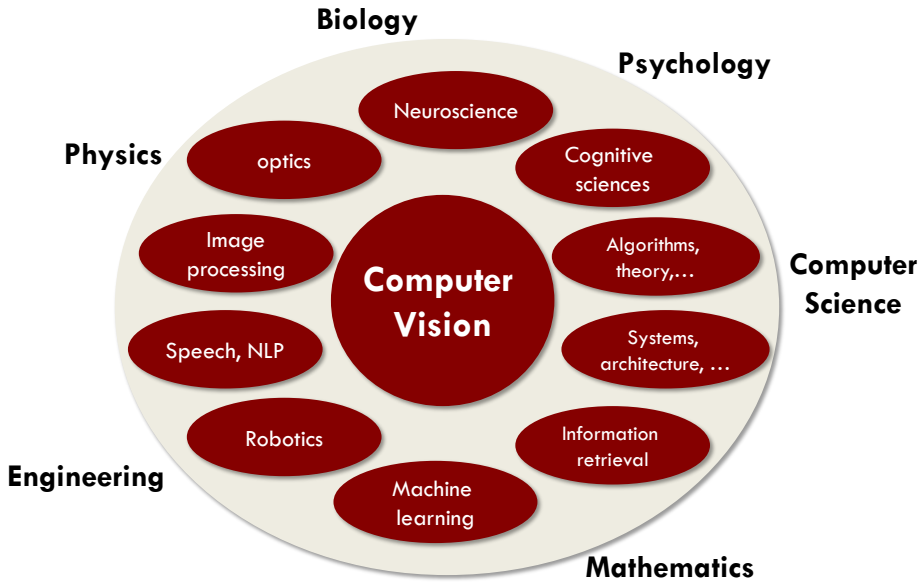
# Machine Vision Technology

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?

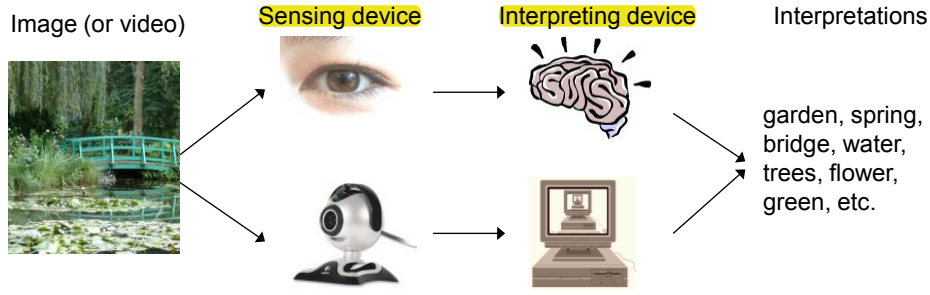
3D Reconstruction

Michaelangelo's David

Recognizing chairs

6

# What is (computer) vision?




7

# The goal of computer vision

- To bridge the gap between pixels and "meaning" 跨越语义鸿沟


What we see

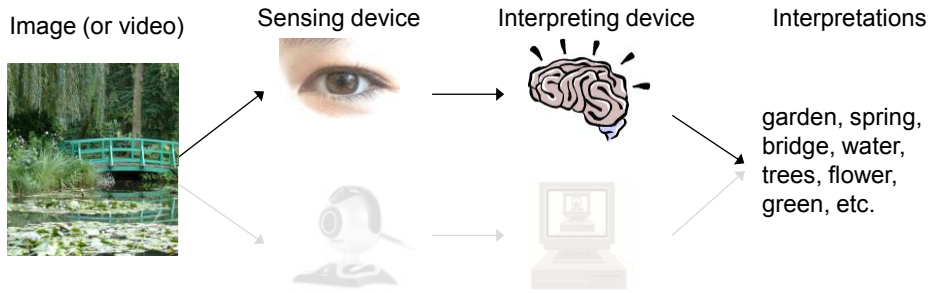
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 3 | 2 | 5 | 4 | 7 | 6 | 9 | 8 |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 1 | 0 | 3 | 2 | 5 | 4 | 7 | 6 |
| 5 | 2 | 3 | 0 | 1 | 2 | 3 | 4 | 5 |
| 4 | 3 | 2 | 1 | 0 | 3 | 2 | 5 | 4 |
| 7 | 4 | 5 | 2 | 3 | 0 | 1 | 2 | 3 |
| 6 | 5 | 4 | 3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

视觉认知是分层次的

Hubel & Wiesel

Dr. Hubel said:

