Efficiency of capturing a phase image using cone-beam x-ray Talbot interferometry

Wataru Yashiro,* Yoshihiro Takeda, and Atsushi Momose

Department of Advanced Materials Science, Graduate School of Frontier Sciences, The University of Tokyo,
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan
*Corresponding author: yashiro@mml.k.u-tokyo.ac.jp

We assesses the efficiency of x-ray Talbot interferometry (XTI), a technique based on the Talbot effect for measuring a wavefront gradient, in terms of how quickly it can capture a high-quality phase image with a large signal-to-noise ratio for a given incident photon number. Photon statistics cause errors in the phase of the moiré fringes and impose a detection limit on the wavefront gradient. The relation between the incident photon number and the detection limit is determined, and a figure of merit of XTI for a monochromatic cone beam is then defined. The dependence of the figure of merit on optical system parameters, such as grating pitch and position, is then discussed. The effects of varying the pattern height and linewidth of the second grating are shown for rectangular and trapezoidal teeth. Finally, we show how to design a practical cone-beam Talbot interferometer for certain boundary conditions. © 2008 Optical Society of America

OCIS codes: 110.6760, 340.7450, 170.0170.

1. INTRODUCTION

X-ray phase imaging has attracted much interest over the past decade [1,2]. For light elements, the cross section of the phase shift of hard x-rays is 3 orders of magnitude larger than that of x-ray absorption that is used in conventional x-ray imaging. The phase shift is advantageous for imaging biological soft tissues. Most phase imaging techniques reported in the literature, however, use high-brilliance x-ray beams from synchrotron sources.

X-ray Talbot interferometry (XTI) has recently been proposed as a novel x-ray phase imaging [3–5] and phase tomography [6–11] technique. XTI, which basically consists of two gratings and an area detector, has potential as a medical diagnostic technique, since it can be used with a compact laboratory x-ray generator [11,12]. In this paper, we investigate how to efficiently use incident photons in phase imaging with cone-beam XTI. A figure of merit representing the efficiency is defined. We demonstrate dependence of it on factors related to the optical system and discuss how to design the optical system of the cone-beam XTI.

XTI is based on the Talbot effect [13–15], a self-imaging phenomenon caused by a periodic object under spatially coherent illumination. The experimental setup of XTI consists of an x-ray source, a sample, two gratings, and an area detector. There are several advantages in using XTI for medical diagnostics. First, it can function with a polychromatic and cone beam [6,8,11] of a compact laboratory x-ray generator. Second, wide-area imaging is possible. The area is determined by the size of the absorption grating. It is, however, difficult to fabricate large absorption gratings for XTI—the pattern height must be more than $20-40$ μm and the pitch must typically be a few micrometers. X-ray lithography and electroplating techniques [16–18] can be used in the fabrication process of such a

grating—large-area patterning across an A4-sized area has already been performed [17]. Third, the optical elements used in XTI do not need to have high mechanical stability. Thus, the cost of the diagnosis apparatus can be kept low.

The spatial coherence length at the first grating has to be comparable to, or more than, one period of the grating to take advantage of the Talbot effect. As stated by the van Cittert–Zernike theorem, the length of spatial coherence at a point is inversely proportional to the angular radius of the x-ray source [19]. Therefore, the Talbot effect requires a small source and/or a long distance between the source and the first grating, both of which diminish the intensity of x rays at the area detector and, as a result, make XTI measurements time-consuming even when polychromatic x rays are used. More recently, a Talbot–Lau-type x-ray interferometer has been proposed that can work with a large incoherent polychromatic x-ray source and a grating in front of the x-ray source, creating an array of individually coherent sources [12,20–22]. Although the large x-ray source reduces spatial resolution, this method allows measurements to be performed over a much shorter time, making it a promising technique for medical diagnostics.

In both Talbot and Talbot–Lau-type interferometry, the question of how to quickly capture a phase image using a low-brilliance laboratory x-ray source is a problem that should be quantitatively studied. In this paper, we investigate the efficiency of cone-beam XTI. The discussion presented here can also be applied to the case of the Talbot–Lau-type interferometer after some modifications. The next section briefly reviews the principle of the cone-beam XTI. In Section 3, we treat the effect of photon statistics analytically and determine the relation between the incident photon number and the detection limit of the wave-

front gradient. From this result, a figure of merit of XTI for a monochromatic cone beam is defined as a measure of efficiency. The figure of merit is expressed by the Fourier coefficients of the moiré fringes measured using an area detector, and it is related to optical system parameters, such as x-ray source size and the position, pitch, and shape of the gratings. Section 4 discusses the dependence of the figure of merit on the parameters. Finally, in Section 5, we discuss how to design an optical system for a cone-beam Talbot interferometer and a Talbot-Lau–type interferometer when certain boundary conditions are given.

2. BACKGROUND

A. Cone-Beam X-Ray Talbot Interferometry

Talbot interferometry is based on the "Talbot effect," a self-imaging phenomenon of a periodic object under spatially coherent illumination. The experimental setup of cone-beam XTI is illustrated in Fig. 1. XTI uses two gratings. Self-images are generated at specific distances from the grating through the Talbot effect when coherent x rays illuminate the first grating. A sample placed in front of or behind the grating deforms the self-images so that they include information on the differential phase shift (wavefront gradient) from the sample. The deformation is formed into a moiré pattern by the second grating (absorption grating) located on a self-image, and this pattern can be captured on an area detector placed just behind the second grating. The differential phase shift of the sample is acquired using a fringe scanning technique that is performed by moving one of the gratings in the direction of its period [3,4]. Phase tomography is also possible by rotating the sample and repeating the measurement [6–9].

Consider a case where the first grating, the pitch of which is d_1, is placed at a distance R_1 from an x-ray source generating a monochromatic cone beam with a wavelength of λ. Here we consider the monochromatic case, but the discussion described below can be easily extended to the polychromatic case (see Section 5). Part of the spherical coordinate systems (the latitude and longitude system) is defined as shown in Fig. 1—the origin is the position of the x-ray source, z is the distance from the origin, and the x and y axes are defined as a longitude and an equator, respectively, of a spherical surface whose center is the origin. The line that passes through the origin
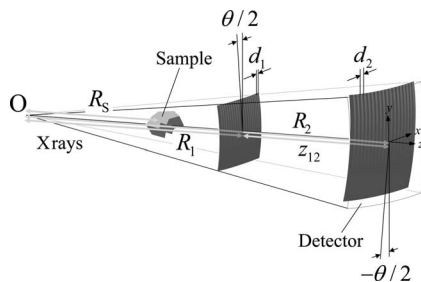
and the intersection point of the x and y axes is defined as the optical axis. For simplicity, we assume that the gratings are set perpendicular to the optical axis and their grooves are parallel to the y axis. The gratings themselves each have a spherical curvature matching a sphere with a radius of curvature equal to the grating's distance from the origin. The surface of the area detector is also on a sphere of radius whose center is the origin. The discussion below can be approximately applied to most practical cases using flat gratings and a flat detector. By expanding the transmission function of the first grating into a Fourier series, the intensity I_{G1} in the $x-y$ plane (the so-called "Fresnel image" [23]) at a distance z_{12} downstream from the grating without a sample is given in the paraxial approximation as follows [14]:

$$I_{\mathrm{G1}}(x,y,z_{12}) \approx I_0 \Delta S t \sum_n \mu_{\bar{n}} b_{\bar{n}}(z_{12}) \exp\!\left(i2\pi \frac{n}{d_2} x \right) \Big/ R_2^2,$$
(1)

where I_0 is the intensity of the x-ray source per unit solid angle per unit time, ΔS is the pixel size of the area detector, and t is the exposure time. In Eq. (1) μ_n is defined by

$$\mu_n = \mu(|npd_1|),$$
(2)

where $\mu(L)$ is the mutual coherence function of x rays [19] in the x direction at the first grating and $b_{\bar{n}}$ is the nth Fourier coefficient of the intensity pattern for completely coherent illumination, given by

$$b_{\bar{n}}(z_{12}) = \sum_{n'} a_{n'+n} a_{n'}^{*} \exp\!\left[-i\pi \frac{((n'+n)^2 - n'^2)\lambda}{d_1^2} \frac{R_1 z_{12}}{R_1 + z_{12}} \right],$$
(3)

where a_n is the nth Fourier coefficient of the complex transmission function of the first grating and R_2 and d_2 are the distance from the source and the pitch of the intensity pattern:

$$R_2 = R_1 + z_{12},$$
(4)

$$d_2 = \frac{R_2}{R_1} d_1.$$
(5)

Here we have assumed for simplicity that the self-image is not compressed (see Appendix A for a compressed self-image).

If the intensity distribution of the source can be expressed as a Gaussian function with standard deviation σ_{s}, the mutual coherence function is given by

$$\mu(L) = \exp[-(L/L_{\mathrm{coh}})^2/2],$$
(6)

where L_{coh} represents the spatial coherence length:

$$L_{\mathrm{coh}} \equiv \lambda R_1/(2\pi\sigma_{\mathrm{s}}).$$
(7)

The distance z_{12} can be expressed using a parameter p in the following form:

$$z_{12} = p \frac{d_1^2}{\lambda} \frac{R_1}{R_1 - p(d_1^2/\lambda)}.$$
(8)

Using p, $b_{\bar{n}}(z_{12})$ can be expressed as

Fig. 1. Experimental setup of the cone-beam x-ray Talbot interferometer, where R_{s}, R_1, and R_2 are the distances of the x-ray source to the sample, the first grating; and the second grating, z_{12} is the distance between the first and the second gratings; and d_1 and d_2 are the pitches of the first and second gratings.

$$b_{\bar{n}}(z_{12}) = \sum_{n'} a_{n'+n} a_{n'}^* \exp[-i\pi p((n'+n)^2 - n'^2)]. \quad (9)$$

The self-images are generated at specific values of p, for example, $1/2, 3/2, 5/2, \ldots$ for a $\pi/2$ phase grating [24].

