Virtual optics and sensing of the retrieved complex field in the back focal plane using a constrained defocus algorithm

MENGQI SHEN,[1] 🅾 **TERRY W. K. CHOW,**[2] **HONG SHEN,**[2] **AND MICHAEL G. SOMEKH**[1,3,*]

[1]*Nanophotonics Research Center, Shenzhen Key Laboratory of Micro-Scale Optical Information Technology & Institute of Microscale Optoelectronics, Shenzhen University, Shenzhen, China*
[2]*Department of Electronics and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China*
[3]*Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, UK*
[*]*mike.somekh@szu.edu.cn*

Abstract: The reflected back focal plane from a microscope objective is known to provide excellent information of material properties and can be used to analyze the generation of surface plasmons and surface waves in a localized region. Most analysis has concentrated on direct measurement of the reflected intensity in the back focal plane. By accessing the phase information, we show that examination in the back focal plane becomes considerably more powerful allowing the reconstructed field to be filtered, propagated and analyzed in different domains. Moreover, the phase often gives a superior measurement that is far easier to use in the assessment of the sample, an example of such cases is examined in the present paper. We discuss how the modified defocus phase retrieval algorithm has the potential for real time measurements with parallel image acquisition since only three images are needed for reliable retrieval of arbitrary distributions.

1. Introduction

Phase retrieval is becoming increasingly important in computational optics and imaging. The advent of more powerful computing and equally importantly convenient ways to detect, quantify and record multiple images has encouraged the development of new and effective algorithms.

A milestone in the development of phase retrieval was the Gerchberg-Saxton algorithm [1] involving multiple iterations between image and Fourier planes with constraints from the known intensity information at each plane applied at each iteration. One significant point about the method is that it can be proved that the residual errors cannot increase at each iteration. This sounds reassuring, however, this technique is highly reliant on the accuracy of the measured intensity and the initial guess and it can easily become stuck into local minimum [2], so that it often stagnates to an unsatisfactory solution. Developments of this method pioneered, by amongst others, Fienup [3] provided better convergence, however, this family of methods cannot ensure convergence in all situations.

More recent methods such as ptychography [4] represent an alternative approach, while still involving an iterative algorithm. Here, several overlapping regions of the desired image are projected and measured in a transform plane. The transforms of these sub-images form the data set that is used to reconstruct the phase of the image. An initial estimate of the phase of the transform of the sub-image is made. This is then inverse transformed to give an estimate of the phase in the region of the sub-image. When this is carried out for each sub-image the field in the overlapping regions needs to match, so a correction is made in the phase estimate to improve the coincidence of the overlapping regions. The algorithm proceeds until the estimates in all the sub-image regions match to within a predetermined level [4]. In this way, a phase

value in the whole image is found that is consistent with the intensity from the detected regions. Essentially, phase is the only free parameter that can be changed, and this is adjusted until consistent measurements are obtained. This method has been shown to be very robust with the ability to reconstruct images with high information content. A significant problem with the usual implementation of ptychography is that a large number of sub-images are required to reconstruct the image, this means that even when the issues with alignment of the different regions are addressed, many sub-images are required, which generally precludes application for rapid real time applications.

Transport of intensity [5] is a well-known non-iterative method that has been successfully used for phase retrieval. It relies on the idea that as the beam propagates the evolution of the intensity depends on the phase profile. This leads to an expression that gives the phase in terms of the intensities at different propagation planes. A practical difficulty with the method is that it requires knowledge of the differentials of the field in the propagation direction which may be difficult to evaluate accurately, moreover, there are certain assumptions in the method that may not be satisfied in all practical cases.

In this paper we want to recover the phase of the light reflected to the back focal plane (BFP) of a microscope objective illuminating a sample. This allows the reflectivity over a large range of incident angles and polarizations to be measured in parallel. The system is shown as a simple conceptual picture (omitting the illumination optics) in Fig. 1(a), showing light reflected from a sample to the BFP which is projected onto a transform plane, which is used to retrieve the phase in the BFP. A more complete optical system is shown later in the description of our experiments. Figure 1(b) shows the BFP distribution of an objective lens, each point on the BFP represents a plane wave with certain incident angle θ and azimuthal angle ϕ at the sample, the central orange dot denotes a normal incident plane wave, and positions at the BFP with same radius have the same incident angle. If the polarization on the BFP is linearly polarized, then the polarization state can vary from pure p-polarization (TM) to pure s-polarization (TE) with the change of ϕ. Any direction at an angle, ϕ, to the horizontal gives a combination of the two polarizations.
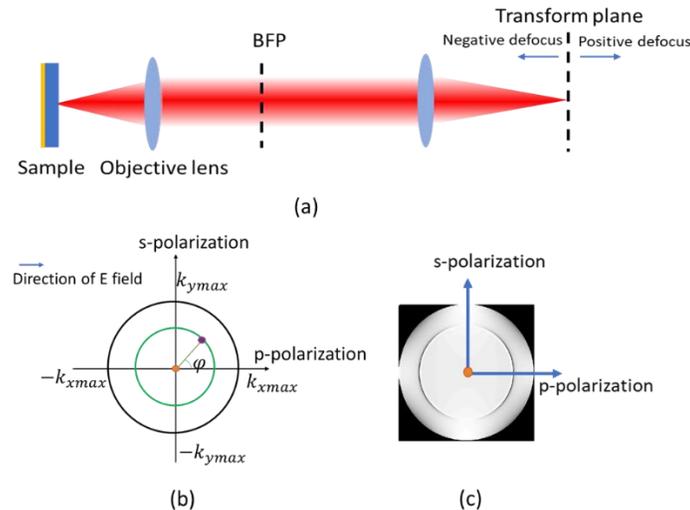
Fig. 1. (a) Simplified schematic of the system showing projection of BFP distribution to a transform plane; (b) relation between different incident angles and polarization states in the BFP distribution of an objective lens; (c) calculated amplitude of BFP distribution reflected from a gold sample in air using a 1.49 NA oil immersion objective lens.

Figure 1(c) shows the calculated BFP amplitude distribution reflected from a 46nm gold film based Kretschmann structure at 633nm wavelength in air.

We will show that phase retrieval measurement in the BFP is a powerful tool for quantitative sample characterization on a microscopic scale that complements phase retrieval applications in imaging. We will also show that acquisition of the phase as well as the amplitude in the BFP allows the application of 'virtual' optics to predict field distributions at different planes with arbitrary masks. This allows one to post-process data to extract different sample parameters in a very convenient manner without the need to repeat the hardware experiment. We will also demonstrate some measurements that are not practical with previous measurement methodologies.

