applied optics

Sub-millimeter depth-resolved digital holography

Joseph van Rooij* and Jeroen Kalkman

Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
*Corresponding author: j.vanrooij@tudelft.nl

We present sub-millimeter full-field depth from focus digital holography of surface topography of rough objects. For each pixel, the depth of the object is calculated from the variance of the intensity image over a set of reconstruction distances. First, we theoretically describe the axial resolution of this method and show that sub-millimeter resolution is feasible. Second, using a digital holography setup without magnifying optics or lateral scanning we experimentally demonstrate 100 µm axial resolution depth ranging and surface topography imaging. This is significantly better than what has previously been reported using digital holography and could make this technique useful for rapid large-area characterization of surface topography of objects.   © 2017 Optical Society of America

1. INTRODUCTION

Surface metrology and absolute distance measurement are essential in many applications; for example, in the field of geosciences, remote sensing aims to reconstruct the surface topology and track changes of the Earth surface over time. On a much smaller scale, optical measurement of surface topography has become vital in many process and quality control methods [1]. A number of established optical surface imaging techniques are available, such as phase shifting interferometry (PSI), white light interferometry (WLI), optical coherence tomography (OCT), digital holographic microscopy (DHM), and focus variation microscopy (FVM). In PSI a controlled phase shift is applied to the reference wave while acquiring the full-field interference pattern. From a series of images, acquired at different phase shifts, the phase can be calculated with 2π ambiguity. Since phase unwrapping is necessary, PSI is problematic with phase discontinuities greater than 2π [2]. WLI is based on the interference signal of a low-coherence (white light) source. The interference between the reference field and object wave field imaged with a lens onto the camera changes as a function of the distance of the scanning reference mirror. The technique can acquire surface topography in full-field at sub-nanometer axial resolution. Furthermore, in contrast to PSI [3], it can be applied to surfaces that are complex in terms of roughness and discontinuities, but has a long acquisition time due to the requirement of axial scanning. Time-domain OCT is similar to WLI but instead of only measuring surface topography, it is used to image the inside of turbid media, such as retinal tissue or skin. Time-domain OCT is based on lateral scanning and the use of a focusing lens to suppress scattered light while

measuring deep in tissue. In Fourier-domain OCT an axial depth scan is calculated with a Fourier transform of an acquired interference spectrum, and as such does not require axial scanning, but is not full-field and therefore lateral scanning is necessary. Meter-scale OCT depth ranging with 15 µm axial resolution has recently been achieved [4]. DHM is a technique that does not use scanning because it reconstructs the complete wave field. The acquisition time is therefore short compared to other techniques while having the possibility to attain sub-nanometer axial resolution. When the reconstructed wave field is in focus, the phase of the wave field is linear to the height of the sample. The reconstructed phase is defined modulo 2π and therefore the height that can be reconstructed in reflection mode without ambiguity is half a wavelength [5]. This can be overcome using phase unwrapping; however, this does not work if abrupt steps greater than half a wavelength are present. Furthermore, for rough objects the wavefront becomes too disturbed to calculate the topography. Localization of rough objects with digital holography has also been demonstrated using a statistical fringe processing technique [6]. In this technique, the object is illuminated from two angles, a digital hologram is captured for each angle, and the two wavefronts are reconstructed at a number of reconstruction distances near the object plane. At each reconstruction plane, a statistical algorithm is applied to the phase difference map for object localization up to a precision of 250 µm. FVM on the other hand is not an interferometric technique. It uses axial focus scanning and exploits the limited depth of focus of the objective lens to extract topology information from focus variation quantified with a focus metric, provided the surface is optically rough [7].

The axial precision of this technique depends on the magnification used; 2.5 × and 100 × magnifications, for example, lead to micrometer and nanometer range precision, respectively [8]. Just like in WLI, the acquisition time is relatively long due to the requirement of axial scanning. For any of these techniques there exists trade-offs between acquisition time, sample properties (e.g., roughness and discontinuities), the presence of scanning, depth range, and axial resolution. Depth from focus digital holography (DFF-DH) attempts to combine the short acquisition time of digital holography (scanless imaging) with the ability to reconstruct topographies with large discontinuities or rough surfaces. In contrast to ordinary imaging where the focus is varied by changing the position of the lens, in DH the in-focus image can be calculated at any depth plane. The DFF-DH method is an image processing approach that estimates the surface location from the optimum of a focus metric calculated from the digitally reconstructed image. For 3D objects the image plane depends on the distance of every part of the object to the camera. By reconstructing the image of the object at different depths, the degree of focus of a particular region in the image reconstruction (calculated with a focus metric) encodes the depth of the object [9]. Because one can reconstruct the complete wave field at any depth from a single digitally captured hologram, this method does not need lateral or axial scanning and has no fundamental limit regarding the depth range that can be measured other than the coherence length of the light source. DFF was first used within the context of digital holography by Ma *et al.* (2004) [10], who recovered object depth for every part of the object in this way from a digital hologram. A similar approach was also used for extended focus imaging by McElhinney *et al.* (2008) [11]. While Tachiki *et al.* (2008) measured an axial accuracy of a centimeter [9], the theoretical understanding of the axial resolution and its fundamental limits remains largely unknown. In this paper we show how the depth resolution of DFF-DH depends on sample properties and on experimental setup parameters. Furthermore, we show experimentally that the precision in the axial direction can be brought down to the 100 µm range in a basic DH setup without any magnification. In the next section, we first give an overview of the basic principles of DFF-DH and a theoretical framework to analyze the precision. Subsequently, we compare our framework with simulations where we show that submillimeter resolution is possible. In the results section we experimentally demonstrate sub-millimeter axial in DFF-DH.