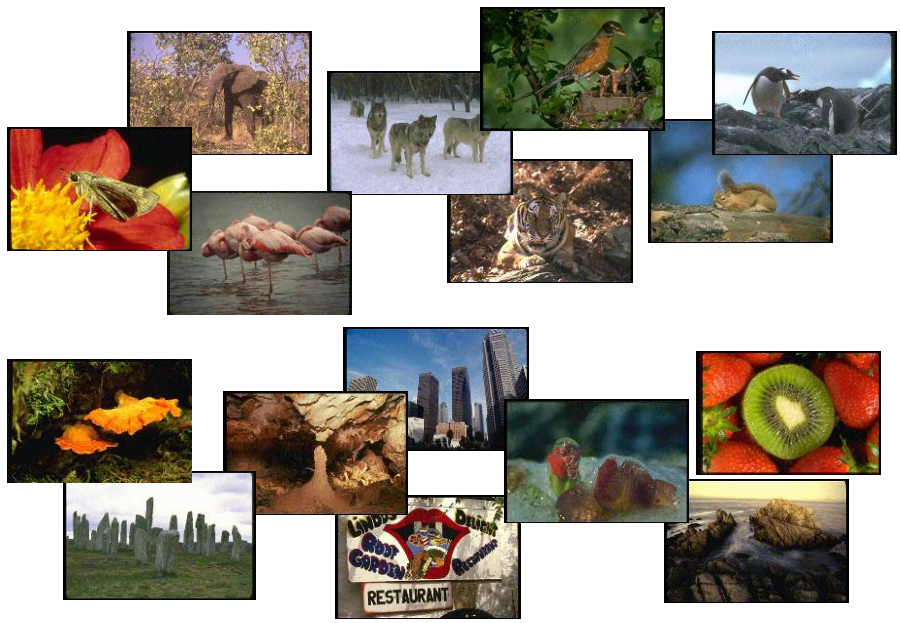
There has been a myth that the brain cannot understand itself. It is compared to a man trying to lift himself by his own bootstraps. We feel that is nonsense. The brain can be studied just as the kidney can.

10

# Human vision is superbly efficient

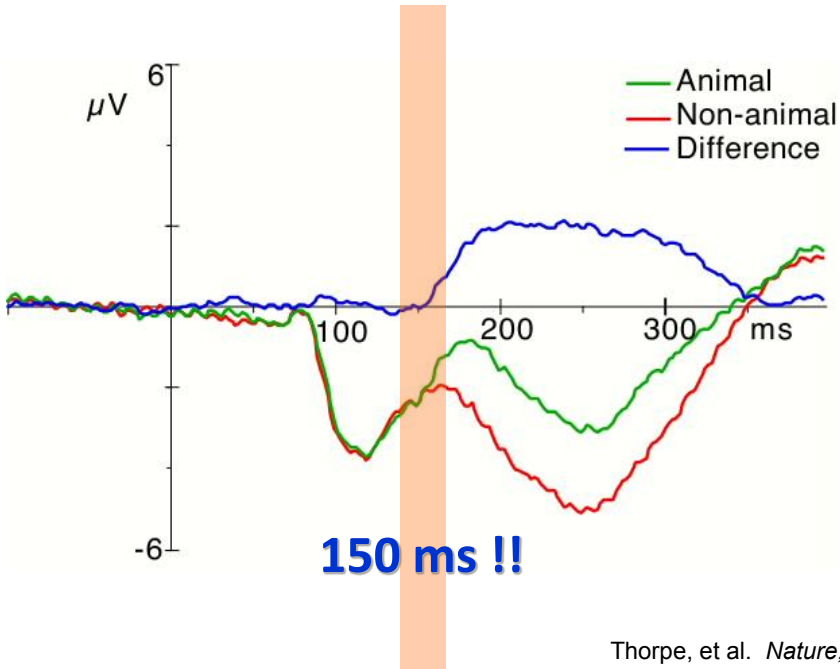
Potter, Biederman, etc. 1970s

11




Thorpe, et al. *Nature,* 1996

12




Thorpe, et al. *Nature,* 1996

13

# Change Blindness 运动视盲

Rensink, O'regan, Simon, etc.

14

# Change Blindness

Rensink, O'regan, Simon, etc.