2. Choice and validation of phase retrieval algorithm

Our aim is to retrieve the amplitude and phase in the BFP in order to perform measurements of the sample properties, such as surface wave propagation and film thickness. Previous work has shown that a single intensity BFP image [6] gives sample information over a wide range of incident angles and input polarization states. We will demonstrate, however, that the acquisition of phase information, allows application of virtual optics which greatly enhances the power of the measurements. Our plan is to develop this instrument to make microscopic real time measurements on samples, for example, to monitor the progress of antibody/antigen binding on a surface. A conventional ptychographic algorithm is well suited to the task of phase retrieval, however, the large number of images required and the necessity to move the probe means that such an approach is too slow for real time applications.

Ideally, a method where all the images required for phase retrieval can be acquired in a single shot is needed, such as described in [7], where, effectively, diffraction patterns from different probe positions are projected onto a single camera, however, the problem with this approach was that crosstalk between different probe positions and limited signal to noise degraded the quality of the reconstruction. This is something of a problem in imaging but completely unacceptable when the primary purpose of the system is to perform quantitative measurements. Recent work has shown that satisfactory ptychographic reconstruction can be achieved with as little as four images, this was shown for both conventional ptychography [8] and Fourier ptychography [9]. The results in [8] give adequate results for a diffuser with 4 probe positions. It appears, however, that the method was not successful when less than 4 probe positions were used. It should also be mentioned that successful phase retrieval can be achieved with a coded aperture and single image with a support constraint [10–12], however, such an approach requires an assumption of sparsity which is not generally applicable.

The approach described in Allen and Oxley [13] achieved phase retrieval by examining the field in the transform plane at different defocused values. In this paper they recommended a minimum of three different defocus positions for reliable retrieval. This is convenient for our purposes since different defocuses can be obtained by placing the detection cameras at different planes. In the present work we did this simply by moving a single camera as described in section 3, however, the method lends itself to insertion of three cameras into the system conjugate with slightly different planes to ensure single shot data acquisition. Future generations of our system will incorporate three cameras in this way, which we expect, in addition, to speeding up data acquisition will reduce sensitivity to, for instance, fluctuations in laser power. Moreover, laser power is attenuated with the neutral density filter in the present implementation, so spreading optical power to three cameras does not present any problem.

The algorithm is shown in Fig. 2 below and where we propagate between different planes and impose the measured intensity distribution as a constraint while retaining the computed phase. The function $P_i(k_r)$ is the phase profile corresponding to propagation between the focus and the ith measurement plane. In both the simulations and experiments, we use diffraction images at the transform plane with defocus distance equal to -0.5μm, -1 μm and -1.5 μm, respectively.

The phase distribution at the different defocus distances are shown in Figs. 3(a), (b) and (c), respectively, these profiles are calculated for an objective with numerical aperture of 1.49 and coupling oil of index 1.518. It is important to point out that different defocuses (including positive values) can be used and the values are not critical as long as they are sufficiently well separated. The phase profiles corresponding to these defocuses are shown in Fig. 3(d).
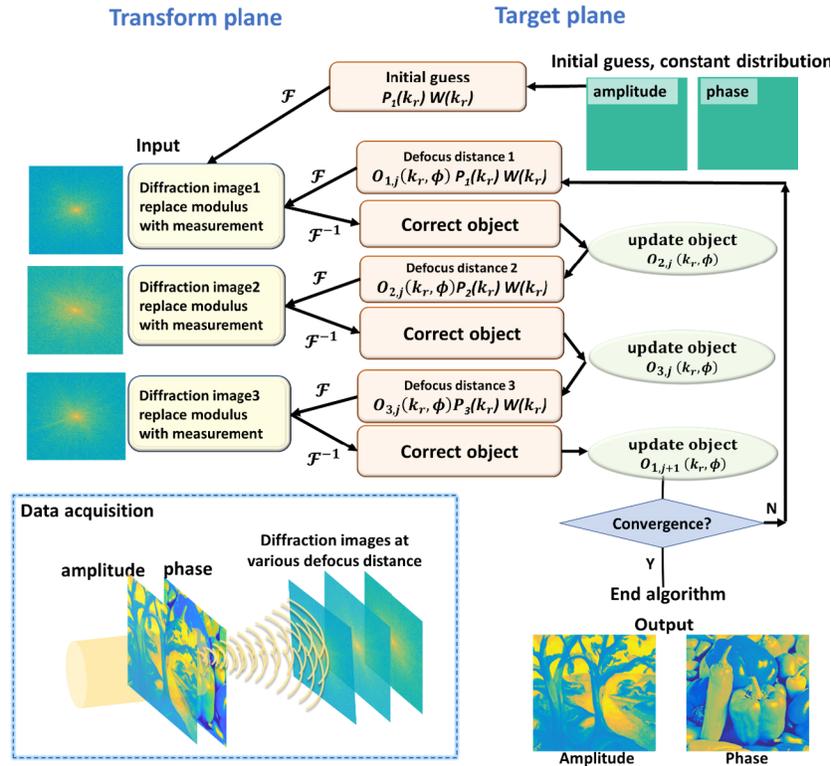
Fig. 2. Schematic of defocus images based three-input algorithm with support constraint.
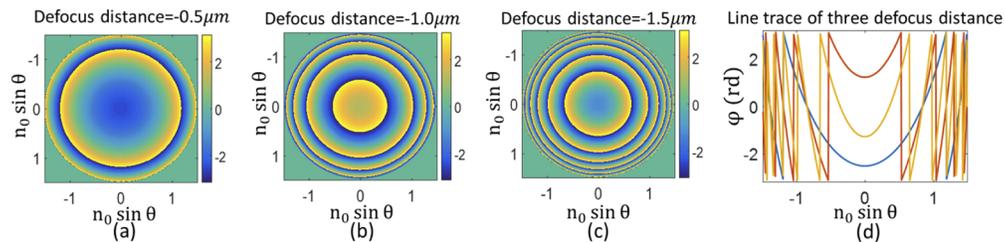
Fig. 3. (a) Phase of probe distribution at defocus distance=-0.5μm; (b) Phase of probe distribution at defocus distance=-1μm; (c) Phase of probe distribution at defocus distance=-1.5μm; (d) comparison of the phase profiles at different defocus distances (blue -0.5μm, yellow -1μm, red-1.5μm). The x-axis on the figures represents the product of the refractive index of the coupling oil (1.518) and the sine of the incident angle.

