SURFACE AND INTERFACE ANALYSIS
Surf. Interface Anal. 2005; **37**: 951–958
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sia.2087

SIA

Scanning electron microscope dimensional metrology using a model-based library

J. S. Villarrubia,* A. E. Vladár and M. T. Postek

National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

Received 7 January 2005; Revised 7 April 2005; Accepted 7 April 2005

The semiconductor electronics industry places significant demands upon secondary electron imaging to obtain dimensional measurements that are used for process control or failure analysis. Tolerances for measurement uncertainty and repeatability are smaller than the spatial resolution of edges in the scanning electron microscope (SEM) that is used to perform the measurements. Image processing techniques, historically used to identify edge locations, are inadequate under these conditions. An alternative approach, based upon Monte Carlo electron transport modeling to assign edge positions, has been developed. The specimen shape is parameterized, and parameters are iteratively adjusted to produce the best least squares fit to the measured image. Because Monte Carlo simulators are too slow to be used directly in such an iterative calculation, the Monte Carlo technique is used relatively few times to construct a library of results for parameters spanning the process space of interest. A function that interpolates the library then becomes a surrogate that is used to rapidly compute the model function as needed. This procedure has yielded measurement results from top-down SEM images that are in good agreement with cross-section measurements and that have as much as a factor of 3 better same-site repeatability than the more traditional techniques. Published in 2005 by John Wiley & Sons, Ltd.

KEYWORDS: critical dimension (CD); model-based library; scanning electron microscopy (SEM); dimensional metrology; linewidth

INTRODUCTION

The dimensions of transistor gates, currently at less than 40 nm according to the International Technology Roadmap for Semiconductors (ITRS),[1] are among the smallest routinely manufactured objects for which strict dimensional control is required. Transistor gates are among the smallest features that are designed for printing. For this reason, and because their dimensions are directly tied to device performance, the gate size is the main 'critical dimension' (CD) that is tracked in order to stabilize the manufacturing process against drift and to provide early warning of more serious process variations. The features themselves are already small and the allowable uncertainty in their measurements is smaller yet. As circuit feature sizes decrease, the amount of memory that can be packed onto a semiconductor chip increases and the microprocessor speeds increase. Both these outcomes represent increased value, so economic incentives push the semiconductor electronics industry to produce the smallest features that are technically feasible. In 1998, Ausschnitt and Lagus estimated that a 10 nm difference in gate size resulted in a microprocessor speed difference that in turn made a $100 difference in the market value of the device.[2] On this basis, they estimated the value of CD control to be $10/nm per microprocessor, a figure they estimated would increase to $20/nm when gate dimensions declined by another factor of 2, as they have today. With microprocessor annual sales in hundreds of millions of units (273 million in 2003; see Ref. 3) the economic value of reduced dimensions is measured in billions of dollars per nanometer. This is a significant incentive for manufacturers to push lithographic and measurement technologies to their limits.

Thus, it is nearly inevitable that there are technical measurement problems to overcome. These are described in more detail in the next section. Following this, we describe a solution based upon a model-based library (MBL) approach to dimensional measurements in the scanning electron microscope (SEM). Results of several recent tests of this method are reviewed and discussed in the final two sections.

*Correspondence to: J. S. Villarrubia, National Institute of Standards and Technology, 100 Bureau Dr. Stop 8212, Gaithersburg, MD 20899-8212, USA. E-mail: John.Villarrubia@nist.gov
Contract/grant sponsors: SEMATECH; NIST Office of Microelectronics Programs.

THE PROBLEM

Routine CD measurements during production must be non-destructive and fast. The requirement for speed means that secondary electron imaging is generally preferred to backscattered electron or other imaging modes since secondary electron count rates are generally higher, producing a better signal-to-noise ratio for a given measurement time. In a secondary electron image, the steep edges of a gate or line appear brighter than horizontal surfaces like the line top or substrate. This is because electrons incident at the top and just inside of the line edge produce secondary electrons that escape the specimen through both the side and the top of the line, instead of only through the latter as is the case for electrons incident far from the edge. This additional brightness is the so-called 'bloom' at an edge (Fig. 1a).

The standard measurement algorithms in use today assign the edge position by using image-processing methods to locate the bloom peak. For example, a tangent line can be drawn at the outside edge of the bloom peak and another through the background farther from the peak, as in Fig. 1(b). The intersection of these two lines is the edge position, according to the regression to baseline method. Other methods might choose the position where the tangent line crosses a different intensity threshold (e.g. half way between the baseline and the maximum) or the position where the slope of the intensity curve is the greatest. Of course, the true edge position might be elsewhere within the bloom peak, the width of which is essentially the microscope's spatial resolution for edges. When lines were wider and measurement tolerances were more relaxed, this uncertainty was not considered significant.

