabu, and M. Ookubo, “Intelligent optical system of a 
9. M. S. Yeung, “Modeling high numerical aperture optical lithography,” 
in Optical/Laser Microlithography, B. J. Lin, ed., Proc. SPIE 922, 
10. M. Yeung, “Modeling aerial images in two and three dimensions,” 