Computer Vision

António J. R. Neves / Paulo Dias

Departamento de Electrónica, Telecomunicações e Informática

Universidade de Aveiro

an@ua.pt / paulo.dias@ua.pt
http://elearning.ua.pt/

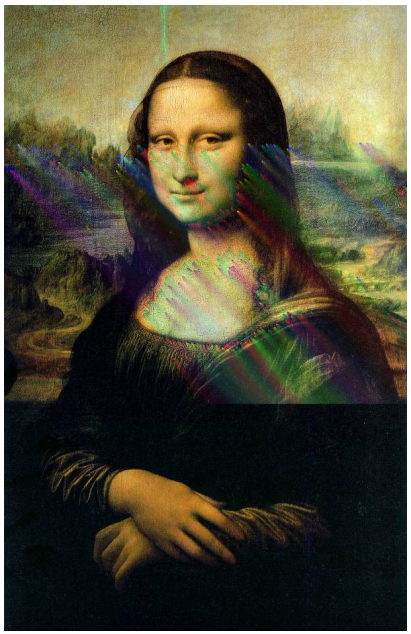
Contents

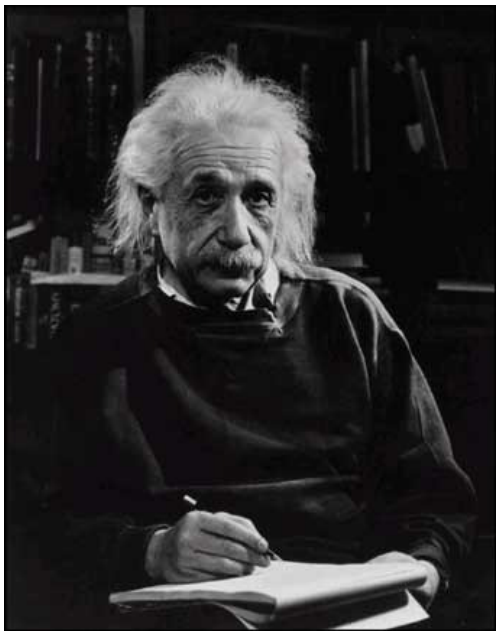
Contents

Motivation

Every image tells a story...

A picture is worth than millions words





Motivation

- Goal of computer vision: perceive the “story” behind the picture.
- Compute properties of the world: 3D shape, names of people or objects, what happened?

Can the computer match human perception?

- Yes and no (mainly no): computers can be better at easy things; humans are much better at hard things.
- But huge progress has been made in the last years: what is considered hard keeps changing.

