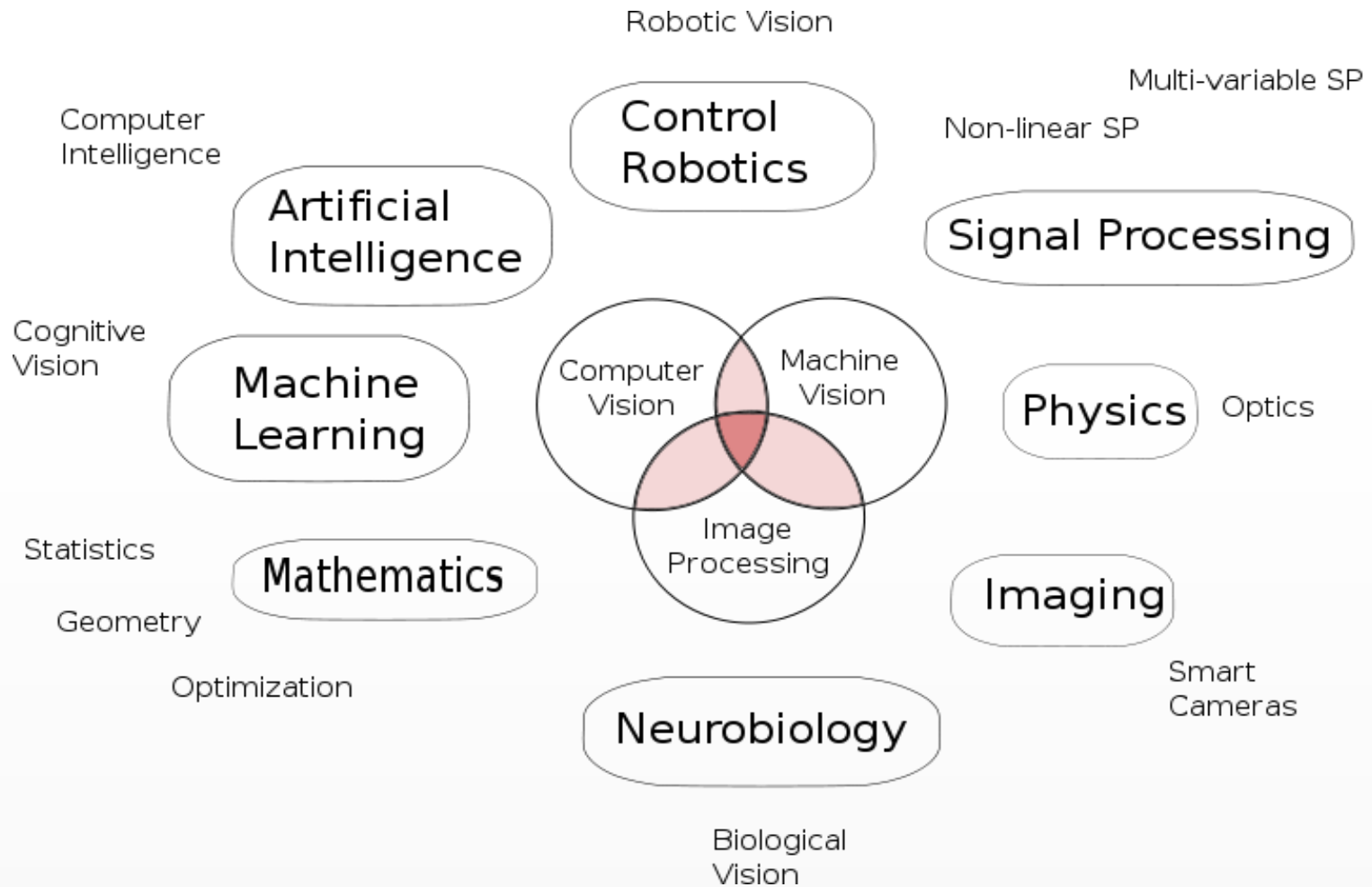


Basic Image Processing Algorithms

PPKE-ITK, 2016

What is this Course About?



Source: Wikipedia

What is this Course About?

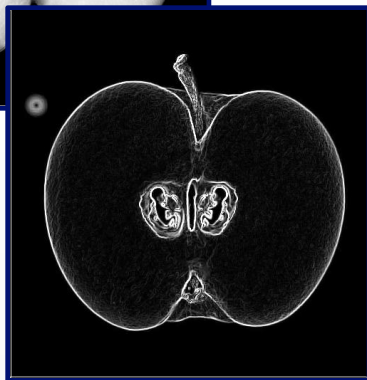
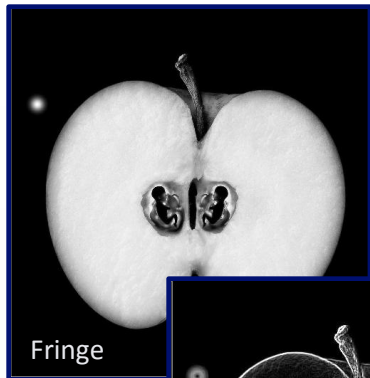
- Machine Vision is the technology and methods used to provide imaging-based automatic inspection and analysis for such applications as automatic inspection, process control, and robot guidance in industry.”



Sources: Wikipedia, <http://automation.com>, <http://www.isquaredt.com/>

What is this Course About?

- Image Processing "is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image."

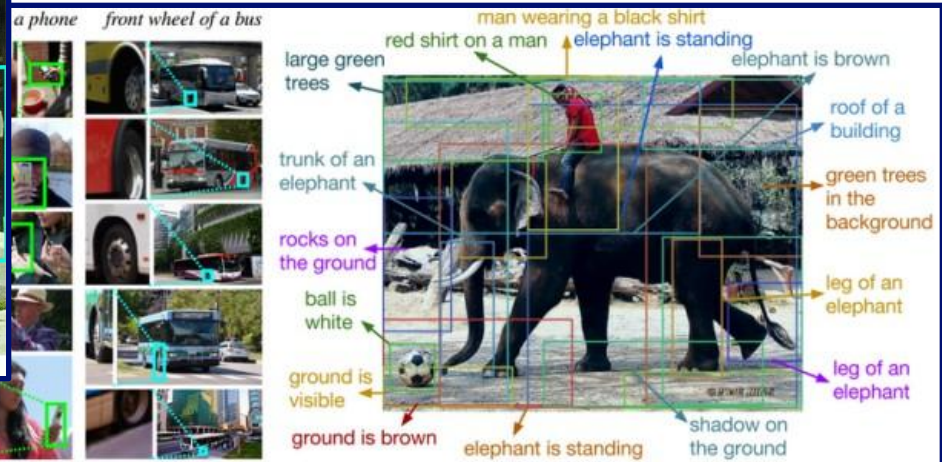
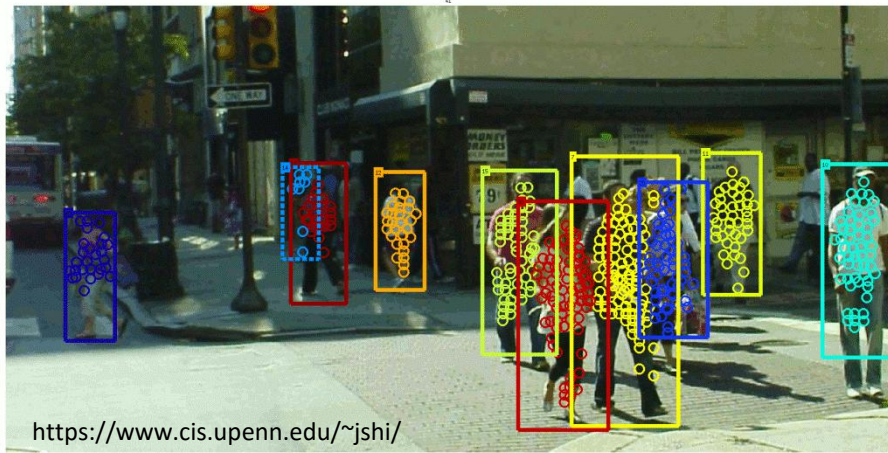


Source: Wikipedia



What is this Course About?

- „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”
„A theme in the development of this field has been to duplicate the abilities of human vision by electronically perceiving and understanding an image.”



