

Optimal Characteristic of Optical Filter for White-Light Interferometry based on Sampling Theory

Hidemitsu Ogawa ⁽¹⁾ and Akira Hirabayashi ⁽²⁾

(1) Toray Engineering Co., Ltd., 1-45, Oe 1-chome, Otsu, Shiga, 520-2141, Japan.

(2) Yamaguchi University, 2-16-1, Tokiwadai, Ube City, Yamaguchi 755-8611, Japan.

hidemitsu-ogawa@kuramae.ne.jp, a-hira@yamaguchi-u.ac.jp

Abstract:

White-light interferometry is a technique of profiling surface topography of objects such as semiconductors, liquid crystal displays (LCDs), and so on. The world fastest surface profiling algorithm utilizes a generalized sampling theorem that reconstructs the squared-envelope function $r(z)$ directly from an infinite number of samples of the interferogram $f(z)$. In practical measurements, however, only a finite number of samples of the interferogram $g(z) = f(z) + C$ with a constant C are acquired by an interferometer. We have to estimate the constant C and to truncate the infinite series in the sampling theorem. In order to reduce both the truncation error and the estimation error for C , we devise an optimal characteristic of the optical filter installed in the interferometer in the sense that the second moment of the square of the interferogram is minimized. Simulation results confirm the effectiveness of the optimal characteristic of the optical filter.

1. Introduction

White-light interferometry is a technique of profiling surface topography of objects such as semiconductors, liquid crystal displays (LCDs), and so on. It is attractive because of its advantages including non-contact measurement and unlimited measurement range in principle [1, 2, 3, 5, 6, 8, 9]. From the viewpoint of sampling theory, white-light interferometry has the following two interesting features. First, a signal to be processed, a white-light interferogram, $f(z)$, is a bandpass signal. Second, a signal to be reconstructed from sampled values of $f(z)$ is not the interferogram itself, but its squared-envelope function $r(z)$. This type of sampling theorem is called a generalized sampling theorem [4, 10, 11].

The present authors also derived such a sampling theorem [9]. Based on the theorem, the world fastest surface profiling algorithm were proposed and installed in commercial systems [5]. The sampling theorem is expressed in a form of infinite series and uses samples of the interferogram $f(z)$. In practical measurements, however, only a finite number of samples of the interferogram $g(z) = f(z) + C$ with a constant C are acquired by an interferometer. Hence, in the algorithm, the constant C is estimated from the samples, and the infinite series is truncated with the number of samples. If both the truncation error and the estimation error for C were reduced, we can

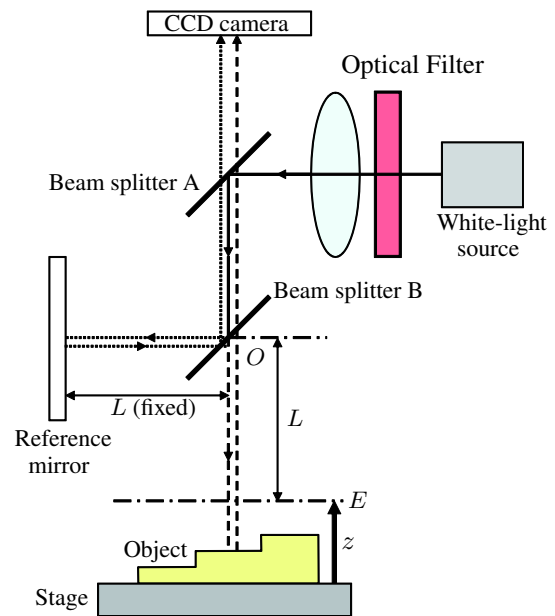


Figure 1: Basic setup of an optical system used for surface profiling by white-light interferometry.

further improve the preciseness of the algorithm. For both error reductions, it is very effective for interferograms to have small side lobes. The waveform of interferograms can be controlled by an optical filter installed in the interferometer.

Hence, in this paper, we devise an optimal characteristic of the optical filter in the sense that the second moment of the square of the interferogram is minimized with a fixed band-width. We show that the optical characteristic is given by a sine curve which has a half of the period as the fixed band-width. We also show that we have a so-called uncertainty principle between the band-width and the second moment. Simulation results confirm the effectiveness of the optimal characteristic of the optical filter.

