

# 電腦視覺

# Computer Vision: from Recognition to Geometry

簡韶逸 Shao-Yi Chien

Department of Electrical Engineering  
National Taiwan University

Spring 2023

# Computer Vision

- Describe the world that the computer see in one or more images and to reconstruct its properties, such as shape, illumination, and color distribution
- Is it hard? An inverse problem



# Computer Vision



[R. C. James]

# Computer Vision

FAILURE PRESS PHOTO EXHIBITION

失敗新聞攝影展 09.01.2018 - 09.10.2018

開幕 Reception 09.01 02:00 p.m.

新北市政府 1 樓大廳東側  
New Taipei City Hall 1F East



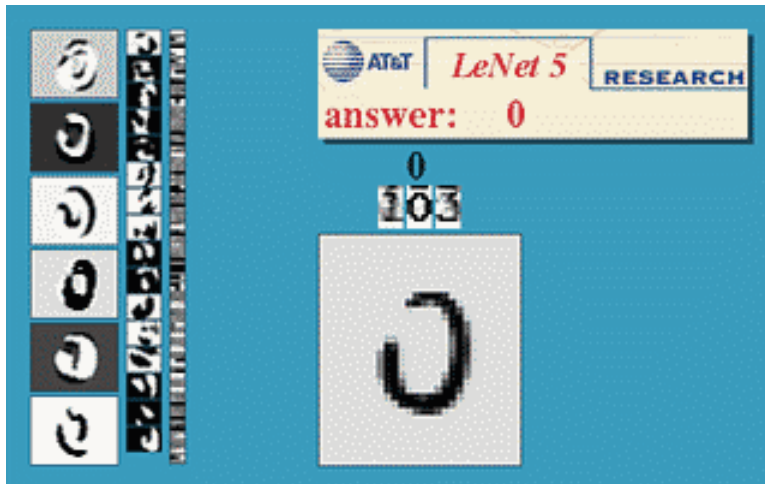
220 新北市板橋區中山路一段 161 號 1 樓

協力贊助

Avocado

# Wide Applications of Computer Vision

- Optical character recognition (OCR)



Digit recognition, AT&T labs

<http://www.research.att.com/~yann/>

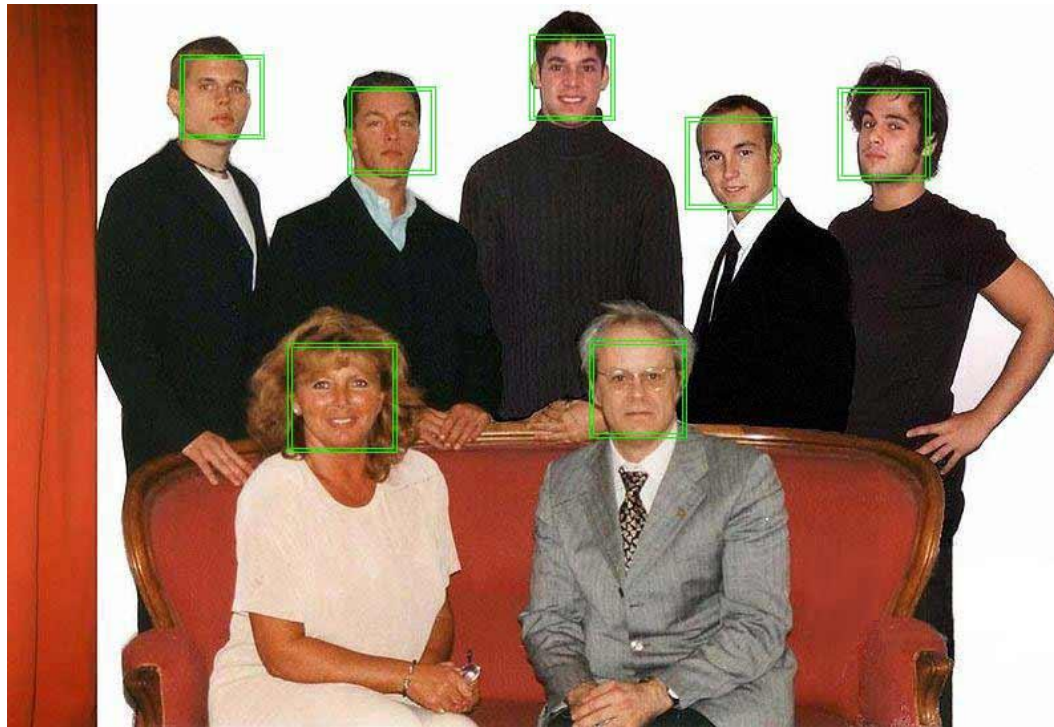


License plate readers

[http://en.wikipedia.org/wiki/Automatic\\_number\\_plate\\_recognition](http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

