

Near Field Holography Using Neural Radiance Fields

JG

May 6, 2025

Abstract

Phase retrieval in near-field holography is a non-linear ill-posed problem as it requires to reconstruct phase and absorption from the measured squared magnitude. Commonly, all projections are processed independently and then a tomographic algorithm is employed to reconstruct the three dimensional volume. Within this paper we present a novel approach to directly solving the 3D tomographic problem from the holographic measurements. Therefore, we are using neural radiance fields, which have shown promising results for CT reconstructions. End-to-end learning, i.e. direct estimation of the object, leads to an inherent regularization by three dimensional consistency since the all estimated phase reconstructions are now coupled.

1 Introduction

To analyze weakly absorbing materials, such as soft-tissues or objects in nanometer scale, with X-ray tomography over the last decades phase contrast methods have been developed. This technique measures the overlay of the probe and the interaction with the object, which leads to small absorptions and strong phase shifts. Since the sensors are only able to measure the squared amplitude of the signal, we have to reconstruct the both, the phase shift and the amplitude from a single measurement. To perform phase shift tomography the common approach is to first reconstruct the projections for all angles and secondly applying tomographic algorithms, such as filtered back-projection (FBP) or simultaneous algebraic reconstruction technique (SART).

2 Method

2.1 Fresnel Propagation

The interaction of an object with X-rays are described by the complex refractive index

$$n(x, y, z) = 1 - \beta(x, y, z) + i\delta(x, y, z), \quad (1)$$

where $(x, y, z) \in \mathbb{R}^3$ denotes the spatial coordinate, $\beta \in \mathbb{R}$ the absorption, $\delta \in \mathbb{R}$ the phase shift, and i the imaginary unit. Assuming that the thin object approximation holds, i.e. only a single scattering event occurs per X-ray (? to verify), the interaction of the X-rays with the object along a line

$$r = x \cos(\theta) + y \sin(\theta) \quad (2)$$

can be described as the integral

$$\Psi_{\text{exit}}(r, z) = \exp \left(-ik \int_0^d \int_0^d (\delta(x, y, z) + \beta(x, y, z)) \chi(x \cos(\theta) + y \sin(\theta) - r) dx dy \right) \Psi_0, \quad (3)$$

where χ describes the dirac-delta function and Ψ_0 denotes a monochromatic wavefield, which is called probe in the following. After the projection approximation the exit wave (?) Ψ_{exit} is propagated following the fresnel progation regime as follows

$$\mathcal{I}_{\text{det}} = \left| \mathcal{F}^{-1} \circ \exp \left(-i\pi \frac{k_x^2 + k_y^2}{\text{Fr}} \right) \circ \mathcal{F} \circ \Psi_{\text{exit}} \right|^2, \quad (4)$$

\mathcal{F} describes the fourier transform, \mathcal{F}^{-1} its inverse, (k_x, k_y) are the coordinates in frequency space, and Fr is the Fresnel number. The Fresnel number is defined as

$$\text{Fr} = \frac{\Delta x^2}{\lambda z_{12}} \quad (5)$$

for parallel beam geometry, where Δx is the pixel size, λ is the wavelength, and z_{12} is the distance from the object to the detector. Given a set of measurements \mathcal{I}_{θ_i} along angles $\theta_i \in [0, \pi)$, $i \in \{1, \dots, n\}$, $n \in \mathbb{N}$ the objective is to find

$$\arg \min_{\beta, \delta} \sum_i \|\mathcal{I}_{\theta_i} - \mathcal{I}_{\text{det}_{\theta_i}}\|_2^2. \quad (6)$$

2.2 Neural Refractive Fields

Recent advancements in computer vision have shown that the use of Implicit Neural Representations leads to state-of-the-art volumetric reconstructions from a set of two dimensional images of the scene. INR employs a neural network, often a simple MLP, to predict opacity and color values for each spatial position. While earlier approaches were cumbersome to train and missed high spatial frequencies, recently proposed hash-encoding resolved both problems. They have also been introduced in the medical domain and have shown promissing results in computer tomography reconstruction, specifically in sparse angle tomography.

The proposed model uses a learnable hash encoding and afterwards 4 MLP layers with 64 neurons each, and a skip connection between the encoder output and the center MLP layer. The networks output is two dimensional, where the first value is strictly positive and represents the absorption, and the second value is strictly negative representing the phase shift. Compared to the voxel count of the three dimensional volume this approach has less learnable parameters, since we assume that the measured object shows some lower dimensional structures. (? das hier ausführen) A schematic representation of the model is depicted in Figure ???). Given an angle θ_i , we are sampling a ray r_j for each pixel on the detector and sample each ray by n evenly spaced points. Each point is now feeded into the neural network, then the ray integral is computed and the fresnel propagation is applied.

wavelength
and k rela-
tion! And
the cone
beam and
parallel
beam stuff
has to be
defined.