



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Vladislav Uzunov

**Light Field Reconstruction and Synthesis from
Irregularly Distributed Data Points**



Julkaisu 987 • Publication 987

Tampereen teknillinen yliopisto. Julkaisu 987
Tampere University of Technology. Publication 987

Vladislav Uzunov

Light Field Reconstruction and Synthesis from Irregularly Distributed Data Points

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 7th of October 2011, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2011

ISBN 978-952-15-2653-4 (printed)
ISBN 978-952-15-2658-9 (PDF)
ISSN 1459-2045

Abstract

A scalar light field is any stationary monochromatic wave in the visible spectrum which propagates in an isotropic, homogeneous, non-dispersive and non-magnetic medium and thus obeys the Helmholtz wave equation. This thesis presents novel methods and algorithms related to the synthesis and reconstruction of a scalar light field, given a set of irregularly-distributed target values within a spatial region of interest. The set of target values and their positions are determined and constrained by some practical application which requires a light field. The light field synthesis task is considered in the thesis as split into two complementary subtasks. The first subtask is to reconstruct numerically a scalar light field whose values at the given locations approximate the target values of the application. The second subtask is to find the optimal configuration of an optical device, e.g. a deflectable micromirror device, which can modulate an incident light wave to produce the desired, reconstructed light field.

The core of the thesis studies problems related to the irregular sampling and reconstruction of scalar light fields. The aim is to represent a continuous light field, which satisfies a certain integral or a differential equation, by a set of discrete, irregularly-distributed samples in a manner which can be stably inverted. This study includes algorithms which reconstruct a continuous field from a discrete set of samples, together with the tools for analyzing the stability and convergence of this reconstruction. Two basic reconstruction approaches are developed. One of them is based on the method of Projection Onto Convex Sets (POCS) and the other one on least squares optimization. The POCS-based approach defines one set per transversal plane or, more generally, one set per sample. The projections from set to set are done by light propagation with the plane-wave decomposition integral. The least squares -based approaches developed in this thesis model the continuous field as expansion over a discrete set of basis functions, defining a linear space of the considered light fields. Two models have been derived – one based on the plane-wave decomposition integral, and the other one on cylindrical harmonic decomposition, involving Bessel functions of the first kind. The reconstruction is done by solving a linear system of equations, obtained by expressing each known sample by equation through the model. The linear system is solved rapidly by an iterative algorithm – the method of conjugate gradients. The weighted least squares is used to accelerate moderately the convergence of the conjugate gradients method. An advantage of the least squares-based approach is the possibility to analyze the stability of the irregular sampling set and the convergence of the iterative solver. Different sample distributions influence the stability of the reconstruction, which are analyzed with the singular values of the linear system.

The light field synthesis part of the thesis studies two optimization algorithms to find the configuration of the deflectable micro-mirror device which synthesizes to best extent a desired light wave, restricted to a transversal plane. The first approach is a precise and extensive search based on the method of simulated annealing. The second approach is a fast and greedy optimization algorithm based on matching pursuit method for signal decomposition. The field is assumed to satisfy the Fresnel diffraction integral. However, the algorithms can be easily adapted to serve any light propagation model.

The main outcome of the thesis is a general framework to be used in light field reconstruction and synthesis applications. The framework components are interconnected together and provide the following functionalities:

- modeling continuous light fields by linear expansions;
- reconstructing continuous light fields by rapid iterative algorithms;
- studying the stability of a given sampling set related to reconstruction;
- modeling the physical inconsistency of the sample values related to an application;
- optimal configuration of an optical device for synthesizing a desired light field on a plane.

The components of this framework are consistent with each other, and yet, each component can be substituted with an alternative. The alternative would still preserve the structure of the framework and would function together with the other components. As such, this framework can be used for evaluation and comparison of alternatives for a certain module. Such flexibility allows adaptation to any application and any inherent innovation.

The developed framework for light field reconstruction and synthesis serves applications related to the following scientific areas: digital holography, computer generated holography and beam shaping. Industrial applications, e.g. lithography, interferometry, microscopy and particle manipulation benefit from the ability to manipulate a light wavefront. Free view-point, enabled by the ability to capture and display three-dimensional scenes, improves environmental applications in education, medicine, TV and the gaming industry, computer aided design, traffic control, amongst others.

Preface

The work presented in this thesis has been carried out at the Department of Signal Processing at Tampere University of Technology during the years 2006-2010. Part of the research was carried out at the Department of Electrical and Electronics Engineering at Bilkent University, Turkey, during my visits there in 2006 and 2007.

Above all, I would like to express my gratitude to my supervisor Prof. Karen Egiazarian for his constant support, bright ideas and valuable advice. I am also indebted to my co-supervisor, Dr. Atanas Gotchev, for introducing me to the exciting world of science from the very basics, through the polishing of advanced scientific concepts to the accomplishments of scientific publishing. I am grateful to Prof. Onural and Prof. Ozaktas from Bilkent University, Turkey, for acquainting me with the subject of this thesis and sharing their valuable knowledge. Many special thanks go to my collaborators and co-authors Gökhan Bora Esmer and Erdem Ulusoy from Bilkent University, Turkey, for their valuable support in all our publications and for becoming good friends in a pleasant working atmosphere.

The thesis was financially supported by the Tampere Graduate School in Information Science and Engineering (TISE), the FP7 3DTV NoE project (Grant 511568) and by the Academy of Finland funded project “High-resolution digital holography: A modern signal processing approach”.

Many thanks to all my colleagues for creating a nice and friendly atmosphere and their support in difficult moments. It is this little help, advice or a joke which cheers you up in the morning and makes for a whole productive day. An incomplete list of their names includes Alessandro, Annamaria, Aram, Atanas, Danilo, Dima Paliy, Dima Rusanovskiy, German, Giacomo, Ilian, Kivanc Kose, Kostadin, Mihail, Polat, Robert, Stanislav, Susanna Minasyan, Prof. Vladimir Katkovnik, and many others.

Finally, I would like to thank my parents, Ivan and Bistra, and my sister Violeta for all the care, moral and practical support. And last, but far from least, a warm hug to Lyupka for the love.

Tampere, September 2011

Vladislav Uzunov

Contents

Abstract	i
Preface	iii
List of Publications	ix
List of Abbreviations	xi
List of Symbols	xiii
1 Introduction	1
1.1 Background and motivation	1
1.2 Thesis objectives	5
1.3 Thesis outline	6
2 Preliminaries	7
2.1 Scalar Diffraction Theory	7
2.1.1 Light waves in space and frequency domain	8
2.1.2 Rayleigh-Sommerfeld diffraction integral	9
2.1.3 Plane-wave decomposition integral	12
2.1.4 Fresnel approximation	16

2.1.5	Conclusions	18
2.2	Sampling	19
2.2.1	Classical sampling theory	19
2.2.2	Frame theory	20
2.2.3	Irregular sampling theory and frames	26
2.2.4	Reconstruction from irregular samples	27
2.2.5	Conclusions	32
2.3	Sampling of diffraction fields	33
2.3.1	Degrees of freedom of a diffraction field	33
2.3.2	Results based on the Nyquist theorem	36
2.3.3	Results based on local bandwidth	38
2.3.4	Generalizations for a volume	40
2.3.5	Conclusions	41
2.4	Light field reconstruction and synthesis	43
2.4.1	Applications of LF reconstruction and synthesis	44
2.4.2	Light wave modulation elements	46
2.4.3	LF synthesis on a target plane	48
2.4.4	LF reconstruction inside a volume	51
2.4.5	Conclusions	55
3	Methods for LF reconstruction and synthesis	57
3.1	Problem formulation and hypotheses	57
3.2	LF reconstruction from samples on irregular grids	61
3.2.1	POCS-based methods	61
3.2.2	Least squares-based methods	65
3.3	Light field synthesis with DMD	76
3.3.1	Rectangular aperture illuminated by a plane wave	77
3.3.2	Mathematical formulation	79
3.3.3	Light field synthesis methods	80

4 Summary of results

List of Publications

This thesis is a compound, based on the following publications. In the text, these publications are referred to as **P1**, **P2** and so on.

- P1** E. Ulusoy, **V. Uzunov**, L. Onural, H.M. Ozaktas, A. Gotchev, “Three-dimensional monochromatic light field synthesis with a deflectable mirror array device,” in *Photon Management II: SPIE Proceedings*, vol. 6187, Strasbourg, France, April 3-7 2006, pp. 1J-1 – 1J-12.
- P2** **V. Uzunov**, A. Gotchev, G.B. Esmer, H.M. Ozaktas, L. Onural, “Non-uniform sampling of diffraction field,” in *Proceedings of The 2006 International Workshop on Spectral Methods and Multirate Signal Processing, SMMSP2006, TICSP Series*, vol. 34, Florence, Italy, 2-3 September 2006, pp. 191-197.
- P3** G.B. Esmer, **V. Uzunov**, L. Onural, H.M. Ozaktas, A. Gotchev, “Diffraction field computation from arbitrarily distributed data points in space,” in *Signal Processing: Image Communication*, vol. 22, No. 2, 2007, pp. 178-187.
- P4** G.B. Esmer, **V. Uzunov**, L. Onural, A. Gotchev, H.M. Ozaktas, “Reconstruction of scalar diffraction field from distributed data points over 3D space,” in *Proceedings of the International Conference: True-Vision, Capture, Transmission, and Display of 3D Video, 3DTV-CON 2007*, Kos Island, Greece, 7-9 April 2007, pp. 1-4.
- P5** **V. Uzunov**, G.B. Esmer, A. Gotchev, L. Onural, H.M. Ozaktas, “Bessel functions-based reconstruction of non-uniformly sampled diffraction fields,” in *Proceedings of the International Conference: True-Vision, Capture, Transmission, and Display of 3D Video, 3DTV-CON 2007*, Kos Island, Greece, 7-9 April 2007, pp. 1-4.
- P6** **V. Uzunov**, A. Gotchev, K. Egiazarian, “Convergence and error analysis of diffraction field iterative non-uniform sampling schemes,” in *Proceedings of the 2008 International Workshop on Local and Non-Local Approximation in Image Processing, LNLA2008, TICSP Series*, vol. 44, Lausanne, Switzerland, 23-24 August 2008, pp. 213-220.
- P7** **V. Uzunov**, A. Gotchev, K. Egiazarian, “Regularized reconstruction of irregularly sampled scalar light fields,” in *Proceedings of Information Optics (WIO), 2010 9th Euro-American Workshop on*, 12-16 July 2010, pp. 1-3.

- P8 V. Uzunov**, A. Gotchev, K. Egiazarian, “On the stability of reconstruction of irregularly sampled diffraction fields,” in *Advances in Optical Technologies*, vol. 2010, Article ID 138024, 2010, 12 pages.

List of Abbreviations

LF	Light Field
DOE	Diffractive Optical Element
SLM	Spatial Light Modulator
POCS	Projection Onto Convex Sets
2D	Two-dimensional
3D	Three-dimensional
3DTV	Three-dimensional Television
CGH	Computer Generated Hologram
CAD	Computer Aided Design
DMD	Digital Micromirror Device
RGB	Red, Green, Blue color system
PWD	Plane-Wave Decomposition
R-S	Rayleigh-Sommerfeld
FFT	Fast Fourier Transform
FrFT	Fractional Fourier Transform
DFT	Discrete Fourier Transform
SVD	Singular Value Decomposition
CG	Conjugate Gradients
NDF	Number of Degrees of Freedom
LB/MLB	Local Bandwidth / Maximum Local Bandwidth
LS	Least Squares

List of Symbols

(x, y, z)	Spatial Cartesian coordinate system.
j	Imaginary unit.
u	Scalar light field.
k_x	Angular frequency along x .
k_y	Angular frequency along y .
k_z	Angular frequency along z .
$\nabla^2 f$	Laplacian differential operator, applied on the function f : $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$.
λ	Wavelength of a monochromatic wave field.
$k = \frac{2\pi}{\lambda}$	Wave number of a monochromatic wave field.
$\mathcal{F}\{u\}$	Fourier transform of the function u : $\mathcal{F}\{u\}(k_x) = \int_{-\infty}^{\infty} u(x)e^{-jk_x x} dx$.
$\mathcal{F}^{-1}\{a\}$	Inverse Fourier transform of the function a : $\mathcal{F}^{-1}\{a\}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} a(k_x)e^{jk_x x} dk_x$.
$rect(x)$	Rectangular pulse $rect(x) = \begin{cases} 1 & \text{for } x \leq 1, \\ 0 & \text{for } x > 1. \end{cases}$
$sinc(x)$	The sinc function $sinc(x) = \frac{\sin x}{x}$.
$\delta(x)$	Dirac delta $\delta(x) = \begin{cases} \infty & \text{for } x = 0, \\ 0 & \text{for } x \neq 0 \end{cases}$, $\int \delta(x) dx = 1$.
\mathbb{Z}	The set of integer numbers $-\infty, \dots, -1, 0, 1, 2, 3, \dots, \infty$.
A^*	Complex-conjugate of a scalar A ; complex-conjugate transpose for a matrix A ; adjoint of an operator A , $A^* : \langle Af, g \rangle = \langle f, A^*g \rangle$.
$\langle f, g \rangle$	Inner product between functions f and g .
$\ f\ $	L_2 norm of f .
\mathbf{A}^\dagger	Pseudo-inverse of a matrix \mathbf{A} .

Chapter 1

Introduction

1.1 Background and motivation

This thesis is devoted to the generation and reconstruction of a desired, application-oriented, optical Light Field (LF) distribution within a spatial region of interest. Typically, the light field distribution is generated by modulating an incident wave through an optical system, which propagates beyond the system in a homogeneous medium within the region of interest. The design and configuration of the optical system is based on the knowledge of a physically consistent light field. Therefore, it is of major importance to reconstruct a light field which satisfies to an optimal extent the constraints imposed by the application at hand.

Light field reconstruction and synthesis finds applications in various social, industrial and scientific areas. Recently, much attention has been paid to applications related to *three-dimensional* (3D) *imaging* [1] and, more specifically, digital holography. Primary tasks are the numerical reconstruction of digitally recorded holograms of real 3D scenes and Computer Generated Holograms (CGH) of 3D synthetic objects and scenes [2]. Potential applications of 3D capture and display include education, medicine, the game industry, cultural heritage, Computer Aided Design (CAD), traffic control, military applications, etc. Education, training and simulation benefit from 3D display to achieve more realistic and detailed visualization. In medicine, high-quality 3D visualizations and telepresence make diagnosis and monitoring during treatment easier. Archaeological discoveries can be precisely captured and reconstructed for simulation in museums and on-line portals. CAD systems for different industrial purposes can visualize and describe the designed products and their interior without projections and cross-sections. The game and entertainment industry challenge the user with completely new freedom of movement and interactivity provided by the 3D visualization. Traffic control and the military industry benefit from 3D navigation, education and training.

Light field reconstruction and synthesis serve directly a wide area of applications in industrial and research optics called *light beam shaping*. They include lithography,

interferometry, microscopy and particle manipulation [3]. These applications specify the light beam shape and properties which have to remain unchanged for a long propagation distance or appear periodically. In lithography, the focus of the plotting beam must be kept for large axial extent. Beams with large depth of focus are convenient to serve as reference beams in interferometry. In microscopy, these beams can be used as scanning beams to obtain images with extended depth of focus. Dashed beams which contain peaks appearing periodically along their longitude are used in measurement and multi-focal imaging. Dark beams have intensity minima along the propagation direction which can be used for particle trapping and guiding.

Consider the example 3DTV display application represented in Fig.1.1. A 3D scene available in an abstract representation has to be re-created at the viewer's display end. The 3D scene specifies an artificial object or data of some real object, recorded by some means. A physically consistent light field is numerically reconstructed from the abstract scene representation by digital processing. According to the Rayleigh-Sommerfeld diffraction integral, the complex-valued light field distribution on a certain plane carries sufficient information for the field in the whole region of interest [4]. This information is used to configure an optical system which re-creates the scene by modulating an incident light wave. A modulated monochromatic incident wave is able to re-produce only a monochromatic scene. A full color scene can be achieved by modulating an incident wave for each color component in e.g. an RGB color system [1].

The optical modulation systems are primarily based on diffractive media which modulate either the phase or the amplitude of an incident beam. Early works optimize the structure of a binary printed transparency to reconstruct a planar pattern in the far field [2]. Current state-of-the-art Diffractive Optical Elements (DOEs) are able to span a large part of the space of physically consistent light fields [3]. They use refraction and their surface is fabricated by a lithographic process. Spatial Light Modulators (SLMs) are electronically reconfigurable and thus able to serve real-time applications [5, 6, 7]. However, they reproduce the degrees of freedom of a light wave to a smaller extent than fixed DOEs do [7]. SLMs are based on liquid crystals, micro-electromechanical systems and magneto- and acousto-optic modulators [5]. Digital Micromirror Devices (DMDs) are a type of reconfigurable devices which modulate an incident wave by reflection. These devices consist of a large array of electronically controlled micromirrors which can take one of a few preset angles [8]. They are very attractive as they provide higher contrast ratio, brightness and optical efficiency compared to liquid crystal -based SLMs. A pioneering work uses DMD to reconstruct a binary hologram [9]. A holographic 3D image projection by DMD in a gel medium is proposed in [10].

The light modulator configuration is optimized by means of computation to synthesize the light field specified on a plane. Many works specify the desired light field as planar scene, often as intensity only, and concentrate on the constraints imposed by the DOE/SLM [2, 11]. However, this assumption is not always reasonable, especially for beam shaping applications. Some of these works try to control the phase and amplitude at the target plane separately by adding another DOE/SLM [12, 13, 14]. The structure of the DOE is commonly optimized by iterative projection methods derived on the base of the Gerchberg-Saxton algorithm [15, 16, 17]. Recently, the

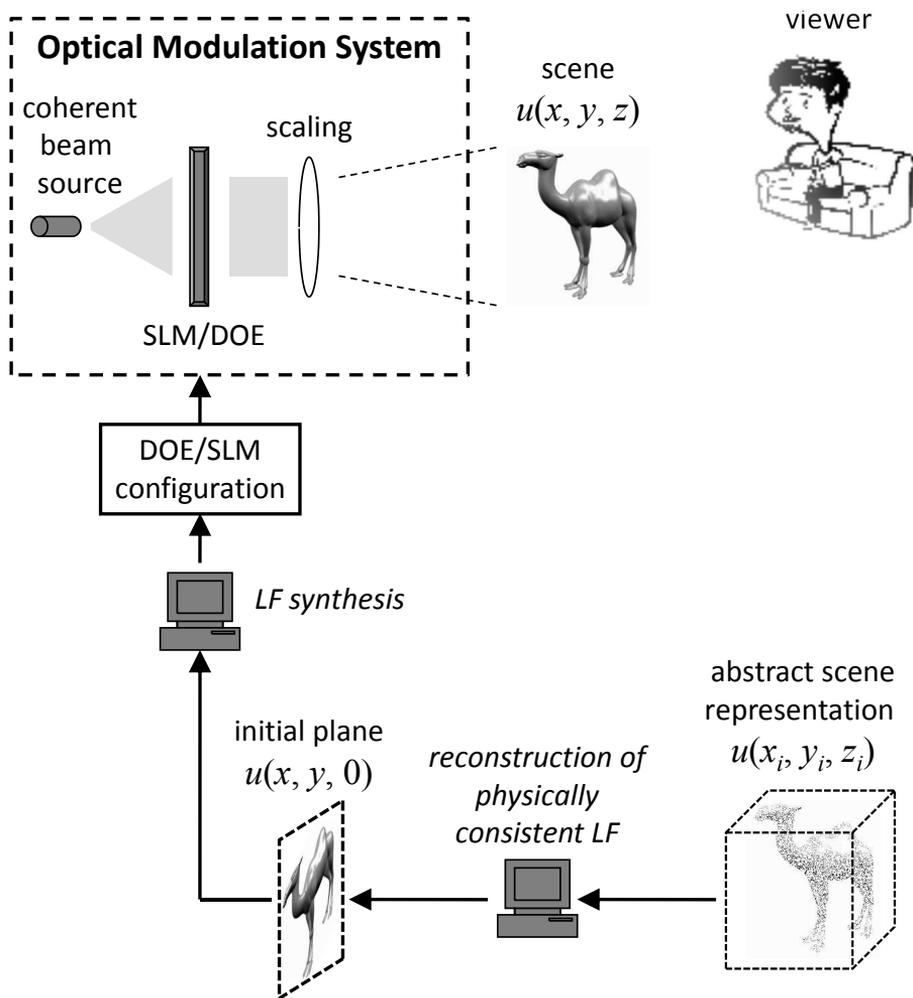


Figure 1.1: Simplified block diagram of a 3D scene display system from an abstract point cloud representation.

rapid growth of the computational resources has drawn interest towards optimization of the DOE structure to approximate directly a volumetric scene specification [3, 18, 19, 20, 21, 22, 23, 24]. A more general approach is to decouple the optimization of the DOE/SLM structure from the data specification. This involves a reconstruction of a physically consistent light field which best satisfies the volumetric specification. As a second stage, a reference plane for this field can be targeted by any type of fixed or reconfigurable modulation device which is appropriate for the application at hand. Therefore, light field reconstruction from the specifications provides a common approach to carry out the processing in any application.

Commonly, the volume of interest is specified on an ensemble of equally spaced planes, at a regular grid of points on each plane [3, 19, 20, 21, 22, 23]. Often, oversampling is used to cover the fine details and to enable efficient projections [3, 21, 23]. This results in large dimensionality, which in many cases is redundant. A minimum amount of samples per plane can be achieved by maximizing the sampling step separately for each plane, based on space and bandwidth assumptions and longitudinal distance [25, 26, 27, 28]. The distance between consecutive planes can be optimized, as well [23, 29]. This results in a fixed non-uniform, but regular 3D grid [23, 29]. However, such a 3D grid is still redundant with respect to the degrees of freedom of a light field, which can be specified only on a single plane. Moreover, some applications might provide fixed positions of known data points, generally inconvenient to be fit to either uniform or a fixed non-uniform grid. Computer graphics specify synthetic 3D scenes on a set of irregularly distributed points such as point clouds, 3D meshes, NURBS, etc [30]. These use higher point density to describe fine details and lower point density for uniform areas in the scene. Traditionally, 3D scenes are captured by recording the interference pattern of their light field with a reference beam [2]. Numerical or optical reconstruction of the scene light field requires the same reference beam, perfectly in-phase. Some recent works try to overcome the need of an additional beam by making multiple recordings of the same scene at different distances [31, 32, 33, 34, 35]. Light field reconstruction from irregularly distributed samples also covers the case of a sensing device having different orientation for each recording. Hence, data specification on irregular grid of points is both dimensionally efficient and application oriented.

To the best of our knowledge, the problem of irregular sampling and reconstruction of monochromatic light fields has not yet been addressed. These light fields are special type of signals where the signal along the propagation dimension depends upon the signal in the transversal dimension(s). However, the theory of sampling conventional signals has been well investigated and contains many results for one- and multidimensional signals [36, 37]. Many works investigated theoretical conditions on the recoverability of one dimensional band-limited signals [38, 39, 40, 41]. Strict conditions for multidimensional signals have not yet been found [37, 42, 43]. However, there is a myriad of methods to recover a signal from irregularly distributed samples. A large group of reconstruction algorithms are based on iterative projections [44, 45, 46, 47]. Many methods for light field reconstruction and synthesis from both volumetric and planar data use iterative projections, as well [3, 21, 22, 23, 32, 48, 49]. Therefore, adaptation of the basic method of Projection Onto Convex Sets (POCS) [16, 50] for light field reconstruction from irregularly distributed data points seems promising. Another mainstream of signal reconstruction algorithms from irregularly distributed samples

is based on finite dimensional modeling of the signal space and iterative matrix inversion, accelerated for structured matrices [37, 42, 43, 51]. Diffraction science offers few continuous models to describe monochromatic light fields [4, 52, 53]. Rayleigh-Sommerfeld diffraction integral and its frequency domain alternative – Plane-Wave Decomposition (PWD) integral provide exact scalar calculation [4, 54]. The Fresnel integral restricts the field of calculations only to small angles and the Fourier integral can be used for the far field [2, 4, 53]. Numerical computations are usually carried out between parallel planes and use discrete-to-discrete models, mostly derived on the base of the Fresnel integral. These use fixed uniform grids on the input and output planes whose steps are either the same or depend on the distance between the planes [2, 55]. However, such a grid does not fit the irregularly distributed sample positions. On the other hand, numerical computations require a description of the space of physically consistent fields with a finite amount of parameters. Hence, a discrete-to-continuous model of diffraction is required to address the irregular sampling and reconstruction problem. Possibly, such a model should not have restrictions so that the light field reconstruction provides exact results. Moreover, such a model should allow for reconstruction from physically inconsistent data samples, which are likely for 3D display and beam shaping applications [3].

1.2 Thesis objectives

The ultimate goal of this thesis is to develop a general and efficient framework which serves a wide range of light field synthesis applications. The generality is targeted by decoupling the light field synthesis from the prior light field reconstruction within a volumetric region of interest. The focus of the thesis is on numerical methods which serve the reconstruction and synthesis. The major part of the thesis investigates a dimensionally efficient reconstruction, while the synthesis concerns computational methods to optimize a certain type of modulation device. The objectives of the thesis are summarized as follows:

1. Develop efficient computational methods for light field reconstruction from data points, irregularly distributed within a region of interest
 - (a) Develop finite-dimensional, continuous expansion type of models for monochromatic light field distribution, based on exact scalar diffraction integrals
 - (b) Develop numerical reconstruction techniques which offer a trade-off between complexity, memory and accuracy
 - (c) Investigate the influence of the amount, density and distribution of the specified data points on the stability of the developed techniques
 - (d) Adapt the methods to serve for a wide range of input data scenarios, including physical inconsistency of the data values
2. Develop computational methods for light field synthesis with a DMD

1.3 Thesis outline

The thesis is organized as follows. Chapter 2 includes some needed preliminaries, ranging from diffraction science and sampling theory to the state-of-the-art of the considered topics. Section 2.1 presents a brief background of diffraction theory. It summarizes the characteristics which define considered light fields and discusses popular diffraction integrals with their inherent assumptions and limitations. Section 2.2 describes the available tools in sampling theory in relation to this thesis. These are mainly about irregular sampling, based on frame theory. Section 2.3 reviews the main results for the sampling of diffraction fields. It presents two common measures of the degrees of freedom of diffraction fields and compares them to the results available in the literature. The chapter concludes with a survey of the current trends in LF synthesis in Section 2.4. It also summarizes the main application areas as a motivation of the thesis. A survey of available LF synthesis methods is used in Section 3.1 to identify the gaps which are addressed by the contributions of the thesis. Section 3.1 also includes the generic hypotheses which are used to motivate the research. In addition, Chapter 3 summarizes the scientific content of the publications in this thesis. Section 3.2 presents the methods for reconstruction of a physically consistent light field from a set of data points, irregularly distributed within a volume of interest. Section 3.3 presents the preliminaries and optimization algorithms for light field synthesis with DMD presented in publication **P1**.

Chapter 2

Preliminaries

Light field reconstruction and synthesis problems, as considered in this thesis, form a highly interdisciplinary subject, which combines knowledge from diffraction science, sampling theory and theory of inverse problems. Therefore, it is essential to present these scientific fields with all relevant details to form basic knowledge and discuss open issues. In addition, the reader should be familiarized with the area of application of light field reconstruction and synthesis problems, where basic sciences meet to form the novel approaches described in this thesis. Hence, the specifics of this application area should be introduced together with a survey of the mainstream state-of-the-art approaches, presented in detail. This chapter includes the basics of the involved scientific disciplines together with the state-of-the-art, relevant to the considered problems. It starts with the basics of diffraction science and modern irregular sampling theory in Section 2.1 and Section 2.2, respectively. Section 2.3 surveys popular approaches to sampling of the diffraction fields as these are widely used in various light field reconstruction and synthesis problems. Finally, Section 2.4 introduces the area of light field reconstruction and synthesis problems and a survey of the state-of-the-art.

2.1 Scalar Diffraction Theory

The phenomenon known as *diffraction* is important in branches of physics and engineering which deal with wave propagation. The diffraction theory discussed in this section is applicable in a wide variety of fields, e.g. acoustics and radio transmission. However, here the focus is within the range of physical optics. Knowledge on diffraction and its inherent limitations is essential in order to fully understand the properties of three-dimensional imaging and optical data processing systems. Diffraction is normally taken to refer to various phenomena which occur when a wave encounters an obstacle. It is described as the apparent bending of waves around small obstacles and the spreading out of waves past small openings.

This section introduces the reader to the scalar light fields considered in the thesis. The presentation starts by reviewing the properties of such fields in the space and frequency domain. The section reviews the integral formulations which describe a scalar diffraction field. These are well known and mostly used for analytical purposes. However, these integrals are used as a base to derive many discrete diffraction models which form the core of the thesis. The Rayleigh-Sommerfeld (R-S) integral and its small-angle approximation are considered in both the space and frequency domain. The discussions of each space/frequency domain model concern the inherent assumptions made for its derivation, their scope of validity, and their advantages and drawbacks in terms of discretization and numerical computations.

2.1.1 Light waves in space and frequency domain

Light waves are electromagnetic waves which are most generally described by the Maxwell's wave equations as vector functions having three components – one for each spatial dimension x , y and z [4]. The consideration in this thesis is limited to the propagation of monochromatic light waves in media as free space and air. Such waves have only a single wavelength λ and the media are isotropic, homogeneous, non-dispersive and non-magnetic. In such a case, the waves are accurately described as the scalar wave functions of the type [4]:

$$U(x, y, z, t) = A(x, y, z) \cos(2\pi\nu t + \phi(x, y, z)) = \text{Re} \{ u(x, y, z) e^{-j2\pi\nu t} \}, \quad (2.1)$$

where t is the time, ν is the temporal frequency, A is the amplitude and ϕ is the phase of the monochromatic wave disturbance $U(x, y, z, t)$. The complex function $u(x, y, z) = A(x, y, z) e^{-j\phi(x, y, z)}$ depends on position only. This function $u(x, y, z)$ describes completely a monochromatic wave, because the time dependence $e^{-j2\pi\nu t}$ is purely deterministic. These considerations simplify Maxwell's equations to a single time-independent differential wave equation [4]:

$$\nabla^2 u(x, y, z) + k^2 u(x, y, z) = 0, \quad (2.2)$$

where the constant $k = 2\pi/\lambda$ is the wave number of the monochromatic field. This equation is known as the *Helmholtz* wave equation. In the rest of the thesis the time-dependent wave disturbance $u(x, y, z, t)$ is not considered any further, as the phasor $u(x, y, z)$ is sufficient to describe a scalar monochromatic diffraction field.

Certain properties of diffraction fields are more evident in the frequency domain and hence often used. Therefore, it is important to derive a frequency domain property, exhibited by any function which satisfies Helmholtz Eq. 2.2. The Fourier transform of a three dimensional scalar function $u(x, y, z)$ is defined as [56]:

$$a(k_x, k_y, k_z) = \iiint_{-\infty}^{\infty} u(x, y, z) e^{-j(k_x x + k_y y + k_z z)} dx dy dz, \quad (2.3)$$

where k_x , k_y and k_z denote the angular frequencies for x , y , and z coordinates, respectively. Applying this integral to Helmholtz wave Eq. 2.2, and using the derivative property of the Fourier transform [56], one can derive the relation:

$$a(k_x, k_y, k_z) (k^2 - k_x^2 - k_y^2 - k_z^2) = 0. \quad (2.4)$$

The term in the brackets equals zero for frequencies k_x , k_y and k_z which lie on the sphere with radius $k = 2\pi/\lambda$, centered at the origin. Eq. 2.4 is satisfied for frequencies outside this sphere only if the Fourier transform $a(k_x, k_y, k_z)$ is zero for these frequencies. Hence, the Fourier transform of any function satisfying the Helmholtz wave equation is non-zero only on the sphere

$$k_x^2 + k_y^2 + k_z^2 = k^2 = \left(\frac{2\pi}{\lambda}\right)^2. \quad (2.5)$$

This sphere is known as *Ewald's sphere* [21, 57]. Often, the considerations are simplified to only one transverse dimension x instead of the plane (x, y) . In this case a similar derivation leads to the conclusion that the two-dimensional (2D) Fourier transform $a(k_x, k_z)$ of the monochromatic scalar wave field $u(x, z)$ is restricted to a circle $k_x^2 + k_z^2 = k^2$ in the frequency plane.

2.1.2 Rayleigh-Sommerfeld diffraction integral

Many practical situations require computations of a scalar light field which satisfies the Helmholtz wave equation. As the differential Eq. 2.2 is not convenient, an equivalent integral equation is needed. A common derivation [4, 58] of such an integral starts by applying Green's identity [59] on two functions which satisfy Eq. 2.2 with a spherical wave as an auxiliary function. The derivation further takes into account the Sommerfeld radiation condition [4]. The result of the derivation is known as the Rayleigh-Sommerfeld diffraction integral:

$$\begin{aligned} u(x, y, z) &= \frac{1}{2\pi} \iint_{-\infty}^{\infty} u(x', y', 0) \frac{ze^{jkr}}{r^2} \left(\frac{1}{r} - jk\right) dx' dy', \\ r &= \sqrt{(x - x')^2 + (y - y')^2 + z^2}. \end{aligned} \quad (2.6)$$

This integral computes the light field $u(x, y, z)$ at any spatial point (x, y, z) from its values on an “*initial*” transversal plane $(x, y, z)|_{z=0}$, aligned to the origin $z = 0$. Assume a hypothetical screen at the plane $(x, y, z)|_{z=0}$, whose transparency follows the field values $u(x, y, 0)$, $x, y \in (-\infty; \infty)$ at this plane. Such a screen would reproduce $u(x, y, 0)$ at the plane $(x, y, z)|_{z=0}$ when illuminated by a normally oriented unit-amplitude and zero-phase plane wave. Thus, the result of Eq. 2.6 is somehow intuitive, since the screen is the obstacle which actually shapes the light from the source(s) to produce that field. In an abstract set-up, the result of Eq. 2.6 can be generalized for the field values from any transversal plane $(x, y, z)|_{z=z_T}$, $z_T \neq 0$. This plane is assumed as “*initial*” and, hence, the calculation of $u(x, y, z)$ should be aligned to this plane by taking the relative value for the longitudinal distance as $z - z_T$ instead of z in Eq. 2.6. Therefore, such an abstract set-up can be used to calculate diffraction immediately behind the screen. However, when a physically-existing screen is illuminated by a light source, the screen is never infinitely thin and there are always internal reflections along the rim of the aperture. This fact produces erroneous results when Eq. 2.6 is used to calculate diffraction within a few wavelengths behind the screen [60, 61, 62]. Another note is that Eq. 2.6 can be used

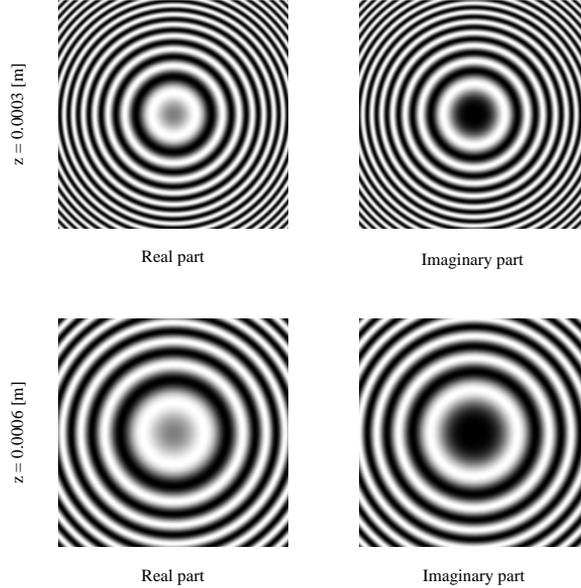


Figure 2.1: R-S impulse response: $h_z(x, y)$ shown on 256×256 grid inside a window of size $10^{-4} \times 10^{-4}$ m for two propagation distances: $z = 3.10^{-4}$ m (top row) and $z = 6.10^{-4}$ m (bottom row)

to calculate diffraction only in positive direction, i.e. for $z > 0$. Simple calculation shows that $u(x, y, -z) = -u(x, y, z)$ by Eq. 2.6, which contradicts the known physical phenomenon. In summary, knowledge of the field on a certain transversal plane $(x, y, z)|_{z=z_T}$ is sufficient to describe the field at any point $P = (x, y, z)$ in the positive three-dimensional half-space $(x, y, z)|_{z>z_T}$ behind that plane.

An important perspective of the integral in Eq. 2.6 is that it can be considered as a convolution integral. Indeed, it can be rewritten as

$$u(x, y, z) = \iint_{-\infty}^{\infty} u(x', y', 0) h_z(x - x', y - y') dx' dy' = (u(\cdot, \cdot, 0) * h_z)(x, y), \quad (2.7)$$

where the second term of the convolution is the function

$$h_z(x, y) = \frac{ze^{jk\sqrt{x^2+y^2+z^2}}}{2\pi(x^2 + y^2 + z^2)} \left(\frac{1}{\sqrt{x^2 + y^2 + z^2}} - jk \right). \quad (2.8)$$

Eq. 2.7 shows that the field $u(x, y, z)$ at any transversal plane $(x, y, z)|_{z=\text{const}}$ is calculated as a two-dimensional convolution between the field $u(x, y, 0)$ at the initial plane and the function $h_z(x, y)$ given in Eq. 2.8. Therefore, light propagation by z in the positive direction is represented as a linear shift-invariant system [63] with impulse response $h_z(x, y)$, which outputs the field $u(x, y, z)$ at the plane $(x, y, z)|_{z=\text{const}}$ when the input is the field $u(x, y, 0)$ from the initial plane $(x, y, z)|_{z=0}$.

The diffraction integral of Eq. 2.6 or Eq. 2.7 is used for many analytical purposes, but not so often for numerical computation of diffraction fields in practice. Numerical computations require discrete data of finite size and hence the integral must be discretized over the support of the integration. The support of the integration is determined by the support of the field at the initial plane $(x, y, z)|_{z=0}$ and the support of the target volume which is desired to be computed. Computation on a domain of infinite support is rarely of practical interest, but a finite domain might still require integration of infinite support because the exponential integrand (or the impulse response of Eq. 2.8) has infinite support. However, if the support of $u(x, y, 0)$ is assumed to be finite, the integration is required only within the sub-region of the plane $(x, y, z)|_{z=0}$ where $u(x, y, 0) \neq 0$. The discretization of the integral in Eq. 2.6 or Eq. 2.7 turns out to be the major problem when using the R-S integral for numerical computation. The straightforward approach takes a Riemann sum instead of the integral. This transforms Eq. 2.7 into a discrete convolution, which is correct if at least one of the integrands in the continuous convolution is band-limited [64]. According to Eq. 2.5, the Fourier spectrum of $u(x, y, z)$ is supported on a sphere of radius $\frac{2\pi}{\lambda}$, and hence $u(x, y, 0)$ is band-limited to $\frac{2\pi}{\lambda}$. However, the required sampling rate, governed by the Nyquist sampling theorem [64], is too high for practical purposes. The impulse response of Eq. 2.8 is not band-limited either. It is an oscillatory chirp-like function of infinite support, where the frequency of the oscillations increases together with the radial distance from the origin, and decreases for larger propagation distance z (cf. Fig. 2.1). Fortunately, for input and output planes of finite support, the convolution in Eq. 2.7 involves only a sub-range of the full support of $h_z(x, y)$, whose size equals the sum of the supports at the input and output planes [63]. Hence, the required portion of $h_z(x, y)$ does not contain oscillations at large radial distances, or equivalently, it is band-limited. Therefore, it can be sampled with respect to the involved bandwidth and the convolution of Eq. 2.7 can be discretized. Note that longer propagation distance along z decreases the bandwidth. This leads to a sub-region of the 3D half-space $z > 0$ where the discretized Eq. 2.7 produces correct results. The discrete convolution can already be computed with a fast $\mathcal{O}(N \log N)$ algorithm, based on the Fast Fourier Transform (FFT) [64]. This is the traditional approach, commonly used for calculation [65]. Shen and Wang [66] enhance this approach by introducing weights in the discrete convolution, whose values are found by the Simpson's rule for numerical integration [67]. Naskov and Logofătu [68] use different sampling intervals on the input and output planes in order to widen the spatial range of accuracy of the discrete convolution. As a result, they obtain a scaled convolution which is still computed fast in $\mathcal{O}(N \log N)$ calculations by the use of the Fractional Fourier Transform (FrFT) [69]. There are also some non-standard techniques which are able to compute the R-S diffraction integral in a much wider spatial range, usually with a limited precision and much slower computation [70, 71]. These use different approaches to approximate the oscillatory part of the R-S integral and describe it in terms of Fresnel integrals, which can be solved by a fast FFT-based approach [65]. D'Arcio et al. use polynomial and parabolic phasor to approximate the impulse response. The approximation turns out to be valid in a region which can be described as a union of small sub-regions with rectangular or parabolic shape [71]. Veerman et al. use different manipulations on the integrals to split them into a smooth and oscillatory part [70]. The smooth part is computed conventionally, while the oscillatory part is approximated by bicubic splines and subsequent Fresnel

integrals. This technique shows uniform computation error over a very wide range in space.

2.1.3 Plane-wave decomposition integral

The diffraction field at any point is described in terms of the 2D function $u(x, y, 0)$ in Eq. 2.6 and Eq. 2.7 from Subsection 2.1.2. Subsection 2.1.1 shows that the spatial frequencies k_x , k_y and k_z in the 3D Fourier transform of a diffraction field obey the relation in Eq. 2.5, which leaves two free coordinates out of all three. Indeed, the 3D diffraction field turns out to be described fully by a 2D function in both cases. The field has already the 2D spatial domain representation in Eq. 2.6 and Eq. 2.7. An alternative 2D representation in the frequency domain turns the convolution of Eq. 2.7 into multiplication, according to the multiplication property of the Fourier transform [56]. Such a relation might be beneficial from the computational and/or analytical point of view. However, the derivation of the Fourier transform of the impulse response $h_z(x, y)$ from Eq. 2.8 is not straightforward. One possible derivation starts by applying the Helmholtz Eq. 2.2 on the 2D inverse Fourier transform for the field distribution on any transversal plane, as follows:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2 \right) \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} a_z(k_x, k_y) e^{j(k_x x + k_y y)} dk_x dk_y = 0, \quad (2.9)$$

where $a_z(k_x, k_y)$ is the 2D Fourier transform of the field distribution at a transversal plane, standing at distance z from the origin. Putting the term in the brackets inside the integral and taking the derivation leads to

$$\frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \left(a_z(k_x, k_y) (k^2 - k_x^2 - k_y^2) + \frac{\partial^2 a_z(k_x, k_y)}{\partial z^2} \right) e^{j(k_x x + k_y y)} dk_x dk_y = 0. \quad (2.10)$$

As the integral is inverse Fourier transform, the term inside the brackets can be considered as the Fourier transform of some 2D function which equals zero. Hence, this term itself equals zero, yielding the homogeneous second order differential equation

$$\frac{\partial^2 a_z(k_x, k_y)}{\partial z^2} + a_z(k_x, k_y) (k^2 - k_x^2 - k_y^2) = 0. \quad (2.11)$$

This equation can be solved with respect to $a_z(k_x, k_y)$, if considered as an one-dimensional function of z . The initial condition can be taken as $a_z(k_x, k_y)|_{z=0} = a_0(k_x, k_y)$ since the goal of the whole derivation is to relate $a_z(k_x, k_y)$ to the Fourier transform $a_0(k_x, k_y)$ of the field $u(x, y, 0)$ on the initial plane $(x, y, z)|_{z=0}$ at $z = 0$. The solution of the differential Eq. 2.10 can be found as [72]:

$$a_z(k_x, k_y) = \begin{cases} a_0(k_x, k_y) e^{\pm jz \sqrt{k^2 - k_x^2 - k_y^2}} & \text{for } k_x^2 + k_y^2 < k^2; \\ a_0(k_x, k_y) e^{\mp z \sqrt{k_x^2 + k_y^2 - k^2}} & \text{for } k_x^2 + k_y^2 \geq k^2. \end{cases} \quad (2.12)$$

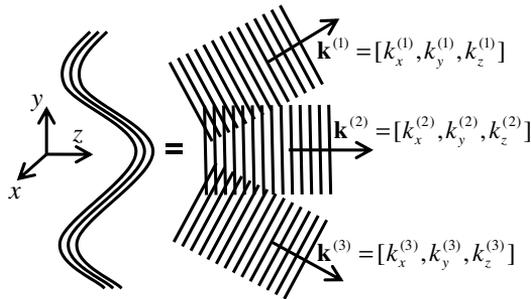


Figure 2.2: Plane-wave decomposition. A diffraction field is decomposed as a superposition of plane waves, each traveling in direction $\mathbf{k} = [k_x, k_y, k_z]$ determined by the spatial frequencies.

If only $a_z(k_x, k_y) = a_0(k_x, k_y)e^{jz\sqrt{k^2 - k_x^2 - k_y^2}}$ is considered as a solution, then the inverse Fourier transform for the field distribution on the plane $(x, y, z)|_{z=\text{const}}$ is

$$u(x, y, z) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} a_0(k_x, k_y) e^{jz\sqrt{k^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)} dk_x dk_y. \quad (2.13)$$

This integral is known as the *plane-wave decomposition* integral. It represents the 3D diffraction field $u(x, y, z)$ as a superposition of plane waves $e^{jz\sqrt{k^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)}$, each weighted by the complex amplitude $a_0(k_x, k_y)$ (cf. Fig. 2.2). Each plane wave has the frequency $k = \frac{2\pi}{\lambda}$ and closes angles θ_x , θ_y and θ_z with the x , y and z axes, respectively. The angles θ_x , θ_y and θ_z are determined from the frequencies k_x , k_y and $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$ (cf. Fig. 2.3). The angle θ_z ranges within $[-\frac{\pi}{2}; \frac{\pi}{2}]$ because $\cos \theta_z = \frac{k_z}{k}$ and $k_z = \sqrt{k^2 - k_x^2 - k_y^2} > 0$ for the considered solution of the differential Eq. 2.11. Therefore, the plane waves $e^{jz\sqrt{k^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)}$ travel in a positive direction along z . If $a_z(k_x, k_y) = a_0(k_x, k_y)e^{-jz\sqrt{k^2 - k_x^2 - k_y^2}}$ is considered as the solution of the differential Eq. 2.11, then $u(x, y, z)$ is decomposed in terms of the plane waves $e^{-jz\sqrt{k^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)}$. Their angle θ_z ranges within $[\frac{\pi}{2}; \frac{3\pi}{2}]$ and their direction is negative with respect to the z axis. However, this is in contradiction to the Sommerfeld radiation condition, which ensures only outgoing waves from a planar diffraction screen (the initial plane $(x, y, z)|_{z=0}$ in this case). Therefore, only waves $e^{jz\sqrt{k^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)}$ with positive direction are included in the PWD integral of Eq. 2.13.

The plane-wave decomposition integral implicitly includes the solutions of Eq. 2.11 for frequencies which satisfy $k_x^2 + k_y^2 \geq k^2$ (cf. Eq. 2.12). Note that these can be obtained from the solutions for $k_x^2 + k_y^2 < k^2$ when the square root of the negative value $k^2 - k_x^2 - k_y^2$ is written as $\sqrt{k^2 - k_x^2 - k_y^2} = j\sqrt{k_x^2 + k_y^2 - k^2}$. The respective decomposition waves in Eq. 2.13 are then $e^{\mp z\sqrt{k_x^2 + k_y^2 - k^2}} e^{j(k_x x + k_y y)}$, with a minus sign for waves traveling in positive z direction and plus sign for waves traveling in

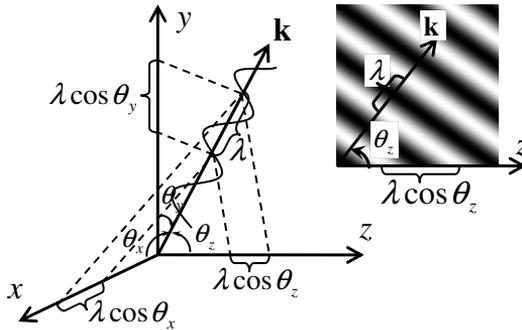


Figure 2.3: Plane wave and directional angles. A plane wave has a fundamental period of λ in its propagation direction $\mathbf{k} = [k_x, k_y, k_z]$. Along the coordinate axes x , y and z the plane wave is harmonic with periods $\lambda \cos \theta_x$, $\lambda \cos \theta_y$ and $\lambda \cos \theta_z$. Hence the frequencies are $k_x = \frac{2\pi}{\lambda \cos \theta_x}$, $k_y = \frac{2\pi}{\lambda \cos \theta_y}$ and $k_z = \frac{2\pi}{\lambda \cos \theta_z}$.

negative z direction. These waves are so called *evanescent waves* [4]. The argument of the quantity $e^{\mp z \sqrt{k_x^2 - k_y^2 - k^2}}$ is always negative, as a minus sign is taken for positive z direction, and the plus sign – for negative z direction. Therefore, the evanescent waves drop off after a few wavelengths of propagation and carry no energy. This is why some authors omit them from consideration, taking the integration in Eq. 2.13 within the disc $k_x^2 + k_y^2 < k^2$. The evanescent waves are not negligible in the vicinity of the initial plane $(x, y, z)|_{z=0}$, but they are much more accurately described by vector theory through Maxwell equations [4].

Eq. 2.13 already provides the desired 2D frequency domain representation of a diffraction field. Moreover, Eq. 2.13 provides representation of finite support for fields whose distribution $u(x, y, 0)$ on the initial plane is band-limited. This serves as an alternative to Eq. 2.6, which has finite support description for fields originating from spatially limited functions $u(x, y, 0)$. Hence, Eq. 2.13 is preferred for band-limited functions $u(x, y, 0)$ and Eq. 2.6 preferred for spatially limited functions $u(x, y, 0)$. Eq. 2.13 also provides a finite support representation for fields with spatially limited $u(x, y, 0)$ for propagation distance of more than a few wavelengths. Even though $a_0(k_x, k_y)$ is of infinite support in this case, the evanescent waves suppress its components outside the disc $k_x^2 + k_y^2 < k^2$. However, such support is often too wide to be used for calculations.

The solution of Eq. 2.11 which corresponds to plane waves propagating in the forward direction can be rewritten as

$$a_z(k_x, k_y) = a_0(k_x, k_y) \cdot H_z(k_x, k_y) \text{ with } H_z(k_x, k_y) = e^{jz \sqrt{k^2 - k_x^2 - k_y^2}}. \quad (2.14)$$

This equation suggests that Eq. 2.14 and Eq. 2.7 are Fourier transforms of each other when $z = z_T$. Hence, the propagation process from the initial plane of the field to another transversal plane at a distance z can be modeled as a shift-invariant system with impulse response $h_z(x, y)$ as in Eq. 2.8 and transfer function $H_z(k_x, k_y)$ as in Eq. 2.14, where $h_z(x, y)$ and $H_z(k_x, k_y)$ form a Fourier transform pair. Moreover, the relation in Eq. 2.14 implies that such a propagation model can be used to calculate

diffraction between any parallel diffraction planes, say at distances z_1 and z_2 ($z_2 > z_1$), because:

$$\begin{aligned} H_{z_2}(k_x, k_y) &= e^{jz_2\sqrt{k^2-k_x^2-k_y^2}} = e^{j(z_2-z_1)\sqrt{k^2-k_x^2-k_y^2}} \cdot e^{jz_1\sqrt{k^2-k_x^2-k_y^2}} \\ &= H_{z_2-z_1}(k_x, k_y) \cdot H_{z_1}(k_x, k_y). \end{aligned} \quad (2.15)$$

This relation corresponds to a serial connection of two shift invariant systems. Hence, the R-S integral can be iterated as well:

$$h_{z_2}(x, y) = h_{z_2-z_1} * h_{z_1}(x, y). \quad (2.16)$$

Eq. 2.15 gives a very useful hint: $a_{z_1}(k_x, k_y)$ can be calculated from $a_{z_2}(k_x, k_y)$ by the inverse transfer function $H_{z_2}(k_x, k_y)/H_{z_2-z_1}(k_x, k_y) = H_{z_1-z_2}(k_x, k_y)$. In other words, Eq. 2.14 can be used to calculate backward propagation. In the general case of backward propagation at distance z , the transfer function is $H_{-z}(k_x, k_y) = e^{-jz\sqrt{k^2-k_x^2-k_y^2}}$. This corresponds to plane waves traveling in a backward direction in terms of the PWD integral. However, back-propagation calculations must be made with care. If the initial function $u(x, y, 0)$ is not band-limited, then the transfer function $H_z(k_x, k_y)$ multiplies the components of $a_0(k_x, k_y)$ for frequencies lying outside the disc $k_x^2 + k_y^2 < k^2$ with the decaying factor $e^{-z\sqrt{k_x^2+k_y^2-k^2}}$. The exact inverse $H_{-z}(k_x, k_y)$ reconstructs these frequency components by multiplication with $e^{z\sqrt{k_x^2+k_y^2-k^2}}$, so theoretically it fully recovers $a_0(k_x, k_y)$ from $a_z(k_x, k_y)$. However, information about the high-frequency components of $a_0(k_x, k_y)$ may be lost in numerical computations if the respective components of $a_z(k_x, k_y)$ fall below the precision limit. Hence, $a_0(k_x, k_y, 0)$ can be recovered safely from $a_z(k_x, k_y)$ by back-propagation with $H_{-z}(k_x, k_y)$ only if its spectrum $a_0(k_x, k_y)$ is supported within the disc $k_x^2 + k_y^2 < k^2$. However, back-propagation can be used to calculate diffraction from the plane at z_2 to the plane at z_1 when $z_2 > z_1 \gg \lambda$, because the high frequency components of $a_{z_2}(k_x, k_y)$ and $a_{z_1}(k_x, k_y)$ are already suppressed.

It should be noted that the derived PWD integral in Eq. 2.13 is valid under certain conditions, mainly related to the existence of the Fourier transform of the field distribution along a transversal line [54]. As a starting point, the derivation assumes that the field distribution $u(x, y, z)$ at the plane $(x, y, z)_{z=\text{const}}$ has a 2D Fourier transform $a_z(k_x, k_y)$ which exists and is invertible. Another assumption is the existence and invertibility of the initial condition $a_0(k_x, k_y)$ for the differential Eq. 2.11. Hence, both $u(x, y, 0)$ and $a_0(k_x, k_y)$ as 2D functions must obey the following properties [56]: (1) they should be piecewise continuous; and (2) they and their first partial derivatives should decay with the same order of magnitude as $\frac{1}{R}$ outside some disc of radius R_0 ($R = \sqrt{x^2 + y^2}$ for $u(x, y, 0)$ and $R = \sqrt{k_x^2 + k_y^2}$ for $a_0(k_x, k_y)$). $H_z(k_x, k_y)$ and $h_z(x, y)$ are associated as a Fourier transform pair throughout the derivation, and it is straightforward to show that they obey the same properties. Therefore, $a_z(k_x, k_y)$ obeys the same properties, since it is a product of $a_0(k_x, k_y)$ and $H_z(k_x, k_y)$. Hence, the inverse Fourier transform of $a_z(k_x, k_y)$ exists and the same conditions are valid for $u(x, y, z)$ as a 2D function. In addition, $u(x, y, z)$ must obey the conditions as a 3D function and the Sommerfeld radiation condition.

Finally, the issues related to eventual numerical computations through the PWD integral should be discussed. The simple Fourier domain product form of Eq. 2.14 suggests

very efficient numerical implementation through the FFT algorithm [66, 73, 74, 75]. It goes by taking the Discrete Fourier Transform (DFT) of the sampled initial plane $u(x, y, 0)$, multiplication by the sampled $H_z(k_x, k_y)$ and an inverse DFT. Hence, this algorithm is beneficial for band-limited functions $u(x, y, 0)$, whose Fourier transform $a_0(k_x, k_y)$ contains a finite number of coefficients when sampled. Sampling $a_0(k_x, k_y)$ corresponds to the periodization of $u(x, y, 0)$ with a period which is inversely proportional to the sampling step. The periodic replicas $u(x, y, 0)$ overlap, because a band-limited function never has a finite spatial support. If, instead, $u(x, y, 0)$ is considered to be essentially space- and band-limited, the overlap between the adjacent replicas of $u(x, y, 0)$ can be minimized by properly decreasing the size of the sampling step for $a_0(k_x, k_y)$. Note that a denser sampling of $a_0(k_x, k_y)$ increases the dimensionality of computations as the essential support of $a_0(k_x, k_y)$ is covered by a larger number of essentially non-zero coefficients. Free-space propagation along z tends to spread the support of the spatially concentrated $u(x, y, 0)$. Therefore, the overlap between the adjacent periodic replicas of $u(x, y, z)$ becomes already significant after some distance z . This means that the maximal computation distance z is limited by the sampling density for $a_0(k_x, k_y)$. If needed, the maximal distance can be increased by increasing the sampling density and, respectively, the dimensionality of computation.

2.1.4 Fresnel approximation

The R-S diffraction integral of Eq. 2.6 can be simplified by taking into account some approximations. Often the diffraction radius $r = z\sqrt{1 + \frac{(x'-x)^2}{z^2} + \frac{(y'-y)^2}{z^2}}$ is a few orders of magnitude higher than the wavelength λ , and then $\frac{1}{r} - jk \approx -jk$. As a next step, r is approximated using the binomial expansion $\sqrt{1+b} = 1 + \frac{b}{2} - \frac{b^2}{8} + \dots$ with $b = \frac{(x'-x)^2}{z^2} + \frac{(y'-y)^2}{z^2}$. The term r^2 in the denominator of Eq. 2.6 is approximated only with the first term of the expansion $r^2 \approx z$. However, the r appearing in the exponent is much more sensitive to errors, because it is multiplied by the very large number k and the value of the exponent might change significantly when its argument changes even by a fraction of a radian. If the first two terms in the binomial expansion remain, the error is dominated by the third term ($b^2/8$) of the binomial expansion. Hence, one needs $\frac{k}{8z^3} [(x-x')^2 + (y-y')^2]_{max}^2 \ll 1$ or $z^3 \gg \frac{\pi}{4\lambda} [(x-x')^2 + (y-y')^2]_{max}^2$. This is the case when the calculated diffraction involves only small angles. After these approximations, the R-S integral of Eq. 2.6 transforms to

$$u(x, y, z) = \iint_{-\infty}^{\infty} u(x', y', 0) \frac{e^{jkz}}{j\lambda z} e^{\frac{jk}{2z}((x-x')^2 + (y-y')^2)} dx' dy'. \quad (2.17)$$

This integral is known as the *Fresnel approximation* to the R-S diffraction integral. Note that the diffraction between parallel planes can still be embodied in a linear shift-invariant system, similar to that of Eq. 2.7 and Eq. 2.8. The impulse response in this case is

$$h_z(x, y) = \frac{e^{jkz}}{j\lambda z} e^{\frac{jk}{2z}(x^2 + y^2)}. \quad (2.18)$$

Its Fourier transform can be directly calculated as [4]:

$$H_z(k_x, k_y) = \mathcal{F}\{h_z(x, y)\} = e^{jkz} e^{-\frac{jz}{2k}(k_x^2 + k_y^2)} = e^{jkz} \left(1 - \frac{k_x^2}{2k^2} + \frac{k_y^2}{2k^2}\right). \quad (2.19)$$

The convolution $u(x, y, z) = (u(\cdot, \cdot, 0) * h_z)(x, y)$ transforms to the product $a_z(k_x, k_y) = a_0(k_x, k_y)H_z(k_x, k_y)$ in the Fourier domain, which suggests a practical computational algorithm for the Fresnel integral:

$$u(x, y, z) = \mathcal{F}^{-1}\{\mathcal{F}\{u(x, y, 0)\} H_z(k_x, k_y)\}. \quad (2.20)$$

The forward and inverse Fourier transforms are implemented by the FFT algorithm, and hence the diffraction is computed in logarithmic complexity. Note that such an implementation is really similar to the computational algorithm based on the PWD integral, as described in Subsection 2.1.3.

The implementation from Eq. 2.20 shows very similar limitations to the algorithm, based on the PWD integral. Still, the Fresnel $H_z(k_x, k_y) = e^{jkz} \left(1 - \frac{k_x^2}{2k^2} + \frac{k_y^2}{2k^2}\right)$ can be considered just an approximation to the PWD $H_z(k_x, k_y) = e^{jz\sqrt{k^2 - k_x^2 - k_y^2}}$. In fact, this approximation corresponds to $k\sqrt{1 - \frac{k_x^2}{k^2} - \frac{k_y^2}{k^2}} \approx k\left(1 - \frac{k_x^2}{2k^2} - \frac{k_y^2}{2k^2}\right)$, where only the first two terms in the binomial expansion of $k\sqrt{1 - \frac{k_x^2}{k^2} - \frac{k_y^2}{k^2}}$ are retained. The error is negligible when $\frac{k_x^2}{k^2} = \cos^2\theta_x \ll 1$ and $\frac{k_y^2}{k^2} = \cos^2\theta_y \ll 1$, where θ_x and θ_y are the angles of propagation of the respective plane wave in the plane-wave decomposition integral in Eq. 2.13. Hence, the Fresnel approximation, done either in the spatial or frequency domain, is valid when the calculated diffraction involves small angles. This implies that the high-frequency details of the field at the output plane are lost. Hence, only a small portion of the field bandwidth needs to be covered, for which only a small number of discretized $H_z(k_x, k_y)$ is sufficient.

In addition, the implementation of the Fresnel integral in Eq. 2.20 requires the same sampling step at the input and output planes, because their Fourier transforms use the same computational grid. In many cases different sampling steps are convenient, because the support of the input and output apertures is different and the size of computations (amount of non-zero samples) must be the same for any FFT-based approach. One way to achieve different sampling steps goes through rewriting the Fresnel integral of Eq. 2.17 in the form of a single-step Fourier transform. At first, the exponent under the integral is decomposed as a product of exponents, after opening the squared brackets in the argument:

$$u(x, y, z) = \frac{e^{jkz}}{j\lambda z} e^{\frac{j\pi}{\lambda z}(x^2 + y^2)} \iint_{-\infty}^{\infty} u(x', y', 0) e^{\frac{j\pi}{\lambda z}(x'^2 + y'^2)} e^{-j\frac{2\pi}{\lambda z}(x'x + y'y)} dx' dy', \quad (2.21)$$

where $k = \frac{2\pi}{\lambda}$ is used. The notations $\omega_x = \frac{2\pi x}{\lambda z}$ and $\omega_y = \frac{2\pi y}{\lambda z}$ help to obtain the Fourier transform-like integral:

$$u(\omega_x, \omega_y, z) = \frac{e^{jkz}}{j\lambda z} e^{\frac{j\lambda z}{4\pi}(\omega_x^2 + \omega_y^2)} \iint_{-\infty}^{\infty} u(x', y', 0) e^{\frac{j\pi}{\lambda z}(x'^2 + y'^2)} e^{-j(\omega_x x' + \omega_y y')} dx' dy'. \quad (2.22)$$

The integral can be seen as the Fourier transform of the function $u(x', y', 0)e^{\frac{j\pi}{\lambda z}(x'^2+y'^2)}$, where the angular frequencies are ω_x and ω_y . Eq. 2.22 suggests that the Fresnel diffraction can be implemented by the algorithm:

$$u(\omega_x, \omega_y, z) = \frac{e^{jkz}}{j\lambda z} e^{\frac{j\lambda z}{4\pi}(\omega_x^2 + \omega_y^2)} \mathcal{F} \left\{ u(x', y', 0) \cdot e^{\frac{j\pi}{\lambda z}(x'^2+y'^2)} \right\}. \quad (2.23)$$

The Fourier transform can be calculated by an N -point FFT algorithm which computes N coefficients in the frequency domain from N samples of an input. The spacing steps \mathcal{W}_x and \mathcal{W}_y between the coefficients in the frequency domain are related to the sampling steps X' and Y' by the relations $X' = \frac{2\pi}{N\mathcal{W}_x}$ and $Y' = \frac{2\pi}{N\mathcal{W}_y}$ [56, 63, 64]. Expressed for the sampling steps X and Y on the output plane, these relations are:

$$X' = \frac{\lambda z}{NX} \quad \text{and} \quad Y' = \frac{\lambda z}{NY}. \quad (2.24)$$

In practice, either X' and Y' at the input plane or X and Y at the output plane are often fixed, for example, by the physical restrictions of some sensor or by the Nyquist criterion for the function on the respective plane [63, 64]. Eq. 2.24 implies that the free sampling step is determined based on the propagation distance z and wavelength λ . However, it can be controlled additionally by the computation size N [55].

2.1.5 Conclusions

This section has presented information about scalar diffraction fields, related to the thesis and the attached publications. Subsection 2.1.1 has set the range of considered light fields in this thesis – the ones which satisfy the Helmholtz Eq. 2.2. An important feature of these light fields is that their spectrum is nonzero only on Ewald's sphere as defined in Eq. 2.5. This fact is used in publications **P5**, **P7** and **P8**. Subsection 2.1.2 has described the most accurate scalar model for diffraction – the R-S integral together with the limits of its practical use for computations. The frequency domain alternative of R-S integral – the PWD integral – has been presented in Subsection 2.1.3. It is used in most of the publications – **P2**, **P3**, **P4**, **P6**, **P7** and **P8**. The derivation of the PWD integral in Eq. 2.9 - Eq. 2.13 is used to explain and underline the connection between the Helmholtz equation and the plane waves constituting the decomposition in the PWD integral. The subsequent discussion has explained the relation between spatial frequencies, direction and directional cosines of the plane wave as a basis function in the PWD integral. The limits of validity of the PWD integral and practical computation issues have been discussed as well. Subsection 2.1.4 has presented the Fresnel approximation of the R-S integral in both the spatial and frequency domain. The presentation included conditions for validity and implementation algorithms related to discrete computations. These are relevant to the literature, surveyed in Chapter 3. Moreover, the few existing results on sampling of the diffraction field, discussed in Section 2.3, have been derived for the case of Fresnel diffraction.

2.2 Sampling

Sampling refers to a general mechanism for converting a continuous function into a sequence of numbers in a manner which can be stably inverted. In the context of this thesis, such conversion is beneficial for the convenience of computations and processing related to various optical applications. In comparison, analog computations performed with optical and electronic systems are faster than digital algorithms implemented on microprocessors, but less flexible.

The purpose of this section is to introduce the reader to modern sampling theory in terms of functional spaces using Riesz bases and frames. The section starts with a brief review of the classical uniform sampling of band-limited functions. The main results from frame theory are described as a background for the necessary and/or sufficient density conditions for reconstruction of a function from a set of irregularly distributed samples. In this context, the most common and widely used reconstruction approaches are described together with the factors influencing their performance.

2.2.1 Classical sampling theory

Classical sampling theory investigates the case when a continuous function $f(x)$ belongs to the space of band-limited functions [76]:

$$B_\Omega = \{f : \|f\| < \infty \text{ and } \mathcal{F}\{f(x)\} = 0 \text{ for } k_x \notin [-\Omega, \Omega]\}. \quad (2.25)$$

A well known result states that such a function can be represented by the sequence of its sample values $\{f(nX)\}_{n \in \mathbb{Z}}$, taken at uniformly distributed locations $x = nX$ in the spatial domain, provided the relation:

$$X \leq \frac{\pi}{\Omega}. \quad (2.26)$$

This relation is often referred to as the Nyquist sampling theorem [77], even though the result was shown earlier by Whittaker [78]. Later Kotelnikov [79] showed the same criterion in the Soviet literature and Shannon generalized the results [76]. The set of functions $\left\{\frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX))\right\}_{n \in \mathbb{Z}}$ forms an orthogonal set for the space B_Ω and, therefore, spans it. The decomposition of every function $f(x) \in B_\Omega$ over this set consists of the samples:

$$\begin{aligned} f(nX) &= \left\langle f(x), \frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX)) \right\rangle = \int_{-\infty}^{\infty} f(x) \frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX)) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k_x) \text{rect}\left(\frac{k_x}{\Omega}\right) e^{jk_x nX} dk_x = \mathcal{F}^{-1}\{F(k_x)\}|_{x=nX}, \end{aligned} \quad (2.27)$$

for all $n \in \mathbb{Z}$. The function $f(x)$ can be recovered from its samples by the following reconstruction formula:

$$\begin{aligned} \hat{f}(x) &= \frac{\Omega}{\pi} \sum_{n=-\infty}^{\infty} f(nX) \text{sinc}(\Omega(x - nX)) = \frac{\Omega}{\pi} \text{sinc}(\Omega x) * \sum_{n=-\infty}^{\infty} f(nX) \delta(x - nX) \\ &= \mathcal{F}^{-1}\left\{\text{rect}\left(\frac{k_x}{\Omega}\right) \sum_{m=-\infty}^{\infty} F\left(k_x - m \frac{2\pi}{X}\right)\right\}, \end{aligned} \quad (2.28)$$

where $F(k_x) = \mathcal{F}\{f(x)\}$. Eq. 2.28 is, in fact, a low-pass filtering operation with the ideal reconstruction filter with impulse response $h_{id}(x) = \frac{\Omega}{\pi} \text{sinc}(\Omega x)$ and frequency response $H_{id}(k_x) = \text{rect}\left(\frac{k_x}{\Omega}\right)$. The set $\left\{\frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX))\right\}_{n \in \mathbb{Z}}$ can be regarded as both sampling and reconstruction set in the case when the samples are taken at an interval $X = \frac{\pi}{\Omega}$. However, this set is not anymore orthogonal when the sampling interval is smaller $X < \frac{\pi}{\Omega}$ even though it still spans B_Ω . In this case, one can use instead the orthogonal set $\left\{\frac{1}{X} \text{sinc}\left(\frac{\pi}{X}(x - nX)\right)\right\}_{n \in \mathbb{Z}}$ for the reconstruction in Eq. 2.28, if it is performed in the spatial domain. Eq. 2.27 also suggests that the set of functions $\left\{\text{rect}\left(\frac{k_x}{\Omega}\right) e^{jk_x n \frac{\pi}{\Omega}}\right\}_{n \in \mathbb{Z}}$ forms a basis for the space of finite support functions $L^2_{[-\Omega, \Omega]} = \{F : \|F(k_x)\| < \infty, F(k_x) = 0 \text{ for } k_x \notin [-\Omega, \Omega]\}$, in the same way as the set $\left\{\frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX))\right\}_{n \in \mathbb{Z}}$ forms a basis for B_Ω .

The notation $\hat{f}(x)$ in Eq. 2.28 is used to underline the fact that the reconstruction is exact only when $F(k_x)$ is supported on $[-\Omega, \Omega]$ and the sampling interval is chosen as $X \leq \frac{\pi}{\Omega}$. The reconstruction $\hat{f}(x)$ is still a good approximation of $f(x)$ when $f(x)$ is essentially band-limited, i.e. when most of the energy of $F(k_x)$ is still supported within $[-\Omega, \Omega]$. The approximation $\hat{f}(x)$ is rather crude for non band-limited functions and/or larger sampling intervals.

2.2.2 Frame theory

The theory described in this subsection serves as a background to describe the theory of irregular sampling and reconstruction of a function. These are results related to the stability of a sampling set, its density and various reconstruction algorithms with the factors influencing their convergence. The consideration is limited to concepts which either form the basics of the theory or are relevant to the results of this thesis.

2.2.2.1 Frame expansion

Sampling theory often considers that the functions to be sampled belong to a Hilbert space. Hilbert spaces are derived from Banach spaces, where the norm metric on a pair of functions from the space is associated with the inner product between the functions [80]. A *Riesz basis* [80, 81, 82, 83] for a Hilbert space \mathbb{H} is a set of linearly independent elements $\{\varphi_n\}$ which span \mathbb{H} , i.e. any element f in the space \mathbb{H} can be expressed as a linear combination of φ_n of the form

$$f = \sum_n c_n \varphi_n. \quad (2.29)$$

The notion of bases in finite-dimensional spaces implies that the number of representative elements is the same as the dimension of the space. Moreover, the linear independence of the set implies that the set of expansion coefficients $\{c_n\}$ is unique for each element in the space \mathbb{H} . *Frames* are generalizations of Riesz bases, where the elements in the representative set $\{\varphi_n\}$ are no more linearly independent, but have some redundancy. Hence, the representation of any space element in Eq. 2.29 is no

longer unique, which opens the opportunity to choose the best representation for the problem at hand.

Frames can formally be defined in the following way: a sequence $\{\varphi_n\}$ in a Hilbert space \mathbb{H} is a frame if there are two constants $0 < A \leq B < \infty$, such that for all f in \mathbb{H} ,

$$A\|f\|^2 \leq \sum_n |\langle \varphi_n, f \rangle|^2 \leq B\|f\|^2. \quad (2.30)$$

The inequalities in Eq. 2.30 are called *frame conditions*, where A and B are *frame bounds*. Each member of the sequence $\{\varphi_n\}$ is referred to as a *frame element*. The frame bounds A and B show how much the energy of the function f might differ from the energy of its decomposition over the frame $\{\varphi_n\}_n$. The energy of f and the energy of its decomposition would be measured by the same physical units in an eventual application and, hence, the frame bounds A and B are dimensionless. Later on in this subsection, it becomes evident that the numerical reconstruction of f from its decomposition is more stable when A and B are closer to each other. In this context, an orthogonal basis has equal frame bounds $A = B$ and is also called a *tight* frame. A frame is called *exact* if, upon removing one of its elements, it is no longer a frame. Exact frames are equivalent to Riesz bases. For example, an *orthonormal basis* is an exact tight frame with frame bounds $A = B = 1$.

