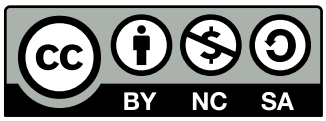
Aalto University
School of Electrical
Engineering

Master's programme in Electronics and Nanotechnology

Neural Dispersive Hologram for Computational Submillimeter-wave Imaging

Sihan Shao

Master's Thesis
2024



Aalto University
School of Electrical
Engineering

Author Sihan Shao

Title Neural Dispersive Hologram for Computational Submillimeter-wave Imaging

Degree programme Electronics and Nanotechnology

Major Microwave Engineering

Supervisor Prof. Zachary Taylor

Advisor M.Sc. (Tech.) Samu-Ville Pälli and D.Sc. (Tech.) Aleksi Tamminen

Date 11 July 2024 **Number of pages** 72 **Language** English

Abstract

Conventional submillimeter-wave imaging systems struggle to achieve scalability, real-time performance, and low complexity simultaneously. Microwave computational imaging offers a solution by using frequency diversity to image scenes through various measurement modes, shifting complexity from hardware to software. Typically, these systems first design the element that creates diverse illumination and then tune reconstruction algorithm parameters for good image quality.

In contrast to the sequential design approach, this thesis proposes a joint optimization method for a frequency-diverse phase hologram and the reconstruction neural network parameters in submillimeter-wave imaging. The goal is to optimize the hologram pattern for enhanced frequency diversity and its corresponding neural network for improved image quality.

First, a literature review of computational imaging systems from optics to millimeter-wave is presented. Additionally, differentiable imaging theory and recent work in the joint optimization of computational imaging systems are studied. A PyTorch-based Fourier-optics simulation codebase is developed to model the imaging physics process, mapping the reflectivity of objects to the reflected frequency response. This codebase analyzes the impact of joint optimization on efficiency, frequency diversity, and reconstruction quality, resulting in two optimal holograms with minimal unit sizes of 1 mm and 2 mm for manufacturing.

To validate the performance of the proposed computational imaging system with holograms, a quasi-optics setup operating at 220-330 GHz is built. While few similarities were observed between simulated and measured field patterns, scanning the fields near the actual hologram and the corresponding comparisons suggest that the designed holograms work as expected. The measured frequency diversity of the designed holograms exceeded that of both the simulations and the previous design. The comparison of measured and simulated frequency responses, along with imaging experiments, indicates partial validity of the current imaging physics model. Future work should focus on developing a more accurate model that accounts for non-ideal factors to better match the measured and simulated frequency responses.

Keywords submillimeter wave, computational imaging, differentiable design, neural network, phase hologram

Preface

The research for this thesis was done at the Department of Electronics and Nanoengineering in Aalto University School of Electrical Engineering as part of the MilliScan project.

First, my warmest thanks go to my supervisor, Prof. Zachary Taylor, for giving me the opportunity to work on this interesting and challenging project, and for his knowledge and thoughtful discussions throughout this thesis. Secondly, I am genuinely grateful to my advisors, Ph.D. student Samu-Ville Pälli and D.Sc. (Tech) Aleksi Tamminen, for their guidance and for sharing their knowledge throughout this thesis. I would also like to extend my gratitude to the THz research group, where I have enjoyed the work and the atmosphere ever since. Special thanks also go to Antti Kuhlberg for his assistance with the hologram manufacturing for this thesis.

I would also want to thank the people at Aalto University for their encouraging and interesting teaching and support. Finally, I want to thank my family for their endless love and support.

Otaniemi, 11 July 2024

Sihan Shao

Contents

Symbols and abbreviations

Symbols

α	learning rate of neural network
β	learning rate of hologram
β_0	average wavenumber of the medium
c	speed of light in vacuum $\approx 3 \times 10^8$ [m/s]
δ	signed value difference
δ_h	loss tangent of the hologram material
$\epsilon_{r,h}$	relative permittivity of the hologram material
$f(\cdot)$	forward model of imaging system
f_x	spatial frequency of x portion
f_y	spatial frequency of y portion
$G(\cdot)$	computational algorithms
$G(\cdot,\cdot)$	scalar Green function between two positions
$G_{\text{learn}}(\cdot)$	learning-based computational algorithms
g	indirect measurement
g_l	Gumbel noise for each category l
h	height map of the hologram
$\hat{h}$	updated height map of the hologram
H	transfer function of angular spectrum method
$\mathbf{H}$	measurement matrix
$\mathbf{J}$	regularization term
k_0	free-space wave number
λ	wavelength
$\mathcal{L}$	loss function of neural network
M	number of measurement frequencies (modes)
n	refractive index
n_δ	deviation from the average refractive index
$\nabla \times \mathbf{A}$	curl of vectorin $\mathbf{A}$
N	number of scene pixels
P	plenoptics function
q	quantization operator of hologram
Q	predetermined quantization set of hologram
r_f	radius from the beam's central axis
ρ	correlation coefficient
σ	sigmoid function
σ_i	i singular values
t_k	stepsize
Θ	all possible parameters of neural network
$\hat{\Theta}$	updated parameters of neural network
ϕ	phase map of the hologram
$\hat{\phi}$	updated phase map of the hologram

Π_Q	quantization projection operation of hologram
$\mathbf{r}$	spatial position of field
$\mathbf{r}_\mathrm{A}$	position at the aperture plane
$\mathbf{r}_\mathrm{f}$	position at the feed plane
$\mathbf{r}_\mathrm{R}$	position at the receiver antenna
$\mathbf{r}_\mathrm{S}$	position at the imaging scene
$\mathbf{r}_\mathrm{T}$	position at the transmitting antenna
$\mathbf{\Phi}$	angle of incident field
U_1	quadratic approximation of off-axis parabolic Mirror
$\mathbf{U}$	unitary matrix of SVD
$\mathbf{V}$	unitary matrix of SVD
Σ	diagonal matrix containing singular values of SVD
$\mathbf{T}_\mathrm{A}$	impulse response function at aperture plane
$\mathbf{E}$	electric field strength
$\mathbf{E}_0$	incident electric field from aperture plane
$\mathbf{E}_\mathrm{A}$	electric field at aperture plane
$\mathbf{E}_\mathrm{feed}$	electric field at the feed plane
$\mathbf{E}_\mathrm{hol}$	electric field modulated by the hologram
$\mathbf{E}_\mathrm{in}$	electric field incident on the hologram
$\mathbf{E}_\mathrm{s}$	scattered electric field between aperture and scene
$\mathbf{E}_\mathrm{T}$	total electric field between aperture and scene
x	cartesian coordinate
x_i	i input variable in computational graph
y	cartesian coordinate
y_i	i output variable in computational graph
z	cartesian coordinate
$\mathbf{f}$	scene reflectivity vector
$\hat{\mathbf{f}}$	reconstructed scene reflectivity vector
$\mathbf{g}$	measurement vector
$\mathbf{n}$	noise vector
v_i	i intermediate variable in computational graph
$\hat{P}$	reconstructed plenoptics function
$\mathcal{U}$	uniform noise in height map of hologram
ω_x	beam waist radius of x portion
ω_y	beam waist radius of y portion

Operators

$\mathbf{A}^T$	transpose of matrix $\mathbf{A}$
$\mathbf{A} * \mathbf{B}$	convolution of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{Ax}$	dot product of vectors $\mathbf{A}$ and $\mathbf{x}$
$\nabla \times \mathbf{A}$	curl of vectorin $\mathbf{A}$
$\frac{\partial}{\partial x_i}$	partial derivative with respect to variable x_i
$\sum_i$	sum over index i
$\int_S$	integral over area S
$\mathcal{F}\{\cdot\}$	Fourier transform
$\mathcal{F}^{-1}\{\cdot\}$	Fourier transform

Abbreviations

1-D	one-dimensional
2-D	two-dimensional
3-D	three-dimensional
AD	automatic differentiable
ADMM	alternating direction method of multipliers
ASM	angular spectrum method
CASSI	Coded aperture snapshot spectral imaging
CNN	convolution neural network
CS	compressive sensing
CSPI	compressive single-pixel imaging
DCNN	deconvolution neural network
DOE	diffractive optical element
DUN	deep unrolling network
EM	electromagnetic field
FCMI	frequency-diverse computational microwave imaging
FoV	field of view
FSSs	frequency-selective surfaces
GAN	generative adversarial network
GHz	gigahertz
HSI	hyper-spectrum information
HQS	half quadratic splitting
ISTA	iterative shrinkage/thresholding algorithm
MAE	mean absolute error
MNIST	Modified National Institute of Standards and Technology
NN	neural network
OAP	off-axis parabolic
OEWG	open-ended waveguide
OUI	an object under imaging
RAM	radar absorbing material
RoI	region of interest
RF	Radio frequency
Rx	receiver
SLM	spatial light modulator
SNR	signal-to-noise ratio
SPD	single-pixel detector
SVD	singular value decomposition
Tx	transmitter
VNA	vector network analyzer
VNAX	vector network analyzer extension
WR	waveguide rectangular
Q-factor	quality factor

1 Introduction

Radio frequency (RF) imaging, including millimeter and submillimeter-wave, offers the significant advantage of penetrating materials opaque to visible wavelengths. These capabilities make it a prime choice for applications such as security screening [1], through-wall imaging [2], nondestructive testing [3], and biomedical imaging [4]. The objective of various techniques in this frequency range is to restore a spatial reflectivity map of an object through analyzing scattered field measurements. Submillimeter-wave image formation often relies on electronic or mechanical beam scanning, coupled with quasi-optical methods to transfer energy from the object to sensors. In mechanically scanned systems, the focus on the object plane is directed by mirrors with mechanical drives. However, submillimeter-wave receivers are notably large and challenging to scale into dense arrays similar to those in digital cameras [5]. This limitation usually requires a scanning method, either electronic or optomechanical, to achieve a usable field of view and resolution. Additionally, incorporating complex control circuitry or mechanical scanners increases the system's size, cost, and complexity.

Microwave imaging that utilizes diverse illumination of quasi-randomized fields has shown significant promise in achieving cost-effective, high-resolution, real-time imaging [6, 7, 8, 9]. Unlike traditional beam-scanning methods, these approaches utilize active or passive aperture to generate a set of diverse field distributions and acquire scene measurements that contain ample information for image reconstruction. These methods facilitate more adaptable hardware designs. Frequency-diverse antennas [9], programmable metasurfaces [8], frequency-selective surfaces (FSSs) [10], or phase holograms [6] are typically utilized to create random fields. However, these approaches are either heuristic or depend on proxy metrics for diverse field distributions instead of evaluating imaging quality after post-processing. Finding an optimal computational imaging system remains challenging without a true end-to-end approach that jointly optimizes the imaging model parameters and the data processing algorithm.

Jointly designing computational optics systems and post-processing algorithms has been widely explored in recent years, such as extended depth of field imaging [11], super-resolution imaging [12], snapshot hyperspectral imaging [13, 14], high-dynamic-range imaging [15]. These methods apply automatic differentiation and efficiently integrate the imaging physics model into the workflow of data-driven algorithms to find optimal computational imaging systems. However, there are only a few works utilizing these approaches in frequency-diverse computational imaging. Based on previous research about hologram [6, 16, 17], this thesis introduces a new paradigm for designing dispersive hologram computational imaging system: end-to-end optimization of a hologram concerning the imaging quality of a reconstruction algorithm, using automatic differentiation. An image formation model based on fully-differentiable wave propagation is developed and employed to optimize both the system parameters and the image processing algorithm parameters.

This thesis aims to create an optimal hologram for a sub-millimeter frequency-diverse imaging system by integrating physics and data-driven-based models through automatic differentiation. This effort is well-supported by a literature review of recent computational imaging methods across the optics and millimeter-wave spectra. The

imaging model in this thesis is based on the Fourier optics simulations. The hologram is optimized end-to-end using a data-driven approach that conducts an objective function to the reconstructed image. Before manufacturing the final designs, comprehensive simulations were performed to evaluate the viability of the proposed method and to understand how different design parameters affect the hologram's efficiency, frequency diversity, and manufacturability. Finally, the performance of the imaging system with the manufactured holograms is validated through submillimeter-wave measurements within the 220-330 GHz frequency range.

The structure of this thesis is as follows: Chapter 2 reviews the literature on computational imaging methods across the optics and millimeter-wave spectrum, showcasing examples of previously developed imaging systems. It also delves into the basics of differentiable imaging and highlights recent work in the joint design of imaging systems and algorithms. Additionally, this chapter introduces two primary reconstruction techniques used in computational imaging: numerical optimization and neural networks. Chapter 3 begins by explaining the mathematical theory behind the frequency-diverse imaging model and introduces the simulation method based on Fourier optics. In addition, the quantization-aware differentiable optimization framework for dispersive hologram imaging is proposed in this chapter, which is applied together with the optimization of the reconstruction algorithm. Chapter 4 presents simulation results of the proposed method, detailing the performance of various hologram designs. It also introduces the designs selected for manufacturing. Chapter 5 describes the actual imaging setup with the designed holograms and presents experimental results for the fabricating holograms, including a comparison to the simulations. Finally, Chapter 6 summarizes the work, summarizing the thesis's contributions and findings.

2 Background

2.1 Computational imaging

Computational imaging is distinguished from traditional imaging by its physical processes [18]. Traditional imaging typically involves a direct one-to-one mapping from the object scene to the sensor. In contrast, computational imaging estimates scene images from indirect measurements using sophisticated algorithms, thereby reducing the hardware burden. A typical computational imaging system consists of two main stag y in Fig.

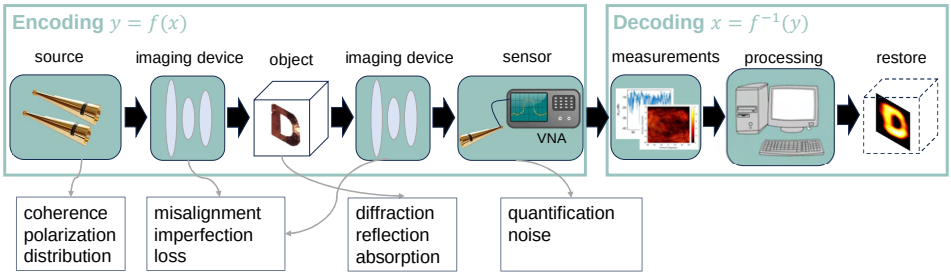


Figure 2.1: A diagram representing a standard computational imaging system. The encoding system includes sources, electromagnetic (EM) propagation, wave-object interaction, and measurement sensors. The decoding process conducts reconstruction of the target from measurements.

For physical encoding, assuming the electric-magnetic field can be will described a function [19] with multiple dimensions of information as below:

$$L = P(\mathbf{r}, \mathbf{\Theta}, t, \lambda), \tag{2.1}$$

where $\mathbf{r} = (x, y, z)$ is the spatial position, $\mathbf{\Theta} = (\theta, \phi)$ is the angle of incident field, t is the time of arrival of the EM-field, and λ is the wavelength. Additional wave properties, such as polarization, coherence, and phase, can also be included in this function. Each dimension of the EM field (space, time, angle, spectrum, polarization, phase, etc.) provides different information about the scene. For example, the spatial information of the field may reflect the distribution of objects in one scene, and the spectrum is strongly related to material properties. The encoding phase converts electromagnetic waves from an object into indirect measurements. This process includes the interaction of waves with both the entire imaging system and the object, and it differs from traditional imaging systems by considering the underlying precise physical model of imaging system $f(\cdot)$ and utilizing decoding stage $G(\cdot)$ to estimate the desired physical quantity from the measurements.

The concept of computational imaging, described by this function, involves using various modulation methods to acquire different projections of the multi-dimensional light field information. This information is then reconstructed using

advanced algorithms, such as numerical methods or neural networks. Next, some typical computational imaging applications based on this concept will be introduced.

2.1.1 Coded aperture snapshot spectral imaging

Coded aperture snapshot spectral imaging (CASSI) aims to encode the 3-D hyperspectrum information (HSI) into a single 2-D compressive image. Subsequently, the corresponding HSI is reconstructed using a reconstruction algorithm [20], which can be represented as:

$$P(x, y, \lambda) \xrightarrow{f(\cdot)} g(x, y) \xrightarrow{G(\cdot)} \hat{P}(x, y, \lambda), \tag{2.2}$$

where $P(x, y, \lambda)$ is the sliced plenoptic function of object scene, $f(\cdot)$ is a coded imaging system, $g(x, y)$ denotes the indirect measurements and $G(\cdot)$ presents the computational decoding algorithms.

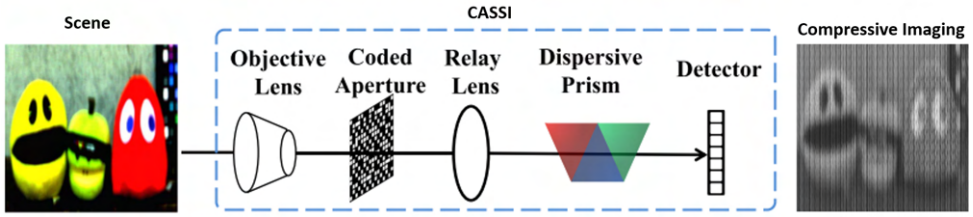


Figure 2.2: CASSI pipeline from [21]. The scene information undergoes three types of optical processing, i.e., spatial modulation, spectral dispersion, and spectral summation.

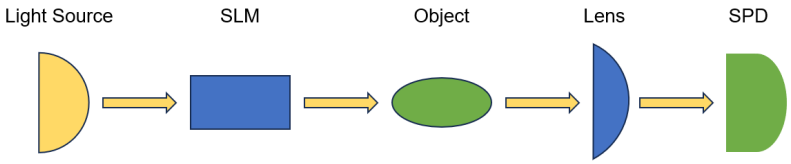
In the CASSI system, as shown in Fig. 2.2, the scene information is initially projected onto the coded aperture, which acts as a spatial modulator. The spatially modulated information is subsequently modified by a dispersive element before it impinges onto the detector. Finally, the compressive measurements in the detector are realized by the summation of the dispersive field over the detector's spectrum range sensitivity.

2.1.2 Compressive single-pixel imaging

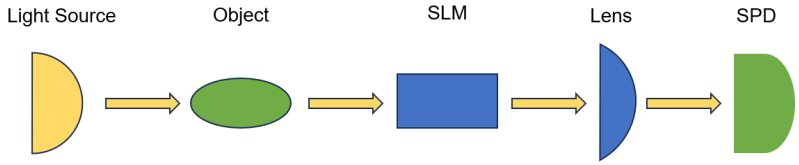
Compressive single-pixel imaging typically involves modulating the light coming from the scene with a series of patterns generated on a spatial light modulator (SLM). Then, the single-pixel detector (SPD) is utilized to acquire the modulated intensity measurements. Each pattern corresponds to a different measurement, and the patterns are designed in such a way that they allow for the efficient reconstruction of the image using compressive sensing algorithms, which can be expressed as:

$$P(x, y) \xrightarrow{f(\cdot)} g(n) \xrightarrow{G(\cdot)} \hat{P}(x, y), \tag{2.3}$$

where (x, y) represents the 2-D plenoptic function of object scene, $f(\cdot)$ is specific coded imaging system and $g(n)$ is the 1-D measurements corresponding to different measurement patterns.



a) Active SPI: modulated light is focused onto the object, and the reflected/transmitted light is collected by an SPD.



b) Passive SPI: the image of the object is sampled by the SLM, and the sampled information is detected by an SPD.

Figure 2.3: Schematics of two types of SPI imaging from [22].

The single-pixel camera [22] mainly consists of four components: a light source with a single frequency, a focusing lens, an SLM, and an SPD. SPI operates in two basic modes, which are forward modulation and backward modulation, as depicted in Fig 2.3. Both techniques can recover the image of the object using the measured signal from the SPD, with a series of binary-coded patterns generated on the SLM that change over time. However, SLMs at submillimeter wavelengths are mostly inaccessible, which limits the capability of single-pixel imaging in this spectrum.

2.2 Frequency-diverse computational microwave imaging

Frequency-diverse computational microwave imaging (FCMI) draws inspiration from compressive single-pixel imaging (CSPI) by encoding scene information using a series of complex field patterns across a broad bandwidth. This approach ensures that the aperture generates various field patterns based on the input frequency. Unlike CSPI, where patterns change over time on an SLM to project the scene into 1-D measurements, FCMI leverages frequency-dependent pattern variation from the aperture to encode the spatial information of the object into the frequency domain.

2.2.1 Frequency diversity via metamaterial aperture

The operation of metamaterial apertures, which generate diverse radiation patterns at microwave frequencies, is introduced. The first frequency-diverse metasurface in Fig. 2.4 was implemented using the diversity of metamaterial elements' resonance frequencies [23]. In this framework, achieving frequency diversity (18-26 GHz) relies on the sharp resonance of metamaterial radiators and the sharpness of the element resonance can be quantified by its quality factor (Q-factor). As the Q-factor of the metamaterial resonators increases, the number of diverse field patterns also increases. However, in practical applications, the Q-factor cannot be increased indefinitely. Strong

resonance leads to significant dielectric and conduction losses [24], which limits the achievable Q-factor. Therefore, a cautious optimization procedure is essential to determine the optimal balance between the Q-factor and radiation efficiency for improved imaging.

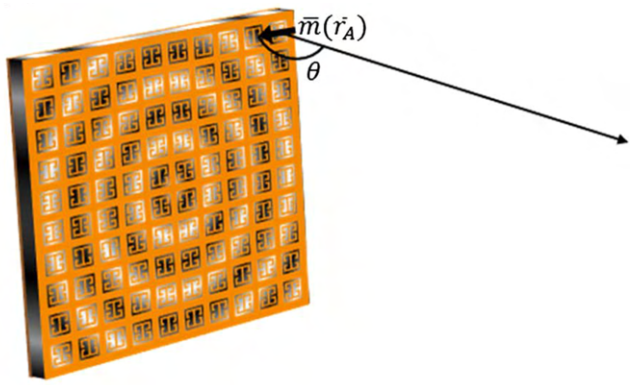


Figure 2.4: The frequency-diverse aperture consists of multiple resonating circuits with varying resonant frequencies distributed across the aperture, resulting in a complex field distribution. (from [23]).

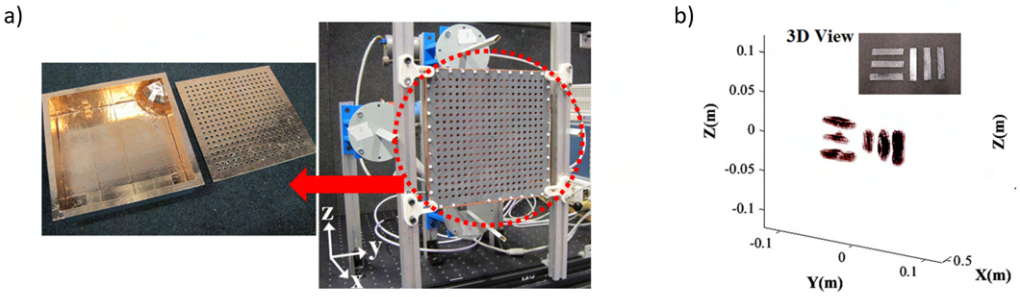


Figure 2.5: a) Overall view of the cavity imaging system including the transmitting cavity. b) Example reconstructed image (from [9])

To resolve the conflict between Q-factor and efficiency, an alternative approach [9] utilizes spatially varied guided modes within the feeding structure to activate the metamaterial components. This technique's primary advantage is its ability to sample the guided mode with weakly resonant meta-components while reducing dielectric and conduction losses. This is achieved without imposing strict limitations on radiation efficiency and Q-factor. The excellent imaging performance of the cavity-backed metasurface is due to its adaptation into planar 2D configurations using a printed cavity as shown in Fig. 2.6 [26, 25, 27]. This setup involves a metasurface printed on a dielectric substrate, with its edges surrounded by conducting walls to create a printed cavity. The cavity is less than half a wavelength thick, supporting single-mode excitation in the broadside direction. Simulations of the field distribution inside this

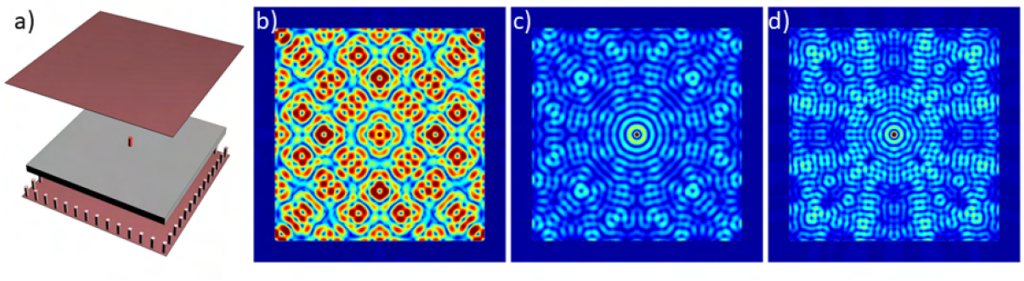


Figure 2.6: a) Frequency-diverse aperture based on a printed cavity. Simulated magnetic field inside the cavity at example frequencies: b) 17.5 GHz c) 22GHz d) 26.5GHz (from [25])

cavity in Fig. 2.6 demonstrate the ability to engineer highly diverse mode patterns at frequencies of 17.5 GHz, 22 GHz, and 26.5 GHz.

