
Machine vision

Pak Alexander Alexandrovich
Candidate of tech. sciences
a.pak@kbtu.kz

Глаза того не зрят чего не видит разум
О Bedel

Lecture 0: Introduction

Course overview

- Image Formation
- Motion Vision
- Shape from Shading
- Binary Image Processing
- Photogrammetry/Stereo
- Object Representation Alignment
- Computational Vision
- Fourier Transform
- Convolution
- Line Finding
- Pattern Recognition

Course language platform: **python, matlab**

Course basics

- Number of credits: 3 (0/1/2); 60 hours total; 15 hours per week;
- Your own project

Course arrangement

- Every student must have github repo for the course
- Class representative must share students id, name, email, github repo.

What is it?

- «To see means to know what is where by looking»
David Marr, Vision, 1982
- "Turing Test" - the computer must answer any question about the image that a person can respond to
- What does it really mean? • Vision - the source of semantic information about the world • Vision - a source of information about the distances and sizes of objects

The birth of AI



On computable numbers, with an app to the
Entscheidungsproblem (1936)
Computing machinery and intelligence (1950)

Difference robot vs human



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

Source: S. Narasimhan

Object extraction



Object classification

- вне помещения
- город
- Пекин, Китай
- Пл. Тяньаньмэнь



Object features extraction



Measurements

Stereo Vision



Distance

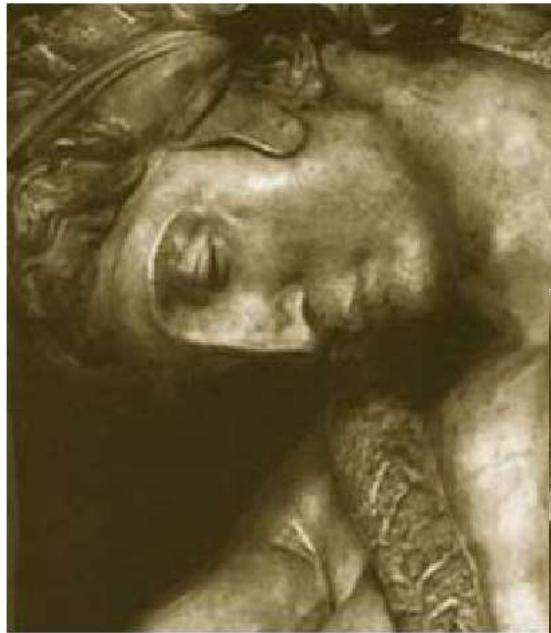


3D

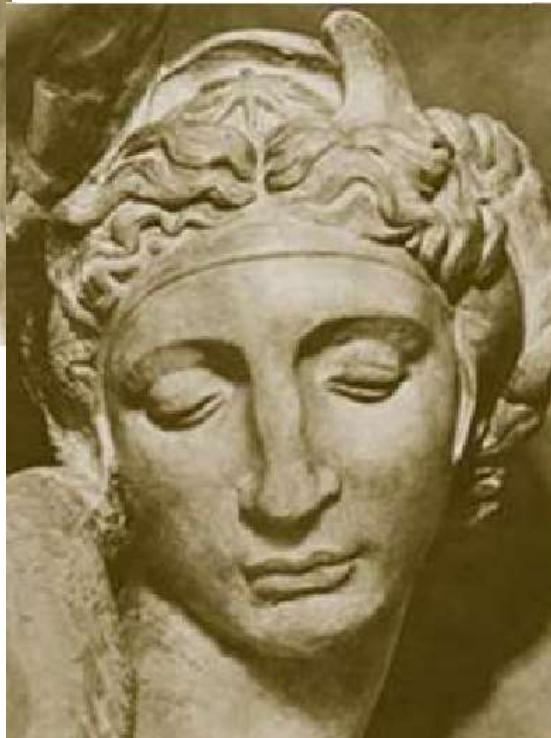


Slide: Svetlana Lazebnik

Foreshortening



Michelangelo 1475-1564

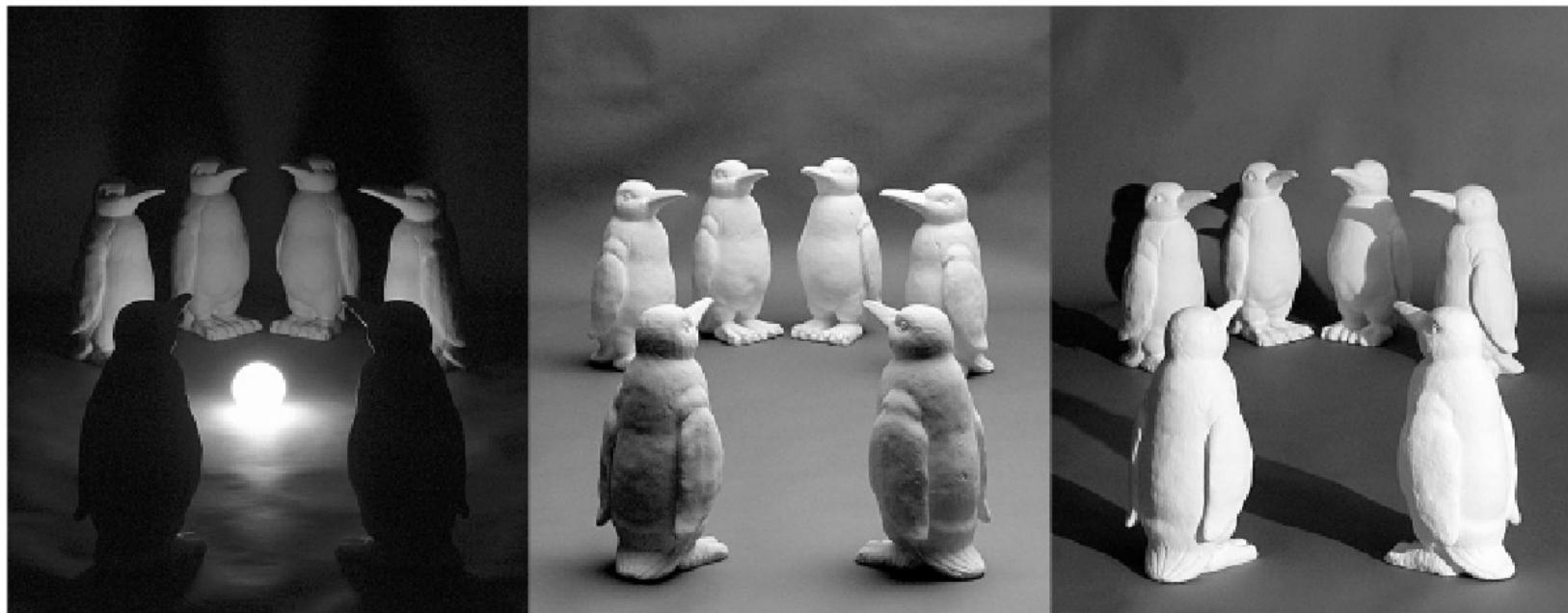


slide credit: Fei-Fei,
Fergus & Torralba

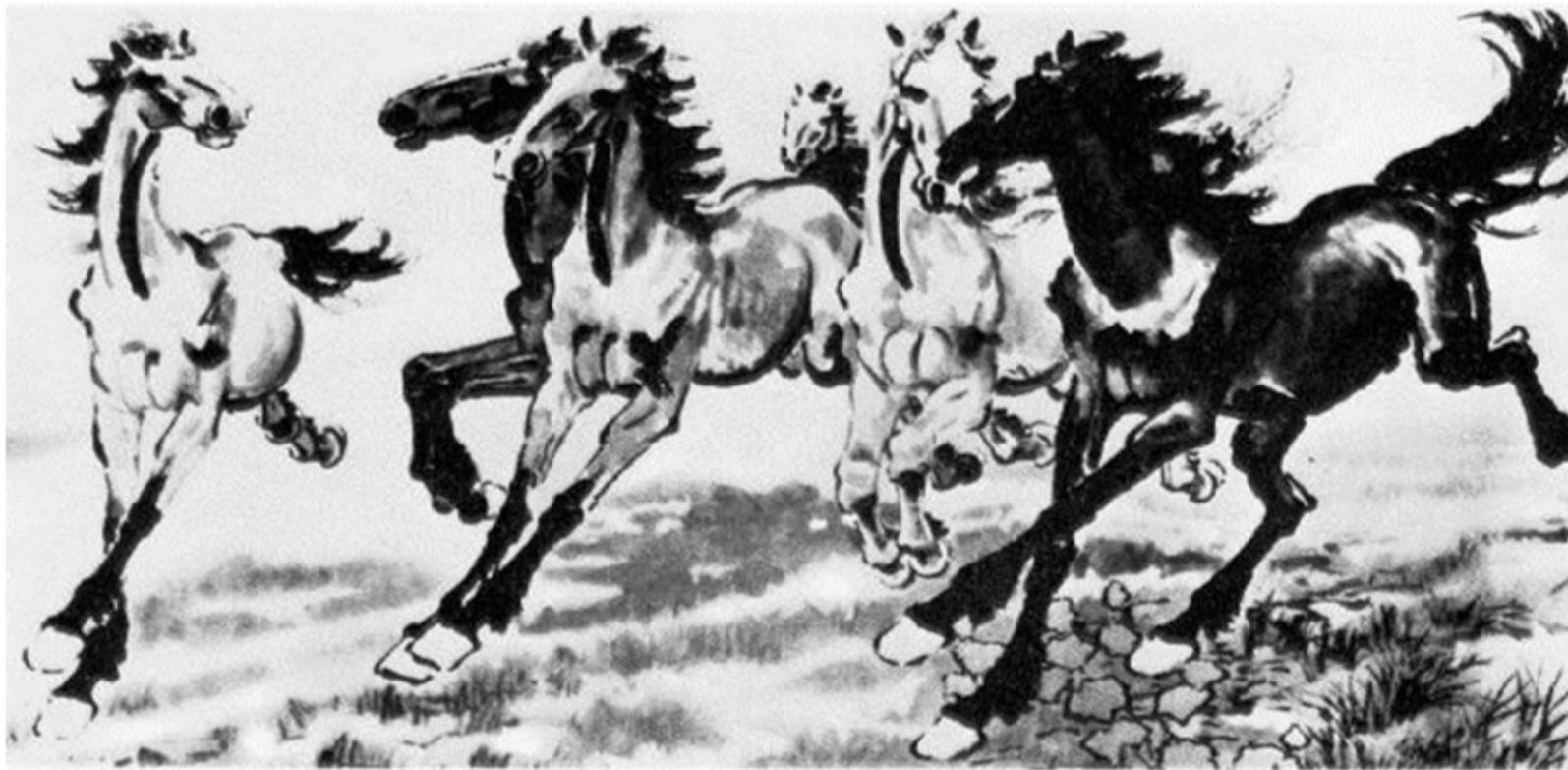
Size



Lightening



Deformation

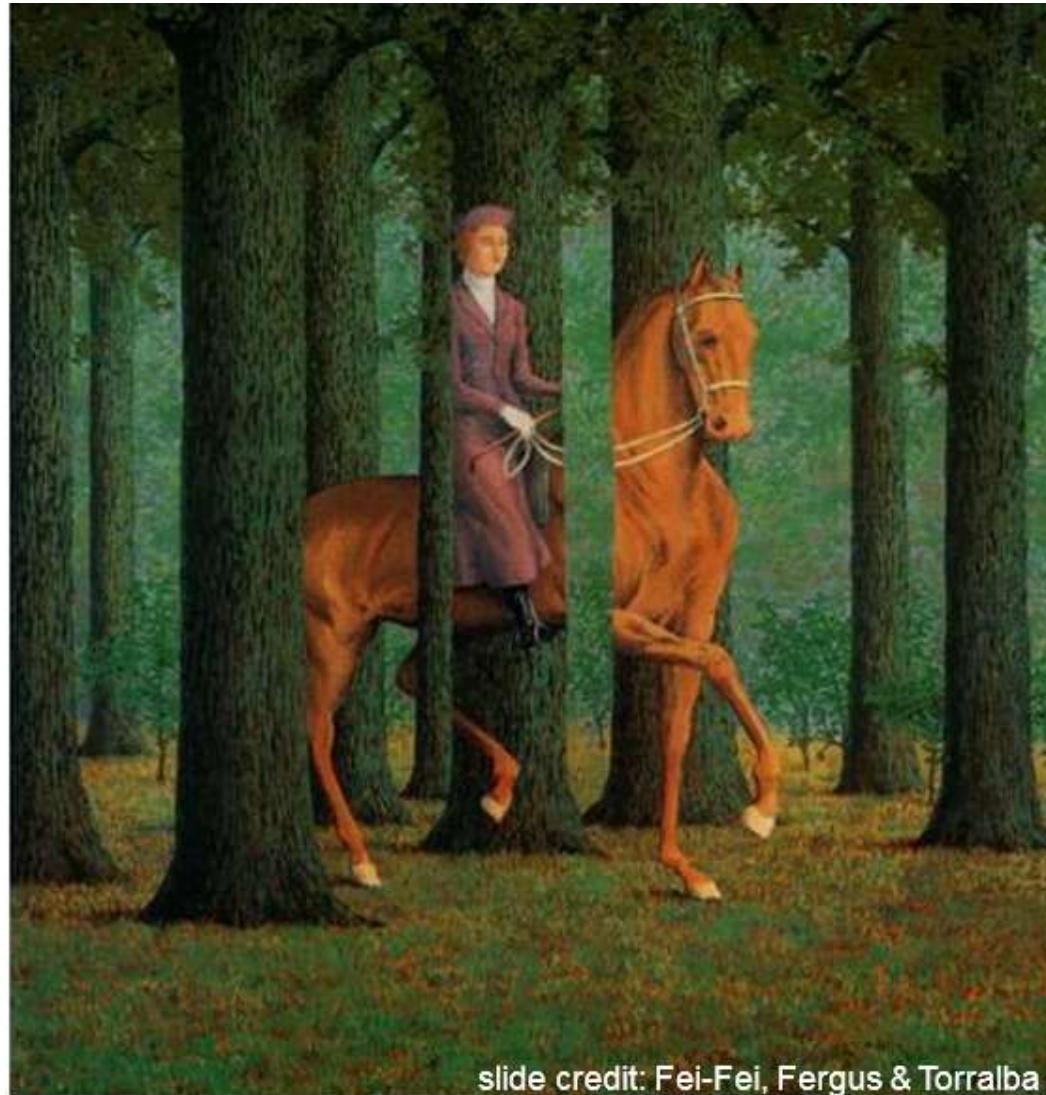


Xu, Beihong 1943

Slide credit: Fei-Fei, Fergus & Torralba

Overlapping

Magritte, 1957



slide credit: Fei-Fei, Fergus & Torralba

Masking



©2007 National Geographic Society. All rights reserved.
Photograph by Tim Laman

NATIONAL
GEOGRAPHIC

© 2007 National Geographic Society. All rights reserved.

Motion



Interclass variability



Slide credit: Fei-Fei, Fergus & Torralba

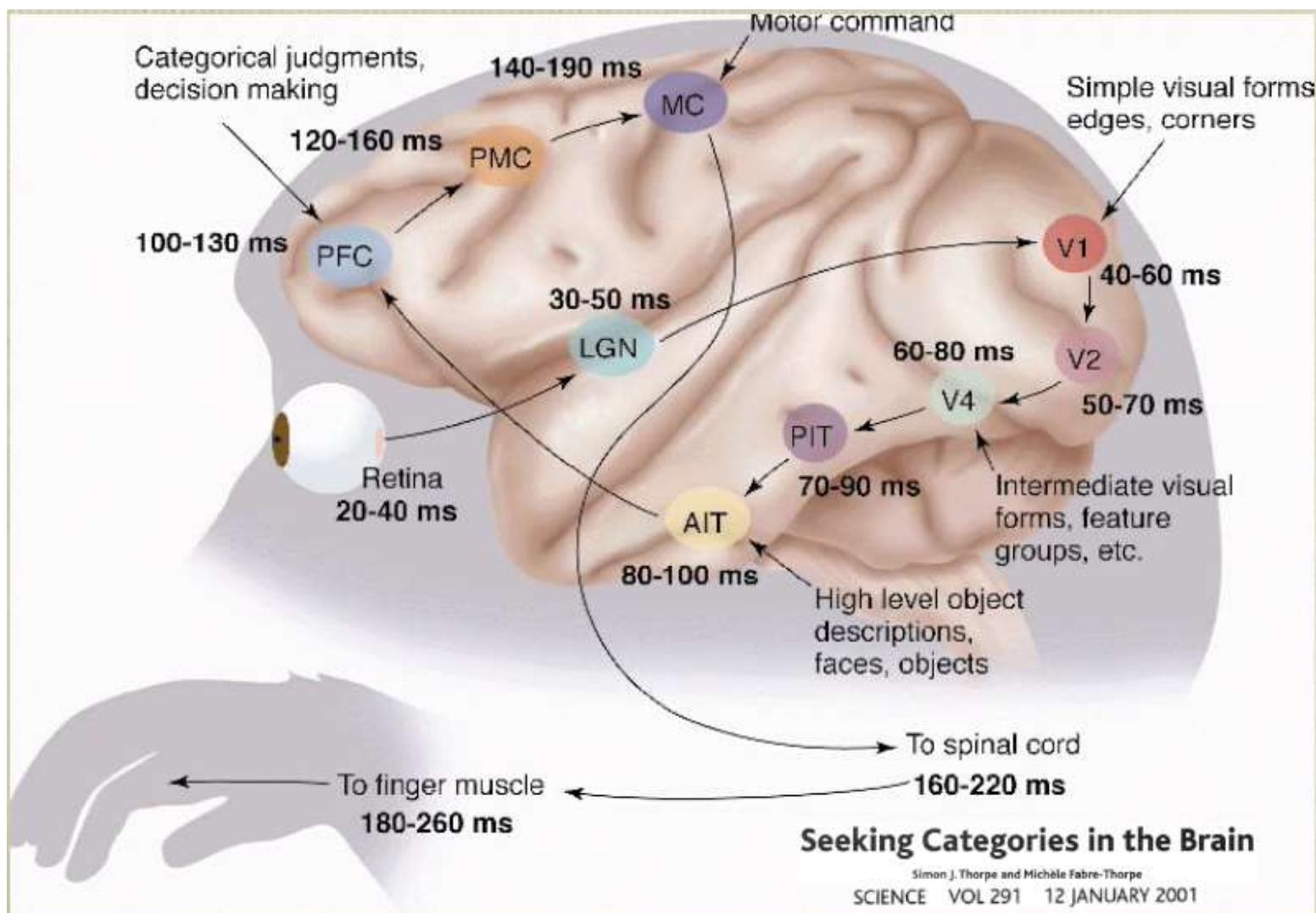
Perspective distortions: people



Hierarchy

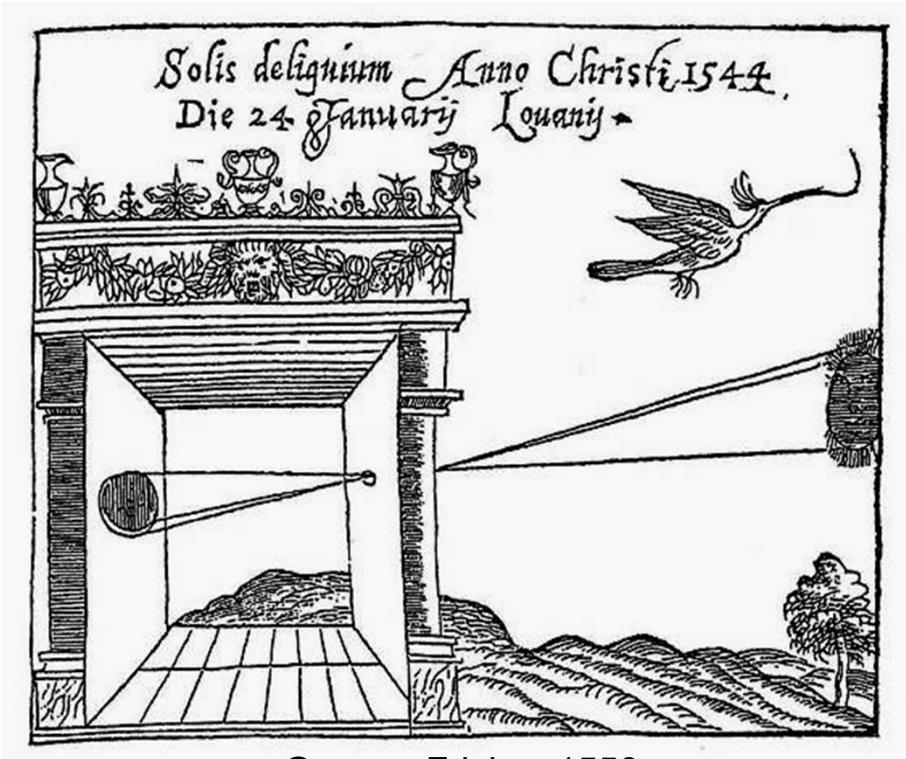
- Image processing: Input and Output are images
- Image analysis: Only 2D
- Pattern recognition: Identification and classification
- Computer vision: 3D from 2D
- Photogrammetry: Distances by 2D images
- Machine vision: industrial tasks

Human vision



- Let's build the model of an image

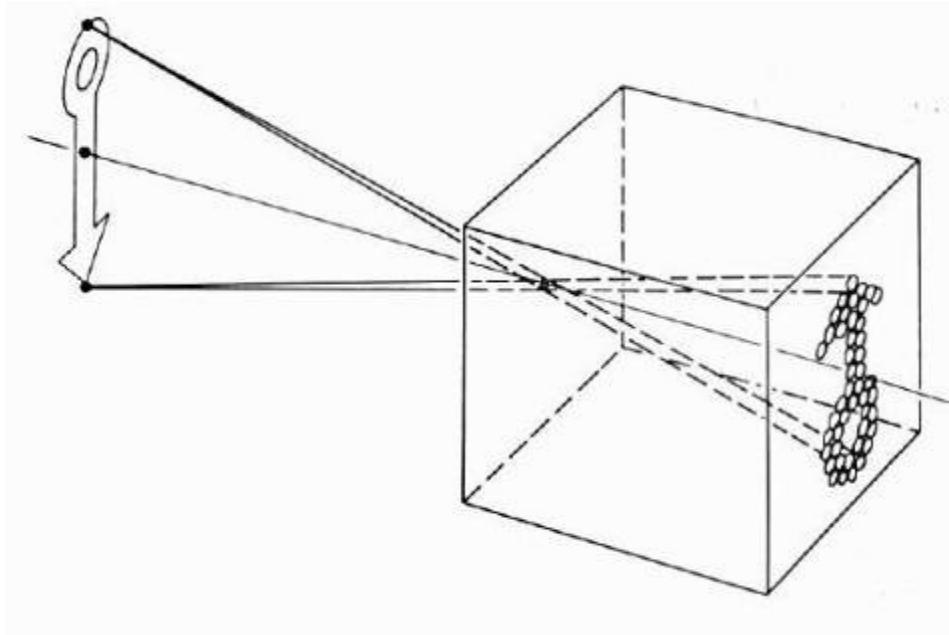
Camera-obscura



Gemma Frisius, 1558

- The principle was known to Aristotle (384-322 BC)
- Helped artists: described by Leonardo da Vinci (1452-1519)

