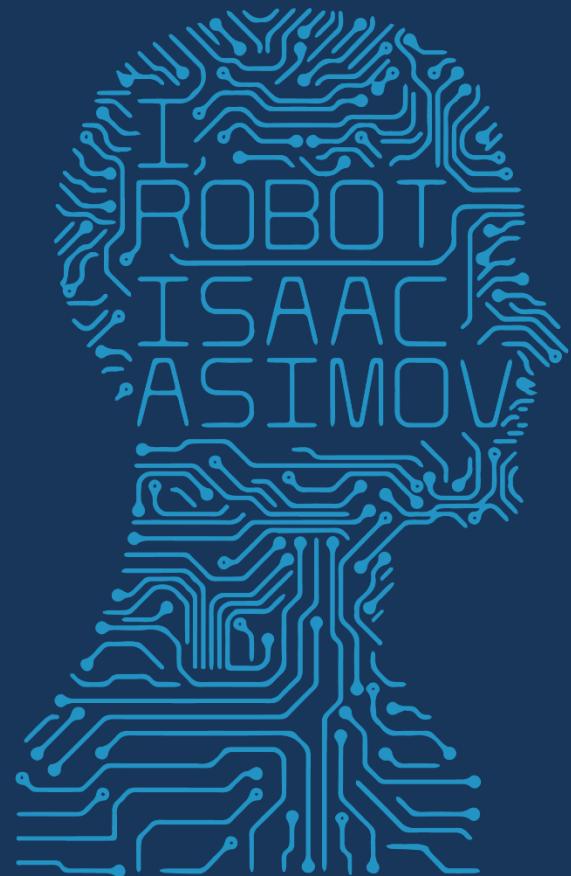


1950

FUTURE VISION



1950
FUTURE VISION



2017 MWF 1PM 368
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?

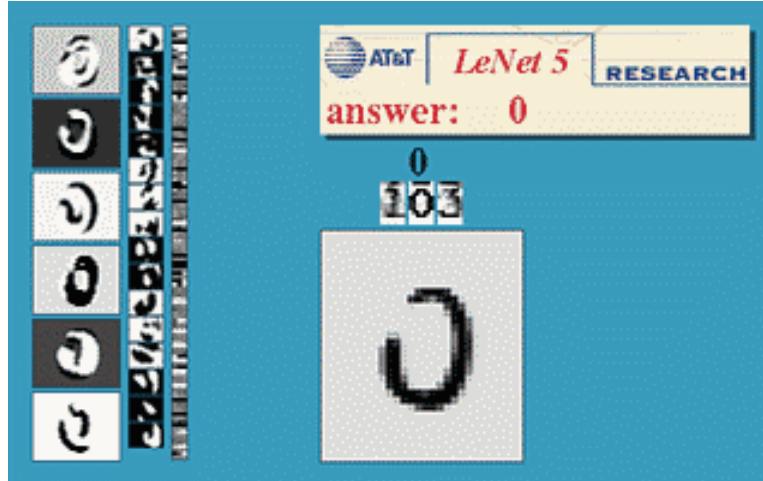
Think-Pair-Share

- 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
- 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
- Transportation: assisted driving (everything), face tracking/iris dilation for drunkenness, drowsiness
- 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

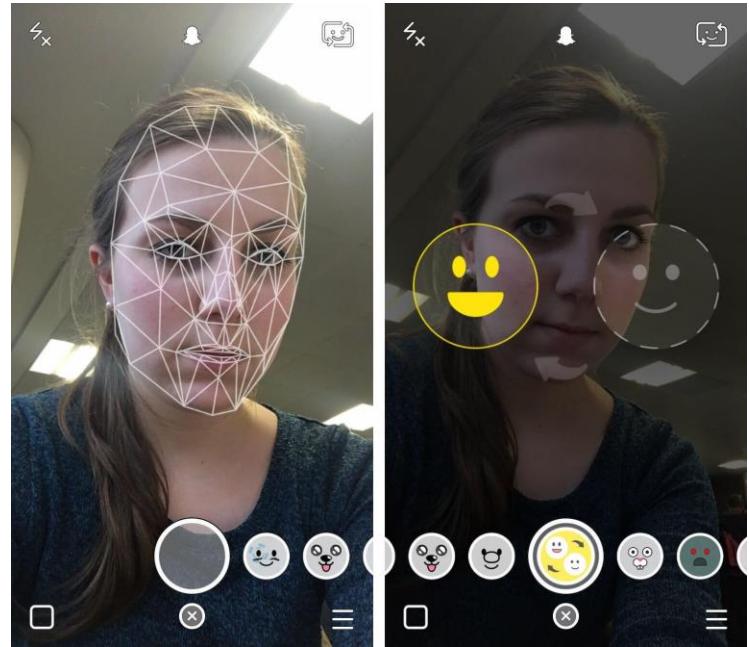


Digit recognition, AT&T labs
<http://www.research.att.com/~yann/>

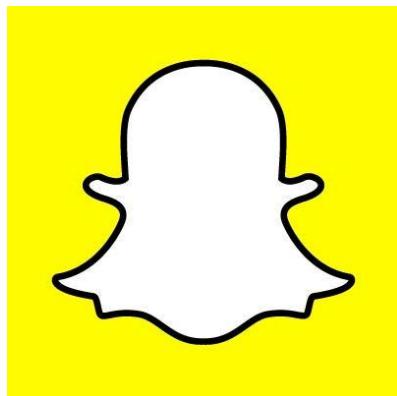


License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

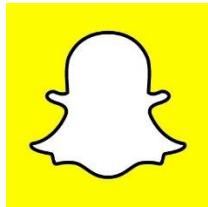
Face detection



- Almost all digital cameras detect faces
- Snapchat face filters



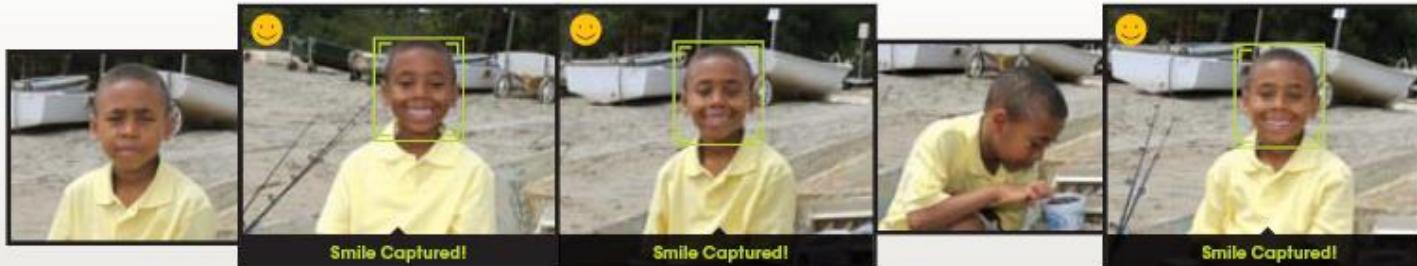




Smile detection

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](#)

