Transport of intensity equation: a new approach to phase and light field

Chao Zuo*[a], Qian Chen[a], Anand Asundi[b]

[a]*Jiangsu Key Laboratory of Spectral Imaging & Intelligence Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*
[b]*Centre for Optical and Laser Engineering, School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore*
surpasszuo@163.com

ABSTRACT

Phase is an important component of an optical wavefield bearing the information of the refractive index, optical thickness, or the topology of the specimen. Phase retrieval is a central problem in many areas of physics and optics since the phase of a wavefield is not accessible directly. The most well-established method for obtaining quantitative phase is through interferometry, such as digital holography. However, this class of methods relies on coherent illumination, therefore, plagued with problems of speckle that prevent the formation of high quality images. On a different note, quantitative phase can be retrieved by transport-of-intensity equation (TIE) using only object field intensities at multiple axially displaced planes. TIE has been increasingly investigated during recent years due to its unique advantages over interferometric techniques: it is non-interferometric, works with partially coherent illumination, computationally simple, no need to phase unwrapping, and does not require a complicated optical system. In this paper, we will review some recent new developments in TIE phase retrieval: including its numerical solution, treatment of boundary problem and the low-frequency artifacts, and configurations for dynamic phase imaging. We also reexamine TIE in terms of phase-space optics, demonstrating the effect of partially coherent illumination on phase reconstruction, and connecting it to light field imaging at the geometry optics limit.

Keywords: transport-of-intensity equation; 3D measurement; phase retrieval;

1. INTRODUCTION

When it comes to "phase measurement" or "quantitative phase imaging", many people will automatically connect them with "laser" and "interferometry". Indeed, conventional quantitative phase imaging and phase measurement techniques rely heavily on the superposition of two beams with a high degree of coherence: complex interferometric device, stringent requirement on the environmental stability, and associated laser speckle noise greatly limit its applications in the field of microscopic imaging. On the other hand, in spite of the insufficiency for interferometry, mircoscopists and biologists tend to use partially coherent illuminations (like Kohler's illumination in a conventional microscope) to achieve optimum spatial resolution, higher signal-to-noise ratio, and better image quality. To compensate for the inability to get the quantitative phase information, they also combine various types of sophisticated phase contrast imaging technology to make phase specimens become visible.

Although phase objects cannot be observed directly, they incessantly, and implicitly manifest their existence: the twinkling of stars in the night, the distorted scene outside the window in the rain, the network of bright lines at the bottom of a swimming pool in the sunshine. These are all manifestations of phase, implying the inextricably ties between phase and intensity of light wave. In 1983, Teague [1] first establish the quantitative relationship between the longitudinal intensity variation and phase of transporting light with use of a second-order elliptic partial differential equation, so called transport of intensity equation. With only intensity measurements at several distances along its propagation direction, the phase can be quantitatively retrieved by solving the equation deterministically, without resorting to interferometry or separated reference wave.

Holography, Diffractive Optics, and Applications VI, edited by Yunlong Sheng, Chongxiu Yu, Changhe Zhou,
Proc. of SPIE Vol. 9271, 92710H · © 2014 SPIE · CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2071713

Fig. 1, Phase measurement, interferometric (phase = fringe) and non-interferometric (phase->intensity).

2. TRANSPORT OF INTENSITY EQUATION

Assume that a paraxial beam is propagating along the z-axis, and let the complex amplitude of the object be written as $\sqrt{I(\boldsymbol{r})}\exp[ik\varphi(\boldsymbol{r})]$. The derivative of intensity in the light propagation direction z contains phase information that can be retrieved via TIE [1]:

$$-k\frac{\partial I(\boldsymbol{r})}{\partial z}=\nabla\cdot\left[I(\boldsymbol{r})\nabla\phi(\boldsymbol{r})\right], \tag{1}$$

where k is the wave number $2\pi/\lambda$, $\boldsymbol{r}$ is the position vector representing the spatial coordinates (x,y). $I(\boldsymbol{r})$ is the in-focus image intensity. ∇ is the gradient operator over $\boldsymbol{r}$ which is normal to the beam propagation direction. The TIE can be derived from Helmholtz equation under paraxial approximation, or using Fresnel wave propagation formulation with small propagation distance approximation, or alternatively, using the Poynting theorem (energy conservation law). Expanding the RHS of Eq. (1), we obtain

$$-k\frac{\partial I}{\partial z}=\nabla I\cdot\nabla\phi+I\nabla^2\phi, \tag{2}$$

In the above expression, the $\boldsymbol{r}$ dependence has been omitted for notation simplicity and note that the Hamilton ∇ applies only to the lateral coordinates (x,y). The first term on RHS of Eq. (2) is called prism term which stands for the longitude intensity variation due to the local wavefront slope. The second term is called lens terms which represent the intensity variation caused by the local wavefront curvature. It is seen that TIE links the longitudinal intensity derivative with the slope and curvature of the wavefront which produces the change in intensity as the wavefront propagates [2].

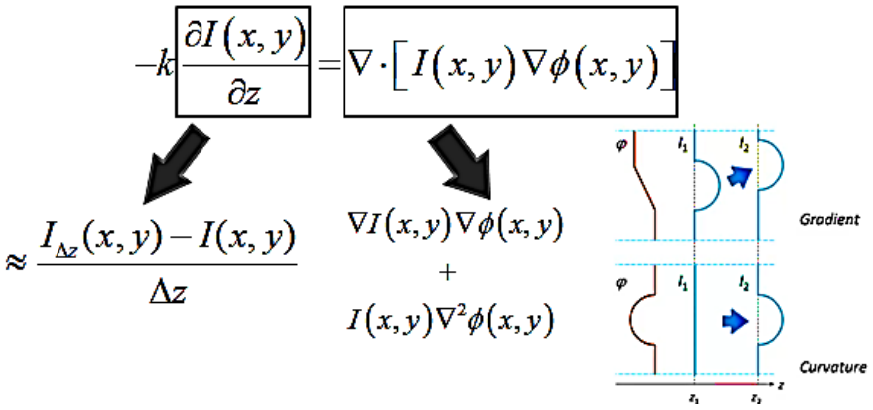

Fig. 2 The meaning of the transport-of-intensity equation.

3. SOLUTION

TIE is an elliptic partial differential equation for the phase function ϕ. With a specified boundary condition, one may find a unique solution provided that the intensity distribution is strictly positive [3]. That is, the phase can be uniquely determined by solving TIE with determined intensity and longitudinal intensity derivative (by simply capturing a minimum of two intensity images at different z planes and using them to estimate $\partial I/\partial z$). Though everything looks quite easy and straightforward, the rigorous implementation of the TIE phase retrieval tends to be difficult because the corresponding boundary conditions are not easy to obtain in practice. Such boundary conditions are usually based on *a priori* knowledge [4], for example, if the sample is isolatedly placed in the center of the camera field of view (FOV), surrounded by a unperturbed plane wave, then one can safely define some simplified boundary conditions, e.g., the homogeneous Dirichlet/Neumann boundary conditions, or the periodic boundary conditions. In this case, the most popular fast Fourier transform (FFT) based TIE solver [5-7] works well because it implies periodic boundary conditions due to the cyclic nature of the discrete Fourier transform. However, this situation is rather restrictive and does not reflect general experimental conditions. When the actual experimental situation violates those imposed assumptions, e.g., objects are located at the image borders, severe boundary artifacts will appear, seriously affecting the accuracy of the phase reconstruction [4, 8, 9]. To our knowledge, the boundary error problem is one major obstacle for the TIE to gaining extensive applications in the field of high-precision phase measurement.