The model of Camera-obscura

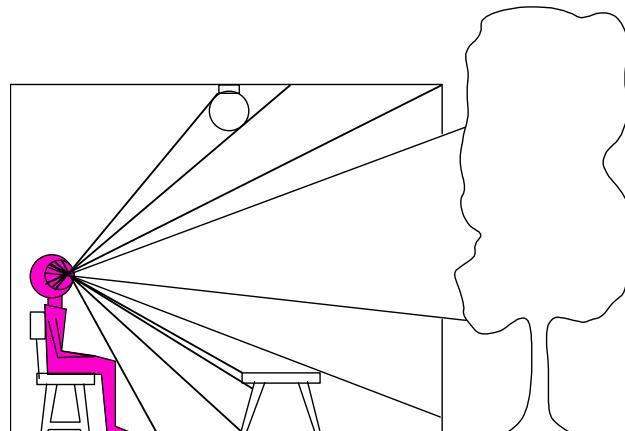


Camera-obscura:

- Captures a beam of rays passing through one point
- The point is called the center of projection (focal point / focal point)
- The image is formed on the picture plane (Image plane)

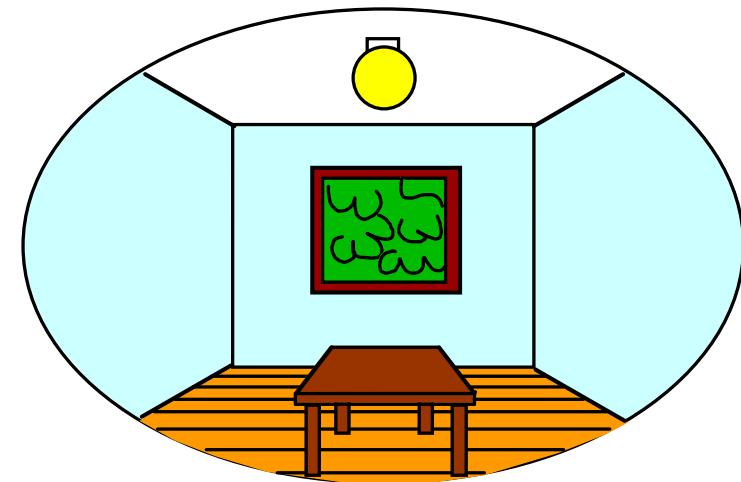
Dimension Reduction Machine

3D world



Point of observation

2D picture



What do we lose?
Angles
Distances and lengths

Perspective distortions

- The extreme columns seem thicker
- These distortions are caused not by lens error!
- The problem was noted by da Vinci

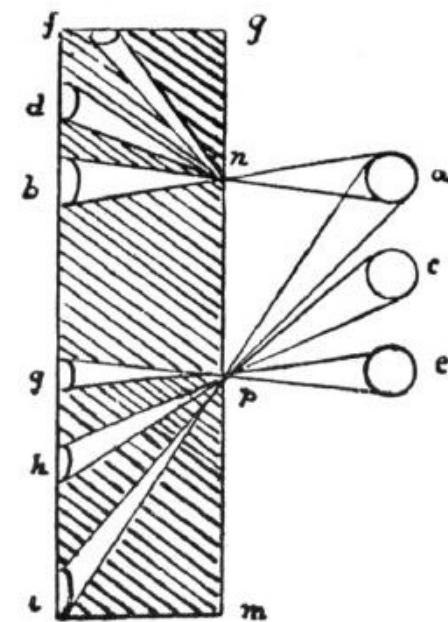
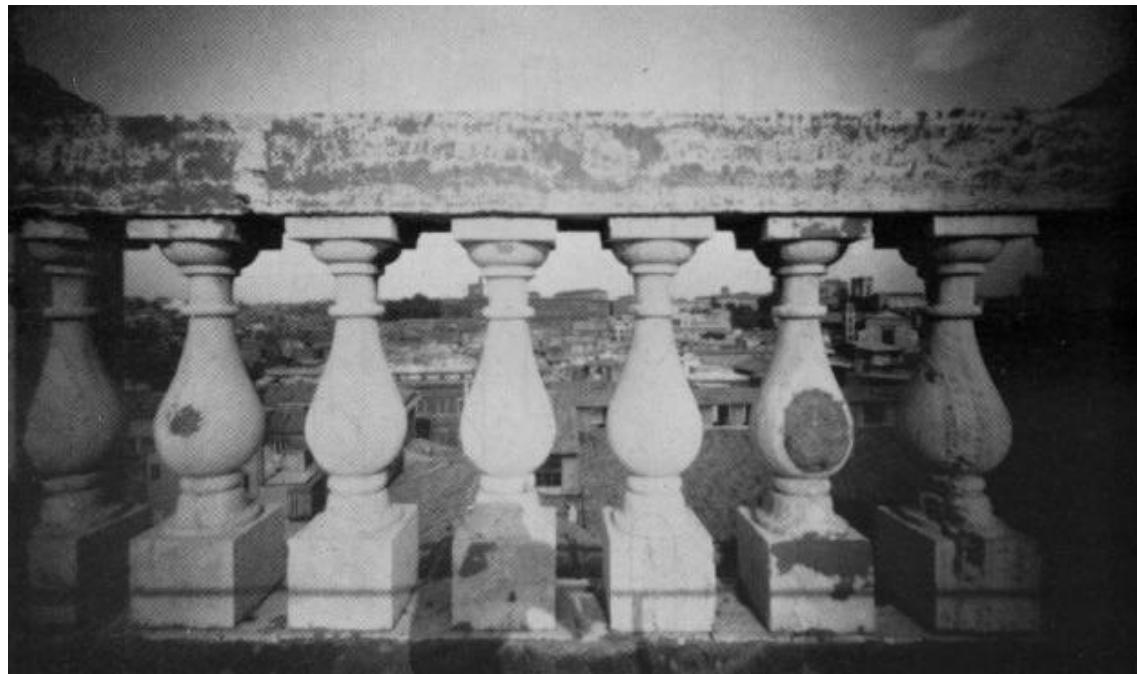
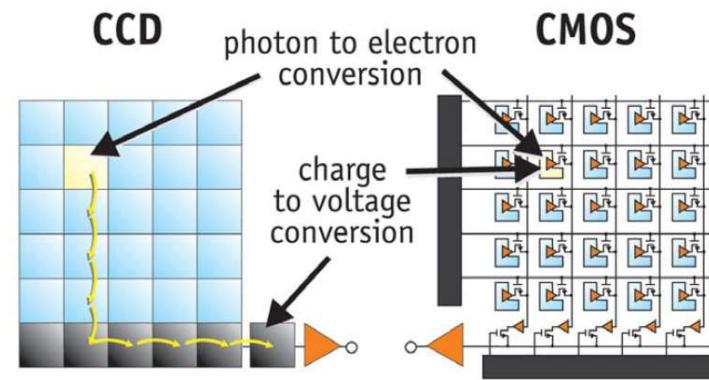


Photo camera

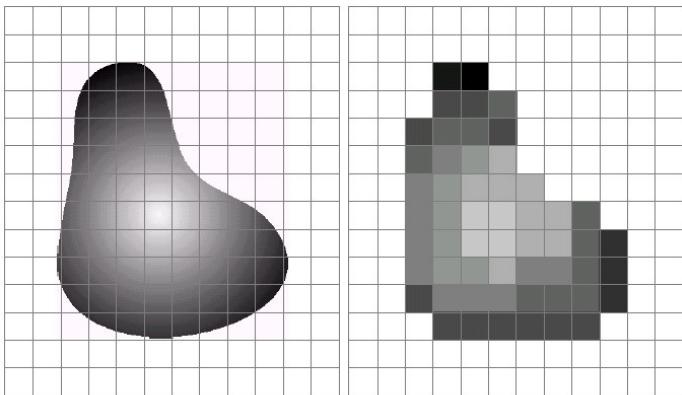


- The same mechanism as a camera-obscura
- The rays are focused with a lens
- The image is formed on the photosensitive film

Digital camera



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.



a

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

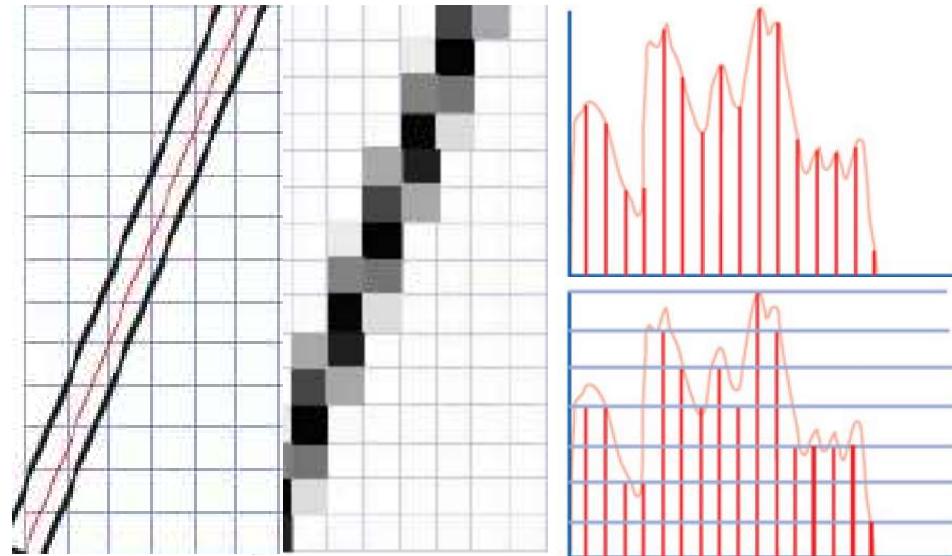


Image (def.)

The optical image is a picture obtained as a result of the passage through the optical system of rays propagating from the object and reproducing its contours and details.

Physical encyclopedic dictionary.

The intensity (brightness) function of the channel,
given on a 2-dimensional grid (matrix)

$$I = g(x, y), \{x \in [x_0, x_1], y \in [y_0, y_1]\}$$

A discrete representation is used

$$I = g(i, j), \{i = \overline{1, n}, j \in \overline{1, m}\}$$

10 events in the history of photography

<http://listverse.com/history/top-10-incredible-early-firsts-in-photography/>



1957 year, 176*176 pixels

Colored photos??



