TECHNICAL UNIVERSITY OF LIBEREC

Faculty of Mechatronics, Informatics and Interdisciplinary Studies

DIPLOMA THESIS

Liberec 2010 **Bc. Pavel Psota**

TECHNICAL UNIVERSITY OF LIBEREC

Faculty of Mechatronics, Informatics and Interdisciplinary Studies

Study program: N2612 – Electrical engineering and informatics

Specialization: Mechatronics

Application for digital holographic interferometry measurement

Aplikace pro měření digitální holografickou interferometrií

DIPLOMA THESIS

Author: Pavel Psota

Tutor for dissertation: Ing. Vít Lédl, Ph.D.

Dissertation Counsellor: Ing. Jan Václavík

zadání - originál

Declaration

Byl jsem seznámen s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu TUL.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Diplomovou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím diplomové práce a konzultantem.

Datum

Podpis

Acknowledgments

I want to thank my supervisor Ing. Vít Lédl, Ph.D. for leading and help with this thesis. Thanks also to Ing. Jan Václavík for valuable advice and to all people who supported me.

Abstract

For several decades, holographic interferometry performs an indispensable role among the measuring techniques. It is an universal technique that can be adapted to a wide range of measurements from deformation measurements through vibrations measurements to the temperature or pressure distribution measurements in gases. Technology is constantly developing as well the holographic measurement.

Firstly, holograms were recorded analog and then these holograms were evaluated by the experimenter. Afterwards, the recorded holograms were evaluated by computers and nowadays the hologram can be recorded electronically on the CCD or CMOS array and object wave is reconstructed using sophisticated algorithms. In digital holographic interferometry the wave fields are compared and the results are displayed.

This whole process is quite demanding and it requires a lot of different mathematical operations. Some of them proceed automatically and others require an operator intervention. The MATLAB GUI application was developed for this purpose. It contains all the necessary operations and also it is equipped by a number of graphics and other tools to facilitate the work with data. The application meets the requirements very well and greatly facilitates the accelerated implementation of the method of holographic interferometry.

Keywords: digital holography, holographic interferometry, contactless measurement, MATLAB GUI

Abstrakt

Holografická interferometrie již několik desítek let plní nenahraditelnou úlohu mezi měřicími technikami. Je to univerzální technika, kterou lze adaptovat na velmi širokou škálu různých měření, od deformací těles přes vibrace až po teplotní nebo tlaková rozložení v plynech. Technika se neustále vyvíjí a rozšiřují se i její možné aplikace.

Nejprve byly hologramy zaznamenány analogově a následně vyhodnoceny experimentátorem. Později zaznamenané hologramy byly vyhodnoceny počítači a dnes již hologramy mohou být zaznamenány elektronicky pomocí CCD nebo CMOS senzorů a objektová vlna je rekonstruována sofistikovanými algoritmy. V digitální holografické interferometrii jsou vlnová pole porovnána a výsledky zobrazeny.

Tento proces od začátku až dokonce je poměrně náročný a vyžaduje desítky různých matematických operací, některé probíhají zcela automaticky a jiné vyžadují zásah operátora. Pro tyto účely byla vytvořena aplikace v prostředí MATLAB GUI zastřešující všechny potřebné operace a navíc vybavená množstvím grafických prvků a nástrojů usnadňujících práci s daty. Vytvořená aplikace splňuje požadavky na ní kladené velmi dobře a výrazně usnadnila, zrychlila provádění celé metody holografické interferometrie.

Klíčová slova: digitální holografie, holografická interferometrie, bezkontaktní měření, MATLAB GUI

Contents

List of Figures

Table of Symbols

Symbol	Symbol unit	Name
c	[$m.s^{-1}$]	speed of light
$\boldsymbol{d}$	[m]	CCD – object distance
$\boldsymbol{d}$	[m]	displacement vector
$\boldsymbol{d_{2\pi}}$	[m]	phase cut
$\boldsymbol{e}$	[m^{-1}]	sensitivity vector
E	[$V.m^{-1}$]	electric field intensity
f	[m]	complex amplitude transmittance
g		wave field in the image plane
$g(\tau)$	[1]	complex degree of temporal coherence
$g(r_1,r_2)$	[1]	complex degree of spatial coherence
g_{12}	[1]	normalized cross-correlation function
$G(\tau)$	[1]	temporal coherence function
$G(r_1,r_2)$	[1]	mutual intensity
H	[$A.m^{-2}$]	magnetic field intensity
h	[1]	complex amplitude transmittance
I	[$W.m^{-2}$]	intensity of light
j	[1]	imaginary unit
k	[m^{-1}]	wave number
$\boldsymbol{k}$	[m^{-1}]	wave vector
$\boldsymbol{l_c}$	[m]	coherent length
M	[1]	number of pixels of CCD in horizontal direction
$\boldsymbol{n}$	[1]	refractive index

N	[1]	number of pixels of CCD in vertical direction
N_F	[1]	Fresnel number
r		reference wave
t	[s]	time
U		complex amplitudes
U_0		real amplitude of a wave
x	[1]	axis
y	[1]	axis
z	[1]	axis
δ	[m]	optical path difference
θ	[°]	angle with optical axis
λ	[m]	wavelength
τ_c	[s]	coherence time
υ	[m^{-1}]	spatial frequency
φ	[rad]	phase
ω	[rad. s^{-1}]	angular velocity
$\Delta\xi$	[m]	pixel size of CCD in vertical direction
$\Delta\eta$	[m]	pixel size of CCD in horizontal direction
Δ φ	[rad]	interference phase
$\mathcal{F}^{-1}$		inverse Fourier transform operator
*		convolution operator

Introduction

With its many applications holography is one of the most interesting discoveries in modern optics. Its scientific importance is emphasized by awarding Nobel Prize in 1971 to its inventor Denis Gabor. He created the word holography from the Greek words "holos" meaning complete and "graphien" meaning to write.

A hologram is the photographically or otherwise recorded interference pattern between a wave field scattered from the object and a coherent background called a reference wave. It contains the information about the entire three dimensional wave field. This information is coded in the form of interference stripes, usually with high spatial frequencies. The object wave can be reconstructed by illuminating the hologram with the reference wave. Such a reconstructed wave is de facto indistinguishable from the original object wave and there can be recognized a three dimensional image with all effects of perspective and depth focus.

Perhaps the most important application of holography is a so called Holographic Interferometry developed in the sixties of the last century. Here, two or more wave fields are compared interferometrically. Such technique allows the measurement of changes of the wave field phase and thus the change of any physical quantity that affects the phase. Holographic Interferometry permits to measure the deformation distribution, to determine the refractive index changes, the temperature field distribution and many others.

The development of the computing technology allowed transferring the recording process or the reconstruction process into the computer. Recording of the hologram numerically way led to a so called Computer Generated Holography, which generates holograms by numerical method and afterwards these holograms are reconstructed optically. Vice versa numerical reconstruction was initially based on sampling of holograms recorded on a photographic plate. Such digitized holograms were reconstructed numerically.

Inexorably evolving computing technology (increasing processing speed and memory capacity of computers, as well as CCD and CMOS cameras having higher resolution) led to direct recording of holograms with the CCD or the CMOS cameras. It was a big step forward because this method - Digital holography - enables a full digital recording and processing of holograms without any intermediate step, such as photographic recording. Digital Holography can also be applied to interferometry

(Digital Holographic Interferometry) which offers much more possibilities than conventional Holographic Interferometry. Beyond classical advantage in digital signal processing like storing or filtering data, the major advantage is hidden in the possibility to directly calculate the phases of stored light waves from the digital holograms, without generating phase shifting holograms, which had to be done in conventional Holographic Interferometry.

The numerical reconstruction and subsequent further processing of reconstructed wave fields require a large number of mathematical operations and algorithms from very different branches of physics like photonics or image processing. The goal of the diploma thesis is to develop such an application that would comprise of all the necessary operations from the numerical reconstruction, through the calculation and the processing of phases, to the displaying of changes of physical quantities, like the deformation or the refraction index.

The whole text is organized into three chapters: First chapter, Fundamentals of Holography, presents theory necessary for Digital Holography and Digital Holography Interferometry, starting with the wave theory, describing such effects as interference, coherence, diffraction, and how these are employed in Digital Holography. Further the meaning of interference phase in Digital Holographic Interferometry and its connection to measurements of other physical quantities is explained. In second chapter, Application for Deformation Measurement using Digital Holographic Interferometry, detailed description of a single function included in the application is described step by step. And the last chapter, Experiment, is devoted to application of the software to a wide range of experimental data.

1 Fundamentals of Holography

In this chapter the physical basics of holography and holographic interferometry are described. The primary phenomena of holography are interference and diffraction. Such effects are based on the wave nature of light. So this chapter begins with a description of the wave theory of light. Terms for understanding the recording and reconstruction of holograms and the effect of holographic interferometry are presented.

1.1 Light wave

Light is an electromagnetic wave of such a wavelength that is visible to the human eye, see Fig. 1.

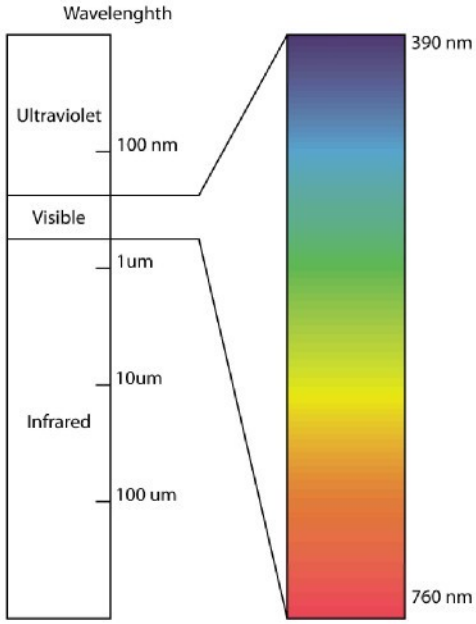

Fig. 1: Electromagnetic spectrum with highlighted optical wavelengths and its corresponding colors.

Generally it can be described by a set of partial differential equations called Maxwell´s equations. Considering the propagation of light waves in vacuum without bound charge and any bound current we obtain:

$$\begin{aligned} \nabla \times \mathbf{H} &= \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}, \\ \nabla \times \mathbf{E} &= -\mu_0 \frac{\partial \boldsymbol{H}}{\partial t}, \\ \nabla . \mathbf{E} &= 0, \\ \nabla . \mathbf{H} &= 0, \end{aligned} \tag{1.1}$$

where $(\nabla \times)$ is an operator of curl, $(\nabla .)$ is an operator of divergence, $\boldsymbol{E}, \boldsymbol{H}$ is a vector of intensity of electric resp. magnetic field, $\mu_0 = 4\pi \times 10^{-7} Hm^{-1}$ is a vacuum permeability and finally $\varepsilon_0 = 8{,}85 \times 10^{-12} Fm^{-1}$ is a vacuum permittivity.

Using properties of curl (vector identity): $\nabla \times (\nabla \times \boldsymbol{E}) = \nabla . (\nabla . \boldsymbol{E}) - \nabla^2 \mathbf{E}$ and the Maxwell´s equations (1.1) we can derive a wave equation:

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0, \tag{1.2}$$

where $\mathbf{c} = \sqrt{\frac{1}{\mu_0 \varepsilon_0}} = 299\,792\,458\, ms^{-1}$ is a constant - speed of light, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator and $\boldsymbol{x, y, z}$ and $\boldsymbol{t}$ are the Cartesian spatial coordinates resp. the temporal coordinate.

Transverse waves oscillate perpendicular to the direction of propagation, so the vector notation must be used. The wave may oscillate horizontally, vertically, or in any combination of these. Such effect is known as polarization. Fortunately, for most application we can assume a wave oscillating in only one direction. Such a wave is called plane polarized. For a plane-polarized wave propagating in the z-direction the following scalar wave equation is sufficient:

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = 0. \tag{1.3}$$

The most important solution of such wave equation is a harmonic wave. If we consider oscillating source of electromagnetic waves in position $\boldsymbol{z = 0}$:

$$E(0, t) = E_0 \cos(\omega t + \varphi),$$

then for the point in the distance $\mathbf{z}$ from the source it follows that:

$$E(z, t) = E_0 \cos(\omega t - kz + \varphi), \tag{1.4}$$

where E_0 is a real amplitude of the wave and the term $(\omega t - kz + \varphi)$ gives the phase of the wave. The wave number k is $= \frac{2\pi}{\lambda}$, where λ is a wavelength.

The use of trigonometric functions leads to difficult calculations, which can be avoided by using the Euler´s formula:

$$e^{j\varphi} = cos\varphi + jsin\varphi, \tag{1.5}$$

where $j = \sqrt{-1}$ is an imaginary unit. Therefore the equation for harmonic wave in complex domain U (where function U represents any of components of electric intensity $\boldsymbol{E}$ or any of components of magnetic intensity $\boldsymbol{H}$) can be written as:

$$\begin{aligned} U(z,t) &= U_0 e^{j(\omega t - kz + \varphi)} = \\ &= E_0 cos(\omega t - kz + \varphi) + jE_0 sin(\omega t - kz + \varphi) \end{aligned} \tag{1.6}$$

Hence it can be seen that $E(z,t) = Re\{U(z,t)\} = \frac{1}{2}[U(z,t) + U^*(z,t)]$. During most of the phenomena in physical optics the frequency does not change so we can omit the term $e^{-j(\omega t)}$ and we can consider only the so called complex amplitude independent on time:

$$U(z) = U_0 e^{j(kz - \varphi)}, \tag{1.7}$$

or in general form:

$$U(\boldsymbol{r}) = U_0 e^{j(\boldsymbol{kr} - \varphi(\boldsymbol{r}))}, \tag{1.8}$$

where $\boldsymbol{r} = (x,y,z)$ is a position vector, $\boldsymbol{k} = (k_x, k_y, k_z)$ and simultaneously $|\boldsymbol{k}| = \frac{2\pi}{\lambda} = k$ is a wave number.

After establishing of equation $E(\boldsymbol{r},t) = U(\boldsymbol{r},t) = U(\boldsymbol{r})e^{-j(\omega t)}$ to wave equation (1.2) we obtain differential equation:

$$(\nabla^2 + \mathrm{k}^2)U(\boldsymbol{r}) = 0, \tag{1.9}$$

called Helmholtz equation. All the waves described by the complex amplitude $U(\boldsymbol{r})$ must obey Helmholtz equation.

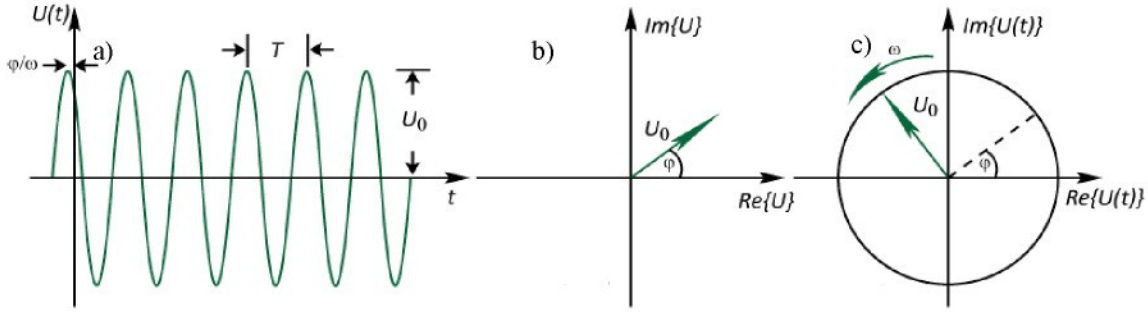

Fig. 2: Representations of a monochromatic wave at a fixed position $\boldsymbol{r}$*: a) the wave function* $\boldsymbol{U(t)}$ *is a harmonic function of time; (b) the complex amplitude* $\boldsymbol{U(r)}$ *is a fixed phasor; (c) the complex wave function* $\boldsymbol{U(t)}$ *is a phasor rotating with angular velocity* $\boldsymbol{\omega}$.

1.2 Intensity of light

The only parameter of light wave which is directly able to affect sensors – eye, photodiode, CCD target, etc. - is the intensity. Intensity I is defined by the energy flux through an area per time. From the Maxwell equations we get:

$$I = \varepsilon_0 c E^2. \tag{1.10}$$

Speed of light and permittivity of vacuum are constants so we just use:

$$I \approx U^2 = E^2. \tag{1.11}$$

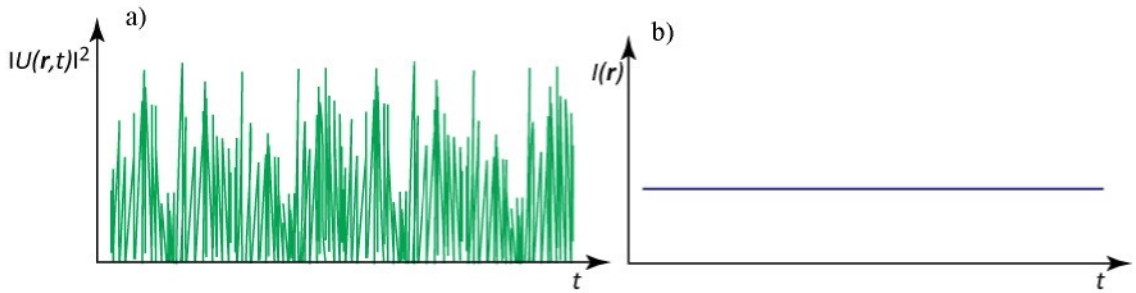

Fig. 3: Intensity of a light wave: a) immediate changes of intensity with very high frequency; b) intensity averaged by a detector which does not vary with time.

We can see that the intensity has a nonlinear dependence on the complex amplitude strength. No detector is able to follow fast changes of intensity because of the very high frequency of light. The measured value of intensity is always a result of integration over a measuring time T_m and the immediate value of intensity is not measurable, see Fig. 3. We can measure only the mean value. If $T_m \gg T = \frac{2\pi}{\omega}$ we define I as:

$$I = UU^* = |U|^2, \tag{1.12}$$

where * denotes complex conjugate. The intensity of a general wave field is:

$$I(\boldsymbol{r}) = \langle U(\boldsymbol{r})U(\boldsymbol{r})^* \rangle = \frac{1}{T_m} \int_{-\frac{T_m}{2}}^{\frac{T_m}{2}} U(\boldsymbol{r},t)U^*(\boldsymbol{r},t)dt. \tag{1.13}$$

1.3 Interference of coherent waves

Interference appears if two or more coherent light waves occur in a space. The interfering waves must be coherent for the interference to appear. Otherwise the interference pattern changes very quickly with time and space and the result is just the mean value of intensity as was mentioned above.

The resulting wave function is described by a sum of each single wave. This basic principle of superposition results from the linearity of the wave equation (1.2). In the case of monochromatic waves with the same frequencies the principle of superposition for complex amplitudes (1.8) is still valid. It conforms to the linearity of Helmholtz equation (1.9). If we consider two monochromatic waves with complex amplitudes $U_1(\boldsymbol{r})$, $U_2(\boldsymbol{r})$, the resulting wave is also monochromatic (with the same frequency) and its complex amplitude is:

$$U(\boldsymbol{r}) = U_1(\boldsymbol{r}) + U_2(\boldsymbol{r}) \tag{1.14}$$

According to formula (1.12) intensities of interfering waves are $I_1 = |U_1|^2$ resp. $I_2 = |U_2|^2$ and for the resulting intensity we get:

$$\begin{aligned} I(\boldsymbol{r}) &= |U(\boldsymbol{r})|^2 + |U_1(\boldsymbol{r}) + U_2(\boldsymbol{r})|^2 = \\ &= |U_1(\boldsymbol{r})|^2 + |U_2(\boldsymbol{r})|^2 + U_1(\boldsymbol{r})U_2{}^*(\boldsymbol{r}) + U_1{}^*(\boldsymbol{r})U_2(\boldsymbol{r}) \end{aligned} \tag{1.15}$$

Using $U_1 = \sqrt{I_1}e^{j\varphi_1}$ and $U_2 = \sqrt{I_2}e^{j\varphi_2}$ where φ_1, φ_2 are corresponding wave phases, for the detectable intensity we obtain an equation:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}cos\varphi, \tag{1.16}$$

where $\varphi = \varphi_2 - \varphi_1$. This relation (1.16) called interference equation can also be interpreted geometrically as the phasor diagram which demonstrates that the magnitude of the phasor U is sensitive to the phase difference φ, not only to the magnitudes of the constituent phasors U_1, U_2, see Fig. 4b. It can be seen that the superposition of two waves does not correspond to sum of their intensities. Moreover, the additional component $2\sqrt{I_1 I_2}cos\varphi$ appears. This term can be positive or negative and it influences the brightness of certain parts of interference pattern.

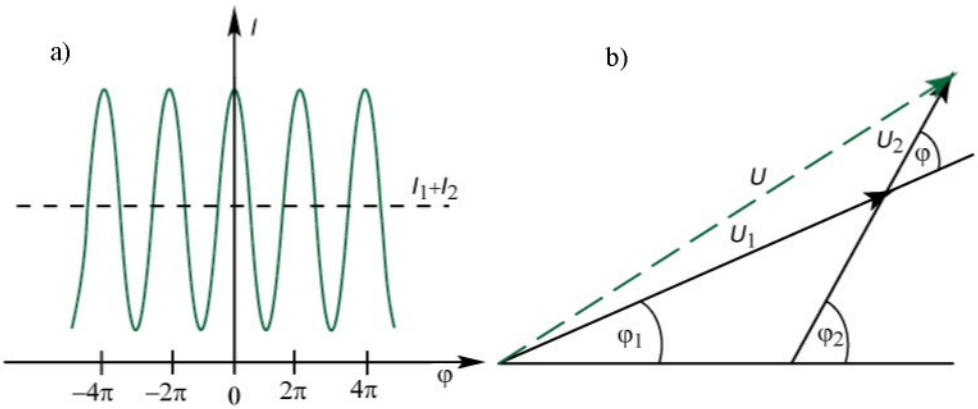

Fig. 4: Interference of two waves: a) relation of total intensity I *on the phase difference* φ*; b) phasor diagram for the superposition of two waves* U_1, U_2 *and phase difference* $\varphi = \varphi_2 - \varphi_1$.

