3-D Shape Measurement of a Transparent Film by White Light Interferometry

Katsuichi KITAGAWA

Toray Engineering Co.
katsuichi_kitagawa@toray-eng.co.jp

Abstract: In vertical scanning white-light interferometry, two peaks appear in the interference waveform when a transparent film exists on the surface to be measured. We developed an algorithm to detect these peaks with high speed and accuracy, and developed a film profiler that can simultaneously measure the surface and back surface profiles of the transparent film and the thickness distribution of the film. This system is applicable to transparent films with an optical thickness of more than 1 μm and has been used effectively in the manufacturing process of semiconductors and LCDs.

1. Introduction

The surface profiling method using optical interferometry is widely used in industry as a fast, accurate, non-contact, non-destructive method for measuring micro-surface topography. However, when the surface to be measured is covered by a transparent film, the interference light from the backside becomes a disturbance and prevents accurate measurement of the surface profile.

In this paper, a signal processing algorithm to measure the surface profile of a target surface covered by a transparent film in white light interferometry (also called vertical scanning interferometry) is presented, and a film profiler, SP-500F, with its built-in signal processing algorithm is described.

2. Measurement Principle

Using white light with a short coherence length as the light source, the Z-axis of the interference microscope is scanned vertically as shown in Fig. 1.

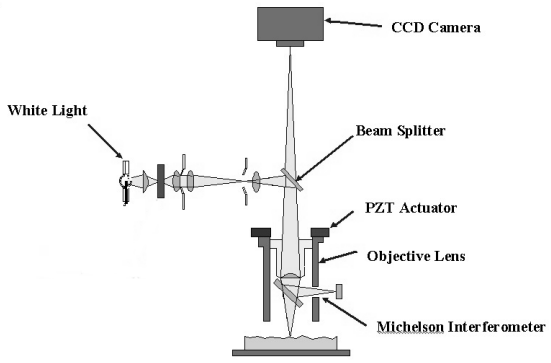


Fig. 1 Optical system of interference microscope

The intensity change at a point in the image (called the interferogram) is shown in Fig. 2, and the maximum modulation of the interference fringes is achieved at the point where the optical path difference between the measurement and reference light paths becomes zero.

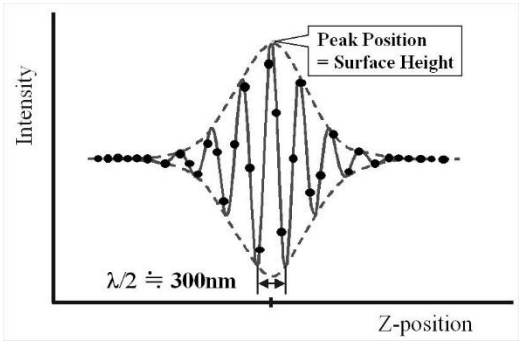


Fig. 2 Interferogram
(Black dots are sampling points)

As algorithms for detecting the peak position of the modulation, the smoothing method [1], the Hilbert transform method [2], the Fourier transform-position method [3], the center-of-gravity method [4], and the waveform restoration method [5] have been proposed. However, most of them are based on the assumption of a single peak. Therefore, if a sample with a transparent film on the surface is measured, the reflected light from the surface is superimposed on the reflected light from the backside, making it impossible to measure the surface profile correctly.

It has long been known that if the thickness of a transparent film is thicker than its coherence length, two peaks, a surface reflection and a backside reflection, appear in the interferogram, so that the thickness of the film can be measured and the film shape can be measured from these peak locations. [6]-[10]

However, those methods reported so far require visual determination of peak positions [6][7], or use a complex calculation algorithm [9][10], and now no automatic measuring devices for practical applications are found in the market.

The author tried to develop a practical algorithm that detects plural peak positions in the interferogram to enable fully automatic measurement of thickness and profiles of a film-covered surface.

3. Algorithms and Experimental Results

3.1 Test Data

Using the previously reported surface profiler SP-500 [5][11], intensity values (interferograms) were collected as shown in Fig. 3. The measurement conditions were as follows.

(1) Illumination: To separate the peaks on the surface of the film from those on the backside, the light source is a halogen lamp and no filter is used, because the shorter coherence length is better. In this case, the coherence length is about 1 μm.

