Optical Proximity Effects in Sub-micron Photomask CD Metrology

Nicholas Doe, Richard Eandi
Advanced Imaging Systems, Zygo Corporation
650 N. Mary Avenue
Sunnyvale, CA 94404

Abstract

Understanding how optical proximity effects (OPE) influence critical dimension (CD) measurements of photomasks and wafers in semiconductor manufacturing has been a subject of intense interest and investigation for many years. OPE, caused by the convolution of the intensity profiles of adjacent lines, introduces errors in the determination of the line edge position, and in turn the linewidth. This paper models several imaging systems using the Optical Transfer Function analysis method and discusses some results from an ongoing study to devise methods for calibrating CD mask metrology tools, and evaluates several different imaging objectives and line measurement algorithms as to their sensitivity to the influences of OPE in the measurement of binary masks.

Introduction

Optical proximity effect is a purely optical phenomenon. It is the result of an imperfect optical imaging system that causes adjacent line edges to spread spatially, such that the tail of the adjacent edge spread function, if sufficiently close, influences the shape and thereby the determination of the position of the line edge in question. For example, Figure 1 depicts and compares the intensity profile of an isolated one micron line with that of a one micron line with another edge in near proximity. Note the resulting shift in the linewidth as defined by the distance between edges at the 50% threshold level.

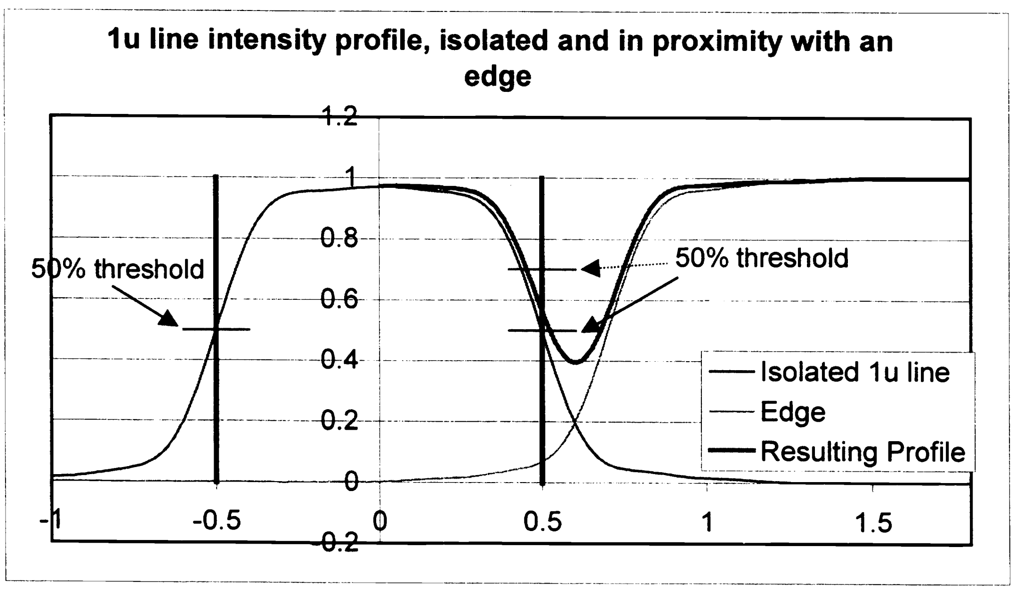

Figure 1. Influence of a nearby edge on the I-Line intensity profile of a 1µm line.

In *16th European Conference on Mask Technology for Integrated Circuits and Microcomponents*, Uwe F. W. Behringer, Editor, Proceedings of SPIE Vol. 3996 (2000) • 0277-786X/00/$15.00

The Optical Transfer Function (OTF)[1] analysis method is used to model our imaging system and to characterize and quantify the influence of OPE. The OTF is a combination of a Modulation Transfer Function, M(f) and a Phase Transfer Function, $\Phi(f)$, where f is the spatial frequency. That is

$$OTF = M(f) \cdot \exp[j\Phi(f)]$$

The Modulation Transfer Function (MTF) describes in quantitative terms the relationship of object and image and provides a framework for analyzing the performance of imaging systems. To illustrate, consider the imaging system shown in Figure 2.

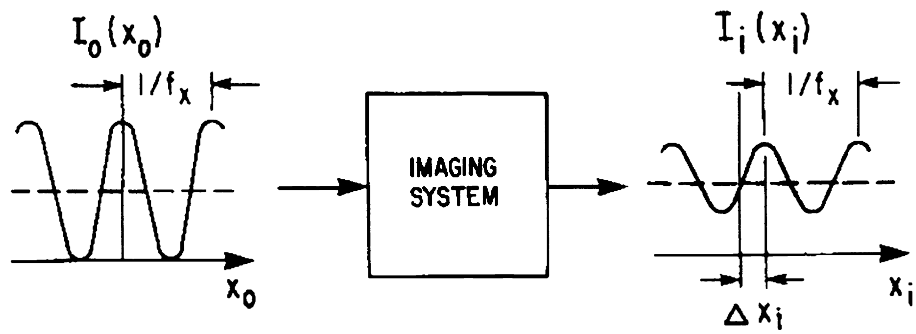

Figure 2. Imaging a sinusoidal object with spatial period 1/f

An object with a sinusoidal spatial variation of intensity, $I(x_0)$ along the x_0 axis, is imaged with unit magnification, producing a sinusoidal image intensity $I_i(x_i)$. The peak to peak distance along the sinusoid is its spatial period; the reciprocal of the period is the spatial frequency. Notice that the spatial frequency of the image and object are the same, whereas the contrast, or modulation, defined as

$$\textbf{Modulation} = (I_{max} - I_{min})/(I_{max} + I_{min})$$

is typically smaller in the image than the object, and that the image is displaced by Δx. If one images several such test objects with different spatial frequencies, one finds that the image modulation and position shift will vary as a function of spatial frequency. The modulation transfer function is defined as

M(f) = image modulation/ object modulation

The MTF expresses how the imaging process alters the contrast as a function of spatial frequency. The phase transfer function, or system phase response, $\Phi(f)$, expresses how the imaging process alters the phase shift of the image as a function of spatial frequency. The above demonstrates the method for a given spatial frequency f. It can be shown[2] that, except for very low spatial frequencies, any spatial intensity profile can be analyzed using the OTF method by representing the intensity profile as a Fourier series.

Model

To model the OPE, we consider an infinite array of lines and spaces as shown in figure 3, where the line width is a and the space is b (the pitch is a+b).

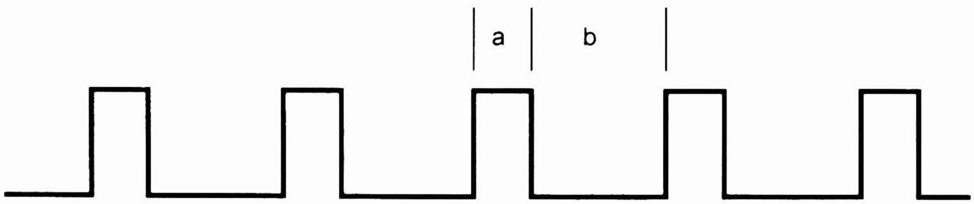

Figure 3. An infinite line/space array.

The model used in this analysis can be summarized as:

1. Record an image of an isolated line and generate an average intensity profile.
2. Calculate the complex Optical Transfer Function (OTF) for the microscope using the intensity profile.
3. Multiply the Fourier series representing the line array by the OTF to give a Fourier series for the image.
4. Evaluate the image Fourier series.