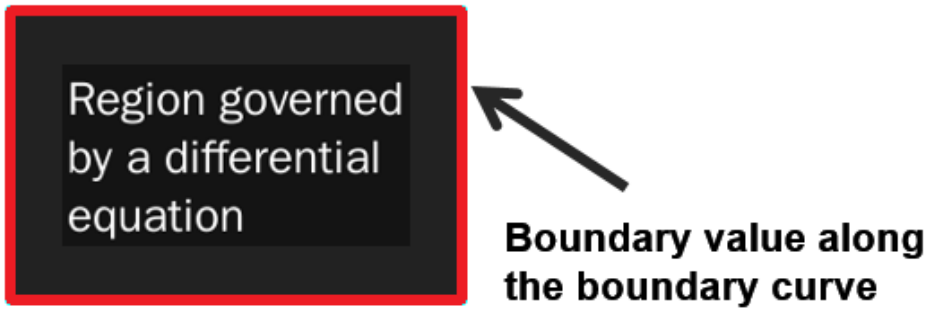

Fig. 3 The boundary conditions for a partially differential equation.

The most well-known fast TIE solver is proposed by Paganin and Nugent [7], which involves the use of FFT to implement the inverse Laplacian operator. Despite its popularity, it implies periodic boundary conditions, forcing the phase outside the domain to be periodically extended. When the opposite boundaries of the true phase distribution match poorly, this periodic assumption can cause serious boundary error across the reconstructed phase [4, 10]. It can be expected since generally the phase at the right boundary has nothing to do with the phase at the left boundary. Alternatively, Volkov *et al.* [8] suggest to symmetrically extend the intensity images into a four times larger size before using Paganin and Nugent's FFT-based solver to remedy the boundary error problem. This method has proven to work well in certain circumstances. Actually, taking the data outside the domain as a reflection of the data inside is equivalent to imposing two special cases of the (homogeneous) Dirichlet and the Neumann boundary conditions [4]. In spite of the value of this technique, it does not guarantee the elimination of the boundary artifacts. It can still create significant boundary artifacts since the assumed boundary conditions may not coincide with the ground truth data. Figure 4 shows an example of boundary artifacts arising from a simulated complex object extending across the image boundary, reconstructed using the Paganin and Nugent's standard FFT-based method (Fig. 4(d)), the even symmetric extension method (Fig. 4(e)), and the odd symmetric extension method (Fig. 4(f)). Note severe boundary errors can be obviously observed in all results because of the fact that none of the three classes of boundary conditions ((3)-(5)) can be fulfilled completely. Such kind of artifacts does not just appear at image borders but propagates inside the domain and degrades the reconstruction accuracy prevailingly.

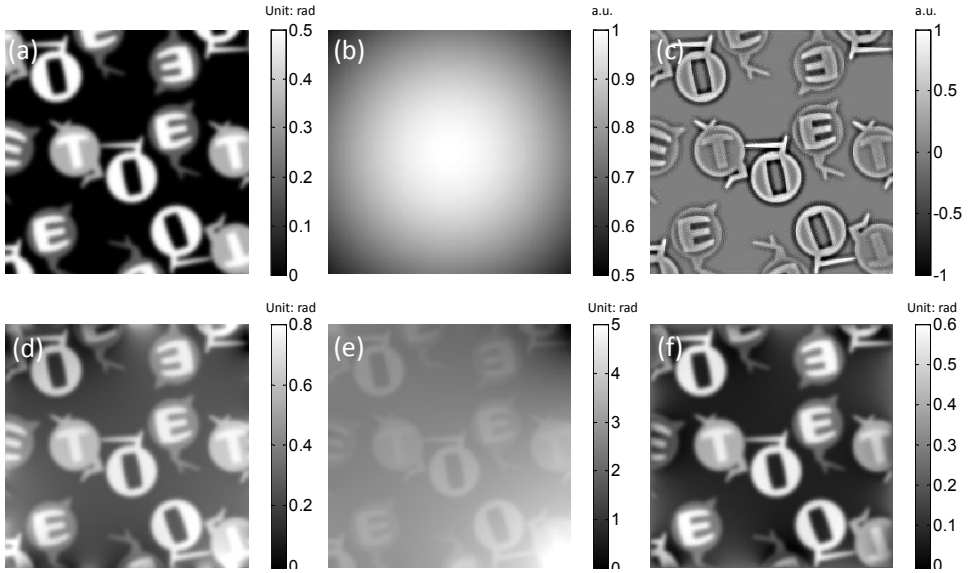

Fig. 4. Phase retrieval for a complex object covering the image boundary. (a) Phase distribution. (b) Intensity distribution. (c) Axial intensity derivative. (d) Phase retrieved by the FFT-based method (periodic boundary conditions). (e) Phase retrieved by the even symmetrization method (Zero Dirichlet boundary conditions). (f) Phase retrieved by the odd symmetrization method (Zero Neumann boundary conditions).

To this end, we recently developed a new method to solve the TIE with experimentally measurable boundary conditions, though an efficient numerical algorithm with use of fast discrete cosine transform (DCT) [10]. The new method has several substantial advantages: First, it clearly defines the TIE phase retrieval as an ***inhomogeneous Neumann boundary*** value problem, and the corresponding boundary values can be directly measured around a hard-edged aperture located in the in-focus plane. Second, it is based on an elegant formulation that requires no *a priori* knowledge about the test object and special-purpose detection scheme to explicitly extract the boundary signals. Finally, it retains the major advantages of the FFT-based solvers – fast and computationally simple for the rectangular domain. The first step of the approach is to introduce a hard-edged aperture in the object plane (or its conjugated plane). Then the in-focus intensity image captured by the camera can be obviously represented as

$$I = A_\Omega I_0 = \begin{cases} I_0 & \mathbf{r} \in \bar{\Omega} \\ 0 & others \end{cases}, \tag{3}$$

where I_0 is the intensity when there is no aperture; A_Ω is the aperture function (A_Ω = 1 when $\mathbf{r} \in \bar{\Omega}$, A_Ω = 0 when $\mathbf{r} \notin \bar{\Omega}$). The function of the aperture is to generate the required inhomogeneous Neumann boundary conditions for solving the TIE. Substituting Eq. (3) into the TIE (Eq. (1)), it can be derived that for this particular intensity distribution, the intensity transport can be written in the following form [10]

$$-k\frac{\partial I}{\partial z} = A_\Omega \left(I_0 \nabla^2 \phi + \nabla I_0 \cdot \nabla \phi \right) - I \frac{\partial \phi}{\partial n} \delta_{\partial\Omega}. \tag{4}$$

where $\delta_{\partial\Omega}$ is the Dirac delta function around the aperture edge. Equation (4) suggests that the axial intensity derivative signals (longitudinal variations of the intensity) consist of two non-overlapping components: (1) the intensity variation inside the domain due to the phase slope and curvature as if the aperture is not present; (2) a delta-function-like signal sharply peaked at the aperture boundary, which provides the exact Neumann boundary conditions for the TIE (Eq. (4)). Since the whole axial intensity derivative (the LHS of Eq. (4)) is experimentally measurable through finite difference scheme (note the aperture must be smaller than the image FOV so that all the boundary signals can be captured), and the two RHS terms do not overlap in space, there is enough information to solve the TIE uniquely without requiring *a prior* knowledge of the boundary conditions. The inhomogeneous boundary value problem has been proven to be ***well-posed***, because it automatically satisfies the following compatibility condition (derived from integrating both sides of Eq. (4) over all space) [10]:

$$\iint_{\bar{\Omega}} -k\frac{\partial I(\mathbf{r})}{\partial z}\mathrm{d}\mathbf{r} = \iint_{\Omega} -k\frac{\partial I(\mathbf{r})}{\partial z}d\mathbf{r} - \int_{\partial\Omega} I(\mathbf{r})\frac{\partial \phi(\mathbf{r})}{\partial n}ds = 0, \tag{5}$$

The physical picture described by Eq. (5) is just the ***energy conservation law*** – the total energy (intensity) of an isolated region is constant and unchangeable, and the loss of energy inside the region arising from energy flow across the region boundary. A detailed derivation of the DCT-based solver can be found in [10], and the final solution of the phase takes the following form

$$\phi(\mathbf{r}) = -k\nabla_{DCT}^{-2}\nabla_{DCT}\left[I^{-1}(\mathbf{r})\nabla_{DCT}\nabla_{DCT}^{-2}\frac{\partial I(\mathbf{r})}{\partial z}\right], \tag{6}$$

where ∇_{DCT} and ∇_{DCT}^{-2} are respectively the gradient and inverse Laplacian operator calculated through DCT-based approaches (see Appendix C of [10] for details). It should be noted that all the source data and related computations must be strictly limited to the rectangular-shaped area $\bar{\Omega}$, which includes both the aperture boundary and the region inside it (so that all the boundary signals can be enclosed in $\bar{\Omega}$). In this way, the measured intensity derivative can be treated as one entity, without requiring special-purpose detection scheme to explicitly extract the boundary signals (which has been known to cause serious difficulties [11, 12]).

