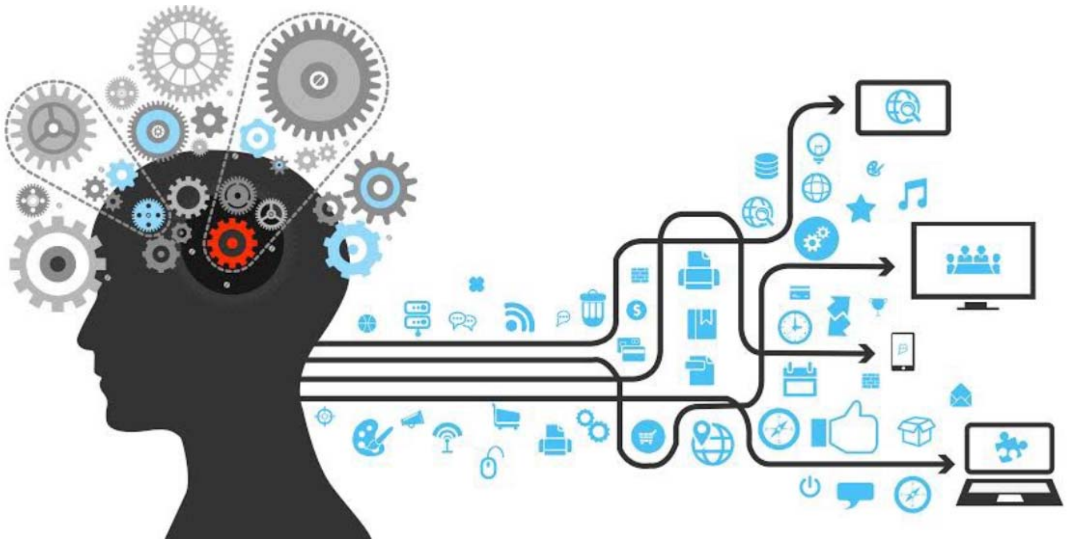
Computer Vision

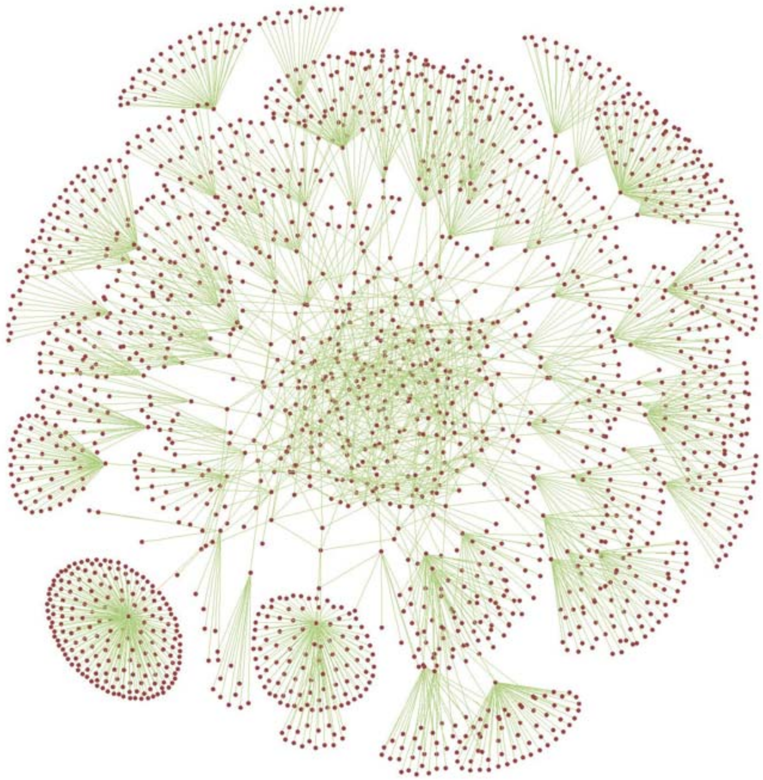
Early vision: Just one image



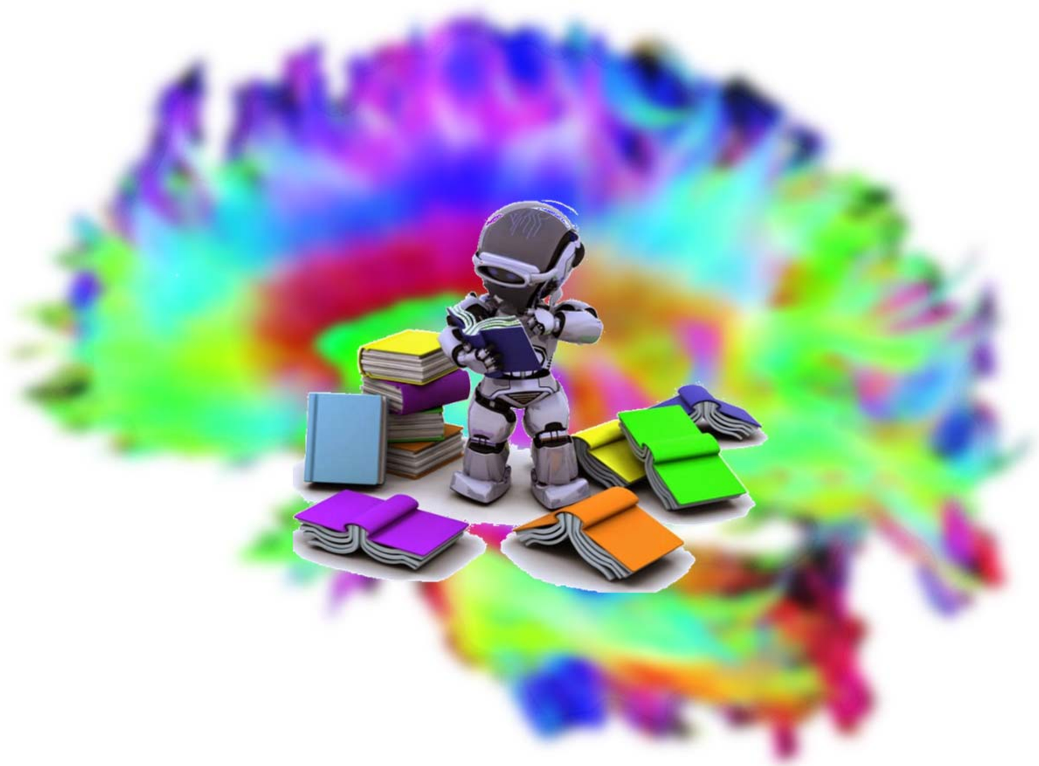
School of Electronic &
Electrical Engineering

Sungkyunkwan University

Hyunjin Park

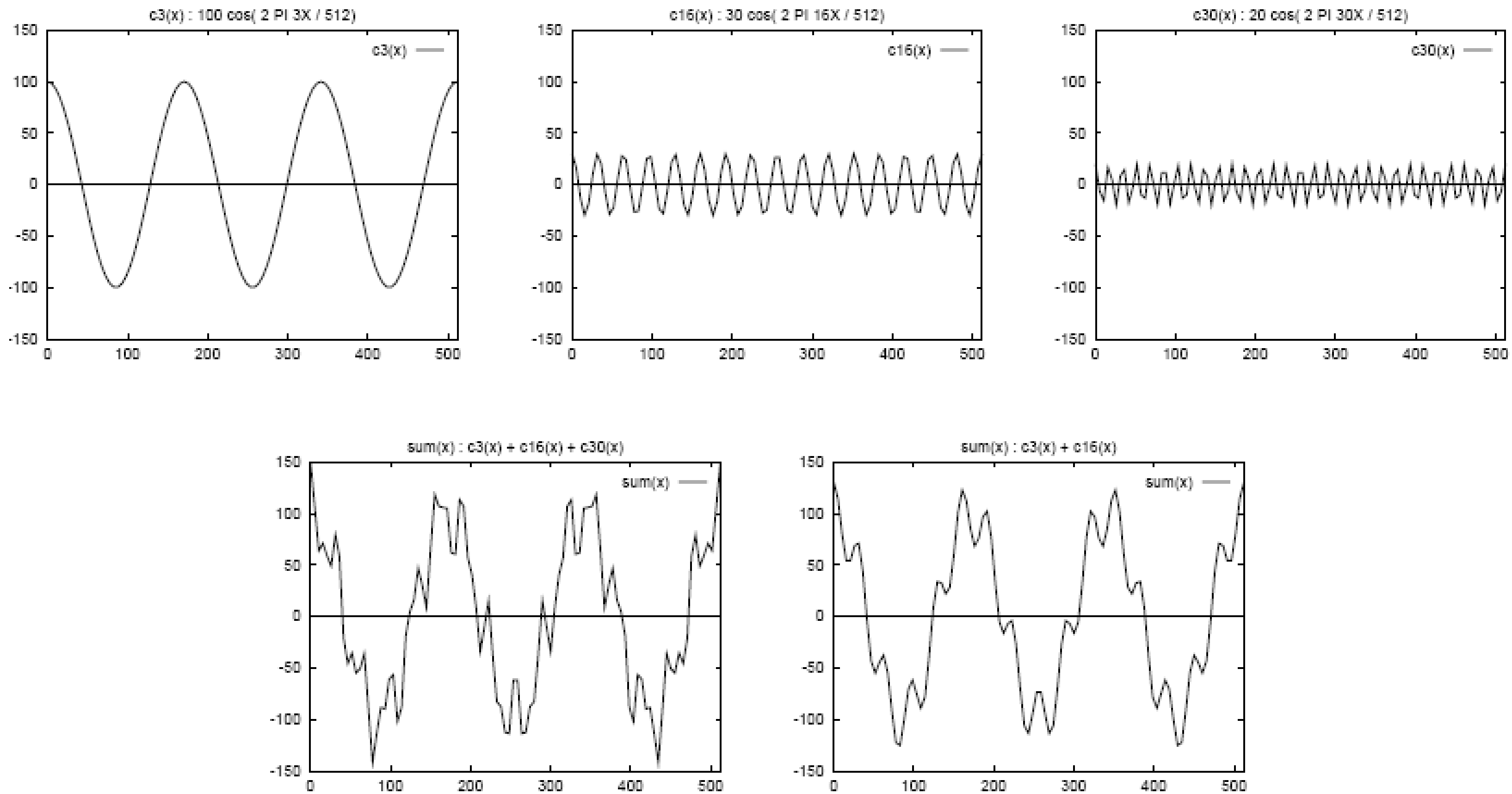


Spatial Frequency Analysis



Combination of signals

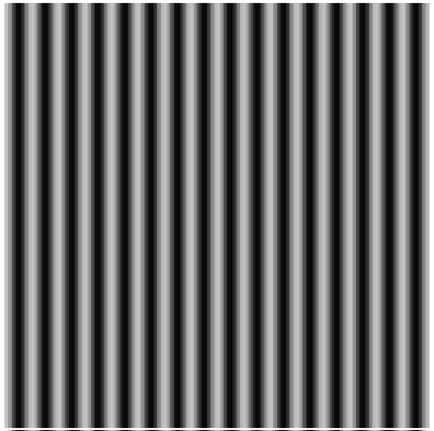
- A signal: Combination of sinusoids with different frequencies



