Basic Image Processing

PPKE-ITK

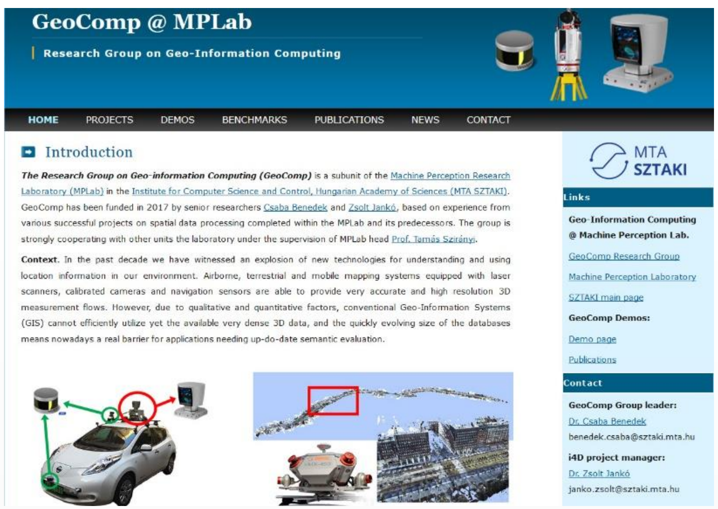
Basic Image Processing Algorithms

- Course responsible and lecturer:
  - Csaba Benedek, PhD
    - Address:
      - *Primary*: MTA SZTAKI (Institute for Computer Science and Control, Hungarian Academy of Sciences), Machine Perception Laboratory, 1111 Budapest, Kende utca 13-17, room 306
      - *Secondary*: PPKE ITK room 408 (only for preliminary fixed meetings)
      - E-mail: benedek.csaba@itk.ppke.hu
- Course webpage:
  - http://kep.itk.ppke.hu/

Personal site: http://mplab.sztaki.hu/~bcsaba

About myself…

- **Research:** computer vision, pattern recognition, 3D sensors (laser scanning), biometrics (gait recognition)
- MTA SZTAKI – leader of Research Group on Geo-Information Computing (GeoComp) @ Machine Perception Laboratory
- PPKE ITK – associate professor (in part time, since 2015)

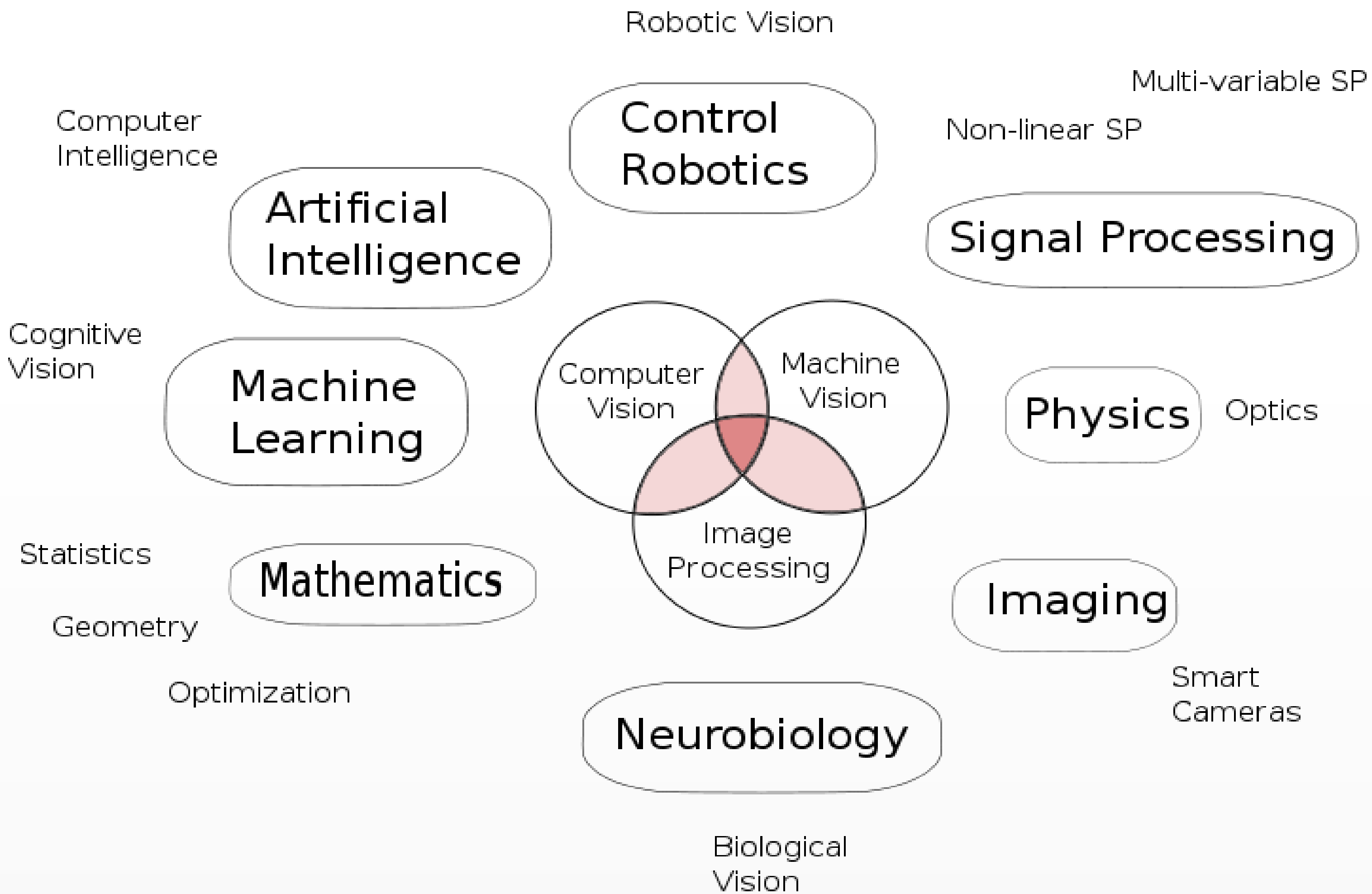


SZTAKI GeoComp Research Group

http://mplab.sztaki.hu/geocomp

PPCU Archaeological GIS Laboratory

What is this Course About?



Source: Wikipedia

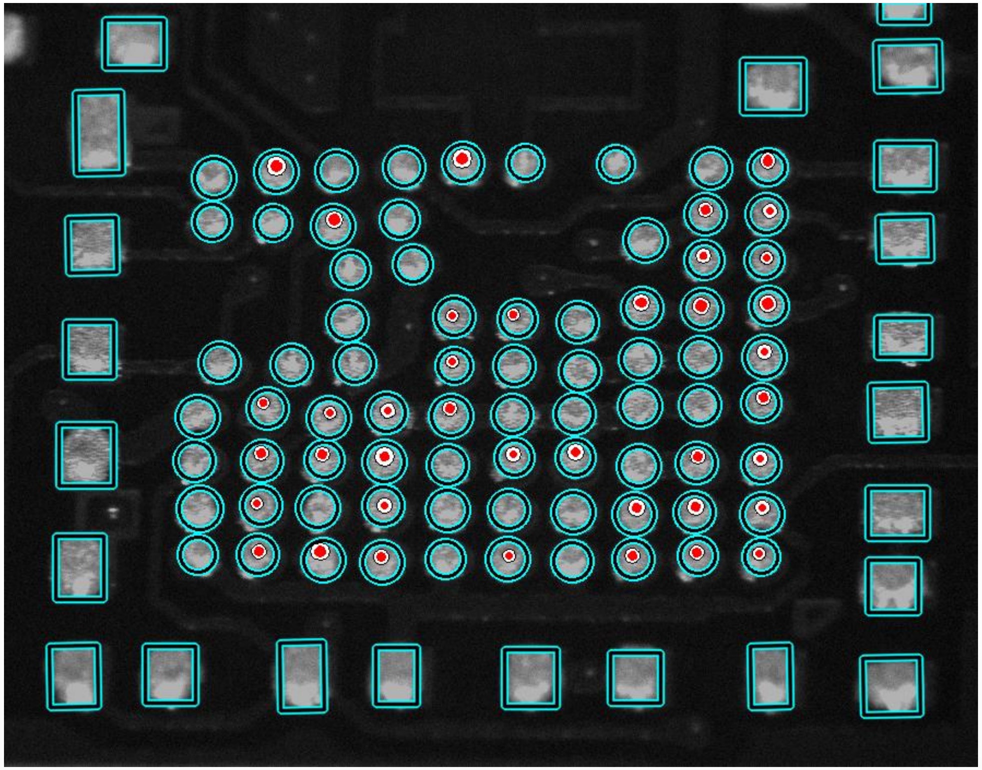
What is this Course About?

- **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/

Machine vision - example

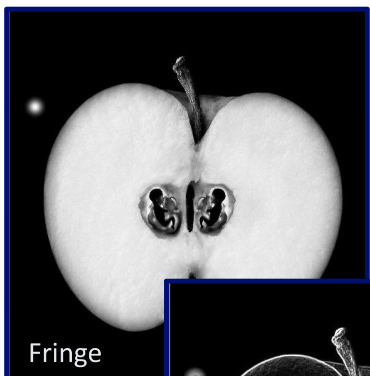
- Optical analysis of scooping artifacts in printed circuit boards

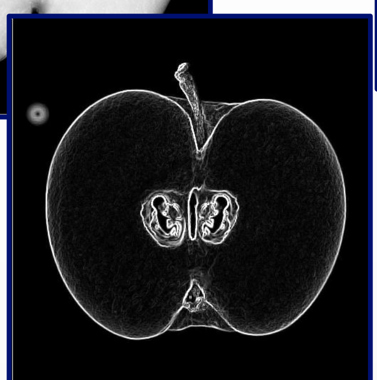


Cs. Benedek, O. Krammer, M. Janóczki and L. Jakab: ”Solder Paste Scooping Detection by Multi-Level Visual Inspection of Printed Circuit Boards”, *IEEE Trans. on Industrial Electronics*, vol. 60, no. 6, pp. 2318 - 2331, 2013

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.”