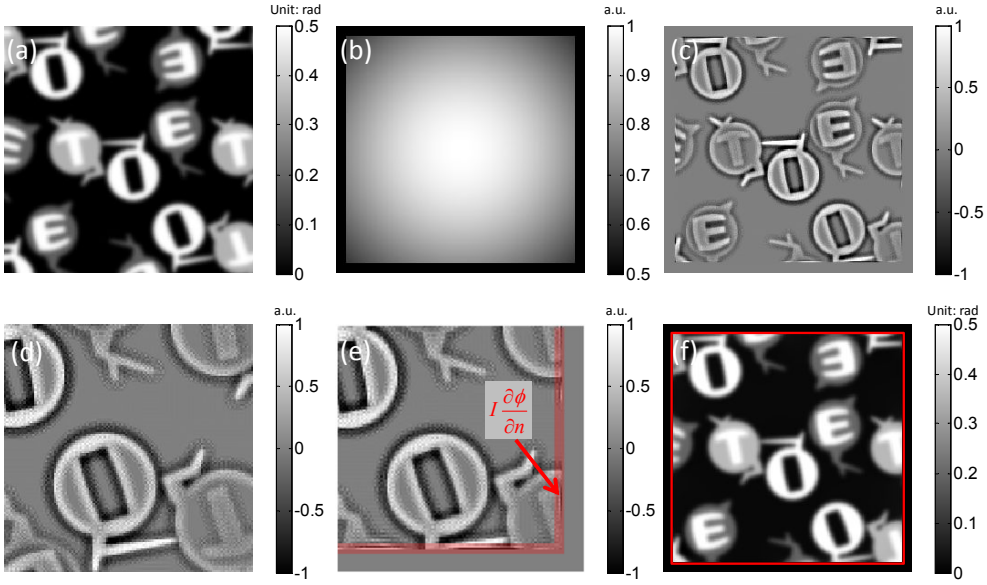

Fig. 5. Phase retrieval for a complex object covering the image boundary with the DCT-based method. (a) Phase distribution. (b) Intensity distribution (with a square aperture). (c) Axial intensity derivative. (d) Enlarged region corresponding to the lower right quarter of Fig. 1(c). (e) Enlarged region corresponding to the lower right quarter of Fig. 3(c), boundary signals can be clearly observed along the aperture edge (red shaded areas). (f) Phase retrieved by the DCT-based method (inhomogeneous Neumann boundary conditions). The red box in (f) outlines the aperture edge.

Figure 5 shows the same sample as in Fig. 4 but reconstructed through the DCT-based approach. A rectangular aperture is introduced in the in-focus plane (Fig. 5(b)) and the corresponding axial intensity derivative is shown in Fig. 5(c). From the enlarged regions shown in Figs. 5(d) and 5(e), it can be clearly seen that the aperture generates additional boundary signals (corresponds to the second term on the RHS of Eq. (4)), which cannot be observed in the absence of the aperture. With correct boundary conditions as well as the corresponding DCT-based TIE solver, the reconstructed result is free from any boundary artifacts, and closely matches the ground-truth image (Fig. 5(f)). This practical performance of this approach has been verified by applications of the micro-optics characterization [13], confirming that it effectively avoid the boundary error even when objects are located at the image borders.

4. PHASE DISCREPANCY COMPENSATION

It should be note that the standard procedure for solving the TIE includes the introduction of the so-called "Teague's assumption" that the transverse flux is conservative so that an a scalar potential (auxiliary function) ψ exists that satisfying

$$\boldsymbol{j}_{\mathbf{x}}(\mathbf{r}) = -k^{-1} I(\mathbf{r}) \nabla \phi(\mathbf{r}) = -k^{-1} \nabla \psi(\mathbf{r}), \tag{7}$$

In this way, the TIE can be reduced to two following standard Poisson equations

$$\frac{\partial I(\mathbf{r})}{\partial z} = -\frac{1}{k} \nabla^2 \psi(\mathbf{r}), \tag{8}$$

and

$$\nabla \cdot \left[I(\mathbf{r})^{-1} \nabla \psi(\mathbf{r}) \right] = \nabla^2 \phi(\mathbf{r}), \tag{9}$$

This treatment makes the TIE conveniently solvable by Fourier transform-based methods for periodic or some simplified homogeneous boundary conditions [6, 7], or the DCT-based method for inhomogeneous Neumann boundary conditions [10]. However, it is important to remark that the Teague's auxiliary function does not always exist in practical situations since the transverse energy flux may not be conservative, and consequently it would produce results that would not adequately match the exact solution [14, 15]. This discrepancy has not yet been recognized as a significant problem since previously the TIE was mainly applied for weak absorbing specimens, like optical components and biological samples. Although in such cases the methods using Teague's auxiliary function are not exactly accurate, they provide a very good approximation for the exact solution to the TIE because the transverse energy flux is nearly irrotational. However, the phase error is often nonnegligible and may even be relatively large when the measured sample exhibits a strong absorption. This obviously poses a significant problem in the practice of the TIE phase retrieval, which has been generally acknowledged to be quantitative. To this end, we have present the theoretical analysis of the phase discrepancy owing to the Teague's auxiliary function as well as a simple iterative algorithm to compensate such phase errors [16]. The recovered phase is firstly pushed back to Eq. (1) to generate the intensity derivative data, and then followed by a consistency checking procedure with the measured $\partial I/\partial z$, the difference between the two intensity derivatives is called error signal, which is used as the source term for the next round TIE reconstruction. The algorithm usually converges very rapidly and the phase discrepancy can be eliminated after three to five iterations.

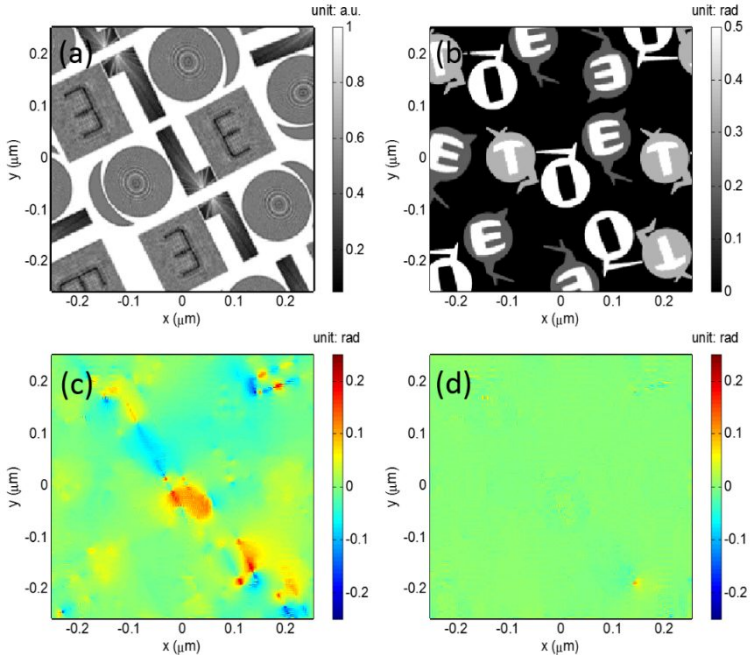

Fig. 6. Simulation results of the iterative compensation algorithm. (a) Intensity distribution. (b) Phase distribution. (c) Phase error before compensation. (d) Phase error after five iterations of the compensation.