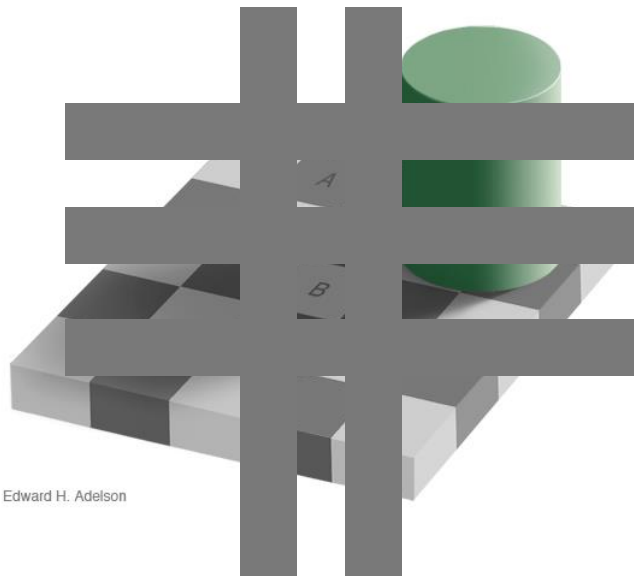
15

# Segmentation 语义分割

16

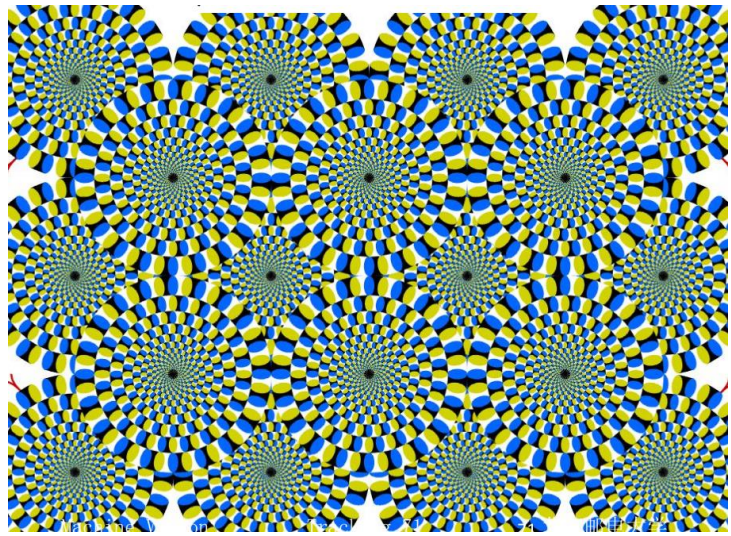
# Perception 认知能力

17




18

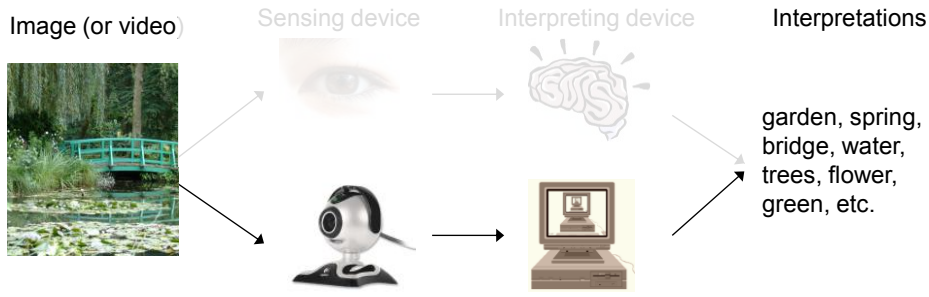
# Motion without movement



19

20

# What is (computer) vision?




21

# The goal of computer vision

- To bridge the gap between pixels and “meaning”


What we see

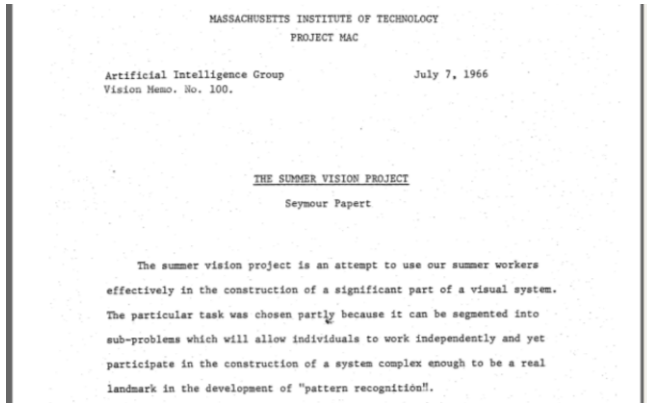
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 3 | 2 | 5 | 4 | 7 | 6 | 9 | 8 |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 1 | 0 | 3 | 2 | 5 | 4 | 7 | 6 |
| 5 | 2 | 3 | 0 | 1 | 2 | 3 | 4 | 5 |
| 4 | 3 | 2 | 1 | 0 | 3 | 2 | 5 | 4 |
| 7 | 4 | 5 | 2 | 3 | 0 | 1 | 2 | 3 |
| 6 | 5 | 4 | 3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