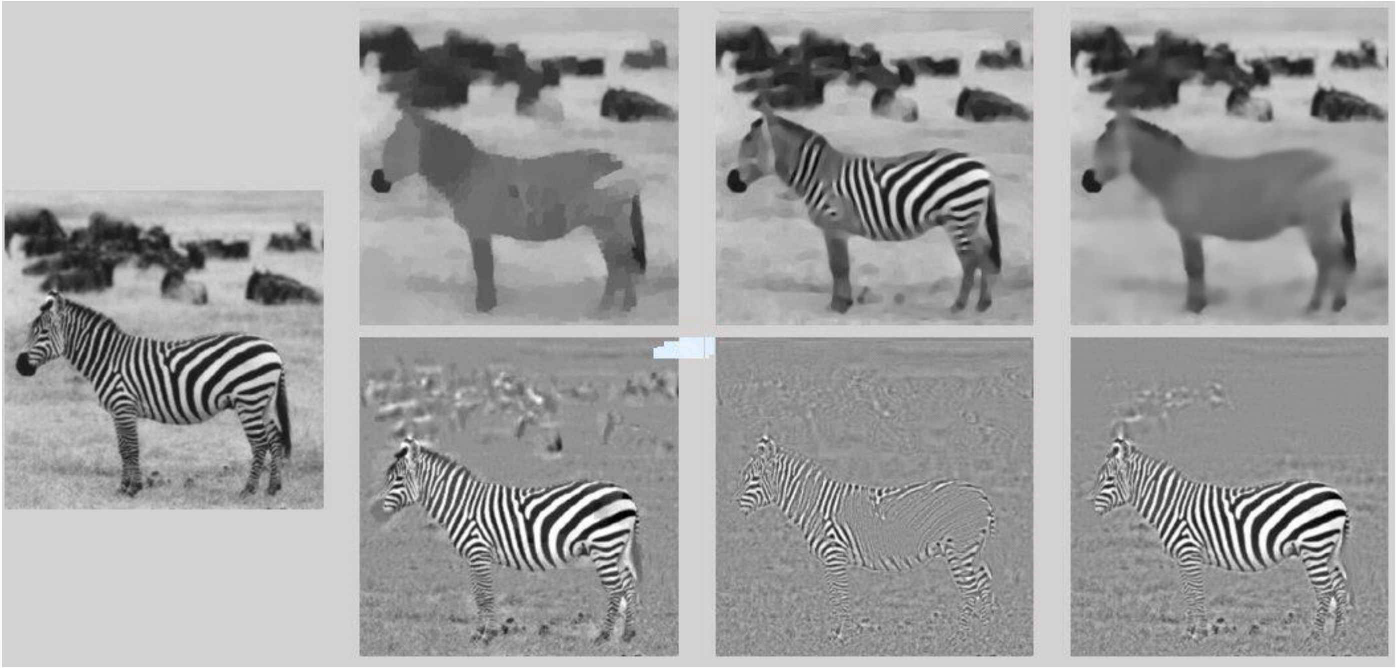


The windmill at Wijk bij Duurstede by Jacob van Ruisdael (1670)

Source: Wikipedia

Image processing example

- Image Decomposition Into Cartoon and Texture Parts



Cartoon (main object shapes)

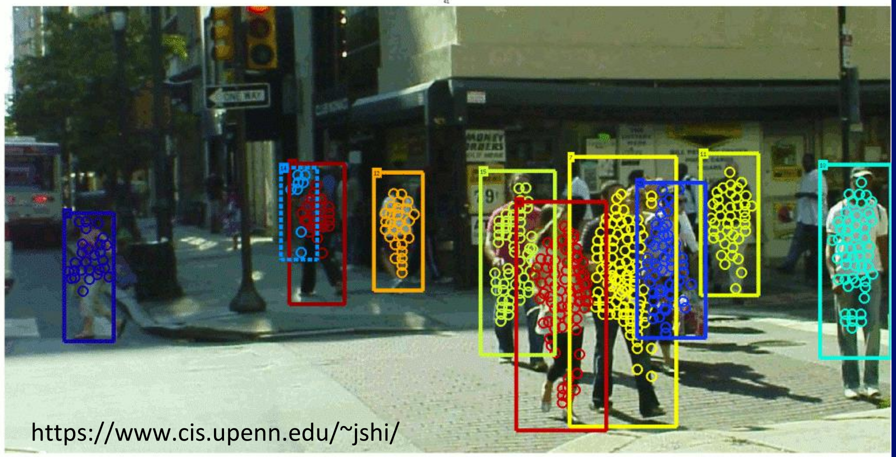
Texture (pattern on the object cover)

Input image Method 1 Method 2 Method 3

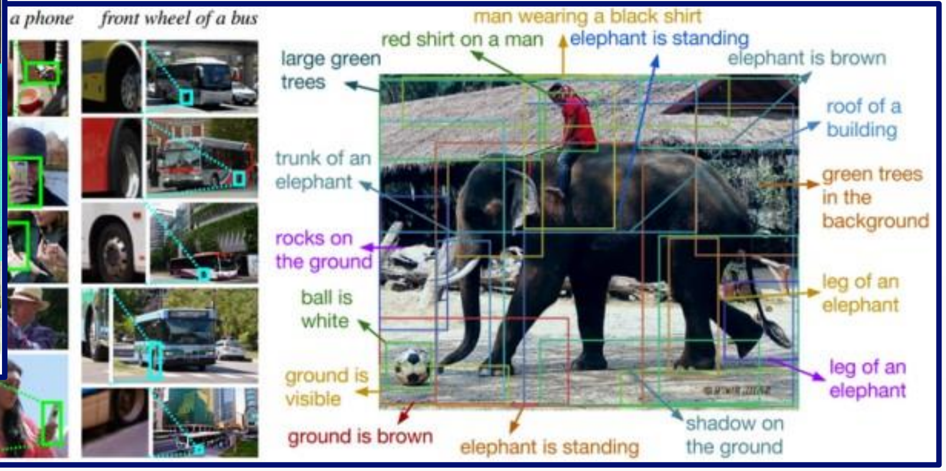
Dániel Szolgay, Tamás Szirányi, „Adaptive Image Decomposition Into Cartoon and Texture Parts Optimized by the Orthogonality Criterion” *IEEE Trans. on Image Processing*, 21 (8). pp. 3405-3415, 2012

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."



https://www.cis.upenn.edu/~jshi/



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

Source: Wikipedia

Various levels of computer vision tasks

- Image classification: assigning a class label to the image

Car: present
Cow: present
Bike: not present
Horse: not present
…

- Object localization: define the location and the category

Location

Category

Video Traffic on the Internet

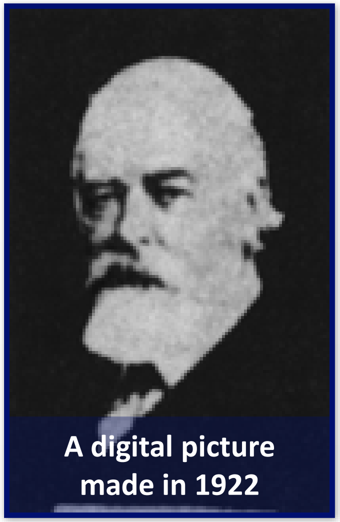
- A few crazy predictions from Cisco for 2020:
  - video traffic will be 82% of all IP traffic (both business and consumer), up from 70% in 2015,
  - it would take more than 5 million years to watch the amount of video that will cross global IP networks each month,
  - every second, a million minutes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

A digital picture produced in 1921



A digital picture made in 1922

  - 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.

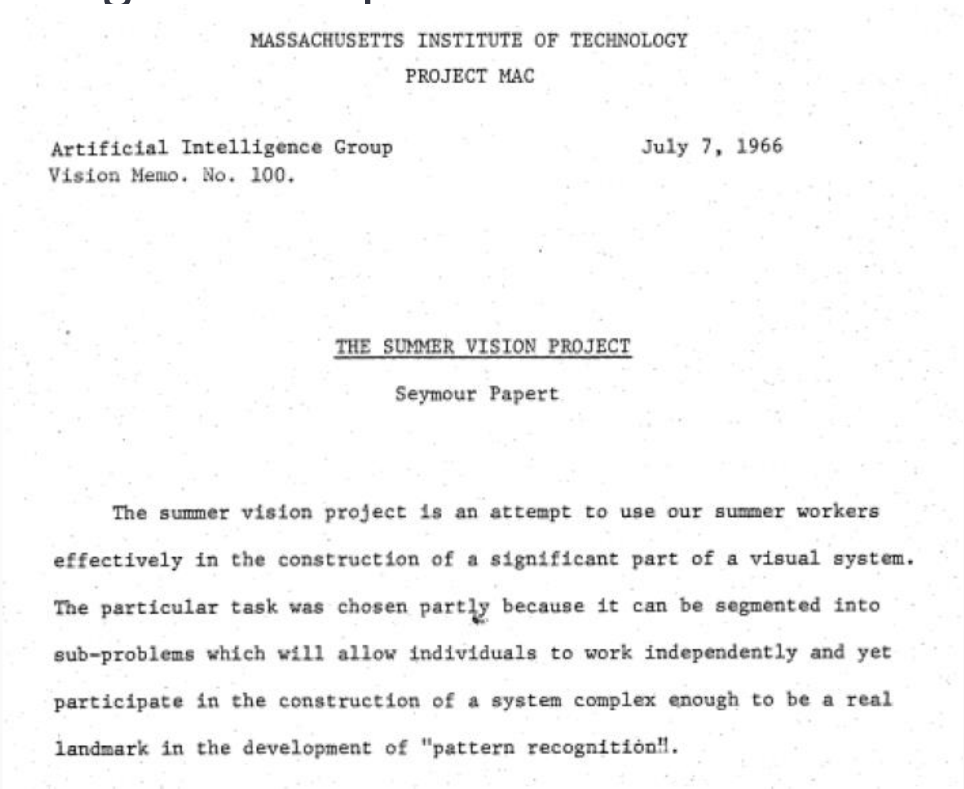
Little bit of History

- 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?

The first picture of the moon by a U.S. spacecraft. (Courtesy of NASA.)

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".

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)

