Simultaneous Measurement of Film Surface Topography and Thickness Variation using Three-wavelength Interferometry -Model-based Separation Algorithm of Two Overlapped Sinusoidal Signals-

Toray Engineering Co. Katsuichi KITAGAWA

Conventional transparent film thickness measurement methods such as spectroscopy are basically capable of measuring a single point at a time, and the spatial resolution is limited. We have developed a novel areal film thickness and topography measurement method by three-wavelength interferometry. It estimates both the profiles of front and back surfaces and the thickness distribution of a transparent film simultaneously by model-based separation of two overlapped signals in the interferogram. The validity of the proposed method is proved by computer simulations and preliminary experiments.

1. Introduction

Spectroscopic methods and ellipsometry are typical methods for measuring the thickness of transparent films. These methods are basically point measurement, and it takes a lot of time to obtain the film thickness distribution. In addition, the horizontal resolution is limited due to the size of the measurement spot.

In order to solve the above problems, we have developed many transparent film measurement algorithms by optical interferometry and have been trying to put them into practical use [1-6].

In this report, we propose a new method for measuring the front-surface profile, the back-surface profile, and the film thickness distribution by three-wavelength interferometry. The validity of the proposed method was confirmed by computer experiments and actual sample measurements.

2. Measurement principle

The interferogram of each pixel during vertical scanning obtained by the three-wavelength interference optical system (Fig. 1) [7] is as shown in Fig. 2(a). This intensity signal is the sum of the surface reflection light component and the back-surface reflection light component of the film, and can be expressed by the following equation.

\[
g(i, j) = a(j) + b_1(j) \cos \left[ 4\pi \left( z(i) - z_1 \right) / \lambda(j) \right] \\
+ b_2(j) \cos \left[ 4\pi \left( z(i) - z_2 \right) / \lambda(j) \right]
\]

where \( i \) is the data number, \( j \) is the wavelength number (B,G,R), \( a(j) \) is the bias component, \( b_1(j) \) is the front amplitude, \( b_3(j) \) is the back amplitude, and \( \lambda(j) \) is the wavelength, \( z \) is the height direction position, \( z_1 \) is the front-surface height, and \( z_2 \) is the back-surface height.

Consider finding the unknown variables by fitting this model function to the observed intensity data by least-squares fitting (minimizing the following evaluation function).

\[
F = \sum_{j=1}^{3} \sum_{i=1}^{N} \left[ g(i, j) - g_{i,j} \right] ^2
\]

where \( g_{i,j} \) is the observed intensity values and \( N \) is the number of the observed points.

However, in this least-squares problem, there are 11 unknown variables, \( a(j) \), \( b_1(j) \), \( b_3(j) \), \( z_1 \), \( z_2 \), while the number of the observed data is totally 9, consisting of three parameters (bias, amplitude, and phase) of the sine wave per each wavelength. Therefore, this problem cannot be solved.

Therefore, we reduce the number of the unknown variables by assuming that the wavelength dependence of the refractive index is small enough in many film types, hence the amplitude ratio \( \alpha = b_1(j)/b_2(j) \) does not depend on the wavelength. Then we obtain:

\[
g(i, j) = a(j) + \alpha b_2(j) \cos \left[ 4\pi \left( z(i) - z_1 \right) / \lambda(j) \right] \\
+ b_2(j) \cos \left[ 4\pi \left( z(i) - z_2 \right) / \lambda(j) \right]
\]

There are nine unknown variables in this model: \( a(j) \), \( b_2(j) \), \( z_1 \), \( z_2 \), \( \alpha \), so this fitting problem can be solved if \( N \geq 3 \).

From \( z_1 \) and \( z_2 \) obtained above, the film thickness \( t = (z_2 - z_1)/n \) and the back-surface height \( b = (s - t) \) are obtained using the refractive index \( n \).

![Fig. 1 Three wavelength color imaging system](Image)

3. Computer experiment

3.1 Experimental method

A theoretical interferogram was created under the following conditions:

- Wavelength: 470nm, 560nm, 600nm
- Surface signal: \( a(B, G, R) = 1 \); \( b(B, G, R) = 1 \)
- Back-surface signal: \( a(B, G, R) = 0.5 \); \( b(B, G, R) = 0.5 \)
- \( z_1 = 0 \)nm, \( z_2 = -200 \)nm
- Observation data interval: 5nm
- Observation data range: ±500 nm

The waveform obtained is shown in Fig. 2(a).

Least-squares fitting used Solver, which is an optimization tool of MS Excel. The initial values were \( z_1 = 50 \)nm, \( z_2 = -250 \)nm, \( \alpha = 2.5 \), \( a(j) = \) intensity average value, \( b(j) = \) [intensity range/2/(1+ \( \alpha \)].

