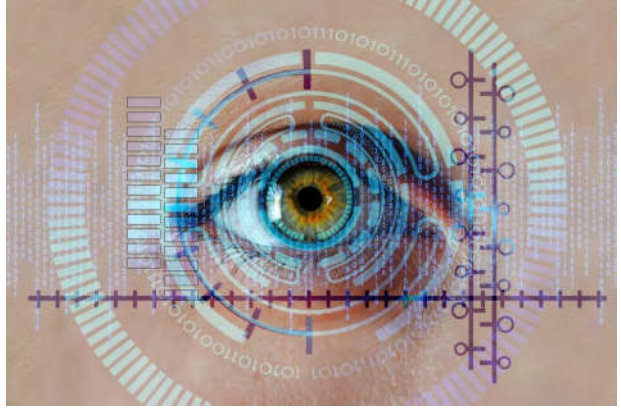
北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

Computer Vision



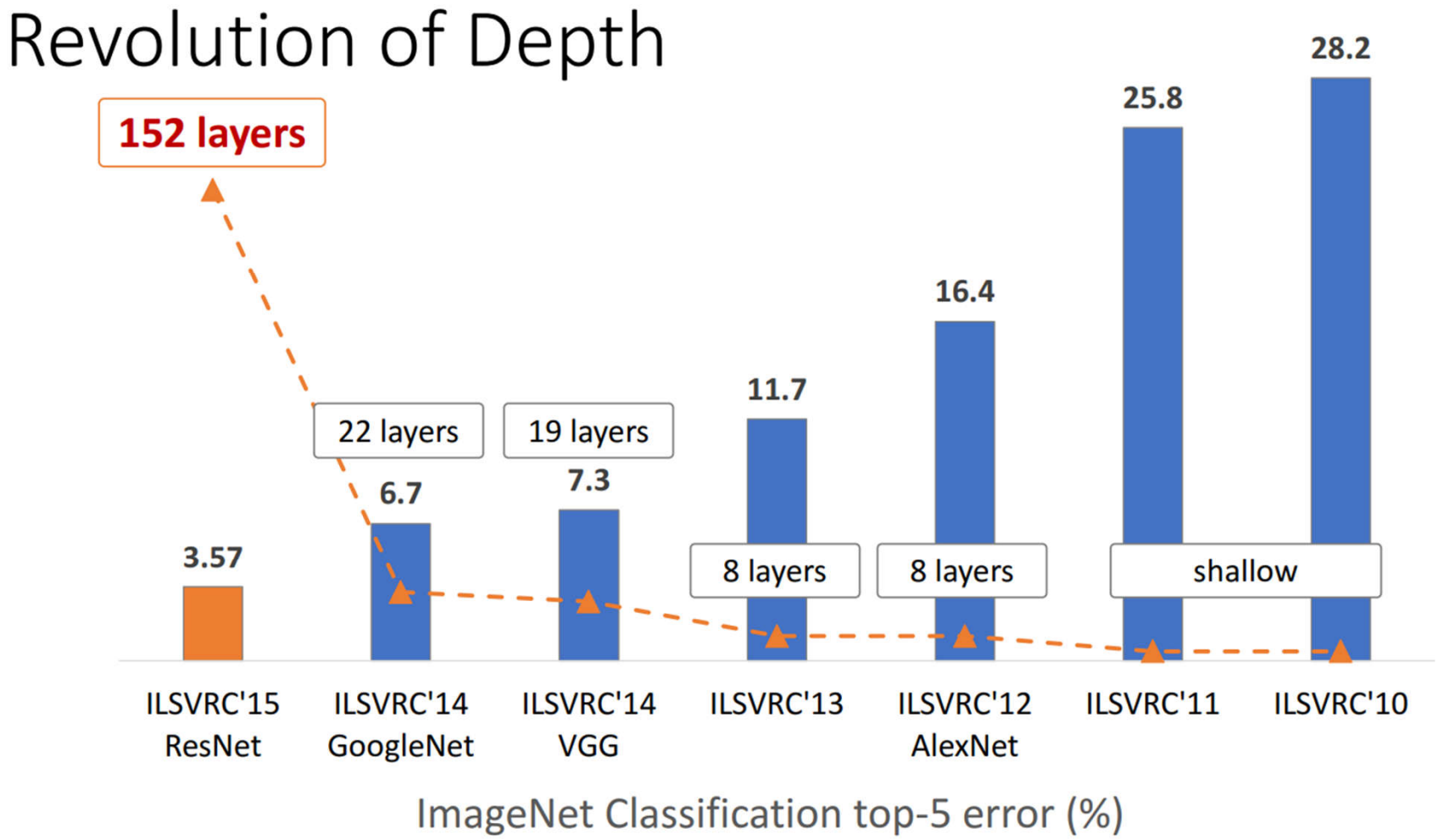
Chapter 1 Introduction

School of Computer Science and Technology

Ying Fu

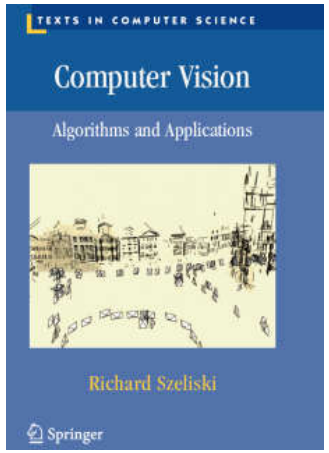
fuying@bit.edu.cn

The development of deep learning



Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun. “Deep Residual Learning for Image Recognition”. CVPR 2016.

About the Course



- Textbook and Reference Material
  - Richard Szeliski. Computer Vision: Algorithms and Applications, Second Edition. Springer, 2021
    https://szeliski.org/Book/
  - David A. Forsyth, Jean Ponce. Computer Vision: A Modern Approach. Second Edition. Publishing House of Electronics Industry, 2012
    https://eclass.teicrete.gr/modules/document/file.php/TM152/Books/Computer%20Vision%20%20A%20Modern%20Approach%20-%20D.%20Forsyth,%20J.%20Ponce.pdf
  - Related papers

About the Course

- Main Conferences
  - Int. Conf. on Computer Vision and Pattern Recognition (CVPR)
  - Int. Conf. on Computer Vision (ICCV)
  - European Conf. on Computer Vision (ECCV)
  - Asian Conf. on Computer Vision (ACCV)
  - Int. Conf. on Pattern Recognition (ICPR)
  - Int. Conf. on Image Processing (ICIP)
  - The Chinese Conference on Pattern Recognition and Computer Vision（PRCV）

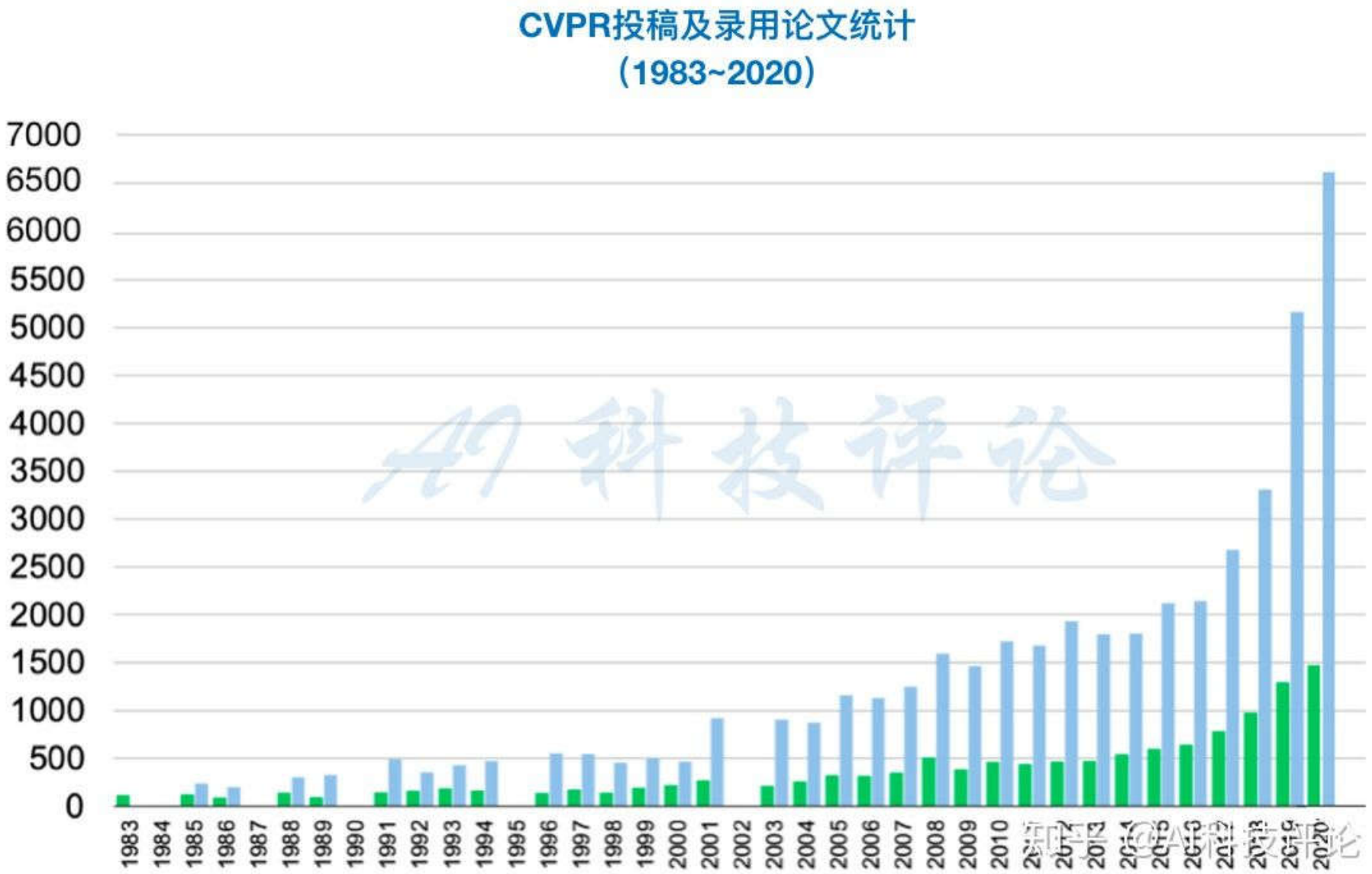
    ……

    https://openaccess.thecvf.com/menu

- Main Journals
  - IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)
  - Int. J on Computer Vision (IJCV)
  - IEEE Trans. on Image Processing (TIP)
  - IEEE Trans. on Neural Networks and Learning Systems(TNNLS)
  - Computer Vision and Image Understanding (CVIU)

    ……

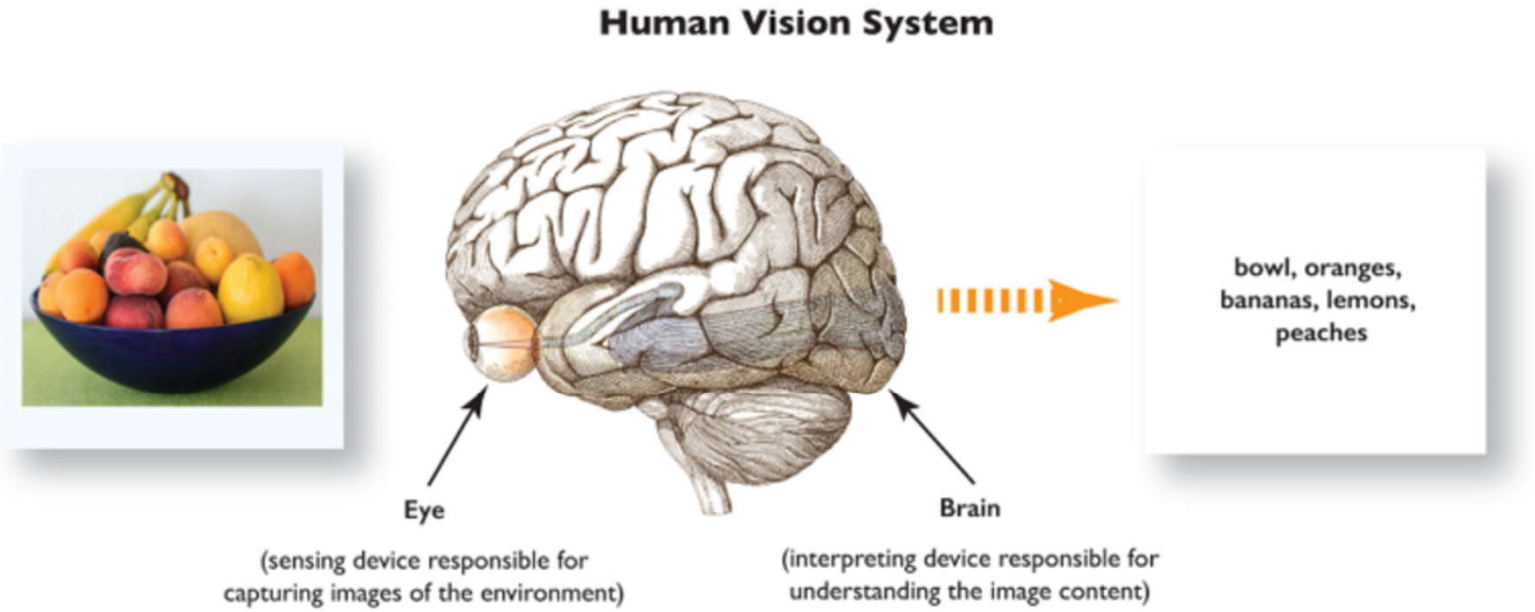
Related material



Outline

- What is computer vision
- The goal of computer vision
- Applications of computer vision
- Challenges
- Course overview