Optical Transfer Function

The intensity profile of an isolated 2μ line is recorded using the KMS-450i metrology tool. Figure 4 below shows the image taken using the 100x lens and i-line (365nm) illumination.

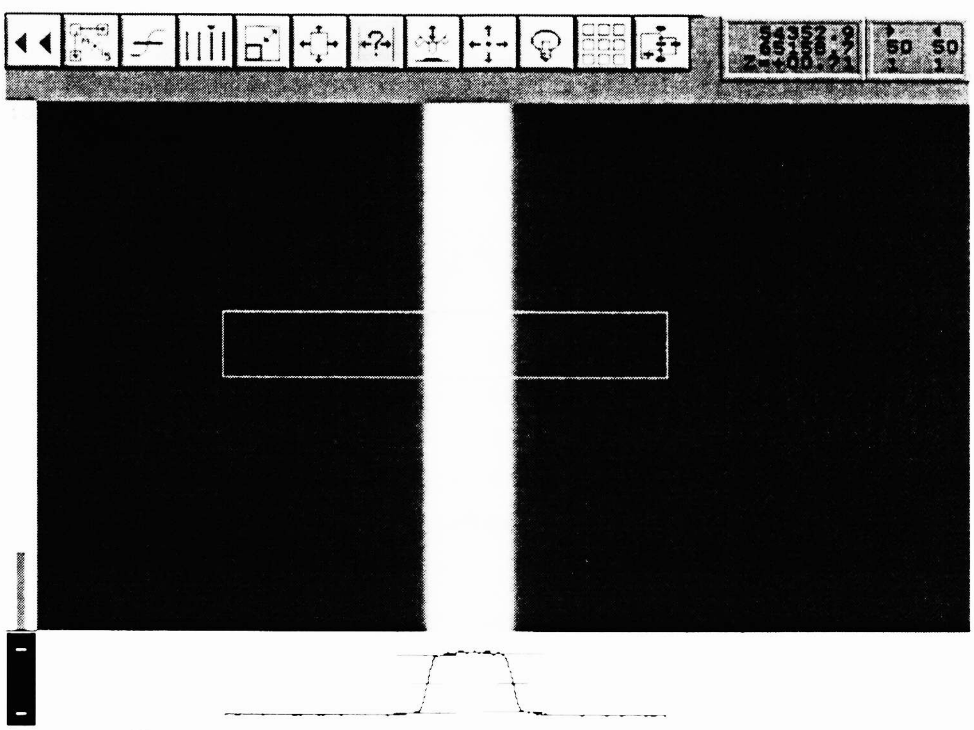

Figure 4. KMS-450 i-line image taken with 100x objective.

The highlighted portion of the image is recorded and averaged over the vertical direction to give a single intensity profile, figure 5.

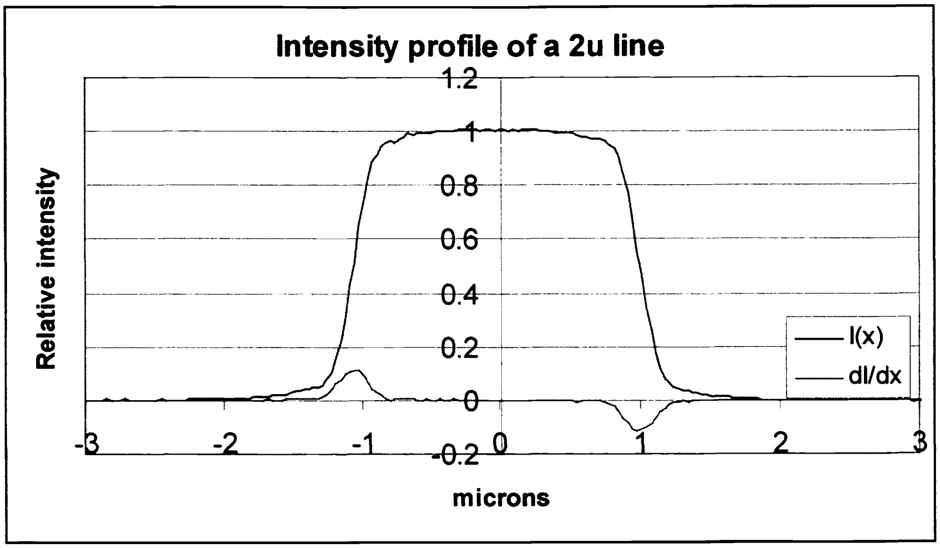

Figure 5. Averaged intensity profile and calculated derivative (dI/dx).

The OTF is the Fourier transform of the image of a line. In a linear system, the image of a line is equivalent to the derivative of the image of an edge.

The derivative: dI(x)/dx is evaluated numerically and is also shown in figure 5. The derivatives for the 2 edges are shifted to zero and the derivative for the falling edge is inverted. This is shown in the figure 6.

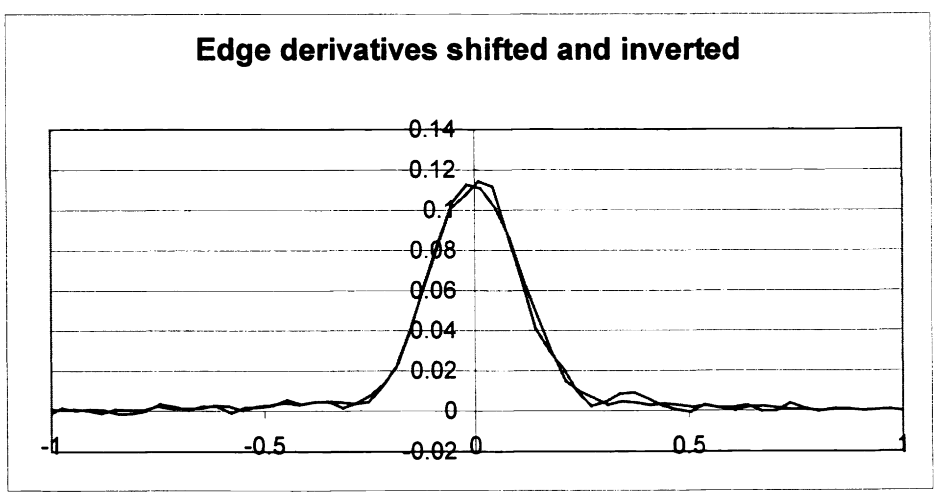

Figure 6. Edge derivatives from figure 5 shifted to zero and normalized.

The OTF is calculated as the Fourier Transform of the average of the edge derivatives as shown in figure 7.

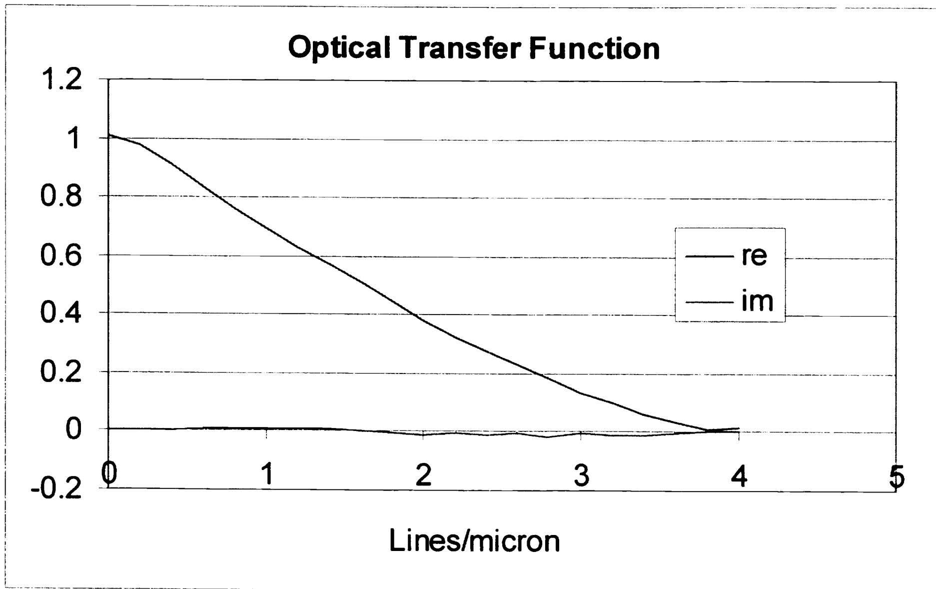

Figure 7. OTF of KMS-450i 100x, i-line transmission.

