19CSE435: Computer Vision

Lecture 01 - 19/01/2022
Introduction

Adopted from Computer Vision Textbook and course materials R_Szeliski

19CSE435 **COMPUTER VISION** **L-T-P-C: 3-0-0-3**

Course Objectives

- This course introduces the geometry of image formation and its use for 3D reconstruction and calibration.
- It introduces the analysis of patterns in visual images that are used to reconstruct and understand the objects and scenes.

Course Outcomes

CO1: *Understand image formation and camera calibration*.

CO2: *Analyse and select image features and apply for image matching*.

CO3: *Understand recognition algorithms through case studies.*

CO4: *Understand the basics of stereo vision.*

Syllabus

Unit 1

Introduction, Image Formation – geometric primitives and transformations, photometric image formation, digital camera, Camera calibration.

Unit 2

Feature Detection and Matching – points and patches, edges, lines, Feature-Based Alignment - 2D, 3D feature based alignment, pose estimation, Image Stitching, Dense motion estimation – Optical flow - layered motion, parametric motion, Structure from Motion.

Unit 3

Recognition – object detection, face recognition, instance recognition, category recognition, Stereo Correspondence – Epipolar geometry, 3D reconstruction.

Text Book(s)

Szeliski R. Computer Vision: Algorithms and Applications Springer. New York. 2010.

Reference(s)

Shapiro LG, Stockman GC. Computer Vision: Theory and Applications. 2001. Forsyth DA, Ponce J. Computer Vision: a modern approach;2012.

Davies ER. Machine vision: theory, algorithms, practicalities. Elsevier; 2004 Dec 22. Jain R, Kasturi R, Schunck BG. Machine vision. New York: McGraw-Hill; 1995 Mar 1

Present Evaluation

Proposed Evaluation: Subject to approval

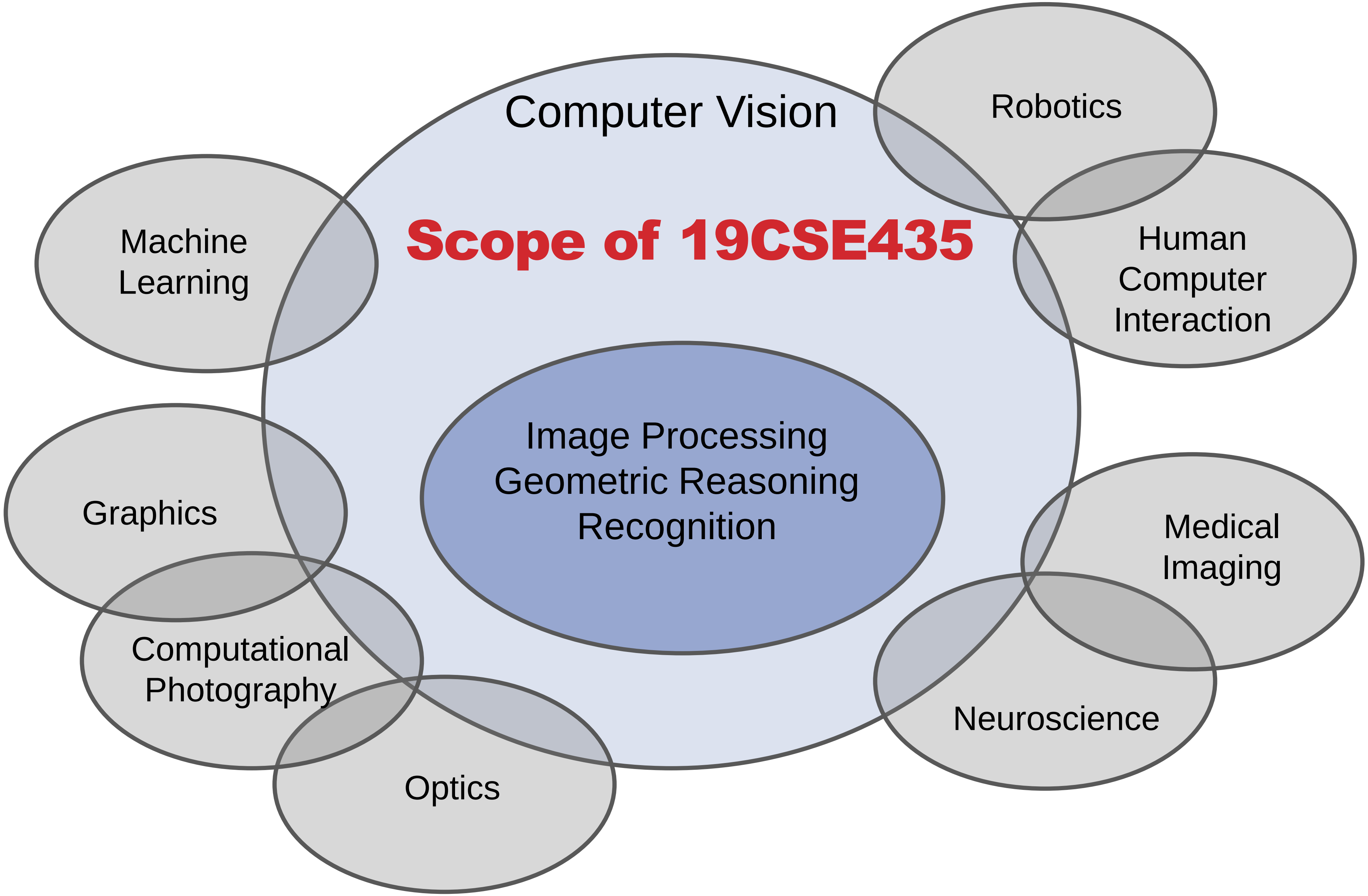
Periodical 1	15		Continuous Assessment (Theory)(CAT)	15	
Periodical 2	15		Continuous Assessment (Lab) (CAL)	20	
*Continuous Assessment (Theory) (CAT)	20		Project	30	
End Semester		50	End Semester		35