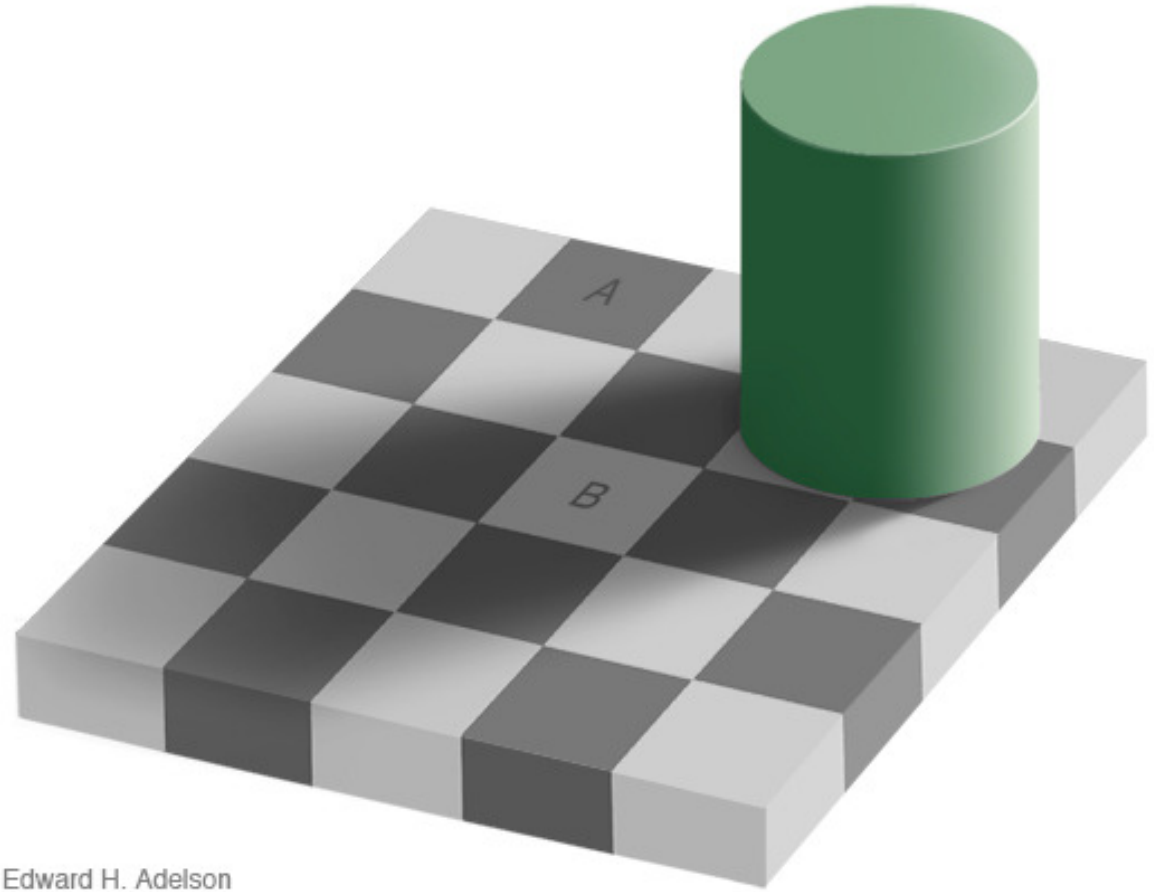
Human vision (1)

- Vision is a complex physical and intellectual human task that stands as a primary interaction tool with the world.
- It is a complex process not completely understood, even after hundreds of years of research.
- The visualization of a physical process involves an almost simultaneous interaction of the eyes and the brain.
- This interaction is performed by a network of neurons, receptors and other specialized cells.

Human vision (2)