(2) Z-axis scanning speed: 2.4 μm/sec (sample point interval = 0.08 μm). This interval corresponds to about a quarter of the interferogram period.

(3) Target sample: An oxide film with a thickness of about 1 μm formed on a Si wafer. Since the refractive index is about 1.5, the optical film thickness is about 1.5 μm. Since the coherence length is about 1 μm, it is thought to be close to the lower limit of peak separation. In addition, in the semiconductor and LCD processes, there are many transparent films of several micrometers in thickness, such as resist films and coating films, so assuming a film of several micrometers in thickness as a measurement target is of practical significance.

(4) Scanning range: 5 μm. The number of images is 5/0.08 = 63, and the measurement (scanning) time when using a standard camera at 30 fps is about 2 seconds.

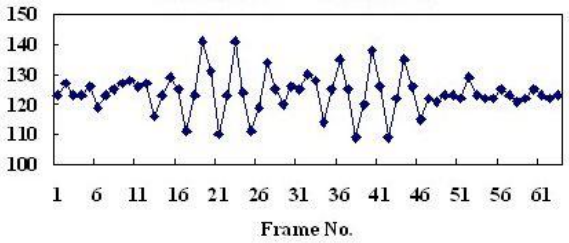


Fig. 3 Interferogram for algorithm testing.

3.2 Algorithm I: Peak Separation Method [12]

The interference amplitude is obtained from the interferogram, which is considered to be a frequency distribution, and the optimal threshold value is obtained by discriminant analysis to separate the peaks, and the peak position is obtained for each separated peak.

3.2.1 Calculation procedure

(1) Step 1: AC component calculation

Fig. 4 is obtained by finding the average value of the intensity data and finding the AC component by (intensity - average value).

(2) Step 2: Amplitude calculation

For the purpose of rectification, squaring the AC component, we obtain Fig. 5.

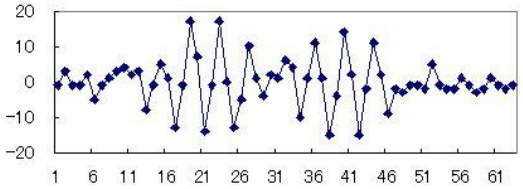


Fig. 4 AC component of intensity signal

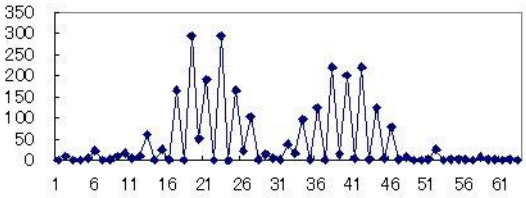


Fig. 5 Squared value of AC components

(3) Step 3: Peak separation

By considering the waveform data with the height on the horizontal axis and the AC component squared value on the vertical axis as the histogram of the observed values and the frequency values on the vertical axis, the optimal classification threshold is obtained by the method of discriminant analysis (Otsu's formula) [8].

That is, given the number of observed levels L and the frequency value $n(i)$, (i = 1, L) of observed levels i, the threshold is k, which is the minimum value of the average intraclass variance expressed by the following equation as an evaluation function.

$$f(k) = \omega_1\sigma_1^2 + \omega_2\sigma_2^2$$

where ω_1 and ω_2: Frequency of each class,

σ_1^2 and σ_2^2: Variance of each class.

In this experiment, the value of the evaluation function looks like the solid line in Fig. 6, with data number 30 being the threshold value.

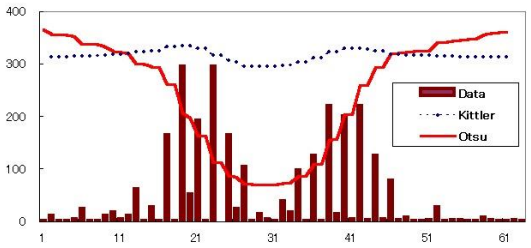


Fig. 6 Peak separation by Otsu's method.
(The solid line is the average intraclass variance.)

