COMP 9517
Computer Vision

Introduction

What is Computer Vision?

Every picture tells a story

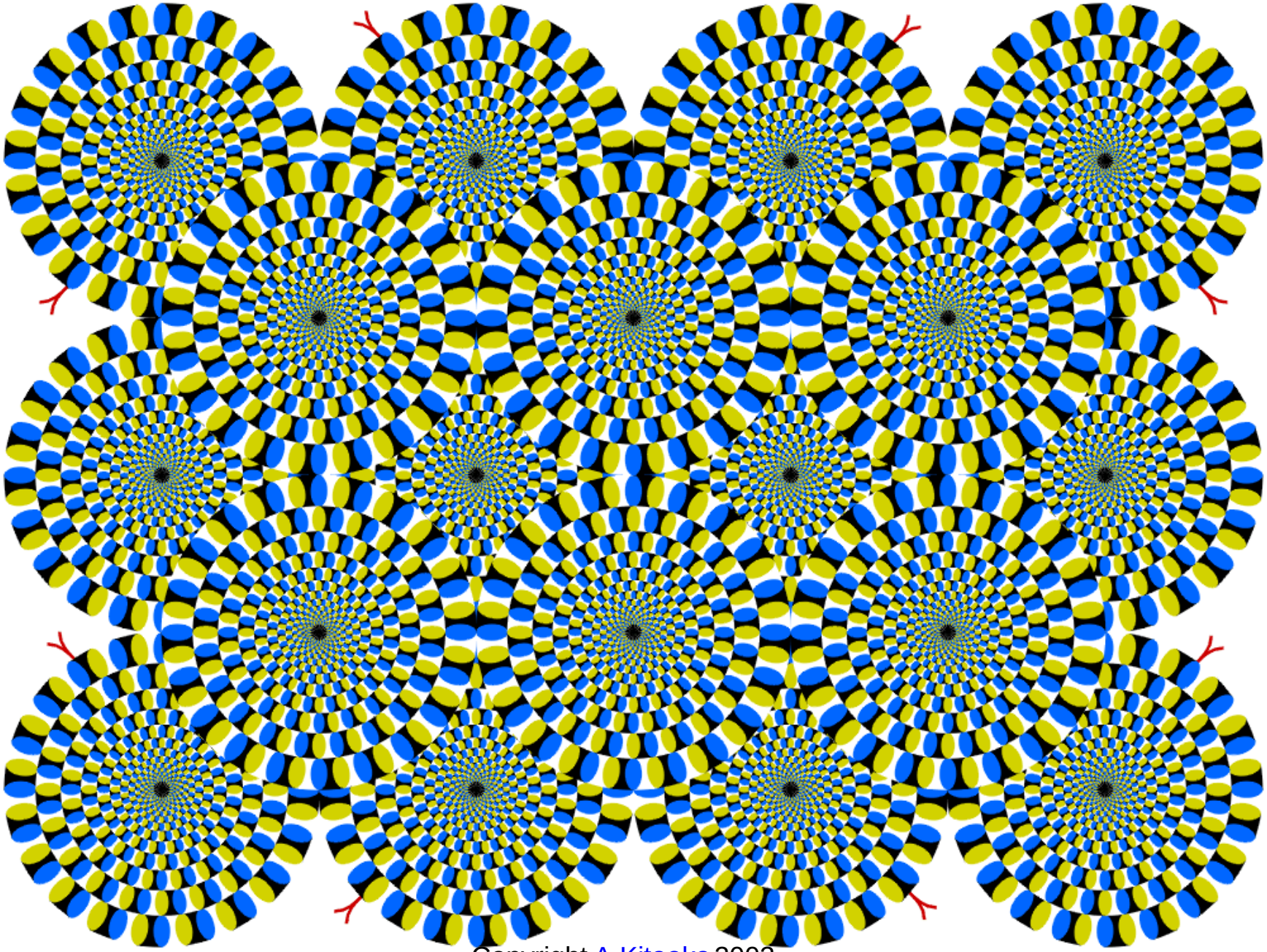
- Goal of computer vision is to write computer programs that can interpret images