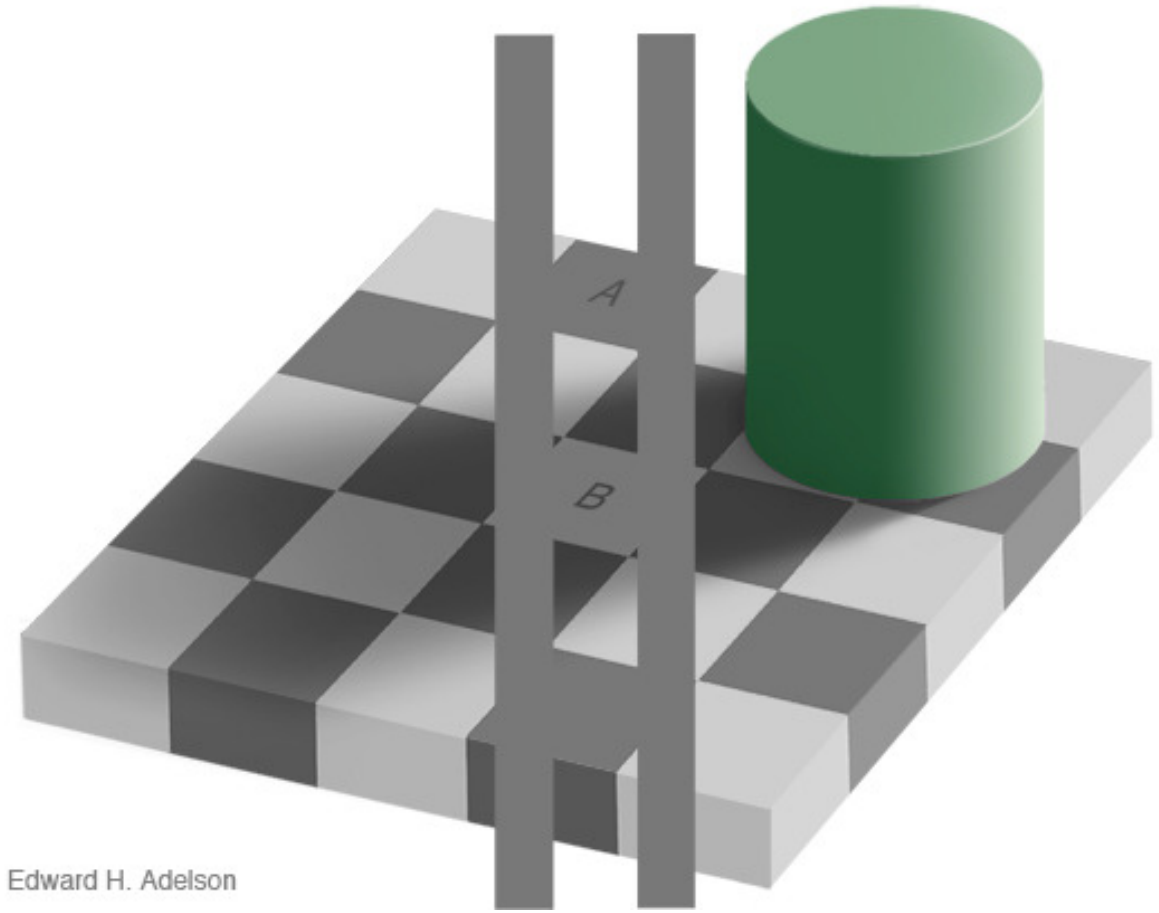
- The human eye is equipped with a variety of optical elements, including the cornea, iris, pupil, a variable lens and the retina.
- Can do amazing things like:
  - Recognize people and objects
  - Navigate through obstacles
  - Understand mood in the scene
  - Imagine stories
- But:
  - Suffers from illusions
  - Ignores many details
  - Ambiguous description of the world
  - Doesn't care about accuracy of world

Illusions (1)



