

# Improving speed and accuracy of phase retrieval applying two in-line hologram recordings.

László R. Orzó<sup>\*</sup>, Márton Zs. Kiss, Ákos Zarándy

Institute for Computer Science and Control, Hungarian Academy of Sciences, 13-17 Kende Street,  
Budapest, HU-1111

## ABSTRACT

In the case of in-line digital holographic microscopes the retrieval of the phase information is indispensable to remove the twin image noise and to characterize the thickness and refractive index distribution of the measured objects. The existing phase reconstructing algorithms, especially, when the goal is to reach high resolution, suffer from really slow convergence. This can prevent their practical applicability. We introduce here a simple method that ensures accurate, high resolution and fast phase retrieval from two holograms recorded in different distances. The so far frequently applied phase retrieval algorithms converge with tolerable speed if the Fresnel number of the system is small enough. This is the consequence that in this case the diffraction of the twin image spreads substantially over the applied object support and thus the support constraint can provide sufficient amount of data for the efficient phase retrieval. However, if the recording measurement Fresnel number is small, the finite aperture of the hologram limits the proper recording of the high spatial frequencies of the object diffraction. In this case the convergence of the phase retrieval will be really fast, but inaccurate. Therefore, we propose to record two different holograms with small and large Fresnel numbers accordingly. From the first one we can reconstruct the low spatial frequencies of the object at a high rate. Subsequently, we can use this partial reconstruction to speed up the reconstruction of the so far not correctly recalled high spatial frequencies, by applying the second hologram.

**Keywords:** Phase retrieval, Gerchberg-Saxton algorithm, Fresnel number, Convergence, Accuracy, Numerical Aperture

## 1. INTRODUCTION

Digital Holography, and especially Digital Holographic Microscopy [1] is widely applied lately, as it can provide simple, robust and cheap measuring architectures [2] with high resolution imaging capabilities [3, 4]. Using this technology, it is also possible to numerically compensate the aberration of the applied optics [5]. This technology can be applied also to implement volumetric imaging, where the hologram of large sample volume is measured [6] and later the image of different objects within the observed volume are numerically reconstructed from this hologram [7]. This way, we can overcome the small depth of focus limitation of the conventional microscopes and the high speed examination of large, sparse sample volumes becomes achievable. Earlier, we applied this technology in several water monitoring applications. Using digital holography, it becomes possible to reconstruct not only the intensity, but also the phase distribution of the measured objects. The phase information gained is essential to estimate the shape and refractive index distribution of the examined sample objects.

Gabor type [8] in-line holographic architectures are frequently applied, as these can provide simple, robust measuring setups. In this case, the illuminating light provides the reference beam and this way this architecture is relatively insensitive to vibrations and does not require the application of a light source with large coherence length. However, the image of the reconstructed object will be biased by the diffractions of the so-called conjugated object (twin image) and the first order terms. These can be reduced or eliminated using appropriate digital phase retrieval or twin image removal algorithms [9, 10, 11]. However, when high resolution reconstruction is needed, these numerical algorithms appears to be really slow.

<sup>\*</sup>orzo@sztaki.mta.hu; phone 36 1 279-6000/7182; fax 36 1 209-5264; analogic.sztaki.hu

Although the corresponding computer technology provided high processing power and the fact that these algorithms are frequently implemented and accelerated by high speed parallel computing hardware (as stream processors (GPGPU) [12]), the phase retrieval algorithms cannot achieve the demanded real time, or close to real time processing speeds.

There are alternative, non-holographic measuring architectures that aim the reconstruction of the phase of the measured object, like the transport of intensity equation [13] or phase shifting interferometry [14] based approaches. However, the first approach, as it is based on paraxial approximation cannot efficiently apply in high resolution measurements of objects that have amplitude and phase modulation too. While, to the implementation of the second measuring setups require either special actuators that can solve the sub wavelength size shifting of the reference wave field.

To avoid the bias of the overlapping twin image and zero order terms, Leith and Upatniek [15] introduced the so called off-axis holographic architectures, where an additional, tilted reference wave is applied. This tilt can ensure the separation of the different terms. However, in the off-axis measuring setups the application of a large coherence length illumination is required and the measuring setup is really sensitive to vibrations. In the digital implementations of these measuring setups the object term can be easily separated in the frequency domain. Nonetheless, only a small part of the sensor aperture can be used *de facto* for the object reconstruction that decreases the space bandwidth product of the setup [16]. Earlier, we introduced a method [17], where a low resolution reconstruction of an off-axis measurement can be efficiently applied to speed up the phase retrieval of the corresponding, high resolution in-line measurement.

