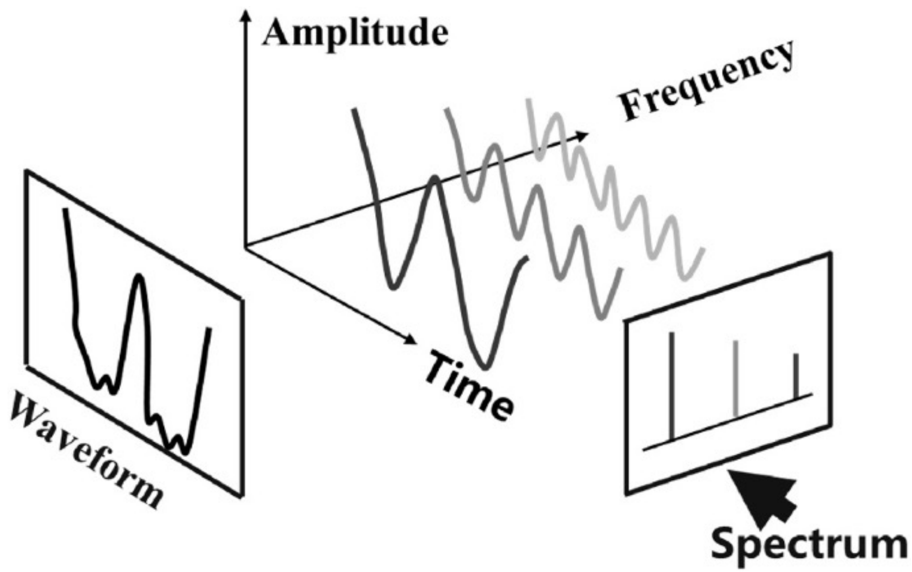


Frequency-domain Operators

Fall 2024

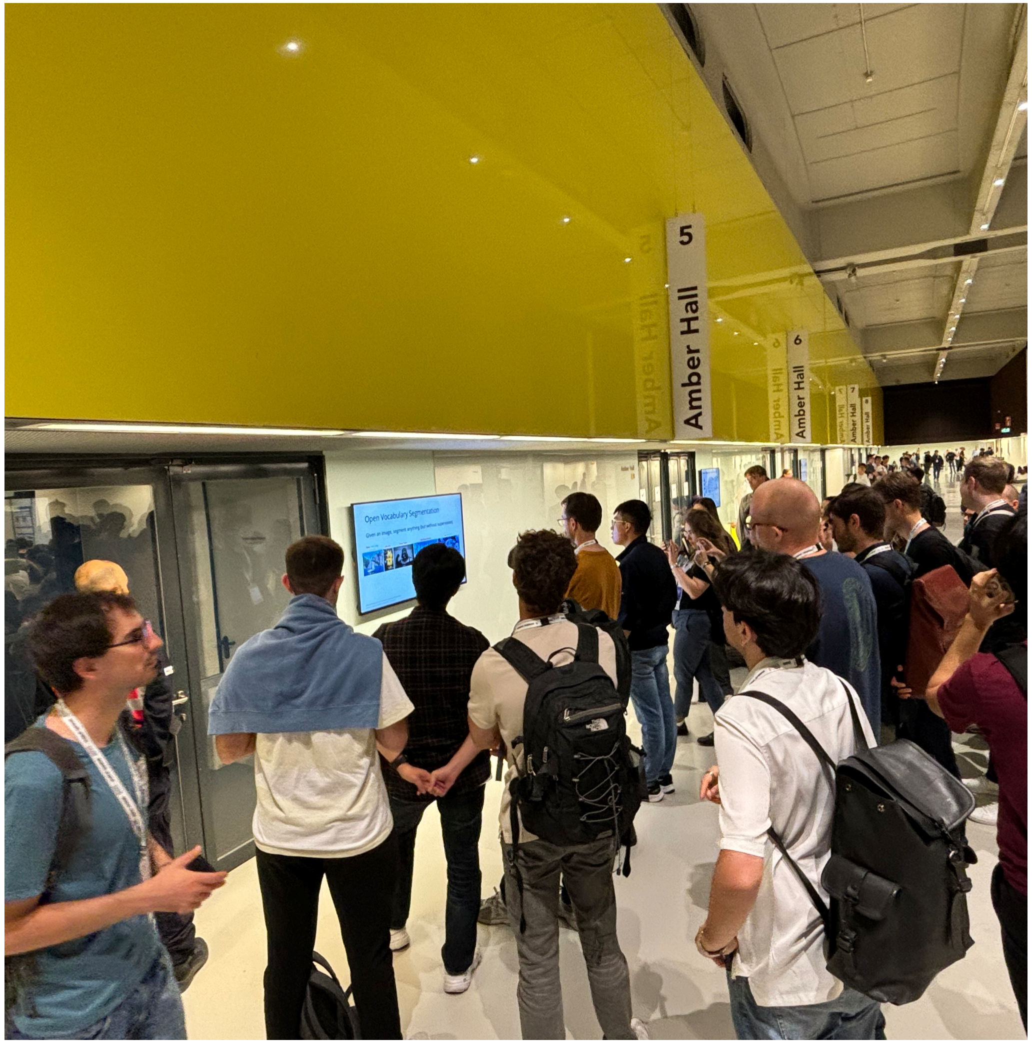
Yi-Ting Chen

ECCV: European Conference on Computer Vision

Computer Vision and Pattern Recognition (CVPR)

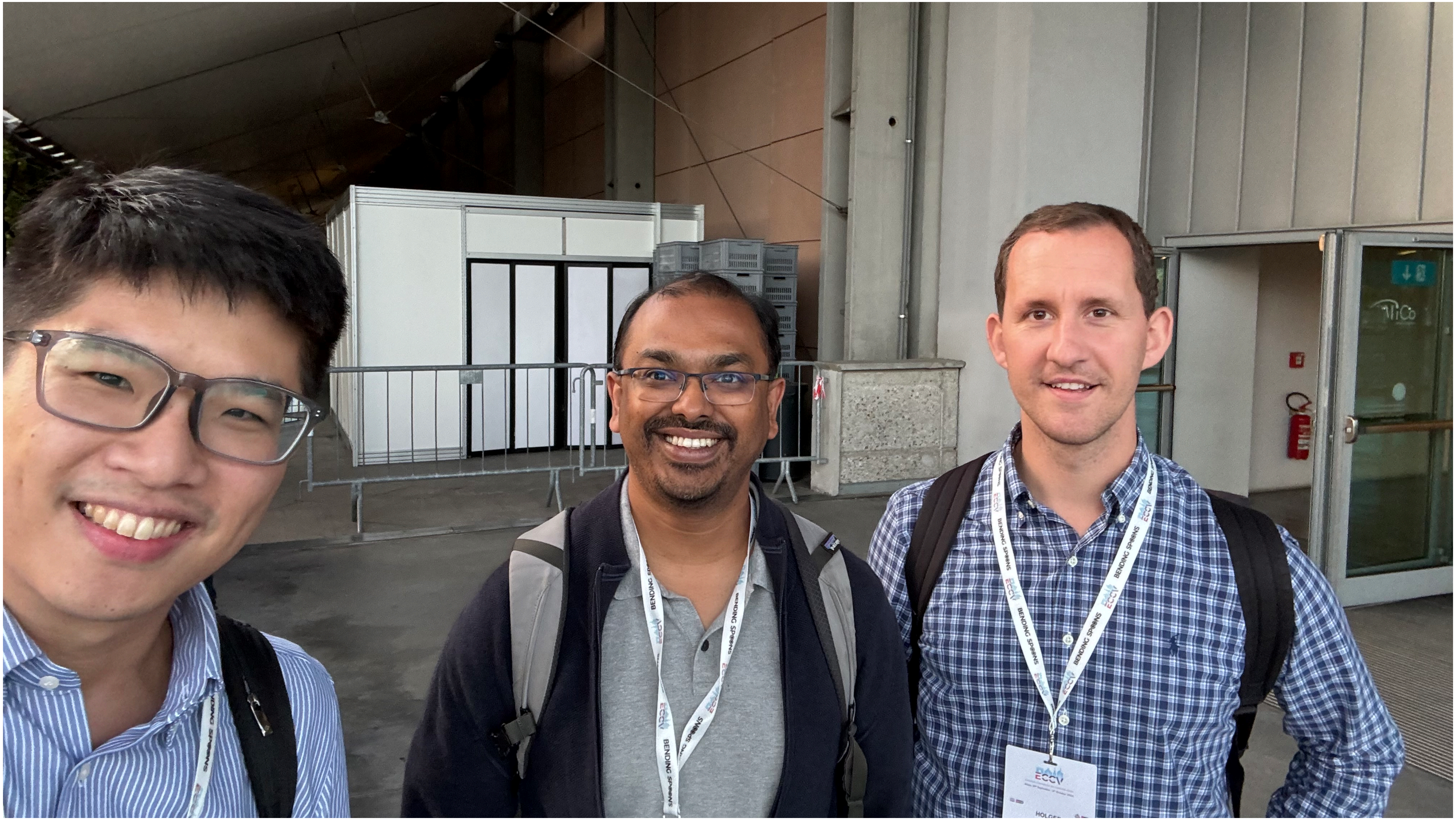
International Conference on Computer Vision (ICCV)

