

写在前面

大概说了一下 CV 与哪些学科相关

CV 实现思路可以参考人类及动物的视觉系统

CV 在现代的一些应用 (人脸识别, 血膜识别, 指纹, OCR)

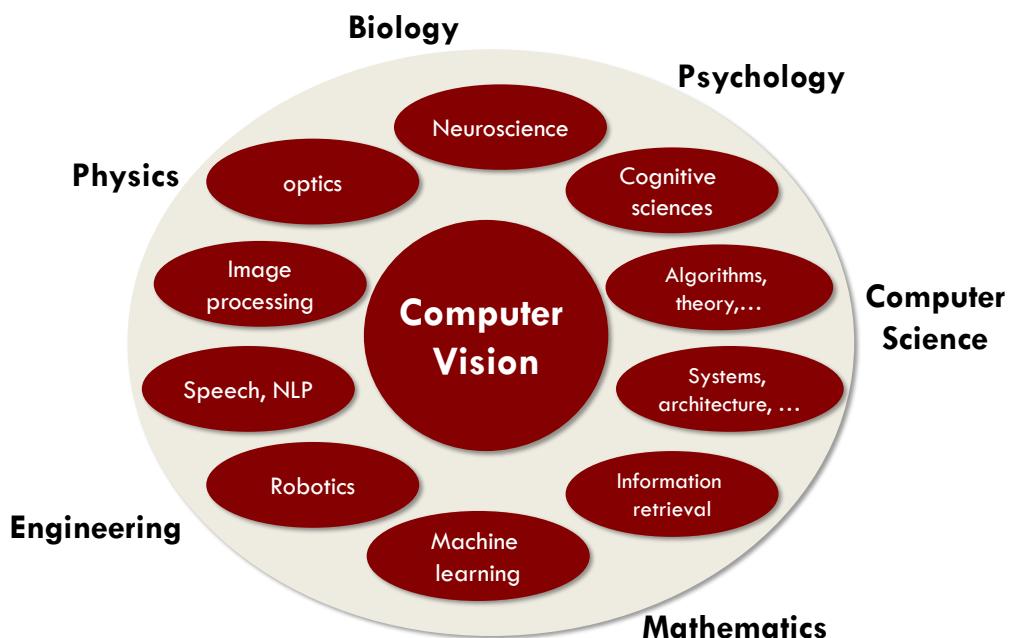
Machine Vision Technology

School of Computer Science
Beijing University of Posts and Telecommunications

1



2



3

Today's agenda

- Introduction to computer vision
- Course overview

4

Quiz?



5

What about this?

3D Reconstruction



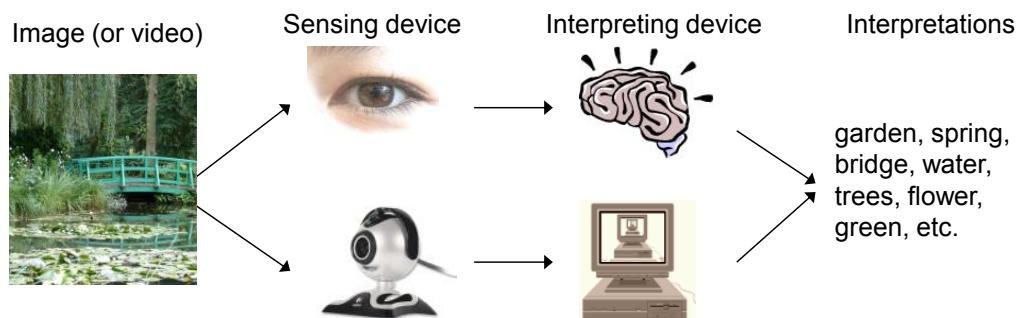
Michaelangelo's David

Recognizing chairs



6

What is (computer) vision?



7

The goal of computer vision

- To bridge the gap between pixels and “meaning”



What we see

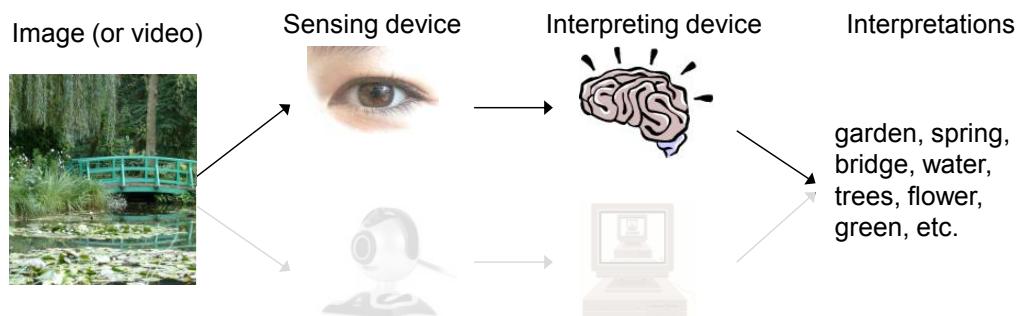
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

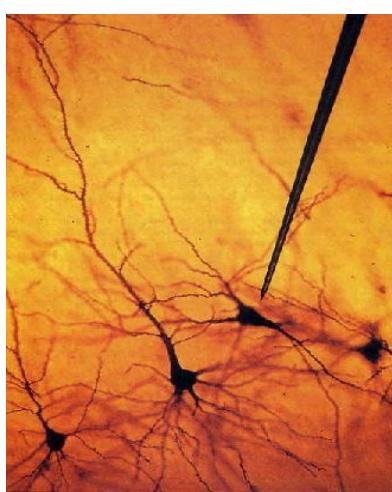
8

What is (computer) vision?



