ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

INGEGNERIA ELETTRONICA, INFORMATICA E
DELLE TELECOMUNICAZIONI

Ciclo XXIV

**Settore Concorsuale di afferenza:** 09/F1 - Campi Elettromagnetici
**Settore Scientifico disciplinare:** ING-INF/02 - Campi Elettromagnetici

# Optical Design for Automatic Identification and Portable Systems

**Presentata da:**
**Ing. Luisa De Marco**

**Relatori:**
**Chiar.mo Prof. Ing. Paolo Bassi**
**Ing. Federico Canini**

**Correlatore:**
**Ing. Marco Gnan, PhD**

**Coordinatore Dottorato:**
**Chiar.mo Prof. Ing. Luca Benini**

Esame finale anno 2012

*Due cose riempiono l'animo*
*con sempre nuovo e crescente*
*stupore e venerazione,*
*quanto più spesso e accuratamente*
*la riflessione se ne occupa:*
*il cielo stellato sopra di me,*
*e la legge morale dentro di me.*

**La citazione è di Immanuel Kant,**
**gliel'ho sussurrata io (chiaramente).**

**-Bassam Hallal-**

# Abstract

This thesis proposes design methods and test tools, for optical systems, which may be used in an industrial environment, where not only precision and reliability but also ease of use is important. The approach to the problem has been conceived to be as general as possible, although in the present work, the design of a portable device for automatic identification applications has been studied, because this doctorate has been funded by Datalogic Scanning Group s.r.l., a world-class producer of barcode readers. The main functional components of the complete device are: electro-optical imaging, illumination and pattern generator systems. For what concerns the electro-optical imaging system, a characterization tool and an analysis one has been developed to check if the desired performance of the system has been achieved. Moreover, two design tools for optimizing the imaging system have been implemented. The first optimizes just the core of the system, the optical part, improving its performance ignoring all other contributions and generating a good starting point for the optimization of the whole complex system. The second tool optimizes the system taking into account its behavior with a model as near as possible to reality including optics, electronics and detection. For what concerns the illumination and the pattern generator systems, two tools have been implemented. The first allows the design of free-form lenses described by an arbitrary analytical function exited by an incoherent source and is able to provide custom illumination conditions for all kind of applications. The second tool consists of a new method to design Diffractive Optical Elements excited by a coherent source for large pattern angles using the Iterative Fourier Transform Algorithm. Validation of the design tools has been obtained, whenever possible, comparing the performance of the designed systems with those of fabricated prototypes. In other cases simulations have been used.

# Contents

# Introduction

Optical systems, devices for producing or controlling light, are commonly used in the daily life. There are simple systems that involve just lenses such as spectacles, contact lenses, etc. , and complicated ones such as cameras, microscopes, lighting devices, etc. , that involve not only lenses but also light sources (laser, LED), detectors (CCD), electronics etc.

These complex systems can be classified as imaging and non-imaging. Imaging systems involve the formation of an image of an object, whereas non-imaging systems are designed for the optimal transfer of light radiation between a source and a target. The present work is concerned with the design of both types of systems.

These devices can be made to be portable. Portable systems involve more problems than non-portable ones and their size (not too large), weight (not too heavy), battery life (long-lasting), mechanical sturdiness become critical topics for a correct design.

Optical systems for Automatic Identification (Auto ID) applications have been considered as case studies. Auto ID is generally known as the automatic data capture using technical devices without human involvement to allocate objects to a class through an identification system.

The Auto ID technologies can be divided in two categories:

- Data carrier technologies: their aim is to collect, store and carry data encoded using right platform; in this category there are optical methods (mainly barcode and Optical Character Recognition, OCR), magnetic storage methods (magnetic stripes) and electronic storage methods (RFID-tag, smart card, chip, smart label, etc.).

- Feature extraction technologies: they could be subdivided in three groups considering the type of the feature, which can involve static (fingerprints, face shape, eye retinal or iris), dynamic (voice, gait, dynamic signature) or chemical-physical properties.

Applications of Auto ID systems are used in the common life for biometric security, logistics, item tracking, inventory control, shipping, health care, warehousing etc. According to the type of application, these systems could be portable
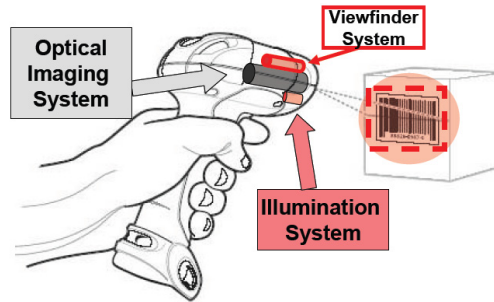
Figure 1: Scketch of a barcode reader and its main components.

or not. An Auto ID portable device is necessary when the user has to carry it easily for identifying the desired object.

The purpose of this work has been the developing of design and characterization tools for improving the performance of Auto ID and portable systems. As an application example, the implemented tools have been used for the optimization of a barcode reader system. The choice of this device comes from the fact that this doctorate has been funded by Datalogic Scanning Group s.r.l., a world-class producer of barcode readers. Since this thesis comes from the collaboration between university and industry, the obtained results have also an industrial importance. Moreover, some sensitive data has been deliberately omitted from this thesis because of obligation to confidentiality.

The barcode reader is composed of three main systems, as schematically shown in Fig. 1:

**Electro-Optical Imaging System** : it is the most important part of the reader because it allows to extract the desired information from the barcode signal.

**Illumination System** : it can increase the irradiance on the barcode plane and help the imaging system to have the proper Signal to Noise Ratio (SNR) for successful decoding of the barcode signal coming from the sensor.

**Pattern generator as viewfinder system** : it helps the user to hit the barcode center. In this way the image can contain all the required information.

This work is divided into two parts: *Imaging Systems* and *Non-Imaging Systems*. In the following these systems and the main features of the tools implemented for improving their performance will be described.

The *Imaging Systems* part describes results concerning electro-optical imaging systems.

For all imaging systems, such as cameras, the image quality is an important parameter for estimating the system performance. For example in Auto ID applications, it is necessary to have an image of the item to be identified as sharp as possible. In other words, the better the ability of the system to capture fine details of the object, the easier the object identification. For this reason, the spatial resolution capability is important for imaging system performance evaluation.

Before designing an imaging system, it is essential to have all the tools for its performance evaluation. Ch. 1 has been focused on this problem. Since standards allow to compare and verify performance of different systems, the implementation of a tool for spatial resolution measurements that follows the ISO 12233 standard on Photography, Electronic Still Picture Cameras, Resolution Measurements [1] has been done. This International Standard defines a method for performing spatial resolution evaluations for imaging systems measuring its Spatial Frequency Response (SFR).

The implemented tool is called SAFARILAB (SFR measurement for a LAB environment). It includes the set-up for measuring the SFR and the SFR evaluation algorithm. Thanks to it, it is possible not only to measure the SFR of an imaging system and consequently to estimate its resolution performance, as described in the standard, but also to reduce the noise that affects the measurement. This tool is now daily used in Datalogic Scanning Group s.r.l. for evaluating the performance of their designed and developed systems.

After setting up SAFARILAB , the work has been focused on the implementation of design tools. The prototypes realized after optimization have been characterized by SAFARILAB . In particular, the design tools have been implemented to be used with ZEMAX® [2], an optical Computer-Aided Design (CAD) tool, a professional CAD for industrial optical design.

The first implemented tool, named Optical Level OptiMization (OLOM ) tool, is described in Ch. 2. It improves ZEMAX® optimization of the optical part of the system. OLOM provides ZEMAX® with a set of optimization constraints tailored to achieving systems having invariant frequency response, according to the required specifications. The joint use of OLOM and ZEMAX® allows to yield a system with the highest and most invariant with defocus frequency response.

The second tool, described in Ch. 3, is named SLALOM , acronym of "System Level AnaLysis and OptiMization", and is divided into two tools that can be run independently: Optimization-SLALOM (O-SLALOM ), for design and optimization of new systems and Analysis-SLALOM (A-SLALOM ) for the analysis of the existing ones. In this case, the optimization is not only of the optical block but of the whole system, including lens, detector and electronics. SLALOM , which is the only contribution to this work developed by other members of the optics group of DEIS (Dipartimento di Elettronica, Informatica e Sistemistica of University of Bologna) eng. Anna Guagliumi, eng. Antonella Liberato and eng. Marco Gnan, PhD, has been used to design an imaging system considering all blocks and starting from a lens optimized at the optical level by OLOM .

For testing and validating these design tools, two different systems, well known in the literature, with the particular ability of extending the Depth Of Field (DoF) have been chosen. The DoF is the range of distances in which the objects in a scene appear acceptably sharp in the formed images. An Extended DoF (EDoF) is an important characteristic especially for Auto ID systems, because in this type of applications, it is better to identify the item independently of its position respect to the device.

The EDoF system to be designed by OLOM is the Cubic Phase Mask (CPM) system [3]. This system consists of a lens, a phase mask with cubic shape, the detector and the electronics. The CPM has been chosen because, thanks to its solid theoretical bases, it allows to validate the design tool comparing the measured results to the expected (theoretical) ones. The prototypes of the designed systems have been realized and the performance has been measured by SAFARiLAB tool.

On the other hand, the EDoF system to be designed by SLALOM is the Quartic Phase Mask (QPM) system [4]. It is conceptually similar to the CPM, but in this case the phase mask added to the lens has a quartic shape. The QPM system is the preferred choice because the phase mask has circular symmetry, instead of the asymmetric shape of CPM. This characteristic allows to use conventional image enhancing techniques and consequently to validate the tool without focusing on the development of a dedicated post-processing.

In both design processes, typical required specifications for a barcode reader have been considered. These tools could be used also with other requirements for different imaging systems.

The second part of the thesis (*Non-Imaging Systems*) describes results concerning illumination and pattern generator systems.

These systems are respectively based on an incoherent source (LED) and a coherent (laser) one. They project the light with a desired shape on the target with a specific Field of View (FoV). The FoV is the range of angles that defines the area to be illuminated.

Two design tools have been implemented for designing the optical part of these systems.

The first tool, presented in Ch. 4, allows to design an illumination system based on a LED. This tool, implemented in MATLAB$^®$ , optimizes the shape of a lens for achieving a custom illumination. The lens profile is described by a $6^{th}$ order polynomial, also called free-form shape. The free-form lens has been optimized, as an example, according to two specifications. First of all the illumination should be concentrated on the desired FoV. Secondly the edge part of the illuminated zone should have more light than the central one. This kind of illumination shape is useful for Auto ID systems, such as barcode readers. Adding this illumination system to an imaging one, it is possible to have more light on the item to be identified and to compensate for both the target angular reflectivity and the angular response of the imaging system. This tool could be used for all

kind of applications that need a custom illumination. Prototypes of the designed free-form lens has been fabricated and the obtained results have been validated, thus confirming that the design tool for free-form lenses works successfully.

The tool implemented in MATLAB® for designing non-imaging systems using coherent sources has been presented in Ch. 5. A procedure to design Diffractive Optical Elements (DOEs) as pattern generator with large diffraction angles (large FoV) once illuminated by a laser source has been set up. The tool is based on the Iterative Fourier Transform Algorithm (IFTA) that has been widely used for pattern generator design [5–7]. IFTA implies fulfillment of the so called paraxial approximation. The implemented tool circumvents this problem. Pattern generator systems can be used in Auto ID applications, for helping the user to hit the item to be identified. A possible example is the viewfinder system present in barcode readers. However, it is also possible to use a pattern generator for other type of applications, as 3D reconstruction. For this purpose, several lines with large angle should be projected on the item to be 3D reconstructed. For validating the implemented algorithm, several prototypes, as example of viewfinder and 3D reconstruction pattern, have been fabricated.

Finally, conclusions will be drawn, considering all the results obtained in this work.

# Part I

# Imaging Systems

# 1

# Imaging System Characterization

In this chapter the implemented tool for measuring the performance of an imaging system will be illustrated. The realization of this tool has been essential before starting the imaging system design.

The ISO 12233 standard [1] defines a method for performing spatial resolution evaluation of an imaging system measuring its SFR. Although many implementations of this standard are available, for example IMATEST (commercially distributed) [8] or IMAGEJ (free) [9], a new implementation of the standard, named SAFARILAB , has been done to enhance some particular features. In SAFARILAB more operational options than in the standard have been added to improve the numerical calculations and to reduce the noise that affects the measurements.

Sec. 1.1 defines preliminarily which parameters must be measured. Sec. 1.2 describes the SAFARILAB tool in detail. Finally, Sec. 1.3 presents the results of some tests done on SAFARILAB to assess the validity of the tool itself.

## 1.1 Imaging System Response

The system which should be characterized from the point of view of its image quality, i.e. its ability to create an "image" which reproduces "faithfully" an "object", can be considered as the cascade of many components: the object itself (a spatial field distribution with temporal features depending on the spectral characteristics of the source which illuminates it), one or more optical devices (lenses or lens assemblies), the detector and some electronics able to reproduce the image on a physical support (a display, for example).

The overall response of the system under test can be described, in the spatial coordinate system, by its Point Spread Function (PSF), the response measured on the image plane when the system is excited by an ideal point source on the object plane [10–12]. The PSF can be described as amplitude or intensity. In this work the PSF is usually considered as intensity. When it will be used as amplitude, this will be specified.

The overall system PSF is the convolution of the PSFs of the various blocks constituting the system under investigation (sampling, s , optics, o, detector, d, electronics, e). One can write, considering also possible time dependencies:

$$g(x,y,t) = g_s(x,y,t) \star g_o(x,y,t) \star g_d(x,y,t) \star g_e(x,y,t). \tag{1.1}$$

If the system is linear, which requires the approximation of temporal and spatial stationarity and paraxial one to be valid, its behavior can be described also in spatial frequency domain using the SFR, which is the system transfer function, calculated as the product of the transfer functions of the various blocks (the Fourier transforms of the relevant PSFs):

$$G(\xi,\eta,\omega) = G_s(\xi,\eta,\omega) \cdot G_o(\xi,\eta,\omega) \cdot G_d(\xi,\eta,\omega) \cdot G_e(\xi,\eta,\omega) \tag{1.2}$$

where each of the $G(\xi,\eta,\omega)$ is the spatial and temporal Fourier transform of the corresponding $g(x,y,t)$. In practice one can measure $g(x,y,t)$ or $G(\xi,\eta,\omega)$. Then, the desired information (for example the optical response or the detector one) must be extracted from the overall system response. This requires knowing the response of the other blocks.

In order to do so, the complete system can be divided in some sub-blocks, optical system, detector, electronics. The response of each sub-block, will be described in the following focusing on the definitions useful for this work. More details can be found in [10, 13].

## 1.1.1  Optical response

The sketch of the optical system is shown in the upper part of Fig. 1.1. It is assumed to be made by a combination of apertures (completely absorbing planes with an hole with known transmittance function, defined as the ratio of the transmitted field amplitude and incident one) and thin ideal lenses (such that the field impinging at some transversal spatial coordinates has, at least approximately, the same spatial coordinates when it comes out from it, affected only by a phase change). In this case, the whole system can be considered as a "blackbox" described by its PSF.

Such an optical system is said to be diffraction limited if a spherical wave emitted by a point source on the object plane is transformed by the optical system into a spherical wave converging in another point on the image plane (see lower part of Fig. 1.1). In a real system, aberrations introduce further distortions which make the output wave not perfectly spherical.

The ratio between the output and input coordinates is the system Magnification (M). It is important to notice that, this holds only for points belonging to a region around the optical axis of the system, where the so called *paraxial approximation*, or *Slowly Varying Envelope Approximation* (SVEA) [10–12].

An optical system can be investigated in the spatial domain or in the spatial frequency domain, and for coherent and incoherent illumination. Considering
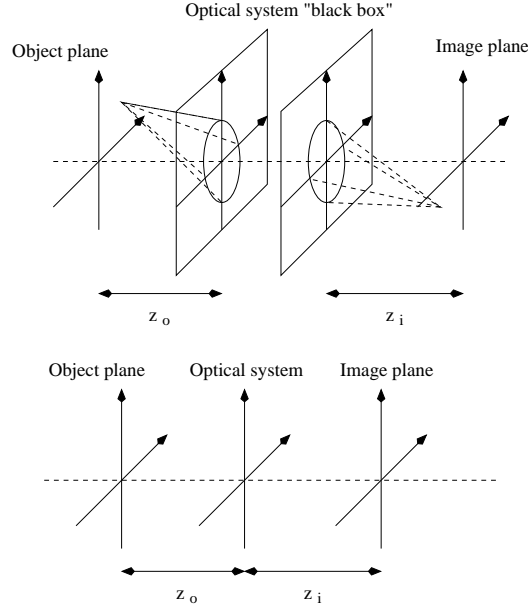
Figure 1.1: Upper part: ideal optical system schematic. Lower part: simplified system schematic.

the spatial domain, if coherent illumination is used, the image amplitude is the convolution between the object field complex amplitude and the PSF of the considered optical system. If incoherent illumination is considered, the image intensity is the convolution between the object intensity and the square of the modulus of the PSF of the optical system.

While considering the spatial frequency domain, in the case of coherent source, instead of considering the convolution of the amplitude PSFs, one can define the transfer function of the system as the product of the Amplitude Transfer Functions (ATFs) , which are the Fourier transforms of the spatially invariant PSFs:

$$H(\xi, \eta) = \iint_\infty h(x, y)\, e^{-j2\pi\,(\xi x + \eta y)}\, dx\, dy. \tag{1.3}$$

If incoherent sources are considered, things become a little bit more complicated, as convolution involves field intensities and not amplitudes. In this case, one can define the intensity normalized Transfer Function as:

$$\mathcal{H}(\xi, \eta) = \frac{\iint_\infty |h(u, v)|^2\, e^{-j2\pi(\xi x + \eta y)}\, dx\, dy}{\iint_\infty |h(u, v)|^2\, du\, dv}. \tag{1.4}$$

It holds then:

$$\mathcal{G}_i(\xi, \eta) = \mathcal{H}(\xi, \eta)\, \mathcal{G}_g(\xi, \eta), \tag{1.5}$$

where $\mathcal{G}_i$ and $\mathcal{G}_g$ are respectively the spectra on the image and object planes. $\mathcal{H}(\xi, \eta)$ is known as Optical Transfer Function (OTF) of the system. Its modulus $|\mathcal{H}(\xi, \eta)|$ is referred to as Modulation Transfer Function (MTF). The system
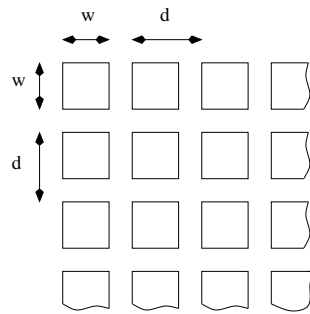
Figure 1.2: CCD structure schematic.

OTF and MTF describe the optical device behavior when incoherent illumination is used. It is important to note that the MTF does not completely characterize the optical system as it lacks of the phase information. Nevertheless, the MTF is the standard way to describe the relationship between the object and image plane intensities in incoherent optical systems.

The MTF can be used to define the spatial resolution of the system, i.e. the minimum distance at which two different point sources in the object plane can be distinguished in the image plane. In the spatial frequency domain this corresponds to the inverse of the spatial frequency value at which the MTF falls below a given threshold [10, 13].

### 1.1.2   Detector

The detector response depends on the technology, it is based on: photographic paper, for example, behaves differently from a CCD array. The former can acquire images with a spatial resolution depending on the sensitivity of the film (size of the sensitive grain), the latter is intrinsically limited by the size of the CCD array elements and their spacing.

A CCD array can be considered as a composition of detectors (supposed square, with size $w \times w$), placed periodically on a grid of size $d \times d$.

Both the values of $w$ and $d$ affect the final result. The effect of the detector pixel size ($w$) is referred to detector footprint MTF. The effect of the distance between the centers of adjacent pixels ($d$) is related to two phenomena which will be referred to as aliasing MTF and sampling MTF. These three effects will be summarized in the following.

**Detector footprint**

The finite size of each element constituting the CCD array makes the measured intensity equal to the integral of the detected one. Considering, for simplicity, the

$1D$ case:

$$\int_0^w I(x)\, dx = I(x) \star \text{rect}\left(\frac{x}{w}\right) = I(x) \star h_{fp}(x)$$

where $I(x)$ is the detected optical intensity and $h_{fp}$ is the PSF of the detector footprint. In the spatial frequency domain it then holds:

$$MTF_{fp}(\xi) = \mathcal{F}\{h_{fp}(x)\} = |\text{sinc}\,(\xi w)| = \left|\frac{\sin \pi \xi w}{\pi \xi w}\right|. \tag{1.6}$$

In the $2D$ case, one must refer to:

$$h_{fp}(x,y) = \text{rect}\left(\frac{x}{w_x}\right) \cdot \text{rect}\left(\frac{y}{w_y}\right)$$

and

$$MTF_{fp}(\xi, \eta) = |\text{sinc}\,(\xi w_x)|\,|\text{sinc}\,(\eta w_y)|. \tag{1.7}$$

**Aliasing**

The first effect of the distance between adjacent pixels is aliasing. Such a phenomenon is related to the maximum detectable spatial frequency by:

$$\xi_N = \frac{1}{2d} \tag{1.8}$$

Such a frequency is known as the Nyquist frequency. The presence of spatial frequencies larger than $\xi_N$ causes an overestimation of lower frequencies (the so called spectrum folding effect). It is then important either to eliminate such frequencies (this needs to operate directly on the optical signal and generally it is not possible) or reducing them to a minimum.

To reduce the negative effect related to the presence of frequencies larger than $\xi_N$, one can take advantage of the finite bandwidth of the optical device before the detector. It is acceptable that the MTF goes to $0$ at $\xi_N$. The price to pay is a further reduction of the detected signal at higher frequencies. Some aliasing must then be accepted if frequencies around $\xi_N$ must be present for any reason [13].

**Sampling**

Sampling itself introduces a further problem as it causes the set-up to become not space invariant, as shown in Fig. 1.3. An ideal line source impinging exactly on a single column of pixels provides a maximum response of the detector. But, if it is slightly displaced, it is detected by two adjacent columns and this reduces the signal level in the involved pixels and degrades the overall MTF.
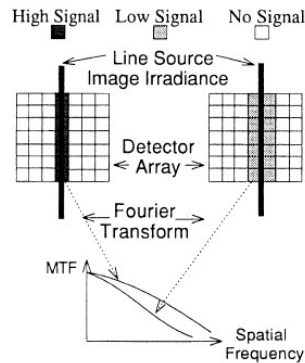
Figure 1.3: Effect of the different relative position of the line source and the pixels of the detector array. From [13].

### 1.1.3  Electronics

Also in the case of the electronic part of the system, there is a frequency response which contributes to the final signal generation. Also these effects can be described by a spectral characteristic. To convert temporal, typical for electronic devices, to spatial frequencies, one can, for example, detect an image as a bar target of known fundamental frequency. This frequency creates a spatial frequency in the image that can be calculated knowing the optical magnification or can be measured directly from the output signal knowing the pixel-to-pixel spacing of the detector array. Inputting the output video signal from the detector array to the spectrum analyzer gives an electrical frequency corresponding to the fundamental image-plane spatial frequency of the bar target [13].

### 1.1.4  Noise

In the imaging system characterization, it is also important to consider the noise. Noise affecting measurement results comes from different sources [14]:

**Shot noise** related to statistical fluctuations of the charges generated by the stream of received photons;

**Fano noise** due to non constant efficiency of the conversion process between photons and electrical charges (generally negligible, however, in the spectral range of interest for imaging system design);

**Fixed pattern noise** related to spatially non uniform response of the pixels constituting the receiver array (it is named "fixed" as it varies if one compares

two different detector arrays, but it is constant in each of them). Fixed pattern noise depends on physical features of the receiver array.

**Read noise** which generally includes all the signal independent noise contributions (dark current, thermal noise, Analog-To-Digital-Converter quantization related effects, for example).

In the following, only the Shot noise contribution will be considered. It can be described by a Poisson statistical distribution with optical intensity dependent mean square $\sigma^2_{SHOT}(I)$ given by:

$$\sigma^2_{SHOT}(I) = I \, \frac{e^{hc/\lambda kT}}{e^{hc/\lambda kT} - 1}, \qquad (1.9)$$

where $I$ is the impinging optical signal intensity, $h = 6.626 \cdot 10^{-34}$ J s is the Planck's constant, $\lambda$ is the free space photon wavelength (in m), $k = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, $c = 2.99 \times 10^{8}$ cm/s is the free space light speed and $T$ is the absolute temperature ($^{\circ}$K) [14].

## 1.2 SAFARILAB tool

As described in Sec. 1.1, the PSF of an optical system defines its behavior. If the system is linear, the PSF evaluation is equivalent to the SFR one, since it is the spatial Fourier Transform of the PSF. In the case of imaging systems with incoherent illumination and considering just the intensities, the SFR reduces to the system MTF.

This section describes firstly the SFR evaluation method, as suggested in the standard, then the SAFARILAB tool, focusing on both its part: the algorithm and the set-up.

### 1.2.1 SFR evaluation

Following the definition, the SFR can be evaluated measuring the response of the system excited by a point source (see Fig. 1.4 (a)), represented by the so called Dirac pulse $\delta(x, y)$. The SFR is then the Fourier transform of the PSF. In this case, the input object is $f(x, y) = \delta(x, y)$ and the output image $g(x, y)$ equals the $PSF(x, y)$, by definition. But this is not a simple solution to work out as problem comes from the need of a true $0D$ *point* source emitting a signal with satisfactory power.

Alternatively, one can consider a $1D$ line source and evaluate the so called system Line Spread Function (LSF) (see Fig. 1.4 (b)). In this case, orienting the line source, for example, along the $y$ direction, the input object is $f(x, y) = \delta(x) \cdot 1(y)$,

where $1(y)$ denotes a function constant along $y$, and the output image is:

$$
\begin{aligned}
g(x,y) = LSF(x) = f(x,y) \star PSF(x,y) &= (\delta(x) \cdot 1(y)) \star PSF(x,y) \\
&= \int_{-\infty}^{\infty} PSF(x,y') \, dy'
\end{aligned}
$$

The MTF along the $\xi$ axis is then the Fourier transform of the LSF:

$$
MTF(\xi,0) = |\mathcal{F}\{LSF(x)\}| \,.
$$

Reorienting the line source one can then determine the MTF along any direction. It must be noticed that $LSF(x) \neq PSF(x,0)$.

Unfortunately, also a line source is not easy to be fabricated as it requires the $1D$ distribution of ideal point sources. The ISO 12233 standard suggests then to consider another source, which is easier to approximate in practice, the so called *knife edge*:

$$
f(x,y) = step(x) \cdot 1(y)
$$

noting that:

$$
step\,(x - x_0) = \int_{-\infty}^{x} \delta(x' - x_0) \, dx' \iff \delta(x - x_0) = \frac{d}{dx} step\,(x - x_0)
$$

one can finally (see Fig. 1.4 (c)) evaluate the so called system Edge Spread Function (ESF). It holds:

$$
\frac{\partial}{\partial x} ESF(x) = \frac{\partial}{\partial x} \int_{-\infty}^{x} LSF(x') \, dx' = LSF(x) = (\delta(x) \cdot 1(y)) \star PSF(x,y).
$$

The PSF is obtained Fourier transforming the LSF obtained, in turn, after the derivative of the ESF.

In other words, the ESF derivative is the LSF and the SFR is the $1D$-Fourier transform of the LSF along the direction of interest ($x$). The standard is then based on the ESF measurement.

The choice to measure the ESF has the further advantage that all the image lines provide an independent measurement of the knife edge image. To improve the measurement quality, the standard suggests to position the knife edge at an angle of about $5°$ (slanted edge). In this case, considering each line on the same reference axis, one gets an irregularly spaced set of data. Groups of data points (the ISO standard fixes to 4 the number of points per group) can then be averaged to provide an oversampled ESF (see Fig. 1.6). Operating in this way, detrimental effects related to the relative position of the knife edge with respect to the sensor pixel matrix (see Fig. 1.3) can be reduced, making the system space-invariant. It is then crucial to have a precise determination of the relative position of all the image lines.
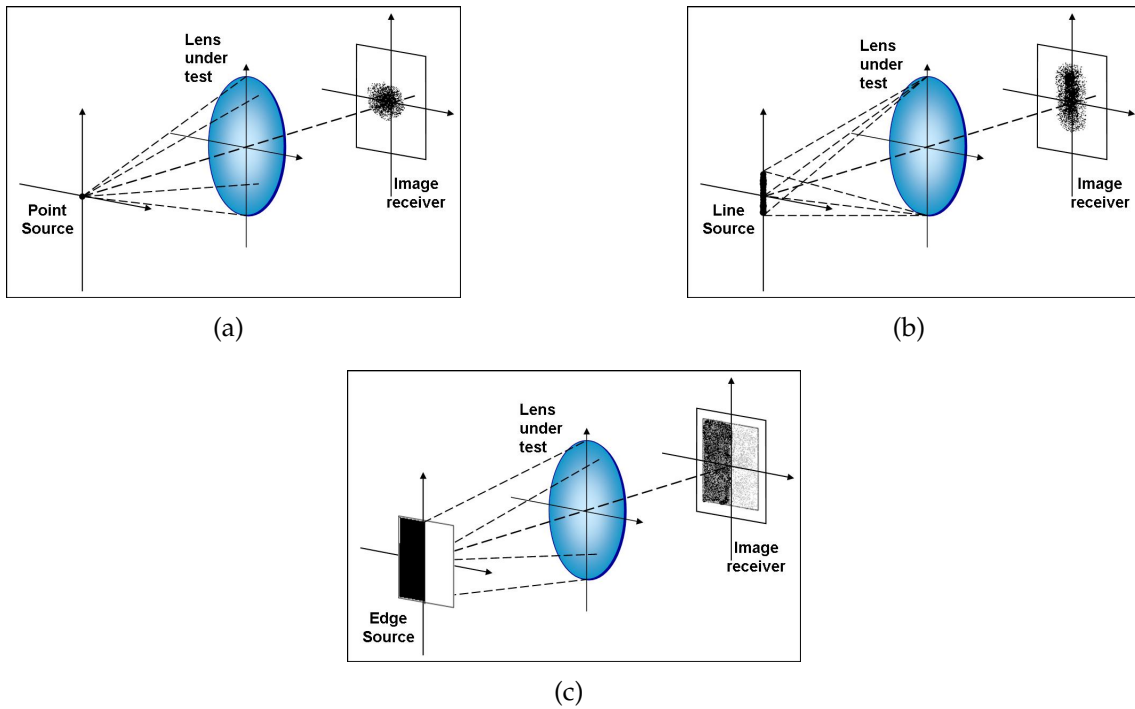
(a)



(b)



(c)

Figure 1.4: Representation of Point Spread Function (PSF) (a), Line Spread Function (LSF) (b) and Edge Spread Function (ESF) (c) of a system. The CCD array and the relevant electronics are represented by the screen where the image is shown.

### 1.2.2   SAFARILAB algorithm

The algorithm for the SFR calculation has been implemented in MATLAB® . The SAFARILAB routine follows the standard, taking particular care in some critical points.

The standard algorithm steps are:

1. Image acquisition and selection of the *Region Of Interest*, ROI, the part of the image which contains only the slanted edge (see Fig. 1.5).

2. Determination of the edge slope (the angle formed by the edge and the vertical axis). This value comes from a linear regression operation on the so called centroids, the peaks of the derivatives of each image line (LSF). This is a critical step as the slope influences the position of the lines on the common reference frame. For this reason, the standard suggests to filter the LSF of each line using a Hamming window, before centroids calculation.

3. Calculation of the oversampled ESF manipulating the image lines as previously described.

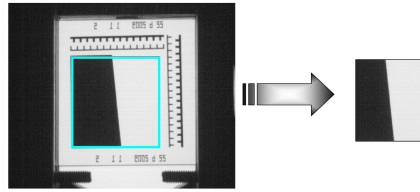4. Calculation of the LSF as derivative of the oversampled ESF.
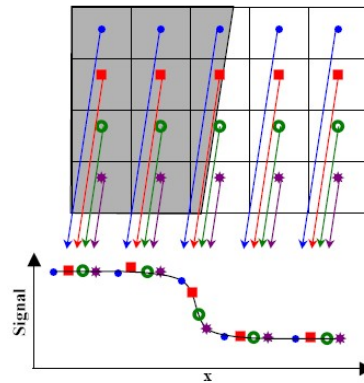
Figure 1.5: Selection of ROI.



Figure 1.6: Projection to put all the detected lines on a single one.

5. Calculation of the SFR as Fourier transform of the LSF.

In SAFARILAB , the edge slope determination is repeated until the angle estimation achieves a precision of $10^{-6}$ degrees.

Furthermore, before Fourier transforming the LSF for calculating the SFR, in SAFARILAB , a zero padding has been added to the LSF. Its size is four times the original ESF length, before oversampling. This operation increase the resolution in the spatial frequency domain.

Two additional options have been implemented in SAFARILAB for reducing the noise that affects the measurements. They have been introduced in the algorithm as shown schematically in red in Fig. 1.7. According to option A, the overall SFR is obtained by averaging a number of SFRs. Each of them is calculated by an acquired image. This allows to increase the measurement robustness. On the other hand, in option B, the SFR comes from the elaboration of an image, average of several images. In this way the measurement uncertainty can be reduced. Sec. 1.3.3 presents the results obtained using these two options individually or together.

In the following, the physical realization of the set-up for using SAFARILAB will be described.
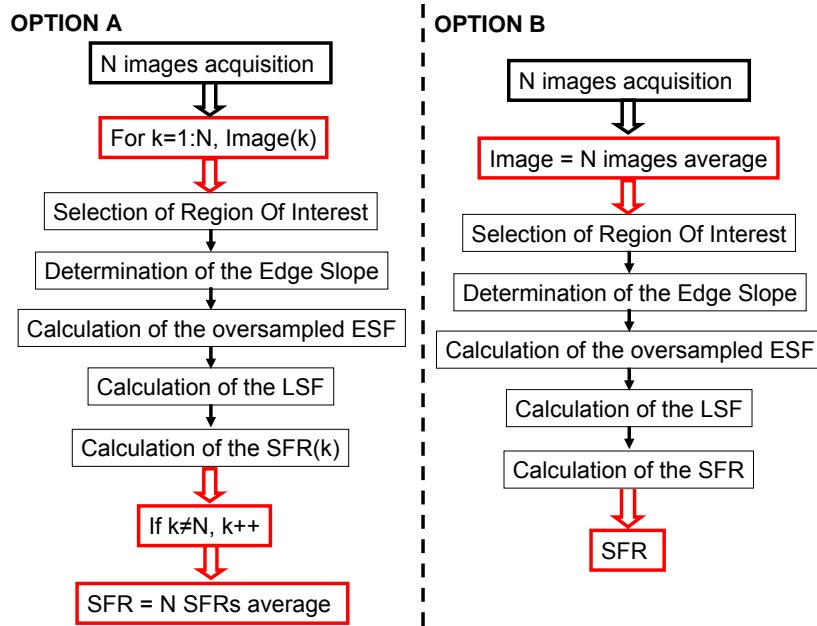
**OPTION A**

```
┌─────────────────────────────┐
│    N images acquisition     │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│    For k=1:N, Image(k)      │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│ Selection of Region Of Interest │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Determination of the Edge Slope │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Calculation of the oversampled ESF │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│    Calculation of the LSF   │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│  Calculation of the SFR(k)  │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│       If k≠N, k++           │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│    SFR = N SFRs average     │
└─────────────────────────────┘
```

**OPTION B**

```
┌─────────────────────────────┐
│    N images acquisition     │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│    Image = N images average │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│ Selection of Region Of Interest │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Determination of the Edge Slope │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Calculation of the oversampled ESF │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│    Calculation of the LSF   │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│   Calculation of the SFR    │
└─────────────────────────────┘
              ⇓
┌─────────────────────────────┐
│            SFR              │
└─────────────────────────────┘
```

Figure 1.7: SAFARILAB options to reduce the noise.

### 1.2.3   SAFARILAB set-up

The SAFARILAB set-up is schematically shown in Fig. 1.8 and consists of an incoherent source (integrating sphere) with a slanted edge mask positioned on its exit port. The lens and the CCD camera are placed in front of it. The CCD camera output is connected to the Personal Computer (PC) making the measurements.

The arrangement of the set-up agrees with the slanted edge measurement technique as described by the reference standard.

The building blocks of the set-up are:

**Source** : The incoherent source is made by a LED emitting at the desired wavelength, to control the input source spectrum, and an "Integrating sphere", providing uniform illumination of the target. Sometimes, for better illumination uniformity, two (or more) integrating spheres can be cascaded between the source (overall input port) and the mask (overall output port). In Fig. 1.9 (a) a picture of the Integrating spheres used in the experiments is shown.

**Mask** : The needed mask is very simple in shape: a slanted edge. In particular, a transmissive mask is used. It is a custom-designed photolithographic mask with very high print quality having feature size as small as $10\mu m$. The target pattern has a usable surface of $25mm \times 25mm$, A picture of the slanted edge mask is shown in Fig. 1.9 (b). It includes alignment lines and features having known dimensions to evaluate the resolution of the image.
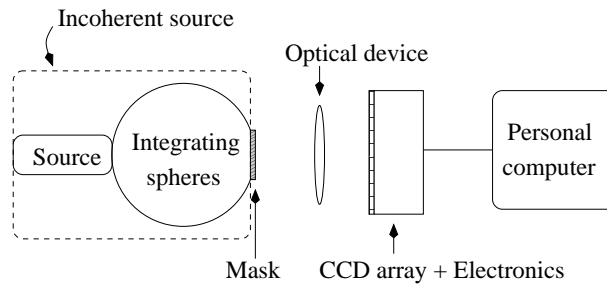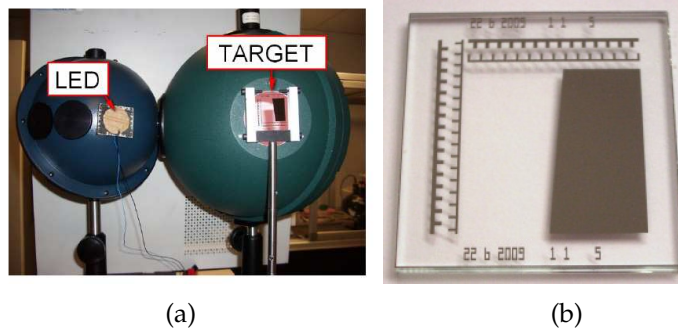
Figure 1.8: Set-up schematic



Figure 1.9: Combination of Integrating spheres used for the measurements (a) and picture of the slanted edge mask (b).

**Lens and Camera** : The lens is the DUT (Device Under Test) of the set-up. The camera can be included in the DUT when the overall SFR response must be determined, including not only optical effects, but also others such as pixel size, electrical noise, etc.

In the following sections the results of the tests done on SAFARILAB using the set-up just described will be presented.

## 1.3 SAFARILAB performance

For checking the SAFARILAB performance, several tests have been done. SAFARILAB has been compared with other available software, firstly elaborating synthetic images, secondly real images. Further tests have been done to characterize the SAFARILAB noise robustness and the measurement repeatability. Finally, the agreement between calculated and measured performance of the optical devices under test has been checked.

### 1.3.1 Results on synthetic images

To check if the SAFARILAB package works correctly, it has been compared with other existing tools available for free (IMAGEJ) or commercially (IMATEST). The first step of this comparison is a test on synthetic images with an analytically known SFR. The generation of a synthetic image requires to chose a function similar to a ESF, with an analytical solutions for its derivative (LSF), and for the Fourier transform of its derivative (SFR). Among the possible functions with an ESF shape, the one chosen to generate the synthetic images is the arctangent since it fulfills all the necessary conditions:

- it has the same qualitative shape of a real slanted edge;

- the analytical formula of the convolution between the function and a rect-angular window (with this operation the finite pixel size influence is simulated) exists (ESF):

$$
\arctan(x) \star \text{rect}\left(\frac{x}{T}\right) = \int_{-\infty}^{\infty} \arctan(x - \tau)\ \text{rect}\left(\frac{\tau}{T}\right)\ d\tau =
$$

$$
= \int_{-\frac{T}{2}}^{\frac{T}{2}} \arctan(x - \tau)\ d\tau =
$$

$$
= -\left[x - \frac{T}{2}\right]\arctan\left[x - \frac{T}{2}\right] + \left[x + \frac{T}{2}\right]\arctan\left[x + \frac{T}{2}\right] +
$$

$$
+ \frac{1}{2}\ln\left[\left(x - \frac{T}{2}\right)^2 + 1\right] - \frac{1}{2}\ln\left[\left(x + \frac{T}{2}\right)^2 + 1\right], \quad (1.10)
$$

where $T$ is the rectangular window size.

- the analytical formula of the convolution derivative exists (LSF):

$$
\frac{d}{dx}\left[\arctan(x) \star \text{rect}\left(\frac{x}{T}\right)\right] = \frac{1}{x^2 + 1} \star \text{rect}\left(\frac{x}{T}\right) \quad (1.11)
$$

- the analytical expression of the Fourier transform $\mathcal{F}$ of (1.11) exists (SFR):

$$
\mathcal{F}\left\{\frac{1}{x^2 + 1} \star \text{rect}\left(\frac{x}{T}\right)\right\} = \pi\, e^{-|2\pi f|} \cdot \text{sinc}(fT). \quad (1.12)
$$

The synthetic images where created using the arctangent function filtered by a rectangular window. To test the different packages in realistic situations, shot noise has been added to the images, having different degrees of contrast. It is simulated changing the dynamic range of the image. Full (0-255) and reduced (30-220) dynamic ranges have been considered.
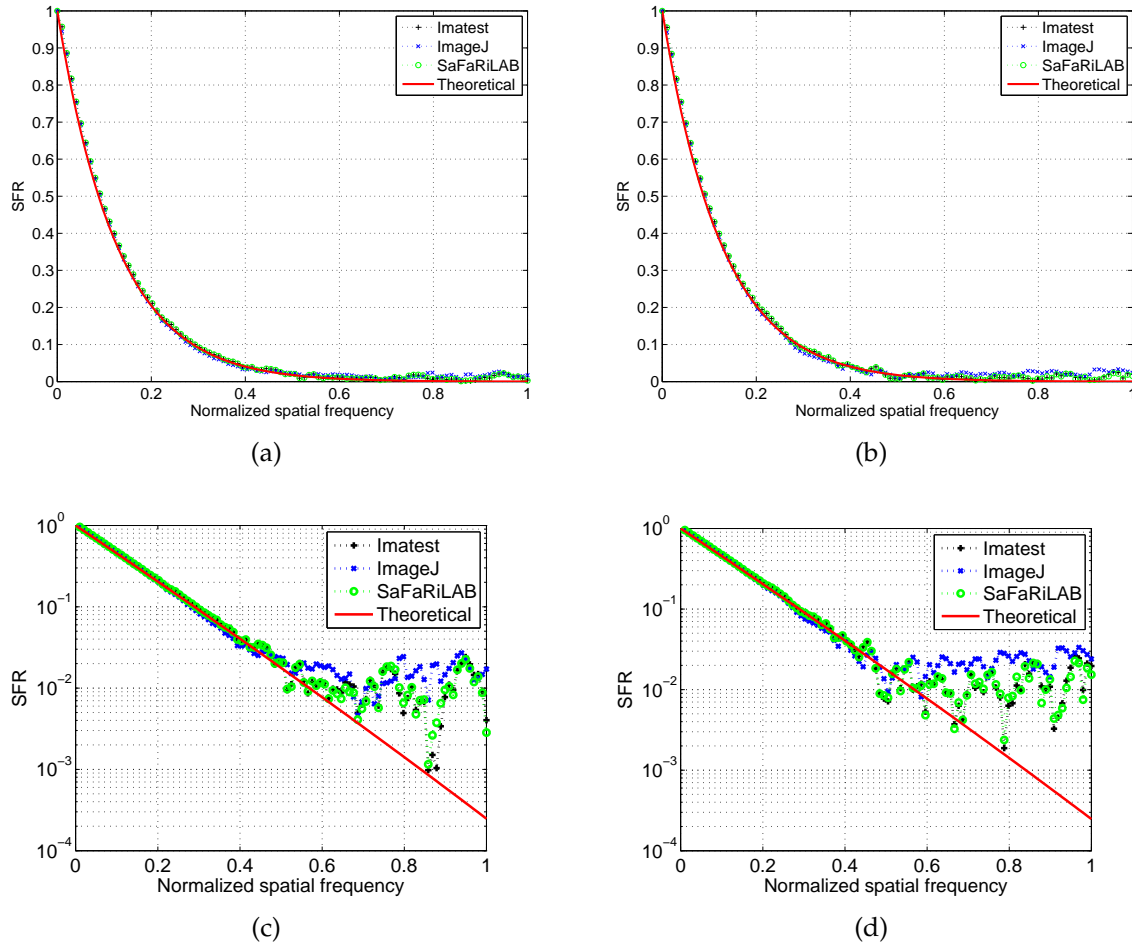
(a) (b)

(c) (d)

Figure 1.10: Test with synthetic noisy images with ranges 0-255 (a) (c) and 30-220 (b) (d). Upper figures with linear scale, lower with logarithmic one.

