Reconstruction algorithms applied to in-line Gabor digital holographic microscopy

Karen M. Molony [a,*], Bryan M. Hennelly [a], Damien P. Kelly [a], Thomas J. Naughton [a,b]

[a] Department of Computer Science, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland
[b] University of Oulu, RFMedia Laboratory, Oulu Southern Institute, Vieriimaantie 5, 84100 Ylivieska, Finland

ABSTRACT

This paper investigates the application of Fresnel based numerical algorithms for the reconstruction of Gabor in-line holograms. We focus on the two most widely used Fresnel approximation algorithms, the direct method and the angular spectrum method. Both algorithms involve calculating a Fresnel integral, but they accomplish it in fundamentally different ways. The algorithms perform differently for different physical parameters such as distance, CCD pixel size, and so on. We investigate the constraints for the algorithms when applied to in-line Gabor digital holographic microscopy. We show why the algorithms fail in some instances and how to alter them in order to obtain useful images of the microscopic specimen. We verify the altered algorithms using an optically captured digital hologram.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Holography [1,2] is an imaging technique made up of two parts, recording and reconstruction. The first part of the imaging system involves the recording of an interference pattern from an object beam and a reference beam. The second part involves reconstructing the object wavefront from this recording. Digital holography differs from conventional holography in that a digital camera (CCD/CMOS) is used in place of photographic film or holographic plates. Reconstruction of the hologram is then performed numerically on a computer [3,4]. The concept of digital holography emerged in the 1960s, as reviewed in Ref. [5]. Numerical reconstruction techniques for optically recorded holograms had been applied by then [6], however, it is only in the past decade that it is has become feasible to record holograms digitally and use them practically [7]. This advance was facilitated by the availability of digital cameras with high spatial resolution and high dynamic range. The output of a numerical reconstruction is, in general, a complex two-dimensional representation of the wavefront at a single distance from the camera plane, and so we refer to the digital capture combined with numerical reconstruction as an imaging system.

The novel microscopic principle originally proposed by Gabor [1] is the simplest realisation of holography and has been coined digital in-line holographic microscopy (DIHM) [8]. This simple set up requires only a narrowband light source, a pinhole, an object to be imaged, and an intensity recording area sensor. It is this optical recording set up that we concern ourselves with in this paper. An assumption made when using this set up is that the object wave is weak with respect to the reference wave, which is an accurate approximation in the case of partially transparent objects [8] or small opaque objects sparsely distributed in the field of view [9]. This limits applicability of DIHM as described in Chapter 9 of [10].

In our analysis of DIHM we consider (i) the diverging spherical wave that illuminates our object, (ii) our spherical reference wave, (iii) DC terms and finally (iv) the twin image. (i) In our DIHM setup we use a microscopic object and a pinhole to spatial filter the illuminating wavefield. We model the field that emerges from our pinhole as a coherent point source that diverges and fully illuminates our object. (ii) The presence of a chirp function due to the diverging spherical reference wave, which would have a complex amplitude of 1 for a plane wave, influences the phase results that can be reconstructed from a hologram. In this paper, different methods for dealing with this chirp function are explored. The presence of a spherical wave reference beam means that the hologram has a convergent quadratic phase present before we try to reconstruct. (iii) A disadvantage of DIHM is the overlapping of the two dc terms and the twin image term during reconstruction [11]. The reference wave intensity can be captured in isolation in the absence of the object. This image can be subtracted from the hologram to achieve a contrast hologram [8,12–14] and thus approximates the removal of the dc term

* Corresponding author. Tel.: +353 874123905.
E-mail address: karen.molony@gmail.com (K.M. Molony).

[8]. (iv) The effect of the twin image on each pixel in the reconstructed intensity is assumed to be small due to the geometry of the system [8]. The two twin images can be found equally separated on either side of the hologram plane. When we take into account the hologram's convergence to one object we can conclude that its twin will undergo an equally dramatic divergence. Therefore the signal produced by the twin is spread out and smeared over the reconstructed object intensity as a small background signal [8]. As we shall see these conditions appear to be satisfied for the experimental results that we present later.

The second stage of holography is reconstruction and it is this stage that is the primary focus of this paper. Digital reconstruction is possible on a computer by numerically calculating a diffraction integral that describes the diffraction in free space from the recorded hologram [11]. In our Gabor set up, which is a lensless system, the diffraction of the illuminated object can be characterised by a Fresnel integral [15], i.e. only small angles of diffraction are involved. The sampling conditions for the Fresnel integral have been formalised [16–19] and numerical approximations of the Fresnel transform (FST) have been applied successfully for digital hologram reconstruction [18,20,21]. Efficient algorithms for calculating free space Fresnel diffraction patterns have been developed [17] and investigated for a converging wave case [22]. These two commonly used algorithms can be described as (i) a direct discretisation of the Fresnel kernel which employs one discrete Fourier transform (DFT) and (ii) the angular spectrum (spectral) method describing the FST as a convolution, usually employing two DFTs. Much work has been carried out understanding and improving these two algorithms. Currently, neither of these well known algorithms are used to reconstruct the complex wavefront at the object plane from a hologram captured specifically via a DIHM set up. The typical DIHM reconstruction approach is to apply a numerical approximation of the Kirchoff–Helmholtz transform bound by the recording aperture [8,23]. In this paper, we investigate the appropriateness of the direct method and the angular spectrum method for the DIHM imaging system by considering the achieved output field of view. We find that the presence of spherical waves that are inherent in DIHM in the planes of the object and the recording CCD complicate the applicability of these algorithms and necessitate the development of a new algorithm, based on the angular spectrum method, with a more suitable output field of view. This is the focus of our paper.

