# Automatic procedure for aberration compensation in digital holographic microscopy and applications to specimen shape compensation 

Tristan Colomb, Etienne Cuche, Florian Charrière, Jonas Kühn, Nicolas Aspert, Frédéric Montfort, Pierre Marquet, and Christian Depeursinge


#### Abstract

We present a procedure that compensates for phase aberrations in digital holographic microscopy by computing a polynomial phase mask directly from the hologram. The phase-mask parameters are computed automatically without knowledge of physical values such as wave vectors, focal lengths, or distances. This method enables one to reconstruct correct and accurate phase distributions, even in the presence of strong and high-order aberrations. Examples of applications are shown for microlens imaging and for compensating for the deformations associated with a tilted thick plate. Finally we show that this method allows compensation for the curvature of the specimen, revealing its surface defects and roughness. Examples of applications are shown for microlenses and metallic sphere imaging. © 2006 Optical Society of America

OCIS codes: $\quad 090.1760,090.1000,180.0180,120.5050,120.3620$.


## 1. Introduction

Digital holographic microscopy (DHM) is a powerful instrument for the study of microscopic samples by retrieving the amplitude and phase of the wave reflected by the sample or transmitted through it. Phase reconstruction is of particular interest since it allows surface topography measurements with nanometer vertical resolution. ${ }^{1}$ Several methods have been proposed for phase measurements in digital holography. In-line techniques ${ }^{2}$ use phase-shifting procedures that require several hologram acquisitions and additional means, such as a piezoelectric transducer (PZT), for controlling the phase of the reference wave. In terms of off-axis configurations, Schnars ${ }^{3}$ demonstrated the possibility of measuring specimen deformations by evaluating the phase difference be-

[^0]tween two states of the sample. However, this doubleexposure technique does not give the absolute phase of the sample.

A solution for absolute phase measurements was proposed by Cuche et al. ${ }^{4}$ in 1999. They introduced the concept of a digital reference wave, ${ }^{1}$ which compensates for the role of the reference wave in an off-axis geometry, and the concept of a digital phase mask, ${ }^{4,5}$ which was introduced for compensating the wavefront deformation associated with the use of a microscope objective (MO). These two quantities are defined numerically as arrays of complex numbers, computed by using a set of parameters called phasereconstruction parameters (PRPs), whose values must be precisely defined so that these computed data fit as closely as possible to their experimental equivalences. This approach allows absolute phase reconstruction with a single hologram acquisition. However, as proposed initially, the procedure for PRP adjustment was not automated and required an a priori knowledge of the PRP values, making the method not user friendly.

To avoid this adjustment, Ferraro et al. ${ }^{6}$ proposed in 2003 the use of a reference hologram, taken in a flat part of the sample, and a subtraction procedure for aberration compensation. This simple technique performs efficiently if the experimental presence of parasitic conditions (vibrations, drifts, etc.) are well controlled, but it fails in providing precise measure-


Fig. 1. Schematic of the holographic microscope for reflection imaging: NF, neutral filter; $\lambda / 2$, half-wave plate; PBS, polarizing beam splitter; BS, beam splitter; BE, beam expander; M, mirror; $\boldsymbol{O}$, object wave; $\boldsymbol{R}$, reference wave. Inset, detail of the off-axis geometry.
ments in most of the situations and is therefore not adapted for most practical applications.
Here we propose a new approach that merges the digital reference wave and the digital phase mask (DPM) in a single entity, whose computation is now performed by an automated, simple, and fast procedure. The method is based on the extraction of reconstructed-phase values along line profiles, located in or around the sample, and in areas that are assumed to be flat and that serve as reference surfaces. The PRPs are then automatically adjusted by applying curve-fitting procedures on the extracted phase profiles. We show that this approach allows for the compensation of the tilt aberration due to the off-axis geometry, the quadratic-phase curvature introduced by the MO, as well as the wavefront deformations induced by any of the microscope components, even those caused by higher-order aberrations, as demonstrated by the introduction of a thick tilted glass plate in the experimental setup.

## 2. Experimental Configurations

Two main configurations exist for the implementation of DHM: the first one (Fig. 1) for reflection imaging and the second one (Fig. 2) for transmission imaging. In both cases the basic architecture is that of a modified Mach-Zehnder interferometer. The light source depends on applications (in Refs. 1 and 4 a $\mathrm{He}-\mathrm{Ne}$ laser is used, but lower-coherence sources can also be used). ${ }^{7}$ The combination of a neutraldensity filter, a half-wave plate, and a polarizing beam splitter is used for the adjustment of the intensities in the reference and object arms. We can notice here that a modified setup ${ }^{8,9}$ with two references


Fig. 2. Schematic of the holographic microscope for transmission imaging. Inset, detail of the off-axis geometry.
arms with orthogonal polarization was developed to record the polarization state of the object wave.

In both configurations a MO collects the object wave $\boldsymbol{O}$ transmitted or reflected by the specimen and produces a magnified image of the sample behind the CCD camera at a distance $d$ of $\sim 5 \mathrm{~cm}$ (Fig. 3). As explained in detail in Ref. 4, this situation can be considered to be equivalent to a holographic configuration without a MO, with an object wave $\boldsymbol{O}$ emerging directly from the magnified image of the sample and not from the sample itself. A high level of accuracy can be obtained by using a MO with a high numerical aperture (NA). Indeed, the role of this high-NA MO is to provide a simple means for adapting the sampling capacity of the camera to the information content of the hologram. As illustrated in Fig. 3, a lens or MO achieves a reduction of the $K_{x}, K_{y}$ components of the $\boldsymbol{K}$ vector components in the sample plane perpendicular to the optical axis. The reduction factor is given by the magnification of the MO. The new components $K_{x}^{\prime}, K_{y}{ }^{\prime}$ of the $\boldsymbol{K}^{\prime}$ wave vector of the beam after they have crossed the MO can be made as small as necessary to satisfy the Shannon theorem to the sampling


Fig. 3. Use of a lens or MO to match the sampling capacity of a CCD placed in the plane of the hologram with the spatial spectrum of the object.
capacity, dictated by the pixel size of the camera. Using a high-magnification objective allows the match to be optimized. At the same time, by maximizing the NA, the transverse resolution can be pushed to the limit of diffraction and submicrometer resolution can be easily achieved (ordinarily better than 600 nm ).
At the exit to the interferometer, the interference between the object wave $\boldsymbol{O}$ and the reference wave $\boldsymbol{R}$ creates the hologram intensity

$$
\begin{equation*}
I_{H}(x, y)=|\boldsymbol{R}|^{2}+|\boldsymbol{O}|^{2}+\boldsymbol{R} \boldsymbol{O}^{*}+\boldsymbol{R}^{*} \boldsymbol{O} . \tag{1}
\end{equation*}
$$

A digital hologram is recorded by a black-and-white CCD camera and transmitted to a computer. The digital hologram $I_{H}(k, l)$ is an array of $N \times N$ (usually $512 \times 512$ or $1024 \times 1024$ ) 8 -bit-encoded numbers resulting from the two-dimensional (2D) sampling of $I_{H}(x, y)$ by the CCD camera:

$$
\begin{equation*}
I_{H}(k, l)=\int_{k \Delta x-\Delta x / 2}^{k \Delta x+\Delta x / 2} \int_{l \Delta y-\Delta y / 2}^{l \Delta y+\Delta y / 2} I_{H}(x, y) \mathrm{d} x \mathrm{~d} y, \tag{2}
\end{equation*}
$$

where $k, l$ are integers and $\Delta x, \Delta y$ define the sampling intervals in the hologram plane (pixel size).