In this paper, we propose a new method to accelerate the convergence of the conventional phase retrieval algorithms. The algorithm applies two in-line holograms measured from different distances. There were some earlier attempts to apply two holograms to achieve phase retrieval [18], but these are quite different from our approach. First, we outline the relationship between the achievable accuracy of the hologram reconstruction and the actual parameters of the corresponding measurement setup. Next, we examine how the measurement setup parameters define the speed and accuracy of the phase retrieval using some conventional algorithms. Based on these results we propose the application of an additional in-line hologram recording and introduce a simple modification of the existing algorithm to achieve high speed accurate phase reconstruction. We show how the algorithm can be optimized further from the points of view of the required computation. Finally, we try to show that this kind of approach is essential, when the goal is to get efficient, high resolution and accurate phase retrieval of all the objects within the measured hologram aperture.

## 2. LIMITATIONS OF THE HOLOGRAM RECONSTRUCTION ACCURACY

In digital holography, only those parts of the diffracted wave field are recorded that are within the applied sensor aperture. Therefore, if we would be able to fulfill perfect phase retrieval by reconstructing the complex amplitude distribution of the hologram, the missing, not recorded components will still bias the reconstruction. As we aim to achieve phase retrieval of all the objects within the measured hologram, we have to estimate the aperture caused limitations.

We made simple simulations to estimate how on the numerical aperture of the hologram recording determines the reconstruction error. We used in the simulation a simple test image, which have amplitude and phase modulation too (see in Fig. 1 a, and b) and used angular spectrum method to simulate the wave field propagation. We used appropriate padding to avoid numerical errors of the propagation [19].

We digitize (e.g. 8 bits) and crop different size of apertures from the resulting, complex diffracted wave field to simulate a hologram a recording with perfectly retrieved phases. From this simulated complex hologram recording we reconstruct the original object by the reverse propagation of the wave field. This way, we can determine the error caused by the finite aperture of the hologram. For example, we simulated 512x512 size hologram recordings (the pixel size and the wavelength were 1.67  $\mu\text{m}$  and 0.532  $\mu\text{m}$  accordingly) with different propagation distances.

In the next figure we can see the amplitudes of the simulated holograms from distances of 2 mm and 4 mm (see in Fig. 1. c, and e, accordingly). We can recognize that the reconstruction error is large for large sensor object distance and small for small distances.

