

# **Discrete diffraction transform ( $DDT$ ) for propagation, reconstruction and design of wavefield distributions**

Vladimir Katkovnik, Jaakko Astola, Karen Egiazarian

Signal Processing Institute, University of Technology of Tampere,

P. O. Box 553, Tampere, Finland, [katkov@cs.tut.fi](mailto:katkov@cs.tut.fi).

A discrete diffraction transform ( $DDT$ ) is a novel discrete wavefield propagation model which is aliasing free for a pixel-wise invariant object distribution. For this class of the distributions the model is precise and has no typical discretization effects because it corresponds to accurate calculation of the diffraction integral. A spacial light modulator ( $SLM$ ) is a good example of the system where a pixel-wise invariant distribution appears. Frequency domain regularized inverse algorithms are developed for reconstruction of the object wavefield distribution from the distribution given in the sensor plane. The efficiency of developed frequency domain algorithms is demonstrated by simulation. ©

2008 Optical Society of America

*OCIS codes:* 070.2025, 100.3010, 100.3190

## 1. Introduction

The wavefield reconstruction from intensity and phase measurements is one of the basic problems in digital wavefield imaging and holography [1]. In a typical holography scenario one distinguishes input/object and output/sensor planes (see Fig.1 ). The object plane is a source of light radiation/reflection

propagating along the optical axis. The sensor plane is parallel to the object plane with a distance  $z$  between the planes. The problem is to reconstruct the wavefield in the input/object plane from data in the output/sensor plane. Some of the usual algorithms mimic the optical procedure when the hologram is illuminated with the reference wave. In digital versions of this type of the methods a recorded hologram is multiplied with a complex-valued distribution modeling the reference wave and the diffraction field is calculated. Three images can be obtained simultaneously in this method: real object in-focus, virtual object out-of-focus and low frequency zero order term. Further processing is required in order to separate these images. The advantage of the procedure is that a single record of the hologram is required [1].

Different types of the algorithms can be used if we assume that wavefield information is available at the sensor plane. The corresponding complex-valued (intensity-phase) measurements can be obtained, in particular, using the phase-shifting techniques, where few phase-shifting holograms should be recorded [1], [2].

A contribution of this paper concerns a few aspects of digital modelling for optical wavefield propagation, reconstruction and design. The first is a

novel discrete diffraction transform ( $DDT$ ) model of the wavefield propagation presented in Section 3. The model is accurate for a pixel-wise invariant object distribution, i.e. for a distribution which is piece-wise constant on rectangular elements of pixel's size of the digital sensor. This discrete modelling gives accurate and aliasing free results for any pixel's size. Here we refer to the aliasing appearing as a result of discretization in the standard discrete approximations of the wavefield propagation integrals. This precise modelling of the wavefield propagation imitates an "ideal" physical experiment as it is presented by the integral propagation equations. This sort of accurate modeling is useful for reliable design and computer testing of digital holography algorithms because these design and tests can be based on the precisely modelled observations. For an arbitrary continuous object distribution the proposed wavefield propagation modelling is approximate with the accuracy depending on the accuracy of the piece-wise fitting of the object wavefield distribution. In this case  $DDT$  continues to be accurate but for the piece-wise approximation of the object distribution. A spacial light modulator ( $SLM$ ) generating a pixel-wise invariant wavefield distribution is a good example of the optical system where the developed digital modelling

is relevant.

The *DDT* approach is different from the standard digital models usual for the holographic data. It is appeared from a signal/image processing view of the problem with intention to apply and develop the signal/image techniques for digital holography. The applications presented in this paper concern a wavefield reconstruction considered in two settings: object distribution reconstruction from a wavefield distribution given in the sensor plane (Section 4) and object distribution design with the goal to obtain a desirable wavefield distribution at the sensor plane (Sections 6). The developed frequency domain regularized inverse and recursive regularized inverse techniques show a very good performance in simulation experiments where the observed data are obtained according to the proposed accurate forward propagation modelling. The approach presented in this paper is originated from [3].

## **2. Standard models of wavefield propagation**

Consider wavefield propagation modelling for a setting shown in Fig. 1. Let  $u_z(x, y)$  be a complex-valued  $2D$  wavefield defined in the sensor plane  $z$  of the  $3D$  space  $(x, y, z)$  as a function of the lateral coordinates  $x$  and  $y$ . According to the scalar diffraction theory there is a linear operator which links this wave-

field with the object wavefield  $u_0(x, y)$  at  $z = 0$  as  $u_z(x, y) = \mathcal{D}_z\{u_0(x, y)\}$ , where  $\mathcal{D}_z$  stands for a diffraction operator with a distance parameter  $z$ . The thorough theory of this operator representation can be found in [4].

The diffraction operator can be given as a convolution

$$u_z(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_z(x - \xi, y - \eta) u_0(\xi, \eta) d\xi d\eta, \quad (1)$$

where the kernel  $g_z$  is shift invariant and has a form of the first Rayleigh-Sommerfeld solution of the Maxwell-Helmholtz equation  $\nabla^2 u + k^2 u = 0$ . Here  $k = 2\pi/\lambda$  is the wave number and  $\lambda$  is the wavelength.

This kernel is as follows [5]

$$g_z = z \frac{\exp(j2\pi r/\lambda)}{j\lambda \cdot r^2}, \quad r = \sqrt{x^2 + y^2 + z^2}. \quad (2)$$

It is shown in [4] that the operator  $\mathcal{D}_z$  is invertible, and this inverse operator also can be presented as a convolution with a shift-invariant kernel. If the diffraction wavefield  $u_z(x, y)$  is given then the wavefield at the input object plane  $z = 0$  can be reconstructed using the inverse operator  $\mathcal{D}_z^{-1}$ , i.e.  $u_0(x, y) = \mathcal{D}_z^{-1}\{u_z(x, y)\}$ .

The terms *diffraction transforms* are used in [4] for these forward  $\mathcal{D}_z$  and backward  $\mathcal{D}_z^{-1}$  operators. Thus, the reconstruction problem is reduced to calculation of the inverse transform and the corresponding wavefield.

One of the fundamental results of the scalar diffraction theory (with two different proofs presented in [4] and [5]) is that the 2D integral Fourier transform of the kernel  $g_z$  has a form

$$G_z(f_x, f_y) = \mathcal{F}\{g_z(x, y)\} = \begin{cases} \exp(j2\pi \frac{z}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}), & (f_x, f_y) \in D_\lambda, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Here  $f_x$  and  $f_y$  are the Fourier transform frequencies and the disc  $D_\lambda = \{(f_x, f_y) : (\lambda f_x)^2 + (\lambda f_y)^2 < 1\}$  defines the area where the transfer function  $G_z$  of the operator is non-zero.

In the frequency domain the forward and inverse diffraction transforms are given by

$$U_z(f_x, f_y) = G_z(f_x, f_y)U_0(f_x, f_y), \quad (4)$$

$$U_0(f_x, f_y) = G_z^*(f_x, f_y)U_z(f_x, f_y), \quad (5)$$

where  $U_z(f_x, f_y) = \mathcal{F}\{u_z(x, y)\}$  and  $G_z^*(f_x, f_y)$  is a complex conjugate of  $G_z(f_x, f_y)$ ,  $G_z^*(f_x, f_y) = G_{-z}(f_x, f_y)$ . It proves that the inverse diffraction transform can be obtained from the forward one by replacing  $z$  by  $-z$  in the kernel and in the transfer function.

Thus, the diffraction transform with the kernel  $g_z$  is a lowpass filter with

the amplitude characteristic  $|G_z(f_x, f_y)| = 1$  and with the phase  $\psi_{G_z} = 2\pi \frac{z}{\lambda} \sqrt{1 - (\lambda\omega_x)^2 - (\lambda\omega_y)^2}$  on the disc  $D_\lambda$ .

As  $|G_z(f_x, f_y)| = 1$  the diffraction transform is not only invertible but well-conditioned on  $D_\lambda$ . This inverse does not require regularization typical for ill-conditioned inverse problems.

If the spectrums  $U_z(f_x, f_y)$ ,  $U_0(f_x, f_y)$  are given, the corresponding wave-fields are calculated by the inverse integral Fourier transform for  $U_z(f_x, f_y)$  having the form

$$u_z(x, y) = \int \int U_z(f_x, f_y) \exp(-j2\pi(f_x x + f_y y)) d\omega_x d\omega_y. \quad (6)$$

Calculating the inverse Fourier transform from (5) we derive an analytical expression for the kernel of the inverse diffraction transform as  $g_{-z}$ , indeed

$$u_0(x, y) = \mathcal{D}_z^{-1}\{u_z(x, y)\} = \int \int g_{-z}(x - \xi, y - \eta) u_z(\xi, \eta) d\xi d\eta = \mathcal{D}_{-z}\{u_z(x, y)\}. \quad (7)$$

Discretization of the *integral diffraction transforms* defined by the formulas (1), (6) is a natural idea in order to derive the digital models for the forward and backward wavefield propagation. It is well known that this discretization is far from being trivial.



The principal difficulty of discretization in the space domain (1) follows from the fact that the kernel  $g_z$  is modulated by a strongly fluctuated high-frequency harmonic factor  $\exp(j2\pi r/\lambda)$ . The discrete sampling of the rate at least twice higher than the highest frequency component of the integrand is a standard remedy. Obviously it can result in an unacceptably high sampling rate.

The integration in the frequency domain in (6) looks as an attractive alternative because for  $|f_x| \ll 1/\lambda$ ,  $|f_y| \ll 1/\lambda$  the phase  $\psi_{G_z} \simeq 2\pi \frac{z}{\lambda} - \pi z \lambda (f_x + f_y)$  is slowly varying in  $f_x$ ,  $f_y$ . However, the formula (3) is derived assuming that the sensor domain is infinite in the lateral coordinates. A finite (usually small size) sensor means that the infinitely extended sensor plane is multiplied by the  $2D$  rectangular window function, which has a magnitude of one for the sensor area and zero elsewhere. In the frequency domain this multiplication causes the convolution of the spectrum  $U_z(f_x, f_y)$  in (6) with the  $2D$  sinc function.