Note the following convenient formulae. From Eqs. (4) and (8), we get

$$R_2 = \frac{R_1^2}{R_1 - p(d_1^2/\lambda)}. \quad (10)$$

The ratio $\kappa_{2,1}$, defined by R_2/R_1 (or d_2/d_1) is given by

$$\kappa_{2,1} = \frac{R_1}{R_1 - p(d_1^2/\lambda)} \quad (11)$$

$$= \frac{R_2 + p(d_2^2/\lambda)}{R_2}. \quad (12)$$

The distance z_{12} can be also written using the product $d_1 d_2$ from Eqs. (5), (8), and (10):

$$z_{12} = \frac{p}{\lambda} d_1 d_2. \quad (13)$$

We generate a moiré pattern by superimposing the second grating, which is an absorption grating with a pitch of d_2, on the self-image, and putting a sample on the z axis. If the grooves of the first and the second gratings are respectively inclined from the y axis by $\theta/2$ and $-\theta/2$ ($\theta \ll 1$) about the optical axis and a sample is placed at a distance R_s from the source, the moiré pattern can be approximated as follows:

$$I(x,y,p)/I_0 \Delta St \approx \frac{1}{R_2^2} \sum_n \mu_n b_n(z_{12}) c_n \exp\left[i2\pi n \right.$$
$$\left. \times \left(\frac{-y\theta + \chi}{d_2} + \zeta_s \varphi_{x_s}(x_s, y_s) \right) \right], \quad (14)$$

where c_n is the nth Fourier coefficient of the intensity transmission function of the absorption grating before the inclination, χ is the displacement of the second grating in the x direction, and the factor ζ_s is

$$\zeta_s = \begin{cases} p d_s/\lambda & (R_s \leq R_1) \\ z_s/d_2 & (R_s \geq R_1) \end{cases}. \quad (15)$$

Here z_s and d_s are defined by $R_2 - R_s$ and $d_2/\kappa_{2,s}$, where $\kappa_{2,s}$ is the magnification factor given by R_2/R_s, and the coordinates x_s and y_s are defined by $x/\kappa_{2,s}$ and $y/\kappa_{2,s}$ (the $x_s - y_s$ coordinate system is defined at $z = R_s$). The factor $\varphi_{x_s}(x_s, y_s)$ in Eq. (14) is the beam deflection angle at the sample

$$\varphi_{x_s}(x_s, y_s) = \frac{\lambda}{2\pi} \frac{\partial \Phi(x_s, y_s)}{\partial x_s}, \quad (16)$$

where $\Phi(x_s, y_s)$ is the phase shift caused by a sample with a refractive index $1 - \delta(x_s, y_s, z)$:

$$\Phi(x_s, y_s) = -\frac{2\pi}{\lambda} \int \delta(x_s, y_s, z) \mathrm{d}z. \quad (17)$$

In obtaining Eq. (14) we assumed that $\Phi(x_s, y_s)$ is a slowly varying function of x_s; $\Phi(x_s - npd_s, y_s)$ can be approximated by

$$\Phi(x_s - npd_s, y_s) \approx \Phi(x_s) - npd_s \frac{\partial \Phi(x_s, y_s)}{\partial x_s}, \quad (18)$$

and the intensity is averaged over one period of the second grating. In addition, in Eq. (17), we assumed that the size of the sample is sufficiently small in the z direction and that the diffraction by the sample is negligible, so that the phase shift caused by the sample is approximately given by a simple projection of $\delta(x_s, y_s, z)$. Moreover, although we assumed that the absorption by the sample is negligible, the effect of absorption can be easily built into the equation (see Section 3).

A practical laboratory x ray generator for XTI will likely have a micrometer-sized source [10,11]. If the detector has a spatial resolution sufficiently larger than the source size, the effect of penumbra will be small except very close to the source. In this case, a smaller R_s provides a higher effective spatial resolution at the sample. Thus, the sample should be placed in front of or just behind the first grating, because otherwise there would be no advantage in terms of spatial resolution or resolution of the beam deflection angle at the sample. Hereafter, we consider only the case of $R_s \leq R_1$.

B. Phase Retrieval
The phase shift caused by the sample [Eq. (17)] can experimentally be obtained as follows. Equation (14) indicates that the moiré pattern is a contour map $\Psi(x, y)$

$$\Psi(x, y) \equiv 2\pi \left(\frac{y\theta}{d_2} + \zeta_s \varphi_{x_s}(x_s, y_s) \right) \quad (19)$$

and that the moiré fringes are bent due to $\varphi_{x_s}(x_s, y_s)$. If the first-order term $\mu_1 b_1 c_1$ is dominant in Eq. (14), a fringe scanning technique for two-beam interferometry can be used; $\Psi(x, y)$ can be approximated as

$$\Psi(x, y) \approx \arg\left[\sum_{m=0}^{M-1} I_m(x, y, p) \exp\left(-2\pi i \frac{m}{M} \right) \right], \quad (20)$$

where $I_m(x, y, p)$ $(m = 1, 2, \ldots, M)$ is $I(x, y, p)$ in Eq. (14) when $\chi = md_2/M$ $(M \geq 3)$ [25]. Note that the higher-order terms in Eq. (14) cause a systematic error in Eq. (20) [7]. Although we do not touch on the details here, a five-step fringe scan $(M = 5)$ can be used to avoid the systematic error; in this case, when the first or second grating has Ronchi ruling [26] (the linewidth being given by half of the pitch of the grating), the 9th-order term is the lowest one that causes a systematic error.

From experimentally obtained $\Psi(x, y)$, we get

$$\varphi_{x_s}(x_s, y_s) \approx \frac{\lambda}{2\pi \Delta x_s} (\Psi(x, y) - \Psi_0(x, y)), \quad (21)$$

where $\Psi_0(x, y)$ is $\Psi(x, y)$ with no sample and $\Delta x_s \equiv pd_s$. We can obtain the phase shift $\Phi(x_s, y_s)$ from Eq. (21) by

integrating $\varphi_{x_s}(x_s, y_s)$ in the x_s direction.

For simplicity we consider a case where $\theta = 0$. From Eq. (21) the precisions, $\Delta\varphi_{x_s}$ and $\Delta\Psi$, in determining $\varphi_{x_s}(x_s, y_s)$ and $\Psi(x, y)$ are related by

$$\Delta\Psi = 2\pi\Delta x_s \Delta\varphi_{x_s}/\lambda. \tag{22}$$

Equation (22) means that, if $\Delta\Psi$ is given, there is a trade-off between Δx_s and $\Delta\varphi_{x_s}$, where $\Delta\varphi_{x_s}$ gives the detection limit of the wavefront gradient [7]. An experimentally obtained $\Psi(x, y)$ image will have random errors, resulting from photon statistics, readout noise, and/or the dark current of the area detector, etc; $\Delta\Psi$ is determined by such errors. As for the statistical error, $\Delta\Psi$ should depend on the number of photons detected by the detector for capturing a $\Psi(x, y)$ image, which is determined by the number of incident photons and the design of the optical system, as formulated by Eq. (14). Thus, to investigate how to efficiently use the incident photons efficiently for capturing a $\Psi(x, y)$ image, the effect of the statistical error on $\Delta\Psi$ must be clarified. In the next section, we treat this effect analytically.

3. FIGURE OF MERIT OF THE CONE-BEAM XTI

In this section, we analytically formulate $\Delta\Psi$ by using the Fourier coefficients of the moiré fringe and define a figure of merit, η, that is inversely proportional to $\Delta\Psi$.

A. Effect of Photon Statistics on $\Delta\Psi$

1. Analytical Calculation
We analytically investigate how random errors in a measurement of the wavefront gradient cause errors in Ψ. This will allow us to clarify how efficiently the incident photons can be used in phase imaging with cone-beam XTI. When χ in Eq. (14) is md_2/M, the number of photons $N_{mes,m}(x, y, p)$ detected by a pixel of the area detector, located just behind the second grating, can be expressed by

$$N_{mes,m}(x, y, p) = I_m(x, y, p) + N_m \tag{23}$$

$$= \Delta SI_0 t \sum_n q_n \exp(i2\pi n(\zeta_s \varphi_{x_s}(x_s, y_s)$$

$$+ m/M)) + N_m. \tag{24}$$

Here t is the exposure time for capturing the moiré image, q_n is the nth-order Fourier coefficient of $I(x, y, p)/\Delta SI_0 t$ given by Eq. (14), that is,

$$q_n \equiv \mu_n b_n(z_{12}) c_n/R_2^2, \tag{25}$$

and the relative inclination angle θ in Eq. (14) is set to be zero for simplicity. The origin of x can be chosen without loss of generality so that q_1 is a real number. The term N_m represents the sum of random errors:

$$N_m = N_{st,m} + \varepsilon, \tag{26}$$

where $N_{st,m}$ is the error due to photon statistics and ε represents other random errors independent of m. Here we consider how $N_{st,m}$ generates errors in $\Psi(x, y)$ [as defined by Eq. (19)]. Note that if ε is dominant, $\Delta\Psi$ becomes [27,28]

$$\Delta\Psi = \frac{\varepsilon}{q_1 M \Delta St I_0} \frac{1}{\sqrt{2}}. \tag{27}$$

Since $N_{st,m}$ follows a Poisson distribution, the mean square of $N_{st,m}$ is proportional to the number of photons:

$$\langle N_{st,m} \rangle = 0, \tag{28}$$

$$\langle N_{st,m} N_{st,m'} \rangle = \begin{cases} I_m(x, y, p) & (m = m') \\ 0 & (m \neq m') \end{cases}, \tag{29}$$

where $\langle \rangle$ means ensemble average. Let us apply the fringe scanning method described in Section 2 to $N_{mes,m}(x, y, p)$ [see Eq. (20)]:

$$\arg\left[\sum_{m=0}^{M-1} N_{mes,m}(x, y, p)\exp\left(-2\pi i \frac{m}{M}\right) \right]$$

$$= \arg\left[\sum_{m=0}^{M-1} (I_m(x, y, p) + N_{st,m})\exp\left(-2\pi i \frac{m}{M}\right) \right] \tag{30}$$

$$= \arg[\mathcal{I} + \mathcal{N}_{st}], \tag{31}$$

where

$$\mathcal{I}(x, y, p) \equiv \sum_{m=0}^{M-1} I_m(x, y, p)\exp\left(-2\pi i \frac{m}{M}\right), \tag{32}$$

$$\mathcal{N}_{st} \equiv \sum_{m=0}^{M-1} N_{st,m} \exp\left(-2\pi i \frac{m}{M}\right). \tag{33}$$

Assuming that the signal-to-noise ratios in the measured images are sufficiently large, i.e., $|\mathcal{I}(x, y, p)| \gg |\mathcal{N}_{st}|$, and using Eq. (20), we can rewrite the right-hand side of Eq. (31) as

$$\arg[\mathcal{I} + \mathcal{N}_{st}] \approx \arg[\mathcal{I}] + \arg[1 + \mathcal{N}_{st}/\mathcal{I}] \tag{34}$$

$$\approx \Psi(x, y) + \Delta\Psi_{st}, \tag{35}$$

where $\Delta\Psi_{st}$ is the error in $\Psi(x, y)$ due to photon statistics:

$$\Delta\Psi_{st} \approx \text{Img}[\mathcal{N}_{st}/\mathcal{I}]. \tag{36}$$

Now let us consider the effect of photon statistics on a $\Psi(x, y)$ image. First, we show that the ensemble average of the left-hand side of Eq. (30) is equal to $\Psi(x, y)$. From Eqs. (28) and (33), we see that $\langle \mathcal{N}_{st} \rangle = 0$; thus,

$$\langle \Delta\Psi_{st} \rangle \approx \frac{1}{2i}[\langle \mathcal{N}_{st} \rangle/\mathcal{I} - \langle \mathcal{N}_{st}^* \rangle/\mathcal{I}^*] \tag{37}$$

$$= 0. \tag{38}$$

Hence

$$\left\langle \arg\left[\sum_{m=0}^{M-1} N_{mes,m}(x, y, p)\exp\left(-2\pi i \frac{m}{M}\right) \right] \right\rangle \approx \Psi(x, y). \tag{39}$$

Next we take the root mean square of $\Delta\Psi_{st}$.

$$\langle \Delta \Psi_{\text{st}}^2 \rangle \approx -\frac{1}{4}[\langle \mathcal{N}_{\text{st}}^2 \rangle/\mathcal{I}^2 - 2\langle |\mathcal{N}_{\text{st}}|^2 \rangle/|\mathcal{I}|^2 + \langle \mathcal{N}_{\text{st}}^{*\,2} \rangle/\mathcal{I}^{*2}]. \quad (40)$$