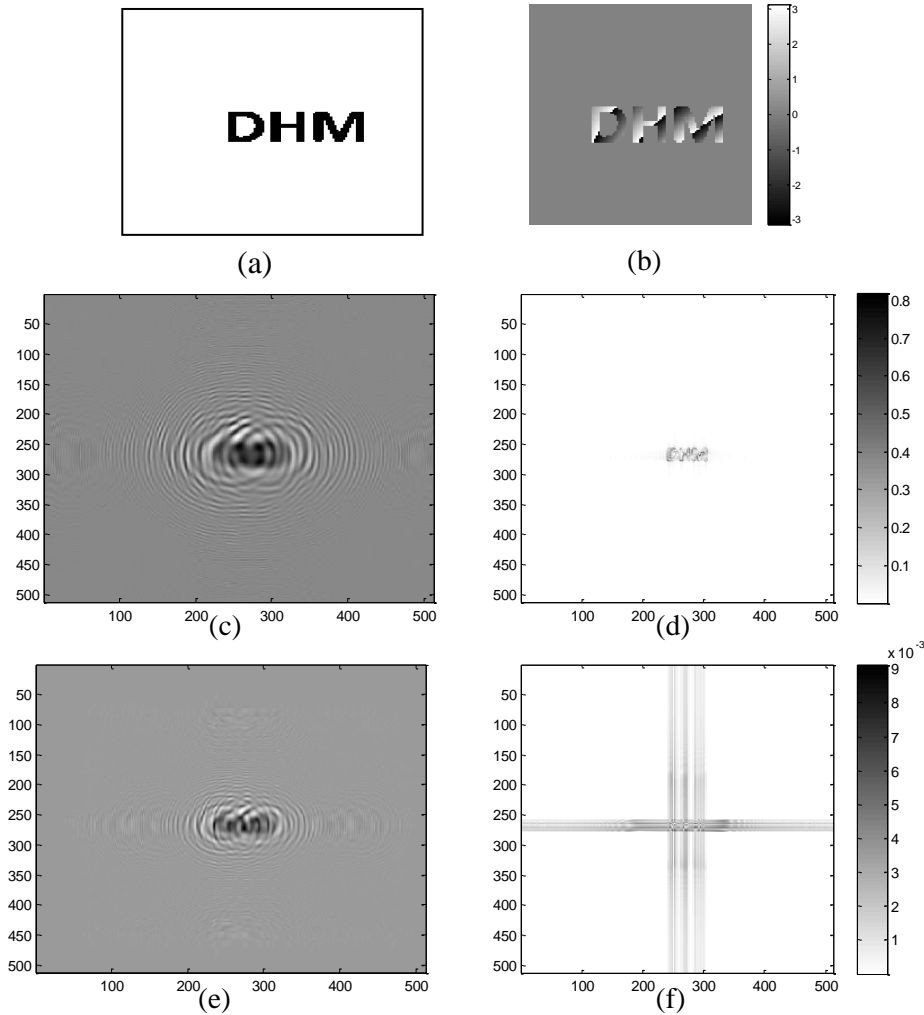


Figure. 1, Amplitude (a) and phase (b) modulation of the test object applied in the simulations. The simulated hologram amplitude of the simulated complex hologram at 4mm (c) and 2mm (e) distances. The corresponding absolute reconstruction errors are shown in (d) and (f).

This is reasonable, considering the assumed reason of the reconstruction error. We must emphasize, that this measured reconstruction error can be unacceptably large for relatively small recording distances as well.

Furthermore, we can show that the reconstruction error is caused by the missing, or biased high spatial frequency components. We compared the low pass filtered reconstruction with the correspondingly filtered original image. The applied low pass filter simply the half of the maximal frequencies were transferred. We can observe that the low spatial frequencies are correctly reconstructed and caused really a small reconstruction error (see in Fig. 2).

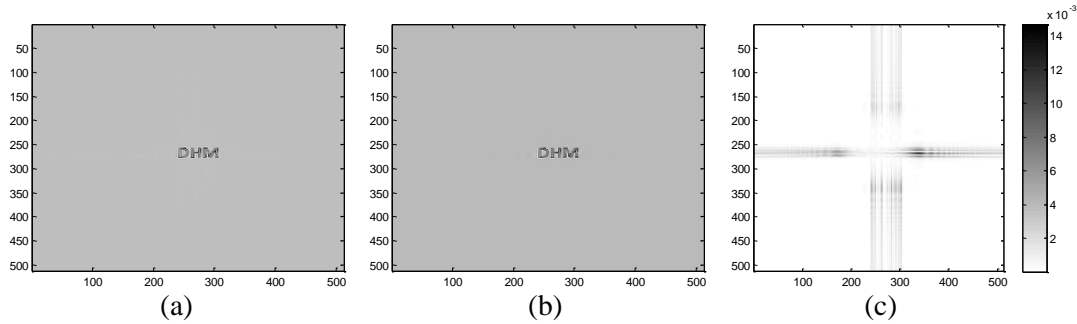


Figure. 2, Amplitude of the filtered test image (a) and filtered amplitude of the reconstruction (b). The reconstruction in these frequencies shows very small absolute error (c), comparing to the not filtered corresponding reconstruction (see in Fig. 1, c).

These results confirm that the hologram aperture limits the correct reconstruction. It seems to be especially restrictive close to the hologram boundaries, where the correct reconstruction required numerical aperture can be ensured only by holograms measured from a really small distance. However, low resolution reconstructions can be achieved even from much larger distance hologram recordings.

### 3. PARAMETERS LIMITING THE PERFORMANCE OF THE PHASE RETRIEVAL ALGORITHMS

In the reconstruction of an in-line hologram the reconstructed object is biased by the diffractions of the conjugate object (twin image) and the zero order terms (see in Fig. 3.).

