AI61201: Visual Computing With AI/ML

Module 4: Frequency Domain Image Analysis and Filtering

Dr. Somdyuti Paul

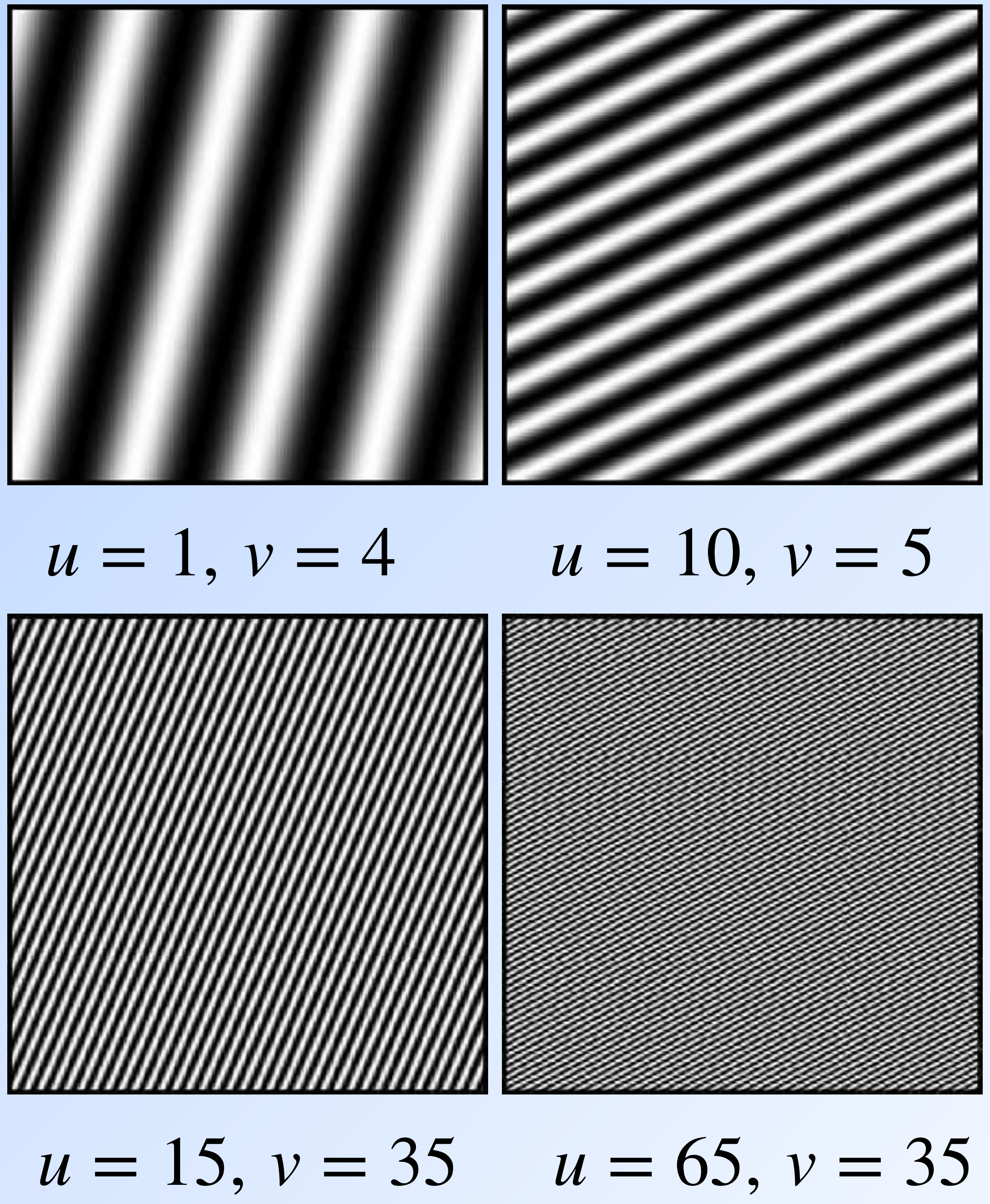
Spatial (2D) Frequency

- Just as a 1D signal can be represented in terms of its constituent frequencies, a 2D signal could be represented in terms of its constituent spatial frequencies.

- A 2D discrete space sinusoid has the following functional form:
  $A\sin[2\pi(Ux + Vy)]$
  where U and V are the spatial frequencies along the horizontal and vertical directions in cycles per pixel.

- The frequency and direction of fastest oscillation is given by:
  $\Omega = \sqrt{U^2 + V^2}$ (radial frequency) and $\theta = \tan^{-1}\dfrac{V}{U}$.

- It is often convenient to represent the spatial frequencies in the units of cycles/image by scaling U and V with the image dimensions:
  $A\sin[2\pi(\dfrac{u}{M}x + \dfrac{v}{N}n)]$



2D sinusoids at different spatial frequencies

Discrete Space Fourier Transform

- The discrete space Fourier transform (DSFT) is a continuous function that describes the frequency composition of a discrete spatial signal.

- The DSFT is given by:

$$F(U,V) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f(x,y)e^{-j2\pi(Ux+Vy)}$$

- The DSFT is complex valued: $F(U,V) = R(U,V) + jI(U,V)$
  - Magnitude spectrum: $|F(U,V)| = \sqrt{R^2(U,V) + I^2(U,V)}$
  - Phase spectrum: $\angle F(U,V) = \tan^{-1}\dfrac{I(U,V)}{R(U,V)}$

Discrete Space Fourier Transform

- The Fourier transform of the discrete space signal is continuous and periodic.
- The discrete space signal can be recovered using the inverse DSFT (IDSFT) as follows:

$$f(x, y) = \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} F(U, V) e^{j2\pi(Ux+Vy)} dUdV$$

- f and F constitute a Fourier transform pair and is represented as $f \xrightarrow{\mathscr{F}} F$.

Properties of DSFT

- **Linearity**
  $af_1(x,y) + bf_2(x,y) \xrightarrow{\mathcal{F}} aF_1(U,V) + bF_2(U,V)$

- **Symmetry**: if $f(x,y)$ is real (e.g. natural images) the DSFT $F(U,V)$ is conjugate symmetric:
  $F(U,V) = F^*(-U,-V)$
  This implies that the magnitude spectrum is even symmetric while the phase spectrum is odd symmetric.

- **Translation:**
  - In the spatial domain: $f(x,y)e^{j2\pi(U_0x+V_0y)} \xrightarrow{\mathcal{F}} F(U-U_0, V-V_0)$
  - In the frequency domain: $f(x-x_0, y-y_0) \xrightarrow{\mathcal{F}} F(U,V)e^{-j2\pi(Ux_0+Vy_0)}$

- **Convolution Theorem:**
  $f(x,y) * h(x,y) \xrightarrow{\mathcal{F}} F(U,V)H(U,V)$

Discrete Fourier Transform

- Since the DSFT is a continuous function of spatial frequency, it is not useful for analyzing and processing digital images.
- A more convenient discrete representation of the DSFT exists in the form of discrete Fourier transform (DFT).
- The DFT is derived by sampling the DSFT at discrete frequencies.
- For signals having a finite spatial dimension of $M \times N$ (such as images), the DSFT becomes:

$$F(U,V) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(Ux+Vy)}$$

- This finite extent spatial signal can be represented by a weighted sum of a finite number of frequency components by sampling frequencies at:

$\dfrac{u}{M}$ for $u = 0,1,\cdots,M-1$ and $\dfrac{v}{N}$ for $v = 0,1,\cdots,N-1$, over one period of the DSFT.

- So, the DFT is given by:

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(\frac{u}{M}x+\frac{v}{N}y)}$$

Discrete Fourier Transform

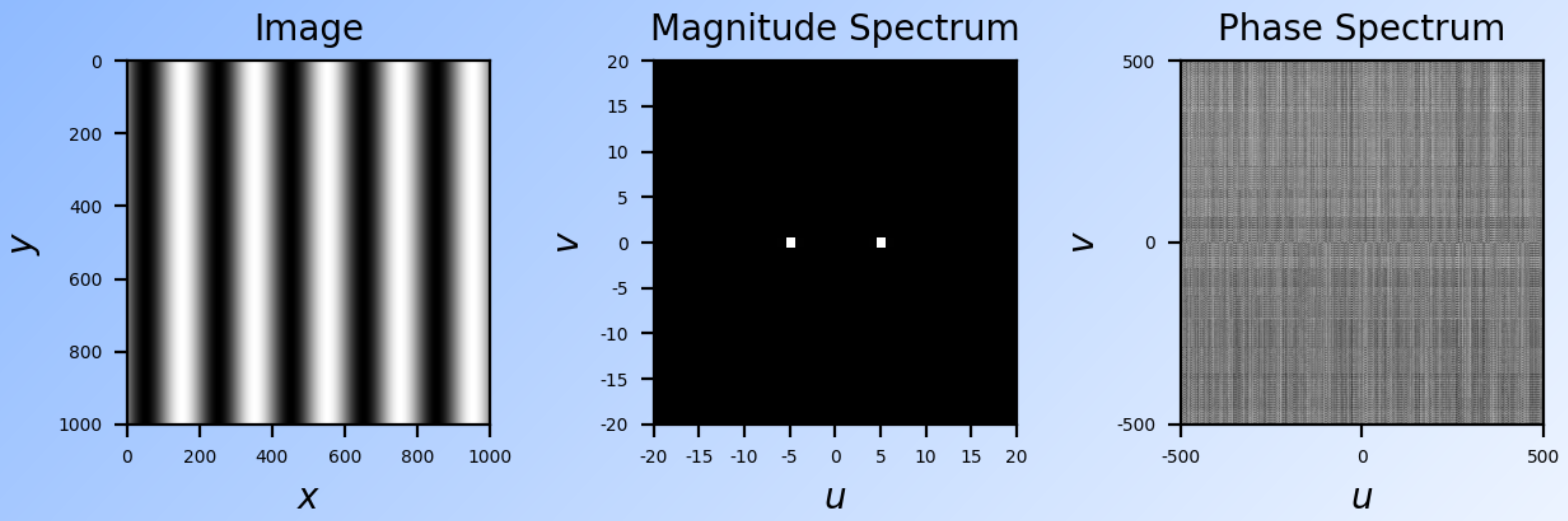
- Thus, the DFT of a $M \times N$ image $f(x, y)$ is a finite extent, complex valued matrix $F(u, v)$ of size $M \times N$.
- The spatial signal $f(m, n)$ could be uniquely recovered by performing the inverse DFT (IDFT) as follows:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{u}{M}x + \frac{v}{N}y)}$$

