COMP 9517
Computer Vision

Introduction

What is Computer Vision?

Every picture tells a story

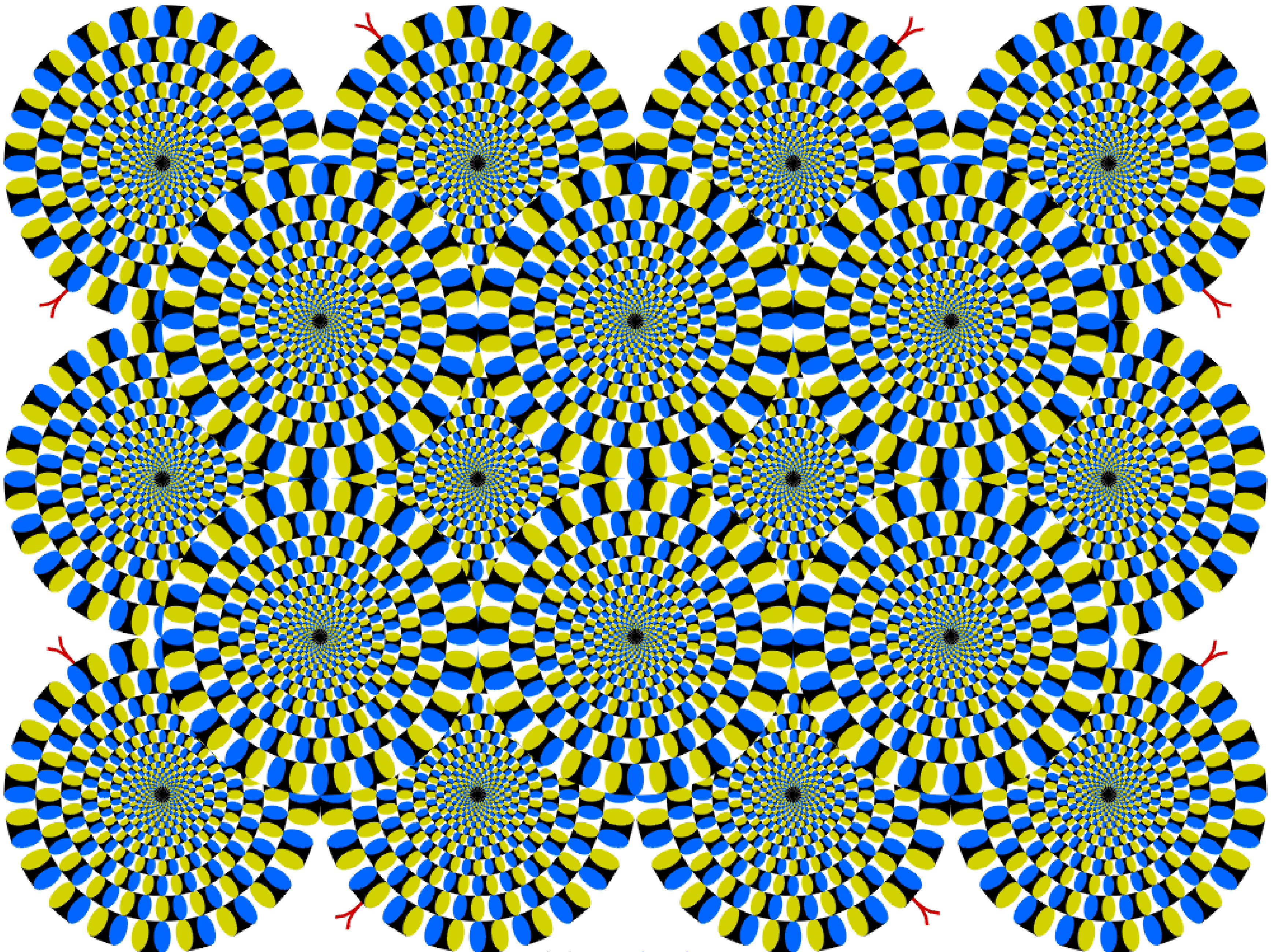
- Goal of computer vision is to write computer programs that can interpret images

Can computers match (or beat) human vision?