Object recognition (in supermarkets)



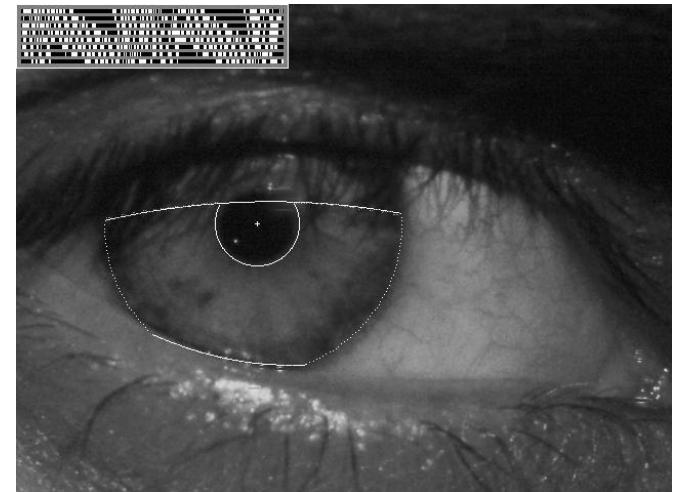
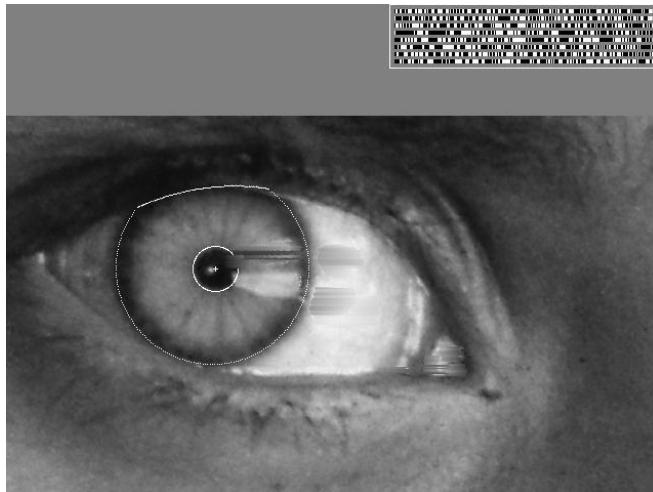
[LaneHawk by EvolutionRobotics](#)

“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... “

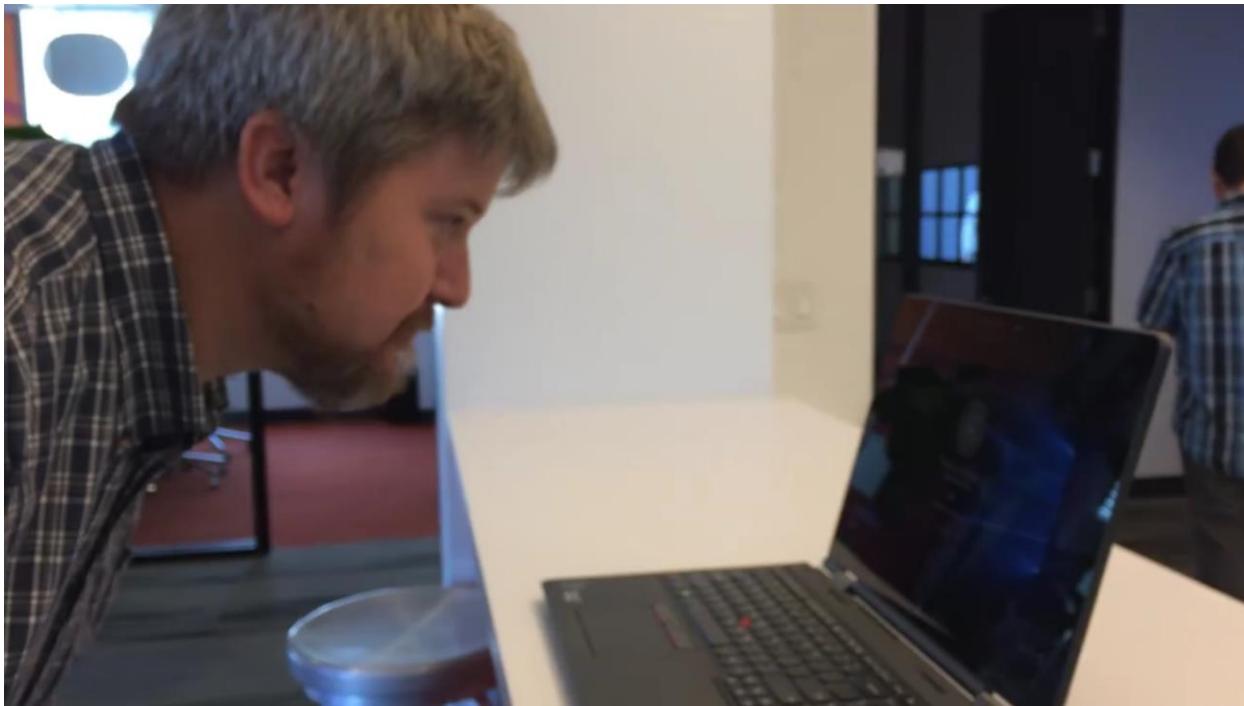
Vision-based biometrics



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



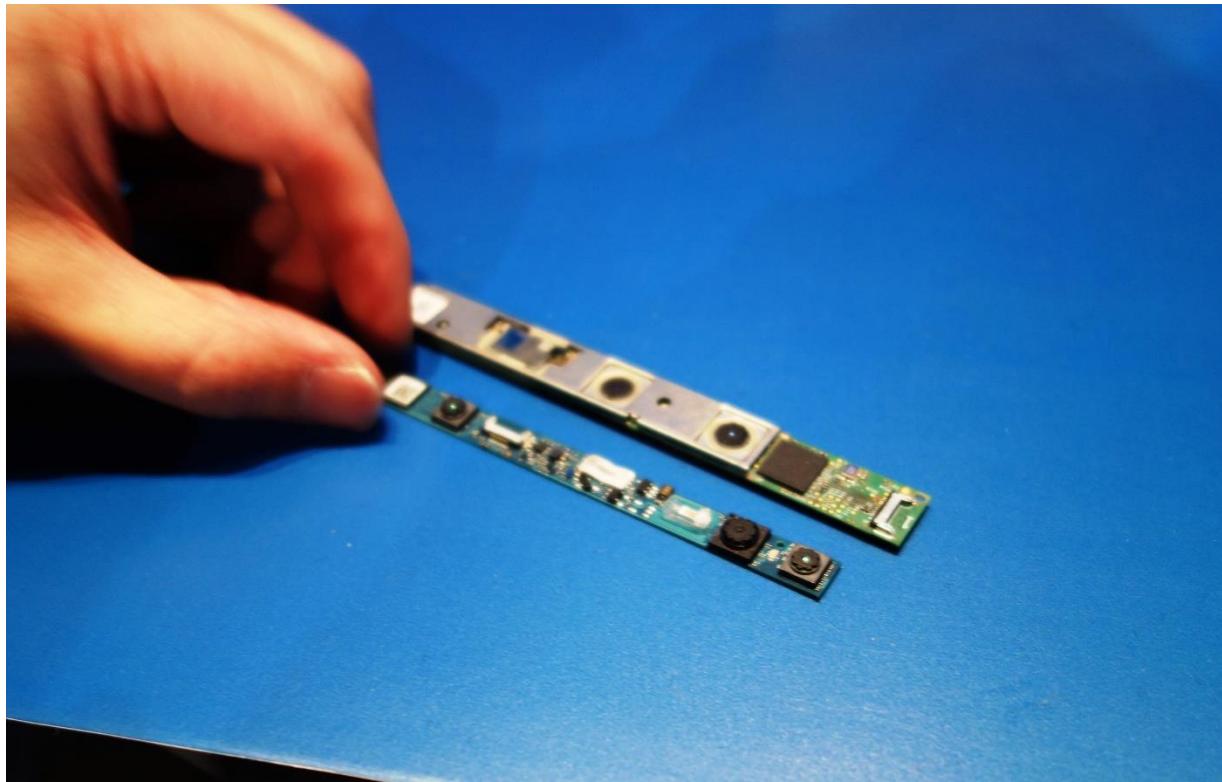
Login without a password...