The outline of the paper is as follows. In Section 2 we review the optical set up for in-line Gabor digital holography. In Section 3 we review the two most commonly used methods for the simulation of free space propagation as described by the FST: the direct and angular spectrum methods. These can be used to reconstruct Fresnel digital holograms. Then we investigate straightforward application of the direct and angular spectrum methods to reconstruct the Gabor digital holograms. The analysis appears to show that problems exist for the angular spectrum algorithm; the output image range is far too great to view a microscopic specimen. By a simple mathematical manipulation we show how to overcome this problem. In Section 4 we show the result using an optically captured digital hologram. Finally, the main conclusions are drawn in Section 5.

2. Recording an in-line Gabor digital hologram

A typical physical set up for Gabor holography is shown in Fig. 1. For simplicity, we consider the one-dimensional case only. A spherical beam emerges from a pinhole with a wavelength λ. After propagating a distance d_1 this diverging spherical field, $r(x; d_1)$, called the reference field, where

$$r(x; z) = \exp[(i\pi x^2)/(\lambda z)], \quad (1)$$

is incident upon a highly transmissive scene [5],

$$g(x) = |g(x)| \exp[i\phi(x)]. \quad (2)$$

Note that for notational simplicity we do not include a $1/\sqrt{i\lambda z}$ in Eq. (1). We approximate the wavefield immediately after the object by following the analysis in Section 8.1.2 of [10]. To employ this analysis we introduce two assumptions. We first assume that the amplitude function, $|g(x)|$, is approximately constant and we replace it with a constant value, C. Secondly we assume that the $\phi(x)$ function does not vary significantly from 2π. Using these assumptions we expand Eq. (2) with a binomial expansion to give

$$g(x) \approx C[1 + i\phi(x) + \ldots]. \quad (3)$$

The field immediately after the object illuminated by $r(x; d_1)$ may be described as

$$t(x; d_1) = r(x; d_1)g(x) \approx C[r(x; d_1) + o(x; d_1)], \quad (4)$$

where $o(x; d_1) = r(x; d_1)f(x)$ and $f(x) = i\phi(x)$. Setting $C = 1$ for simplicity, Eq. (4) becomes

$$t(x; d_1) \approx r(x; d_1) + o(x; d_1). \quad (5)$$

F_z is the FST operator, applied to $\alpha(x)$ with a coordinate system x', that is defined in Chapter 4 of [10] as

$$F_z[\alpha(x)] = 1/\sqrt{i\lambda z} \int_{-\infty}^{+\infty} \alpha(x) \exp\{[i\pi(x - x')^2]/(\lambda z)\}dx. \quad (6)$$

Using this to propagate $t(x; d_1)$ to the camera plane gives

$$T(x') = F_{d_2}[t(x; d_1)] = F_{d_2}[r(x; d_1)] + F_{d_2}[o(x; d_1)]. \quad (7)$$

The intensity of the wave field that is incident on the CCD is our hologram $H(x')$:

$$H(x') = |T(x')|^2 = |R(x') + O(x')|^2, \quad (8)$$

where

$$R(x') = r(x; d_1 + d_2), \quad (9a)$$

and

$$O(x') = F_{d_2}[o(x; d_1)](x'). \quad (9b)$$

Expanding Eq. (8) gives an expression with four terms [7]

$$H(x') = |O(x')|^2 + |R(x')|^2 + O(x')R^*(x') + R(x')O^*(x'). \quad (10)$$

In line with the analysis presented in Section 9.3, Gabor holography, of [10] we assume that $|O(x')| \ll R(x')$ and so the $R(x')$ term dominates Eq. (10). Only the $|O(x')|^2$ term in Eq. (10) has no $R(x')$ part and is, therefore, considered negligible [8]. We drop this term for the remainder of the manuscript.
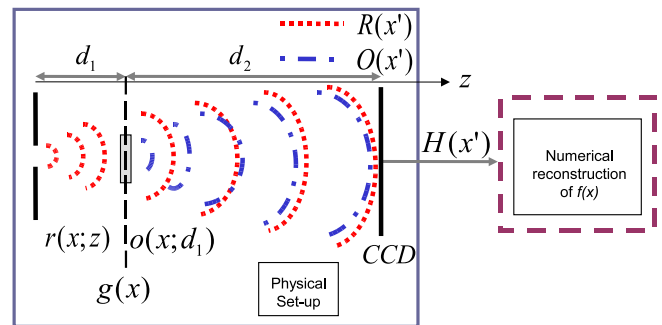

Fig. 1. The Gabor set up with a diverging spherical beam: signal $r(x; z)$ emerging from a pin hole illuminates an object $g(x)$ at a distance d_1. Immediately behind this plane the object wave is $o(x; d_1) = f(x)r(x; d_1)$. The interference pattern, $H(x')$, between the propagated reference wave $R(x')$ and the propagated object wave $O(x')$ is sensed at a distance d_2. This capture is the input to the numerical reconstruction part of the imaging system described in the text.