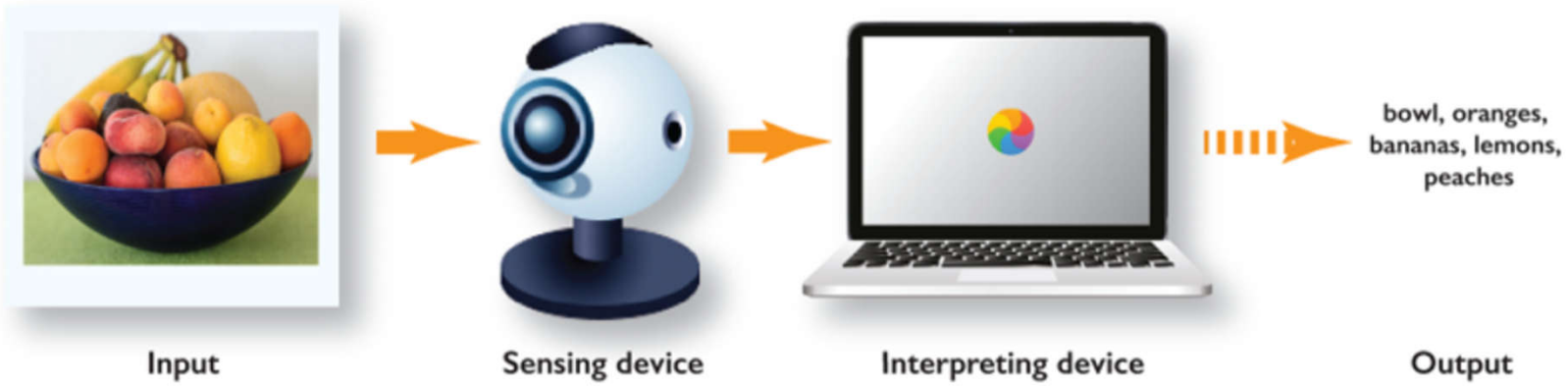
What is Computer Vision



Perceptual

+

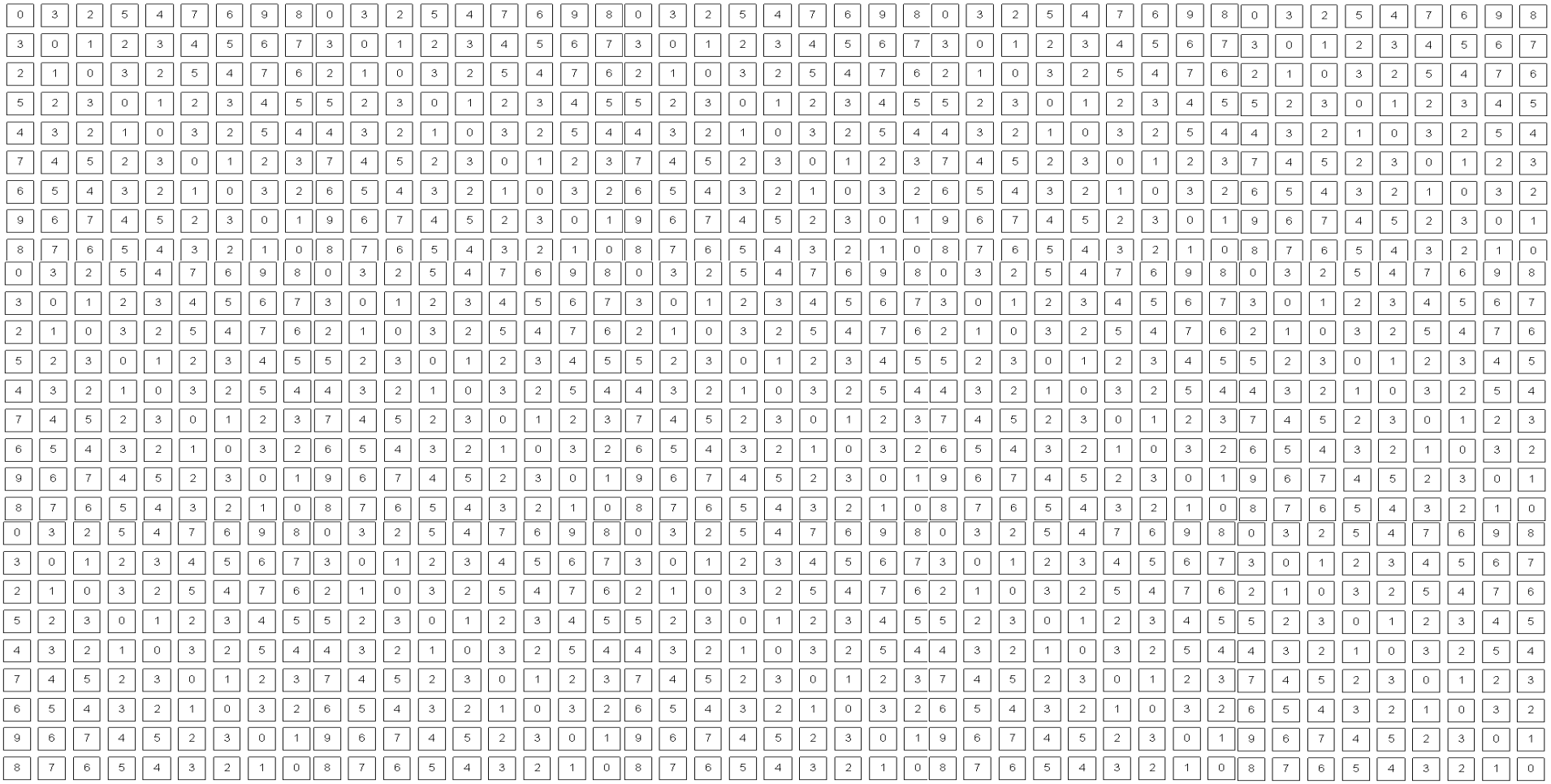
Cognitive



What is Computer Vision

What we see

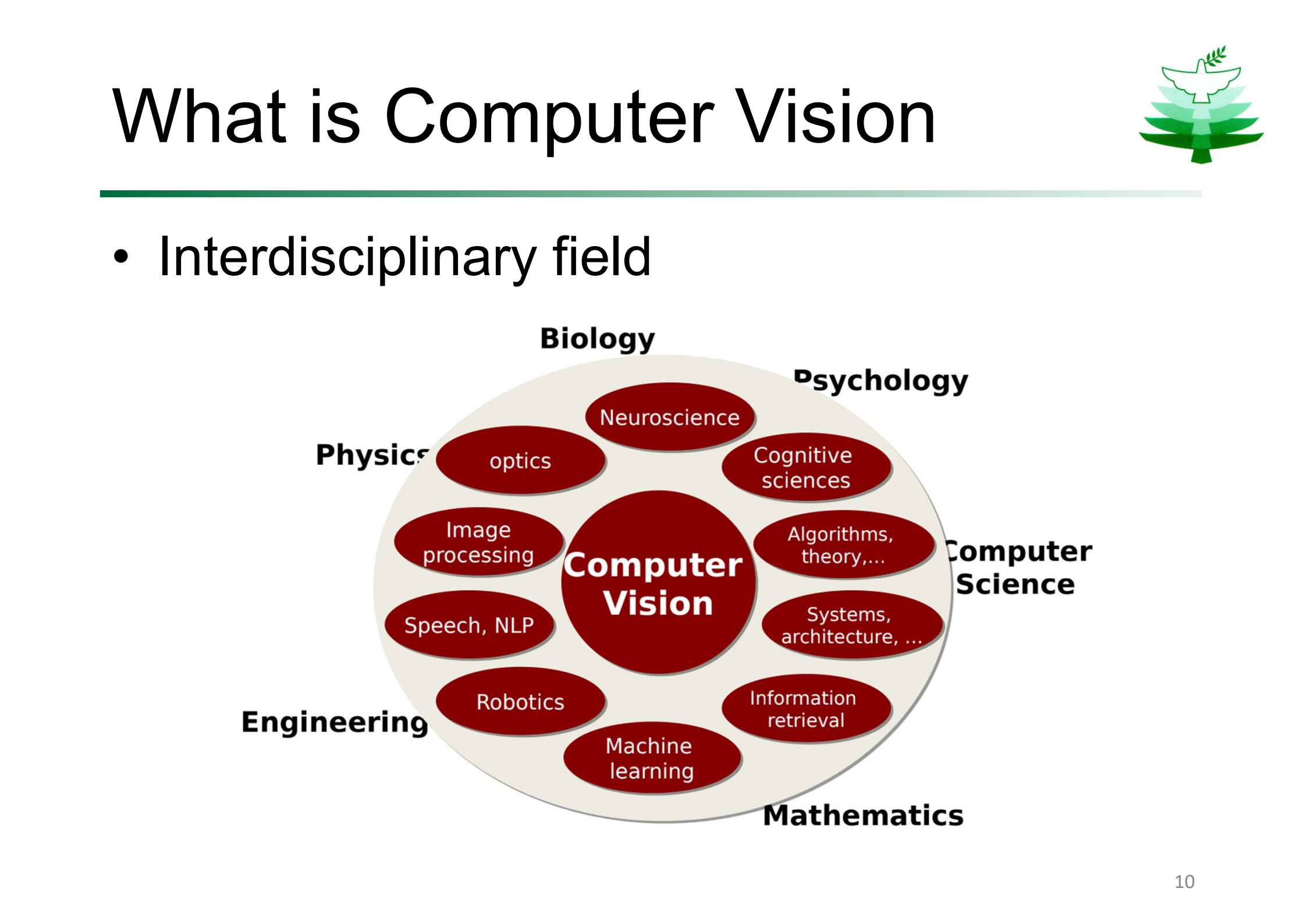
What is Computer Vision



What a computer sees

 510879551@qq.com, 2021/9/14

What is Computer Vision



- Interdisciplinary field

Origins of computer vision

- An MIT undergraduate summer project

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

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".

Outline

- What is computer vision
- The goal of computer vision
- Applications of computer vision
- Challenges
- Course overview

Digital Image

- Digital image can be regarded as two-dimensional matrix or array.
- The basic element of digital image is pixel.
- For intensity image, each pixel has a value (0-255) to represent its brightness.
- For color image, each pixel has a vector to represent its color

Every image tells a story

- Goal of computer vision: perceive the “story” behind the picture
- Compute properties of the world
  - 3D shape
  - Names of people or objects
  - What happened?

The Goal of Computer Vision

- To give computers (super) human-level perception
- Perceive the “story” behind the picture
- 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

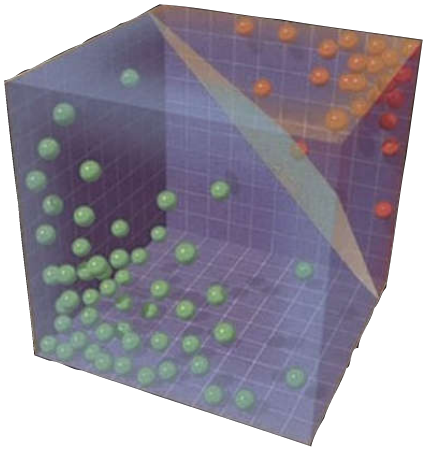
What a computer sees

Can computers match human perception?

- Yes and no (mainly no)
  - computers can be better at “easy” things
  - humans are better at “hard” things
- But huge progress
  - Accelerating in the last five years due to deep learning
  - What is considered “hard” keeps changing

Typical perception pipeline

Input

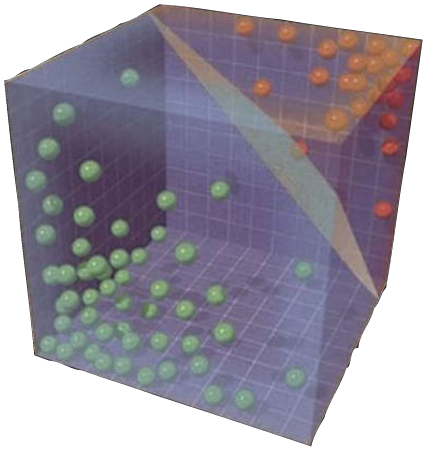


‘fancy math’

output

Typical perception pipeline

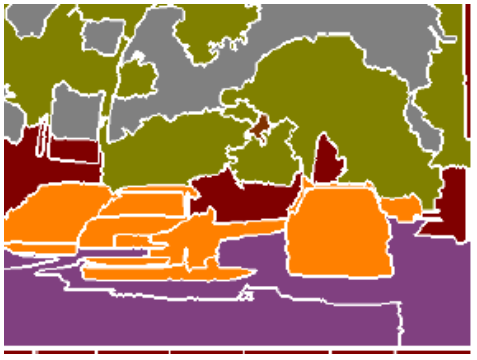
Input



‘fancy math’

output

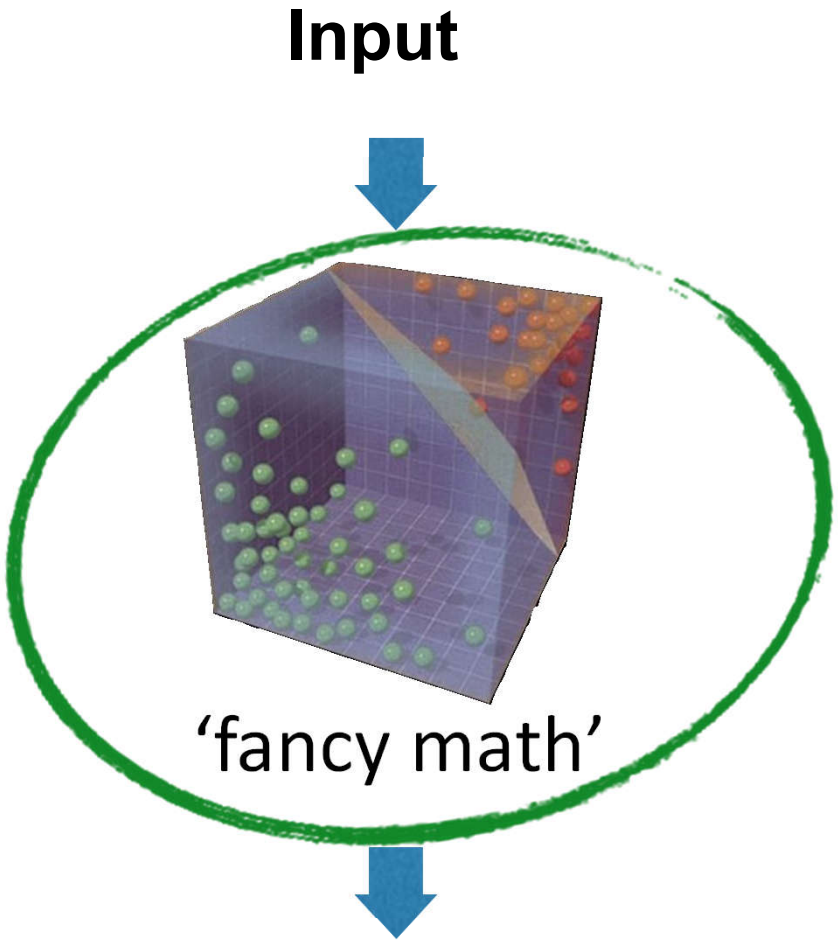
what should we look at?
(image features)



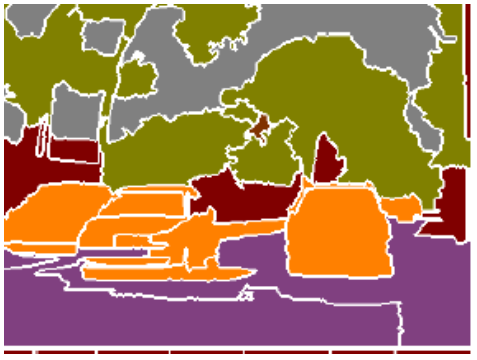
what can we understand?
(semantic segmentation)

Typical perception pipeline

Easy to get lost in the techniques



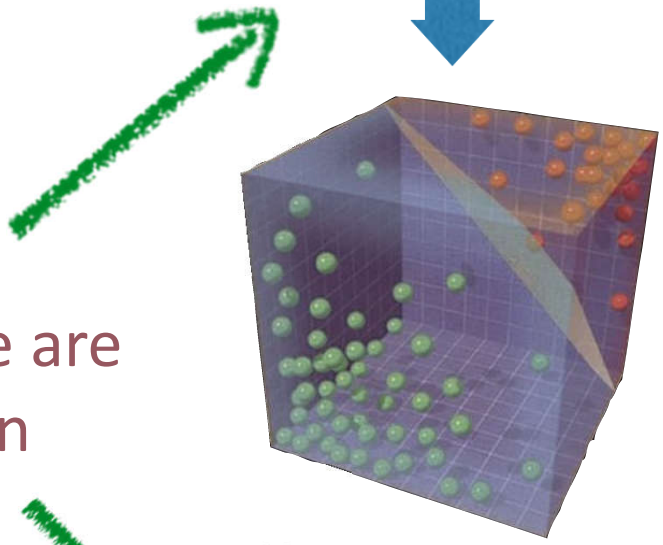
what should we look at?
(image features)



what can we understand?
(semantic segmentation)

Typical perception pipeline

Input

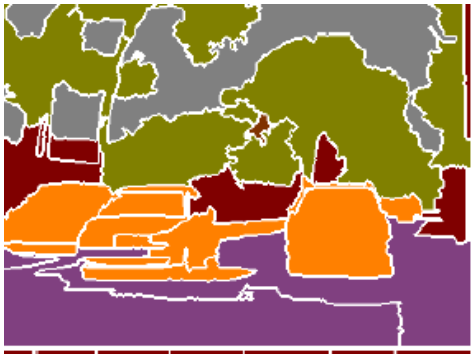


‘fancy math’

output

The parts that we are most interested in

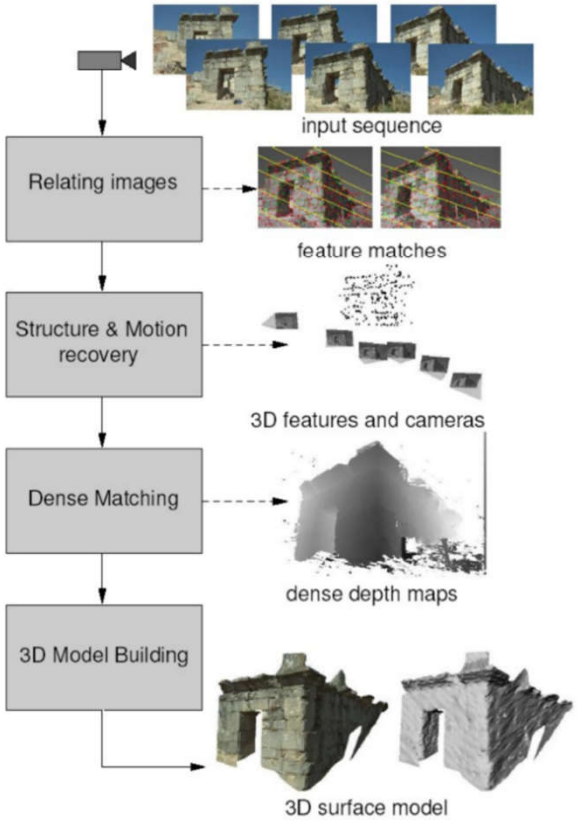
what should we look at?
(image features)



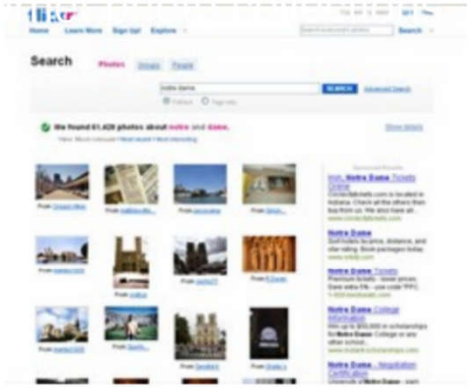
what can we understand?
(semantic segmentation)

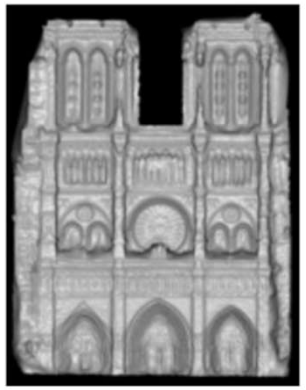
What kind of information can we extract from an image?