9

1981: Nobel Prize in medicine



Hubel & Wiesel

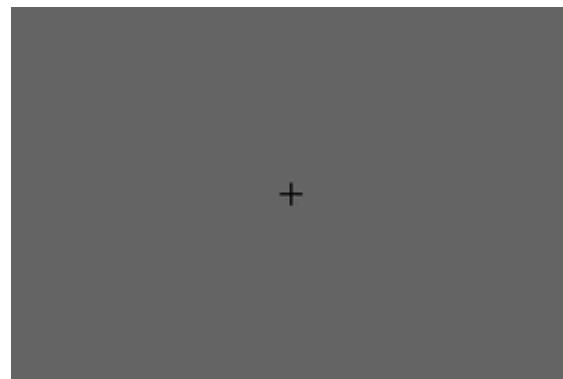
Dr. Hubel said:

There has been a myth that the brain cannot understand itself. It is compared to a man trying to lift himself by his own bootstraps.

We feel that is nonsense. The brain can be studied just as the kidney can.

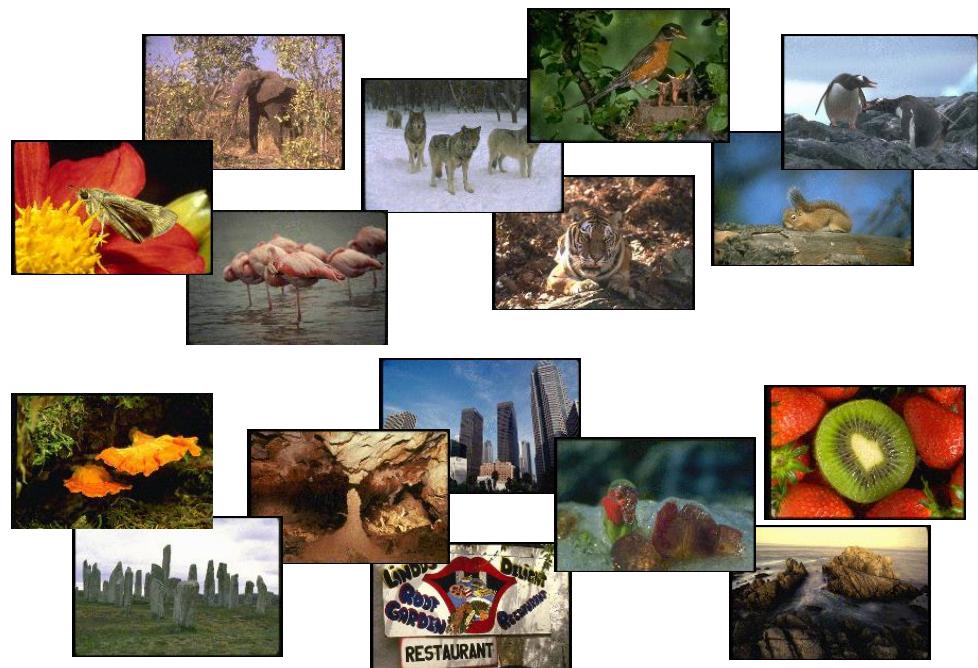
10

Human vision is superbly efficient



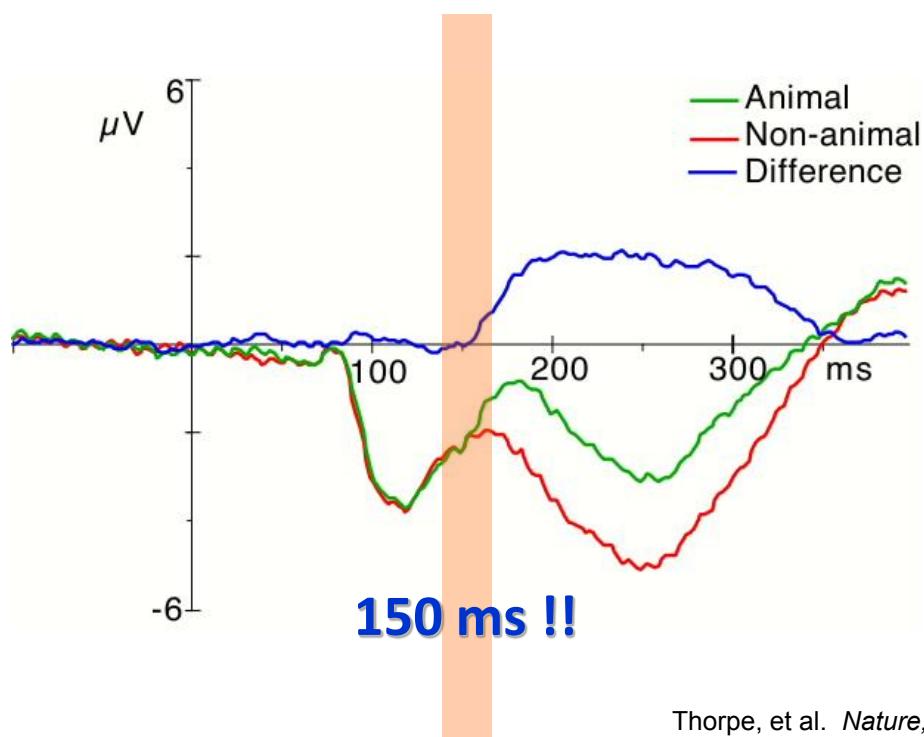
Potter, Biederman, etc. 1970s

11



Thorpe, et al. *Nature*, 1996

12



Thorpe, et al. *Nature*, 1996

13

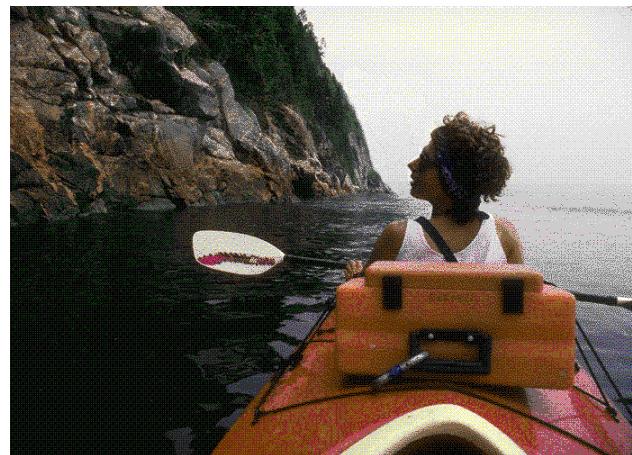
Change Blindness



Rensink, O'regan, Simon, etc.

14

Change Blindness



Rensink, O'regan, Simon, etc.