If $R(x')$ is an ideal spherical wave then $|R(x')|^2$ will simply be a constant value and can be removed using numerical techniques [24]. Alternatively $|R(x')|^2$ can be recorded separately and removed by subtraction which is the approach that we employ. Although the reference spherical wave is modulated by the object, an ideal reference wave at the CCD plane has been assumed during the numerical reconstruction process and a contrast hologram has been achieved [8,12–14]. We note that the reference beam will be modulated by the object, and our assumptions employed by Eqs. (3)–(5) may contribute to error in our reconstructed image after numerical reconstruction. To the best of our knowledge this error has not been considered in previous publications on DIHM [8,12–14]. It is widely assumed that the object wave is weak with respect to the reference wave which is an accurate approximation in the case of highly transmissive objects [10], or opaque objects that are small with respect to the area being imaged [1,8], in order for this approximation to be valid.

Once the two dc terms, $|O(x')|^2$ and $|R(x')|^2$, have been omitted, the two twin image terms remain. As discussed in the introduction we assume that the twin image term is so spread out by the diverging wave that it can be considered insignificant [8]. Only one of the twin images is of interest so we let

$$\widetilde{H}(x') = O(x')R^*(x'),\tag{11}$$

for this analysis.

In this section we have provided a mathematical description of the recording process. However this in itself does not constitute an imaging system. Rather it should be considered as one half of an imaging system. As depicted in Fig. 1, we now implement the second half of the imaging system by processing $\widetilde{H}(x')$ using numerical algorithms. In the next section we examine two numerical algorithms that accomplish this, using the direct method [17] and the angular spectrum method [25,26]. We note that, in effect, we are examining different imaging systems. Both of them have the same first half but different second halves. This assertion becomes clear when we view the numerical reconstruction procedure as an optical system, as illustrated in the figures that follow.

3. Reconstructing an in-line Gabor digital hologram

In the previous section we reviewed the first part of a Gabor in-line holography imaging process: the process of recording an in-line hologram. In this section, we introduce numerical reconstruction algorithms that complete the imaging system. We begin by reviewing a matrix based method for understanding and decomposing algorithms to simulate optical systems including free space propagation as described by the paraxial approximation.

3.1. A matrix description for understanding algorithms

Numerical reconstruction of a hologram generally requires the evaluation of an interference pattern using discrete mathematics. DFTs can be used to efficiently calculate an FST. Two methods for this type of evaluation are considered here, the direct method and the angular spectrum method. In Ref. [27] it was shown that the direct and angular spectrum methods are completely equivalent provided appropriate interpolation of the signal is performed at different stages in the algorithms. All optical systems, including free space propagation as described by the FST, can be represented by an ABCD ray transfer matrix and a framework is provided in Ref. [27] for this type of analysis. Using this framework, a matrix can be decomposed into a product of other matrices that have well understood numerical algorithms associated with them. By implementing these algorithms in the correct sequence of the decomposition one arrives at an overall algorithm that simulates

the original system. Take for example free space propagation as described by the FST: one can derive both the direct and angular spectrum methods in a very convenient manner using this approach. Since this framework is used for the computation presented in this paper which is independent of [27], we must familiarise ourselves with the following matrix operations:

$$\begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix},\tag{12a}$$

$$\begin{bmatrix} S & 0 \\ 0 & 1/S \end{bmatrix},\tag{12b}$$

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ and}\tag{12c}$$

$$\begin{bmatrix} 1 & 0 \\ 1/(-\lambda f) & 1 \end{bmatrix}.\tag{12d}$$

These four matrices represent the ABCD ray transfer matrices for (i) FST, (ii) a magnification system (scaling), (iii) a Fourier transform and (iv) a lens where λ is the wavelength, z is the propagation distance, S is a scaling factor and f is the focal length of the lens [28]. The latter three matrices are well defined [27]. These operations are used throughout this paper.

We refer the reader to Refs. [19,27,29] for a graphical interpretation of these effects based on the Wigner distribution function. We note that the latter three matrices in Eq. (12) have well defined numerical algorithms associated with them, see Section 5 of [27]. The scaling matrix, Eq. (12b), means that we redefine the sampling interval in our algorithm. The Fourier transform matrix, Eq. (12c), means that we implement a DFT algorithm on our samples. The chirp matrix for a positive lens $\exp[(-i\pi x^2)/(\lambda f)]$, Eq. (12d), means that we multiply the samples by a discrete chirp function (f is replaced by $-f$ for a negative or concave lens). If it is possible to decompose our Fresnel matrix, Eq. (12a), into a product of these matrices, Eqs. (12b)–(12d), then we can calculate the FST using an ordered sequence of these well defined numerical processes. The reader should consult Section 5 in Ref. [27] for a complete description on how to numerically implement each matrix and how to track changes to the sampling rate by analysing the space bandwidth product after each matrix operation.

3.2. A brief review of the direct and angular spectrum methods for simulating the FST

The direct method can be represented using the following matrix decomposition

$$\begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/(\lambda z) & 1 \end{bmatrix} \times \begin{bmatrix} \lambda z & 0 \\ 0 & 1/(\lambda z) \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 1/(\lambda z) & 1 \end{bmatrix}.\tag{13}$$

Thus the direct method corresponds to four numerical processes. On the right hand side of Eq. (13), from right to left: we first multiply our sampled signal by a discrete chirp function. Then we implement a DFT. This is followed by a scaling operation which is, in turn, followed by multiplying the resultant output samples by a second discrete chirp function. Appropriate sampling rates to ensure replicas do not overlap with each other can be found using the approach in Ref. [27]. Another possible method for numerically evaluating Eq. (6) is the angular spectrum method. This method comes from the following decomposition

$$\begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -\lambda z & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.\tag{14}$$

The angular spectrum method corresponds to three numerical processes. For the expression right of the equals sign, moving from right to left: we implement a DFT, multiply by a chirp function

and then implement an inverse DFT. Both of these descriptions omit the constant phase term in Eq. (6), $1/\sqrt{i\lambda z}$, which we apply as a final step in both of these methods.

