CSE 152: Introduction to Computer Vision

Have you ever used computer vision?

How? Where?

Think-Pair-Share

Jitendra Malik, UC Berkeley

Three ‘R’s of Computer Vision

Jitendra Malik, UC Berkeley

Three ‘R’s of Computer Vision

“The classic problems of computational vision:

reconstruction

recognition

(re)organization.”

Have you ever used computer vision?

How? Where?

Reconstruction? Recognition? (Re)organization?

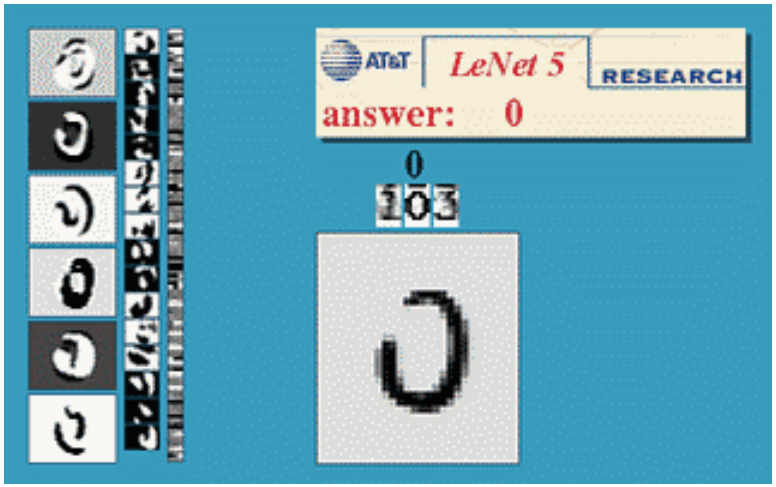
Think-Pair-Share

- Laptop: Biometrics auto-login (face recognition, 3D), OCR
- Smartphones: QR codes, computational photography (Android Lens Blur, iPhone Portrait Mode), panorama construction (Google Photo Spheres), face detection, expression detection (smile), Snapchat filters (face tracking), Google Tango (3D reconstruction),
- Web: Image search, Google photos (face recognition, object recognition, scene recognition, geolocalization from vision), Facebook (image captioning), Google maps aerial imaging (image stitching), YouTube (content categorization)
- VR/AR: outside-in tracking (HTC VIVE), inside out tracking (simultaneous localization and mapping, HoloLens)
- Xbox: Kinect, full body tracking of skeleton, gesture recognition, virtual try-on
- Medical imaging: CAT / MRI reconstruction, assisted diagnosis, automatic pathology, connectomics, endoscopic surgery
- Industry: vision-based robotics (marker-based), machine-assisted router (jig), automated post, ANPR (number plates), surveillance, drones, shopping
- Transportation: assisted driving (everything), face tracking/iris dilation for drunkeness, drowsiness, automated distribution (all modes)
- Media: Visual effects for film, TV (reconstruction), virtual sports replay (reconstruction), semantics-based auto edits (reconstruction, recognition)

Optical character recognition (OCR)

Technology to convert scanned docs to text

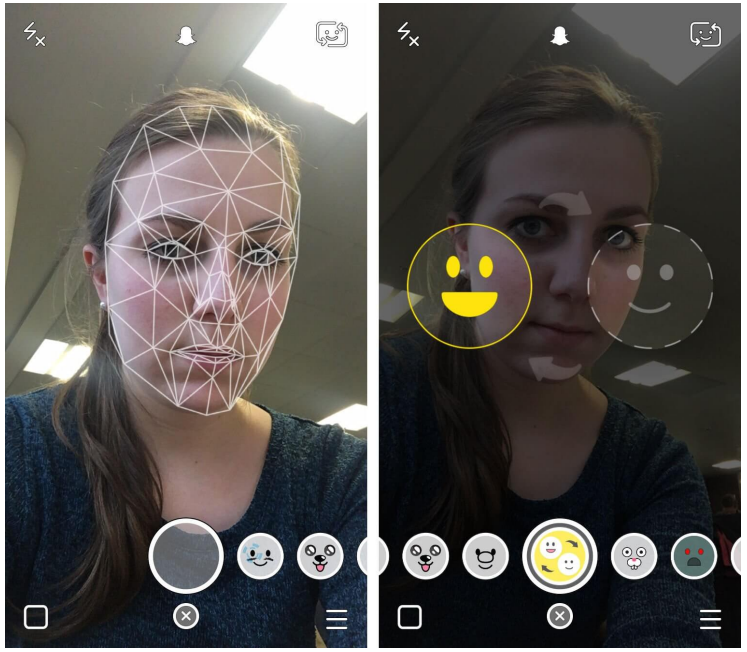
- If you have a scanner, it probably came with OCR software



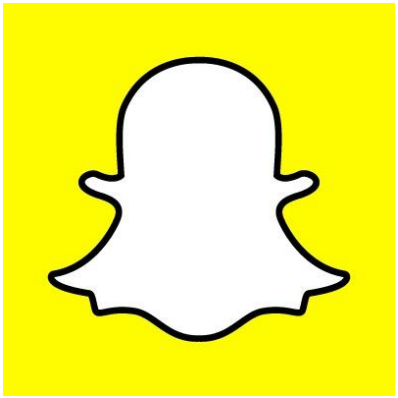
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Face detection



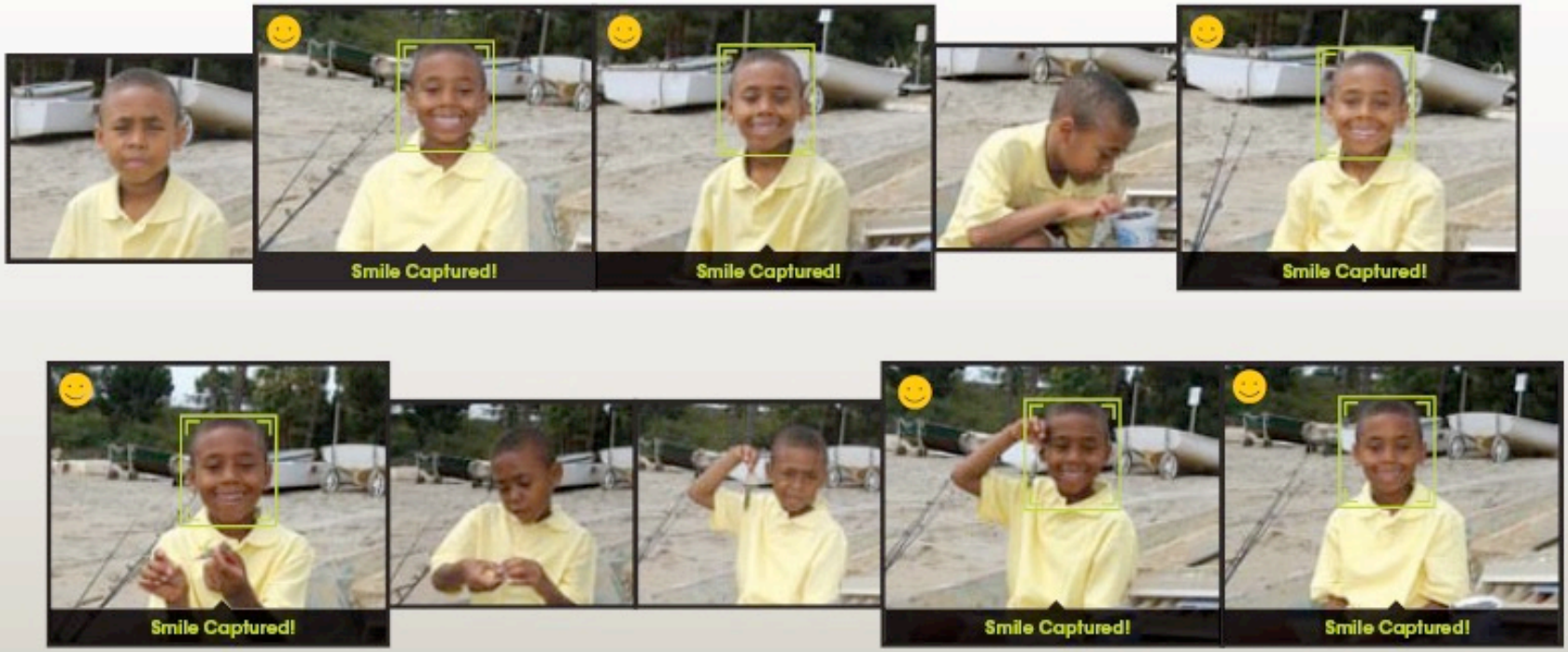
- Almost all digital cameras detect faces
- Snapchat face filters



