EC522 Computational Optical Imaging
Homework No. 2

Shuyue Jia
BUID: U62343813

February 2024

1 Problem 1: Example of geometric optics model based LSI imaging – Defocus and Depth of focus in Imaging

Depth of focus (DOF) is an important imaging metric, which measures the ability to image or filter out the objects that are away from the focal plane. This problem will explore this concept and its quantification from the computational standpoint.

In an ideal camera, the defocus effect can be approximated by an LSI model governed by the convolution model. As such, the output image g captured by the camera is related to the input object's intensity distribution f by

$$g(x, y) = f(x, y) * h(x, y),$$

where $*$ denote the convolution, the defocus point-spread function (PSF) h is characterized by the "circle of confusion" that can be modeled using a geometric optics model, as illustrated in Fig. 1. A point source on the left is imaged onto the camera sensor on the right by a lens. Depending on the axial location of the point source, it can either form a "sharp" point image on the camera or a "defocused" image that resembles a circle – hence the term "circle of confusion".

For simplicity, one can assume a point source can always be focused to a point image by a lens (*i.e.* a focus). However, the camera may not be placed at the correct plane, which results in "defocus". As seen in the geometric relation depicted in Fig. 1, the larger the camera's displacement z from the actual focus is, the larger the defocus circular PSF is, corresponding to more severe blur in the captured image.

Based on the geometrical relation illustrated in Fig. 1, the size of the defocus PSF d can be approximated by

$$d = \theta z,$$

where z is the defocus distance, and θ is the angle of acceptance of the lens, which is an important quantity of an imaging system/camera, and is related to the lens size and the focal length. Often times, the measure of θ is by the numerical aperture (NA), NA $= \sin(\theta/2)$, or the f-number $f/\# = 1/\theta$ of the camera lens.

In the following, we will formulate the forward model of this imaging problem using the tools learned in the lectures.

1.1 (a) Forward model of linear shift invariant (LSI) system in its operator form

1.1.1 (1) Consider a 1D system in which the defocus PSF becomes a "line of confusion". Construct the forward model that relates the output intensity image g with the input object f given the camera's displacement z, and the angle of acceptance θ. (Hint: 1D line-shape PSF can be modeled as a rectangular function.)

In a 1D linear shift-invariant (LSI) system, the forward model can be represented using convolution. In this case, where the defocus PSF becomes a "line of confusion", we can model it as a rectangular function. Let's denote the width of the rectangular function representing the line of confusion as d, the camera's displacement as z, and the angle of acceptance as θ.

The forward model in operator form can be expressed as follows:

$$g(x) = f(x) * h(x),$$

where:
$g(x)$ is the output intensity image.
$f(x)$ is the input object intensity distribution.
$h(x)$ is the line of confusion point spread function (PSF).

The line of confusion PSF, $h(x)$, can be modeled as a rectangular function with width d. Thus, $h(x)$ can be defined as:

$$h(x) = \begin{cases} 1/d & \text{for } -d/2 \leq x \leq d/2, \\ 0 & \text{otherwise.} \end{cases}$$

Substituting this PSF into the convolution equation, we get:

$$g(x) = \frac{1}{d} \int_{-\infty}^{\infty} f(t) \cdot \mathrm{rect}\left(\frac{x-t}{d}\right) dt,$$

where $\mathrm{rect}(x)$ is the rectangular function, defined as:

$$\mathrm{rect}(x) = \begin{cases} 1 & \text{for } |x| \leq 0.5, \\ 0 & \text{otherwise.} \end{cases}$$

This integral represents the convolution operation between the input object intensity distribution $f(x)$ and the line of confusion PSF $h(x)$, resulting in the output intensity image $g(x)$.

So, in summary, the forward model in operator form for the 1D LSI system with a line of confusion PSF is:

$$g(x) = \frac{1}{d} \int_{-\infty}^{\infty} f(t) \cdot \mathrm{rect}\left(\frac{x-t}{d}\right) dt.$$

1.1.2 Find the range space, null space, as well as the spectral representation of the forward operator, the adjoint operator.