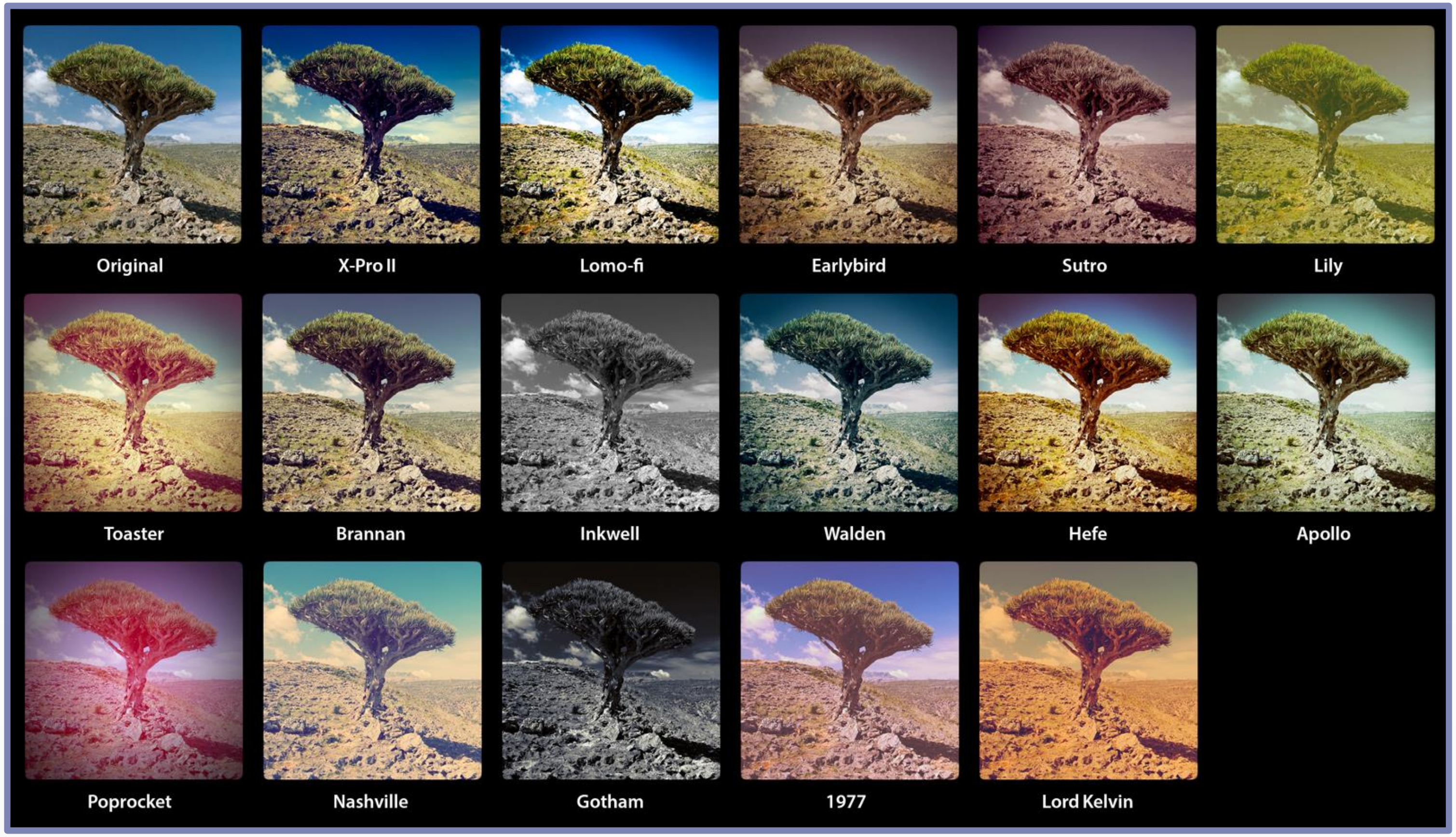
Applications

- 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

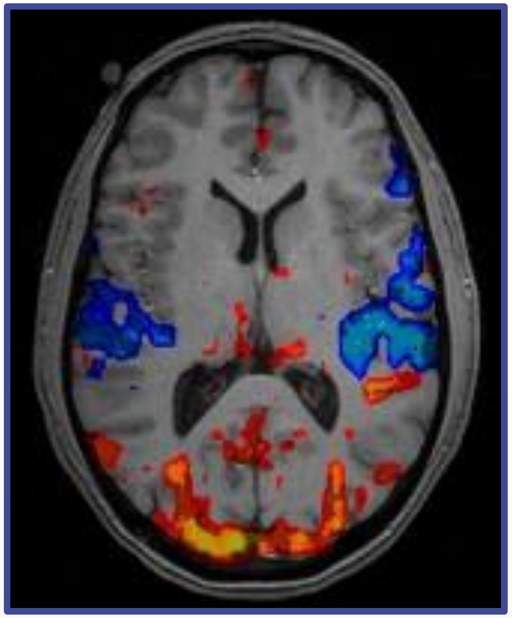
Applications

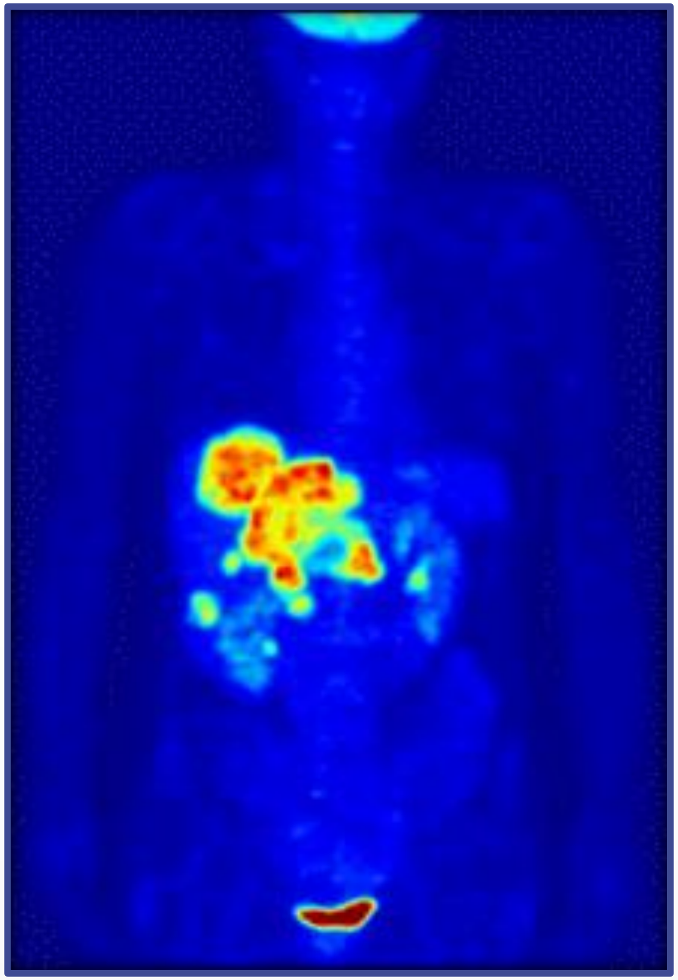
- Panoramic images

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/

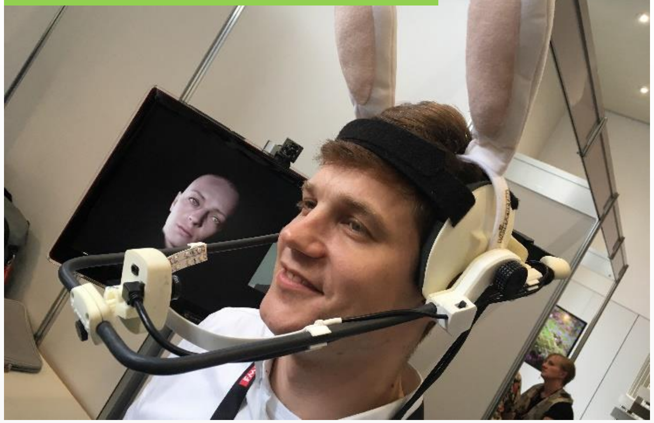
Applications

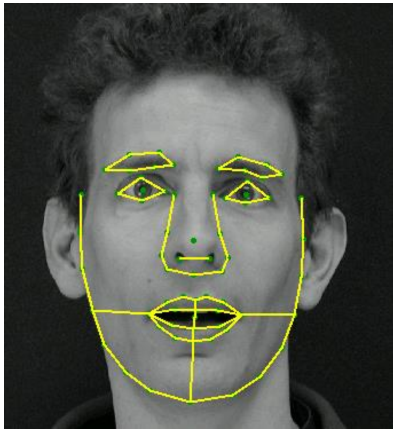
- Facial retargeting:

Entertainment

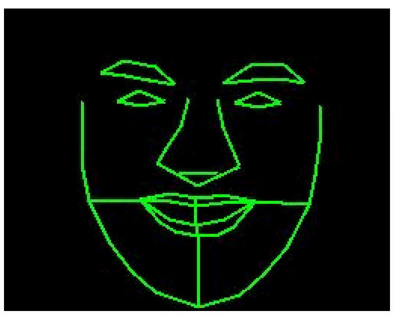
https://www.disneyresearch.com/

FMX 2017, Stuttgart





Active Shape Models (ASM)



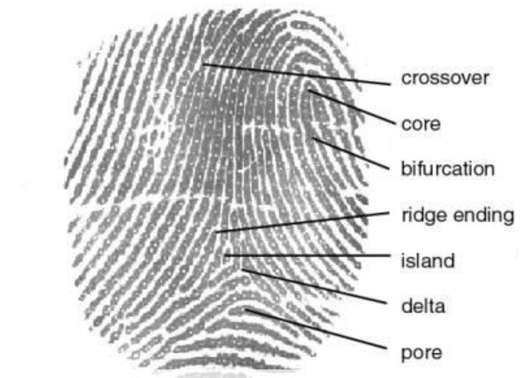
T.F.Cootes, M.Ionita, C.Lindner and P.Sauer, "Robust and Accurate Shape Model Fitting using Random Forest Regression Voting", ECCV 2012

Applications

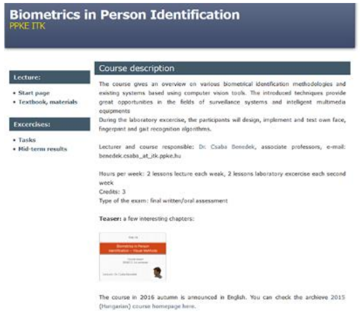
- Emotion recognition:

Applications

- Identity verification and recognition based on biometrics:



http://www.biometrics.gov/


Biometrics in Person Identification
PPKE ITK

Course description

Lecture:
- Start page
- Textbook, materials

Excercises:
- Tasks
- Mid-term results

The course gives an overview on various biometrical identification methodologies and existing systems based using computer vision tools. The introduced techniques provide great opportunities in the fields of surveillance systems and intelligent multimedia equipments
During the laboratory excercise, the participants will design, implement and test own face, fingerprint and gait recognition algorithms.

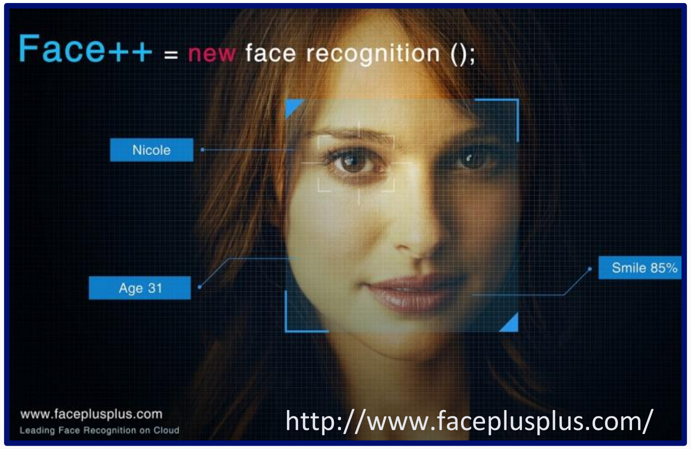
Lecturer and course responsible: Dr. Csaba Benedek, associate professors, e-mail: benedek.csaba_at_itk.ppke.hu

Hours per week: 2 lessons lecture each weak, 2 lessons laboratory excercise each second week
Credits: 3
Type of the exam: final written/oral assessment

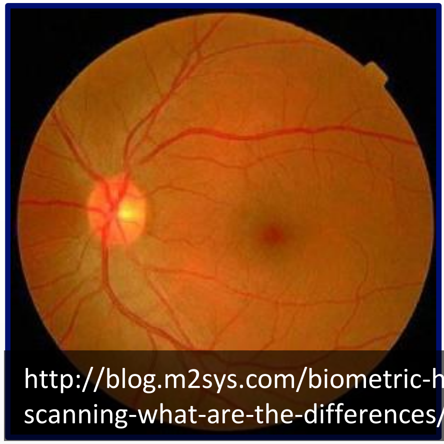
Teaser: a few interesting chapters:

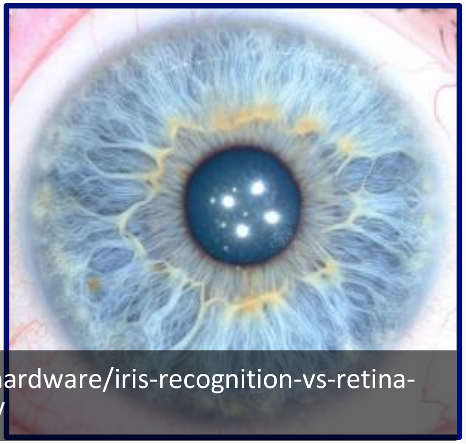
The course in 2016 autumn is announced in English. You can check the archieve 2015 (Hungarian) course homepage here.