- Metric 3D information
  - Vision as measurement device



Pollefeys et al.

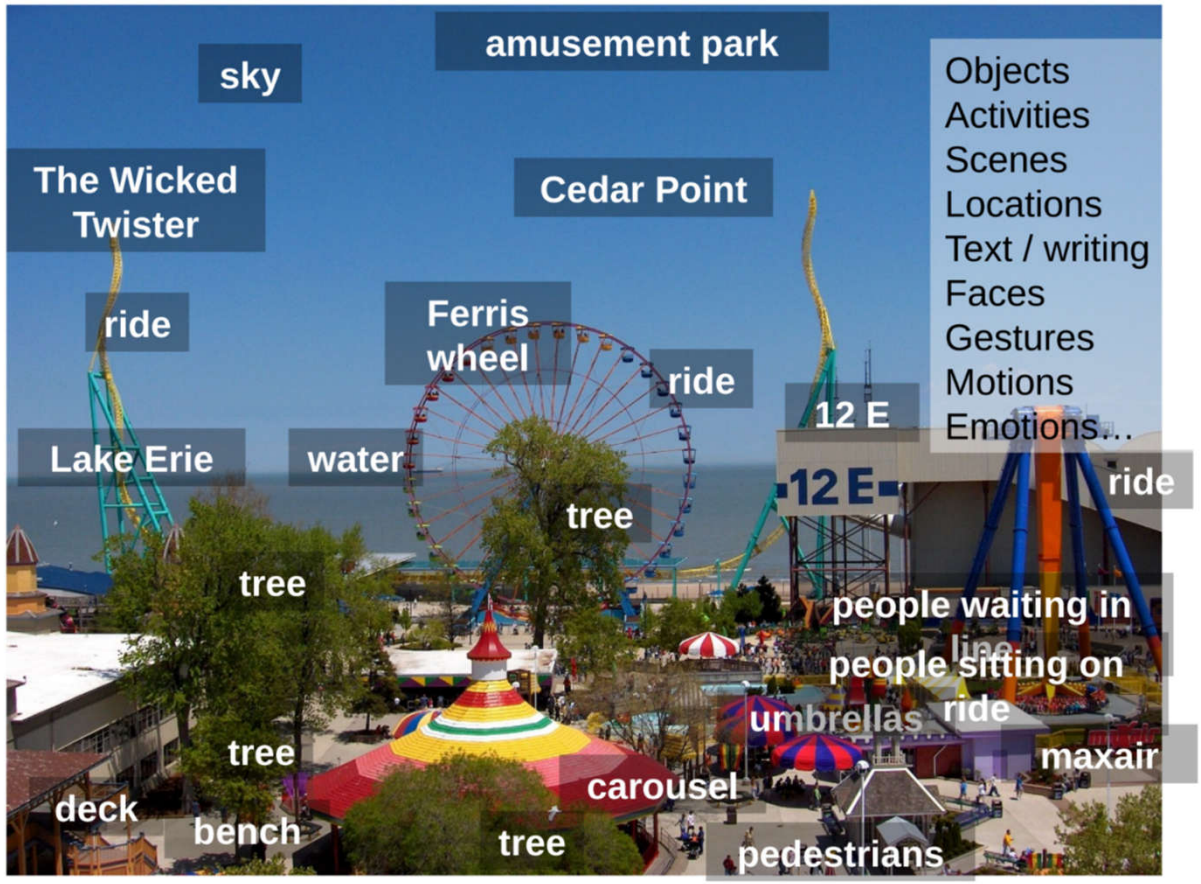




Goesele et al.

What kind of information can we extract from an image?

- Semantic information
  - Vision as a source of semantic information



Slide credit: Kristen Grauman

What kind of information can we expect from an image?

- High-quality photos
  - Improve photos (Computational Photography)

Super-resolution (source: 2d3)

Low-light photography
(credit: Hasinoff et al., SIGGRAPH ASIA 2016)

Depth of field on cell phone camera
(source: Google Research Blog)

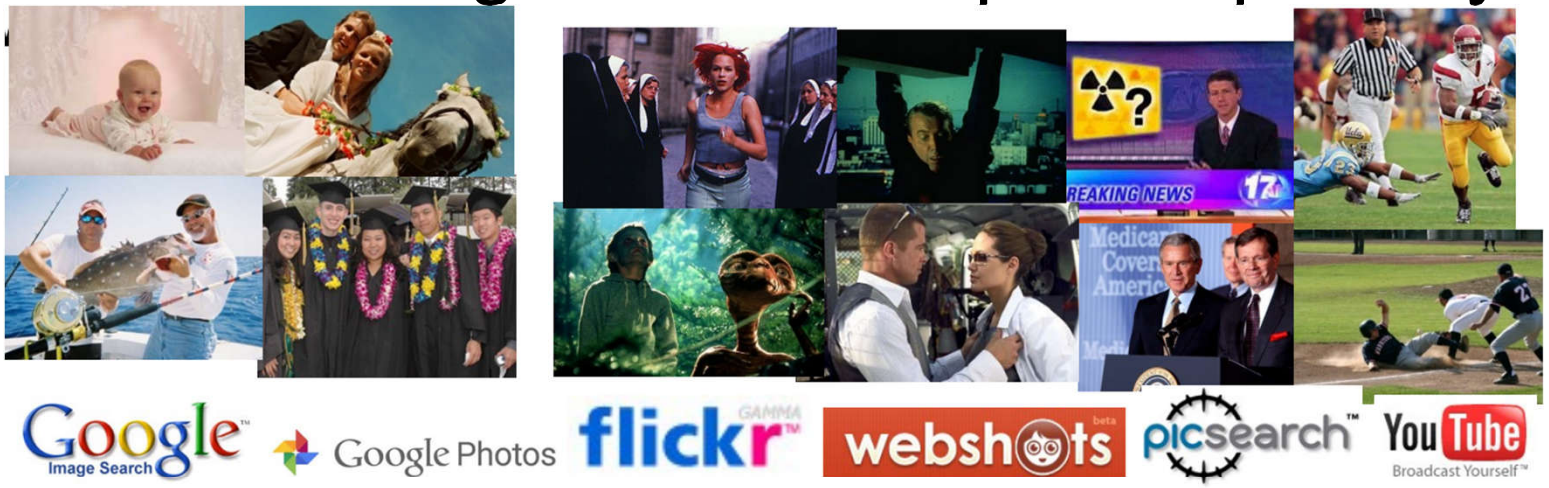
Inpainting / image completion
(image credit: Hays and Efros)

Outline

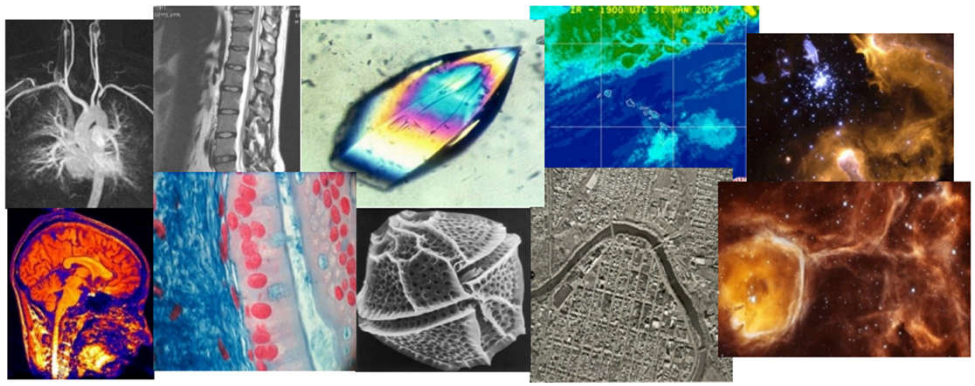
- What is computer vision
- The goal of computer vision
- Applications of computer vision
- Challenges
- Course overview

Why study computer vision?

- Vision is useful: Images/video are everywhere
  - Billions of images/videos captured per day



Surveillance and security



Medical and scientific images

Earth viewers (3D modeling)

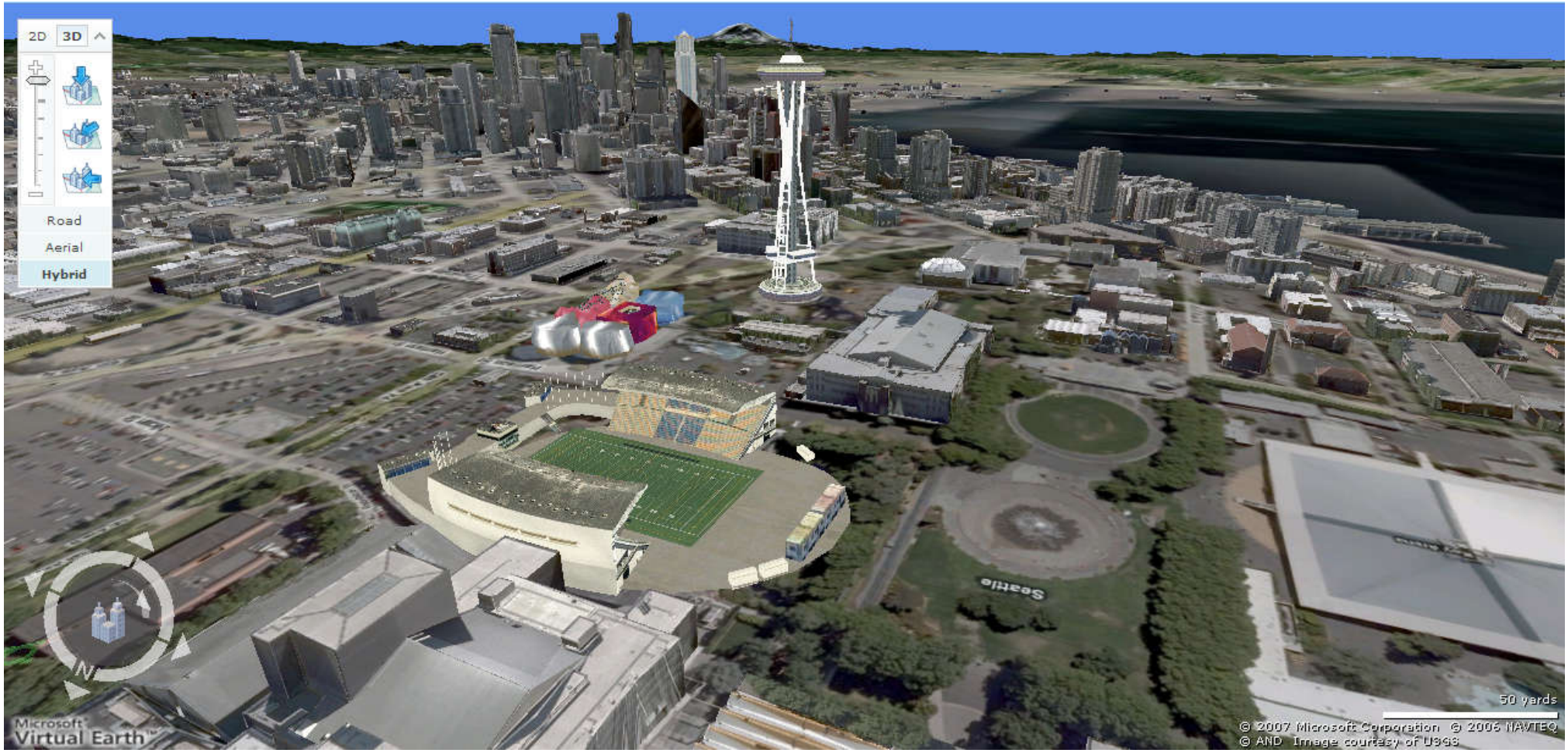
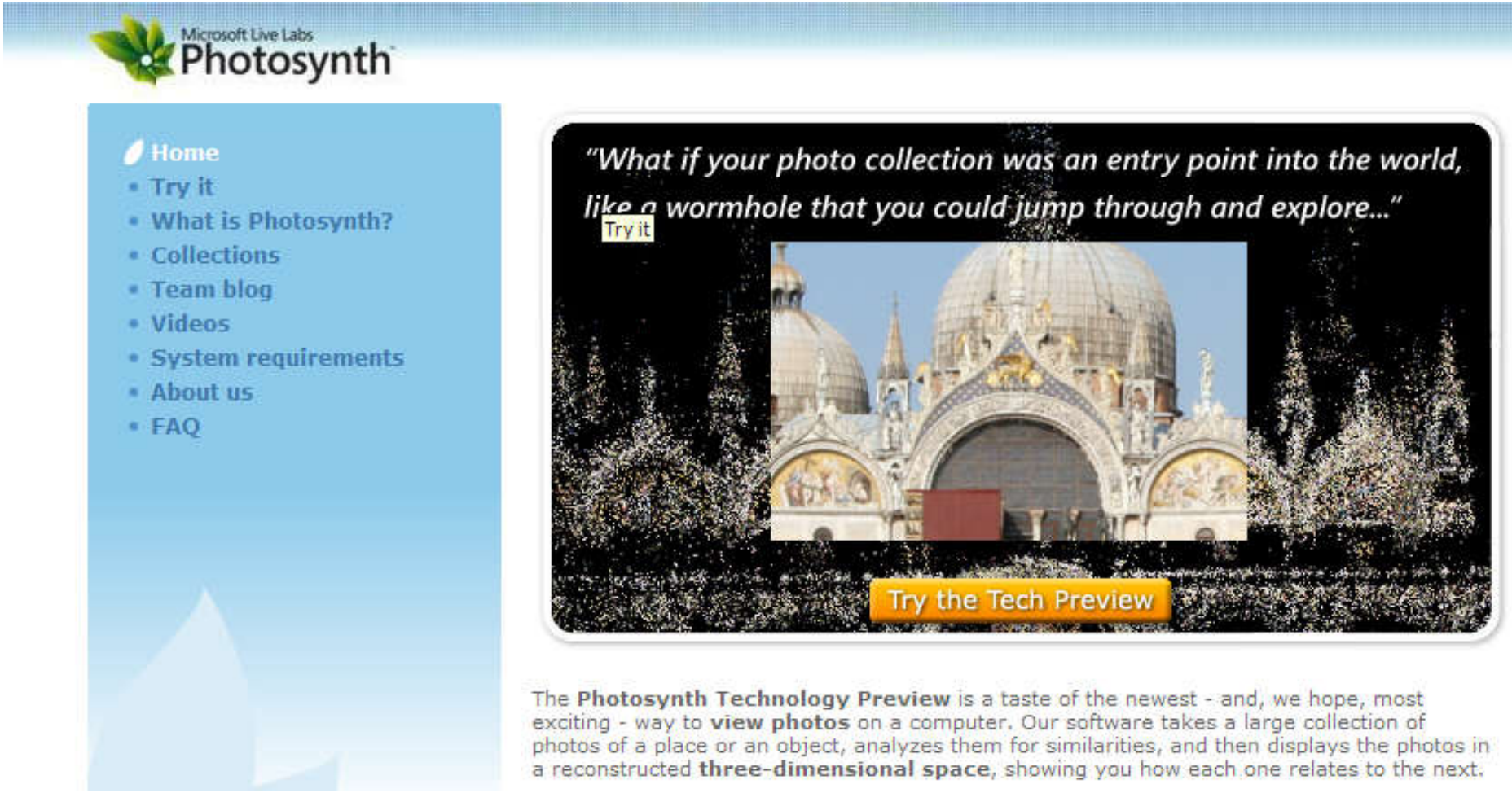