1.4 Coherence

With common light sources such as the sun or a light bulb we observe interference only seldom. Only light of sufficient coherence will exhibit this effect. Coherence is defined more generally by the correlation properties between quantities of an optical field. Light is classified as coherent, incoherent, or in general, partially coherent. The two viewpoints of the general coherence are the temporal and the spatial coherence.

1.4.1 Temporal Coherence

Temporal coherence describes the correlation of a wave with itself (autocorrelation) as it behaves at different points in time. The autocorrelation function of a stationary complex random function $U(t)$ is the average of the product of $U^*(\tau)$ and $U(t+\tau)$ as a function of time delay τ. Result of such a function is so-called temporal coherence function:

$$G(\tau) = \langle U^*(t)U(t+\tau)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} U^*(t)U(t+\tau)\,d\tau. \tag{1.17}$$

It is easy to show that intensity I is equal to $G(\tau)$ when $\tau = 0$, see 1.13.

$$G(0) = \langle U^*(t)U(t)\rangle = \langle |U(t)|^2\rangle = I. \tag{1.18}$$

The temporal coherence function $G(\tau)$ carries information about the intensity I and also the degree of coherence. To dispose of relation on the intensity we can normalize the temporal coherence function to get complex degree of temporal coherence $g(\tau)$:

$$g(\tau) = \frac{G(\tau)}{G(0)} = \frac{\langle U^*(t)U(t+\tau)\rangle}{\langle U^*(t)U(t)\rangle}. \tag{1.19}$$

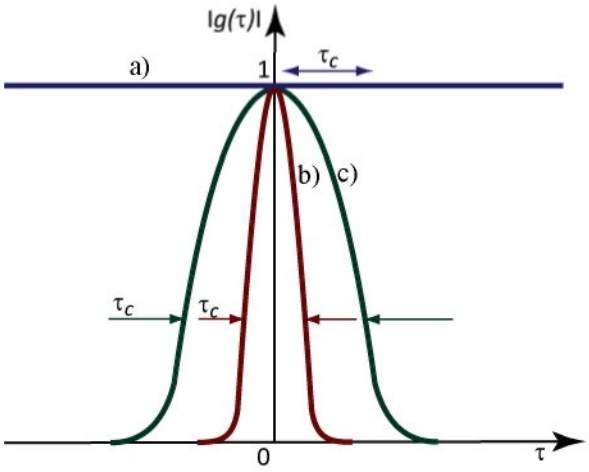

Fig. 5: Three random waves, the magnitudes of their complex degree of temporal coherence $|g(\tau)|$: a) coherent wave with infinite coherent time τ_c; b) ,c) partly coherent light waves with different coherent times τ_c.

The absolute value must be in the $\langle 0,1 \rangle$ interval. If complex degree of temporal coherence $g(\tau)$ decreases monotonically with time delay, the time τ_c at which it falls to a certain value (for example $\frac{1}{e}$ that corresponds to a bandwidth) is the coherence time τ_c. For $\tau < \tau_c$ the light wave is strongly correlated, otherwise for $\tau > \tau_c$ the correlation is weaker. It follows that the coherence time of monochromatic light is infinite, since $|g(\tau)| = 1$ everywhere.

Light, for which the coherence time τ_c is much longer than the differences of the time delays appearing in the optical system can be seen as completely coherent. Thus light is coherent if the distance l_c, is much greater than whole optical path-length. The distance l_c is known as the coherent length:

$$l_c = c\tau_c. \tag{1.20}$$

Effect of coherence can be seen during interference. From equations (1.15) and (1.19) we can derive interference equation for partially coherent waves:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\, Re\{g_{12}\} = I_1 + I_2 + 2\sqrt{I_1 I_2}|g_{12}|cos\varphi, \tag{1.21}$$

where $\varphi = \varphi_2 - \varphi_1$ and $|g_{12}|$ is a magnitude of the normalized cross-correlation. It is obviously recognized for the monochromatic wave $|g_{12}| = 1$, that the interference relation is consistent to interference equation (1.16).

1.4.2 Spatial coherence

Spatial coherence describes correlation of different parts of the same wave front, so the mutual intensity $G(r_1, r_2)$ is sufficient:

$$G(r_1, r_2) = \langle U^*(r_1)U(r_2)\rangle, \tag{1.22}$$

and in its normalized form:

$$g(r_1, r_2) = \frac{G(r_1, r_2)}{\sqrt{I_1(r_1)I_2(r_2)}}. \tag{1.23}$$

The magnitude $|g(\boldsymbol{r_1}, \boldsymbol{r_2})|$ is bounded between zero and one $\mathbf{0} < |g(\boldsymbol{r_1}, \boldsymbol{r_2})| < \mathbf{1}$. If the complex wavefunction $U(\boldsymbol{r}, t)$ is deterministic, $|g(\boldsymbol{r_1}, \boldsymbol{r_2})| = \mathbf{1}$ for all $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$, then the light is completely correlated everywhere.

The spatial coherence of quasi-monochromatic light (its frequencies are strongly peaked about a certain frequency) is described by $|g(\boldsymbol{r_1}, \boldsymbol{r_2})|$ as a function of the distance $|\boldsymbol{r_1} - \boldsymbol{r_2}|$. This function is equal to one when $\boldsymbol{r_1} = \boldsymbol{r_2}$ and falls if $|\boldsymbol{r_1} - \boldsymbol{r_2}|$ increases. The set of points, where the value $|g(\boldsymbol{r_1}, \boldsymbol{r_2})|$ is bigger than certain value (for example $\frac{1}{e}$) is called the coherence area A_c. The coherence area is very important parameter of light wave (similar to coherence length in case of the temporal coherence). It must be considered in relation to other dimensions of our optical system. For example if the coherence area is greater than the size of the aperture through which light is transmitted, the light can be regarded as coherent and vice versa.

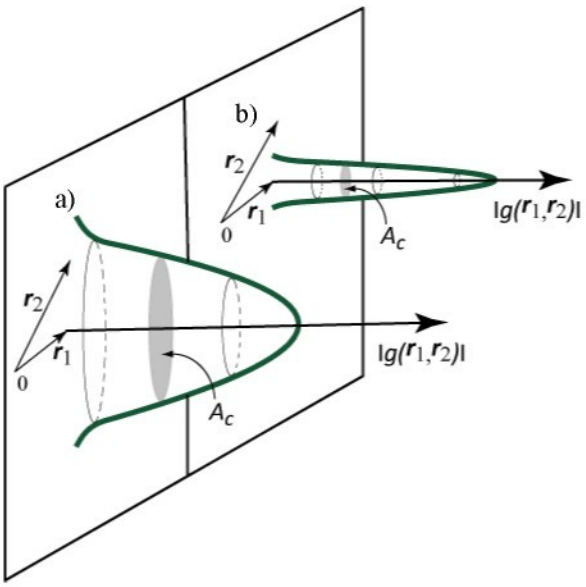

Fig. 6: Two examples of the magnitude of the normalized mutual intensity $|g(\boldsymbol{r_1}, \boldsymbol{r_2})|$ *as a function of* $\boldsymbol{r_1}$ *in the vicinity of a fixed point* $\boldsymbol{r_2}$*. The coherence area* $\boldsymbol{A_c}$ *in (a) is larger than in b).*

1.5 Light waves as a linear system

Consider a plane wave with complex amplitude

$$U(x, y, z) = \mathrm{U}_0 e^{-j(k_x x + k_y y + k_z z)}, \tag{1.24}$$

where k_x, k_y, k_z are components of a wave vector $\boldsymbol{k} = (k_x, k_y, k_z) = 2\pi(v_x, v_y, v_z)$ with spatial frequencies v_x, v_y, v_z in related directions, see Fig. 7. Magnitude of the wave vector is called wave number

$$|\boldsymbol{k}|^2 = {k_x}^2 + {k_y}^2 + {k_z}^2 = \left(\frac{2\pi}{\lambda}\right)^2 = (2\pi)^2({v_x}^2 + {v_y}^2 + {v_z}^2). \qquad (1.25)$$

Furthermore λ denotes wavelength and $\mathbf{U_0}$ is complex envelope. The vector $\boldsymbol{k}$ forms angles

$$\theta_x = arcsin\frac{k_x}{k} = arcsin\, v_x\,\lambda \approx v_x\,\lambda, \qquad (1.26)$$

with the y-z planes and

$$\theta_y = arcsin\frac{k_y}{k} = arcsin\, v_y\,\lambda \approx v_y\,\lambda, \qquad (1.27)$$

with x-z planes. Therefore the angles of the wave vector are related to the spatial frequencies of the harmonic function.

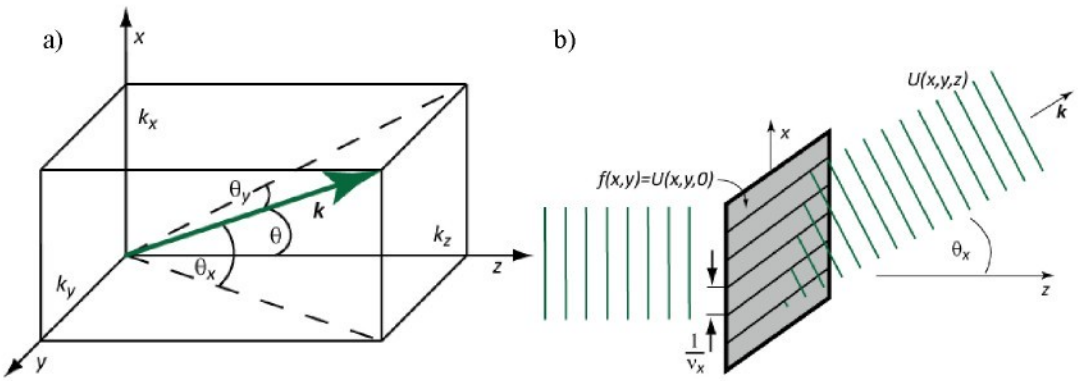

Fig. 7: a) The wave vector $\boldsymbol{k}$ decomposed into components k_x, k_y, k_z and illustration of angles $\theta_x = arcsin\frac{k_x}{k}$ and $\theta_y = arcsin\frac{k_y}{k}$ respectively; b) A thin element f(x,y) with spatial frequency v_x at plane z=0 is corresponding to a plane wave traveling at angle θ_x.

Consider a thin optical element (in our case hologram) with the complex amplitude transmittance (superposition of harmonic functions), see Fig. 7b, then:

$$f(x,y) = \iint_{-\infty}^{\infty} \mathrm{U_0}(v_x, v_y) e^{-j2\pi(v_x x + v_y y)}\, dv_x\, dv_y, \qquad (1.28)$$

in the z = 0 plane. Then the wave is modulated by the harmonic function, so $U(x,y,0) = f(x,y)$. Further the transmitted wave $U(x,y,z)$ is the superposition of plane waves, at $z > 0$

$$U(x,y,z) = \iint_{-\infty}^{\infty} \mathrm{U}_0(v_x, v_y) e^{-j2\pi(v_x x + v_y y)} e^{-jk_z z} dv_x\, dv_y, \tag{1.29}$$

with complex envelope $\mathrm{U}_0(v_x, v_y)$, where $k_z = 2\pi\sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2}$, see 1.25.

Now we examine the propagation of a monochromatic optical wave of wavelength λ and complex amplitude $U(x,y,z)$ in the free space between the planes z = 0 and z = d, called the input and output planes. The complex amplitude of the wave at the input plane $f(x,y) = U(x,y,0)$ is given and from 1.31 we will determine the complex amplitude at the output plane denoting $g(x,y)$, where $g(x,y) = U(x,y,d)$. We regard $f(x,y)$ and $g(x,y)$ as the input and output of a linear system, see Fig. 8.

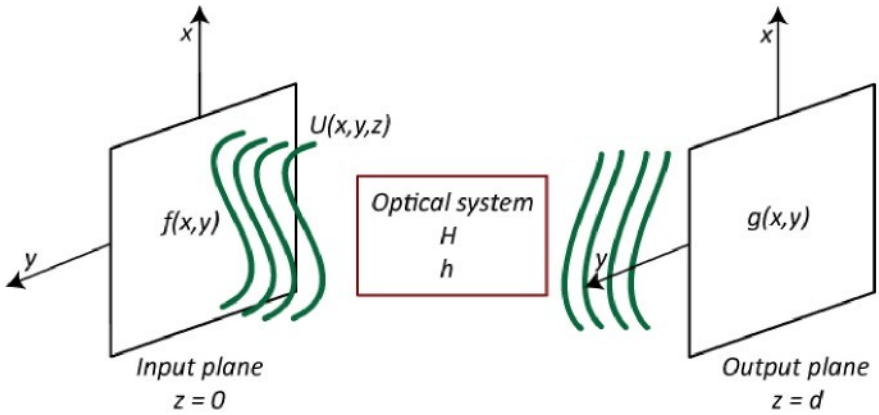

Fig. 8: The transmission of a light wave $U(x,y,z)$ through an optical system between an input plane z = 0 and an output plane z = d. It can be regarded as a linear system, whose input and output are the functions $f(x,y) = U(x,y,0)$ and $g(x,y) = U(x,y,d)$.

The system is linear because $U(x,y,z)$ must satisfy the Helmholtz equation (1.9). The system is shift-invariant because of the invariance of free space to displacement of the coordinate system. A linear shift-invariant system is characterized by its impulse response function $h(x,y)$ or by its transfer function $H(v_x, v_y)$:

$$H(v_x, v_y) = \frac{g(x,y)}{f(x,y)} = e^{-j2\pi d\sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2}}. \tag{1.30}$$

So, the output – input relation can be determined as:

$$g(x,y) = f(x,y)H(v_x, v_y) = \tag{1.31}$$
$$= \iint_{-\infty}^{\infty} \mathrm{U}_0(v_x, v_y) e^{-j2\pi(v_x x + v_y y)} e^{-j2\pi d\sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2}} dv_x\, dv_y.$$

The impulse response function $h(x,y)$ of the system of free space propagation is the response $g(x,y)$ if the input $f(x,y)$ is at point $(0,0)$. It is the inverse Fourier transform of the transfer function $H(v_x, v_y)$:

$$h(x,y) = \mathcal{F}^{-1}\{H(v_x, v_y)\}. \tag{1.32}$$

The expression for the transfer function in (1.32) may be simplified if the input function $f(x,y)$ contains only spatial frequencies that are much smaller than the cutoff frequency $\frac{1}{\lambda}$, which means $\frac{1}{\lambda^2} \gg v_x^2 + v_y^2$. Then the plane-wave components of the propagating light make small angels corresponding to paraxial rays. Then the following formula holds true:

$$\theta^2 = \theta_x^2 + \theta_y^2 \approx \lambda^2(v_x^2 + v_y^2) \Rightarrow \frac{\theta^2}{\lambda^2} = v_x^2 + v_y^2, \tag{1.33}$$

where θ is the angle with the optical axis, see Fig. 7a. The phase factor of the transfer function written as a Taylor series is

$$-j2\pi\frac{d}{\lambda}\sqrt{1-\theta^2} \approx -j2\pi\frac{d}{\lambda}\left(1 - \frac{\theta^2}{2} + \frac{\theta^4}{8} - \cdots\right). \tag{1.34}$$

Omitting the third and higher terms of this expansion, the transfer function $H(v_x, v_y)$ may be approximated by

$$H(v_x, v_y) \approx e^{-j2\pi(\frac{1}{\lambda} - \frac{\lambda}{2}(v_x^2 + v_y^2))d} = e^{-j\frac{2\pi}{\lambda}d} e^{j\pi\lambda d(v_x^2 + v_y^2)}. \tag{1.35}$$

Likewise the output - input relation can be determined as:

$$g(x,y) = f(x,y)H(\upsilon_x, \upsilon_y) = \tag{1.36}$$
$$= e^{-j\frac{2\pi}{\lambda}d} \iint_{-\infty}^{\infty} \mathrm{U}_0(\upsilon_x, \upsilon_y) e^{-j2\pi(\upsilon_x x + \upsilon_y y)} e^{j\pi\lambda d(\upsilon_x^2 + \upsilon_y^2)} d\upsilon_x \, d\upsilon_y.$$

This approximation is known as the Fresnel Approximation. Hence, the impulse response $h(x,y)$ can be derived as:

$$h(x,y) = \mathcal{F}^{-1}\{H(\upsilon_x, \upsilon_y)\} = \frac{j}{\lambda d} e^{-jkd} e^{-jk\frac{x^2+y^2}{2d}}. \tag{1.37}$$

Therefore the free space propagation can be described as a convolution:

$$g(x,y) = f(x,y) * h(x,y), \tag{1.38}$$

where the operator * is a convolution operator.

1.6 Diffraction theory

When an optical wave is transmitted through an aperture in an opaque screen and propagates some distance in free space, its intensity distribution is called the diffraction pattern. If light were propagated as rays, the diffraction pattern would be a shadow of the aperture. Because of the wave nature of light, the diffraction pattern may deviate from the aperture shadow, depending on the distance between the aperture and observation plane, the wavelength, and the dimensions of the aperture, see Fig. 9.

The simplest theory of diffraction is based on the assumption that the impinging wave is transmitted without change at points within the aperture, but is reduced to zero at points on the back side of the opaque part of the screen. If $U(\mathrm{x},\mathrm{y})$ and $f(\mathrm{x},\mathrm{y})$are the complex amplitudes of the wave directly to the left and right of the screen, then:

$$f(\mathrm{x},\mathrm{y}) = U(\mathrm{x},\mathrm{y})p(\mathrm{x},\mathrm{y}), \tag{1.39}$$

where $p(\mathrm{x},\mathrm{y})$ is called the aperture function defined as $p(\mathrm{x},\mathrm{y}) = 1$ inside the aperture and $p(\mathrm{x},\mathrm{y}) = 0$ outside the aperture.

Function $f(x,y)$ is usually given, the complex amplitude $g(x,y)$ at an observation plane at distance d from the screen may be determined using the methods of free space light propagation, see 1.31.

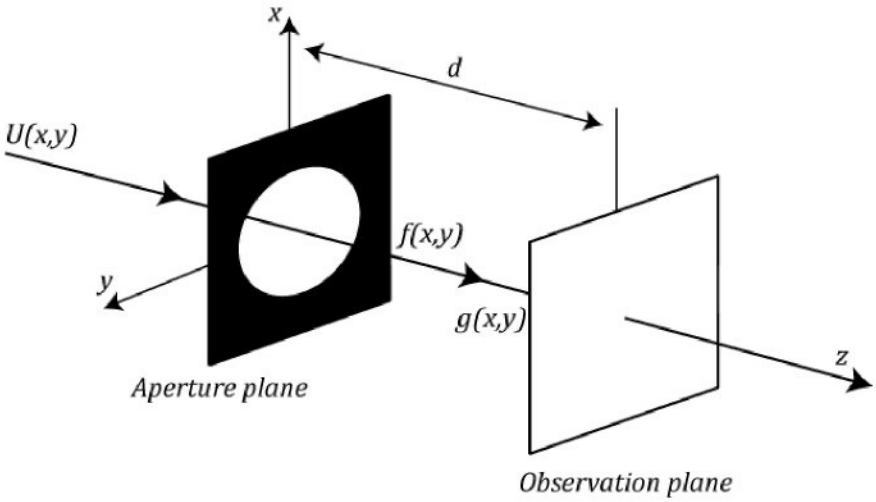

Fig. 9: A wave is transmitted through an aperture of amplitude transmittance $p(x,y)$*, generating a wave with complex amplitude* $f(x,y) = U(x,y)p(x,y)$ *.After propagation in free space in the distance d there is a complex amplitude* $g(x,y)$ *and finally the diffraction pattern is the intensity* $I(x,y) = |g(x,y)|^2$.

The diffraction pattern $I(x,y) = |g(x,y)|^2$ is known as Fraunhofer diffraction or Fresnel approximation. It depends on whether free space propagation is described. The Fresnel approximation (1.34) replaces the secondary spherical waves by waves with parabolic wavefronts, while the Fraunhofer approximation uses even plane wavefronts. For our purpose we use the Fresnel diffraction, which is uniquely less restrictive than the Fraunhofer one, so we will use the formula (1.36). The condition for Fresnel approximation is fulfilled when the third term in 1.34 is much smaller than π:

$$\frac{d\theta^4}{4\lambda} \ll 1. \tag{1.40}$$

If a is the largest distance in output plane and d is the distance, then the largest angle $\theta_m \approx \frac{a}{d}$ and it may be rewritten as:

$$\frac{N_F {\theta_m}^2}{4} \ll 1, \tag{1.41}$$

where $N_F = \frac{a^2}{\lambda d}$ is called Fresnel number. For completeness' the above mentioned Fraunhofer diffraction must fulfil the following condition:

$$N_F \ll 1. \tag{1.42}$$

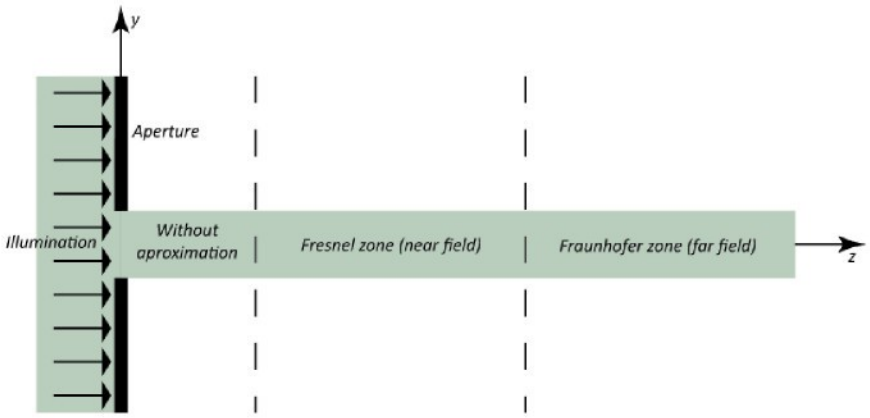

Fig. 10: Relative positions of Fresnel and Fraunhofer zones.

1.7 Holography

Holography involves the recording and reconstruction of optical waves. A hologram is a transparency containing a coded record of the optical wave.

Consider a monochromatic optical wave with complex amplitude $\mathbf{U_0(x,y)}$ in a plane, where $\mathbf{z = 0}$. If a thin optical element with complex amplitude transmittance $\mathbf{h(x,y)}$ equal to $\mathbf{U_0(x,y)}$ exists it would be possible to provide a complete record of the wave. Then the wave could be reconstructed by illuminating the transparency (thin element) with a uniform plane wave of unit amplitude propagating in the $\mathbf{z}$ direction. The transmitted wave would have a complex amplitude in the $\mathbf{z = 0}$ plane $\mathbf{U(x,y) = h(x,y) = U_0(x,y)}$. Then the original wave could be reproduced everywhere in the space $\mathbf{z \geq 0}$.