## 3. Hologram Reconstruction

## A. Former Algorithm

In Ref. 4 the algorithm for hologram reconstruction was formulated as follows:

$$
\begin{align*}
\Psi(\xi, \eta)= & \Phi(\xi, \eta) \frac{\exp (i 2 \pi d / \lambda)}{i \lambda d} \exp \left[\frac{i \pi}{\lambda d}\left(\xi^{2}+\eta^{2}\right)\right] \\
& \times \iint \boldsymbol{R}_{D}(x, y) I_{H}(x, y) \\
& \times \exp \left\{\frac{i \pi}{\lambda d}\left[(x-\xi)^{2}+(y-\eta)^{2}\right]\right\} \mathrm{d} x \mathrm{~d} y . \tag{3}
\end{align*}
$$

This expression describes the Fresnel propagation of the reconstructed wavefront $\Psi$ over a distance $d$, from the hologram plane $0 x y$, to the observation plane $0 \xi \eta$. The digital reference wave $\boldsymbol{R}_{\boldsymbol{D}}$ and the digital phase mask $\Phi$, both introduced in Ref. 4, play major roles in phase reconstruction. $\boldsymbol{R}_{\boldsymbol{D}}$ is defined as a computed replica of the experimental reference wave $\boldsymbol{R}$. If we assume that the hologram was recorded in the off-axis geometry with a plane wave as reference, $\boldsymbol{R}_{\boldsymbol{D}}$ is defined as follows:

$$
\begin{equation*}
\boldsymbol{R}_{\boldsymbol{D}}(x, y)=\exp \left[i \frac{2 \pi}{\lambda}\left(k_{x} x+k_{y} y\right)+i \phi(t)\right], \tag{4}
\end{equation*}
$$

where the parameters $k_{x}, k_{y}$ define the propagation direction and $\phi(t)$ is the phase delay between the object and the reference waves, which may vary during time owing to external perturbations, such as
mechanical vibration. As explained in Ref. 4, for proper phase reconstruction, the $k_{x}$ and $k_{y}$ values must be adjusted so that the propagation direction of the computed wavefront $\boldsymbol{R}_{\boldsymbol{D}}$ fits the propagation direction of the experimental wave $\boldsymbol{R}$.

As explained in Ref. 4, the role of the DPM $\Phi$ is to compensate for the wavefront curvature that appears when a MO is used to improve the magnification and the transverse resolution. In Ref. 4 a simple quadratic model was used for the computation:

$$
\begin{equation*}
\Phi(\xi, \eta)=\exp \left[\frac{-i \pi}{\lambda D}\left(\xi^{2}+\eta^{2}\right)\right], \tag{5}
\end{equation*}
$$

where $D$ is the parameter that can be adjusted to compensate for this curvature. The reconstruction procedure of Eq. (3) involves four parameters: $k_{x}, k_{y}$, and $D$ for phase reconstruction and the reconstruction distance $d$ for image focusing.
This formulation suffers from several drawbacks:
(i) The digital reference wave is located inside the Fresnel integral; therefore changing the $k_{x}$ and $k_{y}$ values induces a translation of the reconstructed image in the observation plane.
(ii) The model used for calculating the DPM is limited to the second order and therefore fails to compensate for higher-order aberrations (e.g., astigmatism and spherical aberration).
(iii) The adjustment of the $k_{x}, k_{y}$, and $D$ parameters is a complex task that requires expertise and $a$ priori knowledge of their values.

We now describe how the reconstruction procedure can be improved to address these drawbacks.

## B. Digital Reference Wave Outside the Integral

We demonstrate here how the digital reference wave $\boldsymbol{R}_{\boldsymbol{D}}$ in Eq. (3) can be replaced by a pseudoreference wave $\boldsymbol{R}^{\prime}$ outside the integral.
First, let us define the 2D Fresnel transform $\mathfrak{F}_{\tau}$ of parameter $\tau$ as

$$
\begin{align*}
\mathfrak{F}_{\tau}[f(x, y)](\xi, \eta)= & \frac{1}{\tau^{2}} \iint f(x, y) \exp \left\{\frac { i \pi } { \tau ^ { 2 } } \left[(x-\xi)^{2}\right.\right. \\
& \left.\left.+(y-\eta)^{2}\right]\right\} \mathrm{d} x \mathrm{~d} y, \tag{6}
\end{align*}
$$

which has the modulation property ${ }^{10}$

$$
\begin{aligned}
& \mathfrak{J}_{\tau}[\exp (i 2 \pi \boldsymbol{v} \boldsymbol{x}) f(\boldsymbol{x})](\boldsymbol{\xi},\eta)=\exp (i 2 \pi \boldsymbol{\nu} \xi) \exp \left(-i \pi \boldsymbol{\nu}^{2} \tau^{2}\right) \\
& \times \mathfrak{F}_{\tau}[f(\boldsymbol{x})]\left(\xi-v_{x} \tau^{2}, \eta-v_{y} \tau^{2}\right),
\end{aligned}
$$

where $\boldsymbol{x}=(x, y), \boldsymbol{\xi}=(\xi, \eta)$, and $\boldsymbol{v}=\left(v_{x}, \nu_{y}\right)$. With this definition, Eq. (3) can be written as

$$
\begin{align*}
\Psi(\xi, \eta)= & -i \Phi(\xi, \eta) \exp (i 2 \pi d / \lambda) \mathfrak{F}_{\tau}\left[\boldsymbol{R}_{\mathbf{D}} I_{H}\right](\xi, \eta), \\
\tau= & \sqrt{\lambda d} \\
= & -i \Phi(\xi, \eta) \exp (i 2 \pi d / \lambda) \mathfrak{F}_{\tau}\left\{\operatorname { e x p } \left[i \frac{2 \pi}{\lambda}\right.\right. \\
& \left.\left.\times\left(k_{x} x+k_{y} y\right)+i \phi(t)\right] I_{H}\right\}(\xi, \eta) . \tag{8}
\end{align*}
$$

Using the property of Eq. (7) with $\nu_{x}=k_{x} / \lambda$ and $\nu_{y}$ $=k_{y} / \lambda$, Eq. (8) becomes

$$
\begin{align*}
\Psi(\xi, \eta)= & -i \Phi(\xi, \eta) \exp (i 2 \pi d / \lambda) \\
& \times \exp \left[i \frac{2 \pi}{\lambda}\left(k_{x} \xi+k_{y} \eta\right)+i \phi(t)\right] \\
& \times \exp \left\{-i \pi \lambda d\left[\left(\frac{k_{x}}{\lambda}\right)^{2}+\left(\frac{k_{y}}{\lambda}\right)^{2}\right]\right\} \\
& \times \mathscr{F}_{\tau}\left[I_{H}\right]\left(\xi+k_{x} d, \eta+k_{y} d\right) . \tag{9}
\end{align*}
$$

Writing

$$
\begin{equation*}
\phi^{\prime}\left(k_{x}, k_{y}, t\right)=\phi(t)-\pi d\left(k_{x}^{2}+k_{y}^{2}\right) / \lambda \tag{10}
\end{equation*}
$$

allows us to use the property of shift invariance of the Fresnel transform to write the wavefront as

$$
\begin{align*}
\Psi(\xi, \eta)= & -i \Phi(\xi, \eta) \exp (i 2 \pi d / \lambda) \boldsymbol{R}^{\prime}(\xi, \eta) \\
& \times \mathfrak{F}_{\tau}\left[I_{H}\right](\xi, \eta) \tag{11}
\end{align*}
$$

where $\boldsymbol{R}^{\prime}$ is the pseudoreference wave

$$
\begin{equation*}
\boldsymbol{R}^{\prime}(\xi, \eta)=\exp \left[i \frac{2 \pi}{\lambda}\left(k_{x} \xi+k_{y} \eta\right)+i \phi^{\prime}\left(k_{x}, k_{y}, t\right)\right] . \tag{12}
\end{equation*}
$$

This means that passing the digital reference wave outside the Fresnel integral is a straightforward operation that conserves the plane-wave nature of the wavefront. In other words, for phase reconstruction with DHM, it means that the effect of the off-axis geometry is the appearance of a tilt aberration in the observation plane. This tilt can be compensated for by multiplying the reconstructed wavefront with a correcting term calculated with the mathematical model of a plane wave. Compared with the former method [Eq. (3)], here the great advantage is that the adjustment of this pseudoreference wave does not produce a translation of the image in the reconstruction plane.

