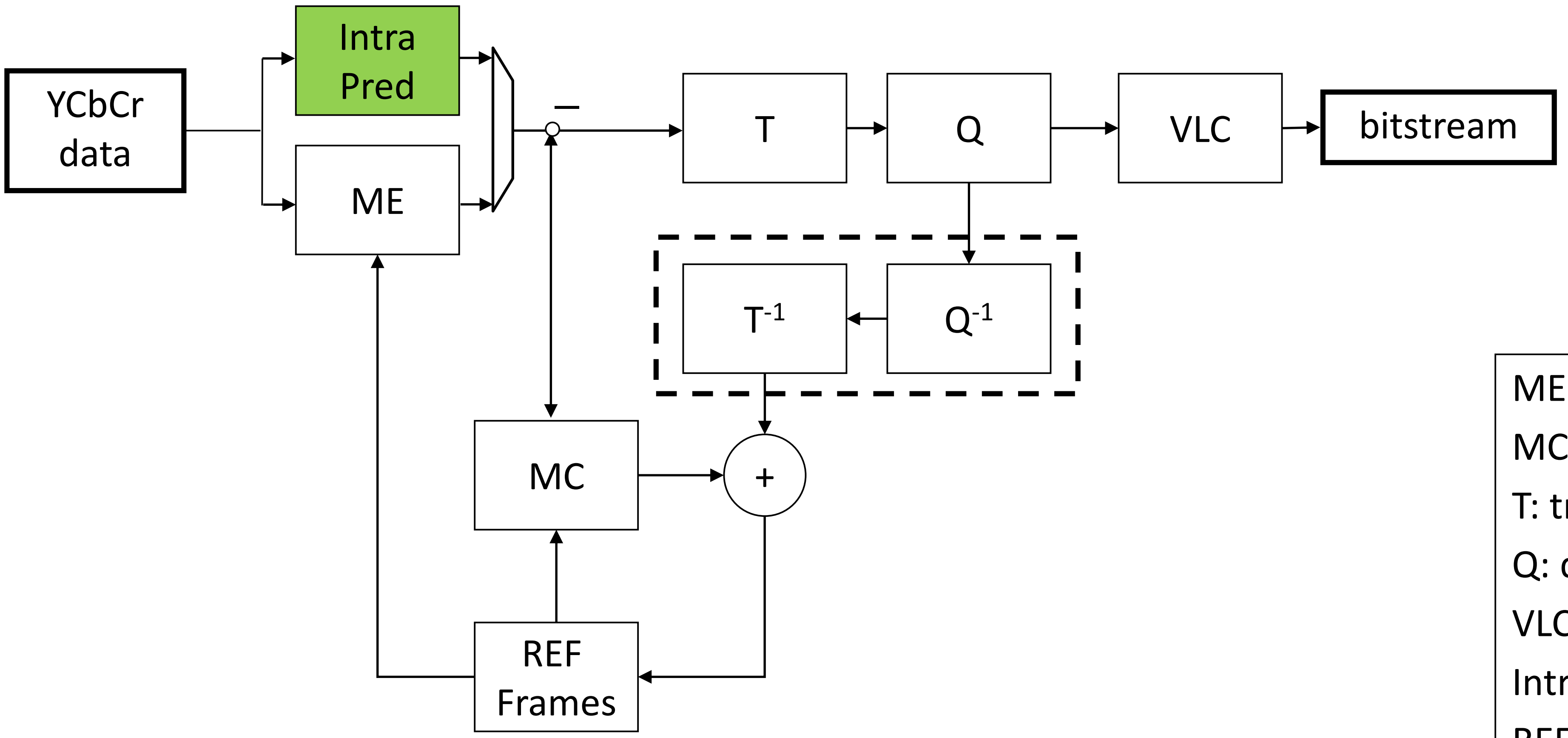
Video Compression

INSTRUCTOR: YAN-TSUNG PENG

DEPT. OF COMPUTER SCIENCE, NCCU

Intra Frame Prediction



ME: motion estimation
MC: motion compensation
T: transform coding
Q: quantization
VLC: variable length coding
Intra Pred: Intra prediction
REF frame: reference frame

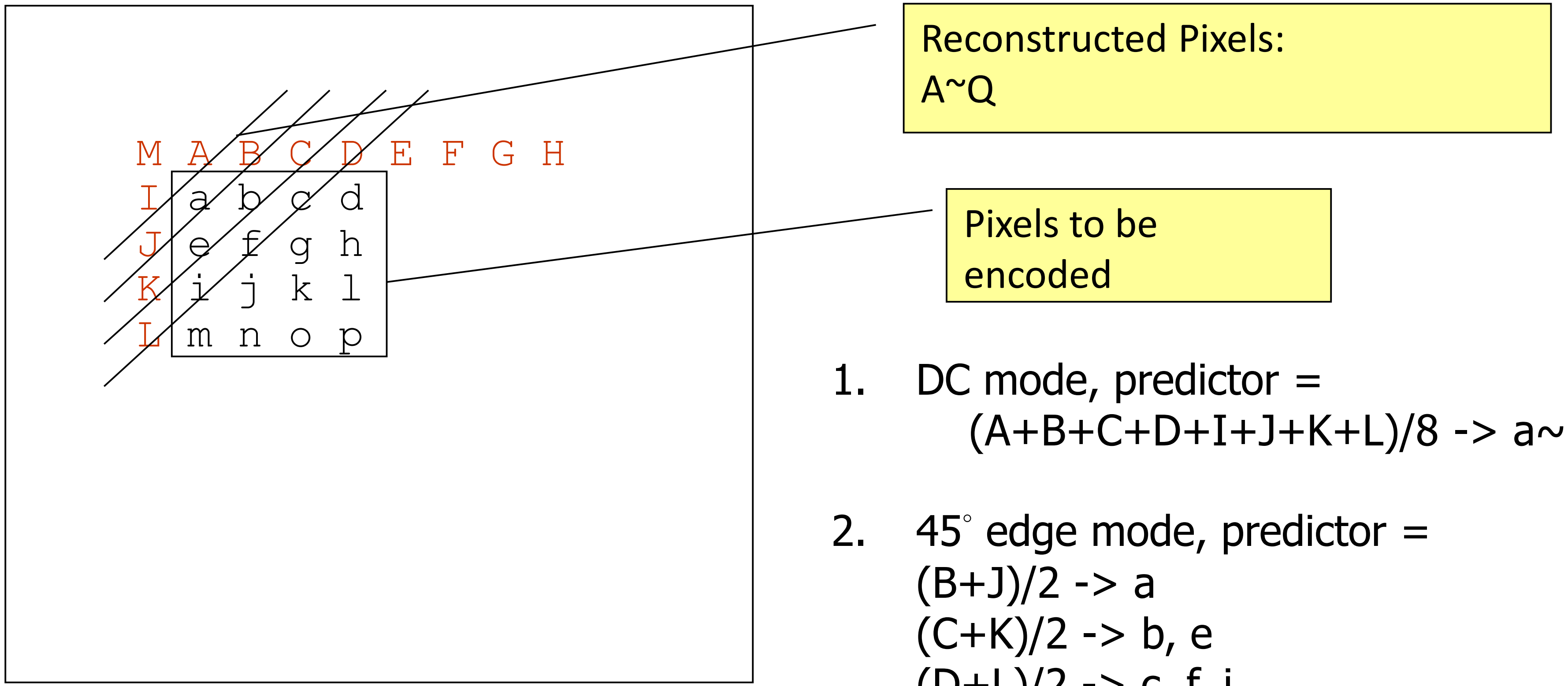
Intra-Frame Prediction

- Using Intra-Frame Prediction of H.264 as an example
- Intra modes for Luma samples
  - 9 modes for 4x4 blocks; 4 modes for 16x16 blocks
- Intra modes for Chroma samples
  - 4 modes for 8x8 blocks

M	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				

a, b, c … p are predicted from A, B, …, M that have been previously encoded

Intra-Frame Prediction

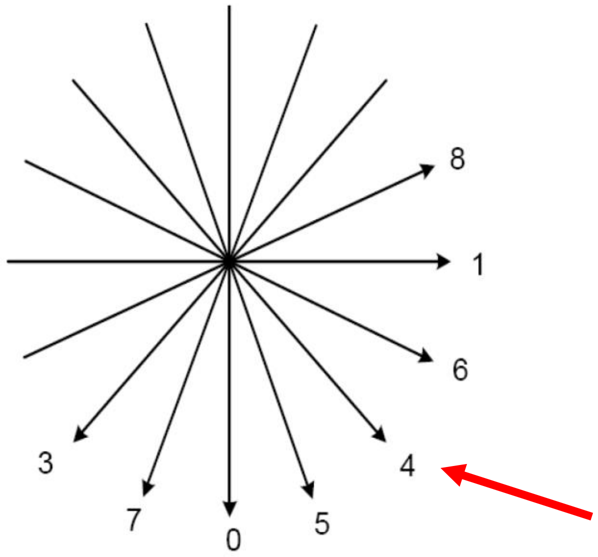


1. DC mode, predictor =
   (A+B+C+D+I+J+K+L)/8 -> a~p

2. 45° edge mode, predictor =
   (B+J)/2 -> a
   (C+K)/2 -> b, e
   (D+L)/2 -> c, f, i

Intra Luma Prediction (4x4)

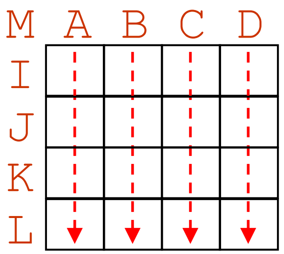
- 4 x 4 Blocks
  - There are 9 modes, 8 of which are shown below.