2.2.2 Frequency diversity via phase hologram

Holograms are diffractive elements designed to reshape EM field [28], which are categorized into amplitude and phase types based on which characteristic they modulate. The structural distinctions between these two types of holograms are depicted in Fig. 2.7. Amplitude holograms feature a metallic grating on a dielectric film, reflecting and diffracting incoming waves according to the grating's size. This process partially blocks the wave, leading to the lower efficiency characteristic of these holograms. The phase hologram is characterized by a complex transmittance function, reflecting a specific thickness profile. The complex transmittance indicates how the phase of the incoming wave will be changed after transmitting through the structure. The structure of holograms varies based on their intended functionality, such as focusing [29], redirecting [30], splitting [31], and imaging [6].

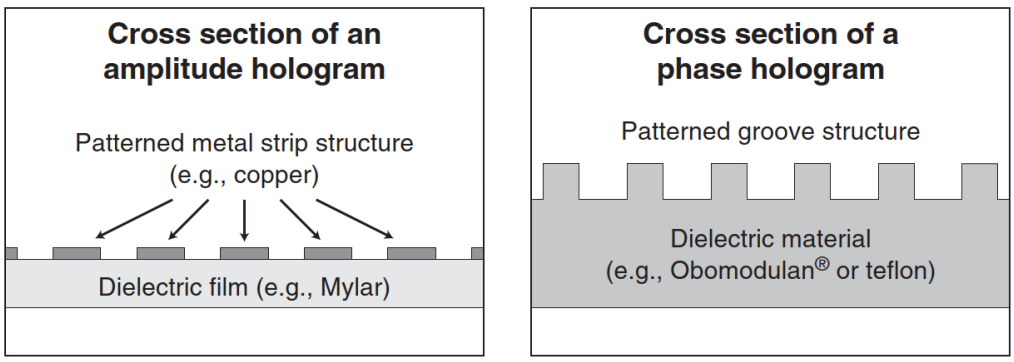


Figure 2.7: Local structure of amplitude and phase holograms.(from [28])

Frequency-diverse phase holograms, unlike conventional radio-wave holograms [32], are designed to generate a spatially varying electric field with frequency at a distance. In [17], a quasi-random synthesizing approach is proposed for frequency-diverse phase holograms that operate at 50–75 GHz (WR-15) and 220–330 GHz

(WR-3.4) bands. In experiments, a trained, fully connected neural network is utilized to localize a corner-cube reflector within the region of interest (RoI) illuminated by a designed dispersive hologram. The results indicate that these frequency-diverse field distributions modulated by the hologram can effectively encode the spatial information of the target.

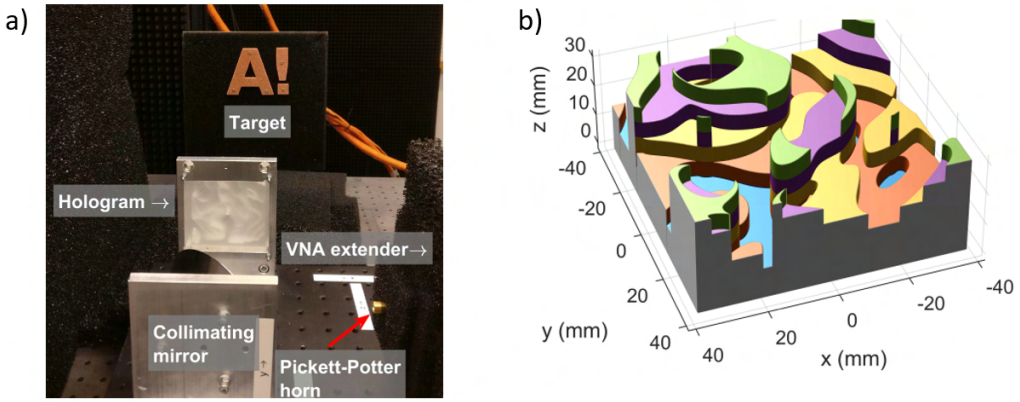
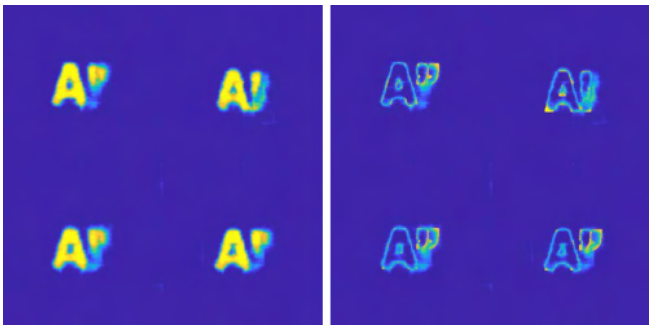
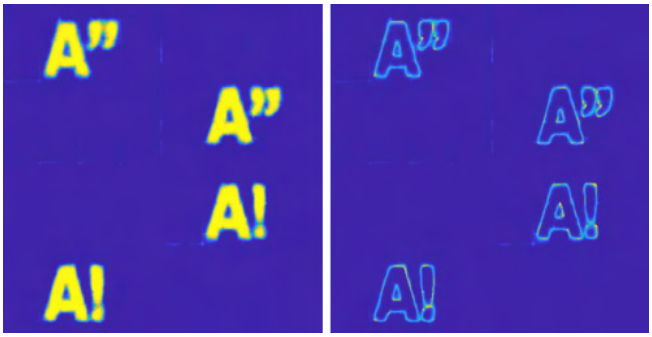


Figure 2.8: a) Photograph of the imaging setup. b) 3-D surface plot of the hologram structure. (from [6])



a) Prediction images and the corresponding absolute value of image error at WR-15.



b) Prediction images and the corresponding absolute value of image error at WR-3.4.

Figure 2.9: Reconstruction results from the neural network from [6].

Additionally, a novel submillimeter-wave imaging technique that combines a dispersive hologram with a neural network is developed in [6] and the system is shown in Fig. 2.8. An off-axis parabolic mirror collimates the Gaussian beam generated by a corrugated horn antenna, and a hologram modulates the collimated beam to illuminate the target in the RoI. A millimeter-wave extender, linked to a vector network analyzer (VNA), captures the hologram's back reflection from the RoI. A trained deconvolution neural network estimates the object image from the reflection signal. Fig. 2.9 demonstrates some reconstruction results at the WR-15 and WR-3.4 bands, respectively.

Despite these studies demonstrating the significant potential of dispersive holograms in submillimeter-wave computational imaging, challenges remain to be addressed. For instance, the designs of dispersive holograms in [6, 17] rely on a quasi-random approach, leaving uncertainty about their optimality for producing varied field distributions. This raises the question: Is there a method to optimize the hologram's thickness profile while considering the reconstruction algorithm's quality? To tackle this issue, this thesis introduces a novel optimization technique for frequency-diverse phase holograms detailed in Section 3.

2.3 Differentiable imaging

Differentiable imaging represents a computational imaging approach in which certain elements of the problem are parameterized as differentiable. This enables the solving of inverse problems through numerical optimization, neural networks, or a blend of both [33]. As illustrated in Fig. 2.10, differentiable imaging models involve extra parameters of the encoding system compared with conventional computational imaging systems presented in Fig. 2.1. These optimizable parameters in imaging systems improve reconstruction quantity and can be used for further improvement or design of ir ot of diffe h in this

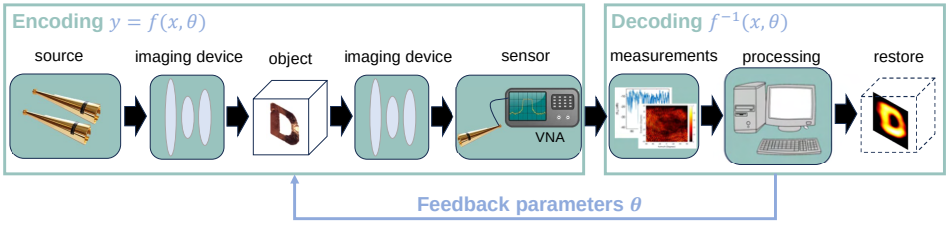


Figure 2.10: A schematic of a differentiable imaging system includes both encoding and decoding processes, integrating an additional parameter θ specific to the imaging system. This parameter can also provide feedback to further improve the imaging system.

2.3.1 Differentiable programming

Thanks to the explosion of deep learning techniques, differentiable programming [34] has rapidly developed in recent years. Typical differentiable programming, as graphically illustrated in Fig. 2.11, is composed of 1) a computation graph, 2) an objective function, 3) a computation of the derivatives of the objective function with automatic differentiation, and 4) a gradient-based optimization.

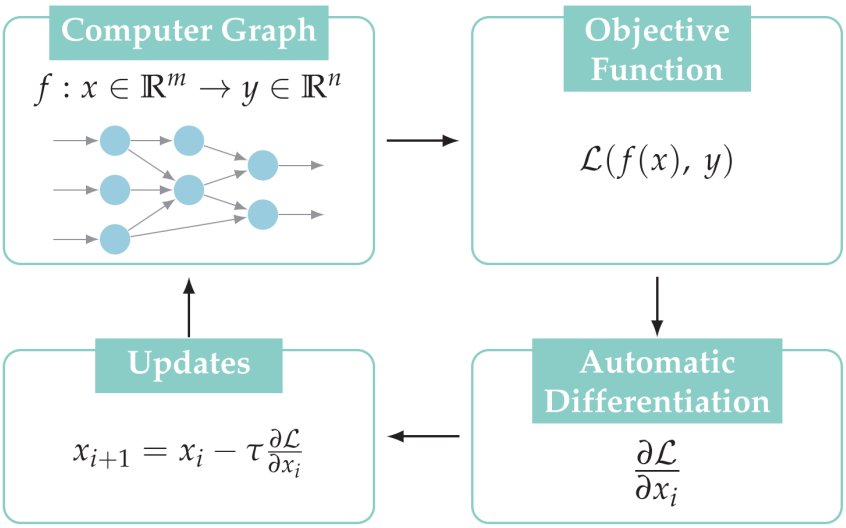


Figure 2.11: A pipeline of a general differentiable programming (from [33]).

To utilize differentiable programming in specific fields, the process begins with constructing a forward computation graph, incorporating necessary physical and mathematical knowledge. This computation graph is a combination of parameterized, differentiable functions, methods, and models that handle the input and output. Differentiation with respect to input parameters and subsequent updates according to the objective function is achieved by computing the derivative of the entire graph. This is done by aggregating the derivatives of basic operations (primitives) using the chain rule for differentiation. The principle in terms of automatic differentiation (AD) [35, 36] is that all numerical computations are fundamentally sequences of finite, elementary operations, each with a known derivative.

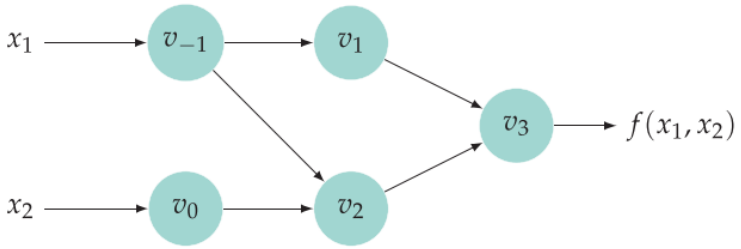


Figure 2.12: Computation graph of the example function. (from [33])

AD can be calculated in two ways, depending on the direction of derivative computation when applying the chain rule: forward mode and backward mode. Forward AD computes numerical derivatives by executing elementary derivative operations simultaneously with the function's evaluation, proceeding from input to output. Conversely, backward AD relies on an expanded forward computational graph to determine the derivatives by traversing the graph in reverse, from output to input.

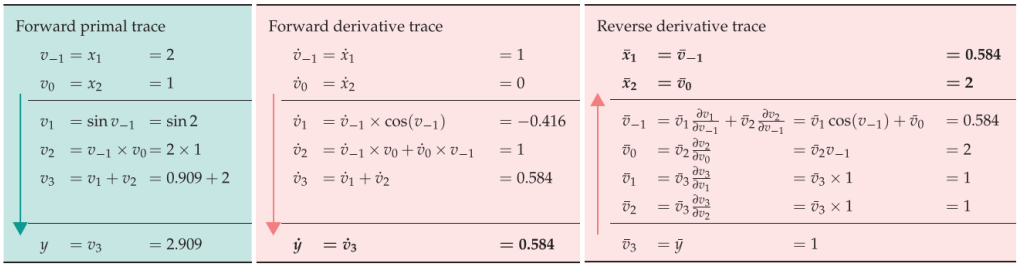


Figure 2.13: AD examples by forward and backward mode, with $y = f(x_1, x_2) = \sin(x_1) + x_1 x_2$ evaluated at $(x_1, x_2) = (2, 1)$ (from [33]) .

To better understand the AD, considering a simple function $y = f(x_1, x_2) = \sin(x_1) + x_1 x_2$, we will demonstrate how to evaluate the derivative $\frac{\partial y}{\partial x_1}$ at point $(x_1, x_2) = (2, 1)$ using forward and backward mode, respectively. The computation graph of this function is represented in Fig. 2.12. For computing the derivative of f in term of x_1 by forward mode, we start by associating with each intermediate variable v_i a derivative

$$\dot{v}_i = \frac{\partial v_i}{\partial x_1}. \tag{2.4}$$

Evaluating the primal v_i and their corresponding derivatives $\dot{v}_i$ from top to bottom gives us the final derivative of $\dot{v}_3 = \frac{\partial y}{\partial x_1}$. In the general case, each forward pass of AD is initialized by setting the derivative of one of the input variables $\dot{x}_i$ to 1 and the rest to zero. This approach allows a full derivative matrix for $f : \mathbb{R}^n \to \mathbb{R}^m$ to be computed in n evaluations. It indicates that forward AD is an efficient and simple method for calculating the derivatives of functions with a small number of input variables.

The backward mode is evaluated in reverse order, beginning at the end of the forward primal trace evaluation and proceeding towards the top with an adjoin, as illustrated in Fig. 2.13:

$$\bar{v}_i = \frac{\partial y_j}{\partial v_i}, \tag{2.5}$$

in the backward mode of automatic differentiation (AD), the process starts by initializing $\bar{v}_3 = \bar{y} = \frac{\partial y}{\partial y} = 1$, indicating the derivative of the output variable with respect to itself. From this point, derivatives such as $\frac{\partial y}{\partial x_1}$ and $\frac{\partial y}{\partial x_2}$ can be efficiently obtained in a single reverse pass. This efficiency makes backward AD particularly suited for evaluating the derivatives of functions with a few number of output variables, like gradient-based optimization and deep learning applications.

This example illustrates the flexibility of AD in extending calculations to more nodes, variables, and operations. When incorporating classical control flows, like if

statements, the computational graph diverges into branches, with gradients calculated only within the relevant branch. Consequently, AD enables derivative computation for mathematical functions and general-purpose programming encompassing control flows, loops, recursions, etc.

2.3.2 Existing work of differentiable imaging

Differentiable imaging offers significant potential in addressing complex imaging challenges through its dual capacity for modeling and design within computational imaging frameworks. This section gives a brief overview of how differentiable imaging can be a pivotal solution for these issues.

Differentiable imaging facilities the joint design of imaging systems and algorithms, enabling a compatible development process that accounts for both hardware capabilities and software strategies. In diffractive snapshot hyperspectral imaging, co-design methods for a single diffractive optical element (DOE) coupled with a data-driven spectral reconstruction network were proposed to restore spectral information from learned spectrally-varying point spread functions [13, 14], shown in Fig. 2.14. Moreover, a broad range of imaging applications based on an optimizable DOE with a differentiable simulator are realized, such as extended depth of field imaging [11], super-resolution imaging [12], and high-dynamic-range imaging [15], which can provide more insights into various computational imaging pipelines.

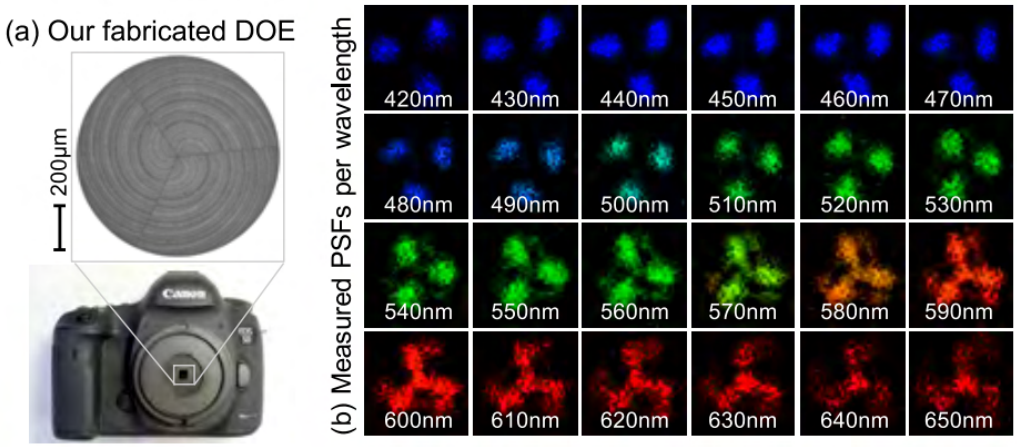


Figure 2.14: a) 3D profile of the fabricated DOE. b) Measured PSFs of the fabricated DOE in the wavelength range from 420nm to 650nm in [13].

By incorporating differentiable programming, it becomes possible to model system imperfections or misalignment, allowing for the simulation and adjustment of various real-world factors that might adversely affect imaging quality. For instance, in holographic display systems, a differentiable wave propagation model, when combined with a parameterized neural network, can accurately learn and compensate for non-idealities, significantly reducing the discrepancy between simulated and actual image formation processes [37]. Furthermore, differentiable imaging is instrumental in

system calibration, offering a systematic method to refine and optimize imaging systems. By automatically adjusting parameters, it ensures the alignment of expected and observed outcomes [38].

While most studies on differentiable imaging focus on the optical spectrum, only a few works extend its application to the microwave range. This is challenging due to the complexity of full-wave simulation methods used in microwave computational imaging frameworks, which require advanced physical and mathematical understanding to implement differentiable programming effectively. However, the submillimeter-wave range offers a promising bridge, as many of the simulation theories applicable in optics remain valid. To the best of my knowledge, this thesis firstly presents an effort to apply differentiable imaging within the realm of frequency-diverse computational imaging in the submillimeter-wave domain, which aims to jointly design the hologram and reconstruction networks.

2.4 Reconstruction methods

Image reconstruction, also known as the decoding process, aims to reconstruct objects from measurements. It primarily addresses the ill-condition inverse problem in computational imaging frameworks. Numerical optimization and neural networks are the two main techniques employed to address ill-condition inverse problems. Numerical optimization typically begins with an analytical theoretical model f that closely represents the actual imaging system. The goal is then to determine x by minimizing a selected error metric that measures the discrepancy between the numerical forward model's predictions and the actual measurements. In contrast, data-driven neural networks aim to learn the implicit representation parameters by minimizing loss functions based on training datasets. This method aims to find out an inverse mapping from measurements to the signal, expressed as $G_{\text{learn}}(\cdot) : y \rightarrow x$. Recently, neural networks have shown great potential to solve highly ill-condition problems in imaging.

2.4.1 Numerical optimization

In a canonical mathematical formulation, classical image reconstruction can be formulated as the solution of a linear equation from the measurement $\mathbf{g}$ to objects $\mathbf{f}$,

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{2.6}$$

where $\mathbf{H}$ is a compact forward operator, and $\mathbf{n}$ represents the additive system noise. It is typically an ill-posed problem. Given the measurements $\mathbf{g}$, general numerical optimization methods usually reconstruct the object $\mathbf{f}$ by solving the following convex problem:

$$\min_{\mathbf{f}} \frac{1}{2}||\mathbf{H}\mathbf{f} - \mathbf{g}||_2^2 + \lambda J(\mathbf{f}), \tag{2.7}$$

where λ is the hyper-parameter to weight the regularization term $J(\mathbf{f})$. The data fidelity term guarantees consistency accords with the forward operator. The regularization

term $J(\mathbf{f})$ incorporates the prior knowledge of one image (e.g., sparsity in a certain domain). Half quadratic splitting (HQS) [39], alternating direction method of multipliers (ADMM) [40], and iterative shrinkage/thresholding algorithm (ISTA) [41] are commonly used optimization algorithms to solve above inverse problem.

Here, the ISTA algorithm with l_1-regularization [41] is used to solve an inverse problem through iterative closed-form derivation. Given $\mathbf{H} \in \mathbb{R}^{m\times n}$, $\mathbf{g} \in \mathbb{R}^m$, find $\mathbf{f} \in \mathbb{R}^n$ by solving,

$$\min_{\mathbf{f}\in\mathbb{R}^n} \frac{1}{2}||\mathbf{Hf} - \mathbf{g}||_2^2 + \lambda||\mathbf{f}||_1, \tag{2.8}$$

where the data fidelity term is a quadratic function which can be handled by gradient descent in the form as,

$$\mathbf{f}_{k+1} = \min_f \left\{\frac{1}{2t_k}||\mathbf{f} - (\mathbf{f}_k - t_k\nabla f(\mathbf{f}_k))||_2^2\right\}, \tag{2.9}$$

where $t_k > 0$ us a suitable stepsize, k denotes iteration counter, $f(\mathbf{f}) = \frac{1}{2}||\mathbf{Hf} - \mathbf{g}||_2^2$, and $\nabla f(\mathbf{x}_k)$ is gradient of f with respect to $\mathbf{f}_k$ which can be represented as,

$$\nabla f(\mathbf{f}_k) = \mathbf{H}^T(\mathbf{Hf}_k - \mathbf{g}). \tag{2.10}$$

Then we introduce the l_1-regularization term into equation (2.9) and let $\mathbf{z} = \mathbf{f}_k - t_k\nabla f(\mathbf{f}_k)$, the quadratic function is changed to

$$\mathbf{f}_{k+1} = \min_f \left\{\frac{1}{2t_k}||\mathbf{f} - \mathbf{z}||_2^2 + \lambda||\mathbf{f}||_1\right\}. \tag{2.11}$$

To find the minimum of $\mathbf{z}$, we differentiate element-wise with respect to $\mathbf{f}$ and set the result equal to zero as,

$$\begin{aligned}
&\varphi(f,z) = \sum_{i=1}^{n} \lambda|f_i| + \frac{1}{2t_k}(f_i - z_i)^2,\\
&\frac{\partial\varphi}{\partial f_i} = \lambda\text{sign}(f_i) + \frac{1}{t_k}(f_i - z_i) = 0,\\
&z_i = \lambda\text{sign}(f_i) + f_i = \begin{cases} f_i - \lambda & f_i < 0,\\ f_i + \lambda & f_i > 0. \end{cases}
\end{aligned} \tag{2.12}$$

Then we may obtain the following piecewise representation for x_i in term of z_i,

$$f_i = \begin{cases} z_i + \lambda & z_i < -\lambda,\\ 0 & |z_i| \le \lambda,\\ z_i - \lambda & z_i > \lambda. \end{cases} \tag{2.13}$$

In short, ISTA [41] solves the inverse problem by iterating between the following update steps,

$$\begin{aligned}
\mathbf{z}_{k+1} &= \mathbf{f}_k - t_k\mathbf{H}^T(\mathbf{Hf}_k - \mathbf{g}),\\
\mathbf{f}_{k+1} &= \text{sign}(\mathbf{z}_{k+1})\max(|\mathbf{z}_{k+1}| - \lambda, 0).
\end{aligned} \tag{2.14}$$

2.4.2 Neural network

Comparing to numerical optimization which needs to design closed-form with specific $\mathbf{H}$ and regularization $J(\mathbf{f})$, neural network-based methods take an end-to-end approach to solve the inverse problem by learning the inverse mapping from a dataset of ground-truth images and their corresponding measurements $\{(f_n, g_n)\}_{n=1}^{N}$, which can be expressed as,

$$G_{\text{learn}} = \min_{G_\theta, \theta \in \Theta} \sum_{n=1}^{N} \mathcal{L}(f_n, G_{\text{learn}}(g_n, \theta)) + \mathbf{J}(\theta), \tag{2.15}$$

where Θ is the set of all possible parameters, $\mathcal{L}$ is the loss function that evaluates the reconstruction quality from the measurements, $\mathbf{J}$ is a regularization term on the parameters to prevent overfitting. As the training step is done, G_{learn} can be used to reconstruct a new image from corresponding measurements.