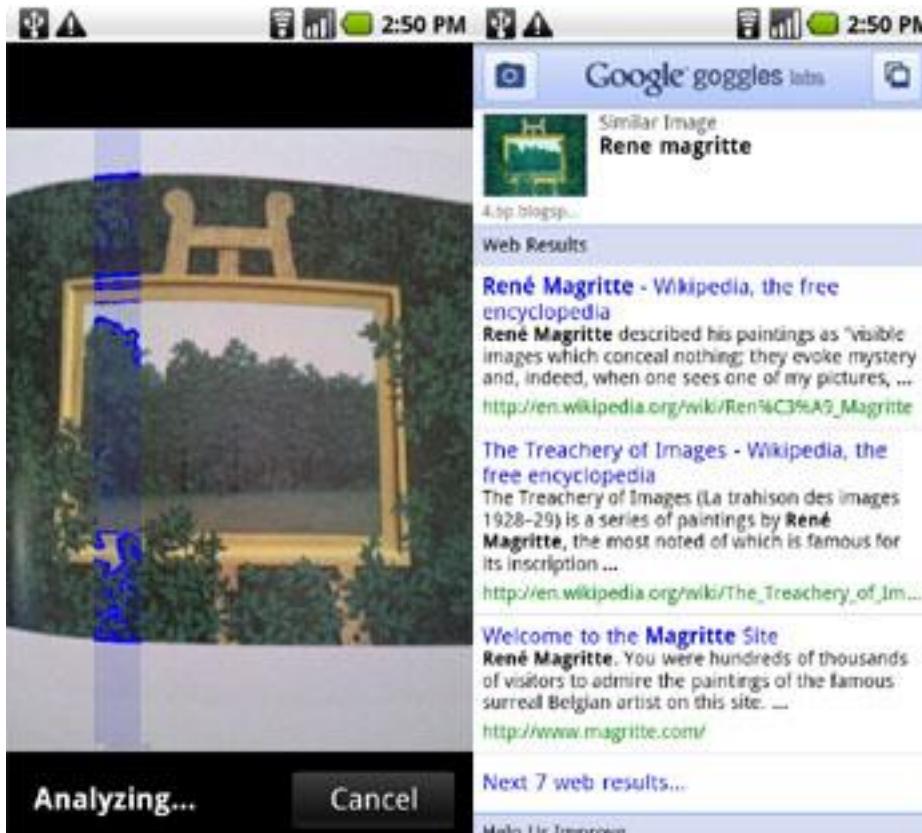
Login without a password...



Login without a password...



Object recognition (in mobile phones)



Google Goggles

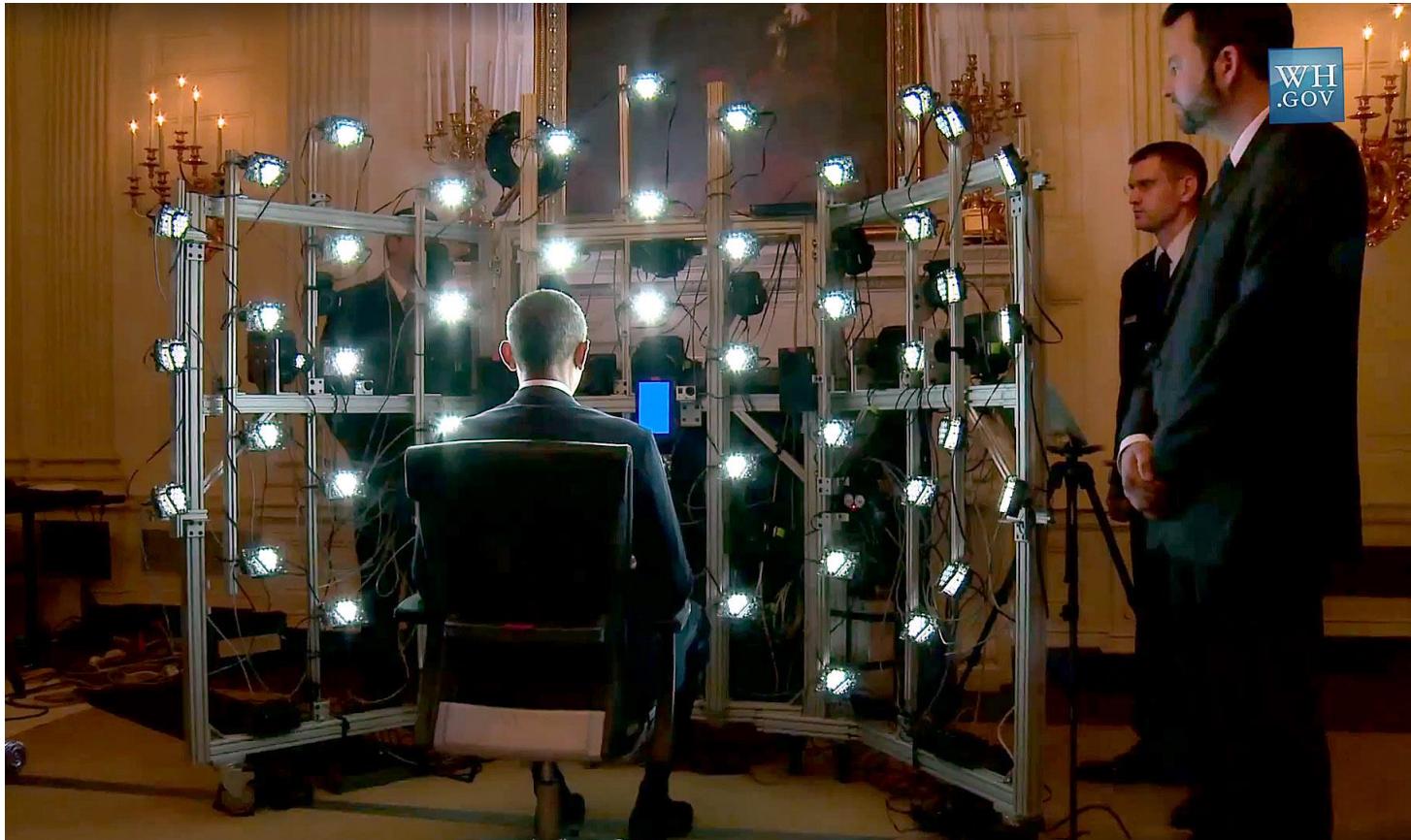
Wikipedia: “In a May 2014 update to Google Mobile for iOS, the Google Goggles feature was removed due to being ‘of no clear use to too many people.’” : (

