Research Article

Better 3D inspection with structured illumination: signal formation and precision

ZHENG YANG, ALEXANDER KESSEL, AND GERD HÄUSLER*

Institute of Optics, Information and Photonics, Friedrich-Alexander-University Erlangen-Nuremberg, Staudtstrasse 7/B2, 91058 Erlangen, Germany
Corresponding author: Gerd.Haeusler@physik.uni-erlangen.de

3D metrology faces increasing demands for higher precision and larger space-bandwidth-speed product (number of 3D points/s). In this paper we consider structured-illumination microscopy as a means for satisfying these demands, developing a theoretical model of the signal formation for both optically smooth and optically rough surfaces. The model allows us to investigate physical limits on precision and to establish rules that allow sensor parameter optimization for greatest precision or highest speed.   © 2015 Optical Society of America

OCIS codes: (120.3940) Metrology; (120.4630) Optical inspection; (110.2990) Image formation theory; (030.4280) Noise in imaging systems; (030.6140) Speckle.

1. INTRODUCTION

This paper is motivated by the demanding requirements of 3D inspection systems for electronic interconnects, where both optically rough and specular surfaces occur at the same sample, the required precision is around 1 μm, the required lateral resolution is in the 5 μm regime, and the depth of the measuring field is in the 1 mm range. Both white-light interferometry (WLI) and confocal microscopy (CM) satisfy these requirements, but industry has an additional need: the 3D sensor should deliver more than 10^8 3D points/s. No sensor satisfies all these requirements at the same time. In this paper we explore an alternative method, structured-illumination microscopy (SIM), which satisfies the original four requirements and in addition has the potential for very high speed. We begin by investigating the limits of precision for SIM. To do so, we develop a model of the physical signal formation for both smooth and rough surfaces. From this model, rules are derived to optimize the sensor parameters for precision or, alternatively, for speed.

2. BACKGROUND

Optical 3D sensors are now widely used in the inspection of precision industrial products, the predominant application being in the measurement of the height profile of an object. There is a wide spectrum of commercially available sensors with sufficient precision for many applications. The major issues of concern today are a large lateral field, high lateral resolution, and speed. Assuming that the field size is $\Delta x \Delta y$ m^2 and the bandwidth is $1/(\delta x \delta y)$ m^{-2} (δx, δy: the smallest laterally resolvable structure size in the x and y directions), a single

3D height map will include $(\Delta x \Delta y)/(\delta x \delta y)$ 3D points. The speed H/s specifies that H height maps can be acquired per second, and the "space-bandwidth-speed product" SBSP $= H(\Delta x \Delta y)/(\delta x \delta y)$ (1/s) thus characterizes the number of acquired 3D points per second. Of course, it is not just the SBSP that makes a sensor useful, precision δz and measuring range Δz both being important. The channel capacity (CC) of the sensor, given by Eq. (1), gives a useful measure (in bit/s) of sensor performance in terms of speed and dynamic range [1]:

$$CC = SBSP \log_2 \left(1 + \frac{\Delta z}{\delta z} \right). \qquad (1)$$

It should be emphasized that all 3D sensors need E ($E \geq 1$) exposures to achieve one height map. The information efficiency $1/E$ is thus generally less than unity [2,3].

We now discuss candidates for the application mentioned above, to understand the origin of the SBSP limit or the CC limit.

Not discussed will be the paradigm of 3D metrology: triangulation with fringe projection [4]. This method cannot, generally, measure specular surfaces. Moreover, it lacks depth of field for high lateral resolution requirements.

White light interferometry [5,6] is nearly ideal in terms of precision and lateral resolution, being capable of measuring both smooth and rough surfaces [7] with a field of view of up to 200 mm [8]. As a unique feature, the precision is independent of the imaging aperture and from the working distance, and narrow boreholes can be thus measured [9]. Smooth surfaces can be measured with nanometer precision, while the precision on rough surfaces corresponds to the surface

roughness [7]. The principal disadvantage of WLI is that it is slow. The depth scan has to acquire the so-called correlogram, which is modulated with a period of $\lambda/2 - 400$ nm. Even if the sampling theorem may be violated to a certain extent [10], the large number E of necessary z-steps reduces the efficiency dramatically and limits the speed of the z-scan to a few hundred micrometers per second, severely limiting the achievable SBSP. Hýbl and co-workers [2,3,11] has proposed a means for overcoming this limitation, but the method has not yet been commercially implanted.

Confocal microscopy in its different implementations [12–16] can be much faster than WLI in depth scan because there is no high-frequency carrier signal. So CM might be a candidate for high-speed measurement.

In SIM, a pattern (commonly a sinusoidal fringe) is projected onto the object surface through a microscope objective (see Fig. 1). The illuminated surface is imaged by the same objective onto an image detector in the conjugate image plane. While the object is scanned in depth (z-scan), the local fringe contrast $C(x, y, z)$ is measured, e.g., via phase shifting [17,18]. From the maximum of the contrast curve $C(x, y, z)$, the surface shape $z(x, y)$ can be found. The major advantage of SIM, compared to WLI, is the scanning speed: only three sampling points over the width (FWHM) of the contrast curve are necessary to interpolate the location of the maximum.

The contrast variation along the depth direction of incoherently projected sinusoidal fringe patterns was first exploited for the acquisition of 3D data, as described in [19]. This concept was subsequently improved by phase shifting [17,18]. Extension of the depth of field or optical sectioning are obvious applications of this method.

However, this method did not receive much attention until Gustafsson et al. [20] demonstrated the twofold improvement of lateral resolution. They used coherent generation of sinusoidal fringes close to the diffraction limit, in connection with fluorescence (avoiding any speckle noise). Today, different

modifications of SIM for 3D metrology are commercially available, not just for biological samples but also for technical surfaces. In [21], instead of a sinusoidal grating, multiple thin lines are projected and laterally shifted during the depth scan. The local height is calculated from the (maximal) gradient of the observed line intensity. Schwertner [22] avoid the lateral line scan by using a full-field sinusoidal grating. The contrast curve is calculated from two exposures (with phase 0 and π) at each z-position. A similar two-phase stepping method was used by Kristensson et al. for fast imaging of spray, injected into a combustion chamber [23]. In [24], a spatial convolution technique is applied to calculate the focus information.