5. AXIAL INTENSITY DERIVATIVE ESTIMATION

The TIE requires to know the in-focus intensity I and axial intensity derivative $\partial I/\partial z$. Experimentally, I is easy to obtain. However, the intensity derivative along the optic axis cannot be directly measured. Conventionally, it is approximated by a finite difference between two out-of-focus images, recorded symmetrically about the in-focus plane with $\pm\Delta z$ defocus distance.

$$\frac{\partial I(\mathbf{r})}{\partial z} \approx \frac{I_{\Delta z}(\mathbf{r}) - I_{-\Delta z}(\mathbf{r})}{2\Delta z}. \tag{10}$$

The phase errors originate from the noise in the captured intensities and the nonlinear error component related to the finite difference approximation. Since the low-frequency phase structure shows less phase contrast than the high-frequency phase structure as the wave propagation [17-19], the TIE is very sensitive to low-frequency artifacts, especially when the defocus distance is chosen too small. Using a larger separation provides a better signal to noise ratio; nevertheless, the breakdown of the linear approximation that underlies the finite difference approximation induces nonlinearity errors that will reduce the phase resolution. To obtain a compromise between nonlinearity error and the low-frequency noise, there exists an optimal distance which is dependent on both the maximum physically significant frequency of the object and the level of noise [20-22]. A priori knowledge of these two aspects is difficult not be known in advance.

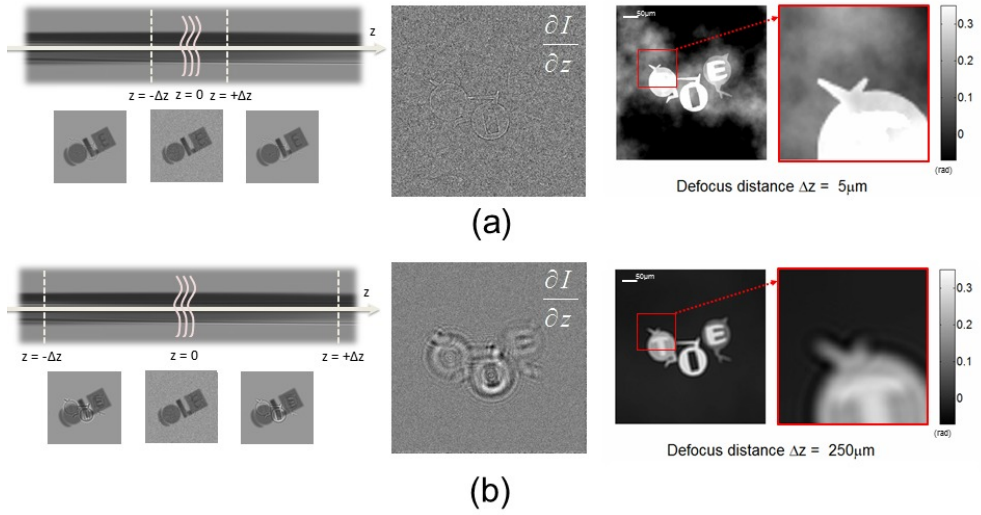

Fig. 7. Effect of the noise and nonlinearity errors on TIE phase retrieval. (a) Small defocus is sensitive to low-frequency noise; (b) large defocus causes phase blurring.

To overcome this difficulty, one may simply increase the number intensity measurements at multiple planes [22, 23]. With more intensity measurements $I_{j\Delta z}(\mathbf{r}), j = -n,...,0,...,n$, the longitudinal intensity derivative can be represented by their linear combination:

$$\frac{\partial I(\mathbf{r})}{\partial z} \approx \sum_{j=-n}^{n} \frac{a_j I_{j\Delta z}(\mathbf{r})}{\Delta z}. \tag{11}$$

Thus, it offers more flexibility for improving the accuracy and noise resistance in derivative estimation. The only difference in these multiple planes derivative estimation methods lies in the coefficients a_i in Eq. (11), and it has been found that all these method can be unified into Savitzky-Golay differentiation filter (SGDF) with different degrees if the finite difference [Eq. (11)] is viewed from the viewpoint of the digital filter [21]. The different viewpoint from the digital filter in signal processing provides great insight into the behaviors, the shortcomings, and the performance of these existing intensity derivative estimation algorithms, revealing that in traditional multiple planes derivative estimation methods, the contradiction between low-frequency artifacts and high-order nonlinearity errors is still irreconcilable. Based on the unified framework, we have propose an optimal frequency selection (OFS) method efficiently solves the contradiction between low-frequency artifacts and high-order nonlinearity errors, improving the noise resistance and

phase reconstruction accuracy significantly. The OFS is a three-step process with the basic idea that selecting the optimal components of the phase recovered from SGDFs with various degrees in spatial frequency domain to account for both phase SNR and nonlinearity error: First, estimate the intensity derivatives with SGDFs with different degrees through Eq. (11). Second, reconstruct the phase distributions with TIE using the estimated intensity derivatives. Finally, extract the optimal frequency components from these phase distributions using a complementary filter bank in spatial frequency domain and then recombine them into a composite phase. Recently, the concept of the OSF has been extended to unequally (exponentially) spaced defocus planes [24], and explicitly taking the influence of spatial coherence of the illumination into account [25].

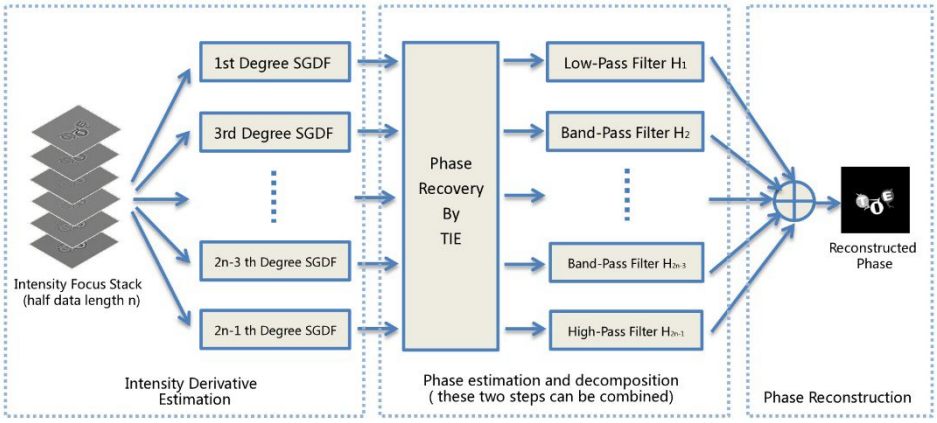

Fig. 8. The flow chart for the optimal frequency selection scheme.

6. PHASE SPACE REPRESENTATION

The fact that TIE is not restricted to the purely coherent regime but works also with a partially coherent source seems to be very well known. However, partially coherent field does not have a well-defined phase since the field experiences statistical fluctuations over time, which brings up an important question about what "phase" is being measured in such scenarios. Phase-space optics provides an elegant description of TIE under partially coherent illumination. Let $U(\mathbf{x})$ be a stationary and ergodic quasi-monochromatic (optical frequency $\omega = c/\lambda$, where c is the speed of light and λ is the wavelength) paraxial scalar field with arbitrary coherence, where $\mathbf{x}$ is the two-dimensional spatial vector. This partially coherent field can be characterized by the Wigner distribution function (WDF) [26]

$$W_\omega(\mathbf{x},\mathbf{u}) = \int \Gamma\left(\mathbf{x}+\frac{\mathbf{x}'}{2}, \mathbf{x}-\frac{\mathbf{x}'}{2}\right)\exp\left(-i2\pi\mathbf{u}\mathbf{x}'\right)d\mathbf{x}', \tag{12}$$

where $\mathbf{x}$ and $\mathbf{u}$ are the two-dimensional spatial and spatial frequency vectors, respectively, Γ is the mutual intensity related to the ensemble average $\Gamma(\mathbf{x}_1,\mathbf{x}_2) = \langle U(\mathbf{x}_1)U^*(\mathbf{x}_2)\rangle$. With use of the Liouville transport equation [26] , we can derive a generalized transport of intensity equation (GTIE) which is valid under partially coherent field

$$\frac{\partial I(\mathbf{x})}{\partial z} = -\nabla_{\mathbf{x}}\cdot\iint \lambda\mathbf{u}W_\omega(\mathbf{x},\mathbf{u})d\mathbf{u}d\omega. \tag{13}$$

