Lecture 1
Introduction to Computer Vision

Lin ZHANG, PhD
School of Software Engineering,
Tongji University
Spring 2022

Course Info

Contact Information

Room 408L, Jishi Building
Email: cslinzhang@tongji.edu.cn

TA: Fengyi ZHANG, 2131507@tongji.edu.cn

Course information can be found at

http://cslinzhang.gitee.io/home/

Materials

- Major materials
  - My slides
- References
  - Some papers
  - Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision (2nd Edition), Cambridge University Press, 2004
  - D.A. Forsyth and J. Ponce, Computer Vision: A Modern Approach (2nd Edition), Pearson Education, Inc., 2013
  - Milan Sonka, Vaclav Hlavac, and Roger Boyle, Image Processing, Analysis, and Machine Vision, Thomson, 2008

Examination

- Homework 30%: 3 times, and each time 10%.
- Project 20%: 2 or 3 people for one group
- Final paper exam: 40%
- Attendance: 10%; Being absent >=1/3 lectures, you will fail this course
- Bonus 5%: being active in class and answering my questions correctly

Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview

What is vision?

"The plain man's answer (and Aristotle's too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is " *David Marr*, Vision 1982

David Marr (1945.1.19 – 1980.11.17), was a British neuroscientist and psychologist. The Marr Prize, one of the most prestigious awards in computer vision, is named in his honor.

What is computer vision?

To bridge the gap between pixels and “meaning”

What we see

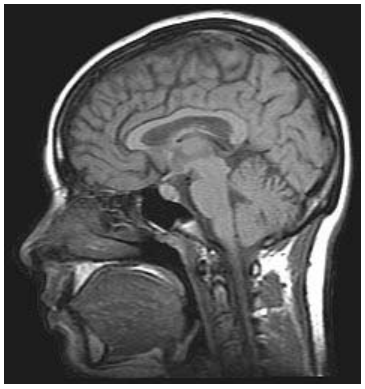
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

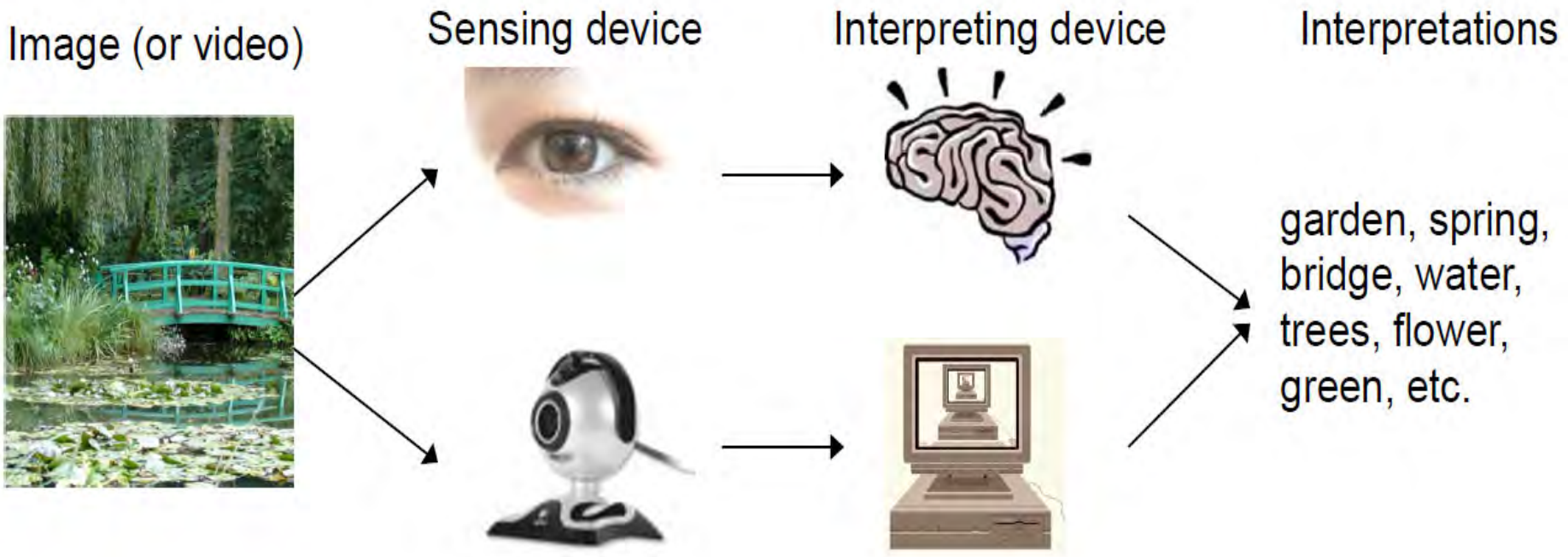
What is computer vision?

- Computer vision is the science and technology of machines that can see
- Concerned with the theory for building artificial systems that obtain information from images
- The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner



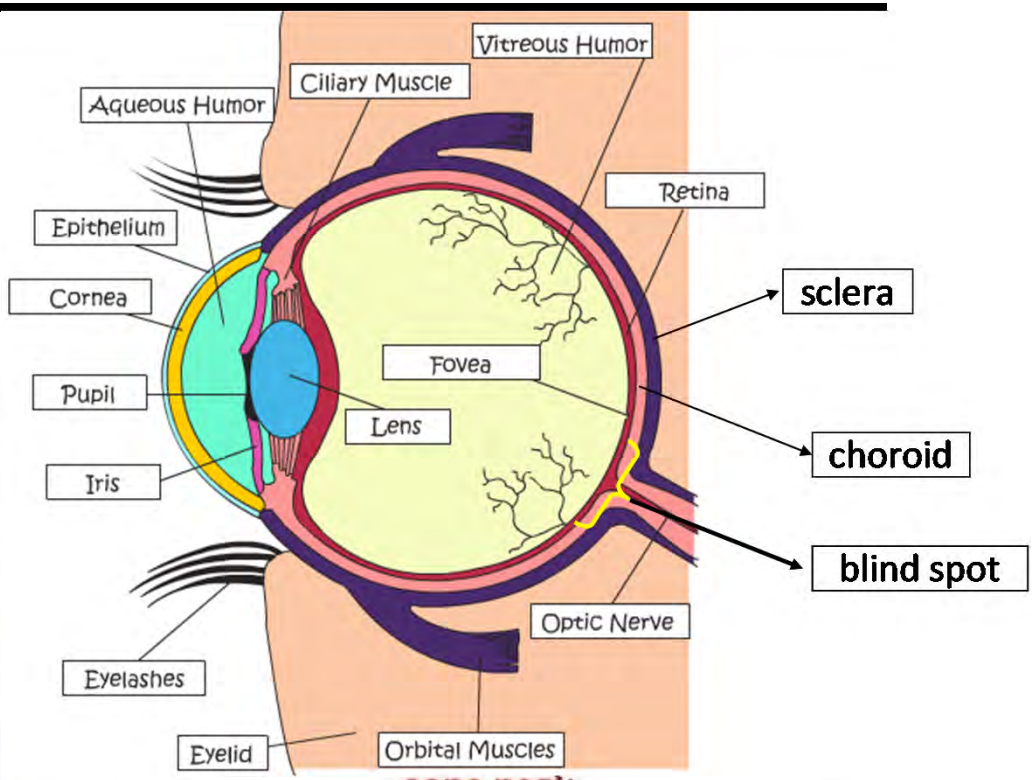


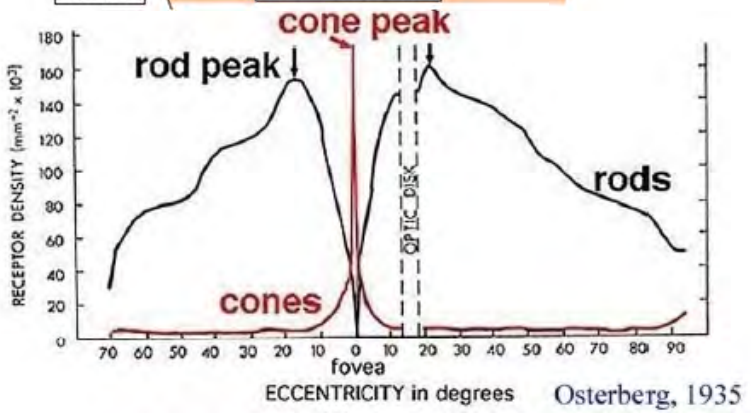
What is computer vision?