15

Segmentation

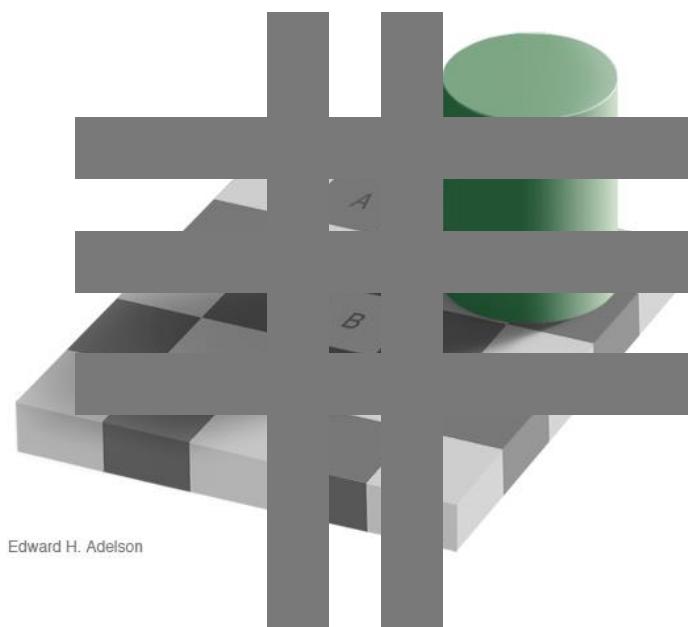


16

Perception

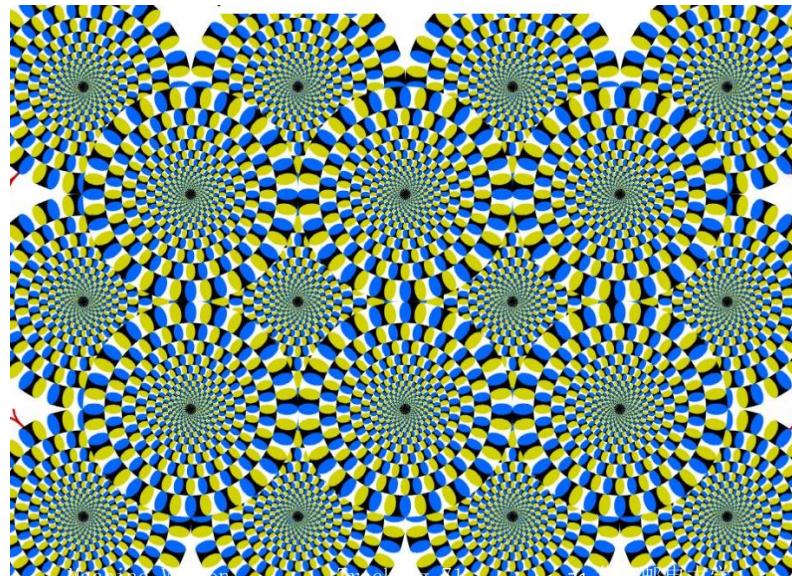


17

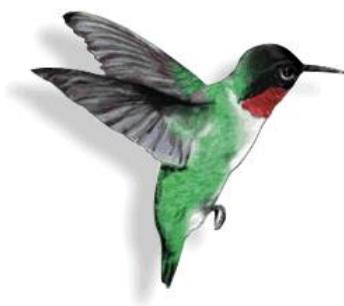


18

Motion without movement

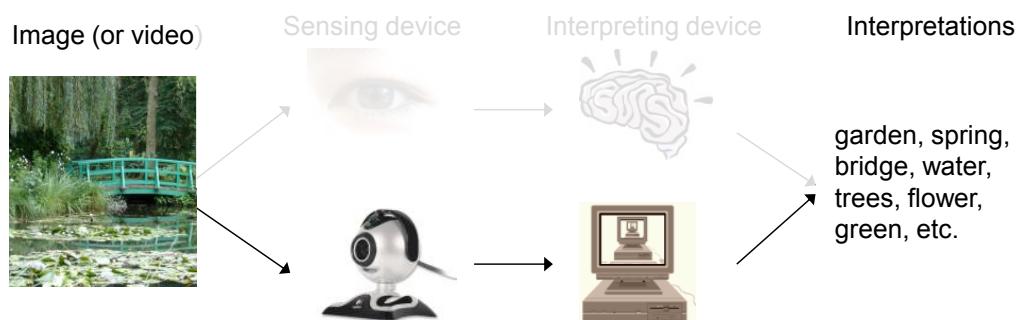


19



20

What is (computer) vision?



21

The goal of computer vision

- To bridge the gap between pixels and “meaning”



What we see

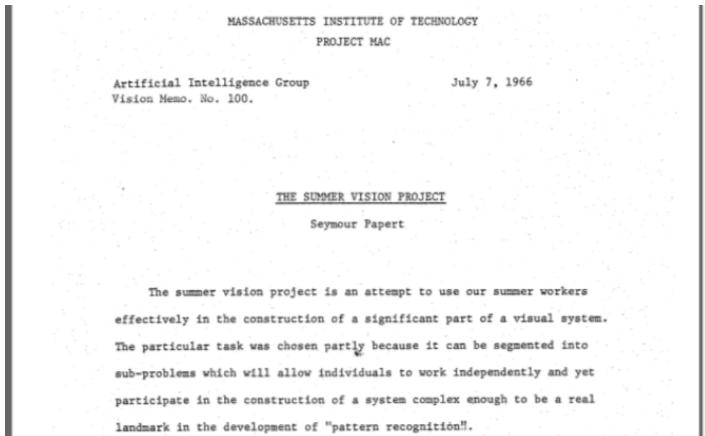
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

22

Origins of computer vision: an MIT undergraduate summer project



23

Pioneer in the field : David Marr (1945-1980)

Computational theory: What is the goal of the computation (task) and what are the constraints that are known or can be brought to bear on the problem?