The coefficients c_n in the linear expansion of Eq. 2.29 are found by inner products of f with the elements from another frame set $\{\tilde{\varphi}_n\}$, called *dual frame* of $\{\varphi_n\}$, as $c_n = \langle f, \tilde{\varphi}_n \rangle$. The dual frame also spans the space \mathbb{H} and the linear expansion can be written for each frame:

$$f = \sum_n \langle f, \tilde{\varphi}_n \rangle \varphi_n = \sum_n \langle f, \varphi_n \rangle \tilde{\varphi}_n. \quad (2.31)$$

The same expansion is also valid for Riesz bases. The only difference between frames and Riesz bases is the uniqueness of the representation. In this sense, each Riesz basis has a unique dual, while a frame has a different dual set for each possible expansion. However, one of the representations minimizes the norm of the frame expansion coefficients and has the form [80, 81]:

$$f = \sum_n \langle f, S^{-1}\varphi_n \rangle \varphi_n, \quad (2.32)$$

where $\tilde{\varphi}_n = S^{-1}\varphi_n$ is the *canonical dual frame*. S is the *frame operator* defined as

$$Sf = \sum_n \langle f, \varphi_n \rangle \varphi_n. \quad (2.33)$$

It is easy to verify that this operator is positive and self-adjoint. Moreover, its largest and smallest eigenvalues are equal to the frame bounds B and A , respectively [80]. In this sense the frame bounds are related to the numerical stability of computing the inverse of the frame operator, which is needed to find the decomposition coefficients. Note that the canonical dual frame $\{\tilde{\varphi}_n\}$ has S^{-1} for a frame operator:

$$\tilde{S}f = \sum_n \langle f, S^{-1}\varphi_n \rangle S^{-1}\varphi_n = S^{-1} \sum_n \langle S^{-1}f, \varphi_n \rangle \varphi_n = S^{-1}SS^{-1}f = S^{-1}f. \quad (2.34)$$

Hence, the largest and smallest eigenvalues of \tilde{S} are A^{-1} and B^{-1} , respectively. Consequently, the lower and the upper bounds of the canonical dual $\{\tilde{\varphi}_n\}$ are B^{-1} and A^{-1} , respectively.

The representations in Eq. 2.31 - Eq. 2.33 have a convenient matrix notation, used in practice in frame reconstruction algorithms. They are used to derive some of the algorithms for light field reconstruction from a set of irregularly distributed data points. Let Φ and $\tilde{\Phi}$ be matrices whose columns are the frame elements and the dual frame elements, respectively:

$$\Phi = [\varphi_1, \varphi_2, \dots, \varphi_n, \dots] \quad \text{and} \quad \tilde{\Phi} = [\tilde{\varphi}_1, \tilde{\varphi}_2, \dots, \tilde{\varphi}_n, \dots]. \quad (2.35)$$

As the frame might contain either a finite or infinite number of elements, the matrices Φ and $\tilde{\Phi}$ have either a finite or infinite number of columns. The frame expansion Eq. 2.31 can be written as

$$f = \Phi \mathbf{c} = \Phi \tilde{\Phi}^* f, \quad (2.36)$$

where the expansion coefficients are written in a column vector \mathbf{c} . The frame operator S and its inverse are

$$Sf = \Phi \tilde{\Phi}^* f \quad \text{and} \quad S^{-1}f = (\Phi \tilde{\Phi}^*)^{-1}f, \quad (2.37)$$

because $S^{-1}Sf = f$ and $SS^{-1}f = f$. Therefore, the canonical dual frame is $\tilde{\varphi}_n = S^{-1}\varphi_n = (\Phi \tilde{\Phi}^*)^{-1}\varphi_n$, and consequently the canonical dual frame matrix is

$$\tilde{\Phi} = S^{-1}\Phi = (\Phi \tilde{\Phi}^*)^{-1}\Phi. \quad (2.38)$$

Eq. 2.37 and Eq. 2.38 define a linear operator in a matrix form which is used for the reconstruction of light fields from a set of irregularly distributed samples. Eq. 2.38 suggests that the canonical dual frame can be calculated by inverting the frame operator matrix $\Phi \tilde{\Phi}^*$ and subsequent multiplication by the frame matrix Φ [81, 82]. Taking the conjugate transpose of Eq. 2.38 yields:

$$\tilde{\Phi}^* = \Phi^* (\Phi \tilde{\Phi}^*)^{-1} = \Phi^\dagger, \quad (2.39)$$

where \dagger designates *pseudo-inverse operator* [81, 82, 84].

The frame bounds A and B in Eq. 2.30 are related to issues of stability of the expansion in Eq. 2.31. Assuming that the function f and the frame elements φ_n are energy normalized, A and B are the bounds of the energy $\sum_n |\langle \varphi_n, f \rangle|^2$ of an eventual coefficient set associated with the frame. This energy determines how stably the frame expansion performs in numerical applications. For example, any additive noise in the function f might be boosted by B when calculating the coefficients $|\langle \varphi_n, f \rangle|$, while useful information might be attenuated by A and thus lost in noise.

2.2.2.2 Frame reconstruction

Any practical application of frame theory meets one of the two most common problems in signal expansion – either to find the expansion coefficients $\{c_n\}$ or to reconstruct

f . Note that the coefficients $\{c_n\}$ can be calculated in a straightforward manner from f if the dual frame $\{\tilde{\varphi}_n\}$ is known (cf. Eq. 2.31). It is also trivial to reconstruct the f from $\{c_n\}$ when the frame $\{\varphi_n\}$ is known. Here, we concentrate on approaches to find the solution of the following non-trivial inverse problems:

1. reconstruct f given the coefficient set $\{c_n\}$ and the dual frame $\{\tilde{\varphi}_n\}$;
2. reconstruct $\{c_n\}$ given f and the frame $\{\varphi_n\}$.

Obviously, it is impossible to tackle the case when neither the frame, nor its dual are known. Both problems 1 and 2 can be put in the same generic form, whose numeric solutions are discussed below.

Frame algorithm. The frame algorithm is based on the Neumann series expansion of the inverse of a non-expansive linear operator L [85]. More precisely, a linear bounded operator on a Banach space \mathbb{B} , which satisfies

$$\|g - Lg\| \leq \gamma \|g\| \quad \text{for all } g \in \mathbb{B}, \quad (2.40)$$

can be inverted through the series

$$L^{-1} = \sum_{i=0}^{\infty} (Id - L)^i. \quad (2.41)$$

Here Id denotes the identity operator. Consequently, an element $g \in \mathbb{B}$ can be reconstructed from Lg through the iterations

$$\hat{g}_{i+1} = \hat{g}_i + L(g - \hat{g}_i). \quad (2.42)$$

After i iterations the error can be estimated as $\|g - \hat{g}_i\| \leq \gamma^{i+1} \|g\|$ [47]. In fact

$$\begin{aligned} \|g - \hat{g}_i\| &= \|(Id - L)(g - \hat{g}_{i-1})\| \leq \|Id - L\| \|g - \hat{g}_{i-1}\| \\ &\leq \|Id - L\|^i \|g - \hat{g}_0\| \leq \gamma^{i+1} \|g\|, \end{aligned} \quad (2.43)$$

because the norm of the operator $Id - L$ is γ (cf. Eq. 2.40) and $\hat{g}_0 = Lg$.

This abstract formulation solves problem 1 by substituting g with f and L with the dual frame operator $\frac{2AB}{A+B}\tilde{S}f = \sum_n c_n \tilde{\varphi}_n$. In fact, this is a reconstruction from the coefficients c_n as they are the input for the frame operator. The factor $\frac{2AB}{A+B}$ is introduced for optimal convergence. The convergence is determined by the norm of $Id - \frac{2AB}{A+B}\tilde{S}$. As this operator is self-adjoint, its norm is estimated as [47, 80]:

$$\begin{aligned} \left\| Id - \frac{2AB}{A+B}\tilde{S} \right\| &= \left\langle \left(Id - \frac{2AB}{A+B}\tilde{S} \right) f, f \right\rangle = \|f\|^2 - \frac{2AB}{A+B} \sum_n |\langle f, \tilde{\varphi}_n \rangle|^2 \\ &\leq \|f\|^2 - \frac{2A}{A+B} \|f\|^2 = \frac{B-A}{A+B} \|f\|^2. \end{aligned} \quad (2.44)$$

The inequality comes from the lower frame bound B^{-1} condition for $\{\tilde{\varphi}_n\}$. The convergence is slow when A is much smaller than B . Moreover, the frame algorithm

requires good estimates of A and B in order to work optimally. When such estimates are not available, the factor $\frac{2AB}{A+B}$ can be substituted by an empirical estimate.

The frame algorithm solves problem 2 in a quite impractical manner. The only chance to estimate the coefficients is to iterate the frame algorithm for each element in the canonical dual set $\{\tilde{\varphi}_n\} = S^{-1}\{\varphi_n\}$. Hence, now g is replaced by $\tilde{\varphi}_n$ and L – by S . The generic frame algorithm, run for every element from the dual frame, converges with the same factor $\gamma = \frac{B-A}{A+B}$ [47, 80].

Least squares, pseudo-inverses and iterative solvers. Eq. 2.36 suggests that the space element f and the coefficient vector \mathbf{c} obey the relations

$$\mathbf{c} = \tilde{\Phi}^* f \quad \text{and} \quad f = \Phi \mathbf{c}. \quad (2.45)$$

In the context of problems 1 and 2, the first equation can be used to estimate f from \mathbf{c} and $\tilde{\Phi}$, and the second equation can be used to estimate \mathbf{c} from f and Φ . In general cases these systems might be under-determined, or the known reference might be noisy and lie outside of the range space of $\tilde{\Phi}^*$ (or respectively Φ). Then, it is convenient to state problems 1 and 2 as the *Least Squares (LS) problems*:

$$\hat{f} = \arg \min_f \left\| \mathbf{c} - \tilde{\Phi}^* f \right\| \quad \text{and} \quad \hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \|f - \Phi \mathbf{c}\|, \quad (2.46)$$

respectively. The straightforward solution of these problems is given by the pseudo-inverses:

$$\hat{f} = (\tilde{\Phi}^*)^\dagger \mathbf{c} = \left(\tilde{\Phi} \tilde{\Phi}^* \right)^{-1} \tilde{\Phi} \mathbf{c} \quad \text{and} \quad \hat{\mathbf{c}} = \Phi^\dagger f = \Phi^* (\Phi \Phi^*)^{-1} f. \quad (2.47)$$

Note that the solution of problem 1 is provided by the left pseudo-inverse of $\tilde{\Phi}^*$, while the solution of problem 2 is provided by the right pseudo-inverse of Φ . In this manner the substitution of Eqs. 2.45 in the respective Eqs. 2.47 guarantees that f is always reconstructed precisely.

The drawback of the least squares approach is its high computational costs, which makes it suitable only when the number of sampling points is small. In the full-rank case, pseudo-inversion is done by QR -decomposition, while for under-determined cases the Singular Value Decomposition (SVD) is preferable [84]. In both cases the pseudo-inversion has cubic complexity. The high computational complexity issue of least squares problem can be tackled by iterative methods for least squares [86]. These are methods that build up the solution step by step, by taking small increments toward the minimum of the residual $\|b - Rg\|$ at each step. Among all iterative least squares solvers, one of the most rapid convergence shows the *method of Conjugate Gradients (CG)*. It requires a Hermitean operator R in order to converge to the minimum, where $b = Rg$. However, this requirement is normally relaxed by applying the CG to the modified problem $R^*b = R^*Rg$. The resultant algorithm is referred to as conjugate gradient on normal equations [86]. It iterates as follows:

1. initialize \hat{g}_0 arbitrary, $r_0 = R^*b - R^*R\hat{g}_0$ and $d_0 = r_0$

2. for $i = 1$ to n_{it}

$$(a) \alpha = \frac{\langle r_i, r_i \rangle}{\langle R d_i, R d_i \rangle}$$

$$(b) \hat{g}_{i+1} = \hat{g}_i + \alpha d_i$$

$$(c) r_{i+1} = r_i - \alpha R^* R d_i$$

$$(d) d_{i+1} = r_{i+1} + \frac{\langle r_{i+1}, r_{i+1} \rangle}{\langle r_i, r_i \rangle} d_i$$

end

The solution is updated each time with a small portion α along the search direction d_n (step 2b). The basic idea of CG is to build the search directions d_n conjugate to each other, so that the solution is found in at most N steps for finite dimensional problems of size N . By conjugate it is meant that the directions are orthogonal to each other, where the orthogonality is measured with respect to the operator R^*R of the least squares problem $-\langle R d_i, R d_m \rangle = \langle d_i, R^* R d_m \rangle = 0$. The value of α is chosen in such a manner that the current error $e_{i+1} = \hat{g}_{i+1} - g$ is conjugate to the previous direction d_i (step 2a). This makes the residual r_{i+1} orthogonal to all previous search directions. The new direction d_{i+1} is built from this residual r_{i+1} so as to be conjugate to all previous directions (step 2d).

Translated in terms of problem 1, the solution is found by CG with the substitutions $g = f, R = \tilde{\Phi}^*$ and $b = \mathbf{c}$. This use of CG can be considered as an alternative to the frame algorithm for $\{\tilde{\varphi}_n\}$ since $\tilde{S} = \tilde{\Phi} \tilde{\Phi}^* = R^* R$ and the increment in step 2a can be computed also as $\alpha = \langle r_i, r_i \rangle / \langle d_i, \tilde{S} d_i \rangle$. Similarly, problem 2 is solved by CG with the substitutions $g = \mathbf{c}, R = \Phi$ and $b = f$. Unlike the matrix pseudo-inversion, CG solves problems 1 and 2 either in discrete or continuous spaces and frames. An advantage of the CG over the frame algorithm is that it does not involve the frame bounds or their estimates for optimal convergence. However, the convergence of the CG algorithm still depends on these bounds indirectly. The convergence factor after i iterations of CG is [86]

$$2 \left(\frac{\sqrt{v_{max}} - \sqrt{v_{min}}}{\sqrt{v_{max}} + \sqrt{v_{min}}} \right)^i, \quad (2.48)$$

where v_{max} and v_{min} are the minimum and maximum eigenvalues of the iteration operator, which is $\tilde{S} = \tilde{\Phi} \tilde{\Phi}^*$ in problem 1 and $\Phi^* \Phi$ in problem 2. Hence, $v_{max} = A^{-1}$ and $v_{min} = B^{-1}$ for problem 1. The eigenvalues of $\Phi^* \Phi$ coincide with the eigenvalues of $S = \Phi \Phi^*$ and are equal to the squared singular values of Φ , because both $\Phi^* \Phi$ and $S = \Phi \Phi^*$ are the product of Φ and its adjoint Φ^* . Therefore, $v_{max} = B$ and $v_{min} = A$ for problem 2. Straightforward calculation shows that the CG convergence factor for both problems 1 and 2 is

$$2 \left(\frac{\sqrt{B} - \sqrt{A}}{\sqrt{A} + \sqrt{B}} \right)^i. \quad (2.49)$$

Hence, CG used for frame calculations is faster than the frame algorithm, especially when there is a big difference between the frame bounds A and B .

2.2.3 Irregular sampling theory and frames

The theory of frames plays the important role of providing a mathematical framework for sampling. The connection between frames and sampling is made through a relation between the frame bound condition in Eq. 2.30 and the energy of the signal samples. Suppose that a signal f is sampled at a set of sampling points $\{x_n\}$. The energy of f is $\|f\|^2 = \int |f(x)|^2 dx$, whereas the sample energy equals $\sum_n |f(x_n)|^2$. A reasonable sampling set must ensure sample energy comparable to the energy of the continuous signal. This means that positive constants A and B should exist, such that

$$A\|f\|^2 \leq \sum_{n=1}^r |f(x_n)|^2 \leq B\|f\|^2 \quad (2.50)$$

for all f belonging to the considered functional space. If the inequalities in Eq. 2.50 are satisfied, then the set of samples $\{x_n\}$ is referred to as *a stable set of sampling* [38, 40, 42, 47, 51]. This requirement coincides with the frame bounds condition in Eq. 2.30, if the samples $\{f(x_n)\}$ are treated as the coefficients of the frame expansion in Eq. 2.31. This implies that the samples are obtained by sampling f with the dual frame $\{\tilde{\varphi}_n\}$. Hence, the term *stable set of sampling* refers to the numerical stability associated with sampling a signal f and its reconstruction from the set of samples.

In the case of sampling, the dual frame is usually constructed as translates of a single function $\tilde{\varphi}(x)$, that is, $\tilde{\varphi}_n = \tilde{\varphi}(x - x_n)$. The reconstruction formula can be written as:

$$f = \sum_n c_n \varphi_n = \sum_n \langle f, \tilde{\varphi}(x - x_n) \rangle \varphi(x - x_n), \quad (2.51)$$

where the reconstruction frame $\{\varphi_n\} = \{\varphi(x - x_n)\}$ is also formed as translates of a single function $\varphi(x)$. This is often the case, even though not a necessary requirement. The frames $\{\varphi_n\}$ and $\{\tilde{\varphi}_n\}$ are also called *reconstruction* and *sampling frames*, respectively. Functions with limited spatial support require a finite amount of samples and thus a finite-dimensional reconstruction frame $\{\varphi_n\}$. However, some classes of signals with infinite support can also be reconstructed by a finite-dimensional reconstruction frame.

In the literature, there are results whose primal concern is the sampling of band-limited functions, i.e. functions belonging to the space B_Ω as defined in Eq. 2.25. Naturally, the set of sampling functions can be selected as $\{\tilde{\varphi}_n(x) = \text{sinc}(\Omega(x - x_n))\}_{n \in \mathbb{Z}}$. Here the question arises when do these functions constitute a basis for B_Ω or at least a frame. In the Fourier domain, such a *sinc*-based set corresponds to the set of functions $\{\text{rect}\left(\frac{k_x}{\Omega_x}\right) e^{jk_x x_n}\}_{n \in \mathbb{Z}}$ and similarly there is the question of when $\{\text{rect}\left(\frac{k_x}{\Omega_x}\right) e^{jk_x x_n}\}_{n \in \mathbb{Z}}$ constitutes a basis or frame for $L^2_{[-\Omega, \Omega]}$. The answer must be searched for around the properties of the set of sampling points $\{x_n\}$, since they determine the separate elements in the sampling frame/basis. Paley and Wiener[87] were searching for the answer by representing the points x_n as perturbations around the regular Nyquist grid $n\pi/\Omega$. Respectively, the set of functions $\{\text{rect}\left(\frac{k_x}{\Omega_x}\right) e^{jk_x x_n}\}_{n \in \mathbb{Z}}$

is related to the orthogonal Fourier series basis $\{rect\left(\frac{k_x}{\Omega_x}\right)e^{jk_x n \frac{\pi}{\Omega}}\}_{n \in \mathbb{Z}}$ through the Paley-Wiener theorem [80]. Hence, the condition

$$\max_{n \in \mathbb{Z}} \left| x_n - n \frac{\pi}{\Omega} \right| \leq D < \frac{\pi}{4\Omega} \quad (2.52)$$

ensures that the set $\{\tilde{\varphi}_n(x) = sinc(\Omega(x - x_n))\}_{n \in \mathbb{Z}}$ is a Riesz basis for B_Ω with dual basis $g_n(x)_{n \in \mathbb{Z}}$ defined by:

$$g_n(x) = \frac{g(x)}{(x - x_n)g'(x_n)}, \text{ where } g(x) = \prod_{n=1}^{\infty} \left(1 - \frac{x}{x_n}\right) \left(1 - \frac{x}{x_{-n}}\right). \quad (2.53)$$

The constant D is the infimum of the deviation of the non-uniformly distributed samples $\{x_n\}_{n \in \mathbb{Z}}$ from the regularly distributed samples at the Nyquist rate $n \frac{\pi}{\Omega}$. The precise bound $D < \pi/(4\Omega)$ is due to Kadec [41]. The explicit dual basis in Eq. 2.53 is due to Levinson [88]. This basis is still impractical for numerical application due to the infinite product in the Lagrange interpolation function. This result has been reformulated and interpreted many times [89, 90, 91].

Note that the density of the samples, required by the condition in Eq. 2.52, coincides with the Nyquist rate π/Ω . Moreover, the condition permits only a slight perturbation around the regular Nyquist grid. Weakening the basis $\{\tilde{\varphi}_n(x) = sinc(\Omega(x - x_n))\}_{n \in \mathbb{Z}}$ to a frame, Duffin and Schaeffer [38] found the following weaker condition:

$$\max_{n \in \mathbb{Z}} \left| x_n - \epsilon n \frac{\pi}{\Omega} \right| \leq D < \infty, \quad (2.54)$$

where $0 < \epsilon < 1$ is an oversampling factor and with the additional requirement that $|x_n - x_m| > 0$ for $n \neq m$. This result was proven in the forward direction by Duffin and Schaeffer [38], in the backward direction by Landau [92], and revised by Feichtinger and Gröchenig [93]. The application of this result is limited, because the frame bounds A and B , which determine the stability of the frame, are not explicitly known. An estimate of the upper bound B is D/ϵ [38], while for A only its existence is known.

The generalization of these results to two dimensions is done in a separable manner [47]. The function $f(x, y)$ is considered to belong to the space $B_\Omega = B_{\Omega_x} \otimes B_{\Omega_y}$ of functions, band-limited both along x and y . The family of functions $\{\tilde{\varphi}_{n,m}(x, y) = \tilde{\varphi}_n(x) \otimes \tilde{\varphi}_m(y) = sinc(\Omega_x(x - x_n)) \cdot sinc(\Omega_y(y - y_m))\}_{n,m \in \mathbb{Z}}$ is a Riesz basis or a frame if the sampling sets $\{x_n\}$ and $\{y_m\}$ satisfy the condition Eq. 2.52 for Riesz basis or the condition Eq. 2.54 for a frame. The Fourier domain alternative of this Riesz basis (frame) is the set $\{rect\left(\frac{k_x}{\Omega_x}\right)e^{jk_x x_n} rect\left(\frac{k_y}{\Omega_y}\right)e^{jk_y y_m}\}_{n,m \in \mathbb{Z}}$.

2.2.4 Reconstruction from irregular samples

This subsection reviews the basic numerical approaches to the reconstruction of a function from an irregularly distributed sampling set. Most of them can be stated through the general formulation of the iterative Neuman series inversion of a linear operator. An accelerated numerical solution is based on discrete modeling of a band-limited function and inversion through the method of conjugate gradient. The reconstruction algorithms based on the POCS method are presented in a general form.

2.2.4.1 Marvasti (frame) method

The set $\{\tilde{\varphi}_n(x) = \text{sinc}(\Omega(x - x_n))\}_{n \in \mathbb{Z}}$ forms a frame for B_Ω , provided that the condition in Eq. 2.54 is satisfied. Therefore, a function $f \in B_\Omega$ can be reconstructed from its irregularly distributed samples $f(x_n) = \frac{\Omega}{\pi} \langle f(x), \text{sinc}(\Omega(x - x_n)) \rangle$ by iterating the frame operator $\tilde{S}f = \sum_n f(x_n) \text{sinc}(\Omega(x - x_n))$, according to the frame algorithm in Eq. 2.42. This algorithm was suggested by Marvasti et al. [94, 95]. Still, the range of application of this algorithm is limited, because the frame bounds A and B which determine the convergence of the frame algorithm are not explicitly known. Hence, the relaxation parameter for convergence of the frame algorithm is estimated empirically [94]. Without the use of a relaxation parameter in the frame algorithm there is no guaranteed rate of convergence, even if the sampling points satisfy the Nyquist criterion. Moreover, the algorithm is very sensitive to irregularities in the sampling set, such as large gaps and/or clusters.

2.2.4.2 Sauer-Allebach-Aldroubi method

The Sauer-Allebach-Aldroubi method is based directly on the iterative algorithm for finding the inverse of an operator of Eq. 2.42 based on the Neumann series. The basic idea of this method is to interpolate the error $f - f_i$ of the i -th iteration from the irregular grid of points to a continuous function, which is then projected onto the space of interest \mathbb{H} . Specifically, the algorithm proceeds as follows [37, 96]:

$$\hat{f}_{i+1}(x) = \hat{f}_i(x) + P_{\mathbb{H}} \left(\sum_{n \in \mathbb{Z}} (f(x_n) - \hat{f}_i(x_n)) q_{x_n}(x) \right). \quad (2.55)$$

The interpolation is done through the so-called *quasi-interpolant* $q_{x_n}(x)$. The operator $P_{\mathbb{H}}$ denotes orthogonal projection onto the space \mathbb{H} . In the discrete case, the interpolation is done from an irregular sampling grid to a regular one. The set of quasi-interpolants $\{q_{x_n}(x)\}_n$ must satisfy the partition of unity condition and should be as simple as possible to compute [46]. In this sense, it is convenient to choose $q_{x_n}(x)$ to act only locally on a few neighboring samples. The simplest possible choice, yet working well in practice [46], is the rectangular function $q_{x_n}(x) = \text{rect} \left((x - \frac{\bar{x}_{n+1} - \bar{x}_n}{2}) / (\bar{x}_{n+1} - \bar{x}_n) \right)$ acting between the mid-points $\bar{x}_{n+1} = \frac{x_{n+1} - x_n}{2}$ and \bar{x}_n of the given irregular samples. Other options are to take $q_{x_n}(x)$ as piecewise linear, cubic, etc.

The projection $P_{\mathbb{H}}$ clearly depends on the space \mathbb{H} . For the space of band-limited functions B_Ω , $P_{\mathbb{H}}$ can be implemented efficiently by filtering in the Fourier domain. Aldroubi [46] considers weighted shift-invariant space and Wiener-Amalgam spaces \mathbb{H} . The weighted shift-invariant space is characterized by the shift-invariance of the space-generating set $\{\varphi_n(x)\}_n$ and weighting function which controls the smoothness and/or decay of the functions within the space. Wiener-Amalgam spaces are generalizations of shift-invariant spaces which ensure that the functions locally belong to L^∞ and a projection onto the spaces imply a better local fit. In practice, the shape of the

space-generating functions can be adapted to a particular impulse response of a device related to an application. In general, these might be polynomial splines [97] or general wavelet families [98, 99, 100].

The requirement which ensures the convergence of the algorithm is derived from the norm of the operator $Id - L$, according to Eq. 2.43. The convergence factor equals the product of $\|P_{\mathbb{H}}\|$ and norm of the difference between the function and its quasi-interpolated version [46].

2.2.4.3 Adaptive weights method

The frame algorithm for the reconstruction of band-limited functions is prone to irregularities in the sampling grid such as gaps and clusters. Moreover, its convergence is undetermined, because the lower frame bound A is unknown. Both of these problems are remedied by an ad-hoc weighting approach, which turns out to have strong theoretical background [47, 45, 93, 96]. The essence of this approach is to pay more attention to more “informative” samples by introducing a scalar weight for each sample. Samples which are densely distributed in a small region are considered to be highly correlated and thus each sample carries a small portion of information about the function. On the contrary, samples near the gaps in the sampling grid would be less correlated and thus carry more information. Naturally, each sample x_n is weighted with a weight $\sqrt{w_n}$ which depends on the width of the gap between the mid-points \bar{x}_n and \bar{x}_{n+1} :

$$w_n = \bar{x}_{n+1} - \bar{x}_n = \frac{x_{n+1} - x_{n-1}}{2}. \quad (2.56)$$

Then reconstruction from the weighted samples $\{\sqrt{w_n}f(x_n)\}_{n \in \mathbb{Z}}$ can be done by the *weighted frame* $\{\sqrt{w_n}\text{sinc}(\Omega(x - x_n))\}_{n \in \mathbb{Z}}$ for B_{Ω} [45, 47, 96]. The frame bounds of this set can be explicitly estimated depending on the bandwidth Ω [47, 96] and the maximal gap τ between two consecutive samples:

$$A = \left(1 - \frac{\tau\Omega}{\pi}\right)^2 \quad \text{and} \quad B = \left(1 + \frac{\tau\Omega}{\pi}\right)^2 \quad \text{with} \quad \tau = \max_{n \in \mathbb{Z}}(x_{n+1} - x_n). \quad (2.57)$$

The reconstruction algorithm can now be done by iterating the frame algorithm with the frame operator S_w .

$$S_w f(x) = \sum_{n \in \mathbb{Z}} w_n f(x_n) \text{sinc}(\Omega(x - x_n)). \quad (2.58)$$

The proof of this result uses the fact that the set $\left\{\frac{1}{\sqrt{w_n}}P_{B_{\Omega}}\text{rect}\left(\frac{x}{\bar{x}_{n+1} - \bar{x}_n} - \frac{1}{2}\right)\right\}_{n \in \mathbb{Z}}$ is a frame for the space B_{Ω} , where $P_{B_{\Omega}}$ is orthogonal projection onto B_{Ω} [47, 96]. Hence, this frame can be used to reconstruct $f(x)$ from its local averages \bar{f}_{x_n} around the sample points x_n because they coincide with the scalar product of the function $f(x)$ and the frame elements:

$$\begin{aligned} \left\langle f(x), P_{B_{\Omega}}\text{rect}\left(\frac{x}{\bar{x}_{n+1} - \bar{x}_n} - \frac{1}{2}\right) \right\rangle &= \left\langle P_{B_{\Omega}}f(x), \text{rect}\left(\frac{x}{\bar{x}_{n+1} - \bar{x}_n} - \frac{1}{2}\right) \right\rangle \\ &= \left\langle f(x), \text{rect}\left(\frac{x}{\bar{x}_{n+1} - \bar{x}_n} - \frac{1}{2}\right) \right\rangle = \bar{f}_{x_n}. \end{aligned} \quad (2.59)$$

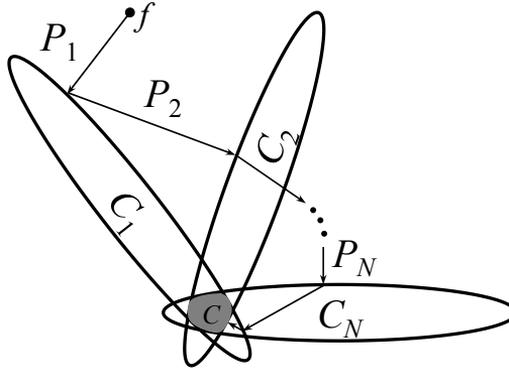


Figure 2.4: The serial POCS algorithm

This is an important practical result since the sampling devices usually use spatial averaging to estimate the sample value. For example, the sampling device in imaging is commonly a CCD sensor.

This adaptive weights method has an accelerated implementation [37, 42, 43, 51]. It is based on arranging the known samples and the unknown function generation coefficients in a special manner to obtain a matrix equation, where the matrix is Toeplitz. The equation is solved by the iterative CG method described in Subsection 2.2.2.2. The matrix-vector product inside CG is efficiently computed through FFT in logarithmic complexity.

2.2.4.4 POCS Method

POCS is an iterative optimization approach for finding an element of a feasible region defined by the intersection of a number of convex constraints, starting with an arbitrary infeasible point [16, 101]. Initially, it has been applied for various signal restoration problems in signal processing and holography [15, 17, 102] and later – for irregular sampling of band-limited signals [103, 104].

The applications of POCS use a priori information to constrain the size of the feasible solution subset in the space of interest \mathbb{H} . This information includes physical constraints related to the application and/or desired properties of the solution. Each constraint is used to define a single set C_n as all functions $f \in \mathbb{H}$ which obey this constraint (cf. Fig. 2.4). Formally, the feasible region $C \subset \mathbb{H}$ is defined as the intersection of all N constraint sets $C_n \subset \mathbb{H}$: $C = \bigcap_{n=1}^N C_n$. A solution for the problem at hand is any function which lies in this intersection. Starting from an arbitrary point f within the space \mathbb{H} , its direct projection onto the set C might be difficult to find, as it should satisfy many constraints. The idea is to use its projection $P_n f$ onto each set $C_n, n = 1, \dots, N$ separately, as those projections are easier to find. A projection $P_n f$ of $f \in \mathbb{H}$ onto the set C_n is defined through a distance measure $E(\cdot, \cdot)$:

$$P_n f = \arg \min_{g \in C_n} E(f, g). \quad (2.60)$$

A common choice for $E(\cdot, \cdot)$ is the Euclidean distance $E(f, g) = \|f - g\|$ or weighted Euclidean distance $E(f, g) = \|W(f - g)\|$ for some positive bounded weighting function W . Sometimes the projection operation is modified to admit relaxation, i.e. $P_n f$ is replaced by the relaxed operator $P_{n, \mu_n} f$ defined by [16, 17]

$$P_{n, \mu_n} f = P_n f + \mu_n (P_n f - f), \quad (2.61)$$

where μ_n is a real relaxation parameter satisfying $|\mu_n| < 1$.

There are two possible ways to iterate from an initial point $f_0 \in \mathbb{H}$ to a point in the feasible set C . One of them is to update the solution at each iteration by applying the projections $P_{n, \mu_n} f$ in a *serial* manner (cf. Fig. 2.4):

$$\hat{f}_{i+1} = P_{N, \mu_N} f \dots P_{2, \mu_2} f P_{1, \mu_1} \hat{f}_i. \quad (2.62)$$

The second approach applies the projections in a *parallel* manner:

$$\hat{f}_{i+1} = \sum_{n=1}^N w_n P_{n, \mu_n} \hat{f}_i, \quad (2.63)$$

where $w_n, n = 1, \dots, N$ are positive real scalars which satisfy $\sum_{n=1}^N w_n = 1$. They play the role of weights which determine the contribution of each set to the solution. Provided that all sets are convex and their intersection is non-empty, both serial and parallel POCS methods converge to a point in the intersection [16]. In cases when the intersection is empty, the serial projections method cycles around the sets, while the serial projections method finds an element in \mathbb{H} which minimizes the weighted average distance to all sets according to the weights $w_n, n = 1, \dots, N$. The situation is different in the presence of non-convex sets. If there are only two sets, then both parallel and serial methods converge [105]. When $N > 2$, only the serial algorithm converges, but not necessarily to a feasible element of C [106].

In many practical cases, the constraints are determined by processes which are not convenient for numerical interpretation. In such cases, the projections become too complicated to be handled with reasonable computational resources [107]. Often, the remedy is to transform the problem into a domain where the projection onto a particular set can be obtained in a simple manner. In addition, a different distance measure $E(\cdot, \cdot)$ can be introduced for a problematic set. The conditions under which such an approach converges are discussed in [50].

The use of POCS method for the reconstruction of band-limited functions $f(x) \in B_\Omega$ from a finite number of irregular samples $\{f(x_n)\}_{n=1}^s$ can be outlined as follows. All band-limited functions which share the same sample value $f(x_n)$ constitute a convex set $C_n = \{\forall g \in B_\Omega : g(x_n) = f(x_n)\} \subset B_\Omega$. Hence, a solution which belongs to the intersection of all convex sets C has the samples $\{f(x_n)\}_{n=1}^s$. The projection P_n onto each set C_n is calculated explicitly as [103]:

$$P_n g(x) = g(x) * \text{sinc}(\Omega x) + (f(x_n) - g(x_n) * \text{sinc}(\Omega x_n)) \text{sinc}(\Omega(x - x_n)). \quad (2.64)$$

2.2.5 Conclusions

The purpose of this section has been to introduce the reader to the concepts of modern sampling theory and irregular sampling, thus providing the background for the papers which form the core of this thesis. Subsection 2.2.1 has represented the sampling problem as decomposition over the space band-limited functions B_Ω spanned by the basis set $\{\frac{\Omega}{\pi} \text{sinc}(\Omega(x - nX))\}_{n \in \mathbb{Z}}$. Therefore, irregular sampling theory has been supplemented with the basics of frame and Riesz basis decompositions at first. These are Eqs. 2.29-2.34 and reformulated in a discrete manner as matrix-vector equations in Eqs. 2.35-2.39. Subsection 2.2.2 has further described the basic inversion tasks related to a frame/Riesz basis decomposition and the basic reconstruction methods. The frame algorithm has been derived in terms of Neumann series expansion of a non-expansive linear operator. The derivation has related the convergence of the algorithm to the frame bounds. Moreover, this derivation is the base of the convergence analysis of the POCS-based algorithm in **P6**. LS-based reconstruction approaches, implemented through the iterative CG method, are done based on matrix-vector expansion equations. Their description supplements publications **P2**, **P4**, **P5**, **P6**, **P7** and **P8**. Finally, Subsection 2.2.3 has translated all the frame theory results to the case of irregular sampling of signals. The conditions in Eq. 2.52 and Eq. 2.54 for recoverability of a function from a set of irregular samples suggest that the rate of the irregular sampling set should not deviate much from the rate of a uniform sampling set which satisfies the Nyquist condition in Eq. 2.26. The section has concluded with the main reconstruction algorithms. The POCS-based method from Subsection 2.2.4.4 has described briefly the theory behind the reconstruction algorithm for diffraction fields, developed in publications **P3**, **P4** and **P6**. The adaptive weights method from Subsection 2.2.4.3 provides the idea of the weighted LS reconstruction approach in **P6**.

2.3 Sampling of diffraction fields

The purpose of this section is to present the main theoretical results concerning sampling schemes for a scalar diffraction field. The discussion starts with the concept of degrees of freedom of the field as a base for evaluation of the redundancy of the ongoing sampling schemes. Two measures for degrees of freedom, namely the space-bandwidth product and Wigner distribution are presented. The initial results about sampling of α -limited scalar light fields, introduced by Gori [25] and revisited by Onural [26], are described first. These results concern the sampling of the diffraction field over a transverse plane. The second approach is the generalized sampling method of Stern and Javidi [108, 109], based on local bandwidth in the Wigner domain. The generalization of these results for sampling a volume of the diffraction field [3, 23], rather than a plane, conclude the section.

2.3.1 Degrees of freedom of a diffraction field

Sampling is the representation of a function in terms of discrete set of scalar values in an invertible manner. Before discussion about the sampling schemes of a diffraction field, it is interesting to pay attention to the theoretically sufficient amount of values which are needed to represent the field. This amount is referred to as *Number of Degrees of Freedom (NDF)* of the diffraction field. Here we present two alternative measures for it – the space-bandwidth product and the Wigner distribution.

Recall from Section 2.1 that a scalar light field $u(x, y, z)$ can be reconstructed completely at any point (x, y, z) in the half-space $z > 0$, from the knowledge of the field $u(x, y, z_T)$ at any transversal plane $(x, y, z)|_{z=z_T}$. This statement is valid for the Rayleigh-Sommerfeld integral, PWD integral and their Fresnel approximations. This plane is referred to as the field-generating plane. Without losing generality, assume that $z_T = 0$ and call the field-generating plane “initial” as well. For the sake of simplicity throughout the discussion in this section, assume that there is only one transverse dimension x , resulting in a two-dimensional field $u(x, z)$. All results concerning the sampling are presented for the 2D case and generalizations for the 3D case are straightforward. The 2D field $u(x, z)$ is completely determined by the knowledge of the initial field-generating line $u(x, 0)$. As $u(x, 0)$ is sufficient, the NDF of the diffraction field $u(x, z)$ must coincide with the NDF of the one-dimensional function $u(x, 0)$.

2.3.1.1 Space-bandwidth product

The amount of discrete values with respect to the conventional sampling representation of $u(x, 0)$ is related to the spatial and frequency support of $u(x, 0)$. As $u(x, 0)$ is a diffraction field, it is band-limited to $\frac{2\pi}{\lambda}$ (cf. Subsection 2.1.1). In practice, finite spatial extent Δ_x occurs often due to finite aperture. However, $u(x, 0)$ cannot be both space and frequency limited in theory. Fortunately, the theory admits functions

$f(x)$ whose essential part of the energy is finitely supported both in the spatial and Fourier domain, i.e.

$$|f(x)|^2 < \epsilon_x \text{ for } |x| > \frac{\Delta_x}{2} \text{ and } |F(k_x)|^2 < \epsilon_{k_x} \text{ for } |k_x| > \Omega, \epsilon_x, \epsilon_{k_x} \rightarrow 0, \quad (2.65)$$

where $F(k_x) = \mathcal{F}\{f(x)\}$. In fact, $u(x, 0)$ is essentially space limited and essentially band-limited. The field $u(x, z)$ is always slightly perturbed near the rim of the aperture and hence not strictly space limited, but rapidly vanishing in the shadow of the aperture. In the frequency domain, the PWD integral includes rapidly vanishing evanescent waves for frequencies beyond $\frac{2\pi}{\lambda}$. For the sake of generality, one can assume further that the essential frequency support of $u(x, 0)$ is for $\Omega \leq \frac{2\pi}{\lambda}$. An essentially band-limited function can be uniformly sampled according to the Nyquist interval $X = \frac{\pi}{\Omega}$, since the overlap of the adjacent replicas in the Fourier transform is negligible. The amount of essentially non-zero samples within the spatial support is then

$$\text{NDF} = \frac{\Delta_x}{X} = \frac{\Delta_x \Omega}{\pi}. \quad (2.66)$$

Alternatively, discretization can be done in the Fourier domain with interval $\frac{2\pi}{\Delta_x}$ [56], which results in the same NDF. This shows that either in the space or frequency domain $u(x, 0)$ can be recovered from NDF with essentially non-zero discrete values. Even when applying non-uniform sampling, the theorems from Subsection 2.2.3 imply a sampling density equal to the Nyquist rate, which would result to the same NDF on the average.

2.3.1.2 Wigner distribution

Eq. 2.66 derives an important, but not the most general result. This is observed in the joint space-frequency plane, through the Wigner distribution of the function $f(x) = u(x, 0)$ [110]:

$$\begin{aligned} W_f(x, k_x) &= \int_{-\infty}^{\infty} f\left(x + \frac{x'}{2}\right) f^*\left(x - \frac{x'}{2}\right) e^{-jk_x x'} dx' \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F\left(k_x + \frac{k'_x}{2}\right) F^*\left(k_x - \frac{k'_x}{2}\right) e^{jk'_x x} dk'_x, \end{aligned} \quad (2.67)$$

where $F(k_x) = \mathcal{F}\{f(x)\}$. The Wigner distribution satisfies the energy relations:

$$|f(x)|^2 = \int W_f(x, k_x) dk_x \text{ and } |F(k_x)|^2 = \int W_f(x, k_x) dx. \quad (2.68)$$

The inversion from the Wigner distribution $W_f(x, k_x)$ to the function $f(x)$ is unique, apart from a constant phase factor:

$$\int W_f(x/2, k_x) e^{jk_x x} dk_x = f(x) f^*(0), |f(0)|^2 = \int W_f(0, k_x) dk_x. \quad (2.69)$$

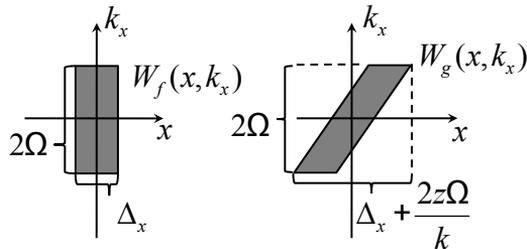


Figure 2.5: Wigner distributions $W_f(x, k_x)$ and $W_g(x, k_x)$ of two functions f and g , which occupy the same area in the space-frequency plane, but have different spatial supports, Δ_x and $\Delta_x + \frac{2z\Omega}{k}$ respectively.

The unique correspondence between the function $f(x)$ and its Wigner distribution $W_f(x, k_x)$ in addition to the energy relations of Eq. 2.68 suggest how to estimate the spatial and frequency support of $f(x)$ from $W_f(x, k_x)$. If the region in the (x, k_x) plane, where $W_f(x, k_x)$ is essentially non-zero, is rectangular, then the NDF, as defined by Eq. 2.66, equals the area of this region divided by 2π . In a more abstract manner, this area equals the number of independent non-zero terms in the Gabor expansion of $f(x)$ [108, 109, 111, 112, 113].

The difference between the area under the nonzero part of $W_f(x, k_x)$ and the space-bandwidth product of $f(x)$ can be clarified through the following example. Consider another function $g(x)$, whose Wigner distribution $W_g(x, k_x)$ is not rectangular (Fig. 2.5). The example is constructed such that the support $|g(x)|^2 = \int W_g(x, k_x) dk_x$ of g along x is larger than the support of f , but the area under the essentially non-zero part of $W_f(x, k_x)$ and $W_g(x, k_x)$ is the same. Therefore, both g and f have the same amount of independent non-zero terms in their Gabor expansion. Still, g has larger spatial support than f and, correspondingly, larger space-bandwidth product as defined by Eq. 2.66. Hence, the space-bandwidth product in Eq. 2.66 coincides with the NDF of $f(x)$ only when the $W_f(x, k_x)$ occupies a rectangular region in the space-frequency plane. A more precise definition for NDF of a function $f(x)$ is the area of the region in the (x, k_x) -plane where $W_f(x, k_x)$ is essentially non-zero:

$$\text{NDF} = \frac{1}{2\pi} \iint_{|W_f(x, k_x)| > \epsilon} dx dk_x. \quad (2.70)$$

The example in Fig. 2.5 is constructed so that $f(x) = u(x, 0)$ corresponds to a diffraction field, restricted to the initial line $(x, z)|_{z=0}$, and $g(x) = u(x, z)$ corresponds to the Fresnel diffraction field, restricted to a transversal line at a distance z . Such a situation is commonly met in practice when Fresnel diffraction between two parallel planes must be computed. Eq. 2.20 suggests that the Fourier transforms of $f(x)$ and $g(x)$ are related through $G(k_x) = F(k_x)e^{jk_x z} e^{-\frac{jz}{2k} k_x^2}$. Therefore, $W_g(x, k_x)$ can be

expressed in terms of $F(k_x)$ as follows:

$$\begin{aligned}
W_g(x, k_x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F\left(k_x + \frac{k'_x}{2}\right) e^{jk_x z} e^{-\frac{jz}{2k}\left(k_x + \frac{k'_x}{2}\right)^2} \times \\
&\quad F^*\left(k_x - \frac{k'_x}{2}\right) e^{-jk_x z} e^{\frac{jz}{2k}\left(k_x - \frac{k'_x}{2}\right)^2} e^{jk'_x x} dk'_x \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} F\left(k_x + \frac{k'_x}{2}\right) F^*\left(k_x - \frac{k'_x}{2}\right) e^{jk'_x x} e^{-\frac{jz}{2k}\left(k_x^2 + k_x k'_x + \frac{k'^2_x}{4} - k_x^2 + k_x k'_x - \frac{k'^2_x}{4}\right)} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} F\left(k_x + \frac{k'_x}{2}\right) F^*\left(k_x - \frac{k'_x}{2}\right) e^{jk'_x\left(x - \frac{z}{k}k_x\right)} = W_f\left(x - \frac{z}{k}k_x, k_x\right). \quad (2.71)
\end{aligned}$$

Eq. 2.71 suggests that Fresnel propagation of a function $f(x)$ in positive z direction shears its Wigner distribution function along x , as shown in Fig. 2.5. Hence, a function $f(x)$, which is essentially band-limited to 2Ω , increases its essential spatial support from Δ_x to $\Delta_x + 2z\Omega/k$ after propagation at a distance z . Consequently, the space-bandwidth product in Eq. 2.66 increases, unlike the area under its Wigner distribution according to Eq. 2.70. Therefore, the Wigner distribution is a more general tool than the space-bandwidth product for measuring the NDF of a diffraction field $u(x, z)$.

2.3.2 Results based on the Nyquist theorem

A very frequent case in practice is the calculation of Fresnel diffraction between two parallel lines $(x', z)|_{z=0}$ and $(x, z)|_{z=\text{const}}$. Numerical computation of $u(x, z)$, restricted to the line $(x, z)|_{z=\text{const}}$, from $u(x', 0)$ requires sampling on both lines. Already for decades, the scientific community has used approaches based on the Nyquist theorem, summarized by Gori [25] and generalized by Onural [26].

Gori considers cases when either the function on the initial line $u(x', 0)$, or its Fresnel pattern $u(x, z)$ on distance z are spatially limited. Consider first the case when $u(x, z)$ is finitely supported on the interval $x \in [-\frac{x_0}{2}, \frac{x_0}{2}]$. Recall from Eq. 2.23 that $u(x, z)$ can be considered as a scaled and modulated version of the Fourier transform of $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ if the $\frac{2\pi x}{\lambda z}$ is the angular frequency. The scaling and modulation do not change the support of $u(x, z)$ and hence the Fourier transform of $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ is finitely supported on the interval $x \in [-\frac{\pi x_0}{\lambda z}, \frac{\pi x_0}{\lambda z}]$. Therefore, the Nyquist criterion from Eq. 2.26 can be applied for $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ with the substitution $X = X'$ for the sampling interval and $\Omega = \frac{\pi x_0}{\lambda z}$ to obtain:

$$X' \leq \frac{\lambda z}{x_0}. \quad (2.72)$$

Such a sampling interval X' guarantees that $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ can be recovered from its samples and, consequently, $u(x', 0)$ can be recovered. The number N of necessary samples depends on the spatial support of $u(x', 0)$. If $u(x', 0)$ is considered to be

essentially space-limited within $x' \in [-\frac{\Delta_x}{2}, \frac{\Delta_x}{2})$ and essentially band-limited within $k_x \in [-\Omega, \Omega)$, then N can be related to the NDF of $u(x', 0)$. The spatial support of $u(x, z)$ is $x_0 = \Delta_x + \frac{2z\Omega}{k}$ according to Eq. 2.71, since $u(x, z)$ is the Fresnel transform of $u(x', 0)$. The number of required samples for $u(x', 0)$ is

$$N = \frac{\Delta_x}{X'} = \frac{\Delta_x x_0}{\lambda z} = \frac{\Delta_x(\Delta_x k + 2z\Omega)}{k\lambda z} = \text{NDF} + \frac{\Delta_x^2}{\lambda z}. \quad (2.73)$$

This computation directly shows the amount of excess samples $\frac{\Delta_x^2}{\lambda z}$ depending on the propagation distance.

As a second case, Gori considers that $u(x', 0)$ is spatially supported on a finite interval $x' \in [-\frac{\Delta_x}{2}, \frac{\Delta_x}{2})$. The term $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ is also supported on $x' \in [-\frac{\Delta_x}{2}, \frac{\Delta_x}{2})$. It is the inverse Fourier transform of $j\lambda z u(\omega_x, z)e^{-jkz}e^{\frac{j\lambda z}{4\pi}\omega_x^2}$ according to Eq. 2.23, where $\omega_x = \frac{2\pi x}{\lambda z}$. Sampling $u(x, z)$ with interval X means sampling $j\lambda z u(\omega_x, z)e^{-jkz}e^{\frac{j\lambda z}{4\pi}\omega_x^2}$ with interval $\frac{2\pi X}{\lambda z}$. This results in periodization of $u(x', 0)e^{\frac{j\pi}{\lambda z}x'^2}$ [56] with period $\frac{\lambda z}{X}$ and the adjacent periodic replicas do not overlap, provided that

$$X \leq \frac{\lambda z}{\Delta_x}. \quad (2.74)$$

Note that this criterion is exactly the same as the criterion in Eq. 2.72. This enables employing $u(x, z)$ and $u(x', 0)$ as an FFT pair, according to the discussion at the end of Subsection 2.1.4. Moreover, this brings the same excess over the NDF when $u(x', 0)$ is essentially space-limited within $x' \in [-\frac{\Delta_x}{2}, \frac{\Delta_x}{2})$ and essentially band-limited within $k_x \in [-\Omega, \Omega)$:

$$N = \frac{\Delta_x + 2z\Omega/k}{X} = \frac{(\Delta_x + 2z\Omega/k)\Delta_x}{\lambda z} = \text{NDF} + \frac{\Delta_x^2}{\lambda z}. \quad (2.75)$$

Onural generalizes the result of Eq. 2.72 and Eq. 2.74 for Fresnel diffraction between two planes for more general uniform sampling grids as quincunx or hexagonal [26]. He completes the work of Gori by discussing also the case of sampling $u(x, z)$ when $u(x', 0)$ is strictly band-limited within $[-\Omega, \Omega)$. Eq. 2.20 implies that $u(x, z)$ is also band-limited within $[-\Omega, \Omega)$ and can be conventionally sampled according to the Nyquist criterion with sampling interval $X \leq \frac{\pi}{\Omega}$. Note that this interval is useful when the convolutional algorithm from Eq. 2.20 is used to compute Fresnel propagation, while the criterion in Eq. 2.72 and Eq. 2.74 is required when the Fourier transform-like algorithm from Eq. 2.23 is used for computation. The criterion $X \leq \frac{\pi}{\Omega}$, used for sampling $u(x, z)$ when $u(x', 0)$ is essentially space-limited within $x' \in [-\frac{\Delta_x}{2}, \frac{\Delta_x}{2})$ and essentially band-limited within $k_x \in [-\Omega, \Omega)$, requires

$$N = \frac{\Delta_x + 2z\Omega/k}{X} = \frac{(\Delta_x + 2z\Omega/k)\Omega}{\pi} = \text{NDF} + \frac{\lambda z \Omega^2}{\pi^2} \quad (2.76)$$

samples. Note that now the excess over NDF grows with the distance z in contrast with Eq. 2.73 and Eq. 2.75. This fact can be used to select either Eq. 2.20 or Eq. 2.23 for Fresnel diffraction computations, depending on the problem at hand.

2.3.3 Results based on local bandwidth

The sampling criteria, formalized by Gori [25] and Onural [26], have an interesting interpretation in terms of Local Bandwidth (LB). Analyzed in the Wigner domain, the sampling problem engenders the generalized sampling theorem by Stern and Javidi [108, 109]. The LB is also employed to produce an optimal non-uniform, but regular sampling grid on the Fresnel plane [29]. In essence, these results are built up by respecting the Nyquist criterion in terms of LB, rather than global bandwidth. Such strategy is useful for diffraction fields restricted to a Fresnel plane, since their LB is often smaller than their global bandwidth.

Stern and Javidi [108, 109] use uniform sampling grid over a line $(x, z)|_{z=\text{const}}$ where the Fresnel approximation is valid. The sampling step X is uniform over the whole line, respecting the Nyquist criterion for the Maximum Local Bandwidth (MLB): $X \leq \frac{2\pi}{\text{MLB}}$. A rigorous derivation of this result is made by the use of the Wigner distribution. Sampling a function $f(x)$ with an interval X effectively periodizes its Wigner distribution $W_f(x, k_x)$ by $\frac{\pi}{X}$ along the frequency axis k_x and samples it along x with an interval $\frac{X}{2}$ [114]:

$$W_{f_s}(n, k_x) = \frac{\pi}{T} \sum_l (-1)^{nl} W_f \left(\frac{nX}{2}, k_x - \frac{l\pi}{X} \right), \quad (2.77)$$

where $W_{f_s}(n, k_x)$ is the Wigner distribution of the sampled function $f_s(x)$. Periodization with $\frac{\pi}{X}$ implies that the sampling period X should obey $\frac{\pi}{X} \geq \text{MLB}$, so that the adjacent periodic replicas of $W_f(x, k_x)$ do not overlap. Note that this is twice higher than the Nyquist rate when MLB is equal to the global bandwidth. However, even when taking $\frac{2\pi}{X} \geq \text{MLB}$, the signal can be reconstructed. At sampling interval $X \leq \frac{2\pi}{\text{MLB}}$, each value $W_{f_s}(n, k_x)$ is result of the overlap by at most two adjacent replicas $W_f \left(\frac{nX}{2}, k_x - \frac{l\pi}{X} \right)$ and $W_f \left(\frac{nX}{2}, k_x - \frac{(l+1)\pi}{X} \right)$. Roughly speaking, ‘‘half’’ of the signal information for recovering $W_f(x, k_x)$ is lost due to this overlap. However, the samples in $W_{f_s}(n, k_x)$ occur at $\frac{Xn}{2}$, i.e. there are twice as many samples than required by the Nyquist theorem occupying the spatial support of $f(x)$. A rigorous analysis on how the whole information about the signal is recovered from $W_{f_s}(n, k_x)$ when $X \leq \frac{2\pi}{\text{MLB}}$ can be found in [108, 109].

Eq. 2.71 can be used to determine the MLB of the Fresnel diffraction field $u(x, z)$, restricted to a line $(x, z)|_{z=\text{const}}$ from the spatial support Δ_x and bandwidth 2Ω of the diffraction field $u(x, 0)$ on the reference line. The relation $W_{u(x,z)}(x, k_x) = W_{u(x,0)} \left(x - \frac{z}{k} k_x, k_x \right)$ suggests that a rectangular region, occupied by $W_{u(x,0)}(x, k_x)$, is sheared to a parallelogram, occupied by $W_{u(x,z)}(x, k_x)$. The vertical sides of the rectangle at $x = \pm \frac{\Delta_x}{2}$ are mapped to tilted lines on the parallelogram such that $x - \frac{z}{k} k_x = \pm \frac{\Delta_x}{2}$ or $k_x = \frac{z}{k} \left(x \mp \frac{\Delta_x}{2} \right)$. Depending on the value of z , the top left corner $\left(x = -\frac{\Delta_x}{2} + \frac{z}{k} \Omega, \Omega \right)$ of the parallelogram might appear before or after the bottom right corner $\left(x = \frac{\Delta_x}{2} - \frac{z}{k} \Omega, \Omega \right)$ along the x -axis (Fig. 2.6). The first case occurs for $z \leq \frac{\Delta_x k}{2\Omega}$ and determines $\text{MLB} = 2\Omega$ (Fig. 2.6 left). The second case occurs for $z > \frac{\Delta_x k}{2\Omega}$ and determines the MLB as the space along k_x between the left

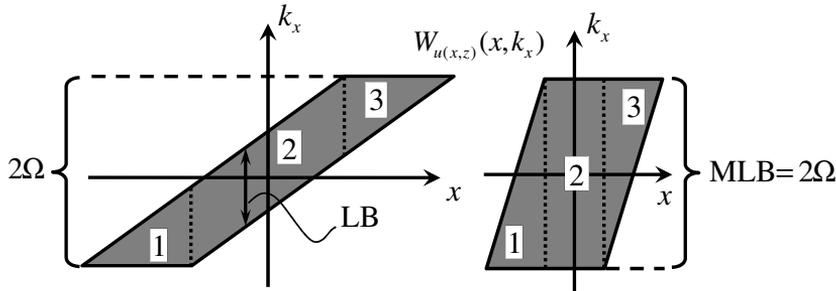


Figure 2.6: Wigner distribution $W_{u(x,z)}(x, k_x)$ of the Fresnel diffraction field $u(x, z)$, restricted to the line $(x, z)|_{z=\text{const}}$ for two different diffraction distances: $z \leq \frac{\Delta_x k}{2\Omega}$ (left) and $z > \frac{\Delta_x k}{2\Omega}$ (right).

$k_x^{(left)}(x) = \frac{z}{k} \left(x + \frac{\Delta_x}{2}\right)$ and right $k_x^{(right)}(x) = \frac{z}{k} \left(x - \frac{\Delta_x}{2}\right)$ sides of the parallelogram (Fig. 2.6 right):

$$\text{MLB} = k_x^{(left)} - k_x^{(right)} = \frac{z}{k} \left(x + \frac{\Delta_x}{2}\right) - \frac{z}{k} \left(x - \frac{\Delta_x}{2}\right) = \frac{k\Delta_x}{z}. \quad (2.78)$$

The first case requires conventional Nyquist sampling with interval $X \leq \frac{\pi}{\Omega}$, because $\text{MLB} = 2\Omega$. The second case requires $X \leq \frac{2\pi}{\text{MLB}} = \frac{\lambda z}{\Delta_x}$ and thus it is equivalent to the criterion in Eq. 2.74. Note that $z \leq \frac{\Delta_x k}{2\Omega}$ is the same as $\frac{\lambda z \Omega^2}{\pi^2} \leq \frac{\Delta_x^2}{\lambda z}$ and the convolutional algorithm in Eq. 2.20 provides less excess on the NDF in comparison with the Fourier transform-like algorithm in Eq. 2.23, according to Eq. 2.76 and Eq. 2.75. The Fourier transform-like algorithm in Eq. 2.23 is less redundant when $z \leq \frac{\Delta_x k}{2\Omega}$. Hence, depending on the propagation distance z , either the convolutional or the Fourier transform-like algorithm is preferred as the less computationally expensive algorithm. In this context, an intuitive algorithm to recover $u(x, 0)$ from the samples of $u(x, z)$ is to compute discrete back-propagation followed by continuous reconstruction by low-pass filtering.

Both of these sampling approaches are redundant with respect to the NDF of the diffraction field. This is due to the fact that they use an uniform sampling step within the spatial support of $u(x, z)$ and thus do not take full advantage of the local bandwidth. For the sake of no redundancy, VanderLugt suggests a non-uniform, but regular sampling grid, where the sampling step depends on the position of the respective sample [29]. The sample spacing is determined according to the Nyquist criterion for the $\text{LB}(x)$ at position x : $X(x) = \frac{2\pi}{\text{LB}(x)}$. Such a sampling scheme is illustrated in Fig. 2.7, where regions 1, 2 and 3 correspond to regions 1, 2 and 3 of the Wigner distribution of $u(x, z)$ in Fig. 2.6. The LB is constant $\text{LB} = 2\Omega$ or $\text{LB} = \text{MLB} = \frac{k\Delta_x}{z}$ in region 2 and the sampling step is uniform $X_2(x) = \frac{\pi}{\Omega}$ or $X_2(x) = \frac{\lambda z}{\Delta_x}$, respectively. The LB decays linearly towards the outer end region 1 and region 3, $\text{LB}(x) = k_x^{(left)}(x) + \Omega = \frac{z}{k} \left(x + \frac{\Delta_x}{2}\right) + \Omega$ for region 1 and $\text{LB}(x) = \Omega - k_x^{(right)}(x) = \Omega - \frac{z}{k} \left(x - \frac{\Delta_x}{2}\right)$ for region 3. The average amount of samples, used inside each region, can be calculated as the x -support of the region times its

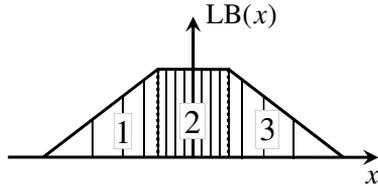


Figure 2.7: Optimum non-uniform sampling according to $LB(x)$ along the support of Fresnel diffraction field $u(x, z)$ which is restricted to the line $(x, z)|_{z=\text{const}}$. The thin vertical lines indicate the positions of the samples.

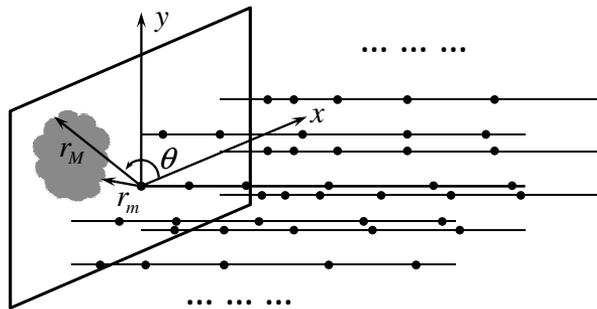


Figure 2.8: Finite aperture with minimal r_m and maximal r_M radial distances to a particular line in the volumetric grid.

average LB divided by 2π . This amount equals the area of each region divided by 2π . Therefore, the average total amount of samples is equal to $\frac{1}{2\pi}$ times the area of the parallelogram occupied by its Wigner distribution, or equivalently, to the NDF of the diffraction field. Hence, this sampling strategy is optimal as it requires no redundancy in the samples. However, the original paper by VanderLugt does not discuss, or even imply, an eventual approach to recover the continuous $u(x, z)$ from the set of non-uniform samples [29].

2.3.4 Generalizations for a volume

Many light field reconstruction and synthesis applications are related to 3D imaging [1, 2] and light beam shaping [3]. These commonly specify a desired light field within an application-oriented volume of interest, thus requiring depth. Therefore, any related numerical computations require sampling of the whole volume of interest. To the best of the author's knowledge, the only volumetric sampling strategy has been developed by Piestun et al. [3, 23, 49]. There, the volume is sampled with an ensemble of longitudinal lines, parallel to the optical axis. Each line is sampled non-uniformly, but regularly. The spacing between lines is chosen on the base of Eq.2.74.

The sampling grid on each line is determined in a manner which guarantees that the light field along the line can be reconstructed from the samples. Without loss

of generality, consider that the line is the optical axis $(x, y, z) = (0, 0, z)$. Assume that the field $u(x, y, 0)$ is known over an arbitrary aperture in the reference plane $(x, y, z)|_{z=0}$ (Fig. 2.8). The Fresnel integral from Eq.2.17 can be written for the optical axis in cylindric coordinate system (r, θ, z) as:

$$\begin{aligned} u(0, 0, z) &= \frac{e^{jkz}}{j\lambda z} \int_0^\infty \int_0^{2\pi} ru(r, \theta, 0) e^{j\frac{k}{2z}r^2} d\theta dr = -\frac{jve^{j\frac{k^2}{2v}}}{2\pi} \int_0^\infty f(r) e^{jvr^2} dr^2 \\ &= -\frac{je^{j\frac{k^2}{2v}}}{2\pi} v \mathcal{F}^{-1}\{\bar{f}(p)\} = \frac{e^{j\frac{k^2}{2v}}}{2\pi} \mathcal{F}^{-1}\left\{\frac{d\bar{f}(p)}{dp}\right\}, \end{aligned} \quad (2.79)$$

where $v = \frac{k}{2z}$, $p = r^2$ and $\bar{f}(p) = f(r) = \int u(r, \theta, 0) d\theta$. The function $\frac{d\bar{f}(p)}{dp}$ is the Fourier transform of the function $g(v) = 2\pi e^{-j\frac{k^2}{2v}} u(0, 0, \frac{k}{2v})$. Therefore, the support of $\frac{d\bar{f}(p)}{dp}$ can be used in the Nyquist criterion to determine the sampling step V for sampling $g(v)$ uniformly over v . Note that $u(0, 0, z)$ can be uniquely determined from $g(v)$. The support of $\frac{d\bar{f}(p)}{dp}$ is the same as the support of $\bar{f}(p)$, which, in turn, can be determined from the shape of the aperture. Assume that the aperture opens from r_m to r_M with respect to the origin (Fig. 2.8). Then the function $f(r)$ is nonzero between r_m and r_M and the support of $\bar{f}(p)$ and $\frac{d\bar{f}(p)}{dp}$ is $r_M^2 - r_m^2$. The sampling step is $V = \frac{2\pi}{r_M^2 - r_m^2}$. The variable z is inversely proportional to v and therefore the uniform grid along v is transformed to a non-uniform, but regular grid along z with variable step size Z . Two consecutive samples along v , spaced by V , occur at $v = \frac{k}{2z}$ and $v = \frac{k}{2(z-Z)}$. This relation can be used to determine Z from the Nyquist criterion for V as:

$$Z \leq \frac{4\pi z^2}{k(r_M^2 - r_m^2) - 4\pi z}. \quad (2.80)$$

Note that the distances r_m and r_M between the aperture and the line are individual for each line in the volumetric grid. Hence, the non-uniform grid along the line is individual for each line. When needed, the individual grids can be aligned by selecting the sampling step Z according to the maximal value of $r_M^2 - r_m^2$. Moreover, Eq.2.74 determines z -dependent spacing between the lines in the volumetric grid. As the lines need to stay parallel to the z -axis, the z -dependency in Eq.2.74 is relaxed by selecting the smallest z in the volume of interest. Such a 3D grid is very redundant with respect to the degrees of freedom of the light field. Not only are the samples per plane redundant, but many planes are taken into computation for the sake of specifying the desired scene in depth.

2.3.5 Conclusions

This section has surveyed the main results related to the sampling of diffraction fields. This detailed survey has been made in order to motivate the idea of reconstructing a diffraction field from a set of samples irregularly distributed within a region of interest. The main motivation is to have a sampling scheme which is able to sample

the diffraction field in a volume and at the same time to be non-redundant, i.e. does not require many more samples than the NDF of the field. The sampling scheme in Subsection 2.3.4 has been derived for a volume of interest, but is very redundant as the required samples exceed by multiple times the NDF of the field. On the other hand, the techniques presented in Subsection 2.3.2 and Subsection 2.3.3 exceed the NDF by just a little, but distribute the samples over a single transversal plane of the field in an uniform manner.

The measure for efficiency is the reason why the section has started by describing two measures for NDF of a function – the space-bandwidth product in Subsection 2.3.1.1 and the area under the Wigner distribution in Subsection 2.3.1.2. The space-bandwidth product is accurate enough when the function occupies a rectangle in the time-frequency plane. This is a reasonable assumption if the function is the light field distribution on the initial line, right behind an aperture. However, this rectangle shears along x after some propagation distance z . In this case, the area under the Wigner distribution of the propagated function is a more accurate measure for the NDF of the diffraction field. Section 2.3.2 has presented sampling schemes for the cases when either the diffraction pattern or its bandwidth is limited. The results are based on the Nyquist condition and the amount of needed samples are comparable to the NDF of the field, with a slight excess over the NDF. The results presented in Subsection 2.3.3 are derived on the basis of the maximal local bandwidth, measured by the Wigner distribution. Even though this analysis is more rigorous, the results are essentially the same as those in Subsection 2.3.2. An advantage of the local bandwidth analysis is that it admits a strictly non-redundant sampling scheme, which distributes the required samples on the plane in a non-uniform, but regular manner.

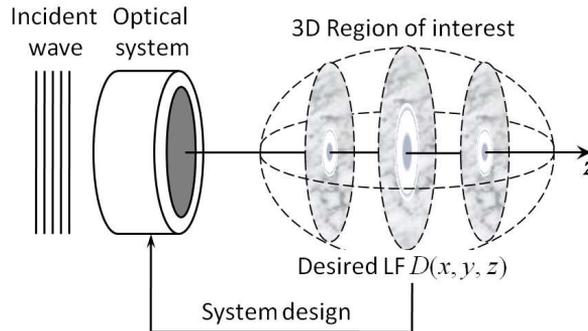


Figure 2.9: Light field synthesis within a 3D volume of interest. The optical system is illuminated by an incident wave to produce the desired light field distribution $D(x, y, z)$.

2.4 Light field reconstruction and synthesis

Light field synthesis refers to the generation of desired optical light-field distribution within a region of interest in 3D space. An incident light wave enters an optical system to be modulated, and then propagates in an homogeneous medium beyond the system where the desired 3D optical field distribution is formed, as illustrated in Fig. 2.9.

The modulation takes place in optical systems that contain classical optical elements such as lenses, prisms, beam splitters, optical fibers and apertures in addition to modulation elements. The modulation elements can be fixed or electronically reconfigurable DOEs/SLMs which modulate the incident wave by masking with a certain relief structure and/or electronically controlled transparency. The DMDs are the type of SLMs considered in this thesis.

The requirements for a 3D light field come from the specifics of the application at hand, e.g. recorded data in digital holography, synthetic objects in computer generated holography, desired shape for a light beam. Mathematically described, the requirements might not be feasible to produce a physically consistent light field which satisfies the Helmholtz wave equation. In addition, the light field may be constrained by finite bandwidth, aperture and energy. Therefore, the primal goal is to search for a field which approximates the requirements to the best extent.

The configuration of the modulation element has the greatest influence on synthesized 3D light field distribution. This configuration is commonly found by a numerical optimization procedure. As discrete computations, these require a discrete representation of the light field and the desired, application-oriented, 3D distribution. While the problem of generating arbitrary wave fronts on one plane has already been well understood, the generation of light distributions within a 3D domain has only recently been attempted. In the following, approaches for solving either the planar or the volumetric case are surveyed. Most of these are primarily based on variations of the POCS algorithm.

This section is organized as follows. First, the significance and specifics of light field synthesis and reconstruction is motivated by a review of potential applications in 3D imaging and light beam shaping. Subsection 2.4.2 describes the specifics for fixed and electronically reconfigurable DOEs/SLMs. Subsection 2.4.3 presents the most popular techniques used for LF synthesis on a transversal plane behind the modulation element. Subsection 2.4.4 summarizes the state-of-the-art LF reconstruction algorithms within a volume.

2.4.1 Applications of LF reconstruction and synthesis

2.4.1.1 Digital holography

Holography is considered to be the most realistic approach for 3D scene capture [2, 55, 115]. Conventional recording sensors such as CCD, CMOS, or analog thin films are able to capture only the amplitude of light on a certain plane. However, within any transversal plane, the phase of the wavefront is the one which carries the scene depth information [4]. In holography, the phase is recorded indirectly by interfering with the object field $O = |O|e^{j\phi_O}$ with a known reference wave $R = |R|e^{j\phi_R}$ onto the plane of the sensor [2, 55, 115]. In this notation, the (x, y) -dependency of O and R is omitted for the sake of readability. The recorded intensity on the sensor is

$$I = |O + R|^2 = OO^* + RR^* + RO^* + OR^*. \quad (2.81)$$

The object light field can be re-synthesized once it is known on a desired reference plane. This can be reconstructed from the last term OR^* in Eq. 2.81 by multiplying with the known wave $\frac{1}{|R|}e^{j\phi_R}$ and back-propagating O to the desired reference plane. Depending on the conditions, a different diffraction model is used for the back-propagation and the holography is referred to as Fresnel, Fraunhofer or image plane hologram [2, 55]. R-S or PWD diffraction is used rarely, when the precision of computations is necessary.

There are various approaches to extract OR^* out of the intensity mixture I [2, 55, 115, 116]. The two most common approaches are known as *temporal phase shifting* and *spatial phase shifting* [55]. The temporal phase shifting approach takes K recordings of the interference $I_i = |O + R_i|^2, i = 1, \dots, K$ at subsequent time instants, each with reference wave R_i . The reference waves R_i are chosen to have different constant phase offsets $\phi_{R,i}$: $R_i(x, y) = |R(x, y)|e^{j\phi_{R,i}}$. The term OR^* is extracted by adding the intensities I_i in an appropriate manner. This is possible, provided that $\sum_i e^{j\phi_{R,i}} = 0$ and $\sum_i e^{j2\phi_{R,i}} = 0$ [2]. One commonly used choice is $\phi_{R,i} = \frac{2\pi i}{K}$ for $K \geq 3$. When $K = 4$, the term OR^* is extracted as $OR^* = (I_1 + jI_2 - I_3 - jI_4)/4$.