Human vision

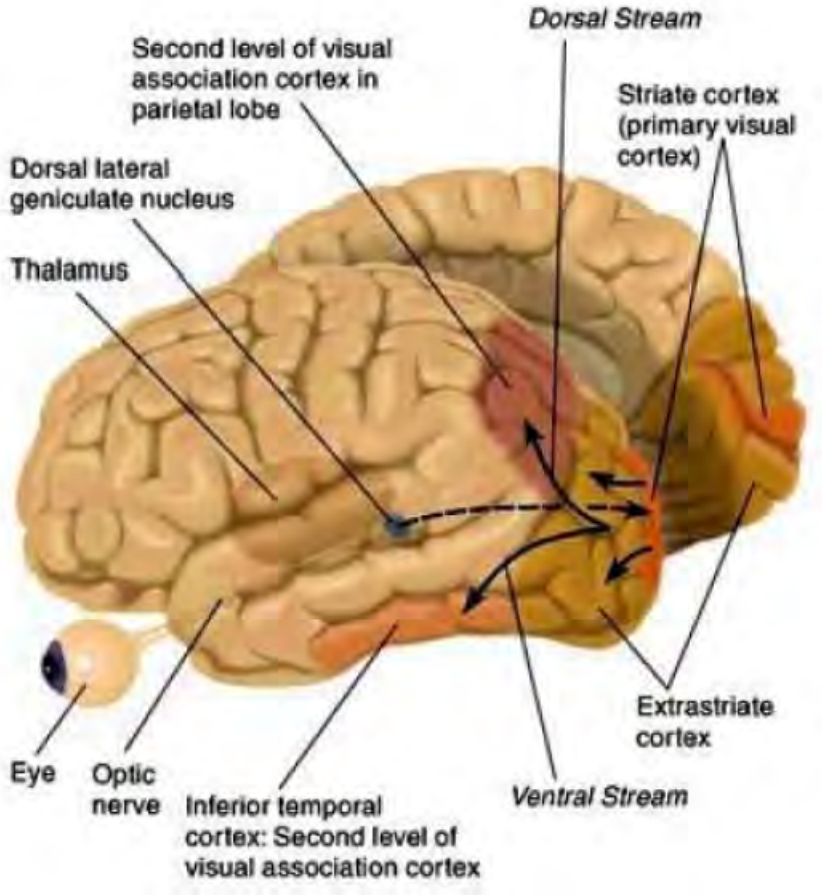
Our Eyes



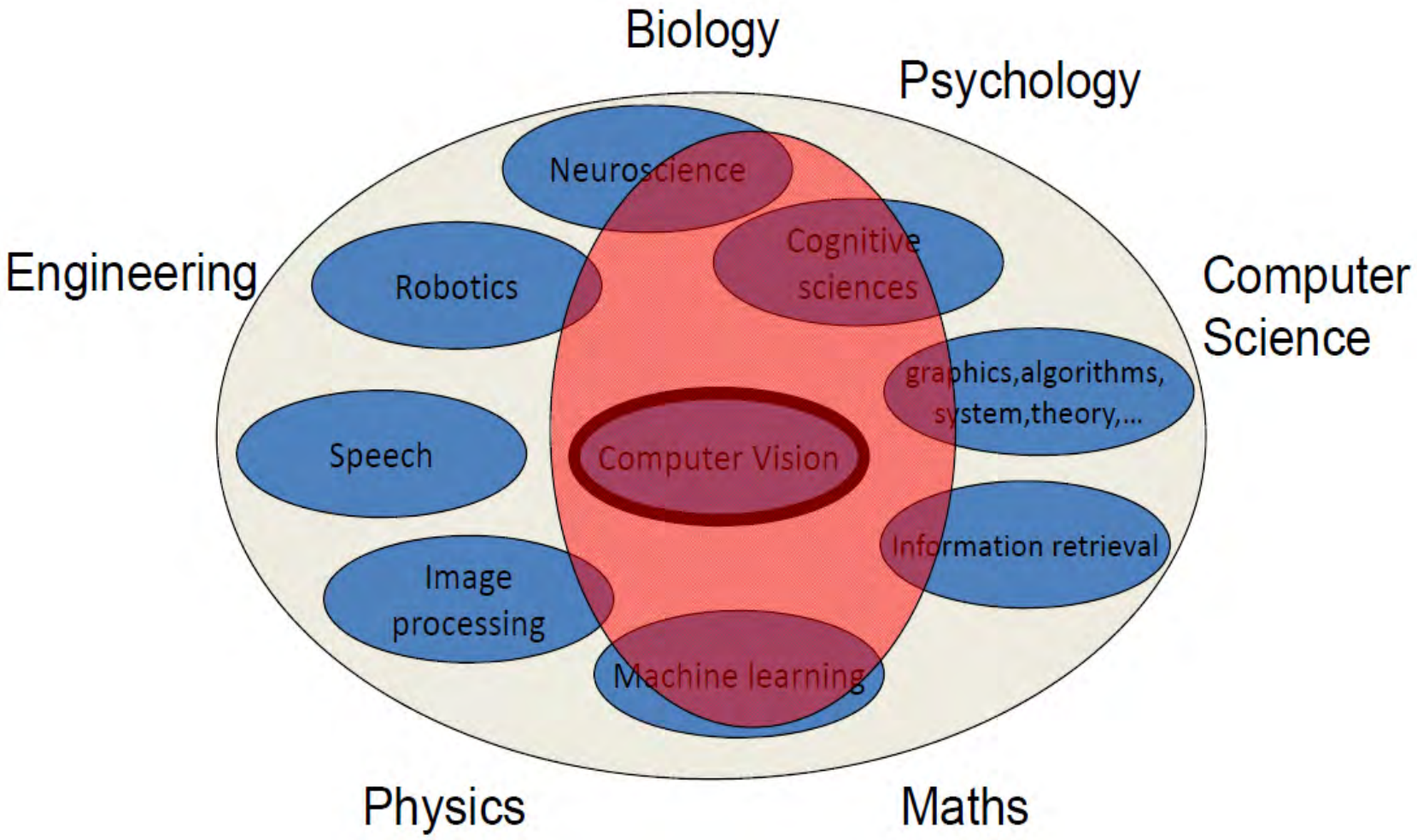


Human Visual System (HVS)

- Optical Receptors
  - Image formation
- Visual Pathway
  - Encoding
  - Representation
- Primary Visual Cortex
  - Interpretation



What is it related to?



Source: Feifei Li

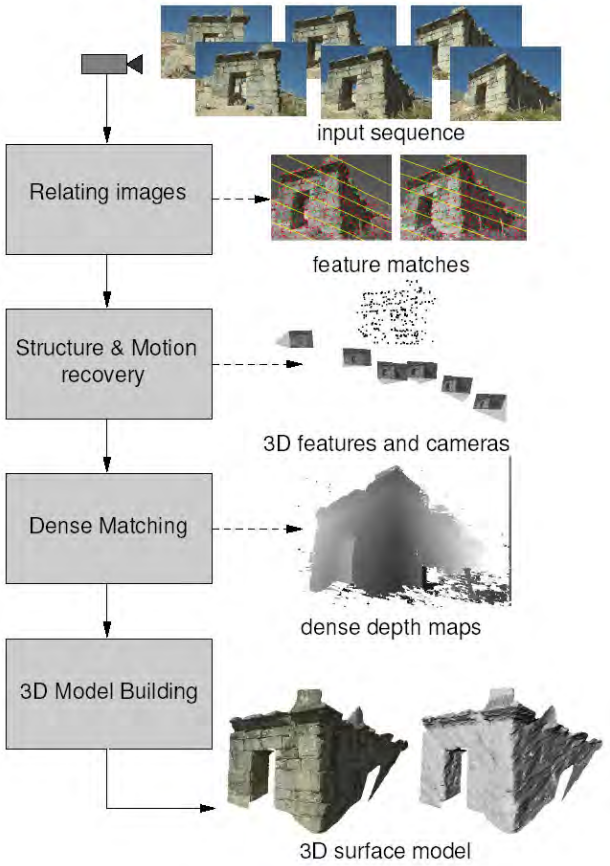
Vision as a measurement device

Real-time stereo

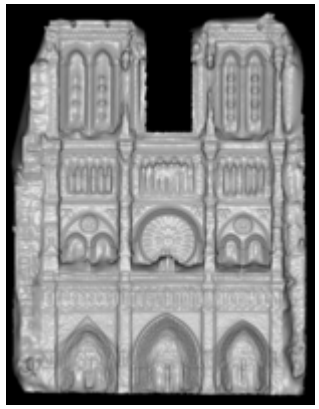
NASA Mars Rover

Pollefeys et al.

Structure from motion



Reconstruction from Internet photo collections



Goesele et al.

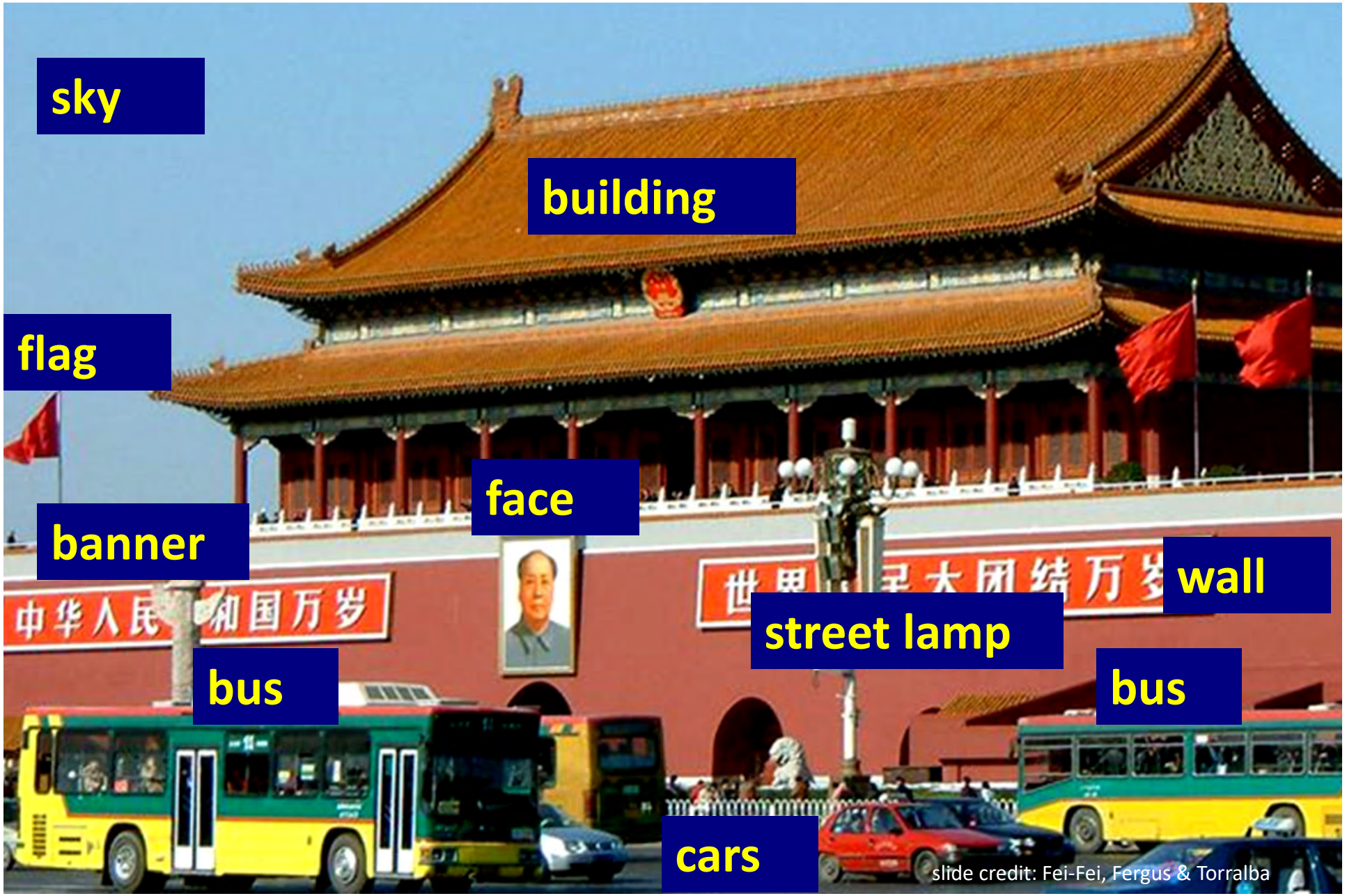
Vision as a measurement device

3D Reconstruction using RGB-D input (C. Guo, L. Zhang et al, ICASSP 2022)