3D from images



Building Rome in a Day: Agarwal et al. 2009

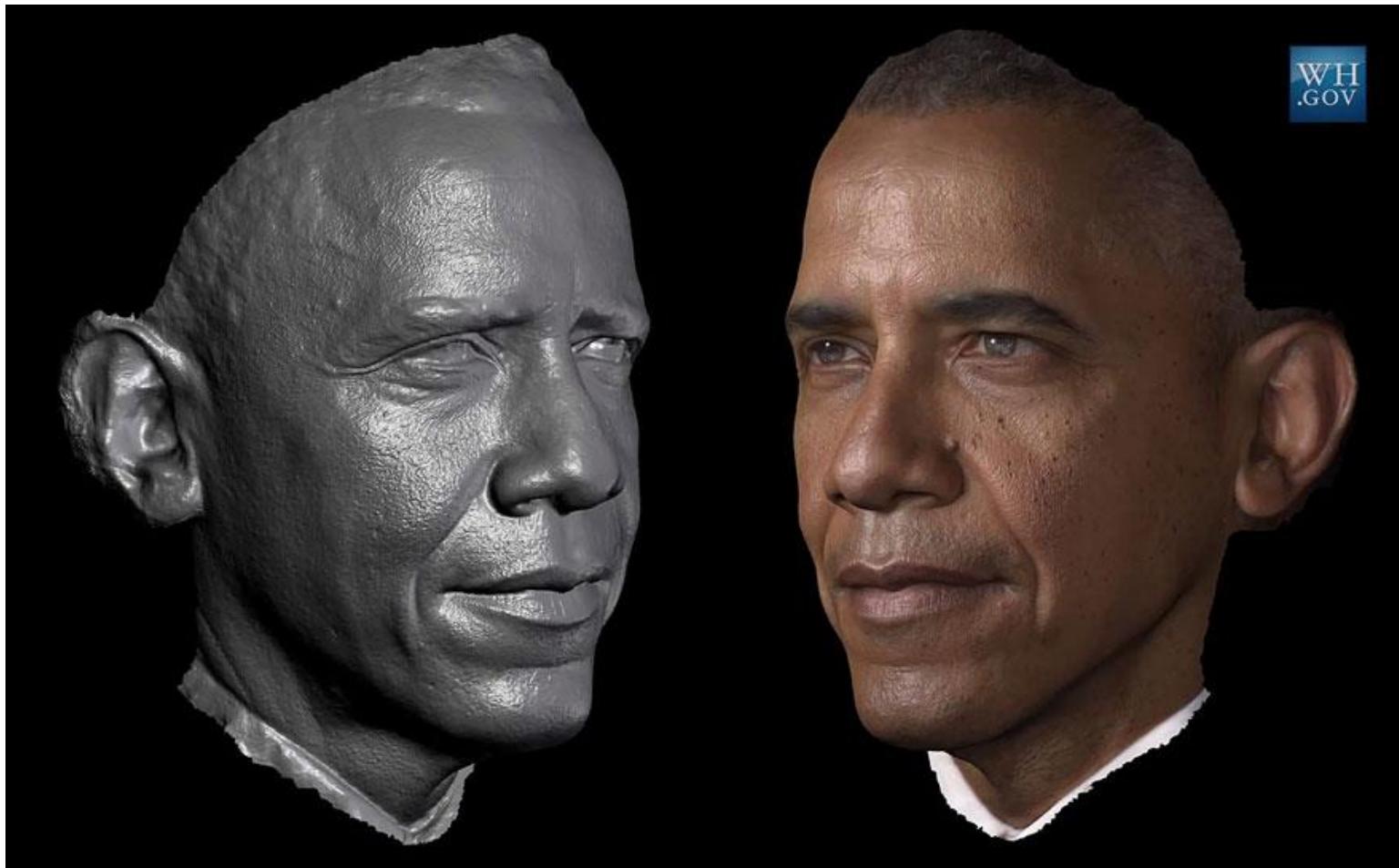
Human shape capture



Human shape capture



Human shape capture



Human shape capture

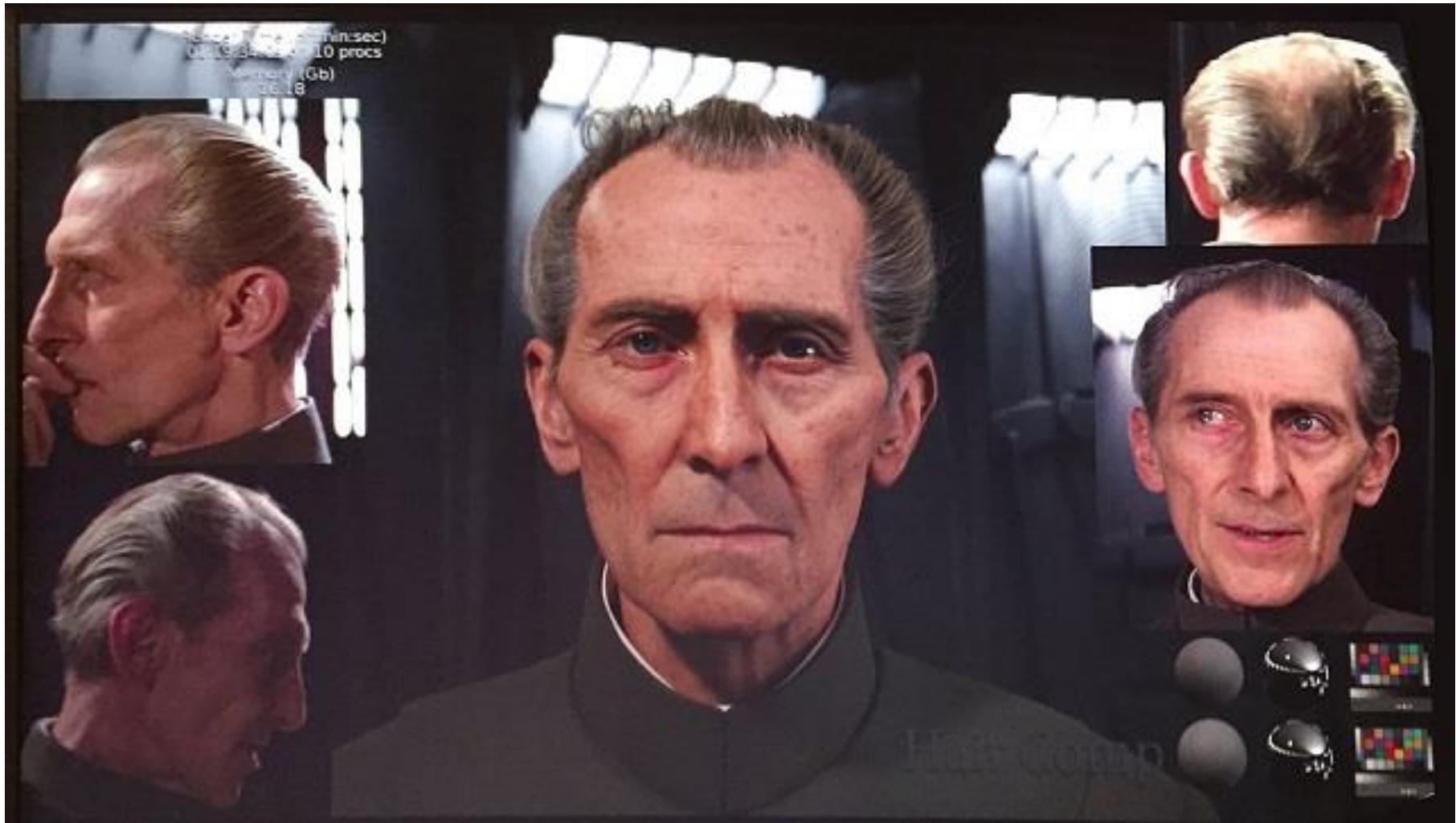


Special effects: shape capture



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=8CTJL5IUjHg>
- 3D: <http://www.youtube.com/watch?v=7QrnwoO1-8A>
- Robot: <http://www.youtube.com/watch?v=w8BmgtMKFbY>



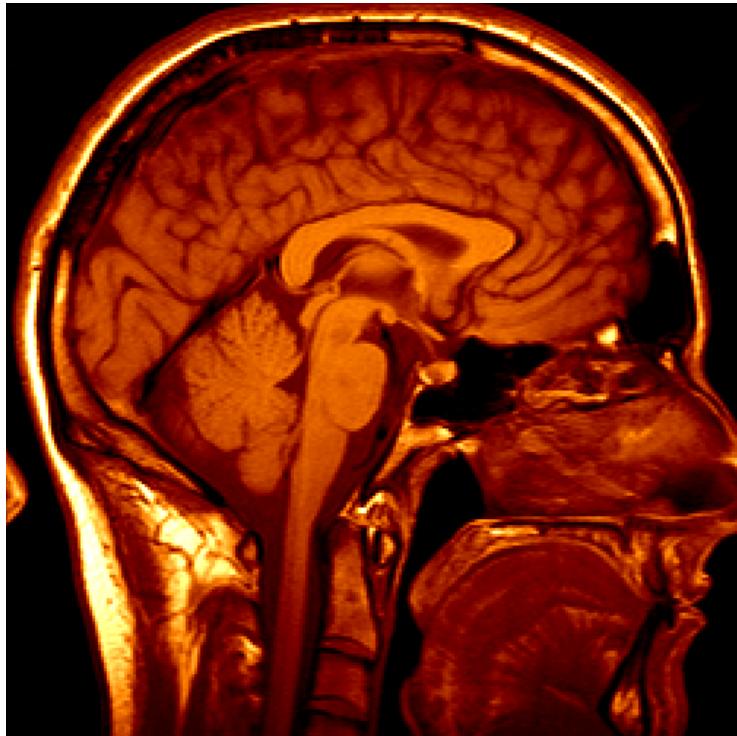
Sports



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

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

Medical imaging



3D imaging
MRI, CT



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