Here

$$\langle |\mathcal{N}_{\text{st}}|^2 \rangle \approx \sum_{m=0}^{M-1} \sum_{m'=0}^{M-1} \langle N_{\text{st},m} N_{\text{st},m'} \rangle \exp\left(-2\pi i \frac{(m-m')}{M}\right) \quad (41)$$

$$= \sum_{m=0}^{M-1} I_m(x,y,p) \quad (42)$$

$$= M\Delta SI_0 t \sum_l q_{lM} \exp(i2\pi lM \zeta_s \varphi_{x_s}(x_s,y_s)), \quad (43)$$

and

$$\langle \mathcal{N}_{\text{st}}^2 \rangle \approx \sum_{m=0}^{M-1} \sum_{m'=0}^{M-1} \langle N_{\text{st},m} N_{\text{st},m'} \rangle \exp\left(-2\pi i \frac{(m+m')}{M}\right) \quad (44)$$

$$= \sum_{m=0}^{M-1} I_m(x,y,p) \exp\left(-2\pi i \frac{2m}{M}\right) \quad (45)$$

$$= \Delta SI_0 t \sum_n q_n \exp(i2\pi n(\zeta_s \varphi_{x_s}(x_s,y_s)))$$
$$\times \sum_{m=0}^{M-1} \exp\left(2\pi i \frac{(n-2)m}{M}\right) \quad (46)$$

$$= M\Delta SI_0 t \sum_l q_{lM+2} \exp(i2\pi(lM+2)\zeta_s \varphi_{x_s}(x_s,y_s)), \quad (47)$$

where l is an integer. Hence, if the sums of the terms of q_n ($|n| \geqslant 2$) in Eqs. (43) and (47) are sufficiently small compared with the terms of q_0 and $q_{\pm 1}$, and if $M \geqslant 4$ [29],

$$\langle |\mathcal{N}_{\text{st}}|^2 \rangle \approx M\Delta SI_0 t q_0 \quad (48)$$

and

$$\langle \mathcal{N}_{\text{st}}^2 \rangle \approx 0. \quad (49)$$

In this case, Eq. (40) can be written as

$$\langle \Delta \Psi_{\text{st}}^2 \rangle \approx \frac{1}{2}[M\Delta SI_0 t q_0/|\mathcal{I}|^2] \quad (50)$$

$$\approx \frac{1}{2} \frac{M\Delta SI_0 t q_0}{(M\Delta SI_0 t q_1)^2} \quad (51)$$

$$= \frac{1}{2} \frac{q_0}{M\Delta SI_0 t q_1^2}, \quad (52)$$

where $\mathcal{I}$ is given by [7]

$$\mathcal{I} = \Delta SI_0 t \sum_n q_n \exp(i2\pi n(\zeta_s \varphi_{x_s}(x_s,y_s)))$$
$$\times \sum_{m=0}^{M-1} \exp\left(2\pi i \frac{(n-1)m}{M}\right) \quad (53)$$

$$\approx M\Delta SI_0 t q_1. \quad (54)$$

Therefore,

$$\Delta \Psi = \sqrt{\langle \Delta \Psi_{\text{st}}^2 \rangle} \quad (55)$$

$$\approx \frac{\sqrt{q_0}}{\sqrt{2M\Delta SI_0 t}} \frac{1}{q_1} \quad (56)$$

$$= \frac{\sqrt{q_0}}{q_1} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{I_{\text{total}}\Delta S}}, \quad (57)$$

where $I_{\text{total}} \equiv MI_0 t$, which is the total incident photon number per unit steradian used for obtaining a $\Psi(x,y)$ image. To take into account absorptions of the air, the grating substrate, and the sample, i.e., $A_{\text{air}}(R_2)$, A_{sub}, and A_{sample}, and the detector efficiency, ϵ, the right-hand side of Eq. (57) can be multiplied by $1/\sqrt{A_{\text{air}}(R_2)A_{\text{sub}}^2 A_{\text{sample}}\epsilon}$. Moreover, $\Delta \Psi$ can be written so as to take into account ε in Eq. (26):

$$\Delta \Psi \approx \frac{\sqrt{q_0 I_{\text{total}}\Delta S + \varepsilon^2}}{q_1 I_{\text{total}}\Delta S} \frac{1}{\sqrt{2}}. \quad (58)$$

This is the general expression for $\Delta \Psi$. Hereafter, we assume that ε is negligible for simplicity.

2. Numerical Calculation
The analytical result described above was verified in a Monte Carlo simulation. We took into account the statistical error of the counts detected by each pixel of the area detector and simulated moiré images. Figure 2 shows simulation examples. The wavelength was fixed at 0.3 Å, and σ_s, which represents the source size, was set to be 5 μm. The gratings were rectangular, made of gold, and had Ronchi ruling. The first grating was a $\pi/2$ phase grating, whose pattern height was 4.05 μm for the wavelength. The second was an absorption grating, whose pattern height was 100 μm. The other parameters were as follows: $d_1 = 4.31$ μm, $d_2 = 5.00$ μm, $R_1 = 2.26$ m, and $R_2 = 2.62$ m ($p = 1/2$). As will be described in Section 4, these parameters, in fact, correspond to an optimal set of d_1, d_2, R_1, and R_2.

The upper figures are simulated moiré images that would be captured by an area detector with a pixel size (ΔS) of 25 μm $\times$ 25 μm placed immediately behind the second grating. The simulation assumed 100% detection efficiency and ignored the absorptions of the air and grating substrates. The left and right images correspond to total photon numbers of 10^{14} and 10^{12} photons/sr (9.1 $\times 10^3$ and 9.1×10^1 photons/pixel). The samples were polystyrene spheres with diameters of 0.25 mm (upper left), 0.50 mm (upper right), 0.75 mm (lower left), and 1.00 mm (lower right), located just in front of the first grating ($R_s = R_1$). The relative inclination angle θ was set
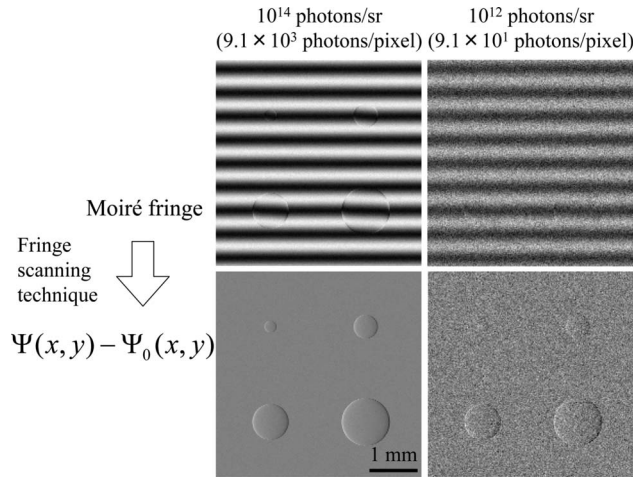
Fig. 2.   Examples of Monte Carlo simulation of the cone-beam XTI. The upper figures are simulated moiré images that would be captured by an area detector placed immediately behind the second grating [corresponding to $m=0$ in Eq. (23) and $\varepsilon=0$ in Eq. (26)]. The left and right figures correspond to the cases of $I_0=10^{14}$ and 10^{12} photons/sr (corresponding to 9.1×10^3 and 9.1×10^1 photons/pixel, respectively). The samples in the figures are polystyrene spheres with diameters of 0.25 mm (upper left), 0.50 mm (upper right), 0.75 mm (lower left), and 1.00 mm (lower right), which are put just in front of the first grating ($R_s=R_1$). The lower figures are $\Psi(x,y)-\Psi_0(x,y)$ images, each of which corresponds to the case of the upper one. Here a $\Psi(x,y)$ image was obtained from five moiré images ($M=5$) by using the fringe scanning technique [see Eq. (20)].

to be 0.5°. The lower figures are $\Psi(x,y)-\Psi_0(x,y)$ images obtained from the upper images [$\Psi_0(x,y)$ being a constant gradient caused by the inclination of the two gratings]. Here a $\Psi(x,y)$ image was obtained from five moiré images ($M=5$) by using the fringe scanning technique described in Subsection 2.B.

From such $\Psi(x,y)-\Psi_0(x,y)$ images as shown in Fig. 2, $\Delta\Psi$, defined by the root mean square of the error in $\Psi(x,y)-\Psi_0(x,y)$, can be calculated. Figure 3 shows the dependence of $\Delta\Psi$ on incident photon number. To avoid the effect of absorption by the samples, the areas of the $\Psi(x,y)-\Psi_0(x,y)$ images that are not affected by the samples were used to calculate $\Delta\Psi$. The abscissa corresponds to $1/\sqrt{I_{\text{total}}}$. Crosses correspond to the results in Fig. 2. The solid black line was calculated with Eq. (57), and it is in good agreement with the results of the simulation. Other symbols are results of cases for different linewidths of the second grating (l_2) and different grating distances from the source. Triangles and open circles correspond to $l_2=d_2/4$ and $l_2=3d_2/4$, respectively; these are examples where even-order Fourier coefficients are not zero. Gray filled circles and gray filled squares correspond to $R_2=2\times2.62$ m and $R_2=0.5\times2.62$ m for $p=1/2$ and $d_2=5.00$ μm, respectively [d_1 and R_1 are given by Eq. (12)]; these are examples where μ_n and R_2 in Eq. (14) change. The broken black, dotted black, solid gray, and broken gray lines corresponding to these four cases were calculated using Eq. (57) and are also in good agreement with the results of the simulation. These results indicate that Eqs. (48) and (49) are correct in the above cases. The reason why the case of $l_2=3d_2/4$ provides a better result than that of $l_2=d_2/4$ is described in Subsection 4.B. The as-
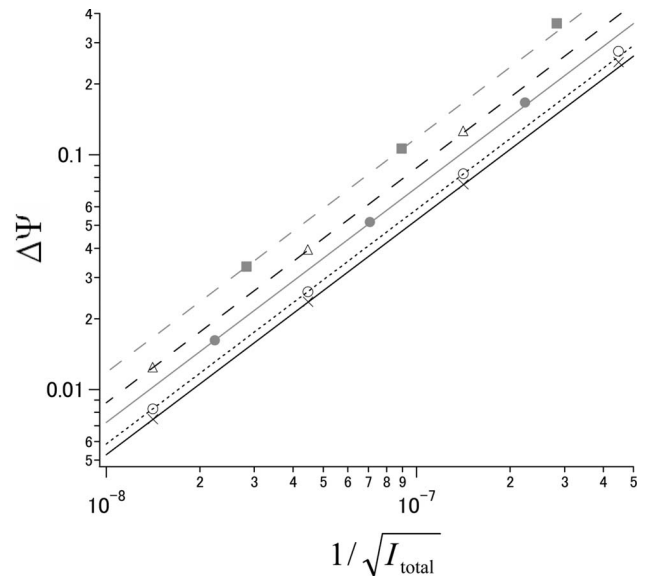

Fig. 3.   Dependence of $\Delta\Psi$ on the incident photon number. The abscissa corresponds to $1/\sqrt{I_{\text{total}}}$. Crosses are results corresponding to Fig. 2. The solid black line was calculated with Eq. (57). Other symbols are results of the simulation in the cases where the linewidth of the second grating (l_2) or distances of the two gratings from the source are changed: triangles, $l_2=d_2/4$ and $R_2=2.62$ m; open circles, $l_2=3d_2/4$ and $R_2=2.62$ m; gray filled circles, $l_2=d_2/2$ and $R_2=2\times2.62$ m; gray filled squares, $l_2=d_2/2$ and $R_2=0.5\times2.62$ m. Broken black, dotted, solid gray, and broken gray lines were calculated with Eq. (57) for the above four cases, respectively.

sumptions used to obtain Eqs. (48) and (49) should provide even more accurate results for a larger M, wherein the contribution of the higher-order terms in Eqs. (43) and (47) is smaller.