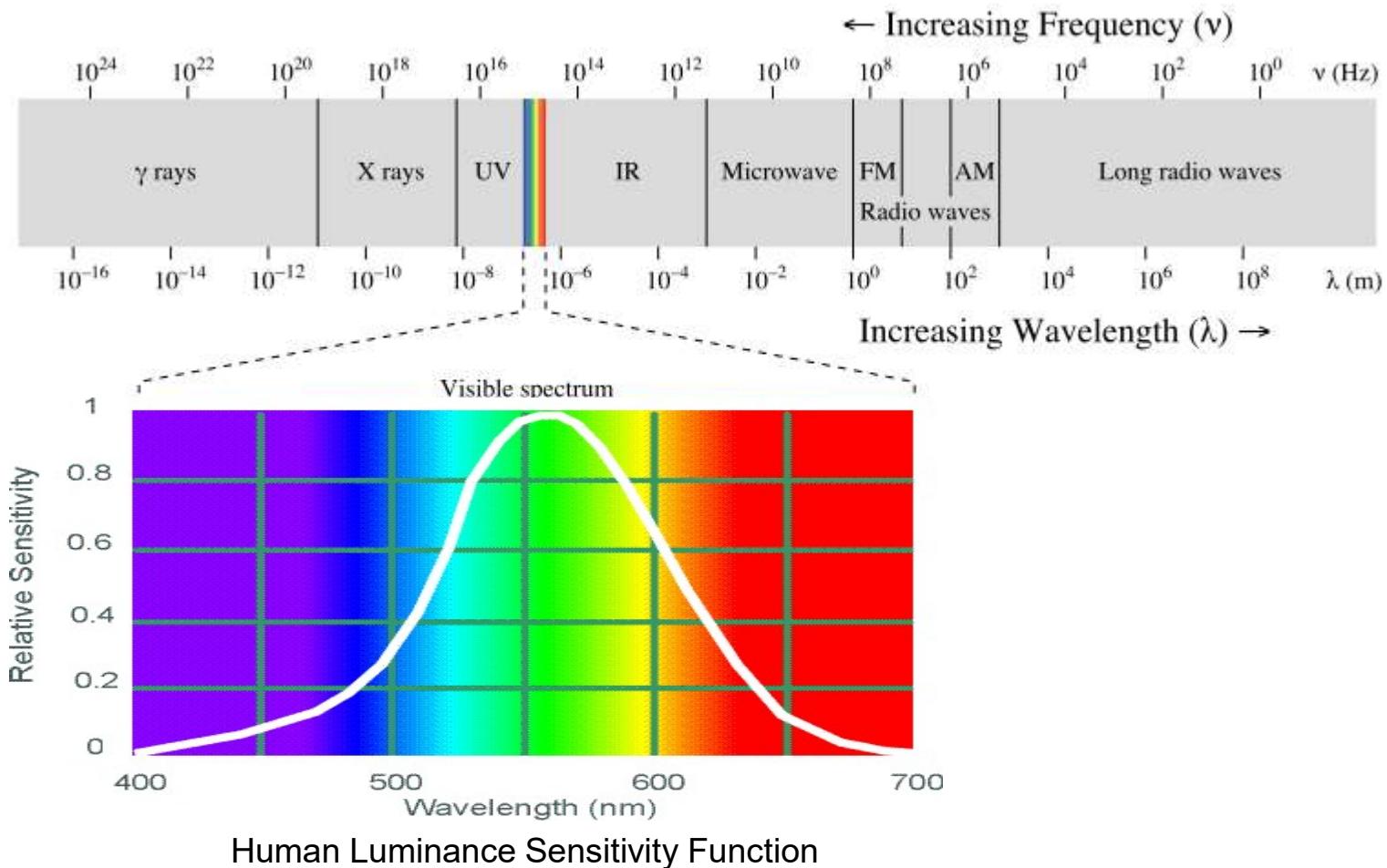
So what is a color?



Wassily Kandinsky (1866-1944), Murnau Street with Women, 1908

- Color is psychological feature of our vision, emerging during the observation of objects and light, (S. Palmer, *Vision Science: Photons to Phenomenology*)
- Color is the result of the interaction of light, scene and our visual system

Electromagnetic spectrum



Why do we see light in this range?
Because it is such a range of solar radiation

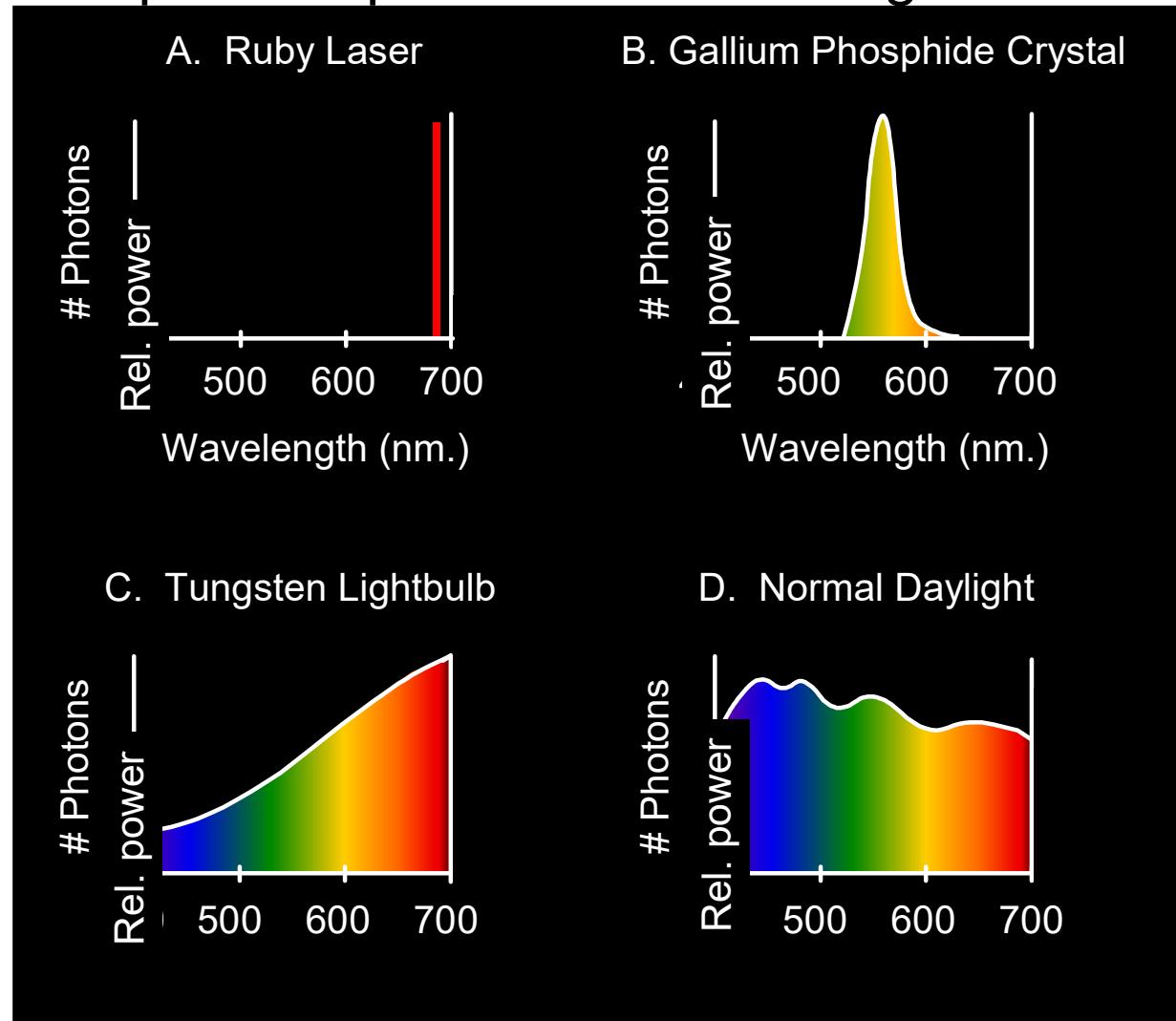
The physics of light

Any light source can be fully described by the spectrum: the amount of radiated energy per unit time for each wavelength in the interval 400-700 nm.



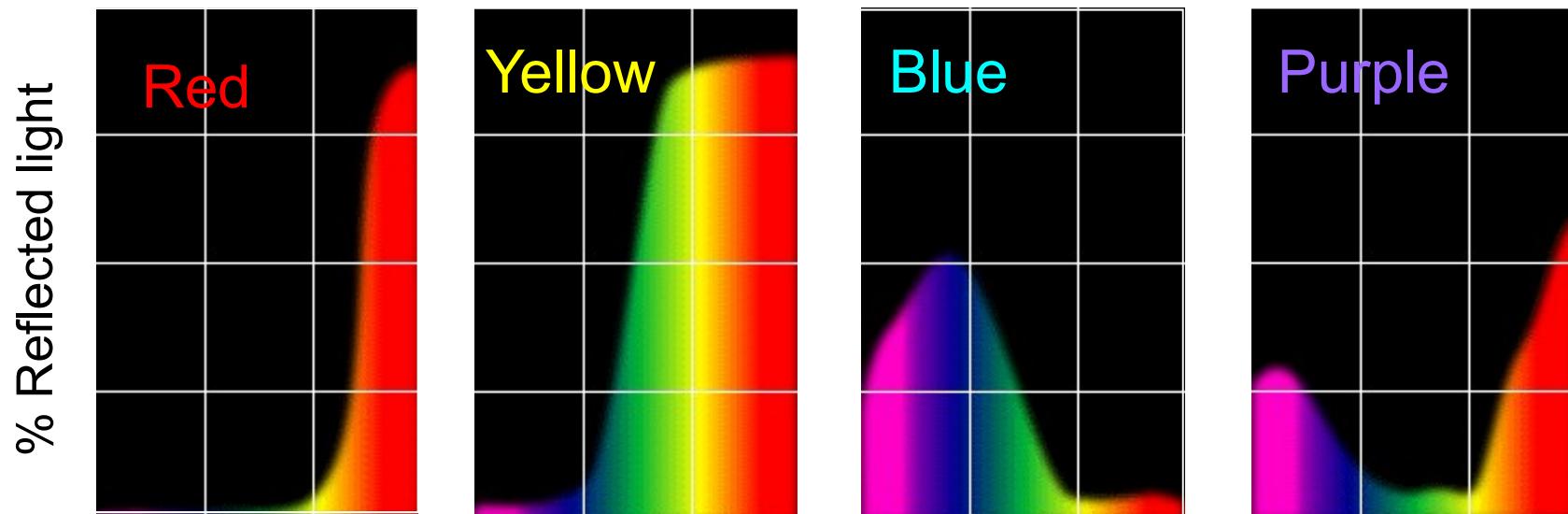
The physics of light

Examples of spectra of different light sources



The physics of light

The examples of reflected light

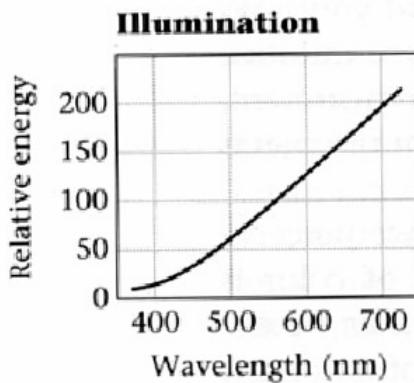


The length of wave (nm)