## C. Generalized Digital Phase Mask

Equation (11) describes the reconstruction algorithm as the Fresnel transform of the hologram intensity $I_{H}$ multiplied by the product of the DPM $\Phi(\xi, \eta)$ with the pseudo-digital reference wave $\boldsymbol{R}^{\prime}$. As $\Phi(\xi, \eta)$ and $\boldsymbol{R}^{\prime}(\xi, \eta)$ appear now out of the Fresnel integral, they can be merged in a single entity, and the reconstruction algorithm becomes

$$
\begin{equation*}
\Psi(\xi, \eta)=-i \exp (i 2 \pi d / \lambda) \Gamma(\xi, \eta) \mathscr{F}_{\tau}\left[I_{H}\right](\xi, \eta) \tag{13}
\end{equation*}
$$

where, according to Eqs. (5) and (12), we have

$$
\begin{align*}
\Gamma(\xi, \eta)= & \exp \left[\frac{i \pi}{\lambda}\left(2 k_{x} \xi+2 k_{y} \eta-\frac{\xi^{2}+\eta^{2}}{D}\right)\right. \\
& \left.+i \phi^{\prime}\left(k_{x}, k_{y}, t\right)\right] \tag{14}
\end{align*}
$$

This new formulation of the DPM involves four reconstruction parameters: $k_{x}$ and $k_{y}$ for the compensation of the tilt aberration due to the off-axis geometry, the phase offset $\phi^{\prime}$ for compensating the phase delay between the object and the reference waves, and $D$ for compensating for a quadratic wavefront curvature.

This expression for the DPM covers a limited range of situations. To allow the correction of higher-order aberrations, we introduce here a generalized polynomial formulation:

$$
\begin{equation*}
\Gamma(\xi, \eta)=\exp \left[-i \sum_{\alpha=0}^{H} \sum_{\beta=0}^{V} P_{\alpha, \beta} \xi^{\alpha} \eta^{\beta}\right] \tag{15}
\end{equation*}
$$

where $P_{\alpha, \beta}$ define a set of reconstruction parameters and $H$ and $V$ define the polynomial orders in the horizontal and vertical directions, respectively. The physical constants ( $\lambda, D$, and $\pi$ ) are suppressed from the definition of the $P_{\alpha, \beta}$ parameters. But the corresponding physical quantities can be evaluated if necessary. For example,

$$
\begin{align*}
& P_{0,0}=-\phi^{\prime} \\
& P_{1,0}=-\frac{2 \pi}{\lambda} k_{x}, \quad P_{0,1}=-\frac{2 \pi}{\lambda} k_{y} \\
& P_{2,0}=P_{0,2}=\frac{\pi}{\lambda D} \tag{16}
\end{align*}
$$

## D. Discrete Formulation

The numerical calculation of Eq. (13) can be performed efficiently, following the discrete formulations of the Fresnel integral. As explained in Ref. 1, a discrete formulation of Eq. (13) involves a discrete Fourier transform (DFT) and can be derived directly as follows:

$$
\begin{align*}
\Psi^{\prime}(m, n)= & A \Gamma(m, n) \exp \left[\frac{i \pi}{\lambda d}\left(m^{2} \Delta \xi^{2}+n^{2} \Delta \eta^{2}\right)\right] \\
& \times \operatorname{DFT}\left\{I _ { H } ( k , l ) \operatorname { e x p } \left[\frac { i \pi } { \lambda d } \left(k^{2} \Delta x^{2}\right.\right.\right. \\
& \left.\left.\left.+l^{2} \Delta y^{2}\right)\right]\right\}_{m, n} \tag{17}
\end{align*}
$$

where $m$ and $n$ are integers $(-N / 2 \leq m, n<N / 2)$, $A=\exp (i 2 \pi d / \lambda) /(i \lambda d)$. The sampling intervals in the
observation plane, $\Delta \xi$ and $\Delta \eta$ are defined as follows:

$$
\begin{equation*}
\Delta \xi=\Delta \eta=\lambda d /(N \Delta x) \tag{18}
\end{equation*}
$$

The discrete formulation of the DPM is simply

$$
\begin{equation*}
\Gamma(m, n)=\exp \left[-i \sum_{\alpha=0}^{H} \sum_{\beta=0}^{V} P_{\alpha, \beta}(m \Delta \xi)^{\alpha}(n \Delta \eta)^{\beta}\right] \tag{19}
\end{equation*}
$$

For computational purposes, Eq. (17) can be further simplified by introducing the quadratic phase term that multiplies the DFT inside the DPM and by suppressing the definition of the sampling intervals $\Delta \xi$ and $\Delta \eta$, which are related to the following physical scales:

$$
\begin{align*}
\Psi(m, n)= & A \Gamma(m, n) \operatorname{DFT}\left\{I_{H}(k, l)\right. \\
& \left.\times \exp \left[\frac{i \pi}{\lambda d}\left(k^{2} \Delta x^{2}+l^{2} \Delta y^{2}\right)\right]\right\}_{m, n} \tag{20}
\end{align*}
$$

with

$$
\begin{equation*}
\Gamma(m, n)=\exp \left(-i \sum_{\alpha=0}^{H} \sum_{\beta=0}^{V} P_{\alpha, \beta} m^{\alpha} n^{\beta}\right) \tag{21}
\end{equation*}
$$

In this case the physical definition of the parameters becomes

$$
\begin{align*}
& P_{0,0}=-\phi^{\prime} \\
& P_{1,0}=-\frac{2 \pi}{\lambda} k_{x} \Delta \xi, \quad P_{0,1}=-\frac{2 \pi}{\lambda} k_{y} \Delta \eta \\
& P_{2,0}=P_{0,2}=\frac{\pi}{\lambda}\left(\frac{1}{D}-\frac{1}{d}\right) \Delta \xi^{2} \tag{22}
\end{align*}
$$

A further simplification is possible by suppressing all physical constants $(A, \pi, \lambda, \Delta x, \Delta y, \Delta \xi, \Delta \eta)$, except the distance $d$, which is important for image focusing. The reconstructed wavefront becomes

$$
\begin{align*}
\Psi^{\prime}(m, n)= & \Gamma^{\prime}(m, n) \operatorname{DFT}\left\{I_{H}(k, l)\right. \\
& \left.\times \exp \left[i / d\left(k^{2}+l^{2}\right)\right]\right\}_{m, n} \\
= & \Gamma^{\prime}(m, n) \Omega(m, n), \tag{23}
\end{align*}
$$

where $\Omega(m, n)$ is the reconstructed wavefront in the Fresnel approximation at distance $d$ without phase correction, and the polynomial DPM is

$$
\begin{equation*}
\Gamma^{\prime}(m, n)=\exp \left(-i \sum_{\alpha=0}^{H} \sum_{\beta=0}^{V} P_{\alpha, \beta} m^{\alpha} n^{\beta}\right) \tag{24}
\end{equation*}
$$

The above-described operations can similarly be adapted to the convolution formalism.