The OFT has the characteristic shape for a partially coherent microscope[3] and is almost entirely real. Had there been any asymmetry in the image, this would have manifest itself in the imaginary part of the OTF. From the OTF we would expect to be able to resolve features with a pitch of about 3 line pairs per micron (with 20% modulation) or about a pitch of 0.33μ. For CD metrology we would expect to need at least 50% modulation, so we would expect to be able to measure features of about 1.8 line/micron or 0.55μ pitch or features of about 0.3μ.

Evaluate the Fourier Series

The complex Fourier series for the infinite line/space array is[4]:

$$I(x)=\frac{ka}{2\pi}+\frac{1}{\pi}\sum_{n=-\infty}^{n=\infty}\frac{1}{n}\sin\left(\frac{nka}{2}\right)\exp(-jnkx) \qquad \text{where} \quad k=\frac{2\pi}{a+b}.$$

Equation 1.

To generate the image profile from the microscope, we multiply the Fourier series for the line/space array by the OTF and evaluate the resulting series.

$$I(x)=F(0)\frac{ka}{2\pi}+\frac{1}{\pi}\sum_{n=-\infty}^{n=\infty}\frac{1}{n}F(nk)\sin\left(\frac{nka}{2}\right)\exp(-jnkx)$$

Equation 2.

where $F(\omega)$is the complex OTF. The OTF is the Fourier transform of a real series (the intensity profile is real, hence its derivative is real) and so $F(\omega)$ will be complex conjugate (real part is an even function, imaginary part is an odd function), exp(-jnkx) is also complex conjugate, so equation 2 can be simplified to:

$$I(x) = F_r(0)\frac{ka}{2\pi} + \frac{2}{\pi}\sum_{n=1}^{n=\infty}\frac{1}{n}\sin\left(\frac{nka}{2}\right)\left(F_r(nk)\cos(nkx) - F_i(nk)\sin(nkx)\right)$$

Equation 3.

where all the components are now real and the resulting intensity profile is real.

Figure 8 shows the Fourier series for an infinite array of 1μ lines and 1μ spaces, the OTF and the series for the resulting image. The image components are the product of the OTF and the array components. Note that imaginary part is much larger than would be found in a real microscope and that the first component is at 0.5 lines/micron, corresponding to the pitch of 2μ.

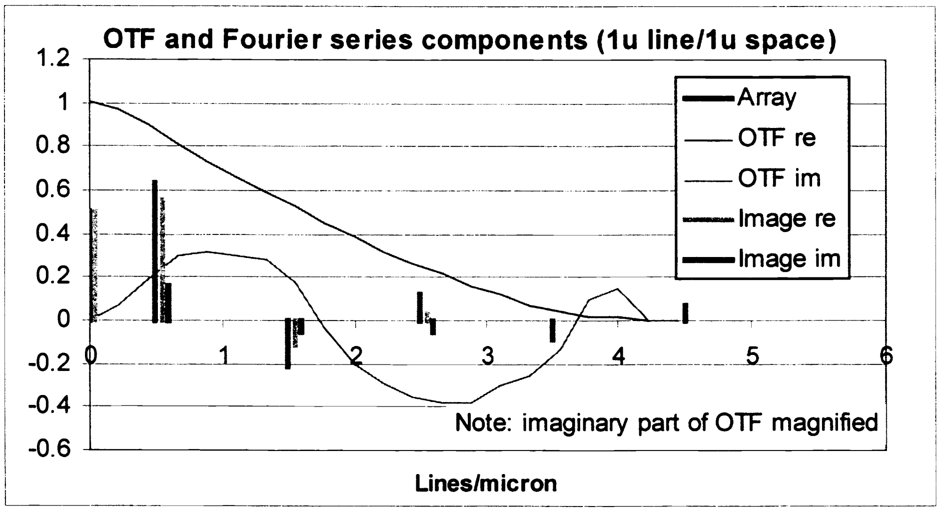

Figure 8. Fourier series components for an array of 1μ lines and 1μ spaces.

The resulting Fourier series is now evaluated to generate the image profile in figure 9. Note that the imaginary part of the OTF gives rise to an asymmetrical profile.

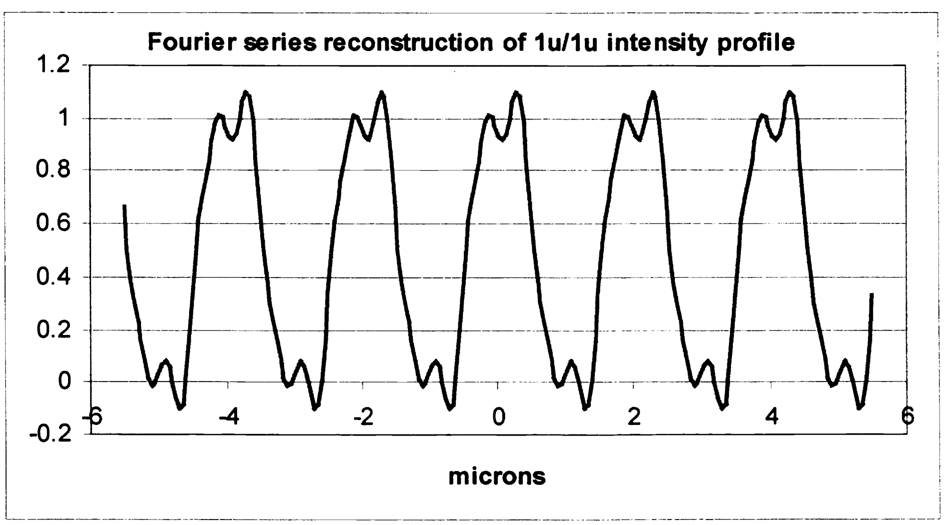

Figure 9. Resulting evaluation of the Fourier series.

Model Verification

Figures 10 to 13 show the comparisons between the model and actual intensity profiles, based on the OTF of figure 7.

