Tesla: http://www.youtube.com/watch?v=mCj_C1NOVxw

BLG453E COMPUTER VISION

Fall 2017

http://www.bostondynamics.com/atlas
Min: 1:10

Istanbul Technical University
Computer Engineering Department

Instructor: Prof. Gözde ÜNAL

Teaching Assistant: Enes ALBAY

http://www.wired.com/story/why-john-deere-just-spent-dollar305-million-on-a-lettuce-farming-robot/
http://computervisiononline.com/blog/put-away-pesticides

Computer vision: a scientific discipline that studies how computers can efficiently perceive, process, and understand visual data such as images and video

A few examples…

What you will learn in this course

Fundamentals of computer vision

- Vision applications
- Digital images, Pointwise Image Processing
- Spatial Transforms = Filtering
- Geometric Transforms (=Coordinate Transformations)
- Feature extraction, including edges, corners, keypoints
- Segmentation: Clustering
- Video analysis, Motion Detection and Estimation
- Object Recognition: Dimensionality Reduction with Principal Component Analysis

⇨ **Assignments with Coding in Python**

References

Numerous books exist on computer vision. Here are a few recommendations (You can find them online or in the library):

Concise Computer Vision: An Introduction into Theory and Algorithms, Springer. R. Klette

Digital Image Processing, R.C. Gonzalez, R.E. Woods, Pearson Prentice Hall 2008

Image Processing, Analysis, and Machine Vision , Hlavac et al.

Interactive Teaching Platform TopHat will be used in class in our course

- **Prerequisites:**

-- You need to register for an account at TopHat. Your TA Enes Albay will send a message about this in the second week.

Contact him with any questions

--http://tophat.com/

What we expect from you for this course Spend TOTAL of 7 hours each week

- **2 hours:** Attend Lectures, Respond to Online Quizzes and Attendance Quiz during the lectures
- **1 hour:** Study the related material afterwards
- **4 hours:** Work on your homework assignments: both algorithm development and Python programming

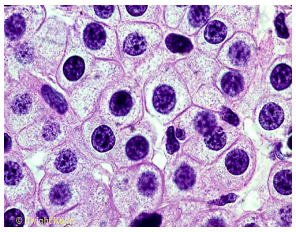
Machine Vision: Self-Driving Cars

https://www.youtube.com/watch?v=mCj_C1NOVxw

Tesla Video min: 0:54

Google's self-driving cars
https://www.google.com/selfdrivingcar/how/

Machine Vision in Medical Diagnosis



e.g. Cell Counting

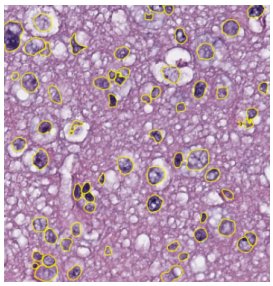


Image guided surgery https://www.youtube.com/watch?v=tJTR4ty0BW4

Computer Vision and Artificial Vision : an important component within AI

ARTIFICIAL INTELLIGENCE:

the science of making computers do things that require intelligence when done by humans.

Summer Vision Project at MIT (1966) https://dspace.mit.edu/handle/1721.1/6125

The Summer Vision Project

Download

Author: Papert, Seymour A.

Citable URI: http://hdl.handle.net/1721.1/6125
Date Issued: 1966-07-01
Abstract:
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which allow individuals to work independently and yet participate in the construction of a system complex enough to be real landmark in the development of "pattern recognition". The basic structure is fixed for the first phase of work extending to some point in July. Everyone is invited to contribute to the discussion of the second phase. Sussman is coordinator of "Vision Project" meetings and should be consulted by anyone who wishes to participate. The primary goal of the project is to construct a system of programs which will divide a vidisector picture into regions such as likely objects, likely background areas and chaos. We shall call this part of its operation FIGURE-GROUND analysis. It will be impossible to do this without considerable analysis of shape and surface properties, so FIGURE-GROUND analysis is really inseparable in practice from the second goal which is REGION DESCRIPTION. The final goal is OBJECT IDENTIFICATION which will actually name objects by matching them with a vocabulary of known objects.
URI: http://hdl.handle.net/1721.1/6125

Complexity of Computer Vision: Various

Where are these objects in this complicated scene: bird, helmet, lamp, plate, map ….