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



Copyright A.Kitaoka 2003

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)

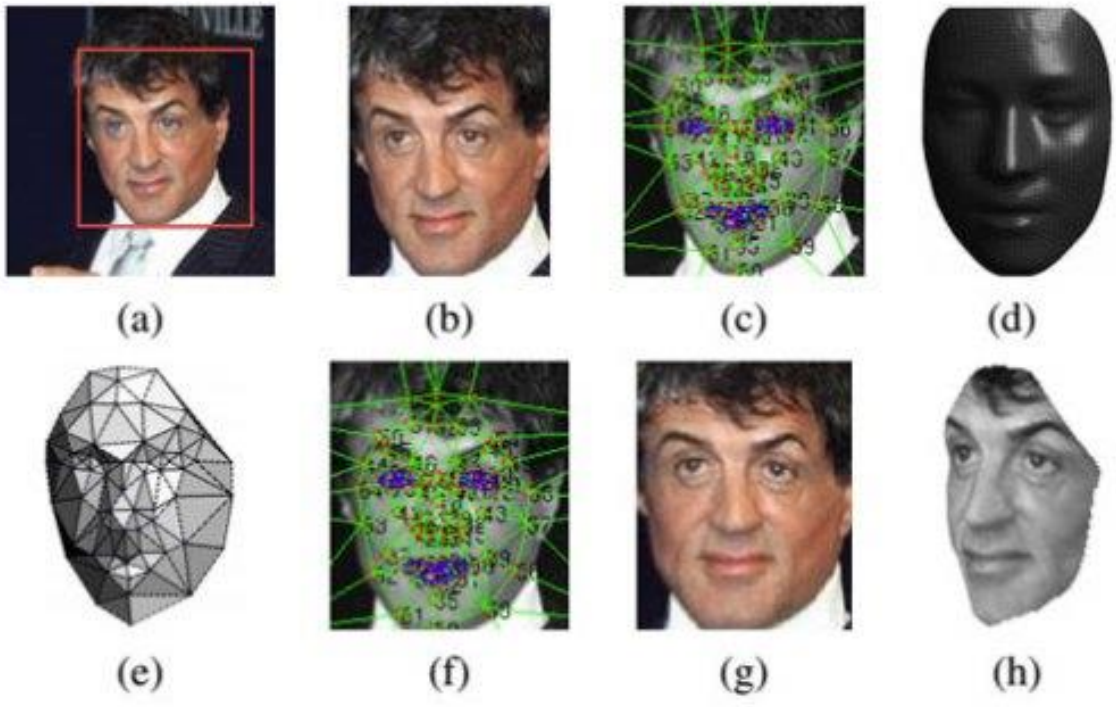
Photo credits - Google

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