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

23

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

提出了总体性框架

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

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

**Hardware implementation**: How are the representations and algorithms mapped onto actual hardware,
硬件
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》

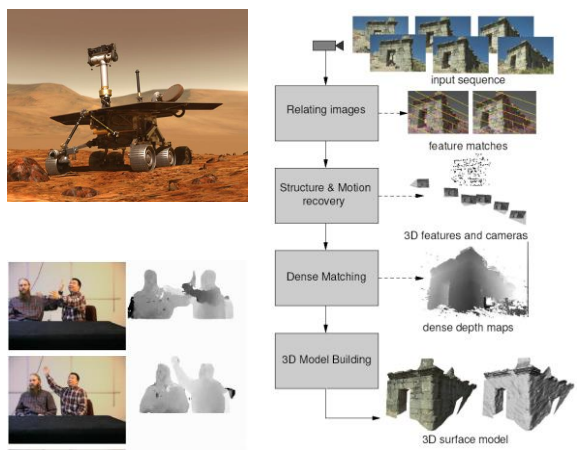
24

# What kind of information can we extract from an image?

- Metric 3D information
- Semantic information

25

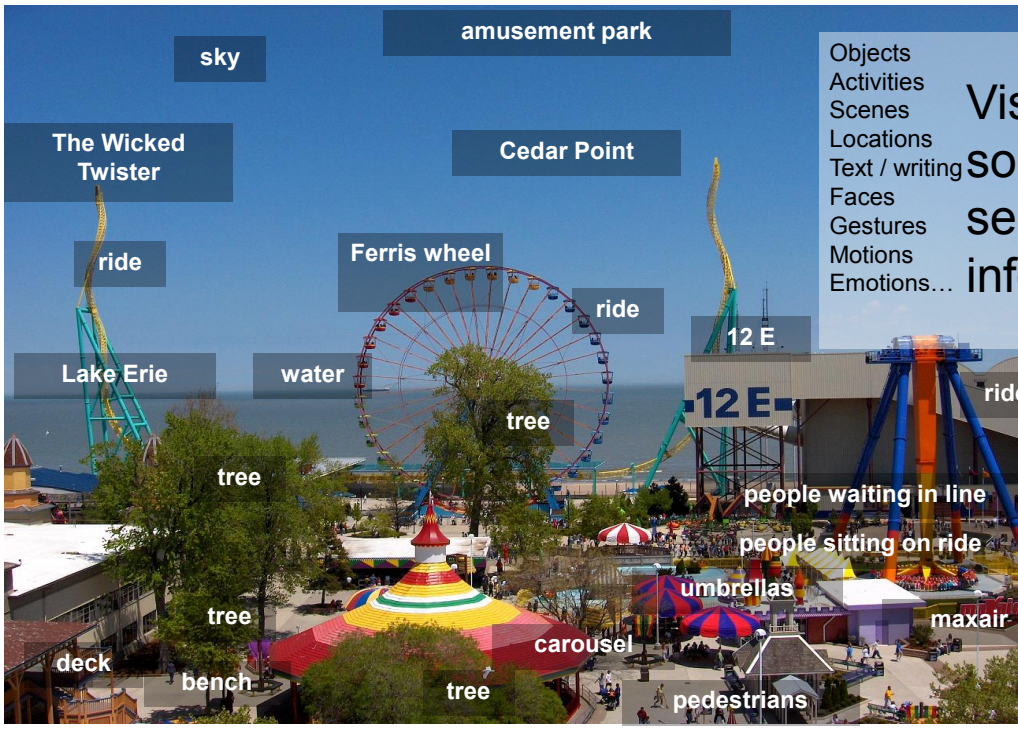
# Vision as measurement device




Pollefeys et al.

Goesele et al.