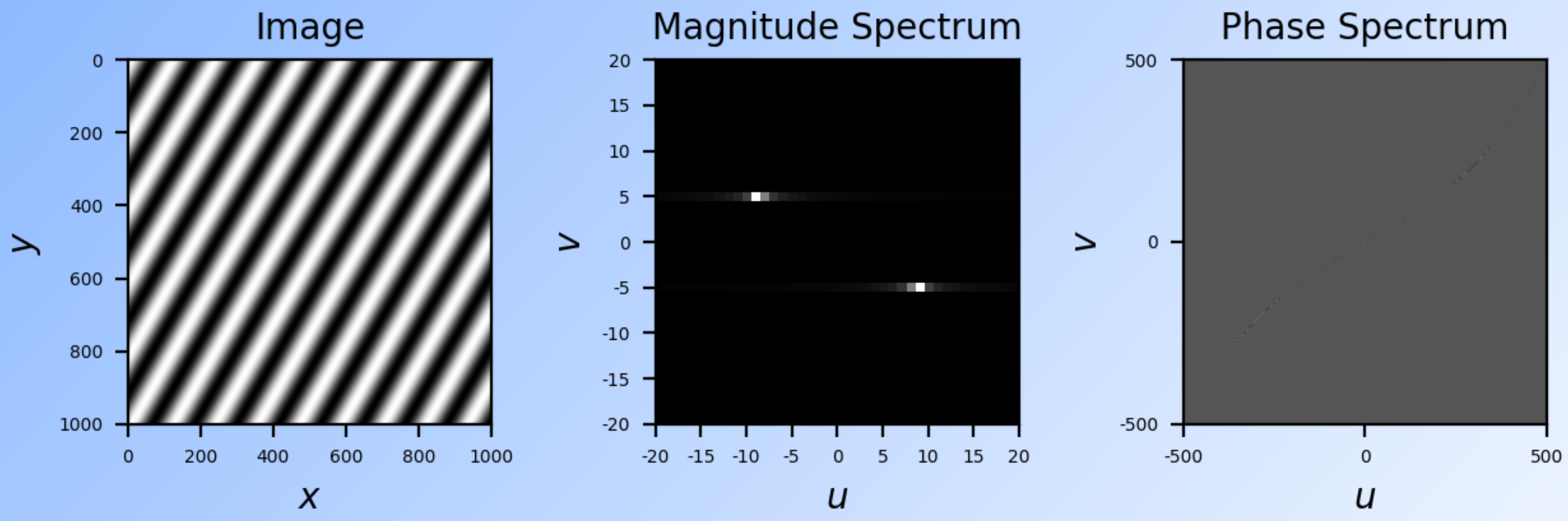
<u>DFT visualization:</u>

- The magnitude spectrum and the phase spectrum of DFT can be visualized as images.
- A logarithmic transformation is often applied to the magnitude spectrum to highlight high frequency components.
- The origin of the DFT is usually shifted from the top left corner to the center of the image in order to display it in the conventional manner, i.e. instead of displaying $F(u, v)$, we display $F(u - M/2, v - N/2)$.

