Computer Vision II: Recognition

Shenghua Gao

Email: gaoshh@shanghaitech.edu.cn

Today’s Class

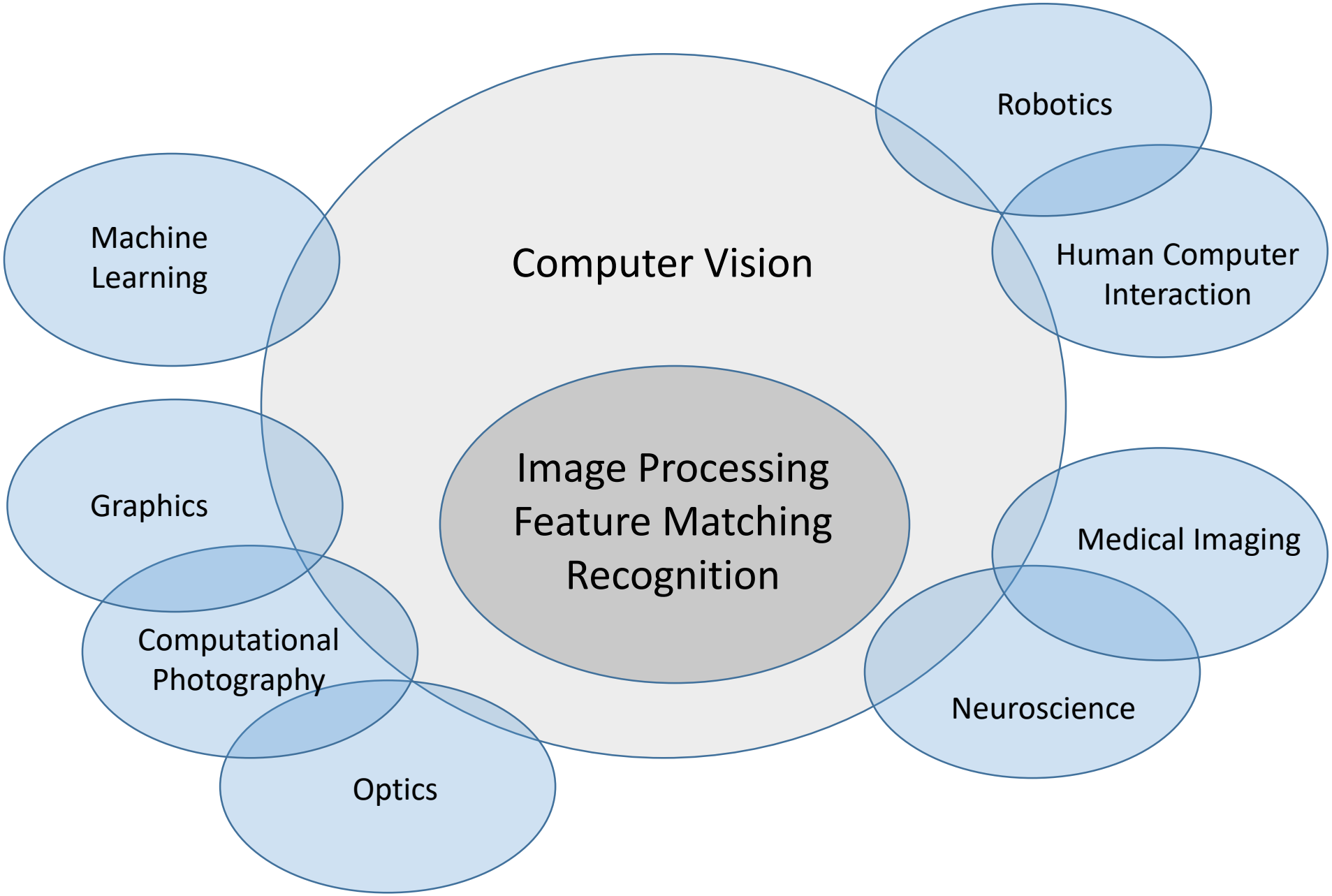
- Specifics of this course
- What is Computer Vision?
- What is the applications of computer vision?
- Course Contents.

Relationship with several nearby courses

- Computer Graphics: Models to Images
- Computer Vision: Images to Models

  Computer vision I: Image formation, 3D reconstruction

  Computer vision II: Recognition


Robotics
Machine Learning
Computer Vision
Human Computer Interaction
Image Processing
Feature Matching
Recognition
Graphics
Medical Imaging
Computational Photography
Neuroscience
Optics

Prerequisite courses

I) **Digital Image Processing/Computer Vision I (undergraduate courses*)**
   I) **If you have never attended these courses before, I strongly recommend you register the Digital Image Processing course first for graduates..**

IV) Programing skills (OpenCV, Matlab, Python)

Course Evaluation

- **Homework (Implement about 5 papers, 20%)** (about 2 weeks per paper*)
  - saliency detection
  - Image Segmentation
  - Image Classification
  - Object Detection
  - Pose estimation
- 上课签到（**course attendance,10%)**
- Course project (70%).
  - 每人一个project，独立完成，要有技术创新或者应用创新
  - 提交一个PPT和一个report，report按照ICPR的模板和要求

课件发布：助教发布于piazza

课程讨论：上课时间微信群讨论，助教负责课程的签到

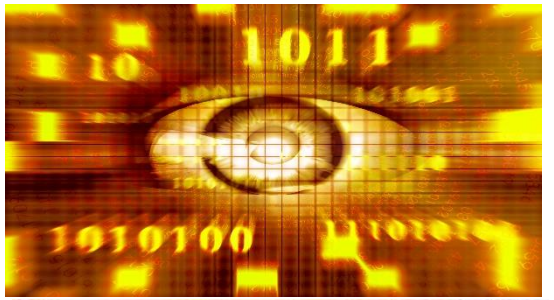
视频上传：学校的网盘

2020春计算机视觉II

该二维码7天内(3月3日前)有效，重新进入将更新

What is Computer Vision?