<http://cs.stanford.edu/people/karpathy/>

Source: Wikipedia

Video Traffic on the Internet

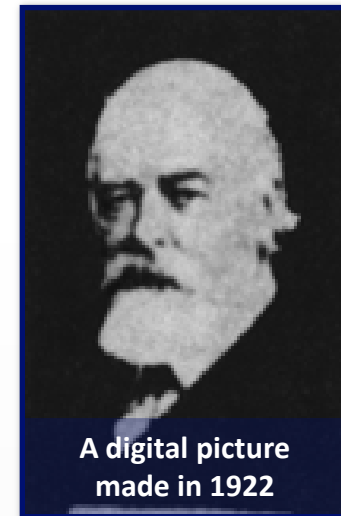
- ◎ A few crazy predictions from Cisco for 2020:
 - video traffic will be 82% of all IP traffic (both business and consumer), up from 70% in 2015,
 - it would take more than 5 million years to watch the amount of video that will cross global IP networks each month,
 - every second, a million minutes of video content will cross the network.



<http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html>

Little bit of History

- ◎ Early 1920s - Bartlane cable picture transmission system
 - Used to transmit newspaper images via submarine cable between London and New York.
 - Took about three hours to send an image, first systems supported 5 gray levels



- But these images were not created with computer, hence not considered as a result of digital image processing.
- The real era of digital images started only after computers got powerful enough for the task.

Little bit of History

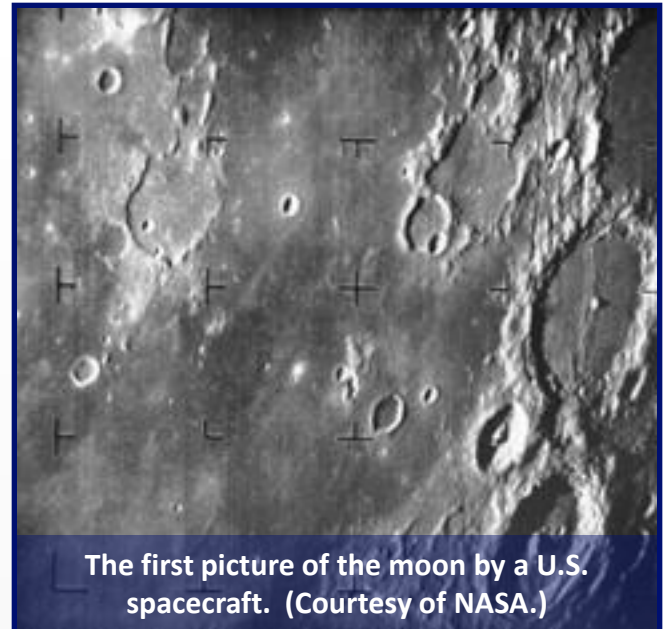
◎ In the early 1960s:

- 1964: NASA's Jet Propulsion Laboratory began working on computer algorithms to improve images of the Moon.
- images were transmitted by Ranger 7 probe.
- corrections were desired for distortions inherent in on-board camera

◎ In the late 1960s and early 1970s:

- medical imaging (CT),
- remote Earth resources observations,
- and astronomy

◎ So far this was image processing, what about computer vision?



Little bit of History

- ◉ In 1966, Marvin Minsky (MIT) asked his student to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

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 will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

Little bit of History

Neural Networks:

- ⦿ 1958: Frank Rosenblatt introduced the Perceptron model
- ⦿ 1970's: Backpropagation algorithm for larger network training
- ⦿ 1980's: Appearance of Convolutional NN
- ⦿ 1998: First success of CNN:
["Gradient-based learning applied to document recognition"](#)
- ⦿ Still other methods are favored over NN
- ⦿ 2012: Imagenet classification challenge won by deep CNN*
Their error rate was 15.3%, whereas the second closest was 26.2%
- ⦿ Since 2012 we are witnessing the golden age of CV

*Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105)

Applications

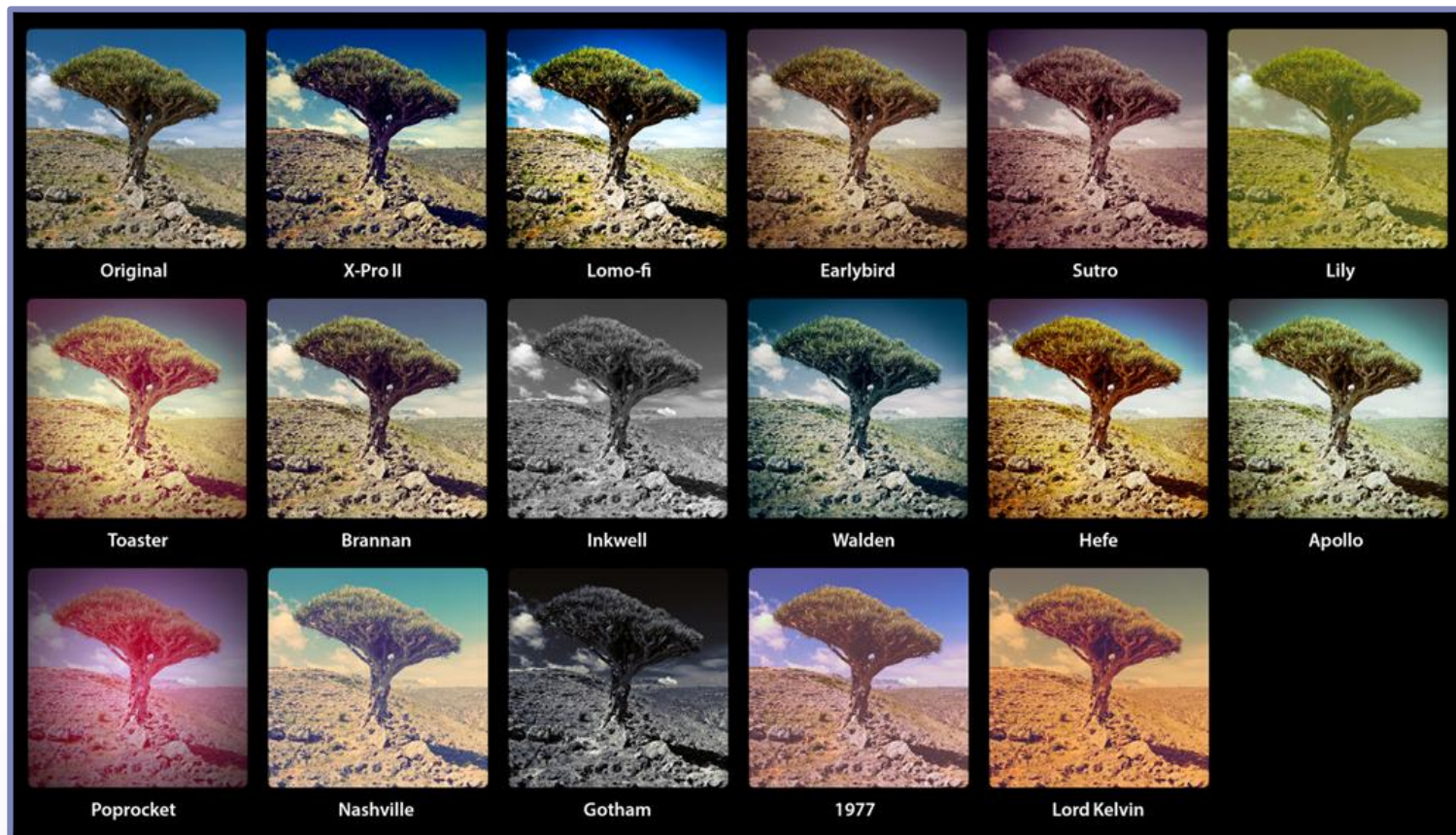
- Early Hubble Space Telescope images were distorted by a flawed mirror and could be sharpened by deconvolution.



