

计算机视觉基础

齐志

东南大学 信息科学与工程学院

- 1.什么是计算机视觉？**
- 2.为什么要学习计算机视觉？**
- 3.怎样学习计算机视觉？**

东南大学 信息科学与工程学院

What is computer vision?



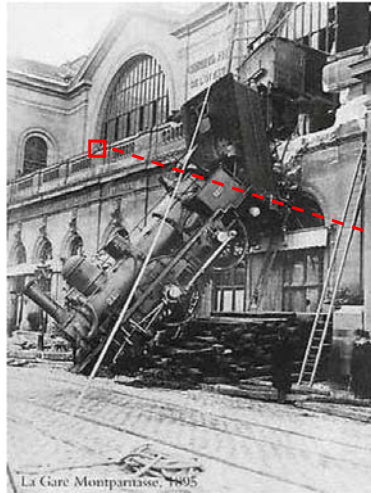
a computer + a connected camera? **No!**

What is computer vision?



what we see is that every
picture tells a story.

What is computer vision?

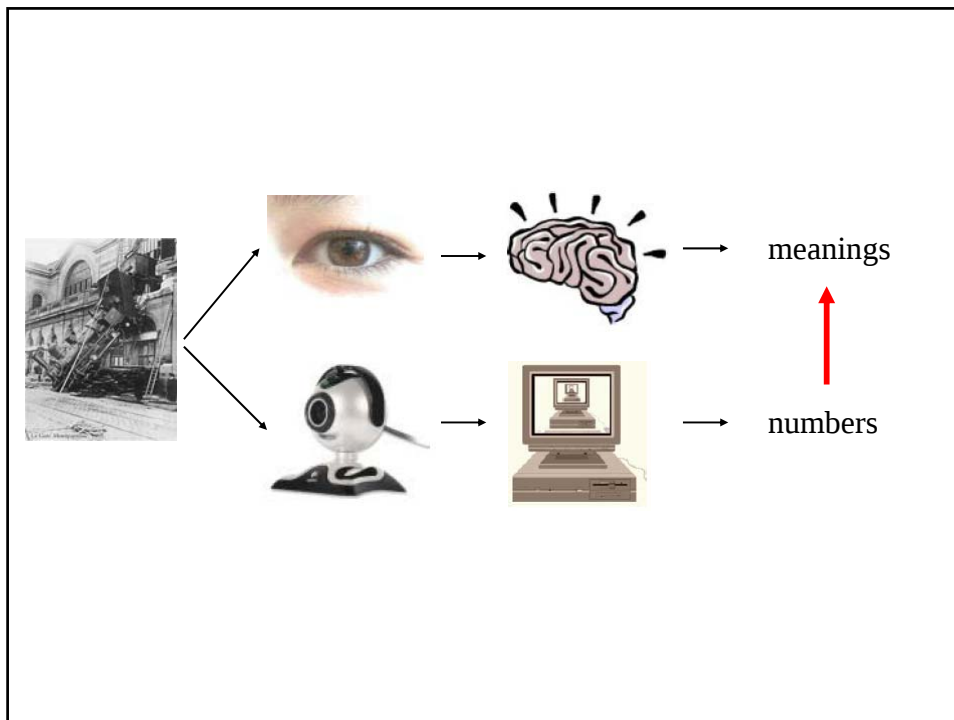


What a computer can see is a matrix containing numbers.

216	198	169	169	171	205
210	196	173	172	173	203
206	195	176	174	175	202
195	184	167	174	178	192
192	182	166	176	181	193
186	178	164	182	189	197

What is computer vision?

The goal of computer vision is to bridge the gap between meanings and numbers.