The discrete modeling of the diffraction transform is a subject of many publications. The review of this area is beyond the scope of this paper. However, we wish to mention that the discrete space domain modelling for hologra-

phy is discussed in details in [1] and the accuracy of the frequency domain approach is analyzed in [6]. The fast Fourier transform (*FFT*) based algorithm with a detailed accuracy analysis is presented in [7]. The stationary phase method applied for integration in (6) results in discrete models with a non-uniform sampling in the frequency domain (e.g. [8]). A number of recent developments concern continuous and discrete Fresnel transforms and their multiresolution versions (e.g. [9], [10], [11]).

One of the simple standard convolution based discrete models is obtained from the continuous domain Fourier representations (4)-(5). For square object and sensor arrays of  $N \times N$  pixels with pixel size  $\Delta \times \Delta$  this discrete model follows from (4)-(5), where *FFT* is used instead of the integral Fourier transform (e.g. [1], [2]):

$$\bar{U}_z(v_x, v_y) = \bar{G}_z(v_x, v_y)\bar{U}_0(v_x, v_y), \quad (8)$$

$$\bar{U}_0(v_x, v_y) = \bar{G}_z^*(v_x, v_y)\bar{U}_z(v_x, v_y). \quad (9)$$

where  $\bar{U}_z = \mathcal{FFT}\{u_z\}$ ,  $\bar{U}_0 = \mathcal{FFT}\{u_0\}$  and  $\bar{G}_z = \mathcal{FFT}\{g_z\}$  are calculated over the  $N \times N$  arrays, and  $v_x$  and  $v_y$  are the *FFT* frequencies.

### 3. Discrete diffraction transform

The standard techniques, in particular mentioned above, consider the discrete models as approximation for underlying continuous ones given in continuous variables and in integral forms. The sampling and aliasing issues are of importance for these approximations as they define preconditions for reconstruction of continuous signals from sampled ones.

Motivation for the discrete model developed in this paper is different. Let us assume that the input distribution is discrete defined by a pixel-wise constant object wavefield distribution and the output is also discrete as defined by the outputs of the sensor pixels. For these pixel-wise constant modelling we accurately integrate the propagation equation (1). Thus, we arrive to the *discrete-to-discrete* modeling where the inputs are pixel values of the object distribution and the output are the pixel values of the digital sensor output. This model is precise as the accurate integration of (1) is assumed.

In this development the standard assumptions concerning the sampling and bandlimitedness can be omitted as they are replaced by the hypothesis that the object plane distribution is piece-wise constant or well approximated by this sort of distributions. A relevant interesting discussion on similar issues

of signal modelling for wavefield propagation can be seen in [9].

We name a proposed discrete propagation model *discrete diffraction transform* (*DDT*). It is represented in space and frequency domains. We give the corresponding formulas assuming that the object and sensor arrays are of  $N_0 \times N_0$  and  $N_z \times N_z$  pixels, respectively and the pixels are square of the equal size  $\Delta \times \Delta$ .

### 3.A. Space domain DDT

Let us derive the formulas for *DDT*. Assume that the distribution  $u_0$  allows a pixel-wise constant approximation, i.e. the distribution is piece-wise invariant on rectangular elements of pixel's size of the digital sensor. Then the integral (1) can be represented as

$$u_z(x, y) = \sum_{s, t=-N_0/2}^{N_0/2-1} u_0[s, t] \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} g_z(x - s\Delta + \xi, y - t\Delta + \eta) d\xi d\eta, \quad (10)$$

$$u_0[s, t] = u_0(s\Delta + \xi, t\Delta + \eta), \quad -\Delta/2 \leq \xi, \eta < \Delta/2,$$

where the sum is calculated over the square array of  $N_0 \times N_0$  pixels.

Let the output signal of a sensor's pixel be the mean value of the distribution impinging on this pixel:

$$u_z[k, l] = \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} u_z(k\Delta + \xi', l\Delta + \eta') d\xi' d\eta'. \quad (11)$$

Inserting (10) into (11) we arrive to the *space domain DDT* given as a discrete convolution

$$u_z[k, l] = \sum_{s, t=-N_0/2}^{N_0/2-1} a_z[k-s, l-t] u_0[s, t], \quad (12)$$

$$k, l = -N_z/2 + 1, \dots, N_z/2 - 1,$$

where the kernel is calculated as

$$a_z[k, l] = \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} g_z(k\Delta + \xi' + \xi, l\Delta + \eta' + \eta) d\xi d\eta d\xi' d\eta', \quad (13)$$

$$k, l = -N_a/2 + 1, \dots, N_a/2 - 1, \quad N_a = N_0 + N_z.$$

The kernel  $a_z$  in (12) is a smoothed (double-smoothed) version of the original kernel  $g_z$  in (1). The smoothing in (11) takes into consideration discretization for both the object and sensor arrays.

It is obvious that for a pixel-wise constant object distribution  $u_0$  the model (12)-(13) is accurate, i.e. it gives the precise distribution for the sensor plane. Note, that this statement is valid provided that the pixel grids in the object

and sensor planes are identical and perfectly aligned.

For small  $\Delta$ , such that  $\Delta/z \rightarrow 0$  and  $\Delta/\lambda \rightarrow 0$

$$a_z[k, l] \rightarrow_{\Delta \rightarrow 0} \Delta^2 g_z(k\Delta, l\Delta). \quad (14)$$

It means for the kernel smoothing effect becomes negligible for small  $\Delta$  and  $a_z$  can be calculated using instead of (13) the simple formula (14), where  $g_z$  is given in (2).

The space domain model (12) can be rewritten in the vector-matrix form

$$\mathbf{u}_z = \mathcal{A}_z \cdot \mathbf{u}_0, \quad (15)$$

where  $\mathbf{u}_z \in R^{N_z^2}$  and  $\mathbf{u}_0 \in R^{N_0^2}$  are vectors formed from  $u_z[k, l]$  and  $u_0[k, l]$ , respectively, and the  $\mathcal{A}_z$  matrix  $N_z^2 \times N_0^2$  is formed from  $a_z[k, l]$ . If  $\mathcal{A}_z$  is a full rank matrix and  $N_z \geq N_0$  a perfect reconstruction of  $\mathbf{u}_0$  from  $\mathbf{u}_z$  is possible

$$\mathbf{u}_0 = (\mathcal{A}_z^H \mathcal{A}_z)^{-1} \mathcal{A}_z^H \mathbf{u}_z, \quad (16)$$

where  $\mathcal{A}_z^H$  is Hermitian conjugate of  $\mathcal{A}_z$ .

However, for usual sensor sizes (say  $512 \times 512$  or  $1024 \times 1024$ ) the dimension of the space domain model (12) becomes too high for practical calculations for both the forward prediction of  $\mathbf{u}_z$  from  $\mathbf{u}_0$  and the backward inverse with reconstruction of  $\mathbf{u}_0$  from  $\mathbf{u}_z$ . Note that the matrix  $\mathcal{A}_z$  is non-sparse, and the

sparsity cannot be applied for more efficient calculations.

### 3.B. Frequency domain DDT

*FFT* cannot be used directly for calculation of the convolution (12) even when  $N_0 = N_z = N$  because in this case while the input  $u_0$  and output  $u_z$  have the same size  $N \times N$  the support size of the kernel  $a_z$  is much larger  $(2N - 1) \times (2N - 1)$ .

In order to make *FFT* applicable we use zero-padding of the variables  $u_0$  and  $u_z$  extended to the extended size  $N_a \times N_a$ ,  $N_a = N_0 + N_z$ , covering the support of the kernel  $a_z$ . In what follows the wave-tilde ( $\tilde{\phantom{x}}$ ) means the corresponding variables extended by zero-padding to the size  $N_a \times N_a$ . The proposed frequency domain algorithm works with these extended size variables and exploits the advantage of *FFT* for fast and accurate calculation of the convolution (12).

The *frequency domain* calculation of *DDT* is produced according the following four step algorithm:

- Define  $FFT$  for  $u_0$  extended to the size  $N_a \times N_a$

$$\begin{aligned} \tilde{U}_0(v_x, v_y) = \mathcal{FFT}\{\tilde{u}_0\} &= \sum_{k,l=-N_0/2}^{N_0/2-1} u_0[k, l] W^{v_x k} W^{v_y l} = \\ &\sum_{k,l=-N_a/2}^{N_a/2-1} \tilde{u}_0[k, l] W^{v_x k} W^{v_y l}, \end{aligned} \quad (17)$$

$$W = \exp(-j2\pi/(2N)), \quad v_x, v_y = -N_a/2, \dots, N_a/2 - 1;$$

- Take  $FFT$  of the  $DDT$  kernel to obtain the  $DDT$  transfer function

$$\begin{aligned} \tilde{A}_z(v_x, v_y) = \mathcal{FFT}\{\tilde{a}_z\} &= \sum_{u,v=-N_a/2+1}^{N_a/2-1} a_z[u, v] W^{v_x u} W^{v_y v} = \\ &\sum_{u,v=-N_a/2}^{N_a/2-1} \tilde{a}_z[u, v] W^{v_x u} W^{v_y v}; \end{aligned} \quad (18)$$

- Calculate  $FFT$  for  $u_z$  extended to the size  $N_a \times N_a$

$$\tilde{U}_z(v_x, v_y) = \tilde{A}_z(v_x, v_y) \tilde{U}_0(v_x, v_y), \quad (19)$$

$$\tilde{u}_z[k, l] = \mathcal{FFT}^{-1}\{\tilde{U}_z(v_x, v_y)\},$$

$$k, l = -N_a/2, \dots, N_a/2 - 1;$$

- Calculate  $u_z$  of the original size  $N_z \times N_z$

$$u_z[k, l] = \tilde{u}_z[k, l], \quad k, l = -N_z/2, \dots, N_z/2 - 1. \quad (20)$$

The following proposition states that the space domain formula (12) and the frequency domain algorithm (17)-(20) give the identical results.



**Proposition 1.** *The formulas (12) and (20) give identical  $u_z[k, l]$ ,  $k, l = -N_z/2, \dots, N_z/2 - 1$ , for any  $u_0[s, t]$ ,  $s, t = -N_0/2, \dots, N_0/2 - 1$ , and for any even  $N_z, N_0$ .*

The proof is given in Appendix 1.

