CDTM Drones Lecture 2016

Basics of Computer Vision
Part 1
Introduction

Seyed-Ahmad Ahmadi
Dept. Neurology, LMU Klinikum Großhadern

Outline of this part

- About myself
- Present the problem at hand
- To understand basics of computer vision
- History of CV
- Domains of CV
- „What works today“

About myself

- 2001-2008: Double Master Electrical Engineering/Information Technology @ TUM / Georgia Tech
- 2003-2005: CDTM
  - Trendreport for FIFA World Cup 2006: Handy E-Tickets with „emotional value“
  - MPD: Location-based Services via crowd-based gamification, „Doggy Style“
  - E-lab: consultancy of a one-man company developing a digital patient file for clinics
- Gradually defected to computer science

About myself

- 2005: Bachelor Thesis with CAMP chair (Prof. Navab, I-16): Surgical workflow analysis in minimally invasive surgery
- 2008: Master thesis at CAMP: Surgical activity detection through body+hand motions, measured with accelerometers
- 2008-2012: PhD thesis on Ultrasound for early detection of Parkinson‘s Disease
- 2012 (ongoing): Post-Doc at Klinikum Grosshadern, Dept. Neurology; medical image computing for neurological movement disorders
- From Nov. 2016: Post-Doc at DSGZ for Big Data Mining and Deep Learning in sensor data of vertigo/balance disorder patients

Problem setting of this course

- (Semi-)autonomous flight of a parkour
- Parrot AR.Drone has 2 cameras and ultrasound sensors for landing
  → Video is primary sensor of AR.Drone
  → Computer Vision/Image Processing for navigation
- Use a sophisticated artifical neural network for robust object detection and tracking, following it down the race track

Aim of this lecture

- To understand basics of image processing and computer vision
- To better understand digital images
  - Color, Intensity distributions, frequencies
- To help you understand the underlying mechanics (filters/convolutions) of convolutional neural networks (CNN) which you will use for robust object tracking and (semi-)autonomous drone flight
- To give you perspective of how advanced this technology actually is

What is Computer Vision?

- „**Computer vision** is a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information, *e.g.*, in the forms of decisions.“
- „As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images, [...], such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner.“

History of Computer Vision

- According to
  - Efstratios Gavves and James Hays,
    http://www.egavves.com/a-brief-history-of-computer-vision/
    http://cs.brown.edu/courses/cs143/
- Early days (1960‘s/70‘s):
  - Understanding simple geometric objects in synthetic worlds, early neural networks experiments
- Middle Ages (1980‘s-mid 90‘s):
  - More geometry, more mathematical rigor, modeling, first image filters and image processing methods (Photoshop), simple text recognition (OCR)

History of Computer Vision

- Golden Years (mid 1990‘s-2010):
  - Face recognition (eigenfaces), statistical analyses, automatic detection of meaningful points in images, better features and feature descriptors (SIFT), compact and fast representations of content in images (dictionaries, bag of words)
  - Better Machine Learning methods! Most notably, Support Vector Machines (SVMs) and Ensemble Learners (Boosting, Forests)
  - First public databases and challenges
  - Open source code/libraries (OpenCV)
  - Social Media, beginning of Big Data era
  - Modern CPUs and GPUs, parallelization
  - First meaningful video processing methods

History of Computer Vision

- Today
  - Abundance of data and processing power/ memory
  - Modern AI (Deep Learning, neural-networks AI), with performance on par or better (!) than human performance
  - Big Tech companies investing in powerful Computer Vision and AI algorithms&projects (Google, Facebook, Microsoft, Baidu etc.)

History of Computer Vision

- Tomorrow
  - Higher-level interpretations, beyond object detection... Interaction of objects, moving objects/videos, towards natural language interfaces for search engines

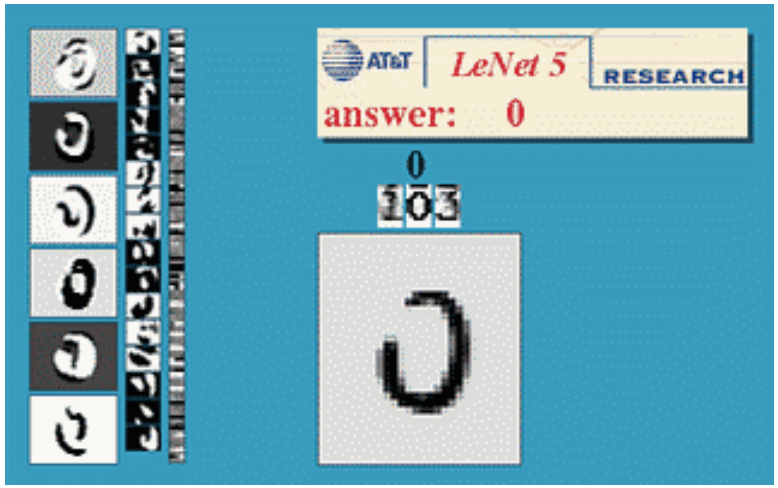
Domains of Computer Vision

- Image enhancement
- Transformations
- Filtering, Fourier and wavelet transforms and image compression
- Color vision
- Feature extraction
- Pose estimation
- Registration
- Visual Recognition
- Etc...