Image from Microsoft's Virtual Earth
(see also: Google Earth)

(Demo)

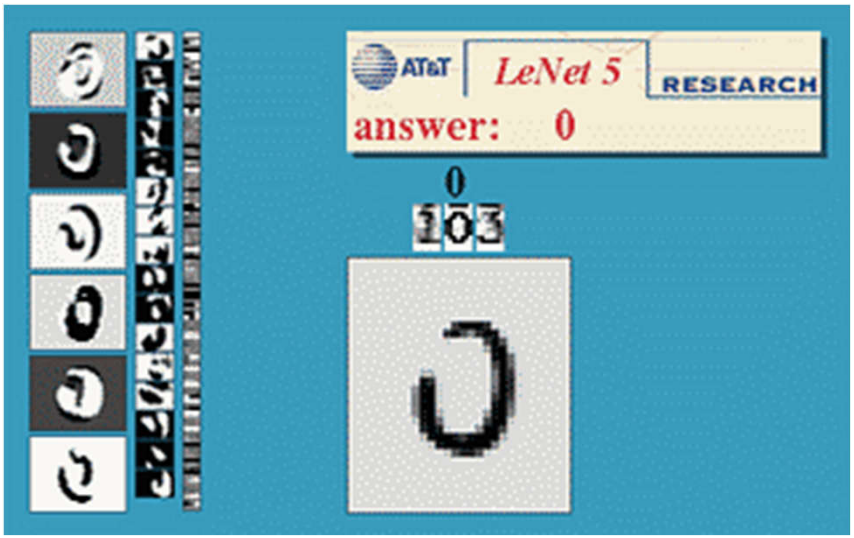


http://labs.live.com/photosynth/

Based on Photo Tourism technology developed here in CSE!
by Noah Snavely, Steve Seitz, and Rick Szeliski

Optical character recognition (OCR)

- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software



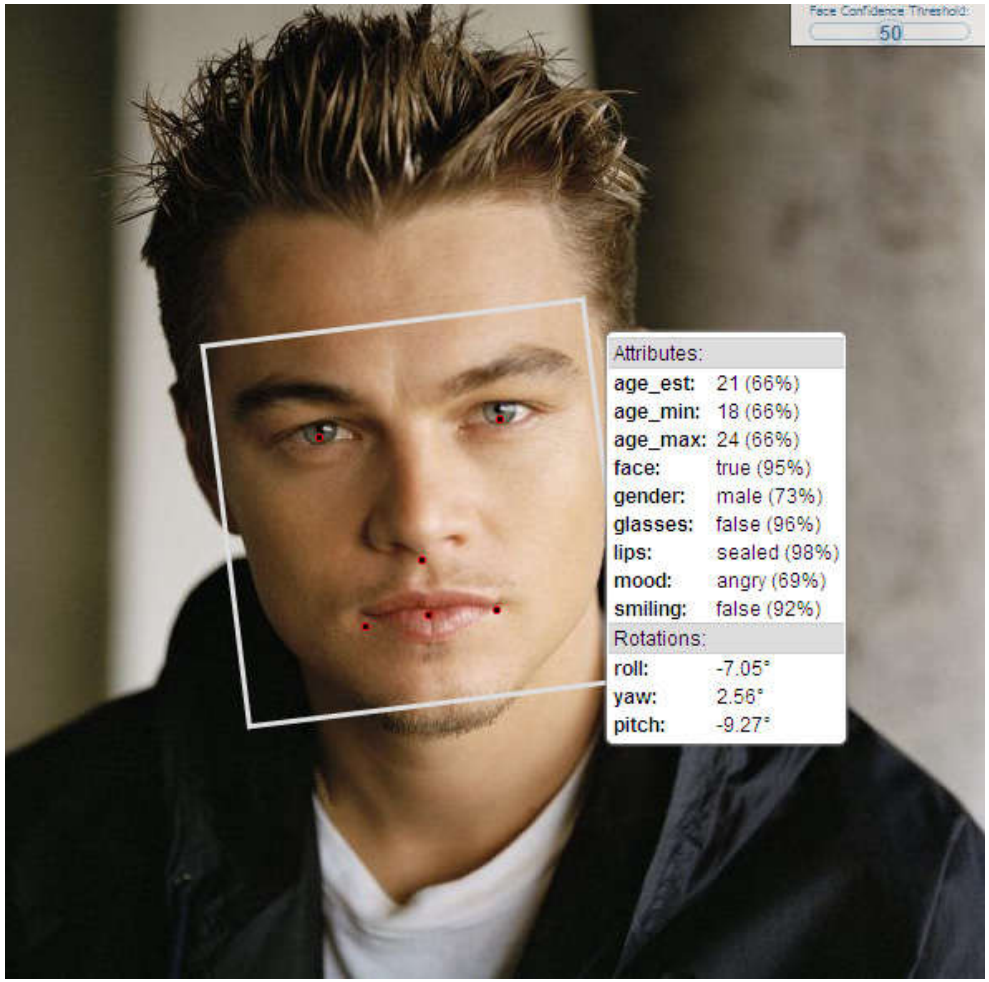
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

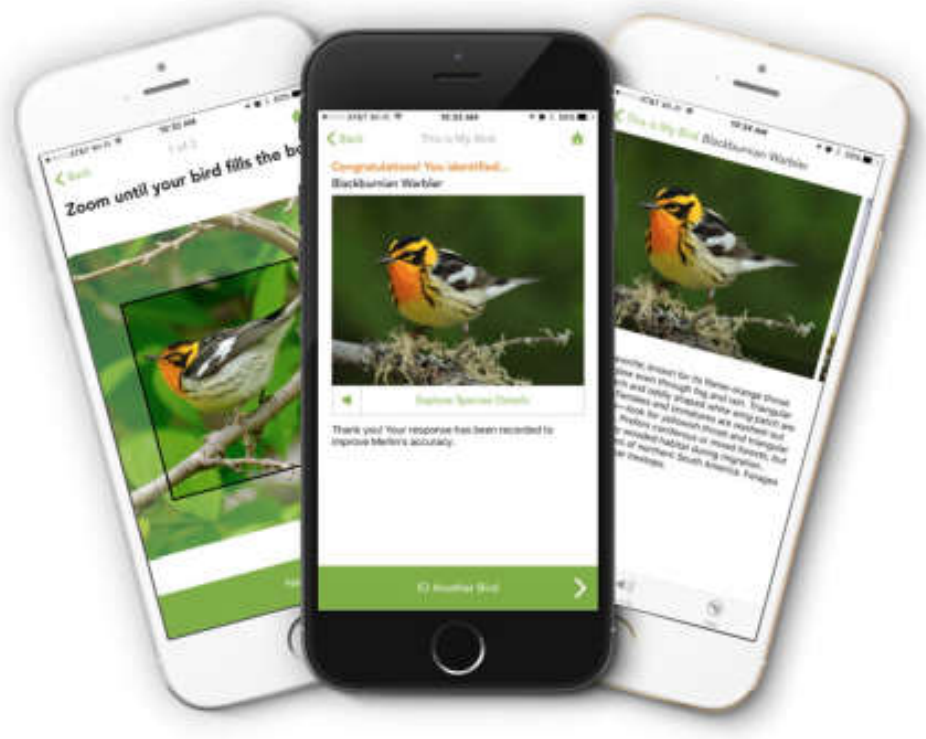
Face detection

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

Face analysis and recognition



Bird identification



Merlin Bird ID (based on Cornell Tech technology!)

Object recognition (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… “

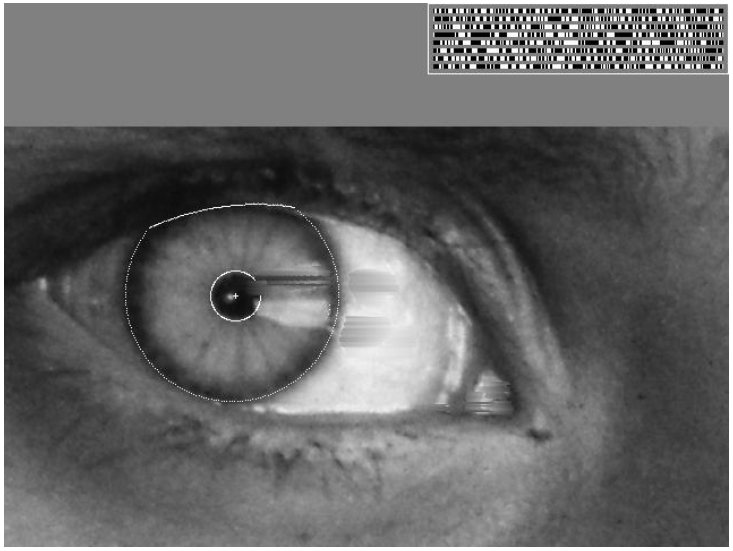
Face recognition

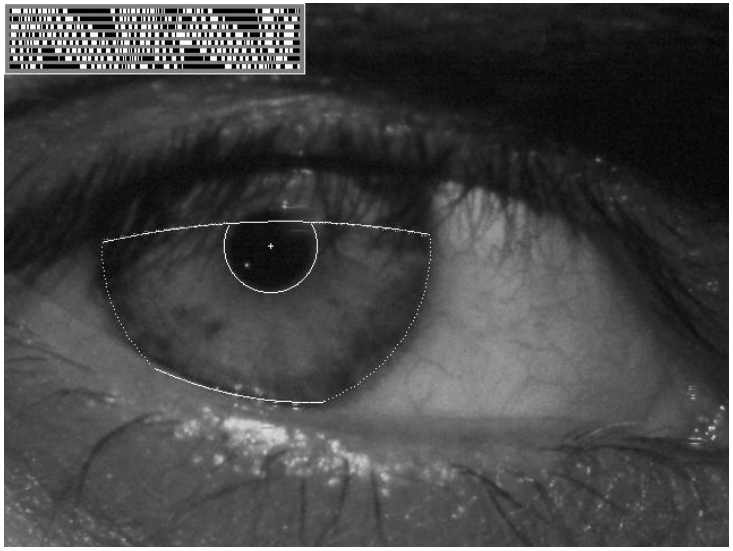
- Apple iPhoto software

https://www.apple.com/macos/photos/

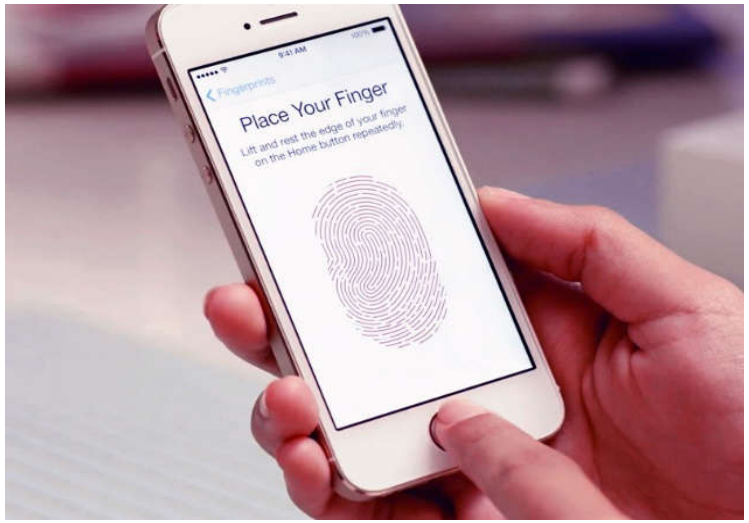
Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story





Login without a password



Fingerprint scanners on many new smartphones and other devices

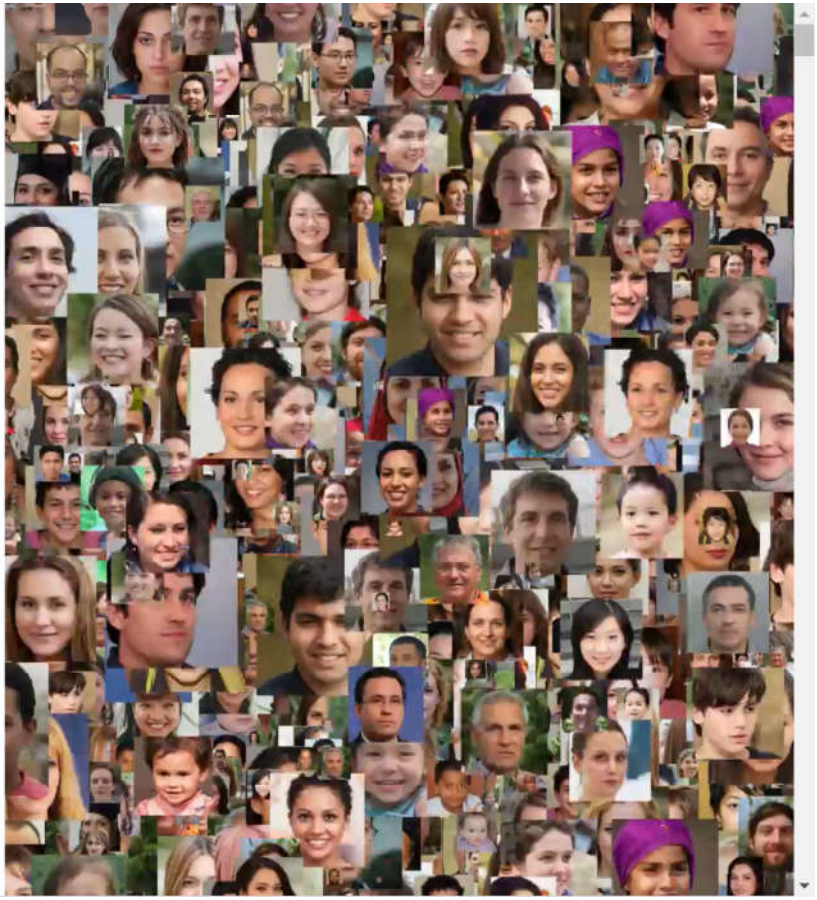
Face unlock on Apple iPhone X
See also http://www.sensiblevision.com/

The New York Times

Account

The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.