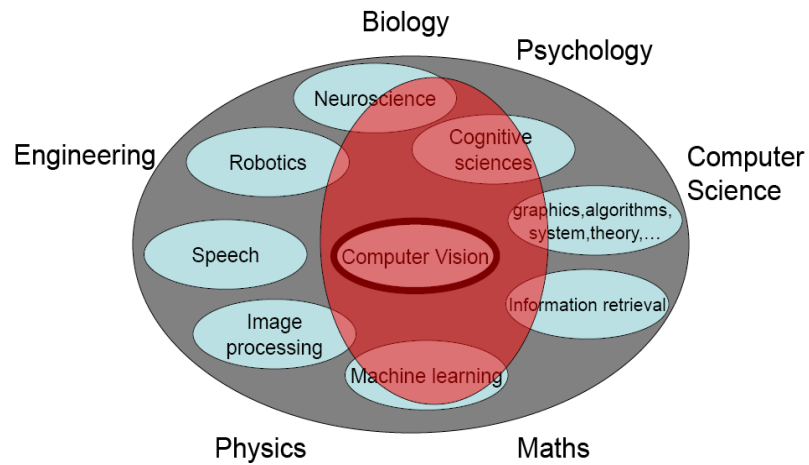


What is it related to?

■ Closely related fields of study

- Artificial intelligence and machine learning (CS)
- Spatial statistics (Math/Stats)
- Image processing (EE/CE)
- Algorithms and optimization (CS/OR)
- Optics and reflectance models (Physics)
- Human visual perception and cognition (Psych)
- Animal vision (Neuroscience)

What is it related to?



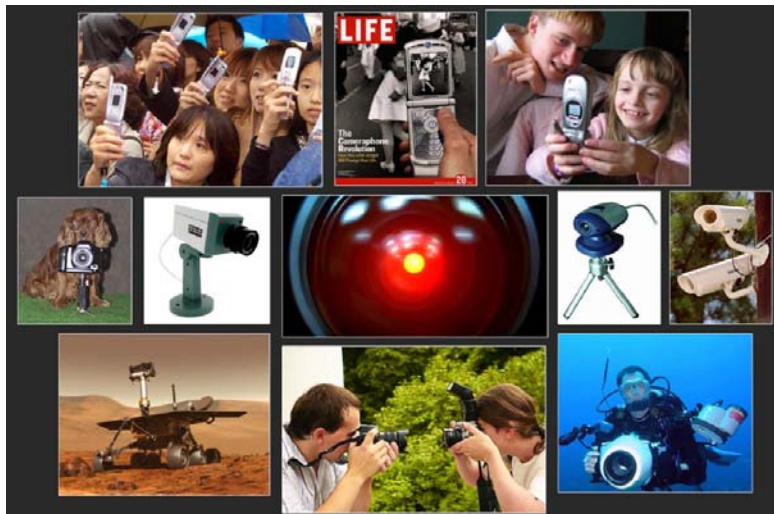
What is it related to?

- Basically, it's an interdisciplinary science concerned with the theory behind artificial systems that extract information from images.

Why Computer vision?

- It is everywhere.
- It is useful.
- It is difficult.

- It is everywhere. Wherever there's a camera, there is a potential computer vision application.



■ It is useful for its wide range of application areas

- Controlling processes
- Detecting events
- Organizing information
- Modeling objects or environments
- Interaction
- ...

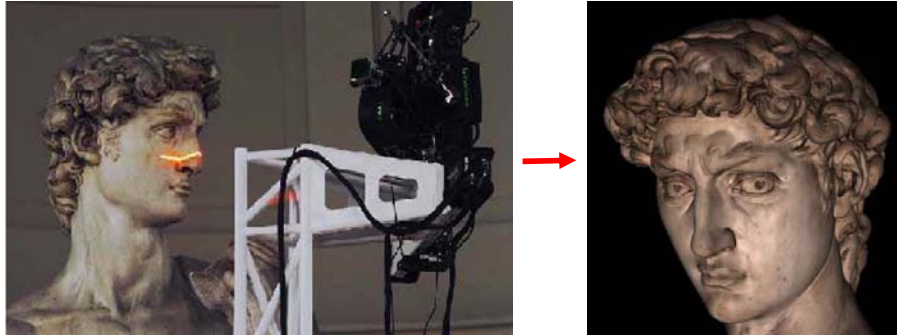
Next, show some state of the art examples