# Wide Applications of Computer Vision

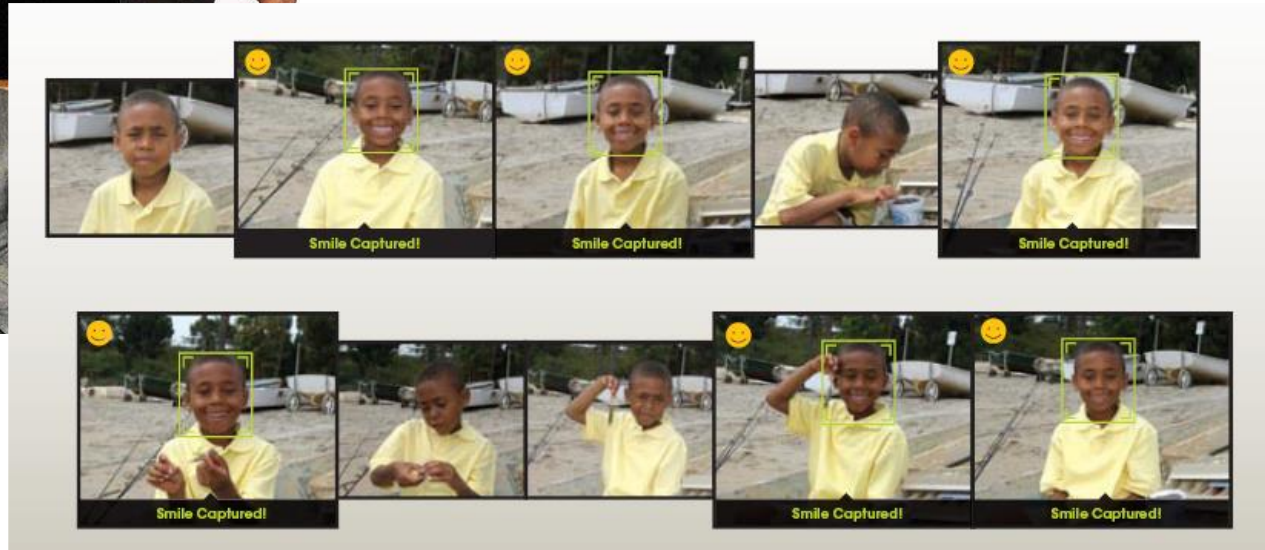
- Face detection: in all digital cameras and smart phones





# Wide Applications of Computer Vision

- Face detection: in all digital cameras and smart phones



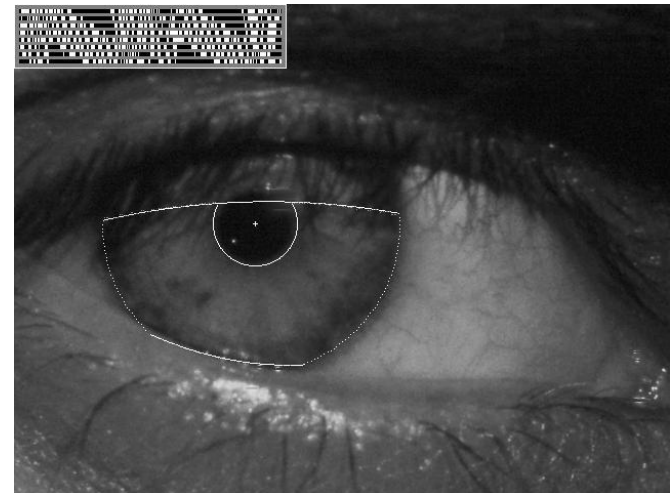
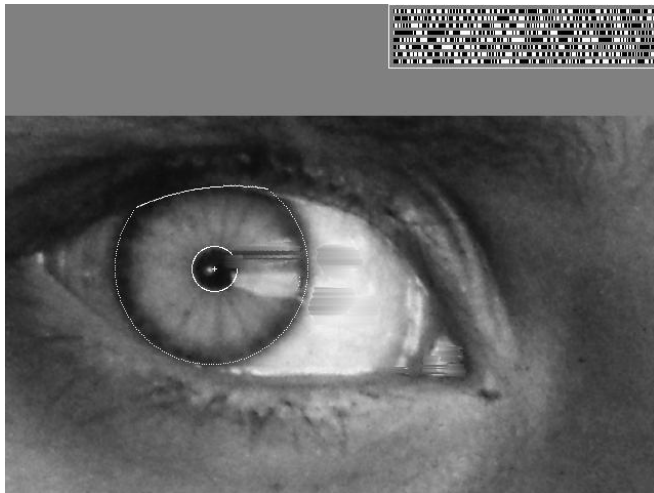
[Sony]

# Wide Applications of Computer Vision

- Iris recognition  
(Vision-based biometrics)