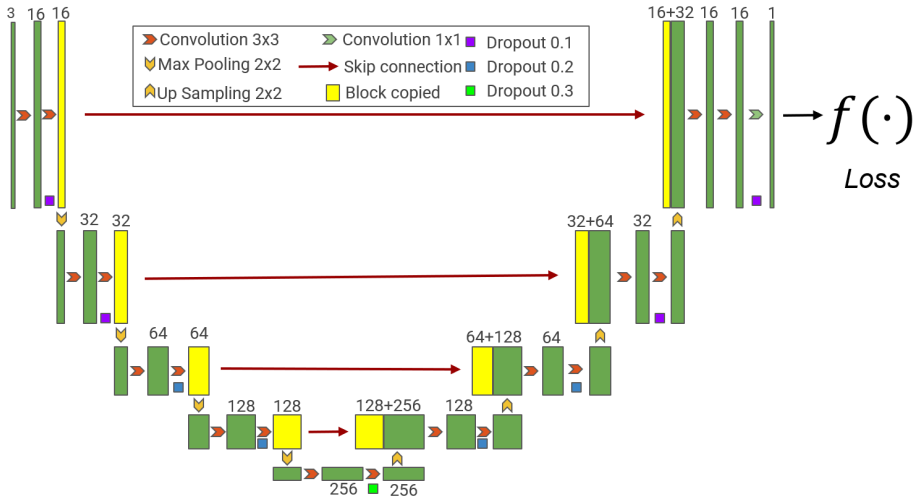


Figure 2.15: An illustration of a typical CNN architecture (U-Net) with a loss function used for training [42].

Neural network-based methods commonly used in computational imaging are divided into convolution neural network (CNN), deep unrolling network (DUN), and generative adversarial network (GAN). A standard CNN comprises convolutional, pooling, activation, and transposed convolution layers. For demonstration, Fig. 2.15 shows a typical CNN architecture. Convolutional layers perform dot product operations between kernels and feature maps. Pooling layers reduce the dimensionality of feature maps, while activation layers introduce non-linearity. Transposed convolution layers increase the size of feature maps. The weight refinement of CNN is achieved through the back-propagation algorithm, beginning with the loss function. Recent successful architectures of CNN include ResNet [43], DenseNet [44], and U-Net [42], which are widely adapted as reconstruction methods in computational imaging [45]. However, CNNs for reconstruction only use data pairs of ground-truth images and corresponding measurements without incorporating prior knowledge of the forward physics process.

It is challenging for a neural network to learn the inversion process from measurements to images with limited data, leading to a lack of generalization ability.

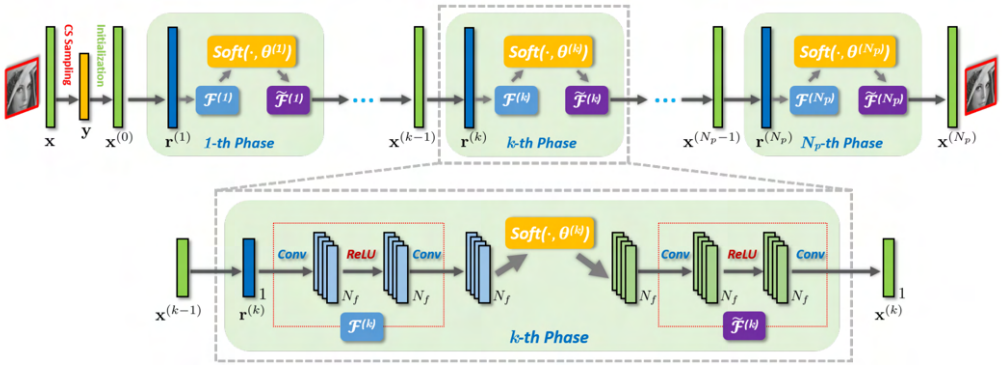


Figure 2.16: An illustration of ISTA-Net [46]. Specifically, ISTA-Net is composed of N_p phases, and each phase strictly corresponds to one iteration in ISTA.

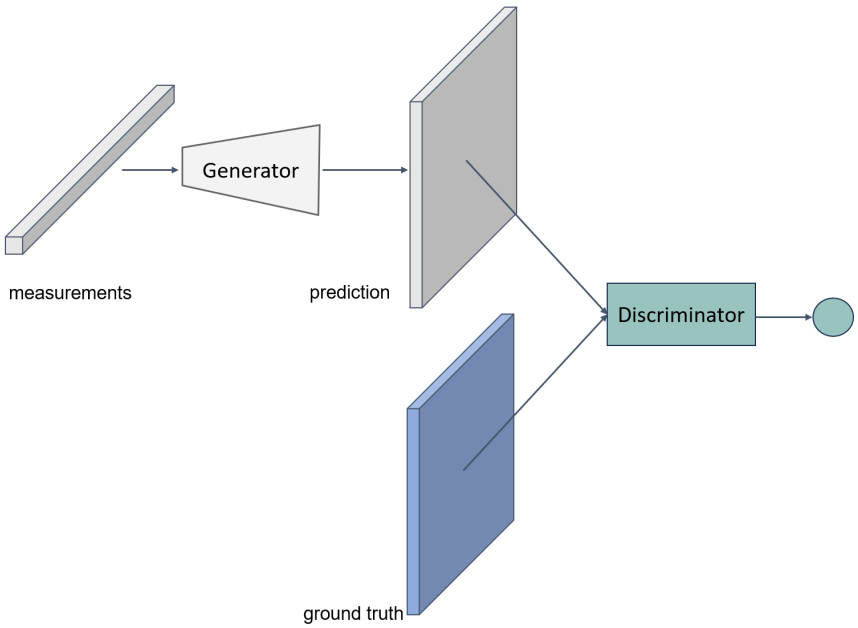


Figure 2.17: The GAN framework for inverse imaging problems. Given an observed measurement, the generator outputs a prediction for the output image, and the discriminator determines whether its input was synthesized by the generator, or comes from the training data.

To address issues associated with CNNs, deep unrolling networks (DUNs) are proposed for solving inverse imaging problems. DUNs combine numerical optimization with neural networks, unrolling each iterative step of numerical optimization into a network layer. DUNs aim to efficiently utilize the advantages of these two paradigms, embedding the forward operator within the network to maintain image consistency with the measured data. For example, Yang et al. [47]. introduced ADMM-CSNet, a deep architecture that unrolls the ADMM solver into iterative steps, with

all parameters learned discriminatively through end-to-end training for compressed sensing applications. Similarly, Zhang et al. [46] transformed ISTA into a deep network, devising an effective approach for solving the proximal mapping with nonlinear transformations. Fig. 2.16 depicts the ISTA-Net framework. It firstly embeds the physics operator A into the gradient descent step to ensure the consistency between measurement and signal domain, then designs $\mathcal{F}(\cdot)$ as a combination of two linear convolutional layers separated by an activation function (ReLU) to learn a richer representation for images. All cited studies illustrate the importance of incorporating the physics prior into the deep learning network, significantly enhancing its generalization capability and convergence performance.

Generative adversarial networks consist of two networks: a generator and a discriminator. In computational imaging, the generator tries to learn a mapping between the measurements and images, whereas the discriminator endeavors to distinguish between the generator's output and the actual data, illustrated in Fig. 2.17. In other words, the discriminator is trained to distinguish between samples from the data distribution and those generated by the generators, while the generator must fool the discriminator. This training scheme allows GANs to learn the transformation between probability densities of measurements and images. GANs have already been used for inverse imaging problems, e.g., single-pixel imaging [48].

3 End-to-end Optimization of Dispersive Hologram with Back-end Neural Network

3.1 Mathematical theory of imaging model

Consider a frequency-diverse imaging system featuring a complex transmitting aperture aligned along z-direction, fed by a single antenna source. The projected electric field scatters off objects in the scene, and the same antenna captures the back-scattered measurements. As depicted in Fig. 3.1, points on the aperture plane are denoted by

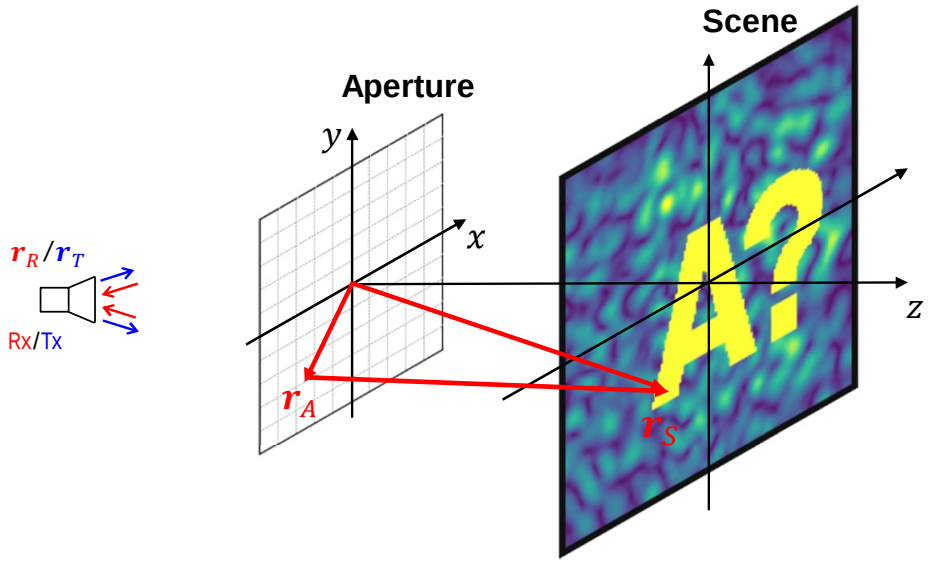


Figure 3.1: The geometry of the proposed imaging system.

The Tx antenna is approximated as a point source located at $\mathbf{r}_\mathrm{T}$, represented by a Dirac delta function $\mathbf{I}_\mathrm{T}(\mathbf{r}_\mathrm{T})\delta(\mathbf{r}-\mathbf{r}_\mathrm{T})$. The field distribution $\mathbf{E}_\mathrm{A}(\mathbf{r}_\mathrm{A})$ across the aperture plane is given by,

Mathematical description of the proposed imaging system.

$$\mathbf{E}_\mathrm{A}(\mathbf{r}_\mathrm{A}) = \mathbf{I}_\mathrm{T} \cdot \mathbf{T}_\mathrm{A}(\mathbf{r}_\mathrm{A}, \mathbf{r}_\mathrm{T}), \tag{3.1}$$

where $\mathbf{T}_\mathrm{A}(\mathbf{r}_\mathrm{A}, \mathbf{r}_\mathrm{T})$ is the impulse response function describing the field at the aperture plane $\mathbf{r}_\mathrm{A}$ due to the point source, encompassing all physical effects from the source to the aperture.

The field distribution $\mathbf{E}_0(\mathbf{r}_\mathrm{S})$ is computed by a convolution integral:

$$\mathbf{E}_0(\mathbf{r}_\mathrm{S}) = \int_\mathrm{S} \mathbf{E}_\mathrm{A}(\mathbf{r}_\mathrm{A}) \frac{\partial}{\partial z} G(\mathbf{r}_\mathrm{S}, \mathbf{r}_\mathrm{A}) d^2\mathbf{r}_\mathrm{A}, \tag{3.2}$$

where $G(\mathbf{r_1}, \mathbf{r_2})$ is scalar Green's function and $\frac{\partial}{\partial z} G(\mathbf{r_S}, \mathbf{r_A})$ denotes the parallel scalar fields-to-fields propagator.

The incident field from the aperture to the object-free region is described as a solution of the Helmholtz equation in free space. In the context of imaging, the object (region of interest) in an inhomogeneous medium is represented by a spatially varying perturbation as follows:

$$\nabla^2 \mathbf{E}_\mathrm{T} + \beta(\mathbf{r}_\mathrm{S})^2 \mathbf{E}_\mathrm{T} = 0, \tag{3.3}$$

where the total field $\mathbf{E}_\mathrm{T}$ is the sum of the incident field $\mathbf{E}_0$ and scattered field $\mathbf{E}_\mathrm{S}$ as

$$\mathbf{E}_\mathrm{T} = \mathbf{E}_0 + \mathbf{E}_\mathrm{S}, \tag{3.4}$$

In the context of electromagnetic fields, disregarding the polarization effects, $\beta(\mathbf{r}_\mathrm{S})$ can be expressed as a scalar function representing the medium's refractive index:

$$\beta(\mathbf{r}_\mathrm{S}) = \beta_0 n(\mathbf{r}_\mathrm{S}) = \beta_0[1 + n_\delta(\mathbf{r}_\mathrm{S})], \tag{3.5}$$

where β_0 is the average wavenumber of the medium, and $n(\mathbf{r}_\mathrm{S})$ is the refractive index, given by,

$$n(\mathbf{r}_\mathrm{S}) = \sqrt{\frac{\mu(\mathbf{r}_\mathrm{S})\epsilon(\mathbf{r}_\mathrm{S})}{\mu_0 \epsilon_0}}. \tag{3.6}$$

where $n_\delta(\mathbf{r}_\mathrm{S})$ refers to the deviation from the average refractive index, and it is presumed that this deviation has finite support, meaning $n_\delta(\mathbf{r}_\mathrm{S})$ is zero outside the object of interest.

Extending the above equation and neglecting second-order terms in n_δ, in accordance with the first Born approximation (where $n_\delta \ll 1$), we arrive at:

$$\nabla^2 \mathbf{E}_\mathrm{T} + \beta_0^2 \mathbf{E}_\mathrm{T} = -2\beta_0 n_\delta(\mathbf{r}_\mathrm{S}) \mathbf{E}_\mathrm{T}, \tag{3.7}$$

in this equation, the term on the right-hand side characterizes our scene, which can be reformulated as $f(\mathbf{r}_\mathrm{S}) = 2\beta_0^2 n_\delta(\mathbf{r}_\mathrm{S})$ symbolizing our estimated object function. Furthermore, the propagation of the incident field $\mathbf{E}_0$ from the aperture to the scene plane adheres to the Helmholtz equation in a vacuum, given by:

$$\nabla^2 \mathbf{E}_0 + \beta_0 \mathbf{E}_0 = 0. \tag{3.8}$$

Similarly, the scattered field $\mathbf{E}_\mathrm{S}$ satisfies the wave equation as follows:

$$\nabla^2 \mathbf{E}_\mathrm{S} + \beta_0 \mathbf{E}_\mathrm{S} = -f(\mathbf{r}_\mathrm{S}) \mathbf{E}_\mathrm{S}, \tag{3.9}$$

which is obtained by substituting equation (3.3) and equation (3.4) into the equation (3.7). To compute the back-scattered field distribution across the aperture, the scene is conceptualized as two-dimensional sources, integrated with the source-to-field propagator $G(\mathbf{r}_\mathrm{A}, \mathbf{r}_\mathrm{S})$ through the principle of superposition:

$$\mathbf{E}_\mathrm{S}(\mathbf{r}_\mathrm{A}) = -\int_s G(\mathbf{r}_\mathrm{A}, \mathbf{r}_\mathrm{S}) f(\mathbf{r}_\mathrm{S}) \mathbf{E}_0(\mathbf{r}_\mathrm{S}) d^2\mathbf{r}_\mathrm{S}, \tag{3.10}$$

having determined the scattered field across the aperture, it is now feasible to compute the measurement g, defined as the field at the detector:

$$g = \int_\mathrm{S} \mathbf{E}_\mathrm{S} T_\mathrm{A}(\mathbf{r}_\mathrm{R}, \mathbf{r}_\mathrm{A}) d^2\mathbf{r}_\mathrm{A}, \tag{3.11}$$

assuming that the field's propagation into the antenna utilizes the same transfer function T_A as the field's emission from an antenna, in accordance with the reciprocity principle of the transfer functions. Referring back to an earlier equation (3.1) and normalizing the source to $|I_f| = 1$, we derive:

$$\mathbf{E}_{\mathrm{A}}(\mathbf{r}_{\mathrm{A}}) = T_{\mathrm{A}}(\mathbf{r}_{\mathrm{A}}, \mathbf{r}_{\mathrm{f}}) = T_{\mathrm{A}}(\mathbf{r}_{\mathrm{R}}, \mathbf{r}_{\mathrm{A}}). \tag{3.12}$$

Subsequently, by substituting the relevant equations (3.10) and (3.12) into the equation (3.11) and reformulating the Green's function,

$$G(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) = \frac{\partial}{\partial z} G(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) D^{-1}(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}), \tag{3.13}$$

where $D(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) = z(\frac{j\beta_0}{R} - \frac{1}{R^2})$ and $R = |\mathbf{r}_{\mathrm{S}} - \mathbf{r}_{\mathrm{A}}|$. Then assigning the $G(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}})$ to the outside integration and rearranging the order of integration yield:

$$\begin{aligned} g &= -\int_{\mathrm{S}} f(\mathbf{r}_{\mathrm{S}}) \mathbf{E}_0(\mathbf{r}_{\mathrm{S}}) \Big[\int_{\mathrm{S}} \frac{\partial}{\partial z} G(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) \mathbf{E}_{\mathrm{A}}(\mathbf{r}_{\mathrm{A}}) D^{-1}(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) d^2\mathbf{r}_{\mathrm{A}} \Big] d^2\mathbf{r}_{\mathrm{S}}, \\ &= -\int_{\mathrm{S}} f(\mathbf{r}_{\mathrm{S}}) \mathbf{E}_0(\mathbf{r}_{\mathrm{S}}) \Big[\int_{z} E_0(\mathbf{r}_{\mathrm{S}}) D^{-1}(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}}) dz \Big] d^2\mathbf{r}_{\mathrm{A}}, \end{aligned} \tag{3.14}$$

with the assumption of narrow field of view ($z \approx R$), it becomes feasible to disregard the comparatively negligible term $\frac{1}{R^2}$ in the estimation of $D^{-1}(\mathbf{r}_{\mathrm{S}}, \mathbf{r}_{\mathrm{A}})$, which can be approximated as $j\beta_0$. Consequently, the equation (3.14) transforms into:

$$g = \frac{j}{\beta_0} \int_{\mathrm{S}} f(\mathbf{r}_{\mathrm{S}}) \mathbf{E}_0^2(\mathbf{r}_{\mathrm{S}}) d^2\mathbf{r}_{\mathrm{S}}. \tag{3.15}$$

Following the discretization of the scene and aperture, the measurement g from the equation (3.15) can be articulated as a summation:

$$g = \frac{j}{\beta_0} \sum_{\mathbf{r}_{\mathrm{S}}} \mathbf{E}_0^2(\mathbf{r}_{\mathrm{S}}) f(\mathbf{r}_{\mathrm{S}}). \tag{3.16}$$

For a frequency-diverse imaging system, equation (3.16) can be succinctly represented in the form of a matrix equation:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{3.17}$$

where $\mathbf{g} \in \mathbb{C}^{\mathrm{M}\times 1}$ symbolizes the back-scattered measurements received by the antenna, the vector $\mathbf{f} \in \mathbb{C}^{\mathrm{N}\times 1}$ denotes the unknown reflectivity distribution of the object, the matrix $\mathbf{H} \in \mathbb{C}^{\mathrm{M}\times\mathrm{N}}$ represents the measurement matrix, which is constructed based on the field distribution at the scene plane. Moreover, $\mathbf{n} \in \mathbb{C}^{\mathrm{M}\times 1}$ is the measurement noise, modeled as a Gaussian distribution with zero mean, and its characteristics are dictated by the system's signal-to-noise ratio (SNR) level. The measurement matrix dimensions M and N denote the number of diverse electric fields corresponding to each frequency and the number of pixels discretizing the field and scene.

3.2 Imaging components and propagators modeling

In this section, we investigate the interaction of the electric field within our frequency-diverse imaging system, concentrating on two scenarios for modeling electric length. The first involves free propagation of the wavefront through a medium, while the second occurs as the wavefront traverses an optical instrument or component.

3.2.1 Mathematical models of imaging components

A. Gaussian Beam. To represent an mm-wave transceiver equipped with a horn antenna, the electric field at the feed plane is modeled as a Gaussian beam:

$$\mathbf{E}_{\text{feed}}(\mathbf{r}_f) = e^{-(\frac{x^2}{\omega_x^2}+\frac{y^2}{\omega_y^2})}, \tag{3.18}$$

here, $r_f = \sqrt{x^2 + y^2}$ denotes the radius from the beam's central axis and ω_x and ω_y are the beam waist radius of x and y portions. In cases where the aperture radius greatly exceeds the Gaussian beam's waist (i.e., $\omega_0 \ll R_{\text{A}}$), the beam approximates a point source, spreading out to form a spherical wavefront akin to that from a point source.

B. 90° Off-axis Parabolic Mirror. The parabolic mirror can be treated as an ideal quadratic focusing element with some approximations [49]. This phase delay is modeled via a quadratic approximation to the spherical wave, related to the lens's focal length and coordinates:

$$U_1(x, y) = P(x, y)e^{-\frac{k_0}{2f}(x^2+y^2)}, \tag{3.19}$$

where β_0 is the wavenumber and f is the focal length. $P(x, y) = circ(R)$ is the common pupil function. To better model the wave interaction with an off-axis parabolic mirror, we use the physical-optics method mentioned in [6, 17] to obtain the field distributions at the hologram plane.

C. Dispersive Phase Hologram. The phase modulation in a hologram stems from its phase profile, a function of the varying height map $h(x, y)$ of a transmissive substrate:

$$\phi(x, y) = \frac{2\pi}{\lambda}(\sqrt{\epsilon_{\text{r, h}}} - 1)h(x, y), \tag{3.20}$$

where λ is the wavelengths and $\epsilon_{\text{r, h}}$ is the relative permittivity of the hologram material, respectively. The total modulation for the incident electric field $\mathbf{E}_{\text{in}}$ on the hologram is expressed as:

$$\mathbf{E}_{\text{hol}}(x, y) = \mathbf{E}_{\text{in}}e^{-\frac{1}{2}k_0 h(x,y)\delta_{\text{h}}\sqrt{\epsilon_{\text{r, h}}}}e^{-j\phi(x,y)}, \tag{3.21}$$

the first exponential term represents the losses due to the hologram's thickness, and δ_{h} is the material's loss tangent.

3.2.2 Mathematical models of optical propagation

For precise modeling of the electric field's free-space propagation, we apply the Rayleigh-Sommerfeld diffraction theory based on Fourier optics:

$$E(x,y,z) = \iint E(x',y',0)\frac{e^{jk_0r'}}{r'}\frac{z}{r'}\left(\frac{1}{2\pi r'} + \frac{1}{j\lambda}\right)dx'dy', \tag{3.22}$$

where $r' = \sqrt{(x-x')^2 + (y-y')^2 + z^2}$, λ is the wavelength and k_0 is the free space wavenumber. The diffraction integral is reformulated into a convolution form using the propagation kernel $h(x,y,z)$ as follows:

$$E(x,y,z) = E(x,y,0) * h(x,y,z), \tag{3.23}$$

where the symbol $*$ denotes two-dimensional convolution over x and y. The propagation

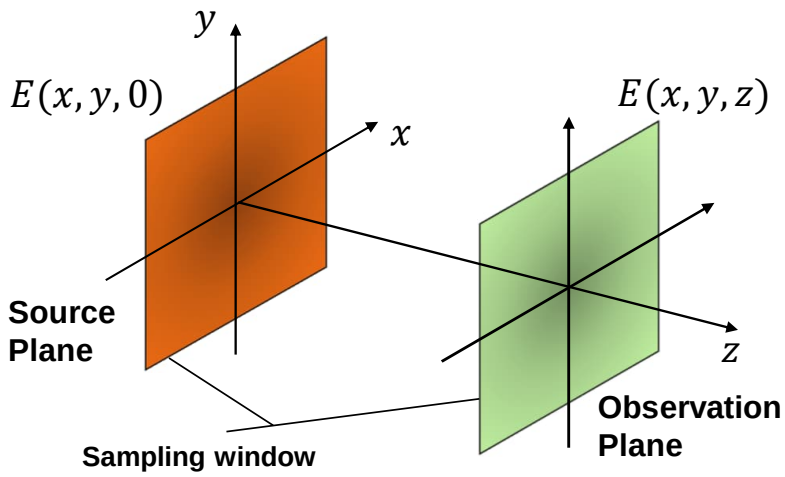


Figure 3.2: Definition of the coordinate system and the geometry of the model.

Recognizing that convolution in the spatial domain equates to multiplication in the frequency domain, the final wavefront field can be expressed as:

$$G(f_x, f_y, z) = G(f_x, f_y, 0)H(f_x, f_y, z). \tag{3.25}$$

The angular spectrum $G(f_x, f_y, 0)$, and the transfer function $H(f_x, f_y, z)$, are given by:

$$\begin{aligned} G(f_x, f_y, 0) &= \iint E(x,y,0)e^{-j2\pi(f_x x + f_y y)}dxdy \\ &= \mathcal{F}\{E(x,y,0)\}, \end{aligned} \tag{3.26}$$

and

$$H(f_x, f_y, z) = e^{[j2\pi \frac{z}{\lambda}\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}]}. \tag{3.27}$$

As a result, the final wavefront is articulated as:

$$E(x, y, z) = \mathcal{F}^{-1}\{\mathcal{F}\{E(x, y, 0)\}[f_x, f_y] \cdot H(f_x, f_y, z)\}[x, y], \tag{3.28}$$

this formulation is known as the angular spectrum method (ASM).