To our knowledge, despite the commercial implementations of SIM, there is no theoretical model in the technical literature that describes the limits of precision for specular and rough surfaces. Our intent in this and a subsequent paper is to help users build "better" SIM systems through the use of such a model. In this paper we develop the theory of the signal generation and determine the physical limits on precision. In the second paper, which is devoted to measurements with large SBSPs, we explain how to avoid the crucial stop-and-go depth scan and at the same time achieve significantly improved information efficiency for speed improvement. Eventually a novel lateral scanning concept will be demonstrated that allows for high-speed SIM with virtually arbitrary field size and with no sacrifice of lateral and longitudinal resolution.

3. SIGNAL FORMATION

SIM is in essence an active focus searching method. As illustrated in Fig. 1, a sinusoidal pattern is displayed by a ferroelectric-liquid-crystal-on-silicon (FLCoS) device with SXGA format (1280 × 1024 pixels) and projected onto the object with an incident light microscope with Leica apochromatic microscope objectives (white-light illumination). The focus plane of the observation is conjugate to the focus plane of the projection, where a CCD camera with 1392 × 1040 pixels is used as the detection unit. The maximum contrast of the grid image is highest when the object surface is in focus. Upon scanning the object in the z-direction, we acquire M phase-shifted grating images [25] at each z-position allowing the calculation of the contrast function $C(z)$ given by

$$C(I_1, I_2 \ldots I_M)$$

$$= \frac{2\sqrt{\left(\sum_{i=1}^{M} I_i \cos \frac{(i-1)2\pi}{M}\right)^2 + \left(\sum_{i=1}^{M} I_i \sin \frac{(i-1)2\pi}{M}\right)^2}}{\sum_{i=1}^{M} I_i}.$$

$$(2)$$

At least three images with three phase steps of 120° are necessary for a local contrast evaluation. In practice, four phase shifts are commonly used to reduce artifacts caused from residual second harmonics of a nonlinear fringe [26]. Higher phase stepping number is necessary, if further harmonics are present [27]. With $M = 4$, Eq. (2) assumes the form

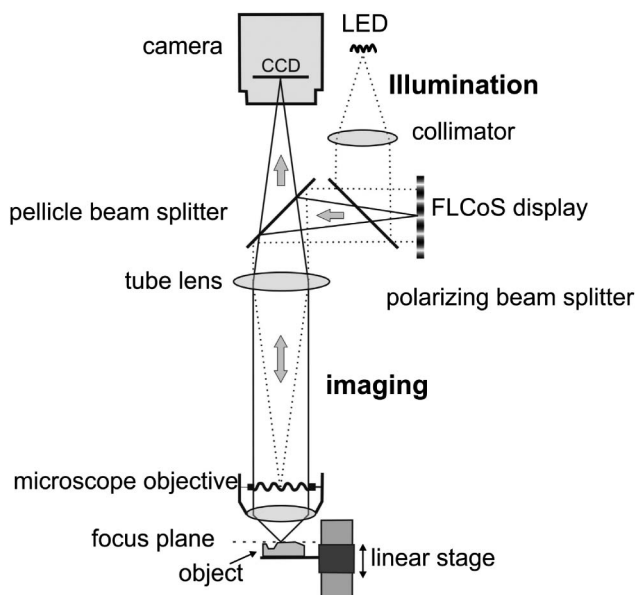$$C(I_1, I_2, I_3, I_4) = \frac{2\sqrt{(I_1 - I_3)^2 + (I_2 - I_4)^2}}{I_1 + I_2 + I_3 + I_4}. \qquad (3)$$

Fig. 1. SIM implemented in a bright-field microscope.

Eventually, the local "height" $z(x, y)$ in each pixel is calculated by localizing the peak of the contrast curve $C(z)$.

To investigate the limits of SIM, we first deal with the physics of the signal formation and derive the relation between the contrast curve and sensor parameters.

The model of the signal formation is derived by means of scalar diffraction theory and linear systems theory. In Fig. 2, the setup of Fig. 1 is unfolded. For easier understanding, the complete observation system is mirrored with respect to the object. Depending on the surface type (specular or scattering), the contrast has to be analyzed differently. For a specular surface, such as a planar mirror, a mirror image of the projected sinusoidal grating is imaged onto the detector. If the object is out of focus by a distance z_o, the observed grating is defocused with effective defocusing distance $2z_o$. The effective pupil function of the total system (projection and observation) including defocusing distance z_o is

$$P^{12}_{\text{spec}}(\nu, \mu, z_o) = u_o \, \exp[i\pi\lambda 2z_o(\nu^2 + \mu^2)], \tag{4}$$

where ν and μ denote spatial frequencies in the x and y directions. Note that the effective pupil function is just the product of the pupil functions for projection and observation, because reflection is a coherent process for a specular object.

The incoherent optical transfer function (OTF) [27] is the normalized autocorrelation of the pupil function P^{12}_{spec}:

$$\text{OTF}_{\text{spec}}(\nu', \mu', z_o) = \frac{B_{\text{spec}}(\nu', \mu', z_o)}{B_{\text{spec}}(0, 0, 0)}, \tag{5}$$

with

$$B_{\text{spec}}(\nu', \mu', z_o) = AC[P^{12}_{\text{spec}}(\nu, \mu, z_o)]. \tag{6}$$

For a specularly reflecting object, the magnitude of the OTF is the same as the contrast function C_{spec}:

$$C_{\text{spec}}(\nu', \mu', z_o) = |\text{OTF}_{\text{spec}}(\nu', \mu', z_o)|. \tag{7}$$

At a rough surface, as shown in Fig. 2 (bottom), the imaging process is different: a rough object, defocused by z_o, scatters a blurred grid image. Although intensity images observed by the camera suffer from coherent (speckle) noise, we nevertheless may assume significant incoherence in the contrast map. In particular, if the illumination aperture is filled with the image of the light source, the diameter of the spatial coherence function at the object is small, in accord with the Van-Cittert–Zernike

Fig. 2. Top: unfolded SIM setup for specular surface. Bottom: unfolded SIM setup for a scattering surface.

theorem. As a consequence, speckle contrast will be significantly reduced by defocusing during the depth scan. Furthermore, speckle noise is strongly correlated in phase-shifted images and will be largely reduced after the contrast calculation (see Section 4.B for details). We may therefore neglect speckle noise to a certain extent in the modeling of SIM with rough surfaces.