26

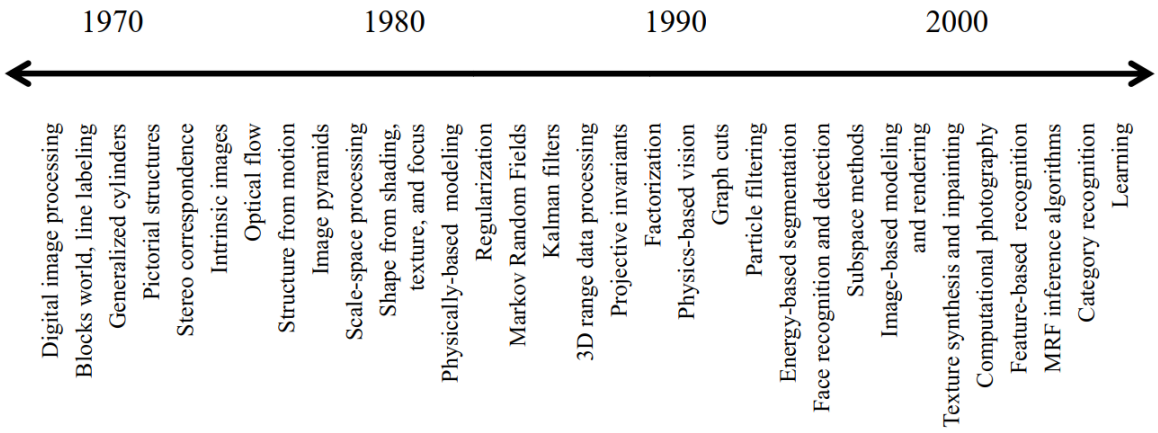



Vision as a source of semantic information

Slide credit: Kristen Grauman

27

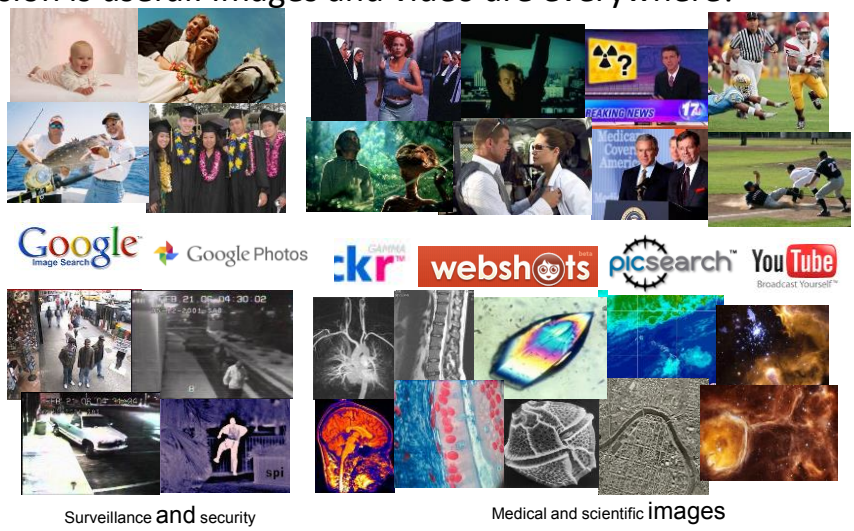
# Most active topics of research in computer vision




28

# Why study computer vision?

- Vision is useful: Images and video are everywhere!