Figure 4 shows the convergence of the algorithm on simulated data used to validate the method. It can be seen that the algorithm as described in [13] gives modest convergence but the sum

of the squared error (SSE, defined in appendix 1) is rather large meaning that it stagnates well before an optimum solution is reached. Since we are trying to recover the BFP distribution the support of the recovered data is very well defined being determined by the numerical aperture (NA) of the microscope objective shown in Fig. 1. It is thus easy to apply a window function related to the support given by:

$$W(k_r) = \begin{cases} 1, & k_r \in \text{illumination region} \\ 0, & k_r \notin \text{illumination region} \end{cases}. \tag{1}$$

In Fig. 4, we compare the convergence between iterative defocus algorithm with support and without support. We use Figs. 4(a) and (b) as target amplitude and phase distributions. In Fig. 4(c), dashed lines and solid lines represent the convergence without support and with support, respectively. The black, purple and green curves show results from the defocus iterative method employed three inputs, four inputs and five inputs, respectively. It can be seen that with a greater number of inputs from transform plane, the convergence rate is higher and the reconstructed result improves, the SSE of the final reconstructed results are -23dB, -30dB and -40dB, respectively. However, as mentioned before a larger number of inputs is impractical for real time measurements. The solid blue, red and yellow curves show results for the defocus iterative algorithm with precise support to match the BFP, support 5% larger diameter than the true support of the BFP and support with 10% larger diameter, respectively. It can be seen that application of even an approximate estimate of the support leads to rapid convergence with a vanishingly small SSE (-310dB) as shown in Fig. 4. A less precise measure of the support requires slightly more iterations but with even a 10% error in the size of the window convergence is complete as shown in Fig. 4. The defocus algorithm with support constraint has worked on all the image sets tested so far, provided the separation of the input planes was greater than 0.5µm. It is worth discussing the situation that arises when different separations between the input planes are used. For good convergence we find that the mutual (complex) correlation coefficients between the profiles should be close to zero, for the values chosen above this is the case. When separations of only 100 nm were used there is substantial correlation, so poor convergence (SSE of -15 dB) was achieved. For very large differences between the probe planes the correlation values are still close to zero so performance is essentially the same as for the values chosen in this work. It is more convenient experimentally to use more modest defocus values. The use of only two defocus positions (omitting the intermediate defocus of -1µm) gave an SSE of -35.6 dB, confirming the observations in [13] recommending more than two defocus positions. The use of three probe positions gives good robustness with relatively convenient experimental implementation. The pseudo code for the algorithm is shown in Appendix 1.
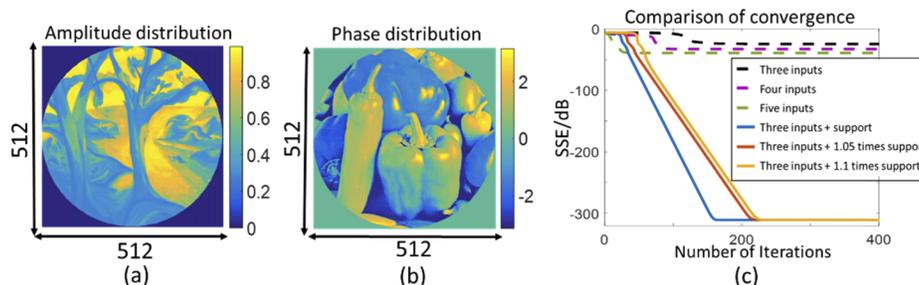
Fig. 4. (a) Amplitude distribution at the target plane; (b) phase distribution at the target plane; (c) comparison of convergence between algorithms.

In addition to the convenience of the defocus based algorithm and the fact that probe scanning is not necessary another useful feature of the algorithm is that the data for the algorithm is recovered at a defocus plane so that the recorded intensity is better balanced, nevertheless, in order to extend the dynamic range, images with two different exposures were used in the experiments.

In view of these results we use the iterative defocus algorithm, which is similar to that described in (Allen & Oxley, 2001) with the addition of the support constraint ($W(k_r)$ as shown in Fig. (2). This makes the convergence far more effective and, $W(k_r)$, is a natural constraint to apply when trying to recover the BFP. It is worth clarifying terminology at this point the BFP we wish to reconstruct is called the 'target' plane and the plane where the measurement takes place is called the 'transform' plane. In these experiments this is conjugate to a defocused image plane.

3. Experimental system and reconstruction results

Figure 5 shows the schematic of the experimental system. The sample was illuminated by linear polarized light with λ=633nm, light information reflected from sample was collected by the objective lens (Nikon, CFI Apochromat TIRF, oil immersion, 100x, NA=1.49). A polarizer was placed in the detection arm to select E_x and E_y components. A camera (Thorlabs, CS2100M,1960×1080, 16bits) was placed at three different negative defocus positions close to the transform plane.
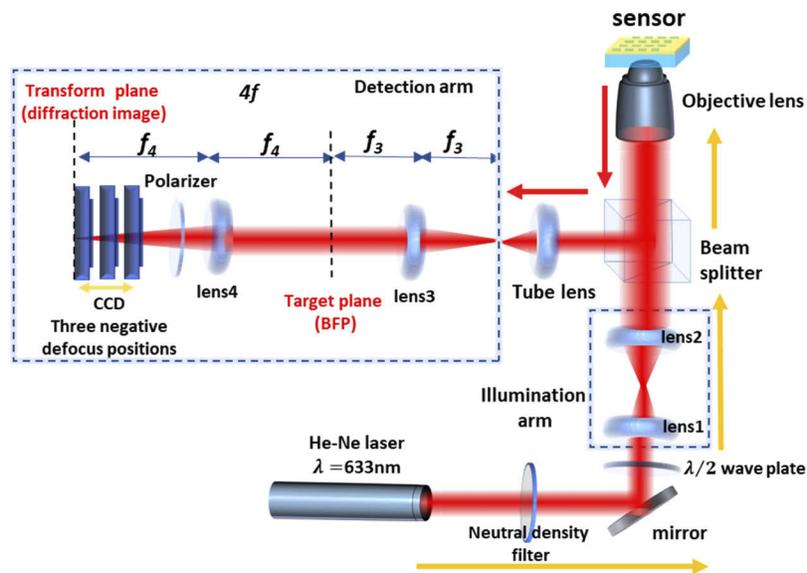
Fig. 5. Schematic of the system used to recover the phase of the back focal plane, showing three different defocus positions of the image plane detector. One additional arm (not shown) was inserted for direct observation of the intensity in the back focal plane. This was not used in any of the reconstructions and simply used for comparison.

3.1. Kretschmann structure based surface plasmon mode reconstruction

We now apply the constrained defocus algorithm to recover the E_x component of the BFP distribution from a gold sample through the high NA oil immersion objective system, as shown in Fig. 6(a).