Smile detection

The Smile Shutter flow

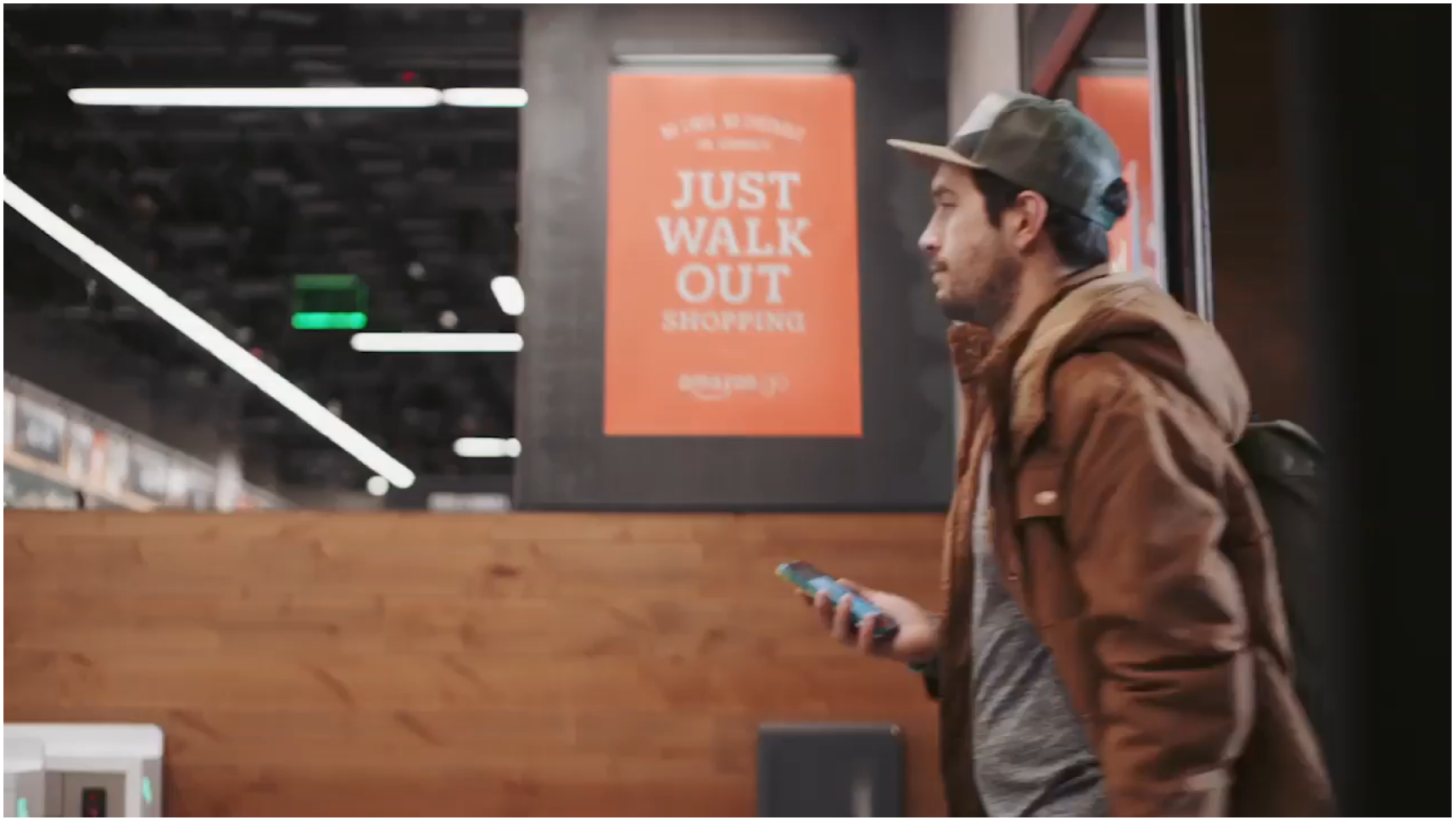
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Sony Cyber-shot® T70 Digital Still Camera