Surveillance and security

Medical and scientific images

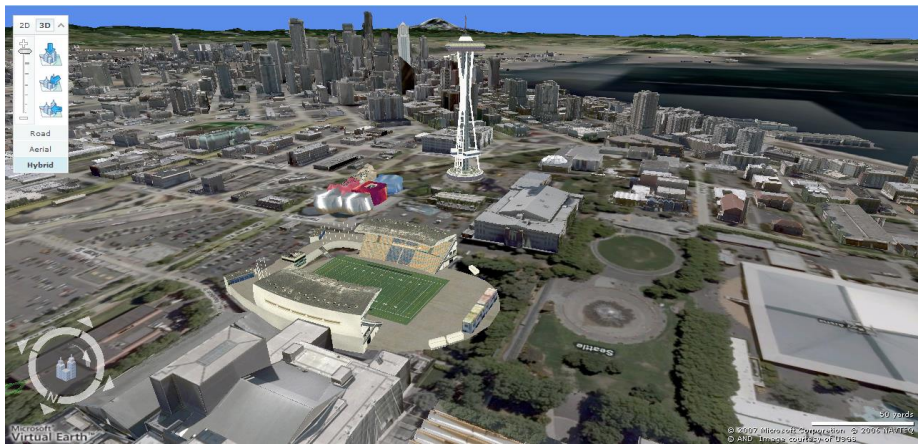
29

# Special effects: shape and motion capture

Source: S. Seitz

30

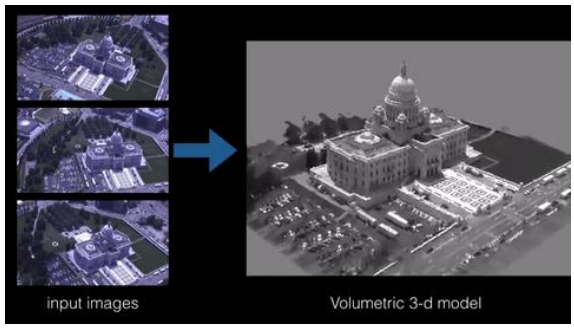
# 3D urban modeling

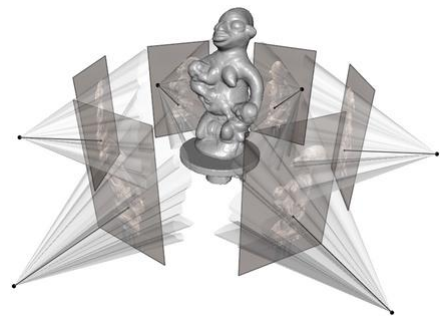



Bing maps, Google Streetview

31

# 3D modeling



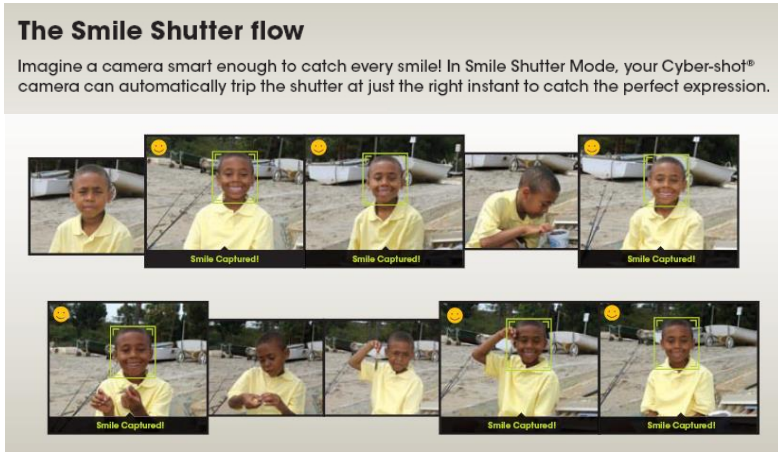
32

# Face detection


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

33

# Smile detection




Sony Cyber-shot® T70 Digital Still Camera

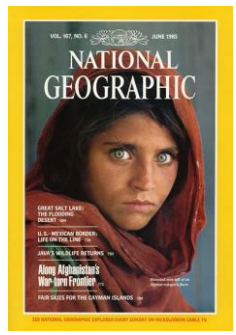
34

## Face recognition: Apple iPhoto software


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

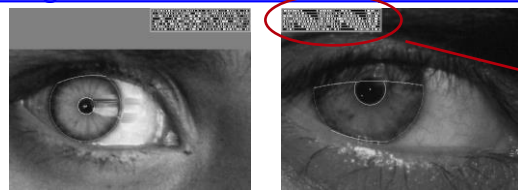
35

## Biometrics




How the Afghan Girl was Identified by Her Iris Patterns

瞳孔采集相对困难



虹膜识别: 展开比较相似度

Source: S. Seitz

36

# Biometrics

指纹采集: 依据指纹中存在不同类型的点，例如分岔点、汇聚点、断点等。对其进行定义，然后根据点的空间位置判断是


Fingerprint scanners on many new laptops, other devices

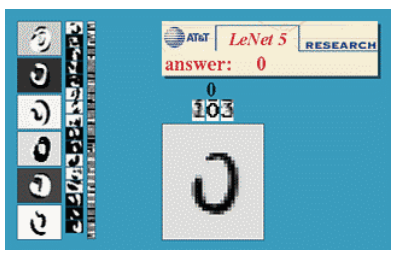
Face recognition systems now beginning to appear more widely
iphone X just introduced face recognition