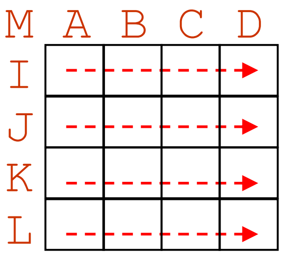
Mode 2 is the DC mode, where the predictor = (A+B+C+D+I+J+K+L)/8

Directions of Prediction

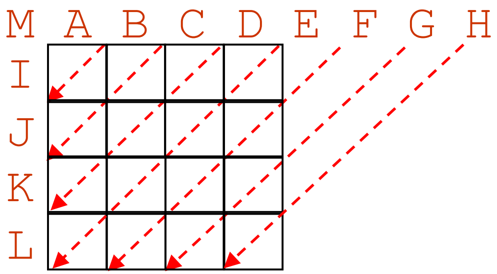
Intra Luma Prediction



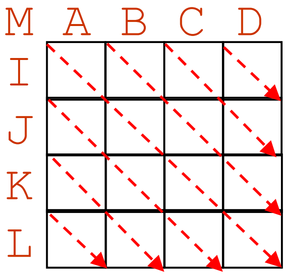
Mode 0
Vertical prediction



Mode 1
Horizontal prediction



Mode 3



Mode 4
Plane prediction

Bharathi S.H. et al

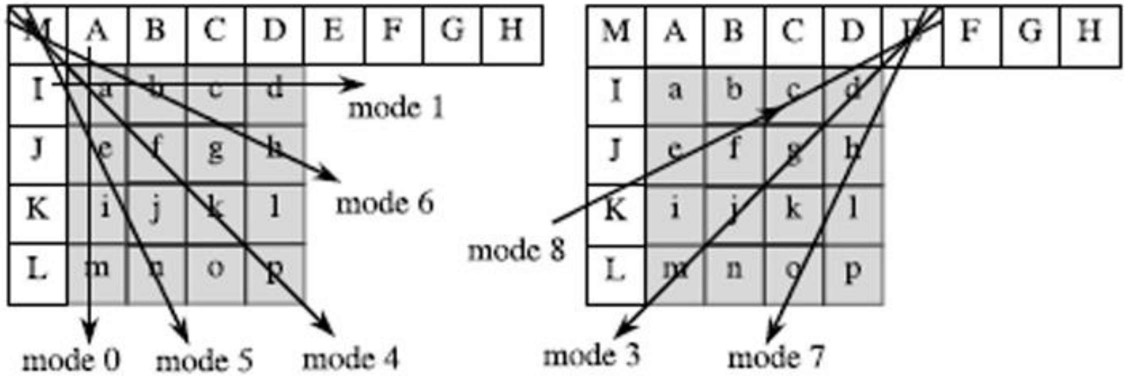


Figure 3: Nine Modes of 4×4 Intraprediction in H.264/AVC.

Bharathi et al., "Implementation of Horizontal and Vertical Intraprediction Modes or H.264 Encoder," International Journal of Electronics and Communication Engineering, 2011

Intra Luma & Chroma Prediction (16x16)

- For Intra 16x16 Blocks, it only has 4 modes
  - Mode 0: Vertical prediction
  - Mode 1: Horizontal prediction
  - Mode 2: DC prediction
  - Mode 4: Plane prediction
- For Intra Chroma prediction, it uses the same mode but for 8x8 Chroma blocks

Mode Decision Process

16x16 Luma Macroblock

Intra Modes (for I, P, and B frames)

- 4 16x16 Modes
- 9 4x4 Modes

Inter Modes (only for P and B frames)

- Choose a Reference frame
- MB partitions:
  16x16, 16x8, 8x16,
  8x8, 8x4, 4x8, 4x4
- Use of integer, half, and quarter pixel motion estimation

Computationally intensive

Mode decision is done through Rate Distortion (RD) Optimization

In the encoding process, the mode will be chosen with the least RD cost

🔶 Rate

- Refers to the **number of bits** needed to encode the block.
- Lower rate = better compression.
- Depends on:
  - Prediction residual size
  - Transform and quantization
  - Entropy coding (e.g. VLC)

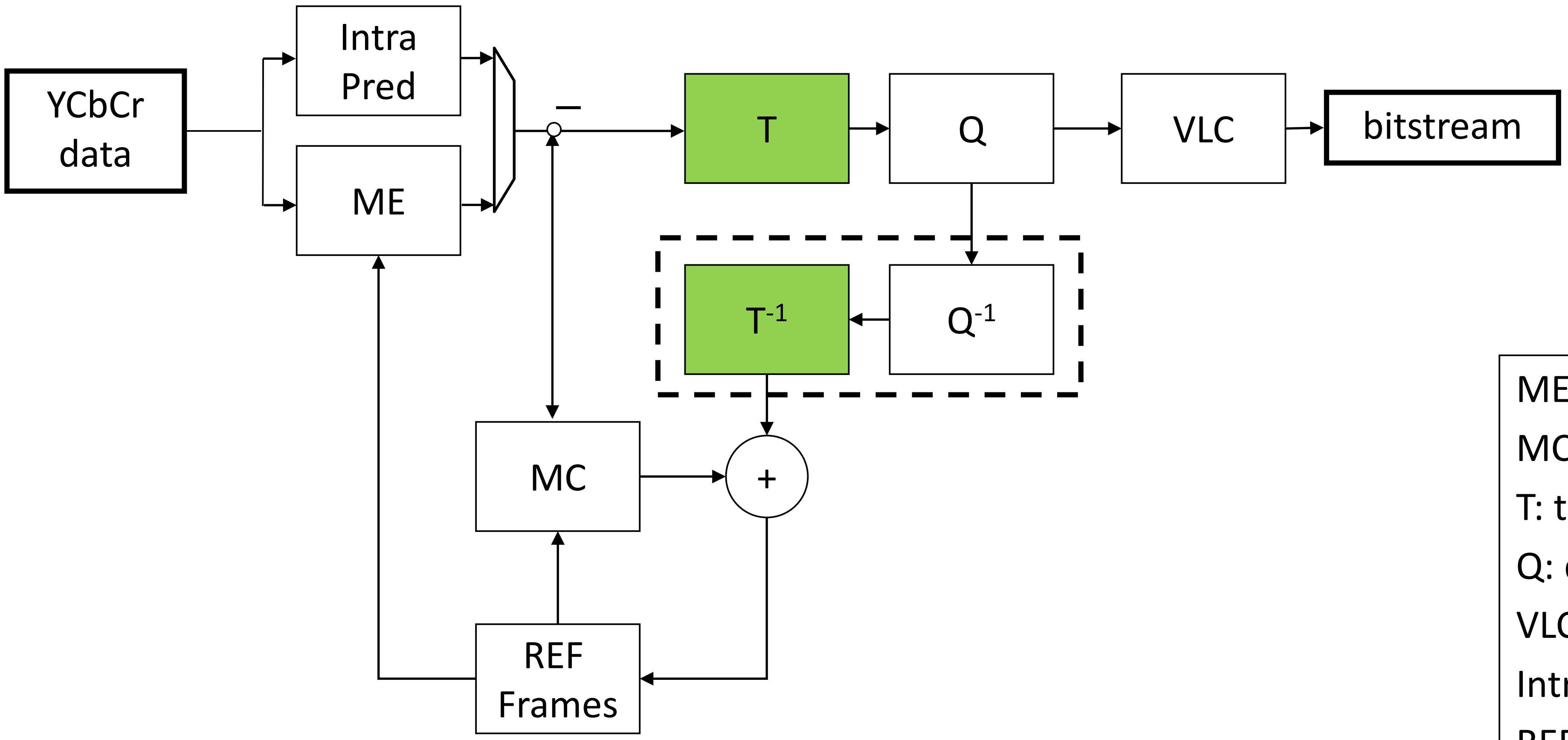
🔷 Distortion

- Measures the **difference between the original block and the reconstructed block**.
- Lower distortion = better visual quality.
- Common metric: **Mean Squared Error (MSE)** or **Sum of Absolute Differences (SAD)**.

⚖️ Trade-off in RD Optimization

- We want both **low rate** (small file size) and **low distortion** (high quality).
- But compressing more usually increases distortion.
- RD optimization balances both using a cost function:
$J = R + \lambda D$