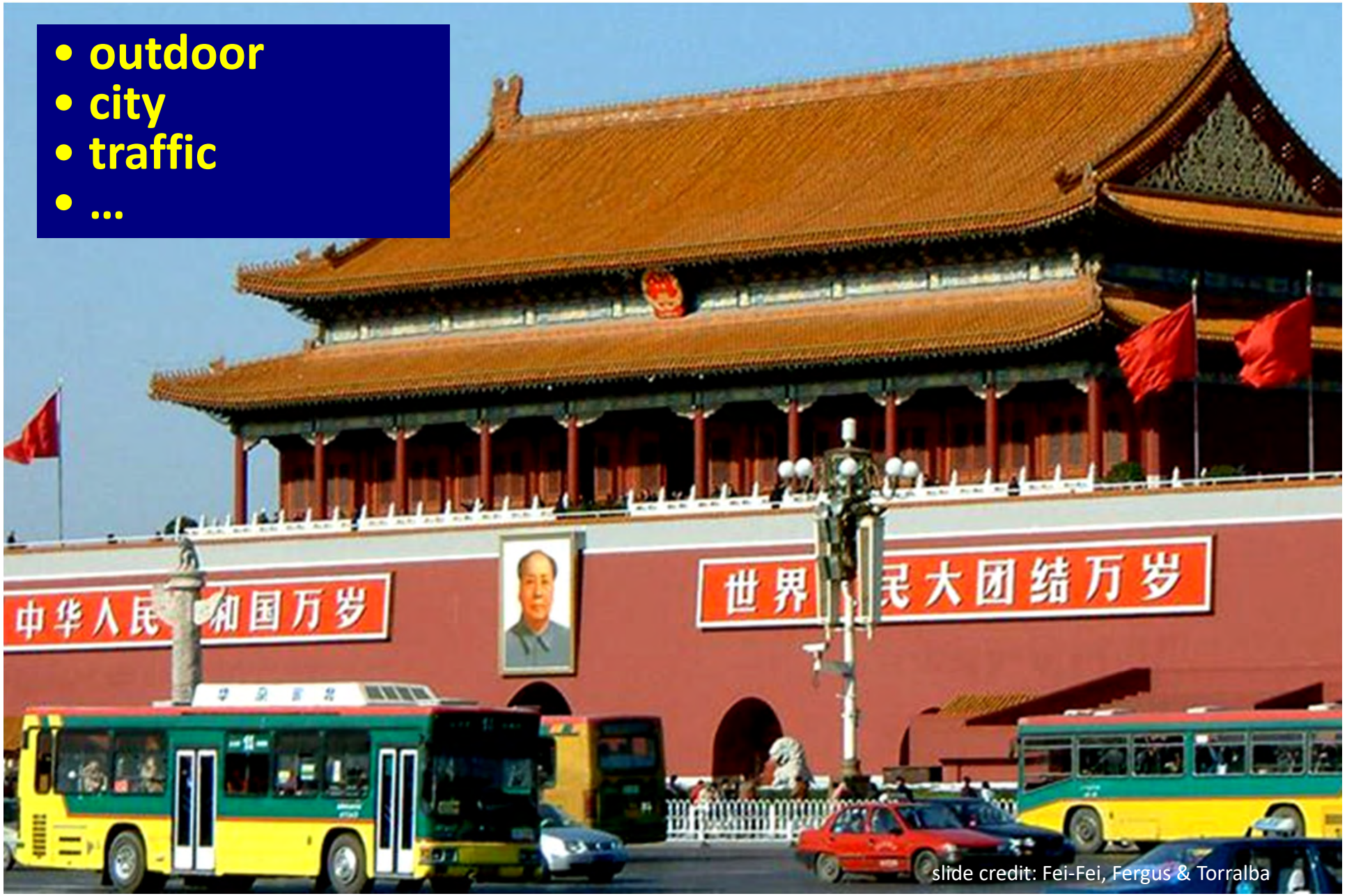
Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

Object categorization



Scene and context categorization



A little story about computer vision

- In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”

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

Marvin Minsky
(1927-2016)
Turing Award (1969)

Today

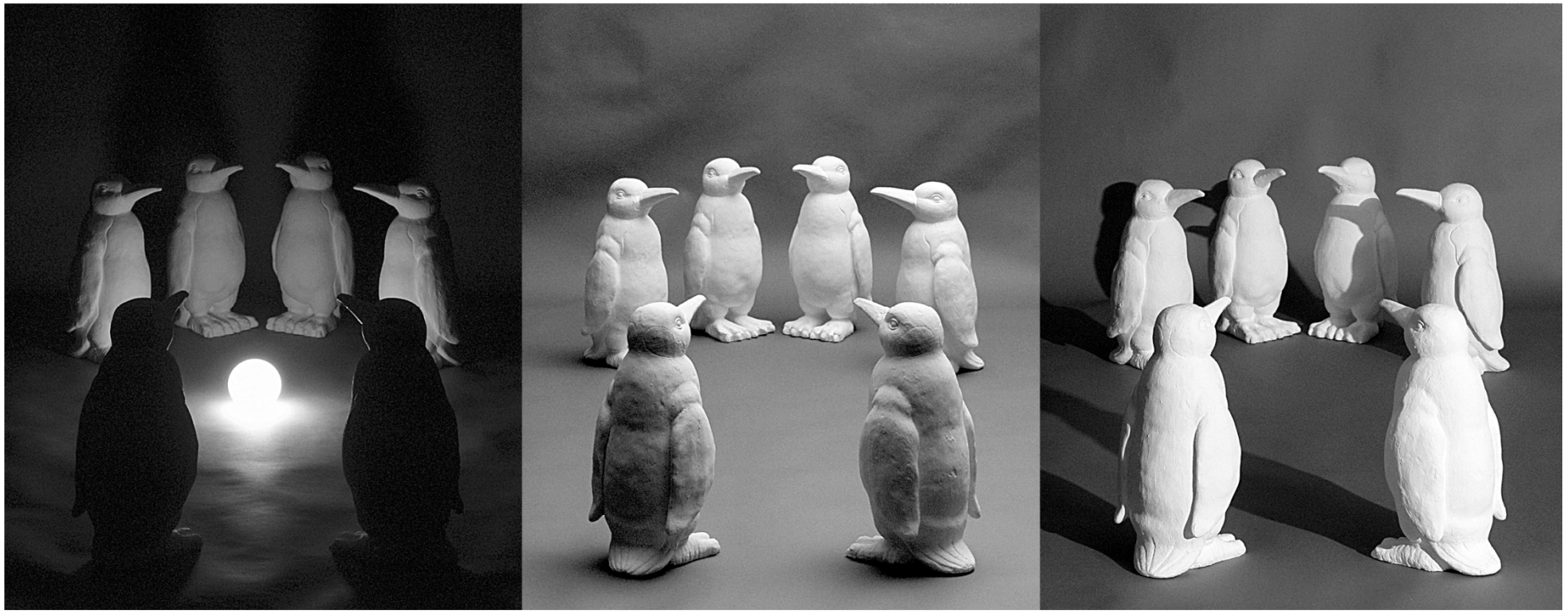
- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview

Why computer vision is difficult?

Challenges: viewpoint variation

Why computer vision is difficult?

Challenges: illumination



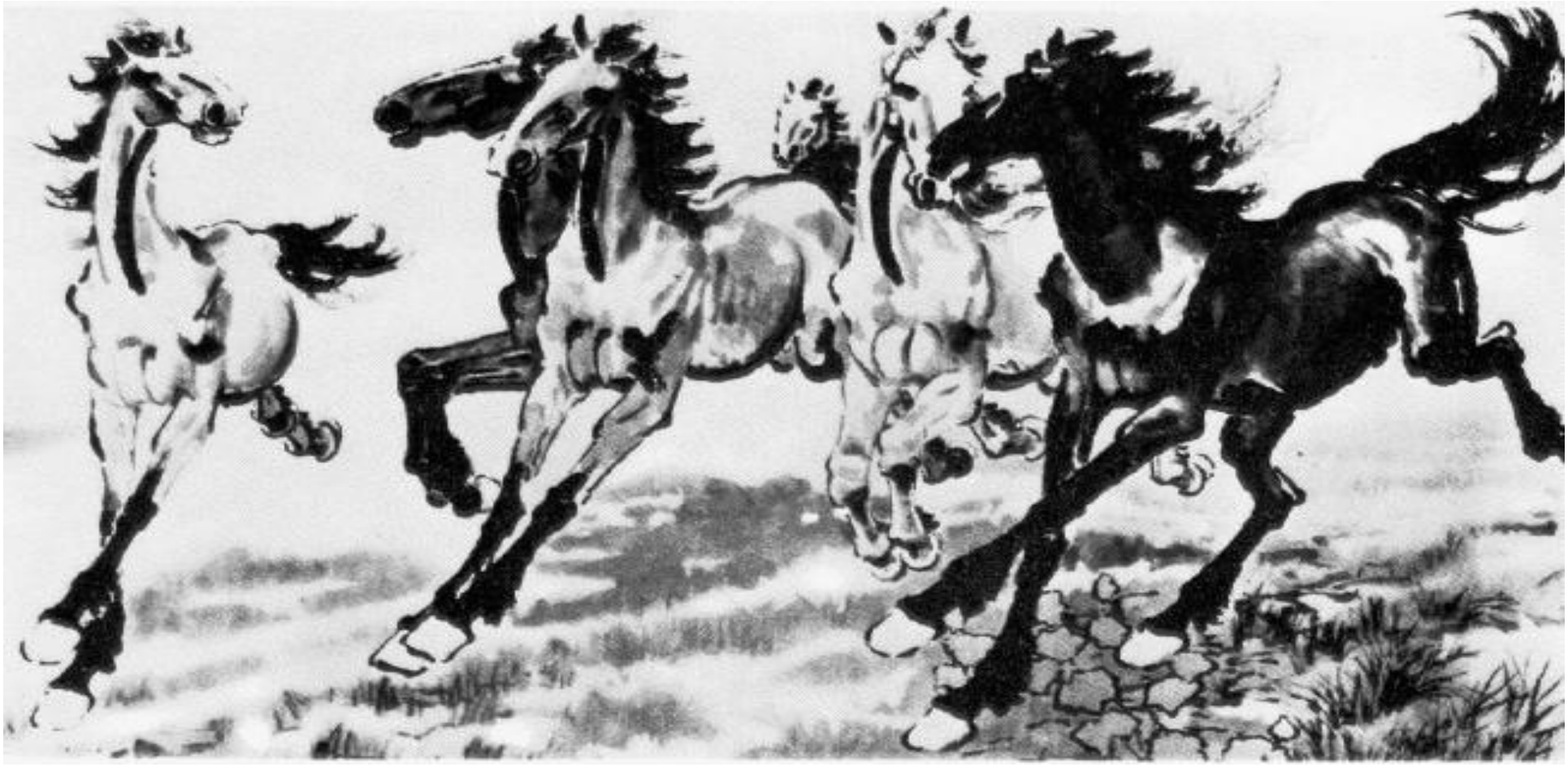
Why computer vision is difficult?