Note the only assumption employed in deriving Eq. (13) is the paraxial field to be stationary and ergodic, thus it is general enough to cover various optical fields with arbitrary spatial and temporal coherence. The spectral dependence ω can be eliminated by assuming the field to be quasi-monochromatic, and thus the GTIE for partially spatially coherent fields can be written as

$$\frac{\partial I(\mathbf{x})}{\partial z} = -\lambda\nabla_{\mathbf{x}}\cdot\int \mathbf{u}W(\mathbf{x},\mathbf{u})d\mathbf{u}. \tag{14}$$

In completely coherent case, the field can be fully described by the 2D complex amplitude $U(\mathbf{x}) = \sqrt{I(\mathbf{x})}\exp[i\phi(\mathbf{x})]$, where ϕ is the (conventional) phase. From the time(space)-frequency analysis perspective, the completely coherent field can be regarded as a mono-component signal, and the first conditional frequency moment of WDF (instantaneous frequency) relates to the transverse gradient of the phase of the complex field [27]:

$$\frac{\int \mathbf{u} W(\mathbf{x},\mathbf{u})\, d\mathbf{u}}{\int W(\mathbf{x},\mathbf{u})\, d\mathbf{u}} = \frac{1}{2\pi}\nabla\phi(\mathbf{x}). \tag{15}$$

Substitution of Eq. (15) into Eq. (14) leads to coherent TIE

$$\frac{\partial I(\mathbf{x})}{\partial z} = -\frac{1}{k}\nabla\cdot\left[I(\mathbf{x})\nabla\phi(\mathbf{x})\right]. \tag{16}$$

The partially coherent field does not have a well-defined phase since the field experiences statistical fluctuations over time. However, the phase-space representation on LHS of Eq. (15) is still valid, leading to a new meaningful and more general definition of "phase". Here we refer the new "phase" $\hat{\phi}(\mathbf{x})$ defined by Eq. (15) as the generalized phase of partially coherent fields to distinguish it from its coherent counterpart. The generalized phase behaving precisely as the conventionally defined phase and directly reduces to conventional phase when the field is fully coherent. It is seen from Eq. (15) that the generalized phase is a scalar potential whose gradient yields the conditional frequency moment of the WDF [28]. It is clear from a distribution point of view, that quantity is the average frequency at a particular location.

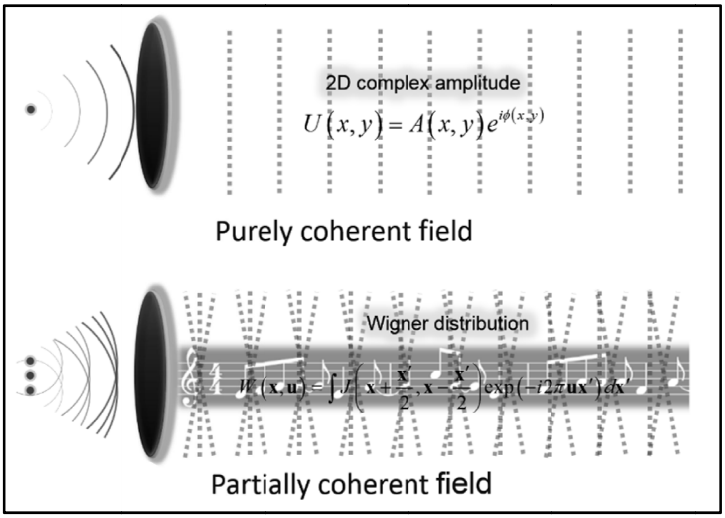

Fig. 9. For purely coherent light, it can be fully characterize by the 2D complex function. While the (spatially) partially coherent field can be viewed as a suitable superposition of several independent coherent modes, thus, requires 4D coherence function (e.g. Wigner distribution function) to fully describe.

7. PHASE RETRIEVAL UNDER PARTIALLY COHERENT ILLUMINATION

Since the partial coherence is explicitly considered in GTIE, it can be used to analyze the phase retrieval under partially coherent illumination. It should be emphasized that the major concern in such scenario is the well-defined phase shift introduced by the specimen rather than the generalized phase of the partially coherent field itself. This leads to the nature choice of treating the contributions of the incident illumination and specimen separately by considering the transmitted field as a product of the illumination function $U_{in}(\mathbf{x})$ and the sample transmission function $T(\mathbf{x}) = \sqrt{\tau(\mathbf{x})}\exp\left[i\phi(\mathbf{x})\right]$, where $\tau(\mathbf{x})$ and $\phi(\mathbf{x})$ are the amplitude and the phase functions of the specimen. It is thus not hard to find the WDF of the resultant field just leaving the object can be represented as the object transmittance WDF $W_T(\mathbf{x},\mathbf{u})$ windowed by the illumination WDF $W_{in}(\mathbf{x},\mathbf{u})$

$$W_{out}(\mathbf{x},\mathbf{u}) = \int W_T(\mathbf{x},\mathbf{u}')W_{in}(\mathbf{x},\mathbf{u}-\mathbf{u}')\, d\mathbf{u}' = W_T(\mathbf{x},\mathbf{u}) \underset{\mathbf{u}}{\otimes} W_{in}(\mathbf{x},\mathbf{u}), \tag{17}$$

which is a mere convolution over the spatial frequency variable $\mathbf{u}$. Equation (17) reveals that the partially coherent incident illumination blurs the object WDF along the spatial frequency dimension. By taking it into LHS of Eq. (15), it is then readily to find the generalized phase of the transmitted field $\hat{\phi}_{out}(\mathbf{x})$ should satisfy the following expression

$$\frac{\int \mathbf{u} W_{out}(\mathbf{x},\mathbf{u})\,d\mathbf{u}}{\int W_{out}(\mathbf{x},\mathbf{u})\,d\mathbf{u}} = \frac{\int \mathbf{u} W_{T}(\mathbf{x},\mathbf{u})\,d\mathbf{u}}{\int W_{T}(\mathbf{x},\mathbf{u})\,d\mathbf{u}} + \frac{\int \mathbf{u} W_{in}(\mathbf{x},\mathbf{u})\,d\mathbf{u}}{\int W_{in}(\mathbf{x},\mathbf{u})\,d\mathbf{u}}, \tag{18}$$

or equivalently,

$$\nabla_{\mathbf{x}} \hat{\phi}_{out}(\mathbf{x}) = \nabla_{\mathbf{x}} \left[\hat{\phi}_{in}(\mathbf{x}) + \phi(\mathbf{x}) \right]. \tag{19}$$

This representation shows the generalized phase accrues upon propagation through the object, behaving precisely as would the conventionally defined phase. In general, the determination of object phase requires two independent measurements, performed respectively with and without the presence of the specimen. The sample-free measurement is used to characterize the incident beam $\hat{\phi}_{in}(\mathbf{x})$ and is subsequently subtracted from the total generalized phase $\hat{\phi}_{out}(\mathbf{x})$ to get the net phase introduced by the object only. However, if the illumination is chosen judiciously to directly nullify $\hat{\phi}_{in}(\mathbf{x})$,

$$\int \mathbf{u} W_{in}(\mathbf{x},\mathbf{u})\,d\mathbf{u} = 0. \tag{20}$$

The total generalized phase $\hat{\phi}_{out}(\mathbf{x})$ directly gives $\phi(\mathbf{x})$ and one single measurement is sufficient to recover the object phase even though the illumination is not fully coherent. For purely coherent illumination, Eq. (20) simply means the incident illumination is an on-axis plane wave. For the illumination created by an incoherent extended source, the condition Eq. (21) means the primary source distribution must be symmetrical about the optical axis.