Note, that both the space and frequency domain *DDT* models are symmetrized with respect to the optical axis.

The frequency domain *DDT* is essentially different from its frequency integral counterpart (6) in two important aspects. *DDT* accurately takes into consideration a finite size of the sensor and as a result the *DDT* operator in general is ill-conditioned while the module of the transfer function  $U_z(f_x, f_y)$  in (6) is equal to one and the corresponding operator is well-conditioned. In the frequency domain the conditioning of the *DDT* is measured as the ratio of the maximal-to-minimal value of the module of the transfer function [13]

$$cond_{\tilde{A}} = \max_{v_x, v_y} |\tilde{A}_z(v_x, v_y)| / \min_{v_x, v_y} |\tilde{A}_z(v_x, v_y)|. \quad (21)$$

Let the sensor and object planes have the size  $0.01 \times 0.01$  m and the wavelength  $\lambda = .632$  nm. Fig.2 shows the dependence of the conditioning number  $cond_{\tilde{A}}$  with respect to the distance  $z$  for  $N_0 = N_z = 512$  ( $\Delta = 0.01/512$  m) and  $N_0 = N_z = 1024$  ( $\Delta = 0.01/1024$  m). The conditioning number is ap-

proximately linear growing function of  $z$ . The larger  $N$  and smaller  $\Delta$  result in larger conditioning numbers. For the distances about  $z = 1$  the conditioning numbers are of the order  $10^4$ .

The amplitude and phase frequency characteristics of the transfer functions  $\tilde{G}_z(v_x, v_y)$  for  $\Delta = 0.01/1024$  m and small  $\Delta \rightarrow 0$ ,  $z = 0.75$  m are demonstrated in Fig. 3 and Fig. 4, respectively. The images are centered with respect to the zero-frequency  $v_x = v_y = 0$ .

These amplitude characteristics are typical for lowpass filters. The amplitude characteristic of  $\tilde{G}_z(v_x, v_y)$  with  $\Delta = 0.01/1024$  m visually is smoother than that for small  $\Delta \rightarrow 0$  when the smoothing effects of the averaging over the pixel surface disappear. The phase characteristics of  $\tilde{A}_z$  and  $\tilde{G}_z$  are very different even though their images are similar in structure.

### 3.C. Fresnel approximation

If  $z \gg x, y$  the Fresnel approximation of the kernel  $g_z$  is valid,

$$g_z \simeq \frac{1}{j\lambda \cdot z} \exp[j(\frac{2\pi z}{\lambda} + \frac{\pi}{\lambda z}(x^2 + y^2))]. \quad (22)$$

Then the formula (13) can be simplified and allows the following factorization

$$a_z[k, l] \simeq \frac{\exp(j2\pi z/\lambda)}{j\lambda \cdot z} \rho_{z,\lambda}[k] \rho_{z,\lambda}[l], \quad (23)$$

where

$$\begin{aligned} \rho_{z,\lambda}[k] &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \exp(j\frac{\pi}{\lambda z}(k\Delta + \xi' + \xi)^2) d\xi d\xi' = \\ &\Delta \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \exp(j\frac{\pi}{\lambda z}(k\Delta + \xi'\Delta + \xi\Delta)^2) d\xi d\xi' = \\ &2\Delta \int_{-1}^1 (1 - |v|) \exp(j\frac{\pi}{\lambda z}(k\Delta + \Delta v)^2) dv. \end{aligned} \quad (24)$$

The proof that the double integral in (24) can be calculated as a single variable integral is given in Appendix 2. For the Fresnel approximation the transfer function (18) is of the form

$$\tilde{A}_z(v_x, v_y) = \frac{\exp(j2\pi z/\lambda)}{j\lambda \cdot z} \Lambda_{z,\lambda}(v_x, N_a) \cdot \Lambda_{z,\lambda}(v_y, N_a), \quad (25)$$

where

$$\begin{aligned} \Lambda_{z,\lambda}(v, N) &= \sum_{k,l=-N/2}^{N/2-1} \rho_{z,\lambda}[k] W^{vk}, \\ W &= \exp(-j2\pi/N_a), \quad v = -N_a/2, \dots, N_a/2 - 1. \end{aligned} \quad (26)$$

The kernel  $a_z$  (13) and the transfer function  $\tilde{A}_z$  (18) are defined for the general wavefield propagation model. The Fresnel approximation allows significantly simplify calculations because the kernel  $a_z[k, l]$  in (23) is calculated

through the product of one-dimensional  $\rho_{z,\lambda}[k]$  and  $\rho_{z,\lambda}[l]$ . In a similar way the 2D transfer function  $\tilde{A}_z(v_x, v_y)$  in (25) is factorized in the product of  $\Lambda_{z,\lambda}(v_x, N_a)$  and  $\Lambda_{z,\lambda}(v_y, N_a)$ .

The integration in (24) can be produced numerically with any given accuracy for any wavelength  $\lambda$  and  $z$ . Note that the system is linear, the kernel  $a_z$  does not depends on the input signal, thus the kernel is calculated only ones for the given parameters and can be used for any input signals.

It is useful to note that for small  $\Delta/z$  and  $\Delta/\lambda$  we do not need to calculate the integrals (13) and (24) because  $\rho_{z,\lambda}[k] \rightarrow \Delta \exp(j \frac{\pi}{\lambda z} (k\Delta)^2)$ . In this case the frequency domain algorithm (17)-(20) is similar to the integration *FFT* algorithm proposed in [7] for the squared input/output arrays  $u_0, u_z$ , where these arrays are zero padded to the size  $(N_a - 1) \times (N_a - 1)$ ,  $N_a = 2N_0$ , in order to obtain the circular convolution and the transfer function  $\tilde{A}_z$  is calculated as *FFT* of the kernel  $g_z$  defined on the extended grid  $(N_a - 1) \times (N_a - 1)$ .

#### 4. Wavefield reconstruction

It is well known that a reconstruction of the object wavefield  $u_0$  from  $u_z$  is an inverse problem. Let us start from the discussion of specific features of this problem considered in the extended size variables what makes the problem

essentially different from the standard settings.

#### 4.A. Reconstruction as a boundary inverse problem

It is convenient to consider this inverse reconstruction in both the space and frequency domains in parallel. Using the representation similar to (15) the extended size inverse problem can be formulated in the algebraic form as

$$\tilde{\mathbf{u}}_z = \widetilde{\mathcal{A}}_z \cdot \tilde{\mathbf{u}}_0, \quad (27)$$

where  $\tilde{\mathbf{u}}_z, \tilde{\mathbf{u}}_0 \in R^{N_a^2}$  and  $\widetilde{\mathcal{A}}_z = \begin{pmatrix} \mathcal{A}_z^{11} & \mathcal{A}_z^{12} \\ \mathcal{A}_z^{21} & \mathcal{A}_z^{22} \end{pmatrix}$  is a structured  $N_a^2 \times N_a^2$  matrix.

All submatrices in  $\widetilde{\mathcal{A}}_z$  are formed from  $\tilde{a}_z[k, l]$ .

The conditions when the system (27) has a unique solution for  $\tilde{\mathbf{u}}_0$  are formulated in the following proposition.

**Proposition 2.** *Let  $N_z \geq N_0$  and*

$$\Lambda_{z,\lambda}(f, N_a) \neq 0, \quad \forall f = -N_a/2, \dots, N_a/2 - 1, \quad (28)$$

*then the system (27) has a unique solution.*

The proof of the proposition is given in Appendix 3.

The conditions (28) are necessary and sufficient for non-singularity of the extended size *DDT* operator. It is difficult to proof for the discrete function  $\Lambda_{z,\lambda}(f, N_a)$  that it never takes zero value. However, these conditions can be

easily checked numerically for the finite set of the frequencies  $f$  in (28). In our experiments for variety of  $\lambda$ ,  $\Delta$  and  $N_a$  we never met a situation when  $\Lambda_{z,\lambda}(f, N_a)$  takes a zero value for  $f$  from the interval  $[-N_a/2, \dots, N_a/2 - 1]$ . Thus, practically the system (27) always has a unique solution. This statement says that for any given  $\tilde{\mathbf{u}}_z$  the corresponding unique  $\tilde{\mathbf{u}}_0$  always can be found. The reconstruction of  $\mathbf{u}_0$  from (27) requires a further detailing.

Let the vectors  $\tilde{\mathbf{u}}_z, \tilde{\mathbf{u}}_0$  in (27) be structured as follows  $\tilde{\mathbf{u}}_z = \begin{pmatrix} \mathbf{u}_z \\ \mathbf{u}'_z \end{pmatrix}$  and  $\tilde{\mathbf{u}}_0 = \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}'_0 \end{pmatrix}$ , where  $\mathbf{u}_z$  and  $\mathbf{u}_0$  are the main variables from (15), and  $\mathbf{u}'_z \in R^{N_a^2 - N_z^2}$  and  $\mathbf{u}'_0 \in R^{N_a^2 - N_0^2}$  are the vectors of the idle variables complementing the vectors  $\tilde{\mathbf{u}}_z, \tilde{\mathbf{u}}_0$  to the extended size. It is obvious that for these structured vectors  $\mathcal{A}_z^{11} = \mathcal{A}$  from (15). It follows from (20) that  $\mathbf{u}'_0 = 0$ .

In the inverse problem of our interest the vector  $\mathbf{u}_z = \mathbf{u}_z^\#$  is given, and the vectors  $\mathbf{u}_0$  and  $\mathbf{u}'_z$  are unknown. Thus we arrive to the equation

$$\begin{pmatrix} \mathbf{u}_z^\# \\ \mathbf{u}'_z \end{pmatrix} = \widetilde{\mathcal{A}}_z \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{0} \end{pmatrix}, \quad (29)$$

with a total number of the equations  $N_a^2$ , and  $N_0^2$  unknowns in  $\mathbf{u}_0$  and  $(N_a^2 - N_0^2)$  unknowns in  $\mathbf{u}'_z$ . Thus, there is the accurate balance of the number of

the unknowns and the equations at hand.

