Theoretical and Experimental Study of Ultra-smooth Surface Based on Gaussian Topography

YUMING ZHU

ABSTRACT

All natural surface heights conform to the law of Gaussian distribution. Thanks to this law, researchers can measure the topography of different surfaces. To function, modern precision optical instruments need extremely smooth surfaces whose RMS is usually controlled under 0.5nm. A Surface that meets this condition is called ‘Ultra smooth Surface.’ Researchers normally use X-ray reflectance test, Atomic Force Microscope (AFM) scanning and Contour scanning when they characterize smooth surfaces. However, not all the optimal conditions of the research instruments can be met in real experiments. Hence, when scientists want to measure ultra-smooth surfaces, results of these methods are often limited. Furthermore, important statistics such as the fractal index and the lateral correlation length cannot be obtained, and redesigning the current measurement devices doesn’t help much to increase accuracy.

X-ray scattering not only empowers scientists with scattering models that determine the surface roughness of a given object, but also offers a PSD curve that functions on a wider spectrum. In this experiment, I put two 15mm*15mm silicon chips on the X-ray scattering device, used the glancing incidence of 0.1 degrees and 0.15 degrees respectively, and analyzed the data. Data from the Detecting Scan is analyzed by the First Order Vector Perturbation theory, and a Power Spectrum Density (PSD) is obtained. Adding on to that the Global Optimization Algorithm—Simulate Anneal Arithmetic, the PSD function is fully fitted, Thus, the surface roughness sigma, the fractal index h, and the lateral correlation length, a, is obtained. Comparing the PSD function of X-ray scattering with results obtained by the Atomic Force Microscope and Contour scanning, this project proved the validity and strength of using the new method of X-ray scattering to measure surface roughness.

INTRODUCTION

The methods to characterize ultra-smooth surfaces are vital for modern optics. The stability and isotropy of silicon chip makes it easy to be experimented, and thus an optimal choice for characterization. The approximate average surface roughness of ultra-smooth silicon chips is 0.1nm to 1nm, too small that it cannot be measured by visible-light interference. The only method that satisfies high sensitivity, high accuracy and large view is using short-wave X-rays that falls into the same order of magnitude with surface height.

Yuming Zhu, Shanghai Foreign Language School, Shanghai, China.
The research project uses specialized copper Kα X-ray that has a wavelength of \(\lambda=0.154\text{nm}\). Since the subject is surface roughness, this project selects the method of X-ray scattering, and validates the result with AFM and contour scanning.

Since the X-ray has a short wavelength and carries high energy, the reflectance of X-ray is low. Thus, the angle of total reflection is small. At glancing incidence, this experiment collects the information of the electro-magnetic wave resulted from the interference of X-ray and target subject. In this project, I created a statistics model of surface roughness according to Born Assimilation, derived the correlation function of the difference of surface height, and fitted the data to deduce the surface parameters. By using the method of detecting scan and processing the data with Fourier transformation, the correlation between the following statistics: surface roughness \(\sigma_r\), fractal index \(h\), horizontal correlation length \(\xi\) is found.

**THEORIES**

**Fractal Theory**

In 1973, Mandelbrot first proposed the concept of fraction. The character of fraction means that a part of the figure resembles the entire figure. In this project, fraction is used to describe the surface height of an ultra-smooth surface. Since the surface of the silicon has self-similarity, the fraction index is an important parameter of surface topology. The fractal theory provides the analysis part of the experiment with a more accurate result. In 1988, Sinha established the theory of X-ray and neutron scattering. To calculate the X-ray diffuse scattering, a statistic model must be established according to the surface or interface. The correlation function of a single surface or interface is as follows.

\[
g(X, Y) \leq [z(x', y') - z(x, y)]^2 \geq 2\sigma^2 - C(X, Y)
\]  

(1)
\[(X, Y) = (x'-x, y'-y)\]  

\[R = \sqrt{X^2 + Y^2}\]  

\[C(X, Y) = C(R)\]  

\[C(R) = \sigma^2 e^{-\left(\frac{R}{a}\right)^{2h}}\]

The algorithm to analyze the fractal surface is as follows.

\[\zeta_n = \sum_{j=1}^{n} \alpha_{nj} U_j, \quad n = 1, 2, ..., N\]  

\[\alpha_{n1} = \frac{c_{n-1}}{\sigma}, \quad n = 1, 2, ..., N\]  

\[\alpha_{nj} = \frac{1}{\sigma} \left( C_{n-j} - \sum_{i=1}^{j-1} \alpha_{ni} \alpha_{ji} \right), \quad j = 2, ..., n-1, n = 2, ..., N\]  

\[\alpha_{nn} = \sqrt{\sigma^2 - \sum_{i=1}^{n-1} \alpha_{ni}^2}, \quad n = 2, ..., N\]