B. Figure of Merit η
The performance of cone-beam XTI for a given $I_{\text{total}}\Delta S$ increases when the proportionality coefficient in Eq. (57) decreases. Hence it is natural to define a figure of merit for cone-beam XTI that is inversely proportional to the coefficient. The proportionality coefficient, η_Ψ, is given by

$$\eta_\Psi = \sqrt{2}\,\frac{q_1}{\sqrt{q_0}} \qquad (59)$$

$$= \frac{\mu_1}{R_2}\frac{b_1}{\sqrt{b_0}}\frac{c_1}{\sqrt{c_0}}\sqrt{2}. \qquad (60)$$

Hence $\Delta x_s\Delta\varphi_{x_s}$ in Eq. (22) can be written as

$$\Delta x_s\Delta\varphi_{x_s} = \frac{\lambda}{2\pi\eta_\Psi\sqrt{I_{\text{total}}\Delta S}}. \qquad (61)$$

Note that the second and third factors in Eq. (60) are determined by the shape of the first and second gratings and do not depend on the pitch or position relative to the source, whereas the first factor is determined by the x-ray source size and the position and pitch of the first grating. Hence, the position of the grating affects only the first factor. This fact is useful for designing the optical system of the cone-beam XTI (see Section 5).

If the point-spread function (PSF) of the area detector is known (the width of which represents the spatial resolution of the detector), the effective width of the PSF at the sample, $\Delta_{\mathrm{Det,s}}$, is given by

$$\Delta_{\mathrm{Det,s}} = \kappa_{2,\mathrm{s}}^{-1}\Delta_{\mathrm{Det}}, \qquad (62)$$

where Δ_{Det} is the width of the PSF of the detector. Here we assumed that the effect of penumbra can be neglected for simplicity. From Eq. (61) we get

$$\Delta\varphi_{x_{\mathrm{s}}} = \frac{\lambda}{2\pi\Delta x_{\mathrm{s}}\eta_{\Psi}\sqrt{I_{\mathrm{total}}\Delta S}}. \qquad (63)$$

Multiplying both sides by $\Delta_{\mathrm{Det,s}}$, we get

$$\Delta_{\mathrm{Det,s}}\Delta\varphi_{x_{\mathrm{s}}} = \frac{\lambda\Delta_{\mathrm{Det,s}}}{2\pi\Delta x_{\mathrm{s}}\eta_{\Psi}\sqrt{I_{\mathrm{total}}\Delta S}} \qquad (64)$$

$$= \frac{\lambda\Delta_{\mathrm{Det}}}{2\pi pd_{2}\eta_{\Psi}\sqrt{I_{\mathrm{total}}\Delta S}}. \qquad (65)$$

Hence, it is convenient to define a figure of merit η as follows:

$$\eta = \frac{pd_{2}\eta_{\Psi}}{\lambda}, \qquad (66)$$

$$= \frac{pd_{2}}{\lambda}\cdot\frac{\mu_{1}}{R_{2}}\cdot\frac{b_{1}}{\sqrt{b_{0}}}\cdot\frac{c_{1}}{\sqrt{c_{0}}}\cdot\sqrt{2}. \qquad (67)$$

Using η, we can rewrite Eq. (65) as

$$\Delta_{\mathrm{Det,s}}\cdot\Delta\varphi_{x_{\mathrm{s}}} = \frac{\Delta_{\mathrm{Det}}}{2\pi\eta\sqrt{I_{\mathrm{total}}\Delta S}}. \qquad (68)$$

Again, there is a trade-off between $\Delta_{\mathrm{Det,s}}$ and $\Delta\varphi_{x_{\mathrm{s}}}$, the product of which takes a minimum when η takes a maximum.

Note that when the spatial resolution of the detector is taken into account, the intensity given by Eq. (14) is averaged by the PSF of the detector. Here we briefly discuss how the averaging in the x direction affects the result of a fringe scan (averaging in the $x-y$ plane can be similarly treated). The term $\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}})$ in Eq. (14) around $x_{\mathrm{s}}=x_{\mathrm{s}0}$, where $x_{\mathrm{s}0}$ is the x coordinate of the center of an effective pixel at the sample, is given by $\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}0})+\varphi'_{x_{\mathrm{s}}}(x'_{\mathrm{s}})$, where $\varphi'_{x_{\mathrm{s}}}(x'_{\mathrm{s}})\equiv\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}0}+x'_{\mathrm{s}})-\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}0})$ and $x'_{\mathrm{s}}\equiv x_{\mathrm{s}}-x_{\mathrm{s}0}$. We assume that the effective PSF is an even function around $x'_{\mathrm{s}}=0$. If $\varphi'_{x_{\mathrm{s}}}(x'_{\mathrm{s}})$ is an odd function of x'_{s} around $x'_{\mathrm{s}}=0$, the result of a fringe scan using the measured intensities is $\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}}0)$. Therefore, if $\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}})$ is a slowly varying function of x_{s}, $\varphi'_{x_{\mathrm{s}}}(x'_{\mathrm{s}})$ is approximately given by a linear function of x'_{s}, and the result of a fringe scan should agree with the average $\varphi_{x_{\mathrm{s}}}(x_{\mathrm{s}})$.

It may also be convenient to relate η to the resolution of electron density. We can write δ in Eq. (17) in the following form that neglects anomalous scattering terms:

$$\delta(x_{\mathrm{s}},y_{\mathrm{s}},z) = \frac{r_{\mathrm{e}}\lambda^{2}}{2\pi}\rho(x_{\mathrm{s}},y_{\mathrm{s}},z), \qquad (69)$$

where r_{e} is the classical electron radius and $\rho(x_{\mathrm{s}},y_{\mathrm{s}},z)$ is the electron density of the sample. Then, from Eqs. (16) and (17), we get

$$\Delta\varphi_{x_{\mathrm{s}}} = \frac{r_{\mathrm{e}}\lambda^{2}}{2\pi}\left(\frac{\partial D(x_{\mathrm{s}},y_{\mathrm{s}})}{\partial x_{\mathrm{s}}}\right)_{\mathrm{DL}}, \qquad (70)$$

where $(\ )_{\mathrm{DL}}$ means a detection limit and $D(x_{\mathrm{s}},y_{\mathrm{s}})$ is a projection of the electron density:

$$D(x_{\mathrm{s}},y_{\mathrm{s}}) = \int\rho(x_{\mathrm{s}},y_{\mathrm{s}},z)\mathrm{d}z. \qquad (71)$$

Hence, from Eqs. (68) and (70), we have

$$\Delta_{\mathrm{Det,s}}\left(\frac{\partial D(x_{\mathrm{s}},y_{\mathrm{s}})}{\partial x_{\mathrm{s}}}\right)_{\mathrm{DL}} = \frac{\Delta_{\mathrm{Det}}}{r_{\mathrm{e}}\lambda^{2}\eta\sqrt{I_{\mathrm{total}}\Delta S}}. \qquad (72)$$

The detection limit $(D)_{\mathrm{DL}}$ of a real electron density $D(x_{\mathrm{s}},y_{\mathrm{s}})$ is approximately

$$(D)_{\mathrm{DL}} \approx \Delta_{\mathrm{Det,s}}\left(\frac{\partial D(x_{\mathrm{s}},y_{\mathrm{s}})}{\partial x_{\mathrm{s}}}\right)_{\mathrm{DL}}. \qquad (73)$$

Therefore, $(D)_{\mathrm{DL}}$ is

$$(D)_{\mathrm{DL}} = \frac{\Delta_{\mathrm{Det}}}{r_{\mathrm{e}}\lambda^{2}\eta\sqrt{I_{\mathrm{total}}\Delta S}}, \qquad (74)$$

or

$$(D)_{\mathrm{DL}} = \frac{\Delta_{\mathrm{Det}}}{\eta'\sqrt{I_{\mathrm{total}}\Delta S}}, \qquad (75)$$

where η' in Eq. (75) is

$$\eta' = r_{\mathrm{e}}\lambda^{2}\eta \qquad (76)$$

$$= r_{\mathrm{e}}\lambda pd_{2}\eta_{\Psi}. \qquad (77)$$

Thus, $(D)_{\mathrm{DL}}$ is determined by η or η' if the x-ray source and the area detector are given. Here η' should be convenient for comparing the performances of optical systems designed for wavelengths. For a given wavelength, the next problem is how to maximize η so that incident photons can be efficiently used. In the next section we will investigate how η depends on the parameters of the optical system of a monochromatic cone-beam XTI.

4. DEPENDENCE OF η ON PARAMETERS SPECIFYING THE OPTICAL SYSTEM

We analyze how the figure of merit η defined by Eq. (67) depends on parameters specifying the optical system of the monochromatic cone-beam XTI. First, we examine the effect of the positions and pitches of the gratings on η for a given wavelength λ. The dependence of η on parameters specifying the shape of the second grating is then discussed. Finally, we show that there are lower limits of $\Delta_{\mathrm{Det,s}}\Delta\varphi_{x_{\mathrm{s}}}$ and $(D)_{\mathrm{DL}}$.

A. Position and Pitch of Gratings

As described in Subsection 3.B, η_Ψ can be divided into two factors: μ_1/R_2 and $(b_1/\sqrt{b_0})(c_1/\sqrt{c_0})$. The former factor μ_1/R_2 depends on the x-ray source size and the pitch and position of the gratings. The latter depends on the shapes of the first and second gratings and does not depend on the pitch or position of the gratings. Thus, by considering only the former factor, we can assess the effect of the pitch and position of the gratings. For example, if d_2 is given, μ_1 increases with increasing R_2 [from Eqs. (2) and (5)–(7)]. On the other hand, the intensity of the x rays per pixel of the area detector decreases as the distance R_2 between the source and the second grating increases. Thus, there is a distance R_2 at which μ_1/R_2 takes a maximum, and the position gives the optimal η_Ψ.

Here we show how to optimize d_1, R_1, and R_2 for given d_2, p, and λ. As a result, it will be shown that the optimal $(pd_2/\lambda)(\mu_1/R_2)$ is independent of d_2, p, and λ. We assume that the compression ratio α, defined by the ratio of the pitch of the first grating to that of its self-image in the case of a plane wave, is 1 (no compression). The following discussion can be easily extended to the case of $\alpha \geqslant 2$ (see Appendix A).

The factor $(pd_2/\lambda)(\mu_1/R_2)$ in Eq. (67) can be rewritten as follows:

$$\frac{pd_2}{\lambda}\frac{\mu_1}{R_2} = \frac{pd_2}{\lambda R_2}\mu(|pd_1|), \tag{78}$$

$$= \frac{1}{2\pi\sigma_s}\frac{p}{k}\exp\left[-\left(\left(\frac{p}{k}\right)^2 \bigg/ 2\right)\right], \tag{79}$$

where k is a ratio of L_{coh} to d_1:

$$k = \frac{L_{\text{coh}}}{d_1} \tag{80}$$

$$= \frac{\lambda R_1}{2\pi\sigma_s d_1} \tag{81}$$

$$= \frac{\lambda R_2}{2\pi\sigma_s d_2}. \tag{82}$$

Thus, $(pd_2/\lambda)(\mu_1/R_2)$ is a function of only p/k and σ_s. Figure 4 plots $(p/k)\exp[-((p/k)^2/2)]$ in Eq. (79) as a function of k/p. The curve has a maximum and is inversely proportional to k/p for a sufficiently large k/p.