AutoCars - Uber bought CMU's lab





Industrial robots



Vision-guided robots position nut runners on wheels

Vision in space

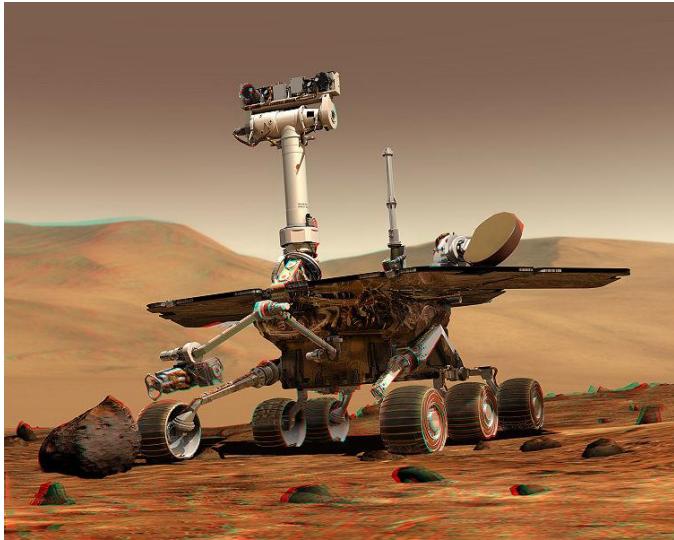


[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

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.

Mobile robots



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



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

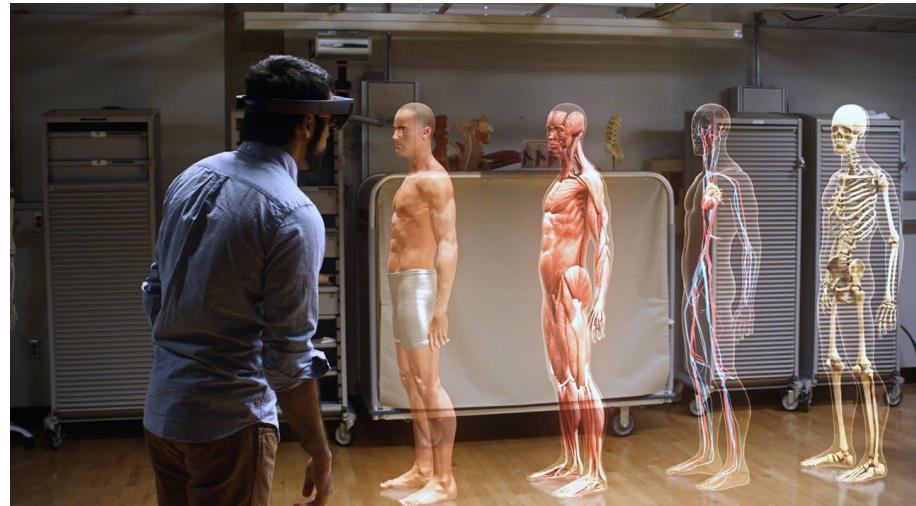


amazon
Prime Air

Saxena et al. 2008
[STAIR](#) at Stanford



Augmented Reality and Virtual Reality



MS HoloLens, Oculus, Magic Leap,

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?