Representations and algorithms: How are the input, output, and intermediate information represented, and which algorithms are used to calculate the desired result?

Hardware implementation: How are the representations and algorithms mapped onto actual hardware, e.g., a biological vision system or a specialized piece of silicon? Conversely, how can hardware constraints be used to guide the choice of representation and algorithm?

『Vision: A Computational Investigation into the Human Representation and Processing of Visual Information』

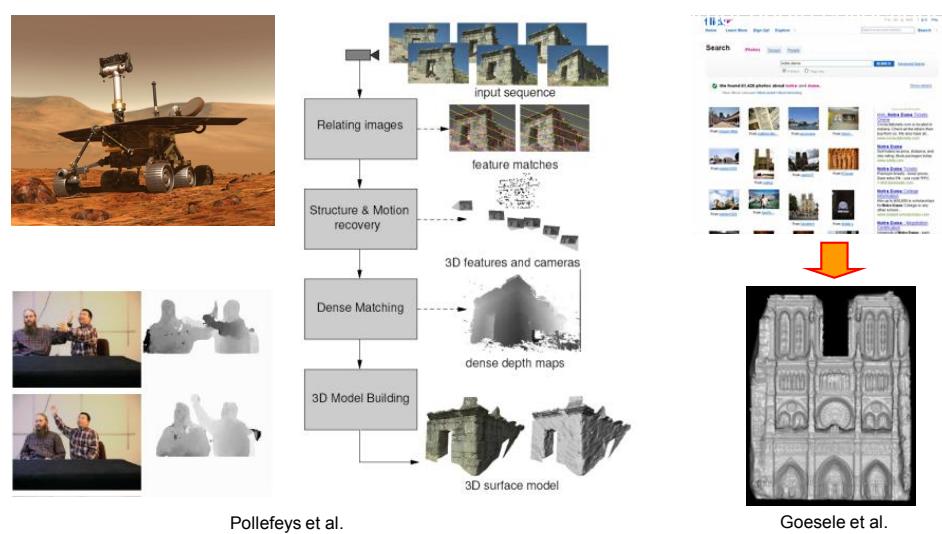
24

What kind of information can we extract from an image?