The equation (29) belongs to the class of the *boundary problems* as the unknown  $\mathbf{u}'_z$  is in the left side of the equation and the unknown  $\mathbf{u}_0$  in the right side of the equation. (29) is resolved with respect to  $\mathbf{u}'_z$ . We can invert  $\widetilde{\mathcal{A}}_z$  in order to resolve this equation with respect to  $\mathbf{u}_0$

$$\begin{pmatrix} \mathbf{u}_0 \\ \mathbf{0} \end{pmatrix} = \widetilde{\mathcal{A}}_z^{-1} \begin{pmatrix} \mathbf{u}_z^\# \\ \mathbf{u}'_z \end{pmatrix}. \quad (30)$$

The boundary inverse problem can be solved iteratively using (29) and (30) and starting from the initialization of  $\mathbf{u}_z'^{(0)} = 0$  :

$$\begin{pmatrix} \hat{\mathbf{u}}_0^{(r)} \\ \hat{\mathbf{u}}_0'^{(r)} \end{pmatrix} = \widetilde{\mathcal{A}}_z^{-1} \begin{pmatrix} \mathbf{u}_z^\# \\ \hat{\mathbf{u}}_z'^{(r-1)} \end{pmatrix}, \quad (31)$$

$$\begin{pmatrix} \hat{\mathbf{u}}_z^{(r)} \\ \hat{\mathbf{u}}_z'^{(r)} \end{pmatrix} = \widetilde{\mathcal{A}}_z \begin{pmatrix} \hat{\mathbf{u}}_0^{(r)} \\ \mathbf{0} \end{pmatrix}, \quad r = 1, \dots \quad (32)$$

Here  $\mathbf{u}_0'^{(r)}$  and  $\hat{\mathbf{u}}_z'^{(r)}$  are the idle variables, while  $\hat{\mathbf{u}}_0^{(r)}$  is a variable of the interest. In what follows we use the *FFT* domain version of the equation (31)-(32) in order to obtain the fast and efficient implementation of the recursive procedure (31)-(32), where the inverse  $\widetilde{\mathcal{A}}_z^{-1}$  is replaced by the regularized

inverse because  $\widetilde{\mathcal{A}}_z$  is ill-conditioned.

#### 4.B. Regularized inverse

Let us start from a simple algorithm for the inverse. Assume that  $\tilde{U}_z(v_x, v_y)$  in (19) is known as given by the observations and introduce the following quadratic criterion defining the accuracy fit of the model  $\tilde{A}_z \tilde{U}_0$  to the given  $\tilde{U}_z$ :

$$J = \|\tilde{U}_z - \tilde{A}_z \tilde{U}_0\|^2 + \alpha \|\tilde{U}_0\|^2. \quad (33)$$

Here  $\|\cdot\|^2$  is the Euclidean norm calculated over the *FFT* frequencies  $v_x, v_y$ ,  $\|\tilde{U}_0\|^2$  and  $\alpha > 0$  are a regularization penalty term and a regularization parameter, respectively. This regularization is of special importance when the *DDT* is ill-conditioned and/or the observations  $\tilde{U}_z$  are noisy [12], [13], [14].

Minimization of  $J$  with respect to  $\tilde{U}_0$  gives the minimum condition in the form  $\partial J / \partial \hat{U}_0^* = 0$  and the following solution (e.g. [14])

$$\hat{U}_0 = \tilde{A}_z^* \cdot \tilde{U}_z / (|\tilde{A}_z|^2 + \alpha^2). \quad (34)$$



Finally we obtain

$$\tilde{u}_0 = \mathcal{F}\mathcal{F}\mathcal{T}^{-1}\{\hat{U}_0\}, \quad (35)$$

$$\hat{u}_0[k, l] = \tilde{u}_0[k, l], \quad k, l = -N_0/2, \dots, N_0/2 - 1.$$

Here  $\hat{u}_0$  is a *regularized inverse* estimate of the object wavefield distribution. The parameter  $\alpha > 0$  controls a level of smoothing in this estimate. Smaller  $\alpha$  means a lower smoothing in  $\hat{u}_0$  with possibly high level of noisy and calculation error components but a lower level of the bias systematic errors. Larger  $\alpha$  means stronger smoothing in  $\hat{u}_0$  with lower level of noisy and calculation error components but possibly higher level of the bias systematic errors [12], [13], [14].

The frequency domain calculations in (34) are produced for the extended size  $N_a \times N_a$  of  $u_0$  and  $u_z$ . The zero-padding of the observations  $u_z$  usually results in some ringing effects along the bounds of the reconstructed object distribution. An improved version of (34)-(35) what we call an *recursive regularized inverse* allows to rid of these effects or at least diminish them significantly.

#### 4.C. Recursive regularized inverse

We propose the following recursive procedure defining the frequency domain and regularized version of the procedure (31)-(32). This algorithm is composed from backward (inverse) and forward transformations performed recursively for  $r = 1, 2, \dots$  :

(a) for the backward propagation equation (31) we produce the following successive calculations

$$\tilde{u}_z^{(r-1)}[k, l] = u_z^\# [k, l], \quad k, l = -N_z/2, \dots, N_z/2 - 1, \quad (36)$$

$$\tilde{U}_z^{(r)}[v_x, v_y] = \mathcal{F}\mathcal{F}\mathcal{T}\{\tilde{u}_z^{(r-1)}[k, l]\},$$

$$\tilde{U}_0^{(r)}[v_x, v_y] = \tilde{A}_z^*/(|\tilde{A}_z|^2 + \alpha^2) \cdot \tilde{U}_z^{(r)}[v_x, v_y]$$

$$\tilde{u}_0^{(r)}[s, t] = \mathcal{F}\mathcal{F}\mathcal{T}^{-1}\{\tilde{U}_0^{(r)}[v_x, v_y]\}, \quad (37)$$

$$\tilde{u}_0^{(r)}[s, t] = 0, \text{ for all } s, t \notin [-N_0/2, \dots, N_0/2 - 1]; \quad (38)$$

(b) for the forward propagation equation (32) we produce the following successive calculations

$$\tilde{U}_0^{(r)}[v_x, v_y] = \mathcal{F}\mathcal{F}\mathcal{T}\{\tilde{u}_0^{(r)}[s, t]\}, \quad (39)$$

$$\tilde{U}_z^{(r)}[v_x, v_y] = \tilde{A}_z(v_x, v_y)\tilde{U}_0^{(r)}[v_x, v_y],$$

$$\tilde{u}_z^{(r-1)}[k, l] = \mathcal{F}\mathcal{F}\mathcal{T}^{-1}\{\tilde{U}_z^{(r-1)}[v_x, v_y]\};$$

(c) finally the estimate of the object distribution  $u_0$  follows from (38):

$$\hat{u}_0^{(r)}[s, t] = \tilde{u}_0^{(r)}[s, t], \quad s, t = -N_0/2, \dots, N_0/2 - 1. \quad (40)$$

The procedure (36)-(39) is initiated by  $\tilde{u}_z^{(0)} = 0$ .

The estimates  $\hat{u}_0^{(r)}$  (40) are obtained from the backward equations, while the forward equations are used for predictions  $\tilde{u}_z^{(r-1)}$  of the observed data from the obtain estimate. The multiple experiments show that the algorithm (36)-(40) is quite efficient and provides a significant improvement of the regularized inverse estimate (34)-(35), which is the first step of this recursive algorithm.

The algorithm (36)-(40) allows easily incorporate a priori information on the object distribution. If the object is an amplitude modulator with the phase of the object distribution equal to zero, then the equation (37) is replaced by

$$\tilde{u}_0^{(r)}[s, t] = |\mathcal{F}\mathcal{F}\mathcal{T}^{-1}\{\tilde{U}_0^{(r)}[v_x, v_y]\}|. \quad (41)$$

If the object is a phase modulator of the amplitude equal to 1 then the equation (37) is replaced by

$$\tilde{u}_0^{(r)}[s, t] = \exp(-j \cdot \text{angle}(\mathcal{F}\mathcal{F}\mathcal{T}^{-1}\{\tilde{U}_0^{(r)}[v_x, v_y]\})). \quad (42)$$

This specification of the estimated object distribution improves the convergence and the accuracy of the algorithm.

## 5. Simulation experiments for wavefield reconstruction

Let us illustrate the performance of the proposed inverse algorithm for different applications. In these numerical experiments we use the *Baboon* intensity distribution ( $512 \times 512$ ) imaged in Fig 5. This distribution denoted in what follows as  $u^\#$  is real valued non-negative,  $0 \leq u^\# \leq 1$ . The parameters of the experiment:  $\lambda = .632 \text{ nm}$ ,  $z = 0.5 \text{ m}$ ,  $N_0 = N_z = N = 512$ , the square sensor and object have the same size  $0.01 \times 0.01 \text{ m}$  with the pixel size parameter  $\Delta = 0.01/N \text{ m}$ .

### 5.A. Sampling conditions

The frequency of the varying on  $x$  harmonic factor  $\exp[j\frac{\pi}{\lambda z}(x^2 + y^2)]$  of the kernel (2) is calculated as  $2\pi x/\lambda z$ . The maximum values of  $x$  is equal to half of the sensor linear size,  $x_{\max} = 0.01/2 = 0.005$ . Then the upper bound for the

frequency is equal to  $2\pi x_{\max}/\lambda z$ . The Nyquist requirement for the non-aliasing sampling follows from the inequality  $2\pi x_{\max}/\lambda z < \pi/\Delta$ , i.e.  $\Delta < \Delta_{crit} = \lambda z/2x_{\max}$ . For the given sensor size,  $\Delta_{crit} = 3.16 \cdot 10^{-5}$ . It follows that for the selected parameters of the experiments  $\Delta = 0.01/512 = 1.95 \cdot 10^{-5} < \Delta_{crit}$ , i.e. the Nyquist condition holds.

In order to demonstrate that *DDT* gives the aliasing free reconstructions we start from the case when the Nyquist condition is violated. Let  $\Delta$  have a double size,  $\Delta = 2 \cdot 1.95 \cdot 10^{-5} > \Delta_{crit}$ . Assume that the true object distribution is real-valued  $u_0 = u^\#$ .

Fig. 6 shows the amplitude and phase wavefield distributions at the sensor plane calculated according to the forward propagation model (17)-(20). It is emphasized once more that this modeling gives the precise prediction for the sensor plane distribution corresponding to the accurately calculated integral (1).