## 4. Phase-Mask Adjustment

The method for aberration compensation is based on four main steps:
(i) Defining a mathematical model for calculating the DPM. As explained previously, the mathematical model is here the complex function $\Gamma^{\prime}(m, n)$ of constant amplitude, whose phase is defined by a 2D polynomial function. The polynomial coefficients $P_{\alpha, \beta}$ define the PRPs.
(ii) Measuring the PRPs $P_{\alpha, \beta}$ by evaluating the contributions of the off-axis geometry and of the experimental setup. As explained in the next section in detail, this task is achieved by fitting polynomial curves along selected profiles extracted from the area of the reconstructed-phase distribution where the specimen contribution is assumed to be a constant.
(iii) Using the measured values of the PRPs $P_{\alpha, \beta}$ to calculate the 2D polynomial phase mask $\Gamma^{\prime}(m, n)$.
(iv) Multiplying the reconstructed wavefront $\Omega(m, n)$ by the DPM $\Gamma^{\prime}(m, n)$ to obtain a corrected wavefront from which we can obtain the phase distribution associated with the specimen.

As explained in the next section, a few iterations of this process may be necessary to retrieve the correct phase distribution, especially if one starts from completely unknown values of the PRPs.

## A. Adjustment of First- and Second-Order Parameters

We now describe the iterative procedure in more detail, and we first restrict the case to the adjustment of the first five PRPs $\left(P_{0,0}, P_{1,0}, P_{0,1}, P_{2,0}, P_{0,2}\right)$, which are necessary to define a second-order phase mask without crossed terms.

To illustrate the procedure, we used a hologram of a U.S. Air Force test target taken in a reflection setup (see Fig. 1) with a $40 \times \mathrm{MO}$. The hologram is presented in Fig. 4. We can see the curvature of the fringe pattern due to the presence of the MO .

After hologram apodization ${ }^{11}$ and spatial frequency filtering, ${ }^{12}$ the reconstructed wavefront $\Omega(m, n)$ is computed and the region of interest (ROI), corresponding to the location of the real image, is delineated inside the image plane.

Figure 5(a) presents the ROI of the phase image of the wavefront $\Omega(m, n)$. As no phase mask is applied and as the phase is defined modulo $2 \pi$, the phase seems to be randomly distributed. In fact, this phase reconstruction contains the information that will be used to adjust automatically the PRP. Indeed, the reconstructed phase of Fig. 5(a) is the addition of three contributions:
(i) The absolute phase of the sample.
(ii) The tilt aberration due to the off-axis geometry.
(iii) The aberrations or wavefront deformations induced by the setup, particularly the curvature produced by the MO.

The method for PRP adjustment consists of evaluating the last two contributions in the area where the


Fig. 4. Digital hologram of a U.S. Air Force test target recorded with a $40 \times \mathrm{MO}$ in a reflection setup.
specimen contribution is known to be constant or, in other words, in a flat reference area located near or on the sample.
The first operation consists of defining two perpendicular lines [see Fig. 5(a)], one for each direction $m$ and $n$, which are, respectively, the horizontal and vertical directions (the origin is defined in the center of the ROI). Along each of these two lines, phase profiles are extracted. With strongly aberrated wavefronts such as those presented in Fig. 5(a), the phase variation along the selected profiles exceeds $2 \pi$ several times. This produces phase jumps (between $-\pi$ and $\pi$ ) that are suppressed by applying a standard one-dimensional (1D) phase-unwrapping algorithm. The resulting continuous-phase profiles are then fitted with a 1D polynomial function. Let us assume a correction limited to the second order, and let us define

$$
\begin{align*}
& Y_{h}=a_{0}+a_{1} x+a_{2} x^{2}, \\
& Y_{v}=b_{0}+b_{1} x+b_{2} x^{2}, \tag{25}
\end{align*}
$$

as the two curves fitted along the horizontal and the vertical profiles, respectively. In this case the DPM becomes simply

$$
\begin{equation*}
\Gamma^{\prime}(m, n)=\exp \left[-i\left(a_{1} m+b_{1} n+a_{2}{ }^{2} m^{2}+b_{2}{ }^{2} n^{2}\right)\right], \tag{26}
\end{equation*}
$$

where $a_{1}=P_{1,0}$ and $b_{1}=P_{0,1}$ correct the tilt aberration along the horizontal and the vertical directions, respectively, and where $a_{2}=P_{2,0}$ and $b_{2}=P_{0,2}$ correct a second-order wavefront curvature along the horizontal and the vertical directions, respectively.
In principle, this procedure should work immedi-


Fig. 5. Automatic adjustment of a second-order phase mask. The left column presents phase images obtained for different values of the PRPs, which are evaluated by fitting a second-order 1D polynomial function on unwrapped phase data extracted along two profiles [horizontal and vertical lines in (a)-(d)]. The right column presents plots of unwrapped phase data, and fitted curves, for the horizontal direction. In the presence of strong aberrations and with zero parameter values (a), the phase unwrapping and fitting procedures fail to define directly the correct PRP values, which are reached after two iterations (c). The correct PRPs can be obtained only with the profile position located in flat regions where the absolute phase contributions of the specimen are constant, as in (d).
ately, but as shown in Fig. 5(b), residual aberrations may persist. The reason for this is that, without correction [Fig. 5(a)] or with initial PRPs that are too far from the correct values, the phase distribution varies so rapidly that several phase jumps may occur over distances smaller than 1 pixel. In this case the phaseunwrapping procedure cannot work properly. To avoid this effect, we apply the procedure described
above iteratively. Starting from initial values $P_{\alpha, \beta}{ }^{(0)}$ provided by a first evaluation, the curve-fitting procedure is applied iteratively. At each iteration, the coefficients of the fitted curve are added to the PRP from the previous iteration. Let us define the fitting curves of the two profiles defined on the reconstructed phase image at iteration $i$ :

$$
\begin{gather*}
Y_{h}{ }^{(i)}=a_{0}{ }^{(i)}+a_{1}{ }^{(i)} x+a_{2}{ }^{(i)} x^{2}, \\
Y_{v}{ }^{(i)}=b_{0}{ }^{(i)}+b_{1}{ }^{(i)} x+b_{2}{ }^{(i)} x^{2} . \tag{27}
\end{gather*}
$$

The new PRPs become

$$
\begin{align*}
& P_{j, 0}{ }^{(i)}=P_{j, 0}{ }^{(i-1)}+a_{j}{ }^{(i)}, \\
& P_{0, j}{ }^{(i)}=P_{0, j}{ }^{(i-1)}+b_{j}{ }^{(i)}, \tag{28}
\end{align*}
$$

where $j=1,2$. The procedure converges rapidly to the optimal values, and fewer than four iterations are sufficient in most cases [see Fig. 5(c)].

## B. Profile Positioning

A key point for retrieving correct and accurate phase distributions is that the phase profiles, along which the polynomial functions are fitted, must be extracted from the image area in which the sample contribution is known to be constant. As shown in Fig. 5(d), if the profile is located on a nonflat area (step-height difference or phase change on the sample), the PRPs are not adjusted correctly.

