Ángel F. Doval and Cristina Trillo, "Dimensionless formulation of the convolution and angular spectrum reconstruction methods in digital holography," Proc. SPIE 7387, "Speckle 2010: Optical Metrology," 73870U (September 13, 2010)

Copyright 2010 Society of Photo Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

http://dx.doi.org/10.1117/12.870049

Dimensionless formulation of the convolution and angular spectrum reconstruction methods in digital holography

Ángel F. Doval and Cristina Trillo

Universidad de Vigo. Departamento de Física Aplicada. E.T.S. de Ingenieros Industriales. Campus Universitario. 36310 Vigo, Spain.

ABSTRACT

The evaluation of the Rayleigh-Sommerfeld diffraction formula by means of numerical convolution and angular spectrum filtering are two of the most usual reconstruction methods in digital holography. Both of them are normally implemented by using a discrete Fourier transform and a sample of, respectively, the free space impulse response function and the corresponding transfer function. In this communication we propose a modified formulation of the sampled free space impulse response and transfer functions in terms of five dimensionless parameters: the wavelength to horizontal pixel size ratio, the reconstruction distance to horizontal field size ratio, the field and pixel aspect ratios and the number of pixels in the horizontal direction. This formulation simplifies the task of comparing and finding equivalences between holographic reconstruction situations with different distance, wavelength, field and pixel sizes. The reconstruction range for each of the methods is expressed in terms of the aforementioned dimensionless parameters by analyzing the resolution limits for the impulse response and the transfer function, respectively. This notation makes very simple to decide which of the two methods should be used for given conditions as well as to tailor range extension strategies based on the effects of hologram manipulations such as zero padding or pixel splitting. The details of the implementation of the convolution and angular spectrum algorithms with the proposed formulation are disclosed paying particular attention to the consequences of the sacrificial zero-padding required to avoid aliasing in Fourier-transform based cyclic convolution.

Keywords: Digital holography, diffraction, dimensionless parameters

1. INTRODUCTION

The first step of the reconstruction process in digital holography is obtaining a numerical complex field $u(x, y, 0)$ at the hologram plane either by multiplying a real-valued digital hologram by a numerical reconstruction beam[1] —which very often is an orthogonal plane wave with amplitude one and phase zero— or by directly taking a complex-valued digital hologram as, for example, in phase-shifting digital holography.[2] Then, the Rayleigh-Sommerfeld diffraction formula[3] or one of its approximations (Fresnel's or Fraunhofer's) is used to propagate the field from the hologram plane ($z_0 = 0$) to a given reconstruction distance z, either forwards ($z > 0$) or backwards ($z < 0$), to eventually get a reconstructed field $u(x, y, z)$.

Since the direct evaluation of the Rayleigh-Sommerfeld integral is complicated, two formally equivalent methods[4] —the convolution method (CA) and the angular spectrum method (AS)— are widely used to simplify its calculation when no approximation is allowed. In the convolution method it is interpreted as a convolution

$$u(x, y, z) = h(x, y, z) * u(x, y, 0) \tag{1}$$

of the field $u(x, y, 0)$ at the hologram plane with the impulse response function

$$h(x, y, z) = \frac{\exp\left\{ j\dfrac{2\pi}{\lambda} z \left[1 + \left(\dfrac{x}{z}\right)^2 + \left(\dfrac{y}{z}\right)^2 \right]^{\frac{1}{2}} \right\}}{j\lambda z \left[1 + \left(\dfrac{x}{z}\right)^2 + \left(\dfrac{y}{z}\right)^2 \right]} \tag{2}$$

A.F.D.: E-mail: adoval@uvigo.es

Speckle 2010: Optical Metrology, edited by Armando Albertazzi Goncalves Jr., Guillermo H. Kaufmann, Proc. of SPIE Vol. 7387, 73870U · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.870049

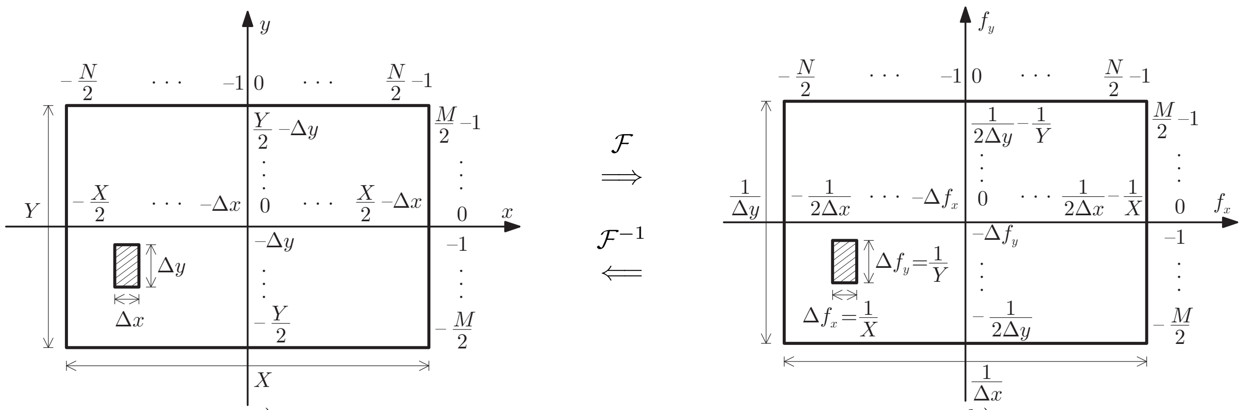


Figure 1. Coordinate systems for: a) the hologram, the impulse response function and the reconstruction, b) the Fourier transforms and the transfer function.

On the other hand, the angular spectrum method makes use of the theory of linear systems to understand the propagation as a linear filtering of the angular spectrum $\mathcal{F}[u(x,y,0)]$ of the original field, formalized as

$$u(x,y,z) = \mathcal{F}^{-1}\left\{H(f_x,f_y,z)\cdot\mathcal{F}[u(x,y,0)]\right\} \tag{3}$$

with the transfer function[3]

$$H(f_x,f_y,z) = \begin{cases} \exp\left\{j\frac{2\pi}{\lambda}z[1-\lambda^2(f_x^2+f_y^2)]^{\frac{1}{2}}\right\} & \text{if } \sqrt{f_x^2+f_y^2} < \frac{1}{\lambda} \\ 0 & \text{if } \sqrt{f_x^2+f_y^2} \geq \frac{1}{\lambda} \end{cases} \tag{4}$$

Both approaches are numerically implemented by using the discrete Fourier transform, commonly in the form of a fast Fourier transform (FFT) algorithm. It is well-known[5,6] that the sampling conditions of the impulse response and transfer functions limit the application range of such implementations to long reconstruction distances for the convolution method and to short ones for the angular spectrum method. Nevertheless the expressions of the resolution limits found in the literature are generally restricted to one dimensional analysis or, at most, to square holograms with square pixels; as they grow in generality, these expressions involve many parameters and it becomes difficult to establish comparisons between different reconstruction situations.