The object plane reconstructions are shown in Fig.7. Fig.7a and Fig.7b show the amplitude distributions in the object plane obtained by the *DDT* regularized inverse (34)-(35) and by the standard *FFT* (8)-(9) algorithm, respectively. The periodical pattern typical for the aliasing effects destroys the

standard reconstruction shown in Fig.7b completely, while the *DDT* based algorithm demonstrates in Fig.7a the quite accurate reconstruction of the *Baboon* test-distribution. This result confirms the ability of the *DDT* algorithm to perform and perform quite well even when the Nyquist condition is violated.

In all forthcoming experiments we assume that  $\Delta = 1.95 \cdot 10^{-5} < \Delta_{crit}$ , i.e. the Nyquist condition holds. Then the standard *FFT* works but qualitatively and numerically the *DDT* technique enables better results. For quantitative numerical characterization of the accuracy of the reconstruction we use the root-mean-squared-error (*RMSE*)

#### 5.B. Amplitude modulation

Let the true wavefield distribution in the object plane be real-valued  $u_0 = u^\#$ . The results obtained by the standard *FFT* and *DDT* algorithms can be seen in Fig.8. Here for the *DDT* reconstruction we use the recursive regularized *DDT* algorithm (36)-(40). The results are given for the ten iteration of this algorithm. Comparison of Fig.8a versus Fig.8b is in the favor of the recursive regularized *DDT* algorithm. It demonstrates a better visual quality as well as a lower *RMSE* value.

The regularized inverse estimate in Fig.7a is a first iteration of the recursive regularized inverse algorithm. The ringing bands parallel to the boundaries in Fig.7a are typical for the regularized inverse deconvolution algorithms. Comparing this image with given in Fig.8a we can note that the recursive regularized inverse allows to get rid of these disturbing effects and yields about twice better *RMSE* value.

#### 5.C. Phase modulation

Consider a reconstruction of the object phase distribution assuming that the true object distribution has a module equal to one and the phase proportional to  $u^\#$ ,  $u_0 = \exp(-j\pi u^\#)$ . For this reconstruction we use the ten iterations of the algorithm (36)-(40). The results obtained by the standard *FFT* and *DDT* algorithms are shown in Fig.9. While numerically ( $RMSE = .186$  for *DDT* algorithm and  $RMSE = .26$  for the *FFT* algorithm) the difference between the estimates is not large the visual advantage of the *DDT* algorithm is quite obvious as it gives sharper and clearer imaging with no ringing effects along boundaries.

## 6. Object wavefield design

Assume that the object distribution  $u_0$  should be selected in such a way that the wavefield distribution at a given distance  $z = z^\#$  would be equal or at least close to a desired wavefield distribution  $u_{z^\#}^\#$ . It is not difficult to realize that formally this design problem is identical to the considered above reconstruction one. Indeed, let us replace the observed distribution  $u_z^\#$  by the corresponding desired one then the regularized inverse and the recursive regularized inverse algorithms give the unknown object distribution  $u_0$ .

Let us exploit this approach for design of the object phase distribution such that the amplitude distribution at the sensor plane would have a desirable amplitude distribution. It is one of the typical setting for programming of the phase SLM used for generation of a given light intensity distribution.

For this problem we use the algorithm (36)-(40) where the equation (37) is replaced by (42). In this way the algorithm varies only the phase of the object distribution trying to get the amplitude at the sensor plane close to  $u_{z^\#}^\#$  desired at the distance  $z^\#$ .

Fig.10 illustrates the obtained results. Fig.10a shows the obtained phase distribution for the object plane, while Fig.10b demonstrates the quality of



this design as it shows the predicted amplitude distribution at the sensor plane corresponding the phase object with the phase distribution shown in Fig.10a. The curve in Fig.11 demonstrates the convergence of the algorithm for this test. The results in Fig.10 are achieved after 50 iterations of the algorithm. Numerically as well as visually (Fig.10b versus Fig 5) the results are quite good.

## 7. Conclusion

This paper concerns two aspects of discrete holography: discrete modelling of the forward wavefield propagation and inverse problems. The important property of the proposed *DDT* is that it gives the accurate prediction for the forward propagation for a pixelwise constant object distribution. This prediction is free of the aliasing effects typical for all standard discrete models.

The averaging of the observed intensities over the sensor pixel is not a new idea and discussed in a number of publications (e.g. [1], [2]). However, to the best of our knowledge nobody used it for averaging both in the sensor and object domains, and, what is more important, to derive the operators corresponding to these averaged variables. The exact forward modelling allows to give a reliable tool for evaluation of the alternative algorithms by computer

simulation.

The developed inverse technique is applied for two problems: distribution reconstruction in the object plane from observations given in the sensor plane and the design of the object plane distribution in order to obtain the desirable distribution in the sensor plane. Simulation experiments show a better accuracy and better imaging of the developed reconstruction techniques versus the standard *FFT* alternative.

The developed inverse algorithms similar to the standard *FFT* algorithm are not able to get a good quality reconstruction with sharp focussing for all values of the parameters  $\Delta$  and  $z$ . However, numerically and visually the *DDT* based algorithm demonstrate a better performance.

One of the constructive ideas to improve the proposed frequency domain inverse algorithms is to use a sensor of the size larger than the size of the object aperture. First promising results in this direction are reported in [15].

Finally we wish to note that *DDT* as a discrete model for wavefield propagation can be generalized to the pixel size different in the object and sensor planes and applied to various optical settings used in digital holography.

## 8. Acknowledgement

We thank the anonymous referee for very fruitful and neat comments. We would like also to express our thanks to Claass Falldorf from BIAS (Bremen, Germany) for comments and discussion. This work was supported by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence program 2006 – 2011).

## References

1. Th.Kreis, *Handbook of Holographic Interferometry (Optical and Digital Methods)* (Wiley-VCH GmbH&Co.KGaA, Weinheim, 2005).
2. I. Yamaguchi, J. Kato, S. Ohta, and J. Mizuno, "Image formation in phase-shifting digital holography and applications to microscopy," *Appl. Opt.*, vol. 40, no. 34, 6177–6186 (2001).
3. V. Katkovnik, J. Astola, K. Egiazarian, "Wavefield reconstruction and design as discrete inverse problems," *Proceedings of 3D TV Conference*, Istanbul (2008).
4. G. S. Sherman, "Integral-transform formulation of diffraction theory," *J. Opt. Soc. Am.*, vol. 57, no. 12, 1490-1498 (1967).
5. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, Inc, New York, Second Edition, 1996).
6. L. Onural, "Exact analysis of the effects of sampling of the scalar diffraction field," *J. Opt. Soc. Am. A*, vol. 24, no. 2, 359-367 (2007).

7. F. Shen and A. Wang, "Fast-Fourier-transform based numerical integration method for the Rayleigh–Sommerfeld diffraction formula," *Appl. Opt.*, vol. 45, no. 6, 1102-1110 (2006).
8. J. M. Arnold, "Phase-space localization and discrete representations of wave fields," *J. Opt. Soc. Am. A*, vol. 12, no. 1, 111-123 (1995).
9. L. Onural, "Some mathematical properties of the uniformly sampled quadratic phase function and associated issues in digital Fresnel diffraction simulations," *Opt. Eng.* vol. 43, no. 11, pp. 2557–2563 (2004).
10. I. Aizenberg and J. Astola, "Discrete generalized Fresnel functions and transforms in an arbitrary discrete basis," *IEEE Trans on Signal Processing*, vol. 54, no.11, pp. 4261-4270 (2006).
11. M. Liebling, Th. Blu, and M. Unser, "Fresnelets: new multiresolution wavelet bases for digital holography," *IEEE Trans on Image Processing*, vol. 12, no. 1, pp. 29-43 (2003).
12. Tikhonov A.N. and V.Y. Arsenin. *Solution of ill-posed problems* (Wiley, New York, 1977).
13. M. Bertero and P. Boccacci. *Introduction to inverse problems in imaging* . (IOP Publishing Ltd, 1998).
14. V. Katkovnik, K. Egiazarian, J. Astola, *Local Approximation Techniques in Signal and Image Processing* (SPIE PRESS, Bellingham, Washington, 2006).
15. V. Katkovnik, J. Astola, K. Egiazarian, "Numerical wavefield reconstruction in phase-shifting holography as inverse discrete problem," submitted to *EUSIPCO 2008* (2008).

## Appendix 1. Proof of Proposition 1.

Calculate  $u_z[k, l]$  from the formulas (19)-(20) using the inverse *FFT*:

$$u_z[k, l] = \frac{1}{N_a^2} \sum_{v_x, v_y=-N_a/2}^{N_a/2-1} W^{-v_x k} W^{-v_y l} \tilde{A}_z(v_x, v_y) \tilde{U}_0(v_x, v_y), \quad (43)$$

$$k, l = -N_z/2, \dots, N_z/2 - 1.$$

Inserting the expressions for  $\tilde{A}_z(v_x, v_y)$  and  $\tilde{U}_0(v_x, v_y)$  and changing the orders of summations we find

$$u_z[k, l] = \frac{1}{N_a^2} \sum_{u, v=-N_a/2+1}^{N_a/2-1} \sum_{u', v'=-N_0/2}^{N_0/2-1} a_z[u, v] u_0[u', v'] \times \quad (44)$$

$$\sum_{v_y, v_x=-N_a/2}^{N_a/2-1} W^{-v_x(k-u-u')} W^{-v_y(l-v-v')}.$$

Noting that because of

$$\frac{1}{N_a^2} \sum_{v_y, v_x=-N_a/2}^{N_a/2-1} W^{-v_x(k-u-u')} W^{-v_y(l-v-v')} = \delta(k-u-u') \delta(l-v-v')$$

we obtain from (44)

$$u_z[k, l] = \sum_{u, v=-N_a/2+1}^{N_a/2-1} \sum_{u', v'=-N_0/2}^{N_0/2-1} a_z[u, v] u_0[u', v'] \delta(k-u-u') \delta(l-v-v'). \quad (45)$$

This last formula gives the convolution (12) provided that for each  $k, l = -N_z/2, \dots, N_z/2 - 1$  there are  $u, u'$  and  $v, v'$  such that  $u = k - u'$  and  $v = l - v'$  provided that  $u, v = -N_a/2 + 1, \dots, N_a/2 - 1$  and  $u', v' = -N_0/2, \dots, N_0/2 - 1$ .