For the formalization of our forward imaging model, it is represented in matrix form·

here,
plane
using
confi

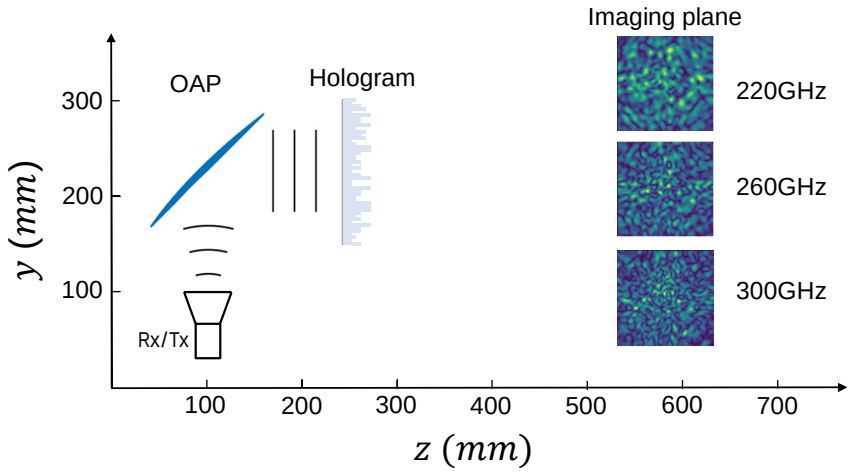


Figure 3.3: Frequency-diverse imaging system and simulated E-field of the scene.

3.3 Differentiable optimization framework

In subsections 3.1 and 3.2, a wave-based image formation paradigm based on Fourier optics is delineated. This model incorporates diffraction and the effects of various elements when imaging an object. Subsequent sections will elucidate the methodology for incorporating this model into contemporary deep learning frameworks, such as PyTorch, ensuring the differentiability of both forward and inverse models. The capacity for differentiability within our optimization schema facilitates the combination of physical principles with neural network architectures, thereby offering potential so ve of ıd co

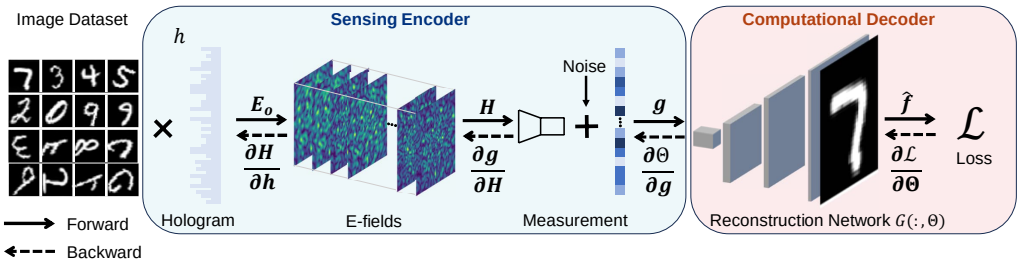


Figure 3.4: The proposed end-to-end differentiable pipeline.

We have formulated a framework called NeuralHolo aimed at optimizing a frequency-diverse imaging system equipped with a phase hologram in Pytorch, utilizing stochastic gradient techniques. The hologram's height map $h(x, y)$ and the parameters Θ of the network $G(\cdot;\Theta)$ are designated as optimization variables. Each forward pass involves simulating the measurement matrix $\mathbf{H}$ in accordance with the wave-based image formation model as per equation (3.29). This simulated matrix is subsequently applied to a set of images, augmented by noise n to simulate measurement uncertainty. Image reconstruction employs a post-processing algorithm, such as a neural network, utilizing the image formation model $\mathbf{H}$:

$$\hat{\mathbf{f}} = G(\mathbf{g};\Theta) \quad \text{with } \mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}. \tag{3.30}$$

A differentiable loss function $\mathcal{L}$ such as mean absolute error in relation to the ground-truth image, is then defined on the reconstructed images:

$$\mathcal{L} = \frac{1}{K}\sum ||\hat{\mathbf{f}} - \mathbf{f}||_1, \tag{3.31}$$

where K represents the total pixel count in the images.

During the backward pass, the error is initially back-propagated to the reconstruction network for updating the parameters Θ:

$$\hat{\Theta}^{(n)} = \Theta^{(n-1)} - \alpha(\frac{\partial \mathcal{L}}{\partial \Theta})^T \mathcal{L}, \tag{3.32}$$

subsequently, it is conveyed to the simulation of the measurement matrix, traversing through the diffraction process, to the phase hologram itself, employing the chain rule:

$$\begin{aligned}\hat{h}^{(n)} &= h^{(n-1)} - \beta(\frac{\partial \mathcal{L}}{\partial h})^T \mathcal{L}, \\ &= h^{(n-1)} - \beta(\frac{\partial \mathcal{L}}{\partial \Theta^{(n-1)}} \cdot \frac{\partial \Theta^{(n-1)}}{\partial \mathbf{g}} \cdot \frac{\partial \mathbf{g}}{\partial \mathbf{H}} \cdot \frac{\partial \mathbf{H}}{\partial h})^T \mathcal{L}.\end{aligned} \tag{3.33}$$

In this process, α and β signify the learning rates for the neural network and the hologram's height map, respectively, in the n-th iteration.

3.3.1 Optimization of heavily-quantized hologram

Currently, most differentiable frameworks typically require continuous values for the hologram phase profile optimization. However, these methods do not consider the physical quantization in hologram fabrication due to the discrete material layers produced with 3D printers, resulting in a mismatch between imaging hardware and reconstruction algorithms in real-world applications. Therefore, we need to develop an approach that optimizes a heavily-quantized hologram, aligning it with specific manufacturing requisites.

A pivotal challenge is the formulation of a differentiable quantization operator q and its integration into the optimization workflow. To address this, we introduce surrogate gradient methods, drawing inspiration from the realm of quantized neural networks [50]. In this approach, the forward pass employs the precise quantization operator q, whereas the back-propagation pass leverages the gradient of a differentiable proxy function $\hat{q}$. This adaptation enables the modification of the gradient descent process for the hologram height map, as delineated in equation (3.33):

$$\hat{\Phi}^{(n)} = \Phi^{(n-1)} - \beta(\frac{\partial \mathcal{L}}{\partial \Theta^{(n-1)}} \cdot \frac{\partial \Theta^{(n-1)}}{\partial \mathbf{g}} \cdot \frac{\partial \mathbf{g}}{\partial \mathbf{H}} \cdot \frac{\partial \mathbf{H}}{\partial q} \cdot \frac{\partial \hat{q}}{\partial \Phi})^T \mathcal{L}. \tag{3.34}$$

The quantization operator q can be defined as:

$$q : \mathbb{R}^{M \times N} \rightarrow \mathbf{Q}^{M \times N}, \quad \phi \rightarrow q(\phi) = \Pi_{\mathrm{Q}}(\phi), \tag{3.35}$$

where Π represents the projection operation, which aligns continuous phase values with their nearest counterparts within a predetermined quantization set Q. Following optimization, the phase map is then converted into a height map using equation (3.20).

Our differentiable proxy function utilizes the Gumbel-softmax method [51, 52], incorporating a continuous relaxation of categorical variables. This method is mathematically represented as:

$$\begin{aligned}\hat{q}(\Phi) &= \sum_{\mathrm{l}=1}^{L} Q_{\mathrm{l}} \cdot \mathcal{G}_{\mathrm{l}}(\mathbf{score}(\Phi, Q)), \\ \mathcal{G}_{\mathrm{l}}(z) &= \frac{\exp\left((z_{\mathrm{l}} + g_{\mathrm{l}})/\tau\right)}{\sum_{\mathrm{l}=1}^{L} \exp\left((z_{\mathrm{l}} + g_{\mathrm{l}})/\tau\right)}, \\ \mathbf{score}_{\mathrm{l}}(\Phi, Q) &= \sigma(w \cdot \delta(\Phi, Q_{\mathrm{l}}))(1 - \sigma(w \cdot \delta(\Phi, Q_{\mathrm{l}}))).\end{aligned} \tag{3.36}$$

here $g_l \backsim \text{Gumbel}(0, 1)$ denotes Gumbel noise for each category $l = 1, \ldots, L$, which correspond to quantized phase levels. σ is a sigmoid function, δ signifies the signed value difference, and w is a scale factor.

Three parameters are crucial during training: the temperature parameter of Softmax τ, the width parameter w, which corresponds to the interval of discrete levels, corresponding to the interval of discrete levels, and a scale factor applied to the score function.

3.3.2 Reconstruction algorithm

The simulated reflection measurement was mapped into an image through a deconvolution neural network (DCNN). Here, we just use a very simple and lightweight NN, which consists of six deconvolution blocks to transform the 1D spectrum reflection to 128×128 reflectivity map of the target. Each block contains a transposed convolution, batch-normalization, ReLU, and dropout layer. This provides a baseline to check the feasibility of the network for the proposed approach. The architecture of the used DCNN is shown in Fig. 3.5.

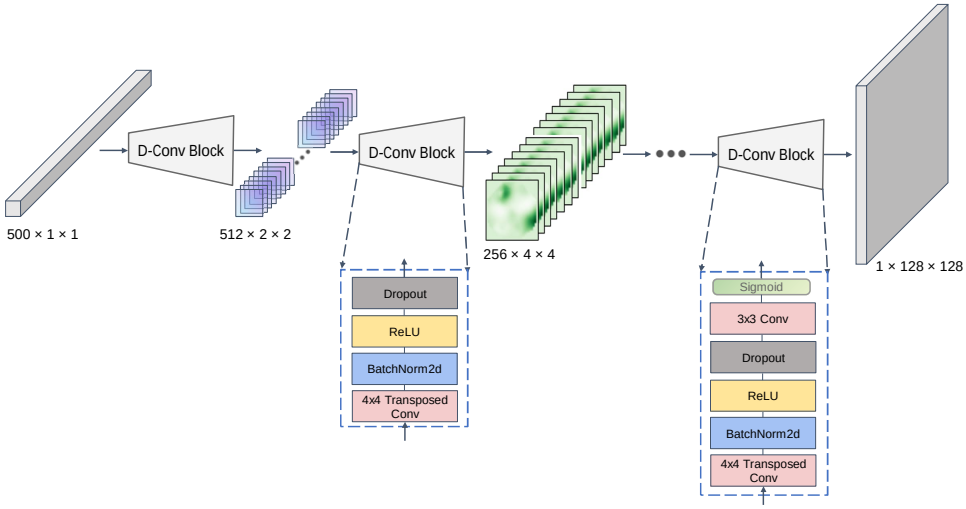


Figure 3.5: The architecture of proposed DCNN. The last block contains a 3×3 convolution and a sigmoid function to set the range of reflectivity map in $[0, 1]$.

We utilize two loss functions to optimize the quality of reconstruction results: mean absolute error (MAE) and dice loss [53], which can be expressed as,

$$\begin{aligned} \mathcal{L}_{\text{total}} &= \alpha \cdot \mathcal{L}_{\text{mae}} + (1 - \alpha) \cdot \mathcal{L}_{\text{dice}}, \\ \mathcal{L}_{\text{mae}} &= \frac{1}{K} \sum ||\hat{\mathbf{f}} - \mathbf{f}||_1, \\ \mathcal{L}_{\text{dice}} &= 1 - \frac{2 \cdot \sum(\hat{\mathbf{f}} \cdot \mathbf{f}) + s}{\sum \hat{\mathbf{f}}^2 + \sum \mathbf{f}^2 + s}, \end{aligned} \tag{3.37}$$

where α is the weighted coefficient of each loss function and here we set it to 0.8 for better performance and s is the smoothing term to stabilize the loss calculation.

3.3.3 Overview of framework and optimization workflow

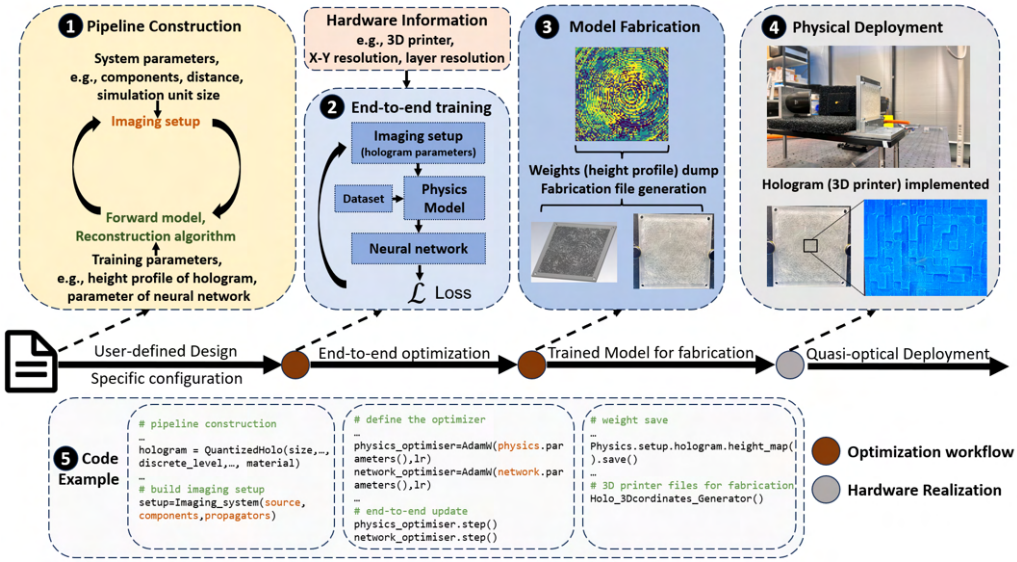


Figure 3.6: End-to-end design workflow overview. 1) Pipeline construction. The design configuration automatically generates the entire end-to-end pipeline, including the imaging setup and training parameters. 2) End-to-end training with the parameter of physical hardware and reconstruction neural network in a data-driven manner. 3) The trained hardware model is saved to create a standard file for 3D printing. 4) Physical deployment for further imaging experiments.

NeuralHolo framework (Table.1) consists of three major components to simplify and accelerate the process of end-to-end design, including a) versatile programming modules for precise physics modeling, b) forward computational model for mapping the target to measurement with customized noise, c) various neural network architecture modules for image reconstruction.

The electric field is the most important feature for precise physics modeling. To ensure flexibility and extensibility in our framework, we introduced a datatype for electric fields that stores spectral and spatial information, along with an easy visualization method. The `ElectricField` class contains a 4-dimensional tensor, with each dimension corresponding to the field components in the xy plane, the wavelength information, and the spatial profile. Both phase and amplitude distributions can be visualized for each predefined wavelength using built-in methods. To model the entire imaging physics process, mathematical modeling modules for various components are required - 1) source beam modeling with flexible wavelength settings and beam profiles such as Gaussian beam and plane wave. 2) precise light diffraction, which includes two categories: the angular spectrum method (ASM) and the Rayleigh-Sommerfeld convolution method (RSC), to handle accurate wave propagation simulation under different distance limitations [54], 3) quantized hologram with an optimizable profile and flexible discrete levels, and 4) general elements such as collimated lenses and amplitude apertures with different shapes. It should be noted that all these modules

Table 1: Overview of the NeuralHolo modules and partial front-end APIs.

Modules for physics modeling	
Class	Description
`ElectricField`	Define a datatype for scalar or vector electric fields using a 4D tensor, including spectrum information and physical size.
`Gaussian_beam`	Define a Gaussian beam with a customized beam waist and wavelengths, returning an ElectricField datatype.
`Plane_wave`	Define a plane wave with customized wavelengths, returning an ElectricField datatype.
`ASM_prop`	Compute the scalar diffraction fields for a given ElectricField using the angular spectrum method with a specific propagation distance.
`RSC_prop`	Compute the scalar diffraction fields for a given ElectricField using the Rayleigh-Sommerfeld convolution method.
`QuantizedHologram`	Define a quantized hologram with an optimizable profile and flexible discrete levels, and compute the modulated electric fields.
`ThinLensElement`	Transparent lens of variable size and focal length.
`ApertureElement`	Rectangular or circular mask of variable size.
Forward computational model	
`Physics`	Contain the imaging forward operator, noise model and sensor model.
Reconstruction algorithms	
`Network`	Customized neural network to reconstruct targets from simulated measurements.

are developed with the `ElectricField` class to process complex fields.

After building the precise physics model of the system, the forward computational model that maps the target information to measurements is utilized based on a specific imaging formation model, such as Eq. 3.29. This `Physics` class contains the imaging forward operator $\mathbf{H}$, a noise model $\mathbf{n}$, and a sensor model $\eta(\cdot)$. In our case, the noise model is Gaussian noise, and the sensor model is identity mapping. The imaging forward operator $\mathbf{H}$ based on the system's physics model and specific imaging formation model, which allows for accurate mapping between the object $\mathbf{f}$ and the measurement $\mathbf{g}$, considering system noise. Finally, the simulated measurements from the `Physics` class are used as inputs for the neural network to reconstruct the corresponding target images. The NeuralHolo framework is built on the conventional automatic differentiation engine PyTorch, allowing for end-to-end design of both the hologram and neural network with GPU acceleration.

Fig. 3.6 shows the end-to-end workflow of the proposed computational imaging system. Using the user-defined design configuration, the entire pipeline is automatically generated in NeuralHolo. This includes imaging components, propagation distance, simulation unit size, and optimized parameters such as the hologram profile and neural network. Then the entire pipeline is trained end-to-end with a specific dataset to find the optimal hologram pattern and corresponding reconstruction neural network. Finally, the optimized hologram profile is converted to standard 3D printing files for manufacturing and further imaging experiments. The code of NeuralHolo framework will be released at https://version.aalto.fi/gitlab/shaos3/NeuralDispersiveHologram.

4 Synthetic Simulations

This section conducts synthetic simulations to evaluate the effectiveness of our approach. The following parts describe the datasets and hyper-parameters used in this synthetic simulation, the comparison of different design parameters, and the performance evaluation of different holograms. Finally, the hologram prototypes chosen for manufacturing are introduced.

4.1 Dataset and hyper-parameters

In the simulation experiment, we chose the classic dataset Extended MNIST (EMNIST) to build our target dataset. This dataset consists of 112,800 training images and 18,800 test images, including both handwritten numbers and letters, each of size 28×28. According to the needs of demanding experiments, we modify the initial image size to 200×200 and regard it as an image that is composed of multiple scattering points with random values between $(0, 1)$ scattering coefficients. The neural network input is the simulated reflection measurement of the target scene based on our differentiable physics imaging process proposed in the previous section.

The parameters of the simulated system are shown in Table 2. The operation bandwidth of our system is 220 - 330 GHz, and the frequency sampling interval is 0.2 GHz. The setup is fed by Pickett-Potter horn antennas, and radiation is collimated by a parabolic mirror with a 0.127 m effective focal length. The distance of the imaging plane between the scene and the hologram is 300 mm, and the resolution of each frequency pattern and field of view is 200×200 with a 1 mm interval size. The measurement matrix H is $550 \times (200 \times 200)$. The hologram uses the high-temperature resin with relative dielectric permittivity of $\epsilon_r = 2.66 + j0.03$ at 300 GHz [55], and its aperture size is 80 mm $\times$ 80 mm with 2 mm pixel size and 5 discrete levels. Each pixel in the hologram is an optimizable parameter during the training process.

Table 2: Main parameters of simulation imaging system.

Parameters	Values
Operation bandwidth	220 ~ 330 GHz
Frequency sampling interval	0.2 GHz
Distance of imaging plane	300 mm
Simulation resolution	200×200 mm
Region of interest	128×128 mm
Sampling interval	1 mm
Hologram's dielectric permittivity	$2.66 + j0.03$
Hologram's pixel size	1×1 mm
Hologram's discrete level	5
Hologram's aperture size	80×80 mm

The phase map of the hologram and the reconstruction NN are trained in an end-to-end fashion. We train the whole model for 400 epochs with a batch size of

64 using the stochastic gradient method with the AdamW optimizer [56]. The initial learning rates for the hologram and neural network are 0.02 and 0.001, respectively, with decay rates of 99.5 % for each epoch. Peak signal-to-noise ratio (PSNR) is used to evaluate the reconstructed imaging quality, which is defined as

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}_{\mathbf{f}}^2}{\mathrm{MSE}(\mathbf{f},\hat{\mathbf{f}})}\right), \tag{4.1}$$

where $\mathrm{MAX}_{\mathbf{f}}$ is the maximum pixel value of the ground truth images, usually 255 or 1 for unnormalized or normalized images, MSE is the mean squared error between the label and reconstruction image. All simulation experiments are built on the Triton high-performance computing cluster at Aalto University with NVIDIA Tesla V100 GPUs and implemented by PyTorch.

4.2 Simulation results

4.2.1 E-field distribution on the imaging plane

The E-field distribution modulated by the hologram at the imaging plane is crucial for the proposed imaging system, as it directly correlates with the scattering coefficients of the scene target. In this section, we firstly demonstrate the 2-D height map of holograms and then compare the E-field distributions from optimized, randomly generated, and previously proposed holograms as detailed in [16], which are shown in Fig. 4.1 and 4.2, respectively.

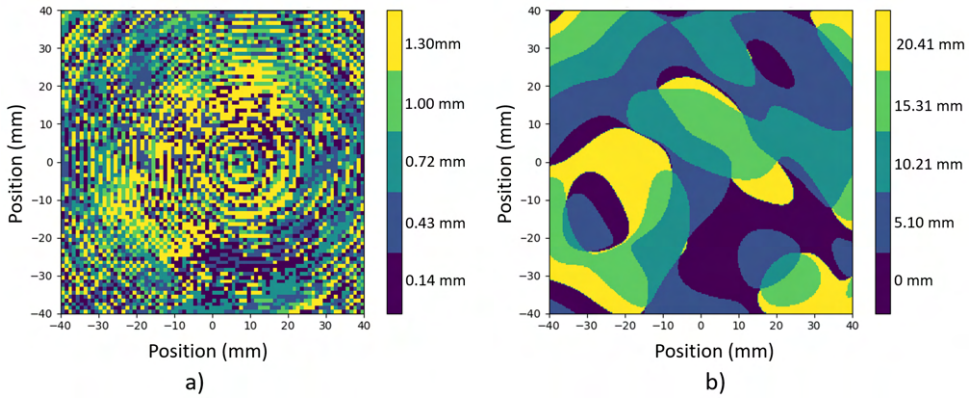
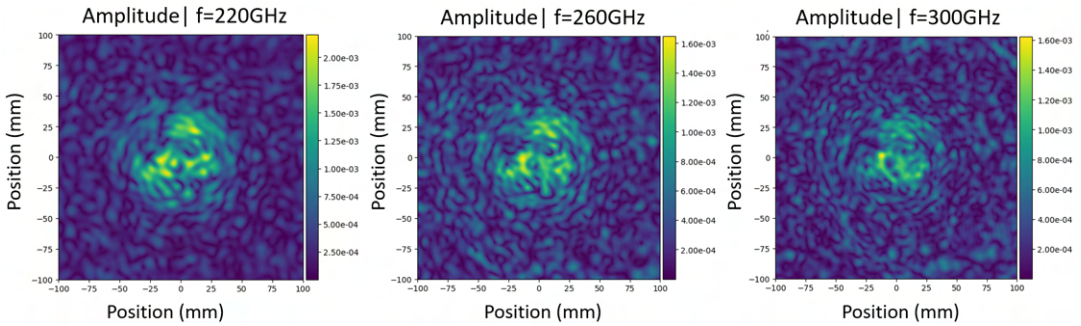
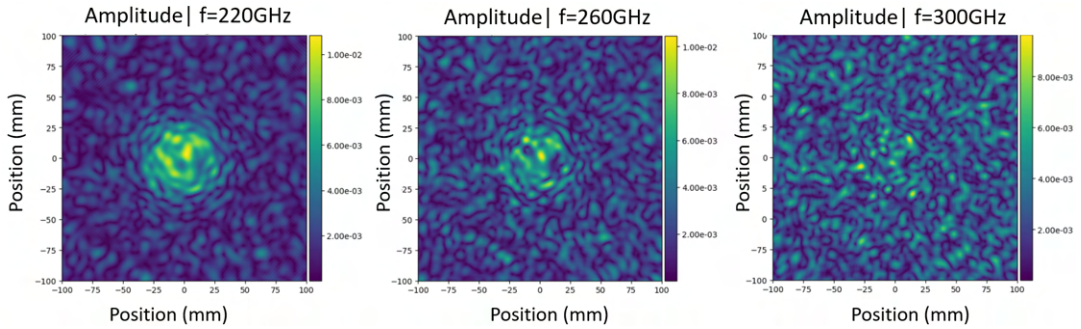


Figure 4.1: The height map of a) optimized and b) previous holograms.