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

Why computer vision is difficult?

Challenges: deformation



Xu, Beihong 1943

Why computer vision is difficult?

Challenges: background clutter

Why computer vision is difficult?

Challenges: Motion

Why computer vision is difficult?

Challenges: object intra-class variation






Today

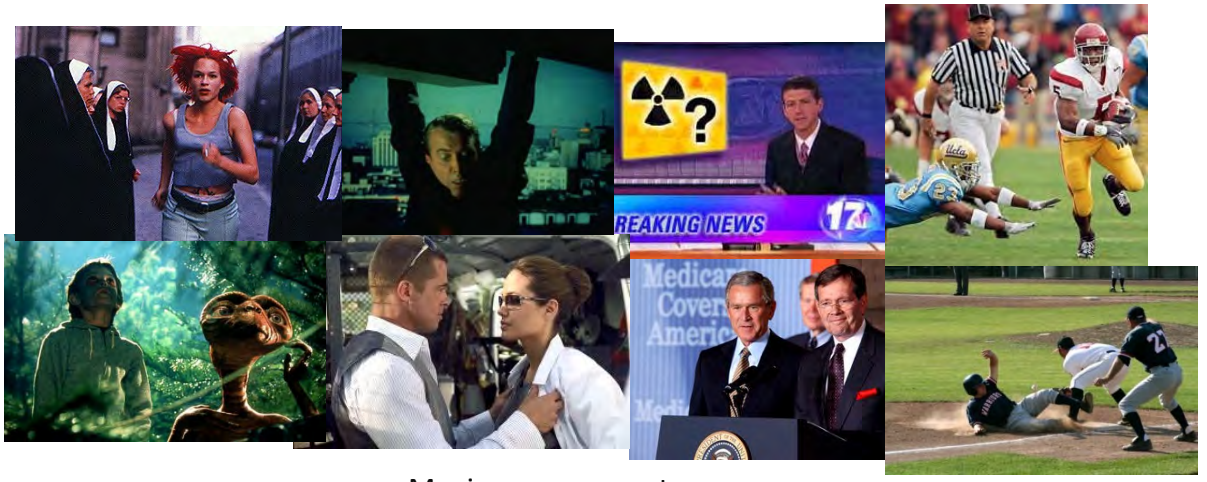
- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview

Why study computer vision?

Why study computer vision?

Vision is useful: Images and video are everywhere!

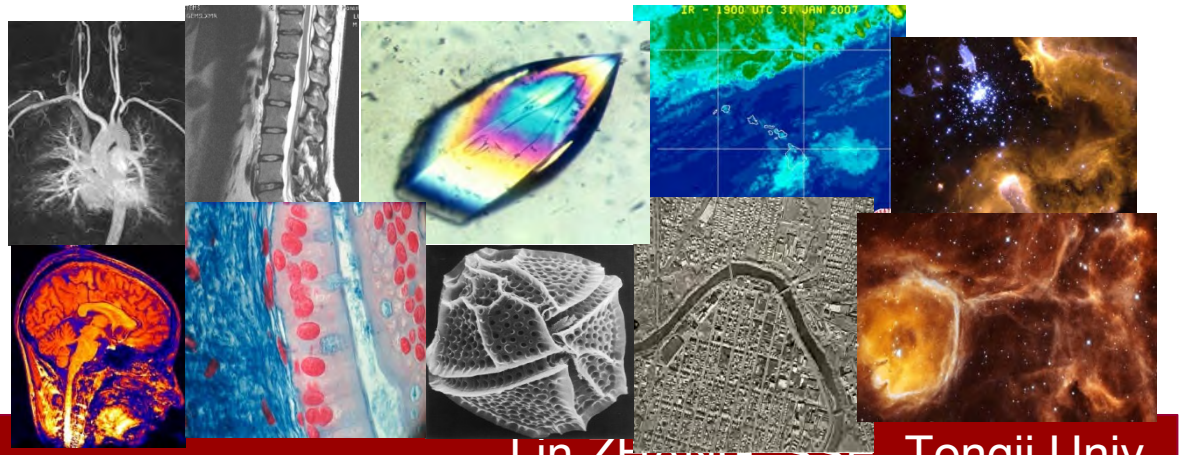
Personal photo albums



Movies, news, sports

Google Image Search Picasa™ flickr™ webshots picsearch™ YouTube Broadcast Yourself™

Surveillance and security



Medical and scientific images

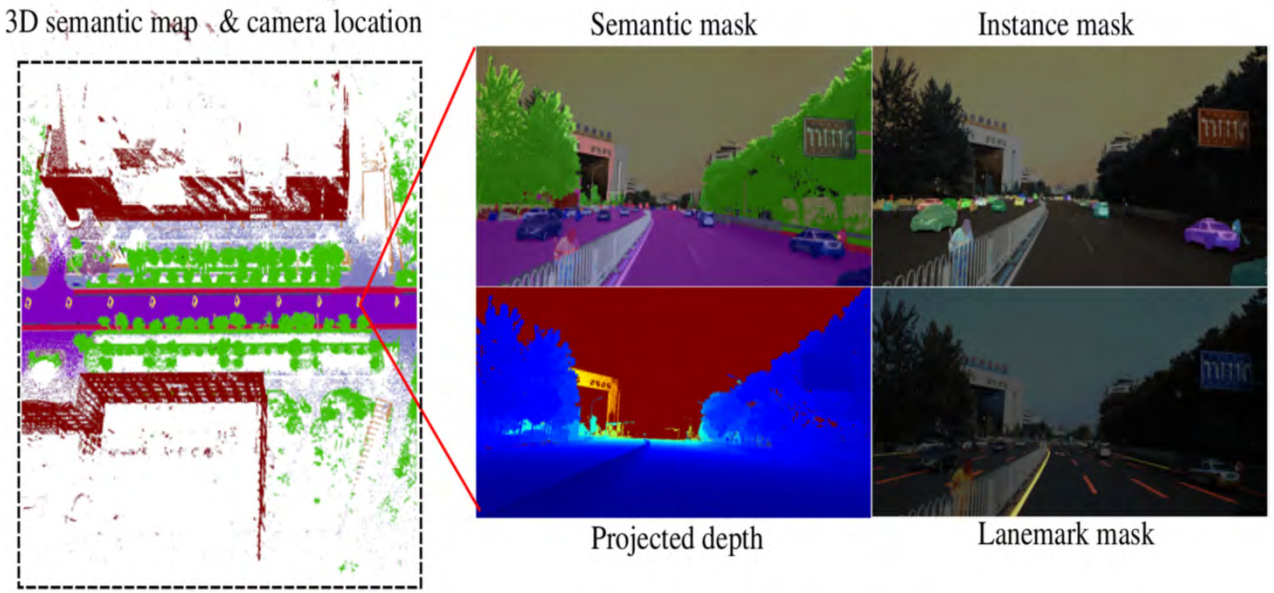
Why study computer vision?

Vision is useful: Images and video are everywhere!

Structure from motion

Bundler: Structure from Motion (SfM) for Unordered Image Collections (https://www.cs.cornell.edu/~snavely/bundler/#S3)

Automotive safety

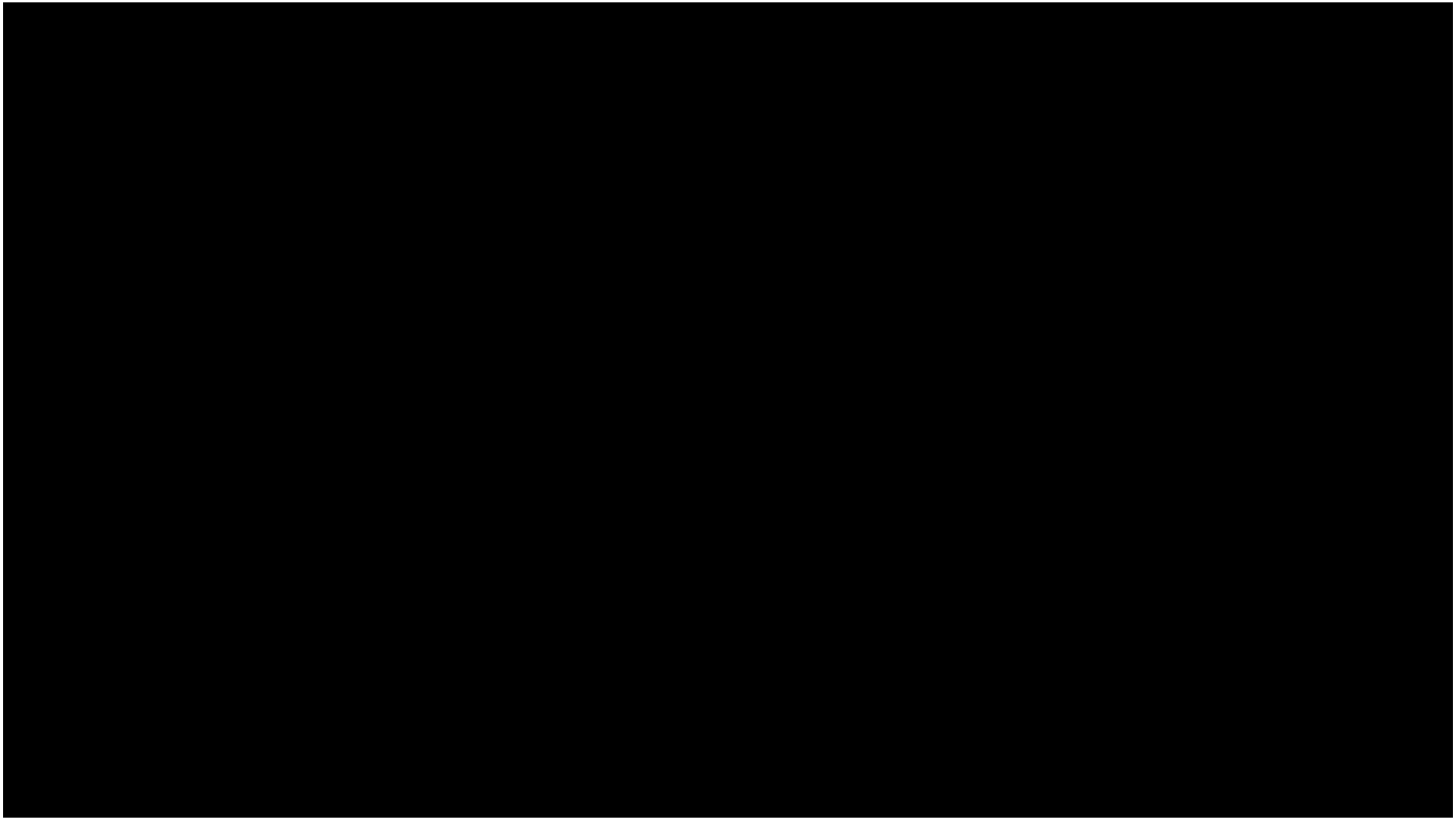


ApolloScape Open Dataset for Autonomous Driving[1]