The results are shown in Fig. 1.10, where the frequency axis is normalized, as in the whole thesis, to the Nyquist frequency. Noise addition produces oscillations, evidenced in figures with logarithmic scale, in the SFRs calculated by all three softwares. Nevertheless, all these SFRs are comparable to the theoretical curve. It can be noted that SAFARILAB (without averaging options) and IMATEST have the same behavior, while IMAGEJ presents different oscillations. This trend will be evidenced also in the experimental tests.

## 1.3.2 Results on real images

Tests were then done on experimentally measured images. The set-up is the one described in Sec. 1.2.3. The camera is made by the optical system to be evaluated, the CCD sensor and the related electronics.

The SFR results obtained testing the softwares with images captured at dis-
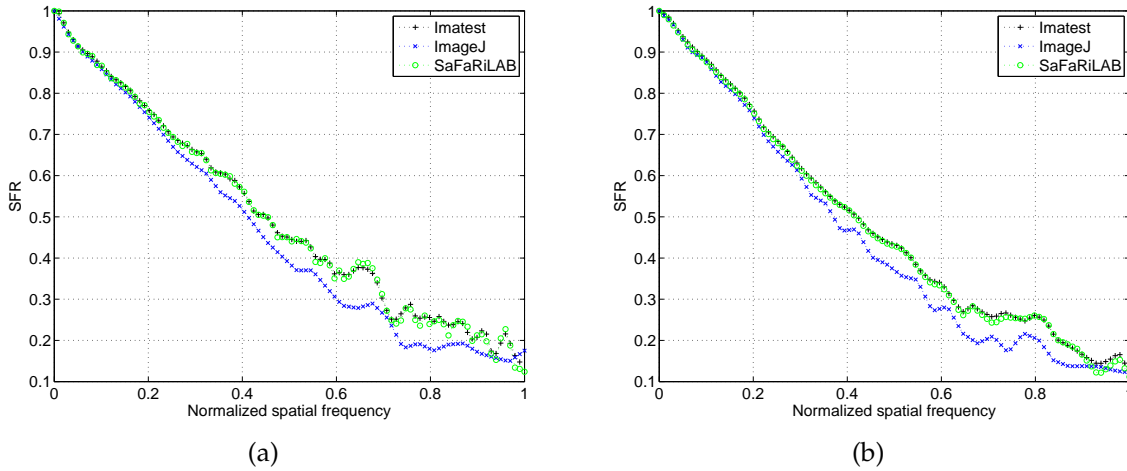
Figure 1.11: SFR obtained with experimental images taken at 115 mm (a) and 190 mm (b) from the target.

tances between the camera and the object of 115 mm and 190 mm are shown respectively in Fig. 1.11 (a) and (b). Results obtained using SAFARILAB , without noise filtering options, and IMATEST are similar, while the IMAGEJ SFR is below the other two curves. Although all the software follow the standard, different implementations led to slightly different results. In spite of this, one can note that SAFARILAB provides results comparable to those of existing packages.

### 1.3.3 Noise test

To evaluate the SAFARILAB noise robustness, several measurements have been done using a camera that allows to control the SNR by setting the gain and the exposure time. The SNR is defined as the difference between the mean value of the intensities detected in an area belonging to the white part of the ROI and the one evaluated in the black part, over the standard deviation of the values of the white region. In the following the results obtained using SAFARILAB with option A or B and A together are shown. These combinations of options are chosen as example, but it is possible to combine them as desired.

During the first test, the camera has been put at 60 mm from the target, setting gain to 1 and the exposure time to 32 ms, and 50 images have been captured. For each of them, the SFR has been calculated. In order to filter the noise, using option A of SAFARILAB , the 50 SFRs have been averaged and the standard deviation $\sigma$ has been calculated.

Fig. 1.12 (a) shows the test results obtained using option A. They demonstrates the noise robustness of this method. The uncertainty range $\pm 3\sigma$ is as small as 0.03.

Fig. 1.12 (b) shows the test results obtained using the cascade of SAFARILAB option B and A. Each of the 50 images have been created as the average of other
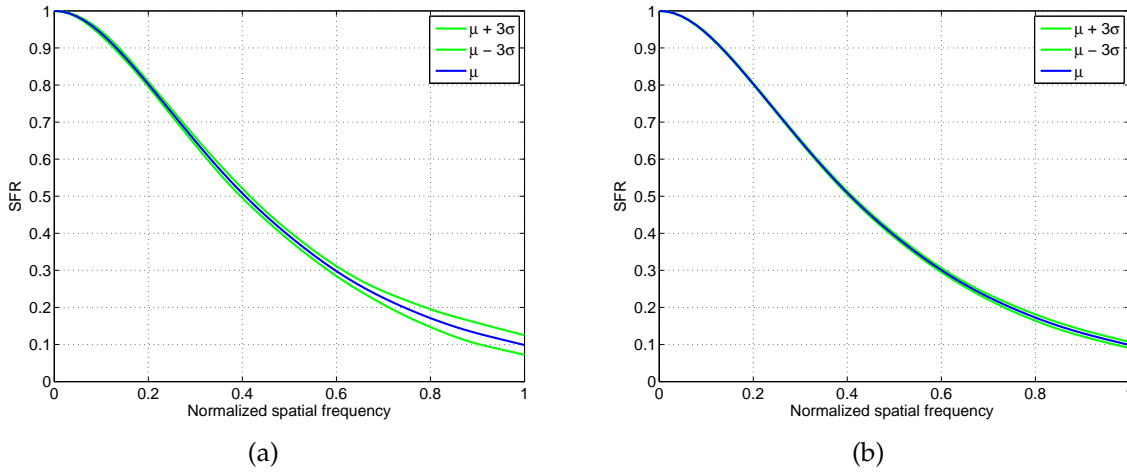
Figure 1.12: First test at 60 mm from the target, gain = 1, exposure time = 32 ms. Results found using SAFARILAB option A (a) and option B + A (b).

16 ones (option B), before calculating each of the 50 SFRs to be averaged (option A). Averaging the images, the uncertainty range is lower than in Fig. 1.12 (a). The maximum value of $3\sigma$ is now 0.009. Thanks to this further filtering operation the uncertainty has been reduced by about $1/3$.

In the second test, the same distance from the target and the same image contrast have been kept. The SNR have been decreased setting the gain to 4 and the exposure time to 8 ms. Fig. 1.13 reports the SFR measured using SAFARILAB option A (a) and option B + A (b). The uncertainty range is almost twice that of the previous test.

In a third test, the detector has been placed at 200 mm from the target with gain set to 1 and exposure time set to 32 ms, the better case between first and second tests. In Fig. 1.14 the results of this test are shown. In this case, the value of $3\sigma$ varies from about 0.03 to 0.02.

In conclusion, the SFR averaging operation (option A) and the image averaging operation (option B) reduce the effects of noise, as expected.

### 1.3.4    Measured vs designed results

An important step in the validation of the developed software is the comparison between the measured and designed results. Since the measured SFR includes not only the effect of the optics but also those of the electronic part of the system, before comparing the results it is important to include these contributions in the designed SFRs (or to deconvolve them from the measured ones).

To describe the system performance, the minimum resolution $R$ has been calculated. $R$ is inversely proportional to the spatial frequency at which the SFR reaches a threshold value (typically under 30% of its maximum depending on
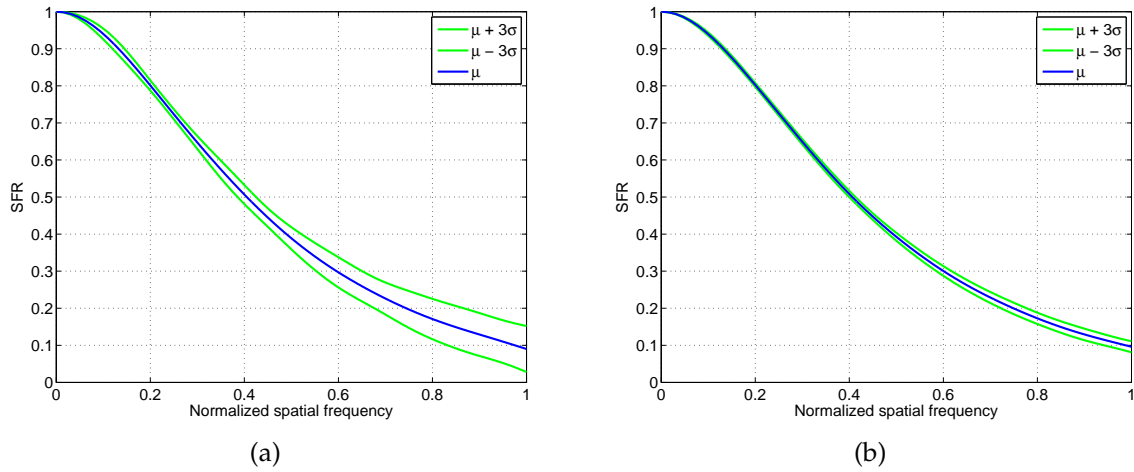
(a)  (b)

Figure 1.13: Second test at 60 mm from the target, gain = 4, exposure time = 8 ms. Results found using SᴀFᴀRɪLAB option A (a) and option B + A (b).
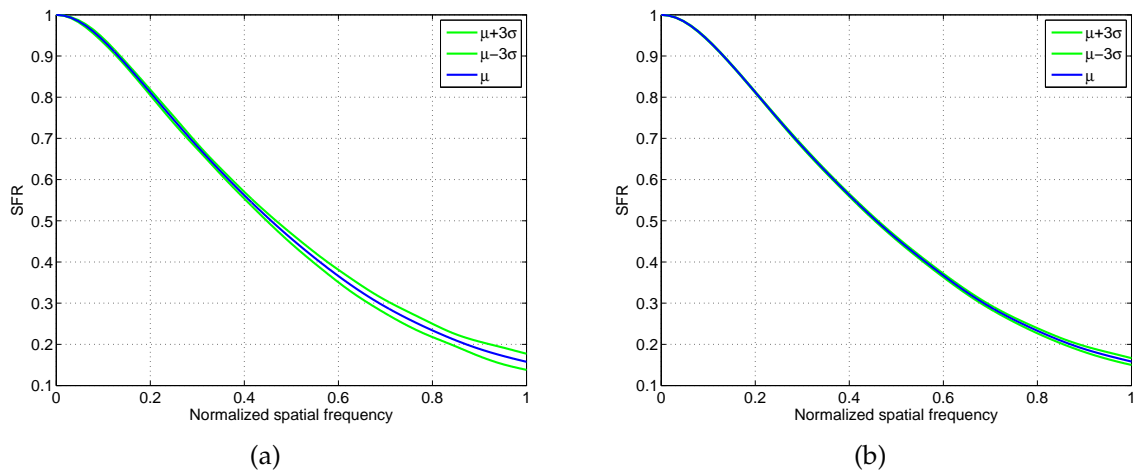


(a)  (b)

Figure 1.14: Third test at 200 mm from the target, gain = 1, exposure time =32 ms. Results found using SᴀFᴀRɪLAB option A (a) and option B + A (b).

system requirements). The minimum resolution is found as a function of the object distance according to:

$$R = \frac{1}{2\, f_{th}\, M} \qquad (1.13)$$

where $f_{th}$ is the limit frequency to obtain the threshold value of the SFR and M is the system magnification depending on the object distance.

The overall SFR of a known system (a lens triplet) has been measured for a range of camera-target distances of $200\mathrm{mm}$. The MTF of the lens is calculated for the same distances using ZEMAX® . Finally, the effect of the detector finite pixel size is included multiplying the MTFs of the lens by the footprint MTF (Eq. (1.7)).

For each SFR, MTF (optics) and MTF (optics with sensor), the resolution $R$ is calculated for a defined threshold. Figure 1.15 shows the resolution expressed in mils ($1\mathrm{mils} = 0.0254\mathrm{mm}$) as a function of the camera-target distance. The measured resolution of the camera (red curve) is compared to that of the lens (cyan curve) and to that of the system composed by the lens and the sensor (blue curve). The measured resolution is calculated by averaging 20 SFRs, using SAFARILAB option A, and the relevant uncertainty bars are also reported. The curve that represents the lens and the sensor system is just within the error bars of the measured resolution curve, thus confirming the quality of the results.
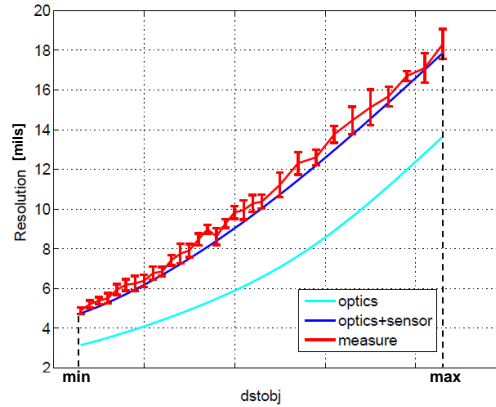


Figure 1.15: Comparison between performances of real system, only lens and lens plus sensor.

## 1.4  Conclusion

After having defined the system performance parameter and how to measure it, the realization of the tool named SAFARILAB has been described. It evaluates the Spatial Frequency Response of an optical system complying with the ISO 12233 standard, the reference standard for this kind of measurements.

Its performance has been successfully compared with those of other available software dedicated to this task. An experimental set-up has also been realized to perform the measurements and the results show excellent behavior in terms of repeatability and ability to filter out random noise effects, thanks to the options added to SAFARILAB .

Finally, the tool has been used to compare the measured optical SFR and that designed with ZEMAX® . This verification is an essential step to validate the whole optical system design process.

In the next chapters SAFARILAB will be used extensively for the characterization of optical imaging system performance.

# 2

# Imaging System Design: Optical level

In this chapter, a tool, named Optical Level OptiMization (OLOM ) tool, will be described. It works only at the optical level of the system, ignoring the effects of the other parts. This tool, working together with the ZEMAX® optimization engine, allows to set constraints on the MTFs following required specifications.

For testing and validating the whole design tool, a Cubic Phase Mask (CPM) system has been used. It is an Extended Depth Of Focus (EDoF) system, suitable for Auto ID applications. The reasons of this choice will be presented in Sec. 2.1, after a literature review.

After the detailed descripion of OLOM tool (Sec. 2.2), in Sec. 2.3 the design of the system with CPM will be described.

Finally the realization (Sec. 2.4) and characterization (Sec. 2.4) of the designed CPM system will be presented. The performance of the imaging system will be calculated using SAFARILAB (previously described in Ch. 1), in order to demonstrate the agreement between the measured and the expected results and to evaluate the contributions given by OLOM .

## 2.1 Imaging system choice

A common way to extend the DoF of a system is the autofocus method. Autofocus systems are provided by a motor which automatically adjusts the focus. Even if these systems have the potential to extend the DoF, the presence of mobile parts and the low response speed are not suitable for Auto ID and portable applications.

For this reason, solutions for EDoF without mobile parts proposed in the literature, have been considered. In the following advantages and disadvantages of these solutions will be presented, analyzing which of them could be not only suitable for extending an imaging system DoF, but also for testing and validating

OLOM tool.

## 2.1.1   Special Shape Lenses

Solutions with *Special Shape Lenses* have been considered first. Conceptually, a way to have an EDoF is to assemble many lenses in a single one, making adjacent radial regions with different focal lengths. Each lens has a different focal distance. At that particular distance all the other lenses produce defocused images, with no or negligible effect on the final result. So the final DoF comes from the total contribution of the various lenses. The simplest case of these lenses is the one with conic shape, the so called *axicon* [15, 16] or *linear axicon* or *Annular Linear Axicon*, ALA. If the focal length is set controlling the lens shape not along the radius but along the azimuth angle, the so called *Light Sword Optical Element* (LSOE) can be generated [17–20].

Comparisons among axicons and LSOEs show that imaging systems based on axicons transfer a wide range of nonzero spatial frequencies with a small contrast. LSOEs, on the contrary, allow a much larger amount of the incident energy of light in the main maximum, providing images with better contrast. Though the central nonzero domain of the MTF of LSOEs is smaller than the axicon one, its values are higher within this range.

One can conclude that axicons perform better if digital processing to restore the final image is allowed, while LSOE look better for real-time imaging with extended depth of focus. On the other hand, the DoF extension demonstrated by both of them is in the order of the centimeters, which may not be enough for the devices developed in this work.

Another type of Special Shape Lenses is the *Photon Sieve*, a modified Fresnel Lens with holes rather than annular sections as apertures. It was proposed in [21] to focus soft x-rays for imaging. This lens is not suitable for imaging system, in which it is important to maximize the system light transmission.

## 2.1.2   Lens Multiplexing

Secondly, solutions based on *Lens Multiplexing* have been considered. They are based on the idea of mixing different lenses to increase the overall DoF and have been realized in different ways.

In [22], Iemmi and coworkers proposed to combine several diffractive lenses with different focal lengths spatially multiplexing them in a random scheme onto the single final lens. After designing a number of lenses with different focal lengths and having partitioned them in subparts, the final lens was composed assigning to each subpart the corresponding one of a randomly chosen lens of the original set, to reduce transversal sidelobes of the lens PSF [23]. EDoF comes from the combined effect of many different elementary lenses on the axial irradiance distributions.

In Furlan [24], the multiplexing approach is used to modify a Fresnel lens into a Fractal Zone Plate (FraZP) in order to reduce the high chromatic aberration that affects Fresnel lens systems. In presence of polychromatic light in the visible range, the FraZP produces a sequence of subsidiary foci around each major focus following the fractal structure of the FraZP itself. These subsidiary foci provide an extended depth of focus for each wavelength that partially overlaps with the other ones, creating an overall extended depth of focus that is less sensitive to chromatic aberration. However in imaging applications, it is important to maximize transmission of the mask, which makes zone plates not suitable.

The idea of spatially multiplexing lenses led also to the so called *Composite Phase Masks* (CoPM). Many small Fresnel Lenses with slightly different focal lengths are superimposed to a larger Fresnel lens and locally correct the focal distance of the larger Fresnel lens. The CoPM were developed by the group of Zalevsky [25, 26], by physically mixing many lenses into one. CoPM lenses appear complicated to be implemented in a general environment, where cost and robustness are primary issues. Moreover, there are no published results concerning the comparison between experimental and simulated data, which would help in understanding the advantage of this implementation with respect to possible different ones.

### 2.1.3   Wavefront Coding

The two approaches illustrated so far operate on the phase front of the optical beam so that the detected image has a spot which does not change for a large excursion in the longitudinal direction. A quite different approach can be however followed. As anticipated in Sec. 1.1, if the whole system can be considered linear and can then be described by its transfer function, it is possible to integrate the design of the optical part with that of the final transducer. System linearity allows to split the overall system transfer function in subparts the product of which is always the desired one. So, if one deliberately introduces a known distortion in the phase front of the incoming beam, such a distortion can be removed with a suitable filter later on. The advantage of this way of working is that, if this distortion is larger than that due to the expected diffraction, a single digital filter, adapted to the introduced predistortion, can be used to deconvolve the image and reconstruct it. Such a procedure is known as *"Wavefront Coding"* (WFC).

The system sketch shown in Fig. 2.1 evidences the predistortion Phase Mask in the Optical Subsystem block and the inverse filter implemented in the signal processing block of the final electronic subsystem.

In [3], the shape of a Phase Mask that satisfies an EDoF condition, has been found from the study of the Ambiguity Function [27–29], seen as the system OTF developed in the spatial frequency domain [3, 30–32].

The phase profile obtained applying EDoF constraints and the stationary phase
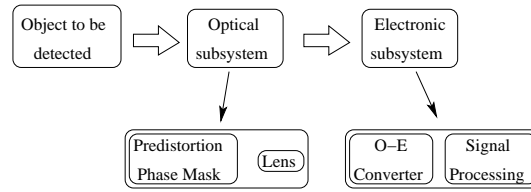
Figure 2.1: System sketch when Wavefront coding is considered.
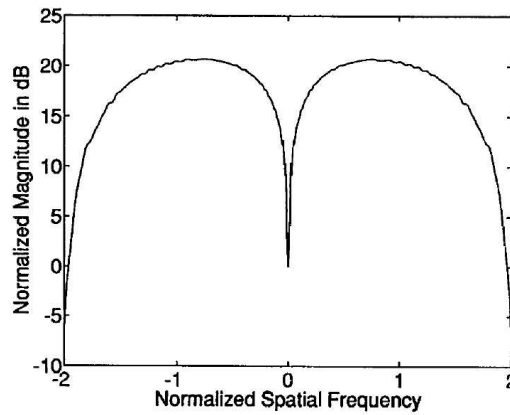


Figure 2.2: Transfer function of a digital filter to deconvolve CPM distortion effects. From [3].

approximation to the Ambiguity Function in the spatial frequency domain is cubic [3], and it can be described by the equation:

$$f(x,y) = \alpha(x^3 + y^3), \tag{2.1}$$

where $\alpha$ is the Cubic Phase Mask (CPM) parameter to be optimized.

The characteristics of the filter used in the digital elaboration of the acquired image to deconvolve the effects of the CPM come from a least square optimization of various PSF calculated for different defocus values. An example of the filter is shown in Fig. 2.2. It is smooth and takes advantage of the lack of zeroes in the OTF in the spectral region of interest. It has unitary value in the origin and amplifies components at higher spatial frequencies to improve the contrast of the reconstructed image.

Recently there has been renewed interest in the analytical study of the performance of an optical system with a CPM. A simplified expression has been proposed [33] for the MTF of a diffraction-limited optical system having a CPM and subject only to defocus. The approximated expression of MTF is employed to find the optimal configuration for imaging tasks [34]: the optimal cubic coefficient and the image plane distance are expressed in simple closed forms as functions of the considered distance range, the aperture size and the focal length of the diffraction-limited optical system.

It is also possible, applying the stationary phase method in the spatial co-ordinate system, to find another solution in an analytical way, stating that the phase mask has in this case a logarithmic distribution (Logarithmic Phase Mask, LPM [35, 36]).

Furthermore, after the seminal work published in [37], Robinson and coworkers [4, 38, 39] have proposed the introduction of a quartic contribution in a system to achieve EDoF (Quartic Phase Mask, QPM). They have also discussed the problem of using a joint optimization of the optical and electronic parts of the system to obtain a better design.

Other different kinds of Phase Masks have been proposed in the last years. Their common feature is that they come heuristically from the optimization of some parameters of particular phase distributions (exponential [40], polynomial [41], rational [42, 43], freeform [44], cubic sinusoidal [45], sinusoidal [46] etc. ). While, for CPM and LPM, the function shape was the result of a theoretical approach, in the other cases the phase function shape is *a-priori* chosen and parameters are optimized according to some criteria.

## 2.1.4 EDoF solutions choice

On the whole, WFC based systems appears to have the best performance of the optical system and the best overall EDoF potential (by including electronic post processing).

Among all the proposed phase masks for WFC, the ones with cubic (CPM) and quartic (QPM) profiles ($f(\rho, \theta) = \alpha \rho^4$) has been deeply studied. The advantages and disadvantages of using these EDoF solutions for validating the tools that provide an optimization at optical level (OLOM ) and at system level (SLALOM , Ch. 3) have been investigated. They are presented in the following.

First of all, the way CPM and QPM are introduced into an optical system is different. The cubic one requires the use of a phase mask. On the other hand, a quartic contribution can be generated through an optimized use of the spherical aberration, a fourth order aberration and an intrinsic feature of any optical system. This results in having a simpler realization.

Secondly, the behavior of the MTF in the two cases presents some differences. Fig. 2.3 shows the MTF with respect to the object distance $z_o$ considering only a simple paraxial lens. Without aberrations, the MTF is very sensitive to the object position and halves the optical bandwidth at $20$mm out of focus ($z_o = 60$mm).

With the introduction of the two aberrations, the MTFs do not depend on longitudinal distance of the original optical system, though in a different way. When cubic aberration is present, the MTF has a large degree of invariance over the whole distance range ($z_o$ from $60$ to $150$mm), but also shows lower values. The cubic aberration results in non-rotationally symmetric MTF: at $45°$ the MTF is on average half of that at $0°$. The MTF of the system with quartic (spherical)
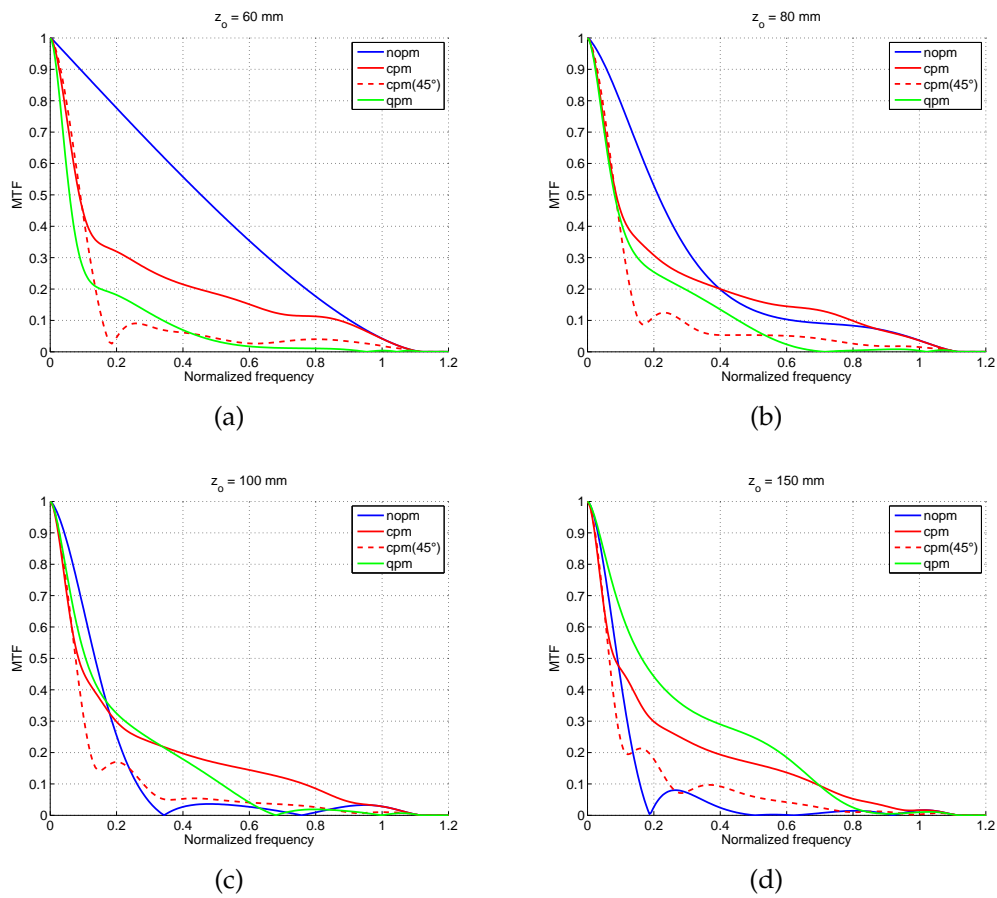
Figure 2.3: MTF of the optical system without phase mask (blue curves), with cubic mask (red curves) and with quartic phase mask (green curves) at object distance $z_o$ (a) 60mm, (b) 80mm, (c) 100mm, (d) 150mm.

aberration lies between the MTF of the system with cubic aberration at $0°$ and that at $45°$. Also, it has smaller bandwidth and less invariance with object distance.

Fig. 2.4 shows the PSF of the three systems at object distance $z_o = 150$mm. It is chosen this distance because it is the worst case for the system without phase mask and it is more visible the contribution given by CPM and QPM. Using them the system behavior doesn't change for different distances. Note the system with cubic aberration has a PSF with large asymmetry.
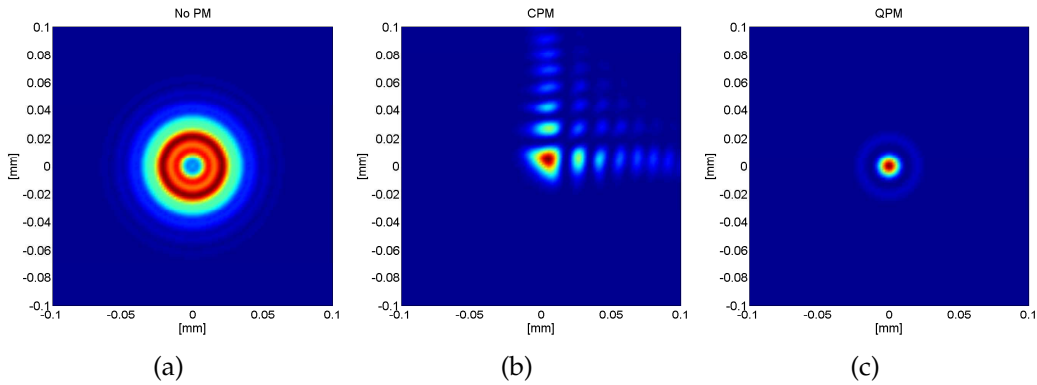


Figure 2.4: PSF of the optical system at $z_o = 150$mm: (a) without phase mask, (b) with cubic mask and (c) with quartic mask.

Fig. 2.5 shows the images created by the three optical system evaluated as the convolution of the image in Fig. 2.5(a) and the PSF of Fig. 2.4. In the case of cubic aberration, the details along the lobe directions of the PSF are lost, whereas, for the other two systems, the images appear very similar.

One can conclude that both the cubic and quartic (spherical) aberrations lead to MTFs which can satisfy the criterion of being longitudinally invariant.

For CPM, the post processing is a critical part, because it needs to create an ad-hoc filter as shown in Fig. 2.2, while for QPM it is possible to use conventional image enhancing techniques, because its PSF is similar to the one of a lens without masks.

In conclusion, the CPM solution has been chosen for testing and validating the OLOM tool, because it is the type of mask that has the most solid theoretical bases, resulting from calculations based on the Ambiguity Function. Furthermore, thanks to Bagheri studies, it is possible to perform very fast calculations, using analytical formulas and not digital elaborations. OLOM tool has been used for optimizing the CPM $\alpha$ parameter to achieve EDoF required specifications, ignoring the post-processing phase.

On the other hand, spherical aberration (or QPM) EDoF solution has been chosen for testing and validating SLALOM tool, because, thanks to its circularly symmetric PSF, it allows the use of a conventional post-processing. Therefore, it is possible to be concentrated only on the optimization tool implemented at
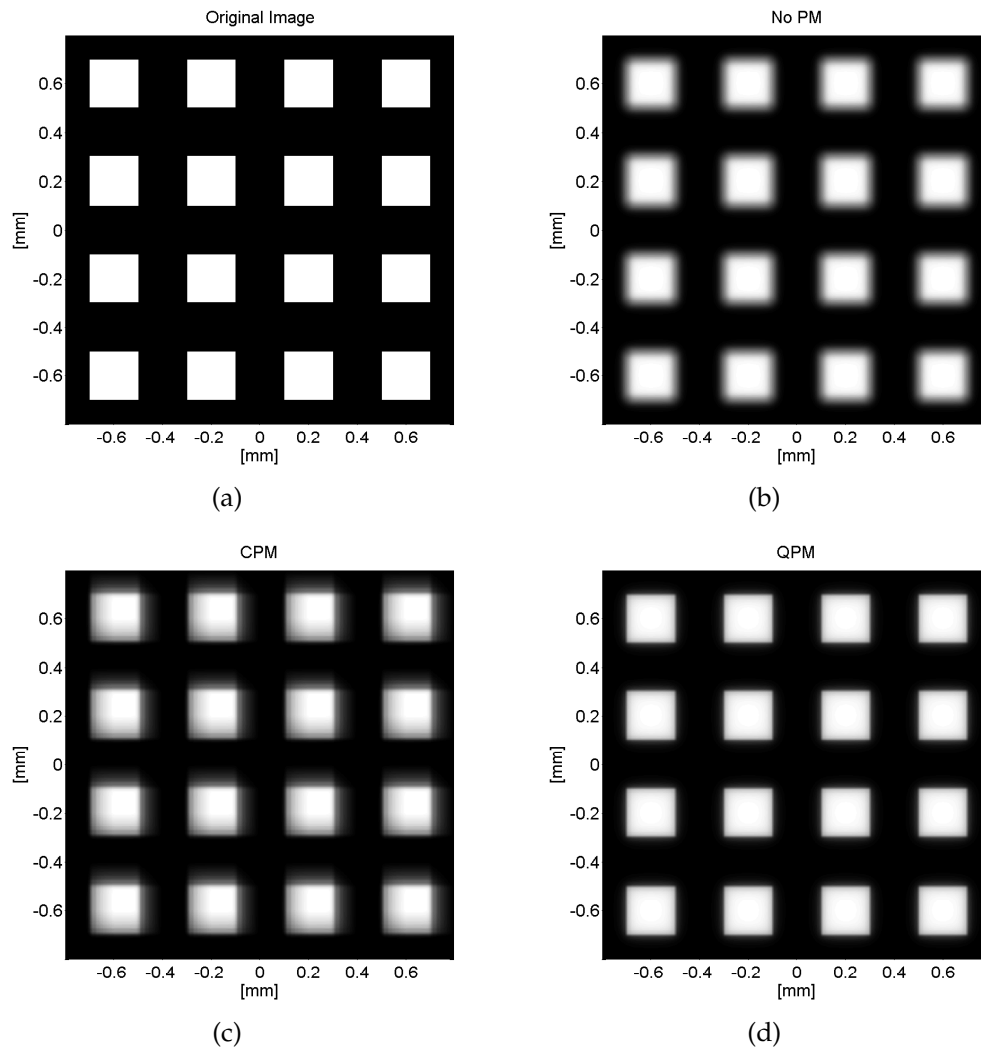
Figure 2.5: (a) Original image and images created by the optical system (b) without phase mask, (c) with cubic mask and (d) with quartic mask at $z_o = 150$mm.

system level, using a conventional filter in the post-processing and not on the construction of an ad-hoc one.

In the next sections, OLOM will be described and its application to a CPM system will be presented, while the optimization of a spherical aberration system using SLALOM will be illustrated in Ch. 3.

## 2.2 Optical Level OptiMization (OLOM ) tool

In this section, the features of the OLOM tool will be detailed. It is an algorithm able to optimize the starting design at an optical level, using the optimization engine of ZEMAX® . Optical level means that optimization is just of the optical part, ignoring the other parts of the system (detector, noise, post-processing etc.).

OLOM allows to integrate the constraints typical of wavefront coding systems into the Merit Function (MF) of ZEMAX® [36].

As said in Sec. 2.1.3, the advantage of WFC systems is that the OTFs (and consequently the MTFs) do not depend on the defocus parameter. This means that the MTFs at different distances are invariant.

The optimization, focused on increasing the extension of DoF made by the invariant MTFs, is divided in two phases. Firstly, the difference between the in-focus MTF and all the MTFs at the defocused planes taken into account are minimized. This allows to have the MTFs as invariant as possible on the desired planes. Secondly, in order to have the in-focus MTF as high as possible, the difference between it and the in-focus diffraction limited function is minimized. Therefore the final result of the optimization should be a system with the highest value of the MTF as invariant with defocus as possible. The two minimizations can be defined as follows:

$$\underset{p}{\arg\min} \ \left( \text{MTF}_{focus}(f) - \text{MTF}_{defocus}^{(k)}(f) \right), \tag{2.2}$$

$$\underset{p}{\arg\min} \ \left( \text{MTF}_{diff-lim}(f) - \text{MTF}_{focus}(f) \right), \tag{2.3}$$

where $p$ is the group of parameters to be optimized, $\text{MTF}_{focus}$ is the in-focus MTF, $\text{MTF}_{defocus}^{(k)}$ is one of the $K$ out of focus MTFs considered with $k = 1...K$ and $\text{MTF}_{diff-lim}$ is the diffraction limited MTF evaluated on the in-focus plane. It can be calculated as [13]:

$$\text{MTF}_{diff-lim}(f) = \frac{2}{\pi} \left( \cos^{-1} \left( \frac{f}{f_{\text{cutoff}}} \right) - \left( \frac{f}{f_{\text{cutoff}}} \right) \sqrt{1 - \left( \frac{f}{f_{\text{cutoff}}} \right)^2} \right), \tag{2.4}$$

where $f_{\text{cutoff}} = 1/(\lambda \cdot (F/\#))$ is the cutoff frequency, product of the chosen wavelength and of the system F-number.

The optimization process in ZEMAX® works minimizing the MF, a numerical representation of how closely an optical system meets a specified set of goals. For setting a MF in ZEMAX® , it is possible to use a list of operands which individually represent different constraints or goals for the system. The construction of the MF is the most critical step to achieve the required specifications. Once the merit function is complete, the optimization algorithm in ZEMAX® will attempt to make its value as small as possible.

The MF used by OLOM follows the criteria explained by Eqs. (2.2) and (2.3), calculated for different values of four parameters: frequency, field, wavelength and object distance.

In table 2.1, an example of one MTF constraint construction referring to the Eq. (2.2) is shown. For example Wave=1 or Field=1 mean that it is considered the first wavelength or the first field defined by the user. The steps shown in table

| #: Oper | Cfg # | - | - | - | - | - |
|---------|-------|-------|------|--------|--------|-------|
| 1: CONF | 1 | - | - | - | - | - |
| #: Oper | Wave | Field | Freq | - | - | Value |
| 2: MTFA | 1 | 1 | $f_1$ | - | - | 0.9 |
| #: Oper | Cfg # | - | - | - | - | - |
| 3: CONF | 2 | - | - | - | - | - |
| #: Oper | Wave | Field | Freq | - | - | Value |
| 4: MTFA | 1 | 1 | $f_1$ | - | - | 0.8 |
| #: Oper | OP#1 | OP#2 | - | Target | Weight | Value |
| 5: DIFF | 2 | 4 | - | 0 | 1 | 0.1 |

Table 2.1: Example of a MTF constraint referring to the Eq. (2.2).

are:

1. set of the configuration 1 (it represents the in-focus plane);

2. extraction of the MTF value (0.9 in the example) for the first field, the first wavelength and the frequency $f_1$;

3. set of the configuration 2 (it represents one of the defocused planes);

4. extraction of the MTF value (0.8 in the example) for the same field, wavelength and frequency used before;

5. difference of the two MTF values identified by the operand number. The target value for this difference is equal to 0 and the weight is set to 1.

In the example of table 2.1 the `MTFA` operand is used: it returns the average value between the tangential MTF and the sagittal MTF one for the specified field,

wavelength and frequency. It is also possible to directly use `MTFT` and `MTFS` to extract the tangential and sagittal MTF separately.

The implementation of Eq. (2.3) is similar to the previously described one. The only difference is that the MTF values are evaluated on the same configuration corresponding to the in-focus plane. The diffraction limited MTF, also named geometrical MTF, can be asked to ZEMAX® using the operand `GMTA`. However, for simplifying the implementation, it has been calculated using Eq. (2.4) directly in MATLAB® during the MF generation. An example of the MTF constraints referring to Eq. (2.4) is written in table 2.2, considering the first configuration as the one referring to the in-focus plane.

| #: Oper | Cfg # | - | - | - | - | - |
|---------|-------|-------|-------|--------|--------|-------|
| 1: CONF | 1 | - | - | - | - | - |
| #: Oper | Wave | Field | Freq | Target | Weight | Value |
| 2: MTFA | 1 | 1 | $f_1$ | 0.93 | 1 | 0.9 |

Table 2.2: Example of a MTF constraint referring to the Eq. (2.3).

Since the combinations of the considered variables are usually a lot, the MF is automatically generated by a MATLAB® script and then loaded in the ZEMAX® MF editor. In particular, wavelengths and fields are directly set in ZEMAX®, whereas frequences and object distances depend on the specifications required by the considered application.

In the following a possible example is presented, considering a particular Auto ID application that requires EDoF: the barcode reader system.

For barcode identification applications, the goal is to correctly read the different types of barcodes when they are positioned in a specific range of distances from the reader. The barcode types are characterized by their resolution, i.e. the minimum module size (thinnest bar, module, for 1D barcode as shown in Fig. 2.6 (a) and smallest element for 2D barcode), expressed in mils.

A barcode located at a distance $z$ from the imaging system is readable if the MTF calculated at $f_{res}$ is greater than a specific threshold, usually set under $30\%$ (see Fig. 2.6 (b)). The frequency $f_{res}$ is calculated as:

$$f_{res} = \frac{1}{2r_{\mathrm{mm}}} \cdot \frac{1}{M(z)},$$ (2.5)

where $r_{\mathrm{mm}}$ is the spatial resolution in $\mathrm{mm}$ and $M$ is the system magnification, dependent on the distance $z$.

An example of specifications is shown in table 2.3. In the first column there are the barcode resolutions in mils and in the others the minimum and maximum distances that specify the desired readable range.