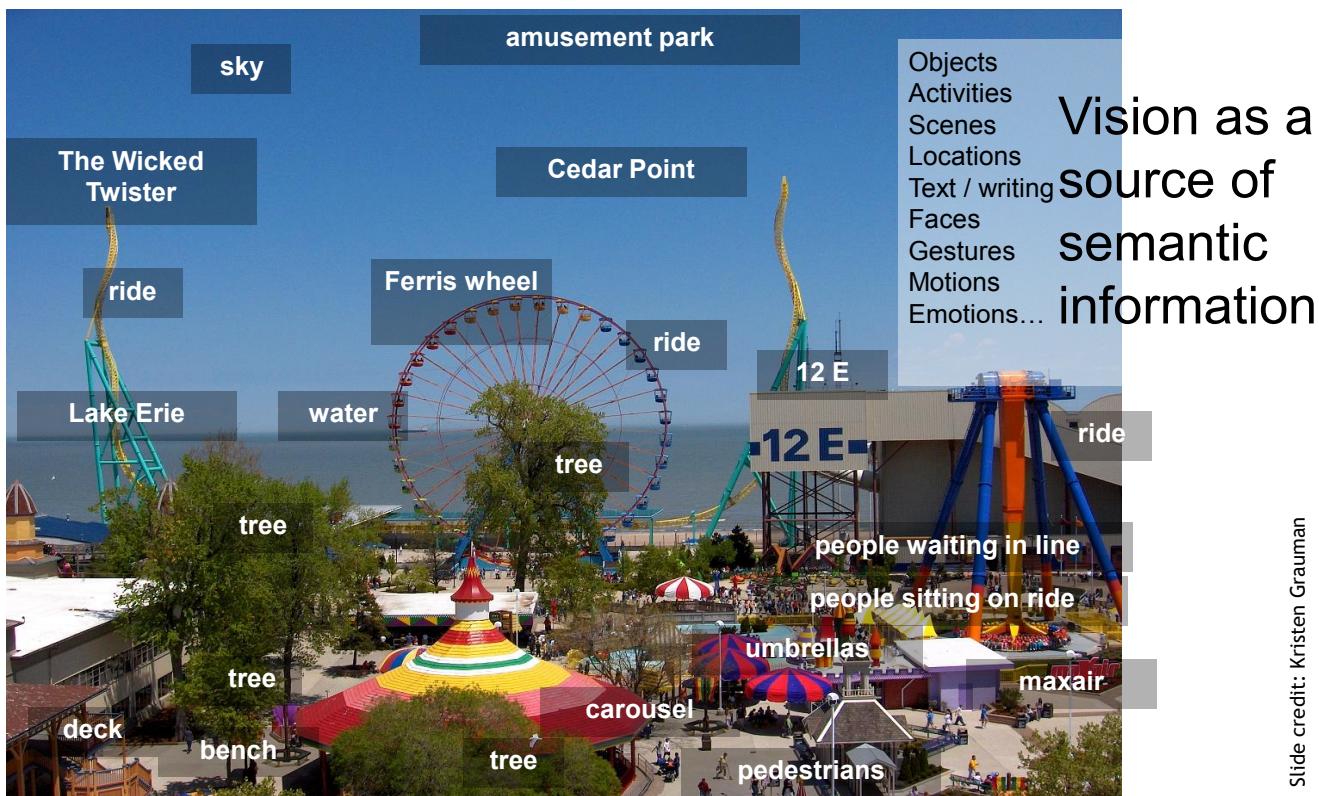
- Metric 3D information
- Semantic information

25

Vision as measurement device



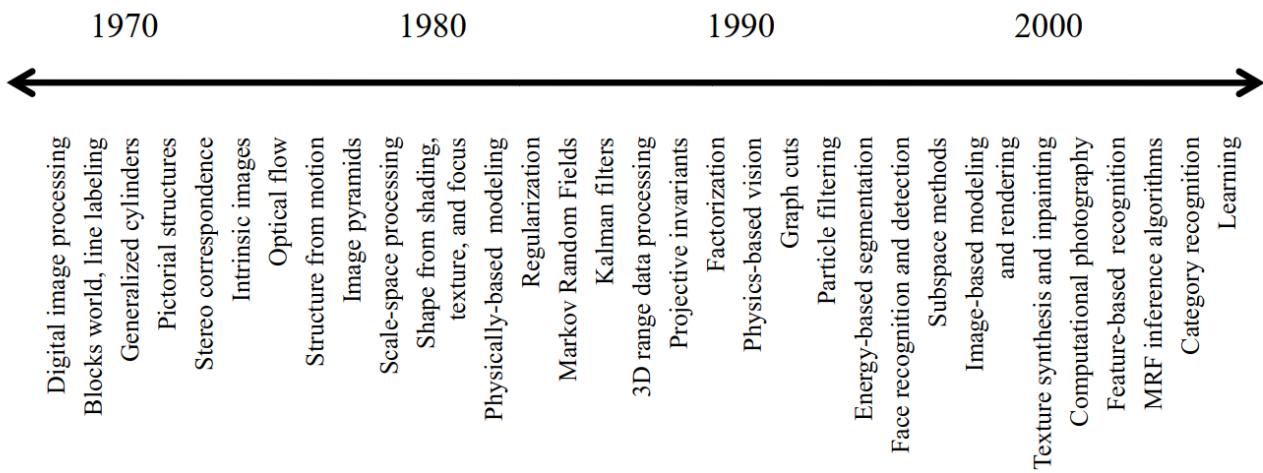
26



27

Slide credit: Kristen Grauman

Most active topics of research in computer vision



28

Why study computer vision?