- Yes and no (but mostly no!)
  - humans are much better at "hard" things
  - computers can be better at "easy" things

Human perception has its shortcomings…

[Sinha and Poggio, *Nature*, 1996]



Current State of the Art

- Here are some examples

3D Reconstruction

Varcity Project (https://varcity.ethz.ch/) recreates 3D city models using social media photos

Smart Image / Video Surveillance

LIGHTHOUSE (https://www.light.house/) – “It's an interactive assistant that tells you everything you want to know about your life at home.”

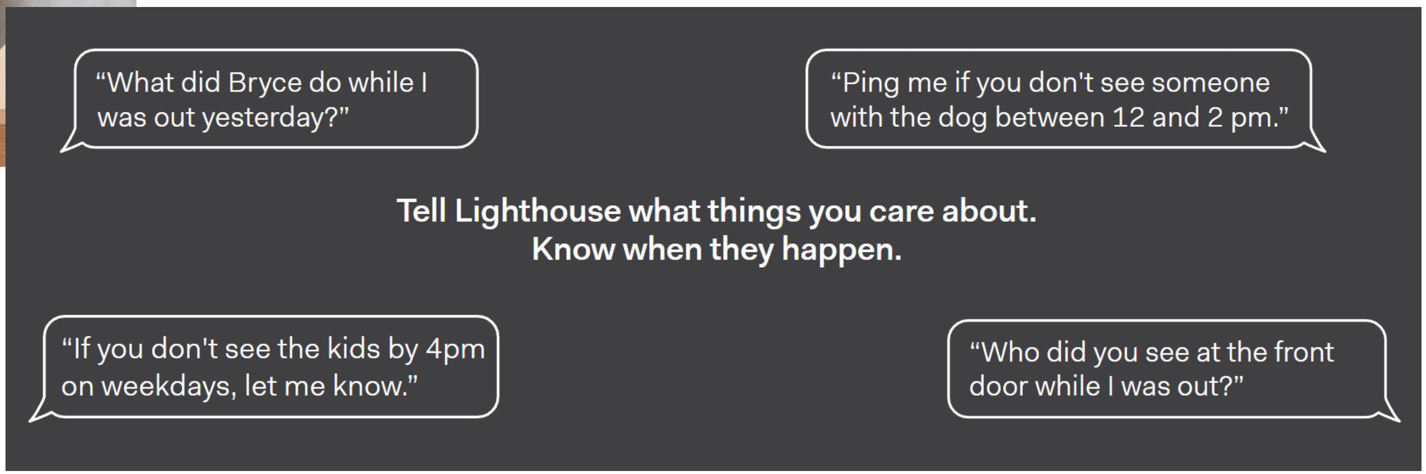


Image Classification and Captioning

Google's Show and Tell open sources image captioning model in TensorFlow Deep Learning Platform (https://github.com/tensorflow/models/tree/master/im2txt)