The question is how to make a transparency $\mathbf{h(x,y)}$ from the original wave $\mathbf{U_0(x,y)}$. The main problem is that optical detectors like photographic emulsions, human eye or CCD sensor are responsive only to the optical intensity $\mathbf{|U_0(x,y)|^2}$ and therefore they are insensitive to the phase $\boldsymbol{arg\{U_0(x,y)\}}$. But the phase is evidently important and cannot be disregarded. To record the phase of $\mathbf{U_0(x,y)}$, a code that transforms phase into intensity must be found. Then the recorded information could be optically decoded in order to reconstruct the wave. Such a code is based on combination

of the original wave - the object wave $\mathbf{U_0}$ with a known reference wave U_r and recording their interference pattern in the $\mathbf{z = 0}$ plane. The intensity of the sum of the two waves is recorded and a transparency of complex amplitude transmittance $\mathbf{h}$ is obtained. The transmittance is then defined as:

$$h \approx |U_o + U_r|^2 = |U_o|^2 + |U_r|^2 + U_o U_r^* + U_o^* U_r = \\ = I_0 + I_r + 2\sqrt{I_0 I_r}\cos(\varphi_r - \varphi_0) \tag{1.43}$$

where I_r and I_0 are the intensities of the reference wave and the object wave in the $\mathbf{z = 0}$ plane. The transparency h called a hologram carries a coded information about the intensity and phase of the wave U_o and the hologram h is highly sensitive to the difference between the phases of the two waves.

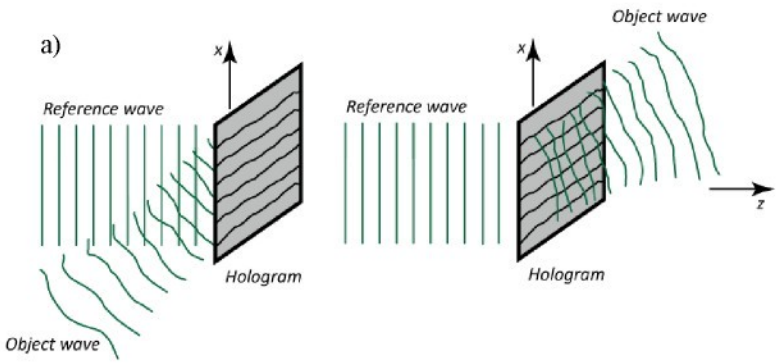

Fig. 11: The hologram: a) transparency on which the interference pattern between the object wave and the reconstruction wave is recorded; b) the object wave is reconstructed by illuminating of the hologram with the reference wave.

To decode the information in the hologram and reconstruct the object wave, the reference wave U_r must be used to illuminate the hologram. The result is a wave in the hologram plane $\mathbf{z = 0}$ with complex amplitude:

$$U = hU_r \approx U_r I_r + U_r I_0 + I_r U_0 + {U_r}^2 {U_0}^*. \tag{1.44}$$

The term $I_r U_0$ is the original wave multiplied by the intensity of the reference wave I_r. If I_r is uniform this term constitutes the desired reconstructed wave. The term ${U_r}^2 {U_0}^*$ is a conjugated version of the original wave modulated by ${U_r}^2$ and the terms $U_r I_r$, $U_r I_0$ represent the reference wave, modulated by the sum of the intensities of the two waves I_r and I_0.

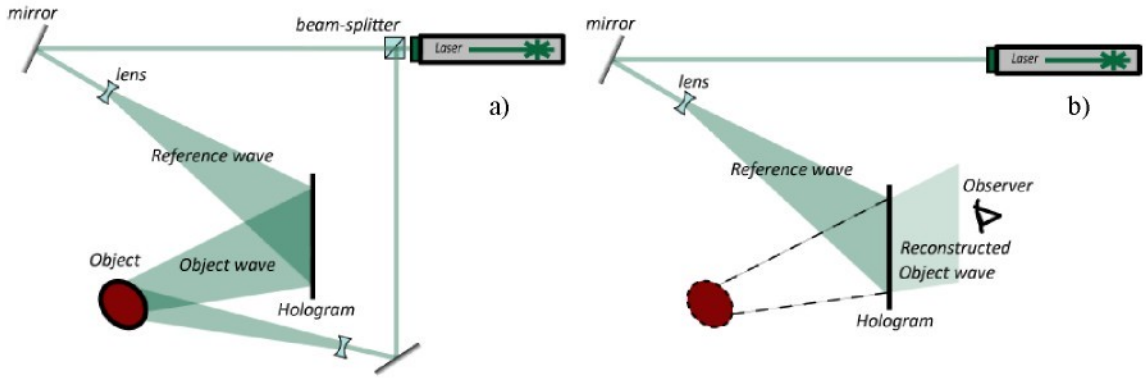

Fig. 12: The holographic setup: a) The object wave U_o scattered from the object and the reference wave U_r interfere on the recording medium. So the hologram is created; b) Object wave is reconstructed by illumination the hologram with the reference wave. The object appears in the same position relative to the hologram as it was positioned during recording.

1.8 Digital holography

The aim of digital holography is to record holograms, which are later stored in a computer memory and then can be reconstructed numerically afterwards. The digital recording medium is usually CCD camera. In this chapter main differences between digital and conventional recording holographic setups and the following reconstruction process are introduced.

1.8.1 Digital holography recording

The basic principle of digital hologram recording is the same as in the conventional holography. Only the recording medium is different. The hologram is the microscopic interference pattern generated by the coherent superposition of an object and a reference wave as was mentioned in the previous chapter.

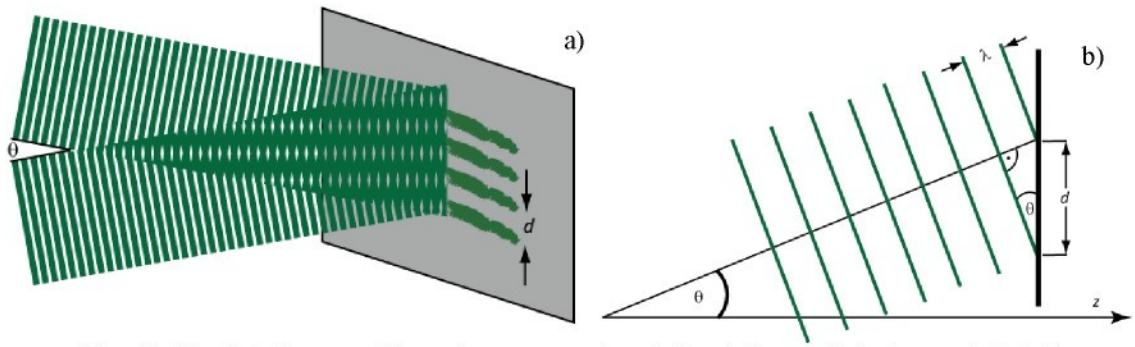

Fig. 13: The interference of two plane waves at angle θ: a) the result is sine period d; b) simplification for easier determination of period d, see 1.45

The spatial frequency of such interference pattern is defined by the angle between these two waves:

$$d = \frac{\lambda}{sin\theta}. \tag{1.45}$$

A meaningful sampling of the intensity distribution constituting the hologram is only guaranteed if the sampling theorem is followed. The sampling theorem requires that the spatial period **d** must be sampled with more than two pixels:

$$d > 2\Delta\xi, \tag{1.46}$$

where $\Delta\xi$ is the pixel size. For small values of angle θ we can consider $sin\theta \approx \theta$. Hence the limit value for the angle θ_{max}, which is the maximum angle formed by reference and object wave when the sampling theorem is followed, can be determined as:

$$\theta_{max} \approx \frac{\lambda}{2\Delta\xi} \tag{1.47}$$

Naturally, the sampling theorem must also be followed for conventional holography, but due to the resolution of digital and analog recording medium, the meeting the conditions of sampling theorem is much more exacting for digital holographic setup.

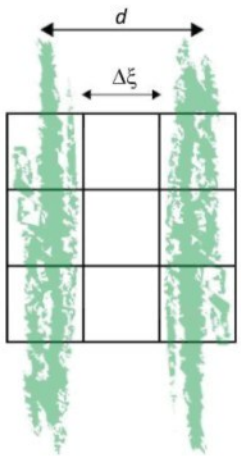

Fig. 14: The detailed illustration of a part of the CCD target: it can be recognized that the distance of interference frings d must be at least greater than pixel extension $\Delta\xi$ to fulfil the sampling theorem.

If we consider geometry shown in Fig. 15a with the object width d_0 in the x direction placed symmetrically to the optical axis, and a reference wave propagating along the optical axis and impinging orthogonally onto the CCD, we can calculate the maximum object width d_0 for a given distance d:

$$tan\theta = \frac{\frac{d_0}{2} + \frac{N\Delta\xi}{2}}{d}. \tag{1.48}$$

Together with the condition of maximum angle θ_{max} we obtain the following:

$$\frac{\lambda}{2\Delta\xi} > \frac{\frac{d_0}{2} + \frac{N\Delta\xi}{2}}{d}. \tag{1.49}$$

From that we can derive a formula:

$$d_0 < \frac{\lambda d}{\Delta\xi} - N\Delta\xi, \tag{1.50}$$

which describes the limit for a maximal width d_0 of the object with respect to distance d and simultaneously the sampling theorem for such parameters is fulfilled.

In practical applications of holographic metrology we often have objects with large surfaces. If so, the sampling theorem does not have to be fulfilled. In such cases the wave field scattered from the surface of object can be reduced by using a negative lens, see Fig 15b.

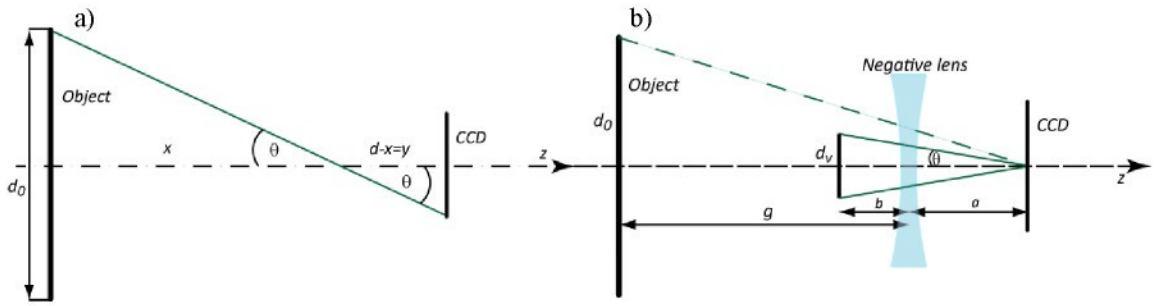

Fig. 15: The geometry for recording of digital hologram: a) holographic setup without a lens; b) holographic setup with reduced angle θ by using a negative lens.

The object wave field impinging onto the CCD target seems to come from the small virtual image of the object of width d_v and not from the large object with width d_0.

From the lens formula $\frac{1}{f}=\frac{1}{g}-\frac{1}{b}$, magnification formula $Z=\frac{d_v}{d_0}=\frac{f}{g-f}$ and the geometry $tan\theta=\frac{d_v}{2(a+b)}$ we can derive the term for distance between the lens and the CCD, see Fig. 15b.

$$a=\frac{fg}{g-f}-\frac{fd_0}{(g-f)2tan\theta} \tag{1.51}$$

Now, the distance between the CCD and the small virtual image of the object is $d=a+b$ instead of the original distance $d=a+g$. It can be seen, that $\theta_{reduced} \ll \theta_{lensless}$ and it is much easier to fulfill the sampling theorem.

1.8.2 Digital holography reconstruction

Although there exist other possibilities of reconstruction [11], here we will concentrate on the numerical calculation of the wave field. The theoretical tool for the numerical reconstruction is the diffraction theory.

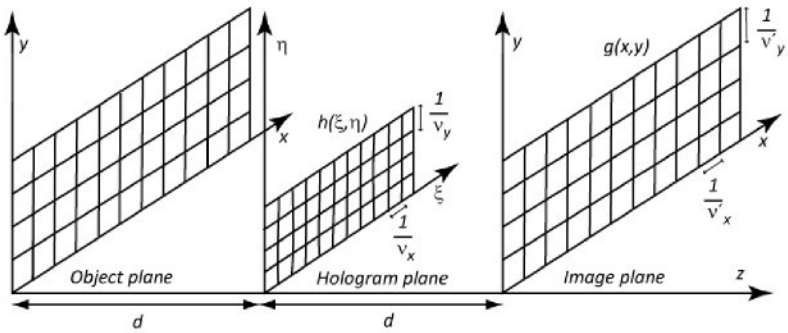

Fig. 16: The geometry for digital holography reconstruction.

Let the geometry for the numerical description is as in Fig. 16. At distance *d* from the object surface we have the hologram plane with coordinates (ξ,η). At distance *d* from this hologram plane we also have the image plane known as the observation plane, where the real image can be reconstructed. The coordinates of this plane are again (*x*, *y*). Consider the hologram with spatial frequencies $v_x=\frac{\xi}{\lambda d}$ and $v_y=\frac{\eta}{\lambda d}$, and their intensity distribution is $h(\xi,\eta)$. Such a hologram originates by interference of reference and object wave. To get real image of the object the reference wave must be complex conjugated $r^*(\xi,\eta)$, see 1.43. The reconstruction requires the illumination of the hologram with the reference wave. This process is now modeled numerically by a

multiplication of the digital hologram h with the reference wave r. From the formula (1.36) and with the use of mentioned hologram $h(\xi,\eta)$, we can derive the relation for the wave field in the image plane:

$$g(x,y) = \frac{e^{-j\frac{2\pi}{\lambda}d}}{\lambda d}\iint_{-\infty}^{\infty} h(\xi,\eta)\, r^*(\xi,\eta)\, e^{\frac{-j2\pi}{\lambda d}(\xi x+\eta y)}\, e^{\frac{j\pi}{\lambda d}(\xi^2+\eta^2)} d\xi\, d\eta, \qquad (1.52)$$

then after substitution $v'_x = \frac{x}{\lambda d}$ and $v'_y = \frac{y}{\lambda d}$, where v'_x, v'_y are spatial frequencies of relevant directions in image plane, we obtain the following:

$$g(x,y) = \frac{e^{-j\frac{2\pi}{\lambda}d}}{\lambda d}\iint_{-\infty}^{\infty} h(\xi,\eta)\, r^*(\xi,\eta)\, e^{-j2\pi(\xi v'_x+\eta v'_y)}\, e^{\frac{j\pi}{\lambda d}(\xi^2+\eta^2)} d\xi\, d\eta. \qquad (1.53)$$

With these substitutions we can see that the wave field $g(x,y)$ is calculated by a Fourier transform of the digital hologram $h(\xi,\eta)$ multiplied by the reference wave $r^*(\xi,\eta)$ and also by the so called chirp function $e^{\frac{j\pi}{\lambda d}(\xi^2+\eta^2)}$.

Because the reconstruction will be calculated numerically, we need a discrete form of the formula (1.53). With the use of the equations:

$$\xi = k\Delta\xi \; where \; 1 < k < N, \qquad (1.54)$$

$$\eta = l\Delta\eta \; where \; 1 < l < M,$$

for sampling of hologram plane and image plane:

$$x = n\Delta x \; where \; 1 < n < N,$$

$$y = m\Delta y \; where \; 1 < m < M, \qquad (1.55)$$

we obtain the reconstruction equation:

$$g(n\Delta x, m\Delta y) = \tag{1.56}$$
$$= \frac{e^{-j\frac{2\pi}{\lambda}d}}{\lambda d}\sum_{k=1}^{N}\sum_{l=1}^{M} h(k\Delta\xi, l\Delta\eta)r^*(k\Delta\xi, l\Delta\eta)e^{\frac{j\pi}{\lambda d}((k\Delta\xi)^2+(l\Delta\eta)^2)}e^{-j2\pi\left(\frac{kn}{N}+\frac{lm}{M}\right)},$$

where $\Delta\xi$, $\Delta\eta$ are CCD pixel extensions, Δx, Δy are the image plane pixel sizes and n, m, k, l are integers. In practical applications the pixel numbers N, M and the pixel size $\Delta\xi \times \Delta\eta$ is given by the CCD by default. Image plane distances $\Delta x, \Delta y$ are given by the size of a frame, where the reconstructed object wave is displayed.

Formula (1.56) defines the practical discrete finite calculation method for the reconstruction of the wave field coded in a digital hologram. The result $g(n\Delta x, m\Delta y)$ is a numerical representation of a complex optical wave field. Intensity $I(n\Delta x, m\Delta y)$ and phase $\varphi(n\Delta x, m\Delta y)$ distribution can be determined as:

$$I(n\Delta x, m\Delta y) = |g(n\Delta x, m\Delta y)|^2, \tag{1.57}$$

$$\varphi(n\Delta x, m\Delta y) = arctan\frac{Im\{g(n\Delta x, m\Delta y)\}}{Re\{g(n\Delta x, m\Delta y)\}}. \tag{1.58}$$

This is a real advantage compared to the optical reconstruction, because by the optical way we only obtain the intensity distribution.

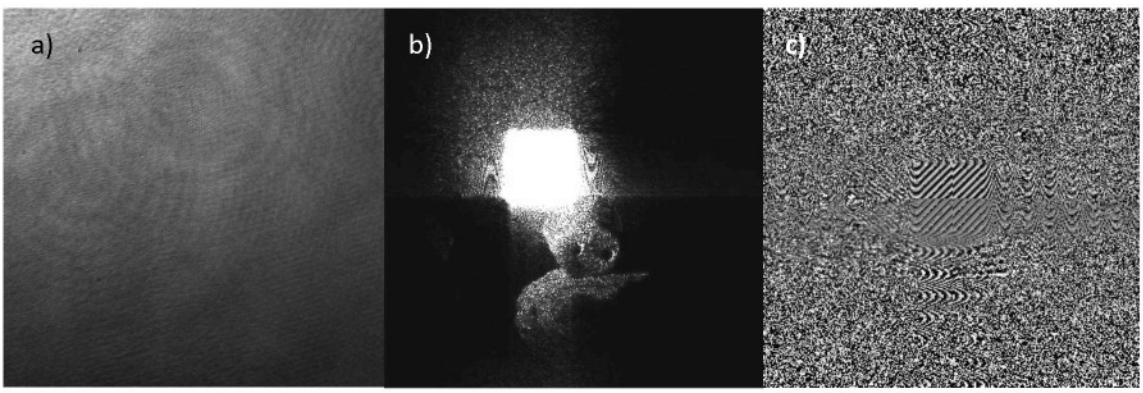

Fig. 17: the digital holography reconstruction: a) the digital hologram; b) the intensity distribution including: the real image (the dice), the virtual image (not focused), and the d.c. term (the bright square in the middle); c) the noisy phase distribution.

In the digital case we also have the possibility to calculate the phase modulo 2π. At first sight it seems the phase calculation has no importance, because for rough object surfaces it varies stochastically. But we will see that the phase calculation brings a real advantage when it comes to applications in digital holographic interferometry.

We should not forget that the reconstructed wave field includes information about the object wave (real image) as well as about its conjugated image (virtual image) and reference wave (d.c.-term), see Fig. 17. This is the result of the reconstruction interference equation (1.44). The real image corresponds to a wavefront converging to a sharp image, while the virtual image belongs to a divergent wavefront that seems to be not in focus. However the virtual image can also be reconstructed. Here the non-conjugated reference wave must be used, see (1.43). This can be done by the numerical calculation of the field in the $(z = -d)$ plane. Such a plane corresponds to the object plane according to Fig. 16. For more detailed information about the virtual image reconstruction, see [2].

Another component, which can be seen in the intensity distribution image as a bright central square, is a d.c.-term. It is much brighter than the reconstructed real or virtual images. The physical meaning of the bright central square is a representation of the zero-order diffraction of the reference wave (whereas real and virtual image represents the first or the minus first diffraction order). Similarly it can also be the projection of the illuminated CCD array. The bright square can be regarded as the undiffracted part of the reference wave or (from the calculation point of view) the *d.c. - term* of the Fresnel hologram. A mutual location of these components is given by the holographic setup and by angle of the reference and the object wave. If the real image is overlapped by the virtual one or by the d.c.-term we have to suppress these phenomena. We can change the holographic setup to have all components well separated or we can suppress them by numerical way instead [5].

1.9 Conventional Holographic Interferometry

The interference occurs if two mutually coherent waves are superposed and holography was introduced as a method for recording and reconstructing optical wave fields. Now these concepts can be combined to define holographic interferometry as the interferometric comparison of two or more wave fields.

There exist several kinds of holographic interferogram recording methods, but for our purposes we will focus at the so called Double Exposure Method. Other methods

like Time Average Holographic Interferometry or Real Time Holographic Interferometry are based on different principles and the processing of the data recorded in this way is not included in the developed application.

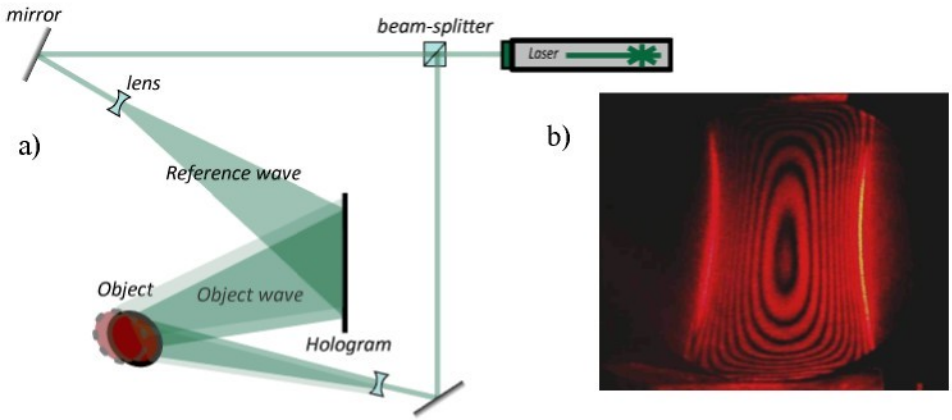

Fig. 18: The double exposure method: a) two different states of object are recorded on the same recording medium; b) the reconstructed holographic interferogram.