(4) Step 4: Peak position detection

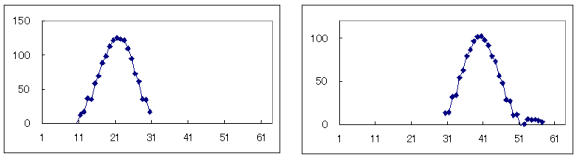
Using data number 30 as a boundary, divide the data into left and right peak regions. For each of the separated peak regions, the peak position detection is performed as before. Here, the low-filtering method, which is the most common peak location detection method, was used to smooth the waveform. The formula

used is a second-order polynomial smoothing method [15] with 19 points, which is expressed as

$$y(n) = (1/W) \Sigma w(i)\, x(n + i)$$

where $i = 0, \pm1, \pm2, ..., \pm9$; W is the sum of $w(i)$ with the normalization coefficient; $w(i)$ is the weight coefficient, and in the case of the 19-point method, from $w(0)$ onwards, 269, 264, 249, 224, 189, 144, 89, 24, -51, and -136.

This results in a smoothed curve, as shown in Fig. 7, and the two peak locations could be obtained as data numbers 21 and 40. Furthermore, by fitting quadratic curves to three data points near the peaks, we obtained the peak positions with a resolution less than the sample point interval by interpolation.



(a) Left peak (b) Right peak

Fig. 7 Each peak smoothing result after peak separation

(5) Step 5: Calculation of the heights and thickness

Let z_{p1} and z_{p2} be the heights converted from the peak position, and let the refractive index of the film be n. Then,

a) Surface height of transparent film = z_{p1}

b) Film thickness $D = (z_{p1} - z_{p2})/n$

c) Height at the back of the film $z_p = z_{p1}$-D

In this test, from the peak positions of 21.17 and 39.61, sample point interval of 0.08 μm, and the refractive index $n = 1.5$, we can obtain the surface height = 5 - 21.17 x 0.08 = 3.31 μm, the thickness = (39.61 - 21.17) x 0.08/1.5 = 0.98 μm, and the height of the back surface = 2.33 μm.

3.2.2 Comparison of peak separation algorithms

In image processing, in addition to the above Otsu's discriminant analysis method, the method of Kittler et al. [14] is another algorithm for finding the binarization threshold based on the gray-level histogram.

In the Kittler method, the threshold is k, which is the minimum evaluation function expressed by the following equation

$$p(k) = \omega_1 \log(\sigma_1/\omega_1) + \omega_2 \log(\sigma_2/\omega_2)$$

where ω_1 and ω_2: Frequency of each class,

σ_1^2 and σ_2^2: Variance of each class.

We compared the two methods for various cases. In the case of this experimental data, there is no difference between them as shown in Fig. 6. Figure 8(a) - (c) shows examples of other cases, where the Kittler method has an advantage when the variance of the two peaks is different, as shown in Fig. 8(c). However, in the case of an interference waveform, there is usually no difference in the variance of the peaks. Therefore, we adopted the Otsu's method, which can obtain good results even in the case of Fig. 8(b).

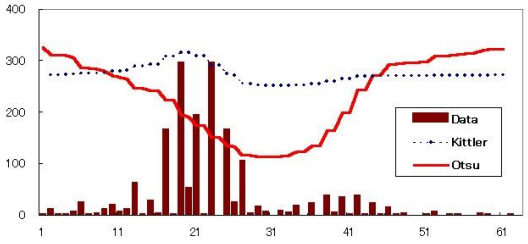


Fig. 8(a) Comparison of the Otsu's method and the Kittler method when the difference in peak height is large

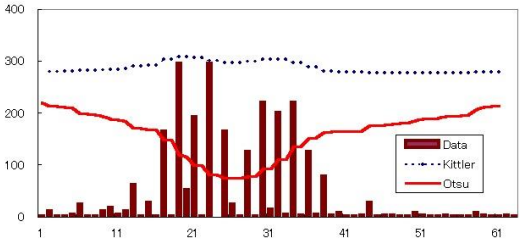


Fig. 8(b) Comparison of the Otsu's method and the Kittler method when the peak interval is small

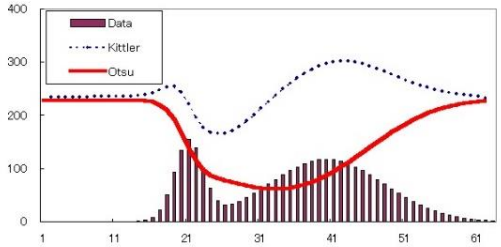


Fig. 8(c) Comparison of the Otsu's method and the Kittler method when there is a difference in variances

3.3 Algorithm II: Function Fitting Method [12]

This method finds the peak position by fitting an approximate function to the peak of the interference amplitude data.