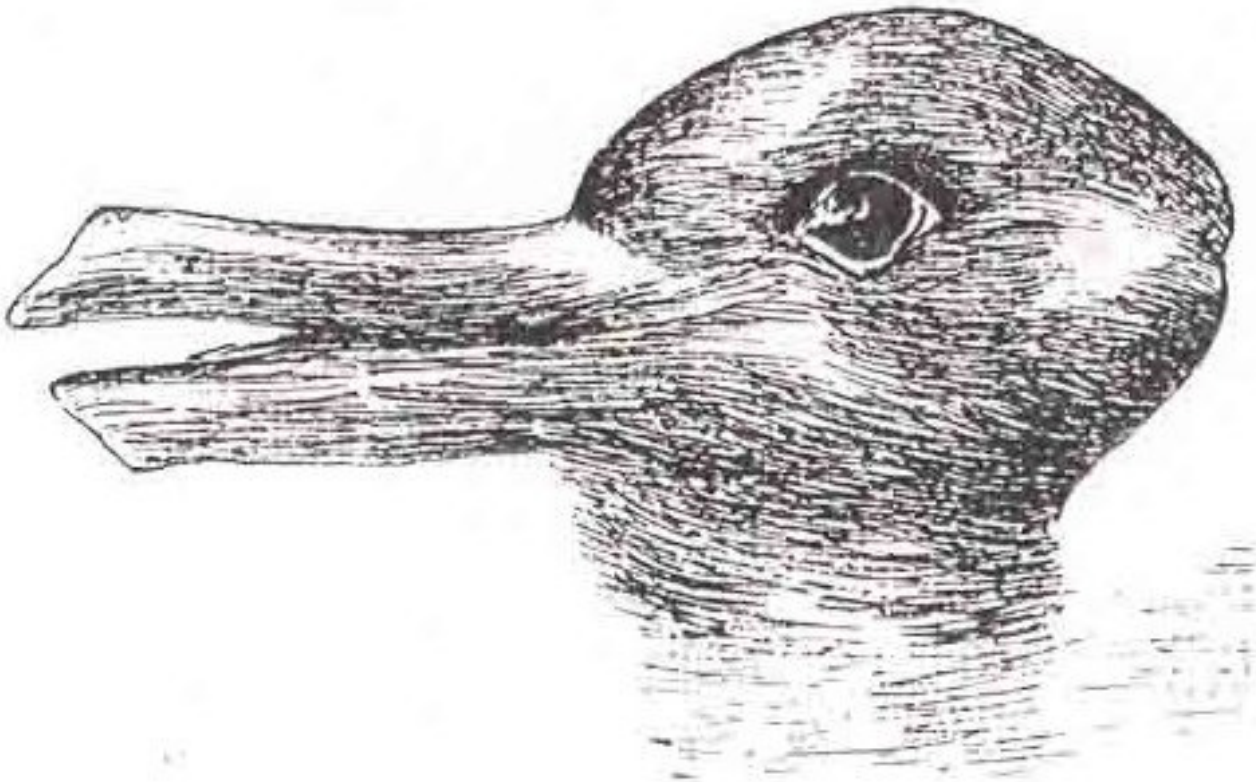
Edward H. Adelson

Illusions (2)



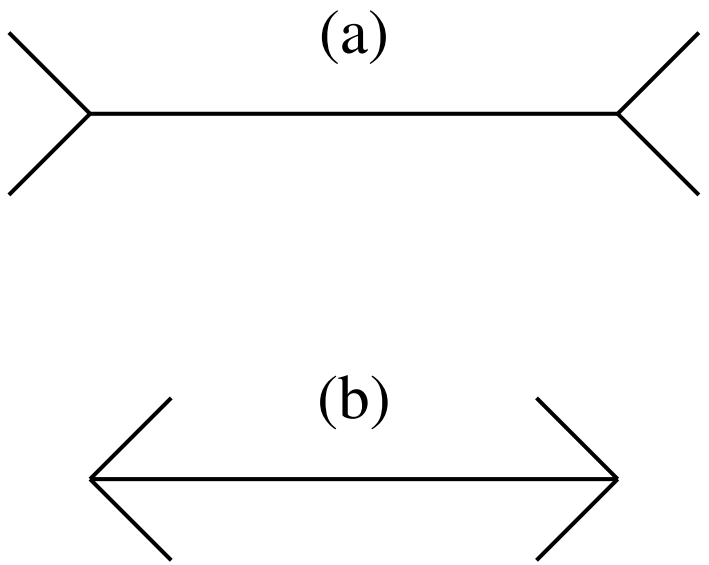
Edward H. Adelson

Illusions (3)

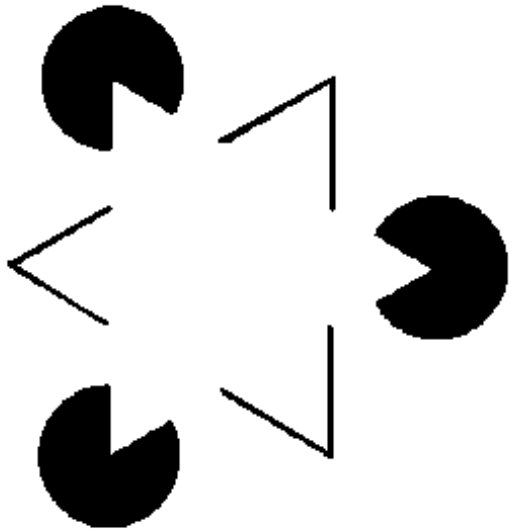


Other illusions...