- Vision is useful: Images and video are everywhere!

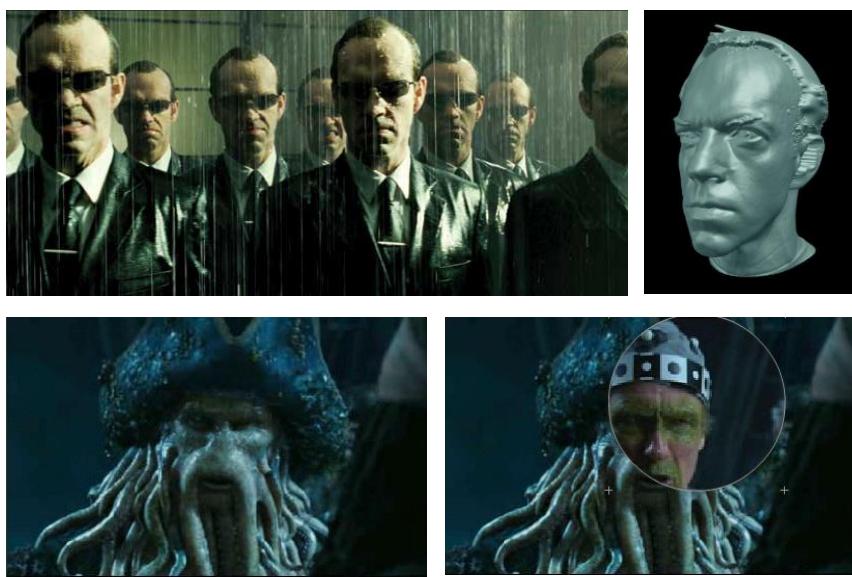


Surveillance and security

Medical and scientific images

29

Special effects: shape and motion capture



Source: S. Seitz

30

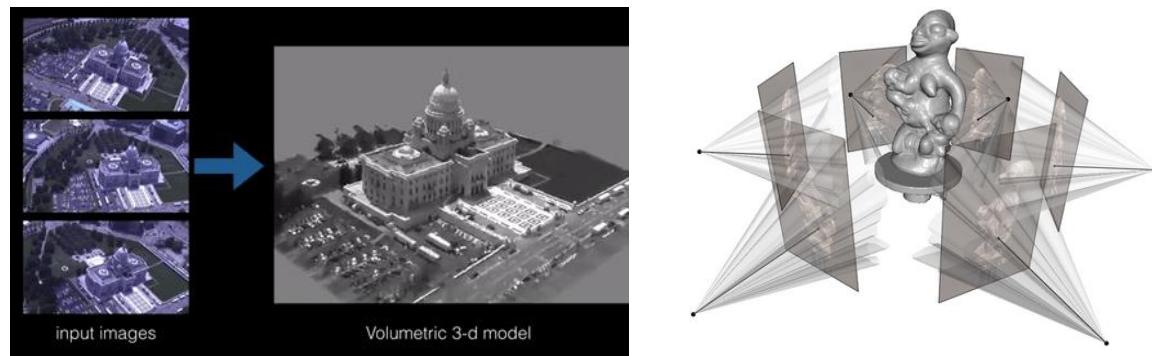
3D urban modeling



[Bing maps](#), Google Streetview

31

3D modeling



32

Face detection



- Many digital cameras now detect faces
 - Canon, Sony, Fuji, ...

33

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



[Sony Cyber-shot® T70 Digital Still Camera](#)

34

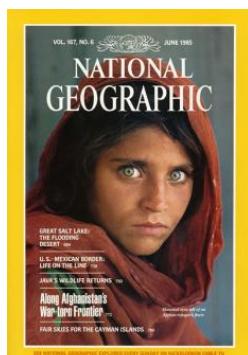
Face recognition: Apple iPhoto software



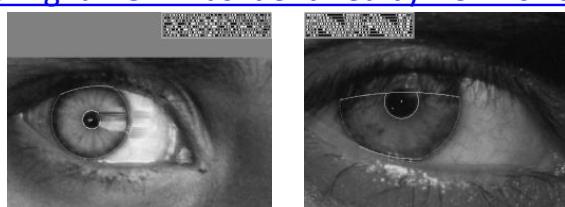
<http://www.apple.com/ilife/iphoto/>

35

Biometrics



[How the Afghan Girl was Identified by Her Iris Patterns](#)