Source: <http://opticalengineering.spiedigitallibrary.org/article.aspx?articleid=1077064>

Applications

⦿ Instagram filters



Source: http://www.thephoblographer.com/2013/01/17/instagrams-presets-come-to-lightroom/#.U5SDT_nv-PQ

Applications

◎ Panoramic images

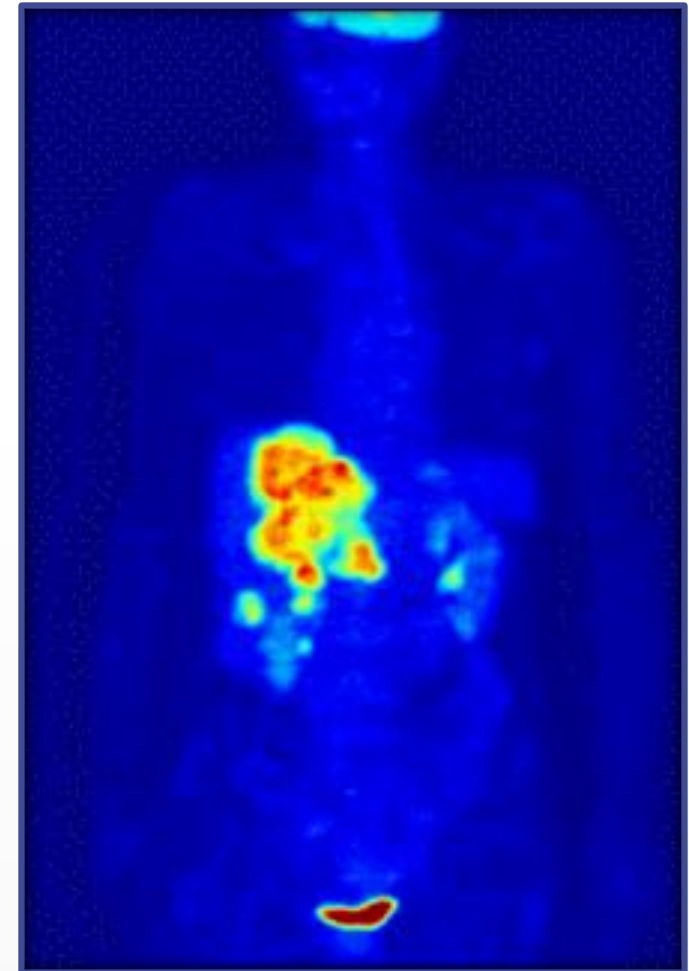
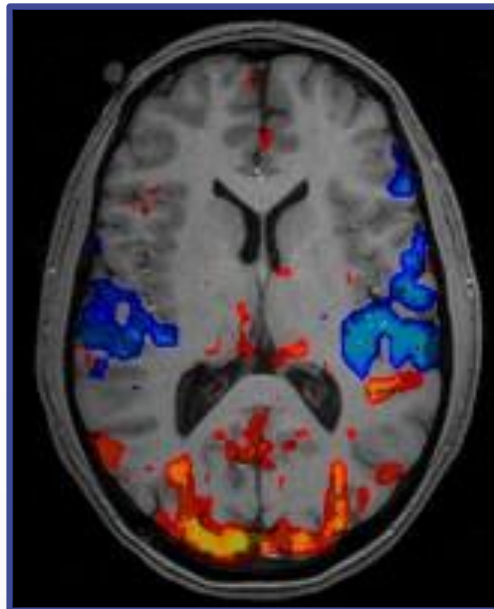


Source: <http://en.wikipedia.org/wiki/User:Diliff>

Applications

◎ Medical image processing:

- Ultrasound
- 3D imaging, MRI, CT, PET
- Image guided surgery
- ...



Source: http://en.wikipedia.org/wiki/Positron_emission_tomography

Applications

- Face detection is a standard feature in smart phones/cameras:



Source: <http://www.imore.com/>

Applications

◎ Facial retargeting:

Entertainment



◎ Emotion recognition:



Applications

◎ Advanced driver assistance systems (ADAS)

- Automatic parking
- Collision avoidance system
- Driver Monitoring System
- Emergency driver assistant
- Lane departure warning system
- Lane change assistance
- Pedestrian protection system
- Traffic sign recognition
- ...



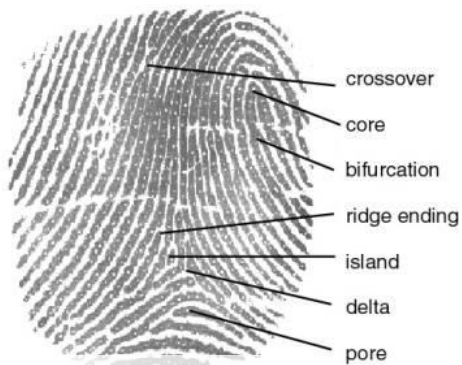
„So last week Uber announced several sensor laden, self-driving Volvo XC 90 SUVs would hit the crowded streets of Pittsburgh. That is five years earlier than a similar plan from Ford (F) and at least three years ahead of Tesla’s (TSLA) master plan. Most important, it uproots the plans of Alphabet (GOOGL), the company Uber fears most.”

2016.08.26.

<http://www.forbes.com/>

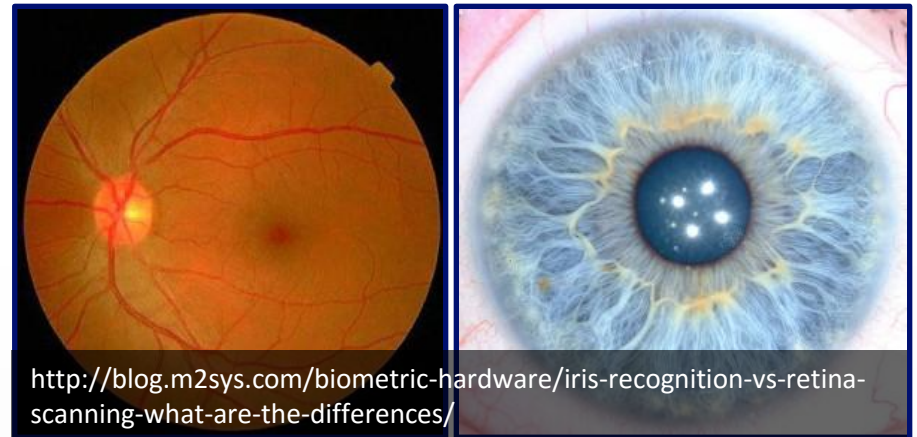
Applications

Identity verification and recognition based on biometrics:

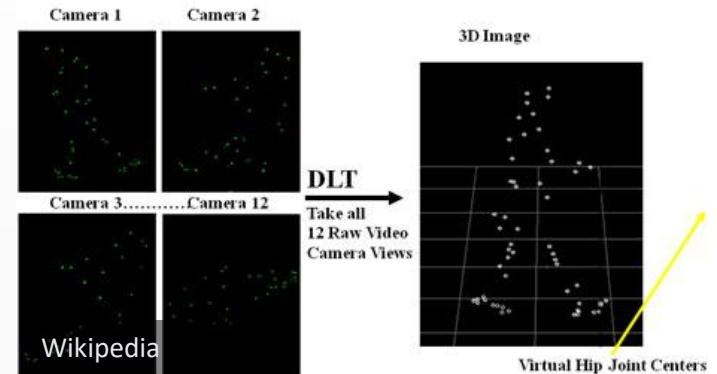
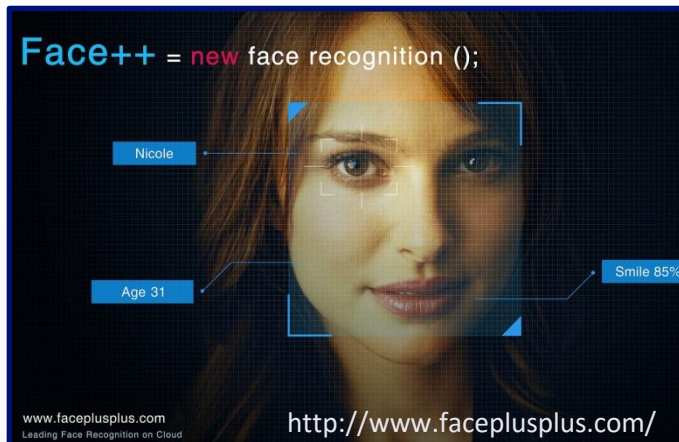


<http://www.biometrics.gov/>

Figure 1

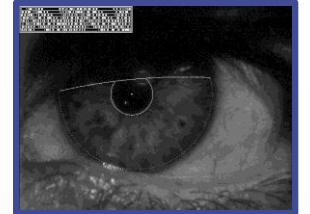
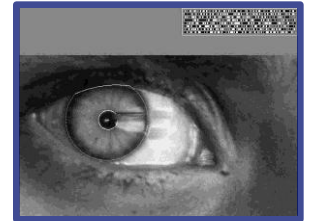