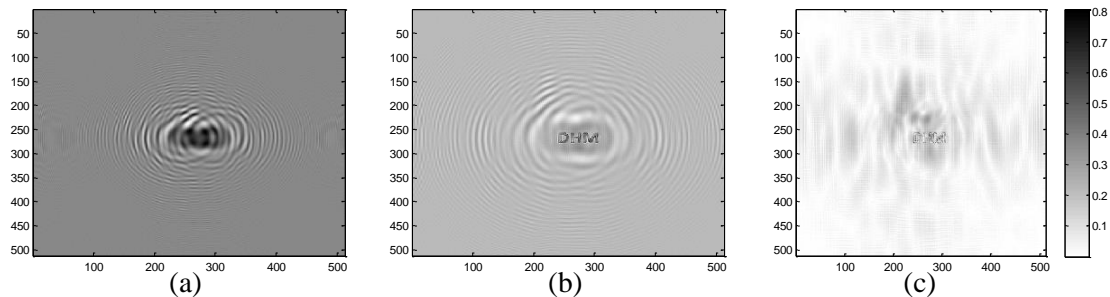


Figure. 3. The reconstructed object (b) from simulated in-line hologram (a) has a large reconstruction error (c) due to the overlapping zero order and twin image noises. Object distance, pixel size and the simulation wavelength were 4 mm, 1.67  $\mu\text{m}$ , and 0.532  $\mu\text{m}$  accordingly.

To remove these terms, we have to retrieve the lost phase distribution of hologram. To achieve this, numerical algorithms are applied. Here we consider the frequently used, modified Gerchberg-Saxton and (its nonlinear filtering-based, accelerated form) the Fienup's Hybrid Input-Output algorithms. These algorithms iteratively apply the hologram amplitude constraint [2] and a kind of object constraint to reconstruct the hologram phase distribution. The hologram amplitude is known from the measured hologram intensity distribution. If the measured objects have pure amplitude or phase modulation, this property can be efficiently used as an object constraint and the phase retrieval is usually swift and accurate. However, usually this is not the case. Most frequently, the finite support property of the measured objects is utilized. Here, the algorithms keep the amplitude and phase modulation of the reconstructed object within the known support, otherwise it is replaced by the estimated background value. The speed of the convergence of these algorithms depends on the recording setup geometry. Systems having a relatively small Fresnel number the convergence of the phase retrieval are usually fast. It seems reasonable, as the finite support constraint provides additional information for the phase retrieval, when a considerable part of the twin image diffractions is spread beyond the support. We made simulations to estimate the effect of different parameters of the hologram recording geometry on the speed of

convergence and the reconstruction error. Here we applied the same test image as we used in the previous section and a loosely defined object support was applied. The error of the reconstructions was measured only within the supported area. It needs to be mentioned that a more precise phase retrieval is achievable applying tighter and more accurate object support. Therefore, the phase retrieval is frequently improved by updating and refining the applied support, according the actual result of the phase retrieval process [20].

We regard the phase retrieval correct if we were able to achieve reconstruction with no visually perceptible error. In this case, the root mean square error is less than one percent and the maximal error is also small. Otherwise the phase reconstruction is biased, or it still has not reached the proper solution.