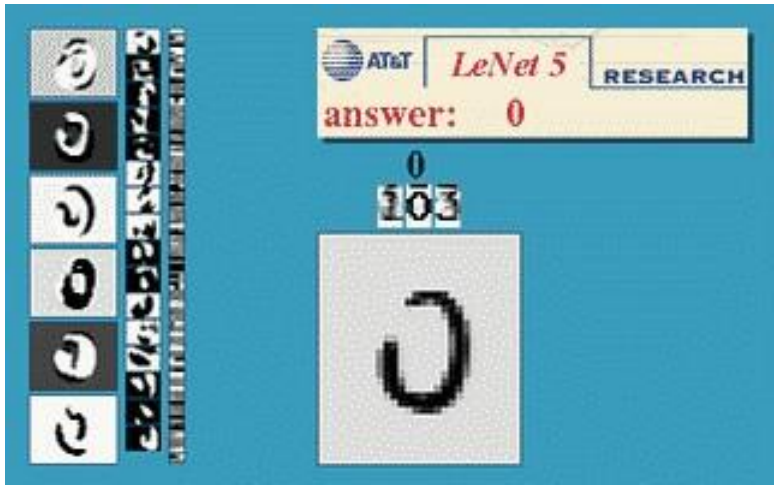
Login without a password…

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



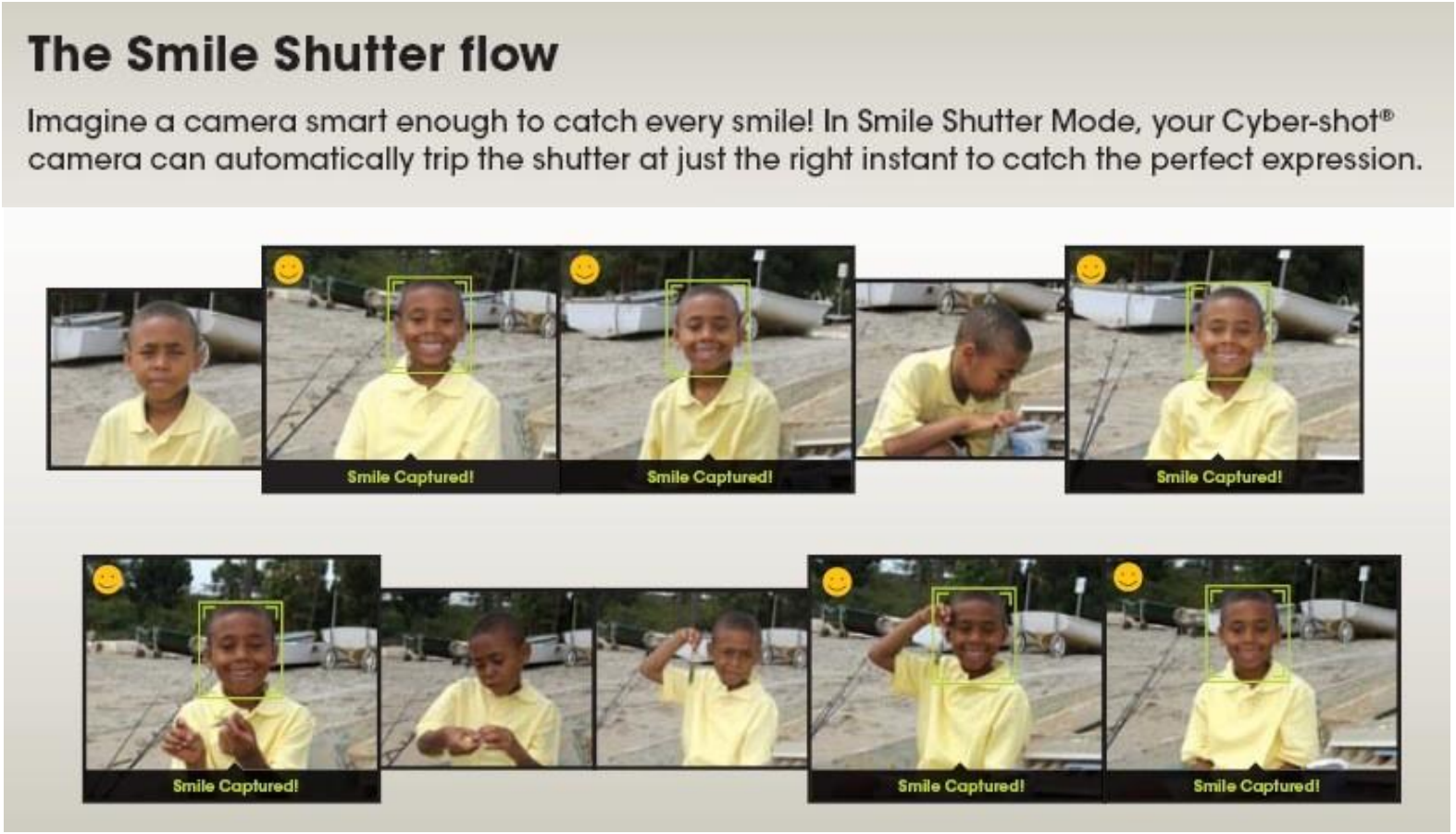
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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?