Let us test these requirements. We can see that

$$-N_z/2 - (N_0/2 - 1) \leq k - u' \leq N_z/2 - 1 + N_0/2,$$

i.e.

$$-N_a/2 + 1 \leq k - u' \leq N_a/2 - 1.$$

It shows that for any  $k = -N_z/2, \dots, N_z/2 - 1$  and any  $u' = -N_0/2, \dots, N_0/2 - 1$  always exist  $u = k - u'$  such that  $u = -N_a/2 + 1, \dots, N_a/2 - 1$ .

In a similar way one can check that

$$-N_a/2 + 1 \leq l - v' \leq N_a/2 - 1$$

and always exist  $v = l - v'$  such that  $v = -N_a/2 + 1, \dots, N_a/2 - 1$ .

Then the formula (45) can be rewritten in the form (12)

$$u_z[k, l] = \sum_{u', v' = -N_0/2}^{N_0/2-1} a_z[k - u', l - v'] u_0[u', v'],$$

$$k, l = -N_z/2, \dots, N_z/2 - 1.$$

It proves that the formula (20) gives the result identical to (12). Note that

$N_a = N_0 + N_z$  is a minimal array size when these formulas are equivalent.

Appendix 2. Proof that

$$\frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} f(k\Delta + \xi' + \xi) d\xi d\xi' = 2\Delta \int_{-1}^1 f(k\Delta + \Delta x)(1 - |x|) dx. \quad (46)$$

Consider the integral

$$\begin{aligned} J &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} f(k\Delta + \xi' + \xi) d\xi d\xi' = \\ &\Delta \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} f(k\Delta + \Delta u + \Delta u') du du', \end{aligned}$$

where  $f$  is a function of two variables and  $u = \xi/\Delta$ ,  $u' = \xi'/\Delta$ . Further,

change the variables of integration according to the following formula

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} u \\ u' \end{pmatrix}. \quad (47)$$

With  $\varphi = \pi/4$  this orthogonal transformation means a rotation of the integration area on the angle  $\pi/4$  as shown in Fig.12.

For  $\varphi = \pi/4$  the new variables are defined as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u \\ u' \end{pmatrix} \quad (48)$$

and according to Fig 12 the following sequence of transformations can be done

$$\begin{aligned}
J &= \Delta \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} f(k\Delta + \Delta u + \Delta u') du du' = \\
&\Delta \int_0^{1/\sqrt{2}} f(k\Delta + \Delta\sqrt{2}x) \int_{-1/\sqrt{2}+x}^{1/\sqrt{2}-x} dy dx + \\
&\Delta \int_{-1/\sqrt{2}}^0 f(k\Delta + \Delta\sqrt{2}x) \int_{-1/\sqrt{2}-x}^{1/\sqrt{2}+x} dy dx = \\
&\Delta \int_0^{1/\sqrt{2}} f(k\Delta + \Delta\sqrt{2}x) 2(1/\sqrt{2} - x) dx + \\
&\Delta \int_{-1/\sqrt{2}}^0 f(k\Delta + \Delta\sqrt{2}x) 2(1/\sqrt{2} + x) dx = \\
&2\Delta \int_{-1/\sqrt{2}}^{1/\sqrt{2}} f(k\Delta + \Delta\sqrt{2}x) 2(1/\sqrt{2} - |x|) dx = \\
&2\Delta \int_{-1}^1 f(k\Delta + \Delta x') (1 - |x'|) dx'.
\end{aligned}$$

The last expression proves (46).

### Appendix 3. Proof of Proposition 3.

If the conditions (28) hold there is the one-to-one link between the extended object and sensor wavefield distributions. In the frequency domain this link



means that

$$\tilde{U}_z(v_x, v_y) = \tilde{A}_z(v_x, v_y) \tilde{U}_0(v_x, v_y), \quad (49)$$

$$\tilde{U}_0(v_x, v_y) = \tilde{A}_z^{-1}(v_x, v_y) \tilde{U}_z(v_x, v_y),$$

$$k, l = -N_a/2, \dots, N_a/2 - 1.$$

For the spatial domain representation in the form (29) it means that the matrix  $\widetilde{\mathcal{A}}_z$  is non-singular.

## List of Figure Captions

Fig. 1. Principal setup of wavefield propagation and reconstruction.

Fig.2 The conditioning number versus the distance  $z$  for the transfer function  $\tilde{A}_z$  with  $N = 512$ ,  $N = 1024$  and  $\lambda = .632 \text{ nm}$ .

Fig. 3. Module frequency characteristics of the discrete diffraction transfer function: (a)  $\Delta \rightarrow 0$  and (b)  $\Delta = .01/1024$ ,  $z = 0.75 \text{ m}$ .

Fig. 4. Phase frequency characteristics of the discrete diffraction transfer function: (a)  $\Delta \rightarrow 0$  and (b)  $\Delta = .01/1024$ ,  $z = 0.75 \text{ m}$ .

Fig.5 Image of the *Baboon* test-distribution.

Fig. 6. Wavefield distribution at the sensor plane: (a) amplitude distribution, (b) phase distribution.

Fig. 7. Object wavefield reconstruction (amplitude distribution), double pixel size  $\Delta = 3.9 \cdot 10^{-5}$ : (a) inverse regularized *DDT* reconstruction,  $RMSE = .108$ , (b) standard *FFT* algorithm. The standard algorithm fails with a pattern of clear aliasing effects. *DDT* shows a good quality aliasing free reconstruction.

Fig. 8. Wavefield reconstruction (amplitude distribution) : (a) the recursive regularized inverse *DDT* technique,  $RMSE = .051$  after 10 iterations, (b)

the standard *FFT* technique,  $RMSE = 0.086$ . The *DDT* algorithm shows sharper and better resolution imaging.

Fig. 9 Phase distribution reconstruction: (a) the recursive regularized inverse *DDT* technique,  $RMSE = .185$  after 10 iterations, (b) the standard *FFT* technique,  $RMSE = 0.26$ . The *DDT* algorithm shows sharper and better imaging with no boundary ringing effect seen in the *FFT* reconstruction.

Fig. 10 Phase distribution design for the object plane in order to obtain a desirable amplitude distribution in the sensor plane: (a) designed phase distribution for the object plane, (b) the corresponding amplitude distribution in the sensor plane.

Fig. 11 Convergence of the recursive regularized inverse phase design algorithm:  $RMSE$  versus the number of iterations.

Fig. 12 The integration area in the original and transformed (rotated) coordinates.

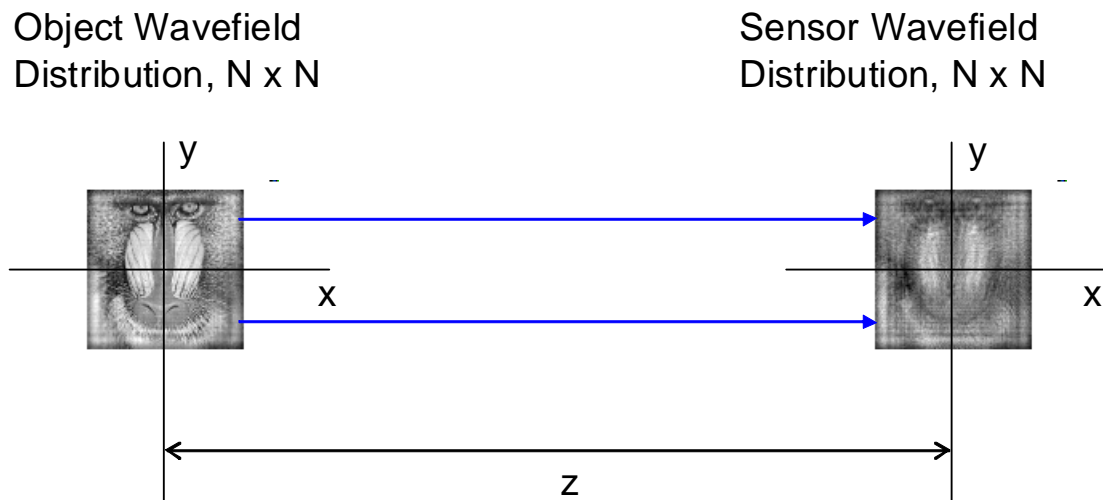


Fig. 1. Principal setup of wavefield propagation and reconstruction.