8. PHASE AND LIGHT FIELD

For completely coherent fields, the phase-space representation is highly redundant because the complex field is defined only over the 2D plane $\mathbf{x}$. The 2D intensity and the reconstructed phase distribution gives total knowledge about the complex field so that the behavior of the field can be perfectly predicted. Such complete knowledge permits a variety forms of coherent optical imaging systems such as the Zernike phase contrast and the differential interference contrast imaging to be computationally emulated without resort to actual optical hardware [29, 30]. If the phase varies slowly such that the approximations $\phi(\mathbf{x}+\mathbf{x}/2)-\phi(\mathbf{x}-\mathbf{x}/2) \approx \mathbf{x}\cdot\nabla\phi(\mathbf{x})$ is valid, the phase space redundancy becomes more apparent since the signal occupies only a single slice in phase space

$$W(\mathbf{x},\mathbf{u}) = I(\mathbf{x})\,\delta\left[\mathbf{u} - \frac{1}{2\pi}\nabla\phi(\mathbf{x})\right]. \tag{21}$$

The form of WDF given above now is a true energy probability distribution in phase-space, telling us the geometric ray or energy flow at single position travels only along single direction described by the phase normal. This is an advantageous feature to allow phase measurement simply by measuring the directions of the rays, using the Shack-Hartmann sensor [31].

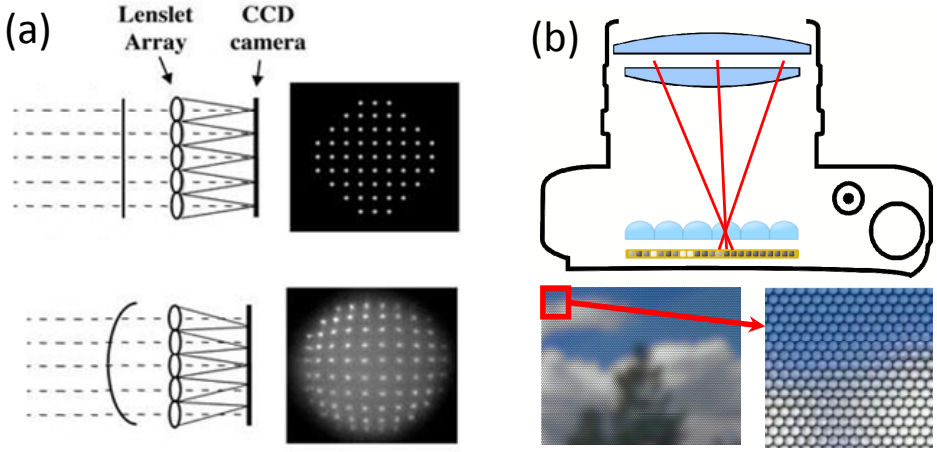

Fig. 10. Principle of the Shack-Hartmann sensor and Light field imaging. (a) Shack-Hartmann sensor. (b) Light field imaging.

The Shack-Hartmann sensor is an array of lenslets, each of which brings the incident field to a focus. When normally incident plane waves are shone onto the sensor, each lenslet brings the light to a focus at the center of its associated detector. If an aberrated wave is incident onto the sensor, then the location of each spot will be displaced by a vector proportional to the average phase gradient $\nabla\phi(\mathbf{x})$ over the lenslet. This displacement may be sensed by, for example, a quadrant detector. The resulting signals should be integrated to create an estimate of the phase distribution. Though its conceptual simplicity, the disadvantage of the Shack-Hartmann sensor is obvious: the spatial resolution over the wavefront is limited by the size and number of the lenslets.

For partially coherent field (include the condition of the totally incoherent fields), the 4D WDF is generally non-redundant. From the geometrical optics perspective, the geometric ray at single position does not travel only in one direction; instead, it fans out to make a 2D distribution, which account for the higher dimensionality of the partially coherent field. The light field camera, as a counterpart of the Shack-Hartmann sensor in computer graphics community, allows joint measurement of the spatial and directional distribution of light [32]. The "light field" is a term commonly used in the computer graphics literature to represent a collection of light rays in geometric optics. It is parameterized by a four-variable function $L(\mathbf{x},\boldsymbol{\theta})$, taking into account both the geometrical position of the rays $\mathbf{x}$ and also their directions $\boldsymbol{\theta}$. It approaches the WDF at geometric optics limit [33]. Light field imaging enables us to apply ray-tracing techniques to compute synthetic photographs, change the focus and perspective view flexibly. However, it requires elaborate optical setups and significantly sacrifices spatial resolution (traded for angular information) as compared to conventional imaging technique.

TIE provides a simple and deterministic way to recover phase from a set of defocused images. However, acquiring the through-focus image stack is usually time-consuming, as the sample stage or camera has to be moved between image captures. Though several configurations have been developed to eliminate the mechanical motion (see next session), light field imaging enables a totally new way to collect the entire image stack in single capture, as the intensity images at an arbitrary focal plane can be computationally reconstructed from the raw light field image. This undoubtedly suggests one viable way to convert the light field to the phase.

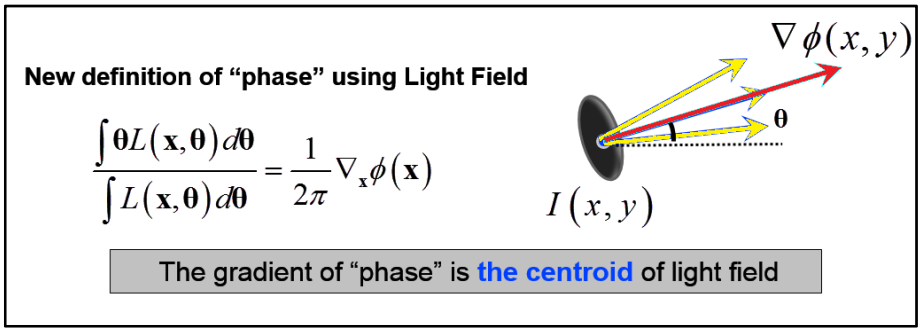

Fig. 11. "Phase" is the centroid of the light field in geometric optics.

It would be more interesting to consider this problem from a different perspective: Eq. (15) suggests that the generalized phase is a scalar potential whose gradient yields the conditional frequency moment of the WDF. Under geometric optics approximation $L(\mathbf{x},\boldsymbol{\theta}) \approx W(\mathbf{x},\lambda\mathbf{u})$, Eq. (15) becomes

$$\frac{\int \boldsymbol{\theta} L(\mathbf{x},\boldsymbol{\theta})\, d\boldsymbol{\theta}}{\int L(\mathbf{x},\boldsymbol{\theta})\, d\boldsymbol{\theta}} = \frac{1}{2\pi}\nabla\phi(\mathbf{x}). \tag{22}$$

It is clear from the above definition that the quantity on the LHS is the centroid of the light field - the average direction of light at one given position, suggesting that the phase gradient can be easily recovered by a simple centroid detection scheme applied to the raw light field image. This is similar with the standard procedure in the Shack-Hartmann method. Compared to the first method, which first reconstructs the through-focus stack, then solves the TIE explicitly; the second method, which employs the definition of the generalized phase here, is inherently much easier and straightforward.