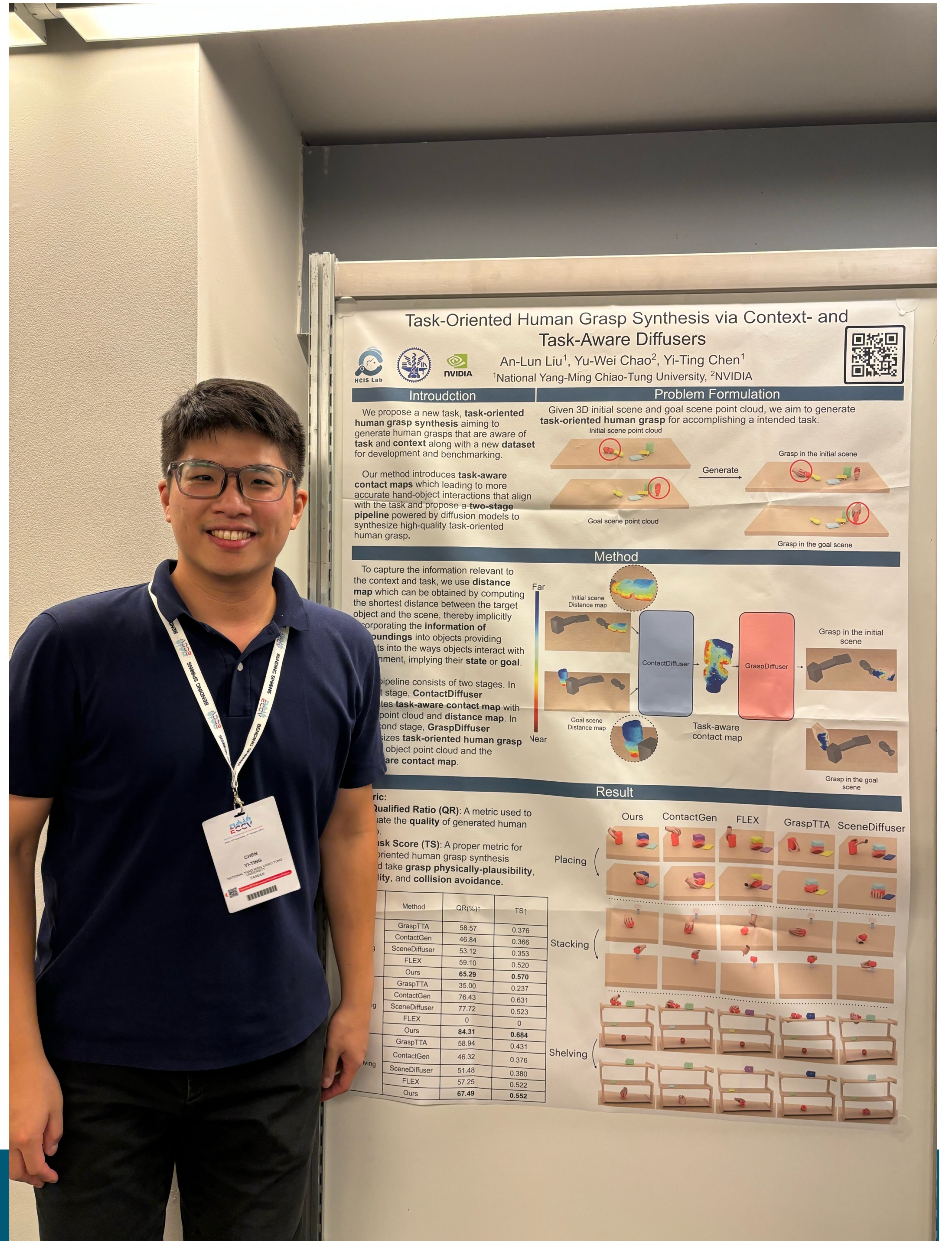

THULE


5
Amber Hall
6
Amber Hall

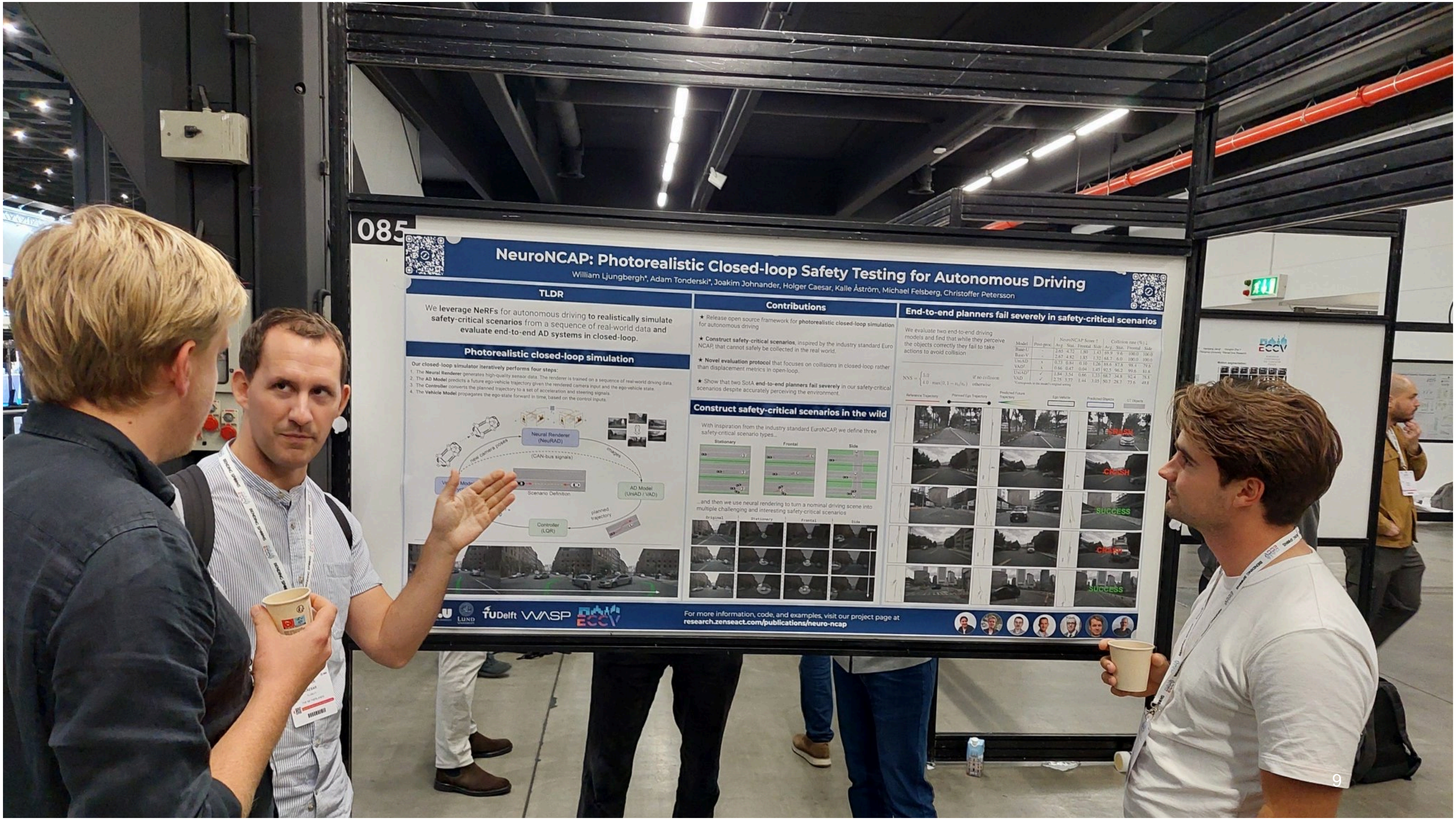

Brown Hall
3
Brown Hall

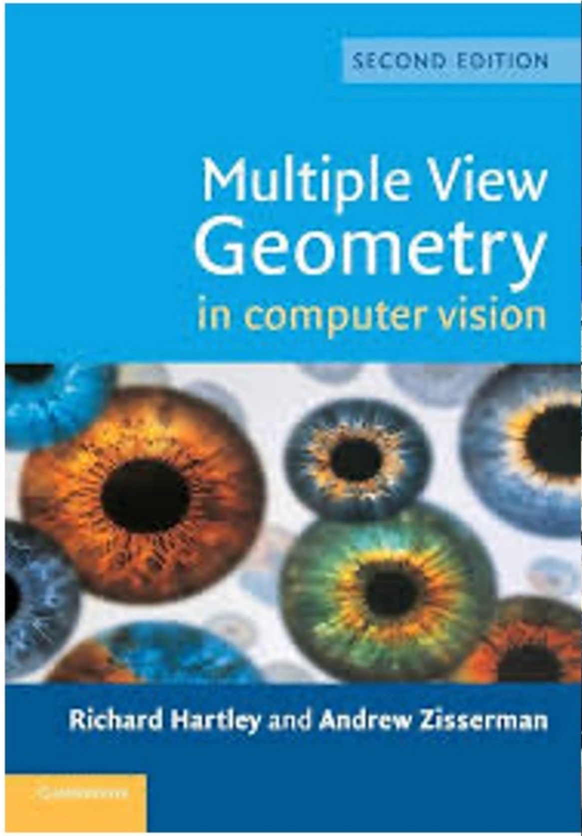

How to perform end-to-end forecasting?
Valeo4Cast — A Modular Approach
ROAD++
valeo.ai
Valeo4Cast: A Modular Approach to End-to-End Forecasting
A modular approach that gains flexibility and superior performance
Detection
Tracking
Forecasting
BEVFusion [4]
AB3DMOT [5]
MTR [6]
end-to-end forecasting
State-of-the-art results on Argoverse 2
mAPf
Ground truth
True positive
False positive
Valeo4Cast
Autonomous Drone-Person Tracking and Following in Uniform Appearance Scenarios
Introduction
Methodology
Conditional Unscented Autoencoders for Trajectory Prediction
BOSCH
ECCV


BENDING SPOONS
ECCV



Digital Image Processing


085
NeuroNCAP: Photorealistic Closed-loop Safety Testing for Autonomous Driving
William Ljungbergh*, Adam Tonderski*, Joakim Johnander, Holger Caesar, Kalle Åström, Michael Felsberg, Christoffer Petersson
TLDR
We leverage NeRFs for autonomous driving to realistically simulate safety-critical scenarios from a sequence of real-world data and evaluate end-to-end AD systems in closed-loop.
Photorealistic closed-loop simulation
Contributions
Construct safety-critical scenarios in the wild
End-to-end planners fail severely in safety-critical scenarios
CRASH
SUCCESS
TUDelft
WASP
ECCV
For more information, code, and examples, visit our project page at research.zenseact.com/publications/neuro-ncap

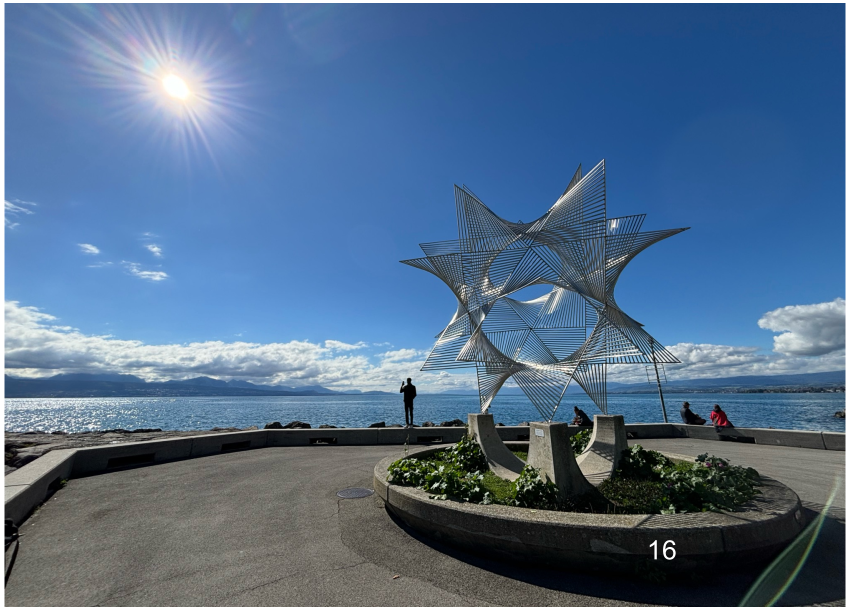

SECOND EDITION
Multiple View
Geometry
in computer vision
Richard Hartley and Andrew Zisserman


intel
captions
eo Generation
1500
amazon | science
1400
nexdata
Meta
VOXEL51
Making visual AI a reality
1600
1700
1800
Google
ZOOX
Hello Seattle
OpenEQA
Imagine Flash
Next-gen AI Product: OPPO Air Glass 3
Shop. Save. Skip the checkout line