2. Surface Profiling by White-Light Interferometry

Figure 1 shows a basic setup of an optical system used for surface profiling by white-light interferometry. It uses the Michelson interferometer. A beam from a white-light source passes through an optical filter. The beam is re-

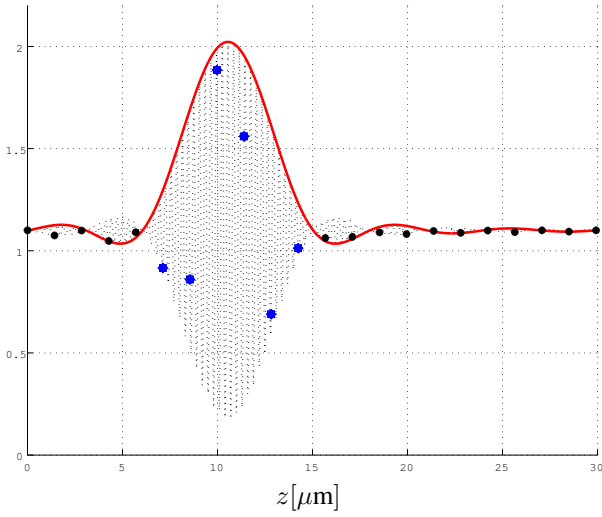


Figure 2: An example of a white-light interferogram $g(z)$ and its sampled values.

flected by the beam splitter A, and divided into two portions by the beam splitter B at the point O . One of the portions indicated by the dotted line is transmitted to a reference mirror, whose distance from the point O is L . The other portion indicated by the dashed line is transmitted to a surface of an object being observed. The height of the surface from the stage at the point P is denoted by z_p . E is a virtual plane whose distance from the point O is L . z is the distance of the plane E from the stage.

The two beams reflected by the object surface and the reference mirror are recombined and interfere. As the interferometer scans along the z -axis, the resultant beam intensity varies as is shown in Fig. 2 by the dotted line. It is called a white-light interferogram or simply an interferogram and denoted by $g(z) = f(z) + C$, where C is a constant. Its peak appears in the right side in Fig. 2 if the height z_p is high, while it appears in the left side if z_p is low. Hence, the maximum position of the interferogram provides the height z_p .

The intensity is observed by a charge-coupled device (CCD) video camera with a shutter speed of 1/1000 second. It has, for example, 512×480 detectors. Each of them corresponds to a point on the surface to be measured. Since the CCD camera outputs the intensity, for example, every 1/60 second, we can utilize only discrete sampled values of the interferogram shown by '•' in Fig. 2. We have to estimate the maximum position of the interferogram from these sampled values.

It is known that the envelope function $m(z)$ shown by the solid line in Fig. 2, or its square $r(z)$, has the same peak as the interferogram and they are much smoother than the interferogram. Hence, usually these functions are used for detection of the peak instead of the interferogram. In this paper, we use the latter $r(z)$, which we call the squared-envelope function.

3. Sampling theorem for squared-envelope functions

Since the interferogram $f(z)$ is a bandpass signal, it can be reconstructed from its samples by using the sampling the-

orem for bandpass signals [7]. It is interesting that, since the squared-envelope function $r(z)$ is the sum of squares of $f(z)$ and its Hilbert transform, the squared-envelope function is also reconstructed from samples of $f(z)$, not those of $r(z)$. Indeed, the following result was established [9, 5]. The center wavelength and the bandwidth of the optical filter in Fig. 1 are denoted by λ_c and $2\lambda_b$, respectively. Let k_l and k_u be angular wavenumbers defined by

$$k_l = \frac{2\pi}{\lambda_c + \lambda_b}, \quad k_u = \frac{2\pi}{\lambda_c - \lambda_b}. \quad (1)$$

Two parameters $\omega_l = 2k_l$ and $\omega_u = 2k_u$ are also used.

Proposition 1 [5] (*Sampling theorem for squared-envelope functions*) Let I be an integer such that

$$0 \leq I \leq \frac{\omega_l}{\omega_u - \omega_l}, \quad (2)$$

and ω_b be any real number that satisfies

$$\begin{cases} \frac{\omega_u}{2} \leq \omega_b & (I = 0), \\ \frac{\omega_u}{2(I+1)} \leq \omega_b \leq \frac{\omega_l}{2I} & (I \neq 0). \end{cases} \quad (3)$$