The pupil function for the projection is in this case given by

$$P^1_{\text{rough}}(\nu, \mu, z_o) = u_o \, \exp[i\pi\lambda z_o(\nu^2 + \mu^2)], \tag{8}$$

where the defocusing term now is z_o, instead of $2z_o$. The (incoherent) OTF of the projection is again given by the autocorrelation of the pupil function,

$$\text{OTF}_{\text{rough}}(\nu', \mu', z_o) = \frac{B_{\text{rough}}(\nu', \mu', z_o)}{B_{\text{rough}}(0, 0, 0)}, \tag{9}$$

with

$$B_{\text{rough}}(\nu', \mu', z_o) = AC[P^1_{\text{rough}}(\nu, \mu, z_o)]. \tag{10}$$

The total OTF is the square of the term in Eq. (9), yielding

$$C_{\text{rough}}(\nu', \mu', z_o) = |\text{OTF}_{\text{rough}}(\nu', \mu', z_o)|^2. \tag{11}$$

The formalism above is principally textbook linear systems theory. Hopkins [28] derived an analytical solution for the defocused OTF. Stokseth gave a more easy-to-use numerical approximation [29] that we will further exploit. The simplified contrast curves on both surfaces types are then given by

$$C_{\text{spec}}(\bar{\nu}, z_o) = \begin{cases} \left| 2(1 - 0.69\bar{\nu} + 0.0076\bar{\nu}^2 + 0.043\bar{\nu}^3) \right. \\ \left. \times \left[\frac{J_1(\alpha - 0.5\alpha)}{\alpha - 0.5\alpha} \right] \right| & |\bar{\nu}| < 2 \\ 0 & |\bar{\nu}| \geq 2 \end{cases} \tag{12}$$

$$C_{\text{rough}}(\bar{\nu}, z_o) = C_{\text{spec}}\left(\bar{\nu}, \frac{z_o}{2}\right)^2, \tag{13}$$

with $\bar{\nu} = 2(\nu'^2 + \mu'^2)^{1/2}/\nu_{\text{cutoff}}$ and $\alpha = (2\pi/\lambda)\sin^2(2uz_o\bar{\nu})$, where $\bar{\nu}$ denotes the normalized spatial frequency $0 \leq |\bar{\nu}| \leq 2$, J_1 is the Bessel function of first kind first order, and u is the aperture angle in the object space.

Figure 3 displays contrast curves $C(z_o)$ for specular and rough surfaces. The contrast curves for a specular surface are narrower (the ratio is 35/52), have a higher maximum, and exhibit some ringing. The width of the contrast curves goes with the inverse square of the aperture, i.e., as the Rayleigh depth of field.

According to Laboureux and Häusler [30], the statistical uncertainty for the localization essentially depends on the curvature κ_0 at the maximum of $C(z_o)$ and on the noise. The curvature κ_0 can be mathematically calculated from the sensor parameters,

$$\kappa_0 \approx \frac{-3.921 C_{\text{max}} \sin^4 u\bar{\nu}(1 - 0.5\bar{\nu})^2}{\beta_1^2 \lambda^2}, \tag{14}$$

with C_{max} for the maximal contrast value of a contrast curve, $\beta_1 = 0.35$ for a specular object, and $\beta_1 = 0.52$ for a rough object.

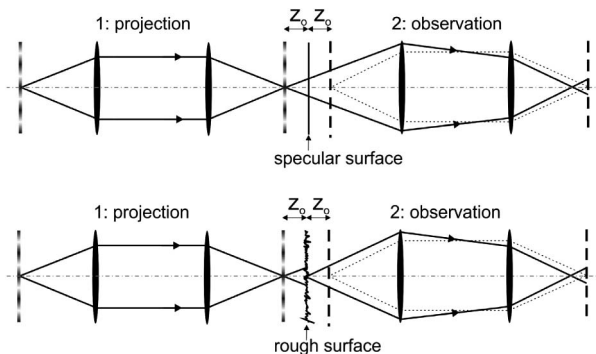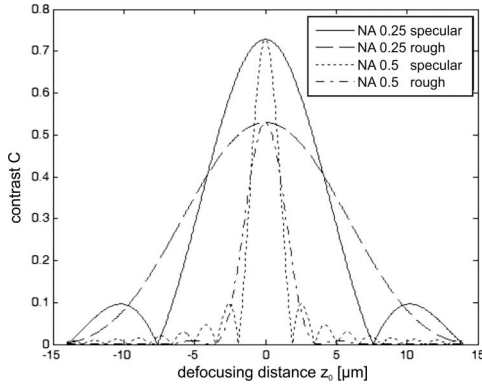**Fig. 3.** Theoretical contrast curves $C(z_o)$ of a fringe pattern with $\nu' = 0.2\nu_{cutoff}$, calculated by Eqs. (12) and (13), for a specular and for a rough surface, each for two apertures, NA = 0.5 and NA = 0.25.

It is obvious that a higher and narrower contrast curve (larger κ_0) is advantageous for better localization. This point is discussed in the following section.

4. PRECISION

Precision is important in metrology. It represents the closeness between the results of repeated measurements under stipulated conditions [31]. Precision can be measured in the space domain. A paradigm example is taking a single measurement of a planar mirror: assuming ergodicity, the precision is determined by calculating the standard deviation of the height map. For SIM, the precision is fundamentally limited by the localization uncertainty of the contrast maximum. Laboureux and Häusler [30] investigated the ultimate lower limit of the registration precision of two one-dimensional curves in the presence of noise. We adapted his results for our task. As denoted in Eq. (15), the uncertainty of registration is dependent on the power spectral density N_0 of the (white) noise, the curvature κ_0 at the curve maximum, and the full width of the data domain ΔT where the measurement values are selected for evaluation:

$$\sigma_z^2 = \frac{N_0}{1/12\kappa_0^2\Delta T^3} = \frac{\sigma_C^2\Delta s}{1/12\kappa_0^2\Delta T^3}, \quad (15)$$

where σ_C is the standard deviation of the contrast noise, Δs is the sampling distance in depth direction, and $N_0 = \sigma_C^2\Delta s$. The algorithms used in [30] and those for the height evaluation of SIM are similar; both methods use the least square method, so we can adapt Eq. (15) for SIM and derive a relation between the height uncertainty σ_z, the contrast noise level σ_C, and the sensor parameters. From Eqs. (14) and (15), we obtain

$$\sigma_z = \frac{\beta_2\sigma_C\lambda}{\sqrt{N}C_{max}\left(\frac{\Delta T}{FWHM}\right)\sin^2 u\overline{\nu}(1-0.5\overline{\nu})}, \quad (16)$$

where $\beta_2 = 0.31$ for a specular object, $\beta_2 = 0.46$ for a rough object, σ_C is the standard deviation of the contrast noise, and N is the number of selected measurement values to be input to the evaluation algorithm.

We move on now to understand the fundamental origins of the precision. For this purpose all possible noise sources are analyzed and the dominant source of noise is identified.