2. THEORY

A. Depth from Focus Digital Holography
In digital holography the image is numerically calculated from an interferogram, instead of it being formed optically with a lens. For an explanation of the basic principles of digital holography, we refer the reader to Schnars and Jueptner (2005) [12]. Digital holography suffers from a limited depth of field just like most other imaging systems. If the reference wave $R(x, y)$ is a plane wave, as we will assume throughout this paper, then the reconstruction distance z_r equals the recording distance z_0 and the reconstructed image appears in focus [13]. Reconstructing the image at other distances than z_0 leads to defocus blurring of the image due to the limited depth of field. The degree of

focus in the image depends on the reconstruction distance and therefore encodes the distance of the object to the sensor. Since the degree of focus is at its maximum when $z_r = z_0$, one needs to quantify the degree of focus with a focus metric and find the optimum value as a function of the reconstruction distance. For each pixel in the object image we can repeat this process and calculate its distance to the sensor plane z_0 to obtain a height map of the object. The degree of focus in the image as a function of reconstruction distance z_r depends on the numerical aperture (NA) of the imaging system. In DH, the NA is inversely proportional to the recording distance and proportional to the dimensions of the sensor.

B. Focus Metric Definition
The degree of focus is quantified using image-based metrics. These metrics are calculated from the image and have their maximum when the image is in focus, and decrease rapidly when the image is out of focus. Different focus metrics exist, see for example Tian *et al.* (2007) for an overview [14]. Image variance is a focus metric that is simple to calculate, and has been proven to be a good depth measure [15]. The variance of a digital image I of $n \times m$ pixels is given by

$$\text{var}(I) = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} (I(i, j) - \bar{I})^2, \qquad (1)$$

where $\bar{I}$ is the mean intensity of the image. In order to derive an analytic model for the variance as a function of reconstruction distance, we will use a continuous form, which is given by

$$\text{var}(I) = \iint (I(x, y) - \mu)^2 \, dx \, dy, \qquad (2)$$

where μ is the mean of the image and is given by

$$\mu = \iint I(x, y) \, dx \, dy. \qquad (3)$$

C. Depth from Focus Precision Analysis
To obtain a theoretical description, we neglect the details of surface scattering processes, following the assumption of Nikolaev *et al.* (2016) in their linear theory approach to FVM [7], and consider the object surface to be a perfect plane with a reflection coefficient that has a random distribution. The object will be considered to have a reflection coefficient that has a "white" frequency distribution, i.e., all spatial frequencies have equal power. Due to the linearity of the digital holographic imaging process with respect to the object field, the reconstructed object field is described by propagation of every spatial frequency of the object into the image space. For each spatial frequency in the object, the focus curve is periodic due to the Talbot effect, and we will term such a focus curve a Talbot curve. In our specific case, it turns out that the focus curve for the object is a sum of the individual Talbot curves for all spatial frequencies. In order to keep the expressions concise we consider a one-dimensional input, although a generalization to two dimensions is straightforward. We first derive an expression for the image of a reflecting planar object with a single spatial frequency and random phase at recording distance z_0 as input to the holographic imaging system, which we reconstruct at reconstruction distance z_r. Neglecting the finite extent

of the aperture, we consider an object described by the field reflection:

$$t_n = \frac{1}{2}[1 + m\cos(2\pi n\xi/L) + \phi_n], \qquad (4)$$

where ϕ_n is a random phase term, n/L is the spatial frequency, m is an amplitude factor, and ξ is the lateral spatial coordinate in the input plane. In principle n can be any number, although in practice it is an integer value due to discrete sampling of the detector. The reconstructed holographic wave field is calculated by propagating the input wave field to the hologram plane, and then to the reconstruction plane, using the transfer function Fresnel diffraction method approach according to the treatment of Goodman (1996) [16]. The resulting reconstructed wave field is

$$U_n(\xi, z_r) = \frac{2 + 2me^{-\frac{i\pi\lambda n^2(z_0+z_r)}{L^2}}\cos)}{4N}. \qquad (5)$$

The intensity of the reconstructed wavefront is then

$$I_n(\xi, z_r) = \frac{1}{4N^2}\left(m\cos\left(\frac{2\pi n\xi}{L} + \phi_n\right)\right.$$
$$\left. + 2\cos\left(\frac{\pi\lambda n^2(z_0+z_r)}{L^2}\right)\right)$$
$$\times m\cos\left(\frac{2\pi n\xi}{L} + \phi_n\right) + \frac{1}{4N^2}. \qquad (6)$$