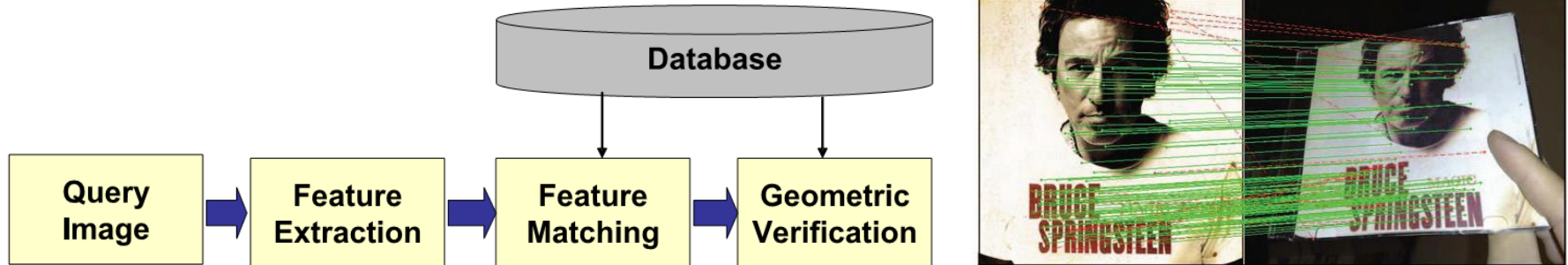
*“How the Afghan Girl was Identified by Her Iris Patterns”* Read the [story](#)



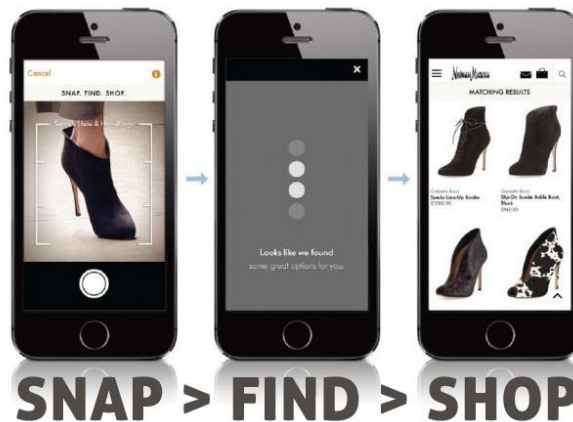


# Wide Applications of Computer Vision

- Object recognition



[Girod et al. 2011]



[slyce.it]

# Wide Applications of Computer Vision

- Shape capture



*The Matrix* movies, ESC Entertainment, XYZRGB, NRC

# Wide Applications of Computer Vision

- Motion capture



*Pirates of the Carribean,*  
Industrial Light and Magic

# Wide Applications of Computer Vision

- Computer vision in sports



Hawk-Eye: helping/improving referee decisions



Intel: [freeD technology](#)



# Wide Applications of Computer Vision

- Smart cars: [ADAS](#)

The screenshot displays the Intel Mobileye website with a focus on automotive safety and computer vision. The main banner features a top-down view of a car with four camera fields of view: rear, forward, and two side cameras, all labeled. The headline reads "Our Vision. Your Safety." Navigation tabs for "manufacturer products" and "consumer products" are at the top. To the right, a "News" section lists articles about Volvo's first collision warning system and a new collision warning system. Below the main banner, three product/application tiles are shown: "EyeQ Vision on a Chip" with an image of the chip, "Vision Applications" showing a pedestrian detection box, and "AWS Advance Warning System" with a car icon and a distance reading of "0.8". Each tile includes a "read more" link. On the far right, an "Events" section lists upcoming appearances at Equip Auto in Paris and SEMA in Las Vegas, also with a "read more" link.