Video Encoder Diagram



ME: motion estimation

MC: motion compensation

T: transform coding

Q: quantization

VLC: variable length coding

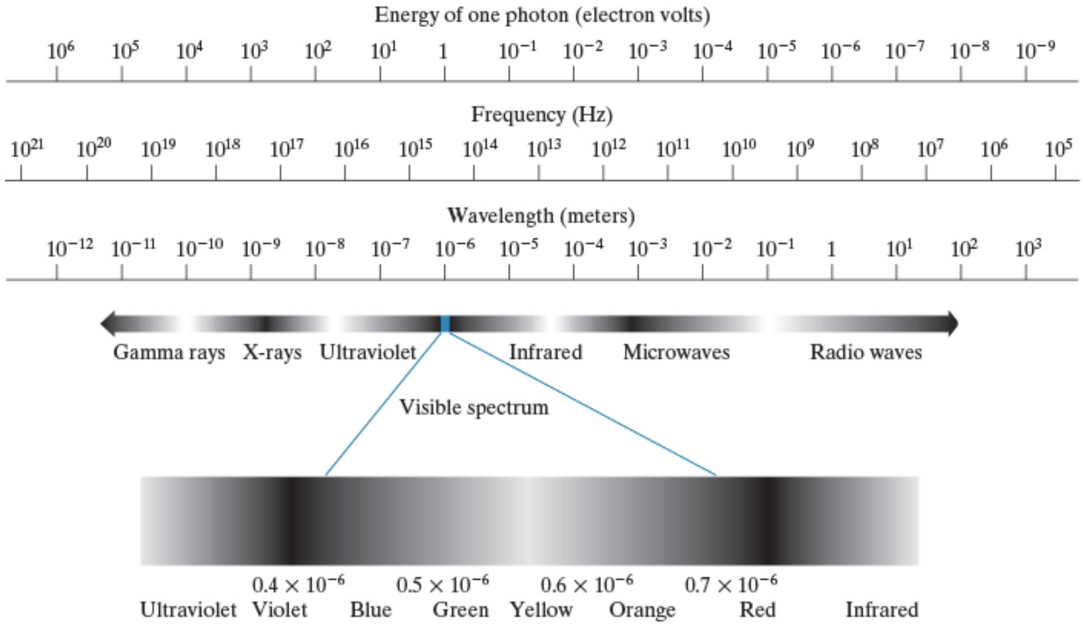
Intra Pred: Intra prediction

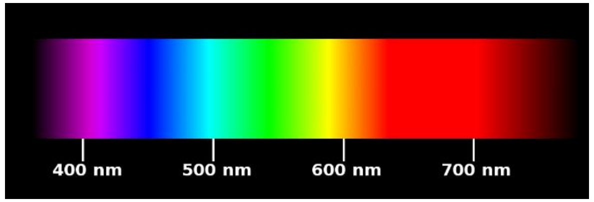
REF frame: reference frame

Transform Coding

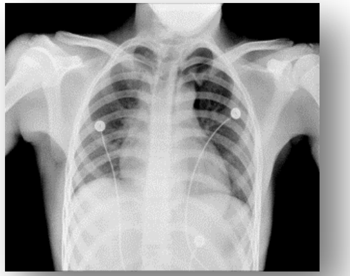
- Transform coding is a fundamental part of modern video coding standards..
- **Why It's Needed**:
  Raw spatial-domain video data is hard to compress—energy is spread evenly across pixels.
- **Core Idea**:
  Transform coding *decorrelates* data to concentrate energy, allowing less important information to be discarded with minimal visual impact.
- Common Transform Techniques:
  - DCT (Discrete Cosine Transform) – used in H.26x standards
  - DWT (Discrete Wavelet Transform) – used in JPEG-2000
- We will be focusing on DCT since it is adopted in video coding standards.

Frequency Data





Bone scan by gamma-ray imaging
(Courtesy of G.E. Medical Systems)



Chest X-ray
(Courtesy of Dr. David R. Pickens, Vanderbilt University Medical Center)

Icons made by Freepik from www.flaticon.com

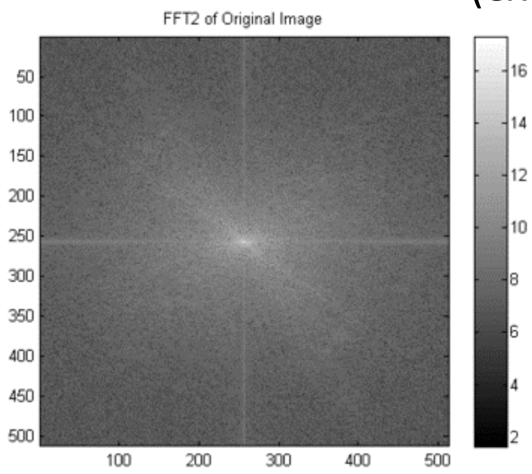
Image source: Christopher S. Baird.

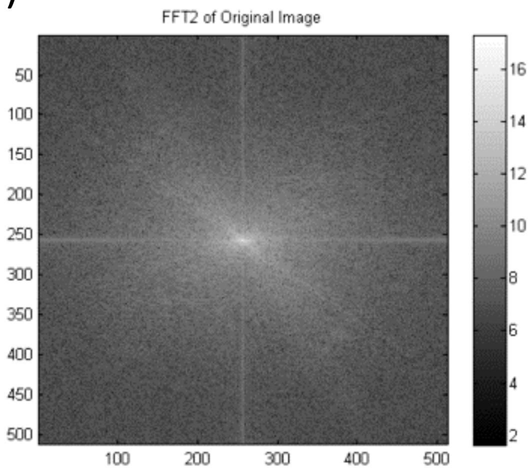
Fourier Series

- Based on the theory of Fourier series, any continuous functions can be decomposed as an infinite sum of trigonometric periodic functions, such as sines and cosines.
- The theory is the orthogonality relationships of the sine and cosine functions.
- Since a video frame (an image) can be considered as a 2D intensity function, we can use Fourier series to decompose it.

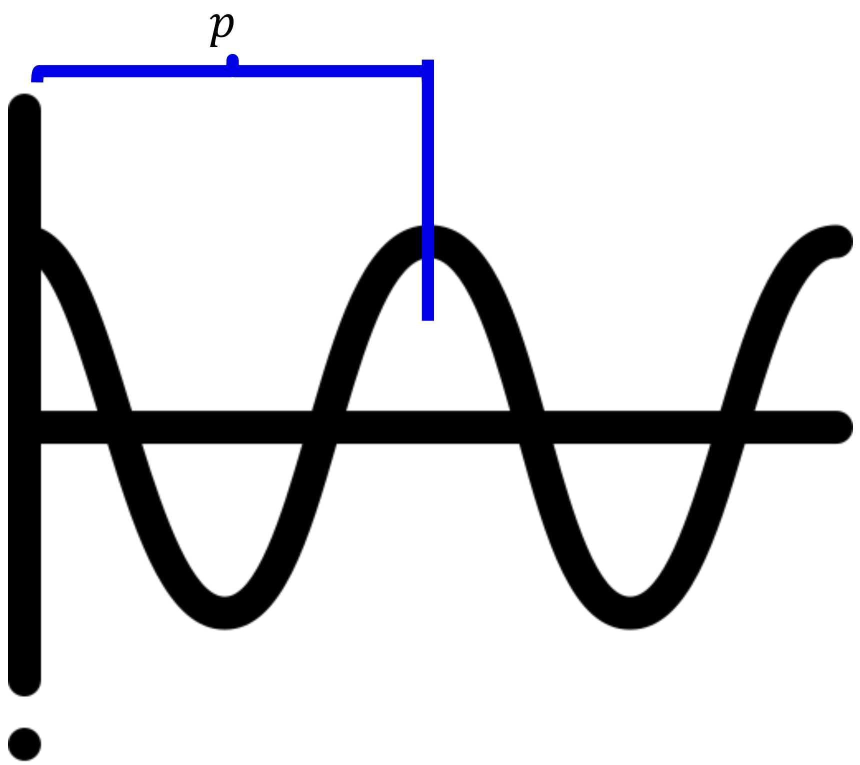
Compression vs Frequency Data

Threshold up
(ex: keep 1%)





Periodic Function

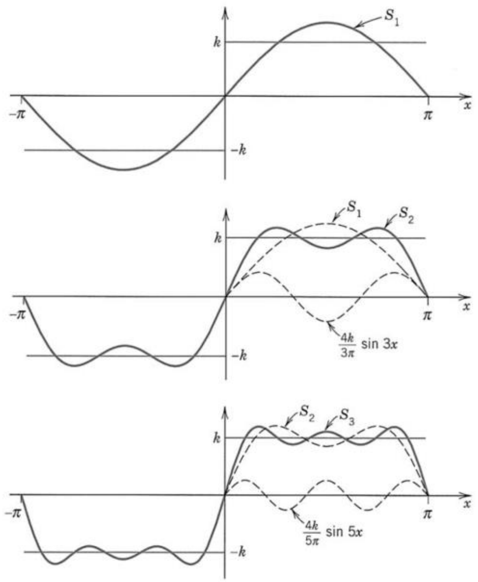


- Assume there is a periodic function $f(x)$ with its period p, x is the spatial variable or time variable, representing different spatial locations or times
- Based on Fourier series, any continuous periodic function can be decomposed by sine and cosine functions with different periods.
- $f(x) = a_0 + \sum_{i=1}^{\infty}(a_i \cos ix + b_i \sin ix)$, where the equation at the right side is called Fourier series, and a_i and b_i are Fourier coefficients.

Fourier Analysis

- General functions can be approximated by sums of trigonometric functions.
- The decomposition process is called Fourier transformation.

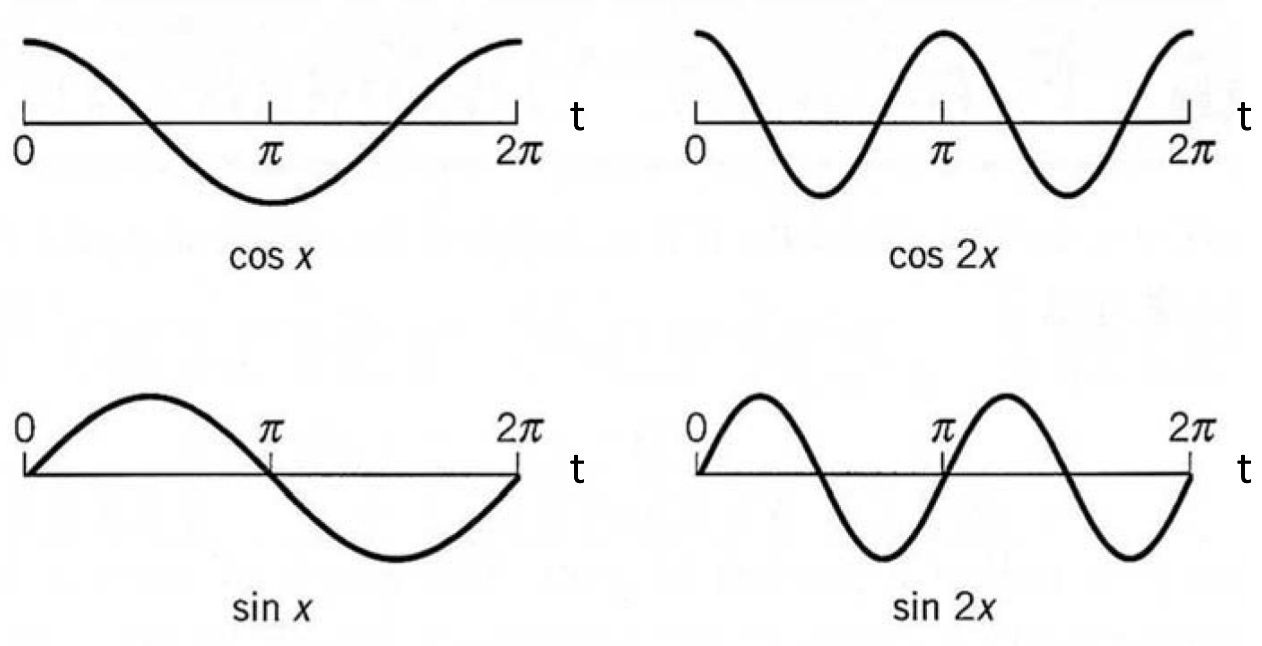
$$f(x) = a_0 + \sum_{i=1}^{\infty}(a_i \cos ix + b_i \sin ix)$$