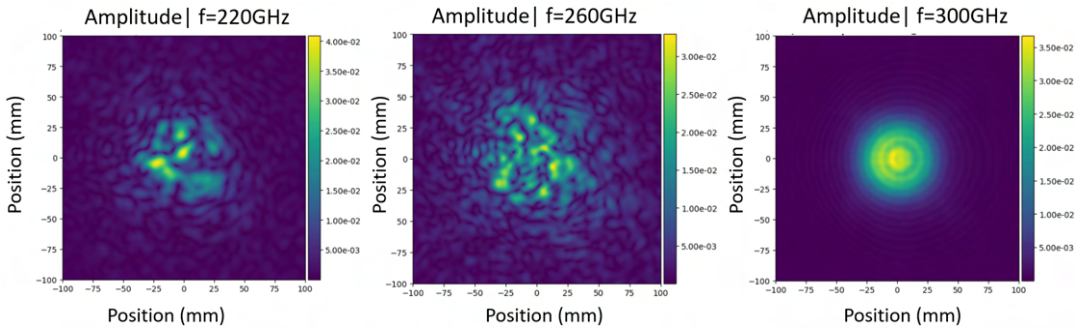
From Fig. 4.2, it is evident that the optimized hologram achieves broader amplitude distributions, allowing for most of the reflected energy to be utilized in encoding scene information from a large field of view (FoV). Additionally, the optimized hologram tends to flatten the amplitude distribution in the region of interest and maintain diversity across all frequency ranges. In contrast, randomly generated holograms lack this property. These characteristics enhance the system's ability to encode information in the spatial domain and improve efficiency in the frequency domain.



a) E-fields from an optimized hologram.



b) E-fields from a randomly-generated hologram.



c) E-fields from the previously-proposed hologram.

Figure 4.2: Visualization of E-field distributions corresponding to a) 220, b) 260, and c) 300GHz.

Compared to the previously proposed hologram, which employs a quasi-random approach with a specific design frequency that does not have the ability of phase modulation in each pixel, this method is less effective. The non-iterative nature of the earlier method makes it difficult to get the hologram in a single attempt. Consequently, most of the energy still concentrates in the center area in this design, leading to weaker capabilities in capturing and utilizing energy from wider angles.

4.2.2 Frequency diversity

The frequency diversity of a hologram refers to the lack of correlation in the spatial field distribution in the RoI across operation frequencies. More similar illumination patterns

at various frequencies result in a diminishing of the overall information extractable from the measurements. In this context, we employed two different metrics to evaluate the frequency diversity of the measurement matrix: singular value decomposition (SVD) and the correlation coefficient. To fairly compare the performance of different holograms, we set the pixel size and discrete levels of the optimized hologram to 1 mm and 5, which is the same as the previously proposed hologram.

The SVD is a mathematical tool used to assess the diversity of measurement matrix produced by holograms. For the measurement matrix $\mathbf{H} \in \mathbb{C}$, the definition of SVD is,

$$\mathbf{H}_{\mathrm{M\times N}} = \mathbf{U}_{\mathrm{M\times N}} \Sigma_{\mathrm{M\times N}} \mathbf{V}^{*}_{\mathrm{M\times N}}, \tag{4.2}$$

where $\mathbf{U}$ and $\mathbf{V}$ denote complex unitary matrices, $(\cdot)^*$ is conjugate transpose of a matrix, and Σ is a rectangular diagonal matrix with non-negative real numbers on the diagonal $\sigma_1, \sigma_2, \ldots, \sigma_m$ in descending order. The diversity of field distributions can be assessed by examining the decay rate of the singular value spectrum. A rapidly decreasing spectrum implies a high correlation between measurements, reflecting low diversity. Conversely, a flat spectrum signifies high diversity among the measurements. Fig. 4.3 shows the normalized singular value spectrum for various holograms across the frequency bands. Both the optimized and randomly generated holograms exhibit better frequency diversity compared to the previously proposed one, as our new approach allows for more detailed hologram profiles. Additionally, the optimized hologram demonstrates a slower decay in its singular value spectrum compared to the randomly generated one, showing the effectiveness of our optimization method.

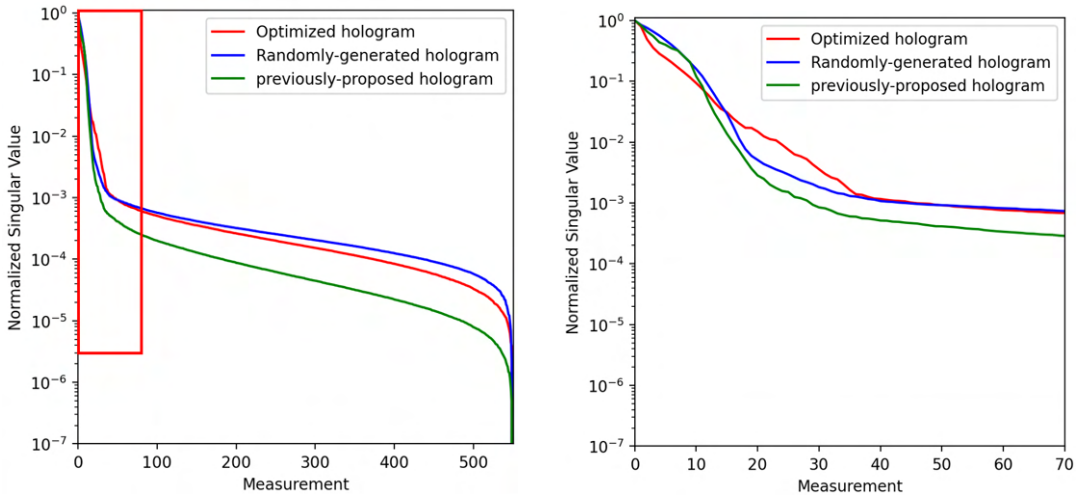


Figure 4.3: a)Normalized singular value comparison of three hologram designs. b) The zoomed-in window presents the difference in the initial decay of singular values.

Additionally, the correlation coefficient quantifies the linear dependence between the measurement matrix's responses at two different frequencies. This coefficient is computed using the complex-valued field vectors, which include the electric field

values at all sampling points within RoI. The definition of correlation coefficient of two electric field patterns is,

$$\rho(H_n, H_m) = \left| \frac{\mathrm{cov}(H_n, H_m)}{\sigma_{H_n} \sigma_{H_m}} \right|, \tag{4.3}$$

where H_n and H_m are complex-valued vectors of n-th and m-th row of the measurement matrix, which is proportional to the square of the corresponding electric field at all sampling points in the RoI, respectively. σ_{H_n} and σ_{H_m} are the standard deviations of H_n and H_m and $\mathrm{cov}(H_n, H_m)$ is the covariance of H_n and H_m. To effectively demonstrate the correlation between various frequency points, the correlation coefficients are presented as 2-D surface plots, shown in Fig. 4.4. In these plots, the x-axis and y-axis represent the frequencies being compared, and the absolute value of the correlation coefficient between the fields at these frequencies fills the plot. A correlation coefficient of 1 denotes perfect linear dependence, whereas a value of 0 indicates no linear relationship between the two vectors. Low correlation coefficient values of the unity diagonal in the surface plots suggest higher frequency diversity in the field patterns, while values approaching 1 imply lower diversity. Fig. 4.4 shows the correlation coefficient plots of the simulated measurement matrix. From Fig. 4.4, we observe some counter-intuitive results that both previously proposed and randomly generated holograms show lower correlation than optimized hologram. The yellow diagonals are thinner than that of the optimized hologram and also exhibit faster decay away from the diagonal, more visible in the cross-section in Fig.4.4 (d)-(f). The contradictory results from SVD and correlation coefficients indicate these two methods may not fully represent the effectiveness of our optimized hologram.

It is important to note that our method jointly optimizes the hologram concerning the output of reconstruction results rather than using some proxy metrics such as SVD and correlation coefficient. Therefore, training the same reconstruction neural network on the same dataset with different types of holograms and comparing the final reconstruction results could be a better way to evaluate the performance of the proposed method. Table 3 shows the PSNR values on the EMNSIT test dataset of different holograms with the same trained reconstruction NN. The end-to-end optimized hologram has an improvement of 0.32 dB compared to the previously proposed hologram. However, the randomly generated hologram suffers from the lower reconstruction results, which only has 13.60 dB in PSNR. It indicates that this randomly generated hologram can not efficiently encode the target information into the measurements even if it has a slower decay of the SVD curve and lower correlation coefficients shown in Fig.4.3 and 4.4.

To better illustrate the optimization process of our proposed method, we visualize the loss curves for these holograms on the validation dataset in Fig. 4.5. Comparing the loss curves of the previously proposed hologram and the optimized hologram, we observe that the loss values for the optimized hologram are initially higher but show a quicker decline before the 50-th epoch. But after the 50-th epoch, the loss curve of the previously proposed hologram decreases more slowly than that of the optimized hologram, indicating that our proposed end-to-end optimization framework can find a more optimal hologram and yield superior results. Furthermore, based

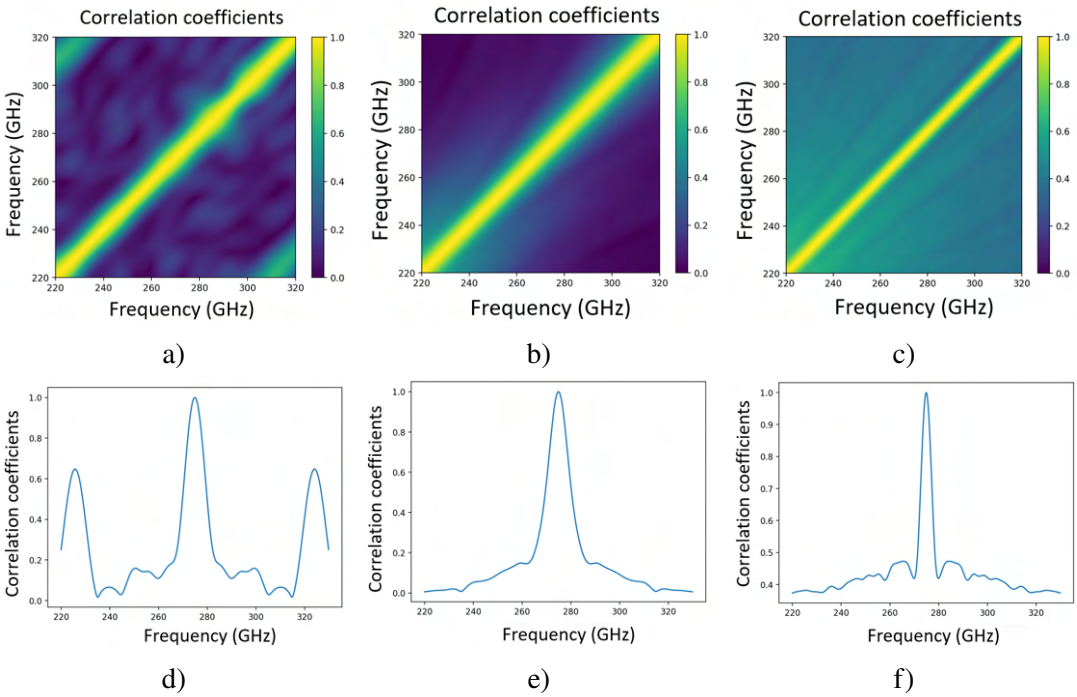


Figure 4.4: a)-c) Correlation coefficients of frequency-diverse field patterns at 220–330 GHz. a) previously proposed hologram, b) randomly-generated hologram c) optimized hologram. d)-f) Cross sections of correlation coefficient plots at 275 GHz corresponding to three different holograms.

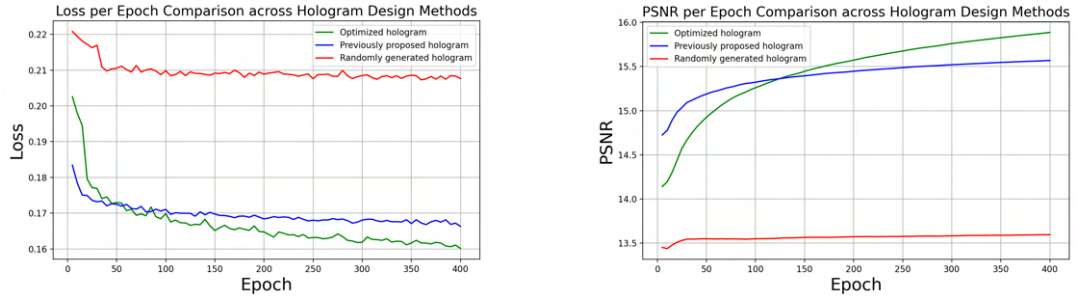
on the comparison of reconstruction results and E-field distributions shown in Fig. 4.2, we can observe some similar properties between the previously proposed and end-to-end optimized holograms. Both E-field distributions from two holograms are more concentrated in the region of interest defined in Table 2, which indicates high efficiency to encode information in this area.

Table 3: Quantitative comparison of the estimation results on EMNIST dataset among three different holograms.

Hologram type	Pixel size	Discrete level	PSNR (dB) ↑
previously proposed	1 mm	5	15.57
randomly generated	1 mm	5	13.60
end-to-end optimized	1 mm	5	15.89

4.2.3 Reconstruction results on test dataset

To evaluate the imaging performance of our proposed method, we simulate the imaging process as described in section 3.1. This involves obtaining measurements of a scene at the imaging plane and applying the deep learning algorithm outlined in section



a) loss curves b) PSNR curves

Figure 4.5: Comparison of a) loss curves and b) PSNR curves with different hologram design methods.

3.3.2 to reconstruct the target. Fig. 4.6 presents the reconstruction results from the MNIST test dataset.

The reconstruction results demonstrate that the proposed algorithm can successfully rebuild various target scenes from the 1D measurement. Specifically, the algorithm effectively recovers targets with different rotations, and it can also restore targets of various sizes and shapes. These capabilities indicate that our hologram can generate diverse field distributions, encoding substantial information into the 1D measurement and allowing the algorithm to effectively recover the target information from these measurements. However, the reconstruction results reveal some information loss, such as incorrect amplitude prediction at the edges of the field of view. This issue arises because less energy is distributed to the edges of the field distribution at the imaging plane, resulting in reduced information encoding at the edges.

4.2.4 Robustness for manufacturing uncertainty

Although high manufacturing precision is achievable using Formlabs' 3D printers to design hologram profiles, perfectly replicating the simulated holograms is still impossible. To assess the robustness of the proposed hologram design approach against manufacturing uncertainties, we introduce uniform noise $\mathcal{U}$ with various tolerances u, denoted as $\mathcal{U} \in [-u, u]$, to the height map of the designed hologram. This simulation of fabrication noise allows us to evaluate the reconstruction quality of the algorithm.

In Figure 4.7, we present the designed hologram with various manufacturing tolerances of u=100um and u=200um, respectively. Given that the layer thickness resolution for Formlabs' 3D printers is 25 μm, which is significantly smaller than these tolerances. It is notable that our algorithm can still correctly reconstruct the target even with the higher manufacturing tolerance of u=200 μm. Although the simulated measurements vary slightly, they maintain the overall trend across different tolerances, demonstrating the robustness of our hologram design against manufacturing uncertainties.

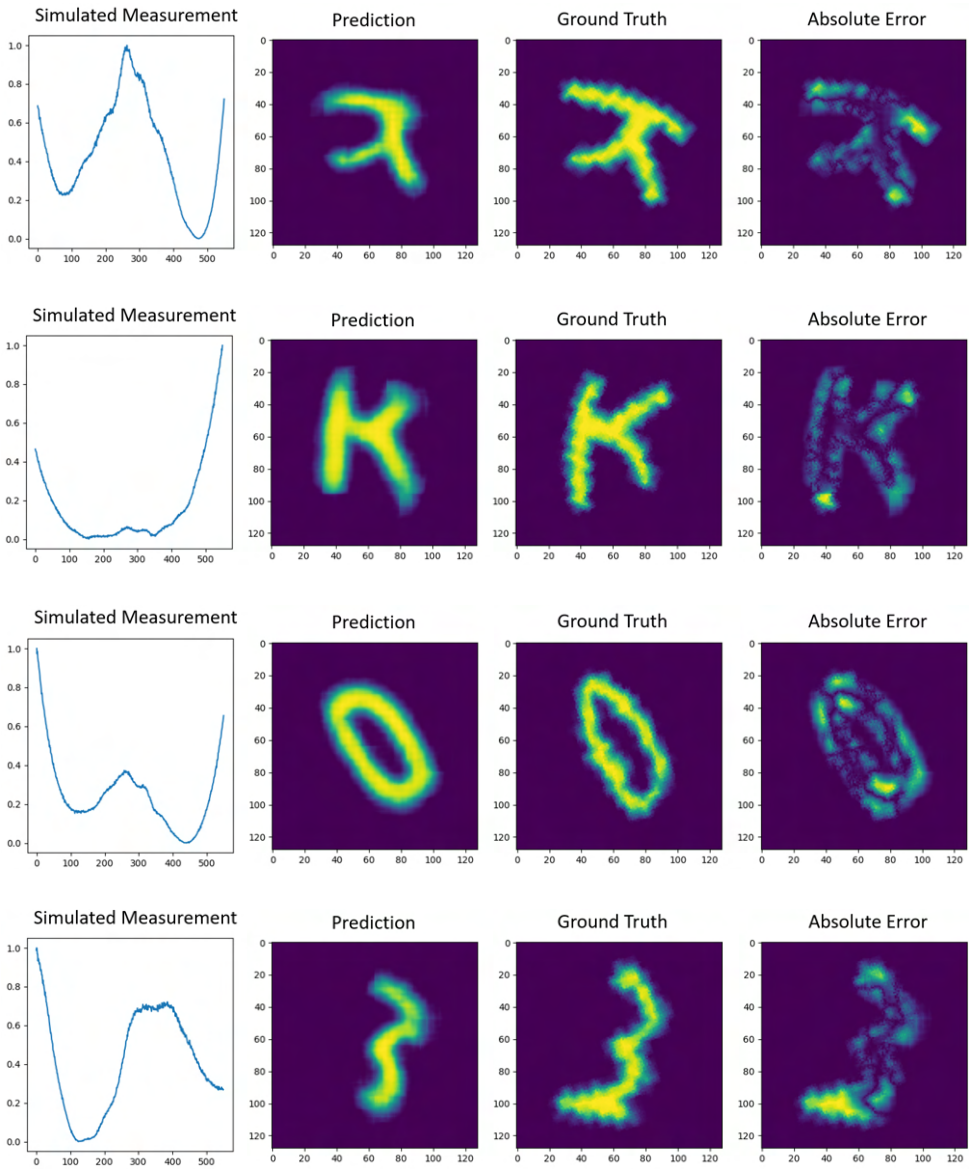
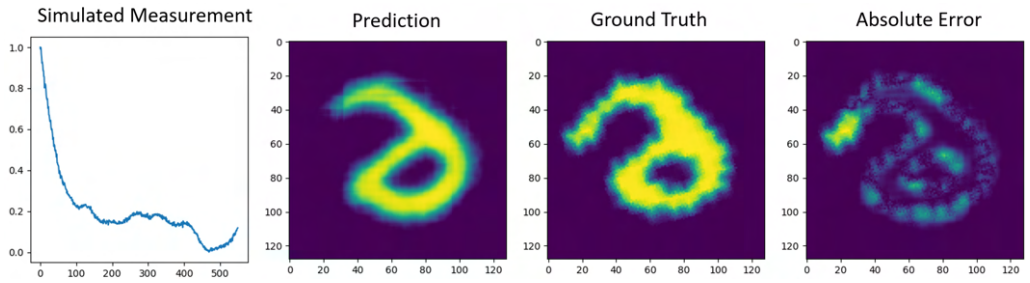


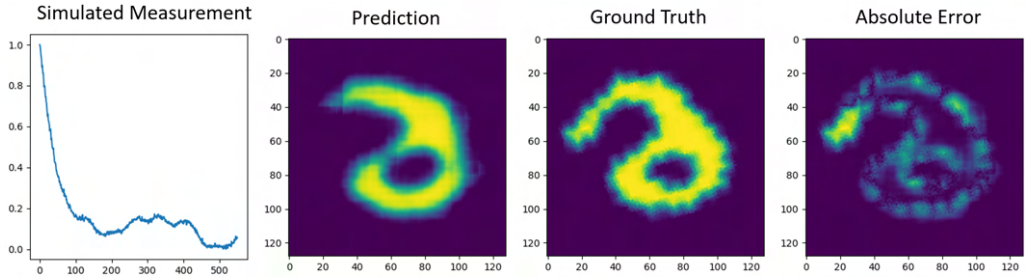
Figure 4.6: The visual comparison of the reconstruction results is displayed as follows: The first column shows the normalized simulated measurement used as the input for the neural network (NN). The second and third columns represent the NN's predictions and the ground truth, respectively. The last column illustrates the absolute error between the prediction and the ground truth.

4.2.5 Generalization of reconstruction algorithm

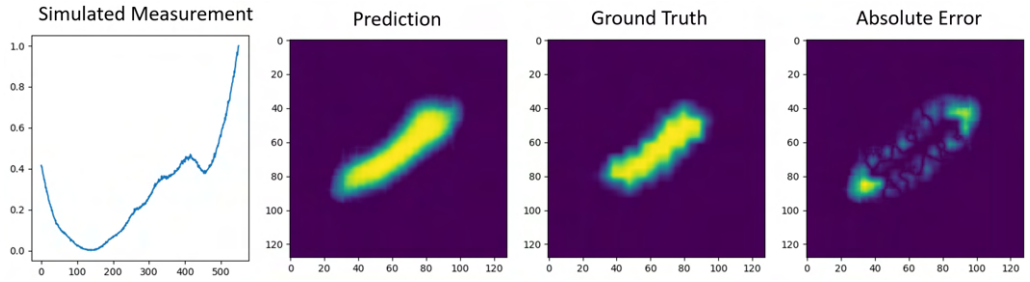
To verify the generalization capability of our proposed algorithm, we tested it on geometries of various shapes and sizes that were not included in the training or testing datasets. Fig. 4.8 shows the reconstruction results for targets that are outside of the EMNIST dataset commonly used. It is evident that our algorithm accurately predicts the location of each target, regardless of their sizes and shapes, demonstrating that



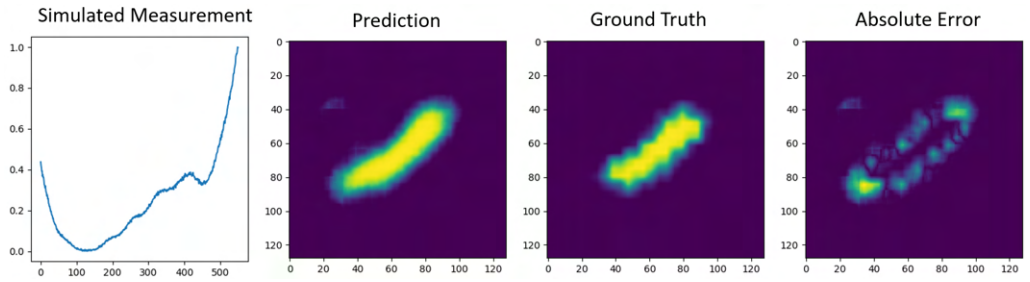
a) With manufacturing tolerance $u = 100$ μm



b) With manufacturing tolerance $u = 200$ μm



c) With manufacturing tolerance $u = 100$ μm



d) With manufacturing tolerance $u = 200$ μm

Figure 4.7: The visual comparison of the reconstruction results with different manufacturing tolerances.

our optimized system can encode information into 1D measurements from general targets that it has never seen before. However, accurately recovering the exact size and shape remains challenging, highlighting some limitations of our algorithm. These limitations could potentially be addressed by utilizing a larger dataset featuring a more diverse range of target shapes and sizes.