ST-GERMAIN
HUGO


EINFÜHRUNGSMATINEE
ZUR PREMIERE VON
FIN DE PARTIE

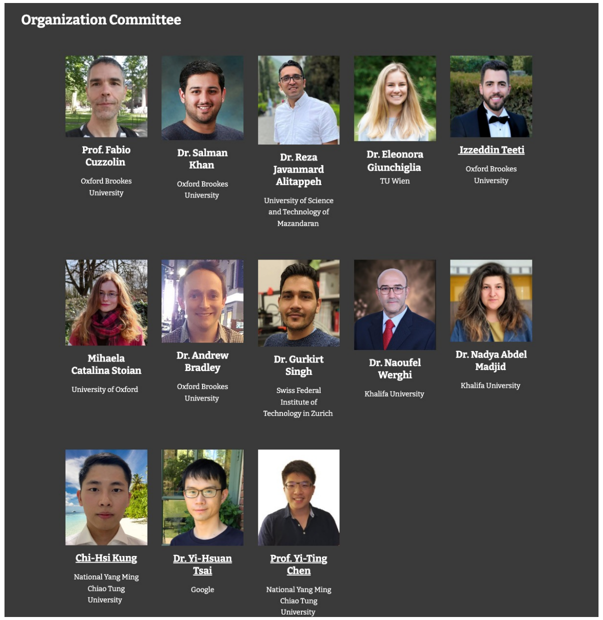


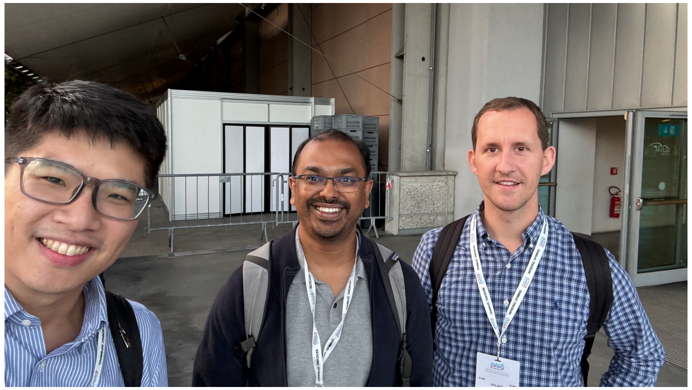


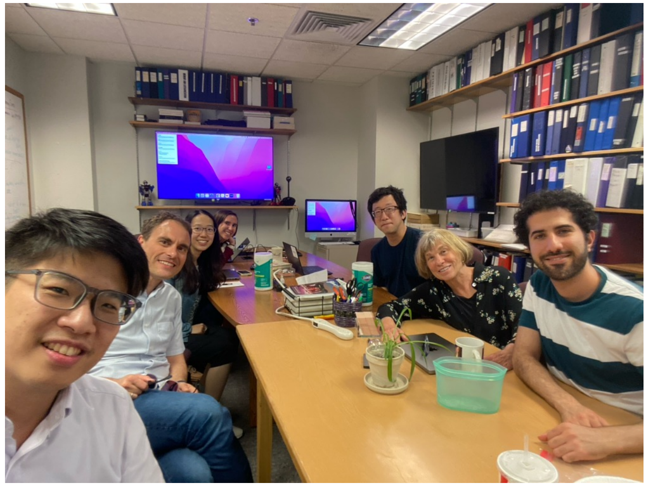
Work Hard and Smart, Play Harder!!

7 countries this year ;-)

One work for many purposes







Kung et al., Action-Slot: Visual Action-centric Representation for Atomic Activity Recognition in Traffic Scenes, CVPR, 2024

1-D Linear Time Invariant System

$$x(t) \longrightarrow \boxed{T} \longrightarrow y(t)$$

Linear System

Additivity: $T\{x_1(t) + x_2(t)\} = T\{x_1(t)\} + T\{x_2(t)\} = y_1(t) + y_2(t)$

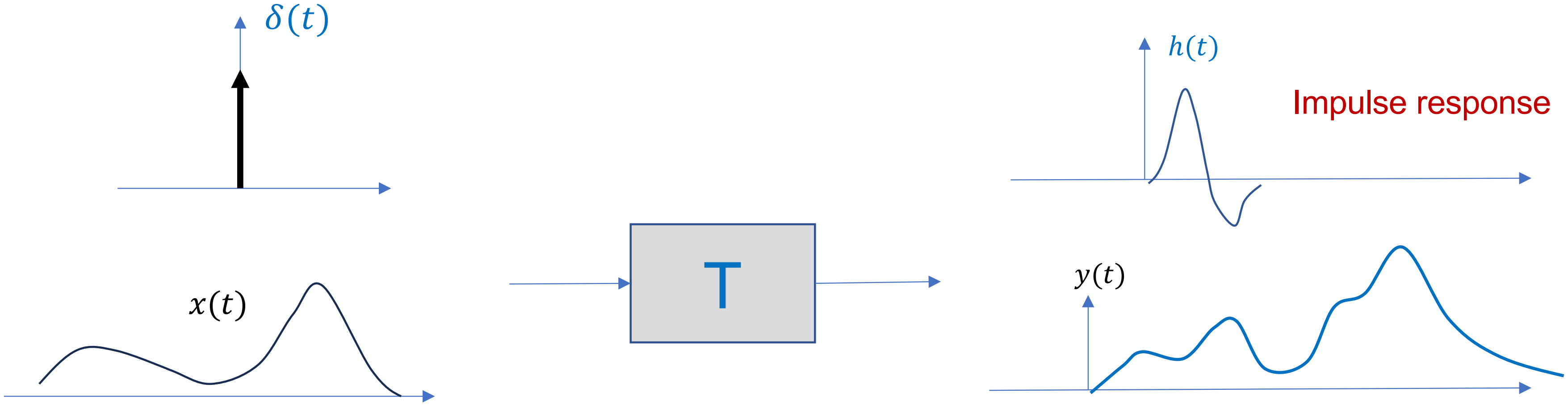
Homogeneity: $T\{ax(t)\} = aT\{x(t)\} = ay(t)$

Time-Invariant System

If $y(t) = T\{x(t)\}$, then $y(t - t_0) = T\{x(t - t_0)\}$.

Summary

$$h(t) = T\{\delta(t)\}$$



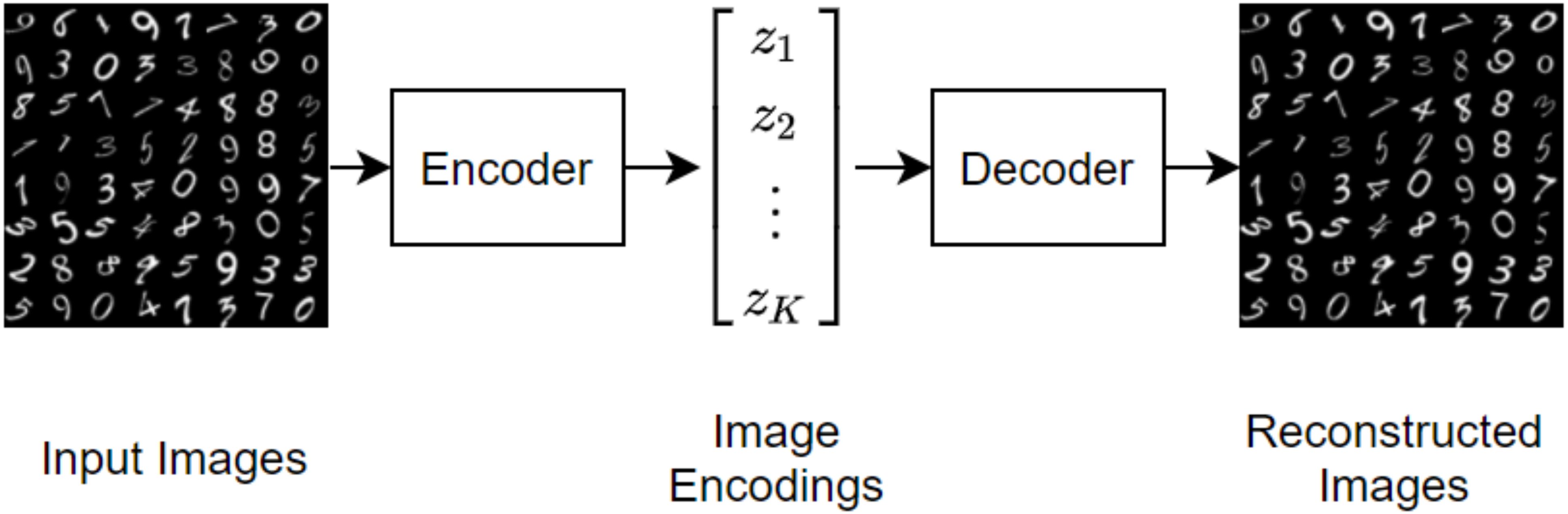
$$x(t) = \int_{-\infty}^{\infty} x(\tau)\delta(t-\tau)d\tau$$

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$ Convolution

Eigenfaces

https://en.wikipedia.org/wiki/Eigenface#:~:text=An%20eigenface%20(%2F%CB%88a%C9%AA%C9%A1,Alex%20Pentland%20in%20face%20classification.

Variational Autoencoder (VAE)



https://en.wikipedia.org/wiki/Variational_autoencoder

Frequency-Domain Analysis (1/3)

$$e^{j2\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft)$$

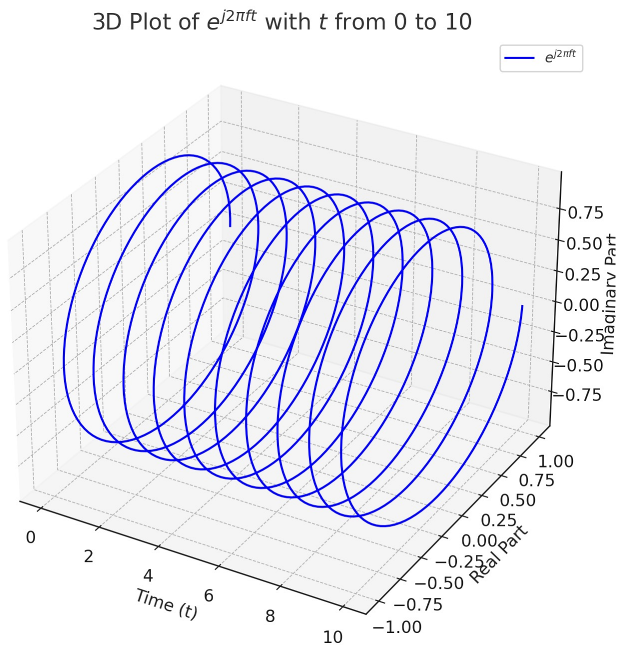
$$x(t) = e^{j2\pi ft} \longrightarrow \boxed{T} \longrightarrow y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$

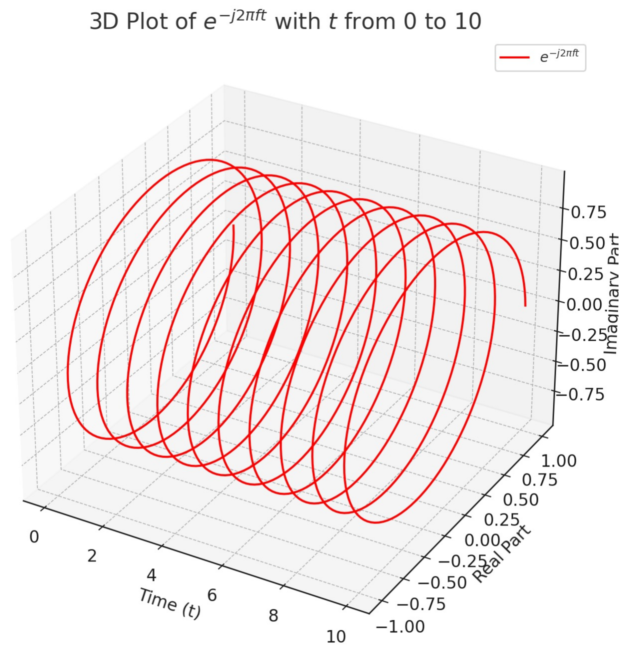
$$\begin{aligned}
&= \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau \\
&= \int_{-\infty}^{\infty} e^{j2\pi f(t-\tau)}h(\tau)d\tau \\
&= e^{j2\pi ft}\int_{-\infty}^{\infty} e^{-j2\pi f\tau}h(\tau)d\tau \\
&\equiv H(f)e^{j2\pi ft}
\end{aligned}$$

Frequency-Domain Analysis (1/3)

$$e^{j2\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft)$$

$$e^{-j2\pi ft} = \cos(2\pi ft) - j\sin(2\pi ft)$$





Frequency-Domain Analysis (1/3)

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$
$$\equiv H(f)e^{j2\pi ft}$$

H(f): Frequency Response

This is a complex-valued function that characterizes how the system responds to a sinusoid of frequency f. It typically includes both a gain (magnitude) and a phase shift.

$$y(t) = |H(f)|e^{j(2\pi ft+\arg(H(f)))}$$

Can we use this charateristic to decompose our signal?