manufacturer products consumer products

## Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

**EyeQ** Vision on a Chip

**Vision Applications**  
Road, Vehicle, Pedestrian Protection and more

**AWS** Advance Warning System

0.8

**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**

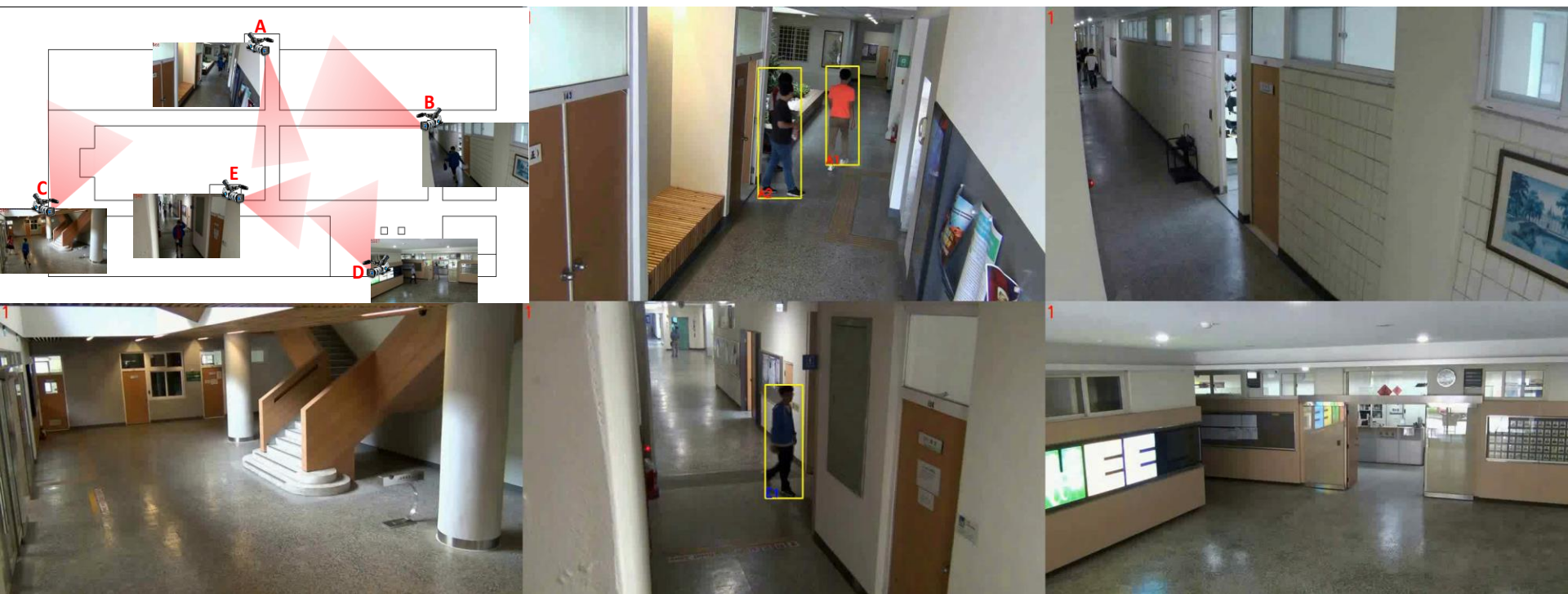
- > Mobileye at Equip Auto, Paris, France
- > Mobileye at SEMA, Las Vegas, NV
- > read more

[Intel Mobileye]



# Wide Applications of Computer Vision

- Surveillance system



Ref: Chih-Wei Wu, Meng-Ting Zhong, Yu Tsao, Shao-Wen Yang, Yen-Kuang Chen, and Shao-Yi Chien, "Track-clustering Error Evaluation for Track-based Multi-camera Tracking System Employing Human Re-identification," *CVPR 2016 Workshop*.

# Wide Applications of Computer Vision

- Vision-based interaction



[Microsoft Xbox]

# Wide Applications of Computer Vision

## DodecaPen: Puppy



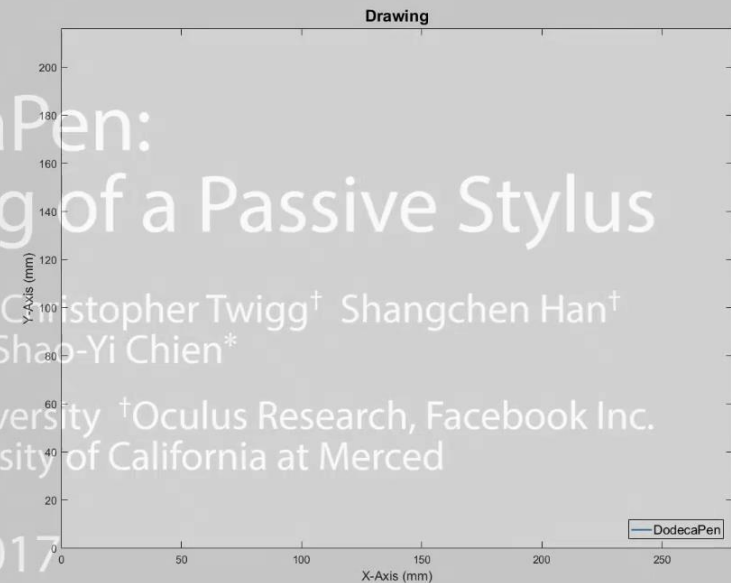
### DodecaPen: Accurate 6DoF Tracking of a Passive Stylus

Po-Chen Wu<sup>\*†</sup> Robert Wang<sup>†</sup> Kenrick Kin<sup>†</sup> Christopher Twigg<sup>†</sup> Shangchen Han<sup>†</sup>  
Ming-Hsuan Yang<sup>†</sup> Shao-Yi Chien<sup>\*</sup>

<sup>\*</sup>Media IC & System Lab, National Taiwan University <sup>†</sup>Oculus Research, Facebook Inc.