Let ω_c be a real number defined by

$$\omega_c = (2I + 1)\omega_b. \quad (4)$$

Let Δ be a sampling interval given by

$$\Delta = \frac{\pi}{2\omega_b}, \quad (5)$$

and $\{z_n\}_{n=-\infty}^{\infty}$ be sample points defined by

$$z_n = n\Delta. \quad (6)$$

Then, it holds that

1. When z is a sample point z_j ,

$$r(z_j) = \{f(z_j)\}^2 + \frac{4}{\pi^2} \left\{ \sum_{n=-\infty}^{\infty} \frac{f(z_j + 2n\Delta)}{2n + 1} \right\}^2. \quad (7)$$

2. When z is not any sample point,

$$r(z) = \frac{2\Delta^2}{\pi^2} \left[\left(1 - \cos \frac{\pi z}{\Delta}\right) \left\{ \sum_{n=-\infty}^{\infty} \frac{f(z_{2n})}{z - z_{2n}} \right\}^2 + \left(1 + \cos \frac{\pi z}{\Delta}\right) \left\{ \sum_{n=-\infty}^{\infty} \frac{f(z_{2n+1})}{z - z_{2n+1}} \right\}^2 \right]. \quad (8)$$

To apply Proposition 1 for surface profiling, we have the following difficulties. In the proposition, an infinite number of sampled values $\{f(z_n)\}_{n=-\infty}^{\infty}$ of the interferogram $f(z)$ are used. In practical applications, however, only a finite number of sampled values $\{g(z_n)\}_{n=0}^{N-1}$ of the interferogram $g(z) = f(z) + C$ are available. Hence, we have to truncate the infinite series in Proposition 1 and approximate the sampled values $f(z_n)$ by $g(z_n) - \hat{C}$, where \hat{C} is an estimate of C . For example, the average of $g(z_n)$ is used as \hat{C} . Now, we are suffered from the truncation error as well as the estimation error for \hat{C} . Both of these errors severely affect our final goal of precise estimation of z_p .

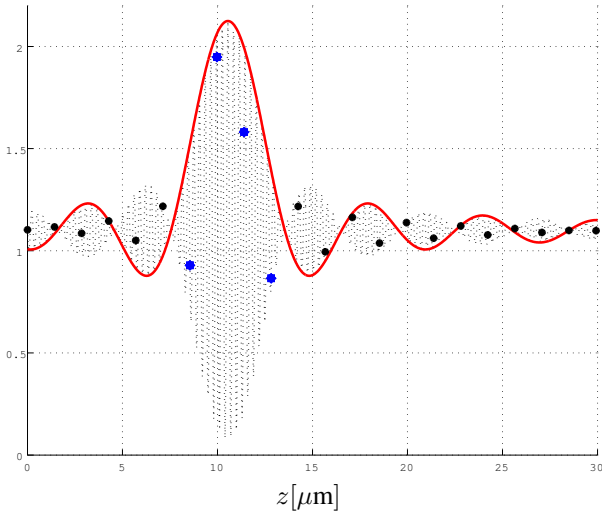


Figure 3: A white-light interferogram $g(z)$ when $\psi(k)$ is rectangular.

4. Optimal characteristics of optical filter

To reduce both of the errors, the following observation is crucial. As you can see in Fig. 2, only a few number of samples are located in the main lobe of $g(z)$ while the rest of them are in side lobes. The latter mostly vanishes once the constant C is estimated precisely. This implies that, the smaller the side lobes are, the smaller the truncation error is. Smaller side lobes also lead us to better estimations of C as shown experimentally in Section 5.

Fortunately, we can control the waveform of the interferogram by the optical filter in the interferometer. Let $a(k)$ be its characteristic in terms of an angular wavenumber k . The support of $a(k)$ is the interval $k_l < k < k_u$. Averaged attenuation rates of two beams along the dashed and the dotted lines in Figure 1 are denoted by $q_o(k)$ and $q_r(k)$, respectively. Let $\psi(k)$ be

$$\psi(k) = \begin{cases} 2\{a(k)\}^2 q_o(k) q_r(k) & (k > 0), \\ 0 & (k \leq 0). \end{cases} \quad (9)$$

It is also supported on the same interval as $a(k)$:

$$\psi(k) = 0 \quad (k < k_l, k > k_u). \quad (10)$$

The function $\psi(k)$ is related to the interferogram $f(z)$ as

$$f(z) = \int_{k_l}^{k_u} \psi(k) \cos 2k(z - z_p) dk. \quad (11)$$

Equation (11) clearly shows that we can control $f(z)$ by $a(k)$ through $\psi(k)$.