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

# Google maps: Annotate all houses and streets


Avenue des Sapins

Goodfellow et al. 2014

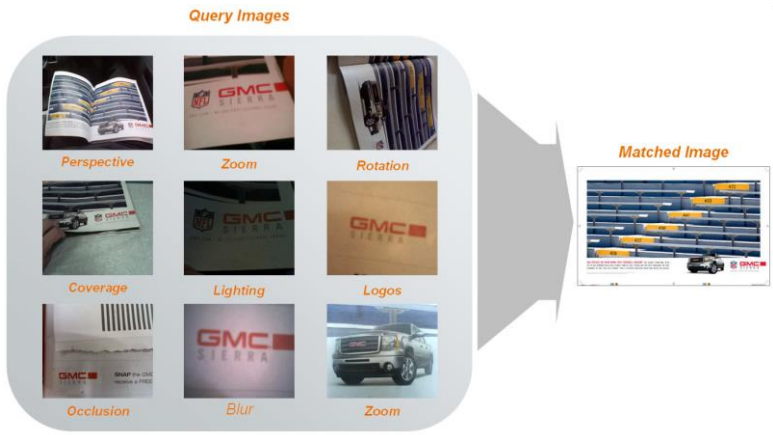
39

# Toys and Robots

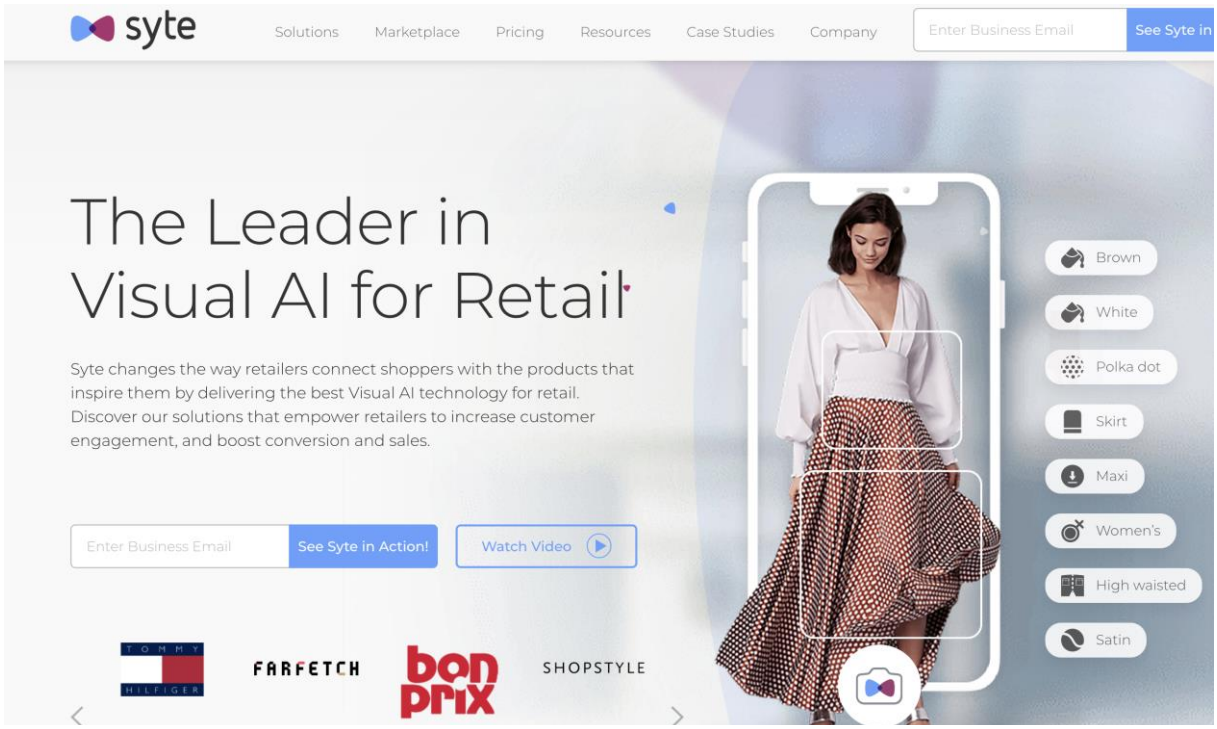
40

# Mobile visual search: iPhone Apps

snaptell kooaba

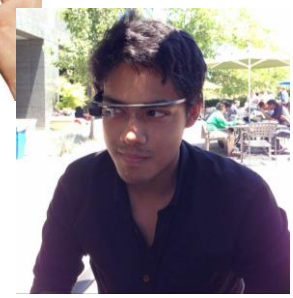



41




42

# Snapstacles and Google glasses



- That's Ranjay in undergrad ->

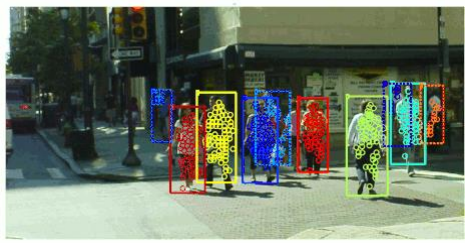
43

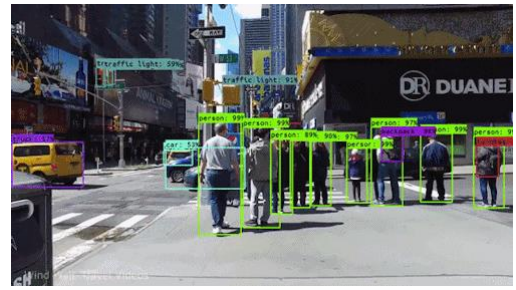
# Automotive safety


- Mobileye

44

# Detection and tracking





45

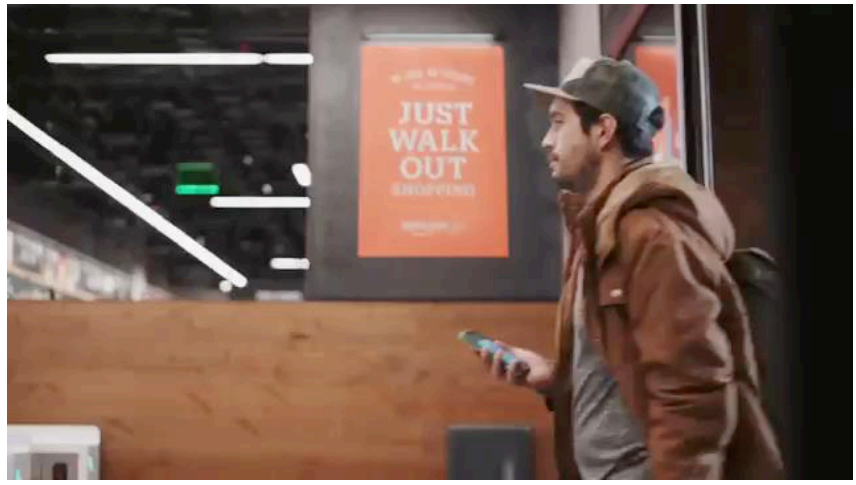
# Vision in supermarkets


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