In the double exposure method of conventional holographic interferometry two wavefronts scattered by the same object are recorded consecutively onto the same holographic medium. These two wavefronts correspond to different states of the object, one in an initial condition and one after the change of a physical parameter. Consider the complex amplitude of the first wavefront in certain position *r* as:

$$U_1(\boldsymbol{r}) = \sqrt{I_1(\boldsymbol{r})}e^{j\varphi(\boldsymbol{r})} \tag{1.59}$$

and the complex amplitude of the changed wavefront:

$$U_2(\boldsymbol{r}) = \sqrt{I_1(\boldsymbol{r})}e^{j(\varphi(\boldsymbol{r})+\Delta\varphi(\boldsymbol{r}))}. \tag{1.60}$$

From the interference equation (1.15) the stationary intensity distribution is:

$$I(\boldsymbol{r}) = I_1(\boldsymbol{r}) + I_2(\boldsymbol{r}) + 2\sqrt{I_1(\boldsymbol{r})I_2(\boldsymbol{r})}\cos|\Delta\varphi(\boldsymbol{r})|, \tag{1.61}$$

where $\Delta\varphi(\mathbf{r})$ is the interference phase difference (or shortly interference phase) and represents the change of physical value. If the intensity of the first and the second wave

fields are equal $(I_1(\boldsymbol{r}) = I_2(\boldsymbol{r}))$, we can convert (1.61) into the resulting intensity distribution:

$$I(\boldsymbol{r}) = 2I_1(\boldsymbol{r})(1 + \cos|\Delta\varphi(\boldsymbol{r})|). \tag{1.62}$$

An example of such a holographic interferogram is shown in Fig. 18b. Bright centers of fringes are the contours, where the interference phase is an even integer multiplied by π, while the dark centers of fringes correspond to odd integer multiplied by π. From such an interference pattern we can determine the interference phase $\Delta\varphi$.

1.10 Digital Holographic Interferometry

The main advantage in digital holographic interferometry is that instead of a reconstruction of a double exposure hologram and evaluation of the resulting intensity field, the reconstructed phase fields can be directly calculated. The calculated interference phase is modulated by the value 2π. Further the interference phase modulo 2π can be demodulated. These demodulated interference phase distributions already carry the information about a change of a certain physical quantity.

1.10.1 Interference Phase Modulo 2π

In digital holographic interferometry the double exposure technique is realized digitally. For this purpose the two digital holograms $h_1(k\Delta\xi, l\Delta\eta)$, $h_2(k\Delta\xi, l\Delta\eta)$ representing two deformation states of the object surface are recorded one by one and then reconstructed numerically (see the chapter Digital Holography Reconstruction). Complex wave fields obtained in this way:

$$\begin{aligned} g_1(n,m) &= g_{r1}(n,m) + jg_{i1}(n,m), \\ g_2(n,m) &= g_{r2}(n,m) + jg_{i2}(n,m), \end{aligned} \tag{1.63}$$

can be compared numerically. There are two ways to process these wave fields numerically. Partly we can calculate phase distributions of the wave fields separately as:

$$\varphi_1(n,m) = arctan\frac{Im\{g_1(n,m)\}}{Re\{g_1(n,m)\}} = arctan\frac{g_{i1}(n,m)}{g_{r1}(n,m)}, \tag{1.64}$$

$$\varphi_2(n,m) = arctan\frac{Im\{g_2(n,m)\}}{Re\{g_2(n,m)\}} = arctan\frac{g_{i2}(n,m)}{g_{r2}(n,m)}$$

and then the interference phase modulo 2π can be determined by a point to point subtraction according to:

$$\Delta\varphi(n,m) = \tag{1.65}$$
$$= \begin{cases} \varphi_2(n,m) - \varphi_1(n,m) + 2\pi \;\; if\; \varphi_2(n,m) - \varphi_1(n,m) < -\pi \\ \varphi_2(n,m) - \varphi_1(n,m) - 2\pi \;\; if\; \varphi_2(n,m) - \varphi_1(n,m) \geq +\pi \\ \varphi_2(n,m) - \varphi_1(n,m) \;\; if\; \varphi_2(n,m) - \varphi_1(n,m) \in \langle -\pi\,, +\pi \rangle. \end{cases}$$

Here we have assured that the phase values are in the range $\langle -\pi\,, +\pi \rangle$, which is required by the range of the inverse trigonometric function arctan.
Another, more direct, way of the interference phase calculation is:

$$\Delta\varphi(n,m) = arctan\frac{Im\{g_2(n,m)g_1{}^*(n,m)\}}{Re\{g_2(n,m)g_1{}^*(n,m)\}} =$$
$$= arctan\frac{g_{r1}(n,m)g_{i2}(n,m) - g_{r2}(n,m)g_{i1}(n,m)}{g_{r1}(n,m)g_{r2}(n,m) + g_{i2}(n,m)g_{i1}(n,m)}. \tag{1.66}$$

The intensity distribution of the resulting complex wave field can also be calculated as:

$$I(n,m) = |g_2(n,m)g_1{}^*(n,m)|^2. \tag{1.67}$$

The intensity distribution of this wave field corresponds to the interference pattern of conventional optical holographic interferometry, see Fig. 18b.

1.10.2 Interference Phase Filtering

Although the noise included in the interference phase distributions generated by digital holographic interferometry is low, there are still some distortions. Therefore a smoothing of the interference phase by digital filtering should be performed to ensure a reliable phase demodulation.

A simple low pass filtering is not applicable to phase distributions modulo 2π. *I*t will cut the sharp edges at the 2π-jumps. It would complicate the demodulation process, which is based on finding these jumps. The Median filter does not preserve the edge

sharpness either. Fortunately there exists a filter that permits smoothing the phase while keeping the 2π-jumps at full amplitude. For this filter the interference phase must be divided into two orthogonal parts:

$$s(n,m) = sin\big(\Delta\varphi(n,m)\big),$$
$$c(n,m) = cos\big(\Delta\varphi(n,m)\big). \tag{1.68}$$

Although $\Delta\varphi(n,m)$ is modulated by the value 2π, $s(n,m)$ and $c(n,m)$ contain no 2π discontinuities. They both can be smoothed by a conventional low pass or median filter. After filtering of $s(n,m)$ and $c(n,m)$ the filtered version of $\Delta\varphi(n,m)$ is determined by:

$$\Delta\varphi_{filt}(n,m) = arctan\frac{s_{filt}(n,m)}{c_{filt}(n,m)}, \tag{1.69}$$

where index "*filt*" denotes the filtered version of the corresponding field. Now the low pass filtered interference phase $\Delta\varphi_{filt}(n,m)$ has kept the 2π jumps at full amplitude and simultaneously the noise is suppressed. The comparison of different filters is shown in Fig 19.

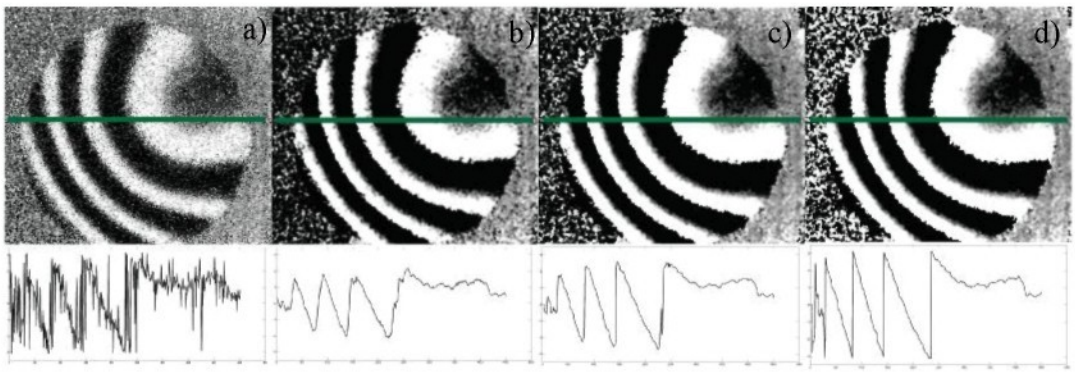

Fig. 19: Interference phase modulo 2π filtering and function values along the green line: a) non-filtered pattern; b) averaging convolution filter with kernel 7x7; c) median filter with kernel 7x7; d) sofisticated filter using median filter with kernel 7x7 in orthogonal directions.

1.10.3 Interference Phase Demodulation

From the formula (1.66) we can see that the interference phase distributions has only values between $-\pi$ and $+\pi$ (see Fig 20). Unfortunately in most practical applications a continuous interference phase distribution is expected. The process of

resolving the 2π discontinuities by adding a step function consisting only of 2π steps is called phase demodulation or phase unwrapping.

There exist two basic ways how to demodulate the interference phase. A more direct possibility of demodulation [10] (a one-dimensional interference phase distribution $\Delta\varphi(x)$) is done by checking the phase differences of neighboring pixels $\boldsymbol{\Delta\varphi(n+1) - \Delta\varphi(n)}$. If this difference is less than $-\pi$, an additional value 2π is added to $\boldsymbol{\Delta\varphi}$ from $\boldsymbol{n+1}$ onwards. If the difference is greater than $+\pi$ one more 2π is subtracted from $\boldsymbol{\Delta\varphi}$ starting at pixel $\boldsymbol{n+1}$. This easy algorithm strongly depends on numerical differentiation, which will amplify the influence of noise in the phase data. If a wrong difference occurs it can lead to an erroneous 2π term added, or subtracted, or missing a necessary 2π term operation. Such a phase error spreads up to all following pixels.

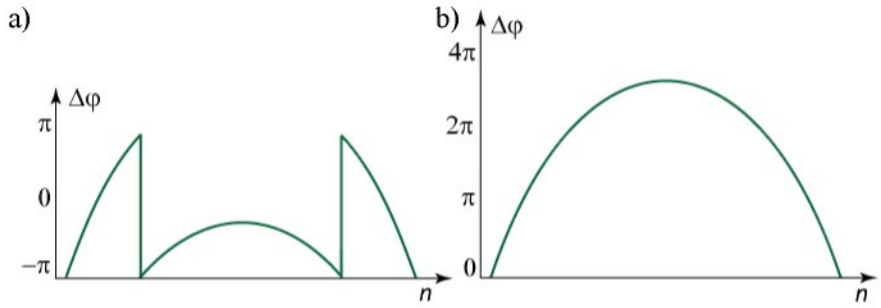

Fig. 20: Interference phase: a) modulo 2π; b) demodulated interference phase.

To avoid the difficulties with a possible spreading of erroneous phase, path independent demodulation procedures are recommended [13]. The following algorithm considers the distribution of interference phase modulo 2π as a graph, where the points are the nodes, and the arcs are the edges between neighboring points. For each edge a value $\boldsymbol{d_{2\pi}(\Delta\varphi_1, \Delta\varphi_2)}$ is defined by the phase values $\boldsymbol{\Delta\varphi_1, \Delta\varphi_2}$of the two nodes it connects:

$$d_{2\pi}(\Delta\varphi_1, \Delta\varphi_2) = min\{|\Delta\varphi_1 - \Delta\varphi_2|, |\Delta\varphi_1 - \Delta\varphi_2 + 2\pi|, |\Delta\varphi_1 - \Delta\varphi_2 - 2\pi|\}. \tag{1.70}$$

The demodulation now proceeds along the paths, where the values of $\boldsymbol{d_{2\pi}(\Delta\varphi_1, \Delta\varphi_2)}$ are minimal. Along these paths the probability of an erroneous demodulation is also minimal. Points with wrong phase are deselected this way as well as masked regions. Finally points, which possess an erroneous phase remain isolated in the final

interference phase distribution and can be easily recognized. Such points cannot be reached along any path.

1.11 Holographic Interferometry as a Measurement Tool

The calculation of the interference phase distribution is usually not the final goal of holographic interferometry. It is only an intermediate step in the determination of the physical quantity of interest like deformation in the case of diffusely reflecting opaque object, refractive index distribution or temperature field in the case of so called phase object (transparent objects, which do not affect the amplitude of an optical wavefront passing through, but only change the phase of this wavefront).

1.11.1 Deformation distribution measurement in relation to the interference phase

The measurements of the deformation of diffusely reflecting opaque object surfaces in holographic interferometry is based on the displacement of each surface point $\boldsymbol{R}$, which influences the optical path difference $\boldsymbol{\delta(R)}$. This is the difference between the paths from the source point S of the illuminating wavefront over the surface point $\boldsymbol{R}$ to the observation point B before and after change of the deformation state. The interference phase $\boldsymbol{\Delta\varphi(R)}$ is related to this path difference as:

$$\Delta\varphi(R) = \frac{2\pi}{\lambda}\delta(R) \tag{1.71}$$

Consider illumination point S and observation point B (see Fig. 21). If the object is deformed, the surface point $\boldsymbol{R}$ moves from position $\boldsymbol{R_1}$ to a new position $\boldsymbol{R_2}$.

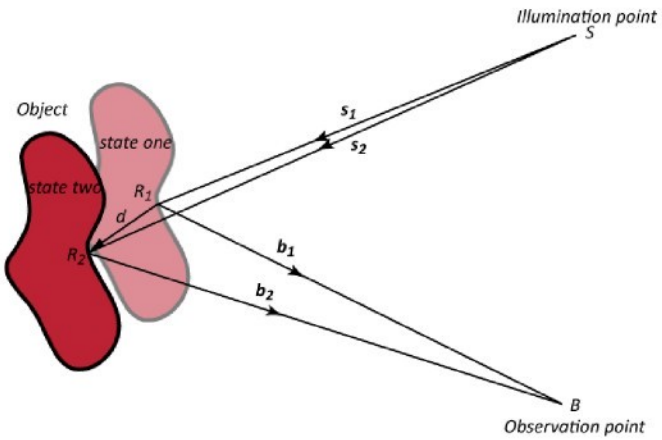

Fig. 21: Geometry of the holographic setup for measurement of deformation.

This change of position is called the displacement vector $\boldsymbol{d} = R_2 - R_1$.

Now we can determine the optical path difference as follows:

$$\delta(R) = |S\,R_1| + |R_1B| - |S\,R_2| - |R_2B| = \\ = \boldsymbol{s_1 S\,R_1 + b_1\,R_1B - s_2 S\,R_2 - b_2\,R_2B}, \tag{1.72}$$

where $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$ are unit vectors in the direction of illumination, $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ are unit vectors in the observation direction. Further $\boldsymbol{S\,R_i}$ and $\boldsymbol{R_iB}$ are the vectors from S to R_i or R_i to B.

From the Fig. 21 we can also determine the equation for the displacement vector

$$\boldsymbol{d}(R) = \boldsymbol{R_1B} - \boldsymbol{R_2B} = \boldsymbol{S\,R_2 - S\,R_1}. \tag{1.73}$$

Because deformations of an object are much smaller than the distances in holographic setup (the displacement $\boldsymbol{d}$ is in the micrometer range, whereas the distances $|S\,R_i|, |R_iB|$ are in range of meters), unit vectors in the direction of illumination s_1, s_2 can be considered as parallel. Unit vectors in the observation direction b_1, b_2 are oriented in the same way. Hence we can use $\boldsymbol{b = b_1 = b_2}$, $\boldsymbol{s = s_1 = s_2}$ and using (1.73) on (1.72) we get:

$$\delta(R) = \boldsymbol{d}(R)[\boldsymbol{b}(R) - \boldsymbol{s}(R)]. \tag{1.74}$$

Now we can define the so called sensitivity vector:

$$\boldsymbol{e}(R) = \frac{2\pi}{\lambda}[\boldsymbol{b}(R) - \boldsymbol{s}(R)], \tag{1.75}$$

The sensitivity vector is defined by the geometry of the holographic setup. With the use of (1.71), the equation for interference phase in relation on displacement vector $\boldsymbol{d}$ can be expressed as:

$$\Delta\varphi(R) = \boldsymbol{d}(R)\boldsymbol{e}(R). \tag{1.76}$$

It is evident we obtain the optimal geometry, in which the setup has maximal sensitivity, if the object is illuminated perpendicular to its surface.

It means that $\boldsymbol{b(P) = (1,0,0)}$, $\boldsymbol{s(P) = (-1,0,0)}$ and we can determine sensitivity vector for optimal holographic setup as follows:

$$\boldsymbol{e(R) = \frac{4\pi}{\lambda}}. \tag{1.77}$$

In such a case the deformation in point R can be calculated by the following equation:

$$\boldsymbol{d(R) = \frac{\lambda}{4\pi}\Delta\varphi(R)}. \tag{1.78}$$

1.11.2 Refractive Index measurement

The refractive index distribution measurement of phase objects is also based on the comparison of the interference phase between two different states.

Consider the double exposure method, when the first recording is done with the refractive index distribution $\boldsymbol{n_1(x,y,z)}$, and the second one with the different refractive index distribution $\boldsymbol{n_2(x,y,z)}$. If the recording wave impinging in the z direction onto the phase object surface (plane with coordinates x, y) (see Fig. 22) then the path difference $\boldsymbol{\delta(x,y)}$ is:

$$\delta(x,y) = \int \Delta n(x,y,z)dz, \tag{1.79}$$

where $\boldsymbol{\Delta n(x,y,z) = n_2(x,y,z) - n_1(x,y,z)}$ and with the use of formula (1.71) the interference phase can be calculated as:

$$\Delta\varphi(x,y) = \frac{2\pi}{\lambda}\int \Delta n(x,y,z)dz. \tag{1.80}$$

The simplest case is an object with refractive index varying in only one direction (direction y for example). Consider the object with the length *l* in the *z*-direction and a constant refractive index along *l*. Then the interference phase is:

$$\Delta\varphi(y) = \frac{2\pi}{\lambda}(n(y) - n_1)l. \tag{1.81}$$

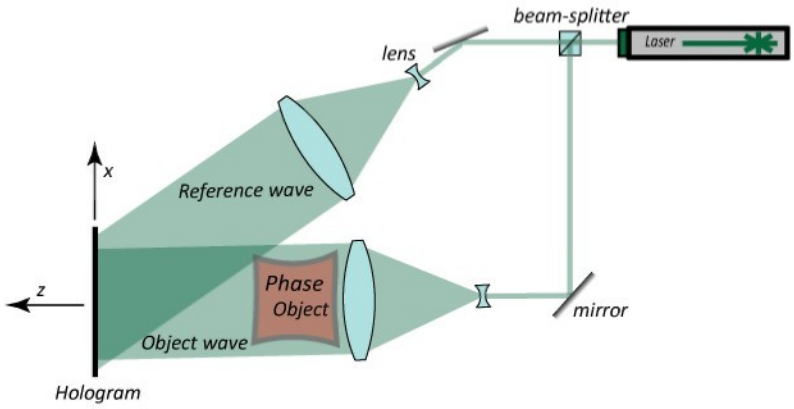

Fig. 22: Holographic setup for measurement of the phase object refractive index.

2 Application for Deformation Measurement using Digital Holographic Interferometry

The first chapter has already shown the digital holographic interferometry includes many operations and calculations which are interconnected. For this purpose an application called DEMETER (DEformation MEasurement by inTERferometry), which effectively solves all processing steps one by one, was developed. Beyond the data processing and necessary numerical calculations the application also uses many graphical tools for data visualization and analysis. User interface is intuitive and leads the user from the beginning (loading of a digital hologram) to the end (a deformation field display). DEMETER was developed using programming language MATLAB and of its all functions and possibilities will be described in this chapter.

2.1 MATLAB Environment

MATLAB (from MATrix LABoratory) is a numerical computing environment and computing language of the fourth generation. MATLAB allows matrix manipulations, 2-D and 3-D plotting of functions and data, algorithms implementation, creating user interfaces, and interfacing with programs written in other languages, including C, C++, and Fortran. The MATLAB family contains more than 90 modules, which can facilitate the data processing. The main structure of MATLAB is displayed in Fig. 23. For the development of our application we have used mainly the Image Processing Toolbox and The Graphical User Interface.

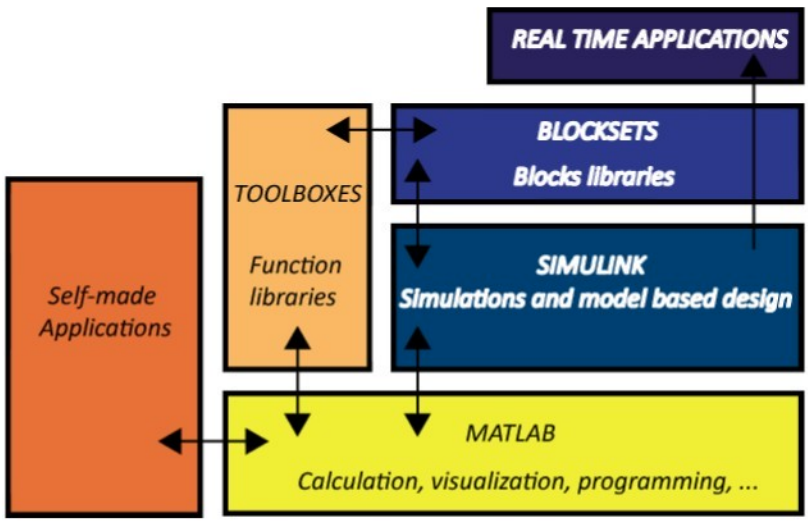

Fig. 23: The structure of MATLAB and SIMULINK

2.2 Structure of DEMETER

The whole application is divided into four logical panel tabs. In the first panel tab, called Hologram, the basic operations with digital hologram like loading, cropping or histogram displaying are possible. These are necessary for the next step, reconstruction, which is located in another panel tab. Here the basic parameters of the

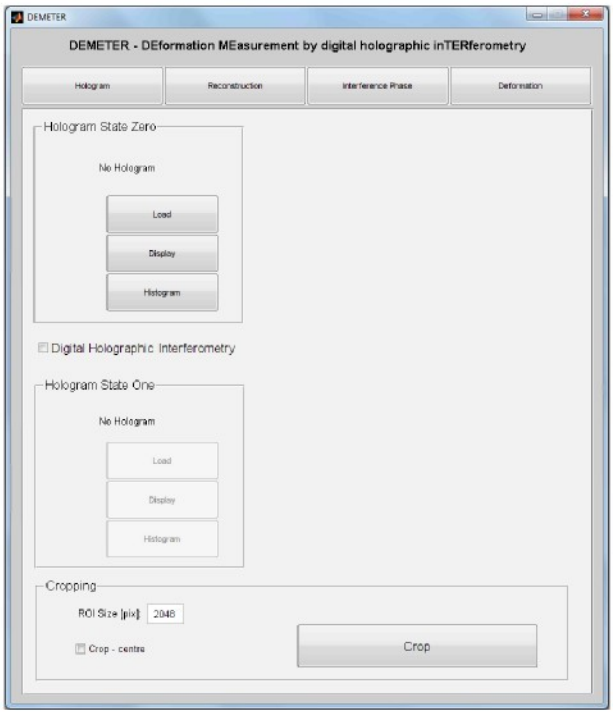

Fig. 24: DEMETER – opening screen.

holographic setup are set up. The numerical algorithms to suppress the d.c. term and the twin image can also be used. The third tab - interference phase - is focused on interference phase calculation, processing and demodulation. Finally the last tab panel, Deformation, is aimed at the deformation field calculation, results analysis and visualization. In the following text individual tab panels and their functions will be described in detail.