Frequency-Domain Analysis (2/3)

Fourier Transform Pair

spectrum $X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$ where $e^{-j2\pi ft} = \cos(2\pi ft) - j\sin(2\pi ft)$

Signal Decomposition $x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df$ where $e^{j2\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft)$

$$X(f) = Re\{X(f)\} + jIm\{X(f)\} = |X(f)| e^{j\angle X(f)}$$

$|X(f)|$ Magnitude

$\angle X(f)$ Phase

Frequency-Domain Analysis (3/3)

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} df$$

$$y(t) = \int_{-\infty}^{\infty} X(f) T\{e^{j2\pi f t}\} df$$
$$= \int_{-\infty}^{\infty} X(f) H(f) e^{j2\pi f t} df$$

$$\Rightarrow \quad Y(f) = X(f)H(f)$$

2-D Linear Shift-Invariant System

$$I(x,y) \longrightarrow \boxed{h(x,y)} \longrightarrow O(x,y) = h(x,y) * I(x,y)$$

2-D FT

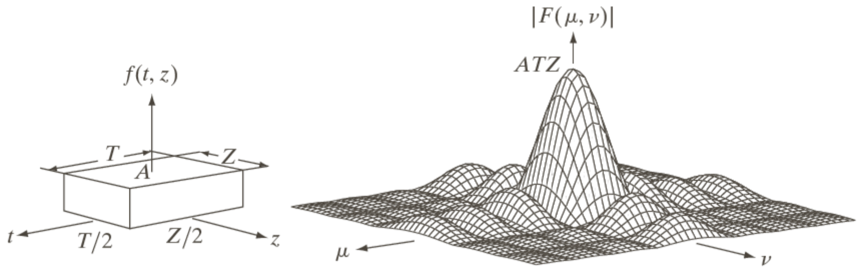
$$I(u,v) \qquad H(u,v) \qquad O(u,v) = H(u,v)I(u,v)$$

h(x,y): Point Spread Function (PSF)
H(u,v): optical transfer function (OTF)

2-D Fourier Transform

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, e^{-j2\pi(ux+vy)} dxdy$$

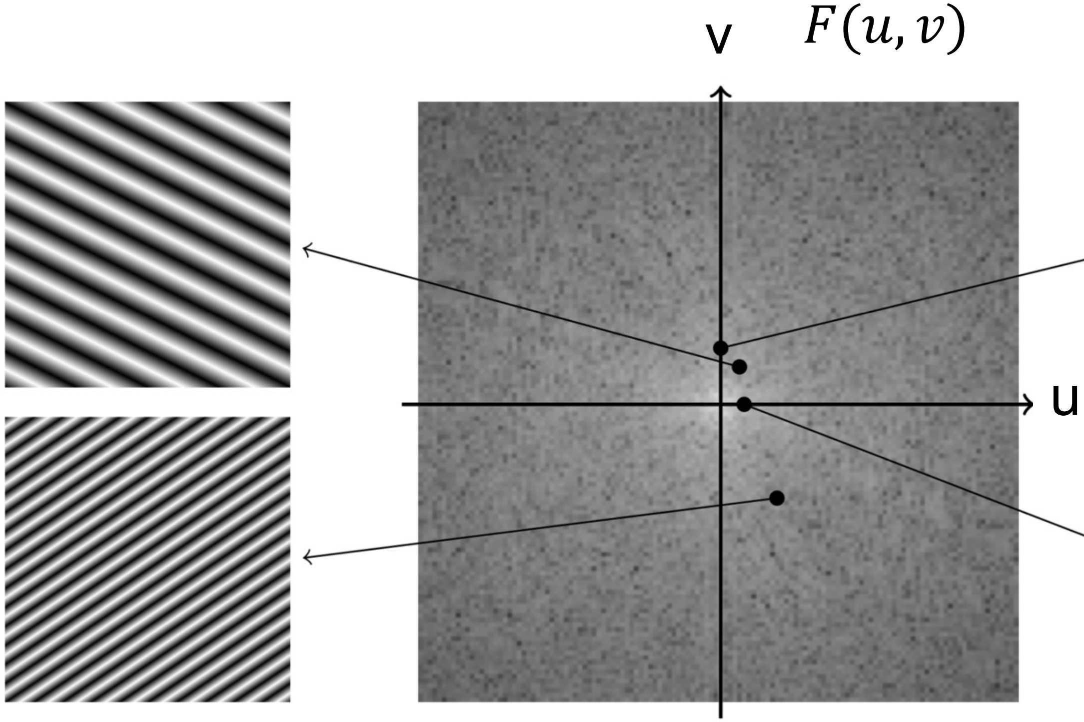
$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v)\, e^{j2\pi(ux+vy)} dudv$$



a b

FIGURE 4.13 (a) A 2-D function, and (b) a section of its spectrum (not to scale). The block is longer along the t-axis, so the spectrum is more "contracted" along the μ-axis. Compare with Fig. 4.4.

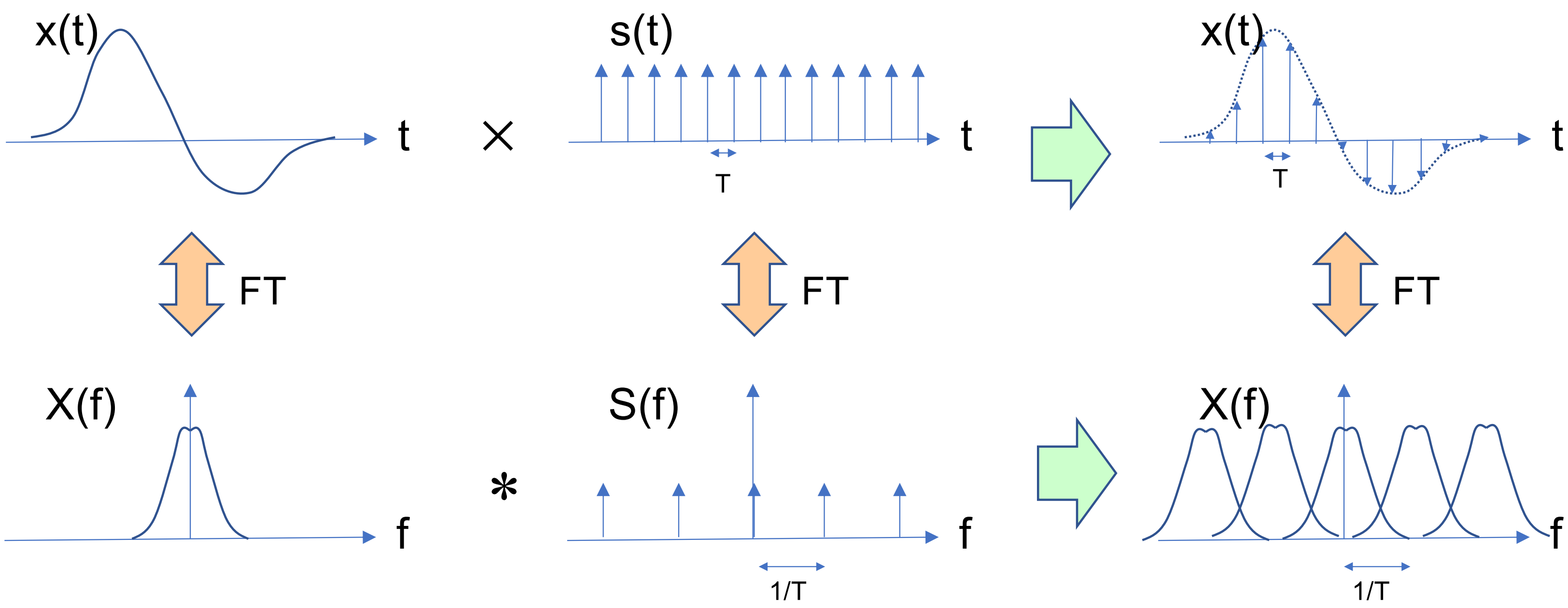
$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)\, e^{j2\pi(ux+vy)} du dv$$



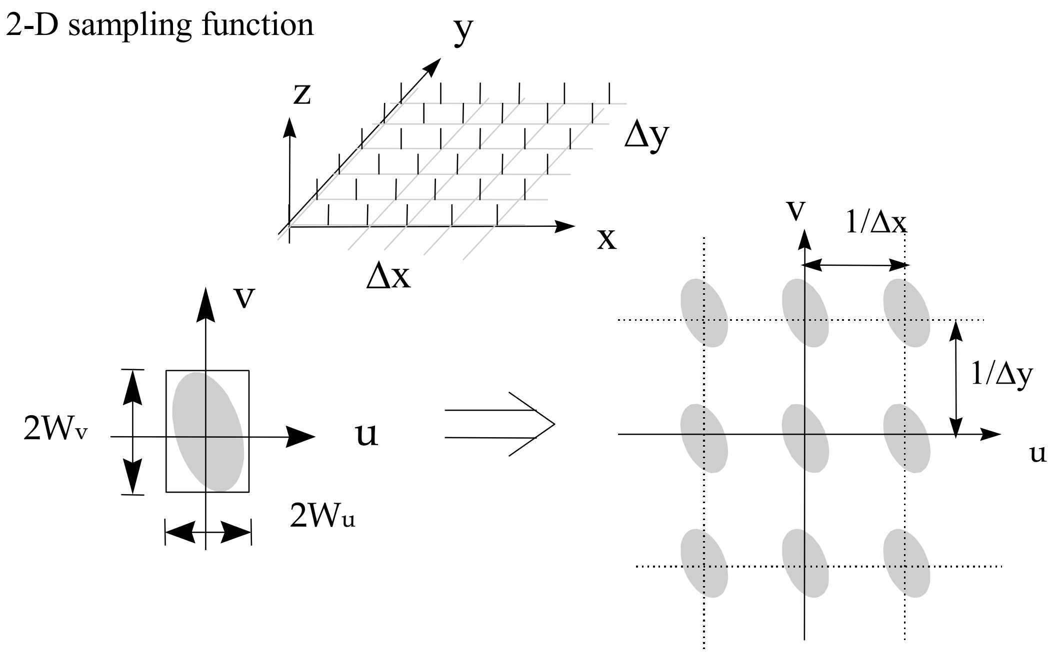
https://commons.wikimedia.org/wiki/File:2D_Fourier_Transform_and_Base_Images.png