Combination of signals

- Constructing an image: Combination of basis images

Basis images



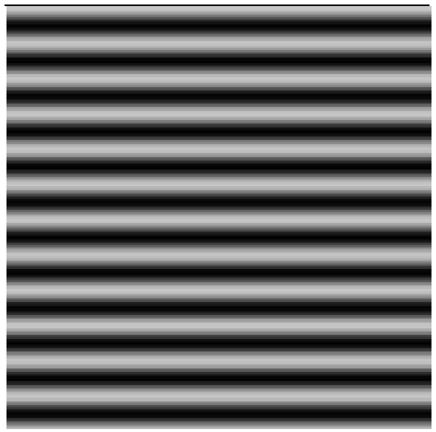
E1

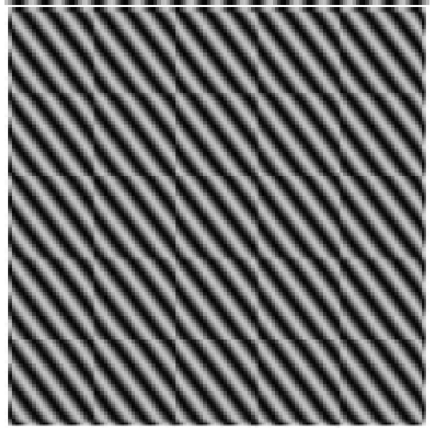
$$E1 = 100\cos\left(2\pi\frac{16}{512}x\right) + 100$$

$$E2 = 100\cos\left(2\pi\frac{12}{512}x\right) + 100$$

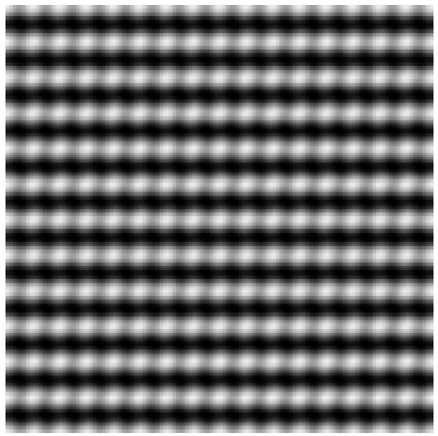
$$E3 = 100\cos\left(2\pi\left(\frac{16}{512}x + \frac{12}{512}x\right)\right) + 100$$

E2



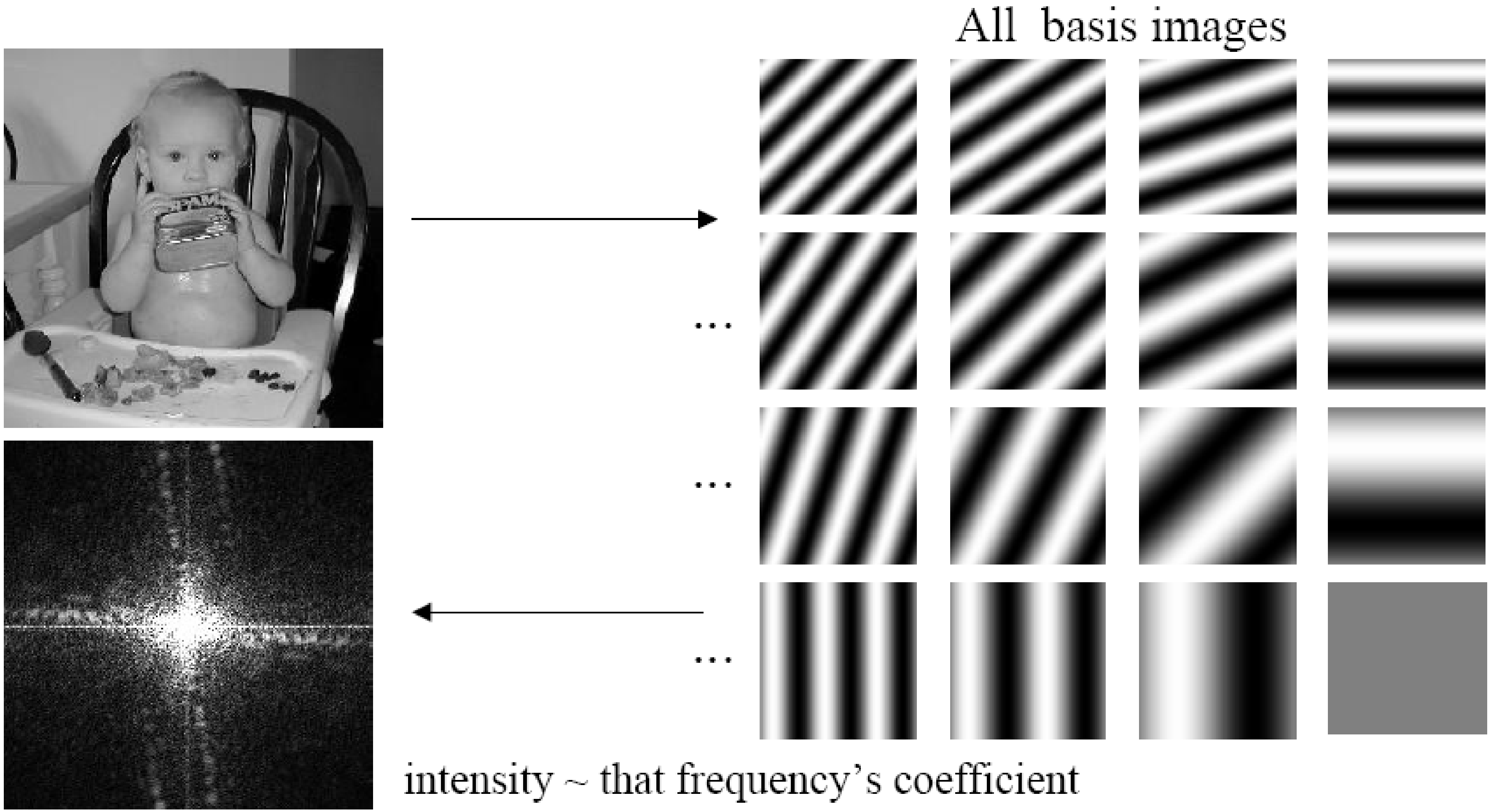


E3



30E1 + 100E2 + 10E3

Analyzing an image



Fourier

- Fourier series
  - Periodic signals can be expressed as a sum of complex exponentials
  - Periodic signal with T_0

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{jw_0 nt}$$

$$X_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{-jw_0 nt} dt$$

- Fourier transform
  - Generalization to nonperiodic signals
  - Periodic signal with $T_0 = \infty$

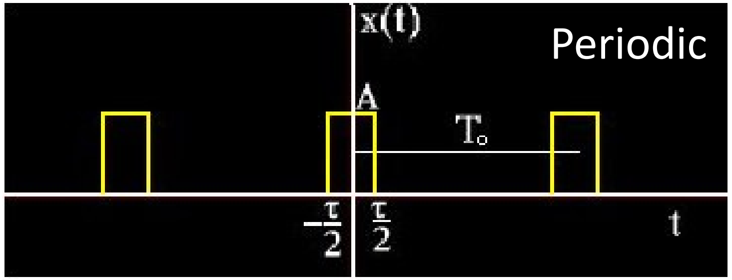
$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(w) e^{jwt}\, dw$$

$$X(w) = \int_{-\infty}^{\infty} x(t)\, e^{-jwt} dt$$

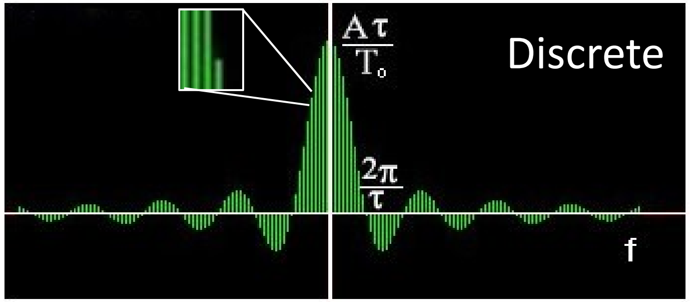
Fourier

- Fourier series

Time

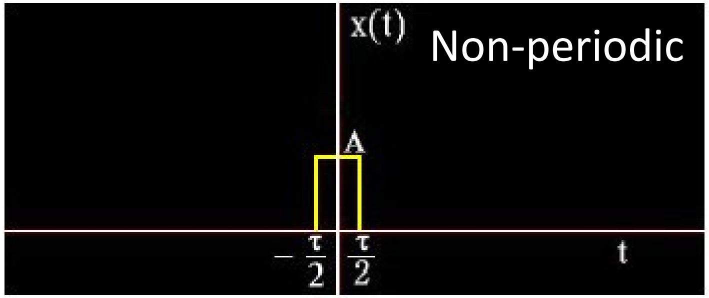


Frequency



- Fourier transform

Time

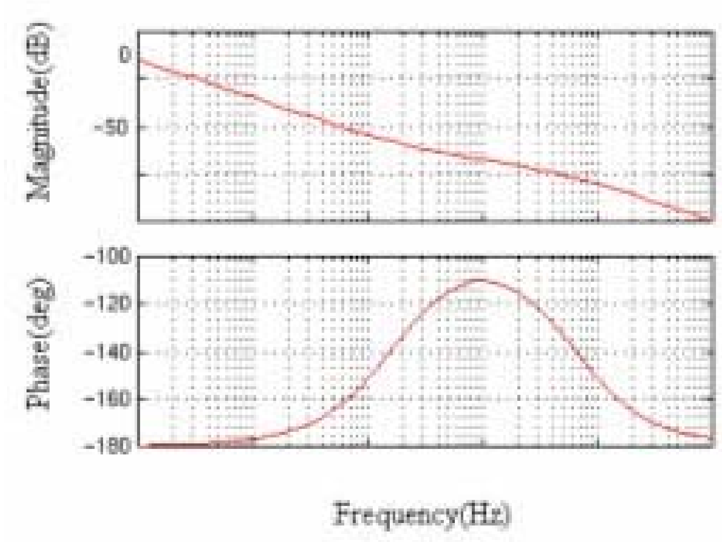


Frequency

Magnitude & Phase

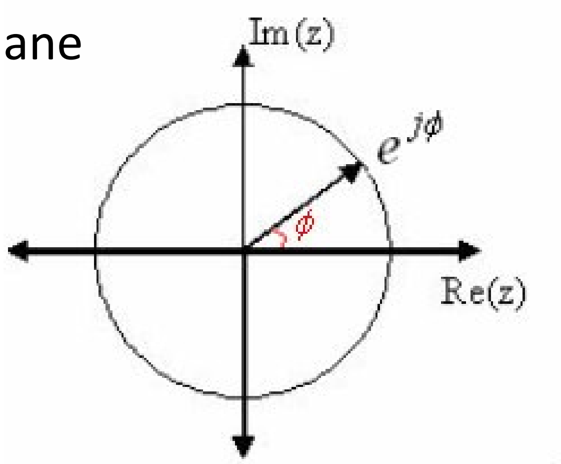
- Fourier transform of a real function is complex
  - Difficult to plot and visualize
  - Instead, think of the magnitude and phase of the transform
- Signal = magnitude + phase
- Magnitude: Magnitude of the complex transform
- Phase: Phase of the complex transform

$$X(w) = |X(w)| \cdot e^{j\varphi} \Rightarrow \begin{cases} 20\log_{10}|X(w)| \\ \varphi \end{cases}$$



