

lec 1

Computer Vision:

- a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information.
- it often uses Image Processing techniques
- is the inverse of Computer Graphics

Related Fields:

• Image Processing:

- Image in, Image out
- low level techniques
- Quantitative measurements

• Computer vision:

- Extracting symbolic descriptions
- higher level techniques
- Semantic output
(quantitative or qualitative)

• Pattern recognition:

- classification

• Photogrammetry:

- Concerned with accurately measuring properties from images

Low-Level vision:

Photo manipulation

- Size → Color → Exposure → X-Pro II

Feature extraction

- Edges → oriented gradients → segments

Mid-level vision!

- Image \leftrightarrow Image
 - Panoramas
 - Image \leftrightarrow World
 - Multi-view stereo
 - Structure from motion
 - Structured light
 - LIDAR
 - Image \leftrightarrow Time
 - Optical flow
 - Time lapse
-

High-level vision!

- Semantics
 - Image classification
 - object detection
 - Segmentation

- Applications

- Retrieval
 - Robots
 - Super-resolution
 - Game playing
-

CV techniques

- Image processing
- Feature extraction + ML
- Deep learning

chp 2

eyespot: a photoreceptive protein (the simplest type of eyes)

→ photosensitive proteins with no other surrounding structure.

→ no nerves, brain or processing

→ Very low acuity

pit eyes: photosensitive cells in pits

→ Block some light

→ Very common

→ more information about where light direction

→ Very simple

→ low acuity

Focusing

viewing a

1. distant object

→ nearly parallel light rays

→ Flattened lens for weak reflection

2. near object

→ Diverging light rays

→ Round lens for strong

reflection

للأجسام البعيدة تكون العين مسطحة

للأجسام القريبة تكون العين دائرية

Complex eyes:

- many different styles and mechanisms
- better resolution
- Rare 6 of 33 animal phyla
- Image forming - high enough acuity to perceive shapes and objects

Photoreceptive Cells (~120 million in retina)

Rods:

- 120 million
- Sensitive to 1 photon
- can pool response
- Slow response time
- only operate in low-light conditions
- Saturate quickly in lots of light

Cones:

- 6 million
- need many photons to activate
- Fast response time
- Fine details
- Fast Change over time
- responsible for most daytime vision
- Mostly packed into one region: Fovea

Fovea:

- Small circle on the retina 1.5 mm
- Contains many Cones (200,000)
- Highest visual acuity
- Reading: text is centered in fovea

peripheral vision area:

- Few cones * low acuity * low perception of color
- lots of rods * good at night

Fixational eye movement

- Receptors adjust, lose sensitivity over time
- Eye keeps moving to expose new parts to light

Microsaccades

- short linear movement

Ocular drift

- Constant slow movement

Microtremors

- tiny vibrations
- Synchronized between eyes

Ganglia transmit info to brain

- 1 million of them
- Connect to different kinds of photoreceptors
- sensitive to different things

M Cells

- depth
- movement
- orientation

P Cells

- color
- shape
- details

Incident Spectral ظروف الإضاءة	\times	Percent Reflectance of skin surface لونه الجلد	$=$	Reflected spectral لونه الجلد انعكاس الضوء
---	----------	--	-----	---

Photoreceptors and light wavelengths:

Rods: Peak around 498nm

Cones:	Short: 4	420 nm	blue	} → كودهم صيفر في ال retina
	Medium: 4	530 nm	green	
	Long: 4	560 nm	red	

RGB is a cube

Hue: what color

Saturation: how much color

Value: how bright

Chp3 - Image Basics

Storage:

Row major:

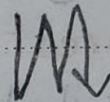
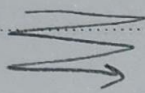
→ HW

→ Rows first then columns

Column major:

→ WH

→ Columns first then rows



HWC:

→ Channels interleaved

CHW:

→ Channels separated

Color spaces:

RGB:

→ Cube

→ Red - Green - blue

HSV:

→ Cylinder

→ Hue - Saturation - Value

Convert RGB → HSV: (3d tensors)

→ tilt cube

→ Squash colors

→ expand base

→ Smooth cylinder

Note: to enhance an image duplicate saturation

→ Value manipulates → exposure

Image resizing:

Nearest neighbor

- What it (Pixel) sounds like.
- Image looks blocky.
- Z (Channel) is still int.

Common pitfall: integer division rounds down in C

- For structured image

Triangle interpolation

- For less structured image

~~is pretty good for grids~~

Bilinear interpolation:

- Pretty good for grids
- Find the closest pixels in a box
- ~~is~~ Smoother than NN
- more complex

Image Resize Algorithm:

For each pixel in new image?

- Map to old image coordinates
- Interpolate value
- Set new value in image

Chp 4

Interpolation: the insertion of an intermediate value into a series by estimating it from surrounding known values.

Resolution: refers to the number of pixels in an image.

Print Resolution: refers to the number of dots per inch (dpi).

Steps for ~~know~~ enlarging an image

- Map the original and desired coordinates
- Find all the original coordinates for ~~each~~ each new coordinate

Use an interpolation method to find each of the new pixel values.