The frequencies and the planes to be used in the optimization are the minimum and maximum distances and the $f_{res}$ frequencies.
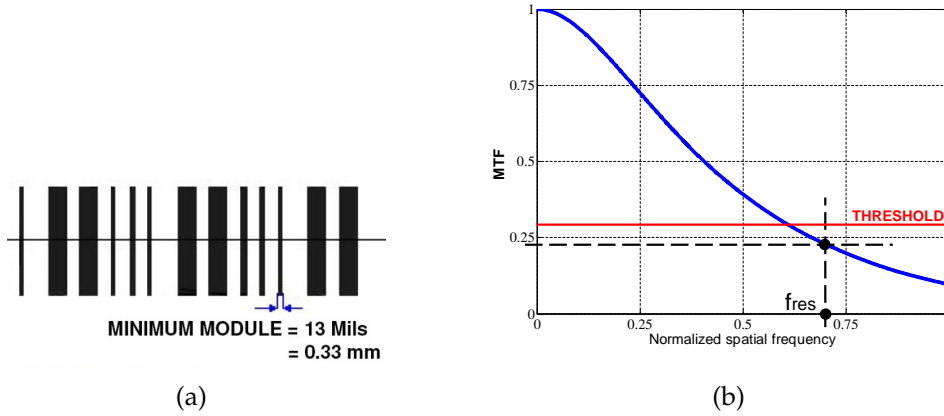
        (a)                                            (b)

Figure 2.6: (a) Example of the minimum module size of a barcode. (b) Representation of an MTF specification.

| Resolution [mils] | Dist. min [mm] | Dist. max [mm] |
|:---:|:---:|:---:|
| ... | ... | ... |

Table 2.3: Example of a specifications table.

In the next section, the use of OLOM will be demonstrated, considering the optimization of a CPM system.

## 2.3   CPM system Design using OLOM

In Sec. 2.1 different types of phase masks, each designed to allow EDoF, have been reviewed. After the literature analysis, the CPM solution for EDoF has been chosen. It is the type of mask that has the most solid theoretical bases, resulting from the optimization of the MTF.

The CPM system to be designed is made by joining a CPM to an existing lens, referred to as original lens, previously optimized. The goal is to validate the OLOM tool for the optimization of the CPM, without starting from scratch designing a totally new system.

Before the optimization of the system, the best position to put the CPM with respect to the original lens has been chosen.

According to the theory, the CPM is to be placed at the exit pupil plane of the lens. The exit pupil is a virtual aperture image of the aperture stop and only rays which pass through it can exit the system.

In the original lens, the exit pupil plane is before the first surface and is therefore a virtual plane. Note, however, that in such lens the first surface is the stop. Since the exit pupil and the stop plane are conjugated planes, their wavefront
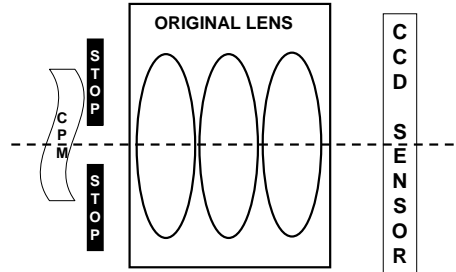
Figure 2.7: Scheme of the system including the CPM.

| Resolution [mils] | Dist. Min [mm] | Dist. Max [mm] |
|:---:|:---:|:---:|
| 13 | 20 | 500 |
| 20 | 20 | 780 |
| 40 | 20 | 1500 |

Table 2.4: Specifications on the required resolution capability of the optical system at the object plane.

aberrations are proportional. Thanks to this relationship it is possible to obtain the required phase distortion on the exit pupil placing a scaled phase mask directly on the stop plane.

After having chosen to place the phase mask at the stop of the original lens, the system to be optimized results as schematically shown in Fig. 2.7. In the following, the optimization of this system will be presented.

### 2.3.1 Optimization

The variables to be optimized are: $\alpha$, the only CPM parameter (cubic phase: $f(x,y) = \alpha(x^3 + y^3)$), and Back Focal Length (BFL), the distance between the last original lens surface and the sensor plane.

The required specifications for this CPM system are expressed in terms of the desired DoF extension for a barcode reader, used as an example of application.

A part of the system requirements for low resolutions are listed in table 2.4. The distances listed in the table are between the object position and the nose of the reader. This is a convention used for the barcode readers.

Therefore, the optimization process looks for the value of $\alpha$ and BFL that can guarantee the performance described in table 2.4. The required MTF thresholds are not listed in the table because they are specific for the type of application.

In order to construct the MF, OLOM described in 2.2 has been used. It allows on one hand to minimize the difference between the in-focus MTF and the defo-

cus ones. On the other hand also the difference between the in-focus MTF and the corresponding diffraction limited is minimized.

Since the number of constraints depends on the frequencies where these differences are evaluated, it is important to choose the lowest possible number of spatial frequencies in order to minimize the CPU time.

To determine these spatial frequencies, the data shown in table 2.4 have been used. Frequencies have been calculated according to Eq. 2.5, reported here for simplicity:

$$f_{res} = \frac{1}{2r_{mm}} \cdot \frac{1}{M(z)}.$$

$M$ is the system magnification, expressed as function of $z$ and calculated as:

$$\frac{1}{M(z)} = \gamma \cdot z + \delta,$$

being $\gamma$ and $\delta$ coefficients obtained from a linear regression on about ten couples $[1/M, z]$ resulting from the original lens design.

This led to $6$ spatial frequencies, one for each distance and resolution. These frequencies can be anyway reduced to $4$ in the optimization process. This was done simply considering only the more stringent requirement and the relevant spatial frequency when, at the same distance, different resolutions are required.

The total number of MTF to be evaluated at these frequencies is given by the product of:

- the number of the considered object distances (including also the focal distance of the original lens);

- the number of the design wavelengths;

- the number of the fields considered to evaluate the system performance.

Finally, the tangential and sagittal MTFs have been considered separately.

The optimization led to $\alpha = -0.035$ as the optimal value of the CPM coefficient and $BFL = 4.571$mm . In the following, the results obtained thanks to this optimization will be shown.

## 2.3.2   Optimization results

In this section, the performance of the CPM system is compared to the one of the original lens. This allows to evaluate not only the ability of CPM to extend the DoF, but also to validate the implemented optimization.

The figures, in the following, show the range, minimum and maximum distances in mm, where a given barcode, represented by its resolution in mils, can be successfully read (i.e. when the MTF threshold is achieved).
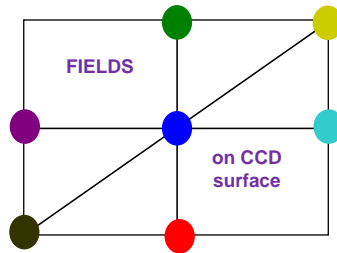
Figure 2.8: Fields on the CCD surface.

Fig. 2.9 refers to the tangential field, while Fig. 2.10 refers to the sagittal field. Fig. 2.11 summarizes these results, presenting the overall lens performance, i.e. the worst case between the tangential and the sagittal ones. In all the plots, for each barcode resolution ($13, 20$ and $40$mils), 7 results are reported. They are labeled by the relevant FoV: the central one, the two vertical ones, the two horizontal ones and the two diagonal ones. They are shown in the label that represents the fields on the CCD surface as scketched in Fig. 2.8.

As one can see, for the lens without the CPM, results are the same for positive and negative values of the field because the original lens has a symmetric PSF. They differ when the CPM is included because the system PSF becomes asymmetric.

Even if not all the specifications listed in Tab. 2.4 are achieved, the CPM optimized using OLOM tool shows an improvement in terms of DoF especially for the central field and for low resolutions ($40$mils) and high resolutions, which have been analyzed but are not reported here.

In the next sections, the realization of this CPM system and its characterization will be shown.

## 2.4 CPM system realization

After an analysis of the production capability of phase mask suppliers, the Nanocomp company have been chosen.

The process used for the CPM fabrication is direct etching to glass.

After choosing the fabrication process, it is important to verify its requirements and limitations. In this case they are the bulk thickness and the total height of the cubic shape. The CPM design is able to satisfy both of them.

Moreover, Nanocomp fabrication has a tolerance on the etch depth: they declare $\pm300$nm at the depth of $3\mu$m, scaling proportionally to the depth. Therefore, it is important to do a tolerance analysis of the phase mask.
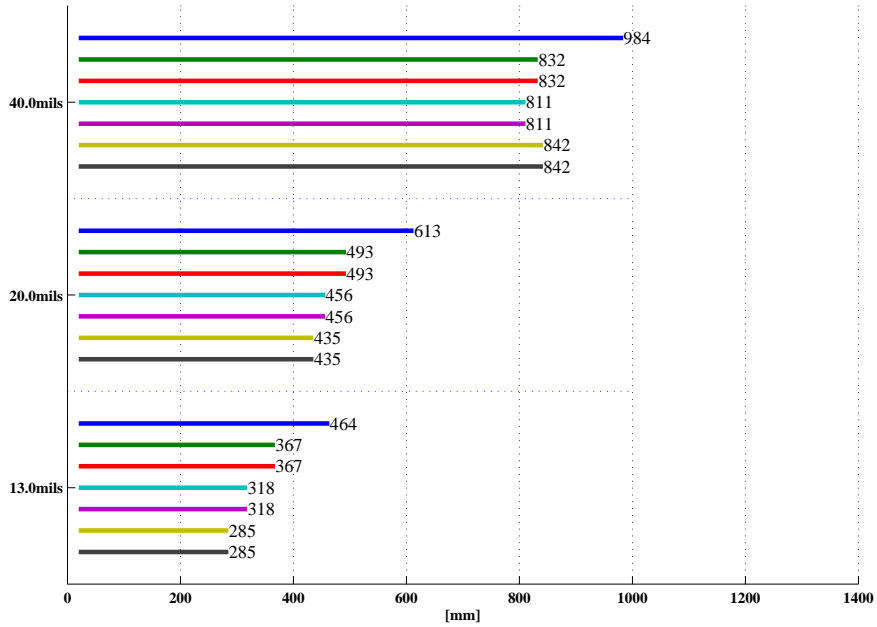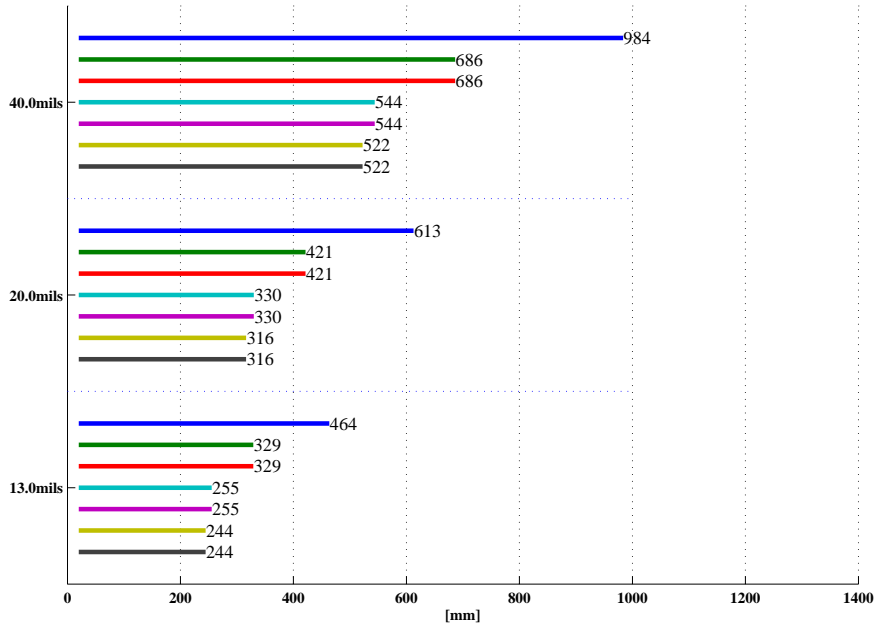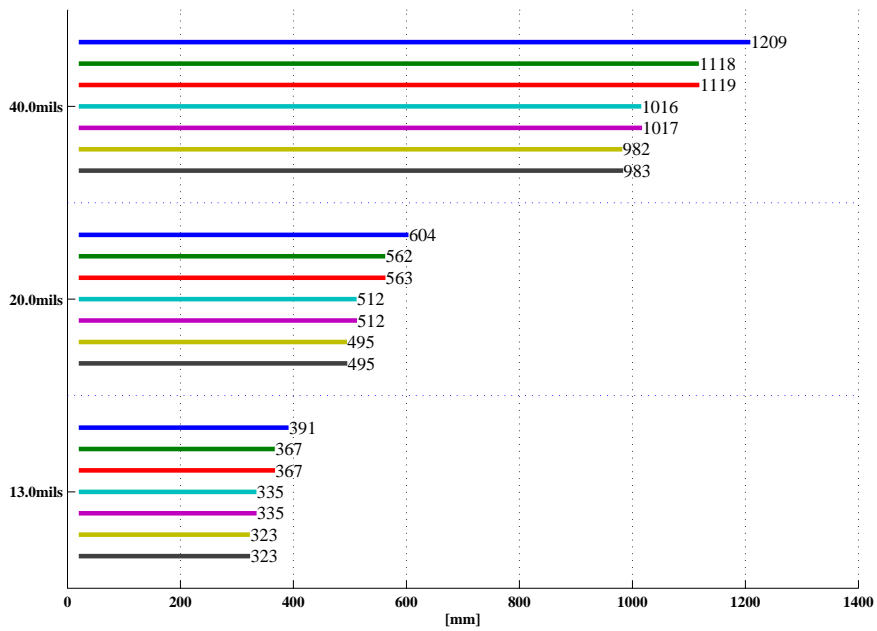
(a)
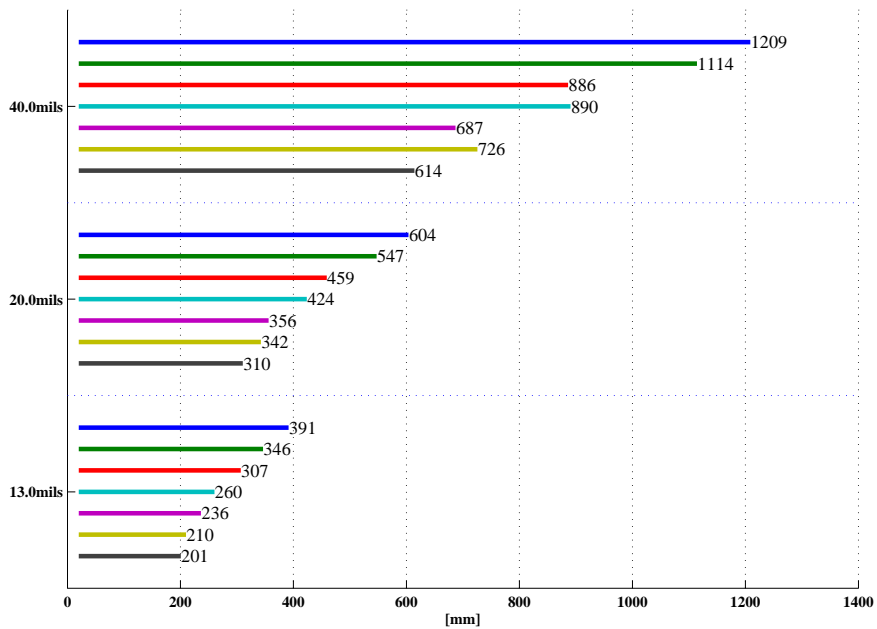


(b)

Figure 2.9: DOF on the tangential plane for the lens: (a) without CPM; (b) with CPM ($\alpha = -0.035$).
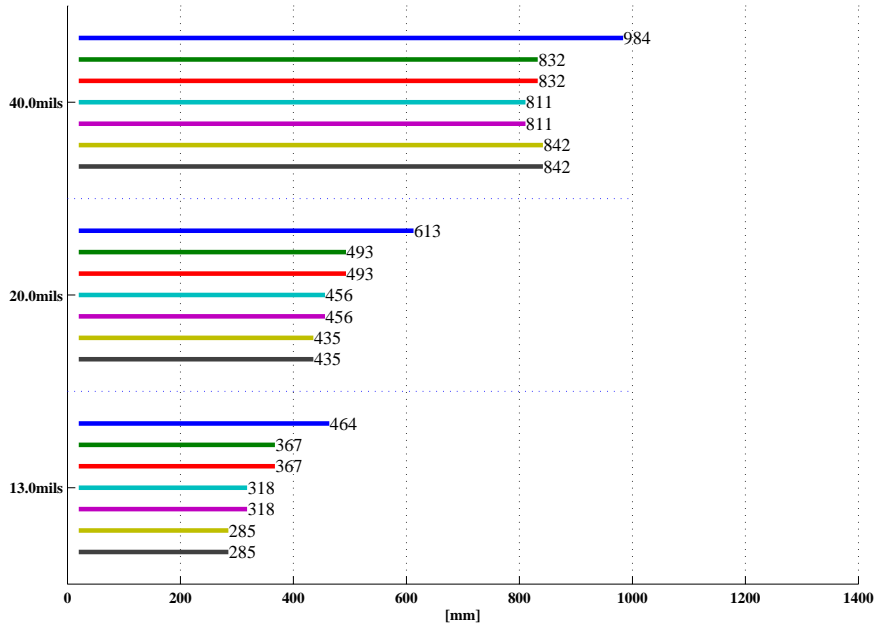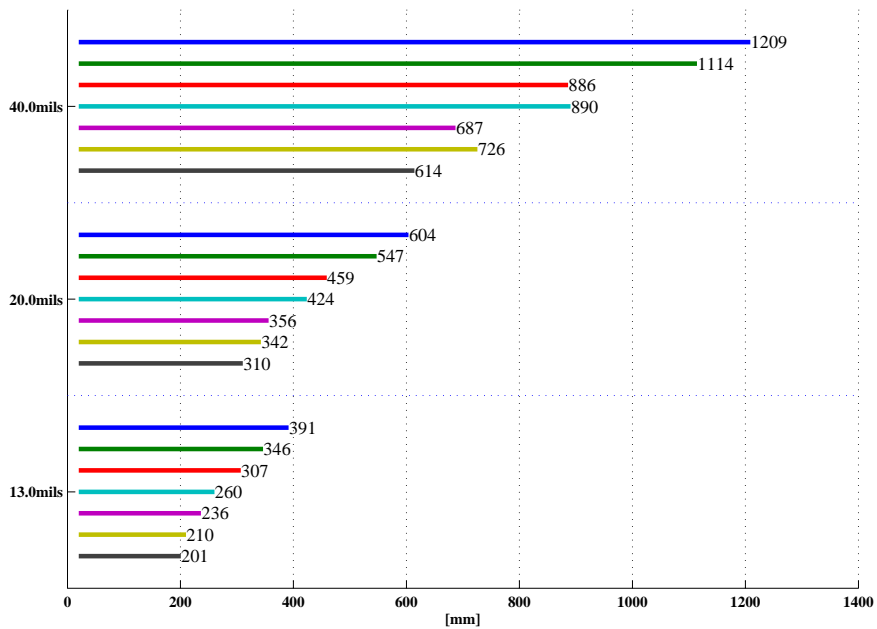
Figure 2.10: DOF on the sagittal plane for the lens: (a) without CPM; (b) with CPM ($\alpha = -0.035$).

(a)



(b)

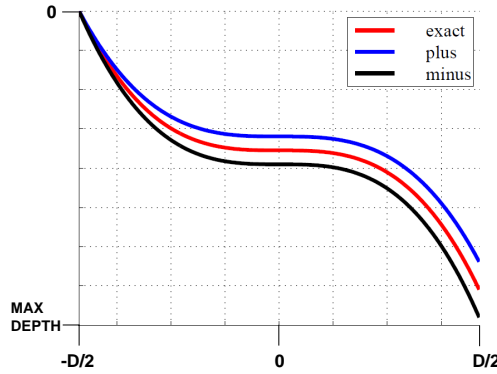Figure 2.11: Overall DOF for the lens: (a) without CPM; (b) with CPM ($\alpha = -0.035$).

Figure 2.12: CPM with exact $\alpha$ (red) and $\alpha \pm tol$ (blue and black). D represents the CPM diameter.

### 2.4.1   Fabrication tolerance analysis

For evaluating the tolerances of the phase mask, the performance of two systems with a phase mask profile calculated adding and subtracting the fabrication tolerance, has been studied. Fig. 2.12 shows the mask profiles computed with the two maximum deviations from the ideal profile. The effect of these fabrication errors on the lens performance are shown in Fig. 2.13 and Fig. 2.14. Fig. 2.13 reports the results for a positive error ($\alpha + tol = -0.0385$) compared to the exact case ($\alpha = -0.035$), while Fig. 2.14 reports results for a maximum negative error ($\alpha - tol = -0.0315$) and the exact solution. The achieved DoF is on average 5%-9% different with respect to the one obtained with the nominal value of $\alpha$.

All three cases have been fabricated. In the following, their features will be described.

### 2.4.2   Fabricated samples characterization

In this section, the characteristics of the phase masks that NanoComp has fabricated and their profilometric and optical characterization will be described.

The complete pattern, defined on the whole substrate, is made out of nine sample-patterns arranged in $3 \times 3$ matrix array. Three different patterns are fabricated, having three values of $\alpha$: 0.0385, 0.035, 0.0315. The different sample-patterns are referred to as SP1, SP2, SP3. In the matrix array, the SP1, SP2 and SP3 are repeated three times each. The cubic-etched area is connected to the non-etched region by an annulus having width of $50\mu$m. The etch depth in this relaxation region is constant along each radius and is equal to the cubic etch depth at the edge of the circle. Each phase mask is surrounded by four markers. The marker patterns are shown in Fig. 2.15. They are gratings arranged in the shape of a cross. The gratings have $4\mu$m period, $50\%$ duty cycle and $100\mu$m width. The overall size of the cross is $500\mu$m $\times 500\mu$m.
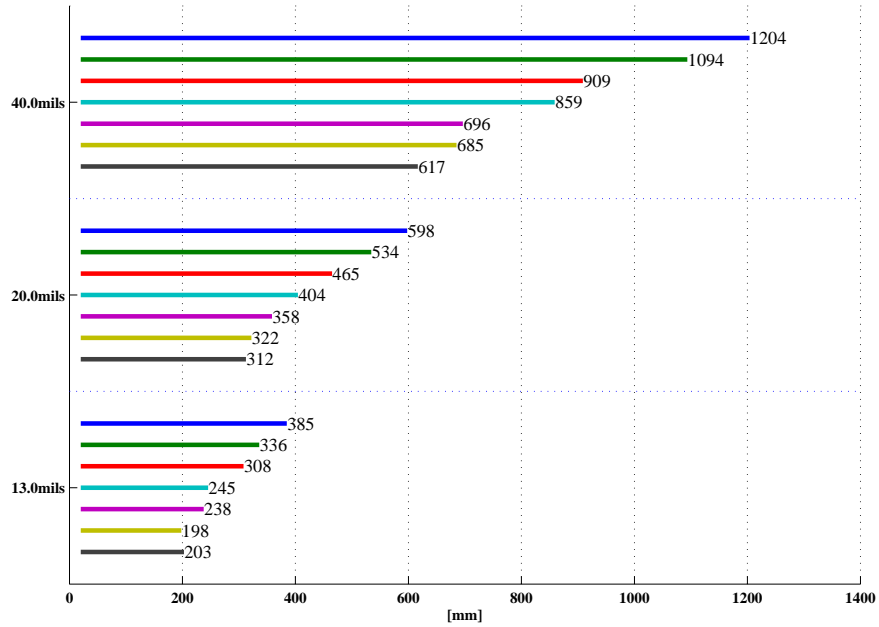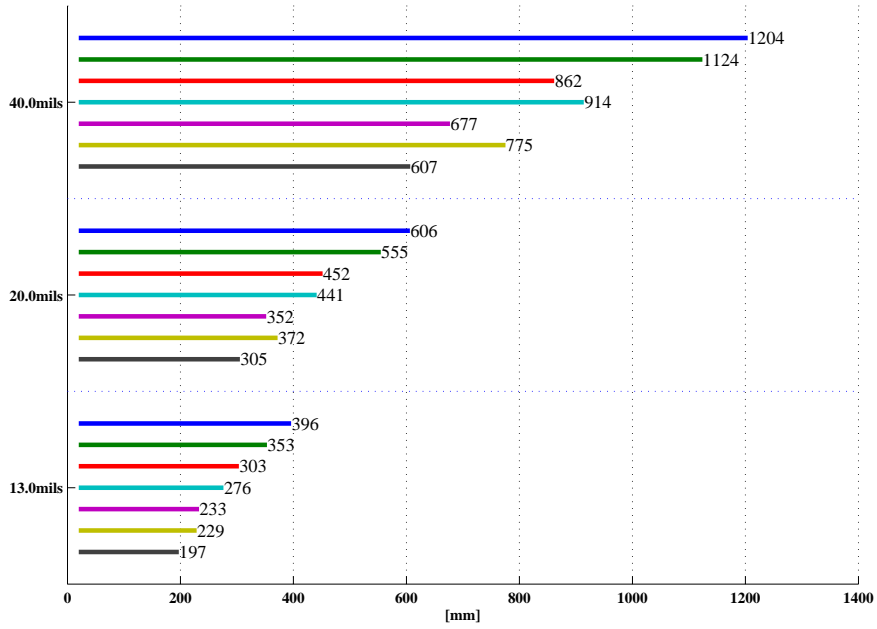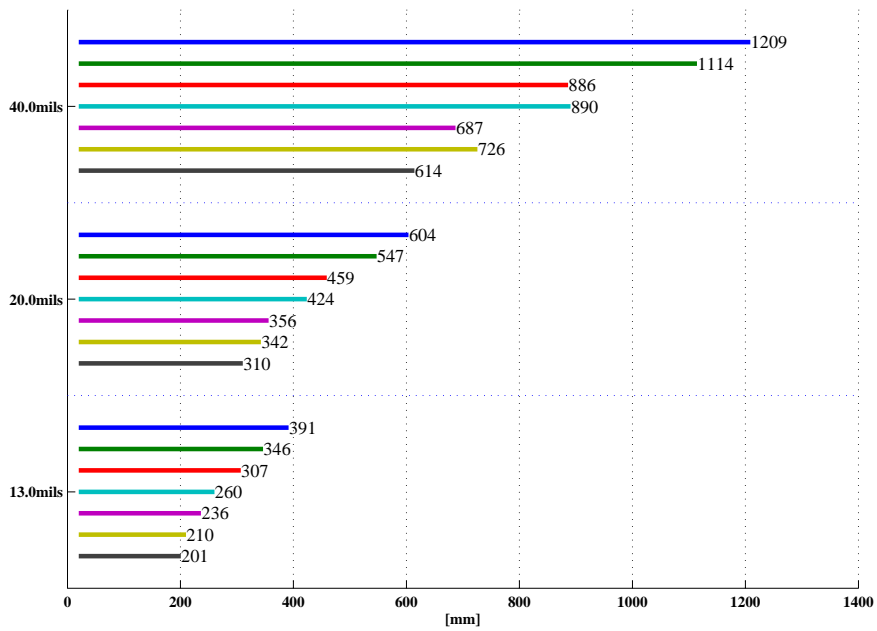
(a)
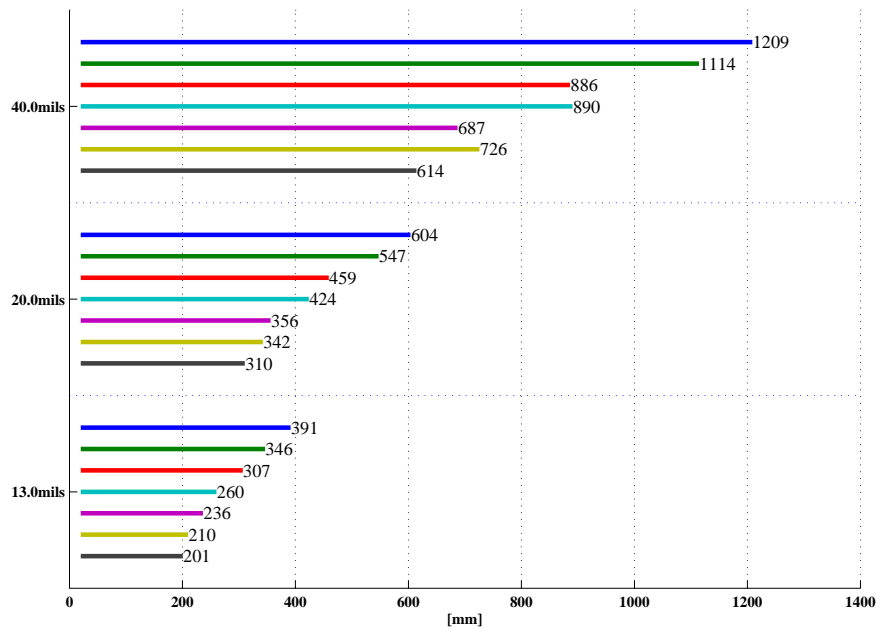


(b)

Figure 2.13: Overall DOF of the CPM system designed with: (a) $\alpha + tol = -0.0385$; (b) $\alpha = -0.035$.

(a)



(b)

Figure 2.14: Overall DOF of the CPM system designed with: (a) $\alpha - tol = -0.0315$; (b) $\alpha = -0.035$.
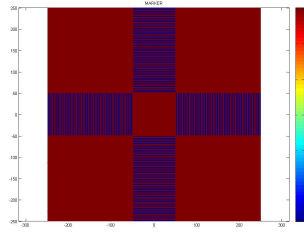
Figure 2.15: Color map representing the etching depth of the cross gratings.
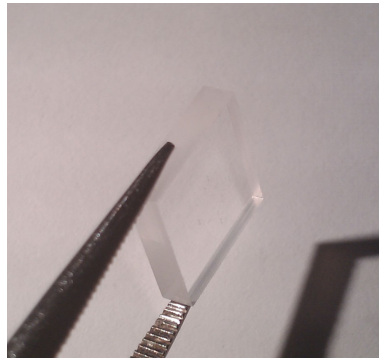


Figure 2.16: Picture of a phase mask sample

Fig. 2.16 shows a picture of a phase mask sample. In Fig. 2.17(a) the phase mask picture and the same phase mask surrounded by four alignment crosses, both taken by the microscope are shown.

**Profilometric characterization**

The real depth profile of all fabricated samples was characterized by NanoComp using a profilometer. The experimental depth curve along the horizontal direction ($y = 0$) is shown in Fig. 2.18 together with the nominal profile lines for three phase masks (one for each type): the optimal profile (type A) is shown in (a) as a continuous blue line, the optimal profile plus tolerance (type B) is indicated by the dashed black line (see (b)) and the optimal profile minus tolerance (type C) is indicated by the dashed red line (see (c)). The agreement between expected and obtained results is very good demonstrating the accuracy of the technological processes. After having verified the CPM features, their optical characterization will be presented.

**Optical characterization**

In order to test the behavior of the phase mask samples, the pattern that they project when illuminated by a coherent laser beam has been considered. In the

<div align="center">(a)</div>



<div align="center">(b)</div>

<div align="center">Figure 2.17: A phase mask with the alignment crosses.</div>



<div align="center">(a)</div>



<div align="center">(b)</div>
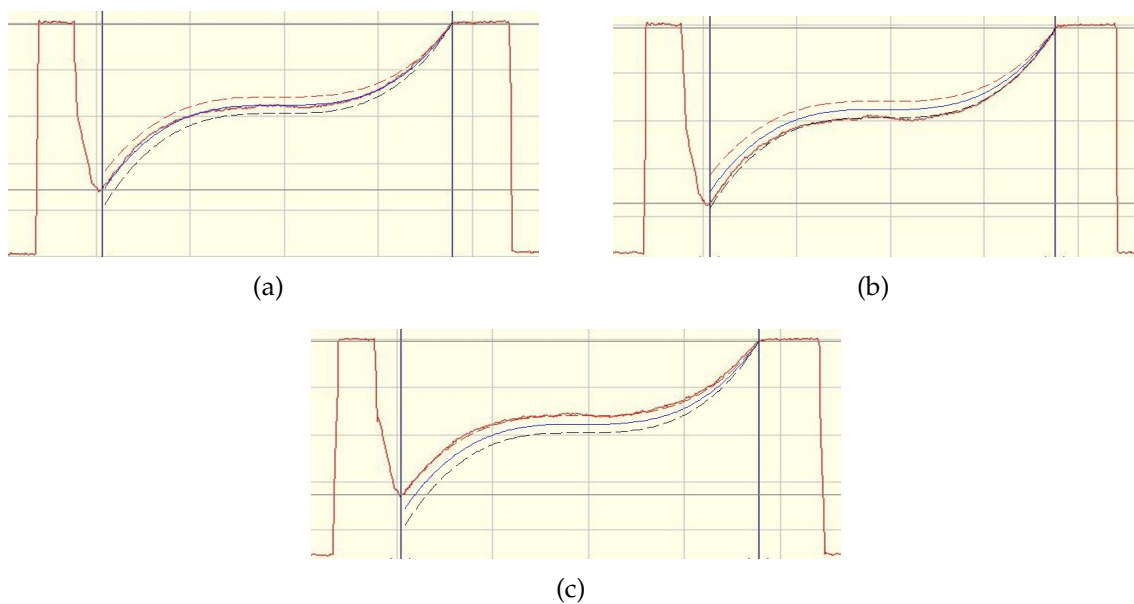


<div align="center">(c)</div>

Figure 2.18: Characterized etched depth of masks: (a) the optimal profile, (b) the optimal profile plus tolerance, (c) the optimal profile minus tolerance.
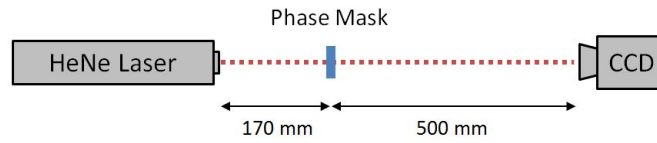
Figure 2.19: Sketch of the experimental set-up for evaluation of the PSF of the phase mask.
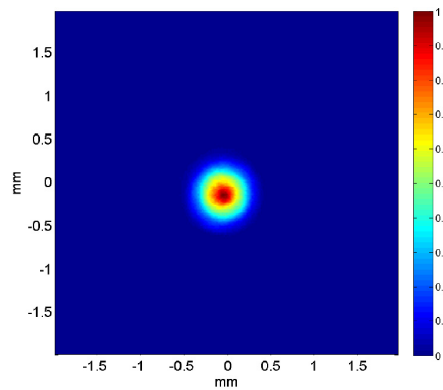


Figure 2.20: Measured intensity of the laser beam impinging on the phase mask.

following the results of this experiment will be presented.

The experimental set-up consisted in a He-Ne laser ($\lambda = 633$nm) illuminating the CPM; the beam exiting the CPM is imaged by a CCD sensor ($240 \times 240$ pixels, pixel size $16.6\mu$m) placed at $500$mm from the PM itself. The experimental set-up is sketched in Fig. 2.19.

The basic idea behind the experiment is that, since the CCD plane can be considered at far-field distance from the CPM and the whole system can be considered linear, the projected field can be well approximated by the convolution between the PSF of the CPM and the source beam. If the source is an ideal point source at infinity, the detected signal is the PSF itself.

The first problem comes from the fact that the source illuminating the CPM is not a point source, but a laser, providing a gaussian beam, as shown in Fig. 2.20.

Simulations can explain clearly the differences between the two cases. In particular, Fig. 2.21(a) shows the pattern projected by the CPM when illuminated by an electromagnetic plane wave (corresponding to the ideal point source at infinity) obstructed by a stop as large as the CPM profile itself. In this case the pattern has large resemblance to the PSF reported in [3] for optical systems with CPMs. However, if the illumination is given by a Gaussian beam such as that shown in Fig. 2.20, the resulting pattern is that of figure 2.21(b).

The comparison between the simulated and the experimental data can then be done only in a qualitative way. Deconvolution of the real source features to obtain the PSF would require a precise knowledge (both in amplitude and phase) of the
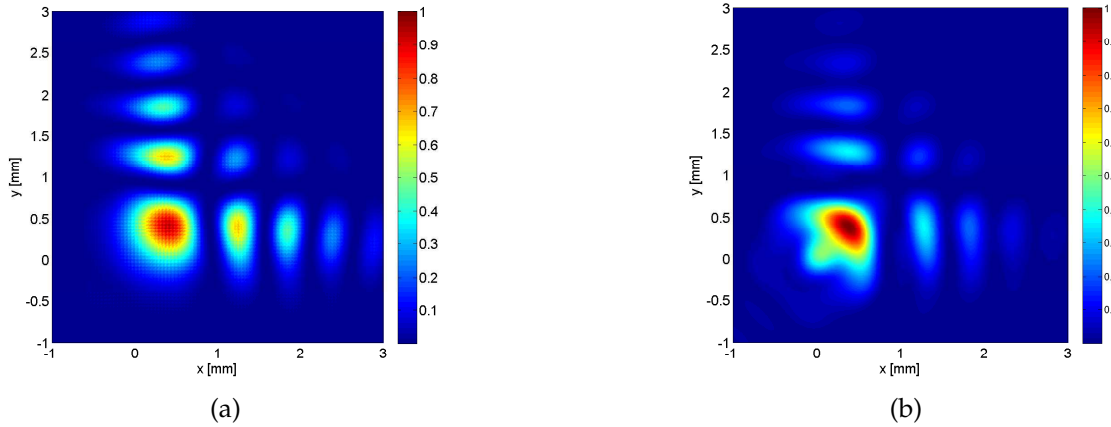
(a)           (b)

Figure 2.21: Simulation of projected pattern of phase mask of type A as illuminated by (a) a flat wavefront as wide as the mask and (b) a Gaussian wavefront larger than the mask

input beam. Moreover the set-up proved to be very sensitive to the alignment between the input beam and the CPM.

Fig. 2.22 shows the experimental patterns projected by three CPM samples (one of each type). They are all quite similar, with the previously shown main lobe deformation.

In the following the assembling of the CPMs and the lenses will be presented.

### 2.4.3 CPM system assembling

After verifying the CPM features, three phase masks (one of each type) were assembled into full systems.

The procedure consists of two steps: the assembling of the lenses and the aperture stop into the lens barrel and the gluing of the CPM to the barrel.

The first step is a standard lens mounting and was carried out by a specialized technician.

In the second step the CPM was placed on top of the lens with glue at the interface. The alignment was made using an optical microscope, aligning the phase mask pattern to the diaphragm. After alignment, the glue was fixed by exposition to ultraviolet light for few seconds. Fig. 2.23 shows the resulting component.

Fig. 2.24 shows optical micrographs of the diaphragm as seen through the phase mask side of the lens. By analyzing the relative position of the phase mask pattern and the stop rim, it was found that for lens with CPM type A the misalignment is $\approx 8\mu m$, for type B the misalignment is not noticeable and for type C the misalignment is $\approx 10\mu m$. Such small misalignment values do not result into evident effects.

Fig. 2.25 shows how a small round pinhole is imaged by the optical system

(a)



(b)



(c)

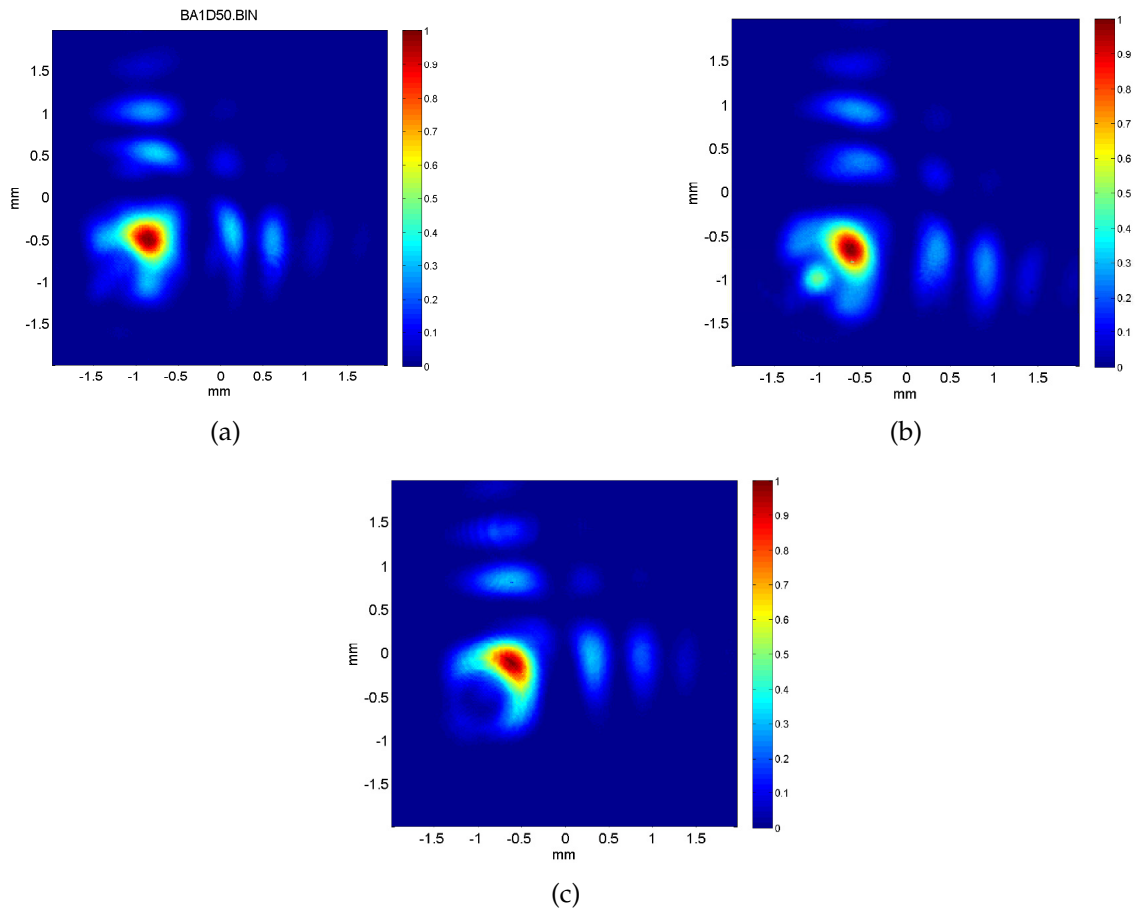Figure 2.22: Projected pattern of phase masks (a) type A, (b) type B, (c) type C



Figure 2.23: Phase mask sample glued to the lens barrel.
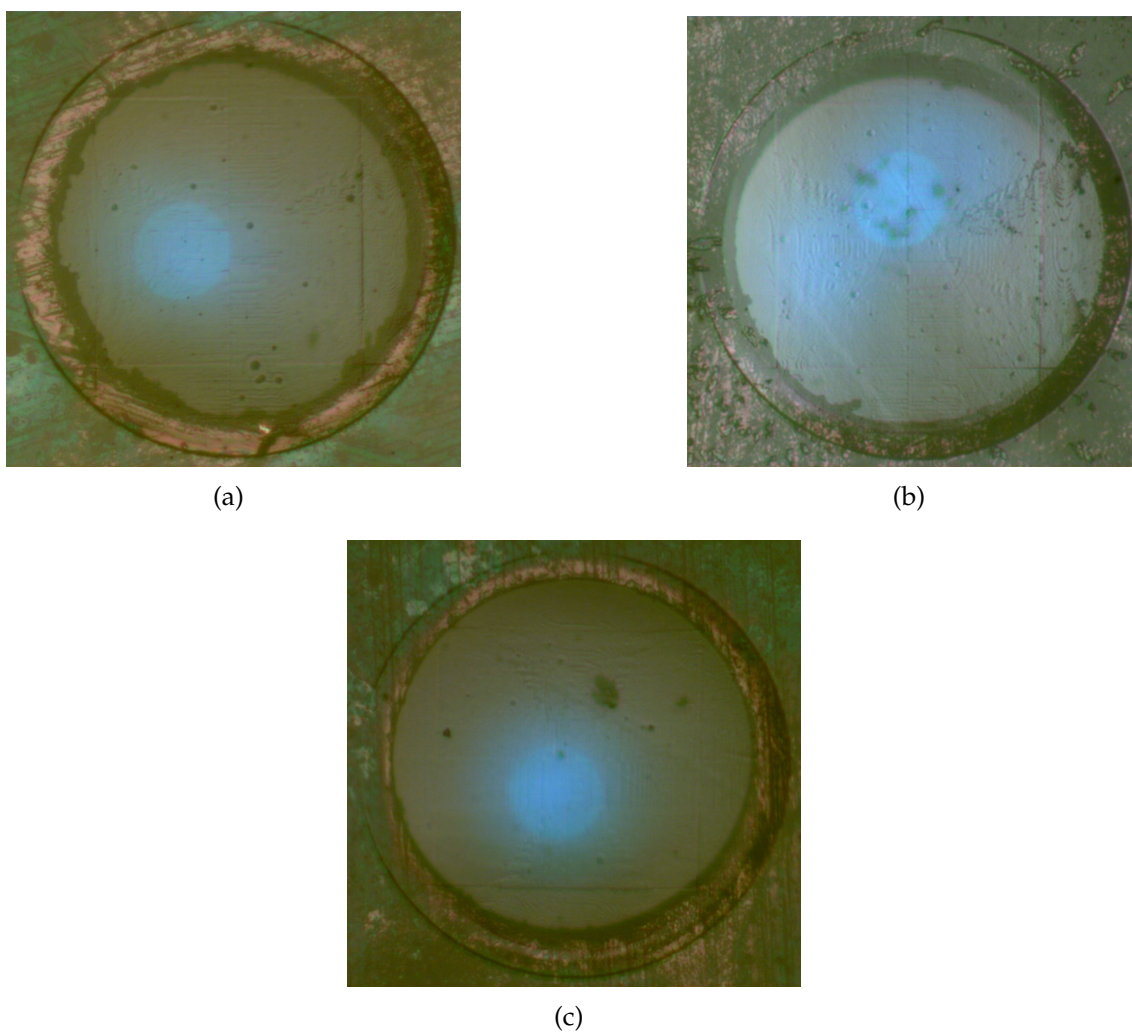
(a)

(b)

(c)

Figure 2.24: Optical micrographs of the lens diaphragm as seen through the phase mask (a) type A, (b) type B, (c) type C.
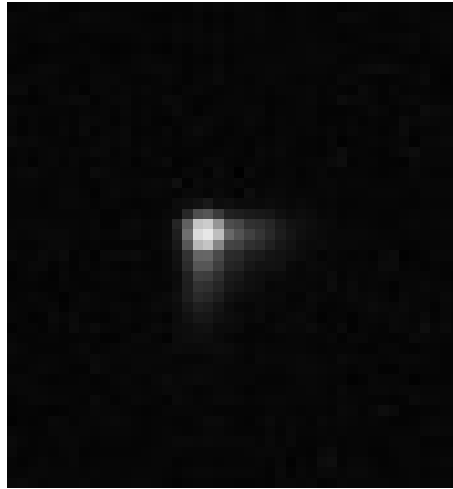
Figure 2.25: Image of a pinhole

with the phase mask. Since the pinhole is very small, this image is very similar to the PSF of the optical system itself. Its shape is analogous to those reported in [3] for similar systems.

