High speed phase retrieval of in-line holograms by the assistance of corresponding off-axis holograms

László Orzó∗
Computational Optical Sensing and Processing Laboratory, Institute for Computer Science and Control, Hungarian Academy of Sciences H-1111 Budapest, Hungary
∗orzo.laszlo@sztaki.mta.hu

http://www.analogic.sztaki.hu

Abstract: Retrieving correct phase information from an in-line hologram is difficult as the object wave field and the diffractions of the zero order and the conjugate object term overlap. The existing iterative numerical phase retrieval methods are slow, especially in the case of high Fresnel number systems. Conversely, the reconstruction of the object wave field from an off-axis hologram is simple, but due to the applied spatial frequency filtering the achievable resolution is confined. Here, a new, high-speed algorithm is introduced that efficiently incorporates the data of an auxiliary off-axis hologram in the phase retrieval of the corresponding in-line hologram. The efficiency of the introduced combined phase retrieval method is demonstrated by simulated and measured holograms.

© 2015 Optical Society of America

OCIS codes: (090.1995) Digital holography; (100.2000) Digital image processing; (100.3010) Image reconstruction techniques; (100.5070) Phase retrieval.

References and links
In-line hologram records the intensity of the diffracted wave field, the diffractions of the complex objects, an in-line holographic architecture, introduced originally by Gabor, provides simple, high-resolution reconstructions, as the whole surface of the recorded hologram is exploited. As an applied light source and it is resistant against vibrations. Furthermore, an in-line setup can ensure an easily implementable measuring setup. It does not demand large coherence lengths from the application fields from metrology to digital holographic microscopy. These properties make holography efficiently utilizable in diverse applications. In principle, holography makes it possible to record the whole diffracted wave field of the measured objects. This way, either the reconstruction of the shape and refractive index distribution, or the volumetric reconstruction and 3D tracking of the measured objects become realizable. This way, either the reconstruction of the shape and refractive index distribution, or the volumetric reconstruction and 3D tracking of the measured objects become realizable. These properties make holography efficiently utilizable in diverse application fields from metrology to digital holographic microscopy.

1. Introduction

In principle, holography makes it possible to record the whole diffracted wave field of the measured objects. This way, either the reconstruction of the shape and refractive index distribution, or the volumetric reconstruction and 3D tracking of the measured objects become realizable. These properties make holography efficiently utilizable in diverse application fields from metrology to digital holographic microscopy.
jugate object (twin image) and zero order terms overlap the reconstructed object wave field [8]. Therefore, usually it is not easy to reconstruct the complex wave field of the measured objects from an in-line hologram.

To overcome this limitation of the in-line systems, Leith and Upatnieks [9] introduced the off-axis hologram recording architecture. In this setup, an appropriately tilted reference beam is applied. The tilted reference can be considered as a carrier wave modulation of the object wave field. This way, by the application of proper spatial frequency filtering the object term is separable from the zero order and conjugate object terms. However, this filtering considerably limits the bandwidth of the reconstruction. To achieve large resolution object reconstructions, appropriate optical magnification has to be applied before recording the hologram, that in turn compromises the reachable field of view. This is one of the main drawbacks of this setup. There has been made attempts that were able to partially ameliorate this pitfall using smart optical arrangements [10]. From the achievable resolution point of view the application of an in-line setup is preferred. Furthermore, an off-axis setup requires the application of light source of larger coherence lengths than that of an in-line setup, and the system is really sensitive to vibrations [11].

In the case of digital holography the holograms are recorded by area scan sensors (CMOS or CCD) and numerical algorithms are applied to realize the hologram reconstruction. In these cases, the resolution issues become even more critical, as the physical parameters of the applied digital sensors limit the area and resolution of the recorded holograms.

If an object has special, e.g. finite support property [12], there is some way to retrieve its whole complex wave field from a measured in-line hologram. Several numerical methods have been introduced [13–16] so far to solve this phase retrieval task. Unfortunately, the convergence speed of these algorithms is really low, especially in the case of systems of high Fresnel number \( N_f = a^2/L\lambda > 10 \), where \( a \) and \( L \) denote the characteristic size of the object and the reconstruction distance respectively, and this way usually only incomplete reconstructions can be accomplished.