To have smaller side lobes, we can easily arrive at the following idea: we design $\psi(k)$ so that it minimizes the second moment of the square of the interferogram $f(z)$:

$$J[\psi] = \int_{-\infty}^{\infty} (z - z_p)^2 \{f(z)\}^2 dz. \quad (12)$$

Now, we are at the point to show our main result in this paper. Let k_a be $(k_u - k_l)/2$.

Theorem 1 Among second continuously differentiable functions $\psi(k) \in C^2[k_l, k_u]$ satisfying

$$\psi(k) = 0 \quad (k \leq k_l, k \geq k_u), \quad (13)$$

$$\psi(k) \geq 0 \quad (k_l < k < k_u), \quad (14)$$

$$\int_{k_l}^{k_u} \{\psi(k)\}^2 dk = 1, \quad (15)$$

$\psi(k)$ that minimizes the criterion $J[\psi]$ is given by

$$\psi(k) = \frac{1}{\sqrt{k_a}} \sin \frac{\pi(k - k_l)}{2k_a}. \quad (16)$$

The minimum value J_0 is given by

$$J_0 = \left(\frac{\pi}{2}\right)^3 \frac{1}{(2k_a)^2} = \frac{\pi \Delta^2}{2}. \quad (17)$$

The following two results are direct consequence of Theorem 1.

Corollary 1 The optimal characteristic $a(k)$ under the criterion $J[\psi]$ is given by

$$a(k) = \left(\frac{\sin \pi(k - k_l)/2k_a}{2\sqrt{k_a} q_o(k) q_r(k)} \right)^{1/2}. \quad (18)$$

Corollary 2 The optimal waveform of the interferogram $f(z)$ is given by

$$f(z) = m(z) \cos(k_u + k_l)(z - z_p), \quad (19)$$

where

$$m(z) = \frac{4\pi\sqrt{k_a} \cos 2k_a(z - z_p)}{\pi^2 - 16k_a^2(z - z_p)^2}. \quad (20)$$

The interferogram shown in Fig. 2 was the optimal one given by Eqs. (19) and (20) while that shown in Fig. 3 is generated from a rectangular $\psi(k)$ given by

$$\psi(k) = \begin{cases} 1/\sqrt{k_u - k_l} & (k_l < k < k_u), \\ 0 & (\text{otherwise}). \end{cases}$$

Though this $\psi(k)$ is not continuously second differentiable, the conditions (13) ~ (15) are satisfied. In both figures, $\lambda_c = 600[nm]$ and $\lambda_b = 30[nm]$ were used. We can see that the side lobes in Fig. 2 are much smaller than those in Fig. 3. The sampling interval used in both figures is $\Delta = 1.425[\mu m]$, which is the maximum among those satisfying Eqs. (2) ~ (5). We have six samples in the main lobe in Fig. 2 while only four samples are located there in Fig. 3 (these samples are displayed by relatively large dots compared to samples in side lobes). In a nutshell, the optimal characteristic results in fewer samples in the small side lobes. This results in small errors on the truncation and the estimation of C , which we demonstrate in the next section through computer simulations.

Before proceeding simulations, let us make a final remark in this section.

Corollary 3 Let σ^2 be the value of $J[\psi]$. Then, the following uncertainty principle holds:

$$\sigma^2 (2k_a)^2 \geq \left(\frac{\pi}{2}\right)^3, \\ \frac{\sigma^2}{\Delta^2} \geq \frac{\pi}{2}.$$

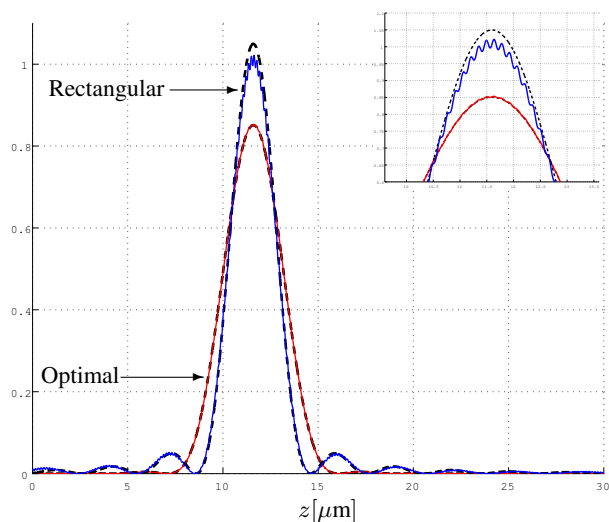


Figure 4: Squared-envelope functions (the dashed lines) and reconstructed functions (the solid lines) from samples of $g(z)$ for both of the optimal and the rectangular $\psi(k)$.