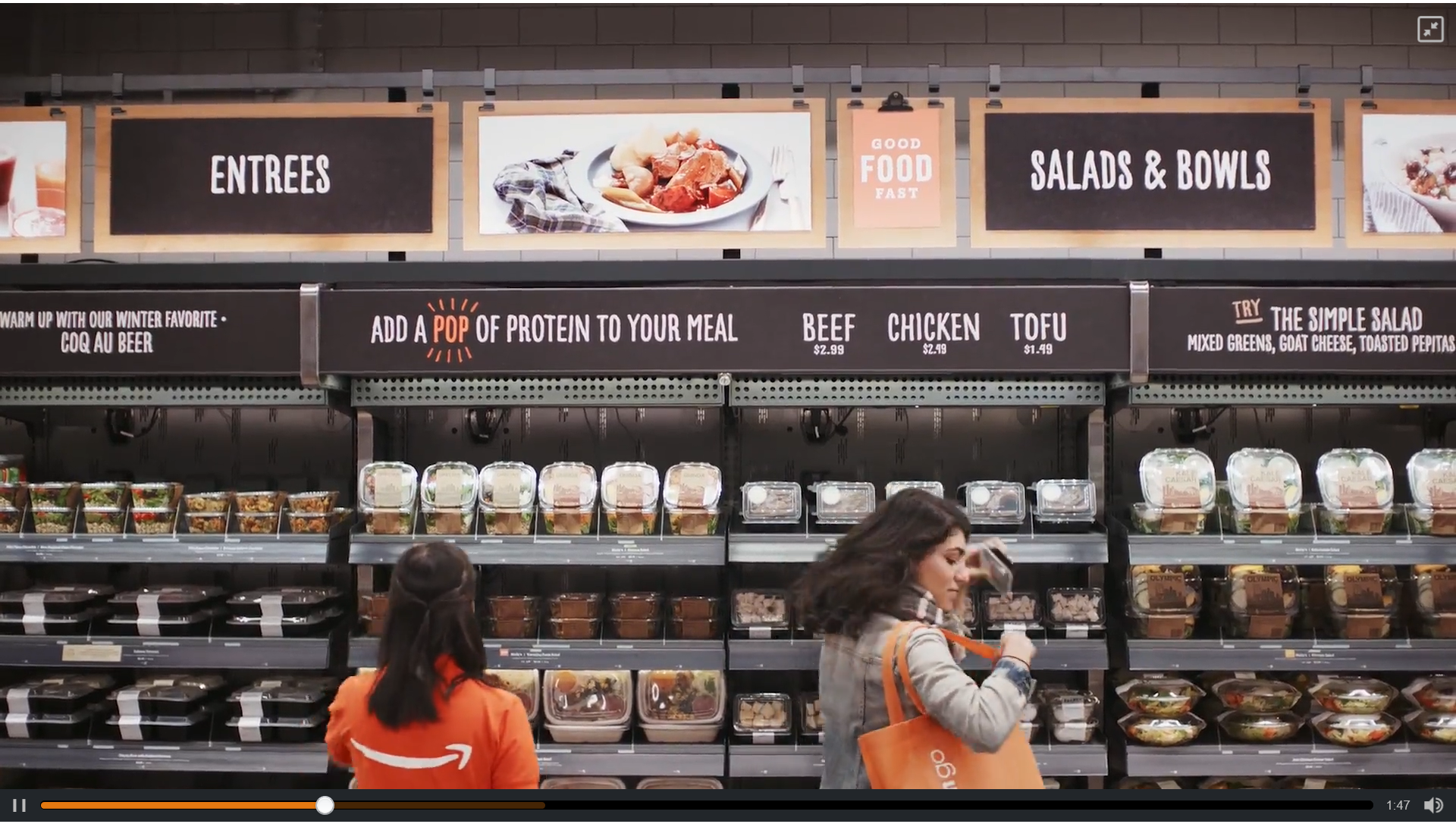
JH

Object recognition (in supermarkets)



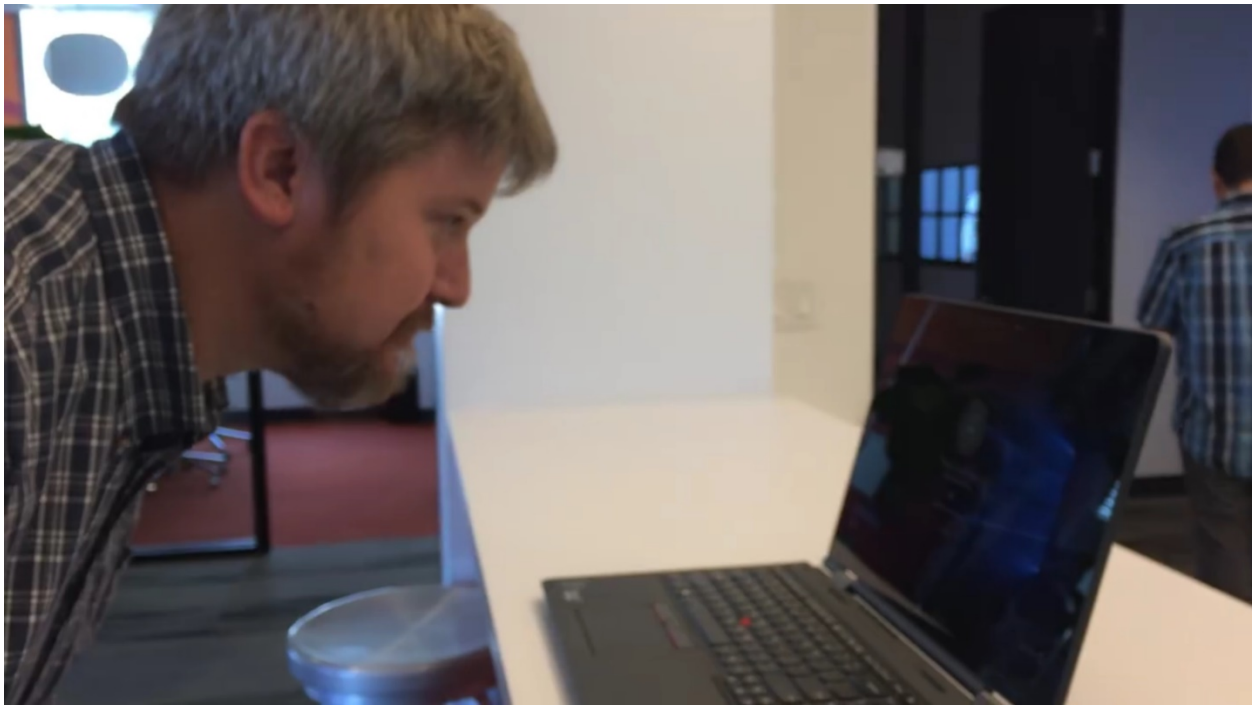
How does it work?