The function $(p/k)\exp[-((p/k)^2/2)]$ has a maximum at

$$k = p. \tag{83}$$

This means that R_1 is optimal when $L_{\text{coh}} = pd_1$ [see Eq. (80)]; that is, the optimal distance $R_{1,\text{op}}$ is given by
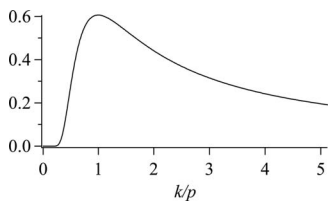

Fig. 4.   Curve $(p/k)\exp[-((p/k)^2/2)]$ plotted as a function of k/p.

$$R_{1,\text{op}} = \frac{2\pi\sigma_s pd_1}{\lambda}. \tag{84}$$

The optimal position of the second grating, $R_{2,\text{op}}$, is also obtained from Eqs. (82) and (83):

$$R_{2,\text{op}} = \frac{2\pi\sigma_s pd_2}{\lambda}. \tag{85}$$

Note that Fig. 2 in Subsection 3.A.2 was, in fact, calculated under this condition. Using Eqs. (12) and (85), we can express $R_{1,\text{op}}$ in terms of d_2:

$$R_{1,\text{op}} = \frac{p}{\lambda}\frac{1}{1/(2\pi\sigma_s d_2) + 1/(2\pi\sigma_s)^2}. \tag{86}$$

The optimal d_1 can be obtained from the ratio $R_{1,\text{op}}/R_{2,\text{op}}$:

$$d_1 = \frac{d_2}{1 + d_2/2\pi\sigma_s}. \tag{87}$$

Figure 5 plots the optimal R_1, R_2, and d_1 as functions of d_2 for $p = 1/2$ and $3/2$ at $\lambda = 0.3$ Å. They are monotonically increasing functions of d_2, and d_1 becomes independent of both p and λ if R_1 is proportional to p/λ (see the Appendix B). This fact is practically useful: *A set of d_1 and d_2, optimized for a given set of p and λ, is optimal for another set of p and λ.*

The maximum of $(pd_2/\lambda)(\mu_1/R_2)$ is as follows:

$$\left(\frac{pd_2}{\lambda}\frac{\mu_1}{R_2}\right)_{\text{max}} = \frac{\exp\left(-\dfrac{1}{2}\right)}{2\pi\sigma_s}, \tag{88}$$

which is independent of d_2, p, and λ. At the optimal position, we have

$$\Delta_{\text{Det,s}}\Delta\varphi_{x_s} \propto \exp\left(\frac{1}{2}\right)\frac{\sigma_s\Delta_{\text{Det}}}{\sqrt{I_{\text{total}}\Delta S}}. \tag{89}$$

That is, $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ is determined by the product of $\sigma_s/\sqrt{I_0(\sigma_s)Mt}$ and $\Delta_{\text{Det}}/\sqrt{\Delta S}$, where the former factor is related to the x-ray source and the latter is related to the area detector.

B. Effect of Second Grating's Shape

As described above, the shapes of the gratings can be treated independently of the pitch and position of the gratings. Here we investigate the effect of the parameters that define the shape of the second grating on $c_1/\sqrt{c_0}$, to which η is proportional. The effect of parameters that define the shape of the self-image generated by the first grating can be treated in the same manner.

Figure 6(a) shows a dependence of $c_1/\sqrt{c_0}$ on the pattern height h_2 and linewidth l_2 of the second grating. The horizontal and vertical axes correspond to h_2 and l_2/d_2, respectively. Here $\lambda = 0.3$ Å, and the second grating is assumed to be rectangular. The numbers in the figure are the values of $c_1/\sqrt{c_0}$. Here $c_1/\sqrt{c_0}$ takes a maximum value around $l_2/d_2 = 1/2$, for each height. In fact, the optimal values of l_2/d_2 are not on the axis of $l_2/d_2 = 1/2$ (the dashed line in the figure). This is due to the asymmetry of
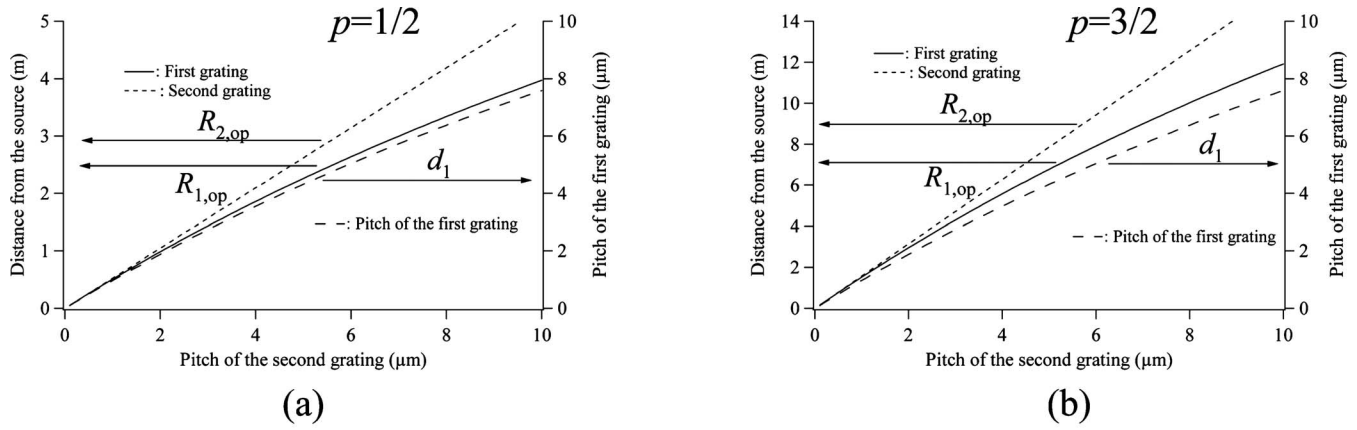
Fig. 5.   Optimum distances of the first and second gratings from the source, $R_{1,op}$ and $R_{2,op}$, and the optimal pitch of the first grating d_1 that maximize $(pd_2/\lambda)(\mu_1/R_2)$ are plotted as a function of the pitch of the second grating d_2 when $\lambda = 0.3$ Å: (a) $p = 1/2$ and (b) $p = 3/2$. The plot of d_1 is, in fact, independent of λ and p.

c_0 around the axis of $l_2/d_2 = 1/2$. That is, the average transmittance of the second grating (corresponding to c_0) when $l_2/d_2 - 1/2$ is positive is smaller than when $l_2/d_2 - 1/2$ is negative, which makes the plot of $c_1/\sqrt{c_0}$ asymmetric.

The linewidth l_2 that maximizes $c_1/\sqrt{c_0}$ for each height h_2 is the solution to the following equation:

$$\tan \pi \left(1 - \frac{l_2}{d_2} \right) = \frac{2\pi \left[(1 - l_2/d_2) + T(h_2)(l_2/d_2) \right]}{1 - T(h_2)}, \quad (90)$$

where $T(h_2)$ is the transmittance of the line of the absorption grating. The optimal linewidth $l_{2,op}$ is approximately

given by $d_2/2$ when h_2 is small. When $T(h_2) \rightarrow 0$, $l_{2,op}$ can be found by solving

$$\tan \pi \left(1 - \frac{l_{2,op}}{d_2} \right) = 2\pi (1 - l_{2,op}/d_2). \quad (91)$$

The solution to Eq. (91) is $l_{2,op}/d_2 = 0.63$. In fact, it can be shown that the transmission function given by a rectangular second grating with $T(h_2) = 0$ and $l_2/d_2 = l_{2,op}/d_2$ is the most optimal transmission function (see Appendix C). The optimal value of $c_1/\sqrt{c_0}$ is 0.48, which is 6.7% larger than that of a Ronchi grating with $T(h_2) = 0$. The transmission function given by a rectangular second grating
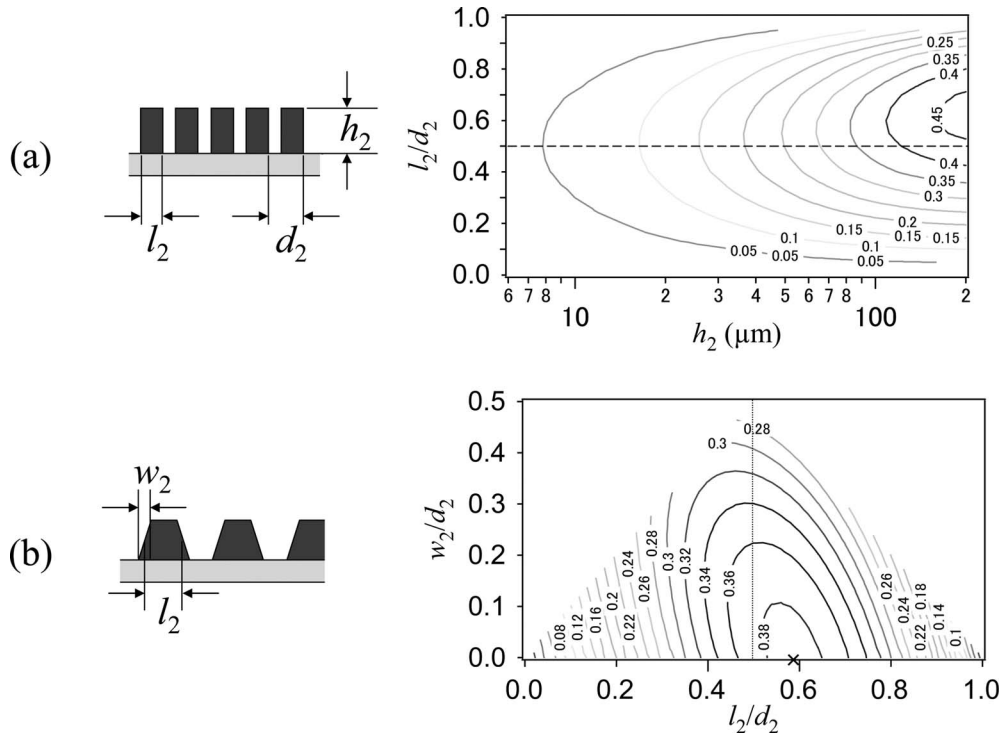

Fig. 6.   (a) Dependence of $c_1/\sqrt{c_0}$ on the pattern height h_2 and linewidth l_2 (normalized by d_2) of the rectangular second grating. (b) Plot of $c_1/\sqrt{c_0}$ for a trapezoidal model as functions of the averaged linewidth l_2 and the width of slope w_2 (normalized by d_2), where the pattern height h_2 was fixed at 100 μm.

with $l_2/d_2 = l_{2,\text{op}}/d_2$ is the optimal one for a finitely large pattern height h_2.

Let us demonstrate another useful result. Figure 6(b) shows the simulated result for a trapezoidal grating, where $c_1/\sqrt{c_0}$ is plotted as a function of average linewidth l_2 and slope width w_2 (defined in the figure) normalized by d_2; $\lambda = 0.3$ Å, and the pattern height of the grating was 100 μm. The factor $c_1/\sqrt{c_0}$ takes a maximum value at the point indicated by the cross on the axis of l_2/d_2. When the linewidth is given on the left of the cross, on the vertical dotted line, for example, $c_1/\sqrt{c_0}$ takes a maximum value at $w_2/d_2 \neq 0$. That is, a grating with edges that are not 90° is better than a rectangular shape for this linewidth. We see from Fig. 6(b) that the performance degradation is gradual as w_2/d_2 increases. This means that a grating with a shallower edge angle does not severely affect the performance of the Talbot interferometer.