In the next section, the results obtained by the characterization of three CPM systems (one of each type) will be presented.

## 2.5   CPM system characterization

In this section specific characterization results will be shown. Firstly, the MTF at infinity of the systems with and without CPM will be discussed. Secondly, the systems performance, before and after the CPM assembling, characterized measuring the SFR with the SAFARILAB tool will be presented. This demonstrates also the versatility of SAFARILAB (see Ch. 1), that is suitable for characterizing all kind of optical systems.

### 2.5.1   MTF at infinity measurement

In the following the results of the characterization of the MTF at infinity of lens with and without phase masks will be described. This comparison allows a first quantitative evaluation of the effect of introducing the optimized CPM in the optical system.

The instrument for the evaluation of the MTF of the object at infinity is shown in Fig. 2.26(b). A light source at $523\text{nm}$ (green) with an amplitude shape in the form of a cross (see Fig. 2.26(a))passes through the lens under investigation and is collected by a calibrated imaging system. In this way, it is possible to calculate the LSFs in the two directions using the captured image of the line source in the

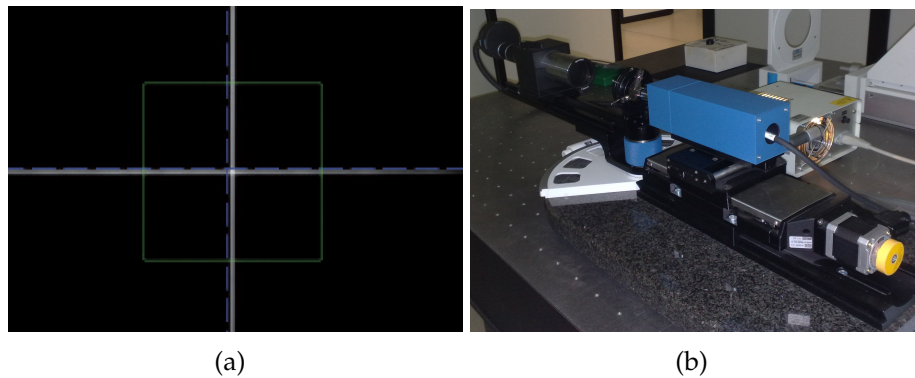(a)                                              (b)

Figure 2.26: (a) Light source with an amplitude shape in the form of a cross; (b) Experimental set up to evaluate the lens MTF at infinity.

sagittal and tangential planes and to generate the MTF by Fourier transforming the LSFs.

The MTF were characterized for tangential and sagittal planes with object at $0°, \pm 8°, \pm 16°$. The lenses have been characterized before and after the phase masks mounting in order to highlight the effect of the CPM optimization.

At any fixed source (object) angle, the instrument can displace its sensing stage automatically, as to characterize the MTF over a set of back focal distances (or back focal length, BFL).

As a first direct qualitative view of the effect of the phase mask, Figs. 2.27 and 2.28 show the MTF curves for all the BFL values taken into account, considering the lens with and without CPM type A and the central field. Similar results were obtained for the other CPM systems. A comparison of the curve bundles shows that the phase mask has a large impact on the optical MTFs: on average, the curves are lowered, but the zero positions are shifted to higher frequencies, near the optical cutoff. Finally, their dependency on the BFL is made less strong. It is noted that the experimental tangential and sagittal MTF are slightly different: this is due to an imperfect alignment caused by the lens shape that does not fit the holder clamps. Finally, the simulated data reproduces the experiment very closely.

In order to have a more compact view of the results, the MTF is plotted at two fixed spatial frequencies for different BFL expressed as shift from the in-focus BFL represented by the zero value. Considering lenses with and without phase mask type A, Figs. 2.29 and 2.30 show the MTF values for the field of $0°$ at $1/3$ and $2/3$ of the normalized frequency (Nyqist frequency, $f_N$), respectively, with simulated data on top of experimental ones. Figs. 2.31 and 2.32 show the MTF values for the field of $8°$ at $1/3$ and $2/3$ of $f_N$, respectively, with simulated data on top of experimental ones. Figs. 2.33 and 2.34 show the MTF values for the field of $16°$ at $1/3$ and $2/3$ of $f_N$, respectively, with simulated data on top of experimental ones.

The effect of the CPM is to lower the MTF peak sensibly and to make the
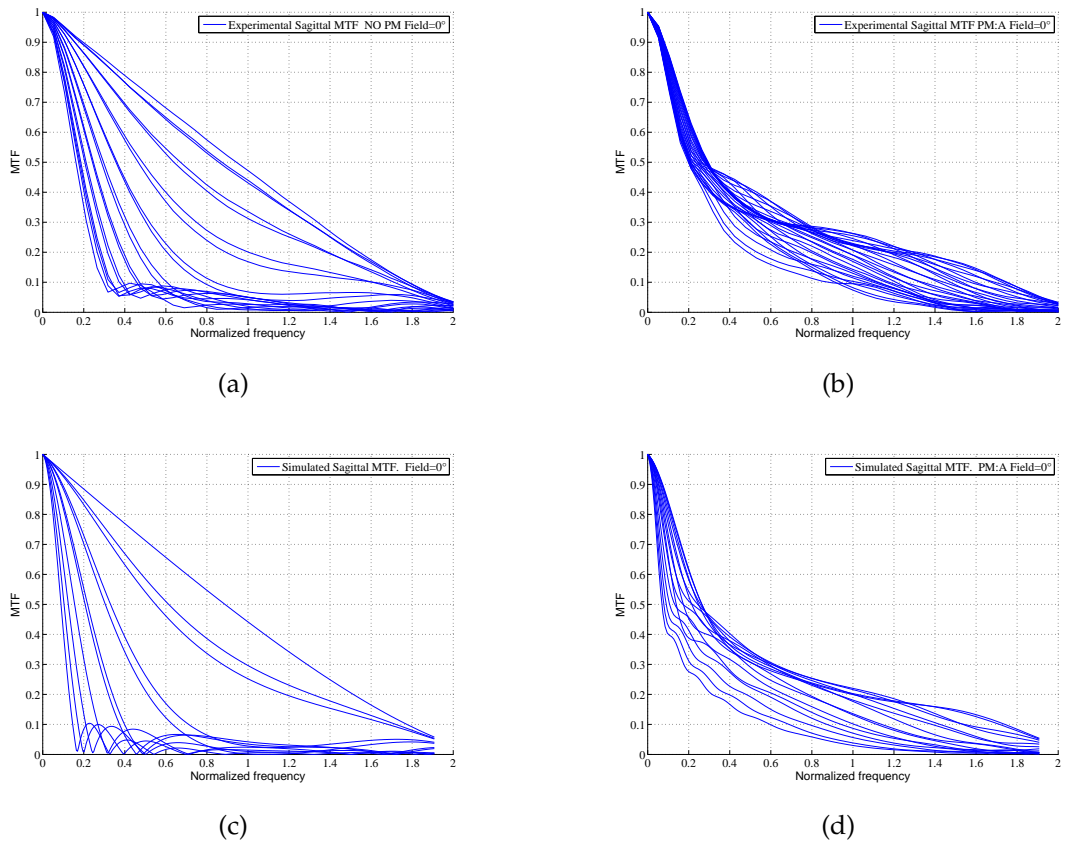
Figure 2.27: Measured (upper figures) and simulated (lower figures) sagittal MTF at different back focal distances without (left side figures) and with (right side figures) the phase mask.
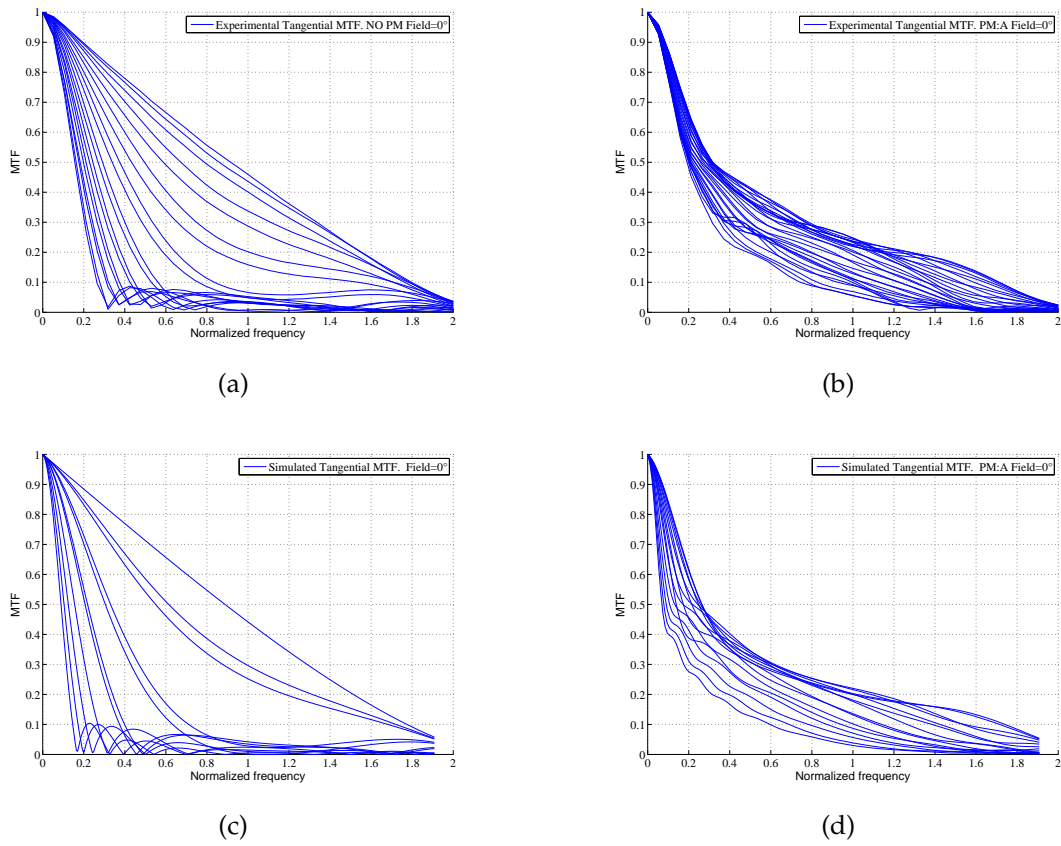
Figure 2.28: Measured (upper figures) and simulated (lower figures) for the tangential MTF at different longitudinal distances without (left side figures) and with (right side figures) the phase mask.
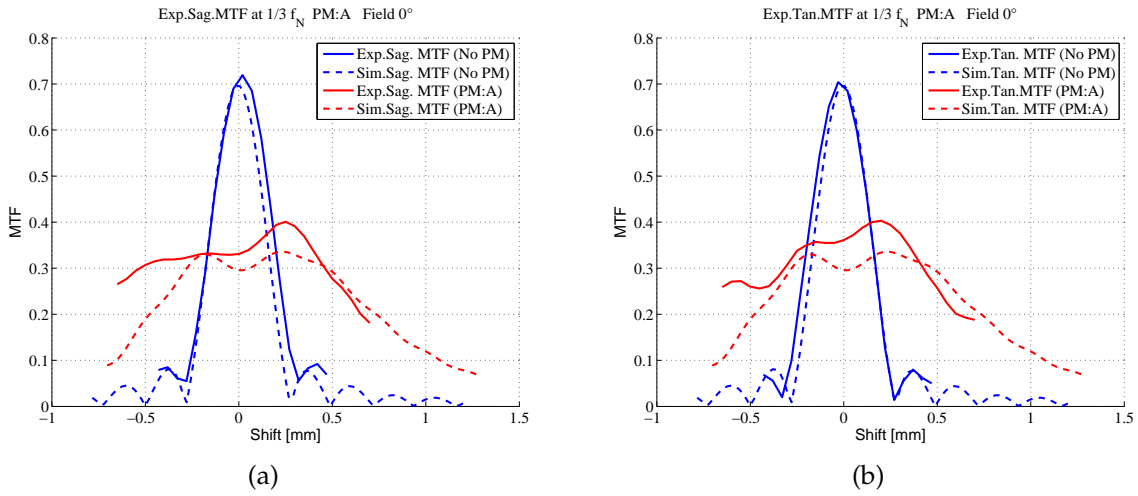
Figure 2.29: Experimental and simulated MTF at $1/3$ of $f_N$ at $0°$

curves flatter over a BFL range almost two times wider than that of the lens without the CPM.

Note again how well the simulated data reproduce the experiment.

At angled fields, the experimental MTFs are much lower than the simulated ones. This may be due to misalignments, to lower power incident on the sensor that makes the measurement more uncertain and to the additional lateral shift of the focused spot that cannot be completely followed by the sensing stage of the instrument.

These plots clearly highlight the effect of extension of the in-focus region induced by the optimized CPM and show that the experimental results fit those obtained by simulations well, thus confirming the quality and the reliability of the design procedure.

The results of this type of measurement suggest that when the object source is at finite distance, the SFR is less sensitive to its varying position for any fixed BFL. In the next section, results of the SFR measurements will be reported and commented.

## 2.5.2   SFR measurement using SAFARILAB

SFR measurements have been performed also with object and image planes at finite distance. Results can then provide information more suitable for verifying the performance of devices in real operating conditions. Also in this case, the lenses have been characterized before and after the phase masks mounting in order to highlight the effect of the CPM optimization.

Before starting to measure the SFRs of the system with and without CPM, it is important to mount the lens in front of the sensor. The distance between
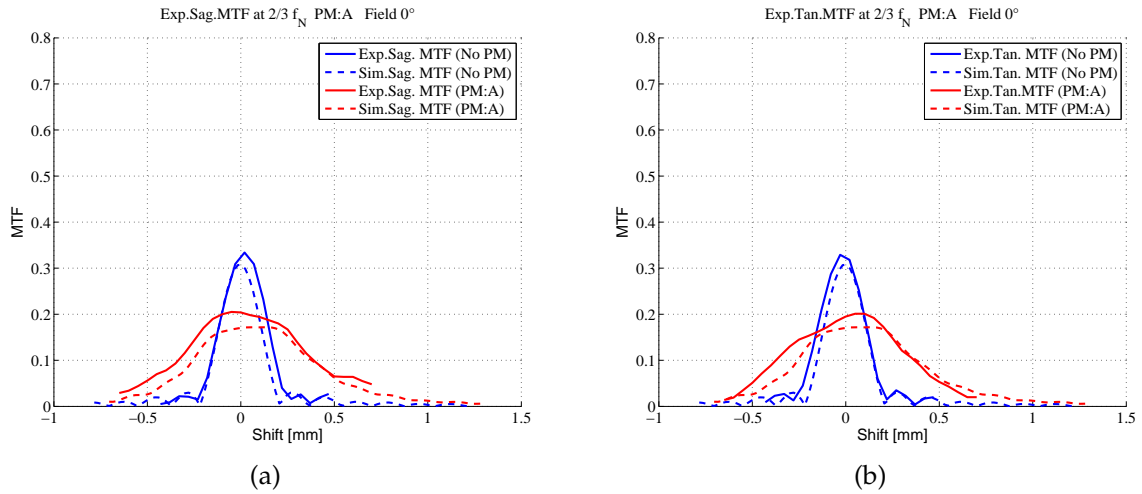
Figure 2.30: Experimental and simulated MTF at $2/3$ of $f_N$ at $0°$
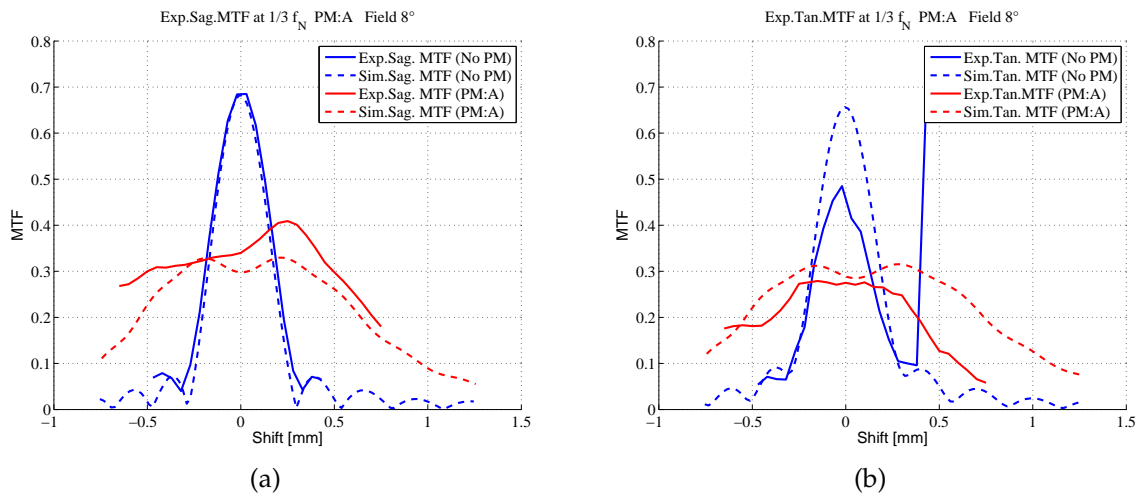


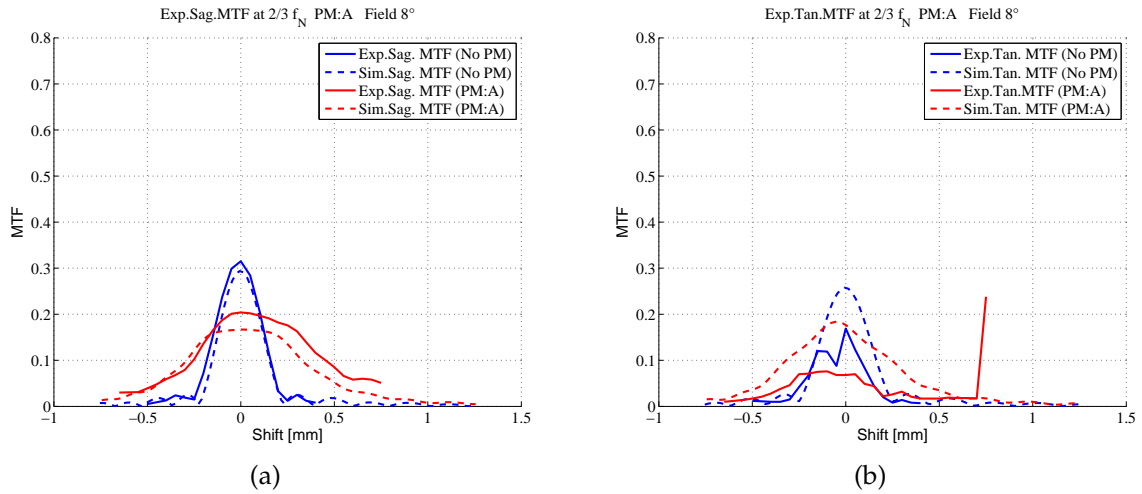Figure 2.31: Experimental and simulated MTF at $1/3$ of $f_N$ at $8°$

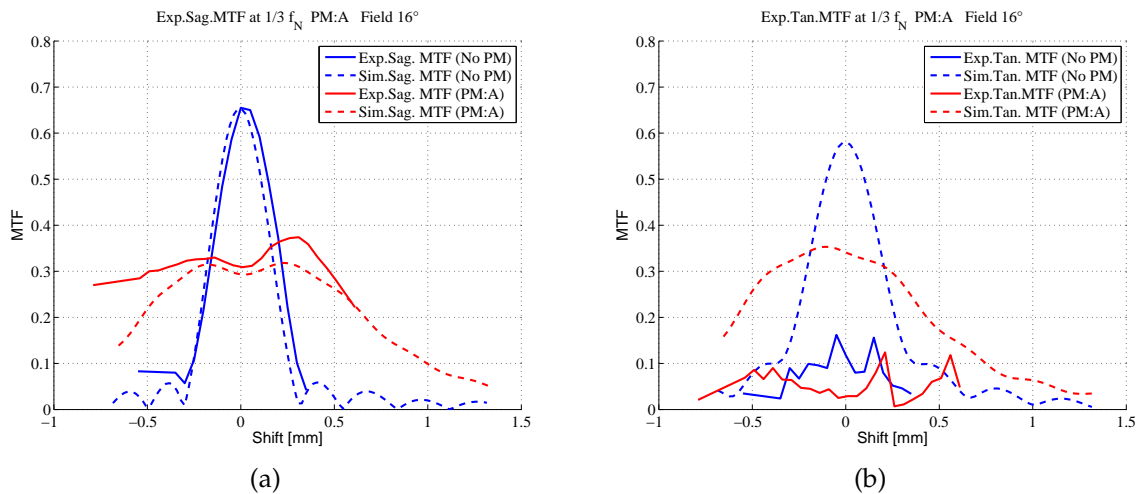Figure 2.32: Experimental and simulated MTF at $2/3$ of $f_N$ at $8°$



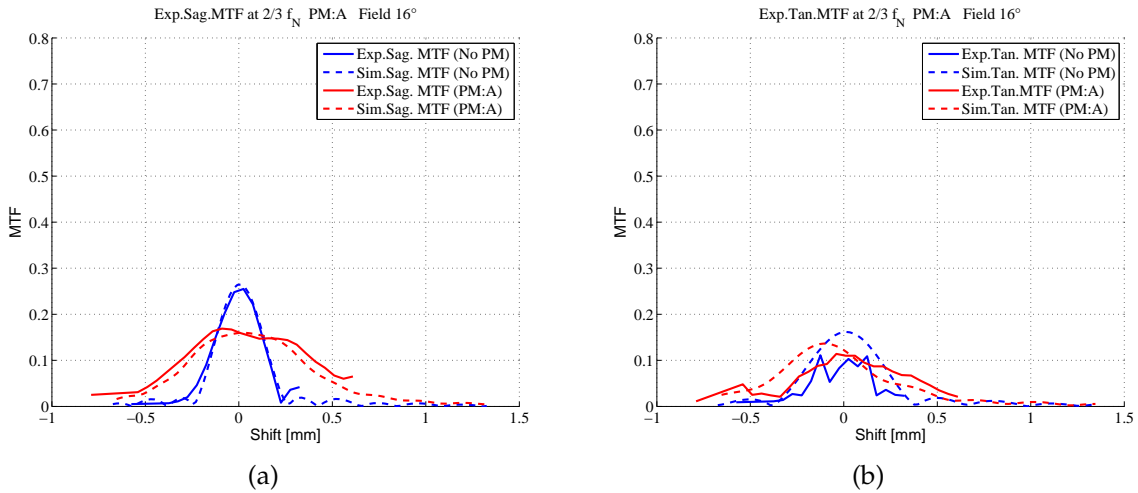Figure 2.33: Experimental and simulated MTF at $1/3$ of $f_N$ at $16°$

Figure 2.34: Experimental and simulated MTF at $2/3$ of $f_N$ at $16°$

them should be the BFL found by the optimization process. It was necessary to evaluate it experimentally.

The optimal working position of the lenses within the holder is then found, before CPM mounting maximizing the SFR for an object positioned at the focus plane of the original lens. The position could be adjusted in steps of $100\mu m$ making use of gauged shims. The calculation of the SFRs has been done with SAFARILAB tool, described in Ch. 1.

After assembling the lens and the sensor together, the SFRs of the system with and without CPM have been measured. For this purpose a slanted-edge has been moved from a distance of $20mm$ to a distance of $1500mm$. For each object distance, the SFR curves are calculated using option B and A of SAFARILAB to reduce the presence of noise. Each image for the SFR evaluation was the average of 10 images, and each SFR was the average of 10 SFRs. For this set of measurements only the central field was inspected.

After the characterization, further ZEMAX$^®$ simulations were run to match the experimental results. In particular, since the lens has no fixed position within the optical chamber, the BFL was found "a-posteriori", as the one that matches the experiments the most.

In order to be compared with the experimental data, the simulated MTF has to be multiplied by the detector footprint MTF that is expressed by Eq. (1.6).

Fig. 2.35 shows simulations and measurement results at all distances. The effect of the mask is evident: it reduces the sensitivity of the MTF to the object distance, reduces its average value at all frequencies and shifts the zero toward the cutoff.

The simulations reproduce the experiment closely.

Considering an MTF threshold usually under $30\%$, the MTF curves are trans-
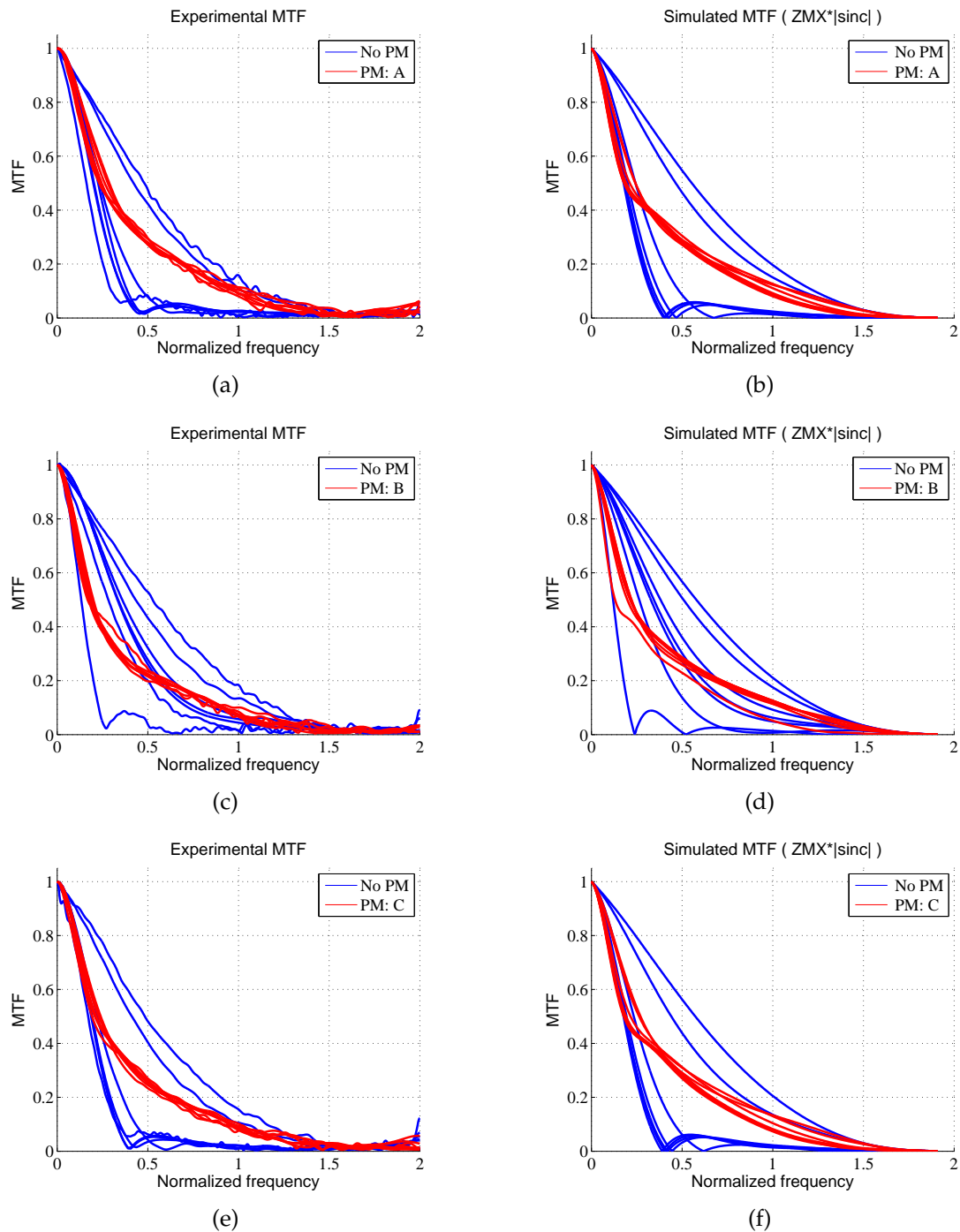
Figure 2.35: Comparison between simulations and measurements results with and without phase mask
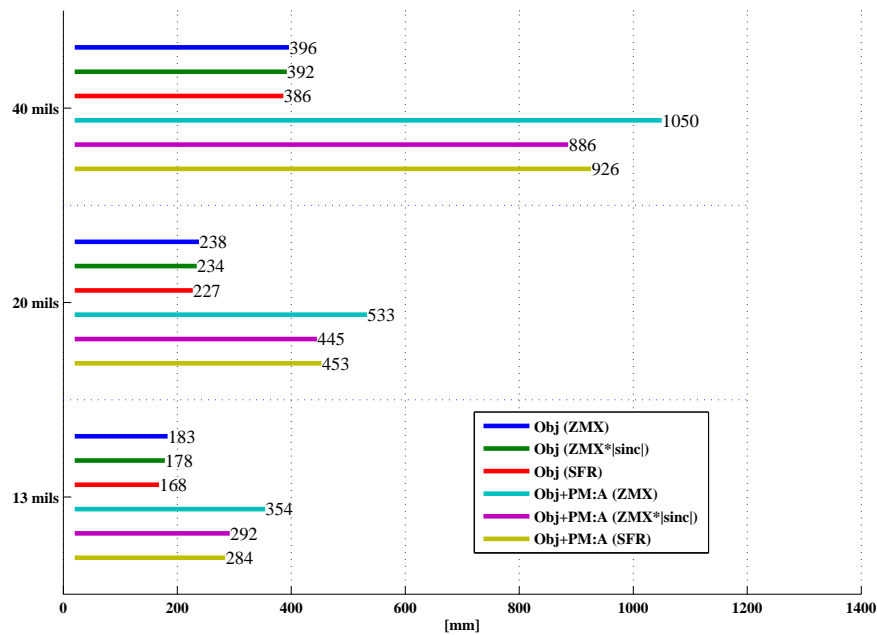
Figure 2.36: Simulated and experimental resolution bars with and without phase mask type A.

lated into information on the resolution on the object plane just considering the central field. Figs. 2.36, 2.37, 2.38 show the DoF bars with and without the phase masks type A, B, C respectively. The DoF for simulated ZEMAX® data with and without the detector effect and for the experimental SAFARILAB data has been calculated. The DoF is increased almost doubled for these resolutions shown.

## 2.6 Conclusion

In this chapter, the optical design of an imaging system with EDoF have been described.

After a literature review, it appeared that the so called Wavefront Coding based systems show the best performance for the DoF extension of an optical system. The Cubic Phase Mask is the type of mask that has the most solid theoretical bases, resulting from the optimization of the MTF. For this reason, it has been chosen for validating OLOM tool that allows to design WFC systems with MTF invariance.

After having presented the main feature of this tool, the optimization of a CPM system has been described. The considered requirements were the ones of a barcode reader, taken into account as an example of Auto ID and portable system application. The simulated performance of the optical part of the system with and without CPM has demonstrated the ability of the CPM optimized by OLOM

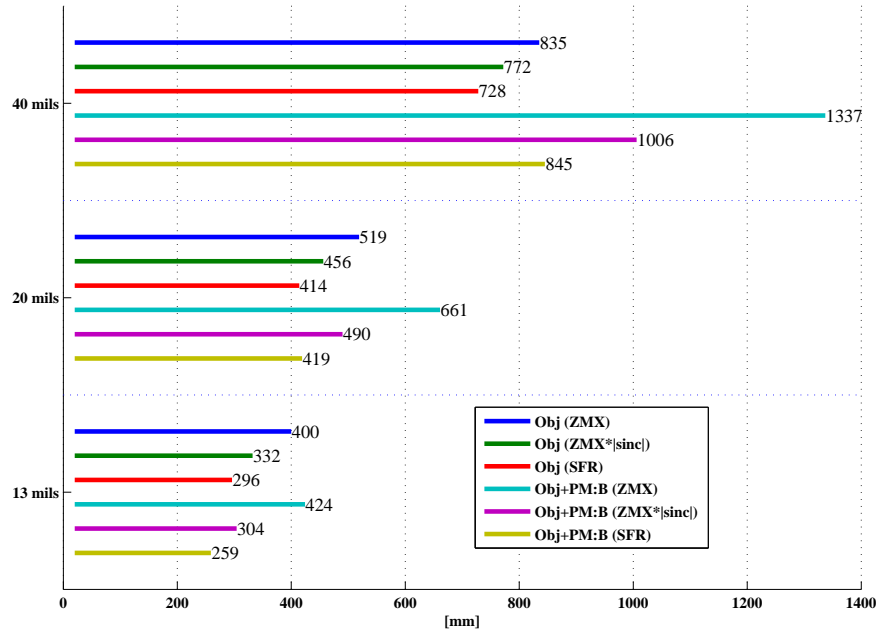Figure 2.37: Simulated and experimental resolution bars with and without phase mask type B.
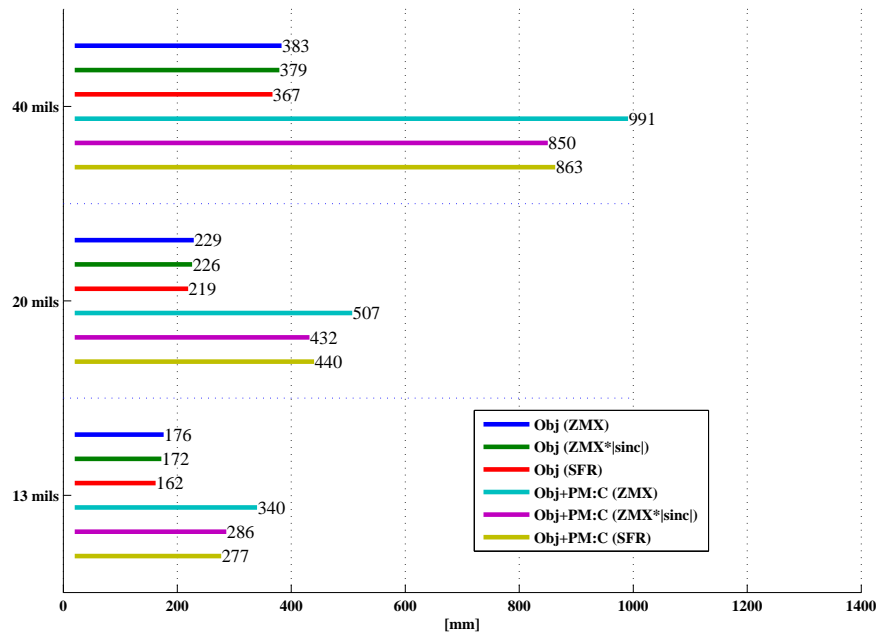


Figure 2.38: Simulated and experimental resolution bars with and without phase mask type C.

tool in extending the DoF.

After designing the CPM system and simulating its results, it has been realized. A fabrication tolerance analysis have been shown that adding and subtracting the tolerance value to the optimized cubic shape, the differences in terms of DoF are around the 5%-9% of the exact CPM DoF. All three cases have been fabricated and characterized firstly with a profilometer. The results demonstrate the accuracy of the technological processes. Secondly, an optical characterization has been done. The beam shapes produced by the CPMs agreed with the expected ones. Lastly, after assembling the CPM with the stop and the lens, the misalignment values have been measured. They have not resulted into evident effects.

Finally, the CPM system performance has been studied. Firstly the MTFs at infinity of the lens with and without system have been measured for different BFL values. The results showed the agreement between the simulated and the measured data, but also confirmed the EDoF produced by the optimized CPM. Secondly, SFRs of the system have been measured using SAFARILAB . Measurements demonstrated again that the CPM optimized using OLOM enhances the performance of a lens, extending its DOF. The whole design procedure is reliable, since the results obtained by theoretical means are confirmed by the characterized behavior of the realized prototypes.

In conclusion, OLOM tool have been validated and its ability to increase the performance of a system in terms of DoF has been demonstrated.

However, the designed CPM system does not achieve all the required specifications. This is due to the fact that only the optical part is considered during the optimization, and not the one of the whole system. Moreover, the electronic post-processing is not taken into account during the evaluation of the system performance (Sec. 2.3.2). In order to take into account the whole system and not only the optical part, an optimization and analysis tool that allows to design the lens at a system level has been developed. In the next chapter, the results obtained using this tool, named SLALOM , will be shown and discussed.

# 3

# Imaging System Design: System level

In this chapter, the analysis and design of optical systems including also the effect related to the presence of electronics and detection will be described. The study is done not only at the optical level, as seen in the previous chapter, but also at the system level, considering all the system components. In this way, the design is achieved by optimization of the overall system at once, rather than by sequential optimization of its constituting elements.

This study is done using a novel software developed within the research group: the package has been named SLALOM , acronym of "System Level AnaLysis and OptiMization" and is based on two tools that can be run independently: Optimization-SLALOM (O-SLALOM ), for design and optimization of new systems, and Analysis-SLALOM (A-SLALOM ) for the analysis of the existing ones.

This framework will be described first (Sec. 3.1). Then results of the application of O-SLALOM and A-SLALOM to the design of a lens with EDoF obtained exploiting spherical aberration will be presented (Sec. 3.2). The choice of spherical (quartic) aberration has been made (see Sec. 2.1) because it provides a PSF with circular symmetry that allows to use a conventional post-processing.

## 3.1 SLALOM : System Level Analysis and Optimization tool

In this section, the features of SLALOM will be described. The software can be applied to any electro-optical imaging device and in this work it will be targeted to the design of the vision system of a barcode reader.

An electro-optical vision system for the acquisition of information encoded in the form of a barcode can be considered as the cascade of three major blocks (schematically shown in Fig. 3.1):
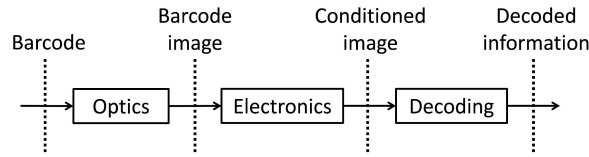
Figure 3.1: Schematic of the barcode reader as an electro-optical system.

- the optical subsystem, composed by the lens. Its behavior is modeled by ZEMAX® which provides its PSF or SFR;

- the electronic subsystem, made by the photo-detector that translates the optical image into an electric signal also adding noise;

- the decoder. Its specifications are on the SFR and on the SNR which depend on the code to be read and on its distance from the reader.

The optical subsystem allows a barcode printed over a support material and illuminated by a source to be imaged over a sensitive medium (typically CCD or CMOS photosensors). The electronic subsystem moves the image on to the electrical domain so that it is digitally re-conditioned (post-processed) and finally interpreted (decoded) to extract the information.

The design of the electro-optical system is classically made by designing the three constituting parts independently. However, in a complex system the best performance of a single element may not correspond to that of the complete processing chain [38]. Therefore, as seen in Ch. 2, the optimization of the system only at the optical level may not be enough to achieve the required performance. The system optimization should then consider the parameters of the whole system (i.e. belonging to both optical and electronic subsystems) in order to maximize the decoding rate.

For this purpose, the design tool is implemented as a C-based extension to ZEMAX® that includes the modeling of the electronic subsystem not encompassed by ZEMAX® itself. Fig. 3.2 shows how the general system description is adapted with respect to the functionality of (a) O-SLALOM and (b) A-SLALOM . Even if they have different input and output interfaces, the two tools share a large part of core routines which provide model of the same electro-optical system. For example, the optimization could be seen as an iteration of analysis. These tools have been called with different names just to emphasize their purpose to be used.

O-SLALOM is used iteratively by ZEMAX® , therefore it has to respond to its command and provide output that is compatible to its optimization engine. The purpose of the tool is to evaluate the SFR and the SNR of images generated by the system blocks (cascade of optical system, noisy photosensor and digital reconstruction) and to compare them with the requirements for decoding given by the decoding library.
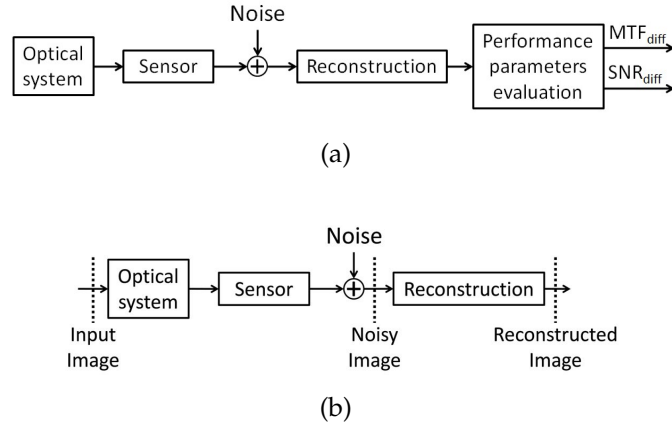
Figure 3.2: Schematics of O-SLALOM (a) and A-SLALOM (b).