- See more: http://biometrika.itk.ppke.hu/



http://www.faceplusplus.com/





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

Applications

- Video surveillance:

Thesis topic option @ PPKE/SZTAKI
http://mplab.sztaki.hu/geocomp
Contact: Csaba Benedek

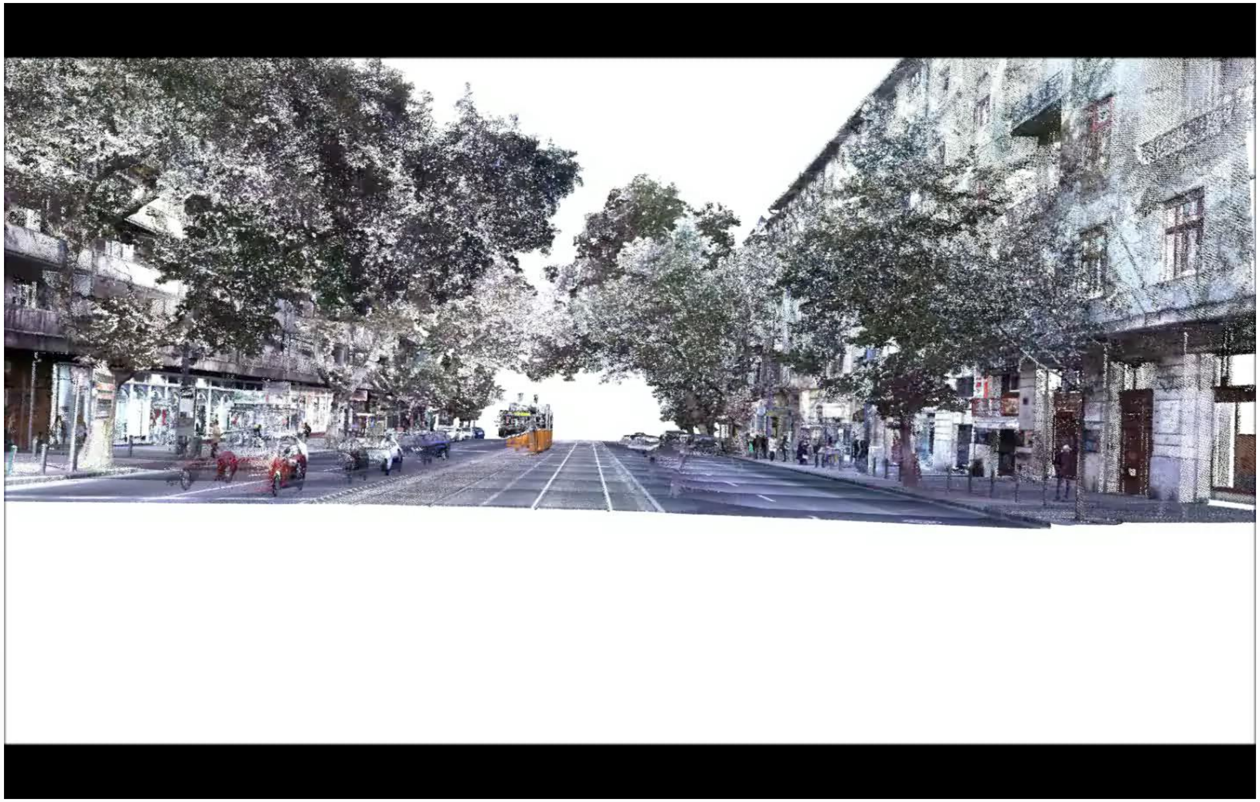
Applications
3D scene understanding via laser scanning (Lidar) data

Automated data filtering, correction, vectorization and interpretation

Thesis topic option @ PPKE/SZTAKI
http://mplab.sztaki.hu/geocomp
Contact: Csaba Benedek

Laser scanning

Point cloud classification



- pedestrian
- moving car
- parking car
- tall vehicle
- column/bole
- vegetation
- road
- building wall

B. Nagy, and Cs. Benedek: "3D CNN Based Semantic Labeling Approach for Mobile Laser Scanning Data", ***IEEE Sensors Journal***, to appear 2019

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
  - ...


https://www.tesla.com/


© Google

UBER

Thesis topic option @ PPKE/SZTAKI
http://mplab.sztaki.hu/geocomp

Lidar and image data fusion for self-driving cars in SZTAKI

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:
  - Csaba Benedek: benedek.csaba at itk.ppke.hu
  - Miklós Koller: koller.miklos at itk.ppke.hu
  - Márton Bese Naszlady : naszlady.marton.bese at itk.ppke.hu

Administration and Course Requirements

- ***Lectures/Seminars***:
  - Start time: 8:15 AM
  - The attendance is obligatory
  - In every lecture between Week 2 and 12 there will be a short test from the previous lecture's topic
    - Each time 2 short questions for a maximum of 2 points (in total 22 points)
    - > 60% (13.5/22 points) is a requirement to take the final exam
    - Recovery option: if (and only if) you fail to pass the threshold by the end of the semester, you will have a chance on Week 13 to correct your two worst/missing results.
    - no other occasions during the semester to replace missing tests!
  - 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 to the online submission system on time
  - You have to understand your code and be able to explain it.
  - Submitting all tasks in time is a requirement of passing the course.
  - You can collect maximum 11 points, threshold is 7 (details clarified by Miklós and Márton, see Lab01 „General course info")

Administration and Course Requirements

- ***Assignments***:
  - During the semester we will hand out 4 longer programming assignments.
  - You have to complete and submit your solution via the online submission system
  - Maximum 22 points, threshold for passing: 14
- ***Oral exam*** at the end of the semester:
  - To be able to participate to the exam you have to pass **all three threshold** for the (i) short tests, (ii) programing practices and (iii) assignments.
  - ***Mid term points count in 30% into the final grade, 70% is coming from the oral exam***
- ***Offered Grade – only in expetional cases:***
  - On Week 13, you should write a 12th short test from the previous lecture, and you can recover 1 test from the previous ones
  - 90% rule: minimum 22 (out of 24) points in total from the tests, 10 points from the programing practices and 20 from the assignments

  -> You get a 5 as final mark
- ?

Human Vision

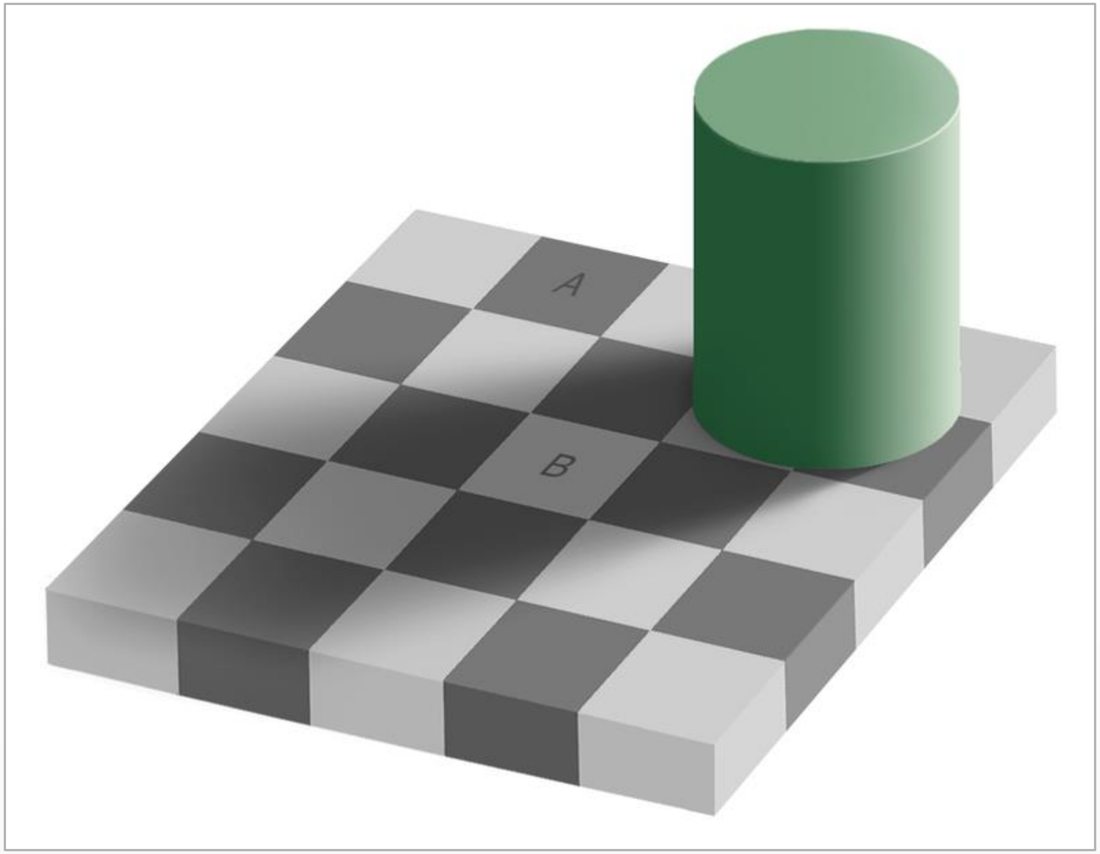
- The **human visual system:**
  - Gives us the ability to process visual stimulus, to be able to detect and interpret information from visible light (build a representation of the surrounding environment).
- The ultimate goal of computer vision is to build a system that is capable of seeing as a human 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

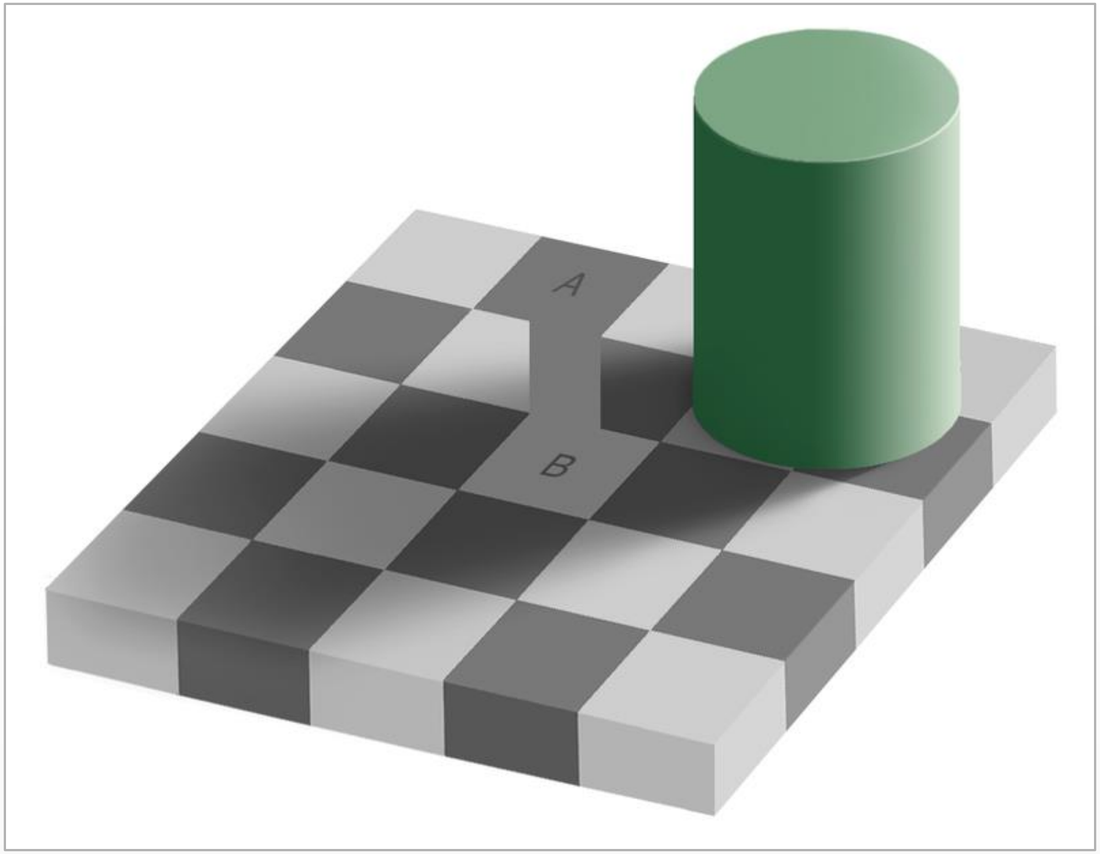
- Lateral inhibition + assumptions tailored for the natural visual world