<sup>†</sup>Vision and Learning Lab, University of California at Merced

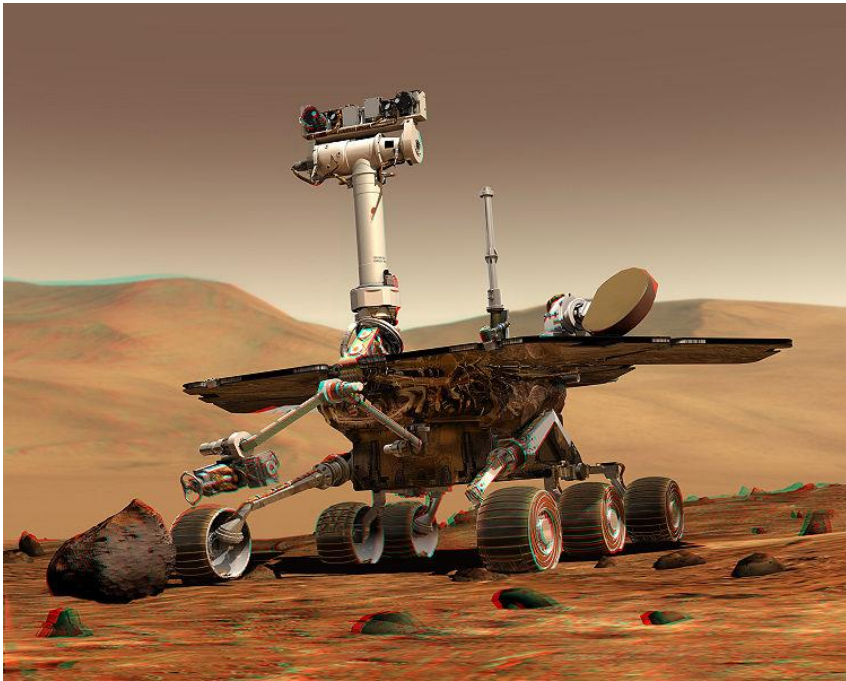
UIST 2017





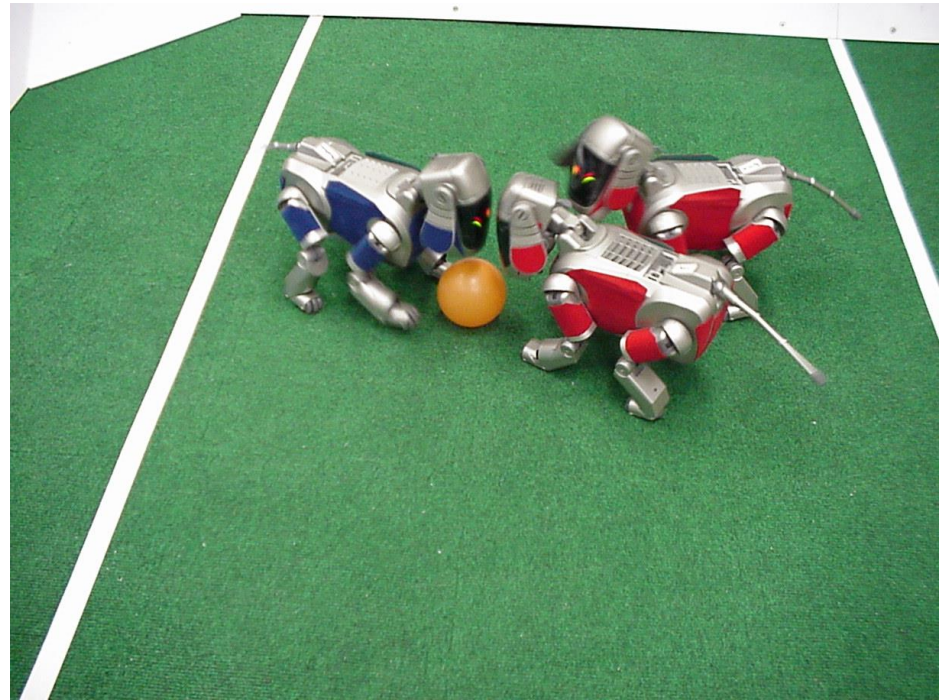
# Wide Applications of Computer Vision

- Robotics



NASA's Mars Spirit Rover

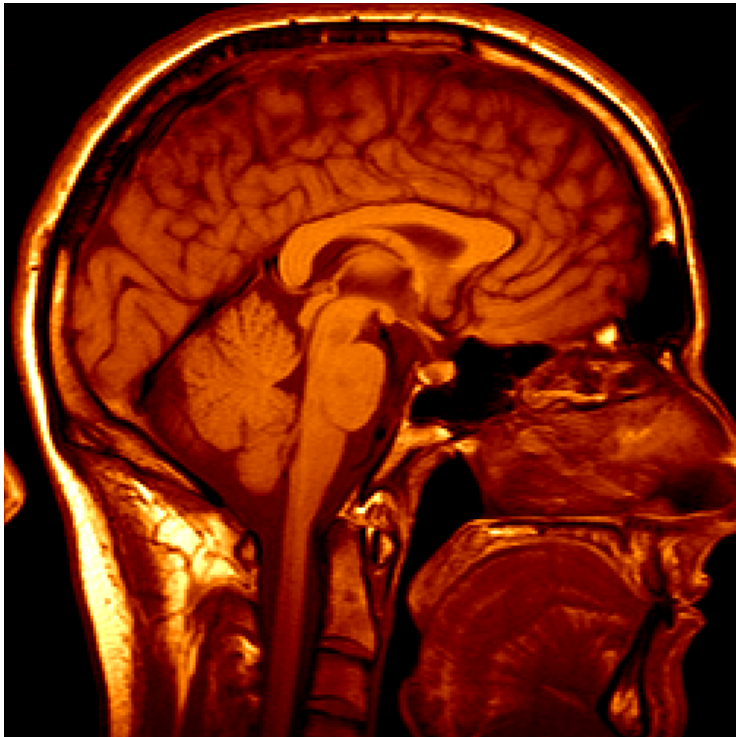
[http://en.wikipedia.org/wiki/Spirit\\_rover](http://en.wikipedia.org/wiki/Spirit_rover)



<http://www.robocup.org/>

# Wide Applications of Computer Vision

- Medical image



3D imaging  
MRI, CT



Image guided surgery  
[Grimson et al., MIT](#)