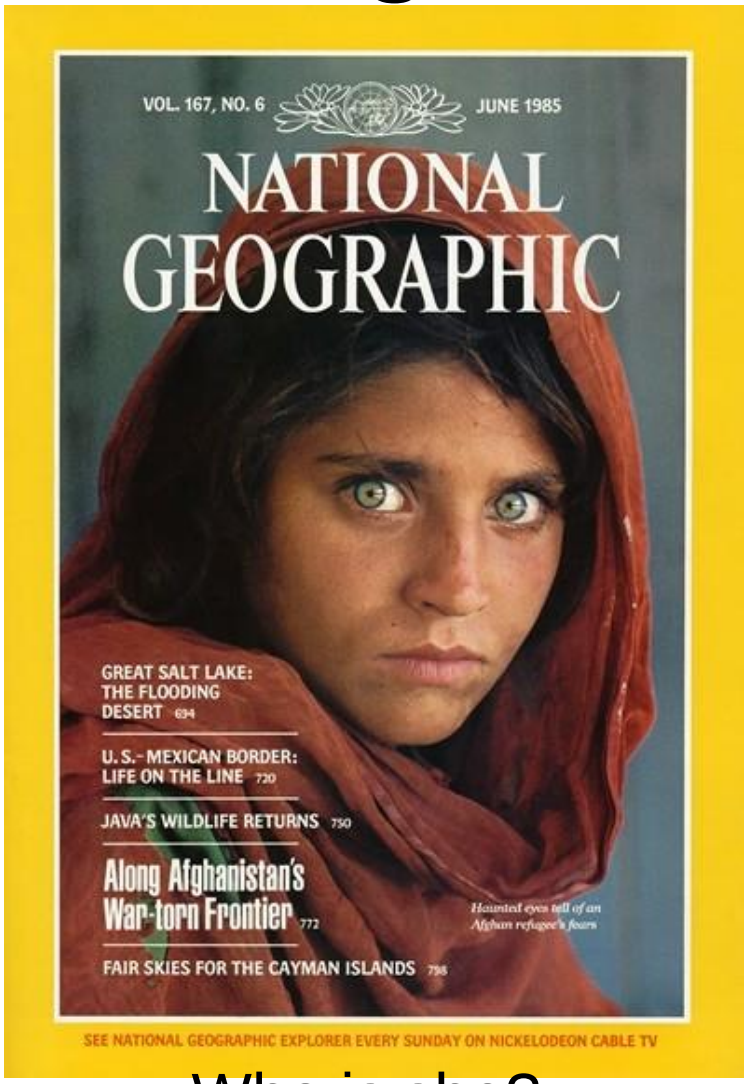
[Sony Cyber-shot® T70 Digital Still Camera]

Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

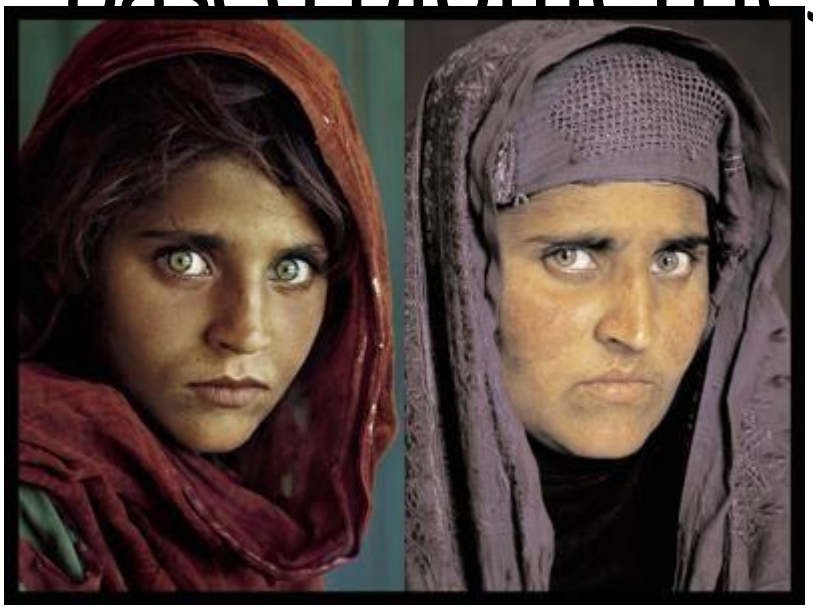
"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you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

Face recognition

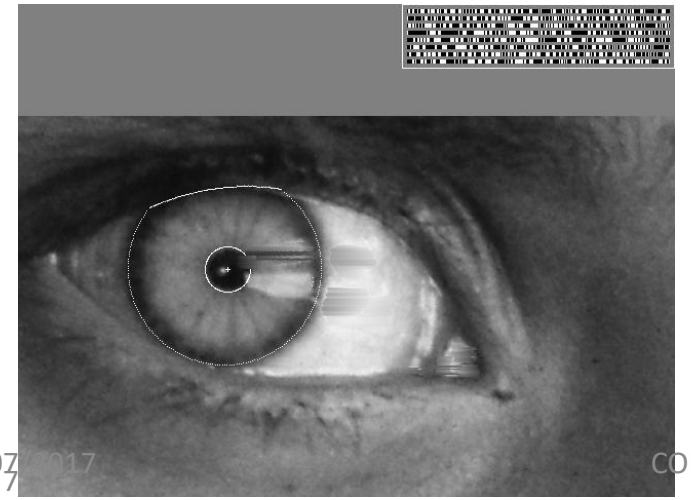


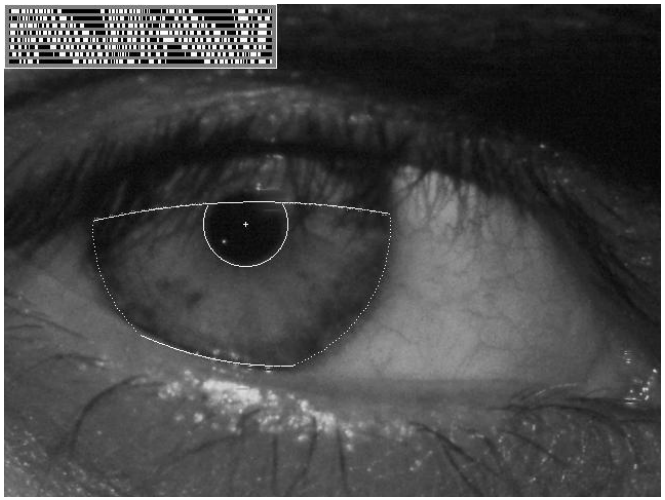
Who is she?

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
  - Point & Find, Nokia

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic
Click here for interactive demo

Sports

Sportvision first down line
How do they superimpose the first-down line on the field on televised football games?
Nice explanation on www.howstuffworks.com

Slide content courtesy of Amnon Shashua

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.