3.3.1 In the case of no transparent film

The AC component squared data is considered as envelope data, and the Gaussian function is fitted by the least-squares method.

The Gaussian function has three parameters: surface height (peak position) x_0, peak width w, and peak amplitude a. The Gaussian function is expressed as

$$g(x, x_0, a, w) = a \cdot \exp[-\pi \{(x - x_0)/w\}^2]$$

In this case, the sum of the squared error is expressed as

$$f(x_0, w, a) = \Sigma [q_i - g(x, x_0, a, w)]^2$$

where $q_i = i$-th actual data,

$x_i = i$-th position.

Finding the three unknown parameters which give the minimum is a non-linear optimization problem, and there are many calculation algorithms and commercial software available. In this experiment, Solver, an

optimization software provided with Excel, was used.

Although the initial values of the unknown parameters need to be selected appropriately, the optimum value was obtained stably by setting (1) peak position = maximum square position; (2) peak amplitude = (maximum square value)/2. Since the peak width is a device parameter that is almost determined by the optical system, such as the illumination source, and the coherence length, we obtained the approximate value beforehand theoretically or experimentally and used it as the initial value.

3.3.2 In the case of transparent film

Since two peaks appear in the interference waveform when the surface of the object is covered by a transparent film, we adopt the sum of two Gaussian functions $g(x, x_0, a, w) + g(x', x'_0, a', w')$ as the model function.

(1) Step 1 to 2: Same as in the Algorithm I, compute the square of the AC component.

(2) Step 3: Function fitting

The sum of two Gaussian functions is used to perform a least-squares fit. The results of this test are shown in Fig. 9 and Table 1. We obtained peak positions 21.14 and 39.84. These values are very close to the results of the algorithm I. How to find the initial values is shown in the next section.

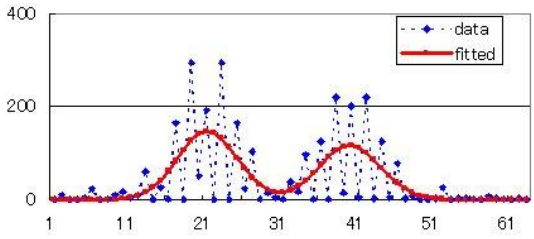


Fig. 9 Raw data and function fitting results

Table 1 Parameter search results

	Initial values		Estimated values	
	Peak 1	Peak 2	Peak 1	Peak 2
Position x	21	39	21.14	39.84
Amplitude a	167	117	147.35	116.98
Width w	13	13	9.69	10.09

(3) Step 4: Calculate the heights and thickness of the front and back surfaces

The same equations as Step 5 of the Algorithm I are used to find the heights and thickness of the film.

3.3.3 Calculation of initial values

In order to set the initial parameter values of the two peaks appropriately, we investigated how to obtain a sufficiently smooth envelope.

(1) Squared-value phase-shifted addition:

In this experiment, since the sample point interval is close to a quarter of the fundamental period of the interference waveform, one-step shifting of the sample value corresponds to shifting the phase of the waveform by 90°. The sum of the square amplitudes of the original and phase-shifted data is the square amplitude of the approximate sinusoidal waveform, as shown in the following equation.

$$\begin{aligned} E^2 &= q(i) + q(i+1) \\ &= f(x)^2 + f(x+\pi/4)^2 \\ &= [A\sin(x)]^2 + [A\sin(x+\pi/4)]^2 \\ &= [A\sin(x)]^2 + [A\cos(x)]^2 \\ &= A^2 \end{aligned}$$

By this method, the envelope is obtained by a simple operation. This method is effective when the sample point interval is 1/4, 1/8, 1/16, etc. We can say that this method is a simplified version of the Hilbert transform-based peak detection method [2]. The results of this test are shown in Fig. 10.