Think-Pair-Share



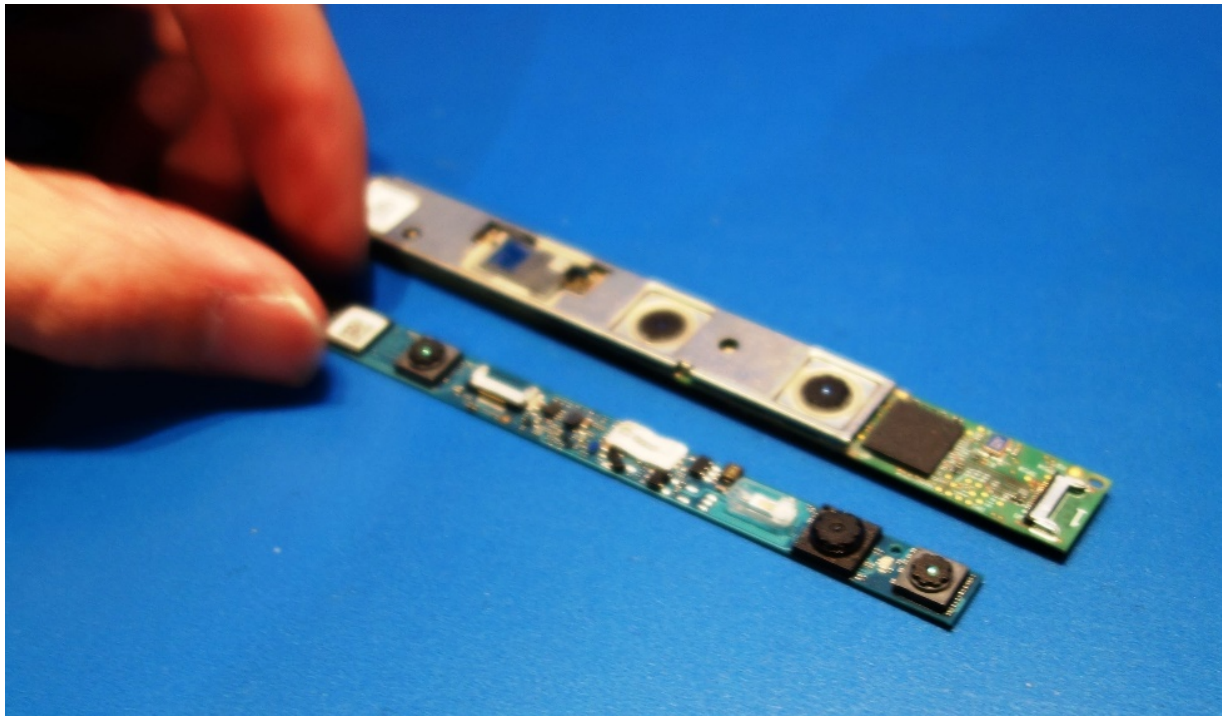

GOOD

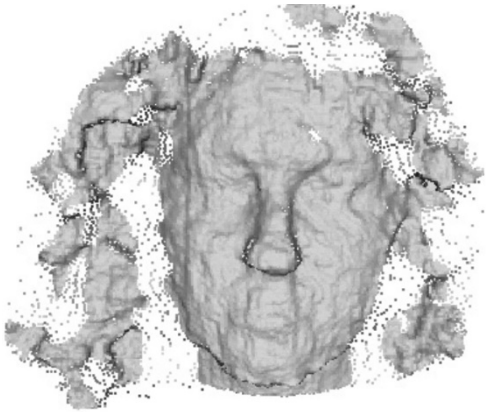
Login without a password…



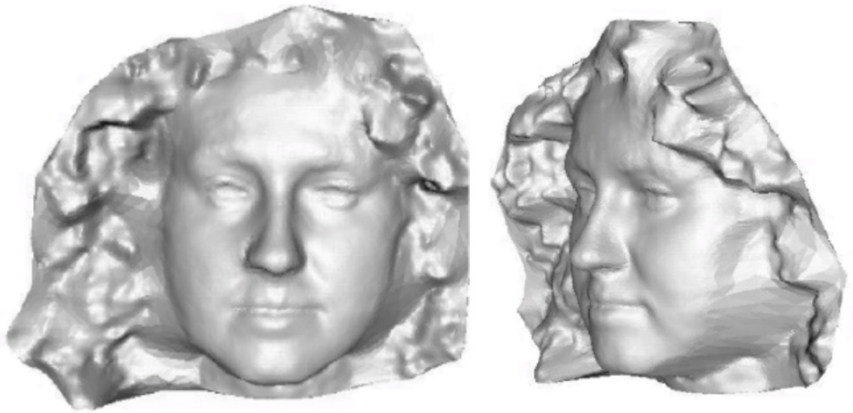
Login without a password…

Login without a password…





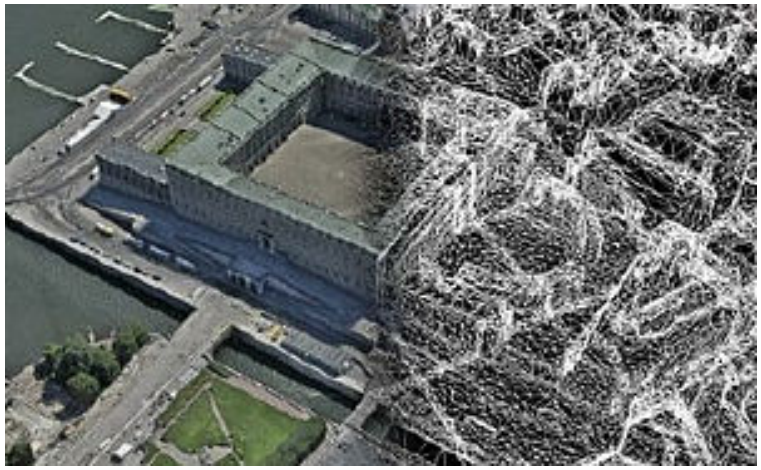
Single depth frame



Reconstructed 3D mesh

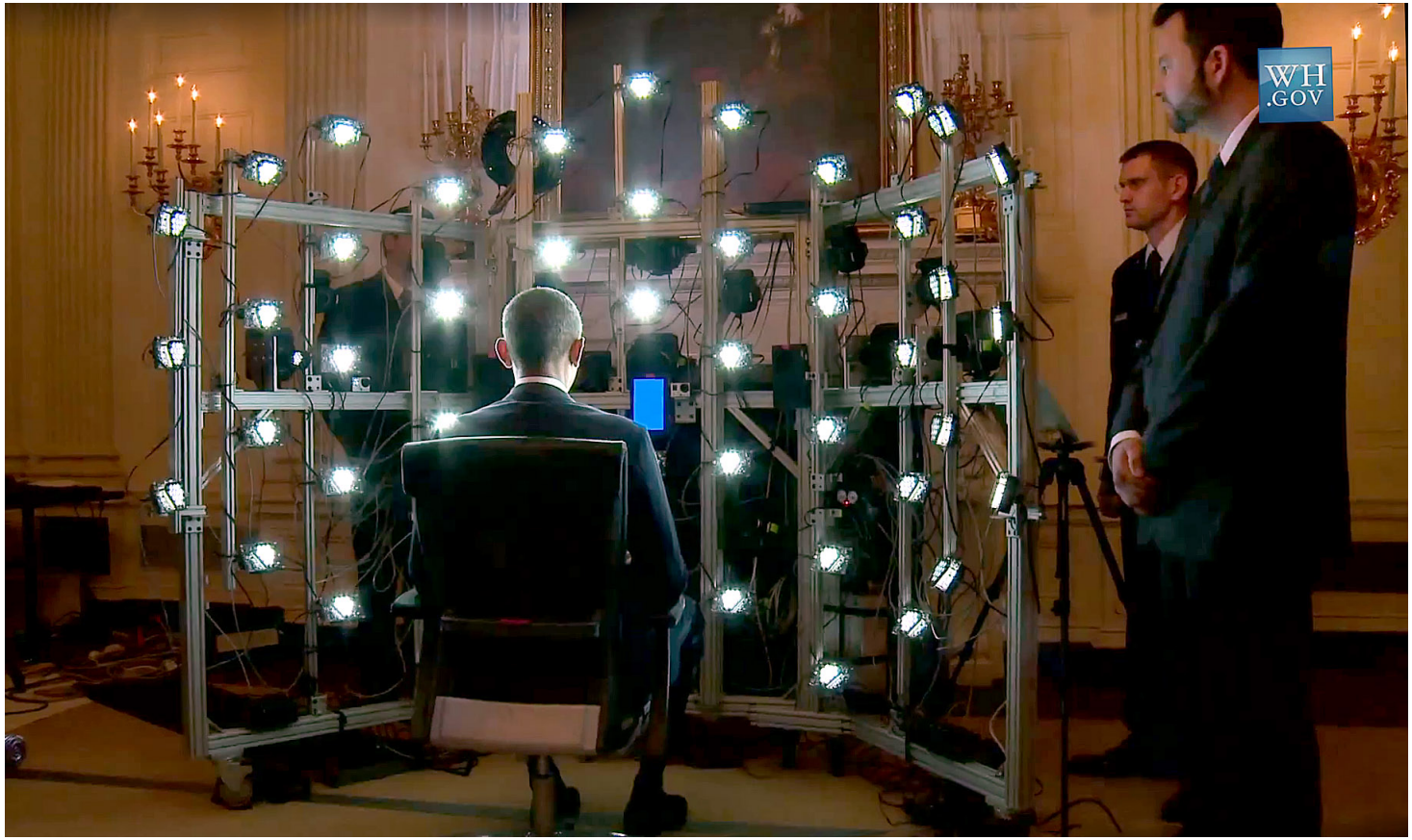