What is Computer Vision?

Computer Graphics: Models to Images

Comp. Photography (Image Processing): Images to Images

Computer Vision: Images to Models (Interpretation)


Computer Vision
Scope of 19CSE435
Robotics
Machine Learning
Human Computer Interaction
Image Processing
Geometric Reasoning
Recognition
Graphics
Medical Imaging
Computational Photography
Neuroscience
Optics

Computer Vision

Make computers understand images and video or any visual data.

What kind of scene?

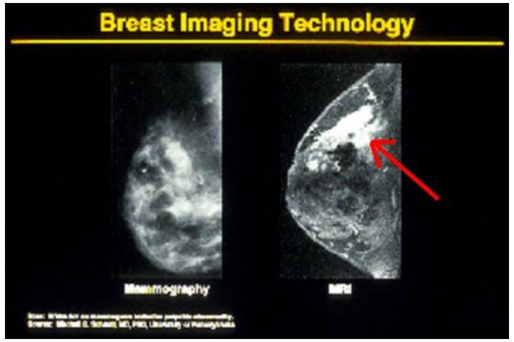
Where are the cars?

How far is the building?

…

Why computer vision matters

Safety

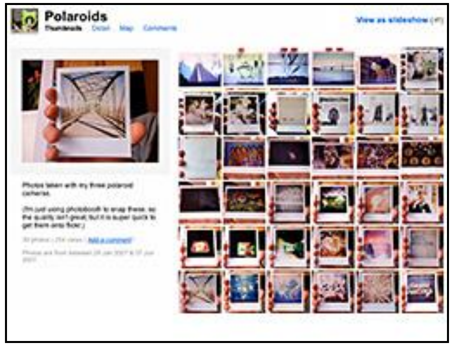


Health

Security

Comfort

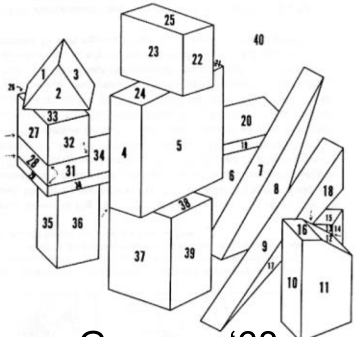
Fun



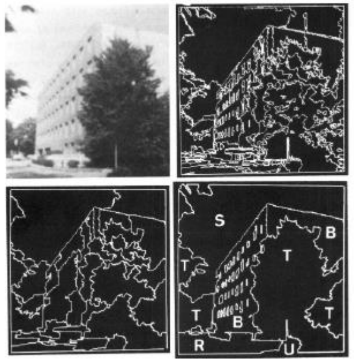
Access

Ridiculously brief history of computer vision

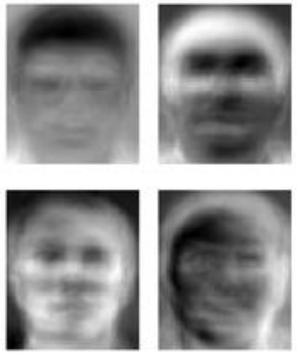
- **1966: Minsky assigns computer vision as an undergrad summer project**
- **1960's: interpretation of synthetic worlds**
- **1970's: some progress on interpreting selected images**
- **1980's: ANNs come and go; shift toward geometry and increased mathematical rigor**
- **1990's: face recognition; statistical analysis in vogue**
- **2000's: broader recognition; large annotated datasets available; video processing starts**
- **2010's: Deep learning with ConvNets**
- **2020's: Widespread autonomous vehicles?**
- **2030's: robot uprising?**



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

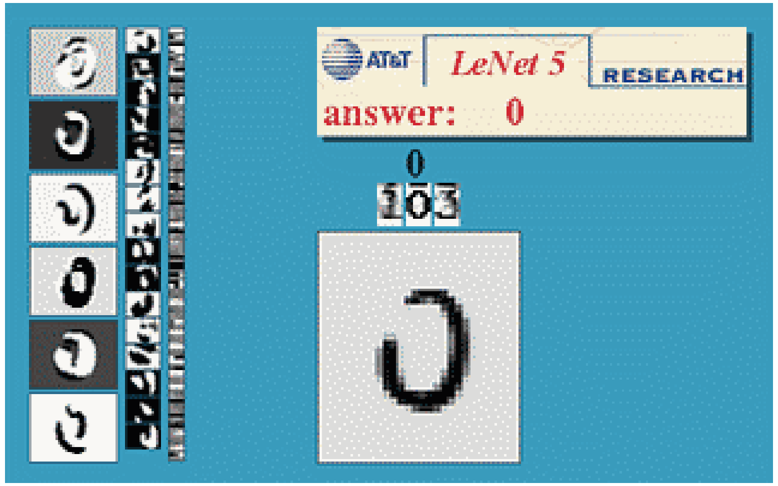
How vision is used now

- **Examples of real world applications**

Optical character recognition (OCR)