However, measurement errors must be small compared to feature size. With transistor gates now narrower than 40 nm, the industry's most demanding measurement requirements[1] are for a 3-standard-deviation repeatability and total measurement bias approaching 0.5 nm and 3 nm respectively. These tolerances are marked in Fig. 1(b) for comparison with the edge bloom width. It is evident from this comparison that today's tolerances are significantly tighter than the instrument's spatial resolution for edges. The most obvious effect of edge assignment error is that the computed gate width differs from the correct value. If the average measurement error, or bias, for a right edge is ε, then the bias for the corresponding mirror image left edge would be $-\varepsilon$ and the measurement error in the difference between the two would be 2ε; that is, the left/right mirror symmetry ensures

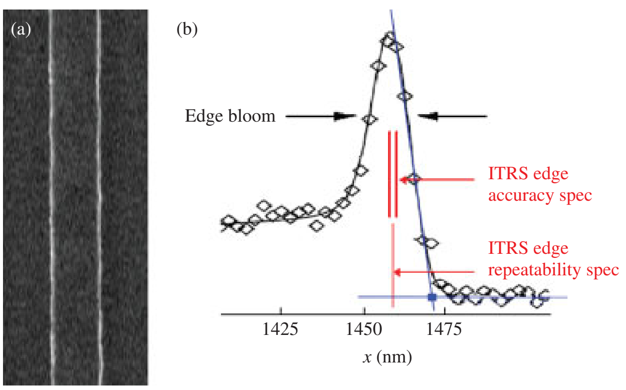

Figure 1. (a) Illustration of an edge bloom. The line's left and right edges are brighter than the background. (b) The intensity profile across such an edge exhibits this characteristic peak with a finite width. ITRS measurement requirements[1] are shown at the same scale. The regression to baseline method assigns the edge position to the intersection between the background and a line tangent at the outside edge of the peak.

that the measurement errors add rather than cancel when gate width is determined.

Sometimes – for example, the manufacturing process control example mentioned in the introduction – only relative measurements are needed. We wish to know only whether the gate size has changed from one manufacturing lot to the next. In such cases, if the measurement bias described in the previous paragraph were constant, it might not be of great importance. Unfortunately, it is not likely to be constant. The size of the error is expected to depend upon secondary specimen characteristics. A secondary characteristic is a feature of the specimen that we do not wish to measure. From the point of view of a measurement of the bottom width of a gate, for example, the rounding of specimen corners or slopes of the edges are secondary characteristics. Although these secondary characteristics can change without changing the actual gate width, any such change may very well affect the image and thereby change the *measured* gate width. Measurement bias as a function of sidewall angle for four edge assignment methods is reproduced from a simulation study[4] in Fig. 2. These results indicate that the rate of change in bias with sidewall angle is 1 nm/1° to 2 nm/1° for the most commonly used methods of edge assignment, a result that was also corroborated experimentally.[5] Therefore, reduction of edge assignment errors is important for relative as well as absolute measurements.

MODEL-BASED LIBRARY APPROACH

We have been implementing an alternative measurement approach. The essential components are (1) a physics-based model that simulates the image for a given parameterized

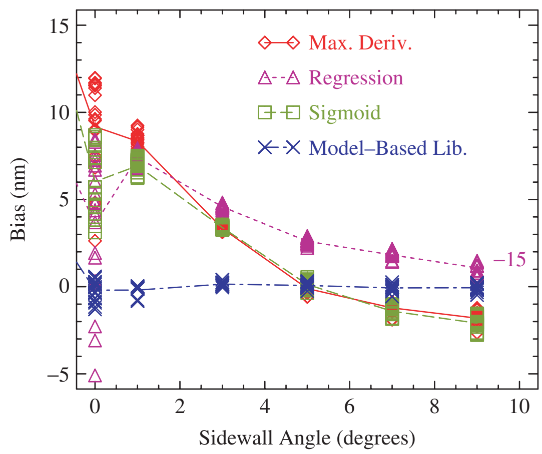

Figure 2. Bias (i.e. average error) in widths determined using four different edge assignment methods generally changes as the specimen sidewall angle changes. The methods were applied to simulated noisy images, so the true width was known. Curves are for a fixed noise level and beam size. Some other parameters (e.g. e-beam cone angle, line separation, corner radius) were varied – hence the scatter in the data within each curve. (For details, see Ref. 4.) Note that the simulation used the same model as the MBL method, so the favorable MBL result is meant to indicate performance that is possible when the model matches reality, not to prove that it does. (From Ref. 4).

specimen shape, composition, and imaging conditions and (2) a capability to invert the model to determine the parameters, given its measured image. Since, as we will see, the model in component (1) is a time-consuming Monte Carlo simulation and the inversion in component (2) requires many evaluations of the model, a third component is of great practical benefit. This component (3) is the approximation of the model by interpolation of a library of simulation results. These three components are now explained in more detail.

Forward modeling

The first component is an imaging simulator. This simulator embodies a model of the physics of the interaction between the SEM and the specimen in a computer code. For our modeling, we use the MONSEL Monte Carlo–based simulator.[6] This simulator uses Mott elastic scattering, includes explicit treatment of Möller (sometimes also called inelastic Mott in the nonrelativistic limit that we use) and plasmon-mediated generation of secondary electrons, and approximates electron energy losses with a modified Bethe continuous energy loss formula. Incident electrons and all generations of secondary electrons are followed individually until they either escape the specimen or their energies fall below the work function, making escape impossible. Secondary electrons (i.e. electrons with kinetic energies less than 50 eV) that leave from the upper surface of the specimen are counted, and the ratio of the count to the number of incident electrons (the yield) is reported. According to one recently added option, an electron that leaves from the upper surface is always deemed to have escaped the sample entirely. This option approximates an idealized critical dimension SEM. These instruments are typically operated with high external electric fields that are intended to pull all escaping low-energy electrons into the detector, regardless of their initial trajectories. As per an older option, an electron that leaves the upper surface is assumed to travel in a trajectory that is approximately straight on the scale of specimen topography, and it is deemed to have escaped if that straight trajectory does not reintersect another part of the specimen. This option allows for the possibility that electrons will be recaptured and is appropriate when electric fields outside the specimen are low. We currently approximate situations intermediate between these extremes by interpolating between these two solutions. Let us call the secondary electron yield predicted by such a Monte Carlo simulation for an idealized zero-diameter beam, $Y_M(x;g_i)$. Here x denotes the landing position of the incident electron beam and g_i denotes a set of parameters, both geometrical (line height, edge positions, edge angles, radii of corner curvatures, etc.) and instrumental (incident beam energy and angle, etc.). Then our model, $M(x;g_i,b_i)$, of a linescan produced by an SEM measurement is