DFT Visualization

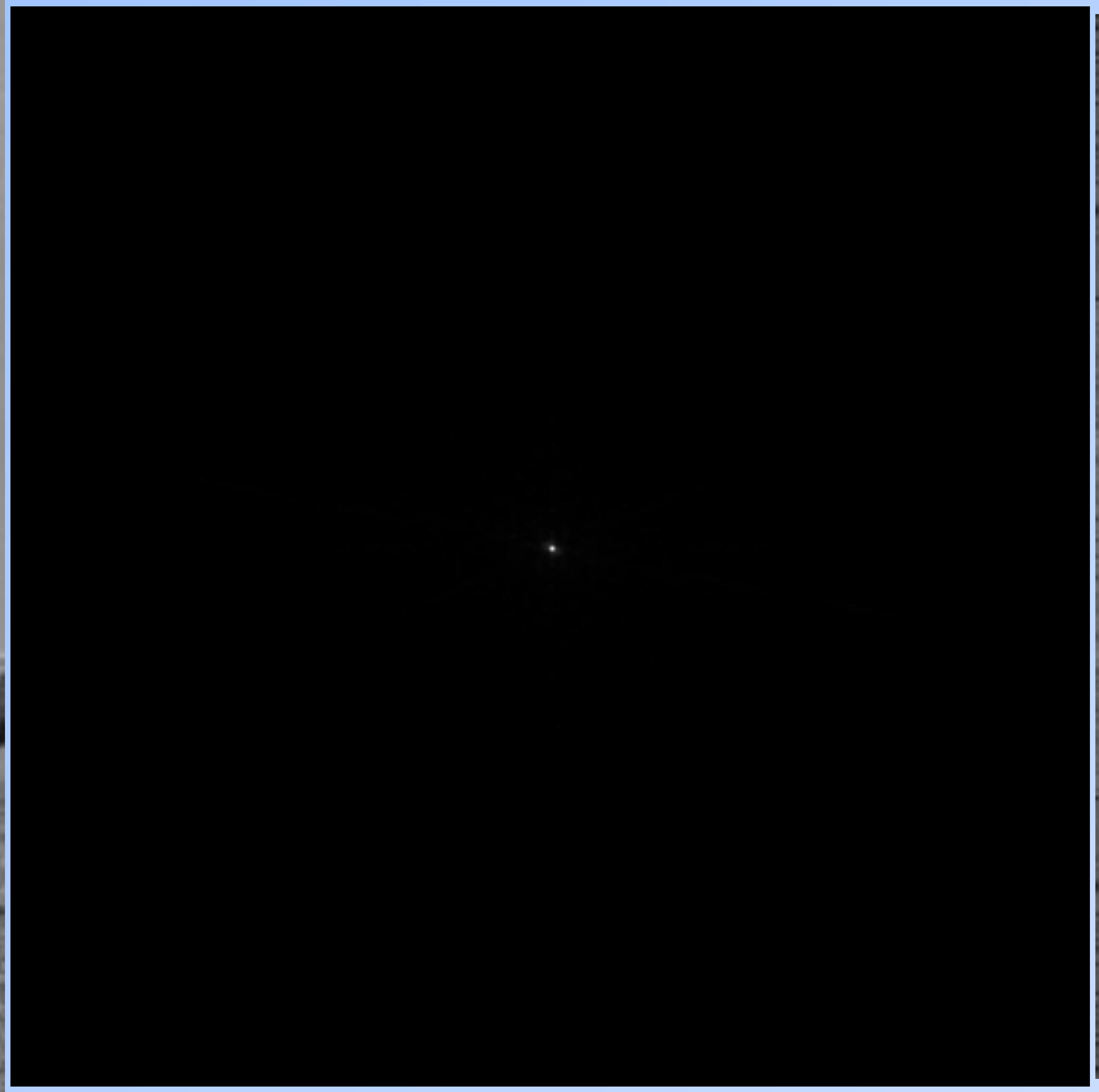


DFT Visualization

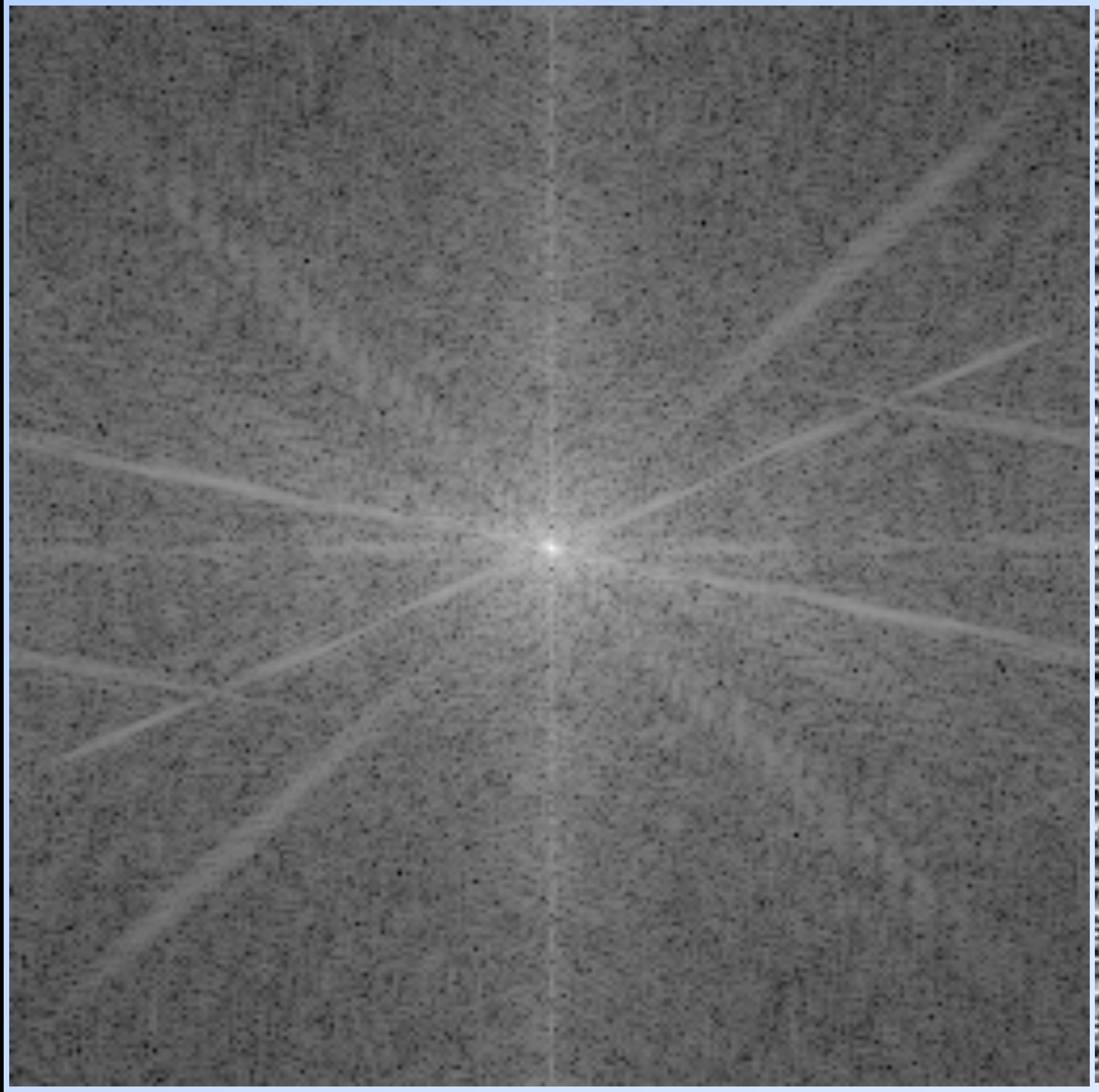


DFT Visualization

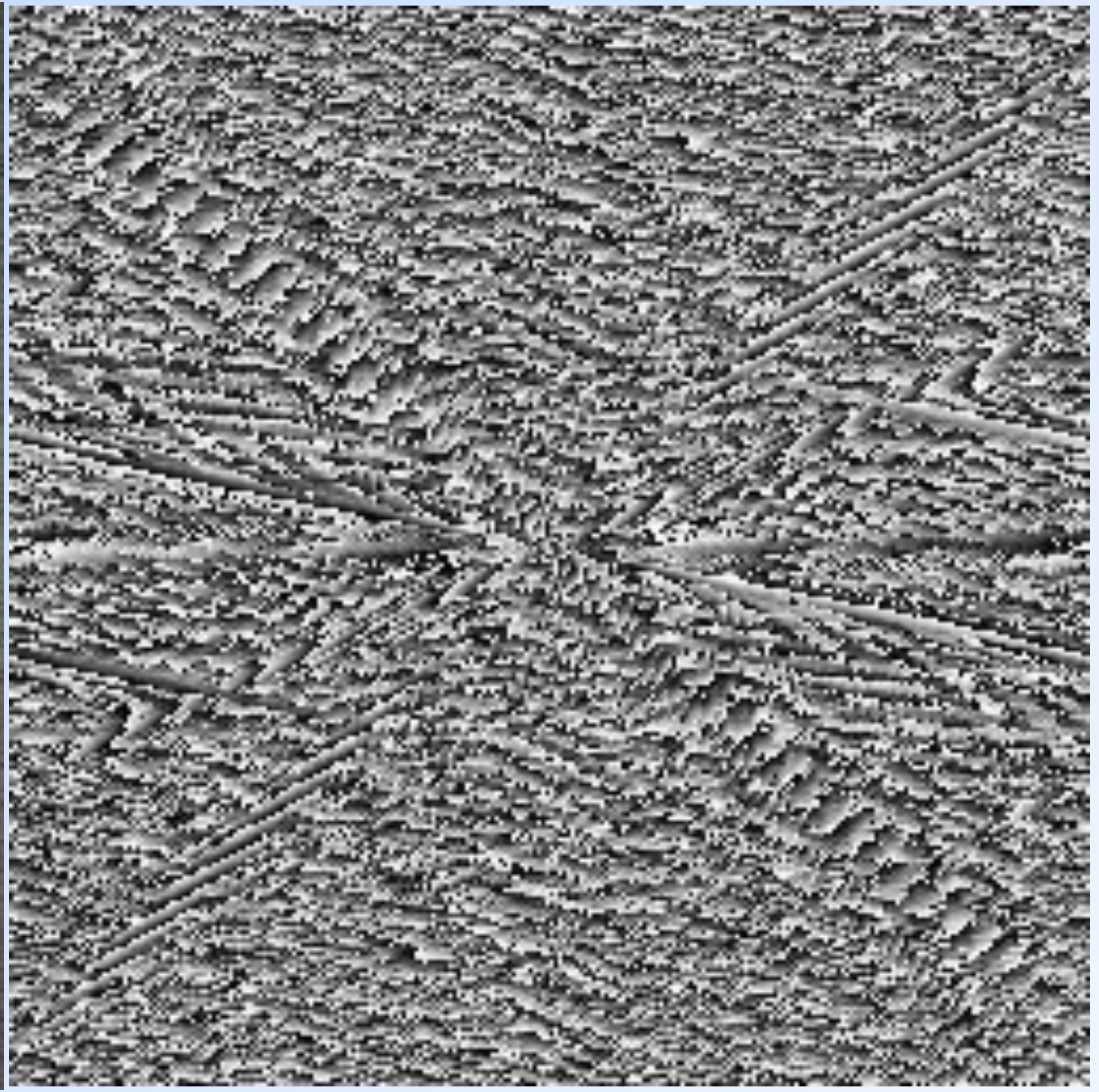
Image



Magnitude Spectrum



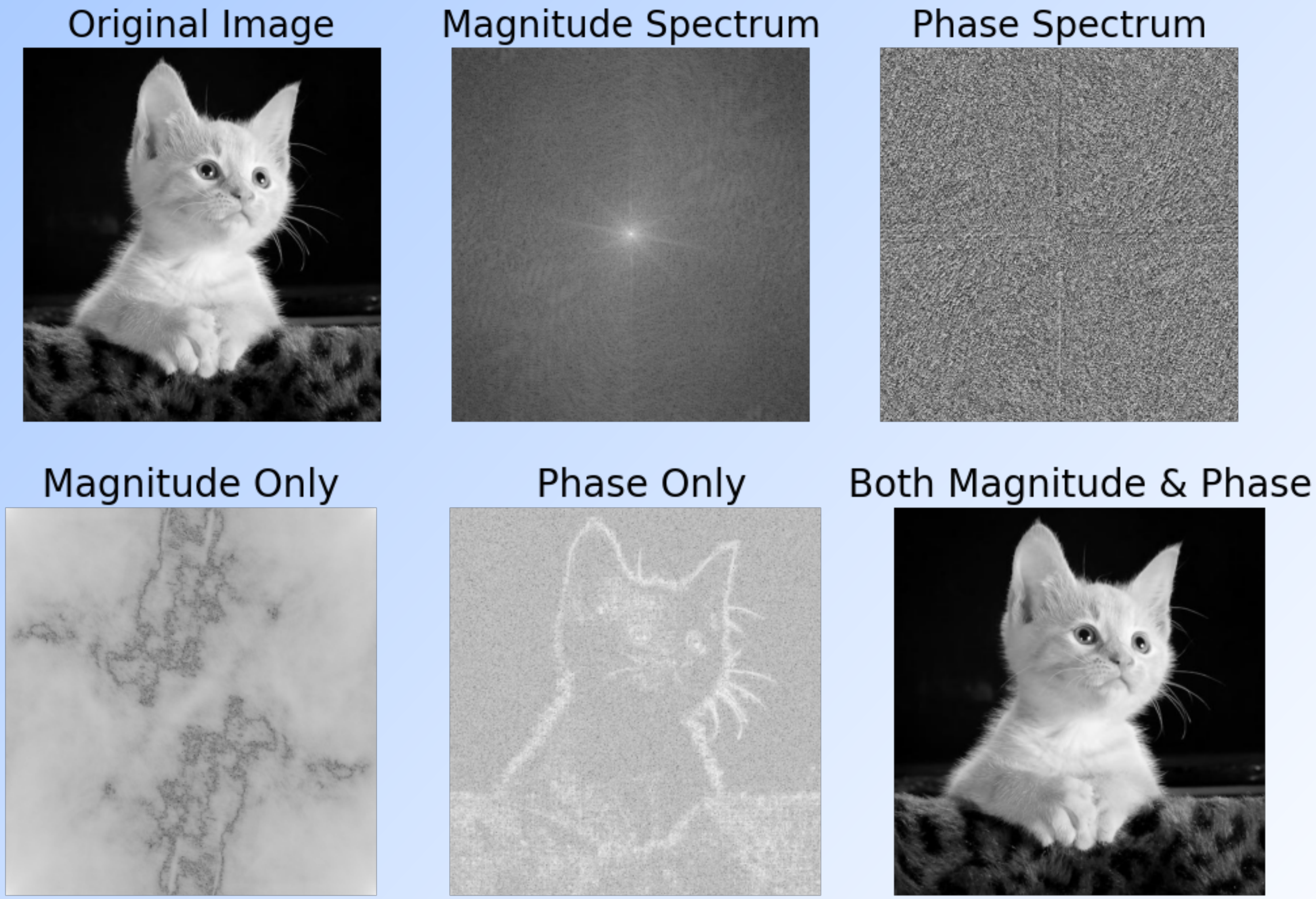
Magnitude Spectrum
after log transformation



Phase Spectrum

Importance of DFT Phase

- The magnitude spectrum provides information about the strength of various frequency components,
- the phase spectrum provides critical information about the spatial arrangement and interaction of the frequency components
- The DFT phase contains information about the image structure
- Although visually uninformative, the phase information is vital for image reconstruction.



DFT Properties

- The DFT is linear, invertible and conjugate symmetric (same as DSFT).

- **Periodicity in the Frequency Domain**: the DFT is periodic with periods of M and N along the u and v dimensions, respectively.
  $F(u + kM, v + lN) = F(u, v)$ for $0 \leq u \leq M - 1$ and $0 \leq v \leq N - 1$

- **Periodicity in the Spatial Domain:** similarly, the IDFT implies that the image is periodic in the spatial domain (a counterintuitive effect that arises due to sampling the DSFT).
  $f(x + kM, y + lN) = f(x, y)$ for $0 \leq x \leq M - 1$ and $0 \leq y \leq N - 1$

- **Duality**: The duality property of DFT can be stated as follows:
  If $f(x, y) \leftrightarrow F(U, V)$ then $F(x, y) \leftrightarrow MNf((-U)_M, (-V)_N)$

Circular Convolution

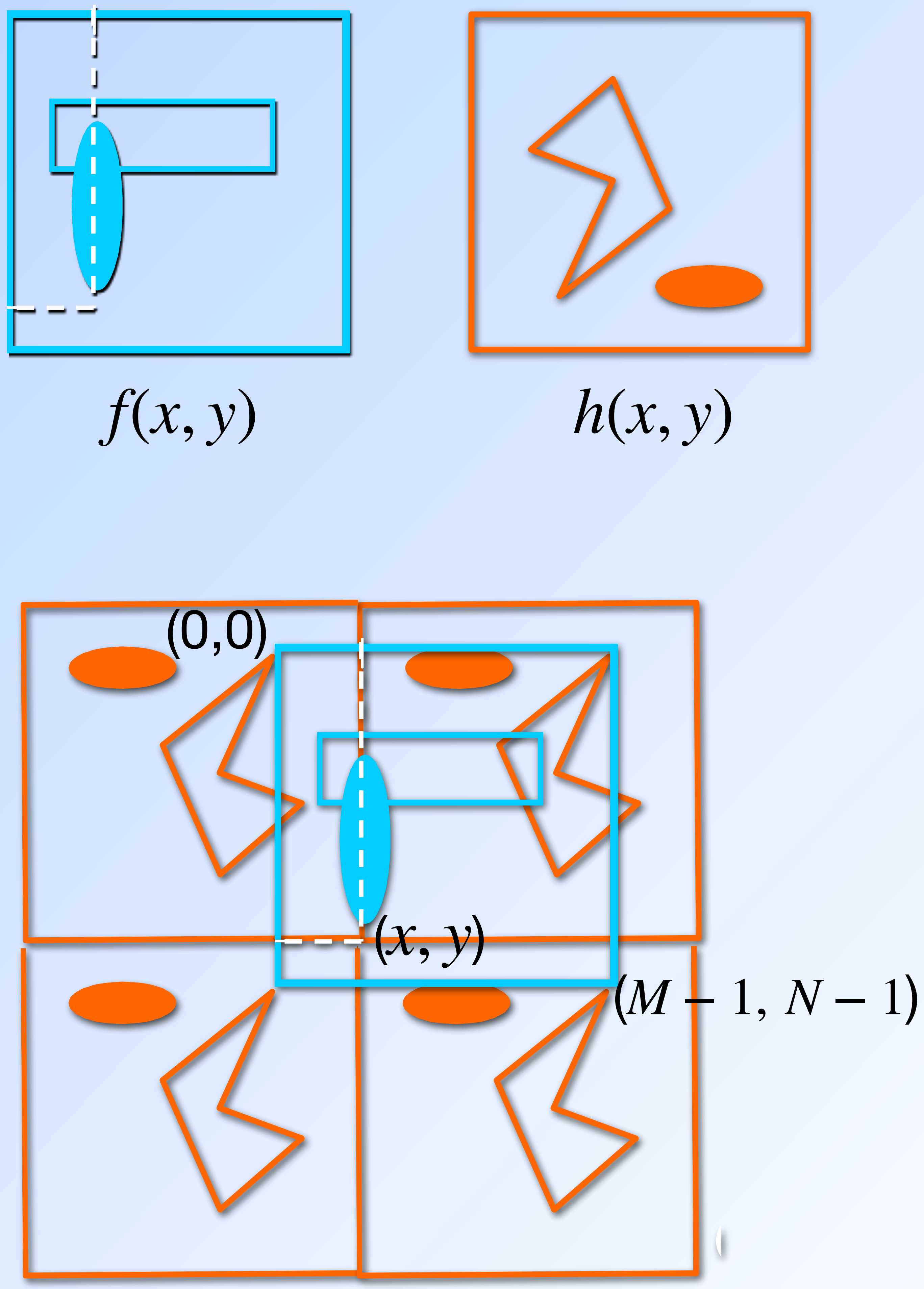
Circular Convolution

The multiplication of the DFTs of two images correspond to the circular convolution of the images in the spatial domain, i.e.

