On the single point resolution of on-axis digital holography

Corinne Fournier,¹,³∗ Loïc Denis ² and Thierry Fournel¹

¹Laboratoire Hubert Curien, CNRS UMR 5516, Université Jean Monnet, 18 rue Pr Benoit Lauras, F-42000 Saint-Etienne, France
²Centre de Recherche Astronomique de Lyon, UMR 5574 CNRS/ENS Lyon/UCBL, Lyon, France
³TELECOM Saint-Etienne, 23 rue du Dr P. Michelon, F-42023 Saint-Etienne, France

*Corresponding author: corinne.fournier@telecom-st-etienne.fr

On-axis Digital Holography is becoming widely used for its time-resolved 3D imaging capabilities. A 3D volume can be reconstructed from a single hologram. Digital Holography (DH) is applied as a metrological tool in experimental mechanics, biology, fluid dynamics and therefore the estimation and the improvement of the resolution are current challenges. However the resolution depends on several experimental parameters such as the recording distance, the sensor definition, the pixel size and also on the location of the object in the field of view. This paper derives resolution bounds in DH by using estimation theory. The single point resolution expresses the standard deviations on the estimation of the spatial coordinates of a point source from its hologram. Cramér-Rao lower bounds give a lower limit for resolution. The closed-form expressions of the Cramér-Rao lower bounds are obtained for a point-source located on and out of the optical axis. The influence of the 3D location of the source, the numerical aperture and the signal-to-noise ratio are studied. © 2010 Optical Society of America

OCIS codes: 000.0000, 999.9999.

1. Introduction

On-axis digital holography is a metrological tool widely used in experimental mechanics, biology or fluid dynamics [1] and therefore the estimation and the improvement of resolution are key issues of this field [2–5]. As the resolution depends on several experimental parameters
(e.g. sensor definition, fill factor, recording distance) and on the image processing algorithm used to perform the 3D reconstruction, experimenters are in need of methodologies to tune the experimental setup and to select the reconstruction that will provide the best achievable resolution.

First of all, let us recall that two definitions of resolution can be distinguished [6]: (1) two point resolution is defined as the system’s capability to determine the separation of two point-sources (2) single point resolution is defined as the system’s capacity to determine the location of a point-source. The first definition is used when the imaging goal is to separate two objects or structures inside an object (for example in astronomy, in microscopy, ...). The second definition, directly linked to accuracy of measurement, is more suited for metrological applications.

Rayleigh criterion is the most cited resolution criterion in literature. It is based on a two point resolution (definition (1)) for diffraction limited systems. According to the Rayleigh criterion, two point sources are just resolved if the central maximum of the intensity diffraction pattern produced by one point source coincides with the first zero of the intensity diffraction pattern produced by the other. For an imaging system with a square aperture and numerical aperture \( \Omega \), classical Rayleigh resolution limits (lateral \( \delta_x, \delta_y \) and axial \( \delta_z \)) are well known [7]:

\[
\delta_x = \delta_y = \frac{\lambda}{\Omega}, \quad \delta_z = \frac{\lambda}{\Omega^2}
\]  

Rayleigh criterion is based on the assumption of a known PSF and negligible noise. It should be pointed out [6] that, with numerical tools, if the PSF was to be known, the two-component model of the PSF could be fitted numerically to the observations with respect to the component locations and amplitudes. In a noiseless situation, a perfect fit would result, leading to a virtually unlimited resolution in spite of diffraction. It is obvious that the study of the practically achievable resolution in imaging should take into account the noise level (whatever definition of resolution is chosen).

Resolution limits in DH have been discussed by many authors [2,5,8]. The commonly used approach for resolution estimation is to evaluate Rayleigh resolution by estimating the width of the point spread function of the digital holographic system in the reconstructed planes. The effects of diffraction are taken into account but also the sampling and pixel integration effects.

In this paper, we choose to study single point resolution (more adapted to metrological applications), taking into account the effect of the noise level in the image. We present a methodology based on parameter estimation theory (see e.g., [9]) to estimate resolution. This type of approach has already been applied to many application fields [10–15]. Here we apply it to estimate the single point resolution in on-axis DH.