https://www.clarifai.com/demo

Watch Prof. Fei Fei Li talk: Girl vs Computer: 0:18 – 0:32 vs 14:29 -15:08
https://www.youtube.com/watch?v=40riCqvRoMs&list=PLzjJ7xfFxm6nTKTfB1xwE70efksbY6jbC

AI, Machine Vision and Machine Learning

- Machine learning is the core technology of artificial intelligence (AI) that provides computers with the ability to learn by programming them with inputting a lot of data.
- Machine learning focuses on the development of computer programs that can teach themselves to grow and change when exposed to new data.
- Supervised
- Unsupervised
- Reinforcement learning

Recently: Machine Learning has become a crucial component of Computer (Machine Vision) algorithms

Learning Outcomes of the Course

Students will be able to:

1. Discuss the main problems of computer (artificial) vision, its uses and applications
2. Design and implement various image transforms: point-wise transforms, neighborhood operation-based spatial filters, and geometric transforms over images
3. Define and construct segmentation, feature extraction, and visual motion estimation algorithms to extract relevant information from images
4. Construct least squares solutions to problems in computer vision
5. Describe the idea behind dimensionality reduction and how it is used in data processing
6. Apply object recognition approaches to problems in computer vision

Week 1 and Week 2 Topics and Learning Objectives:
Introduction to Computer Vision
Pointwise Image Processing
Image Intensity Transformations, Image Histograms, Image Enhancement

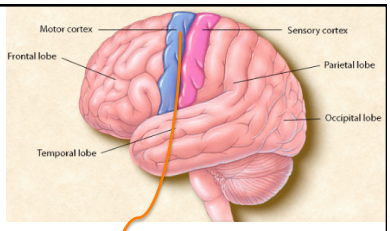
At the end of Week 1&2: Students will be able to:

1. Discuss the main problems of computer (artificial) vision, its uses and applications
2. Design and implement various image transforms: point-wise transforms, neighborhood operation-based spatial filters, and geometric transforms over images

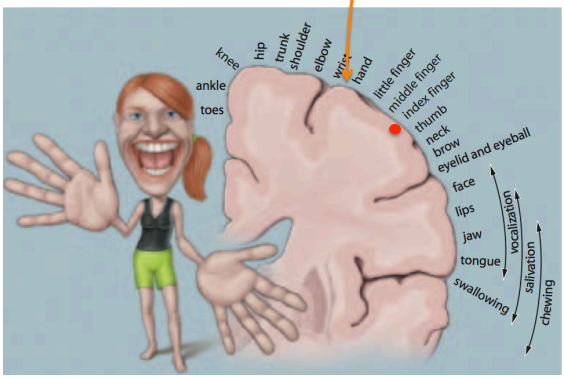
Visual Circuits in our brain? Mapping Brain Circuits

The details of our body map were studied in detail by a Canadian neurosurgeon Wilder Penfield in 1930-40s (Nobel prize recipient)

Below is a drawing of the body part and the corresponding location in the brain cortex: e.g. if you stimulate the red point, the patient feels a sensation in the index finger



Every point on your body surface has a corresponding point on this map

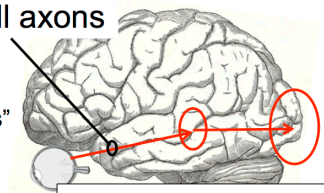


e.g. professional guitar players or violinists ?

Vision: Image Formation

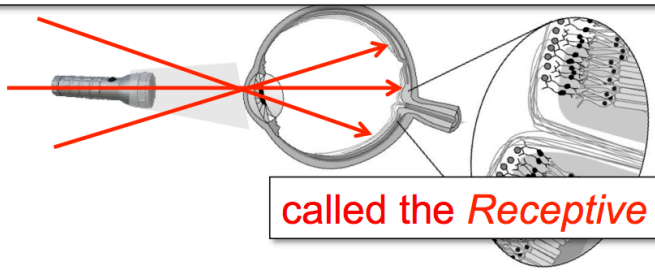
Retinal ganglion cell axons

Microscopic in diameter
Travel many centimeters
Same "nearest-neighbors"



Preserves the map!