Frequency and Period

Definition:

- Frequency – #occurrences of a periodic event per unit of time.
- Period – Time of one cycle of a periodic event
- The reciprocal of Frequency is Period
  - Frequency = $\frac{1}{\text{Period}}$



The period of $\sin x$ is $2\pi \equiv$ The frequency of $\sin x$ is $\frac{1}{2\pi}$

How about $\sin nx$?

Orthogonality of sines and cosines

- sine and cosine functions with different frequencies are orthogonal
- As we know, two vectors a, b being orthogonal means their inner product equals 0 ($a \cdot b = 0$)
- Similarly, two functions f, g being orthogonal means their inner product also equals to 0 ($\int_{-\infty}^{\infty} f(x)g(x)\,dx = 0$)
- Since the period of $\sin nx$ and $\cos nx$ are both $\frac{2\pi}{n}$, their being orthogonal between $[-\pi, \pi]$ means they are orthogonal between $[-\infty, \infty]$

Orthogonality of sines and cosines

- $\int_{-\pi}^{\pi} \sin nx \sin mx \, dx = 0, if\ n \neq m$
- $\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = 0, if\ n \neq m$
- $\int_{-\pi}^{\pi} \sin nx \cos mx \, dx = 0$
- $\int_{-\pi}^{\pi} \sin nx \sin nx \, dx = \pi$ and $\int_{-\pi}^{\pi} \cos nx \cos nx \, dx = \pi$
  - $\int_{-\pi}^{\pi} \sin^2 nx \, dx = \int_{-\pi}^{\pi} \frac{1}{2} - \frac{\cos 2nx}{2} dx = \frac{x}{2} - \frac{\sin 2nx}{4n}\Big|_{-\pi}^{\pi} = \pi$
  - $\int_{-\pi}^{\pi} \cos^2 nx \, dx = \int_{-\pi}^{\pi} \frac{\cos 2nx}{2} + \frac{1}{2} dx = \frac{\sin 2nx}{4n} + \frac{x}{2}\Big|_{-\pi}^{\pi} = \pi$

- $2\sin a \sin b = -\cos(a+b) + \cos(a-b)$
- $2\cos a \cos b = \cos(a+b) + \cos(a-b)$
- $2\sin a \cos b = \sin(a+b) + \sin(a-b)$
- $\int_{-\pi}^{\pi} \cos nx \, dx = 0, if\ n \geq 1$
- $\int_{-\pi}^{\pi} \sin nx \, dx = 0, if\ n \geq 1$
- $\cos 2x = \cos^2 x - \sin^2 x$
- $\sin^2 x + \cos^2 x = 1$
- $\frac{d\sin x}{dx} = \cos x$
- $\frac{d\cos x}{dx} = -\sin x$

Trigonometric Identities

- Proof
- $2\sin a \sin b = -\cos(a+b) + \cos(a-b)$
- $2\cos a \cos b = \cos(a+b) + \cos(a-b)$
- $2\sin a \cos b = \sin(a+b) + \sin(a-b)$

❑ Using Euler's formula: $e^{i\theta} = \cos\theta + i\sin\theta$

$$e^{i(a+b)} = \cos(a+b) + i\sin(a+b) = (\cos a + i\sin a)(\cos b + i\sin b)$$

$$= (\cos a \cos b - \sin a \sin b) + i(\cos a \sin b + \sin a \cos b)$$

$$\Rightarrow \cos(a+b) = \cos a \cos b - \sin a \sin b$$

$$\Rightarrow \sin(a+b) = \cos a \sin b - \sin a \cos b$$

Fourier Coefficients

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx$$

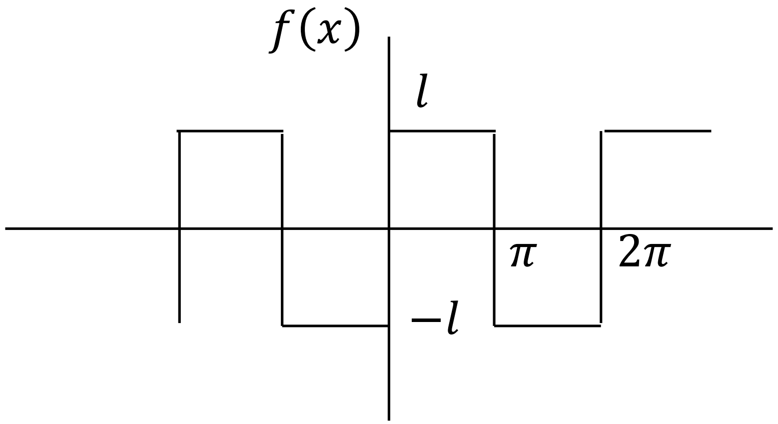
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx$$

Example

$$f(x) = \begin{cases} -l, & if\ \ (2n-1)\pi < x < 2n\pi; \\ l, & if\ 2n\pi < x < (2n+1)\pi. \end{cases} \qquad n \in Z$$

Period of $f(x)$ is 2π, which means $f(x) = f(x + 2\pi)$



$$f(x) = \begin{cases} -l, & if -\pi < x < 0; \\ l, & if\ 0 < x < \pi. \end{cases}, f(x) = f(x+2\pi)$$

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx = 0$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} -l\cos nx\,dx + \int_{0}^{\pi} l\cos nx\,dx\right] = \frac{2l}{\pi}\frac{\sin nx}{n}\Bigg|_0^{\pi} = 0$$

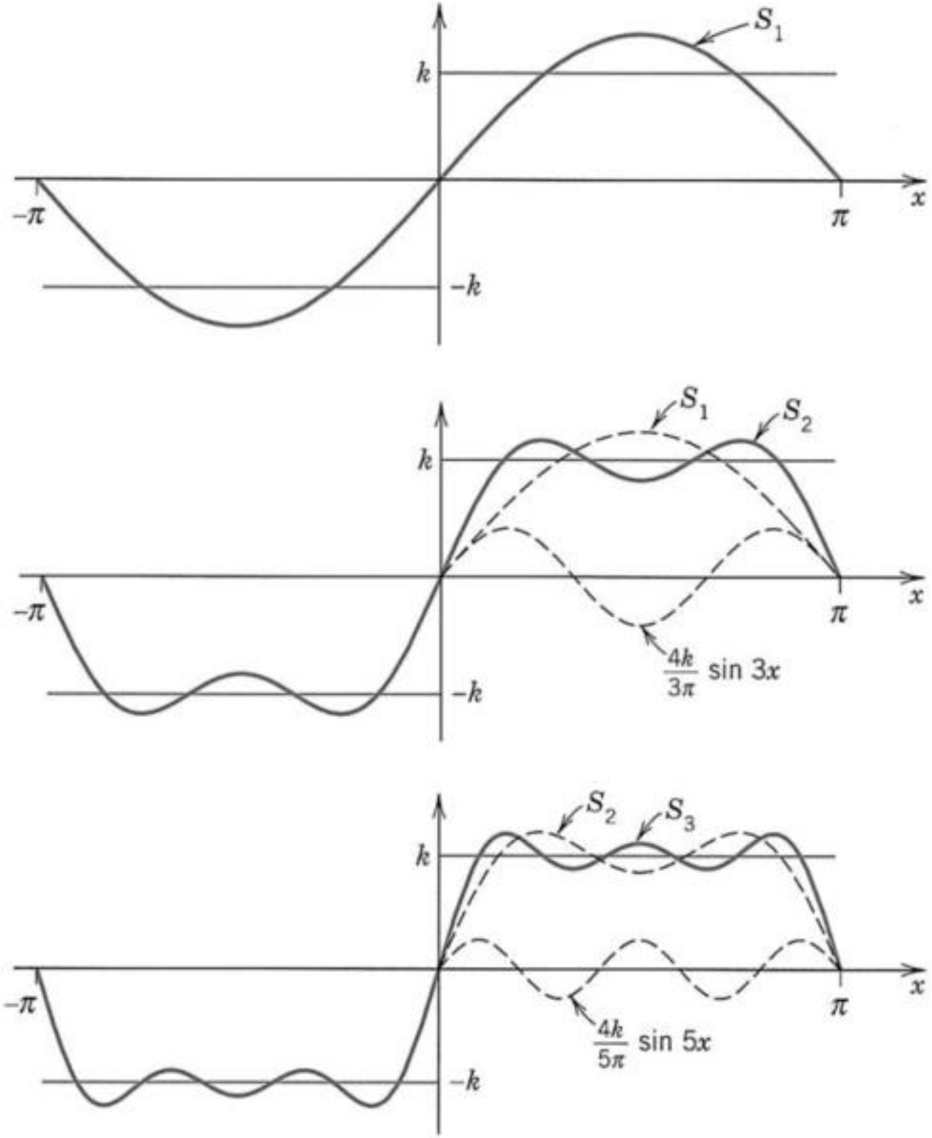
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} -l\sin nx\,dx + \int_{0}^{\pi} l\sin nx\,dx\right] = -\frac{2l}{\pi}\frac{\cos nx}{n}\Bigg|_0^{\pi}$$

$$= \frac{2l}{n\pi}(1 - \cos n\pi) = \frac{2l}{n\pi}(1 - (-1)^n), \qquad n = 1,2, \ldots$$