Technology to convert scanned docs to text

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



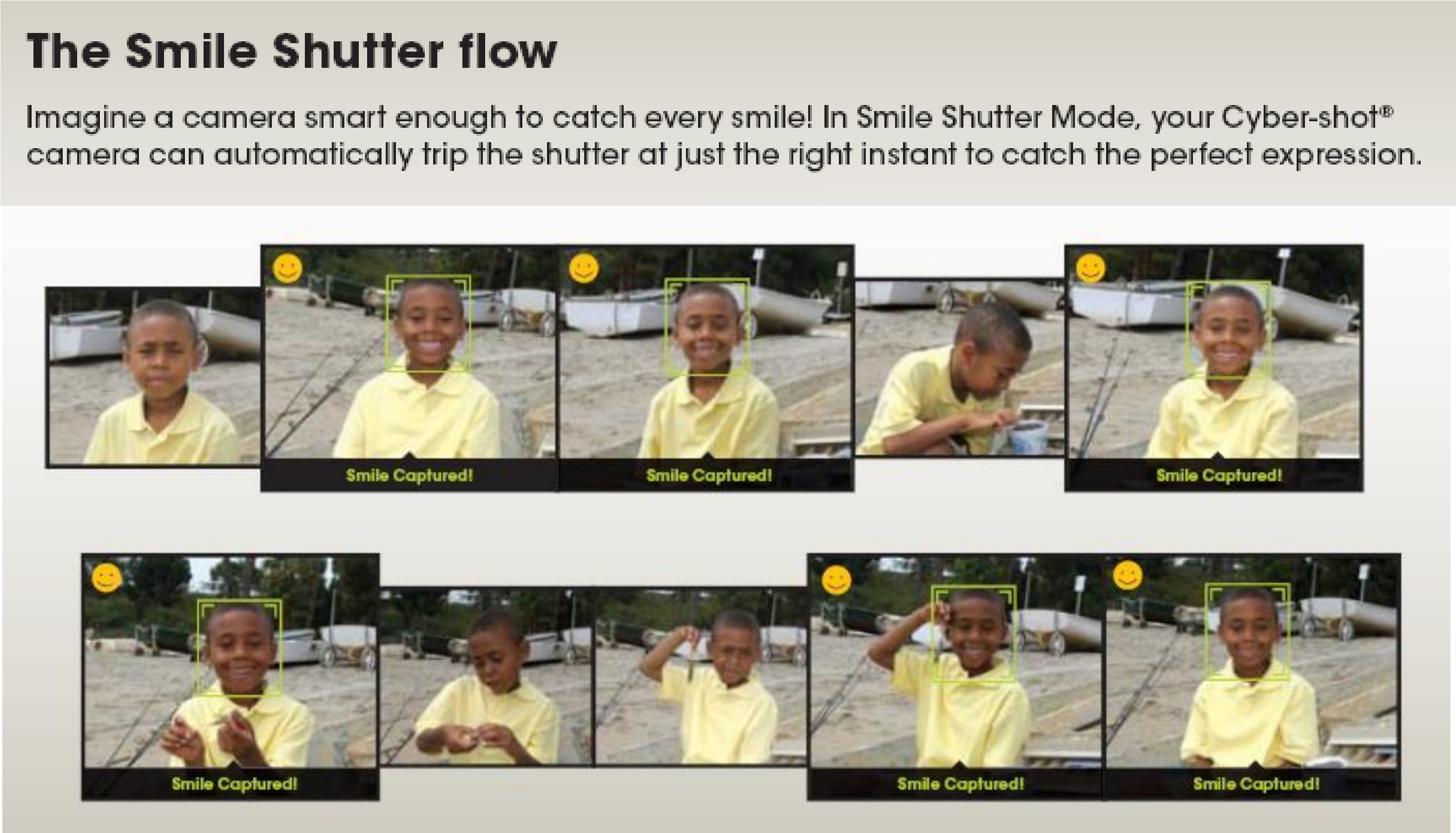
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Face detection