This reduces to $|t_n|^2$ for reconstruction distances

$$z_r = -z_0 + \frac{2L^2 k}{\lambda}, \qquad (7)$$

which means that the input grating is replicated at fixed distances for an integer value of k, also called "self-imaging." This occurs for periodic inputs in general and is a manifestation of the Talbot effect [16] appearing in holographic reconstruction. The variance for one spatial frequency as a function of reconstruction distance is subsequently found by calculating the integral in one dimension [Eq. (2)] over the integration range 0 to L, which yields

$$\text{var}_n(z_r) = a + b\cos\left(\frac{2\pi\lambda n^2(z_0+z_r)}{L^2}\right), \qquad (8)$$

where m is assumed to be equal to one, and

$$a = \frac{9L(2L^2 - 4L + 3)}{128N^4}, \qquad (9)$$

and

$$b = \frac{L}{16N^4}. \qquad (10)$$

For every spatial frequency n/L the variance is thus periodic as a function of reconstruction distance z_r. For this reason we will refer to such a curve as a Talbot curve. We now assume that the object is composed of many spatial frequencies. Since in an experimental setting the spatial frequencies are discretely sampled, we describe the reflection of the object as

$$t_N = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{2}[1 + m\cos(2\pi n\xi/L) + \phi_n]. \qquad (11)$$

According to the superposition principle, the reconstructed field intensity is

$$I_N(\xi, z_r) = \left|\sum_{n=1}^{N} U_n(\xi, z_r)\right|^2, \qquad (12)$$

$$= \sum_{n=1}^{N}|U_n|^2 + \sum_{n=1}^{N}\sum_{m=1}^{N} U_n U_m^*(1 - \delta_{n,m}). \qquad (13)$$

Due to the random phase term ϕ_n, the first term in Eq. (13) and the last term are independent random variables. Using the property that $\text{var}(a + b) = \text{var}(a) + \text{var}(b)$ if a and b are independent random variables, we can write the variance of the reconstructed field intensity as

$$\text{var}(I_N) = \sum_{n=1}^{N}\text{var}(|U_n|^2) + \text{var}\left(\sum_{n=1}^{N}\sum_{m=1}^{N} U_n U_m^*(1 - \delta_{n,m})\right). \qquad (14)$$

In the appendix we show that this expression reduces to a linear sum of Talbot curves [see Eq. (8)]:

$$\text{var}_N(z_r) = A + B\sum_{n=1}^{N}\cos\left(\frac{2\pi\lambda n^2(z_0+z_r)}{L^2}\right), \qquad (15)$$

with A and B constants. Equation (15) provides a physical understanding of the width of the focus curve. The sum of cosines gives rise to a peak at $z_r = -z_0$, the virtual image plane. For larger N, either because of a larger numerical aperture or a larger number of contributing spatial frequencies in the input, the peak width becomes smaller. The result in Eq. (15) gives a limit for the precision that can be achieved by depth from focus digital holography in terms of the peak width. In Fig. 1, the focus curve of Eq. (15), for $z_0 = 70.9$ mm (real image), $\lambda = 633$ nm, $L = 6.7$ mm, and $N = 1024$ is plotted as a function of reconstruction distance z_r. In the preceding analysis it has been assumed that the object has an ideal flat power
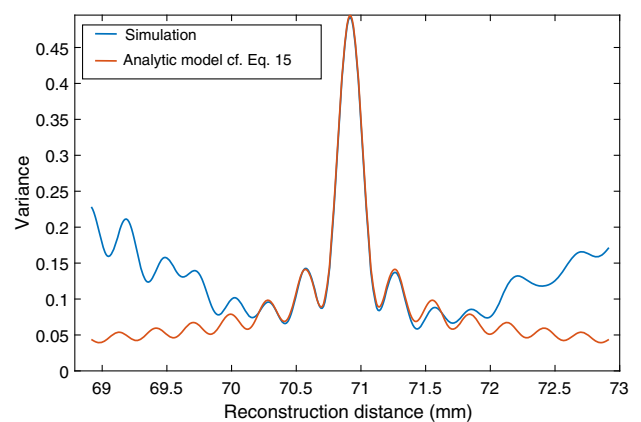
Fig. 1. Comparison between the focus (variance) curves resulting from the Fresnel diffraction simulation (blue) and the analytic model Eq. (15) (red). The difference between the two calculations is due to the finite aperture used in the simulation.

spectrum, and the phases ϕ_n for spatial frequencies n/L were assumed to be statistically independent (ideal white noise input). For a real object the spatial frequency distribution of the object may be less ideal leading to fewer terms contributing to the summation in Eq. (15) and therefore resulting in a wider peak. Finally, the summation is limited by the numerical aperture and sampling of the imaging system. The numerical aperture of the system limits the highest spatial frequency to reach the sensor, while the pixel pitch of the sensor limits the highest spatial frequency that can be sampled properly. At the critical recording distance, the maximum amount of spatial frequencies is imaged without aliasing. This distance is given by $z_{\mathrm{crit}} = N_{\mathrm{pix}} \Delta \xi^2 / \lambda$, where N_{pix} is the number of sensor pixels, $\Delta \xi$ is the pixel pitch, and λ is the recording wavelength. For an amplitude grating with spatial frequency n/L, the angle under which the first diffraction order propagates from the normal is

$$\sin(\theta) = \frac{\lambda n}{L}. \tag{16}$$

The numerical aperture is determined by the sensor dimension $N_{\mathrm{pix}} \Delta \xi$ and recording distance z_0 and limits the maximum diffraction angle θ_{max} that can be imaged by the system, according to

$$\theta_{\mathrm{max}} = \tan^{-1}\left(\frac{N_{\mathrm{pix}} \Delta \xi}{2 z_0}\right), \tag{17}$$

which follows from the imaging geometry.