- The human visual system exhibits a considerable cognitive component, influenced by memory, context, and intention:



Which is the longer one?

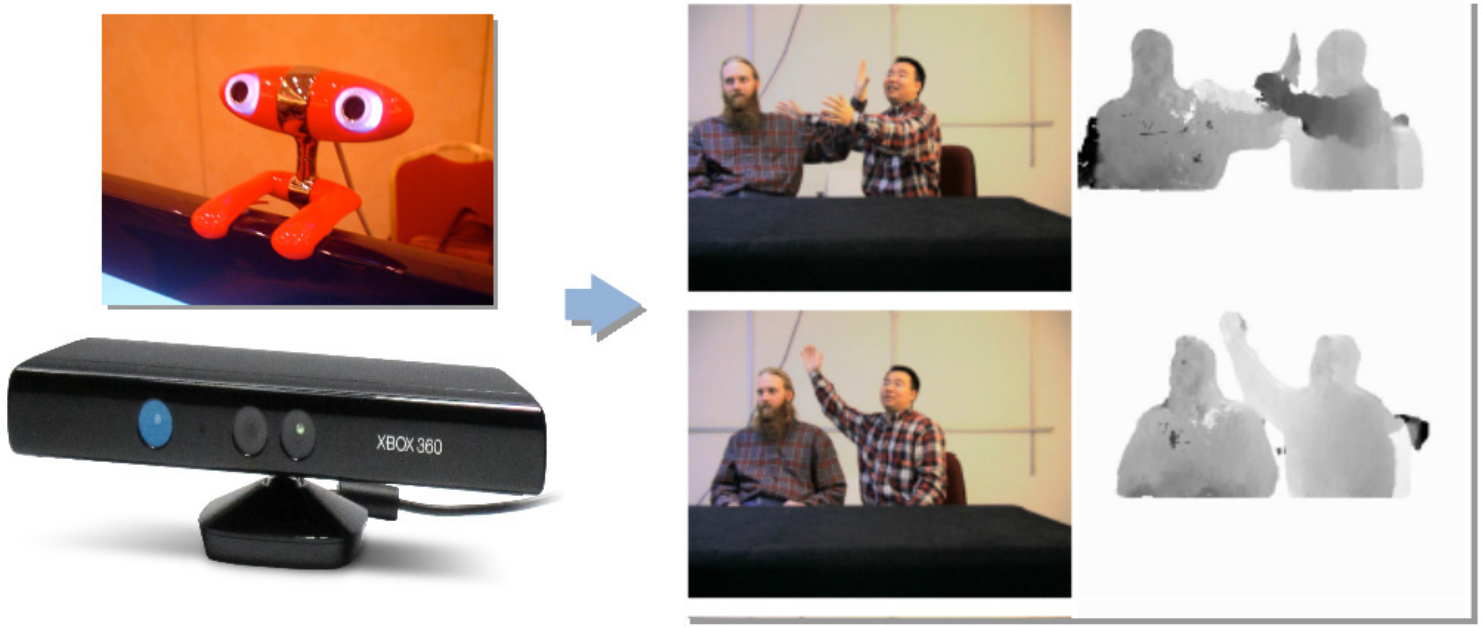


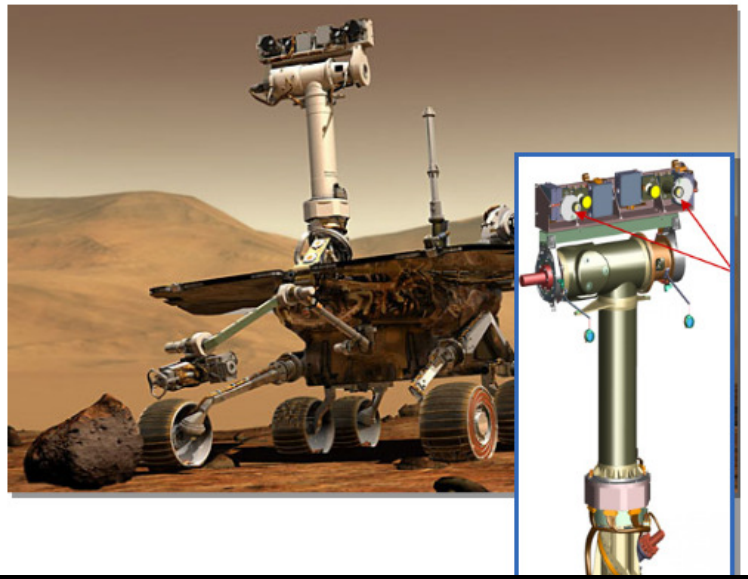
A triangle?

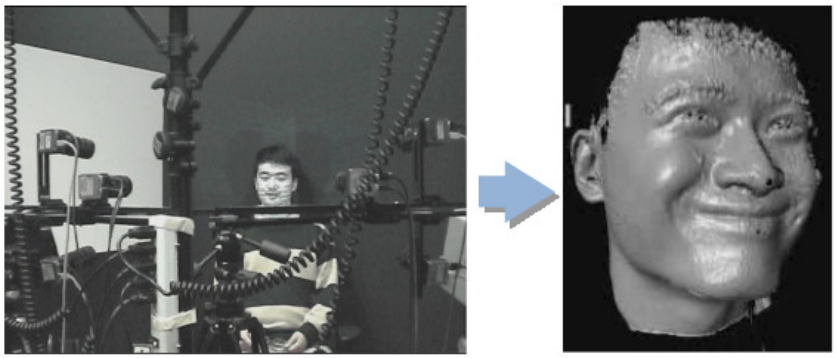
Contents

Goal of Computer Vision

Compute 3D shape of the world

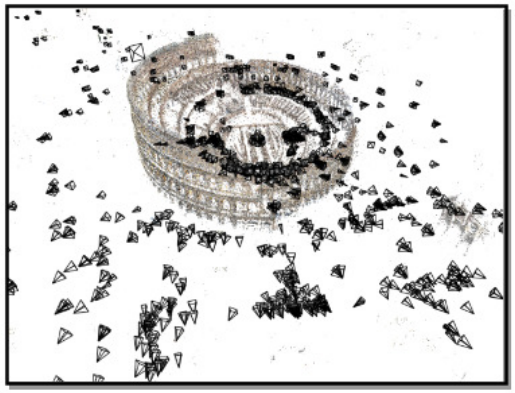






Compute 3D shape of the world (2)

Internet Photos (“Colosseum”)



Reconstructed 3D cameras and points

Dense 3D model

Optical character recognition (OCR)