A-SLALOM is used in independent single runs directly by the user or by a command script. Its purpose is to create the barcode images generated by the system blocks (sequence of optical system, noisy photosensor and digital reconstruction). The images can then be further processed for decoding (by a separate tool) as to complete the whole barcode reader functionality.

Before presenting O-SLALOM and A-SLALOM , it is important to describe the general features of the electro-optical system on which they are both based. It is composed of the common blocks of the two schematics of Fig. 3.2: a sensor located after the optical system creates a noisy image which must be electronically restored to be "readable".

**Optical system**

The optical system (first block of Figs. 3.2(a) and 3.2(b)), fully modeled by ZEMAX®, is composed by lenses and glass surfaces that create an image of the object on the image plane. Its parameters are set as variable during the optimization and will be better discussed in Sec. 3.2.

**Illumination system**

The illumination system (not shown in Fig. 3.2 as it is separated from the barcode reader) and the environment lighting where the images are taken are also important for successful decoding.

Illumination conditions are described by radiometric quantities, in particular the *irradiance* ($E_E$), measured in Watt per square meter ($W/m^2$), gives the power of electromagnetic radiation per unit area incident on the object plane as a function of the distance from the reader:

$$E_E(z) = \frac{I_{E0}}{z^2} + E_{E,room} \tag{3.1}$$

where $z$ is the distance from the emitting source measured in m, $I_{E0}$ is the radiant intensity (measured in W) given by the illumination system and $E_{E,room}$ is the irradiance due to the room light, assumed constant in the whole space.

These radiometric quantities correspond to photometric ones: the *luminous intensity* $I_{V0}$ (measured in lumen, lm) and the *illuminance* $E_{V,room}$ (lm/m$^2$). The conversion coefficient $k$ is referred to as "luminous efficacy" and depends on the source of the considered radiation. The relations are:

$$I_{E0} = \frac{I_{V0}}{k_{LED}}; \qquad E_{E,room} = \frac{E_{V,room}}{k_{room}} \qquad (3.2)$$

where $k_{LED}$ is the luminous efficacy of the LED source and $k_{room}$ is the luminous efficacy related to room light.

The irradiance on the object plane $E_O$ as a function of the distance from the reader is evaluated using the formula:

$$E_O(z) = 0.9 \left( \frac{I_{V0}}{k_{LED} \, z^2} + \frac{E_{V,room}}{k_{room}} \right) \qquad (3.3)$$

where the coefficient $0.9$ considers the reflectivity of the paper where the input image is printed.

**Sensor and noise**

The sensor (second block of Figs. 3.2(a) and 3.2(b)) converts the incident radiating power into current. Since the detector surface is an array of pixels, the image provided by the detector is a sampled version of the image provided by the lens.

Each pixel of area $A$, integrates the incident irradiance $E_I$ by a time $\tau$ allowing the formation of $n_e$ electrons according to the relation

$$n_e = QE(\lambda) \, E_I \, \tau \, A \, \frac{\lambda}{hc} \qquad (3.4)$$

where $QE$ is the detector quantum efficiency, $\lambda$ is the optical wavelength, $h$ is the Plank constant and $c$ is the speed of light.

Assuming a linear detector, which is the case of the CCD, used in these experiments, the sensor converts the number of electrons into gray levels (expressed in *Digital Number*, $DN$) through the linear relation

$$DN = \frac{n_e}{K_{ADC}} \qquad (3.5)$$

where $K_{ADC}$ is the conversion constant. The value of the constant depends also on the number of bits by which $DN$ is given.

Unfortunately, the detector also introduces noise. Noise in imaging sensors comes from the sum of four contributions: shot noise, Fano noise, fixed pattern

noise and read noise, as described in Ch. 1. In the following, only the first contribution, the shot noise, is considered.

The sensors considered in this work are characterized by three photon transfer relationships: irradiance [W/m²] versus signal [$DN$], irradiance [W/m²] versus shot noise power [W] and irradiance [W/m²] versus SNR.

**Reconstruction**

Reconstruction (third block of Figs. 3.2(a) and 3.2(b)) should enhance the details of the image removing the effects of the used mask. Reconstruction shall then be a sharpening of the black-white transitions, blurred by the low-pass transfer function of the electro-optical system.

Among the many different image reconstruction algorithms proposed in the literature [47], Wiener filters have been chosen since such filters balance reduced computational complexity and good effectiveness: they are in fact linear and make use of the knowledge of the transfer function to be inverted, of the signal to be reconstructed and of the noise affecting the image.

Given a linear space-invariant system with impulse response $h(x)$, the output signal $s_{out}(x)$ is:

$$s_{out}(x) = s_{in}(x) \star h(x) + n(x) \tag{3.6}$$

where $s_{in}(x)$ is the input signal, $n(x)$ is the additive noise and "$\star$" denotes convolution. In the spatial frequency domain, $f$, this expression is written using the spectra of the signals and the transfer function $H(f)$ of the system:

$$S_{out}(f) = S_{in}(f) \, H(f) + N(f). \tag{3.7}$$

The reconstructed signal is found applying a filter to the output according to the formula:

$$S_r(f) = R(f) \, S_{out}(f) = R(f)\Big(S_{in}(f) \, H(f) + N(f)\Big). \tag{3.8}$$

The purpose is to find the reconstruction filter $R(f)$ that gives the best estimation of the input signal knowing only the output one. Without any noise ($N(f) = 0$), this filter is a simple inversion of the system transfer function: $R(f) = 1/H(f)$.

The Wiener filter provides the best estimation of the signal $S_{in}(f)$ in presence of noise, according to the Minimum Mean Square Error (MMSE) criterion:

$$W(f) = \arg \left\{ \min_{R(f)} \left\{ \mathbb{E} \left| S_r(f) - S_{in}(f) \right|^2 \right\} \right\} \tag{3.9}$$

where $\arg$ stands for the argument of the minimum and $\mathbb{E}$ is the mean function.

Substituting Eq. (3.8) in (3.9) and differentiating the result to find the minimum, the expression of the filter becomes:

$$W(f) = \frac{1}{H(f)} \cdot \frac{|H(f)|^2}{|H(f)|^2 + \dfrac{N(f)}{S_{in}(f)}}. \tag{3.10}$$

The leftmost fraction in this equation takes into account the inversion that has to be made if there is no noise, while the rightmost fraction avoids the enhancement of frequencies where the noise is the most important contribution.

A more general version of Eq. (3.10) considers also a target function and, in a 2D reference frame, is given by:

$$W_{2D}(f_x, f_y) = \frac{H_{target}(f_x, f_y)}{H(f_x, f_y)} \cdot \frac{|H(f_x, f_y)|^2}{|H(f_x, f_y)|^2 + \dfrac{N(f_x, f_y)}{S_{in}(f_x, f_y)}} \qquad (3.11)$$

where:

- $H_{target}(f_x, f_y)$ is the target transfer function. In the most simple case, it is given by a two-dimensional window function (for example a raised cosine window) with a carefully chosen cutoff frequency.

- $H(f_x, f_y)$ is the real transfer function including the sensor sampling effect. It is given by:

$$H(f_x, f_y) = H_{opt}(f_x, f_y) \left| \mathrm{sinc}\,(w f_x)\,\mathrm{sinc}\,(w f_y) \right| \qquad (3.12)$$

  where $H_{opt}(f_x, f_y)$ is the two-dimensional transfer function of the real optical system and $w$ is the pixel pitch of the sensor.

- $N(f_x, f_y)$ is the noise power spectrum.

- $S_{in}(f_x, f_y)$ is the signal spectrum before the reconstruction filter.

The filter is then normalized to have the central value $W_{2D}(0, 0)$ equal to $1$, so that it does not change the mean value of the image.

As a useful simplification, both the power spectrum of the noise and of the signal are assumed to be flat over the relevant bandwidth of the system. In this case one gets:

$$\frac{N(f_x, f_y)}{S_{in}(f_x, f_y)} = \frac{N}{S_{in}} = \frac{N}{A S_O} \qquad (3.13)$$

where $S_O$ is the signal power on the object plane and $A$ is the power attenuation between the object and the image plane of the optical system.

Ideally, the Wiener filter used to reconstruct an image should match the transfer function of the optoelectronic system that generates it, since reconstruction using non-matched filters generates artifacts that may impair the decoding process.

However, in the considered case the transfer function is closely related to the OTF, which creates two problems that prevent the use of a perfect matched filter to reconstruct all the images:

1. The system transfer function depends on the distance of the object plane. In principle, one should then have a matched filter for each considered distance. Since the system is designed to reduce the dependence of the OTF on the object plane distance, only a limited number of filters is necessary (usually 5-10).

2. The system transfer function depends also on the field that is considered. The most general reconstruction algorithm should be implemented as space-variant. For reduction of the computation complexity, the reconstruction has been implemented as space invariant, and only the central field has been considered for the creation of the kernel of the Wiener filter.

In the following, for the sake of brevity, the planes where the filters are calculated will be referred to as "Wiener planes". Different criteria can be used to choose the positions of the Wiener planes. For example they can be located observing the rate of variation of the MTF with the object distance or choosing positions that avoid the presence of spurious artifacts during the optimization. The choice of the number and the position of the Wiener planes has an effect on the reconstruction quality and on the overall decoding performance and must then be done with care, depending on the situation. Alternatively the Wiener plane positions can also be considered as variable during the optimization of the whole system.

Once a set of Wiener planes is defined, the actual filter used for the reconstruction is evaluated on the Wiener plane the F-number of which is the most similar to that of the object plane.

In the next sections, O-SLALOM and A-SLALOM will be described.

### 3.1.1 O-SLALOM : System Level Optimization tool

The purpose of O-SLALOM is to evaluate the SFR and the SNR of the image after it has passed through the entire system and to compare them with the requirements for decoding given by the decoding library. This process is repeated optimizing the parameters until they achieve the target values.

To evaluate how good the system performance is (fourth block of Fig. 3.2(a)), the differences between the SFR and SNR obtained after the system and the ones required by the decoding library are used:

$$MTF_{diff} = SFR_{eval} - MTF_{spec} \tag{3.14}$$

$$SNR_{diff} = SNR_{eval} - SNR_{spec} \tag{3.15}$$

where $SFR_{eval}$ is the value of the SFR at the frequency given by the specifications and $SNR_{eval}$ is the SNR of the reconstructed image and does not depend on the spatial frequency.

The SNR is defined as (Fig.3.3):

$$SNR = \frac{S_{out}}{\sigma_{out}} \tag{3.16}$$

where $S_{out}$ is the difference between the mean high gray level and the mean low gray level and $\sigma_{out}$ is the maximum standard deviation of the noise. Because of the dependence of the standard deviation on the mean value of the signal, typical of Poisson processes, $\sigma_{out}$ always corresponds to the standard deviation of the noise in the white part of the image.
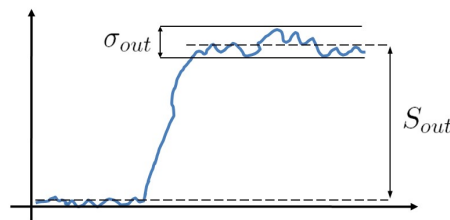


Figure 3.3: SNR definition.

The specifications are given for various barcode resolutions. For each requirement the differences are calculated by the C program and given to ZEMAX® as the results of the simulation. All these differences have to be kept positive and maximized in order to obtain the best performance of the optics.

O-SLALOM implements the model of the system in two ways according to the domain in which calculations are performed: the *spatial domain* and *spatial frequency domain* or, simply, *frequency domain*.

In the following their general features will be described.

**Spatial domain model**

The spatial domain model refers to simulations done working directly on images, as sketched in Fig. 3.4. The output image is the result on the input image of the cascaded effect of each block.

Knowing the position and the size of the input image (a slanted edge), ZEMAX® creates a bitmap version of the edge as it appears after passing through the optical system, using the so called "Image Simulation" tool. Also the sensor sampling



Figure 3.4: Simulation in the spatial domain.

Figure 3.5: Simulation in the spatial frequency domain.



Figure 3.6: Scheme for the SNR evaluation in the frequency domain. $W_{2D}(f_x, f_y)$ is the Wiener filter used in the reconstruction.

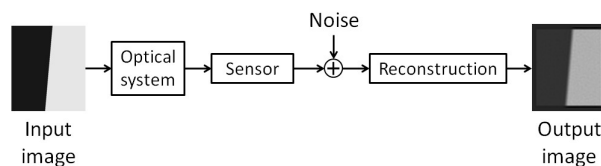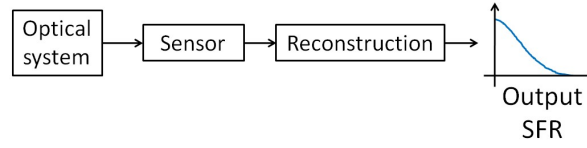effect is modeled by ZEMAX$^{\circledR}$ , while the noise is added by the C program. The reconstruction is then done with a convolution between the noisy image and the kernel of the Wiener filter, calculated as the Inverse Fourier Transform of the filter defined in the frequency domain. The SFR is found using SAFARILAB (see Ch.1) on the edge image. The SNR is directly calculated on the image according to Eq. (3.16) and does not depend on the analyzed frequency.

Even though the approach in the spatial domain is the most close to reality, it involves operations on images (e.g. convolutions, use of SAFARILAB , etc.) that slow down the optimization process. Moreover, working in the spatial domain can cause problems related to the correct sampling of the functions that describe the optical system.

**Frequency domain model**

The spatial frequency domain model follows the scheme shown in Fig. 3.5.

It does not handle images directly since it describes the system using its transfer function. For example the optical system is completely characterized by ZEMAX$^{\circledR}$ using the OTF. The effect of the sensor is taken into account by the C-based program that adds the appropriate spectral folding. The reconstruction is then done by direct multiplication in the frequency domain between the obtained transfer function and the Wiener filter. In this way the SFR of the whole system is obtained without using convolutions and SAFARILAB .

A critical point of this approach may be the model of noise to be used to evaluate the SNR with Eq. (3.16). Considering the scheme of Fig. 3.6, it is possible to find $S_{out}$ and $\sigma_{out}$ following these steps:

- The signal $S_{in}$ is the difference between the high and low gray levels entering the filter. These numbers are calculated from the illumination on the object plane using the characterization of the sensor (the signal $S_{out}$ after the reconstruction is equal to the signal $S_{in}$ before it since the Wiener filter

$W_{2D}(f_x, f_y)$ is normalized to be equal to $1$ at zero frequency). In formulae:

$$S_{out} = S_{in}|W_{2D}(0,0)|^2 = S_{in} = DN_{high} - DN_{low} \qquad (3.17)$$

where $DN_{high}$ and $DN_{low}$ are the digital numbers corresponding to high and low gray levels of the signal.

- The noise present in the image is described by a Poisson random process or a Gaussian one depending on the small or large number of electrons arriving on the sensor. If the noise power spectrum is flat over the bandwidth relevant to the system, an Additive White Gaussian Noise (AWGN) can be considered. In this case the variance $\sigma_{out}^2$ of the signal after the filter is found as [48]:

$$
\begin{aligned}
\sigma_{out}^2 &= \mathcal{N}_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |W_{2D}(f_x, f_y)|^2 \, df_x \, df_y \approx \\
&\approx \mathcal{N}_0 \sum_{k=1}^{N_x} \sum_{h=1}^{N_y} |W_{2D}(f_{xk}, f_{yh})|^2 \, \Delta f_x \, \Delta f_y
\end{aligned}
\qquad (3.18)
$$

where $\mathcal{N}_0$ is the power spectrum density of the AWGN that enters into the reconstruction filter and $W_{2D}(f_x, f_y)$ is the Wiener filter used in the reconstruction. $\mathcal{N}_0$ is known from the characterization of the sensor and is given by:

$$\mathcal{N}_0 = \frac{\sigma_{in}^2}{(N_x \, \Delta f_x)(N_y \, \Delta f_y)}$$

where $\sigma_{in}^2$ is the variance of the noise entering the filter.

### 3.1.2   A-SLALOM : System Level Analysis tool

The purpose of A-SLALOM is to create an image of the chosen input as it appears after passing through the entire electro-optical system. Even if the idea behind this tool is more general, it has been specialized for simulating barcode images. The sequence of operations is described as follows.

Firstly, the optical system is characterized through its PSF, calculated as the inverse Fourier Transform of the OTF provided by ZEMAX® . The image after the optical system is found as the convolution between the PSF and the barcode input image, created directly by A-SLALOM knowing the code resolution, position and template (the sequence of bars and spaces). Then, the sensor sampling effect and the noise are added by the C program. Finally, the reconstructed image is found as the convolution between the noisy image and the kernel of the Wiener filter, evaluated in the same way as for O-SLALOM .

In order to allow the study of the reconstruction filters and the tuning of their parameters, both the images before and after the reconstruction step are saved as outputs of the program.

Even if the scheme of the program is the same shown in figure 3.4, many differences are present concerning how the output image is evaluated. In A-SLALOM the convolution between the input image and the PSF of the optical system is done directly by the C program rather than by ZEMAX® . The use of a FFT-based algorithm to evaluate the PSF — which is not implemented by the ZEMAX® image simulation tool — allows to balance accuracy and speed even if a high degree of aberrations is present. Moreover, in the C program the simulation parameters are automatically chosen, whereas they have to be set manually in ZEMAX® : this makes A-SLALOM suitable for many subsequent automated simulations. On the other hand, A-SLALOM generates images of linear barcodes assuming space-invariance, absence of distortion and constant illumination over the field of view.

## 3.2   Spherical aberration system design using SLALOM

After describing the features of SLALOM , this section describes its use and the results of the optimization and analysis of a system with a high degree of spherical aberration.

Firstly, some preliminary parameters related to the reconstruction step are found using A-SLALOM and comments on the problem of reconstruction are reported.

Then, the design procedure is illustrated: the starting designs are described together with the merit functions used to optimize them.

Finally the optimized designs are shown and their performance discussed. Note that all the results are found using O-SLALOM in the frequency domain, since this allows a faster and more accurate optimization.

### 3.2.1   Tuning of the reconstruction filter parameters

Before running the optimization tool, the values of some parameters of the reconstruction filters have been set by using A-SLALOM . They are: the number of points for sampling the Wiener kernel used to reconstruct the image, the cut-off frequency of the window chosen as target function and its roll-off factor (if a raised cosine window is used). These parameters have been fixed before performing the overall system optimization in order to speed up calculations. The simulation for estimating the reconstruction filter parameters has been done considering an optical system with a high degree of spherical aberration by simulating codes with different resolutions at various object distances. Many values of the previously cited parameters have been considered and the best ones have been found using information about the decodability of the obtained images, trying also to balance efficiency and accuracy. All the results are found assuming the

object to be at the Wiener planes in order to avoid artifacts that may impair the results.

The obtained values are not reported since they depend on the considered optical system: they are not the optimum ones for any configuration, but only the best candidates for the investigated one. Different configurations will require another tuning process.

This study gives the opportunity to make a brief analysis of the effect of reconstruction based on the found optimal values.

All the results are obtained for a lens triplet with a high degree of spherical aberration optimized in the frequency domain using O-SLALOM . The Wiener planes are located at the distances where the specifications are given, in order to avoid spurious effects during the optimization.

Fig. 3.7 shows the decodability as a function of the object distance for codes with different resolutions, both for noisy images (blue line) and reconstructed images (red line).

In all cases the reconstruction has a positive effect:

- in Fig. 3.7(a) barcodes are correctly read also between $20\mathrm{mm}$ and $40\mathrm{mm}$;

- in Fig. 3.7(b) the depth of field increases from $50\mathrm{mm}$ up to $90\mathrm{mm}$;

- in Fig. 3.7(c) one finds that all the noisy images cannot be decoded, whereas the reconstruction allows a depth of field of $190\mathrm{mm}$ between $110\mathrm{mm}$ and $300\mathrm{mm}$ (even if some problems are still present between $200\mathrm{mm}$ and $220\mathrm{mm}$).

These data show that the reconstruction has a large positive impact on the decodability of the images independently of the object distance. In real cases, however, the decoding capability depends on the amount of artifacts introduced by the reconstruction, which in turn depends on the matching between the Wiener filter actually applied and the OTF of the system at the relevant object distance.

However, it is important to notice that at least one decodable image for each distance may be enough: the real depth of field of the system is given by the union of the distances where the noisy and the reconstructed images are decoded.

### 3.2.2   Starting designs

A major assumption taken before design was that of starting not from scratch with totally new lenses but draw on an already existing design, which has a limited depth of field that has to be extended and adapted to the purpose of this work, enhancing the lens spherical aberration. Since this aberration goes like the $4^{th}$ power of the lens radius, the design needs at least one aspherical surface with non-zero $4^{th}$ order term. This can be done by modifying the lenses of the original design or by adding new optical elements to it.

In Fig. 3.8 the system scheme is shown. The original lens is different from the one used in Ch. 2. It is composed by three lenses and one of them is aspherical.
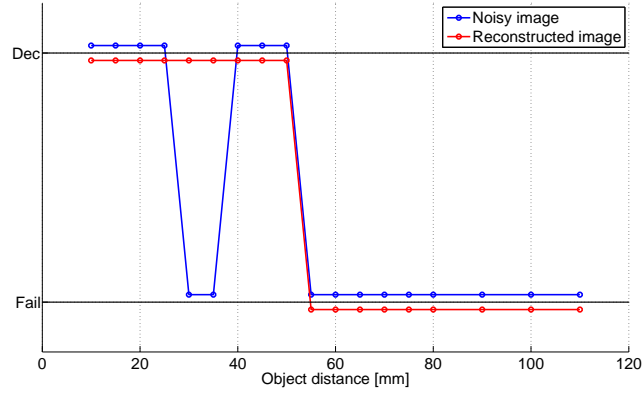
(a)
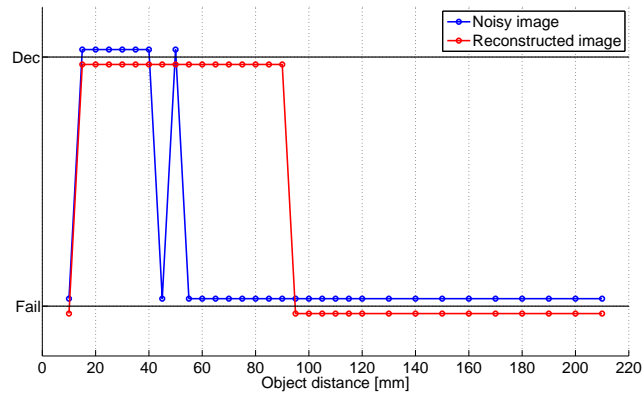
(b)
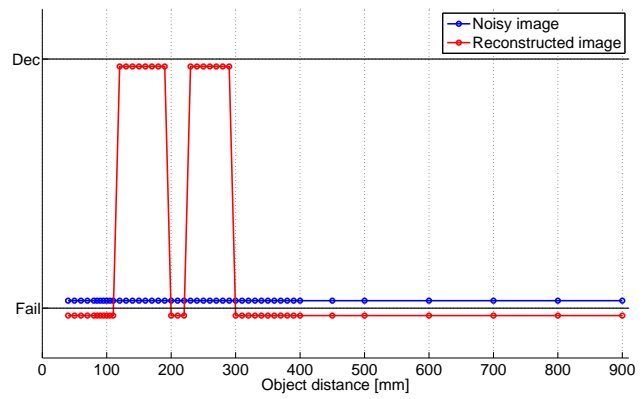
(c)

Figure 3.7: Decodability of different code resolutions for the noisy (blue line) and reconstructed (red line) images.

Figure 3.8: Representation of the two system design possibilities.

The first design option consists in modifying a surface of one of the three lenses, as shown in red in Fig. 3.8. The second design has been created by adding an aspherical lens, as shown in blue in Fig. 3.8. Both possibilities are described in the following.

**First design**

The most immediate solution is the modification of one surface of the original lens, so that it can generate the spherical aberration needed for EDoF.
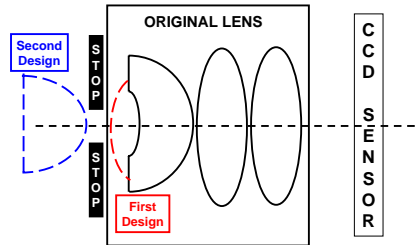
The surface that can mostly contribute to the total spherical aberration is the front surface of the aspheric lens, because of its position and of the high angle of incidence of the incoming rays. Therefore, thickness, conic constant and $4^{th}$ order term of the lens front surface were set as variable together with the distance between this aspheric lens and the followed glass one. Also the BFL is considered as a variable.

**Second design**

Another possible design adds a surface with non-zero $4^{th}$ order term to the original lens. Since a high degree of spherical aberration is needed, an aspheric lens (implemented in ZEMAX® as an "even asphere" surface) has been considered.

An important decision is about the lens material. Being aspheric, it is easier and cheaper to manufacture it in plastic instead of glass. Key factors are Abbe number and index of refraction. A lower Abbe number corresponds to increase the chromatic aberration. A lower index of refraction helps to avoid total reflection and high angle of transmission/incidence. Nine plastic materials were studied and the best one was found to be polycarbonate.

The last decision concerns the location of the polycarbonate aspheric lens. As long as it may be desirable that the structure of the starting lens is not modified, a new optical element can be introduced only before or after the external barrel.

Since the stop is the entrance hole of the barrel itself and the angle of incidence of incoming rays is greater, the best location for the new lens is before the barrel.

Thickness, $4^{th}$ order term of the polycarbonate lens back surface together with distance between this lens and the barrel are set as variable. Such as in the previous design, BFL is variable too.

### 3.2.3   Optical pre-design using OLOM

Once the geometries have been chosen, the values of their parameters are the result of the optimization.

Before running the O-SLALOM process that permits to optimize the all electro-optical system, it is better to pre-optimize the optical part of the considered systems alone. The results will be the starting points for the O-SLALOM optimization. This is possible thanks to the OLOM tool described in Sec. 2.2.

Two different Merit Functions (MFs) have been implemented ad-hoc for each design.

The specifications of the merit functions that involve the optical performance, include:

1. Optical system constraints such as spherical aberration, that should be increased, relative illumination, maximum distortion, Half Field of View (HFoV);

2. Fabrication constraints in order to control the mechanical constraints that lenses and other glass surfaces need to satisfy to be feasible;

3. MTF constraints, set by OLOM , to satisfy the specifications about the distances where each barcode needs to be read;

The first three types of constraint are the same used in the next section added to the O-SLALOM specifications.

Considering the layouts of the two starting designs, distinct MFs have been written and different optimization processes have been run. This is the direct consequence of the fact that the two lenses have not the same number of components.

In the following, the main similarities and differences between the two merit functions will be illustrated.

**Optical system constraints**

For both designs, it is necessary to include in the MF the following optical system constraints:

1. **Spherical aberration.** Its measurement unit is given in wavelengths and the limits which bound the allowed range are $5$ and $10$. These values have been suggested by experience and have been confirmed by optimization

results as able to guarantee both enough defocus for EDoF and successful decoding.

2. **Relative illumination.** Calculated at the primary wavelength for the outer field, it must be larger than a fixed threshold.

3. **Maximum distortion.** Also this parameter is calculated at the primary wavelength for the outer field and must be limited under a fixed threshold.

4. **Half Field Of View (HFoV).** It is the radial max field at primary wavelength. The barcode reader is required to have a specific HFoV.

The weights of all these optimization variables differ and must be carefully chosen. The weight of an operand is in fact a measure of the importance of that operand relative to the others. During the optimization process, the weight of the same operand has been changed according to their tendency to violate the constraints.


**Fabrication constraints**

The fabrication constraints are important for controlling the mechanical part of the system and its feasibility.

They are different for the two designs, because they have different optical component to fabricate.

The fabrication constraints are:

1. **Practical feasibility.** The structures obtained by optimization must be fabricated. Radius, thickness and sag as well as the distance from the next element of the spherically aberrated surface may differ and their values may prevent or allow fabrication.

2. **Length of the optical barrel.** It is the distance from the first to the last surface of the three lenses. It is an important device mechanical constraint though it plays a practical role only in the first lens since the second one has a fixed length.

3. **Geometric constraints.** They are different since the structures are different (the number of lenses is not the same, for example). Referring to the first design, constraints on central and edge thickness are set on the first lens of original lens while air-space thickness between lenses has to be positive and lower than a maximum value dependent on the optical barrel length. Referring to the second design, the same constraints are set on the lens added before the barrel while air-space thickness between the lens and the barrel has to be positive and lower than a fixed maximum value.

Also for these constraints the weight choices has to be done accurately.

**MTF constraints**

In order to find the highest value of the MTF for a system that is as invariant with defocus as possible, the constraints on the MTF are set using OLOM , as described in the previous chapter (Sec. 2.2).

The used wavelengths and fields are related to the system to be optimized, whereas frequencies and object distances to be considered depend on the required specifications.

In particular the frequencies ($f_I$) to be considered in OLOM are calculated in the way described in Sec. 2.2 using Eq. (2.5). However, other frequencies have been added in order to take into account also the Print Ratio (PR) of the code. They are calculated as $f_I$/PR.

After setting all these constraints, the variables to be optimized have been chosen. They are: referring to the first design the first lens surface and the BFL — and referring to the second design the second surface of the added aspheric lens and the BFL.

The results of this pre-optimization process are used in the following as starting points for the system design using O-SLALOM .

## 3.2.4   System level design using O-SLALOM

Thanks to the previous optimization, a good starting point has been found even if it can not satisfy all the system constraints. The purpose of O-SLALOM is to optimize the whole system considering not only the optical part but also the others.

In the following the results of the optimization will be described. The Wiener planes have been placed at the distances where the specifications are given in order to avoid spurious effects during the optimization, while the parameters used in the reconstruction are those already found using A-SLALOM .

The MF used during the optimization is the same constructed before, substituting the OLOM part with the one related to O-SLALOM . In this case the specification table considers a combination of MTF and SNR constraints.

**First design**

The optimization has made the first lens of the starting design thinner, with a different shape and has moved it away from the second lens.

The most important limiting factors in reducing the merit functions have been found to come from the O-SLALOM constraints on MTF and SNR.

Results concerning MTF are shown in Fig. 3.9 which reports the variation of optical system MTF versus normalized spatial frequency as a function of the object distance. The black arrow highlights that the MTF curves tend to a tight bun-
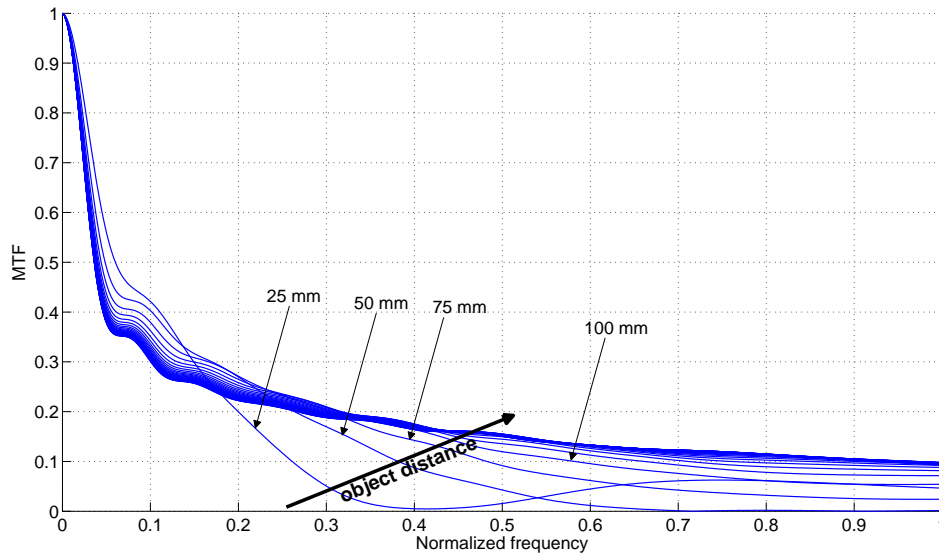
Figure 3.9: MTFs of the first optimized designs evaluated for various object distances between 25mm and 750mm.

dle (without zero points) as the object distance increases. The behavior desired to obtain EDoF is then basically achieved.

The performance of the whole electro-optical system can be evaluated by using A-SLALOM that simulates the barcode reading. As an example, Fig. 3.10 shows the decodability of a set of barcode images with a given resolution. Part (a) refers to the starting design, part (b) refers to the optimized one. Blue and red lines correspond, respectively, to images before and after reconstruction. In this case a significant extension of the DOF is obtained.

**Second design**

In this case spherical aberration turns out to be properly controlled. Also for this design, the O-SLALOM constraints on MTF and SNR are the most important limiting factors to the reduction of the merit function.

Fig. 3.11 shows the variation of optical MTF versus normalized spatial frequency as a function of the object distance. In this case, the curves do not have zero points, which is advantageous for the system. The black arrow also highlights that as the nose distance increases, the MTF curves tend to a tight bundle. The overall behavior is then similar to that obtained in the first design, with the further advantage of the absence of zeroes in the MTF at any distance.

The performance of the whole electro-optical system has been again evaluated using A-SLALOM and an example is shown in Fig. 3.12.. As for the previous analysis, the top figure refers to the starting design whereas the bottom one refer
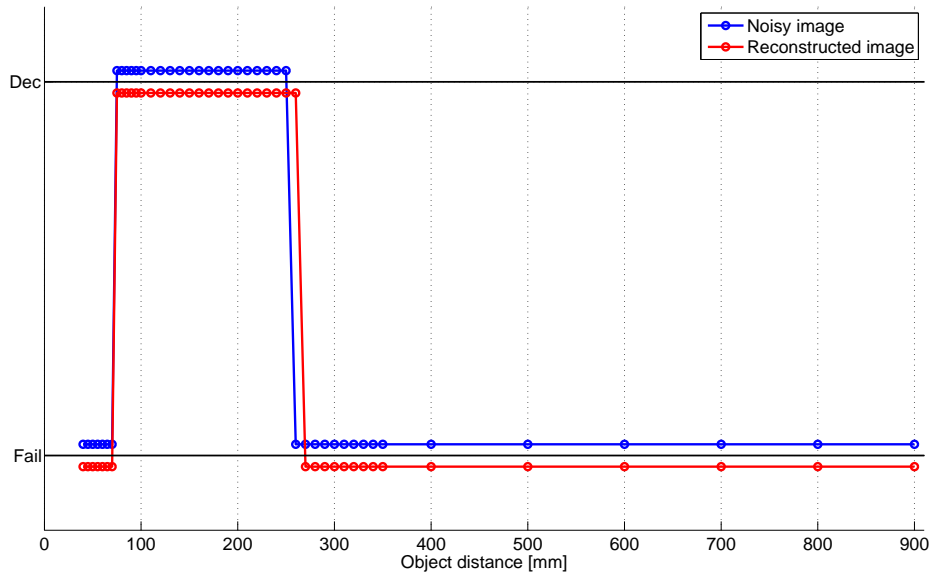
(a)
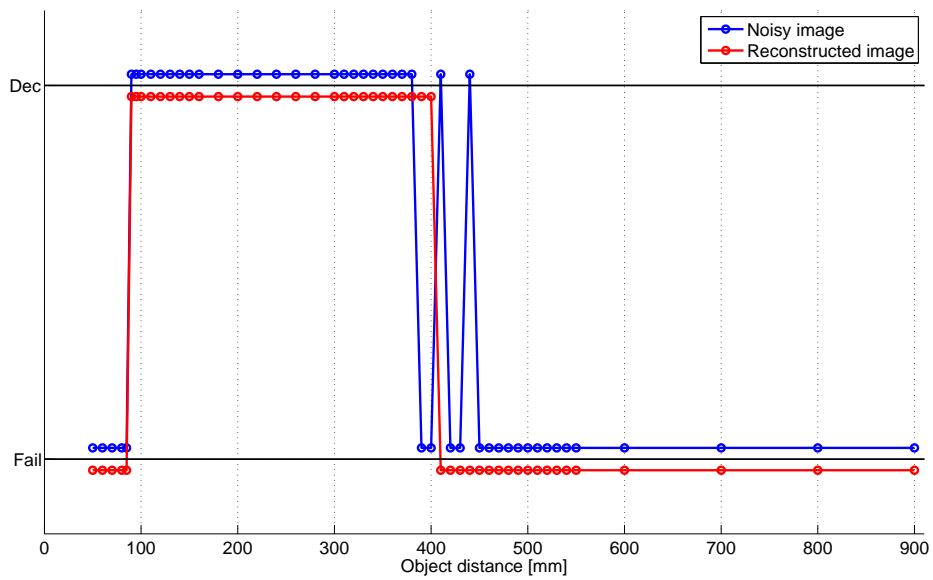


(b)

Figure 3.10: Decodability as a function of the object distances for the first starting design (a) and for the first optimized design (b).

Figure 3.11: MTFs of the second optimized designs evaluated for various object distances between $25$ and $750$mm.

to the optimized design, the blue lines correspond to images before reconstruction while the red lines refers to images after reconstruction. Also in this case an extension of the DOF is demonstrated.
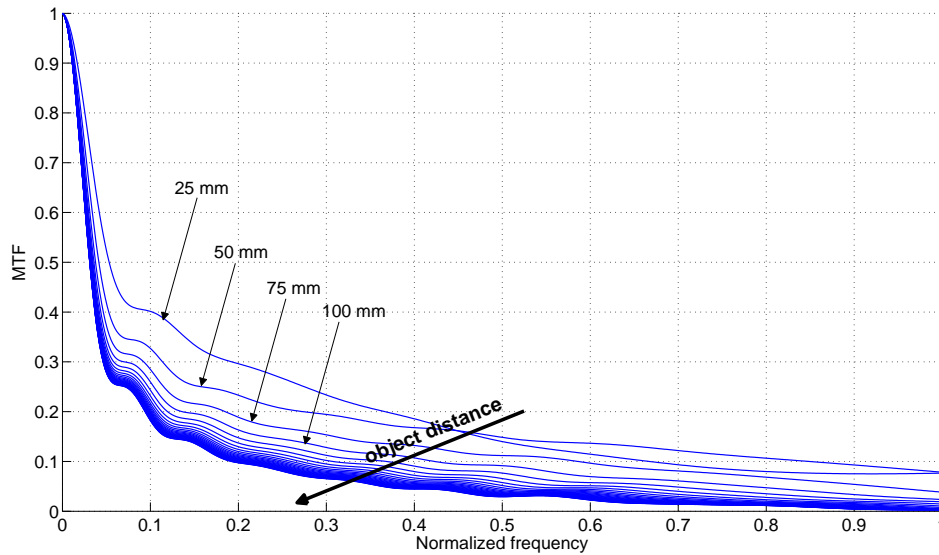
## 3.2.5   Merit function sensitivity to optimized parameters

In order to validate the results obtained by means of the optimization just described, including fabrication, spherical aberration and O-SLALOM constraints, an analysis of the merit function sensitivity to the optimized parameters has been made, considering the second optimized design.

This is useful both to verify that the final merit function has been optimized to an absolute minimum and to investigate the system sensitivity to the variable parameters in view of its possible fabrication.

The system parameters that were varied during the analysis are the thickness of the polycarbonate lens ($THIC_2$), the $4^{th}$ order term of its back surface ($PAR2_3$), its distance from the barrel ($THIC_3$) and the back focal length ($BFL$).

Each parameter is sampled in nine equidistant points in the range $\pm 10\%$ off the optimal values. Fig. 3.13 shows how the merit function changes varying only one parameter at a time while the other three are fixed at their optimized value.