This condition represents a limitation for the application of the method. However, as described above, this condition can be satisfied for a large variety of specimens, especially microscopic ones.

The first and most simple case occurs when the sample itself comprises the flat reference surface. This case is valid for most microfabricated or nanofabricated structures. With thin films, the reference surface can be the film itself or the substrate on which the deposition has been made. With flat layers (formed by deposition, evaporation, chemical attack, etc.) or with all semiconductor devices, the substrate can be chosen as a flat reference surface.

With biological specimens and organisms (cells and tissues in particular) observed in air or in liquids, the standard preparations used for investigations with other microscopy methods are also well suited. Most of these preparations involve plastic, glass, semiconductor, or metallic surfaces that offer opportunities to define a reference surface near the specimen. For example, with cells cultured on plastic or glass surfaces, the reference surface can be defined on the substrate between the cells. Finally, if the specimen is not specifically attached to a flat surface, it is sufficient in most cases to put the specimen on a flat sample holder, which will act as a reference surface. For transmission imaging, the sample holder must be transparent (e.g., glass or plastic plate). For reflection
imaging, the sample holder must be reflective (e.g., a metallic or dielectric surface).

Finally, if a specific ROI on a sample does not allow a flat reference surface to be defined, it is possible to adjust the PRPs elsewhere on the sample or on the sample holder. The sample can then be translated to obtain an image of the desired ROI.

## C. Multiprofile Procedure

As explained in the preceeding paragraph, the procedure requires that profiles be extracted on flat areas. It may be that only small areas of the field of view are flat. In this case the procedure for PRP adjustment may lack precision because the accuracy of the curvefitting procedure depends on the lengths of the extracted profiles. In other words, the longer the extracted profiles are, the better the procedure for parameter adjustment is.

Figure 6 presents the reconstruction of a wafer of quartz microlenses imaged with a $10 \times \mathrm{MO}$ in a transmission setup (see Fig. 2). Since the lenses with a diameter of $230 \mu \mathrm{~m}$ are distributed densely on the wafer, the flat areas around the lenses are small and do not provide sufficiently long profiles [see Figs. 6(a) and 6(b)]. To solve this problem, one can apply a multiprofile procedure.
For simplicity, and because the procedure is identical for both the horizontal and the vertical profiles, we restrict the discussion here to the second-order correction along the horizontal direction. Let us define $N_{h}$ horizontal profiles located at different places $p$ on or around the specimen, in which a portion of the reference surfaces can be defined. The $p$ th profile (1 $\left.\leq p \leq N_{h}\right)$ has $n_{p}$ points. Let us also define $Y(x, p)$ as the polynomial fitted on the $p$ th profile:

$$
\begin{equation*}
Y(x, p)=a_{0}(p)+a_{1}(p) x+a_{2}(p) x^{2} \tag{29}
\end{equation*}
$$

Because of noncompensated aberrations, different profile positions provide different phase offsets $a_{0}(p)$, but the superior-order coefficients should be identical as long as the corresponding aberrations are homogeneously distributed in the field of view. A first approach to define the parameters $a_{j}$ is simply to calculate the mean value:

$$
\begin{equation*}
a_{j}=\frac{1}{N_{h}} \sum_{p=1}^{N_{h}} a_{j}(p) \tag{30}
\end{equation*}
$$

With this method, each profile has the same weight to compute the coefficient $a_{j}$. In other words, the weight attributed to each phase value measured on the profile depends on the length of the profile: The shorter the profile length is, the larger the point weight is. Therefore another method was developed to restore an equal share of the points in the parameter estimation.
Let us define a system of equations involving each point of each profile. The experimental phase value $Y\left(k_{p}, p\right)$ at the position $x\left(k_{p}, p\right)$ corresponding to the $k_{p}$ th point of the $p$ th profile ( $1 \leq k_{p} \leq n_{p}$ ) should be the


Fig. 6. Phase reconstructions of a wafer of quartz microlenses recorded in transmission with a $10 \times \mathrm{MO}$ (lens diameter, $230 \mu \mathrm{~m}$ ). Panels (a)-(c) illustrate a multiprofile PRP adjustment procedure. The number of extracted profiles (white lines) increases from (a) two profiles to (c) nine profiles. (d) The perspective view of the 2D unwrap of (c). The black lines on (c) indicate the position of the profile used for the shape-compensation procedure. (e), (f) The perspective views obtained by adjusting the second-order shapecompensation parameters on the centered and lower-right microlenses, respectively.
solution of the following linear equation:

$$
\begin{equation*}
Y\left(k_{p}, p\right)=a_{0}(p)+a_{1} x\left(k_{p}, p\right)+a_{2} x^{2}\left(k_{p}, p\right), \tag{31}
\end{equation*}
$$

where $a_{0}(p)$ are the respective phase offsets of the $p$ th profile, as presented before, and $a_{1}, a_{2}$ are the unknown coefficients of the second-order polynomial to be determined. Equation (31) defines therefore an
overdetermined linear system with $N_{h}+2$ unknown factors for $N_{\text {eq }}$ equations ( $N_{\text {eq }}=\sum_{p=1}^{N_{h}} n_{p}$ ) that can easily be solved by computing in the least-squares sense the solution of

$$
\begin{equation*}
\boldsymbol{X} \times \boldsymbol{A}=\boldsymbol{Y}, \tag{32}
\end{equation*}
$$

where $\boldsymbol{A}$ is the vector of the $N_{h}+2$ coefficients,

$$
\boldsymbol{A}=\left[\begin{array}{lllll}
a_{0}(1) & \cdots & a_{0}(p) & \cdots & a_{0}\left(N_{h}\right) \tag{33}
\end{array} a_{1} a_{2}\right],
$$

$\boldsymbol{Y}$ is the vector of the $N_{\text {eq }}$ phase values,

$$
\begin{align*}
\boldsymbol{Y}= & {\left[Y(1,1) \ldots Y\left(n_{1}, 1\right) \cdots Y(1, p) \ldots\right.} \\
& \left.\ldots Y\left(n_{p}, p\right) \cdots Y\left(1, N_{h}\right) \ldots Y\left(n_{N_{h}}, N_{h}\right)\right], \tag{34}
\end{align*}
$$

and $\boldsymbol{X}$ is the matrix
$\boldsymbol{X}=$

$$
\left[\begin{array}{ccccccc}
1 & 0 & 0 & \cdots & 0 & x(1,1) & x^{2}(1,1)  \tag{35}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
1 & 0 & 0 & \cdots & 0 & x\left(n_{1}, 1\right) & x^{2}\left(n_{1}, 1\right) \\
0 & 1 & 0 & \cdots & 0 & x(1,2) & x^{2}(1,2) \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
0 & 1 & 0 & \cdots & 0 & x\left(n_{2}, 2\right) & x^{2}\left(n_{2}, 2\right) \\
& & \vdots & & & \vdots & \vdots \\
0 & \cdots & 1 & 0 & 0 & x(1, p) & x^{2}(1, p) \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 1 & 0 & 0 & x\left(n_{p}, p\right) & x^{2}\left(n_{p}, p\right) \\
& & \vdots & & & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 1 & x\left(1, N_{h}\right) & x^{2}\left(1, N_{h}\right) \\
\vdots & & & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 1 & x\left(n_{N_{h}}, N_{h}\right) & x^{2}\left(n_{N_{h}}, N_{h}\right)
\end{array}\right] .
$$

Whatever least-squares linear-fit method might be used to compute $\boldsymbol{A}$ and therefore to extract the values of $a_{1}$ and $a_{2}$. The iterative procedure explained for the standard method [see Eqs. (27) and (28)] can be applied to the multiprofile procedure. As presented by Fig. 6, the accuracy of the PRP adjustment increases with the number of profiles. A 2D unwrap of Fig. 6(c) provides a perspective-view representation of the wafer's lenses [Fig. 6(d)].