Magnitude & Phase

Z-plane

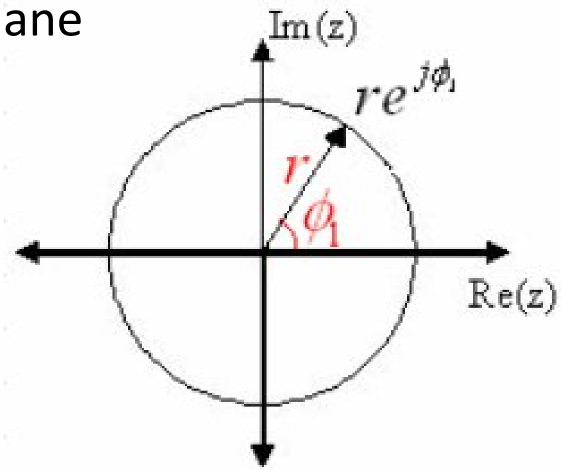


$$e^{j\phi} = \cos\phi + j\sin\phi$$

$$\text{Magnitude} = \sqrt{\cos^2\phi + \sin^2\phi} = 1$$

$$\text{Phase} = \tan^{-1}\frac{\sin\phi}{\cos\phi} = \phi$$

Z-plane



$$re^{j\phi_1} = r\cos\phi_1 + jr\sin\phi_1$$

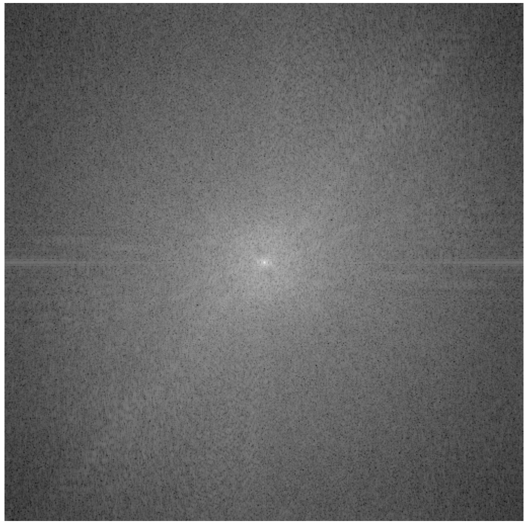
$$\text{Magnitude} = r$$

$$\text{Phase} = \tan^{-1}\frac{r\sin\phi_1}{r\cos\phi_1} = \phi_1$$

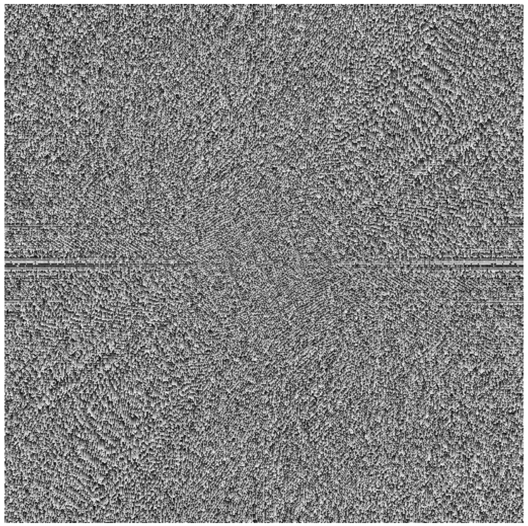
Magnitude & Phase

- Curious fact
  - All natural images have about the same magnitude transform
  - Hence, phase seems to matter, but magnitude largely doesn't
- Demonstration
  - Take two pictures, swap the phase transforms, compute the inverse
  - What does the result look like?

Magnitude & Phase

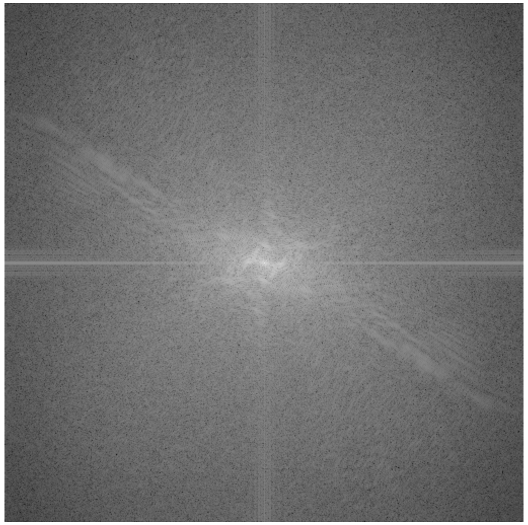


Magnitude transform
of the picture

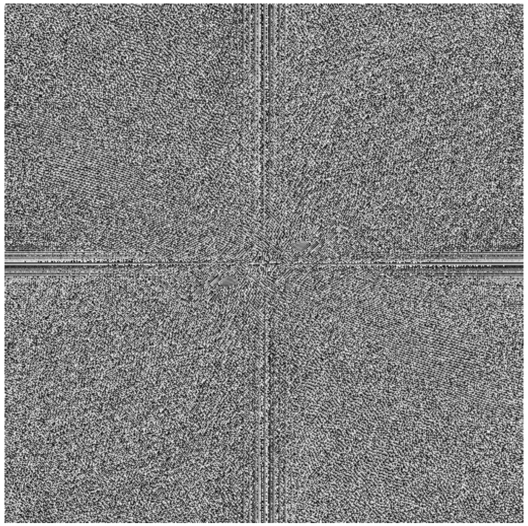


Phase transform
of the picture

Magnitude & Phase



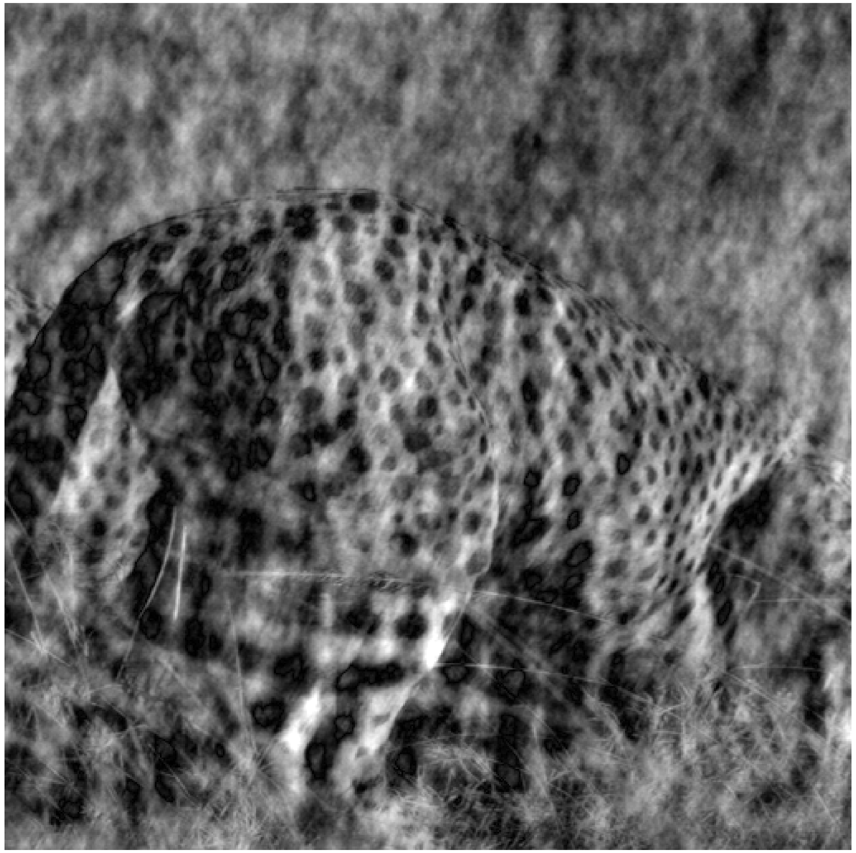
Magnitude transform
of the picture



Phase transform
of the picture

Magnitude & Phase

Reconstruction with
zebra phase and cheetah magnitude



Reconstruction with
cheetah phase and zebra magnitude

Discrete-time signal

- Discrete-time Fourier series (DFS or DTFS)
  - Discrete-time version of the Fourier series
  - Applicable to periodic signals
  - Produce a discrete frequency spectrum

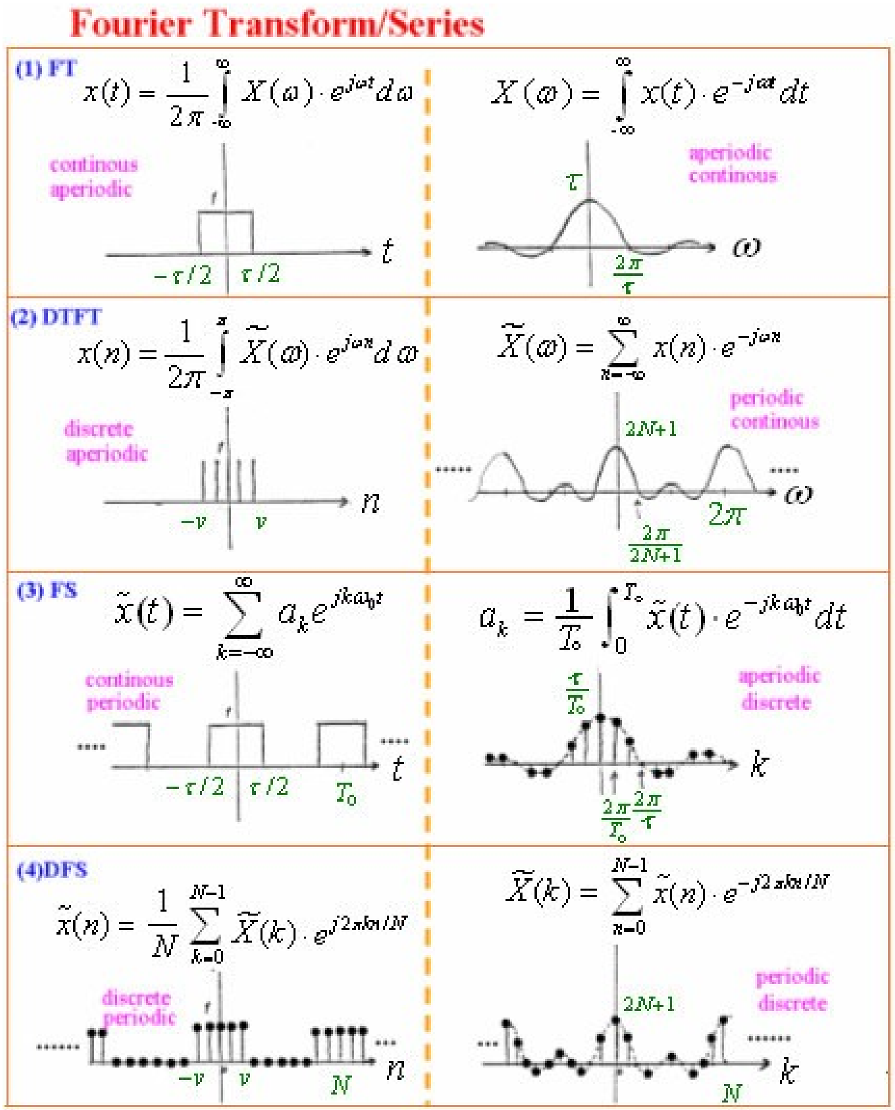
- Discrete-time Fourier transform (DFT or DTFT)
  - Discrete-time version of the Fourier transform
  - Applicable to aperiodic signals
  - Produce a continuous frequency spectrum

Discrete-time signal

- DFS $\xrightarrow{one\ period}$ DFT

- DFT is not only discrete for both time and frequency domain but also a finite sequence