2.2.1 Hologram

At first the digital hologram must be loaded from a file. The digital hologram is usually stored in computer memory as an image in typical image format as .jpeg, .bmp, .png, .tiff, .gif or .raw. The button "Load" displays a modal dialog box that lists files in the current folder and enables us to select or enter the name of a file. The current folder

path is saved, which brings more comfort for repeated using of "Load" button callback, because we do not have to search the current folder again.

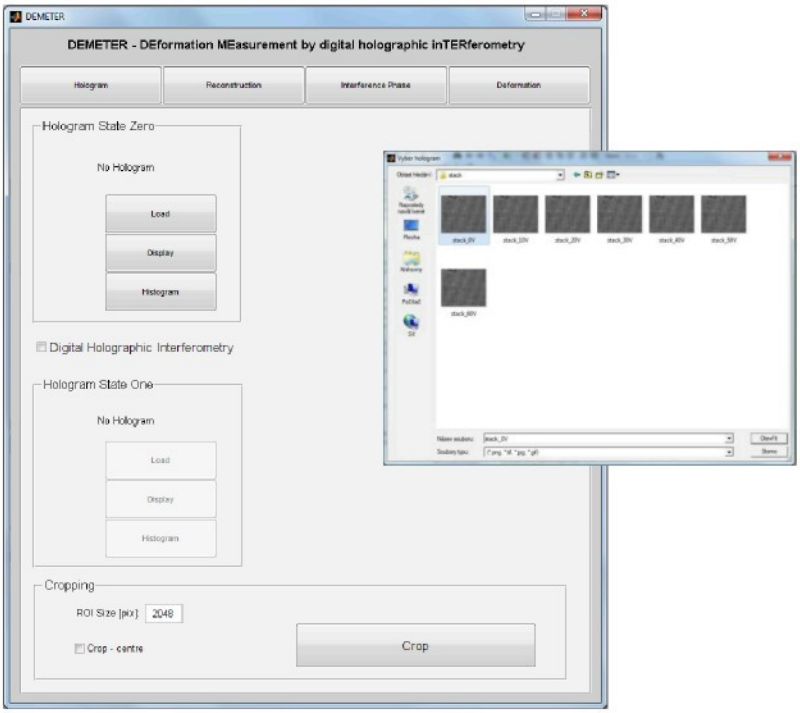

Fig. 25: Tab panel Hologram with modal dialog box.

Now the digital hologram is ready for processing and can be displayed in its current size using the "Display" button. Statistical analysis of the digital hologram histogram can be displayed by pressing the "Histogram" button. It shows the distribution of data by counting the number of data values within a range between 0 - 255 values and displaying each range as a rectangular column, see Fig. 56. The height of a bin represents the number of values that fall within each range. It helps us to decide if the shutter of the digital camera used for digital holographic recording was set optimally.

Digital holography interferometry is based on a comparison of two different wave fields, which correspond to a certain state of object. Therefore the checkbox "Digital Holographic Interferometry" makes available the buttons "Load", "Display" and "Histogram" for the second digital hologram.

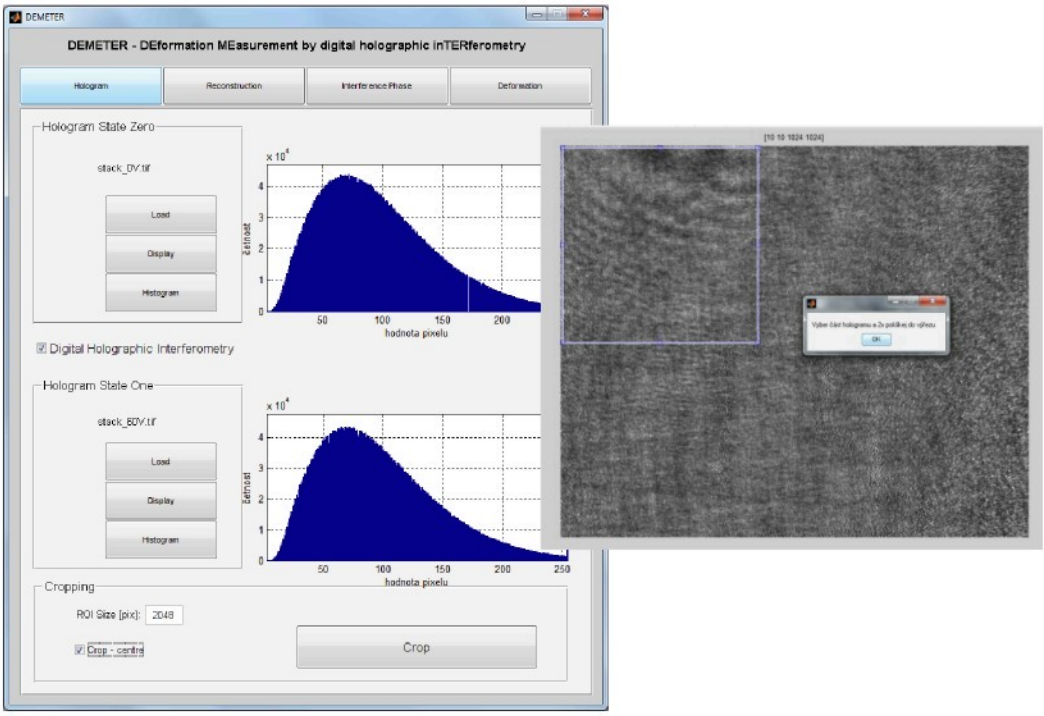

Fig. 26: Tab panel hologram: display of digital hologram histograms and crop tool.

In reconstruction of wave fields the Fast Fourier Algorithm, which needs $\mathbf{2^n}$ values, is used. It is recommended to crop the digital holograms using $\mathbf{2^n}$ squared region. The button "Crop" serves to this effect. If the check box "Crop - centre" is selected, the cropped image is the squared array placed in the centre of initial image. Its size can be set by "edit text" "ROI Size". If the check box "Crop - centre" is not selected, the Crop Image Tool with region of interest (its size is also given by "edit text" "ROI Size" value) is displayed. The region is a moveable, resizable rectangle which we can position interactively using the mouse. Then the cropping is done by double clicking inside the region. The centre cropping in the majority of cases is sufficient for processing of a digital hologram.

2.2.2 Reconstruction

The reconstructed wave field results from the formula (1.56). To simplify the calculation of the reconstruction equation, the inverse discrete Fourier transform can be used [2]:

$$f(n,m) = \sum_{k=1}^{N}\sum_{l=1}^{M} F(k,l)e^{j2\pi\left(\frac{kn}{N}+\frac{lm}{M}\right)}, \qquad (2.0)$$

then the reconstruction equation can be presented as:

$$g(n\Delta x, m\Delta y) = \frac{e^{-j\frac{2\pi}{\lambda}d}}{\lambda d}\mathcal{F}^{-1}\left\{h(k\Delta\xi, l\Delta\eta)r^*(k\Delta\xi, l\Delta\eta)e^{\frac{j\pi}{\lambda d}((k\Delta\xi)^2+(l\Delta\eta)^2)}\right\}, \tag{2.1}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform, d is CCD – object distance, λ is the wavelength of used laser and $\Delta\xi \times \Delta\eta$ is the CCD pixel size in orthogonal directions.

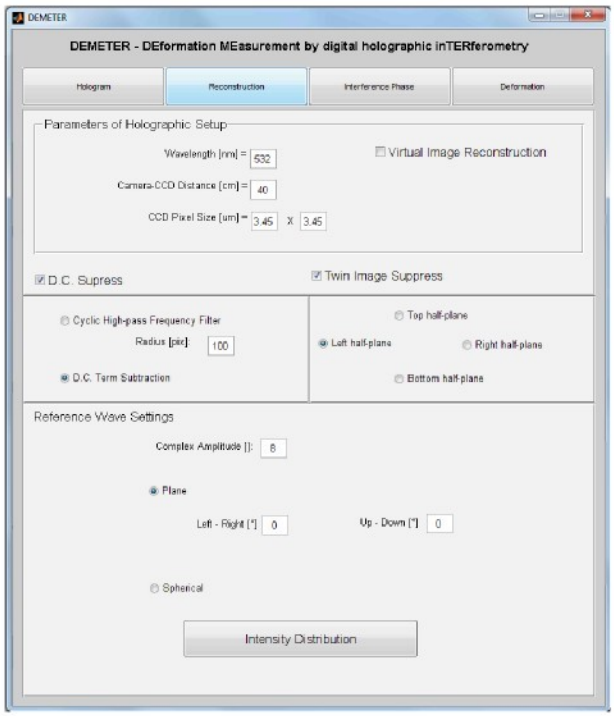

Fig. 27: The tab panel reconstruction.

Utilization of the Fast Fourier Transform algorithm (FFT) leads to the acceleration of the computing process and to the reduction of computing time. Variables of reconstruction formula (2.1) are set by corresponding "edit text", see Fig. 27.

As was mentioned in the chapter about digital holography reconstruction, the reconstructed wave also the twin image and the d.c. term. It is undesirable in case if the twin image or the d.c. term overlaps the real image.

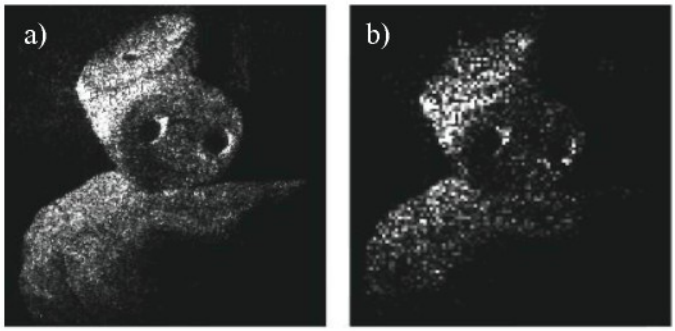

Fig. 28: The resolution of intensity distribution: a)reconstructed from digital hologram of size 1024x1024 pixels; b) 256x256 pixels

2.2.2.1 D.C. Term and Twin Image Suppression

The d.c. term is the zero order diffraction image and therefore it is represented by low spatial frequencies of the digital hologram. Such low frequencies can be suppressed by high-pass filtering of the digital hologram. There exist many ways how to do it. In DEMETER two different ways with comparable results are possible. If the "Cyclic High-pass Frequency Filter" radiobutton is selected, low frequencies in the spectrum of the digital hologram are set to zero. The region, where the frequencies are suppressed, is the circle with centre in the coordinate origin of the digital hologram spectrum. The radius of the circle is adjustable using "edit text" "Radius". Otherwise we can select the radiobutton "D.C. Term Subtraction". There the average intensity h_{ave} of the digital hologram h is calculated as follows:

$$h_{ave} = \frac{1}{NM}\sum_{k=1}^{N}\sum_{l=1}^{M} h(k,l), \tag{2.2}$$

where $N \times M$ is the size of the digital hologram. N is the number of rows and M is the number of columns. This average intensity is now subtracted from each stored hologram intensity value, yielding the modified digital hologram h_m

$$h_m(k,l) = h(k,l) - h_{ave}. \tag{2.3}$$

By this subtraction of the d.c. term we have only suppressed the spatial frequency at the point [0,0]. Therefore other high-pass filters suppressing the smallest spatial frequencies can be employed to achieve desirable effects.

This high-pass filter is defined by the subtraction of the averages over each 3 × 3 pixel neighborhood from the original digital hologram:

$$h_F(k,l) = h(k,l) - \\ -\frac{1}{9}\begin{bmatrix} h(k-1,l-1)+h(k-1,l)+h(k-1,l+1)+ \\ +h(k,l-1)\ +h(k,l)+h(k,l+1)+ \\ +h(k+1,l-1)+h(k+1,l)+\ h(k+1,l+1) \end{bmatrix}, \quad (2.4)$$

where $\boldsymbol{h_F}$ is the filtered hologram and $\boldsymbol{h}$ is the initial digital hologram or the modified digital hologram $\boldsymbol{h_m}$.

After the digital hologram processing we can easily use the filtered hologram $\boldsymbol{h_F}$ instead of initial hologram $\boldsymbol{h}$ in the reconstruction formula (2.1).

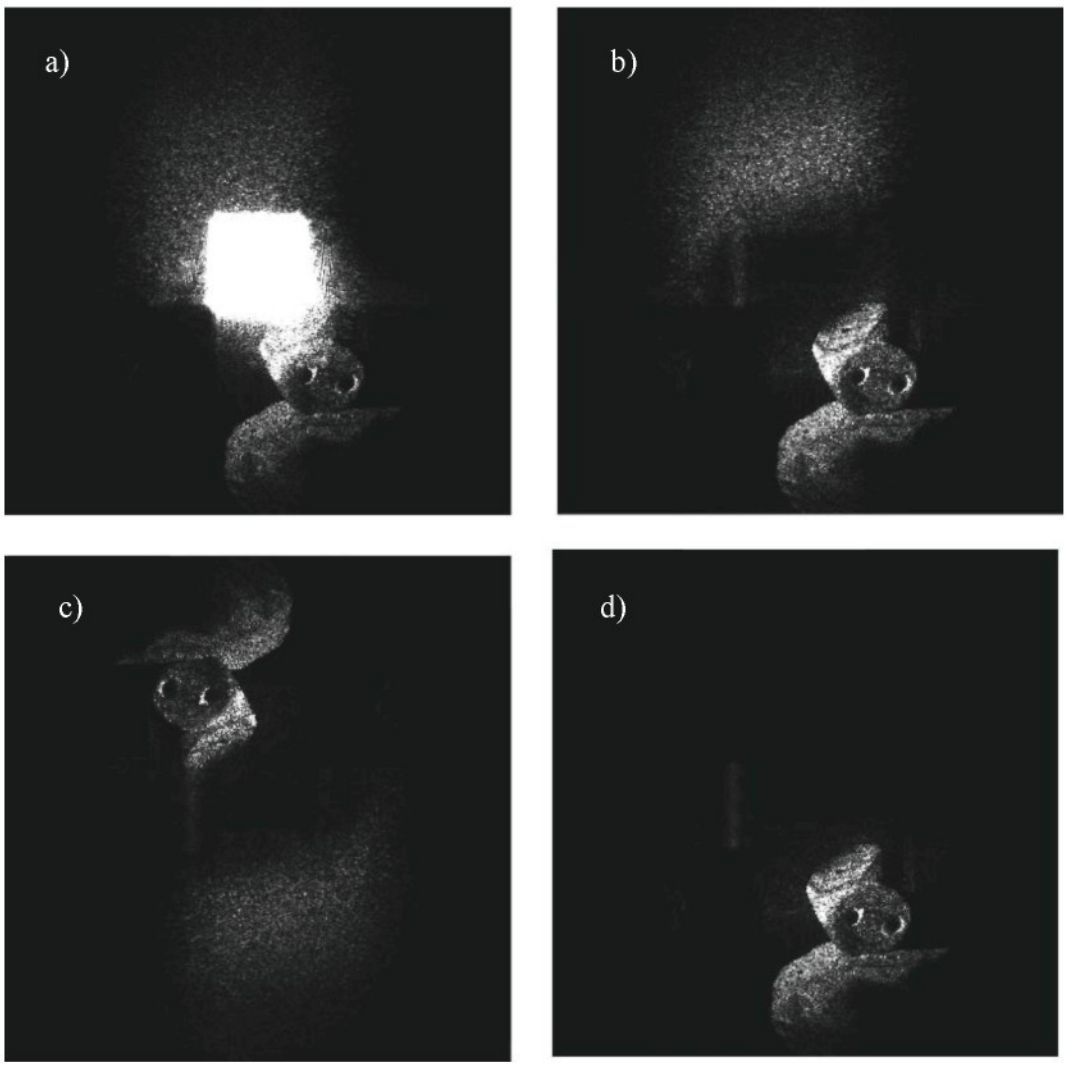

Fig. 29: The reconstructed intensity distribution: a) without any adjustments; b) after d.c. term suppression; c) twin image reconstruction with d.c. term suppression; d) intensity distribution after d.c. term suppression and twin image suppression.

Another component of the reconstructed wave field is the twin image represented by the conjugated part of the complex field. The twin image is not focused in the distance d (the observer plane) but it is focused in distance $-d$ (the object plane). By selecting the checkbox "Virtual Image", the twin image will be focused, see Fig. 29c.

Usually the presence of the twin image is undesirable, so we want to suppress it. If the object is placed completely outside the optical axis, so that the twin image in the reconstructed field will not overlap the real one, then the amplitude spectrum of the digital hologram will consist of two symmetrically placed partial spectra. One of these contains the frequencies leading to the real image. The other one contains those responsible for the virtual image. So we can set the frequencies of the virtual image to zero and in the reconstructed frame only the real image remains. In DEMETER we can select which half-plane, where the twin image is located, we want to suppressed, see Fig. 29d.

2.2.2.2 *Reference Wave*

During the reconstruction, the digital holograms are multiplied with the numerical models of the reference waves before the diffracted field in the image plane is calculated. The most frequent reference waves used in praxis are the plane wave and the spherical wave. This choice is also possible in DEMETER. The most commonly used in digital holography is the plane reference wave. In general the reference wave is not normally impinging onto the CCD and can be described as:

$$r^*(k,l) = U_r e^{\frac{j\pi}{\lambda d}(k\sin\theta_1\Delta\xi + l\sin\theta_2\Delta\eta)}, \tag{2.5}$$

where U_r is the complex envelope, θ_1 is the angle to normal in horizontal direction and θ_2 is the angle to normal in vertical direction. These parameters are adjustable using corresponding edit texts.

The second case to be considered in DEMETER is a spherical wave. The numerical expression for the Fresnel approximation of the spherical reference wave is:

$$r^*(k,l) = U_r e^{\frac{j2\pi d}{\lambda}} e^{\frac{j\pi}{\lambda d}((k\Delta\xi)^2 + (l\Delta\eta)^2)}. \tag{2.6}$$

The effects of using different reference waves in the digital holographic process are shown in Fig. 30.

Fig. 30: The effect of reference wave direction: a) normally impigning; b) $\theta_1 = 0.5°$ *in horizontal direction c)* $\theta_1 = 0.5°$ *in horizontal direction and* $\theta_2 = -0.2$ *in vertical direction.*

2.2.3 Interference Phase

In this section of application we want to calculate and process the interference phase modulo 2π. The result should be demodulated interference phase without 2π jumps.

After the reconstruction process of two digital holograms we have two complex fields corresponding to certain state of the object:

$$g_1(n,m) = g_{r1}(n,m) + jg_{i1}(n,m), \qquad (2.7)$$
$$g_2(n,m) = g_{r2}(n,m) + jg_{i2}(n,m).$$

To reduce the noise, components of these complex fields can easily be filtered by low-pass filter:

$$g_r^{filt}(n,m) = g_r(n,m) * K, \qquad (2.8)$$
$$g_i^{filt}(n,m) = g_i(n,m) * K,$$

where $K = \frac{1}{n^2}\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$ is the matrix with n equal to the number of columns or number of rows and symbol $*$ is the convolution operator. The variable n can be set by edit text "Squared Kernel Size", see Fig. 31. If user sets the edit text "Squared Kernel Size" to one, no filtering process is applied.

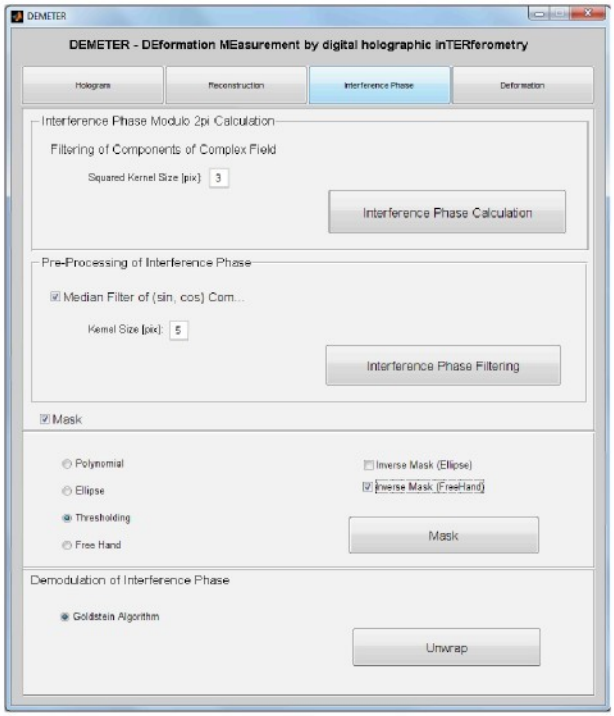

Fig. 31: The tab panel Interference Phase.

The filtered version of complex fields is:

$$g_1{}^{filt}(n,m) = g_{r1}^{filt}(n,m) + jg_{i1}^{filt}(n,m),$$
$$g_2{}^{filt}(n,m) = g_{r2}^{filt}(n,m) + jg_{i2}^{filt}(n,m). \tag{2.9}$$

The comparison of this effect of the filter is shown in Fig. 33. Now the interference phase modulo 2π can be calculated using (1.66), where we can use the filtered components of complex fields $g_1{}^{filt}$, $g_2{}^{filt}$ as well as the initial components g_1, g_2. The interference phase is calculated for all components of reconstructed field, but we are only interested in the part of the field, where the object is located. Other pixels are redundant and can be cropped. The cropping lowers the computing time during further operations with the field. For this purpose a crop tool can be utilized. This tool is based on the same principle as for the digital hologram cropping.