Combining the last two equations and solving for the maximum spatial frequency $n/L = N_{\mathrm{max}}/L$, we find that

$$\frac{N_{\mathrm{max}}}{L} = \sin\left(\tan^{-1}\left(\frac{N_{\mathrm{pix}} \Delta \xi}{2 z_0}\right)\right) \frac{1}{\lambda}. \tag{18}$$

Using the critical sampling distance $z_0 = z_{\mathrm{crit}}$ this becomes

$$\frac{N_{\mathrm{max}}}{L} = \frac{1}{2 \Delta \xi \sqrt{\frac{\lambda^2}{4 \Delta \xi^2} + 1}} \approx \frac{1}{2 \Delta \xi}. \tag{19}$$

Hence, the summation in Eq. (15) is limited to the term with the maximum spatial frequency N_{max}/L that is imaged with the system. At the critical sampling distance, this is equal to the Nyquist frequency and in the summation of Eq. (15) limits the DFF-DH minimum peak width.

3. DIGITAL HOLOGRAPHY SIMULATIONS

We validate the analytic focus curve model of Eq. (15) as a sum of Talbot curves by comparing it to the well-known method of Fresnel diffraction calculations, namely the transfer function method. We implemented the transfer function methodcode in MATLAB 2016; we refer the reader to the work of Voelz (2011) [17] for more information. The input image in the simulation is given by Eq. (11), where the input object support L is 200 pixels, the number of pixels in the recording plane is $N = 1000$, the number of unique spatial frequencies in the input image is $L/2 - 1$, the wavelength is $\lambda = 633$ nm, the pixel size is $\Delta \xi = 6.45$ µm, and the recording distance z_0 at $N \Delta \xi^2 / \lambda$ (critical recording distance). These are parameters used in a typical experimental DH realization.

The digital hologram is calculated and the real image is reconstructed in 400 steps over the distance $z_0 - 2$ mm to $z_0 + 2$ mm. The variance of the reconstructed image is calculated at each step. The variance as a function of reconstruction distance is shown in Fig. 1.

Since there is amplitude scaling in the diffraction calculation that is not present in the analytic model, the simulated focus curve is scaled along the vertical axis to match the peak variance of the analytic model. The simulated focus curve describes the analytic model of Eq. (15) well around the center $z_r = z_0$, and deviates toward the edges due to finite aperture effects that occur in the simulation, but are not accounted for in the analytic model. Both approaches lead to the same focus curve peak width of approximately 450 µm, indicating the possibility of sub-millimeter axial resolution with DFF-DH without magnification.

4. MATERIALS AND METHODS

A. Digital Holography Setup

The digital holography setup is shown in Fig. 2 and consists of a Michelson interferometer with the light illuminating the object normal to the surface. The light source is a HeNe laser with a wavelength of 633 nm and an output power of the order of 3 mW. Two lenses (Thorlabs, LD2568 and LA1979) are used to expand and collimate the illuminating laser beam respectively to a width (FWHM) of approximately 15 mm; besides that no lenses or objectives are used in the imaging process. A shearing interferometer (Thorlabs, SI254) is used to assure proper collimation of the reference beam, so that the distance in reconstruction (virtual) space equals the physical recording distance. The object is placed approximately 70 mm away from the sensor plane. The mirror in the reference arm is mounted onto a piezoelectric transducer controlled by a computer for phase-shifting the digital hologram, where we capture four holograms with a phase shift of $\pi/2$ of the reference beam between each subsequent hologram. From a linear combination of these holograms a complex hologram is formed where the zeroth and out-of-focus conjugate order are removed [18]. In this way we can use the full image plane and maximize
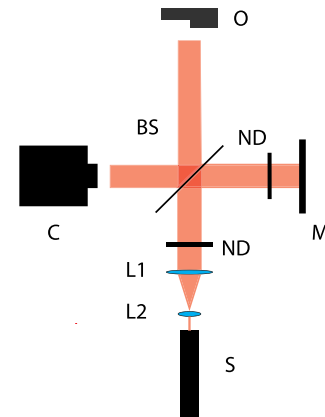
Fig. 2. Michelson interferometer setup for acquiring the digital holograms. S, HeNe laser (633 nm); BS, pellicle beam-splitter; C, CCD camera (1344 × 1024 pixels); M, piezo mounted mirror; O, object; L, lens; ND, variable neutral density filter.

the lateral resolution in the reconstructed image. We use a pellicle beam splitter (Thorlabs, BP233) in order to avoid ghost images in the reconstructions (that resulted in degraded reconstructions in earlier research on DFF-DH [9]). Due to the sensitivity of the pellicle beam splitter membrane to vibrations from sound or airflow, the setup is enclosed in a box in order to minimize movement of the membrane. A variable neutral density filter is placed between the beam splitter and the piezo mounted mirror and right after the laser aperture in order to control saturation of the camera and the ratio of the intensity between the reference and the object beam, which is close to 1. The digital holograms are captured by a CCD camera (ORCA ER, Hamamatsu) with 1344×1024 square pixels with a pixel size of 6.45 µm and acquisition speed of approximately 10 frames a second.