- Most of data is given as an aperiodic discrete-time signal



2D DFT properties

Property	Expression(s)
Fourier transform	$F(u,v) = \dfrac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)}$
Inverse Fourier transform	$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi(ux/M+vy/N)}$
Polar representation	$F(u,v) = \lvert F(u,v)\rvert e^{-j\phi(u,v)}$
Spectrum	$\lvert F(u,v)\rvert = \left[R^2(u,v) + I^2(u,v)\right]^{1/2}$, $\quad R = \text{Real}(F)$ and $I = \text{Imag}(F)$
Phase angle	$\phi(u,v) = \tan^{-1}\left[\dfrac{I(u,v)}{R(u,v)}\right]$
Power spectrum	$P(u,v) = \lvert F(u,v)\rvert^2$
Average value	$\bar{f}(x,y) = F(0,0) = \dfrac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)$
Translation	$f(x,y)e^{j2\pi(u_0x/M+v_0y/N)} \Leftrightarrow F(u-u_0, v-v_0)$ $f(x-x_0, y-y_0) \Leftrightarrow F(u,v)e^{-j2\pi(ux_0/M+vy_0/N)}$ When $x_0 = u_0 = M/2$ and $y_0 = v_0 = N/2$, then $f(x,y)(-1)^{x+y} \Leftrightarrow F(u-M/2, v-N/2)$ $f(x-M/2, y-N/2) \Leftrightarrow F(u,v)(-1)^{u+v}$

2D DFT properties

Conjugate symmetry	$F(u,v) = F^*(-u,-v)$ $\lvert F(u,v)\rvert = \lvert F(-u,-v)\rvert$
Differentiation	$\dfrac{\partial^n f(x,y)}{\partial x^n} \Leftrightarrow (ju)^n F(u,v)$ $(-jx)^n f(x,y) \Leftrightarrow \dfrac{\partial^n F(u,v)}{\partial u^n}$
Laplacian	$\nabla^2 f(x,y) \Leftrightarrow -(u^2+v^2)F(u,v)$
Distributivity	$\Im[f_1(x,y)+f_2(x,y)] = \Im[f_1(x,y)] + \Im[f_2(x,y)]$ $\Im[f_1(x,y)\cdot f_2(x,y)] \neq \Im[f_1(x,y)]\cdot\Im[f_2(x,y)]$
Scaling	$af(x,y) \Leftrightarrow aF(u,v),\, f(ax,by) \Leftrightarrow \dfrac{1}{\lvert ab\rvert}F(u/a, v/b)$
Rotation	$x = r\cos\theta \qquad y = r\sin\theta \qquad u = \omega\cos\varphi \qquad v = \omega\sin\varphi$ $f(r,\theta+\theta_0) \Leftrightarrow F(\omega,\varphi+\theta_0)$
Periodicity	$F(u,v) = F(u+M,v) = F(u,v+N) = F(u+M,v+N)$ $f(x,y) = f(x+M,y) = f(x,y+N) = f(x+M,y+N)$
Separability	See Eqs. (4.6-14) and (4.6-15). Separability implies that we can compute the 2-D transform of an image by first computing 1-D transforms along each row of the image, and then computing a 1-D transform along each column of this intermediate result. The reverse, columns and then rows, yields the same result.

2D DFT properties

Property	Expression(s)
Computation of the inverse Fourier transform using a forward transform algorithm	$\frac{1}{MN} f^*(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F^*(u, v) e^{-j2\pi(ux/M + vy/N)}$ This equation indicates that inputting the function $F^*(u, v)$ into an algorithm designed to compute the forward transform (right side of the preceding equation) yields $f^*(x, y)/MN$. Taking the complex conjugate and multiplying this result by MN gives the desired inverse.
Convolution†	$f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x - m, y - n)$
Correlation†	$f(x, y) \circ h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f^*(m, n) h(x + m, y + n)$
Convolution theorem†	$f(x, y) * h(x, y) \Leftrightarrow F(u, v) H(u, v)$; $f(x, y) h(x, y) \Leftrightarrow F(u, v) * H(u, v)$
Correlation theorem†	$f(x, y) \circ h(x, y) \Leftrightarrow F^*(u, v) H(u, v)$; $f^*(x, y) h(x, y) \Leftrightarrow F(u, v) \circ H(u, v)$

2D DFT properties

Some useful FT pairs:

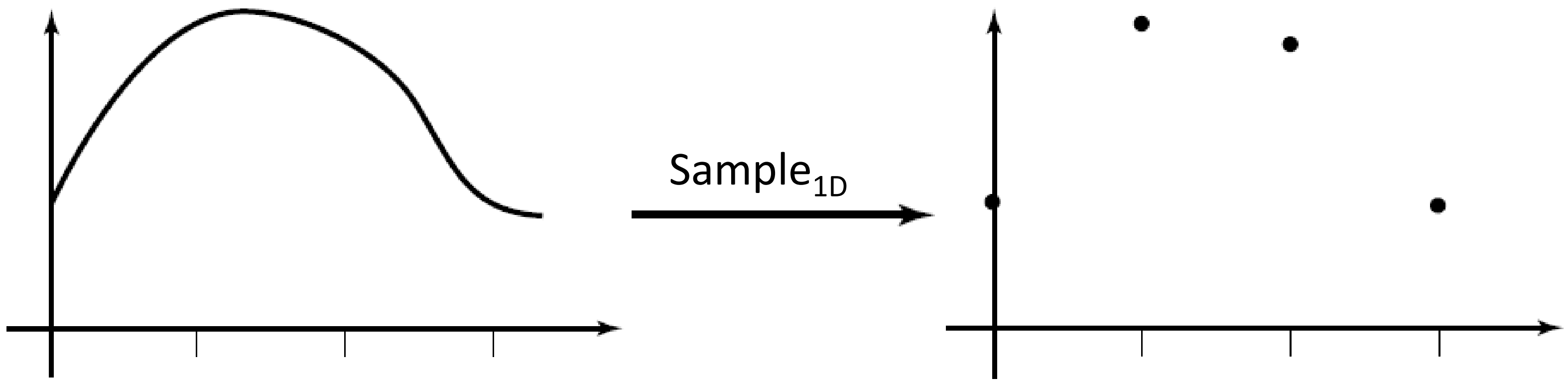
Impulse	$\delta(x, y) \Leftrightarrow 1$
Gaussian	$A\sqrt{2\pi}\sigma e^{-2\pi^2\sigma^2(x^2+y^2)} \Leftrightarrow Ae^{-(u^2+v^2)/2\sigma^2}$
Rectangle	$\mathrm{rect}[a, b] \Leftrightarrow ab\dfrac{\sin(\pi ua)}{(\pi ua)}\dfrac{\sin(\pi vb)}{(\pi vb)}e^{-j\pi(ua+vb)}$
Cosine	$\cos(2\pi u_0 x + 2\pi v_0 y) \Leftrightarrow \dfrac{1}{2}\left[\delta(u + u_0, v + v_0) + \delta(u - u_0, v - v_0)\right]$
Sine	$\sin(2\pi u_0 x + 2\pi v_0 y) \Leftrightarrow j\dfrac{1}{2}\left[\delta(u + u_0, v + v_0) - \delta(u - u_0, v - v_0)\right]$

† Assumes that functions have been extended by zero padding.

Sampling

- Go from continuous world to discrete world

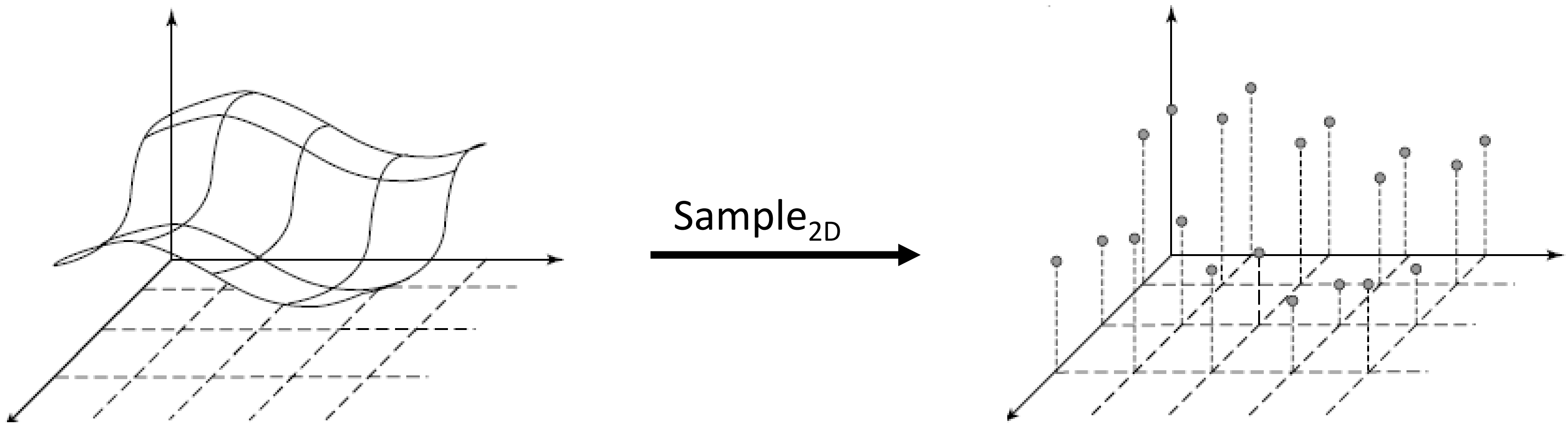
- Samples are typically measured on regular grid



- Sampling in 1D takes a continuous function and replaces it with a vector of values, consisting of the function's values at a set of sample points

Sampling

- Sampling in 2D does the same thing, only in 2D



Sampling of continuous signal

- In general, some information is lost as a result of sampling

- If the continuous-time signal is band-limited and the sampling rate is sufficiently high ($f_s \geq 2f_{max}$), little or no information loss will occur

Sampling of continuous signal

$$x_a(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_a(j\Omega)e^{j\Omega t}d\Omega$$

$$X_a(j\Omega) = \int_{-\infty}^{\infty} x_a(t)e^{-j\Omega t}dt$$

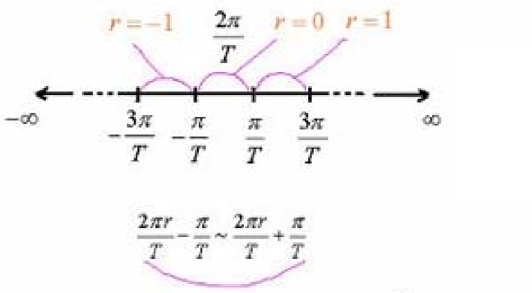
Periodic sampling of $x_a(t)$ with period of T

$$\begin{tikzcd} & x(n) = x_a(nT) \arrow[dl, "\text{F.T.}"'] \arrow[dr, "\text{F.T.}"] & \\ X\left(e^{jw}\right) \arrow[rr, leftrightarrow] & & X_a(j\Omega) \end{tikzcd}$$

How they related?