It is well known that the direct method produces a scaled output distribution, whereas the angular spectrum method has a constant output field of view [17,22,26]. Accordingly the direct method is more applicable for a diffraction field further from the diffraction plane, and the angular spectrum method is applicable for short distances [10,30]. In what follows we define the resulting output window width for each of the algorithms given the system parameters d_1, d_2, λ, N and dx where d_1 and d_2 are shown in Fig. 1, λ is the wavelength of the light, N is the number of pixels across the width of the CCD and dx is the distance between the centers of adjacent pixels of the CCD. We then use this criteria alone for determining which algorithm is suitable rather than performing a time consuming interpolation step.

Of particular importance in the present discussion is the relationship between input and output window width for both the direct and angular spectrum methods. The output window width, Δ_{out}, with respect to the input window width, Δx, depends on the method used to evaluate the Fresnel integral. The direct method, Eq. (13), evaluates Eq. (6) discretely by sampling the integral. The Fourier transform part of the Fresnel integral can be made to reduce to a discrete Fourier transform by imposing the following relation [17]

$$\Delta x \Delta_{out} = \lambda z N. \tag{15}$$

Δ_{out} widens as z increases. For the angular spectrum method, Eq. (14), the DFT and DFT^{-1} mean that the input and output distributions are identical [17]:

$$\Delta x = \Delta_{out}. \tag{16}$$

3.3. A matrix description of the recording set up

Now that we have established the two main numerical techniques for simulating Fresnel transformation we focus on finding a matrix description for the recording set up and its corresponding inverse. The idea here is that we will be able to describe the recording system matrix as a product (or decomposition) of the matrices associated with the numerical algorithms that we have at our disposal, e.g. the Fresnel matrix is associated with the two algorithms we have just discussed. Thus we will develop algorithms that can undo the recording process and bring us back to the object plane. This is of course our image plane. Our goal is to calculate the field $f(x)$ from $\tilde{H}(x')$. The matrix, M_p, describing the relationship between $f(x)$ and $\tilde{H}(x')$ can be derived by looking at Fig. 1 and is defined as

$$M_p = \begin{bmatrix} 1 & 0 \\ -1/[\lambda(d_1+d_2)] & 1 \end{bmatrix} \times \begin{bmatrix} 1 & \lambda d_2 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 1/(\lambda d_1) & 1 \end{bmatrix}. \tag{17}$$

The rightmost matrix represents the illuminating chirp, the middle one represents the FST using a distance d_2, and the remaining matrix represents the reference beam chirp that multiplies by the object wave field at the CCD. Our goal therefore is to find a decomposition of the inverse of this matrix, M_p^{-1} where $M_p^{-1} \times M_p = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ such that the Gabor set up followed by our algorithm forms a one to one imaging system. This decomposition will be in terms of the four matrices already defined above in Eq. (12a). This time we can say that the Fresnel matrix has a well defined algorithm associated with it since we have just shown how to derive two in this section.

In the following subsections 3.4 and 3.5 we derive two decompositions of M_p^{-1} that employ the Fresnel matrix, the lens matrix and the scaling matrix. The first of the two approaches calculates

$f(x)$ using the actual distance d_2 for FST evaluation. The second approach calculates $f(x)$ using a virtual distance that is related to the physical distance, d_2, by a scaling factor,

$$S = (d_1 + d_2)/d_1, \tag{18}$$

for evaluating the FST.

The necessity for more than one algorithm to reconstruct the image can be explained as follows. The first approach is essentially a direct inverse of the system. We undo the reference chirp by multiplying by the conjugate chirp. Then we propagate back to the object plane using one of our two Fresnel algorithms, and finally we undo the illuminating chirp by multiplying by its conjugate. This somewhat obvious and straightforward method works well when the direct method is used as the algorithm to calculate the Fresnel part. However, the angular spectrum method fails when it is used because the output size is too wide. From the above discussions we know the output range of the angular spectrum method will be equal to the input range, i.e. the physical CCD size which is usually in the order of 10 mm. This is an impractical output image size if our microscopic object is of the order of 0.1 mm. Thus, we need a second approach based on a different matrix decomposition of our system that leads to a new algorithm that does not have this problem. In the following two subsections we discuss these ideas in more detail.

3.4. The straightforward matrix decomposition: actual distance (AD) approach

In this approach we implement M_p^{-1} by simply implementing the inverse of each of three matrices in Eq. (17) one by one. Thus, the first step required in order to reconstruct $f(x)$ from $\tilde{H}(x')$ is to remove the chirp function $R^*(x')$. We achieve this numerically by multiplying $\tilde{H}(x')$ by $R(x') = \exp\{(i\pi x'^2)/[\lambda(d_1+d_2)]\}$ to give $O(x')$. So now the task remains to reconstruct $f(x)$ from $O(x')$.

Recovering $o(x; d_1)$ from $O(x')$ can be performed as per Eq. (9b) using either the direct method or the angular spectrum method with a distance of $-d_2$. $f(x)$ is then recovered from $o(x; d_1)$ by removing the additional chirp term, i.e. multiplying $o(x; d_1)$ by $r^*(x; d_1)$ to obtain the final image, $f(x)$. A schematic of this process is shown in Fig. 2.