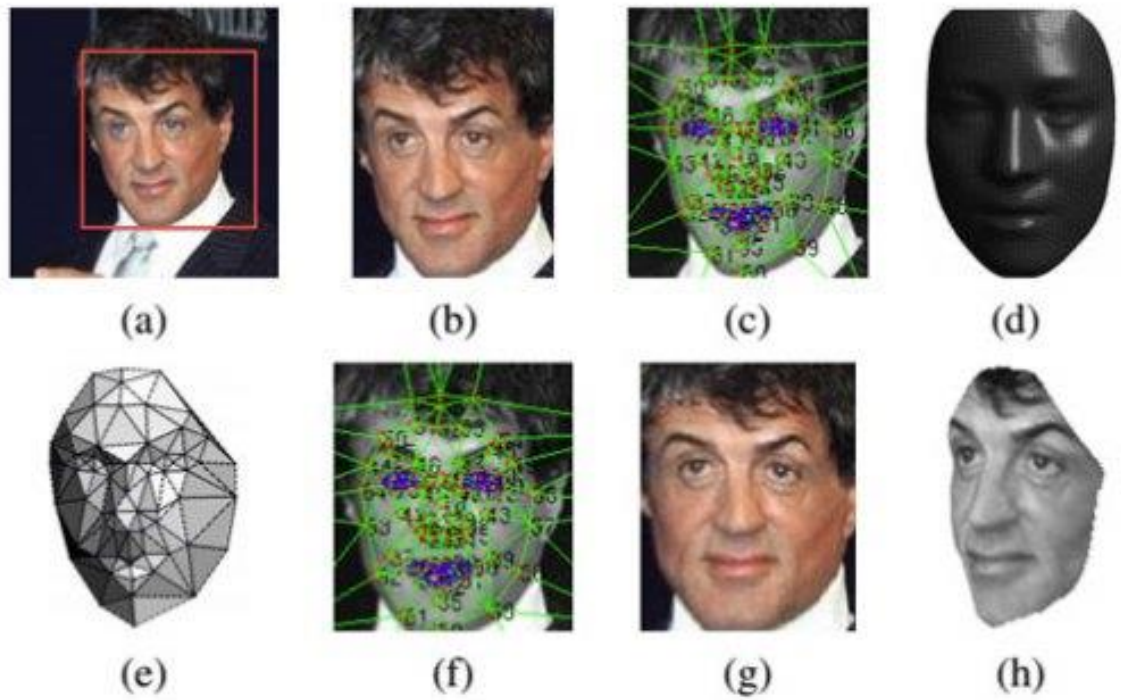
Photo credits - Google

Event Detection and Self Controlled Motion

Iris Automation (http://www.irisonboard.com/) - Safer drone operation with intelligent collision avoidance

Face Detection

Facebook's DeepFace Project Nears Human Accuracy In Identifying Faces (https://research.fb.com/publications/deepface-closing-the-gap-to-human-level-performance-in-face-verification/)



'DeepFace,' an algorithm capable of identifying a face in a crowd with 97.25 percent accuracy

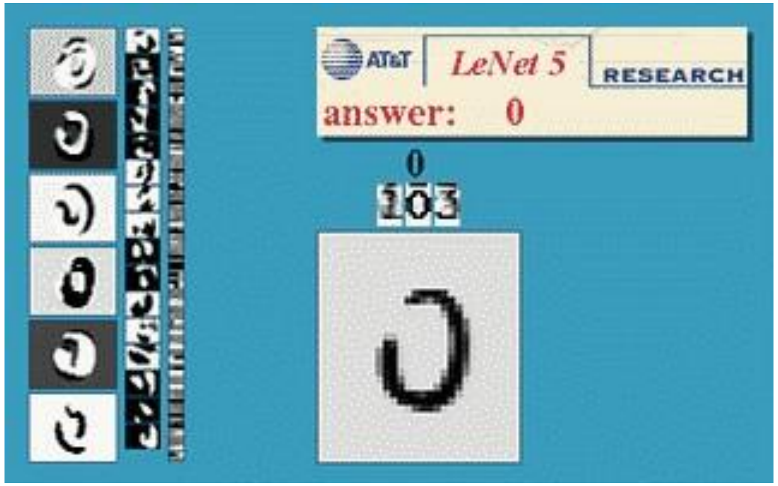
Login without a password…

Fingerprint scanners on
many new laptops,
other devices

Windows Hello makes login is
as easy as looking at your PC

Optical character recognition (OCR)

- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software



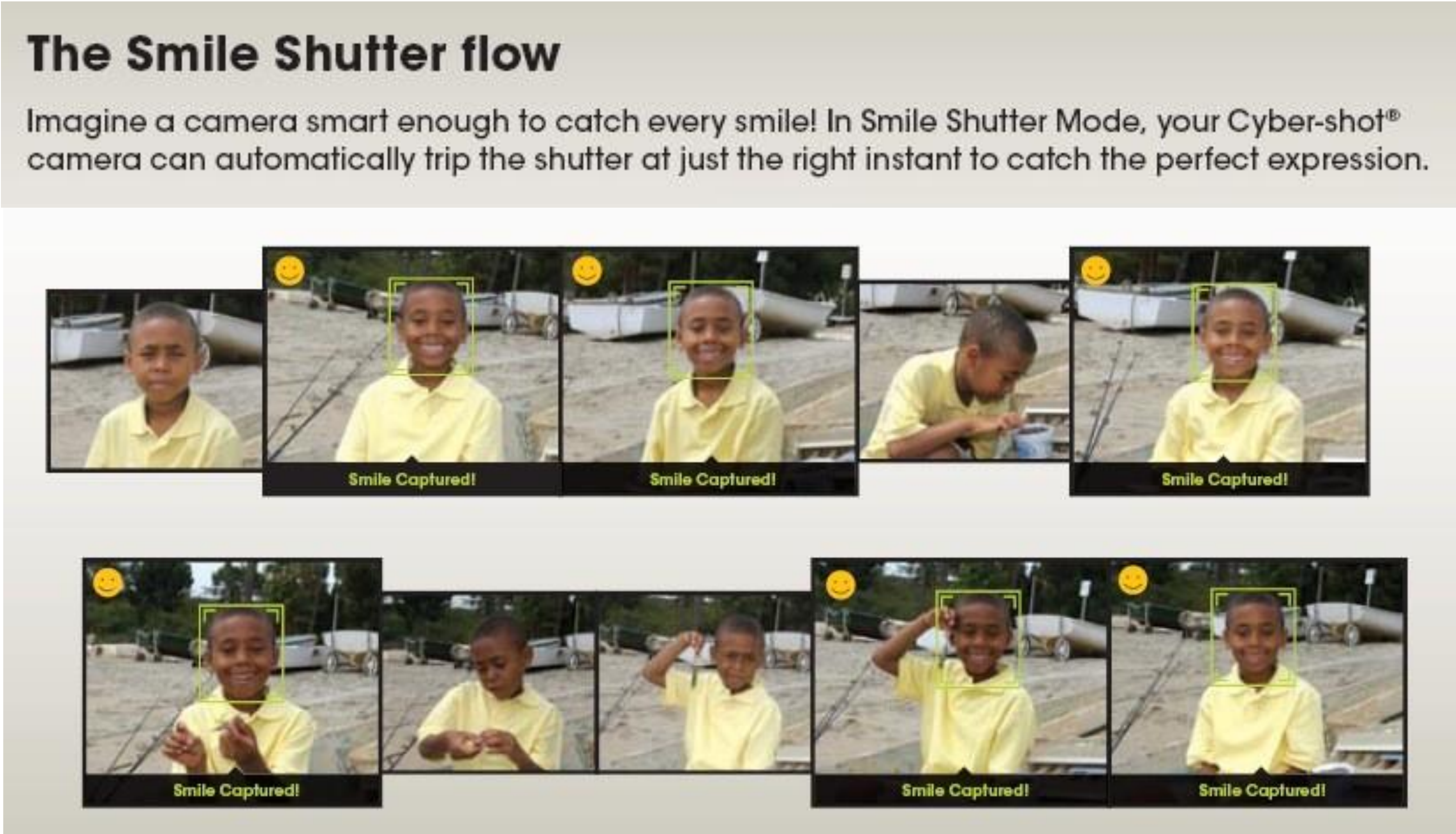
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Face detection

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

Smile detection



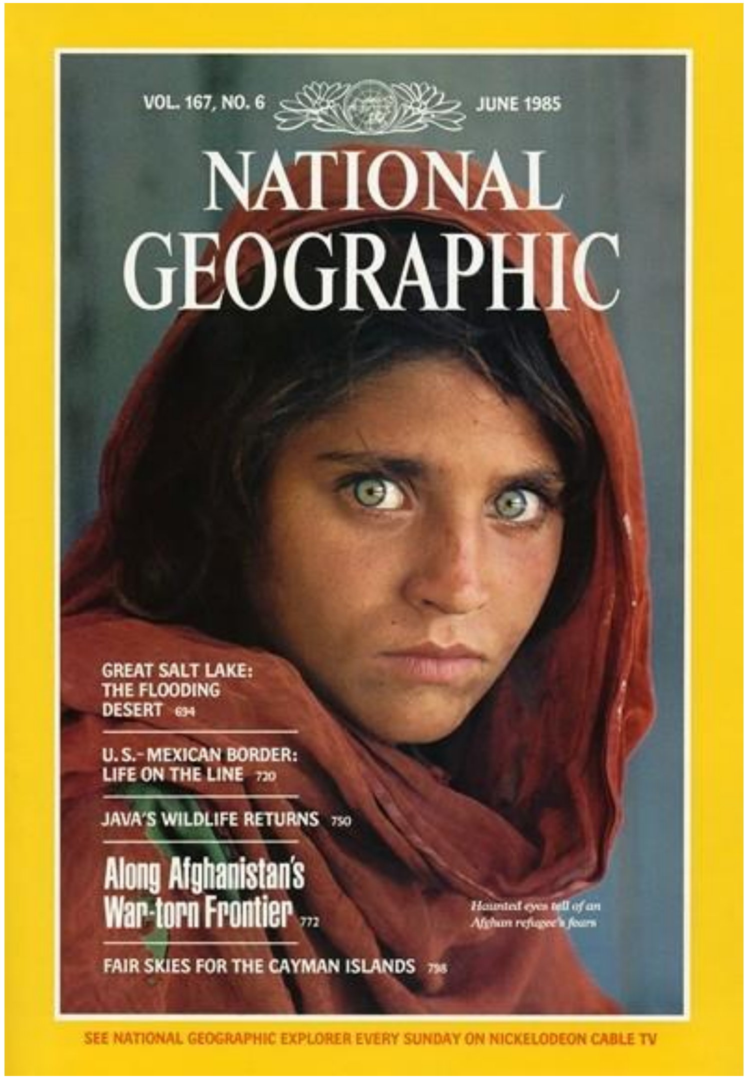
[Sony Cyber-shot® T70 Digital Still Camera]