In this communication we propose a modified formulation, established in terms of dimensionless parameters, for the reconstruction ranges of both reconstruction methods as well as for the sampled free space impulse response and transfer functions. This formulation covers arbitrary image and pixel aspect ratios and simplifies the tasks of finding the most convenient reconstruction method and of comparing and finding equivalences between holographic reconstruction situations with different distance, wavelength, field and pixel sizes.

2. THEORY

Let us assume that a digital hologram has been recorded with a sensor whose size is $X \times Y$ and has $N \times M$ pixels with size $\Delta x \times \Delta y$ as shown in Fig. 1-a. These six parameters are not independent since $X = N\Delta x$ and $Y = M\Delta y$. The indices of the pixels (n,m) and their physical coordinates (x,y) with respect to the centre of the sensor are

$$-\frac{N}{2} \leq n \leq \frac{N}{2}-1 \quad ; \quad x = n\Delta x \quad \Rightarrow \quad -\frac{X}{2} = -\frac{N}{2}\Delta x \leq x \leq \left(\frac{N}{2}-1\right)\Delta x = \frac{X}{2}-\Delta x \tag{5}$$

$$-\frac{M}{2} \leq m \leq \frac{M}{2}-1 \quad ; \quad y = m\Delta y \quad \Rightarrow \quad -\frac{Y}{2} = -\frac{M}{2}\Delta y \leq y \leq \left(\frac{M}{2}-1\right)\Delta y = \frac{Y}{2}-\Delta y \tag{6}$$

When a FFT algorithm is applied to a field sampled in the aforementioned conditions, the resulting frequency space has also $N \times M$ pixels but their size is $\Delta f_x \times \Delta f_y = \frac{1}{X} \times \frac{1}{Y} = \frac{1}{N\Delta x} \times \frac{1}{M\Delta y}$ (see Fig. 1-b). The indices of the pixels (n, m) and their physical coordinates in terms of spatial frequencies (f_x, f_y) are

$$-\frac{N}{2} \leq n \leq \frac{N}{2} - 1 \quad ; \quad f_x = n\Delta f_x \quad \Rightarrow \quad -\frac{1}{2\Delta x} = -\frac{N}{2}\Delta f_x \leq f_x \leq \left(\frac{N}{2} - 1\right)\Delta f_x = \frac{1}{2\Delta x} - \Delta f_x \tag{7}$$

$$-\frac{M}{2} \leq m \leq \frac{M}{2} - 1 \quad ; \quad f_y = m\Delta f_y \quad \Rightarrow \quad -\frac{1}{2\Delta y} = -\frac{M}{2}\Delta f_y \leq f_y \leq \left(\frac{M}{2} - 1\right)\Delta f_y = \frac{1}{2\Delta y} - \Delta f_y \tag{8}$$

2.1 Implementation of the Rayleigh-Sommerfeld formula by use of the Fast Fourier Transform

The usual definitions of the direct and inverse fast Fourier transforms[7] and their relation with the sampled values of the continuous Fourier transform are

$$\mathrm{FFT}(g) = \sum_{p=0}^{M-1}\sum_{q=0}^{N-1} g(q\Delta x, p\Delta y)\exp\left(i2\pi\frac{pm}{M}\right)\exp\left(i2\pi\frac{qn}{N}\right) \tag{9}$$

$$\mathcal{F}\left[g\right]_{(n\Delta f_x, m\Delta f_y)} \approx \Delta x\Delta y\,\mathrm{FFT}(g) \tag{10}$$

$$\mathrm{FFT}^{-1}(G) = \frac{1}{MN}\sum_{p=0}^{M-1}\sum_{q=0}^{N-1} G(q\Delta f_x, p\Delta f_y)\exp\left(-i2\pi\frac{pm}{M}\right)\exp\left(-i2\pi\frac{qn}{N}\right) \tag{11}$$

$$\mathcal{F}^{-1}\left[G\right]_{(n\Delta x, m\Delta y)} \approx \Delta f_x\Delta f_y\,MN\,\mathrm{FFT}^{-1}(G) = \frac{1}{\Delta x\Delta y}\,\mathrm{FFT}^{-1}(G) \tag{12}$$

The FFTW library,[8] which is the basis of the Fourier transform functions in the current versions of both Matlab® and Octave, uses this convention. This formulation of the FFT assumes that the signal is periodic in both x and y and that the origins of the coordinates for both the direct and reciprocal spaces are located at a corner of the sampled field. But in digital holography, as in general optics, it is usual to assume that the origin, i.e. the optical axis, is located at the centre of the sampled hologram and of its transform. Therefore quadrants 1st and 3rd as well as 2nd ant 4th must be shifted to change from the optical to the FFT reference system and *vice versa*. We will hereafter denote this quadrant shifting as $\mathrm{FSh}(g)$, where g is either the sampled field or its transform. To keep the dimensional consistency and comply with the coordinate systems depicted in Fig. 1, the sampled Fourier transforms must be calculated as

$$\mathcal{F}\left[g\right] \approx \Delta x\Delta y\,\mathrm{FSh}\{\mathrm{FFT}[\mathrm{FSh}(g)]\} \tag{13}$$

$$\mathcal{F}^{-1}\left[G\right] \approx \frac{1}{\Delta x\Delta y}\,\mathrm{FSh}\{\mathrm{FFT}^{-1}[\mathrm{FSh}(G)]\} \tag{14}$$

2.1.1 Convolution method

The discretized impulse response of free space, Eq. (2), can be written, according to the reference system in Fig. 1-a, as

$$h(n\Delta x, m\Delta y, z) = \frac{\exp\left\{j\frac{2\pi}{\lambda}z\left[1 + \left(\frac{n\Delta x}{z}\right)^2 + \left(\frac{m\Delta y}{z}\right)^2\right]^{\frac{1}{2}}\right\}}{j\lambda z\left[1 + \left(\frac{n\Delta x}{z}\right)^2 + \left(\frac{m\Delta y}{z}\right)^2\right]} \tag{15}$$