A. Specular Surfaces

At specular surfaces, the essential sources of noise are photon noise, electronic (camera) noise, statistical error of the translation stage, and nonlinearity of the sinusoidal pattern.

The first source is fundamental, whereas the other sources are caused by technology. Assuming that the technological errors are avoidable, the best achievable precision on specular surfaces is fundamentally limited by photon noise. The photon noise in all phase-shifted intensity images I_i from which the contrast is calculated is random and uncorrelated. The standard deviation σ_C of the contrast noise can be expressed by Eq. (17) (see Appendix A for detailed derivation). According to Eq. (17), σ_C^2 is inversely proportional to the squared SNR of the photon noise and the number M of phase shifts. Furthermore, it depends on the contrast C of the observed sinusoidal fringes:

$$\sigma_C^2 = \left(\frac{\partial C}{\partial I_1}\right)^2\sigma_{I_1}^2 + \left(\frac{\partial C}{\partial I_2}\right)^2\sigma_{I_2}^2 + ...\left(\frac{\partial C}{\partial I_M}\right)^2\sigma_{I_M}^2$$

$$\approx \frac{4}{SNR_{shot}^2 M}\left(1-\frac{C^2}{2}\right). \quad (17)$$

We combine Eqs. (16) and (17) and get the ultimate relation for the dependence of the height precision on sensor parameters. We use only the contrast values within the FWHM range (in other words, $\Delta T/FWHM = 1$):

$$\sigma_z = 0.62\frac{1}{\sqrt{N}}\frac{1}{\sqrt{M}}\frac{1}{SNR_{shot}}\frac{\lambda}{\sin^2 u}\frac{\sqrt{\frac{1}{C_{max}^2}-0.5}}{\overline{\nu}(1-0.5\overline{\nu})}. \quad (18)$$

In Fig. 4, the theoretical precision σ_z calculated from Eq. (18) is compared to measurements with various microscope objectives. The test sample was a planar mirror and the precision σ_z was
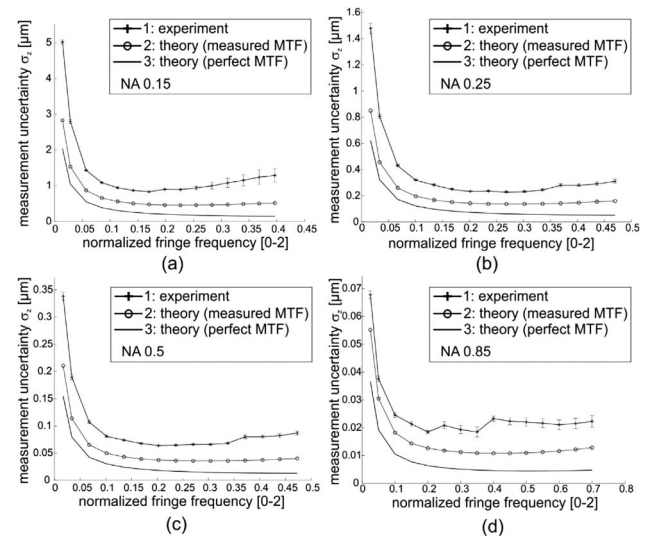

Fig. 4. Comparison of the theory of photon-noise limited measurement uncertainty and experimental results for specular objects. The experiments are performed on a planar mirror, with (a) $5 \times /0.15 -$, (b) $10 \times /0.25 -$, (c) $20 \times /0.5 -$, and (d) $50 \times /0.85 -$ microscope objective, where $\lambda = 0.5$ μm, $N = 3$, $M = 4$. A signal-to-noise ratio $SNR_{shot} \approx 100$ is assumed.

evaluated from nine patches each with 100×100 pixels. The precision is presented as a function of the chosen normalized fringe frequency $\bar{\nu}$. Curve 3 is calculated from Eq. ([18](#)) for a diffraction-limited objective. In this case, the maximum contrast $C_{\max}$ in Eq. ([18](#)) is the diffraction-limited modulation transfer function (MTF) from Eq. ([5](#)), for defocusing distance $z_o = 0$ and fringe frequency $\bar{\nu}$. Curve 2 is again calculated from Eq. ([18](#)); however, the (nondiffraction-limited) real maximum contrast $C_{\max}$ was measured and used for the calculation. Eventually, curve 1 displays the experimentally determined precision, taken from the measurement of a planar mirror.

Since the experimental setup is affected by additional noise sources, such as camera noise and motor noise, we are not surprised that the experiment (curve 1) displays greater uncertainty than does curve 2. However, the trend of the results displays considerable consistency between curve 1 and curve 2.

Figure [4(a)](#) shows the results for a numerical aperture NA = 0.15. The precision (for curve 1 and 2) exhibits a soft minimum for frequencies around 10% of the cutoff frequency, whereas the theoretical optimum frequency is at 25%. This is reasonable, since for the nondiffraction-limited objective the best precision is achieved for a lower frequency than for the diffraction-limited lens. It is useful to know that the minimum is not sharp. There is thus no fine tuning of the projected fringe frequency necessary. It is further noteworthy that the ultimate limit of the precision for the ideal lens is about 2× better than for the real lens.

As illustrated in Fig. [4(a)](#), a theoretical height uncertainty of around 200 nm can be achieved with a $5 \times /0.15$ microscope objective and standard video equipment. The precision can be greatly improved, down to a few nanometers, with higher NA: see, e.g., Fig. [4(d)](#). Further improvement is possible by more samples for the depth scan and more phase shifts.

If we calculate the contrast from N samples within the contrast function, and from M phase shifts, we profit from some averaging and get the factors $1/N^{1/2}$ and $1/M^{1/2}$.

The term $\lambda/\sin^2 u$ corresponds to the Rayleigh depth of field. The term $(1/C_{\max}^2 - 0.5)^{1/2}/[\bar{\nu}(1 - 0.5\bar{\nu})]$ shows a global minimum around one fourth of the cutoff frequency for diffraction-limited optics. This term is softly varying around the optimal frequency, so we can choose the fringe frequency within a comparably wide range.

B. Rough Surfaces

If the object surface displays a roughness greater than $\lambda/4$ within the width of the point spread function (PSF), speckle noise will be dominant [32], rather than photon noise. Because illumination and observation use the same lens, the illumination aperture cannot be larger than the observation aperture. Hence, there is significant spatial coherence, even with a large light source.