$$M(x;g_i,b_i) = sY_{\mathrm{M}}(x-x_0;g_i) * B(x;b_i) + y_0 \qquad (1)$$

Here the convolution with $B(x;b_i)$ accounts for the nonideal distribution of electrons in a real beam. It represents the distribution of electrons as a function of the distance, x, from the targeted landing spot. The b_i represent some parameters of this function. Currently, we use

$$B(x;b) = \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{x^2}{b^2}\right) \qquad (2)$$

This is a normalized Gaussian with b as its standard deviation. The parameters s and y_0 in Eqn (1) account for the fact that the SEM does not output the absolute yield. Rather, the measured secondary electron count is typically scaled and offset (a contrast and brightness adjustment) to make best use of the available range of the instrument's analog-to-digital converter. The parameter x_0 produces a lateral shift of the model result that allows it to be matched with the features' positions in the image.

Inversion

If the parameters s, y_0, b_i, and g_i that describe the instrument and specimen are known, then Eqn (1) provides a way to calculate the image. In practice, however, we usually want to do the reverse. We have a measured image, $I(x)$, and we need to solve for any unknown parameters, particularly those that describe the specimen. We do this in the usual least squares fashion by finding parameters that minimize

$$\chi^2 = \sum_x [I(x) - M(x;g_i,b_i)]^2 \qquad (3)$$

There are a number of standard algorithms to solve this minimization problem by iterative adjustment of the parameters. We use the method of Levenberg-Marquardt.[7–9] This method approaches a gradient (steepest descents) search for the minimum of the χ^2 surface when far from the minimum and converts smoothly to the analytical solution for a parabolic surface as the minimum is approached. It is efficient in the sense that it converges with far fewer evaluations of the model function than many alternative methods. Nevertheless, the number of iterations required increases with the number of parameters, and many of our fits require tens to hundreds of evaluations per image linescan. Therefore, computation time of the model function is an important consideration.

Using a model-based library

Iterative minimization of an expression (Eqn (3)) that contains a Monte Carlo simulation, each evaluation of which may require an appreciable fraction of an hour, is unacceptably slow. Accordingly, we developed an MBL method that is an extension of a concept published for SEM by Davidson and Vladár[10] in 1999. As they implemented it, the part of the model that we have called $Y_{\mathrm{M}}(x-x_0;g_i) * B(x;b_i)$ in Eqn (1) was evaluated for an assumed beam shape and for all combinations of a discrete sampling of the parameters, $\{x_0, g_i\}$. (The beam parameters, b_i, could have been included in this sampling, if desired.) Each of the curves in this 'library' was fit to the measured linescan by the best choice of s and y_0. (Solving for these only is a quickly solvable linear least squares problem.) The result was a sampling of the χ^2 function on a discrete regular grid in a chosen neighborhood. The parameters that correspond to the minimum χ^2 on this grid are assumed to be the best estimate of the target geometry.

The method produced a predicted geometry for a photoresist line that agreed well with a cross section of the line.[10]

However, in this version of the method, the output parameters cannot be between the sampled grid points, so measurement resolution is limited by grid spacing. Tabulation of all χ^2 values in a neighborhood is not as efficient as the more targeted search provided by standard nonlinear least squares methods, but these methods require the model function to be available for any parameter values, not just those on a discrete grid. The number of simulations required to build the library is of order M^N, where M is the number of values for each parameter and N is the number of parameters. If the library must include simulations of the entire specimen (i.e. the entire $Y_M(x - x_0; g_i) * B(x; b_i)$, function), this rapidly becomes very large for specimens more complex than an edge. For example, if each line has 2 wall angles, 2 top corner radii, and a width, a linescan of 10 lines will have $N = 59$ (5 parameters for each of 10 lines plus 9 line separations).

The situation is greatly improved by two extensions we have made to the method: image stitching and interpolation of the library. In our implementation, the library contains only individual edges rather than the entire structure. The yield function for right edges at $x = 0$ is computed by MONSEL at values of geometrical parameters only for a single edge (Fig. 3). For example, the sidewall angle parameter could vary from 0° to 10° in 1° steps and the corner radius from 0 nm to 100 nm in 10-nm steps (although there is no requirement that the steps be all the same size).