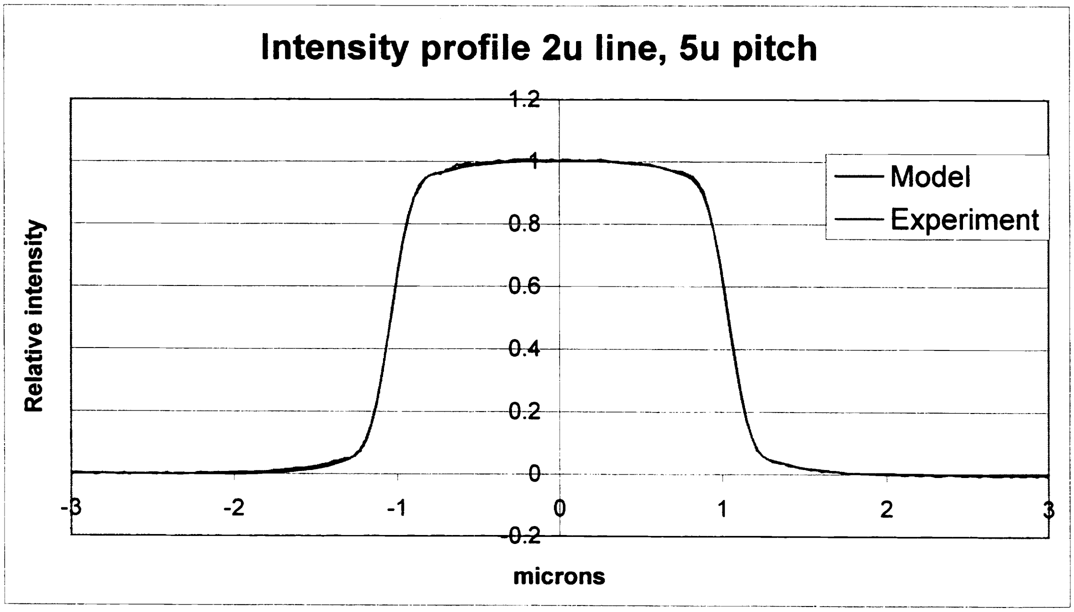

Figure 10. Comparison of OTF model with experiment of a 2μ line, 5μ pitch array.

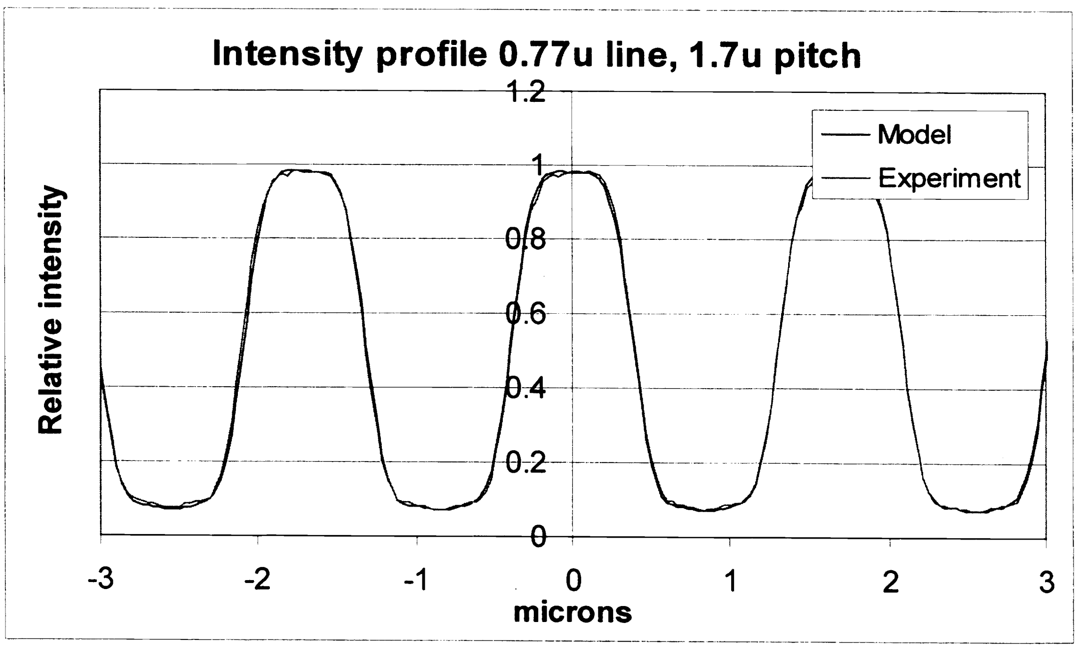

Figure 11. Comparison of OTF model with experiment of a 0.77μ line, 1.7μ pitch array.

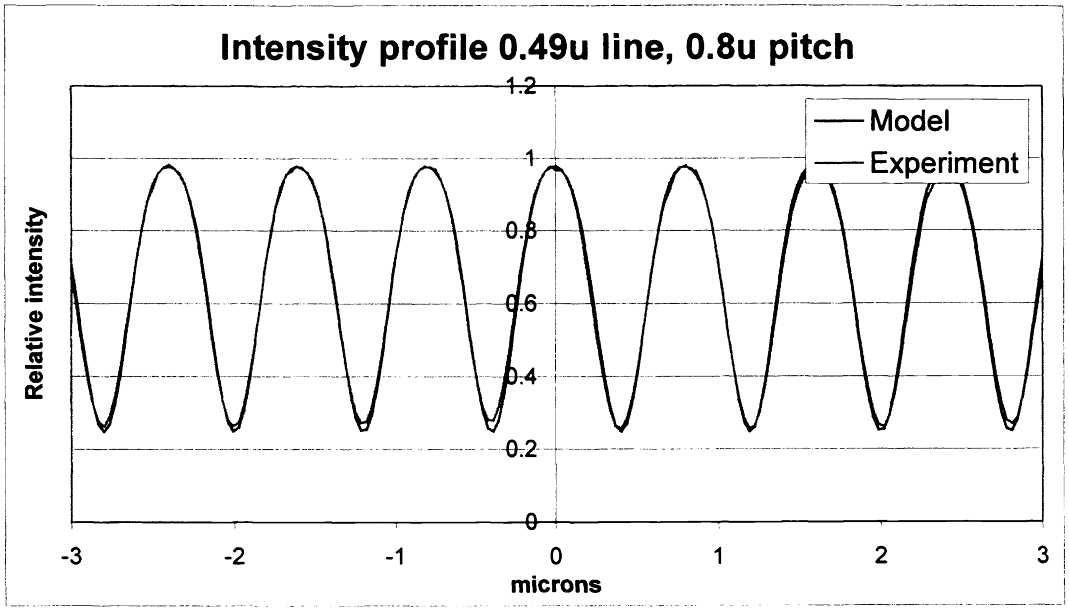

Figure 12. Comparison of OTF model with experiment of a 0.49μ line, 0.8μ pitch array.