Vision--based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET

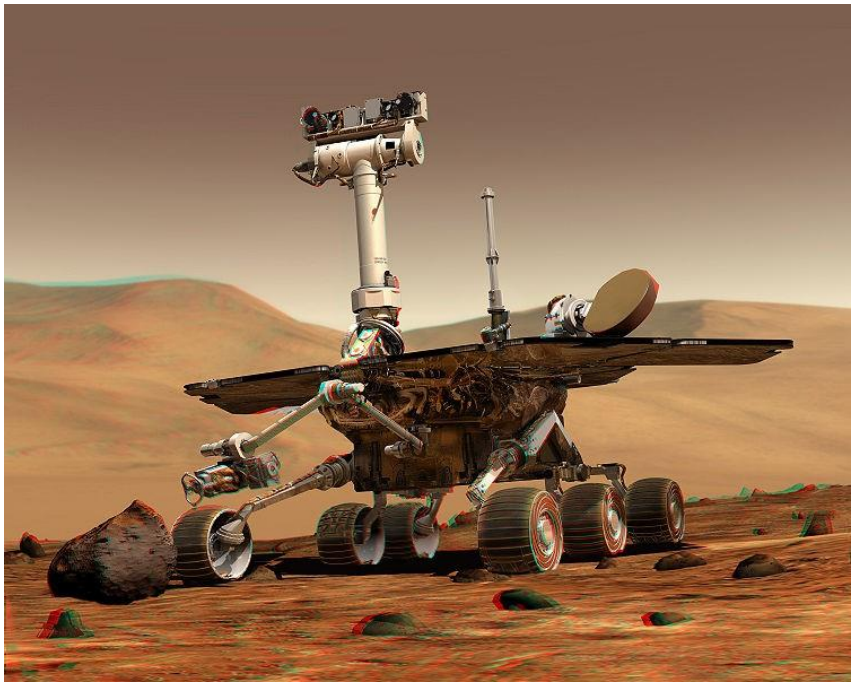
Camera tracking a crowd, based on this work.

Vision in space

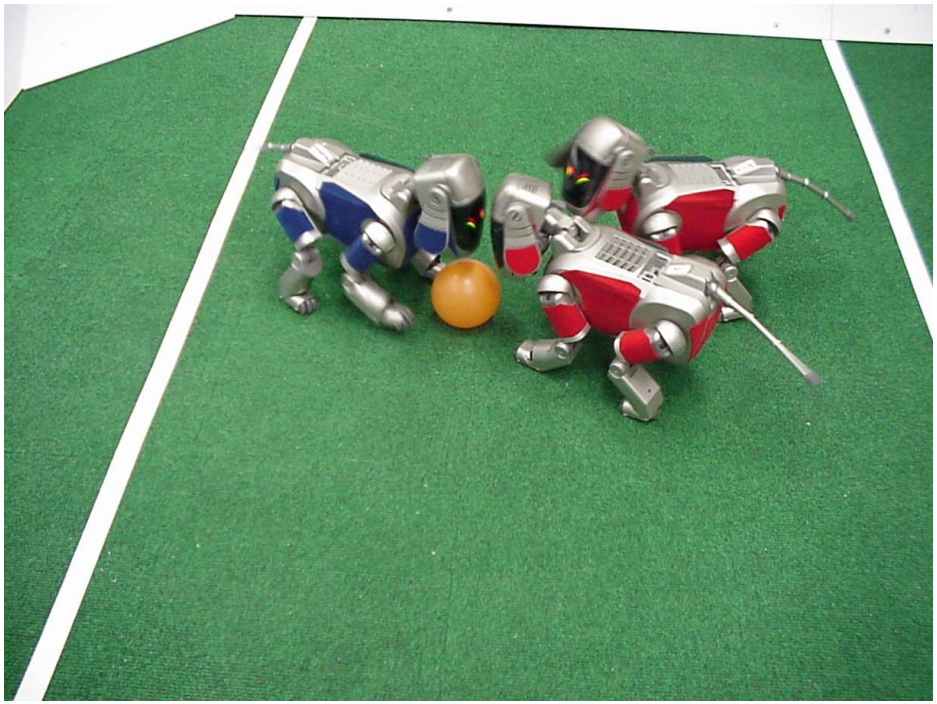
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