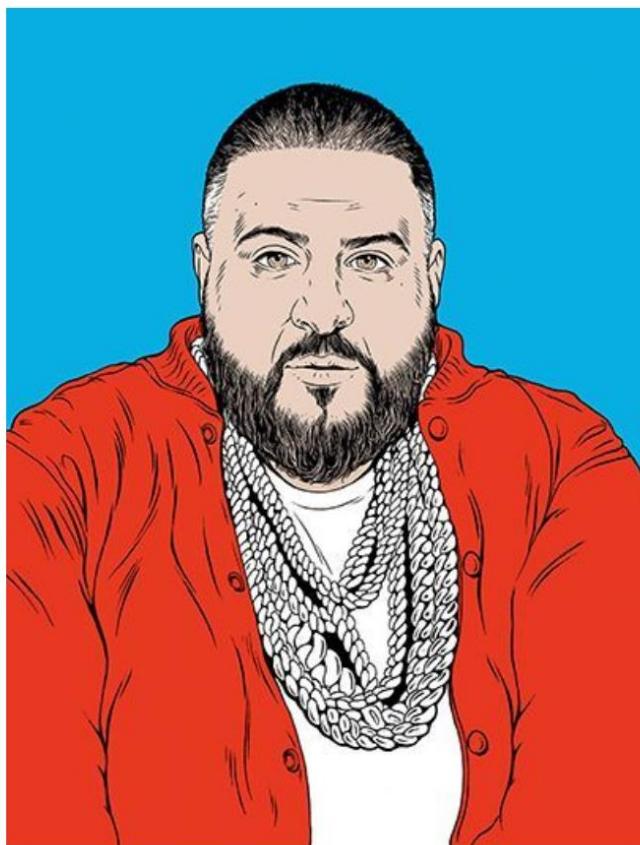
With enough training data, computer vision nearly matches human vision at most recognition tasks

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

WIRED

100

WHO'S SHAPING THE DIGITAL WORLD?



DJ Khaled

Credit [Louise Zergaeng Pomeroy](#)

73. DJ Khaled

Snapchat icon; DJ and producer

Louisiana-born Khaled Mohamed Khaled, aka DJ Khaled, cut his musical chops in the early 00s as a host for Miami urban music radio WEDR. He proceeded to build a solid if not dazzling career as a mixtape DJ and music producer (he founded his label We The Best Music Group in 2008, and was appointed president of Def Jam South in 2009).

69. Geoffrey Hinton

Psychologist, computer scientist; researcher, Google Toronto

British-born Hinton has been dubbed the "godfather of deep learning". The Cambridge-educated cognitive psychologist and computer scientist started being an ardent believer in the potential of neural networks and deep learning in the 80s, when those technologies enjoyed little support in the wider AI community.

But he soldiered on: in 2004, with support from the Canadian Institute for Advanced Research, he launched a University of Toronto programme in neural computation and adaptive perception, where, with a group of researchers, he carried on investigating how to create computers that could behave like brains.

Hinton's work – in particular his algorithms that train multilayered neural networks – caught the attention of tech giants in Silicon Valley, which realised how deep learning could be applied to voice recognition, predictive search and machine vision.

The spike in interest prompted him to launch a free course on neural networks on e-learning platform Coursera in 2012. Today, 68-year-old Hinton is chair of machine learning at the University of Toronto and moonlights at Google, where he has been using deep learning to help build internet tools since 2013.

66. PewDiePie

YouTuber, Brighton

In just six years, 26-year-old Swedish YouTuber Felix Kjellberg has risen from dropout who funded his videos by working at a hot-dog stand to a \$12 million global internet sensation. Kjellberg - aka [PewDiePie](#) - has built a quasi-cult fandom by uploading longish videos (Is This Game Too Sexual?!; Worst Game on the Planet!; What Happened To Resident Evil 7??) where he fools around and cracks jokes while playing video games.

In May 2016, Kjellberg's YouTube channel passed 45 million subscribers and, with 12 billion views, it's now the most-watched channel in the history of the video-sharing website. In the gaming industry, it is thought that he can single-handedly determine the success of a game by simply mentioning it in one of his videos.

Kjellberg - who lives in Brighton with his girlfriend and fellow YouTuber Marzia Bisognin, aka CutiePieMarzia - has tried to use his online clout for good, asking his fanbase (dubbed the "Bro Army") to raise money for charities on several occasions.

Recently, he teamed up with The Walking Dead creator Robert Kirkman to produce his first web series - Scare PewDiePie. The series, which features the YouTuber cavorting in horror-game-like situations, debuted on YouTube Red in February 2016.

63. Yann Lecun

Director of AI research, Facebook, Menlo Park

LeCun is a leading expert in deep learning and heads up what, for Facebook, could be a hugely significant source of revenue: understanding its user's intentions.

63. Yann Lecun

Director of AI research, Facebook, Menlo Park

LeCun is a leading expert in deep learning and heads up what, for Facebook, could be a hugely significant source of revenue: understanding its user's intentions.

61. Taylor Swift

Entertainer, Los Angeles





Credit [Google DeepMind](#)

Google-backed startup DeepMind Technologies has built an **artificial intelligence** agent that can learn to successfully play 49 classic Atari games by itself, with minimal input.



8. Demis Hassabis

Co-founder and CEO, DeepMind, London

In March 2016, DeepMind's AI [AlphaGo](#) beat the Go world champion [Lee Se-dol](#). The Google-owned startup is moving machine learning forward at a pace that could affect every industry, from healthcare to commerce.

Computer Vision and Nearby Fields

Derogatory summary of computer vision:
“Machine learning applied to visual data.”