Slide credits:
James Hays, Brown University,

„What works today“

- Optical Character Recognition (OCR)



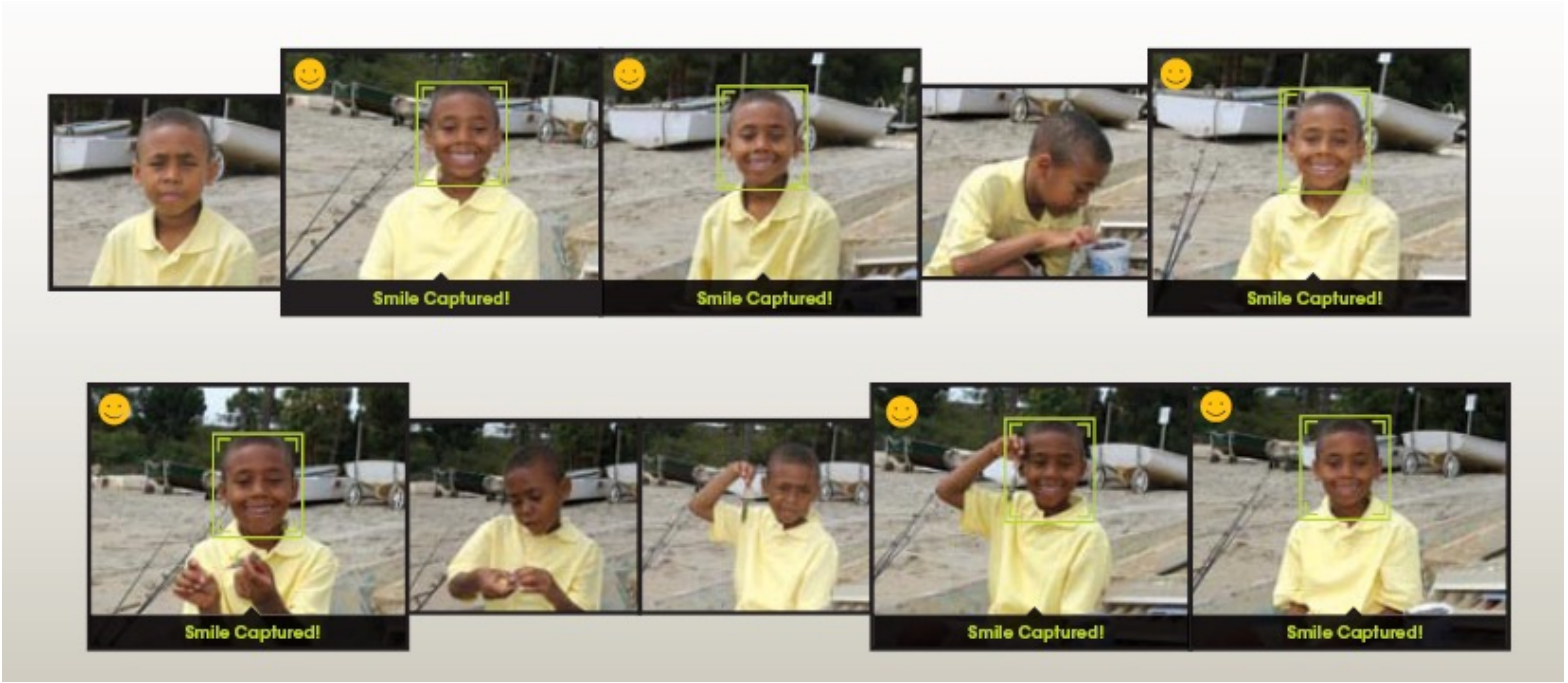
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Slide credits:
James Hays, Brown University,

„What works today“

- Face/Smile recognition in cameras (embedded)