$F(u,v)H(u,v) \xrightarrow{\mathscr{F}^{-1}} f(x,y) \circledast h(x,y)$, where

$$f(x,y) \circledast h(x,y) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} f(p,q) h[(x-p)_M, (y-q)_N] \quad \text{(Circular convolution)}$$

where $(z)_N = z$ mod N.



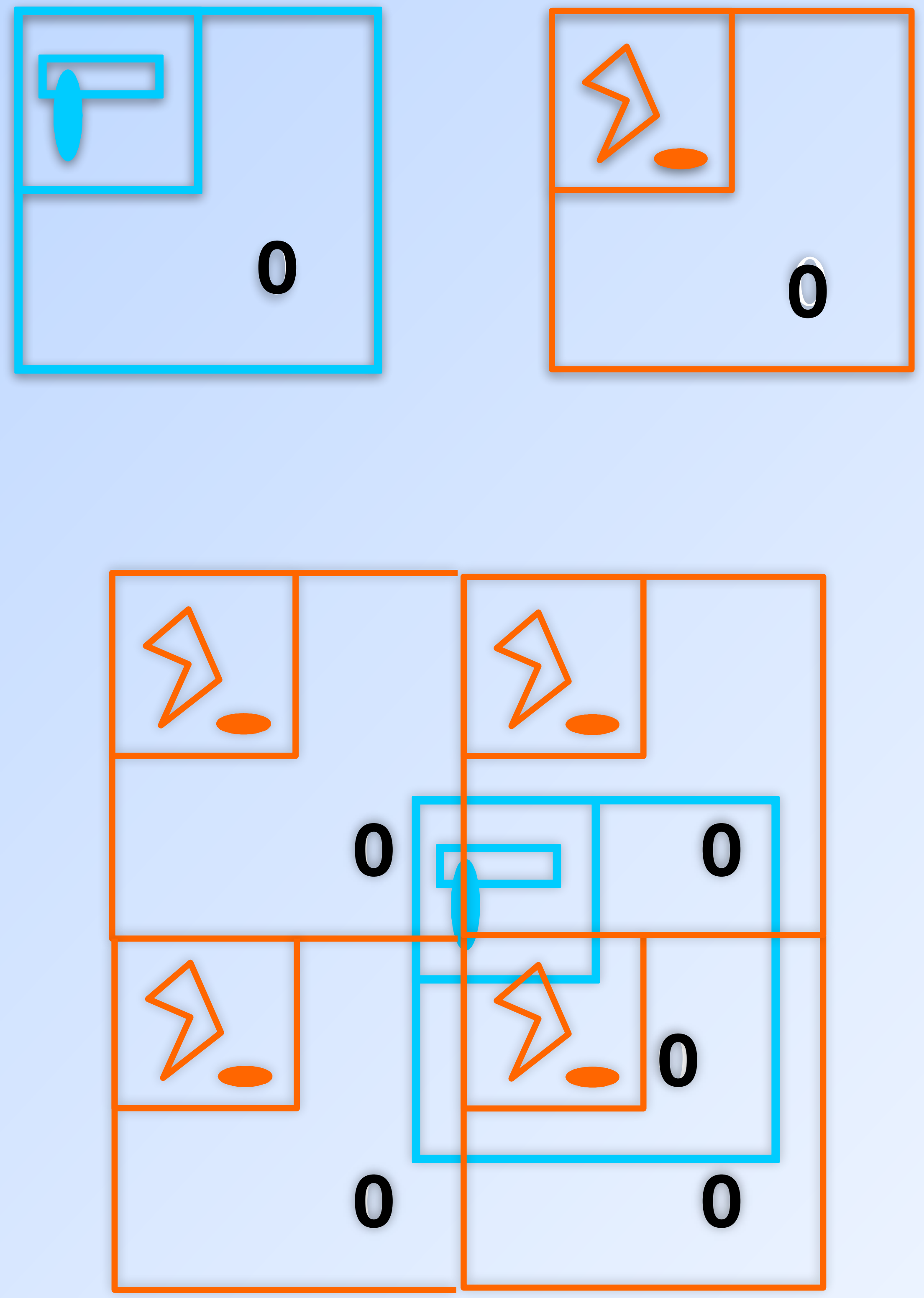
Circular Convolution

- In most applications, the linear convolution output is desirable

- The linear convolution result can be recovered from the product of the two DFTs by appropriately padding the inputs in the spatial domain.

- The size of the full linear convolution of the same images is $(M + P - 1) \times (N + Q - 1)$

- Thus, to make the size of the circular convolution correspond to that of the linear convolution, both the images must be zero padded to a size of $(M + P - 1) \times (N + Q - 1)$, before multiplying their DFTs.



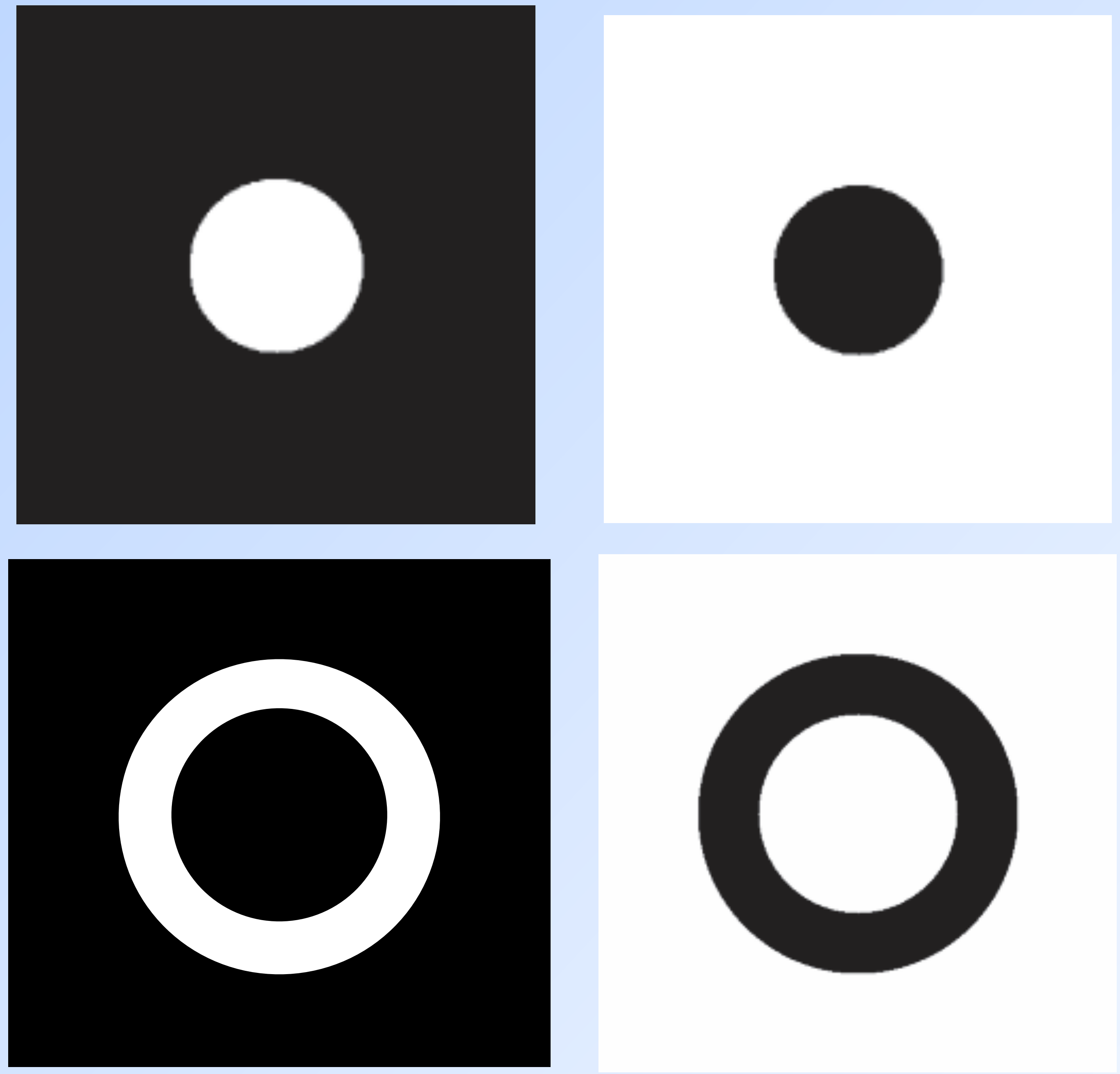
Linear Convolution Using Zero-Padded Circular Convolution

Important DFT Pairs

Function	$f(x,y)$	$F(u,v)$
Delta	$\delta(x,y)$	1
Cosine	$cos(2\pi(u_0x+v_0y))$	$\frac{MN}{2}\left[\delta(u-u_0M, v-v_0N)+\delta(u+u_0M, v+v_0N)\right]$
Rectangle	$rec(a,b)$	$ab\frac{\sin(\frac{\pi ua}{M})}{\frac{\pi ua}{M}}\frac{\sin(\frac{\pi vb}{N})}{\frac{\pi vb}{N}}e^{j\pi(\frac{u}{M}+\frac{v}{N})}$
Gaussian	$e^{-(\frac{x^2+y^2}{2\sigma^2})}$	$2\pi\sigma^2e^{-2\pi^2\sigma^2(u^2+v^2)}$

Image Filtering in the Frequency Domain

- A digital filter can be characterized by its impulse response $h(m, n)$ or its frequency response $H(u, v)$.
- The frequency response of a system specifies how the system affects each frequency component of its input.
- The link between filtering in the spatial and frequency domains is established by the convolution theorem.
- Frequency domain filtering consists of modifying the DFT of an image $F(u, v)$ using a filter transfer function $H(u, v)$ that selectively attenuates or amplifies the contribution of some frequencies, to alter the appearance of the image $f(x, y)$ in the spatial domain.