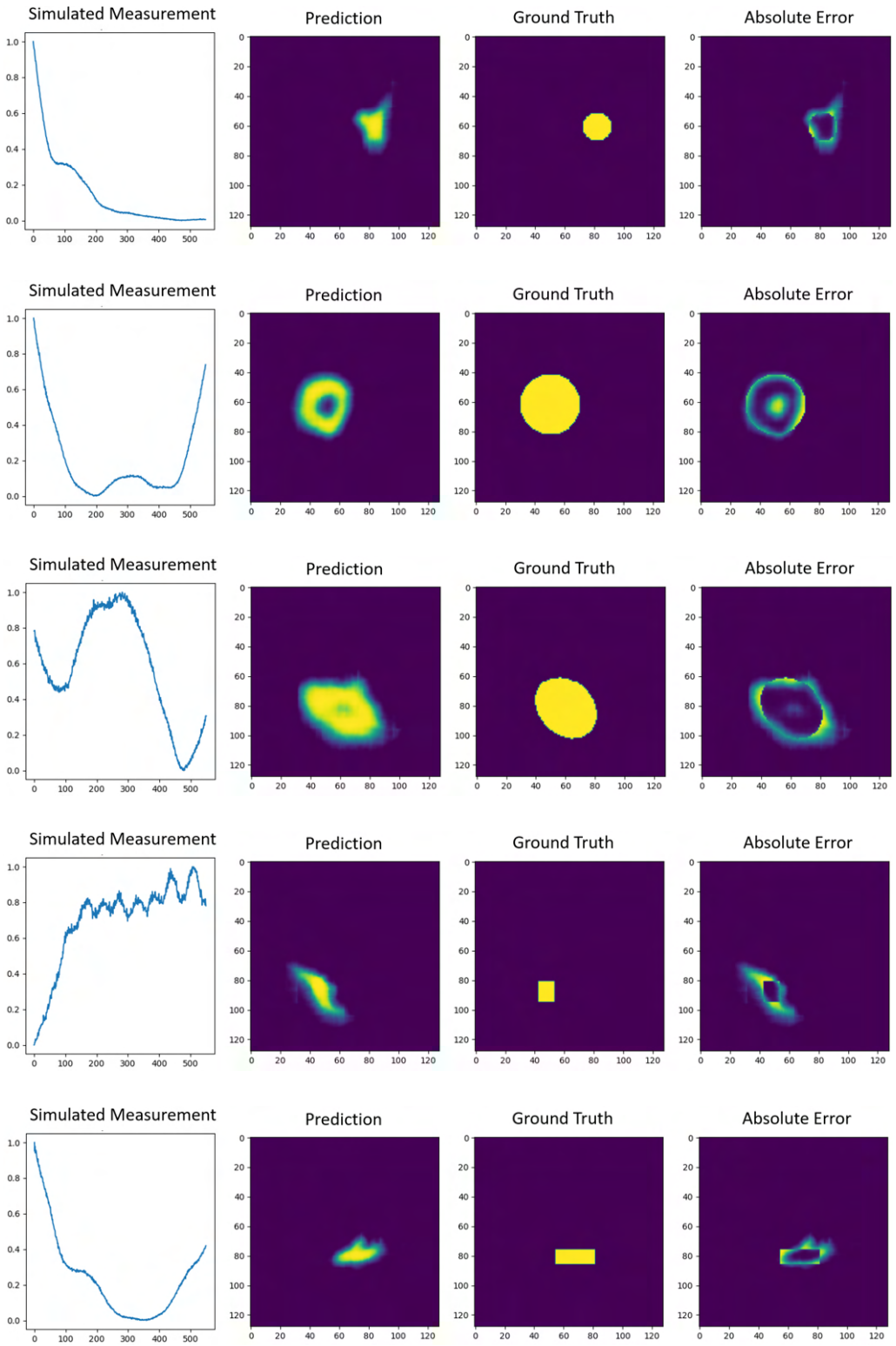


Figure 4.8: The visual comparison of the reconstruction results for targets that are outside the dataset distribution.

4.3 Comparison of design parameters

In this section, we examine the impact of design parameters on the hologram structure. These parameters include the quantized discrete levels and pixel size of the hologram. This study aims to provide a fundamental understanding of how these design parameters affect hologram properties, enabling the choice of an optimal balance between performance and manufacturing complexity.

Table 4: Quantitative comparison of the reconstruction results on EMNIST dataset with different design parameters

Aperture size	Pixel size	Levels	PSNR ↑
80 × 80 mm	2 mm	3	13.78
80 × 80 mm	1 mm	3	15.03
80 × 80 mm	2 mm	4	14.05
80 × 80 mm	1 mm	4	15.39
80 × 80 mm	2 mm	5	14.07
80 × 80 mm	1 mm	5	15.89
80 × 80 mm	2 mm	6	14.28
80 × 80 mm	1 mm	6	16.99

Table 4 displays the impact of various design parameters on the final reconstruction results for the EMNIST dataset. It shows that the PSNR increases with more discrete levels, indicating that the hologram has a greater capability to modulate the field distributions across the illumination frequency band. Similarly, the trend in pixel size demonstrates that smaller pixels correspond to a greater ability for phase modulation, which in turn improves the final results. Fig. 4.9 shows the normalized singular values for different design parameters. It reveals that the singular values become flatter with an increase in discrete levels and a decrease in the pixel size of the hologram.

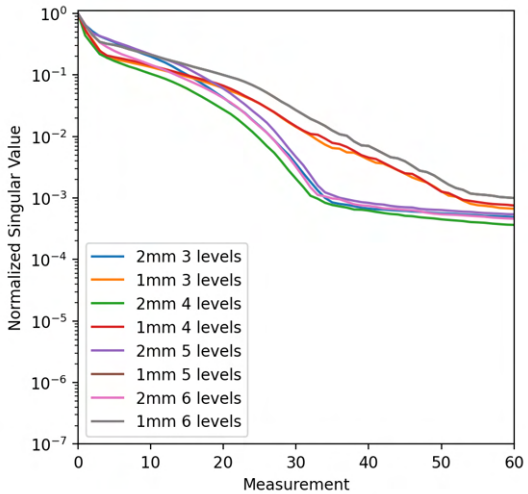


Figure 4.9: The normalized singular values for different design parameters.

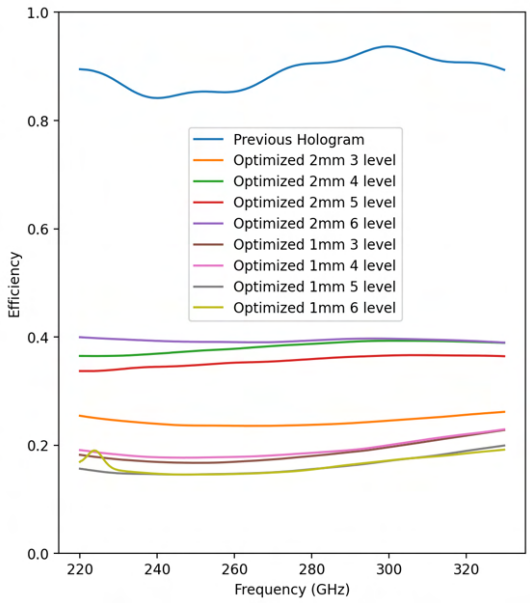


Figure 4.10: The impact of various parameters on the simulated efficiency of the hologram. The upper blue line shows the efficiency of the previous hologram using the material of Rexolite 1422 ($\epsilon_r = 2.53, \tan\delta = 0.0005$).

The power efficiency of a hologram is defined as the ratio of the power within the RoI with the hologram to the power incident on the hologram. Fig. 4.10 depicts the influence of various design parameters on the hologram's efficiency across its operation frequency. The line in blue in Fig. 4.10 represents the efficiency of a previously designed hologram, which maintains an overall efficiency above 0.8. The relatively high efficiency of this hologram is attributed to its smoother surface profile, which minimizes diffraction effects. Additionally, the difference in materials used for the two types of holograms, with loss tangents of 0.0005 and 0.03 respectively, also contributes to the higher efficiency. This figure also shows a clear decrease in efficiency as the pixel size of the hologram decreases. The efficiency of optimized holograms with a 2 mm pixel size remains above 0.3, whereas those with a 1 mm pixel size fall below 0.3. The low efficiency of the optimized holograms is due to the larger loss tangents of the materials used and the high diffraction effects, which result from their detailed structures. These high diffraction effects make some energy outside the region of interest, reducing efficiency. This phenomenon becomes evident when comparing the efficiency of optimized holograms with varying pixel sizes. Specifically, a hologram with a 1 mm pixel size, which introduces more detailed structures and higher diffraction effects, shows a significant decrease in efficiency.

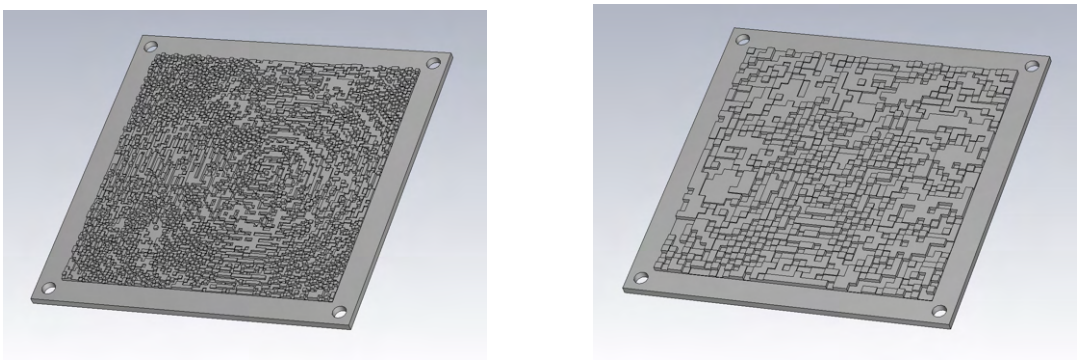
Considering the reconstruction results and power efficiency across different design parameters, it is evident that a more detailed hologram profile correlates with enhanced modulation capability, resulting in improved reconstruction outcomes. However, this also leads to lower power efficiency due to increased diffraction effects. Therefore, designing an optimized hologram that balances greater modulation ability with high efficiency will be the focus of our future work. It is possible to achieve this goal by

introducing an energy regularization term in the loss function, forcing most of the energy to hit the region of interest during the optimization process.

4.4 Manufactured prototypes

The surface profiles of the designed holograms are depicted in Fig. 4.11. Both holograms feature five discrete levels of thickness. Prototype 1 exhibits a more detailed surface structure compared to Prototype 2. This difference in structure is primarily due to the pixel sizes defined before the optimization process.

Both designs were manufactured using Formlabs' 3D printer with high-temperature resin. Fig. 4.11 includes a 2 mm substrate made of the same material, on which the surface structure is printed. Four holes are placed at the corners for mounting in the actual measurement setup. The x-y resolution and layer thickness of the 3D printer are both 25 µm, which demonstrates the manufacturability of the two holograms. Two manufactured holograms are presented in Fig. 4.12.



a) Prototype # 1 b) Prototype # 2

Figure 4.11: Surface profiles of holograms chosen for manufacturing. a) Designed hologram with 1 mm pixel size and b) hologram with 2 mm pixel size.

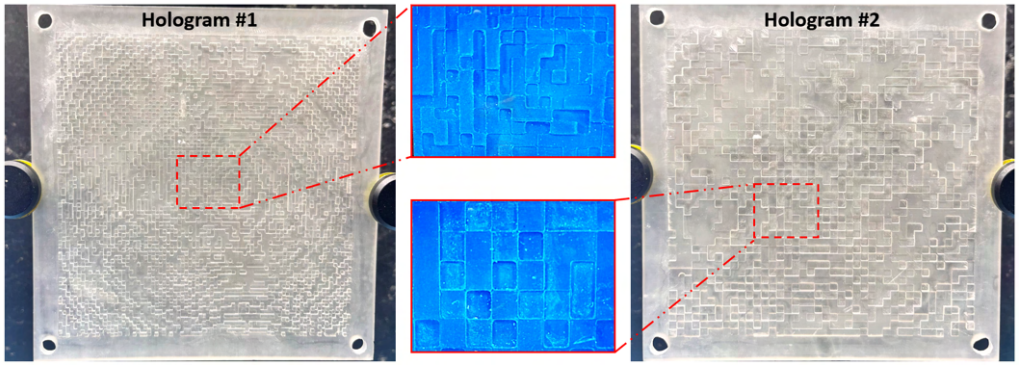


Figure 4.12: Manufactured holograms used in measurements. Left is a hologram with a 1 mm pixel size, and the right is that with a 2 mm pixel. Blue figures represent zoom-in areas under a microscope.

5 Physical Experiment

This section details the measurement setup and procedures, as well as the hologram measurements. It presents the measured field patterns and frequency diversity for both manufactured holograms across the operation frequency band, comparing these with the simulated results. Finally, we will compare the measured reflected signals with the simulated ones to verify the accuracy of the imaging system modeling.

5.1 Measurement setup

The 220-330 GHz measurement setup is shown in Fig. 5.1. It consists of submillimeter-wave extension modules (Virginia Diodes WR3.4-VNAX) connected to a Keysight N5225A PNA Microwave Network Analyzer, which serves as both a signal generator and analyzer. The transceiver VNAX extension modules are coupled with a Pickett-Potter horn antenna, which feeds into an off-axis parabolic (OAP) mirror to collimate the beam. The OAP mirror has a diameter of 76.2 mm and an effective focal length of 127 mm. The distance from the hologram surface to the open-ended waveguide at the receiver is 300 mm. An open-ended waveguide (OEWG) for the corresponding frequency range is used as a probe antenna and is coupled to the VNAX receiver module to perform x-y scanning in the measurement plane. Flat surfaces near the measurement setup are covered with radar-absorbing material (RAM) to prevent multipath propagation and unwanted reflections.

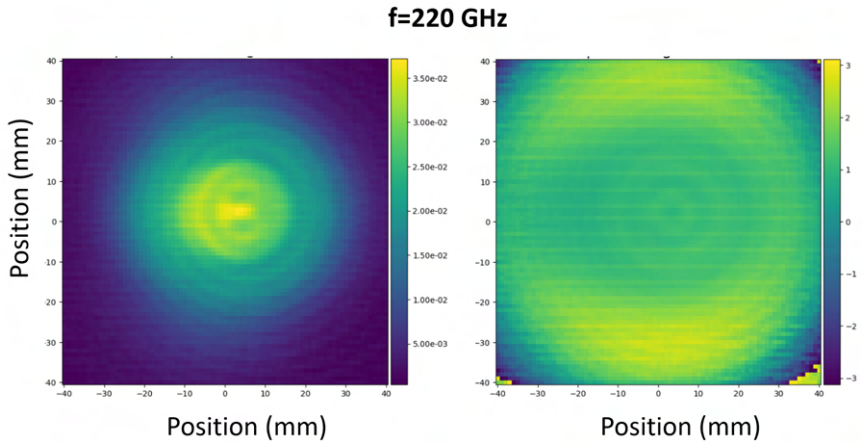
Before measurements, calibration is performed to align the VNAX transceiver module and the horn antenna with the OAP mirror by scanning the field without the hologram. The hologram is secured to its mount once the scanning plane shows sufficient planarity. The setup components and the probe antenna at the scanning area origin are aligned using a line laser. After calibration, the scanning plane area is defined as 0.2×0.2 m^2 with a sampling interval of 1 mm, resulting in 201×201 pixels.

Figure 5.1: Photography of the imaging setup at 220-330 GHz. 1) WR3.4-VNAX transceiver; 2) Pickett-Potter horn antenna; 3) OAP mirror; 4) holograms; 5) OEWG probe antenna connected with to WR3.4-VNAX receiver. 6) RAM

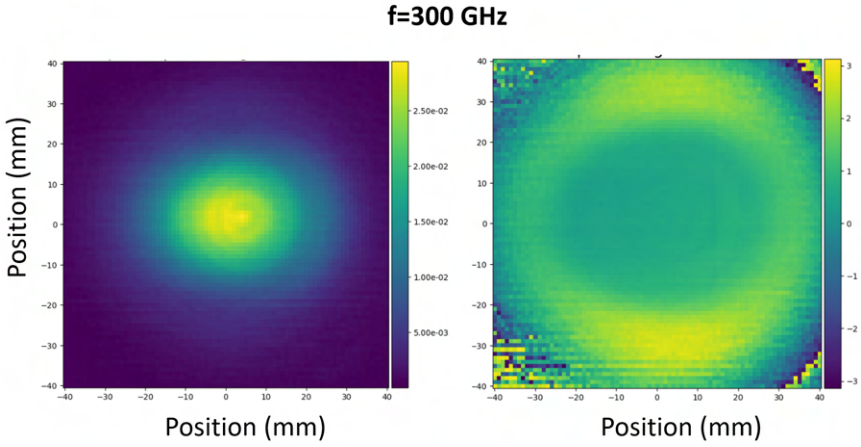
5.2 Measurements

5.2.1 Field pattern

In this section, the field measurements will be visualized and analyzed, including the collimated field at the aperture, the modulated fields after the holograms at a distance of 6 mm, and the fields at the target plane. It is noted that the distance between the hologram and the parabolic mirror was adjusted when scanning near the hologram and at the target plane due to the limitations of the optical table and scanning setup. The non-ideal effects of each component in this system, including the Gaussian beam source and holograms, will also be discussed.



a)



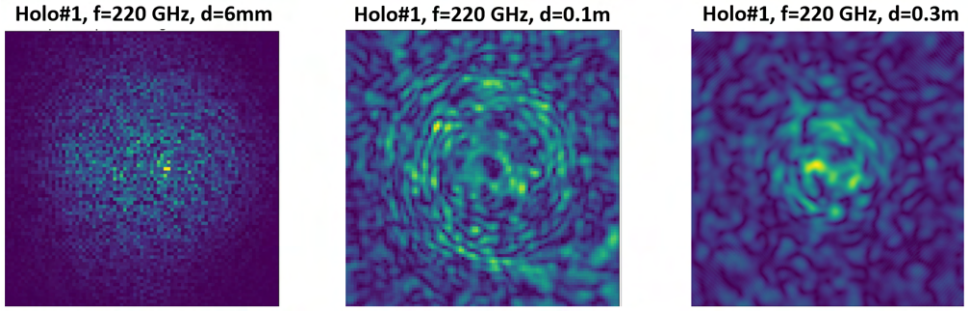
b)

Figure 5.2: Measured collimated fields at a) 220 GHz and b) 300 GHz.

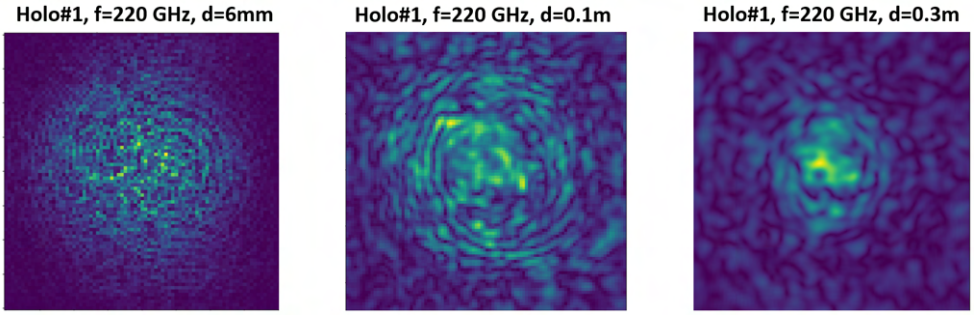
The collimated fields at the aperture plane were measured over an area of 80 × 80 mm^2 with a 1 mm step size, as shown in Fig. 5.2. It is evident that the Gaussian beams are collimated after the parabolic mirror across the operational frequency band, despite some calibration errors. Additionally, the beam waist radius in the x and y directions becomes unequal with increasing frequency, as observed in the amplitude

distributions of each field. To compare the simulated and realistic holograms and understand their interaction with the incident field, the fields near the hologram's surface have been measured. Meanwhile, the measured collimated fields are used to simulate transmission through the ideal (simulated) hologram. Using these fields near the hologram, the propagated fields at 0.1 m and 0.3 m can be calculated using the ASM and compared. Figs. 5.3 and 5.4 present the comparison for hologram #1 and #2 at 220 GHz and 300 GHz, respectively. It is evident that the fields at 6 mm from both the simulated and realistic holograms produce similar field distributions at both frequencies and distances. For example, circular curves around the field origin and one maximum point can be observed for hologram #1 at both frequencies with d=0.1 m and d=0.3 m. However, comparing the field patterns in Fig. 5.4, produced by hologram #2 with a 2 mm pixel size, the field patterns from hologram #1 with a 1 mm pixel size show more differences. This is because the detailed structures in hologram #1 are comparable to or even smaller than the band wavelengths (1.36 mm - 0.9 mm), making the scattering effect relatively significant along with the phase modulation. Overall, the similarity between the simulated and measured fields near the hologram and their ASM-computed field patterns at various distances indicates that the 3D printing technique can be used to manufacture high-quantized holograms with sufficient precision.

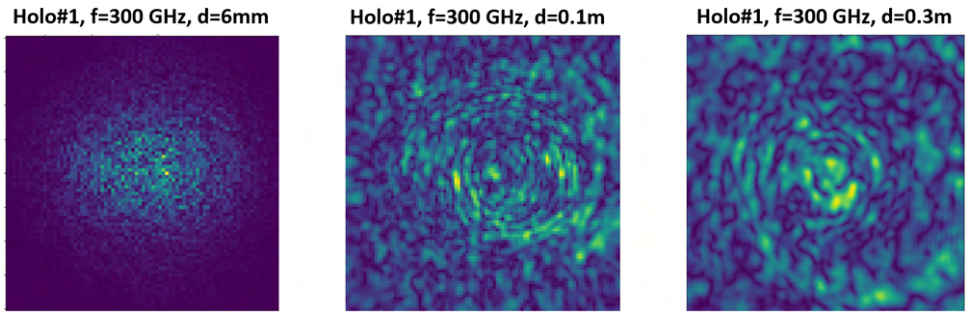
To explore the factors causing differences between the simulated and measured field patterns, we compare the simulated fields, ASM-computed fields from measured fields near realistic holograms, and measured fields in Fig. 5.5. It is important to note that the measured fields near realistic holograms and at the target plane are obtained from two configurations with the same components and calibration. The only difference is the distance between the mirror and the hologram, which the distance for measuring fields near the realistic hologram is shorter than that for measuring fields at the target plane. Repeated installations occur due to the limitations of the optical table and scanning setup in our lab, which prevent us from scanning the fields near the hologram and the target plane simultaneously. From Fig. 5.5 a)-b), similarities between the simulated fields and ASM-computed fields from measured fields near realistic holograms can be observed. For example, a maxima point around the field origin at 220 GHz and circular curves at the bottom edge at 300 GHz are distinguishable. Based on the discussion related to Fig. 5.3 and 5.4, the modulation introduced by the hologram is sufficiently accurate. Thus, a non-ideal Gaussian beam and misalignment between the antenna and parabolic mirror are the main factors causing the discrepancies between the left and middle plots in Fig. 5.5. This conclusion is further supported by comparing the middle and right plots. The difference between these two measurements is the distance between the parabolic mirror and the hologram. The effects of the non-ideal Gaussian beam and the non-planarity of the collimated field can be neglected at a shorter distance but become more pronounced at a longer distance. This makes the incident field on the hologram deviate from being regarded as a plane wave and could significantly affects the phase modulation of the designed hologram, leading to inconsistencies between the simulated and measured field patterns at the target plane.



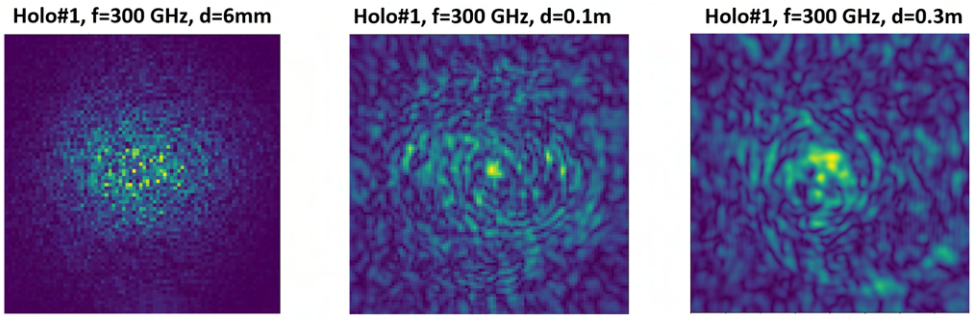
a) Measured collimated field with the simulated hologram #1.



b) Measured field near the hologram #1.

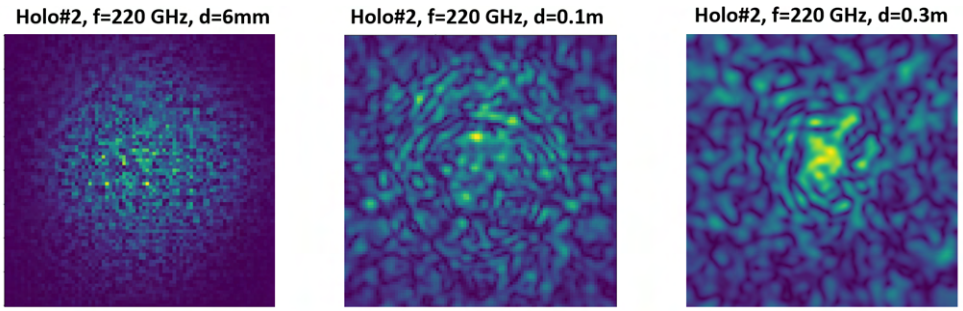


c) Measured collimated field with the simulated hologram #1.