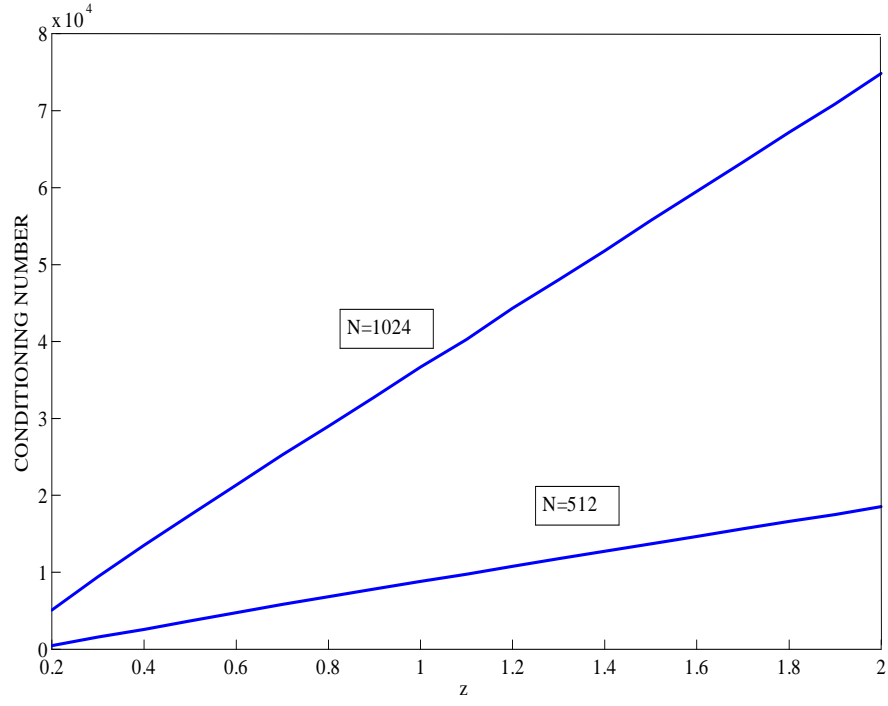
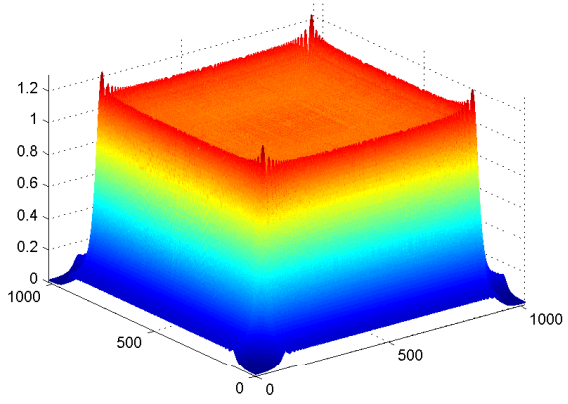
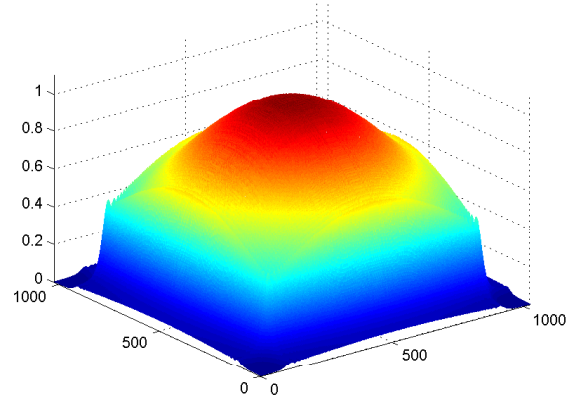


Fig. 2. The conditioning number versus the distance  $z$  for the transfer function  $\tilde{A}_z$  with  $N = 512$ ,  $N = 1024$  and  $\lambda = .632 \text{ nm}$ .

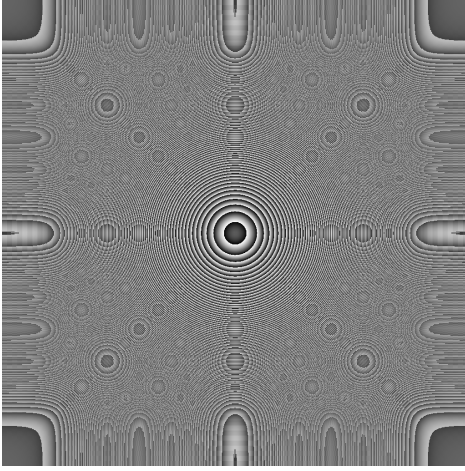


(a)

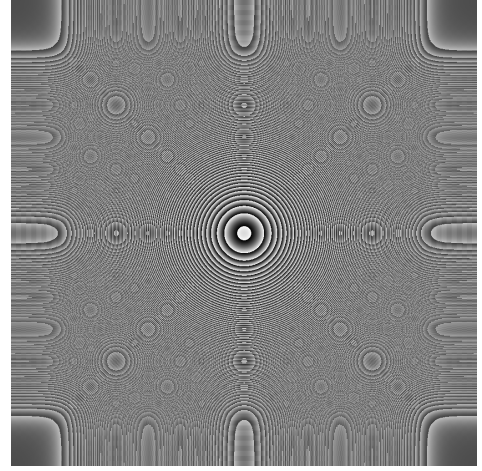


(b)

Fig. 3. (Color online) Module frequency characteristics of the discrete diffraction transfer function: (a)  $\Delta \rightarrow 0$  and (b)  $\Delta = 0.01/1024$ ,  $z = 0.75$  m.



(a)



(b)

Fig. 4. Phase frequency characteristics of the discrete diffraction transfer function: (a)  $\Delta \rightarrow 0$  and (b)  $\Delta = 0.01/1024$ ,  $z = 0.75$  m.

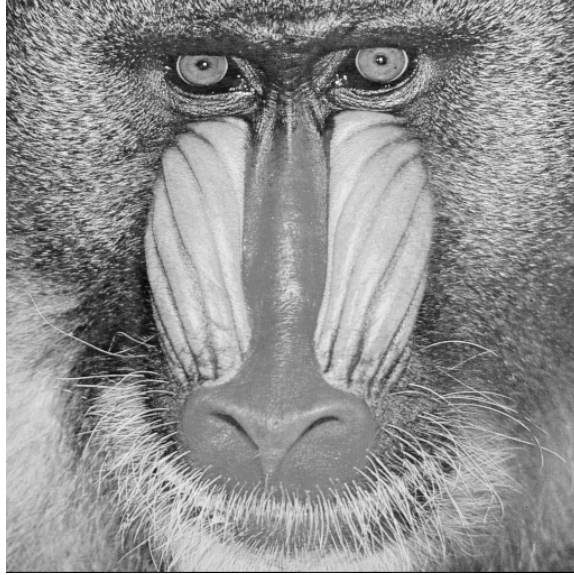
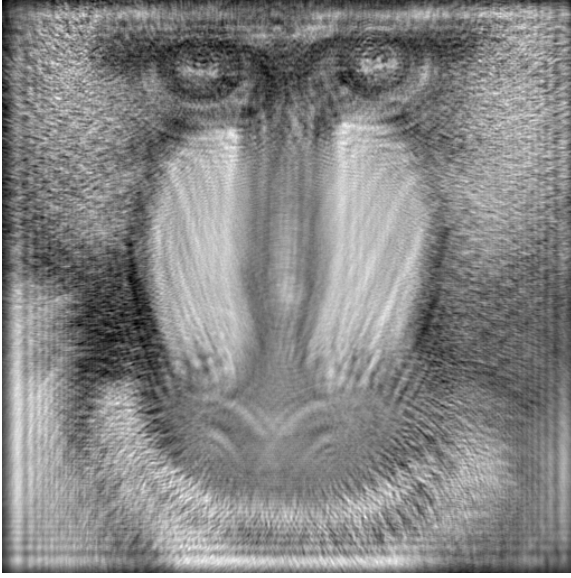
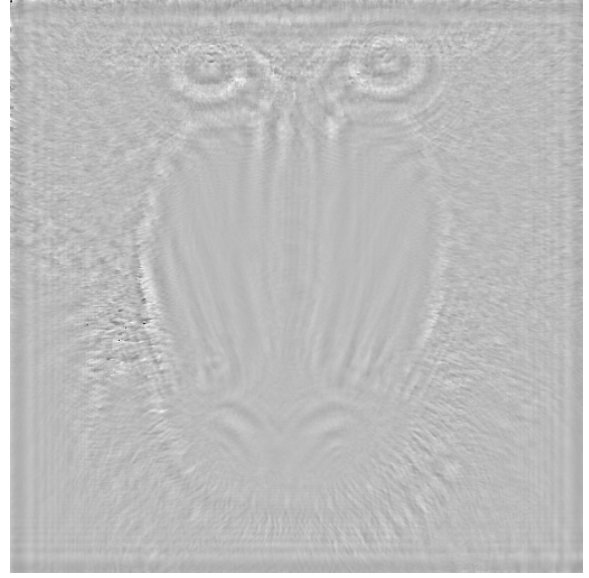


Fig. 5. Image of the *Baboon* test-distribution.



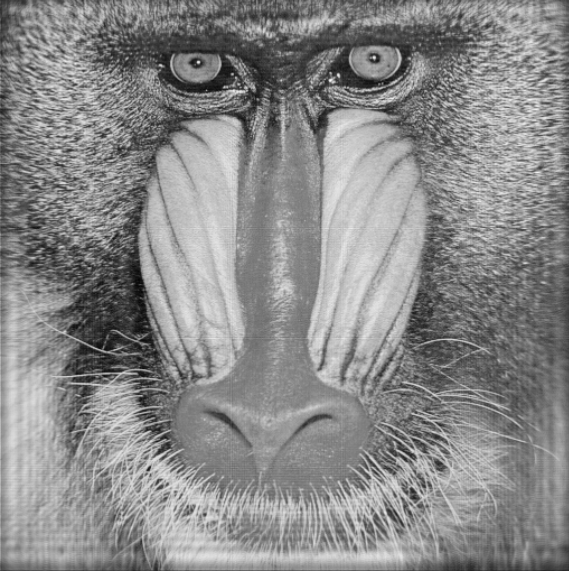


(a)

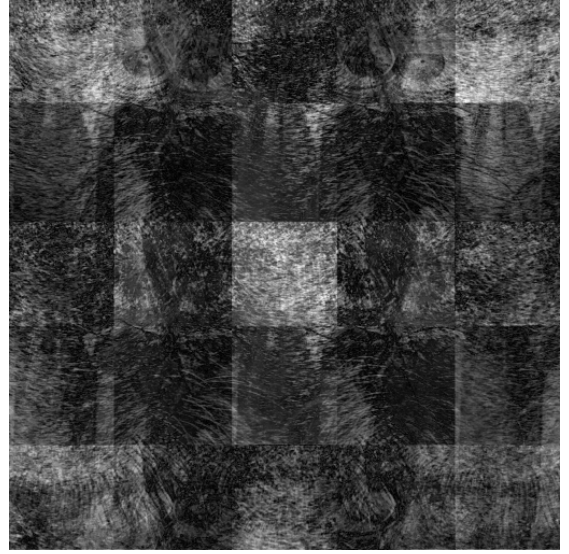


(b)

Fig. 6. Wavefield distribution at the sensor plane: (a) amplitude distribution, (b) phase distribution.