- What are examples of computer vision being used in the world?



Computer Vision

Make computers understand images and video.

What kind of scene?

Where are the cars?

How far is the building?

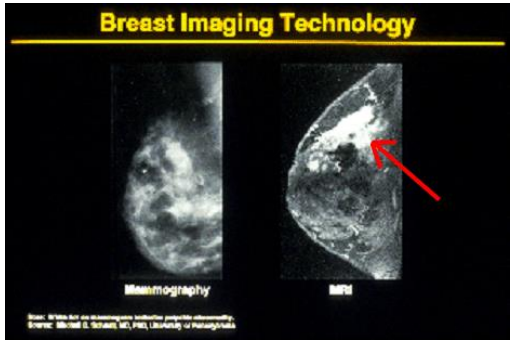
…

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of Macaque brain
  - More human brain devoted to vision than anything else

Why computer vision matters

Safety

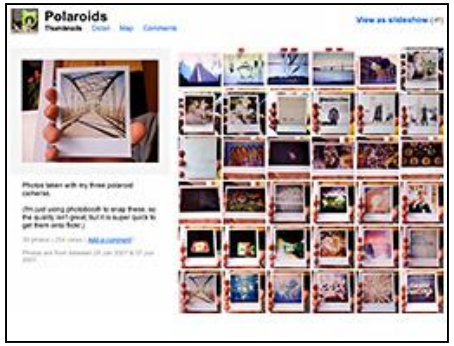


Health

Security

Comfort

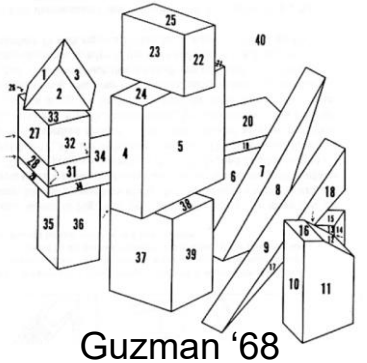
Fun



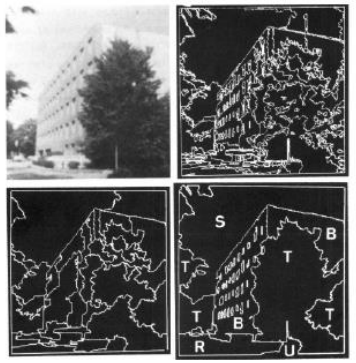
Access

Ridiculously brief history of computer vision

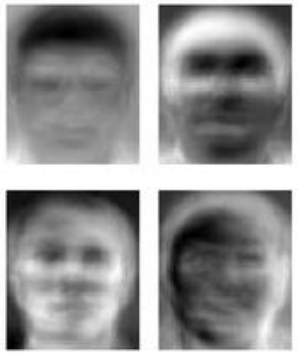
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: deep learning greatly boosts the performance of computer vision.
- 2030's: robot uprising?



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

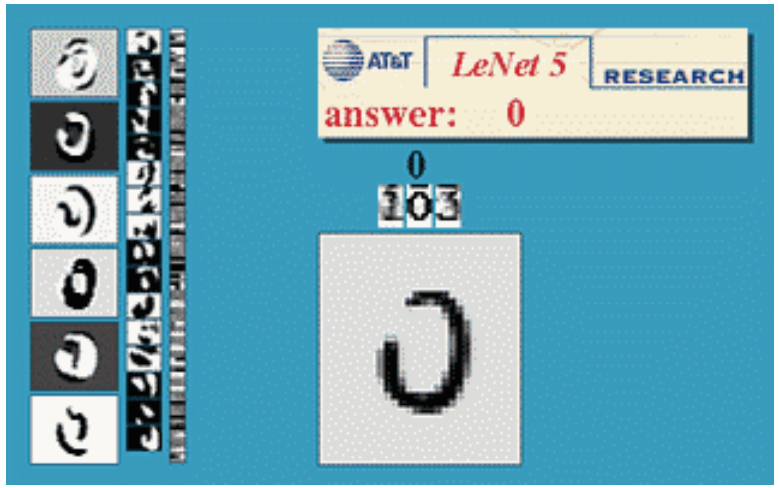
The Applications of Computer Vision

- Optical character recognition
- Face detection
- Image editing
- Medical image analysis