The numerical propagation by use of the convolution method is generally implemented by using FFT convolution, $h * u = \mathcal{F}^{-1}\left[\mathcal{F}\left[h\right]\cdot\mathcal{F}\left[u\right]\right]$, which is significantly faster than the direct convolution algorithm. Taking into account Eqs. (13) and (14) the reconstruction formula results

$$u(n\Delta x, m\Delta y, z) = \frac{1}{\Delta x\Delta y}\,\mathrm{FSh}\Big(\mathrm{FFT}^{-1}\Big\{\mathrm{FSh}\Big[\Delta x\Delta y\,\mathrm{FSh}(\mathrm{FFT}\{\mathrm{FSh}[h(n\Delta x, m\Delta y, z)]\}) \times \Delta x\Delta y\,\mathrm{FSh}(\mathrm{FFT}\{\mathrm{FSh}[u(n\Delta x, m\Delta y, 0)]\})\Big]\Big\}\Big) \tag{16}$$

which, by use of the shift properties of the Fourier transform, can be reduced to

$$u(n\Delta x, m\Delta y, z) = \Delta x\Delta y\,\mathrm{FSh}(\mathrm{FFT}^{-1}\{\mathrm{FFT}[h(n\Delta x, m\Delta y, z)]\cdot \mathrm{FFT}[u(n\Delta x, m\Delta y, 0)]\}) \tag{17}$$

2.1.2 Angular spectrum method

The discretized transfer function of free space, Eq. (4), can be written, according to the reference system in Fig. 1-b, as

$$H(n\Delta f_x, m\Delta f_y, z) = \exp\left(j\frac{2\pi}{\lambda}z\{1-\lambda^2[(n\Delta f_x)^2+(m\Delta f_y)^2]\}^{\frac{1}{2}}\right) \tag{18}$$

and the numerical propagation by angular spectrum filtering is discretized as

$$u(n\Delta x, m\Delta y, z) = \frac{1}{\Delta x\Delta y}\,\mathrm{FSh}\Big(\mathrm{FFT}^{-1}\Big\{\mathrm{FSh}\Big[H(n\Delta f_x, m\Delta f_y, z) \times \Delta x\Delta y\mathrm{FSh}(\mathrm{FFT}\{\mathrm{FSh}[u(n\Delta x, m\Delta y, 0)]\})\Big]\Big\}\Big) \tag{19}$$

which, by making use one more time of the shift properties of the Fourier transform, can be reduced to

$$u(n\Delta x, m\Delta y, z) = \mathrm{FFT}^{-1}\{\mathrm{FSh}[H(n\Delta f_x, m\Delta f_y, z)]\cdot \mathrm{FFT}[u(n\Delta x, m\Delta y, 0)]\} \tag{20}$$

2.2 Sampling criteria for the impulse response and transfer functions

Both the digital hologram and, according to the chosen reconstruction method, the discretized impulse response or transfer function must comply with the conditions of the sampling theorem to yield an aliasing-free holographic reconstruction. Since the modulus of the transfer function H is constant and the modulus of the impulse response function $|h|$ changes are significantly slower than its phases', the local spatial frequencies of both functions can be approximated as

$$f_{x_i} \approx \frac{1}{2\pi}\left|\frac{\partial\phi}{\partial x_i}\right| \tag{21}$$

The values of these frequencies must be lower than or equal than the Nyquist critical frequency f_{cx_i} for each of the axes ($x_i \in \{x, y\}$)

$$f_{x_i} \approx \frac{1}{2\pi}\left|\frac{\partial\phi}{\partial x_i}\right| \le \frac{1}{2\Delta x_i} = f_{cxi} \Rightarrow |\Delta\phi| \approx \Delta x_i\left|\frac{\partial\phi}{\partial x_i}\right| \le \pi \tag{22}$$

with Δx_i the sampling interval in the direction of axis x_i, i.e., the local phase change between adjacent samples must not be larger than pi.

2.2.1 Convolution method: impulse response function

The local phase change of the discretized impulse response function $h(n\Delta x, m\Delta y, z)$ between adjacent samples (pixels) along the x axis is

$$\Delta\phi \approx \Delta x\frac{\partial\phi_h}{\partial x} = \Delta x\frac{\partial}{\partial x}\left\{\frac{2\pi}{\lambda}z\left[1+\left(\frac{x}{z}\right)^2+\left(\frac{y}{z}\right)^2\right]^{\frac{1}{2}}\right\} = \frac{2\pi\Delta x}{\lambda\sqrt{1+\left(\frac{z}{x}\right)^2+\left(\frac{y}{x}\right)^2}} \tag{23}$$

and, according to Eq. (22), the full resolution of $h(n\Delta x, m\Delta y, z)$ requires that

$$|\Delta\phi| \approx \frac{2\pi\Delta x}{\lambda\sqrt{1+\left(\frac{z}{x}\right)^2+\left(\frac{y}{x}\right)^2}} \le \pi \quad \Rightarrow \quad \frac{2\Delta x}{\lambda\sqrt{1+\left(\frac{z}{x}\right)^2+\left(\frac{y}{x}\right)^2}} \le 1 \tag{24}$$

This phase change reaches its maximum values for $x = x_{\max} = -X/2$ and $y = y_{\max} = -Y/2$, therefore, this necessary condition for resolution becomes

$$\frac{2}{\frac{\lambda}{\Delta x}\sqrt{1+\left(\frac{Y}{X}\right)^2+\left(2\frac{z}{X}\right)^2}} \leq 1 \Rightarrow \frac{|z|}{X} \geq \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{Y}{X}\right)^2+1\right]} \tag{25}$$

A similar reasoning applied to axis y yields a second necessary condition

$$\frac{2}{\frac{\lambda}{\Delta y}\sqrt{1+\left(\frac{X}{Y}\right)^2+\left(2\frac{z}{Y}\right)^2}} \leq 1 \Rightarrow \frac{|z|}{Y} \geq \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta y}\right)^2} - \frac{1}{4}\left[\left(\frac{X}{Y}\right)^2+1\right]}$$
$$\Rightarrow \frac{|z|}{X} \geq \sqrt{\frac{\left(\frac{Y}{X}\right)^2\left(\frac{\Delta y}{\Delta x}\right)^2}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{X}{Y}\right)^2+1\right]} \tag{26}$$