Fig. 3.13(a) shows that the dependence of the merit function on the lens thickness is small, since the curve is almost flat. This loose dependence shows that the polycarbonate lens thickness is not a critical parameter for the system though it can represent a problem for the optimization tool, because it runs into more
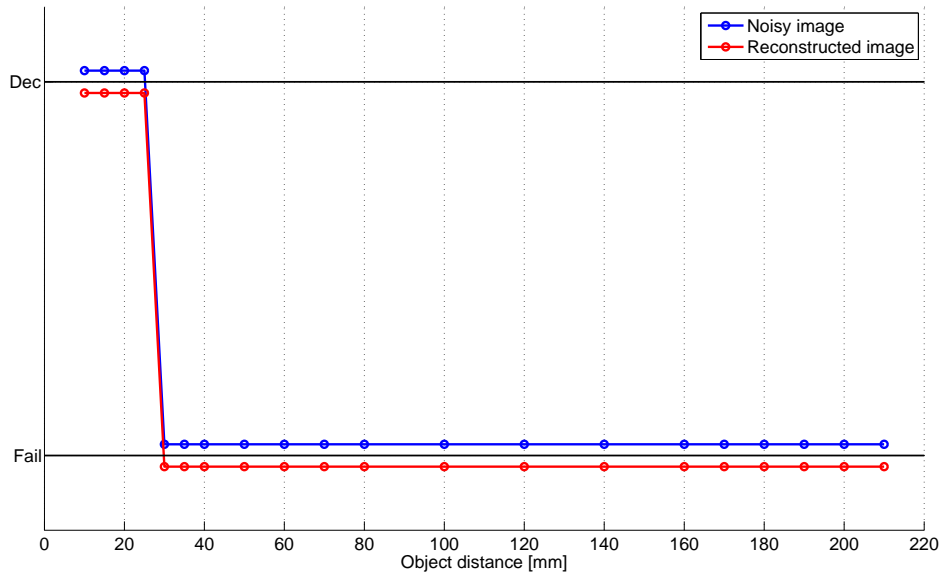
(a)
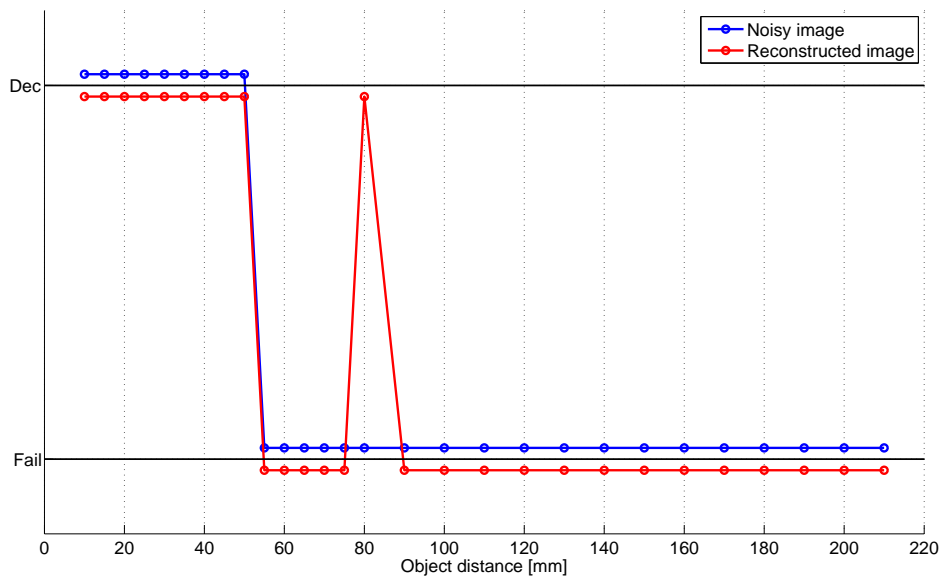


(b)

Figure 3.12: Decodability as a function of the object distances for the second starting design (a) and for the second optimized design (b).
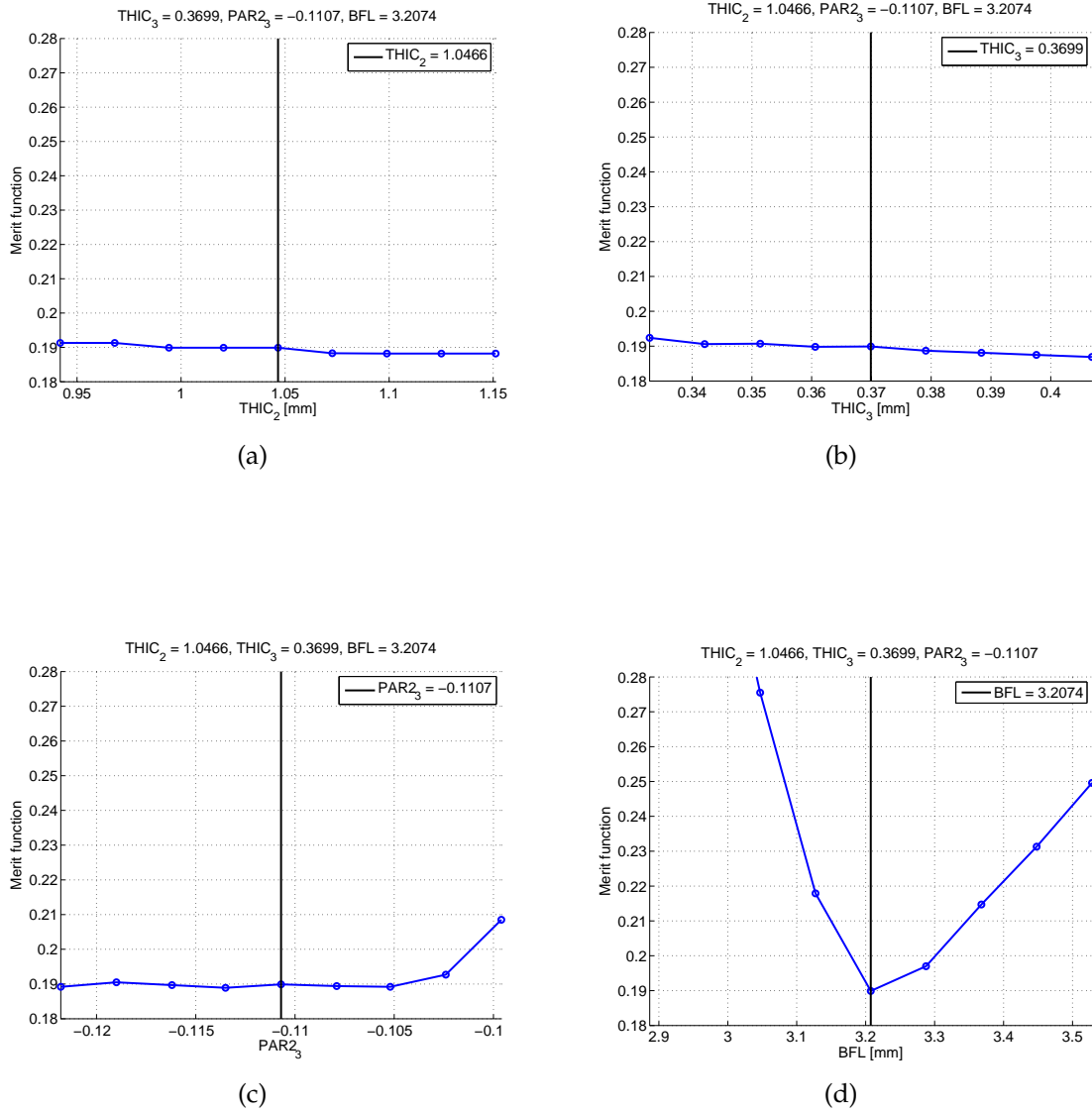
Figure 3.13: Merit function sensitivity to: (a) thickness of the first lens, (b) distance of this lens from the stop, (c) $4^{th}$ order term and (d) back focal length.
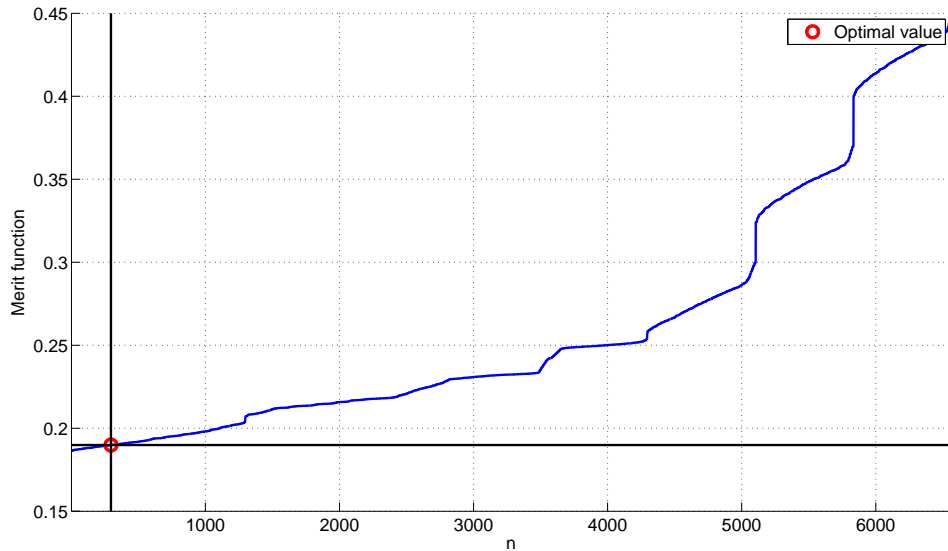
Figure 3.14: MF of optical systems obtained from 6561 different combinations of the optimized parameters.

difficulties in finding a minimum.

Similar conclusions are valid for the distance between the lens and the barrel, see Fig. 3.13(b) and for the $4^{th}$ order term of the lens back surface, see Fig. 3.13(c): these parameters are not crucial for the system performance. However, the variation of the lens sag needs a further comment. Shifting the $4^{th}$ order term toward zero, the merit function tends to rise. This behavior is correctly explained by theory. Making the $4^{th}$ order term vanish means decreasing the total amount of the system spherical aberration, with the resulting fall of the MTFs and the violation of required specifications. That implies the increase of the merit function.

Fig. 3.13(d) shows that the back focal length is the most critical parameter among those taken into account. The optimized value is a minimum of the curve, and this validates the optimization result. Increasing this value by $10\%$ results in a merit function increase larger than $30\%$, whereas decreasing it by $10\%$ makes the merit function rise more than $120\%$. Note that the BFL is a compensator parameter typically adjusted during the assembly.

Fig. 3.14 shows a summary of the analysis just described, considering the 6561 different optical systems obtained from the combinations of the four parameters. The merit function of the second optimized design, circled in red, is actually among the lowest values. Nothing can be said about the feasibility or the reading performance of the systems with lower merit function, so the second optimized design has been considered to be a very good one on the whole.

## 3.3   Conclusion

In this chapter, the general features of SLALOM tool for the optimization and the analysis of an imaging system have been presented and results of its application have been described.

In the first part of the chapter, the structure of SLALOM has been explained with respect to both the optimization task performed by O-SLALOM and to the analysis task performed by A-SLALOM . The two tools leverage on an extensive library of C modules that are implemented for the modeling of optical systems (in conjunction with ZEMAX$^{®}$ ), photosensors (taking into account pixel sampling and shot noise generation), and image restoration (via digital Wiener filtering). These modules allow the use of ZEMAX$^{®}$ beyond the optical domain.

In the second part of the chapter, the use of O-SLALOM for the optimization of a system based on the presence of a large spherical aberration have been presented. Firstly, the reconstruction filter parameters to be used in the optimization process have been found using A-SLALOM . Thanks to OLOM tool described in Ch. 2, a good starting point for the two designs considered has been found.

Finally, the two lenses pre-designed have been optimized by O-SLALOM . The results obtained after the optimization process have been considered. The features and the ability of both designs to extend the depth of field with respect to the corresponding starting designs have been validated the SLALOM tool in both its parts. Moreover, also the sensitivity of the merit function has been studied. Results show a low sensitivity to all the optimized parameters but the BFL: even if this low sensitivity can create problems during the optimization of the system because a sharp miminum of the MF is not present, it represents a good result taking into consideration that possible errors in the following fabrication step do not change the overall performance of the system.

In conclusion, results have shown that O-SLALOM can be successfully used for system-level design since it allows to obtain the almost absolute minimum of the merit function. However, results also show that spherical aberration, although capable of impacting on the lens DoF, does not allow to extend it so much as to meet all the required specifications. Furthermore, A-SLALOM is not only useful for analyzing the imaging systems and simulating their behavior but it also allows to find reconstruction parameter for the optimization process.

# Part II

# Non-Imaging Systems

# 4

# Illumination System Design

The illumination system is a non-imaging system that consists of an incoherent source (LED) and an optical part.

In this chapter, an analytical and simple tool for designing the optical part, such as a free-form lens, to achieve a custom illumination will be described. The term free-form refers to the shape of a lens that has more degrees of freedom than those allowed to traditional optical surfaces. For validating the tool, a free-form lens has been optimized according to two requirements: the illumination should be concentrated on the desired FoV and the edge of the illuminated zone must receive more light than the central one.

Being made out of a single lens attached to an LED source, such an illumination system can have sufficiently low weight and energy consumption to be employed in portable devices. The required illumination pattern is especially useful to imaging system because it compensates for both the target angular reflectivity and the angular response of the imaging system. For example, in a barcode reader, it helps the imaging system to have the proper SNR to successfully decode the barcode signal coming from the sensor. The free-form lens could be designed by this tool for satisfying also other specifications required by applications that need custom illumination.

After illustrating briefly the background of the free-form lens design problem (Sec. 4.1), the tool algorithm will be described (Sec. 4.2). In Sec. 4.3, the results of the free-form design will be shown. The mechanical and optical characterization of the fabricated lens will be presented in Sec. 4.4.

## 4.1   Free-form lens design background

In the last decade, the free-form lens design has become a hot topic especially for applications that need a flexibility that cannot be achieved by classical azimuth symmetric optical systems.

For non-imaging applications, free-form lenses are employed in combination

with high power LED sources [49] for street or ambient illumination [50], where uniform illumination is required on a given surface, or bar code reading [51], where the illumination pattern is tailored to diverse and non-uniform shapes.

Original procedures to design free-form lenses were proposed [52,53]. However, they are limited by the use of point sources and by assuming the illumination profiles to be uniform.

The inclusion of real sources that are not Lambertian [49,50], sets new challenges in writing the equations to be optimized in the design procedure and in studying source algorithms [54–57].

Fabrication introduces further problems such as alignment — whenever more than one curved surface is relevant to the design of the lens — and tolerances, whenever the designed lens has a complex shape, e.g. if it is not continuous [53].

A general design procedure should take into account the mentioned fabrication issues and also allow to target custom illumination patterns. A choice to tackle these problems is to use dedicated commercial software that can handle a large set of optical design problems (such as the widespread ZEMAX® [2], ASAP® [58], OSLO® [59], TRACEPRO® [60]). This solution may anyway be expensive when routine design of such lenses is not needed.

In the next section, a simple algorithm for the design of optical free-form surfaces is illustrated.

The tool, I have developed, has various attractive features: it can optimize any analytical function describing the lens surface shape; it has the potential to obtain any desired radiation pattern; it handles real sources such as high power LED sources rather than ideal uniform or Lambertian point sources, using a specific source algorithm. Finally, it was programmed in MATLAB®, which guarantees versatility and minimum programming effort.

## 4.2 Free-form lens design algorithm

In this section, the proposed algorithm is described.

A plano-convex free-form lens is considered. Letting one of the two surfaces of the device remain planar reduces the degrees of freedom available for design, but has a positive impact on the ease of fabrication and alignment, which is important in practice. The convex surface has been chosen to be described by a $6^{th}$ order polynomial:

$$z = f(x,y) = c_1\,x^2 + c_2\,y^2 + c_3\,x^4 + c_4\,x^2\,y^2 + c_5\,y^4 + c_6\,x^6 + c_7\,x^4\,y^2 + c_8\,x^2\,y^4 + c_9\,y^6 \quad (4.1)$$

where the nine $c_i$ coefficients are the unknowns to be determined by optimization. Lower or higher order polynomials can be considered, depending on design specifications.

The geometry of the structure is sketched in Fig. 4.1. Note that the flat surface is in front of the source plane, set at $z = 0$.
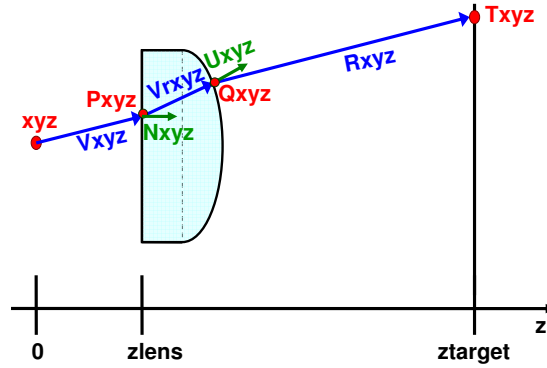
Figure 4.1: Definition of the free-form surface parameters.

In a cartesian reference frame with coordinates $i = x, y, z$, $V_i$ indicates the ray direction cosines out of the source, $P_i$, the coordinates on the free-form lens input plane, $V_{ri}$, the ray direction cosines within the thick lens, $N$, the direction cosines of the normal to the input surface, $Q_i$, the coordinates on the free-form lens output surface, $R_i$, the ray direction cosines out of the free-form lens, $U_i$, the direction cosines of the normal to the free-form surface lens and $T_i$, the coordinates on the final (target) surface.

The source to be considered is a real LED, modeled using ray tracing. Data sheets provided by manufacturers, include the so-called "ray-file", which contains the description of the source ray distribution pattern, i.e. the source angular irradiance distribution. In practice, the ray-file lists the coordinates of the point $i = x, y, z$ where each ray originates, its direction cosines $V_i$ and the ray flux $P$. This description of the source allows to consider any real source, with its own radiation pattern.

Since the ray-file generally contains from tens to hundreds of thousands of rays, it is necessary to reduce their number in order to allow a better performance of the implemented algorithm. In the following a possible procedure to do this will be explained.

One can first notice that the ray spatial distribution depends on the source radiation pattern and on the distance between the source and the lens. Some of the rays will then not impinge on the input surface of the lens and can immediately be eliminated. However, the majority still must be considered.

To further reduce their number, one can note that, if the phase fronts are smooth enough to guarantee that not too rapid changes in the direction cosines of adjacent surface points occur, rays can be grouped dividing the input surface of the lens in subsurfaces. Each group has a constant ray flux through the input and output sections. This way of working takes advantage of the ray optics hypothesis that rays are normal to the beam phase fronts. Each subarea is associated to

a set of direction cosines, calculated as the average of the direction cosines of all the rays impinging on that region.

Such grouping operation should consider also the shape of the lens input surface. In the considered case, however, this surface is flat, simplifying the operation. This source algorithm allows not only to increase the computational efficiency, but also to consider a real source as a lot of point sources, without reducing the algorithm reliability. In other words the results are not affected by using this new source instead of the real one.

A typical result is shown in Fig. 4.2 for a square lens with $2\text{mm}$ side. Subfigure (a) shows schematically how the source algorithm works. It is done subdividing the whole lens input area in vertical slices with the same overall intensity of rays (red points). Each slice is then partitioned in subareas following the same rule. For each subarea, the centroid (black point) is calculated and it represents all the rays impinging on that area.

Subfigure (b) shows the surface subdivided in bin, in color scale. The green points represent the centroids of each subsurface, as explained before. The color scale represents the central ray intensity (from white, maximum, to black, minimum). Each area is anyway associated with the same overall flux, this is confirmed by the fact that the darker areas are the largest ones. This way, a much smaller number of "equivalent" rays is considered, which is useful in the next optimization procedure. In the example, $10 \times 10$ ray bundles (or bins) are considered. The matrix size could be changed, if necessary.



(a)                                                             (b)
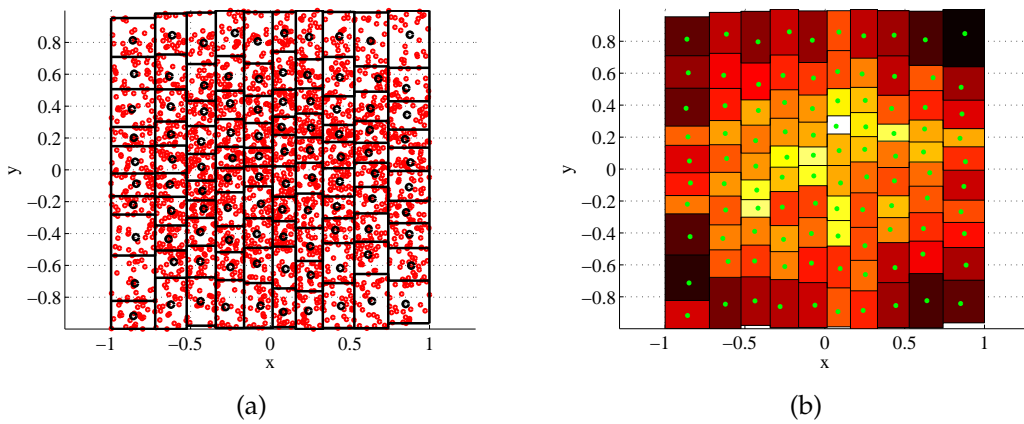
Figure 4.2: Example of input surface subdivided in bins with the same flux (coordinate $z = zlens$ of Fig. 4.1) for a square lens with $2\text{mm}$ side in two different representation, with the rays (red points) (a) and in color scale (b).

The algorithm is based on the assumption that the free-form surface creates an isomorphism between the lens input surface and the target surface. The target surface is subdivided into the same number of parts as the input one. On this surface each area is related to the expected local flux. For example, if uniform

target illumination is desired, the rectangular final surface shall be subdivided into equal parts. If not, subareas must be determined to fit the required specifications: the smaller the area, the larger the local expected flux. This way of setting the illumination pattern features guarantees wide flexibility of the algorithm. A few simple tests have been run preliminarily to tune the choice of the parameters (for example the number of bins to reduce the number of ray flux tubes) depending on the required intensity pattern distribution. After having decided how to subdivide the input plane (source) and the output one (target), it is necessary to define their relationship.

The ray trajectories from the input lens surface to the target plane are calculated using the vectorial Snell's law. This can be done considering the gradient of the surface

$$\nabla g(x, y, z) = \left( -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, +1 \right) \tag{4.2}$$

where $g(x, y, z) = z - f(x, y) = 0$ and $f(x, y)$ is the polynomial given by equation (4.1). Its knowledge leads to the direction cosines

$$U_x = -\frac{\frac{\partial f}{\partial x}}{||\nabla g||}, \qquad U_y = -\frac{\frac{\partial f}{\partial y}}{||\nabla g||}, \qquad U_z = \frac{1}{||\nabla g||} \tag{4.3}$$

being $||\nabla g|| = \sqrt{1 + (\nabla f)^2}$. The knowledge of this surface allows to calculate also the direction cosines of the normal to the surface.

The coefficients of the polynomial function are calculated minimizing the root mean square error between the real arrival points ($T_i, i = x, y, z)$) and the expected ones ($T_{i,exp}, i = x, y, z)$):

$$e = \langle \sqrt{(T_{x,exp} - T_x)^2 + (T_{y,exp} - T_y)^2 + (T_{z,exp} - T_z)^2} \rangle \tag{4.4}$$

where $\langle \cdot \rangle$ is the average of the argument.

The optimization procedure starts by assuming a parabolic surface (i.e. by setting to $0$ all the coefficients of Eq. (4.1) except those of $x^2$ and $y^2$) and calculating the related output ray position and slope on the target and the first value of the error. Then, all the coefficients are progressively varied at the following optimization iterations, until convergence is reached. The optimization is performed using Newton-like algorithms implemented in MATLAB® , which can find local minima of a user-defined function.

## 4.3   Free-form lens design

This section describes how the proposed algorithm is used to design a lens obeying to the following specifications: *i*) the free-form lens operates at the wavelength of $633\text{nm}$ and is fabricated in acrylic; *ii*) the field of view (FoV) along the

two axes should be $FoV_x = 35°$ and $FoV_y = 22°$; *iii*) at least $80\%$ of the ray flux entering the lens should be in the given FoV range; *iv*) the intensity distribution should provide corners illumination at least $30\%$ higher than that at the center of the target; *v*) the lens is at $0.8$mm from the external surface of the LED package; *vi*) the target is at $200$mm from the source plane.

The lens size should be large enough to collect as much power as possible, with possible size constraints set by mechanical (mounting) problems. In this case, a surface of $3.5$mm $\times$ $3.5$mm positioned at $0.8$mm from the LED package surface can collect about $70\%$ of the ray flux emitted by the LED used. This can be considered a satisfactory value. This percentage is an hundred times the ratio between the power ray flux impinged on the first lens surface and the total power LED flux.

However, internal reflections of rays impinging on the lateral lens surface are possible. To exclude them from calculations, a further guard ring ($0.5$mm in this case) around the lens has been considered (leading to an overall size of $4.5$mm $\times$ $4.5$mm and an overall ray flux of $80\%$ impinging on the first surface lens). The choice of using the guard ring is beneficial to reduce lens side scattering effects in fabricated devices, which have also a detrimental effect on performance.

In the following the results obtained after the lens optimization made by the proposed algorithm will be shown.

## 4.3.1 Optimization results

The division in bins of the target plane obeying the specifications obtained after the optimization process is shown in Fig. 4.3. To check the design, TRACEPRO®
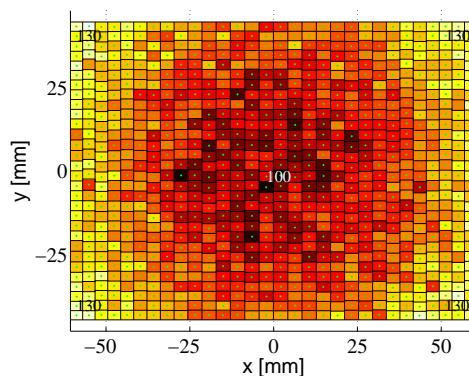


Figure 4.3: Division in bins of the target at $z = 200$mm to allow an illumination with $30\%$ power increase at the target corner compared to the center.

package to simulate the flux distributions that can be obtained by the overall system has been used. TRACEPRO® traces all the source rays from the input to the target plane. Fig. 4.4 shows the irradiance profile at $200$mm from the source

plane. One can see that, at the corners, the flux is larger than it is at the center. In
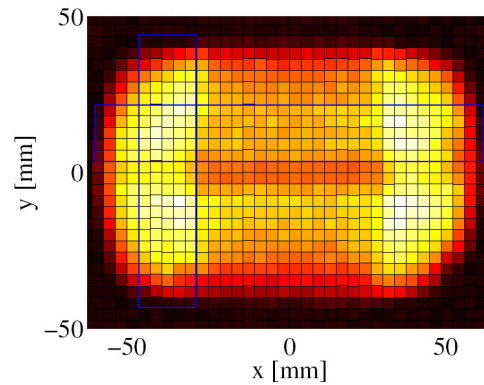


Figure 4.4: Intensity distribution at 200mm as resulting after reducing in bins the TRACEPRO® ray table.

particular, the horizontal (Fig. 4.5(a)) and vertical (Fig. 4.5(b)) irradiance profiles are shown. They are calculated as the average of the lines passing through the local maxima at the four corners, shown as blue rectangle in Fig. 4.4. The flux is normalized with respect to the average intensity of a central portion of the whole profile. One can see that the flux is more than 30% larger than that at the center, thus satisfying the specifications.



(a)



(b)
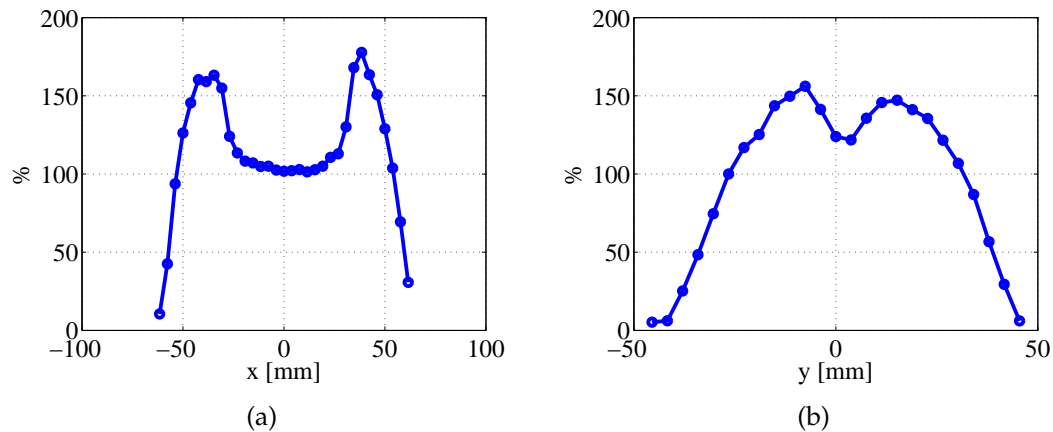
Figure 4.5: Normalized horizontal (a) and vertical (b) irradiance profiles at 200mm.

Finally, the specification on efficiency has been verified. It is attained, since 80% of the flux entering the lens does fall into the area defined by the specified FoV. This efficiency is calculated as the ratio between the power flux of the rays impinging on the target inside the specified FoV and the power flux of the rays entering the first surface lens.

## 4.4  Free-form lens characterization

A prototype of the designed lens has been fabricated using acrylic material. A picture of the sample is shown in Fig. 4.6.



Figure 4.6: Picture of one of the fabricated samples of free-form surface lens (actual size: $4.5\text{mm} \times 4.5\text{mm}$).

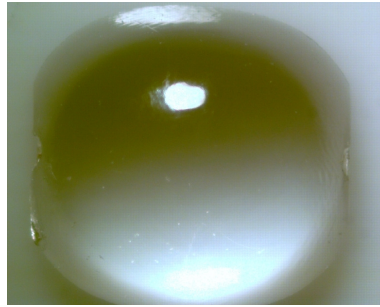The lens was characterized, both mechanically — to check if the prescribed shape had been obtained — and optically, to check its performance.

### 4.4.1  Surface profile characterization

In order to characterize the surface shape, a Digital Electronics Automation (DEA) touch-probing tool shown in Fig. 4.7 has been used. The tool can evaluate the coordinates of a set of points of the lens surface. Positioning of the sample in the tool is critical. The sample has a size comparable to the test tip and the tool must have direct access to the bottom flat surface, as this is the reference for all the other measurements. Moreover, the touching tip must have access to the lateral surfaces of the lens. An ad-hoc sample holder has then be designed. It is shown in Fig. 4.8.

The surface has been characterized measuring $900$ points. The result is shown in Fig. 4.9(a) with the best fitting surface described by a polynomial given by equation (4.1), the coefficients of which have been calculated using the measured points.

The comparison between this surface and the designed one evidenced that the fabricated lens has a curvature slightly larger than the designed one especially at the lens border. For example, Fig. 4.9(b) shows one of the principal sections of the free-form lens as designed (blue solid line) and fitted (red dashed line). The maximum difference at the guard ring border —which limits the area collecting the most of the input ray flux — is of $85\mu\text{m}$, which is about $4\%$ of the lens height ($\sim 2\text{mm}$) at this coordinate, an error compatible with the fabrication technology.

To evaluate the performance differences between the designed lens and the fitted surface one, TRACEPRO® was used to simulate their behavior. Fig. 4.10
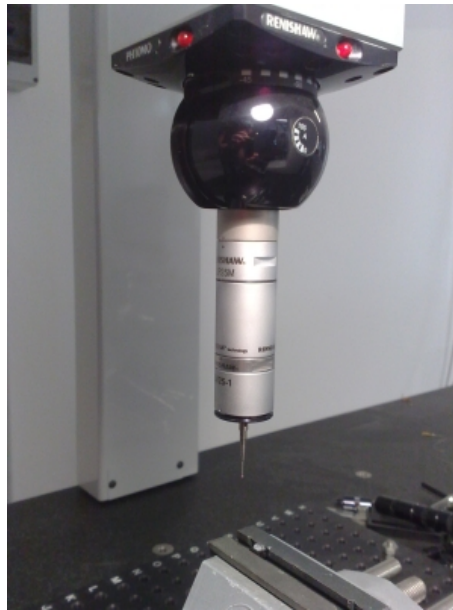
Figure 4.7: Digital Electronics Automation (DEA) machine.
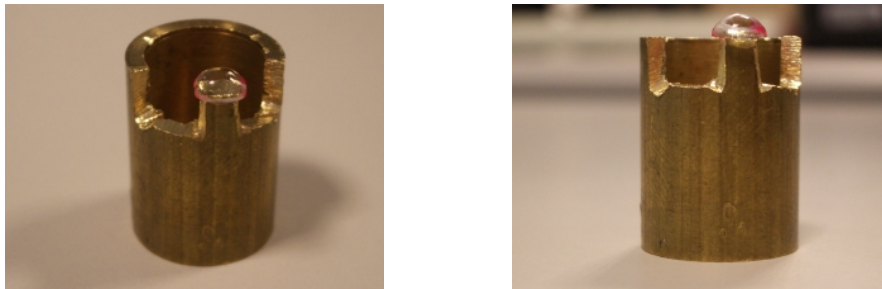


Figure 4.8: Front (left) and side (right) view of the lens glued on the holder.

reports the results of such a comparison. The horizontal (a) and vertical (b) irradiance profiles at $200\mathrm{mm}$ from the lens on lines passing through the maxima, as explained before, are shown considering the designed lens (blue line with circle markers) and the fitted surface (red line with square markers). The curves agree quite well. The peak value is more than $30\%$ with respect to the average value of the image center.

## 4.4.2 Irradiance measurement

After the mechanical characterization, the optical performance has been measured. The illumination pattern has been projected on a diffusing screen and detected by a calibrated camera. The setup is shown in Fig. 4.11 (a). The camera considered is a barcode reader characterized to allow direct reading of the irradiance profile simply from the captured image intensity. It is possible to use any

(a)                                                         (b)
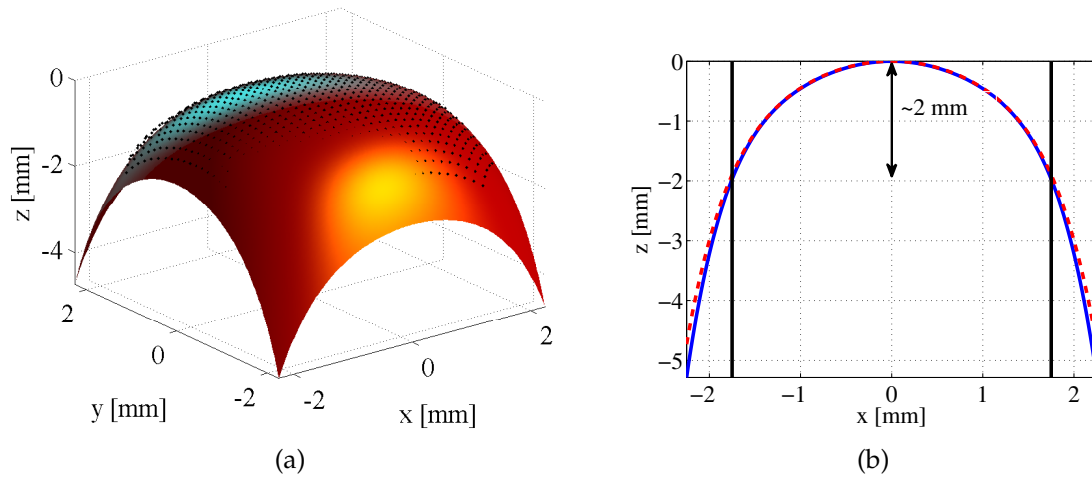
Figure 4.9: (a) Measured surface points and best fitting surface. (b) Lens shape sections: designed (blue solid line) and fitted from measured data (red dashed line). The vertical black lines indicate the interfaces between the inner useful part of the lens and the outer frame.



(a)                                                         (b)
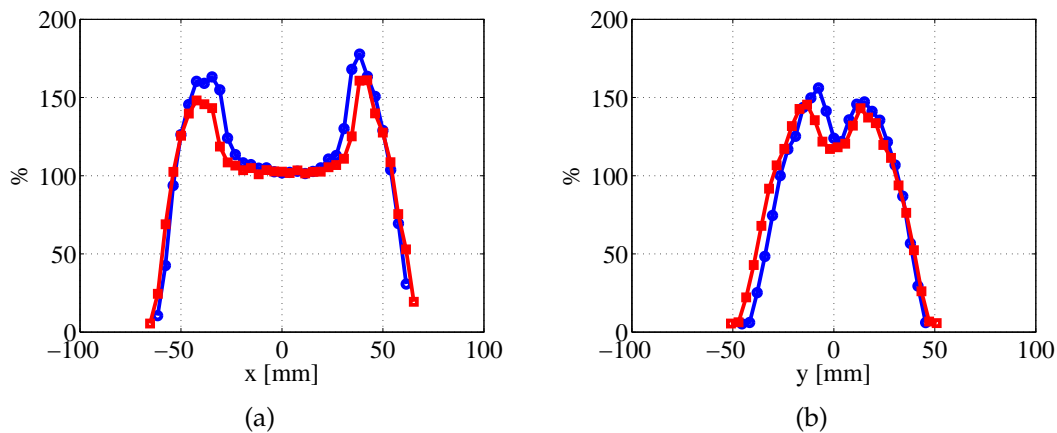
Figure 4.10: Horizontal (a) and vertical (b) irradiance profiles at $200\text{mm}$. Plots are normalized with respect to the value at the image center. The blue line with circle markers refers to the designed lens, the red one with square markers refers to the surface obtained from the best fit of the experimental surface points.

type of characterized camera. The measurement steps are the following ones:

- The diffuser is positioned at $130\text{mm}$ from the reader.

- On the other side of the screen, the LED with the freeform lens mounted (see Fig. 4.11 (b)) are positioned at $200\text{mm}$ from the screen;

- 10 images are captured with the LED switched off;

(a)             (b)

Figure 4.11: (a) Irradiance measurement set up. (b) Lens mounting on the LED .

- 10 images are then captured with the LED switched on.

The desired image comes from the difference between the average of the images captured with the LED on and the average of the images captured with the LED off, corrected by the camera characterization data. These data take into account the measured intensity, lower at the image border than in the center, due to the camera angular response.

After measuring the irradiance profile of the fabricated lens, the agreement between experimental results and theoretical predictions has to be checked. The simulated behavior on a screen at $200\mathrm{mm}$ of the designed and the fitted surface lenses used is the one shown in the previous section. Fig. 4.12 reports the overall irradiance patterns, in three cases. Fig. 4.12(a) shows the irradiance of the designed lens, Fig. 4.12(b) shows the irradiance of fitted surface lens. Finally Fig. 4.12(c) shows the measured irradiance.

Fig. 4.13 reports horizontal (a) and vertical (b) sections of the previous patterns, always taken on lines through intensity maxima. The agreement between theoretical predictions and measurements appears very good especially in the regions of maxima.

## 4.5 Conclusion

A free-form lens design tool in MATLAB® has been proposed and successfully demonstrated optimizing a plano convex lens with curved surface described by a $6^{th}$ order polynomial function to provide a custom illumination pattern. The designed lens has been fabricated and tested. Measured results agree with theoretical predictions and confirm the quality of the proposed algorithm.

The algorithm is general, simple and effective. It allows the design of freeform lenses described by an arbitrary analytical function and is able to provide custom illumination conditions for all kind of applications.
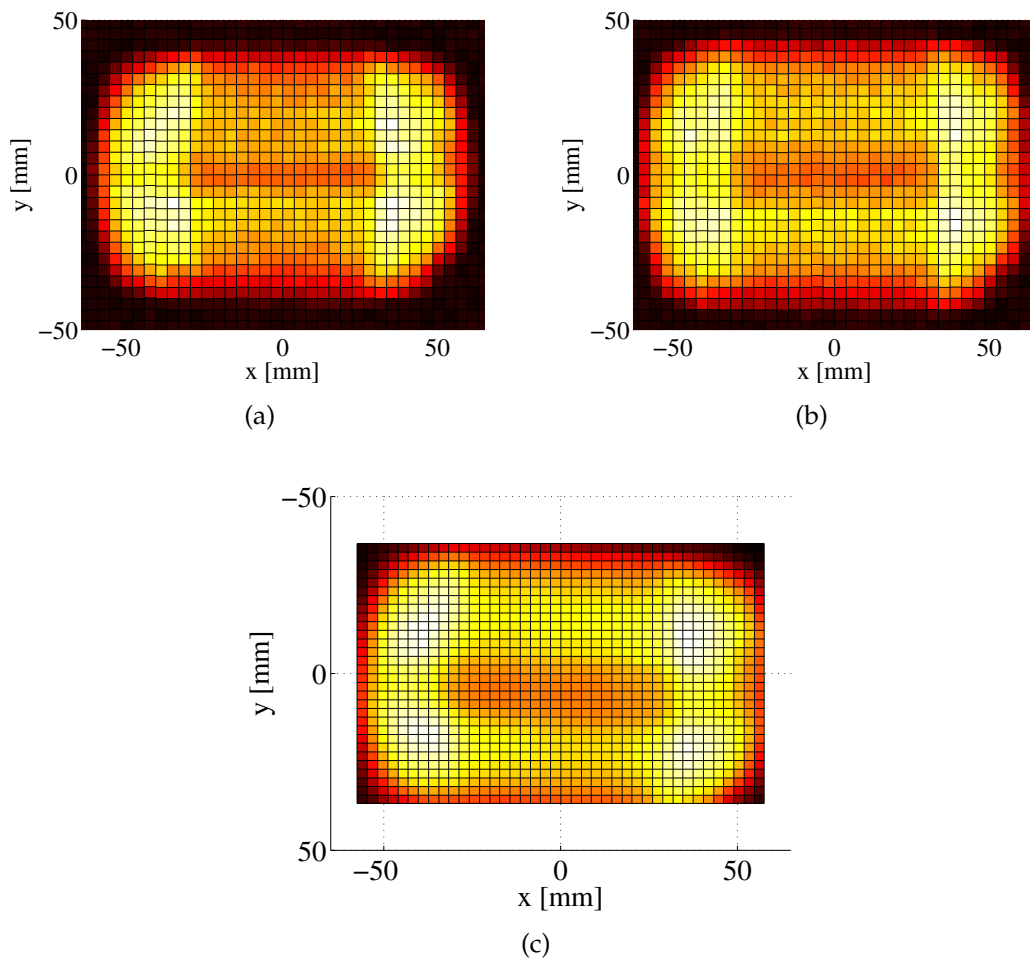
Figure 4.12: Irradiance of the designed lens (a), irradiance of the fitted surface lens (b) and measured irradiance (c).
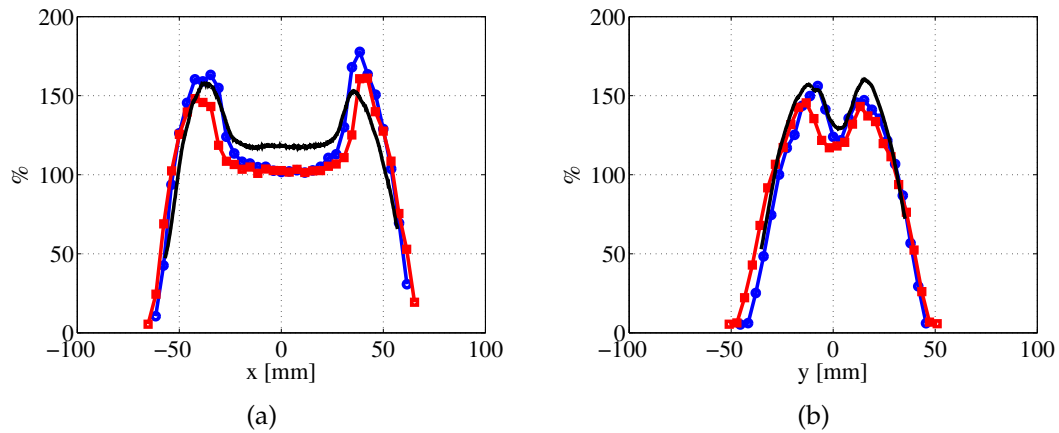
Figure 4.13: Horizontal (a) and vertical (b) irradiance profiles at $200\text{mm}$ for designed lens (blue line with circle markers), for fitted surface lens (red line with square markers) and with measured data (black solid line).

Another advantage of the tool is that it is realized in MATLAB® and allows not to use expensive commercial optical design tools, which is important when free-form lens design is not routinely done and cost issues become relevant.

# 5

# Pattern Generator Design

The non-imaging system with a coherent source (laser) is the pattern generator system.

Such a system can be employed as a pattern generator for various types of applications, such as 3D reconstruction, laser range finders, movement detection and robot vision. Regarding Auto ID applications, it can be used as a viewfinder for helping the user point a device towards the item to be identified by projecting a rectangle-like illumination pattern. Usually for both types of applications, the required pattern angle is large, such as $40°$, because it is necessary to hit the whole item to be identified or object to be 3D reconstructed etc.

For this reason, a tool to design Diffractive Optical Elements (DOEs) as a pattern generator with large diffraction angles (large FoV) has been implemented.

The results, described in this chapter, were obtained at the Friedrich-Alexander University of Erlangen, in Germany, in the group of prof. Norbert Lindlein. The aim of this collaboration was twofold. It was not only to set up a design procedure, but also to study the phase only DOE fabrication process and understand how it can influence the design phase.

After a brief introduction to the design problem (Sec. 5.1), phase only DOE characteristics (Sec. 5.2) and its fabrication process (Sec. 5.3) will be presented. The procedure set up for pattern generator design will be described in Sec. 5.4. Finally, results will be discussed in Sec. 5.5.

## 5.1  Introduction to the design problem

It is often assumed that the DOEs are designed for beam shaping [61]. DOEs which are designed with the help of a computer and written by lithographic methods are called Computer Generated Hologram (CGH) [62].

When a DOE is illuminated by a coherent source, it projects a diffraction pattern that can be described as the Fresnel or the Fraunhofer integral of the amplitude and phase distribution of the DOE depending on the distance $z$ of the

illuminated plane from the DOE itself. In the particular case of very large distance (far field condition), the diffraction pattern can be approximated well by the Fourier transform of the imaginary exponential of the phase function ($e^{j\Psi(x,y)}$) implemented by the DOE. So, knowing the desired pattern which should be projected at the desired distance from the DOE, the features of the DOE can be obtained calculating, in the paraxial approximation, the Fourier transform of the image itself.