This approach can described through the following decomposition

$$M_p^{-1} = \begin{bmatrix} 1 & 0 \\ -1/(\lambda d_1) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -\lambda d_2 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 1/[\lambda(d_1+d_2)] & 1 \end{bmatrix}. \tag{19}$$

The matrices, from right to left, represent: multiplication by the conjugate of chirp $(d_1 + d_2)$, FST $(-d_2)$ which is achieved by either the direct method or the angular spectrum method, and multiplication by conjugate of chirp $(-d_1)$. So for the AD approach, the distance used for the FST is the actual distance, d_2.
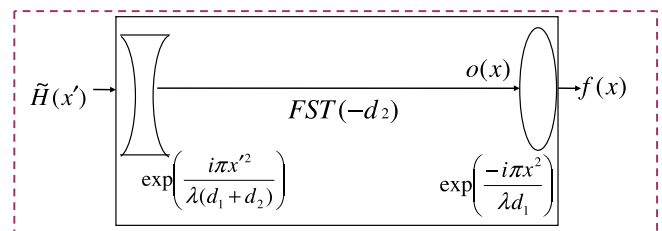

Fig. 2. Schematic of AD approach. This optical system is simulated, with $\tilde{H}(x')$ as input, in order to get to the image plane of the system in Fig. 1.

In a previous section we discussed how the output window width, Δ_{out}, to the FST is different for the direct method and the angular spectrum method as shown in Eqs. (15) and (16), respectively. The resulting window size at the output of a given reconstruction algorithm depends on the input window width and the method used to approximate the FST. In our case, where we are reconstructing an in-line Gabor hologram, the input window width will be given by $\Delta x = \Delta_{CCD}$, i.e. the physical width of the CCD. The AD output window widths are straightforward to calculate as the actual distance d_2 is substituted for z in Δ_{out} for the output window width using the direct method. The output window width for the angular spectrum method is simply the input window width and no scaling needs to be considered for either method.

In order to fully understand the importance of the output window widths we now illustrate the output widths from the AD approach for different values of $z = d_2$. We take a value for the camera width $\Delta_{CCD} = 12$ mm, with a number of pixels $N = 2008$, and the wavelength of the illuminating light to be $\lambda = 405 \times 10^{-6}$ mm. These values are the actual values of the physical parameters in the results section that follows. In Fig. 3 we show the output widths for the AD method when the direct and angular spectrum methods are employed. The window width is plotted against an increasing d_2 for a fixed $d_1 = 1.2$ mm. In these graphs it can be seen that the output window width is constant for the angular spectrum method, but that it expands from zero for the direct method as d_2 increases. In general, our object will be quite small and will be placed a short distance from the camera. In Fig. 3 we illustrate this point by showing a small object window 7 mm at a distance from the camera equal to $d_2 = 12$ mm. This is consistent with the microscopic specimen imaged in the results section of this paper. It is clear that the output window width for the angular spectrum case is far too large, i.e. our object will appear very small in our image. This necessitates the need for a new approach if the angular spectrum method is to be used for practical in-line Gabor digital hologram reconstructions.
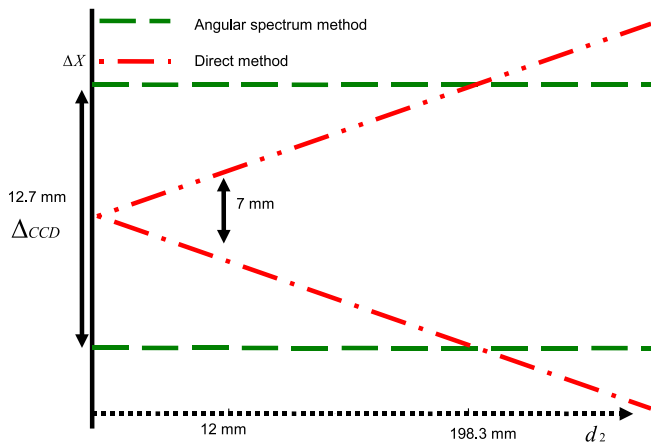

Fig. 3. For the AD approach, given a fixed $d_1 = 1.2$ mm, this figure shows the change in output window size (field of view) of reconstructions obtained by the direct method and the angular spectrum method for different values of d_2. The window width of the CCD, Δ_{CCD}, is 12.7 mm, which is the output window width for the angular spectrum method irrespective of d_2. The output window width of the direct method at the object plane, $d_2 = 12$ mm, is 7 mm and the object can be observed. The field of view observed by the angular spectrum method is very large compared to the size of small object. This requires one to rescale or zoom to view the object comfortably, and is wasteful of computer memory if applied to the whole field of view. However, the direct method achieves a small enough field of view to see our object. We observe a common output window width for the two methods when $d_2 = 198.3$ mm.

3.5. The second matrix decomposition: scaled down distance (SDD) approach

As in the AD approach, $f(x)$ is calculated from $\widetilde{H}(x')$. Again, the chirp function $R^*(x')$ is removed by multiplying $\widetilde{H}(x')R(x')$ which leaves $O(x')$. $f(x)$ is reconstructed from $O(x')$ using the numerical process shown in Fig. 4. This time $o(x; d_1)$ is not recovered directly from $O(x')$ as the chirp function is removed prior to the FST and so the input to the FST is scaled by $1/S$, where S is defined in Eq. (18),

$$F_{-d_2/S}[f(x)](x'/S). \tag{20}$$

For the SDD approach the distance used to evaluate the FST is the actual distance, d_2, divided by S. Removal of the two chirp functions cancel each other and so they are not shown in the subsequent discussion. A detailed derivation of this result is provided in Appendix A using wave equations.