Each photoreceptor only "sees" one place



called the *Receptive Field*

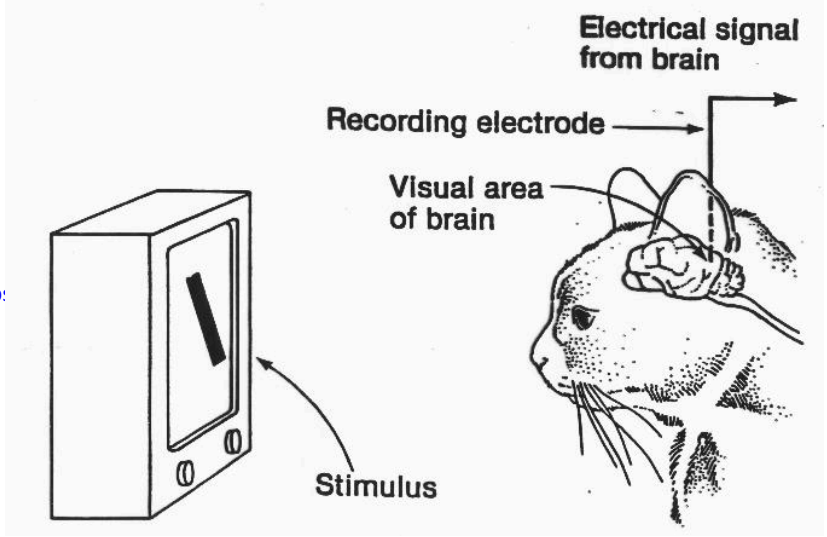
The receptive field of a photoreceptor is a region of the visual field to which the photoreceptor is sensitive.

Brain Visual Maps

Hubel and Wiesel (1959): The Nobel Prize in Physiology or Medicine 1981 *for their discoveries concerning information processing in the visual system"*

Series of visual stimuli presented to a cat, responses are recorded from a neuron in visual cortex

http



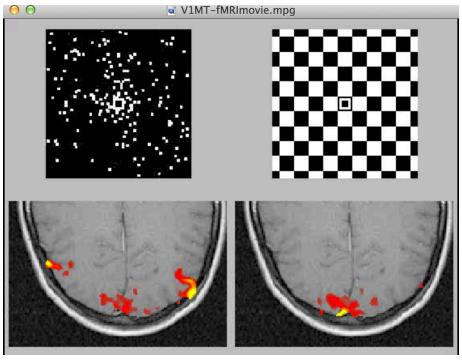
Mapping the Brain: Imaging Studies

About half of the entire cerebral cortex in primates is devoted to processing of visual information!

Vision is hiearchical

Magnetic Resonance Imaging (MRI)

Functional MRI



SENSE OF VISION

* One of the most important senses in life

- Boom in visual sensors (e.g. smart phones with camera…)
- By 2019, 80% of the World's Internet Traffic Will Be Video

(CISCO Forecast[1] 2020)

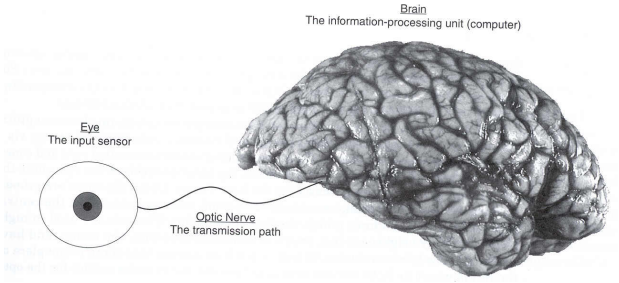
What is t*he aim of* Computer Vision?

To make the computers see!

[1] http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

Computer Vision

- Visual System: a collection of devices that transform measurements of light into information about spatial and material properties of a scene
- Among the devices:
  - Photosensitive sensors (camera or retina)
  - Computational mechanisms (computer or brain) to extract information from raw sensory readings



Computer Vision

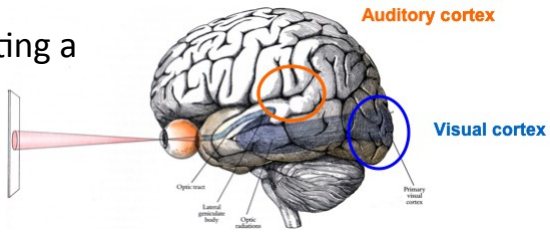
- Is it hard for a computer to see?