Source: S. Seitz

36

Biometrics



Fingerprint scanners on
many new laptops,
other devices



Face recognition systems now beginning
to appear more widely
iphone X just introduced face recognition

37

Optical character recognition (OCR)

Technology to convert scanned docs to text

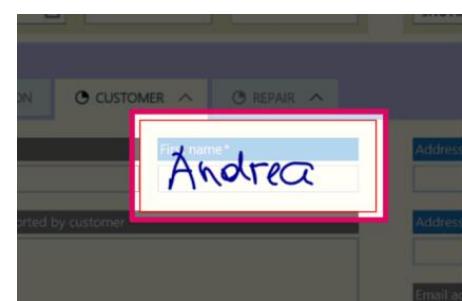
- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs



License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition



Source: S. Seitz

38

19

Google maps: Annotate all houses and streets



Avenue des Sapins

Goodfellow et al. 2014

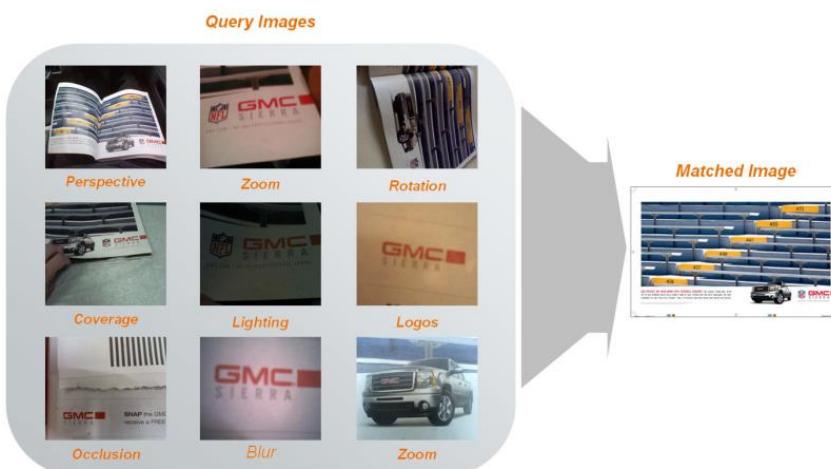
39

Toys and Robots



40

Mobile visual search: iPhone Apps



41

syte

Solutions Marketplace Pricing Resources Case Studies Company Enter Business Email See Syte in

The Leader in Visual AI for Retail

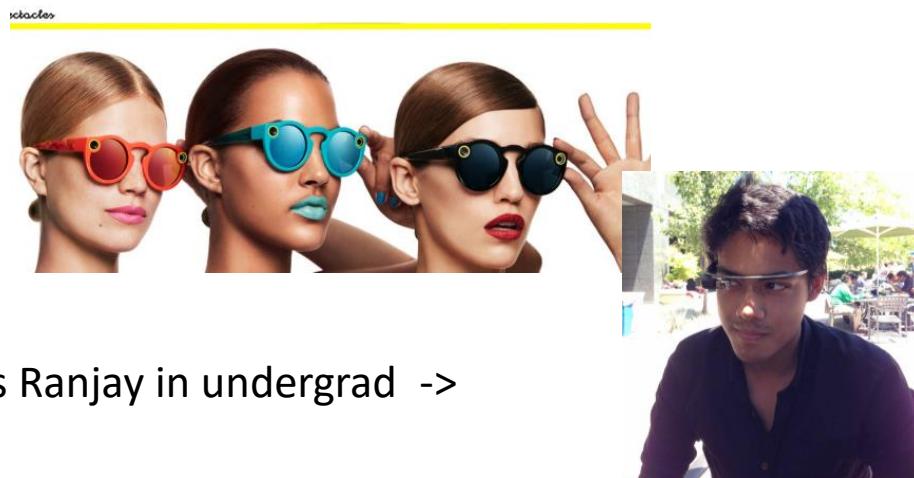
Syte changes the way retailers connect shoppers with the products that inspire them by delivering the best Visual AI technology for retail. Discover our solutions that empower retailers to increase customer engagement, and boost conversion and sales.

Enter Business Email See Syte in Action! Watch Video

TOMMY HILFINGER FARFETCH bon PRIX SHOPSTYLE

42

Snapstacles and Google glasses



- That's Ranjay in undergrad ->

43

Automotive safety

News

- > [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- > [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)
- > [all news](#)

Events

- > [Mobileye at Equip Auto, Paris, France](#)
- > [Mobileye at SEMA, Las Vegas, NV](#)
- > [read more](#)

- [Mobileye](#)

44

Detection and tracking



45

Vision in supermarkets



LaneHawk by EvolutionRobotics

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it...”