5. Simulations

We compare the optimal and the rectangular characteristics $\psi(k)$ by computer simulations. We first sample the interferograms $g(z)$ generated from both $\psi(k)$ with the sampling interval $\Delta = 1.425\mu\text{m}$. Then, the averages for each sample values are computed for the estimation of C . Finally, we reconstruct the squared-envelope functions $r(z)$ by using a finite number of $g(z_n) - \hat{C}$ instead of $f(z_n)$ in Proposition 1. The reconstructed functions are shown in Fig. 4 by the solid lines as well as the original squared-envelope functions by the dashed lines for both of the optimal and the rectangular $\psi(k)$. The small window in the top-right side shows the magnified image around the peak. We can see that the reconstructed function for the optimal $\psi(k)$ provides a much better result than that for the rectangular $\psi(k)$. We also notice that the latter oscillates severely.

The normalized truncation errors for the optimal and the rectangular $\psi(k)$ are 0.45% and 4.68%, respectively. The former is less than 10% of the latter. When $C = 1.10$, its estimation results are 1.10 and 1.06 for the optimal and the rectangular $\psi(k)$, respectively. Finally, errors for the estimation of z_p are $0.05\mu\text{m}$ and $0.06\mu\text{m}$ for the optimal and the rectangular $\psi(k)$, respectively. Even though the difference is not so significant, the oscillation of the reconstructed squared-envelope function for the rectangular $\psi(k)$ may cause difficulties for fast search of the maximum position.

We repeated the same simulations for thirty two values of z_p from $10\mu\text{m}$ to $20\mu\text{m}$. Then, averages of estimation errors were $0.0496\mu\text{m}$ and $0.0541\mu\text{m}$ for the optimal and the rectangular, respectively. They are almost the same value. However, the averages of truncation errors were 0.35% and 4.67% for the optimal and the rectangular $\psi(k)$, respectively. The former is less than 7% of the latter. These results show the effectiveness of the optimal characteristics of the optical filter.

6. Conclusion

In this paper, we devised an optimal characteristic of the optical filter that minimizes the second moment of the square of the interferogram so that both of the truncation error and the estimation error for the constant in the interferogram are reduced. We showed that the optimal characteristic is given by a sine curve which has a half of the period as the band-width of the optical filter. Simulation results showed that the truncation error for the optimal characteristic is less than 7% of that for the rectangular one. The estimation error of the constant for the optimal characteristic was also smaller than the rectangular one. Even though the difference on the estimation error of the maximum position was not so significant, reconstructed functions for the optimal characteristic was much smoother than those for the rectangular one. These results showed the effectiveness of the optimal characteristic. Our future tasks include to produce a prototype of the optical filter with the optimal characteristic.

References:

- [1] P.J. Caber. Interferometric profiler for rough surfaces. *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. Surface profiling by analysis of white-light interferograms in the spatial frequency domain. *Journal of Modern Optics*, 42(2):389–401, 1995.
- [4] O.D. Grace and S.P. Pitt. Sampling and interpolation of bandlimited signals by quadrature methods. *The Journal of the Acoustical Society of America*, 48(6):1311–1318, 1969.
- [5] A. Hirabayashi, H. Ogawa, and K. Kitagawa. Fast surface profiler by white-light interferometry by use of a new algorithm based on sampling theory. *Applied Optics*, 41(23):4876–4883, 2002.
- [6] G.S. Kino and S.S.C. Chim. Mirau correlation microscope. *Applied Optics*, 29(26):3775–3783, 1990.
- [7] A. Kohlenberg. Exact interpolation of band-limited functions. *Journal of Applied Physics*, 24:1432–1436, 1953.
- [8] K.G. Larkin. Efficient nonlinear algorithm for envelope detection in white light interferometry. *Journal of Optical Society of America*, 13(4):832–843, 1996.
- [9] H. Ogawa and A. Hirabayashi. Sampling theory in white-light interferometry. *Sampling Theory in Signal and Image Processing*, 1(2):87–116, 2002.
- [10] D.W. Rice and K.H. Wu. Quadrature sampling with high dynamic range. *IEEE Transactions on Aerospace and Electronic Systems*, AES-18(4):736–739, 1982.
- [11] W.M. Waters and B.R. Jarrett. Bandpass signal sampling and coherent detection. *IEEE Transactions on Aerospace and Electronic Systems*, AES-18(4):731–736, 1982.