225	217	213	214	211	220	229	227	226	230	230	226	227	229	230	230	224	229	234	231	228	238	232	233
241	234	229	234	229	230	229	227	223	229	225	225	226	232	229	227	220	220	224	229	224	226	226	228
233	226	236	236	231	229	226	226	227	222	220	220	215	221	229	217	228	227	221	220	214	221	219	224
220	217	222	227	231	222	220	223	223	224	225	218	208	216	218	221	223	227	222	215	219	214	217	213
211	213	213	224	222	216	218	216	214	219	222	225	208	209	218	219	224	228	227	222	223	218	208	209
212	213	211	215	219	213	213	207	207	212	218	222	213	215	215	214	228	226	235	238	224	211	201	203
217	214	216	218	218	223	214	210	204	203	217	213	214	222	216	220	223	224	220	222	220	211	205	202
212	212	217	216	215	220	217	209	215	210	208	210	207	214	224	217	216	214	209	207	216	214	217	222
215	216	210	212	209	211	211	209	214	213	209	206	191	193	203	203	201	196	205	204	209	209	217	219
211	214	209	212	213	218	221	223	221	215	206	192	203	217	222	198	209	219	199	191	194	200	199	198
215	210	203	201	210	216	222	219	211	207	204	200	216	227	237	230	231	225	198	192	192	187	190	199
213	212	206	197	201	202	202	217	209	215	217	222	230	212	206	213	214	203	197	196	196	199	204	197
220	213	203	204	206	203	209	236	240	229	223	237	236	223	211	200	197	192	192	189	189	188	185	188
215	213	202	205	209	198	208	214	228	228	217	207	209	204	209	203	192	194	188	182	186	189	179	183
196	204	201	200	196	198	214	206	203	208	202	197	194	190	190	187	192	193	195	189	189	191	191	191
212	206	204	211	217	211	201	198	198	191	188	187	188	190	195	197	195	199	206	216	202	192	187	189
224	219	222	208	204	204	200	192	187	192	194	185	176	183	188	192	195	201	211	218	204	198	196	196
211	215	216	211	206	210	207	203	201	212	216	200	198	194	186	181	184	193	203	198	192	195	202	202
202	202	213	216	203	202	197	194	198	205	199	199	195	195	198	185	185	191	191	182	177	184	192	196
195	201	201	198	194	195	193	196	194	196	189	191	192	187	194	187	189	197	196	181	173	179	193	207
211	199	201	197	197	192	197	199	190	194	198	199	206	200	191	196	191	193	198	194	187	185	191	206
212	197	207	201	202	206	205	203	207	207	197	203	207	202	196	196	196	191	196	198	202	198	186	192
186	194	193	193	201	213	204	205	214	209	200	203	201	200	195	206	215	211	210	208	208	202	190	187
164	170	175	190	194	198	197	199	202	205	204	209	207	196	203	225	235	232	229	226	216	202	191	185

* Each image frame is just a collection of positive numbers that measure the amount of light incident on a particular location (or pixel) on a photosensitive surface

Computer Vision

Is it as simple as connecting a camera to a computer ?



- ❑ Camera → Computer
- ❑ Retina → Brain

- Nowadays, a digital camera can deliver many many “frames” per second to a computer

Computer Vision

- Take the same exact scene:
  - change the viewpoint
  - Change the illumination (sun, candle light, neon lamp etc or its angle)

We have no difficulty in interpreting the scene having the same objects!

Frog pictures from Osadchy-Keren CVIU-04

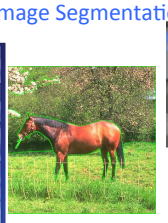
How can we interpret these pixel values?

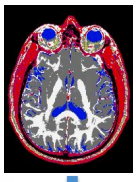
to tell whether we are looking at an apple or an eye or a face or a tree?

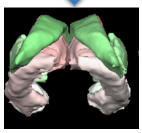
Computer Vision: numerous areas

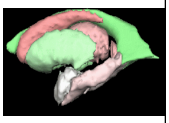
Image / Object Recognition

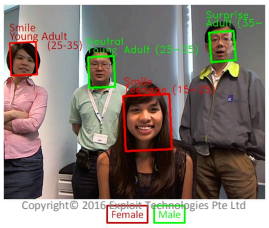
Image Segmentation



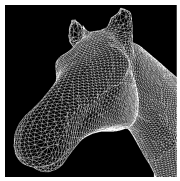


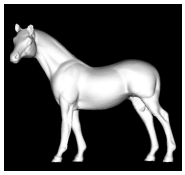






Constructing 3D geometry/
models from multiple2D images





Computer / Machine Vision in Medical Diagnosis Intervention and Monitoring

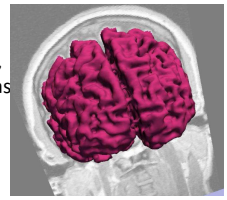


Image Analysis

Segmentation, Registration, Diagnostic, Prognostic diseas specific analysis, markers, measures, ...