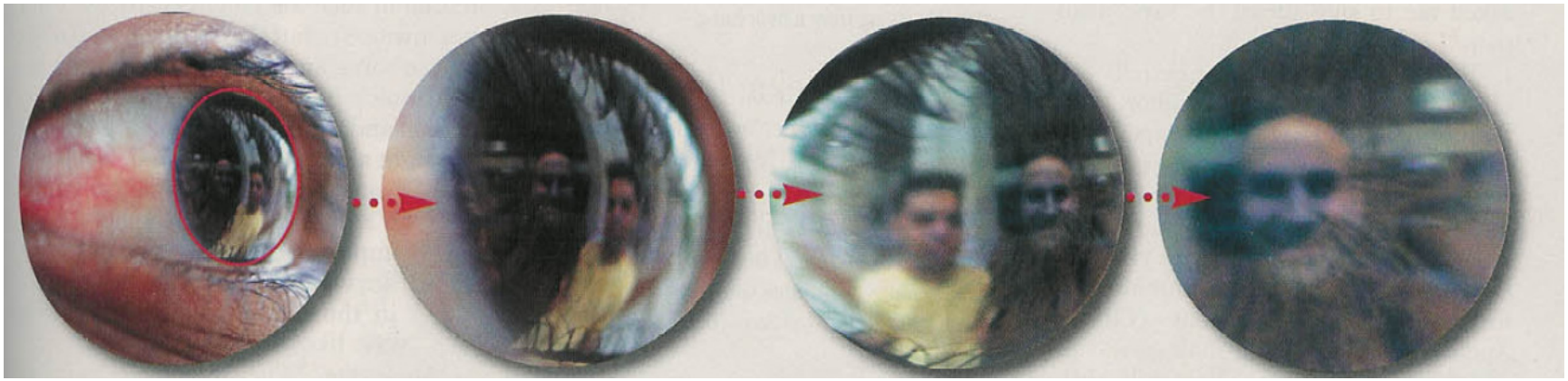
Digit recognition, AT&T labs

License plate readers

Automatic check processing

Sudoku grabber

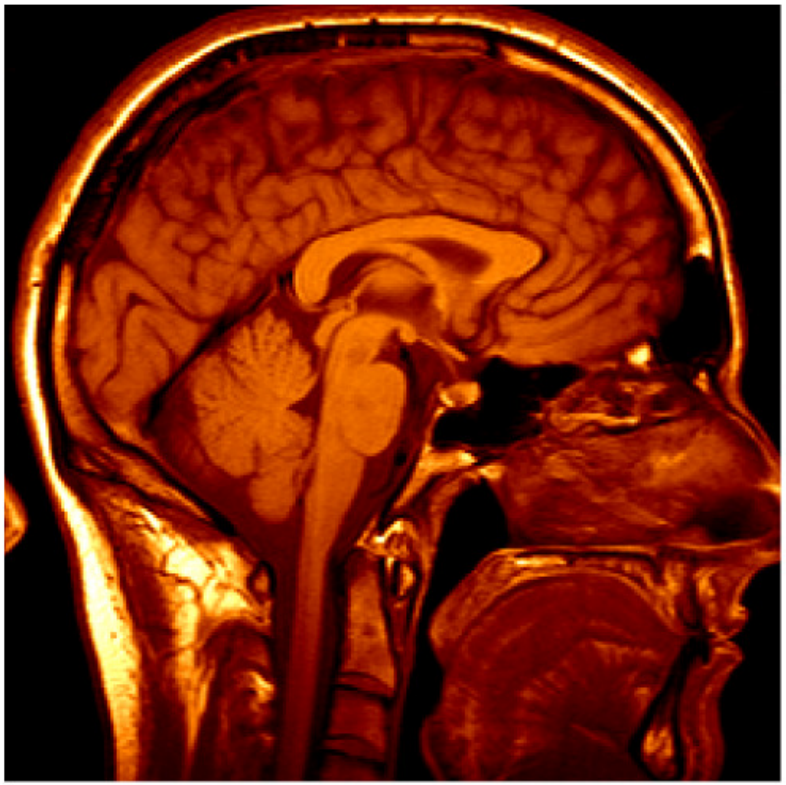
Forensics



Face detection / recognition

http://www.face-rec.org/

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Space

The Heights of Mount Sharp
http://www.nasa.gov/mission_pages/msl/multimedia/pia16077.html
Panorama captured by Curiosity Rover, August 18, 2012 (Sol 12)

Vision systems (JPL) uses for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

Smart Cars

- Mobileye
  - Vision systems currently in high-end models

Vision based interaction

Ex: camera-based IR tracking.

Assistive technologies

Sports


OAK 7 4TH
TEN 14 11:04


Day 5: Swimming - Men's 4X200M Final
UNITED STATES
650M
5:38.01
07:13 / 14:25
Highlights of the men's 4x200m relay final on Day 5.