B. Sample Preparation and Characterization

We use a brass reflector with four different step heights (squares) separated at heights approximately 200 µm apart as a test target to demonstrate sub-millimeter axial resolution. The sample surface is made a rough surface by briefly sandblasting the object. The total area of the square sample is 25 mm^2. For a reference measurement of the surface height topology we use a white-light interferometer (WLI, Bruker ContourGT-K). We used the smallest magnification available on the WLI, which was $2.5 \times$. Due to the field of view (2 mm) being significantly smaller than the lateral sample dimensions (5 mm), stitching is necessary which is done automatically with the accompanied *Vision* software. Tilt is corrected for by fitting a plane to one of the four surfaces and correcting the entire profile for this tilt. The acquisition time for the full surface at the lowest magnification is of the order of 30 min. The roughness parameters S_a (average roughness) and S_q (root mean square roughness) are measured with the WLI for the sample and are given by 3.6 and 11.6 µm, respectively.

C. Digital Holographic Reconstruction

Even though the lateral sample dimensions are smaller than the sensor and thus the transfer function approach can be used, there is the possibility of the sample holder, which extends to beyond the sensor plane, appearing in the reconstructed image. For this reason reconstruction of the digital holograms was performed using the single fast Fourier transform (S-FFT) method [19] since this does not restrict the reconstruction plane to the dimensions of the sensor. Since noise appears as high-frequency information in the reconstructed images, we apply a spatial Gaussian filter to each reconstructed image. We found that a standard deviation σ of approximately 0.7 pixels gave the most accurate reconstructions. We calculate the focus metric per pixel for each reconstruction distance by taking a window of N by N pixels around the center pixel. A trade-off exists in DFF-DH between the axial resolution and the window size and thus the lateral resolution [9]. We used a window size of 61 by 61 pixels (0.2 mm^2) and calculated the variance of the windowed image over reconstruction distances $z_r = z_0 - 2$ mm to $z_0 + 2$ mm. The reconstruction distance where the variance peaks for a particular pixel is our estimate for the distance from the recording plane to the object pixel. Finally, we corrected for surface tilt of the entire object with respect to the recording

plane by fitting a plane to two diagonally opposite surfaces and calculating the average tilt. This tilt is subtracted from the actual distances. The time for the algorithm to calculate a complete depth image on a quad core computer is around half a minute.
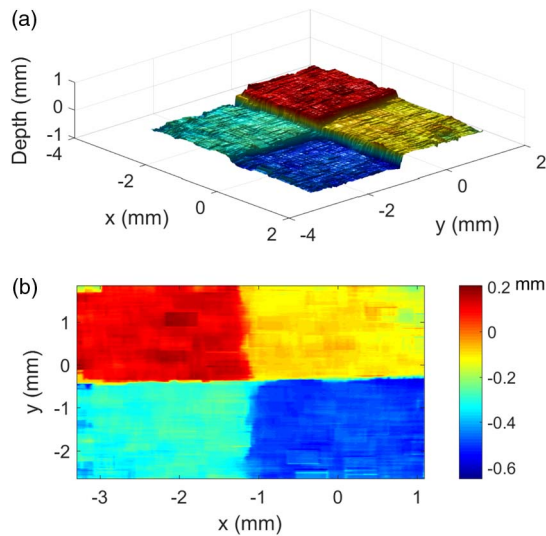
5. RESULTS

Figures 3(a) and 3(b) show the reconstructed DH intensity images at reconstruction distances $z_r = 69.8$ and $z_r = 70.4$ mm, respectively. At $z_r = 69.8$ mm, the lower right quadrant is in focus and shows the fine details of the surface, whereas the upper left quadrant has a blurred appearance because it is out of focus. At $z_r = 70.4$ mm the opposite occurs. Figure 4 shows focus metric curves obtained from the experimental data at the locations indicated in Fig. 3. Based on the optimum of
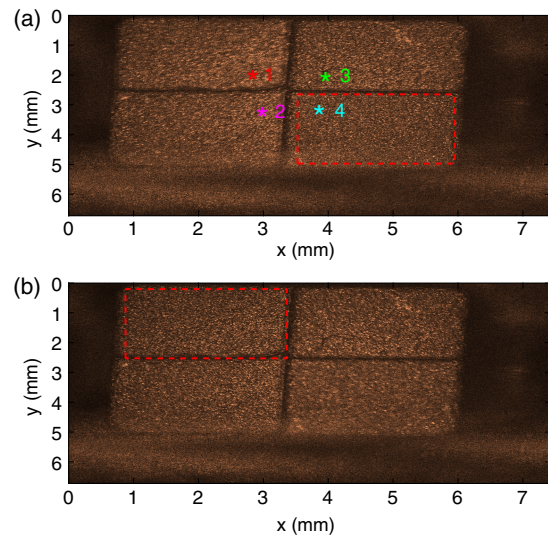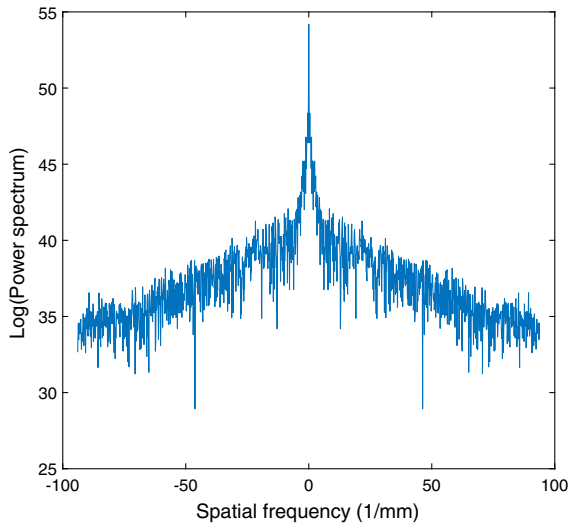