Sampling

Whittaker-Shannon Sampling Theorem



2-D Sampling



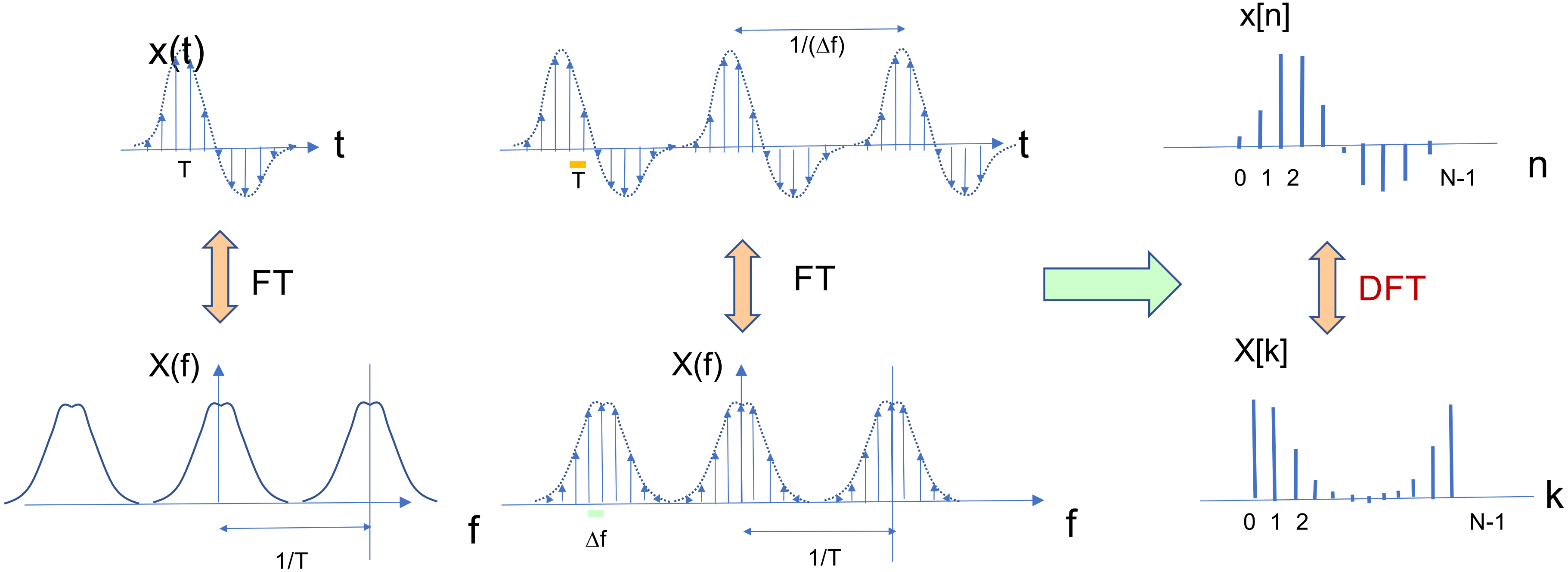
Credit: Prof. Sheng-Jyh Wang NYCU EE

Aliasing

a b c

FIGURE 4.17 Illustration of aliasing on resampled images. (a) A digital image with negligible visual aliasing. (b) Result of resizing the image to 50% of its original size by pixel deletion. Aliasing is clearly visible. (c) Result of blurring the image in (a) with a 3×3 averaging filter prior to resizing. The image is slightly more blurred than (b), but aliasing is not longer objectionable. (Original image courtesy of the Signal Compression Laboratory, University of California, Santa Barbara.)

Discrete Fourier Transform (1/2)



Discrete Fourier Transform (2/2)

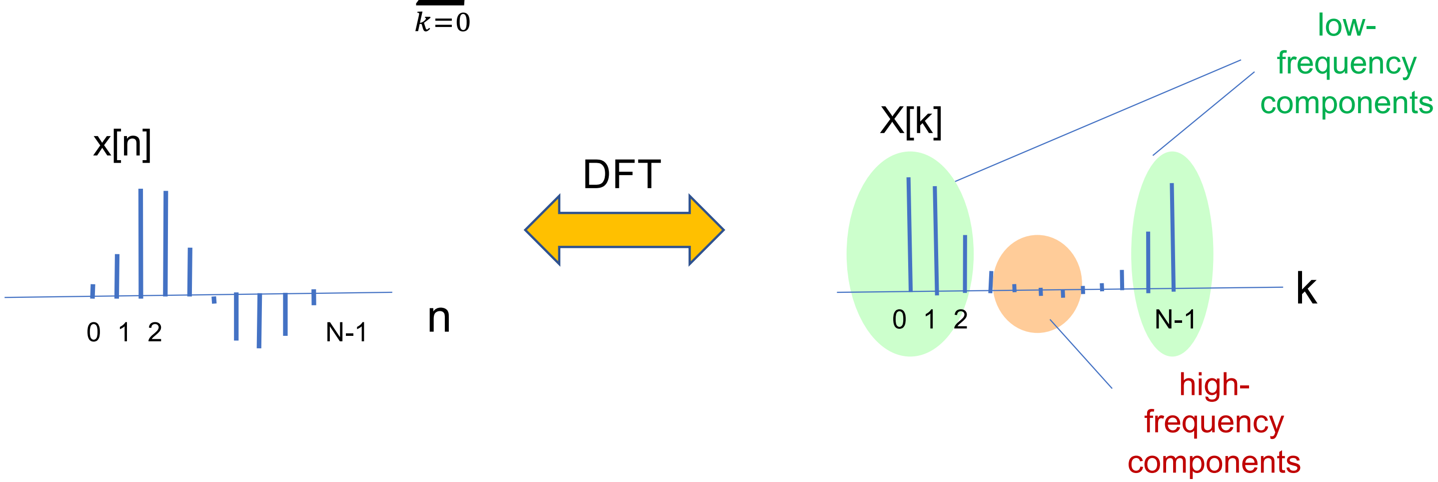
Fourier Transform Pair

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt$$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} df$$

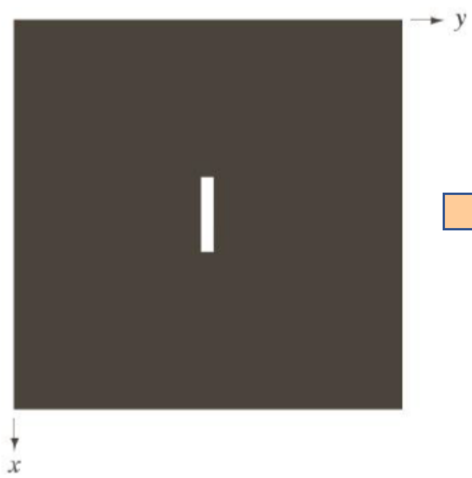
$$X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi k n}{N}} \qquad k = 0,1,2,..,N-1$$

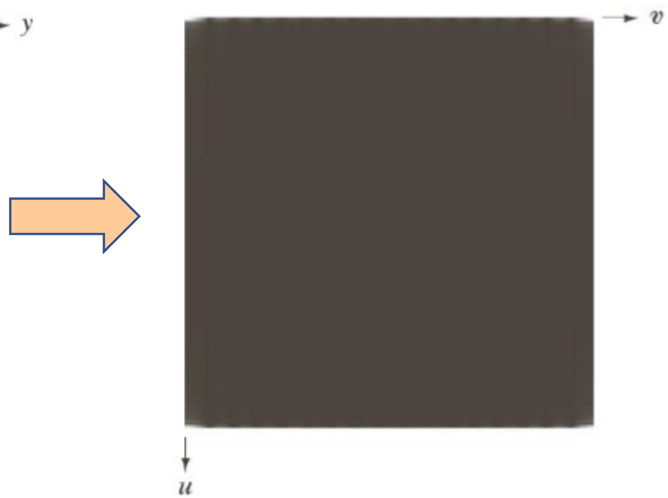
$$x[n] = \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi k n}{N}} \qquad n = 0,1,2,..,N-1$$

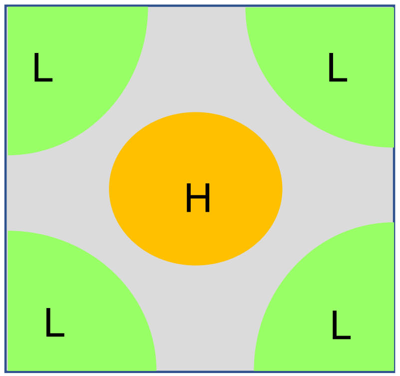


2-D DFT

$$F[u,v] = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[m,n]\, e^{-j2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)} \qquad f[m,n] = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F[u,v]\, e^{j2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$



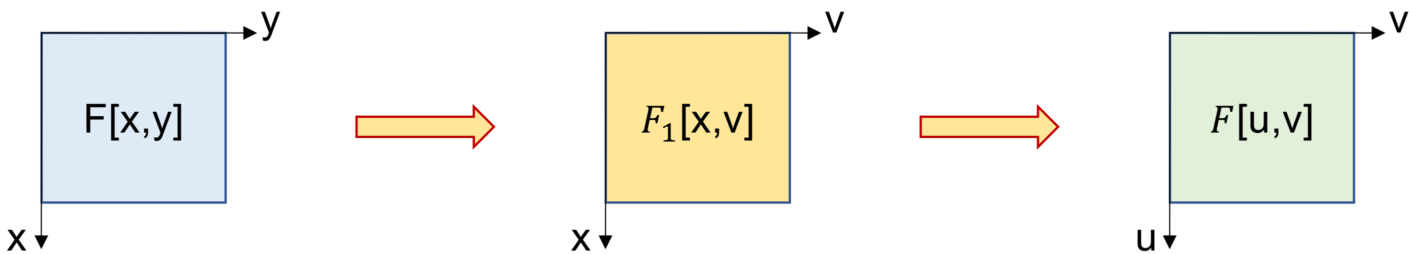




Properties of DFT - Separability

$$F(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi\frac{ux+vy}{N}}$$

$$= \frac{1}{N}\sum_{x=0}^{N-1} e^{-j2\pi\frac{ux}{N}}\{\sum_{y=0}^{N-1}\frac{1}{N}f(x,y)e^{-j2\pi\frac{vy}{N}}\}$$



$$\mathrm{F}[u,v] = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[m,n]\, e^{-j2\pi(\frac{um}{M}+\frac{vn}{N})}$$

$$\mathrm{f}[m,n] = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F[u,v]\, e^{j2\pi(\frac{um}{M}+\frac{vn}{N})}$$

- **Average Value** $\bar{f}[x,y] = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f[x,y] = F[0,0]$