Sampling of continuous signal

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\boxed{X(e^{jw})}e^{jwn}\,dw = x(n) = x_a(nT) = \frac{1}{2\pi}\int_{-\infty}^{\infty}X_a(j\Omega)e^{j\Omega nt}\,d\Omega$$



, which has the interval of $\frac{2\pi}{T}$ for all r

$$= \frac{1}{2\pi}\boxed{\sum_{r=-\infty}^{\infty}\int_{(2r-1)\pi/T}^{(2r+1)\pi/T}}X_a(j\Omega)e^{j\Omega nt}\,d\Omega$$

Let $\Omega \to \Omega + \frac{2\pi r}{T}$

$$= \frac{1}{2\pi}\sum_{r=-\infty}^{\infty}\int_{-\pi/T}^{\pi/T}X_a\left(j\Omega + j\frac{2\pi r}{T}\right)e^{j\Omega nt + j2\pi nr}\,d\Omega$$

$$= \frac{1}{2\pi}\sum_{r=-\infty}^{\infty}\int_{-\pi/T}^{\pi/T}X_a\left(j\Omega + j\frac{2\pi r}{T}\right)e^{j\Omega nt}\,d\Omega$$

Let $\Omega = w/T, \because w = \Omega T$

$$= \frac{1}{2\pi}\sum_{r=-\infty}^{\infty}\int_{-\pi}^{\pi}X_a\left(j\frac{w}{T} + j\frac{2\pi r}{T}\right)e^{jnw}\frac{1}{T}\,dw$$

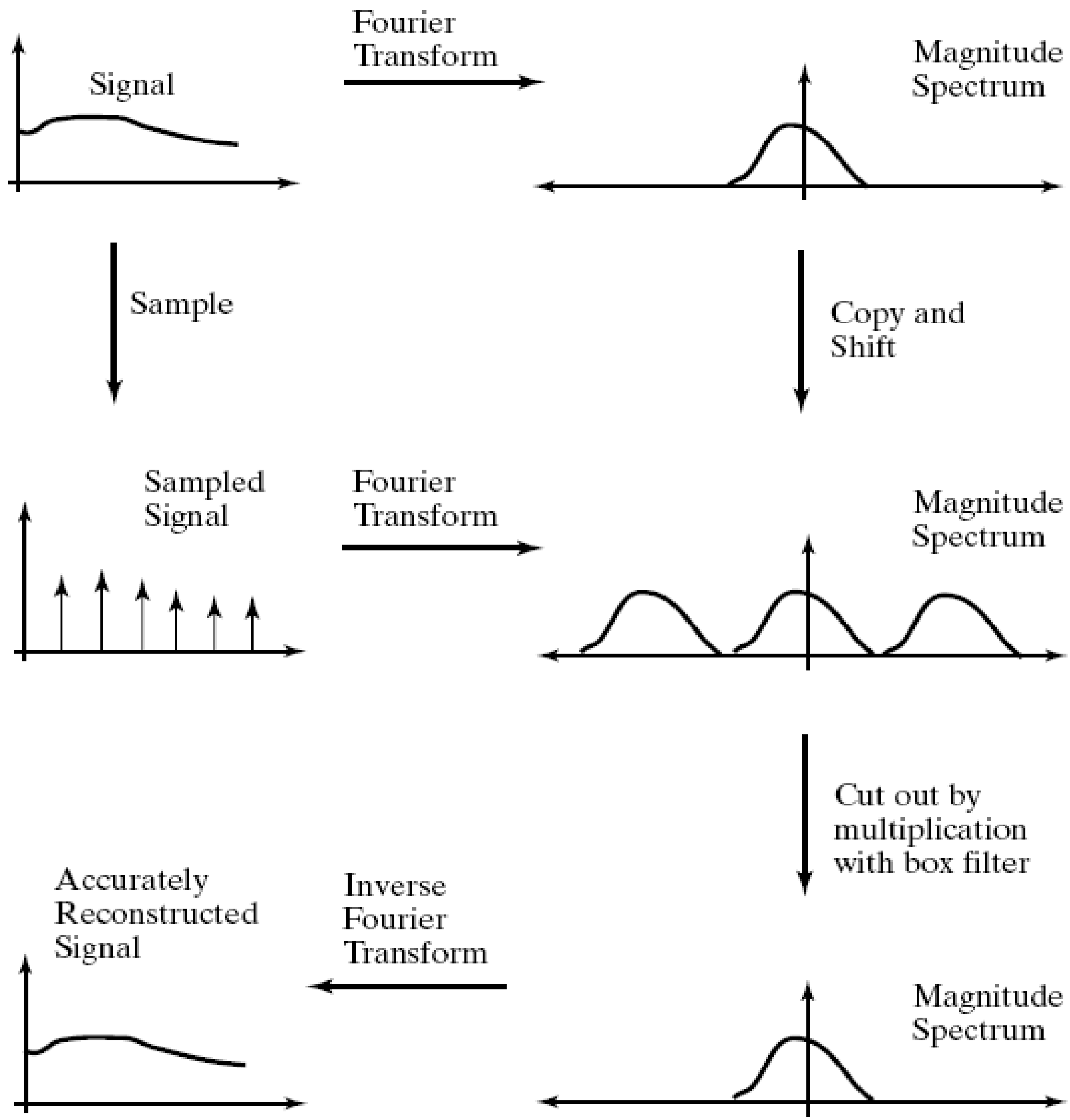
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\boxed{\frac{1}{T}\sum_{r=-\infty}^{\infty}X_a\left(j\frac{w}{T} + j\frac{2\pi r}{T}\right)}e^{jnw}\,dw$$

Sampling of continuous signal

- The relationship of Fourier transform of digital signal and that of analog signal

$$X(e^{jw}) = \frac{1}{T}\sum_{r=-\infty}^{\infty} X_a\left(j\frac{w}{T} + j\frac{2\pi r}{T}\right)$$

Signal reconstruction



Signal reconstruction

Reconstruction of $x_a(t)$ from $x(n)$

$X(e^{jw}) = \frac{1}{T} X_z(j\Omega)$ for $-\frac{\pi}{T} \le \Omega \le \frac{\pi}{T}$

$$x_a(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_a(j\Omega) e^{j\Omega t} d\Omega$$

$w = \Omega T$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} T \sum_{k=-\infty}^{\infty} x(n) e^{-j\boxed{w}n}\, e^{j\Omega t} d\Omega$$

$$= \sum_{n=-\infty}^{\infty} x(n) \left[\frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} e^{j\Omega(t-nT)} d\Omega\right]$$

$$= \sum_{n=-\infty}^{\infty} x(n) \frac{\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)}$$

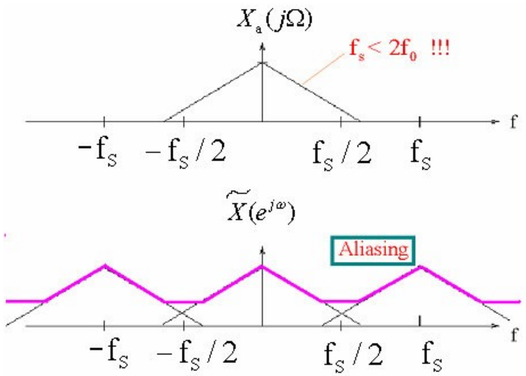
$$\boxed{x_a(t) = \sum_{n=-\infty}^{\infty} x(n) * sinc(t-nT)}$$

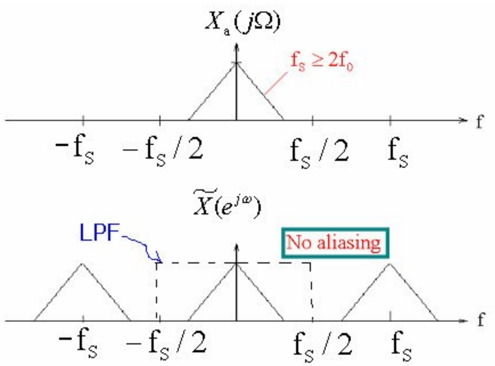
Convolution $x(n)$ with $sinc$ function

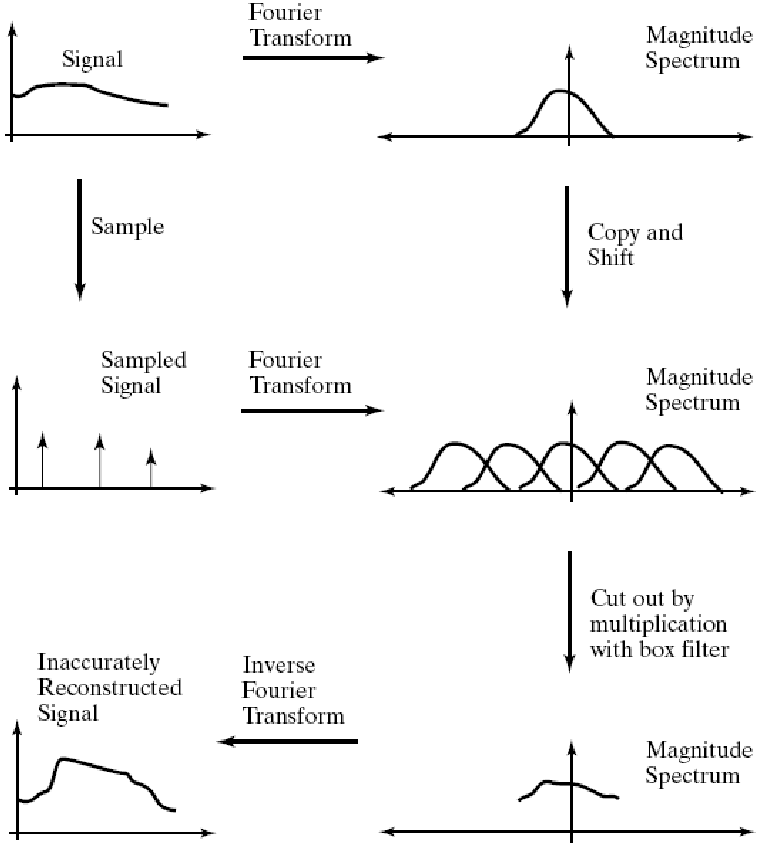
, where $x(n) = x_a(nT)$, $sinc(t-nT) = \frac{\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)}$

Aliasing

- Nyquist's theorem
  - ***The sampling frequency must be at least twice the highest frequency*** present for a signal to be reconstructed from a sampled version
- $f_s \geq 2f_0$

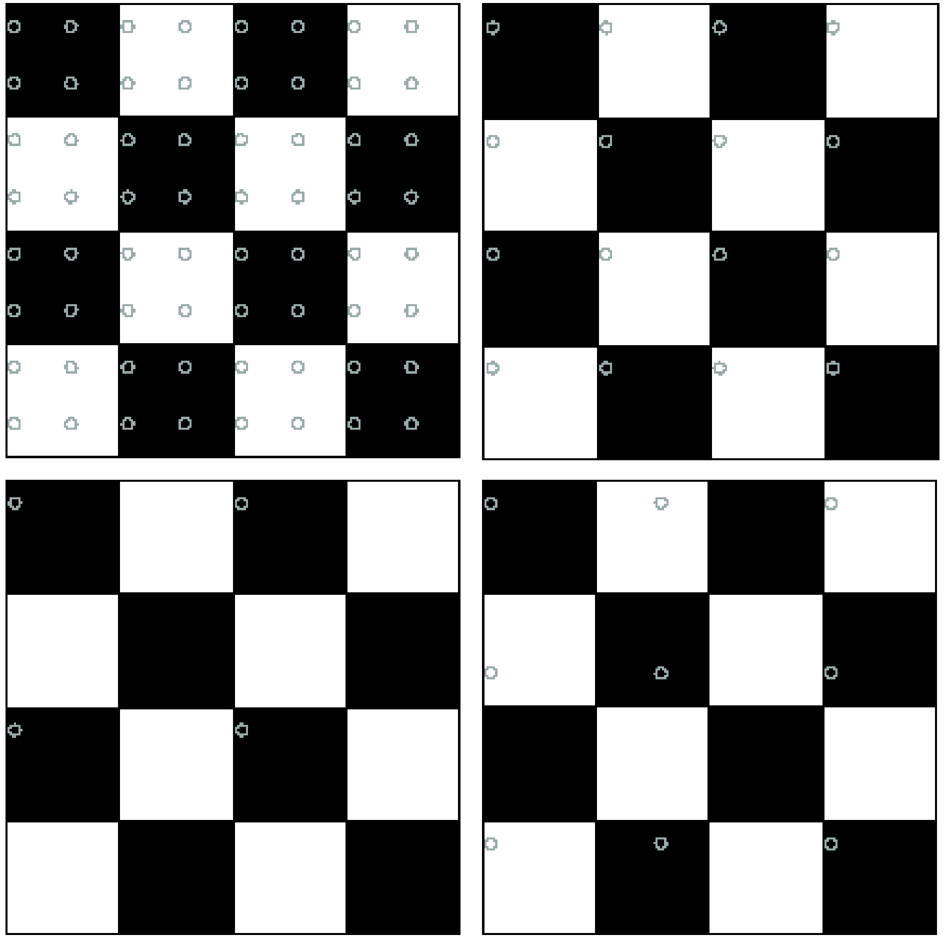






Loss of information

- Resample the checkerboard by taking one sample at each circle
  - Top left: New representation is reasonable
  - Top right: Also yields a reasonable representation
  - Bottom left: All black
  - Bottom right: Checks are too big