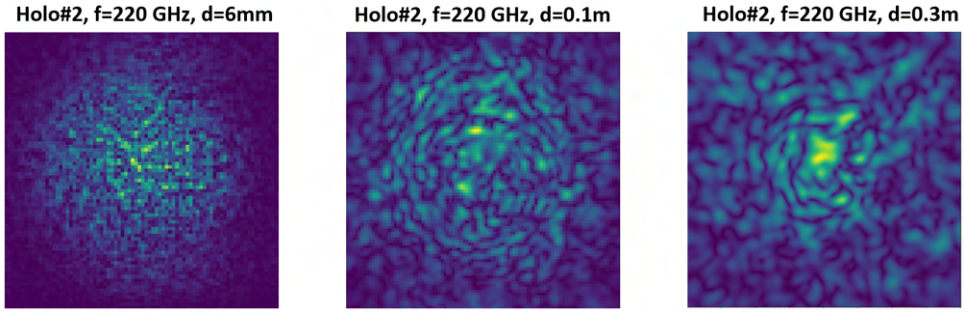


d) Measured field near the hologram #1.

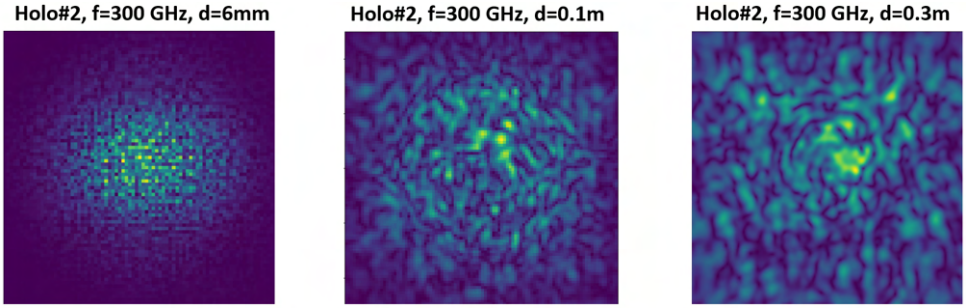
Figure 5.3: a) and c) present the amplitude patterns of the measured collimated fields transmitting through the simulated hologram #1 at 6 mm, 0.1 m, and 0.3 m, computed by ASM. b) and d) demonstrate the measured patterns near the realistic hologram #1 at 6 mm, with the field patterns at 0.1 m and 0.3 m computed by ASM. The FoVs from left to right are 80×80, 100×100, 200×200 mm^2, respectively.



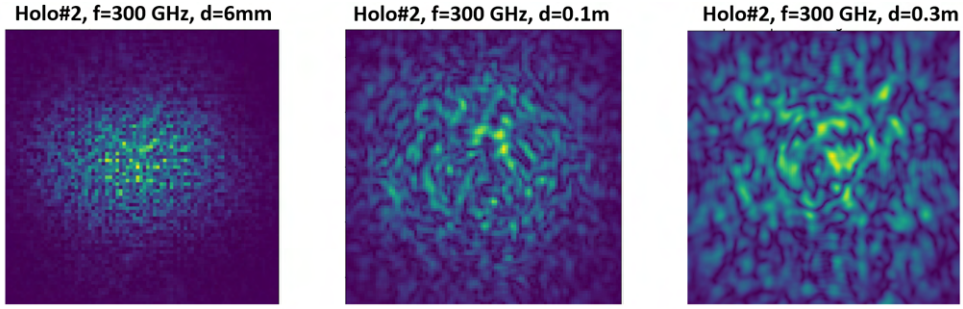
a) Measured collimated field with the simulated hologram #2.



b) Measured field near the hologram #2.

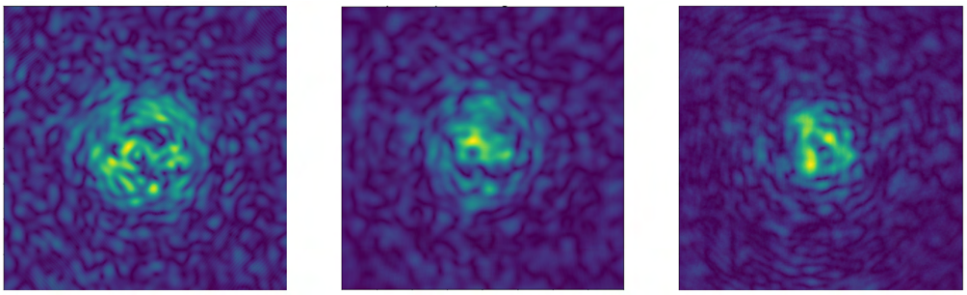


c) Measured collimated field with the simulated hologram #2.

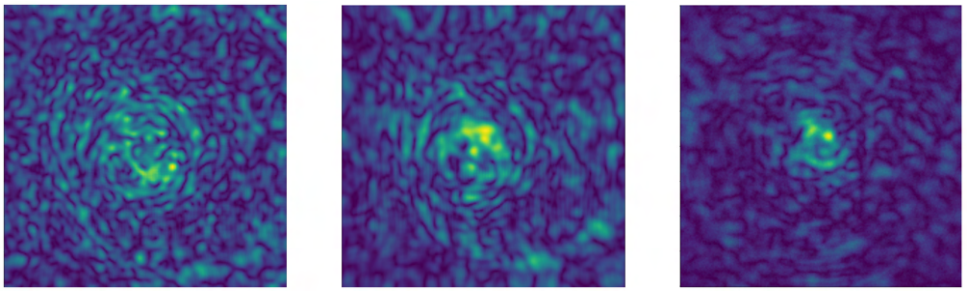


d) Measured field near the hologram #2.

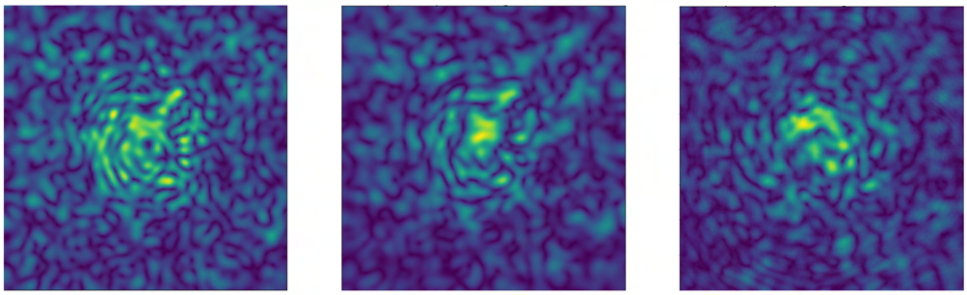
Figure 5.4: a) and c) present the amplitude patterns of the measured collimated fields transmitting through the simulated hologram #2 at 6 mm, 0.1 m, and 0.3 m, computed by ASM. b) and d) demonstrate the measured patterns near the realistic hologram #2 at 6 mm, with the field patterns at 0.1 m and 0.3 m computed by ASM. The FoVs from left to right are 80×80, 100×100, 200×200 mm^2, respectively.



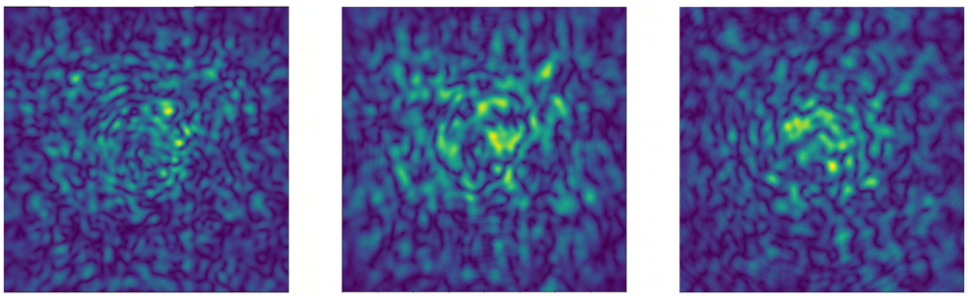
a) Hologram #1 at 220GHz



b) Hologram #1 at 300GHz



c) Hologram #2 at 220GHz



d) Hologram #2 at 300GHz

Figure 5.5: Comparison of simulated fields (left), ASM-computed fields from measured fields near realistic holograms (middle), and measured fields (right) at the target plane for hologram #1 and #2. The FoV is 200×200 mm^2.

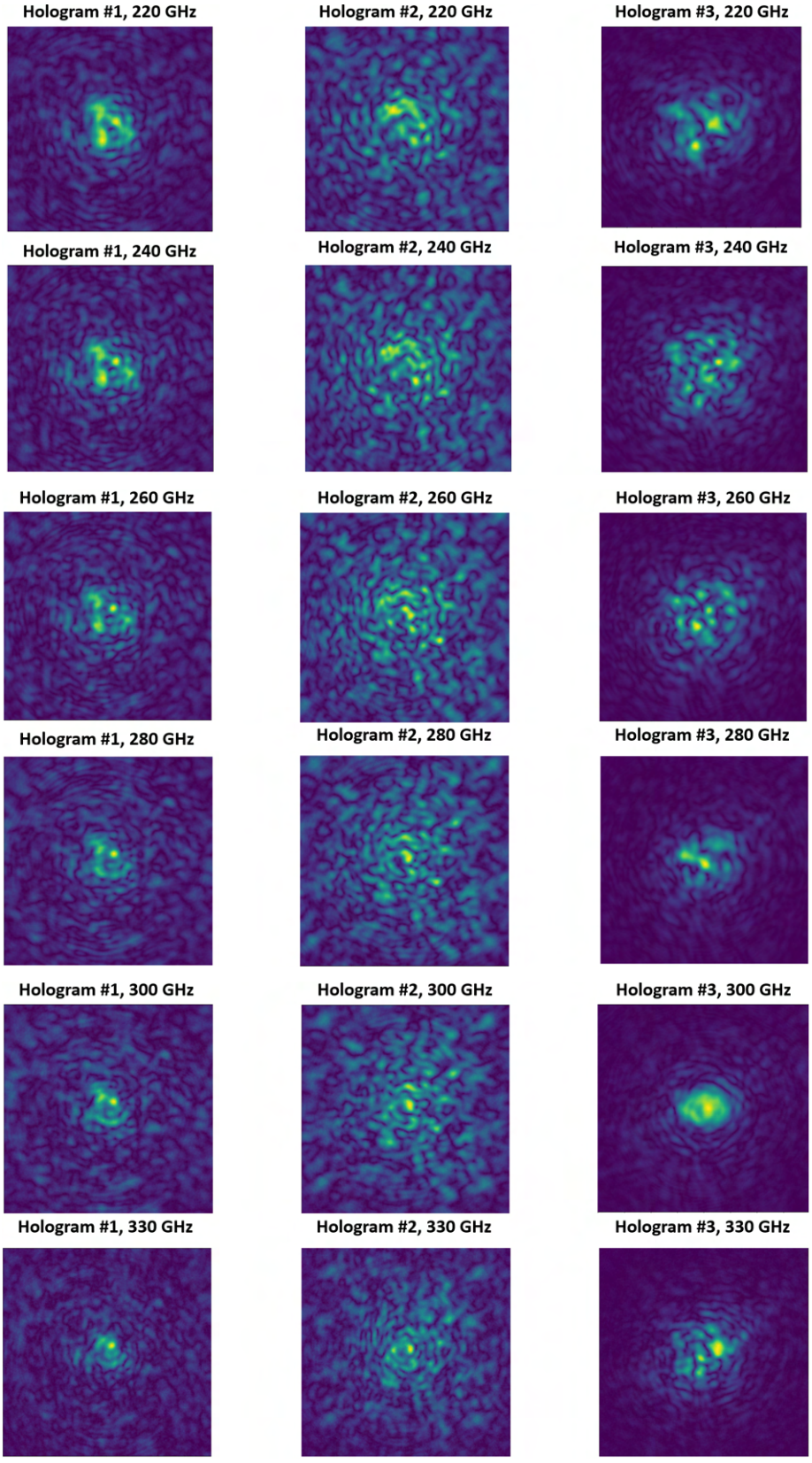


Figure 5.6: Measured electric field amplitude patterns from three designed holograms with different frequencies. Hologram #1 and #2 have 1 mm and 2 mm pixel size, respectively, and hologram #3 is from previous work.

5.2.2 Frequency diversity

Fig. 5.7 shows the normalized singular value spectra of all measured and simulated complex field patterns for the three holograms at 220-330 GHz, respectively. Similar trends from measurements and simulations can be observed, but the value at which the curves from measurements eventually flatten is higher than that from simulations. This indicates less correlation and better frequency diversity between frequency points in the measured data compared to the simulated data. One factor causing the higher flattening value in the measured data could be manufacturing and measurement uncertainties, such as non-ideal Gaussian beams and device-based phase and amplitude instability. However, compared to the simulation results, the normalized singular value of hologram #1 dropped slightly and is now close to that of hologram #3. In contrast, hologram #2 demonstrates the greatest singular value curve among them.

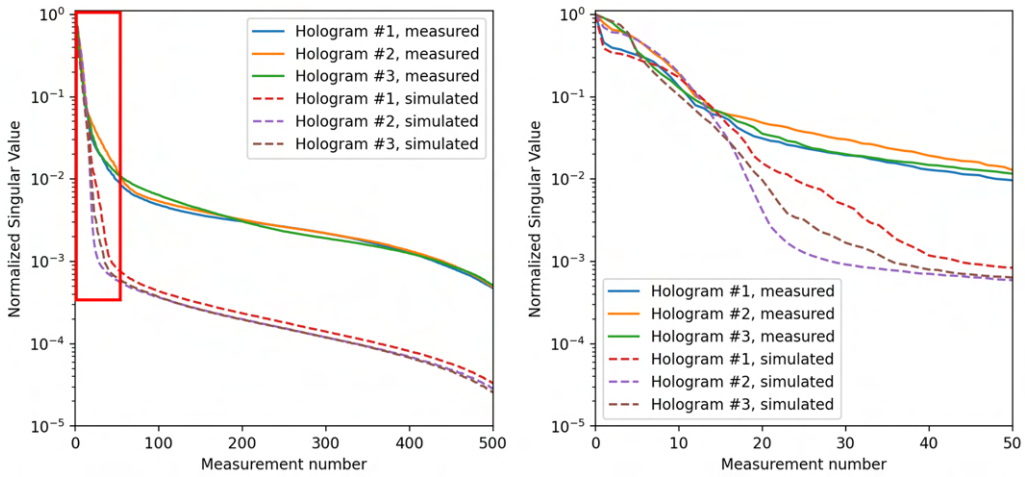
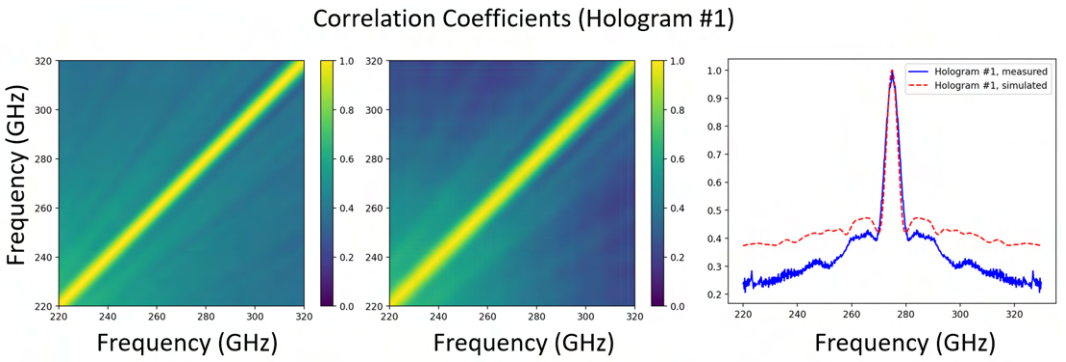
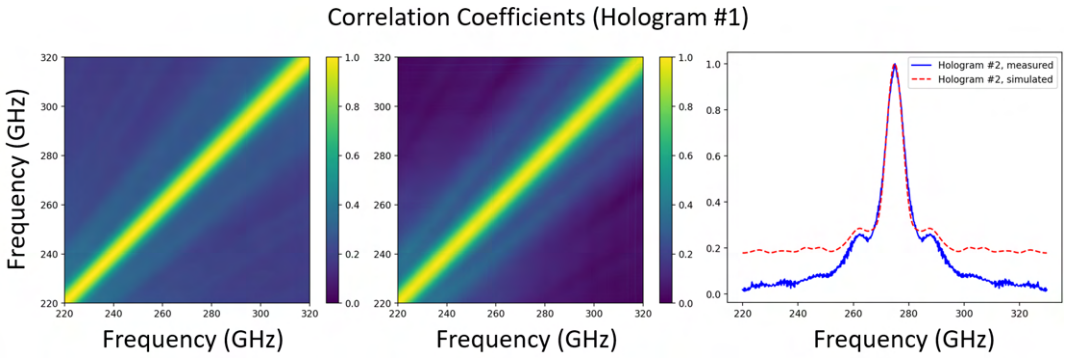


Figure 5.7: Left: Normalized singular value spectra of 500 measured field patterns for both holograms #1 and #2. Right: The majority of frequency points are cropped for clarity.

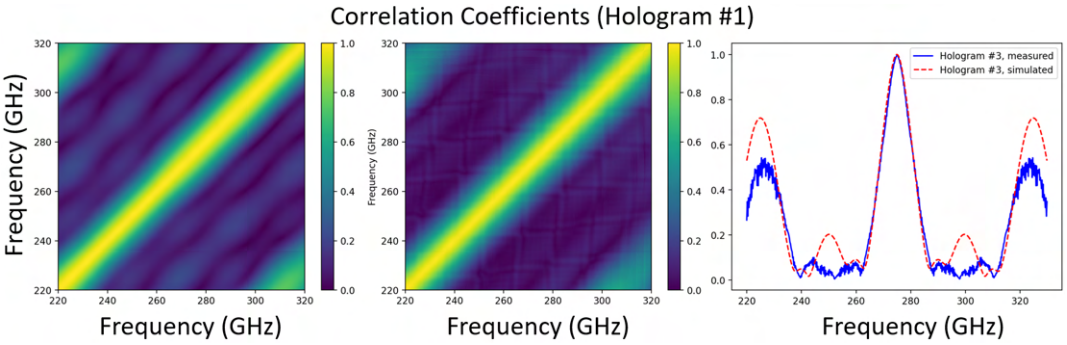
Additionally, the correlation coefficients for measured and simulated fields are shown in Fig. 5.8. The correlation coefficients demonstrate good agreement between the measured and simulated data along the yellow diagonal, while the values from the sidelobes exhibit a difference of ~ 0.2, maintaining a similar trend. The yellow diagonal of holograms #1 and #2, designed using the proposed method, is thinner than that of hologram #3. This indicates that the holograms designed in this thesis have lower correlation fields than the previous hologram. Moreover, the values for hologram #2 quickly drop below 0.2, showcasing its superiority over the others.



a) Hologram #1



b) Hologram #2



c) Hologram #3

Figure 5.8: Correlation coefficients of field patterns for holograms a) #1, b) #2, and c) #3 at 220-330 GHz are shown. From left to right, the simulated correlation coefficient matrix, measured correlation coefficient matrix, and cross sections are presented, respectively.

5.3 Validation of imaging system model

5.3.1 Signal extraction via time-gating

For the proposed imaging system, the horn antenna receives not only the signal reflected by the target but also multiple reflections between the hologram and the horn antenna. Therefore, a time-gating operation needs to be applied to the raw reflected signals to extract the desired signals. In Fig. 5.9, the raw S_{11} is first converted to a time-series signal using the inverse Fast Fourier Transform (iFFT). From the time domain signals in Fig. 5.9 c), two peaks (at around 3.34 ns and 6.68 ns) can be observed, corresponding to multiple reflections from the incident surface of the hologram. Given that the distance between the hologram and the imaging plane is $d = 0.3$m, the arrival time, which also represents the peak of the time-gating window, from the target can be computed as:

$$t_{\mathrm{p}} = 3.34\ \mathrm{ns} + \frac{2 \cdot \sqrt{d^2 + r_{\mathrm{p}}^2}}{c_0}, \tag{5.1}$$

where r_{p} is the distance of the target from the center of the imaging plane, and c_0 is the speed of light. The width of the window can be expressed as $[(1-\Delta)\cdot t_{\mathrm{p}}, (1+\Delta)\cdot t_{\mathrm{p}}]$ and Δ is set to 0.15 in this experiment to avoid the overlap of the second reflection. After the time gating, the time domain signals are converted back to S_{11} in the frequency domain using the Fast Fourier Transform (FFT) to obtain the reflected signals from the target.

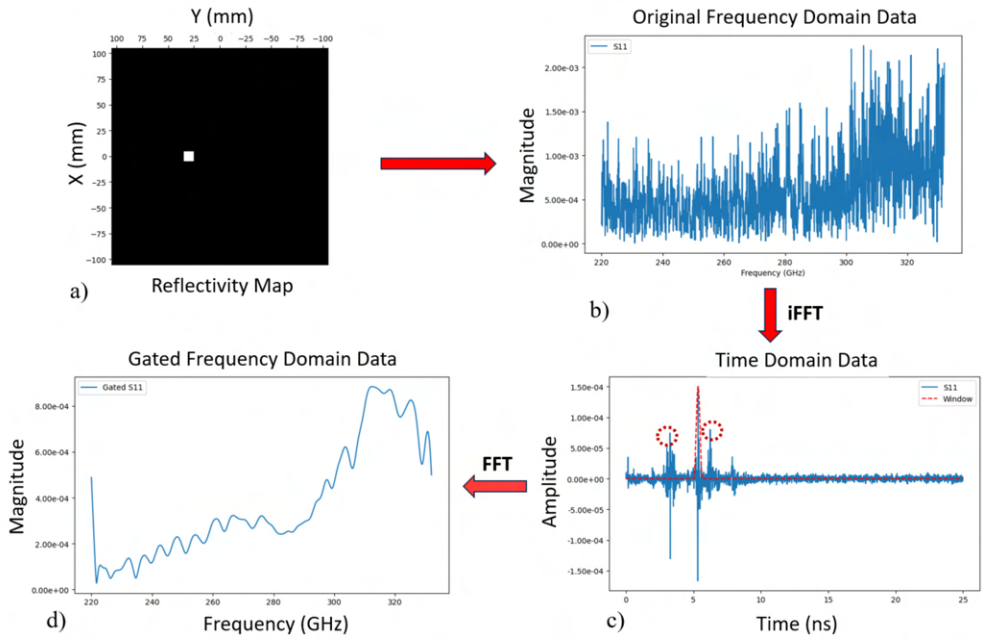


Figure 5.9: The process of time gating. The red dotted line presents the calculated time gating window, and two dotted circles denote the first and second reflections between the hologram and horn antenna.

5.3.2 Comparison of measured and simulated reflected signals

To verify the imaging model proposed in Section 3.1, we compared the directly measured reflected signals with the simulated ones computed by the constructed measurement matrix at the same position. In the experimental setup, a corner cube with a diameter of 12.7 mm and 10×30 mm^2 metallic patch are respectively fixed to a 2-D translation stage and automatically scanned over a 200×200 mm^2 imaging plane. The corner cube was used instead of a 10×10 mm^2 metallic patch because it reflects more energy, improving the signal-noise-ratio of the reflected signals. The receiver collects reflected signals from 21×21 different positions for both targets. To establish the measurement matrix $\mathbf{H}$, a near-field probe scans in a parallel plane close to the hologram aperture [7]. The aperture fields over the operation frequency are measured by this near-field scanning system, and the field distributions $\mathbf{E}_0$ at the imaging plane can be extrapolated based on ASM. Finally, the $\mathbf{H}$ matrix is calculated based on Eq. 3.16, and the corresponding simulated reflected signals are obtained by multiplying the measurement matrix by the reflectivity map, which matches the size and position of the target.