The thickness of the gold film was determined to be 46nm and the literature value of the gold permittivity was -12.33 + 1.21i at 633nm wavelength [14]. The calculated amplitude and phase distribution from this structure is shown in Fig. 6(b) and (c), respectively. In the experiment, we
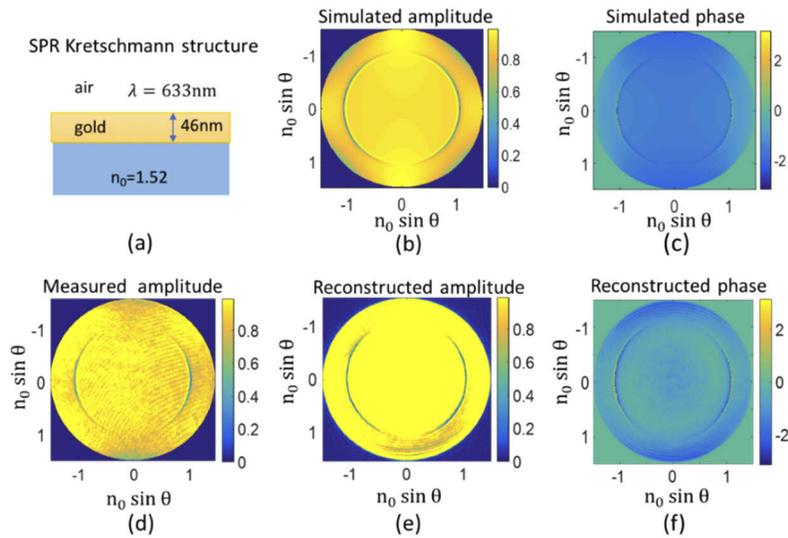
Fig. 6. (a) Structure of sample illuminated through the coverslip; (b) simulated amplitude distribution of BFP; (c) simulated phase distribution of reflected BFP; (d) measured amplitude distribution; (e) reconstructed amplitude distribution of BFP; (f) reconstructed phase distribution of BFP.

used the different wavefronts corresponding to different defocus positions along the optical axis as the 'probe', the camera was placed at three negative defocus positions corresponding to values of -0.5µm, -1µm and -1.5µm respectively at the focus of the object. The axial magnification between the sample plane and the target plane means that the actual displacement of the cameras was -20 mm, -40mm and -60mm respectively. Since one camera was used a linear stage was needed to change the camera position axially, the system is being modified to use three cameras in parallel. Figure 6(e) shows the directly measured amplitude distribution in the experiment for comparison. The recovered amplitude distribution and phase distribution using the algorithm are shown in Figs. 6(e) and (f), respectively. Both the recovered amplitude and phase show good agreement with the calculated data.

We now use the phase of the reconstructed signal to perform two separate tasks (i) as a sensor and (ii) to reconstruct image plane distributions with filtering at arbitrary defocus planes.

3.1.1. Sensing

One of the main applications of surface plasmon resonance (SPR) measurement is to sense small changes in the surface due to analyte binding or refractive index change. There is evidence within the biosensing community that phase measurement confers superior sensitivity compared to amplitude only measurement largely on account of the sharper transition [15–16]. We demonstrate below how the recovered phase may be used to measure the thickness of a deposited layer. Two different thickness layers of Al_2O_3 were deposited on top of gold film of Kretschmann configuration. The structure is shown in Fig. 7(a). Figure 7(b) shows the simulated results of phase transition corresponding to the generation of the SP. The blue, red and yellow curves represent the uncoated gold response, the 5nm Al_2O_3 layer and the 10nm Al_2O_3 layer respectively. The calculated resonance positions units of $n_0 \sin\theta$ are 1.038, 1.058 and 1.08. Figure 7(d) represent the reconstructed phase information. The reconstructed resonance position of blue, red and yellow lines are 1.038, 1.06 and 1.076 units of $n_0 \sin\theta$, respectively. These values are converted to a layer thickness using the refractive index value of 1.766, which corresponds to

Al_2O_3 layer thickness of 5.8nm and 9.8nm. The comparison of thickness of the Al_2O_3 layer obtained from the SPR measurement and the sputtering machine are shown in Fig. 7(c). These values are within the known thickness errors of the sputtering machine (Kurt J. Lesker, PVD75) used to fabricate the layers.
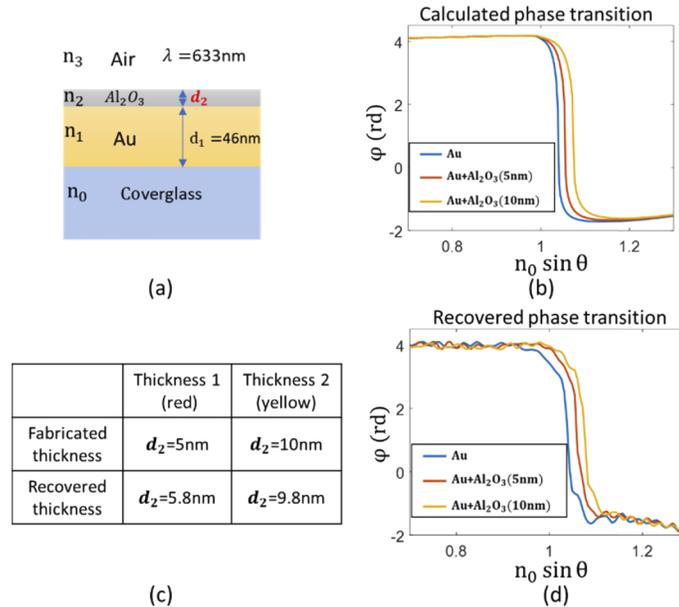
Fig. 7. (a) Structure of sample layer; (b) simulated phase transition; (c) comparison of fabricated layer thickness (obtained from sputterer) and reconstructed layer thickness; (d) reconstructed phase transition.

3.1.2. Virtual optics propagation

The amplitude and phase of the BFP has been constructed from three images in the transform plane, however, with phase retrieval we may now use the recovered complex BFP to generate the field at the image plane at any value of defocus. Moreover, we may apply any virtual mask to generate a desired virtual distribution, this may be used to select different modes generated at different incident angles and polarizations.

These reconstructed distributions demonstrate another important advantage of phase measurement for sensing applications. In a recent paper [17] we showed that examining the leakage radiation in the image plane on a uniform sample allowed one to recover similar information to an intensity BFP distribution. We showed, however, a crucial advantage of image plane measurement, that in the presence of a sample with varying properties it could recover local values of the surface wave k-vector, whereas an intensity BFP measurement would reconstruct some averaged value rather than a localized measurement. Clearly, then acquisition of the BFP phase allows one to transform this complex field to an image plane distribution to replicate direct image plane measurement with far superior localization compared to intensity only BFP measurement.