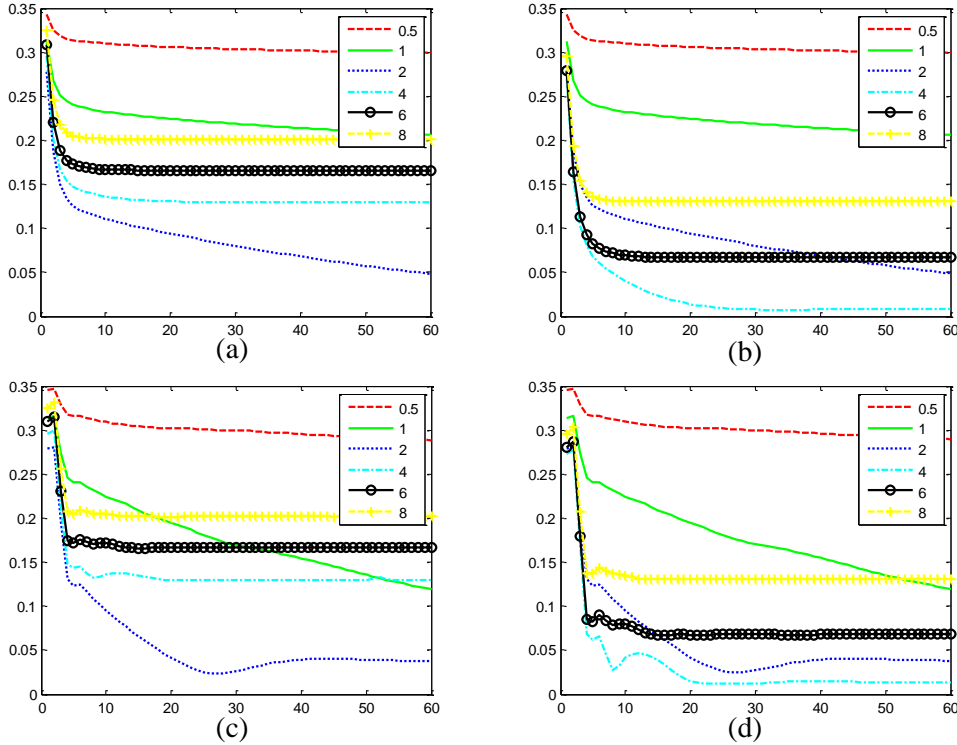


Figure 3. The evaluation of the root mean square error of the phase retrieval of holograms recorded from different distances (0.5, 1, 2, 4, 6, 8mm) using the modified Gerschgberg-Saxton (hologram sizes were 512x512 (a) and 1024x1024 (b)) and Fienup's HIO algorithms (hologram sizes were 512x512 (c) and 1024x1024 (d)) accordingly.

The root mean square error evaluation can be used not only to estimate the speed of convergence, but also to characterize the achievable accuracy of the reconstruction. The Fresnel Numbers ( $FN = a^2 / \lambda d$ ), corresponding to the simulated distances 0.5, 1, 2, 4, 6, 8 mm were 67.1, 33.5, 16.8, 8.4, 5.5, 4.4 accordingly. We can recognize from these data that for small recording distances (large FN) both algorithms show really slow convergence, while for large distances (small FN) they show rapid convergence, but really pure accuracy. This property is the consequence of the reconstruction characteristics shown in the previous section. There is an intermediate distance range, where acceptable speed and relatively accurate reconstruction can be achieved. However, their performance depends considerably on the simulated hologram size, the actual numerical aperture.

In practical applications we aim the accurate, high resolution phase retrieval not only for the objects in the center of the hologram, where the required aperture is available, but for objects within the majority of the measured surface.

We found, that the phase retrieval is either fast and inaccurate or accurate, but unacceptably slow. In the second case, even after hundreds of iterations, the algorithm still does not reach the desired, correct solution. We examined how the applied algorithm reconstructs the low spatial frequency components of the objects as well. Astonishingly, we found -

similar to the results shown in Fig. 2, - that the phase reconstruction of the low spatial frequency components converges to the accurate solution (see later in Fig. 4. e and Fig. 5. b).

#### 4. PROPOSED METHOD

How can we overcome these limitations of the phase retrieval? We found that the reason of the slow convergence of large Fresnel number systems is, primarily, the slow convergence of the low spatial frequency components. The slow convergence of these terms is the consequence that in their regards, the application of the support constraint is pointless. On the other hand, we can define these low septal frequency components from an additional measurement easily.

Therefore, we propose to make an additional in-line hologram measurement and apply it to ameliorate the phase retrieval process. For the high-resolution reconstruction, we have to ensure high numerical aperture on the majority of the hologram surface. This can be ensured by small recording distance hologram (large Fresnel number).

Furthermore, we need a distant hologram, which can provide the low spatial frequency reconstruction at high speed.

In the proposed algorithm we apply conventional phase retrieval for the distant hologram. It will converge in a few steps and use these results to initiate the phase retrieval of the close (high Fresnel Number) hologram. The achieved results can be seen in the next figure (see in Fig. 4.).