Alternative measuring techniques were also introduced to ensure the exact reconstruction of the object wave field without the resolution compromise of the off-axis systems. These methods, like the phase-shifting interferometry [17], phase diversity based methods [18], transport of intensity algorithms [19], or aperture synthesis based techniques [20], require the recording of several holograms. Accordingly, they are not applicable, when moving objects are to be measured. Furthermore, these methods require complex measuring setups with intricate and expensive additional devices (e.g. implementing the required phase shifts), and usually demands accurate adjustment and precise calibration of the components. There are eloquent numerical phase reconstruction techniques, where special regularization terms are applied to retrieve the phase of the object from one or more measured holograms. However, these techniques require large number of iterations and the applied regularization constraints the amplitude and phase distribution of the reconstructions [21, 22].

We investigated whether the data of the off-axis hologram reconstruction can be exploited somehow in the phase retrieval of a corresponding in-line hologram. Our objective was to overcome the bandwidth restrictions of the off-axis systems and to ameliorate the phase retrieval of the in-line systems simultaneously. An analogous approach has been suggested earlier in a telescopic application [23], where the complete, but low resolution reconstruction of an image is applied to improve the efficiency of the phase retrieval of the same image at a considerably larger resolution. Similar approach was introduced lately, where the off-axis measurement data was applied for initialization of the phase retrieval process using a transport of intensity type algorithm [24]. We introduce here, an efficient phase retrieving method that integrate the off-axis measurement data into the phase retrieval process and combines the high achievable resolution.
of the in-line setups with the correct phase reconstruction capabilities of the off-axis holograms.

Our work is organized as follows: In Section 2, the operation, the most important properties
and limitations of the so far applied in-line hologram phase retrieval algorithms are delineated.
In Section 3, the special properties and the resolution constraints of the off-axis hologram re-
constructions are outlined. In Section 4, a simple rationale is provided why it is reasonable to
combine the two hologram reconstruction algorithms and a simple method is introduced that
incorporates the off-axis reconstruction data into the in-line hologram phase retrieval process.
The operation and efficiency of this extended algorithm is demonstrated on simulated and meas-
ured holograms. Finally, we conclude our results and discuss the required further development
and research directions.

2. Phase retrieval of in-line holograms

The propagation of the object wave field to the hologram recording plane defines the wave field
of the in-line hologram (see Eq. (1)).

\[ w_{ilz} = p_z \{ r(1 - o) \} = r_{il} - o_z = a_{ilz}(x,y)e^{i\phi_{ilz}(x,y)} \]  

(1)

For the sake of simplicity, the reference beam is regarded as a unit magnitude \( r = 1 \) plane
wave, while the object is characterized by its opacity function \( o = a_o(x,y)\exp(i\phi_o(x,y)) \). As
here, the recording of Gabor type holograms are assumed, therefore the objects cause only small
perturbation \( a_o(x,y) << 1 \) of the reference wave field. The free space wave field propagation
is denoted by \( p_z \). The hologram measures the intensity of this wave field.

\[ h_{il} = |w_{il}|^2 = |r_{il} - o_z|^2 = |r_{il}|^2 - r_{il}o_z - o_z\overline{r_{il}} + |o|^2 \]  

(2)

Reconstruction of the object is achieved simply by multiplying the in-line hologram with the
conjugated reference wave field (its amplitude and phase can be straightforwardly specified at
this plane) and propagating this reconstructed wave field back to the object plane. However, the
reconstructed object wave field is corrupted by the diffractions of the zero order and twin image
terms (see the third and fourth terms in Eq. (3)).

\[ o_{ilr} = p_{-z} \{ h_{il} \} = 1 - o - p_{-z} \{ \tilde{\sigma} \} + p_{-z} \{ |o|^2 \} \]  

(3)

As the measured in-line hologram defines the amplitude distribution of the hologram wave
field \( h_{il} = a_{il}^2(x,y) \) our goal is to unravel its phase distribution \( \exp(i\phi_{ilr}(x,y)) \). Exploiting
the finite support property of the objects this task becomes solvable. If an object has pure
amplitude or pure phase modulation, these information can be also efficiently applicable in
the phase retrieval process. The phase retrieval of in-line holograms are usually based on the
modified Gerchberg-Saxton (GS) algorithm [25].