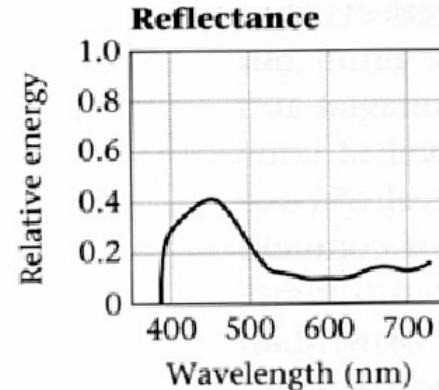
Interactions of light and objects



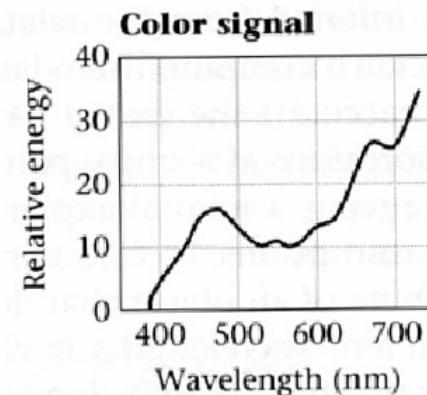
The visible color is the result of interaction of the emitted light specter and surface



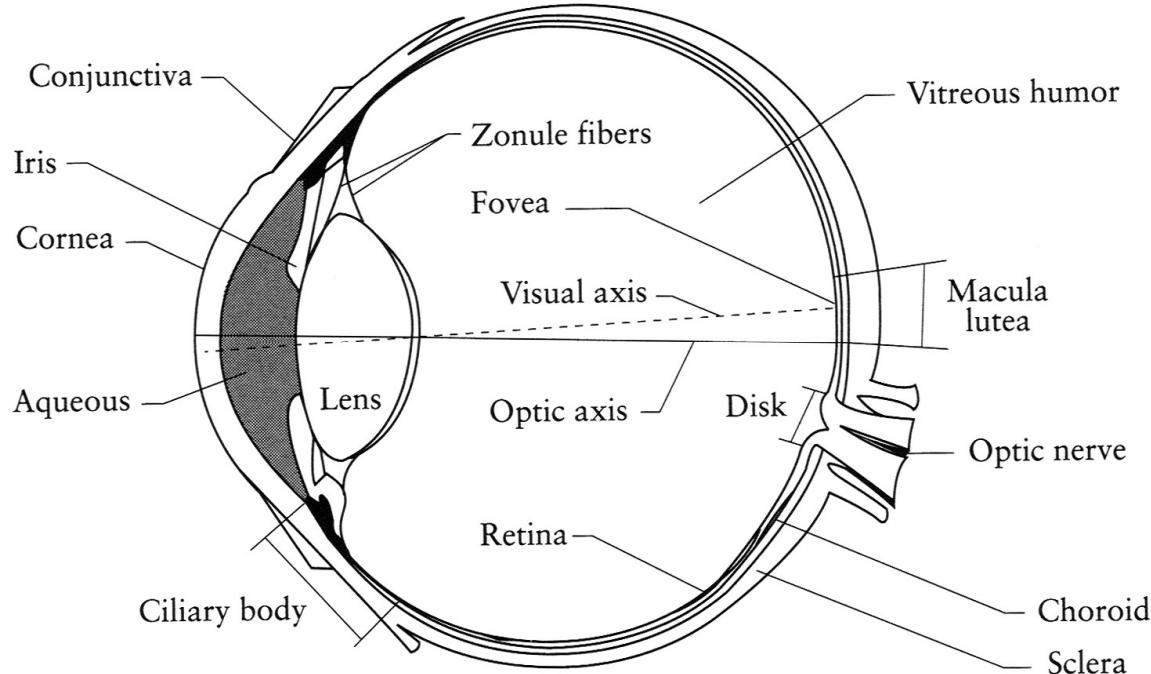
• *



||



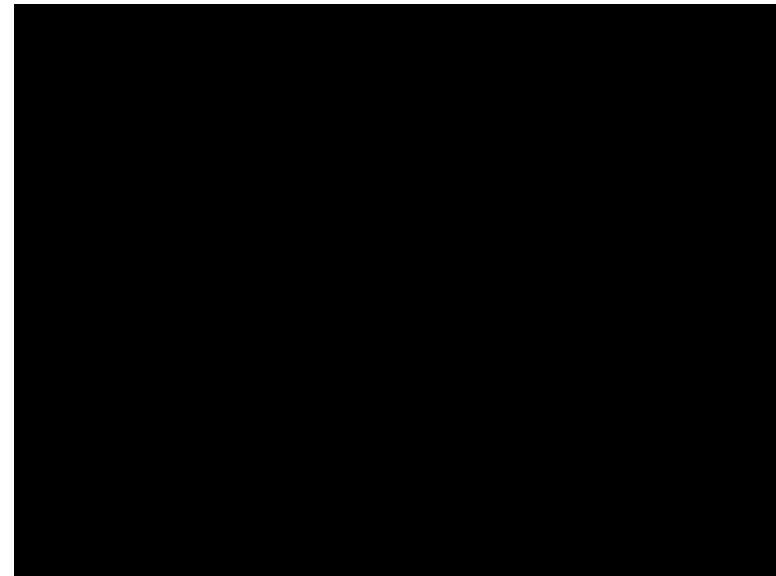
Human eye



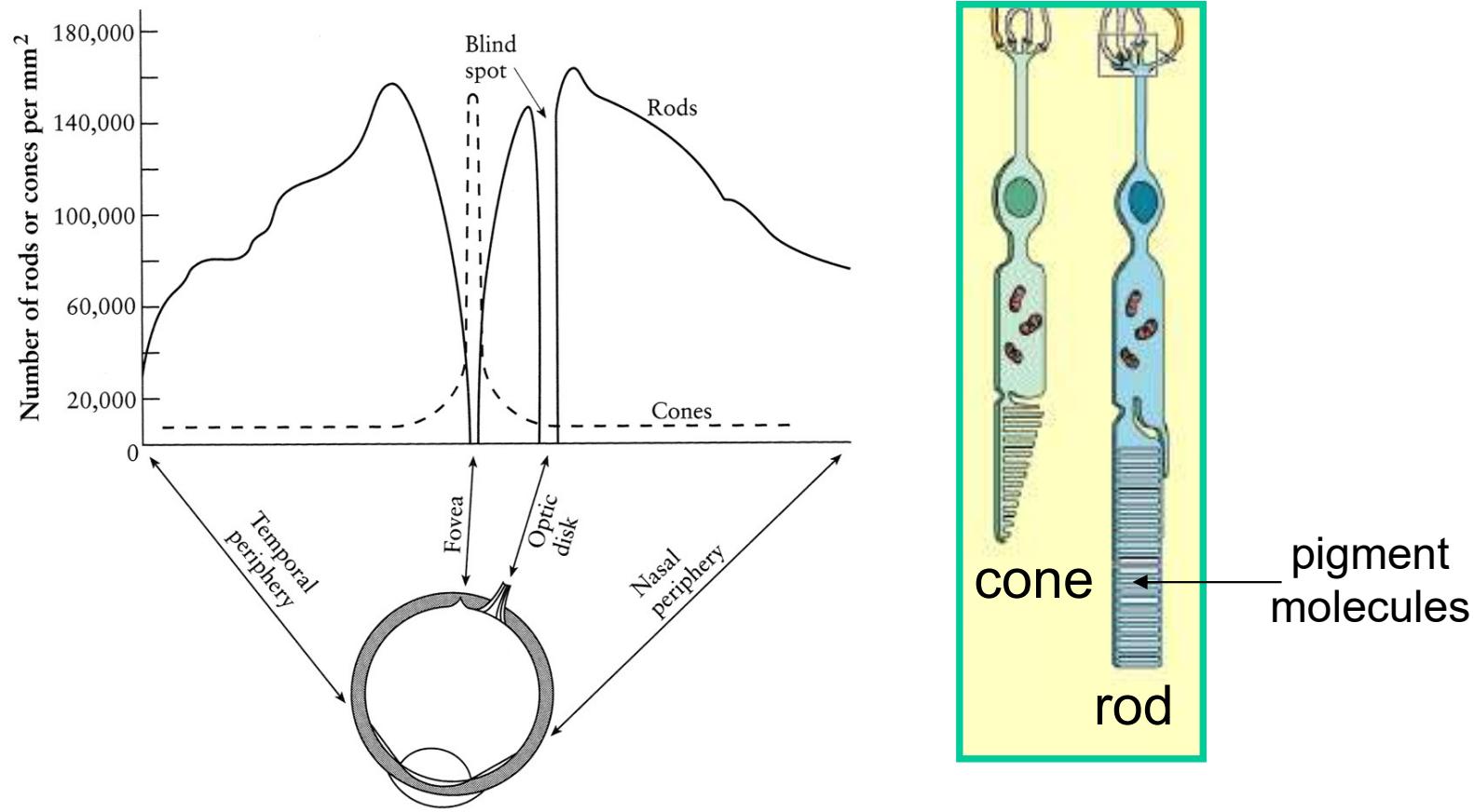
The eye is like photocamera!

- **Радужка (Iris)** – colored film with radial muscles
- **Зрачок(pupil)** - hole (аппертура), diameter controlled by iris
- **Линза(lens)** – changes its form with muscles
- Where is matrix?
 - Cells-photoreceptors on the retina

Тест на внимательность

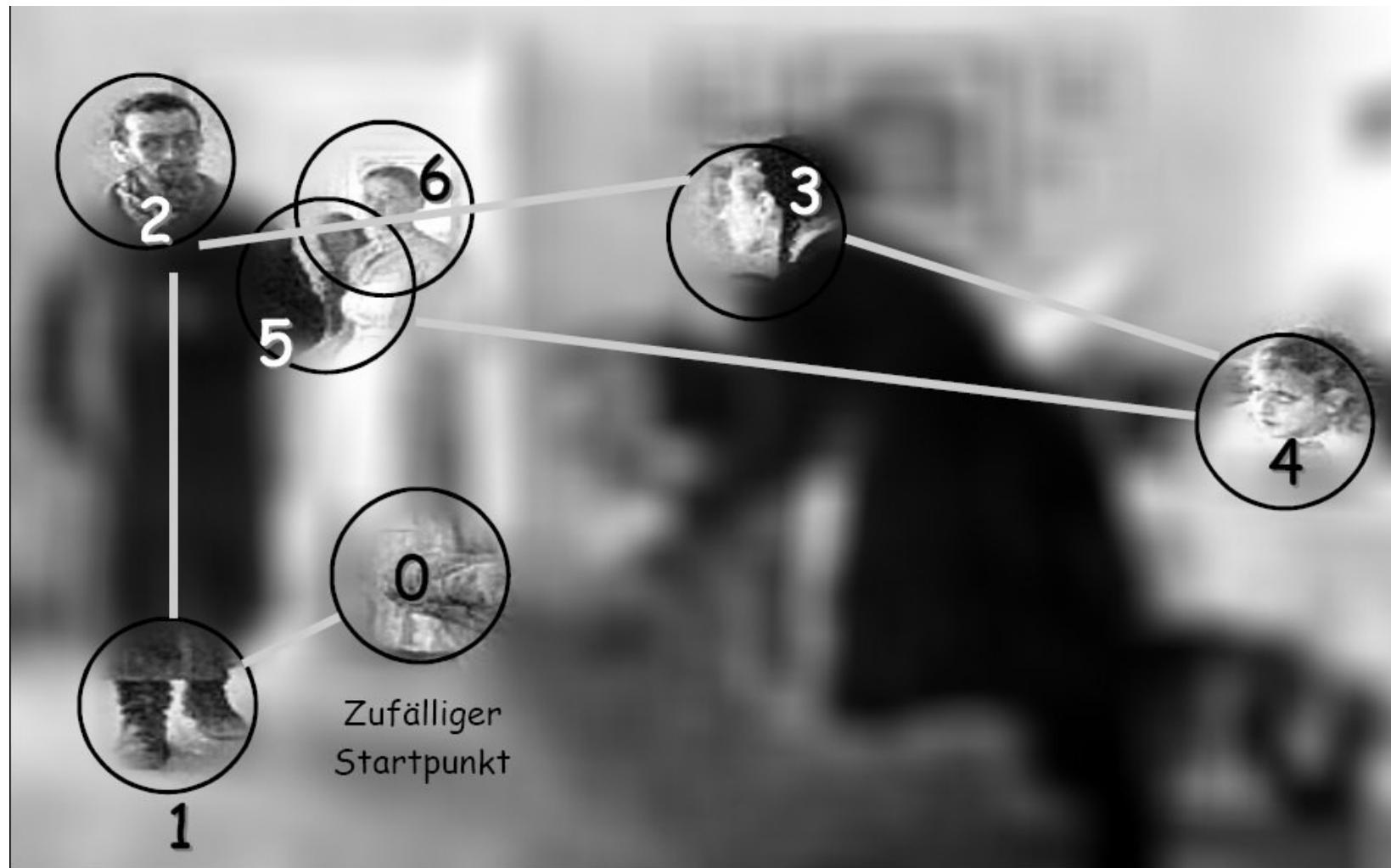


Density of rods(палочек) and cones (колбочек)



Rods and cones are distributed unevenly

What we actually see



Source: A. Yarbus

Eye movements



Free examination.