The resolution estimation problem can be formulated as the estimation of a point source
location \((x, y, z)\) from the analysis of its hologram (data). As the hologram of a point source can be modeled by a parametric function (parameters \(x, y, z\)) and a model of noise, the single point resolution is estimated by mean of the variance on the parameters. Cramér-Rao lower bound (CRLB) \([16, 17]\), gives a lower bound on the variance of the estimators. This bound is reached asymptotically (for large data) by the maximum likelihood estimator. Let us note that the maximum likelihood estimator has already been used with success in DH to detect objects of simple geometrical form leading to enhanced accuracy \([18–20]\), and a widening of the field of view classically limited to the hologram borders.

The paper is organized as follows. In Section 2, we describe a parameter estimation approach to estimate a point source location \((x, y, z)\). The single point resolution is then derived from the CRLB. In section 3, we give the analytical expressions of the variance calculated on the optical axis. In section 4, analytical expressions of single point resolution in the whole field of view are given and illustrated by resolution maps. Finally, in section 5, the influence of sampling and pixel integration are discussed.

2. Statistical estimation of single point resolution

In this section, we formulate single point resolution from the view point of statistical parameter-estimation theory. We consider here the estimation problem of a point source location \((x, y, z)\) from the analysis of its hologram. The accuracy of the estimates can be determined by calculating the standard deviation of the estimates. The CRLB gives a lower bound for the variance of any unbiased estimator of the parameter to be reconstructed (e.g., \(x, y, z\)). CRLB can thus be used to calculate the theoretical resolution limit.

2.A. Single point hologram model

The (optical) system response to a point source can be described by a parametric model. We will consider the case of a hologram of a single point source recorded on an idealized square sensor of side \(L\) (sampling and quantization effects are neglected) (see Figure 1). For a point \((x_k, y_k)\) located on the sensor, the intensity of the diffraction pattern (parametric model \(g_{\theta}\)) under Fresnel approximation is a radial chirp function and depends on the axial point source location \(z\) and lateral point source location \((x, y)\):

\[
g_{\theta}(x_k, y_k) = A \sin \left( \pi \frac{(x - x_k)^2 + (y - y_k)^2}{\lambda z} \right)
\]

with \(\theta = (x, y, z)\) the vector parameters, \(\lambda\) the wavelength of the light, omitting the proportionality factor and the offset level. Due to noise, the measured image is a perturbed version of the model. We consider in first approximation the noise as white Gaussian.
Fig. 1: Illustration of the parametric model \( g_\theta(x_k, y_k) \) of an on-axis digital hologram of a point source (\( x = 0.2L, y = 0.4L, z = 100\text{mm}, \lambda = 0.532\mu\text{m}, \Omega = 8.6.10^{-3} \))

2.B. Cramér-Rao lower bound

According to the Cramér-Rao inequality, the covariance matrix of any unbiased estimator \( \hat{\theta} = \{\hat{\theta}_i\} \) of the unknown vector parameter \( \theta^* \) is bounded from below by the inverse of the so-called Fisher information matrix:

\[
\text{var} \left( \hat{\theta}_i \right) \geq \left[ I^{-1}(\theta^*) \right]_{i,i} \tag{3}
\]

where \( I(\theta^*) \) is the \( n \times n \) Fisher information matrix. It is connected with the curvature of the log-likelihood function \( \ln p(d; \theta) \) [9]:

\[
[I(\theta)]_{i,j} \overset{\text{def}}{=} -E \left[ \frac{\partial^2 \ln p(d; \theta)}{\partial \theta_i \partial \theta_j} \right]. \tag{4}
\]