Yield curves for structures containing multiple lines are constructed using the edge library multiple times. For each edge in the target, the curve with the same edge geometry parameters is retrieved from the library. The library is interpolated when the required parameter values lie between entries. If a left edge is required, it is formed from its right edge counterpart by reflection. This edge is then positioned at the required location. By this stitching procedure, the yield curve $Y_M(x - x_0; g_i)$ can be constructed for any combination of lines and trenches. A relatively small single-edge library can be used to simulate structures with a parameter space of much larger dimension. In the 10-line example given in the preceding text, a 2-dimensional library suffices for a 59-dimensional parameter space. Stitching is justified when either (a) the shape of an edge's yield curve does not depend upon the edge's position within the image or (b) when any such positional dependence is taken into account by the introduction of a suitable parameter in the model. An example of (a) would be isolated lines that are too wide for electrons generated near one edge to leave the line via the other. In this case, only bloom separation but not bloom shape is a function of linewidth. This should be true in Si for linewidths down at least to 35 nm with 1 keV incident electrons, and smaller than that with lower energies.[11] An example of (b) might be edges on opposite sides of a trench such as will be discussed in the next section when we discuss Fig. 6. Such edges do interact because the intervening vacuum does not impede electrons. However, it is easy to account for the interaction by explicitly including edge separation in the structure modeled by MONSEL. Knowing an approximate separation, such as can be determined directly from the unmodeled image, is good enough because the bloom shape varies rather slowly as the separation changes.

Standard methods (e.g. Ref. 8, Sec. 3.6) are used to linearly interpolate the multidimensional library when required. Interpolation provides a continuous rather than

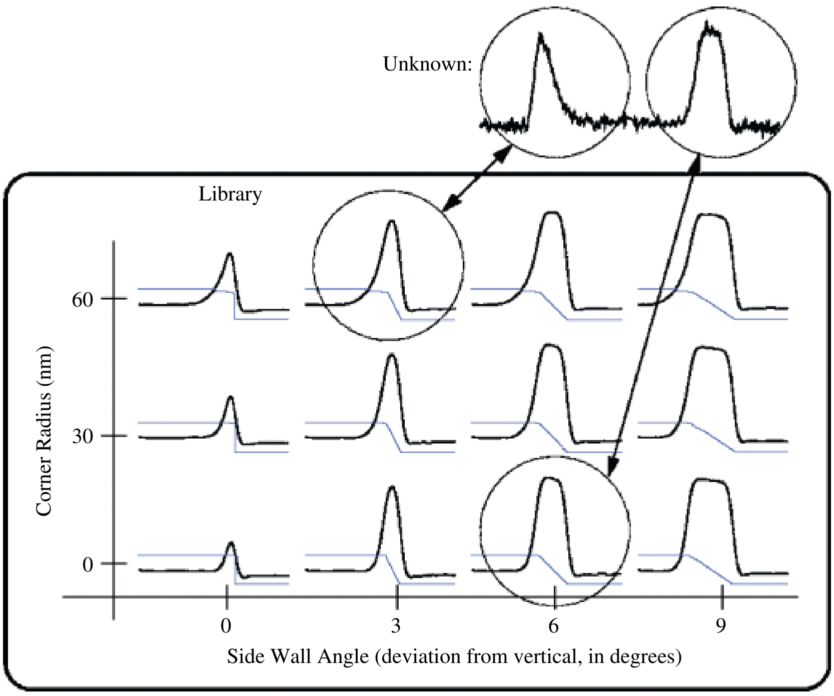

Figure 3. Schematic illustrating measurement using an edge library. Linescans computed for edges with varying parameters (wall angle and corner radius here) comprise the library. Parameters for the unknown specimen are determined by matching, with interpolation as needed. From Ref. 12.

discrete model function, permitting the use of efficient least squares algorithms. Parameter increments can be chosen as needed to give an accurate interpolation, without regard to measurement resolution since this resolution is no longer limited by the grid interval. Fortunately, the yield function appears to change slowly and continuously enough that parameters need not be too tightly spaced. Accuracy can be checked by comparing the interpolation results to results of a full calculation at a sampled set of points. We obtained quite good agreement between the interpolation and full Monte Carlo simulation even when sidewall angle spacings in excess of 1° were used.[11] Since a 10° range has, in our experience, been enough to cover the expected manufacturing process variation, the library can contain relatively few entries.

Interpolation and stitching of an MBL can be considered a surrogate model function, replacing the full Monte Carlo calculation for the purpose of iterative least squares. After the comparatively modest initial time investment to construct the library, it offers large speed improvements (a factor of 10^5 is typical in our experience) with negligible difference in output values.

TESTS OF THE APPROACH

Accuracy

Does the accuracy of the MBL method meet the requirements of the ITRS? In order to answer this question, it is necessary to compare the MBL measurement to the result of some other more trusted measurement technique. Unfortunately, this is difficult. At the nanometer accuracy level desired by the ITRS, other techniques have their own important measurement artifacts. For example, atomic force microscope images of lines are wider than the actual lines by an amount determined by the tip size.[13] The corrected width is uncertain by at least the uncertainty with which the tip size can be determined.