To find the range space, null space, and the spectral representation of the forward operator and the adjoint operator, let's first define the operators explicitly.

Given the forward operator A, which represents the convolution operation with the line of confusion PSF $h(x)$, and the input signal $f(x)$, we have:

$$g(x) = \frac{1}{d} \int_{-\infty}^{\infty} f(t) \cdot \mathrm{rect}\left(\frac{x-t}{d}\right) dt.$$

To find the range space and null space of the forward operator A, we need to determine which signals $g(x)$ can be produced from $f(x)$ and which signals $f(x)$ produce $g(x) = 0$.

1. **Range Space**:

$$R(A) = \{g(x) = f(x) * h(x), x \in \chi\}.$$

The range space of the forward operator A consists of all possible output signals $g(x)$ that can be obtained by convolving the input signal $f(x)$ with the line of confusion PSF $h(x)$. Since convolution is a linear operation, any signal that can be represented as a convolution of $f(x)$ with $h(x)$ lies in the range space of A.

2. **Null Space**:

$$N(A) = \{g(x) = f(x) * h(x) = 0, x \in \chi\}.$$

The null space of the forward operator F consists of all input signals $f(x)$ that produce zero output signal $g(x) = 0$ when convolved with the PSF $h(x)$. In other words, any signal $f(x)$ that is orthogonal to the PSF $h(x)$ lies in the null space of A.

3. **Adjoint operator A^***:

$$< f(x) * h(x), g(x) >_{g(x)} = < f(x), A^* g(x) = h^*(x) * g(x) >_{f(x)}.$$

It represents the convolution operation with the adjoint PSF $h^*(x)$. Since convolution is a linear operation, the adjoint operator A^* can be represented by convolution with the complex conjugate of the PSF $h^*(x)$.

4. **The spectral representation** of an operator is obtained through its eigenvalues and eigenvectors. For the forward operator A, we can represent it in the Fourier domain to simplify the analysis. Let $\mathcal{F}$ denote the Fourier transform operator.

$$G(u) = \mathcal{F}\{g(x)\}, \quad H(u) = \mathcal{F}\{h(x)\},$$

$$A(u) = \int_{-\infty}^{+\infty} G(u) \cdot H(u) \cdot e^{i \cdot 2\pi \cdot ux} du.$$

The **range space and null space of** A^* can be determined similarly to A, considering the adjoint PSF $h^*(x)$ instead.

5. **Range Space of adjoint operator** A^*:

$$R(A^*) = \{g(x) = f(x) * h^*(x), x \in \chi\} = N(A)^{\perp}.$$

6. **Null Space of adjoint operator** A^*:

$$N(A^*) = \{g(x) = f(x) * h^*(x) = 0, x \in \chi\} = R(A)^{\perp}.$$

7. **For the adjoint operator** A^***, the spectral representation** is given by:

$$A^*(u) = \int_{-\infty}^{+\infty} G(u) \cdot H^*(u) \cdot e^{i \cdot 2\pi \cdot ux} du,$$

where $H^*(u)$ is the complex conjugate of $H(u)$.

1.2 (b) Based on the analysis in (a) and what we learned about null space, explain the following observations in practice

1.2.1 (3) Why the larger the defocus distance z is, the "harder" it is to recover the in-focus object given only a single defocused image

Short Answer:
Generally speaking, because the inverse mapping is not unique, and the inverse process is ill-posed! The larger the defocus distance z, the larger the null space of the forward operator becomes, meaning that there are more possible input signals $f(x)$ that result in the same defocused image $g(x)$.

The observation that it becomes "harder" to recover the in-focus object as the defocus distance z increases, given only a single defocused image, can be explained by considering the null space of the forward operator and its implications for signal recovery.