C. Lower Limits of $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ and $(D)_{\text{DL}}$

As shown above, the upper limit of $c_1/\sqrt{c_0}$ is 0.48. For example, if a rectangular $\pi/2$ phase grating with Ronchi ruling is used as the first grating and its absorption is negligible, $b_1/\sqrt{b_0}$ is $2/\pi$ [24]. Thus, we can put a lower limit on $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ for the $\pi/2$ phase grating: From Eqs. (67), (68), and (88), and by taking into account the absorptions of the air, grating substrate, and sample, $A_{\text{air}}(R_2)$, A_{sub}, and A_{sample}, and the detection efficiency of the detector, ϵ, we can obtain

$$\Delta_{\text{Det,s}}\Delta\varphi_{x_s} \geqslant 3.8 \frac{\sigma_s}{\sqrt{I_{\text{total}}}} \frac{\Delta_{\text{Det}}}{\sqrt{\Delta S}} \frac{1}{\sqrt{A_{\text{air}}(R_2)A_{\text{sub}}^2 A_{\text{sample}}\epsilon}}. \tag{92}$$

Similarly, the lower limit of $(D)_{\text{DL}}$ becomes

$$(D)_{\text{DL}} \geqslant 3.8 \frac{1}{r_e\lambda^2} \frac{\sigma_s}{\sqrt{I_{\text{total}}}} \frac{\Delta_{\text{Det}}}{\sqrt{\Delta S}} \frac{1}{\sqrt{A_{\text{air}}(R_2)A_{\text{sub}}^2 A_{\text{sample}}\epsilon}}. \tag{93}$$

The factor $\Delta_{\text{Det}}/\sqrt{\Delta S}$ is specific to the individual detector. A smaller width of the PSF makes $(D)_{\text{DL}}$ better for a given ΔS. The factors σ_s, I_{total}, Δ_{Det}, $A_{\text{air}}(R_2)$, A_{sub}, A_{sample}, and ϵ depend on λ. Hence, taking into account their dependence on λ, we can determine the optimal wavelength that minimizes $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ or $(D)_{\text{DL}}$.

5. DISCUSSION

A. How to Design the Optical System of a Cone-Beam XTI

1. Design for a Given Monochromatic X-Ray Source and Area Detector

We discuss how to design the optical system of the monochromatic cone-beam XTI from the results presented in Sections 3 and 4. Here we consider the case where we know the parameters of a monochromatic x-ray source of size σ_s and wavelength λ and an area detector with a PSF of Δ_{Det} (representing the spatial resolution) and a pixel size of ΔS. We also know the shape of two gratings. The problem is how to determine d_1, d_2, R_s, R_1, and R_2 for given set of (p, α). There are two conditions related to the

Talbot effect [for example, Eqs. (10) and (A1)]. In addition, three conditions are necessary to determine a set of d_1, d_2, R_s, R_1, and R_2. The optimization condition [see Eq. (85) and (A12) in Appendix A] that maximizes η in Eqs. (68) and (74) should be the first condition. For the second condition, for example, we can give the effective spatial resolution $\Delta'_{\text{Det,s}}$ at the sample. If the effect of penumbra of the x-ray source is negligible, $\Delta'_{\text{Det,s}}$ can be represented by $\Delta_{\text{Det,s}}$. If the effect of penumbra is not negligible, the effective spatial resolution $\Delta'_{\text{Det,s}}$ at the sample is given by

$$\Delta'_{\text{Det,s}} = \sqrt{\{(1 - \kappa_{2,s}^{-1})\sigma_s\}^2 + (\Delta_{\text{Det}}\kappa_{2,s}^{-1})^2}. \tag{94}$$

In any case, if $\Delta'_{\text{Det,s}}$ is given, we can determine $\kappa_{2,s}$ $(= R_2/R_s)$. As the last condition, we can take d_2. In practical cases, we could use a graph similar to Fig. 5, plotting d_1, d_2, R_s, $R_{1,\text{op}}$, and $R_{2,\text{op}}$ as functions of d_2.

The experimental space that can be used for cone-beam XTI also sets critical conditions on the design of the optical system. We can check the length of the optical system along the optical axis and the space around the sample from the plots of R_s, $R_{1,\text{op}}$, and $R_{2,\text{op}}$. Such plots could show the upper limit of d_2 $(d_2 = \Delta_{\text{Det}}/(\alpha^2 p))$ and lower limits of d_1 and d_2 as determined from the fabrication technology of the gratings. Thus, for each set of (p, α), we may determine some sets of d_1, d_2, R_s, R_1, and R_2.

The above discussion assumed a fixed source size (σ_s). In practical cases, where σ_s can be varied, we should measure the dependence of I_0 on the source size. In Eq. (89) $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ is proportional to $\sigma_s/\sqrt{I_0(\sigma_s)}$. If $I_0(\sigma_s)$ is proportional to σ_s^2, $\Delta_{\text{Det,s}}\Delta\varphi_{x_s}$ does not depend on σ_s. Practically, this is not the case. In terms of cooling efficiency, a smaller σ_s may be better for a given $I_0(\sigma_s)/\sigma_s^2$.

2. Design for the Talbot–Lau-Type Interferometry

Next, we show an example of how to design the optical system of a Talbot–Lau-type interferometer. If a grating G_0 with a pitch of d_0 (the "source grating") is placed at the x-ray source, we have

$$\frac{R_1}{d_0} = \frac{z_{12}}{d_2}. \tag{95}$$

Taking the compression ratio into account and replacing d_2 in Eq. (13) by αd_2, we can find z_{12}

$$z_{12} = \frac{p d_1(\alpha d_2)}{\lambda}. \tag{96}$$

From Eqs. (95) and (96), we find that

$$\frac{1}{d_0} = \frac{p}{\lambda} \frac{\alpha d_1}{R_1} \tag{97}$$

$$= \frac{p}{\lambda} \frac{\alpha^2 d_2}{R_2}. \tag{98}$$

From Eqs. (95) and (97), we can relate d_0, d_1, and d_2:

$$\frac{1}{d_0} = \frac{\alpha}{d_1} - \frac{1}{d_2}. \tag{99}$$

Similar to Eq. (94), the effective spatial resolution $\Delta'_{\mathrm{Det,s}}$ at the sample is given by

$$\Delta'_{\mathrm{Det,s}} = \sqrt{\{(1 - \kappa_{2,\mathrm{s}}^{-1})\Sigma_{\mathrm{s}}\}^2 + (\Delta_{\mathrm{Det}}\kappa_{2,\mathrm{s}}^{-1})^2}, \tag{100}$$

where Σ_{s} is the total size of the x-ray source. If Σ_{s} and Δ_{Det} are given, we can determine $\kappa_{2,\mathrm{s}}$ for a given spatial resolution $\Delta'_{\mathrm{Det,s}}$. In addition, from Eq. (68), we get $\Delta\varphi_{x_{\mathrm{s}}}$ as

$$\Delta\varphi_{x_{\mathrm{s}}} = \frac{\kappa_{2,\mathrm{s}}}{2\pi\eta\sqrt{(2.35\sigma_{\mathrm{s}}/d_0)I_{\mathrm{total}}\Delta S}}. \tag{101}$$

Here I_0 in I_{total} (remember $I_{\mathrm{total}} \equiv I_0 M t$) is the incident photon number per unit steradian from the total area of the x-ray source, and $2.35\sigma_{\mathrm{s}}$ roughly represents the aperture width of G_0 [30]. Hence, the optimization problem for a given $\kappa_{2,\mathrm{s}}$ (a given effective spatial resolution at the sample) is how to maximize $\eta\sqrt{\sigma_{\mathrm{s}}/d_0}$. From Eq. (98) and Eq. (A9) in Appendix A, we have

$$\frac{\sigma_{\mathrm{s}}}{d_0} = \frac{\alpha p}{2\pi k}. \tag{102}$$

Furthermore from Eq. (102) and Eqs. (A7) and (A8) in Appendix A, we have

$$\eta\sqrt{\frac{\sigma_{\mathrm{s}}}{d_0}} \propto \left(\frac{\alpha p}{k}\right)^{1/2} \exp\left[-\left(\left(\frac{\alpha p}{k}\right)^2 \Big/ 2\right)\right]\frac{\sqrt{2\pi}}{d_0}, \tag{103}$$

where $x^{1/2}\exp(-x^2/2)$ has an extreme at $x = \sqrt{1/2}$. Therefore, the right-hand side of Eq. (103) has a maximum at $k = \sqrt{2}\alpha p$ for a given d_0. The optimal positions $R_{1,\mathrm{op}}$ and $R_{2,\mathrm{op}}$ are given by

$$R_{1,\mathrm{op}} = \sqrt{2}\frac{2\pi\sigma_{\mathrm{s}}\alpha p d_1}{\lambda}, \tag{104}$$

$$R_{2,\mathrm{op}} = \sqrt{2}\frac{2\pi\sigma_{\mathrm{s}}\alpha^2 p d_2}{\lambda}. \tag{105}$$

Therefore, from Eqs. (97) and (104), the condition that maximizes $\eta\sqrt{\sigma_{\mathrm{s}}/d_0}$ is

$$\frac{\sigma_{\mathrm{s}}}{d_0} = \frac{1}{2\pi\sqrt{2}}. \tag{106}$$

At the optimal position,

$$\Delta\varphi_{x_{\mathrm{s}}} \propto \frac{\exp(1/4)}{\sqrt{2.35 \times 2\pi(1/2)^{1/2}}}\frac{d_0\kappa_{2,\mathrm{s}}}{\sqrt{I_{\mathrm{total}}\Delta S}}. \tag{107}$$

This means that a smaller pitch of d_0 gives a smaller $\Delta\varphi_{x_{\mathrm{s}}}$.

Let us discuss how to determine the set of d_0, σ_{s}/d_0, d_1, d_2, R_{s}, R_1, and R_2 for a given set of (p, α) (we assume that Σ_{s} and Δ_{Det} are given). We already imposed five conditions: Eqs. (10), (95), and (A1); a given $\kappa_{2,\mathrm{s}}$ (a given effective spatial resolution $\Delta'_{\mathrm{Det,s}}$); and the optimization condition Eq. (106). We have to impose two other constraints to determine the set of parameters. For example, if $\Delta\varphi_{x_{\mathrm{s}}}$ for

an exposure time is given as the first constraint, d_0 can be determined from Eq. (107). As a result, from Eqs. (99), (104), and (105), d_1, R_1, and R_2 can be plotted as a function of d_2 as in the case of the cone-beam XTI. Finally, taking into account boundary conditions such as $\alpha^2 p d_2 \leq \Delta_{\mathrm{Det}}$, the maximum length of the optical system, and the minimum space around the sample, we can find the solution parameter set.