Speckle noise in each phase-shifted image I_i is highly correlated. This correlation must be taken into account in a quantitative evaluation of the precision. To keep the analysis as easy as possible, we rewrite Eq. ([3](#)) as Eq. ([19](#)), the contrast being written as a function of A and B:

$$C(I_1, I_2 ... I_M) = \frac{A - B}{A + B}, \tag{19}$$

where

$$A = \frac{1}{M}$$
$$\times \left(\sum_{i=1}^{M} I_i + 2\sqrt{\left(\sum_{i=1}^{M} I_i \cos \frac{(i-1)2\pi}{M}\right)^2 + \left(\sum_{i=1}^{M} I_i \sin \frac{(i-1)2\pi}{M}\right)^2} \right)$$
$$= I_{\text{mean}} + CI_{\text{mean}}, \tag{20}$$

and

$$B = \frac{1}{M}$$
$$\times \left(\sum_{i=1}^{M} I_i - 2\sqrt{\left(\sum_{i=1}^{M} I_i \cos \frac{(i-1)2\pi}{M}\right)^2 + \left(\sum_{i=1}^{M} I_i \sin \frac{(i-1)2\pi}{M}\right)^2} \right)$$
$$= I_{\text{mean}} - CI_{\text{mean}}. \tag{21}$$

In Eqs. ([20](#)) and ([21](#)), A and B are numerically equivalent to the maximal and minimal intensities of the sinusoidal fringes.

We know from the experiments that speckle noise in all phase-shifted images is highly correlated. So we assume in a first approach that it acts like spatial multiplicative noise $\rho(x, y)$, with a standard deviation $\sigma_\rho \in [0, 1]$. The intensity error ΔI_i can be expressed by

$$\Delta I_i = I_i \rho(x, y). \tag{22}$$

From Eq. ([22](#)), the random error ΔA and ΔB in A and B can be approximated as in Eq. ([23](#)):

$$\Delta A = A\rho(x, y), \quad \Delta B = B\rho(x, y). \tag{23}$$

Defining the speckle contrast $S_C = \sigma_\rho/I_{\text{mean}}$ ($I_{\text{mean}} = 1$) and taking into account Eqs. ([20](#)) and ([21](#)), the standard deviation of ΔA and ΔB can be written as follows:

$$\sigma_A = S_C(I_{\text{mean}} + CI_{\text{mean}}), \quad \sigma_B = S_C(I_{\text{mean}} - CI_{\text{mean}}). \tag{24}$$

According to the rule of the combined uncertainty from correlated input quantities, defined in guide to the expression of uncertainty in measurement (GUM) [33], the standard deviation σ_C of the contrast noise in Eq. ([19](#)) can be expressed by Eq. ([25](#)) (see Appendix [B](#) for detailed derivation):

$$\sigma_C^2 = \left(\frac{\partial C}{\partial A}\right)^2 \sigma_{I_A}^2 + \left(\frac{\partial C}{\partial B}\right)^2 \sigma_{I_B}^2 + 2\frac{\partial C}{\partial A}\frac{\partial C}{\partial B}\sigma_A\sigma_B r(\Delta A, \Delta B)$$
$$= 1/2 S_C^2 (1 - C^2)^2 (1 - r(\Delta A, \Delta B)). \tag{25}$$

We now switch to the realistic case, where the speckle noise is no longer perfectly correlated in the different phase-shifted images. The correlation coefficient between ΔA and ΔB is $r(\Delta A, \Delta B)$. For complete correlation ($r = 1$), the errors ΔA and ΔB cancel each other. In practice, r is dependent on the fringe frequency and has a value around 80%. It must be determined experimentally.

As in the case of Eq. ([18](#)) for specular surfaces, we obtain Eq. ([26](#)) for the height uncertainty at rough surfaces by a combination of Eqs. ([16](#)) and ([25](#)). Experiments show that the speckle noise of the contrast curve, at different depth positions, is highly correlated as well, so smoothing over N sampling points will not greatly affect the height uncertainty. In other

words, there is little advantage to sampling the contrast curve in many narrow steps. That is why there is no N-dependence of Eq. ([26](#)). This is an important result, since we aim for high speed:

$$\sigma_z = 0.2 S_C \frac{\lambda}{\sin^2 u} \frac{\frac{1}{C_{max}} - C_{max}}{\bar{\nu}(1 - 0.5\bar{\nu})} \sqrt{1 - r(\Delta A, \Delta B)}. \quad \textbf{(26)}$$

Equation ([26](#)) exhibits quite interesting consequences. First, the height precision σ_z depends on three terms, telling us something about the physics of SIM at rough surfaces. The well-known term $S_C \lambda / \sin 2u$ is generic for triangulation: Dorsch *et al.* [34] found that the depth precision in a laser triangulation sensor cannot be better than $S_C \lambda / (2\pi \sin^2 u)$. The term $(1/C_{max} - C_{max})/[\bar{\nu}(1 - 0.5\bar{\nu})]$ reflects the influence of the MTF of the lens. This term diverges for high fringe frequencies and is approximately constant for medium frequencies.

Now to the term $[1 - r(\Delta A, \Delta B)]^{1/2}$. Figure [5(a)](#) shows the measured correlation coefficient $r(\Delta A, \Delta B)$ as a function of the fringe frequency. In the medium range, we find $r \sim 0.9$. It might be noted that for $r = 0.93$, Eq. ([26](#)) delivers the same precision for SIM as Dorsch *et al.* [34] found for laser triangulation ($u = $ triangulation angle):

$$\sigma_z \approx \frac{S_C \lambda}{2\pi \sin^2 u}. \quad \textbf{(27)}$$

To validate the theory leading to Eq. ([26](#)), SIM measurements of a roughness standard with $R_a = 0.8$ μm were performed with a microscope objective $5 \times /0.15$. The standard deviation σ_z was evaluated from nine height patches as in the experiments shown in Fig. [4](#). The results (see curve 1) are compared with the theory (see curve 2) in Fig. [5(b)](#). The correlation coefficient $r(\Delta A, \Delta B)$ was experimentally determined, as shown in Fig. [5(a)](#). As mentioned, speckle noise is highly correlated (for lower frequencies, other noise sources become more important, and the correlation drops in this region). To calculate the theoretical precision, C_{max} (~MTF) and the speckle contrast $S_C \sim 0.6$ were measured and the values were put into Eq. ([26](#)). The comparison shows that theory and experiment are in considerable agreement in the medium frequency range $\bar{\nu} > 0.05$. As mentioned, the model is not appropriate for very low frequencies.

As for specular surfaces, again there is no sharp minimum for the precision, so there is a wide range of suitable fringe

frequencies. For high lateral resolution, of course, high frequencies are advantageous.