- **Periodicity**

$$f[x,y] = f[x+kN, y+lN]$$

$$F[u,v] = F[u+kN, v+lN]$$

$$k,l = 0,1,2,\ldots$$

- **Translation**

$$f[x-x_0, y-y_0] \Leftrightarrow F[u,v]e^{-j2\pi(\frac{ux_0+vy_0}{N})}$$

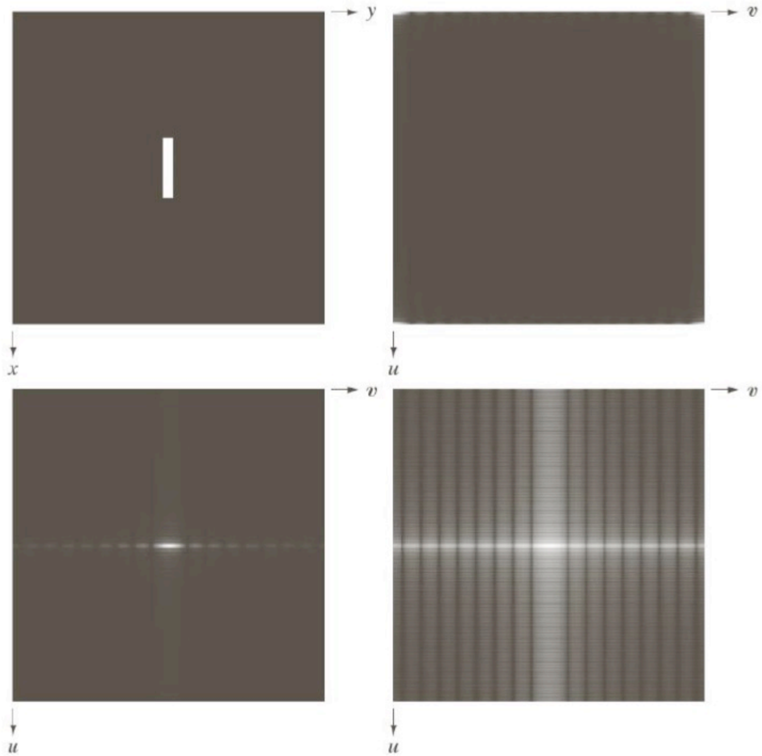
$$f[x,y]e^{j2\pi(\frac{u_0x+v_0y}{N})} \Leftrightarrow F[u-u_0, v-v_0]$$

If $u_0 = v_0 = N/2$

$$f[x,y](-1)^{(x+y)}$$

DFT

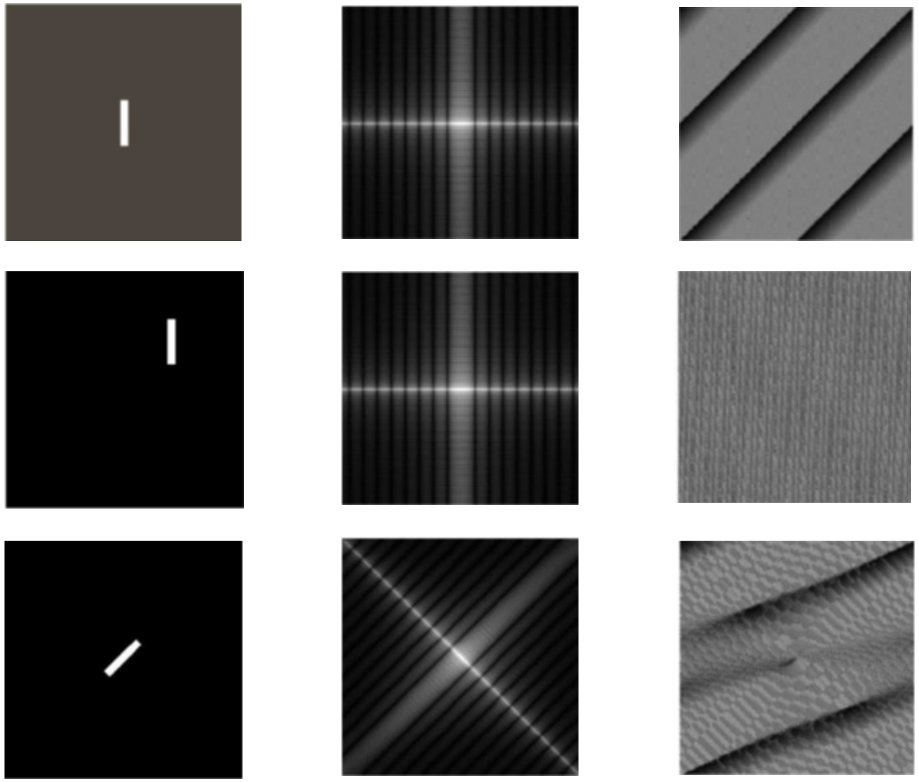
$$F[u-\frac{N}{2}, v-\frac{N}{2}]$$



Log function

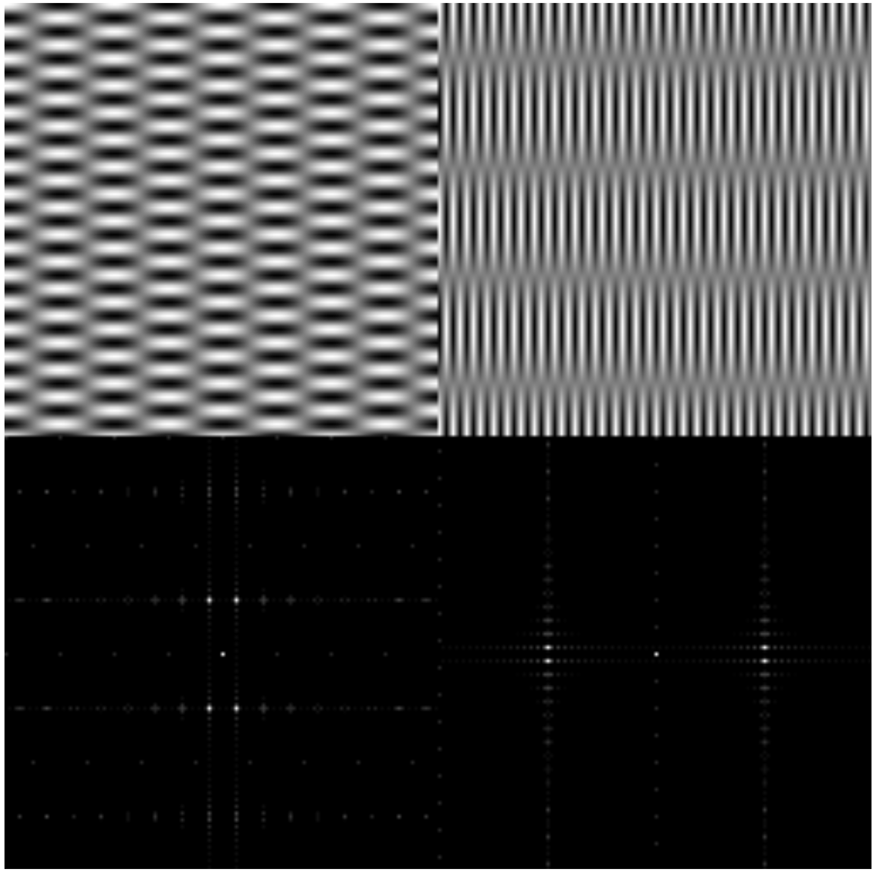
Rotation

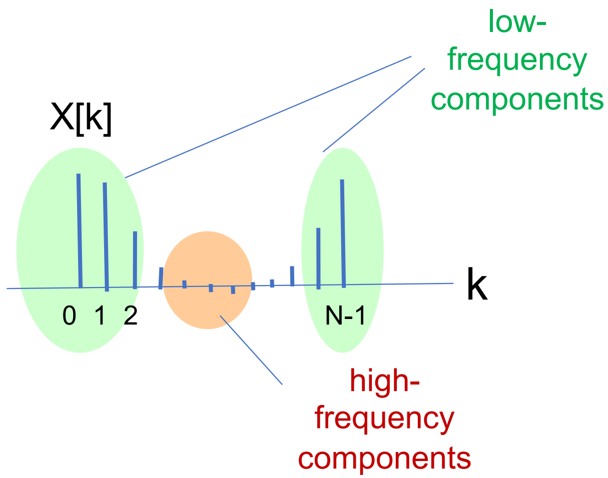
In principle, $f[r, \theta + \theta_0] \leftrightarrow F[w, \phi + \theta_0]$



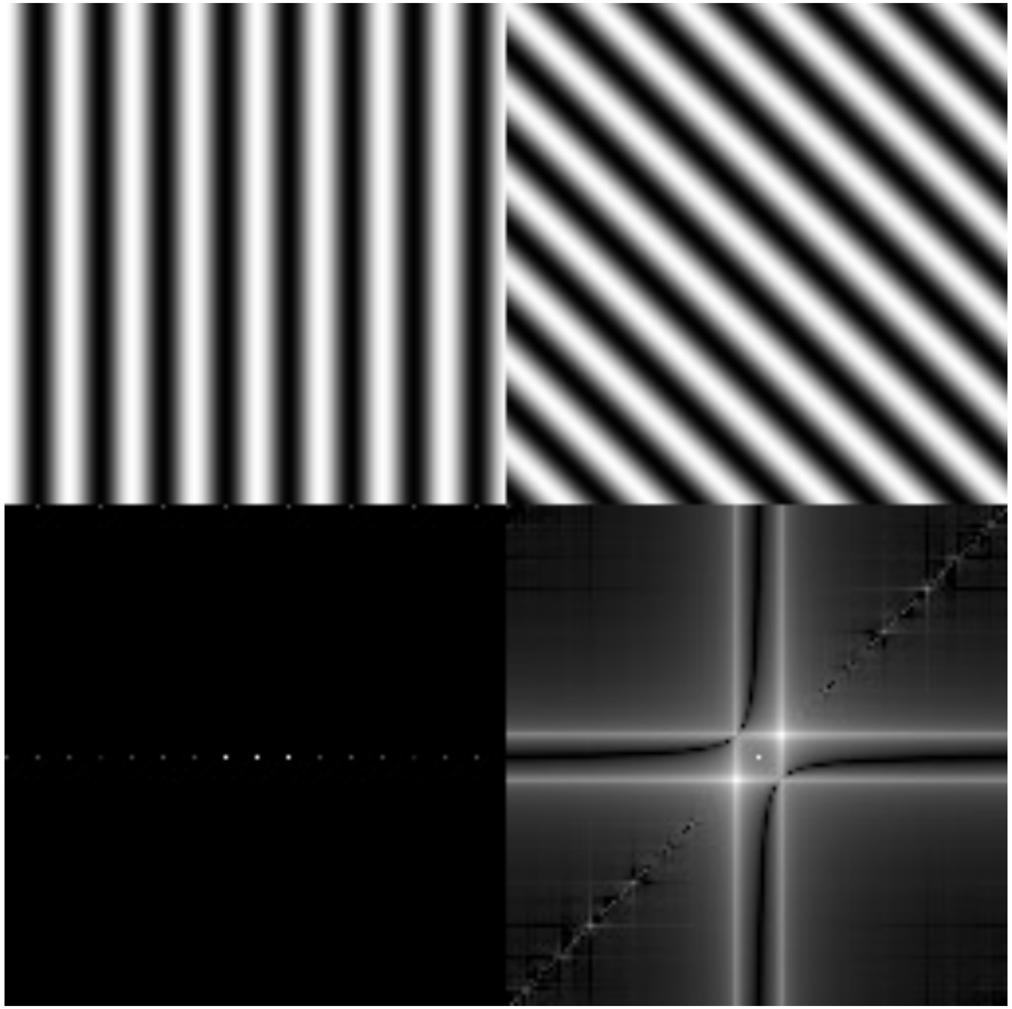
Ref: Gonzalez and Woods, “Digital Image Processing”