- 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

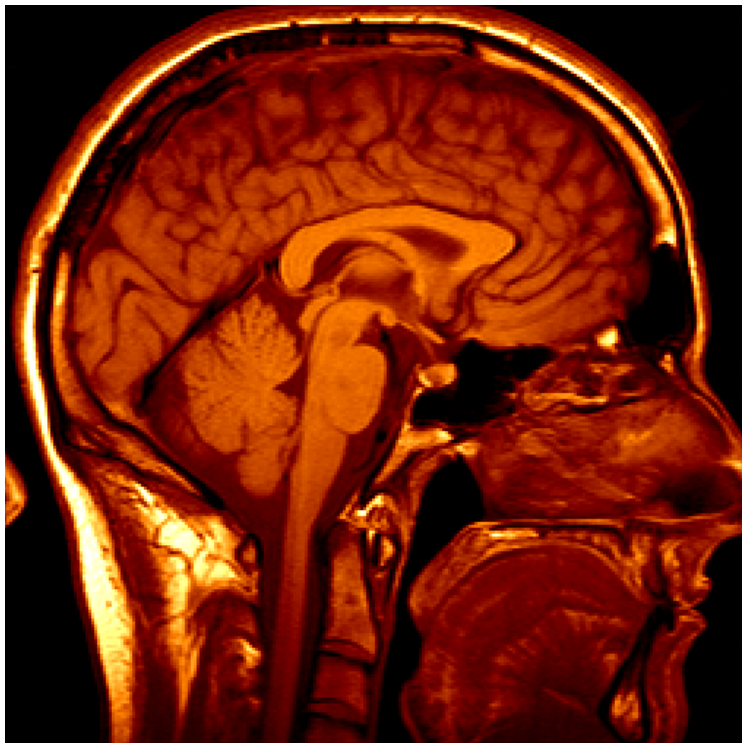


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

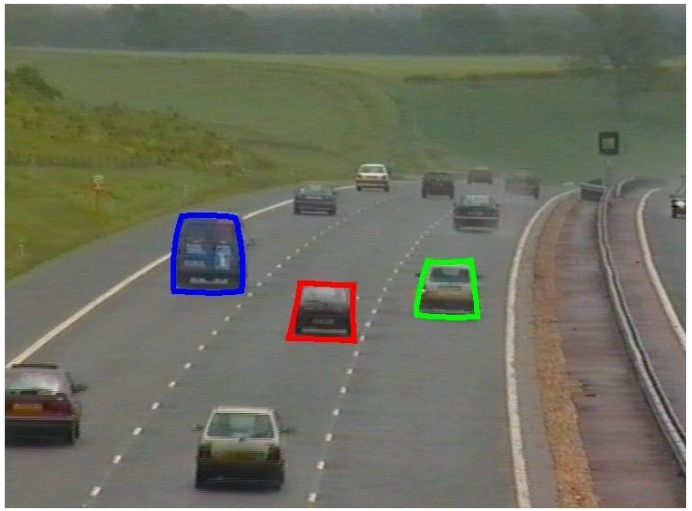
Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