The steps of the GS algorithm are detailed below.

1. The measured hologram defines the amplitude distribution of the object wave field at the
hologram plane, while its phase is initially set to zero.

\[ w_{il,0} = \sqrt{h_{il}} = a_{il}(x,y)e^{0} \]  

(4)

2. Simulate the propagation of this wave field to the object reconstruction plane (\(-z\)). In this
paper, the angular spectrum method is applied for the numerical simulation of the wave
field propagation, as this technique provides correct results for large Fresnel number
systems [26, 27].
\[ w_z = p_z \{ w_0 \} = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ w_0 \} e^{i k_z(u,v)z} \right\}, \tag{5} \]

where,

\[ k_z(u,v) = \begin{cases} \frac{2\pi}{\lambda} \sqrt{1 - \lambda u^2 - \lambda v^2}, & \text{if } (1 - \lambda u^2 - \lambda v^2) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{6} \]

Here, \( \mathcal{F} \) denotes Fourier transform, while \( \lambda, u \) and \( v \) the wavelength and the spatial frequencies of \( x \) and \( y \) directions respectively.

3. Only the twin image and zero order terms diffract beyond the object support. Therefore, if we eliminate them, their contribution to the wave field reconstruction decreases. Consequently, the reconstructed wave field beyond the support is replaced by the mean background value.

\[ 1 - \tilde{o}_n = \begin{cases} p - z \{ w_{il,n-1} \} & \text{if } (x,y) \in \chi_{il} \\ B & \text{otherwise} \end{cases} \tag{7} \]

The mean background value and the support are denoted by \( B \) and \( \chi_{il} \) respectively.

4. Propagate this estimation of the object wave field back to the hologram plane.

\[ \tilde{w}_{il,n} = p_z \{ 1 - \tilde{o}_n \} \tag{8} \]

5. Replace the estimated hologram wave field amplitude by the measured hologram intensity defined one, while the obtained phase estimation is preserved.

\[ w_{il,n+1} = a_{il}(x,y) \frac{\tilde{w}_{il,n}}{|\tilde{w}_{il,n}|} \tag{9} \]

Steps 2-5 can be iterated until the phase recovery converges.

Fienup provided a simple modification of the original GS algorithm (Hybrid input output algorithm) that prevents the otherwise frequently occurring stagnation of the algorithm and considerably speeds up its convergence [13, 28]. However, in the recording geometries applied and tested in this paper (i.e. large Fresnel number systems) no considerable improvement has been detected. Therefore here, the conventional GS algorithm is applied. There are other phase retrieval algorithms that are principally based on the reconstruction of the twin image [15, 29], but their efficient application is limited to systems of relatively small Fresnel number.

Nonetheless, whatever phase retrieval algorithm is applied so far, large number of iterations was required to reconstruct the correct wave field of the object.

Analyzing the algorithm, it can be recognized that the better the support separates the diffractions of the conjugate and zero order terms from the reconstruction of the object the faster will be the convergence. It is usually easy to satisfy in systems where the object is far from the hologram plane (systems of small Fresnel number). In this case the conjugated object (twin image) diffracts considerably beyond the support and therefore the algorithm swiftly converges. However, due to the small numerical aperture of the system the achievable resolution is confined [30].

From the same reason, it is important to find the tightest possible object support [31]. However, due to the overlapping twin image diffractions, it is not easy to find. Therefore, the support estimation is frequently updated, tightened during the phase retrieval process [32].
It can be recognized also, that mainly the high spatial frequency components of the twin image diffract beyond the support. Although, this property is accurate only, when Fraunhofer approximation (far field diffraction) is hold, but it can be regarded as a good estimation for diffractions of moderate and small distances.