We now show how the BFP may be converted to an image plane distribution for a sample supporting surface plasmons and will later use a similar approach for a dielectric sample supporting a large number of surface wave modes. Table 1 shows the comparison between calculated field distribution and reconstructed field distribution in virtual optics using the results in Figs. 6(e) and (f).

Table 1. Comparison between calculated data and reconstructed data using various mask at different propagation positions.

The field distributions at defocus values of 4 μm and -4μm are shown, and the scale in the transform plane is in microns. Three different masks were applied: 1) no mask, only the illumination aperture window was applied; 2) a 'phi' mask, where only angles between ±45 degrees are passed through the aperture centered around TM polarization direction to enhance SP propagation; 3) an arc shape mask, with predominantly TM polarization and incident angles associated with SP excitation. In Table 1, the virtual propagation results using phi mask and arc shape mask show excellent agreement with the simulation data. In the first row with no mask applied, agreement is still reasonable but some noise is present in reconstructed results. We can see that for the phi and arc masks at positive defocus there is a great deal of energy propagating away from the axis because phase matching at positive defocus supports propagation away from the axis and negative defocus towards the axis [18].

3.2. E_y-component of the field in the back focal plane

Figure 8 shows the recovered BFP of a gold sample obtained by selecting the electric field along y-direction by rotating the polarizer by 90 degrees. In this case, there is no difference between horizontal direction and vertical directions in the amplitude distribution, and phase distribution shows four quadrants with each quadrant shifted in phase by π rads compared to the adjacent quadrant. Figures 8(a) and (b) are the calculated amplitude and phase, respectively. Figures 8(c), (d) and (e) are the measured amplitude, reconstructed amplitude, and reconstructed phase, respectively. We can see that the recovered BFP is consistent with the simulated result. The center is somewhat noisy because the signal tends to zero along the *x*- and *y*- axes. The results, show, however, the ability to recover signal around singularities.

3.3. Reconstruction of a dielectric multilayer structure

One of the advantages of the application of phase retrieval to measurement is to visualize modes in lossless/transparent structures. In the transparent structure, there is no dip in the reflection spectrum, it is therefore difficult to make good observations of the different modes with an intensity only method since intensity changes rely on interference between TE and TM polarizations. Moreover, other methods such as confocal interferometry [19] are not applicable
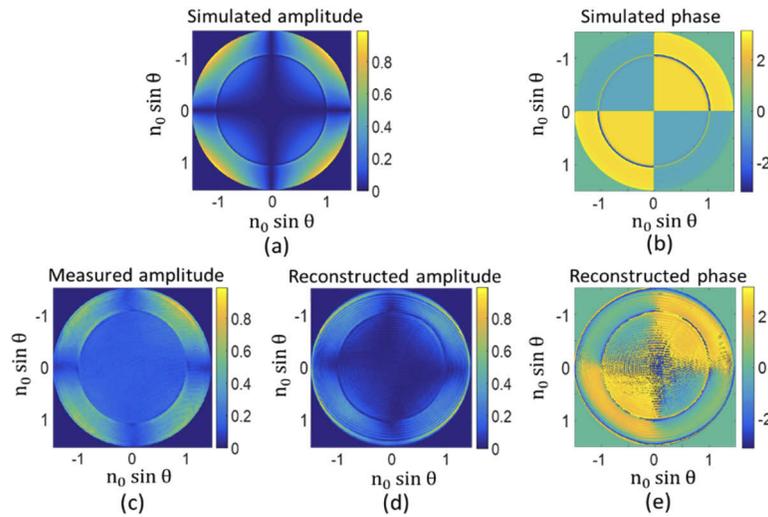
Fig. 8. (a) Simulated amplitude distribution in the BFP with cross-polars; (b) simulated phase distribution in the BFP; (c) directly measured amplitude distribution in the BFP; (d) reconstructed amplitude distribution in the BFP; (e) reconstructed phase distribution in the BFP.

because they require a strong reflection at normal incidence to form the reference beam, and with these transparent samples the normal reflection is very small. Since phase information is related to the mode, as shown in the sensing section, recovering phase transitions provides a direct means to monitor mode information. Figure 9(a) shows the lossless sample structure, fabricated by depositing LiF (Lithium Fluoride) and PMMA (poly(methyl methacrylate)) layer on top of the cover glass. The refractive index of cover glass, LiF and PMMA at 633nm wavelength are 1.52, 1.37 and 1.49, respectively. The thickness of LiF and PMMA are 400nm and 550nm, respectively. Figures 9(b) and (c) are the simulated E_x component of amplitude and phase distributions of structure in Fig. 9(a). Figures 9(d), (e) and (f) are the measured amplitude, recovered amplitude and recovered phase of BFP, respectively. It can be seen that general features are recovered and the phase information matches the simulated data better than the amplitude. This is partly due to the performance of the reconstruction process and also likely to be due to the fact that there are some interface losses in the sample which affects the intensity signal more strongly than the phase.