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

Illusions

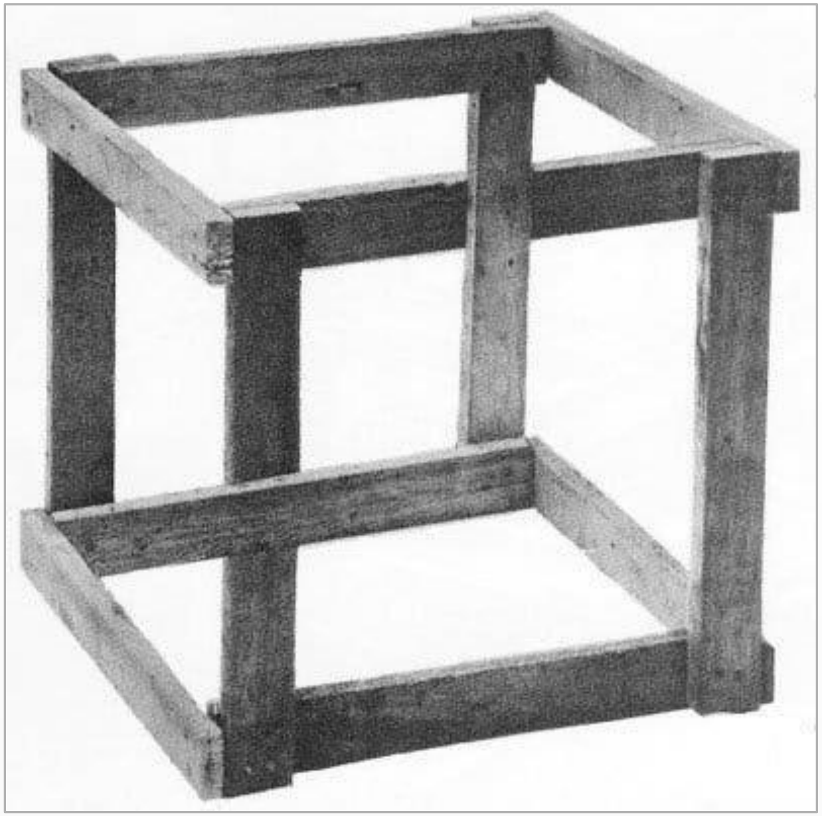
- Lateral inhibition + assumptions tailored for the natural visual world



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

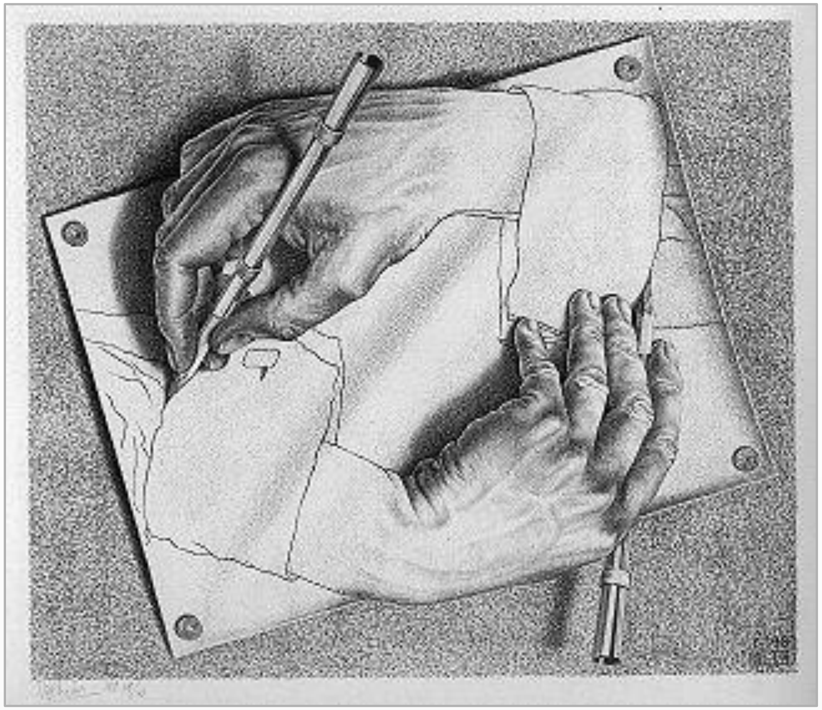
Illusions

- 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)

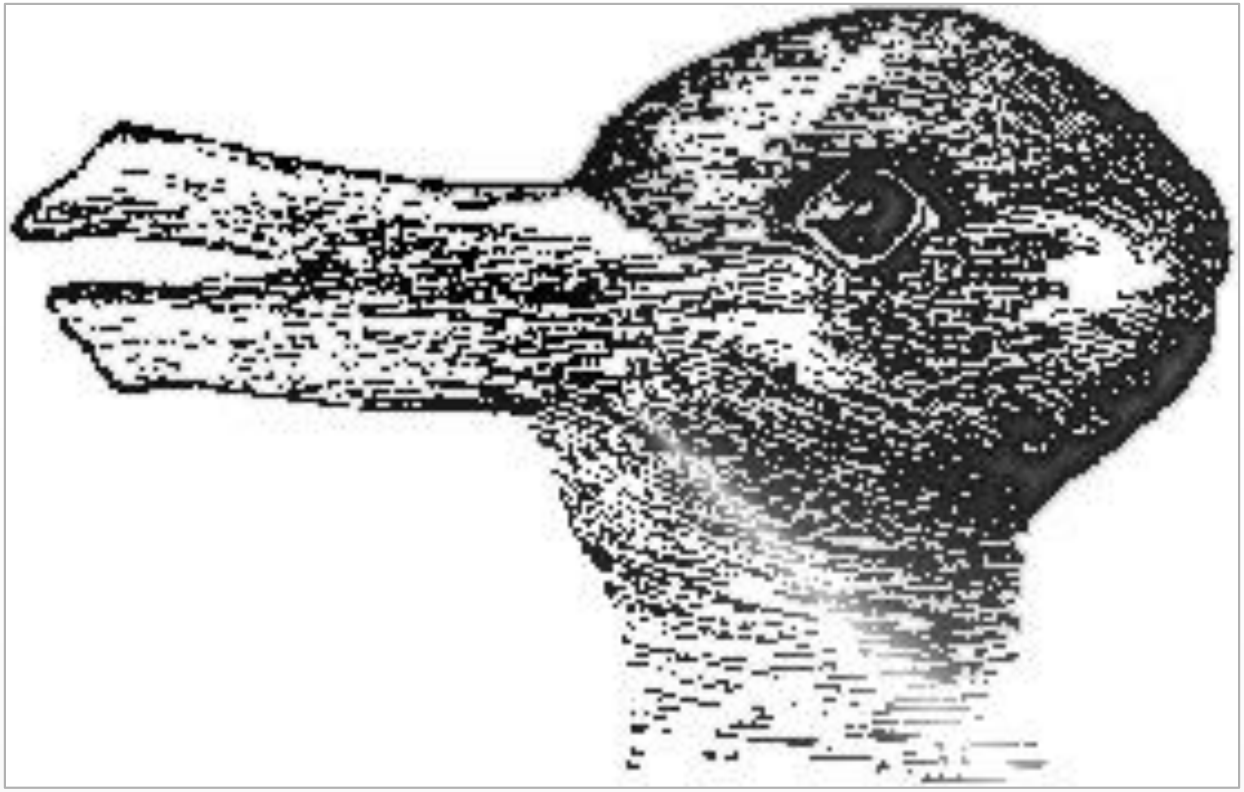
Illusions

- Uncertainty due to incomplete knowledge of the environment

Charles Allan Gilbert
All Is Vanity (1892)

Illusions

- Uncertainty due to incomplete knowledge of the environment

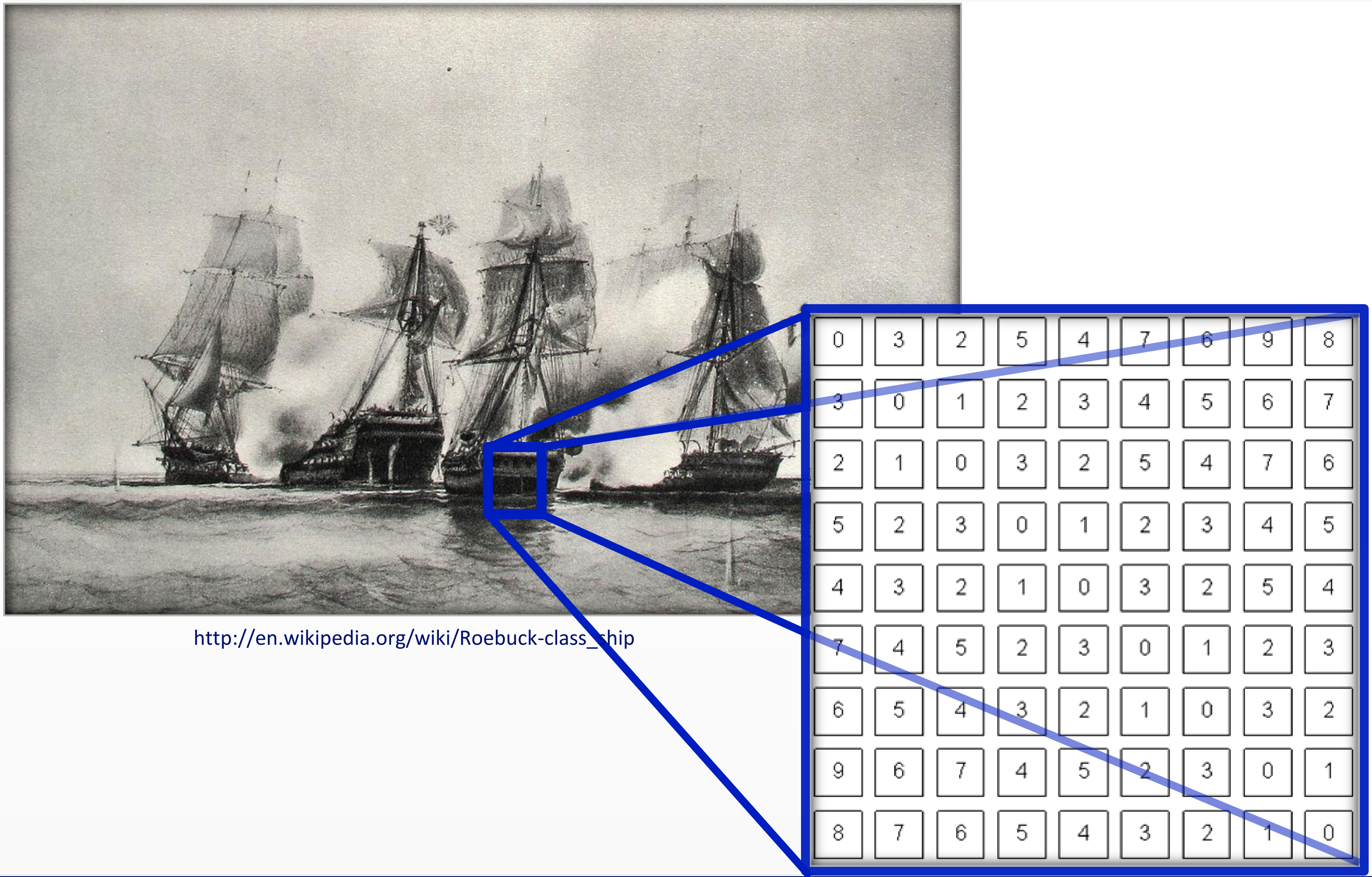


Ludwig Wittgenstein
Rabbit and Duck (1892)

Illusions

Frog & horse

What does a computer „see”?