During the GS iterations, the diffracted components beyond the support are eliminating quickly, while the remaining error is concentrated mainly within the support. Removal of this, apparently low spatial frequency bias takes a lot of time, as the support constraint cannot considerably help its removal during phase retrieval process. This seems to be the main reason of the apparent slow convergence of the so far applied phase retrieval algorithms. That is, the applied algorithm swiftly reconstructs the high spatial frequencies of the object (and that of the conjugated object, the twin image), while the low spatial frequencies remains noticeably biased. These properties of the phase retrieval are demonstrated in Fig. 1. First, the in-line hologram of a complex test object – that has both amplitude and phase modulation – is generated. Next we applied the GS algorithm on it to retrieve the phase of the object. It turns out that even after 25 iterations the phase reconstruction is still considerably distorted. It should be emphasized, that the residual bias remains appreciable mainly within the support.

The in-line hologram is simulated by the propagation of the object wave field to the hologram plane using the angular spectrum method. Only a quantized (8 bits), sensor sized, small section of this wave field intensity is regarded as the measured in-line hologram. During the phase retrieval process we apply a simple round shaped support that loosely covers the object. Here, our goal was not to optimize the speed of the phase reconstruction, but to show how the algorithm works if a roughly estimated, approximate support is applied and to provide a reference for the subsequent convergence speed comparisons.

Fig. 1. To test the efficiency of the GS phase retrieval algorithm a simulated in-line hologram is generated (b), using a test object that has amplitude and phase modulations (a). The phase retrieval after 25 iterations is shown in (d). During the phase retrieval process, the error of the reconstruction (c) is eliminating in the not supported region (red circle in (d)) swiftly, but remains substantial within it.
3. Off-axis hologram reconstruction

In the case of off-axis holography an appropriately tilted reference beam modulates the diffracted object wave field. The off-axis hologram is:

\[ h_{oa} = |r_{il} - o_z + r_{oa}|^2 = h_{il} - r_{oa}o_z - \overline{r_{oa}}o_z + r_{il}\overline{r_{oa}} + |r_{oa}|^2 + r_{oa}r_{il}, \]  

(10)

where the reference beam is \( r_{oa} = a_{oa} \exp(i2\pi(\alpha_{ox}x + \alpha_{oy}y)) \). This way, in the frequency space the object, the conjugate object and the zero order terms become separable.

\[ H_{oa} = \mathcal{F}\{h_{oa}\} = \mathcal{F}\{h_{il}\} + \mathcal{F}\{|r_{oa}|^2\} + a_{oa}\delta(\alpha_{ox} - u, \alpha_{oy} - v)\mathcal{F}\{r_{il} - o_z\} + a_{oa}\delta(u - \alpha_{ox}, v - \alpha_{oy})\mathcal{F}\{r_{il} - o_z\}. \]

(11)

The reconstruction can be simply fulfilled by the application of a proper band pass filter \([33]\). The size of the applicable filter is constrained by the extent of the zero order term (first term in Eq. (11)), the bandwidth of the off-axis setup, and finally the tilt parameters of the applied off-axis reference beam. The reconstruction is based on the object term of the off-axis hologram:

\[ H_{oa,o} = \begin{cases} \frac{1}{r_{oa}}H_{oa}(u,v)\delta(\alpha_{ox} - u, \alpha_{oy} - v) & \text{if } (u,v) \in \chi_{oa,c} \\ 0 & \text{otherwise}, \end{cases} \]

(12)

where \( \chi_{oa,c} \) denotes the applied spatial frequency filter. It should be noted that the first term can overlap with the forth term in Eq. (11) and this way slightly contaminates the object term estimation. That is,

\[ H_{oa,o} \simeq \begin{cases} \mathcal{F}\{r_{il} - o_z\} & \text{if } (u - \alpha_{ox}, v - \alpha_{oy}) \in \chi_{oa,c} \\ 0 & \text{otherwise}. \end{cases} \]

(13)

The applied filter considerably limits the bandwidth and so the resolution of the reconstruction. The properties of the off-axis reconstruction are demonstrated in Fig. 2. It can be recognized that although a correct reconstruction of the object phase is achieved, the high resolution details are missing.

4. Combined phase retrieval

As we demonstrated in the previous sections; the numerical reconstruction of the off-axis hologram is able to recall the low spatial frequency parts of the object wave field, while the phase retrieval of the in-line hologram prefers the fast recovery of just the high spatial frequency components. It seems to be really promising if we combine these methods. Here, we show a simple and straightforward extension of the modified GS algorithm, that incorporates the off-axis reconstruction data in the phase retrieval process of the in-line hologram.