Shape and motion capture

Object detection

Vision as sensor in robotics

Computer Vision

- Vision is a complex physical and intellectual human task that stands as a primary interaction tool with the world.
- Computer vision is a field that includes methods for acquiring, processing, analyzing, and understanding images ...
- Computer vision applications are increasing:
  - surveillance;
  - machine inspection;
  - medicine;
  - robotics;
  - entertainment;
  - media.
- The main goal: make computer vision converge towards human vision. Can we ever accomplish that?

Why is computer vision difficult?

Viewpoint variation

Illumination

Scale

Why is computer vision difficult? (2)

Intra-class variation

Motion (Source: S. Lazebnik)

Background clutter

Occlusion

Objectives

- Computer vision seeks to develop algorithms that replicate one of the most amazing capabilities of the human brain - inferring properties of the external world purely by means of the light reflected from various objects to the eyes.
- With vision, it is possible to determine how far away objects are, how they are oriented with respect to the subject, and in relationship to various other objects.
- It is possible to guess their colors and textures and recognize them.
- It is possible to segment regions of space corresponding to particular objects and track them over time.
- In this course, we will study some of the concepts and algorithms used in Computer Vision to achieve the referred tasks . . .

Topics to be covered

- Image acquisition and representation
  - digital cameras, digital images, color spaces, ...
- Low-level image processing
  - neighbors, filtering, histograms, contours, morphological operators ...
- Digital camera calibration
  - intrinsic parameters, distortion correction, color calibration, ...
- Stereo image processing
  - camera calibration, 3D reconstruction, ...
- 3D imaging
  - 3D cameras, point clouds, ...
- Video processing
  - egomotion, tracking, optical flow, ...
- High-level image processing
  - template matching, pattern recognition, descriptors, ...

Organization of te classes

- 4 classes low-level image processing
- 4 classes camera calibration and 3D image processing
- 4 classes high-level image processing
- 2 classes project development

Contents

Methods

- Expository lectures
  - some invited talks
- Laboratory work
  - use of digital cameras
  - OpenCV - open source based C/C++/Phyton Computer Vision library (`http://docs.opencv.org/`)
- Homework

Contents

Grading

- Theoretical component: Final exam 40 %
- Practical component: 60 %
  - In class challenges 25 %
  - Final project + presentation 75 %

- The platform `code.ua.pt` will be used as repository of the software developed, as well as all the files that the groups need or produce.
- Each group should create a project for the course `vc1617` and subversion or git can be used as repository of the source code developed during the classes and for the final project.

Reports

- The template of the DETI journal should be used to produce your reports (available in the course website).
- Academic dishonesty cannot be condoned. Academic dishonesty, as a general rule, involves one of the following acts:
  - Cheating on an examination or quiz.
  - Substituting for another person during an examination or allowing such substitution for one's self.
  - Plagiarism. This is the act of appropriating passages from the work of another individual, either word for word or in substance, and representing them as one's own work. This includes any submission of written work other than one's own.
  - Collusion with another person in the preparation or editing of assignments submitted for credit, unless such collaboration has been approved in advance by the instructor.

Contents

Bibliography

Textbook

- Richard Szeliski, Computer Vision: Algorithms and Applications, Springer-Verlag London, 2011 (Available online: `http://szeliski.org/Book/`).

Other references

- Making Things See, Greg Borenstein, O'Reilly 2012
- Learning OpenCV: Computer Vision in C++ with the OpenCV Library, Gary Bradski, Adrian Kaehler, O'Reilly 2012
- Machine vision: Theory, algorithms, practicalities, E. R. Davies, Morgan Kaufmann 2005.
- Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Prentice Hall, 2007
- Image Processing: Analysis and Machine Vision, Milan Sonka et al., Chapman & Hall, 2007
- D. Forsyth, J. Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2002.