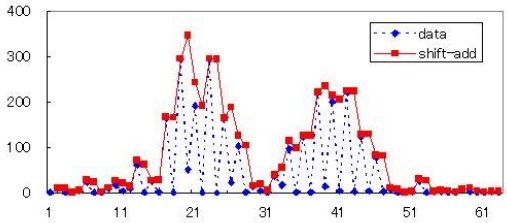


Fig. 10 Amplitude values by squared phase shift addition

(2) Convex envelope:

It is possible to eliminate small concave points by the following three-point neighborhood operation.

$$q(i) = \max[q(i), \{q(i-1) + q(i+1)\}/2]$$

The results of extending this idea to five points and applying it to the amplitude values in Fig. 10 are shown in Fig. 11.

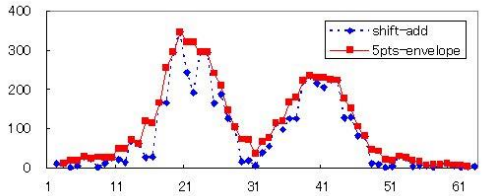


Fig. 11 Convex envelope results

(3) Smoothing:

The waveform is smoothed by the low-pass filter described in the previous section 3.2.1(4). The results of convolution of the coefficients for the five points quadratic approximation are shown in Fig. 12. This results in a sufficiently smooth envelope waveform, and we obtain (21, 333) and (39, 234) as the approximate values of the peak position and height. These were used as the initial values in the previous section.

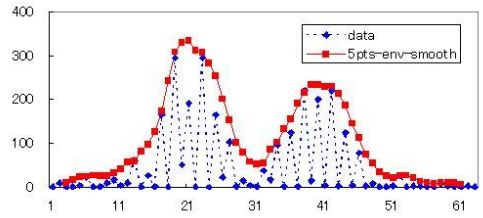


Fig. 12 Smoothing results

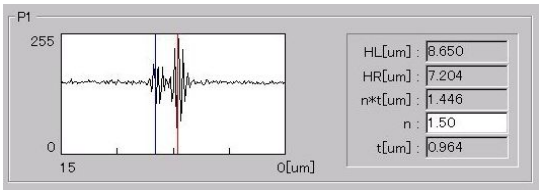
3.4 Comparison of Two Algorithms

Comparing the above two algorithms, the peak-separation method seems to be superior in terms of stability, computational cost, and multi-layered film support, and was adopted for the following commercial product. The function-fitting method will be superior for the thin films with overlapping peaks.

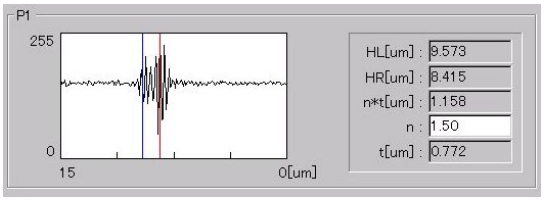
4. Measurable Film Thickness Range

Since this measurement method assumes that the interferogram peaks are separable, there is a lower limit of the film thickness that can be measured.

The interferogram of an oxide film formed on an Si wafer and an example of measurement is shown in Fig. 13. The optical thickness of about 1.1 μm (actual thicness 0.77 μm) is measurable, and the lower limit of measurable thickness is considered to be about 1 μm.



(a) Optical film thickness = 1.45 μm



(b) Optical film thickness = 1.16 μm

Fig. 13 Interferograms of oxide films

5. Film Profile Measuring System SP-500F

The surface profiler SP-500 [11] was installed with software for transparent films based on the above algorithm I. The main specifications are shown in Table 2.

Table 2 Main specifications of the SP-500F film profiler

Measurement method	White light interferometry
Measurement function	Simultaneous measurement of surface and backside shape and film thickness distribution of transparent films
Height range	0 to 100μm [Option: 350μm].
Thickness Range	1 to 100μm in optical film thickness
Field of view	Variable by lens (max. 30mmφ)
Measurement time	Depends on film thickness; several seconds/FOV
Display unit	1nm
Measurement reproducibility	σ = a few to a few-tens nm
Number of measurement pixels	Standard 512 × 480/FOV
Horizontal resolution	Depends on lens magnification and number of camera pixels

6. Examples of Measurements

6.1 Film Thickness Step

Figure 14 shows the results of measuring the oxide film steps of 4 μm and 1 μm thickness formed on the Si wafer. The back side of the wafer is correctly measured as a flat surface.