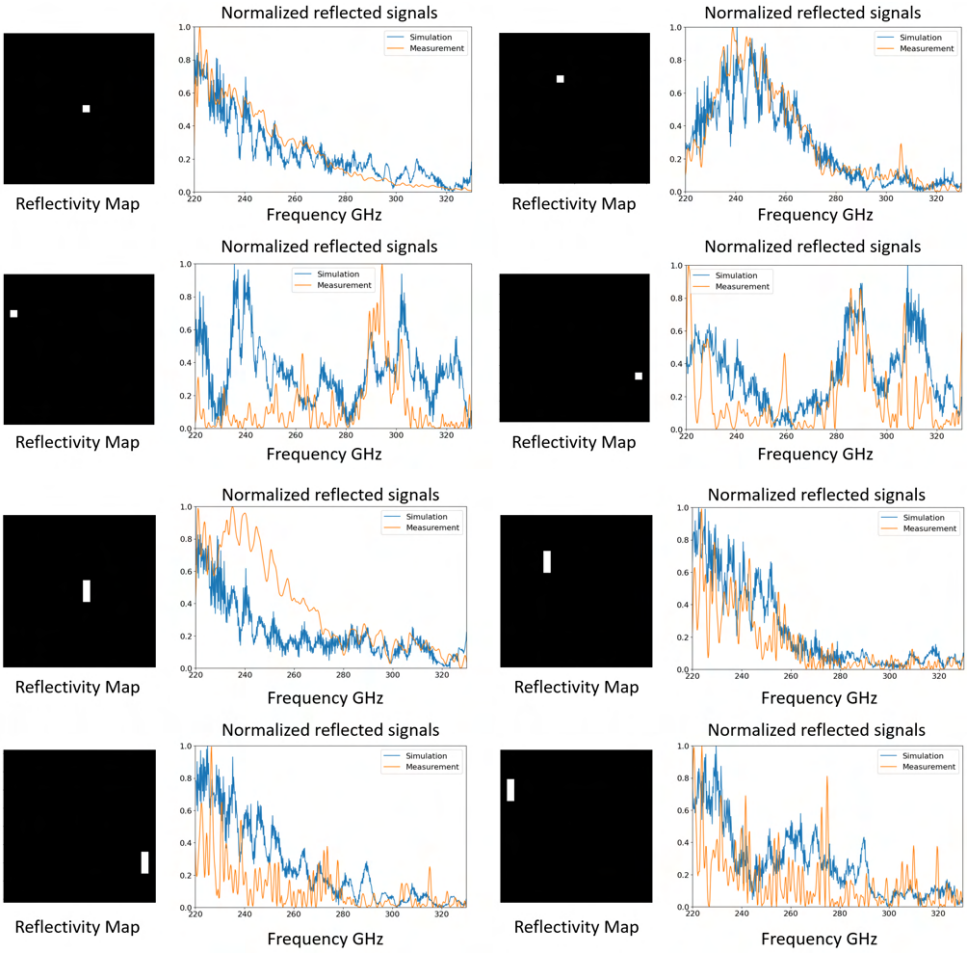


Figure 5.10: Simulated and measured S_{11} for hologram #1.

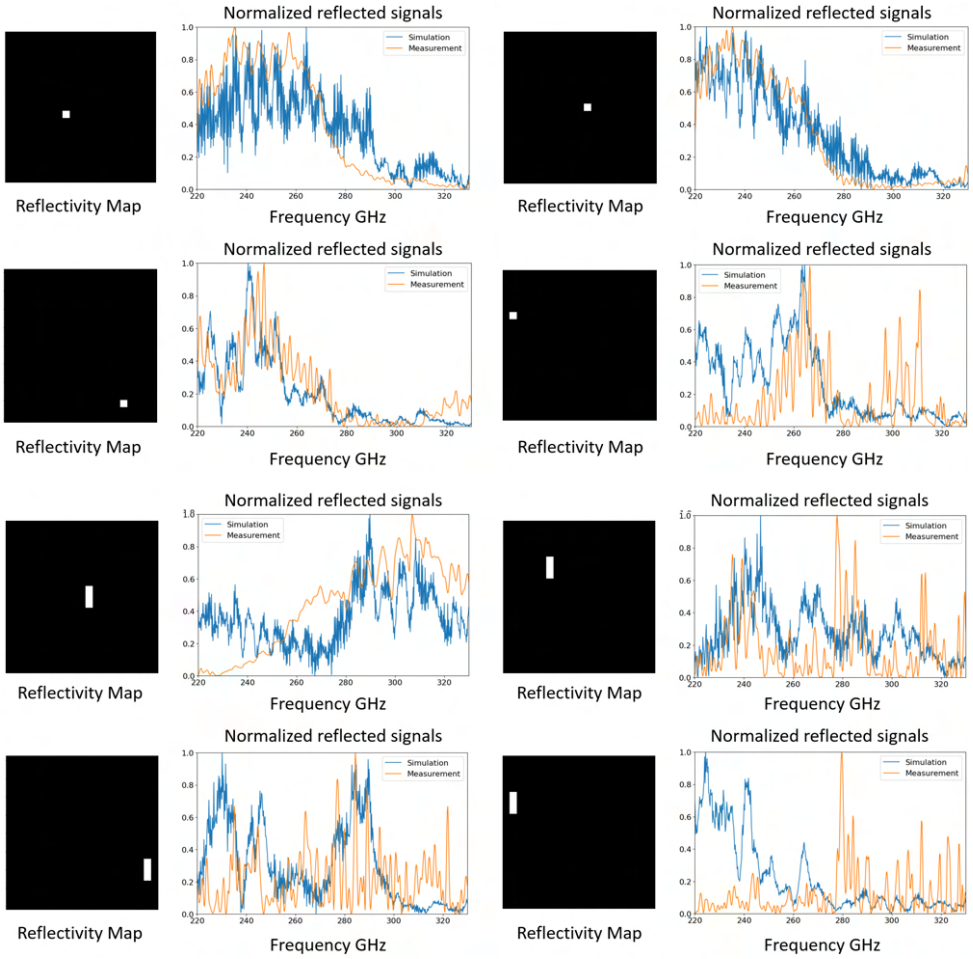


Figure 5.11: Simulated and measured S_{11} for hologram #2.

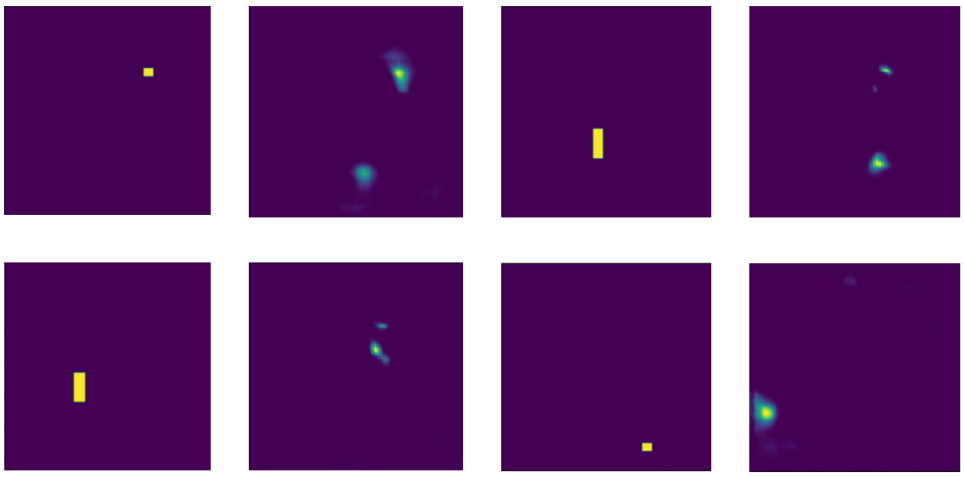
Fig.5.10 and 5.11 show the measured and simulated S_{11} for two designed holograms targeting the same object at the same position, respectively. From Fig. 5.10, it is evident that the measured and simulated S_{11} show a similar trend in the paraxial area for both the corner cube and the metallic patch. However, the discrepancy of the curves becomes significant for objects located at the edge of the field of view. Similar results can be observed from Fig. 5.11, with some exceptions. For example, the measured and simulated S_{11} show a similar trend for the corner cube located on the right-down side, while a significant mismatch is found for the metallic patch in the center area. The similarities in the paraxial area indicate that the proposed imaging model is valid under certain conditions. The simulated reflected signals, calculated by multiplying the measurement matrix and the object's reflectivity map, assume that all energy is collected by the receiver. In the paraxial area, the corner cube reflects all energy to the receiver, resulting in consistency between the measurements and simulations for both holograms. Although the metallic patch in the center area does not reflect all energy, most of it can still be received by the horn antenna, showcasing some similarity. However, as the object moves closer to the edge of the field of view, part of the energy is not reflected back to the antenna due to the large incident angle, resulting

in significant variance between the simulation and measurement. Additionally, the field distribution changes rapidly in adjacent pixels, so even minor mismatches in the position or size of the measured and simulated targets in the paraxial area can cause inconsistencies.

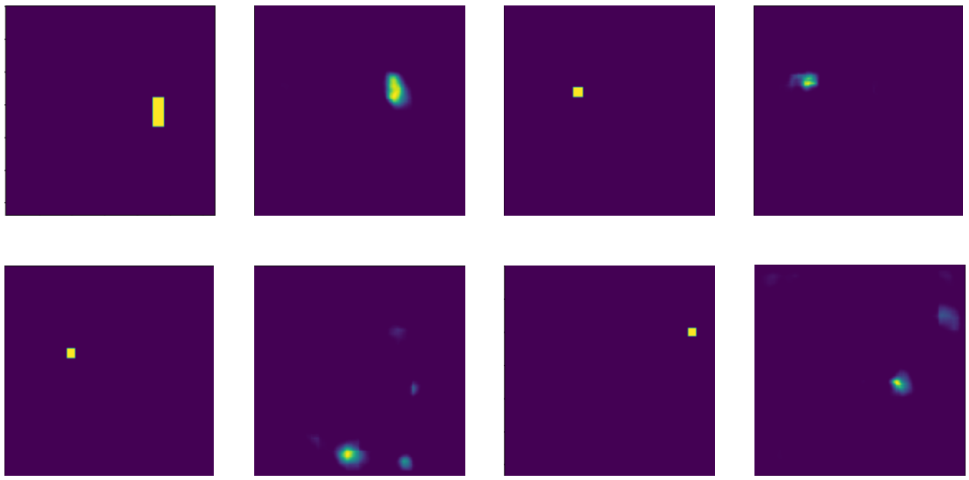
In conclusion, the proposed imaging system model demonstrates validity, with consistent results between measurements and simulations in the paraxial area. However, variations in energy reflection, especially at the edge of the field of view, and minor mismatches in target position or size can lead to significant discrepancies. Moreover, the diversity of measurements from different positions also indicates the potential imaging capability of this system.

5.3.3 Experimental imaging

One of the goals of this thesis is to reconstruct images from real measurements. However, the trained neural network discussed in Section 4.2 cannot be directly applied to real measurements due to the domain shift in measurement matrix between the simulated and real imaging systems [57]. Therefore, re-training the neural network is essential for the imaging experiments. Utilizing limited measured data pairs for training is not advisable due to the risk of over-fitting. As stated in Section 4.2.5, the neural network trained on the simulated measurement matrix and EMNIST dataset demonstrates generalization ability for some unseen targets. A method based on the constructed measurement matrix is proposed and validated in this section. The neural network will be trained with this constructed measurement matrix and an additional dataset containing targets of different sizes and shapes. It will then be tested on real measurement data shown in Fig. 5.12. It is evident that the position of some targets can be reconstructed from the 1D measurements by the neural network, which was completely trained on the synthetic dataset for both holograms. However, inaccurate predictions still exist due to the domain shift between the simulated and real measurement matrices, especially for some targets located at the edge of the field of view. Overall, the neural network trained on the synthetic dataset cannot accurately reconstruct both the shape and position of real targets from the real measurements. However, some good reconstruction results indicate that the simulated measurement matrix can partially represent the real physical process of the imaging system to some extent.



a) Reconstruction results from the hologram #1.



b) Reconstruction results from the hologram #2.

Figure 5.12: Reconstruction results from the real measurements.

6 Conclusions

6.1 Summary

This thesis proposes an end-to-end design paradigm for a sub-millimeter frequency-diverse computational imaging system, which jointly optimizes a heavily-quantized hologram as the encoder and a reconstruction network as the decoder. The approach starts by building an imaging physics model with a dispersive hologram, based on the Fourier-optics method, to map the reflectivity of an object to 1-D reflected frequency response. This imaging model is then integrated into a deep learning framework to optimize the profile of the dispersive hologram using neural networks in a data-driven manner. Additionally, a differentiable quantization method is presented for hologram design to reduce the mismatch between simulation and manufacturing.

The main research questions of this thesis were to explore 1) whether the end-to-end design method improves the frequency diversity of the hologram in the 220-330 GHz range, 2) whether the manufactured holograms with detailed structures can produce the desired field distributions as shown in simulations, and 3) the correctness of the proposed imaging physics model. A Pytorch-based Fourier-optics simulation codebase was developed and used to optimize both hologram surface patterns and neural networks with different design parameters. This codebase enables multi-wavelength parallel simulation with GPU acceleration, significantly improving efficiency compared to the original physical-optics simulation.

In the simulation, the effect of hologram design parameters, including minimal detail size and the number of discrete levels, on frequency diversity, efficiency, and reconstruction quality across the 220-330 GHz range was studied. Based on the simulation results, end-to-end designed holograms demonstrate better performance in frequency diversity and reconstruction quality compared to previous methods. However, the efficiency of these holograms decreased due to significant diffraction effects introduced by the more detailed structures. Finally, two hologram designs with minimal detail sizes of 1×1 mm^2 and 2×2 mm^2 were selected for manufacturing using Formlabs' 3D printing machine.

To validate the research questions proposed in this thesis, a quasi-optical setup was built to measure the electric field and reflected frequency response of both holograms at 220-330 GHz. The frequency diversity analysis of the measured fields shows a better performance of end-to-end designed holograms. However, different design paradigms did not significantly improve the flatness of the singular value spectrum, indicating a possible upper-bound limitation of phase modulation for frequency diversity. In addition, the measured field patterns did not match the simulated patterns in the first measurement. This discrepancy is explored and attributed to the non-planar phase of the collimated incident wave. The designed holograms, with their detailed structures, function as phase modulators and are sensitive to the phase of the incident wave. To study how waves interact with manufactured holograms, considering the non-ideal incident wave, the aperture fields with and without holograms were scanned and analyzed. The sufficient similarities verify the accuracy of the hologram modeling. Moreover, a scattering effect is observed from the hologram with a minimal size of

1 mm due to its comparable size with the wavelength. Finally, the simulated and measured reflected frequency responses are compared and some consistent results indicate the partial validity of imaging physics modeling in the paraxial condition. The imaging experiments also demonstrate that the neural network fully trained on the synthetic dataset with the simulated measurement matrix cannot accurately predict both the shape and position of targets from real measurements. However, some good reconstruction results indicate that the simulated measurement matrix can partially represent the real physical process of the imaging system to some extent.

6.2 Limitation and future work

The practical imaging performance of computational imaging systems is greatly limited by the accuracy of the forward physics model and the reliability of the inverse reconstruction algorithm. In this thesis, the imaging physics model does not fully represent the imaging process of a realistic system. Consequently, the optimized neural network cannot be directly used to reconstruct images from real measurements. Additionally, due to the simplification of our imaging model, the end-to-end optimization method may not identify the most optimal hologram pattern for the corresponding neural network. In the future, more accurate imaging models and the uncertainty of the imaging process should be explored to better match simulated and measured results and enhance the potential of end-to-end design.

References

[1] B. Gonzalez-Valdes, Y. Álvarez, Y. Rodriguez-Vaqueiro, A. Arboleya-Arboleya, A. García-Pino, C. M. Rappaport, F. Las-Heras, and J. A. Martinez-Lorenzo, "Millimeter wave imaging architecture for on-the-move whole body imaging," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 6, pp. 2328–2338, 2016.

[2] Y. Wang and A. E. Fathy, "Advanced system level simulation platform for three-dimensional uwb through-wall imaging sar using time-domain approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 5, pp. 1986–2000, 2011.

[3] S. Kharkovsky and R. Zoughi, "Microwave and millimeter wave nondestructive testing and evaluation-overview and recent advances," *IEEE Instrumentation & Measurement Magazine*, vol. 10, no. 2, pp. 26–38, 2007.

[4] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE microwave magazine*, vol. 12, no. 7, pp. 78–94, 2011.

[5] X. Gu, A. Valdes-Garcia, A. Natarajan, B. Sadhu, D. Liu, and S. K. Reynolds, "W-band scalable phased arrays for imaging and communications," *IEEE communications magazine*, vol. 53, no. 4, pp. 196–204, 2015.

[6] A. Tamminen, S.-V. Pälli, J. Ala-Laurinaho, and Z. D. Taylor, "Millimeter- and submillimeter-wave imaging through dispersive hologram and deep neural networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 6, pp. 3281–3290, 2022.

[7] Q. Zhang, Q. Xi, C. Zhu, B. Zhang, J. Wang, H. Yu, Y. Li, Z. Zhu, X. Li, and L. Ran, "Microwave imaging enhanced by duplexed random phase modulations," *IEEE Transactions on Microwave Theory and Techniques*, 2023.

[8] H. Fu, F. Dai, and L. Hong, "Metasurface aperture design for far-field computational microwave imaging beyond rayleigh diffraction limitations," *IEEE Transactions on Microwave Theory and Techniques*, 2023.

[9] T. Fromenteze, O. Yurduseven, M. F. Imani, J. Gollub, C. Decroze, D. Carsenat, and D. R. Smith, "Computational imaging using a mode-mixing cavity at microwave frequencies," *Applied Physics Letters*, vol. 106, no. 19, 2015.

[10] Q. Zhang, Q. Xi, C. Ma, B. Zhang, T. Zhou, Y. Li, Z. Zhu, X. Li, and L. Ran, "A microwave imager implemented with a space-fed randomized aperture," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 9, pp. 4180–4190, 2021.

[11] V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, "End-to-end optimization of optics and image processing

for achromatic extended depth of field and super-resolution imaging," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 1–13, 2018.

[12] B. Zhang, X. Yuan, C. Deng, Z. Zhang, J. Suo, and Q. Dai, "End-to-end snapshot compressed super-resolution imaging with deep optics," *Optica*, vol. 9, no. 4, pp. 451–454, 2022.

[13] D. S. Jeon, S.-H. Baek, S. Yi, Q. Fu, X. Dun, W. Heidrich, and M. H. Kim, "Compact snapshot hyperspectral imaging with diffracted rotation," 2019.

[14] L. Li, L. Wang, W. Song, L. Zhang, Z. Xiong, and H. Huang, "Quantization-aware deep optics for diffractive snapshot hyperspectral imaging," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 19 780–19 789.

[15] Q. Sun, E. Tseng, Q. Fu, W. Heidrich, and F. Heide, "Learning rank-1 diffractive optics for single-shot high dynamic range imaging," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1386–1396.

[16] S.-V. Pälli, "Distributed beamsteering at millimeter waves for imaging and localization," Master's thesis, 2020.

[17] S.-V. Pälli, A. Tamminen, J. Ala-Laurinaho, and Z. D. Taylor, "Design and characterization of phase holograms for standoff localization at millimeter and submillimeter waves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 1, pp. 907–918, 2021.

[18] J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Advances in Optics and Photonics*, vol. 10, no. 2, pp. 409–483, 2018.

[19] E. H. Adelson, J. R. Bergen *et al.*, "The plenoptic function and the elements of early vision," *Computational models of visual processing*, vol. 1, no. 2, pp. 3–20, 1991.

[20] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2013.

[21] L. Wang, T. Zhang, Y. Fu, and H. Huang, "Hyperreconnet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2257–2270, 2018.

[22] M. Li, A. Mathai, L. Yandi, Q. Chen, X. Wang, and X. Xu, "A brief review on 2d and 3d image reconstruction using single-pixel imaging," *Laser Physics*, vol. 30, no. 9, p. 095204, 2020.

[23] G. Lipworth, A. Mrozack, J. Hunt, D. L. Marks, T. Driscoll, D. Brady, and D. R. Smith, "Metamaterial apertures for coherent computational imaging on the physical layer," *JOSA A*, vol. 30, no. 8, pp. 1603–1612, 2013.

[24] A. D. Yaghjian and S. R. Best, "Impedance, bandwidth, and q of antennas," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 4, pp. 1298–1324, 2005.

[25] M. F Imani, T. Sleasman, J. N. Gollub, and D. R. Smith, "Analytical modeling of printed metasurface cavities for computational imaging," *Journal of Applied Physics*, vol. 120, no. 14, 2016.

[26] O. Yurduseven, V. R. Gowda, J. N. Gollub, and D. R. Smith, "Multistatic microwave imaging with arrays of planar cavities," *IET Microwaves, Antennas & Propagation*, vol. 10, no. 11, pp. 1174–1181, 2016.

[27] ——, "Printed aperiodic cavity for computational and microwave imaging," *IEEE Microwave and Wireless Components Letters*, vol. 26, no. 5, pp. 367–369, 2016.

[28] J. Salo, J. Meltaus, E. Noponen, M. M. Salomaa, A. Lönnqvist, T. Koskinen, V. Viikari, J. Säily, J. Häkli, J. Ala-Laurinaho *et al.*, "Holograms for shaping radio-wave fields," *Journal of Optics A: Pure and Applied Optics*, vol. 4, no. 5, p. S161, 2002.

[29] S. Banerji and B. Sensale-Rodriguez, "A computational design framework for efficient, fabrication error-tolerant, planar thz diffractive optical elements," *Scientific reports*, vol. 9, no. 1, p. 5801, 2019.

[30] P. Komorowski, P. Czerwińska, M. Surma, P. Zagrajek, R. Piramidowicz, and A. Siemion, "Three-focal-spot terahertz diffractive optical element-iterative design and neural network approach," *Optics Express*, vol. 29, no. 7, pp. 11 243–11 253, 2021.

[31] S. Banerji, Y. Shi, V. S.-E. Su, U. Ghosh, J. Cooke, Y. L. Kong, L. Liu, and B. Sensale-Rodriguez, "Inverse designed thz spectral splitters," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 5, pp. 425–428, 2021.

[32] J. Meltaus, J. Salo, E. Noponen, M. M. Salomaa, V. Viikari, A. Lonnqvist, T. Koskinen, J. Saily, J. Hakli, J. Ala-Laurinaho *et al.*, "Millimeter-wave beam shaping using holograms," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 4, pp. 1274–1280, 2003.

[33] N. Chen, L. Cao, T.-C. Poon, B. Lee, and E. Y. Lam, "Differentiable imaging: A new tool for computational optical imaging," *Advanced Physics Research*, vol. 2, no. 6, p. 2200118, 2023.

[34] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: a survey," *Journal of Marchine Learning Research*, vol. 18, pp. 1–43, 2018.

[35] L. B. Rall and G. F. Corliss, "An introduction to automatic differentiation," *Computational Differentiation: Techniques, Applications, and Tools*, vol. 89, pp. 1–18, 1996.

[36] M. Bücker, *Automatic differentiation: applications, theory, and implementations*. Springer, 2006.

[37] Y. Peng, S. Choi, N. Padmanaban, and G. Wetzstein, "Neural holography with camera-in-the-loop training," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 6, pp. 1–14, 2020.

[38] C. Wang, N. Chen, and W. Heidrich, "Towards self-calibrated lens metrology by differentiable refractive deflectometry," *Optics Express*, vol. 29, no. 19, pp. 30 284–30 295, 2021.

[39] R. He, W.-S. Zheng, T. Tan, and Z. Sun, "Half-quadratic-based iterative minimization for robust sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 2, pp. 261–275, 2013.

[40] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[41] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[44] Y. Zhu and S. Newsam, "Densenet for dense flow," in *2017 IEEE international conference on image processing (ICIP)*. IEEE, 2017, pp. 790–794.

[45] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, 2017.

[46] J. Zhang and B. Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1828–1837.

[47] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 3, pp. 521–538, 2018.

[48] W. Feng, S. Zhou, S. Li, Y. Yi, and Z. Zhai, "High-turbidity underwater active single-pixel imaging based on generative adversarial networks with double attention u-net under low sampling rate," *Optics Communications*, vol. 538, p. 129470, 2023.

[49] F. G. Paul, "Quasioptical systems: Gaussian beam quasioptical propogation and applications," 1998.

[50] Z.-G. Liu and M. Mattina, "Learning low-precision neural networks without straight-through estimator (ste)," *arXiv preprint arXiv:1903.01061*, 2019.

[51] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," *arXiv preprint arXiv:1611.01144*, 2016.

[52] S. Choi, M. Gopakumar, Y. Peng, J. Kim, M. O'Toole, and G. Wetzstein, "Time-multiplexed neural holography: a flexible framework for holographic near-eye displays with fast heavily-quantized spatial light modulators," in *ACM SIGGRAPH 2022 Conference Proceedings*, 2022, pp. 1–9.

[53] Z. Wang, T. Popordanoska, J. Bertels, R. Lemmens, and M. B. Blaschko, "Dice semimetric losses: Optimizing the dice score with soft labels," *arXiv preprint arXiv:2303.16296*, 2023.

[54] W. Zhang, H. Zhang, C. J. Sheppard, and G. Jin, "Analysis of numerical diffraction calculation methods: from the perspective of phase space optics and the sampling theorem," *JOSA A*, vol. 37, no. 11, pp. 1748–1766, 2020.

[55] D. Liao, M. Wang, K. F. Chan, C. H. Chan, and H. Wang, "A deep-learning enabled discrete dielectric lens antenna for terahertz reconfigurable holographic imaging," *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 4, pp. 823–827, 2022.

[56] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[57] O. Kilim, A. Olar, T. Joó, T. Palicz, P. Pollner, and I. Csabai, "Physical imaging parameter variation drives domain shift," *Scientific Reports*, vol. 12, no. 1, p. 21302, 2022.