Liang et al. 2014

3D from images



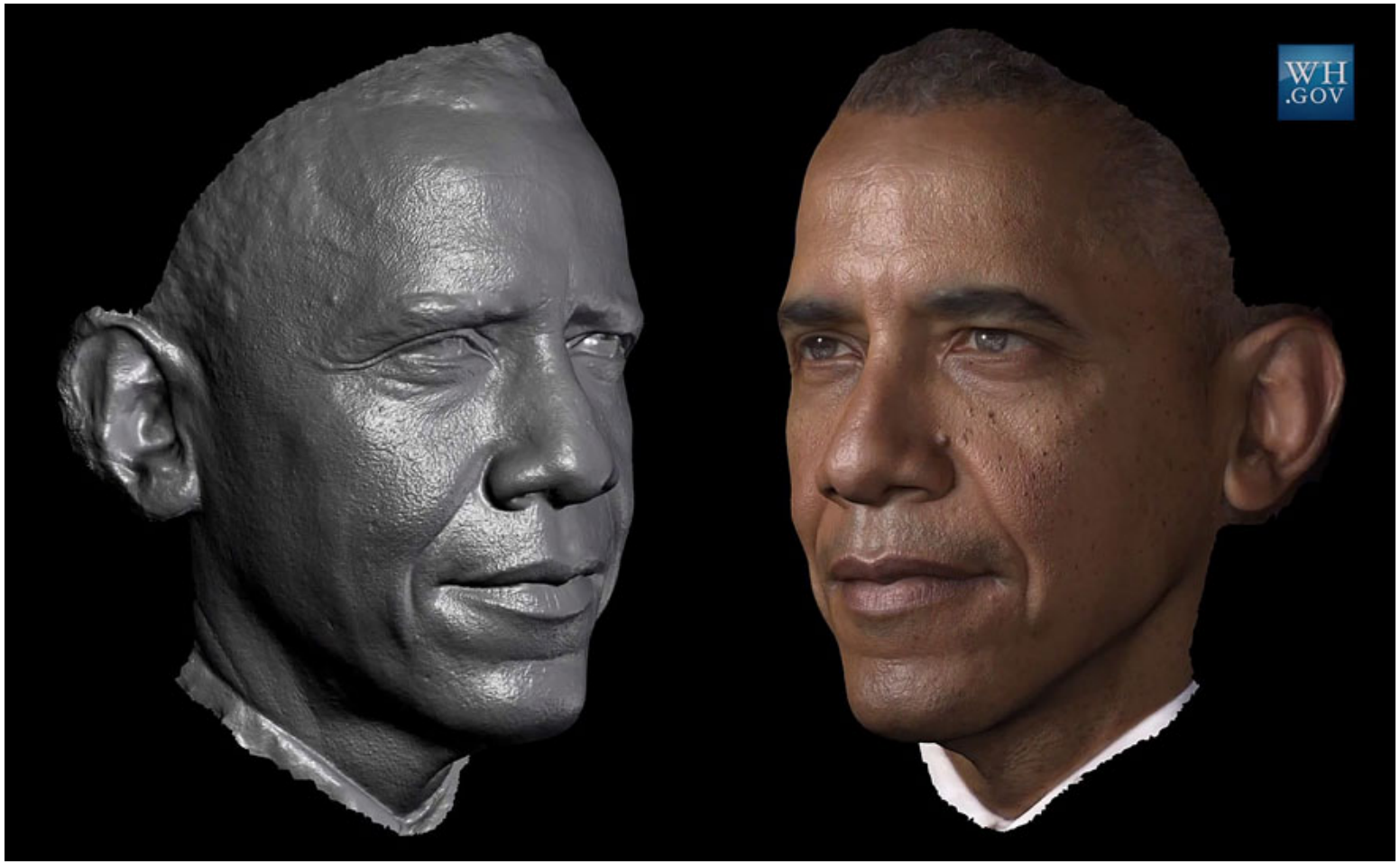
Building Rome in a Day: Agarwal et al. 2009

Human shape capture



Human shape capture

Human shape capture



Human shape capture

Special effects: shape capture

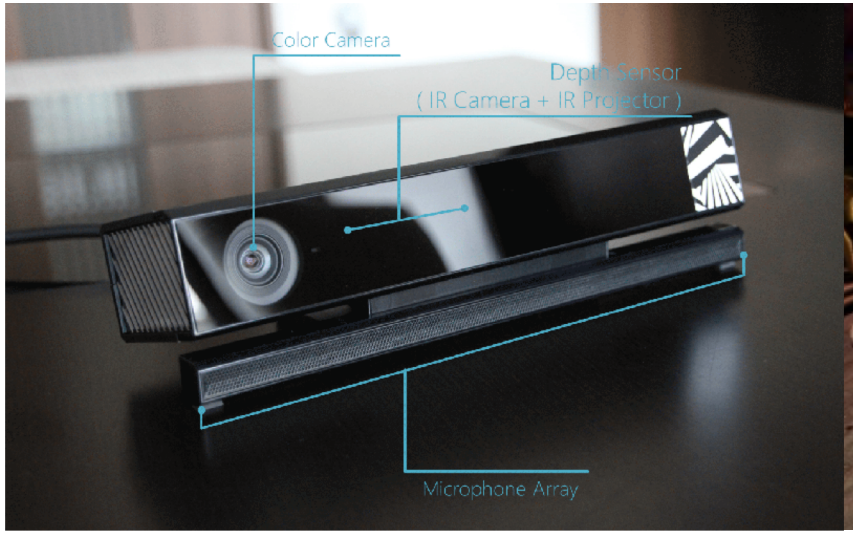
Star Wars: Rogue One – Peter Cushing / Admiral Tarkin