Vision in Space



[NASA'S Mars Exploration Rover Spirit](#)

3D Scanning



[The Digital Michelangelo Project](#), which builds the precise 3D model with the accuracy of 0.29mm

Vision-based Biometric



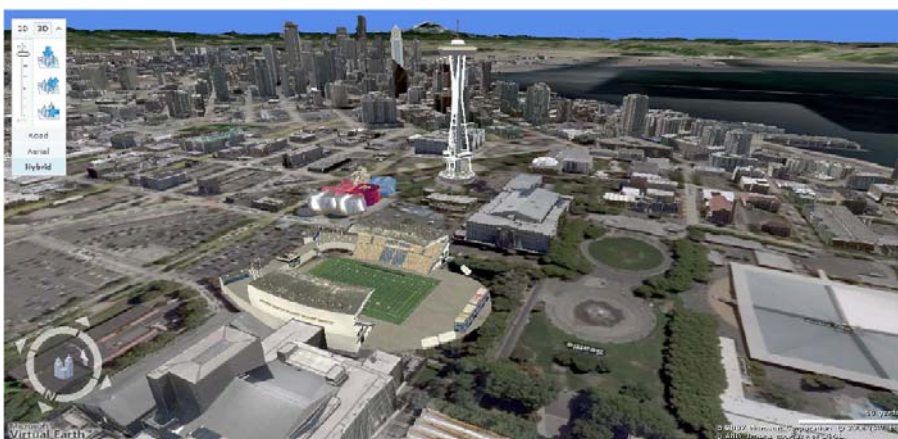
[How the afghan girl was identified by her iris patterns](#)

Optical Character Recognition (OCR)



Technology to convert scanned documents to text, which is popular in today's scanner.

3D Urban Modeling



Bing maps, Google Streetview

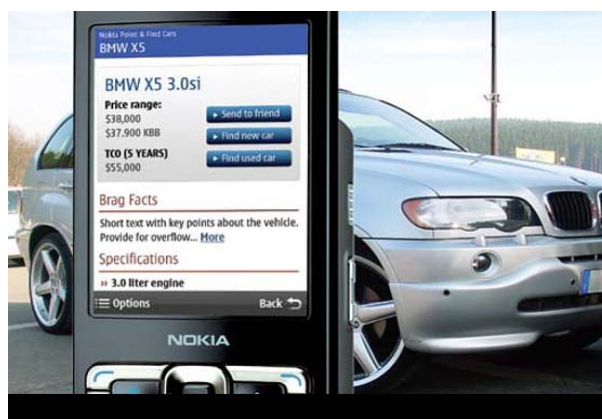
Smart Cars – Automobile Safty



Vision systems in high-end BMW, GM, Volvo models

"In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians."

Mobile Object Recognition



Recognize the object through the mobile camera immediately.