1
Estimate material circumstances
of the family



2
Give the ages of the people.



3
Surmise what the family had
been doing before the arrival
of the unexpected visitor.



4
Remember the clothes
worn by the people.



5
6
Remember positions of people and
objects in the room.

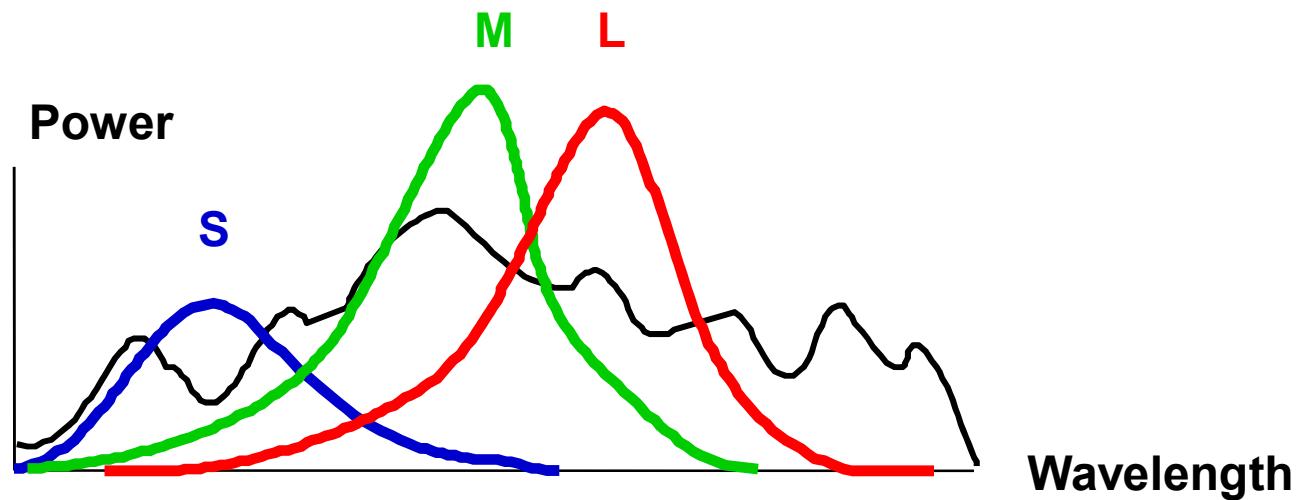


7
Estimate how long the visitor had
been away from the family.

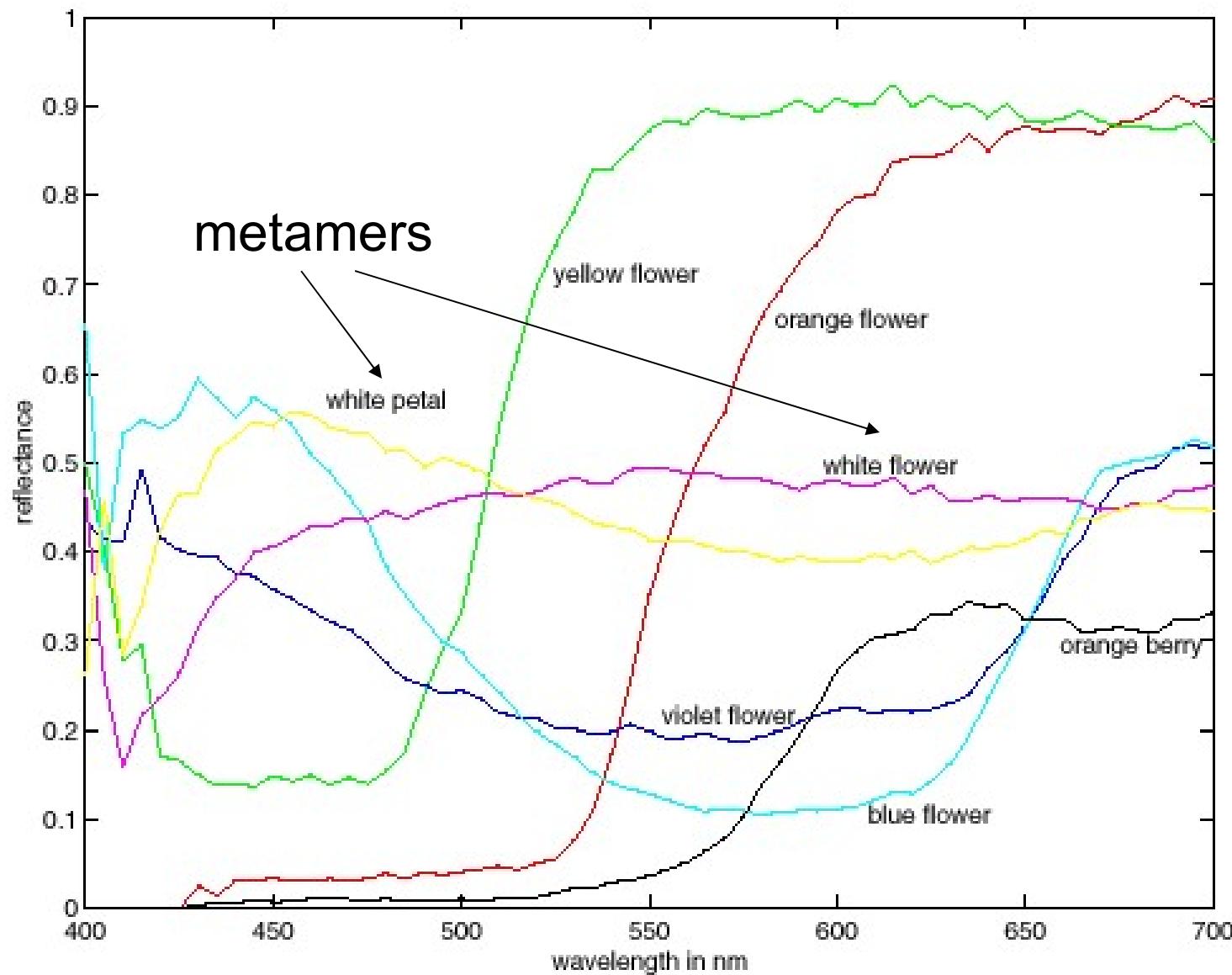
3 min. recordings
of the same
subject

Source: A. Yarbus

Color reception



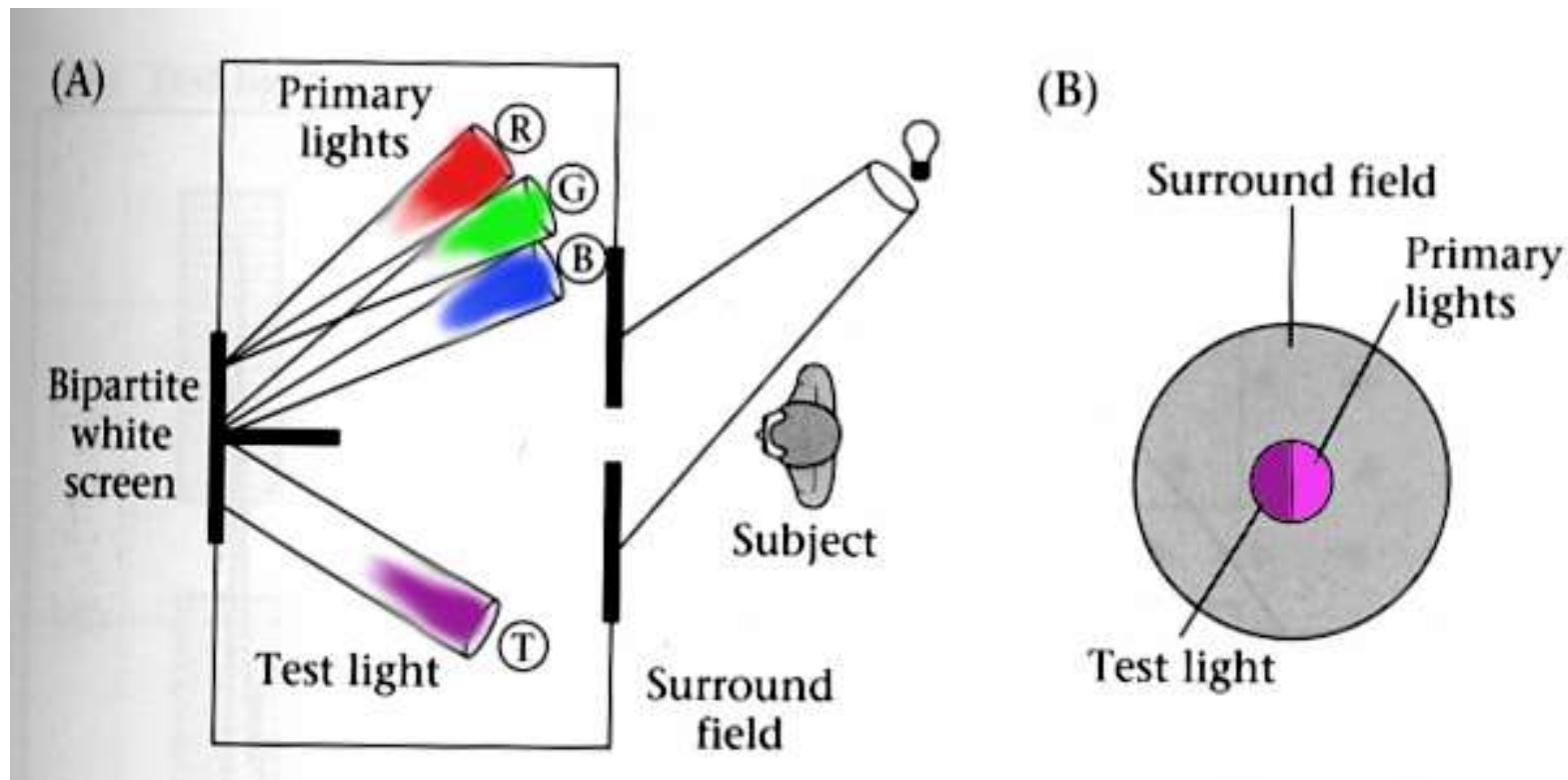
Spectra of some objects



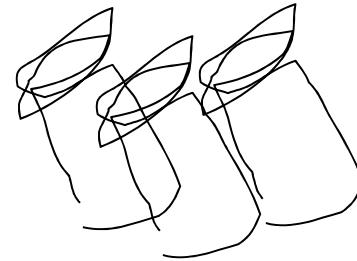
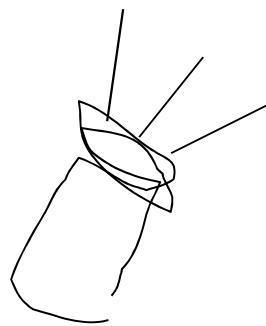
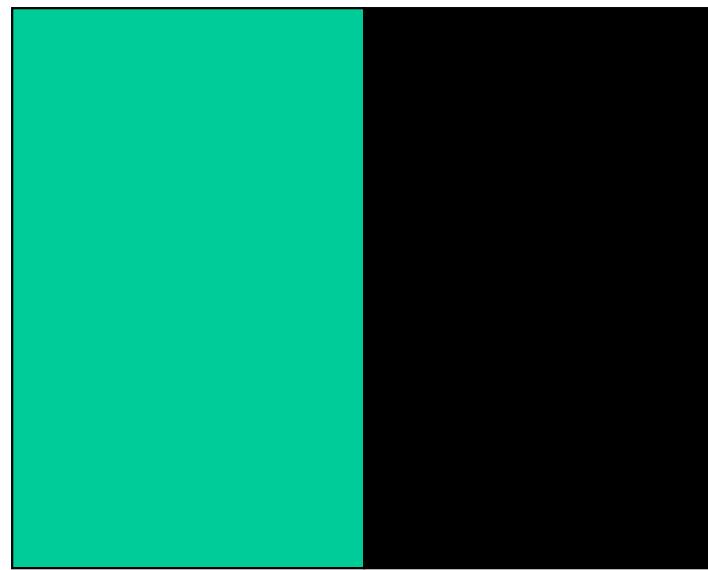
Standardization of color perception