Special effects: shape capture

Special effects: motion capture

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



JH

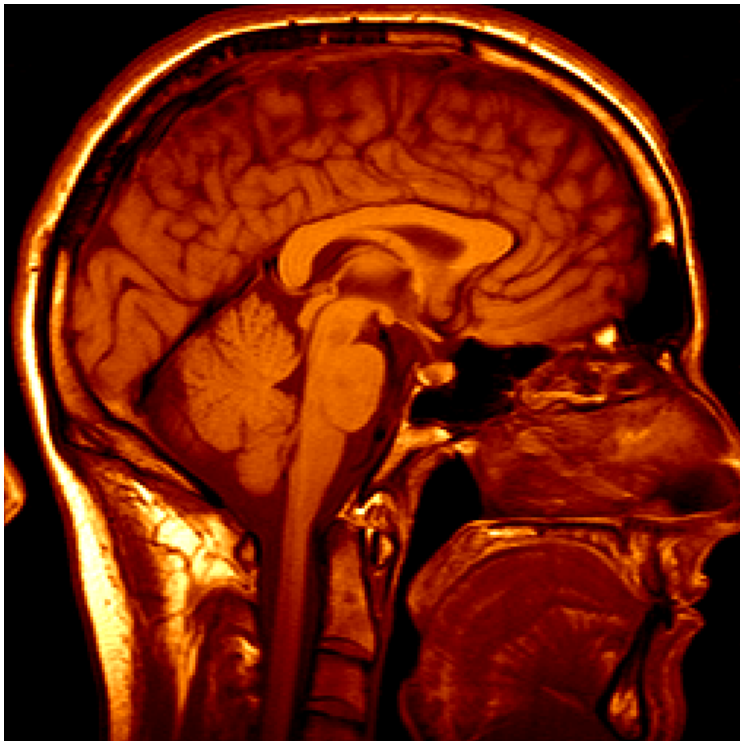
Sports

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

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

JH

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

JH

AutoCars - Uber bought CMU's lab


Google Street View
Explore the world at google.com/streetview

Industrial robots

Vision-guided robots position nut runners on wheels

JH

Vision in space

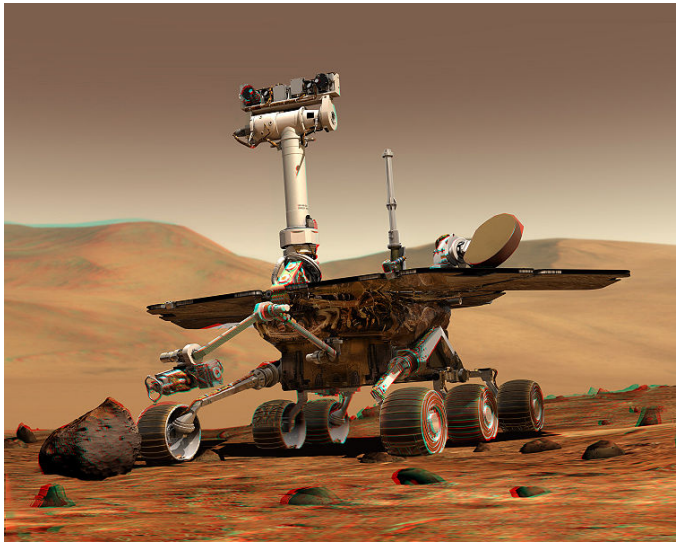
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

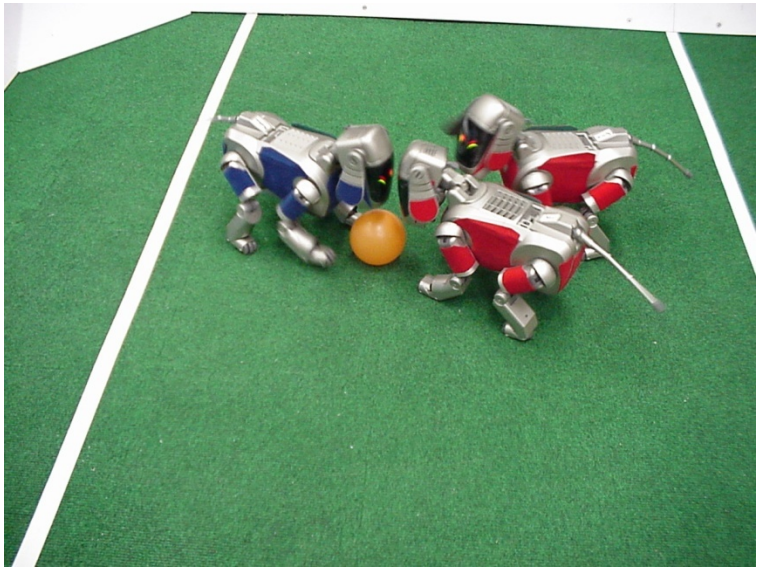
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

JH

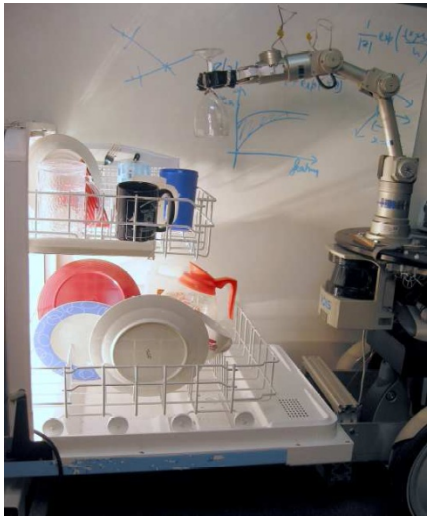
Mobile robots



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



http://www.robocup.org/

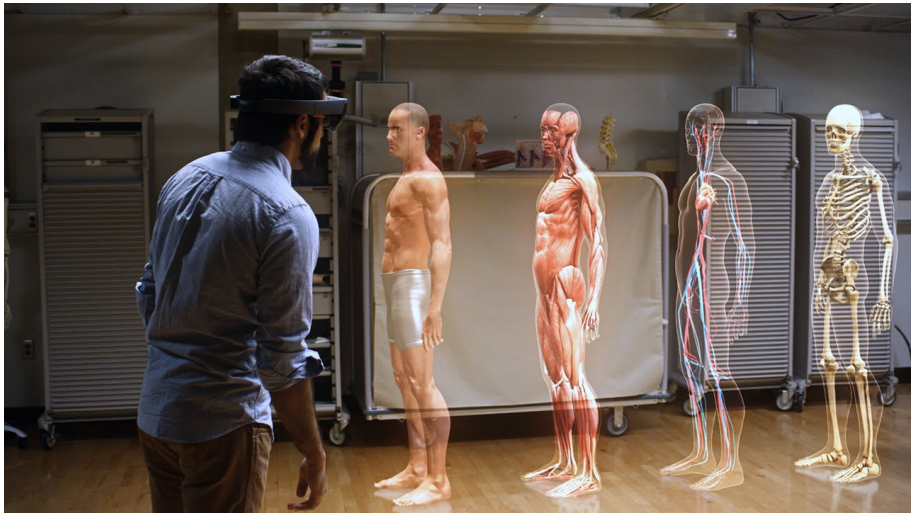


Saxena et al. 2008
STAIR at Stanford

amazon Prime Air

JH

Augmented Reality and Virtual Reality



MS HoloLens, Oculus, Magic Leap,
ARCore / ARKit

Jitendra Malik, UC Berkeley

Three ‘R’s of Computer Vision

“[Further progress in] the classic problems of computational vision:

reconstruction

recognition

(re)organization

[requires us to study the interaction among these processes].”