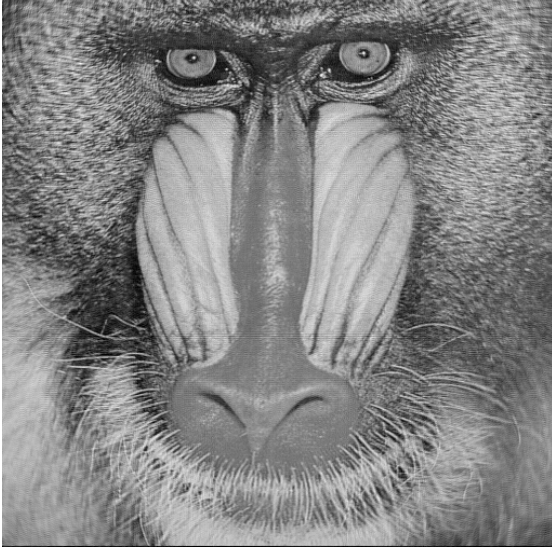


(a)

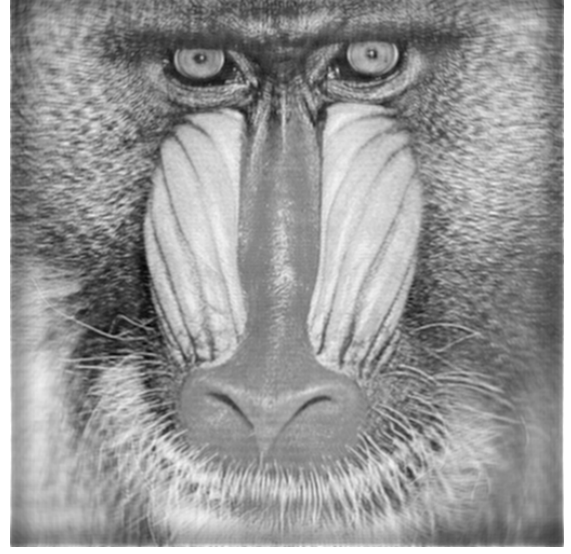


(b)

Fig. 7. Object wavefield reconstruction (amplitude distribution), double pixel size  $\Delta = 3.9 \cdot 10^{-5}$  m,  $z = 0.5$  m: (a) inverse regularized *DDT* reconstruction,  $RMSE = .108$ , (b) standard *FFT* algorithm. The standard algorithm fails with a pattern of clear aliasing effects. *DDT* shows a good quality aliasing free reconstruction.

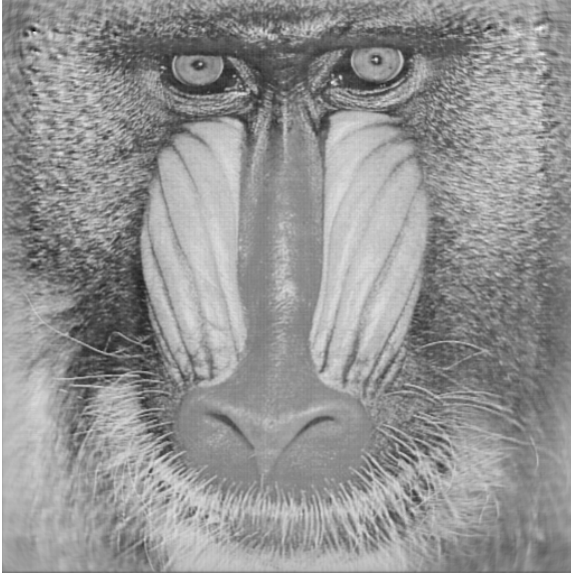


(a)

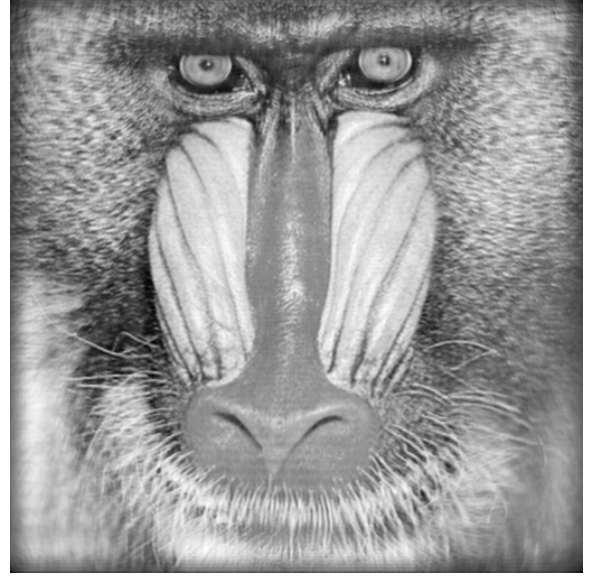


(b)

Fig. 8. Wavefield reconstruction (amplitude distribution) : (a) the recursive regularized inverse *DDT* technique,  $RMSE = .051$  after 10 iterations, (b) the standard *FFT* technique,  $RMSE = 0.086$ . The *DDT* algorithm shows sharper and better resolution imaging.

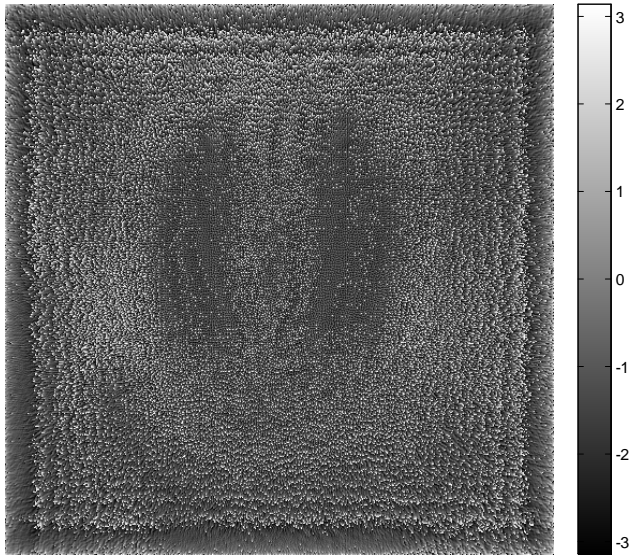


(a)

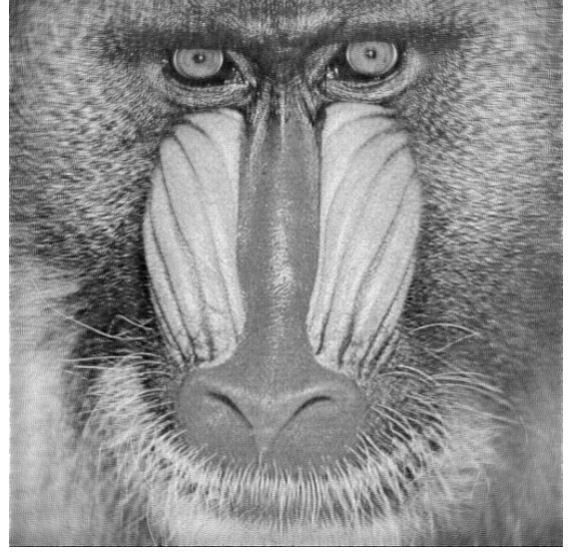


(b)

Fig. 9. Phase distribution reconstruction: (a) the recursive regularized inverse *DDT* technique,  $RMSE = .185$  after 10 iterations, (b) the standard *FFT* technique,  $RMSE = 0.26$ . The *DDT* algorithm shows sharper and better imaging with no boundary ringing effect seen in the *FFT* reconstruction.



(a)



(b)

Fig. 10. Phase distribution design for the object plane in order to obtain a desirable amplitude distribution in the sensor plane: (a) designed phase distribution for the object plane, (b) the corresponding amplitude distribution in the sensor plane,  $RMSE = 0.07$ .

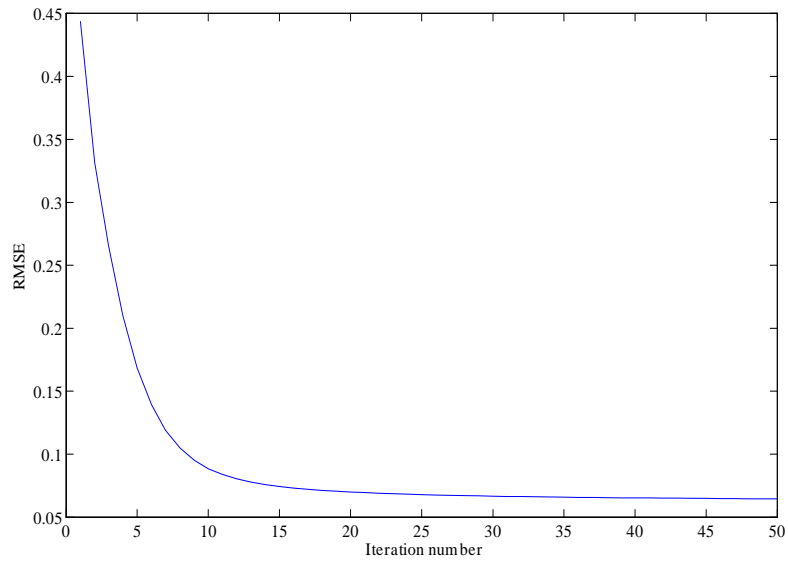


Fig. 11. Convergence of the recursive regularized inverse phase design algorithm:  $RMSE$  versus the number of iterations.

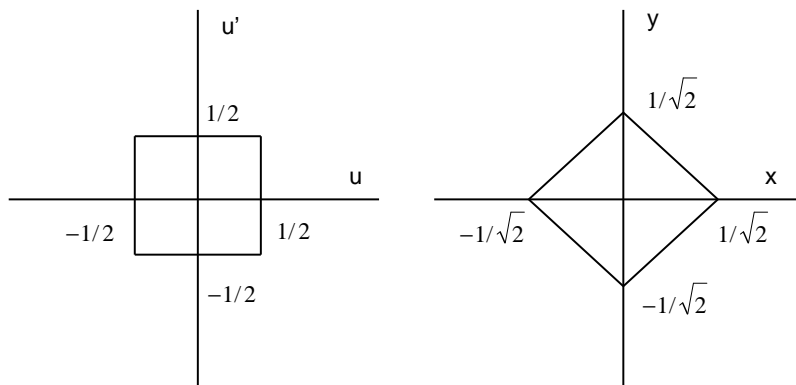


Fig. 12. The integration area in the original and transformed (rotated) coordinates.