- ✓ holistic semantic dense point cloud for each site
- ✓ Stereo
- ✓ per-pixel semantic labelling
- ✓ lanemark labelling
- ✓ instance segmentation
- ✓ 3D car instance
- ✓ high accurate location for every frame in various driving videos

[1] X. Huang et al., The ApolloScape Open Dataset for Autonomous Driving and Its Application, IEEE T-PAMI, vol. 42, pp. 2702-2719, Oct. 2020

AVP (Automated Valet Parking)

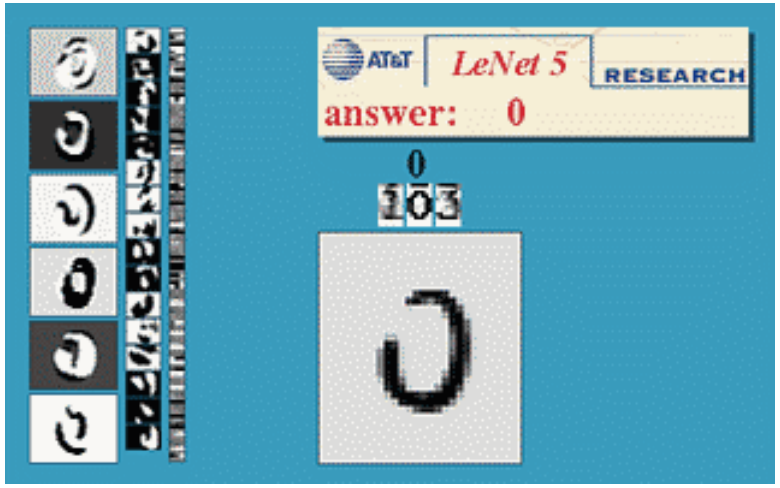


2021年4月，威马W6车型已量产百度AVP系统

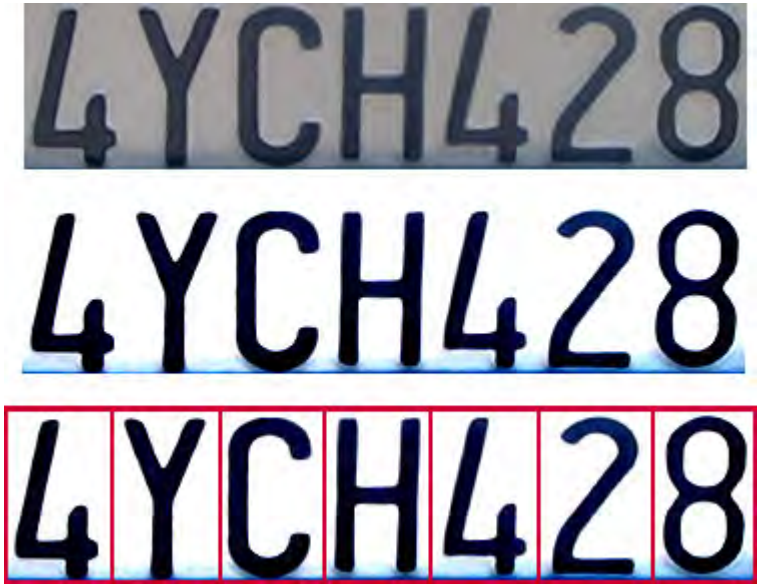
Optical character recognition (OCR)

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/

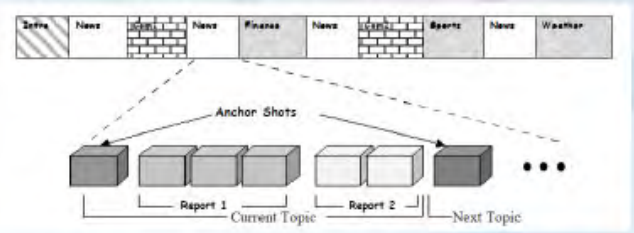


License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Videos based applications

视频拆分系统

- 检测片头片尾
- 检测新闻故事边界
- 检测广告和看点



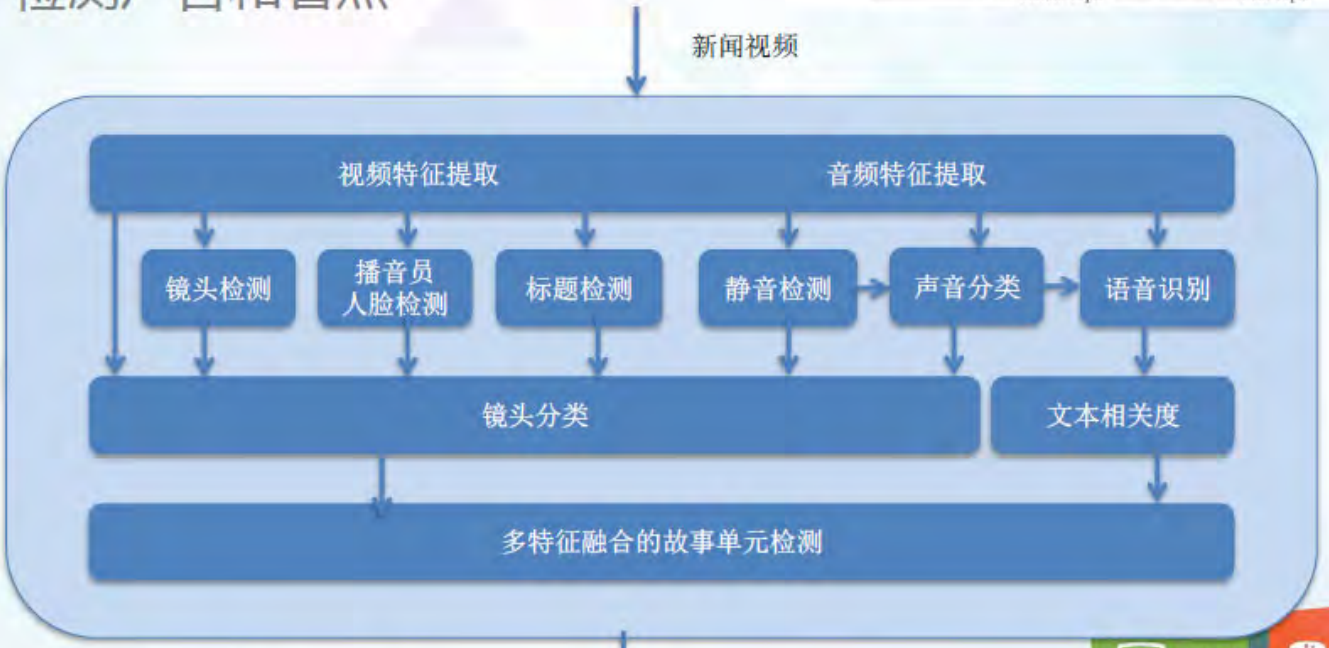
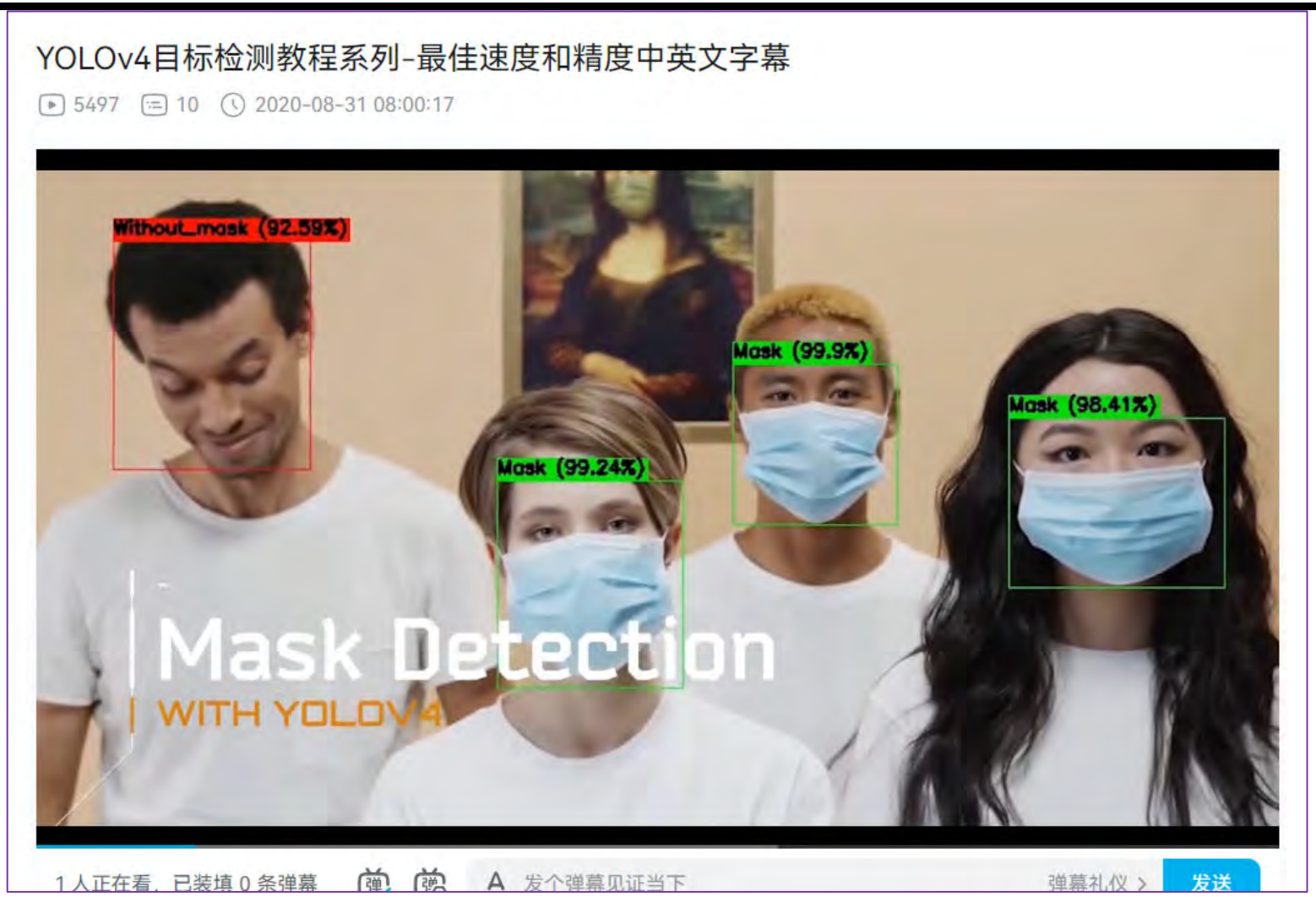


