Thickness Measurement of Thin Film for End-Point Detector(EPD) of Spin Etcher Using the White Light Interferometry.

NOHYU KIM* , YOUNG CHUL CHANG* AND JUNG KEUN CHO**

**Dept. of Mechatronics Engineering, Korea University of Technology and Education, Chunan, Chungnam 330-860, Korea*
phone: +82-41-560-1126, fax: +81-41-560-1253

*** DNS Korea, Chunan, Chungnam 330-290, Korea*
phone : +82-41-620-8360

1. Introduction

End-point detector (EPD) is the thin film measurement equipment developed to control the amount of etching thickness and the remaining time to stop the etching process, which is very critical in semiconductor industry to assure the quality and performance of production line of wafers. Traditional EPD system widely used in plasma etcher called also as dry etching technology, measures the film thickness of the wafer that sits still in plasma gas without any motion or any chemicals on the wafer. However, in the spin etcher the wafer is etched using chemicals while it rotates at a very high speed wafers. Spin etcher is now considered as one of the most promising technology due to its etching speed and low cost [1].

Since approximately 40 percent of whole semiconductor processes is etching and rinsing process, major leading companies including SEZ in Australia and DNS in Japan have been developing a new concept processing method, Spin Etcher, to replace the traditional wet bench. Spin etcher is very effective and cost saving for 12" wafer processing line, so that it is expected to substitute the conventional wet process in major semiconductor manufacturers. EPD system is vitally necessary to ensure the performance and effectiveness of Spin etcher. However EPD for the spin etcher requires special capabilities very different from that for the plasma etcher in its functions, one of which is a real-time measurement to enable to monitor the remaining thickness of the thin film during etching process. The another one which makes it difficult to measure the thin film thickness with high accuracy and reliability is that the interferencce signal from the thin film would deteriorate very much because of the rotational motion of the wafer and chemicals. For the development of EPD for the spin etcher in the operating condition, three technologies are considered to be most important, which are the interferometry over a wide range of light spectrum, noise reduction including filter technique, and optimization technique [2,3].

In this paper an adaptive filtering software for spin etcher is developed to reduce the noise produced in the wide band white light interferometry and its performance tested.

2. Theory

2.1 Interference model in a thin film

When the light is incident to a transparent thin film as shown in Fig. 1, the total intensity I of the reflected light I_1 from the surface of the thin film and the reflected light I_2 from the substrate can be expressed without the attenuation

$$I = I_1 + I_2 + \frac{1}{2Z}\left[Aae^{i\phi} + Aae^{-i\phi}\right] = I_1 + I_2 + \frac{Aa}{2Z}\cos\phi \quad (1)$$

where Z is the impedance of the thin film, A and a are the amplitudes of two waves, and ϕ the phase difference of two wave. Since $I_1 + I_2$ is always same, the equation (1) shows that the intensity of the sum of the two reflected waves varies according to the phase difference of two waves.

2.2 Thin film thickness measurement techniques with white light

For the measurement of thin film thickness the white light has been used in many applications. The reason is because the monochromatic light may generate much more noise and errors in the interference signal, especially the phase information. Contrary to the colored light, the white light produces a special interference variation over the entire range of the spectrum, which shows the repetition of a constructive and destructive interference of each frequency component of the white light.

Based on this interference phenomenon from the thin film and geometric conditions, the thickness of the thin film, h, is then calculated as the followings depending on the thickness of the thin film[4,5].

$$h = m*(4n_1(\sigma_m))^{-1} \quad \text{for } h > \lambda \quad (2)$$

$$h = (2n_1\sigma)^{-1} \text{ or } h = (4n_1\sigma)^{-1} \quad \text{for } h \approx \lambda \quad (3)$$

$$h = \frac{\sqrt{f(\sigma)}}{4\pi n_1} \quad \text{for } h < \lambda \quad (4)$$

The above equations are basically derived from an ideal thin film model like Fig. 1. However, the interference pattern obtained from the thin film of the wafer rotating with a chemical fluid in spin etcher is corrupted by a lot of noise induced by the vibration and the chemicals. Thus a special filtering using the adaptive algorithm and optimization technique are applied to determine the film thickness.

3. Adaptive filtering

Many sources of noise can contaminate and blur the light intensity from the wafer. The noise may come from the fluctuations, vibration of the rotating table holding the wafer, and from statistical noise induced by photon statistics. Other noise, often confused with signals, is caused by an irregular flow distribution. In order to overcome these noise problems, a data processing technique to reduce the noise is needed. In this study an additive noise is considered.