Fig. 3. Reconstructed intensity image of the brass reflector object at (a) $z_r = 69.8$ and (b) $z_{1r} = 70.4$ mm. For (a) the lower right square is in focus, for (b) the upper left square is in focus.
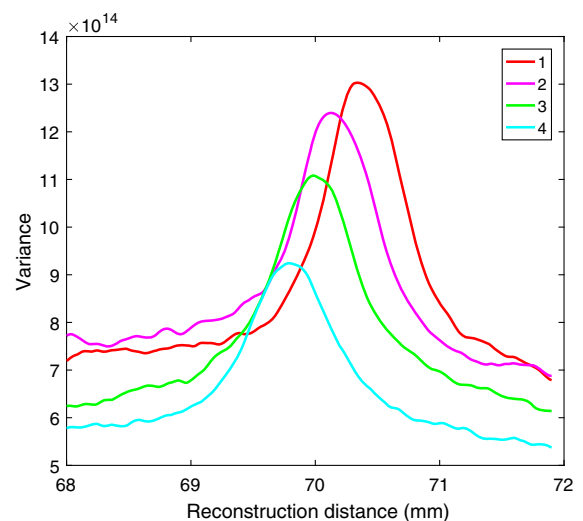
Fig. 4. DFF curves obtained from the experimental data at the points indicated in Fig. 3 at different planes.
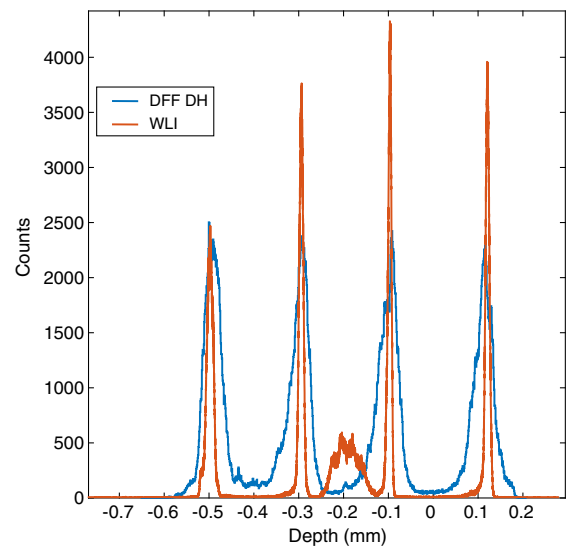
(a)

(b)

Fig. 5. (a) Reconstructed DFF-DH depth map of the brass reflector object and (b) top view.

Fig. 6. Cross section along the vertical axis of the logarithm of the object image power spectrum when reconstructed in focus along $x = 4$ mm in Fig. 3.

Fig. 7. Histogram of the distribution of depth values from the DFF-DH reconstruction of Fig. 5 (blue), together with the depth distribution of the reference WLI measurement (red).

To obtain an estimate of the accuracy of our approach we determined the distribution of reconstructed depth values for all pixels in the image. The result is shown in Fig. 7 as a histogram together with a comparison of the distribution obtained with the white-light interferometer. Good agreement is obtained with the DFF-DH method shown here. Assuming that the WLI data represents the actual height distribution of the object we estimate that from the width of the DFF-DH histogram our method has a depth resolution of approximately 100 µm.

6. DISCUSSION

We presented sub-millimeter DFF-DH imaging of surface topography without the use of magnifying or lateral scanning optics. In the theoretical analysis presented in this paper it is assumed that all spatial frequencies in the object surface are equally present with randomly distributed phases. This results in a lower limit of the depth resolution of DFF-DH systems in terms of the peak width of the focus curve.

The focus curves obtained from the experimental data do not show the oscillating behavior in the sidelobes of the curve. Moreover, the experimental focus curve was significantly broader than the theoretical limit. Both effects are likely caused by the fact that the object plane was not perfectly parallel to the camera, but instead is slightly tilted, and the spatial frequency spectrum of the object does not correspond to a white noise signal.