The spatial phase shifting approach introduces a linear offset in the phase of the reference wave by tilting it with respect to the optical axis [2, 55]. Therefore, this approach is often referred to as “off-axis” holography. The reference wave is usually selected as a plane wave. A tilt at angle θ_x around the y -axis with respect to the

optical z -axis introduces a linear phase shift $R(x, y) = |R|e^{jkx \sin \theta_x}$. Analogously, a tilt at angle θ_y around the x -axis with respect to the optical z -axis introduces a linear phase shift $R(x, y) = |R|e^{jk_y \sin \theta_y}$. Generally, the tilted plane wave has the expression $R(x, y) = |R|e^{jk(x \sin \theta_x + y \sin \theta_y)}$ whose 2D Fourier transform is a spike at $(k_x, k_y) = (k \sin \theta_x, k \sin \theta_y)$. In the Fourier domain, the term $\mathcal{F}\{RR^*\}$ is a spike at the origin $(k_x, k_y) = (0, 0)$, $\mathcal{F}\{OO^*\}$ is centered at the origin with a twice larger extent than $\mathcal{F}\{O\}$ and the terms $\mathcal{F}\{OR^*\}$ and $\mathcal{F}\{RO^*\}$ are shifted replicas of $\mathcal{F}\{O\}$ at $(k_x, k_y) = (-k \sin \theta_x, -k \sin \theta_y)$ and $(k_x, k_y) = (k \sin \theta_x, k \sin \theta_y)$, respectively. Hence, $\mathcal{F}\{OO^*\}$ and $\mathcal{F}\{OR^*\}$ do not overlap, provided that $k \sin \theta_x$ and $k \sin \theta_y$ are at least 1.5 times the support of $\mathcal{F}\{O\}$ along k_x and k_y , respectively. Moreover, the sensor should record frequencies four times larger than the bandwidth of the object. When these are satisfied, $\mathcal{F}\{OR^*\}$ is extracted from $\mathcal{F}\{I\}$ by e.g. band-pass filtering.

2.4.1.2 Computer generated holography

Computer generated holography is a general term which encompasses the light field reconstruction and synthesis of a 3D scene, given its abstract representation in mathematical or computer graphics terms [1, 2, 3, 10]. This subsection presents the framework for reconstruction of the wave front of a computer generated hologram for subsequent synthesis with a modulation device. Digital holography can be considered as a part of computer generated holography which is related to the light field reconstruction and display of real world objects, recorded by a sensor.

The objects constituting the 3D scene are not necessary real, might be related to any application involving computer generated imagery: from CAD to gaming, virtual reality -based training and testing, medical imaging, cultural heritage. Commonly, the 3D scenes are assumed to be available in geometrical, or vector graphics models [1]. These represent a scene as a composition of elementary diffracting elements such as point scatters, curve segments, slits, etc. The most general representation of 3D scene or object is provided by a point cloud. A point cloud is a set of vertices of (x, y, z) coordinates, typically intended for the representation of the external surface of an object. In computer graphics, point clouds often serve as a base of surface models such as polygon meshes or NURBS. In computer generated holography, a point cloud is very convenient to pose the light field reconstruction problem as reconstruction from irregularly distributed samples, considered in publications **P2**, **P3**, **P4**, **P5**, **P6** and **P7**, **P8**. A polygon mesh is a collection of vertices, edges and faces that defines the shape of a polyhedral object. NURBS is a mathematical model for representing curves and surfaces, defined by its order and a set of weighted control points. A wire frame model is created by specifying each edge of the physical object where two mathematically continuous smooth surfaces meet, or by connecting an object's constituent vertices.

The light wave front of a 3D scene represented by a geometrical model is usually calculated as a superposition of elemental wave fronts which correspond to each element of the geometrical model [2, 10]. Such a technique offers a crude approximation which offers reasonable results for simple scenes and large size of the models. Better techniques are based on direct 3D light field reconstruction, as discussed in Subsection 2.4.4.

2.4.1.3 Light beam shaping

A light field undergoes diffraction effects upon propagation, which tend to induce variations in its profile. For example, a diffraction pattern which stays in focus at a certain plane, undergoes dramatic changes away from this plane. However, many practical applications drive their interest towards light beams whose profile does not change, or maintains certain properties for a long diffraction distance. Some general cases with their applications are listed below [3].

3D curves. 3D light curves have peak intensity of the light distribution concentrated around some 3D curve, following any trajectory in space. Non-diffracting beams have a prominent intensity peak propagating for long distances along the axis without significant divergence. These beams are used in measurement and alignment systems. Non-diffracting beam array is an array of light spots which keep their intensity and width along propagation. Another typical 3D light curve is the dashed beam, which consists of axial intensity peaks that appear periodically along propagation. Such beams find application in measurement and multi-focal imaging.

Dark beams. Dark beams are beams where a dark region with a certain pattern and axial extent is important. This situation appears in optical systems that are designed for specific purposes such as optical inspection systems, optical tweezers, and atom trapping and guiding.

Rotating beams. A special case of generalized propagation invariance occurs when the transverse intensity distribution is invariant upon propagation, except for a continuous rotation about the optical axis and a scale change. A typical example is screw wave front dislocations.

Extended depth of focus. This is a 3D light field where a complex 2D structure might appear in focus for a long propagation distance. For example, in a microscope only one plane and its very near proximity produce a sharp image, while objects lying in other planes appear blurred. In microlithography, a mask image pattern must be projected onto the wafer surface in focus. Extended depth of focus of the image will relax the wafer alignment and planarization.

Super-resolving lenses. These are synthesized phase masks whose point spread function is narrower in comparison with conventional optical lenses. This might be used to improve the resolution in, e.g. microscopic imaging.

2.4.2 Light wave modulation elements

This subsection describes different elements used to modulate an incident wave in order to produce a desired light distribution. An ideal wave modulation element would modify accurately the amplitude and phase of an incident light wave in a continuous manner with a high resolution on a large useful area. Thus, the full number of degrees of freedom of the desired light field can be reproduced. Unfortunately, there is no

such ideal device and the technological limitations must be taken into account. The greatest limitations are coupling between phase and amplitude, limited aperture, resolution and dynamic range. In the following, the technological principles of commonly available modulation elements are discussed with respect to their limitations.

Fixed DOE/SLM. The earliest available form of DOEs are binary printable media, used to synthesize computer generated holograms in the Fraunhofer region [2, 117]. The desired resolution of the Fourier transform of the synthesized object is achieved by dividing the media into square cells. A fully transparent opening inside each cell modulates both the phase and amplitude of the desired object. The amplitude is modulated by the size of the opening, while the phase is modulated by a shift of the opening within the cell. In this manner the size of the cell limits the freedom of the coupled amplitude-phase modulation. A higher degree of freedom can be achieved by combining and/or rearranging two three or four cells to contribute to a single resolution element in the Fourier transform [118, 119]. These methods are known as *phase detour* methods.

Current state-of-the-art DOEs modulate an incident light wave by a mask containing an amplitude pattern or a relief structure. Usually, these operate on the phase of the incident wave, leaving the intensity unchanged. The diffractive surface of DOEs is split into an array of cells, each designed to transform the phase of the coherent illuminating wave by a specified amount. The required phase change at each cell is achieved by accurately producing a sub-wavelength multi-level depth profile. Upon illumination, each cell is considered to be emitting a spherical wavefront with a specified phase retardation. These diffracted wavefronts interfere in the reference plane to produce the required light distribution.

DOEs are usually fabricated from transparent silicon-based substrate. The multi-level relief structure on the surface is performed by multi-step lithographic process, including photo-lithography or direct e-beam writing, and etching. The size of one resolution cell reaches size up to 200-500 nm with 1-5 inch wafer plate. The power efficiency reaches above 90%.

Electronically reconfigurable DOE/SLM. Electronically reconfigurable SLMs are the most widely used optical modulation elements for light field synthesis applications. Their main advantage over fixed elements is the possibility to reconfigure them during run-time. Therefore, a single SLM can be used for many purposes in an application, often providing the possibility for real-time performance.

The available SLM technologies are primarily based on liquid crystals. The molecules of a liquid crystal may be oriented in a crystal-like way, where different orientation leads to different optical properties. In essence, liquid crystals change their refractive index depending on the molecule orientation, which can be controlled by applying an electrical charge.

There are two basic technologies for reconfigurable elements, differing in the way the electrical charge is applied to the liquid crystal. The first group of SLMs are optically addressed, while the second group of SLMs are electronically addressed by a silicon chip. The optically addressed SLMs use an incoherent light image to control the

orientation of the liquid crystal molecules. The intensity of the incoming incoherent light is sensed by a photo-sensor, which transforms the intensity into electrical charge directly over the liquid crystal. The resolution achieved by this method reaches 100 cycles/mm. The major disadvantage is the variable contrast and sensitivity across the device. In electronically addressed SLMs, the voltage to the liquid crystal is applied by an electrode on one side and a silicon chip on the other. The silicon chip is divided into pixel cells, to change the local electric field across the liquid layer and thus switch the pixels “on” or “off”. Amplitude modulation in each pixel can be performed by altering the time the pixel is in the “on” state. With this technology, modern commercially available SLMs have a pixel pitch size of $8-30\mu\text{m}$, with around 90% fill factor. The size of the matrices reaches 1920×1200 pixels.

Liquid crystal SLMs can be addressed to modulate light in three common configurations [120]. One of the configurations aims at high contrast in the amplitude at the expense of spiral dependency between the phase and the amplitude in the complex plane. The second configuration sacrifices contrast in order to modulate only the amplitude of the incident light wave. In fact, there is always a small phase modulation. The third configuration is nearly phase-only. There are commercially available SLMs which provide over 2π phase modulation with rather linear input/output response. A few attempts have been made to couple the phase with amplitude towards a larger domain of pixel values in the complex plane. Birch et al. use phase detour methods to codify a complex number by a group of neighboring pixels [121]. However, this comes at the expense of effective SLM resolution. Stolz et al. have studied in detail the properties of the coupled phase and amplitude domains using another complex number codification scheme [122]. The domain of pixel values in the complex plane is significantly enlarged by the use of more than one SLM at the expense of more sophisticated control [6, 120].

2.4.3 LF synthesis on a target plane

This subsection summarizes popular approaches to the synthesis of a desired complex light field distribution $d(x, y) = |d(x, y)|e^{j\phi_d(x, y)}$ on a reference plane $(x, y, z)|_{z=\text{const}}$. The reference plane can be considered as any plane from the volume of interest where the desired 3D LF distribution $D(x, y, z)$ should be generated. This means that the approaches considered here assume knowing the distribution $d(x, y)$, which, when propagated within the volume of interest, produces the desired 3D light distribution $D(x, y, z)$. The synthesis of $d(x, y)$ is done by complex wave front modulation elements, whose design and/or configuration is facilitated by computer based optimization algorithms.

Phase modulation elements are simple to design and configure, because they have a quite linear input/output characteristic and achieve resolution close to the light wavelength. Such resolution is sufficient for light field reconstruction and hence for direct phase modulation without limiting the degrees of freedom of the modulated wavefront. In addition, phase elements have high diffraction efficiency, i.e. the modulated wave front energy is close to the incident wave energy. Provided with a light wave which already satisfies the desired amplitude $|d(x, y)|$, a phase element placed

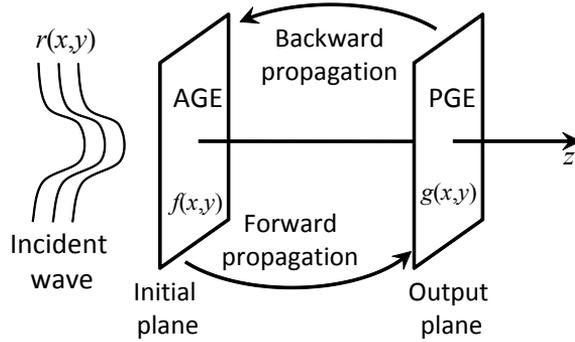


Figure 2.10: LF synthesis by two modulation elements: one to control the phase (PGE) and one to control the amplitude (AGE) of the incident wave.

in the output plane $(x, y, z)|_{z=\text{const}}$ can produce the desired phase $\phi_d(x, y)$. For sake of clarity, abbreviate this element as Phase Generation Element (PGE). The PGE leaves a remaining challenge in the generation of the desired amplitude $|d(x, y)|$. The desired amplitude $|d(x, y)|$ is usually produced by modulating an incident light wave $r(x, y) = |r(x, y)|e^{j\phi_r(x, y)}$ with another modulation element, abbreviated as Amplitude Generation Element (AGE), which is placed in the initial plane $(x, y, z)|_{z=0}$ (Fig. 2.10). Due to technological limitations, the available laser sources often produce the incident light wave $r(x, y)$ as a Gaussian beam, even though it is possible to use any light field distribution in theory. If AGE is also a phase modulation element, then its transmittance $e^{j\phi_A(x, y)}$ must be optimized to generate right before the PGE a wave front $g(x, y)$ at the output plane $(x, y, z)|_{z=\text{const}}$, whose amplitude approximates to the best extent the desired amplitude $|d(x, y)|$. The wave front $g(x, y)$ is the result of modulating the incident wave by the phase only transmittance $e^{j\phi_A(x, y)}$ and propagation on distance z to the plane $(x, y, z)|_{z=\text{const}}$:

$$g(x, y) = \int |r(x', y')| e^{j(\phi_r(x', y') + \phi_A(x', y'))} h_z(x, y; x', y') dx' dy'. \quad (2.82)$$

The propagation response $h_z(x, y; x', y')$ corresponds to the impulse response of the R-S integral in Eq. 2.8, the Fresnel impulse response in Eq. 2.18, or the Fourier transform kernel of the Fraunhofer approximation [4]. In practice, the long propagation distance for the Fraunhofer approximation is shortened by an appropriate lens-based architecture [12, 13, 14]. Some authors even use the fractional Fourier transform for propagation, implemented also by appropriate lens configuration [12].

A direct method for synthesis of the desired amplitude $|d(x, y)|$ decomposes it into elementary geometrical objects: points, lines, circles, discs, rectangles elliptic disks [11]. A phase modulation pattern is calculated analytically for each type of elementary object, depending on the object's parameters. The desired amplitude $|d(x, y)|$ is synthesized by superposing the phase patterns of the elementary objects which constitute $|d(x, y)|$. This superposition mixture is directly generated by the phase element. An evident shortcoming of this approach is the necessity of decomposing $|d(x, y)|$ into elementary objects, which might be a complex optimization itself. In

addition, the authors report that the diffraction efficiency is difficult to estimate for a certain object.

One of the first, yet widely used, algorithms for optimization of the AGE transmittance $e^{j\phi_r(x,y)}$ subject to a desired output magnitude $|d(x,y)|$ is the *Gerchberg-Saxton algorithm* [15, 17, 102]. In essence, it iterates between the initial plane $(x,y,z)|_{z=0}$ and the output plane $(x,y,z)|_{z=\text{const}}$ by numerical propagation and enforces the incident wave $r(x,y)$ and desired amplitude $|d(x,y)|$ restrictions, respectively. That is, each iteration modifies the propagated $g(x,y)$ to have the desired amplitude as $\hat{g}(x,y) = |d(x,y)|e^{j\phi_g(x,y)}$. The modified $\hat{g}(x,y)$ is back-propagated to the initial plane as $f(x,y) = \int \hat{g}(x',y')h_{-z}(x,y;x',y')dx'dy'$ where the incident wave is enforced as $\hat{f}(x,y) = |r(x,y)|e^{j\phi_f(x,y)}$. Under this enforcement, the phase $\phi_f(x,y)$ is preserved by selecting the phase of the AGE as $\phi_A(x,y) = \phi_f(x,y) - \phi_r(x,y)$. The basic idea of the iterations is to preserve the phase correspondence between the LF distributions on the initial and output planes, while enforcing their amplitudes.

Note that this approach resembles the serial POCS method, which was described in Subsection 2.2.4.4. In this case there are two sets C_r and C_d associated with the incident wave $r(x,y)$ and the desired output LF distribution $d(x,y)$, respectively. The sets can be defined formally as:

$$\begin{aligned} C_r &= \{\forall f \in DF_0 : |f(x,y)| = |r(x,y)|\} \text{ and} \\ C_d &= \{\forall g \in DF_z : |g(x,y)| = |d(x,y)|\}. \end{aligned} \quad (2.83)$$

The space DF_z can be defined as the space of the two dimensional functions which are diffraction fields at a transversal plane $(x,y,z)|_{z=\text{const}}$. The set C_r consists of all diffraction fields at the initial plane $(x,y,z)|_{z=0}$ whose amplitude is restricted to the amplitude $|r(x,y)|$ of the reference wave. Note that the phase of the $\phi_r(x,y)$ does not impose a further restriction, because the AGE transmittance adds a voluntary phase $\phi_A(x,y)$ to $\phi_r(x,y)$. The set C_d consists of all diffraction fields at the output plane $(x,y,z)|_{z=\text{const}}$ whose amplitude is restricted to the desired amplitude $|d(x,y)|$. The projection onto each set is done by forward or backward propagation and amplitude enforcement. Note that the sets defined as in Eq. 2.83 are not convex. A simple calculation shows that a set C_r does not include every convex combination $\alpha|r|e^{j\phi_1} + (1-\alpha)|r|e^{j\phi_2} = |r|(\alpha e^{j\phi_1} + (1-\alpha)e^{j\phi_2}) \neq |r|e^{j\phi}$, $0 < \alpha < 1$ of two elements $|r|e^{j\phi_1}$ and $|r|e^{j\phi_2}$ which belong to C_r . Fortunately, there are only two sets and the convergence of the algorithm is assured, even though the final solution might be a local minimum [48]. The propagation in the projections can be implemented according to either R-S integral, Fresnel integral, Fraunhofer approximation, or fractional Fourier transform. All of these have frequency domain alternatives which lead to efficient FFT-based numerical implementations, as discussed in Section 2.1. Different propagation operators may contribute to improving the convergence and the diffraction efficiency [12, 13, 48].

Fienup [123] shows that the Gerchberg-Saxton algorithm and some of its modifications are equivalent to fast implementations of gradient search methods. As such, the Gerchberg-Saxton algorithm is sensitive to the starting point and is often trapped to local minima. Moreover, the technology of the available modulation element might not always have linear input/output characteristic, good resolution and phase-only mode.

For example, a cheap DOE has low resolution and coarse step-wise input/output characteristic, with only a few available levels. Hence, the possible states of the modulated output wave are discrete rather than continuous. The mirrors of the DMD, considered in publication **P1**, take one of a few pre-defined angles and its output does not directly modulate the phase or amplitude of incident wave front. Yet another example is a deformable mirror with a continuous surface which is controlled by a discrete set of actuators [124]. In this case the set of possible output LF distributions is discrete rather than continuous. The LF distribution which approximates best $d(x, y)$ is produced by appropriately configuring the modulation element. The search for best configuration is an NP-hard combinatorial problem and requires constrained optimization algorithm. Many constrained optimization algorithms can act as global optimization algorithms when the set of possible solutions is continuous.

Genetic algorithms are a class of global/constrained optimization algorithms which resemble the mechanisms of nature survival [125]. A genetic algorithm operates on a population of potential solutions iteratively and optimizes by the principle of survival of the fittest, according the problem specifics. At each iteration, a new generation of potential solutions is produced from the fittest solutions in the current generation. The reproduction is based on crossover (random information exchange) between two solutions in the current generation and mutation (random information permutation) of their offspring. These are the operations which allow the algorithm to explore new regions where better solutions may lie. The crossover and mutation are done with a certain probability, which decays for later iterations. The performance of a genetic algorithm in a beam shaping application was demonstrated by controlling a deformable mirror [124]. A common problem with the genetic algorithms and other global algorithms is the slow convergence rate and excessive computations for evaluation of the fitness of all candidate solutions. A modification which speeds up the convergence is the hybrid global/local approach, suggested for the design of a phase-only DOE in beam shaping [14]. The basic idea is to optimize each candidate solution in the new generation locally before the next fitness evaluation and selection for reproduction.

2.4.4 LF reconstruction inside a volume

This subsection summarizes approaches which design and configure light wave modulation elements directly from the desired 3D light distribution $D(x, y, z)$ specified within the volume of interest. Some of these approaches concentrate only on the reconstruction of a complex light field distribution $d(x, y)$ on a reference plane $(x, y, z)|_{z=\text{const}}$ which, when propagated within the volume of interest, would produce $D(x, y, z)$. Provided $d(x, y)$ one can use any of the approaches described in 2.4.3, or similar, to design and configure the light modulation element.

2.4.4.1 Physical limitations of the light fields

Depending on the application at hand, the desired 3D LF distribution $D(x, y, z)$ is not guaranteed to be physically consistent. Therefore, it is important to summarize the limitations which determine a physically consistent light field. A scalar light

field $u(x, y, z)$ is a complex-valued function which satisfies the Helmholtz Eq. 2.2. Consequently, the 3D Fourier transform of $u(x, y, z)$ is supported on the Ewald's sphere $k_x^2 + k_y^2 + k_z^2 = k^2$ with radius $k = \frac{2\pi}{\lambda}$. In addition, only part of this sphere may be occupied due to the resolution limits of eventual modulation system. Often, the part of the Ewald's sphere for frequencies $|k_x| \leq \Omega_x < k$, $|k_y| \leq \Omega_y < k$ and $\sqrt{k^2 - \Omega_x^2 - \Omega_y^2} \leq |k_z| \leq k$ is sufficient for the application at hand and used for the computations. In the spatial domain, the transversal extent of $u(x, y, z)$ is also limited by the finite aperture of the modulation device. To be consistent with the Fourier transform theory, $u(x, y, z)$ is assumed to be essentially limited in the frequency domain and essentially limited along any transversal plane $(x, y, z)|_{z=\text{const}}$. Another restriction follows from the PWD integral of Eq. 2.13. The Fourier transform of the field restricted to the transversal plane $(x, y, z)|_{z=\text{const}}$ is $a_0(k_x, k_y)e^{jz\sqrt{k^2 - k_x^2 - k_y^2}}$. Therefore, the field $u(x, y, z = \text{const})$ and $a_0(k_x, k_y)e^{jz\sqrt{k^2 - k_x^2 - k_y^2}}$ obey Parseval's energy identity:

$$\begin{aligned} \iint |u(x, y, z = \text{const})|^2 dx dy &= \iint |a_0(k_x, k_y)e^{jz\sqrt{k^2 - k_x^2 - k_y^2}}|^2 dk_x dk_y \\ &= \iint |a_0(k_x, k_y)|^2 dk_x dk_y. \end{aligned} \quad (2.84)$$

Hence, the energy of the field on any transversal plane must be the same.

In addition, one may impose constraints associated with the application at hand if these can be easily placed inside the optimization procedure. Such constraints may be useful to narrow down the search space for the reconstructed field. For example, non-diffracting beams have plane wave components which have the same k_z [24]. This effectively selects a ring on Ewald's sphere, where the 3D Fourier transform of $u(x, y, z)$ is non-zero. A generalization of the non-diffracting beams are self-imaging beams, which consist of more than one ring in the Fourier domain. Under certain conditions, the separate non-diffracting beams which correspond to the separate rings go in and out of phase periodically along z . This creates the periodicity of the pattern along z .

2.4.4.2 Related work

There are approaches to light field reconstruction and synthesis performed by a direct calculation, without the need of an optimization procedure. Such methods often limit critically the set of possible light fields which can be reconstructed by inherent assumptions and crude approximations. Rosen and Yariv demonstrated the reconstruction of arbitrary paraxial focal lines with Fraunhofer diffraction [18]. A desired focal line is designed as a superposition of straight line segments, each of which corresponds to a radial harmonic function on the reference plane. Two other works also approximate a Fourier hologram from the data of multiple-angle planar recordings [126, 127]. Shabtay et al. use the method of calculus of variations to reconstruct a light field from a desired 3D LF distribution within the volume of interest [20]. A drawback of this method is that the desired shape must be integrated along z , either numerically or analytically.

Most widely used approaches to light field reconstruction are iterative and are based on generalized projection methods [50]. These methods differ in the way they specify the desired 3D LF distribution within the volume of interest, the way they form the sets and the way the projections are implemented.

A straightforward light field reconstruction approach defines the set of physically consistent light fields as all 3D functions $u(x, y, z)$, whose 3D Fourier transform is non-zero only on Ewald's sphere $k_x^2 + k_y^2 + k_z^2 = k^2$ with radius $k = \frac{2\pi}{\lambda}$. The projection of any 3D function $f(x, y, z)$ on this set is done by computing its 3D Fourier transform and setting all the frequency components to zero except those which lie on Ewald's sphere. The desired 3D LF distribution within the volume of interest constrains another set of all possible functions which have specific shape within the volume of interest. The projection of a light field onto this set would take inverse 3D Fourier transform from Ewald's sphere and restrict the desired values within the volume of interest. This rather elegant LF reconstruction approach suffers some practical inconveniences. If the function inside the volume of interest is specified on a uniform rectangular lattice, it also transforms to a uniform rectangular lattice in the Fourier domain. Such a lattice can not host precisely a spherical surface where the 3D Fourier transform of the LF must be non-zero. The original paper for this approach does not mention or imply any sampling strategy inside the volume of interest [21]. Moreover, the 3D FFT assumes that the light field is periodic along the coordinates x , y and z . Finite aperture can always be assumed and the field can be effectively periodized along x and y . However, periodicity along z is a critical restriction which limits the set of reconstructible light fields.

A few related works try to reconstruct a light field from a set of CCD intensity measurements [31, 32, 33]. Upon illumination, the object produces a light field whose intensity is captured by a CCD sensor on a few parallel planes [31, 32], sometimes by two or more CCD images per plane [33]. The light field is reconstructed using an algorithm based on serial POCS method. A set is defined for each CCD measurement plane as all Fresnel light fields whose intensity is measured by the CCD. Hence, the sets are not anymore convex. The projection is made by propagation with the Fresnel transform to the plane of the measurement, where the measured intensity is forced and the result is back-propagated to the initial plane. As the sets are not convex, the convergence of the method is not theoretically guaranteed. In practice, the convergence depends on the precise alignment of the CCD planes and rapid convergence is achieved by adding more CCD measurements. This requires long measurement time, which is inconvenient in many applications. Short measurement times are achieved by putting a phase modulation SLM in the Fourier domain of the imaging system [34, 35]. The SLM is used to simulate the transfer function of propagation and thus the CCD position does not need to be changed and aligned for each measurement.

This approach and many others fall within a general framework for light field reconstruction and synthesis, formalized by Piestun and Shamir [3] and used for various beam shaping and 3D display problems [22, 23, 24, 49, 50, 128, 129, 130]. The aim is to find the light field $u(x, y, z)$ which lies closest to a desired 3D LF distribution $D(x, y, z)$ within the volume of interest, and, eventually satisfies other desired constraints associated with the field properties. The overall procedure can be outlined as follows:

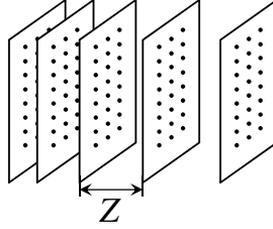


Figure 2.11: Sampling grid within the volume of interest for the general iterative LF reconstruction approach.

1. Define a volume of interest and desired light field distribution $D(x, y, z)$ inside. If needed, $D(x, y, z)$ contains only magnitude.
2. Define tolerance distribution $\eta(x, y, z)$ for the desired field values
3. Transform $D(x, y, z) \pm \eta(x, y, z)$ to mathematical constraints. This is done by the volumetric sampling strategy described in Subsection 2.3.4. In effect, this strategy leads to an ensemble of parallel planes, with a uniform rectangular sampling grid at each plane (Fig. 2.11). The sampling steps along x and y are the same for each plane and are determined according to Eq. 2.74, with the smallest z used inside the volume of interest. The distance Z between two consecutive planes is selected according to Eq. 2.80. On such sampling grid, one constraint set C_l is defined as all possible light fields which have the specified values on the grid at a certain plane.
4. Define constraints for the technological limitations of the SLM/DOE if a direct, one-stage LF synthesis is targeted.
5. Run an iterative optimization procedure to solve the problem:

$$\hat{u}(x, y, z) = \arg \min_{u \in C_l} E_w(u(x, y, z), D(x, y, z)), \quad (2.85)$$

where $E_w(\cdot, \cdot)$ is a weighted distance measure between $u(x, y, z)$ and $D(x, y, z)$ with different weights w_l for each constraint. This sets a hierarchy among the constraints, relaxing the important ones more.

In general, one could use any optimization procedure to solve Eq. 2.85. Global optimization procedures, e.g. genetic algorithms or simulated annealing, are avoided as computationally intensive. The least squares methods are often inconvenient due to the non-linearity, non-convexity and non-intersecting constraints. Serial or parallel POCS, described in Subsection 2.2.4.4 are preferred as easy to implement. The projection of the function on the reference plane $d(x, y)$ onto the set C_l , defined on the grid of the l -th plane $(x, y, z)|_{z=z_l}$, can be formally defined as:

$$P_l\{d(x, y)\} = BP_{z_l}\{FP_{z_l}\{d(x, y)\}(nX, mY) := D(nX, mY, z_l)\}, \quad (2.86)$$

where FP_{z_l} and BP_{z_l} denote forward and backward propagation on distance z_l , respectively, according to R-S, PWD, Fresnel or Fraunhofer diffraction. The set of points

$\{(nX, mY)\}_{m,n \in \mathbb{Z}}$ define the uniform rectangular grid on the plane $(x, y, z)|_{z=z_l}$ with sampling steps X and Y along the coordinates x and y , respectively. The parallel POCS implementation directly relates the weights of the sets to the weights w_l of the distance function $E_w(\cdot, \cdot)$. Also different weights can be used for the separate iterations. It is possible to use hybrid methods which perform serial projections onto a few sets in parallel with the projections onto the rest of the sets. Also in the parallel mode different weights can be used for the separate iterations.

2.4.5 Conclusions

This section has discussed all matters relevant to the topic of the thesis – LF reconstruction and synthesis. The section start with a brief introduction to the problem, motivated by the potential application areas which have been summarized in Subsection 2.4.1. Subsection 2.4.2 outlined the basic technical constraints which the fixed or electronically reconfigurable SLMs impose on the range of synthesized light fields. These constraints play a leading role for the derivation of any of the algorithms for LF synthesis on a target output plane which have been summarized in Subsection 2.4.3. Yet these methods are derived based on the assumption that the desired complex LF distribution on the plane is available. As discussed in Subsection 2.4.1, such an assumption is not valid in many applications. Therefore, Subsection 2.4.3 presents the current state-of-the-art in LF reconstruction based on volume specifications. The technical discussion of these methods serves as a basis of comparison with the methods, developed in this thesis.

Chapter 3

Methods for LF reconstruction and synthesis

This chapter summarizes the main results and contributions of the publications which constitute this thesis. Section 3.1 identifies the current issues in the LF synthesis and reconstruction literature and summarizes the hypotheses, followed by the publications. The section underlines the gap filled by the contributions in the publications, and serves as a starting point to find practical situations which would benefit from the designed methods. Section 3.2 presents the relations and the essence of the methods for LF reconstruction from a set irregularly distributed samples which appear in publications **P2**, **P3**, **P4**, **P5**, **P6**, **P7**, **P8**. Section 3.3 summarizes the methods for light field synthesis with DMD from publication **P1**.

3.1 Problem formulation and hypotheses

The possibility to control light is attractive in various industrial, measurement and 3D imaging applications. Many of these require the generation of a desired application-oriented light field within a volume of interest. The situation is formally depicted in Fig. 3.1. An incident light wave is modulated by an optical system to produce a light field $u(x, y, z)$ within the volume of interest. Optimally, the produced light field $u(x, y, z)$ follows the desired 3D scene, as described by the function $D(x, y, z)$. In theory, light fields are described as complex scalar functions $u(x, y, z)$ in three dimensions (x, y, z) . Yet, the two-dimensional complex function $d(x, y) = u(x, y, z)|_{z=\text{const}}$, which describes the field on any transversal “reference” plane, carries sufficient information about the field behavior in the whole 3D space. The depth information is carried by the phase $\phi_d(x, y) = \arg d(x, y)$ of such a two-dimensional function.

Due to this fact, many scientists search for methods to synthesize a complex wave function $d(x, y)$ over a target reference plane which is transversal to the propagation

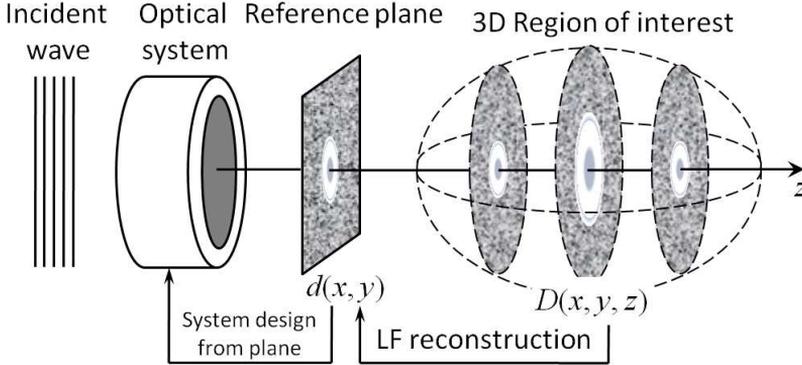


Figure 3.1: Two-step light field synthesis. LF synthesis can be based on LF reconstruction prior to simpler modulation system design.

direction. These methods assume that $d(x, y)$ is already available and concentrate on the design/configuration of the light wave modulation element, as surveyed in Subsection 2.4.3. However, many practical situations specify the desired scene $D(x, y, z)$ in the volume of interest, rather than a plane. Therefore, “planar” light field synthesis methods are not directly applicable in such situations.

As discussed in Subsection 2.4.4, the depth information of the scene $D(x, y, z)$ is commonly specified on multiple planes, at a regular grid of samples per plane. In some cases the planes are formed by multiple CCD recordings. For applications where such CCD recordings are not possible/available, the desired data points are specified on the regular grid shown in Fig. 2.11. This grid was derived based on the considerations illustrated in Fig. 2.8, which define a non-uniform, but fixed-position grid. To the best of our knowledge, no other sampling strategy within a volume was discussed in the literature. Note that both grids require an amount of samples which exceeds the NDF of the reconstructed light field by a very large factor. Moreover, the sample positions defined in any of these grids are not convenient to host data points, whose position is fixed by an application. A typical example is CGH, where the data points of the desired scene are the vertices of a 3D mesh or a point cloud, distributed in a very irregular manner.

Finally, the alternative to design/configuration of the light wave modulation element directly from a volume light field specification is more complicated than the planar case. An eventual method has to deal both with the limitations of the modulation device and the desired constraints on the light field over the whole volume of interest. Such methods exist, but often they are limited to specific area of application.

The publications in the thesis follow a rather general, two-stage approach to synthesize a LF inside a volume. This approach decouples the *design of the light wave modulation element* from the *reconstruction of a physically consistent light field*. By physically consistent light field is meant any wave which satisfies the Helmholtz wave equation and other restrictions such as essentially finite bandwidth and/or spatial extent, etc. At first, the desired 3D specification $D(x, y, z)$ is used for the reconstruction of the

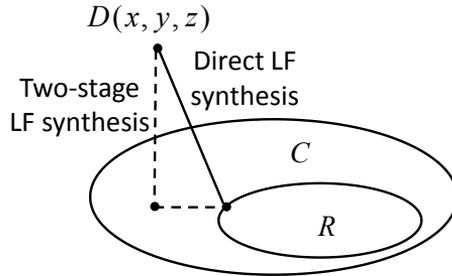


Figure 3.2: LF sets for reconstruction and synthesis. The desired light field $D(x, y, z)$ can be directly projected to the set R of fields reproducible by a DOE/SLM, or, first to the set C of physically consistent light fields and then to R .

light field distribution on a reference plane, its Fourier spectrum, or a more general abstract light field representation. At a second stage, this light field representation is used as input for the design and configuration of the modulation element.

The benefit of decoupling the LF synthesis from the LF reconstruction is two-fold. The desired 3D scene $D(x, y, z)$ can be described efficiently without much redundancy compared to the NDF of the reconstructed light field, as explained in Hypothesis 1. This fact suggests better numerical efficiency of the LF reconstruction methods. The second benefit of the two-stage LF synthesis approach is that it is more universal and technology-independent. The information about the reconstructed physically consistent light field can be stored and later produced under the available technology and conditions. Moreover, different LF reconstruction methods can be combined with the same light wave modulation element.

The eventual accuracy and complexity of the two-stage LF synthesis approach is discussed at a higher level of abstraction with the help of Fig. 3.1. The set of light fields which can be reproduced with a certain DOE/SLM is denoted by R and the set of all physically consistent light fields is denoted by C . The desired 3D scene $D(x, y, z)$ is not always physically consistent and lies outside the set C in the general case. Any light field which is reproduced by the DOE/SLM is physically consistent at the same time and hence R is a subset of C . Therefore, the two-stage LF synthesis approach yields the same light field as a single-stage approach if a global optimization procedure is used in both approaches. Hence, the point in R which lies at minimal distance from the desired $D(x, y, z)$ can be found by first finding the closest point on C , which is then projected onto R . Often, the optimization procedures used to find a minimal distance projection is local rather than global. In such a case, the direct and the two-stage LF synthesis produce different results and a general statement of which result is preferable cannot be made. Moreover, it is also arguable whether the direct or the two-stage approach uses less computational resources. The direct approach uses a single optimization procedure and intuitively should be faster. However, it works under harder constraints than the two optimization procedures of the two-stage approach and thus may converge more slowly.

Hypothesis 1. The reconstruction of a physically consistent light wave which best approximates the desired 3D scene $D(x, y, z)$ should be formalized in a manner which

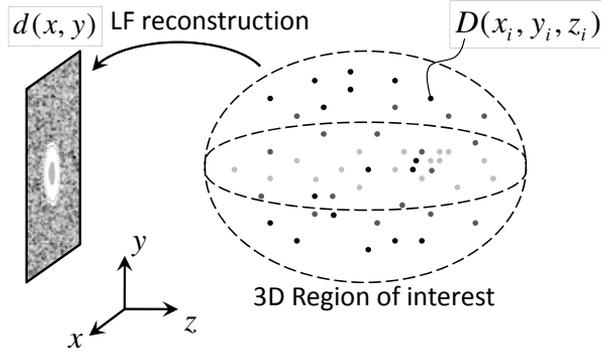


Figure 3.3: Irregular sampling of a LF. A desired 3D scene $D(x, y, z)$, specified on a set of irregularly distributed points $\{(x_i, y_i, z_i)\}_{i=1}^s$, is used to reconstruct a light field distribution $d(x, y)$ on a reference plane.

suits the specifics of any application. The core content of the thesis consists of methods for LF reconstruction which follow the *irregular sampling* scenario. A desired scene $D(x, y, z)$ is specified through complex values $D(x_i, y_i, z_i), i = 1 \dots s$ on a set of points $\{(x_i, y_i, z_i)\}_{i=1}^s$, irregularly distributed within a volume of interest (Fig. 3.3). The LF reconstruction problem is stated as to *find a representation which carries sufficient information about the physically consistent light field $u(x, y, z)$ from the desired complex values $\{D(x_i, y_i, z_i)\}_{i=1}^s$* . Such a representation can be the complex LF distribution $d(x, y) = u(x, y, z)|_{z=\text{const}}$ on the reference plane, or its Fourier transform, or the parameters of a more abstract light field model. The set of values $\{D(x_i, y_i, z_i)\}_{i=1}^s$ is considered to carry *a small amount of redundant information*, i.e. the amount s of samples is slightly greater, but comparable to NDF of the reconstructed light field. The values $\{D(x_i, y_i, z_i)\}_{i=1}^s$ may specify a light beam shape, a desired computer generated object or recorded LF values by some sensing device. For beam shaping applications, the designer has the freedom of choosing the point locations. The point locations can be chosen in a manner which ensures the stability of the reconstruction and at the same time is sufficient to specify some desired properties of the reconstructed light field, such as resolution, aperture, etc. Moreover, some applications, e.g. computer-generated holography, deal with synthetic 3D scenes commonly described by a point cloud, mesh or any other geometrical model which specifies the positions of the sampling points $\{(x_i, y_i, z_i)\}_{i=1}^s$ itself. The values $\{D(x_i, y_i, z_i)\}_{i=1}^s$ may also represent recorded data of the light field of an illuminated object, captured by e.g. a holographic technique.

Hypothesis 2. *The light field synthesis of the reconstructed physically consistent light field $d(x, y) = |d(x, y)|e^{j\phi_d(x, y)}$ can be done separately by any available light wave modulation element.* Naturally, the extent to which $d(x, y)$ on a reference plane is approximated depends on the technological limitations of the modulation element itself. Publication **P1** *optimizes the configuration of a DMD to reproduce $d(x, y)$ when illuminated by a plane wave.* The DMD consists of a 2D array of square micro-mirrors, each of which can be deflected into one of a few pre-defined angles. The DMDs are not a standard device for 3D LF synthesis and their use demonstrates the range and flexibility of the two-stage approach to LF synthesis, represented in Fig. 3.1.

3.2 LF reconstruction from samples on irregular grids

Publications **P2**, **P3**, **P4**, **P5**, **P6**, **P7** and **P8** address the problem of light field reconstruction. The problem is cast as reconstruction from a set of field samples, irregularly distributed within the region of interest. Such a consideration facilitates the use of many results derived in the irregular sampling theory of conventional 1D and 2D functions. Note that a light field is not a conventional 3D function since it is completely described by the 2D complex function on a reference plane.

Formally stated, the problem is to reconstruct a physically consistent light field $u(x, y, z)$, given the set of desired complex values $\{D(x_i, y_i, z_i)\}_{i=1}^s$. The form of the reconstruction can be the complex LF function $d(x, y) = u(x, y, z)|_{z=\text{const}}$ on the reference plane, or its Fourier transform, or the parameters of a more abstract light field model. The desired values $D(x_i, y_i, z_i), i = 1 \dots s$ are irregularly distributed within some volume of interest and their amount s should be slightly exceeding, but of the same order of the NDF of, the reconstructed light field.

The methods for LF reconstruction developed in publications **P2**, **P3**, **P5**, **P6**, **P7** and **P8** simplify the consideration to one transverse dimension x only, keeping also the longitudinal dimension z , as the one which enables the depth specifications. In this context, the problem is formulated so as to reconstruct a field restricted to the reference line $d(x)$, provided a set of irregularly distributed desired field samples $\{D(x_i, z_i)\}_{i=1}^s$. Generalizations from 2D consideration over (x, y) to 3D consideration over (x, y, z) is often straightforward. Comparison of the two basic reconstruction methods in 3D is considered in publication **P4**.

Two basic approaches for LF reconstruction from a set of irregularly distributed samples have been developed. One of them is based on the serial POCS method. The other one is based on the method of least squares approximations and is computationally implemented by the method of conjugate gradients. The LS-based approach has a good theoretical background in irregular sampling theory, which relates the stability and speed of the reconstruction to the properties of the sample distribution.

3.2.1 POCS-based methods

Publications **P3**, **P4** and **P6**, attached in this thesis, discuss POCS-based methods for light field reconstruction from irregularly distributed data points. POCS has been applied to various problems in holography and image restoration, where a priori information is used to constrain the size of the feasible solution set [12, 15, 17, 102]. Moreover, it has been applied for signal reconstruction from irregularly distributed samples [103, 104]. Therefore, it is anticipated that POCS can be applied in the considered LF reconstruction problem. The basic form of the POCS-based reconstruction algorithm is developed in **P3** for the 2D case and the generalization to 3D is presented in **P4**. Publication **P6** generalizes the basic 2D POCS in a manner whereby the irregularly distributed samples can have free positions and discusses its convergence.

3.2.1.1 Basic POCS method

POCS is a computational approach to find an element of a feasible region defined by the intersection of a number of convex sets, starting with an arbitrary infeasible point. Usually, the sets are defined by some constraints related to the application. The constraints in the considered LF reconstruction problem are the desired irregularly distributed samples $\{D(x_i, z_i)\}_{i=1}^s$. Here it is convenient to define a set $C_l, l = 1, \dots, L$ related to the samples $\{D(x_i, z_i)\}_{i:z_i=z_l}$ which lie on a transversal line $(x, z)|_{z=z_l}$. The set is defined as all diffraction fields $u(x, z)$ which have the desired samples $\{D(x_i, z_i)\}_{i:z_i=z_l}$:

$$C_l = \{\forall u(x, z) \in DF : u(x_i, z_l) = D(x_i, z_l), \forall i : z_i = z_l\}, l = 1, \dots, L. \quad (3.1)$$

The notation DF in Eq. 3.1 is used for the space of diffraction fields, i.e. all functions $u(x, z)$ which satisfy the Helmholtz wave equation. The set C_l is related to a transversal line in order to have computationally efficient projections from one set to another. The projection P_l of a function $f(x, z)$ onto the set C_l is done by substituting the known data points $\{D(x_i, z_i)\}_{i:z_i=z_l}$ on the line $(x, z)|_{z=z_l}$ and propagation to ensure that the resultant function lies in the space DF :

$$P_l f(x, z) = \mathcal{F}^{-1} \left\{ e^{-jz_l \sqrt{k^2 - k_x^2}} \mathcal{F} \left\{ f(x, z_l) + \sum_{i:z_i=z_l} (D(x_i, z_l) - f(x_i, z_l)) \delta(x - x_i, z - z_l) \right\} \right\}, l = 1, \dots, L. \quad (3.2)$$

The substitution with the known data samples in Eq. 3.2 is represented as just an addition of the difference $D - f$ at the positions $\{(x_i, z_i)\}_{i:z_i=z_l}$ of the samples which define the set C_l . The propagation model is the PWD implemented in the Fourier domain. Note that the resultant function after the resubstitution is back-propagated on distance $-z_l$ to the initial line $(x, z)|_{z=0}$. In general, the propagation can be made to any other transversal line $(x, z)|_{z=z_p}$, e.g. related to the next set C_p . In such a case the propagation kernel must be $e^{j(z_p - z_l) \sqrt{k^2 - k_x^2}}$ instead of $e^{-jz_l \sqrt{k^2 - k_x^2}}$. Propagation at any distance, in a backward or forward direction, which is done after substitution with the known data samples ensures that the resultant function is a diffraction field.

The Fourier domain propagation in Eq. 3.2 can be implemented efficiently by the FFT algorithm. The FFT acts on uniformly distributed data points in space/frequency. Therefore, the input and output lines of the propagation should be available on uniform sampling lattices, aligned to each other. The projections from the set C_{l-1} to the next set C_l use propagation from line $(x, z)|_{z=z_l}$ to the line $(x, z)|_{z=z_{l+1}}$ which defines the next set C_{l+1} . The projection onto the set C_L of the last line $(x, z)|_{z=z_L}$ propagates to the line $(x, z)|_{z=z_1}$ of the first set C_1 . Hence, passing through all sets requires having the propagation data points placed on each line on uniform lattices which are all aligned. Correspondingly, the resubstitution of the known data samples $\{D(x_i, z_i)\}_{i:z_i=z_l}$ for the projection onto C_l must be hosted by the uniform lattice on line $(x, z)|_{z=z_l}$ used for propagation. This arrangement of the known data points would be inconvenient for applications where their position cannot be selected during

the design process. For example, meshes describing objects used in computer generated holography have the vertices fixed to certain locations. In an attempt to tackle this situation, one might select very dense sampling lattices on all lines. In any case, the step of the lattices should be smaller or equal to the smallest non-zero difference between the x -locations of any two samples from the set $\{D(x_i, z_i)\}_{i=1}^s$. However, such an approach might increase the complexity of the computations much above the desired one, e.g. the one defined by the degrees of freedom in the desired field.

The POCS-based reconstruction algorithm presented here is generalized to 3D in **P4**. The generalization is straightforward. The convex sets $C_l, l = 1, \dots, L$ are defined based on the known samples $\{D(x_i, y_i, z_i)\}_{i:z_i=z_l}$ which lie in the same transversal plane $(x, y, z)|_{z=z_l}$ rather than a line. Then the projection P_l is done by resubstitution of these known samples on their locations at the plane $(x, y, z)|_{z=z_l}$ and propagation to the plane $(x, y, z)|_{z=z_{l+1}}$ which defines the next set C_{l+1} . The algorithm is completed again by iteration from set to set, passing through all the sets. It still has the drawback that all known data samples $\{D(x_i, y_i, z_i)\}_{i=1}^s$ should be hosted on the uniform sampling lattices used for propagation.

3.2.1.2 Generalization to free sample positions

Publication **P6** modifies the basic POCS-based algorithm in a manner that the given desired samples $\{D(x_i, z_i)\}_{i=1}^s$ may have free positions and do not need to be hosted by any uniform lattices. However, the propagation used in the projections benefits from efficient computations based on the FFT. The modification is done by defining the sets and the projections in a slightly different form.

The projections in the modification do not propagate from one line to another, but rather propagate the *new information* from each particular convex set $C_l, l = 1 \dots L$ to the reconstructed function $d(x) = u(x, 0)$ on the initial line $(x, z)|_{z=0}$. If the propagation is done by FFT in a straightforward manner, it still requires uniform lattices on the input $(x, z)|_{z=z_l}$ and output $(x, z)|_{z=0}$ lines which host the given data samples $\{D(x_i, z_i)\}_{i:z_i=z_l}$. However, one might define the set C_l as all known samples $u(x_i, z_i)$ with $z_i = z_l$, and which can be hosted by an N -point uniform lattice \mathbf{x}_l with a fixed step. Note that more set(s) C_m must be defined for the same line if the lattice \mathbf{x}_l does not host all samples with $z_i = z_l$. The sets C_m correspond to the same distance $z = z_l$, but are defined by different lattices \mathbf{x}_l and \mathbf{x}_m . All lattices have the same sampling step, but they are slightly displaced. In the general case, there might be no more than one sample per line which defines the lattice \mathbf{x}_l and the set C_l , respectively. In such a case, the number of convex sets L will coincide with the number of sampling points s and each projection will result in an update caused by a single point. However, if two or more points per line happen to belong to uniform grids, the algorithm should benefit from this. Again, in the general case each grid \mathbf{x}_l is not aligned to the grid at the initial line $(x, z)|_{z=0}$ and is assumed to be centered at some point $\chi_l \neq 0$ and encompassing a spatial interval of length T along x :

$$\mathbf{x}_l = \{(\mathbf{x}_l)_k\}_{k=0}^{N-1} = \chi_l + \{-T/2 + kT/N\}_{k=0}^{N-1}. \quad (3.3)$$

Now the formal definition of a set $C_l, l = 1 \dots L$ can be stated as:

$$\begin{aligned} C_l &= \{\forall u(x, z) \in DF : u(\mathbf{p}_l) = D(\mathbf{p}_l)\}, \\ \mathbf{p}_l &= \{(x_i, z_i), i = 1 \dots s : z_i = z_l, x_i \in \mathbf{x}_l\}, \end{aligned} \quad (3.4)$$

where \mathbf{p}_l is the set of points which are hosted by the grid \mathbf{x}_l and DF denotes the space of diffraction fields.

Now it suffices to define the projection P_l of a diffraction field $u(x, y)$ onto the set C_l . The given samples which define the set must be enforced on their positions \mathbf{p}_l and this information must be propagated back to the reconstructed function $d(x)$ on the initial line $(x, z)|_{z=0}$. However, the sampling lattice for $d(x)$ at the initial line $(x, z)|_{z=0}$ might not be aligned to all lattices $\mathbf{x}_l, l = 1, \dots, L$ simultaneously. Therefore, the new information from a set C_l is aligned to the origin by the use of the shift theorem for the Fourier transform [56]. An offset χ_l of the lattice is alternatively considered as shift of the function $u_{\chi_l}(x, z) = u(x + \chi_l, z)$. Hence, its Fourier transform along x is modulated by the factor $e^{jk_x \chi_l}$ and the lattice offset is compensated by the inverse factor $e^{-jk_x \chi_l}$. The projection is formally represented as:

$$\begin{aligned} P_l u(x, z) &= \mathcal{F}^{-1} \left\{ \mathcal{F} \{u(x, 0)\} + e^{-jk_x \chi_l} e^{-jz_l \sqrt{k^2 - k_x^2}} \times \right. \\ &\quad \left. \times \mathcal{F} \left\{ \sum_{(x_i, z_i) \in \mathbf{p}_l} (D(x_i, z_l) - u(x_i, z_l)) \delta(x - x_i, z - z_l) \right\} \right\}, \end{aligned} \quad (3.5)$$

where

$$d(x_i, z_l) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{u(x, 0)\} e^{jz_l \sqrt{k^2 - k_x^2}} \right\} \Big|_{x=x_i}. \quad (3.6)$$

This generalization of the POCS-based reconstruction algorithm can be outlined as follows:

1. initialize $\hat{d}_{0,0}$ arbitrary
 2. for $n = 1$ to n_{it}
 - for $l = 1$ to L
 - (a) use $\hat{u}_{n,l}(x, 0) = d_{n,l}(x)$ in Eq. 3.6 to predict the field $\hat{u}_{n,l}(x_i, z_l)$ at the points $\{\forall (x_i, z_i) \in \mathbf{p}_l\}$ of the given samples which define the set C_l
 - (b) calculate $\hat{d}_{n,l+1}(x) = \hat{u}_{n,l+1}(x, 0)$ by the projection $\hat{u}_{n,l+1} = P_l \hat{u}_{n,l}$ of the predicted field $\hat{u}_{n,l}(x, z)$ onto the set C_l .
- end

Publication **P6** derives the condition and rate of convergence of this version of the POCS algorithm. The derivation goes through an expression for the error $e_{n,l}$, obtained at the n -th iteration for the set C_l , in terms of the previous error $e_{n,l-1}$, obtained at the n -iteration for the previous set C_{l-1} . The error $e_{n,l}$ is measured by

the difference between the predicted field at the current step $\hat{u}_{n,l}(x, z)$ and the true underlying field $u(x, y)$, which must be reconstructed, and whose samples $\{u(x_i, z_i)\}_{i=1}^s$ approximate best the desired samples $\{D(x_i, z_i)\}_{i=1}^s$. The convergence of the algorithm is ensured if the norm of the error decreases after each projection, that is $\|e_{n,l}\| \leq \|e_{n,l-1}\|$. The derivation in **P6** shows that the norms of the consecutive errors are related through the norm of an operator, which incorporates the forward and backward propagation, lattice offset compensation and resubstitution of the known data samples done for each projection. The L_2 norm of this operator is:

$$\|u(x, z_l) - u(\mathbf{p}_l)\|_2^2 = \|u(x, z_l)\|_2^2 - \|u(\mathbf{p}_l)\|_2^2. \quad (3.7)$$

The norm of the the field restricted to the transversal line $(x, z)|_{z=z_l}$ can be normalized to unity before the calculations. Therefore, the norm of the difference $u(x, z_l) - u(\mathbf{p}_l)$ is always less than or equal to unity. Hence, the norm of the operator which relates the two consecutive errors $e_{n,l-1}$ and $e_{n,l}$ is less than unity and the algorithm converges.

3.2.2 Least squares-based methods

LS-based LF reconstruction methods use a finite-dimensional model which represents the scalar light field by a linear combination of generating functions, similarly to the frame decomposition in Eq. 2.29. Consequently, the LF reconstruction problem transforms to finding the parameters of the model, provided the desired data values $\{D(x_i, z_i)\}_{i=1}^s$. Respectively, the light field synthesis becomes defined as an inverse problem. The parameters of the decomposition from Eq. 2.29 are the coefficients c_n , provided that the waveforms φ_n are known.

This subsection starts with the presentation of the finite-dimensional models derived in **P2** and **P5**. The relation of the models to the number of degrees of freedom is also described, as in **P8**. The focus of this subsection is on the basic LS reconstruction approach, based on the CG method. Its convergence is discussed in terms of matrix singular values and frame bounds. These are the factors which define the stability of a particular sampling set. Upon reconstruction, an unstable sampling set leads to a boost of the inconsistency of the sample values. The Tikhonov approach, used in **P7** and **P8**, is discussed briefly as a regularized solution. Another elegant approach to precondition the CG method is to use the weighted LS as in **P6**, with adaptive weights which depend on the stability of the sampling set.

3.2.2.1 Finite-dimensional models

Fourier generators. The first model is based on discretization of the PWD integral. The model assumes that the field $u(x, 0)$ on the initial line $(x, z)|_{z=0}$ has essentially limited support in the spatial and frequency domain. Denote the spatial support of $u(x, 0)$ by Δ_x and the frequency support by 2Ω , as in Section 2.2 and Section 2.3. The function $u(x, 0)$ can be periodized along x since its spatial support is essentially limited. In the frequency domain, this periodization reflects in the discretization of

the Fourier transform $a(k_x)$ of $u(x, 0)$ at a uniform rate. Note that the period T of the periodization must be larger than the spatial extent Δ_x of $u(x, 0)$, i.e. $T \geq \Delta_x$. Eventually, $u(x, 0)$ is represented within a finite spatial support $x \in [-\frac{T}{2}, \frac{T}{2})$ by a discrete set of Fourier coefficients a_m . In essence, this periodization samples the PWD integral at points $k_x = \frac{2\pi m}{T}$, which leads to the discrete expansion:

$$u(x, z) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j \frac{2\pi}{T} \left(x m + z \sqrt{\frac{T^2}{\lambda^2} - m^2} \right)}. \quad (3.8)$$

This discrete, *Fourier generators*-based model describes the field $u(x, z)$ continuously at any spatial point as a superposition of the basis functions $\varphi_m(x, z) = e^{j \frac{2\pi}{T} \left(x m + z \sqrt{\frac{T^2}{\lambda^2} - m^2} \right)}$. These are a discrete set of plane waves, orthogonal to each other. Note that Eq. 3.8 includes a finite number of M basis elements. These correspond to the M discretized a_m of $a(k_x)$, taken at the rate $k_x = \frac{2\pi}{T}$, which fall within the support 2Ω of the Fourier spectrum $a(k_x)$. Therefore, M , Ω and T should obey the relation $M = \lfloor \frac{\Omega T}{\pi} \rfloor$.

The discrete model in Eq. 3.8 defines a functional space. It is spanned by the set of mutually orthogonal basis functions $\varphi_m(x, z) = e^{j \frac{2\pi}{T} \left(x m + z \sqrt{\frac{T^2}{\lambda^2} - m^2} \right)}$. Any function $u(x, z)$ which belongs to this space exhibits the properties of a scalar diffraction field within the limits of one transversal period T . The period T is also the parameter which determines the basis functions $\varphi_m(x, z)$ and thus the space spanned by them. In general, any function which belongs to the space spanned by $\{\varphi_m(x, z)\}_m$ with one value of T does not belong to the space spanned by $\{\varphi_m(x, z)\}_m$ with another value of T .

The generalization of the model in Eq. 3.8 to the 3D case is straightforward. Another transversal dimension y brings an extra term in the plane waves $e^{j \left(z \sqrt{k^2 - k_x^2} + x k_x + y k_y \right)}$ [4]. Now the 2D function $u(x, y, 0)$ must be assumed space-limited also along y , as discretization of k_y corresponds to periodization along y . The model of Eq. 3.8 is derived in **P2**, generalized to 3D in **P4** and discussed in **P6**, **P7** and **P8**.

Bessel-Fourier generators. Another expansion which expresses the 2D light field $u(x, z)$ in terms of a set of basis functions is derived in publication **P5**. As discussed in Subsection 2.1.1, the Fourier spectrum of a 3D light field $u(x, y, z)$ is supported on the Ewald's sphere with radius k , described by Eq. 2.5. The Ewald's sphere is a circle in two dimensions and the 2D Fourier transform $a(k_x, k_z)$ of $u(x, z)$ is non-zero only on this circle. Hence, a change of the spatial frequencies k_x and k_z to polar form $(k_x, k_z) = (k \sin \theta, k \cos \theta)$ represents the 2D Fourier transform $a(k_x, k_z)$ as the one-dimensional function:

$$a(k \sin \theta, k \cos \theta) = a \left(\frac{2\pi}{\lambda} \sin \theta, \frac{2\pi}{\lambda} \cos \theta \right) \equiv C(\theta), \quad (3.9)$$

where the rotation angle varies the interval $\theta \in [0, 2\pi)$. After this change of the variables, The 2D light field $u(x, z)$, expressed in terms of the inverse 2D Fourier

transform integral, becomes the one-dimensional integral:

$$u(x, z) = \frac{2\pi}{\lambda} \int_0^{2\pi} C(\theta) e^{j \frac{2\pi}{\lambda} (x \sin \theta + z \cos \theta)} d\theta. \quad (3.10)$$

A change of the spatial coordinates x and z in polar form as $x = r \sin \phi$ and $z = r \cos \phi$ simplifies the integral in Eq. 3.10 to

$$u(r, \phi) = \frac{2\pi}{\lambda} \int_0^{2\pi} C(\theta) e^{j \frac{2\pi}{\lambda} r \cos(\theta - \phi)} d\theta. \quad (3.11)$$

The function $C(\theta)$ is defined on the interval $\theta \in [0, 2\pi)$. It can be considered periodic outside this interval and, consequently, described by the complex Fourier series:

$$C(\theta) = \sum_{m=-\infty}^{\infty} c_m e^{jm\theta}. \quad (3.12)$$

Inserting Eq.3.12 into Eq.3.11, changing the order of summation and integration and making the substitution $\alpha = \theta - \phi$, the field can be expressed as:

$$u(r, \phi) = \frac{4\pi^2}{\lambda} \sum_{m=-\infty}^{\infty} c_m e^{jm(\phi + \frac{\pi}{2})} J_m \left(\frac{2\pi}{\lambda} r \right). \quad (3.13)$$

Here, $J_m(t)$ denotes the m -th Bessel function of the first kind [131]:

$$J_m(t) = \frac{1}{2\pi e^{jm\pi/2}} \int_0^{2\pi} e^{jm\alpha} e^{jt \cos \alpha} d\alpha.$$

More details on this derivation are presented in publication **P5**. The model from Eq. 3.13 describes the field continuously at any spatial point as a superposition of a discrete set of basis functions $\psi_m(r, \phi) = e^{jm(\phi + \frac{\pi}{2})} J_m(kr)$, $m \in \mathbb{Z}$, referred to as *Bessel-Fourier (B-F) generators*. Hence, the discrete set of scaling coefficients c_m describes completely the continuous field. However, no explicit discretization is done during the derivation. Finite dimensions in the summation are achieved by truncation to M nonzero coefficients $c_m \neq 0$, $m = -\lfloor \frac{M}{2} \rfloor, \dots, \lfloor \frac{M-1}{2} \rfloor$. However, such a truncation should be conformed with the characteristics of the function $C(\theta)$. The set of basis functions $\{\psi_m(r, \phi)\}_m$ is orthogonal and separable, as derived in **P5**. Therefore, they also span a functional space where each function of the space satisfies the Helmholtz wave equation. As such, the functions which belong to this space are scalar diffraction fields. However, the space spanned by a truncated set of M functions $\psi_m(r, \phi) = e^{jm(\phi + \frac{\pi}{2})} J_m(kr)$, $m = -\lfloor \frac{M}{2} \rfloor, \dots, \lfloor \frac{M-1}{2} \rfloor$ is a subspace of the space of all possible scalar diffraction fields.

Each $\psi_m(r, \phi)$ is a Bessel function along radial direction and exponentially oscillating along the angular direction. Generalization of this model to 3D uses spherical harmonics instead of cylindrical ones $e^{jm(\phi + \frac{\pi}{2})}$ and spherical Bessel functions of the first kind along the radial direction r [132].

3.2.2.2 Dimensionality of the models and NDF

The amount M of the non-zero weighting coefficients determines the dimensionality of any forward or inverse problem related to computation or reconstruction of a light field. Therefore, it is important to relate this amount to the physical properties of the approximating field. The NDF of the light field $u(x, z)$ is equal to the NDF of the generating function of the reference line $u(x, 0)$. In terms of the Wigner distribution, the NDF is the area in the time-frequency plane under which the Wigner distribution of the function is essentially nonzero [110]. If the shape of this area is nearly rectangular, the NDF can be measured alternatively by the space-bandwidth product of $u(x, 0)$ according to Eq. 2.66. The following discussion is derived in publication **P8**.

Fourier generators. The amount of required coefficients M_F for this model is directly related to the bandwidth 2Ω of $u(x, 0)$:

$$M_F \frac{2\pi}{T} = 2\Omega = \frac{2\pi \text{NDF}}{\Delta_x},$$

$$\text{hence } M_F = \left\lceil \text{NDF} \frac{T}{\Delta_x} \right\rceil. \quad (3.14)$$

Eq. 3.14 suggests that M_F is directly proportional to NDF. This result is expected as the Fourier generators discretize the frequency band of $u(x, 0)$. The proportionality coefficient $\frac{T}{\Delta_x}$ defines the excess which needs to be taken. The period T must be larger than Δ_x so that the periodic replicas of $u(x, 0)$ do not overlap. However, choosing T close to Δ_x does not guarantee that an overlap will not occur on a line at further distance z . Discrete frequency axis k_x defines a discrete Fourier spectrum of $u(x, z)$ at any line $(x, z)|_{z=\text{const}}$, keeping $u(x, z)$ periodic along x with the same period T . On the other hand, spatially limited pattern $u(x, 0)$ tends to spread its transversal support when propagated along z . Therefore, given maximal distance z_{max} according to the specified region of interest, one must ensure that $T > \Delta_x^{(z_{max})}$, where $\Delta_x^{(z_{max})}$ is the support of $u(x, z = \text{const} = z_{max})$. The relative increase of this spatial support $\Delta_x^{(z_{max})}/\Delta_x$ is smaller (closer to 1) for larger Δ_x at fixed distance z_{max} .

Bessel-Fourier generators. The coefficients c_m in the Bessel-Fourier model are the Fourier series coefficients of $C(\theta)$. Their amount M_B can be determined from the frequency support of $C(\theta)$. Direct comparison of Eq. 3.11 and Eq. 2.13, written for the 2D field $u(x, z)$, yields:

$$C(\theta) = a(k \sin \theta, k \cos \theta) = a(k \sin \theta). \quad (3.15)$$

Hence, the frequency support of $a(k \sin \theta)$ coincides with the frequency support of $C(\theta)$ and is used to estimate the number M_B .

The duality property of the Fourier transform $\mathcal{F}\{a(k_x)\} = u(-2\pi x, 0)$ can be used to obtain the frequency support of $a(k_x)$ as $\Delta_x/2\pi$. However, the frequency support of $a(k \sin \theta)$ with respect to θ is not the same, even though it is related to $\Delta_x/2\pi$. The highest frequency in the Fourier transform of $a(k_x)$ is $\Delta_x/4\pi$ – the same as the frequency of the harmonic $\cos(\frac{\Delta_x}{2} k_x)$. Therefore, the frequency support

of $\cos(\frac{\Delta_x}{2}k \sin \theta) = \cos(\frac{\pi \Delta_x}{\lambda} \sin \theta)$ can be used to estimate the frequency support of $a(k \sin \theta)$. In communication theory, such a harmonic function is recognized as a special case of frequency modulated signal $\cos(2\pi f_c \theta + \frac{f_\Delta}{f_m} \sin(2\pi f_m \theta))$. Its frequency support is approximated as $2(f_m + f_\Delta)$ according to Carson's rule [133]. The frequency support of $\cos(\frac{\pi \Delta_x}{\lambda} \sin \theta)$ and $a(k \sin \theta)$ and $C(\theta)$ respectively is estimated as $2(\frac{1}{2\pi} + \frac{\Delta_x}{2\lambda}) \approx \frac{\Delta_x}{\lambda}$. Hence, the dimensionality of the Bessel-Fourier model, required to describe a field with transversal support Δ_x on the reference line, is

$$M_B = \left\lfloor \frac{\Delta_x}{\lambda} \right\rfloor = \left\lfloor \text{NDF} \frac{\pi}{\Omega \lambda} \right\rfloor. \quad (3.16)$$

This result suggests that the Bessel-Fourier model requires a small amount of generators when the spatial support of the field is comparable to the wavelength. From another point of view, if NDF is assumed to be fixed, then the excess in M_B compared to NDF is determined by the ratio between π/Ω and λ . In fact, π/Ω is the size of the finest detail structure in $u(x, 0)$.

3.2.2.3 Discrete models, frames and sampling sets

The finite dimensional models in Eq. 3.8 and Eq. 3.13 represent the continuous-domain light field $u(x, z)$ (resp. $u(r, \phi)$) as a decomposition on a set of mutually orthogonal basis functions $\{\varphi_m(x, z)\}_m = \left\{ e^{j \frac{2\pi}{T} (xm + z \sqrt{\frac{T^2}{\lambda^2} - m^2})} \right\}_m$ (resp. $\{\psi_m(r, \phi)\}_m = \left\{ e^{jm(\phi + \frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda} r\right) \right\}_m$). Knowledge of the coefficients $\{a_m\}_m$ (resp. $\{c_m\}_m$) which weight the basis functions are sufficient to recover the continuous light field at any desired point. Each coefficient a_m (resp. c_m) can be found by an inner product between the field and the respective basis function φ_m (resp. ψ_m), because the basis is orthogonal.

In practice, any calculations related to these models must be carried out in a discrete fashion, on a sampling grid of representative points. Most probably, there exists a sampling grid for each model where the basis functions still maintain their mutual orthogonality after sampling. However, the primal concern of this thesis is to relax any sample point distribution $\{(x_i, z_i)\}_{i=1}^s$ defined by the set of points $\{D(x_i, z_i)\}_{i=1}^s$ desired to be reconstructed in any potential application. Often, the applications are not flexible with regard to the sample locations. The orthogonality of the sets $\{\varphi_m\}_m$ and $\{\psi_m\}_m$ might not remain after sampling on an arbitrary irregular grid $\{(x_i, z_i)\}_{i=1}^s$.

Eq. 3.8 and Eq. 3.13 can be written for each point (x_i, z_i) (resp. (r_i, ϕ_i)) with $(x_i, z_i) = (r_i \sin \phi_i, r_i \cos \phi_i)$ from the irregular grid $\{(x_i, z_i)\}_{i=1}^s$ defined by the desired sample values $\{D(x_i, z_i)\}_{i=1}^s$:

$$u(x_i, z_i) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j \frac{2\pi}{T} \sqrt{\frac{T^2}{\lambda^2} - m^2} z_i} e^{j \frac{2\pi}{T} m x_i}, \quad (3.17)$$

$$u(r_i, \phi_i) = \frac{4\pi^2}{\lambda} \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} c_m e^{jm(\phi_i + \frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda} r_i\right). \quad (3.18)$$

After irregular sampling, the sets $\{\varphi_m(x_i, z_i)\}_m = \left\{ e^{j\frac{2\pi}{T}(x_i m + z_i \sqrt{\frac{T^2}{\lambda^2} - m^2})} \right\}_m$ and $\{\psi_m(r_i, \phi_i)\}_m = \left\{ e^{jm(\phi_i + \frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda} r_i\right) \right\}_m$ of the discrete-domain functions do not form an orthogonal basis, but might form a frame. In fact, we are interested in the sampling distributions for which these sets form frames. Note that these frames differ from the continuous-domain sampling and reconstruction frames discussed in Subsection 2.2.2. These frames span spaces similarly to the orthogonal bases in Eq. 3.8 and Eq. 3.13. However, the spaces here are vector spaces rather than the functional spaces spanned by the bases in Eq. 3.8 and Eq. 3.13. The vector spaces are derived from these functional spaces by picking only the values of the functions at the locations required by the sampling set $\{(x_i, z_i)\}_{i=1}^s$.

The properties of the frames $\{\varphi_m(x_i, z_i)\}_m$ and $\{\psi_m(r_i, \phi_i)\}_m$ are used to determine the stability of the sampling sets $\{(x_i, z_i)\}_{i=1}^s$. Recall that the stability of a set of irregularly distributed samples is measured by relating the energy of the samples to the energy of the continuous-domain function being sampled. In this case the energy of the samples is simply calculated as $\sum_{i=1}^s |u(x_i, z_i)|^2$. The continuous-domain light field has finite energy, which is the same on each transversal line $(x, z)|_{z=\text{const}}$. It is equal to the energy of the field on the initial line $\|u(x, 0)\|^2$ which, in turn, should be equal to the energy of its Fourier transform $a(k_x)$. The coefficients $\{a_m\}_m$ are obtained by uniform discretization of $a(k_x)$ and therefore their energy is a measure of the energy of the continuous field $u(x, z)$. Hence, the stability of the sampling set $\{(x_i, z_i)\}_{i=1}^s$ is measured in terms of the Fourier generators model from Eq. 3.8 as:

$$A_F \sum_m |a_m|^2 \leq \sum_{i=1}^s |u(x_i, z_i)|^2 \leq B_F \sum_m |a_m|^2. \quad (3.19)$$

The B-F model from Eq. 3.13 represents the Fourier transform $C(\theta)$ of the continuous field in terms of its Fourier series according to Eq. 3.12. Therefore, the energy of the coefficients $\{c_m\}_m$ should be equal to the energy of $C(\theta)$. The function on the Ewald's circle $C(\theta)$ is the 2D Fourier transform of the continuous field $u(x, z)$ and thus the energy of $C(\theta)$ measures the energy of $u(x, z)$. Hence, the stability of the sampling set $\{(r_i, \phi_i)\}_{i=1}^s$ can be measured in terms of the B-F generators model from Eq. 3.13 as:

$$A_B \sum_m |c_m|^2 \leq \sum_{i=1}^s |u(r_i, \phi_i)|^2 \leq B_B \sum_m |c_m|^2. \quad (3.20)$$

The upper and lower bounds in Eq. 3.19 and Eq. 3.20 determine the stability of a sampling set $\{(x_i, z_i)\}_{i=1}^s$. The closer the lower and upper bound are, the better the stability of the set. In fact, these bounds are proportionally related to the bounds of an actual continuous-domain sampling frames. The sampling frame according to the Fourier generators model can be obtained as the ensemble of translates

$\{\alpha(x - x_i, z - z_i)\}_{i=1}^s$ of the function $\alpha(x, z) = \sum_m e^{j\frac{2\pi}{T}(x+z\sqrt{\frac{T^2}{\lambda^2}-m^2})}$. The sampling frame according to the B-F generators model can be obtained as the ensemble of translates $\{\beta(r - r_i, \phi - \phi_i)\}_{i=1}^s$ of the function $\beta(r, \phi) = \sum_m e^{jm(\phi+\frac{\pi}{2})} J_m(\frac{2\pi}{\lambda}r)$. It is possible to verify that $u(x_i, z_i) = \langle u(x, z), \alpha(x - x_i, z - z_i) \rangle$ and $u(r_i, \phi_i) = \langle u(r, \phi), \beta(r - r_i, \phi - \phi_i) \rangle$. Even though these two sampling frames seem to be convenient for analytical purposes, they are not convenient for numerical computations. Moreover, their consideration steps beyond the contents of the publications enclosed in this thesis. The subsequent discussion shows how the frame bounds of the discrete frames $\{\varphi_m(x_i, z_i)\}_m$ and $\{\psi_m(r_i, \phi_i)\}_m$ can also be used to analyze the stability of the sampling set.