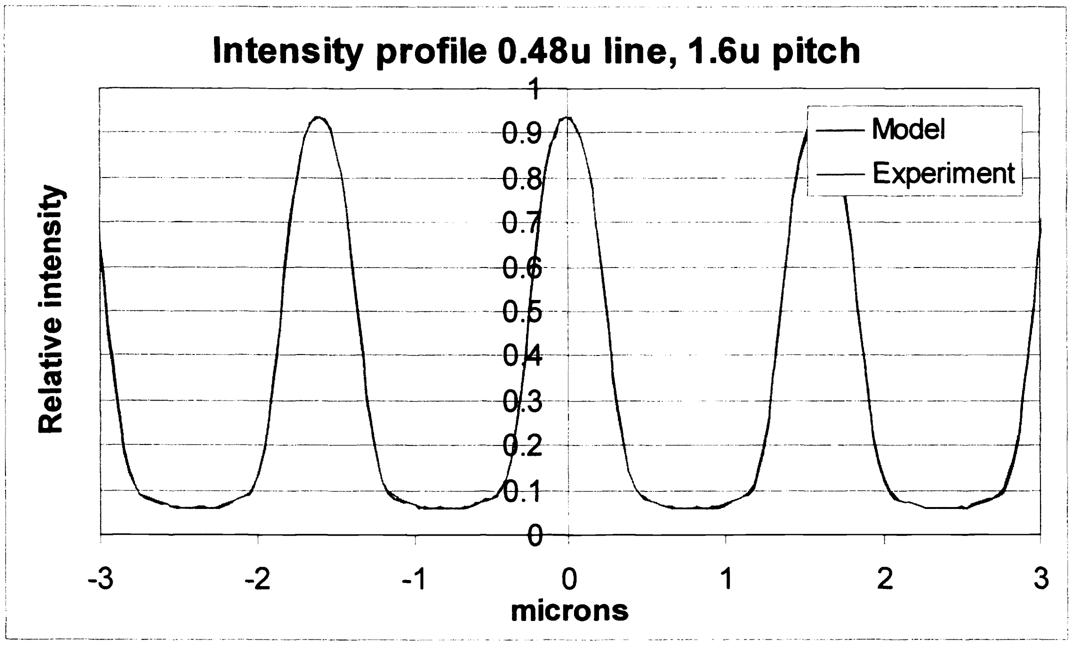

Figure 13. Comparison of OTF model with experiment of a 0.48μ line, 1.6μ pitch array.

An additional DC offset was added to the model profiles to match the experimental data. This offset may be due to light scattering in the system or some light passing through the chrome on the mask.

The agreement between the model, based on a single image capture, and the experimental data is excellent. We can feel confident that the OTF gives a very good representation of the imaging performance of the KMS-450i on binary masks. We now have a tool for evaluating the performance of the KMS-450i metrology system for different arrays of lines and spaces. The DC offset is not important to the algorithm evaluations as all the algorithms normalize the profiles based on the minimum and maximum values within the profile and hence are insensitive to offset.

CD Metrology

In this section we investigate the OPE on CD metrology performance for the threshold and maximum gradient algorithms. We introduce a new algorithm to compensate for the OPE and investigate its performance using the model.

Two algorithms that are commonly used for CD metrology are the Threshold and Maximum Gradient algorithms. The algorithms are illustrated in figure 14. For the threshold algorithm, each edge is considered independently. We normalize the profile based on the local minimum and maximum to the edge, fit a line to the edge and find the position where the line crosses the threshold and take this point as a measure of the edge position. For the Gradient algorithm, we numerically differentiate the profile and find the position of the maximum gradient and use this as the edge position.

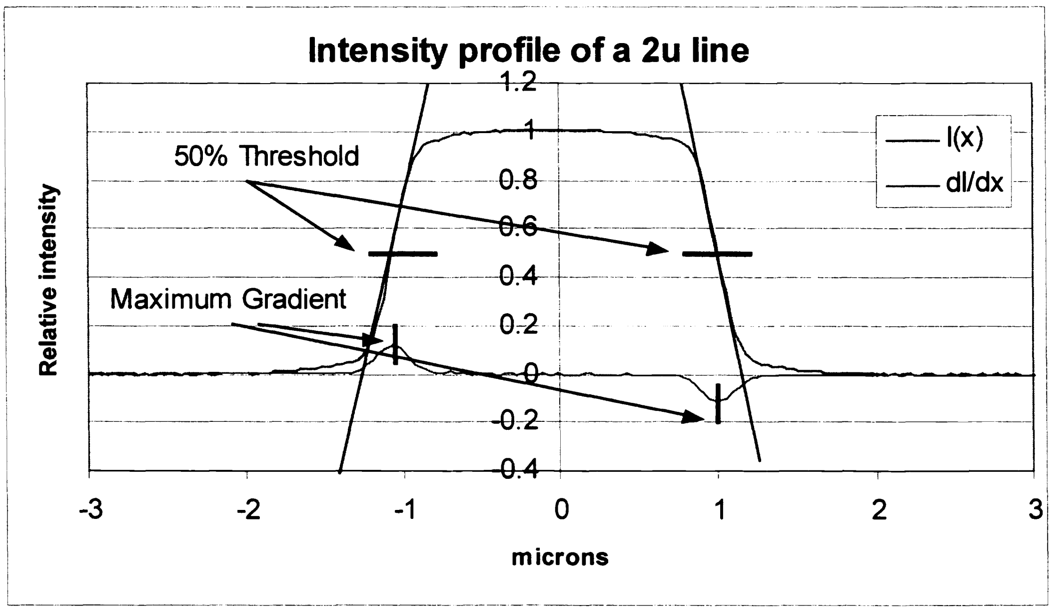

Figure 14. Threshold and maximum threshold algorithms.

Measurement Linearity on Isolated Features

It is useful to determine the linearity of the algorithms on isolated lines and spaces. To simulate isolated features, we use a pitch of 5μ in the model described. We then apply the algorithms to the calculated intensity profiles. Figure 15 shows the CD measurements and the deviations from the actual feature sizes for isolated lines. The curves for isolated spaces are identical due to the high degree of symmetry with this lens (essentially no imaginary component to the OTF).

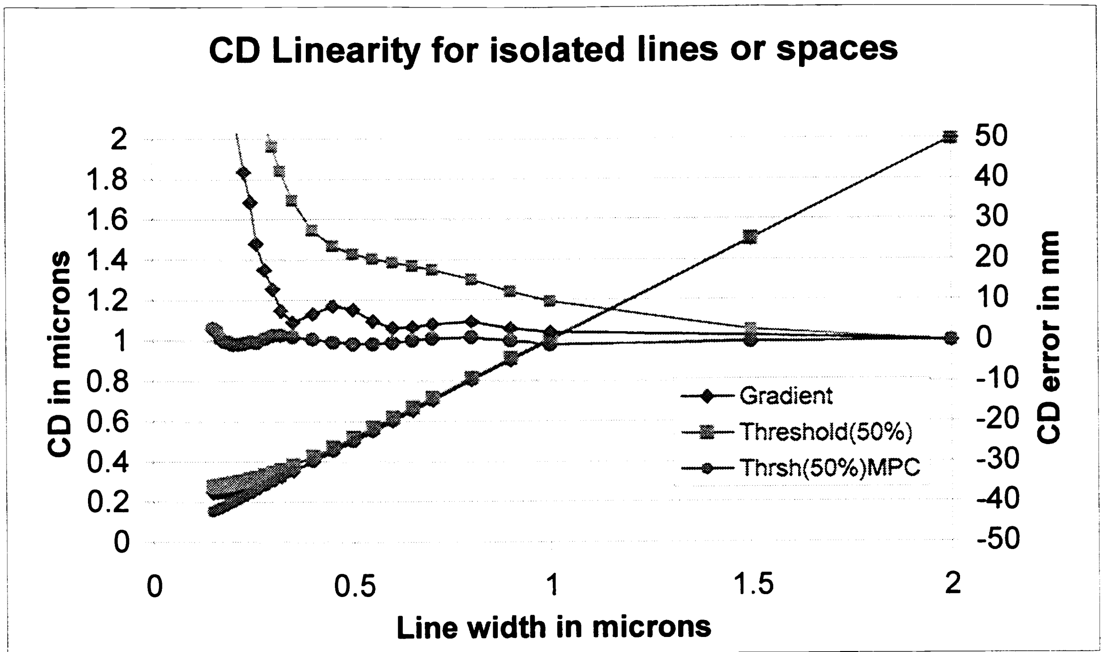

Figure 15. CD linearity for isolated lines and spaces. Errors are measured – actual values.