If $\left(\frac{Y}{X}\right)^2\left(\frac{\Delta y}{\Delta x}\right)^2 < 1 \Leftrightarrow \frac{Y}{X}\frac{\Delta y}{\Delta x} < 1$ Eq. (25) imposes a more strict condition to $|z|/X$ than Eq. (26) and *vice versa*. Therefore, when the convolution method is used, the resolution condition for the impulse response is

$$\frac{|z|}{X} \geq \left(\frac{|z|}{X}\right)_{\min} = \begin{cases} \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{Y}{X}\right)^2+1\right]} & \text{if} \quad \frac{Y}{X}\frac{\Delta y}{\Delta x} \leq 1 \\ \sqrt{\frac{\left(\frac{Y}{X}\right)^2\left(\frac{\Delta y}{\Delta x}\right)^2}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{X}{Y}\right)^2+1\right]} & \text{if} \quad \frac{Y}{X}\frac{\Delta y}{\Delta x} \geq 1 \end{cases} \tag{27}$$

2.2.2 Angular spectrum method: transfer function

Assuming that

$$1-\lambda^2(f_x^2+f_y^2) \geq 0 \Rightarrow \lambda < \frac{2\Delta x\Delta y}{\sqrt{\Delta x^2+\Delta y^2}} \tag{28}$$

the local phase change of the discretized transfer function $H(n\Delta f_x, m\Delta f_y, z)$ from one pixel to the next along the f_x direction is

$$\Delta\phi \approx \Delta f_x \frac{\partial \phi_H}{\partial f_x} = \Delta f_x \frac{\partial}{\partial f_x}\left\{\frac{2\pi}{\lambda} z[1-\lambda^2(f_x^2+f_y^2)]^{\frac{1}{2}}\right\} = -\frac{2\pi z\Delta f_x}{\sqrt{\frac{1}{(\lambda f_x)^2} - 1 - \left(\frac{f_y}{f_x}\right)^2}} \tag{29}$$

and, to get full resolution, Eq. (22) requires that

$$|\Delta\phi| = \frac{2\pi|z|\Delta f_x}{\sqrt{\frac{1}{(\lambda f_x)^2} - 1 - \left(\frac{f_y}{f_x}\right)^2}} \leq \pi \Rightarrow \frac{2|z|\Delta f_x}{\sqrt{\frac{1}{(\lambda f_x)^2} - 1 - \left(\frac{f_y}{fx}\right)^2}} \leq 1 \tag{30}$$

The maximum value of $|\Delta\phi_H|$ is reached for $f_x = f_{x\max} = -1/(2\Delta x)$ and $f_y = f_{y\max} = -1/(2\Delta y)$ with $\Delta f_x = 1/X$, therefore, the necessary resolution condition in f_x becomes

$$\frac{\frac{|z|}{X}}{\sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{\Delta x}{\Delta y}\right)^2+1\right]}} \leq 1 \Rightarrow \frac{|z|}{X} \leq \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}\left[\left(\frac{\Delta x}{\Delta y}\right)^2+1\right]} \tag{31}$$

and the corresponding necessary condition along the f_y axis is

$$\frac{\frac{|z|}{Y}}{\sqrt{\frac{1}{\left(\frac{\lambda}{\Delta y}\right)^2}-\frac{1}{4}\left[\left(\frac{\Delta y}{\Delta x}\right)^2+1\right]}} \le 1 \Rightarrow \frac{|z|}{Y} \le \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta y}\right)^2}-\frac{1}{4}\left[\left(\frac{\Delta y}{\Delta x}\right)^2+1\right]}$$
$$\Rightarrow \frac{|z|}{X} \le \frac{Y}{X}\frac{\Delta y}{\Delta x}\sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2}-\frac{1}{4}\left[\left(\frac{\Delta x}{\Delta y}\right)^2+1\right]} \tag{32}$$

If $\frac{Y}{X}\frac{\Delta y}{\Delta x} < 1$ the condition in Eq. (32) is stricter than Eq. (31) and *vice versa*. Therefore, with the angular spectrum method, the sufficient condition to have a fully resolved discretized transfer function is

$$\frac{|z|}{X} \le \left(\frac{|z|}{X}\right)_{\max} = \begin{cases} \frac{Y}{X}\frac{\Delta y}{\Delta x}\sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2}-\frac{1}{4}\left[\left(\frac{\Delta x}{\Delta y}\right)^2+1\right]} & \text{if} \quad \frac{Y}{X}\frac{\Delta y}{\Delta x} \le 1 \\ \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2}-\frac{1}{4}\left[\left(\frac{\Delta x}{\Delta y}\right)^2+1\right]} & \text{if} \quad \frac{Y}{X}\frac{\Delta y}{\Delta x} \ge 1 \end{cases} \tag{33}$$

2.3 Introducing the dimensionless parameters

The lengths involved in the resolution conditions for the convolution method, Eq. (27), and for the angular spectrum method, Eq. (33), have been conveniently grouped in the previous expressions to show that they are functions of just four dimensionless parameters, namely

the wavelength to horizontal pixel size ratio $$\Lambda = \frac{\lambda}{\Delta x} \tag{34}$$

the reconstruction distance to horizontal field size ratio $$Z = \frac{z}{X} \tag{35}$$

the field aspect ratio $$\Upsilon = \frac{Y}{X} \tag{36}$$

the pixel aspect ratio $$\Delta = \frac{\Delta y}{\Delta x} \tag{37}$$

Consequently, the full resolution condition for the impulse response function h (convolution method) admits the following dimensionless formulation

$$|Z| \ge |Z|_{\min} = \begin{cases} \sqrt{\frac{1}{\Lambda^2}-\frac{1}{4}(\Upsilon^2+1)} & \text{if} \quad \Upsilon\Delta \le 1 \\ \sqrt{\frac{\Upsilon^2\Delta^2}{\Lambda^2}-\frac{1}{4}(\Upsilon^2+1)} & \text{if} \quad \Upsilon\Delta \ge 1 \end{cases} \tag{38}$$

and the corresponding dimensionless full resolution condition for the transfer function H (angular spectrum method) is

$$|Z| \le |Z|_{\max} = \begin{cases} \Upsilon\Delta\sqrt{\frac{1}{\Lambda^2}-\frac{1}{4}\left(\frac{1}{\Delta^2}+1\right)} & \text{if} \quad \Upsilon\Delta \le 1 \\ \sqrt{\frac{1}{\Lambda^2}-\frac{1}{4}\left(\frac{1}{\Delta^2}+1\right)} & \text{if} \quad \Upsilon\Delta \ge 1 \end{cases} \tag{39}$$