Ideal Low Pass, High Pass and Band Pass and Band Reject Filter Responses.

Frequency Domain Filtering Steps

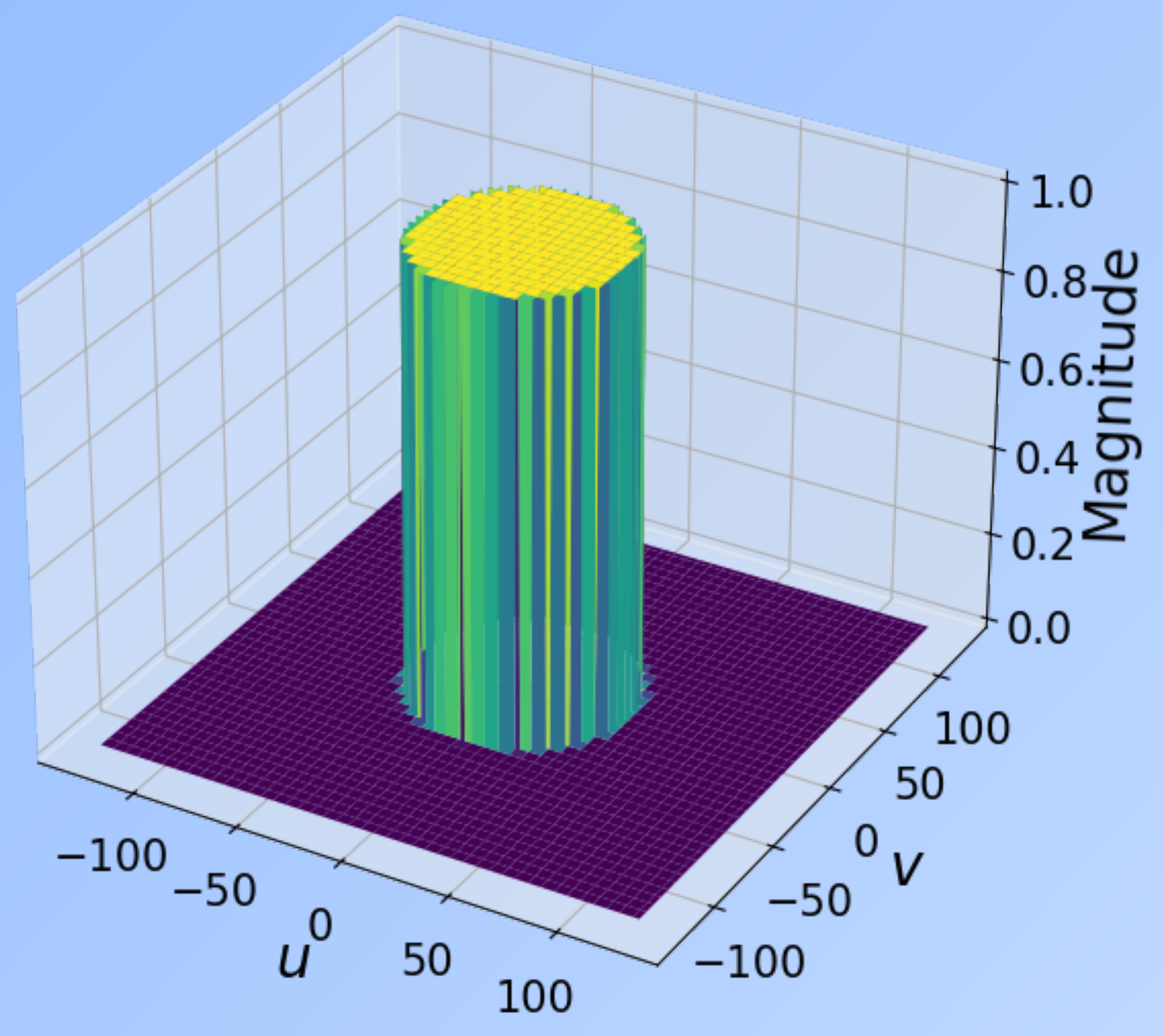
1. *Padding:* Given an input image $f(x, y)$ of size $M \times N$, and filter kernel get padded image $f_p(x, y)$ of size $2M \times 2N$.
2. *Shifting*: Multiply $f_p(x, y)$ with $(-1)^{x+y}$ to center its DFT.
3. DFT of image: $(-1)^{x+y} f(x, y) \xrightarrow{\mathcal{F}} F_p(u, v)$
4. Filter design: construct the symmetric filter transfer function $H(u, v)$ of size $2M \times 2N$.
5. Filtering: obtain the element-wise product $G(u, v) = F_p(u, v)H(u, v)$.
6. IDFT: $g_p(m, n) = \text{Real}\left(\mathcal{F}^{-1}\left(G(u, v)\right)\right)(-1)^{(x+y)}$
7. Padding removal: obtain the unpadded filtered image $g(m, n)$ by extracting the top $M \times N$ region from $g_p(m, n)$.

Low Pass, High Pass and Bandpass Filtering

- An image can be processed using ideal low pass, high-pass and bandpass filters.

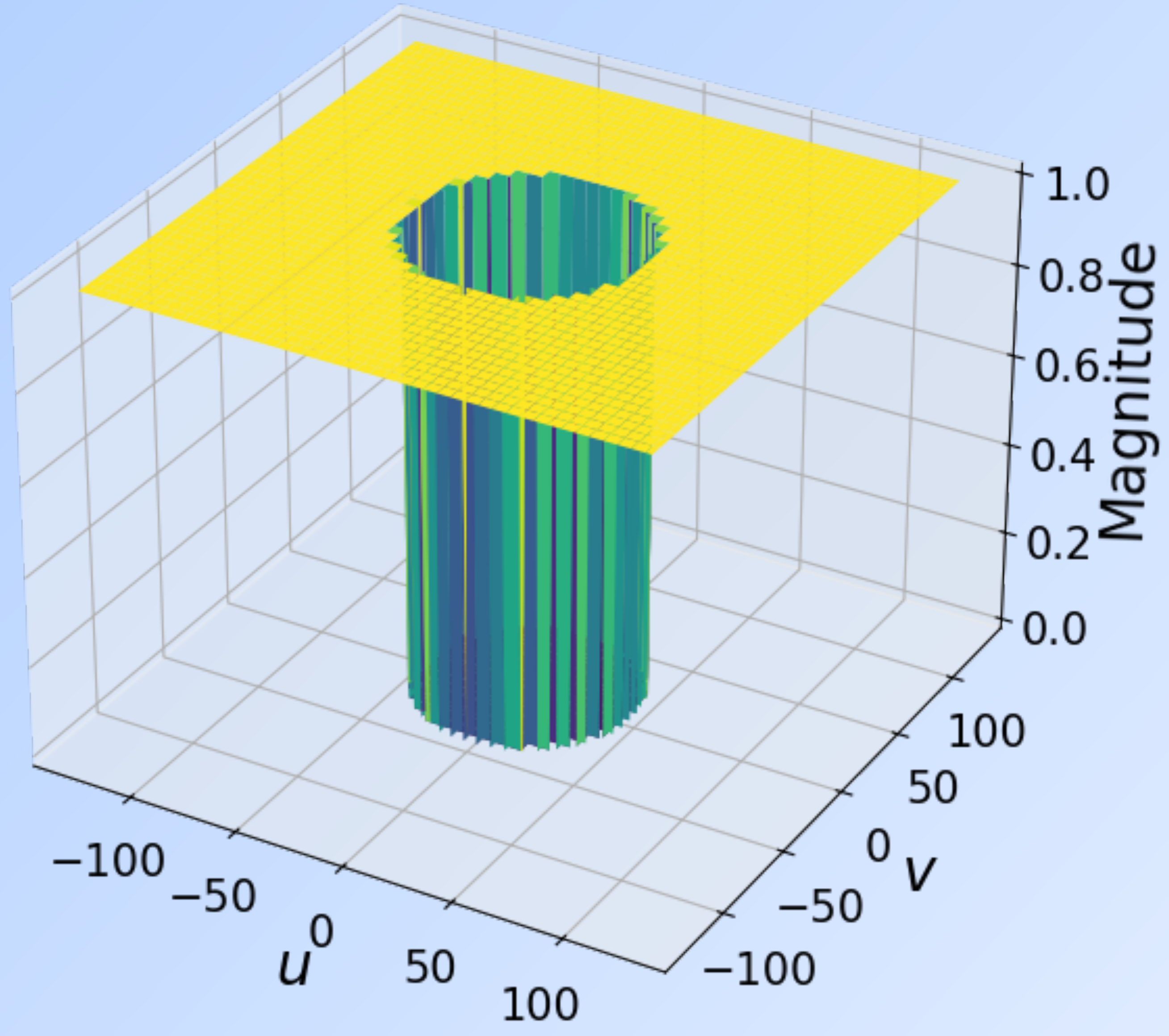
Ideal low Pass Filter

$$H(u,v) = \begin{cases} 1 & \text{if } \sqrt{u^2+v^2} \le D_0 \\ 0 & \text{if } \sqrt{u^2+v^2} > D_0 \end{cases}$$