In the context of the provided forward model, the null space of the forward operator consists of all input signals $f(x)$ that result in a zero output signal $g(x) = 0$ when convolved with the line of confusion PSF $h(x)$. In other words, signals in the null space of the forward operator do not contribute to the defocused image captured by the camera.

As the defocus distance z increases, the defocused image becomes more severely blurred, and the information about the in-focus object becomes increasingly spread out over a larger area in the defocused image. This spreading of information is reflected in the larger size of the line of confusion PSF $h(x)$, which grows linearly with z according to $d = \theta z$.

Now, if we try to recover the in-focus object from the defocused image using a single defocused image, we encounter difficulties because **the spreading of information due to defocus results in a loss of high-frequency components and details in the captured image**. Additionally, since the defocused image contains contributions from a wide range of object distances, **the information about the in-focus object may be mixed with contributions from other object distances, leading to ambiguity in the reconstruction process**.

In essence, **the larger the defocus distance z, the larger the null space of the forward operator becomes, meaning that there are more possible input signals $f(x)$ that result in the same defocused image $g(x)$**. This ambiguity makes it challenging to uniquely recover the in-focus object from the defocused image, especially when using only a single defocused image.

In practice, to mitigate this difficulty and improve the recovery of the in-focus object from defocused images, techniques such as multi-image fusion, depth estimation, and deblurring algorithms may be employed, which leverage information from multiple defocused images or additional depth information to enhance the reconstruction quality and reduce ambiguity.

1.2.2 (4) Why the larger the NA is, the harder it is to recover the in-focus object given only a single defocused image.

Short Answer:
It is because of the properties of the line of confusion PSF and its impact on the imaging process. As the NA increases, the angle of acceptance θ of the lens also increases, leading to a smaller defocus PSF d. when trying to recover the in-focus object from a defocused image, the smaller defocus PSF associated with a larger NA results in a defocused image that contains more high-frequency components and details of the in-focus object, which blur the image again!

In the provided forward model, the line of confusion PSF represents the blurring effect introduced by defocus, and its size d is proportional to the defocus distance z and inversely proportional to the numerical aperture (NA) according to the formula $d = \theta z$, where θ is related to the NA.

As the NA increases, the angle of acceptance θ of the lens also increases, leading to a smaller defocus PSF d. Conversely, a smaller NA results in a larger defocus PSF. This relationship indicates that a larger NA produces a sharper image, with less spread of light due to defocus.

Now, **when trying to recover the in-focus object from a defocused image, the smaller defocus PSF associated with a larger NA results in a defocused image that contains more high-frequency components and details of the in-focus object, which blur the image again!!!!** However, since the defocused image still suffers from blurring due to defocus, these high-frequency components may be smeared out and mixed with contributions from nearby object distances.

In contrast, a smaller NA produces a larger defocus PSF, leading to more severe blurring in the defocused image. In this case, although the defocused image may contain less high-frequency information about the in-focus object, the blurring effect is more pronounced, making it easier to detect the presence of defocus but potentially harder to extract detailed information about the in-focus object.

Therefore, the relationship between the NA and the difficulty of recovering the in-focus object from a single defocused image can be understood in terms of the trade-off between the level of blurring introduced by defocus and the amount of high-frequency information preserved in the defocused image. While a larger NA produces a sharper defocused image with more high-frequency details, it also leads to a more subtle blurring effect that can make it challenging to distinguish between in-focus and defocused regions.

2 Problem 2: Example of wave optics model based LSI imaging – Digital holography

Holography is a 3D imaging technique, in the sense that it allows recreate the 3D scene (optically or digitally) from its single 2D measurement. In this problem, we will explore the general idea of in-line (Gabor) holography for 2D imaging and understand the unique feature about holography using the tools we have learned so far.

