

Multi-Scale Analysis of Line Edge Roughness Based on Wavelet Transform

Fei Wang,* Ning Li*

* School of Mechanical Engineering, Harbin Institute of Technology, Harbin, P. R. China

The problem of quantifying line edge roughness (LER) in the semiconductor industry has become critical with the rapid development of the microelectronic manufacturing technology [1]. Atomic force microscope (AFM) has better sensitivity in characterizing the line structure topography and has a 3D imaging ability, which is very useful to analyze roughness information at different height positions. So it can be used to quantify LER by taking advantage of the high spatial resolution approaching the atomic scale. Here, we introduce a multi-scale analysis method based on wavelet transform to deal with the experiment data.

Wavelet transform gives a novel tool to observe the measured sample in different scale: larger scale can reflect the approximate original signal, while smaller scale gives detail of original signal. For 1D continuous time signals, the continuous time wavelet transform is given by

$$WT_f(a, \tau) = f^* \psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \int_R f(t) \psi^* \left(\frac{t-\tau}{a} \right) dt \quad (1)$$

where a and τ represents the scale index and the position shift. The position-scale parameter can be discretized using the dyadic grid sampling. Wavelet analysis can divide a function into different frequency components with each component having a resolution, which is matched to its scale. For the decomposed level is J in partial discrete wavelet transform the multi-scale analysis of series X and the energy decomposition can be obtained:

$$X = \sum_{j=1}^J D_j + S_J, \quad \|X\|^2 = \sum_{j=1}^J \|D_j\|^2 + \|S_J\|^2 \quad (2)$$

where D_j is the detail content at scale $a_j=2^{j-1}$, S_j is the approximation content at scale 2^j . The definition of scale is a standardization quantity without unit. In the practice application the sampling must be considered, that is to say, the standard scale a_j is comparable to the physical scale $a_j \cdot \Delta$. The variance of X is defined by:

$$\sigma_x^2 = \frac{1}{N} \sum_{t=0}^{N-1} (X_t - \bar{X})^2 = \frac{1}{N} \|X\|^2 - \bar{X}^2 \quad (3)$$

where \bar{X} is sample mean value, N is the number of sample points. The relationship between the decomposed coefficients and the amplitude parameter satisfies the following equation [2]:

$$\sigma_x^2 = \frac{1}{N} \sum_{j=1}^J \|D_j\|^2 + \frac{1}{N} \|S_J\|^2 - \bar{X}^2 \quad (4)$$

The LER series can be decomposed into the detail content at each scale and the approximation content. The contributions of line edge roughness at each scale can be estimated through the analysis of the signal energy at each scale. Thus the amplitude variance of LER in transverse direction and the spatial frequency of sample points in lengthways direction are together used in the LER characterization. The energy coefficient e is defined by:

$$e_j = \frac{1}{N} \|D_j\|^2, \quad j = 1, 2, \dots, J \quad (5)$$

Hence, the energy distribution of LER can be characterized by calculating the energy coefficients based on multi-scale analysis. When the decomposed level is J , the approximation content contains sufficient information about large scale fluctuation, while the high frequency components which is

not be neglected for contributions to LER are decomposed into the corresponding detail content at the level $j(j=1, 2, \dots, J)$. The detail content above level J is believed as putting together into the approximate content at the level J due to their little influence on LER. The characteristic length l determines the correlation range of LER feature, which can be estimated using the physical scale $a_j \Delta$ (nm) at the decomposed level J . That means the signal at the decomposed level J is enough approximate to represent the original signal without higher frequency components beyond our care about. Therefore, besides the amplitude parameter σ , the multi-scale analysis method of LER can provide the energy coefficient e and the characteristic length l to indicate the spatial property of LER.