Smoothing filters in frequency domain

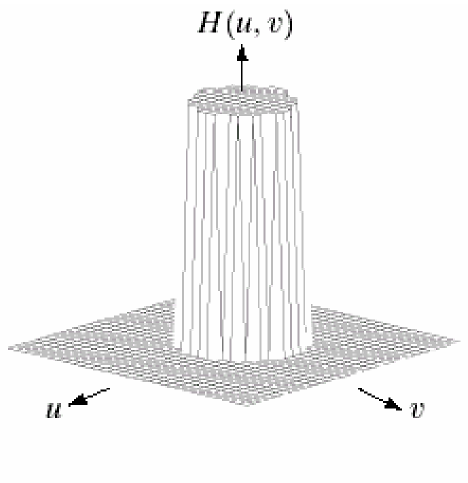
- High frequency: Edge, noise, sharp transition in intensity
- Low-pass filter (LPF): Attenuate high frequency component
- Spatial domain: Blur, take average

$$G(u,v) = H(u,v) \cdot F(u,v)$$

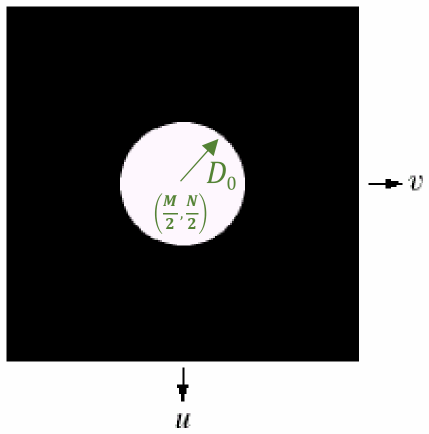
- Filter
  - Ideal: Very sharp transition
  - Butterworth: Its sharpness depends on the filter order
  - Gaussian: Very smooth transition
  - Chebyshev 1,2
  - Elliptic

Non-linear phase
∴ Usually not used for image filtering

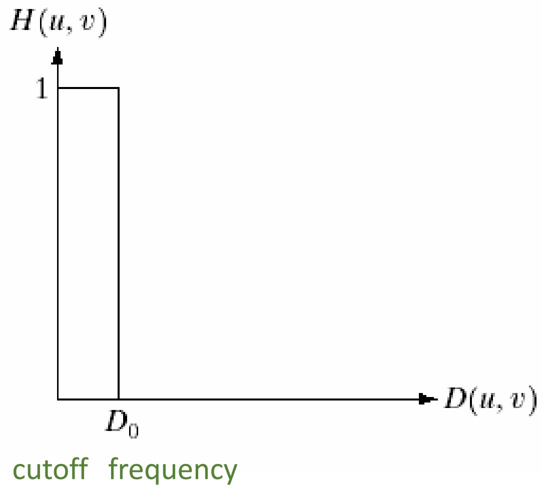
Ideal LPF (ILPF)



ILPF transfer function



Filter displayed as an image



Filter radial cross section

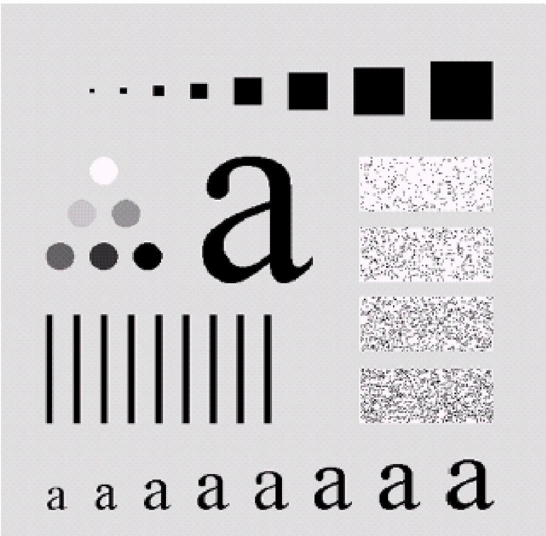
Distance from any point (u, v) to the center: $D(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}$

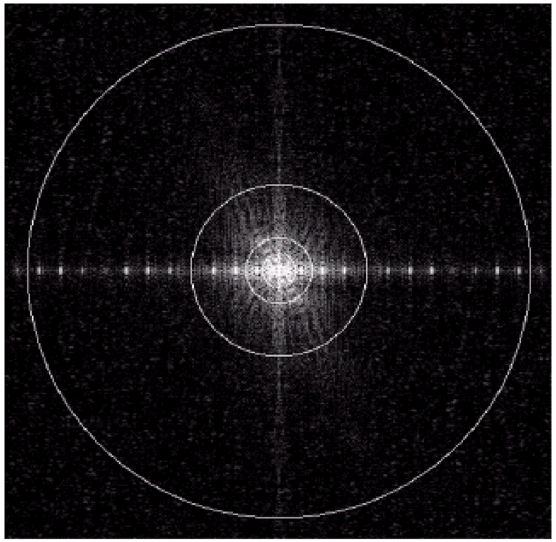
Total image power: $P_T = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v)$

Specifying the cutoff with alpha of power: $\alpha[\%] = \frac{\sum P(u,v)}{P_T}$

Ideal LPF (ILPF)

Fourier transform





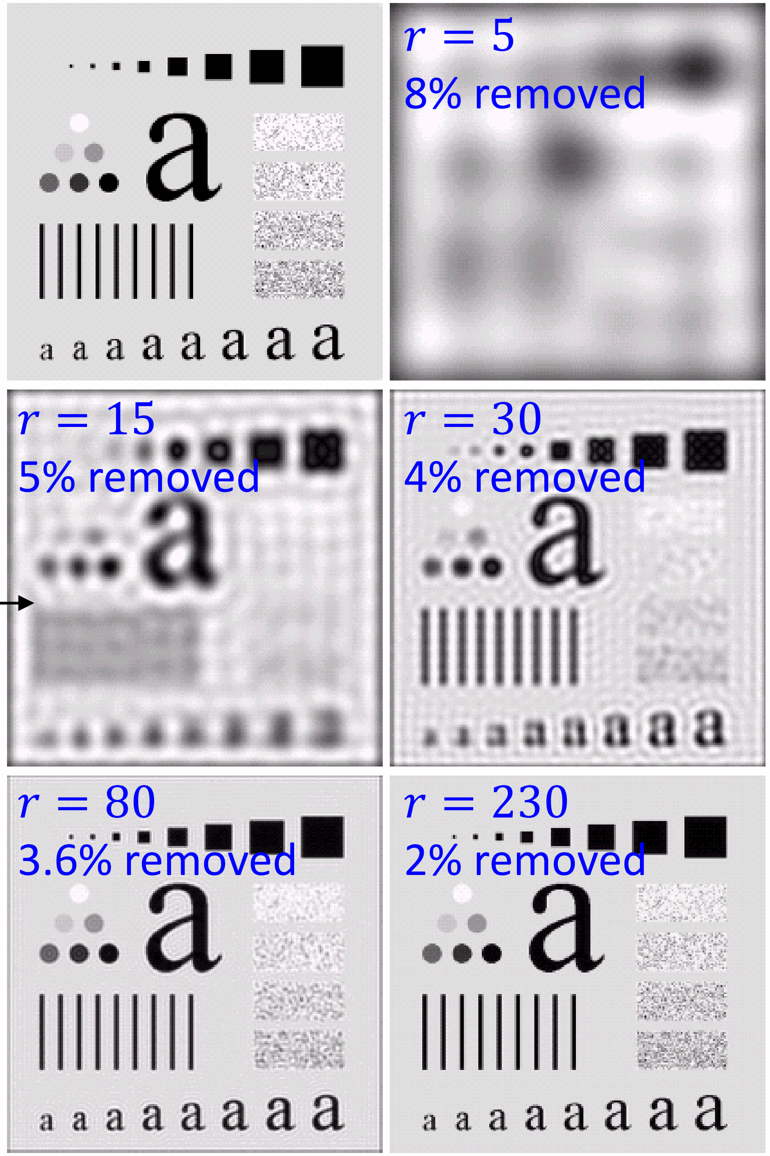
$r = \boxed{5,} \ldots, 230$

$\alpha = \boxed{92,} \ldots, 99.5\%$

Only 8% of P_T was removed

Ideal LPF (ILPF)

ringing

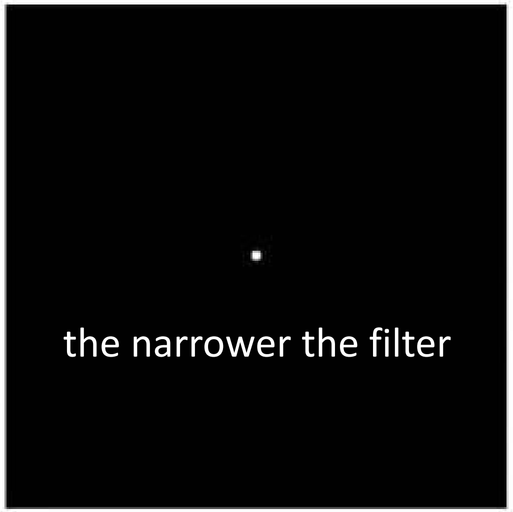


Most of the sharp detail information is gone in the 8% power removed

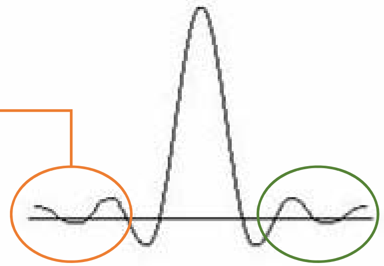
Ideal LPF (ILPF)