It is quite understandable that the phase can be recovered from the light field since the 4D light field is inherently in a higher dimensionality which totally cover the 2-D phase information. But it would be quite interesting to consider whether the inverse of this process can be realized: with the knowledge of phase, is it possible to recover the whole light field? Apparently, the answer should be NO for the general cases. However, under certain conditions, we can indeed covert the phase to the light field. Besides the trivial case when the field is purely coherent, let us consider that the sample is illuminated by a spatial stationary illumination that is generally true for the experimental arrangements in

optical microscopy [26]. The fully spatially incoherent primary source (usually at the condenser aperture plane for an optical microscope) featured by the intensity distribution $\left|P_c(\mathbf{x})\right|^2$ and the positional cross-spectral density $\Gamma(\mathbf{x}+\mathbf{x}'/2,\mathbf{x}-\mathbf{x}'/2)=\left|P_c(\mathbf{x})\right|^2\delta(\mathbf{x}')$ produces the illumination WDF $W_{in}(\mathbf{x},\mathbf{u})$ at the far-zone

$$W_{in}(\mathbf{x},\mathbf{u})=\left|P_c(\mathbf{u})\right|^2. \tag{23}$$

Equation (23) is in fact an expression of the Van Cittert-Zernike theorem. If the phase varies slowly such that the approximations $\phi(\mathbf{x}+\mathbf{x}/2)-\phi(\mathbf{x}-\mathbf{x}/2)\approx\mathbf{x}\cdot\nabla\phi(\mathbf{x})$, the resultant field just leaving the object can be represented as

$$L(\mathbf{x},\boldsymbol{\theta})=cI(\mathbf{x})\left|P_c\left[\boldsymbol{\theta}-\frac{1}{k}\nabla_{\mathbf{x}}\phi(\mathbf{x})\right]\right|^2, \tag{24}$$

where c is a constant ensuring $I(\mathbf{x})=\int L(\mathbf{x},\boldsymbol{\theta})d\boldsymbol{\theta}$. Note the last step we used the approximation $L(\mathbf{x},\boldsymbol{\theta})\approx W(\mathbf{x},\lambda\mathbf{u})$. Equation (24) represents exactly the geometric optical behavior of the specimen: for each incident ray, it leaves the specimen from the same position but its direction is shifted as a function of the phase gradient of the object. The specimen, can be regarded as a spread-less system, does not change the angular distribution of the incident field, which is fully determined by the source intensity distribution. Thus, with the knowledge of the source intensity distribution and the phase of the object $\phi(\mathbf{x})$ (can be retrieved from TIE with only 2 images), the 4D light-field can be fully characterized.

9. DYNAMIC PHASE IMAGING

TIE requires a series of images captured at different focal depths, which is usually realized by translating the camera or the object manually or mechanically. This not only complicates the image acquisition process, but also prolongs the measurement time, precluding real-time observation of dynamic process. To this end, we introduce two schemes to extend transport of intensity quantitative phase imaging to dynamics applications. Both the two schemes are built upon a commercial microscope by introducing an additional image-relay system to the camera port of microscope. The first scheme (called tunable lens based TIE (TL-TIE) [30]) introduce an electronically tunable lens (ETL) in the Fourier plane of the image-relay system, which allows for rapid electronically-controlled re-focusing with constant magnification without altering the original high imaging quality of the microscope. The second scheme called single-shot quantitative phase microscopy (SQPM) [34] introduces a Michelson-like architecture with a spatial light modulator between the image-relay system, so that two laterally separated images at different focus distance can be obtained simultaneously by a single camera exposure, enabling TIE phase recovery to be performed at frame rates that are only camera limited.

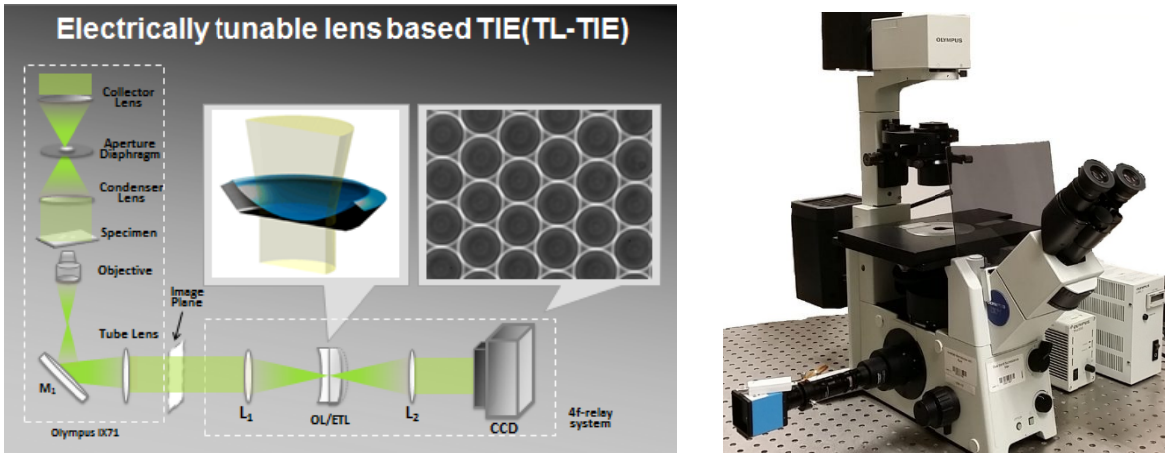

Fig. 12 Schematic and the current prototype system of the tunable lens based TIE (TL-TIE) system.

These schemes offer the possibility to extend TIE phase imaging to the study of fast moving objects and structural changes in dynamic processes. By using the numerical reconstruction implemented by the computer software, real-time transport of intensity imaging can be realized. We have successfully applied the technology in the characterization of microlens array, investigations of cellular dynamics and drug-induced morphology changes. Figure 13 shows the 3D rendering of phase maps of one macrophage cell at different stages of apoptosis.

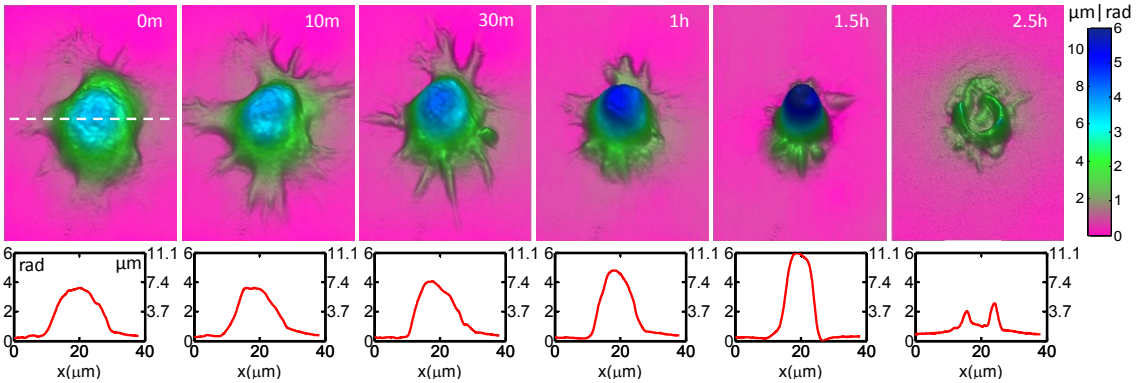

Fig. 13. Morphological changes of a macrophage during chemical-induced apoptosis. The upper row shows the color-coded 3D phase distributions with phase values and cell thickness cross sections (lower row) in different periods after chemical treatment.

10. DIRECT CONTINUOUS RECOVERY FOR DIGITAL HOLOGRAPHY

One of the great advantages of TIE is able to directly reconstruct the continuous phase of the wave, without the need for phase unwrapping [35, 36]. Taking advantage of this appealing feature, we can combine it with digital holography. Numerical focusing is a unique capability of digital holography, where a single hologram is used to calculate the optical field at any number of image planes, emulating the focusing control of a conventional imaging system. This inspires us to properly combine these two techniques so that merits of both methods can be accrued. One can directly recover the continuous phase in digital holography based on solving the TIE from the normalized wavefield obtained during the numerical reconstruction of the hologram, and meanwhile, eliminate the tilt and quadratic phase aberration inherent in digital holography without cumbersome physical or numerical compensation procedure [37]. But it should be mentioned that though no need for unwrapping is an advantageous feature of the TIE, the famous "2π ambiguity problem" is since the TIE implies the phase is a continuous function [38].

11. CONCULSIONS AND FURTHER DIRECTIONS

In this paper, we have introduced and reviewed some recent developments in TIE computational imaging, which can be broadly categorized as follows:

- Fast, exact numerical solution and the boundary conditions problem
- Axial intensity derivative estimation and the noise problem
- Extension of transport of intensity equation to partially coherent fields
- Relations between "phase" and light field
- High speed through-focus stack collection for dynamic imaging

This new approach to phase and light field endows the conventional microscope with the ability of 3D quantitative phase imaging, enabling us to obtain 3D morphology and optical properties of phase specimens with nanometer-scale and millisecond temporal resolution. The TIE-based imaging provides the following unique advantages：

- It is non-interferometric, works with partially coherent illumination, like the built-in halogen lamp in a conventional microscope.
- It does not require a complicated optical system, can well be compatible with the conventional bright field microscope.
- It is a single beam method and less stringent to environmental instability.
- It directly recovers the continuous phase without 2π discontinuity, so there is no need to phase unwrapping.
- It can provide high-quality, real-time phase measurement with high resolution in full four-dimension (lateral (x and y) and vertical (z) and temporal (t)).