Image Guided Intervention and Surgery

3D Slicer: MGH

http://www.highpointregional.com/sites/www/Uploads/images/CancerCenter/interradiology_004.jpg

Industrial Computer Vision

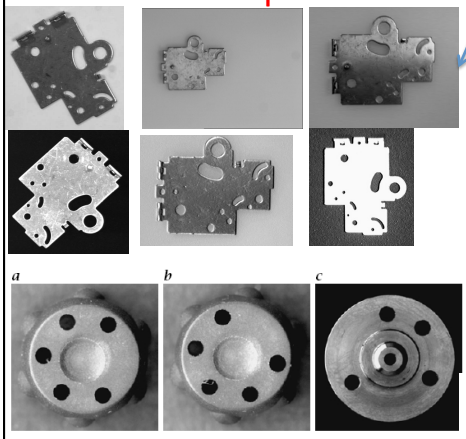


Figure 1.2: Industrial parts that are checked by a visual inspection system fo the correct position and diameter of holes (courtesy of Martin von Brocke, Rober Bosch GmbH).

Circuit board inspection

Pose estimation of parts for industrial robots

Have to be Robust to: Position; Orientation; Lighting

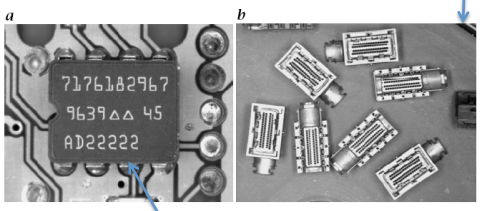


Figure 1.10: Industrial inspection tasks: **a** Optical character recognition. **b** Connectors (courtesy of Martin von Brocke, Robert Bosch GmbH).

Classification: label recognition on a chip

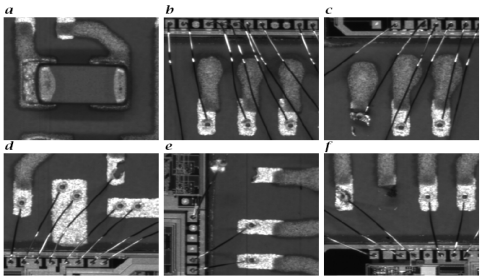
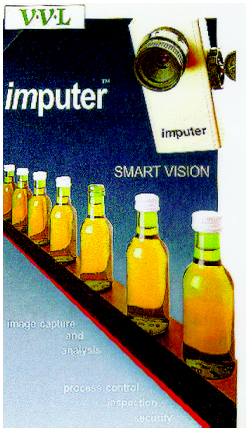
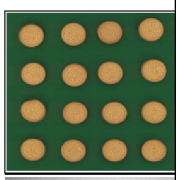


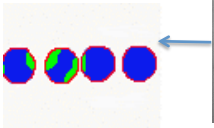
Figure 1.11: Errors in soldering and bonding of integrated circuits. Courtesy of Florian Raisch, Robert Bosch GmbH).

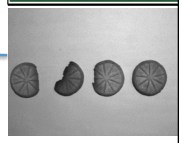
Food Industry

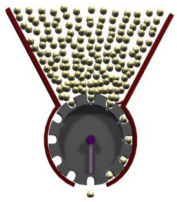


- fruit and vegetable sorting and grading
- automatic portioning,
- inspection for foreign objects,
- verify the size and shape of contents
- quality inspection of containers,
- fill level inspection,
- closure inspection