New York Times, Jan. 18, 2020

by Kashmir Hill

Researchers warn peace sign photos could expose fingerprints

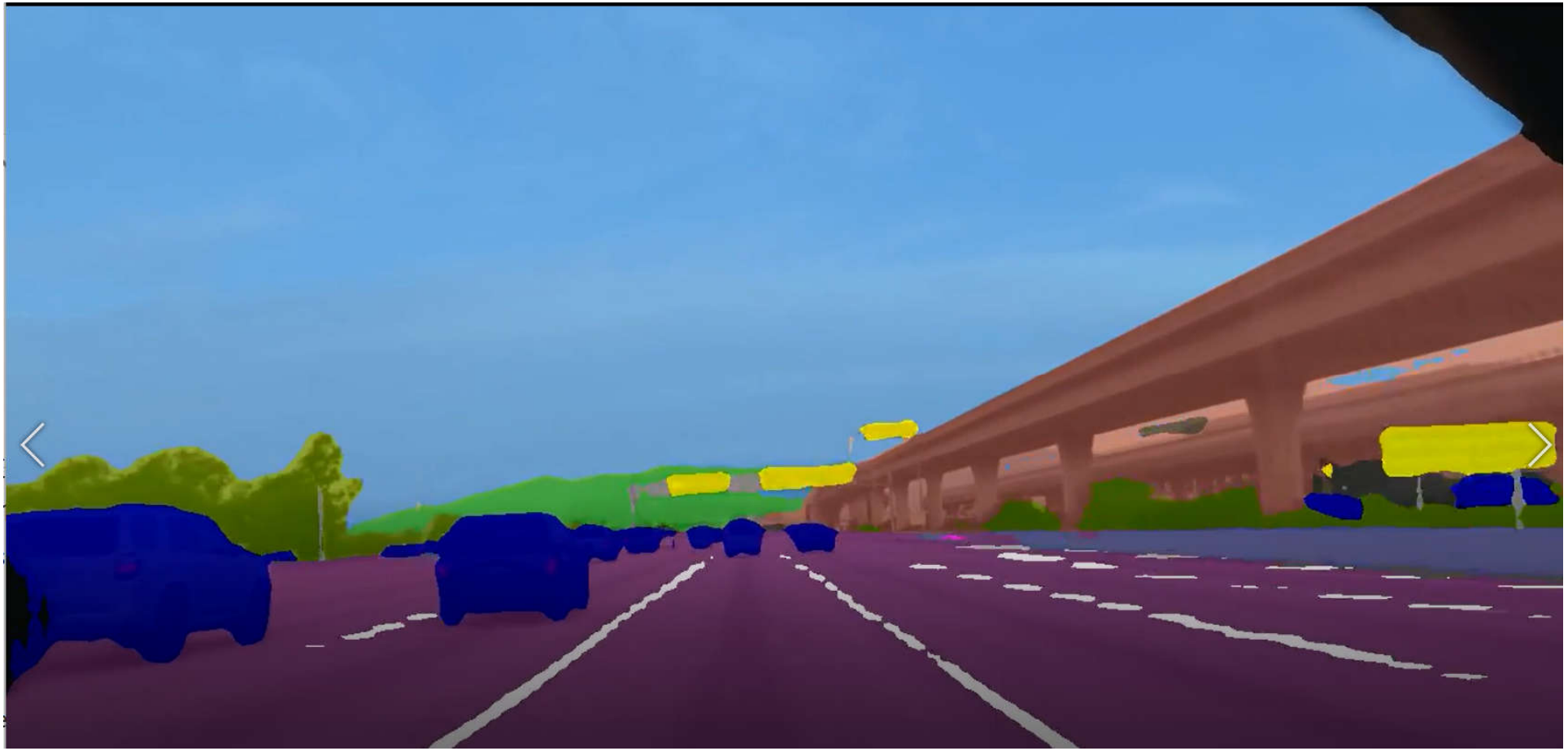
But the likelihood of anyone actually using images to recreate prints is pretty slim.

Jamie Rigg, @jmerigg
01.13.17 in Security

Getty

Semantic Segmentation

- （Demo）

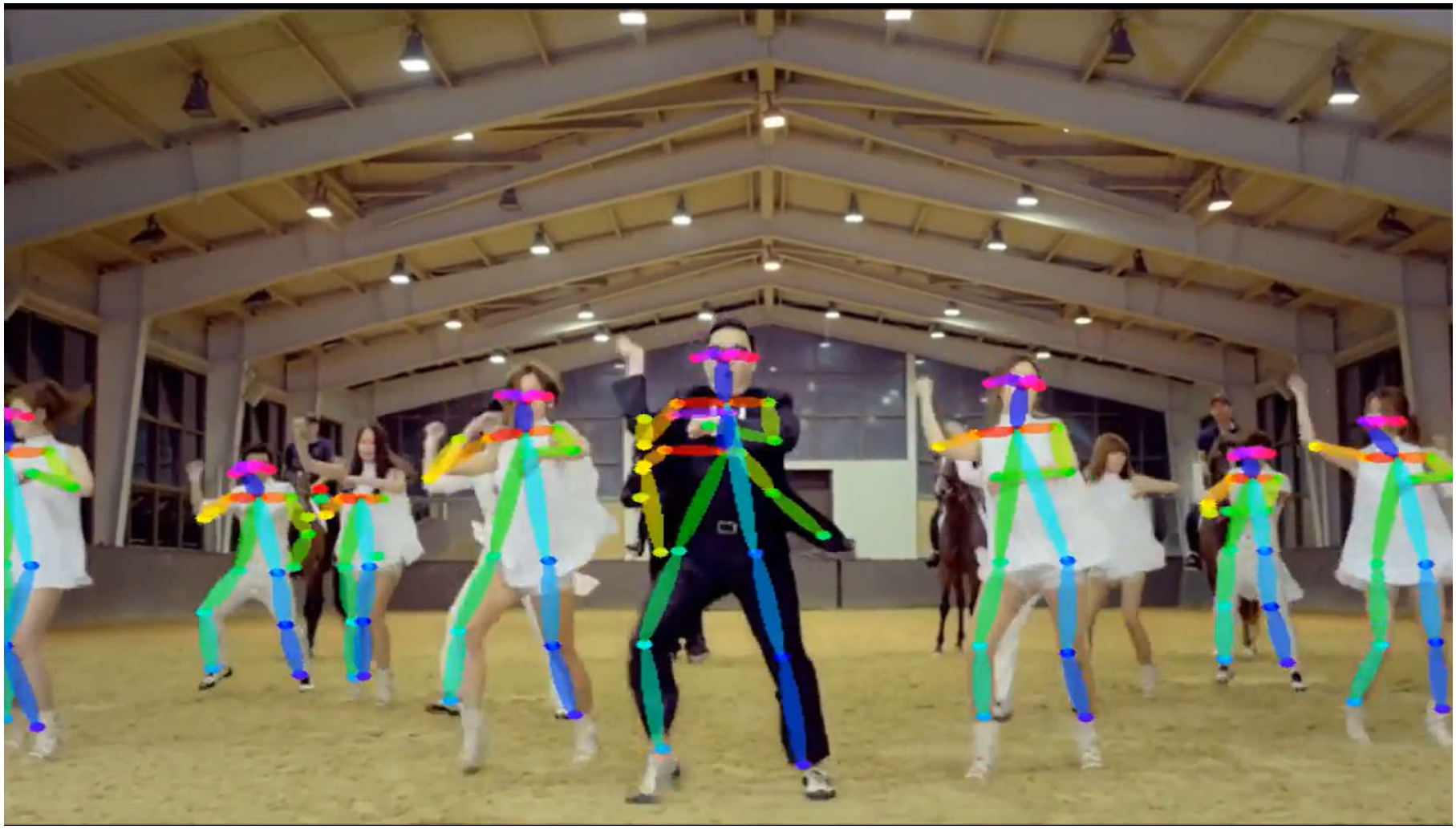


Instance Segmentation

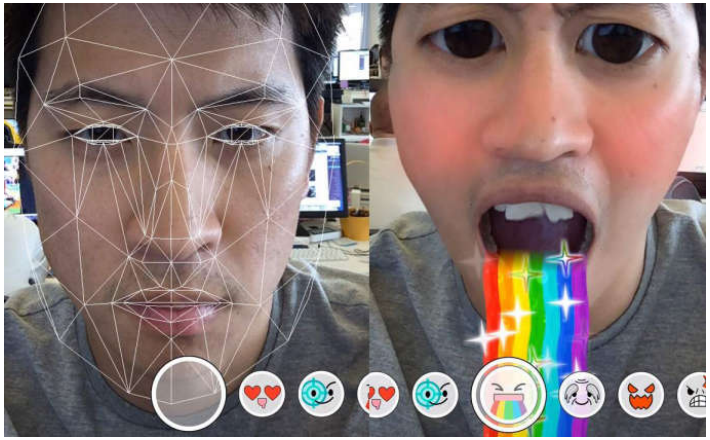
- （Demo）

Pose Estimation

- （Demo）



3D face tracking with consumer cameras



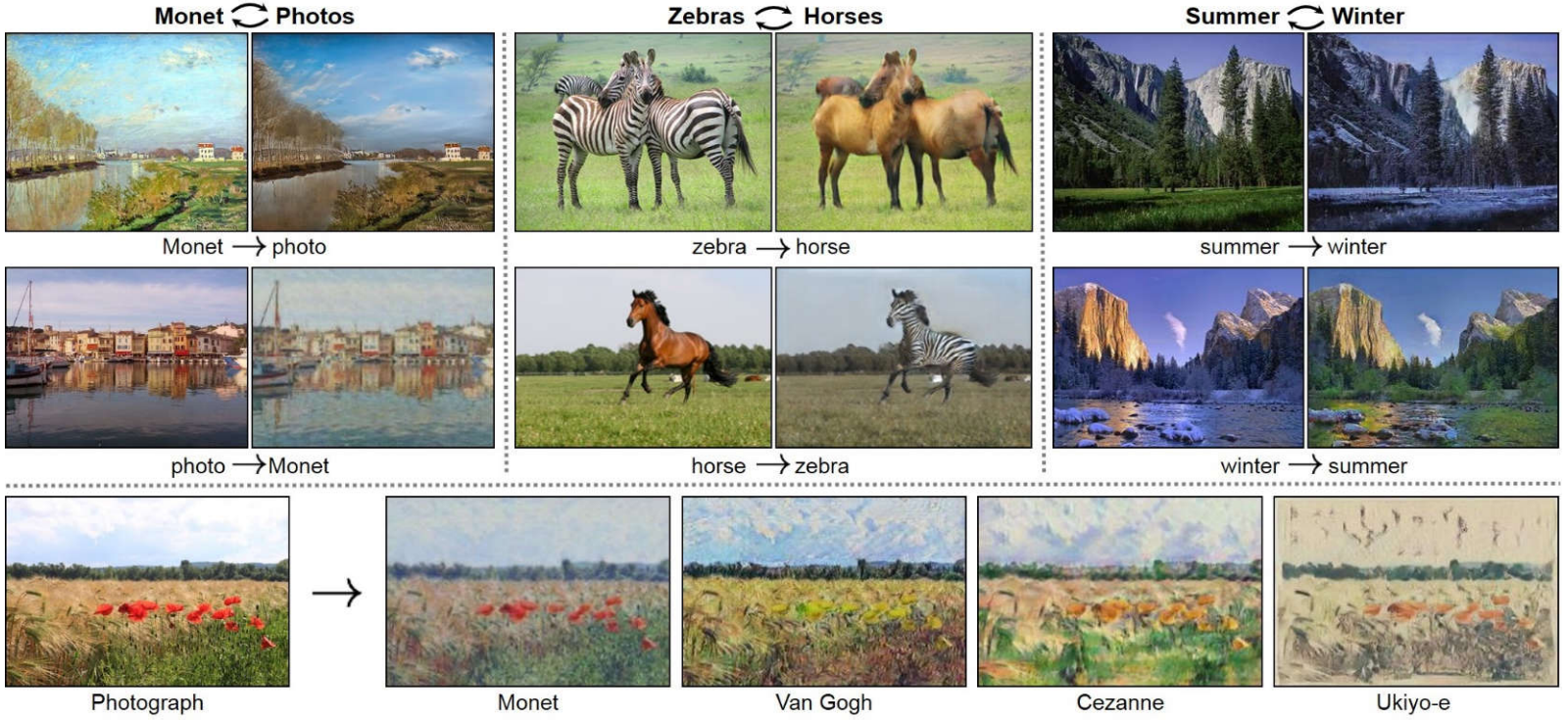
Snapchat Lenses

Face2Face system (Thies et al.) (demo)

Face synthesis

Karras, et al., *Progressive Growing of GANs for Improved Quality, Stability, and Variation*, ICLR 2018 (Demo)

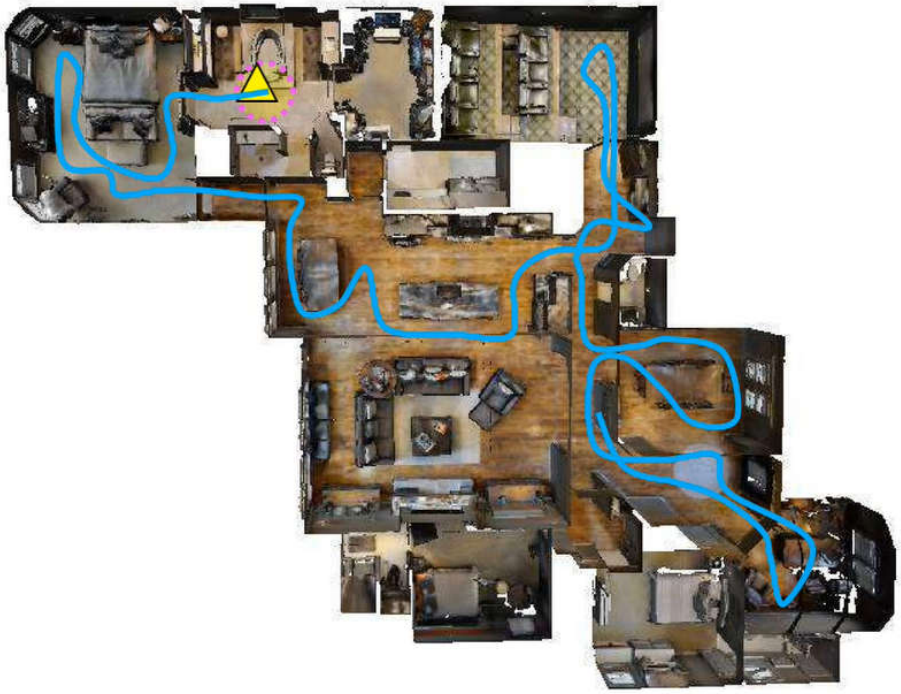
Image synthesis



Zhu, et al., *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*, ICCV 2017 (Demo)

Navigation

- （Demo）



Both trajectories are considered same in terms of the success signal.

Smart cars

- Mobileye(Demo)
  - Vision systems currently in high-end BMW, GM, Volvo models

Vision-based interaction (and games)

Digimask: put your face on a 3D avatar.