<http://blog.m2sys.com/biometric-hardware/iris-recognition-vs-retina-scanning-what-are-the-differences/>



Applications

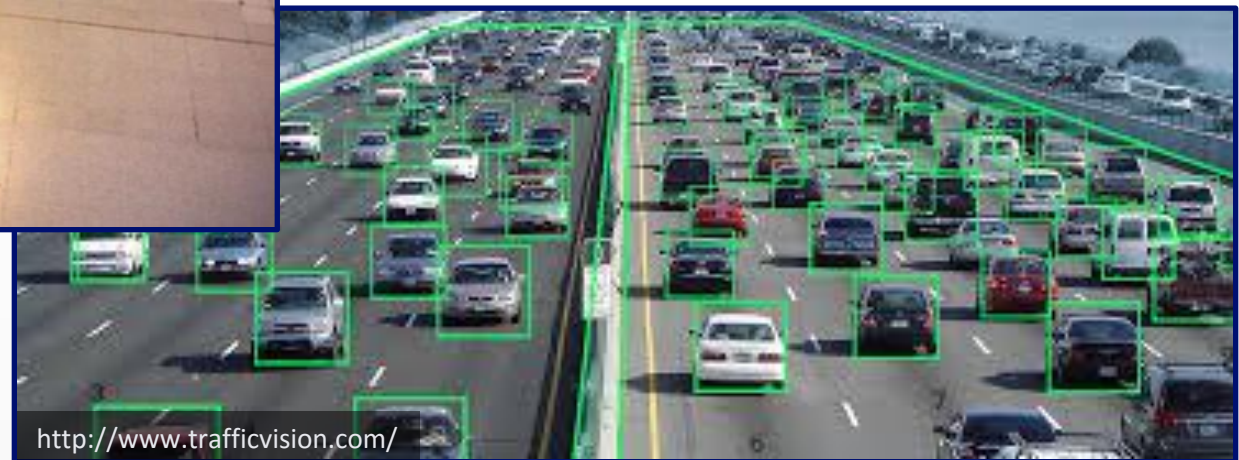
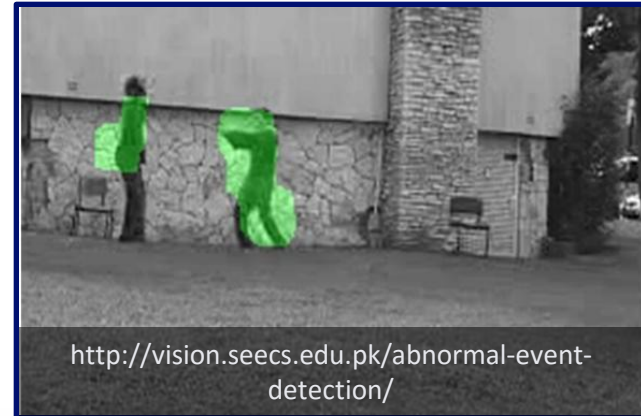
- ◎ Story of the Afghan girl:



Source: <http://www.cl.cam.ac.uk/~jgd1000/afghan.html>

Applications

◎ Video surveillance:



And There are Many More...

Administration

Administration and Course Requirements

◎ Web:

- <http://kep.itk.ppke.hu>

◎ Mailing list:

- <https://lists.ppke.hu/cgi-bin/mailman/listinfo/kepelemzes>
- You will be subscribed automatically

◎ Contacts:

- Daniel Szolgay: szolgay.daniel@itk.ppke.hu
- Miklós Koller: koller.miklos@itk.ppke.hu
- Márton Naszlady : naszlady.marton.bese@hallgato.ppke.hu
- Ágnes Szabó: szabo.agnes@hallgato.ppke.hu
- Csaba Farkas: farkas.csaba@hallgato.ppke.hu

Administration and Course Requirements

◎ **Seminars:**

- Start time?
- The attendance is obligatory
- In every seminar there will be a short test (with 1-2 questions):
 - > 60% in total is a requirement to take the final exam
 - > 90% in total is a necessary condition to get offered grade
- No midterm or final test.

◎ **Lab practice:**

- The attendance is obligatory
- There will be programming tasks (the default language is Matlab) that you have to complete and submit via e-mail by the deadline.
- You have to understand your code and be able to explain it.
- Late submission policy: will be clarified by Miklós.
- Submitting all assignments in time is a requirement of passing the course.

Administration and Course Requirements

◎ **Assignments:**

- During the semester we will hand out 3-4 longer programming assignments.
- You have to complete and submit your solution via e-mail by the deadline.

◎ **Oral exam** at the end of the semester:

- To be able to participate to the exam you have to complete all programming practice and assignment
- Have to complete the small tests at least at the required minimum level.

◎ **Offered Grade:**

- > 90% in total from the tests
 - All assignments, programming tasks are accepted
 - 1 scientific paper processed (details will be given later)
- > You get a 5 for the final mark

◎ ?

Human Vision

◎ The **human visual system**:

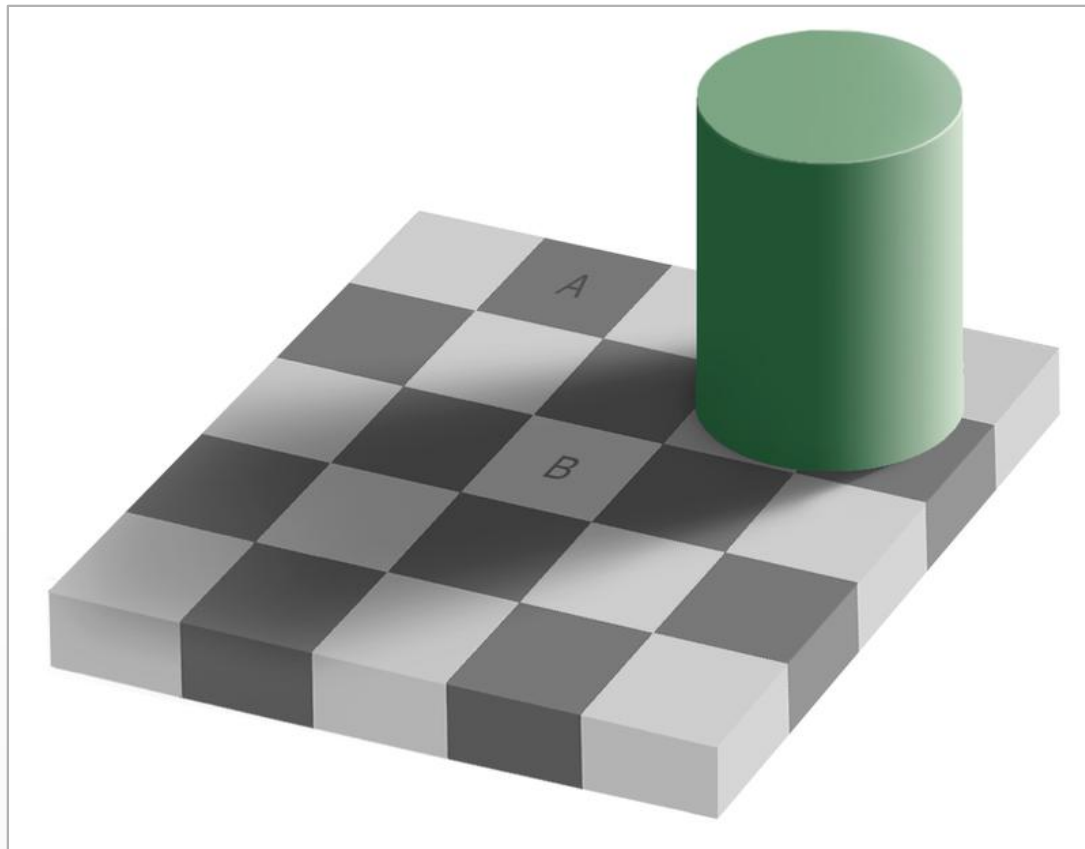
- Gives us the ability to process visual stimulus, to be able to detect and interpret information from visible light (build a representation of the surrounding environment).
- ◎ The ultimate goal of computer vision is to build a system that is capable of seeing as a human can. (or even better)
 - ◎ It is not easy! Human vision was trained through the many years of the evolution. It can perform complex tasks (face/facial expression recognition) easily.
 - ◎ For a computer it is still an unsolved problem: there is a gap between how a human and how a computer sees an image.
 - ◎ Yet human vision is fallible. Illusions and ambiguities are encountered all the time.