A matrix decomposition of this approach is given by

$$M_p^{-1} = \begin{bmatrix} 1 & (-\lambda d_2)/S \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1/S & 0 \\ 0 & S \end{bmatrix}. \tag{21}$$

From right to left these matrices represent: scale $(1/S)$, and FST $(-d_2/S)$.

For the SDD approach, the input window size is scaled by $1/S$. For the direct method d_2/S is substituted for z in Δ_{out} which cancels out the scaling factor. In fact, it is for this reason that the SDD approach using the direct method is exactly the same as the AD method using the direct method. For the angular spectrum method the scaled input window width is the output window width. The output window widths for the SDD approach are calculated by multiplying by the scaling factor on the system, $1/S$, for both methods and replacing z in equation Δ_{out} with $d_2 S$ for the direct method.

Now we illustrate the output widths from the SDD approach for different values of $z = d_2$. Once again we take a value for the camera width $\Delta_{CCD} = 12.7$ mm, with a number of pixels $N = 2008$, and the wavelength of the illuminating light to be $\lambda = 405 \times 10^{-6}$ mm. We take $d_1 = 0.12$ mm and $d_2 = 12$ mm. These values are the actual values of the physical parameters in the results section that follows. In Fig. 5 we show the output widths for the SDD method when the direct and angular spectrum methods are employed for a small object window, 0.7 mm, at a distance, $d_2 = 12$ mm, from the camera.

It is clear that the output window width for the angular spectrum case is now a suitable size for viewing the object.

4. Results

An example hologram, see Fig. 6, of a real world object captured using a Gabor digital holographic microscope set up, is used to illustrate the performance of each system. The object is a slice through the head of a drosophila melanogaster, commonly known
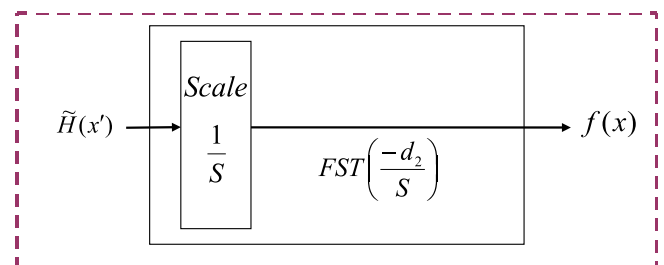

Fig. 4. Schematic of SDD approach. This optical system is simulated, with $\widetilde{H}(x')$ as input, in order to get to the image plane of the system in Fig. 1.
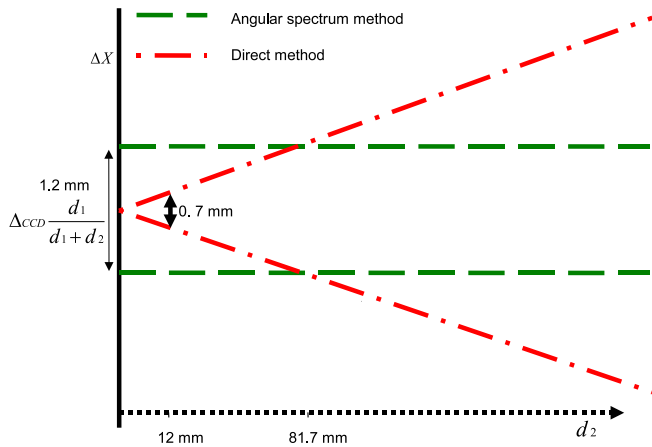
Fig. 5. For the SDD approach, given a fixed $d_1 = 1.2$ mm, this figure shows the change in output window size of reconstructions obtained by the direct method and the angular spectrum method for different values of d_2. The output window size achieved by the angular spectrum method is $\Delta_{CCD}/S = 1.2$ mm, for all values of d_2, which is now sufficiently small to view comfortably our rescaled small object, which was too small in the field of view using this method and the AD approach. An output window width of 0.7 mm is achieved by the direct method which is scaled down with respect to the AD approach. However the same scale function applied to the reconstruction distance when using the SDD approach cancels out this observed effect and the same view of our small object can be achieved by both the AD and SDD approach.
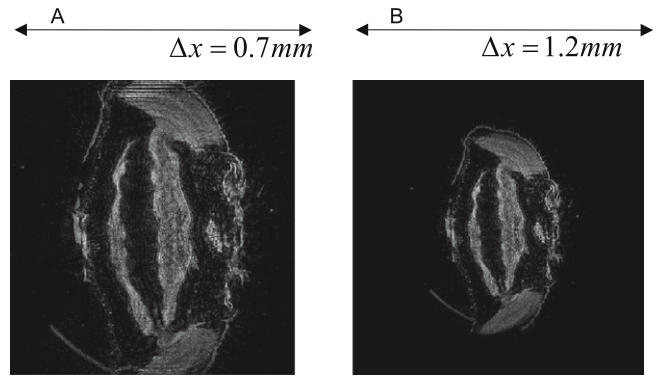

Fig. 6. Optically captured hologram of a drosophila melanogaster head, minus the intensity image of the reference beam $|R(x')|^2$ when no object was present in the optical path, courtesy of Prof. H. Juergen Kreuzer and his collaborators.

Fig. 7. Reconstruction results for the AD approach: (A) intensity image from the direct method, and (B) intensity image from the angular spectrum method. The output window width for each result is shown. We can see in (A) that the AD approach using the direct method is suitable for viewing our object. However, in (B) it is clear that the field of view, 12.7 mm, is too large to comfortably view our small object using the AD approach and the angular spectrum method. In fact multiple small copies of our object can be observed in this case.