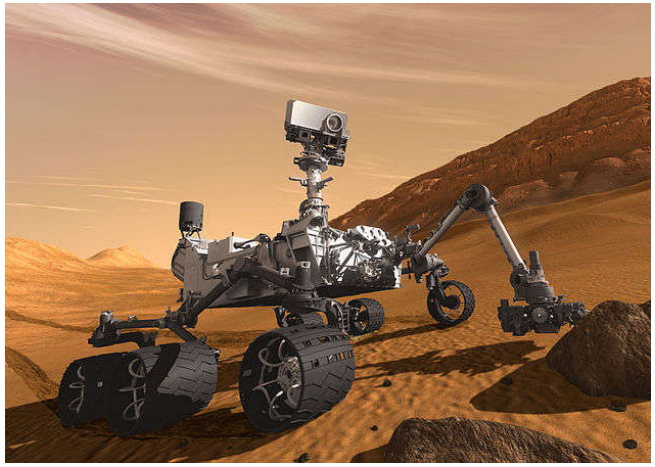
Leap, LeapMotion
(Demo)

Vision in space

- Vision systems (JPL) used for several tasks. (For more, read "Computer Vision on Mars" by Matthies et al.)
  - Panorama stitching
  - 3D terrain modeling
  - Obstacle detection, position tracking

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Robotics



NASA's Mars Curiosity Rover
https://en.wikipedia.org/wiki/Curiosity_(rover)

Amazon Picking Challenge
http://www.robocup2016.org/en/events/amazon-picking-challenge/

Amazon Prime Air

Amazon Scout

Movie

Movie

- Shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Source: S. Seitz

Movie

- Motion capture

Pirates of the Carribean, Industrial Light and Magic

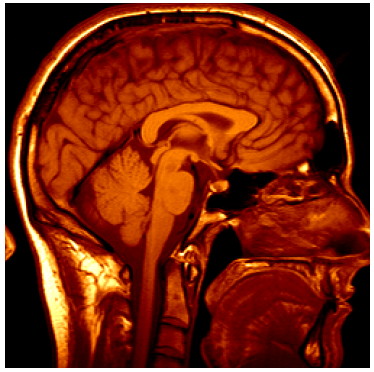
Source: S. Seitz

Sports

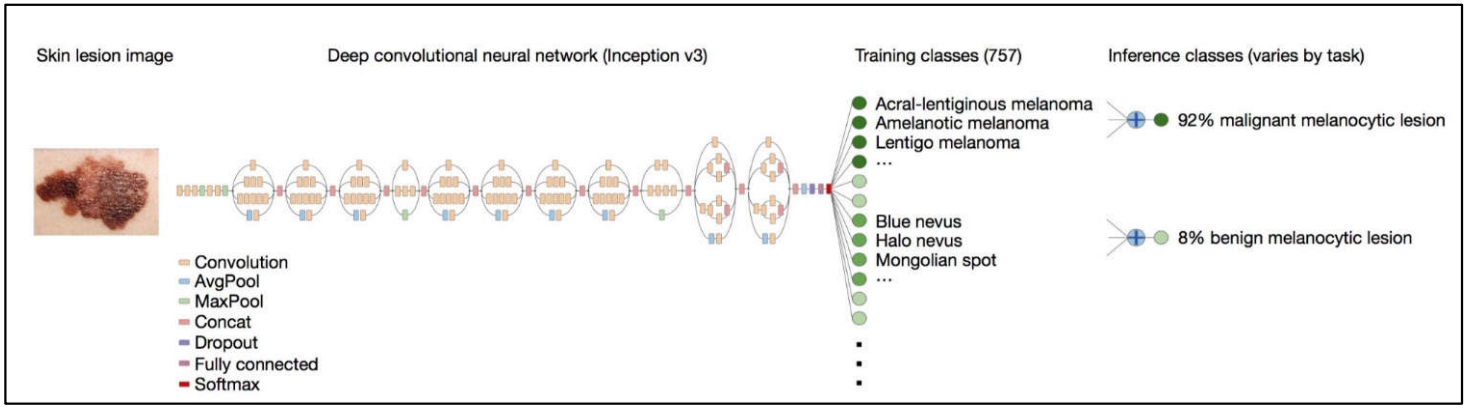
Sportvision first down line
Nice [explanation](#) on www.howstuffworks.com

Source: S. Seitz

Medical imaging



3D imaging
(MRI, CT)



Skin cancer classification with deep learning
https://cs.stanford.edu/people/esteva/nature/

Current state of the art

- You just saw many examples of current systems.
  - Many of these are less than 5 years old
- Computer vision is an active research area, and rapidly changing
  - Many new apps in the next 5 years
  - Deep learning powering many modern applications
- Many startups across a dizzying array of areas
  - Deep learning, robotics, autonomous vehicles, medical imaging, construction, inspection, VR/AR, …

Outline

- What is computer vision
- The goal of computer vision
- Applications of computer vision
- Challenges
- Course overview

Why is computer vision difficult?

Viewpoint variation

Illumination

Scale

Why is computer vision difficult?

Intra-class variation

Motion (Source: S. Lazebnik)

Background clutter

Occlusion

Challenges: local ambiguity

slide credit: Fei-Fei, Fergus & Torralba

But there are lots of visual cues we can use…

Source: S. Lazebnik

Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a given 2D image

  - We often must use prior knowledge about the world's structure

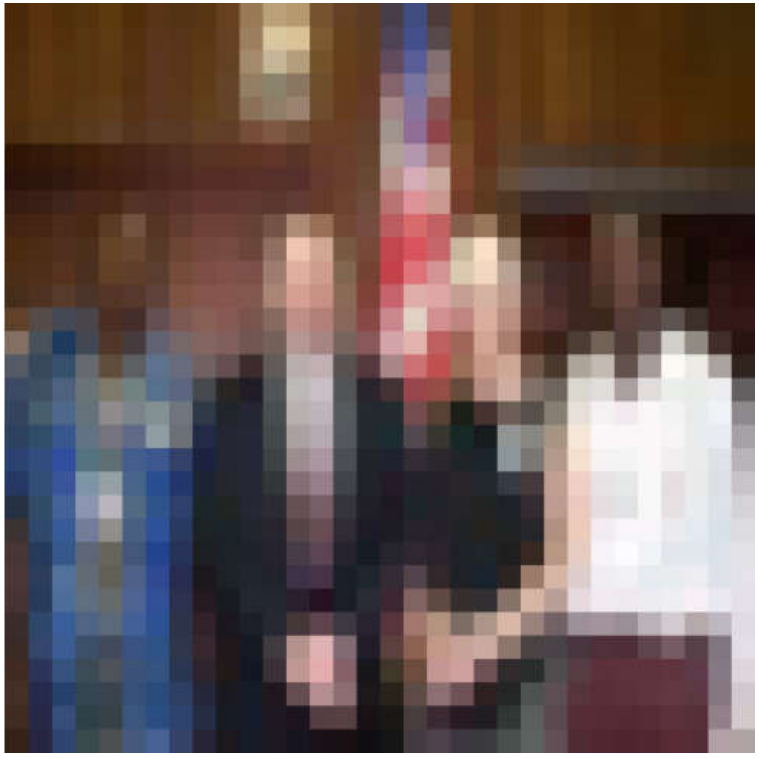
The state of computer vison and AI

- We are really, really far.

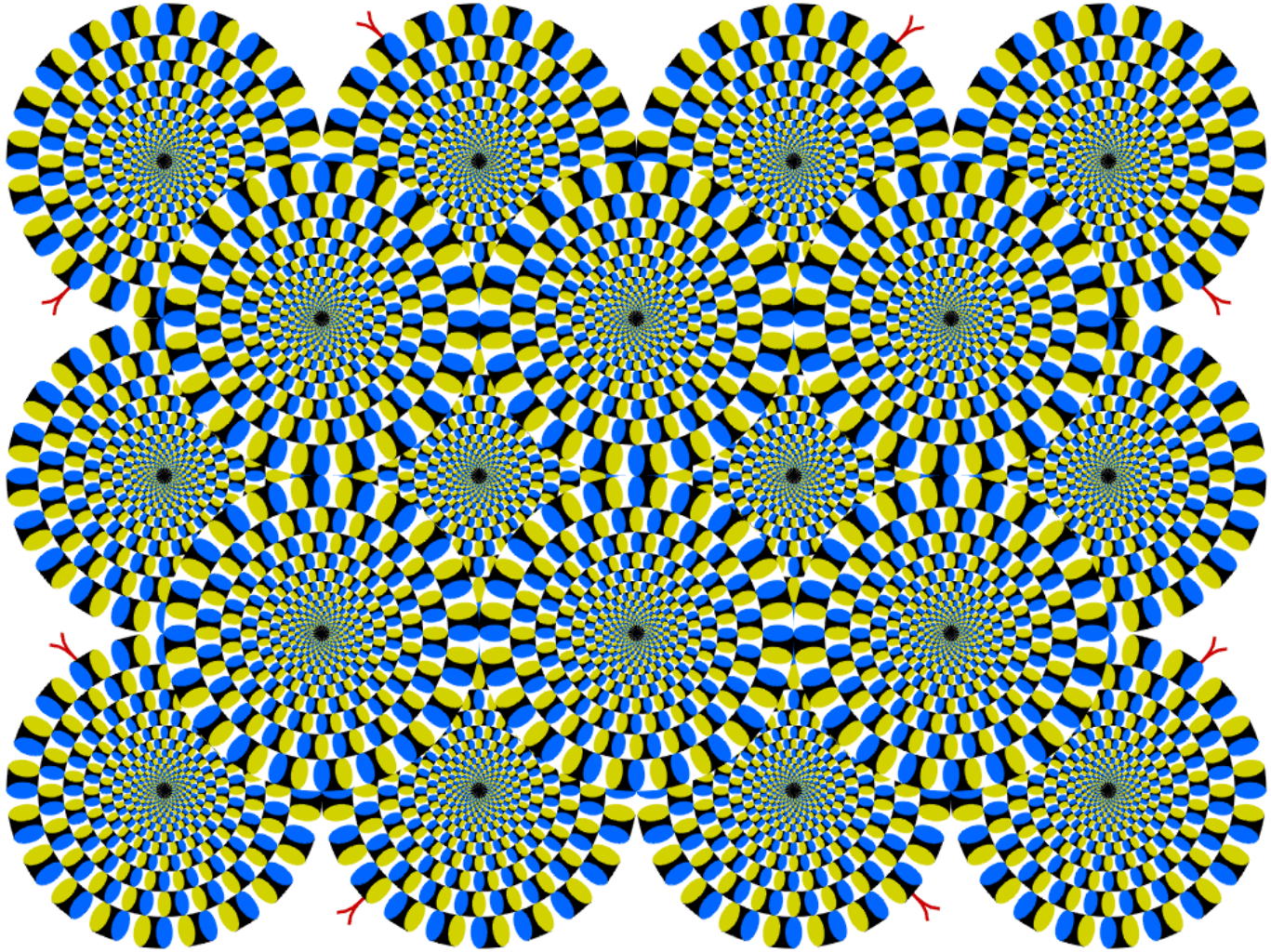
The picture above is funny.