Illusions

- ⦿ The visual system is optimized to process natural images (through evolution)
- ⦿ It is faced with an ill-posed problem:
 - Ambiguity due to projection from 3D to 2D image
 - Uncertainty due to incomplete knowledge of the environment
 - Uncertainty due to noise in photoreceptors and neurons
- ⦿ The visual system relies on a set of assumptions to solve this ill-posed problem
 - Assumptions presumably learned via evolution
 - Assumptions tailored for the natural visual world
 - Assumptions cause illusions/failures under impoverished conditions
- ⦿ Illusions can provide insights into the brain's assumptions.

Illusions

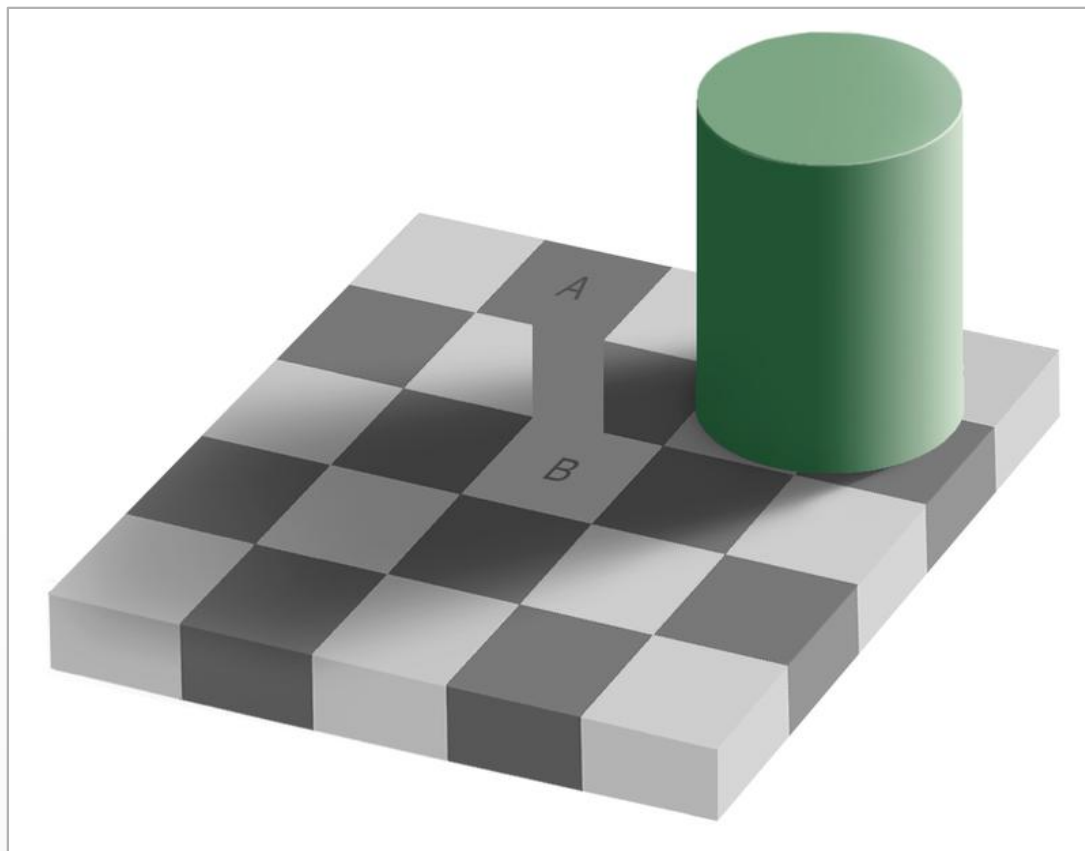
- ⦿ Lateral inhibition + assumptions tailored for the natural visual world



http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

Illusions

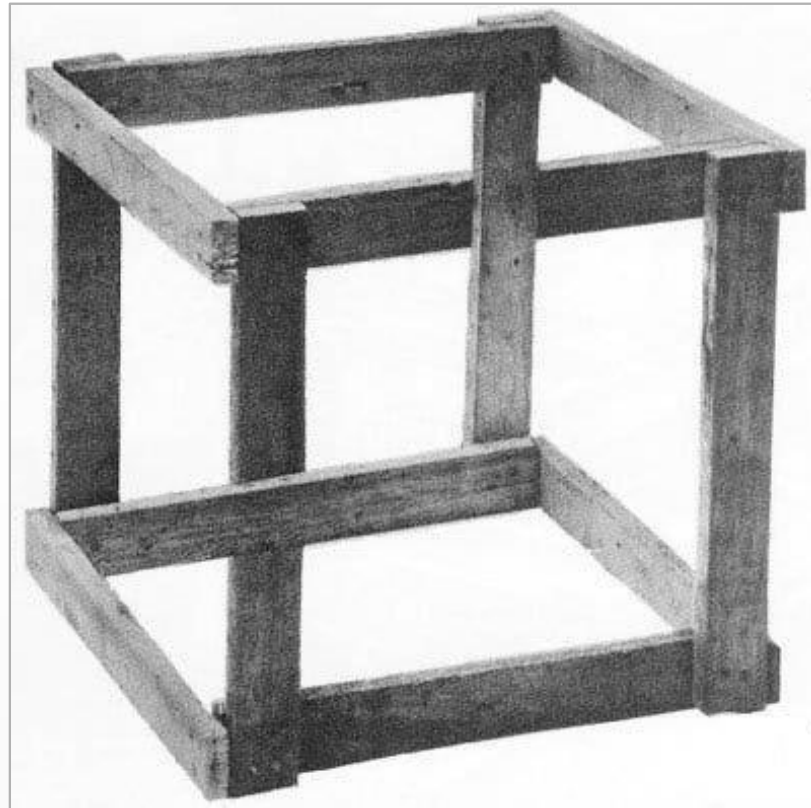
- ⦿ Lateral inhibition + assumptions tailored for the natural visual world