Vision-based Interaction and Games



Xbox, kinect

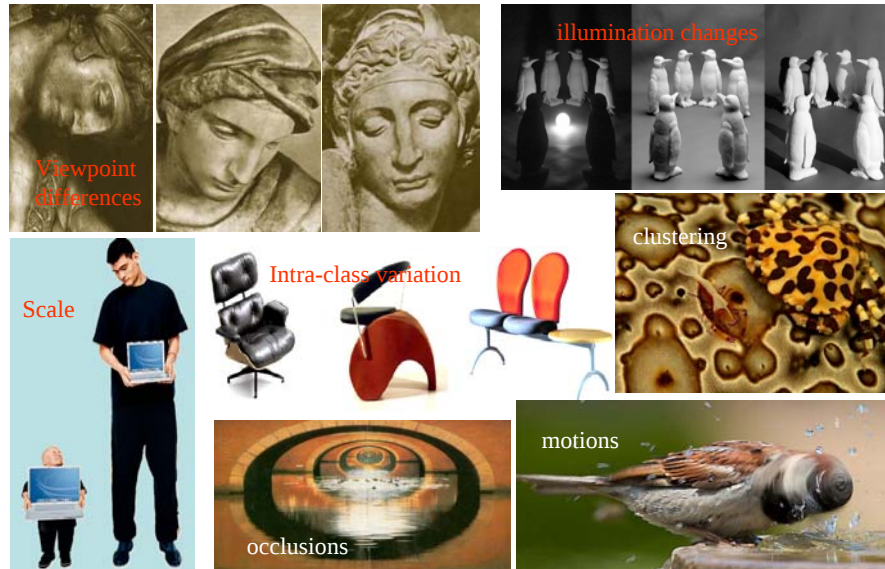


Sony, Eyetoy

■ It is difficult.

- Half of primate cerebral cortex is devoted to visual processing [Felleman and van Essen 1991]
- Achieving human-perception is a “AI-complete” problem

Difficulties due to various factors



How to study Computer vision?

- Textbook:
 - Computer Vision: A Modern Approach,
by D. Forsyth and J. Ponce
 - Computer Vision: Algorithms and Applications,
by Richard Szeliski

- Prerequisites

- Data structures
- A good working knowledge of C, C++, or Matlab
- Linear algebra
- Vector calculus
- Fundamental statistics

- Course does **not** assume prior imaging experience in computer vision, image processing, graphics, etc.

- Grading Policy

- Assignment (60%)
- Final report + Class participation (40%)

■ Basically, it's an interdisciplinary science concerned with the theory behind artificial systems that extract information from images.

■ Thus focus on algorithms

■ Course outline (keywords that may appear)

Image formation	Visual motion/optical flow
Filtering	Parametric motion
Edge detection	Structure from motion
Corner detection	Stereo
Interest points – SIFT features	Markov Random Fields for stereo
Matching, Chamfer and Hausdorff	MRF inference
Distance transforms	Image segmentation
3D camera geometry	Face recognition, subspace methods
Multiview geometry	Object category recognition
Image panoramas	Flexible template models
Image stitching/mosaicing	Tracking by matching

Then show some examples.

Image fomration

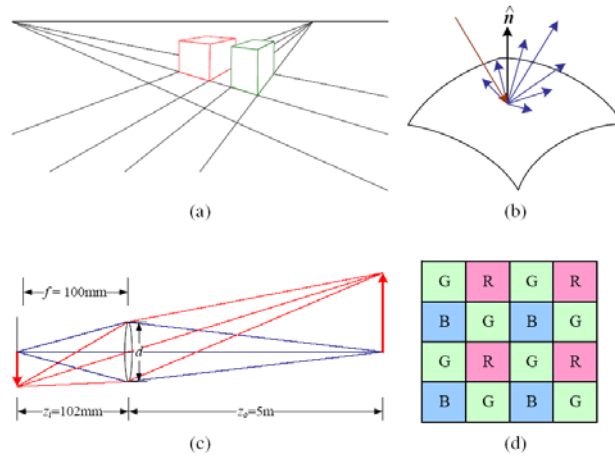


Figure 2.1: A few components of the image formation process: (a) perspective projection; (b) light scattering when hitting a surface; (c) lens optics; (d) Bayer color filter array.