Among the possible design algorithms, the so called Iterative Fourier Transform Algorithm (IFTA) has been widely used for CGH design [5–7], and turns out to be a very effective design tool. However, such an algorithm, being based on iterative application of the classical Fourier Transform, implies fulfillment of the so called paraxial approximation, if no additional relations between the coordinates in the hologram plane and the coordinates in the far field are taken into account. Unfortunately, this is not true for the large FoV that the device to be designed in this work should have. Ways to circumvent this problem must therefore be found.

A significant amount of study has been given to methods for non-paraxial propagation [63], [64], [65], [66]. All of them are too computationally complex to be used in an iterative method. On the other hand, IFTA is attractive as it is conceptually and computationally simple. Therefore, since the purpose of this work consists in designing a wide angle CGH in a simple way, IFTA has been adapted as to compensate for DOE producing non paraxial holograms. This is the first important topic of the activity.

A second problem comes from the DOE discretization, a critical step in the IFTA. In the results there could be undesirable effects due to the choice of the phase quantization rule. This problem is widely studied [67], [68]. The solution proposed by Wyrowski [69] that consists in a stepwise quantization, also named soft quantization, has been chosen for the design tool. It allows to avoid the stagnation problem that occurs frequently in the IFTA.

A third problem comes from the fact that the DOE has to be realized by laser lithography and the requirement of large FoV may suffer from the limits imposed by the available technological apparatus. Before analyzing these issues and describing the solutions adopted to overcome these problems, the DOE general characteristics and the fabrication process will be presented.

## 5.2   Phase only DOE characteristics

A phase only DOE is a diffractive element designed using digital computations. The computed phase pattern can be realized either using multilevel micro-reliefs (binary optics) or continuous micro-reliefs. Reliefs have features in the range from sub-micron size to millimeter size and relief amplitudes of a few microns. The modulation of the surface profile locally modifies the phase of the impinging

radiation, thus resulting in a phase modulation of the field. The electromagnetic radiation emitted from the different zones of the DOE, which is then a grating, interferes to form the desired wavefront. The shape of the grating reliefs determines the efficiency of the element, which is the amount of light that goes into a particular diffraction order [70–72].

DOEs emerged from holography in the 1989 thanks to Swanson and Veldkamp [73] and in 1993 thanks to Herzig and Dändliker [74].

A phase only DOE is then a particular case of phase mask. In order to find the relationship between a generic phase distribution $\Phi(x, y)$, which can for example be expressed as a polynomial with coefficients $a_{mn}$:

$$\Phi(x, y) = \frac{2\pi}{\lambda_0} \sum_{m=0}^{M} \sum_{n=0}^{N} a_{mn} \, x^m \, y^n$$

and the characteristics of the phase only DOE, one can first notice that changes of integer multiples of $2\pi$ in the phase have no effect. In other words one can wrap the phase in the interval $[0, 2\pi]$ and get:

$$\Psi(x, y) = \Phi(x, y)\Big|_{\text{mod}2\pi}.$$

The complex transmittance of the phase only DOE is then given by

$$t(x, y) = e^{j\Psi(x,y)}, \tag{5.1}$$

where j is the imaginary unit. The surface profile $h(x, y)$ for a thin phase only DOE in transmission is then related to the phase profile $\Psi(x, y)$ by

$$h(x, y) = \frac{\lambda_0}{n(\lambda_0) - 1} \frac{\Psi(x, y)}{2\pi}$$

where $n$ is the refractive index of the phase only DOE plate material and $\lambda_0$ is the free space design wavelength.
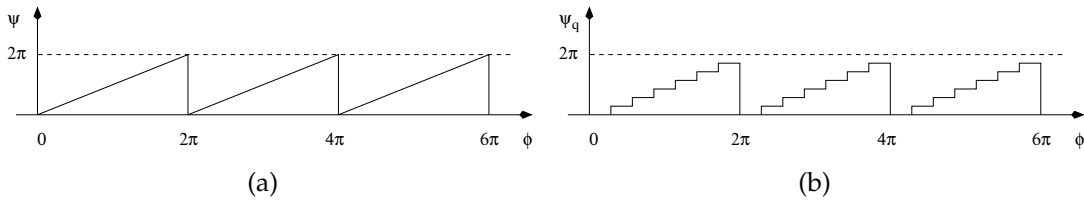


Figure 5.1: (a) Wrapped phase $\psi$ vs unwrapped phase $\phi$; (b) Quantized wrapped phase $\psi_q$ vs unwrapped phase $\phi$.

Once the phase distribution $\Phi(x, y)$ or, equivalently, the reduced $\Psi(x, y)$ is known, its values must be sampled to an integer number of discrete values if

the previously mentioned binary optical approach is to be used. Considering, for formal simplicity, only the 1D case, the continuous phase $\psi(x)$ is approximated by a staircase function with uniformly distributed values, with steps $\Delta\psi = 2\pi/N$. Fig. 5.1 shows the wrapped phase $\psi(x)$ vs the unwrapped one $\phi(x)$, (a), and the quantized wrapped phase $\psi_q(x)$ vs the unwrapped one $\phi(x)$, (b). One can then note that, although the phase profile is a function of the spatial coordinates, it is not convenient to consider this dependence explicitly, since the technology used to realize such phase distributions will realize uniform steps of the phase change and not uniform steps in the position of the sampled function.

The design and fabrication of a phase only DOE requires several steps. At the system level one should perform:

1. an analysis of the global system requirements,

2. a partitioning of the opto-electro-mechanical characteristics of the basic constituting functional elements of the system into the desired micro-electro-mechanical system;

3. a further optical definition, deciding whether refractive/reflective, waveguides/DOEs components should be used;

4. a tolerance analysis related to the phase only DOE fabrication, operation, packaging realization phases.

Considering specifically the phase only DOE element, one must then:

1. understand the physics of the problem and build the corresponding mathematical model;

2. decide if the problem can be solved analytically or numerically;

3. identify the optimization problem, with constraints and cost function;

4. write a suitable computer code and run it up to convergence;

5. transfer the code result to a real device which can be physically fabricated;

6. fabricate and characterize the DOE, to check if expected results are obtained.

The beam shaping problem considered has no analytical solution. It is possible to use only the numerical approach.

A numerical-type phase only DOE is calculated and optimized as a two dimensional matrix of regularly spaced complex data sampled over the $xy$ space. In the present case, the problem of obtaining a specific pattern on a screen is a typical Fourier phase only DOE design problem. The optimization algorithm then aims to find the values of the elements of the phase only DOE matrix that will form the desired image as accurately as possible, in terms, for example, of

diffraction efficiency, root mean square error in reconstruction, signal-to-noise ratio, uniformity of the obtained pattern, etc.

The degrees of freedom that allow a phase only DOE optimization process to evolve and converge are of two types: the first considers properties of the phase only DOE itself while the second considers properties which must hold on the reconstruction (image) plane. Since a phase only DOE is basically a matrix of sampled values of a given phase distribution, the allowed degrees of freedom are therefore the size of the matrix (i.e. the number of cells in the DOE), the size of the cells and the number of phase levels (quantization levels). These features must be chosen with care as a trade-off between accuracy of pattern reconstruction, CPU time, fabrication time and, which is most important in this work, technological fabrication constraints.

Phase only DOE design is based on iterative optimization algorithms. Such algorithms can be grouped in four different categories [71, 75–77]:

1. algorithms which perform a single pixel change in the phase only DOE at each iteration and analyze the effects of this change on the reconstruction. These algorithms are known as "unidirectional" algorithms as they use a single forward propagation at each step;

2. algorithms performing one, two or even more global transformations at each iteration, using forward and backward propagators and projecting the result on several different constraints defined by the relevant application, fabrication, technology, operation, etc. These include, for example, Gerchberg-Saxton algorithms [5] and the Ferwerda and Yang-Gu [78, 79] algorithms. They are known as IFTAs and bidirectional algorithms;

3. algorithms based on evolutionary programming, the so called "genetic algorithms";

4. algorithms that perform global optimization, related to encoding techniques such as error diffusion.

The choice of the optimization algorithm is closely related to the type of application and the performance required. For the pattern generator design tool, the IFTA has been chosen, as previously explained.

## 5.3 DOE fabrication process

The most important technological approach for the fabrication of micro-optics such as DOEs is lithography [72]. Two types of lithographic fabrication procedures can be distinguished (see Fig. 5.2): scanning lithography and mask lithography.

In scanning lithography no mask is used but local variation of the photoresist exposure is achieved in a so-called direct-writing process. To this purpose a laser or an electron-beam source is scanned over the substrate, while the beam intensity and exposure time of the beam are modulated.

In mask lithography the pattern of the component is encoded as an amplitude distribution in a lithographic mask. Uniform illumination of the mask is used to expose a photosensitive coating on the substrate.

In both cases, after the exposure of the photoresist layer, a development step converts the exposed photoresist into a surface profile. In a further processing step (etching) the surface profile of the photoresist pattern is transferred into the substrate.

In the following, only the scanning lithography will be described in detail because it is the one used in this work.
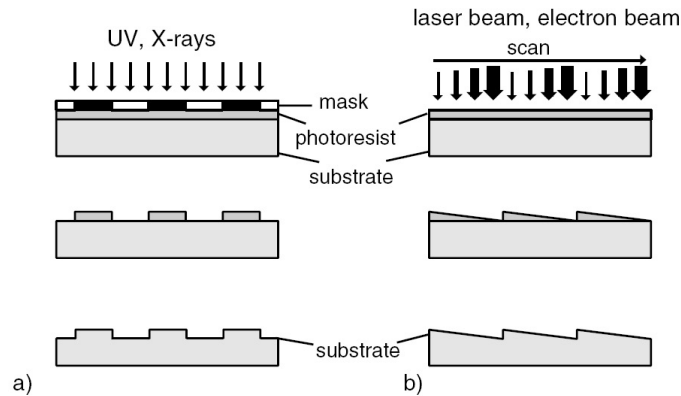


Figure 5.2: Approaches to lithographic fabrication: a) mask lithography, b) scanning lithography. From [72].

The three main phases of scanning lithography are sample preparation, pattern writing and pattern transfer.

## 5.3.1 Sample preparation

The first step of the scanning lithographic fabrication process is preparation of the uniform thin layer of photoresist material to be written by the scanning laser beam on the polished face of the substrate. The cheapest and fastest way to generate these layers is the so called "spin coating". The substrate is fixed, for example on a vacuum chuck, which is rotated at a rate of typically 3000-6000 rotations per minute ($r/min$). If some coating material (for example a diluted photosensitive polymer) is deposited onto the substrate, the centrifugal forces due to rotation spread the material to form a layer over the substrate. Its width and uniformity are controlled by the spinning speed, the viscosity of the material, the temperature of the substrate, the environment and the overall rotation time. The rest

of the material is thrown off the substrate. Post baking removes the solvent and leaves a uniform and solid coating on the substrate.

If the wafer is free of dust particles and the spinning speed is high enough, this coating method is simple and yields good quality coatings with highly uniform layer thickness. Problems may come in non-rotation-symmetric substrates or for textured substrates, where results may not be so uniform.

### 5.3.2 Pattern writing

Once the photosensitive layer has been deposited, the pattern of the micro-optical components can be directly written on it using a focused laser beam. Fig. 5.3 schematically shows the setup of a typical laser scanner. A laser with an emission wavelength is modulated (for example by an acousto-optic modulator, AOM) and focused onto the wafer coated with a layer of photoresist. The focused beam is scanned over the whole wafer area to write the whole pattern. Reasonable scanning speed can be achieved using a combination of direct deflection of the laser beam (e.g. an acousto-optic deflector AOD) and scanning of the wafer using an $xy$ stage.
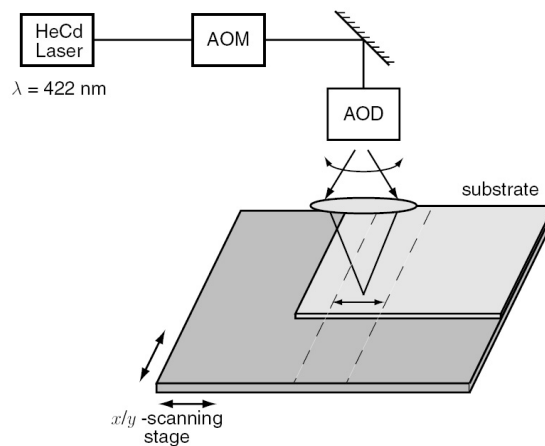


Figure 5.3: Schematic setup for laser beam writing: AOM: acousto-optic modulator; AOD: acousto-optic deflector. From [72].

The critical parameters of the process are the synchronization between the modulator and the deflector, the relative position of the $xy$ stage, and the precise focusing of the laser beam. The acousto-optic modulation can be realized using a deflector and a spatial filtering system. When deflected, the laser beam misses a pinhole (the spatial filter) and is thus blocked. The modulation is electronically synchronized with the beam deflection and the $xy$ scanning of the substrate. The position of the $xy$ stage is controlled interferometrically.

The most critical issue for the minimum feature size is the exact focusing of the laser beam. This must not be affected by the roughness of the photoresist

layer, which causes a variation of the distance between the focusing lens and the layer. A common solution to this problem uses a focusing lens floating on an air cushion of constant pressure. This way it is possible to achieve a constant minimum spot size depending on the illumination wavelength.

There are two procedures for pattern writing: the binary and the gray-tone lithographic techniques, as schematically shown in Fig. 5.4.



Figure 5.4: Lithographic techniques: (a) binary; (b) gray tone. From `http://www.optik.uni-erlangen.de/odem/index.php?lang=e&type=12`.
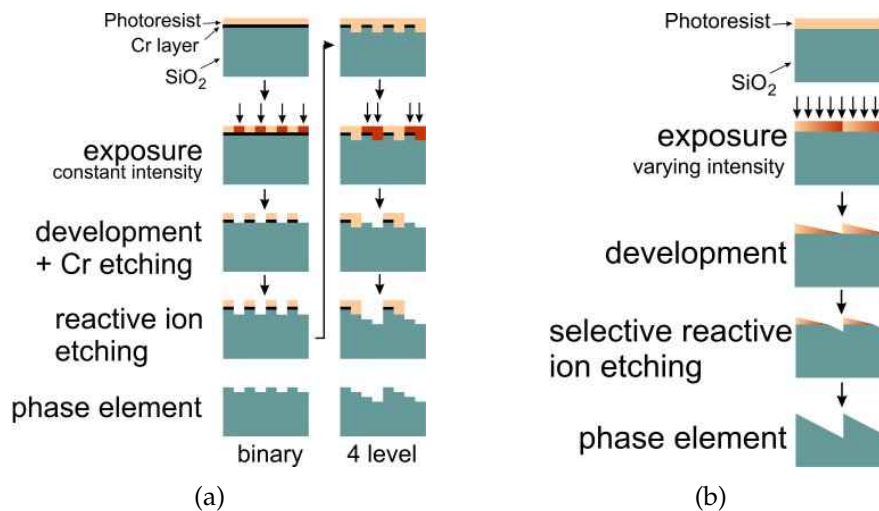
Fig. 5.4 (a) shows the steps of the binary lithographic process. The first one is the exposure of the photoresist on the top of a thin $Cr$ layer (sample prepared as explained before), using a laser beam with uniform intensity. After exposure, the substrate is developed. This way, the exposed areas of the (positive) photoresist layer are removed by the developer. The next step is the pattern transfer made with a reactive ion etching, explained in the following. This process produces a DOE with $2$ levels. More levels can be obtained repeating the whole process. For example, if a $4$ level DOE is desired, the process has to be repeated twice, aligning two different exposure masks. For good DOE performance, the alignment accuracy is very important.

Fig. 5.4 (b) shows the schematic of the gray tone lithographic process. In this case the analog response of a photo emulsion is used to generate a gray-level profile varying the writing laser beam intensity using, for example, a laser beam scanner. This process is similar to the just described binary lithography, but with three main differences.

The first difference is during the exposure phase and consists in the use of a beam with varying intensity instead of a constant one. Secondly, in the development step, there are different developer concentration for a binary or a graytone lithography. For the binary one, a low developer concentration is used to have

a high contrast that produces a binary profile. On the other hand for having a graytone profile, it is necessary to have a low contrast, and consequently a high developer concentration. The third difference consists in having a single process, independently of the number of levels in the graytone lithography, and a multi process in the binary lithography (one process each two levels) that implies the importance of accurate alignment step.

### 5.3.3   Pattern Transfer

Transferring the computer-generated pattern into the photosensitive layer on the substrate wafer is done etching the photoresist with a suitable chemical process, which progressively removes the photoresist layer. The thinner the layer, the deeper the etching of the substrate, for a given etching time.

The pattern transfer process used is the reactive ion etching which transfers the photoresist structure into the substrate material. The etching process occurs in a vacuum chamber filled with the etching gases to a working pressure of about $10^{-6}$bar where a plasma of chemically active species is generated applying an HF-field. For a binary diffractive element to be realized in quartz, $CHF_3$ (fluoride) ions are used, while oxygen is the etching material for the photoresist layer. The etching rate of the quartz and the resist is controlled by the concentration of the two etching gases. The ratio of these concentrations, the so called etching selectivity $S$ and the time of etching, determines the final height of the structure.

## 5.4   Design tool

The design tool developed is based on the IFTA, a simple and computationally fast method. As said in Sec. 5.1, this algorithm consists in bouncing back and forth between two planes related by a Fourier transform, applying different constraints at each iteration defined by type of application, fabrication, technologies, operation, etc.

The IFTA [5] allows to calculate the amplitude and phase field distribution $(f(x, y) = a(x, y) e^{j\Psi(x,y)})$ on the CGH plane which, illuminated by a plane wave or any beam shape, provides a field $F(x, y) = b(x, y) e^{j\Phi(x,y)}$ on the image plane in the far field. Within the limits of the paraxial approximation, $F$ is then the Fourier transform of $f$.

The block diagram of the algorithm is illustrated in Fig. 5.5. It starts in the image plane from $F' = be^{j\Phi}$ where $\Phi$ is a random phase and $b$ is the desired amplitude and applies a Fourier transform (direct or inverse, depending on the sign conventions) to determine the corresponding field distribution $f'(x, y)$ in the CGH plane. At this stage, it is important to apply the fabrication constraints. First of all, if a phase plate must be realized, the field modulus must be 1 in the CGH domain. In addition, for multilevel surface-relief elements, the fabrication con-
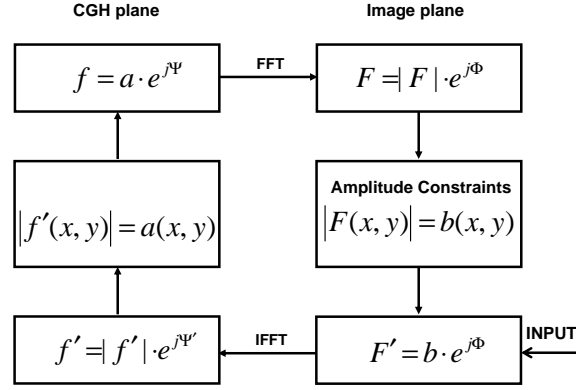
CGH plane                           Image plane

$$f = a \cdot e^{j\Psi}$$   —FFT→   $$F = |F| \cdot e^{j\Phi}$$

$$|f'(x,y)| = a(x,y)$$   **Amplitude Constraints** $$|F(x,y)| = b(x,y)$$

$$f' = |f'| \cdot e^{j\Psi'}$$   ←IFFT—   $$F' = b \cdot e^{j\Phi}$$   ←INPUT

Figure 5.5: Schematic representation of the iterative Fourier transform algorithm.

straints include the discretization of the phase function to $N$ phase levels. Such constraints allow to obtain a physically realizable phase only DOE. The new field is then (inversely or directly) Fourier transformed to obtain the corresponding field in the image plane $F(x,y)$. The next step consists in an amplitude correction, constraining the $F$ modulus equal to the desired amplitude $b(x,y)$. This process is repeated until the desired image or a maximum number of iteration is achieved.

The discretization is a critical step because in the results there could be undesirable effects due to the choice of the phase quantization rule. There are two main quantization rules: the direct one and the soft one. A direct quantization is given by:

$$\exp(-j\hat{\Psi}(u)) = \begin{cases} \exp(-j\pi) & \Psi(u) + \pi \leq 0.5\Delta \\ \vdots \\ \exp(j(-\pi + n\Delta)) & (n-0.5)\Delta < \Psi(u) \leq (n+0.5)\Delta \\ \vdots \\ \exp(j\pi) & (N-0.5)\Delta < \Psi(u) + \pi \end{cases} \quad (5.2)$$

where $\hat{\Psi}(u)$ is the quantized phase of $\Psi$, $\Delta = 2\pi/N$, $N$ is the number of quantization levels and $n = 0, \ldots, N$.

Applying this quantization rule to the IFTA, it happens frequently that $\Psi(u)$ at each iteration does not change enough to overcome the threshold levels causing a stagnation problem.

In order to avoid this stagnation, Wyrowski in [69] proposed an improved quantization rule, the soft quantization. He has introduced a stepwise quantization in the iterative procedure.

At $i^{th}$ iteration, the quantization rule is:

$$\exp(-j\hat{\Psi}_i(u)) = \begin{cases} \exp(-j\pi) & \Psi(u) + \pi \leq 0.5\Delta \cdot \epsilon^{(s)} \\ \vdots & \\ \exp(j(-\pi + n\Delta)) & (n - 0.5\epsilon^{(s)})\Delta < \psi(u) \leq (n + 0.5\epsilon^{(s)})\Delta \\ \vdots & \\ \exp(j\pi) & (N - 0.5\epsilon^{(s)})\Delta < \psi(u) + \pi \\ \exp(j\Psi_i(u)) & otherwise \end{cases}$$

(5.3)

with

$$0 < \epsilon^{(1)} < \epsilon^{(2)} < \ldots \epsilon^{(s)} \ldots < \epsilon^{(S)} = 1$$

(5.4)

This second quantization rule, called soft or stepwise, has been chosen to avoid the stagnation increasing the capability of the IFTA to change the phase distribution during the iterative procedure.

The IFTA loop described is iterated until the field on the image plane coincides, within a decided error, with the desired one or if the maximum number of iterations is achieved.

The aim of this work consists in designing a wide angle CGH. The main challenge is to introduce in the IFTA a correction for DOE producing non paraxial holograms. In the next section this correction will be described.

## 5.4.1   Paraxial correction

The desired pattern to be created by the DOE is in the far field of the DOE itself. The relationship between the two fields is assumed to be a Fourier transform [10]. Unfortunately, this algorithm stands on the paraxial approximation, i.e. on the assumption that the error introduced considering $\sin \Omega \simeq \Omega$ is negligible.

However, in this case, where the diffraction angles are large, this hypothesis is no longer true. This causes a pincushion distortion error in the image plane as shown schematically in Fig. 5.6. Considering the ray with angle $\Omega$, it lands on the image plane at point $E$; the segment $OE$ corresponds to $\tan \Omega$. Use of $\Omega$ instead of $\sin \Omega$, applying the definition of $\Omega$ in radiants, is possible only if $\Omega$ is small. This means that $OE$ coincides with the ray $OP$ landing on the point $P$ on the sphere with radius $d$, being $z = d$ the distance between the object and the image plane. However, if $\Omega$ is large and since the algorithm makes reference to the image plane, point $P'$, the projection of $P$ on the image plane, is used instead of $E$. It is visible that $P'$ is low shifted with respect to $E$. This effect is the mentioned pincushion distortion of the image.

It has been investigated how the paraxial approximation induces an error in the design procedure and how one can compensate such an error keeping the IFTA basically unchanged.
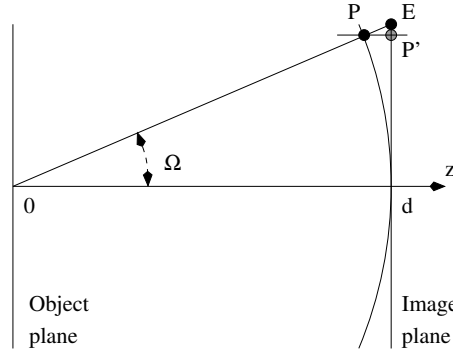
Figure 5.6: Interpretation of the paraxial approximation: instead of the exact point $E$ on the image plane ($\tan \Omega$ in the expression of the spatial frequencies), the point $P$, on a sphere with radius $z = d$, ($\Omega$ in the expression of the spatial frequencies) is considered. Point $P$ has projection $P'$ on the image plane.

A plane wave with complex amplitude propagating in free space is given by:

$$U(x, y, z) = A\, e^{-j\mathbf{k} \cdot \mathbf{r}} = A\, e^{-j(k_x x + k_y y + k_z z)}$$

The propagation vector $\mathbf{k}$ (with modulus $k = 2\pi/\lambda$) makes angles $\theta_x = \sin^{-1}(k_x/k)$ and $\theta_y = \sin^{-1}(k_y/k)$ with the $yz$ and $xz$ planes, respectively (see Fig. 5.7).

The complex amplitude in the $z = 0$ plane can then be written

$$U(x, y, 0) = f(x, y) = A\, e^{-j(k_x x + k_y y)} = A\, e^{-j2\pi(\nu_x x + \nu_y y)}$$

having introduced the spatial frequencies $\nu_x = k_x/2\pi$ and $\nu_y = k_y/2\pi$. The angles of the wavevector are therefore related to the spatial frequencies of the harmonic function by

$$\theta_x = \sin^{-1}(\lambda \nu_x), \quad \theta_y = \sin^{-1}(\lambda \nu_y). \tag{5.5}$$

At a distance $z = d$ from the plane $z = 0$, it holds

$$U(x, y, d) = f(x, y)\, e^{-jk_z d}$$

being

$$
\begin{aligned}
k_z &= \sqrt{k^2 - k_x^2 - k_y^2} = k\sqrt{1 - \frac{k_x^2 + k_y^2}{k^2}} \\
&= k\sqrt{1 - \sin^2\theta_x - \sin^2\theta_y} = k\sqrt{1 - \lambda^2\nu_x^2 - \lambda^2\nu_y^2}.
\end{aligned}
\tag{5.6}
$$

The square root is real if $k^2 > k_x^2 + k_y^2$. If not, the argument becomes negative and the square root imaginary, corresponding to evanescent waves not observed in the far field. If $k_x \ll k$ and $k_y \ll k$, so that the wavevector $\mathbf{k}$ is paraxial, angles $\theta_x$ and $\theta_y$ are small ($\sin\theta_x \approx \theta_x$ and $\sin\theta_y \approx \theta_y$) and

$$\theta_x = \lambda \nu_x, \quad \theta_y = \lambda \nu_y. \tag{5.7}$$

The angles of inclination of the wavevector become now directly proportional to the spatial frequencies of the corresponding harmonic function. Equation (5.6) can also be simplified and one can write:

$$k_z = k\sqrt{1 - \sin^2\theta_x - \sin^2\theta_y} \simeq k\sqrt{1 - \theta_x^2 - \theta_y^2} \qquad (5.8)$$

and $\theta_x \simeq x/d$, $\theta_y \simeq y/d$. The wavevector has now components $k_x \simeq (x/d)k$ and $k_y \simeq (y/d)k$ and spatial frequencies are $\nu_x \simeq x/(\lambda d)$ and $\nu_y \simeq y/(\lambda d)$.

Without paraxial and far field approximation, the spatial frequencies can be written as:

$$\nu_x = \frac{\sin\theta_x}{\lambda}, \qquad \nu_y = \frac{\sin\theta_y}{\lambda}. \qquad (5.9)$$
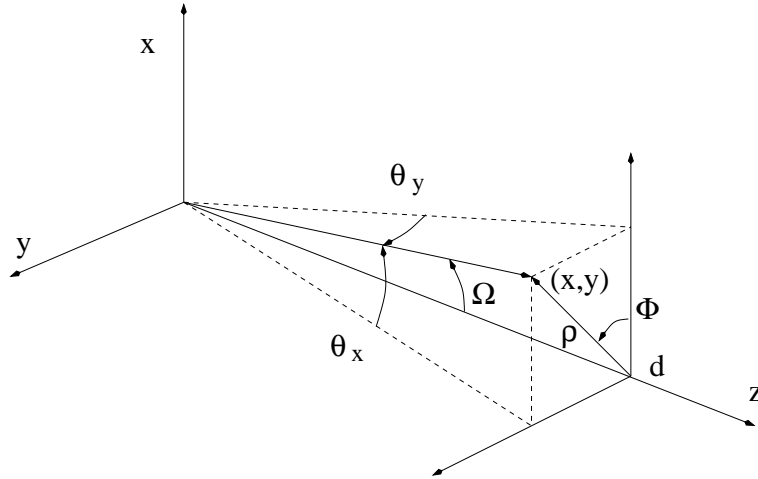


Figure 5.7: Representation of the parameters: $\rho$, $\phi$ (polar coordinates on the xy plane) and $\theta_x$, $\theta_y$, $\Omega$ (angular coordinates from $z = 0$ to $z = d$ planes).

With reference to Fig. 5.7 one can write:

$$\sin\theta_x = \frac{x}{\sqrt{x^2 + y^2 + d^2}}.$$

Substituting $x$ with $\rho\cos\phi$ and $x^2 + y^2$ with $rho$, one gets:

$$\sin\theta_x = \frac{\rho\cos\phi}{\sqrt{\rho^2 + d^2}},$$

and finally, substituting $\rho$ with $d\tan\Omega$:

$$\sin\theta_x = \frac{d\tan\Omega\cos\phi}{\sqrt{(d\tan\Omega)^2 + d^2}}.$$

In a similar way, one can write the same expression for $\sin \theta_y$ using $y = \rho \sin \phi$ as:

$$\sin \theta_y = \frac{d \tan \Omega \sin \phi}{\sqrt{(d \tan \Omega)^2 + d^2}}.$$

In conclusion, the expressions of the spatial frequencies without the paraxial approximation are:

$$\nu_x = \frac{1}{\lambda} \frac{d \tan \Omega \cos \phi}{\sqrt{(d \tan \Omega)^2 + d^2}}, \qquad \nu_y = \frac{1}{\lambda} \frac{d \tan \Omega \sin \phi}{\sqrt{(d \tan \Omega)^2 + d^2}}. \qquad (5.10)$$

If the paraxial approximation is considered, $\tan \Omega \simeq \Omega$ and these equations reduce to:

$$\nu_x = \frac{1}{\lambda} \frac{d \Omega \cos \phi}{\sqrt{(d \Omega)^2 + d^2}}, \qquad \nu_y = \frac{1}{\lambda} \frac{d \Omega \sin \phi}{\sqrt{(d \Omega)^2 + d^2}}. \qquad (5.11)$$

The paraxial approximation produces an error, which can be anyway compensated, if a suitable predistortion is introduced on the desired amplitude in the image plane. In practice, $\Omega$ can be written as $\arctan(\tan \Omega)$ but with $\tan \Omega \simeq \Omega$ approximation, it reduces to $\arctan \Omega$. To compensate the error introduced by the use of paraxial approximation in the case of large angle, $\arctan \Omega$ instead of $\Omega$ has to be considered in equations (5.11). To have a correct image pattern, spatial frequency axes should then simply be locally stretched to obey the following equations:

$$\nu_x = \frac{1}{\lambda} \frac{d \arctan \Omega \cos \phi}{\sqrt{(d \arctan \Omega)^2 + d^2}}, \qquad \nu_y = \frac{1}{\lambda} \frac{d \arctan \Omega \sin \phi}{\sqrt{(d \arctan \Omega)^2 + d^2}}. \qquad (5.12)$$

To do so, for each point of the image plane, sampled on a fine enough grid, the spatial frequency must be calculated first. This would correspond to a point in the grid of the DOE. Such a point is anyway moved to a different one to introduce the desired predistortion. So, though the Fourier transform introduces an error, such an error is compensated by this predistortion of the signal.

In Fig. 5.8 an example of the target compensation is shown. A square with a FoV of $40°$ has been chosen as target. The blue square is the pattern desired, the green one is the pincushion-distorted image because of paraxial approximation and the red one is the predistorted square. Using the red square as target in the IFTA, a pattern projected in the reality as the desired one (blue square in the figure) has been obtained.

In the next section, results of the DOEs designed by IFTA with this compensation, fabricated and measured will be illustrated.

## 5.5   Results

In this section, the results obtained using the design tool and fabricating the designed DOEs in the laboratories of Friedrich Alexander University will be presented.
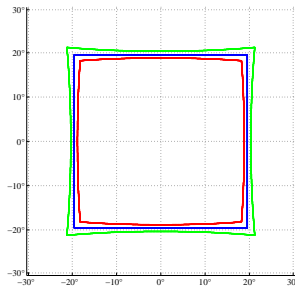
Figure 5.8: The blue square is the desired pattern, the green one is the pincushion-distorted image because of paraxial approximation and the red one is the predistorted square.

First of all, for testing the design tool, a graytone lithographic technique, illustrated in Fig. 5.4 (b), without etching of the substrate has been chosen. In this way the samples are rewritable, just removing the photoresist layer.

After having analyzed all the project problems, the final sample using the binary lithographic technique, illustrated in Fig. 5.4 (a), including the ion etching phase has been fabricated.

In the following, both these samples will be described, focusing on their design, fabrication and characterization.

## 5.5.1 Graytone test sample

Several DOEs have been fabricated using the graytone lithography without etching. This technique has been chosen for two reasons. First of all, the graytone lithography allows to fabricate multilevel DOEs in a single writing process. Secondly, it allows to re-employ the same substrate for further experiments just by removing the exposed photoresist layer. This process is quite simple and quick to be used for just testing the implemented algorithm.

**Design**

In order to test the paraxial compensation, a DOE projecting a simple image, similar to ones used as viewfinder, has been chosen. It is shown as the blue curve of Fig. 5.9 It is a double set of squares and circles, the former with a FoV of $20°$ ($\Omega = 10°$, value at which the paraxial hypothesis is fulfilled) and the latter with a FoV of $40°$ ($\Omega = 20°$, value at which the paraxial hypothesis is no longer valid). One expects that the smaller square/circle does not have differences before and after the paraxial compensation for the smaller angle, while such a compensation should be important at $40°$. In particular, a barrel distortion should be appreciated, just the opposite distortion caused by paraxial approximation. The paraxial
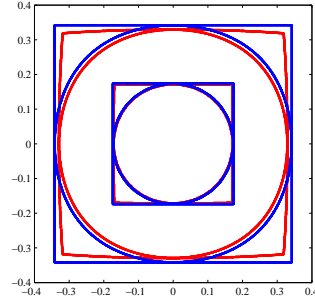
Figure 5.9: Image used for the paraxial compensation test. The blue curves are the desired image while the red curves are the compensated image. The smaller square/circle correspond to a FoV of $20°$, the bigger ones to a FoV of $40°$.

compensated target is shown in Fig. 5.9 by the red curve. The compensation is created by shifting each point of the desired image by a quantity calculated by Eq. (5.12).

The following samples have been obtained using a red HeNe laser operating at 633nm. The spatial resolution on the image plane is $\Delta x_i/d = \Delta y_i/d = 0.0006$. This sampling is enough to appreciate differences between $\sin\theta$ and $\theta$ for $\theta > 9°$. The DOE, on the object plane is a matrix of $N \times N = 1281 \times 1281$ samples. Their spacing can be calculated using the following formula:

$$\Delta x_o = \Delta y_o = \frac{\lambda}{N}\frac{d}{\Delta x_i} = \frac{\lambda}{N}\frac{d}{\Delta y_i} = \frac{633 \cdot 10^{-9}}{1280}\frac{1}{0.0006}\text{m} \simeq 0.82\mu\text{m}.$$

Such a spacing is larger than $0.5\mu$m, the minimum pixel size of the lithographic machine, and not too close to it, to avoid critical working conditions. The pixel size that the machine can use is $0.8\mu$m. The use of this pixel size instead of the exact one $0.82\mu$m causes some errors in the FoV, as explained in the following.

This target image (red curves) has been used as the desired amplitude $F_{des}$ in the image plane. The IFTA algorithm illustrated in Sec. 5.4 has been used. The introduced constraints are:

1. in the image plane, the field modulus is set equal to $F_{des}$;

2. in the DOE plane, the modulus is set to $1$, i.e. a plane wave is assumed to illuminate it.

3. 4 levels achieved using soft quantization.

After about $100$ iterations the IFTA provided the DOE phase profile. Fig. 5.10 shows the simulated field intensity in the image plane. It is the same as the red curve in Fig. 5.9.
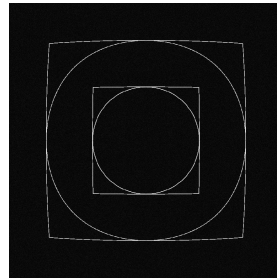
Figure 5.10: Image resulting after IFTA iterations.

**Fabrication**

The laser lithography machine used for the fabrication phase is called DWL 425. It is similar to the one described in Sec. 5.3. Its setup is shown in Fig. 5.11. It is characterized by: a camera system for overlay alignment and metrology, an interferometric controlled air bearing $xy$ stage, an AOM with $128$ intensity levels during exposure, the possibility of fabricating binary and continuous structures and replaceable writing heads. The laser lithography machine technical data are: minimum feature size of $0.5\mu$m, edge roughness $< 80.0$nm, overlay accuracy $< 250.0$nm, auto focus range $90\mu$m, interferometric position control $< 10$nm.
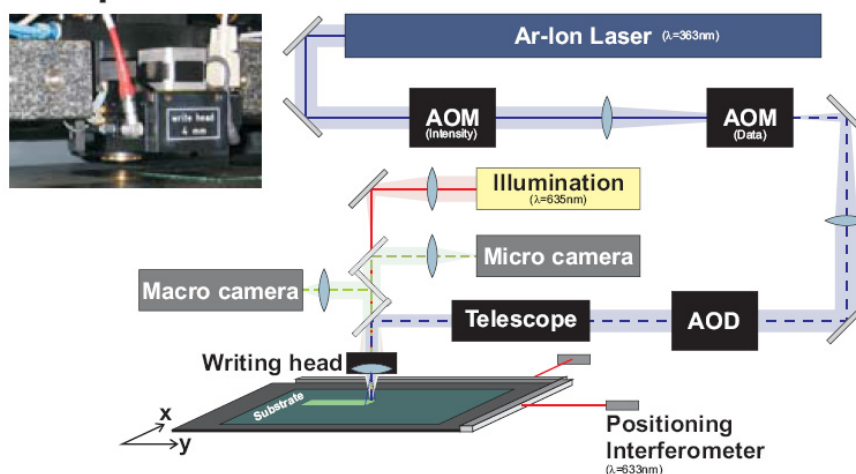


Figure 5.11: Lithographic system setup schematic. From http://www.optik.uni-erlangen.de/odem/.

The data come from the IFTA are not in the right format for the lithography machine. Thus, the IFTA data have been converted in the machine format with the following parameters: pixel size of $0.8\mu$m, instead of $0.82\mu$m, as in the design and $10 \times 10$ matrix of optimized DOE repetition for extending the zone to be

illuminated by the user. In this way the user has not to aim the small single DOE (about $1\text{mm} \times 1\text{mm}$). The sample test to be fabricated has one $10 \times 10$ matrix of the DOE designed with paraxial compensation (CPA), and another $10 \times 10$ matrix without the compensation (PA).

In the following, the details of the fabrication process are described. The fabrication phases are three:

1. the sample preparation, made in the chemical room consists in:

   - cleaning the sample in subsequent steps using remover, deionized water, $N_2$, aceton, isopropanol and $N_2$ (see Fig. 5.12 (a));

   - drying the sample in the hot plate (see Fig. 5.12 (b));

   - spin coating of photoresist (as described in 5.3) (see Fig. 5.12 (c)):

     - putting on the sample the HMDS (adhesion promoter);
     - spinning ($4000\text{r}/\text{min}$);
     - put on the sample the resist (AZ1518);
     - spinning ($4000\text{r}/\text{min}$);

   - putting the sample on the hot plate (soft bake) for $4\text{min}$ at $110°\text{C}$ (see Fig. 5.12 (d)).

2. the exposure phase, made in the cleaning room, in details consists in:

   - putting the sample in the laser lithography machine activating the air cushion of constant pressure and covering the not used hole (this is important to fix firmly the sample stable on the plate);

   - letting the machine measure the sample center;

   - setting parameters in the machine software (number of samples, positions, laser intensity, focus)

   - starting exposure (it takes $2 - 3\text{h}$).

3. the development phase, made in the chemical room, consists in:

   - mixing the developer AZ351B with deionized water (ratio 1:2);

   - putting the sample in the developer for about $20\text{s}$;

   - cleaning it quickly with deionized water;

   - drying it with $N_2$.

After development, the etching process should be made, but, as explained before, this sample has been chosen just as test.
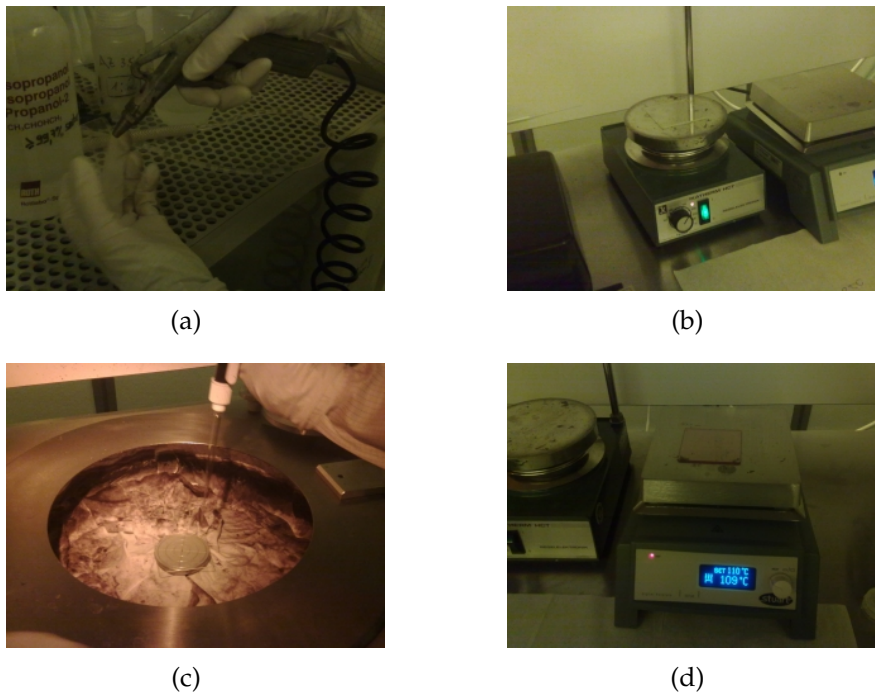
(a)



(b)



(c)



(d)

Figure 5.12: Pictures of some steps during the sample preparation process: (a) dry the sample with $N_2$; (b) dry the sample in the hot plate; (c) spin coating of photoresist; (d) put the sample on the hot plate (soft bake) for $4\text{min}$ at $110°\text{C}$.

**Characterization**

To characterize the test sample, its profile has been observed with the microscope and the shape and the angle associated to the projected pattern have been measured.

Fig. 5.13 shows a portion ($64\mu\text{m} \times 64\mu\text{m}$) of the sample CPA, taken with the microscope.