This essay simulates 2000 points that conforms to a standard distribution with a standard deviation of 1. Figure 2 shows the simulation results of how fractal index (Left: \(\sigma=1\)nm, \(a=20\)μm, \(N=2000\), \(L=200\)μm, \(\Delta X=L/N=10\)nm) influences the surface roughness and how horizontal correlation length (Right: \(\sigma=1\)nm, \(h=0.95\), \(N=2000\), \(L=200\)μm, \(\Delta X=L/N=10\)nm) influences the surface roughness.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma)</td>
<td>surface height</td>
<td>(U_j)</td>
<td>random numbers that conform to standard distribution</td>
</tr>
<tr>
<td>(a)</td>
<td>lateral correlation length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(h)</td>
<td>fractal index</td>
<td>(C(R))</td>
<td>Self-correlation function</td>
</tr>
<tr>
<td>(R)</td>
<td>random distance between points</td>
<td>(\zeta_n)</td>
<td>The height of the (N^{th}) point</td>
</tr>
</tbody>
</table>
X-ray Scattering Theory

Measuring the light distribution of X-ray scattered from rough surfaces, the X-ray scattering method obtains the surface statistics without directly touching the sample. X-ray scattering not only scans fast, but also measures fine details of both the surface and the interior of the sample. In addition, it measures a large surface area. The scattering formula is as follows.

\[
\Phi(\theta, \varphi) = \frac{1}{Q_i} \cdot \frac{dQ_i}{d\Omega} = \frac{1}{\sin \theta_0} \cdot \frac{\left< |A(q)|^2 \right>}{\int d^2 \mathbf{r}} \tag{10}
\]

\[
\bar{A}(q) = e^{\mathcal{R}(q^2) \sigma^2} \int d^2 \mathbf{r} \left[ e^{i \mathbf{r}} \left( -1 \right) e^{-iq \cdot \mathbf{r}} \right] \tag{11}
\]

First-order Vector Perturbation Theory (FOVPT)

The afore-mentioned scalar quantity can only be applied in the condition of glancing incidence, and cannot obtain the surface height information. The first order vector perturbation theory can obtain the PSD information of the target surface. In the process of characterization of the roughness statistics of optical instruments, the PSD function is the most perceptual expression. When \( \phi = 0 \), the FOVPT can be expressed as follows.

\[
\Pi(\theta) = \frac{1}{W_0} \frac{dW_{\text{scatt}}}{d\theta} = \frac{k^3 \left| 1 - \varepsilon \right|^2} {16\pi \sin \theta_0 \sqrt{\cos \theta_0 \cos \theta}} \frac{\left| f(\theta_0) \mu(\theta) \right|^2}{PSD_{1D}(p)} \tag{12}
\]

\[
p = \frac{1}{\lambda} \left| \cos \theta - \cos \theta_0 \right| \tag{13}
\]
\[ t(\theta) = \frac{2 \sin \theta}{\sin \theta + \sqrt{\varepsilon - \cos^2 \theta}} \]  

(14)

Figure 3. Simulation Results at Different T. Left: \( Y=14.991X^2+8.644X+0.887 \) at \( T=1000 \). Middle: \( Y=14.989X^2+8.492X+7.344 \) at \( T=100 \). Right: \( Y=15.024X^2+9.438X+3.805 \) at \( T=10 \).

\[ \varepsilon = (1 - \delta + i\beta)^2 \]  

(15)

\[ PSD_{10}(p) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(h + \frac{1}{2})}{\Gamma(h)} \frac{\sigma^2 a}{(1 + a^2 p^2)^{h + \frac{1}{2}}} \]  

(16)

Simulated Annealing Algorithm (SAA)

The Simulated Annealing Algorithm is a heuristic optimization algorithm. Conforming to the Metropolis criterion, a procedure that uses the aforementioned algorithm first set up a target function \( M \) analogical to the intrinsic energy \( E \) initially found in the object awaiting to be annealed. Meanwhile, the temperature \( T \) is simulated by the control parameter \( a_1 \). The SAA solves the problem with seven steps. The initial temperature of the SAA determines how the probability of finding the optimal solution. Figure 3 shows the simulation results of a 3-parameter quadrilateral polynomial using different SAA parameters.

- Left: \( T=1000; \ T_0=0.01; \ a_1=0.95; \ \text{maxstep}=1000; \ L=100; \)
- Middle: \( T=100; \ T_0=0.01; \ a_1=0.95; \ \text{maxstep}=1000; \ L=100; \)
- Right: \( T=10; \ T_0=0.01; \ a_1=0.95; \ \text{maxstep}=1000; \ L=100; \)

Using the least square method, the quadrilateral polynomial is fitted. Hence, the initial starting temperature should be high enough for optimal results.
The control parameter $a_1$ is the most important parameter (see Figure 4). If $a_1$ is too close to 1, the optimal choice will likely be eliminated. If $a_1$ is too small, finding the optimal choice will be too long.