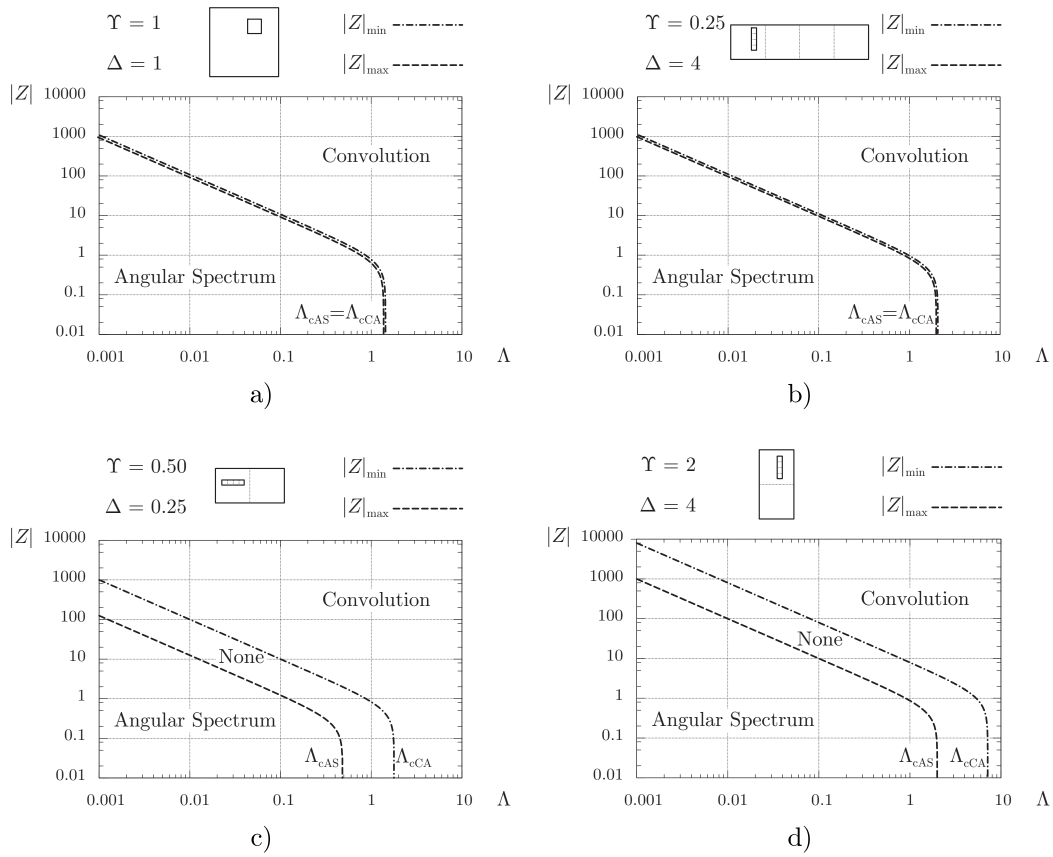


Figure 2. Examples of dimensionless resolution limits graphs for: a) square field $\Upsilon = 1$ and square pixels $\Delta = 1$, b) rectangular field and rectangular pixels with $\Upsilon\Delta = 1$, c) *idem* with $\Upsilon\Delta \leq 1$, d) *idem* with $\Upsilon\Delta \geq 1$. The resolution ranges for the *convolution* and *angular spectrum* methods are labelled, as well as the region where *none* of both methods is able to yield non-aliased results.

Given a pair of values for the field Υ and pixel Δ aspect ratios, which are generally determined by the characteristics of the sensor used to record the digital hologram, expressions (38) and (39) render the $|Z|$ and Λ aliasing-free reconstruction ranges with the corresponding methods. These resolution ranges become apparent when the resolution limits $|Z|_{\text{min}}$ and $|Z|_{\text{max}}$ are plotted versus Λ as shown in Fig. 2.

It follows from Eq. (38) that for the convolution method the impulse response function h is fully resolved regardless of the reconstruction distance when $|Z|^2_{\text{min}} \leq 0$

$$|Z|^2_{\text{min}} \leq 0 \quad \Leftrightarrow \quad \Lambda \geq \Lambda_{\text{cCA}} = \left\{ \begin{array}{ll} \dfrac{2}{\sqrt{\Upsilon^2+1}} & \text{if} \quad \Upsilon\Delta \leq 1 \\ \dfrac{2\Upsilon\Delta}{\sqrt{\Upsilon^2+1}} & \text{if} \quad \Upsilon\Delta \geq 1 \end{array} \right\} \quad \Rightarrow \quad h \quad \text{is fully resolved} \quad \forall Z \tag{40}$$

i.e., the convolution method yield non-aliased results for any reconstruction distance whenever the wavelength is Λ_{cCA} times or more longer than the horizontal pixel size Δx. On the other hand, for the angular spectrum method Eq. (39) establishes that the transfer function H cannot be fully resolved for any distance if $|Z|^2_{\text{max}} \leq 0$

$$|Z|^2_{\text{max}} \leq 0 \quad \Leftrightarrow \quad \Lambda \geq \Lambda_{\text{cAS}} = \frac{2\Delta}{\sqrt{\Delta^2+1}} \quad \Rightarrow \quad H \quad \text{is undersampled} \quad \forall Z \tag{41}$$

therefore, the angular spectrum method results will be contaminated with aliasing for any reconstruction distance if the wavelength is Λ_{cAS} times or more longer than the width Δx of the pixels.

2.4 Dimensionless formulation of the impulse response and transfer functions

Once the four dimensionless parameters (34)-(37) that determine the resolution conditions for the discretized free-space impulse response h and transfer H functions have been identified, these functions themselves can be expressed in terms of those and a fifth dimensionless parameter, namely

the number of pixels in the horizontal (x) direction $$N = \frac{X}{\Delta x} \tag{42}$$

to get a full dimensionless formulation for both numerical propagation methods.