# Wide Applications of Computer Vision

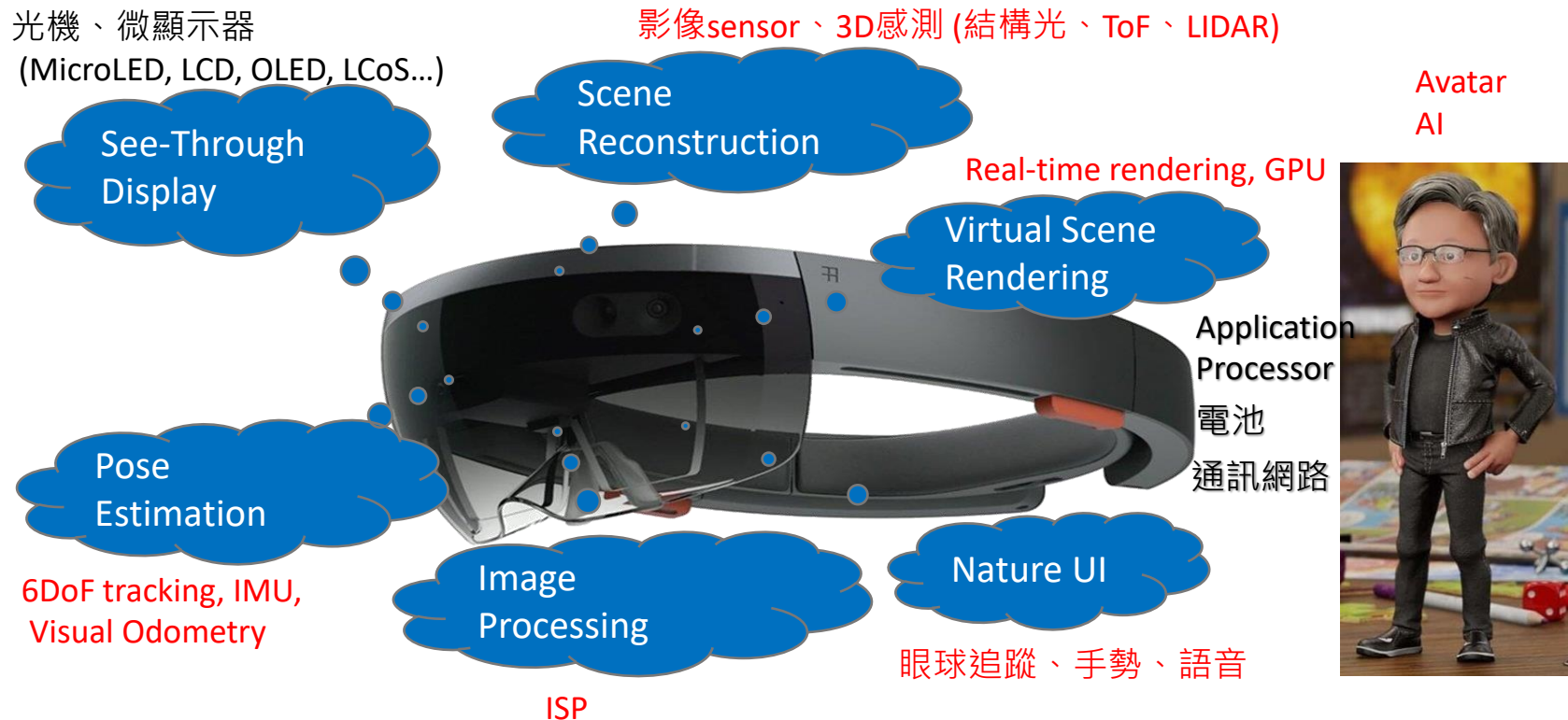


# Wide Applications of Computer Vision



# Wide Applications of Computer Vision

- AR/VR devices for the metaverse



# Wide Applications of Computer Vision



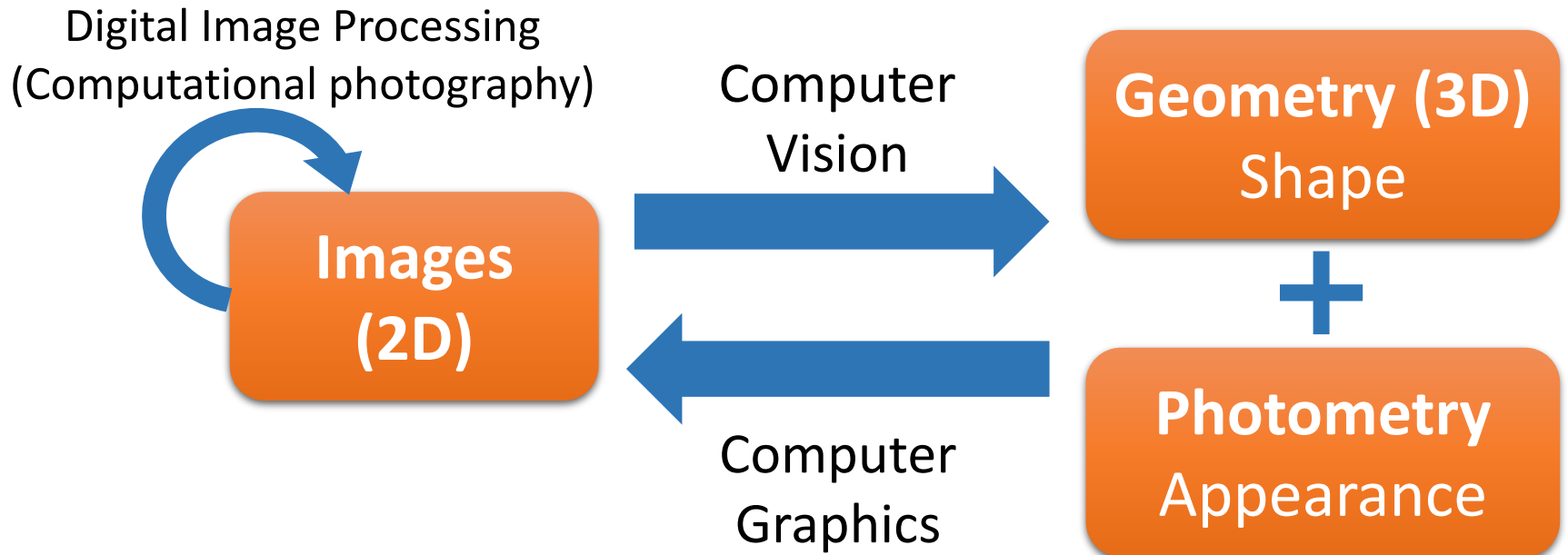
[Midjourney]

# Important Near-Future Applications

- AR/VR, Metaverse
- Autonomous vehicle
- Robot
- IoT: AIoT (AI+IoT), IoVT (Internet-of-Video-Things)
- Medical imaging
- Large-scale video analysis
- Computational photography/image synthesis
- Industrial automation
- ...



# Related Fields



- The boundaries between digital image processing/computer vision/computer graphics become vague nowadays

# About this Course...

- Provide a comprehensive introduction to the field of computer vision (CV)
  - From classical methods to deep learning based methods
  - From recognition to geometry
  - No experiences in CV and image processing are required
- The two courses, **Computer Vision** and **Deep Learning for Computer Vision**, can give you a complete view of modern CV techniques
- Grading
  - Four homeworks: 60%
  - Class/talk participation: 10%
  - Group final project: 30%

# Course Website

- Course website

- <https://cool.ntu.edu.tw/courses/26914>
- <http://media.ee.ntu.edu.tw/courses/cv/23S/>

- TA

- Yu-Kai Chen (陳昱愷) [chenyukai@media.ee.ntu.edu.tw](mailto:chenyukai@media.ee.ntu.edu.tw)
- Tzu-Chieh Liu (劉子傑) [tzujliu@media.ee.ntu.edu.tw](mailto:tzujliu@media.ee.ntu.edu.tw)
- Yi-Hsun Lee (李奕勳) [smilel6g84@media.ee.ntu.edu.tw](mailto:smilel6g84@media.ee.ntu.edu.tw)
- Yung-Wei Fan (范詠為) [ywfan@media.ee.ntu.edu.tw](mailto:ywfan@media.ee.ntu.edu.tw)