5. PRECISION, SPEED, AND NUMBER OF DEPTH SCANNING STEPS

From the foregoing we can derive all necessary features of our system. According to customer specifications, we start by selecting the NA of the lens and get the optimal fringe frequency as well as the achievable precision by using the model. To estimate the best possible speed we calculate the minimal number S_z of necessary z-scanning steps. The FWHM width of the contrast function is given by Eqs. ([12](#)) and ([13](#)) for smooth and rough objects, respectively. As noted above, the contrast curve needs a sampling distance $\Delta s \leq \Delta s_{max} = $ FWHM/2. With a given object depth Δz, the minimum number of necessary scanning steps is given by

$$S_z = \frac{2\Delta z}{\text{FWHM}} + 1 + 2. \quad \textbf{(28)}$$

The additive term appears because we need to cover the edge of the object as well. For a rough object with 500 μm depth and an NA = 0.15 lens, we need $S_z = 50$ steps at the optimal fringe frequency $\bar{\nu} = 0.5$. Since we use a 4-phase shift, the number E of exposures is 200. With a given camera frame rate F (and the given precision), the total measuring time cannot be shorter than E/F.

If the required precision is lower, it is appropriate to select a smaller aperture, and exploit a wider contrast curve. It is not appropriate to use a high aperture and a low fringe frequency, because the lateral resolution will be unnecessarily reduced. Of course, the number of scanning steps can be increased over the minimal number S_z, in order to improve the precision via averaging. As already mentioned, this is not efficient for rough objects. For smooth objects it is helpful, because photon noise is uncorrelated in the different exposures.

6. SUMMARY AND CONCLUSION

The paper was motivated by the search for a 3D metrology principle that has the potential for exceptionally high SBSP and for high precision. Moreover, high lateral resolution for both rough and specular surfaces is required. All these features are specifically necessary, e.g., for the inspection of electronic interconnects or for micro-electro-mechanical systems (MEMS). The candidates for smooth and rough surface inspection are CM, WLI, and SIM. The performance of WLI and CM is well known, but SIM is not yet well established for technical applications. This is why a physical model for the signal generation of SIM was developed. The model allows estimation of the fundamental limit of the precision and its dependence on system parameters, such as aperture, fringe frequency, and the number of depth sampling steps. With this knowledge it is possible to optimize SIM for precision or speed.

We summarize the results by one example: with an NA = 0.15 lens, the optimal fringe frequency is in the range of 150 lp/mm. With a diffraction-limited lens, the ultimate limit of the precision for a rough objects is about 2 μm, in agreement with the well-known limit of coherent triangulation. The precision limit on smooth surfaces is in the range of 200 nm, with
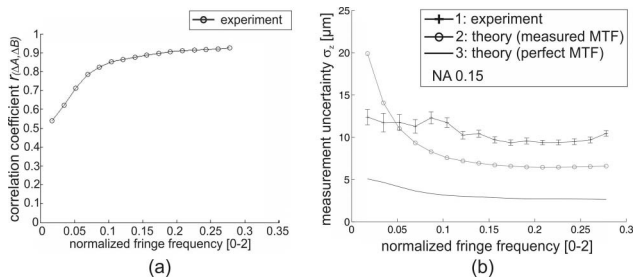

Fig. 5. Precision limit due to speckle noise. (a) Experimentally determined correlation coefficient r with $5 \times /0.15$ micro-objective. (b) Comparison of theoretical and experimental height uncertainty (again with $5 \times /0.15$ micro-objective, $\lambda = 0.5$ μm, and speckle contrast $S_C = 0.6$.

standard video SNR ~ 100. With a very high aperture, NA = 0.85, the precision limit for smooth surfaces is about 5 nm. The comparison of our theoretical model with experiments displays a qualitative agreement, within the limits of experimental imperfections.

With the required precision and the object depth, given by the customer, our model enables the design of the optimal system parameters: the aperture, fringe frequency, number of scanning steps, and maximum possible speed.

In a companion paper, we introduce a novel method, an apparatus, and measurements that exhibit significantly faster scanning and a macroscopic field with no loss of precision and lateral resolution, suggesting that SIM is a candidate measuring method with very high SBSP.

APPENDIX A: DERIVATION OF EQ. ([17])

The intensity of ith measured camera image with M phase steps is

$$I_i = I_m \left\{ 1 + C \cos \left[(i-1)\frac{2\pi}{M} + \varphi_0 \right] \right\}, \tag{A1}$$

where I_m is the mean intensity and φ_0 is the starting phase.

Using Eqs. ([A2]) and ([A3]) we can unambiguously calculate the contrast C and the starting phase φ_0 from phase-shifted camera images $I_1, I_2 \dots I_M$,

$$C(I_1, I_2 \dots I_M)$$

$$= \frac{2\sqrt{\left(\sum_{i=1}^M I_i \cos \frac{(i-1)2\pi}{M} \right)^2 + \left(\sum_{i=1}^M I_i \sin \frac{(i-1)2\pi}{M} \right)^2}}{\sum_{i=1}^M I_i}, \tag{A2}$$

$$\tan \varphi_0 = \frac{\sum_{i=1}^M I_i \sin \frac{(i-1)2\pi}{M}}{\sum_{i=1}^M I_i \cos \frac{(i-1)2\pi}{M}}. \tag{A3}$$

Equation ([A2]) can be written as function of T_1, T_2, T_3, and T_4:

$$C(I_1, I_2 \dots I_M) = \frac{T_4}{T_3} = \frac{2\sqrt{T_1^2 + T_2^2}}{T_3}, \tag{A4}$$

with

$$T_1 = \sum_{i=1}^M I_i \cos \frac{(i-1)2\pi}{M}, \tag{A5}$$

$$T_2 = \sum_{i=1}^M I_i \sin \frac{(i-1)2\pi}{M}, \tag{A6}$$

$$T_3 = \sum_{i=1}^M I_i, \tag{A7}$$

$$T_4 = 2\sqrt{T_1^2 + T_2^2}. \tag{A8}$$

Because the noises in T_3 and T_4 are correlated to each other, the variance of contrast noise σ_C^2 can be formulated by

$$\sigma_C^2 = \left(\frac{\partial C}{\partial T_3} \right)^2 \sigma_{T_3}^2 + \left(\frac{\partial C}{\partial T_4} \right)^2 \sigma_{T_4}^2$$

$$+ 2\frac{\partial C}{\partial T_3}\frac{\partial C}{\partial T_4} \sigma_{T_3}\sigma_{T_4} r(\Delta T_3, \Delta T_4)$$

$$= \left(\frac{T_4}{T_3^2} \right)^2 \sigma_{T_3}^2 + \left(\frac{1}{T_3} \right)^2 \sigma_{T_4}^2 - 2\frac{T_4}{T_3^3} \sigma_{T_3}\sigma_{T_4} r(\Delta T_3, \Delta T_4). \tag{A9}$$