Object recognition (in supermarket)

[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… "

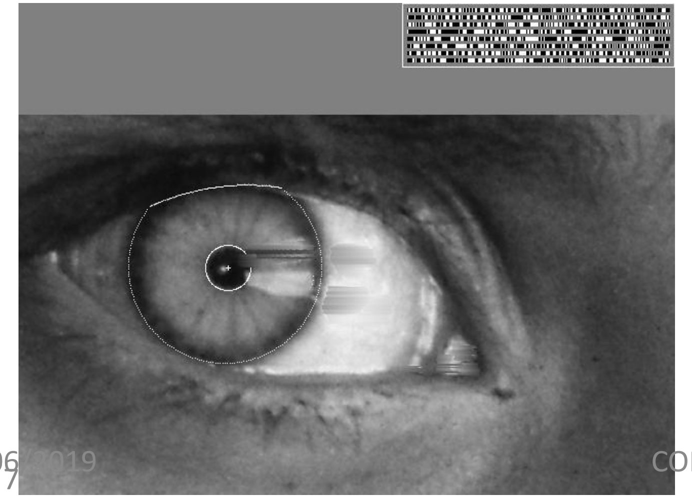
Face recognition

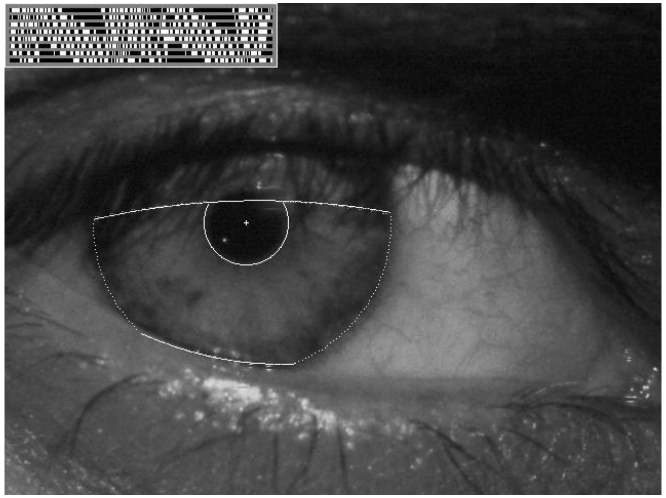


Who is she?

Vision--based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story at http://www.cl.cam.ac.uk/~jgd1000/afghan.html





Object recognition (in mobile phones)

- This is becoming real:
  - Microsoft Research

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

Smart cars

- [Mobileye https://www.mobileye.com/](https://www.mobileye.com/)
  - Vision systems currently in **313 car models from 27 OEM partners by 2017**