Applying the multi-scale analysis based on wavelet transform to the LER data in Fig.1, we can make LER into the sum of S_j and D_i ($j=1, 2, \dots, J$). An overall view of the structures involved in the LER can be observed in Fig.2. The approximations up to level 4 give the general trend of the LER, and the irregular structures above level 4 can be ignored because it has little impact on LER. So the characteristic length l is about 44.56nm, which means the edge points can be considered uncorrelated when their distance is longer than $l/3$ measuring data sets with the same scan length (2854nm) are selected for calculating the energy coefficients e_j at each scale, the result is shown in Fig.3. Obviously the energy of the detail content is mostly in the scale 2 and 3, that is to say, the spatial frequency components at the physical scale of (11.2-22.3)nm have large contributions to LER.

References

[1] E. Gogolides, V. Constantoudis, G. P. Patsis, et al. Microelectronic Engineering. 83 (2006) 1067.
 [2] D. B. Percival, T. Walden. Wavelet Methods for Time Series Analysis, Beijing, (2006) China Machine Press:104-135.

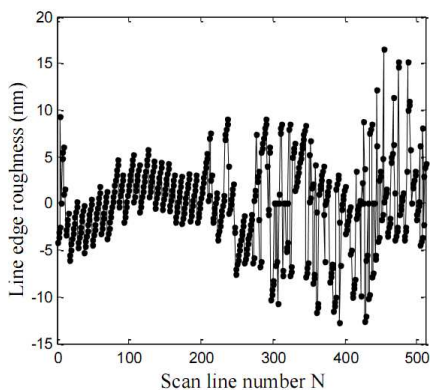


FIG. 1 Extracted edge roughness features of left line of one sample from AFM image

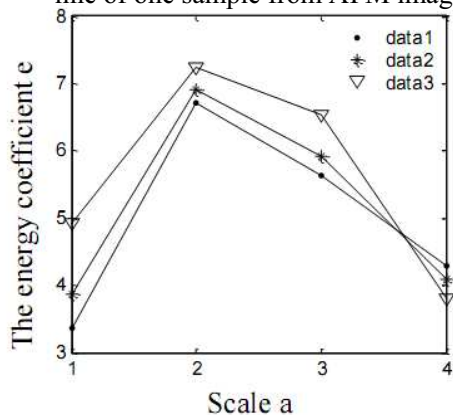


FIG. 3. The distribution of energy coefficients e_j at the scale j ($j=1,2,3,4$) of 3 measuring data sets

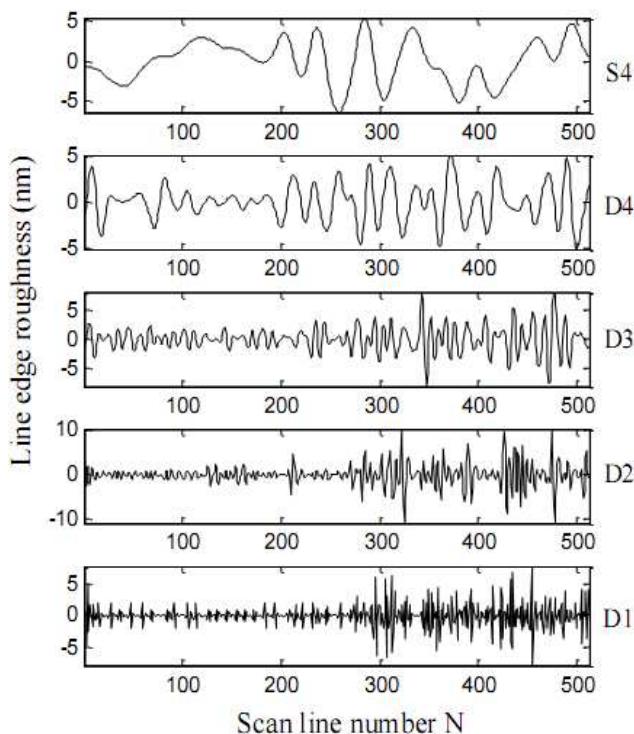


FIG. 2. The multi-scale analysis of LER based on wavelet transform series