On the other hand, the optimal depth is determined from the maximum value of the peak, which is much better defined than the peak width. Since the DFF-DH method in this paper relies only on finding the maximum, this partially compensates for the broadening of the focus curve in the experiment; the measured width is approximately 750 µm, while the DFF depth resolution is around 100 µm. As a result the surface

the focus curves, the surface is reconstructed as shown in Fig. 5. The experimental focus curve has a full width at half-maximum of approximately 750 µm, which is significantly broader than the width of the simulated curve in Fig. 1. We attribute this to the fact that the frequency content of the object is limited and the object plane was slightly tilted with respect to the sensor plane. In that sense, the focus curve in Fig. 1 represents a limiting case with perfect parallel orientation and uniform spatial frequency power spectrum in the input image. The spatial frequency content of the experimental input is estimated by taking the Fourier transform of the optimal in focus intensity image. The logarithm of the power spectrum in Fig. 6 shows that indeed the power spectrum is not flat, but instead shows a significant drop off in power for high spatial frequencies.

height estimation precision is significantly below the width of the focus curves. Although using the maximum value of the focus metric to estimate the surface position has a physical basis, it does not utilize other information of the focus curve, such as the peak width and shape, to improve the depth estimate. Fitting a polynomial or point spread function curve, as applied in FVM [8], could further improve the axial resolution.

While this paper mainly was dedicated to analyzing the axial resolution of DFF-DH, lateral resolution also plays an important role in applications. DFF-DH is based on an area of pixels to calculate the focus metric at the center pixel, thus smoothing the depth map in the transversal direction. In principle there is a relation between the number of pixels needed for obtaining a depth estimate and the axial precision, thereby resulting in an intrinsic trade-off between axial and lateral resolution. We observed that a window size of 61 pixels gave good results. Using more sophisticated methods, mentioned earlier to estimate the depth, will likely enable the use of smaller window sizes.

In our configuration, four digital holograms were acquired for phase shifting digital holography. This was to make full use of the image plane; however, this is not a strict requirement. A setup with a larger sensor can produce equivalent results from a single hologram capture if the reconstructed diffraction orders are properly separated spatially in the reconstruction plane.

The results in this work were obtained with a basic form of DFF-DH, without using additional magnifying optics. Magnifying optics can be incorporated into the DFF-DH system in order to improve lateral and axial resolution, which goes at the expense of the field of view. Since in digital holography lateral resolution depends also on the aperture of the sensor and indirectly on the pixel size, resolution in DFF-DH also can be improved without magnifying optics by using a larger sensor while retaining the field of view. We envision this technique can be useful for characterizing objects in art, process control and computer vision, or in general when absolute distance measurements are needed in a very short acquisition time.

7. CONCLUSIONS

We demonstrated that the DFF curve as a function of reconstruction depth can be approximated by the sum of Talbot curves for every spatial frequency. Analytic theory and numerical simulation indicated that the axial precision in terms of the peak width can be brought down to well below the millimeter range using a basic digital holography setup without scanning, imaging optics, or magnification. Submillimeter depth resolution was experimentally demonstrated in a DFF-DH setup using a rough reflecting step-profile where our method was able to discriminate between layers that were approximately 200 μm apart. This is approximately 100 times better than what was previously demonstrated in DFF-DH [9].

APPENDIX A

In this appendix we show that the focus curve of a white noise input to the DH imaging system is given by the sum of Talbot curves [see Eq. (8)] for all spatial frequencies.

Using Eq. (8), the first term of Eq. (13) can be written as

$$\sum_{n=1}^{N} \text{var}(|U_n|^2) = Na + b\sum_{n=1}^{N} \cos\left(\frac{2\pi\lambda n^2(z_0 + z_r)}{L^2}\right). \quad \text{(A1)}$$

Using Eq. (5), the second term of equation Eq. (13) can be rewritten as

$$\sum_{n=1}^{N} \sum_{m=1}^{N} U_n U_m^*(1 - \delta_{n,m})$$

$$= \frac{1}{4N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\cos\left(\frac{\pi\lambda(n^2 - m^2)(z_0 + z_r)}{L^2}\right)$$

$$\times \cos\left(\frac{2\pi n\xi}{L} + \phi_n\right)\cos\left(\frac{2\pi m\xi}{L} + \phi_m\right)(1 - \delta_{n,m})$$

$$+ \frac{N-1}{2N^2}\sum_{n=1}^{N}\cos\left(\frac{\pi\lambda n^2(z_0 + z_r)}{L^2}\right)\cos\left(\frac{2\pi n\xi}{L} + \phi_n\right)$$

$$+ \frac{N(N-1)}{4N^2}. \quad \text{(A2)}$$

Using the independence property for the variance, we obtain for the variance of this expression

$$\text{var}(I_n) = \text{var}\left(\sum_{n=1}^{N}\sum_{m=1}^{N} U_n U_m^*(1 - \delta_{n,m})\right)$$

$$= \text{var}\left(\sum_{n=1}^{N}\sum_{m=1}^{N}\cos\left(\frac{\pi\lambda(n^2 - m^2)(z_0 + z_r)}{L^2}\right)\right.$$

$$\times \cos\left(\frac{2\pi n\xi}{L} + \phi_n\right)\cos\left(\frac{2\pi m\xi}{L} + \phi_m\right)\frac{(1 - \delta_{n,m})}{2\,N^2}\right)$$

$$+ \text{var}\left(\frac{N-1}{2\,N^2}\sum_{n=1}^{N}\cos\left(\frac{\pi\lambda n^2(z_0 + z_r)}{L^2}\right)\right.$$

$$\times \cos\left(\frac{2\pi n\xi}{L} + \phi_n\right)\right). \quad \text{(A3)}$$