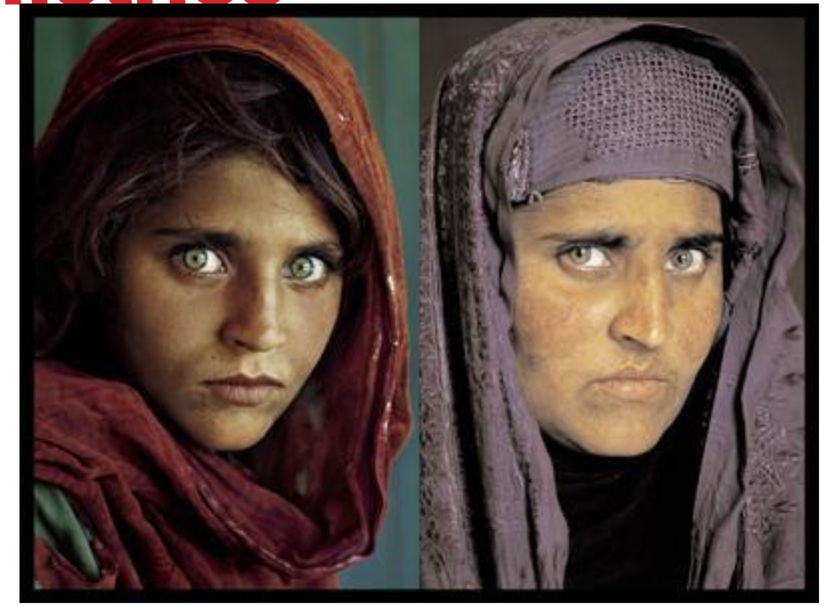
- **Digital cameras detect faces**

Smile detection

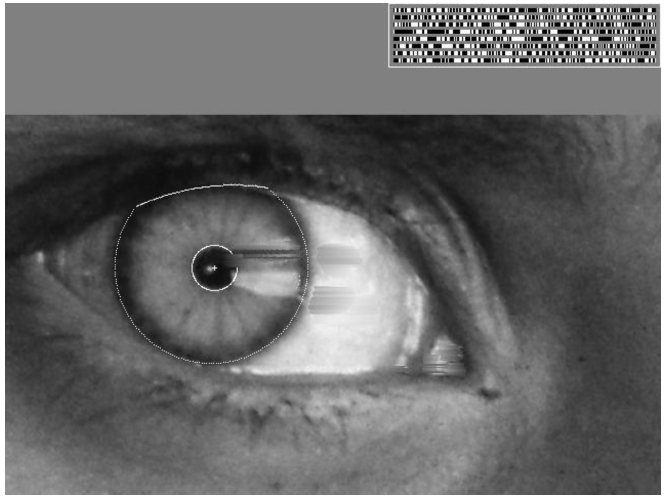


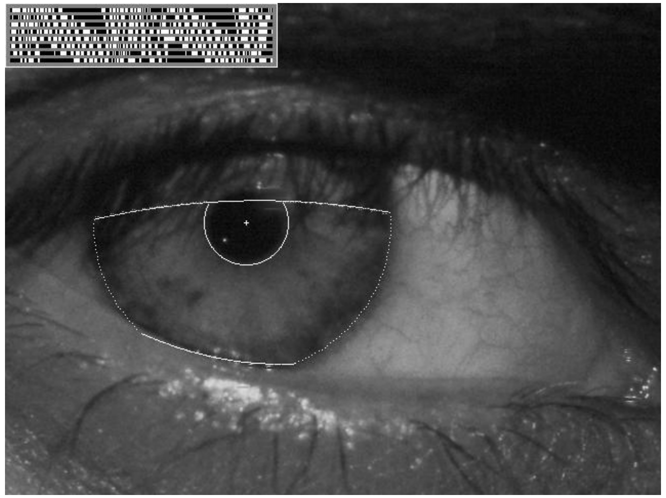
[Sony Cyber-shot® T70 Digital Still Camera](#)