In order to extract accurate and reliable modal information we are interested in the signals along the pure TE and TM polarization directions (vertical and horizontal). There is, of course, a lot of information at other azimuthal angles, and we use a least squares method which optimizes the signal to noise ratio, this is similar to the one used for intensity BFP [6], the algorithm used for complex data is presented in Appendix 2. The resulting recovered reflections for the two polarizations are shown in Fig. 10. Figures 10(a) and (c) are the amplitude and phase information of TM mode, Figs. 10(b) and (d) are the amplitude and phase line trace of TE mode. It can be seen that the phase information gives clear mode information while the amplitude is ambiguous and can be easily confused with noise. The most immediately noticeable thing with the phase reconstruction is that several modes are clearly indicated, from small excitation angles to large excitation angles, the resonant modes are labelled as TM_1, TM_2, TM_3 in Fig. 10(c) along *p*-polarized direction, and TE_1, TE_2, TE_3 along the *s*-polarization direction in Fig. 10(d). There are only weak hints of modes below about 60-65 degrees in the amplitude maps obtained with either direct measurement or the amplitude obtained from phase reconstruction; these do not
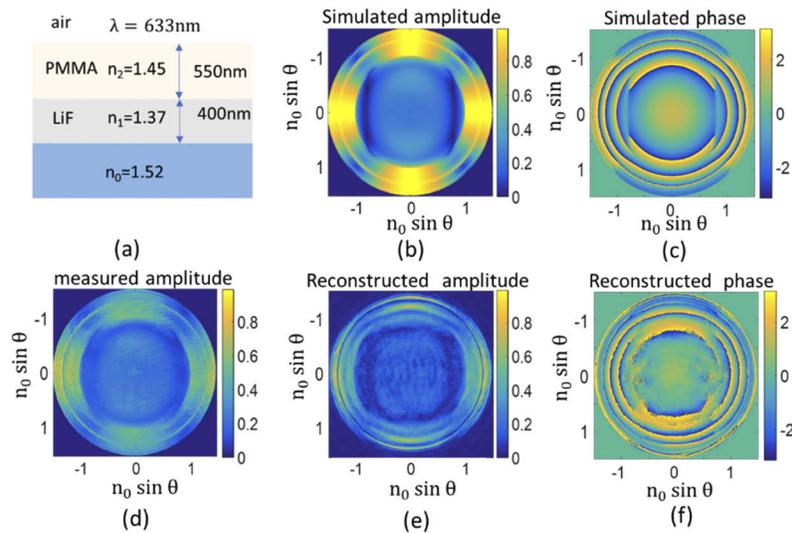
Fig. 9. (a) Structure of lossless guided wave structure; (b) simulated amplitude distribution in the BFP; (c) simulated phase distribution in the BFP; (d) directly measured amplitude distribution; (e) reconstructed amplitude distribution in the BFP; (f) reconstructed phase distribution in the BFP.

allow for reliable measurement of the mode position. These can be identified once the modes are known from the phase data but cannot be used for either positive identification or accurate measurement as they are rather broad and indistinct on account of the large coupling loss.
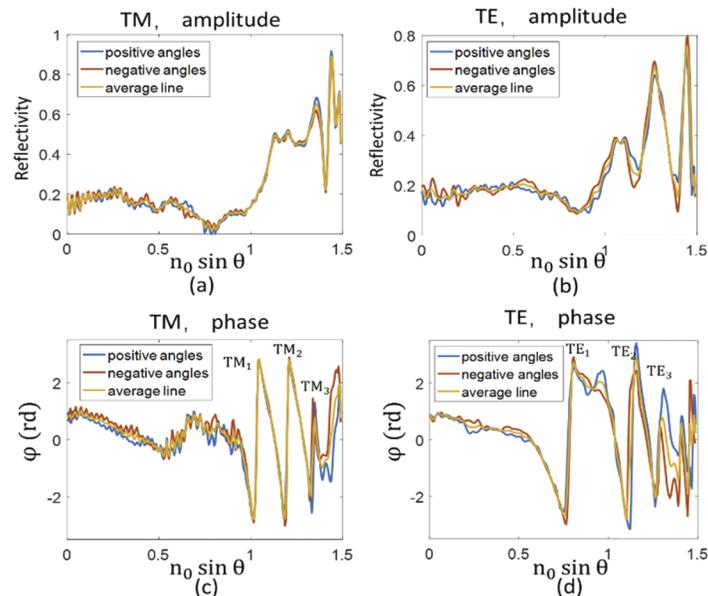
Fig. 10. (a) Reflectivity of TM mode; (b) reflectivity of TE mode; (c) phase information of TM mode; (d) phase information of TE mode.

The phase shift in the reflection (or transmission) coefficient is the true signature of the presence of a surface wave because it represents lateral flow of energy. The dips in the amplitude

only occur in lossy materials and are a manifestation of resonant field enhancement. The phase shifts are present in both lossy (metallic) and lossless (dielectric) structures. The dips that can be observed in the amplitude response in our dielectric structure are thus a manifestation of non-ideality of the structures, primarily local roughness. It is not surprising therefore that this information is rather weak.

Table 2. Comparison of mode positions between measured results and simulated data from nominal sample dimensions.

TM		TE	
Measured ($n_0 \sin \theta$)	Simulated ($n_0 \sin \theta$)	Measured ($n_0 \sin \theta$)	Simulated ($n_0 \sin \theta$)
1.032	0.9992	0.9102	0.8589
1.198	1.186	1.157	1.139
1.326	1.338	1.321	1.309

Table 2 shows the incident angle for the excitation of different wave modes measured from the phase plots of Fig. 10. It is interesting that several modes are identified for TE polarization including one below the critical angle (*c.*41 degrees). The amplitude information confirms the phase measurement but is of no value on making reliable measurements of the modal positions. The mode position difference between the measured results and simulated data can arise from the error of the thickness of each layer or the deviation of the refractive index from the design values. Among all the modes, the one along TE direction below critical angle shows the largest deviation with 5.6% difference. The deviations of the other modes are all less than 3.2%. This is also because the mode at smaller angles are much more sensitive to the change of bulk material, layer thickness and environment than modes with larger angles [6]. These results are all within the tolerances of the sample fabrication and may be used to refine the true sample dimensions.

With the complex field information retrieved form the multimode structure, we can extract the information of each mode separately, and propagate it to any observation position in virtual optics. We specifically choose two defocus positions of $4\mu m$ and $-4\mu m$ for illustration and applying arc masks to select each mode with certain angle and polarization for comparison. We can see from Table 3 that reconstructed results show good agreement with the simulated data.

Table 3. Comparison of TE and TM modes obtained from the measured reconstructed phase, compared with calculated values obtained using the nominal parameters of Fig. 1.