The first term is equal to

$$\text{var} = \frac{L}{32\,N^4}\sum_{n=1}^{N}\sum_{m=1}^{N}\cos\left(\frac{2\pi\lambda(n^2 - m^2)(z_0 + z_1)}{L^2}\right) + 1. \quad \text{(A4)}$$

The last term is equal to

$$\text{var} = \frac{L(N-1)^2}{16\,N^4}\sum_{n=1}^{N}\left(\cos\left(\frac{2\pi\lambda n^2(z_0 + z_1)}{L^2}\right) + 1\right). \quad \text{(A5)}$$

From an analogy with Fourier analysis we can observe that the peak width of $\sigma^2(I_N)$ is limited by the higher spatial frequencies caused by the n^2 terms in the argument of the cosine, and not by the difference terms $n^2 - m^2$, which correspond to lower spatial frequencies. We thus approximate the variance curve as a linear sum of the Talbot curves resulting from the n^2/L^2 frequencies in the variance curve:

$$\text{var}_N(z_r) = A + \sum_{n=1}^{N} B\cos\left(\frac{2\pi\lambda n^2(z_0 + z_r)}{L^2}\right), \quad \text{(A6)}$$

where

$$A = N\left(a + \frac{L(N-1)^2 + L}{16N^4}\right) \qquad \text{(A7)}$$

and

$$B = \frac{L(N-1)^2 + L}{16N^4}. \qquad \text{(A8)}$$

REFERENCES

1. R. Leach, *Optical Measurement of Surface Topography* (Springer, 2011).
2. P. de Groot, "Phase shifting interferometry," in *Optical Measurement of Surface Topography*, R. Leach, ed. (Springer, 2011), Chap. 8, pp. 167–186.
3. P. de Groot, "Coherence scanning interferometry," in *Optical Measurement of Surface Topography*, R. Leach, ed. (Springer, 2011), Chap. 9, pp. 187–208.
4. Z. Wang, B. Potsaid, L. Chen, C. Doerr, H.-C. Lee, T. Nielson, V. Jayaraman, A. E. Cable, E. Swanson, and J. G. Fujimoto, "Cubic meter volume optical coherence tomography," Optica **3**, 1496–1503 (2016).
5. J. K. Tristan Colomb, "Digital holographic microscopy," in *Optical Measurement of Surface Topography*, R. Leach, ed. (Springer, 2011), Chap. 10, pp. 209–235.
6. T. J. T. Abregana and P. F. Almoro, "Object localization using the statistical behavior of volume speckle fields," Opt. Eng. **55**, 121720 (2016).
7. N. Nikolaev, J. Petzing, and J. Coupland, "Focus variation microscope: linear theory and surface tilt sensitivity," Appl. Opt. **55**, 3555–3565 (2016).
8. F. Helmli, "Focus variation instruments," in *Optical Measurement of Surface Topography*, R. Leach, ed. (Springer, 2011), Chap. 7, pp. 131–166.
9. M. L. Tachiki, M. Itoh, and T. Yatagai, "Simultaneous depth determination of multiple objects by focus analysis in digital holography," Appl. Opt. **47**, D144–D153 (2008).
10. L. Ma, H. Wang, Y. Li, and H. Jin, "Numerical reconstruction of digital holograms for three-dimensional shape measurement," J. Opt. A **6**, 396–400 (2004).
11. C. P. McElhinney, B. M. Hennelly, and T. J. Naughton, "Extended focused imaging for digital holograms of macroscopic three-dimensional objects," Appl. Opt. **47**, D71–D79 (2008).
12. U. Schnars and W. Jueptner, "Digital holography," in *Digital Holography* (Springer, 2005), pp. 41–69.
13. J. Chang Li and P. Picart, "Digital off-axis Fresnel holography," in *Digital Holography* (Wiley, 2013), pp. 165–236.
14. Y. Tian, K. Shieh, and C. F. Wildsoet, "Performance of focus measures in the presence of nondefocus aberrations," J. Opt. Soc. Am. A **24**, B165–B173 (2007).
15. M. Subbarao and G. Surya, "Depth from defocus: a spatial domain approach," Int. J. Comput. Vis. **13**, 271–294 (1994).
16. J. Goodman, *Introduction to Fourier Optics*, 2nd ed. (MaGraw-Hill, 1996).
17. J. Breckinridge and D. Voelz, *Computational Fourier Optics: A MATLAB Tutorial*, SPIE Press monograph (Society of Photo Optical, 2011).
18. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," Opt. Lett. **22**, 1268–1270 (1997).
19. N. Verrier and M. Atlan, "Off-axis digital hologram reconstruction: some practical considerations," Appl. Opt. **50**, H136–H146 (2011).