The log-likelihood function in the case of additive Gaussian white noise (with variance \( \sigma_b^2 \)) is given by:

\[
\ln p(d; \theta) = -\frac{1}{2\sigma_b^2} \sum_{x_k} \sum_{y_k} [d(x_k, y_k) - g_\theta(x_k, y_k)]^2 + C \tag{5}
\]

where \( d \) represents the data (pixel values), \( g_\theta \) the parametric model depending on the parameter vector \( \theta \) and \( C \) a constant. Using (2), (4), (5) and neglecting sampling and quantization effects, Fisher matrix becomes:

\[
[I(\theta)]_{i,j} = \frac{1}{\sigma_b^2} \int_{\text{sensor}} \int \left( \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_i} \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_j} \right) \, dx_k \, dy_k \tag{6}
\]
The CRLB is asymptotically (for large samples) reached by maximum likelihood estimators. In digital holography, where the signal is distributed on the whole sensor, estimation is performed using a large set of independent identically distributed measurements (typically more than one million). The maximum likelihood estimator approaches the theoretical resolution limit given by the CRLB. In practice the log-likelihood function is maximized by numerical optimization. This amounts in this case to minimizing the squared difference between measurements and model. Note that if the minimization technique fails to reach the global minimum or if the noise level is too high, the resulting estimation error will exceed CRLB.

3. Single point resolution on the optical axis

This section aims to study the variance of the parameters \((x, y, z)\) in the simple case of a point source located on the optical axis (see Figure 1). This case is of interest because it leads to an analytical expression of the resolution which will be compared to the classical Rayleigh formula (1).

3.A. Analytical form of the Fisher information matrix

Neglecting the effect of sampling, the Fisher information matrix \(I\) can be calculated from (6). For a square sensor of side \(L\), \(I\) has the form:

\[
[I(\theta)]_{i,j} = \frac{1}{\sigma_b^2} \frac{1}{L^2} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} \left( \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_i} \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_j} \right) \bigg|_{x=0} \bigg|_{y=0} dx_k dy_k 
\]

Using a symbolic computation software, the integral (7) can be calculated with the expression of the model (2). It gives an analytical expression composed of many terms that can be significantly simplified assuming:

\[
L^2/Lz \gg 1 
\]

The simple expressions obtained are:

\[
I_{xx} = I_{yy} \approx \pi^2 A^2 \frac{L^2}{\sigma_b^2 6 \lambda^2 z^2}, \quad I_{zz} \approx \pi^2 A^2 \frac{7 L^4}{\sigma_b^2 360 \lambda^2 z^4} 
\]

Let us note that the assumption (8) is always verified in classical conditions. It means that more than 2 oscillations of the chirp function are recorded on the sensor. Due to the parity of \[\left| \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_i} \right|_{x=0} \bigg|_{y=0}\] with respect to \(x_k\) and \(y_k\), diagonal terms (\(I_{xy}, I_{xz}\) and \(I_{yz}\)) are null. This simple case of a point source located on the optical axis and square sensor leads to a diagonal Fisher information matrix.
3.B. Analytical form of the covariance matrix

As Fisher matrix is diagonal, the covariance matrix is also diagonal. Standard deviations on lateral ($\hat{\sigma}_x$, $\hat{\sigma}_y$) and axial ($\hat{\sigma}_z$) measurements are given by:

$$\hat{\sigma}_x = \hat{\sigma}_y = \frac{\lambda}{\Omega SNR} c_1$$
$$\hat{\sigma}_z = \frac{\lambda}{\Omega^2 SNR} c_2$$

(10)

with $\Omega$ the numerical aperture ($\Omega = \frac{L}{2z}$), $c_1 = \frac{\sqrt{6}}{2\pi} \approx 0.4$, $c_2 = \frac{3\sqrt{10/7}}{2\pi} \approx 0.6$ and $SNR = \frac{A}{\sigma_b}$.

Let us make some remarks about these expressions:

- Each error is inversely proportional to the signal to noise ratio.
- The covariance matrix is diagonal: the errors on the estimation of $x$, $y$ and $z$ are not correlated.
- The resolutions on $x$ and $z$ are proportional to classical resolution formulas (1) with proportionality constants depending on the SNR. However let us notice that the proportionality factors are different. Whereas the Rayleigh criterion gives $\delta_z = \delta_x / \Omega$ our study leads to $\hat{\sigma}_z = (c_2/c_1) \hat{\sigma}_x / \Omega$ with $c_2/c_1 \approx 1.5$.