http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

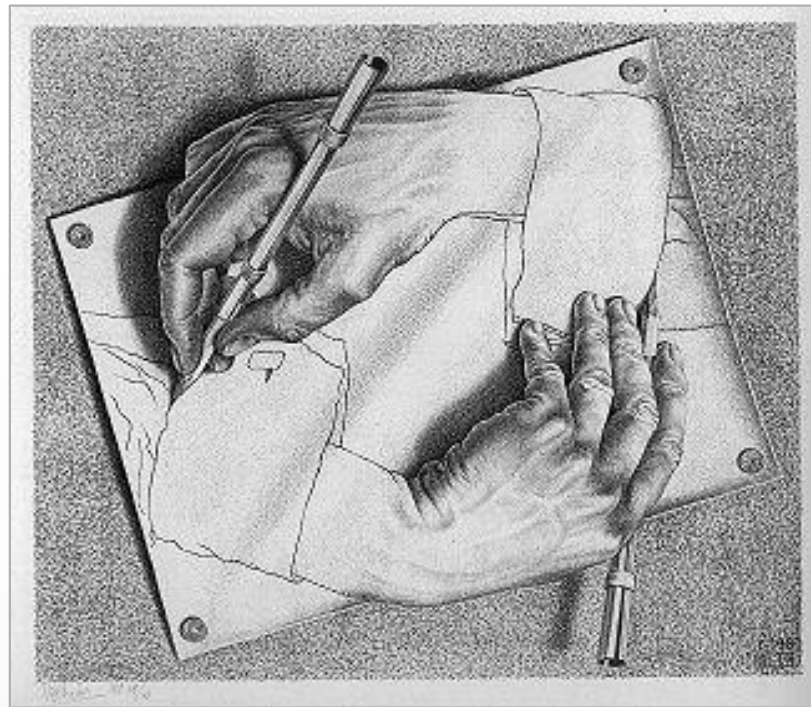
Illusions

- ◉ Ambiguity due to projection from 3D to 2D image



Illusions

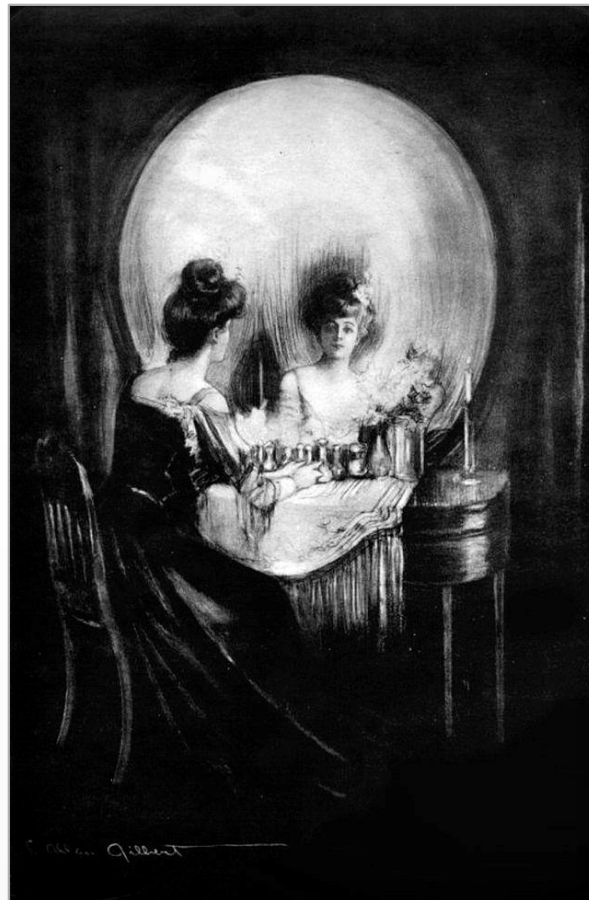
- ◉ Ambiguity due to projection from 3D to 2D image



Maurits Cornelis Escher
Drawing Hands (1948)

Illusions

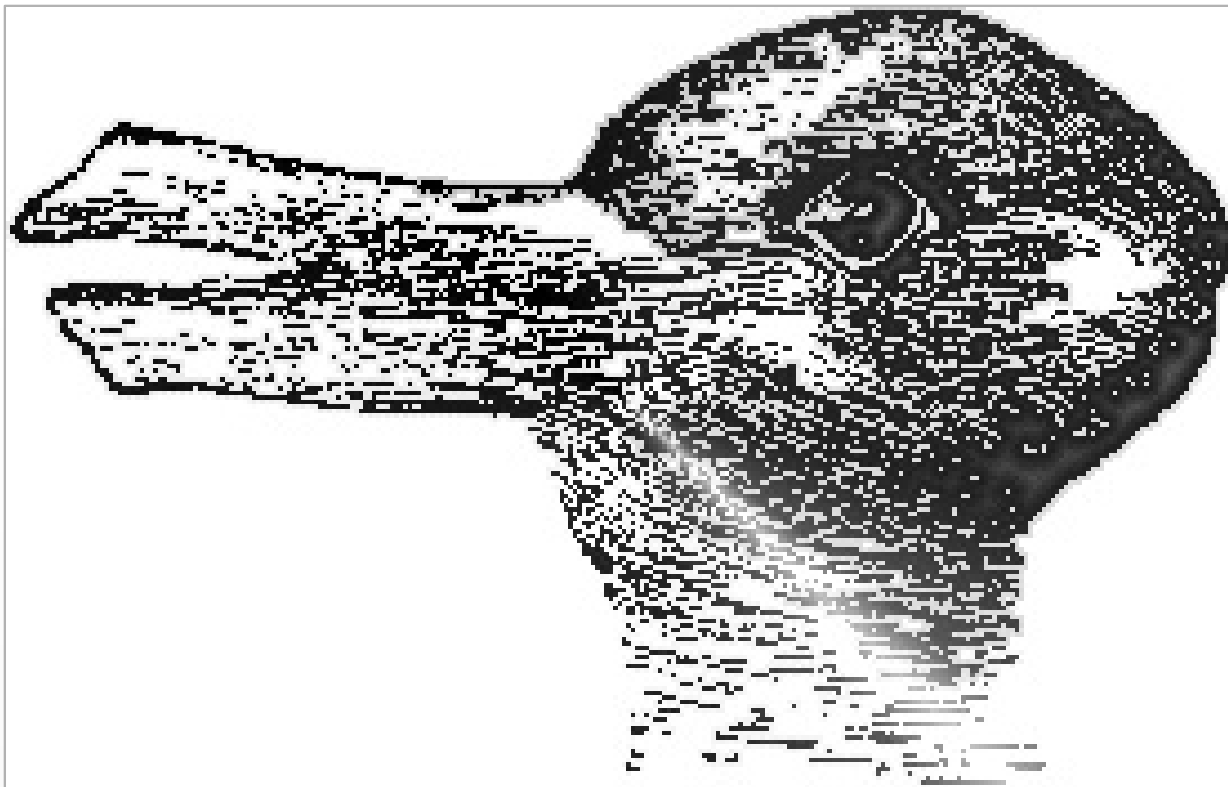
- Uncertainty due to incomplete knowledge of the environment



Charles Allan Gilbert
All Is Vanity (1892)

Illusions

- Uncertainty due to incomplete knowledge of the environment



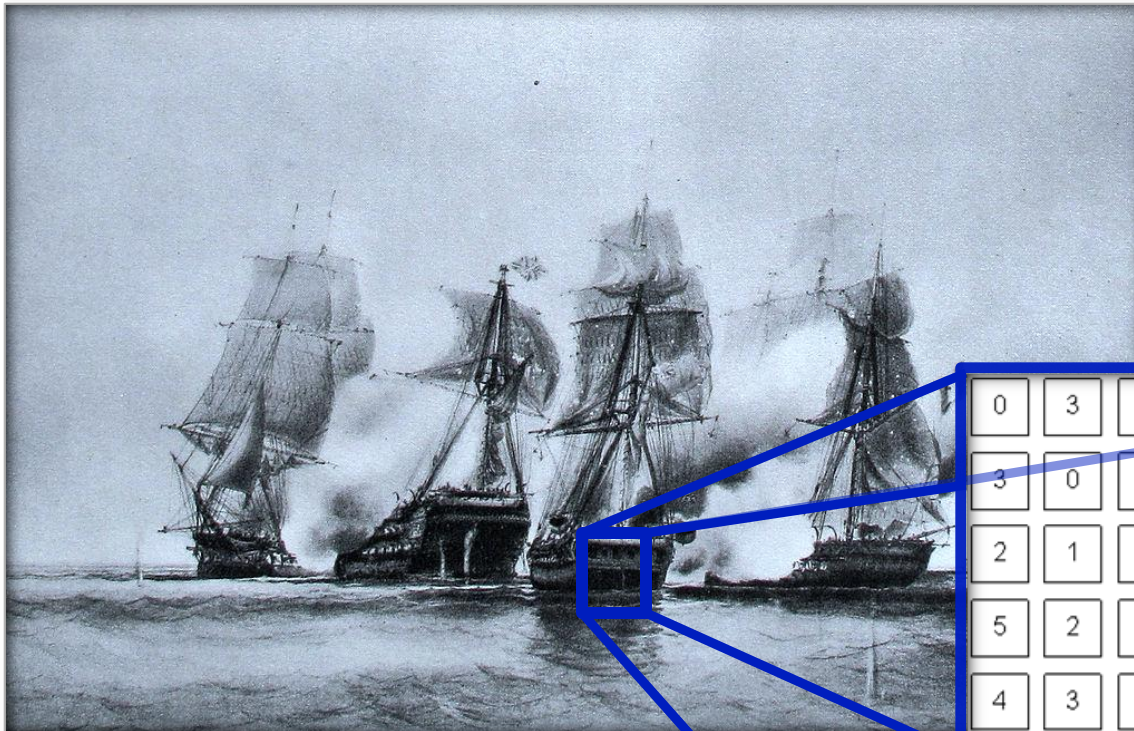
Ludwig Wittgenstein
Rabbit and Duck (1892)

Illusions

- ⦿ Uncertainty due to missing information:



What does a computer „see”?