It is standard practice for users to measure isolated features and generate a lookup table or Multipoint Calibration (MPC) table which can be used to correct for these deviations in the CD measurement. The graph shows the same data corrected using an MPC correction. For this simulation, a polynomial was used to give the MPC correction, the uncorrected data was used to generate the MPC polynomial. The correction is, of course, excellent, the small deviations are due to the finite number of terms in the polynomial.

The gradient algorithm looks very promising as an alternative to the threshold algorithm remaining much more linear as the features become smaller.

OPE on CD Measurements

To investigate the Optical Proximity Effect, we now consider a line/space array with a fixed line width and varying width for the space. Figures 16, 17 and 18 show the variation in CD of a 0.3μ, 0.5μ and 1.0μ line in an array as the space is varied from 0.15μ to 3μ for different algorithms.

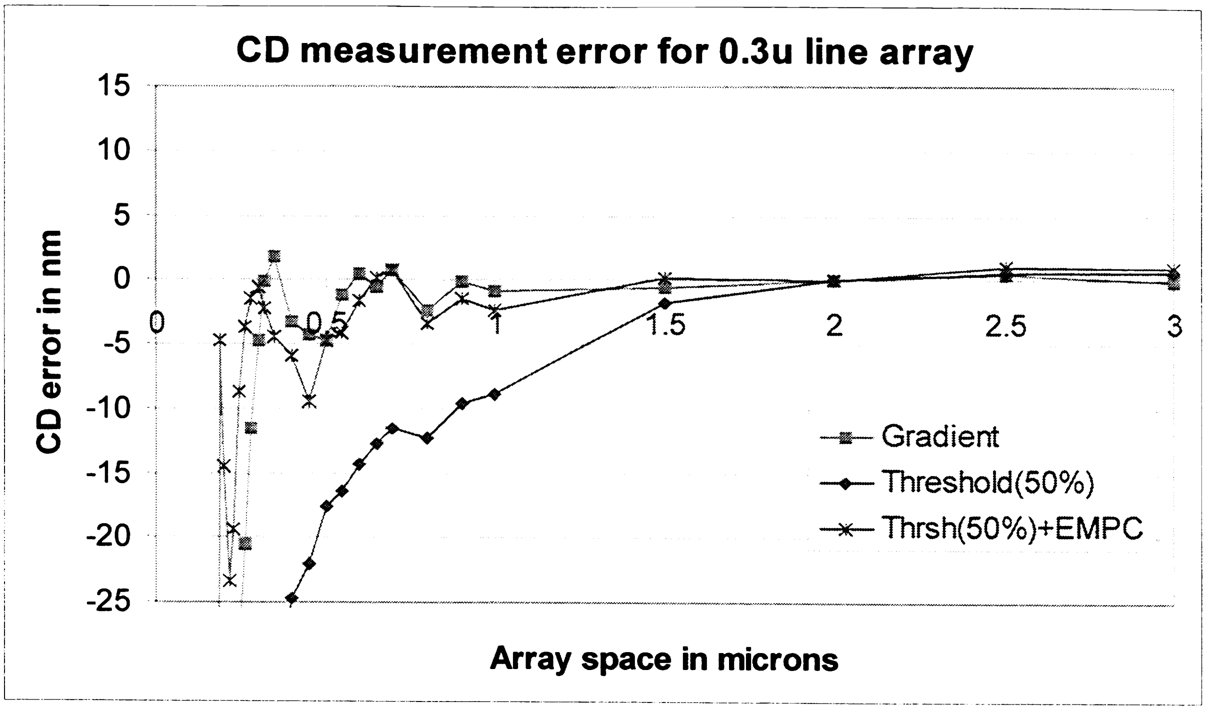

Figure 16. CD measurement of a 0.3μ line in a line/space array as a function of the width of the space.

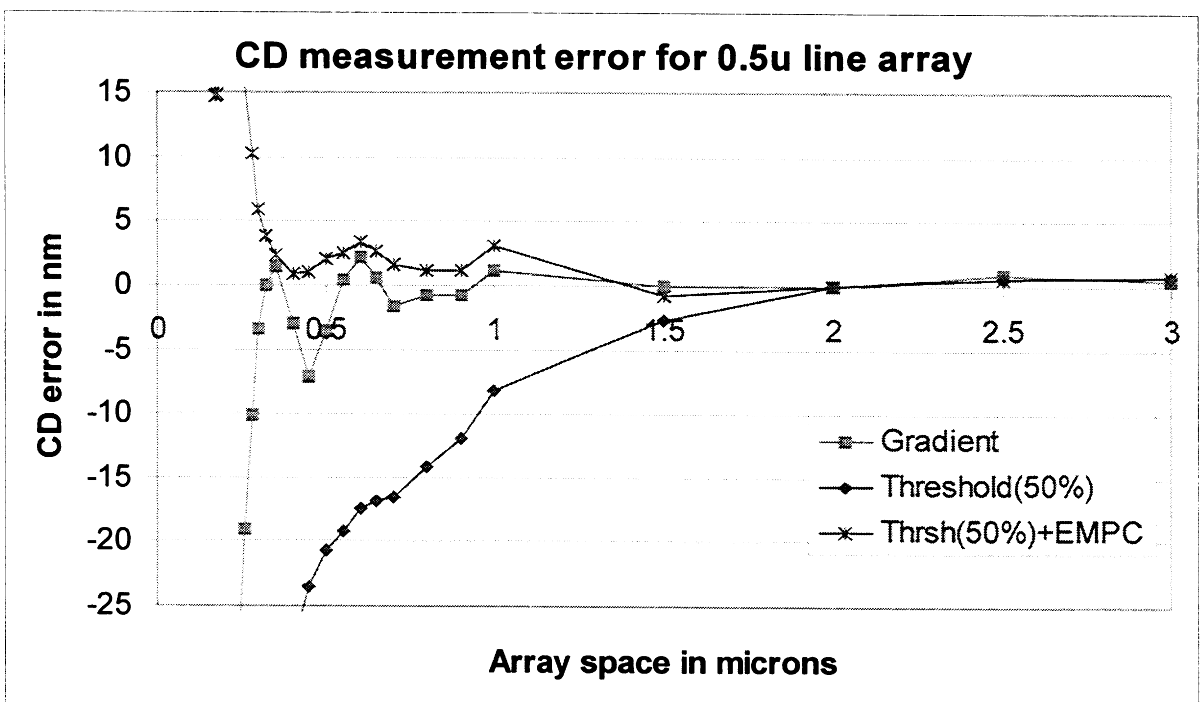

Figure 17. CD measurement of a 0.5μ line in a line/space array as a function of the width of the space.