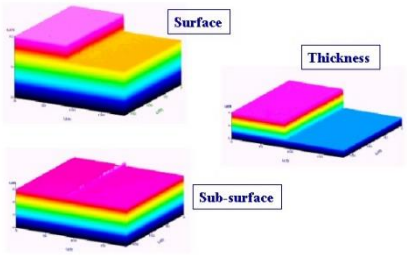


Fig. 14 Film thickness step measurement results

6.2 Thickness-Gradient Film

The results of the thickness-gradient oxide film formed on the Si wafer are shown in Fig. 15. Large concave-shaped surface and backside profiles, and gradient thickness distributions were measured.

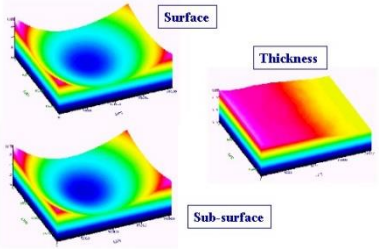
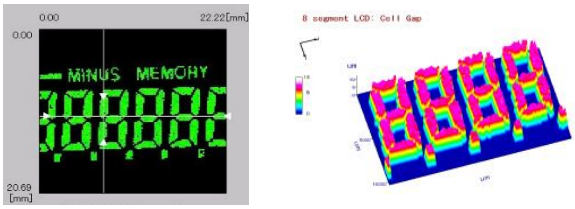


Fig. 15 Thickness-gradient film measurement results

6.3 LCD Cell Gap

The measurement result of the cell gap of the 8-segment LCD is shown in Fig. 16. A wide-field interference optical system with a field of view size of 20 mm was used, and the distribution of the gap (about 8 μm) of the glass substrate was measured.



(a) 2D display (b) 3D display

Fig. 16 Cell gap measurement results

6.4 Refractive Index Estimation

We attempted to measure the thickness of the color resist coated on a glass substrate. Since the refractive index of the resist film was not known, the shape of the film was measured by removing a part of the coated film. The step (= physical film thickness t) of the stripped area was about 1.6 μm and the optical thickness ($n*t$) was about 3.0 μm. From this, the refractive index n is inferred to be about 1.9.

The correctness of the refractive index can also be confirmed by the fact that the surface of the substrate and the backside of the film are flush with the surface of the delaminated area. The shape of the backside of the film when the refractive index is changed is shown in Fig. 17. Visually, it can be estimated that $n = 1.90 \pm 0.02$.

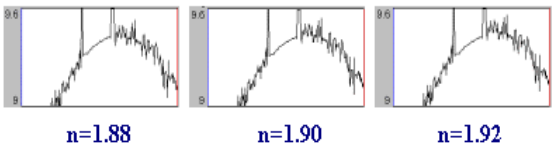


Fig. 17 Back surface profiles calculated with different refractive indexes. Peeled area is located at center left.

6.5 Water Droplets

The measurement results of the water droplet on the glass substrate are shown in Fig. 18.

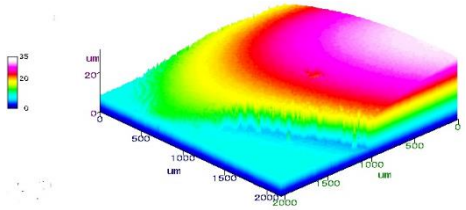


Fig. 18 Results of water droplet measurement

6.6 Measurement Reproducibility

The results of repeated measurements of the oxide film samples are shown in Fig. 19. The average value = 4.609 μm; standard deviation σ = 0.017 μm; CV value = 0.36% were obtained.

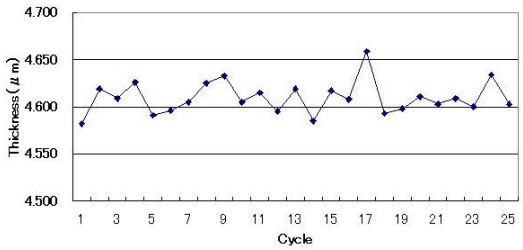


Fig. 19 Reproducibility of film thickness measurement

In practical cases, the area-average thickness is often important, and therefore, on a resist film, averaging was made in the area consisting of approximately 4000 pixels, which obtained the mean value = 1.526 μm; standard deviation (σ) = 3.4 nm and CV value = 0.2%.