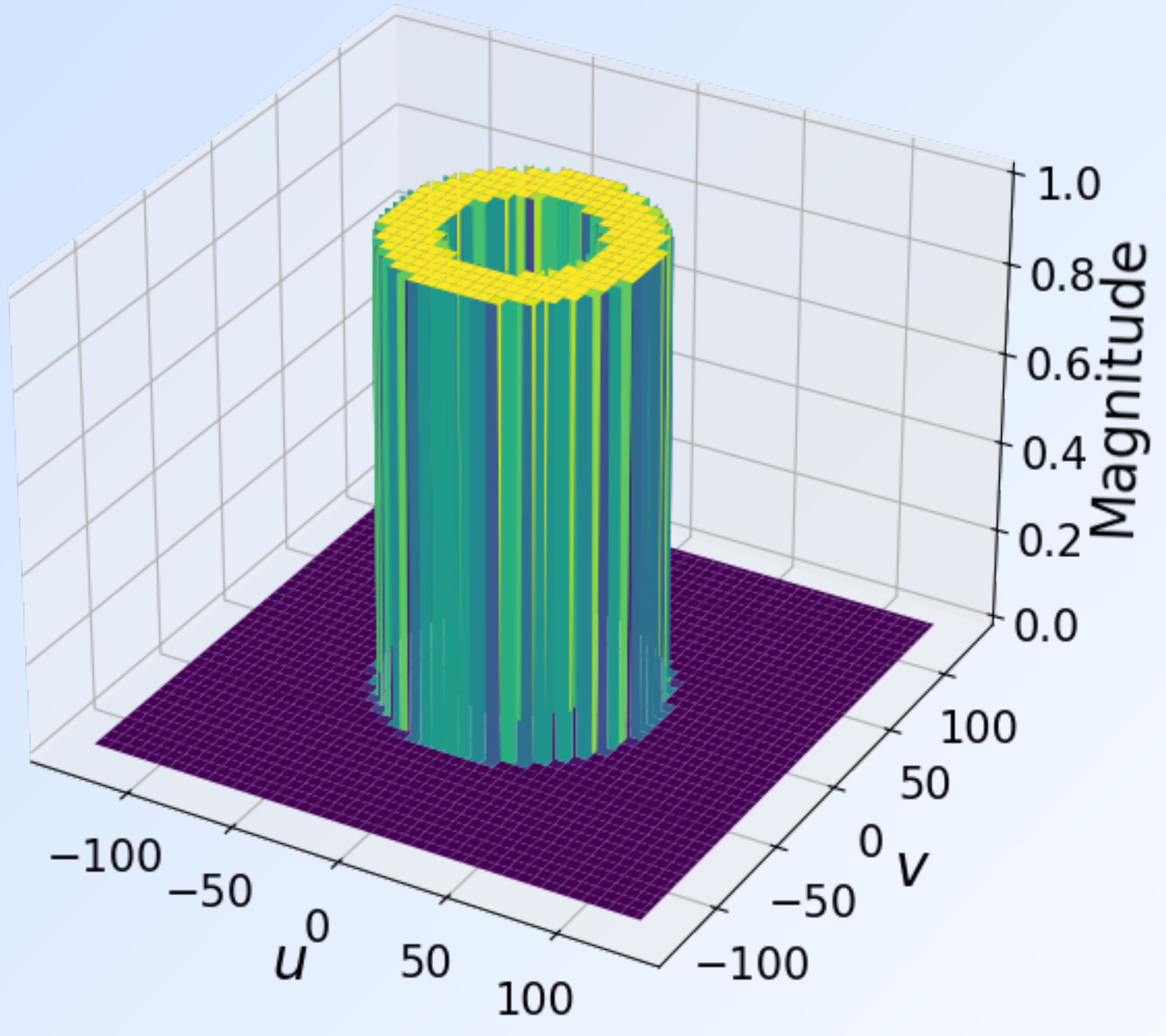
Ideal High Pass Filter

$$H(u,v) = \begin{cases} 0 & \text{if } \sqrt{u^2+v^2} \le D_0 \\ 1 & \text{if } \sqrt{u^2+v^2} > D_0 \end{cases}$$

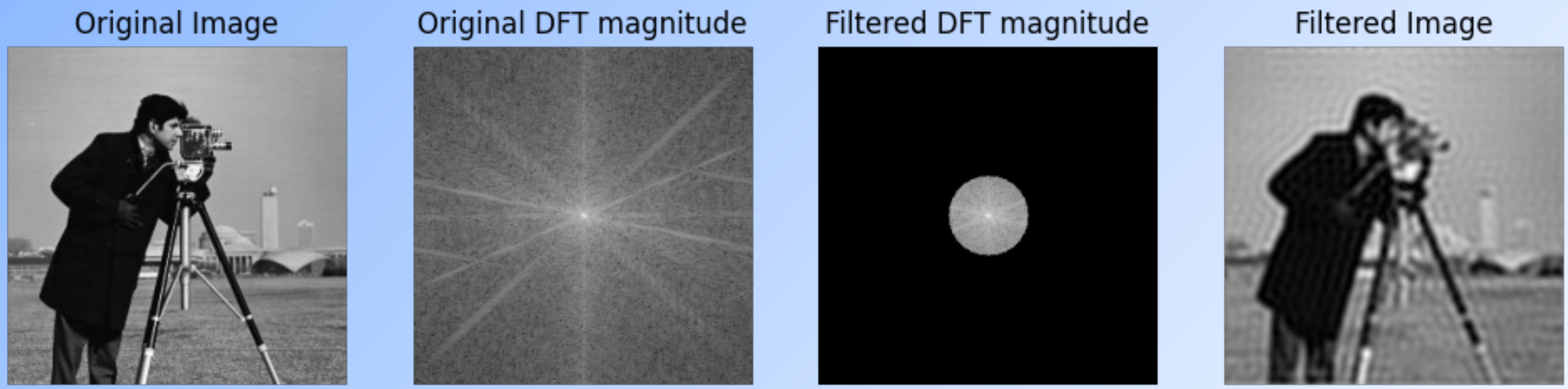


Ideal Bandpass Filter

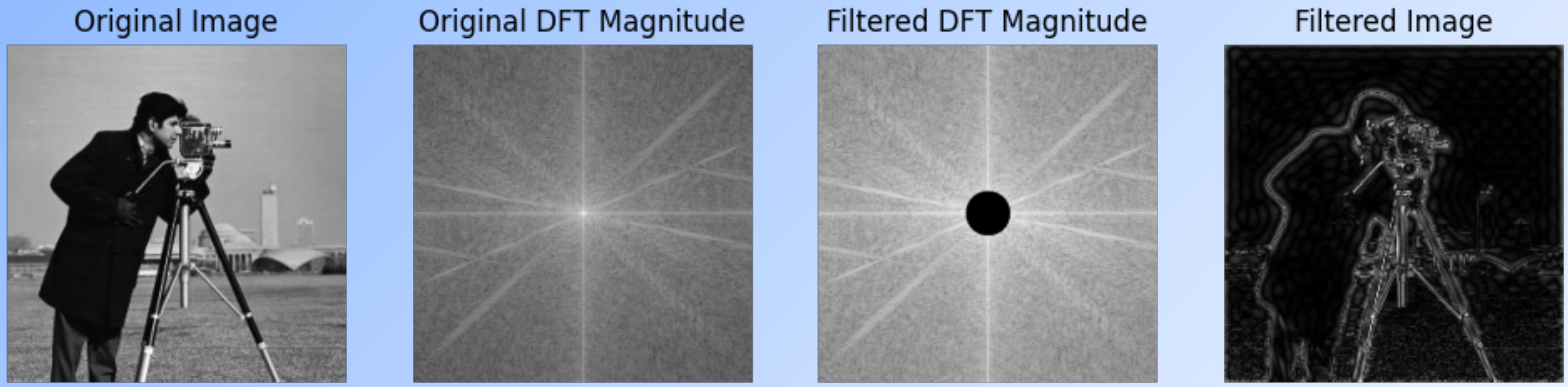
$$H(u,v) = \begin{cases} 1 & \text{if } D_1 \le \sqrt{u^2+v^2} \le D_2 \\ 0 & \text{otherwise} \end{cases}$$



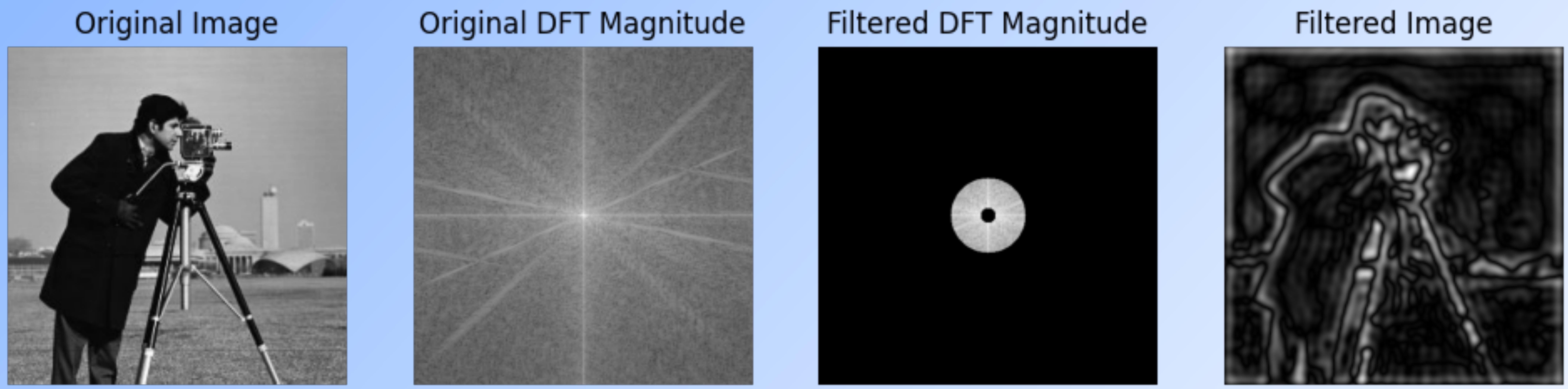
Low Pass, High Pass and Bandpass Filtering



Low Pass, High Pass and Bandpass Filtering

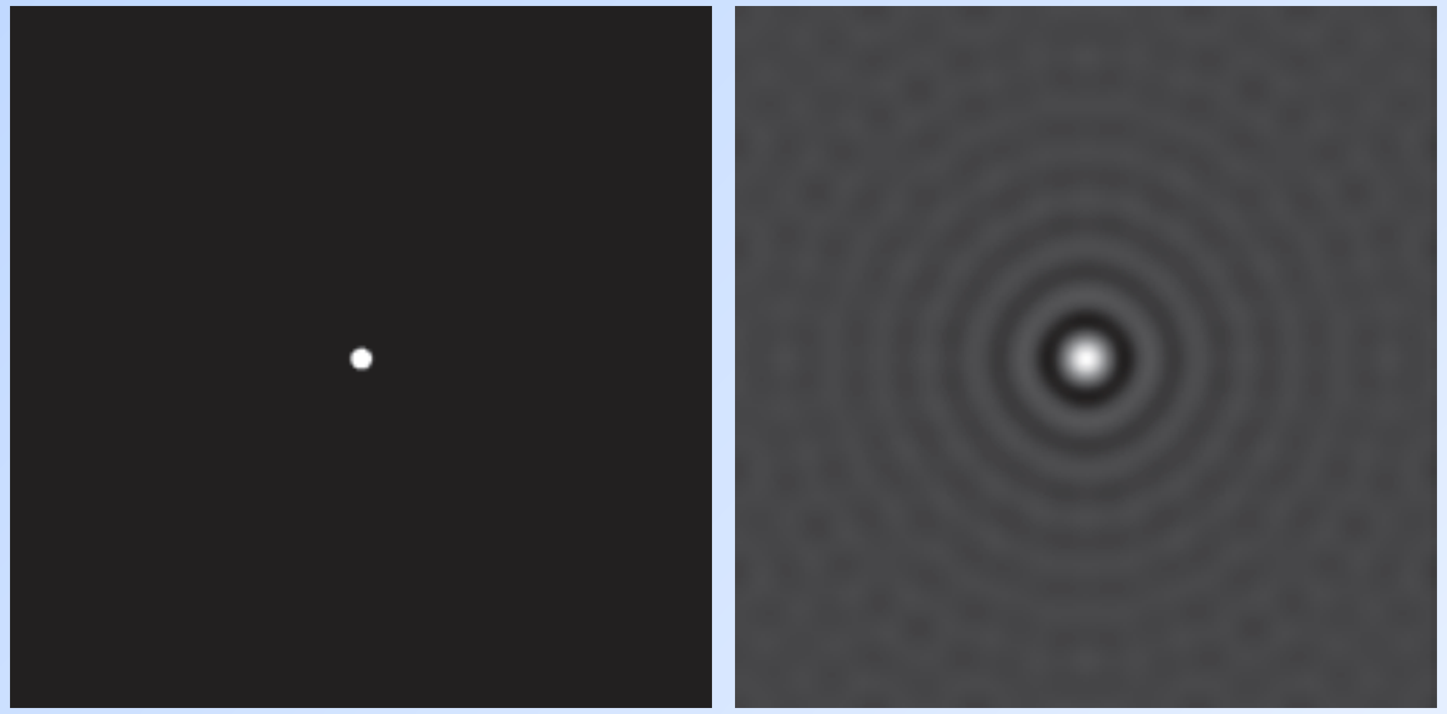


Low Pass, High Pass and Bandpass Filtering



Limitations of Ideal Filters

- Ideal filters that have sharp cut-off frequencies lead to oscillations in the spatial domain which produces unpleasant “ringing” artifacts.
- The inverse Fourier transform of the ideal low pass filter (LPF) resembles a sinc function.
- Convolution of an input image with the sinc-like function in the spatial domain causes ringing due to interference between the minor lobes of the sinc-like spatial pattern.
- The ringing effect is more pronounced when the cut-off frequency of the frequency domain filter is lower.
- Ideal filters are also hard to implement (practically impossible).