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

Scope of CSCI 1430

Image Processing
Geometric Reasoning
Recognition
Deep Learning

Machine Learning

Graphics

Computational
Photography

Optics

Robotics

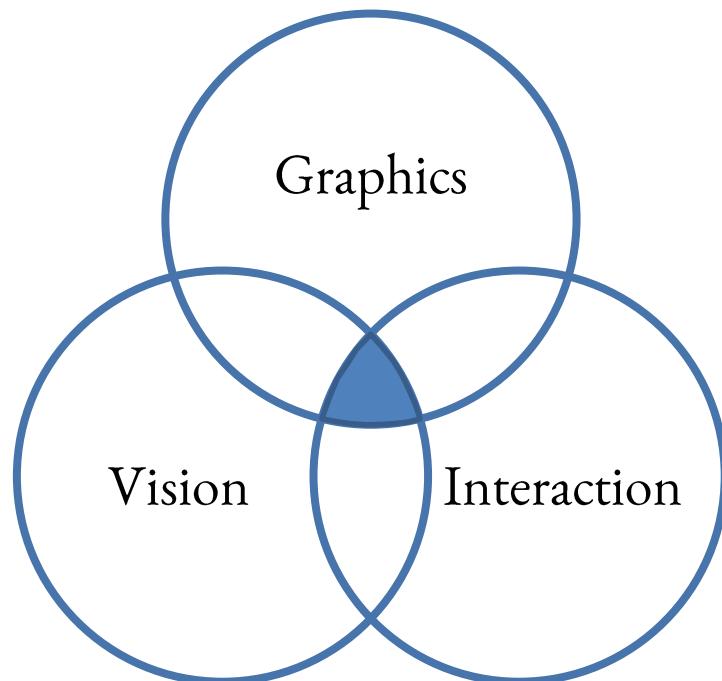
Human Computer
Interaction

Medical Imaging

Neuroscience

COURSE STAFF

I work in here.



Start



End









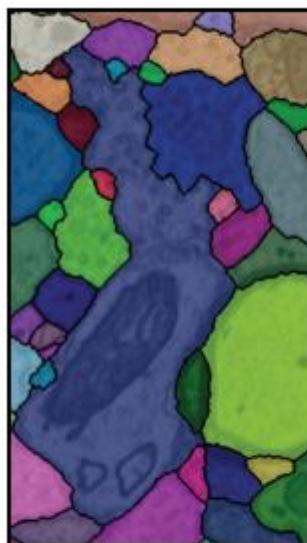
(a) Supervoxel Segmentation



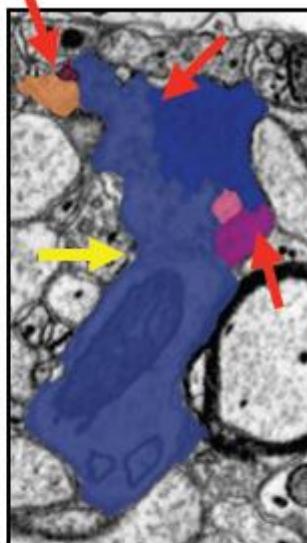
(b) User Interactions



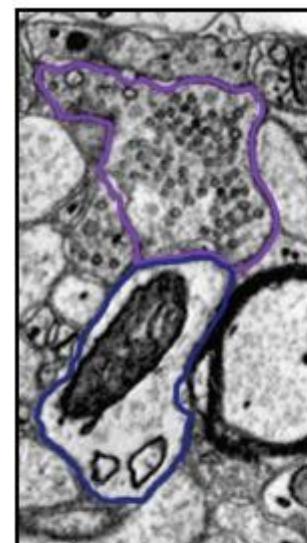
(c) Multi-label Segmentation



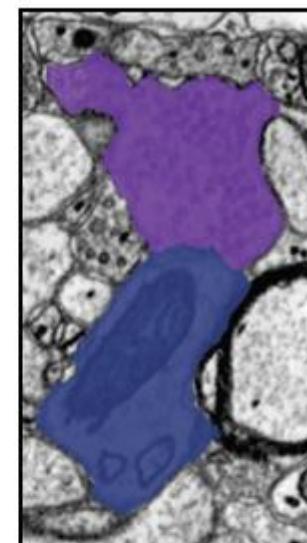
Initial
Segmentation



Merge- and Split
Errors



Correct
Borders



Fixed
Segmentation

Instructor: James Tompkin



HARVARD

John A. Paulson
School of Engineering
and Applied Sciences



mpipi

Max Planck Institute
Germany



University College London
UK



Massachusetts
Institute of
Technology



Render pixels?
Capture pixels?
Interact with pixels?
I am probably interested.

Watch my research overview video!

james_tompkin@brown.edu

Office: CIT 547

TA introduction

Eric Xiao

Daniel Nurieli

Eleanor Tursman

Martin Zhu

Jackson (Jack) Gibbons



CS 143 – James Hays

- Continuing his course – many materials, the courseworks, from him + previous staff – serious thanks!
- If you see 'JH' on a slide, then it is one of his.

Prerequisites

- **Linear algebra**, basic calculus, and probability
- Programming, data structures

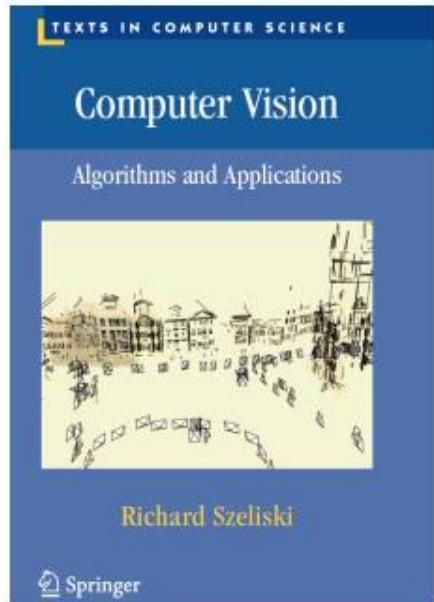
Contact

- Course runs *quiet hours* – 9pm to 9am.
 - No contact! We will ignore you.
- Piazza first
 - TAs are not on there all the time; set hours.
- cs1430tas@lists.brown.edu second
- TA office hours: Moon Lab 227, Tues-Thurs 7-9pm
- James: Tues 1-2pm

Textbook

Computer Vision: Algorithms and Applications

© 2010 [Richard Szeliski](#), Microsoft Research



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

Textbook

Deep Learning

An MIT Press book

Ian Goodfellow and Yoshua Bengio and Aaron Courville

- 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

- 100% programming projects (6 total)

Project 0 (do this immediately)

- Install MATLAB
- Go Help, go through ‘Getting Started with MATLAB’ tutorial. Familiarize.
- Go through second tutorial (image operands)
- Come to us in office hours (next week) if you have trouble.