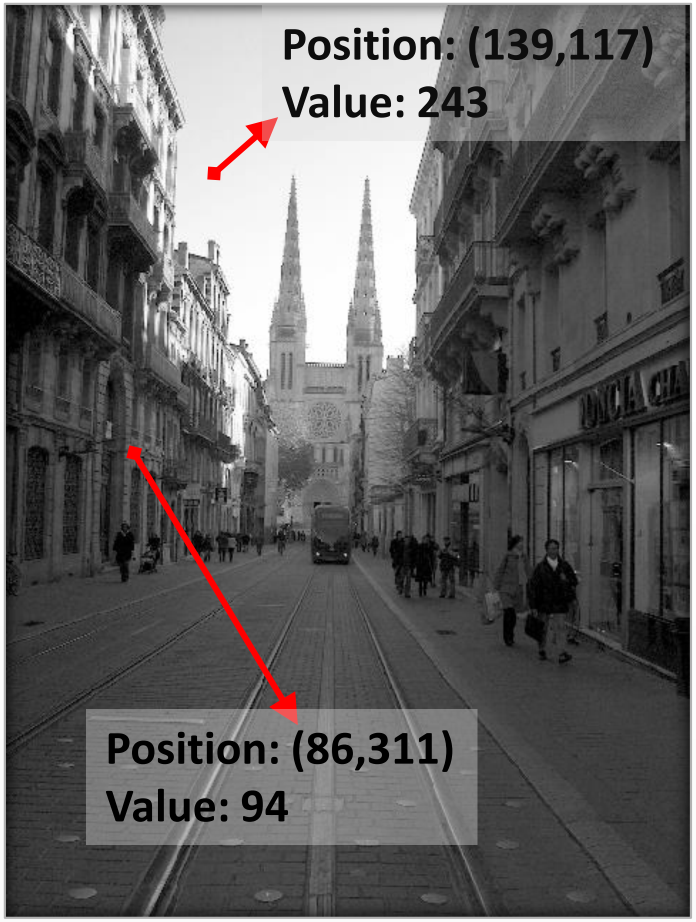


0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

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

What does a computer „see”?

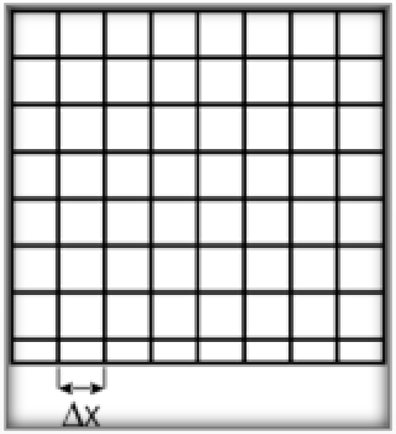
- A **digital image** is discreet representation of a continuous measurement, usually a **2 or 3 dimensional array**.
- An element of this array is a **pixel** (picture element).
- A pixel has a **position** (its coordinates on the image) and an **intensity value**.
- A digital image is discretised both 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.

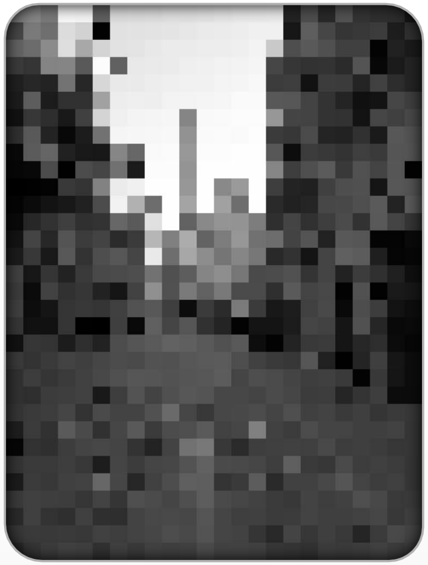
- In case of 2D signals (images) a grid is used for sampling
- The grid points will be represented as pixels.
- The frequency of the sampling defines:
  - How many grid points we have?
  - What is the resolution of the image?
  - How detailed the discretised image is?



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 discreet 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

The Histogram of an Image

- **Histogram**:
  h(k) = the number of pixels on the image with value *k.*

Original Image*

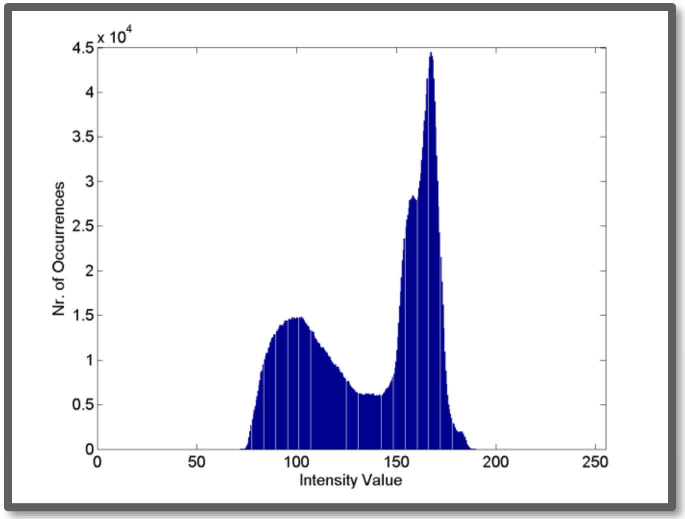


Image Histogram

- The histogram normalized with the total number of pixels gives us the ***probability density function*** of the intensity values.

* Modified version of Riverscape with Ferry by Salomon van Ruysdael (1639)

Image representation with different gray level depths

2 bit
(4 values)

3 bit
(8 values)

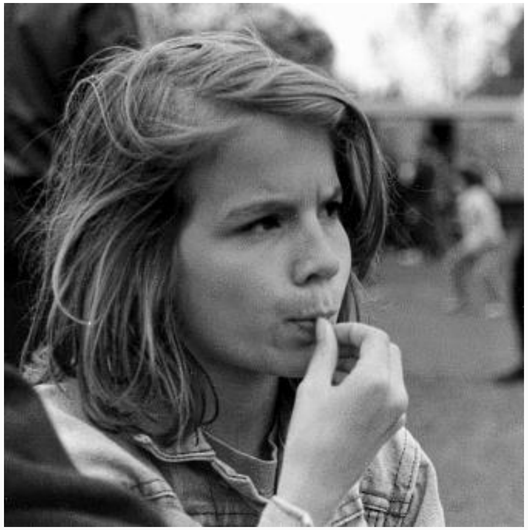
4 bit
(16 values)

6 bit
(64 values)

Slide credit ® Vladimir Székely, BME

Dithering

- Mapping to reduced bit number
  - Trivial solution: truncating the digital word representing the gray value of each pixel
    - pixel=pixel>>4



Original image

4 gray levels

2 gray levels

Dithering with white noise

- Quantitatization for a single bit
  - Each pixel becomes black or white
  - Probability of a given pixel being black or white depends on its original gray level
  - For a grayscale image f let the domain of gray levels of 0…F_{max}
    - A given pixel receives 1 (white) color, if the following condition holds:

$$f \geq random(\ )\cdot F_{\max}$$

- where random() is a randomly generated number from [0,1]

Slide credit © Prof. Vladimir Székely, BME

Dithering with white noise

- With some re-arrangement:

$$f \geq random(\ )\cdot F_{\max}$$

$$f + (1 - random(\ ))\cdot F_{\max} \geq F_{\max}$$

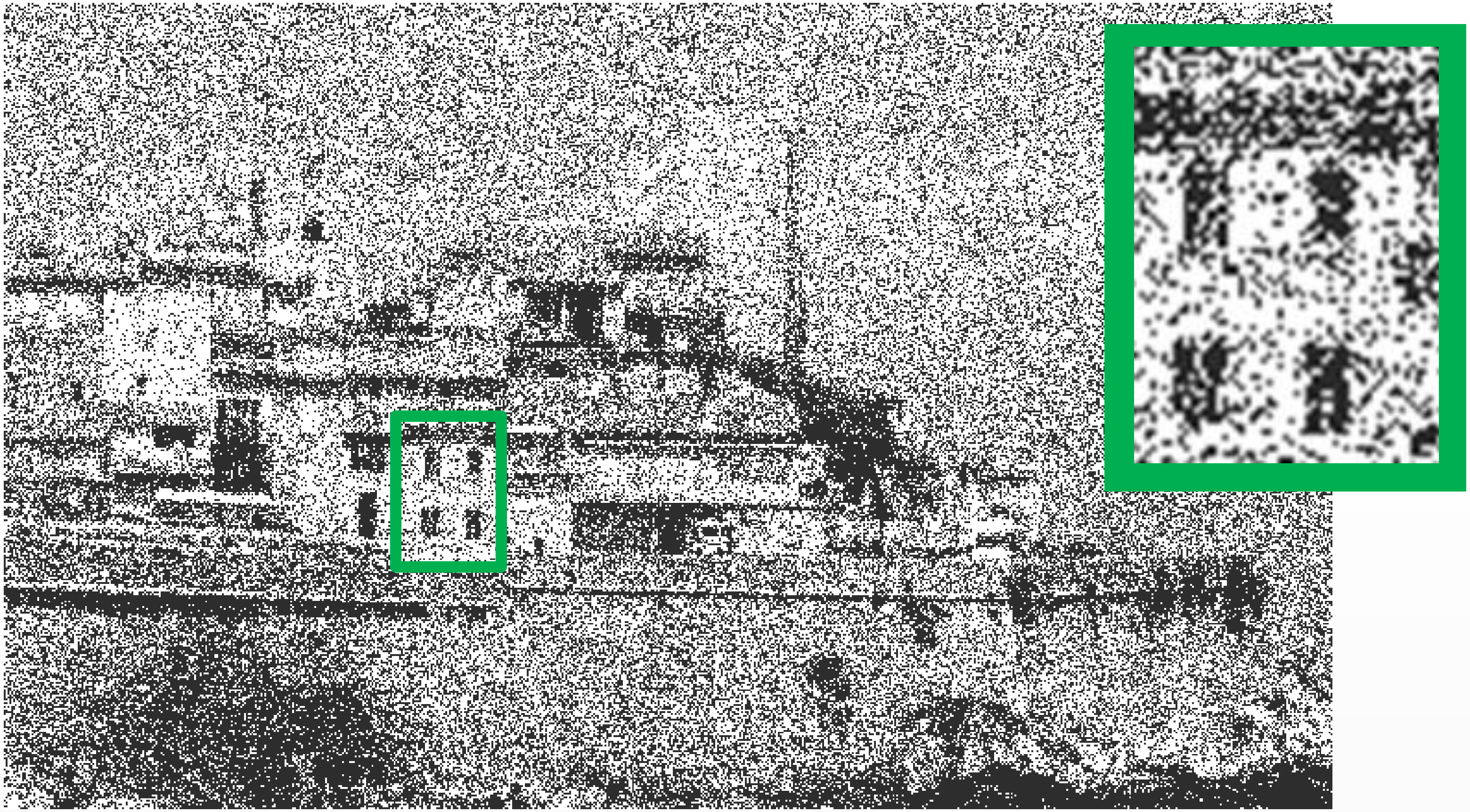
- The algorithm:

$$f + random(\ )\cdot F_{\max} \geq F_{\max}$$

  - we add random numbers („noise”) with values between 0 és F_{max} to the original grayscale image
  - Pixels with values above F_{max} receive label 1 (white), the other ones 0 (black)