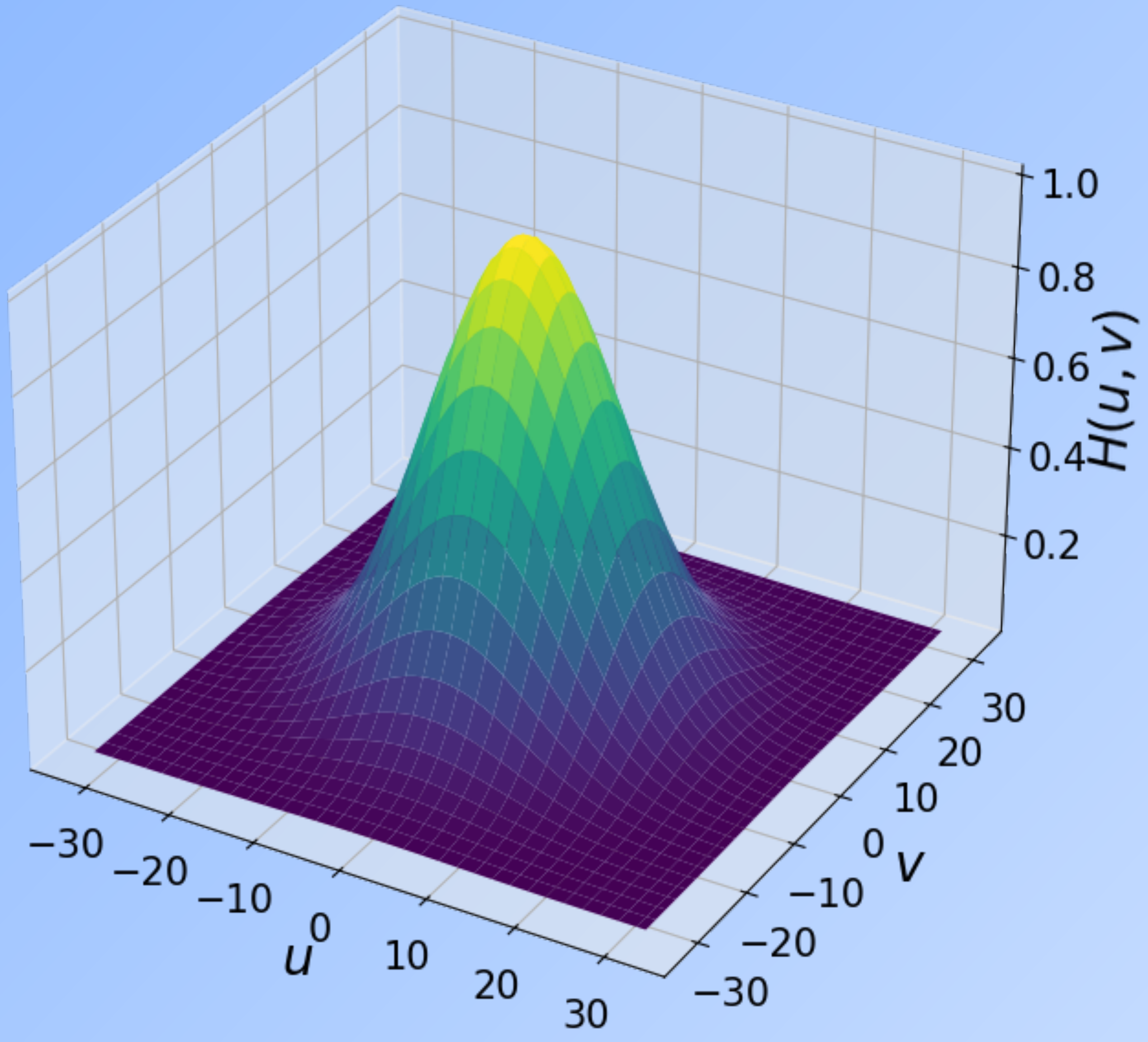
Frequency domain (left) and spatial domain (right) representation of the ideal LPF.

Gaussian Filters

- As an alternative to ideal filters, Gaussian filters could be applied to filter images.