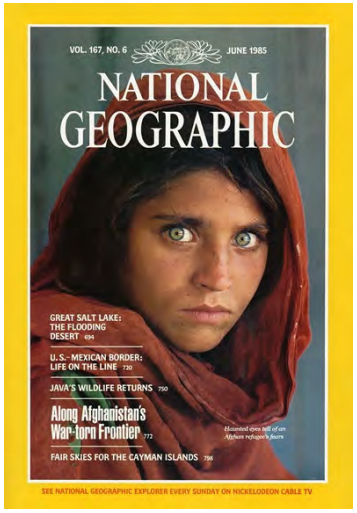
图 视频拆分标注系统

Face and mask detection

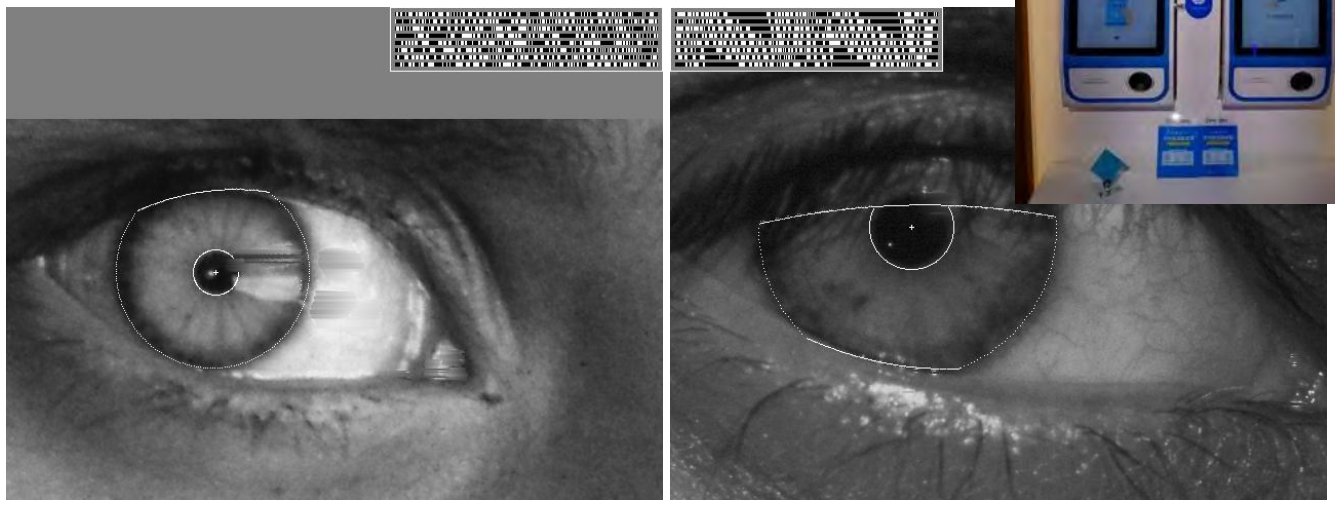


https://www.bilibili.com/video/av926885838?p=15

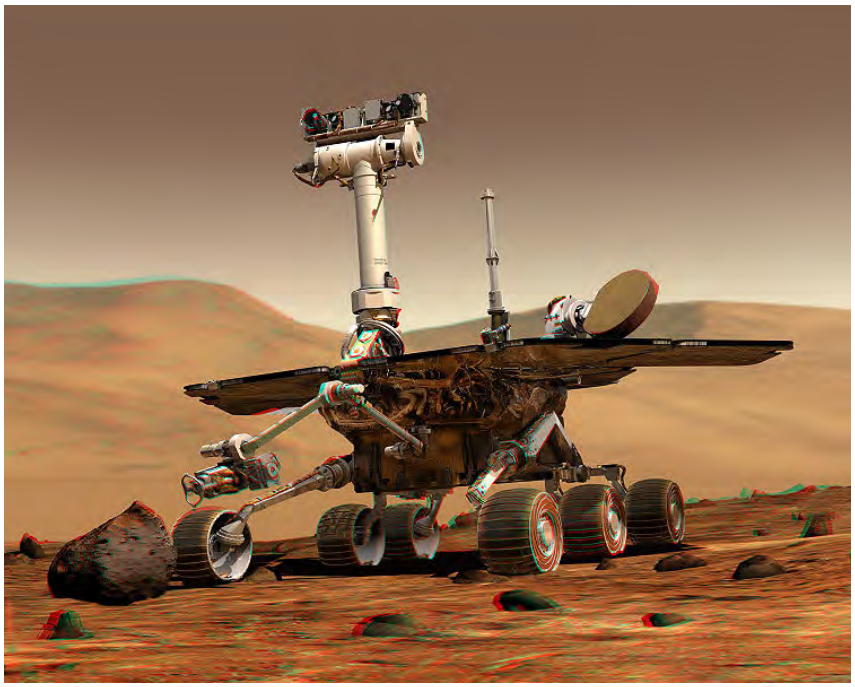
Vision-based biometrics



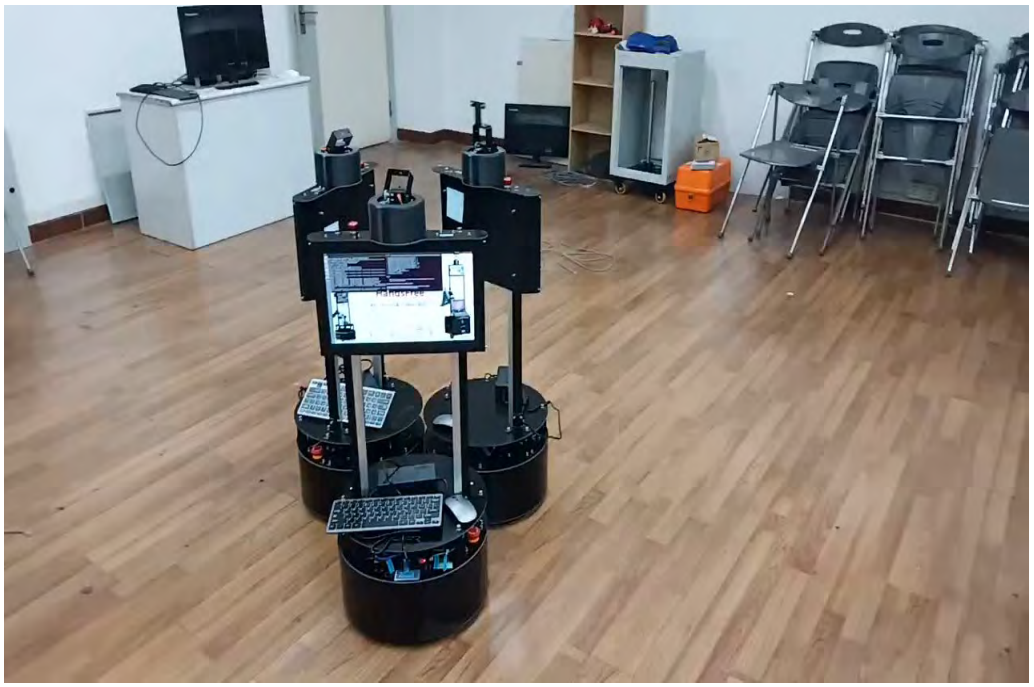
"How the Afghan Girl was Identified by Her Iris Pattern



Robotics



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



Robotics in MAP (Machine Autonomous Perception) lab

Work by our students: Photosynth

Project products of students from 2009 Media&Arts

Recent work of our group

Lin Zhang, Lida Li, Anqi Yang et al., "Towards contactless palmprint Recognition: A novel device, a new benchmark, and a collaborative representation based identification approach", Pattern Recognition, vol. 69, pp. 199-212, 2017

Recent work of our group

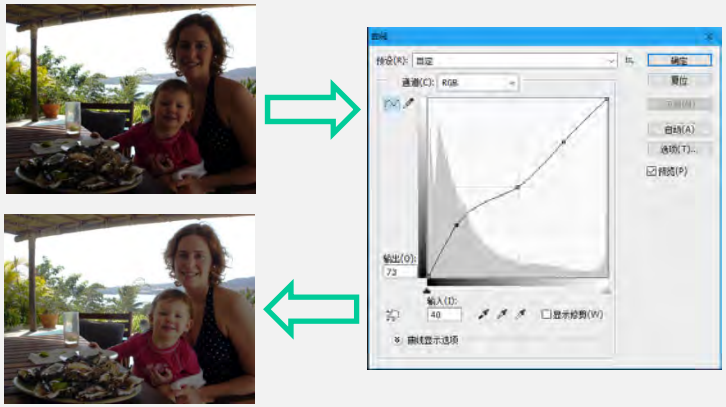
Yingyi Zhang, Lin Zhang et al., Pay by showing your palm: A study of palmprint verification on mobile platforms, in: Proc. ICME, pp. 862-867, 2019.