We chose to make comparisons of MBL to cross-sectional SEM, which is a standard practice in the semiconductor industry. For such a measurement, the specimen is cleaved at right angles to the line. The specimen is repositioned to view the line from the end. (The viewing direction is indicated by an arrow in Fig. 4(a).) There are a number of important error sources in cross-sectional SEM that such a comparison must account for: (1) The cross section provides a width at a single position along the line. It is generally difficult to accurately identify this position with a position in the top-down image. To the extent that the measurements occur at different locations, differences in linewidth from place to place are a source of error. This generally (i.e. as long as the registration uncertainty is larger than the roughness correlation length) produces an uncertainty that is equal to the linewidth roughness. (2) The cross-section SEM image is subject to edge bloom and must be modeled in a way similar to the top-down image. The modeling for this configuration is arguably simpler. Since the electron beam travels along the length of the line, parallel to the sidewalls, more of the relevant geometry is known *a priori*. A library-based approach is not required. Once the beam scans past the edge of the line, there is no obstacle to the electrons for a considerable distance. In effect, there is no substrate. On the one hand, this removes any complications that would otherwise be introduced by scattering there. On the other hand, it means that (3) the cross-section geometry is more sensitive to the instrument's depth of field (caused by divergence of the electrons from the point of best focus) than is the top-down geometry.[14] In our first test, lines in polycrystalline Si were cleaved and then imaged top down near the cleavage plane (Fig. 4(a)). An area of the image slightly away from the cleaved edge – a distance sufficient to eliminate bloom effect from the edge – was analyzed by MBL. A typical fit of the model to a measured linescan is shown at the top of Fig. 4(b), with the corresponding library line profile at the bottom. The scale of the top-down and cross-section images was matched independently of this measurement by imaging a nearby periodic array of lines and requiring the periodicity to be the same. The MBL line profile

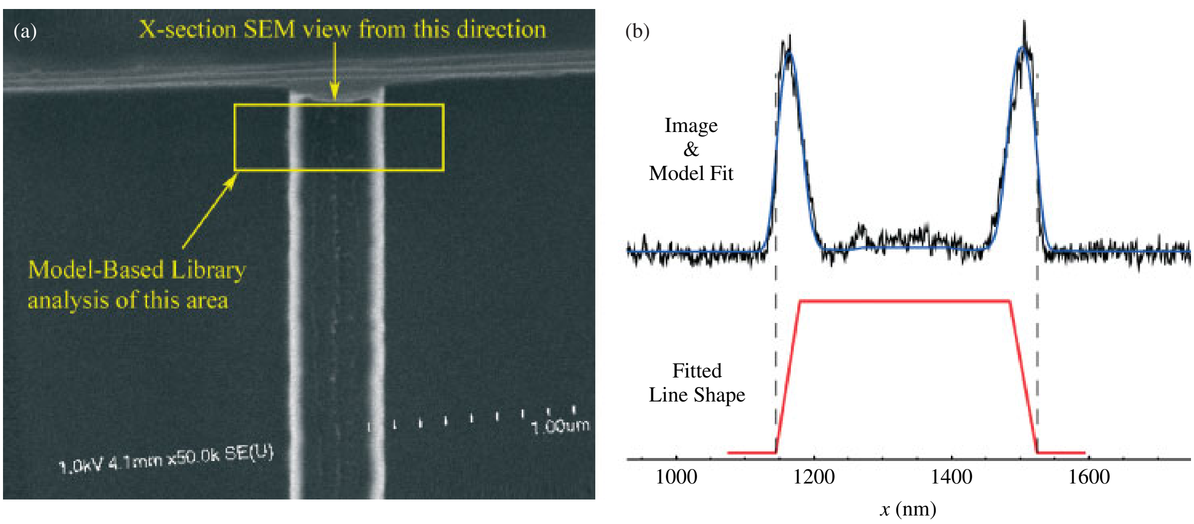

Figure 4. (a) Geometry for comparison of MBL to cross-section measurements. A cleaved polycrystalline Si line is imaged top down and the area (box) near the cleaved edge is analyzed. The specimen is also imaged (resulting in Fig. 5(a)) in cross section from the direction indicated. (b) A typical model fit to a linescan is shown above the simulation input line shape that produced it. From Ref. 12.

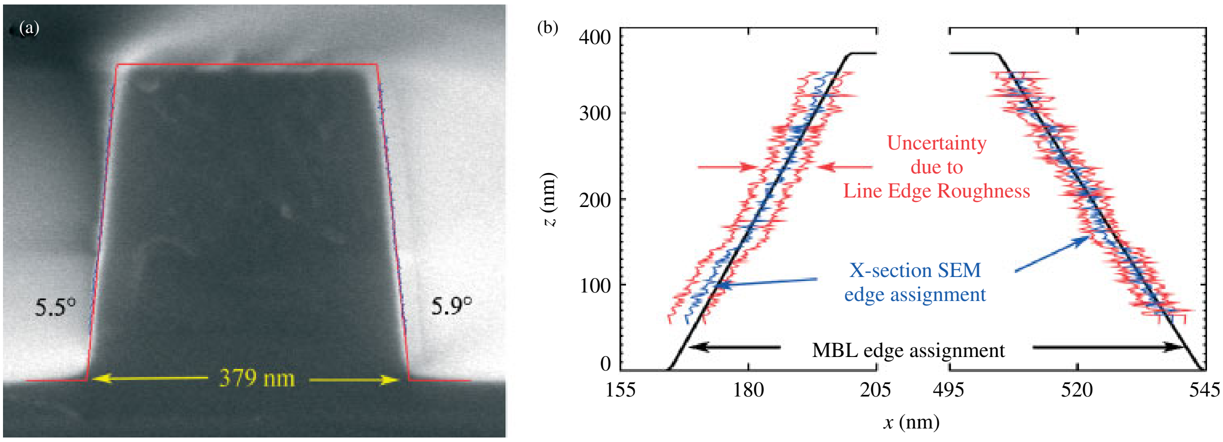

Figure 5. (a) Cross-section image of the line, with its assigned edge (noisy curve) upon which the MBL profile is superimposed for comparison. Edge regions are expanded in (b) for a better view. Uncertainty bands due to line roughness are shown. From Ref. 12.

is overlaid on the cross-section image in Fig. 5(a). (This is the smooth trapezoidal line.) The edge assignment from the cross-section image (a rougher line) is also shown. The edge regions are expanded in Fig. 5(b), where uncertainty bands of ± 3.6 nm, equivalent to ± 2 standard deviations of line edge roughness are shown. The average widths determined by the two methods differed by 1.4 ± 4 nm. The left edge angles differed by 0.2 ± 0.4 and the right edge angles differed by 0.1 ± 0.3. Details are in Ref. 12.