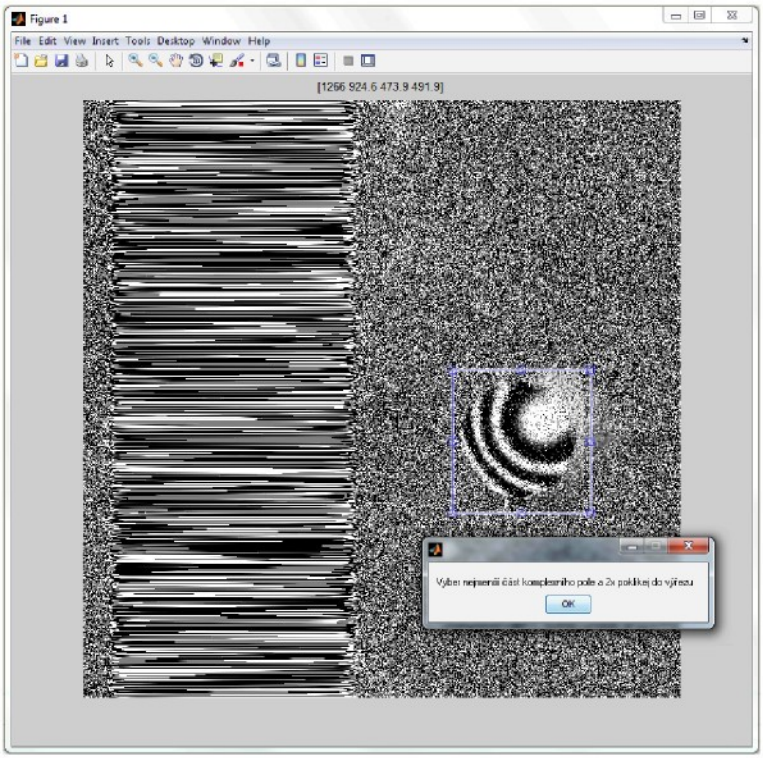

Fig. 32: The crop tool.

After clicking the "Interference Phase calculation" button, the calculated interference phase distribution with the cropping rectangle is displayed, see Fig. 32. There we can crop the redundant pixels and after that the cropped or filtered interference phase is displayed, see Fig. 33.

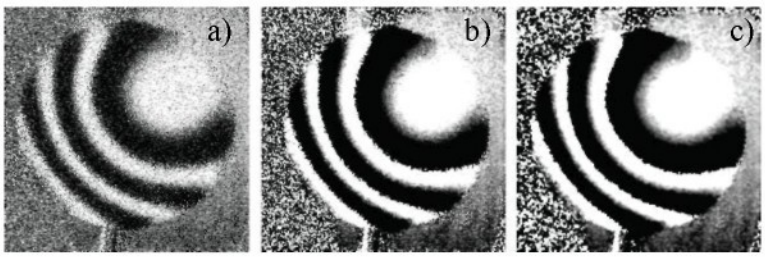

Fig. 33: The cropping effect and filtering of components of the complex field: a) non-filtered version; b) filtered field by averaging filter with matrix of size 3x3pixels; c)field filtered by averaging filter with matrix of size 5x5 pixels.

In DEMETER there is another possibility how to pre-process the interference phase. The sophisticated filtering is strongly recommended before the demodulation process, as was described in detail in chapter "Interference Phase Filtering". To do that, the checkbox "Median Filter of (sin, cos) Components" must be selected. There the interference phase is divided into two orthogonal components as can be seen from formula (1.68), and these components are then filtered. To filter the

components $s(n,m)$, $c(n,m)$ is used median filter. Median filter replaces the value of each pixel by the median of values of all pixels in the neighborhood. The size of the neighborhood can be set in edit text "Kernel Size". Consequently the filtered version of interference phase is given by (1.69). In DEMETER the whole process of sophisticated filtering of interference phase is done by clicking the "Interference Phase Filtering" button and the desired filtering effect is shown in Fig. 34.

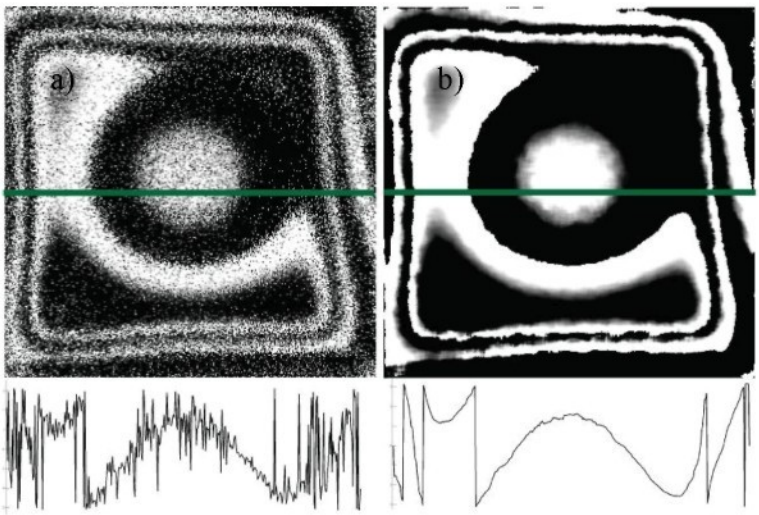

Fig. 34: The effect of advanced interference phase filtering: a) non-filtered version and its function values along the green line (in bottom part); b) the filtered version and its function values along the green line.

Although we have already discarded some redundant pixel by phase field cropping, usually there are still some left. These pixels are given by the object shape in the rectangle region, see Fig. 32, where the round object is placed in rectangle region. To eliminate these redundant pixels and to obtain an image without noisy background the masking can be used. First of all the mask is created. It is a binary image with the same size as the interference field. The pixels of the mask, corresponding to the redundant pixels of the interference phase, are set to zero, other pixels are set to one. Then the masked interference phase can be calculated from:

$$\Delta\varphi_{mask}(n,m) = \Delta\varphi(n,m) \cdot Mask(n,m), \tag{2.10}$$

where $Mask$ is the mask field and symbol $\cdot$ indicates pointwise product.

In DEMETER there are four possibilities how to create the mask. If the object has a clearly defined geometrical shape, the mask type "Ellipse" for oval-shaped objects or the mask type "Polynomial" for objects with straight edges can be used.

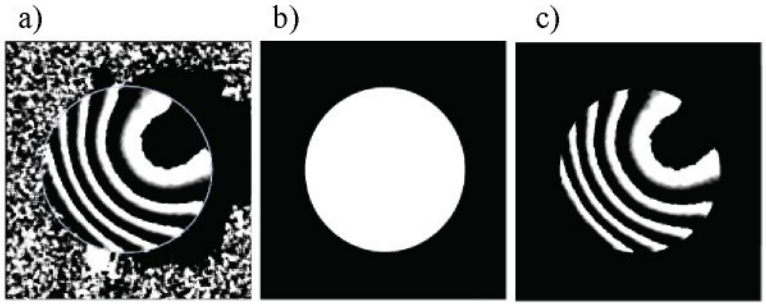

Fig. 35: Masking process: a) initial image; b) "ellipse" mask; c) finial image after applying the mask.

The "Thresholding" or "FreeHand" masks serve for more complicated shapes of objects.

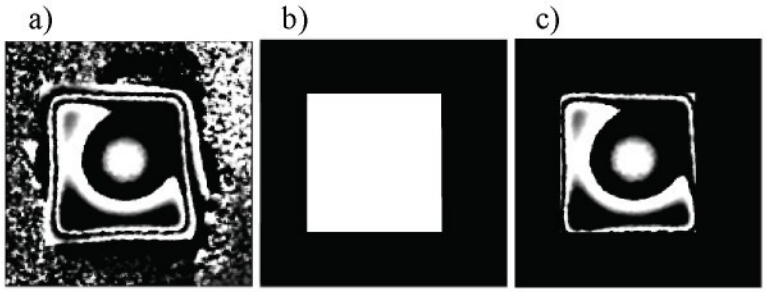

Fig. 36: Masking process: a) initial image; b) "polynomial" mask; c) finial image after the mask is applied.

If user selects the "Thresholding" mask, the threshtool is displayed, see Fig 37. There we can continuously change the threshold value and find its optimized value, when the redundant pixels are set to zero. While changing the thresh, all pixels in the interference phase with luminance greater than the threshold value are replaced with the value one (white) and all other pixel values are replaced with the value zero (black), see Fig. 38.

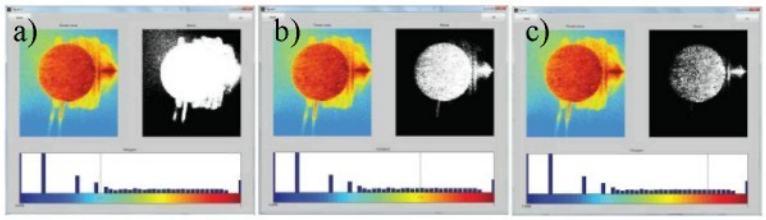

Fig. 37: Thresh tool: how does the change of threshold value inside the histogram plot affects the binary mask.

The last kind of mask is the "FreeHand" mask. It is very versatile, because we can distinguish any shape we want by an interactive mouse dragging, see Fig. 39.

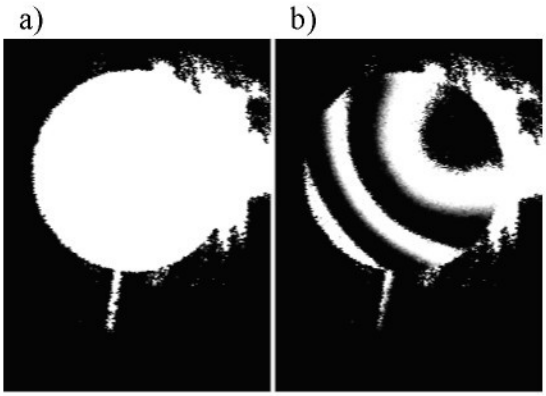

Fig. 38: Masking process: a) "threshtool" mask; c) finial image after the mask is applied.

Moreover the possibility of applying the inverse mask is available. Inverse mask can eliminate redundant parts of the object, e.g. a screw, see Fig. 39. We can choose "ellipse" or "FreeHand" inverse mask.

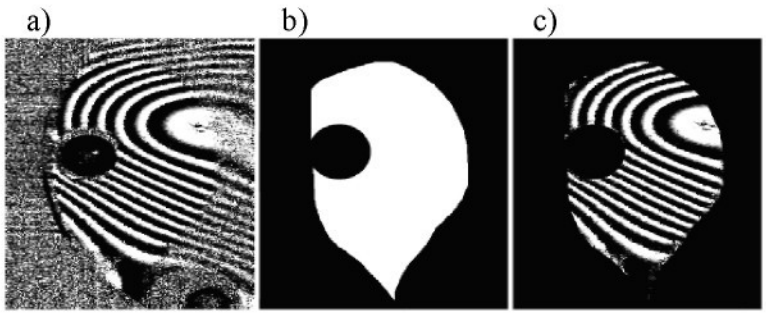

Fig. 39: Masking process: a) initial image; b) "FreeHand" mask for the periphery detection and „ellipse" inverse mask inside the region for the screw elimination; c) finial image after the mask was applied.

Now the interference phase is prepared enough for demodulation process. The goal of demodulation is to unwrap the modulated interference phase and obtain continuous interference phase without 2π jumps. Existing phase demodulating techniques start from the fact that it is possible to estimate the neighboring pixel differences from the modulated phase, when these differences are less than π. Now, the demodulated phase can be reconstructed by adding a constant.

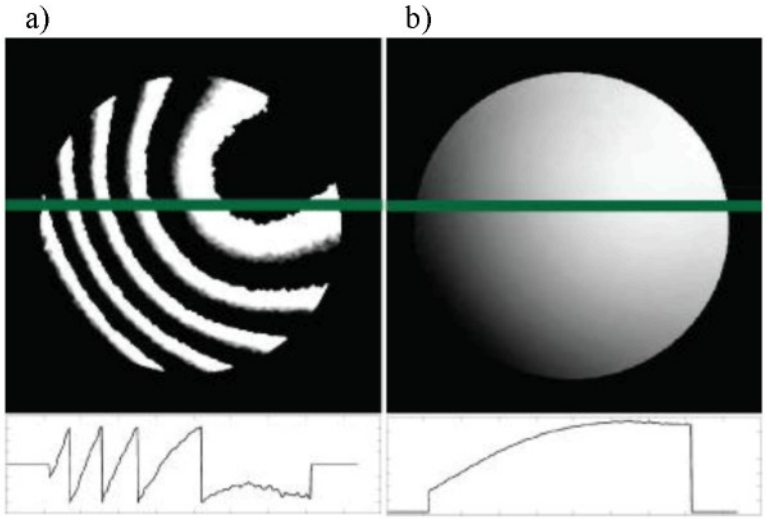

Fig. 40: Demodulation process: a) initial masked interference phase distribution modulo 2π *and function values along the green line in the bottom part; b) interference phase distribution after demodulation and function values along the green line.*

The demodulating methods differ in the way, how they overcome the difficulty of errors being present in the phase field. The errors cause the estimated modulated phase differences to be inconsistent and their demodulated value depends on the integration path.

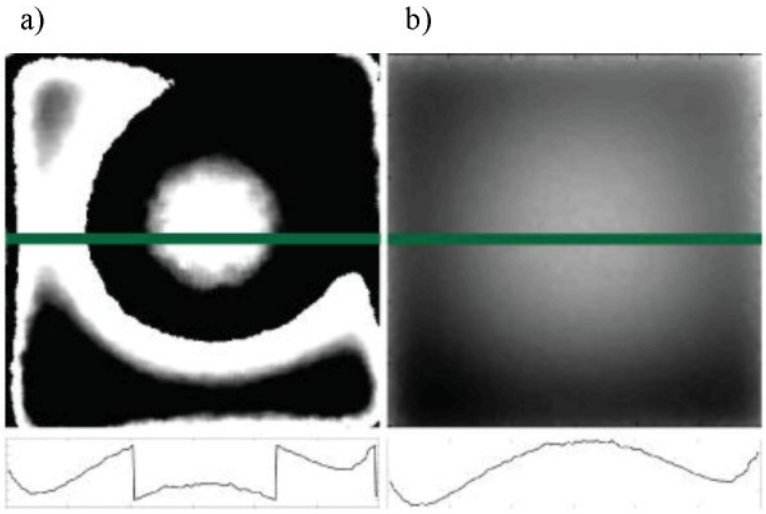

Fig. 41: Demodulation process: a) initial masked interference phase modulo 2π *distribution and function values along the green line are shown in the bottom part of the figure; b) interference phase distribution after demodulation and function values along the green line.*

For demodulation process in DEMETER, "Goldstein algorithm" is used. This algorithm demodulates the modulated phase by integrating the estimated neighboring

pixel differences avoiding the regions, where these estimated differences are inconsistent.

The theoretical basis of the method was described in chapter „Interference Phase Demodulation“. The main structure of this algorithm was taken from website [4] and later it was upgraded for our own purposes. Demodulation is done by clicking the “Unwrap Phase” button. The effect of demodulation process is shown in Fig. 40 and Fig. 41.

2.2.4 Deformation

In this tab panel functions for calculations, processing and displaying of resulting deformation distribution are available.

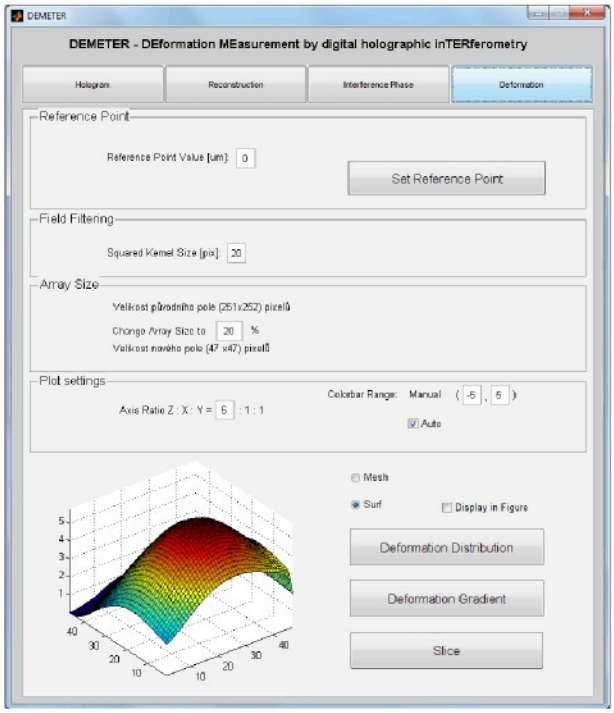

Fig. 42: The tab panel Deformation.

At first we can set the reference point, which belongs to the deformation distribution corresponding to known value of the deformation. In DEMETER we can set a value anywhere in the deformation field, but usually we set a zero reference value in point, where the value of deformation is also zero (e.g. near a rigid frame).

According to (1.78) the continuous demodulated phase has direct relation to the deformation, so it can be easily calculated. It is done after clicking the “Deformation distribution” button and then the resulting deformation distribution is displayed. If the

radio button "Mesh" is selected, a colored wireframe mesh is drawn, where the color is proportional to the surface deformation is drawn. On the other hand, if the radio button "Surf" is selected a three-dimensional shaded surface is created. Both plots are partly displayed in the user interface and they can also be displayed in their own figure if the checkbox "Display in Figure" is checked, see Fig. 42, Fig. 43, Fig. 44.

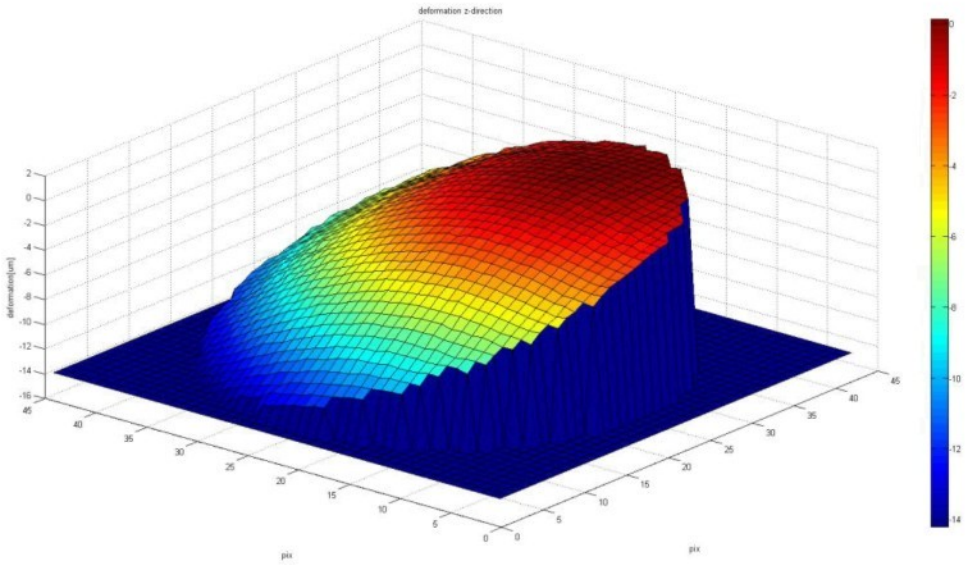

Fig. 43: The surf plot of the deformed surface.

Because the calculated deformation field is not absolutely without noise, the averaging convolution filter can be applied as:

$$d_{filt}(n,m) = d(n,m) * K,$$

where $K = \frac{1}{n^2}\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$ is the matrix with n equal to number of columns and number of rows and the symbol $*$ is the convolution operator. The matrix size n can be set by edit text „Squared Kernel Size"and filtering is done by „Deformation distribution" button click. So we obtain the smoothed version of the deformation distribution. If user sets the value in the edit text „Squared Kernel Size" to one, no filter will be applied.

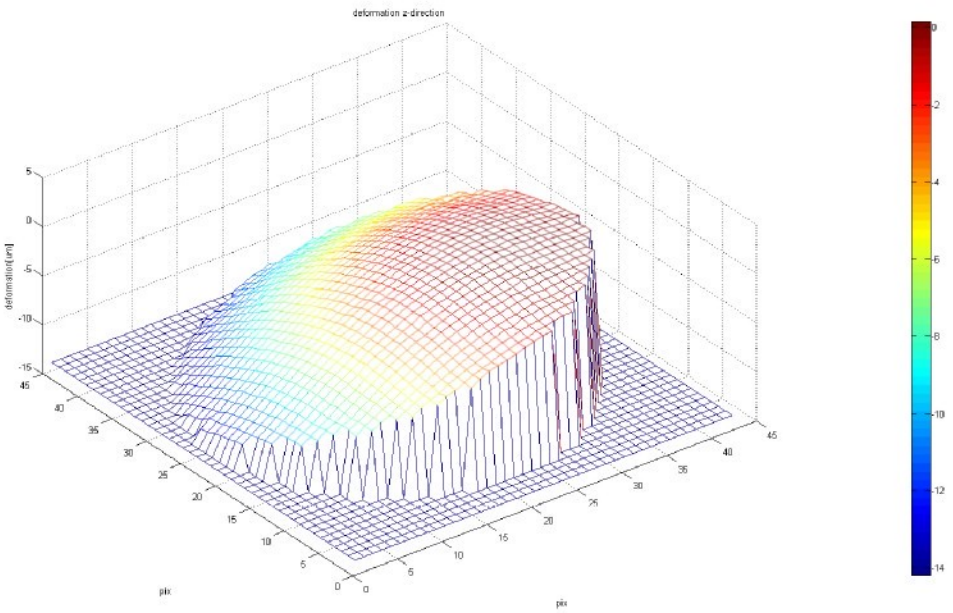

Fig. 44: The mesh plot of the deformed surface.

The user can also interpolate or decimate the deformation field. The decimation appears advantageous in case of large fields. There the pixels density is unnecessarily high. The size of the deformation field can be changed by edit text "Change array size". The effect of decimation and interpolation can be seen in Fig. 45a,b.