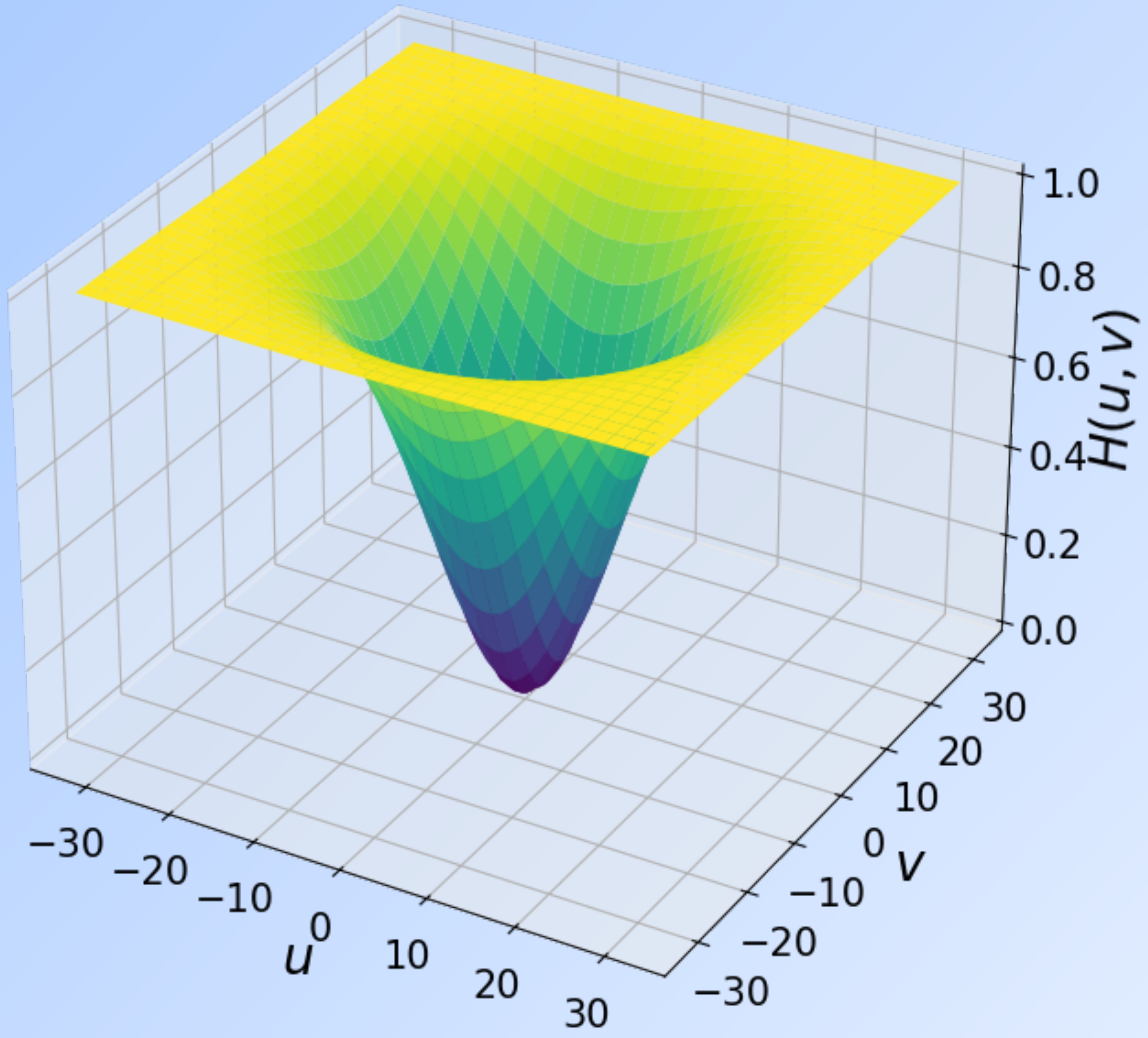
Gaussian low pass filter

$$H(u,v) = e^{-\frac{(u-\mu_u)^2 + (v-\mu_v)^2}{2\sigma^2}}$$

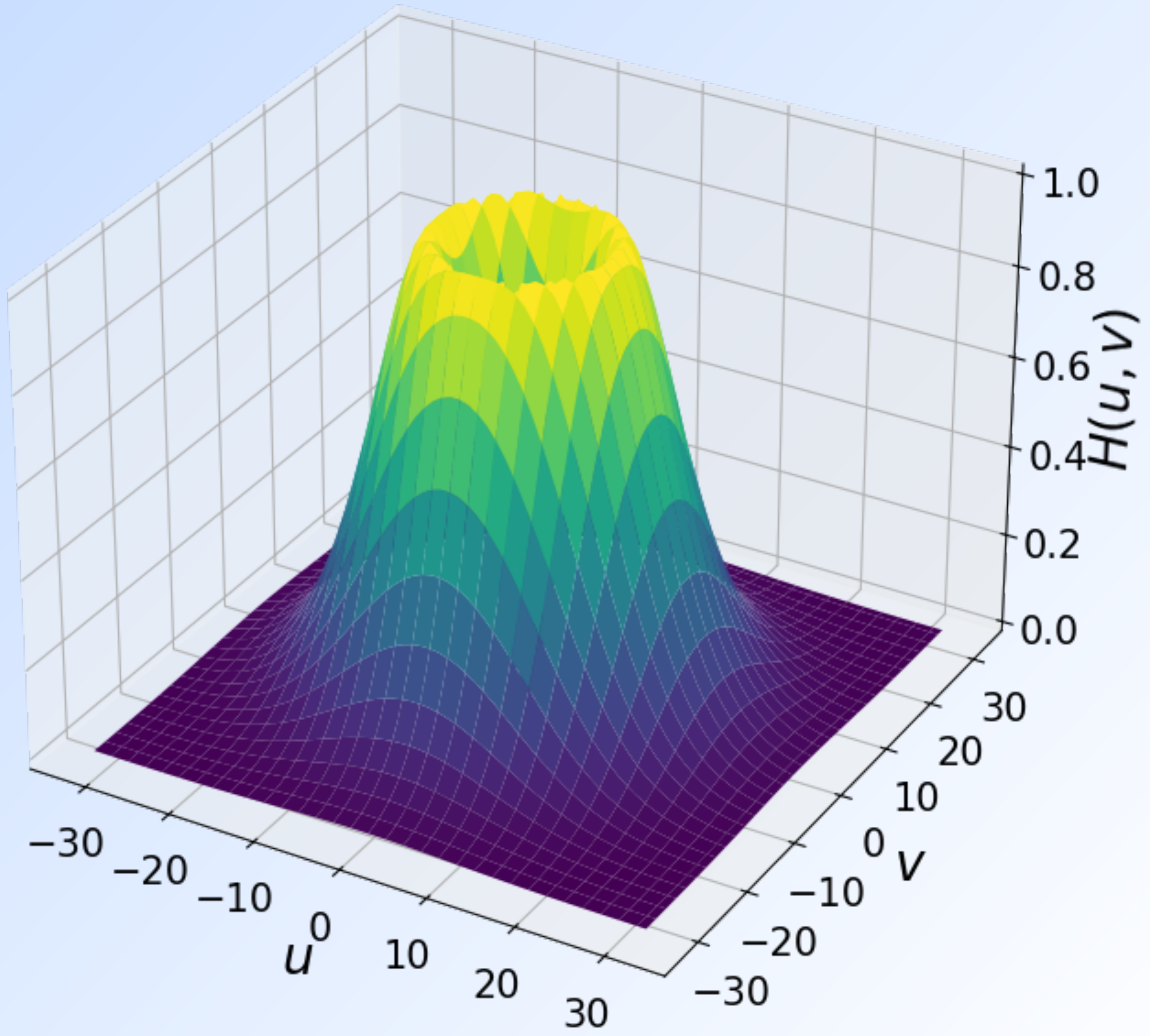


Gaussian high pass filter

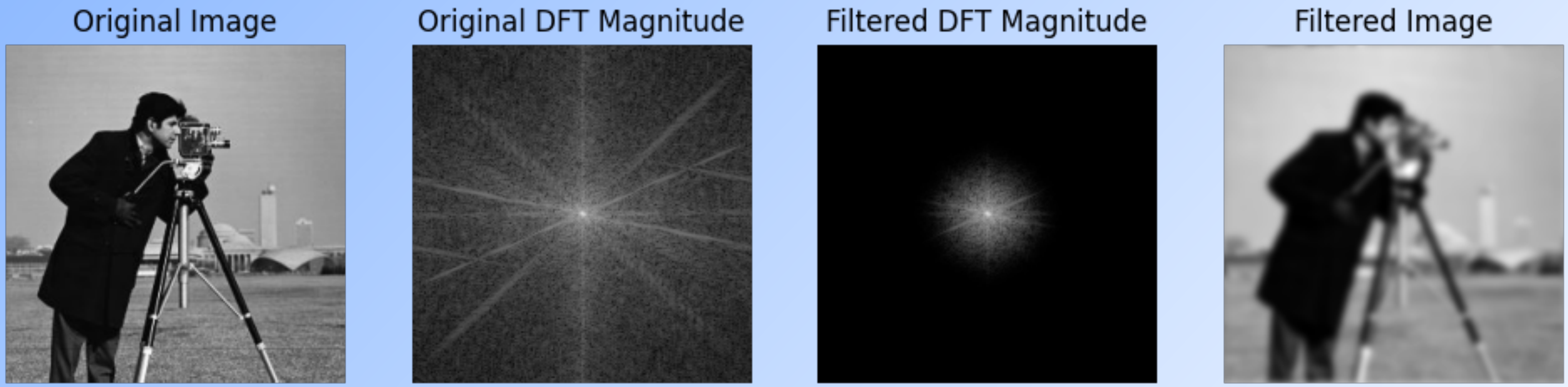
$$H(u,v) = 1 - e^{-\frac{(u-\mu_u)^2 + (v-\mu_v)^2}{2\sigma^2}}$$



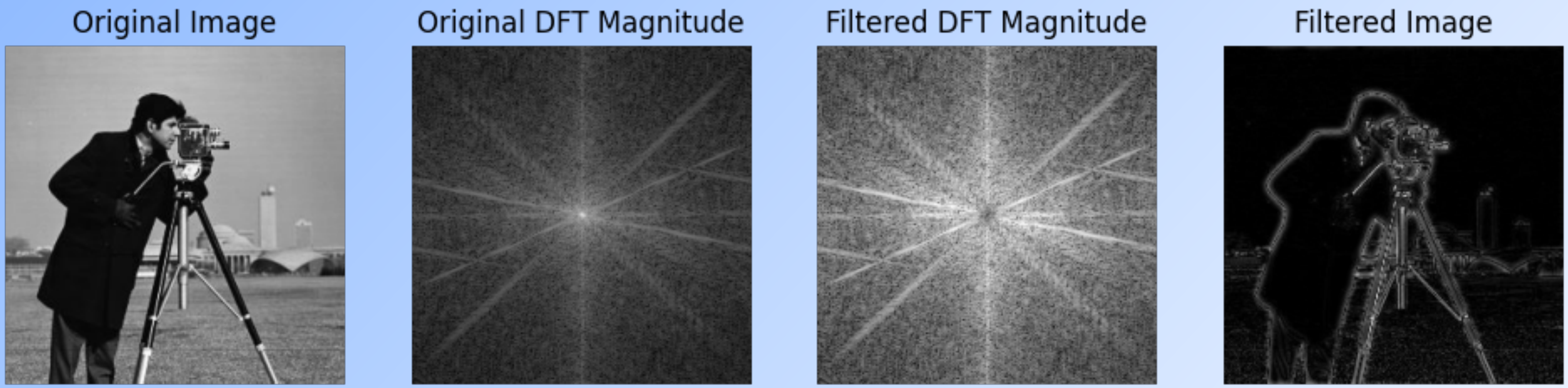
Gaussian band pass filter



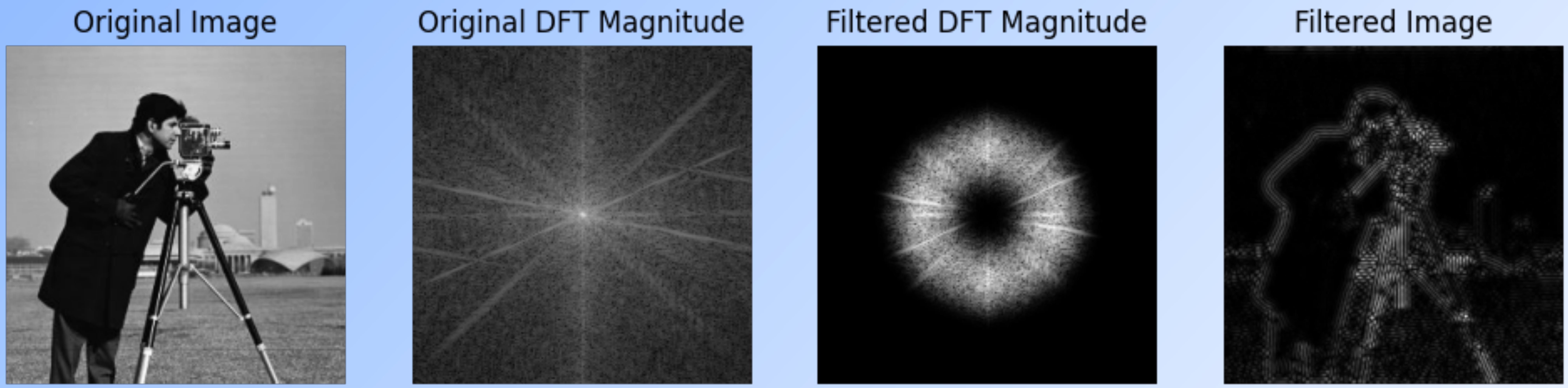
Gaussian Filters



Gaussian Filters



Gaussian Filters



Sampling and Aliasing: A Frequency Domain Perspective

- A 2D signal $f(x, y)$ could be sampled by multiplying it with a 2D periodic impulse train: $f_s(x, y) = f(x, y) \cdot \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - m\Delta T_x, y - n\Delta T_y)$, where ΔT_x and ΔT_y are the sampling periods along the horizontal and vertical directions.

- At each sample location, the value of the sampled function is the impulse weighted by the value of $f(x, y)$ at that location:
$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(x - k_1\Delta T_x, y - k_2\Delta T_y)dxdy = f(k_1\Delta T_x, k_2\Delta T_y)$ (Sifting property of impulse function)

Sampling and Aliasing: A Frequency Domain Perspective

- The Fourier transform of the sampled signal is given by:

$$\mathscr{F}(f_s(x,y)) = \mathscr{F}\left(f(x,y) \cdot \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - m\Delta T_x, y - n\Delta T_y)\right)$$

$$= \mathscr{F}\left(f(x,y)\right) * \mathscr{F}\left(\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - m\Delta T_x, y - n\Delta T_y)\right) \text{ (Using duality)}$$

$$= F(U,V) * \frac{1}{\Delta T_x \Delta T_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(U - \frac{m}{\Delta T_x}, V - \frac{n}{\Delta T_y})$$

$$= \frac{1}{\Delta T_x \Delta T_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} F(U - \frac{m}{\Delta T_x}, V - \frac{n}{\Delta T_y}) \text{ (Using sifting property of the impulse function)}$$

Sampling and Aliasing: A Frequency Domain Perspective

- Thus, sampling causes shifted versions of $F(U,V)$ i.e. $F(U-\frac{m}{\Delta T_x}, V-\frac{n}{\Delta T_y})$ to be infinitely repeated in the frequency domain, at intervals of $(\frac{1}{\Delta T_x}, \frac{1}{\Delta T_y})$ along the two frequency dimensions.
- Aliasing occurs when the repetitions overlap.
- Let the maximum frequencies present in the signal along the horizontal and vertical directions be U_{max} and V_{max}. Then to avoid aliasing, we must have:
$\frac{1}{\Delta T_x} - U_{max} > U_{max}$ and $\frac{1}{\Delta T_y} - V_{max} > V_{max}$.
- Let $f_{s_x} = \frac{1}{\Delta T_x}$ and $f_{s_y} = \frac{1}{\Delta T_y}$ be the sampling frequencies in the horizontal and vertical directions.
- Thus, by analyzing the spectrum of the sampled signal in the frequency domain, we derived the result of the Nyquist Theorem, i.e to avoid aliasing, $f_{s_x} \geq 2U_{max}$ and $f_{s_y} \geq 2V_{max}$.