双目立体视觉技术

# Vision-based interaction (and games)


Microsoft's Kinect

Sony EyeToy

Assistive technologies

48

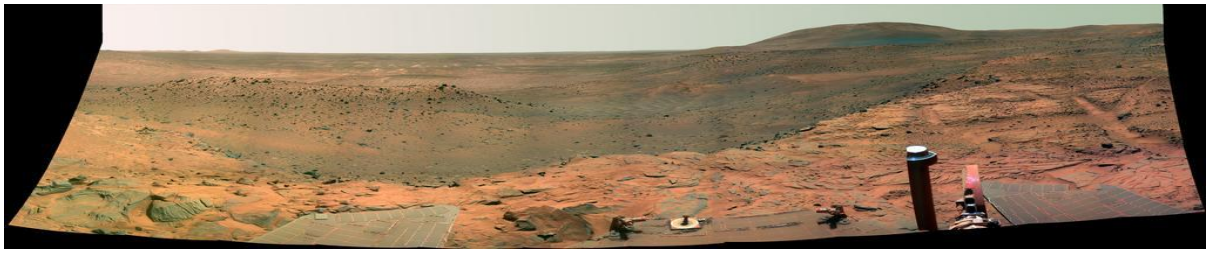
# Augmented Reality


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.

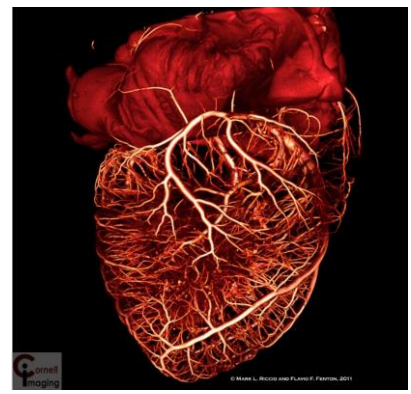
Vision systems (JPL) used for several tasks

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

Source: 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<br>Edges & Fitting<br>Local features<br>Texture | Segmentation<br>Clustering | Recognition<br>Detection | Motion<br>Tracking | | Camera<br>Model | Camera<br>Calibration | Epipolar<br>Geometry | SFM |
| 10 | 4 | 4 | 2 | | 2 | 2 | 2 | 2 |

60