2.4.1 Convolution method: dimensionless impulse response function

The free-space impulse response function Eq. (2) and, consequently, its discretized version Eq. (15) have dimensions of inverse area. Nevertheless, the discrete convolution reconstruction formula Eq. (17) leads us to a simple and effective way to define a "dimensionless impulse response function"

$$\bar{h}(n\Delta x, m\Delta y, z) = \Delta x \Delta y \; h(n\Delta x, m\Delta y, z) = \frac{\exp\left\{ j\frac{2\pi}{\lambda} z \left[1 + \left(\frac{n\Delta x}{z}\right)^2 + \left(\frac{m\Delta y}{z}\right)^2 \right]^{\frac{1}{2}} \right\}}{j\frac{\lambda z}{\Delta x \Delta y} \left[1 + \left(\frac{n\Delta x}{z}\right)^2 + \left(\frac{m\Delta y}{z}\right)^2 \right]} \tag{43}$$

which is perfectly fitted to be used together with the fast Fourier transform and has the same resolution limits than h. By substituting $\bar{h}$ into the reconstruction formula Eq. (17) it simplifies to

$$u(n\Delta x, m\Delta y, z) = \mathrm{FSh}(\mathrm{FFT}^{-1}\{\mathrm{FFT}[\bar{h}(n\Delta x, m\Delta y, z)] \cdot \mathrm{FFT}[u(n\Delta x, m\Delta y, 0)]\}) \tag{44}$$

and after convenient manipulation to introduce the dimensionless parameters that we have chosen into Eq. (43) we get its final formulation

$$\bar{h}(n\Delta x, m\Delta y, z) = \frac{\exp\left\{ j2\pi \frac{NZ}{\Lambda} \left[1 + \frac{1}{N^2 Z^2}(n^2 + \Delta^2 m^2) \right]^{\frac{1}{2}} \right\}}{j\frac{NZ\Lambda}{\Delta} \left[1 + \frac{1}{N^2 Z^2}(n^2 + \Delta^2 m^2) \right]} \tag{45}$$

2.4.2 Angular spectrum method: transfer function

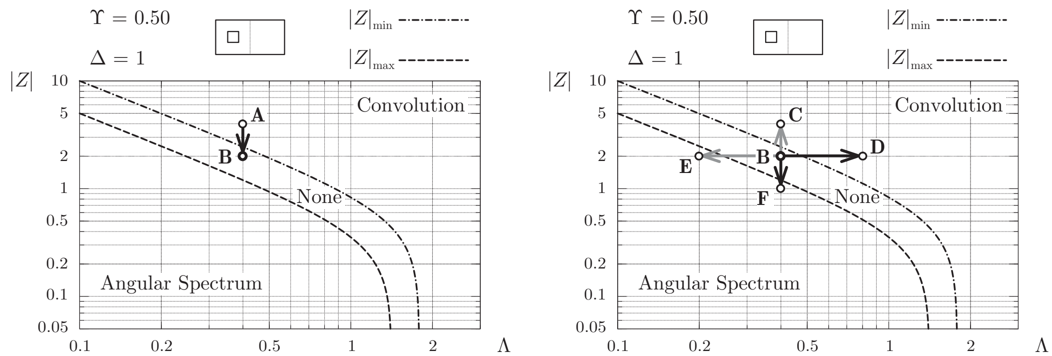
The free-space transfer function, Eq. (4), and its discretized version, Eq. (18), are intrinsically dimensionless. By convenient manipulation, our dimensionless parameters can be introduced into Eq. (18) to yield

$$H(n\Delta f_x, m\Delta f_y, z) = \exp\left\{ j2\pi \frac{NZ}{\Lambda} \left[1 - \frac{\Lambda^2}{N^2}(n^2 + \frac{m^2}{\Upsilon^2}) \right]^{\frac{1}{2}} \right\} \tag{46}$$

which can be used directly with Eq. (20) to apply the angular spectrum reconstruction method.

3. DISCUSSION

The resolution conditions for the impulse response $\bar{h}$, Eq. (38), and transfer H, Eq. (39), functions establish that both of them cannot be simultaneously resolved for any combination of the wavelength ratio Λ and the reconstruction distance ratio Z, as shown in Fig. 2. The resolution regions for the convolution and angular spectrum methods never overlap and they are sometimes separated by a no-resolution band; therefore, only one or none of the two methods will yield aliasing-free results for each particular situation. The limits of the reconstruction regions in the dimensionless resolution domain are given by the values of the field Υ and pixel Δ aspect ratios. Two situations become apparent: if $\Upsilon\Delta = 1$ —i.e., the field and pixel aspect ratios are reciprocal $\Delta = 1/\Upsilon$— the resolution limits for $\bar{h}$ and H are the same and therefore either the one or the other method will



a) b)

Figure 3. Use of the dimensionless resolution limits graph to analize: a) the effect of the sacrificial zero-padding (**A** original working point, **B** after zero-padding). b) some reconstruction-range extension strategies (**B** working point after sacrificial zero-padding, **C** hologram cropping, **D** pixel splitting, **E** pixel merging, **F** supplementary zero-padding).

yield an aliasing-free reconstruction for any working point (see Fig. 2-a and -b); on the other hand, if $\Upsilon\Delta \neq 1$ there is a no-resolution band between the two resolution regions where none of the functions ($\bar{h}$ and H) is resolved (see Fig. 2-c and -d).

Let us assume that a digital hologram generated with a laser of wavelength λ has been recorded with a given sensor with $N \times M$ pixels and size $X \times Y$, as stated in section 2, fills completely the sensor field and is critically sampled, i.e., its spectrum fills completely the frequency space. Its field Υ and pixel Δ aspect ratios are determined by the characteristics of the sensor

$$\Upsilon = \frac{Y}{X} \quad ; \quad \Delta = \frac{\Delta x}{\Delta y} = \frac{MX}{NY} \tag{47}$$

For any given distance z from the hologram plane, the position of the working point into the resolution domain (e.g., point **A** in Fig. 3-a), which is defined by the values of Λ and Z,

$$\Lambda = \frac{\lambda}{\Delta x} = \frac{M\lambda}{X} \quad ; \quad Z = \frac{z}{X} \tag{48}$$

must be analyzed before proceeding with the reconstruction either to choose the proper method accordingly or to modify the digital hologram to move the working point to the desired region. Generally speaking, for a given value of Λ the convolution method will fit long reconstruction distances and the angular spectrum method short ones, nevertheless, only the convolution method will yield aliasing-free results for large values of Λ; given a reconstruction distance Z, the convolution method will work properly for large values of the wavelength Λ and the angular spectrum method for short ones.

3.1 Effect of the sacrificial zero-padding

The digital holographic reconstruction process —Eqs. (44) and (20)— involves a convolution implemented by using a fast Fourier transform algorithm which, due to its cyclic nature, may yield aliased results even if both the digital hologram field u and the discretized impulse response $\bar{h}$ or the transfer function H are properly sampled. The most usual procedure to avoid aliasing in cyclic convolution[7] is to pad the signal with zeros up to doubling its size along each dimension before applying the FFT and remove the padded region, where aliasing effects concentrate, after the convolution process. In digital holography, this means that the edges of the original digital hologram have to be padded with zeros to build a new hologram with $X' = 2X$, $Y' = 2Y$, $N' = 2N$, $M' = 2M$ and the same pixel size $\Delta x \times \Delta y$ and aspect ratios Υ and Δ which is effectively reconstructed. The actual working point becomes

$$\Lambda' = \frac{\lambda}{\Delta x} = \Lambda \quad ; \quad Z' = \frac{z}{X'} = \frac{z}{2X} = \frac{Z}{2} \tag{49}$$

(e.g., point **B** in Fig. 3-a) and the corresponding impulse response $\bar{h}$ or transfer function H must be calculated with $N' = 2N$.