Slide credit © Prof. Vladimir Székely, BME

Dithering with white noise

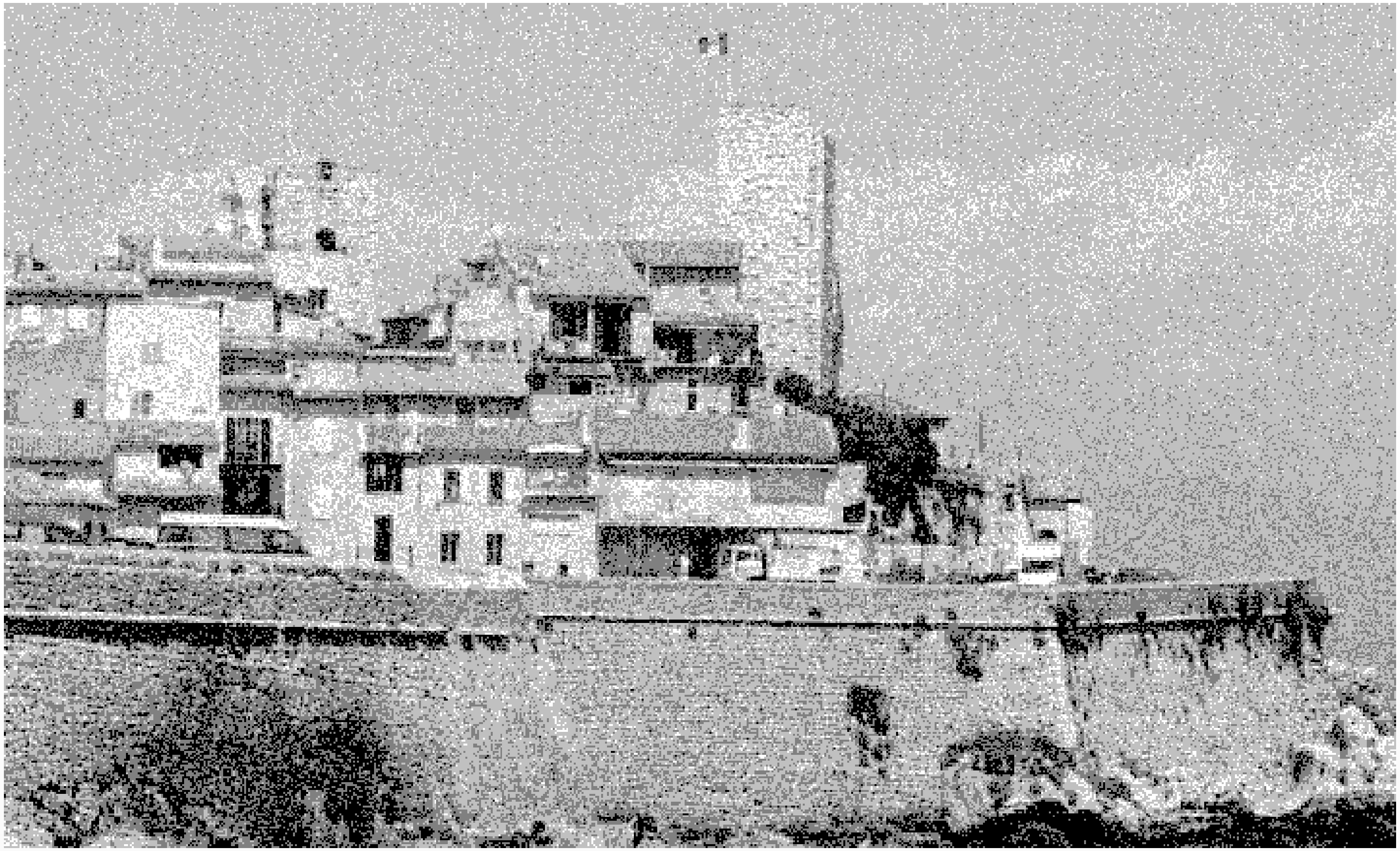


Mapping to 1 bit

Dithering with white noise – multiple bits

- If the gray value of the output image can be represented by multiple (2,3,…) bits, the result of dithering can be significantly enhanced.
- Let *n* be the number of different gray levels in the output image
- Algorithm
  - Add random numbers (noise) between 0 and $F_{max}/(n-1)$ to the image pixel values
  - Pixels with the modified gray levels larger than $k \cdot F_{max}/(n-1)$ but not larger than $(k+1) \cdot F_{max}/(n-1)$ receive the output value k .

Dithering with white noise



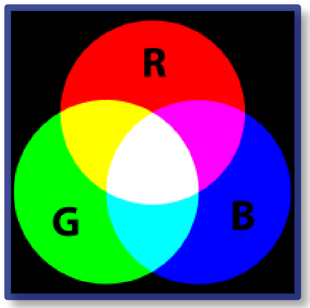
Mapping to 2 bits

Color Images

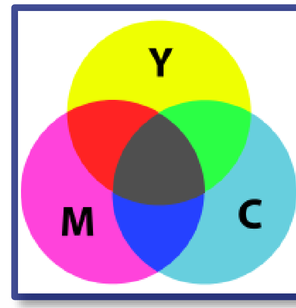
- Humans can distinguish thousands of color shades and intensities, but only a few dozens of gray.
- Color can be a useful descriptor for image segmentation, tracking, detection,…

Color models

- Two main color mixing models:
  - Additive model:



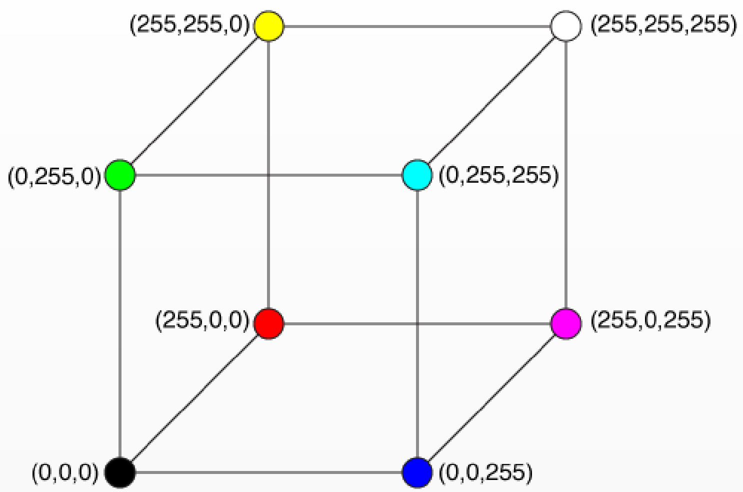
Subtractive model:



- Color spaces:
  - 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

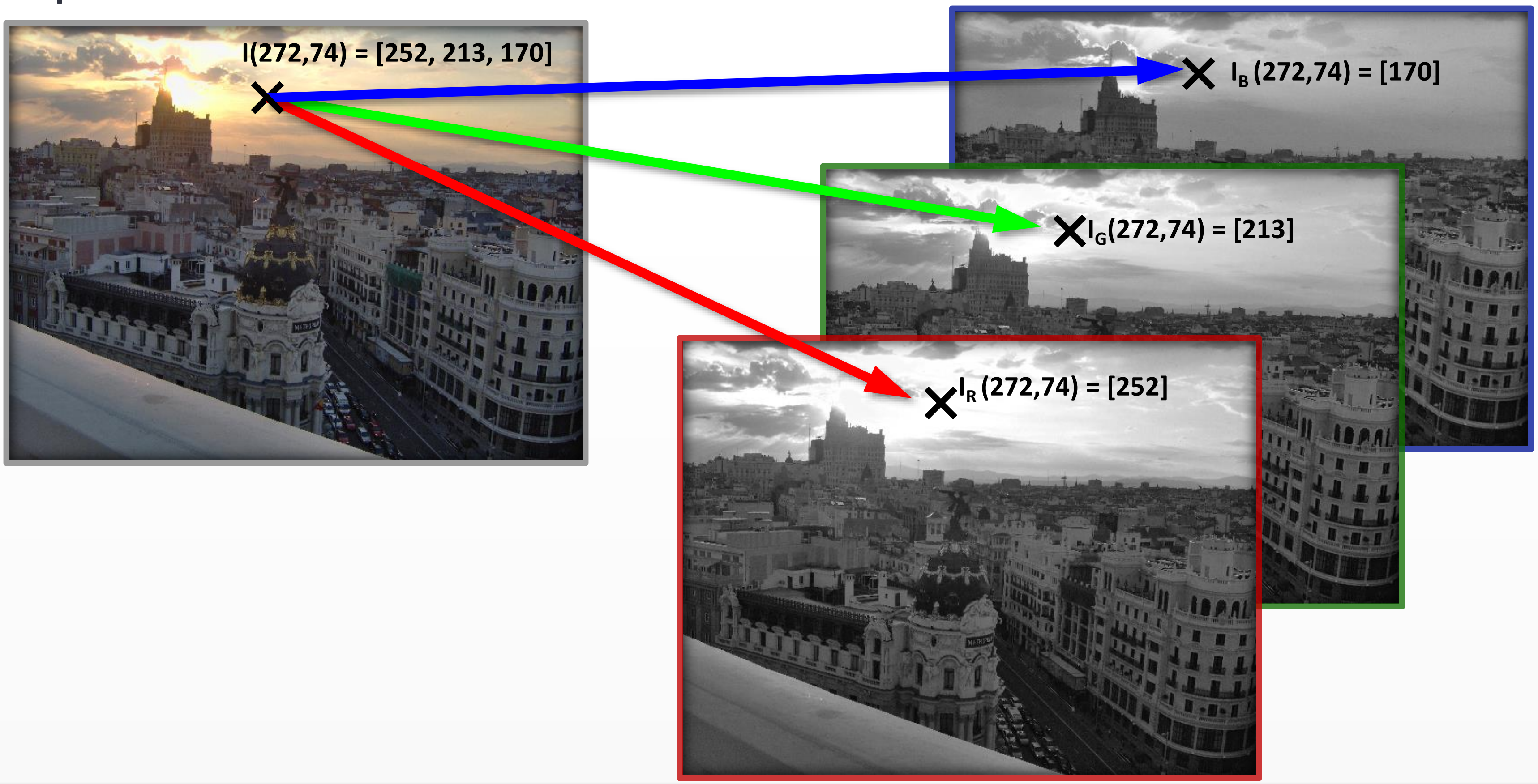
Color Spaces

- **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

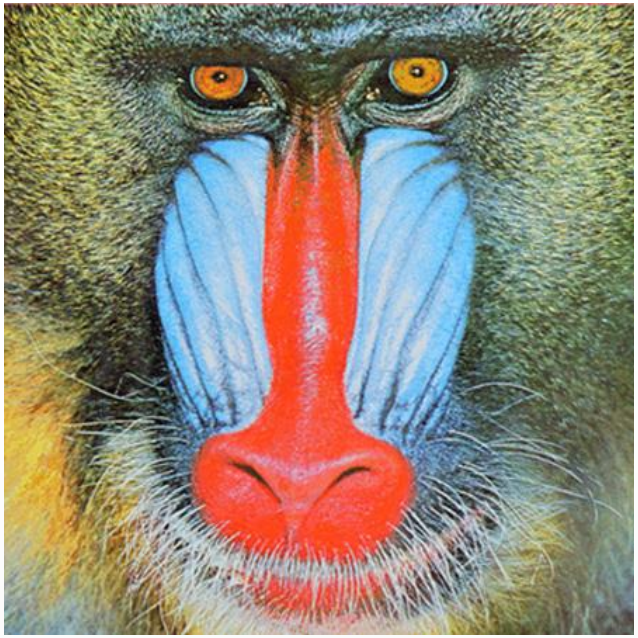


Color Images

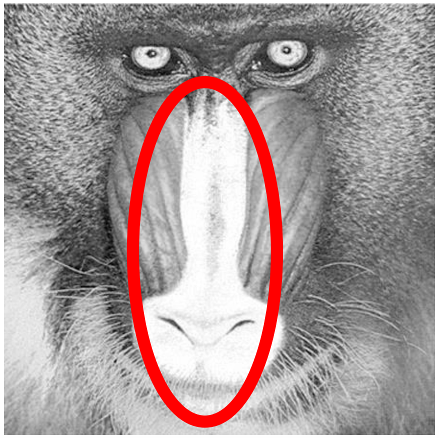
- Color images are formed by combination of different color planes



