Image Processing

INT 302

Siyue Yu

Some questions

- Why image/video processing is important?
- How the human vision system works ?
- What is light ?
- What is image ?

What is image?

What is image ?

- A monochrome image is a two-dimensional light intensity function $f(x,y)$
  - x, y denote spatial coordinates
  - $f(x,y)$ is proportional to the brightness (or *gray level*) of the image at that point

- A digital image is a function $f(x,y)$ which has been discretized both in space and in brightness
  - It is a matrix of picture elements, called *pixels*

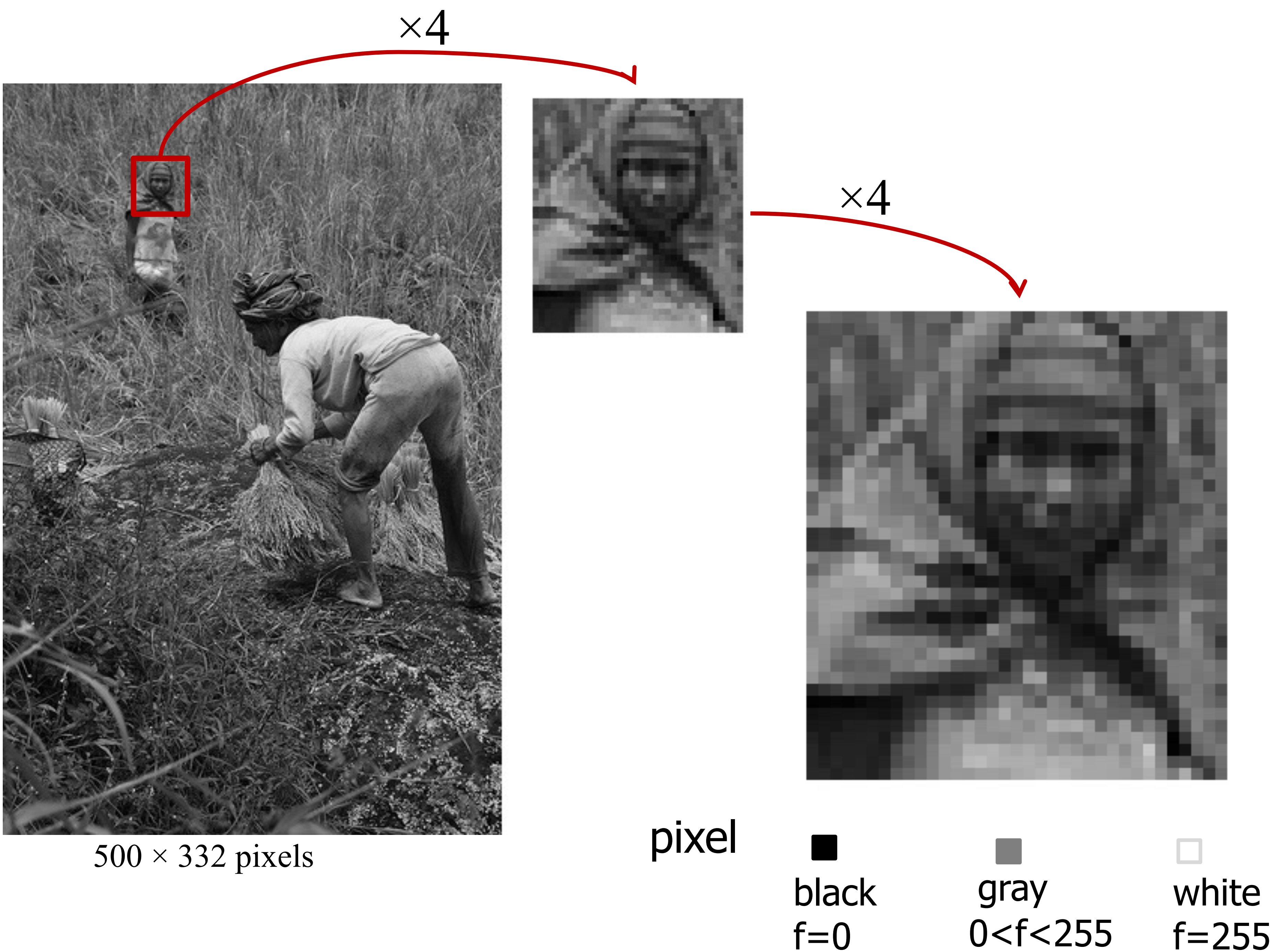
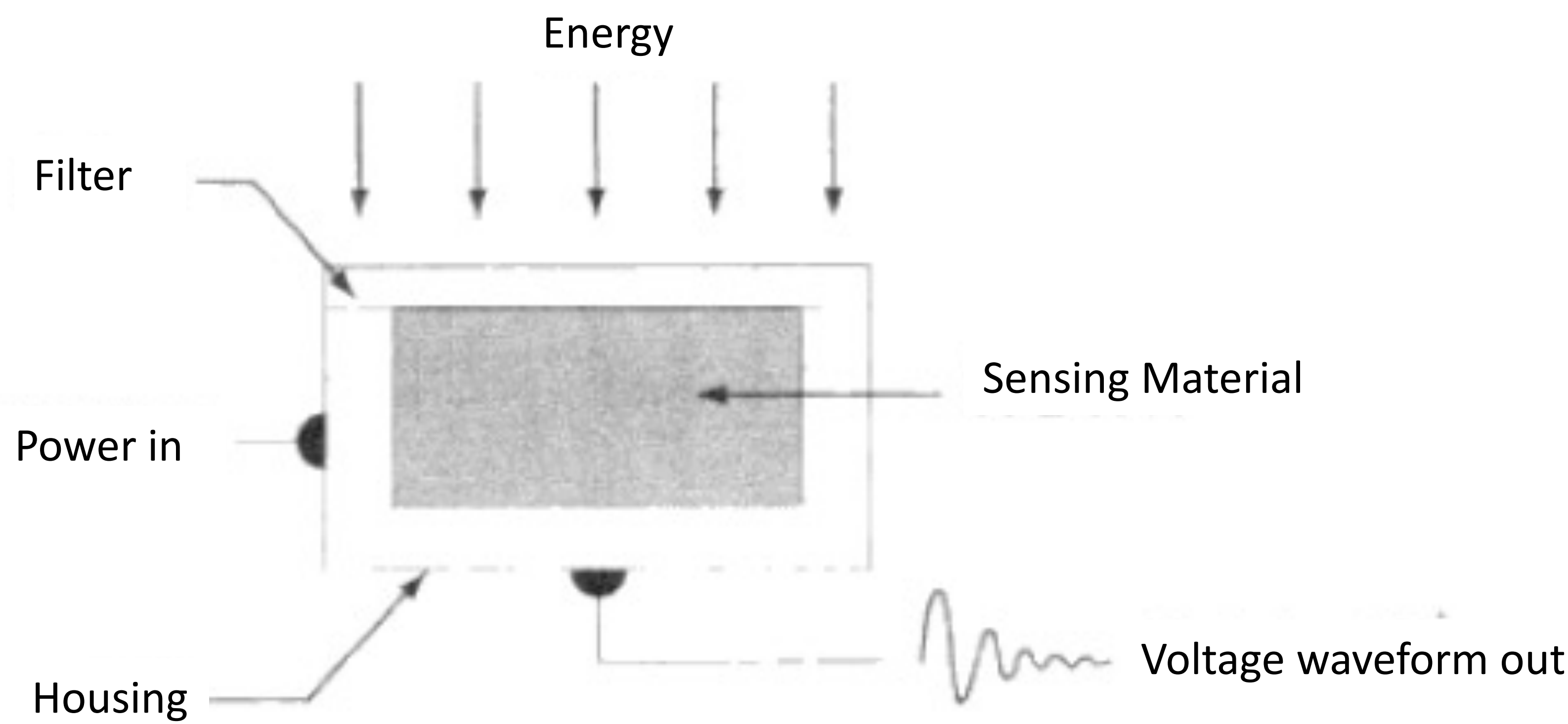

×4
×4
500 × 332 pixels
pixel
black
f=0
gray
0<f<255
white
f=255

Image acquisition

- Two elements necessary:
  - Physical device sensitive to a band in the electromagnetic spectrum (e.g., x-ray, infrared, visible,…), which produces an electrical signal proportional to the level of energy sensed
  - Analog-to-digital converter (if the physical device is not able to produce digital output directly)

Image sensing

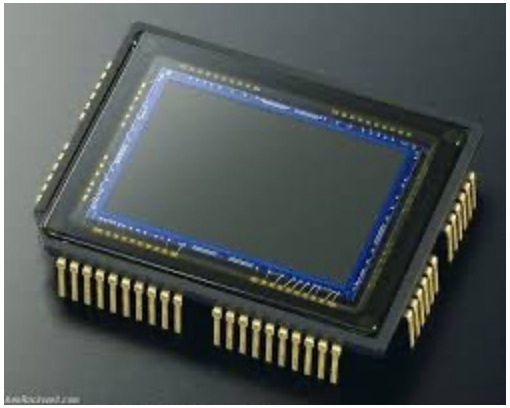


Example: X-ray imaging

- X-ray source directed to an object
- Medium sensitive to x-ray placed on the other side of the object (film, TV camera, discrete detectors for digital imaging)
- The sensitive medium generate an image of tissues having various degrees of x-ray absorption

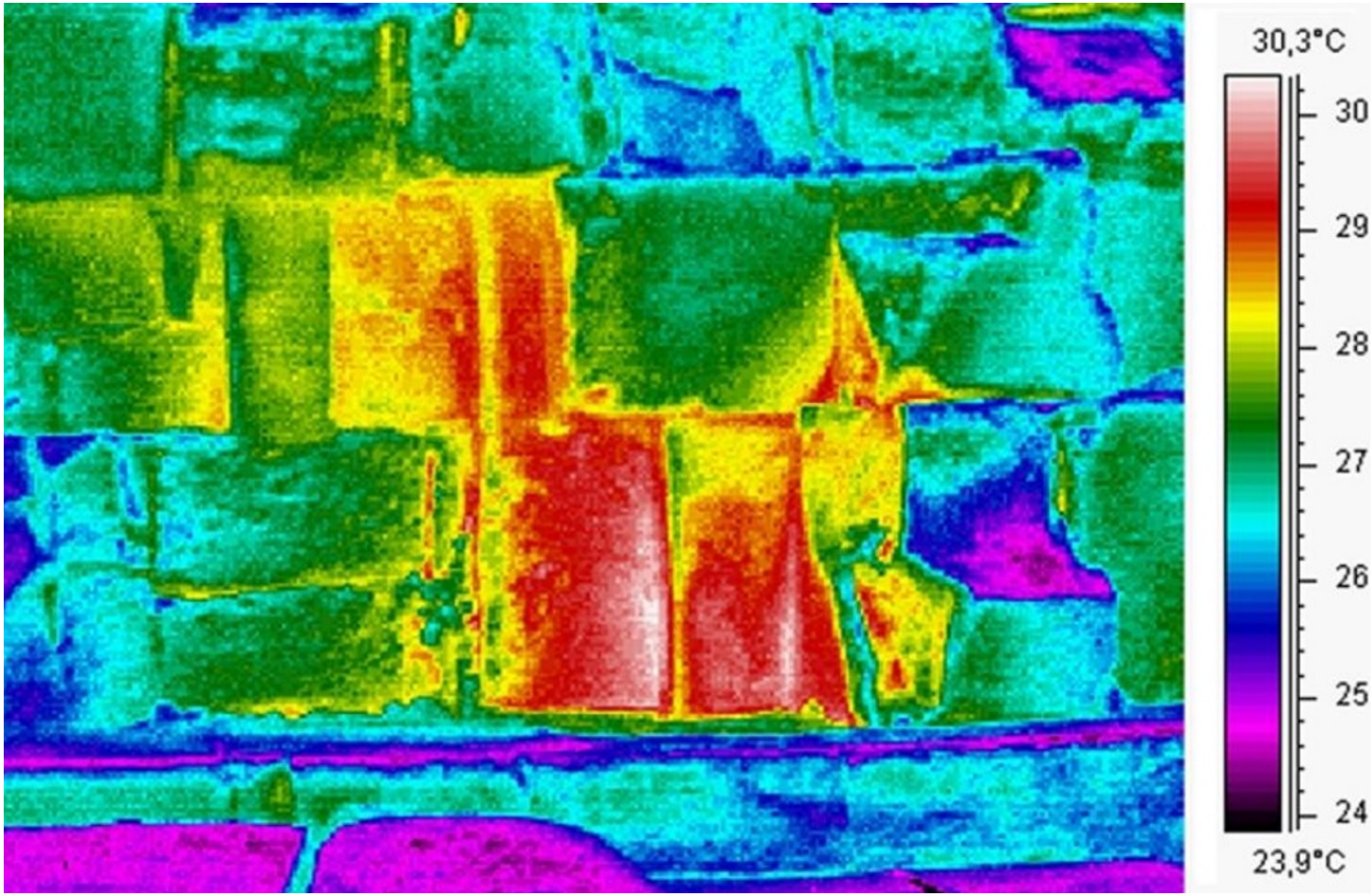
Example: Solid state imaging sensors

- Discrete silicon imaging elements, for example technology based on CCDs
- Voltage output proportional to incident light intensity



Example: Thermal imaging

- Thermal scans of the pyramid of Cheops were performed during the whole day, from sunrise to sunset



Digital Image Storage

- Stored in two parts
  - header
    - width, height … cookie.
      - Cookie is an indicator of what type of image file
  - data
    - uncompressed, compressed, ascii, binary.
- File types
  - JPEG, BMP, PPM, …

PPM, Portable Pixel Map

- Cookie
  - Px
    - Where x is:
    - 1 - (ascii) binary image (black & white, 0 & 1)
    - 2 - (ascii) grey-scale image (monochromatic)
    - 3 - (ascii) colour (RGB)
    - 4 - (binary) binary image
    - 5 - (binary) grey-scale image (monochromatic)
    - 6 - (binary) colour (RGB)

PPM example

- PPM colour file RGB

P3

feep.ppm

4 4

15

```
 0  0  0    0  0  0    0  0  0   15  0 15
 0  0  0    0 15  7    0  0  0    0  0  0
 0  0  0    0  0  0    0 15  7    0  0  0
15  0 15    0  0  0    0  0  0    0  0  0
```

RGB Images in Matlab

- An RGB color image is an *M x N x 3* array of color pixels
- Each color pixel is a triplet corresponding to the red, green and blue components of the RGB image
- Let `fR`, `fG`, and `fB` represent three RGB component images. To form an RGB image: `rgb_image = cat(3, fR, fG, fB)`
- To extract a component image: `fR = rgb_image(:, :, 1);`

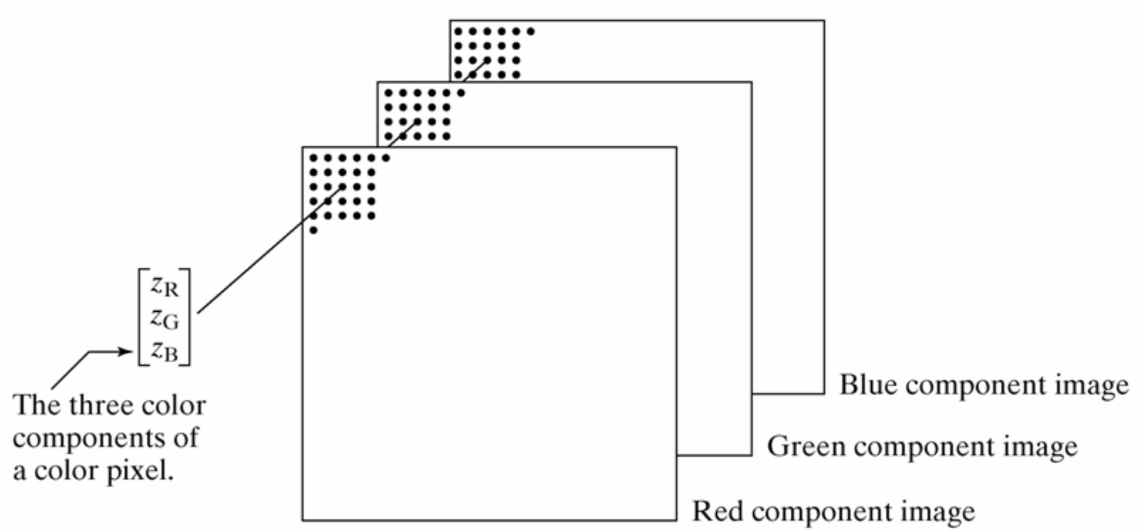


Image File Formats (1/3)

The American National Standards Institute (ANSI) sets standards for voluntary use in US. One of the most popular computer standards set by ANSI is the American Standard Code for Information Interchange (ASCII) which guarantees all computers can exchange text in ASCII format

BMP – Bitmap format from Microsoft uses Raster-based 1~24-bit colors (RGB) without compression or allows a run-length compression for 1~8-bit color depths

GIF – Graphics Interchange Format from CompuServe Inc. is Raster-based which uses 1~8-bit colors with resolutions up to 64,000*64,000 LZW (Lempel-Ziv-Welch, 1984) lossless compression with the compression ratio up to 2:1

Some Image File Formats (2/3)

- **Raw** – Raw image format uses a 8-bit unsigned character to store a pixel value of 0~255 for a Raster-scanned gray image without compression. An R by C raw image occupies R*C bytes or 8RC bits of storage space
- **TIFF** – Tagged Image File Format from Aldus and Microsoft was designed for importing image into desktop publishing programs and quickly became accepted by a variety of software developers as a standard. Its built-in flexibility is both a blessing and a curse, because it can be customized in a variety of ways to fit a programmer's needs. However, the flexibility of the format resulted in many versions of TIFF, some of which are so different that they are incompatible with each other
- **JPEG** – Joint Photographic Experts Group format is the most popular lossy method of compression, and the current standard whose file name ends with ".jpg" which allows Raster-based 8-bit grayscale or 24-bit color images with the compression ratio more than 16:1 and preserves the fidelity of the reconstructed image
- **EPS** – Encapsulated PostScript language format from Adulus Systems uses Metafile of 1~24-bit colors with compression
- JPEG 2000

Some Image File Formats (3/3)

- Common image formats include:
  - 1 sample per point (B&W or Grayscale)
  - 3 samples per point (Red, Green, and Blue) –
  - 4 samples per point (Red, Green, Blue, and "Alpha", a.k.a. Opacity)

- For most of this course we will focus on grey-scale images
  - Gray = 0.30 * R + 0.59 * G + 0.11 * B (one transfomraion)

Image resolution and quantization

Image resolution

- Image resolution quantifies how lines can be close to each other and still be visibly resolve
  ➔ Higher resolution means more image detail

- Since the image is represented by an $N \times M$ matrix
  - The more N and M is increased, the better the image resolution

800 × 531 pixels ≈0.4M pixels (the format is roughly 4 × 6)

http://www.flickr.com/photos/tammam-tillo/sets/

400 × 266 pixels

200 × 133 pixels

100 × 67 pixels

100 × 67 pixels enlarged (8 × zoom in)

800 × 531 pixels, original size

Measure of spatial resolution

- *SPI (samples per inch)*: used e.g. for scanners and image sampling devices

  (*scanned image resolution*)

- *PPI (pixels per inch):* the number of pixels displayed in an inch of image (*display resolution*)
  – IPhone 8 plus: 5.5 inch, resolution 1920*1080, 401ppi

- *DPI (dots per inch):* a measure of the resolution of a printer. It refers to the dots of ink used by a printing device (*printer resolution*)

dots-per-inch resolution

1250 dpi roughly books resolution   300 dpi (glossy brochures :175 dpi)

150 dpi (magazines : 133 dpi)   72 dpi (newspaper : 75 dpi)

Question?

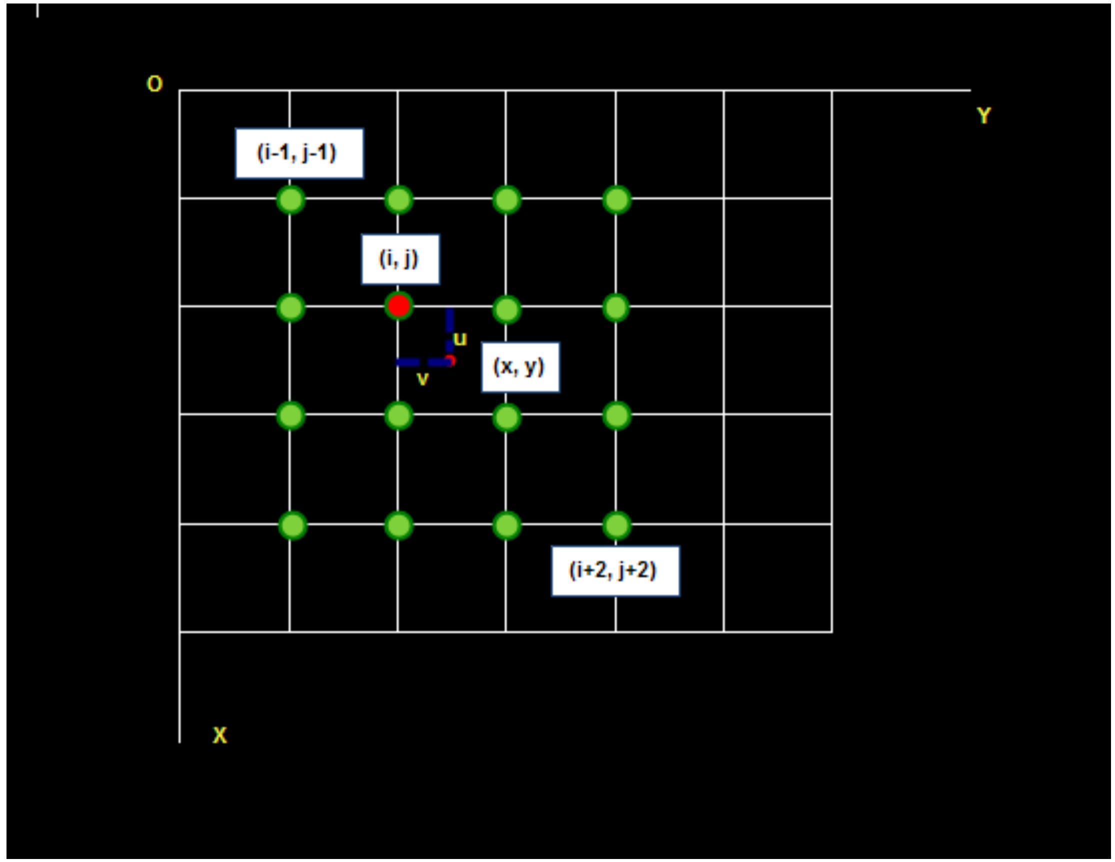
- How to get a High-resolution image from a low-resolution image?

Image Interpolation

1. Nearest Neighbor Interpolation
   (one point)

2. Bilinear interpolation
   (4 points)

   f(x,y)=f(i+u, j+v)
   = (1-u) * (1-v) * f(i, j)
   + (1-u) * v * f(i, j+1)
   + u * (1-v) * f(i+1, j)
   + u * v * f(i+1, j+1)



3. Bicubic interpolation
   (16 points)

Bilinear interpolation

- $f(i, j+v) = [f(i, j+1) - f(i, j)] * v + f(i, j)$
- $f(i+1, j+v) = [f(i+1, j+1) - f(i+1, j)] * v + f(i+1, j)$
- $f(i+u, j+v) = (1-u) * (1-v) * f(i, j) + (1-u) * v * f(i, j+1) + u * (1-v) * f(i+1, j) + u * v * f(i+1, j+1)$

Image Interpolation

- Bicubic interpolation

$f(x, y) = f(i+u, j+v) = ABC$

$$A = \begin{pmatrix} S(1+v) \\ S(v) \\ S(1-v) \\ S(2-v) \end{pmatrix}^T$$

$$B = \begin{pmatrix} f(i-1,\ j-1) & f(i-1,\ j) & f(i-1,\ j+1) & f(i-1,\ j+2) \\ f(i,\ j-1) & f(i,\ j) & f(i,\ j+1) & f(i,\ j+2) \\ f(i+1,\ j-1) & f(i+1,\ j) & f(i+1,\ j+1) & f(i+1,\ j+2) \\ f(i+2,\ j-1) & f(i+2,\ j) & f(i+2,\ j+1) & f(i+2,\ j+2) \end{pmatrix}$$

$$C = \begin{pmatrix} S(1+u) \\ S(u) \\ S(1-u) \\ S(2-u) \end{pmatrix}$$

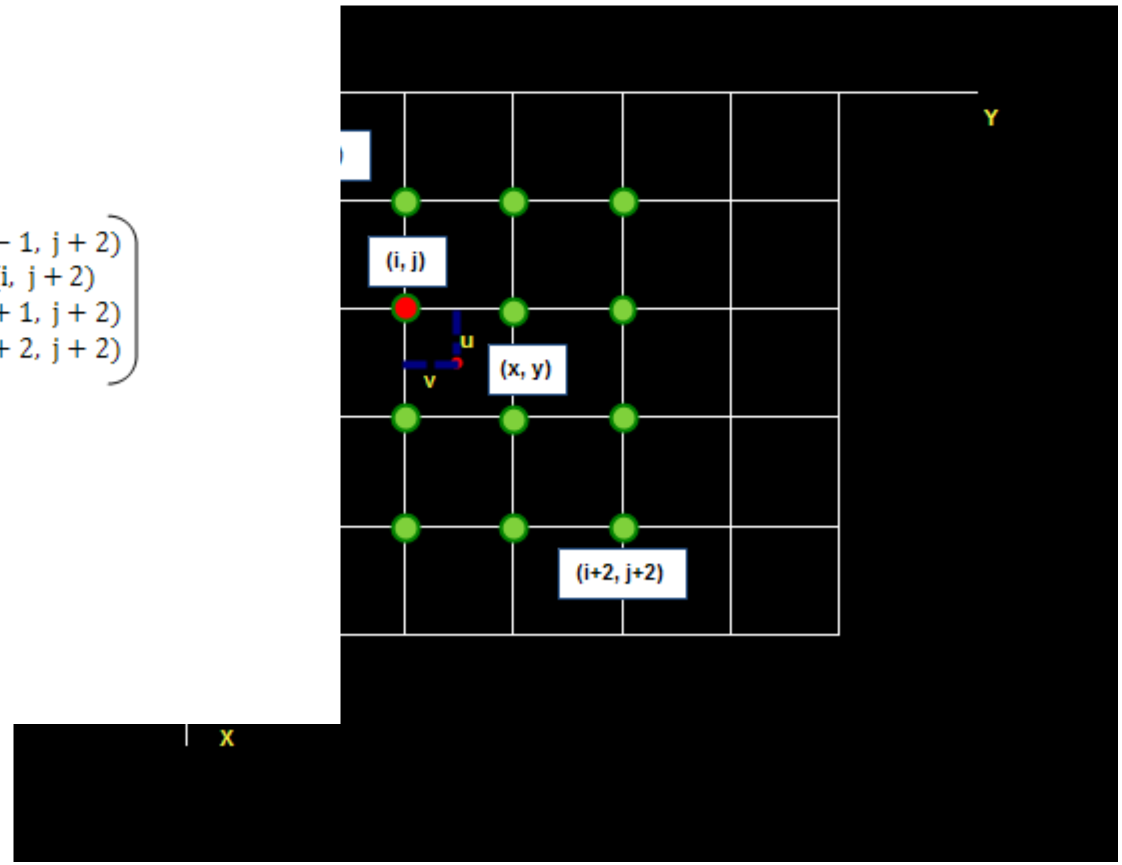
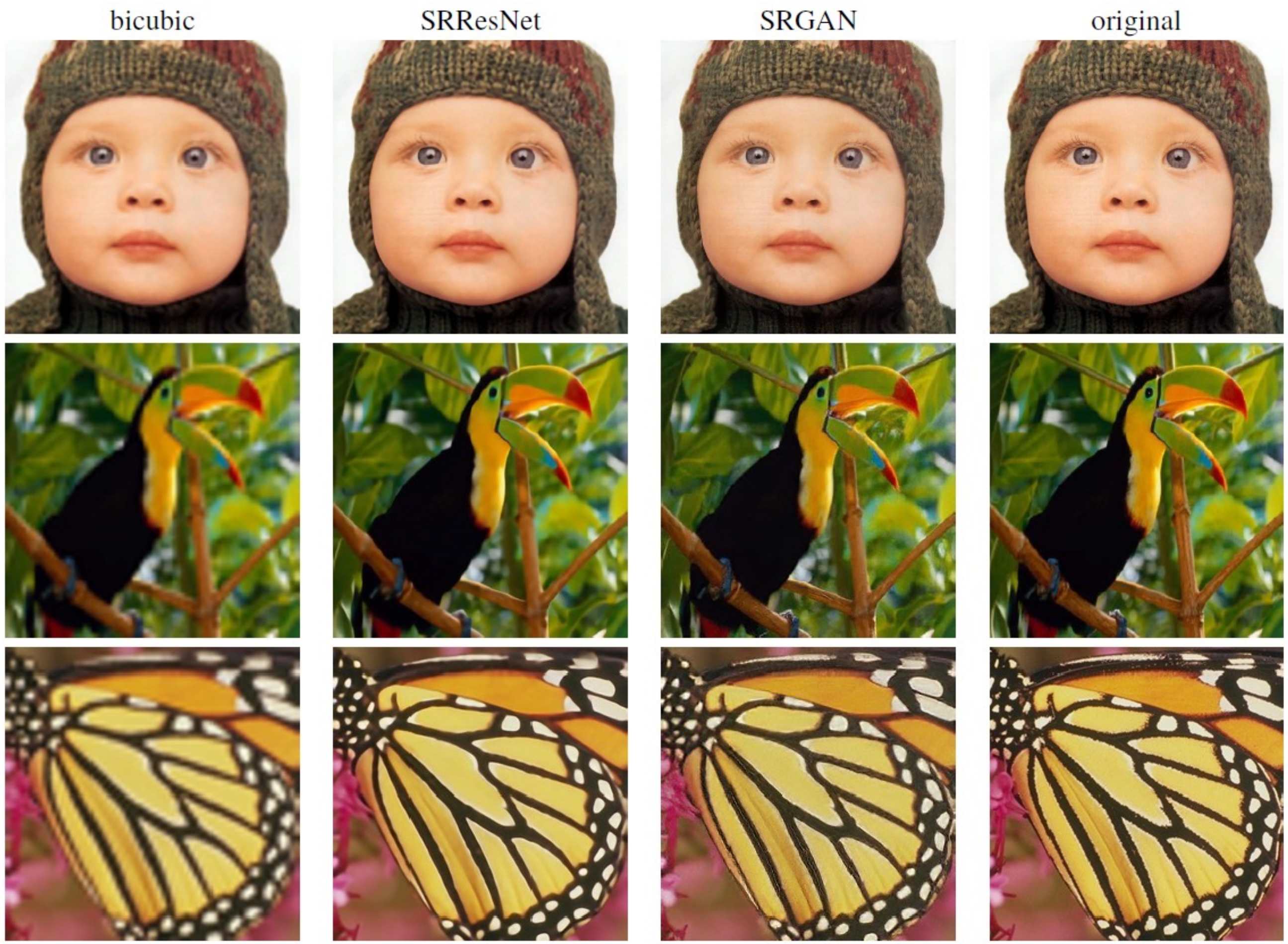
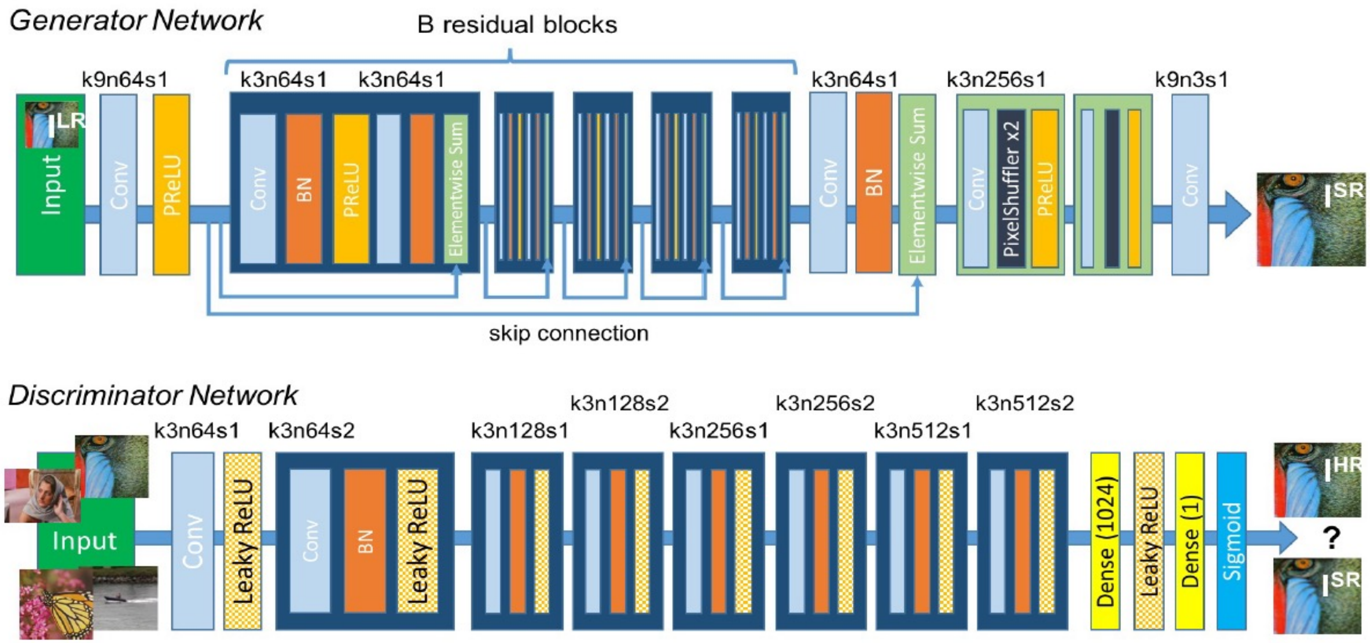


Image Super-Resolution

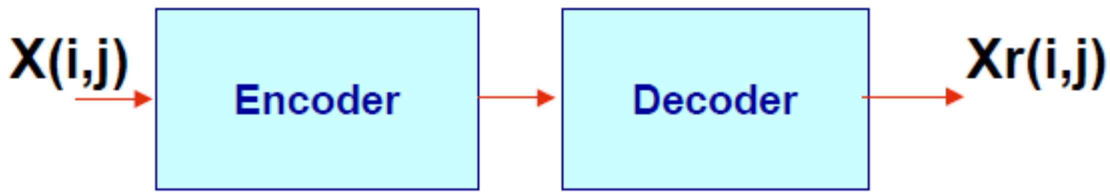


Framework of GAN



Quality metrics of images

- Peak signal-to-noise ratio (PSNR)



$$SNR = 10\log_{10} \frac{\frac{1}{NM}\sum_{i,j}\left|X(i,j)\right|^2}{\frac{1}{NM}\sum_{i,j}\left|X(i,j)-X_R(i,j)\right|^2} = 10\log_{10} \frac{\frac{1}{NM}\sum_{i,j}\left|X(i,j)\right|^2}{MSE}$$

$$PSNR = 10\log_{10} \frac{255^2}{\frac{1}{NM}\sum_{i,j}\left|X(i,j)-X_R(i,j)\right|^2}$$

code

Quality metrics of images

- Salt an paper noise (27.3dB)

code

Quality metrics of images

- AWGN noise (27.3dB)

Image quantization

Bit-depth of image

- The number of bits used to indicate the “color” of a single pixel, or the number of bits used for each color (8-bit-per-channel) component of a single pixel.
- If m denotes the number of bits necessary to store each monochrome “gray” pixel, then the total number of gray levels is

$$G = 2^m$$

- The more m is increased, the more levels of gray in the image could be represented

Bit-depth of image

G= 256

From left to right:
m = 8, 7, 4, 3, 2, 1

code

Number of storage bits for typical N and m values

x2

N \ m	1	2	3	4	5	6	7	8
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	392,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608

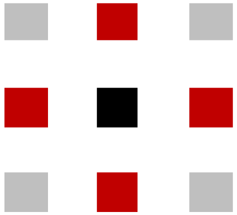
x4

- The storage requirements increase rapidly with N *(width=height)* (and with m*:bit*)

Basic relationships among pixels
- Neighbors-

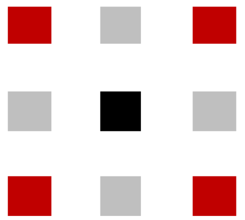
Basic relationships among pixels

- *Neighbors*: a pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are: $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$, these are called the *4-neighbors* of p and denoted as $N_4(p)$
- Some 4-neighbors can lie outside the image
  - Which ones?