As shown in Eq. (22), the parameter $P_{0,0}$ is a phase offset that permits us to compensate for the variation of $\phi^{\prime}$ originating from the phase delay $\phi(t)$ between the object and the reference waves and the phase shift introduced by the adjustment of $k_{x}$ and $k_{y}$ [see Eq. (10)]. This parameter is the last one to be adjusted during aberration compensation. $P_{0,0}$ can be defined by taking one of the $a_{0}$ of $\boldsymbol{A}$ [Eq. (36)] that are all equal when the PRPs are adjusted.

## 5. Higher-Order and Crossed-Term Corrections

As described above, the procedure restricted to the second order corrects the so-called tilt and defocusing aberrations. Astigmatism, which is one of the most common aberrations, is already partially corrected, because the wavefront curvature is compensated for differently along the horizontal and vertical directions. But the method fails if the principal axes defining the astigmatism are not along these directions. Furthermore, other standard aberrations, such as coma, spherical, or higher-order aberrations, cannot be properly addressed by a second-order correction.

A first approach for higher-order corrections consists simply of increasing the polynomial order of the fit along the horizontal and the vertical directions and of applying the same procedure explained above to compute the PRPs. In the case of a single-profile procedure, Eq. (25) becomes

$$
\begin{align*}
& Y_{h}=a_{0}+a_{1} x+a_{2} x^{2}+\cdots+a_{\alpha} x^{\alpha}+\cdots+a_{H} x^{H}, \\
& Y_{v}=b_{0}+b_{1} x+b_{2} x^{2}+\cdots+b_{\beta} x^{\beta}+\cdots+b_{V} x^{V}, \tag{36}
\end{align*}
$$

where $H$ and $V$ are the polynomial order for the horizontal and the vertical directions, respectively. The computed PRPs are therefore $P_{\alpha, 0}=a_{\alpha}$ and $P_{0, \beta}=$ $b_{\beta}(\alpha \leq H, \beta \leq V)$.

In the case of a multiprofile procedure, increasing the polynomial order entails keeping the same vector $\boldsymbol{Y}$ of the measured phase value [see Eq. (34)] and writing the vector of the $\left(N_{h}+H\right)$ unknown factors as

$$
\begin{equation*}
\boldsymbol{A}=\left[a_{0}(1) \cdots a_{0}(p) \cdots a_{0}\left(N_{h}\right) a_{1} a_{2} \cdots a_{H}\right] . \tag{37}
\end{equation*}
$$

The matrix $\boldsymbol{X}$ is constructed as in Eq. (35) by increasing its order. Using the same general least-squares linear fit allows for computing the PRPs $P_{\alpha, 0}$ and $P_{0, \beta}(\alpha \leq H, \beta \leq V)$.

However, according to the general formulation of the DPM [Eq. (28)], this procedure is effective only for the computation of $P_{\alpha, 0}$ and $P_{0, \beta}$ terms. For the crossterms correction, it is necessary to extract other profiles in the image.

## A. $P_{1,1}$ Computation

Let us consider first the second-order crossed-term parameter $P_{1,1}$, which can be evaluated by defining two profiles of slopes $\pm 1$ along the diagonals of the image: $n=m+\tau(y=x+$ cst $)$ and $n=-m+v(y$ $=x-\mathrm{cst})$. The second-order phase component of the DPM along the first profile is

$$
\begin{align*}
\phi_{+1}(m)= & P_{1,0} m+P_{0,1}(m+\tau)+P_{2,0} m^{2}+P_{0,2}(m+\tau)^{2} \\
& +P_{1,1} m(m+\tau) \\
= & \left(P_{2,0}+P_{0,2}+P_{1,1}\right) m^{2}+\left(P_{1,0}+P_{0,1}\right. \\
& \left.+2 P_{0,2} \tau+P_{1,1} P_{0,2}\right) m+\left(P_{0,1} \tau+P_{0,2} \tau^{2}\right) \\
= & a_{2}{ }^{(+1)} m^{2}+a_{1}{ }^{(+1)} m+a_{0}{ }^{(+1)} . \tag{38}
\end{align*}
$$

Similarly, along the other profile we have

$$
\begin{align*}
\phi_{-1}(m) & =\left(P_{2,0}+P_{0,2}-P_{1,1}\right) m^{2}+\cdots \\
& =a_{2}{ }^{(-1)} m^{2}+a_{1}{ }^{(-1)} m+a_{0}{ }^{(-1)}, \tag{39}
\end{align*}
$$

which yields

$$
\begin{equation*}
P_{1,1}=\left(a_{2}{ }^{(+1)}-a_{2}^{(-1)}\right) / 2, \tag{40}
\end{equation*}
$$

where $a_{2}^{(+1)}$ and $a_{2}^{(-1)}$ are the second-order coefficients of the polynomial fit along the diagonal profiles $y=x+\operatorname{cst}$ and $y=-x+$ cst, respectively. Henceforth we note $a_{i}^{(+j)}$ as the $i$ th-order coefficient of the polynomial fit along a profile of slope $j: y=j x+$ cst.
B. $\quad P_{1,2}$ and $P_{2,1}$ Computations

To compute the third-order crossed terms $P_{1,2}$ and $P_{2,1}$, we need to consider two additional profiles at $n$ $=2 m+\tau(y=2 x+$ cst $)$ and $n=-2 m+v(y$ $=-2 x+\mathrm{cst})$. Along the two profiles of slopes $\pm 1$, the relation to the third order holds:

$$
\begin{align*}
\phi_{+1}(m) & =\left(P_{3,0}+P_{0,3}+P_{1,2}+P_{2,1}\right) m^{3}+\cdots \\
& =a_{3}{ }^{(+1)} m^{3}+\cdots, \\
\phi_{-1}(m) & =\left(P_{3,0}-P_{0,3}+P_{1,2}-P_{2,1} m^{3}+\cdots\right. \\
& =a_{3}{ }^{(-1)} m^{3}+\cdots . \tag{41}
\end{align*}
$$

For the profiles of slopes $\pm 2$, we have

$$
\begin{align*}
\phi_{+2}(m) & =\left(P_{3,0}+8 P_{0,3}+4 P_{1,2}+2 P_{2,1}\right) m^{3}+\cdots \\
& =a_{3}{ }^{(+2)} m^{3}+\cdots, \\
\phi_{-2}(m) & =\left(P_{3,0}-8 P_{0,3}+4 P_{1,2}-2 P_{2,1}\right) m^{3}+\cdots \\
& =a_{3}{ }^{(-2)} m^{3}+\cdots . \tag{42}
\end{align*}
$$

Eliminating the terms $P_{3,0}$ and $P_{0,3}$ in these equations yields

$$
\begin{align*}
& P_{1,2}=\left(a_{3}{ }^{(+2)}+a_{3}{ }^{(-2)}-a_{3}{ }^{(+1)}-a_{3}^{(-1)}\right) / 6, \\
& P_{2,1}=\frac{2}{3}\left(a_{3}^{(+1)}-a_{3}{ }^{(-1)}-\frac{a_{3}^{(+2)}-a_{3}^{(-2)}}{8}\right) . \tag{43}
\end{align*}
$$

By analogy, the definition of other profiles and the solution of the corresponding equation systems allows the computation of higher-order crossed terms.