| Defocus distance | $4\mu m$ | | $-4\mu m$ | |
mask	Simulated	Reconstructed	Simulated	Reconstructed
	TM mode			
TM_1				
TM_2				
TM_3				
	TE mode			
TE_1				
TE_2				
TE_3				
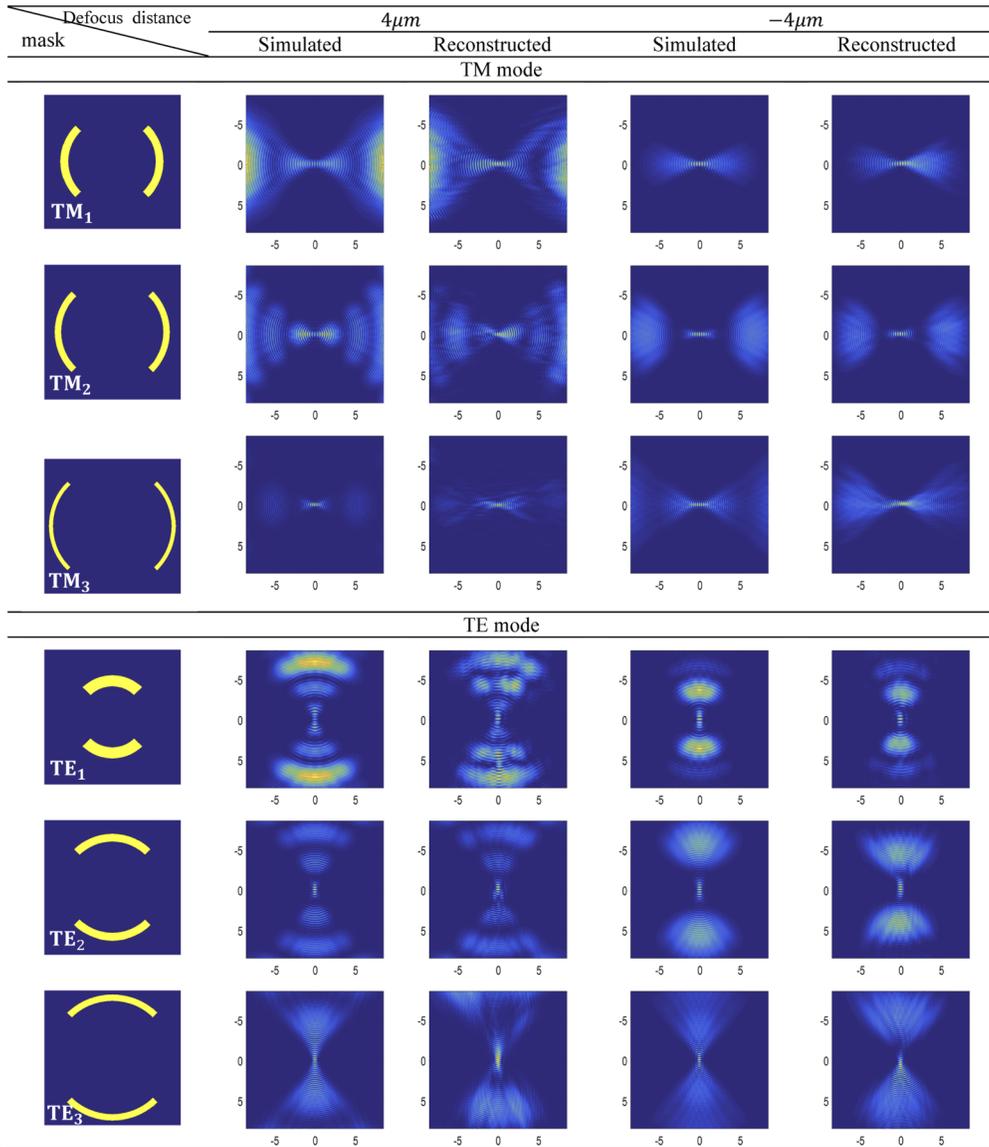
4. Conclusion and discussion

In this paper we have modified an algorithm based on defocus to recover the phase as well as the amplitude of the field in the BFP. The purpose of the algorithm is to develop a stable and robust method that allows reconstruction with the minimum number of input images, so that the system may be adapted for single shot operation, and hence real time monitoring of biological processes. The use of three defocused images and a support constraint from the known extent of the image in the BFP provides extremely reliable convergence on both actual and simulated data.

There are a vast number of measurement possibilities that become available when the complex field of the BFP is acquired and we give a flavor of some of the different possibilities here. We show that the phase allows reconstruction of a virtual image plane at any defocus with considerable advantages in terms of measurement localization as described in [17]. We also show how phase reconstruction combined with a noise reduction algorithm in the BFP is very effective in recovering data on transparent samples where previous techniques were not successful, moreover, it is clear that the phase measurement provides the basis for the mode identification. Separation of the modes with a virtual mask allows one to identify and visualize their propagation independently. Such measurements would be prohibitively difficult without phase retrieval as some modes are very close and their position is not known until the measurements are complete.

Routine recovery of phase information in the BFP provides further opportunities to produce new generations of sensors for high lateral resolution and multi-point measurement.

Appendix 1: Reconstruction algorithm

Input data: Acquire images at three defocus plane and determine a rough estimate of the window function $W(k_r)$.

Computational operations on input data

(1) Initial estimate of the phase and amplitude of the object, $O_{i,j}(k_r, \phi)$, can be random or uniform, in this algorithm, uniform is better. The subscript i is the ith transform image corresponding to the ith defocus position, j represents the jth iteration.

(2) The exit-wave $\psi_{i,j}(k_r, \phi)$ is the product of object $O_{i,j}(k_r, \phi)$, the defocused wave front $P_i(k_r)$ and the illumination window $W(k_r)$.

(3) Fourier transform $\psi_{i,j}(k_r, \phi)$, $\Psi_{i,j}(u) = \text{F}[\psi_{i,j}(k_r, \phi)]$.

(4) Replace the modulus of the resulting computed inverse Fourier transform with the measured diffraction amplitude ($|\Psi_{measured}(u)|$), $\Psi'_{i,j}(u) = |\Psi_{i,measured}(u)|\angle\Psi_{i,j}(u)$.

(5) Inverse Fourier transform $\Psi'_{i,j}(u)$, $\psi'_{i,j}(k_r, \phi) = \text{F}^{-1}[\Psi'_{i,j}(u)]$.

(6) Update function: $O_{i+1,j}k_r, \phi = O_{i,j}k_r, \phi + P_i(k_r)^*(\psi'_{i,j}(k_r, \phi) - \psi_{i,j}(k_r, \phi))$.

(7) Calculate the SSE. If target SSE is reached EXIT, the SSE is defined as:

$$SSE = 10log_{10}\left\{\frac{\sum\{|O_j(k_r, \phi)| - |O(k_r, \phi)|\}^2}{\sum\{|O(k_r, \phi)|\}^2}\right\}.$$

Where $|O(k_r, \phi)|$ is the measured amplitude of the target, which can be as a reference to evaluate the accuracy of the retrieved field. $\sum$ means the summation of points in target plane. For simulated data the SSE is readily available; for experimental implementation, however the algorithm is stopped after a fixed number of iterations or when the intensity correction in stage (4) does not change.

(8) Return to (2) with new defocused position (if appropriate).