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/

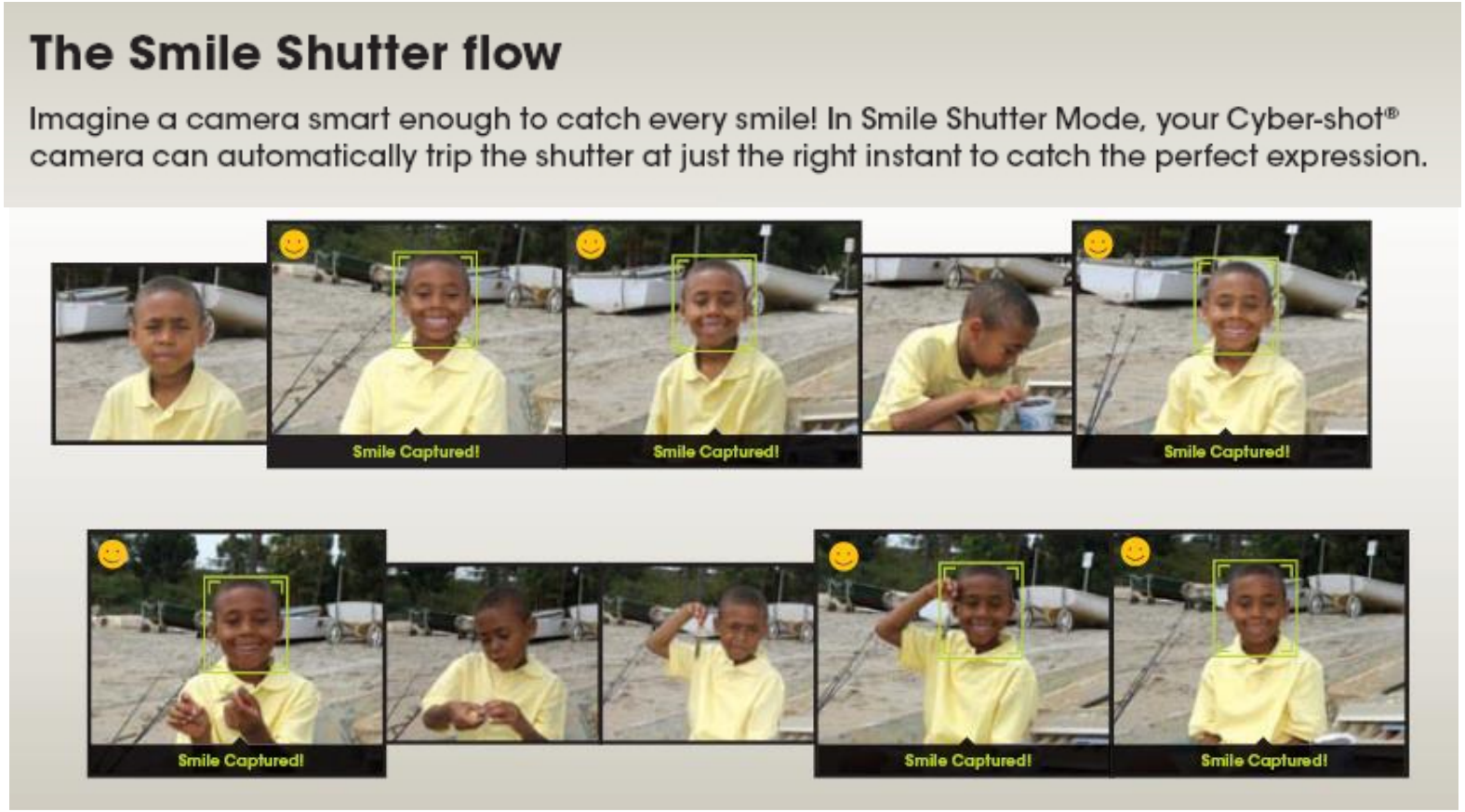
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

车牌识别、银行卡识别等

Face detection

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

Smile detection



[Sony Cyber-shot® T70 Digital Still Camera]

Face expression transfer

- Face2face technology
- https://www.youtube.com/watch?v=ohmajJTcpNk

3D from thousands of images

街景、三维重建

Building Rome in a Day: Agarwal et al. 2009

Object recognition (in supermarkets)

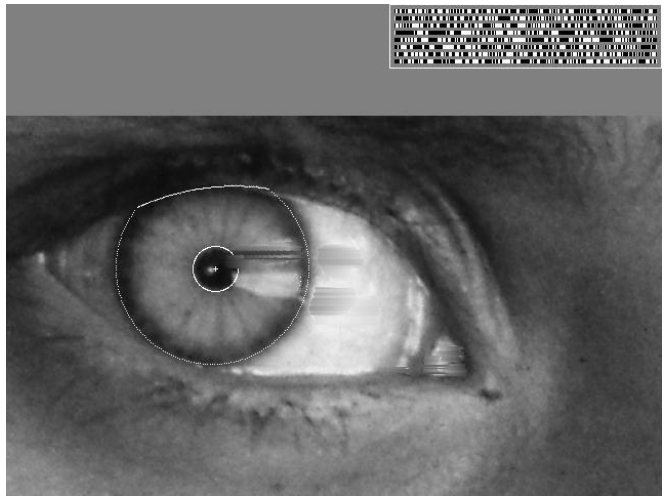
LaneHawk by EvolutionRobotics

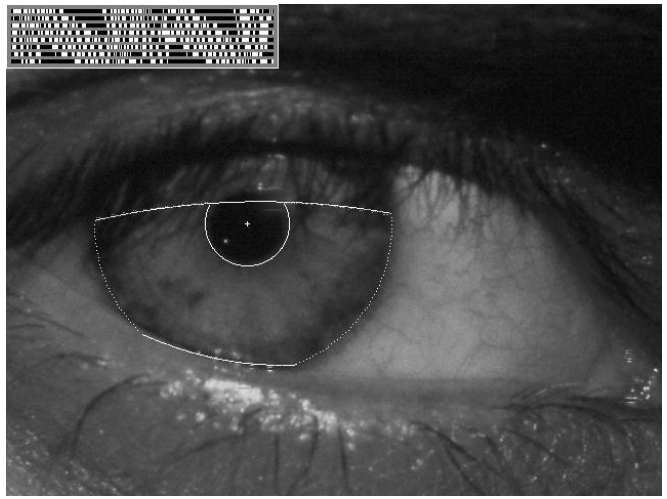
“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… “