Sony Cyber-shot® T70 Digital Still Camera

Slide credits:
James Hays, Brown University,

„What works today“

- 3D from thousands of images https://youtu.be/GdPeydPbM0g

Slide credits:
James Hays, Brown University,

„What works today“

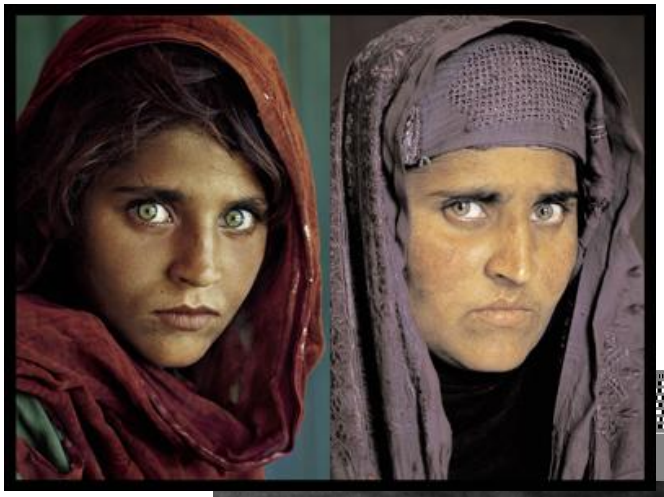
- Object recognition in supermarkets

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… “

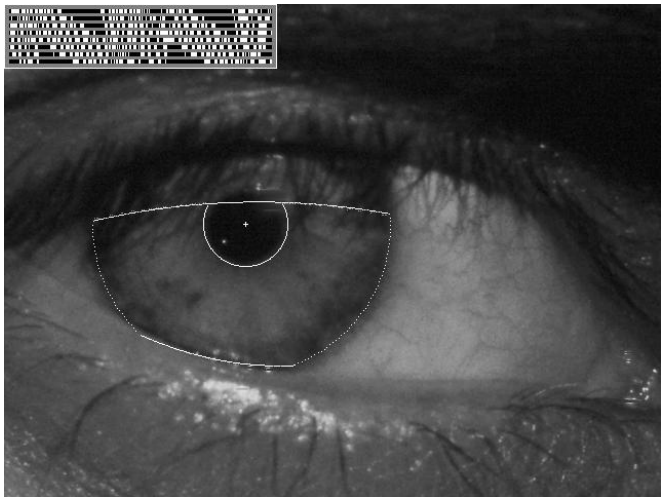
Slide credits:
James Hays, Brown University,

„What works today“

- Biometrics (iris, fingerprint etc.)



"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story wikipedia



Slide credits:
James Hays, Brown University,

Object recognition / AR (in mobile phones)

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

[Google Goggles]()

Slide credits:
James Hays, Brown University,

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Slide credits:
James Hays, Brown University,

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Slide credits:
James Hays, Brown University,

Sports

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

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

Slide credits:
James Hays, Brown University,

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.

Slide credits:
James Hays, Brown University,

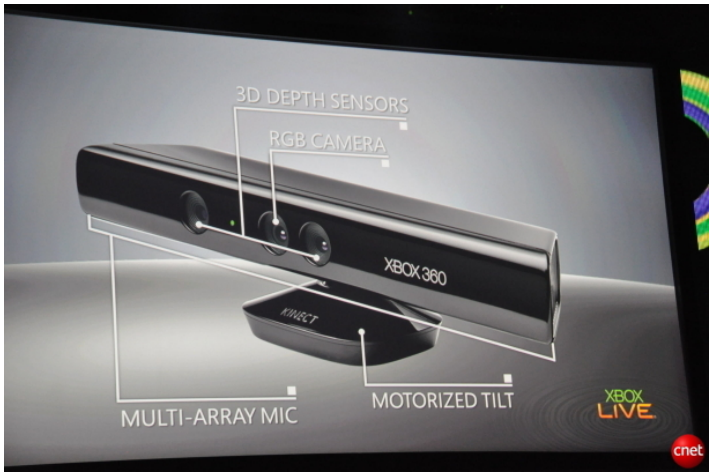
Google cars

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)

Slide credits:
James Hays, Brown University,

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



Slide credits:
James Hays, Brown University,

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.

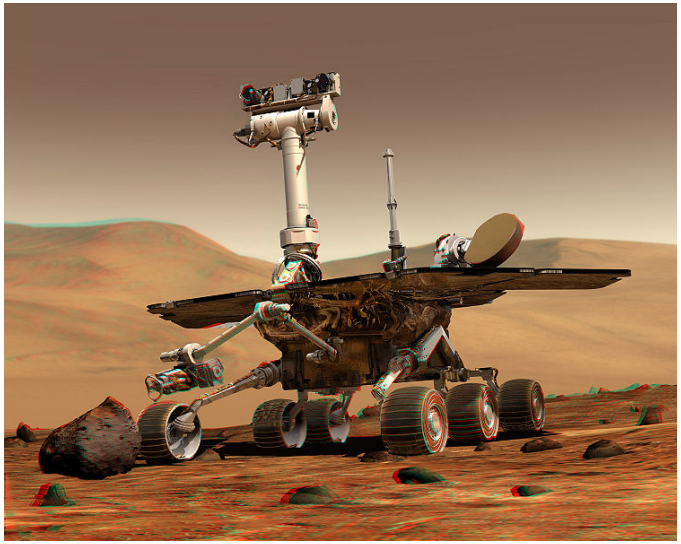
Slide credits:
James Hays, Brown University,