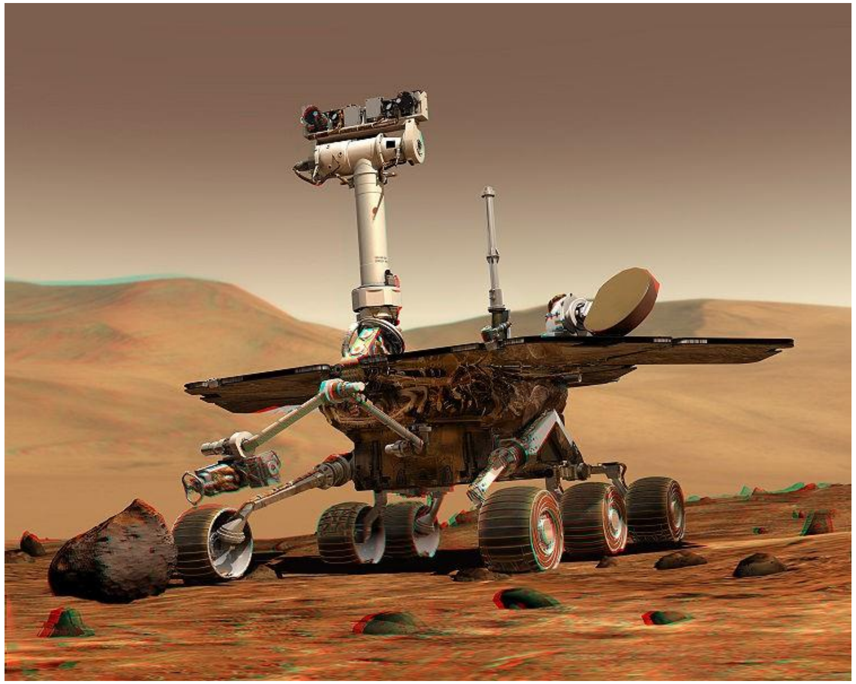
Slide content courtesy of Amnon Shashua

Vision in space

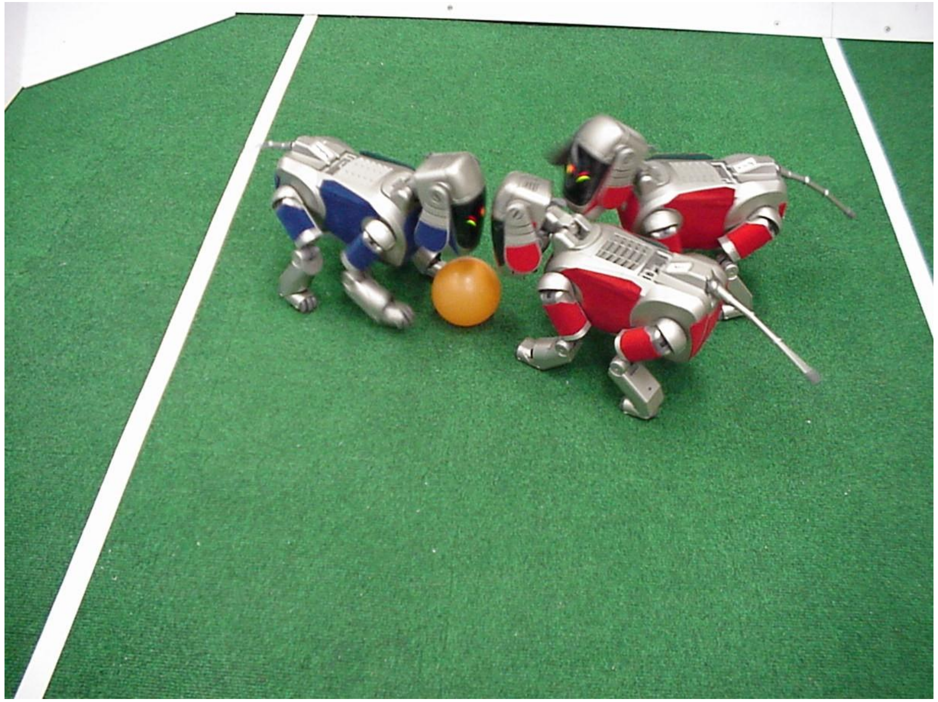
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

- 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.