Amazon 离线超市

Vision-based biometrics

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





Login without a password…

Fingerprint scanners on many new laptops, other devices

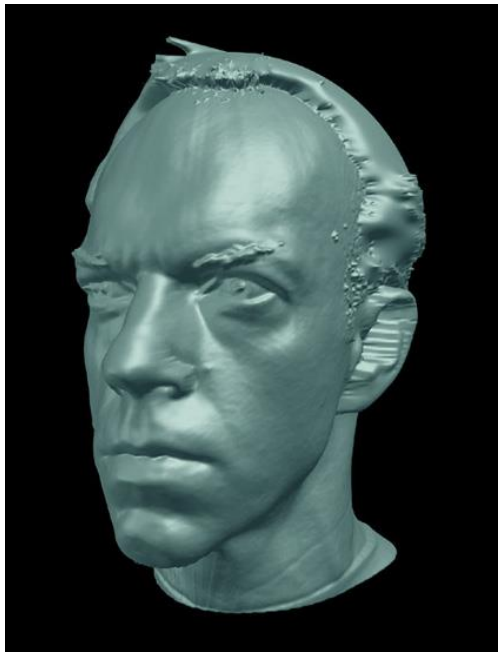
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

[Point & Find](), [Nokia]()

[Google Goggles]()

Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

http://www.sportvision.com/video.html

运动员行为分析

Slide content courtesy of Amnon Shashua

Smart cars

- Google Driverless Car
- Tesla Model S
- Minieye

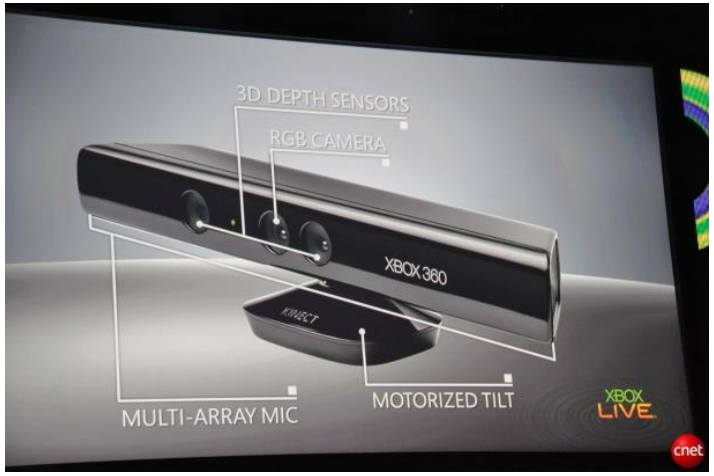
无人驾驶

Google cars

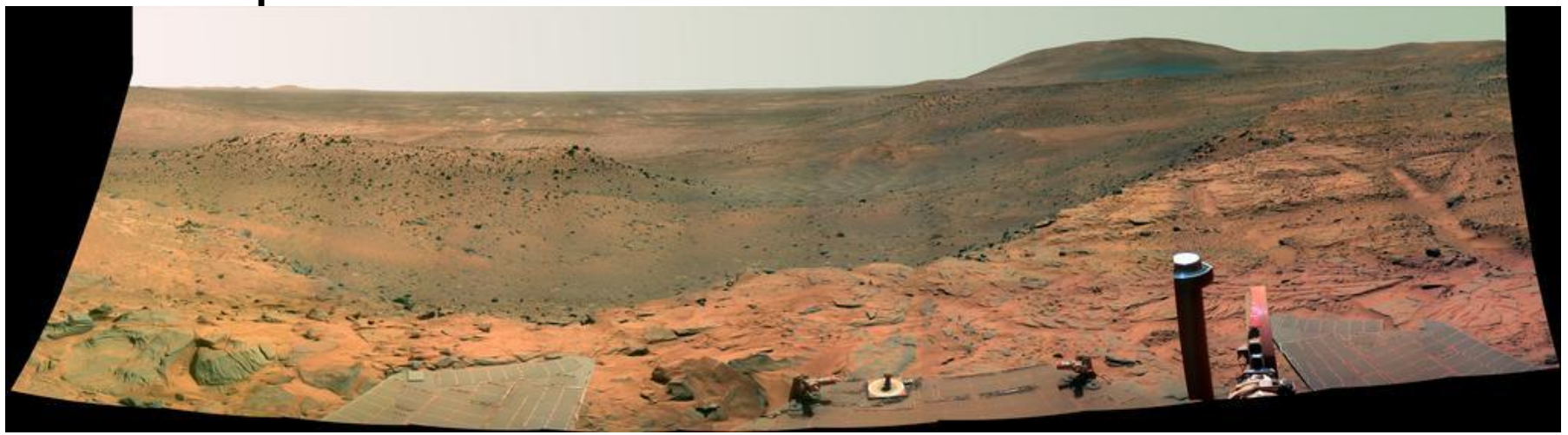
Oct 9, 2010. "Google Cars Drive Themselves, in Traffic". *The New York Times*. John Markoff
June 24, 2011. "Nevada state law paves the way for driverless cars". *Financial Post*. Christine Dobby
Aug 9, 2011, "Human error blamed after Google's driverless car sparks five-vehicle crash". *The Star* (Toronto)