Frequency domain

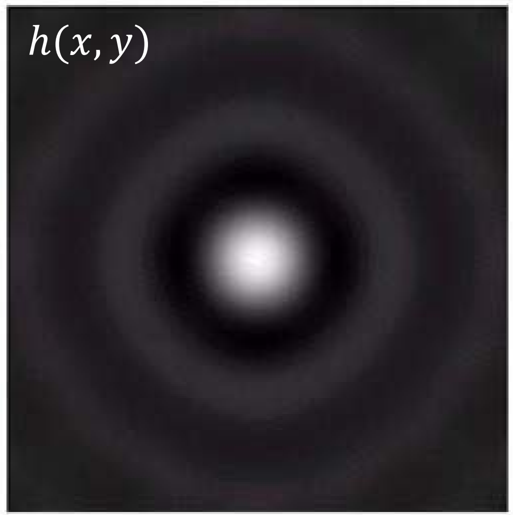
Spatial domain



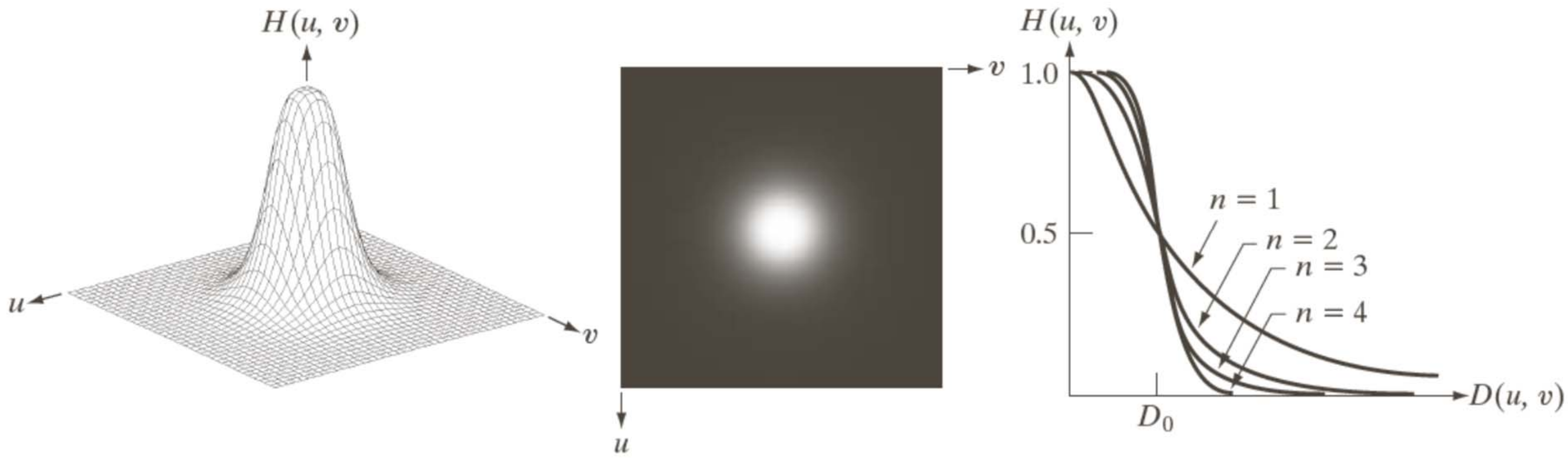
negative
∴ scaling is required
after filtering (convolution)



$$IDFT\big(H(u,v)\big)$$



Butterworth LPF (BLPF)



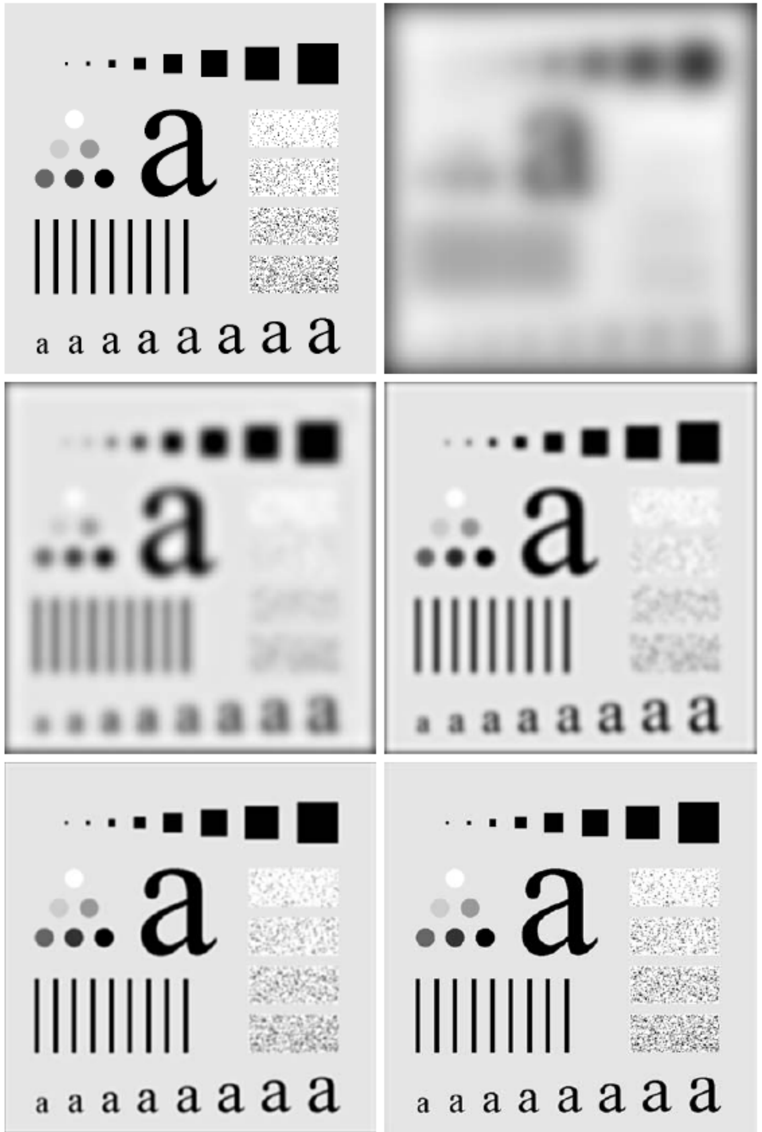
BLPF transfer function   Filter displayed as an image   Filter radial cross section

Transfer function of BLPF of order n, and with cutoff frequency at a distance D_0 from the origin

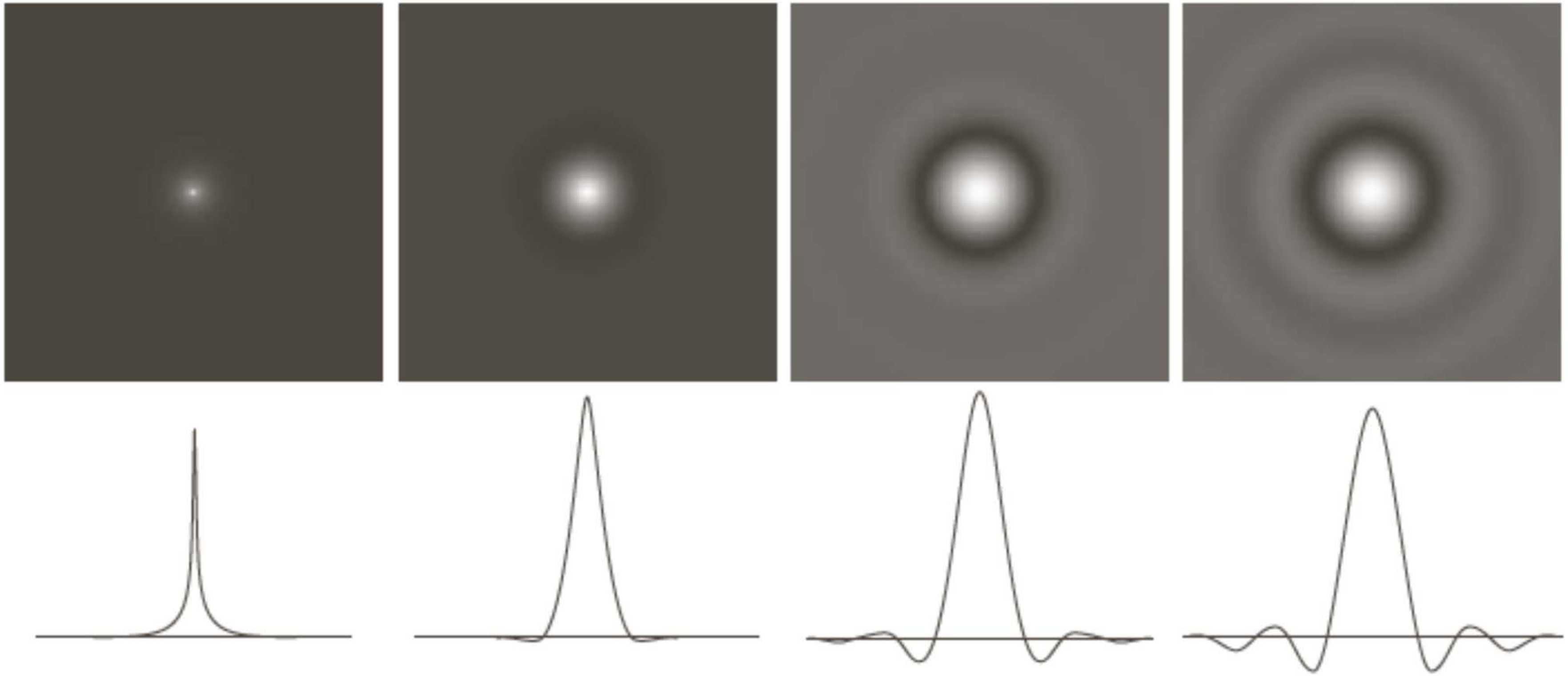
$$H(u,v) = \frac{1}{1 + \left[\frac{D(u,v)}{D_0}\right]^{2n}}$$

Butterworth LPF (BLPF)

less
ringing

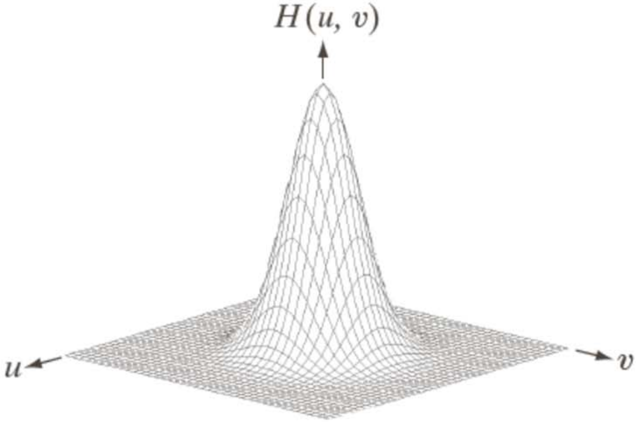


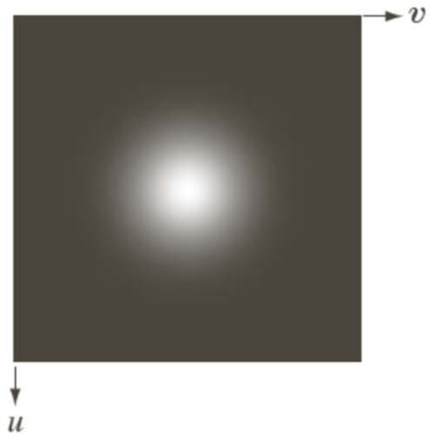
Butterworth LPF (BLPF)

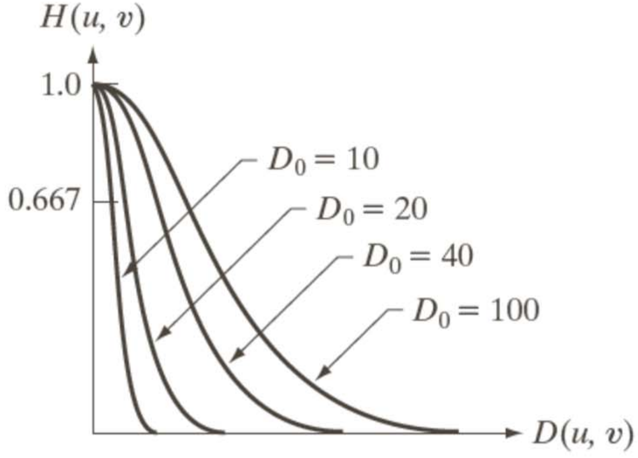


BLPFs of order 1, 2, 5, and 20
Ringing increases as filter order goes up

Gaussian LPF (GLPF)