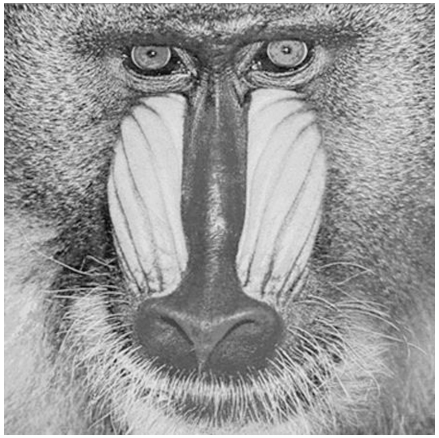
RGB channels



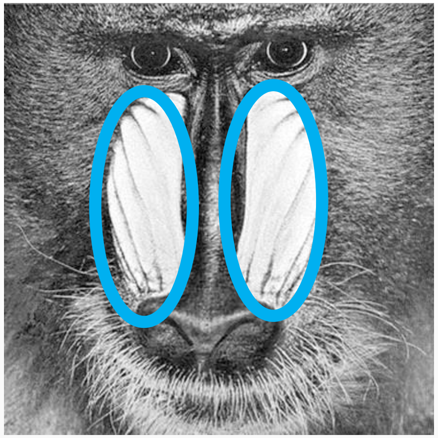
RGB color image



Red channel



Green channel



Blue channel

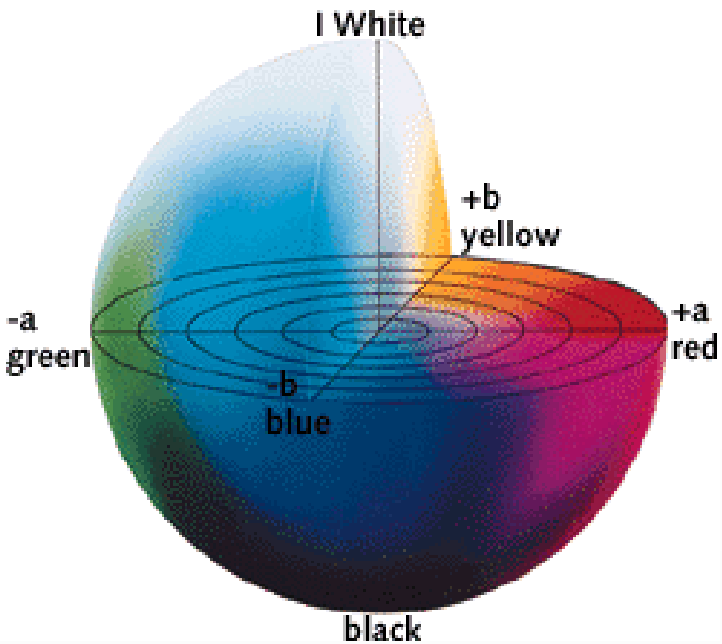
Color Spaces

- **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

CIELAB, Lab, L*a*b

- One luminance channel (L)
 and two color channels (a and b).
- In this model, the color differences which you perceive correspond to Euclidian distances in CIELab.
- The a axis extends from green (-a) to red (+a) and the b axis from blue (-b) to yellow (+b). The brightness (L) increases from the bottom to the top of the three-dimensional model.

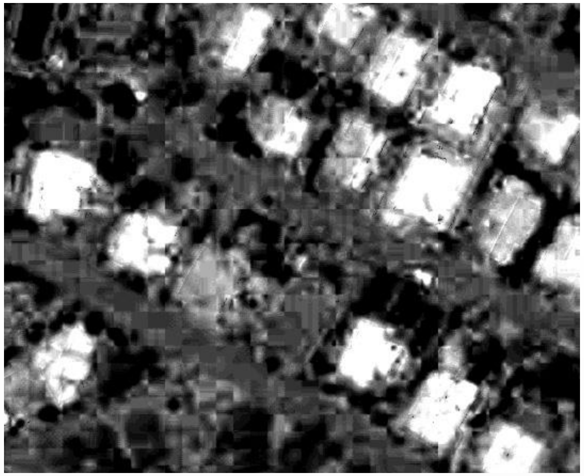


CIE L*a*b*

- Color filtering in CIE L*a*b*: where are the red roods?

RGB image

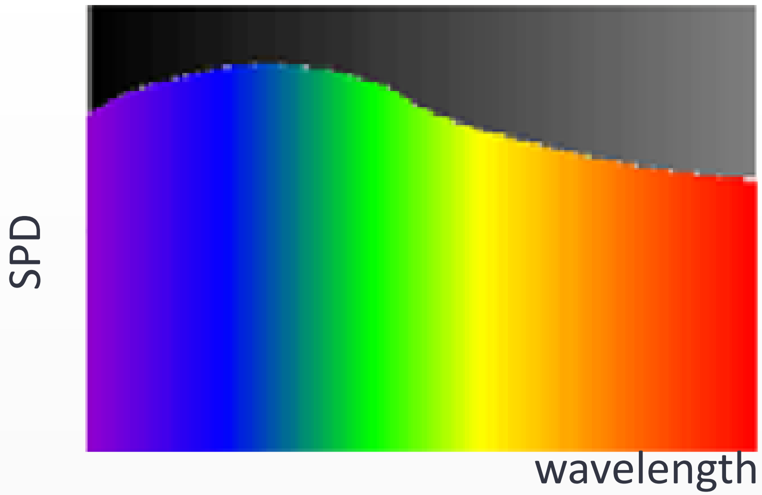
R channel of the RGB image



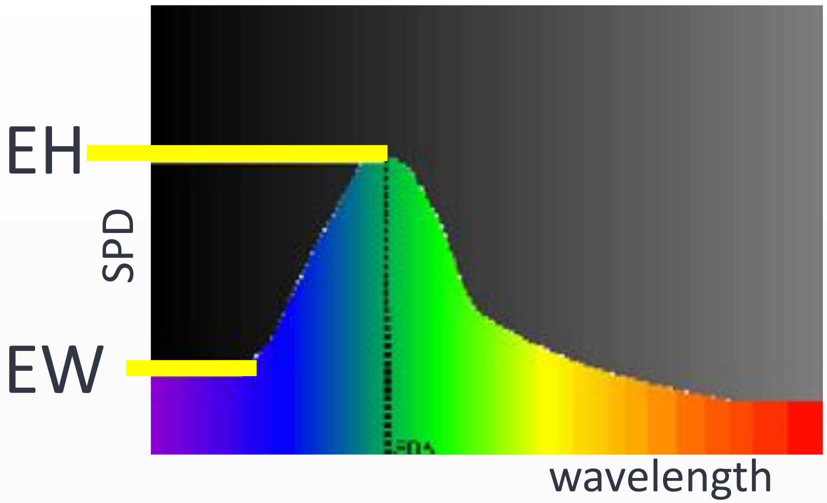
a* channel of the CIE L*a*b* image

HSI color space - fundamentals

- **Hue:** dominant wavelength of the Spectral Power Distribution: EH
- **Saturation:** relative purity or the amount of white light in the mixture EH-EW
- **Intensity:** indicates the dominant wavelength in the mixture of light waves: EW



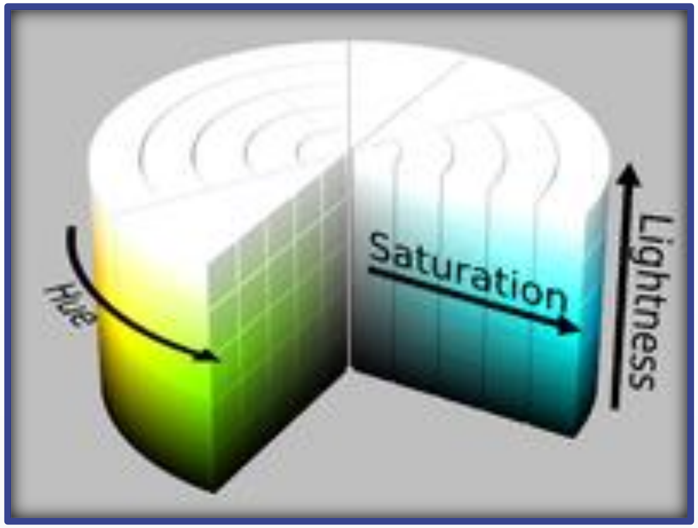
white light

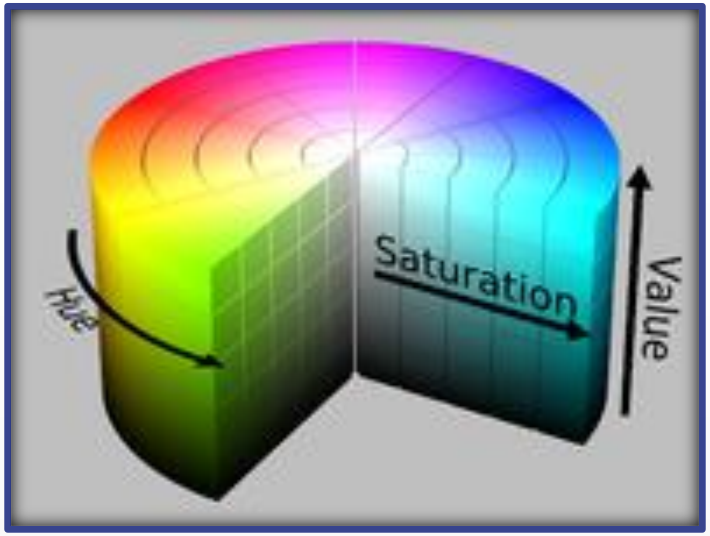


green light

Color Spaces

- **HSI/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
    - **Intensity/Lightness** or **Value**: the height





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

Sources

Fundamentals of Digital Image and Video Processing lectures by Aggelos K. Katsaggelos
Slides of Prof. Vladimir Székely (BME)