Fig. 8. Reconstruction results for the SDD approach: (A) intensity image from the direct method, and (B) intensity image from the angular spectrum method. The output window width for each result is shown. These results show that the SDD approach provides a suitable output window width for viewing our small object with either the direct method or the angular spectrum method.
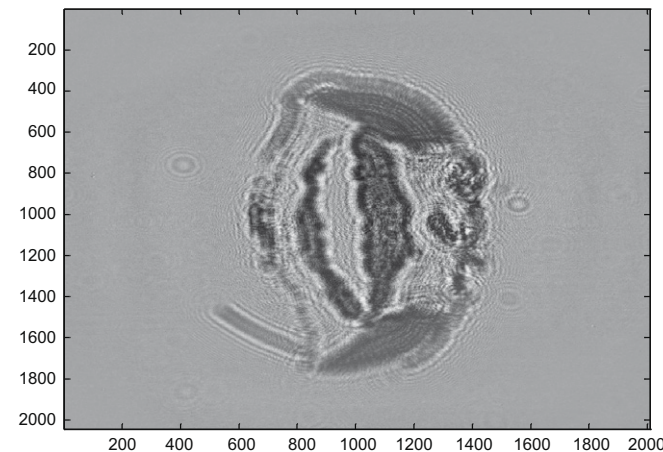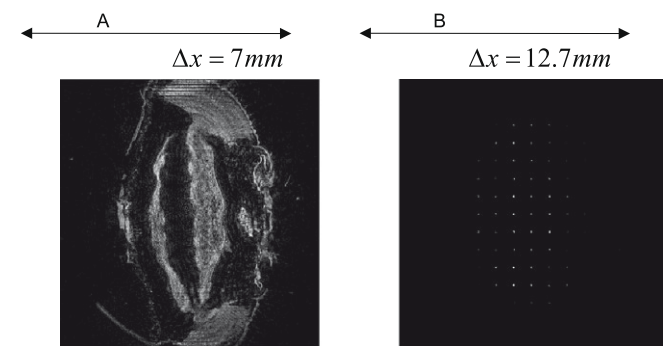
as a fruit fly. The associated parameters are given in the previous section. The reference image $|R(x')|^2$ has been subtracted, thus the hologram can be described as $|O(x')^2| + O(x')R^*(x') + O^*(x')$ $R(x')$. As previously stated, $|O(x')|^2$ is ignored as it is such a weak term, and the spread out twin image, $O^*(x')R(x')$, can also be considered negligible. We thus assume that our hologram can be described simply as $\widetilde{H}(x') = O(x')R^*(x')$. The verification for the two methods are given in Figs. 7 and 8.

5. Conclusion

Numerical algorithms for the reconstruction of Gabor in-line holograms using Fresnel integral calculation have been presented. The direct method and the angular spectrum method were used to evaluate Fresnel diffraction patterns for the DIHM system considered. Imaged objects are magnified by the scaling factor of the imaging system which is the distance of the light source to the imaging device divided by the distance from the light source to the sample. Therefore microscopic objects are often placed close to the light source.

Two approaches for numerically reconstructing holograms in this configuration have been investigated here. The first, the AD approach, is a direct numerical inverse of the physical recording system. This approach uses the physical distance from the object to the imaging device as an input parameter for the Fresnel integral evaluation. It was shown that this approach is applicable using the direct method but that using the angular spectrum method with this approach is impractical due to the output image size. Therefore a second and new approach was developed, the SDD approach. This new approach is also an inverse of the physical imaging system. However the process is not reversed, rather the order of implementation is manipulated. A scaled distance, the physical distance from the object to the imaging device scaled down by the scaling factor of the imaging system, is used as the input parameter for the Fresnel integral evaluation. This parameter results in the input window width being scaled by the same factor. For the direct method this scaled window size cancels out with the scaled distance and so an identical result to the previous approach is obtained. The scaled input window width for the angular spectrum method renders the output window width suitable for viewing the object.

These two approaches combined with the same physical Gabor set up comprises two separate DIHM imaging systems. The direct method for Fresnel integral evaluation is suitable for use with either imaging system. The angular spectrum method, however, is only applicable for the SDD approach.

Acknowledgements

Appendix A

A full mathematical description of the SDD approach using wave theory is presented here.

$$R(x') \times \widetilde{H}(x') = O(x') \tag{A.1}$$

$$O(x') = F_{d_2}\{f(x)\exp[(i\pi x^2)/(\lambda d_1)]\}(x') \tag{A.2}$$

$$O(x') = \exp\left[(i\pi x'^2)/(\lambda d_2)\right] \int_{-\infty}^{+\infty} \{f(x)\exp\left[(i\pi x^2)/(\lambda d_1)\right]\} \times \exp\left[(i\pi x^2)/(\lambda d_2)\right]\exp[(-i2\pi xx')/(\lambda d_2)]dx \tag{A.3}$$

$$O(x') = \exp\left[(i\pi x'^2)/(\lambda d_2)\right] \int_{-\infty}^{+\infty} f(x) \times \exp\left[(i\pi x^2 S)/(\lambda d_2)\right]\exp[(-i2\pi xx')/(\lambda d_2)]dx \tag{A.4}$$

where $S = (d_1 + d_2)/d_1$ as before.