A schematic of the in-line holography is shown in Fig. 2. To record a hologram, a coherent light source (*e.g.* laser) is used to illuminate the 3D scene. Accordingly, the formation of the hologram needs to be modeled using the wave optics model (as opposed to the geometric optics model which does not account for the effect of interference). The hologram (*i.e.* the intensity image captured by the camera) is the result from the interference between the unperturbed illumination (*i.e.* the reference beam) and the light diffracted from the 3D object.

Using a wave optics model, the formation of the hologram from a 2D object at a depth z can be approximated using the following linear shift invariant (LSI) model

$$g_{\text{out}}(x, y) = g_{\text{in}}(x, y; z) * h(x, y; z),$$

where $*$ denotes the 2D convolution, g_{out} is the signal term of interest contained in the hologram measurement, g_{in} is the object signal and is a complex valued function, and h is the point spread function (PSF) and is also a complex valued function. The form of h can be found by the free-space propagation and wave diffraction theory, which has the following approximated form,

$$h(x, y; z) = \frac{1}{i\lambda z} \exp\left\{ik\frac{x^2 + y^2}{2z}\right\},$$

and the corresponding transfer function (i.e. the 2D Fourier transform of the PSF $h(x, y; z)$ at a given depth z).

$$H(u, v; z) = \exp\left\{-i\pi\lambda z\left(u^2 + v^2\right)\right\},$$

where $k = 2\pi/\lambda$ is a constant (*i.e.* the wavenumber), λ is the wavelength of the laser, x, y denote the lateral coordinates and z denotes the axial direction along which the laser propagates from, and u, v denote the spatial frequency coordinates, according to the following 2D Fourier transform definition

$$H(u,v) = \iint h(x,y)\exp\{-i2\pi(ux+vy)\}dxdy.$$

In the following, we will formulate the forward model of this imaging problem using the tools learned in the lectures.

2.1 (1) Construct the forward model in the linear operator form.

To construct the forward model in linear operator form, we need to express the formation of the hologram as a linear operation between the input object signal and the output hologram signal.

Given the LSI model:

$$g_{\text{out}}(x,y) = g_{\text{in}}(x,y;z) * h(x,y;z).$$

We can express this operation as a linear operator A acting on the input signal g_{in} to produce the output signal g_{out}. Let's denote the input signal g_{in} as a vector $\mathbf{g}_{\text{in}}$ and the output signal g_{out} as a vector $\mathbf{g}_{\text{out}}$. Similarly, the PSF $h(x,y;z)$ can be represented as a matrix $\mathbf{H}$.

The convolution operation $g_{\text{in}}(x,y;z) * h(x,y;z)$ can be equivalently represented as matrix-vector multiplication. Therefore, we can write the forward model in linear operator form as:

$$\mathbf{g}_{\text{out}} = \mathbf{H} \cdot \mathbf{g}_{\text{in}}.$$

Here, $\mathbf{H}$ represents the linear operator corresponding to the PSF $h(x,y;z)$. Then, we will have the forward model in the linear operator form:

$$\mathbf{g}_{\text{out}} = \int_{-\infty}^{+\infty} g_{\text{in}}\left(x',y';z\right) h(x-x',y-y';z)dx'dy',$$

$$\mathbf{g}_{\text{out}} = \int_{-\infty}^{+\infty} g_{\text{in}}\left(x',y';z\right) \frac{1}{i\lambda z}\exp\left\{ik\frac{(x-x')^2+(y-y')^2}{2z}\right\}dx'dy'.$$

2.2 (2) Find the range, null space, spectral representation and adjoint of the LSI system in the linear operator form.