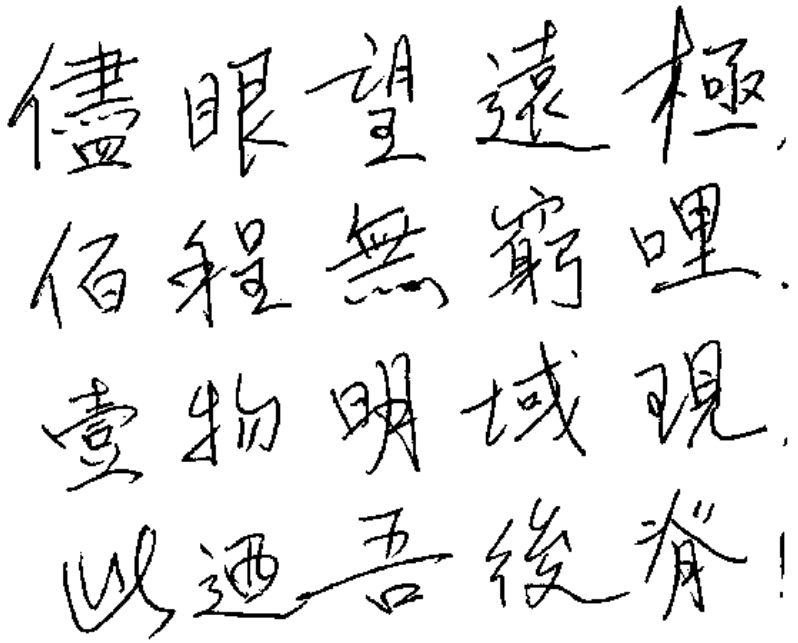
- Video Surveillance
  - Traffic Monitoring
  - Object tracking
  - Action recognition, driving, stopping, etc
  - Vehicle speed
  - Counting
  - Challenge: occlusion, illumination changes and non--linear speed



- Image/video retrieval
  - Content--based retrieval
  - Search engine
  - Challenge: big data volume, semantic



- Text Recognition
  - Converting information from paper documents into digital form
  - Challenge: semantic interpretation



I looked as hard as I could see, beyond 100 plus infinity an object of bright intensity-- it was the back of me!

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
  - combine the information available in multiple images into a coherent whole
  - 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.

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?

Computer Vision Processes

- Low level processes
  - use little knowledge of image content
  - include image compression, noise filtering, edge extraction, ...
  - use data which resemble the input image, eg. matrix of picture elements
- High level processes
  - based on knowledge, goals, plans
  - use Artificial Intelligence methods
  - simulate human cognition and decision making based on information in the image
  - cognitive processes, geometric models, goals, plans,...

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 to program well in any of C/C++, Java, 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 Matlab
- be familiar with vector calculus and linear algebra, or be willing to learn them independently

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

Course Changes

1. Labs redesigned: 6 compulsory labs: to provide more coverage of concepts and algorithms
2. Automarking of assignments: to address marking consistency, better utilisation of lab time
3. Changes in assessment
   - lab work will be assessed, to provide more feedback
   - frequent short quizzes replaced by 2 (longer) tests
   - fewer reports to write!
4. Consultations redesigned
   - Contact consultants by multiple modes
   - Track project consultations via meeting logs
   - Watch out for times and modes on class webpage
   - Forums will continue, useful to all when used by you responsibly!!

Weekly Class Structure

1. Lectures, 12-2 PM most weeks, except in test / presentation weeks- watch out for notices in class, as well as class web page
2. Labs: 2-3 PM in weeks 2, 3, 4, 5, 6, 9
3. Consultations: outside regular class times, schedule will be announced on class web page; some consultations in lab hours, subject to availability of time; group project consultations may require appointments with your assigned tutor

ALL changes will be announced on class web page

Assessment

- 2 assignments, online submission, automarking
- 2 class tests (<= 1 hour each)
- Group project, groups of 3
- Lab mark for in-lab performance

Assessment Details

Assessment Type	*Marks*	*Release Date*	*Due in*
Assignment 1	10%	Week 2	Week 4
Assignment 2	15%	Week 4	Week 6
Test 1	20%		Week 7
Test 2	20%		Week 11
Project (multiple stages)**	30%	Week 6	
• *Interim Project Presentation*			Week 8
• *Demo*			Week 13
• *Project report*			Fri of Week 13
Lab mark	5%		

For Reading

- Chapter 1, Szeliski
- Chapter 1, Shapiro and Stockman
- Appendices A and B, Shapiro 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
- Images and Videos credited where possible

For Reading

- Chapter 1, Szeliski
- Chapter 1, Shapiro and Stockman
- Appendices A and B, Shapiro 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
- Images and Videos credited where possible