When lines are close to each other there are proximity effects in an SEM image. These arise from competing effects. On the one hand, slow electrons that leave one line and might have escaped to the detector may be captured by a nearby neighboring line. This shadowing effect tends to decrease the detected signal. On the other hand, fast backscattered electrons that leave the first line and hit the second may generate additional secondary electrons that would not otherwise have existed, tending to increase the signal. For this reason, modeling a dense array of lines is a different problem from modeling an isolated line. An MBL result for such an array of polycrystalline Si lines is shown overlaid on a cross-sectional image in Fig. 6. There are two kinds of edges in the sample. *Interior* edges face a nearby neighboring edge across a trench. The first and last edge in the pattern are *exterior* edges, which do not have nearby neighbors. Both these kinds of edges were modeled and included in the library, so edge separation is a library parameter. Unlike the other library parameters, we know in advance which edges are interior and which are exterior, so the separation for each edge was pinned to the appropriate value. Pinning is not mandatory of course. The separation could have been treated as a free parameter like all the others. However, as we will discuss later, limiting fitting parameters to the minimum number necessary is a good practice. There are five lines in the image, therefore five differences between the cross-section and MBL linewidth determinations. The average of these differences was 0.7 nm, indicating relatively little systematic bias between the measurements. The standard deviation of the differences was 3 nm, an amount of random error consistent with the linewidth roughness of these lines.

A similar comparison was also performed[14] for lines in Sumitomo PAR-810 photoresist.[†] Resist is an important target in the semiconductor industry, but is potentially problematic inasmuch as it is an insulator that may charge during imaging, and it shrinks when exposed to an electron beam. Results of this comparison are shown in Fig. 7. Top-down images were performed first. Then the specimens were cross sectioned and imaged. The MBL widths averaged 3.5 nm larger than the cross-section widths. If the lines shrink during e-beam exposure, the MBL widths should be larger, since the cross-sectioned specimen has received electron dose twice (once each during the top-down and cross-section images). The magnitude of the difference is consistent with previously observed shrinkage rates.[15,16] The methods had a random difference with standard deviation of 5 nm, 3 nm of which is expected from the observed linewidth roughness.

Repeatability

The tightest CD tolerances in the ITRS, now approaching 0.5 nm, are not for bias, but for measurement tool

[†]Certain commercial equipment or materials are identified in this report in order to describe the experimental and analytical procedures adequately. Such identification does not imply recommendation or endorsement by NIST nor does it imply that the items identified are necessarily the best available for the purpose.

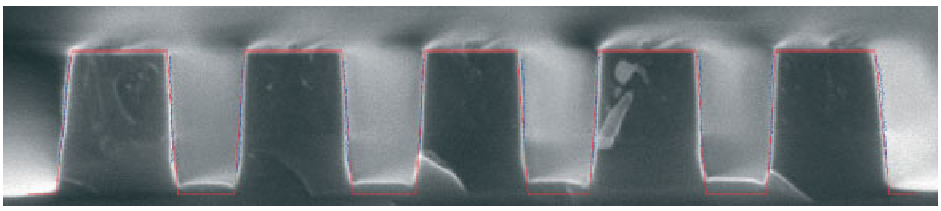

Figure 6. MBL result (smooth trapezoids) for a dense array of lines, superimposed upon a cross-section image.

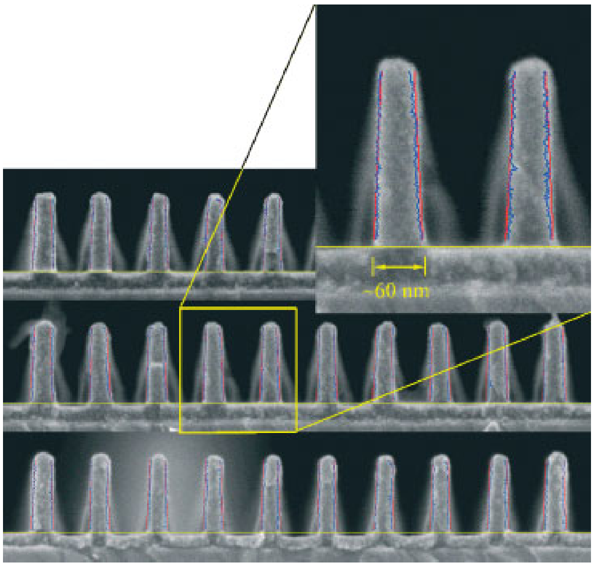

Figure 7. MBL results on an insulating photoresist surface compared to cross section. The lower, middle, and upper arrays are examples of the range of match quality that was observed. They correspond from bottom to top in 10th, 50th, and 90th percentile in χ^2. The inset shows an enlargement of the median match. From Ref. 14.