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

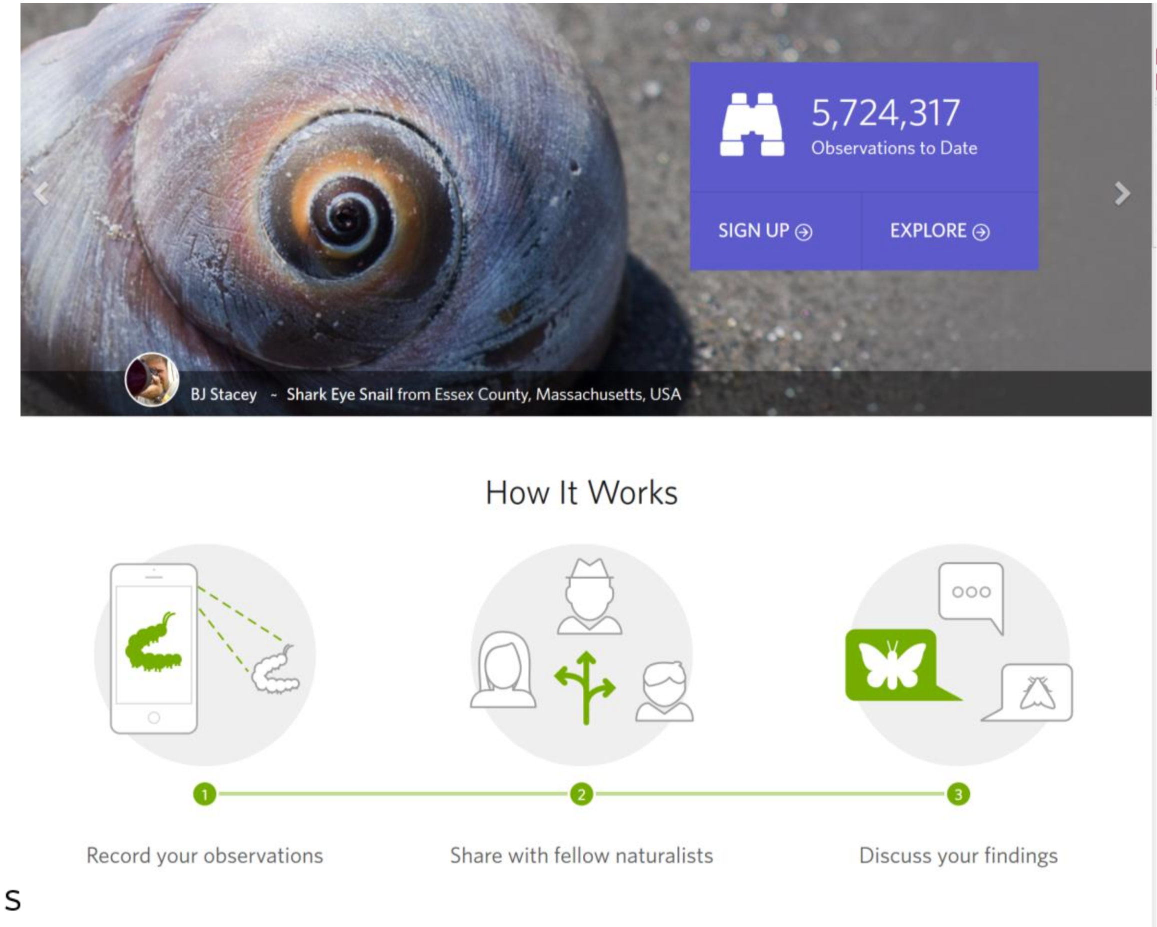
Face recognition systems now
beginning to appear more widely
http://www.sensiblevision.com/