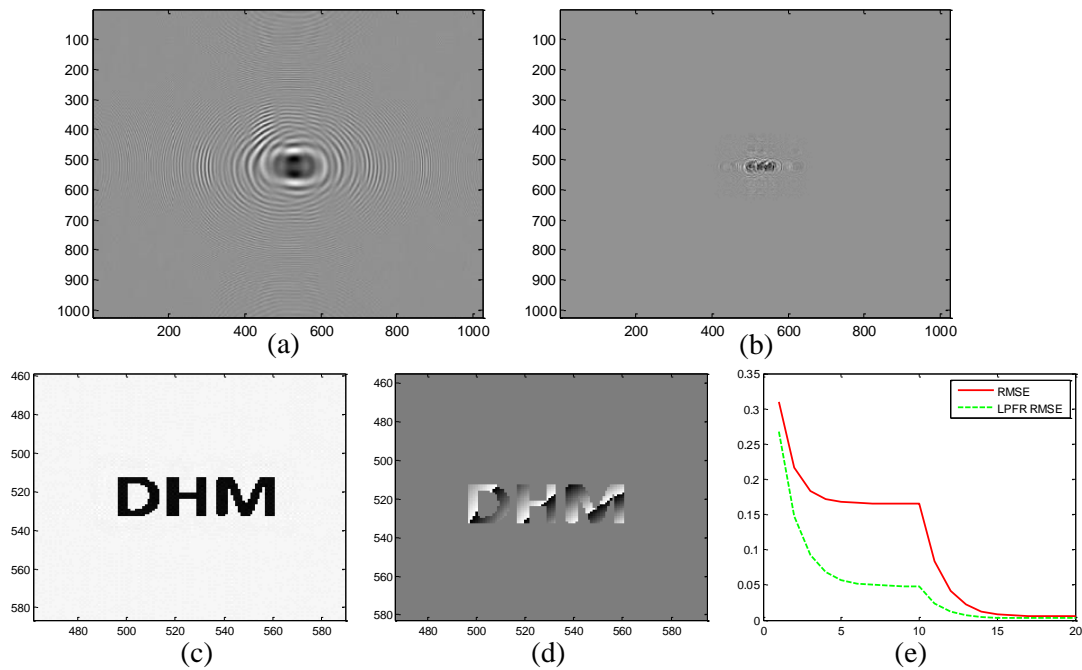


Fig. 4. The proposed combined phase retrieval applies two holograms, recorded from different distances (8mm (a) and 2 mm (b) accordingly). The reconstruction achieved in two consecutive 10 steps processing stages as it can be recognized the root mean square error evaluation plot (e) and results in correct amplitude (c) and phase (d) reconstruction in 20 steps of iteration (e). The solid line (in e) shows the RMSE evaluation of the phase retrieval process, while the dashed line denotes the evaluation of the low pass filtered component.

We were able to achieve considerable improvement, from the points of view both speed and accuracy..

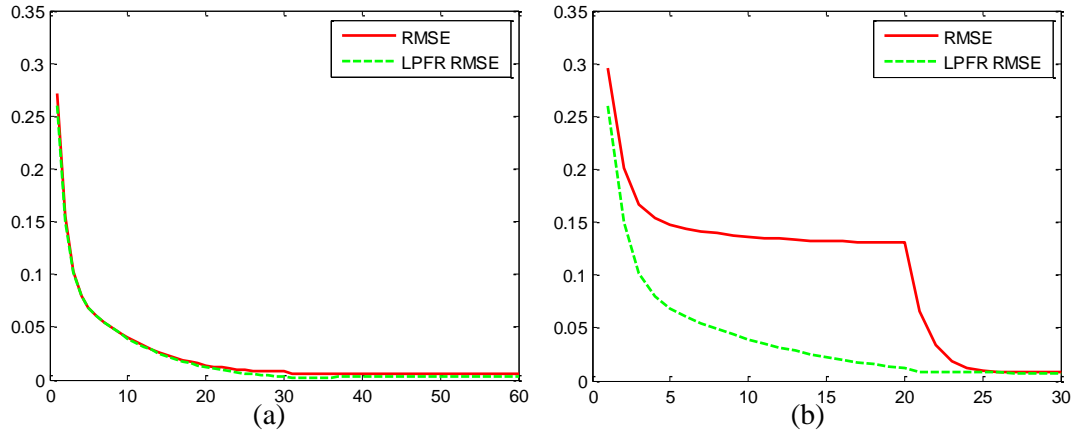


Figure. 5, Evaluation of the RMSE of the GS phase retrieval form 1024x1024 (a) and 512x512 (b) size of holograms using first 4 mm, and later 1 mm recording distances. The solid (red) lines denote the evaluations of the root mean square error during reconstruction, while the dashed (green) line denote the evaluation of the error of the low pass filtered component.

We can recognize that for 4 mm distance if the object has adequate aperture size (see in Fig. 4. a) the phase reconstruction is fair. However, it would fail, when the object has a smaller aperture (see in Fig. 4. b, and e, the first stage). However, the low pass filtered component show additional convergence (denoted by the dashed line in Fig. 4.). These low spatial frequency data can be applied efficiently to the phase retrieval from a much smaller distance recorded hologram (extremely large Fresnel number) as well.