By using the relations $r(\Delta T_3, \Delta T_4) \approx T_4/T_3$ and $\sigma_{T_3}\sigma_{T_4} \approx \sigma_{T_3}^2$, which were found from numeric simulation, Eq. ([A9]) yields Eq. ([A10]):

$$\sigma_C^2 = \left(\frac{1}{T_3} \right)^2 \sigma_{T_4}^2 - \left(\frac{T_4}{T_3^2} \right)^2 \sigma_{T_3}^2. \tag{A10}$$

Following the derivation routine of Eq. ([A9]) and setting $r(\Delta T_1, \Delta T_2) \approx 0$ (known from numeric simulation) we obtain

$$\sigma_{T_4}^2 = \left(\frac{2}{\sqrt{1 + \left(\frac{T_2}{T_1} \right)^2}} \right)^2 \sigma_{T_1}^2 + \left(\frac{2}{\sqrt{1 + \left(\frac{T_1}{T_2} \right)^2}} \right)^2 \sigma_{T_2}^2. \tag{A11}$$

Combing Eqs. ([A11]) and ([A9]) leads to

$$\sigma_C^2 = \left(\frac{2}{T_3\sqrt{1 + \left(\frac{T_2}{T_1} \right)^2}} \right)^2 \sigma_{T_1}^2 + \left(\frac{2}{T_3\sqrt{1 + \left(\frac{T_1}{T_2} \right)^2}} \right)^2 \sigma_{T_2}^2$$

$$- \left(\frac{2\sqrt{T_1^2 + T_2^2}}{T_3^2} \right)^2 \sigma_{T_3}^2. \tag{A12}$$

By taking Eqs. ([A2]), ([A3]) and Eqs. ([A5])–([A8]) into account, we can replace T_1 and T_2 by these terms, which are functions of fringe contrast C and start phase φ_0,

$$\sigma_C^2 = \left(\frac{2}{T_3\sqrt{1 + \tan \varphi_0^2}} \right)^2 \sigma_{T_1}^2$$

$$+ \left(\frac{2}{T_3\sqrt{1 + \tan \left(\frac{\pi}{2} - \varphi_0 \right)^2}} \right)^2 \sigma_{T_2}^2 - \left(\frac{C}{T_3} \right)^2 \sigma_{T_3}^2$$

$$= \left(\frac{2 \cos \varphi_0}{T_3} \right)^2 \sigma_{T_1}^2 + \left(\frac{2 \sin \varphi_0}{T_3} \right)^2 \sigma_{T_2}^2 - \left(\frac{C}{T_3} \right)^2 \sigma_{T_3}^2. \tag{A13}$$

Due to the uncorrelated noise in I_i, $\sigma_{T_3}^2$, $\sigma_{T_3}^2$, and $\sigma_{T_3}^2$ can be derived in Eqs. ([A14])–([A16]) according to the Gaussian law of error propagation,

$$\sigma_{T_1}^2 = \sum_{i=1}^M \cos \left(\frac{(i-1)2\pi}{M} \right)^2 \sigma_{I_i}^2 = \frac{1}{\alpha} I_m \frac{M}{2}, \tag{A14}$$

$$\sigma_{T_2}^2 = \sum_{i=1}^M \sin \left(\frac{(i-1)2\pi}{M} \right)^2 \sigma_{I_i}^2 = \frac{1}{\alpha} I_m \frac{M}{2}, \tag{A15}$$

$$\sigma_{T_3}^2 = \sum_{i=1}^{M} \sigma_{I_i}^2 = \frac{1}{\alpha} I_m M, \tag{A16}$$

where the standard deviation of shot noise is calculated with $\sigma_{I_i} = \sqrt{I_i/\alpha}$ and α is the conversion coefficient between photons and digital signal.

By combining Eqs. (A13)–(A16), we obtain Eq. (A17) describing the relation between variance of contrast noise and the corresponding parameters,

$$
\begin{aligned}
\sigma_C^2 &= \left(\frac{2\cos\varphi_0}{T_3}\right)^2 \frac{1}{\alpha} I_m \frac{M}{2} + \left(\frac{2\sin\varphi_0}{T_3}\right)^2 \frac{1}{\alpha} I_m \frac{M}{2} \\
&\quad - \left(\frac{C}{T_3}\right)^2 \frac{1}{\alpha} I_m M \\
&= \frac{2 I_m M}{\alpha (I_m M)^2} \left(\cos\varphi_0^2 + \sin\varphi_0^2 - \frac{C^2}{2}\right) \\
&= \frac{2}{\alpha I_m M} \left(1 - \frac{C^2}{2}\right).
\end{aligned}
\tag{A17}
$$

If we define that the maximal SNR_{shot} is given by twice the mean fringe intensity,

$$SNR_{shot} = \sqrt{2\alpha I_m}, \tag{A18}$$

the standard deviation σ_C of shot-noise-limited contrast noise on specular surfaces can be written as in Eq. (A19),

$$\sigma_C = \frac{2}{SNR_{shot}\sqrt{M}} \sqrt{1 - \frac{C^2}{2}}. \tag{A19}$$

APPENDIX B: DERIVATION OF EQ. (25)

By deriving Eq. (19), the variance of contrast noise on rough surfaces can be formulated in Eq. (B1),

$$
\begin{aligned}
\sigma_C^2 &= \left(\frac{\partial C}{\partial A}\right)^2 \sigma_A^2 + \left(\frac{\partial C}{\partial B}\right)^2 \sigma_B^2 + 2\frac{\partial C}{\partial A}\frac{\partial C}{\partial B}\sigma_A\sigma_B r(\Delta A, \Delta B) \\
&= \left(\frac{2B}{(A+B)^2}\right)^2 \sigma_A^2 + \left(\frac{-2A}{(A+B)^2}\right)^2 \sigma_B^2 \\
&\quad - 8\frac{AB}{(A+B)^4}\sigma_A\sigma_B r(\Delta A, \Delta B).
\end{aligned}
\tag{B1}
$$