46

Amazon Go



47

Vision-based interaction (and games)



Microsoft's Kinect



Sony EyeToy



Assistive technologies

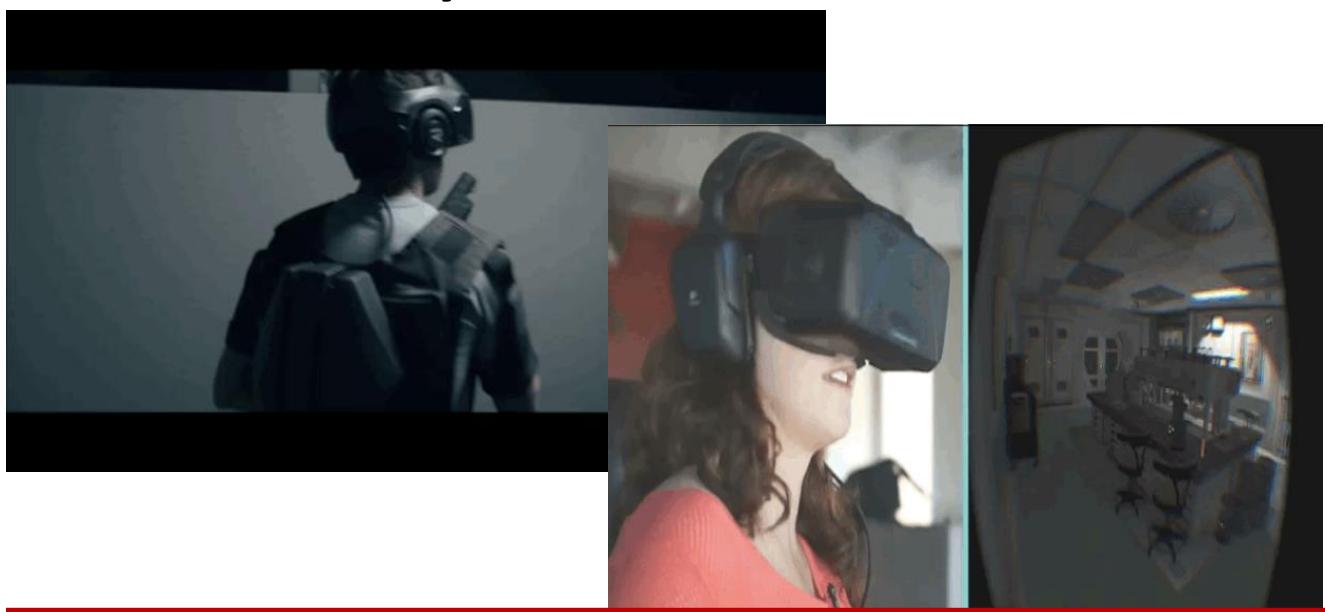
48

Augmented Reality



49

Virtual Reality



50

Vision for robotics, space exploration



[NASA's Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

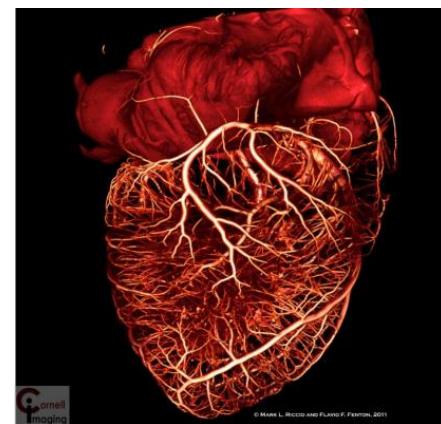
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "[Computer Vision on Mars](#)" by Matthies et al.

ce: S. Seitz

51

Medical image recognition



52

Vision for meteorology



53

Course Overview

54

Syllabus organized from

- **CS131 Computer Vision: Foundations and Applications**

http://vision.stanford.edu/teaching/cs131_fall1920/index.html

- **CS231A: Computer Vision, From 3D Reconstruction to Recognition**

<http://web.stanford.edu/class/cs231a/>

- **CS 543/ECE 549: Computer Vision** <http://slazebni.cs.illinois.edu/spring19/>

- **CS 376: Computer Vision** <http://vision.cs.utexas.edu/376-spring2018/>

55

Grading policy

- Attendance: 10%
- Homework 1: 15%
- Homework 2: 15%
- Classroom test(open-book): 60%

56

Why should you take the class?

- Become a vision researcher
 - CVPR 2020 conference
 - ICCV 2020 conference
- Become a vision engineer in industry
 - Perception team at Google AI
 - Vision at Google Cloud
 - Vision at Facebook AI
 - Vision at SenseTime
 - Vision at MEGVII
 - Vision at Tencent AI Lab
- General interest

58

Overall Philosophy

Breadth

- Computer vision is a huge field
- It can impact every aspect of life and society
- It will drive the next information and AI revolution
- Pixels are everywhere in our lives and cyber space
- This course is meant as an broad overview course, we will not cover all topics of CV
- Lectures are mixture of detailed techniques and high level ideas
- Speak our “language”

Depth

- Computer vision is a highly technical field, i.e. know your math!
- Master bread-and-butter techniques: face recognition, corners, lines, features, optical flows, clustering and segmentation

59

Roadmap

Machine Vision Technology							
Semantic information				Metric 3D information			
Pixels	Segments	Images	Videos	Camera		Multi-view Geometry	
Convolutions Edges & Fitting Local features Texture	Segmentation Clustering	Recognition Detection	Motion Tracking	Camera Model	Camera Calibration	Epipolar Geometry	SFM
10	4	4	2	2	2	2	2

60