We can see that the speed of the convergence is a bit slower for the low spatial frequencies (comparing the solid and dashed lines, in Fig. 5. b), which is caused by probably, that the error of the high spatial frequencies masks that of the low spatial frequencies. In practical applications, where the convergence is measured by the change of the result of the consecutive steps, the application of low pass filtering is needed.

## 5. OPTIMIZATION OF THE PHASE RETRIEVAL PROCESS

For efficient implementation of the phase retrieval, we tried to optimize the convergence of the algorithms. We can recognize that in the first stage of the proposed method the phase retrieval of the low spatial frequencies of the measured objects is aimed. To achieve this, we do not really need the same resolution reconstruction as in the final reconstruction. Applying a reduced resolution the required computation will be considerably faster. We can measure a lower resolution hologram or proper sub-sampling a high resolution one. This way not only the required number of iterations decreases during the phase retrieval process, but the first stage of the phase retrieval will be accelerated as the required FFT operations are done on a smaller resolution image. In the next figure (see in Fig. 6), we demonstrate that we can achieve the same reconstruction accuracy, using a smaller resolution distant hologram.

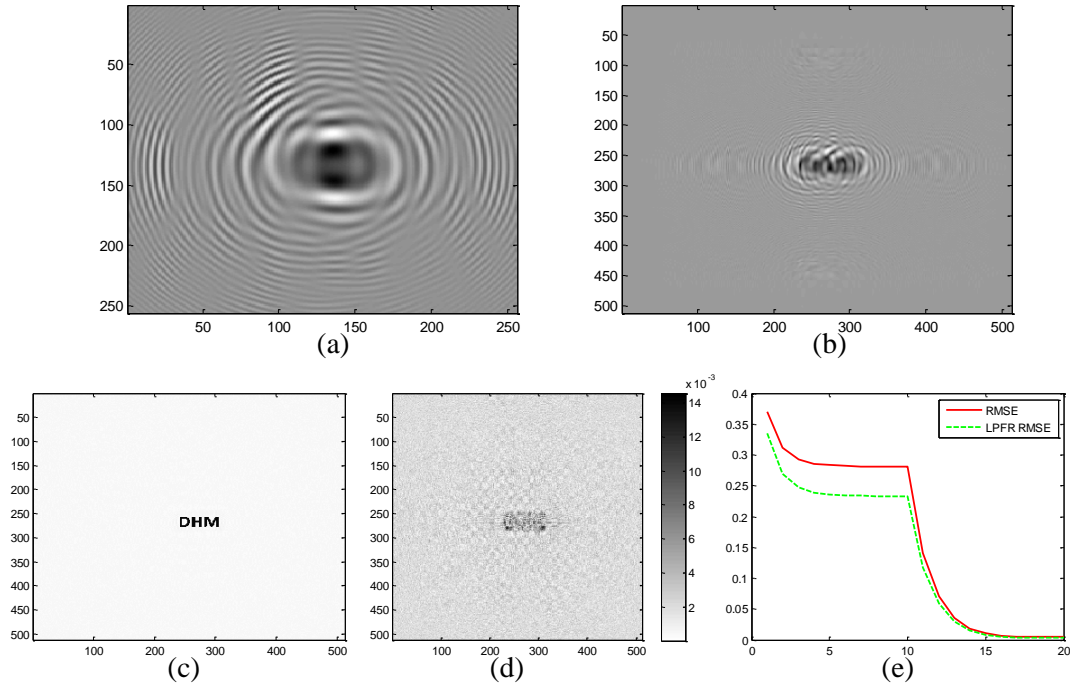


Fig. 6. Phase retrieval using (a) a low resolution (256x256, pixel size 3.34 $\mu$ m) hologram (a) recorded from 12mm distance and from a high resolution (512x512, pixel size 1.67 $\mu$ m) one recorded (b) from 2mm. The algorithm converges rapidly (e) to a reconstruction (c) that has small absolute error (d).

## 6. CONCLUSIONS:

We considered the aperture requirement of the high-resolution hologram reconstruction and estimated the parameters of the recorded holograms, where the support constraint based phase retrieval algorithms can be efficiently utilized. We found that these requirements are contradictory. This way, accurate, high resolution phase retrieval can be achieved only for a small part of the measured hologram, where the required aperture is available. We show that by the application of two properly set hologram recordings we can alleviate the above mentioned conflict. We proposed a simple algorithm that improves considerably both the accuracy and speed of the phase retrieval and ensure the precise reconstruction of the phase distribution in a much larger area of the hologram.