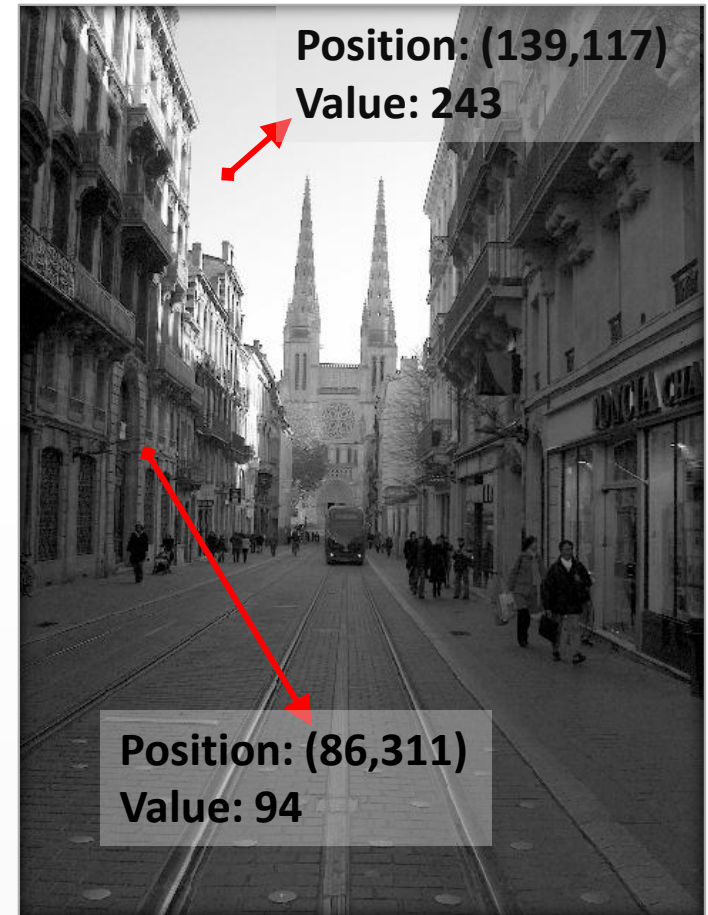


http://en.wikipedia.org/wiki/Roebuck-class_ship

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

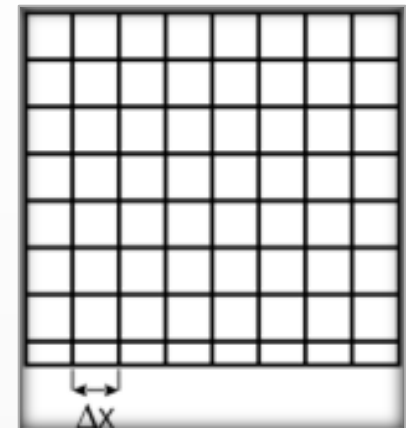
What does a computer „see”?

- ⊙ A **digital image** is discrete representation of a continuous measurement, usually a **2 or 3 dimensional array**.
- ⊙ An element of this array is a **pixel** (picture element).
- ⊙ A pixel has a **position** (its coordinates on the image) and an **intensity value**.
- ⊙ A digital image is discretised both in space and intensity:
 - Spatial discretisation is referred to as **sampling**.
 - Intensity discretisation is referred to as **quantization**.



Sampling

- ⦿ **Sampling** is the reduction of a continuous signal to a discrete signal.
- ⦿ A finite set of values (called **samples**) are selected to represent the original continuous signal.
- ⦿ In case of 2D signals (images) a grid is used for sampling
- ⦿ The grid points will be represented as pixels.
- ⦿ The frequency of the sampling defines:
 - ⦿ How many grid points we have?
 - ⦿ What is the resolution of the image?
 - ⦿ How detailed the discretised image is?



Sampling

- ⦿ Sampling usually leads to information loss.
- ⦿ The sampling frequency determines how much information we lose.
- ⦿ We have to decide what is the smallest detail that we want to keep:

24x32



48x64



120x160



480x640



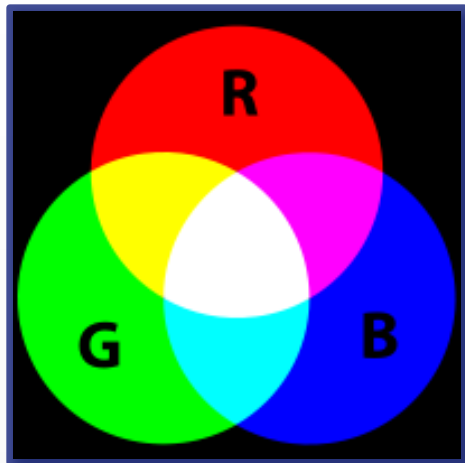
Quantisation

- ◎ Intensity discretisation is referred to as **quantization**.
- ◎ The digital image quality is highly depending on how many bits we use for coding the discrete intensity values:
 - Binary: each pixel is coded on 1 bit (zero or one, black or white)
 - Gray scale coded on 2/4/8/16/24/32 bits

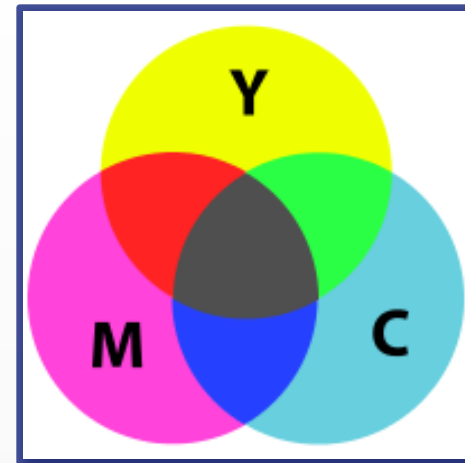


Color Images

- ⊙ Humans can distinguish thousands of color shades and intensities, but only a few dozens of gray.
- ⊙ Color can be a useful descriptor for image segmentation, tracking, detection,...
- ⊙ Two main color mixing models:
 - Additive model:



Subtractive model:



Color Images

◎ Color Characteristics:

- **Brightness:** used to describe color sensation (it is similar to intensity of achromatic light)
- **Hue:** it indicates the dominant wavelength in the mixture of light waves
- **Saturation:** relative purity or the amount of white light in the mixture

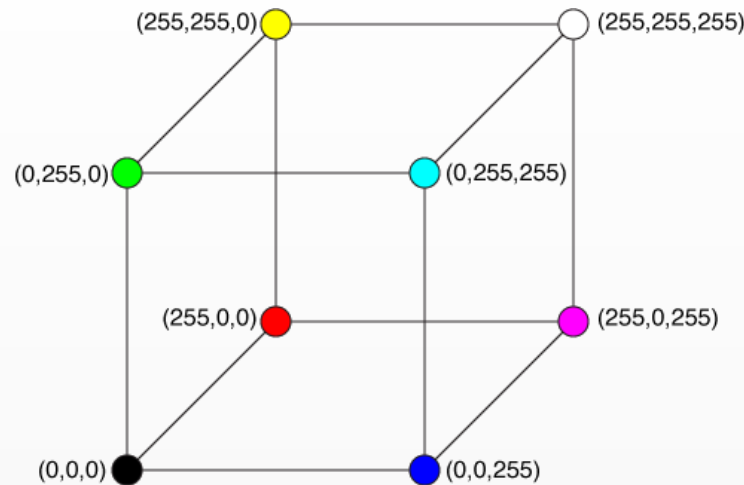
◎ Color Models:

- They specify a coordinate system and a subspace within that system, where each color is represented by a single point.
 - RGB
 - CMY, CMYK
 - HSL/HSV/HSI
 - YUV, YCbCr