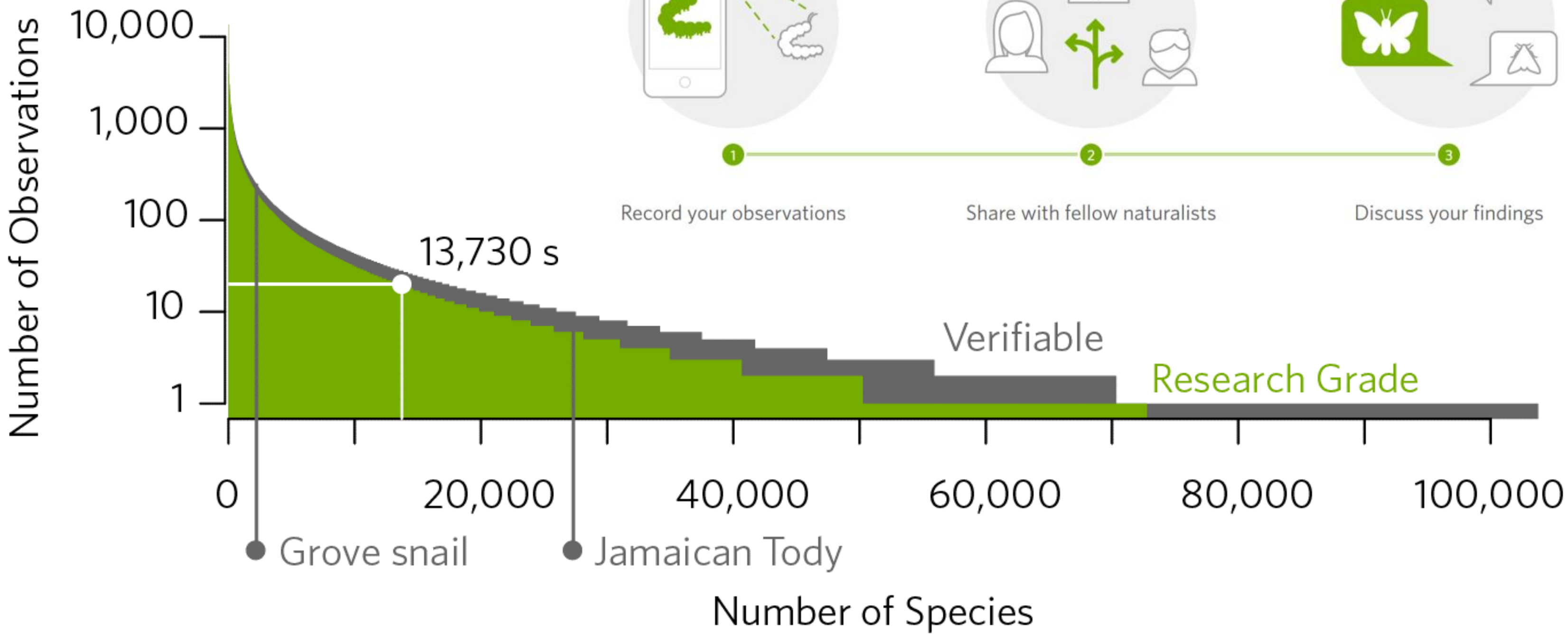
Object recognition (in mobile phones)

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

[Google Goggles]()

iNaturalist





Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

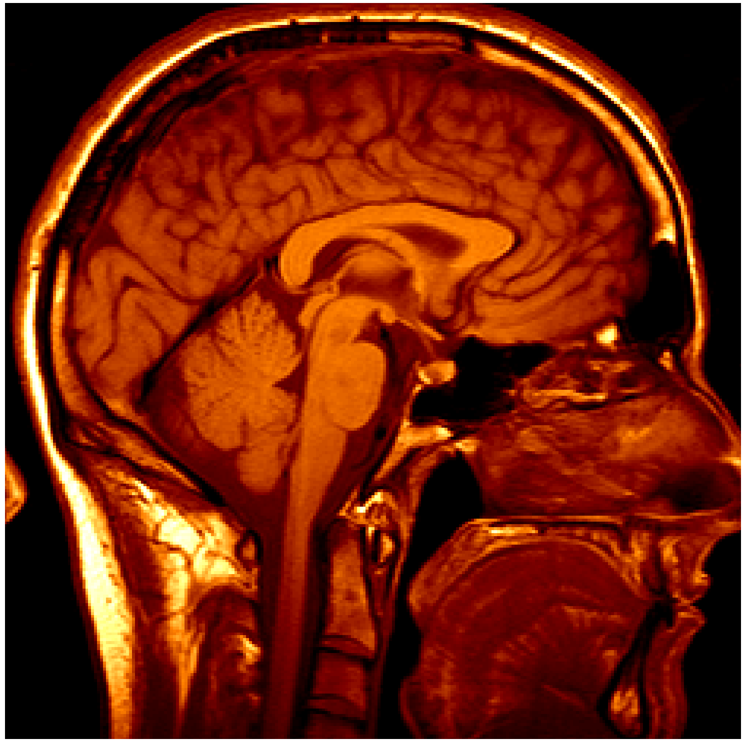
Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Sports