4.1. Proposed algorithm

The proposed method simply introduces a new step in the conventional GS algorithm. After step 4, the low spatial frequencies of the estimated hologram wave field spectrum are replaced by the off-axis object term defined ones.

\[ \tilde{\omega}_{iloa,n} = \mathcal{F}^{-1}\begin{cases} \mathcal{F}\{r_{il} - o_z\} & \text{if } (u - \alpha_{ox}, v - \alpha_{oy}) \in \chi_{oa,c} \\ \mathcal{F}\{\tilde{\omega}_{il,n}\} & \text{otherwise} \end{cases}. \]

(14)

This way the exact, low spatial frequency data of the off-axis object term are correctly integrated into the in-line phase retrieval.
Fig. 2. The reconstruction of an off-axis hologram (a) is based on spatial frequency filtering. The applied filter – denoted by a red circle in the inverted frequency spectrum of the off-axis hologram (b) – limits the reconstruction bandwidth, and eliminates the high frequency details of the reconstruction (c). The inverted absolute error of the reconstruction (d) demonstrates the low pass filtering property of the reconstruction and the biasing effects of the components of the zero order term that overlaps the filtered region (arrows).

4.2. Simulation results

To test the algorithm performance we investigated its operation and phase retrieval capabilities on simulated and measured holograms. There are two parameters that have to be adjusted to ensure the proper integration of the two different hologram measurements: The first one is the tilt angle of the off-axis reference beam. The second one is the relative intensity of the two different hologram reconstructions. The tilt angle of the reference beam in the off-axis system can be estimated either from the position of the reconstructed object or from an empty off-axis hologram, where the position of the DC part of the ‘object’ term in the frequency space defines it correctly. The relative weights of the two holograms can be determined by comparing the DC term of the in-line hologram intensity with that of the object term of the off-axis hologram. Otherwise, measuring the ratio of some extended dark and bright regions in the different hologram reconstructions can corroborate the validity of these estimations.

Figure 3 demonstrates that the proposed algorithm can properly reconstruct the object using the simulated in-line and off-axis holograms.

It can be recognized that the speed of the convergence is high. After a few iterations, the error of the reconstruction becomes order of magnitude smaller than it is achieved by the earlier approaches. It worth mentioning that the fast and correct phase retrieval is achieved with the application of rough support estimation in a system that can be characterized by a large Fresnel number ($N_f = a^2/L\lambda > 50$). We can appreciate the efficiency of the proposed algorithm, comparing its convergence speed with that of the conventional phase retrieval algorithm (see Fig. 1(c)).
4.3. Convergence properties

The same way as the GS algorithm (Fig. 4(a)), the speed of convergence of the introduced algorithm increases with the tightening of the applied supports (see Fig. 4(b)).

The convergence speed of the introduced algorithm depends on the bandwidth of the off-axis filter too. Increasing the size of the filter, accelerated convergence is expected. However, as the filter bandwidth is increasing, it overlaps more with the components of the zero order terms. This overlap impures the estimation of the off-axis object term. This way it contaminates the off-axis reconstruction (tagged with arrows in Fig. 2(c)), and the combined reconstruction (see...
Fortunately, the zero order term corresponds to the measured, known in-line hologram (see in Eq. (10)).

\[
\tilde{h}_{oa} = h_{oa} - h_{il} = -r_{oa}\bar{a}_z - r_{oa}\bar{r}_{oa} + |r_{oa}|^2 + r_{oa}\bar{r}_{il}.
\] (15)

This way we can remove its overlapping parts completely from the object term of the off-axis hologram.

\[
\tilde{R}_{out,o} = \begin{cases} 
\mathcal{F}\{r_{il} - a_z\} & \text{if } (u - \alpha_{oa,x}, v - \alpha_{oa,y}) \in \chi_{oa,c} \\
0 & \text{otherwise.}
\end{cases}
\] (16)

Applying this improved off-axis object term estimation in the algorithm (see in Eq. (14)), the expected filter size dependence of the convergence speed (see Fig. 5(b)).