Robotics

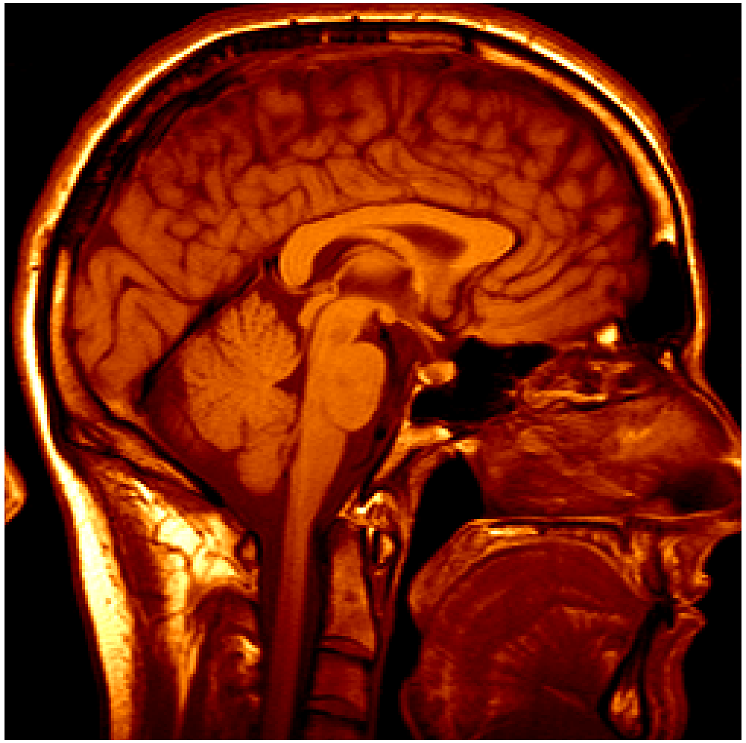


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Medical imaging

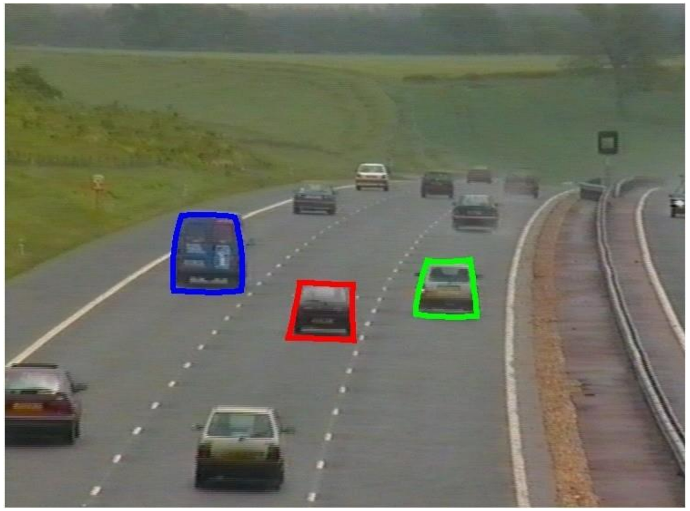


3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Video Surveillance

- Traffic Monitoring
- Object tracking
- Action recognition, driving, stopping, etc
- Vehicle speed
- Counting
- Challenge: occlusion, illumination changes and non--linear speed



Goals of Computer Vision

- Extract useful information from images: both metric and semantic
- Complexity of visual data is a challenge
- Recent progress due to higher processing power, memory, storage capacity
- Image->measurements->model->algorithms for learning and inference

Computer Vision Topics

- Requires a solid understanding of camera and of the physical process of imaging to:
  - obtain simple inferences from individual pixel values
  - enforce some order on groups of pixels to separate them from each other or infer shape information
  - recognise objects using geometric information or probabilistic techniques
  - combine the information available in multiple images into a coherent whole

Critical Issues

- **Sensing**: how do sensors obtain images of the world?
- **Encoded Information**: how do images yield information of the scene, such as colour, texture, shape, motion, etc.?
- **Representations**: what representations are appropriate to describe objects?
- **Algorithms**: what algorithms process image information and construct scene descriptions?

Low Level Vision

- Almost entirely **digital image processing**
  - **Sensing**: image capture and digitisation
  - **Pre-processing**: improve image quality: suppress noise, enhance object features, edge extraction
  - **Image segmentation**: separate objects from background, partition image into objects of interest
  - **Description**: compute features which differentiate objects-- also called *feature extraction*
  - **Classification***:* assign labels to image segments (regions)

High Level Vision

- About knowledge construction, representation and inference
  - **recognition**: identification of objects
  - **interpretation**: assign meaning to groups of recognized objects
  - **scene analysis**

Assumed Knowledge

Before commencing this course, students should:

- know how to program well in Python, or be willing to learn it independently
- be familiar with data structures and algorithms, and basic statistics
- be able to learn to use and integrate software packages, including OpenCV, Scikit-learn, Keras
- be familiar with vector calculus and linear algebra, or be willing to learn them independently

Self-assess BEFORE deciding to stay/enroll in the course.