7. Conclusion

In the vertical scanning white interferometry, an algorithm was devised to automatically separate the interference signal into the signal from the surface reflection of the transparent film and the signal from the backside reflection. Using this algorithm, we have developed a non-contact Film Profiler, SP-500F, which can simultaneously measure the surface height, back surface height and film thickness distribution within the field of view of a camera in a few seconds.

This development enables the following measurement, which conventionally could hardly be available:

(1) Measurement of a surface profile covered with a transparent film, and

(2) 2D measurement of film thickness distribution.

It has already been effectively used in semiconductor and LCD manufacturing processes.

References

[1] P. J. Caber: Interferometric profiler for rough surface, Applied Optics, 32 (19), 3438-3441 (1993)

[2] S. S. C. Chim and G. S. Kino: Three-dimensional image realization in interference microscopy, Applied Optics, 31 (14), 2550-2553 (1992)

[3] P. de Groot and L. Deck: Three-dimensional imaging by sub-Nyquist sampling of white-light interferograms, Optics Letters, 18 (17), 1462-1464 (1993)

[4] R. Dändliker, E. Zimmermann and G. Frosio: Electronically scanned white-light interferometry: a

novel noise-resistant signal processing, Optics Letters, 17 (9), 679-681 (1992)
[5] K. Kitagawa, H. Ogawa and A. Hirabayashi: Fast Surface Profiler using Generalized Sampling Theorem, Proc. of Vision Engineering Workshop (ViEW) 2001, 141-146 (Yokohama, Dec. 2001) (in Japanese).
[6] P. A. Flournoy, R. W. McClure and G. Wyntjes: White-Light Interferometric Thickness Gauge, Applied Optics, 11 (9), 1907-1915 (1972)
[7] T. Tsuruta and Y. Ichihara: Accurate measurement of lens thickness by using white-light fringes, Proc. ICO Conf. Opt. Methods in Sci. and Ind. Meas., Tokyo, 369-372 (1974)
[8] B. S. Lee and T. C. Strand: Profilometry with a coherence scanning microscope, Applied Optics, 29 (26) 3784-3788 (1990)
[9] O. Nakamura and K. Toyoda: Coherence Probing of the Patterns under Films, Kougaku, 21 (7), 481-484 (1992). (in Japanese)
[10]M. Itoh, R. Yamada, R. Tian, M. Tsai and T. Yatagai: Broad-Band Light-Wave Correlation Topography Using Wavelet Transform, Optical Review, 2 (2), 135-138 (1995)
[11]Toray Inspection Equipment Homepage
http://www.scn.tv/corp/torayins/
[12]Patent Pending
[13]N. Otsu: An Automatic Threshold Selection Method Based on Discriminant and Least Square Criteria, Trans. IECE J63-D, No.4, 349-356 (1980) (in Japanese)
[14]M. Takagi and H. Shimoda ed., Handbook of Image Analysis, Tokyo Univ. Press, pp.503-504 (1991) (in Japanese)
[15]S. Minami: Wave Data Processing for the Scientific Measurements, CQ Shuppan, p.91 (1997) (in Japanese)

Katsuichi Kitagawa: Graduated from the Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo in 1964, and joined Toray Industries, Inc. in the same year. Since 1989, he has been engaged in R&D of inspection equipment for semiconductors and FPDs using image processing and laser optics. Since 2000, he has been a Senior Adviser, Technology in Toray Engineering Co., Ltd. Received the Technology Award from SICE (Society of Instrument and Control Engineers) in 2001. He is a member of SICE and the Technical Committee on Image Application Technology of the Japan Society for Precision Engineering (JSPE).

This paper received the ViEW Odawara Award (Best Paper Award)

This is an English translation of the paper:
Katsuichi Kitagawa: "3-D Shape Measurement of a Transparent Film by White Light Interferometry", Proc. of Vision Engineering Workshop (ViEW) 2003, 87-92 (Yokohama, Dec. 2003).
(original in Japanese; translated by the author in Oct. 2020; revised in Jan. 2026)