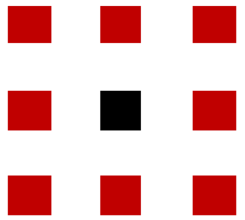


Basic relationships among pixels

- *Diagonal neighbors*: have coordinates $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$, and are denoted by $N_D(p)$; some of $N_D(p)$ can fall outside the image



- $N_D(p) \cup N_4(p) = N_8(p)$ (set of all neighbors)



Basic relationships among pixels -connectivity-

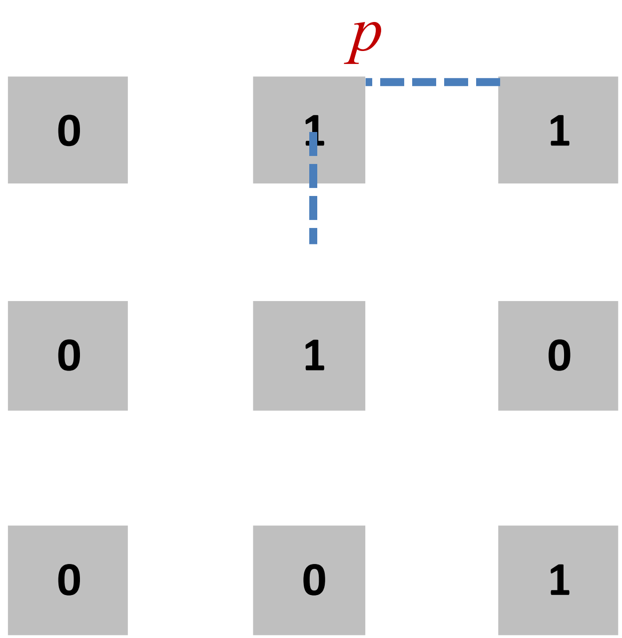
Connectivity

- Important concept to establish e.g. boundaries of objects
- Two pixels are connected if:
  - They are adjacent in some way (for example they are 4-neighbors)
  - Their gray levels are within a set of gray level values, V, used to define connectivity

4-connectivity

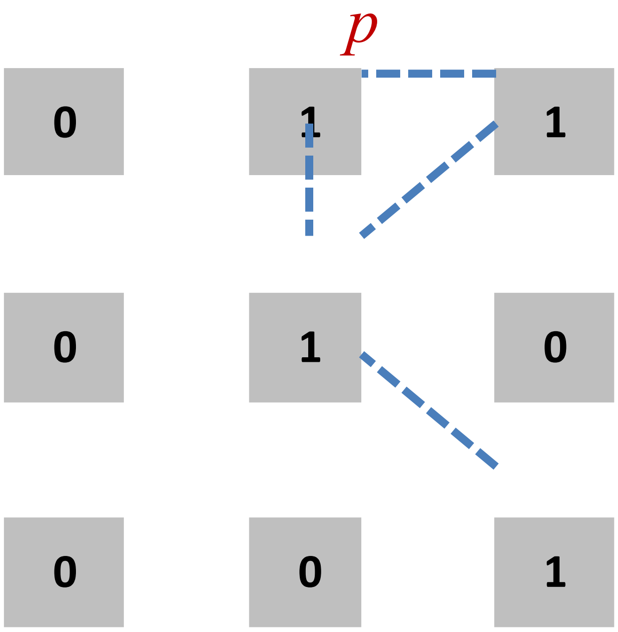
- If pixel $q \in N_4(p)$ and pixels p and q are with values in V.

Binary image, *V={1}*



8-connectivity

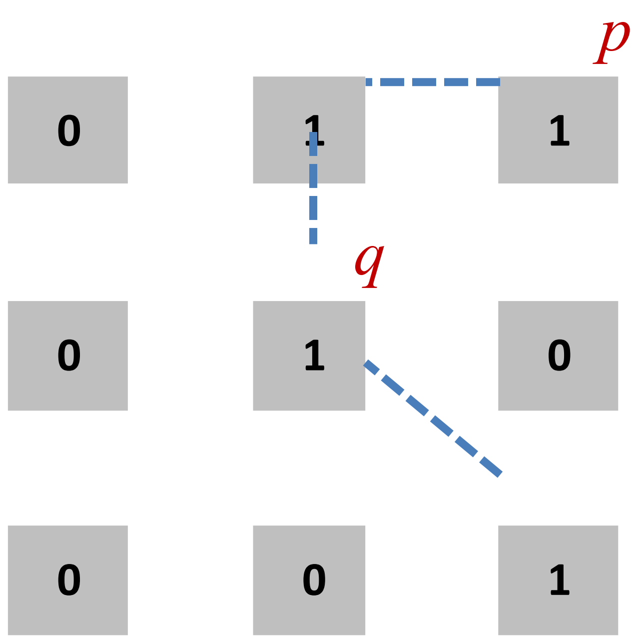
- If pixel $q \in N_8(p)$ and pixels p and q are with values in V.



Mixed (m) connectivity

- If two pixels p and q with values in V and if :
  - $q \in N_4(p)$ **or**
  - $q \in N_D(p)$ and $N_4(p) \cap N_4(q)$ is empty (i.e., they don't have any common N_4 neighbor)

Binary image, *V={1}*



Basic relationships among pixels
-distance measures-

Distance measures

- Pixels p and q with coordinates (x_1,y_1) and (x_2,y_2) respectively
- Euclidean distance:

$$D_e(p,q) = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$$

- Pixels with distance from p less than or equal to r are contained in a disk of radius r centered at p
- D_4 distance:

$$D_4(p,q) = | x_1 - x_2 | + | y_1 - y_2 |$$

- Pixels with D_4 distance from p less than or equal to r form a diamond centered at p

Distance measures

- *D_8 distance* or *chessboard distance* :

$$D_8(p,q) = \max(|x_1 - x_2|, |y_1 - y_2|)$$

- Pixels with D_8 distance from p less than or equal to r form a square centered at p

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

D4 ≤ 2

D8 ≤ 2

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Thanks