$n: odd$ $\quad b_{2k+1} = \dfrac{4l}{(2k+1)\pi}, k = 0, 1, 2 \ldots$ $\qquad f(x) = \dfrac{4l}{\pi}(\sin x + \dfrac{1}{3}\sin 3x + \cdots)$

$n: even$ $\quad b_{2k} = 0$

$$f(x) = \frac{4l}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \cdots\right)$$

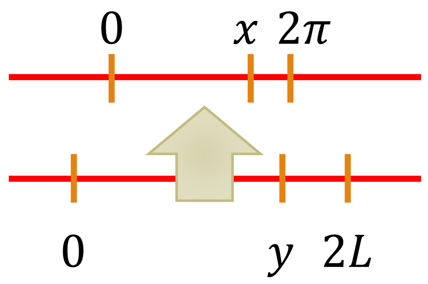


Fourier Series Expansion on the Interval $[-L, L]$

Suppose that we have a periodic function $f(y)$ with arbitrary period $2L$ (generalizing the special case $p = 2\pi$)

Since $f(x)$ with period $2\pi \rightarrow$ change the variable to make its period $2L$

To chnage to the new period $y = 2L$ from $x = 2\pi$, we can imagine transforming y back to x as



$$\frac{2\pi}{x} = \frac{2L}{y} \rightarrow x = \frac{\pi}{L} y$$

$$f(y) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos n\frac{\pi}{L} y + b_n \sin n\frac{\pi}{L} y\right)$$

$$\int_{-\pi}^{\pi} \sin^2 nx\, dx = \int_{-\pi}^{\pi} \cos^2 nx\ dx = \pi \rightarrow \quad \int_{-L}^{L} \sin^2 n\frac{\pi}{L} y\, dy = \int_{-L}^{L} \cos^2 n\frac{\pi}{L} y\, dy = L$$

$$\int_{-L}^{L} \sin^2 n\frac{\pi}{L} y\, dy = \int_{-L}^{L} \frac{1}{2} - \frac{\cos 2n\frac{\pi}{L} y}{2} dy = \frac{y}{2} - \frac{\sin 2n\frac{\pi}{L} y}{4n\frac{\pi}{L}}\Bigg|_{-L}^{L} = L$$

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos \frac{n\pi}{L} x + b_n \sin \frac{n\pi}{L} x)$$

$$a_0 = \frac{1}{2L} \int_{-L}^{L} f(x)\, dx$$

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{n\pi}{L} x\, dx$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin \frac{n\pi}{L} x\, dx \qquad n = 1,2,3, \ldots$$

We can use the complex number with Euler formula to simplify Fourier Series:

$$e^{inx} = \cos nx + i\sin nx \rightarrow \begin{cases} \cos nx = \dfrac{e^{inx} + e^{-inx}}{2} \\ \sin nx = \dfrac{e^{inx} - e^{-inx}}{2i} \end{cases}$$

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \frac{e^{inx} + e^{-inx}}{2} + b_n \frac{e^{inx} - e^{-inx}}{2i}) = a_0 + \sum_{n=1}^{\infty} \left(\frac{e^{inx}}{2}(a_n - ib_n) + \frac{e^{-inx}}{2}(a_n + ib_n) \right)$$

$So, let\ f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$, where $c_0 = a_0$

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + c_{-n} e^{-inx} \qquad c_n = \frac{a_n - ib_n}{2} \qquad c_{-n} = \frac{a_n + ib_n}{2}$$

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx \, dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx \, dx$$

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

$$c_n = \frac{a_n - ib_n}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx, \text{where n} = \pm 1, \pm 2, \ldots$$

If the period of $f(x)$ is $2L$

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi}{L}x}$$

$$c_n = \frac{a_n - ib_n}{2} = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-i\frac{n\pi}{L}x}\, dx, \text{where n} = \pm 1, \pm 2, \ldots$$

If $f(x)$ is not a periodic function, we can assume its period is ∞.

For this, we can have various frequencies for the sine and cosine functions.

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi}{L}x}$$

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-i\frac{n\pi}{L}x}\,dx, \text{where n} = \pm 1, \pm 2, \ldots$$

Let $u_n = \frac{n\pi}{L}$, $\Delta u = \frac{\pi}{L}$, and $F(s) = \int_{-L}^{L} f(x)e^{-isx}\,dx$

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-i\frac{n\pi}{L}x}\,dx = \frac{1}{2L}F(u_n)$$

$\rightarrow f(x) = \sum_{n=-\infty}^{\infty} \frac{F(u_n)}{2L} e^{iu_nx} = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} F(u_n)\, e^{iu_nx}\,\Delta u$

Let $L \rightarrow \infty$, $\Delta u \rightarrow 0$, $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(u)\, e^{iux}\,du$

$\Rightarrow$

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(u)\, e^{iux}\,du$$

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-iux}\,dx$$

, or

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F(u)e^{iux}\,du$$

$$F(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-iux}\,dx$$

Linearity

- **Linearity**.

$x(t)$ → $y(t)$

$\alpha x(t)$ → $\alpha y(t)$

Homogeneity

$x_1(t)$ → $y_1(t)$

$x_2(t)$ → $y_2(t)$

$x_1(t) + x_2(t)$ → $y_1(t) + y_2(t)$

Additivity

- **Fourier Transform** is **linear**.

Fourier Transform

Fourier Transform

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(u) e^{iux}\, du$$

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-iux}\, dx$$

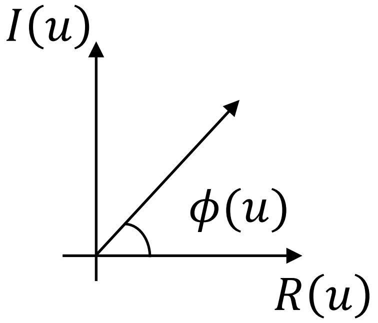
Inverse
Fourier Transform

Fourier Transform

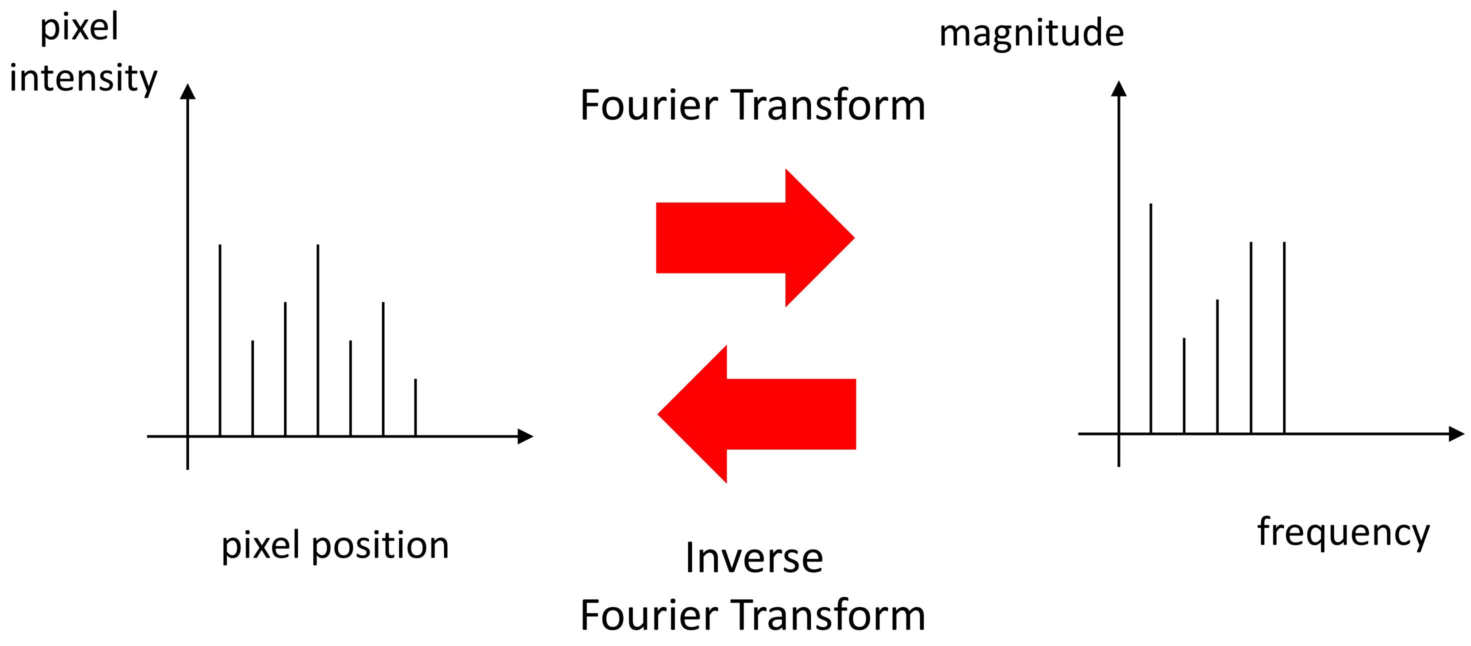
- A function after Fourier transform consists of both real and imaginary parts
  - $F(u) = R(u) + i\, I(u)$
  - $F(u) = |F(u)|e^{i\phi(u)}$, where $\phi(u) = \tan^{-1}\frac{I(u)}{R(u)}$ and $|F(u)| = \sqrt{R^2(u) + I^2(u)}$