Recent work of our group

图像曝光度的自动校正

研究目标

- 图像曝光度失真的校正
- 利用无监督学习解决马尔可夫随机场问题，设计端到端的自动化算法
- 部署至移动端
- 实时？视频曝光度校正？



图像曝光度校正原理——调整图像α曲线

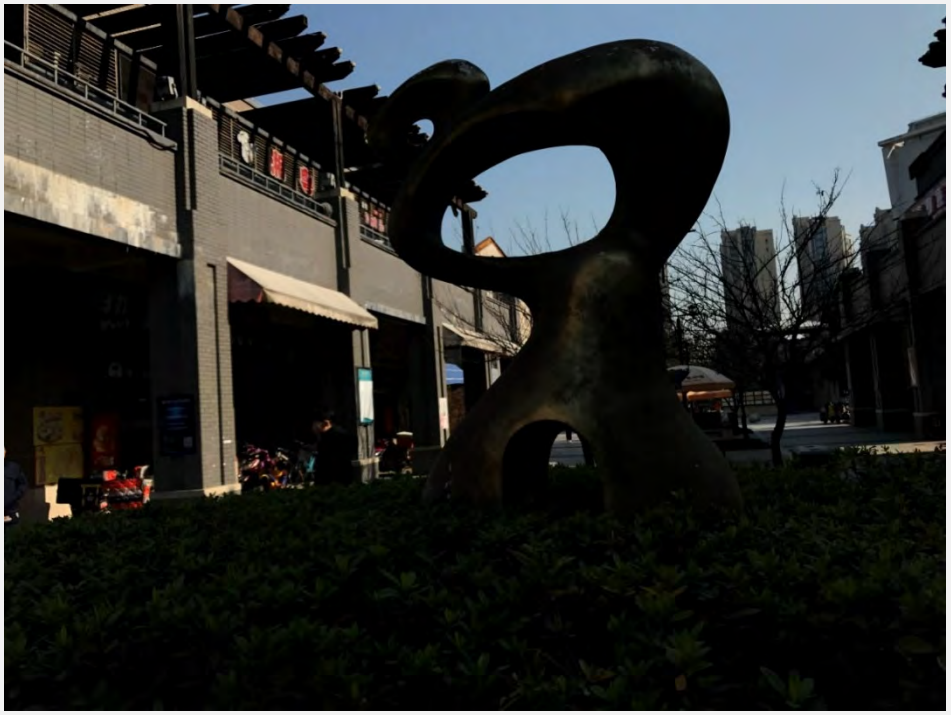
图像曝光度校正典型示例

我们的贡献

- ✓ 首次提出了MRF as CNN算法，利用神经网络解决图像曝光度校正问题
- ✓ 部署至移动端，开发具有实用价值的移动应用

Recent work of our group

图像曝光度的自动校正



Recent work of our group

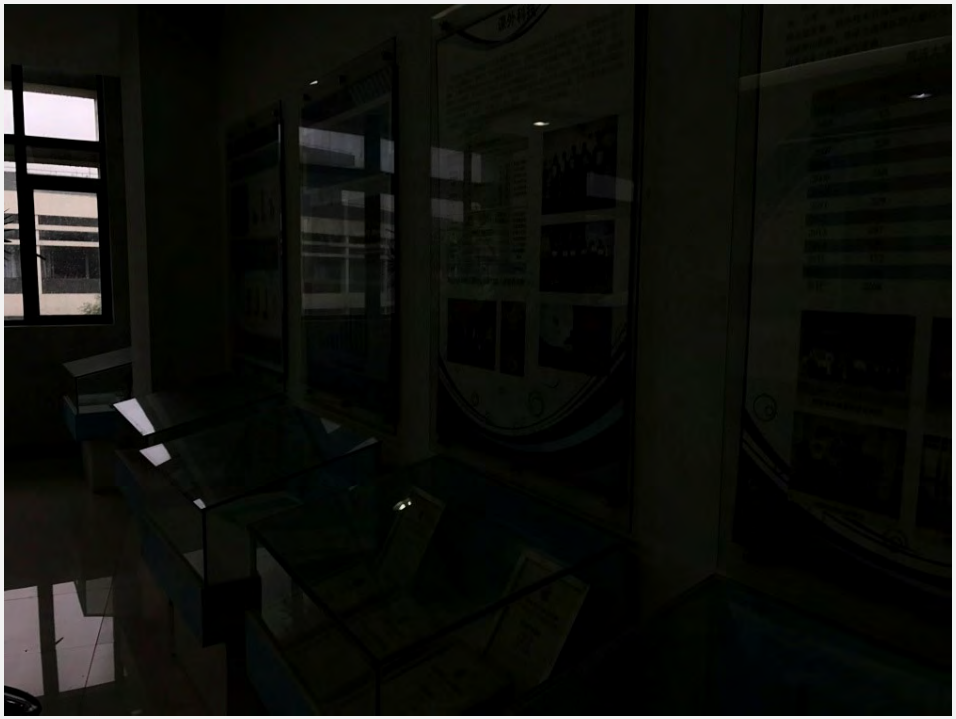
图像曝光度的自动校正

Recent work of our group

图像曝光度的自动校正

Recent work of our group

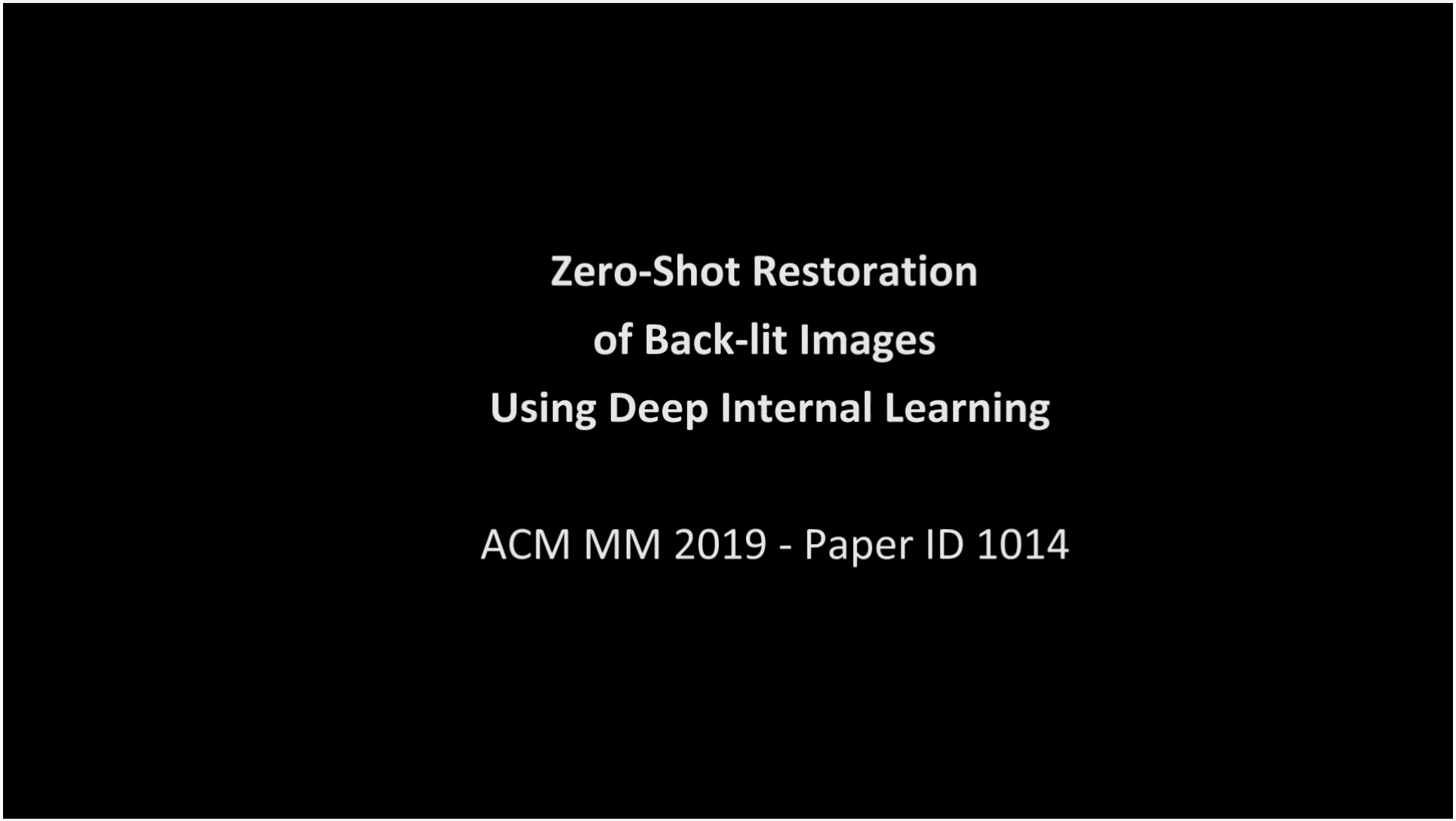
图像曝光度的自动校正



Recent work of our group

图像曝光度的自动校正

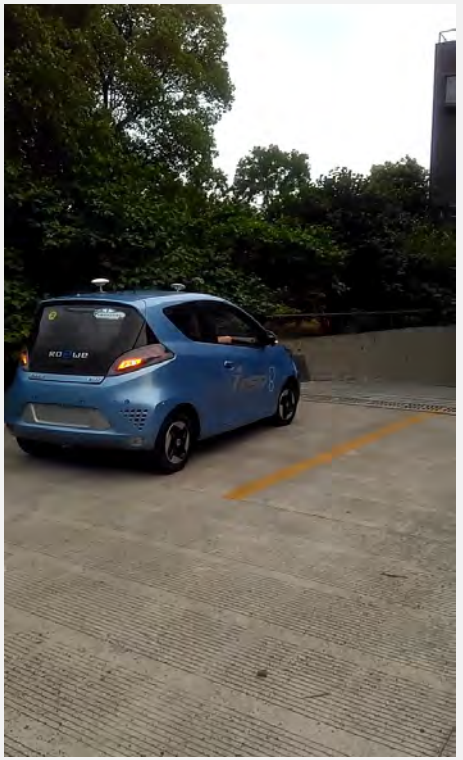
Recent work of our group



Lin Zhang, Lijun Zhang et al., "Zero-Shot Restoration of Back-lit Images Using Deep Internal Learning", ACM Int'l Conf. Multimedia, 2019

Recent work of our group

Short-range Self-parking



Lin Zhang, Junhao Huang et al., "Vision-based parking-slot detection: A DCNN-based approach and a large-scale benchmark dataset", IEEE Trans. Image Processing, vol. 27, no. 11, pp. 5350-5364, 2018.