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



Vision in space



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

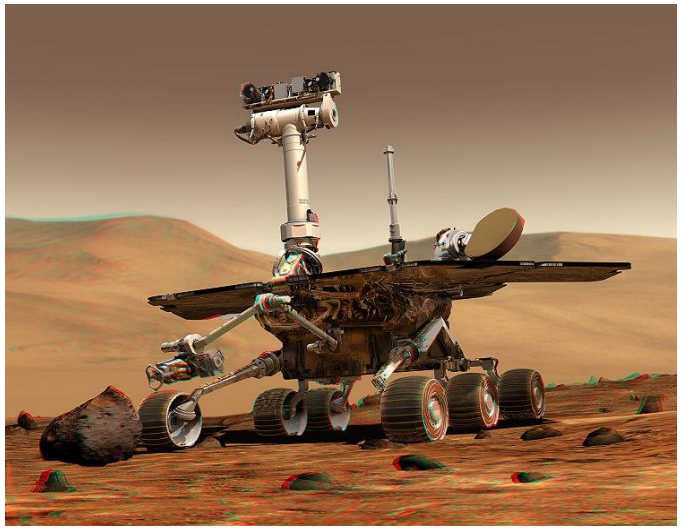
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.

Industrial robots

Vision-guided robots position nut runners on wheels

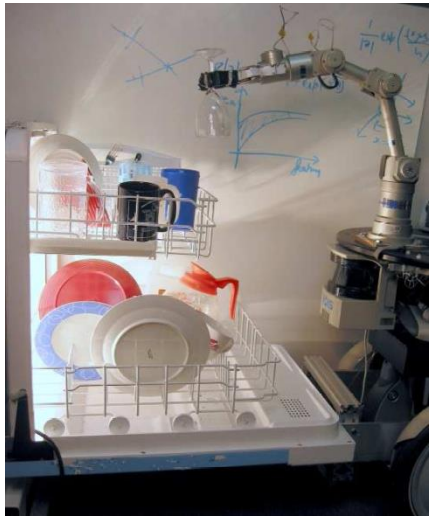
工业机器人

Mobile robots



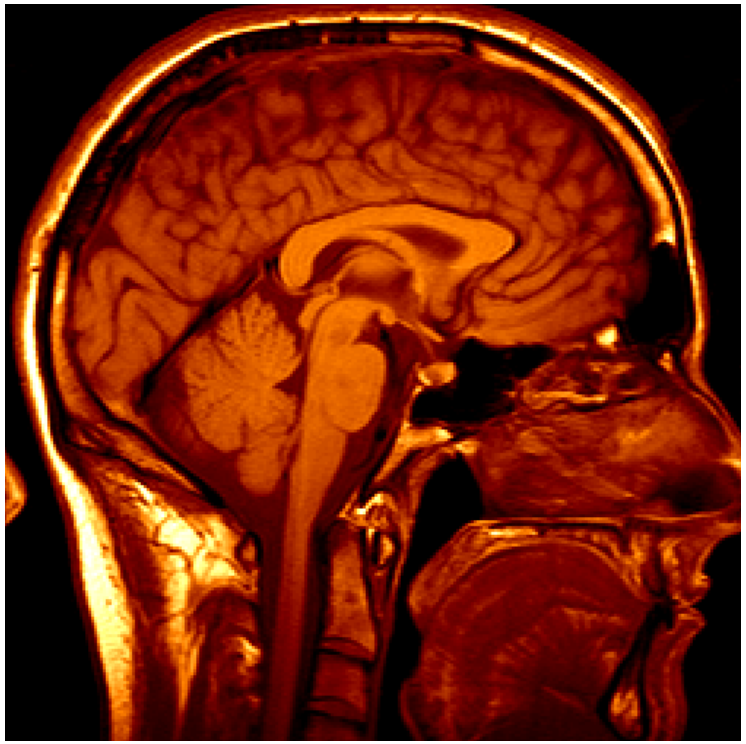
NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

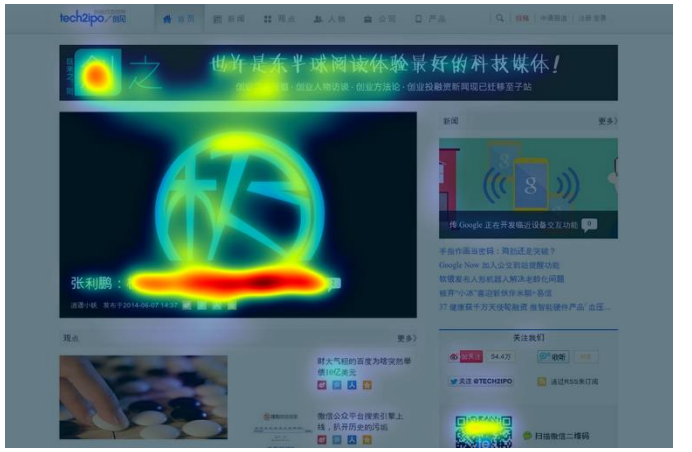
Image guided surgery
Grimson et al., MIT