3.2.2.4 LS reconstruction and CG

The light field reconstruction problem consists of finding the unknown field-generating coefficients c_m or a_m , given the irregularly distributed samples $D(x_i, z_i), i = 1, \dots, s$. Each of the expansions in Eq. 3.17 and Eq. 3.18, considered for a single sample $u(x_i, z_i)$ (resp. $u(r_i, \phi_i)$), represents a linear equation for the M unknown coefficients $\{a_m\}_m$ or $\{c_m\}_m$. When considered for all the available samples, each set of equations defines a linear system of equations. These systems are expressed in a matrix form as:

$$\mathbf{u} = \mathbf{A}\mathbf{h}, \quad (3.21)$$

where the unknown vector $\mathbf{h} = \mathbf{a} = [a_{-\lfloor \frac{M}{2} \rfloor}, a_{-\lfloor \frac{M}{2} \rfloor + 1}, \dots, a_{\lfloor \frac{M-1}{2} \rfloor}]^T$ or $\mathbf{h} = \mathbf{c} = [c_{-\lfloor \frac{M}{2} \rfloor}, c_{-\lfloor \frac{M}{2} \rfloor + 1}, \dots, c_{\lfloor \frac{M-1}{2} \rfloor}]^T$ contains the field generating coefficients. The light field samples are ordered in the vector $\mathbf{u} = [u(x_1, z_1), u(x_2, z_2), \dots, u(x_s, z_s)]^T$. \mathbf{A} is the reconstruction matrix, which has two different forms depending on the discrete model. For the Fourier generators model \mathbf{A} is

$$\mathbf{A} = \mathbf{R} = \{\mathcal{R}_{p,q}\} = \{e^{j\frac{2\pi}{T}\sqrt{\frac{T^2}{\lambda^2} - (q - \lfloor \frac{M}{2} \rfloor - 1)^2}} z_p e^{j\frac{2\pi}{T}(q - \lfloor \frac{M}{2} \rfloor - 1)x_p}\},$$

$$p = 1, \dots, s, q = 1, \dots, M \quad (3.22)$$

and for the Bessel-Fourier generators model \mathbf{A} is

$$\mathbf{A} = \mathbf{J} = \{\mathcal{J}_{p,q}\} = \{e^{j(q - \lfloor \frac{M}{2} \rfloor - 1)(\phi_p + \frac{\pi}{2})} J_{q - \lfloor \frac{M}{2} \rfloor - 1}(kr_p)\}$$

$$p = 1, \dots, s, q = 1, \dots, M. \quad (3.23)$$

The matrix \mathbf{A} in Eq. 3.21 relates the coefficients of the discrete model, stored in the vector \mathbf{h} , to the irregularly distributed samples of the corresponding light field, stored in \mathbf{u} . However, the application at hand often requires data point values $\{D(x_i, z_i)\}_{i=1}^s$ which do not obey any of the discrete models in Eq. 3.8 or Eq. 3.13. This means that there is no such set of coefficients which, when fed into the model, would produce the desired values at the desired locations. Therefore, the remaining possibility is to reconstruct a set of irregularly distributed light field samples $\hat{\mathbf{u}}$ which approximate the desired values $\{D(x_i, z_i)\}_{i=1}^s$. This statement is formalized as the least squares optimization problem:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} \|\mathbf{D} - \mathbf{A}\mathbf{h}\|, \quad (3.24)$$

where $\mathbf{D} = [D(x_1, z_1), \dots, D(x_s, z_s)]^T$ is the vector of desired data samples. The coefficient vector $\hat{\mathbf{h}}$ produces the approximating vector $\hat{\mathbf{u}} = \mathbf{A}\hat{\mathbf{h}}$ when fed into Eq. 3.21. The approximating light field vector $\hat{\mathbf{u}}$ lies closest to the vector of desired values \mathbf{D} in the s -dimensional range space of the frame matrix \mathbf{A} . This is the space of light fields which obey one of the discrete models in Eq. 3.8 or Eq. 3.13 and are sampled on the irregular grid of points $\{(x_i, z_i)\}_{i=1}^s$.

The LS problem defined in Eq. 3.24 is linear because the operator \mathbf{A} acts linearly on the unknown vector \mathbf{h} . Therefore it can be solved by the CG method, described in Subsection 2.2.2.2. Written for this specific case, the CG iterates as follows:

1. initialize $\hat{\mathbf{h}}_0$ arbitrary, $\mathbf{r}_0 = \mathbf{A}^*\mathbf{D} - \mathbf{A}^*\mathbf{A}\hat{\mathbf{h}}_0$ and $\mathbf{d}_0 = \mathbf{r}_0$
2. for $n = 1$ to n_{it}
 - (a) $\alpha = \frac{\langle \mathbf{r}_n, \mathbf{r}_n \rangle}{\langle \mathbf{A}\mathbf{d}_n, \mathbf{A}\mathbf{d}_n \rangle}$
 - (b) $\hat{\mathbf{h}}_{n+1} = \hat{\mathbf{h}}_n + \alpha \mathbf{d}_n$
 - (c) $\mathbf{r}_{n+1} = \mathbf{r}_n - \alpha \mathbf{A}^*\mathbf{A}\mathbf{d}_n$
 - (d) $\mathbf{d}_{n+1} = \mathbf{r}_{n+1} + \frac{\langle \mathbf{r}_{n+1}, \mathbf{r}_{n+1} \rangle}{\langle \mathbf{r}_n, \mathbf{r}_n \rangle} \mathbf{d}_n$
- end
3. reconstruct the light field values $\hat{\mathbf{u}}$ as $\hat{\mathbf{u}} = \mathbf{A}\hat{\mathbf{h}}$ according to the considered discrete model

The iteration operator is now the matrix $\mathbf{A}^*\mathbf{A}$, the unknown is the vector of coefficients $\hat{\mathbf{h}}$ and the target is the vector of desired field samples \mathbf{D} . The current residual at the n -th iteration is $\mathbf{r}_n = \mathbf{A}^*\mathbf{D} - \mathbf{A}^*\mathbf{A}\hat{\mathbf{h}}_n$, and the current search direction is denoted by the vector \mathbf{d}_n . The convergence of the CG algorithm depends on the ratio between the largest v_{max} and the smallest v_{min} eigenvalues of the iteration operator. These are the the squared of the largest σ_{max} and the smallest σ_{min} singular values of the matrix \mathbf{A} . According to Eq. 2.48, n iterations decrease the reconstruction error by a factor of at least:

$$2 \left(\frac{\sigma_{max} - \sigma_{min}}{\sigma_{max} + \sigma_{min}} \right)^n = 2 \left(\frac{\sigma_{max}/\sigma_{min} - 1}{\sigma_{max}/\sigma_{min} + 1} \right)^n. \quad (3.25)$$

The closer the singular values σ_{max} and σ_{min} are, the closer the error decay factor is to zero and the convergence of the CG algorithm is faster. Note that the column index q of the matrix \mathbf{A} coincides with the frame element index m in Eq. 3.17 and Eq. 3.18. This means that the columns of the matrix \mathbf{A} are the frame elements $\varphi_m(x_i, z_i)$ (resp. $\psi_m(r_i, \phi_i)$) of the discrete frames. Hence, \mathbf{A} plays the role of the frame matrix Φ from Eq. 2.35. Therefore, the smallest and largest singular values of \mathbf{A} coincide with the square root of the lower and upper bounds of the frame used in \mathbf{A} , respectively. By definition, the smallest and largest singular values of a matrix \mathbf{A} are found as:

$$\begin{aligned} \sigma_{min}^2 &= \min_{\mathbf{h}: \|\mathbf{h}\|=1} \|\mathbf{A}\mathbf{h}\|^2 = \min_{\mathbf{h}: \|\mathbf{h}\|=1} \langle \mathbf{A}\mathbf{h}, \mathbf{A}\mathbf{h} \rangle = \min_{\mathbf{h}: \|\mathbf{h}\|=1} \langle \mathbf{u}, \mathbf{u} \rangle = \\ &= \min_{\mathbf{h}: \|\mathbf{h}\|=1} \sum_{i=1}^s |u(x_i, z_i)|^2 = A \|\mathbf{h}\|^2, \end{aligned} \quad (3.26)$$

$$\begin{aligned}
 \sigma_{max}^2 &= \max_{\mathbf{h}: \|\mathbf{h}\|=1} \|\mathbf{A}\mathbf{h}\|^2 = \max_{\mathbf{h}: \|\mathbf{h}\|=1} \langle \mathbf{A}\mathbf{h}, \mathbf{A}\mathbf{h} \rangle = \max_{\mathbf{h}: \|\mathbf{h}\|=1} \langle \mathbf{u}, \mathbf{u} \rangle = \\
 &= \max_{\mathbf{h}: \|\mathbf{h}\|=1} \sum_{i=1}^s |u(x_i, z_i)|^2 = B \|\mathbf{h}\|^2.
 \end{aligned} \tag{3.27}$$

The lower A and upper B frame bounds coincide with A_F and B_F from Eq. 3.19 when the frame $\{\varphi_m(x_i, z_i)\}_m$ is used in \mathbf{A} or with A_B and B_B from Eq. 3.20 when the frame $\{\psi_m(r_i, \phi_i)\}_m$ is used in \mathbf{A} . The singular values σ_{max} and σ_{min} of the matrix \mathbf{A} coincide with the square root of the sampling set bounds. This explains one aspect of the term “stability” of a sampling set. The sampling set and an underlying continuous-domain model define a frame which spans a vector space according to Eq. 3.17 or Eq. 3.18. The frame has the same bounds as the sampling set. This frame is used to reconstruct a desired element from the vector space with the CG algorithm or the frame algorithm. A stable sampling set is one whose lower and upper bounds are close to each other and the reconstruction algorithm converges rapidly.

3.2.2.5 Stability of the sampling set and physically inconsistent data

This subsection explains probably the most influential aspect of the term “stability” of the sampling set. The large difference between the bounds A and B slows down the convergence of the CG algorithm, used for reconstruction. More importantly, the reconstruction algorithm boosts any, even negligibly small, physical inconsistency in the desired sample values $\{D(x_i, z_i)\}_{i=1}^s$ if A is much smaller than B . Here, the reasons for that are explained and the regularization strategy used in **P7** and **P8** is outlined.

Often in a practical application, the desired sample values $\{D(x_i, z_i)\}_{i=1}^s$ cannot be simulated as samples of a scalar light field which satisfies the Helmholtz wave equation. In such a situation, the vector of the desired values \mathbf{D} does not belong to the space spanned by the frames $\{\varphi_m(x_i, z_i)\}_m$ or $\{\psi_m(r_i, \phi_i)\}_m$. However, the desired values are imposed only at the locations $\{(x_i, z_i)\}_{i=1}^s$ and the physical inconsistency can be modeled only there. The inconsistency might be of various origins and difficult to predict in the common case. A reasonable and rather general assumption is to model the inconsistency as an additive Gaussian term $\varepsilon = [\varepsilon_1, \dots, \varepsilon_s]^T$:

$$\mathbf{D} = \mathbf{A}\mathbf{h} + \varepsilon. \tag{3.28}$$

The entries of the random vector ε are drawn from zero-mean normal distribution $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. A naive direct reconstruction approach strongly amplifies the random term ε so that it prevails over the desired coefficient vector estimate $\hat{\mathbf{h}}$ if there is a big difference in the singular values of \mathbf{A} . This defines the reconstruction problem as an inverse problem [134].

Any linear system can be solved by taking the pseudo-inverse with the help of SVD [84]. SVD decomposes an $s \times M$ matrix \mathbf{A} in terms of its singular values and singular vectors:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*, \tag{3.29}$$

where \mathbf{U} and \mathbf{V} are $s \times s$ and $M \times M$ unitary matrices, whose columns are the left and right singular vectors of \mathbf{A} , respectively. The matrix $\mathbf{\Lambda}$ is a $s \times M$ diagonal matrix with the value ordered singular values of \mathbf{A} :

$$\mathbf{\Lambda} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\}, \sigma_{max} = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n = \sigma_{min} \geq 0, \quad (3.30)$$

where $n = \min(s, M)$. The Moore-Penrose pseudo-inverse can be used to find a solution $\hat{\mathbf{h}}$, provided the desired data vector \mathbf{D} [84]:

$$\begin{aligned} \hat{\mathbf{h}} &= \mathbf{A}^\dagger \mathbf{D} = \mathbf{V} \mathbf{\Lambda}^\dagger \mathbf{U}^* \mathbf{D}, \\ \mathbf{\Lambda}^\dagger &= \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\}. \end{aligned} \quad (3.31)$$

The pseudo-inverse is able to deal with the inverse problems which break Hadamard's existence condition [134] when \mathbf{A} has small condition number $\kappa = \frac{\sigma_{max}}{\sigma_{min}}$. If the solution does not exist, applying the pseudo-inverse maps \mathbf{D} to the vector $\hat{\mathbf{h}}$ which lies closest to it in the range space of \mathbf{A} which is known as the LS solution. However, if the condition number of \mathbf{A} is large, the pseudo-inversion does not produce a sensible solution. The norm of \mathbf{A}^\dagger is equal to its largest singular value, which is σ_{min}^{-1} . If the SVD pseudo-inverse is applied to the left side of the inconsistent vector $\mathbf{D} = \mathbf{A}\mathbf{h} + \varepsilon$, the noise norm in the resultant reconstruction is boosted by a factor of σ_{min}^{-1} . In the case when the condition number of \mathbf{A} is large, σ_{min}^{-1} is large as well, and the noise is dominant in the reconstruction.

The LS optimization problem tries to find a solution which yields a residual $\mathbf{D} - \mathbf{A}\mathbf{h}$ with minimum norm. The Tikhonov regularization approach tries to minimize also the norm of the solution:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} \left\{ \|\mathbf{A}\mathbf{h} - \mathbf{D}\|^2 + \delta \|\mathbf{h}\|^2 \right\}. \quad (3.32)$$

Tikhonov regularization balances between the contradictory requirements for small residual and small norm of the solution [134]. The penalty term δ is used to tune this balance. The optimal value of δ can be determined automatically by Morozov's discrepancy principle [134]. The minimum norm requirement ensures the smoothness of the solution and thus robustness to noise, while minimizing the residual ensures that the solution is close to the target. This regularization approach was used in **P7** and **P8**.

Written in SVD terms, the Tikhonov solution can be shown to be equivalent to [134]:

$$\begin{aligned} \hat{\mathbf{h}} &= \mathbf{V} \mathbf{\Lambda}_\delta^\dagger \mathbf{U}^* \mathbf{D}, \\ \mathbf{\Lambda}_\delta^\dagger &= \text{diag} \left\{ \frac{\sigma_1}{\sigma_1^2 + \delta}, \dots, \frac{\sigma_n}{\sigma_n^2 + \delta} \right\}. \end{aligned} \quad (3.33)$$

This form shows improvement over the pseudo-inverse SVD. The small singular values do not boost the random term ε anymore, since they are substituted with a value close to δ . This value is selected in such a manner that the singular vectors corresponding to small eigenvalues are taken to contribute to the solution. On the other hand, large eigenvalues are not modified much by δ and contribute with the accurate weight to the solution, driving $\hat{\mathbf{h}}$ toward the desired value.

The SVD form of the Tikhonov regularized solution in Eq. 3.33 is very convenient for analytical purposes. However, finding the SVD of a matrix requires cubic complexity [84]. A more computationally attractive approach [134] uses the CG method as an iterative matrix solver. The minimizer $\hat{\mathbf{h}}$ of the Tikhonov functional in Eq. 3.32 is equivalent to the solution of the linear system [134]:

$$(\mathbf{A}^* \mathbf{A} + \delta \mathbf{I}) \hat{\mathbf{h}} = \mathbf{A}^* \mathbf{D}. \quad (3.34)$$

The matrix $\mathbf{H} = \mathbf{A}^* \mathbf{A} + \delta \mathbf{I}$ is symmetric and positive definite and thus can be inverted by the CG method. This is the numerical reconstruction approach used in **P7** and **P8**.

The sampling set bounds A and B are determined by the spatial distribution of the desired data points $\{D(x_i, z_i)\}_{i_1}^s$ within the region of interest. Suppose that $u(x, z)$ is the desired light field whose samples approximate the values $\{D(x_i, z_i)\}_{i_1}^s$ at the positions $\{(x_i, z_i)\}_{i_1}^s$. In general, the light field $u(x, z)$ might not spread its energy uniformly within the region of interest, i.e. there are darker areas of shadow and brighter areas with high light field values. Suppose further that a large portion of the samples do not cover the bright parts of the field, leaving large gaps. The energy of these samples is then very small in comparison with the energy of the generating coefficients \mathbf{h} , no matter how densely the samples are distributed outside the gap. This pushes the lower bound A of the sampling set down and consequently increases the condition number κ of the reconstruction matrix \mathbf{A} . Now consider the case when a dense cluster of samples lie on one of the bright areas of the field. The same local information is added several times in Eq. 3.19 or Eq. 3.20, which increases the upper bound B and κ . Thus, the presence of clusters and gaps in the density of the sampling set has significant influence on the bounds of the sampling set $\{u(x_i, z_i)\}_{i_1}^s$. Publications **P7** and **P8** study the influence of this and some other properties of the sample distributions which influence the stability of the sample set. The spectral characteristics of the matrix \mathbf{A} are analyzed in various scenarios depending on the size of the volume of interest, clustering of the data points and amount of input data.

Better bounds and faster convergence of the conjugate gradient algorithm can be achieved if the samples are weighted with some adaptive weights $w_i, i = 1, \dots, s$ satisfying $\sum_i w_i = 1$. Then, the LS problem is to find the coefficient vector \mathbf{h} provided weighted desired values \mathbf{WD} and weighted frame matrix \mathbf{WA} :

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} \|\mathbf{WD} - \mathbf{WA}\mathbf{h}\|, \quad (3.35)$$

where $\mathbf{W} = \text{diag}\{w_1, \dots, w_s\}$ is a diagonal matrix of the sample weights. These weights can be chosen based on the separation between the neighboring samples, e.g. by the reciprocal of the distance to the closest sample. Such an approach is demonstrated in **P6**. In essence, this approach implies weighting the frame elements in the matrix \mathbf{A} . This decreases the difference between the lower and the upper bounds of the frame and the sampling set, respectively.

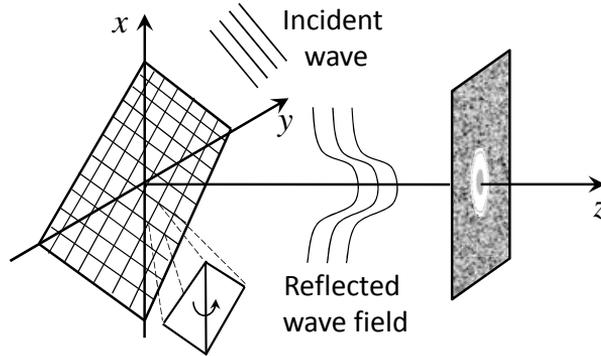


Figure 3.4: Deflectable mirror device.

3.3 Light field synthesis with DMD

This section discusses the generation of a desired complex LF $d(x, y) = |d(x, y)|e^{j\phi_d(x, y)}$ on a reference plane $(x, y, z)|_{z=\text{const}}$ by modulation of an incident light source wave with an optical modulation element. The design and configuration of complex wave front modulation element subject to its technological constraints is greatly facilitated by computer based optimization algorithms. The emphasis in this section is on the use of matching pursuit and simulated annealing algorithms for optimizing the configuration of a DMD which would produce $d(x, y)$ on a desired plane when illuminated by a plane wave. These algorithms constitute the content of publication **P1**.

The DMD is a reflection-mode type of spatial light modulator, which consists of a two-dimensional array of square shaped identical micro-mirrors (Fig. 3.4). Each mirror can be deflected separately, tilting it along its diagonal axis to one of a few pre-fixed angles. The most widely used and commercially available DMD is developed by Texas Instruments [135]. This is micro-electromechanical device, whose chip has on its surface of the order of a million microscopic mirrors. The TI DMD has been used for the synthesis of high quality two-dimensional color images in digital video display systems. Each micro-mirror can be tilted two possible angles of $\pm 10\text{-}12^\circ$, corresponding to “on” and “off” states. In the “on” state, the light from the illuminating light source is reflected to the plane of interest, while in the “off” state, the light is directed onto a side heat sink. Each mirror can modulate the amplitude of the reflected light wave by temporal pulse width modulation. The mirrors themselves are made out of aluminum and are around $10\mu\text{m}$ across with spacing of around $1\mu\text{m}$. The greatest advantage of the DMDs over liquid crystal-based SLMs is the higher brightness and contrast ratio. Occasionally, DMDs have been also used in holographic applications [8]. Kreiss et al. used DMD to synthesize binary amplitude holograms for application in digital holographic interferometry [9]. The generation of holographic stereograms in a gel medium by a DMD is suggested in [10].

Publication **P1** describes LF synthesis by appropriately tilting the mirrors of the DMD and illuminating the device with coherent laser light. The light which is reflected from the mirrors propagates to approximate the intended light field on the

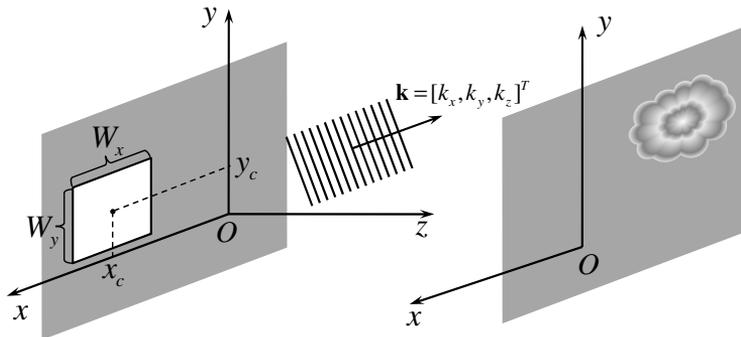


Figure 3.5: Diffraction of a an arbitrary oriented plane wave $e^{j(k_x x + k_y y + k_z z)}$ by a rectangular aperture of size $W_x \times W_y$, lying at the initial plane $(x, y, z)|_{z=0}$ and centered at (x_c, y_c) .

target plane (Fig. 3.4). Note that such an approach concentrates on optimizing the deflection angles for the mirrors, subject to the target light field rather than altering between “on” and “off” states to modulate only the amplitude of the incident light wave. In contrast, either tilt angle of each mirror can contribute to the desired light field. Combining the optimal angles for each mirror enables simultaneous phase and amplitude modulation.

3.3.1 Rectangular aperture illuminated by a plane wave

This section makes analytical derivation of the diffraction field which results when a rectangular aperture is illuminated by an arbitrary oriented plane wave. Such a case sets the background for the approach presented in publication **P1** about light field synthesis by a digital micro-mirror device. The surface of each micro-mirror can be considered to be the rectangular aperture which obscures an incident plane wave. The angle of this plane wave depends on the orientation of the micro-mirror and the DMD illumination wave.

Consider the set-up in Fig. 3.5, where a plane wave $e^{j(k_x x + k_y y + k_z z)}$ illuminates a planar rectangular aperture. Without loss of generality, the aperture is assumed to lie in a transversal plane $(x, y, z)|_{z=0}$ and the problem is to calculate the diffraction pattern at another transversal plane $(x, y, z)|_{z=\text{const}}$ at some distance z . Let the size of the aperture be $W_x \times W_y$ and let it be centered at the point (x_c, y_c) . As discussed in Subsection 2.1.3, the frequency vector $\mathbf{k} = [k_x, k_y, k_z]^T$ determines the orientation of the plane wave, with $|\mathbf{k}|^2 = k_x^2 + k_y^2 + k_z^2 = k^2$. The plane wave within the plane $(x, y, z)|_{z=0}$ of the aperture has the expression $e^{j(k_x x + k_y y)}$. The aperture can be described by the function

$$A(x, y) = \text{rect}\left(\frac{x - x_c}{W_x}\right) \text{rect}\left(\frac{y - y_c}{W_y}\right), \quad (3.36)$$

where $rect(x)$ is a rectangular pulse of width 1 and height 1, centered at the origin. The field $u(x, y, 0)$ generated by the plane wave in the plane of the aperture is the function

$$\begin{aligned}
u(x, y, 0) &= A(x, y)e^{j(k_x x + k_y y)} = rect\left(\frac{x - x_c}{W_x}\right) rect\left(\frac{y - y_c}{W_y}\right) e^{j(k_x x + k_y y)} \\
&= e^{j(k_x x_c + k_y y_c)} rect\left(\frac{\xi}{W_x}\right) rect\left(\frac{\eta}{W_y}\right) e^{j(k_x \xi + k_y \eta)} \\
&= e^{j(k_x x_c + k_y y_c)} v(\xi, \eta),
\end{aligned} \tag{3.37}$$

where $\xi = x - x_c$ and $\eta = y - y_c$. The field $u(x, y, z)$ at the desired plane $(x, y, z)|_{z=\text{const}}$ is the convolution between the field $u(x, y, 0)$ at initial plane and a free space propagation impulse response $h_z(x, y)$:

$$\begin{aligned}
u(x, y, z) &= (u(\cdot, \cdot, 0) * h_z)(x, y) = \iint_{-\infty}^{\infty} u(x', y', 0) h_z(x - x', y - y') dx' dy' \\
&= \iint_{-\infty}^{\infty} v(\xi, \eta) h_z(x - x_c - \xi, y - y_c - \eta) d\xi d\eta \\
&= e^{j(k_x x_c + k_y y_c)} (v(\cdot, \cdot) * h_z)(x - x_c, y - y_c) \\
&= e^{j(k_x x_c + k_y y_c)} \psi_{\mathbf{k}}(\mathbf{x} - \mathbf{x}_c).
\end{aligned} \tag{3.38}$$

In Eq. 3.38, the following vector notations are used: $\mathbf{k} = [k_x, k_y, k_z]^T$, $\mathbf{x} = [x, y, z]^T$ and $\mathbf{x}_c = [x_c, y_c, 0]^T$. The function $\psi_{\mathbf{k}}(\mathbf{x})$ is the field generated when the rectangular aperture is centered at the origin and illuminated by the plane wave with orientation \mathbf{k} . The impulse response $h_z(x, y)$ can be either the R-S impulse response as in Eq. 2.8 or the Fresnel impulse response as in Eq. 2.18. Derivation of $\psi_{\mathbf{k}}(\mathbf{x})$ in terms of the Fresnel impulse response is simpler, yet accurate enough for many practical cases, including the problem considered in publication **P1**. Moreover, a closed form solution can be derived in this case:

$$\begin{aligned}
\psi_{\mathbf{k}}(\mathbf{x}) &= (v(\cdot, \cdot) * h_z)(x, y) \\
&= \iint_{-\infty}^{\infty} rect\left(\frac{\xi}{W_x}\right) rect\left(\frac{\eta}{W_y}\right) e^{j(k_x \xi + k_y \eta)} \frac{e^{jkz}}{j\lambda z} e^{\frac{jk}{2z}((x-\xi)^2 + (y-\eta)^2)} d\xi d\eta \\
&= \frac{e^{jkz}}{j\lambda z} \int_{-W_x/2}^{W_x/2} e^{k_x \xi} e^{\frac{jk}{2z}(x^2 - 2x\xi + \xi^2)} d\xi \int_{-W_y/2}^{W_y/2} e^{k_y \eta} e^{\frac{jk}{2z}(y^2 - 2y\eta + \eta^2)} d\eta \\
&= \frac{e^{jkz}}{j\lambda z} \int_{-W_x/2}^{W_x/2} e^{\frac{jk}{2z}((x - \frac{zk_x}{k}) - \xi)^2} e^{jk_x(\frac{zk_x}{2k} - x)} d\xi \int_{-W_y/2}^{W_y/2} e^{\frac{jk}{2z}((y - \frac{zk_y}{k}) - \eta)^2} e^{jk_y(\frac{zk_y}{2k} - y)} d\eta \\
&= \frac{e^{jkz}}{j} e^{jk_x(\frac{zk_x}{2k} - x)} \mathcal{I}_{W_x}\left(x - \frac{zk_x}{k}\right) e^{jk_y(\frac{zk_y}{2k} - y)} \mathcal{I}_{W_y}\left(y - \frac{zk_y}{k}\right).
\end{aligned} \tag{3.39}$$

The last line of the derivation uses the notation

$$I_W(t) = \frac{1}{\sqrt{\lambda z}} \int_{-W/2}^{W/2} e^{j\frac{\pi}{\lambda z}(t-\tau)^2} d\tau = \frac{1}{\sqrt{2}} \int_{-\sqrt{\frac{2}{\lambda z}}(\frac{W}{2}+t)}^{\sqrt{\frac{2}{\lambda z}}(\frac{W}{2}-t)} e^{j\frac{\pi}{2}\gamma^2} d\gamma. \quad (3.40)$$

The integral $I_W(t)$ can be split into two parts: one with limits from $-\sqrt{\frac{2}{\lambda z}}(\frac{W}{2}+t)$ to 0 and another – from 0 to $\sqrt{\frac{2}{\lambda z}}(\frac{W}{2}-t)$. Each of the resultant integrals has tabulated forms and can be found in many larger computational software packages [136, 137, 138].

As a final result of this discussion, consider the case when the aperture is a planar mirror from the DMD chip. The derivation of the diffracted field away from the mirror should be the same, as the part of the plane wave falling on the mirror is totally reflected, while the rest of the plane wave travels in the direction behind the mirror. The only difference is that the direction of the illuminating plane wave, i.e. the vector $\mathbf{k} = [k_x, k_y, k_z]^T$, must be assumed to follow the direction of the reflection. On the DMD chip, the mirrors are able to change their orientation according to one of a few pre-set angles, say α , while the desired field still falls onto the same target plane. The angle is three-dimensional $\alpha = [\alpha_x, \alpha_y, \alpha_z]^T$ in the general case. Therefore, the solution of Eq. 3.38 for the diffracted field at the target plane must be modified, because the plane wave falls at a different orientation than $\mathbf{k} = [k_x, k_y, k_z]^T$ onto the mirror and the target plane and the initial plane close an angle α . Let us then fix two coordinate systems: one $\mathbf{x} = [x, y, z]^T$ with the target plane and another $\mathbf{x}' = [x', y', z']^T$ with the initial plane. Assume that the coordinate system \mathbf{x} is fixed so that that the target plane is transversal to the Oz axis at some distance z . This is convenient as it coincides with the assumption made when deriving the solution in Eq. 3.38. Let the other coordinate system \mathbf{x}' be fixed so that the tilted initial plane is transversal to the axis Oz' and aligned with the origin $z' = 0$. With this choice, the coordinate systems \mathbf{x} and \mathbf{x}' close an angle α and relate to each other through the rotation matrix \mathbf{R}_α :

$$\mathbf{x}' = \mathbf{R}_\alpha \mathbf{x}. \quad (3.41)$$

In a similar manner, the orientation of the plane wave $\mathbf{k} = [k_x, k_y, k_z]^T$ transforms in the new coordinate system \mathbf{x}' as $\mathbf{k}' = \mathbf{R}_\alpha \mathbf{k}$. Hence the solution for the diffraction field $u(x, y, z)$ at the target plane becomes:

$$u(x, y, z) = e^{j(k_x x_c + k_y y_c)} \psi_{\mathbf{R}_\alpha \mathbf{k}}(\mathbf{R}_\alpha(\mathbf{x} - \mathbf{x}_c)). \quad (3.42)$$

3.3.2 Mathematical formulation

The main aim of publication **P1** was to develop algorithms which determine the optimal configuration of the mirrors on the DMD chip, so that a desired light field $d(x, y)$ at a transversal plane $(x, y, z)|_{z=\text{const}}$ is reproduced to the best extent. The desired field $d(x, y)$ contains amplitude and phase simultaneously, because the mirrors from the DMD chip reflect in a certain manner the incident light wave $r(x, y)$.

The angle of each mirror determines its contribution to the field at the output plane $(x, y, z)|_{z=\text{const}}$. Therefore, the problem of finding the optimal tilt angles for all mirrors subject to the desired field $d(x, y)$ must be described in a formal manner. This subsection formulates the problem of scalar light field synthesis as a constrained optimization problem. This formalism is convenient for the presentation of the possible optimization algorithms.

Assume that the DMD consists of $M \times N$ micro-mirrors, where each mirror can be deflected to one of S possible angles. As the number of mirrors is finite, their index j is selected to be one-dimensional $j = 1, \dots, M \times N$. Let $g_{j,s}(x, y)$ denote the field generated by j -th mirror when it is tilted to s -th possible angle with $s = 1, \dots, S$ and $j = 1, \dots, M \times N$. The respective angle and position of the mirror can be taken into account in Eq. 3.42 to calculate each $g_{j,s}(x, y)$. With this new notation, the total light field, produced by the DMD on the output plane can be expressed as the superposition of the light fields $g_{j,s}(x, y)$ produced by each mirror:

$$g(x, y) = \sum_{j=1}^{M \times N} \sum_{s=1}^S p_{j,s} g_{j,s}(x, y). \quad (3.43)$$

In this linear combination the scalar coefficients $p_{j,s}$ bear certain restrictions. A particular mirror with index j can be turned only to one of its S positions and therefore contributes to the total field $g(x, y)$ only with the field $g_{j,s}(x, y)$ corresponding to this active position s . On the other hand, each mirror on the DMD chip is turned to a certain position, so it inevitably produces one of its fields $g_{j,s}(x, y)$. Hence, the scalar coefficients $p_{j,s}$ take binary values 0 or 1 with the restriction that for the j -th mirror only one coefficient $p_{j,s}$, $s = 1, \dots, S$ can have value 1 and the rest must take value 0. This restriction of the coefficients $p_{j,s}$, together with Eq. 3.43 constrain the minimization problem:

$$\hat{d}(x, y) = \arg \min_g \{ \|d(x, y) - g(x, y)\| \}. \quad (3.44)$$

Eq. 3.44 implies that the optimally synthesized field $\hat{d}(x, y)$ is the field $g(x, y)$ which lies closest to the desired field $d(x, y)$. The field $g(x, y)$ ranges within the space of fields which can be reproduced by the DMD. That is, $g(x, y)$ should satisfy Eq. 3.43 together with the restrictions on the coefficients $p_{j,s}$. Therefore, a candidate optimization algorithm strategy must find the coefficients $p_{j,s}$, $j = 1, \dots, M \times N$, $s = 1, \dots, S$, which generate through Eq. 3.43 the best approximation of $d(x, y)$. The sequel presents the algorithms from publication **P1** which naturally handle such kinds of constrained optimization problems.

3.3.3 Light field synthesis methods

3.3.3.1 Simulated annealing

Simulated annealing is an optimization procedure based on trial and error evaluation of the possible states of the solution. It is a global optimization algorithm which is very

easily applied when the search space is discrete, as in the case for finding the optimal tilt angle for each mirror of the DMD chip. The algorithm avoids local minima by a random search which accepts changes which sometimes might degrade the solution. Taking a degraded solution in a random direction around the current solution widens the search range around a local minimum. This step toward a degraded solution should be large enough with respect to a local minimum and small with respect to a global one. A certain quality metric determines how close to the global minimum the solution lies.

The simulated annealing optimizes the mirror tilt angles on the chip as follows. The initial mirror configuration is selected randomly, and each step of the algorithm alters the state of all mirrors. The quality of the current solution is evaluated by a quality metric Q , which is the reciprocal of the mean squared error $\|d(x, y) - g(x, y)\|^2$. When handling each mirror during the iterations, at first the quality metric Q_1 of the solution with the mirror's current state is calculated. Then, its tilt angle is changed to a different state, selected randomly, and the quality metric Q_2 of the solution with the new state is calculated. The effect of the change is evaluated by the ratio r of the new to the old quality metric $r = \frac{Q_2}{Q_1}$. If the change improves the solution, i.e. $r > 1$, the new state of the mirror is accepted. The new state of the mirror might still be accepted even if it degrades the solution, i.e. $r \leq 1$. In this case the change is accepted with probability $rT(i)$. The function $T(i)$ is a weighting function of the current iteration i , which usually is selected to be monotonically decreasing. This possible random change for each mirror is performed on each mirror in a sequential manner. One pass through all the mirrors forms an iteration of the simulated annealing algorithm.

The choice of the weighting function $T(i)$ is essential for the performance of the algorithm. The function should be decreasing on its argument. In such a manner, random degrading changes are more probable during the early iterations, because it is also more likely to be around a local minimum during the early iterations. On the other hand, late iterations are closer to the global optimum and a degrading change of the mirror states should not be tolerated. The implementation in publication **P1** uses a function $T(i)$, which depends not only on the iteration number i , but also on the current mirror position j . The particular choice is $T = \exp\left\{-\frac{(i-1)MN+j}{4MN}\right\}$. The reason behind this choice is that DMD typically contains a large number of mirrors and the synthesized field might improve significantly during a single pass through all the mirrors.

3.3.3.2 Matching pursuit

Originally, the matching pursuit algorithm has been developed as a fast and greedy approach for signal approximation by a linear combination of waveforms, iteratively selected from a large collection [139]. As a greedy approach, matching pursuit is known for its sub-optimal behavior, converging to a local minimum. However, it is very fast, because a few passes through the collection might be sufficient to arrive at a satisfactory solution.

The expression of the output light field, as formulated in Eq. 3.43, resembles a linear signal expansion of $g(x, y)$ over the collection of waveforms the $g_{j,s}(x, y)$, weighted with the coefficients $p_{j,s}$. The difference with the classical case here is that the set of values for the coefficients $p_{j,s}$ is constrained to 0 and 1, where for a certain mirror j only one of the possible $p_{j,s}$ can be 1 and the rest must be 0. The waveforms in this case are the light fields $g_{j,s}(x, y)$ obtained by each mirror turned to certain angle. None of the DMD configurations guarantees to reproduce fully any desired light field $d(x, y)$ which satisfies the Helmholtz wave equation. Therefore, the expansion in Eq. 3.43 is not guaranteed to describe it precisely. However, a similar description in terms of the light fields $g_{j,s}(x, y)$ is:

$$d(x, y) = \sum_{j=1}^{m \leq M \times N} \sum_{s=1}^S p_{j,s} g_{j,s}(x, y) + R_m(x, y) = g(x, y) + R_m(x, y). \quad (3.45)$$

The function $R_m(x, y)$ denotes the residual between the actual synthesized field $g(x, y)$ and the desired one $d(x, y)$. The index j of the first summation in Eq. 3.45 runs up to $m \leq M \times N$. This notation is used for generality, i.e. it also includes the case when the desired field $d(x, y)$ is approximated with fewer mirrors than available on the DMD chip.

Matching pursuit is a recursive adaptive algorithm that builds up the signal representation one element at a time, picking the most contributive element at each step. In terms of the considered optimization problem, the algorithm chooses at each step a mirror turned at a certain position among all mirrors with all possible positions. The chosen mirror and position generate a light field $g_{j_s}(x, y)$ which contributes best to the desired field. The measure for contribution is the portion of the current residual $R_j(x, y)$, as defined by Eq. 3.45, which the generated light field $g_{j_s}(x, y)$ occupies. Formally, the algorithm starts from an initial residual $R_0(x, y) = d(x, y)$ and the mirror and tilt angle selected at each step are the ones which minimize the norm of the residual $\|R_j(x, y)\| = \|R_{j-1}(x, y) - g_{j,s}(x, y)\|$. In this manner, matching pursuit naturally serves the constrained optimization problem at hand. The modification used in publication **P1** is outlined below.

- initialize $R_0(x, y) = d(x, y)$
- for $j = 1$ to MN
 - (a) go through all the mirrors and their positions and pick the most contributive one: $\{i, s\} = \arg \min_{i,s} \|R_{j-1}(x, y) - g_{i,s}(x, y)\|$
 - (b) update residual $R_j(x, y) = R_{j-1}(x, y) - g_{i,s}(x, y)$
- end

Matching pursuit works particularly well when the waveforms in the collection are mutually orthogonal [139]. However, such a statement is not valid for the light fields $g_{j,s}(x, y)$ generated from the mirrors when turned to any of their angles. The algorithm selects the mirror and the positions one by one, resulting in local optimization.

Therefore, in some cases the first few elements might be chosen wrongly, and then the algorithm spends many subsequent iterations without much improvement. Therefore, such an algorithm may end up far from the optimum solution. On the positive side, some fields might be sufficiently approximated by only a few mirrors, chosen greedily during the first steps. In such a case, matching pursuit is preferable as very fast algorithm.

The speed of this algorithm can be traded for achieving better approximation. This is illustrated in publication **P1** by increasing the order of the optimization. It is essentially the same as the procedure above, but uses a combination of two mirrors instead of one. That is, a combination of the two most contributive mirrors is found instead of the most contributive mirror at step (a). Step (b) subtracts the joint contribution of these two mirrors from the current residual.

Chapter 4

Summary of results

This thesis contains a collection of publications related to scalar light field synthesis and reconstruction from a set of data points, irregularly distributed within a region of interest. The problem of light field synthesis is decoupled from the problem of light field reconstruction from the input data points, as described in Fig. 3.1. Detailed discussion of the light field synthesis problem for a particular light modulation device (DMD) is presented in publication **P1**. The main emphasis of the thesis is on the problem of reconstruction of scalar light fields from a set of irregularly distributed data points. Most of the publications on light field reconstruction consider one transverse dimension only. Generalization to 3D is considered in publication **P4**. The reconstruction problem is initiated in publication **P3**, where the first practical reconstruction algorithms are discussed in detail. Publications **P2**, **P4**, **P5**, **P6**, **P7** and **P8** cast the problem as reconstruction from a set of irregularly distributed light field samples. Publication **P5** derives a discrete model based on cylindrical harmonics. Publication **P6** compares the convergence properties of the LS-based and POCS-based reconstruction methods and **P7** and **P8** study the stability of the reconstruction under various data point distribution scenarios.

Publication **P1** considers the problem of scalar light field synthesis by a DMD. This work formulates the problem as a constrained linear optimization problem, which can be solved by suitable global or local combinatorial optimization algorithm. The formulation is done by calculating possible fields, generated by illuminating each mirror at each of its possible tilt angles. Two optimization algorithms are considered: one local – matching pursuit algorithm and one global – simulated annealing. The conventional matching pursuit algorithm is adapted to the constrained optimization problem. In addition, a second-order matching pursuit algorithm is suggested in order to improve the behavior, otherwise prone to falling into local minima. This improvement comes at the expense of increased computational complexity. Simulated experiments are carried out for two possible DMD sizes, where each DMD is configured to synthesize two possible patterns on a desired target plane – one of the patterns belongs to the range space of the device, while the other does not. In all cases, simulated annealing shows better reconstruction than both matching pursuit algorithms, reaching the

global minimum. However, matching pursuit algorithms use less computations, still showing satisfactory reconstructions. The convergence of the simulated annealing algorithm is very much dependent on its starting point. Moreover, its parameters have to be tuned by an ad-hoc approach. The author contributed to this publication mainly with the part related to the matching pursuit algorithm. This includes the development of the idea, its implementation for numerical experiments and analysis of the results. The text of the paper was written and edited jointly with the co-authors, while the author made a significant contribution.

Publication **P3** is the first publication in the collection which considers in detail the problem of light field reconstruction from arbitrarily distributed data points within a region of interest. One transverse dimension is considered and the known data points are taken from a 2D rectangular grid which covers the 2D region of interest. The light field is reconstructed on a reference line, at an uniform grid of points, aligned to the 2D grid of points in the region of interest. The PWD integral is assumed to describe accurately the light field propagation in this publication. Two approaches to light field reconstruction are considered. The first approach reconstructs the light field on an initial line by inverting a linear system with the Moore-Penrose pseudo-inverse. The linear system is obtained by writing a propagation equation from the initial line to each given data point. The second approach is based on the serial POCS method and represents the major novelty of publication **P3**. It defines one set per transversal line in the 2D grid from the region of interest. The projections from set to set are implemented efficiently based on the FFT algorithm by propagating from line to line and re-substituting the known data points. Both algorithms are evaluated to reconstruct a light field, which is generated by a *rect*-function on the initial line. The distribution of points inside the region of interest is done arbitrarily. A fair evaluation is achieved by averaging the reconstruction results over an ensemble of different data point distributions. The experiments reconstruct the light field for a variable number of known data points. The POCS-based reconstruction is able to produce satisfactory results after many iterations, even when the known data points are less than the target resolution, i.e. less than the number N of points which constitute the sampling grid on the initial line. For more than N known data points, the matrix pseudo-inversion-based method reconstructs to practically zero error. For a small number of iterations, the POCS-based approach needs slightly larger than N data points to converge to a zero error. When the number of given data points increases, the computational complexity of the POCS-based approach decreases up to a certain level. The pseudo-inverse reconstruction is more efficient if the number of known data points is less than or slightly larger than N . The author contributed to the theoretical derivation of the POCS-based reconstruction algorithm. He also implemented the POCS-based algorithm and prepared the experiments for the publication. His contribution to the manuscript concerns writing all parts related to the POCS-based algorithm and also organizing the structure and editing of the text jointly with the other co-authors.

Publication **P2** casts the LF reconstruction problem as reconstruction from irregularly distributed LF samples. Such a consideration allows the positions of the known data points to be completely free, not belonging to any predefined uniform grid as the one considered in publication **P3**. The irregular sampling problem is posed through a discrete model of the continuous diffraction field. In essence, this model discretizes the PWD integral. Such a discretization leads to periodization of the light field in

a transversal direction, which requires an essentially space-limited support along the transversal direction. As essentially space-limited patterns tend to spread upon propagation, the maximal transversal size of the region of interest can be used to determine the periodization. Finite dimension of the model is achieved if the frequency support of the diffraction field is also assumed to be essentially limited. The reconstruction problem is posed as an LS problem by writing the model for each known field sample within the region of interest. Publication **P2** suggests solving the LS problem by the iterative method of conjugate gradients because of its fast convergence. Simulations are carried out on two different light fields: a Gaussian beam and a rectangular beam, similarly as in publication **P3**. The Gaussian beam has essentially limited support in space and frequency, while the rectangular beam is not band-limited. However, the reconstruction results for both fields are very similar. The CG procedure converges in a small number of iterations if the known field samples are slightly more than the assumed non-zero coefficients in the discrete model. The number of iterations which lead to a satisfactory reconstruction is further reduced if the number of known samples is increased. Consequently, the computational complexity reduces in a similar manner. However, the iterations are not significantly reduced for very large number of samples, while the complexity per iteration increases. Therefore, it is not beneficial to have a very large number of samples. The author took part in the clarification of the idea behind the problem formulation and solution. Most of the theoretical derivations were also carried out by him. The numerical experiments and results were implemented by the author, who also analyzed the results and prepared the manuscript of the paper.

Publications **P3** and **P2** derive two iterative light field reconstruction approaches: one of them is based on the POCS method and the other one on the CG method. Publication **P4** generalizes these two approaches for the 3D case when there are two transverse dimensions. A comparison of the numerical performance of these approaches is interesting for an eventual practical application. The POCS-based approach propagates efficiently from plane to plane by the use of FFT and assumes that the known data points belong to a predefined 3D uniform grid which is aligned with a uniform grid on the reconstructed reference plane. Such a setup is not in favor of the CG-based reconstruction, which assumes completely free data point positions. The simulations in publication **P4** try to reconstruct a diffraction field which has a *rect* distribution on the reference plane. The simulations vary the number of known data points and the number of iterations for the CG and POCS. The two algorithms have different computational complexity per iteration and, therefore, they can be assessed by the number of complex multiplications. The results show that POCS converges for a smaller number of known samples than CG when the computational complexity of each algorithm is fixed. If a desired reconstruction error is fixed, then POCS again outperforms CG, requiring a smaller number of complex multiplications to achieve the desired error for any amount of known samples. Even though POCS seems to be faster than CG in every way, the CG-based approach has the great advantage that the known samples can occupy freely any positions within the region of interest, not necessarily aligned to a uniform grid. The author implemented the CG-based algorithm and took part in the analysis of the numerical results. Another contribution is to the revision and editing of the manuscript.

The results of publication **P4** inspired a theoretical study and comparison of the

POCS and CG-based reconstruction approaches in publication **P6**. For a fair comparison, the POCS-based reconstruction approach is generalized to free known sample positions. This is done by defining each set by the points on a transversal line, which can be considered to belong to the same uniform grid. This grid can have free offset from the grid on the reference line. In such a set up, there can appear more than one set per line and/or only one known sample defining a set. The projections are done by efficient FFT-based propagation from the initial line to the line defining the set, where the free shift between the grids is compensated according to the shift theorem of the Fourier transform. The error analysis of this POCS-based reconstruction shows that the error reduction after each projection depends on the portion of the energy which is contained in the samples forming the set. The convergence of the CG-based approach depends on the condition number of the matrix obtained when writing the diffraction equation for each point in the irregular sampling set. The entries of this matrix depend only on the positions of the data points. In this sense, a stable sampling set yields a matrix with a good condition number. An unstable sampling can be moderately compensated by introducing adaptive weights to each sample, resulting in a weighted LS-based reconstruction. The simulations in this publication reconstruct a Gaussian beam from a fixed number of known samples. This number is selected to be sufficiently large, i.e. such that its further increase would not significantly speed up the convergence. One of the experiments illustrates the benefit of using adaptive weights for CG-based reconstruction. Another experiment shows that the POCS-based approach converges faster when the sets formed by the known samples are better structured. The author initiated the idea of this publication by deriving the convergence and error analysis for the POCS-based algorithm. The idea of incorporating free sample positions in the POCS-based algorithm came during this derivation. The author also implemented and carried out all experiments which produce the results of the publication. He also prepared the essential part of the manuscript.

Publication **P5** derives an alternative model for the LS-based light field reconstruction approach with the CG method. The model is derived for one transverse dimension only in a polar coordinate system. It decomposes the continuous light field onto a set of functions, which are cylindrical harmonics along the angle and Bessel functions of the first kind along the radius. The biggest advantage of this model is evident during the derivation – the model does not assume finite spatial extent of the field for the inherent discretization. The Fourier spectrum of the field is supported on Ewald’s circle and thus it is periodic. It can be naturally represented in discrete terms by its Fourier series. The Bessel functions of the first kind are used to represent the remaining integral of the inverse Fourier transform. The model is feasible in practice, because there exist efficient and accurate numerical algorithms to calculate Bessel functions [140]. The simulations in this publication aim at reconstructing a Gaussian beam for a different number of known data points within a region of interest. A satisfactory reduction in a small number of CG iterations is achieved when the known data points are slightly more than the dimensionality of the model. The convergence of this model depends only on the positions of the known samples, in a similar manner to the discrete PWD model in publications **P2**, **P4** and **P6**. The Bessel functions-based model leads to faster CG convergence when the sampling set contains points with more diverse positions along the angular direction. The author took part in the derivation

of the discrete model consisting of cylindrical harmonics and Bessel functions. He also proved that the continuous functions constituting the model form an orthogonal set. The author implemented all the numerical experiments and produced all the results. He had a leading role in preparing the manuscript.

Publications **P7** and **P8** investigate the factors which influence the stability of light field reconstruction from a set of data samples, irregularly distributed within a region of interest. These factors are related mainly to the distribution of known data samples within the region of interest. Such a consideration is interesting for many practical applications where the specified data samples are not physically consistent. A consistent light field is represented by the discrete models of publications **P2** and **P5**. The inconsistency of the known data points is represented inside the models as an additive random term, whose values have normal distribution. Such a representation is quite general as it encompasses inconsistency in a wide range of applications. The representation of the known data points poses the light field reconstruction problem as an inverse problem. Both publication **P7** and publication **P8** use basic Tikhonov regularization for the light field reconstruction from irregularly distributed and physically inconsistent data samples. The motivation behind this choice is to illustrate the properties of the modeling. Tikhonov regularized inversion is implemented iteratively by the CG method. In theory, the models from **P2** and **P5** require a certain dimensionality which is sufficient for the representation of a light field with certain NDF, measured by the space-bandwidth product of $u(x, 0)$. This issue is derived in detail in publication **P8**. The required dimensionality of both models is proportional to the NDF of the light field. The excess of the PWD-based model from publication **P2** is determined by the ratio of the minimal and maximal field support within the region of interest. The excess of the the Bessel functions-based model from publication **P2** is determined by the ratio of the finest detail structure of the field to the wavelength. Extensive numerical simulations investigate the factors which affect the stability of the reconstruction for scenarios which require the same dimensionality for both models. The stability is measured by the distribution of the singular values of the matrices which correspond to the models. Uniform sample density defines the reconstruction problem as well-posed when described by the PWD-based model and as ill-posed when described by the Bessel functions-based model. However, spatial variations of the density define the reconstruction problem as ill-posed for both models. The more clustered the sample density is, the less uniformly the singular values of the reconstruction matrices are distributed. Consequently, the reconstruction requires regularized inversion. **P7** investigates only a limited number of scenarios, where the level of clustering is simulated by varying the extent of the spatial region occupied by a fixed amount of data samples. In addition, **P8** considers a variable amount of data samples. Moreover, the existence of an ill-posed problem and, respectively, the need for regularization is clearly underlined by comparing the performance of the reconstruction in cases with ill-posed and well-posed data sample distributions, consistent and inconsistent known sample values. In all experiments, the regularization is able to compensate for the ill-posedness of the reconstruction problem, yielding reconstruction error comparable to the level of inconsistency of the samples. The convergence of the CG algorithm becomes slower for sampling sets with more clustered density. Consequently, the reconstruction error in a finite number of iterations is larger. The practical applicability of the reconstruction approach is illustrated by an additional

experiment in **P8**. The experiment is done on a distribution of the field samples, mimicking a scenario which arises for high-resolution light field reconstruction from a multiple low-resolution and limited-aperture CCD recordings. The results demonstrate rapid convergence to an error, comparable to the level of inconsistency of the sample values, where the CCD noise is considered as the source of inconsistency. The author has prepared the major parts of publications **P7** and **P8**. These include the idea of the problem modeling, solution and considered scenarios, their derivations, the implementation of the experiments and analysis of the results, and preparation of the manuscript.

Bibliography

- [1] H. M. Ozaktas and L. Onural, Eds., *Three-Dimensional Television: Capture, Transmission, Display*, Springer, 2008.
- [2] L. Yaroslavsky and M. Eden, *Fundamentals of Digital Optics*, Birkhuser, Boston, 1996.
- [3] R. Piestun and J. Shamir, “Synthesis of three-dimensional light fields and applications,” *Proceedings of the IEEE*, vol. 90, pp. 222–244, 2002.
- [4] J. W. Goodman, *Introduction to Fourier Optics*, Mc-Graw-Hill, New York, 1996.
- [5] J.A. Neff, R.A. Athale, and S.H. Lee, “Two-dimensional spatial light modulators: A tutorial,” *Proceedings of the IEEE*, vol. 78, pp. 826–855, 1990.
- [6] J. Amako, H. Miura, and T. Sonehara, “Wave-front control using liquid-crystal devices,” *Applied Optics*, vol. 32, no. 23, pp. 4323–4329, 1993.
- [7] R. W. Cohn, “Analyzing the encoding range of amplitude-phase coupled spatial light modulators,” *Optical Engineering*, vol. 38, no. 2, pp. 361–367, 1999.
- [8] D. Dudley, W.M. Duncan, and J. Slaughter, “Emerging digital micromirror device (dmd) applications,” in *SPIE Proceedings, 4985, MOEMS Display and Imaging Systems*, 2003.
- [9] T. Kreiss, P. Aswendt, and R. Hofling, “Hologram reconstruction using a digital micromirror device,” *Optical Engineering*, vol. 40, pp. 926–933, 2001.
- [10] M. L. Huebschman, B. Munjuluri, and H. R. Garner, “Dynamic holographic 3-d image projection,” *Optics Express*, vol. 11, pp. 437–445, 2003.
- [11] T. Dresel, M. Beyerlein, and J. Schwider, “Design and fabrication of computer-generated beam-shaping holograms,” *Applied Optics*, vol. 35, pp. 4615–4621, 1996.
- [12] Z. Zalevsky, D. Mendlovic, and R. G. Dorsch, “Gerchberg-saxton algorithm applied in the fractional fourier or the fresnel domain,” *Optics Letters*, vol. 21, pp. 842–844, 1996.

- [13] D. Mendlovic, Z. Zalevsky, G. Shabtay, and E. Marom, "High-efficiency arbitrary array generator," *Applied Optics*, vol. 35, pp. 6875–6880, 1996.
- [14] G. Zhou, X. Yuan, P. Dowd, Y. Lam, , and Y. Chan, "Design of diffractive phase elements for beam shaping: hybrid approach," *J. Opt. Soc. Am. A*, vol. 18, pp. 791–800, 2001.
- [15] R.W. Gerchberg and W.O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [16] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput Math Math Phys*, vol. 7, pp. 1–24, 1967.
- [17] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part I-theory," *IEEE Trans Med Imag*, vol. TMI-1, pp. 81–94, 1982.
- [18] J. Rosen and A. Yariv, "Snake beam: a paraxial arbitrary focal line," *Optics Letters*, vol. 20, pp. 2042–2044, 1995.
- [19] B. Salik, J. Rosen, and A. Yariv, "One-dimensional beam shaping," *J. Opt. Soc. Am. A*, vol. 12, pp. 1702–1706, 1995.
- [20] G. Shabtay, Z. Zalevsky, U. Levy, and D. Mendlovic, "Optimal synthesis of three-dimensional complex amplitude distributions," *Optics Letters*, vol. 25, pp. 363–365, 2000.
- [21] G. Shabtay, "Three-dimensional beam forming and ewald's surfaces," *Optics Communications*, vol. 226, pp. 33–37, 2003.
- [22] U. Levy, D. Mendlovic, Z. Zalevsky, G. Shabtay, and E. Marom, "Iterative algorithm for determining optimal beam profiles in a three-dimensional space," *Applied Optics*, vol. 38, pp. 6732–6736, 1999.
- [23] R. Piestun, B. Spektor, and J. Shamir, "Wave fields in three dimensions: analysis and synthesis," *J. Opt. Soc. Am. A*, vol. 13, pp. 1837–1848, 1996.
- [24] J. Courtial, G. Whyte, Z. Bouchal, and J. Wagner, "Iterative algorithms for holographic shaping of non-diffracting and self-imaging light beams," *Optics Express*, vol. 14, pp. 2108–2116, 2006.
- [25] F. Gori, "Fresnel transform and the sampling theorem," *Optics Communications*, vol. 39, pp. 293–297, November 1981.
- [26] L. Onural, "Sampling of the diffraction field," *Applied Optics*, vol. 39, pp. 5929–5935, November 2000.
- [27] L. Onural, "Some mathematical properties of the uniformly sampled quadratic phase function and associated issues in fresnel diffraction simulations," *Optical Engineering*, vol. 43, pp. 2557–2563, 2004.
- [28] A. Stern and B. Javidi, "Analysis of practical sampling and reconstruction from fresnel fields," *Optical Engineering*, vol. 43, pp. 239–250, 2004.

- [29] A. VanderLugt, "Optimum sampling of fresnel transforms," *Applied Optics*, vol. 29, pp. 3352–3361, August 1990.
- [30] O. Schreer, P. Kauff, and T. Sikora, *3D Video Communication*, Wiley, 2005.
- [31] Y. Zhang, G. Pedrini, W. Osten, and H. Tiziani, "Whole optical wave field reconstruction from double or multi in-line holograms by phase retrieval algorithm," *Optics Express*, vol. 11, pp. 3234–3241, 2003.
- [32] P. Almero, G. Pedrini, and W. Osten, "Complete wavefront reconstruction using sequential intensity measurements of a volume speckle field," *Applied Optics*, vol. 45, pp. 8596–8605, 2006.
- [33] P. Almero, G. Pedrini, and W. Osten, "Aperture synthesis in phase retrieval using a volume-speckle field," *Optics Letters*, vol. 32, pp. 733–735, 2007.
- [34] C. Falldorf, M. Agour, C. von Kopylow, and R. B. Bergmann, "Phase retrieval by means of a spatial light modulator in the fourier domain of an imaging system," *Appl. Opt.*, vol. 49, no. 10, pp. 1826–1830, Apr 2010.
- [35] C. Falldorf, M. Agour, C. von Kopylow, and R. B. Bergmann, "Design of an optical system for phase retrieval based on a spatial light modulator," in *American Institute of Physics Conference Series*, P. K. Rastogi & E. Hack, Ed., Apr. 2010, vol. 1236, pp. 259–264.
- [36] F.A. Marvasti, Ed., *Nonuniform sampling: Theory and Practice*, Kluwer, 2001.
- [37] T. Strohmer, *Efficient Methods for Digital Signal and Image Reconstruction from Nonuniform Samples*, Ph.D. thesis, University of Vienna, 1993.
- [38] R.J. Duffin and A.C. Schaeffer, "A class of nonharmonic fourier series," *Transactions of the American Mathematical Society*, vol. 72, pp. 341–366, 1952.
- [39] J.R. Higgins, "A sampling theorem for irregularly spaced sample points," *IEEE Transactions on Information Theory*, vol. 22, pp. 621–622, 1976.
- [40] A. Aldroubi and K. Gröchenig, "Beurling-landau type theorems for nonuniform sampling in shift invariant spline spaces," *J. Fourier Anal. Applicat.*, vol. 6, no. 1, pp. 93–103, 2000.
- [41] M.I. Kadec, "The exact value of the paley-wiener constant," *Sov. Math. Dokl.*, vol. 5, pp. 559–561, 1964.
- [42] K. Gröchenig and T. Strohmer, "Numerical and theoretical aspects of non-uniform sampling of band-limited images," in *Nonuniform sampling: Theory and Practice*, F. A. Marvasti, Ed. 2001, Kluwer, Chapter 6.
- [43] T. Strohmer, "Computationally attractive reconstruction of bandlimited images from irregular samples," *IEEE Transactions on Image Processing*, vol. 6, pp. 540–548, 1997.
- [44] P. Ferreira, "Interpolation and the discrete papoulis-gerchberg algorithm," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2596–2606, 1994.

- [45] K. Gröchenig, “Reconstruction algorithms in irregular sampling,” *Mathematics of Computation*, vol. 59, pp. 181–194, 1992.
- [46] A. Aldroubi and K. Gröchenig, “Nonuniform sampling and reconstruction in shift-invariant spaces,” *SIAM Review*, vol. 43, pp. 585–620, 2001.
- [47] H.G. Feichtinger and K. Gröchenig, “Theory and practice of irregular sampling,” in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier, Eds. 1993, pp. 305–363, CRC Press, Chapter 8.
- [48] H. Stark, Y. Yang, and D. Gurkan, “Factors affecting convergence in the design of diffractive optics by iterative vector-space methods,” *J. Opt. Soc. Am. A*, vol. 16, pp. 149–159, 1999.
- [49] R. Piestun and J. Shamir, “On-axis computer-generated holograms for three-dimensional display,” *Optics Letters*, vol. 22, pp. 922–924, 1997.
- [50] T. Kotzer, N. Cohen, and J. Shamir, “Generalized projection algorithms with applications to optics and signal restoration,” *Optics Communications*, vol. 156, pp. 77–91, 1998.
- [51] H. G. Feichtinger, K. Gröchenig, and T. Strohmer, “Efficient numerical methods in non-uniform sampling theory,” *Numerische Mathematik*, vol. 69, pp. 423–440, 1995.
- [52] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, John Wiley and Sons, Inc., 1991.
- [53] E. Hecht, *Optics*, Addison Wesley, 1998.
- [54] É. Lalor, “Conditions for the validity of the angular spectrum of plane waves,” *J. Opt. Soc. Am.*, vol. 58, pp. 1235–1237, 1968.
- [55] T. Kreiss, *Handbook of Holographic Interferometry*, John Wiley and Sons, Inc., 2004.
- [56] R.N. Bracewell, *The Fourier Transform and Its Applications*, McGraw-Hill, third edition, 2000.
- [57] A. W. Lohmann, D. Mendlovic, Z. Zalevsky, and G. Shabtay, “The use of ewald’s surfaces in triple correlation optics,” *Optics Communications*, vol. 144, pp. 170–172, 1997.
- [58] M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, Oxford, sixth edition, 1993.
- [59] W. Kaplan, *Advanced Calculus*, Addison-Wesley, fifth edition, 2003.
- [60] N. Mukunda, “Consistency of rayleigh’s diffraction formulas with kirchhoff’s boundary conditions,” *J. Opt. Soc. Am.*, vol. 52, no. 3, pp. 336–337, 1962.
- [61] E. Wolf and W. Marchand, “Comparison of the kirchhoff and the rayleigh-sommerfeld theories of diffraction at an aperture,” *J. Opt. Soc. Am.*, vol. 54, no. 5, pp. 587–594, 1964.

- [62] J. B. Keller, "Geometrical theory of diffraction," *J. Opt. Soc. Am.*, vol. 52, no. 2, pp. 116–130, 1962.
- [63] A. Oppenheim, A. Willsky, and S. Hamid, *Signals and Systems*, Prentice Hall, second edition, 1996.
- [64] A. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, Prentice Hall, third edition, 2009.
- [65] J. Astola and L. P. Yaroslavsky, *Advances in Signal Transforms: Theory and Applications*, Hindawi, first edition, 2007.
- [66] F. Shen and A. Wang, "Fast-fourier-transform based numerical integration method for the rayleigh-sommerfeld diffraction formula," *Appl. Opt.*, vol. 45, pp. 1102–1110, 2006.
- [67] C. Pozrikidis, *Numerical Computation in Science and Engineering*, Oxford U. Press, first edition, 1998.
- [68] V. Nascov and P. C. Logofătu, "Fast computation algorithm for the rayleigh-sommerfeld diffraction formula using a type of scaled convolution," *Appl. Opt.*, vol. 48, no. 22, pp. 4310–4319, 2009.
- [69] H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, *The Fractional Fourier Transform, with Applications in Optics and Signal Processing*, Wiley, first edition, 2001.
- [70] J. A. C. Veerman, J. J. Rusch, and H. P. Urbach, "Calculation of the rayleigh-sommerfeld diffraction integral by exact integration of the fast oscillating factor," *J. Opt. Soc. Am. A*, vol. 22, no. 4, pp. 636–646, 2005.
- [71] L. A. D'Arcio, J. M. Braat, and H. J. Frankena, "Numerical evaluation of diffraction integrals for apertures of complicated shape," *J. Opt. Soc. Am. A*, vol. 11, no. 10, pp. 2664–2674, 1994.
- [72] W. E. Boyce and R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, Wiley, sixth edition, 1996.
- [73] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast fourier transform approach," *J. Opt. Soc. Am. A*, vol. 15, no. 4, pp. 857–867, 1998.
- [74] N. Delen and B. Hooker, "Verification and comparison of a fast fourier transform-based full diffraction method for tilted and offset planes," *Appl. Opt.*, vol. 40, no. 21, pp. 3525–3531, 2001.
- [75] L. Onural, "Exact analysis of the effects of sampling of the scalar diffraction field," *J. Opt. Soc. Am. A*, vol. 24, no. 2, pp. 359–367, 2007.
- [76] C. E. Shannon, "Communication in the presence of noise," *Proceedings of IRE*, vol. 37, pp. 10–21, 1949.
- [77] H. Nyquist, "Certain topics in telegraph transmission theory," *Trans. AIEE*, vol. 47, pp. 617–644, 1928.

- [78] E. T. Whittaker, “On the functions which are represented by the expansions of the interpolation theory,” *Proc. Royal Soc. Edinburgh*, vol. 35, pp. 181–194, 1915.
- [79] V. A. Kotelnikov, “On the carrying capacity of the ether and wire in telecommunications,” in *Material for the First All-Union Conference on Questions of Communication*, Moscow, USSR, 1933, Izd. Red. Upr. Svyazi RKKKA.
- [80] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhuser, first edition, 2002.
- [81] J. Kovacevic and A. Chebira, “Life beyond bases: the advent of frames (part i),” *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 86–104, 2007.
- [82] S.-Ch. Pei and M.-H. Yeh, “An introduction to discrete finite frames,” *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 84–96, 1997.
- [83] D. Han, K. Kornelson, D. Larson, and E. Weber, *Frames for Undergraduates*, American Mathematical Society, first edition, 2007.
- [84] Åke Björck, *Numerical Methods for Least Squares Problems*, SIAM, Amsterdam, Holland, 1990.
- [85] C. Neumann, *Untersuchungen ber das Logarithmische und Newtonsche Potential*, Teubner.
- [86] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, USA, 2003.
- [87] R. E. A. C. Paley and N. Wiener, “Fourier transforms in complex domains,” *AMS Colloquium Publications*, vol. 19, 1934.
- [88] N. Levinson, “Gap and density theorems,” *AMS Colloquium Publications*, vol. 26, 1940.
- [89] F. J. Beutler, “Error-free recovery of signals from irregularly spaced samples,” *SIAM Review*, vol. 8, pp. 328–335, 1966.
- [90] K. Yao and J. O. Thomas, “On some stability and interpolatory properties of nonuniform sampling expansions,” *IEEE Transactions on Circuit Theory*, vol. 14, pp. 404–408, 1967.
- [91] J. R. Higgins, “Five short stories about the cardinal series,” *Bulletin of American Mathematical Society*, vol. 12, pp. 45–89, 1985.
- [92] H. Landau, “Necessary density conditions for sampling and interpolation of certain entire functions,” *Acta Mathematica*, vol. 117, pp. 37–52, 1967.
- [93] H. G. Feichtinger and K. Gröchenig, “Iterative reconstruction of multivariate band-limited functions from irregular sampling values,” *SIAM Journal of Mathematical Analysis*, vol. 23, pp. 244–261, 1992.
- [94] F. Marvasti and M. Analoui, “Recovery of signals from nonuniform samples using iterative methods,” in *Proceedings of International Symposium for Circuits and Systems*, Portland, USA, 1989.

- [95] F. Marvasti, "An iterative method to compensate for the interpolation distortion," *IEEE Transaction on Acoustics, Speech and Signal Processing*, vol. 37, no. 10, pp. 1617–1621, 1989.
- [96] K. Gröchenig, "A discrete theory of irregular sampling," *Linear Algebra and its Applications*, vol. 193, pp. 129–150, 1993.
- [97] Y. Liu and G. G. Walter, "Irregular sampling in wavelet subspaces," *Journal of Fourier Analysis Applications*, vol. 2, pp. 181–189, 1996.
- [98] W. Chen, S. Itoh, and J. Shiki, "Irregular sampling theorems for wavelet subspaces," *IEEE Transactions on Information Theory*, vol. 44, pp. 1131–1142, 1998.
- [99] Y. Liu, "Irregular sampling for spline wavelet subspaces," *IEEE Transactions on Information Theory*, vol. 42, pp. 623–627, 1996.
- [100] X. Zhou and W. Sun, "On the sampling theorem for wavelet subspaces," *Journal of Fourier Analysis Applications*, vol. 5, pp. 347–354, 1999.
- [101] R. Aharoni and Y. Censor, "Block iterative projection methods for parallel computation of solutions to convex feasibility problems," *Linear Algebra Applicat.*, vol. 120, pp. 165–175, 1989.
- [102] R. G. Dorsch, A. W. Lohmann, and S. Sinzinger, "Fresnel ping-pong algorithm for two-plane computer-generated hologram display," *Applied Optics*, vol. 33, pp. 869–875, 1994.
- [103] S.-J. Yeh and H. Stark, "Iterative and one-step reconstruction from nonuniform samples by convex projections," *Journal of the Optical Society of America A*, vol. 7, no. 3, pp. 491–499, 1990.
- [104] H. G. Feichtinger, C. Cenker, M. Mayer, H. Steier, and T. Strohmer, "New variants of the pocs method using affine subspaces of finite codimension, with applications to irregular sampling," in *SPIE92 Conference on Visual Communications and Image Processing*, Boston, USA, 1992, p. 299310.
- [105] A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," *Journal of the Optical Society of America A*, vol. 1, no. 9, pp. 932–943, 1984.
- [106] T. Kotzer, N. Cohen, and J. Shamir, "A projection based algorithm for consistent and inconsistent constraints," *SIAM Journal on Optimization*, vol. 7, no. 2, 1997.
- [107] P. L. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.
- [108] A. Stern and B. Javidi, "General sampling theorem and application in digital holography," in *SPIE Proceedings*, 2004, vol. 5557, pp. 110–123.
- [109] A. Stern and B. Javidi, "Sampling in the light of wigner distribution," *J. Opt. Soc. Am. A*, vol. 21, no. 3, pp. 360–366, 2004.