$|F(u)|$: Fourier spectrum

$\phi(u)$: (Fourier) phase angle



Spatial ⟷ Frequency Domain



Spatial domain signal

Frequency domain signal

Spatial ⟷ Frequency Domain

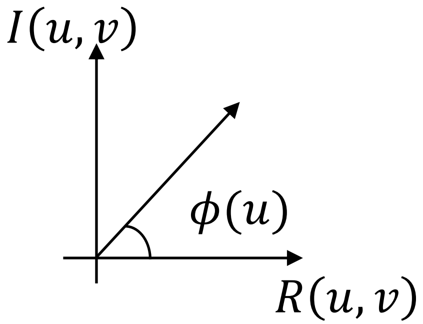
- Considering a frame is a 2D function, we can extend 1D Fourier transform to 2D Fourier transform

$$f(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{i(ux+vy)}\,du\,dv$$

$$F(u,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-i(ux+vy)}\,dx\,dy$$

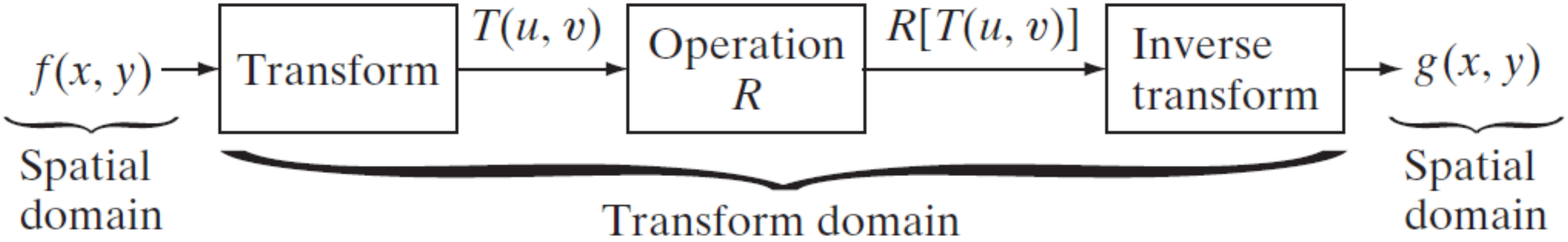
$|F(u,v)|$: Fourier spectrum

$\phi(u,v)$: (Fourier) phase angle



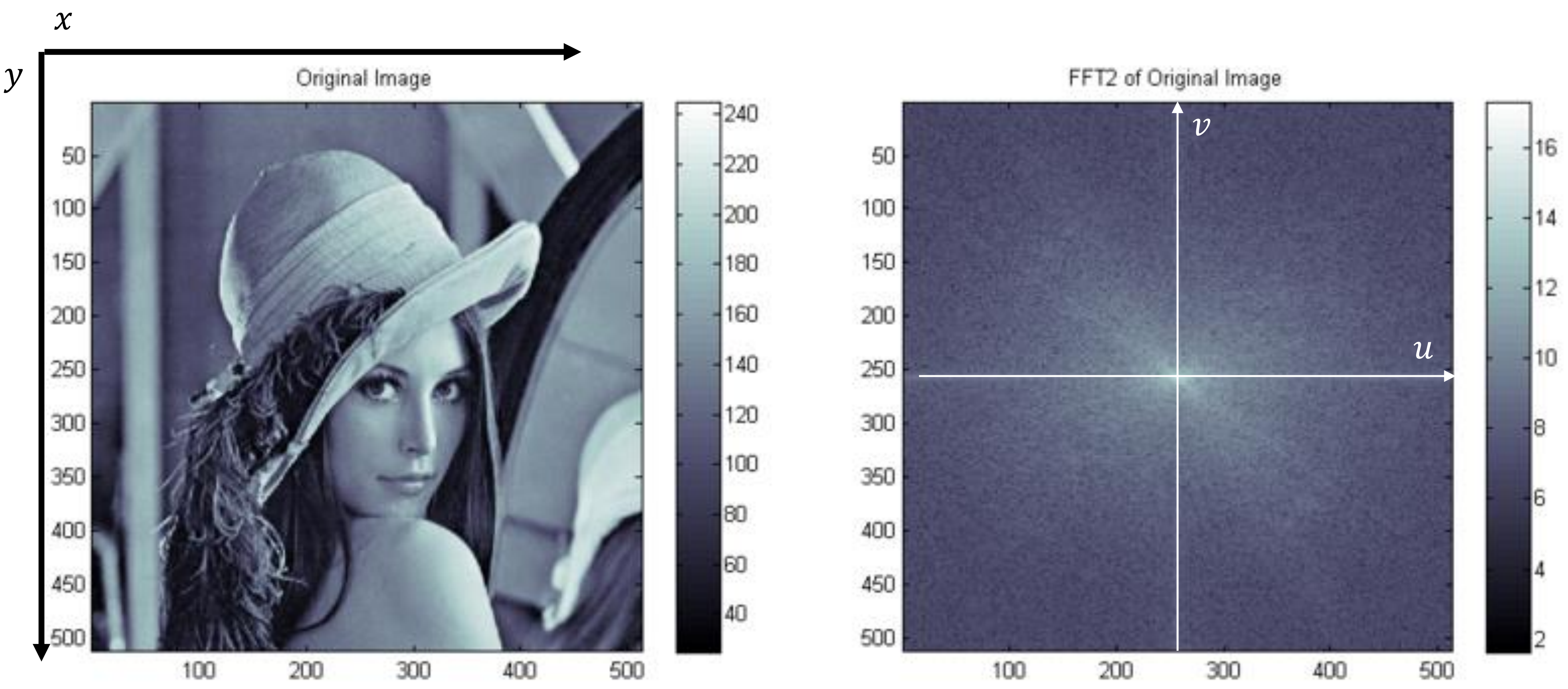
Since a function transformed consists of the real and imaginary part, we can only display its magnitude $|F(u,v)|$

Image Transform



Fourier transform

Spatial Domain -> Frequency Domain



Discrete Fourier Transform

- Since image or video data are discrete, we should do discrete Fourier transform
- Assume we have sampled data as $f(x_0), f(x_0+\Delta x), f(x_0+2\Delta x), \dots, f(x_0+(N-1)\Delta x)$, denoted as $f(0), f(1), f(2), \dots, f(N-1)$.

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-i\frac{2\pi x}{N}u}, u = 0, 1, 2, \dots, N-1$$

$$f(x) = \sum_{x=0}^{N-1} F(u)\, e^{i\frac{2\pi u}{N}x}, x = 0, 1, 2, \dots, N-1$$

$f(x) = \sum_{n=-\infty}^{\infty} F(u)e^{i\frac{n\pi}{L}x}$, with period $2L$

Let $L = \frac{N}{2} \rightarrow$ period N

2D Discrete Fourier Transform

- Since image or video data are discrete, we should do discrete Fourier transform
- Assume we have sampled data as $f(x_0, y_0), f(x_0, y_0 + \Delta y), f(x_0, y_0 + 2\Delta y), \dots, f(x_0 + (M-1)\Delta x, y_0 + (N-1)\Delta y)$, denoted as $f(0,0), f(0,1), f(0,2), \dots, f(M-1, N-1)$.

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{x=0}^{N-1} f(x,y)\, e^{-i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}, u = 0,1,2,\dots, M-1; v = 0,1,2,\dots, N-1$$

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v)\, e^{i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}, x = 0,1,2,\dots, M-1; x = 0,1,2,\dots, N-1$$

2D Discrete Fourier Transform

- Assume $M = N$.

$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\, e^{-i2\pi\left(\frac{ux}{N}+\frac{vy}{N}\right)},\ u,v = 0,1,2,\dots,N-1$$

$$f(x,y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} F(u,v)\, e^{i2\pi\left(\frac{ux}{N}+\frac{vy}{N}\right)},\ x,y = 0,1,2,\dots,N-1$$

$$F(0,0) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) = N\mu_f,\ \text{where } \mu_f = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)$$

In terms of images, this value increases as the image size increases

Separability for Fourier Transform

$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\, e^{-i2\pi\left(\frac{ux}{N}+\frac{vy}{N}\right)},\ u,v = 0,1,2,\dots,N-1$$

$$= \frac{1}{N}\sum_{x=0}^{N-1} e^{-i2\pi\frac{ux}{N}} \sum_{y=0}^{N-1} f(x,y)\, e^{-i2\pi\frac{vy}{N}}$$

$$= \underbrace{\frac{1}{N}\sum_{x=0}^{N-1} F(x,v)e^{-i2\pi\frac{ux}{N}}}_{N \text{ times 1D Fourier transform}},\ \text{where } F(x,v) = N\underbrace{\left(\frac{1}{N}\sum_{y=0}^{N-1} f(x,y)e^{-i2\pi\frac{vy}{N}}\right)}_{N \text{ times 1D Fourier transform}}$$

2D Fourier transform = $2N$ times 1D Fourier transform

Complexity Comparison: 2D vs. 1D Fourier Transform (FT)

Assume the image size is N^2

$$F(u,v) = \frac{1}{N} \underbrace{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1}}_{N \times N} f(x,y)\, e^{-i2\pi\left(\frac{ux}{N}+\frac{vy}{N}\right)}, u,v = 0,1,2,\ldots,N-1$$