The state of computer vison and AI

- humans can tell a lot about a scene from a little information…

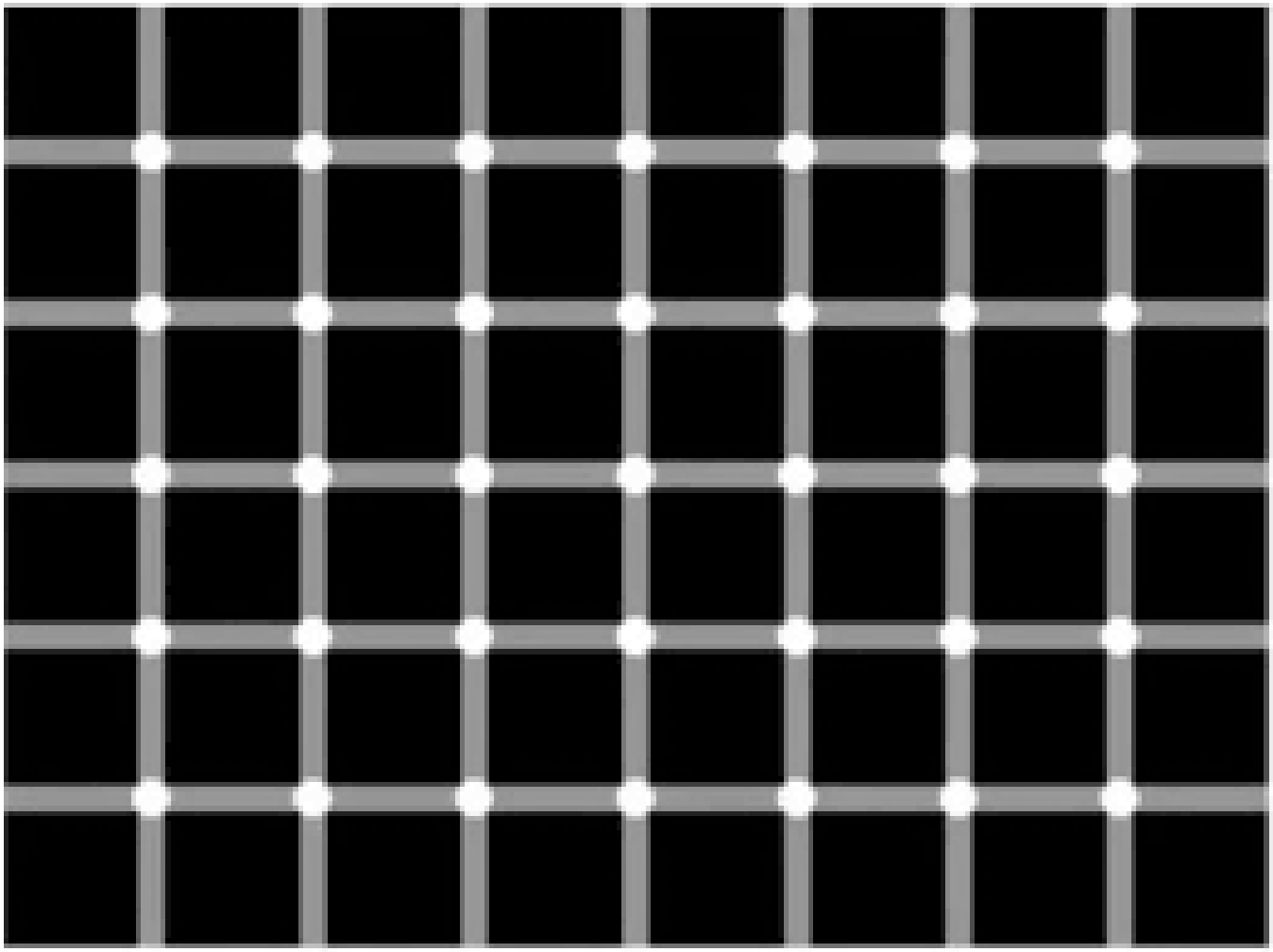


Optical illusions

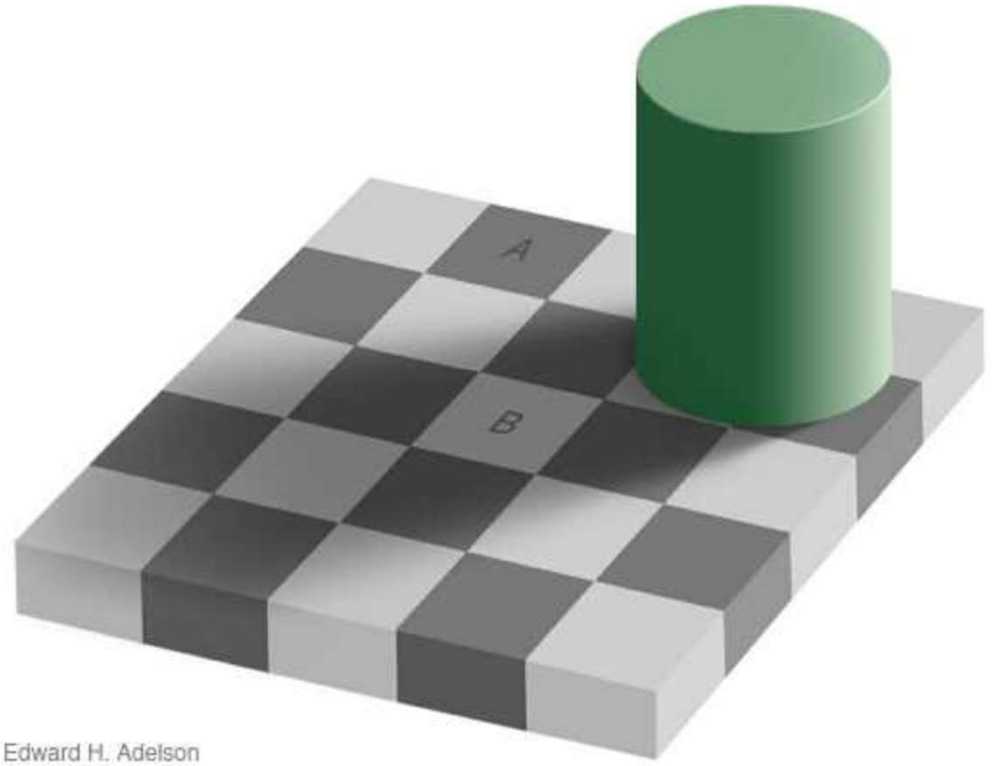


Copyright
A.Kitaoka 2003

Optical illusions

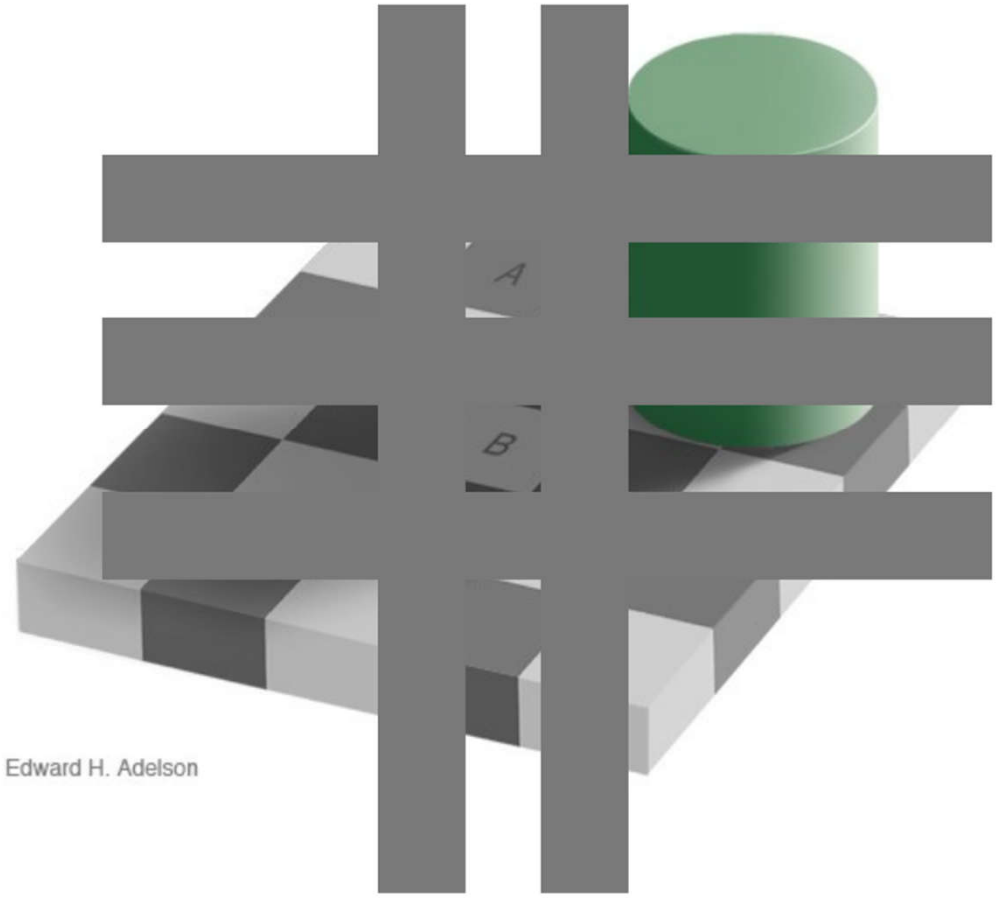


Optical illusions



Edward H. Adelson

Optical illusions

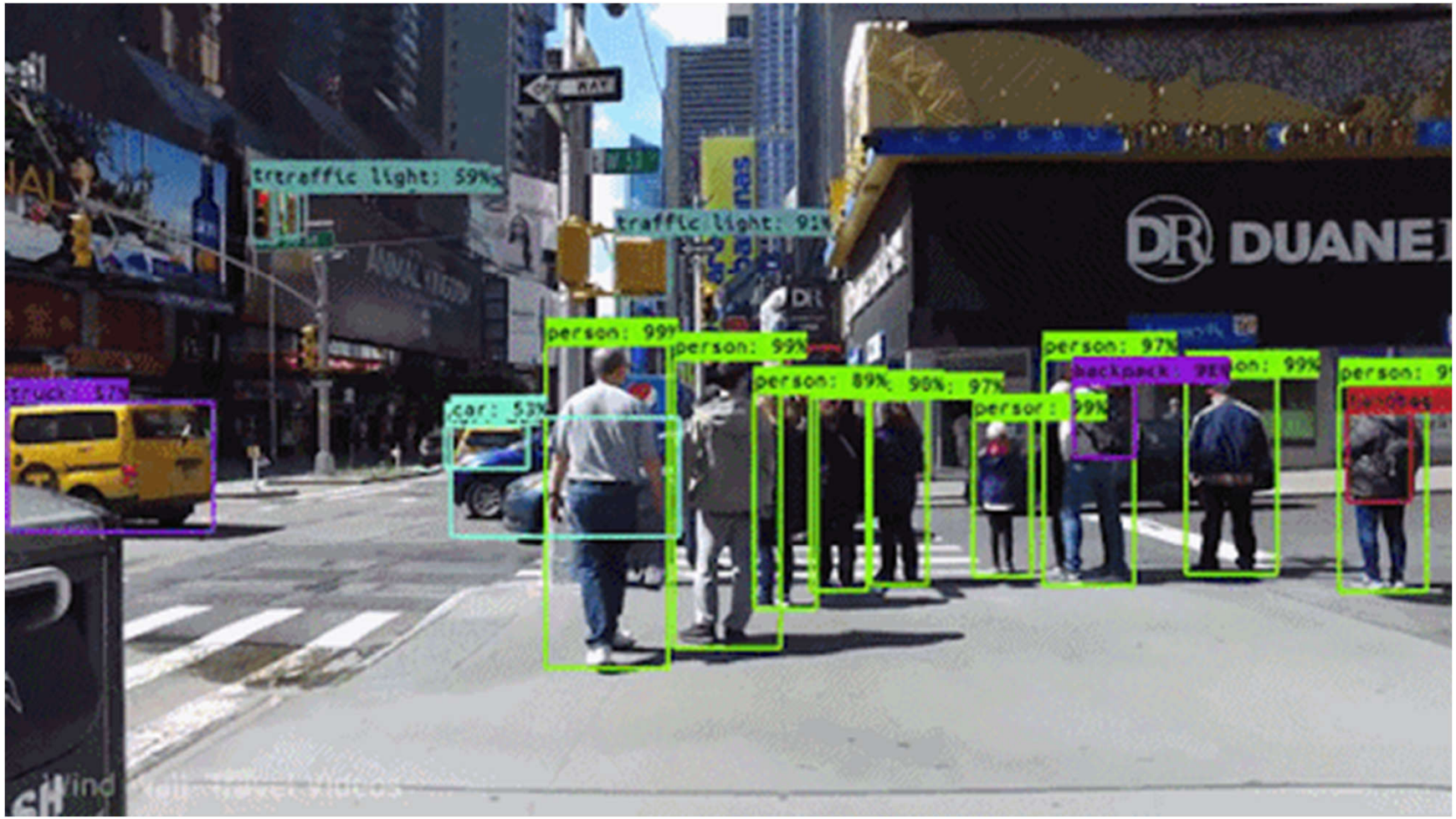


Important note

- In general, computer vision does not work

Important note

- Work except in certain situations/conditions



Outline

- What is computer vision
- The goal of computer vision
- Applications of computer vision
- Challenges
- Course overview

The goal of this course

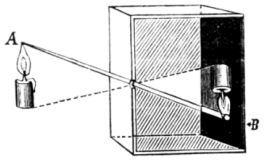
- Project-based course whose goal is to teach you the basics of computer vision – image processing, geometry, recognition – in a hands-on way

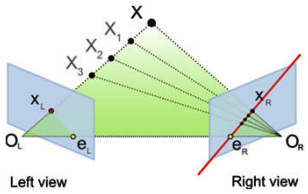
Course requirements

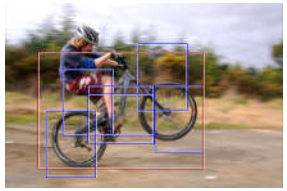
- Prerequisites
  - Data structures
  - Good working knowledge of Python programming
  - Linear algebra
  - Vector calculus

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

Course overview (tentative)







1. Low-level vision
   - image processing, edge detection, feature detection, cameras, image formation
2. Geometry and algorithms
   - projective geometry, stereo, structure from motion, optimization
3. Recognition
   - face detection / recognition, category recognition, segmentation

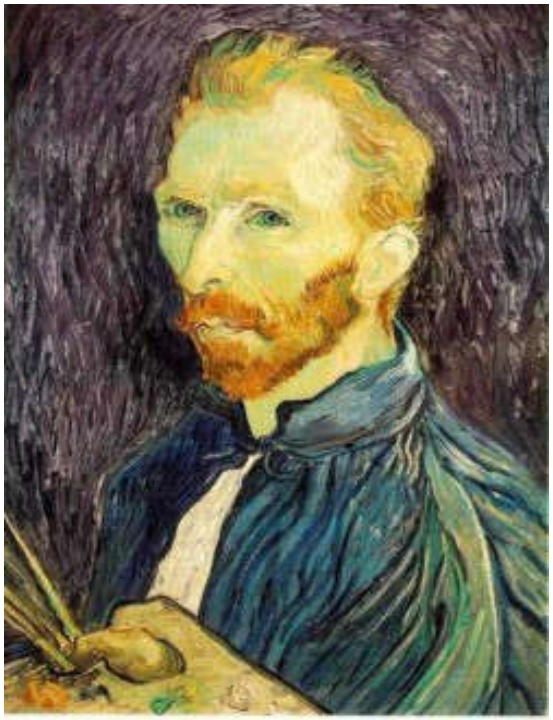
Grading policy

- Homework and participation in class
  - 20%-40% (30%)
- Term project
  - 40%-20% (30%)

    Proposal / report/ survey+Presentation
- Final Exam
  - 40%

Project1: Image Blending

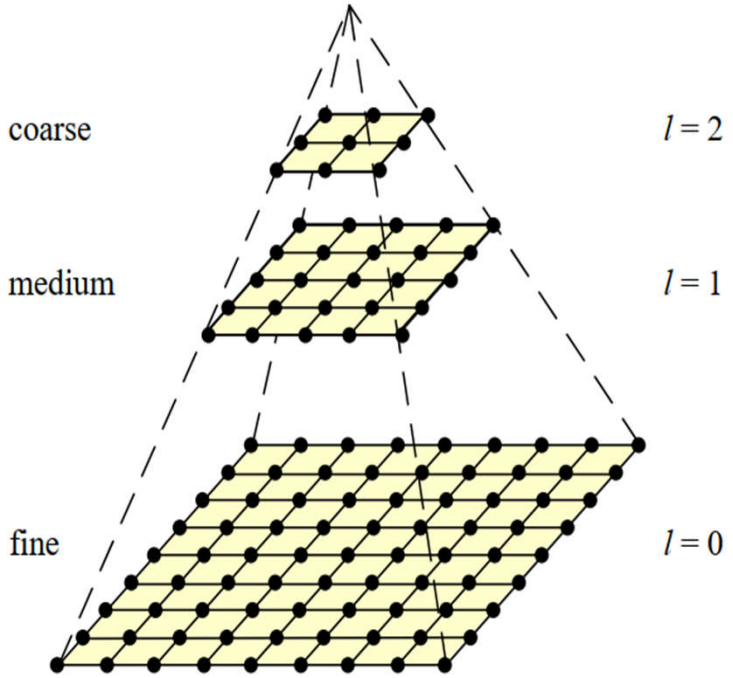
- Blending based on Laplacian pyramid



Gaussian 1/2

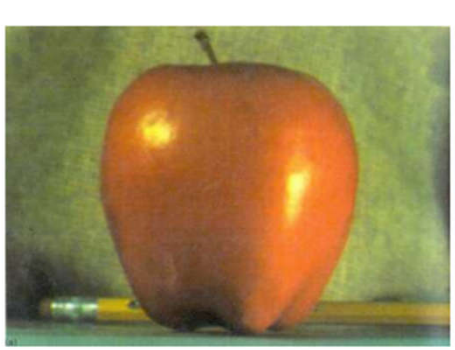
G 1/4

G 1/8

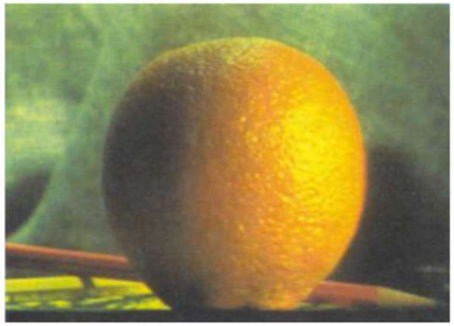


Project1: Image Blending

- Blending based on Laplacian pyramid



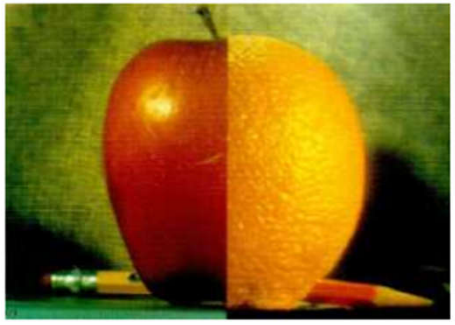
(a)



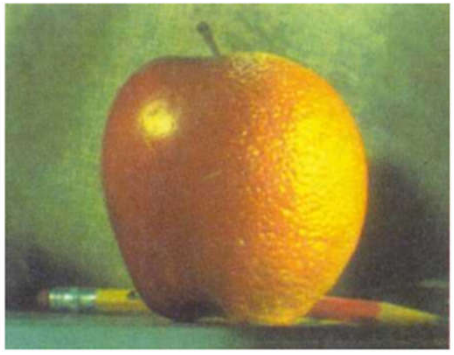
(b)

(e)

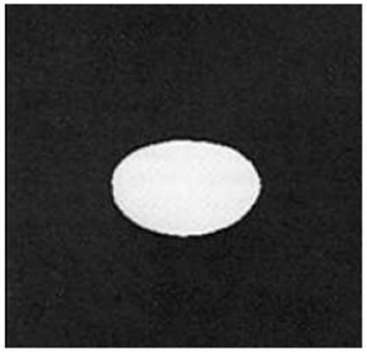
(f)



(c)



(d)

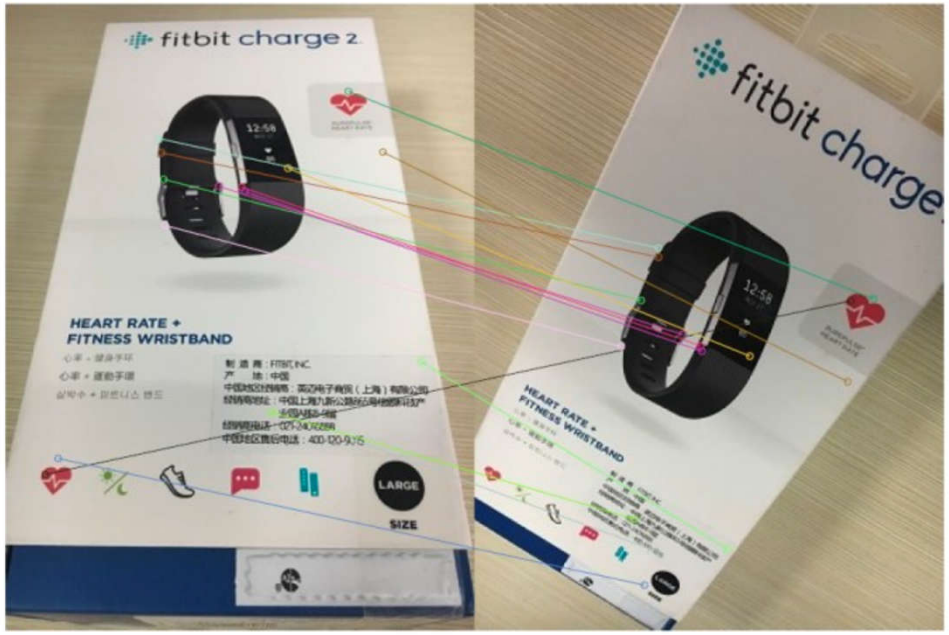


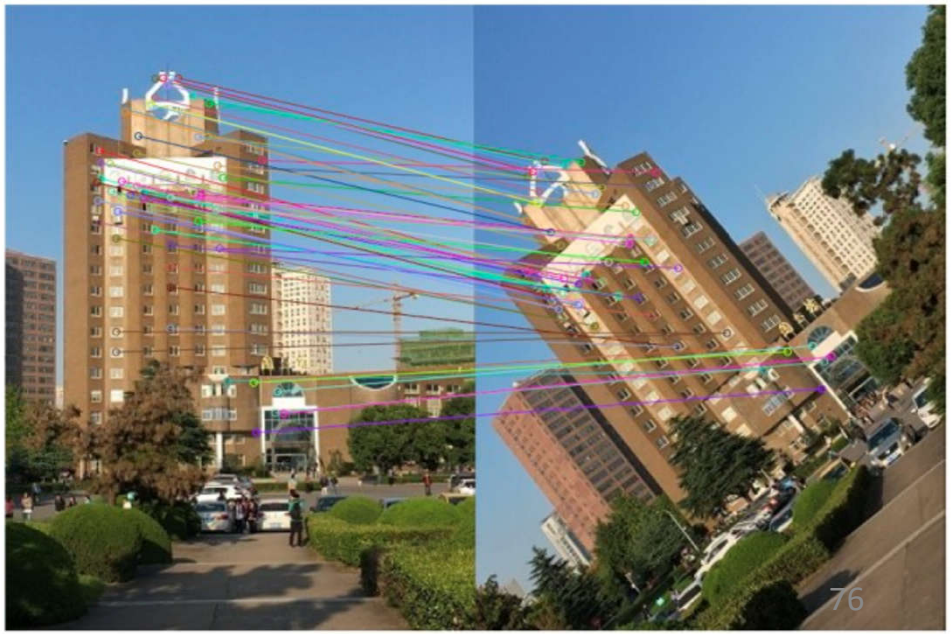
(g)