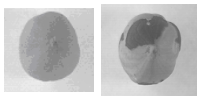


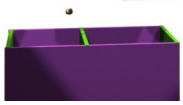






Automatic Nut Inspection system

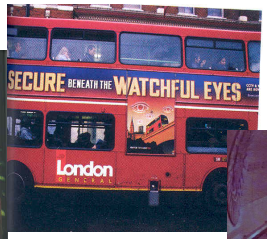


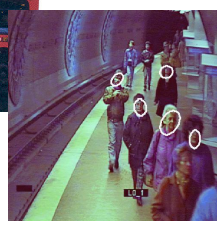


Apricot Quality Inspection

Slide: Prof. Aytul Erçil

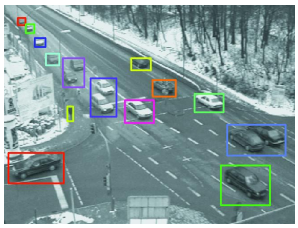
Surveillance Applications

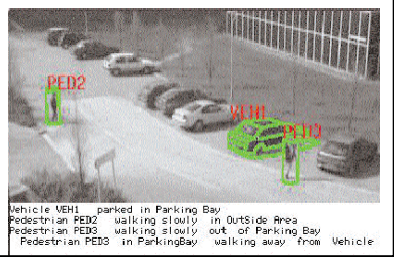




London Residents: the most watched people on Earth
About 5 million cameras placed on corners around the city
People are caught on camera roughly 300 times a day!

Traffic control





Robot Guidance, Navigation

Unmanned Air Vehicles (UAVs)

Honda, Asimo Robot

Satellite Imaging: Remote Sensing

e.g. Land Classification

Recognize:
- Forests
- Buildings
- Agricultural land
- Roads
- Industrial regions
- etc

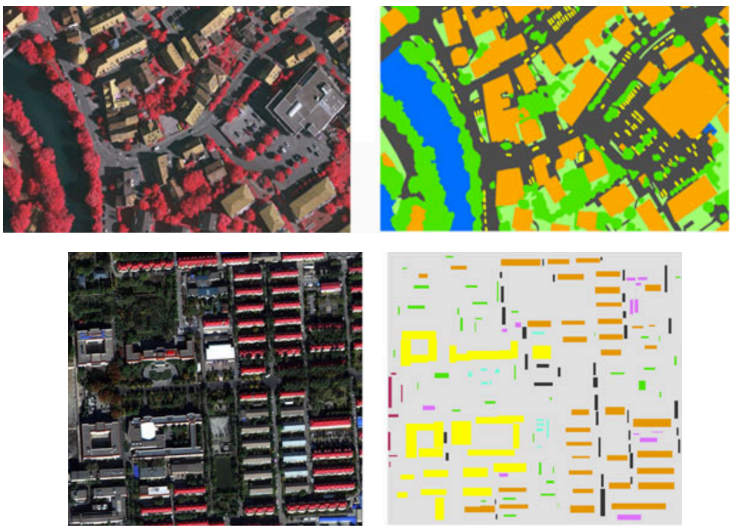


Figure: Zhao et al, "Object-Based Convolutional Neural Network for High-Resolution Imagery Classification", IEEE Journal Selected Topics in Applied Earth Observations and Remote Sensing

Space Imaging

Images from the Hubble Telescope
Ability to image distant galaxies:
Task: Separate galaxies into different classes (shape, color, ...)
Distinguish stars, planets etc from other objects

Figure 1.12: *Hubble deep space image: classification of distant galaxies (http://hubblesite.org/).*

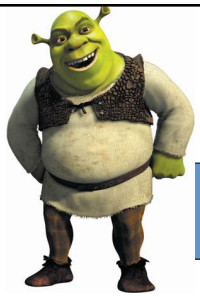


Image Processing

Image

Computer Vision: Inverse problem of image formation; uses images to recover a description of objects in space

Computer Vision, Image Analysis, Pattern Recognition

Ill-posed: we lose 1 dimension going from 3D world to 2D images

Image Formation: how objects give rise to images

Computer Graphics

Object, Scene Model and Description

Shrek:
- has a green face
- wears a vest, a belt,...
- has a huge belly
- is smiling
- is happy,.....etc

Designing vision algorithms: requires appropriate modeling of relevant information in images
Appropriate models: manageable; trade-off between physical complexity -constraints and mathematical simplicity while keeping fidelity to the existing visual information

Wrap-up Introduction to Computer Vision

- We understand the fact that Computer Vision is NOT just transferring data from a camera to a computer
- The field together with machine learning is currently at the intelligence level of a 3 year old human !
- There are a HUGE range of applications that require intelligent processing of visual data

In this course: we will start with the basics: Need to learn LOW-LEVEL PROCESSING first in order to go to image understanding / higher level processing and cognition!

- Next: Pointwise Image Operations
  - Image Intensity Transformations
- Image Coordinate Transformations
- Neighborhood Image Operations (filtering etc)
- Extracting Edges, Corners, other features in images
- Segmentation
- Continue with other more advanced topics