皮肤癌检测、心率检测、CT、MRT， PET，肿瘤检测、蛋白质结构预测、眼病检测等

Course contents

- **saliency detection**
  - **unsupervised saliency detection/salient object discovery)**
- **segmentation**
  - **(unsupervised image segmentation, semantic segmentation)**
- **object detection**
  - **(face detection/general object detection, supervised methods)**
- **object recognition**
  - **(face recognition/general image classification, (un)supervised methods)**
- **Image generation**
  - **Generative adversarial network, style transfer**
- **video processing**
  - **(tracking/event classification(supervised))**

Saliency detection



Application of computer vision

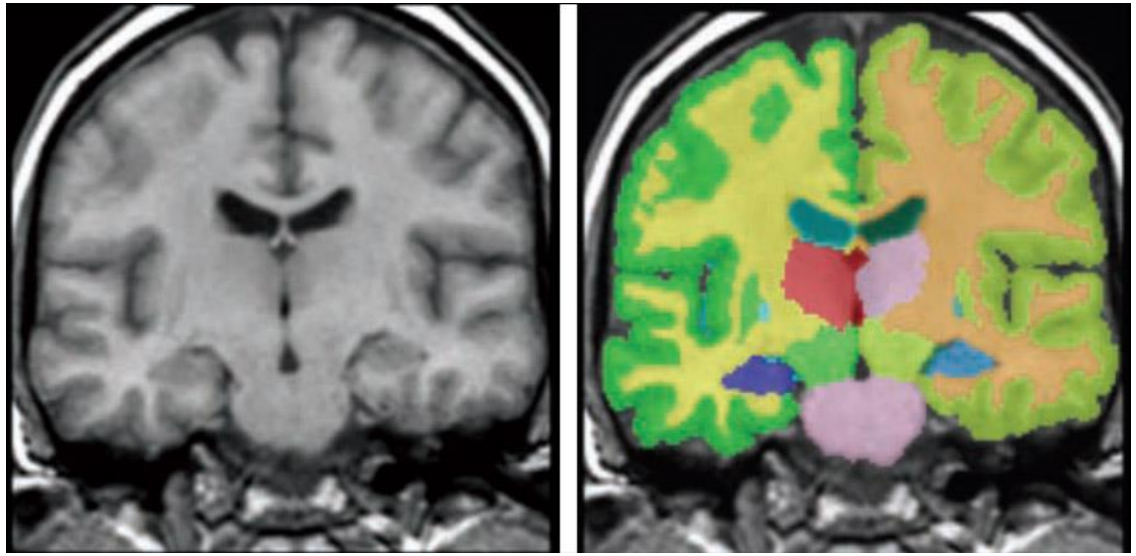
Content based image resizing (Image Retargeting)

https://www.youtube.com/watch?v=6NcIJXTlugc

Website Design

Image/Video Compression

Image Segmentation



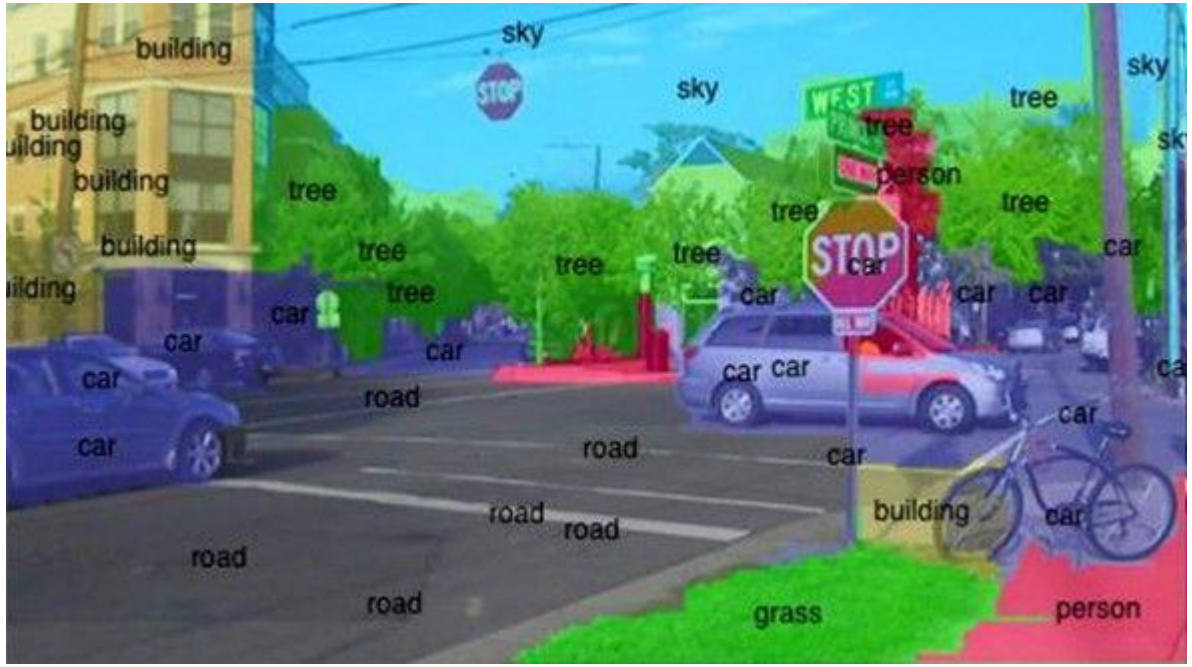
(a) Color Labels (ACA)