3.2 Experimental results

The results are shown in Table 1. Nine unknown variables were correctly estimated including the film thickness of 200 nm. Figure 2(b)(c)(d) shows the signal waveforms of the front and back surfaces extracted from the fitting and the fitted waveform.
Table 1 Fitting results of computer experiment

<table>
<thead>
<tr>
<th>Variables</th>
<th>True</th>
<th>Initial</th>
<th>Estimated</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (B)</td>
<td>1.50</td>
<td>1.57</td>
<td>1.50</td>
<td>0.00</td>
</tr>
<tr>
<td>a (G)</td>
<td>1.50</td>
<td>1.42</td>
<td>1.50</td>
<td>0.00</td>
</tr>
<tr>
<td>a (R)</td>
<td>1.50</td>
<td>1.44</td>
<td>1.50</td>
<td>0.00</td>
</tr>
<tr>
<td>b (B)</td>
<td>0.50</td>
<td>0.45</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>b (G)</td>
<td>0.50</td>
<td>0.34</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>b (R)</td>
<td>0.50</td>
<td>0.29</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>r</td>
<td>2.00</td>
<td>2.50</td>
<td>2.00</td>
<td>0.00</td>
</tr>
<tr>
<td>z (1)</td>
<td>0.0</td>
<td>50.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>z (2)</td>
<td>200.0</td>
<td>-250.0</td>
<td>200.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 2 Fitting results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Initial</th>
<th>Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (B)</td>
<td>1.00</td>
<td>1.64</td>
</tr>
<tr>
<td>a (G)</td>
<td>1.00</td>
<td>1.65</td>
</tr>
<tr>
<td>a (R)</td>
<td>1.00</td>
<td>1.63</td>
</tr>
<tr>
<td>b (B)</td>
<td>1.00</td>
<td>1.65</td>
</tr>
<tr>
<td>b (G)</td>
<td>1.00</td>
<td>1.54</td>
</tr>
<tr>
<td>b (R)</td>
<td>1.00</td>
<td>1.54</td>
</tr>
<tr>
<td>r</td>
<td>0.50</td>
<td>0.65</td>
</tr>
<tr>
<td>x (1)</td>
<td>24.3</td>
<td>21.9</td>
</tr>
<tr>
<td>x (2)</td>
<td>0.0</td>
<td>0.7</td>
</tr>
<tr>
<td>f (R)</td>
<td>0.0200</td>
<td>0.0201</td>
</tr>
<tr>
<td>z</td>
<td>100.0</td>
<td>87.6</td>
</tr>
</tbody>
</table>

Fig. 3 Target image for actual experiment (nominal thickness: 100nm)

4. Actual sample experiment

4.1 Experimental method

(1) Algorithm

In this experiment, the interference signal was sampled from a single image taken with the tilted reference plane. Vertical scanning is not necessary and single-shot measurement is performed. However, the film profile must be constant in the tilt direction.

The model formula in this case is as follows:

\[ F = \sum_{j=1}^{N} \sum_{i=1}^{4} \left[ a(j) + ab(j) \cos(2\pi f_G(j)(x(i) - x_1)) \right]^2 + \sum_{j=1}^{N} \sum_{i=1}^{4} \left[ b(j) + ab(j) \cos(2\pi f_B(j)(x(i) - x_2)) \right]^2 - g_{ij} \]

(4)

where \( f_G \) and \( f_B \) are the fringe frequencies, \( x(i) \) is the horizontal position (pixel), \( x_1 \) is the surface position, and \( x_2 \) is the back-surface position.

Since there is an inverse relationship between the frequency \( f \) and the wavelength \( \lambda \), the frequencies of G and B are expressed by \( f_G = f(R)\lambda d/2\phi \) and \( f_B = f(R)\lambda d/2\theta \). Therefore, there are 10 unknown parameters, \( a(j), b(j), f(R), x_1, x_2, \) and \( \alpha \).

The observed intensity value is normalized by \( \{\text{intensity value-average value}/\{(\text{maximum value-minimum value})/2\}\} + 1 \) to balance the error evaluation among the three wavelengths, and is within the range of approximately \([0, 2]\).

(2) Target of measurement

An image of a commercially available film thickness step wafer (SiO2/Si; refractive index \( n = 1.46 \)) with a nominal film thickness of 100 nm was taken with the optical system in Fig. 1. The illumination light source is a mixture of three wavelengths of 470 nm, 560 nm and 600 nm. The center 4200 pixels on the center line (white line) of the obtained image (Fig. 3) was used as the fitting target data. The interferogram is shown in Fig. 4(a).

4.2 Experimental results

Table 2 shows the fitting results when the initial values are \( a(j) = 1, b(j) = 1, \alpha = 0.5, f(R) = 0.02 \) (visually estimated value), \( x_1 = 24.3, \) and \( x_2 = 0 \). The waveforms are shown in Fig. 4(b)(c)(d). A film thickness of 87.6 nm was obtained, which is close to the manufacturer's value of 82.1 nm measured by an ellipsometer.

5. Discussion

The proposed method separates two superimposed sinusoidal signals of the same frequency individually, and makes it possible to solve a problem that cannot be solved in general, based on some assumptions. The least-squares fitting of this method has many local solutions because the model contains a periodic function. Therefore, in the optimization method based on the gradient, a correct solution may not be obtained depending on the initial value. This problem can be solved by the multi-start method using multiple initial values. From the experiment, the interval of local solutions is expected to be about 200 nm, so the search step should be about 100 nm.

This method can be applied to thin films of 1 μm or less, and in particular, it can be expected to measure ultra-thin films (50 nm or less), which has been difficult with the conventional methods.

6. Summary

We proposed a method for simultaneous measurement of surface shape, back-surface shape, and film thickness distribution. The interference signal model is least-squares fitted to the interference intensity waveform of three wavelengths, and the peak position of the front-surface, the peak position of the back-surface, and the waveform parameters are collectively obtained. The validity of the proposed method was confirmed by computer experiments and actual sample measurements. In the future, we would like to improve the robustness by improving the algorithm and put it into practical use.
This is an English translation of the paper:

(Original in Japanese; translated by the author in Aug, 2020)

The latest paper of this research is: