Multimedia Technology

Lecture 6: Fundamentals about Image Processing

Lecturer: *Dr*. Wan-Lei Zhao

Autumn Semester 2024

Email: wlzhao@xmu.edu.cn,

Outline

Opening Discussion (1)

- From now on, we are going to explore a new field
- Image and Videos
- It is another dimension of Multimedia
- In our daily life, around **80%** information comes from vision
- With the proliferation of digital devices, millions of images/videos are generated each day

Opening Discussion (2)

- It was guessed that vision caused **Cambrian Explosion** which took place in 542 million years ago
- Vision drove all creatures to evolve faster to survive
- With vision, they could search for food easier than before

Brief history about Digital Image (1)

- It was a long dream that one day we could keep what we see in somewhere besides our brain

Figure: Painting by Neanderthal who lived in Europe around 40,000 years ago.

Brief history about Digital Image (2)

Figure: Painting by ancient Egyptian who lived in 5,000 years ago.

- Notice that in these two periods, people could only try to draw what they saw

Brief history about Digital Image (3)

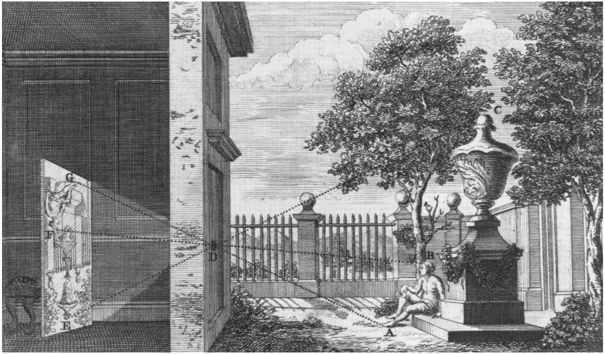


Figure: The known first camera by Aristotle.

- Why it works?
- How the size of aperture impacts the projected image
- How to keep this capture is still a big problem

Brief history about Digital Image (4)

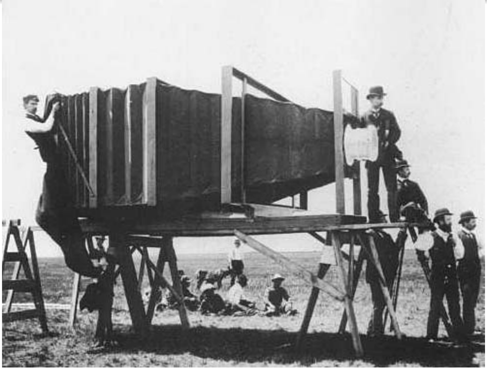


Figure: In 1900 the Chicago & Alton Railroad Train co. , commissioned Lawrence with the manufacture of the largest camera ever made and the largest photo ever shot in order to promote a new train.

- Around 30 years before that event, film was invented
- However, it requires long time of exposure

Basic knowledge about camera

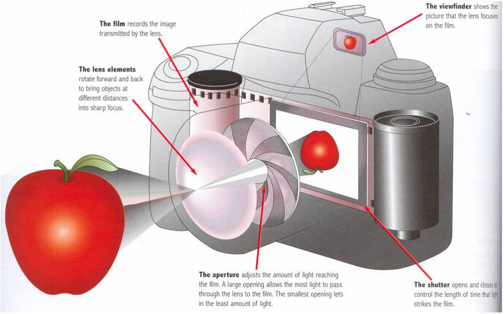


Figure: Structure of a modern camera

- Major components
  1. Lens
  2. Aperture
  3. Film/Sensor field

Basic knowledge about camera: the lens (1)

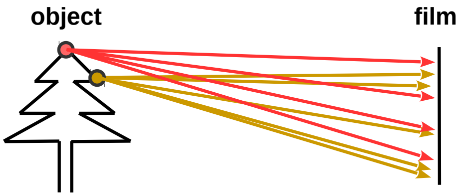


Figure: Try to put film in front of an object, see what you can get

- You get nothing but gray because ambient lights come from all directions

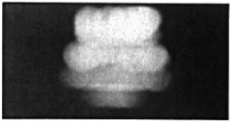
Basic knowledge about camera: the lens (2)

Figure: Remove camera lens

- Why blurry??

Figure: Camera imaging without lens.

Basic knowledge about camera: the lens (3)



(a) 2mm

(b) 1mm

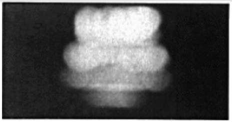
(c) 0.6mm

(d) 0.35mm

Figure: Imaging with different sizes of hole.

- Is it the smaller the better?

Basic knowledge about camera: the lens (4)



(a) 2mm

(b) 1mm

(c) 0.6mm

(d) 0.35mm

(e) 0.17mm

(f) 0.07mm

Basic knowledge about camera: the lens (5)

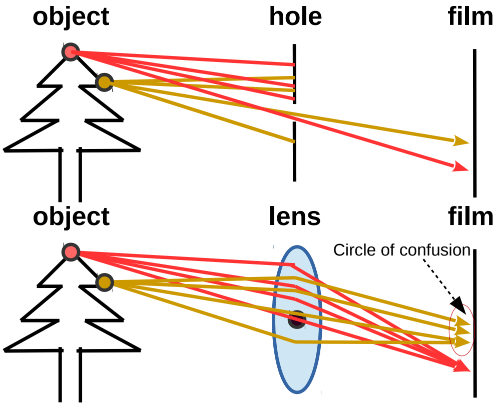
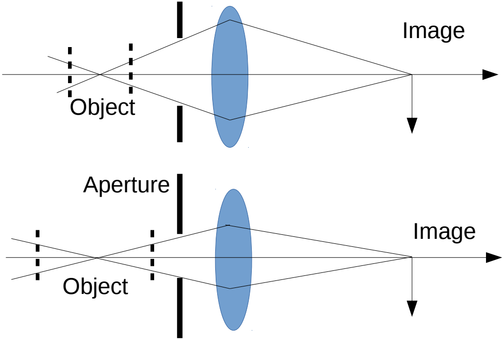


Figure: Function of the lens.

- Confusion is inevitable
- Overall it is beter to integrate lens

Basic knowledge about camera: the aperture (1)



- Large aperture covers relatively short "**depth of field**"
- Small aperture covers wide "**depth of field**" (DOF)
- However, small aperture allows less light pass-through
- It requires longer exposure time

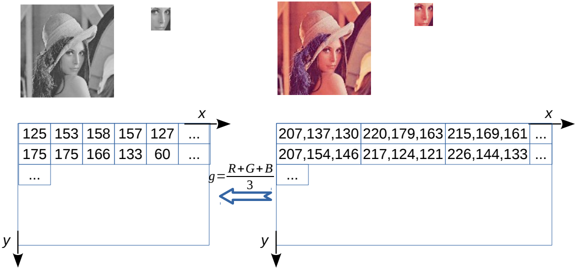
Basic knowledge about camera: the aperture (2)

(a) f=2.8, large aperture

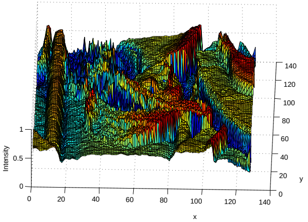
(b) f=22, small aperture

- Large aperture covers relatively short “**depth of field**”

Digital Image



(c) Format of a digital image



(d) Image displayed as function f(x,y) for the gray image

- Image is kept as a matrix, (0∼255)
- Pay attention to orientations for x and y coordinates
- One image is a function f(x,y) with two free variables x and y

Linear Transformation, Fourier Transform and Discrete FT

- Image is a matrix (one channel/gray level)
- We are therefore able to operate it as a matrix
  - Linear Transformations
  - Singular Value Decomposition
- Image is a multi-variable function
- We are therefore able to operate it as multi-variable function
  - Perform Fourier Transform
  - Taylor expansion
  - Search for extremal maximum and minimum values for the function

2D discrete Fourier Transform (1)

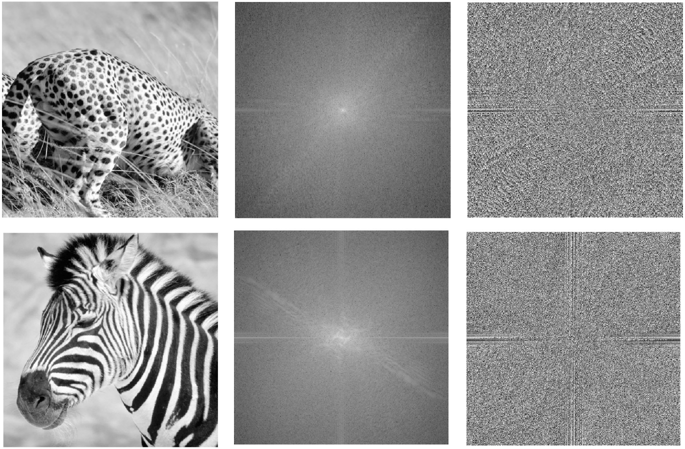
- Forward 2D discrete Fourier Transform

$$F(u,v)=\sum_{y=0}^{h-1}\sum_{x=0}^{w-1} f(x,y)e^{-2\pi i(\frac{uy}{h}+\frac{vx}{w})} \tag{1}$$

- Backward 2D discrete Fourier Transform

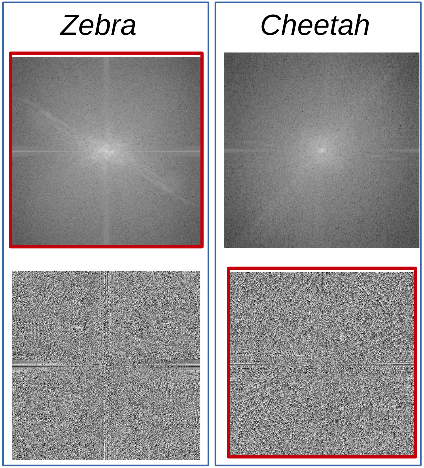
$$f(x,y)=\frac{1}{h\cdot w}\sum_{u=0}^{h-1}\sum_{v=0}^{w-1} F(u,v)e^{2\pi i(\frac{uy}{h}+\frac{vx}{w})} \tag{2}$$

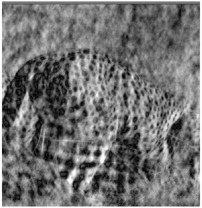
2D discrete Fourier Transform (2)



Input image | Fourier magnitude | Fourier phase

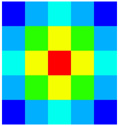
2D discrete Fourier Transform (3)



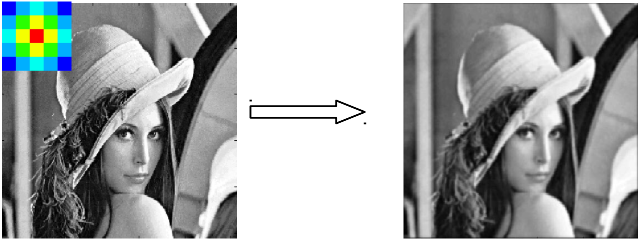


Phase from Cheetah
+
Magnitude from Zebra

Image Convolution



1	2	3	2	1
2	4	5	4	2
3	5	9	5	3
2	4	5	4	2
1	2	3	2	1



- Image Convolution achieves similar effects as 2D Fourier Transform
- One can define various types of convolution templates
- In practice, we do CORRELATION instead

Outline

About Basic Geometric Transformations on Image

- We already know how to process image as a multi-variable function
- We want to see how image is processed as a geometric matrix/map
- Basic linear transformations will be covered
  - Translation
  - Rotation
  - Scaling
  - Affine

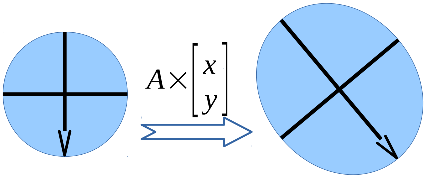
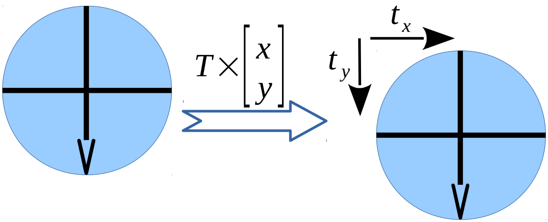