- [110] A. W. Lohmann, R. G. Dorsch, D. Mendlovic, Z. Zalevsky, and C. Ferreira, "Space-bandwidth product of optical signals and systems," *J. Opt. Soc. Am. A*, vol. 13, no. 3, pp. 470–473, 1996.
- [111] A. Walther, "Gabor's theorem and energy transfer through lenses," *Journal of the Optical Society of America*, vol. 57, no. 5, pp. 639–644, 1967.
- [112] J. M. Arnold, "Phase-space localization and discrete representations of wave fields," *Journal of the Optical Society of America A*, vol. 12, no. 1, pp. 111–123, 1995.
- [113] S. Qian and D. Chen, "Decomposition of wigner-ville distribution and time-frequency distribution series," *IEEE Transaction on Signal Processing*, vol. 42, no. 10, pp. 2836–2842, 1994.
- [114] F. Peyrin and R. Prost, "A unified definition for the discrete-time, discrete frequency, and discrete-time/frequency wigner distributions," *IEEE Transaction on Acoustics, Speech, and Signal Processing*, vol. 34, no. 4, pp. 858–867, 1986.
- [115] U. Schnars and W. Jüptner, "Direct recording of holograms by a ccd target and numerical reconstruction," *Appl. Opt.*, vol. 33, no. 2, pp. 179–181, Jan 1994.
- [116] Y. Frauel, T. J. Naughton, O. Matoba, E. Tajahuerce, and B. Javidi, "Three-dimensional imaging and processing using computational holographic imaging," *Proceedings of the IEEE*, vol. 94, no. 3, pp. 636–653, 2006.
- [117] B. R. Brown and A. Lohmann, "Complex spatial filtering with binary masks," *Applied Optics*, vol. 5, no. 6, pp. 967–969, 1966.
- [118] W. H. Lee, "Sampled fourier transform hologram generated by computer," *Applied Optics*, vol. 9, no. 3, pp. 639–643, 1970.
- [119] C. B. Burckhardt, "A simplification of lee's method of generating holograms by computer," *Applied Optics*, vol. 9, no. 8, pp. 1949, 1970.
- [120] R. Tudela, I. Labastida, E. Martn-Badosa, S. Vallmitjana, I. Juvells, and A. Carnicer, "A simple method for displaying fresnel holograms on liquid crystal panels," *Optics Communications*, vol. 214, no. 1-6, pp. 107–114, 2002.
- [121] P. Birch, R. Young, C. Chatwin, M. Farsari, D. Budgett, and J. Richardson, "Fully complex optical modulation with an analogue ferroelectric liquid crystal spatial light modulator," *Optics Communications*, vol. 175, no. 4-6, pp. 347–352, 2000.
- [122] C. Stolz, L. Bigué, and P. Ambs, "Implementation of high-resolution diffractive optical elements on coupled phase and amplitude spatial light modulators," *Applied Optics*, vol. 40, no. 35, pp. 6415–6424, 2001.
- [123] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, 1982.
- [124] P. Yang, Y. Liu, W. Yang, M. Ao, S. Hu, B. Xu, and W. Jiang, "An adaptive laser beam shaping technique based on a genetic algorithm," *Chinese Optics Letters*, vol. 5, pp. 497–500, 2007.

- [125] X. Qi and F. Palmieri, “Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space: Part i. basic properties of selection and mutation,” *IEEE Transactions on Neural Networks*, vol. 5, pp. 102–119, 1994.
- [126] Y. Li, D. Abookasis, and J. Rosen, “Computer-generated holograms of three-dimensional realistic objects recorded without wave interference,” *Applied Optics*, vol. 40, no. 17, pp. 2864–2870, 2001.
- [127] D. Abookasis and J. Rosen, “Computer-generated holograms of three-dimensional objects synthesized from their multiple angular viewpoints,” *Journal of the Optical Society of America A*, vol. 20, no. 8, pp. 1537–1545, 2003.
- [128] R. Piestun and J. Shamir, “Pattern generation with an extended focal depth,” *Applied Optics*, vol. 37, no. 23, pp. 5394–5398, 1998.
- [129] R. Piestun, “Control of wave-front propagation with diffractive elements,” *Optics and Photonics News*, pp. 28–32, 2001.
- [130] R. Piestun and J. Shamir, “Control of wave-front propagation with diffractive elements,” *Optics Letters*, vol. 19, no. 11, pp. 771–773, 1994.
- [131] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, USA, 1972.
- [132] N. J. Moore and M. A. Alonso, “Bases for the description of monochromatic, strongly focused, scalar fields,” *Journal of the Optical Society of America A*, vol. 26, no. 7, pp. 1754–1761, 2009.
- [133] A.B. Carlson, P.B. Crilly, and J.C. Rutledge, *Communication Systems*, McGraw-Hill, fourth edition, 2001.
- [134] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005.
- [135] R. J. Gove, “DMD display systems: the impact of an all digital display,” in *International Symposium of the Society for Information Display*. Texas Instruments White Pages, 1994.
- [136] S. Wolfram, *The Mathematica Book*, Cambridge University Press, fourth edition, 1999.
- [137] E. Pärt-Enander and A. Sjöberg, *The Matlab V Handbook*, Addison Wesley Longman, first edition, 2000.
- [138] F. Garvan, *The MAPLE Book*, Chapman and Hall/CRC, first edition, 2001.
- [139] S. Mallat and Z. Zhang, “Matching pursuit with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, 1993.
- [140] D. E. Amos, “Algorithm 644: A portable package for bessel functions of a complex argument and nonnegative order,” *ACM Transactions on Mathematical Software*, vol. 12, no. 3, pp. 265–273, 1986.

Publication P1

E. Ulusoy, **V. Uzunov**, L. Onural, H.M. Ozaktas, A. Gotchev, “Three-dimensional monochromatic light field synthesis with a deflectable mirror array device,” in *Photon Management II: SPIE Proceedings*, vol. 6187, Strasbourg, France, April 3-7 2006, pp. 1J-1 – 1J-12.

©2006 SPIE. Reprinted with permission.

Three-Dimensional Monochromatic Light Field Synthesis with a Deflectable Mirror Array Device

Erdem Ulusoy¹, Vladislav Uzunov², Levent Onural¹,
Haldun M. Ozaktas¹ and Atanas Gotchev²

¹Bilkent University, Department of Electrical and Electronics Engineering, TR-06800, Ankara, Turkey

²Tampere University of Technology, Institute of Signal Processing, FIN-33101, Tampere, Finland

ABSTRACT

We investigated the problem of complex scalar monochromatic light field synthesis with a deflectable mirror array device (DMAD). First, an analysis of the diffraction field produced by the device upon certain configurations is given assuming Fresnel diffraction. Specifically, we derived expressions for the diffraction field given the parameters of the illumination wave and the tilt angles of the mirrors. The results of the analysis are used in later stages of the work to compute the samples of light fields produced by mirrors at certain points in space. Second, the light field synthesis problem is formulated as a linear constrained optimization problem assuming that mirrors of the DMAD can be tilted among a finite number of different tilt angles. The formulation is initially developed in the analog domain. Transformation to digital domain is carried out assuming that desired fields are originating from spatially bounded objects. In particular, we arrived at a $\mathbf{D}\mathbf{p} = \mathbf{b}$ type of problem with some constraints on \mathbf{p} , where \mathbf{D} and \mathbf{b} are known, and \mathbf{p} will be solved for and will determine the configuration of the device. This final form is directly amenable to digital processing. Finally, we adapt and apply matching pursuit and simulated annealing algorithms to this digital problem. Simulations are carried out to illustrate the results. Simulated annealing performs successful synthesis when supplied with good initial conditions. However, we should come up with systematic approaches for providing good initial conditions to the algorithm. We do not have an appropriate strategy currently. Our results also suggest that simulated annealing achieves better results than MP. However, if only a part of the mirrors can be used, and the rest can be turned off, the performance of MP is acceptable and it turns out to be stable for different types of fields.

Keywords: Light Field Synthesis, Deflectable Mirror Array Device, Fresnel Diffraction, Linear Constrained Optimization Problem, Matching Pursuit Algorithm, Simulated Annealing Algorithm

1. INTRODUCTION

In this work we consider the problem of complex scalar monochromatic light field synthesis with a deflectable mirror array device (DMAD) and we deal with the associated signal processing. The DMAD is a reflection-mode spatial light modulator (SLM), which consists of a two-dimensional array of square shaped identical micro-mirrors. The mirrors can be tilted along their diagonal axes. The deflection of each mirror can be controlled separately. One of the most widely used practical DMADs is the digital micromirror device (DMD) developed by Texas Instruments.¹ This device comprises a large number of mirrors, typically in the order of a million mirrors per chip. This device has been used for the synthesis of high quality two-dimensional color images in digital video display systems.² It has also been occasionally used in holographic applications. In particular, its employment in generation of holographic stereograms was suggested,³ and its usage in digital holographic interferometry to synthesize binary amplitude holograms was proposed.⁴

Here we concentrate on the usage of DMADs to synthesize desired monochromatic light fields within a spatial volume of interest. Piestun and Shamir⁵ formulate the general light field synthesis problem as the optimization of device configurations for the purpose of achieving the most successful creation of a desired light field within

This work is supported by EC within FP6 under Grant 511568 with acronym 3DTV.

a region in space and in time. Our formulation is similar to theirs. In our case, light fields can be synthesized by appropriately tilting the mirrors on the DMAD and illuminating the device with coherent laser light. The light which is reflected from the mirrors will propagate to approximate the intended light field within the target volume (Fig. 1). Since we posed our problem as the synthesis of monochromatic light fields, time dimension is irrelevant to us. We will basically try to replicate the phasor of a given monochromatic field.

Therefore, given the desired field, the key problem is the determination of the tilt angles of the mirrors accordingly. Therefore it falls in the framework of combinatorial problems. In some applications, it may be the case that this inverse problem should be solved in a fast manner compatible with real-time operation. However, the problem has a very high dimensionality due to the large number of mirrors present on the device. Hence, solving this problem is a challenging task that requires the usage of both accurate and fast combinatorial optimization algorithms. An attractive candidate is the simulated annealing algorithm.^{6,7} Its major advantage over other methods is an ability to avoid becoming trapped at local minima. The algorithm employs a random search which not only accepts changes that decrease objective function, but also some changes that increase it with time decaying probability. Another fast algorithm that comes from the field of signal representation is the matching pursuit algorithm.⁸ It is a greedy strategy for signal approximation by a linear combination of waveforms, iteratively selected from a large collection. This algorithm can be easily adapted to the problem at hand by restricting the possible values of the coefficients in the linear combination.

Below, we first provide an analysis of the field produced by this device in sec. 2. Up to a point, we summarize the results we derived in an earlier work.⁹ We also extend the analysis assuming Fresnel diffraction. Second, we pose the light field synthesis problem as a linear constrained optimization problem in sec. 3. Initially, the problem is posed in the analog domain. The developed framework is independent of particular digitization schemes. Later, assuming that fields to be synthesized are arising due to diffraction from spatially bounded objects, we build up a digital optimization problem on top of the analog optimization problem. This problem can directly be attacked on a digital computer. Finally, we describe the adaption and application of matching pursuit and simulated annealing algorithms to solve this digital optimization problem, respectively in sec. 4.1 and sec. 4.2. We test our formulations and illustrate our results through a number of computer simulations, which we describe in sec. 4.3. Our conclusions are stated in sec. 5.

2. ANALYSIS OF LIGHT FIELD GENERATED BY A DMAD

In this section, we firstly summarize the results of one of our earlier works.⁹ In that work, there is an analysis of the light field produced by a DMAD given the tilt angles of the mirrors on the device. However, the results were not specialized for a specific diffraction kernel. Here, we secondly extend the results of that previous analysis for the case of Fresnel diffraction.

We will express the light field produced by the device upon illumination by a plane wave under a certain configuration with respect to the coordinate system shown in Fig. 2. We favored this coordinate system because the mirrors are tilted along their diagonal axes. By this choice, we expect to obtain more manageable expressions at the end. In this coordinate system, tilt axes of mirrors are parallel to y -axis. The mirrors of the DMAD can be indexed with a column vector $\mathbf{i} = [m \ n]^T$ with $m, n \in \mathcal{Z}$, such that the \mathbf{i} 'th mirror is centered around:

$$\mathbf{T}_{\mathbf{i}} = \begin{bmatrix} x_c \\ y_c \\ 0 \end{bmatrix} = \left(W\sqrt{2} + \frac{L\sqrt{2}}{2} \right) \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{i} \quad (1)$$

where $2W$ is the width of the mirrors and L is the length of the spacing between the mirrors.

Let the device be illuminated by an incident plane wave whose functional representation in the coordinate system described above is given by $\exp\{j\frac{2\pi}{\lambda}\boldsymbol{\kappa}^T\mathbf{x}\}$. Here, \mathbf{x} is the position vector given by $\mathbf{x} = [x \ y \ z]^T$ and $\boldsymbol{\kappa}$ is the direction cosines vector of the plane wave given by $\boldsymbol{\kappa} = [\alpha \ \beta \ \gamma]^T$ satisfying the constraint $\alpha^2 + \beta^2 + \gamma^2 = 1$. This incident plane wave will hit the mirrors, and it will be reflected back by each of them. Let us in particular denote the functional representation of the field reflected from the \mathbf{i} 'th mirror by $u_{\boldsymbol{\kappa}}^{(\mathbf{i},\theta_{\mathbf{i}})}(\mathbf{x})$ with respect our coordinate system, assuming that the mirror is tilted by an angle $\theta_{\mathbf{i}}$ in clockwise direction around its

axis of rotation. Note that the dependence of the field on the direction of the incident plane wave is also stressed through inclusion of the κ parameter to the representation. Then, the total field reflected from the DMAD can be obtained by superposing the individual fields reflected from the mirrors as:

$$u(\mathbf{x}) = \sum_{\mathbf{i}} u_{\kappa}^{(\mathbf{i}, \theta_{\mathbf{i}})}(\mathbf{x}) \quad (2)$$

Turning our attention back to the individual fields, $u_{\kappa}^{(\mathbf{i}, \theta_{\mathbf{i}})}(\mathbf{x})$ is given as:

$$u_{\kappa}^{(\mathbf{i}, \theta_{\mathbf{i}})}(\mathbf{x}) = \exp \left\{ j \frac{2\pi}{\lambda} \kappa^T \mathbf{T}_{\mathbf{i}} \right\} \psi_{\mathbf{R}_{\theta_{\mathbf{i}} \kappa}}(\mathbf{R}_{\theta_{\mathbf{i}}}(\mathbf{x} - \mathbf{T}_{\mathbf{i}})) \quad (3)$$

where $\mathbf{T}_{\mathbf{i}}$ is the matrix given by Eq. 1, \mathbf{R}_{θ} is given as:

$$\mathbf{R}_{\theta} = \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \theta & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \sin \theta \\ -\frac{1}{\sqrt{2}} \cos \theta & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \sin \theta \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (4)$$

and the six parameter function $\psi_{\kappa}(\mathbf{x})$ is given as:

$$\psi_{\kappa}(\mathbf{x}) = \psi \left[\begin{matrix} \alpha & \beta & \gamma \end{matrix} \right]^T \left(\begin{bmatrix} x & y & z \end{bmatrix}^T \right) = h_z(x, y) \otimes_{x,y} \text{rect} \left(\frac{x}{2W} \right) \text{rect} \left(\frac{y}{2W} \right) \exp \left\{ j \frac{2\pi}{\lambda} (\alpha x + \beta y) \right\} \quad (5)$$

Here, $h_z(x, y)$ denotes the impulse response of free space propagation system and $\otimes_{x,y}$ denotes 2D convolution in the x and y coordinates.

We refer to Fig. 3 to explain the function defined in Eq. 5 in more detail. In this figure, as seen, there is a single square shaped mirror of width $2W$, and this mirror is placed in a coordinate system such that its aperture can be described by $\text{rect} \left(\frac{x}{2W} \right) \text{rect} \left(\frac{y}{2W} \right)$. Also, this mirror is illuminated by a plane wave, which is represented by $\exp \left\{ j \frac{2\pi}{\lambda} \kappa^T \mathbf{x} \right\} = \exp \left\{ j \frac{2\pi}{\lambda} (\alpha x + \beta y + \gamma z) \right\}$. Hence, the field on the mirrors aperture becomes $\text{rect} \left(\frac{x}{2W} \right) \text{rect} \left(\frac{y}{2W} \right) \exp \left\{ j \frac{2\pi}{\lambda} (\alpha x + \beta y) \right\}$. This function can be convolved with $h_z(x, y)$ to arrive at the description of the field that is reflected from the mirror. Since $\psi_{\kappa}(\mathbf{x})$ is obtained exactly in the same way as suggested by Eq. 5, $\psi_{\kappa}(\mathbf{x})$ actually represents the reflected field in the problem depicted in Fig. 3.

The equations derived up to this point are valid regardless of the choice for the diffraction kernel $h_z(x, y)$. Actually, only the $\psi_{\kappa}(\mathbf{x})$ function of Eq. 5 is influenced by the choice of the diffraction kernel. In the rest of this section, we will derive explicit expressions for this function under the Fresnel approximation for free-space diffraction. The Fresnel kernel for free-space diffraction is given as¹⁰:

$$h_z(x, y) = \frac{e^{jkz}}{j\lambda z} e^{j \frac{k}{2z} (x^2 + y^2)} \quad (6)$$

Above, we pointed the connection between the problem depicted in Fig. 3 and $\psi_{\kappa}(\mathbf{x})$. Therefore, to solve for $\psi_{\kappa}(\mathbf{x})$, we have to solve the problem in Fig. 3 for arbitrary values of κ , i.e. for arbitrary illumination directions, assuming Fresnel diffraction. In the fourth chapter of his book,¹⁰ Goodman solves this problem for the case of normally incident illumination ($\alpha = \beta = 0, \gamma = -1$). In particular, he achieved the following result:

$$\psi \left[\begin{matrix} 0 & 0 & -1 \end{matrix} \right]^T \left(\begin{bmatrix} x & y & z \end{bmatrix}^T \right) = \frac{e^{jkz}}{j} I(x) I(y) \quad (7)$$

where

$$I(x) = \frac{1}{\sqrt{\lambda z}} \int_{-W}^W \exp \left\{ j \frac{\pi}{\lambda z} (x - \zeta)^2 \right\} d\zeta \quad (8)$$

Here, as previously declared, $2W$ is the width of the mirrors. Eq. 8 can be simplified further by setting

$$\alpha_1 = -\sqrt{\frac{2}{\lambda z}} (W + x) \quad (9)$$

$$\alpha_2 = \sqrt{\frac{2}{\lambda z}} (W - x) \quad (10)$$

Using the new variables above, Eq. 8 can be rewritten as

$$I(x) = \frac{1}{\sqrt{2}} \{ [C(\alpha_2) - C(\alpha_1)] + j [S(\alpha_2) - S(\alpha_1)] \} \quad (11)$$

where $C(z)$ and $S(z)$ are the familiar frequently encountered and tabulated Fresnel integrals:

$$C(z) = \int_0^z \cos\left(\frac{\pi t^2}{2}\right) dt \quad (12)$$

$$S(z) = \int_0^z \sin\left(\frac{\pi t^2}{2}\right) dt \quad (13)$$

To extend these results to the case of oblique illumination with small incidence angle (for general α, β, γ satisfying $\alpha^2 + \beta^2 + \gamma^2 = 1$, and α, β sufficiently small so that Fresnel approximation is valid) we essentially went through a procedure similar to that outlined by Goodman,¹⁰ and arrived at the following explicit formula for $\psi_\kappa(\mathbf{x})$:

$$\begin{aligned} \psi \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix}^T \begin{bmatrix} x & y & z \end{bmatrix}^T &= \frac{e^{jkz}}{j} \exp \left\{ j \frac{\pi}{\lambda z} (2\alpha z x - \alpha^2 z^2) \right\} I(x - \alpha z) \\ &\quad \exp \left\{ j \frac{\pi}{\lambda z} (2\beta z y - \beta^2 z^2) \right\} I(y - \beta z) \end{aligned} \quad (14)$$

where $I(x)$ is given by Eq. 8 as before.

3. LINEAR CONSTRAINED OPTIMIZATION PROBLEM FORMULATION

Our main intention is to synthesize monochromatic light fields with a DMAD in a space region of interest to the best extent. As explained in detail in sec. 1, to synthesize a desired complex light field to the best extent, we must determine the optimum tilt angles for the mirrors. Therefore, we should develop a procedure to convert the specifications of the desired light field to these optimum tilt angles. In this section, we formulate the problem of complex monochromatic light field synthesis with a DMAD as a linear constrained optimization problem. A lot of readily available optimization algorithms exist for this class of problems. For the rest of this paper, we assume that the mirrors on the DMAD we work with can be tilted to a finite number of discrete tilt angles, rather than being continuously tunable.

For convenience, we slightly change the notation of the previous section. Let us assume that our DMAD consist of $M \times N$ ($M, N \in \mathcal{Z}^+$) micro-mirrors and each micro-mirror can be tuned to $S \in \mathcal{Z}^+$ different discrete tilt angles. Since the mirrors are finite in number, let us adopt a one-dimensional indexing scheme for the mirrors this time, reserving the letter j for the mirror index, such that $j \in \mathcal{Z}$ and $1 \leq j \leq M \times N$. Let $u_{j_s}(\mathbf{x})$ denote the field generated by j 'th mirror when it is tilted to s 'th position ($s \in \mathcal{Z}$, $1 \leq s \leq S$). Finally, let $b(\mathbf{x})$ represent the desired light field.

With this new notation, the total light field produced by the DMAD can be expressed in the following form:

$$u(\mathbf{x}) = \sum_{j=1}^{M \times N} \sum_{s=1}^S p_{j_s} u_{j_s}(\mathbf{x}) \quad (15)$$

This expression represents the field for a fixed configuration of the device (i.e. the configuration does not change in time) such that the tilt angles for the entire DMAD have been specified. At first glance, this expression seems to be a linear combination of all individual fields generated by the mirrors. However, there are several constraints associated with this linear combination due to the nature of the device. In particular, all the possible fields $u_{j_s}(\mathbf{x})$ cannot be present in this linear combination with nonzero coefficients simultaneously. Only one of the S fields produced by each mirror can be contributing to the total field, since each mirror can be tilted to a single position at any time. On the other hand, each mirror on the chip will in any case be tilted in one of the

S angles in a fixed manner, so will inevitably be producing one of the S fields. All of these remarks imply that in Eq. 15, for each j , strictly one of the coefficients p_{j_s} is one and the others are zero.

Therefore, Eq. 15 is actually a constrained linear combination of the individual fields produced by each mirror. Remember that given $b(\mathbf{x})$ (the desired field), our aim is to determine the set of optimum tilt angles for the mirrors. In other words, for all j , we want to find out the best selections for the nonzero and unity coefficients p_{j_s} , such that $b(\mathbf{x})$ is synthesized to the best extent. At this point, we have set the light field synthesis problem as a linear constrained optimization problem in the analog domain.

Although we have a well-defined optimization problem in the analog domain, we will most likely have to solve this problem in digital domain. Therefore, we have to transform the analog problem into an equivalent digital problem. For this purpose, the first step is the digitization of the information present in the analog light fields $u_{j_s}(\mathbf{x})$ and $b(\mathbf{x})$. There may be various strategies for this task, each of which makes sense under different assumptions and approximations. Here, we base our digitization strategy on the assumption that the light fields we wish to synthesize propagate from spatially bounded objects. We believe this is reasonable in real-world situations. When the Fresnel approximation is considered for diffraction, light fields emerging from spatially bounded objects can be represented fully by their samples taken at finite rates.^{11,12} In the case of arbitrary spatially bounded objects, the number of samples required to fully characterize the diffraction fields might be infinite. In our work, we further assume that a finite number of samples is enough to characterize the fields. We assume that all the fields we wish to synthesize are fully determined by their samples taken at a set of R ($R \in \mathcal{Z}^+$) predetermined space locations. Thus, we can represent $u_{j_s}(\mathbf{x})$ and $b(\mathbf{x})$ by vectors \mathbf{u}_{j_s} and \mathbf{b} respectively. Both \mathbf{u}_{j_s} and \mathbf{b} are of size $R \times 1$.

Finally, we get the following matrix equation as a digital description of our light field synthesis problem:

$$\mathbf{D}\mathbf{p} = \mathbf{b} \quad (16)$$

In the above equation,

- Vector $\mathbf{b} \in \mathcal{C}^{R \times 1}$ represents the given field as stated before.
- $\mathbf{D} \in \mathcal{C}^{(M \times N \times S) \times R}$ represents the behavior of the device. Each column of \mathbf{D} , which is of size $R \times 1$, corresponds to the samples of one of the individual fields produced by one of the mirrors on the device. Specifically, the $\{(j-1) \times S + s\}$ th column of \mathbf{D} is \mathbf{u}_{j_s} .
- $\mathbf{p} \in \mathcal{R}^{(M \times N \times S) \times 1}$ generates the constrained linear combination of the columns of \mathbf{D} , as previously stated through Eq. 15. We seek to solve for this vector actually, which will dictate the choice of tilt positions of the mirrors. The already discussed constraints of the problem reduce the set of possible values for this vector. In particular, the following requirements should be fulfilled: \mathbf{p} is such that for all $k \in \mathcal{Z}$, $1 \leq k \leq M \times N$; the sub-vector $\mathbf{p}[(k-1) \times S + 1 : k \times S]$ is equal to a column of the $S \times S$ identity matrix. Each of these sub-vectors actually determine the configuration of a single mirror. The nonzero element within the sub-vector will correspond to the tilt angle selected for that mirror.

At last, we achieved a discretized linear constrained optimization problem. The problem can be attacked on a digital computer in this form. A myriad of algorithms exist for this type of problems, such as constrained least squares, matching pursuit or simulated annealing, and they can be applied to solve for the optimum tilt angles (the vector \mathbf{p}) given the desired field (the vector \mathbf{b}) and the specifications of the device (the matrix \mathbf{D}).

4. APPLICATION OF MATCHING PURSUIT AND SIMULATED ANNEALING ALGORITHMS

Having put our optimization problem in a suitable format for digital processing, now it is time to adapt and apply solution algorithms. For this purpose, here we consider matching pursuit and simulated annealing algorithms, and report our simulation results. First, we give a brief introduction to these algorithms.

4.1. Matching Pursuit Algorithm

The light field synthesis problem, as formulated by Eq. 15, can be formally interpreted as a signal representation problem. The desired light field $u(\mathbf{x})$ can be considered as a signal that has to be represented by linear combination of waveforms, selected from a dictionary \mathbf{D} . These waveforms can be the light fields $u_{j_s}(\mathbf{x})$ generated by the mirrors of the DMAD. Respectively, p_{j_s} are the coefficients in the representation. The difference in our problem is the restrictions on these coefficient that were imposed in subsection 3. Since we deal with practical light fields that cannot be always synthesized exactly by the DMAD, their approximate representation can be stated as

$$u(\mathbf{x}) = \sum_{j=1}^m \sum_{s=1}^S p_{j_s} u_{j_s}(\mathbf{x}) + R^{(m)}(\mathbf{x}), \quad (17)$$

where $R^{(m)}(\mathbf{x})$ is the signal residual after synthesis by $m < M \times N$ dictionary elements. Mallat and Zhang⁸ have discussed a general method for such approximate decomposition that addresses the sparsity issue directly. Matching pursuit (MP) is a recursive, adaptive algorithm that builds up the signal representation one element at a time, picking the most contributive element at each step. Starting from an initial residual $R^{(0)}(\mathbf{x}) = u(\mathbf{x})$, the element chosen at the j th step is the one which minimizes $\|R^{(j)}(\mathbf{x})\|$ as defined in (17). The residual at stage j is given by $R^{(j)}(\mathbf{x}) = R^{(j-1)}(\mathbf{x}) - p_{j_s} u_{j_s}(\mathbf{x})$.

The MP approach works well for many types of signals. It has been shown to be especially useful for extracting structure from signals which consist of components with widely varying time-frequency localizations.⁸ When stopped after a few steps, it yields an approximation using only a few atoms. The algorithm selects the dictionary elements one by one, resulting in a local optimization. Therefore, in certain cases it might choose wrongly in the first few iterations and end up spending most of its time correcting for any mistakes made in the first few terms. This results in an algorithm that may end up far from the global optimal solution.

Despite of this drawback, MP is a very attractive choice since it is a fast search procedure. This is necessary when one has to deal with DMADs consisting of thousands of mirrors. Another advantage in comparison with most of the traditional optimization methods is that MP can be easily modified to adopt the constraints on the coefficients p_{j_s} . This modification is outlined below.

- initialize $R^{(0)}(\mathbf{x}) = u(\mathbf{x})$
- for $j = 1$ to MN
 - (a) go through all the mirrors and their positions and pick the most contributive one: $\{i, s\} = \arg \min_{i,s} \|R^{(j-1)}(\mathbf{x}) - p_{i_s} u_{i_s}(\mathbf{x})\|$
 - (b) update residual $R^{(j)}(\mathbf{x}) = R^{(j-1)}(\mathbf{x}) - p_{i_s} u_{i_s}$
- end

The orthogonality between the light fields reflected from each mirror of the DMAD is highly dependent on the chosen sampling grid. Therefore, the dictionary elements for our specific case are not necessarily orthogonal and the MP algorithm is expected to show its local optimality behavior. This raises the question of whether higher-order optimizations may be of benefit. To clarify this need some experiments were done also with a "second-order" matching pursuit (MP2). It follows essentially the same strategy as described above. The difference is that at step (a), instead of the most contributive mirror, a combination of the two most contributive mirrors is found. At step (b) the joint contribution of these two mirrors is subtracted from the residual. It should be noted that the complexity of the search for pairs of most contributive mirrors increases to $O((M \times N)^3)$ in comparison to $O((M \times N)^2)$ in the case of first order matching pursuit.

4.2. Simulated Annealing Algorithm

Simulated annealing is a well-known and frequently used optimization algorithm.^{6,7} Since it is widely used, we will directly present its application to our problem. Basically, in our case the algorithm starts with an initial configuration of the mirrors, and we hope to eventually arrive at a state such that the desired field is replicated to a satisfactory extent. In our experiments, as the distance metric, we use the total mean-squared error; and we use the reciprocal of this quantity as the quality measure to assess the level of success within the reconstruction. We go over all the mirrors sequentially. When we are handling a specific mirror during the iterations, we firstly calculate the current value of quality measure, Q_1 . Then, we switch the current tilt angle of the mirror arbitrarily to one of the remaining options, and calculate the new value of the quality measure, Q_2 . We calculate the ratio between these two quantities, $r = \frac{Q_2}{Q_1}$, which signals the effect of the change that is attempted. If the results improve with this modification ($r > 1$), the state of the mirror is updated to the new tilt position. On the other hand, if there is a deterioration in the results ($r \leq 1$), the change is performed with a probability of $rT(t)$, where $T(t)$ is a time-dependent weighting function. As stated before, all the mirrors on the chip will be experimented sequentially with this simple test, and their states will be altered appropriately if necessary. Usually, within the overall run, there will be several passages over the entire chip. Remembering from previous sections that we had $M \times N$ mirrors, if the number of passages over the entire chip is $K \in \mathcal{Z}^+$, there will be a total of $K \times M \times N$ mean squared error computations and comparisons during execution of the algorithm.

The weighting function is the essential part of the algorithm for the purpose of simulating the annealing part. Specifically, when the algorithm is running, this function has high values during the earlier stages, while its value decreases as the number of iterations increase. Since it directly influences the decision on whether a change will be made under degradation of results or not, the interpretation of its effect is as follows: at the beginning, even if results become worse, changes are allowed and enhanced. The aim here is to prevent the algorithm to converge to a local optimum, but rather provide it with a chance to step on a path through which it can arrive at the global optimum. Towards the end of the algorithm though, changes resulting in spoilage are barely allowed, this time to prevent the algorithm from getting out of the road that leads to the global optimum.

In our simulations, we selected $K = 40$. As for the weighting function, we used $T(t) = \exp\left\{-\frac{t}{4 \times M \times N}\right\}$. When the j 'th mirror was processed for the k 'th ($k \in \mathcal{Z}$, $1 \leq k \leq K$) time, this function was evaluated at $t = (k - 1) \times M \times N + j$, and the obtained value is assigned as the weighting factor to be used during the operations.

4.3. Simulation Results

We have carried out several simulations to illustrate and verify the usefulness of our formulations and results developed in the previous sections.

In our simulations, we represented the target field by its samples taken uniformly on a square planar patch residing in front of the DMAD. Through optimization algorithms, we tried to configure the DMAD to reconstruct these samples with minimum error. Our error criteria is the mean squared error of the difference signal between the samples of the original and reconstructed fields.

Our convention for exact placement of the sampling patch relative to the DMAD is illustrated in Fig. 4. The coordinate system here is the same as in Fig. 2. First, we calculated the dimensions of the square region on the $z = 0$ plane in which the DMAD exactly fits. This region is shown as R_1 in Fig. 4, with dimensions $d_1 \times d_1$. The distance of the sampling patch R_2 to the DMAD is taken to be $z = d_1$, while the dimensions of the sampling patch are chosen as $d_2 = 0.75d_1$.

The simulations were carried out separately for two DMADs of sizes 25×25 and 71×71 . In both cases, the mirrors were $16\mu\text{m} \times 16\mu\text{m}$ in size, and they could be tilted to $-12,0$ and 12 degrees (their axes of rotation is parallel to the y -axis in Fig. 4) and the interspacing between the mirrors was $1\mu\text{m}$. For the case of 25×25 DMAD, all the light fields were represented by 11×11 uniformly distributed samples on the sampling patch. The distance of the patch to the mirrors was $z = 300\mu\text{m}$, while the sampling interval in both x and y directions were $dx = dy = 45\mu\text{m}$. When we used instead the 71×71 DMAD, the sampling patch consisted of 25×25 uniformly distributed samples, the distance of the patch to the mirrors turned out to be $z = 853\mu\text{m}$, while the sampling intervals became $dx = dy = 53\mu\text{m}$.

In all the experiments, the DMADs were illuminated by normally incident plane waves of wavelength $\lambda = 633nm$. To arrive at the format suggested by Eq. 16, we firstly prepared the matrix \mathbf{D} by computing the samples of the fields produced by mirrors of the chip on the sampling patch. We made extensive use of the formulas developed in sec. 2 for that purpose. In our experiments with the 25×25 DMAD, \mathbf{D} had a size of 121×1875 . The row size follows since the sampling patch consists of 11×11 samples, resulting in a total of 121 samples. The column size is obtained as $675 \times 3 = 1875$ since there are a total of 625 mirrors and each mirror can be tilted to three different angles. In the case of the 71×71 device, the size of \mathbf{D} became 625×15123 .

With both the 25×25 and 71×71 DMADs, we conducted the simulations for two target fields. One of these fields was randomly selected within the range space of the device, so that the optimization problem had an exact solution. The other one was chosen to be circle function. We tried to reconstruct the samples of this circular field. In this case, we tried to find out the configuration that produced the closest set of samples to the original field in minimum mean-squared sense. The performance of the proposed algorithms was evaluated with the percentage mean squared error error. It is computed by dividing the mean square error of reconstruction to the total energy of the original samples, and multiplying the result by 100.

In the first experiment a 25×25 DMAD was used to to synthesize a field which was in its range space, where the optimal tilt positions of the mirrors were found by using first- and second-order matching pursuit algorithms. Fig. 5 shows the progress of the percentage error along with the number of iterations. (Note that in the case of matching pursuit algorithm, the total number of iterations equal the total number of mirrors.) The dashed curve is associated with first-order matching pursuit, while the solid one shows the error when second-order matching pursuit is used. These curves illustrate the performance of MP as a fast approximation strategy using minimal number of dictionary elements. Both versions of the algorithm arrive at an approximate solution - 0.8% and 0.5% error for first and second order matching pursuit respectively. In both cases, this lowest value is reached after the 400th iteration - less than the total number of mirrors. After a good approximate solution has been found with minimum number of mirrors, adding more mirrors increases the error when they are combined with the mirrors selected so far. This due to the fact that the expansion coefficients in 17 are restricted only to unity and the portion of reflected light from a mirror cannot be controlled to minimize the error further. On the contrary, when it is strictly unity it degrades the approximation when combined with the other mirrors. Indeed, the final errors yielded by the algorithms were 5.5% and 4.3%. Another important observation is that for any number of mirrors the second order MP leads to lower error. These suggests that the use of global optimization algorithm will be of benefit.

We also tried to reconstruct the field above using simulated annealing. This algorithm and the solution it converges is inevitably influenced by the initial estimate begun with. At first, we ran the algorithm a hundred times in which for each run we randomly selected an initial configuration, and among the hundred results, we looked at the most satisfactory one. In this trial, we were able to recover the exact configuration that produced the input field. However, in most of the trials the algorithm is stuck in a local minimum and finds only an approximate solution. Therefore we lack a systematic approach for selection of the initial conditions which will lead in most of the cases to the global optimum.

For the 25×25 DMAD, another group of simulations were performed when the target field had a circular distribution of $1.5circ\left(\frac{\sqrt{\left(\frac{x}{dx}\right)^2 + \left(\frac{y}{dy}\right)^2}}{2}\right)$. This field was not necessarily in the range space of the device. Results with first- and second-order matching pursuit algorithms are plotted in Fig. 5. Still the performance of second-order algorithms is better - the errors are 1.0% (MP) and 0.7% (MP2) percent during the iterations. However, this time the final errors produced by the algorithms had significantly intolerable values - 78.2% (MP) and 74.2% (MP2). We also applied simulated annealing. Firstly, similar to the previous case, we initiated the algorithm a hundred times with randomly selected initial conditions, and searched for the best result. Via this algorithm, the target field was reconstructed up to an error of 13.3%. This shows that if only an approximate solution is theoretically possible, MP is superior to simulated annealing as a greedy algorithm that minimizes the error without trying to recover the underlying structure.

The same sequence of simulations, with some minor changes, were applied to the 71×71 DMAD. We firstly tried to regenerate a field within the range space. This time, due to its high computational complexity, we avoided the application of second-order matching pursuit, but concentrated only on the first-order version. The

propagation of error is demonstrated in Fig. 7. The curve is similar in shape to that in the previous experiments. The final value of the error is 8.8% while the minimum value reached is 0.7%. On the other hand, we observed that when several random initial conditions were applied, simulated annealing recovered the field perfectly. Secondly, we tried to synthesize a circular field distribution of $1.5\text{circ}\left(\frac{\sqrt{\left(\frac{x}{dx}\right)^2+\left(\frac{y}{dy}\right)^2}}{2}\right)$. The error curve of first-order matching pursuit algorithm is plotted in Fig. 8. Similar to the counterpart of this simulation for the case of 25×25 DMAD, the final error is unacceptably high (46.5%) in spite of the very low value (0.7%) the error reaches during the iterations. We also applied simulated annealing a hundred times with different initial conditions each time. The best reconstruction we obtained is shown in Fig. 9, while the corresponding error versus iteration number curve is plotted in Fig. 10. (Note that in the case of simulated annealing, the total number of iterations is equal to the multiplication of the total number of mirrors and the total number of passes over the entire device.) The final error is 5.8%.

Consideration of results leads us to the following observations: We infer from the experiments in which the desired field is within the range space of the DMAD that the simulated annealing algorithm is able to perform perfect reconstruction (given good initial conditions). However, we lack a systematic and quick approach to supply good initial estimates that would enable convergence to the global optimum. Another alternative might be to search for parameters of simulated annealing that would yield better results. Our choices here were based on experimental trials instead of more systematic approaches, so they may not be optimal. On the other hand, when considered on its own, the matching pursuit algorithm produces high error values at the end despite the fact that the error goes very low during the iterations. As explained above this is due to the restriction for the coefficients in the expansion 17. However, the experiments with circular fields suggest that the performance of MP does not depend on whether the desired field is theoretically reproducible by the DMAD. Moreover, the convergence time may be very short. Under the assumption that only part of the mirrors can be used to synthesize the field the performance of the matching pursuit approach is acceptable, and the algorithm itself might turn out to be very handy.

5. CONCLUSION

In this work, we addressed the problem of complex monochromatic light field synthesis with a deflectable mirror array device(DMAD). With their large number of degrees of freedom, DMADs are promising spatial light modulators to be used in light field synthesis applications, but they raise challenging inverse problems: finding the optimal position of the mirrors of the DMAD is a difficult optimization problem. The major contribution of this work is the formulation of the problem as a constrained linear optimization problem which can be numerically solved in a fast manner by using any of the dozens of readily existing algorithms for such problems. Here, the matching pursuit and simulated annealing algorithms were tested. They were chosen since they are fast and can be easily adapted to the constraints of the problem. According to our results, MP shows worse behavior compared to simulated annealing, in the sense that at the end, the configuration suggested by simulated annealing results in a better reproduction of the desired field than the one suggested by matching pursuit. However, if only a part of the mirrors can be used, and the rest can be turned off, the performance of MP is acceptable and it turns out to be stable for different types of fields. On the other hand, the simulated annealing algorithm, with its low computational complexity, exhibits a strong potential to produce good results when it is supplied with good initial conditions. However, at the moment, we are neither aware of systematic approaches for examining, assessing and speculating on the prospective impact of initial conditions to the final result of the algorithm, nor on ways of producing such good initial estimates. These problems must be undertaken in future works.

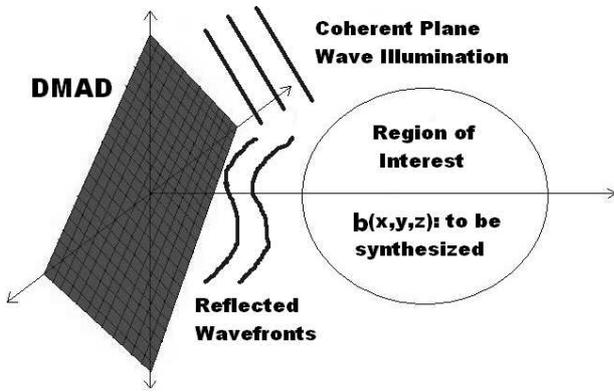


Figure 1. Light field synthesis with a DMAD

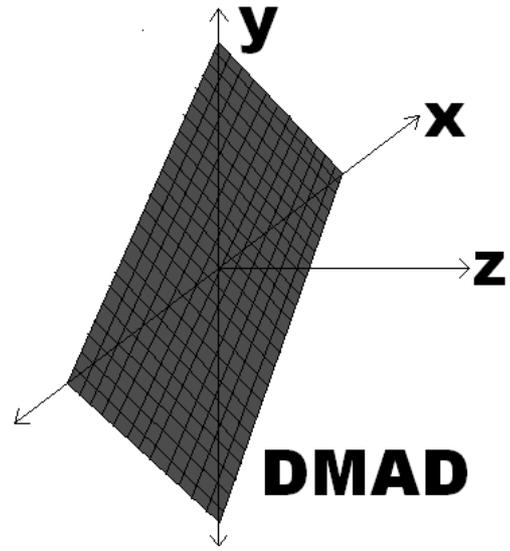


Figure 2. Chosen coordinate system

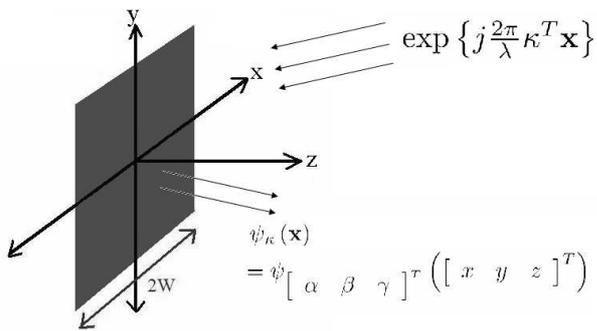


Figure 3. Physical interpretation of $\psi_{\kappa}(\mathbf{x})$ of Eq. 5

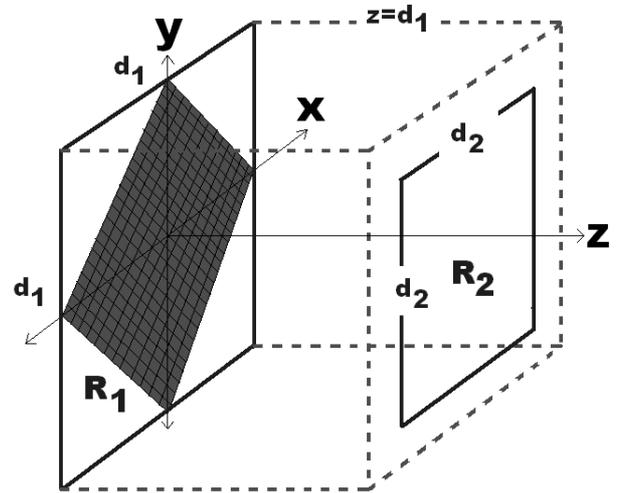


Figure 4. Placement of sampling patch with respect to DMAD

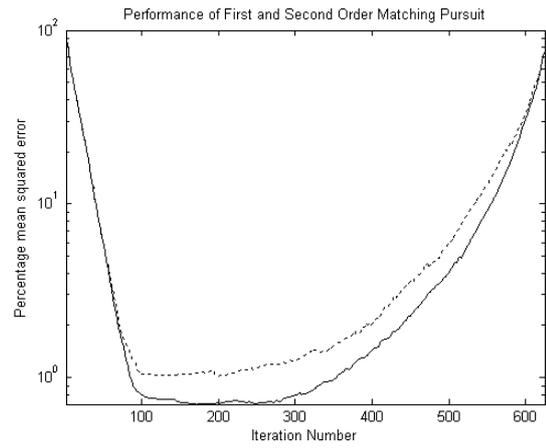
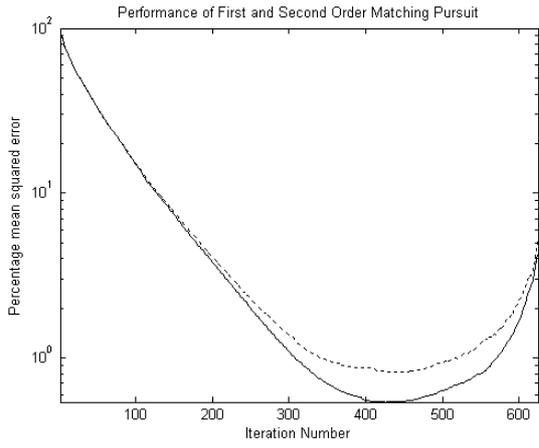


Figure 5. Error of first and second-order MP for field in range space **Figure 6.** Error of first and second-order MP for circular field

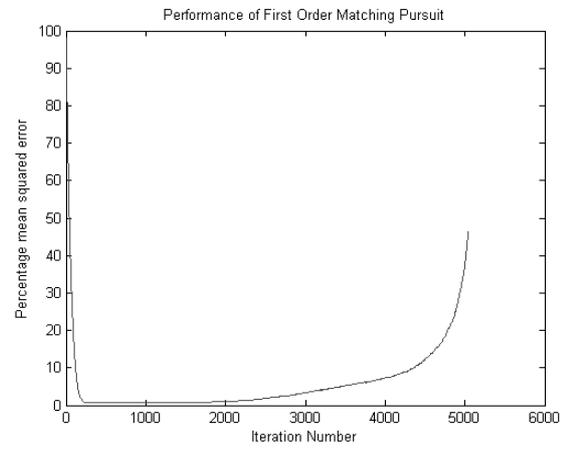
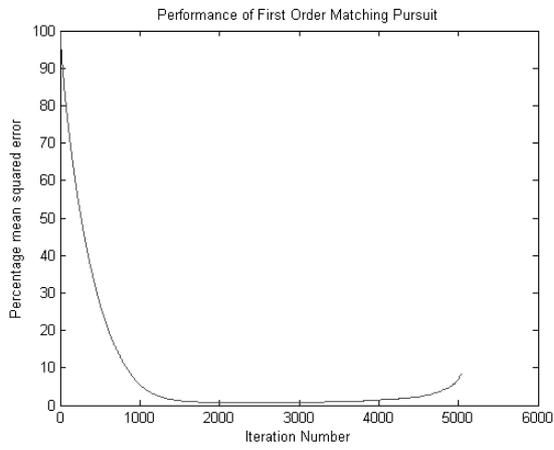


Figure 7. Error of first-order MP for field in range space

Figure 8. Error of first-order MP for circular field

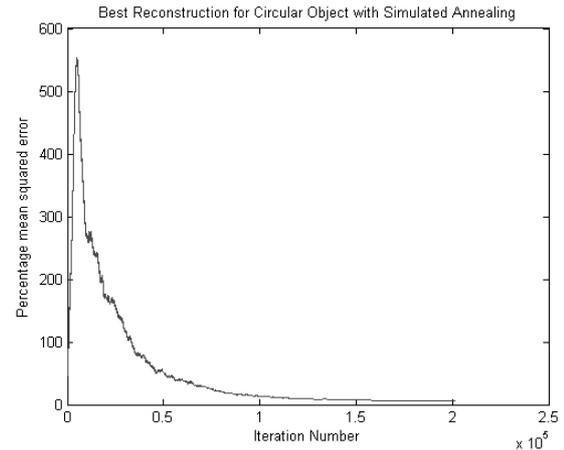
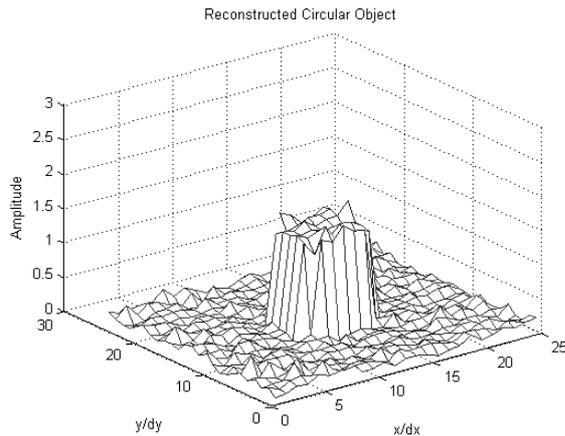


Figure 9. Reconstruction of circular object with simulated annealing **Figure 10.** Error of simulated annealing for circular field annealing

REFERENCES

1. R. J. Gove, "DMD display systems: the impact of an all digital display," in *Int Symp Society for Information Display*, Texas Instruments White Pages, 1994.
2. L. J. Hornbeck, "Digital light processing update: status and future applications," in *Conf on Projection Displays V, Proc SPIE*, **3634**, pp. 158–170, 1999.
3. R. S. Nesbitt, S. L. Smith, R. A. Molnar, and S. A. Benton, "Holographic recording using a digital micromirror device," in *Conf on Practical Holography XIII, Proc SPIE*, S. A. Benton, ed., **3637**, pp. 12–20, 1999.
4. T. Kreis, P. Aswendt, and R. Hofling, "Hologram reconstruction using a digital micromirror device," *Optical Engineering* **40**, pp. 926–933, 2001.
5. R. Piestun and J. Shamir, "Synthesis of three-dimensional light fields and applications," *Proceedings of the IEEE* **90**, pp. 222–244, 2002.
6. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**, pp. 671–681, 1983.
7. P. J. M. V. Laarhoven and E. H. L. Aarts, *Simulated annealing: theory and applications*, D. Reidel, 1987.
8. S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Transactions on Signal Processing* **41**, pp. 3397–3415, 1993.
9. E. Ulusoy, L. Onural, and H. M. Ozaktas, "Analysis of the complex light field generated by a deflectable mirror array device," in *International Conference on Holography, Optical Recording and Processing of Information*, (Varna, Bulgaria), 2005. to appear in Proc SPIE.
10. J. W. Goodman, *Introduction to Fourier Optics*, Mc-Graw-Hill, New York, 1996.
11. L. Onural, "Sampling of the diffraction field," *Applied Optics* **39**, pp. 5929–5935, 2000.
12. F. Gori, "Fresnel transform and sampling theorem," *Optics Communications* **39**, pp. 293–297, 1981.

Publication P2

V. Uzunov, A. Gotchev, G.B. Esmer, H.M. Ozaktas, L. Onural, “Non-uniform sampling of diffraction field,” in *Proceedings of The 2006 International Workshop on Spectral Methods and Multirate Signal Processing, SMMSP2006, TICSP Series*, vol. 34, Florence, Italy, 2-3 September 2006, pp. 191-197.

NON-UNIFORM SAMPLING AND RECONSTRUCTION OF DIFFRACTION FIELD

Vladislav Uzunov¹, Atanas Gotchev¹, G. Bora Esmer², Levent Onural², and Haldun M. Ozaktas²

¹ Tampere University of Technology, Institute of Signal Processing, FIN-33101 Tampere, Finland

² Bilkent University, Department of Electrical and Electronics Engineering,
TR-06800 Bilkent, Ankara, Turkey

ABSTRACT

Reconstruction of the diffraction field from non-uniformly sampled data points is an important signal processing problem in 3D display of a scene. The plane wave decomposition approach is traditionally used to calculate the diffraction between two parallel planes. Here, we use its discrete form to create a finite dimensional model of continuous diffraction fields. The model is used to state the problem of field reconstruction as matrix inversion. When the analysis is carried out in frequency domain, a symmetric reconstruction matrix can be formed. Its inversion is simplified by fast iterative approach, based on the conjugate gradient (CG) method. Numerical experiments demonstrate the applicability and accuracy of the chosen approach. The CG method converges in a small number of iterations and requires much less computational costs compared to direct matrix inversion.

1. INTRODUCTION

The computation of the light field distribution which arises over the entire three-dimensional (3D) space from an abstract representation of a 3D scene is known as the *forward problem* in holography. This problem is important in holographic three-dimensional television (3DTV), since the display device has to be controlled by signals which depend on the optical field of the scene. Solving the forward problem is a challenging task, since the 3D scene may be a complex structure consisting of opaque, transparent or semi-transparent surfaces and lines. In a more general setting, the sampled representation of the scene can be assumed to be in the form of discrete data points, non-uniformly distributed over the 3D space. Thus, the forward problem becomes a problem of diffraction field reconstruction, given a set of irregularly distributed samples. In this work we investigate the problem and suggest a fast numerical solution.

If the optical field is known on a certain plane, it can be calculated at any other parallel plane using a linear, shift-invariant relation [1, 2, 3]. Once the field at any plane is sampled, efficient numerical approaches can be utilized for fast computation of the whole diffraction field [4, 5]. The sampling problem in optics has been investigated in the literature. The field samples at a certain plane

are chosen so that the function which generates the field can be recovered perfectly. Gori [6], showed that a function can be reconstructed from its samples, provided that the function is α -Fresnel limited. Onural [7], derived conditions and approaches for recovery of the optical field in the cases of space-limited and band-limited functions. Stern and Javidi [8], considered in addition the effects of finite aperture size of the recording sensor.

In our earlier work [9], we have addressed the problem of reconstruction of optical field from arbitrary distributed samples, taken from uniformly sampled planes. The problem was analyzed theoretically and solutions which are based on direct matrix inversion and projection onto convex sets (POCS) [10, 11] were suggested. In the present work, we generalize the problem for non-uniform sampling along x and z coordinates and build a Fourier domain reconstruction matrix. Its symmetric structure allows us to use an efficient inversion algorithm, i.e. the conjugate gradient optimization algorithm [12, 13].

2. PROBLEM FORMULATION

We consider optical fields generated by monochromatic (single wavelength) light waves, propagating in free space [1, 2, 3]. The free space as propagation medium is linear, isotropic, and homogeneous. Under these conditions, the optical field on one plane can be accurately related to that on another parallel plane through the Rayleigh - Sommerfeld diffraction integral, which is a linear and shift-invariant relationship [1, 2, 3]. We assume that the distances involved are $r \gg \lambda$, where λ is the wavelength. Under these conditions, the Rayleigh - Sommerfeld diffraction integral and the so-called plane wave decomposition approach are equivalent [14, 15]. In the rest of the paper the plane wave decomposition approach is adopted. For simplicity, the discussions are restricted to one transverse dimension only. Extension to two transverse dimensions is straightforward. The diffraction integral over 2D space arising from the plane wave decomposition approach is

$$u(x, z) = \int_{-\frac{2\pi}{\lambda}}^{\frac{2\pi}{\lambda}} a(k_x) e^{j(k_x x + k_z z)} dk_x, \quad (1)$$

where $u(x, z)$ is the field over 2D space, and $a(k_x)$ is the Fourier spectrum of the field $u(x, 0)$ on the reference

This work is supported by EC within FP6 under Grant 511568 with acronym 3DTV.

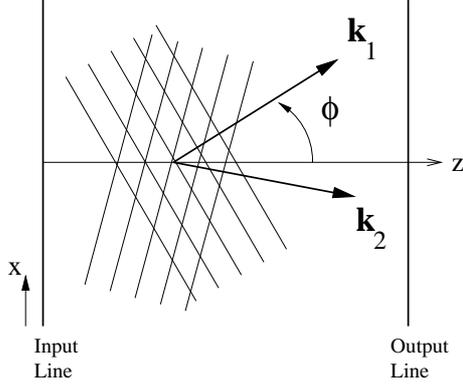


Figure 1. The vectors $\mathbf{k}_1 = [k_{1x}, k_{1z}]$ and $\mathbf{k}_2 = [k_{2x}, k_{2z}]$ are the wave vectors of the plane waves. The field at the output line is equal to the superposition of the plane waves illuminated on the input line according to Eq.1 [9].

line $z = 0$. The x axis is the transverse axis and the z axis is the optical axis along which the field propagates. The variables k_x and k_z are the spatial frequencies of the propagating plane waves along the x and z axes respectively, and form the wave vectors $\mathbf{k} = [k_x, k_z]$ (see Figure 1). The variable k_z is related to k_x by $k^2 = k_x^2 + k_z^2$ where $k = \frac{2\pi}{\lambda}$ is the length of the wave vector. A propagating monochromatic wave with wavelength λ cannot have a harmonic component in the transverse plane with higher frequency than $\frac{2\pi}{\lambda}$ since $k_x = k \cos \phi$ (cf. Figure 1). Therefore, in the case of free space propagation of monochromatic light, the function on the reference line $u(x, 0)$ is assumed bandlimited. This also determines the limits of the integral in Eq. 1.

In order to address numerically the problem of reconstruction of the field from finite number of non-uniformly distributed samples, a *finite dimensional model* of Eq. 1 has to be created. In Eq. 1 every point of the field $u(x, z)$ can be expressed in terms of the spectrum $a(k_x)$ of the field $u(x, 0)$ on the reference line. Consider an *essentially* space-limited function $u(x, 0)$, or, more precisely, a function whose space - bandwidth product is finite, and denote the spatial extent of interest by T . Such a function can be periodized, which is equivalent to discretization of $a(k_x)$. A periodic and bandlimited function is equivalently represented by a finite number of discrete spectral components. That is, the discrete values of the frequency k_x are :

$$k_x = \frac{2\pi m}{T}, \quad (2)$$

where $m = -\lfloor \frac{M}{2} \rfloor, \dots, \lfloor \frac{M-1}{2} \rfloor$. Substituting k_x as given by Eq.2 and $k = 2\pi/\lambda$ in Eq.1 leads to the finite dimensional model of diffraction:

$$\begin{aligned} u(x, z) &= \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j\sqrt{k^2 - k_x^2} z} e^{jk_x x} \\ &= \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j\frac{2\pi}{T}\sqrt{\beta^2 - m^2} z} e^{j\frac{2\pi}{T} m x}, \quad (3) \end{aligned}$$

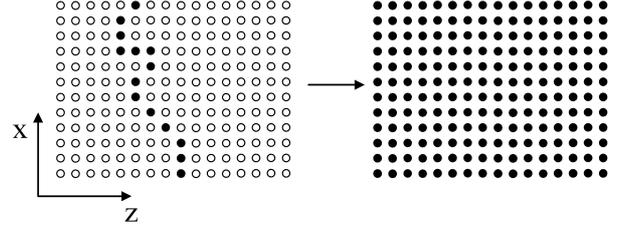


Figure 2. An example of irregularly sampled data representing a curved line. Filled circles represent given data points, blank circles represent unknown samples

where $\beta = \frac{T}{\lambda}$ and $a_m = a(\frac{2\pi}{T}m)$ are the coefficients of the Fourier series expansion of $u(x, 0)$.

The finite dimensional model in Eq.3 can be used to calculate the diffraction field at any point (x, z) from the diffraction field on the reference line $u(x, 0)$, represented by its Fourier coefficients a_m . In the general case of irregular sampling, the diffraction field $u(x, z)$ is given at a finite set of s sampling points $\{(x_i, z_i)\}_{i=1}^s$. Therefore, the irregular sampling and reconstruction problem can be stated as to find the Fourier coefficients a_m given the samples $u(x_i, z_i), i = 1, \dots, s$. A system of s linear equations for a_m can be constructed by writing Eq. 3 for each point in the irregular sampling set $\{(x_i, z_i)\}_{i=1}^s$:

$$u(x_i, z_i) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j\frac{2\pi}{T}\sqrt{\beta^2 - m^2} z_i} e^{j\frac{2\pi}{T} m x_i}, \quad (4)$$

for $i = 1, \dots, s$.

The problem of diffraction field reconstruction from irregularly distributed samples provides a general framework for computation of the diffraction field generated by an arbitrary-shaped object (Figure 2). This problem requires a rather novel solution, since most of the known methods deal with the calculation of the diffraction between parallel or tilted lines only [5, 16, 17].

3. RECONSTRUCTION ALGORITHM

The system in Eq. 4 is linear and can be expressed in a matrix form:

$$\mathbf{u} = \mathbf{R}\mathbf{a}, \quad (5)$$

where $\mathbf{a} = [a_{-\lfloor \frac{M}{2} \rfloor}, a_{-\lfloor \frac{M}{2} \rfloor+1}, \dots, a_{\lfloor \frac{M-1}{2} \rfloor}]^T$ is the unknown vector of the Fourier coefficients and the vector of given samples is $\mathbf{u} = [u(x_1, z_1), u(x_2, z_2), \dots, u(x_s, z_s)]^T$. \mathbf{R} is the reconstruction matrix

$$\mathbf{R} = \{r_{p,q}\}, \quad (6)$$

where $r_{p,q}, p = 1, \dots, s, q = 1, \dots, M$ is the matrix element

$$r_{p,q} = e^{j\frac{2\pi}{T}\sqrt{\beta^2 - (q - \lfloor \frac{M}{2} \rfloor + 1)^2} z_p} e^{j\frac{2\pi}{T} m x_p}. \quad (7)$$

The diffraction field at the point (x_i, z_i) equals the the inner product of the i -th row of \mathbf{R} with \mathbf{a} .

The straightforward approach to solve Eq.5 is to take the pseudo-inverse \mathbf{R}^\dagger of \mathbf{R} :

$$\mathbf{a} = \mathbf{R}^\dagger \mathbf{u}. \quad (8)$$

In the case when $s \geq M$ the system in Eq.5 is over-determined and the pseudo-inverse of \mathbf{R} has the form:

$$\mathbf{R}^\dagger = (\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H, \quad (9)$$

where \mathbf{R}^H is the conjugate transpose of \mathbf{R} . When $s < M$ the system in Eq.5 is under-determined. If the data in \mathbf{u} is consistent, then infinitely many solutions exist. However, there is no way to find among them the true underlying vector \mathbf{a} which generated the field. In this case the pseudo-inverse of \mathbf{R} is

$$\mathbf{R}^\dagger = \mathbf{R}^H (\mathbf{R} \mathbf{R}^H)^{-1}. \quad (10)$$

and Eq. 8 produces a minimum-norm solution for \mathbf{a} . Still, there is no guarantee that this is the solution of interest.

The major drawback of solving the problem by computing the pseudo-inverse is the high computational costs. In the over-determined case pseudo-inversion is usually done by computing QR-decomposition of \mathbf{R} , while in the under-determined case it is recommended to use singular value decomposition (SVD) of \mathbf{R} [12]. Even the most efficient numerical approaches for QR or SVD have cubic complexity [18].

Therefore, we need a fast and numerically stable algorithm to find the (approximate) solution of the following (LS) problem:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathcal{R}^M} \|\mathbf{u} - \mathbf{R}\mathbf{a}\|_2 \quad (11)$$

In linear programming [13] there exists a myriad of iterative algorithms for solving LS problems [13, 12]. However, most of them have slow convergence when applied directly to LS problems with large condition number of the matrix. Nevertheless, an important observation in Eq. 9 and Eq. 10 helps to avoid this difficulty. The costly operation in the equations is the inversion of the matrices $\mathbf{Q}' = \mathbf{R}^H \mathbf{R}$ in Eq. 9 and $\mathbf{Q}'' = \mathbf{R} \mathbf{R}^H$ in Eq. 10. We observe that both \mathbf{Q}' and \mathbf{Q}'' are symmetric:

$$\begin{aligned} (\mathbf{Q}')_{p,q} &= \mathbf{c}_p^H \mathbf{c}_q = (\mathbf{c}_q^H \mathbf{c}_p)^H = (\mathbf{Q}')_{q,p}^H; \\ (\mathbf{Q}'')_{p,q} &= \mathbf{r}_p \mathbf{r}_q^H = (\mathbf{r}_q \mathbf{r}_p^H)^H = (\mathbf{Q}'')_{q,p}^H. \end{aligned} \quad (12)$$

Here \mathbf{c}_p and \mathbf{c}_q are the p -th and q -th columns of \mathbf{R} , and \mathbf{r}_p and \mathbf{r}_q are the p -th and q -th rows of \mathbf{R} respectively. The symmetry of \mathbf{Q}' and \mathbf{Q}'' allows taking advantage of the *conjugate gradient* (CG) method to invert them. This method is known to be one of the most powerful and flexible methods [13]. It converges in at most n iterations for n - dimensional problems [12]. For the over-determined case the dimensionality of the problem is $n = M$, while for the under-determined case it is $n = s$.

A single formulation of the inversion problem can be stated for both under - and over-determined cases:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{R}^n} \|\mathbf{b} - \mathbf{Q}\mathbf{x}\|_2, \quad (13)$$

When $s > M$ this formulation can be derived by multiplying both sides of Eq.5 with \mathbf{R}^H to get

$$\mathbf{R}^H \mathbf{R} \mathbf{a} = \mathbf{R}^H \mathbf{u}. \quad (14)$$

The following substitutions in Eq.14

$$\mathbf{Q} = \mathbf{Q}' = \mathbf{R}^H \mathbf{R}, \quad \mathbf{x} = \mathbf{a} \quad \text{and} \quad \mathbf{b} = \mathbf{R}^H \mathbf{u} \quad (15)$$

lead to $\mathbf{Q}\mathbf{x} = \mathbf{b}$ which can be restated as in Eq. 13.

In the under-determined case the minimum norm solution for \mathbf{a} lies in the row space of the matrix \mathbf{R} [12]. Therefore, it can be written as the linear combination:

$$\mathbf{a} = \sum_{i=1}^s v_i \mathbf{r}(x_i, z_i) = \mathbf{R}^H \mathbf{v}, \quad (16)$$

where \mathbf{v} is the vector of coefficients. Using this fact, Eq. 5 can be rewritten as:

$$\mathbf{R} \mathbf{R}^H \mathbf{v} = \mathbf{u}. \quad (17)$$

Now denote

$$\mathbf{Q} = \mathbf{Q}'' = \mathbf{R} \mathbf{R}^H, \quad \mathbf{x} = \mathbf{v} \quad \text{and} \quad \mathbf{b} = \mathbf{u} \quad (18)$$

to get the same form $\mathbf{Q}\mathbf{x} = \mathbf{b}$, formally stated by Eq. 13. Note that in the over-determined case solving the LS problem in Eq. 13 directly outputs the vector of Fourier coefficients $\hat{\mathbf{a}}$, while in the under-determined case $\hat{\mathbf{x}}$ must be multiplied by \mathbf{R}^H to get $\hat{\mathbf{a}}$.

The CG algorithm for the general case proceeds as follows:

1. compute the matrix \mathbf{R} according to Eq.6 and Eq.7;
2. If $s < M$ compute \mathbf{Q} and \mathbf{b} according to Eq. 18, otherwise compute them according to Eq. 15.
3. initialize $\hat{\mathbf{x}}^{[0]}$ arbitrary, $\mathbf{d}_0 = \mathbf{b} - \mathbf{Q}\hat{\mathbf{x}}^{[0]}$ and $\mathbf{g}_0 = -\mathbf{d}_0$
4. for $n = 1$ to $n_{it} \leq M$
 - (a) $\alpha_n = -\frac{\mathbf{g}_n^T \mathbf{d}_n}{\mathbf{d}_n^T \mathbf{Q} \mathbf{d}_n}$
 - (b) $\hat{\mathbf{x}}^{[n+1]} = \hat{\mathbf{x}}^{[n]} + \alpha_n \mathbf{d}_n$
 - (c) $\mathbf{g}_{n+1} = \mathbf{Q}\hat{\mathbf{x}}^{[n+1]} - \mathbf{b}$
 - (d) $\gamma_n = -\frac{\mathbf{g}_{n+1}^T \mathbf{Q} \mathbf{d}_n}{\mathbf{d}_n^T \mathbf{Q} \mathbf{d}_n}$
 - (e) $\mathbf{d}_{n+1} = -\mathbf{g}_{n+1} + \gamma_n \mathbf{d}_n$
- end
5. If $s < M$ compute $\hat{\mathbf{a}} = \mathbf{R}^H \hat{\mathbf{x}}$, otherwise $\hat{\mathbf{a}} = \hat{\mathbf{x}}$.
6. reconstruct the diffraction field $u(x, z)$ from the estimated Fourier coefficient vector $\hat{\mathbf{a}}$ with Eq. 3.

For the case when $s < M$ the dimensionality of the problem is equal to s . In this case the CG algorithm is unable to restore true field and converges to the local minimum, nearest to the starting point. For the over-determined case the dimensionality of the problem, and hence the complexity of the reconstruction algorithm, depends only on the bandwidth M of the field and does not depend on the number of known samples s . However, the convergence rate of the algorithm is influenced by the vector \mathbf{b} .

Roughly speaking, the equation $\mathbf{b} = \mathbf{R}^H \mathbf{u}$ can be considered as inverse of Eq.4 and hence \mathbf{b} plays the role of an initial estimate of \mathbf{a} . Therefore, the closer \mathbf{b} is to \mathbf{a} , the faster the CG method converges to a solution. The goodness of the estimate \mathbf{b} is proportional to s and hence the convergence rate of the algorithm depends indirectly on s .

4. EXPERIMENTAL RESULTS

The approach outlined above is illustrated by experiments with two different optical fields. A simple bandlimited function on the reference line is constructed first and then the field on the other lines is generated by Eq. 3. The bandwidth of the two functions is bandlimited to 256 non-zero coefficients. The function on the reference line of the first field is a smoothed square pulse located at the center of the line with width approximately 1/8 of the assumed function duration. The function is constructed from a unit-magnitude square pulse, whose spectrum is windowed in Fourier domain with Kaiser window. In this way, the field-generating function is bandlimited and thus consistent with the finite dimensional model of Eq. 3. The function on the reference line of the second field is a Gaussian pulse located at the center of the line with similar width as in the previous function. We will refer to the first field as "rectangular field" and to the second field as "Gaussian field" for short. For the rectangular field the distance from the reference line z ranges between 0.5 and 1.5, while for the Gaussian field z ranges between 0.5 and 4.5. For numerical simulation, the fields are uniformly sampled at $L = 256$ lines with $N = 256$ samples per line. The simulated fields are depicted in Figure 3.

Assessment of the results is based on the normalized error between the original and reconstructed diffraction patterns $u(x, z)$ and $\hat{u}(x, z)$ respectively:

$$e = \frac{\|\hat{u}(x, z) - u(x, z)\|_2}{\|u(x, z)\|_2}. \quad (19)$$

For each of the diffraction fields the error has been tabulated for different number s of known data points as shown in Tables 1 and 2. First 10 columns of the tables show the normalized error when the fields are reconstructed with the CG algorithm. Each column corresponds to a certain number of CG iterations n_{it} . In this way the impact of n_{it} on the performance is illustrated. The last columns in Tables 1 and 2 show the error when the reconstruction is done by directly inverting the matrix \mathbf{R} according to Eq. 8. For each value of s , 15 diffraction patterns have been reconstructed using s randomly selected data points from the initial simulated fields. Each of the 15 reconstructed diffraction patterns corresponds to a different random choice of the positions of the s known data points within the simulated field. The final error reported for each value of s is the average of the error over all 15 choices.

As expected, the error decreases with increasing number of given points. A sufficient number of given points is crucial for the performance of the algorithm. This number turns out to be M since for $s \geq M$ direct inversion

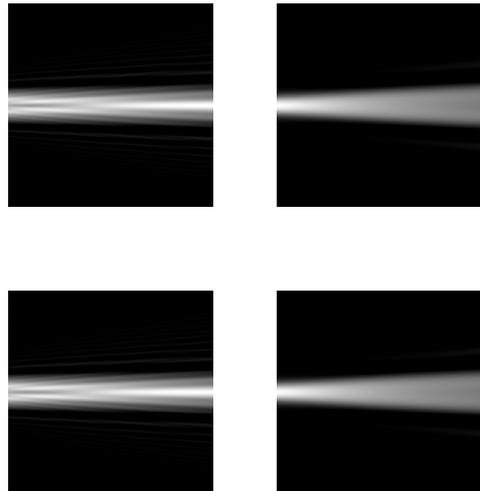


Figure 3. Simulated diffraction fields and reconstructions from $s = 282$ known data points. Top: simulated rectangular field (left) and Gaussian field (right). Bottom: reconstructed rectangular field (left) and Gaussian field (right).

yields practically zero error. However, the CG method needs slightly more points to find a good solution in reasonable number of iterations. When s is lower than M , the error takes higher values and cannot be reduced much by undertaking more iterations. This is expected since we have an under-determined system in Eq. 5. The CG algorithm converges to the minimum-norm solution, as the direct pseudo-inversion does.