State of the art today?

With enough training data, computer vision nearly matches human vision at most recognition tasks

Deep learning has been an enormous disruption to the field. Many techniques are being “deepified”.

Computer Vision and Nearby Fields

Derogatory summary of computer vision:

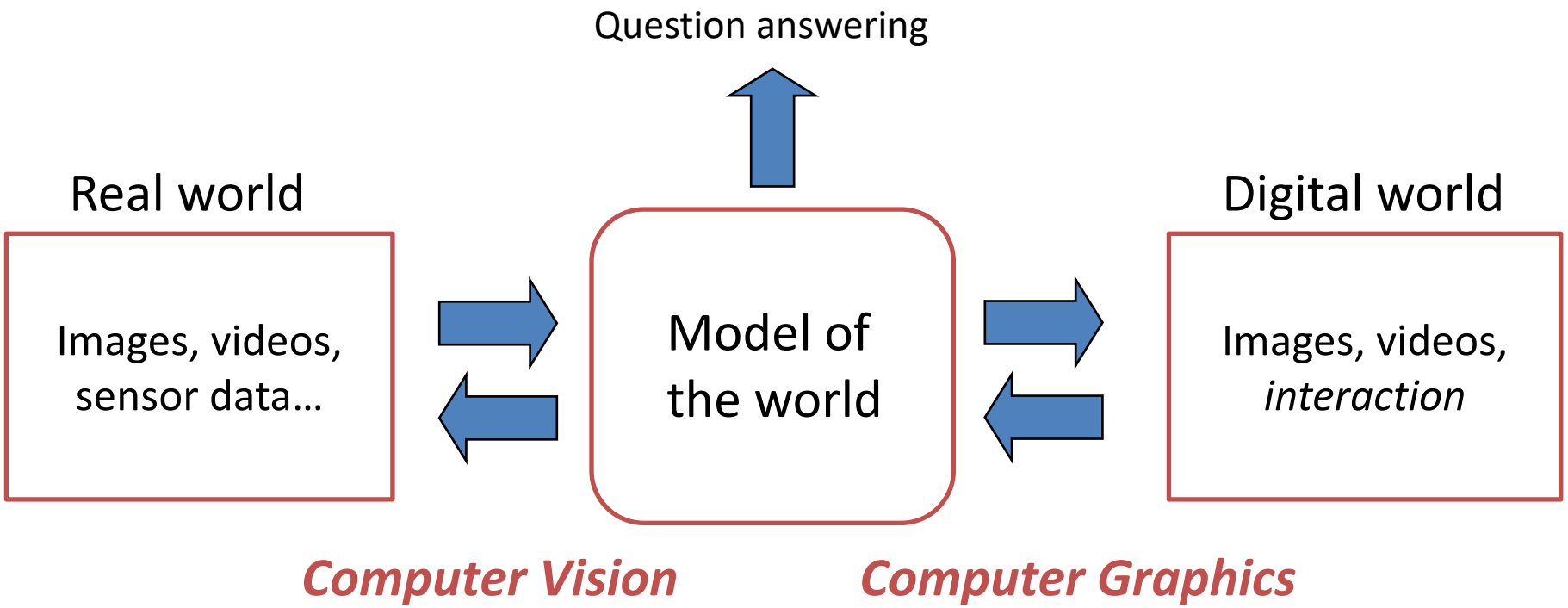
> "Machine learning applied to visual data."

JH

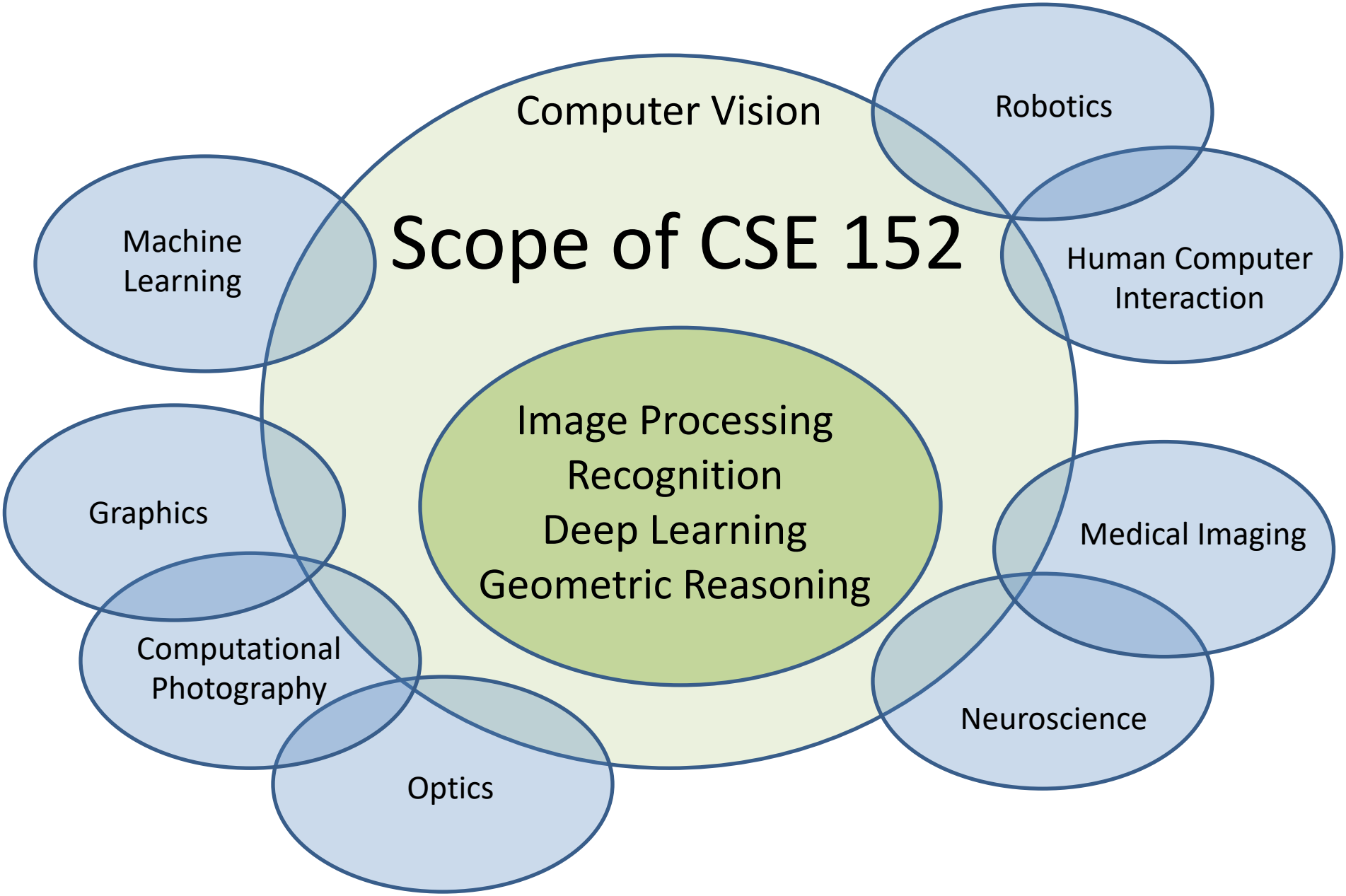
Computer Vision and Nearby Fields

Derogatory summary of computer vision:

"Machine learning applied to visual data."