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

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Smart cars

Slide content courtesy of Amnon Shashua

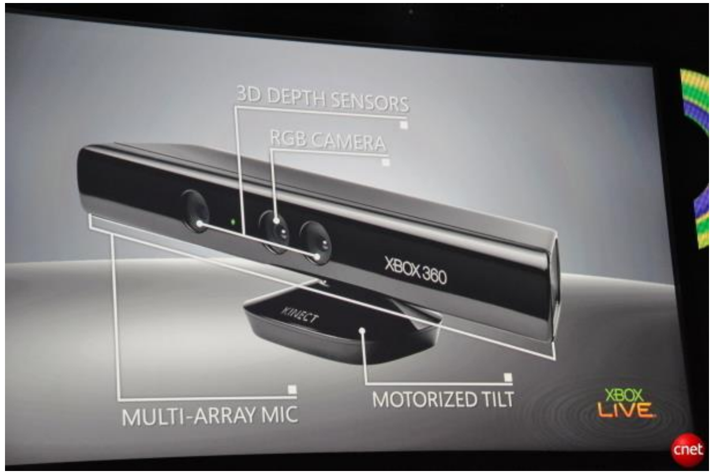
- Mobileye
  - ~~Market Capitalization: 11 Billion dollars~~
  - Bought by Intel for 15 Billion dollars

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



Augmented Reality and Virtual Reality

Magic Leap, Oculus, Hololens, etc.

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.

Amazon Prime Air

https://www.amazon.com/b?node=8037720011

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. More and more techniques are being “deepified”.

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.

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.

62. Richard Branson

Founder, Virgin Group, London

Branson saw his personal fortune grow £550 million when Alaska Air bought Virgin America for $2.6 billion in April. He is pressing on with civilian space travel with Virgin Galactic.

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.

The story of AlphaGo so far

AlphaGo is the first computer program to defeat a professional human Go player, the first program to defeat a Go world champion, and arguably the strongest Go player in history.

AlphaGo's first formal match was against the reigning 3-times European Champion, Mr Fan Hui, in October 2015. Its 5-0 win was the first ever against a Go professional, and the results were published in full technical detail in the international journal, **Nature**. AlphaGo then went on to compete against legendary player Mr Lee Sedol, winner of 18 world titles and widely considered to be the greatest player of the past decade.

AlphaGo's 4-1 victory in Seoul, South Korea, in March 2016 was watched by over 200 million people worldwide. It was a landmark achievement that experts agreed was a decade ahead of its time, and earned AlphaGo a 9 dan professional ranking (the highest certification) - the first time a computer Go player had ever received the accolade.

During the games, AlphaGo played a handful of **highly inventive winning moves,** several of which - including move 37 in game two - were so surprising they overturned hundreds of years of received wisdom, and have since been examined extensively by players of all levels. In the course of winning, AlphaGo somehow taught the world completely new knowledge about perhaps the most studied and contemplated game in history.

Since then, AlphaGo has continued to surprise and amaze. In January 2017, an improved AlphaGo version was revealed as the online player "Master" which achieved **60 straight wins in online fast time-control games** against top international Go players.

In May 2017, Alpha Go took part in The Future of Go Summit in the birthplace of Go, China, to delve deeper into the mysteries of Go in a spirit of mutual collaboration with the country's top players. You can read more about the five day summit **here**.

› **More on The Future of Go Summit in this video**

› **Watch the video here**

BigGAN

2018 Arxiv, ICLR 2019

Large Scale GAN Training for High Fidelity Natural Image Synthesis

Andrew Brock, Jeff Donahue, Karen Simonyan

This person does not exist

https://www.thispersondoesnotexist.com/