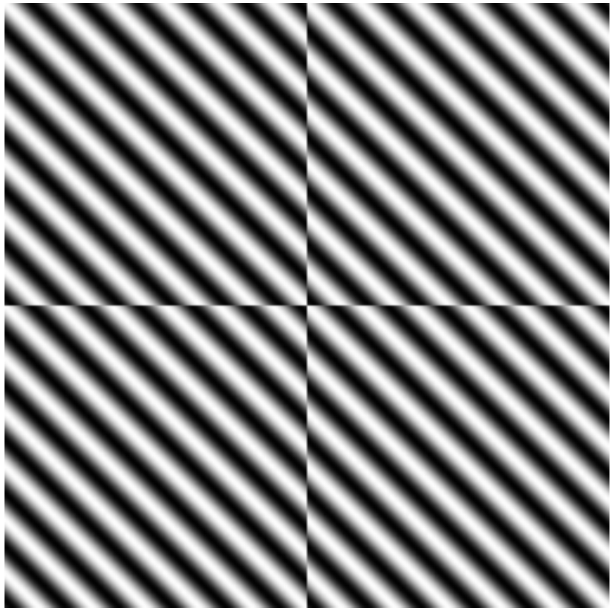
Example

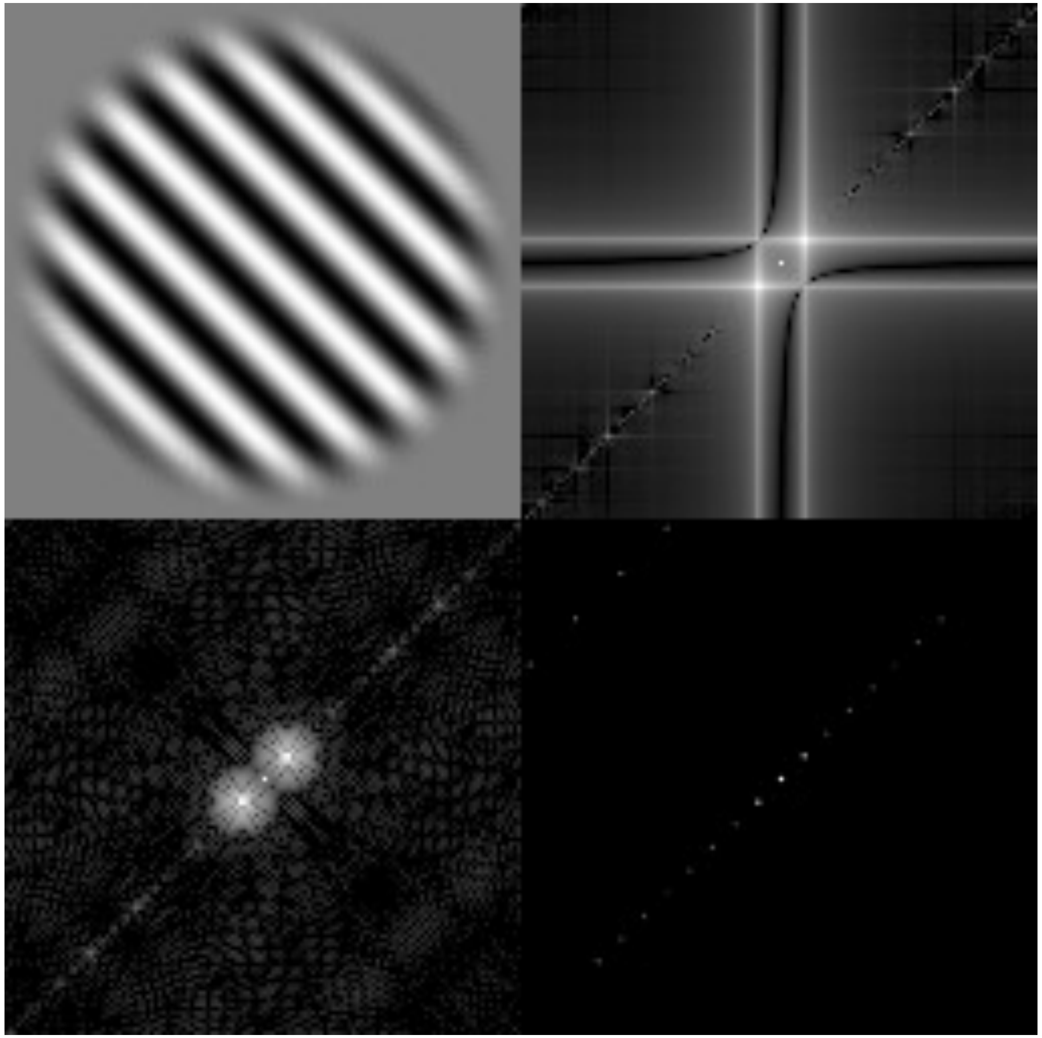




Ref: http://www.cs.unm.edu/~brayer/vision/fourier.html







Frequency-domain Filtering

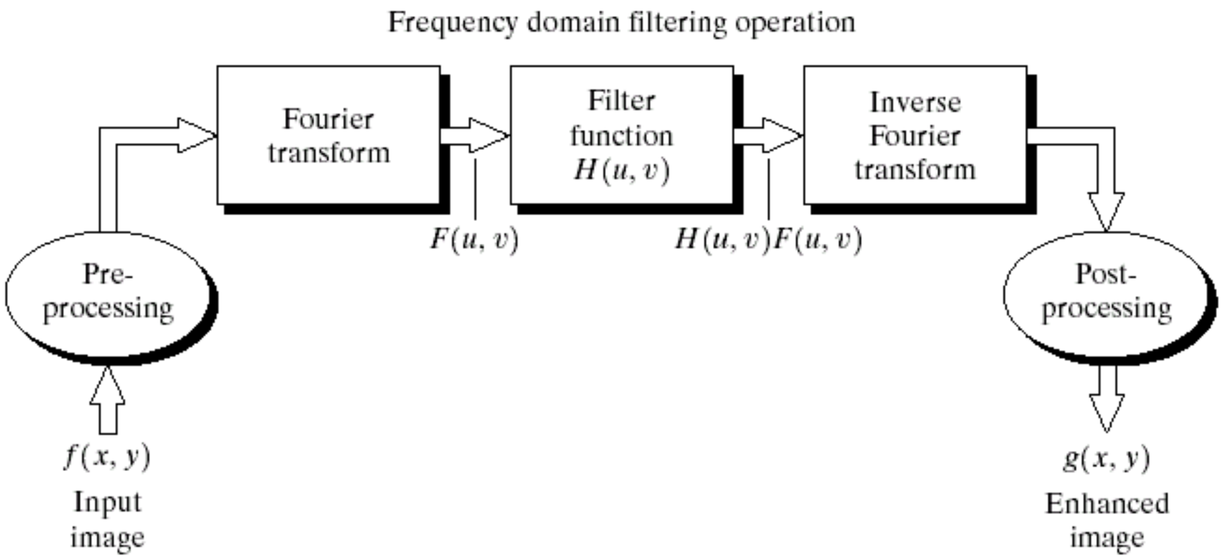
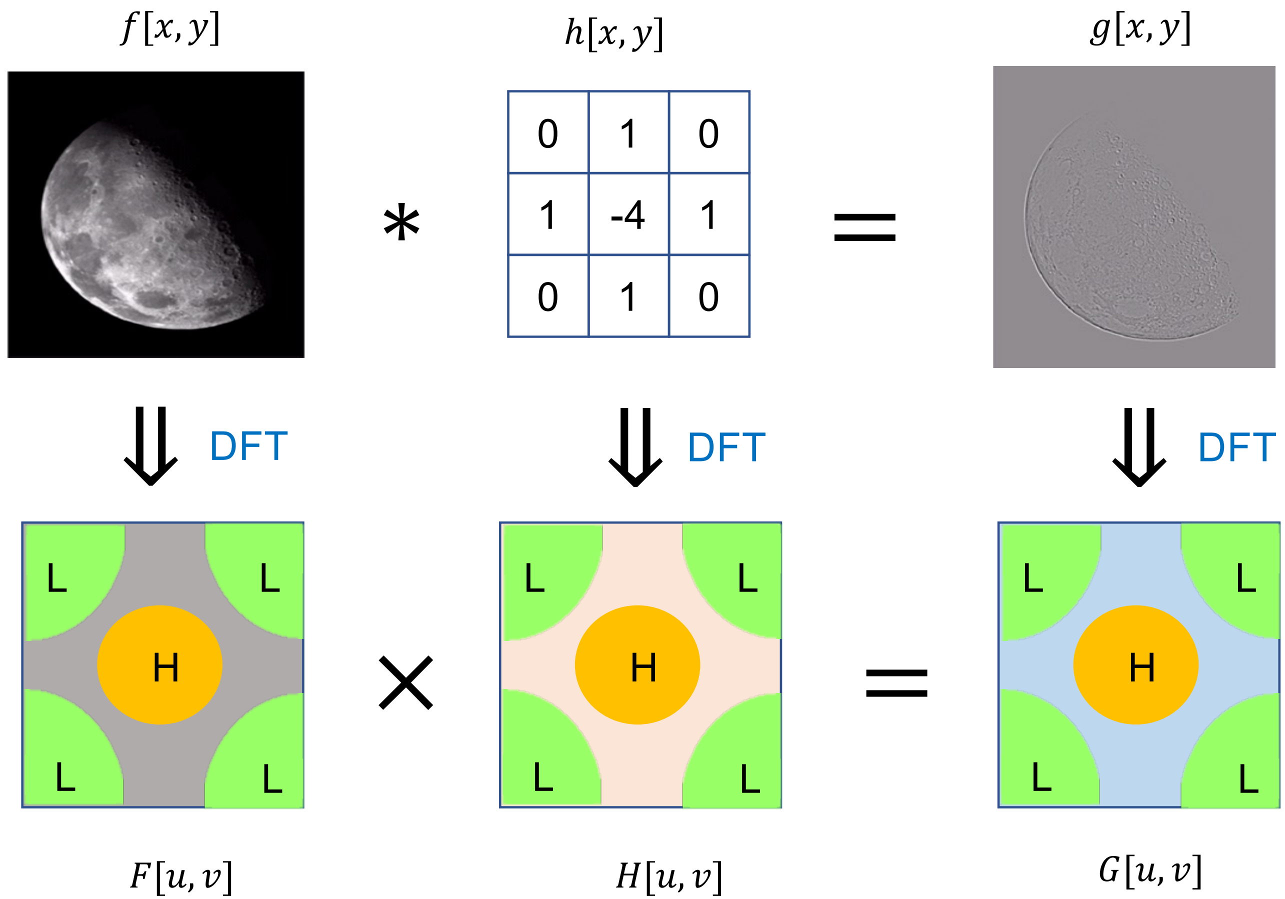
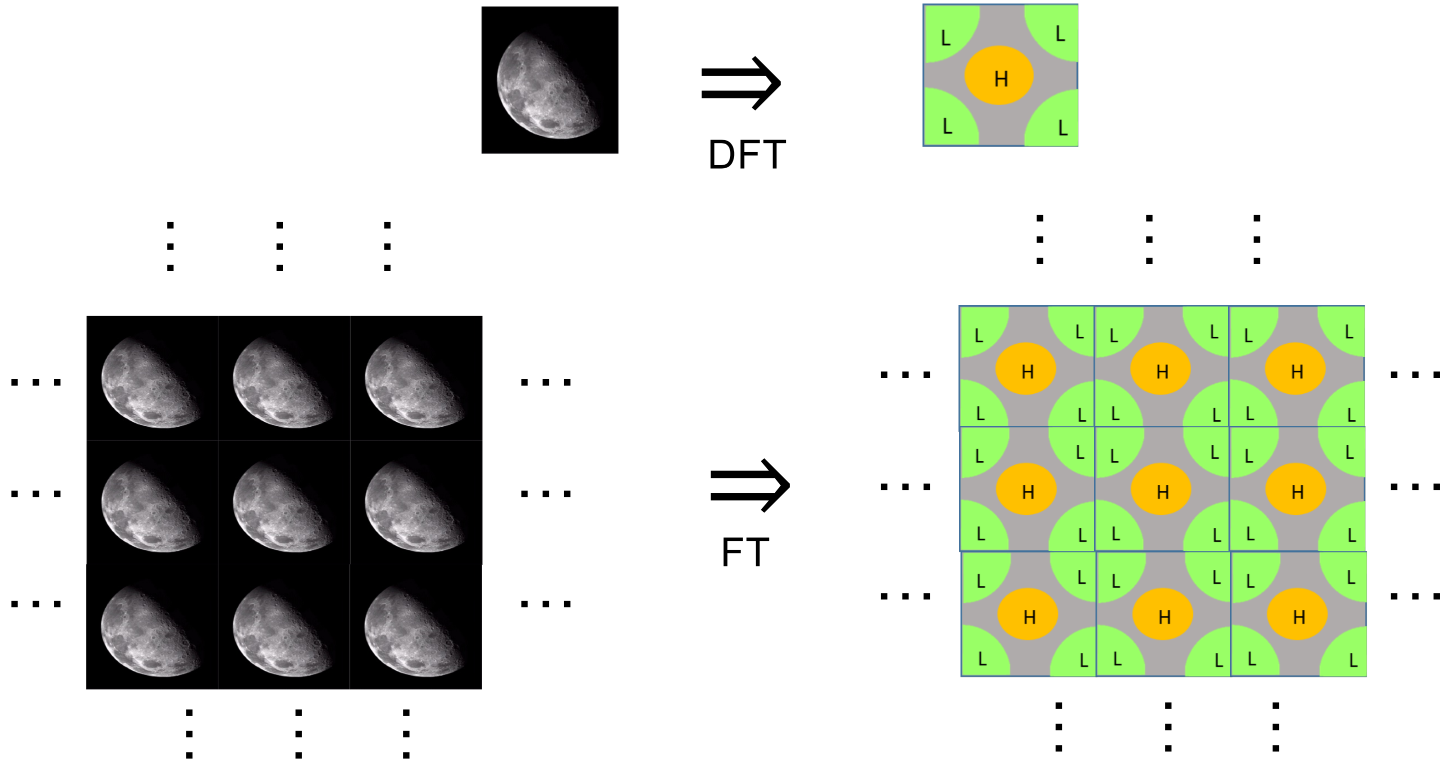