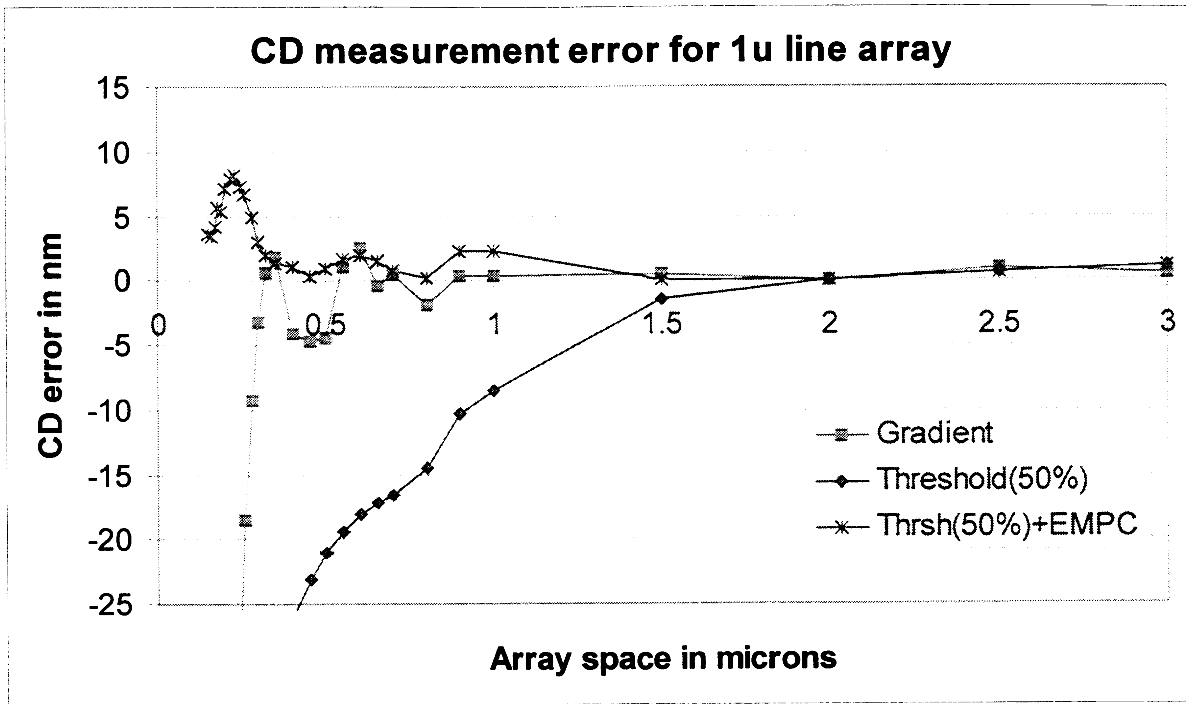

Figure 18. CD measurement of a 1μ line in a line/space array as a function of the width of the space.

The table summarizes the results from the figures above. It shows the minimum space width for a given algorithm, feature size and allowed measurement error.

	Gradient	Threshold(50%)	Threshold(30%)	Threshold(70%)	Threshold(50%)+ EMPC
0.3u 10nm	0.3	0.9	1.2	0.5	0.24
0.5u 10nm	0.3	0.9	1.2	0.55	0.28
1u 10nm	0.28	0.9	1.2	0.6	<0.15
0.3u 5nm	0.3	1.3	1.4	0.9	0.5
0.5u 5nm	0.45	1.3	1.4	0.95	0.31
1u 5nm	0.45	1.3	1.35	0.9	0.28
0.3u 3nm	0.5	1.4	1.5	1.2	0.8
0.5u 3nm	0.5	1.4	1.5	1.4	0.33
1u 3nm	0.5	1.4	1.4	1.2	0.3

For example the minimum allowed distance for measurement of a 0.5μ line using the gradient algorithm and a maximum CD error of 5nm is shown in the 6th row of the 1st column of the table to be 0.45μ. Similarly for the 50% threshold algorithm, the distance is 1.3μ. Results for different thresholds are included but in general 50% is to be preferred as this tends to optimize noise performance.

Extended MultiPoint Calibration

In the figures and table above, results are shown for a new algorithm (EMPC). This algorithm uses the standard multipoint calibration data used to correct for system linearity as a function of feature size to compensate for OPE.

Consider a line, space, line image as shown in figure 19. The position of the image of edge 3 is affected by the position of edge 4.

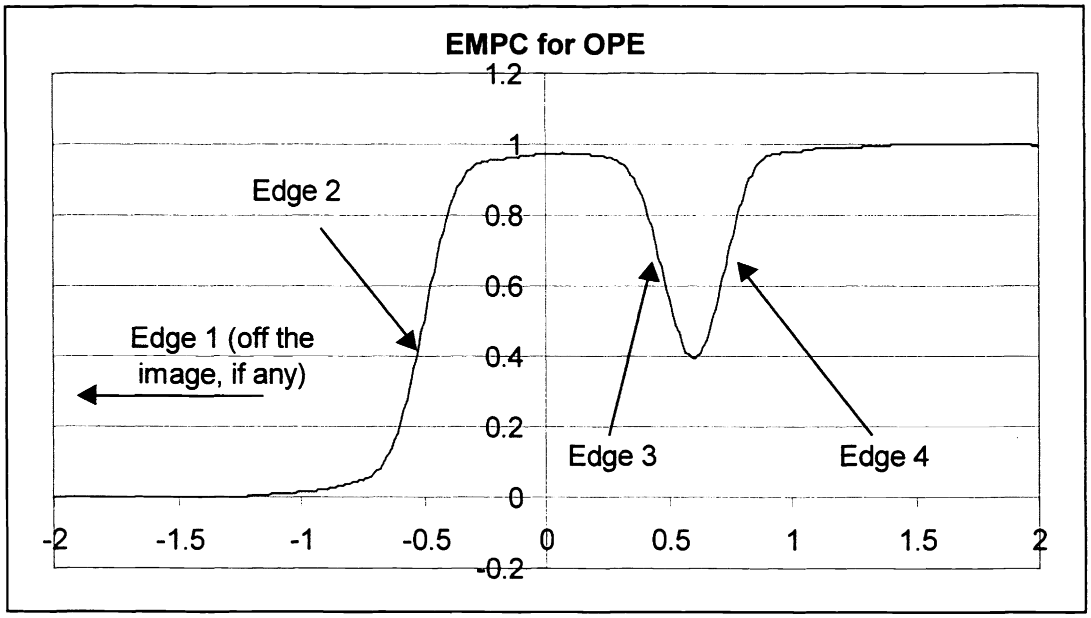

Figure 19. Extended multipoint calibration for OPE.

By inspection edge 2 to edge 3 is a line measurement and edge 3 to edge 4 is a space measurement. Let the standard multipoint calibration be represented by 2 functions; MPCL(line) is the corrected value for a line measurement and MPCS(space) is the corrected value for a space measurement. Let x_1 be the position of edge 1 etc. To first order the correct value for the edge3 to edge 4 space is MPCS(x_4-x_3). Since this is different from (x_4-x_3), the values for x_3 and x_4 are in error. We can assume that the error in x_3 and x_4 are opposite and equal so we can calculate an improved value, x_3', for x_3:

$$\mathbf{x_3' = x_3 - (MPCS(x_4-x_3) - (x_4 - x_3))/2}$$

If the corrected value for the space width is larger than the measured width, we subtract half the correction from x_3 and add half to x_4. If there were an edge to the left of edge 2 we would make a similar correction to give x_2'. Having obtained the corrected values of x_2' and x_3', we can apply the MPC so the final line width is:

$$\mathbf{width = MPCL(x_3'-x_2')}$$

In figures 16, 17 and 18 above, the results of the EMPC are shown to be very successful in compensating for the OPE in the modeled data.

Discussion

Traditionally CD metrology has been thought of in terms of measuring feature sizes. It is, however, instructive to think in terms of measuring EDGE positions, from which the feature sizes can be determined.

Consider an object consisting of an infinite half plane. Our algorithms can determine the position of the edge very accurately. Now consider an second edge approaching the first edge from either side (a negative edge approaching from the clear area corresponds to a line, similarly a positive edge approaching from the dark region is a space). This second edge will affect the apparent position of the first edge. For the threshold algorithm, the main effect is the "tail" of the second edge affecting the local minimum or maximum used for determining the normalizing factors. With the gradient algorithm, the effect of the second edge is not felt until the tail encroaches on the center of the first edge. This explains the improved linearity performance of the gradient algorithm. We can think of the second edge as "pushing" or "pulling" the image of the first edge. The MPC essentially is a measurement of the degree to which the edge has been pulled.