Fig. 5.14 shows the setup to measure the pattern generated by the DOE, consisting in a laser, the DOE, a screen and the camera behind it.

Fig. 5.15 shows the pattern projected by the DOE without paraxial compensation (a) and the pattern obtained using DOE with paraxial compensation (b). The hologram without the paraxial compensation have a pin-cushion distortion, as expected. This distortion is clearly visible in the bigger square.

Fig. 5.16 shows subfigure (a) of Fig. 5.15 without cropping the ghost images. The DOE projects not only the desired hologram but also other replicas. It is not possible to eliminate these replicas as they are intrinsically connected with the image sampling existing in the IFTA.

The holograms sizes, which, knowing the distance between DOE and screen, provides the FOV angles have been measured using the setup shown in Fig. 5.14, at different distances. A polynomial regression allowed to find the desired angles.
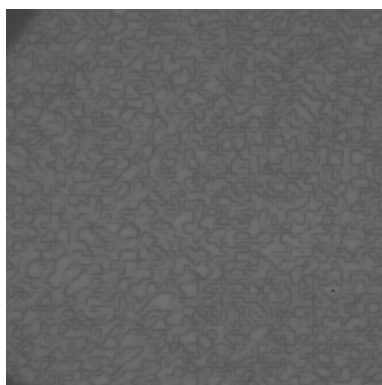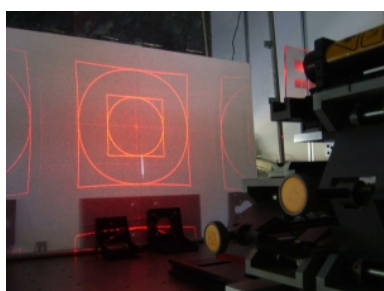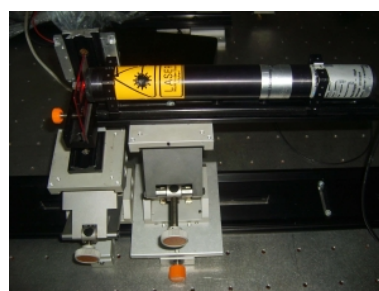
Figure 5.13: Microscope photo of the sample.



| (a) | (b) |

Figure 5.14: The setup for measuring the hologram shape consists of a screen (a), a laser (b), the sample and a camera.

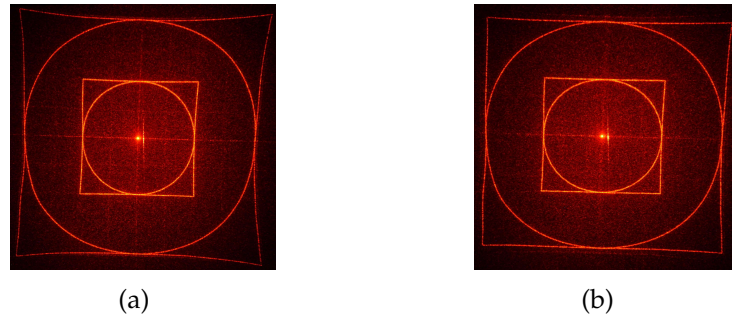(a)                                                    (b)

Figure 5.15: Photos of the holograms projected by the DOE without paraxial compensation (a) and by the DOE with paraxial compensation (b).
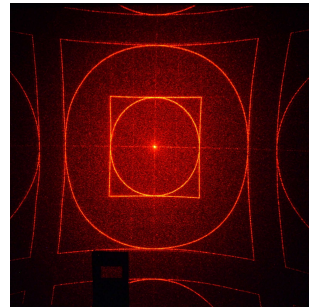


Figure 5.16: Projected pattern with "ghost" images.

Fig. 5.17 shows the angle measurement results. The lines represent the polynomial regression of the measured point: the red ones represent the holograms size (smaller and bigger square) in the case of paraxial compensation; the blue ones represent the holograms size (smaller and bigger square) in the case of paraxial approximation. The red and blue points represent the x-size measured on the image plane at different distances between the screen and the DOE-laser system, with and without paraxial compensation, respectively. The desired angles are $40°$ for the bigger square and $20°$ for the smaller one. The real angles are $41.4°$ and $21.2°$ for the CPA case, $45.2°$ and $21.4°$ for the PA case. In the paraxial compensation case there is a difference of about $1°$-$1.5°$, caused by the difference between the designed DOE pixel size ($0.82\mu m$) and the real one ($0.8\mu m$). Without the paraxial compensation case, the difference is much larger, about $5°$, because of the image distortion.

In conclusion, this sample test confirms that paraxial compensation has been obtained with a FOV of about $40°$.

## 5.5.2   Binary sample

After having tested the paraxial correction, the final sample has been fabricated using binary lithography with etching. It is important to understand how the

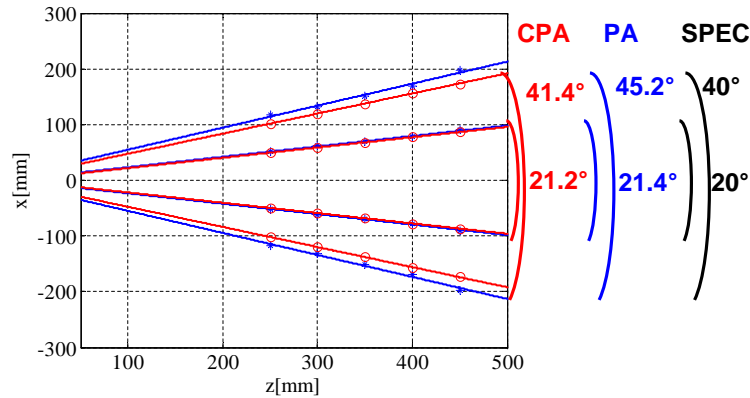Figure 5.17: Angle measurement: the red curves represent holograms size (smaller and bigger square) in the case of paraxial compensation and their related angles; the blue curves represent holograms size (smaller and bigger square) in the case without paraxial compensation and their related angles. In black the spec angles are shown.

fabrication process can influence the results.

In collaboration with the group of Prof. Häusler of the Friedrich-Alexander University, a DOE that projects several lines with large angle used in that group for 3D reconstruction called Flying Triangulation [80], [81] has been designed and fabricated. In this approach a light pattern is projected onto the surface (see Fig. 5.18) and the changes in the pattern due to the surface observed by the camera gives the depth information about the 3D surface shape. With Flying Triangulation it is possible to measure the 3D topography of an object surface on the fly. For measuring large objects, such as sculptures or rooms it is important to have a light pattern with a large diffraction angle.

**Design**

Using IFTA with paraxial correction, two DOEs for large pattern angles used for Flying Triangulation, one with distortion correction and the other without have been designed. The pattern consists of 13 vertical lines with a FoV of about $40° \times 40°$, useful for the application shown in Fig. 5.18. For designing the DOEs, four quantized levels, $1100 \times 1100$ samples, DOE pixel size of $0.8\mu m$ corresponding to $\Delta x_i/d = \Delta y_i/d = 0.00075$, with a wavelength of $660nm$ have been used. This sampling is enough to appreciate differences between $\sin \theta$ and $\theta$ for $\theta > 9.5°$. In this way, the distortion is corrected for large angles. At first the desired target has

(a)
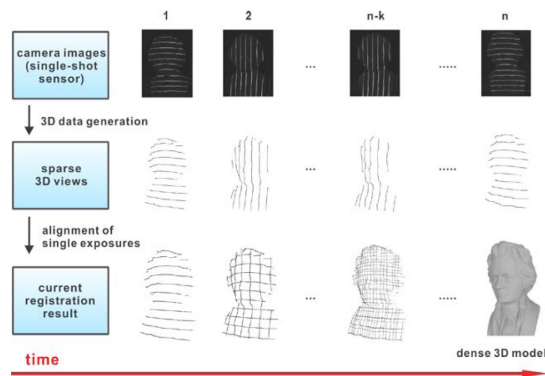


(b)

Figure 5.18: Flying Triangulation measurement (a) and scheme (b). Adapted from [80].

been pre-distorted as shown in Fig. 5.21 (a) obtaining the new target (b). Secondly both targets have been used as amplitude constraints in two different IFTA loops. After 100 iterations the two patterns with (d) and without correction (c) have been obtained. Using IFTA, an amplitude in the image plane with the same shape of the desired one but noisy has been obtained.

**Fabrication**

Two DOEs with optimized phase profiles have been fabricated using the binary technique with etching described in Sec. 5.3. The DOEs have been designed using 4 levels. In this case, the fabrication process is double and each time the alignment step is the most critical part.

The process can be divided into three main phases, the sample preparation, the first etching (for 2 levels) and the second etching (for the other 2 levels).

The standard sample is a fused silica (pure silicon dioxide $SiO_2$) with a chromium coating (thickness of $60$nm) and a thin layer of chromium oxide for antireflection

(dark chromium). The sample has been prepared for having the black and white structure shown in Fig. 5.19. The four squares (only two of them have been used) are for the designed DOEs (6mm × 6mm), the crosses (100µm far from the square corner) and the rows used for identifying the center and the alignment of the sample. They are necessary for having the same alignment condition in the first and in the second etching phase.

First of all, the chromium has been removed in the white areas shown in figure 5.19, in this way:

1. spin coating of photoresist:

    - put on the sample the HMDS (adhesion promoter);
    - rotate plate for 1min (4000r/min);
    - put on the sample the resist (AZ1505);
    - rotate plate for 1min (4000r/min);

2. put the sample on the hot plate for 1min at $110°C$;

3. exposure for having the structure of Fig. 5.19;

4. development:

    - mix the developer AZ351B with deionized water (ratio 1:4);
    - put the sample in the developer for 1min;
    - clean it quickly with deionized water;
    - dry it with $N_2$;

5. chromium etching:

    - put the sample in the chromium etch #1 from Microchemicals for $40s$ ca. (this is a wet chemical etching process);
    - clean it with deionized water;
    - dry it with $N_2$.

It is necessary to make the spin coating of the sample, as explained for the graytone sample preparation. This time the MIR 701 resist that needs a baking of 1 min at $90°$ has been used. It is a thicker resist than the one used for the graytone lithography.

For the pattern writing phase, the sample has been put in the laser lithography machine in a specific position using some drawing pin, activating the air cushion of constant pressure and covering the not used hole. It is important to align the sample with the machine coordinate. The crosses shown before are useful for calculating the tilt angle between the machine coordinate system and the sample coordinate system. Using the camera inside the machine, the three crosses

indicated with a green circle in Fig. 5.19 have been found and the crosses centers have been selected with the green lines as shown on the left of Fig. 5.19. The software calculates the rotating angle and then adjusts the machine coordinate system rotating it. This step has to be repeated until a rotating angle less than $1°$ is reached. After this alignment, the zero of the coordinate system is set in the red point shown in Fig. 5.19. After the exposure has started, the software finds each square to be exposed searching the center of the three crosses signed in Fig. 5.19 with a light blue square. After exposure, it is necessary to put the sample in the developer (concentration 1:4 respect to deionized water) for 1min, then clean and dry it.
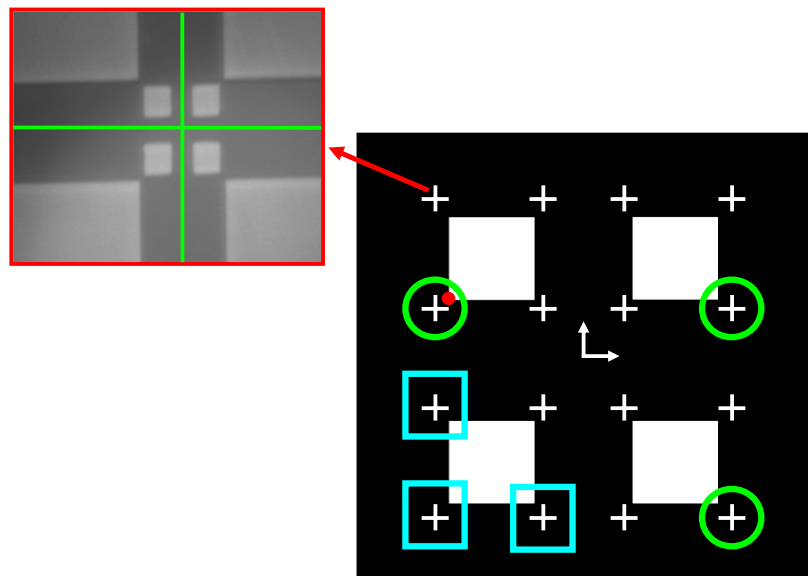


Figure 5.19: Schematically alignment process.

The sample is ready for the first dry etching process, previously described in 5.3. There are several etching machine parameters. In the process used all of them, gas mixture, pressure, power, temperature have been fixed except for the time. This is the parameter used for controlling the level height. If the height achieved is not equal to the one desired, the zero and + 1 order are visible in the pattern projected by the DOE. For this reason the choice of this parameter is fundamental. For the first etching, it is necessary to create a DOE with two levels that induces a phase of $0$ and $\pi$. The high level that corresponds to a phase of $\pi$ is expressed by $\lambda/2(n-1)$, where $\lambda = 660$nm is the laser wavelength and $n = 1.4562$ is the quartz refractive index. With this parameters the height to be achieved is 723nm.

Several etching tests have been done for establishing the right time that can

produce this height. For each test, the height has been measured using the machine shown in Fig. 5.20. At the end, the right time used during the first etching is 21min. The last step is to put the sample in the ultrasonic bath for half an hour for removing the resist.



Figure 5.20: Machine for measuring the high level of the etching test samples.

For having a four level DOE it is necessary to make again the same phases explained for the first two levels: sample preparation, alignment, exposure, development, etching, resist removing. In this case half etching time of the previous etching process has been used to achieve half level height.

Finally the desired DOEs have been fabricated with four levels.

**Characterization**

To measure the pattern generated by the DOEs, the same setup described in previous section (Fig. 5.14) has been used.

Fig. 5.21 (third row) shows the pictures of the pattern projected by the fabricated DOEs illuminated by a laser, designed with (f) and without correction (e). They are compared with the target pictures with and without pre-distortion (first row of Fig. 5.21) and the ones achieved in the simulation after having used the IFTA (second row of Fig. 5.21).

One can note that the pattern produced by the DOE designed without the correction (subfigure (e)) has a pincushion distortion. In (f) there are no distortions. Unfortunately, the lines are not straight because of misalignment in the set up between the laser the DOE, the screen and the camera. The zero order and +1 order are visible because of fabrication errors. For knowing the etching errors, the levels height have been measured with an interferometer. The height that induces a phase of $\pi$ should be 723nm and the measured one is 680nm, the $\pi/2$ should be 361nm but it is 300nm, the $3\pi/2$ should be 1084nm but it is 980nm. These height differences cause the appearance of undesired diffraction orders. Errors could be also caused by a wrong alignment.

The pattern FoV at different distances between the DOE and a screen has been measured. After fitting the data, an error between the desired FoV and the measured one $< 1°$ has been found.

At the end, the aim desired has been achieved. A DOE, designed in a simple way using IFTA, that projects the desired wide angle pattern without distortion has been fabricated and the problems caused by the fabrication errors have been analyzed.

## 5.6 Conclusion

A new method to design DOEs for large pattern angles using IFTA has been described. The IFTA has been illustrated, improved with soft quantization developed by Wyrowski. Moreover, the method for correcting the pattern distortion caused by using large angles in a Fourier Optics approach has been presented. The main DOE fabrication problems have been illustrated. Finally the experimental results have shown that the correction works successfully.
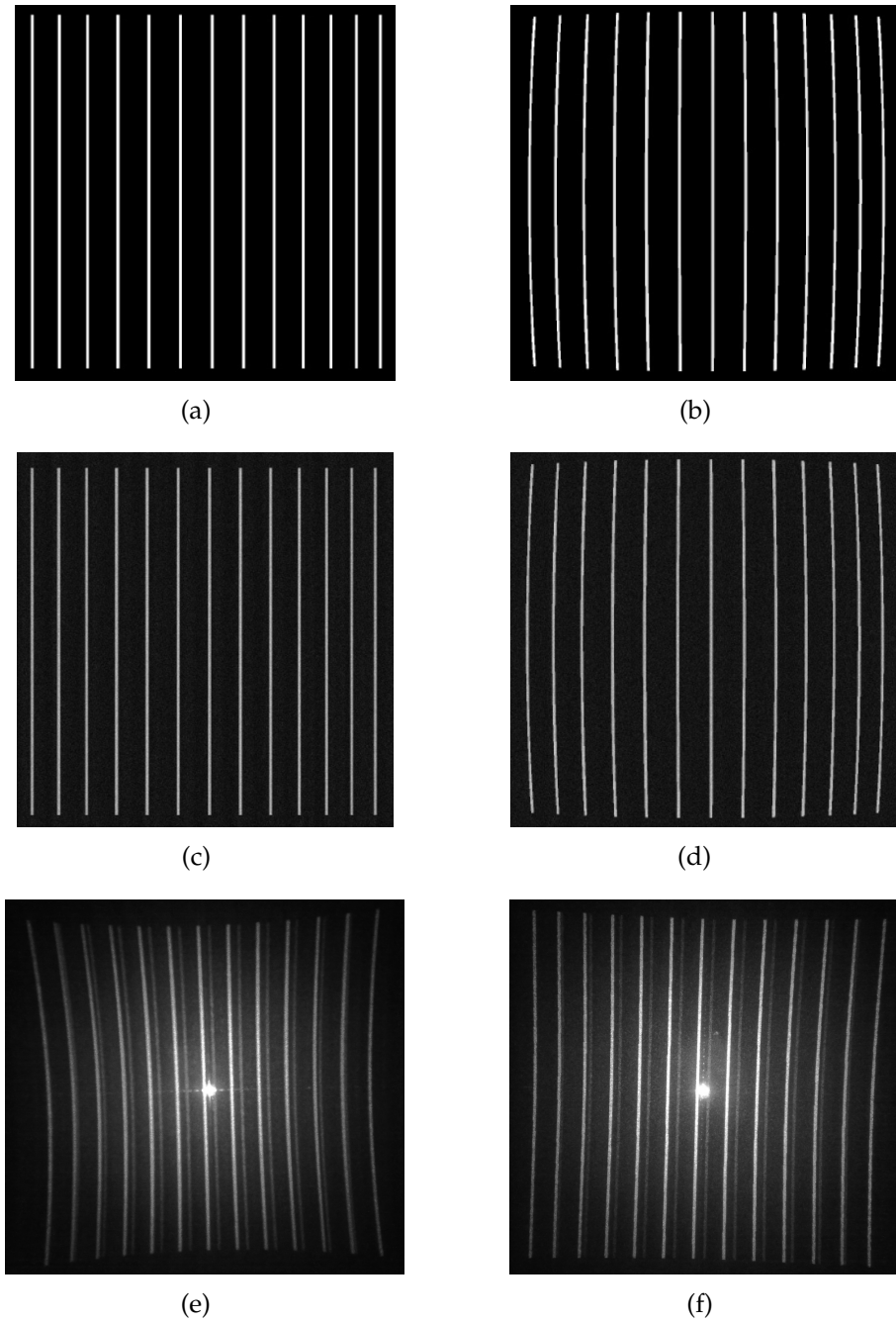
Figure 5.21: In the first row, there are the amplitudes used as constraints (a), (b), in the second row, the amplitudes obtained by IFTA (c), (d) and in the third row pictures of the pattern projected by the fabricated DOEs (e), (f). The first column corresponds to images without distortion correction (a), (c), (e) and the second one with it (b), (d), (f).

# Conclusions

This thesis has tackled the problem of conceiving a design philosophy and the related suitable codes to help designers of an industrial environment to develop their activities without too much concern on software set up. The activity has been carried out in close contact with Datalogic Scanning Group s.r.l., the industry which promoted and funded this PhD activity. This helped in understanding details about the industrial design process, such as how parameters to be optimized have to be chosen, how the required performance can be achieved, how the fabrication tolerance can impact on the design etc.. The approach to the problem has been conceived to be as general as possible for what concern its applications, although in the present work, the design of a portable device for automatic identification (Auto ID) applications has been studied as a case study to provide suitable examples. The complete device has been studied and its main functional components have been identified first: electro-optical imaging system, illumination system and pattern generator system. Several design tools have then been implemented for improving the optical design of all the subparts of the whole device, either increasing the overall performance optimization or proposing new design methods. For validating the design tools, the performance of the designed systems has been measured, thanks to the realization of prototypes, or simulated. In the case of electro-optical system design, tools for its analysis and characterization have been set up. The comparison between the measured or simulated results with the expected ones demonstrates that the design tools work successfully and that the characterization and analysis tools are reliable. For this reason, most of them are now used in Datalogic Scanning Group s.r.l..

Optical design of the electro-optical imaging system has been discussed in the first part of the thesis, while in the second part illumination system and pattern generator system have been dealt.

Since design requires a characterization process to check if the desired performance has been achieved, this problem, applied to the imaging system design, has been addressed first. In particular, Ch. 1 dealt with the characterization of electro-optical imaging system performance. For this purpose, a tool named SA-FARILAB (SFR measurement for a LAB environment) has been set up. It has been used extensively for measuring the performance of the designed system shown in Ch. 2. SAFARILAB evaluates the Spatial Frequency Response (SFR) of an imaging system complying with the ISO 12233 standard, the reference standard for

this kind of measurements. SAFARILAB tool consists of both the algorithm for the SFR evaluation and the experimental set-up. Taking advantage of the degrees of freedom left by the standard, new options have been added to the algorithm to improve the numerical calculations and to reduce the noise that affects the measurements. For validating this tool, several tests have been done. The SAFARI-LAB results obtained for synthetic and experimental tests have shown that they can be successfully compared with those of other freely or commercially available software dedicated to the same task. Further tests have shown the SAFARILAB excellent behavior in terms of repeatability and ability to filter out random noise effects, especially thanks to the options added. The tool has been used to compare the measured optical SFR and that designed with an optical CAD, ZEMAX®. This successful verification is an essential step to validate the whole optical system design process.

The design of an electro-optical system can be made at different levels, such as optimizing just the optical part (optical level) or the whole system including optics, electronics and detection (system level). Both are important because the first has the advantage to optimize just the core of the system improving its performance ignoring all other contributions and generating a good starting point for the optimization of the whole complex system (second approach). On the other hand, the second approach allows to optimize the system taking into account its behavior with a model as near as possible to reality.

In Ch. 2, the design at optical level is presented. The tool implemented is named Optical Level OptiMization, OLOM . It designs the system only at its optical level, ignoring the effects of the other parts. This tool, working together with the ZEMAX® optimization engine, allows to set constraints on the Modulation Transfer Functions (MTFs) of the optical part of the system following required specifications. The Cubic Phase Mask system (CPM), one of the Wavefront Coding (WFC) based systems, has been chosen for validating OLOM , because it shows the best performance for the DoF extension of an optical system (an important characteristic for Auto ID applications) and the most solid theoretical bases, resulting from the optimization of the MTF. The optimization of the CPM system using OLOM has been done considering the requirements of a barcode reader, taken into account as an example of Auto ID and portable system application. The fabrication tolerance analysis of the CPM has shown that adding and subtracting the tolerance value to the optimized cubic shape, the differences in terms of Depth of Field (DoF) are about the $5\%$-$9\%$ of the exact CPM DoF. All the three case have been fabricated. Both the profilometric and optical characterization of the fabricated prototypes have demonstrated the accuracy of the technological processes and the agreement between the CPM realized and the designed ones. The CPM system performance has been measured in two ways. Firstly the MTFs at infinity of the lens with and without system have been measured for different Back Focal Length (BFL). The results have not only shown the agreement between the simulated and the measured data, but also confirmed the Extended Depth of

Field (EDoF) produced by the optimized CPM. Secondly, SFRs of the system have been measured using SAFARILAB . Measurements confirm again that the CPM optimized by OLOM enhances the performance of a lens, increasing its DOF almost doubled. The whole design procedure is reliable, since the results obtained by theoretical means are confirmed by the characterized behavior of the realized prototypes. OLOM tool was validated and its characteristic to increase the performance of a system in terms of DoF has been demonstrated. Thanks to the successful validation of OLOM , it can be used for designing the starting point to be used in the system level optimization.

Ch. 3 describes the main features of the software package named System Level AnaLysis and OptiMization, SLALOM , able to study and design optical systems including also the effect related to the presence of electronics and detection. SLALOM consists of two parts one for the optimization task performed by Optimization-SLALOM , O-SLALOM , and the other for the analysis task performed by Analysis-SLALOM , A-SLALOM . The two tools leverage on an extensive library of C modules that are implemented for the modeling of optical systems (in conjunction with ZEMAX$^{®}$ ), photosensors (taking into account pixel sampling and shot noise generation), and image restoration (via digital Wiener filtering). These modules allow the use of ZEMAX$^{®}$ beyond the optical domain.

The SLALOM tool, set up in the optics group of DEIS, has been used to design a system based on the presence of a large spherical aberration. It provides a Point Spread Function (PSF) with circular symmetry that allows to use a conventional post-processing. Such design has been done first optimizing some of the parameters, such as the reconstruction filter parameters, using A-SLALOM , and then optimizing the whole system using O-SLALOM . A good starting point for the two considered has been found by OLOM to be optimized at system level by O-SLALOM . The results obtained by O-SLALOM optimization, have demonstrated the ability of both optimized designs to extend the depth of field with respect to the corresponding starting designs. The analysis of the merit function sensitivity shows that it is low to all the optimized parameters but the BFL: even if this can create problems during the optimization of the system, it represents a good result taking into consideration possible fabrication errors. In conclusion, results show that O-SLALOM can be successfully used for system level design since it allows to obtain the almost absolute minimum of the merit function. However, results also show that spherical aberration, although capable of impacting on the lens DoF, does not allow to extend it so much as to meet all the required specifications. Furthermore, A-SLALOM is not only useful for analyzing the imaging systems and simulating their behavior but it also allows to find reconstruction parameters for the optimization process.

Ch. 4 shows a tool for the design of non-imaging systems with an incoherent source. It allows the design of free-form lenses described by an arbitrary analytical function and is able to provide custom illumination conditions for all kind of applications. The designed free-form lens has been fabricated and tested. The

measured results satisfy the illumination requirements and confirm the quality of the implemented algorithm. Another advantage of the tool is that it is realized in MATLAB® and allows not to use expensive commercial optical design tools, which is important when free-form lens design is not routinely done and cost issues become relevant.

Ch. 5 shows a tool for the design of non-imaging systems with a coherent source. It consists of a new method to design Diffractive Optical Elements (DOEs) for large pattern angles using the Iterative Fourier Transform Algorithm (IFTA). The IFTA has been improved with soft quantization developed by Wyrowski and with the method for correcting the pattern distortion caused by using large angles in a Fourier Optics approach. The main DOE fabrication problems have been illustrated. Finally the experimental results show that the correction works successfully, even if fabrication errors occurred.

# List of publications

P1. <u>L. De Marco</u>, B. Hallal, F. Canini, M. Gnan, P. Bassi,
"Spatial Frequency Response Characterization of Optical Imaging Systems,"
*Proceedings XVIII RiNEm* Prima Riunione Nazionale URSI, Benevento, 6-10
Settembre 2010.

P2. <u>L. De Marco</u>, A. Guagliumi, M. Gnan, B. Hallal, F. Canini, P. Bassi,
"SAFARiLAB : a rugged and reliable MTF measurement system for industrial environment,"
ATEAS, 21 January 2012.

# Bibliography

[1] "Standard ISO 12233 (Photography - Electronic Still Picture Cameras - Resolution Measurements)." [Online]. Available: http://www.iso.org/iso/catalogue_detail.htm?csnumber=33715

[2] ZEMAX®, a software for optical system design. [Online]. Available: http://www.zemax.com

[3] E. R. Dowsky and W. T. Cathey, "Extended depth of field through wave-front coding," *Applied Optics*, vol. 34, no. 11, pp. 1859–1866, 1995.

[4] M. D. Robinson and D. G. Stork, "Comparing the extended depth-of-field imaging performance of spherical aberration and asymmetric wavefront coding." [Online]. Available: http://rii.ricoh.com/technology_research/publications/dor/COSIsphericalcodingv1.pdf

[5] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik (Stuttgart)*, vol. 35, no. 2, pp. 237–250, 1972.

[6] J. R. Fienup, "Iterative method applied to image reconstruction and to computer-generated holograms," *Optical Engineering*, vol. 19, no. 3, pp. 297–305, 1980.

[7] F. Wyrowski and O. Bryngdahl, "Iterative Fourier-transform algorithm applied to computer holography," *Journal of Optical Society of America A*, vol. 5, no. 7, pp. 1058–1065, 1988.

[8] IMATEST software. [Online]. Available: http://www.imatest.com/products/software/master

[9] IMAGEJ, Image Processing and Analysis in Java. [Online]. Available: http://rsbweb.nih.gov/ij/

[10] J. W. Goodman, *Introduction to Fourier Optics*. McGraw-Hill, 1996.

[11] M. Born and E. Wolf, *Principles of Optics*. Pergamon press, 1980.

[12] P. Bassi, G. Bellanca, and G. Tartarini, *Propagazione ottica libera e guidata*. CLUEB, 1999.

[13] G. D. Boreman, *Modulation Transfer Function in Optical and Electro-Optical Systems*. SPIE press, 2001.

[14] J. R. Janesick, *Photon Transfer*. SPIE press, 2007.

[15] J. H. McLeod, "The Axicon: A New Type of Optical Element," *Journal of Optical Society of America*, vol. 44, no. 8, pp. 592–597, 1954.

[16] J. H. McLeod, "Axicons and Their Uses," *Journal of Optical Society of America*, vol. 50, no. 2, pp. 166–169, 1960.

[17] A. Kołodziejczyk, S. Bará, Z. Jaroszewicz, and M. Sypek, "The Light Sword Optical Element - a New Diffraction Structure with Extended Depth of Focus," *Journal of Modern Optics*, vol. 37, no. 8, pp. 1283–1286, 1990.

[18] G. Mikuła, A. Kołodziejczyk, M. Makowski, C. Prokopowicz, and M. Sypek, "Diffractive elements for imaging with extended depth of focus," *Optical Engineering*, vol. 44, no. 5, p. 058001, 2005.

[19] G. Mikuła, Z. Jaroszewicz, A. Kołodziejczyk, K. Petelczyc, and M. Sypek, "Imaging with extended focal depth by means of lenses with radial and angular modulation," *Optics Express*, vol. 15, no. 15, pp. 9184–9193, 2007.

[20] J. Ares Garca, S. Bará, M. Gomez Garca, Z. Jaroszewicz, A. Kołodziejczyk, and K. Petelczyc, "Imaging with extended focal depth by means of the refractive light sword optical element," *Optics Express*, vol. 16, no. 22, pp. 18371–18378, 2008.

[21] L. Kipp, M. Skibowski, R. Johnson, R. Berndt, R. Adelung, S. Harm, and R. Seemann, "Sharper images by focusing soft X-rays with photon sieves," *Nature*, vol. 414, pp. 184–188, 2001.

[22] C. Iemmi, J. Campos, J. Escalera, O. López-Coronado, R. Gimeno, and M. J. Yzuel, "Depth of focus increase by multiplexing programmable diffractive lenses," *Optics Express*, vol. 14, no. 22, pp. 10207–10219, 2006.

[23] A. Márquez, C. Iemmi, J. Campos, and M. J. Yzuel, "Achromatic diffractive lens written onto a liquid crystal display," *Optics Letters*, vol. 31, no. 3, pp. 392–394, 2006.

[24] W. Furlan, G. Saavedra, and J. Monsoriu, "White-light imaging with fractal zone plates," *Optics Letters*, vol. 32, no. 15, pp. 2109–2111, 2007.

[25] E. Ben-Eliezer, Z. Zalevsky, E. Marom, and N. Konforti, "All optical extended depth of field imaging system," *Journal of Optics A: Pure and Applied Optics*, vol. 5, no. 5, pp. S164–S169, 2003.

[26] E. Ben-Eliezer, E. Marom, N. Konforti, and Z. Zalevsky, "Experimental realization of an imaging system with an extended depth of field," *Applied Optics*, vol. 44, no. 14, pp. 2792–2798, 2005.

[27] A. Papoulis, "Ambiguity function in Fourier optics," *Journal of the Optical Society of America*, vol. 64, no. 6, pp. 779–788, 1974.

[28] A. Brenner, A. W. Lohmann, and J. Ojeda-Castaneda, "The ambiguity function as a polar display of the OTF," *Optics Communications*, vol. 44, no. 5, pp. 323–326, 1983.

[29] H. Bartelt, J. Ojeda-Castaneda, and E. Sicre, "Misfocus tolerance seen by simple inspection of the ambiguity function," *Applied Optics*, vol. 23, no. 16, pp. 2693–2696, 1984.

[30] W. T. Cathey and E. R. Dowski, "New paradigm for imaging systems," *Applied Optics*, vol. 41, no. 29, pp. 6080–6092, 2002.

[31] J. Van Der Gracht, E. R. Dowsky, M. G. Taylor, and D. M. Deaver, "Broadband behavior of an optical-digital focus-invariant system," *Optics Letters*, vol. 21, no. 13, pp. 919–921, 1996.

[32] S. Bradburn, W. T. Cathey, and E. R. Dowsky, "Realizations of focus invariance in optical-digital systems with wavefront coding," *Applied Optics*, vol. 36, no. 35, pp. 9157–9166, 1997.

[33] S. Bagheri, D. Pucci De Farias, and P. E. X. Silveira, "Analytical optimal solution of the extension of the depth of field using cubic-phase wavefront coding. Part I. Reduced-complexity approximate representation of the modulation transfer function," *Journal of the Optical Society of America A*, vol. 25, no. 5, pp. 1051–1063, 2008.

[34] S. Bagheri, P. E. X. Silveira, R. Narayanswamy, and D. Pucci De Farias, "Analytical optical solution of the extension of the depth of field using cubic-phase wavefront coding. Part II. Design and optimization of the cubic phase," *Journal of the Optical Society of America A*, vol. 25, no. 5, pp. 1064–1074, 2008.

[35] S. S. Sherif, W. T. Cathey, and E. R. Dowsky, "A logarithmic phase filter to extend the depth of field of incoherent hybrid imaging systems," *Proceedings of SPIE conference on "Algorithms and Systems for Optical Information Processing V"*, vol. 4471, pp. 272–280, 2001.

[36] S. S. Sherif, W. T. Cathey, and E. R. Dowsky, "Phase plate to extend the depth of field of incoherent hybrid imaging systems," *Applied Optics*, vol. 43, no. 13, pp. 2709–2721, 2004.

[37] C. Fales, F. Huck, and R. Samms, "Imaging system design for improved information capacity," *Applied Optics*, vol. 23, no. 6, pp. 872–888, 1984.

[38] M. D. Robinson and D. G. Stork, "Theoretical foundations for joint digital-optical analysis of electro-optical imaging systems," *Applied Optics*, vol. 47, no. 10, pp. B64–B75, 2008.

[39] M. D. Robinson, G. Feng, and D. G. Stork, "Spherical coded imagers: Improving lens speed, depth-of-field, and manufacturing yield through enhanced spherical aberration and compensating image processing," *Proceedings of SPIE conference on "Novel Optical Systems Design and Optimization XII"*, vol. 7429, p. 74290M, 2009.

[40] Q. Yang, J. Liu, and J. Sun, "Optimized phase pupil masks for extended depth of field," *Optics Communications*, vol. 272, no. 1, pp. 56–66, 2007.

[41] N. Caron and Y. Sheng, "Polynomial phase masks for extending the depth of field of a microscope," *Applied Optics*, vol. 47, no. 22, pp. E39–E43, 2008.

[42] F. Zhou, G. Li, H. Zhang, and D. Wang, "Rational phase mask to extend the depth of field in optical-digital hybrid imaging systems," *Optics Letters*, vol. 34, no. 3, pp. 380–382, 2009.

[43] F. Zhou, R. Ye, G. Li, H. Zhang, and D. Wang, "Optimized circularly symmetric phase mask to extend the depth of focus," *Journal of the Optical Society of America A*, vol. 26, no. 8, pp. 1889–1895, 2009.

[44] Y. Takahashi and S. Komatsu, "Optimized free-form phase mask for extension of depth of field in wavefront-coded imaging," *Optics Letters*, vol. 33, no. 13, pp. 1515–1517, 2008.

[45] F. Zhao, Y. Li, H. Feng, Z. Xu, and Q. Li, "Cubic sinusoidal phase mask: Another choice to extend the depth of field of incoherent imaging system," *Optics and Laser Technology*, vol. 42, no. 4, pp. 561–569, 2010.

[46] H. Zhao and Y. Li, "Optimized sinusoidal phase mask to extend the depth of field of an incoherent imaging system," *Optics Letters*, vol. 35, no. 2, pp. 267–269, 2010.

[47] A. K. Jain, *Fundamental of digital image processing*. Prentice Hall, 1989.

[48] P. Z. Peebles, *Probability, Random Variables and Random Signal Principles*. McGraw-Hill, 1987.

[49] M. R. Krames, O. B. Shchekin, R. Mueller-Mach, G. O. Mueller, L. Zhou, G. Harbers, and M. G. Craford, "Status and Future of High-Power Light-Emitting Diodes for Solid-State Lighting," *IEEE Journal of Display Technology*, vol. 3, no. 2, pp. 160–175, 2007.

[50] L. Sun, S. Jin, and S. Cen, "Free-form microlens for illumination applications," *Applied Optics*, vol. 48, no. 29, pp. 5520–5527, 2009.

[51] D. Vincenzi and G. Oliva, "Tailoring freeform lenses for bar code lighting," *Proceedings of 2007 IEEE Workshop on "Automatic Identification Advanced Technologies"*, pp. 134–139, 2007.

[52] H. Ries and J. Muschaweck, "Tailored freeform optical surfaces," *Journal of the Optical Society of America A*, vol. 19, no. 3, pp. 590–595, 2002.

[53] K. Wang, S. Liu, F. Chen, Z. Liu, and X. Luo, "Effect of manufacturing defects on optical performance of discontinuous freeform lenses," *Optics Express*, vol. 17, no. 7, pp. 5457–5465, 2009.

[54] F. Fournier and J. Rolland, "Optimization of freeform lightpipes for light-emitting-diode projectors," *Journal of the Optical Society of America A*, vol. 47, no. 7, pp. 957–966, 2008.

[55] F. Fournier, W. Cassarly, and J. Rolland, "Fast freeform reflector generation using surce-target maps," *Optics Express*, vol. 18, no. 5, pp. 5295–5304, 2010.

[56] V. Oliker, "On design of free-form refractive beam shapers, sensitivity to figure error, and convexity of lenses," *Journal of the Optical Society of America A*, vol. 25, no. 12, pp. 3067–3074, 2008.

[57] J. Minano, P. Benitez, and A. Santamaria, "Free-form Optics for illumination," *Optical Review*, vol. 16, no. 2, pp. 99–102, 2009.

[58] ASAP®. [Online]. Available:
http://www.breault.com/software/asap.php

[59] OSLO®. [Online]. Available:
http://www.sinopt.com/

[60] TRACEPRO®. [Online]. Available:
http://www.lambdares.com/software_products/tracepro/

[61] H. Aagedal, F. Wyrowski, and M. Schmid, "Paraxial beam splitting and shaping," in *Diffractive Optics for Industrial and Commercial Applications*, J. Turunen and F. Wyrowski, Eds. Akademie Verlag, 1997, ch. 6, pp. 165–188.

[62] W. H. Lee, "III Computer-Generated Holograms: Techniques and Applications," in *Progress in Optics*, E. Wolf, Ed. Elsevier, 1978, vol. 16, pp. 119–232.

[63] G. R. Hadley, "Multistep method for wide-angle beam propagation," *Optics Letters*, vol. 17, no. 24, pp. 1743–1745, 1992.

[64] P. Chamorro-Posada, G. S. McDonald, and G. H. C. New, "Non paraxial beam propagation method," *Optics Comunications*, vol. 192, pp. 1–12, 2001.

[65] A. Sharma and A. Agrawal, "New method for nonparaxial beam propagation," *Journal of the Optical Society of America*, vol. 21, no. 6, pp. 1082–1087, 2004.

[66] J. E. Harvey, A. Krywonos, and D. Bogunovic, "Nonparaxial scalar treatment of sinusoidal phase gratings," *Journal of the Optical Society of America A*, vol. 23, no. 4, pp. 858–865, 2006.

[67] J. R. Fienup and C. C. Wackerman, "Phase-retrieval stagnation problems and solutions," *Journal of the Optical Society of America A*, vol. 3, no. 11, pp. 1897–1907, 1986.

[68] F. Wyrowski, "Iterative quantization of digital amplitude holograms," *Applied Optics*, vol. 28, no. 18, pp. 3864–3870, 1989.

[69] F. Wyrowski, "Diffractive optical elements: iterative calculation of quantized, blazed phase structures," *Journal of Optical Society of America A*, vol. 7, no. 6, pp. 961–969, 1990.

[70] H. P. Herzig, *Micro-optics: Elements, Systems And Applications*. CRC Press, 1997.

[71] B. Kress and P. Meyrueis, *Digital Diffractive Optics*. J. Wiley & Sons, 2000.

[72] S. Sinzinger and J. Jahns, *Microoptics*. Wiley-VCH, 2003.

[73] G. J. Swanson and W. B. Veldkamp, "Diffractive optical elements for use in infrared systems," *Optical Engineering*, vol. 28, no. 6, pp. 605–608, 1989.

[74] H. P. Herzig and R. Dändliker, "Holographic optical scanning elements: Analytic method for determining the phase function," *Journal of Optical Society of America A*, vol. 4, no. 6, pp. 1063–1070, 1987.

[75] J. R. Fienup, "Iterative method applied to image reconstruction and to computer generated holograms," *Optical Engineering*, vol. 19, no. 3, pp. 297–305, 1980.

[76] B. K. Jennison, J. P. Allebach, and D. W. Sweeney, "Iterative approaches to Computer generated holography," *Optical Engineering*, vol. 28, no. 6, pp. 629–637, 1989.

[77] F. Wyrowski, "Design theory of diffractive elements in the paraxial domain," *Journal of Optical Society of America A*, vol. 10, no. 7, pp. 1553–1561, 1993.

[78] X. Tan, B. Y. Gu, G. Z. Yang, and B. Z. Dong, "Diffractive phase elements for beam shaping: a new design method," *Applied Optics*, vol. 34, no. 8, pp. 1314–1320, 1995.

[79] G. Z. Yang, B. Z. Dong, B. Y. Gu, J. Y. Zhuang, and O. K. Ersoy, "Gerchberg-Saxton and Yang-Gu algorithms for phase retrieval in a nonunitary transform system: a comparison," *Applied Optics*, vol. 33, no. 2, pp. 209–218, 1994.

[80] S. Ettl, O. Arold, F. Willomitzer, Z. Yang, and G. Häusler, ""Flying Triangulation" - Acquiring the 360 Topography of the Human Body on the Fly," *Proceedings of the International Conference on "3D Body Scanning Technologies"*, pp. 279–284, 2010.

[81] F. Willomitzer, Z. Yang, S. Ettl, O. Arold, and G. Häusler, "3D face scanning with "Flying Triangulation"," *DGaO-Proceedings*, 2010.