Analysis of Line-edge Roughness in Resist Patterns and Its Transferability as Origins of Device Performance Degradation and Variation

Atsuko Yamaguchi\textsuperscript{a}, Hiroshi Fukuda\textsuperscript{a}, Hiroki Kawada\textsuperscript{b}, and Takashi Iizumi\textsuperscript{b}

\textsuperscript{a}Central Research Laboratory, Hitachi, Ltd. Higashikoigakubo, Kokubunji, Tokyo 185-8601, Japan
\textsuperscript{b}Hitachi High-Technologies Corporation, Ichige 882, Hitachinaka-shi, Ibaraki 312-8504, Japan

General property of line-edge roughness (LER) in resist pattern and its transferability to under-lying layer were investigated. Longer-period components were found to have larger amplitudes in a resist pattern and to remain after dry etching. In addition, long-period LER strongly affects a transistor performance. Long-period LER in resist patterns, therefore, is as important as short-period LER. Metrology of LER was reconsidered to evaluate the both LER properly, and a guideline for choosing measurement parameters was proposed from a viewpoint of device performance estimation.

**Keywords:** Line-edge Roughness, Dry Etching, CD-SEM, Measurement Parameter, Device Performance

1. Introduction

The influence of line-edge roughness (LER) on MOSFET performance is becoming a serious problem as shrinking gate length toward sub-100-nm technology node [1-5]. For example, LER of 5 nm in both edges of a 65-nm-gate, causes more than 10\% fluctuation in gate length. Measurement and control of LER are important issues as addressed in the ITRS roadmap.

Figure 1 explains two types of LER, intra-transistor LER and inter-transistor LER. Intra-transistor LER, i.e., LER with shorter spatial period than a transistor width \(w_x\), causes gate-length fluctuation and degrades device performances of the transistor. On the other hand, inter-transistor LER, i.e., LER with a longer spatial period than \(w_x\), causes variation in the averaged gate-length (\(L_e\)) of the transistor, resulting in variation in device performance. The effects of these two types of LER should be discussed separately. Figure 2 shows an example of calculated threshold-voltage (\(V_{th}\)) shift of a transistor. Even if \(L_e\) equals the designed size, \(V_{th}\) drops 20 mV (point P in Fig. 2) from the designed value (point Q in Fig. 2) when the degree of intra-transistor LER, 3\(\sigma\), is 15 nm. On the other hand, transistor performance is also affected by \(L_x\).

![Fig. 1 Effect of LER on device performance](image)

Solid curve in Fig. 2 shows that \(V_{th}\) drops more than 30 mV (point R in Fig. 2) when \(L_e\) deviates from the designed size by 5 nm even if 3\(\sigma\) equals zero. Inter-transistor LER is the origin of the \(L_e\) deviation, and causes variation in device performance.

Received April 14, 2003
Accepted May 28, 2003
performance. Therefore, it is necessary to examine the property of long-period LER; however, short-period LER has been mainly investigated so far.

![Graph showing calculated degradation of device performance.](image)

Fig. 2 Calculated degradation of device performance.

In this work, characteristics of both types of LER and their transferability by LSI process were investigated. Changes in LER by SiO₂ dry etching and polycrystalline silicon (poly-Si) etching were measured. Furthermore, metrology for LER was discussed. To evaluate short- and long-period components of LER, measurement parameters, such as inspection-area length, measurement-point interval, and number of measurements were examined.

2. Spatial-frequency distribution

First, we investigated the general characteristics of LER in various resist materials, one electron beam (EB) resist, one KrF resist, and six ArF resists. From top-down scanning electron microscope (SEM) images of line patterns, an in-house image-processing program extracts a set of 512 edge points along a line edge and calculates line-edge fluctuation \(|\Delta x_i|, i=0,1,2 \cdots\). In this procedure, edge points are calculated from line-profile data of an SEM image by the threshold method (threshold value is 50%).

The spatial period of LER was analyzed by Fourier transform of the fluctuation data. Typical spatial-frequency (f) distributions are shown in Figs. 3-5. In the high frequency region (Fig. 3), there are two types of frequency distribution. In the first type [Fig. 3(a)], amplitude is proportional to \(1/f\) in the whole range of spatial frequency. The spectra for the EB, KrF, and three of the ArF resists were this type. In the second type of distribution [Fig. 3(b)], amplitude is proportional to \(1/f\) in the high-frequency region, but a peak appears at \(f\) of about 10 \(\mu\text{m}^{-1}\) (spatial period \(T\) is about 100 nm). This peak, an additional component, often appeared in spectra of three of the ArF resists. We speculate that this peak is related to resist material (not to LER in the mask pattern or exposure tools) because the peak frequency does not depend on exposure tools, lithography magnification, or SEM magnification [6]. From the results obtained from 2.7-\(\mu\text{m}\)-long lines (Fig. 4), it is clear that the \(1/f\)-rule cannot describe the spatial-frequency distribution when \(f\) is less than 1 \(\mu\text{m}^{-1}\). Analysis
of much longer pattern edges showed that spatial-frequency distribution of LER is almost constant in low frequency region (Fig. 5).

