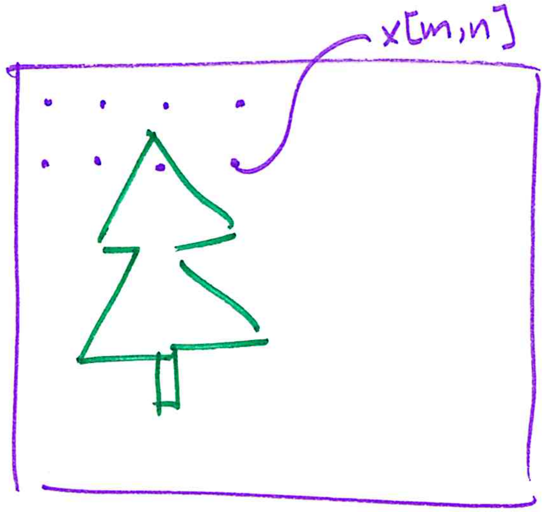
Lecture 21: Image Processing

An image is a 2-dimension signal

$x =$



samples of x = image pixels

$x(u,v)$ = "continuous time" signal

u, v are locations in space / image plane.

we can sample $x(u,v)$ to get a digital image:

$x[m,n] = x(m\Delta, n\Delta)$ where Δ = spacing between samples/pixels (analogous to T_s)

for $m = 0, 1, \ldots, M-1,$
$n = 0, 1, \ldots, N-1$ $\rightarrow$ $M \times N$ pixel image

If we store pixel values $x[m,n]$ for $m = 0, \dots, M-1$
$n = 0, \dots, N-1$,
then we call this a digital image.

Ex. image with 3264×2448 = 8 Megapixels (millions of pixels)

if we use 24 bits to represent each $x[m,n]$ value
(8 bits for red, 8 bit for blue, 8 bits for green)
↳ 256 possible levels of red light brightness

⇒ total # bits = 196 Mb = 24 MB (8 bits = 1 byte)
this is what bitmap (.BMP) does

in contrast, JPEG compression → 3.2 MB

Discrete-time Fourier Transform ~~Series~~

1D

$x[n]$ $n = 0, \ldots, N-1$ is discrete time signal

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi \frac{kn}{N}}$$

where

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{kn}{N}}$$

$\Rightarrow$ $x[n]$ = sum of weighted sinusoids.

2D DFT

$x[m,n]$, $m = 0, \ldots, M-1$, $n = 0, \ldots, N-1$ is digital image.

$$x[m,n] = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} X[k,l] e^{j2\pi \frac{km}{M}} e^{j2\pi \frac{ln}{N}}$$

where

$$X[k,l] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m,n] e^{-j2\pi \frac{km}{M}} e^{-j2\pi \frac{ln}{N}}$$

$\Rightarrow$ $x[m,n]$ = sum of weighted (2d) sinusoids of the form

$$e^{j2\pi \frac{km}{M}} e^{j2\pi \frac{ln}{N}}$$

$\frac{k}{M}$: frequency in vertical direction

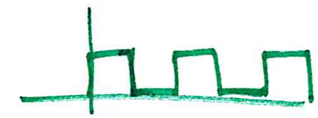
$\frac{l}{N}$ = frequency in horizontal direction

$$X[k,l] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m,n]\, e^{-j2\pi \frac{km}{M}}\, e^{-j2\pi \frac{ln}{N}}$$

$$= \underbrace{\sum_{n=0}^{N-1} \underbrace{\left[\sum_{m=0}^{M-1} x[m,n]\, e^{-j2\pi \frac{km}{M}} \right]}_{\text{this is the DFT of the } n^{th} \text{ column of } x \text{ — call this } Z[k,n]} \cdot e^{-j2\pi \frac{ln}{N}}}_{\text{this is the DFT of the } k^{th} \text{ row of } Z}$$

$\Rightarrow$ 2D DFT $\equiv$ 1D DFT of each column $\rightarrow$ 1D DFT of each row of result.

$X[k, l]$ = complex number that tells us the weight of each 2-d sinusoid building block



in images, edges are sharp changes and correspond to high frequency ~~condu~~ content

- not smooth like sinusoid
- we need many high-frequency sinusoids to represent signal
- sharp changes

$\Updownarrow$

high frequency content

$$X[k, l] = X[M-k, N-l]^*$$

for real-valued images.