For a certain number of CG iterations, there is a number of given points (drop-off value) providing the desired accuracy (error below a small threshold). Figure 4 illustrates this behavior for $n_{it} = 256$ iterations. The number of iterations used to construct this figure is chosen to be equal to the dimensionality of the problem, because this number is theoretically sufficient [12]. As expected, the drop-off value of the curve in Figure 4 is reached for values of s slightly higher than M .

The convergence of the CG method is illustrated in Figure 5 for fixed number of known data points. In general, the error decreases by the number of iterations. The algorithm converges faster for larger values of s as one can see from Tables 1 and 2.

The computational complexity of the proposed algorithm is mainly dependent on the computational complexity of the involved CG method. For each iteration in the CG algorithm, at least two multiplications of the matrix \mathbf{Q} with a vector are required. Their costs dominate over the costs of other operations. For the over-determined case this results in about $2M^2$ complex multiplications per iteration. The under-determined case is not considered here, because it shows poor reconstruction results and therefore

s	Conjugate gradient algorithm										Direct matrix inversion
	number of iterations n_{it}										
	10	20	30	50	100	150	200	300	500	1000	
77	0.7242	0.7161	0.7161	0.7161	0.7161	0.7161	0.7161	0.7161	0.7161	0.7161	0.7161
128	0.8218	0.5402	0.5395	0.5395	0.5395	0.5395	0.5395	0.5395	0.5395	0.5395	0.5395
205	2.8867	0.4141	0.2368	0.2163	0.2154	0.2152	0.2151	0.2150	0.2150	0.2150	0.2150
256	0.9420	0.5081	0.3497	0.2151	0.1882	0.1586	0.1417	0.1167	0.0886	0.0553	0.0000
282	0.5781	0.1810	0.0702	0.0112	0.0059	0.0019	0.0008	0.0002	0.0000	0.0000	0.0000
307	0.3650	0.0654	0.0125	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
333	0.2163	0.0199	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
384	0.0780	0.0017	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
512	0.0067	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 1. Reconstruction results for the rectangular field. The table shows the normalized error e for different numbers of given data points s . When CG algorithm is utilized different error values are shown for increasing number of iterations.

s	Conjugate gradient method										Direct matrix inversion
	number of iterations n_{it}										
	10	20	30	50	100	150	200	300	500	1000	
77	0.7142	0.7082	0.7082	0.7082	0.7082	0.7082	0.7082	0.7082	0.7082	0.7082	0.7082
128	1.0080	0.5255	0.5252	0.5252	0.5252	0.5252	0.5252	0.5252	0.5252	0.5252	0.5252
205	6.3613	0.5706	0.2363	0.2073	0.2072	0.2072	0.2072	0.2072	0.2072	0.2072	0.2072
256	1.9310	1.0258	0.6939	0.4449	0.3889	0.3309	0.2953	0.2458	0.1947	0.1335	0.0000
282	1.2134	0.3593	0.1248	0.0143	0.0058	0.0021	0.0008	0.0000	0.0000	0.0000	0.0000
307	0.7040	0.1050	0.0158	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
333	0.4191	0.0296	0.0017	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
384	0.1425	0.0025	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
512	0.0119	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 2. Reconstruction results for the Gaussian field. The table shows the normalized error e for different numbers of given data points s . When CG algorithm is utilized different error values are shown for increasing number of iterations.

is not of practical interest. To find the solution vector $\hat{\mathbf{a}}$ in n_{it} iterations, one needs $2n_{it}M^2$ complex multiplications plus sM complex multiplications for the initial estimate \mathbf{b} . The overall complexity in terms of complex multiplications is given by

$$2n_{it}M^2 + sM. \quad (20)$$

The number of given points s affects the complexity not only through the second term of Eq.21 but also indirectly through the number of iterations needed to achieve the desired accuracy. The latter dependence is illustrated in Figure 6, for a pre-specified reconstruction error of 0.0005. The curve drops right after $s = M$ and for larger values of s the desired error is reached in a smaller number of iterations. For example, less than 15 iterations are sufficient if starting with $s = 2M$ given points.

In contrast, when matrix inversion is utilized, the costs increase for higher s . The pseudo-inverse is computed according to Eq. 9, by applying QR-decomposition of \mathbf{R} . When efficient numerical approach for QR-decomposition is utilized, e.g. Householder transformations [18], the complexity is estimated as

$$sM^2 - \frac{M^3}{3} + \frac{M^2}{2}. \quad (21)$$

Figure 7 compares the computational complexity of the investigated algorithm with that of direct matrix inversion for certain number s of given samples in terms of complex multiplications. The dotted curve represents the computational complexity for reconstruction by pseudo-inversion, according to Eq. 21. The values for CG method are estimated by Eq. 20, where the number of iterations n_{it} is selected to ensure a normalized reconstruction error below 0.0005. Figure 7 shows that the proposed algorithm becomes to be less computationally expensive for $s > 1.2M$ and for sufficient number of points it might be up to 100 times faster.

Figure 3 shows reconstruction results for values of s close to M , with no visual difference between the simulated and reconstructed fields.

5. CONCLUSION

In this paper we have addressed the problem of reconstruction of diffraction field from irregularly distributed samples. Various forward problems in diffraction, directly related with the 3D scene display can benefit from its solution. Such a problem is calculation of the optical field, generated or reflected by arbitrary curved lines and surfaces. Our aim has been to reduce the computational costs

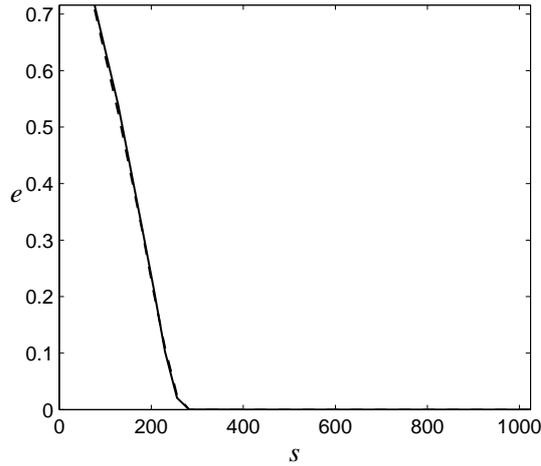


Figure 4. Normalized error e for different number of known samples s when the iterations are $n_{it} = 256$. Solid line stands for the rectangular field and dashed line stands for the Gaussian field.

in comparison with conventional methods for direct matrix inversion. We have taken advantage of the symmetric Fourier domain reconstruction matrix and have implemented a rapidly convergent iterative algorithm, based on the CG method. The numerical experiments show practically perfect reconstruction for small number of iterations. The proposed method outperforms the straightforward pseudo-inversion approach when the number of given samples is slightly larger than M .

6. REFERENCES

- [1] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [2] M. Born and E. Wolf, *Principles of Optics: Electromagnetic theory of propagation, interference and diffraction of light*, Cambridge University Press, New York, 1980.
- [3] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, John Wiley and Sons, Inc., 1991.
- [4] L.P. Yaroslavsky and N.S. Merzlyakov, *Methods of Digital Holography*, Consultance Bureau, New York, 1980.
- [5] L. Onural and P. D. Scott, "Digital decoding of in-line holograms," *Optical Engineering*, vol. 26, pp. 1124–1132, 1987.
- [6] F. Gori, "Fresnel transform and the sampling theorem," *Optics Communications*, vol. 39, pp. 293–297, November 1981.
- [7] L. Onural, "Sampling of the diffraction field," *Applied Optics*, vol. 39, pp. 5929–5935, November 2000.

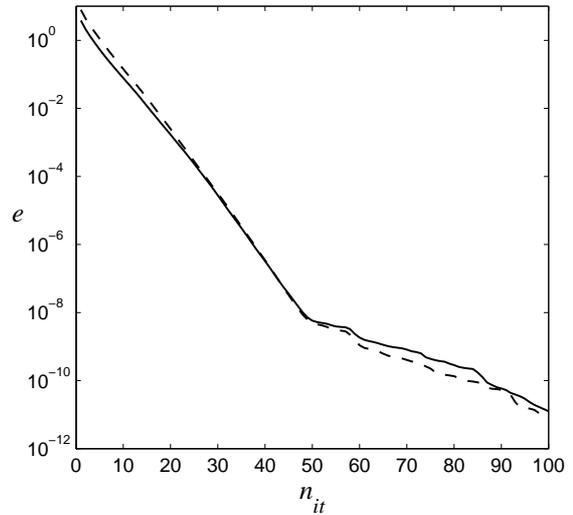


Figure 5. Convergence of the proposed reconstruction algorithm for both fields measured as normalized error e for increasing number of iterations n_{it} . Solid line stands for the rectangular field and dashed line stands for the Gaussian field. The number of known samples is $s = 384$.

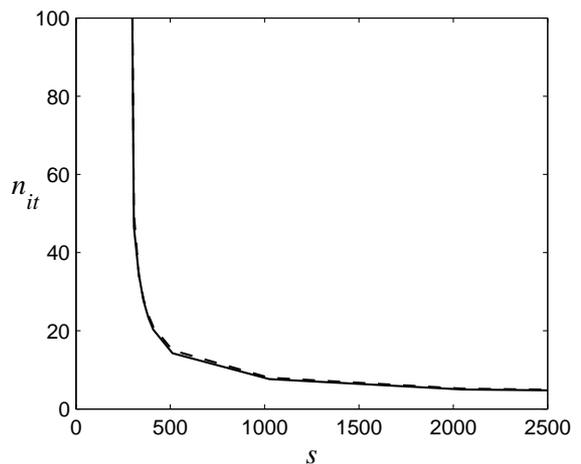


Figure 6. Number of iterations n_{it} needed to achieve error $e < 0.0005$ for different number s of given samples. Solid line stands for the rectangular field and dashed line stands for the Gaussian field.

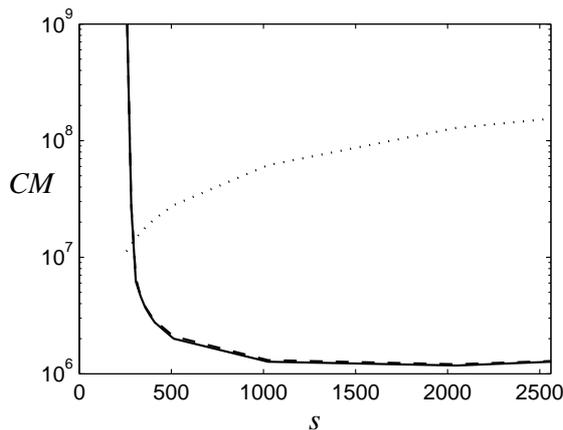


Figure 7. Comparison between the computational complexity of the CG method (dashed and solid curves) and matrix pseudo-inversion according to Eq.8 (dotted line). The curves represent the number of complex multiplications (CM) needed to achieve normalized error $e < 0.0005$. Solid line stands for the rectangular field and dashed line stands for the Gaussian field.

- [8] A. Stern and B. Javidi, "Analysis of practical sampling and reconstruction from fresnel fields," *Optical Engineering*, vol. 43, pp. 239–250, January 2004.
- [9] G. B. Esmer, V. Uzunov, L. Onural, H. M. Ozaktaş, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," *Image Communication*, vol. Special Issue on Three-Dimensional Video and Television.
- [10] R. Aharoni and Y. Censor, "Block iterative projection methods for parallel computation of solutions to convex feasibility problems," *Linear Algebra Applications*, vol. 120, pp. 165–175, 1989.
- [11] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput Math Math Phys*, vol. 7, pp. 1–24, 1967.
- [12] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*, Prentice Hall, Upper Saddle River, New Jersey, 2000.
- [13] S. G. Nash and A. Sofer, *Linear and nonlinear programming*, McGraw-Hill, 1996.
- [14] G. C. Sherman, "Application of the convolution theorem to rayleigh's integral formulas," *J. Opt. Soc. Am.*, vol. 57, pp. 546–547, 1967.
- [15] É. Lalor, "Conditions for the validity of the angular spectrum of plane waves," *J. Opt. Soc. Am.*, vol. 58, pp. 1235–1237, 1968.
- [16] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast fourier transform approach," *J. Opt. Soc. Am. A*, vol. 15, pp. 857–867, 1998.
- [17] K. Matsushima, H. Schimmel, and F. Wyrowski, "Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves," *J. Opt. Soc. A*, vol. 20, pp. 1755–1762, 2003.
- [18] Åke Björck, *Numerical Methods for Least Squares Problems*, SIAM, Amsterdam, Holland, 1990.

Publication P3

G.B. Esmer, **V. Uzunov**, L. Onural, H.M. Ozaktas, A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," in *Signal Processing: Image Communication* , vol. 22, No. 2, 2007, pp. 178-187.

©2007 Elsevier. Reprinted with permission.

Diffraction field computation from arbitrarily distributed data points in space[☆]

G. Bora Esmer^{a,*}, Vladislav Uzunov^b, Levent Onural^a,
Haldun M. Ozaktas^a, Atanas Gotchev^b

^a*Department of Electrical and Electronics Engineering, Bilkent University, TR-06800 Bilkent, Ankara, Turkey*

^b*Institute of Signal Processing, Tampere University of Technology, FIN-33101 Tampere, Finland*

Received 21 November 2006; accepted 29 November 2006

Abstract

Computation of the diffraction field from a given set of arbitrarily distributed data points in space is an important signal processing problem arising in digital holographic 3D displays. The field arising from such distributed data points has to be solved simultaneously by considering all mutual couplings to get correct results. In our approach, the discrete form of the plane wave decomposition is used to calculate the diffraction field. Two approaches, based on matrix inversion and on projections on to convex sets (POCS), are studied. Both approaches are able to obtain the desired field when the number of given data points is larger than the number of data points on a transverse cross-section of the space. The POCS-based algorithm outperforms the matrix-inversion-based algorithm when the number of known data points is large.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Scalar optical diffraction; Plane wave decomposition; Pseudo-matrix inversion; Projection onto convex sets

1. Introduction

Holographic three-dimensional television (3DTV) systems will consist of several parts fulfilling different functions: capturing the 3D scene, representing it in abstract form, its transmission, and finally display. Once the abstract representation of the scene arrives at the display end, it is necessary to first compute the diffraction field that the scene would have created, in order to drive the display device in a manner so that the same field will be

recreated. This paper deals with the solution of the problem of computing the diffraction field from the sampled abstract representation of the scene. This sampled representation consists of known values of the optical field over an irregularly distributed, arbitrary array of discrete points.

Diffraction problems are most commonly formulated such that the optical field at one plane is computed in terms of the field at another plane. This paper deals with computation of the field over all of space in terms of a set of given data points distributed arbitrarily over the space. It is not correct to calculate the optical field arising from such distributed data by means of straightforward superposition. This fact is often ignored and straightforward superposition is employed, which

[☆]This work is supported by EC within FP6 under Grant 511568 with acronym 3DTV.

*Corresponding author.

E-mail address: borahan@ee.bilkent.edu.tr (G.B. Esmer).

amounts to assuming that each data point represents a source [12,10]. Here we properly formulate the problem and consider its solution using two approaches. The first utilizes a direct matrix inversion while the second one is iterative and utilizes the projections onto convex sets (POCS) method [1,9], which has been successfully used for various restoration problems in image processing and holography [7,20,5].

2. Review of diffraction theory

The computation of the light field in a desired region, based on the knowledge of the field in some other region, is the subject of diffraction theory, which is a mature area of knowledge [8,3,17,16]. Despite this, there are gaps in the efficient numerical application of the theory and there seems to be considerable scope for application of signal processing techniques such as sampling, numerical linear algebra, fast transformations, iterative projections, and decomposition algorithms towards obtaining better computation methods.

In most cases, monochromatic (single wavelength) light is used in holography. Moreover, the medium we are interested in is linear, isotropic, and homogeneous. Under these conditions, the optical field on one plane can be accurately related to that on another plane through the Rayleigh–Sommerfeld diffraction integral, which is a linear and shift-invariant relationship [8,3,17].

In this study, we work with fields consisting of propagating waves; hence the input and output fields do not contain any evanescent wave components. Moreover, we assume that the distances involved are $r \gg \lambda$, where λ is the wavelength. Also, we use plane wave decomposition to compute scalar optical diffraction field, because plane wave decomposition and Rayleigh–Sommerfeld diffraction integral are equivalent [18,11]. For simplicity, we restrict our discussions to only one transverse dimension; extension to two transverse dimensions is straightforward. The diffraction integral over 2D space arising from the plane wave decomposition approach is

$$u_d(x, z) = \int_{-2\pi/\lambda}^{2\pi/\lambda} A(k_x) \exp[j(k_x x + k_z z)] dk_x, \quad (1)$$

where $u_d(x, z)$ is the field over 2D space, and $A(k_x)$ gives the complex amplitudes of the harmonic components (Fourier coefficients) of the field

$u_d(x, 0)$ over the input line. (The relationship between $(2\pi)A(k_x)$ and $u_d(x, 0)$ is the 1D Fourier transform (FT) relation.) The variables k_x and k_z are the spatial frequencies of the propagating plane waves along the x - and z -axis, respectively. The x -axis is the transverse axis and the z -axis is the optical axis along which the field propagates.

The variable k_z is related to k_x by $k_z = \sqrt{k^2 - k_x^2}$ where $k = 2\pi/\lambda$. The expression in Eq. (1) can be rewritten as

$$u_d(x, z) = \mathcal{F}^{-1}\{\mathcal{F}[u_d(x, 0)] \exp(j\sqrt{k^2 - k_x^2} z)\}, \quad (2)$$

where \mathcal{F} denotes the FT and \mathcal{F}^{-1} is the inverse FT.

The input field is a bandlimited spatial function, whose bandlimit is within $\pm k$. This is because a propagating monochromatic wave with wavelength λ cannot have a harmonic component in the transverse plane with higher frequency. In certain cases, it may be desirable to further restrict the bandwidth along the transverse direction x . For instance, we may restrict k_x such that $-B \leq k_x < B$, where $B \leq k$. So far, k_x may assume any real value in this interval. To arrive at a feasible numerical framework, we must restrict ourselves to a finite number N of possible values of k_x . These may be chosen as $k_x = l(2B/N)$ where $l = -N/2, \dots, N/2 - 1$ for even N and a similar formula for odd N . Discretizing the transverse frequency will result in a field which is periodic along the transverse direction x , with a fundamental period $X = \pi N/B$. Therefore, a careful choice of simulation parameters is necessary if the consequences of this periodicity effect are to be minimized. Therefore, the field becomes

$$u_d(x, z) = \sum_{m=0}^{N-1} A_m \exp(j\sqrt{k^2 - k_x^2} z) \exp\left(j\frac{2B}{N} mx\right), \quad (3)$$

where

$$k_x = \begin{cases} 2\pi\frac{m}{X}, & m = \left[0, \frac{N}{2}\right), \\ 2\pi\frac{(m-N)}{X}, & m = \left[\frac{N}{2}, N\right). \end{cases} \quad (4)$$

Sampling this periodic and bandlimited function with a sampling period $X_s = \pi/B$, which is its Nyquist rate, yields N samples per period, as expected. Restricting our attention to these samples, $x = nX_s$ and $z = pX_s$, where n is an integer

spanning one period, and p is the distance between the lines along the longitudinal direction z , we can write

$$u_d(nX_s, pX_s) = \sum_{m=0}^{N-1} A_m \exp(j\sqrt{k^2 - k_x^2} pX_s) \exp\left(j\frac{2\pi}{N}mn\right). \quad (5)$$

Each frequency component m , determines the propagation direction of the corresponding plane wave. The angle ϕ , which is shown in Fig. 1, denotes the angle between the z -axis and the propagation direction:

$$\phi_m = \begin{cases} \sin^{-1}\left(\frac{\lambda m}{NX_s}\right), & m = \left[0, \frac{N}{2}\right), \\ \sin^{-1}\left(\frac{\lambda(m-N)}{NX_s}\right), & m = \left[\frac{N}{2}, N\right). \end{cases} \quad (6)$$

Thus $m \in [0, N/2)$ corresponds to the positive ϕ angles, and $m \in [N/2, N)$ corresponds to the negative ϕ angles. The relation between k_z and m is

$$k_z = \begin{cases} \frac{2\pi}{NX_s} \sqrt{\beta^2 - m^2}, & m = \left[0, \frac{N}{2}\right), \\ \frac{2\pi}{NX_s} \sqrt{\beta^2 - (m-N)^2}, & m = \left[\frac{N}{2}, N\right), \end{cases} \quad (7)$$

where $\beta = NX_s/\lambda$. Therefore, the discrete kernel corresponding to the plane wave decomposition becomes,

$$H_p(m) = \begin{cases} \exp\left(j\frac{2\pi}{N} \sqrt{\beta^2 - m^2} p\right), & m = \left[0, \frac{N}{2}\right), \\ \exp\left(j\frac{2\pi}{N} \sqrt{\beta^2 - (m-N)^2} p\right), & m = \left[\frac{N}{2}, N\right) \end{cases} \quad (8)$$

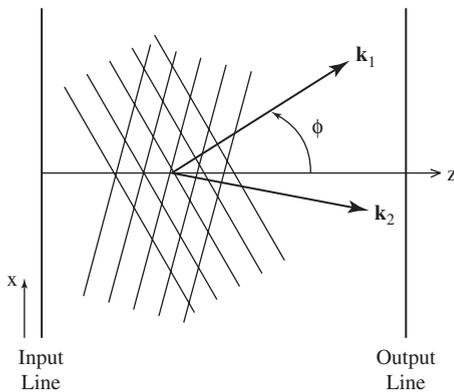


Fig. 1. The vectors \mathbf{k}_1 and \mathbf{k}_2 are the wave vectors of the plane waves.

and the resultant form of the discrete representation of the plane wave decomposition is

$$u(n, p) = \sum_{m=0}^{N-1} A_m H_p(m) \exp\left(j\frac{2\pi}{N}nm\right), \quad (9)$$

where $N \cdot A_m$ is the DFT of the sampled input field. The variable n is restricted to the range $[0, N)$. Therefore, the discrete diffraction field can be expressed as

$$u(n, p) = \text{DFT}^{-1}\{\text{DFT}\{u(n, 0)\}H_p(m)\}. \quad (10)$$

This completes the proper discretization of the diffraction formula.

3. Computation of diffraction patterns from distributed data points over 2D space

Computation of the diffraction pattern at the output arising from a distributed set of points can be time-consuming, due to the z dependency of the kernel. The diffraction relation between the distributed data set and a given line cannot be represented as a shift-invariant system. As already described, calculation of the diffraction field from one line to another is relatively straightforward since the corresponding system is linear and shift-invariant. The distributed points we wish to deal with may represent samples over a curved line, a tilted line, or another shape in 2D space. An illustration is given in Fig. 2. There are several fast methods which can be used to compute the diffraction field between parallel and tilted lines [15,19,4,14,13,6]. However, these methods cannot be used to compute the diffraction field from arbitrarily-distributed data points. In such a case, a simple and naive approach, i.e., calculation of the diffraction field by direct superposition assuming each

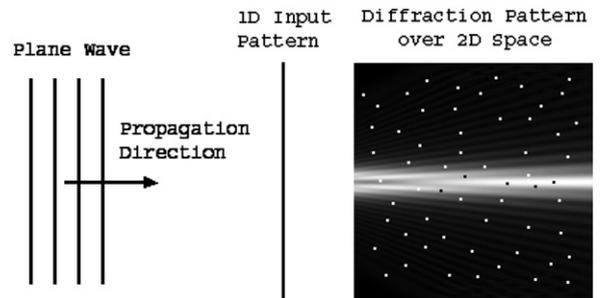


Fig. 2. An illustration of 1D object illumination and the diffraction pattern of the object over 2D space. Dots on the corresponding diffraction pattern represent the locations of a set of distributed known data points over 2D space.

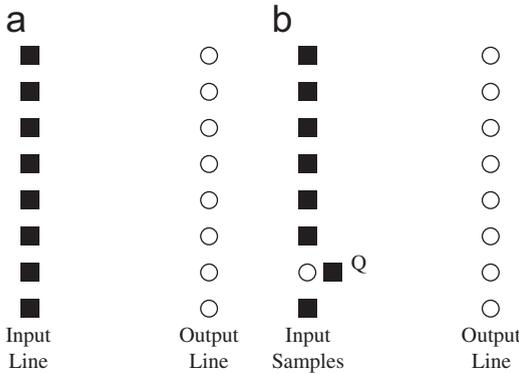


Fig. 3. (a) Parallel input and output lines. (b) Example involving a single displaced known data point.

data point as a source of light, would yield erroneous results. Two simple examples clarify this fact: In Fig. 3, black squares represent given samples while circles represent the missing ones. In the first example, the input and output lines are parallel to each other with no missing points (Fig. 3(a)). The relationship between the fields on these lines can be represented as a linear shift-invariant system, given by Eq. (10). In the second example (Fig. 3(b)), the field at one of the sample points on the input line is not known. Instead, the field at another sample point Q , not located on the input line, is given in order to compensate the missing data. Each point on the input line contributes to the field at Q , and in turn, the field at Q affects the field in other places. Therefore, it is not possible to write a proper superposition including the point Q for the field at unknown regions.

Let us assume a hypothetical straight line, referred to as the reference line, and compute the diffraction field on it. We rewrite Eq. (10) in matrix form:

$$\mathbf{g} = \mathbf{A}\mathbf{f}, \quad (11)$$

where the vectors \mathbf{f} and \mathbf{g} represent the discrete diffraction fields $u(n, 0)$ and $u(n, p)$ of the reference line and some other line, respectively. The rows of \mathbf{A} are obtained from Eq. (10) as

$$\mathbf{A} = \mathbf{W}^{-1}\mathbf{H}_p\mathbf{W}, \quad (12)$$

where the matrix \mathbf{W} is the N -point DFT matrix. The matrix \mathbf{H}_p is the diagonal matrix

$$\mathbf{H}_p = \begin{bmatrix} H_p(0) & 0 & \dots & 0 \\ 0 & H_p(1) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & H_p(N-1) \end{bmatrix}. \quad (13)$$

The matrix \mathbf{A} is unitary. Furthermore, \mathbf{A}^k represents the diffraction field at a distance pk . Proofs of these properties are given in the Appendix.

3.1. Calculation by matrix inversion

Our first approach calculates the diffraction field on the reference line by directly solving the system of linear equations (direct matrix inversion).

Let the vector \mathbf{g}' denote the diffraction data over s distributed data points,

$$\mathbf{g}' = \begin{bmatrix} g'_1 \\ g'_2 \\ \vdots \\ g'_s \end{bmatrix}. \quad (14)$$

Each g'_i is a function of both p_i , the index in the z direction, and n_i , the index in the x direction. The relation between the vector \mathbf{g}' and the reference vector \mathbf{f} is given by

$$\mathbf{g}' = \mathbf{A}_{\mathbf{BF}}\mathbf{f}, \quad (15)$$

where $\mathbf{A}_{\mathbf{BF}}$ is an s by N matrix:

$$\mathbf{A}_{\mathbf{BF}} = \begin{bmatrix} \mathbf{r}(p_1, n_1) \\ \mathbf{r}(p_2, n_2) \\ \vdots \\ \mathbf{r}(p_s, n_s) \end{bmatrix}, \quad (16)$$

where $\mathbf{r}(p_i, n_i)$ is a 1 by N row vector from the matrix \mathbf{A} with $p = p_i$. This row vector provides the diffraction field relation between the field on the reference line and the field on the point specified by p_i and n_i . The vector \mathbf{f} is obtained by

$$\mathbf{f} = \mathbf{A}_{\mathbf{BF}}^+\mathbf{g}', \quad (17)$$

where $\mathbf{A}_{\mathbf{BF}}^+$ is the pseudo-inverse of the matrix $\mathbf{A}_{\mathbf{BF}}$:

$$\mathbf{A}_{\mathbf{BF}}^+ = \begin{cases} (\mathbf{A}_{\mathbf{BF}}^H\mathbf{A}_{\mathbf{BF}})^{-1}\mathbf{A}_{\mathbf{BF}}^H, & s > N, \\ \mathbf{A}_{\mathbf{BF}}^H(\mathbf{A}_{\mathbf{BF}}^H\mathbf{A}_{\mathbf{BF}})^{-1}, & s < N. \end{cases} \quad (18)$$

Here $\mathbf{A}_{\mathbf{BF}}^H$ is the conjugate transpose of $\mathbf{A}_{\mathbf{BF}}$. Knowing the diffraction field on the reference line, we can compute the diffraction field on any other line in the 2D space using Eq. (11).

3.2. Calculation by using projections onto convex sets

Our second approach utilizes an iterative technique, based on the POCS method [1,9]. POCS has

been applied to various problems in holography and image restoration where a priori information is used to constrain the size of the feasible solution set [7,20,5]. It is a computational approach for finding an element of a feasible region defined by the intersection of a number of convex constraints, starting with an arbitrary infeasible point [1,9]. Fig. 4 shows how convergence to the intersection is achieved by iterative projections onto the individual convex sets. In our problem, the constraints are the known data points on consecutive lines. In addition, the data points are known to belong to the same diffraction field and therefore Eq. (10) has to be satisfied among all lines. Therefore, the convex set $C_l, l = 1, \dots, M$ (number of lines), can be defined as all possible diffraction fields having the given data points on a certain line $z = z_l$:

$$C_l = \{\forall f(x, z) : f(x_{i_l}, z_l) = \mathbf{v}_l, f(x, z_j) = \mathbf{A}^{j-l} f(x, z_l), \forall j, x\}, \quad (19)$$

where \mathbf{v}_l is the vector of known values on the line $z = z_l$ and \mathbf{i}_l is the vector with the indices of their positions. \mathbf{A} is the diffraction matrix given by Eq. (12) written for two consecutive lines, and its $(j-l)$ th power \mathbf{A}^{j-l} , is the diffraction matrix from the line $z = z_l$ to the line $z = z_j$. A closed form expression for $f(x, z_j)$ can be written with the help of an arbitrary function $q(x)$:

$$f(x, z_j) = \mathbf{A}_{\mathbf{i}_l}^{j-l} \mathbf{v}_l + \mathbf{A}_{\bar{\mathbf{i}}_l}^{j-l} q(x_{\bar{\mathbf{i}}_l}), \quad (20)$$

where $\bar{\mathbf{i}}_l$ is the vector with the indices of the unknown values on the line $z = z_l$, and $\mathbf{A}_{\mathbf{i}_l}^{j-l}$ and $\mathbf{A}_{\bar{\mathbf{i}}_l}^{j-l}$ are submatrices of \mathbf{A}^{j-l} obtained by taking columns with indices \mathbf{i}_l and $\bar{\mathbf{i}}_l$.

It is straightforward to show that the sets C_l defined by Eq. (19) are convex. Let us assume that the functions $f_1(x, z)$ and $f_2(x, z)$ belong to the set C_l , and $F(x, z) = \alpha f_1(x, z) + (1 - \alpha) f_2(x, z), 0 < \alpha < 1$ is their convex combination. Then for every line $z =$

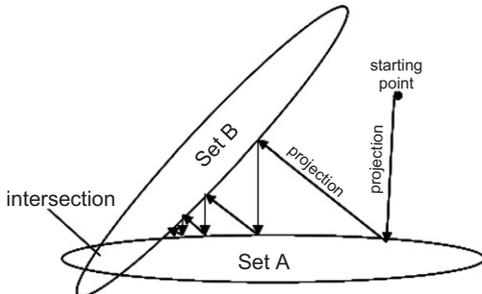


Fig. 4. Projections onto convex sets (POCS).

z_j we have

$$\begin{aligned} F(x, z_j) &= \alpha \mathbf{A}_{\mathbf{i}_l}^{j-l} \mathbf{v}_l + \alpha \mathbf{A}_{\bar{\mathbf{i}}_l}^{j-l} q_1(x_{\bar{\mathbf{i}}_l}) \\ &\quad + (1 - \alpha) \mathbf{A}_{\mathbf{i}_l}^{j-l} \mathbf{v}_l + (1 - \alpha) \mathbf{A}_{\bar{\mathbf{i}}_l}^{j-l} q_2(x_{\bar{\mathbf{i}}_l}) \\ &= \mathbf{A}_{\mathbf{i}_l}^{j-l} \mathbf{v}_l + \mathbf{A}_{\bar{\mathbf{i}}_l}^{j-l} Q(x_{\bar{\mathbf{i}}_l}), \end{aligned} \quad (21)$$

where $Q(x) = \alpha q_1(x) + (1 - \alpha) q_2(x)$. Since $q(x)$ in Eq. (20) can be an arbitrary function, the last line of Eq. (21) becomes of the form of Eq. (20) and therefore $F(x, z)$ belongs to the set C_l . This fact shows that C_l is convex, because a convex combination of any two functions which belong to C_l also belongs to C_l .

A POCS-based algorithm requires iterating from set to set using projections. A straightforward strategy is to propagate from line to line using Eq. (10). Unknown points are generated by the field from the previous line and the known points are kept. Let us assume $g(x, z) \in C_{l-1}$ and $f(x, z) \in C_l$. Then,

$$f(x_{\bar{\mathbf{i}}_l}, z_l) = g(x_{\bar{\mathbf{i}}_l}, z_l)$$

and

$$f(x_{\mathbf{i}_l}, z_l) = \mathbf{v}_l, \quad (22)$$

while projecting from the set C_{l-1} to the set C_l . With this choice, $f(x, z_l)$ differs from $g(x, z_l)$ in the restricted values at positions $x_{\mathbf{i}_l}$ only, and therefore the distance from $g(x, z_l)$ to $f(x, z_l)$ is minimized with respect to all functions in C_l . Moreover, the distance taken for any other line $z = z_j$ is minimized since

$$\begin{aligned} d &= \|f(x, z_j) - g(x, z_j)\|^2 \\ &= \|\mathbf{A}^{j-l}(f(x, z_l) - g(x, z_l))\|^2 \\ &= \|f(x, z_l) - g(x, z_l)\|^2 \end{aligned} \quad (23)$$

and \mathbf{A}^{j-l} is unitary for any j and l . Hence, the squared distance between $g(x, z)$ and $f(x, z)$ is also minimal for all the functions from the set C_l because

$$\|f(x, z) - g(x, z)\|^2 = \sum_j \|f(x, z_j) - g(x, z_j)\|^2.$$

Therefore, the choice for $f(x, z)$ given by Eq. (22) corresponds to the orthogonal projection of $g(x, z)$ onto the set C_l and hence the iterative method will converge [20].

At the first line $z = 0$, iterations are initiated by setting the missing points to arbitrary values. After passing through all lines, the unknown points on the first line are computed using the data on the last line.

The algorithm can be summarized as follows:

1. initialize the first line of the desired field
 $f(x_{i_1}, 0) = \mathbf{v}_0, f(x_{i_2}, 0) = q(x_{i_2}),$ for any $q(x)$
2. for $i = 1$ to n_{it}
 - (a) for $l = 2$ to M
 - i. $f(x, z_l) = \mathbf{A}f(x, z_{l-1})$
 - ii. $f(x_{i_l}, z_l) = \mathbf{v}_l$
 - (b) end
 - (c) $f(x, 0) = \mathbf{A}^{-M+1}f(x, z_M)$
 - (d) $f(x_{i_l}, 0) = \mathbf{v}_0;$
3. end

where n_{it} is the number of iteration.

3.3. Results

For purposes of illustrating and evaluating the two approaches outlined above, we choose a simple function as the optical field on the reference line. Then, the diffraction field at other lines is calculated using Eq. (10). The computation is conducted with $M = 256$ lines with $N = 256$ samples per line and the range of p is $[129, 384]$. The function on the reference line is a unit-magnitude square pulse of 32 samples located at the center of the line. The generated initial data pattern is depicted in Fig. 5(a).

Assessment of the results is based on the normalized error between the original and reconstructed diffraction patterns $u(n, p)$ and $u'(n, p)$, respectively,

$$\frac{\sum_{n=0}^{N-1} \sum_{p=129}^{384} |u'(n, p) - u(n, p)|^2}{\sum_{n=0}^{N-1} \sum_{p=129}^{384} |u(n, p)|^2}. \quad (24)$$

The error has been tabulated for different numbers s of known data points, as shown in Tables 1 and 2. For each value of s , 15 diffraction patterns have been reconstructed using s randomly selected data points from the initial simulated field. Each reconstructed diffraction pattern corresponds to a different random choice of the positions of the s known data points within the field. The final error reported for each value of s is the average of the error over all 15 choices.

The numerical implementations of the proposed algorithms utilize the complex arithmetic operations such as multiplication and addition. Computation times of these operations depend on many implementation details including hardware properties and operating system behavior. Implementations of the described digital algorithms need the execution of large amount of complex arithmetic operations; furthermore, large amount of data fetch and write

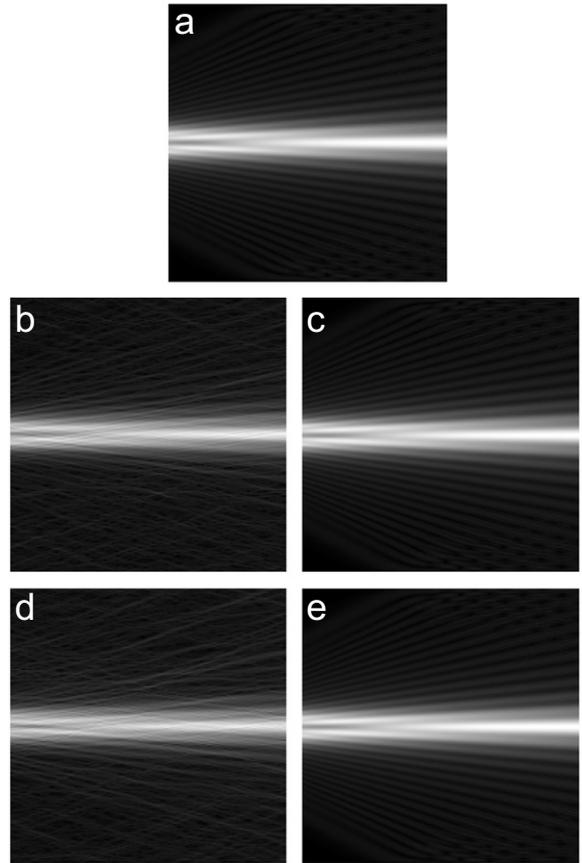


Fig. 5. Initial diffraction field over the entire 2D space; $N = 256$ samples per line (a) and reconstructed diffraction fields from s known data points (b)–(e); (b) matrix inversion method with $s = 230$; (c) matrix inversion method with $s = 282$; (d) POCS algorithm with $s = 230$; (e) POCS algorithm with $s = 282$.

operations are needed. The actual resultant computation time naturally depends on the specifics of the computer architecture and the operating system behavior. Though incomplete, a comparison of the required total number of multiplications may give an idea about the computational complexity of the algorithms.

The numerical results for the matrix inversion method are summarized in Table 1. Higher number of initial given samples yields better reconstruction of the original diffraction pattern, as expected. When this number is equal to or higher than N , the diffraction field is reconstructed perfectly. This approach involves computation of the diffraction field on a reference line by taking the pseudo-inverse of \mathbf{A}_{BF} and computation of the entire diffraction field from that line. Forward and inverse DFTs are required to calculate the diffraction from between

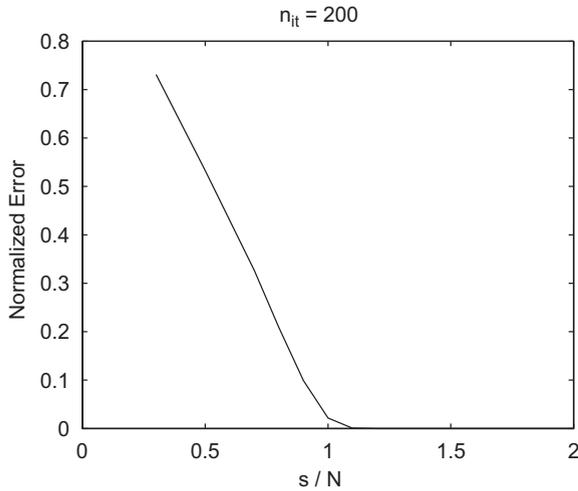


Fig. 6. Normalized errors for different numbers of known samples at 200 iterations.

underdetermined system where the intersection of the sets C_l is a region rather than a point. Therefore, successive projections yield a solution rather far from the original diffraction field. In general, the error decreases by the number of iterations and after a certain number of iterations is reached, the error saturates. The algorithm converges much faster for larger values of s .

The computational complexity of the POCS-based algorithm is determined by the number of iterations n_{it} , the number of parallel lines M , and the number of sample points per line N . For each iteration, M lines are computed by Fourier domain operations (cf. Eq. (10)). $N \log_2 N + N$ complex multiplications are required for calculation of the diffraction field on a line when FFT is used, because FFT algorithm is used twice. This results in a total number of complex multiplications:

$$n_{it}(MN \log_2 N + MN). \quad (27)$$

Note that the number of given points s is not present in the complexity measure given by (27). However, this number influences the complexity indirectly, as it determines the number of iterations needed for achieving the desired accuracy. As shown in Fig. 7, more given points result in less iterations to achieve the same accuracy. The curve drops right after $s = N$ and after that point the desired error can be reached in a reasonable number of iterations. For example, less than 10 iterations are sufficient if starting with $s = 2N$ given points. In contrast, the matrix inversion algorithm gets more complex for higher s . To relate the computational

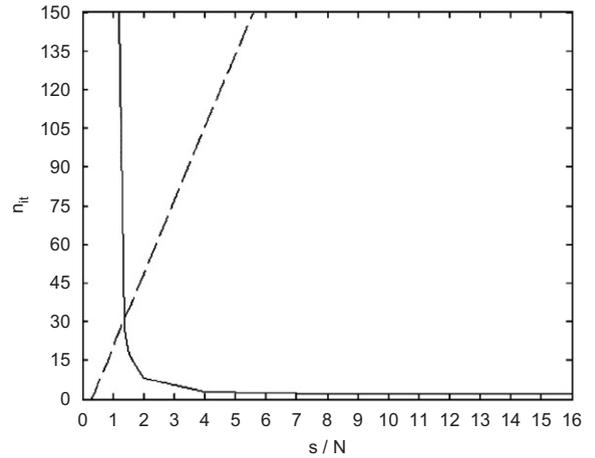


Fig. 7. Evaluation of POCS computational efficiency for different numbers of known samples s . Solid line: number of POCS iterations n_{it} needed to achieve normalized error < 0.0005 . Dashed lines: number of iterations for which POCS and matrix inversion methods give the same computational costs.

complexity of the POCS-based algorithm with that of direct matrix inversion, a second curve (the dashed line) is constructed in Fig. 7. For each s , it gives the number of iterations so that the computational complexity of the two algorithms is the same. Any number of iterations below that curve positions the POCS approach as the more preferable method. In Fig. 7, these iterations become less for $s \geq 1.4N$. Direct comparison of the number of the multiplications that both approaches require for different values of s are presented in Table 3. The values for the matrix inversion method are calculated using Eq. (26). The values for POCS method are found from Eq. (27). The parameter n_{it} , in Eq. (27), is determined by the POCS method that yields normalized error below 0.0005. Again, Table 3 shows that the POCS algorithm is less computationally costly when $s \geq 1.4N$. An additional peculiarity of the POCS approach is that, being an iterative algorithm, it is more robust to computational errors than direct matrix inversion, especially for high values of N .

Fig. 5 shows simulations for values of s close to N , with no visual difference between the results of the two methods.

4. Conclusion

In this work, the computation of the diffraction field from a set of distributed data points which may represent the abstract structure of an object has

Table 3

Comparison between the number of complex multiplications of POCS method and pseudo-inversion method (PINV) for different numbers of known samples s

s	256	282	358	384	410	512	1024	2048	4096	8192	16384
PINV	1.18e+7	1.35e+7	1.85e+7	2.02e+7	2.19e+7	2.86e+7	6.21e+7	1.29e+8	2.63e+8	5.32e+8	1.07e+9
POCS	8.55e+9	1.29e+8	1.61e+7	1.13e+7	9.17e+6	4.72e+6	1.77e+6	1.18e+6	1.18e+6	1.18e+6	1.18e+6

The costs are shown in number of complex multiplications needed to achieve error below 0.0005.

been investigated. First, a discrete formulation of the relevant diffraction theory has been presented. In this formulation, fictionally periodic fields have been used where the period of the field is determined by the size of the field along the transverse direction. Two approaches for obtaining simultaneous solution of the diffraction field have been studied. The first attacks the problem by a direct matrix inversion approach while the second utilizes the POCS method to recover the desired diffraction field iteratively. Both algorithms converge to the desired field when the number of given samples is equal to or larger than the period but they may not converge to the desired field if the number of given data points is lower than the period. Computational complexity which is determined by the number of complex multiplications is another issue addressed in this work. The matrix inversion enjoys lower computational complexity for small numbers of initially given points while POCS is beneficial when there are more given points, when a few number of iterations becomes sufficient to achieve the desired accuracy.

Appendix

The proof that the matrix \mathbf{A} is unitary is given below by showing that $\mathbf{A}^H \mathbf{A} = \mathbf{I}$ as follows:

$$\begin{aligned}
 \mathbf{A}^H \mathbf{A} &= \mathbf{A} \mathbf{A}^H = \mathbf{I} \\
 &= (\mathbf{W}^{-1} \mathbf{H}_p \mathbf{W})(\mathbf{W}^{-1} \mathbf{H}_p \mathbf{W})^H \\
 &= (\mathbf{W}^{-1} \mathbf{H}_p \mathbf{W})(\mathbf{W}^H \mathbf{H}_p^H \mathbf{W}^{-H}) \\
 &= \mathbf{W}^{-1} \mathbf{H}_p \mathbf{N} \mathbf{H}_p^H \mathbf{W}^{-H} \\
 &= \mathbf{W}^{-1} \mathbf{N} \mathbf{W}^{-H} \\
 &= \mathbf{I}.
 \end{aligned} \tag{28}$$

Another property of \mathbf{A} is that

$$\begin{aligned}
 \mathbf{A}^l \mathbf{A}^j &= (\mathbf{W}^{-1} \mathbf{H}_{lp} \mathbf{W})(\mathbf{W}^{-1} \mathbf{H}_{jp} \mathbf{W}) \\
 &= \mathbf{W}^{-1} \mathbf{H}_{lp} \mathbf{H}_{jp} \mathbf{W} \\
 &= \mathbf{W}^{-1} \mathbf{H}_{(l+j)p} \mathbf{W},
 \end{aligned} \tag{29}$$

where $\mathbf{H}_{(l+j)p}$ represents the kernel of the discrete system which is used to calculate the diffraction field at $p(l+j)$.

References

- [1] R. Aharoni, Y. Censor, Block iterative projection methods for parallel computation of solutions to convex feasibility problems, *Linear Algebra Appl.* 120 (1989) 165–175.
- [2] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Amsterdam, Holland, 1990.
- [3] M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Cambridge University Press, New York, 1980.
- [4] N. Delen, B. Hooker, Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast Fourier transform approach, *J. Opt. Soc. Am. A* 15 (1998) 857–867.
- [5] R.G. Dorsch, A.W. Lohmann, S. Sinzinger, Fresnel ping-pong algorithm for two-plane computer-generated hologram display, *Appl. Opt.* 33 (1994) 869–875.
- [6] G.B. Esmer, *Computation of holographic patterns between tilted planes*, Master's Thesis, Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, 2004.
- [7] R.W. Gerchberg, W.O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, *Optik* 35 (1972) 237–246.
- [8] J.W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [9] L.G. Gubin, B.T. Polyak, E.V. Raik, The method of projections for finding the common point of convex sets, *USSR Comput. Math. Math. Phys.* 7 (1967) 1–24.
- [10] M.L. Huebschman, B. Munjuluri, H.R. Garner, Dynamic holographic 3-d image projection, *Opt. Exp.* 11 (2003) 437–445.
- [11] E. Lalor, Conditions for the validity of the angular spectrum of plane waves, *J. Opt. Soc. Am.* 58 (1968) 1235–1237.
- [12] M. Lucente, *Diffraction-specific fringe computation for electro-holography*, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1994 (online), (<http://www.lucente.us/pubs/PhDthesis/contents.html>).
- [13] D. Mas, J. Pérez, C. Hernández, C. Vázquez, J.J. Miret, C. Illueca, Fast numerical calculation of fresnel patterns in convergent systems, *Opt. Commun.* 227 (2003) 245–258.
- [14] K. Matsushima, H. Schimmel, F. Wyrowski, Fast calculation method for optical diffraction on tilted planes by use of

- the angular spectrum of plane waves, *J. Opt. Soc. A* 20 (2003) 1755–1762.
- [15] L. Onural, P.D. Scott, Digital decoding of in-line holograms, *Opt. Eng.* 26 (1987) 1124–1132.
- [16] H.M. Ozaktas, A. Koc, I. Sari, Efficient computation of quadratic-phase integrals in optics, *Opt. Lett.* 31 (2006) 35–37.
- [17] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, Wiley, New York, 1991.
- [18] G.C. Sherman, Application of the convolution theorem to Rayleigh's integral formulas, *J. Opt. Soc. Am.* 57 (1967) 546–547.
- [19] T. Tommasi, B. Bianco, Computer-generated holograms of tilted planes by a spatial frequency approach, *J. Opt. Soc. Am. A* 10 (1993) 299–305.
- [20] D.C. Youla, H. Webb, Image restoration by the method of convex projections: part I-theory, *IEEE Trans. Med. Imag. TMI-1* (1982) 81–94.

Publication P4

G.B. Esmer, **V. Uzunov**, L. Onural, A. Gotchev, H.M. Ozaktas, "Reconstruction of scalar diffraction field from distributed data points over 3D space," in *Proceedings of the International Conference: True-Vision, Capture, Transmission, and Display of 3D Video, 3DTV-CON 2007*, Kos Island, Greece, 7-9 April 2007, pp. 1-4.

©2007 IEEE. Reprinted with permission.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Tampere University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

RECONSTRUCTION OF SCALAR DIFFRACTION FIELD FROM DISTRIBUTED DATA POINTS OVER 3D SPACE

G. Bora Esmer¹, Vladislav Uzunov², Levent Onural¹, Atanas Gotchev² and Haldun M. Ozaktas¹

¹ Bilkent University, Department of Electrical and Electronics Engineering,
TR-06800 Ankara, Turkey

² Tampere University of Technology, Institute of Signal Processing, FIN-33101 Tampere, Finland

ABSTRACT

Diffraction field computation is an important task in the signal conversion stage of the holographic 3DTV. We consider an abstract setting, where the diffraction field of the desired 3D scene to be displayed is given by discrete samples distributed over 3D space. Based on these samples, a model of the diffraction field should be built to allow the field computation at any desired point. In our previous works, we have proved our concepts for the simplistic 2D case. In this paper, we generalize the earlier proposed techniques, namely the projection onto convex sets and conjugate gradient based techniques and test them for their computational efficiency and memory requirements for a specific 3D case.

Index Terms— Scalar Optical Diffraction, Rayleigh - Sommerfeld Diffraction, Plane Wave Decomposition, Projection onto Convex Sets, Conjugate Gradient

1. INTRODUCTION

Holographic 3DTV requires the following basic building blocks: capture, representation, coding, signal conversion and display. In this chain, the display device plays a very important role, as it generates the light field being the optical replica of the captured and abstractly represented 3D scene. This work is related with the signal conversion part which provides the connection between the 3D scene representation and the display end.

In diffraction theory, computation of a diffraction field at a surface due to a given field elsewhere is one of the major problem which has attracted researchers for decades [1]. Computation of the scalar optical diffraction due to an abstract 3D scene is a challenging problem.

In this work, we assume that a 3D scene is described by a set of distributed data points over the 3D space. In our earlier works [2], [3], this problem has been addressed within a 2D space and now it is extended to 3D space. Problems computational complexity and memory management arising with this extended space setting, are commented in more details.

This work is supported by EC within FP6 under Grant 511568 with acronym 3DTV.

2. BASICS OF THE SCALAR OPTICAL DIFFRACTION THEORY

To compute scalar optical fields, we do not use Fresnel or Fraunhofer approximations. Instead, we rely on the Rayleigh-Sommerfeld (R-S) integral as the more general and exact scalar optical diffraction integral. More specifically, we utilize the plane wave decomposition (PWD) since it provides the same result as the R-S diffraction integral [4]. While the latter utilizes spatial domain relations, the former interprets the problem in frequency domain and is more attractive from computational point of view. The notations in this work are generalized from [2] and [3].

Lets assume that an initial diffraction field, $u_a(x, y, 0)$, is given on the plane $z = 0$. 2D Fourier transform (FT) of $u_a(x, y, 0)$ gives the complex coefficients of the plane waves, $A(k_x, k_y)$, that form $u_a(x, y, 0)$,

$$(2\pi)^2 A(k_x, k_y) = \mathcal{F}\{u_a(x, y, 0)\} \quad (1)$$

where \mathcal{F} denotes the 2D FT [5]. Then, the diffraction field for monochromatic waves on another plane which is parallel to $z = 0$ is expressed as

$$u_a(x, y, z) = \int_{-\frac{2\pi}{\lambda}}^{\frac{2\pi}{\lambda}} A(k_x, k_y) \exp[j(k_x x + k_y y + k_z z)] dk_x dk_y \quad (2)$$

where k_x , k_y and k_z are the spatial frequencies of the propagating waves along the directions x , y and z axes, respectively. The x and y axes denote the transversal directions and z is the longitudinal axis which is the optical axis. The variable k_z can be expressed as a function of k_x and k_y , because of dealing with monochromatic waves, $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$, where $k = \frac{2\pi}{\lambda}$.

The spatial frequencies of the diffraction field which is propagating along the optical axis may be restricted to be within $-B \leq k_x, k_y \leq B$, where $B \leq k$. For numerical computations the frequencies k_x and k_y are discretized to N frequency terms each. Consequently, the input signal can be represented by N^2 frequency components. These frequency components are selected as $k_x = n_f \frac{2B}{N}$ and $k_y = m_f \frac{2B}{N}$,

where n_f and m_f are integers and elements of the set $[-N/2, N/2)$. Uniform sampling operation in frequency domain causes periodicity in the transversal spatial domain. The period in both x and y axes is $X = \frac{\pi N}{B}$. Therefore, the expression given by Eq. 2 becomes

$$u_a(x, y, z) = \sum_{n_f=-N/2}^{N/2-1} \sum_{m_f=-N/2}^{N/2-1} A_D(n_f, m_f) \exp(j\sqrt{k^2 - (\frac{2B}{N}n_f)^2 - (\frac{2B}{N}m_f)^2}z) \exp[j\frac{2B}{N}(n_f x + m_f y)] \quad (3)$$

where $A_D(n_f, m_f)$ is a 2D array representing samples of $A(k_x, k_y)$ [6]. Sampling along the x and y axes is accomplished by setting the sampling period $X_s = \frac{\pi}{B}$ to satisfy the Nyquist rate. Therefore, the expression in Eq. 3 can be rewritten as

$$u_a(nX_s, mX_s, pX_s) = \sum_{n_f=-N/2}^{N/2-1} \sum_{m_f=-N/2}^{N/2-1} A(n_f, m_f) H_p(n_f, m_f) \exp[j\frac{2\pi}{N}(n_f n + m_f m)] \quad (4)$$

where $H_p(n_f, m_f)$ is the frequency response of the free space propagation kernel, which is defined as

$$H_p(n_f, m_f) = \exp(j\frac{2\pi}{N}\sqrt{\beta^2 - n_f^2 - m_f^2}p), \quad (5)$$

where $\beta = \frac{NX_s}{\lambda}$ and $p = \frac{z}{X_s}$. Thus the discrete diffraction field becomes,

$$u(n, m, p) = DFT^{-1}\{DFT\{u(n, m, 0)\}H_p(n_f, m_f)\}, \quad (6)$$

where DFT and DFT^{-1} stand for discrete FT and inverse discrete FT, respectively [6].

3. DISCRETE FIELD COMPUTATION FROM DISTRIBUTED DATA

Two methods are presented: projection onto convex sets (POCS) and conjugate gradient (CG). Both of them take the given distributed sample points as an input and give the diffraction field on a reference plane. We choose to implement these fast iterative methods, because direct solution needs much more computation time.

3.1. Projection Onto Convex Sets

First method is based on an iterative approach POCS. Our problem falls in the framework of POCS as proven in [2]. The constraints for the problem are the given samples and the R-S diffraction field relationship. The algorithm utilizes Eq. 6.

The summary of the algorithm is

1. initialize the first line of the desired field $f(n_{i_1}, m_{i_1}, 1) = \mathbf{v}_1$, $f(n_{\bar{i}_1}, m_{\bar{i}_1}, 1) = q(n_{\bar{i}_1}, m_{\bar{i}_1})$, for any $q(n, m)$
2. for $i = 1$ to n_{it}
 - (a) for $l = 2$ to M
 - i. $f(n, m, p_l) = DFT^{-1}\{DFT\{f(n, m, p_{l-1})\}H_1(n_f, m_f)\}$
 - ii. $f(n_{i_l}, m_{i_l}, p_l) = \mathbf{v}_l$
 - (b) end
 - (c) $f(n, m, 1) = DFT^{-1}\{DFT\{f(n, m, p_M)\}H_{-M+1}(n_f, m_f)\}$
 - (d) $f(n_{i_1}, m_{i_1}, 1) = \mathbf{v}_1$;
3. end

where \mathbf{v}_l is the vector of the known samples on the plane $z = p_l X_s$, and i_l and \bar{i}_l are the vector of indices of the known and the unknown values on the plane $z = p_l X_s$, respectively. $q(n, m)$ is an arbitrary function whose samples are used in place of the unknown samples and n_{it} is the number of total iteration [2].

3.2. Conjugate Gradient

The second algorithm is based on Eq. 3 which provides the relation between the complex amplitudes of the plane waves, $A_D(n_f, m_f)$, and the given data samples. The relationship given by Eq. 3 can be expressed as a matrix multiplication,

$$\mathbf{u} = \mathbf{R}\mathbf{a} \quad (7)$$

where the vector \mathbf{a} denotes the complex amplitudes of the plane waves that form the diffraction field on the reference plane. In Eq. 3, these complex amplitudes, $\mathbf{A}_D(n_f, m_f)$, are given as a 2D array, $\mathbf{A}_D = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_N]$. The representation of \mathbf{A}_D is converted into a vector, \mathbf{a} , in Eq. 7 as

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_N \end{bmatrix}. \quad (8)$$

The vector \mathbf{u} in Eq. 7 denotes the diffraction field on the given data points. The \mathbf{R} matrix in Eq 7 is the reconstruction matrix and its elements are

$$\mathbf{r}_{l,i,j} = \exp(j\sqrt{k^2 - (\frac{2B}{N}i)^2 - (\frac{2B}{N}j)^2}z_l) \exp(j\frac{2B}{N}ix_l) \exp(j\frac{2B}{N}jy_l) \quad (9)$$

where x_l and y_l are the locations of the given samples on the plane $z = z_l$. The \mathbf{R} matrix is formed as in Eq. 10.

Multiplication of \mathbf{u} by the pseudo-inverse of the \mathbf{R} will give \mathbf{a} . The pseudo-inversion of the matrix \mathbf{R} is taken by CG algorithm. The method is summarized as

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_{1,1,1} & \dots & \mathbf{r}_{1,N,1} & \mathbf{r}_{1,1,2} & \dots & \mathbf{r}_{1,N,2} & \dots & \mathbf{r}_{1,1,N} & \dots & \mathbf{r}_{1,N,N} \\ \mathbf{r}_{2,1,1} & \dots & \mathbf{r}_{2,N,1} & \mathbf{r}_{2,1,2} & \dots & \mathbf{r}_{2,N,2} & \dots & \mathbf{r}_{2,1,N} & \dots & \mathbf{r}_{2,N,N} \\ \vdots & & & \vdots & & & \ddots & \vdots & & \\ \mathbf{r}_{s,1,1} & \dots & \mathbf{r}_{s,N,1} & \mathbf{r}_{s,1,2} & \dots & \mathbf{r}_{s,N,2} & \dots & \mathbf{r}_{s,1,N} & \dots & \mathbf{r}_{s,N,N} \end{bmatrix}. \quad (10)$$

1. compute the \mathbf{R} by using Eq. 9 and Eq. 10
2. if $s < N^2$ compute $\mathbf{Q} = \mathbf{R}\mathbf{R}^H$ and $\mathbf{b} = \mathbf{u}$, otherwise compute $\mathbf{Q} = \mathbf{R}^H\mathbf{R}$ and $\mathbf{b} = \mathbf{R}^H\mathbf{u}$
3. initialize $\hat{\mathbf{x}}^{[0]}$ arbitrarily, $\mathbf{d}_0 = \mathbf{b} - \mathbf{Q}\hat{\mathbf{x}}^{[0]}$ and $\mathbf{g}_0 = -\mathbf{d}_0$
4. for $n = 1$ to $n_{it} \leq N^2$
 - (a) $\alpha_n = -\frac{\mathbf{g}_n^T \mathbf{d}_n}{\mathbf{d}_n^T \mathbf{Q} \mathbf{d}_n}$
 - (b) $\hat{\mathbf{x}}^{[n+1]} = \hat{\mathbf{x}}^{[n]} + \alpha_n \mathbf{d}_n$
 - (c) $\mathbf{g}_{n+1} = \mathbf{Q}\hat{\mathbf{x}}^{[n+1]} - \mathbf{b}$
 - (d) $\gamma_n = -\frac{\mathbf{g}_{n+1}^T \mathbf{d}_n}{\mathbf{d}_n^T \mathbf{Q} \mathbf{d}_n}$
 - (e) $\mathbf{d}_{n+1} = -\mathbf{g}_{n+1} + \gamma_n \mathbf{d}_n$
5. if $s < N^2$ compute $\hat{\mathbf{a}} = \mathbf{R}^H \hat{\mathbf{x}}$, otherwise $\hat{\mathbf{a}} = \hat{\mathbf{x}}$
6. reconstruct the diffraction field on the reference plane from the estimated complex amplitudes of the plane waves by utilizing inverse FT relation.

4. RESULTS

The outlined algorithms are evaluated by a synthetically generated simple optical field on the reference plane. Then, the diffraction field over the 3D space, due to the diffraction field on the reference plane, is computed according to Eq. 6. The field on the reference plane consists of N by N samples, where $N = 64$. There is an 8 by 8 unit-magnitude rectangular opening located in the middle of the reference plane and the rest of the samples are taken as zero. The 3D space consists of $M = 64$ planes which are uniformly located along the z -axis and there is a 64 by 64 uniform grid on each plane. An illustration of the implemented scenario is given in Figure 1. In the scenario, the distance parameter p between the reference and first plane equals to four and there are M planes in the defined 3D space. Typical results of the algorithms can be seen in Figure 2.

Evaluation of the results is based on two parameters. One of them is the normalized error between the original and the reconstructed diffraction patterns, $u(n, m, 0)$ and $u'(n, m, 0)$:

$$\left(\sum_{n,m=0}^{N-1} |u'(n, m, 0) - u(n, m, 0)|^2 \right) / \left(\sum_{n,m=0}^{N-1} |u(n, m, 0)|^2 \right). \quad (11)$$

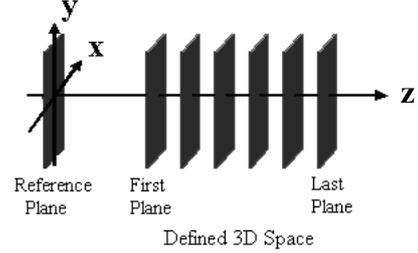


Fig. 1. Implemented scenario.

The other one is the number of complex multiplications required by the algorithms. From the computed field over the 3D space, we randomly take s data points to reconstruct the field on the reference plane. For each value of s , ten different random selections of data points are generated, assessment parameters are computed for each selection and then averaged for each value of s .

Increasing the number of given samples, s , in both algorithms provides faster convergence to the given field on the reference plane, as expected. The curves in Figure 3 show how fast the normalized errors decrease when the number of complex multiplications is fixed to 1.7×10^9 .

When $s \leq N^2$, the solution sets of the algorithms will not be comprised by only the original field. Hence, the solution may converge to pattern which may not be the same as the original one.

The computational complexities of both algorithms are determined by the number of complex multiplications, because we assume that complex multiplication needs more computation time than complex addition, data fetching and writing operations. Both algorithms use the 2D-DFT operation which can be implemented by $N^2 \log_2 N$ complex multiplications if common 2D-FFT algorithms are used. For the POCS algorithm, total number of complex multiplications is

$$n_{it}(2MN^2 \log_2 N + N^2 M). \quad (12)$$

The parameter n_{it} is related to the s , but there is no closed form for it. It can be found from the experiments. For the CG algorithm, the total number of complex multiplications is

$$2n_{it}N^4 + sN^2 + N^2 \log_2 N \quad (13)$$

where the parameter n_{it} is again heuristically estimated. It is found by the numerical experiments for each scenario. The curves in Figure 4 illustrate the necessary complex multiplications in POCS and CG algorithms.

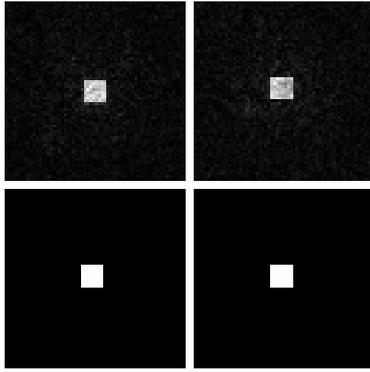


Fig. 2. Layout of the figure $|\frac{a}{c}| |\frac{b}{d}|$ (a) Magnitude of the reconstructed diffraction field on the reference plane obtained by the POCS algorithm when the number of given samples is $0.8N^2$. (b) The same scenario as in (a) when CG is used (c) Reconstructed field by POCS on the same plane when the number of given samples is $2.0N^2$. (d) The same scenario as in (c) when CG is utilized.

5. CONCLUSION

Two effective methods to calculate the scalar optical diffraction field simultaneously due to the arbitrarily distributed sample points over the 3D space are presented. First method utilizes POCS algorithm and the second one is based on CG algorithm. Both of them are iterative methods. POCS needs less memory space than CG. In CG, we have to use large matrices to represent the diffraction field relationship. Implementation of POCS needs less number of complex multiplications than CG in the case, when the given 3D diffraction field samples are taken from an uniform grid with large enough sampling period over the z-axis.

6. REFERENCES

- [1] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [2] G.B. Esmer, V. Uzunov, L. Onural, H.M. Ozaktas, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," *Signal Processing: Image Communication* (2006).
- [3] V. Uzunov, A. Gotchev, G. B. Esmer, L. Onural, and H. M. Ozaktas, "Non-uniform sampling and reconstruction of diffraction field," in *In Workshop on Spectral Methods and Multirate Signal Processing, SMMSP 2006, volume TICSP Series 34*, 191-198, September 2006.

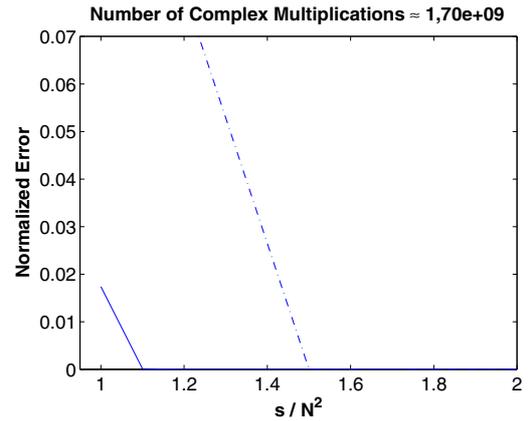


Fig. 3. Convergence of the POCS and the CG algorithms for different number of known samples at approximately 1.7×10^9 complex multiplications. These curves are obtained by averaging the results of 10 simulations. Solid line stands for the POCS, dashed line is used for the CG.

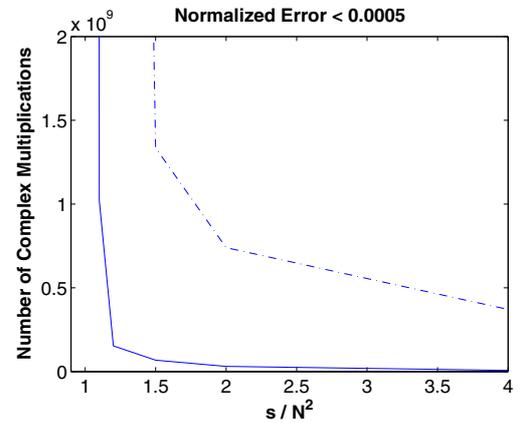


Fig. 4. Number of complex multiplications for POCS and CG methods when the normalized error is limited to 0.0005. The given results are based on the average of 10 simulations. Solid line represents the POCS, dashed line indicates the CG.

- [4] G. C. Sherman, "Application of the convolution theorem to Rayleigh's integral formulas," *J. Opt. Soc. Am.*, vol. 57, pp. 546-547, 1967.
- [5] Levent Onural, "Impulse functions over curves and surfaces and their applications to diffraction," *J. Math. Anal. Appl.*, vol. 322, pp. 18-27, 2006.
- [6] G. Bora Esmer and Levent Onural, "Computation of holographic patterns between tilted planes," in *Holography 2005*, Varna, Bulgaria, 21-25, May 2005, vol. Proc. SPIE 6252.

Publication P5

V. Uzunov, G.B. Esmer, A. Gotchev, L. Onural, H.M. Ozaktas, “Bessel functions-based reconstruction of non-uniformly sampled diffraction fields,” in *Proceedings of the International Conference: True-Vision, Capture, Transmission, and Display of 3D Video, 3DTV-CON 2007*, Kos Island, Greece, 7-9 April 2007, pp. 1-4.

©2007 IEEE. Reprinted with permission.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Tampere University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

BESSEL FUNCTIONS - BASED RECONSTRUCTION OF NON-UNIFORMLY SAMPLED DIFFRACTION FIELDS

Vladislav Uzunov¹, G. Bora Esmer², Atanas Gotchev¹, Levent Onural², and Haldun M. Ozaktas²

¹ Tampere University of Technology, Institute of Signal Processing, FIN-33101 Tampere, Finland

² Bilkent University, Department of Electrical and Electronics Engineering,
TR-06800 Bilkent, Ankara, Turkey

ABSTRACT

A discrete computational model for the diffraction process is essential in forward problems related to holographic TV. The model must be as general as possible, since the shape of the displayed objects does not bear any restrictions. We derive a discrete diffraction model which suits the problem of reconstruction of diffraction fields from a set of non-uniformly distributed samples. The only restriction of the model is the wave nature of the field. The derivation takes advantage of changing the spatial and frequency coordinates to polar form and ends up with a model stated in terms of Bessel functions. The model proves to be a separable orthogonal basis. It shows rapid convergence when evaluated in the framework of the non-uniform sampling problem.

Index Terms— diffraction, non-uniform sampling, polar coordinates, Bessel functions

1. INTRODUCTION

The computation of the light field distribution which arises over the entire three-dimensional (3D) space from an abstract representation of a 3D scene is known as the *forward problem* in holography [1]. This problem is important in holographic three-dimensional television (3DTV), because the display device has to be controlled by signals which depend on the optical field of the scene. In order to solve forward problems, it is relevant to have an accurate and general model that simulates numerically the diffraction process. Since numerical calculations usually are carried out by a computer, the model has to be also discrete, i.e. it must be dependent on finite number of parameters. In addition, the model must be as general as possible, because the nature of the displayed scene is usually not known in advance.

In our earlier work [2, 3] the forward problem is stated as reconstruction of the diffraction field from a set of given samples, non-uniformly distributed in the space. This statement is general enough, because it does not assume any particu-

lar shape for the object. However, the model used to calculate the diffraction bears certain limitations. The plane wave decomposition integral [4] is discretized in order to achieve discrete diffraction model. The discretization is done by sampling the Fourier transform of first plane of the field. Since this sampling is equivalent to periodization of the function on the first plane, this function must have finite spatial extend. It is also assumed to be essentially band-limited in order to obtain a model which depends on a finite number of Fourier coefficients. Moreover, the diffraction field can be considered up to a limited distance along the propagation direction z . The discretization of the spectra on the first line assumes periodization along the transverse directions. On the other hand, a space-limited pattern on the first plane spreads when propagated along the z -axis. The spread is proportional to the distance from the initial plane. After certain distance the adjacent periodic replicas on the same plane start overlapping [5] and the model starts producing erroneous results. One way to overcome this problem is to assume larger periods in the transverse directions when discretizing the Fourier spectra of the initial plane, so that there is more distance between the adjacent replicas. However, this will increase the dimensionality of the model, because larger period in spatial domain corresponds to denser sampling in frequency domain and the essential frequency band is represented by more coefficients. Therefore the model in [2, 3] can be used to calculate accurately the diffraction field up to some distance, which depends on the support of the diffraction pattern and the chosen dimensionality of the model.

Our motivation in this paper is to search for a model which can be used to calculate diffraction patterns with less restrictions. Yet the model must be discrete and finite-dimensional. In order to do so, we start from the wave nature of the diffraction field. The Helmholtz wave equation puts the only restriction on our model - the Fourier transform of the wave is nonzero only on a sphere. We arrive at an elegant model involving Bessel functions of the first kind [6] by writing the spatial and frequency coordinates in a polar form. There exist accurate numerical methods to calculate the Bessel functions [7] which make our model practically feasible.

This work is supported by EC within FP6 under Grant 511568 with acronym 3DTV.