Image Translation

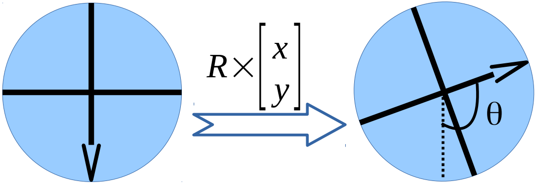
- Move image along x, y directions or both as a whole



$$[x' \quad y' \quad 1]^T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

Image Rotation

- Move image along x, y directions or both as a whole

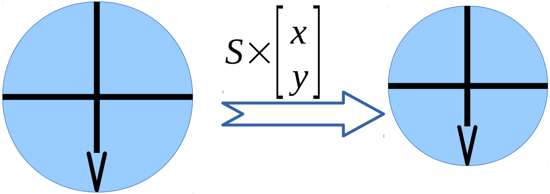


$$[x' \quad y' \quad 1]^T = \begin{bmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

- Notice that $R^{-1} = R^T$

Image Scaling

- To achieve zoom-in or zoom-out effect

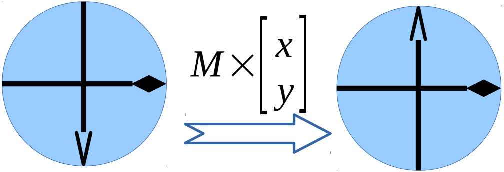


$$[x' \quad y' \quad 1]^T = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

- Notice that the scaling factors for x and y directions, s_x and s_y could be different

Image Reflection

- Also known as mirroring

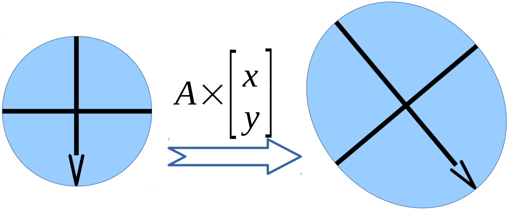


$$[x' \quad y' \quad 1]^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (6)$$

- You cannot achieve this by rotating
- If you mirror x and y both, it is then equivalent to rotating 180 degree

Image Affine Transformation

- There is another one called ‘shear’, check yourself what it is
- Basic transformations are **independent** from each other



$$[x' \quad y' \quad 1]^T = T{\cdot}S{\cdot}R{\cdot}M{\cdot}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{7}$$

- Affine is a combination of these basic transformations

Geometrical Invariance (1)

- If a vision system still recognizes what the object is after the object is geometrically transformed
- We say that this vision system is capable of geometrical invariance
- This is an IMPORTANT concept
- Our vision system is capable of geometrical invariance in various degree

Geometrical Invariance: rotation invariance (1)

- How much our vision achieves rotation invariance

Geometrical Invariance: rotation invariance (2)

- How much our vision achieves rotation invariance

- Verify your answer:)

Geometrical Invariance: scale invariance (1)

- What you can see from the image?

- grass, snow and ??

Geometrical Invariance: scale invariance (2)

- Get closer, what you can see from the image?

- grass, snow and ??

Geometrical Invariance: scale invariance (3)

- Get closer, what you can see from the image?

- grass, snow and sheep, clearly
- Conclusion: our vision is partially scale invariant

Geometrical Invariance: affine invariance (1)

- How much our vision achieves affine invariance

Geometrical Invariance: affine invariance (2)

- How much our vision achieves affine invariance




- Verify your answer:)

References

1. Digital Image Processing (Third Edition), Rafael C. Gonzalez and Richard E. Woods
2. Multiple View Geometry in Computer Vision, Richard Hartley and Andrew Zisserman
3. Computer Vision: Algorithms and Applications, Richard Szeliski

Q & A

Thanks for your attention!