Time complexity of transforming the whole image using 2D FT is $O(N^4)$

$$F(u,v) == \frac{1}{N} \underbrace{\sum_{x=0}^{N-1} F(x,v) e^{-i2\pi\frac{ux}{N}}}_{N}, \text{where } F(x,v) = N\left(\frac{1}{N} \underbrace{\sum_{y=0}^{N-1} f(x,y) e^{-i2\pi\frac{vy}{N}}}_{N}\right)$$

Transforming the whole image using 2D FT requires $O(N^2 \times 2N) = O(N^3)$

Periodicity for Fourier Transform

- Fourier transform is a periodic function

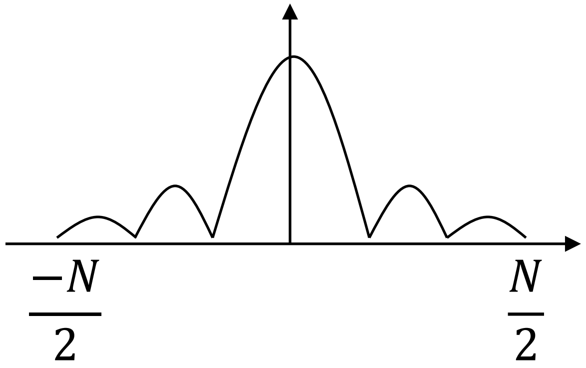
$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-i\frac{2\pi u}{N}x}, u = 0, 1, 2, \dots, N-1$$

$$f(x) = \sum_{x=0}^{N-1} F(u)\, e^{i\frac{2\pi u}{N}x}, x = 0, 1, 2, \dots, N-1$$

Period: N

$$F(u) = F(u+N)$$

$$F(u) = F^*(-u) \rightarrow \text{conjugate symmetry}$$



Translation for Fourier Transform

$f(x) = \sum_{x=0}^{N-1} F(u)\, e^{i\frac{2\pi u}{N}x}, x = 0, 1, 2, \dots, N-1$ $\quad$ $F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-i\frac{2\pi u}{N}x}, u = 0, 1, 2, \dots, N-1$

$$f(x)e^{i\frac{2\pi u_0}{N}x} \Rightarrow \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{i\frac{2\pi u_0}{N}x} e^{-i\frac{2\pi u}{N}x} = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-i\frac{2\pi(u-u_0)}{N}x} = F(u-u_0)$$

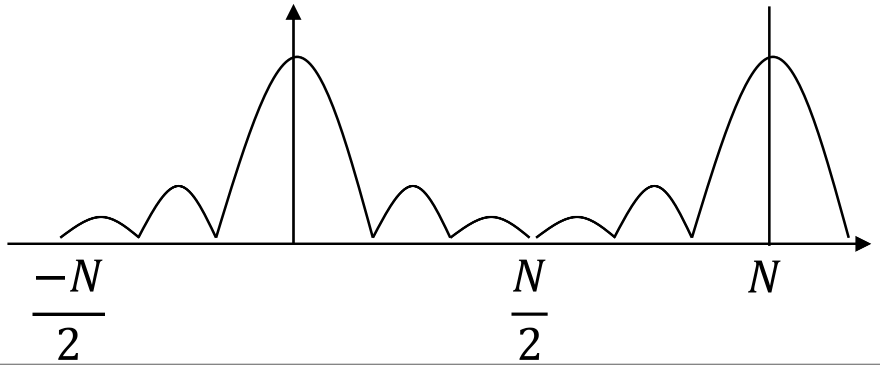
$$F(u)e^{-i\frac{2\pi u}{N}x_0} \Rightarrow \frac{1}{N}\sum_{x=0}^{N-1} F(u)\, e^{i\frac{2\pi u}{N}(x-x_0)} = f(x-x_0)$$

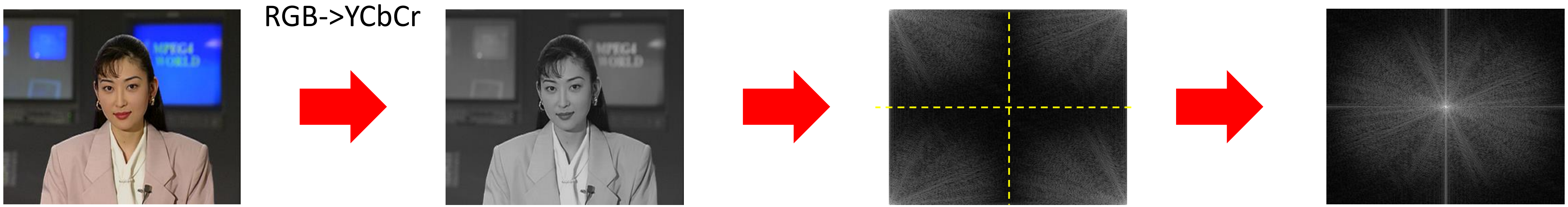
$$\left|F(u)e^{-i\frac{2\pi u}{N}x_0}\right| = |F(u)| \qquad \because \left|e^{-i\frac{2\pi u}{N}x_0}\right| = 1$$

Distributivity for Fourier Transform

- $f(x) + g(x) = \sum_{x=0}^{N-1} \big(F(u) + G(u)\big)\, e^{i\frac{2\pi u}{N}x}, x = 0, 1, 2, \dots, N-1$
- However, $f(x)g(x) \neq \sum_{x=0}^{N-1} \big(F(u)G(u)\big)\, e^{i\frac{2\pi u}{N}x}$

Demonstration



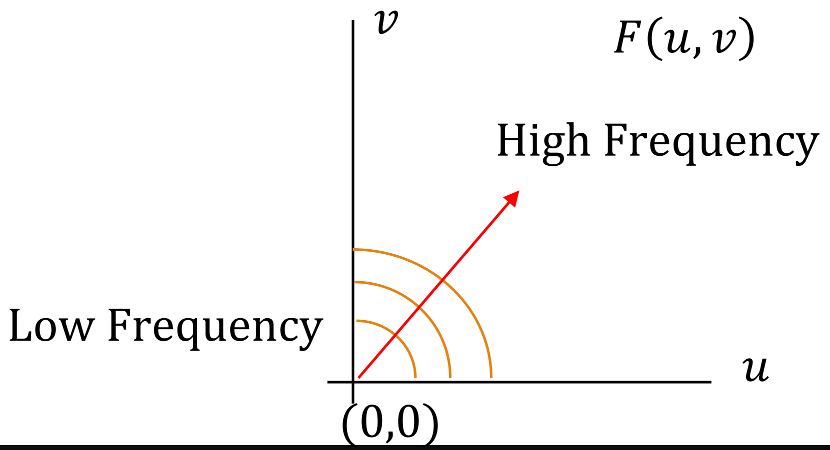


Y

$|F(Y)|$

$|F(Y)|$
shifted to center
(fftshift)

- To demonstrate Fourier transform, we will show the transformed magnitudes
- The magnitudes for low frequencies are large whereas those for high frequencies are extremely small , so we will use log to reduce the magnitude differences as $\log(1 + |F(Y)|)$
- At last, to show it as an image, you should normalize it.



Periodicity for Fourier Transform

- Fourier transform is a periodic function

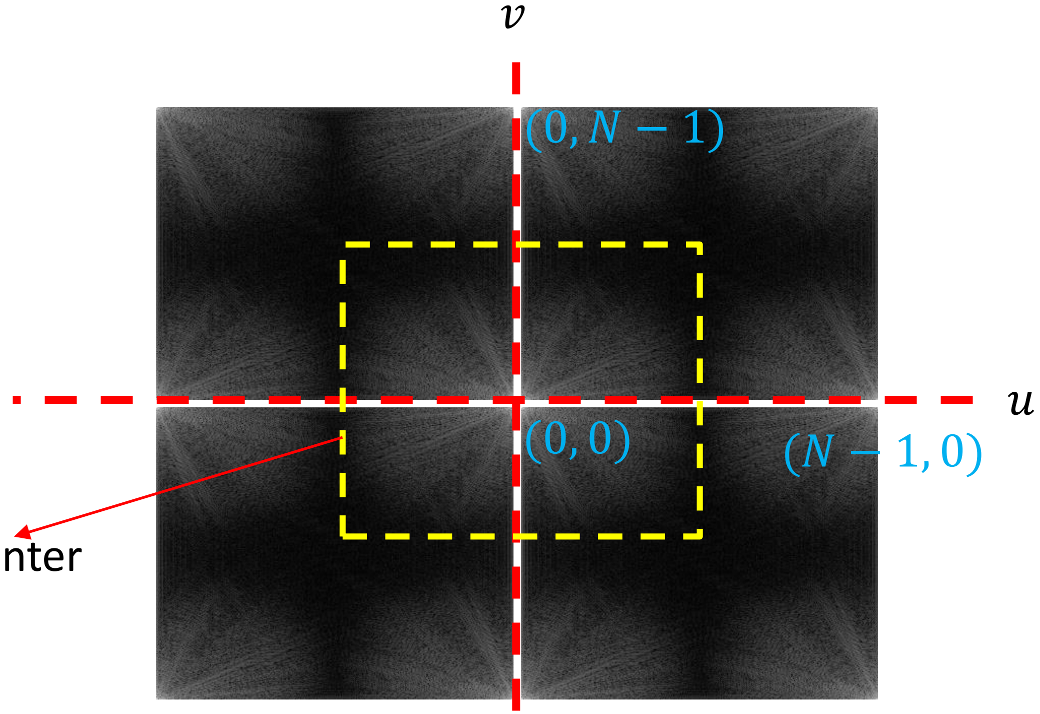
$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\, e^{-i\frac{2\pi(ux+vy)}{N}},\ u,v = 0,1,2,\dots,N-1$$

$$F(u,v) = F(u+N, v+N)$$ Period: N

$$f(x)e^{i\frac{2\pi u_0}{N}x} \Leftrightarrow F(u-u_0)$$

$$F\left(u-\frac{N}{2}\right) \Leftrightarrow f(x)e^{i\pi x} = f(x)(-1)^x$$

$$F\left(u-\frac{N}{2}, v-\frac{N}{2}\right) \Leftrightarrow f(x,y)e^{i\pi(x+y)} = f(x,y)(-1)^{x+y}$$



shifted to center

Discrete Cosine Transform (DCT)