B. Case of Polychromatic Cone Beam

An XTI design that employs a polychromatic cone beam must include the effects of the absorptions of the air and the grating substrate as well as detection efficiency; these depend on the x-ray energy. We give $\Delta\Psi$ as follows:

$$\Delta\Psi = \frac{1}{\eta_{\mathrm{poly},\Psi}}\frac{1}{\sqrt{I_0 M t \Delta S}}, \tag{108}$$

$$\eta_{\mathrm{poly},\Psi} \equiv \sqrt{2}\frac{\int q_1(\lambda,\sigma_{\mathrm{s}})\bar{I}_0(\lambda,\sigma_{\mathrm{s}})A(\lambda)\epsilon(\lambda)\mathrm{d}\lambda}{\sqrt{\int q_0(\lambda)\bar{I}_0(\lambda,\sigma_{\mathrm{s}})A(\lambda)\epsilon(\lambda)\mathrm{d}\lambda}}, \tag{109}$$

where $\bar{I}_0(\lambda,\sigma_{\mathrm{s}})$ is the spectrum of the x-ray source normalized by I_0, $A(\lambda)$ is the absorption $(A_{\mathrm{air}}(R_2,\lambda)A_{\mathrm{sub}}(\lambda)^2 A_{\mathrm{sample}}(\lambda))$, and $\epsilon(\lambda)$ is the detection efficiency. Other random errors that have a standard deviation of ε can be taken into account in the same manner as Eq. (58):

$$\Delta\Psi = \frac{\sqrt{\int q_0(\lambda)I_{\mathrm{total}}\Delta S\bar{I}_0(\lambda,\sigma_{\mathrm{s}})A(\lambda)\epsilon(\lambda)\mathrm{d}\lambda + \varepsilon^2}}{\int q_1(\lambda,\sigma_{\mathrm{s}})I_{\mathrm{total}}\Delta S\bar{I}_0(\lambda,\sigma_{\mathrm{s}})A(\lambda)\epsilon(\lambda)\mathrm{d}\lambda}\frac{1}{\sqrt{2}}. \tag{110}$$

The factor $\eta_{\mathrm{poly},\Psi}$ may be increased so long as the width of the energy band stays in a specific range. For example, as discussed in a previous paper [9], when $p = 1/2$ and $\bar{\lambda} - \bar{\lambda}/2 < \lambda < \bar{\lambda} + \bar{\lambda}/2$, where $\bar{\lambda}$ is the average wavelength of the spectrum and the moiré fringe is constructively added to that at $\bar{\lambda}$. A broader bandwidth, however, affects the accuracy of $\Psi(x,y)$. From Eqs. (16), (17), and (69), we get

$$\frac{\partial\varphi_{x_{\mathrm{s}}}}{\partial\lambda} = -\frac{r_{\mathrm{e}}\bar{\lambda}}{\pi}\int\frac{\partial\rho(x_{\mathrm{s}},y_{\mathrm{s}},z)}{\partial x_{\mathrm{s}}}\mathrm{d}z \tag{111}$$

$$= \frac{2}{\lambda}\varphi_{x_{\mathrm{s}}}. \tag{112}$$

Hence, $\Delta\lambda$, which is a deviation of λ from $\bar{\lambda}$, causes a systematic error $\Delta\Psi_\lambda$ proportional to Ψ:

$$\Delta\Psi_\lambda = \frac{2\Delta\lambda}{\bar{\lambda}}\Psi. \tag{113}$$

Considering that the first Fourier coefficient $q_1(\lambda)$ is used in Eq. (20) for a fringe scan, the systematic error in Ψ due

to a finite bandwidth is, however, canceled out when the spectrum of $q_1(\lambda)I_0(\lambda,\sigma_s)A(\lambda)\epsilon(\lambda)$ is symmetric around $\bar{\lambda}$.

Even when the spectrum is symmetric around $\bar{\lambda}$, adding the intensity of a moiré fringe will be destructive for a too large Ψ [see Eq. (113)]. This fact determines an upper limit of Ψ that can be measured. Remember that $\eta_{\mathrm{poly},\Psi}$ determines a lower limit of Ψ. These conditions determine a dynamic range of detection of $(D)_{\mathrm{DL}}$. That is, if random errors ε in Eq. (110) are negligible,

$$\frac{\Delta_{\mathrm{Det}}}{\eta'_{\mathrm{poly}}\sqrt{I_{\mathrm{total}}\Delta S}} \leqslant (D)_{\mathrm{DL}} \leqslant \frac{\pi}{4}\frac{\Delta_{\mathrm{Det}}}{pd_2 r_e \Delta\lambda}, \qquad (114)$$

where

$$\eta'_{\mathrm{poly}} \equiv pd_2 r_e \bar{\lambda}\,\eta_{\mathrm{poly},\Psi}. \qquad (115)$$

The upper limit was obtained from $\Delta\Psi_\lambda < \pi/2$, and it is a condition for a moiré fringe not to be smeared out. It may be more convenient to write Eq. (114) as

$$C_s\frac{1}{\eta_{\mathrm{poly},\Psi}\sqrt{I_{\mathrm{total}}\Delta S}} \leqslant (D)_{\mathrm{DL}} \leqslant C_s\frac{\pi}{4}\frac{\bar{\lambda}}{\Delta\lambda}, \qquad (116)$$

where

$$C_s \equiv \frac{\Delta_{\mathrm{Det}}}{pd_2 r_e \bar{\lambda}}. \qquad (117)$$

6. SUMMARY

We investigated how to efficiently use incident photons in phase imaging using cone-beam XTI. In Section 3, it was analytically shown how photon statistics cause errors in the phase of the moiré fringe and consequently impose a detection limit on the wavefront gradient. This result was verified using Monte Carlo simulations. From the relation between the incident photon number and the detection limit, a figure of merit of XTI [Eq. (67)] for a monochromatic cone beam, η, was defined as a measure of efficiency.

Section 4 analyzed the dependence of the figure of merit on the parameters of the optical system, such as the pitch and position of the two gratings, for a given wavelength. It was shown that for a given x-ray source size, the two gratings have optimum positions that are independent of parameters related to the grating's shape. We also demonstrated how the figure of merit depends on the parameters of the shape of the second grating for two typical shapes (rectangle and trapezoid). For a given pattern height of the second grating, there is an optimal transmission function that maximizes the figure of merit η. In fact, a rectangular second grating with an optimal linewidth provides the best transmission function (see Appendix C). From these results, it was shown that there is an upper limit to the figure of merit η for a given wavelength; the detection limits of $\Delta_{\mathrm{Det},s}\Delta\varphi_{x_s}$ and $(D)_{\mathrm{DL}}$ are given by Eqs. (92) and (93).

Section 5 discussed the design of a practical optical system for monochromatic cone-beam XTI given certain boundary conditions. The way to design the optical sys-

tem of the Talbot–Lau interferometer was also discussed. The discussion for a monochromatic cone beam was extended into the case of a polychromatic cone beam. This extension will be useful for designing the optical system of polychromatic cone-beam XTI.

APPENDIX A: FIGURE OF MERIT AND OPTIMUM CONDITIONS FOR A COMPRESSED SELF-IMAGE

When a self-image with a compression ratio of α (being an integer), i.e., with a period of $1/\alpha\,d_1 R_2/R_1$, is generated and an absorption grating with the same pitch (d_2) is superimposed on the self-image to generate a moiré fringe,

$$\frac{\alpha d_2}{d_1} = \frac{R_2}{R_1}. \qquad (A1)$$

Equation (14) is replaced with

$$I(x,y,p)/I_0\Delta St \approx \frac{1}{R_2^2}\sum_n \mu_{\alpha n}b_{\alpha n}(z_{12})c_{\alpha n}$$
$$\times\exp\left[i2\pi n\left(\frac{-y\theta+\chi}{d_2} + \zeta_s\varphi_{x_s}(x,y)\right)\right], \qquad (A2)$$

where

$$\zeta_s = \begin{cases} \alpha pd_s/\lambda & (R_s \leqslant R_1) \\ z_s/d_2 & (R_s \geqslant R_1) \end{cases}, \qquad (A3)$$

and d_s is defined by $\alpha d_2/\kappa_{2,s}$ (the other definitions are the same as those in Section 2). Here we have taken into account the fact that the compressed self-image is constructed by the $(n\alpha)$th Fourier coefficient. Then Δx_s in Eq. (22) is given by αpd_s instead of pd_s. Equations (64) and (65) are replaced as follows:

$$\Delta_{\mathrm{Det},s}\Delta\varphi_{x_s} = \frac{\lambda\Delta_{\mathrm{Det},s}}{2\pi\alpha pd_s\eta_\Psi\sqrt{I_{\mathrm{total}}\Delta S}}, \qquad (A4)$$

$$= \frac{\lambda\Delta_{\mathrm{Det}}}{2\pi\alpha p(\alpha d_2)\eta_\Psi\sqrt{I_{\mathrm{total}}\Delta S}}. \qquad (A5)$$

Therefore η should be defined by

$$\eta = \frac{\alpha^2 pd_2}{\lambda}\sqrt{2}\frac{q_\alpha}{\sqrt{q_0}} \qquad (A6)$$

$$= \left(\frac{\alpha^2 pd_2}{\lambda}\frac{\mu_\alpha}{R_2}\right)\frac{b_\alpha}{\sqrt{b_0}}\frac{c_\alpha}{\sqrt{c_0}}\sqrt{2}, \qquad (A7)$$

instead of Eq. (67). The first factor is given by

$$\frac{\alpha^2 pd_2}{\lambda}\frac{\mu_\alpha}{R_2} = \frac{1}{2\pi\sigma_s}\frac{\alpha p}{k}\exp\left[-\left(\left(\frac{\alpha p}{k}\right)^2/2\right)\right], \qquad (A8)$$

where

$$k = \frac{\lambda R_2}{2\pi\sigma_s \alpha d_2}. \tag{A9}$$

This factor has a maximum at

$$k = \alpha p, \tag{A10}$$

and Eqs. (84)–(88) are replaced as follows:

$$R_{1,\mathrm{op}} = \alpha^2 \frac{p}{\lambda} \frac{1}{1/(2\pi\sigma_s d_2) + 1/(2\pi\sigma_s)^2}, \tag{A11}$$

$$R_{2,\mathrm{op}} = \alpha^2 2\pi\sigma_s d_2 \frac{p}{\lambda}, \tag{A12}$$

$$\left(\frac{\alpha^2 p d_2}{\lambda}\frac{\mu_\alpha}{R_2}\right)_{\max} = \frac{\exp\left(-\dfrac{1}{2}\right)}{2\pi\sigma_s}, \tag{A13}$$

$$d_1 = \alpha \frac{d_2}{1 + d_2/2\pi\sigma_s}. \tag{A14}$$

The curve $(\alpha^2 p d_2/\lambda)\,(\mu_\alpha/R_2)$ as a function of $k/\alpha p$ [Eq. (A8)] is plotted in Fig. 4. The maximum of η is proportional to $\exp(-\tfrac{1}{2})/(2\pi\sigma_s)$ and is independent of α, p, d_2, and λ. For example, if a π phase grating is used as reported in [20,31], and the self-image at $p = 1/8\,(\alpha = 2)$ is used, $R_{1,\mathrm{op}}$, $R_{2,\mathrm{op}}$, and d_1/α [Eqs. (A11), (A12), and (A14)] are the corresponding curves in Fig. 5(a).

APPENDIX B: WHEN R_1 IS PROPORTIONAL TO $1/\Lambda$ OR P/Λ

If the distance between the source and the first grating is inversely proportional to λ, d_2/d_1 is independent of λ. Substituting $R_1 = C/\lambda$, where C is a constant that is independent of λ, into Eq. (11), we can obtain the following form:

$$\frac{d_2}{d_1} = \frac{C}{C - p d_1^2}. \tag{B1}$$

Then d_2/d_1 is independent of λ. Similarly, if R is proportional to p/λ, d_2/d_1 becomes independent of both p and λ. Substituting $R_1 = C' p/\lambda$, where C' is a constant that is independent of p and λ, into Eq. (11), we can obtain the following form:

$$\frac{d_2}{d_1} = \frac{C'}{C' - d_1^2}. \tag{B2}$$

For example, if R_1 is given by Eq. (84), d_2/d_1 can be written as follows:

$$\frac{d_2}{d_1} = \frac{1}{1 - d_1/2\pi\sigma_s}. \tag{B3}$$

APPENDIX C: BEST TRANSMISSION FUNCTION OF ABSORPTION GRATING

We show that a second grating with a rectangular shape and with a linewidth of $l_{2,\mathrm{op}}$ [the solution to Eq. (90)] provides the best transmission function if the maximum pattern height $h_{2,\max}$ of the second grating is given. The best transmission function means that it maximizes $c_1/\sqrt{c_0}$, where c_0 and c_1 are the 0th- and first-order Fourier coefficients of the transmission function of the second grating [see Eq. (67)].