## REFERENCES

- [1] Hariharan, P., [*Optical Holography: principles, techniques, and applications*0.1em], no. 20, Cambridge Univ Pr (1996).
- [2] Mudanyali, O., Tseng, D., Oh, C., Isikman, S., Sencan, I., Bishara, W., Oztoprak, C., Seo, S., Khademhosseini, B., and Ozcan, A., "Compact, light-weight and cost-effective microscope based on lensless incoherent holography for telemedicine applications," *Lab on a Chip* **10**(11), 1417 (2010).
- [3] Bishara, W., Su, T., Coskun, A., and Ozcan, A., "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution," *Optics express* **18**(11), 11181–11191 (2010).
- [4] Göröcs, Z. and Ozcan, A., "On-chip biomedical imaging," *IEEE reviews in biomedical engineering* **6**, 29–46 (2013).
- [5] Colomb, T., Cuche, E., Charrière, F., Kühn, J., Aspert, N., Montfort, F., Marquet, P., and Depeursinge, C., "Automatic procedure for aberration compensation in digital holographic microscopy and applications to specimen shape compensation," *Applied optics* **45**(5), 851–863 (2006).



- [6] Moon, I. and Javidi, B., “Three-dimensional identification of stem cells by computational holographic imaging,” *Journal of The Royal Society Interface* **4**(13), 305 (2007).
- [7] Orzó, L., Göröcs, Z., Fehér, A., and Tökés, S., “In-line hologram segmentation for volumetric samples,” *Applied Optics* **52**(1), A45–A55 (2013).
- [8] Gabor, D., “A new microscopic principle,” *Nature* **161**(4098), 777–778 (1948).
- [9] Fienup, J., “Phase retrieval algorithms: a comparison,” *Applied Optics* **21**(15), 2758–2769 (1982).
- [10] Yang, G. Z., Dong, B. Z., Gu, B. Y., Zhuang, J. Y., and Ersoy, O. K., “Gerchberg–Saxton and Yang–Gu algorithms for phase retrieval in a nonunitary transform system: a comparison,” *Applied optics* **33**(2), 209–218 (1994).
- [11] Koren, G., Polack, F., and Joyeux, D., “Iterative algorithms for twin-image elimination in in-line holography using finite-support constraints,” *Journal of the Optical Society of America A* **10**(3), 423–433 (1993).
- [12] Orzó, L., Göröcs, Z., Szatmári, I., and Tökés, S., “GPU implementation of volume reconstruction and object detection in digital holographic microscopy,” in [*Proceedings of IEEE Conference on Cellular Nanoscale Networks and Their Applications (CNNA)*], 1–4, IEEE (2010).
- [13] Gureyev, T. and Nugent, K., “Rapid quantitative phase imaging using the transport of intensity equation,” *Optics Communications* **133**(1-6), 339–346 (1997).
- [14] Yamaguchi, I. and Zhang, T., “Phase-shifting digital holography,” *Optics letters* **22**(16), 1268–1270 (1997).
- [15] Leith, E. N. and Upatnieks, J., “Reconstructed wavefronts and communication theory,” *JOSA* **52**(10), 1123–1128 (1962).
- [16] Stern, A. and Javidi, B., “Space-bandwidth conditions for efficient phase-shifting digital holographic microscopy,” *JOSA A* **25**(3), 736–741 (2008).
- [17] Orzó, L., “High speed phase retrieval of in-line holograms by the assistance of corresponding off-axis holograms,” *Opt. Express* **23**, 16638–16649 (Jun 2015).
- [18] Zhang, Y., Pedrini, G., Osten, W., and Tiziani, H. J., “Reconstruction of in-line digital holograms from two intensity measurements,” *Optics letters* **29**(15), 1787–1789 (2004).
- [19] Matsushima, K. and Shimobaba, T., “Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields,” *Optics Express* **17**(22), 19662–19673 (2009).
- [20] Raupach, S., “Observation of interference patterns in reconstructed digital holograms of atmospheric ice crystals,” *Journal of Atmospheric and Oceanic Technology* **26**(12), 2691–2693 (2009).