- The difference between DCT and Fourier Transform is that DCT only uses cosine functions, so it’s a real function not a complex function.
- 2-D Discrete Cosine Transform
  - Forward Transform (for N x N blocks)

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

  - Inverse Transform (for N x N blocks)

$$f(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)f(u,v) \cos\frac{(2x+1)u\pi}{2N} \cos\frac{(2y+1)v\pi}{2N}$$

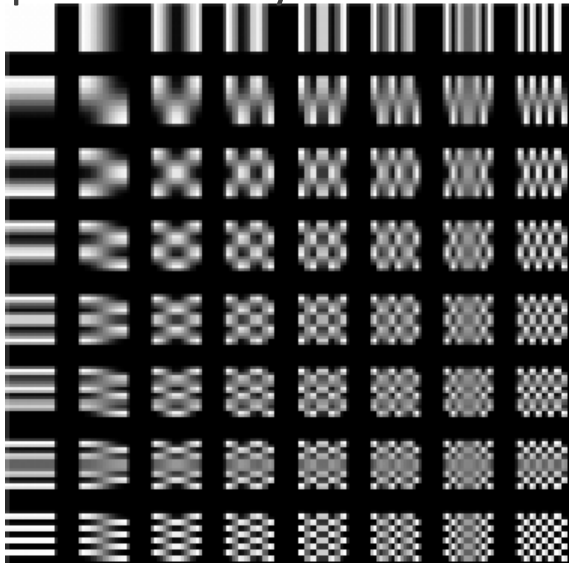
$$u,v = 0,1,\ldots N-1$$

$$C(t) = \begin{cases} \frac{2}{\sqrt{N}}, & t = 0 \\ 2 \cdot \sqrt{\frac{2}{N}}, & t \neq 0 \end{cases}$$

$$x,y = 0,1,\ldots N-1$$

DCT for 8x8 Blocks in Video Compression

- DCT in video compression usually takes the spatial samples in 9 bits (signed values) to produce the coefficients in 12 bits. The dynamic range of the coefficients is [-2048:+2047].
  - Why signed values? -> ME
- It applies to one block at a time
- Any 8x8 image block can be represented by a linear combination of the following basis functions



$$F(u,v)=\tfrac{2}{N}C(u)C(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$$

DC Component

$$F(u,v)=\frac{C(u)C(v)}{4}\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)\cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16}$$

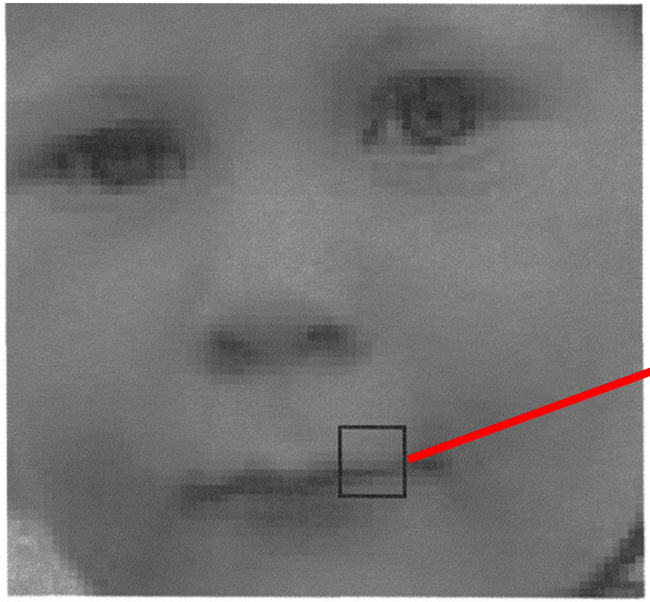
$$C(t)=\begin{cases}\frac{2}{\sqrt{N}}, & t=0\\ 2\cdot\sqrt{\frac{2}{N}}, & t\neq 0\end{cases}$$

- Let $u=0, v=0$

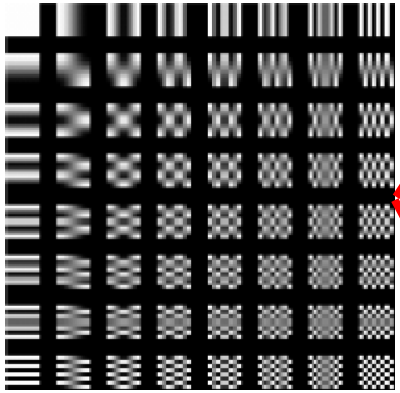
$F(0,0)=\frac{1}{8}\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)$ which stands for the average luma/chroma value for a block, called the DC component

- DC is the most important coefficient among 64 coefficients.

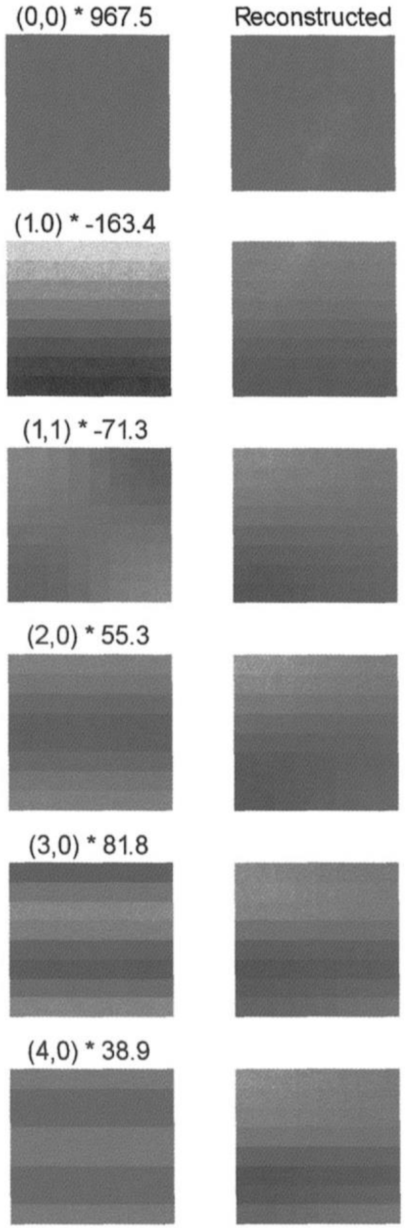
Example of DCT



8x8 DCT



967.5	-6.3	-1	1.3	0.1
-164.4	-71.3	1	.2	-1.6
55.3	13.6	-	.7	-0.6
81.8	-8.9	-	0.1	1
38.9	-12.9	-	0.3	1
-7.6	-8.4	[illegible]	.1	2.1
-14.7	1.4	4	0.2	1.1
-22.3	5	4	.2	-0.3



DCT in Video Compression

- Data decorrelation
- Real number computations
- Separablity (apply 1D DCT)
- Still Works well for error residuals in video compression

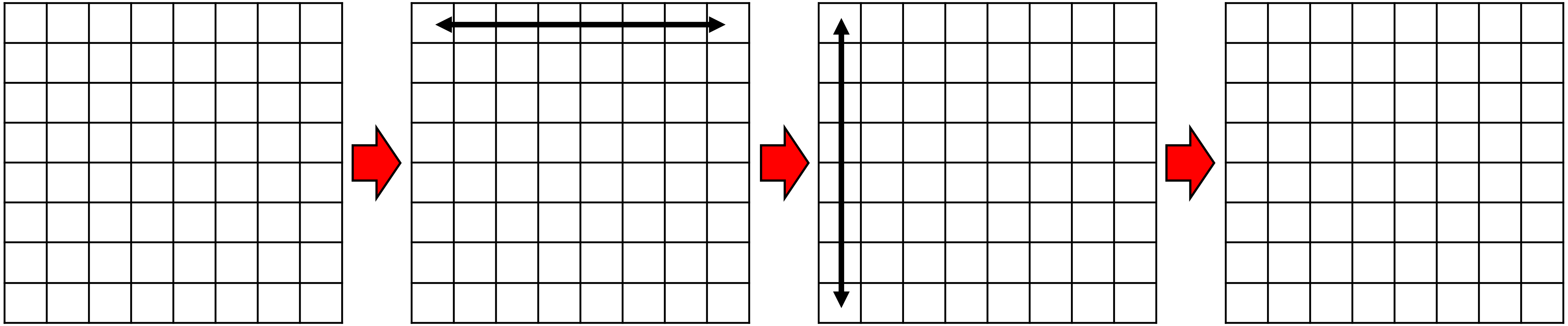
Separable Transform for DCT

- Forward DCT:

$$F(u) = \frac{C(u)}{4} \sum_{x=0}^{7} f(x) \cos \frac{(2x+1)u\pi}{16}$$

$$C(t) = \begin{cases} \frac{2}{\sqrt{N}}, & t = 0 \\ 2 \cdot \sqrt{\frac{2}{N}}, & t \neq 0 \end{cases}$$

Time complexity: $O(2 \times 8^3)$



8x8 block | 1-D FDCT on rows | 1-D FDCT on columns | 8x8 DCT coefficients