If we now think of our edge being approached on either side by an edge, it will be pushed and pulled by both edges depending on their distance. We can use the MPC data to correct for the error in position of the image of the edge. This is the basis for the Extended MultiPoint Calibration algorithm described above.

We have demonstrated that the EMPC algorithm should significantly extend the measurement capabilities of the KMS in the presence of other features (OPE). There is no overhead to the user to use the new algorithm as the MPC data is typically collected as part of the calibration to extend the linearity of the system for isolated features. The EMPC extends the use of the same data to make corrections for dense features.

The gradient algorithm is included in the latest version of the KMS software but appears to be significantly more noisy, with 3σ values several time worse than the threshold algorithm. This increased 3σ value has to be weighed against the improved linearity and OPE susceptibility. The potential advantages warrant investigation into the current implementation to improve the 3σ performance.

Any measurement algorithm is ultimately limited by signal to noise. As the feature sizes and spacings decrease, the signal (or modulation) is reduced and this needs to be considered in our expectations of any algorithm. The only solution to the S/N problem is to use a shorter wavelength or a larger NA lens. Figure 20 shows the OTF for a new 50x ELWD lens for through pellicle metrology and some preliminary results on a DUV system. It is apparent that the DUV optical system is not optimized but the improvements over the i-line lens is obvious. Figure 21 shows the calculated profile for a 0.4μ line on a 1μ pitch based on the OTFs.

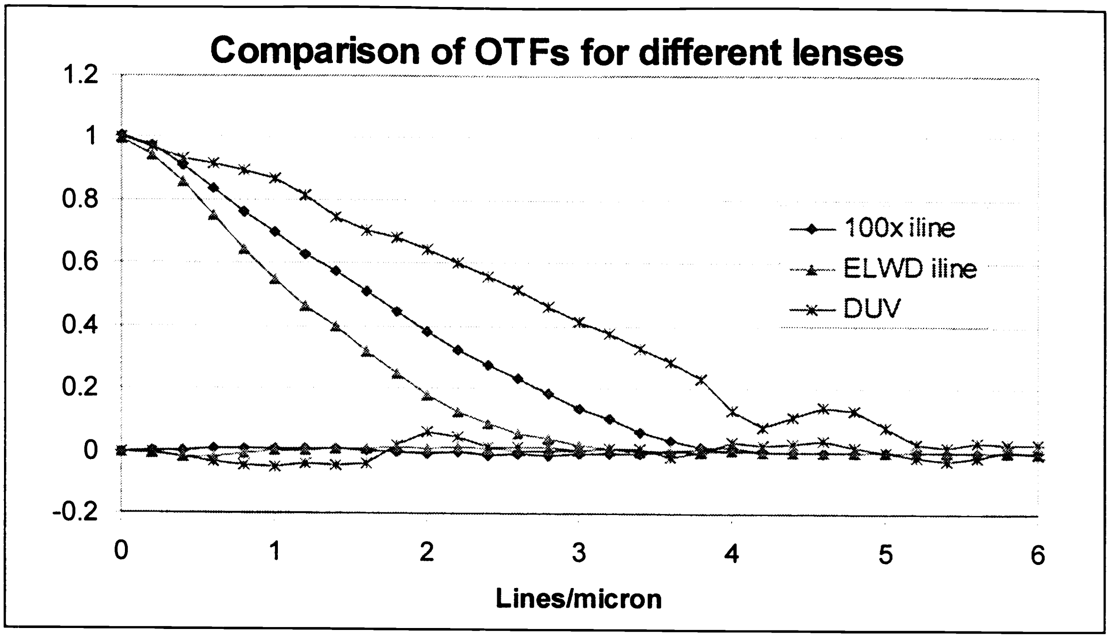

Figure 20. Comparison of OTFs for different objective lenses.

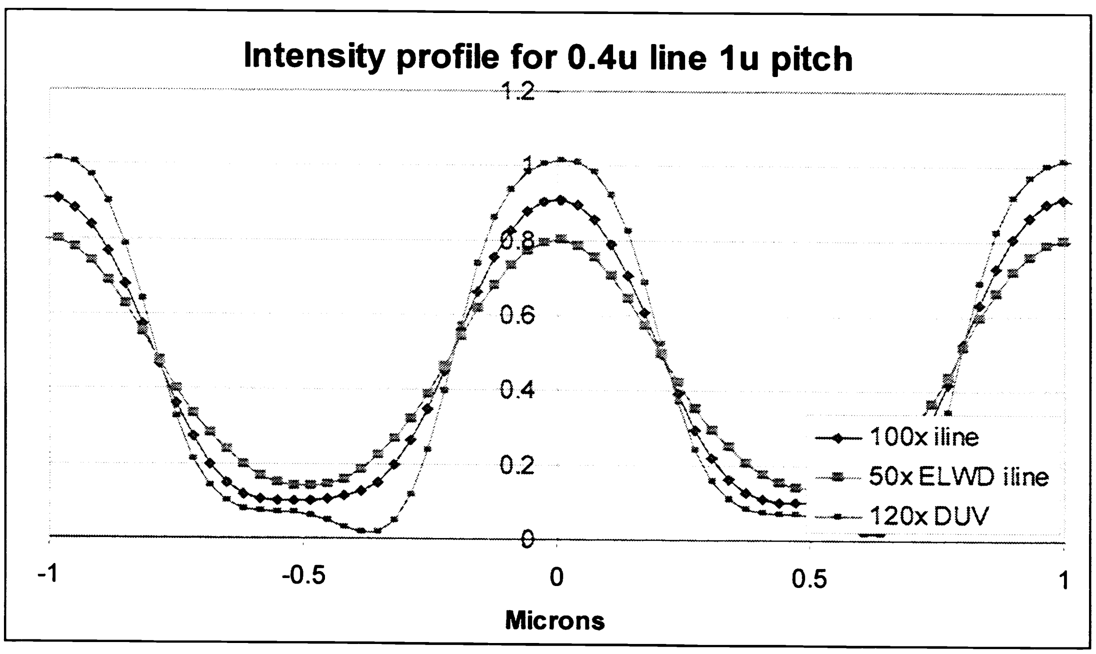

Figure 21. Intensity profiles for lenses shown in figure 20. The improvement with DUV is apparent even in this non-optimized system.

Conclusions

- We have demonstrated that we can calculate the OTF for a real microscope and use this to generate intensity profiles with excellent agreement with experimental data.
- Using the model, we have evaluated the Optical Proximity Effect on CD metrology using different algorithms.
- The Gradient algorithm is less susceptible to OPE and has better linearity for small features.
- We have developed a new algorithm (EMPC) which extends the MPC from isolated to dense features allowing corrections for the OPE.
- The EMPC requires no additional calibration of the KMS metrology system.
- The EMPC is applicable equally to the gradient as to the threshold algorithm.
- We need to implement and demonstrate the EMPC and verify its performance against AFM measurements.
- Although "there is no substitute for horsepower (wavelength)" these investigations should enable us to get more from our current images.

References

1. Inoué S. (1987) Video Microscopy. Plenum Press, New York. Page 467.

2. Schade O.H. (1975) Image Quality – A Comparison of Photographic and Television Systems. RCA Laboratories, Princeton NJ. Page 6

3. Wilson T. (1984) Scanning Optical Microscopy. Academic Press, page 45.

4. Howatson, A.M.(1972) Engineering Tables and Data. Chapman and Hall , page 15.