## C. Procedure Validation

To illustrate the procedure for higher-order and crossed-term corrections, we studied theoretical and experimental aberrated phase distributions. First, a polynomial phase distribution of the fourth order,


Fig. 7. Phase reconstructions for increasing orders of a polynomial phase mask. The computed coefficients are (a) $P_{0,0}$; (b) adding $P_{1,0}, P_{0,1}, P_{2,0}$, and $P_{0,2}$; (c) adding $P_{1,1}$; (d) adding $P_{\alpha, \beta}$ with $\alpha+\beta$ $=3$; and (e) adding $P_{4,0}$ and $P_{0,4}$. Left column, compensation for a computed phase distribution; right column, phase distributions reconstructed from a hologram recorded with a strongly aberrated microscope. The accuracy of the aberration compensation is measured by the standard deviation (Std) of the phase distribution evaluated over the entire field of view.

$$
\begin{equation*}
\phi(m, n)=\sum_{\alpha+\beta=1}^{\alpha+\beta=4} C_{\alpha, \beta} m^{\alpha} n^{\beta} \tag{44}
\end{equation*}
$$

was computed with $-N / 2 \leq m, n<N / 2(N=220)$ and with the coefficients $C_{0,0}=0, C_{1,0}=1, C_{0,1}=-1.1$, $C_{2,0}=-1 e-2, C_{0,2}=-1.1 e-2, C_{3,0}=-1 e-7, C_{0,3}$ $=1.1 e-7, C_{4,0}=-1 e-9, C_{0,4}=1.1 e-9, C_{1,1}=-1 e$ $-4, C_{1,2}=1 e-7, C_{2,1}=1.1 e-7, C_{1,3}=C_{3,1}$ $=C_{2,2}=0$. This computed phase distribution was then wrapped between $-\pi$ and $\pi$ as shown in the left image of Fig. 7(a). The white lines are the selected profiles defined to compute the cross terms.
Second, a strongly aberrated experimental configuration was realized by introducing a tilted thick plate between the beam splitter and the CCD camera in the reflection setup (see Fig. 1). The analytical expressions for the corresponding aberrations are described in Ref. 13. A hologram of a mirror was re-
corded with a $10 \times$ MO. The right image of Fig. 7(a) presents the phase reconstruction of this hologram without aberration compensation.
Figure 7 presents the computed phase reconstructions for the theoretical specimen (left column) and for the experimental specimen (right column) for increasing orders of correction from top to bottom. The offset parameter $P_{0,0}$ is adjusted to have a mean phase value on the entire image that is equal to zero. The efficiency of the correction is evaluated by measuring the standard deviation (Std) of the phase distribution over the entire field of view. As can be seen, the Std decreases rapidly, but the perfect correction for the theoretical specimen $(\operatorname{Std}=0)$ is achieved only when the polynomial order of the DPM is equal to the polynomial order of the generated phase distribution, in which the procedure computes $P_{\alpha \beta}=C_{\alpha \beta}$ for all PRPs.

In the experimental case, the aberrations are compensated for up to a Std of $5^{\circ}$, as shown in the right image of Fig. 7(e). In this case the procedure computes the following coefficients values: $P_{1,0}=3.116$, $P_{0,1}=-3.151, P_{2,0}=-1.056 e-2, P_{0,2}=-1.065 e-$ $2, P_{3,0}=-3.437 e-7, P_{0,3}=3.358 e-7, P_{4,0}=$ $-1.026 e-9, P_{0,4}=1.190 e-9, P_{1,1}=-1.350 e-4$, $P_{1,2}=-4.587 e-7, P_{2,1}=3.407 e-7$. The residual noise may have several origins, such as uncorrected aberrations; surface defects of the mirror or of optical components of the microscope; and spatial phase fluctuations caused by dust, parasitic reflections, and diffraction effects.
A mirror is an ideal sample because long profiles can be taken to evaluate the PRPs. As presented in Fig. 8, coefficients evaluated with a mirror can be used as calibrated values to correct high-order and crossed-term aberrations for further uses of the microscope. Figure 8 presents the phase reconstructions, in 3D perspective views, obtained with different PRPs from a hologram recorder with a U.S. Air Force 1950 resolution test target. Figure 8(a) presents the reconstructed phase by computing the four PRPs $P_{j, 0}$ and $P_{0, j},(j=1$ to 2 , second-order correction) from flat areas on the sample. As for the mirror, we see that these PRPs do not compensate for the aberrations due to the tilted plate. Figure 8(b) presents the reconstructed phase computed with the PRPs, up to order four, obtained previously with the mirror. Phase jumps appear because the tilt has changed between the two samples. Figure 8(c) presents the result after application of a first-order procedure (tilt compensation) on the phase distribution of Fig. 8(b).

## 6. Phase Reconstruction with Specimen Shape Compensation

The principle of the method for compensating the shape of a specimen is identical to the aberrationcompensation method. A second polynomial phase mask $\Gamma_{\text {SCM }}$, called a shape-compensation mask (SCM), is introduced in Eq. (23) to produce a shapecompensated wavefront:

$$
\begin{equation*}
\Psi_{\mathrm{SCM}}(m, n)=\Gamma_{\mathrm{SCM}}(m, n) \Psi(m, n), \tag{45}
\end{equation*}
$$



Fig. 8. Phase reconstructions in perspective view of a U.S. Air Force test target obtained (a) by computing a second-order phase mask with the method presented in Fig. 5, (b) by using the PRPs (cross and noncross terms up to order four) calibrated with a mirror [see Fig. 7(e)], and (c) by adding a first-order (tilt-compensation) procedure to the phase distribution of Fig. 8(b).
where the SCM is defined as

$$
\begin{equation*}
\Gamma_{\mathrm{SCM}}(m, n)=\exp \left(-i \sum_{\alpha=0}^{H \text { SCM }} \sum_{\beta=0}^{V_{\mathrm{SCM}}} P_{\alpha, \beta}{ }^{\mathrm{SCM}} m^{\alpha} n^{\beta}\right), \tag{46}
\end{equation*}
$$

where $P_{\alpha, \beta}{ }^{\text {SCM }}$ defines a new set of coefficients called shape-compensation parameters (SCPs). Here also the accuracy of the compensation can be adapted by changing the polynomial orders $H_{\text {SCM }}$ and $V_{\text {SCM }}$ in the horizontal and vertical directions. The resulting polynomial order of the SCM is defined as $H_{\text {SCM }}+V_{\text {SCM }}$. The SCPs can be adjusted automatically by a curvefitting procedure that is similar to the procedure applied to adjust the PRPs. There is, however, an important difference between the standard topographic imaging mode and the shape-compensation imaging mode, which concerns the position of the


Fig. 9. (a) Phase reconstructions of the vertex of a metallic sphere, (b) 2D unwrapped phase image, (c) digitally flattened phase image revealing surface roughness.
profiles from which phase data are extracted for parameter evaluation using the curve-fitting procedure. For the standard PRP adjustment procedure, the profiles are positioned in areas of the field of view where specimen contributions are constant. For shape compensation, phase data are extracted along profiles defined on the specimen itself. The consequence of this is that the fitting procedures provide coefficient values that include the contribution of the specimen shape, such as its curvature.