Let f_i be the degraded signal of the pixel (i) in a 1-dimensional N x 1 reflection signal. The local mean and variance are calculated over a M x 1 window (one-dimension). If the signal d(x) is corrupted by white additive noise n(x), then

$$f(x) = d(x) + n(x) \quad \text{or} \quad f_i = d_i + n_i \quad (5)$$

where, n(x) is a white random sequence with zero mean and variance σ^2

$$E[n_j] = 0 \quad \text{and} \quad E[n_i n_j] = \sigma^2 \delta_{ij} \tag{6}$$

where δ_{ij} is the kronecker delta and E is the statistical expectation operator which is defined as $E[n(x)] = \frac{1}{T}\int_0^T n(x)dx$. The problem is to estimate d_i, given f_i and the noise statistics, i.e., to approximate d_i with the degraded signal f_i and the noise statistics. Assuming that $\bar{d}_i$ and Q_i are the a-priori mean and variance of d_i, they can be approximated by the local mean and variance of the degraded signal f_i and the noise statistics as follows.

$$\bar{d}_i \cong E[d_i] = E[f_i - n_i] = E[f_i] - E[n_i] = E[f_i] = \bar{f}_i \tag{7}$$

$$Q_i \cong E[(d_i - \bar{d}_i)^2] \cong E[\{(f_i - n_i) - \bar{f}_i\}^2] = E[\{(f_i - \bar{f}_i) - n^2\} = E[(f_i - \bar{f}_i)^2] + \sigma^2 \tag{8}$$

In most filtering techniques, the a-priori mean and variance of d_i are derived from an assumed autocorrelation model or, recursively, from an autoregressive Markov sequence [18]. Either method is an approximation. By this approximation, it is easy to obtain a filtering algorithm either by minimizing the mean-square error or by a weighted least-square estimation. Both methods can be done with the same algorithm. The estimator d'_i of d_i, is computed from

$$d'_i = \bar{d}_i + k_i (f_i - \bar{d}_i) \text{ or } d'_i = k_i f_i + (1-k_i)\bar{d}_i \tag{9}$$

where, the gain is given by

$$k_i = \frac{Q_i}{Q_i + \sigma^2} \tag{10}$$

Since Q_i and σ^2 are both positive, k_i will lie between 0 and 1. A simple intuitive interpretation is that for a low signal-to-noise ratio region Q_i is small compared to σ^2. Then $k_i \approx 0$, and the estimated d'_i is $\bar{d}_i$. Conversely, for a high SNR region, Q_i is much larger than σ^2, $k_i \approx 1$, and the estimated d'_i is f_i, the corrupted signal at the pixel i. As can be seen from the above equation (10), this adaptive filter acts like a high pass filter in a region where the variance of the signal is large, i.e., the image will be sharpened. Similarly, it behaves like a low-pass filter in a region where the variance of a signal is small; in this case the image will be smoothed. The use of different window sizes will greatly affect the quality of processed images. If the window size is too small, the noise filtering algorithm is not effective. If the window is too large, subtle details of the image will be lost in the filtering process. Previous experiments indicate that a [7 x 1] window is a fairly good choice, so that all data in this study are processed by the [7 x 1] window .

12“ wafer specimens prepared to measure using EPD prototype system include the oxide film(500nm and 600nm in thickness) which are fabricated very precisely in the factory. Experiment shows that a good accuracy can be seen form the comparison of the measurement results and the exact thickness of the thin film.

4. Conclusions

Adaptive filter for EPD software program of spin etcher has been developed to measure the thin film thickness of 12" wafers. From the experimental results, the wide band white light from UV to visible can be filtered very effectively to reduce the measurement error due to the vibration and chemical flow in the spin etcher. Interference pattern over a wide range of spectrum provides the abundant information about the thickness and material properties such as refractive indices to remove the noise induced by the spinning motion.

References

[1] M. Quirk and J. Serda, "Semiconductor Manufacturing Technology",Prentice Hall (2001).

[2] S. Kim and G. Kim, "Thickness-profile measurement of transparent thin-film layers by white-light scanning interferometry",Applied Optics, Vol. 38, No. 28,(1999) p. 5968-5973.

[3] H. L. Maynard and N. Hershkowitz,"Thin-film interferometry of patterned surfaces", J. Vac. Sci. Technol. B 13(3), (1995) p.848-857.

[4] Y. Xu and Y. J. Li, "Surface profile measurement of thin films using phase-shifting interferometry", SPIE Vol. 2545, (1995) p.238-248

[5] D. Wang, H. Zhang, Y. Yang and Y. Zhang, " Ultrathin thickness and spacing measurement by interferometry and correction method", SPIE Vol. 2542, (1995) p.129-135

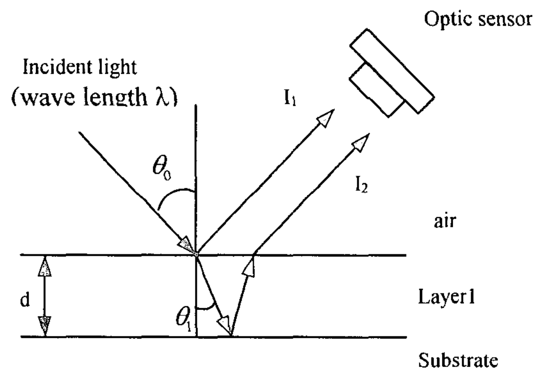

Fig.1 Reflection waves from a thin film