repeatability. Repeatability (or precision as it is known in the industry) is assessed by performing multiple measurements on the same specimen and then calculating 3σ, where σ is the standard deviation of the measurements. This assesses the sensitivity of the measurement to a limited number of random error sources, mainly image noise and, if the specimen is reloaded between measurements, pattern recognition to reacquire the correct measurement target. As we noted in Fig. 2 and the accompanying discussion, it does not assess other sources of random error that may be relevant in an actual process control application. However, in the absence of an easy way to assess accuracy, much of the emphasis of CD-SEM development has been upon measurement repeatability. *A priori*, there is a good chance that an MBL measurement algorithm will have better noise immunity than the image-processing edge assignment methods currently employed. This is because the current methods use a relatively small part of the data within a linescan in order to assign the edge. The regression to baseline method illustrated in Fig. 1(b), for example, fits a tangent line to the linear part of the outside edge of the bloom peak. Since the MBL approach fits the entire peak, it includes a larger number of data points to make the edge assignment, in effect performing more averaging within the linescan.

This was tested by taking four repeated images of a polycrystalline Si specimen. One of these is shown in Fig. 8. This data set was analyzed two ways, once with regression to baseline and once by using MBL to determine the bottom width. Edge assignments are superimposed on the image. The resulting widths are shown in Fig. 9. There are two groups of curves. The upper four are the widths determined by regression to baseline. The lower set is the MBL group. Some of the width variation within each curve is actual linewidth roughness that can be seen in Fig. 8 and some is

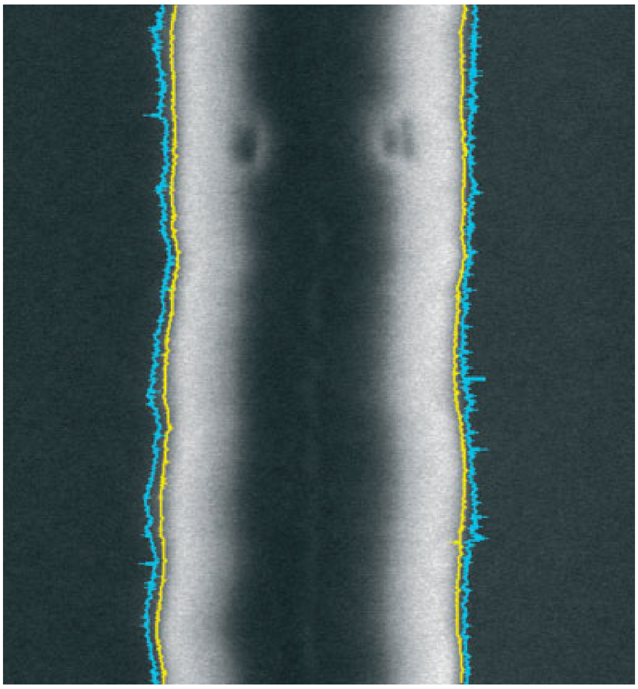

Figure 8. Example of one of four images of the same area that were used to measure repeatability. Edge assignments are shown. The inner smoother curves are assigned by MBL. The outer edge assignments are by regression to baseline.

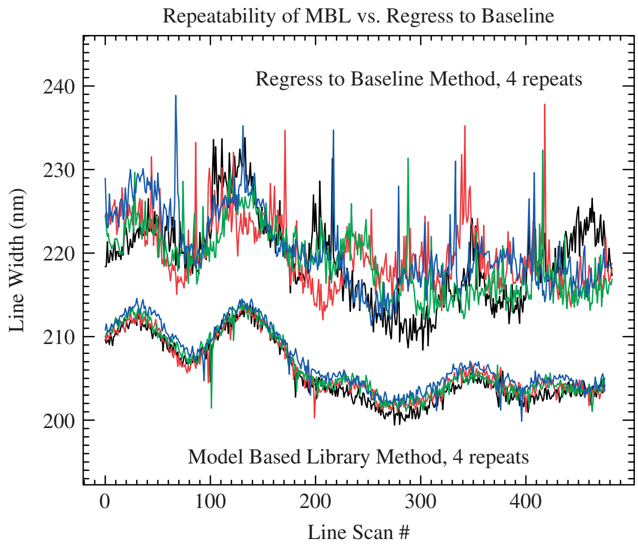

Figure 9. Repeatability of two edge assignment methods. The curves are linewidth as a function of position along the line. The upper group had edges assigned by regression to baseline. The lower group shows bottom widths assigned by MBL.

edge assignment error due to noise. It is evident in the figure that there is more curve-to-curve scatter in the upper set than in the MBL group. This difference was quantified as follows: Curves 2 through 4 within each group in Fig. 9 were shifted left or right as required for maximum correlation with curve 1. In this way, any small drift between images was removed. At the first position along the line, the standard deviation of the four width values was determined. This standard deviation was then averaged with the standard deviations determined at each of the other line positions. The standard deviation so determined for the regression to baseline curves was 8.1 nm. For the MBL curves it was 2.4 nm. Incidentally, the average separation between the two groups is about 15 nm. In light of the agreement between MBL and cross sections, we interpret this to mean that for this line the regression method has about 15 nm of bias.