![Figure 5: Spatial-frequency distribution of LER at low frequency.](image)

**Fig. 5** Spatial-frequency distribution of LER at low frequency.

![Figure 6: General feature of spatial-frequency distribution of LER.](image)

**Fig. 6** General feature of spatial-frequency distribution of LER.

According to these results, schematic spatial-frequency distribution for general LER is given as shown in Fig. 6. (In some cases of LER in ArF resist patterns, the additional component appears at around 10 µm⁻¹.) It should be noticed that the longer-period component has the larger amplitude. Again, this general characteristics are also related resist material, because it does not depend on exposure tools, lithography magnification, or SEM magnification. Results of simple Monte-Carlo simulation also suggested that the 1/f-nature in shorter-period region is based on randomness in resist reactions [7]. For metrology optimization, it is sometimes convenient to approximate the above general spatial-frequency distribution as follows,

\[ A(f) = A_0 \cdot f^{-1}, \]
where \( f > 10 \) µm⁻¹ and

\[ A(f) = \frac{A_0}{f_0} \cdot f^{-n} \]

\[ n = 1 - \log_{10} f_0 \quad (1 \mu m^{-1} < f \leq 10 \mu m^{-1}) \]
\[ 0.25 \quad (f \leq 1 \mu m^{-1}), \]

where \( f \leq 10 \) µm⁻¹. The value of \( f_0 \) was found to be 3-5 by experiment.

3. Transferability of LER

Various processes in LSI manufacturing affect the gate-length distribution. Here we consider gate-delineating process in which gate patterns in the resist layer are transferred onto the SiO₂ layer, and the SiO₂ patterns are transferred onto poly-Si layer. Figures 7(a), (b), and (c) are top-down images of resist patterns, SiO₂ patterns, and poly-Si patterns. The height of the images is 1350 nm. The observed patterns are at exactly the same position on the wafer after each process. The resist pattern [Fig. 7(a)] was delineated by electron-beam writing with a commercially available EB resist. The beam was intentionally blurred (defocused) in order to introduce LER. Figures 8(a), (b), and (c) show detected edge points corresponding to the images in Fig. 7. Although the short-period LER was reduced by the dry-etching processes, long-period structure is almost unchanged.

Spatial-frequency (f) distributions, i.e., f-A(f) graphs of the eight edges in each SEM image were calculated, and their geometric mean of eight values of A(f) was obtained. Figure 9 shows the mean spatial-frequency distributions for the LER data shown in Fig. 8. Higher frequency components are slightly reduced by SiO₂ dry etching. Analysis on more than forty edges before and after SiO₂ etching confirmed a decrease in LER components for a spatial period of 30-70 nm. Looking from the whole inspection area longer than 1 µm, contribution of this spatial range to the total LER is small (because of 1/f nature), and the SiO₂ dry etching process reduced 3σ of LER by only 3-4%. Further analysis on twenty edges showed that subsequent poly-Si dry etching process reduced LER by about 7%. In total, LER (3σ) in the inspection area of about 1 µm was reduced only about 10% through all the dry-etching processes for gate delineation. This means that the most dominant factor that determined LER in the gate pattern is in resist materials and lithography process. It should be noticed that long-period components of LER, which have large amplitudes in the spatial-frequency distribution as shown in Sec. 2, remains after dry etching processes. Because this long-period component causes variation in device characteristics or degrades device-performance, origin of the long-period LER in resist pattern needs to be identified to reduce gate-length fluctuation. Also, to evaluate or to monitor the influence on device
Fig. 7 SEM images before and after SiO$_2$ and poly-Si dry etching.

(a) EB resist patterns. (b) SiO$_2$ patterns. (c) Poly-Si patterns. Images (a) and (b) are observed by S8820 (Hitachi High-technologies) with 800 V of high-voltage. Image (c) is observed by S9260 (Hitachi High-technologies) with 300 V of high-voltage.

Fig. 8 Detected edges of the SEM images before and after SiO$_2$ and poly-Si dry etching. Calculated from the images shown in Fig. 7 with our image-processing program.

Fig. 9 Geometric mean of spatial-frequency distributions.

Performances of long-period LER as well as that of short-period LER, metrology of LER needs to be reconsidered carefully.

We investigated the relation between noise intensity in spatial-frequency distribution and number of scanning repetition, and it was confirmed that the noise appears in the shorter period range than 25 nm (i.e., 40 μm$^{-1}$). Therefore, the observed decrease in LER in dry etching processes is not caused by noise reduction.