Color Spaces

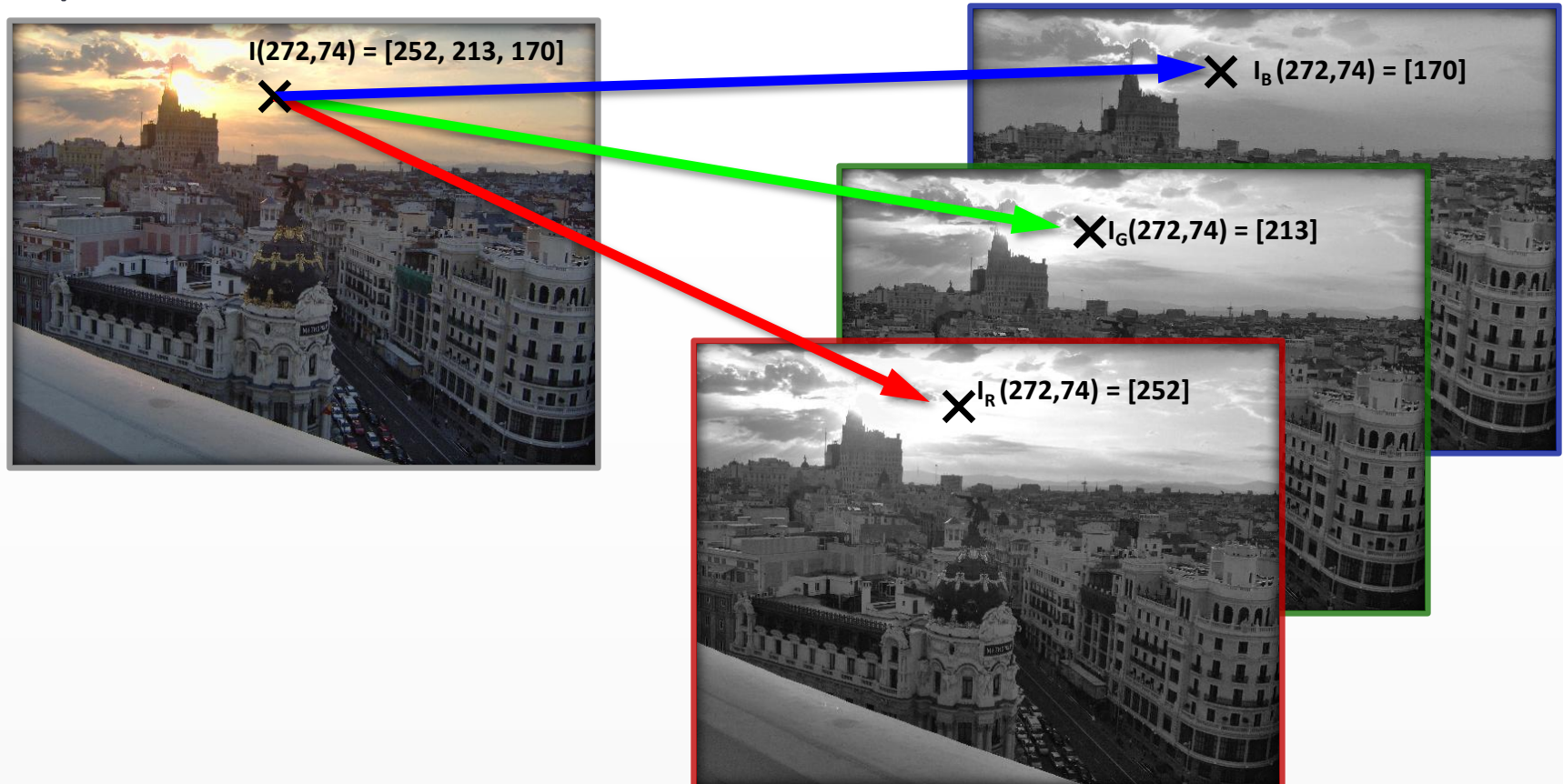
◎ RGB:

- Most common color model
- Channels: Red, Green, Blue
- All components are depending on luminosity
- All channel needs to be coded with the same bandwidth
- Changing the intensity level is not efficient, all 3 channels has to be modified



Color Images

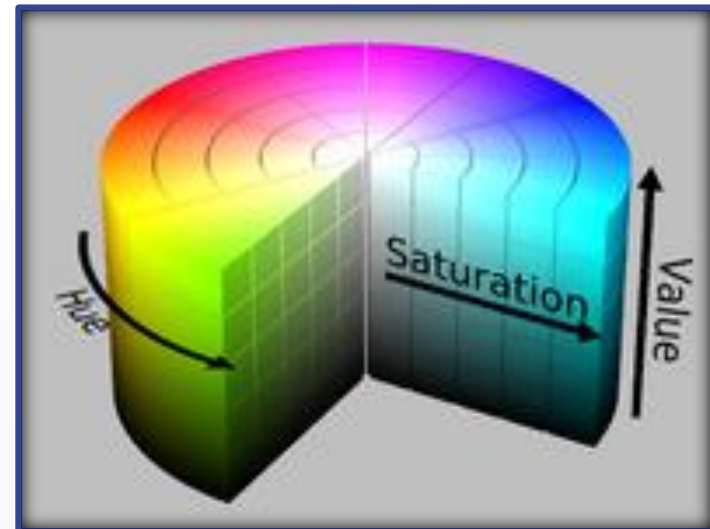
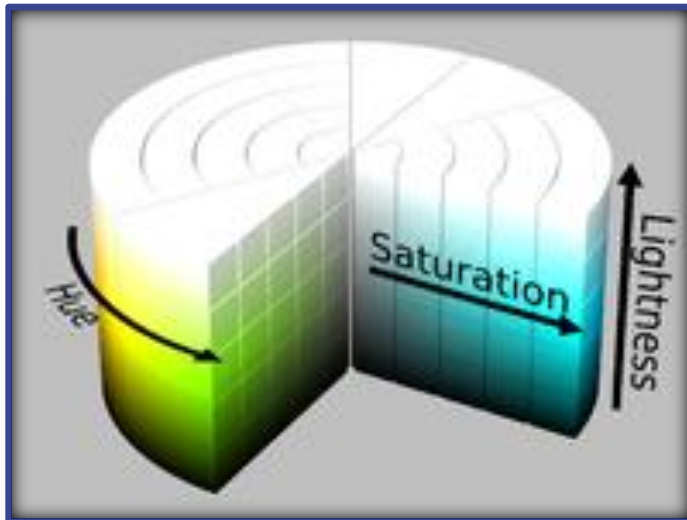
- Color images are formed by combination of different color planes



Color Spaces

◎ HSL, HSV:

- The components are more intuitive
 - **Hue**: the angle around the central vertical axis (defined in degrees)
 - **Saturation**: the distance from the central axis
 - **Lightness** or **Value**: the height



Source: http://en.wikipedia.org/wiki/HSV_color_space

Color Spaces

◎ Y'UV:

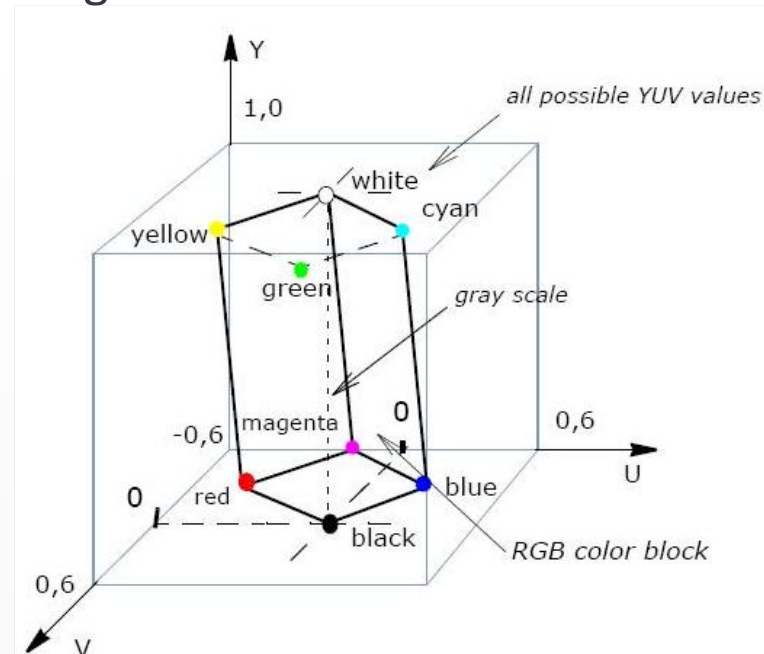
- Used in compression, in the PAL and SECAM composite color video standards.
- Luma is coded in a separate channel (Y')
- Takes human perception into account allowing reduced bandwidth for chrominance components (U,V)

◎ Conversion from RGB to Y'UV:

$$Y' = 0.299 * R + 0.587 * G + 0.114 * B$$

$$U = 0.492 * (B - Y')$$

$$V = 0.877 * (R - Y')$$



Color Spaces

◎ **CMY:**

- used in printing
- Based on the subtractive color model: describes what kind of inks need to be applied, so the reflected light produces the given color.

◎ **CMYK:**

- The black produced by the mixture of CMY is not really black in practice
- Black ink is added as 4th component.

◎ **CIE:**

- the CIE color model is based on how humans perceive color
- was developed to be completely independent of any device
- (CIE stands for *Comission Internationale de l'Eclairage*)

Source: http://dba.med.sc.edu/price/irf/Adobe_tg/models/cie.html

Sources

Fundamentals of Digital Image and Video Processing lectures by Aggelos K. Katsaggelos

Image Processing slides of Csaba Benedek: „Bevezetés és Programozási Környezet Bemutatása.” 2008