2. DIFFRACTION MODEL IN POLAR COORDINATES

We consider optical fields generated by monochromatic (single wavelength) light waves, propagating in free space [4]. For simplicity, the discussions are restricted to one transverse dimension only. Extension to two transverse dimensions is straightforward. The light field $u(x, z)$ can be expressed in terms of its two-dimensional Fourier transform $A(k_x, k_z)$ as

$$u(x, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z. \quad (1)$$

The x axis is the transverse axis and the z axis is the optical axis along which the field propagates. The variables k_x and k_z are the spatial frequencies for x and z respectively. They can be written in polar form as $k_x = k \sin \theta$ and $k_z = k \cos \theta$, where $k \in [0, \infty)$ is the radius and $\theta \in [0, 2\pi)$ is the angle. After the change of the variables the 2D Fourier transform integral of Eq.1 becomes

$$u(x, z) = \int_0^{\infty} \int_0^{2\pi} k A(k \sin \theta, k \cos \theta) e^{j(kx \sin \theta + kz \cos \theta)} dk d\theta. \quad (2)$$

The light field $u(x, z)$ satisfies the Helmholtz wave equation $\nabla^2 u(x, z) + k^2 u(x, z) = 0$. Therefore the Fourier transform $A(k_x, k_z)$ is nonzero only on a circle, centered at the origin with radius $k_0 = \frac{2\pi}{\lambda}$, where λ is the wavelength of the monochromatic light. This implies that the spectra $A(k_x, k_z)$ written in polar coordinates depends only on one variable:

$$A(k \sin \theta, k \cos \theta) = A(k_0 \sin \theta, k_0 \cos \theta) \equiv A(\theta). \quad (3)$$

Now the Fourier integral of Eq.2 becomes one dimensional:

$$u(x, z) = \frac{2\pi}{\lambda} \int_0^{2\pi} A(\theta) e^{j \frac{2\pi}{\lambda} (x \sin \theta + z \cos \theta)} d\theta. \quad (4)$$

The change of the spatial coordinates x and z in polar form as $x = r \sin \phi$ and $z = r \cos \phi$ simplifies the integral in Eq. 4 to

$$\begin{aligned} u(r, \phi) &= \frac{2\pi}{\lambda} \int_0^{2\pi} A(\theta) e^{j \frac{2\pi}{\lambda} (r \sin \phi \sin \theta + r \cos \phi \cos \theta)} d\theta \\ &= \frac{2\pi}{\lambda} \int_0^{2\pi} A(\theta) e^{j \frac{2\pi}{\lambda} r \cos(\theta - \phi)} d\theta \end{aligned} \quad (5)$$

We arrived at a rather simple form that does not impose any restrictions on the generated field. However, this expression is far from a computational formula since it involves integration and does not depend on finite number of elements. The Fourier transform of the field is nonzero only on a circle and

therefore $A(\theta)$ can be considered as 2π -periodic with respect to θ . Hence $A(\theta)$ can be represented by the complex Fourier series:

$$A(\theta) = \sum_{m=-\infty}^{\infty} c_m e^{jm\theta} \quad (6)$$

Inserting Eq.6 in Eq.5 leads to

$$\begin{aligned} u(r, \phi) &= \frac{2\pi}{\lambda} \int_0^{2\pi} \sum_{m=-\infty}^{\infty} c_m e^{jm\theta} e^{j \frac{2\pi}{\lambda} r \cos(\theta - \phi)} d\theta \\ &= \frac{2\pi}{\lambda} \sum_{m=-\infty}^{\infty} c_m \int_0^{2\pi} e^{jm\theta} e^{j \frac{2\pi}{\lambda} r \cos(\theta - \phi)} d\theta. \end{aligned} \quad (7)$$

The integral inside the summation of the last line can be written as a Bessel function of the first kind [6] after making the substitution $\alpha = \theta - \phi$:

$$\begin{aligned} u(r, \phi) &= \frac{2\pi}{\lambda} \sum_{m=-\infty}^{\infty} c_m \int_{-\phi}^{2\pi - \phi} e^{jm(\alpha + \phi)} e^{j \frac{2\pi}{\lambda} r \cos \alpha} d\alpha \\ &= \frac{2\pi}{\lambda} \sum_{m=-\infty}^{\infty} c_m e^{jm\phi} \int_0^{2\pi} e^{jm\alpha} e^{j \frac{2\pi}{\lambda} r \cos \alpha} d\alpha \\ &= \frac{2\pi}{\lambda} \sum_{m=-\infty}^{\infty} c_m e^{jm\phi} 2\pi j^m J_m \left(\frac{2\pi}{\lambda} r \right). \end{aligned} \quad (8)$$

Here by $J_m(t)$ is denoted the m -th Bessel function of the first kind [6], whose integral form is

$$J_m(t) = \frac{1}{2\pi j^m} \int_0^{2\pi} e^{jm\alpha} e^{jt \cos \alpha} d\alpha. \quad (9)$$

The last line of Eq. 8 can be simplified to obtain the final form of the discrete diffraction model:

$$u(r, \phi) = \frac{4\pi^2}{\lambda} \sum_{m=-\infty}^{\infty} c_m e^{jm(\phi + \frac{\pi}{2})} J_m \left(\frac{2\pi}{\lambda} r \right) \quad (10)$$

There exist efficient and accurate numerical algorithms to calculate Bessel functions [7], which can be utilized in the model. Equation 10 can be used to calculate the field at any point (r, ϕ) if the coefficients c_m of the Fourier series of $A(\theta)$ are known. Note that there are no practical restrictions for the desired field. There were no assumptions during the derivation, except that the field is a wave, i.e. it satisfies the Helmholtz wave equation and, consequently, its Fourier transform is nonzero only on the circle with radius $\frac{2\pi}{\lambda}$.

Equation 10 is in fact a representation of the diffraction field $u(r, \phi)$ as a linear combination of the functions:

$$\psi_m(r, \phi) = e^{jm(\phi + \frac{\pi}{2})} J_m \left(\frac{2\pi}{\lambda} r \right). \quad (11)$$

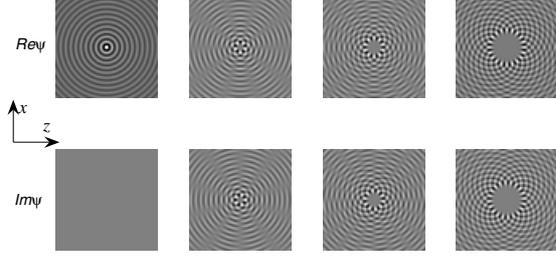


Fig. 1. Plots of some basis functions $\psi_m(r, \phi)$ in cartesian coordinate system (x, z) for $m = 0, 5, 10, 20$ (from left to right). The top and the bottom rows depict the real and imaginary parts, respectively.

It is interesting to note that these functions are *separable*. They are product of a Bessel function which depends only on r and an exponent which depends only on ϕ . The separability property can be used to show that the functions $\psi_m(r, \phi)$ are mutually *orthogonal*:

$$\begin{aligned}
\langle \psi_m(r, \phi), \psi_n(r, \phi) \rangle &= \int_0^{2\pi} \int_0^{\infty} \psi_m(r, \phi) \psi_n^*(r, \phi) d\phi dr \\
&= \int_0^{2\pi} \int_0^{\infty} e^{jm(\phi+\frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda}r\right) e^{-jn(\phi+\frac{\pi}{2})} J_n\left(\frac{2\pi}{\lambda}r\right) d\phi dr \\
&= \int_0^{\infty} J_m\left(\frac{2\pi}{\lambda}r\right) J_n\left(\frac{2\pi}{\lambda}r\right) dr \int_0^{2\pi} e^{j(m-n)(\phi+\frac{\pi}{2})} d\phi \\
&= 0 \text{ for } m \neq n.
\end{aligned} \tag{12}$$

Therefore the set $\{\psi_m(r, \phi)\}_{m=-\infty}^{\infty}$ forms an orthogonal basis. As derived, the basis is suitable for decomposition of signals representing light fields. Figure 1 shows some basis functions for $m = 0, 5, 10$ and 20 . The functions are shown with respect to the cartesian coordinate system (x, z) . As one can see, they are smooth circular oscillatory functions, whose action moves away from the origin when the number m increases.

3. RECONSTRUCTION FROM NON-UNIFORMLY DISTRIBUTED SAMPLES

For numerical computation the infinite summation in the model from Eq. 10 is impractical. A finite dimensional version of the model can be obtained by taking only the partial sum of M basis elements:

$$u(r, \phi) = \frac{4\pi^2}{\lambda} \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} c_m e^{jm(\phi+\frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda}r\right). \tag{13}$$

We will use the polar model of Eq. 13 for the problem of non-uniform sampling of diffraction fields [2]. The problem can be stated for polar coordinates as reconstruction of $u(r, \phi)$ from a finite set of s sampling points $\{(r_i, \phi_i)\}_{i=1}^s$. The field can be calculated if the coefficients c_m are known. A system of s equations for c_m can be constructed by writing Eq. 13 for each point in the irregular sampling set:

$$u(r_i, \phi_i) = \frac{4\pi^2}{\lambda} \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} c_m e^{jm(\phi_i+\frac{\pi}{2})} J_m\left(\frac{2\pi}{\lambda}r_i\right), i = 1 \dots s. \tag{14}$$

This system is linear and can be stated in matrix form:

$$\mathbf{u} = \mathbf{J}\mathbf{c}, \tag{15}$$

where $\mathbf{c} = [c_{-\lfloor M/2 \rfloor}, c_{-\lfloor M/2 \rfloor+1}, \dots, c_{\lfloor M/2 \rfloor}]^T$ is the unknown vector of the coefficients and the vector of given samples is $\mathbf{u} = [u(r_1, \phi_1), u(r_2, \phi_2), \dots, u(r_s, \phi_s)]^T$. \mathbf{J} is the $s \times M$ reconstruction matrix

$$\mathbf{J} = \{\mathcal{J}_{i,l}\} = \{e^{j(l-\lfloor M/2 \rfloor-1)(\phi_i+\frac{\pi}{2})} J_{l-\lfloor M/2 \rfloor-1}\left(\frac{2\pi}{\lambda}r_i\right)\}, i = 1, \dots, s, l = 1, \dots, M. \tag{16}$$

The straightforward approach to solve Eq. 15 is to take the pseudo-inverse. It has two different forms - for the over-determined case ($s > M$) and the under-determined case ($s < M$). In this work the latter case is ignored, because then the field can never be reconstructed, and therefore the results are not of practical interest. The major drawback of solving the problem by computing the pseudo-inverse is the high computational costs which grow when the number of given samples s increases. In linear programming [8] there exists a myriad of iterative algorithms for solving linear systems. Conjugate gradient method applied on normal equations (CGN) [8] is one of the most powerful algorithms for inversion of rectangular matrices. It proceeds as follows:

1. compute the matrix \mathbf{J} according to Eq. 16 and $\mathbf{b} = \mathbf{J}^H \mathbf{u}$;
2. initialize $\hat{\mathbf{c}}^{[0]}$ arbitrary, $\mathbf{g}_0 = \mathbf{b} - \mathbf{J}^H \mathbf{J} \mathbf{c}$ and $\mathbf{d}_0 = -\mathbf{g}_0$
3. for $n = 1$ to n_{it}
 - (a) $\alpha = \frac{\mathbf{g}_n^T \mathbf{g}_n}{\mathbf{d}_n^T \mathbf{J}^H \mathbf{J} \mathbf{d}_n}$
 - (b) $\hat{\mathbf{c}}^{[n+1]} = \hat{\mathbf{c}}^{[n]} + \alpha \mathbf{d}_n$
 - (c) $\mathbf{g}_{n+1} = \mathbf{g}_n + \alpha \mathbf{J}^H \mathbf{J} \mathbf{d}$
 - (d) $\gamma = -\frac{\mathbf{g}_{n+1}^T \mathbf{g}_{n+1}}{\mathbf{g}_n^T \mathbf{g}_n}$
 - (e) $\mathbf{d}_{n+1} = -\mathbf{g}_{n+1} + \gamma \mathbf{d}_n$
- end
4. reconstruct the diffraction field $u(r, \phi)$ from the estimated Fourier coefficient vector $\hat{\mathbf{c}}$ with Eq. 13.

4. EXPERIMENTAL RESULTS

The non-uniform sampling problem outlined in section 3 was simulated in two experiments in order to evaluate the finite dimensional model of Eq. 13. In the experiments the given samples were generated by Eq. 14, where $M = 256$ coefficients c_m were chosen as a Gaussian pulse centered at the origin. The positions of the samples are randomly chosen inside a spatial rectangle, centered at the origin and lying in the half-plane $z > 0$. Assessment of the results is based on the normalized error between the original \mathbf{c} and reconstructed $\hat{\mathbf{c}}$ coefficient vectors - $e = \|\mathbf{c} - \hat{\mathbf{c}}\|_2 / \|\mathbf{c}\|_2$.

Samples whose positions are mirror images with respect to $z = 0$ of the positions of the given samples were added to complete the initial information during the experiments. The values of the new samples are complex conjugate to the original values, as can be seen from Eq. 14. The addition of the new samples speeds up the convergence of CGN.

The goal of the first experiment is to track the behavior of CGN for different numbers of given samples. The number of iterations is kept fixed to the dimensionality of the problem $n_{it} = 256$, because this number is theoretically sufficient [8]. Figure 2(a) illustrates this experiment. As expected, the drop-off value of the curve is reached for values of s slightly higher than M . The second experiment illustrates the convergence of the CGN method (Figure 2(b)). In this experiment the number of iterations n_{it} is increased from zero to 200 while the number of given samples s is kept fixed to 384. This number is chosen to be $1.5M$, because in the first experiment this is the value where the error curve in Figure 2(a) saturates.

During the experiments for each value of s , 50 different vectors $\hat{\mathbf{c}}$ are reconstructed using s randomly chosen data points. Each reconstructed vector corresponds to a different random choice of the positions of the s known samples. The final error estimate for a value of s is an average of the errors of all 50 choices.

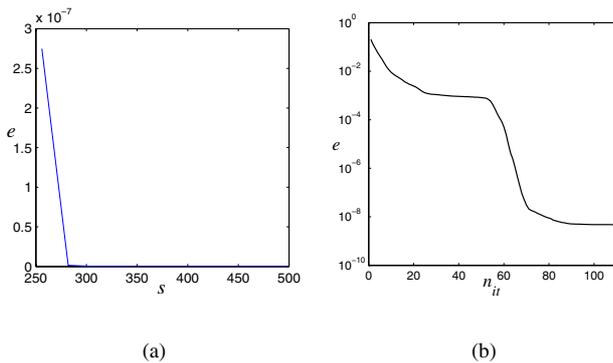


Fig. 2. Experimental results (a) Normalized error e for different number of known samples s when the iterations are $n_{it} = 256$ (b) Convergence of the CGN algorithm when the number of given samples is $s = 384$.

5. CONCLUSION

The main goal of this paper is the derivation of discrete diffraction model suitable for non-uniform sampling and reconstruction of monochromatic light fields. It is practically feasible because its formulation is in terms of Bessel functions, for which there are accurate computational algorithms. The main advantage of the model is that it does not impose any restrictions on the field. Moreover, the model is a separable orthogonal basis where any diffraction field can be decomposed. The model was evaluated for reconstruction of diffraction fields from a set of non-uniformly distributed samples. The reconstruction converges fast for number of given samples slightly larger than the dimensionality of the model.

6. REFERENCES

- [1] L. Onural and H. M. Ozaktas, "Signal processing issues in diffraction and holographic 3dvt," in *Proceedings of 13-th Signal Processing Conference - EUSIPCO*, Antalya, Turkey, 2005.
- [2] V. Uzunov, A. Gotchev, G. B. Esmer, L. Onural, and H. M. Ozaktas, "Non-uniform sampling and reconstruction of diffraction field," in *Proceedings of The 2006 SMMSP Workshop*, Florence, Italy, 2007, pp. 191–197.
- [3] G. B. Esmer, V. Uzunov, L. Onural, H. M. Ozaktas, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," *Image Communication*, vol. Special Issue on Three-Dimensional Video and Television.
- [4] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [5] L. Onural, "Sampling of the diffraction field," *Applied Optics*, vol. 39, pp. 5929–5935, November 2000.
- [6] E. W. Weisstein, "Bessel function of the first kind," From MathWorld – A Wolfram Web Resource, available at <http://mathworld.wolfram.com/BesselFunctionoftheFirstKind.html>.
- [7] D. E. Amos, "A portable package for bessel functions of a complex argument and nonnegative order," *Trans. Math. Software*, 1986.
- [8] S. G. Nash and A. Sofer, *Linear and nonlinear programming*, McGraw-Hill, 1996.

Publication P6

V. Uzunov, A. Gotchev, K. Egiazarian, “Convergence and error analysis of diffraction field iterative non-uniform sampling schemes,” in *Proceedings of the 2008 International Workshop on Local and Non-Local Approximation in Image Processing, LNLA2008, TICSP Series*, vol. 44, Lausanne, Switzerland, 23-24 August 2008, pp. 213-220.

CONVERGENCE AND ERROR ANALYSIS OF DIFFRACTION FIELD ITERATIVE NON-UNIFORM SAMPLING SCHEMES

Vladislav Uzunov, Atanas Gotchev, Karen Egiazarian

Department of Signal Processing, Tampere University of Technology,
P.O.Box 553, FIN-33101 Tampere, FINLAND

ABSTRACT

Reconstruction of a light field from a set of non-uniformly distributed samples is important for holographic type of 3D scene display. The non-uniform sampling problem has been already addressed and rapidly convergent iterative schemes developed, mainly based on the method of Projection Onto Convex Sets (POCS) and matrix inversion by the method of Conjugate Gradients (CG). This paper presents a derivation of the convergence rates for both of the schemes and analyzes the factors which influence the error decay. The presence of clusters and gaps in the distribution of the given samples turns out to have the most significant influence on the convergence. Based on the analysis, a regularization method is derived for the CG-based reconstruction scheme which proves to be powerful even when the local sample density estimate is rough. The POCS-based scheme is generalized for the case of off-grid known samples. Its error analysis leads to similar conclusions for the convergence impact factors as in the CG-based scheme. In addition, the analysis suggests scenarios where the POCS excels, and others where the CG is much better.

1. INTRODUCTION

The computation of the light field over the entire three-dimensional (3D) space from an abstract 3D scene representation is known as the forward problem in holography. Holographic 3D display systems use the light field information to display a scene, and thus benefit from efficient techniques which solve the forward problem. The development of such techniques is a challenging task, since 3D scenes consist of various shapes and shades. In a general setting, the 3D scene information can be assumed to be available as non-uniformly distributed in space light field points. Thus the forward problem can be formulated as to reconstruct a light field from irregularly distributed samples.

A real-time application needs to reconstruct the field for large data sample sets in fast and accurate manner. Previously, we have developed iterative reconstruction schemes, based mainly on the method of Projection Onto Convex Sets (POCS) [1], and matrix inversion through the Conjugate Gradients (CG) method [2]. The POCS-based

scheme assumes that the known field samples belong to a pre-defined regular grid, and utilizes an $O(N \log N)$ -complex propagation algorithm to iterate between successive lines which define the convex sets. The CG-based scheme builds a discrete model of the light field which represents the reconstruction task as matrix inversion problem, and solves it iteratively through CG. However, for certain sample distributions the involved matrix has high condition number and inversion with CG converges slowly.

The main goal of this paper is to perform a theoretical analysis of the convergence rate of the POCS- and CG-based reconstruction schemes. We aim at identifying and analyzing the main factors which influence the convergence and the computational efficiency, and to illustrate their impact. The results can be used to achieve faster convergence and/or better computational efficiency.

2. PROBLEM FORMULATION

In this section, we formulate the problem of reconstruction of a diffraction field from its sampled version, where the sampling coordinates are given on a non-uniform grid. We start with a brief review of diffraction theory equations we need, then we specify a finite dimensional model of diffraction to be used in the reconstruction.

2.1. A finite-dimensional model of diffraction

Diffraction theory studies phenomena occurring during propagation and reflection of light within, and at the borders of different physical media. In general, these phenomena are described through complex-valued functions representing light (optical) wave fields and their distribution across space. In the general case, these are vector functions confined to the Maxwell equations [3]. When the medium is free space and the light source is monochromatic with wavelength λ , the corresponding light field is accurately described by a scalar function. Under the conditions of linearity, isotropy and homogeneity of the free space, the light field at any spatial point can be related to that on a plane through the Rayleigh-Sommerfeld diffraction integral [4]. The same integral can be used to compute the optical field on a plane given the field on another parallel plane, having linear and shift-invariant relationship between the two functions [4, 5]. Under the assumption that the distance r between the parallel planes satisfies $r \gg \lambda$, the Rayleigh-Sommerfeld diffraction inte-

This work was supported by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence program (2006 - 2011)).

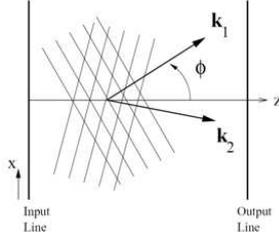


Figure 1. The vectors $\mathbf{k}_1 = [k_{1x}, k_{1z}]$ and $\mathbf{k}_2 = [k_{2x}, k_{2z}]$ are the wave vectors of the plane waves. The field at the output line is equal to the superposition of the plane waves illuminated on the input line according to Eq. 1 [1].

gral can be regarded as an equivalent to the so-called *plane wave decomposition* approach [4, 6, 7].

In the rest of the paper, the plane wave decomposition approach is used. For the sake of simplicity, the discussion is restricted to one transverse dimension only. The diffraction field at any point in the 2D free space is computed as:

$$u(x, z) = \int_{-\frac{2\pi}{\lambda}}^{\frac{2\pi}{\lambda}} a(k_x) e^{j(k_x x + k_z z)} dk_x, \quad (1)$$

where $u(x, z)$ is the field over 2D space, the x axis is the transverse axis and the z axis is the optical axis along which the field propagates. This integral represents the light field $u(x, z)$ as a superposition of plane waves $e^{j(k_x x + k_z z)}$, where each plane wave is oriented at an angle determined by the wave vector $\mathbf{k} = [k_x, k_z]$, as illustrated in Figure 1. The variable k_z is related to k_x by $k^2 = k_z^2 + k_x^2$ where $k = \frac{2\pi}{\lambda}$ is the length of each wave vector. In the superposition integral each plane wave contributes with complex amplitude $a(k_x)$. This superposition of plane waves bears another interpretation, which is more convenient for our purpose. The variables k_x and k_z are the spatial frequencies of the propagating plane waves along the x and z axes respectively. The angles of the plane waves determine different spatial frequencies k_x along the transversal dimension x . Therefore, different plane waves determine harmonics with different frequencies k_x in the transverse plane. Consequently, the complex amplitudes $a(k_x)$ can be considered as the Fourier spectrum of the field $u(x, 0)$ on a reference line with $z = 0$. A propagating monochromatic wave with wavelength λ cannot have a harmonic component in the transverse plane with higher frequency than $\frac{2\pi}{\lambda}$ since $k_x = k \cos \phi$ (cf. Figure 1). Therefore, in the case of free space propagation of monochromatic light, the function on the reference line $u(x, 0)$ is assumed band-limited. This also determines the limits of the integral in Eq. 1.

To elucidate the role of the function $a(k_x)$, let us consider the plane waves in Eq. 1 split as $e^{j(k_x x + k_z z)} =$

$e^{j\sqrt{k^2 - k_x^2} z} e^{jk_x x}$. The plane wave decomposition integral takes the form:

$$u(x, z) = \int_{-\frac{2\pi}{\lambda}}^{\frac{2\pi}{\lambda}} a(k_x) e^{j\sqrt{k^2 - k_x^2} z} e^{jk_x x} dk_x. \quad (2)$$

Now interpret Eq. 2 as inverse Fourier integral where the function $a(k_x)_{z_i} = a(k_x) e^{j\sqrt{k^2 - k_x^2} z_i}$ is the Fourier transform of the field $u(x, z_i)$ at a line slice $z = z_i$. In other words, the field $u(x, z)$ at any particular line $z = z_i$ can be computed from the field on the initial line $u(x, 0)$ by three consecutive steps: (1) Fourier transform to find $a(k_x)$; (2) Fourier-domain multiplication by the transfer function of the free space $e^{j\sqrt{k^2 - k_x^2} z}$; and (3) inverse Fourier transform to get back to spatial domain. This interpretation has both theoretical and numerical importance. Theoretically, it clearly shows the dependence between input, system and output. Numerically, it suggests the utilization of an FFT algorithm, under proper sampling [1].

According to Eq. 1, every point of the field $u(x, z)$ can be expressed in terms of the spectrum $a(k_x)$ of the field $u(x, 0)$ on the reference line. Consider an *essentially* space-limited function $u(x, 0)$ or, more precisely, a function whose space - bandwidth product is finite, and denote the spatial extent of interest by T . Such a function can be periodized, which is equivalent to discretization of $a(k_x)$. A periodic and bandlimited function is equivalently represented by a finite number (M) of discrete spectral components. That is, the discrete values of the frequency k_x are :

$$k_x = \frac{2\pi m}{T}, \quad (3)$$

where $m = -\lfloor \frac{M}{2} \rfloor, \dots, \lfloor \frac{M-1}{2} \rfloor$. Substituting k_x as given by Eq. 3 in Eq. 2 leads to a *finite-dimensional* model of diffraction:

$$u(x, z) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j\frac{2\pi}{T} \sqrt{\frac{T^2}{\lambda^2} - m^2} z} e^{j\frac{2\pi}{T} m x}, \quad (4)$$

where $a_m = a(\frac{2\pi}{T} m)$ are the coefficients of the Fourier series expansion of $u(x, 0)$. This model assumes the field at the reference line $z = 0$ is a trigonometric polynomial of order M , which is easily transferable to any other z coordinate though the transfer function $e^{j\sqrt{k^2 - k_x^2} z}$.

2.2. Non-uniformly sampled diffraction fields

The finite-dimensional model in Eq. 4 can be regarded as a function expansion with respect to M kernels formed as product of exponents along x axis and chirps along z axis, weighted by the coefficients a_m . These M generators determine any particular monochromatic field with wavelength λ and spatial extend of the reference line T . In the general case of irregular sampling, the diffraction field $u(x, z)$ is given at a finite set of s sampling points $\{(x_i, z_i)\}_{i=1}^s$. The irregular sampling and reconstruction problem can be stated as to find the unknown field generating coefficients a_m , given the samples $\{u(x_i, z_i)\}_{i=1}^s$.

Eq. 4 can be written for each point in the irregular sampling set $\{(x_i, z_i)\}_{i=1}^s$:

$$u(x_i, z_i) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j \frac{2\pi}{T} \sqrt{\beta^2 - m^2} z_i} e^{j \frac{2\pi}{T} m x_i}, \quad (5)$$

for $i = 1, \dots, s$.

These s equations form a linear system for the M unknowns a_m . Speaking linear algebra, this system will have a solution if it is of full rank. To characterize the stability of the sampling set with respect to the field generators, we are interested in having a sampling set bounded by the energy of the field at the first line, as

$$A \|a\|^2 \leq \sum_{i=1}^s |u(x_i, z_i)|^2 \leq B \|a\|^2. \quad (6)$$

Such set of sampling points $\{(x_i, z_i)\}_{i=1}^s$ which has bounded energy for any choice of generating coefficients a_m with strictly positive bounds A and B is referred to as *stable sampling set* [8].

In the following we describe and analyze two different approaches to find the coefficients a_m . The first uses iterative algorithm to solve the system of Eq. 5 [2]. The second algorithm iterates from given point to given point and falls in the framework of the method of POCS [9, 10].

3. RECONSTRUCTION BY ITERATIVE MATRIX INVERSION

3.1. Method description

The system in Eq. 5 is linear and can be expressed in a matrix form:

$$\mathbf{u} = \mathbf{R}\mathbf{a}, \quad (7)$$

where $\mathbf{a} = [a_{-\lfloor \frac{M}{2} \rfloor}, a_{-\lfloor \frac{M}{2} \rfloor + 1}, \dots, a_{\lfloor \frac{M-1}{2} \rfloor}]^T$ is the unknown vector of the field generating coefficients and $\mathbf{u} = [u(x_1, z_1), u(x_2, z_2), \dots, u(x_s, z_s)]^T$ is the vector of given samples. \mathbf{R} is the reconstruction matrix

$$\begin{aligned} \mathbf{R} &= \{r_{p,q}\} \\ &= \{e^{j \frac{2\pi}{T} \sqrt{\beta^2 - (q - \lfloor \frac{M}{2} \rfloor + 1)^2} z_p} e^{j \frac{2\pi}{T} (q - \lfloor \frac{M}{2} \rfloor + 1) x_p}\}, \\ &p = 1, \dots, s, q = 1, \dots, M, \end{aligned} \quad (8)$$

The diffraction field at the point (x_i, z_i) equals the inner product of the i -th row of \mathbf{R} with \mathbf{a} .

We need a fast and numerically stable algorithm to find (approximate) solution $\hat{\mathbf{a}}$ for the unknown vector \mathbf{a} and we limit the consideration to the over-determined case $s \geq M$.

Expressing the residual between the true vector and its approximation

$$\mathbf{g} = \mathbf{u} - \mathbf{R}\hat{\mathbf{a}}, \quad (9)$$

and minimizing the L_2 norm of this residual $\|\mathbf{u} - \mathbf{R}\hat{\mathbf{a}}\|_2$, one ends with solving a *Least Squares (LS)* problem. Reaching an LS solution goes through finding the (pseudo-)inverse

of the matrix \mathbf{R} , which operation is of cubic complexity in the general case [11]. We opt for the conjugate gradient method (CG), known as one of the most rapidly convergent and numerically stable algorithms for solving LS problems iteratively [12]. However, it requires a Hermitian and positive definite matrix. When the matrix is rectangular, as in our case, we can consider the following equivalent matrix equation that will produce the same solution as Eq. 7:

$$\mathbf{R}^H \mathbf{R} \mathbf{a} = \mathbf{R}^H \mathbf{u}. \quad (10)$$

Now the matrix to be solved by CG is the Hermitian matrix $\mathbf{R}^H \mathbf{R}$. The residual whose norm is minimized by iterating CG in this case is measured as:

$$\mathbf{g} = \mathbf{R}^H \mathbf{u} - \mathbf{R}^H \mathbf{R} \hat{\mathbf{a}}. \quad (11)$$

CG applied on the matrix $\mathbf{R}^H \mathbf{R}$ has a form, where the matrix $\mathbf{R}^H \mathbf{R}$ is never explicitly computed. This form is called CG method on normal equations (CGN) and is outlined below [12]:

1. initialize $\mathbf{b} = \mathbf{R}^H \mathbf{u}$ $\hat{\mathbf{a}}_0$ arbitrary, $\mathbf{g}_0 = \mathbf{b} - \mathbf{R}^H \mathbf{R} \hat{\mathbf{a}}_0$ and $\mathbf{d}_0 = \mathbf{g}_0$
2. for $n = 1$ to n_{it}
 - (a) $\alpha = \frac{\mathbf{g}_n^H \mathbf{g}_n}{\mathbf{d}_n^H \mathbf{R}^H \mathbf{R} \mathbf{d}_n}$
 - (b) $\hat{\mathbf{a}}_{n+1} = \hat{\mathbf{a}}_n + \alpha \mathbf{d}_n$
 - (c) $\mathbf{g}_{n+1} = \mathbf{b} - \mathbf{R}^H \mathbf{R} \hat{\mathbf{a}}_{n+1}$
 - (d) $\gamma = \frac{\mathbf{g}_{n+1}^H \mathbf{g}_{n+1}}{\mathbf{g}_n^H \mathbf{g}_n}$
 - (e) $\mathbf{d}_{n+1} = \mathbf{g}_{n+1} + \gamma \mathbf{d}_n$

end

3. reconstruct the diffraction field $u(x, z)$ from the estimated field generating coefficient vector $\hat{\mathbf{a}}$ with Eq. 4.

Any iterative algorithm for solving LS problems builds the solution step by step, updating the solution vector each time with a small portion (α) along some search direction \mathbf{d}_n (step 2b). The basic idea of CG is to build the search directions \mathbf{d}_n conjugate to each other, so that after at most M steps the solution will be found. By conjugate is meant that the directions are orthogonal to each other, where the orthogonality is measured with respect to the $\mathbf{R}^H \mathbf{R}$ matrix of the LS problem - $\mathbf{d}_n^H \mathbf{R}^H \mathbf{R} \mathbf{d}_k = 0$. The value of α is chosen in such a manner that the current error $\mathbf{e}_{n+1} = \hat{\mathbf{a}}_{n+1} - \mathbf{a}$ is conjugate to the previous direction \mathbf{d}_n (step 2a). This makes the residual \mathbf{g}_{n+1} orthogonal to all previous search directions. The new direction \mathbf{d}_{n+1} is build from this residual \mathbf{g}_{n+1} as to be conjugate to all previous directions (step 2d-e).

3.2. Convergence analysis

This subsection presents the derivation of the approximation error at the n -th CG iteration in terms of the initial error. Then, factors influencing the final expression can be analyzed with respect to different sampling distributions. The analysis can be used as a base to decrease the impact of the distributions on the convergence rate.

Write the error vector \mathbf{e}_{n+1} in the form:

$$\begin{aligned} \mathbf{e}_{n+1} &= \mathbf{a}_{n+1} - \mathbf{a} = \mathbf{a}_n + \alpha_n \mathbf{d}_n - \mathbf{a} = \mathbf{e}_n + \alpha_n \mathbf{d}_n \\ &= \mathbf{e}_{n-1} + \alpha_{n-1} \mathbf{d}_{n-1} + \alpha_n \mathbf{d}_n = \dots \\ &= \mathbf{e}_0 + \sum_{k=0}^n \alpha_k \mathbf{d}_k. \end{aligned} \quad (12)$$

This equation shows that \mathbf{e}_{n+1} belongs to a space $\mathbf{e}_0 + D_{n+1}$, where $D_{n+1} = \text{span}\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_n\}$. The search directions are built from the residuals (step 2e) and therefore $D_{n+1} = \text{span}\{\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_n\}$. Now denote $\mathbf{Q} = \mathbf{R}^H \mathbf{R}$ and write each residual \mathbf{g}_n in the form:

$$\begin{aligned} \mathbf{g}_n &= -\mathbf{Q}\mathbf{e}_n = -\mathbf{Q}(\mathbf{e}_{n-1} + \alpha_{n-1} \mathbf{d}_{n-1}) \\ &= \mathbf{g}_{n-1} - \alpha_{n-1} \mathbf{Q}\mathbf{d}_{n-1}. \end{aligned} \quad (13)$$

Recall that $\mathbf{d}_{n-1} \in D_n$ and therefore each new subspace D_{n+1} is constructed from D_n and the subspace $\mathbf{Q}D_n$. Hence,

$$\begin{aligned} D_n &= \text{span}\{\mathbf{g}_0, \mathbf{Q}\mathbf{g}_0, \mathbf{Q}^2\mathbf{g}_0, \dots, \mathbf{Q}^{n-1}\mathbf{g}_0\} \\ &= \text{span}\{\mathbf{Q}\mathbf{e}_0, \mathbf{Q}^2\mathbf{e}_0, \mathbf{Q}^3\mathbf{e}_0, \dots, \mathbf{Q}^n\mathbf{e}_0\}. \end{aligned} \quad (14)$$

The subspaces D_n are known as Krylov subspaces. Recall that the error \mathbf{e}_n belongs to the space $\mathbf{e}_0 + D_n$. Then it can be expressed as a linear combination of the spanning elements $\mathbf{Q}^i \mathbf{e}_0$ of this subspace:

$$\mathbf{e}_n = \left(I + \sum_{i=1}^n \phi_i \mathbf{Q}^i \right) \mathbf{e}_0 = P_n(\mathbf{Q})\mathbf{e}_0. \quad (15)$$

$P_n(\mathbf{Q})$ is the polynomial of order n from the parenthesis of the above expression. If \mathbf{v} is an eigenvector of \mathbf{Q} with respective eigenvalue λ , then $\mathbf{Q}^i \mathbf{v} = \lambda^i \mathbf{v}$ and therefore $P_n(\mathbf{Q})\mathbf{v} = P_n(\lambda)\mathbf{v}$. Since the matrix \mathbf{Q} is Hermitian, the eigenvectors can be chosen to form an orthonormal set and the error \mathbf{e}_0 can be expressed as a linear combination of these eigenvectors. Inserting this linear combination in the last form of Eq. 15, \mathbf{e}_n can be expressed as similar linear combination of the eigenvectors, where the coefficients of this combination are multiplied by $P_n(\lambda_i)$. Now the \mathbf{Q} -norm of the error can be written as [13]:

$$\|\mathbf{e}_n\|_{\mathbf{Q}} \leq \min_{P_n} \max_{\lambda \in \Lambda(\mathbf{Q})} P_n^2(\lambda) \|\mathbf{e}_0\|_{\mathbf{Q}}. \quad (16)$$

CG determines alpha in such a manner which minimizes $\|\mathbf{e}_n\|_{\mathbf{Q}}$ within the space $D_n + \mathbf{e}_0$ (step 2a), or, in other words CG finds the polynomial $P_n(\lambda)$ which minimizes the expression in Eq. 16. However, the convergence is determined by the eigenvalue which gives maximum value of $P_n(\lambda)$.

Since our goal is to analyze and estimate the error reduction after n iterations of CG, we need to find the polynomial $P_n(\lambda)$ explicitly. This means that the coefficients ϕ_i in Eq. 15 must be known. However, they depend on the coefficients α_i and γ_i which are not available before the algorithm is run. More general approach is to find a polynomial which minimize Eq. 16 over the interval $[\lambda_{min}, \lambda_{max}]$ between the smallest and largest eigenvalue, rather than on particular set of eigenvalues $\Lambda(\mathbf{Q})$. Such a polynomial is known to be a relation of two Chebyshev polynomials T_n of order n [12]:

$$P_n(\lambda) = \frac{T_n\left(\frac{\lambda_{max} + \lambda_{min} - 2\lambda}{\lambda_{max} - \lambda_{min}}\right)}{T_n\left(\frac{\lambda_{max} + \lambda_{min}}{\lambda_{max} - \lambda_{min}}\right)}. \quad (17)$$

The polynomial in the numerator has maximal value of 1 inside the interval $[\lambda_{min}, \lambda_{max}]$ and therefore the error \mathbf{e}_n can be estimated as

$$\begin{aligned} \|\mathbf{e}_n\|_{\mathbf{Q}} &\leq \left[T_n\left(\frac{\lambda_{max} + \lambda_{min}}{\lambda_{max} - \lambda_{min}}\right) \right]^{-1} \|\mathbf{e}_0\|_{\mathbf{Q}} \\ &= \left[T_n\left(\frac{\kappa + 1}{\kappa - 1}\right) \right]^{-1} \|\mathbf{e}_0\|_{\mathbf{Q}} \\ &= 2 \left[\left(\frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1} \right)^n + \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \right]^{-1} \|\mathbf{e}_0\|_{\mathbf{Q}}, \end{aligned} \quad (18)$$

where $\kappa = \lambda_{max}/\lambda_{min}$ is the condition number of the matrix \mathbf{Q} . The second addend inside the square brackets converges to zero as n grows, so it is common to estimate the error with the weaker inequality

$$\|\mathbf{e}_n\|_{\mathbf{Q}} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \|\mathbf{e}_0\|_{\mathbf{Q}}. \quad (19)$$

This equation suggests that the convergence of the CG algorithm highly depends on the condition number of the matrix \mathbf{Q} . To relate κ to the bounds a stable sampling set (cf. Eq. 6), the energy of the known samples is can be expressed as:

$$\|\mathbf{u}\|_2^2 = \mathbf{u}^H \mathbf{u} = \mathbf{a}^H \mathbf{R}^H \mathbf{R} \mathbf{a} = \mathbf{a}^H \mathbf{Q} \mathbf{a}. \quad (20)$$

For a stable sampling set the energy of the samples is strictly positive and thus the matrix \mathbf{Q} is strictly positive definite, i.e. its eigenvalues are always positive. Since the energy of the samples is bounded for all vectors \mathbf{a} , consider the case when it is an eigenvector \mathbf{v} of \mathbf{Q} :

$$\begin{aligned} A \|\mathbf{v}\|_2^2 &\leq \mathbf{v}^H \mathbf{Q} \mathbf{v} \leq B \|\mathbf{v}\|_2^2, \text{ or} \\ A \|\mathbf{v}\|_2^2 &\leq \lambda \mathbf{v}^H \mathbf{v} \leq B \|\mathbf{v}\|_2^2. \end{aligned} \quad (21)$$

This last equation implies that $A \leq \lambda \leq B$, for all eigenvalues of \mathbf{Q} . Therefore $A \leq \lambda_{min}$ and $\lambda_{max} \leq B$ and consequently the condition number of \mathbf{Q} can be estimated as $\kappa(\mathbf{Q}) \leq B/A$. The constants A and B are determined from the reconstruction function and the sampling set. Suppose that the coefficients \mathbf{a} determine a light field which is a beam centered along the z -axis, and has essentially finite extent along x . Suppose further that the

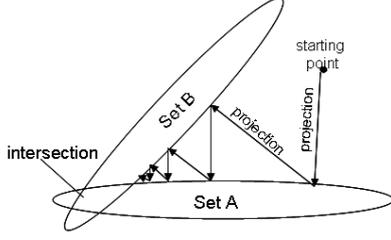


Figure 2. Projection Onto Convex Sets

samples do not cover the extent of the beam, leaving a large gap. The energy of such samples will be very small in comparison with the energy of \mathbf{a} , no matter how dense they are outside the gap. This pushes the lower bound A from Eq. 6 down and consequently increases $\kappa(\mathbf{Q})$. Now consider the case when we have a dense cluster of samples lying on the beam. The same local information will be added several times in Eq. 6, which will increase B and $\kappa(\mathbf{Q})$. Thus, the presence of clusters and gaps in the density of the sampling set has significant influence on the condition number of \mathbf{Q} .

An add-hoc approach to compensate for the clustering is to weight different samples according to their density. Samples from areas with higher density are assigned with small weights and samples from less populated areas are emphasized as more important with higher weights. This weighting can be represented in a matrix form as multiplying the known samples vector \mathbf{u} with a diagonal weighting matrix \mathbf{W} . Then, Eq. 7 is modified as follows:

$$\mathbf{W}\mathbf{u} = \mathbf{W}\mathbf{R}\mathbf{a}. \quad (22)$$

CG can be used to solve this LS problem by further multiplying both sides by \mathbf{R}^H to obtain a symmetric matrix $\mathbf{R}^H\mathbf{W}\mathbf{R}\mathbf{a}$ and target vector $\mathbf{R}^H\mathbf{W}\mathbf{u}$.

4. RECONSTRUCTION BY POCS

As a second approach we consider an iterative technique, developed based on the POCS method. POCS is a computational approach for finding an element of a feasible region defined by the intersection of a number of convex constraints, starting with an arbitrary infeasible point [14, 15]. Figure 2 illustrates how convergence to the intersection is achieved by iterative projections onto the individual convex sets. In our earlier work [1], POCS was used to reconstruct a diffraction field from a set of irregularly distributed samples, which belong to a predefined uniform grid. Here, the same algorithm is reformulated to serve the general case when the given field samples have random positions, unrelated to any particular uniform grid.

4.1. Method description

The POCS-based reconstruction method from [1] considers the known data samples as the constraints that determine the convex sets. Naturally, a convex set is defined

as all possible diffraction fields that have the given data points on a certain line. This is beneficial only in the case these points belong to a uniform grid with sampling interval T/M , and aligned with the starting sampling grid on the reference line. In such a case, the algorithm makes use of an FFT-based computation of diffraction field from line to line. This was the key-point of the algorithm developed in [1]. There, the algorithm was based on iterating from line to line efficiently by the use of FFT, and re-substituting the values of the known data points at each line, as they were always on positions of an M -point regular grid. However, in the case of no structured samples such an iteration is not possible. Consider, e.g. the case of single point per line, or the case of dense points not aligned with respect to the starting grid.

Here, we suggest a modification to overcome these difficulties. We still want to benefit from using FFT. However, instead of propagating from line to line, we propagate the *new information* from each particular convex set $\mathbf{C}_l, l = 1 \dots L$ to the unknown coefficients a_m . In order to use an M -point IDFT for this propagation, we define the set \mathbf{C}_l as all known samples $u(x_i, z_i)$ with $z_i = z_l$, which can be hosted by an M -point uniform grid \mathbf{x}_l . In the general case there might be no more than one sample per line to be considered as belonging to the set \mathbf{C}_l . In such a case, the number of convex sets L will coincide with the number of sampling points s and each projection will result to an update caused by a single point. However, if two or more points per line happen to belong to uniform grids, we should benefit from this. Again, in the general case each such grid \mathbf{x}_l will be not aligned to the initial grid and can be assumed centered at some point $\chi_l \neq 0$ and encompassing a spatial interval of length T along x :

$$\mathbf{x}_l = \{(\mathbf{x}_l)_k\}_{k=0}^{M-1} = \chi_l + \{-T/2 + kT/M\}_{k=0}^{M-1} \quad (23)$$

Now the formal definition of a set $\mathbf{C}_l, l = 1 \dots L$ can be stated as:

$$\begin{aligned} \mathbf{C}_l &= \{\forall f(x, z) \in \mathbf{DF}_M : f(\mathbf{p}_l) = u(\mathbf{p}_l)\}, \\ \mathbf{p}_l &= \{(x_i, z_i), i = 1 \dots s : z_i = z_l, x_i \in \mathbf{x}_l\}, \end{aligned} \quad (24)$$

where \mathbf{p}_l is the set of points which are hosted by the grid \mathbf{x}_l and \mathbf{DF}_M denotes the space of diffraction fields, generated by M nonzero coefficients through Eq. 4.

To complete the POCS method, the projections onto the sets have to be determined. Consider a field $f(x, y)$, generated through Eq. 4 by coefficients $a_m^{(f)}$. Its projection P_l onto \mathbf{C}_l can be defined as substituting its samples $f(\mathbf{p}_l)$ by the known samples $u(\mathbf{p}_l)$ which define the set, and reflecting this substitution onto the coefficients $a_m^{(f)}$. The substitution can be represented as just addition of the difference $u - f$ at the positions \mathbf{p}_l of the samples which define the set:

$$P_l f(x, z) = f(x, z) + \sum_{(x_i, z_i) \in \mathbf{p}_l} (u - f)(x_i, z_i) \delta(x - x_i, z - z_i), \quad (25)$$

where $\delta(x, z)$ is a Kronecker delta. The substitution acts only on line $z = z_l$ and on points \mathbf{p}_l which belong to

the grid \mathbf{x}_l . It is sufficient to propagate the information from the grid \mathbf{x}_l back to the coefficients f_m . According to Eq. 4, the Fourier series coefficients of the field $u(x, z)$ at $z = z_l$ are $a_m e^{j \frac{2\pi}{T} \sqrt{\frac{T^2}{\lambda^2} - m^2} z_l}$. These coefficients are related through M -point DFT with M regular samples on a grid \mathbf{x}_0 , centered at the origin. This can be represented in a matrix form as:

$$\mathbf{H}_{z_l} \mathbf{a} = \mathbf{F} u(\mathbf{x}_0, z_l), \quad (26)$$

where $\mathbf{H}_{z_l}^{-1} = \text{diag}(e^{j \frac{2\pi}{T} \sqrt{\frac{T^2}{\lambda^2} - m^2} z_l})$ and \mathbf{F} is the DFT matrix. Having a sampling grid \mathbf{x}_l for a set \mathbf{C}_l means it is shifted from the origin by χ_l . Alternatively, we can assume shifting the field $u_{\chi_l}(x, z) = u(x + \chi_l, z)$ prior to sampling on \mathbf{x}_0 . Eq. 4 performs this operation by modulation: $a_m e^{j \frac{2\pi}{T} m \chi_l}$. This changes Eq. 26 to:

$$\mathbf{H}_{z_l} \mathbf{E}_{\chi_l} \mathbf{a} = \mathbf{F} u(\mathbf{x}_l, z_l), \quad (27)$$

where $\mathbf{E}_{\chi_l} = \text{diag}(e^{j \frac{2\pi}{T} m \chi_l})$. Eventually,

$$\mathbf{a} = \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} u(\mathbf{x}_l, z_l) \quad (28)$$

is the equation which propagates information from the field line $u(x, z_l)$, sampled with a regular grid \mathbf{x}_l , to the coefficients a_m . Denote by \mathbf{f} an s -dimensional vector which contains the samples $f(x_i, z_i)$, $i = 1 \dots s$, ordered in the same way as the known samples $u(x_i, z_i)$ in the vector \mathbf{u} from Eq. 7. The difference $u - f$ sampled at the grid \mathbf{x}_l can be represented with the following equation:

$$(u - f)(\mathbf{x}_l, z_l) = \mathbf{S}_l (\mathbf{u} - \mathbf{f}), \quad (29)$$

where \mathbf{S}_l is a $M \times s$ permutation matrix which takes the samples $(u - f)(x_i, z_i)$, $(x_i, z_i) \in \mathbf{p}_l$ from the difference vector $\mathbf{u} - \mathbf{f}$ and positions them properly on the grid \mathbf{x}_l . Thus \mathbf{S}_l has value 1 on positions (k, i) , where i is such that $(x_i, z_i) \in \mathbf{p}_l$ and k are the positions of these samples on the grid \mathbf{x}_l .

Eq. 28 written for the sampled difference $(u - f)(\mathbf{x}_l, z_l)$ from Eq. 29 describes the back-propagation of the difference information to the field generating coefficients $a_m^{(f)}$, ordered in a vector $\mathbf{a}^{(f)}$. Consequently, the projection P_l can be re-defined to act on the field generating coefficient vector $\mathbf{a}^{(f)}$:

$$P_l \mathbf{a}^{(f)} = \mathbf{a}^{(f)} + \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l (\mathbf{u} - \mathbf{f}). \quad (30)$$

The POCS-based reconstruction algorithm proceeds as follows:

1. initialize $\hat{\mathbf{a}}_{0,0}$ arbitrary
2. for $n = 1$ to n_{it}
 - for $l = 1$ to L
 - (a) predict the field at the points of the given samples $\{(x_i, z_i)\}_{i=1}^s$ as $\hat{\mathbf{u}}^{(n,l)} = \mathbf{R} \hat{\mathbf{a}}_{n,l}$.
 - (b) calculate $\hat{\mathbf{a}}_{n,l+1} = P_l \hat{\mathbf{a}}_{n,l}$ by projecting the predicted field $\hat{u}^{(n,l)}(x, z)$ through Eq. 30

end

3. reconstruct the diffraction field $u(x, z)$ from the recovered field generating coefficient vector $\hat{\mathbf{a}}$ with Eq. 4.

4.2. Convergence analysis

We attempt on deriving an expression for the error obtained at the n -th iteration in terms of the error at previous iterations. Such expression can be analyzed to find factors which influence the decay of the error norm through the iterations, and the impact of these factors.

The error vector $\mathbf{e}_{n,l+1}$ which is obtained at the n -th iteration, after projecting on the set \mathbf{C}_l can be derived as follows:

$$\begin{aligned} \mathbf{e}_{n,l+1} &= \mathbf{a} - \hat{\mathbf{a}}_{n,l+1} = \mathbf{a} - P_l \hat{\mathbf{a}}_{n,l} \\ &= \mathbf{a} - \hat{\mathbf{a}}_{n,l} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l (\mathbf{u} - \hat{\mathbf{u}}^{(n,l)}) \\ &= \mathbf{e}_{n,l} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l \mathbf{R} (\mathbf{a} - \hat{\mathbf{a}}_{(n,l)}) \\ &= (\mathbf{I} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l \mathbf{R}) \mathbf{e}_{(n,l)}. \end{aligned} \quad (31)$$

In order to ensure convergence, that is $\|\mathbf{e}_{n,l+1}\| \leq \|\mathbf{e}_{n,l}\|$, the iteration matrix $(\mathbf{I} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l \mathbf{R})$ must be *non-expansive* i.e. it does not increase the the norm of a vector, when applied to it. A *strict* non-expansiveness of the matrix is sufficient condition and a non-strict non-expansiveness is a necessary condition for convergence. The norm of the iteration matrix is a measure for expansiveness:

$$\begin{aligned} \|\mathbf{I} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l \mathbf{R}\| &= \\ \|\mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} (\mathbf{F}^{-1} \mathbf{H}_{z_l} \mathbf{E}_{\chi_l} - \mathbf{S}_l \mathbf{R})\| &\leq \\ \|\mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F}\| \|\mathbf{F}^{-1} \mathbf{H}_{z_l} \mathbf{E}_{\chi_l} - \mathbf{S}_l \mathbf{R}\| &= \\ \|\mathbf{F}^{-1} \mathbf{H}_{z_l} \mathbf{E}_{\chi_l} - \mathbf{S}_l \mathbf{R}\|. \end{aligned} \quad (32)$$

The matrices \mathbf{H}_{z_l} and \mathbf{E}_{χ_l} represent propagation of the coefficients and shift in spatial domain along x by χ_l . According to the definition of the matrix norm, the derivation from Eq. 32 continues as:

$$\begin{aligned} \|\mathbf{F}^{-1} \mathbf{H}_{z_l} \mathbf{E}_{\chi_l} - \mathbf{S}_l \mathbf{R}\| &= \\ \max_{\mathbf{a}, \|\mathbf{a}\|=1} \|\mathbf{F}^{-1} \mathbf{H}_{z_l} \mathbf{E}_{\chi_l} \mathbf{a} - \mathbf{S}_l \mathbf{R} \mathbf{a}\| &= \\ \max_{\mathbf{a}, \|\mathbf{a}\|=1} \|u(\mathbf{x}_l, z_l) - \mathbf{S}_l \mathbf{u}\|. \end{aligned} \quad (33)$$

The operation $\mathbf{S}_l \mathbf{u}$ selects the samples $u(x_i, z_i)$, $(x_i, z_i) \in \mathbf{p}_l$ which determine the set \mathbf{C}_l from the vector \mathbf{u} , and puts them on the respective positions of the grid \mathbf{x}_l . Therefore the difference $u(\mathbf{x}_l, z_l) - \mathbf{S}_l \mathbf{u}$ is $u(\mathbf{x}_l, z_l)$ with zeros instead of the set determining samples $u(x_i, z_i)$, $(x_i, z_i) \in \mathbf{p}_l$. For L_2 norm, it can be written as:

$$\|u(\mathbf{x}_l, z_l) - \mathbf{S}_l \mathbf{u}\|_2^2 = \|u(\mathbf{x}_l, z_l)\|_2^2 - \|u(\mathbf{p}_l)\|_2^2. \quad (34)$$

As the norm of the coefficient vector \mathbf{a} is unity, the norm of the M -point uniformly sampled field $u(\mathbf{x}_l, z_l)$ on the line z_l is unity as well. Therefore, the norm of the difference $u(\mathbf{x}_l, z_l) - \mathbf{S}_l \mathbf{u}$ is always less than or equal to unity.

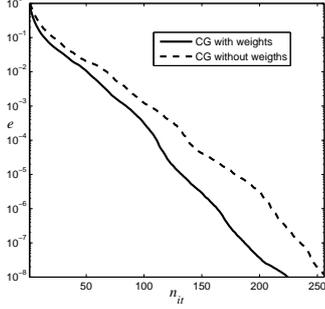


Figure 3. Performance of the CG-based reconstruction on data with (solid line) cluster with and without (dashed line) using weights for regularization.

Hence the iteration matrix $(\mathbf{I} - \mathbf{E}_{\chi_l}^{-1} \mathbf{H}_{z_l}^{-1} \mathbf{F} \mathbf{S}_l \mathbf{R})$ is non-expansive and the algorithm is convergent.

Eq. 34 suggest that if samples forming the set \mathbf{C}_l contain large portion of the energy of $u(\mathbf{x}_l, z_l)$ then the convergence is faster. Consider the case when $u(x, z)$ is a beam, with finite extent along x at the line $z = z_l$. If this beam falls into a place where the samples $u(\mathbf{p}_l)$ form a gap, then these samples will have small energy. Increasing their density outside the gap will not decrease the norm in Eq. 34 and therefore will not speed up the convergence. Thus, taking the maximal value of this norm in Eq. 33 over all possible field generators \mathbf{a} will produce a value close to 1 in the case when there are clusters or gaps in the samples $u(\mathbf{p}_l)$. On the other hand, if these samples are spread along the whole spatial extent T , then for any \mathbf{a} they will take significant part of the energy of $u(\mathbf{x}_l, z_l)$ resulting in lower norm in Eq. 33. Clusters along z are also undesired. Closely related points define closely spaced convex sets. A projection from a set to set will not change much the projected signal and will bring little new information.

The POCS performance highly depends on the structure of the given samples. Fully arbitrary sample positions determine large number of sets (up to s) while samples on few lines and on regular grids determine low number of sets and more efficient projections from set to set.

5. EXPERIMENTS

The theoretical analysis on the convergence of the described approaches is illustrated by three different experiments. The given samples for the experiments were generated by Eq. 4, where $M = 256$ non-zero field generating coefficients a_m were chosen as a Gaussian pulse centered at the origin. Assessment of the results is based on the normalized error between the original \mathbf{a} and reconstructed $\hat{\mathbf{a}}$ coefficient vectors - $e = \|\mathbf{a} - \hat{\mathbf{a}}\|_2 / \|\mathbf{a}\|$.

The goal of the first experiment is to verify the benefit of using adaptive weights to regularize the CG-based method for the case when there are clusters in the known samples. The number of the given samples is chosen to be $s = 1.5M$ such that the system in Eq. 5 is not underdetermined. M samples are chosen to be randomly scattered within a spatial rectangle, centered at the origin and

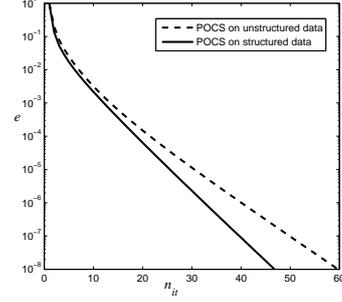


Figure 4. Performance of the POCS-based reconstruction method on data which defines structured (solid line) and unstructured (dashed line) convex sets.

lying in the half-plane $z > 0$, while the rest $0.5M$ samples are chosen to form a cluster. The CG-based reconstruction method is run with and without adaptive weights on the samples for regularization. The error convergence is shown in Figure 3 with solid and dashed lines respectively. The condition numbers of the weighted and unweighted matrices reflects the benefit of using the weights - $\kappa(\mathbf{R}^H \mathbf{W} \mathbf{R}) = 2826$, $\kappa(\mathbf{R}^H \mathbf{R}) = 8031$. A simple approach is used to measure the clustering of the samples and assign the adaptive weights in \mathbf{W} . The spatial rectangle where the known samples are situated is subdivided into rectangular cells with a coarse grid. The number of samples n_c inside each cell c is counted, and each sample which falls inside the cell c is weighted by $1/n_c$. While this method is not very precise, as it does not adapt to the cluster shapes, it is used here only to demonstrate the regularization power of the adaptive weights iterative approach.

The second experiment illustrates the rapid convergence of the POCS-based reconstruction method when structured convex sets can be formed from the known samples. Two sets of $s = 1.5M$ scattered light field samples were generated. The samples for the first set are chosen to be randomly scattered within a spatial rectangle, centered at the origin and lying in the half-plane $z > 0$. The samples for the second set are chosen inside the same spatial rectangle, but they are selected so that they form 4 structured convex sets. Each of these sets contains points which lie on the same distance z , and are irregularly scattered such that an M -point regular grid is able to host them. The error convergence of the POCS-based reconstruction method is shown in Figure 4 with a solid line for the structured data set and dashed line for the completely random data set.

In general, the CG-based reconstruction algorithm converges faster than the POCS-based algorithm. However, when the known samples can be used to form structured convex sets, the POCS-based algorithm uses much less computations. This can be illustrated by choosing irregularly scattered data points which form structured data sets, in the same manner as in the previous experiment. For such data, the error convergence rate in terms of number of

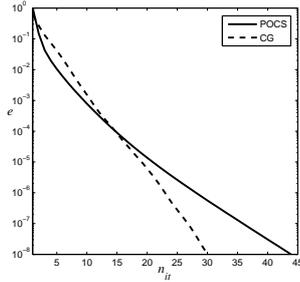


Figure 5. Performance in terms of number of iterations n_{it} for the POCS-based (solid line) and CG-based (dashed line) reconstruction methods on structured data.

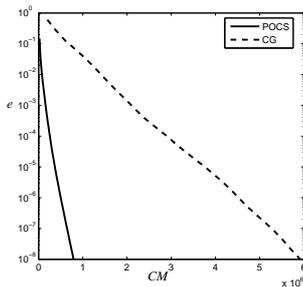


Figure 6. Performance in terms of complex multiplications (CM) for the POCS-based (solid line) and CG-based (dashed line) reconstruction methods on structured data.

iterations is similar for both reconstruction methods (Figure 5). However, one iteration of CG has computational complexity of $O(Ms)$, while the POCS-based algorithm has complexity of $O(M \log M)$ for the projection from one structured set to another, resulting in $O(LM \log M)$ per iteration for L sets. In this experiment the structured sets are $L = 4$, and the benefit of using POCS instead of CG can be seen in Figure 6.

6. CONCLUSION

In this work, we aimed at analyzing the convergence of the two preferred iterative schemes for diffraction field reconstruction from non-uniformly distributed samples. The convergence of the CG-based reconstruction method is highly dependent on the condition number of the non-uniform sample generating matrix. For clustered sample distributions, this matrix has a high condition number which can be reduced by adaptively weighting the samples. This method shows efficiency even when the weight selection method is rather rough. The POCS-based reconstruction method is generalized for the case when the known samples do not belong to a pre-defined regular grid. A detailed theoretical analysis shows that if structured convex sets can be defined from the samples distribution, the reconstruction method is rapidly convergent and efficient implementation exists. Numerical simulations show smaller computational costs than those required by the CG-based approach.

7. REFERENCES

- [1] G. B. Esmer, V. Uzunov, L. Onural, H. M. Ozaktas, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," *Signal Processing: Image Communication*, vol. 22, pp. 178–187, February 2007.
- [2] V. Uzunov, A. Gotchev, G. B. Esmer, L. Onural, and H. M. Ozaktas, "Non-uniform sampling and reconstruction of diffraction field," in *Proceedings of The 2006 SMMSP Workshop*, Florence, Italy, 2007, pp. 191–197.
- [3] E. Hecht, *Optics*, Addison Wesley, 1998.
- [4] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [5] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, John Wiley and Sons, Inc., 1991.
- [6] É. Lalor, "Conditions for the validity of the angular spectrum of plane waves," *J. Opt. Soc. Am.*, vol. 58, pp. 1235–1237, 1968.
- [7] G. C. Sherman, "Application of the convolution theorem to rayleigh's integral formulas," *J. Opt. Soc. Am.*, vol. 57, pp. 546–547, 1967.
- [8] K. Gröchenig and T. Strohmer, "Numerical and theoretical aspects of non-uniform sampling of band-limited images," in *Nonuniform Sampling: Theory and Practice*, F. Marvasti, Ed. Kluwer, 2001.
- [9] R. Aharoni and Y. Censor, "Block iterative projection methods for parallel computation of solutions to convex feasibility problems," *Linear Algebra Applications*, vol. 120, pp. 165–175, 1989.
- [10] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput Math Math Phys*, vol. 7, pp. 1–24, 1967.
- [11] Åke Björck, *Numerical Methods for Least Squares Problems*, SIAM, Amsterdam, Holland, 1990.
- [12] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, USA, 2003.
- [13] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," <http://www.cs.cmu.edu/quake-papers/painless-conjugate-gradient.pdf>.
- [14] R. Aharoni and Y. Censor, "Block iterative projection methods for parallel computation of solutions to convex feasibility problems," *Linear Algebra Applications*, vol. 120, pp. 165–175, 1989.
- [15] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput Math Math Phys*, vol. 7, pp. 1–24, 1967.

Publication P7

V. Uzunov, A. Gotchev, K. Egiazarian, “Regularized reconstruction of irregularly sampled scalar light fields,” in *Proceedings of Information Optics (WIO), 2010 9th Euro-American Workshop on*, 12-16 July 2010, pp. 1-3.

©2010 IEEE. Reprinted with permission.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Tampere University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

Regularized Reconstruction of Irregularly Sampled Scalar Light Fields

Vladislav Uzunov, Atanas Gotchev, Karen Egiazarian

Department of Signal Processing, Tampere University of Technology, P.O.Box 553, FIN-33101 Tampere, FINLAND

Email: vladislav.uzunov@tut.fi

Abstract—This paper addresses the problem of reconstruction of a monochromatic light field based on data points, irregularly distributed within a volume of interest. Such set up is considered to serve a wide area of applications related with three-dimensional display and beam shaping, where physically inconsistent input data is commonly available. Finite-dimensional models of scalar light fields are used to state the reconstruction problem as matrix inversion. Regularized inversion is done by the Tikhonov method, implemented by the iterative algorithm of conjugate gradients. The problem proves to be ill-posed, where the data points inconsistency is fully compensated by the regularized inversion, showing to be attractive for any application.

I. INTRODUCTION

The numerical reconstruction of a physically consistent light field within a volume of interest is important in many light field synthesis applications related with light beam shaping [1] and digital holography as three-dimensional (3D) imaging [2]. Primary tasks of digital holography are numerical reconstruction of digitally recorded holograms and computer generated holograms of synthetic 3D objects and scenes.

According to the Rayleigh-Sommerfeld diffraction integral, the light field on a certain plane is sufficient to compute the wave field at any point within the volume of interest [3]. In contrast, most works specify the volumetric structure of the 3D object on an ensemble of planes, at regular grid of points at each plane, often oversampled [4], [5], [1]. Such a 3D grid is still redundant in the sense that a light field can be uniquely determined from a single plane. Moreover, a fixed sampling grid often does not fit to data points distribution provided by the application at hand. Therefore, a rather general and dimensionally efficient set-up is to specify the volume of interest as ensemble of data points with free positions, often determined by the application at hand. In many cases, the data specifications are not guaranteed to be physically consistent with a light wave. For example, beam shaping and computer generated holography specify a 3D object or scene as a very complex structure of various shape and refraction properties, and thus is unlikely to be a physically consistent light wave. A common approach to treat the data inconsistencies are based on iterative serial or block projection techniques, derived from the Gerchberg-Saxton algorithm [6].

In our work we employ two previously developed finite-dimensional models of diffraction which represent a light field by a linear combination of generating functions[8], [7]. In this

manner the light field reconstruction is cast as an inversion problem. The inconsistencies of the specified data points can be modeled by an additive Gaussian term. Being simple, such a modeling is quite general as it encompasses a wide range of distortions coming from various applications. In many cases the data point distributions contain clusters and gaps which cause high condition number of the inversion matrices. We aim at treating the inconsistencies as inverse problem with a common approach such as the matrix-free Tikhonov regularization, based on conjugate gradient iterations [9]. We study the practical use of finite dimensional models and their performance depending on the input data specifications.

II. PROBLEM FORMULATION

This paper uses two discrete expansions to model a scalar monochromatic light field $u(x, z)$ within a free space volume of interest. For sake of simplicity, only one transversal spatial dimension x is used together with the longitudinal dimension z . The first model [7] discretizes the plane wave angular spectrum [3] to get the expansion:

$$u(x, z) = \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} a_m e^{j \frac{2\pi}{T} z \sqrt{\frac{T^2}{\lambda^2} - m^2}} e^{j \frac{2\pi}{T} m x}. \quad (1)$$

The coefficients a_m are the discretized of the angular spectrum $a(k_x)$ for the discrete spatial frequencies $k_x = \frac{2\pi m}{T}$. Their number M depends on the effective bandwidth of the described light field and the density, determined by the assumed spatial extent T of the field [7]. We refer to this expansion as to *Fourier Series (FS) -based model*.

The second discrete expansion represents the light field u in polar coordinate system (r, ϕ) with the correspondence $x = r \sin \phi$ and $z = r \cos \phi$. A detailed derivation [8] leads to an expansion of the field in terms of circular harmonics:

$$u(r, \phi) = \frac{4\pi^2}{\lambda} \sum_{m=-\lfloor \frac{M}{2} \rfloor}^{\lfloor \frac{M-1}{2} \rfloor} c_m e^{jm(\phi + \frac{\pi}{2})} J_m \left(\frac{2\pi}{\lambda} r \right), \quad (2)$$

where J_m is the m -th order Bessel function of the first kind. The number M of nonzero coefficients c_m is related to the spatial and angular extent of the field [8]. We refer to this expansion as to *Bessel Functions (BF) -based model*.

Naturally, an irregular grid can be specified as the set of s sampling points $\{(x_i, z_i)\}_{i=1}^s$, or, alternatively in polar

coordinates $\{(r_i, \phi_i)\}_{i=1}^s$. The field of interest is specified as sample values $u(x_i, z_i), i = 1, \dots, s$ at these points. Now the light field reconstruction problem can be stated as to find the unknown field-generating coefficients c_m or a_m , given the irregularly distributed samples $u(x_i, z_i), i = 1, \dots, s$. Eq. (2) and Eq. (1) can be written for each point (x_i, z_i) to obtain two alternative sets of equations, which can be expressed in a matrix form as:

$$\mathbf{u} = \mathbf{A}\mathbf{h}, \quad (3)$$

where $\mathbf{h} = \mathbf{a} = [a_{-\lfloor M/2 \rfloor}, a_{-\lfloor M/2 \rfloor + 1}, \dots, a_{\lfloor (M-1)/2 \rfloor}]^T$ or $\mathbf{h} = \mathbf{c} = [c_{-\lfloor M/2 \rfloor}, c_{-\lfloor M/2 \rfloor + 1}, \dots, c_{\lfloor (M-1)/2 \rfloor}]^T$ is the unknown vector of the field generating coefficients and the vector of given samples is $\mathbf{u} = [u(x_1, z_1), u(x_2, z_2), \dots, u(x_s, z_s)]^T$. \mathbf{A} is the reconstruction matrix which has two different forms depending on the discrete model. For the BF-based model \mathbf{A} is

$$\begin{aligned} \mathbf{A} &= \mathbf{J} = \{\mathcal{J}_{p,q}\} \\ &= \left\{ e^{j(q - \lfloor \frac{M}{2} \rfloor - 1)(\phi_p + \frac{\pi}{2})} J_{q - \lfloor \frac{M}{2} \rfloor - 1}(kr_p) \right\}, \\ & p = 1, \dots, s, q = 1, \dots, M \end{aligned} \quad (4)$$

and for the FS-based model \mathbf{A} is

$$\begin{aligned} \mathbf{A} &= \mathbf{R} = \{\mathcal{R}_{p,q}\} \\ &= \left\{ e^{j\frac{2\pi}{T}z_p \sqrt{\frac{T^2}{\lambda^2} - (q - \lfloor \frac{M}{2} \rfloor - 1)^2}} e^{j\frac{2\pi}{T}(q - \lfloor \frac{M}{2} \rfloor - 1)x_p} \right\}, \\ & p = 1, \dots, s, q = 1, \dots, M \end{aligned} \quad (5)$$

The straightforward approach to solve for the coefficient vector \mathbf{h} is to invert the matrix \mathbf{A} . Note that the structure of this matrix is determined only by the positions of the known samples. Some field specifications might use more samples to describe the fine details of a scene and less for the uniform regions. Such clustering of the samples causes large condition number of the matrix \mathbf{A} and inversion becomes numerically unstable. In addition, there might be noise and/or the scene samples can be inconsistent with the physical models of Eq. (2) and Eq. (1). As the inconsistencies might be of various origin and difficult to predict in the common case, it is reasonable to assume that their nature is also random as any eventual noise appeared upon the scene capture. These two factors can be accounted in a common random term ε :

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \varepsilon, \quad (6)$$

where now $\mathbf{y} = [y_1, \dots, y_s]^T$ is the vector of given data samples, \mathbf{h} is the unknown coefficient vector, $\mathbf{A} = \mathbf{J}$ or $\mathbf{A} = \mathbf{R}$ depending on the model chosen for reconstruction, and $\varepsilon = [\varepsilon_1, \dots, \varepsilon_s]^T$ is a random vector drawn from zero-mean normal distribution $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. Upon direct inversion, the high condition number of \mathbf{A} causes strong amplification of the random term and it dominates over the reconstructed coefficients \mathbf{h} .

The Tikhonov regularization approach finds a solution $\hat{\mathbf{h}}$ which minimizes the functional:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{g} \in \mathbb{R}^M} \left\{ \|\mathbf{A}\mathbf{g} - \mathbf{y}\|^2 + \delta \|\mathbf{g}\|^2 \right\}. \quad (7)$$

Tikhonov regularization balances between the contradictory requirements for small residual and small norm of the solution [9]. The penalty term δ is used to tune this balance. Optimal value of δ can be determined automatically by the Morozov's discrepancy principle. The minimum norm requirement ensures smoothness of the solution and thus robustness to noise, while minimizing the residual ensures that the solution is close to the target. A computationally attractive approach [9] is based on iterative matrix solvers. The minimizer $\hat{\mathbf{h}}$ of the Tikhonov functional in Eq. (7) is equivalent to the solution of the linear system [9]:

$$(\mathbf{A}^T \mathbf{A} + \delta \mathbf{I}) \hat{\mathbf{h}} = \mathbf{A}^T \mathbf{y} \quad (8)$$

As the matrix $\mathbf{H} = \mathbf{A}^T \mathbf{A} + \delta \mathbf{I}$ is symmetric and positive definite, it can be inverted with an iterative method. The iterations build the solution step by step, updating the solution vector each time until a desired accuracy is achieved. In many cases the iterations converge fast enough to ensure $\mathcal{O}(N^2)$ complexity. The *conjugate gradient (CG) method* is one of the most rapidly convergent and numerically stable algorithms [9]. The level of the regularization parameter δ can be automatically chosen e.g. according to Morozov's discrepancy principle [9].

