Contents

1 Presentation

1.1 Logistics:

Tuesday 8:15 - 9:30 Friday 9:45 - 11:00

1.2 Grading

- Homework 60%

- Final exam 20%
- Midterm exam 20%

1.3 Work

- Regular quizzes
- Homework/ project -> pre-requisite to take the final exam
  - won't accept any homework if it is late.

1.4 Health excuses

Not accepting health excuses.

2 Introduction to Computer Vision

The goal of computer vision to bridging the gap between pixels and meaning.

2.1 what is vision?

- humans: -> images -> sensing devices -> iterpreting device -> interpretations
- computers: -> images -> cameras -> computers -> interpretation

2.2 what information to extract?

- Metric 3D information
- Semantics

2.3 what is color?

- the result of ityeraction between physical light in the environment and our visual system.
- A psychological property of our visual experiences when we look at objects and lights, not a physical property of those objects or lights.

2.4 two types of light sensitive receptors

- cones: cone-shaped, less sensitive, operate in high light, color vision
- Rods: rod-shaped, highly snesitive, operate at night, gray-scala vision

2.5 trichromacy

- three numbers seem to be sufficient for encoding color
- dates back to 18th centuary.
- don't have to be always RGB

2.6 RGB v. HSV

- RGB: component
- HSV: hue, saturation, value

2.7 white balance:

- it is the process of removing unrealistic color casts, so that objects which appear white in person are rendered white in you photo.
- whne the white balance is not correct, the picture will have n unnatural color "cast".

2.7.1 Film cameras:

different types of film or different filteres for different illumination condition

2.7.2 digital cmeras

- automatic white balance
- custom white balancing using a reference objects

2.7.3 Von Kries adaptation

- multiply each channel by a gain factor
- A more general approach would correspond to an arbitrary 3x3 matrix

2.7.4 Best way: gray card

- take a pciture of a neutral object (gray/white)
- deduct the weight of each channel, use the inverse of the weight to calibrate each channel.
- without gray card, we need to guess
    - using the average of the image, assume it is gray

2.7.5 Brightest pixel assumption (non-saturated)

- highlights usually have the color of the light source
- Use weights inversely proportional to the values of the pixel

2.7.6 Gamut mapping

3 Scale-invariant

4 Detectors

4.1 Harris detector

4.1.1 mathematics

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]$$

This equation is computational expensive, so we can use the first order Taylor

$$\begin{aligned}
&\sum_{x,y}[I(x+u,y+v) - I(x,y)] \\
&= \sum_{x,y}[I(x,v) + uI_x + vI_y - I(x,y)] \\
&= u^2 I_x^2 2uvI_xI_y + v^2 I_y^2 \\
&= [uv](\sum \begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix})[uv]'
\end{aligned}$$

Thus, we can have a bilinear approximation

$$M = (\sum w(x,y) \begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix})$$

and

$$E(u,v) \approx [u,v]M[u,v]'$$

4.1.2 cornerness

We look at the two eigen value of the M matrix. With two big eigenvalues, it means it is a corner. If one is significantly bigger than the other, it is a edge. If both small, it is plain images. We write

$$R = det(M) - k \cdot tr(M)^2$$

, where $det(M) = \lambda_1\lambda_2$, $tr(M) = \lambda_1 + \lambda_2$

4.2