Combining Eq. (B1) with Eqs. (20), (21), and (24) leads to Eq. (B2),

$$
\begin{aligned}
\sigma_C^2 &= \left(\frac{2B}{(A+B)^2}\right)^2 \sigma_A^2 + \left(\frac{-2A}{(A+B)^2}\right)^2 \sigma_B^2 \\
&\quad - 8\frac{AB}{(A+B)^4}\sigma_A\sigma_B r(\Delta A, \Delta B) \\
&= \left(\frac{1-C}{2I_m}\right)^2 S_C^2 (I_m + C I_m)^2 + \left(\frac{1+C}{2I_m}\right)^2 S_C^2 (I_m - C I_m)^2 \\
&\quad - 2\left(\frac{1-C^2}{4I_m^2}\right) S_C^2 I_m^2 (1-C^2) r(\Delta A, \Delta B) \\
&= \frac{(1-C^2)^2 S_C^2}{4} + \frac{(1-C^2)^2 S_C^2}{4} - \frac{(1-C^2)^2 S_C^2}{2} r(\Delta A, \Delta B) \\
&= \frac{(1-C^2)^2 S_C^2}{2} (1 - r(\Delta A, \Delta B)).
\end{aligned}
\tag{B2}
$$

The standard deviation σ_C of speckle-noise-limited contrast noise at scattering surfaces is given by Eq. (B3), as stated above, in Eq. (25):

$$\sigma_C = \frac{1}{\sqrt{2}} S_C (1 - C^2) \sqrt{1 - r(\Delta A, \Delta B)}. \tag{B3}$$

REFERENCES

1. C. Wagner and G. Häusler, "Information theoretical optimization for optical range sensors," Appl. Opt. **42**, 5418–5426 (2003).
2. O. Hýbl, A. Berger, and G. Häusler, "Information efficient white-light interferometry," in *Proc. DGaO* **A29** (2008).
3. O. Hýbl and G. Häusler, "Information efficient white-light interferometry," in *Proc. ASPE 2010 Summer Topical Meeting*, Asheville, North Carolina, June, 2010.
4. V. Srinivasan, H. C. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3-D diffuse objects," Appl. Opt. **23**, 3105–3108 (1984).
5. G. S. Kino and S. S. C. Chim, "Mirau correlation microscope," Appl. Opt. **29**, 3775–3783 (1990).
6. B. S. Lee and T. C. Strand, "Profilometry with a coherence scanning microscope," Appl. Opt. **29**, 3784–3788 (1990).
7. T. Dresel, G. Häusler, and H. Venzke, "3D-sensing of rough surfaces by 'coherence radar,'" Appl. Opt. **31**, 919–925 (1992).
8. X. C. de Lega, J. Biegen, P. de Groot, G. Häusler, and P. Andretzky, "Large field-of-view scanning white-light interferometers," in *Proceedings of ASPE 18th Annual Meeting*, Portland, Oregon, October, 2003.
9. R. Groß, O. Hýbl, C. Richter, B. Wiesner, and G. Häusler, "Reduced measurement uncertainty of white-light interferometry on rough surfaces," Proc. SPIE **6341**, 63410V (2006).
10. P. de Groot and L. Deck, "Three-dimensional imaging by sub-Nyquist sampling of white-light interferograms," Opt. Lett. **18**, 1462–1464 (1993).
11. O. Hýbl, "Information-efficient white-light interferometry," Ph.D. thesis (Friedrich-Alexander-University, 2013).
12. T. Wilson, *Confocal Microscopy* (Academic, 1990).
13. H. Tiziani and H. Uhde, "Three-dimensional image sensing by chromatic confocal microscopy," Appl. Opt. **33**, 1833–1843 (1994).
14. K. Carlsson and R. Hummel, "Confocal imaging for 3-D digital microscopy," Appl. Opt. **26**, 3232–3238 (1987).
15. G. Q. Xiao and G. S. Kino, "A real-time confocal scanning optical microscope," Proc. SPIE **0809**, 107–113 (1987).
16. F. Bitte, G. Dussler, and T. Pfeifer, "3D micro-inspection goes DMD," Opt. Lasers Eng. **36**, 155–167 (2001).
17. M. A. A. Neil, R. Juskaitis, and T. Wilson, "Method of obtaining optical sectioning by using structured light in a conventional microscope," Opt. Lett. **22**, 1905–1907 (1997).
18. T. Wilson, "Microscopy imaging apparatus and method," U.S. patent 6,376,818 (23 April 2002).
19. K. Engelhardt and G. Häusler, "Acquisition of 3D-data by focus sensing," Appl. Opt. **27**, 4684–4689 (1988).
20. M. G. L. Gustafsson, L. Shao, P. M. Carlton, C. J. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," Biophys. J. **94**, 4957–4970 (2008).
21. R. Artigas, A. Pinto, and F. Laguarta, "Three-dimensional micromeasurements on smooth and rough surfaces with a new confocal optical profiler," Proc. SPIE **3824**, 93 (1999).

22. M. Schwertner, "Verfahren und Anordnung zur optischen Abbildung mit Tiefendiskriminierung," German patent DE 102007018048 (16 October 2008).
23. E. Kristensson, E. Berrocal, and M. Alden, "Two-pulse structured illumination imaging," Opt. Lett. **39**, 2584–2587 (2014).
24. J. J. Xu, "3-D optical microscope," U.S. patent 7,729,049 (1 June 2010).
25. Y. Kim, K. Hibino, N. Sugita, and M. Mitsuishi, "Design of phase shifting algorithms: fringe contrast maximum," Appl. Opt. **22**, 18203–18213 (2014).
26. Y. Surrel, "Design of algorithms for phase measurements by the use of phase stepping," Appl. Opt. **35**, 51–60 (1996).
27. J. Goodman, *Introduction to Fourier Optics* (Roberts & Company, 2005).
28. H. H. Hopkins, "The frequency response of a defocused optical system," Proc. R. Soc. London Ser. A **231**, 91–103 (1955).
29. P. A. Stokseth, "Properties of a defocused optical system," J. Opt. Soc. Am. **59**, 1314–1321 (1969).
30. X. Laboureux and G. Häusler, "Localization and registration of three-dimensional objects in space—where are the limits?" Appl. Opt. **40**, 5206–5216 (2001).
31. Standard ISO TC 69/SC 6 N, "Accuracy (trueness and precision) of measurement methods and result—Part 1: introduction and basic principles," 2011.
32. G. Häusler, "Speckle and coherence," in *Encyclopedia of Modern Optics*, B. Guenther, ed. (Elsevier, 2004), pp. 114–122.
33. JCGM 100:2008, "GUM evaluation of measurement data—Guide to the expression of uncertainty in measurement," 2008.
34. R. G. Dorsch, G. Häusler, and J. M. Herrmann, "Laser triangulation: fundamental uncertainty in distance measurement," Appl. Opt. **33**, 1306–1314 (1994).