$$O(x') = \exp\left[(i\pi x'^2)/(\lambda d_2)\right] \int_{-\infty}^{+\infty} f(x) \times \exp\left[(i\pi x^2 S)/(\lambda d_2)\right]\exp\{[(-i2\pi xx'S)/(\lambda d_2)](1/S)\}dx \tag{A.5}$$

$$O(x') = \exp\left[(i\pi x'^2)/(\lambda d_2)\right]\exp\{[(-i\pi x'^2 S)/(\lambda d_2)](S)\} \times \exp\{[(i\pi x'^2 S)/(\lambda d_2)](S)\} \int_{-\infty}^{+\infty} f(x) \times \exp\left[(i\pi x^2 S)/(\lambda d_2)\right]\exp\{[(-i2\pi xx'S)/(\lambda d_2)](1/S)\}dx \tag{A.6}$$

$$O(x') = \exp\{(i\pi x'^2)/(\lambda)[1/d_2 - S/(d_2 S)]\} \times \exp\{[(i\pi x'^2 S)/(\lambda d_2)](S)\} \int_{-\infty}^{+\infty} f(x) \times \exp\left[(i\pi x^2 S)/(\lambda d_2)\right]\exp\{[(-i2\pi xx'S)/(\lambda d_2)](1/S)\}dx \tag{A.7}$$

$$O(x') = \exp(0)\exp\{[(i\pi x'^2 S)/(\lambda d_2)](S)\} \int_{-\infty}^{+\infty} f(x) \times \exp\left[(i\pi x^2 S)/(\lambda d_2)\right]\exp\{[(-i2\pi xx'S)/(\lambda d_2)](1/S)\}dx \tag{A.8}$$

$$O(x') = F_{d_2/S}[f(x)](x'/S) \tag{A.9}$$

$$f(x) = F_{-d_2/S}[O(x')(S/x')] \tag{A.10}$$

References

[1] D. Gabor, Nature 161 (1948) 777.
[2] E.N. Leith, J. Upatnieks, J. Opt. Soc. Am. 54 (1964) 1295.
[3] T.M. Kreis, Handbook of Holographic Interferometry, Wiley-VCH, 2005.
[4] U. Schnars, W.P.O. Jüptner, Digital Holography, Springer, 2004.
[5] L.B. Lesem, P.M. Hirsch, J.A. Jordan jr., Commun. ACM 11 (1968) 661.
[6] J.W. Goodman, R.W. Lawrence, Appl. Phys. Lett. 11 (1967) 77.
[7] U. Schnars, T.M. Kreis, W.P.O. Jüptner, Opt. Eng. 35 (1996) 977.
[8] J. Garcia Sucerquia, W. Xu, S.K. Jericho, P. Klages, M.H. Jericho, H.J. Kreuzer, Appl. Opt. 45 (2006) 836.
[9] M. Kempkes, E. Darakis, T. Khanam, A. Rajendran, V. Kariwala, M. Mazzotti, T.J. Naughton, A.K. Asundi, Opt. Express 17 (2009) 2938.
[10] J.W. Goodman, Introduction to Fourier Optics, Roberts and Company Publishers, 2004.
[11] G. Pedrini, P. Fröning, H. Fessler, H.J. Tiziani, Appl. Opt. 37 (1998) 6262.
[12] W. Xu, M.H. Jericho, I.A. Meinertzhagen, H.J. Kreuzer, Appl. Opt. 41 (2002) 5367.
[13] J. Garcia-Sucerquia, D.C. Alvarez-Palacio, M.H. Jericho, H.J. Kreuzer, Opt. Lett. 31 (2006) 2845.
[14] J. Garcia-Sucerquia, W. Xu, S.K. Jericho, M.H. Jericho, H.J. Kreuzer, Optik 119 (2008) 419.
[15] G. Liu, P.D. Scott, J. Opt. Soc. Am. A 4 (1) (1987) 159.
[16] F. Gori, Opt. Eng. 39 (1981) 293.
[17] D. Mas, J. García, C. Ferreira, L.M. Bernardo, F. Marinho, Opt. Commun. 164 (1999) 233.
[18] A. Stern, B. Javidi, Opt. Eng. 43 (2004) 239.
[19] D.P. Kelly, B.M. Hennelly, W.T. Rhodes, J.T. Sheridan, Opt. Eng. 45 (2006) 088201-1.
[20] Y. Zhang, G. Pedrini, W. Osten, H.J. Tiziani, Opt. Lett. 29 (2004) 1287.
[21] Y. Frauel, T.J. Naughton, O. Matoba, E. Tajahuerce, B. Javidi, Proc. IEEE 94 (2006) 636.
[22] D. Mas, J. Pérez, C. Hernández, C. Vázquez, J.J. Miret, C. Illueca, Opt. Commun. 227 (2003) 245.
[23] H.J. Kreuzer, Holographic Microscope and Method of Hologram Reconstruction, US Patent No. 6411406B1, 2002.
[24] T.M. Kreis, W.P.O. Jüptner, Opt. Eng. 36 (1997) 2357.
[25] D. Mendlovic, Z. Zalevsky, N. Konforti, J. Mod. Opt. 44 (1997) 407.
[26] J. García, D. Mas, R.G. Dorsch, Appl. Opt. 35 (1996) 7013.
[27] B.M. Hennelly, J.T. Sheridan, J. Opt. Soc. Am. A 22 (2005) 917.
[28] S.A. Collins Jr., J. Opt. Soc. Am. 60 (1970) 1168.
[29] J.T. Sheridan, B.M. Hennelly, D.P. Kelly, Opt. Lett. 28 (2003) 884.
[30] C.J. Daly, T.W. Nuteson, N.A.H.K. Rao, IEEE Trans. Antennas Propag. 51 (2003) 700.