DISCUSSION

It should be noted that what we have loosely been calling 'inversion by least squares' is not guaranteed to be unique. The instrument function modeled by Eqn (1) is, in the strictest mathematical sense, not invertible. Such functions are usually many-to-one mappings; that is, there is more than one specimen configuration that can produce the same image. A trivial example would be a specimen with a buried void at a depth not reached by the electron beam and that, therefore, does not influence the image. Other specimen configurations, though distinguishable in principle, may produce images that are so close to one another that they are difficult to distinguish in the presence of noise. Undercut sidewalls (in which the bottom of the line is narrower than the top) are difficult to distinguish from each other or from vertical ones, for example. Since the typically low-energy incident electrons do not penetrate, the only thing to distinguish one undercut specimen from another is the signal that derives from electrons backscattered from the substrate. These create secondary electrons when they strike the (possibly undercut) sidewall. Such images change only slightly when the magnitude of the undercut changes, so sensitivity is low. Nonuniqueness of solutions is not so much an issue with the analysis method as with the amount of information available in the image, regardless of how it is analyzed. Fortunately, the typical application for CD measurements is in manufactured structures with a well-defined target geometry and with relatively small process variation. Only candidate solutions that span the range of reasonable process variation need to be considered. In the MBL approach, this is what determines the region of parameter space that must be spanned by the library. When this restricted solution set can be characterized by a suitably small number of parameters, the solution can be expected to be unique. If the number of parameters becomes large for the information in the image, one begins to notice correlations among sets of parameters such that different combinations yield similar images.[11] For a linescan across a line feature like the one in Fig. 1(a), which has a bloom peak for each of two edges, seven parameters are easily justified. These are background levels common to both edges, and three parameters for each bloom peak. The three parameters needed to describe a peak are the peak position, width, and height above background. In the SEM, these seem to be closely related to specimen edge position, beam size, and sidewall angle, although the relationship is not 1 : 1. (Beam size and sidewall angle each affect both peak height and width, but in different ways such that it is possible to distinguish them.) Beyond these easy parameters, if the signal-to-noise ratio is good enough, it may be possible to justify one or more additional ones, like corner radii, that affect peak symmetry, but sensitivity for these is generally lower because their effect on the image is more subtle. To minimize parameter correlations, it is important to avoid unnecessary parameters. For example, instrument parameters (s, y_0, and the b_i in Eqn (1)) are determined from the image as a whole, rather than allowed to vary from linescan to linescan. Because independent sets of parameters can be derived from parts of the specimen separated by more than the spatial resolution of the instrument (a few nanometers), the number of parameters that can be validly assigned to a sample area can be very large. In the rare cases when there are multiple solutions within the parameter subspace for a particular manufacturing process, the MBL approach offers the opportunity to discover this fact by checking for more than one local minimum within the library. Then the result can be reported with a suitably increased uncertainty. If that uncertainty is too large, it may be possible to resolve the ambiguity by finding imaging conditions under which the candidate structures image differently. There is no reason in principle why the MBL approach cannot be extended to stereo pairs of images (i.e. pairs with different incidence angles), for example.

Acknowledgements

The staff of the failure analysis lab and Benjamin Bunday, both from International SEMATECH, supplied us with cross-section and CD-SEM images for Fig. 7 and Fig. 8.

This work was funded partly by SEMATECH and the NIST Office of Microelectronics Programs.

REFERENCES

1. Semiconductor Industry Association, *International Technology Roadmap for Semiconductors* (2003 edn), SEMATECH: Austin, TX, 2003; Table 117a, p. 10, Metrology Section. Also available at http://public.itrs.net/Files/2003ITRS/Metrology2003.pdf.
2. Ausschnitt CP, Lagus ME. *Proc. SPIE* 1998; **3332**: 212.
3. U.S. Census Bureau, *Semiconductors, Printed Circuit Boards, Other Electronic Components* (2003), U.S. Census Bureau: Washington DC; Table 2, p. 2, From MA334Q: Current Industrial Reports. Also available at http://www.census.gov/industry/1/ma334q03.pdf. Quoted data is for microprocessors having an internal data bus of 32 bits or more shipped in 2003.
4. Villarrubia JS, Vladár AE, Postek MT. *Proc. SPIE* 2003; **5038**: 138.
5. Ukraintsev VA. *Proc. SPIE* 2003; **5038**: 644.
6. Lowney JR, Vladár AE, Postek MT. *Proc. SPIE* 1996; **2725**: 515; Lowney JR. *Scanning Microsc.* 1996; **10**: 667.
7. Marquardt DW. *J. Soc. Ind. Appl. Math.* 1963; **11**: 431.
8. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. *Numerical Recipes in C*. Cambridge University Press: Cambridge, 1988.
9. Bevington PR, Robinson DK. *Data Reduction and Error Analysis for the Physical Sciences* (2nd edn). McGraw-Hill: New York, 1992.
10. Davidson MP, Vladár AE. *Proc. SPIE* 1999; **3677**: 640.
11. Villarrubia JS, Vladár AE, Lowney JR, Postek MT. *Proc. SPIE* 2001; **4344**: 147.
12. Villarrubia JS, Vladár AE, Lowney JR, Postek MT. *Proc. SPIE* 2002; **4689**: 304.
13. Villarrubia JS. *J. Res. Natl. Inst. Stand. Technol.* 1997; **102**: 425. Also available at http://nvl.nist.gov/pub/nistpubs/jres/102/4/j24vil.pdf.
14. Villarrubia JS, Vladár AE, Bunday BD, Bishop M. *Proc. SPIE* 2004; **5375**: 199.
15. Habermas A, Hong D, Ross M, Livesay W. *Proc. SPIE* 2002; **4689**: 92.
16. Sullivan N, Dixson R, Bunday B, Mastovich M, Knutrud P, Fabre P, Brandom R. *Proc. SPIE* 2003; **5038**: 483.