To find the range space, null space, spectral representation, and adjoint of the LSI system represented by the linear operator $\mathbf{H}$, let's first define the operator. **We can express this operation as a linear operator A acting on the input signal g_{in} to produce the output signal g_{out}.**

Given the linear operator $\mathbf{H}$ representing the PSF $h(x, y; z)$, which operates on the input signal $\mathbf{g}_{\text{in}}$ to produce the output signal $\mathbf{g}_{\text{out}}$, the forward model is represented as:

$$\mathbf{g}_{\text{out}} = \mathbf{H} \cdot \mathbf{g}_{\text{in}}.$$

Now, let's analyze each aspect:

1. **Range Space**:

$$R(A) = \{g_{\text{out}}(x, y) = g_{\text{in}}(x, y; z) * h(x, y; z), x \in \chi, y \in Y\}.$$

The range space represents all possible output signals that can be obtained by applying the operator to the input signals. In this case, it represents all possible holograms that can be generated from the input object signals. The range space depends on the properties of the PSF $h(x, y; z)$ and the specific imaging system.

2. **Null Space**:

$$N(A) = \{g_{\text{out}}(x, y) = g_{\text{in}}(x, y; z) * h(x, y; z) = 0, x \in \chi, y \in Y\}.$$

The null space consists of all input signals that result in zero output signals. In other words, it represents the set of input signals that are completely suppressed or do not contribute to the holographic image formation. The null space depends on the properties of the PSF $h(x, y; z)$ and the specific imaging system.

3. **Spectral Representation**:

$$G(u, v) = \mathcal{F}\{g_{\text{in}}(x, y; z)\},$$
$$H(u, v) = \iint h(x, y) \exp\{-i2\pi(ux + vy)\} dxdy,$$
$$A(u, v) = \int_{-\infty}^{+\infty} G(u, v) \cdot H(u, v) \cdot e^{i \cdot 2\pi \cdot ux \cdot vx} dudv.$$

The spectral representation can be obtained by computing its eigenvalues and eigenvectors. These eigenvalues and eigenvectors provide insights into the frequency response and spatial characteristics of the imaging system. The spectral representation helps understand how the operator acts on different frequency components of the input signals.

4. **Adjoint** A^*:

$$
\begin{aligned}
&< g_{\text{in}}(x, y; z) * h(x, y; z), g_{\text{out}}(x, y) >_{g_{\text{out}}(x,y)} \\
= &< g_{\text{in}}(x, y; z), A^* g_{\text{out}}(x, y) = h^*(x, y; z) * g_{\text{out}}(x, y) >_{g_{\text{in}}(x,y;z)}.
\end{aligned}
$$

The adjoint of the operator A, denoted as A^*, is the conjugate transpose of the operator. It represents the reverse operation of the original operator in the context of inner products. The adjoint operator is important in various mathematical and signal processing applications, including reconstruction algorithms and optimization problems.

2.3 (3) Based on what we learned about null space, explain why it is "easier" to recover an object that this is defocused (i.e. z = 0) using holography, as compared to a standard camera (as in Problem 1).

In the context of null space, let's consider the scenario where the object is defocused, *i.e.*, $z = 0$. **In holography, this means that the object is situated at the same plane as the hologram recording, whereas in a standard camera setup, this would correspond to the object being out of focus.**

When an object is defocused in holography, it implies that the input signal $g_{\text{in}}(x, y; z)$ is essentially a plane wave or uniform illumination across the hologram plane. This input signal is relatively simple and contains fewer spatial variations compared to the focused case, where the object is at a distance away from the hologram plane. Consequently, the null space of the linear operator $\mathbf{H}$ for the defocused case would likely be smaller or contain fewer degrees of freedom.

In contrast, in a standard camera setup, when the object is out of focus, the recorded image contains a mix of information from different depths, leading to a more complex representation. The out-of-focus blur spreads the object information across the image, making it harder to recover the original object. In this case, the null space of the imaging system is likely larger due to the loss of depth information and the mixing of spatial frequencies.

Therefore, because holography captures both amplitude and phase information of the object, including its depth, it is "easier" to recover a defocused object using holography compared to a standard camera. The simpler input signal and the ability to capture depth information enable holography to provide better reconstruction of the object, even when it is defocused.