Appendix 2. Algorithm to average around azimuth with different weightings

We start with the general expression for the horizontal (x) component of the returning field as calculated in Fig. 11, this component is selected with the polarizer in Fig. 5.
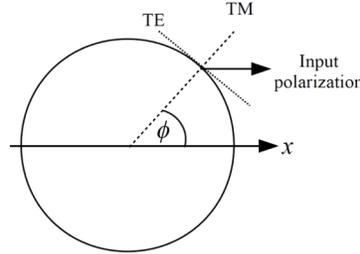
Fig. 11. TM and TE components arising from linear *x* input polarization

For linear polarized light along the *x*-direction, we can see that the component of the field along the TM (*p*-) polarization is weighted by a factor of $\cos\phi$, which is again multiplied by the factor as it is resolved back alon the *x*-component. Similarly, the TE component is weighted by $\sin^2\phi$. We then separate the output along *x* into real and imaginary parts to give two entirely real equations.

$$\mathrm{Re}(E_x) = r'_{p\alpha}\cos^2\phi + r'_{s\alpha}\sin^2\phi. \tag{2}$$

$$\mathrm{Im}(E_x) = r''_{p\alpha}\cos^2\phi + r''_{s\alpha}\sin^2\phi. \tag{3}$$

Where the single superscript denotes the real part and the double subscript the imaginary part of the reflection coefficients. We add the subscript α to account for the fact that the recovered reflectivity values are multiplied by an arbitrary common phase factor that arises from the starting point of the phase retrieval algorithm.

Taking *n* measurements for $\mathrm{Re}(E_x)$ calling each measurement R_n.

We then need to minimize: $\sum_n (R_n - r'_{p\alpha}\cos^2\phi_n - r'_{s\alpha}\sin^2\phi_n)^2$ this is achieved by partially differentiating with respect to $r'_{p\alpha}$ and $r'_{s\alpha}$ and setting each partial differential to zero. This is analogous to the approach used in phase stepping interferometry [20].

This yields a matrix equation as follows:

$$\frac{1}{2}\begin{bmatrix} \sum_n \cos^4\phi_n & \sum_n \cos^2\phi_n\sin^2\phi_n \\ \sum_n \cos^2\phi_n\sin^2\phi_n & \sum_n \sin^4\phi_n \end{bmatrix}\begin{bmatrix} r'_{p\alpha} \\ r'_{s\alpha} \end{bmatrix} = \begin{bmatrix} \sum_n R_n\cos^2\phi_n \\ \sum_n R_n\sin^2\phi_n \end{bmatrix}. \tag{4}$$

Precisely, the same process is used to recover $r''_{p\alpha}$ and $r''_{s\alpha}$, using the data from the imaginary part of E_x. The result of this process is that the recovered values for $r'_{p\alpha}$, $r''_{p\alpha}$, $r'_{s\alpha}$ and $r''_{s\alpha}$ are now effectively averaged over many different azimuthal positions with considerable noise reduction. Figure 8 in the main text shows the effect applying this process which gives a much improved signal to noise ratio compared to a single line. For these experiments 180 azimuthal angles from -90 deg. to +90 deg. were used. The condition number for the 2×2 matrix is 2 which accounts for the excellent noise reduction. The strength of the retrieval algorithm combined with efficient use of redundancy in the BFP data thus allows us to recover stable complex data that would be very difficult to achieve with other methods.

Funding

Disclosures

The authors declare no conflicts of interest.

References

1. R. W. Gerchberg and W. O. Saxton, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," Optik **35**(2), 1–6 (1972).
2. C. Guo, S. Liu, and J. T. Sheridan, "Iterative phase retrieval algorithms. I: Optimization," Appl. Opt. **54**(15), 4698–4708 (2015).
3. J.R. Fienup, "Reconstruction of an object from the modulus of is Fourier transform," Opt. Lett. **3**(1), 27–29 (1978).
4. J. Rodenburg, "Ptychography and Related Diffractive Imaging Methods," Adv. Imaging Electron Phys. **150**, 87–184 (2007).
5. L. Waller, L. Tian, and G. Barbastathis, "Transport of Intensity phase-amplitude imaging with higher order intensity derivatives," Opt. Express **18**(12), 12552–12561 (2010).
6. M. Shen, S. Learkthanakhachon, S. Pechprasarn, Y. Zhang, and M. G. Somekh, "Adjustable microscopic measurement of nanogap waveguide and plasmonic structures," Appl. Opt. **57**(13), 3453–3462 (2018).
7. P. Sidorenko and O. Cohen, "Single-shot ptychography," Optica **3**(1), 9–15 (2016).
8. S. McDermott and A. Maiden, "Near-field ptychographic microscope for quantitative phase imaging," Opt. Express **26**(19), 25471–25480 (2018).
9. J. Sun, C. Zuo, J. Zhang, Y. Fan, and Q. Chen, "High-speed Fourier ptychographic microscopy based on programmable annular illminations," Sci. Rep. **8**(1), 7669 (2018).
10. R. Egami, R. Horisaki, L. Tian, and J. Tanida, "Relaxation of mask design for single-shot phase imaging with a coded aperture," Appl. Opt. **55**(8), 1830–1837 (2016).
11. R. Horisaki, R. Egami, and J. Tanida, "Experimental demonstration of single-shot phase imaging with a coded aperture," Opt. Express **23**(22), 28691–28697 (2015).
12. R. Horisaki, Y. Ogura, M. Aino, and J. Tanida, "Single-shot phase imaging with a coded aperture," Opt. Lett. **39**(22), 6466–6469 (2014).
13. L. Allen and M. Oxley, "Phase retrieval from series of images obtained by defocus variation," Opt. Commun. **199**(1-4), 65–75 (2001).
14. P. B. Johnson and R. W. Christy, "Optical Constants of the Noble Metals," Phys. Rev. B **6**(12), 4370–4379 (1972).
15. A. Kabashin, S. Patskovsky, and A. N. Grigorenko, "Phase and Amplitude Sensitivities in surface plasmon resonance bio and chemical sensing," Opt. Express **17**(23), 21191–21204 (2009).
16. Y. Huang, H. P. Ho, S. K. Kong, and A. V. Kabashin, "Phase-sensitive surface plasmon resonance biosensors: methodology, instrumentation and applications," Ann. Phys. **524**(11), 637–662 (2012).
17. T. W. K. Chow, D. P. K. Lun, S. Pechprasarn, and M. G. Somekh, "Defocus leakage radiation microscopy for single shot surface plasmon measurement," Meas. Sci. Technol. **31**(7), 075401 (2020).
18. S. Pechprasarn, T. W. K. Chow, and M. G. Somekh, "Application of confocal surface wave microscope to self-calibrated attenuation coefficient measurement by Goos-Hanchen phase shift modulation," Sci. Rep. **8**(1), 8547 (2018).
19. B. Zhang, S. Pechprasarn, and M. G. Somekh, "Quantitative plasmonic measurements using embedded phase stepping confocal interferometry," Opt. Express **21**(9), 11523–11535 (2013).
20. J. E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng. **23**(4), 234350 (1984).