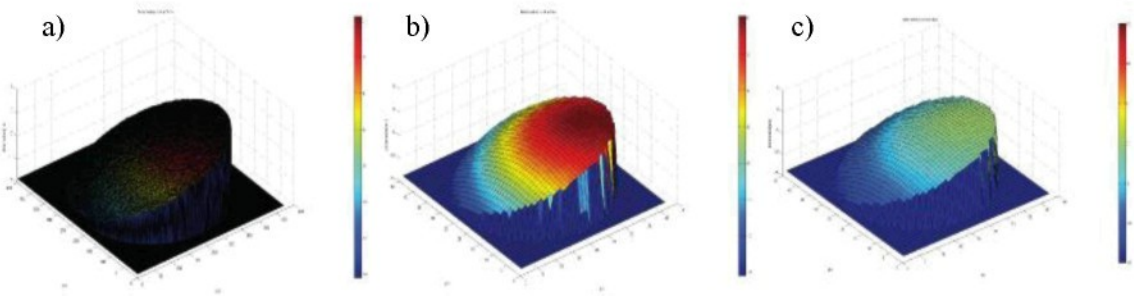

Fig. 45: The effect of decimation and colorbar scaling: a) initial 100% field of deformation; b) 10% decimated field of deformation with colorbar autoscale; c) 10% decimated field of deformation with colorbar manually scaled from -14 to 14μm;

In DEMETER the axis ratio can be adjusted. This can be used to obtain better 3-D plot of deformation field. The colorbar scale can be adjusted manually or automatically according to minimal and maximal value of the deformation field, see Fig.45b,c.

When measuring the deformation we can also be interested in the direction of the greatest rate of increase of the scalar field – the gradient:

$$\nabla d = \frac{\partial d}{\partial n}\boldsymbol{i} + \frac{\partial d}{\partial m}\boldsymbol{j},$$

where d is the deformation distribution, $\boldsymbol{n}, \boldsymbol{m}$ are field coordinates and $\boldsymbol{i}, \boldsymbol{j}$ are unit vectors. The button "Gradient of Deformation" serves for this purpose. It plots gradient vectors as arrows for each corresponding pair of elements in $\boldsymbol{n}$ and $\boldsymbol{m}$, see Fig. 46. Simultaneously the level curve (a curve along which the function has a constant value) is plotted as can be seen in Fig. 46.

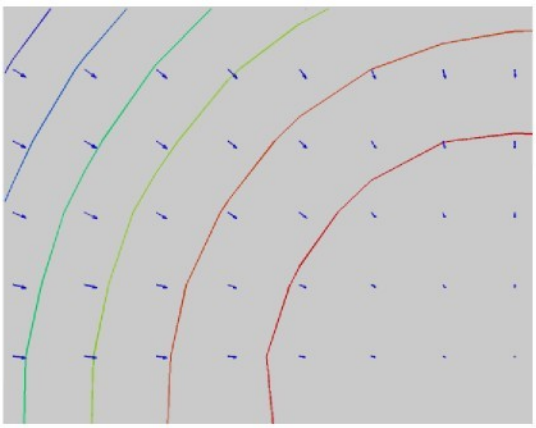

Fig. 46: The gradient field with level lines.

Another useful function for results interpretation is "slice", which permits us to plot function values along a certain line.

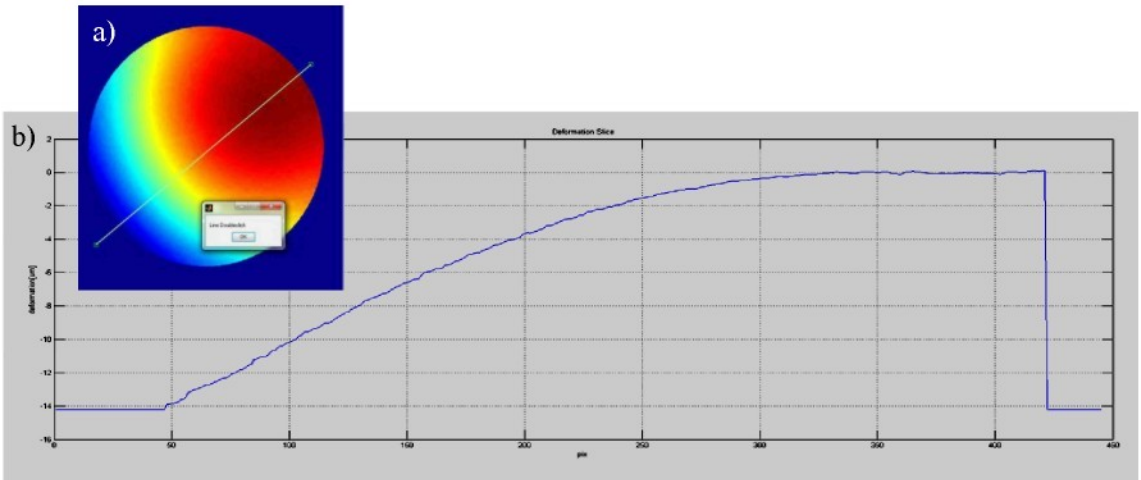

Fig. 47: The slicetool: a) moveable, resizable and draggable line in the deformation field image; b) Function values along the chosen line.

After the "SliceTool" button click the image of deformation distribution is displayed. There we can draw any line by a drag mouse. The line is resizable and draggable and finally the values of deformation are plotted along the line, see Fig. 47.

3 Experiment

In this chapter of my thesis we will demonstrate the applications of DEMETER for measurement of deformation distribution.

3.1 The deformation distribution measurement of piezoelectric actuators

Piezoelectric actuators are transducers that convert electrical energy into a mechanical displacement using an inverse piezoelectric effect, which is the ability of some materials to product mechanical stress in response to applied electric field.

We used two different piezoelectric actuators: round and squared. We powered them by different DC voltage values to obtain different deformation distributions - states of object - and for each of these states we have recorded a digital hologram. The digital hologram recorded for non-supplied piezoelectric actuator was used as a reference (not deformed) state. Then the digital holograms were processed by DEMETER step by step and resulting deformation distribution field were displayed.

3.1.1 Square piezoelectric actuator

The DC voltage, which was applied to generate deformation states, was in range 0 – 60V. Then we changed the polarity of the DC voltage and repeated the measurement. The results are displayed as a surf 3D plot, where the deformation distribution is displayed for the whole surface of object. Moreover, the values of deformation along the white line in the surf plot are plotted in the bottom part of the figure, see Fig. 48. In following the results for single values of DC voltage are introduced:

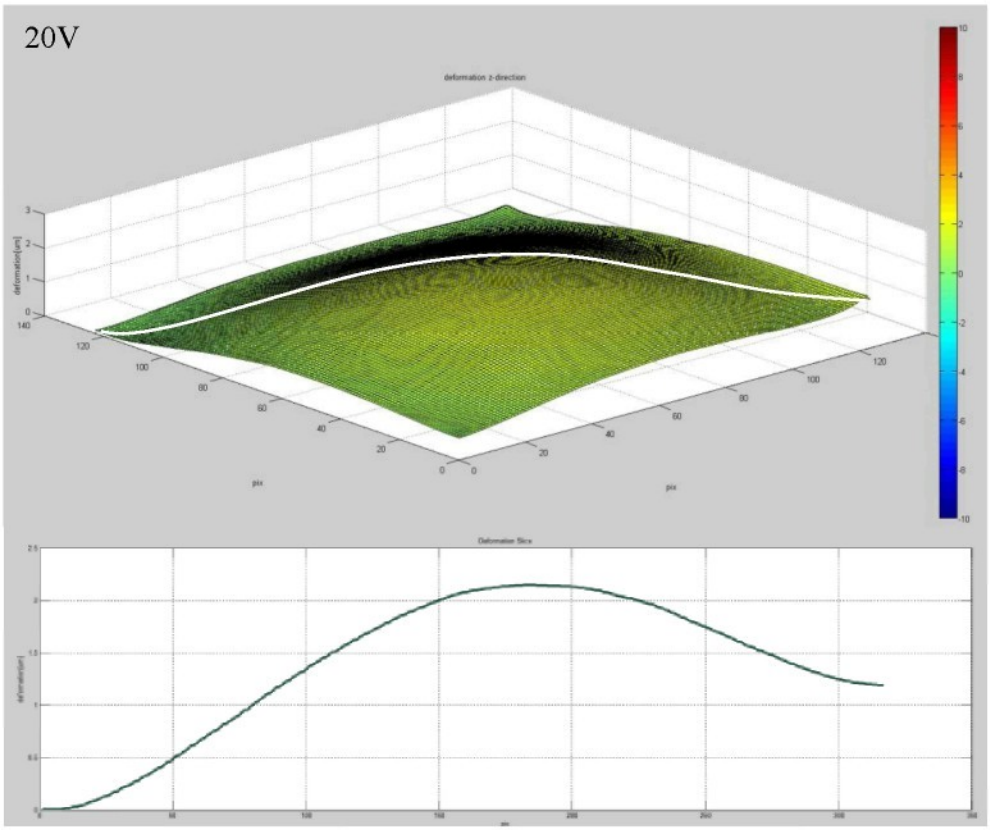

Fig. 48: Deformation distribution for 20V DC voltage.

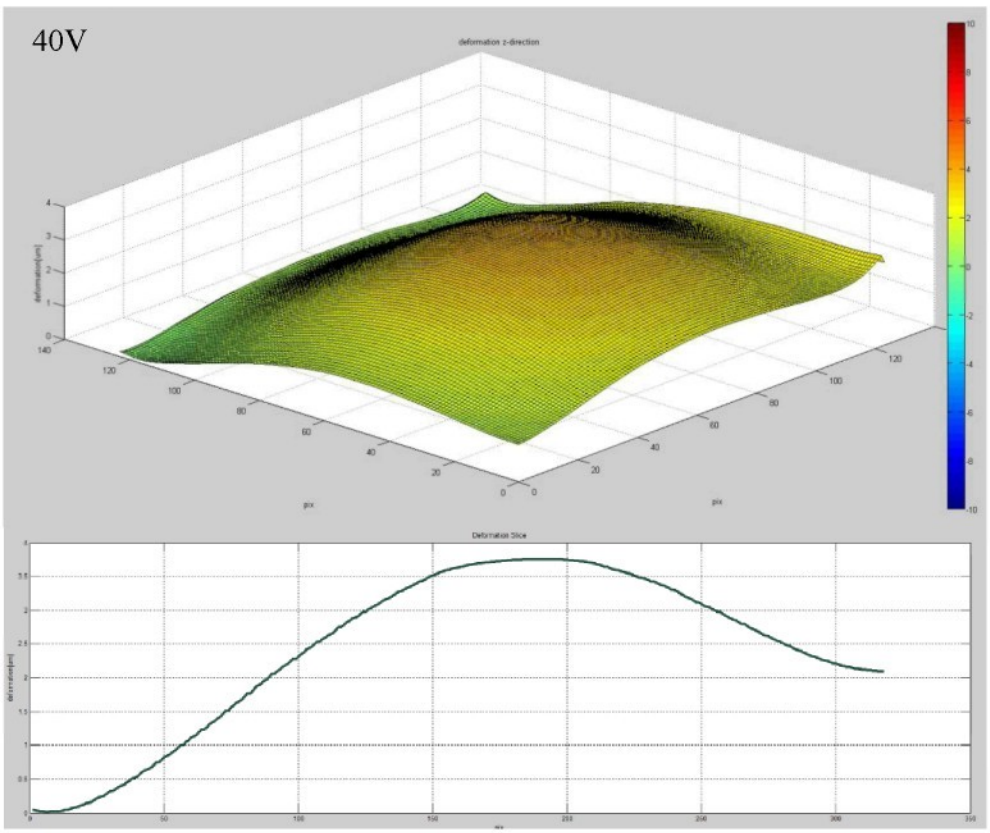

Fig. 49: Deformation distribution for 40V DC voltage.

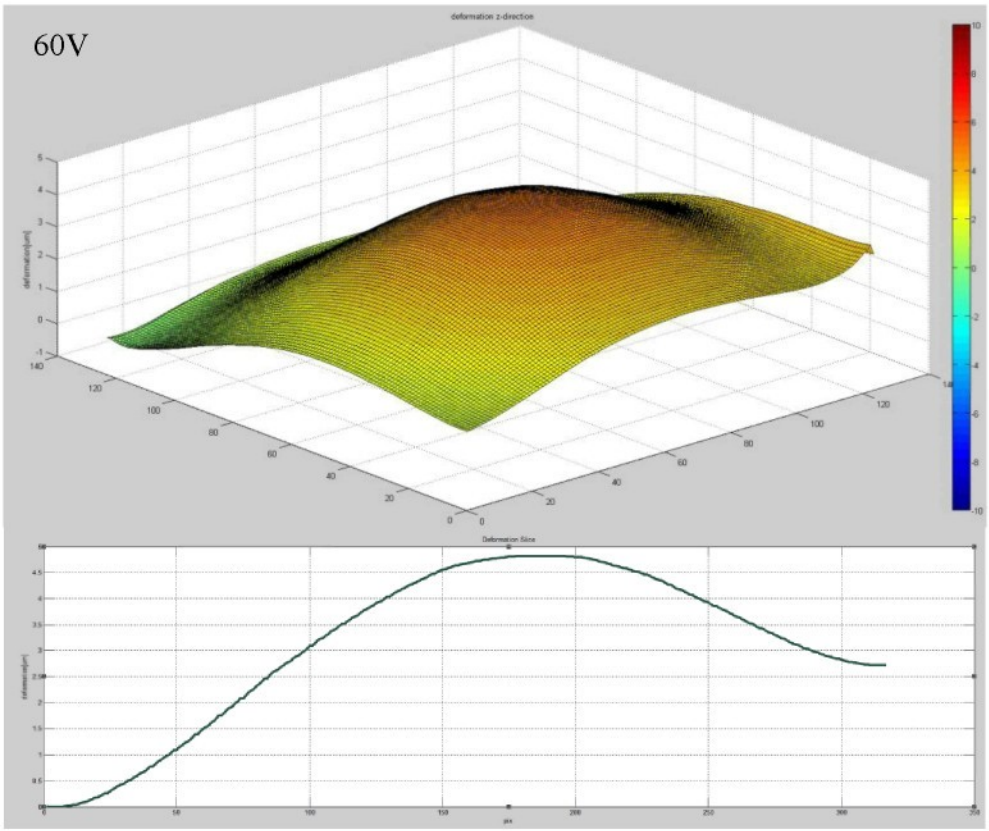

Fig. 50: Deformation distribution for 60V DC voltage.

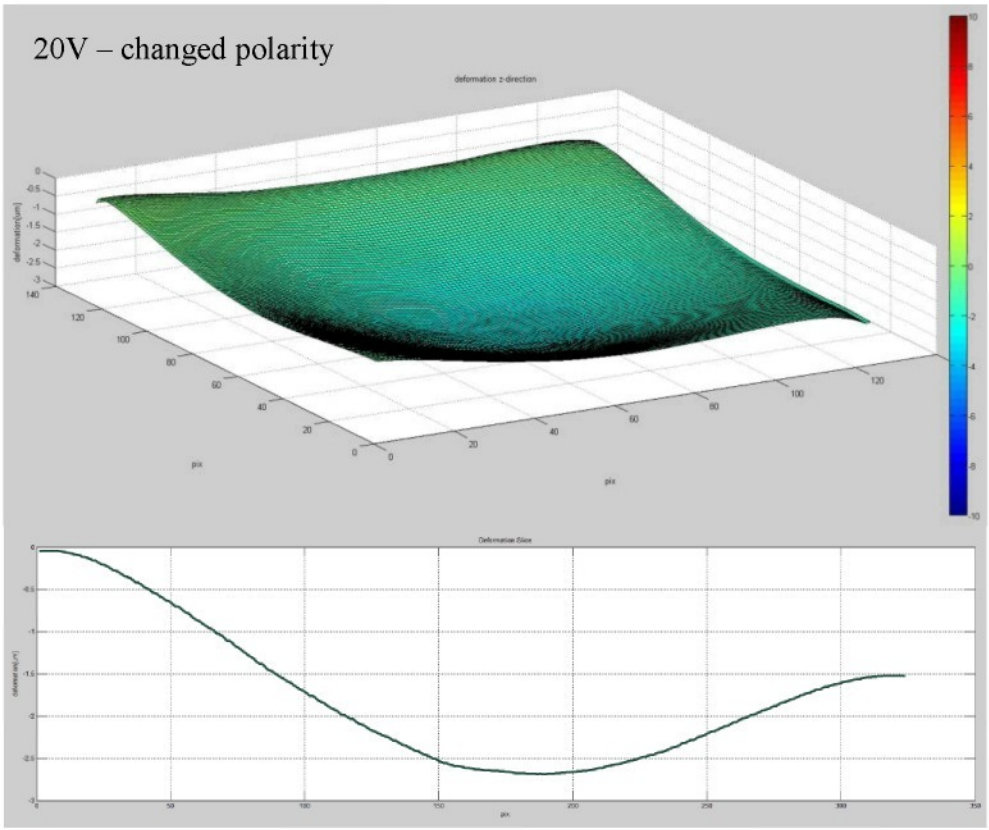

Fig. 51: Deformation distribution for 20V DC voltage with changed source polarity.

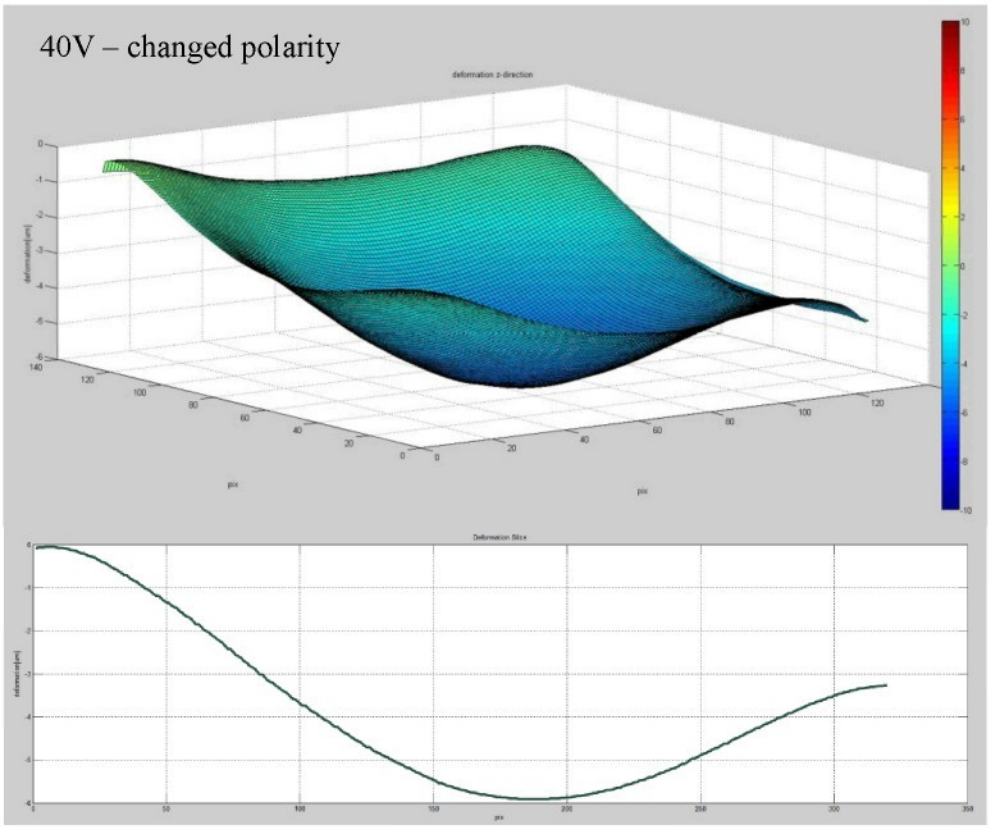

Fig. 52: Deformation distribution for 40V DC voltage with changed source polarity.

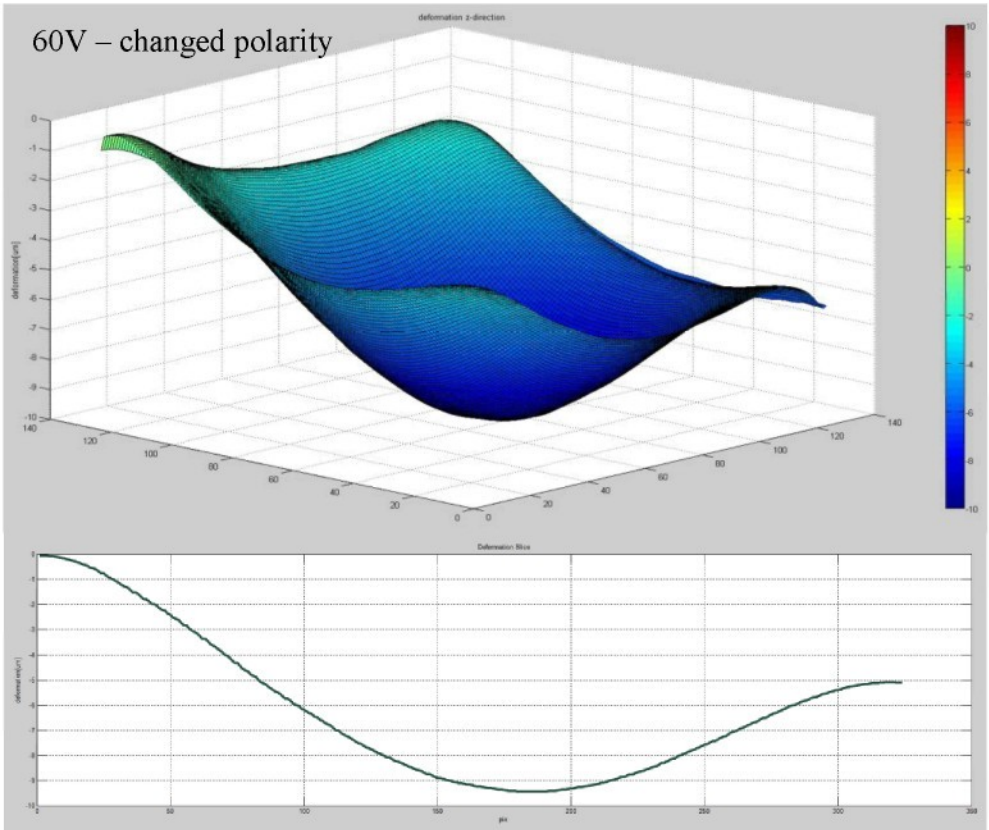

Fig. 53: Deformation distribution for 60V DC voltage with changed source polarity.