Recent work of our group

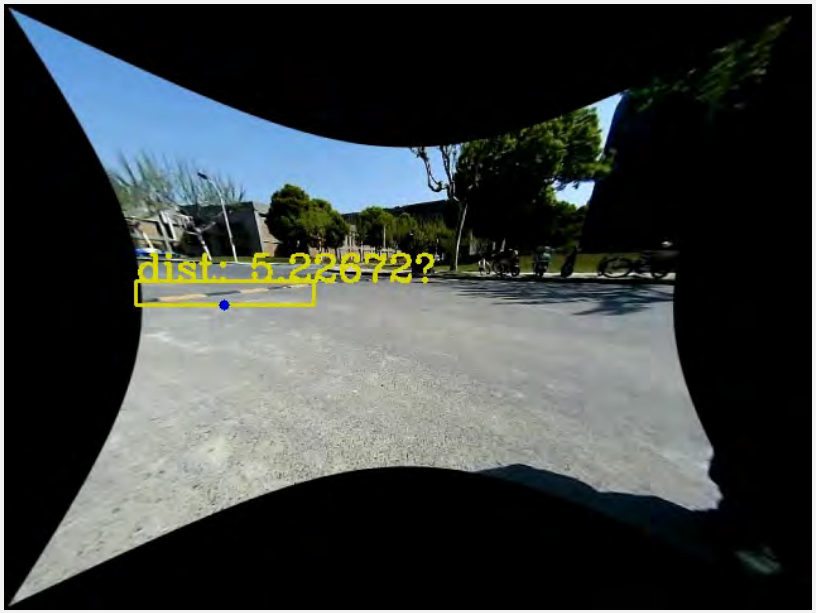
2018年9月，人民网采访报道

在学术界和工业界都产生了较大影响

- 在学术界的影响
  - ✓ CSDN、知乎等多家网络技术媒体对申报人的泊车位检测技术进行了大篇幅介绍和正面评价
  - ✓ 公开的数据集和工具已经被美国南德州大学、韩国汉阳大学、北交大、华南理工、湖南大学等多所研究机构的人员下载使用
- 在工业界的影响
  - ✓ 华为、科大讯飞、纵目、天瞳威视等企业下载使用了我们的数据或复现了我们的算法

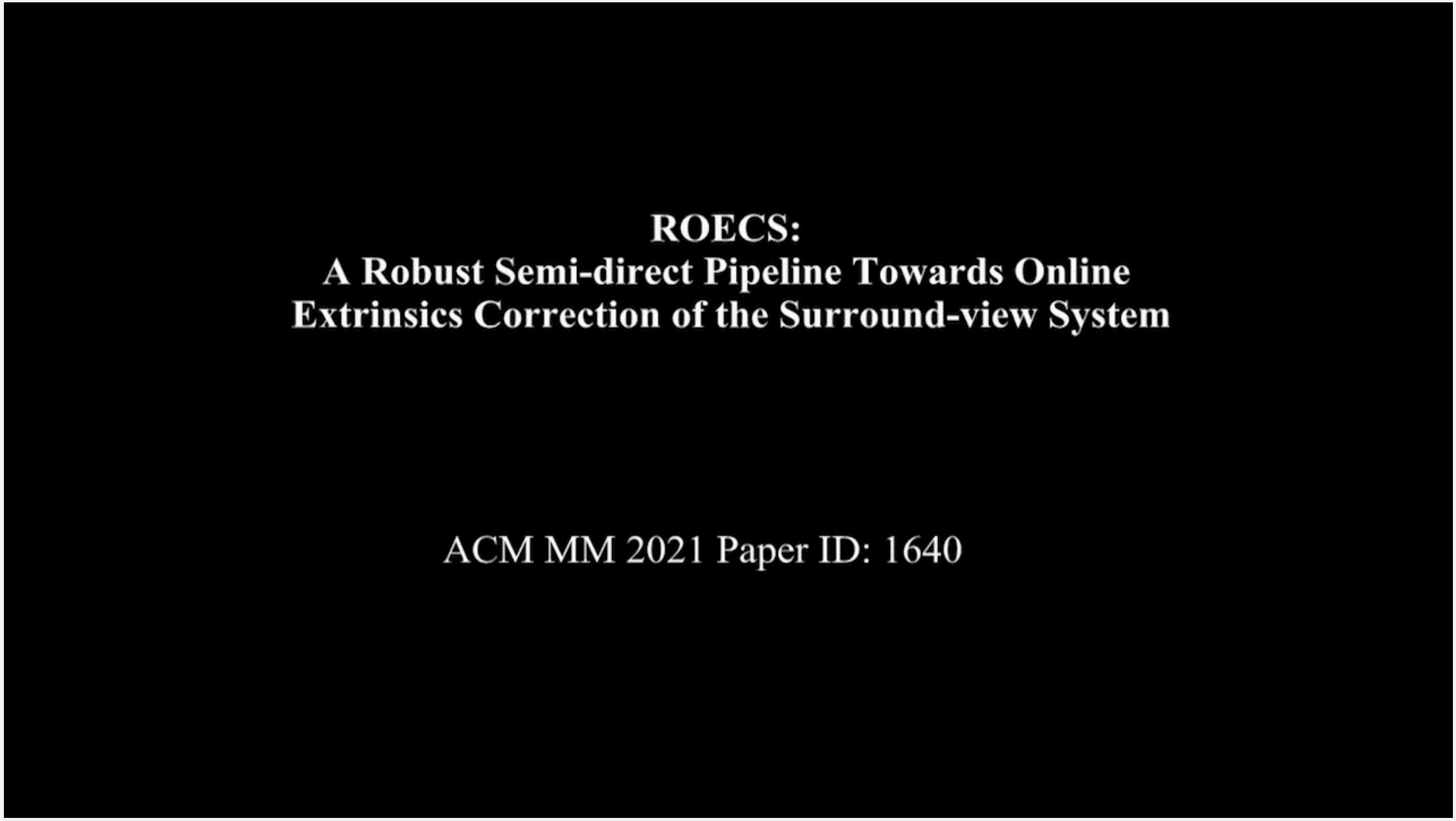
Recent work of our group

Pedestrian and speed-bump detection and distance measurement



Recent work of our group

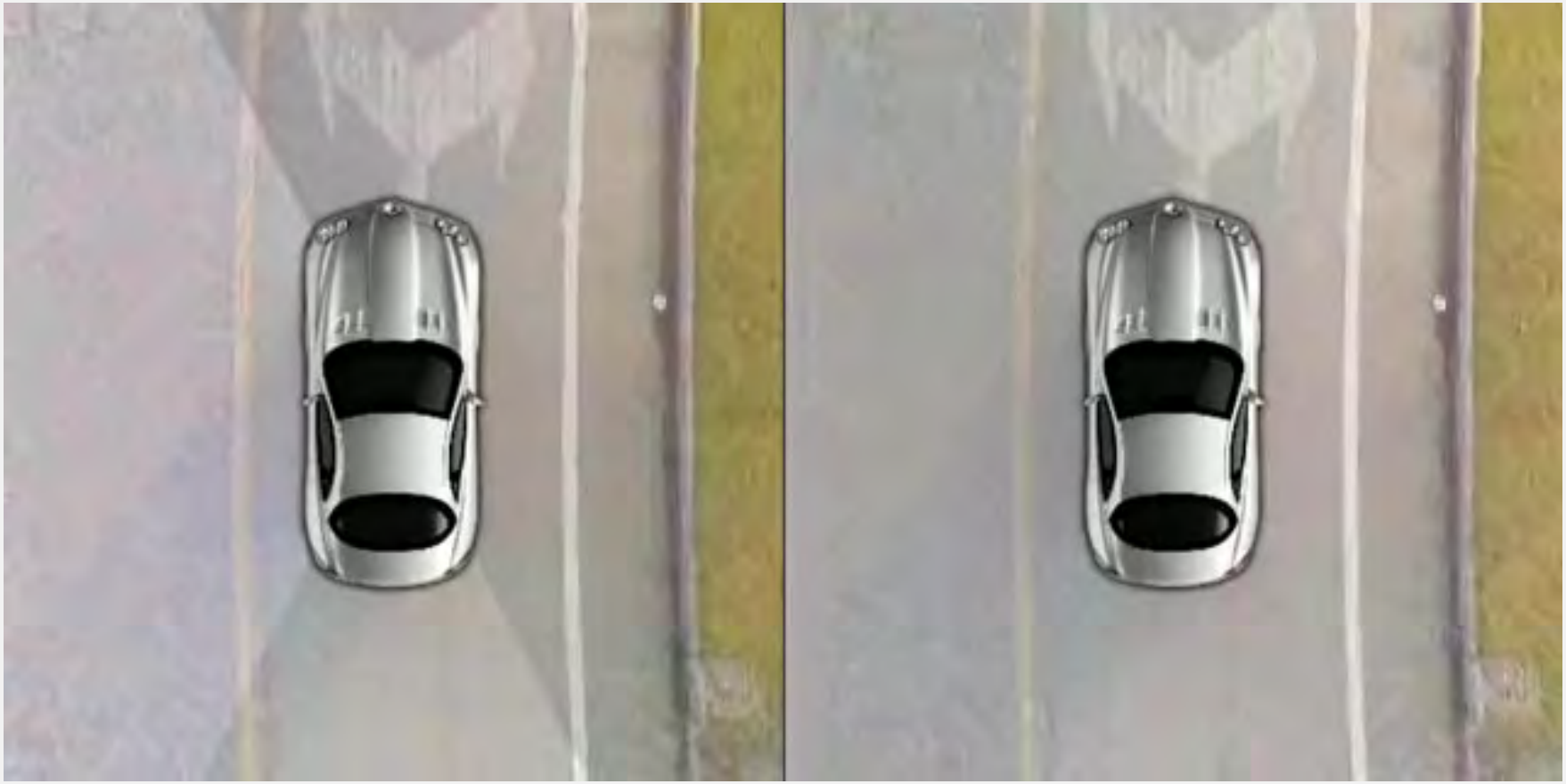
Online optimization of camera poses in a surround-view system



Tianjun Zhang, Lin Zhang et al., "ROECS: A Robust Semi-direct Pipeline Towards Online Extrinsics Correction of the Surround-view System", in Proc. ACM Int'l Conf. Multimedia, 2021.

Recent work of our group

Photometric adjustment in the surround-view



Recent work of our group

Xuan Shao, Lin Zhang et al., "MOFISSLAM: A multi-object semantic SLAM system with front-view, inertial and surround-view sensors for indoor parking", IEEE Trans. CSVT, 2022

Recent work of our group

Lin Zhang, Anqi Zhu et al., "Simulation of atmospheric visibility impairment", IEEE Trans. Image Processing, vol. 30, pp. 8713-8726, 2021.

You can find a good job!

- Many first-class companies now are developing CV related applications, to name a few
  - Google
  - Ali
  - Facebook
  - SenseTime
  - megvii
  - Tencent
  - Baidu
  - DJI
  - Huawei
  - …

Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview

Course content

- **Introduction**
- **Local interest point detectors**
- **Local feature descriptors and matching**
- **Projective geometry**
- **Nonlinear least squares**
- **Measurement using a single camera**
- **Basics for machine learning and its applications**
- **Applications of DCNNs**
- **Introduction to numerical geometry**

The lectures are grouped into four themes

Some tips

- Prerequisites
  - Linear algebra
  - Calculus
  - Matlab Programming
  - C++ Programming
- Knowledge sources
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
  - IEEE Transactions on Image Processing (TIP)
  - International Journal of Computer Vision (IJCV)
  - IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)
  - IEEE International Conference on Computer Vision (ICCV)
  - European Conference on Computer Vision (ECCV)


Q+A