Course Changes in 2019

- Changes in assessment, to address level of difficulty and volume of work
  - Lab questions will be released during the week prior to the lab
  - Group project will contain a larger individual component
  - Replace the class test with final exam
- Tutor training for better experience and communication
- Co-lecturers will be provided more information on course expectations and outcomes

Weekly Class Structure

Week	Time	Topic
Week 1	Monday (4-6pm), Tuesday (12-2pm)	Introduction, image formation, image processing
Week 2	Tuesday (12-2pm)	Image processing
Week 3	Monday (4-6pm), Tuesday (12-1pm)	Feature representation
Week 4	Monday (4-6pm), Tuesday (12-1pm)	Pattern recognition
Week 5	Monday (4-6pm), Tuesday (12-1pm)	Segmentation
Week 6	Monday (4-6pm), Tuesday (12-1pm)	Deep learning
Week 7		No lecture
Week 8	Monday (4-6pm), Tuesday (12-1pm)	Motion and tracking
Week 9	Monday (4-6pm)	Applications
Week 10		No lecture
Week 11	Monday (4-6pm)	Recap

Weekly Class Structure

- Labs: 1-2 PM on Tuesday in weeks 3-6
- Project consultations:
  - On Tuesdays in weeks 7-9
  - Detailed schedule will be announced on class web page
  - Further consultation require appointments with your assigned tutor
- Project demo:
  - On Monday and Tuesday in week 10
  - Detailed schedule will be announced on class web page

ALL changes will be announced on class web page

Assessments

Assessment Type	*Marks*	*Release Date*	*Due in*
Assignment	10%	week 1	week 3
Lab work	5%		weeks 3 - 6
Project	45%*	week 5	
Stage 1 Design and planning			week 7
Stage 2 Implementation			
Stage 3 Demo and report			week 10
* 25% individual and 20% group components			
Exam	40%		exam period

Communication- Modes and Etiquette

- **Online forum (WebCMS3) is your first port of call-** post query of wider interest on lectures, labs, assessments
- **Contact LIC** for late submission, absence, special consideration, assessment deadlines, lab and assessment content
- **Contact Course admin** for issues with enrolment, file submission, group enrolment or any admin matter
- every effort will be made to respond quickly to queries- allow **maximum of 24 hours turnaround**
- Do observe **standards of equity and respect** in dealing with all students and staff- in person, emails, forum posts, all other communication

Late Submissions

- Work submitted up to one day late will incur a penalty of 10% of your mark
- Two days late will incur a penalty of 15%
- Any later may not be accepted at all
- In case of illness or other major issues, contact lecturer PRIOR to deadline, and also submit special consideration application online with supporting documents.

Special Consideration

- UNSW handles special consideration requests centrally. Do not just email the LIC about special consideration.
- Mark calculated in the same way as other students who sat the original assessment
- If you are awarded a Supp and do not attend, then your exam mark will be zero.

Plagiarism Policy

READ the UNSW Policy and Procedure on this.

For the purposes of COMP 9517, plagiarism includes copying or obtaining all, or a substantial part, of the material for your assignment, whether written or graphical report material, or software code, **without written acknowledgement** in your assignment from:

- a location on the Internet
- a book, article or other written document whether electronic or on paper or other medium
- another student, whether in your class or another class
- someone else

Plagiarism (ctd)

- If you copy material from another student or non-student **with acknowledgement**, you will not be penalised for plagiarism, but the marks you get for this will be at the marker's discretion, and will reflect the marker's perception of the amount of work you put into finding and/or adapting the code/text.

The assessments provide opportunities for you to develop important skills. Use these opportunities!

For Reading

- Chapter 1, Szeliski
- Chapter 1, Shapiro and Stockman
- Appendices A and B, Szeliski for background on linear algebra, numerical techniques, statistics

Acknowledgement

- Some images on applications taken from the textbook resources for the text by Szeliski 2010, with original sources credited where possible
- Other images and videos credited where possible