As an impression directly obtained from the SEM images [Figs. 7(a) and (b)] and the viewgraphs of detected edges [Figs. 8(a) and (b)],

Fig. 10 Distributions of edge-area widths.
(a) resist pattern, (b) SiO$_2$ pattern.
Distribution of this ‘white-band width’ obtained from the SEM images (Fig. 10) showed that the width of the white band in the resist pattern [Fig. 10(a)] has a wider distribution than that in the SiO₂ pattern [Fig. 10(b)] (the standard deviations of 3.2 nm and 1.9 nm, respectively). We speculate that the distribution is due to varying three-dimensional structure (surface roughness of resist sidewall) along the edge. The fact that the detected edge-points in the etched pattern almost reflect those in the resist pattern means that the detected edge-points of resist patterns well represent final gate pattern edges.

4. Measurement of LER

4.1 Inspection-area length

In this section, inspection-area length is examined to measure CD (corresponds to the averaged gate-length \( L_0 \)) and \( 3\sigma \) for a single transistor.

The value of \( A(f) \) has a tendency to saturate as the value of \( f \) decreases, but it does not reduce to zero. This means that measured LER increases with measurement length \( L \) along the gate-pattern edge (i.e., inspection-area length). Dependencies of CD and LER (3σ) on length \( L \) are evaluated from experimental data by the following procedure. First, a 2.7-µm-long line feature is segmented into short lines with segmentation length \( L \) of 2.7/2^\( n \) µm, where \( n=1,2,3,\ldots \) consequently, \( 2^\( n \) \) short lengths are made. Second, CD and LER (3σ) of the short lengths are calculated for each segmentation length \( L \).

![Graph showing CD and 3σ dependence on L](image)

Fig. 11 Inspection-area dependence of CD and 3σ.

The obtained \( L \)-dependence of CD and 3σ are shown in Figs. 11(a) and (b). The values of CD and 3σ heavily depend on \( L \), and it is clear that measurement length \( L \) should be identified when measuring LER. Comparing LER results measured with different measurement length \( L \) has no meaning. The solid curve in Fig. 11(a) was calculated statistically according to the assumption that the width data are randomly distributed. However, this assumption is not satisfied because there is a strong correlation between neighboring data. In fact, experimental results [shown in Fig. 11(a)] cannot be described with the calculated curve in the small-\( L \) region. The results also show that 3σ of a short length deviates from the value obtained from the original line, especially when \( L \) is small. Both CD and 3σ seem to saturate when \( L \) is larger than 1 µm. However, this saturation is partially due to the fact that the length of the original long line was 2.7 µm. It is still uncertain whether the values would be constant at larger \( L \). In conclusion, therefore, it is recommended to evaluate LER in a line whose length is equal to the width of the investigated transistor when estimating the influence of LER in a gate pattern on transistor performance.

4.2 Measurement-point interval

In this section, measurement-point interval for accurate edge-point detection is discussed. The effect of measurement-point interval on the accuracy of LER measurement was investigated as follows. The measured original data set was a line-width fluctuation \( \{w_i|i=1,2,\ldots,N\} \) in a line of length \( L \). The standard deviation of this original data was defined as the ideal value of the standard deviation, \( \sigma_0 \). From the original data series, \( \{w_i|i=1,2,\ldots,N\} \), the values were selected with an interval \( M \) (data number). The interval in length \( \Delta y \) is given as the product of 1.318 and \( M \). The standard deviation of the data, \( \sigma(\Delta y) \), was calculated from the selected data group. Error \( p \) is given as

\[
p = \frac{\sigma(\Delta y)}{\sigma_0} - 1. \tag{3}
\]

Figures 12(a), (b), and (c) show the \( \Delta y \) dependence of \( p \) for various \( L \) values obtained from measurement on two lines.

Solid curves in the graphs were calculated as follows. The results of the spatial-frequency analysis showed that the component \( A/|L| \) has the largest contribution to LER in a line of length \( L \). When detecting edge points with an interval of \( \Delta y \), components whose frequency is higher than \( 1/\Delta y \)
cannot be detected. The contribution of this undetectable component to LER is about $A(1/\Delta y)$. Therefore, measurement of LER (such as 3σ) by edge-point detection with an interval of $\Delta y$ may include the deviation, whose ratio $p$ is roughly estimated as $A(1/\Delta y)A(1/L)$, from the ideal value. By using eq. (1) and (2), curves in Figs. 12 were obtained.