FIGURE 4.5 Basic steps for filtering in the frequency domain.

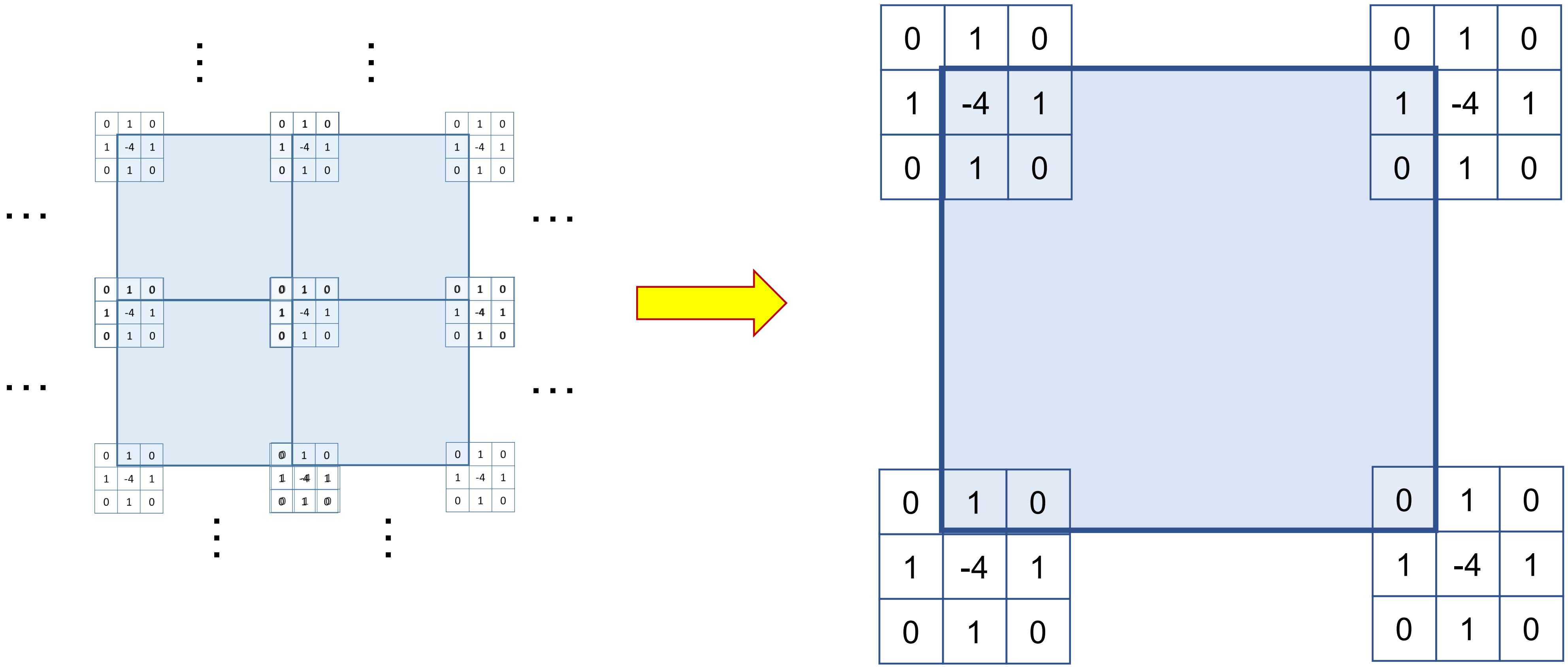
Linear Filtering in Frequency-domain (1/4)



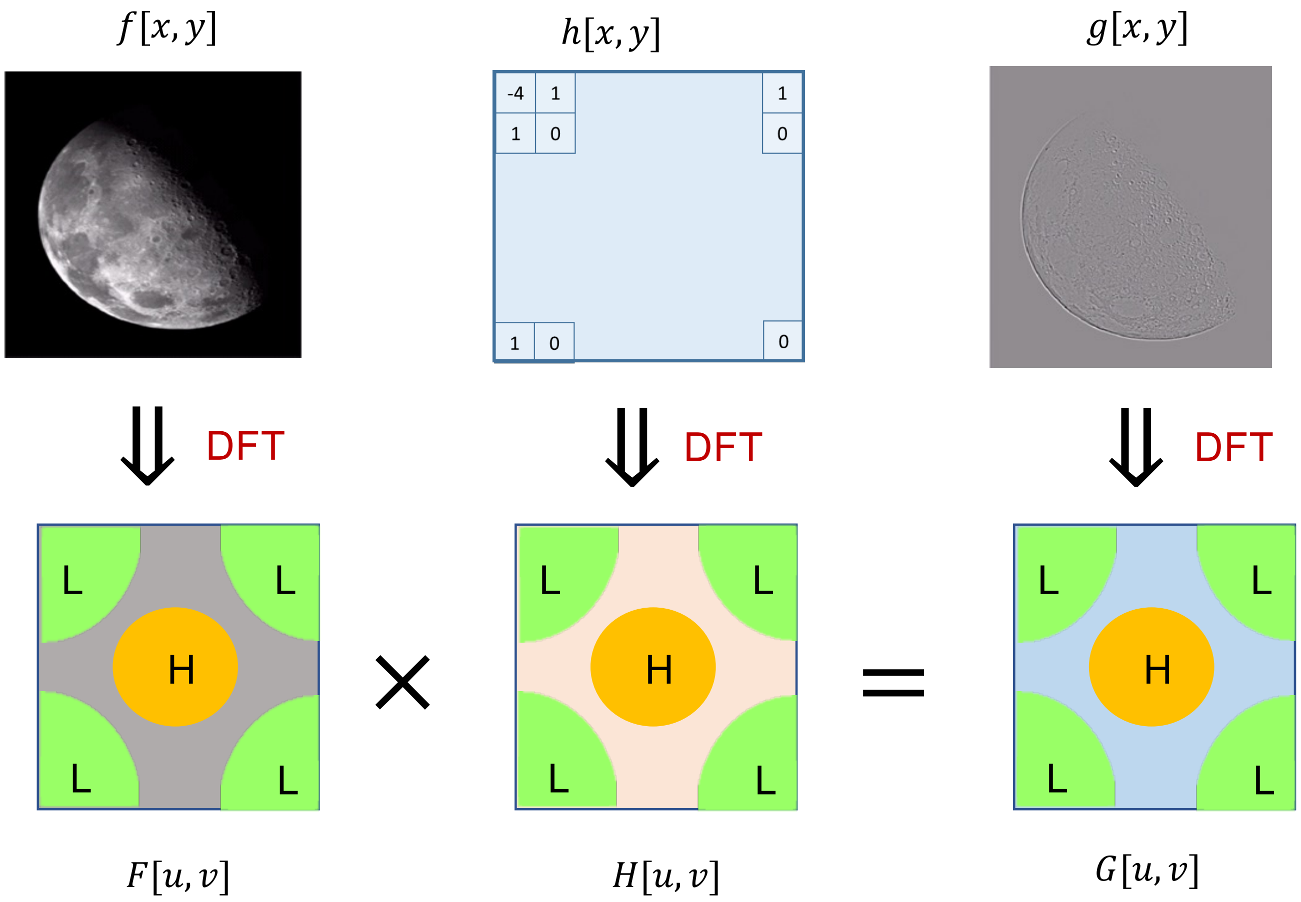
Linear Filtering in Frequency-domain (2/4)



Linear Filtering in Frequency-domain (3/4)



Linear Filtering in Frequency-domain (4/4)



$$\mathrm{F}[u,v] = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[m,n]\, e^{-j2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

Next: Image Enhancement via Spatial-domain and Frequency-domain Operators