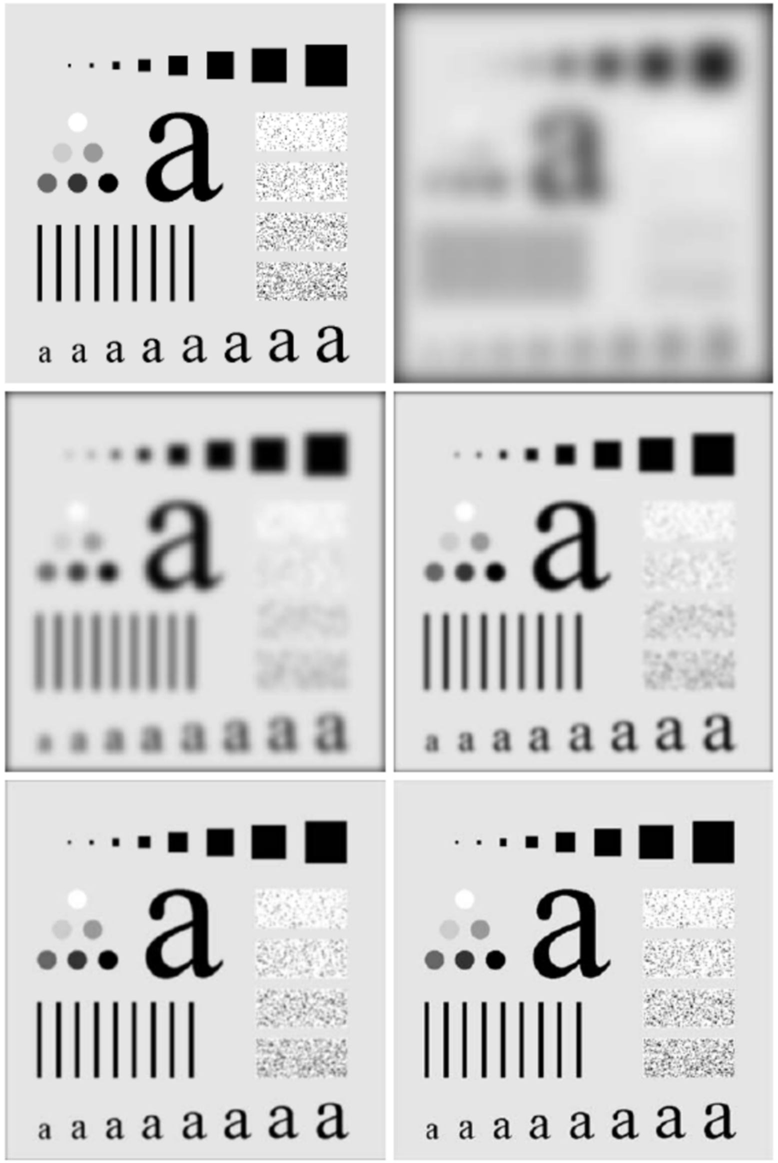


GLPF transfer function   Filter displayed as an image   Filter radial cross section

$$H(u,v) = e^{-\frac{D^2(u,v)}{2D_0^2}}$$

Gaussian LPF (GLPF)

- Gaussian vs. Butterworth
  - Less smoothing for Gaussian for same cutoff as Gaussian is more spread out in frequency domain



LPF

- Low-pass filters
- D_0: Cutoff frequency
- n: The order of Butterworth filter

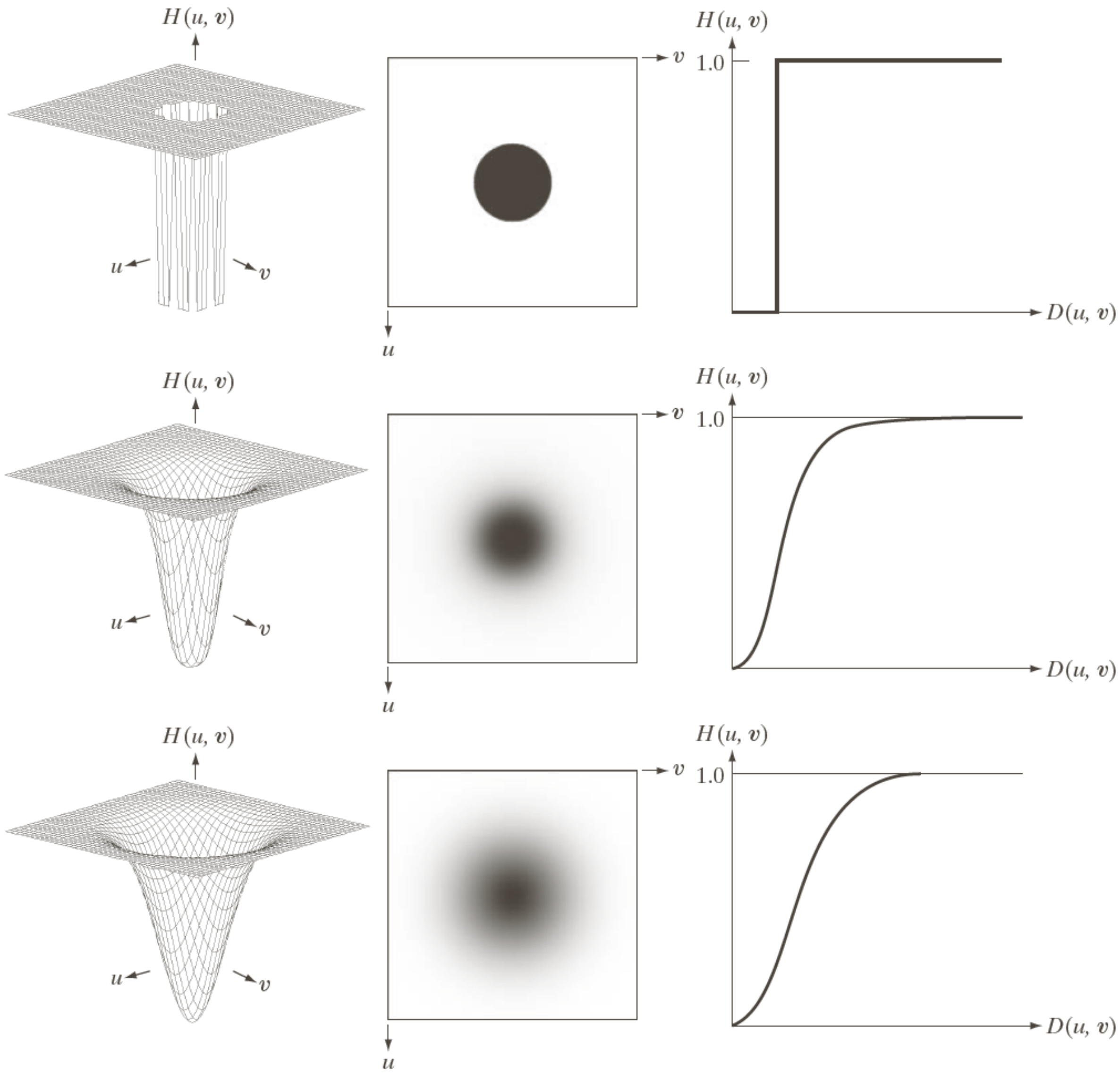
Ideal	Butterworth	Gaussian
$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \le D_0 \\ 0 & \text{if } D(u,v) > D_0 \end{cases}$	$H(u,v) = \dfrac{1}{1 + [D(u,v)/D_0]^{2n}}$	$H(u,v) = e^{-D^2(u,v)/2D_0^2}$

Image sharpening in frequency domain

- High frequency: Edge, noise, sharp transition in intensity
- High-pass filter (HPF): Only keep high frequency and attenuate low frequency
- Spatial domain: take spatial derivative

$$H_{HP}(u, v) = 1 - H_{LP}(u, v)$$

HPF

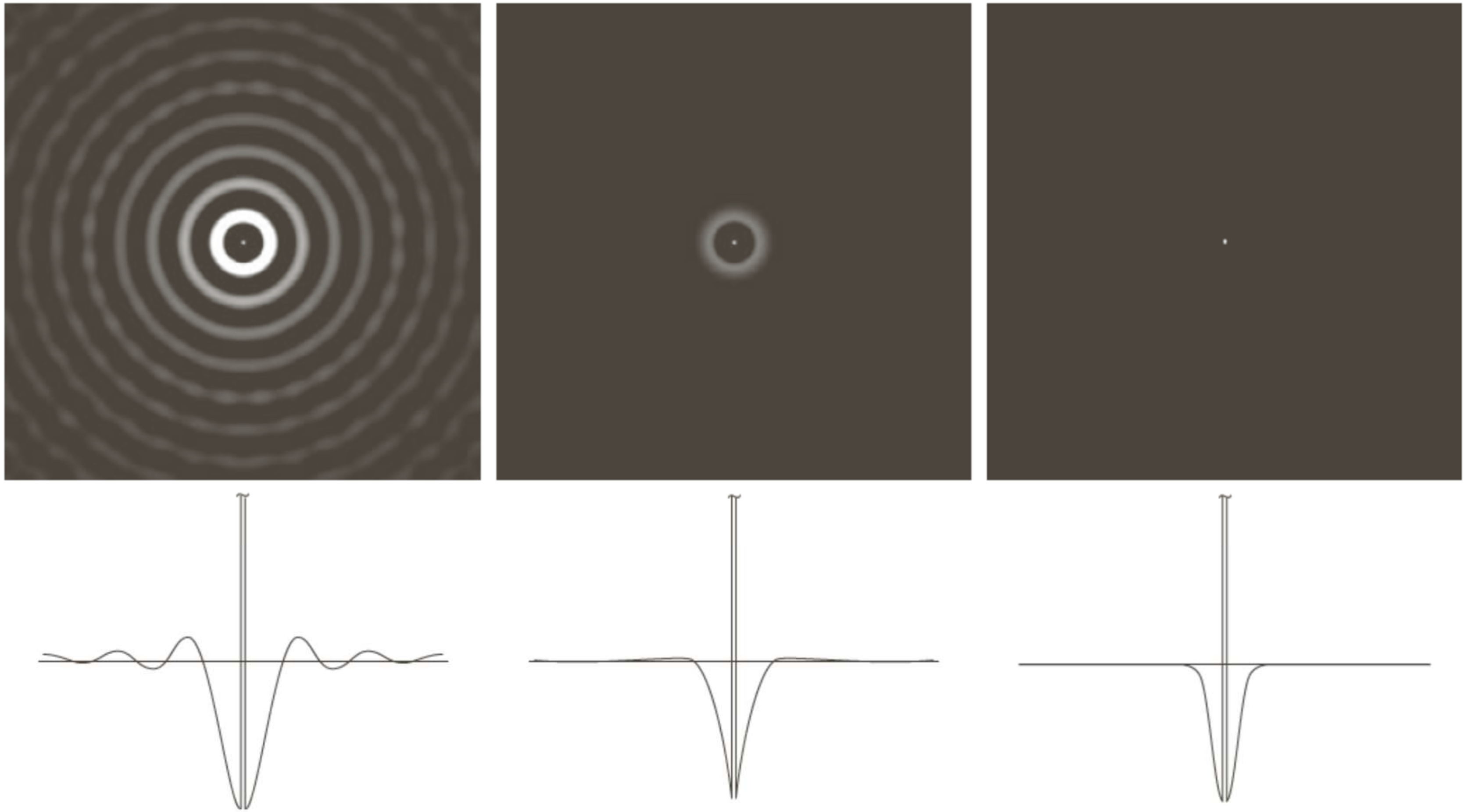


Ideal HPF

Butterworth HPF

Gaussian HPF

HPF



Ideal HPF　　Butterworth HPF　　Gaussian HPF

HPF

Ideal HPF

Butterworth HPF

Gaussian HPF

HPF

- High-pass filters
- D_0: Cutoff frequency
- n: The order of Butterworth filter

Ideal	Butterworth	Gaussian
$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \leq D_0 \\ 0 & \text{if } D(u,v) > D_0 \end{cases}$	$H(u,v) = \dfrac{1}{1 + [D_0/D(u,v)]^{2n}}$	$H(u,v) = 1 - e^{-D^2(u,v)/2D_0^2}$

Filters

- Various filter types