# (Tentative) Schedule: May be Modified...

Week	Date	Topic
1	2/24	Introduction to human vision systems
2	3/3	Camera basic, image formation and basic Image processing
3	3/10	Feature detection and matching
4	3/17	Machine learning basics
5	3/24	Deep learning basics
6	3/31	Recognition
7	4/7	Segmentation
8	4/14	Projective Geometry
9	4/21	Estimation of Transformations
10	4/28	Single Camera Geometry/Camera calibration
11	5/5	Two-View Geometry
12	5/12	Dense motion estimation/stereo
13	5/19	Optical flow + object tracking
14	5/26	3D reconstruction/depth sensing
15	6/2	Structure from motion
16	6/9	Final project presentation

或是延到6/16

# Homeworks

- Four assignments:
  - HW1: Image filters
  - HW2: Detection or recognition
  - HW3: Pose estimation
  - HW4: Stereo matching
- Official language is Python
- Lab0: Python and basic image processing
  - 3/1 18:30--20:00 @ TBD

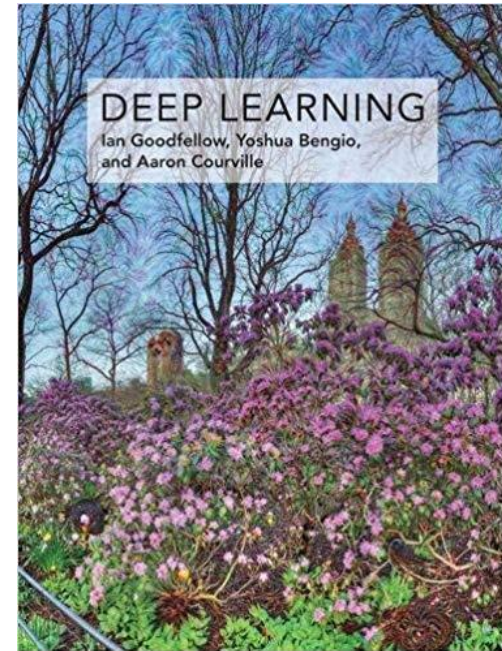
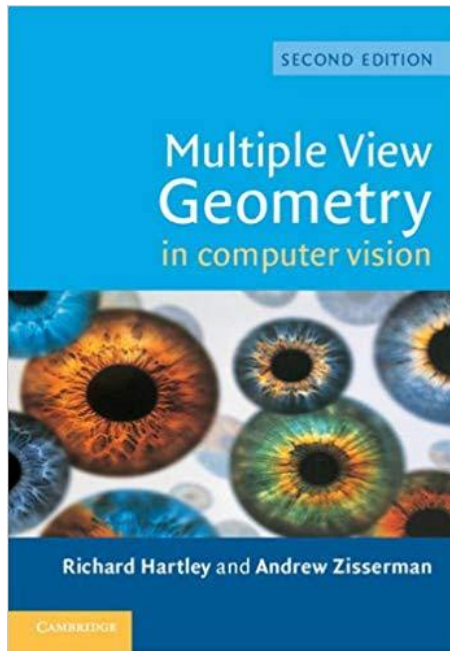
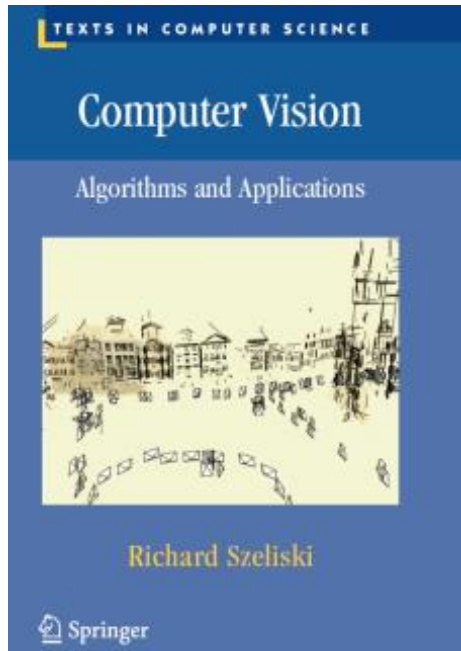


# Final Project

- Will have one or two problems/challenges
- Each team should have 3—4 members
- Project may be supported by industry with awards
- Evaluated by professor, TAs, guest judges from industry, and you (peer review)!
- The problems/challenges will be announced around the week of mid exam

# Reference Materials

- Reference books



<http://szeliski.org/Book/>

- And papers in CVPR, ICCV, ECCV, BMVC, WACV, ACCV, ....

# 加簽規則

- 請慎重考慮.....
- 以教室容量為限，可加簽約40位同學
- 篩選順序
  - 電資學院及重點科技學院(含輔系) > 工學院 > 理學院 > 其他
  - 博班 > 碩二 > 碩一 = 大四 > 大三 > 大二 > 大一
- 想加簽的同學，請於第二節上課之前填寫好下列表單：
  - <https://forms.gle/FQTK4rpkRAVhCG4b9>
- 有選上的同學將在一週內寄送授權碼