3.2 Range extension strategies

After sacrificial zero padding, the working point may fall into the no-resolution band (e.g., point **B** in Fig. 3-b) or into the resolution region of a non desired reconstruction method. The dimensionless resolution limits graph (Fig. 3) can be used to design and analyze strategies to shift the position of the working point towards the desired resolution region. The simplest of these strategies are those involving the change of just one of the four dimensionless parameters involved in the resolution conditions (Υ, Δ, Λ and Z).

There are two obvious strategies which will shift the working point towards the convolution method region. The first of them, shown as **B**→**C** in Fig. 3-b, consists in increasing the value of $|Z'| = |z|/(2X)$ by reducing both X and Y proportionally to keep Υ constant; this is not a good strategy since it means cropping part of the hologram and, therefore, losing information. The second and more convenient, represented as **B**→**D** in Fig. 3-b, is increasing the value of $\Lambda = \lambda/\Delta x$ by reducing Δx and Δy proportionally to keep Δ constant; this can be implemented either by splitting the pixels of the hologram in the direct space or by zero padding its Fourier transform in the reciprocal space.

Two simple strategies are also apparent to shift the working point into the angular spectrum region. The one, **B**→**E** in Fig. 3-b, is reducing the value of $\Lambda = \lambda/\Delta x$ by increasing Δx and Δy proportionally to keep Δ constant, i.e., merging the pixels of the hologram to make them artificially larger; this implies a loss of information as well as a risk of subsampling and, therefore, this method is not recommendable. The other, labelled **B**→**F** in Fig. 3-b, is reducing the value of $|Z'| = |z|/(2X)$ by increasing both X and Y proportionally, to keep Υ constant, with supplementary zero-padding of the hologram.

4. CONCLUSIONS

We have presented a dimensionless formulation for the reconstruction of digital holograms with the convolution method —Eqs. (44) and (45)— and the angular spectrum method —Eqs. (20) and (46)—, as well as for the resolution conditions —Eqs. (38) and (39)— that the free space impulse response and transfer functions must, respectively, satisfy to yield an aliasing-free reconstruction.

Though using dimensionless impulse response and transfer functions, this formulation is dimensionally coherent. The only dimensional term in our reconstruction formulae is the field $u(x, y, 0)$ at the hologram plane and its dimensions are directly transferred to the reconstructed field $u(x, y, z)$.

The dimensionless resolution conditions simplify tasks as comparing holographic reconstruction situations with different distance, wavelength, field and pixel sizes, deciding which of the two methods fits better a given set of reconstruction conditions and analyzing the effects of hologram manipulations such as sacrificial zero-padding or range extension strategies based on cropping, supplementary zero-padding, pixel splitting or merging.

ACKNOWLEDGMENTS

The authors acknowledge funding from the *Ministerio de Ciencia e Innovación* and the European Commission (ERDF) (project number DPI2008-02709) as well as from the *Universidad de Vigo* (contract number 09VIA07).

REFERENCES

[1] Schnars U. and Jueptner W., [Digital holography], Springer, Berlin, Ch. 3 (2005).
[2] Yamaguchi I. and Zhang T., "Phase-shifting digital holography", Opt. Lett., 22(16), 1268-1270 (1997).
[3] Gaskill J.D., [Linear Systems, Fourier Transforms and Optics], Wiley, New York, Ch. 10 (1978).
[4] Kreis T.M., Adams M. and Jüptner W.P.O., "Methods of digital holography: a comparison", Proc. SPIE, 3098, 224-233 (1997).
[5] García J., Mas D., Dorsch G., "Fractional-Fourier-transform calculation through the fast-Fourier-transform algorithm", Appl.Opt., 35(35), 7013-7018 (1996).
[6] Li J., Peng Z. and Fu Y., "Diffraction transfer function and its calculation of classic diffraction formula", Opt. Commun., 280(2), 243-248 (2007).
[7] Press W.H., Flannery B.P., Teukolsky S.A. and Vetterling W.T., [Numerical Recipes in C], Cambridge University Press, Cambridge, Ch. 12 (1988).
[8] Frigo M. and Johnson S.G., "The design and implementation of FFTW3," Proc. IEEE, 93(2), 216-231 (2005).

Corrigendum to: "Dimensionless formulation of the convolution and angular spectrum reconstruction methods in digital holography."

Ángel F. Doval and Cristina Trillo

Universidad de Vigo. Departamento de Física Aplicada. E.T.S. de Ingenieros Industriales. Campus Universitario. 36310 Vigo, Spain.

The condition to maximize the phase change between pixels for the impulse response function in section 2.2.1 (page 5) should be "$x = x_{\max} = -X/2$ and $y = y_{\min} = 0$" instead of "$x = x_{\max} = -X/2$ and $y = y_{\max} = -Y/2$", as was erroneously stated in the paper.[1] Consequently, the following changes should be made:

Section 2.2.1, top of page 5

This phase change reaches its maximum values for $x = x_{\max} = -X/2$ and $y = y_{\min} = 0$, therefore, this necessary condition for resolution becomes

$$\frac{2}{\frac{\lambda}{\Delta x}\sqrt{1+\left(2\frac{z}{X}\right)^2}} \leq 1 \Rightarrow \frac{|z|}{X} \geq \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}} \tag{25}$$

A similar reasoning applied to axis y yields a second necessary condition

$$\frac{2}{\frac{\lambda}{\Delta y}\sqrt{1+\left(2\frac{z}{Y}\right)^2}} \leq 1 \Rightarrow \frac{|z|}{Y} \geq \sqrt{\frac{1}{\left(\frac{\lambda}{\Delta y}\right)^2} - \frac{1}{4}}$$

$$\Rightarrow \frac{|z|}{X} \geq \frac{Y}{X}\sqrt{\frac{\left(\frac{\Delta y}{\Delta x}\right)^2}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}} \tag{26}$$