(h)

Project2: SIFT

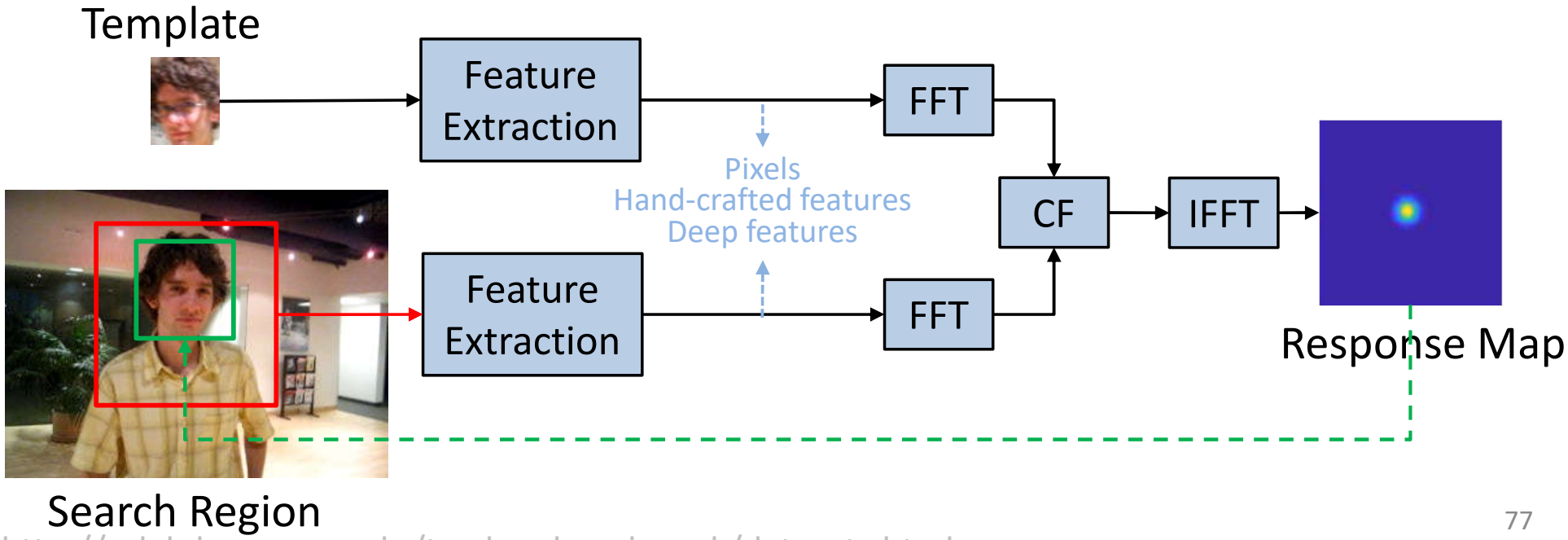
- Implement SIFT (detection and matching)
  - Python or MATLAB
  - CUDA is recommended (optional)
- Compare with OpenCV





Project3: Tracking

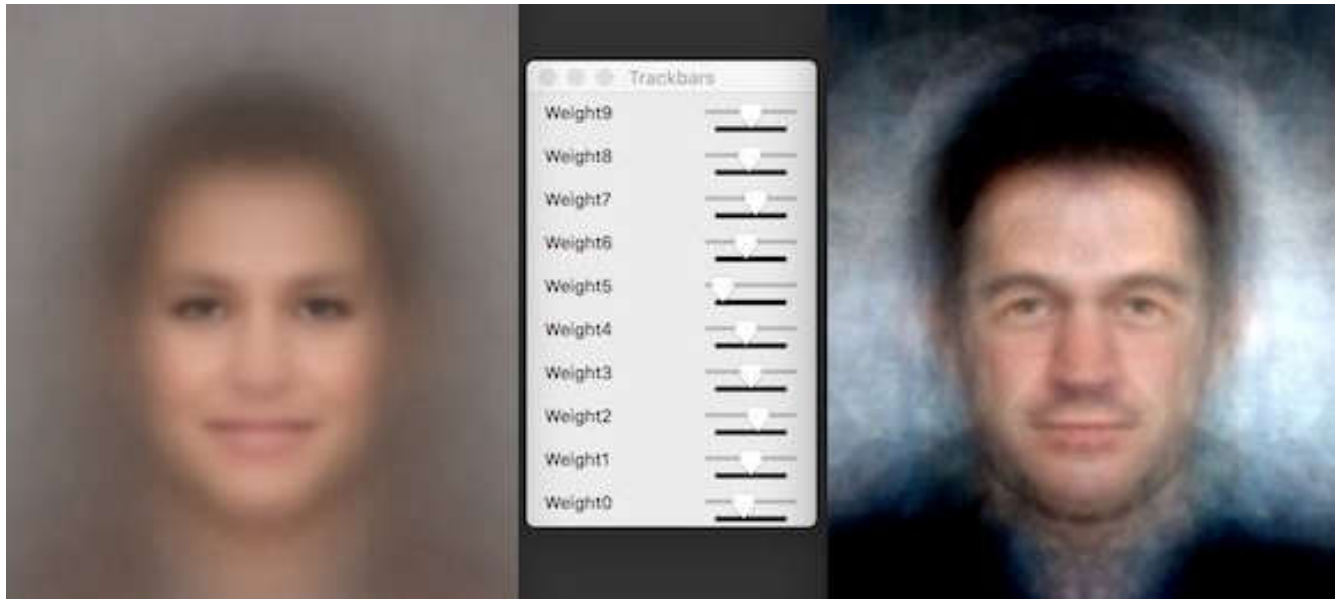
- Single object tracking based on Correlation Filters (CF). Different types of data can be used for CF:
  - Pixels, Hand-crafted features, Deep features



http://cvlab.hanyang.ac.kr/tracker_benchmark/datasets.html

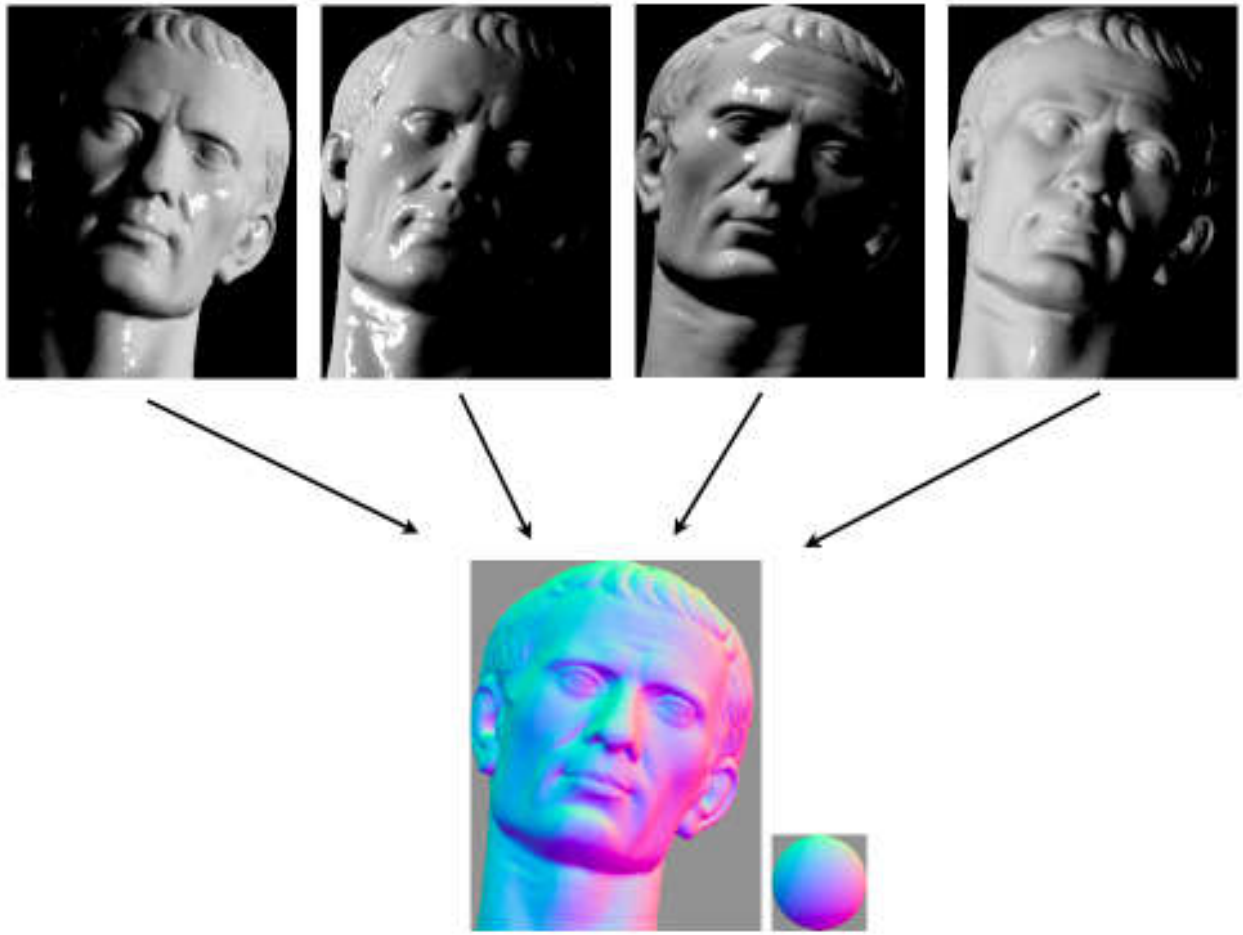
Project4: Eigen Face

- Eigen face
  - with dimensionality reduction (e.g. PCA)
  - for the recognition of different attributes (e.g. smiling vs. unsmiling, or different persons)

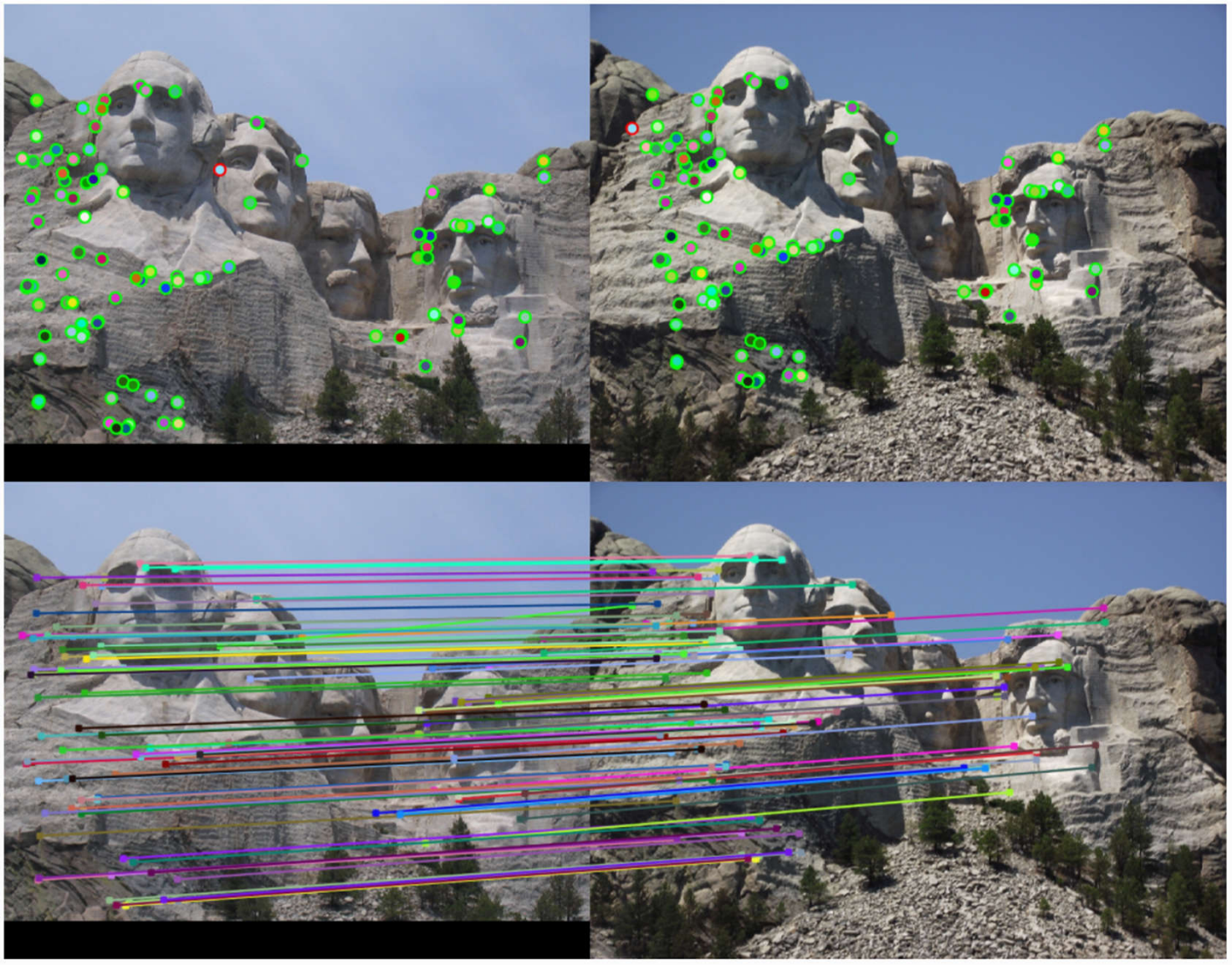


Project5: Photometric stereo

- 3D Reconstruction



Project6: Feature detection and matching



Project

- Image blending
- SIFT implementation
- Tracking (based on Correlation Filters)
- Eigen Face
- Photometric stereo
- Feature detection and matching
- Panorama stitching
- Corner detection
- Face recognition
- Image segmentation (Graphcut)
- Depth estimation (based on stereo Images)

Project

- Object detection
- Image segmentation
- Classification (clustering)
- Recognition (Face/object)
- Deep learning (Hinton)
- Generative adversarial net (GAN)
- variational autoencoder (VAE)
- Object Tracking
- Image Denoising/super-resolution……

Assignments

- Select 3 projects
  - Image blending
  - SIFT implementation
  - Tracking based on correlation filters
- Reproduce code for each project
- Submit code and report

Term Project

- Project: Systematic
- Regular paper/proposal: Novelty
- Presentation: Up to 15% (Extra Credit)
- Submit code, PPT, and paper (Chinese/English, Word/Latex)

Homework 1

- Send an email to liuqk3@163.com
- Title: CV-student ID+name+homework1

e.g.,CV-1120181100 张三 homework1

please attach a word file with the same name

- Content：1.个人简介，不少于100字

2.期望的课程学习目标

3.未来两年目标规划

Schedule

Date	Submission	Date	Submission
Sep. 7th		Nov.2ed	Project 3,Due
Sep.14ed	Homework1,Due	Nov.9th	
Sep.21th		Nov.16th	
Sep.28th		Nov.23rd	Term project, ppt
Oct.5th	Project 1,Due	Nov.30th	
Oct.12th		Dec.7th	Term project, Due
Oct.19th	Project2,Due	Dec.14th	
Oct.26th		Dec.21th	