Despite these advantages, there are many unexplored problems in the application of this novel imaging method. Venue to extend the current study includes:

- Fast numerical solution of TIE for a arbitrarily-shaped region
- Extension of transport of intensity equation to recovery of the 4D coherence function
- Extension of TIE to full 3D refractive index tomography for volumetric objects with low-coherent illumination

ACKNOWLEDGEMENT

This project was supported by the Research Fund for the Doctoral Program of Ministry of Education of China (No. 20123219110016) and the Research and Innovation Plan for Graduate Students of Jiangsu Higher Education Institutions, China (No. CXZZ11_0237). C. Zuo gratefully acknowledges the financial support from China Scholarship Council (No. 201206840009).

REFERENCES

[1] M. Reed Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am., 73(11), 1434-1441 (1983).

[2] K. Ichikawa, A. W. Lohmann, and M. Takeda, "Phase retrieval based on the irradiance transport equation and the Fourier transform method: experiments," Applied Optics, 27(16), 3433-3436 (1988).

[3] T. E. Gureyev, A. Roberts, and K. A. Nugent, "Partially coherent fields, the transport-of-intensity equation, and phase uniqueness," J. Opt. Soc. Am. A, 12(9), 1942-1946 (1995).

[4] J. Martinez-Carranza, K. Falaggis, T. Kozacki *et al.*, "Effect of imposed boundary conditions on the accuracy of transport of intensity equation based solvers," 87890N-87890N (2013).

[5] T. E. Gureyev, and K. A. Nugent, "Phase retrieval with the transport-of-intensity equation. II. Orthogonal series solution for nonuniform illumination," J. Opt. Soc. Am. A, 13(8), 1670-1682 (1996).

[6] T. E. Gureyev, and K. A. Nugent, "Rapid quantitative phase imaging using the transport of intensity equation," Optics Communications, 133(1–6), 339-346 (1997).

[7] D. Paganin, and K. A. Nugent, "Noninterferometric Phase Imaging with Partially Coherent Light," Physical Review Letters, 80(12), 2586-2589 (1998).

[8] V. V. Volkov, Y. Zhu, and M. De Graef, "A new symmetrized solution for phase retrieval using the transport of intensity equation," Micron, 33(5), 411-416 (2002).

[9] J. Frank, S. Altmeyer, and G. Wernicke, "Non-interferometric, non-iterative phase retrieval by Green's functions," J. Opt. Soc. Am. A, 27(10), 2244-2251 (2010).

[10] C. Zuo, Q. Chen, and A. Asundi, "Boundary-artifact-free phase retrieval with the transport of intensity equation: fast solution with use of discrete cosine transform," Optics Express, 22(8), 9220-9244 (2014).

[11] I. W. Han, "New method for estimating wavefront from curvature signal by curve fitting," Optical Engineering, 34(4), 1232-1237 (1995).

[12] T. Gureyev, A. Roberts, and K. Nugent, "Phase retrieval with the transport-of-intensity equation: matrix solution with use of Zernike polynomials," JOSA A, 12(9), 1932-1942 (1995).

[13] C. Zuo, Q. Chen, H. Li *et al.*, "Boundary-artifact-free phase retrieval with the transport of intensity equation II: applications to microlens characterization," Optics Express, 22(15), 18310-18324 (2014).

[14] J. A. Schmalz, T. E. Gureyev, D. M. Paganin *et al.*, "Phase retrieval using radiation and matter-wave fields: Validity of Teague's method for solution of the transport-of-intensity equation," Physical Review A, 84(2), 023808 (2011).

[15] J. A. Ferrari, G. A. Ayubi, J. L. Flores *et al.*, "Transport of intensity equation: Validity limits of the usually accepted solution," Optics Communications, 318(0), 133-136 (2014).
[16] C. Zuo, Q. Chen, L. Huang *et al.*, "Phase discrepancy analysis and compensation for fast Fourier transform based solution of the transport of intensity equation," Optics Express, 22(14), 17172-17186 (2014).
[17] N. Streibl, "Three-dimensional imaging by a microscope," J. Opt. Soc. Am. A, 2(2), 121-127 (1985).
[18] E. D. Barone-Nugent, A. Barty, and K. A. Nugent, "Quantitative phase-amplitude microscopy I: optical microscopy," Journal of Microscopy, 206(3), 194-203 (2002).
[19] C. J. R. Sheppard, "Defocused transfer function for a partially coherent microscope and application to phase retrieval," J. Opt. Soc. Am. A, 21(5), 828-831 (2004).
[20] J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Optimum measurement criteria for the axial derivative intensity used in transport of intensity-equation-based solvers," Optics letters, 39(2), 182-185 (2014).
[21] C. Zuo, Q. Chen, Y. Yu *et al.*, "Transport-of-intensity phase imaging using Savitzky-Golay differentiation filter - theory and applications," Opt. Express, 21(5), 5346-5362 (2013).
[22] 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).
[23] M. Soto, and E. Acosta, "Improved phase imaging from intensity measurements in multiple planes," Appl. Opt., 46(33), 7978-7981 (2007).
[24] Z. Jingshan, R. A. Claus, J. Dauwels *et al.*, "Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes," Optics Express, 22(9), 10661-10674 (2014).
[25] M. H. Jenkins, J. M. Long, and T. K. Gaylord, "Multifilter phase imaging with partially coherent light," Applied Optics, 53(16), D29-D39 (2014).
[26] M. J. Bastiaans, "Application of the Wigner distribution function to partially coherent light," Journal of the Optical Society of America A, 3(8), 1227-1238 (1986).
[27] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," Proceedings of the Ieee, 80(4), 520-538 (1992).
[28] C. Zuo, Q. Chen, and A. Asundi, "Light field moment imaging: comment," Optics letters, 39(3), 654-654 (2014).
[29] D. Paganin, T. Gureyev, S. Mayo *et al.*, "X-ray omni microscopy," Journal of Microscopy, 214(3), 315-327 (2004).
[30] C. Zuo, Q. Chen, W. Qu *et al.*, "High-speed transport-of-intensity phase microscopy with an electrically tunable lens," Optics Express, 21(20), 24060-24075 (2013).
[31] B. C. Platt, "History and principles of Shack-Hartmann wavefront sensing," Journal of Refractive Surgery, 17(5), S573-S577 (2001).
[32] R. Ng, M. Levoy, M. Brédif *et al.*, "Light field photography with a hand-held plenoptic camera," Computer Science Technical Report CSTR, 2(11), (2005).
[33] Z. Zhengyun, and M. Levoy, "Wigner distributions and how they relate to the light field." 1-10.
[34] C. Zuo, Q. Chen, W. Qu *et al.*, "Noninterferometric single-shot quantitative phase microscopy," Optics letters, 38(18), 3538-3541 (2013).
[35] A. Barty, K. A. Nugent, D. Paganin *et al.*, "Quantitative optical phase microscopy," Opt. Lett., 23(11), 817-819 (1998).

[36] C. J. Bellair, C. L. Curl, B. E. Allman *et al.*, "Quantitative phase amplitude microscopy IV: imaging thick specimens," Journal of Microscopy, 214(1), 62-69 (2004).
[37] C. Zuo, Q. Chen, W. Qu *et al.*, "Direct continuous phase demodulation in digital holography with use of the transport-of-intensity equation," Optics Communications, 309(0), 221-226 (2013).
[38] C. Zuo, Q. Chen, and A. Asundi, [Comparison of Digital Holography and Transport of Intensity for Quantitative Phase Contrast Imaging] Springer Berlin Heidelberg, 17 (2014).