III. EXPERIMENTS AND RESULTS

The data for all experiments is simulated by choosing the coefficients in the models of Eq. 1 and Eq. 2 and computing the field sample values at a set of irregularly distributed positions $\{(x_i, z_i)\}_{i=1}^s$. The sample positions are randomly distributed within the region of interest, where a part of them forms a cluster. The data inconsistency with the physical model is simulated by adding a Gaussian term to the calculated sample values. An inverse crime [9] is avoided by simulating the field values with the BF-based model of Eq. 2 and performing the reconstruction with the FS-based model of Eq. 1 in all experiments. The performance of the reconstruction is assessed by the Mean Squared Error (MSE) between the simulated and reconstructed field values, calculated on a regular grid.

Fig. 1 demonstrates the existence of an inverse problem in the presence of cluster. It shows the plots of the singular values of the matrices \mathbf{R} and \mathbf{J} from Eq. 5 and Eq. 4, respectively. The matrices are of size 384×256 which corresponds to reconstruction of 256 unknown coefficients from 384 data samples. Half of the sample positions are chosen to form a cluster of square shape and variable size (cf. Fig. 1). In the next experiment, the field samples are simulated at the same sets of positions by the BF-based model. The samples are simulated by choosing coefficients c_m for the model which would generate a continuous Gaussian beam inside the region of interest. The energy level of the inconsistency term added to the irregular samples is $\sigma = 5\%$ of the energy of the field, computed on the regular grid used to estimate the MSE. Fig. 2 shows the error convergence of the CG-based iterative reconstruction, done by the matrix \mathbf{R} . The value of the regularization parameter δ is chosen according to Morozov's discrepancy principle [9].

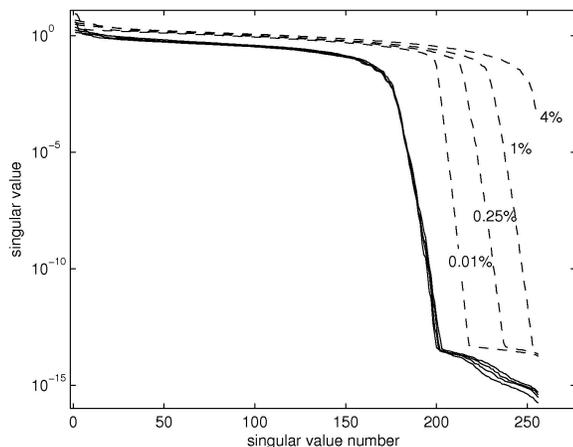


Fig. 1. Singular values of \mathbf{J} (solid lines) and \mathbf{R} (dashed lines) for different size of the cluster as percent of the region of interest.

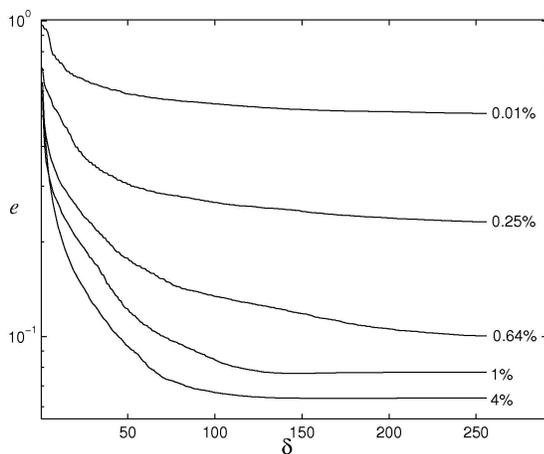


Fig. 2. Convergence of the CG-based Tikhonov regularized reconstruction for different sizes of the cluster (as percent of the region of interest). The samples are simulated with the BF-based model and $\sigma = 5\%$ inconsistency and the reconstruction is done by the FS-based model.

The last experiment shows the performance of the reconstruction for different noise levels σ . The field sample distribution is chosen such that half of the 384 samples form a cluster of 100 times smaller than the volume of interest. Again, the coefficients c_m for the BF-based model used in the simulations correspond to a continuous Gaussian beam inside the region of interest. Fig. 3 shows the reconstruction error as a function of the regularization parameter δ . The curves were built by varying δ in the regularized reconstruction. The optimal values of δ which give the best reconstruction for each noise level are very close to the ones obtained by the Morozov's discrepancy principle.

IV. CONCLUSION

In this paper we use two discrete models of diffraction to demonstrate the performance of regularized reconstruction of scalar light fields from irregularly distributed samples with

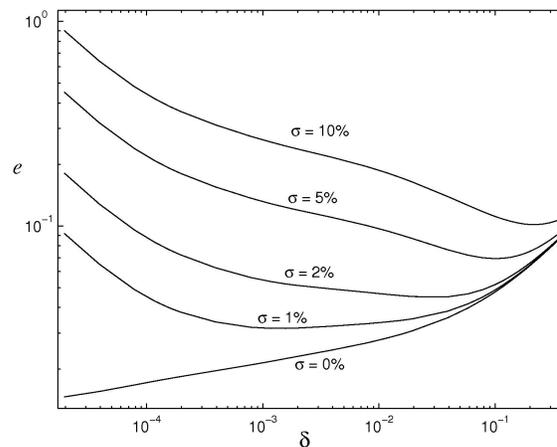


Fig. 3. Regularized reconstruction error e versus the regularization parameter δ for different noise levels σ . The samples form a square cluster 100 times smaller than the volume of interest.

inconsistency. The condition number of the reconstruction matrices grows with the level of clustering of the data points. The iterative reconstruction is able to compensate for the sample inconsistency only when their clustering is not very strong. The reconstruction converges rapidly also for larger cluster size. The regularization compensates the ill-posedness of the problem with reconstruction error close to the data point inconsistency.

REFERENCES

- [1] R. Piestun and J. Shamir, "Synthesis of three-dimensional light fields and applications", Proc. IEEE vol.90, pp. 222–244, 2002.
- [2] H. Ozaktas and L. Onural, eds., Three-Dimensional Television: Capture, Transmission, Display, Springer, 2008.
- [3] J. W. Goodman, Introduction to Fourier Optics Mc-Graw-Hill, New York, 1996.
- [4] G. Shabtay, "Three-dimensional beam forming and Ewald's surfaces", Opt. Commun., vol.226, pp. 33–37, 2003.
- [5] U. Levy, D. Mendlovic, Z. Zalevsky, G. Shabtay, and E. Marom, "Iterative algorithm for determining optimal beam profiles in a three-dimensional space", Appl. Opt., vol. 38, pp. 6732–6736, 1999.
- [6] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures", Optik, vol.35, pp. 237–246, 1972.
- [7] V. Uzunov, A. Gotchev, and K. Egiazarian, "Convergence and error analysis of diffraction field iterative non-uniform sampling schemes", in Proc. of the 2008 LNLA Workshop, Lausanne, Switzerland, 2008, p. 8p.
- [8] V. Uzunov, G. B. Esmer, A. Gotchev, L. Onural, and H. M. Ozaktas, "Bessel functions -based reconstruction of non-uniformly sampled diffraction fields", in Proc. of The 2007 3DTV Conference, Kos, Greece, 2007, pp. 1–4.
- [9] J. Kaipio and E. Somersalo, Statistical and Computational Inverse Problems, Springer, 2005.
- [10] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, Philadelphia, USA, 2003.

Publication P8

V. Uzunov, A. Gotchev, K. Egiazarian, “On the stability of reconstruction of irregularly sampled diffraction fields,” in *Advances in Optical Technologies*, vol. 2010, Article ID 138024, 2010, 12 pages.

Research Article

On the Stability of Reconstruction of Irregularly Sampled Diffraction Fields

Vladislav Uzunov, Atanas Gotchev, and Karen Egiazarian

Department of Signal Processing, Tampere University of Technology, P. O.Box 553, 33101 Tampere, Finland

Correspondence should be addressed to Vladislav Uzunov, vladislav.uzunov@tut.fi

Received 17 August 2010; Accepted 24 October 2010

Academic Editor: Leonid Yaroslavsky

Copyright © 2010 Vladislav Uzunov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper addresses the problem of reconstruction of a monochromatic light field from data points, irregularly distributed within a volume of interest. Such setting is relevant for a wide range of three-dimensional display and beam shaping applications, which deal with physically inconsistent data. Two finite-dimensional models of monochromatic light fields are used to state the reconstruction problem as regularized matrix inversion. The Tikhonov method, implemented by the iterative algorithm of conjugate gradients, is used for regularization. Estimates of the model dimensionality are related to the number of degrees of freedom of the light field as to show how to control the data redundancy. Experiments demonstrate that various data point distributions lead to ill-posedness and that regularized inversion is able to compensate for the data point inconsistencies with good numerical performance.

1. Introduction

Many optical applications require a generation and control of light fields. Digital processing by computers is an attractive way of implementing such operations as it overcomes possible physical limitations of analog devices. However, signal processing has to be suitably coupled with the otherwise naturally continuous optical signals. This coupling requires effective discrete representation of the continuous functions, associated with the light fields. From one point of view, the discrete representation of a light field should preserve the degrees of freedom of the continuous model which describes the physical properties of the field. From another point of view, the discrete representation should admit the requirements on the light field which are imposed by the application. Three-dimensional (3D) imaging deals with the reconstruction of captured optical signals by digital means as CCD arrays and recreation of synthetic or computer generated, 3D data by holographic means [1]. Light beam shaping requires the reconstruction or synthesis of light beams which maintain certain properties along their propagation [2].

This paper addresses the problem of monochromatic light field reconstruction from ensemble of data samples

with free positions, distributed within a volumetric region of interest. Setting the light field specification at a nonuniform irregular grid of points is rather general and can be utilized in a wide range of applications. For example, computer generated holography [3] reconstructs a light field to approximate a synthetic or a captured object, given its 3D abstract model—a point cloud, mesh, NURBS, and so forth [4]. The data provided by such 3D models can be hosted conveniently by an irregular grid. Limited aperture and resolution in digital holography are remedied by multiple CCD recordings in a single or different planes [5, 6]. These are prone to distortions originating from not precisely known CCD measurement locations which can be handled by an irregular grid of points. Reconstruction of a light field in terms of light field generators is required to drive properly a light modulation system for an eventual light field synthesis. For example, a color holographic 3D display system drives three spatial light modulators—one for each red, green and blue color channel [7].

Another advantage of specifying the application constraints as an ensemble of irregularly distributed samples is the control on the input data redundancy. According to the Rayleigh-Sommerfeld diffraction integral, the light field

on a certain plane is sufficient to compute the wave field at any point within the volume of interest [8]. Therefore, a regular grid of samples at a certain transversal plane should be sufficient to reconstruct the whole field. Gori [9] and Onural [10] have derived conditions for the sampling rate when the support of the field at any Fresnel plane is limited, considering the cases when the diffraction field originates from strictly band limited or strictly space limited function. In the general case of a function which is essentially limited both in space and frequency domain, the amount of needed samples should not be higher than the degrees of freedom N of the field, measured by the finite space-bandwidth product. In this case the conditions from [9, 10] require redundant amount of samples, as the spatial spread of the field extends upon propagation along the longitudinal direction z , while the frequency support remains the same. An adhoc approach as in [11] and a strict derivation as in [12, 13] use information about local bandwidth to reach the number of degrees of freedom by the amount of needed samples at any transversal plane along z . Early works [3, 14] use some of these results to specify the object on a plane, loosing the volumetric opportunity. Some recent works commonly specify the volume of interest on an ensemble of equally spaced planes, at a regular grid of points at each plane, often oversampled [2, 15–19]. However, such a 3D grid is still redundant with respect to the number of degrees of freedom of the diffraction field. Moreover, a fixed 3D grid might not fit well to an application-driven data point distribution. In this sense, an irregular grid of points not only fits any potential application but also opens up opportunities for volumetric specifications and data redundancy control.

In any application, the specifications for the light field are not guaranteed to be physically consistent with a light wave. Digital recording of holograms include CCD noise, misalignment, finite aperture and resolution. Light beam shaping and computer generated holography specify complex structures with properties, unlikely to be exhibited by a light wave. The mainstream of articles on light field synthesis treats the data specification inconsistencies by various iterative projection methods, derived on the base of the Gerchberg-Saxton algorithm [20]. Volumetric specifications usually define one constraint set per plane, assign weights and tolerances to the sets and use parallel or serial projections to reach an optimized solution [2, 6, 18, 19, 21, 22]. Other methods define the constraint sets on Ewald's surface [17] or fractional Fourier domain [23]. Alternative approaches to treat the data inconsistency are based on genetic algorithms with application to beam shaping [24, 25]. In our work we employ two finite-dimensional models for diffraction which represent a light field by a linear combination of generating functions [26, 27]. In these models, the inconsistencies of the specified data points can be represented by an additive Gaussian term. While being simple, such a representation is quite general and encompasses a wide range of distortions.

The light field reconstruction problem can be cast as an inverse problem [28] to find the coefficients of the finite-dimensional models from the specified data points [26, 27, 29]. The focus of this paper is on the characteristics of the input data specification which influence the stability of the

finite-dimensional models for light field reconstruction. The spectrum of each model is analyzed in various scenarios depending on the size of the volume of interest, clustering of the data points and amount of input data, as these are important for practical applications. As far as the mathematical approach is concerned, the Tikhonov regularization [28] is used with both models to provide an approximate solution which fits the specified data points. The singular value decomposition- (SVD-) based implementation of Tikhonov regularization is used to illustrate and analyze the limits of the reconstruction. The regularization parameters are determined by Morozov's discrepancy principle. The matrix-free Tikhonov regularization is suggested as practical approach for reconstruction. It iteratively finds a regularized solution based on the method of conjugate gradients (CGs) [30]. The convergence rate and final reconstruction error for various scenarios are presented to assess this approach.

The paper is organized as follows. Section 2 introduces the basics of diffraction and notations related to the considered light fields. Section 3 presents the finite-dimensional models used to represent the field, and discusses their properties with focus on dimensionality issues. Section 4 states the observation model in a matrix form, based on the irregularly distributed inconsistent data samples. Regularized solutions used for approximating the input data are presented in Section 5. Section 6 illustrates the performance of the suggested reconstruction methods in a number of guided experiments.

2. Basics of Diffraction

Consider a light field $u(x, y, z)$, generated by a monochromatic light wave, which propagates in a linear, isotropic and homogeneous media. Under such conditions, the spatial distribution of the complex amplitude u satisfies the homogeneous Helmholtz wave equation [8]:

$$\nabla^2 u + k^2 u = 0, \quad (1)$$

where $k = 2\pi/\lambda$ is the wave number of the monochromatic light. Such waves, emerging from an optical system, satisfy the Sommerfeld radiation condition and are accurately described by the Rayleigh-Sommerfeld diffraction integral [8]. It relates the light field $u(x, y, z)$ at any point (x, y, z) to that on a "reference" plane at $z = 0$ in a linear and shift-invariant relationship. However, the practical use of this relationship for computations is limited as it requires very high sampling rate [31]. Therefore, frequency domain alternatives of the Rayleigh-Sommerfeld diffraction integral are preferred [2, 22, 31].

For the sake of simplicity, the discussion is restricted to one transverse dimension (x) only, resulting in a two-dimensional (2D) scalar function $u(x, z)$ describing the light field. Note that the interesting dimension for the scope the paper is the longitudinal one (z) as it keeps the volumetric properties of the light field. Generalizations to the three dimensional case are straightforward and mentioned

whenever necessary. The function $u(x, z)$ can be represented by its 2D Fourier transform $A(k_x, k_z)$:

$$u(x, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z, \quad (2)$$

where k_x and k_z are the spatial frequencies along x and z directions, respectively. Applying the Fourier transform on (1) and using the derivative property of the Fourier transform, one ends up with the relation $A(k_x, k_z)(k^2 - k_x^2 - k_z^2) = 0$. Hence, $A(k_x, k_z)$ can be nonzero only on the circle with radius $k = 2\pi/\lambda$:

$$k_x^2 + k_z^2 = k^2 = \left(\frac{2\pi}{\lambda}\right)^2. \quad (3)$$

3. Finite-Dimensional Models of Diffraction

3.1. Bessel-Fourier (B-F) Generators. As the spatial frequencies k_x and k_z are limited only to a circle, the Fourier transform $A(k_x, k_z)$ is in fact one-dimensional function:

$$A(k_x, k_z) = A(k \sin \theta, k \cos \theta) \equiv C(\theta), \quad (4)$$

where the frequencies are expressed in polar form as $k_x = k \sin \theta$, $k_z = k \cos \theta$ with $k = 2\pi/\lambda$ and $\theta \in [0, 2\pi)$. Upon a change of the spatial coordinates x and z in polar form as $x = r \sin \phi$ and $z = r \cos \phi$, the Fourier transform in (2) becomes

$$u(r, \phi) = \frac{2\pi}{\lambda} \int_0^{2\pi} C(\theta) e^{j(2\pi/\lambda)r \cos(\theta-\phi)} d\theta. \quad (5)$$

The function $C(\theta)$ is 2π -periodic. Consequently, it can be described by the complex Fourier series as $C(\theta) = \sum_m c_m e^{jm\theta}$. Inserting this into (5) and changing the order of summation and integration, the field can be expressed as

$$u(r, \phi) = \frac{4\pi}{\lambda} \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} c_m e^{jm(\phi+\pi/2)} J_m\left(\frac{2\pi}{\lambda} r\right), \quad (6)$$

where $J_m(t)$ is the m th order Bessel function of the first kind [32] which solves the remaining integral under the summation. Equation (6) assumes that the Fourier series representation of $C(\theta)$ has been truncated to M nonzero coefficients $c_m \neq 0$, $m = -\lfloor M/2 \rfloor, \dots, \lfloor (M-1)/2 \rfloor$ to arrive at finite representation. Detailed derivation of the model can be found in [26]. This model describes the field continuously at any spatial point as a superposition of a discrete set of basis functions $\psi_m(r, \phi) = e^{jm(\phi+\pi/2)} J_m(kr)$, which are mutually orthogonal and separable. Hence, the discrete set of scaling coefficients c_m describes completely the continuous field. Yet, no explicit discretization was done during the derivation.

The model of (6) can be generalized to three dimensions [33]. In 3D, the Fourier transform $A(k_x, k_y, k_z)$ is supported on a sphere with the same radius $k = 2\pi/\lambda$ and therefore it is two-dimensional. Such a function can be expressed as a superposition of spherical harmonics, instead of the cylindrical harmonics $e^{jm(\phi+\pi/2)} J_m(kr)$. In radial direction the Bessel functions $J_m(kr)$ are substituted by the spherical Bessel functions of the first kind.

3.2. Fourier Generators. Another frequency domain representation of the diffraction field $u(x, z)$ is based on a decomposition in plane waves $e^{j(z\sqrt{k^2-k_x^2}+xk_x)}$:

$$u(x, z) = \int_{-2\pi/\lambda}^{2\pi/\lambda} a(k_x) e^{j\sqrt{k^2-k_x^2}z} e^{jk_x x} dk_x, \quad (7)$$

where $a(k_x)$ is the 1D Fourier transform of the field $u(x, 0)$ restricted only to the initial line at $z = 0$. The derivation of this integral is based on solving the Helmholtz equation for the 1D Fourier transform $a_z(k_x)$ of the field $u(x, z)$ restricted to a line at a distance z , given $a(k_x)$ as initial condition. The plane waves $e^{j(z\sqrt{k^2-k_x^2}+xk_x)}$ propagate only in positive direction $z \geq 0$ as the frequencies $k_z = \sqrt{k^2 - k_x^2}$ assume only positive sign. This corresponds to taking only half of the circle $k_z^2 + k_x^2 = k^2$ for $k_z \geq 0$.

As k_x is limited within $[-2\pi/\lambda, 2\pi/\lambda]$, the function $u(x, 0)$ is band limited. It can be further assumed to be essentially space limited or, more precisely, to have a finite space-bandwidth product. Such a function can be periodized with period equal to the transversal extent T of a spatial region of interest. This is equivalent to discretization of $a(k_x)$. A periodic and band limited function is equivalently represented by a finite number (M) of discrete spectral components at values of the frequency $k_x = 2\pi m/T$, $m = -\lfloor M/2 \rfloor, \dots, \lfloor (M-1)/2 \rfloor$. Using these values of k_x in (7) leads to a finite-dimensional model of diffraction:

$$u(x, z) = \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} a_m e^{j(2\pi/T)z\sqrt{(T^2/\lambda^2)-m^2}} e^{j(2\pi/T)mx}, \quad (8)$$

where $a_m = a(2\pi m/T)$ are the coefficients of the Fourier series expansion of $u(x, 0)$. This discrete, Fourier generators-based, model describes the field continuously at any spatial point as a superposition of the basis functions $\varphi_m(x, z) = e^{j(2\pi/T)z\sqrt{(T^2/\lambda^2)-m^2}} e^{j(2\pi/T)mx}$.

The generalization of the model in (8) is straightforward. Another transversal dimension y brings an extra term in the plane waves $e^{j(z\sqrt{k^2-k_x^2}+xk_x+yk_y)}$ [8]. Now the 2D function $u(x, y, 0)$ must be assumed space limited also along y , as discretization of k_y corresponds to periodization along y .

3.3. Dimensionality of the Models and Field Characteristics. The derived models involve a finite number of M basis functions in a linear combination to build up a light field. The amount of the nonzero weighting coefficients determine the dimensionality of any forward or inverse problem related with computation or reconstruction of a light field. Therefore, it is important to relate this amount to the physical properties of the approximating field. The specified region of interest and target detail level can be naturally related to the spatial and frequency content of the field, that is, to the number of degrees of freedom of the field. As the field can be completely described by the function on the initial line $u(x, 0)$, its number of degrees of freedom is equal to the degrees of freedom of $u(x, 0)$. The number of degrees of freedom N is a set of N numbers which describe it completely. In terms of the Wigner distribution, this is the

area in the time-frequency plane under which the Wigner distribution of the function is essentially nonzero [34]. This can be measured by the space-bandwidth product between the spatial Δx and frequency Δk_x extent:

$$N = \frac{\Delta x \Delta k_x}{2\pi}, \quad (9)$$

where the factor 2π renormalizes the angular frequency k_x to an ordinary one. The number N can be related to the number of independent Gabor atoms with unit space-bandwidth product in the Gabor representation of $u(x, 0)$. Alternatively, if $u(x, 0)$ is sampled uniformly along x according to the Nyquist rate $2\pi/\Delta k_x$, then N becomes the number of essentially nonzero samples from which the continuous function $u(x, 0)$ can be reconstructed. In practice, N can often be estimated from the application at hand. For example, it can be related to the physical properties of a sensing device. For synthetic data and other applications, reasonable assumptions about the desired detail level of the target can be used.

3.3.1. Fourier Generators. The amount of required coefficients M_F for this model is directly related with the bandwidth of $u(x, 0)$:

$$\frac{2\pi M_F}{T} = \Delta k_x = \frac{2\pi N}{\Delta x}, \quad (10)$$

hence $M_F = \left\lceil N \frac{T}{2\pi \Delta x} \right\rceil$.

Equation (10) suggests that M_F is directly proportional to N . This result is expected as the Fourier generators discretize the frequency band of $u(x, 0)$. The proportionality coefficient $T/\Delta x$ defines the excess which needs to be taken. The period T must be larger than Δx so that the periodic replicas of $u(x, 0)$ do not overlap. However, choosing T close to Δx does not guarantee that an overlap will not occur on a line at further distance z . Discrete frequency axis k_x defines discrete Fourier spectrum of $u(x, z)$ at any line z , keeping $u(x, z)$ periodic with the same period T . On the other hand, spatially limited pattern $u(x, 0)$ tends to spread its transversal support when propagated along z . Therefore, given maximal distance z_{\max} according to the specified region of interest, one must ensure that $T > \Delta_{z_{\max}} x$, where $\Delta_{z_{\max}} x$ is the support of $u(x, z_{\max})$. The relative increase of this spatial support $\Delta_{z_{\max}} x/\Delta x$ is smaller (closer to 1) for larger Δx at fixed distance z_{\max} .

3.3.2. Bessel-Fourier Generators. The coefficients c_m in the Bessel-Fourier model are the Fourier series coefficients of $C(\theta)$. Their amount M_B can be determined from the frequency support of $C(\theta)$. Direct comparison of (2) and (7) yields

$$A(k_x, k_z) = a(k_x) \delta\left(k_z - \sqrt{k^2 - k_x^2}\right), \quad (11)$$

$$C(\theta) = A(k \sin \theta, k \cos \theta) = a(k \sin \theta) \delta(0).$$

Hence, the frequency support of $a(k \sin \theta)$ coincides with the frequency support of $C(\theta)$ and can be used to estimate the number M_B .

The duality property of the Fourier transform $FT\{a(k_x)\} = u(-2\pi x, 0)$ can be used to obtain the frequency support of $a(k_x)$ as $\Delta x/2\pi$. However, the frequency support of $a(k \sin \theta)$ with respect to θ is not the same, even though it is related to $\Delta x/2\pi$. The highest frequency in the Fourier transform of $a(k_x)$ is $\Delta x/4\pi$ —the same as the frequency of the harmonic $\cos((\Delta x/2)k_x)$. Therefore, the frequency support of $\cos((\Delta x/2)k \sin \theta) = \cos((\pi \Delta x/\lambda) \sin \theta)$ can be used to estimate the frequency support of $a(k \sin \theta)$. In communication theory, such a harmonic function is recognized as a special case of frequency modulated signal $\cos(2\pi f_c \theta + (\Delta f/f_m) \sin(2\pi f_m \theta))$. Its frequency support is approximated as $2(f_m + \Delta f)$ according to the Carson's rule [35]. The frequency support of $\cos((\pi \Delta x/\lambda) \sin \theta)$ and $a(k \sin \theta)$ and $C(\theta)$, respectively, is estimated as $2((1/2\pi) + (\Delta x/2\lambda)) \approx \Delta x/\lambda$. Hence, the dimensionality of the Bessel-Fourier model required to describe a field with transversal support Δx on the reference line is:

$$M_B = \left\lceil \frac{\Delta x}{\lambda} \right\rceil = \left\lceil N \frac{2\pi}{\Delta k_x \lambda} \right\rceil. \quad (12)$$

This result suggests that the Bessel-Fourier model requires small amount of generators when the spatial support of the field is comparable to the wavelength. From another point of view, if N is assumed to be fixed, then the excess in M_B compared to N is determined by the ratio between $2\pi/\Delta k_x$ and λ . In fact, $2\pi/\Delta k_x$ is the size of the finest detail structure in $u(x, 0)$.

4. Irregularly Sampled Diffraction Fields

An irregular grid is specified as the set of s sampling points $\{(x_i, z_i)\}_{i=1}^s$, or, in polar coordinates $\{(r_i, \phi_i)\}_{i=1}^s$, with the correspondence $x_i = r_i \sin \phi_i$ and $z_i = r_i \cos \phi_i$. The sampled field is specified as the sample values $u(x_i, z_i)$, $i = 1, \dots, s$. The light field reconstruction problem is to find the unknown field-generating coefficients c_m or a_m , given the irregularly distributed samples $u(x_i, z_i)$, $i = 1, \dots, s$. Equations (6) and (8) can be written for each point in the irregular sampling set to obtain

$$u(r_i, \phi_i) = \frac{4\pi^2}{\lambda} \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} c_m e^{jm(\phi_i + (\pi/2))} J_m\left(\frac{2\pi}{\lambda} r_i\right), \quad (13)$$

$$u(x_i, z_i) = \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} a_m e^{j(2\pi/T)\sqrt{(T^2/\lambda^2) - m^2} z_i} e^{j(2\pi/T)mx_i},$$

for $i = 1, \dots, s$. Each of these sets of equations forms a linear system for the M unknown coefficients c_m or a_m and can be expressed in a matrix form:

$$\mathbf{u} = \mathbf{A}\mathbf{h}, \quad (14)$$

where $\mathbf{h} = \mathbf{a} = [a_{-\lfloor M/2 \rfloor}, a_{-\lfloor M/2 \rfloor+1}, \dots, a_{\lfloor (M-1)/2 \rfloor}]^T$ or $\mathbf{h} = \mathbf{c} = [c_{-\lfloor M/2 \rfloor}, c_{-\lfloor M/2 \rfloor+1}, \dots, c_{\lfloor (M-1)/2 \rfloor}]^T$ is the unknown

vector of the field generating coefficients and the vector of given samples is $\mathbf{u} = [u(x_1, z_1), u(x_2, z_2), \dots, u(x_s, z_s)]^T$. \mathbf{A} is the reconstruction matrix which has two different forms depending on the discrete model. For the Bessel-Fourier model \mathbf{A} is

$$\begin{aligned} \mathbf{A} = \mathbf{J} &= \left\{ \mathcal{J}_{p,q} \right\} \\ &= \left\{ e^{j(q - \lfloor M/2 \rfloor - 1)(\phi_p + \pi/2)} J_{q - \lfloor M/2 \rfloor - 1}(kr_p) \right\}, \quad (15) \\ p &= 1, \dots, s, \quad q = 1, \dots, M \end{aligned}$$

and for the Fourier model \mathbf{A} is

$$\begin{aligned} \mathbf{A} = \mathbf{R} &= \left\{ \mathcal{R}_{p,q} \right\} \\ &= \left\{ e^{j(2\pi/T)\sqrt{(T^2/\lambda^2) - (q - \lfloor M/2 \rfloor - 1)^2} z_p} e^{j(2\pi/T)(q - \lfloor M/2 \rfloor - 1)x_p} \right\}, \\ p &= 1, \dots, s, \quad q = 1, \dots, M. \quad (16) \end{aligned}$$

The straightforward approach to solve for the coefficient vector \mathbf{h} is to invert the matrix \mathbf{A} . Note that the structure of this matrix is determined only by the positions of the known samples. Some field specifications might use more samples to describe the fine details of a scene and less for the uniform regions. Such clustering of the samples causes large condition number of the matrix \mathbf{A} and inversion becomes numerically unstable. In addition, there might be noise and/or the scene samples can be inconsistent with the physical models of (6) and (8). As the inconsistencies might be of various origin and difficult to predict in the common case, it is reasonable to assume that their nature is also random as any eventual noise appeared upon the scene capture. These two factors can be accounted in a common random term ε :

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \varepsilon, \quad (17)$$

where now $\mathbf{y} = [y_1, \dots, y_s]^T$ is the vector of given data samples, \mathbf{h} is the unknown coefficient vector, $\mathbf{A} = \mathbf{J}$ or $\mathbf{A} = \mathbf{R}$ depending on the model chosen for reconstruction, and $\varepsilon = [\varepsilon_1, \dots, \varepsilon_s]^T$ is a random vector drawn from zero-mean normal distribution $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. Upon direct inversion, the high condition number of \mathbf{A} causes strong amplification of the random term and it dominates over the reconstructed coefficients \mathbf{h} .

5. Regularized Reconstruction

A linear system can be solved by taking the pseudoinverse with the help of SVD [36]. SVD decomposes an $s \times M$ matrix as $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ with \mathbf{U} and \mathbf{V} containing the left and right singular vectors of \mathbf{A} , respectively. $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$, $n = \min(s, M)$, is an $s \times M$ diagonal matrix with the value ordered singular values of \mathbf{A} . The Moore-Penrose pseudoinverse \mathbf{A}^+ can be used to find a solution $\hat{\mathbf{h}}$, provided the desired data vector \mathbf{y} as $\hat{\mathbf{h}} = \mathbf{A}^+\mathbf{y} = \mathbf{V}\mathbf{D}^+\mathbf{U}^T\mathbf{y}$, where $\mathbf{D}^+ = \text{diag}\{d_1^{-1}, d_2^{-1}, \dots, d_n^{-1}\}$ [36].

If the SVD pseudoinverse is applied to the left side of the noisy observation vector $\mathbf{y} = \mathbf{A}\mathbf{h} + \varepsilon$, the noise norm in the resultant reconstruction will be boosted by a factor of d_n^{-1} . In the case when the condition number of \mathbf{A} is large, d_n^{-1} is large as well, and the noise will be dominant in the reconstruction. A common remedy is to apply Tikhonov regularization, that is, to introduce a penalty term δ into the inversion matrix $\mathbf{D}_\delta^+ = \text{diag}\{d_1/(d_1^2 + \delta), \dots, d_n/(d_n^2 + \delta)\}$ [28]. This is equivalent to solving the minimization problem

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{g} \in \mathbb{R}^M} \left\{ \|\mathbf{A}\mathbf{g} - \mathbf{y}\|^2 + \delta \|\mathbf{g}\|^2 \right\}, \quad (18)$$

that is, to minimizing the norms of both residual and solution. Tikhonov regularization balances between the contradictory requirements for small residual and small norm of the solution by the penalty term δ [28]. The minimum norm requirement ensures smoothness of the solution and thus robustness to noise, while minimizing the residual ensures that the solution is close to the target. Optimal value of δ can be determined automatically by the Morozov's discrepancy principle [28].

The SVD form of the Tikhonov regularized solution is very convenient for analysis purposes. However, finding the SVD of a matrix is of cubic complexity [36]. A more computationally attractive approach is based on iterative matrix solvers [28]. The minimizer $\hat{\mathbf{h}}$ of the Tikhonov functional in (18) is equivalent to the solution of the linear system [28]:

$$(\mathbf{A}^T\mathbf{A} + \delta\mathbf{I})\hat{\mathbf{h}} = \mathbf{A}^T\mathbf{y} \quad (19)$$

The matrix $\mathbf{H} = \mathbf{A}^T\mathbf{A} + \delta\mathbf{I}$ is symmetric and positive definite. Therefore, it can be inverted with an iterative method, which builds the solution step by step, updating the solution vector each time until a desired accuracy is achieved [30]. In many cases the iterations converge fast enough to ensure lower complexity than the SVD-based Tikhonov solution. The conjugate gradient method is one of the most rapidly convergent and numerically stable algorithms [30]. It iterates as follows:

- (1) initialize $\mathbf{b} = \mathbf{A}^T\mathbf{y}$, $\hat{\mathbf{h}}$ arbitrary, $\mathbf{r}_0 = \mathbf{b} - \mathbf{H}\hat{\mathbf{h}}$ and $\mathbf{d}_0 = \mathbf{r}_0$.
- (2) For $i = 1$ to i_{\max}
 - (a) $\mathbf{w}_i = \mathbf{H}\mathbf{d}_i$
 - (b) $\beta_i = \|\mathbf{r}_i\| / (\mathbf{d}_i^T \mathbf{w}_i)$
 - (c) $\hat{\mathbf{h}} = \hat{\mathbf{h}} + \beta_i \mathbf{d}_i$
 - (d) $\mathbf{r}_{i+1} = \mathbf{r}_i - \beta_i \mathbf{w}_i$
 - (e) $\gamma_i = \|\mathbf{r}_{i+1}\|^2 / \|\mathbf{r}_i\|^2$
 - (f) $\mathbf{d}_{i+1} = \mathbf{r}_{i+1} + \gamma_i \mathbf{d}_i$
- (3) End.

CG method updates the solution at each iteration with a small portion β_i along the search direction \mathbf{d}_i (step (2c)). The search directions are conjugate to each other, which guarantees convergence in at most M iterations. Practically,

the algorithm converges to a predefined accuracy in much less iterations $i_{\max} \ll M$. The most computationally expensive operation at each iteration is the matrix-vector product at step (2a) which requires M^2 operations. The total complexity of the algorithm is $\mathcal{O}(i_{\max}M^2)$ —lower than $\mathcal{O}(M^3)$ required by the SVD-based Tikhonov solution.

6. Experiments and Results

The light field reconstruction approaches presented in this paper are evaluated by controlled experiments on simulated data. The experiments are described in two subsections. The first subsection describes the details of input data simulation and the criterion for reconstruction performance evaluation, common to all experiments. The second subsection presents experimental results and discusses the benefits of the regularized reconstruction in various scenarios.

6.1. Data Simulation and Performance Evaluation. All experiments are carried out on synthetic data. The data is simulated using the observation model of (17), where the matrix \mathbf{A} is computed according to (15) for the B-F model or according to (16) for the Fourier model. The matrices relate M generating coefficients with s data samples, pseudorandomly distributed within a volume of interest.

The dimensionality M of the reconstruction matrix must be the same for both models so that the reconstruction results are comparable. However, this dimensionality depends on different factors for the different models (cf. Section 3.3). For both models, $M = 256$ nonzero coefficients are able to produce a Gaussian beam inside a square region of interest with size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$. The size of the region of interest is selected so that the adjacent replicas of the field produced by the Fourier generators do not overlap at distance z_{\max} comparable to T . Figure 1 depicts the magnitude of the fields produced by the B-F generators and the Fourier generators by showing the values on a uniform 256×256 grid which spans the whole region of interest.

The samples of the field used for its reconstruction are simulated by (17) on s randomly selected positions within the volume of interest. The amount of points s is chosen to be greater than the amount of nonzero coefficients M . The underdetermined case $s < M$ is not considered interesting as the minimal norm solution might diverge substantially from the true underlying coefficient vector [27, 37]. In some experiments, part of the specified data points are selected to form clusters with size much smaller than the region of interest. The samples are randomly distributed according to uniform distribution both inside and outside of the clusters. The size and position of the clusters and the amount of points inside and outside the clusters define different sample density in different subdomains of the region of interest. Varying these parameters helps to investigate the effect of variable sample density within the light field support. The influence of these parameters on the ill-posedness of the reconstruction problem is examined during the first group of experiments. Based on the results, representative scenarios are selected to test the reconstruction approaches.

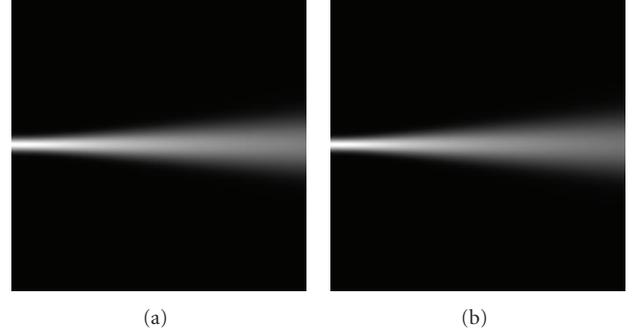


FIGURE 1: A diffraction field representing a Gaussian beam computed by the B-F generators (a) and by the Fourier generators (b) on an uniform 256×256 grid covering a region of interest with size $0.1 \text{ mm} \times 0.1 \text{ mm}$.

In a fair experiment, the reconstruction performance must be compared with the inconsistency of the input data. Hence, the inconsistency level σ and the reconstruction error e must be measured according to the same quantity. Often one model is used to simulate the samples and another to reconstruct the light field, so that an inverse crime does not take place when experimenting on simulated data [28]. The models are characterized by sets of generating coefficients which cannot be compared directly. Instead, the sets of coefficients are used to compute the reconstructed light field within the whole region of interest on a uniform, rectangular grid $\Gamma = \{l\xi, n\zeta\}_{l,n=0}^M$, with $\xi = T/M$ and $\zeta = z_{\max}/M$ being the sampling steps. The error is measured on this uniform grid:

$$e = \frac{\|\mathbf{F}_0 - \mathbf{F}\|}{\|\mathbf{F}_0\|}. \quad (20)$$

Here, $\mathbf{F}_0 = u_0(\Gamma)$ and $\mathbf{F} = u(\Gamma)$ are vectors formed by the field values computed on the grid Γ out of the simulation and reconstruction models, respectively. An inconsistency level σ , comparable to this error, can be selected as a percentage of the energy of the field \mathbf{F}_0 .

6.2. Experiments. The light field reconstruction problem becomes ill-posed and requires regularization when the spectrum of the reconstruction matrix (\mathbf{R} or \mathbf{J}) is widely spread and contains many small values. The structure of the matrices depends only on the specified data point distribution. A first group of experiments investigates the spectra of these matrices for different data point distribution scenarios. A second group of experiments illustrates the need and benefit of regularization even for some cases of well-distributed and physically consistent data points. Finally, the last group of experiments shows the benefits and limitations of the regularization for data distributions which make the reconstruction problem ill-posed. This group includes also an experiment which demonstrates a potential practical application of the considered reconstruction approach.

6.2.1. Matrix Singular Values. This group of experiments examines the spectra of \mathbf{R} and \mathbf{J} for various volume of interest

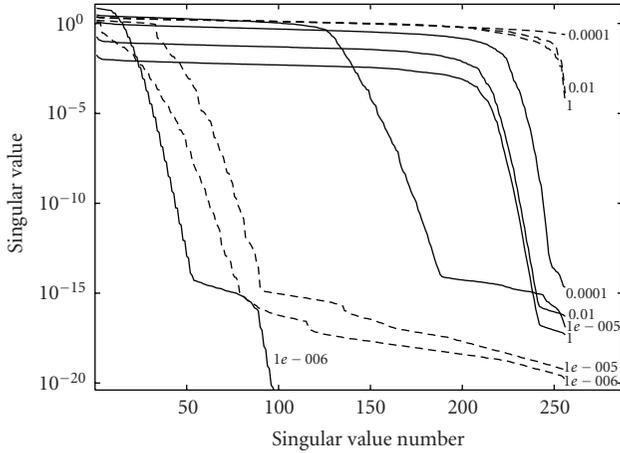


FIGURE 2: Singular values of \mathbf{J} (solid lines) and \mathbf{R} (dashed lines) for different sizes of the region of interest in [m] (number next to each line) and $s = 384$ given data points which do not form a cluster.

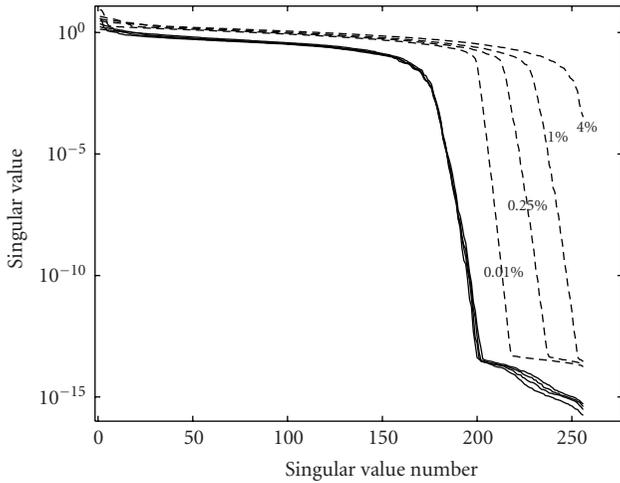


FIGURE 3: Singular values of \mathbf{J} (solid lines) and \mathbf{R} (dashed lines) for different sizes of the cluster as percent of the region of interest.

sizes and specified data point densities. Spatial variations of the density are simulated by forming data point clusters of various size, shape and position.

The plots in Figure 2 represent the singular values of \mathbf{J} and \mathbf{R} for varying size of the region of interest. It has square shape and the amount of irregularly distributed points is selected as $s = 1.5M = 384$ so that the matrices are overdetermined. No clusters are formed in this experiment. The vertical axis in Figure 2 represents the singular value d_i , while the abscissa represents its index $i = 1, 2, \dots, n$ after the singular values have been sorted in a descending order, that is, $d_1 \geq d_2 \geq \dots \geq d_n$. Both \mathbf{J} and \mathbf{R} show their most compact spectra when the dimensionality $M = 256$ fits the size of the region of interest. Driving away from this matching size, the spectrum of both \mathbf{J} and \mathbf{R} spreads more. The matrix \mathbf{R} has well-concentrated singular values, while \mathbf{J} is still ill-conditioned for the matching size. Hence, in the absence of clusters, \mathbf{R} can still be used for reconstruction

without regularization, as the next group of experiments will demonstrate. Irregularly distributed samples in a square region of interest do not have wide angular diversity and the Bessel-Fourier generators do not bring enough information to recover the coefficients from the samples. This causes wide-spread singular values of \mathbf{J} . It is possible to demonstrate that the spectra of \mathbf{J} become more compact if the volume of interest has wider size along x than along z .

Next experiments investigate how the presence of clusters influence the spectra of \mathbf{J} and \mathbf{R} . The basic experiment varies the size of a single square cluster, positioned in the center of a square region of interest of size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$. The experiments use $s = 384$ points, half of which are distributed inside the cluster and the other half outside. The singular value plots in Figure 3 show that the matrix \mathbf{R} becomes ill-conditioned similarly to \mathbf{J} . A smaller cluster spreads the singular values more for both matrices. However, the spectra of \mathbf{J} are less sensitive to the changes of the size of the cluster, while the spectra of \mathbf{R} vary considerably. This is another illustration of the fact that the Bessel-Fourier generators produce a well-conditioned reconstruction matrix when the angular diversity of the samples is higher. On the other hand, this also shows that the Fourier model depends more on the transversal and longitudinal diversity of the sample positions, as the cluster size controls the sample density distribution. Some additional experiments vary the position of a fixed-size cluster within the region of interest and the amount of clusters. The amount of clusters does not influence significantly the spectra of both matrices, provided that the sample density within the clusters is kept constant. However, large amount of clusters spread the general distribution of samples, which improves the spectra of \mathbf{R} . The spectra of \mathbf{J} improve significantly when the cluster position goes closer to the origin, thus providing samples with higher angular diversity. The spectra of \mathbf{R} remain invariant to the position of the cluster within the region of interest.

The last experiments of this group investigate how the amount of given points s influences the singular values of \mathbf{J} and \mathbf{R} . The simulated region of interest has size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$ and contains a single cluster of 100 times smaller size, located in the middle. One of the experiments varies the amount of points outside the cluster from 64 to 1024, while the amount of points within the cluster is kept fixed to 192. The other experiment varies the amount of points inside the cluster from 64 to 1024, while the amount of the samples outside the cluster is kept fixed to 192. The minimum total amount of points s is at least 256 in both experiments, so that the matrices are not underdetermined. The singular value plots of \mathbf{J} and \mathbf{R} show significant influence when the points outside the cluster are varied (Figure 4), while they remain almost invariant to the amount of points within the cluster. More points outside the cluster carry more information for both models and the singular values of the reconstruction matrices become more compact. The Bessel-Fourier generators benefit from new points outside the cluster as they carry new angles, unlike new points within the cluster. The Fourier generators benefit from more points outside the cluster as this increases the overall point density,

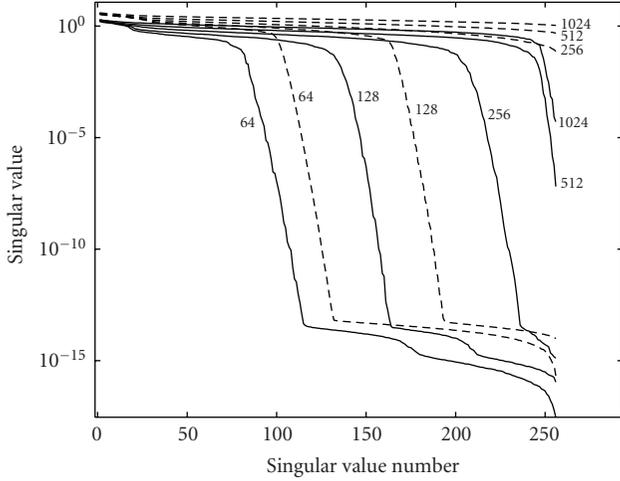


FIGURE 4: Singular values of \mathbf{J} (solid lines) and \mathbf{R} (dashed lines) for different amounts (shown next to each line) of given samples outside a cluster of size 100 times smaller than the region of interest.

TABLE 1: Nonregularized reconstruction errors for different noise levels.

σ , [%]	e , Fourier model	e , B-F model
0	$3.19 \cdot 10^{-13}$	$1.88 \cdot 10^{-5}$
1	0.015	$4.56 \cdot 10^7$
2	0.027	$9.12 \cdot 10^7$
5	0.068	$2.28 \cdot 10^8$
10	0.13	$4.56 \cdot 10^8$

while more points within the cluster keep the overall density the same. However, this benefit vanishes after the overall density goes above a certain level, which is theoretically sufficient to reconstruct any field for the chosen volume of interest and dimensionality M .

6.2.2. Regularization for Unclustered Input Data Points.

This group of experiments aims at demonstrating that an ill-conditioned reconstruction matrix needs regularization for reconstruction even for well-distributed and physically consistent input data. The benefits and limitations of the Tikhonov regularization method are illustrated for different levels of inconsistency of the input data. The data was simulated inside a square region of interest of size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$ with the B-F model. The model uses $M = 256$ nonzero coefficients. The amount of simulated points is $s = 384$ so that the matrices \mathbf{R} and \mathbf{J} are overdetermined. The points do not form a cluster so that \mathbf{R} has compact spectrum and \mathbf{J} does not (Figure 2). At first, nonregularized reconstructions were done for different levels of inconsistency with errors shown in Table 1. The reconstruction based on the Fourier model shows practically zero error for physically consistent data ($\sigma = 0\%$), while the reconstruction with B-F generators shows some very small, yet not zero error. For inconsistent data, the Fourier generators are able to reconstruct a field with an error comparable to the inconsistency level. At the same time the

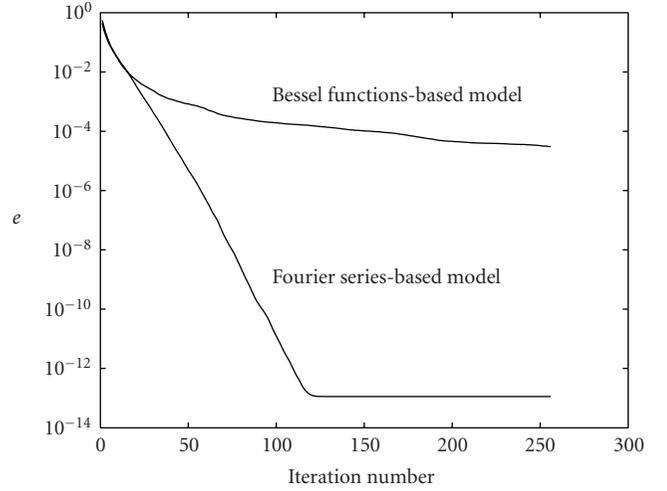


FIGURE 5: Convergence of the CG algorithm for reconstruction from nonclustered and physically consistent field samples, simulated by the B-F generators. The CG algorithm was applied to the nonregularized matrices \mathbf{J} and \mathbf{R} .

B-F generators model is unable to reconstruct the field even for very small inconsistency ($\sigma = 1\%$). This is due to the random term amplification by the reciprocal of the small singular values of \mathbf{J} , as discussed in Section 5. In practice, the reconstruction from consistent data is done by applying an iterative technique based on the CG method on the reconstruction matrices \mathbf{R} and \mathbf{J} . The convergence of CG depends on the condition number d_1/d_n of the matrix [30]. Hence, CG converges fast for \mathbf{R} whose condition number is small and very slow for \mathbf{J} which has high condition number (Figure 5). A regularization strategy is needed when the reconstruction matrix is ill-conditioned. Tikhonov regularization changes the singular values of a matrix by the parameter δ (cf. Section 5). Any change in the spectra of the reconstruction matrix impairs the accuracy of an eventual reconstruction. However, the small singular values are increased and the inconsistency random term ε is not amplified much. An optimal value of δ corresponding to a minimal error is found by Morozov's principle. Such value of δ optimizes the condition number of $\mathbf{J}^T \mathbf{J} + \delta \mathbf{I}$ so that the CG-based Tikhonov regularization converges faster with a minimal final error (Figure 6). The final error is very close to the inconsistency level σ . Thus, the regularization of \mathbf{J} helps to achieve very similar convergence and error as reconstructing with the well-conditioned \mathbf{R} .

6.2.3. Benefits and Limitations of Regularization for Clustered Data Points.

The last group of experiments aim at illustrating the benefits and the limitations of the regularized field reconstruction when the specified data point distribution contains clusters. In such scenarios, both matrices \mathbf{R} and \mathbf{J} have widely spread spectra. In this context, it is interesting to investigate how the factors affecting the spectra of \mathbf{R} and \mathbf{J} influence the regularized reconstruction performance. The factors which have the greatest influence on the spectrum are the size of the cluster and the amount of points outside the

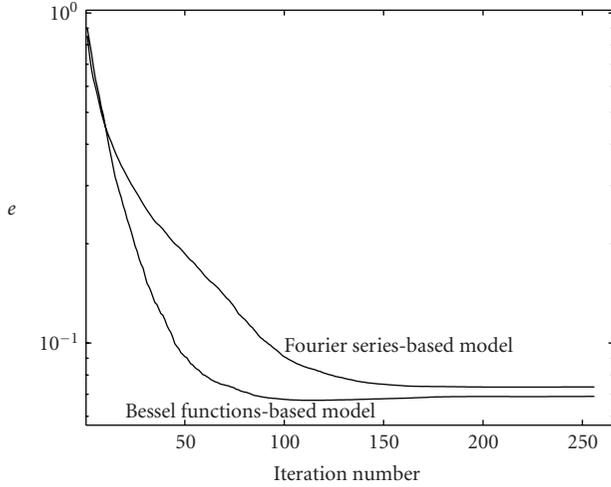


FIGURE 6: Convergence of the CG algorithm for reconstruction from nonclustered field samples, generated by the B-F model. The inconsistency level is 5%. The CG algorithm was applied to the regularized \mathbf{J} and unregularized \mathbf{R} .

cluster. The performance is assessed by the reconstruction error and the convergence rate of the CG-based Tikhonov regularization.

The first experiment of this group investigates the performance of the Tikhonov regularized reconstruction for different levels of inconsistency σ of the input data. The B-F generators are used to simulate the data points inside a square region of interest of size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$. The total amount of points is $s = 384$ and half of them form a cluster of 100 times smaller size than the region of interest. The cluster is located in the center of the region of interest. The field was reconstructed from the known samples by the SVD-based Tikhonov regularization on the matrix \mathbf{R} for different values of the regularization parameter δ . The plots of the reconstruction error e versus δ are shown in Figure 7. The plots show that the Tikhonov regularization is able to compensate completely for the clusterization of inconsistent input data. If the samples are physically consistent, only an approximate solution can be achieved with sufficiently low error rate. Figure 8 shows the convergence rate of the CG-based Tikhonov regularized reconstruction when the level of inconsistency of the input data is 5%. The value of the regularization parameter δ was computed by Morozov's principle. The B-F model leads to slightly faster convergence rate than the Fourier model, demonstrating the existence of an inverse crime. Comparison with Figure 6 shows that both models have similar performance compared to the case of nonclustered data. In this case, the Tikhonov regularization is able to compensate completely for the ill-posedness caused by data clustering.

Next, it is interesting to test if decreasing the size of the cluster further affects the performance of the regularized reconstruction as it affects the spectra of the matrices \mathbf{R} and \mathbf{J} (Figure 3). In this experiment, the data inconsistency level is fixed to $\sigma = 5\%$ and the size of the cluster is varied. All other parameters of the scenario are the same as in the previous

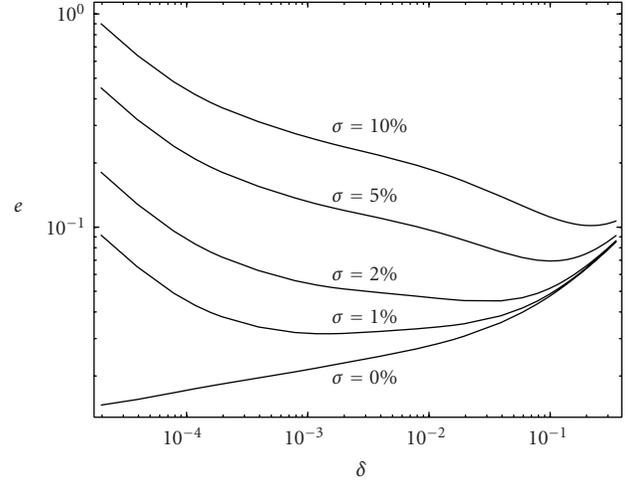


FIGURE 7: Tikhonov regularization reconstruction error e versus the regularization parameter δ for different noise levels σ . The samples are simulated with the B-F generators and the field is reconstructed by the Fourier generators. The samples form a square cluster 100 times smaller than the volume of interest.

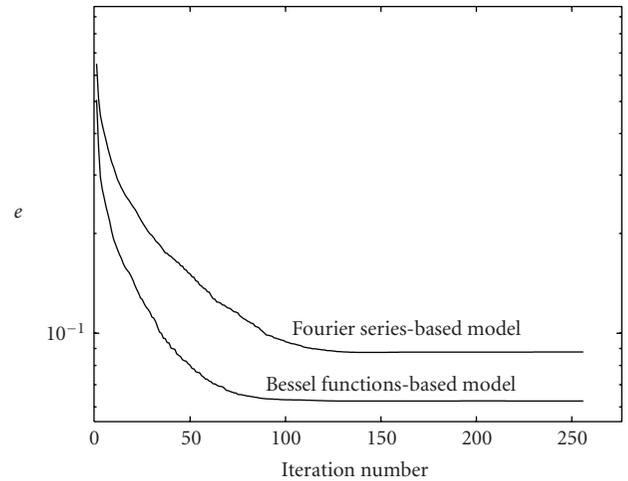


FIGURE 8: Convergence of the CG-based Tikhonov regularized reconstruction from samples generated by the B-F model which form a square cluster of size 100 less than the region of interest. The inconsistency level of the samples is 5%.

experiment. Figure 9 shows the convergence of the CG-based Tikhonov regularized reconstruction involving the Fourier generators. The optimal regularization parameter δ for each curve is obtained by Morozov's discrepancy principle. The optimal δ value is smaller for smaller cluster sizes, and hence the condition number of the reconstruction matrix $\mathbf{A}^T \mathbf{A} + \delta \mathbf{I}$ increases and the convergence of the CG algorithm becomes slower.

Another factor which influences the spectrum of the reconstruction matrices \mathbf{J} and \mathbf{R} is the amount of given data samples outside the cluster (Figure 4). It is interesting to check the influence of this amount on the reconstruction error and convergence of the CG-based reconstruction. The

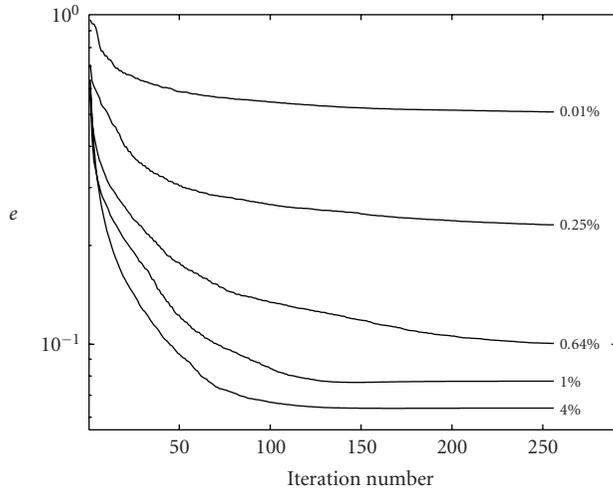


FIGURE 9: Convergence of the CG-based Tikhonov regularized reconstruction for different sizes of the cluster (as percent of the region of interest). The samples are simulated with the B-F model and $\sigma = 5\%$ inconsistency and the reconstruction is done by the Fourier generators.

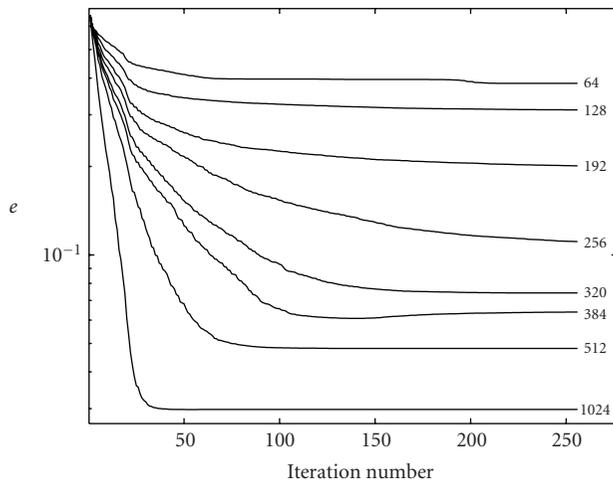


FIGURE 10: Convergence of the CG-based Tikhonov regularized reconstruction with the Fourier generators for different amounts of data points outside a cluster (number next to each line). The samples are simulated with the B-F generators and $\sigma = 5\%$ inconsistency.

B-F generators are used to simulate the data points inside a square region of interest of size $T \times z_{\max} = 0.1 \text{ mm} \times 0.1 \text{ mm}$ with level of inconsistency $\sigma = 5\%$. 192 points form a cluster in the middle of the region of interest with size 400 times smaller than this region. The amount of points outside the cluster is varied from 64 to 1024 so that the total amount of points is always at least 256 and \mathbf{J} and \mathbf{R} are not underdetermined. Figure 10 shows the convergence plots of the CG-based reconstruction done with the Fourier model. The lower reconstruction error clearly indicates the benefit of having more data points outside the cluster, showing lower minimal error rate. The optimal value

of the regularization parameter δ increases together with the amount of points. This decreases the condition number of the regularized reconstruction matrix $\mathbf{A}^T \mathbf{A} + \delta \mathbf{I}$ and speeds up the convergence rate.

The last experiment investigates a sample distribution scenario which occurs for high-resolution light field reconstruction from multiple low-resolution and limited-aperture CCD recordings. Such a scenario is investigated to demonstrate the practical applicability of the proposed reconstruction approach. In addition, this experiment reconstructs a light field which originates from a realistic object as a more sophisticated distribution. The reconstruction is done again with one transverse dimension only for the sake of comparability with the other experiments. The object at the initial line is 256-sample vertical strip from the “Lena” image, which is used as a standard test material in image processing. The bandwidth of this object contains 256 nonzero frequency components, which coincide with the coefficients a_m of the Fourier generators model. The light emerging from such an object is propagated within a region of size $T \times z_{\max} = 0.2 \text{ mm} \times 0.2 \text{ mm}$ which makes the effective bandwidth of the image $\Delta k_x = 256(2\pi/2 \cdot 10^{-4}) \approx 8042.5 \text{ rad/mm}$. The CCD sensors are considered to have twice lower frequency resolution, that is, sampling step of $2(2\pi/\Delta k_x) \approx 0.0016 \text{ mm}$ and size 0.1 mm—twice smaller than the transversal extent T . This setting results in 64 samples per CCD. Samples of the light field are captured by 6 different CCD recordings obtained at 6 different CCD positions on a transversal line, such that the oversampling factor is 1.5. The positions of the CCDs are selected such that the total amount of samples is distributed denser around the origin and sparser towards the side of the line, as shown in Figure 11. In this manner, the sampling approximates a density distribution, optimal in the sense of [11]. The higher sample density at the center forms a cluster which spreads the singular values of the reconstruction matrix \mathbf{R} , as evident from Figure 11. Complex values at the CCDs sample positions are simulated from the coefficients a_m of the Fourier generators model. In practice, complex values can be obtained by, for example, temporal phase shifting for each CCD position. The CCD noise is simulated by adding an inconsistency with level $\sigma = 5\%$ to the simulated CCD sample values. The regularized, CG-based reconstruction shows rapid convergence to an error comparable to the inconsistency level, as illustrated in Figure 12.

7. Conclusion

This paper has addressed the problem of monochromatic light field reconstruction from irregularly distributed samples with physically inconsistent values. The proposed reconstruction method is based on finite-dimensional modeling of the problem and regularized inversion. Our approach encompasses a wide range of applications in the area of 3D display and beam shaping. The dimensionality of both models can be directly related to the number of degrees of freedom of the field with a certain excess. The excess depends on the ratio between the transversal extent and volume of interest for the Fourier model and on the ratio between the finest detail level and wavelength for the B-F

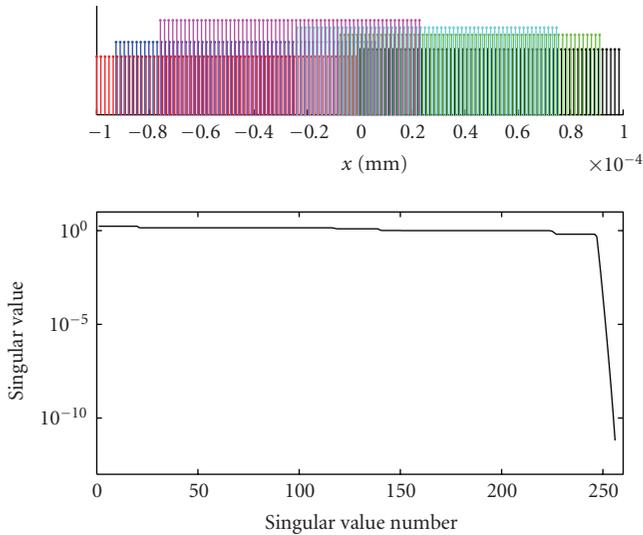


FIGURE 11: Samples from CCD recordings at 6 different CCD positions (shown in different colors and heights) on a transversal line (top) spread the singular values of \mathbf{R} (bottom).

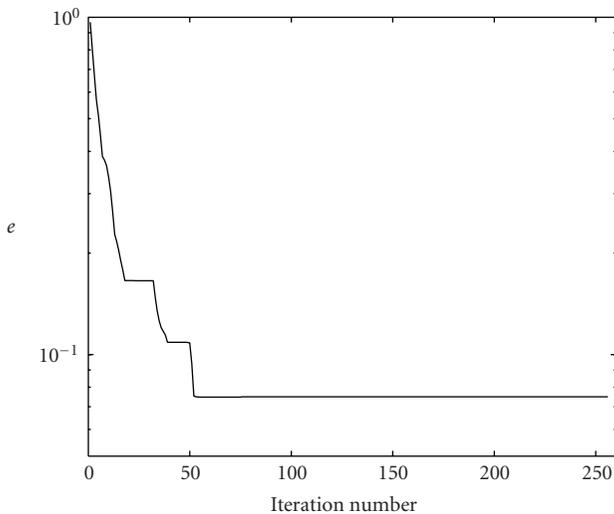


FIGURE 12: Convergence of the CG-based Tikhonov regularized reconstruction of the “Lena” strip light field from samples which imitate data multiple CCD recordings. The samples are simulated with the Fourier generators and $\sigma = 5\%$ inconsistency.

model. Uniform sample density defines the reconstruction problem as well posed when described by the Fourier model and as ill-posed when described by the B-F model. For such density, satisfactory reconstruction can be done with amount of points close to the number of degrees of freedom of the field. However, spatial variations of the density define the reconstruction problem as ill-posed for both models. Regularized reconstruction is able to compensate for the ill-posedness when the sample density variations are not very diverse within the region of interest. The reconstruction error and convergence of the iterative solver are very similar to the well-posed case. Large variations of the sample density

increase the reconstruction error above the level of inconsistency of the input sample data. In this sense, increasing redundancy in the amount of samples away from the number of degrees of freedom of the field brings clear benefit.

Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments and recommendations which helped to improve the manuscript. This work was supported by the Academy of Finland (project No. 213462, Finnish Programme for Centres of Excellence in Research 2006–2011 and project No. 137012 “High-Resolution Digital Holography: A Modern Signal Processing Approach”).

References

- [1] H. M. Ozaktas and L. Onural, Eds., *Three-Dimensional Television: Capture, Transmission, Display*, Springer, Berlin, Germany, 2008.
- [2] R. Piestun and J. Shamir, “Synthesis of three-dimensional light fields and applications,” *Proceedings of the IEEE*, vol. 90, no. 2, pp. 222–244, 2002.
- [3] L. Yaroslavsky and M. Eden, *Fundamentals of Digital Optics*, Birkhäuser, Boston, Mass, USA, 1996.
- [4] O. Schreier, P. Kauff, and T. Sikora, *3D Video Communication*, John Wiley & Sons, New York, NY, USA, 2005.
- [5] T. Kreis and K. Schlüter, “Resolution enhancement by aperture synthesis in digital holography,” *Optical Engineering*, vol. 46, no. 5, Article ID 055803, 2007.
- [6] P. Almero, G. Pedrini, and W. Osten, “Complete wavefront reconstruction using sequential intensity measurements of a volume speckle field,” *Applied Optics*, vol. 45, no. 34, pp. 8596–8605, 2006.
- [7] F. Yaraş, H. Kang, and L. Onural, “Real-time phase-only color holographic video display system using LED illumination,” *Applied Optics*, vol. 48, no. 34, pp. H48–H53, 2009.
- [8] J. W. Goodman, *Introduction to Fourier Optics*, Mc-Graw-Hill, New York, NY, USA, 1996.
- [9] F. Gori, “Fresnel transform and sampling theorem,” *Optics Communications*, vol. 39, no. 5, pp. 293–297, 1981.
- [10] L. Onural, “Sampling of the diffraction field,” *Applied Optics*, vol. 39, no. 32, pp. 5929–5935, 2000.
- [11] A. VanderLugt, “Optimum sampling of Fresnel transforms,” *Applied Optics*, vol. 29, pp. 3352–3361, 1990.
- [12] A. Stern and B. Javidi, “Sampling in the light of Wigner distribution,” *Journal of the Optical Society of America A*, vol. 21, no. 3, pp. 360–366, 2004.
- [13] A. Stern and B. Javidi, “General sampling theorem and application in digital holography,” in *Optical Information Systems II*, vol. 5557 of *Proceedings of SPIE*, pp. 110–123, August 2004.
- [14] T. Dresel, M. Beyerlein, and J. Schwider, “Design and fabrication of computer-generated beam-shaping holograms,” *Applied Optics*, vol. 35, no. 23, pp. 4615–4621, 1996.
- [15] B. Salik, J. Rosen, and A. Yariv, “One-dimensional beam shaping,” *Journal of the Optical Society of America A*, vol. 12, pp. 1702–1706, 1995.
- [16] G. Shabtay, Z. Zalevsky, U. Levy, and D. Mendlovic, “Optimal synthesis of three-dimensional complex amplitude distributions,” *Optics Letters*, vol. 25, no. 6, pp. 363–365, 2000.

- [17] G. Shabtay, "Three-dimensional beam forming and Ewald's surfaces," *Optics Communications*, vol. 226, no. 1–6, pp. 33–37, 2003.
- [18] U. Levy, D. Mendlovic, Z. Zalevsky, G. Shabtay, and E. Marom, "Iterative algorithm for determining optimal beam profiles in a three-dimensional space," *Applied Optics*, vol. 38, no. 32, pp. 6732–6736, 1999.
- [19] R. Piestun, B. Spektor, and J. Shamir, "Wave fields in three dimensions: analysis and synthesis," *Journal of the Optical Society of America A*, vol. 13, no. 9, pp. 1837–1848, 1996.
- [20] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
- [21] H. Stark, Y. Yang, and D. Gurkan, "Factors affecting convergence in the design of diffractive optics by iterative vector-space methods," *Journal of the Optical Society of America A*, vol. 16, no. 1, pp. 149–159, 1999.
- [22] R. Piestun, J. Shamir, B. Weßkamp, and O. Bryngdahl, "On-axis computer-generated holograms for three-dimensional display," *Optics Letters*, vol. 22, no. 12, pp. 922–924, 1997.
- [23] Z. Zalevsky, D. Mendlovic, and R. G. Dorsch, "Gerchberg-Saxton algorithm applied in the fractional Fourier or the Fresnel domain," *Optics Letters*, vol. 21, no. 12, pp. 842–844, 1996.
- [24] P. Yang, Y. Liu, W. Yang et al., "Adaptive laser beam shaping technique based on a genetic algorithm," *Chinese Optics Letters*, vol. 5, no. 9, pp. 497–500, 2007.
- [25] G. Zhou, X. Yuan, P. Dowd, Y. L. Lam, and Y. C. Chan, "Design of diffractive phase elements for beam shaping: hybrid approach," *Journal of the Optical Society of America A*, vol. 18, no. 4, pp. 791–800, 2001.
- [26] V. Uzunov, G. B. Esmer, A. Gotchev, L. Onural, and H. M. Ozaktas, "Bessel functions—based reconstruction of non-uniformly sampled diffraction fields," in *Proceedings of the 1st International Conference on 3DTV (3DTV-CON '07)*, pp. 1–4, Kos, Greece, May 2007.
- [27] V. Uzunov, A. Gotchev, G. B. Esmer, L. Onural, and H. M. Ozaktas, "Non-uniform sampling and reconstruction of diffraction field," in *Proceedings of the International Workshop on Spectral Methods and Multirate Signal Processing (SMMSP '06)*, pp. 191–197, Florence, Italy, September 2006.
- [28] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, Berlin, Germany, 2005.
- [29] V. Uzunov, A. Gotchev, and K. Egiazarian, "Convergence and error analysis of diffraction field iterative non-uniform sampling schemes," in *Proceedings of the International Workshop on Local and Non-Local Approximation in Image Processing (LNLA '08)*, p. 8, Lausanne, Switzerland, 2008.
- [30] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, Pa, USA, 2003.
- [31] F. Shen and A. Wang, "Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula," *Applied Optics*, vol. 45, no. 6, pp. 1102–1110, 2006.
- [32] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, NY, USA, 1972.
- [33] N. J. Moore and M. A. Alonso, "Bases for the description of monochromatic, strongly focused, scalar fields," *Journal of the Optical Society of America A*, vol. 26, no. 7, pp. 1754–1761, 2009.
- [34] A. W. Lohmann, R. G. Dorsch, D. Mendlovic, Z. Zalevsky, and C. Ferreira, "Space-bandwidth product of optical signals and systems," *Journal of the Optical Society of America A*, vol. 13, no. 3, pp. 470–473, 1996.
- [35] A. B. Carlson, P. B. Crilly, and J. C. Rutledge, *Communication Systems*, McGraw-Hill, Chicago, Ill, USA, 4th edition, 2001.
- [36] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Amsterdam, The Netherlands, 1990.
- [37] G. B. Esmer, V. Uzunov, L. Onural, H. M. Ozaktas, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," *Signal Processing*, vol. 22, no. 2, pp. 178–187, 2007.