4. Single point resolution map

Many studies about resolution (using various approaches) are restricted to the optical axis case. However as the truncation of the signal varies with the lateral location of the source $(x,y)$ (Figure 1.b) resolution should depend on $x$ and $y$. The non uniformity of resolution on a plane has been presented in [5]. The authors qualitatively analyze it considering the PSF (in reconstructed planes) shift variant. We will go further in a quantitative study. In this section we first neglect the sampling effect in order to derive analytical expressions of single point resolution map. In the next section we discuss the limitations of these analytical expressions and propose a numerical way to calculate the resolution maps taking into account sampling effects and pixel integration.

4.A. Analytical form of the Fisher information matrix

In the more general case where the point source is not located on the optical axis but at coordinates $(x, y, z)$, analytical expressions of resolution can be obtained using the previous approach. The model $g_\theta$ is in this case translated, this can be interpreted as a shift in integration limits:

$$[I(\theta)]_{i,j} = \frac{1}{\sigma_b^2 L^2} \int_{-\frac{L}{2}-x}^{\frac{L}{2}-x} \int_{-\frac{L}{2}-y}^{\frac{L}{2}-y} \left( \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_i} \frac{\partial g_\theta(x_k, y_k)}{\partial \theta_j} \right) |_{x=0} \ dx_k \ dy_k$$

(11)
Using a symbolic computation software and the hypothesis (8), we obtain analytical form for \( I \):

\[
I_{xx} \approx \pi^2 \frac{A^2}{\sigma_b^2} \frac{L^2}{\lambda^2 z^2} \frac{(1 + 12\bar{x}^2)}{6}
\]

\[
I_{yy} \approx \pi^2 \frac{A^2}{\sigma_b^2} \frac{L^2}{\lambda^2 z^2} \frac{(1 + 12\bar{y}^2)}{6}
\]

\[
I_{xy} = I_{yx} \approx \pi^2 \frac{A^2}{\sigma_b^2} \frac{L^2}{\lambda^2 z^2} \frac{\bar{x} \bar{y}}{2}
\]

\[
I_{zz} \approx \pi^2 \frac{A^2}{\sigma_b^2} \frac{L^4}{\lambda^2 z^4} \frac{(7 + 120(\bar{x}^2 + \bar{y}^2) + 180(\bar{x}^2 + \bar{y}^2)^2)}{360}
\]

\[
I_{xz} = I_{zx} \approx -\pi^2 \frac{A^2}{\sigma_b^2} \frac{L^3}{\lambda^2 z^3} \frac{\bar{x}}{3} \left(1 + \bar{x}^2 + \bar{y}^2\right)
\]

\[
I_{yz} = I_{zx} \approx -\pi^2 \frac{A^2}{\sigma_b^2} \frac{L^3}{\lambda^2 z^3} \frac{\bar{y}}{3} \left(1 + \bar{x}^2 + \bar{y}^2\right)
\]

(12)

with \( \bar{x} = \frac{x}{L} \) and \( \bar{y} = \frac{y}{L} \) the non-dimensional coordinates.

Let us note that \( I \) is no more diagonal and therefore the estimation errors on \( x, y \) and \( z \) are correlated.

4.B. Analytical form of the covariance matrix

The covariance matrix is obtained by inverting the Fisher information matrix. By introducing \( \tilde{\sigma}_x, \tilde{\sigma}_y \) et \( \tilde{\sigma}_z \) obtained in the simple case of a point source located on the optical axis (10), and a non-dimensional radial coordinate \( \tilde{\rho} = \sqrt{\bar{x}^2 + \bar{y}^2} \), the covariance matrix can be written:

\[
\begin{pmatrix}
(1 + 156\frac{\bar{x}^2}{K}) \tilde{\sigma}_x^2 & 156\frac{\bar{x} \bar{y}}{K} \tilde{\sigma}_x \tilde{\sigma}_y & 60\sqrt{\frac{7}{15}} \bar{x} \frac{1+3\tilde{\rho}^2}{K} \tilde{\sigma}_x \tilde{\sigma}_z \\
156\frac{\bar{x} \bar{y}}{K} \tilde{\sigma}_x \tilde{\sigma}_y & (1 + 156\frac{\bar{y}^2}{K}) \tilde{\sigma}_y^2 & 60\sqrt{\frac{7}{15}} \bar{y} \frac{1+3\tilde{\rho}^2}{K} \tilde{\sigma}_y \tilde{\sigma}_z \\
60\sqrt{\frac{7}{15}} \bar{x} \frac{1+3\tilde{\rho}^2}{K} \tilde{\sigma}_x \tilde{\sigma}_z & 60\sqrt{\frac{7}{15}} \bar{y} \frac{1+3\tilde{\rho}^2}{K} \tilde{\sigma}_y \tilde{\sigma}_z & (1 + 60\tilde{\rho}^2 \frac{2-3\tilde{\rho}^2}{K}) \tilde{\sigma}_z^2
\end{pmatrix}
\]

(13)

where \( K = 7 - 36\tilde{\rho}^2 + 180\tilde{\rho}^4 \) depends only on the radial coordinate.

4.C. Standard deviation maps

Standard deviations are given by the square roots of the 3 diagonal terms of the covariance matrix.
\[
\sigma_x = \hat{\sigma}_x \sqrt{1 + \frac{156 \bar{x}^2}{7 - 36 \bar{\rho}^2 + 180 \bar{\rho}^4}} \\
\sigma_y = \hat{\sigma}_y \sqrt{1 + \frac{156 \bar{y}^2}{7 - 36 \bar{\rho}^2 + 180 \bar{\rho}^4}} \\
\sigma_z = \hat{\sigma}_z \sqrt{1 + \frac{60 \rho^2}{7 - 36 \rho^2 + 180 \rho^4}}
\]

\(\bar{x}, \bar{y}\) represent the lateral location of the point source. The accuracy to estimate the location of a point source varies with the lateral location \((\bar{x}, \bar{y})\) of the point source. This variation is not negligible: on the x error map the maximum is 2.0 times the minimum, while on the z error map the maximum is 3.7 times the minimum.

By using non-dimensional variables \(\bar{x}\) and \(\bar{y}\), the obtained error maps are invariant up to a multiplicative constant when \(z\) or \(L\) change. These analytical results have been compared with numerical integrations of equation (11) (see section 5) to make sure that hypothesis (8) is valid. The relative error on the error maps is negligible (lower than 0.3 \%).

5. Influence of sampling and pixel integration

DH reconstruction are in practice transversally limited to the hologram support. Within an inverse problems framework, field extrapolation can be performed to reconstruct fields with
a transversal size up to four times the sensor size [20, 21]. Field extrapolation is challenging since the signal of out-of-field objects recorded on the sensor is weak due to pixel integration and attenuation due to diffraction lobes.

The simplified model (2) used to derive the expressions of resolution (eq. 13-16) considers that the amplitude of the signal is independent of the transversal location of the point source, i.e., the signal is not attenuated even when the point source is located far away from the sensor boundaries. The obtained expressions of resolution can not be used to assess the asymptotical behavior of resolution in extrapolated fields. At the limit of an x or y location at infinity, the formulas give a null variance due to the absence of signal attenuation and sampling considerations.

We study now the impact of sampling and pixel integration in the resolution maps obtained in the previous section. When taking into account pixel integration the model $g_\theta(x, y)$ of the diffraction pattern is a chirp function modulated by two sinc functions [1]:

$$g_\theta(x, y) = g_\theta(x_k, y_k) \text{sinc} \left( \frac{\pi \kappa_x \Delta_x (x - x_k)}{\lambda z} \right) \text{sinc} \left( \frac{\pi \kappa_y \Delta_y (y - y_k)}{\lambda z} \right)$$  \hspace{1cm} (17)$$

with $\Delta_x$ and $\Delta_y$ the sampling periods and $\kappa_x \Delta_x$ and $\kappa_y \Delta_y$ the width and the height of active area of the pixel. The integral of eq (6) cannot be computed in closed form for this model. We therefore consider the definition of Fisher information matrix (4), using the sampling given by the detector:

$$[I(\theta)]_{i,j} = \frac{1}{\sigma^2} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{M} \left( \frac{\partial g_\theta^p(i \Delta_x, j \Delta_y)}{\partial \theta_i} \frac{\partial g_\theta^p(i \Delta_x, j \Delta_y)}{\partial \theta_j} \right)$$  \hspace{1cm} (18)$$

where N is the number of rows and M number of columns of the sensor. Equation (18) can be computed for all $\bar{x}$ and $\bar{y}$ using fast Fourier transforms by noting that it expresses the correlation between a binary mask (defined as equal to one on the sensor and zero outside) with a product of model gradients $\left( \left| \frac{\partial g_\theta^p}{\partial \theta_i} \frac{\partial g_\theta^p}{\partial \theta_j} \right|_{\bar{x}=0, \bar{y}=0} \right)$. The error maps calculated numerically from (18) with the same experimental parameters as in section 4.C are closed to that of Figure 2. The relative differences (expressed in percent) between the two error maps are presented on Figure 3. It shows that the differences in the classical field (limited by the sensor boundaries) are less than 6 % and reaches 16 % in an extended field of two times the sensor size. When the pixel number increases (with constant width $L$) these differences decrease: for $N = 512$ it is less than 1.5 % in the classical field and less than 4 % in the extended field.

Let us note first that, in agreement with intuition, when the pixel integration is taken into account, the errors tend to infinity when the point source move away from the optical axis since the signal tends to zero. Second, when taking into account pixel integration, the shape
Fig. 3: Relative differences (%) between resolution maps (Figure 2) and resolution maps taking into account pixel integration for $z = 100\, \text{mm}$, $\lambda = 0.532\, \mu\text{m}$, $\Omega = 8.6 \times 10^{-3}$, $\sigma_b = 0.1$, $\Delta_x = \Delta_y = 6.7\, \mu\text{m}$, $\kappa_x = \kappa_y = 1$ and $N = 256$. The squares in the center of the figures represent the sensor boundaries.

of the error maps depends on $z$ and $L$ contrary to the ideal sensor case of section 3 and 4 (where sampling is neglected).

6. Conclusion

On-axis digital holography is used in metrological applications, thus the resolution issues are essential in this field. Although DH resolution has already been studied by many authors, quantitative studies have been limited to the resolution on the optical axis. It is well known that resolution varies with the transversal location, but it has never been precisely quantified to our knowledge. In this paper we proposed a methodology to study resolution maps for DH over the whole field of view. By using a statistical framework and a parametric model of the image we calculate resolution maps from CRLB. These resolution limits give the best resolutions achievable by any unbiased estimator. These bounds are reached asymptotically (for large data) by a maximum likelihood estimator such as proposed in [18–20]. Note that the errors that occur when global optimization fails should also be considered (i.e., inability to numerically compute the maximum likelihood estimator) [18].

Without taking into account the effects of sampling, closed form expressions of resolutions can be derived. These expressions, quickly computed, give a good approximation of resolution in the classical field of view (limited by the sensor boundaries) with typical experimental setup parameters. The simple model of image formation can be made more complex (e.g. taking into account the sampling and the pixel integration), in that case numerical integration
of the gradient of the model can be computed by FFT.

The methodology employed here to study resolution in DH can be applied to study many important issues in DH by refining the noise model or changing the image formation model. For example, it can be used to estimate accuracy lower bound of object location in DH. The parametric model depends on the position parameters \((x,y,z)\) of the object but also on the size and shape of the object. All these parameters are correlated and a study of error maps can be used to design an optimal setup (with respect to the accuracy of object location estimation). An other example is to estimate the resolution improvement comparing different setups/ optical techniques as off-axis DH, phase shifting DH or color DH.

Recently, maximum a posteriori (MAP) approaches have been shown to lead to reconstructions with few artifacts [22–24]. The theoretical study of the resolution of the obtained 3D reconstructions is challenging, though. The methodology and results described in the paper provide a lower bound on the resolution achievable with MAP techniques.

Acknowledgments

The authors are grateful to P. Refregier for fruitful discussions.
References