Both conditions, Eq. (25) and Eq. (26), must be simultaneously satisfied by $|z|/X$. Therefore, when the convolution method is used, the resolution condition for the impulse response is

$$\frac{|z|}{X} \geq \left(\frac{|z|}{X}\right)_{\min} = \max\left\{\sqrt{\frac{1}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}}\ ,\ \frac{Y}{X}\sqrt{\frac{\left(\frac{\Delta y}{\Delta x}\right)^2}{\left(\frac{\lambda}{\Delta x}\right)^2} - \frac{1}{4}}\right\} \tag{27}$$

Section 2.3, page 6, Equation (38)

$$|Z| \geq |Z|_{\min} = \max\left\{\sqrt{\frac{1}{\Lambda^2} - \frac{1}{4}}\ ,\ \Upsilon\sqrt{\frac{\Delta^2}{\Lambda^2} - \frac{1}{4}}\right\} \tag{38}$$

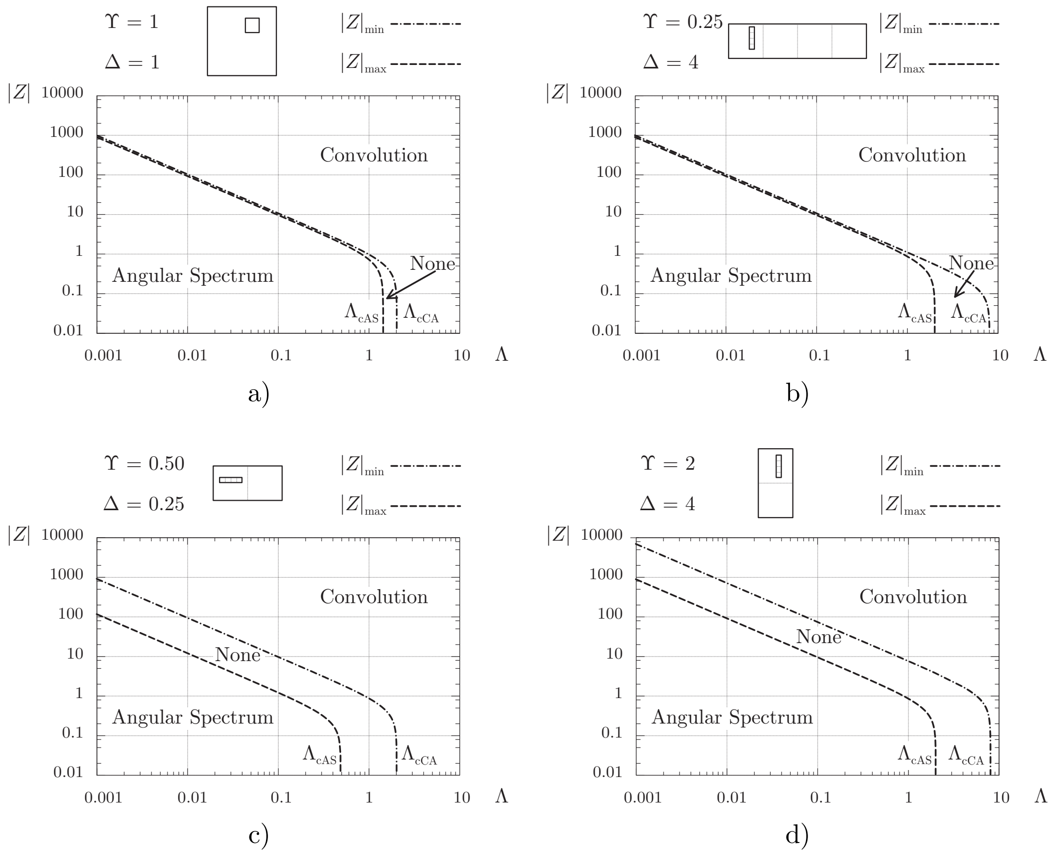
Section 2.3, page 7, Equation (40)

$$|Z|^2_{\min} \leq 0 \quad \Leftrightarrow \quad \Lambda \geq \Lambda_{\mathrm{cCA}} = \max\{2\ ,\ 2\Delta\} \tag{40}$$

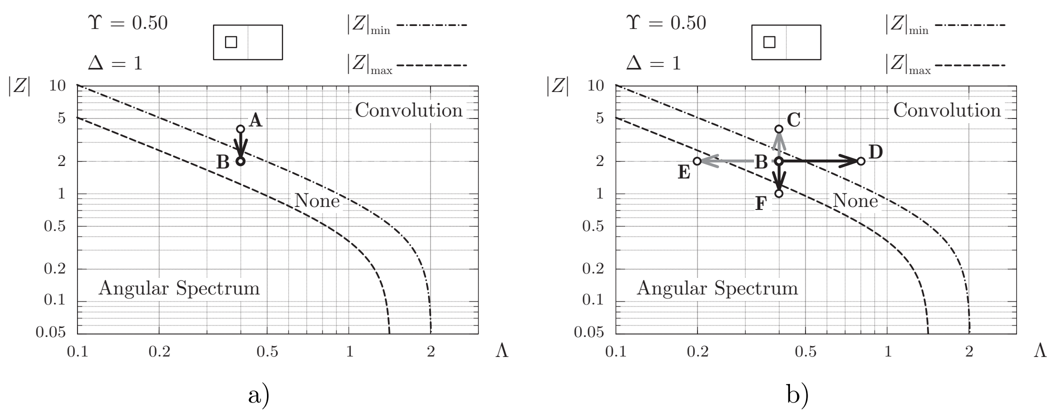
Section 3, page 8 last two lines and page 9 first line should read:

"Two situations become apparent: if $\Upsilon\Delta = 1$ —i.e., the field and pixel aspect ratios are reciprocal $\Delta = 1/\Upsilon$— the resolution limits for $\bar{h}$ and H converge for low values of Λ, such as those found in optical digital holography (typically between $\Lambda = 0.02$ and $\Lambda = 0.2$), therefore in these conditions either the one or the other method will yield an aliasing-free reconstruction for almost any working point (see Fig. 2-a and -b);"

Page 7, Figure 2



Page 9, Figure 3



A further *erratum* has been found in **Section 2.2.2, page 5, Equation (28)**, which should be

$$1 - \lambda^2(f_x^2 + f_y^2) > 0 \Rightarrow \lambda < \frac{2\Delta x \Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} \tag{28}$$

REFERENCES

[1] Doval A.F. and Trillo C., "Dimensionless formulation of the convolution and angular spectrum reconstruction methods in digital holography", in [Speckle2010. Speckle fields forever], Proc. SPIE 7387 (2010).