4.4. Measured holograms

The performance of the proposed algorithm is tested on measured holograms too. A conventional, simple off-axis measuring setup is applied, that is able to record both the in-line and the off-axis holograms (see in Fig 6(c)). In this setup a fiber coupled laser and a fiber splitter is applied to ensure both the illumination and the off-axis reference beams, while afocal optics provides the required moderate (5 times) magnification in the system. This way, the object and sensor positions become more or less freely adjustable [34]. An Air Force Test target is applied as a test object. The wavelength of the applied laser and the pixel size of the sensor was 650nm and 3.5\(\mu\)m respectively. Only a 1024x1024 pixel section of the recorded hologram is processed. The reconstruction distance was 9mm.

The processing of the measured data corroborates the simulation results. That is, the in-line hologram reconstruction is biased considerably by the twin image diffractions (Fig. 6(g)), and the off-axis hologram reconstruction shows reduced resolution (Fig. 6(h)). Using the proposed combined algorithm all the twin image noise of the in-line reconstruction can be eliminated, while its high resolution details are preserved (Fig. 6(i)). In the algorithm only rough estimation of the object support is applied, that was defined by using the off-axis reconstruction. Due to the uneven illumination and the large noise of the experimental setup, the support was fitted by hand (see in Fig. 6(d)). However, in the case of relatively small, compact objects, the support estimation can be obtained automatically. The algorithm converged within a few iterations (Fig. 6(f)). The achievable resolution is confined by the applied optical setup (\(\sim 5\mu m\)). Here we aimed only to demonstrate the operation and applicability of the algorithm and not to optimize the resolution of the hologram recording setup itself. To realize higher resolution reconstructions more sophisticated off-axis measuring setup has to be used [35]. Due to the uneven illumination and the extended and scattered object, the conventional phase retrieval algorithm, using the
Fig. 6. By the application of a simple experimental setup (c) both the in-line (a) and off-axis (b) holograms of the test objects can be recorded. Using the measured holograms, a rough support estimation (d) and the employed off-axis object term (the applied spatial frequency filter is denoted by a red rectangle in the inverted spectrum of the in-line compensated (Eq. (16)) off-axis hologram, we can achieve reconstruction (i), that eliminates the twin image noise of the in-line hologram reconstruction (g) and provides considerably higher resolution than that of the off-axis reconstruction (h). Furthermore the introduced combined algorithm shows swift convergence (f).

rough support estimation, did not converged properly. Conversely, the introduced algorithm did not require special fine tuning neither the object support, nor the spatial frequency filter (see in Fig. 6(e)) of the off-axis hologram reconstruction. The introduced algorithm appears to be simple, fast and robust.

5. Conclusion

A new, efficient phase retrieval method is introduced. By the incorporation of an off-axis object term based constraint into a conventional phase retrieval algorithm, both the convergence speed and the achievable resolution is improved considerably. We have demonstrated on simulated and measured holograms that in a few iterations the phase of the object wave field can be
retrieved correctly even in large Fresnel number systems.

The proposed method can be efficiently applied for Gabor type in-line holograms, where the objects finite support constraint is hold. The introduced algorithm is simple, robust and works suitably even with rough object support estimations.

Although we are able to reconstruct the phase of the in-line hologram correctly we still need two different recorded holograms that can considerably limit the applicability and flexibility of the proposed method. However, it seems to be possible to efficiently separate the zero order, in-line term and the object term within an off-axis hologram. This way it is possible to apply the algorithm for a single recorded off-axis hologram. Details of this segmentation method and the achievable resolution of the reconstructions are subjects of a consecutive paper.

Further investigations are necessary to determine, how the accuracy of the phase retrieval depends on the errors of the different hologram measurements. It is also a question, how to modify or apply the introduced phase retrieval algorithm efficiently, when several different objects are to be reconstructed from a single recorded hologram.

Acknowledgments

This work was supported by the Hungarian National Office for Research and Technology (NKTH 1981822A) project entitled “Water Biology Digital Holographic Microscope (DHM) as an early warning environmental system” and by the framework of Green Industry Innovation Programme of the EEA and Norwegian Financial Mechanisms (grant number: HU09-0078-A1-2013).