What needs be proved is this: *For given c_0 and $h_{2,\max}$ the best shape of the absorption grating that maximizes $c_1/\sqrt{c_0}$ is rectangular*. Then from Subsection 4.B, it can be shown that $c_1/\sqrt{c_0}$ of the grating with a linewidth $l_{2,\mathrm{op}}$ is better than any other rectangular grating. For simplicity, we consider a case where $h_{2,\max} = \infty$ $(T(h_{2,\max}) = 0)$. The following consideration can be applied to the case where $h_{2,\max}$ is finite.

Without loss of generality (including the case of an asymmetric grating), the 0th- and the first-order Fourier coefficients, c_0 and c_1, can be expressed by

$$c_0 = \frac{1}{d_2}\int_{-d_2/2}^{d_2/2} T(x)\mathrm{d}x, \tag{C1}$$

$$c_1 = \frac{1}{d_2}\int_{-d_2/2}^{d_2/2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x. \tag{C2}$$

Equation (C2) is correct when the origin of x is chosen so that $T(x)$ satisfies

$$\int_{-d_2/2}^{d_2/2} T(x)\sin 2\pi\frac{x}{d_2}\mathrm{d}x = 0. \tag{C3}$$

We define $T_{\mathrm{rect}}(x)$ to be a transmission function with a period d_2:

$$T_{\mathrm{rect}}(x) = \begin{cases} 1 & \left(-\dfrac{c_0}{2}d_2 \leq x \leq \dfrac{c_0}{2}d_2\right) \\[2mm] 0 & \left(-\dfrac{d_2}{2} < x < -\dfrac{c_0}{2}d_2,\ \dfrac{c_0}{2}d_2 < x < \dfrac{d_2}{2}\right) \end{cases}. \tag{C4}$$

The 0th-order Fourier coefficient is given by c_0. The first-order Fourier coefficient $c_{1,\mathrm{rect}}$ of $T_{\mathrm{rect}}(x)$ is given by

$$c_{1,\mathrm{rect}} = \frac{1}{d_2}\int_{-(c_0/2)d_2}^{(c_0/2)d_2} \cos 2\pi\frac{x}{d_2}\mathrm{d}x. \tag{C5}$$

Next, $c_{1,\mathrm{rect}}$ is compared with the first Fourier coefficient c_1 of any transmission function $T(x)$, the 0th-order Fourier coefficient of which is c_0:

$$c_{1,\mathrm{rect}} - c_1 = \frac{1}{d_2}\left[\int_{-(c_0/2)d_2}^{(c_0/2)d_2} \cos 2\pi\frac{x}{d_2}\mathrm{d}x\right.$$

$$\left.- \int_{-d_2/2}^{d_2/2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x\right] \qquad (C6)$$

$$= \frac{1}{d_2}\left[\int_0^{(c_0/2)d_2} (1 - T(x))\cos 2\pi\frac{x}{d_2}\mathrm{d}x\right.$$

$$\left.- \int_{(c_0/2)d_2}^{d_2/2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x\right]$$

$$+ \frac{1}{d_2}\left[\int_{-(c_0/2)d_2}^0 (1 - T(x))\cos 2\pi\frac{x}{d_2}\mathrm{d}x\right.$$

$$\left.- \int_{-d_2/2}^{-(c_0/2)d_2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x\right]. \qquad (C7)$$

The terms in the square bracket of the first line in Eq. (C7) are given by

$$\int_0^{(c_0/2)d_2} (1 - T(x))\cos 2\pi\frac{x}{d_2}\mathrm{d}x - \int_{(c_0/2)d_2}^{d_2/2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x$$

$$\geq \cos 2\pi\frac{c_0}{2}\int_0^{(c_0/2)d_2} (1 - T(x))\mathrm{d}x$$

$$- \cos 2\pi\frac{c_0}{2}\int_{(c_0/2)d_2}^{d_2/2} T(x)\mathrm{d}x = \cos 2\pi\frac{c_0}{2}\left(\frac{c_0}{2} - c_{0,+}\right)d_2,$$

$$(C8)$$

where

$$c_{0,+} \equiv \frac{1}{d_2}\int_0^{d_2/2} T(x)\mathrm{d}x. \qquad (C9)$$

Here we used $0 \leq T(x) \leq 1$. Similarly, The terms in the square bracket of the second line in Eq. (C7) are given by

$$\int_{-(c_0/2)d_2}^0 (1 - T(x))\cos 2\pi\frac{x}{d_2}\mathrm{d}x - \int_{-d_2/2}^{(-c_0/2)d_2} T(x)\cos 2\pi\frac{x}{d_2}\mathrm{d}x$$

$$\geq \cos 2\pi\frac{c_0}{2}\int_{-(c_0/2)d_2}^0 (1 - T(x))\mathrm{d}x$$

$$- \cos 2\pi\frac{c_0}{2}\int_{-d_2/2}^{-(c_0/2)d_2} T(x)\mathrm{d}x$$

$$= \cos 2\pi\frac{c_0}{2}\left(\frac{c_0}{2} - c_{0,-}\right)d_2, \qquad (C10)$$

where

$$c_{0,-} \equiv \frac{1}{d_2}\int_{-d_2/2}^0 T(x)\mathrm{d}x. \qquad (C11)$$

Hence

$$c_{1,\mathrm{rect}} - c_1 \geq 0. \qquad (C12)$$

Here we used $c_{0,+} + c_{0,-} = c_0$. Therefore, the first-order Fourier coefficient of a rectangular transmission function is better than any other transmission function for a given c_0; that is, a rectangular transmission function provides the maximum $c_1/\sqrt{c_0}$ for a given c_0.

ACKNOWLEDGMENTS

REFERENCES AND NOTES
1. R. Fitzgerald, "Phase-sensitive x-ray imaging," Phys. Today **53**, 23–26 (2000).
2. A. Momose, "Recent advances in x-ray phase imaging," Jpn. J. Appl. Phys., Part 1 **44**, 6355–6367 (2005).
3. A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki, "Demonstration of x-ray Talbot interferometry," Jpn. J. Appl. Phys., Part 1 **42**, L866–L868 (2003).
4. T. Weitkamp, B. Nöhammer, A. Diaz, and C. David, "X-ray wavefront analysis and optics characterization with a grating interferometer," Appl. Phys. Lett. **86**, 54101–54103 (2005).
5. A. Momose and S. Kawamoto, "X-ray Talbot interferometry with capillary plates," Jpn. J. Appl. Phys., Part 1 **45**, 314–316 (2006).
6. T. Weitkamp, A. Diaz, B. Nöhammer, F. Pfeiffer, T. Rohbeck, P. Cloetens, M. Stampanoni, and C. David, "Hard x-ray phase imaging and tomography with a grating interferometer," Proc. SPIE **5535**, 137–142 (2004).
7. A. Momose, S. Kawamoto, I. Koyama, and Y. Suzuki, "Phase tomography using an x-ray Talbot interferometer," Proc. SPIE **5535**, 352–360 (2004).
8. T. Weitkamp, A. Diaz, C. David, F. Pfeiffer, M. Stampanoni, P. Cloetens, and E. Ziegler, "Quantitative x-ray phase imaging with a grating interferometer," Opt. Express **13**, 6296–6304 (2005).
9. A. Momose, W. Yashiro, Y. Takeda, Y. Suzuki, and T. Hattori, "Phase tomography by x-ray Talbot interferometry for biological imaging," Jpn. J. Appl. Phys., Part 1 **45**, 5254–5262 (2006).
10. A. Momose, W. Yashiro, M. Moritake, Y. Takeda, K. Uesugi, A. Takeuchi, Y. Suzuki, M. Tanaka, and T. Hattori, "Biomedical imaging by Talbot-type x-ray phase tomography," Proc. SPIE **6318**, 63180T (2006).
11. M. Engelhardt, J. Baumann, M. Schuster, C. Kottler, F. Pfeiffer, O. Bunk, and C. David, "High-resolution differential phase contrast imaging using a magnifying projection geometry with a microfocus x-ray source," Appl. Phys. Lett. **90**, 224101 (2007).
12. C. Kottler, C. David, F. Pfeiffer, and O. Bunk, "'A two-directional approach for grating based differential phase contrast imaging using hard x-rays," Opt. Express **15**, 1175–1181 (2007).
13. H. F. Talbot, "Facts relating to optical science," Philos. Mag. **9**, 401–407 (1836).
14. K. Patorski, "The self-imaging phenomenon and its applications," in *Progress in Optics XXVII* (Elsevier, 1989).
15. P. Cloetens, J. P. Guigay, C. De Martino, and J. Baruchel, "Fractional Talbot imaging of phase gratings with hard X rays," Opt. Lett. **22**, 1059–1061 (1997).
16. E. W. Becker, W. Ehrfeld, P. Hagmann, A. Maner, and D. Münchmeyer, "Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron radiation lithography, galvanoforming, and plastic

moulding (LIGA process)," Microelectronics **4**, 35–56 (1986).

17. Y. Utsumi, T. Kishimoto, T. Hattori, and H. Hara, "Large-area x-ray lithography system for LIGA process operating in wide energy range of synchrotron radiation," Jpn. J. Appl. Phys., Part 1 **44**, 5500–5504 (2005).

18. M. Matsumoto, K. Takiguchi, M. Tanaka, Y. Hunabiki, H. Takeda, A. Momose, Y. Utsumi, and T. Hattori, "Fabrication of diffraction grating for x-ray Talbot interferometer," in *High Aspect Ratio Micro Structure Technology Workshop 2005* (Harmst, 2005), p. 22.

19. M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge U. Press, 1999).

20. F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase-retrieval and differential phase-contrast imaging with low-brilliance x-ray sources," Nat. Phys. **2**, 258–261 (2006).

21. T. Weitkamp, C. David, C. Kottler, O. Bunk, and F. Pfeiffer, "Tomography with grating interferometers at low-brilliance sources," Proc. SPIE **6318**, 63180S (2006).

22. F. Pfeiffer, C. Kottler, O. Bunk, and C. David, "Hard x-ray phase tomography with low-brilliance sources," Phys. Rev. Lett. **98**, 108105 (2007).

23. J. T. Winthrop and C. R. Worthington, "Theory of Fresnel images. I. Plane periodic objects in monochromatic light," J. Opt. Soc. Am. **55**, 373–381 (1965).

24. J. P. Guigay, "On Fresnel diffraction by one-dimensional periodic object, with application to structure determination of phase object," Opt. Acta **18**, 677–682 (1971).

25. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, "Digital wavefront measuring interferometer for testing optical surfaces and lenses," Appl. Opt. **13**, 2693–2703 (1974).

26. O. Kafri and I. Glatt, *The Physics of Moiré Metrology* (Wiley-Interscience, 1990).

27. J. H. Bruning, *Optical Shop Testing*, D. Malacara, ed. (Wiley-Interscience, 1978).

28. M. Takeda, "Spatial-carrier fringe-pattern analysis and its applications to precision interferometry and profilometry: An overview," Ind. Metrology **1**, 79–99 (1990).

29. The condition $M \geq 3$ is required for Eq. (20) and the condition $M \geq 4$ is required for Eq. (49). For this reason the result Eq. (57) is correct for $M \geq 4$. If $M = 3$, $\Delta\Psi$ depends on φ_{x_s} because the term of q_{-1} is not negligible in Eq. (47).

30. We assumed that the distribution of intensity from each groove is roughly represented as a Gaussian function of a standard deviation of σ_s. Then the full width at half-maximum (FWHM) of it is approximately given by $2.35\sigma_s$.

31. V. Arrizón and E. López-Olazagasti, "Binary phase grating for array generation at 1/16 of Talbot length," J. Opt. Soc. Am. A **12**, 801–804 (1995).