Furthermore, $\Delta y_{\text{max}}$, the maximum measurement-point interval for the accuracy $p$, can be obtained. To meet accuracy 5% or less, the value of $\Delta y$ where $p$ equals 5% was read from the graphs. The results are plotted in Fig. 13. Solid curves in this figure show the calculated values (The value of $f_o$ was set to 5 $\mu$m$^{-1}$). Results obtained by the $\Delta y_{\text{max}}$-calculation show good agreement with the results obtained by measurements.

The $\Delta y_{\text{max}}$, calculation method mentioned above was based on the $1/f$-distribution [Fig. 3(a)]. Next, we discuss the case that the spatial-frequency distribution has the additional component at around 10 $\mu$m$^{-1}$ [Fig. 3(b)]. First, the case that the value $L$ is near around 10 $\mu$m$^{-1}$ is considered. In this case, the real $A(1/L)$ is larger than the calculated value by the schematic frequency distribution (shown in Fig. 6); consequently, the real value of $p$ becomes smaller than the value calculated as $A(1/\Delta y)A(1/L)$. Therefore, the $\Delta y_{\text{max}}$-calculation method is valid in this case. The second case is that the $\Delta y$ is in the additional-component region. In this case, the undetectable component becomes larger than that calculated with the $\Delta y_{\text{max}}$-calculation method, and the real error may exceed the required $p$ value. The additional component is distributed around 50 to 200 nm, and it was previously found by auto-correlation function analysis that the specific spatial-period of the LER feature appears at 40-300 nm [6]. Therefore, $\Delta y$ should be set below 40 nm even if the calculated value for $\Delta y_{\text{max}}$ is larger than 40 nm. However, the value of $L$ that gives the calculated value of $\Delta y_{\text{max}}$ as 40 nm is 256 $\mu$m. If it is estimated from measured $\Delta y_{\text{max}}$, it becomes dozens of microns, which is unrealistic for normal transistor widths. It can thus be said that the $\Delta y_{\text{max}}$-calculation method shown above is valid even in the case that the spatial-frequency distribution has the additional component.

The number of measurement points required to meet an error target of 5% or less is plotted in Fig. 14. It was calculated as $L$ divided by $\Delta y_{\text{max}}$. To evaluate LER from one SEM image of a 1-$\mu$m-long line, 100 points (at least 70 points, estimated from the measured results) are needed.
4.3 Number of measurements
To obtain variation in transistor performance (caused by inter-transistor LER), measurement of CD and $3\sigma$ should be performed several times. As shown in Figs. 11(a) and (b), both CD and $3\sigma$ values are distributed widely at small $L$ (corresponding to the case of small $w_i$), however, converge at $L$ larger than 1 $\mu$m. (This is partially due to the fact that the length of the original long line was 2.7 $\mu$m). For a relatively short line (edge length smaller than 0.5 $\mu$m), in our experience, reasonable distribution data can be obtained when total summation of inspection-area length becomes longer than 2 $\mu$m. For a longer edge length, more than five measurements are needed.

Thus a guideline for LER measurement is obtained. For example, in estimating $3\sigma$ and $L_{ex}$ for a 200-nm-wide transistor pattern, inspection-area length $L$ must be 200 nm and 2-$\mu$m-long area must be measured in total. The data for a long line is divided into ten data sets, and ten values of $3\sigma$ and are obtained. Device performance degradation for each transistor can be calculated from the values of $3\sigma$ and $L_{ex}$, and furthermore, variation of device performance can be obtained from the ten data sets. Also, a typical value of LER in the resist pattern is given as the average (root mean square) of the ten values of $3\sigma$. According to the method described in Sec. 4.2, the required value of $\Delta y_{max}$ for $p$ of 5% is about 6.1 nm.

5. Summary
Spatial-frequency distribution of LER was investigated and the longer-period components were found to have the larger amplitudes. From the analysis on LER transferability by SiO$_2$ and poly-Si dry etching processes, it was found that the long-period LER in resist patterns were transferred to the underlying layer by etching. Considering the device performance calculation results, it is concluded that long-period LER should be evaluated properly as well as short-period LER. A guideline for accurate LER measurement was proposed from the viewpoint of device-performance estimation. Sufficient measurement length, proper measurement-point interval, and number of measurements are essential for accurately evaluating the both long- and short-period LER.

Acknowledgments
The authors are grateful to Osamu Komuro for his advice on edge detection algorithm, Ryuta Tsuchiya for calculating I-V curves of ideal devices. The authors also wish to acknowledge Takuya Hagiwara, Isao Uchida, Yoshie Azuma, Toshiyuki Yamamoto, Takeshi Kuriyagawa, and Yasushi Goto for supplying the sample wafers. The authors are also grateful to Chie Shishido, Yuji Takagi, Makoto Ezumi, and Shozo Yoneda, and Drs. Toshihiko Tanaka, Tsuneo Terasawa, and Minoru Toriumi for their helpful comments.

References