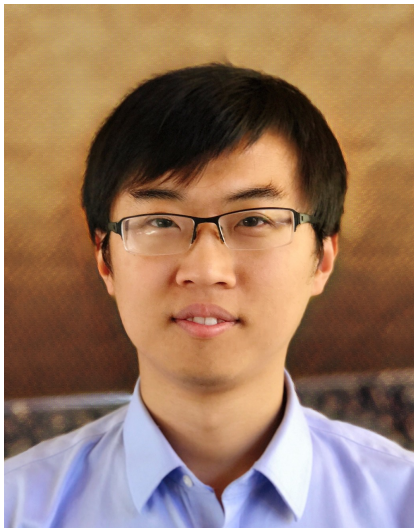


JH

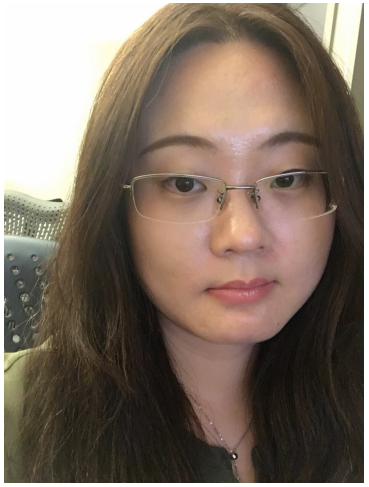

Computer Vision
Robotics
Scope of CSE 152
Machine Learning
Human Computer Interaction
Image Processing
Recognition
Deep Learning
Geometric Reasoning
Graphics
Medical Imaging
Computational Photography
Neuroscience
Optics

JH

BORING COURSE DETAILS



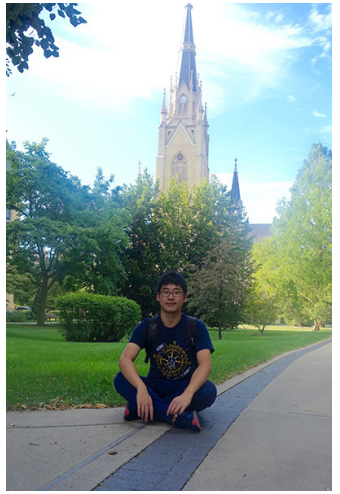
Hao Su



Meng Song

Owen Jow

Zhiran Chen

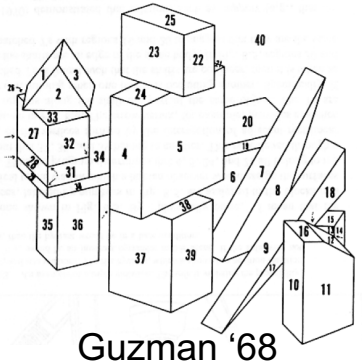


Tongzhou Mu

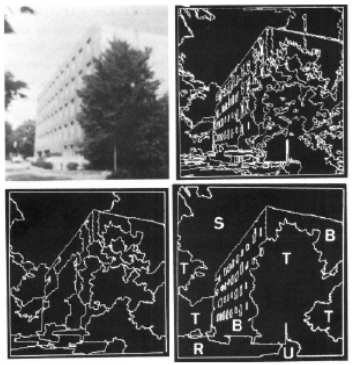
We are here to teach you!

Ridiculously brief history of computer vision

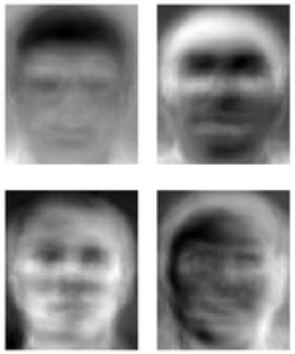
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: Deep learning with ConvNets
- 2027: My very own robot?



Guzman '68



Ohta Kanade '78


Turk and Pentland '91

JH

Course Topics

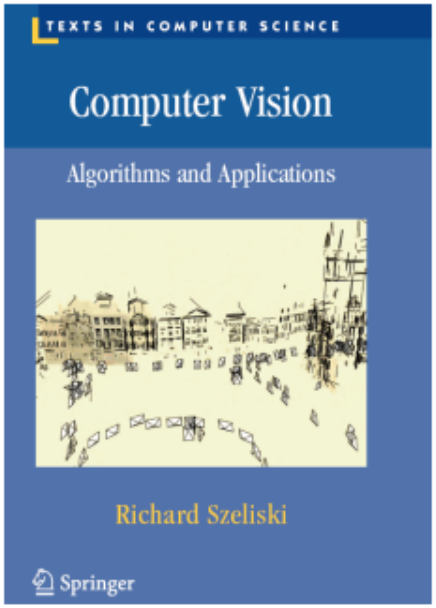
Unit 0	Introduction and Math background
1	Introduction + Exam
2	Color+Math basics
3	Linear algebra
Unit 1	Low-level Vision
4	Frequency Analysis I
5	Frequency Analysis II
6	Filters and Edge
7	Classical Local Features
8	Basics of Deep Feature Learning
Unit 2	Motion
9	Tracking and Optical Flow I
10	Tracking and Optical Flow II
11	Correspondence, RANSAC
Unit 3	3D Vision
12	Mid-term Exam + Stereo
13	Multiview Geometry
14	Structure from Motion
15	Shape from X
Unit 4	Image Understanding
16	Dimension Reduction (PCA)
17	Grouping (K-means, Line fitting, N-cut)
18	Classification
19	Object Detection, Semantic Segmentation
20	Modern Computer Vision and Final Review

Class Intro

- Freshman/Sophomores/Juniors/Seniors/Grads
- Linear algebra
- Probability?
- Graphics course?
- Vision/image processing course before?
- Machine learning?

Textbook

Computer Vision: Algorithms and Applications



http://szeliski.org/Book/

JH

Textbook

- Can I get a PDF of this book?

  No, our contract with MIT Press forbids distribution of too easily copied electronic formats of the book.

- Why are you using HTML format for the web version of the book?

  This format is a sort of weak DRM required by our contract with MIT Press. It's intended to discourage unauthorized copying/editing of the book.

- What is the best way to print the HTML format?

  Printing seems to work best printing directly from the browser, using Chrome. Other browsers do not work as well.

Projects / Grading

- Homework (4 assignments, 40%)
- Mid-term 20%
- Final 40%

- Our homework is MATLAB based