Examples of standard topographic images are presented in Figs. 6(c) and 9(a) for the microlens wafer and the metallic sphere, respectively. The real specimen shapes, without $2 \pi$ ambiguities, are shown in

Figs. 6(d) and 9(b), which present the corresponding phase distributions after application of a standard phase-unwrapping algorithm. For Fig. 9(a), the PRP adjustment was performed by a calibration procedure using a mirror as a specimen before application to the metallic sphere. Indeed, this sample does not comprise flat surfaces, enabling a proper estimation of the PRPs.
Figures 6(e), 6(f), and 9(c) present perspective views of phase images obtained by the shapecompensation procedures. Figures 6(e) and 6(f) were obtained for profile positions [see black lines in Fig. 6(c)] centered on two different microlenses and with the second-order SCM:

$$
\begin{align*}
\Gamma_{\mathrm{SCM}}(m, n)= & \exp \left[-i\left(P_{0,0}{ }^{\mathrm{SCM}}+P_{1,0}{ }^{\mathrm{SCM}} m+P_{0,1}{ }^{\mathrm{SCM}} n\right.\right. \\
& \left.\left.+P_{2,0}{ }^{\mathrm{SCM}} m^{2}+P_{0,2}^{\mathrm{SCM}_{n}}\right)\right] . \tag{47}
\end{align*}
$$

As can be seen, the procedure only partially compensates for the lens shape. A ring pattern appears that results from the difference between the expected spherical shape of the lens and the parabolic model used for calculating the SCM. If desired, the shape compensation can be improved by increasing the polynomial order of the SCM as shown in Fig. 10. Figures 10(a) and 10(b) were obtained, respectively, with third- and fourth-order SCMs computed on the centered lens in Fig. 6(c). The profiles defined by the arrows in Figs. 10(a) and 10(b) and plotted in Fig. 10(c) show that a fourth-order SCM is necessary to compensate for the ring effect. Once obtained, the flattened representation of the lens can be used to evaluate the surface roughness and may also be used to observe fine defects, such as scratches, that would not be apparent on the standard topographic mode owing to their small size with respect to the global lens shape. However, the shape compensation is not yet perfect, as shown in particular near the edges of the microlens in Fig. 10(b). This defect could result from a form defect of the lens or from an inadequate mathematical model used to compensate for the shape of the microlens. However, the properly flattened part of the lens is sufficient to provide reliable measurement over its working surface. The use of more sophisticated mathematical models for the definition of the SCM was also investigated to evaluate the exact influence of the lens shape near its border.
Figure 9(c) presents the flattened representation of the metallic sphere vertex obtained by application of a third-order SCM, with profiles positioned on the center of the field of view. Figure 9(c) illustrates an interesting feature of the method. As can be seen, such a flattened representation of the specimen surface provides straightforward access to the surface state, independently of the specimen shape, and reveals fine surface textures or topographical defects. For instance, parameters such as the mean roughness, in this case $R_{a}=25.0 \mathrm{~nm}$, can be directly evaluated as for a flat specimen.


Fig. 10. Perspective views computed by compensating for the curvature induced by the lens with (a) a third-order SCM and (b) a fourth-order SCM. (c) Phase profiles defined along the directions of the white arrows in (a) and (b).

## 7. Conclusion

In this paper we have shown that a phase mask defined by using a polynomial function enables us to compensate digitally for wavefront deformations in DHM, even in the presence of strong, high-order aberrations. The proposed method uses the hologram itself to evaluate automatically and accurately the values of parameters involved by the DPM by using curve-fitting procedures applied to phase data extracted along profiles defined in the region of the field of view, where specimen contributions are constant. It is also demonstrated that parameter values associated with high-order terms of the DPM can be calibrated by using a flat reference sample.
These results illustrate an interesting feature of digital holographic techniques, which offer unique possibilities for the digital processing of wavefronts. In particular, the possibility of performing accurate phase measurements by using optical systems with strong aberrations offers attractive opportunities for the development of cost-effective solutions dedicated to metrology applications that require interferometric resolutions.
Finally, we demonstrate that the proposed method can be applied to compensate for the shape of a specimen. It allows for imaging and evaluating the surface quality of a specimen with a curved surface. The
procedure has been applied to characterize a microlens array and the vertex of a metallic sphere. These applications show that DHM and the shapecompensation procedure allow us to obtain a flattened representation of curved samples that provides straightforward access to the surface state of the specimen, independently of the specimen's shape.
This research was funded through research grants 2153-067068.01 and 205320-103885/1 from the Swiss National Science Foundation and through research grant 6606 from the Innovation Promotion Agency (KTI/CTI).

## References

1. E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," Opt. Lett. 24, 291-293 (1999).
2. I. Yamaguchi, J. Kato, and H. Matsuzaki, "Measurement of surface shape and deformation by phase-shifting image digital holography," Opt. Eng. 42, 1267-1271 (2003).
3. U. Schnars, "Direct phase determination in hologram interferometry with use of digitally recorded holograms," J. Opt. Soc. Am. A 11, 2011-2015 (1994).
4. E. Cuche, P. Marquet, C. Depeursinge, "Simultaneous amplitude and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," Appl. Opt. 38, 6994-7001 (1999).
5. E. Cuche and C. Depeursinge, "Method for simultaneous amplitude and quantitative phase contrast imaging by adjusting
reconstruction parameters for definition of digital replica of reference wave and aberration parameters correction digitally," Patent WO200020929-A (13 April 2000).
6. P. Ferraro, S. De Nicola, A Finizio, G. Coppola, S. Grilli, C. Magro, and G. Pierattini, "Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging," Appl. Opt. 42, 19381946 (2003).
7. P. Massatsch, E. Cuche, and C. Depeursinge, "Low coherence digital holographic tomography," Novel Opt. Instrum. Biomed. Appl. 5143, 18-21 (2003).
8. T. Colomb, P. Dahlgren, D. Beghuin, E. Cuche, P. Marquet, and C. Depeursinge, "Polarization imaging by use of digital holography," Appl. Opt. 41, 27-37 (2002).
9. T. Colomb, E. Cuche, F. Montfort, P. Marquet, and C. Depeursinge, "Jones vector imaging by use of digital holography: simulation and experimentation," Opt. Commun. 231, 137-147 (2004).
10. M. Liebling, "On Fresnelets, interference fringes, and digital holography," Ph.D. thesis (Swiss Federal Institute of Technology, Lausanne, Switzerland, 2004).
11. E. Cuche, P. Marquet, and C. Depeursinge, "Aperture apodization using cubic spline interpolation: application in digital holographic microscopy," Opt. Commun. 182, 59-69 (2000).
12. E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," Appl. Opt. 39, 4070-4075 (2000).
13. J. Braat, "Analytical expressions for the wave-front aberration coefficients of a tilted plane-parallel plate," Appl. Opt. 36, 8459-8467 (1997).

[^0]:    T. Colomb (tristan.colomb@a3.epfl.ch), F. Charrière, J. Kühn, F. Montfort, and C. Depeursinge are with the Ecole Polytechnique Fédérale de Lausanne (EPFL), Imaging and Applied Optics Institute, CH-1015 Lausanne, Switzerland. E. Cuche and N. Aspert are with Lyncée Tec SA, PSE-A, CH-1015 Lausanne, Switzerland. P. Marquet is with the Centre de Neurosciences Psychiatriques, Département de Psychiatrie, DP-CHUV, Site de Cery, 1008 PrillyLausanne, Switzerland.

    Received 16 May 2005; revised 11 August 2005; accepted 17 August 2005.

    0003-6935/06/050851-13\$15.00/0
    © 2006 Optical Society of America