We want to understand which spectra of light cause the same color sensations in people

Color Matching Experiments

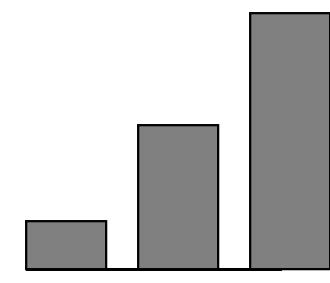
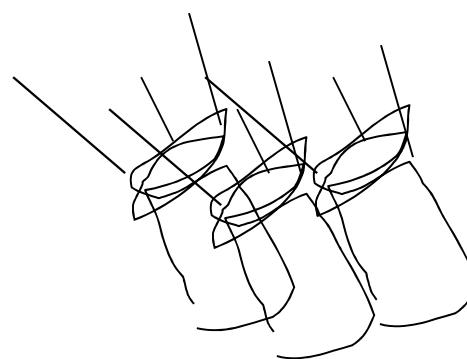
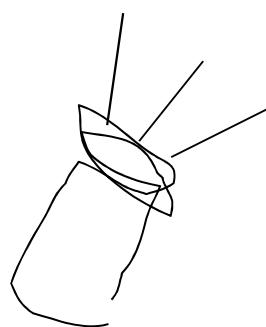
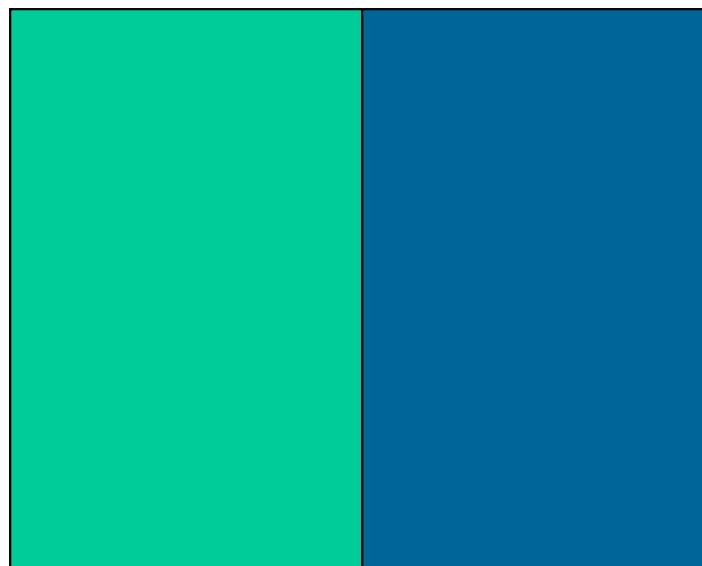


Experiment №1



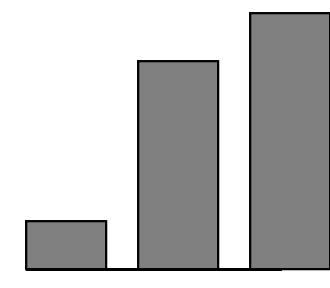
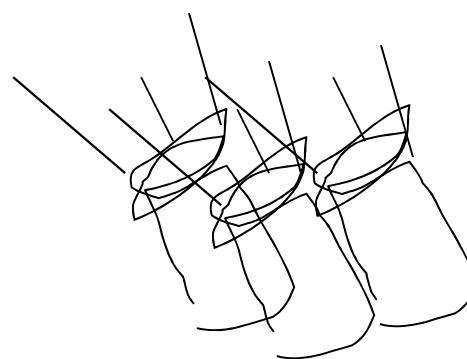
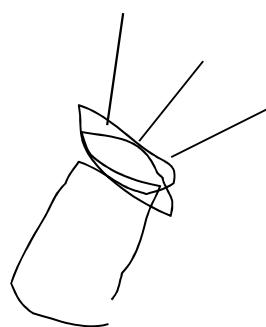
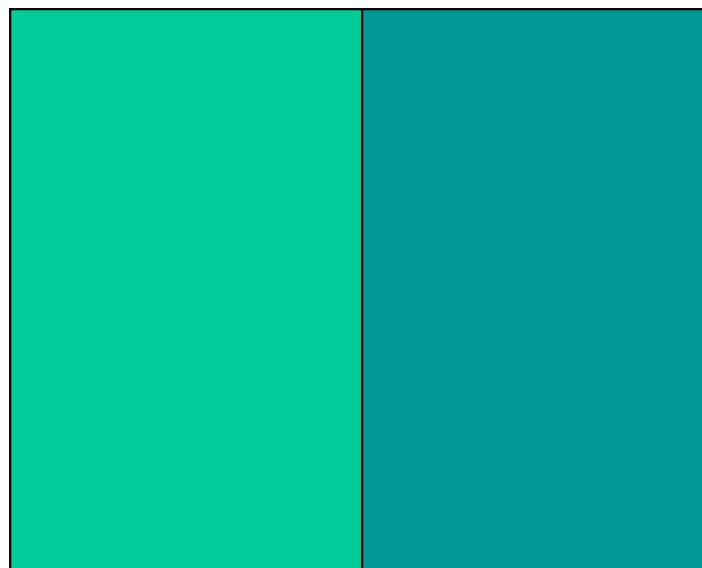
Source: W. Freeman

Experiment №1



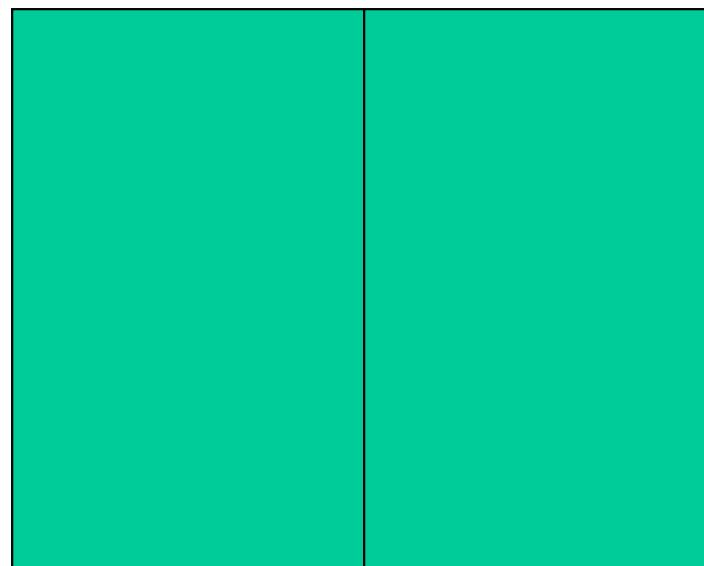
Source: W. Freeman

Experiment №1

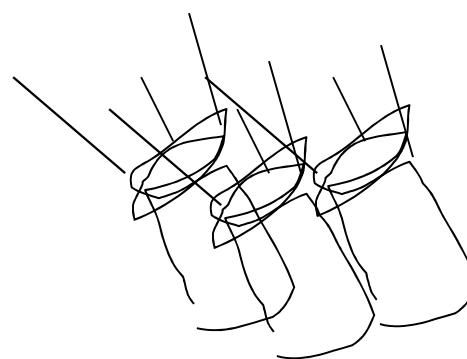
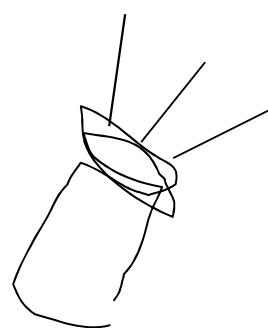
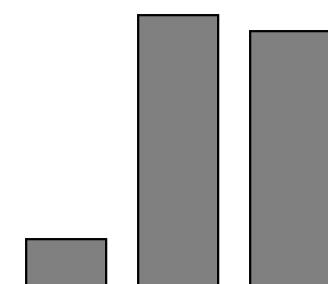


Source: W. Freeman

Experiment №1

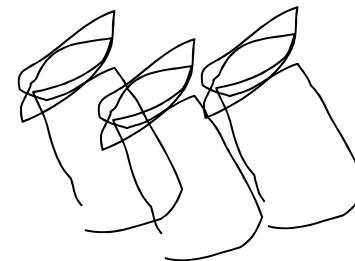
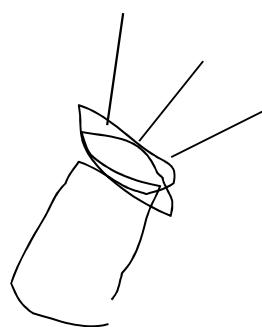