Feature descriptor

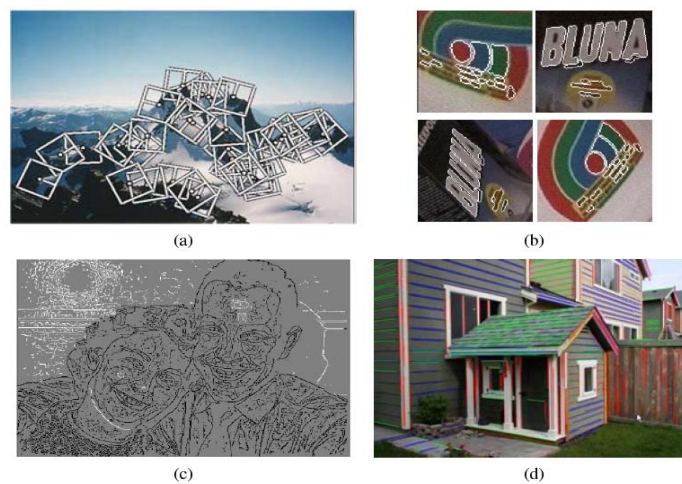


Figure 4.1: A variety of feature detector and descriptors can be used to analyze describe and match images: (a) point-like interest operators (Brown et al. 2005) © 2005 IEEE; (b) region-like

Feature based realtime head tracking



Figure 4.29: Real-time head tracking using the fast trained classifiers of Lepetit et al. (2004) ©

Stereo related topics

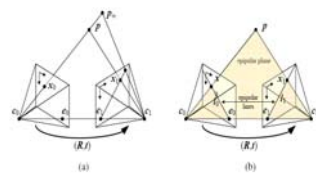
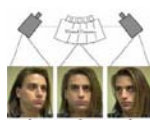


Figure 11.3: Epipolar geometry: (a) epipolar line segment corresponding to one ray; (b)



(d)



(c)



(f)



(g)



(h)

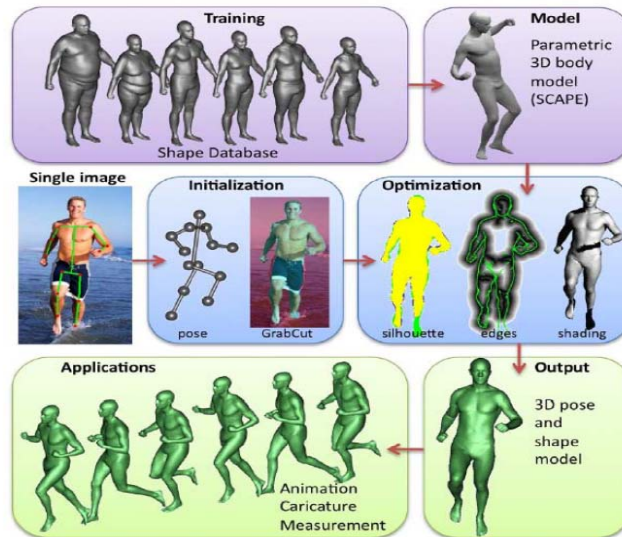


(i)



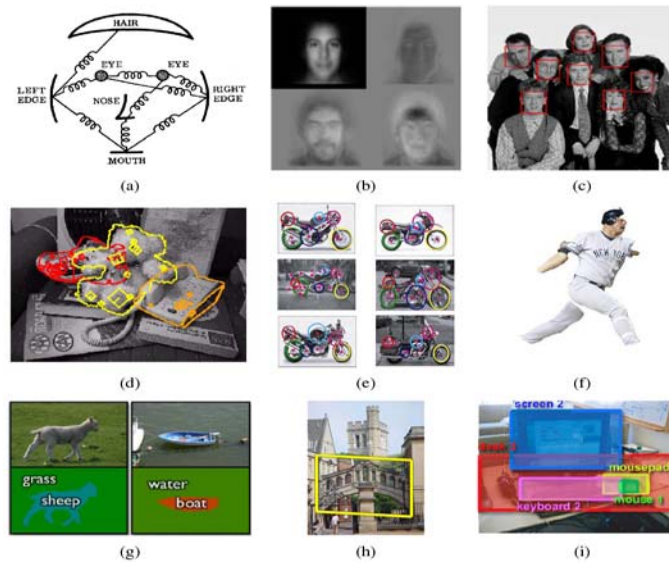
(j)

3D modeling



1: Estimating human shape and pose from a single image using a parametric

Recognitions



Now, how do you feel about this?

Want to join us?

You are sooo... wellcome😊

Thank you!

(Some of the photos used sourced from Seitz, Feifei, Richard, et. al.)

东南大学 信息科学与工程学院