From these results we can estimate the behavior of the actuator with respect to the value of the DC voltage. As was expected the higher the value of voltage the higher the values of deformation we measured. Further the shape of deformation is given by the polarity of the DC voltage source. The shape can be convex (Fig. 48, Fig. 49, Fig. 50) and if we change the polarity, the shape of deformation distribution is concave (Fig. 51, Fig. 52, Fig. 53). For concave deformation distribution the values of deformation are higher than for the convex one while the voltage remains the same value.

3.1.2 Circle piezoelectric actuator

For measurement of deformation distribution of circle piezoelectric actuator we have applied the same procedure as for the squared actuator. Only the range of the applied DC voltage was 0 – 20V and no change of polarity was done.

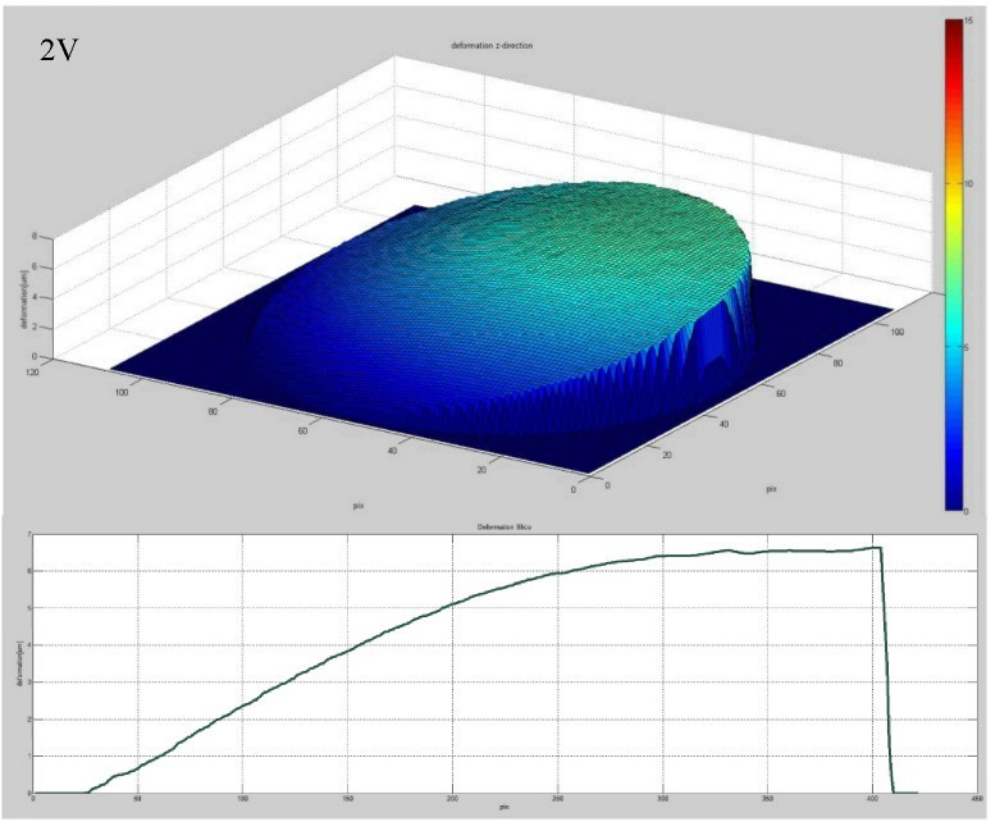

Fig. 54: Deformation distribution for 2V DC voltage.

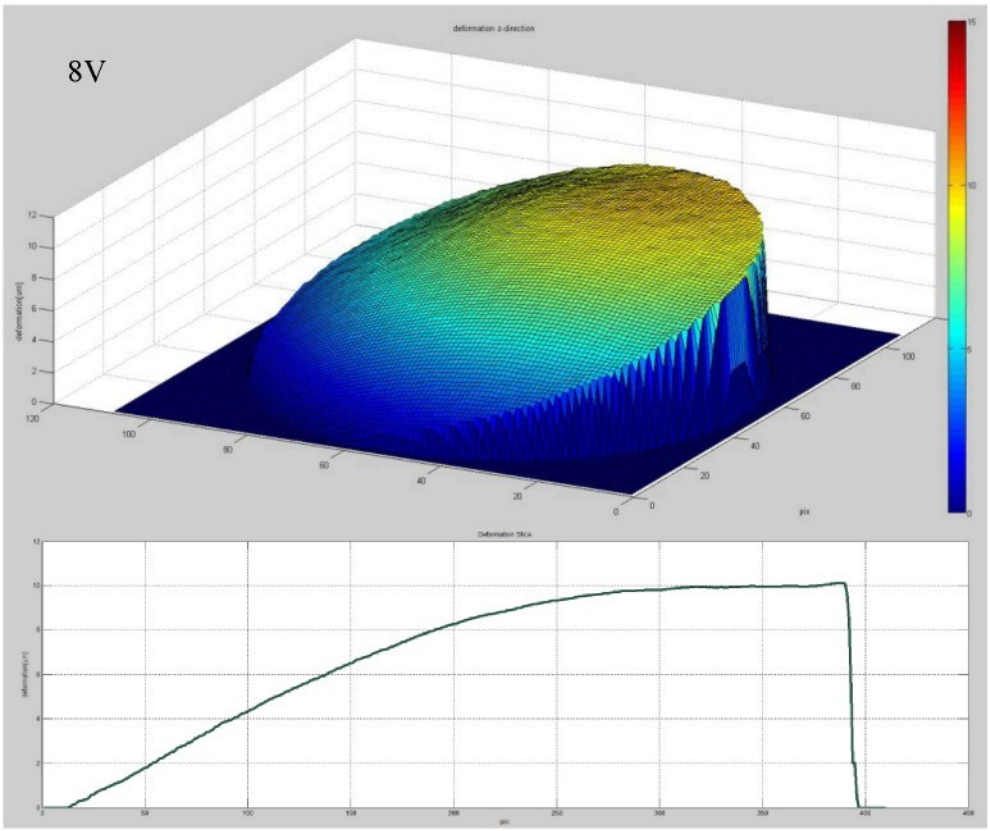

Fig. 55: Deformation distribution for 8V DC.

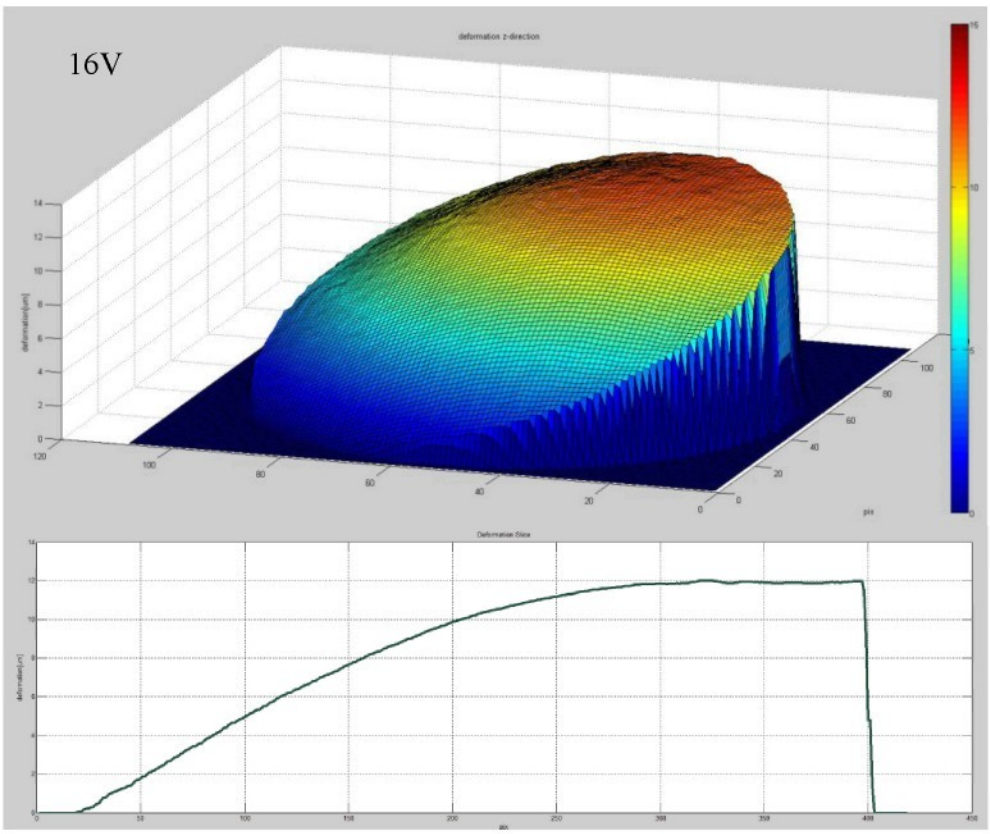

Fig. 56: Deformation distribution for 16V DC.

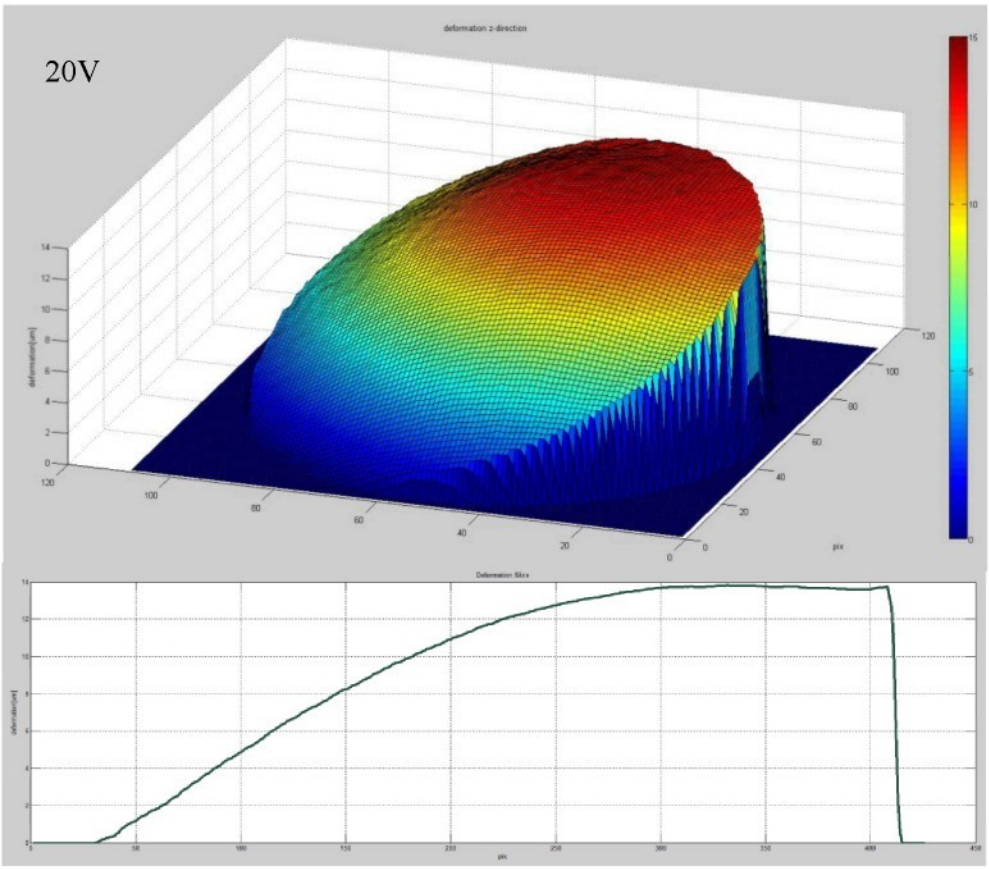

Fig. 57: Deformation distribution for 20V DC.

3.2 The deformation distribution measurement of plate caused by micrometer screw

In this case the deformation was caused by mechanical pressure generated by a micrometer screw. By turning the micrometer screw we got different states of deformation distributions. These states were recorded as a digital hologram and the deformation distribution was evaluate by DEMETER, see Fig. 58.

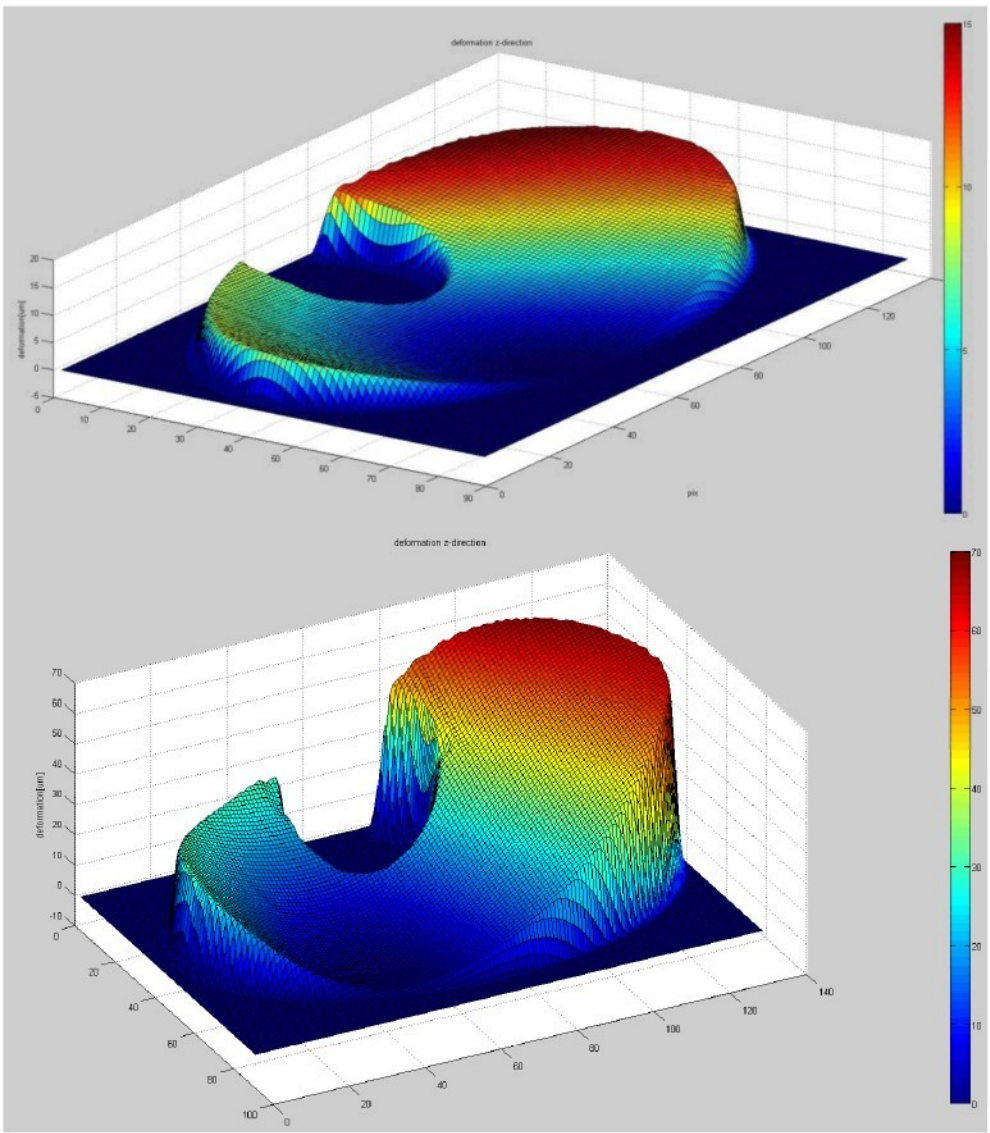

Fig. 58: Two different deformation distribution of a plate surface caused by micrometer screw.

The oval-shaped part with zero value of deformation is the part where the screw and the noisy background were masked. From these results it can be seen, that the deformation distribution is changing with respect to turning of the micrometer screw.

3.3 The phase object interference phase measurement

We have also tried to apply the digital holographic interferometry to phase object as was described in the chapter "refractive index measurement". As a phase object we have used a candle flame and the heated resistor. The red color corresponds to higher temperature.

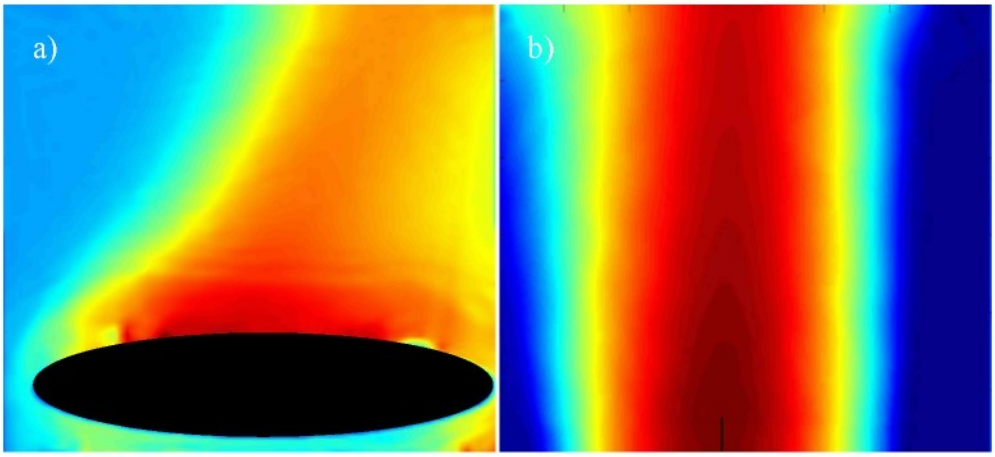

Fig. 59: The phase objects interference phase distributions: a) surrounding of a heated up resistor, the resistor is placed in the oval masked region; b) surroundings of a candlewick, which is placed vertically in the middle of the field.

Conclusion

The goal of this thesis is to outline the theory of digital holographic interferometry and on this basis to develop an application for the measurement of deformation distribution. Deformation measurement, using digital holographic interferometry is contactless, noninvasive, and a full surface method with very high precision, which is in fractions of a wavelength of the laser in use. These properties are very welcome and custom made for the method of a wide usage in engineering and industry.

Recent developments in digital holographic interferometry are mainly connected to the progress of the digital recording technique. A better resolution of digital sensors permits the recording of a very fine structure hologramically. The evaluation process of such holograms consists of many operations from different disciplines like optics or image processing. It leads to a complex and large application. The complexity and modernity of the problematic are probably the main reasons why such software has not been obtainable by purchase or as a freeware until now and, therefore, we had to develop this software ourselves.

The developed application meets requirements, such as robustness, adaptability and applicability to the entire range of different experimental holographic setups for accurate deformation measurement. The examined objects may be of almost any shape or material and the measurement and range of deformation extends from about $5.10^{-4} - 5.10^{2}\mu m$.

In this thesis we have firstly implemented a summary of the fundamentals of digital holographic interferometry, and present some formulas necessary for the development of the application. In the next chapter, the user interface and control of the application is introduced and described in detail. Finally, many experiments for the measurement of deformation distribution were conducted and consequently evaluated by the developed application. Results of these experiments are also included in this thesis.

The measurement of the deformation is not the only possible application of digital holographic interferometry. The measurement of the other physical quantities (strain distribution, stress distribution, refractive index distribution, temperature field distribution etc.) utilizes the double exposure method principle as well. Therefore the application can be easily adapted for the measurement of these quantities. Another

holographic interferometry method is a time-average method, for measurement the amplitudes of oscillating surfaces, in which we are also engaged. Currently, we are working hard to develop applications for the evaluation of other measured quantities.

Since the results, so far, are very encouraging and can withstand the strictest comparison, we will continue in the same way also in the future. The author of the thesis is also the author of a paper dealing with these issues and is the co-author several others.

A CD containing the developed application and the text of the diploma thesis is enclosed.

Author Publications

PSOTA, P.; LÉDL, V.; VÁCLAVÍK, J. MATLAB GUI for Deformation Measurement by Digital Holographic Interferometry. In . *International Conference Technical Computing Prague 2009*. [s.l.] : [s.n.], 2009. p. 9. ISBN 978-80-7080-733-0.

DOLEČEK, R.; LÉDL, V.; KOPECKÝ V.; PSOTA, P; VÁCLACÍK, J.; VÍT, T. Prospects of digital holographic interferometry in heat trasfer measurement. In *Experimental Fluid Mechanics 2009*. [s.l.] : [s.n.], 2009. p. 9. ISBN 978-80-7372-538-9.

References

[1] FUKA, J; HAVELKA, B. *Optics*. Praha : SPN, 1961. 845 p (in Czech).

[2] KREIS, T. *Handbook of Holographic Interferometry : Optical and Digital Methods* . Berlin : Wiley, 2004. 542 p.

[3] *MathWorks : MATLAB and Simulink for Technical Computing* [online]. 1994-2010 [cit. 2010-05-07]. Available at WWW: <http://www.mathworks.com/>.

[4] *MathWorks : 2D phase unwrapping algorithms* [online]. 1994-2010 [cit. 2010-05-11]. MATLAB Central. Available at WWW: <http://www.mathworks.com/matlabcentral/fileexchange/22504-2d-phase-unwrapping-algorithms>.

[5] MCELHINNEY, C., et al. Removing the twin image in digital holography by segmented filtering of in-focus twin image. *Optics and Photonics for Information Processing II* [online]. August 2008, vol. 7072, [cit. 2010-05-07]. Available at: <http://www.digitalholography.eu/varasto/McElhinney%202009%20Proc.pdf>.

[6] MILER, M. *Holography : theoretical and experimental basics and its applications*. Praha :SNTL, 1974. 272 p. (in Czech)

[7] PSOTA, P. *The basis of the numerical reconstruction of digital holograms*. Liberec, 2009. 24 p. Semestral project. Technical University of Liberec.

[8] RASTOGI, P.K. *Holographic Interferometry : Principles and Methods*. Berlin : Springer , 1994. 341 p.

[9] SALEH, B. E. A; TEICH, M. C. *Fundamentals of Photonics*. Praha : MATFYZPRESS, 1994. 1056 p. (in Czech)

[10] SCHNARS , U.; JUEPTNER, W. *Digital Holography*. Berlin : Springer, 2005. 164 p.

[11] TORRES, E.; GUERRERO BERMUDEZ, J. Digital Fourier holography : recording and optical reconstruction. *Revista Colombiana de Fisica*. 2003, 35(1), p. 168–171.

[12] TRAGER, F. *Springer Handbook of Lasers and Optics*. New York : Springer, 2007. 1331 p.

[13] VENEMA, T.M.; SCHMIDT, J.D. Optical phase unwrapping in the presence of branch points. *Optics Express*. May 2008, 16, p. 6985-6998.