EXPERIMENTS

AFM Results

AFM uses the interaction force between atoms to measure the probe and the object. Hence, the surface height can be measured. AFM forms its image by three methods: contact, non-contact, and tapping. In this project, tapping mode is utilized to measure the surface roughness. Samples: two 15mm*15mm square silicon chips (M1 and M2).

<table>
<thead>
<tr>
<th>Table II. AFM RESULTS (Surface height RMS nm).</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1-point1</td>
</tr>
<tr>
<td>0.138</td>
</tr>
<tr>
<td>0.135</td>
</tr>
</tbody>
</table>

Figure 5. PSD Curves of AFM Results (Upper Left: M1-10nm, Upper Right: M1-50nm, Lower Left: M2-10nm, Lower Right: M2-50nm).
X-ray Scattering Results

Figure 6. PSD Curves of X-ray Scattering Results (Upper Left: M1-0.1deg, Upper Right: M1-0.15deg, Lower Left: M2-0.1deg, Lower Right: M2-0.15deg).

<table>
<thead>
<tr>
<th></th>
<th>Surface roughness under 10×</th>
<th>Surface roughness under 50×</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area 1</td>
<td>0.51</td>
<td>0.31</td>
</tr>
<tr>
<td>Area 2</td>
<td>0.57</td>
<td>0.30</td>
</tr>
<tr>
<td>Area 3</td>
<td>0.51</td>
<td>0.47</td>
</tr>
<tr>
<td>M-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area 1</td>
<td>0.49</td>
<td>0.33</td>
</tr>
<tr>
<td>Area 2</td>
<td>0.54</td>
<td>0.32</td>
</tr>
<tr>
<td>Area 3</td>
<td>0.64</td>
<td>0.33</td>
</tr>
</tbody>
</table>
White Light Interferometry Results

The contour scanning device uses interference pattern to obtain the entire 3-dimensional morphology data. Since the contour scanning device cannot obtain extremely accurate data, it can only be utilized to validate the X-ray scattering device.

Comparison of Results

Putting all the afore-mentioned experimental data on a single figure, the different data of chip M1 is compared. The ten times Contour Scanner stands out at the outlier because the PSD function from $10^{-3}$ to $5 \times 10^{-2}$ dissatisfies theoretical condition. In the right part of the data, all measurement methods except X-ray scattering appears to be smooth. This is largely because of the white noise in the data. X-ray scattering tends to perform well on all frequencies.

Same as chip M1, the ten times Contour Scanner measures differently from other methods. Hence, the curve appears to be an outlier. Although the same correct factor is used, the X-ray scattering fits better with other data this time. The X-ray scattering device performed extremely well on both ultra-smooth chips.
Figure 8. Comparison of Results (Left: M1, Right: M2).

TABLE IV. DATA FITTING RESULTS.

<table>
<thead>
<tr>
<th>Fraction index</th>
<th>M1(0.1)</th>
<th>M1(0.15)</th>
<th>M2(0.1)</th>
<th>M2(0.15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface height(nm)</td>
<td>0.686</td>
<td>0.646</td>
<td>0.603</td>
<td>0.589</td>
</tr>
<tr>
<td>Lateral correlation length(nm)</td>
<td>0.135</td>
<td>0.145</td>
<td>0.189</td>
<td>0.197</td>
</tr>
<tr>
<td>131.875</td>
<td>281.983</td>
<td>258.861</td>
<td>270.654</td>
<td></td>
</tr>
</tbody>
</table>

Data Fitting Using Annealing Algorithm

See Table IV.

CONCLUSION & DISCUSSION

The independent variable $p$ falls between 0.01-10/μm.

$$\log\left(PSD_{1D}(p)\right) = \log\left(\frac{2}{\sqrt{\pi}} \frac{\Gamma(h+\frac{1}{2})}{\Gamma(h)} \frac{\sigma^2 a}{(1+a^2 p^2)^{h+\frac{1}{2}}}\right)$$

(17)

Since $h$ appears on the exponential term, it is the crucial parameter in the simulation. Its fit is the closest to the real value. The other important factor, $\sigma$, also falls between 0.09±0.01nm. Although this result has a maximum of 7.4% margin of error, it is already much more accurate than the 1nm range of Contour scanning devices. Meanwhile, the greatest margin of error appears on the horizontal correlation parameter M1. This is because M1 shows the periodicity of surface fluctuation. Therefore, it is normal that this parameter fluctuates, as long as it is in the same order.

The data of X-ray scattering has very small absolute residual values. Combining the data with other measuring methods, it can be concluded that X-ray scattering,
along with Simulated Anneal Algorithm, supersedes Contour Scanning and AFM on both accuracy and stability.

REFERENCES