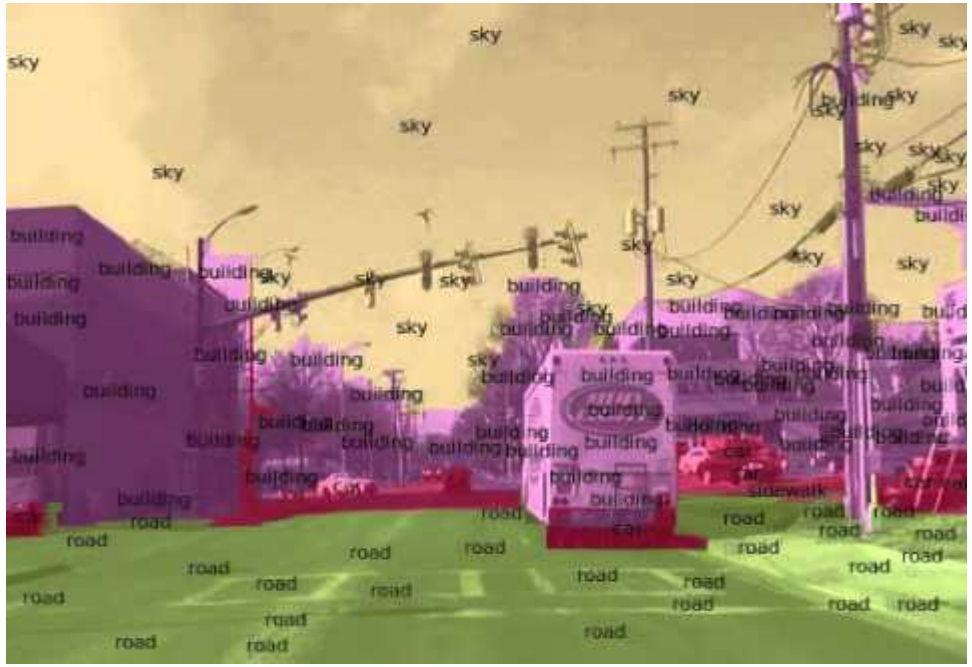
(b) Texture Classes

(c) Crude Segmentation

(d) Final Segmentation

Semantic segmentation





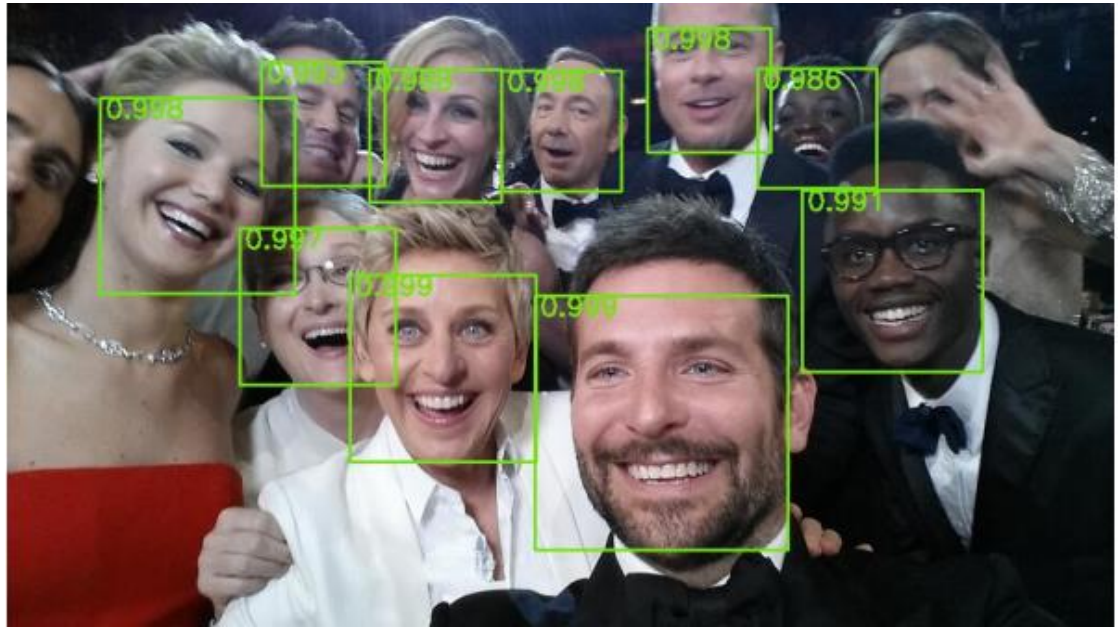
Application

Medical Image Analysis: MRI/CT/PET

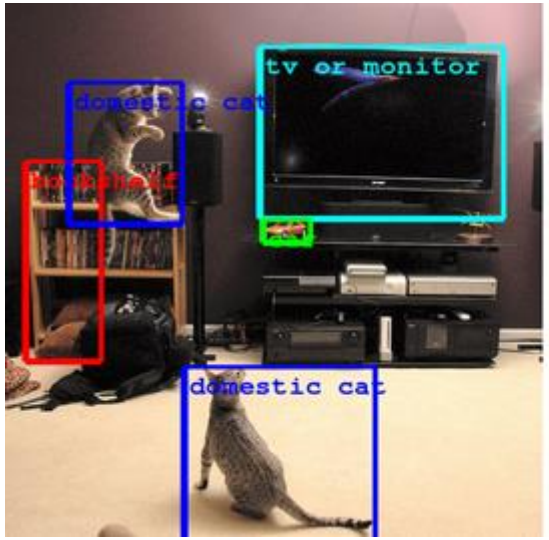
Image retrieval

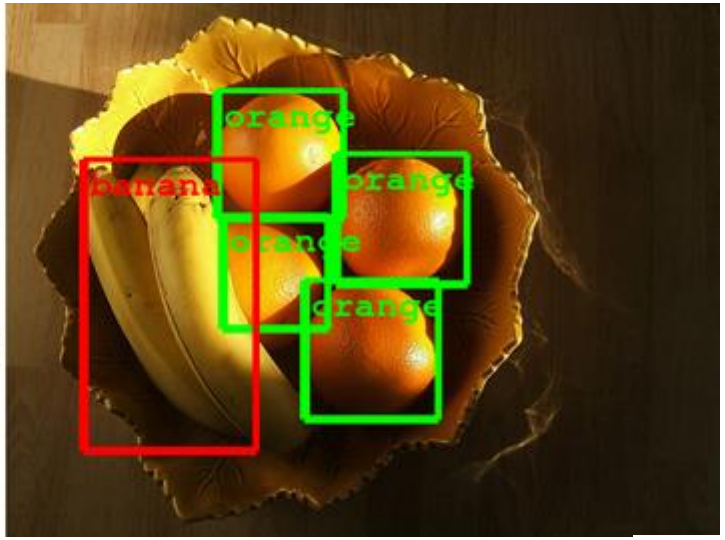
Finding Images by Sketching

Object Detection



Object recognition





Face recognition

Face identification

Who?

A

B

C

...

...

Multiclass classification

Object Identification

Face verification

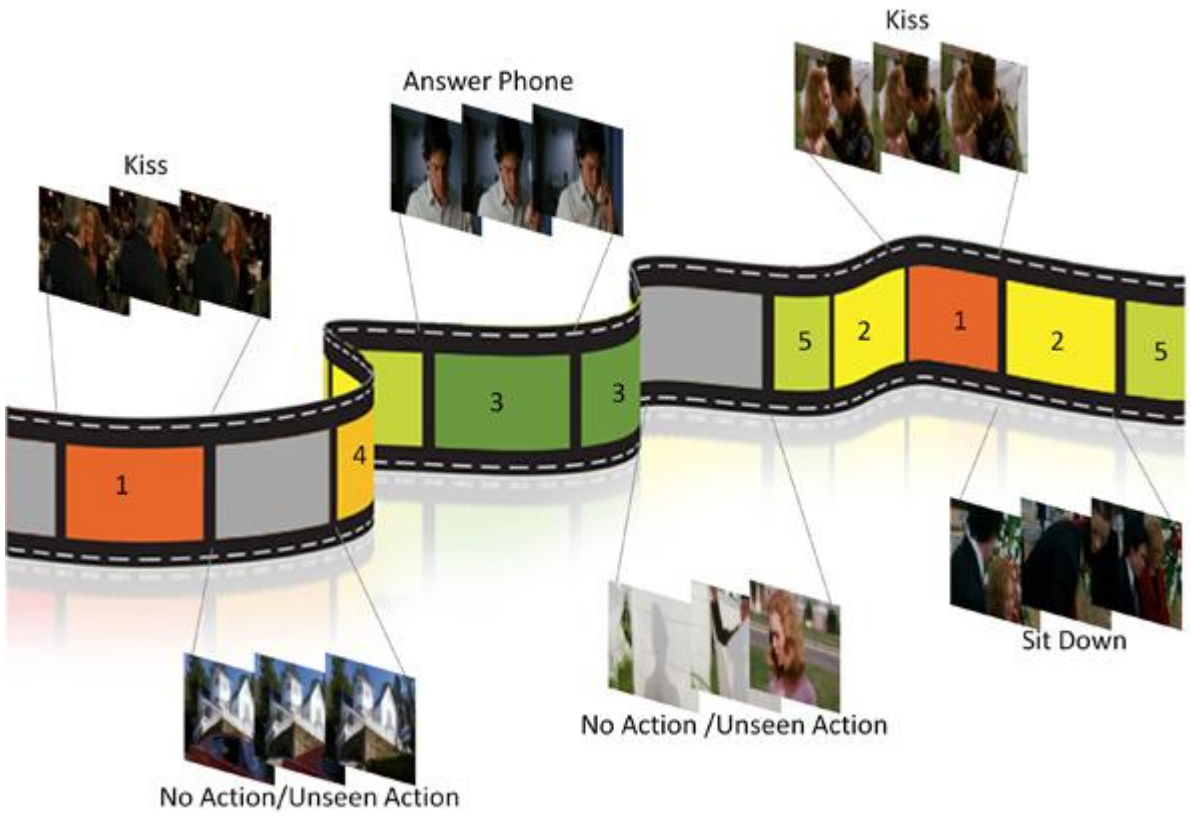
Same person or not.

Binary Result

Generative Adversarial Network

- Face generation
- Style transfer
- Scene generation
- https://www.youtube.com/watch?v=OGGjXG562WU

Video Classification



References

- **1. Computer Vision: A Modern Approach (2nd Edition) , David A. Forsyth , University of California, Berkeley Jean Ponce**
- **2. Computer Vision: Algorithms and Applications (Texts in Computer Science) , by Richard Szeliski (Author), Springer, 2010**
- **3. Hartley and Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2004**
- **4. Stephen E. Palmer, Vision Science: Photons to Phenomenology, MIT Press, 1999**
- **5. Koller and Friedman, Probabilistic Graphical Models: Principles and Techniques, MIT Press, 2009**

Online sources

- Stanford CS231n: deep learning
- Andrew Ng: Deep Learning
- …