Industrial robots

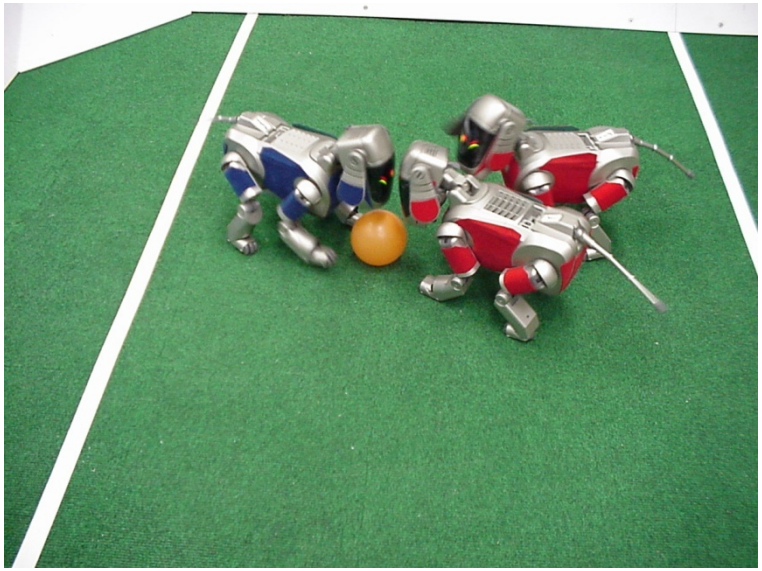
Vision-guided robots position nut runners on wheels

Slide credits:
James Hays, Brown University,

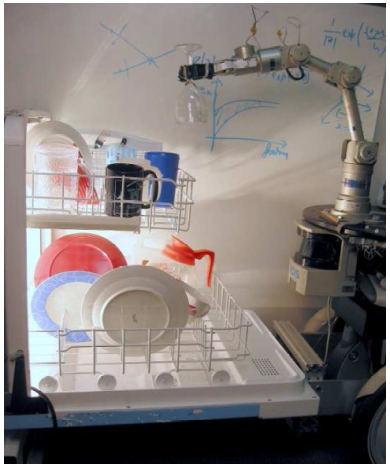
Mobile robots



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



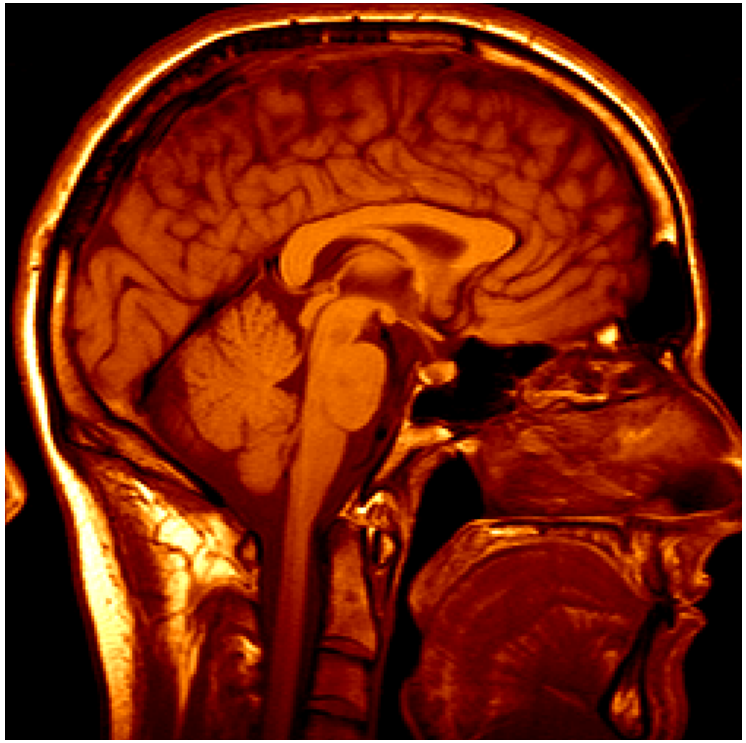
http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Slide credits:
James Hays, Brown University,

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Slide credits:
James Hays, Brown University,

Computer Vision and Nearby Fields

- Computer Graphics: Models to Images
- Comp. Photography: Images to Images
- Computer Vision: Images to Models

Questions?