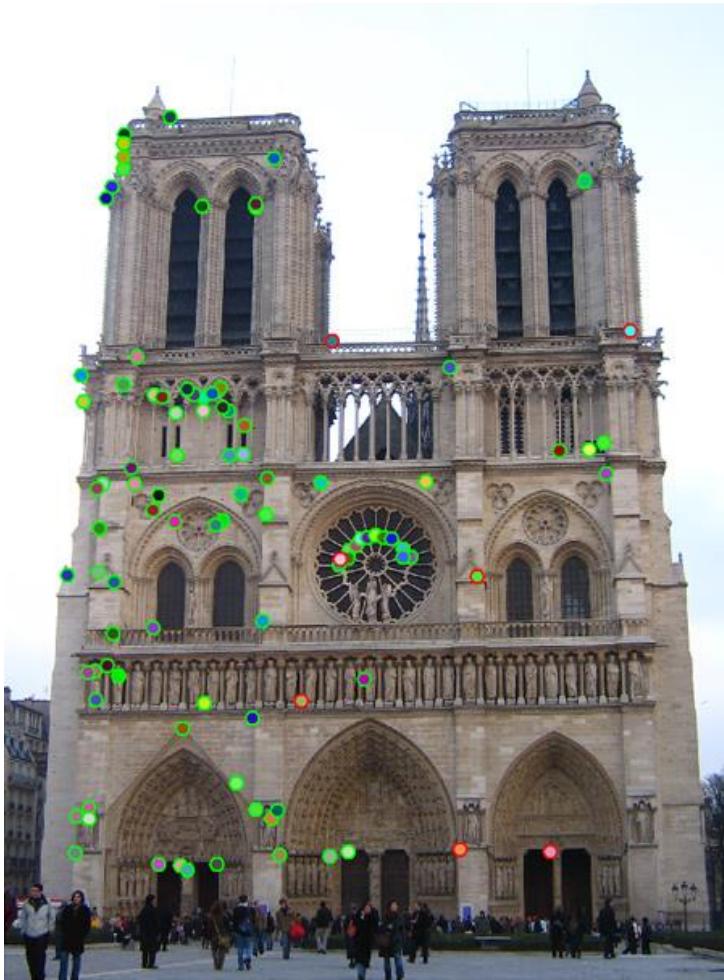
Proj 1: Image Filtering and Hybrid Images

- Implement image filtering to separate high and low frequencies.
- Combine high frequencies and low frequencies from different images to create a scale-dependent image.



Proj 2: Local Feature Matching

- Implement interest point detector, SIFT-like local feature descriptor, and simple matching algorithm.



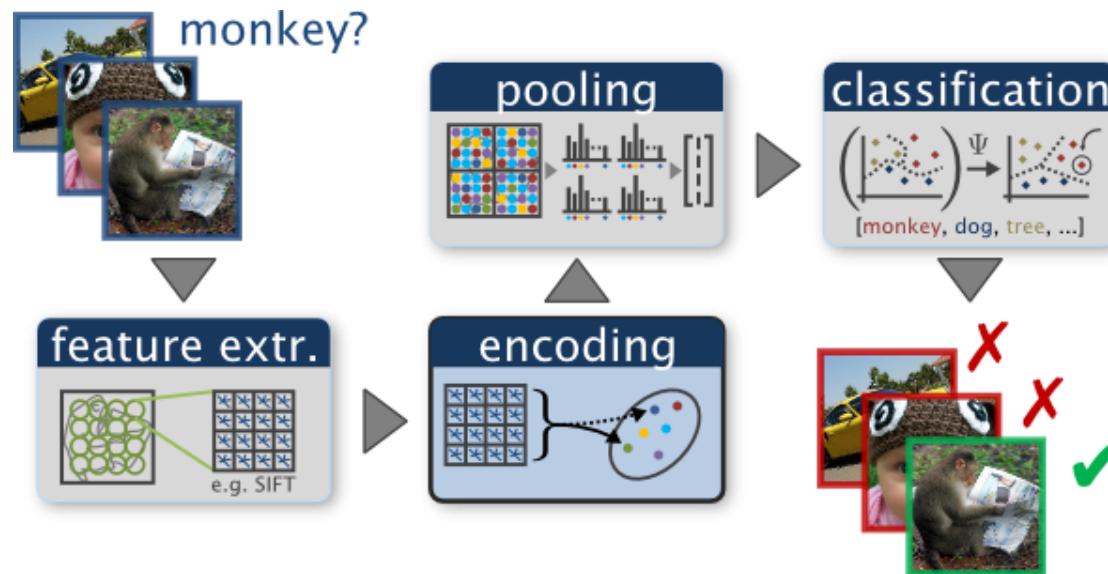
Proj 3: Multi-view Geometry

- Recover camera calibration from feature point matches.
- Foundation for almost all measurement in computer vision.



Proj 4: Scene Recognition with Bag of Words

- Quantize local features into a “vocabulary”, describe images as histograms of “visual words”, train classifiers to recognize scenes based on these histograms.



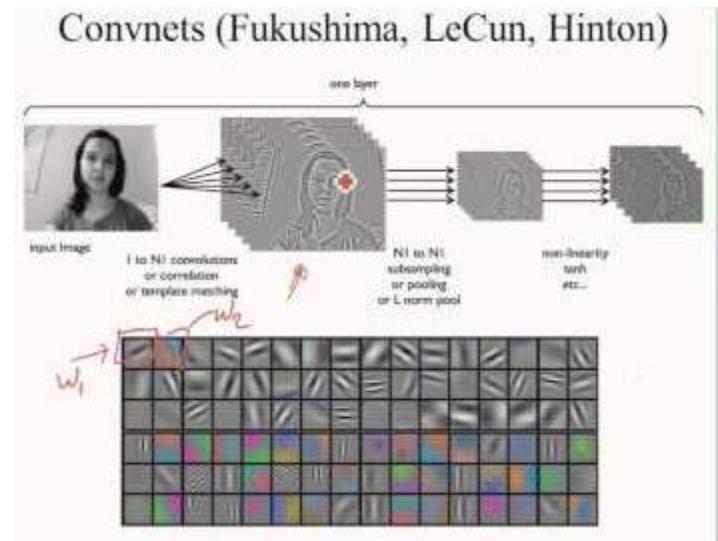
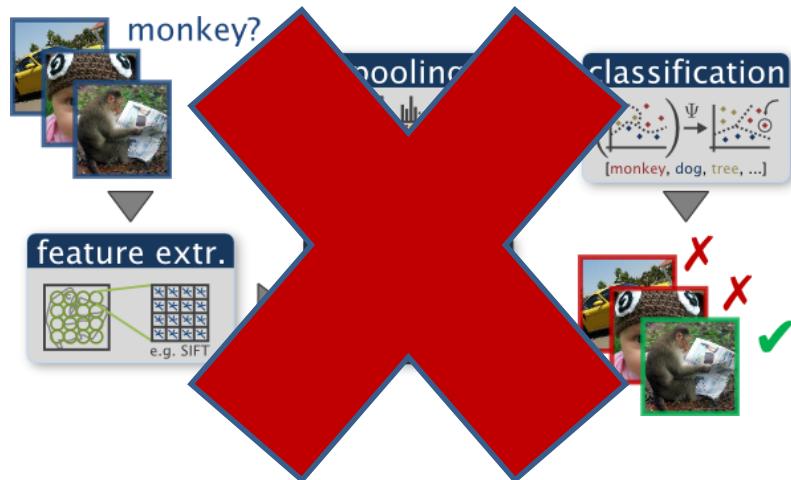
Proj 5: Object Detection with a Sliding Window

- Train a face detector based on positive examples and “mined” hard negatives, detect faces at multiple scales and suppress duplicate detections.



Proj 6: Convolutional Neural Net

- Proj 4 again, but state of the art.



Any questions?!

- Friday: Light and Color