Basic colors



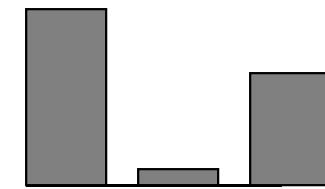
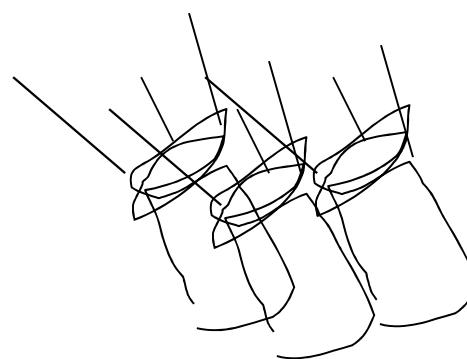
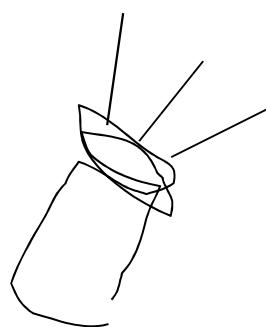
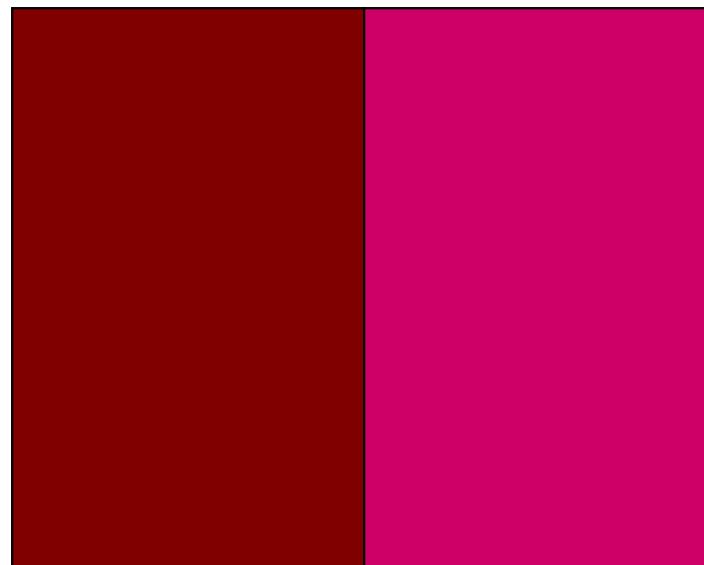
p_1 p_2 p_3
Source: W. Freeman

Experiment №2



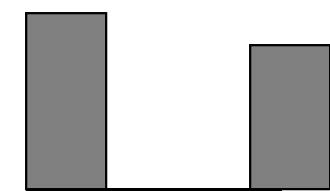
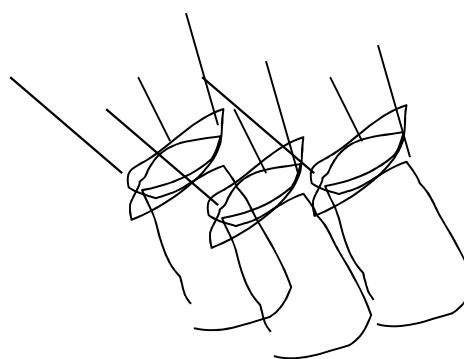
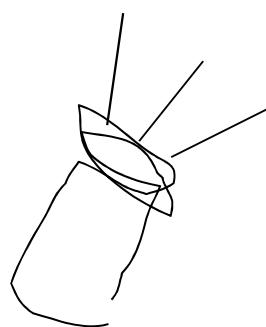
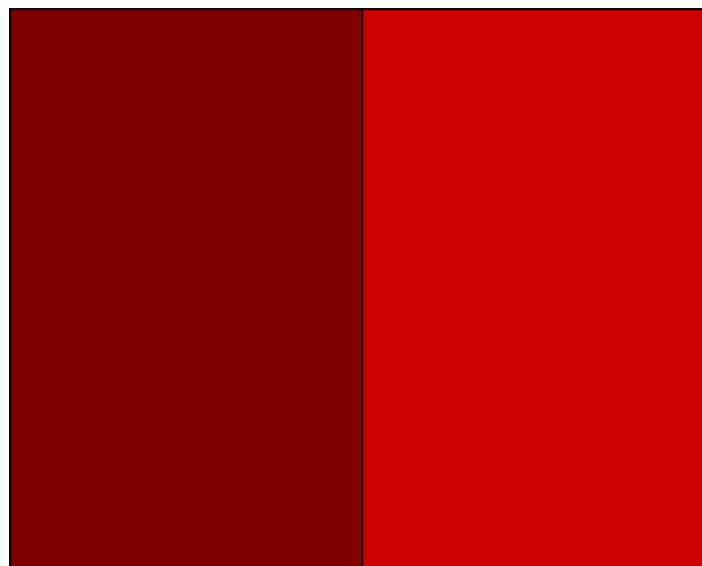
Source: W. Freeman

Experiment №2



Source: W. Freeman

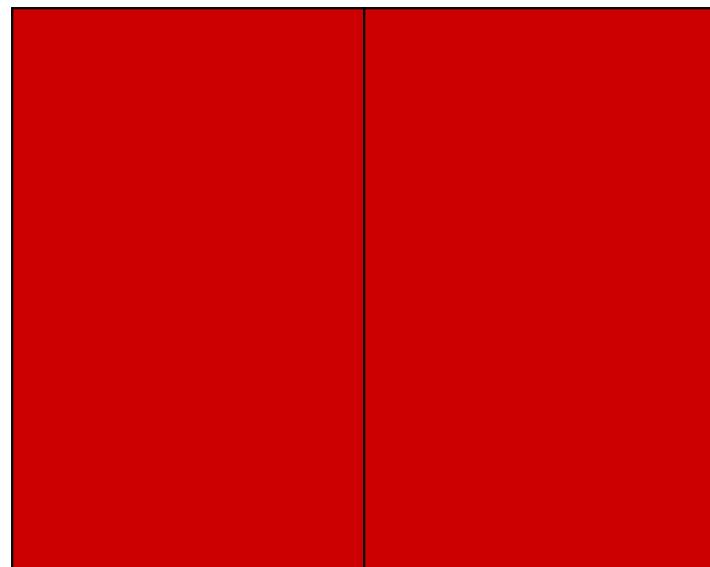
Experiment №2



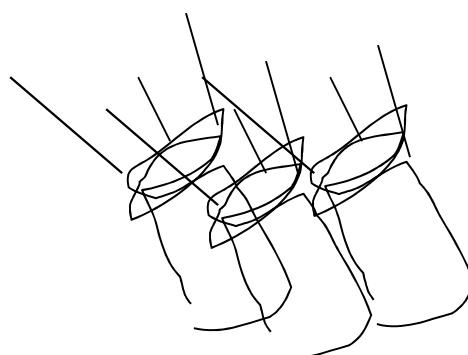
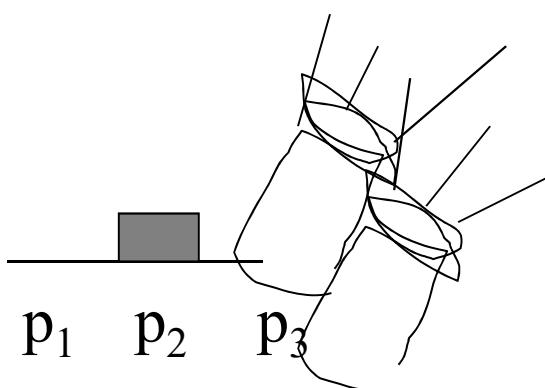
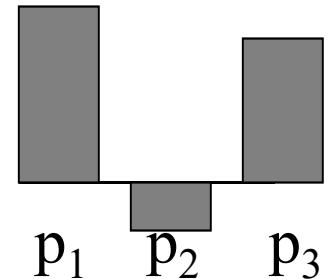
Source: W. Freeman

Experiment №2

We call m the
"negative" weight
of the base color, if
the color should be
added to the light
to be compared.



Basic colors



Source: W. Freeman

Trichromatic theory

In experiments on color matching, most people have 3 primary colors to match any color

Primary colors should be independent

For the same spectrum, and the same basic colors, people choose the same weights

Exceptions: color blindness

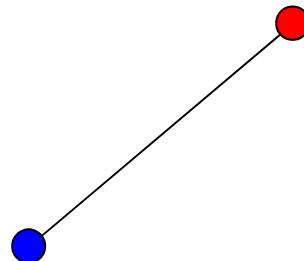
Trichromatic theory

Three numbers are enough to describe the color

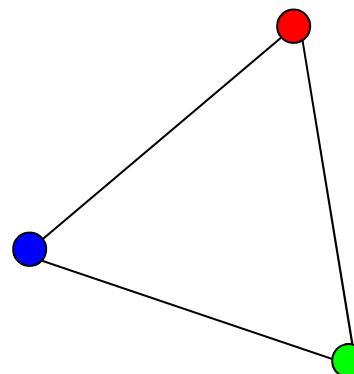
History dates back to the 18th century (Thomas Jung)

Linear color spaces

- Determined by the choice of 3 primary colors
- The "color coordinates" are specified by the weights of the primary colors required for matching
- Each coordinate is encoded in 1-2 bytes
- Comparison functions: weights necessary for comparison with coherent light sources



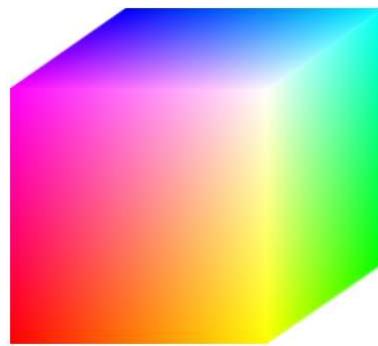
Merge 2 colors



Merge 3 colors

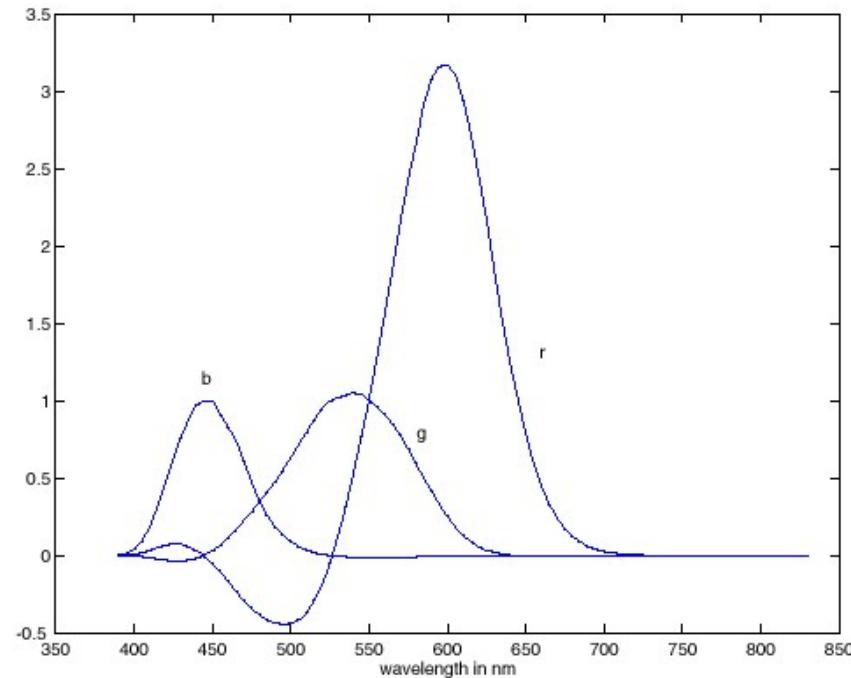
Linear color models: RGB

- The main colors are monochromatic (in the monitor they correspond to three kinds of phosphors)
- Subtraction is necessary to match certain wavelengths

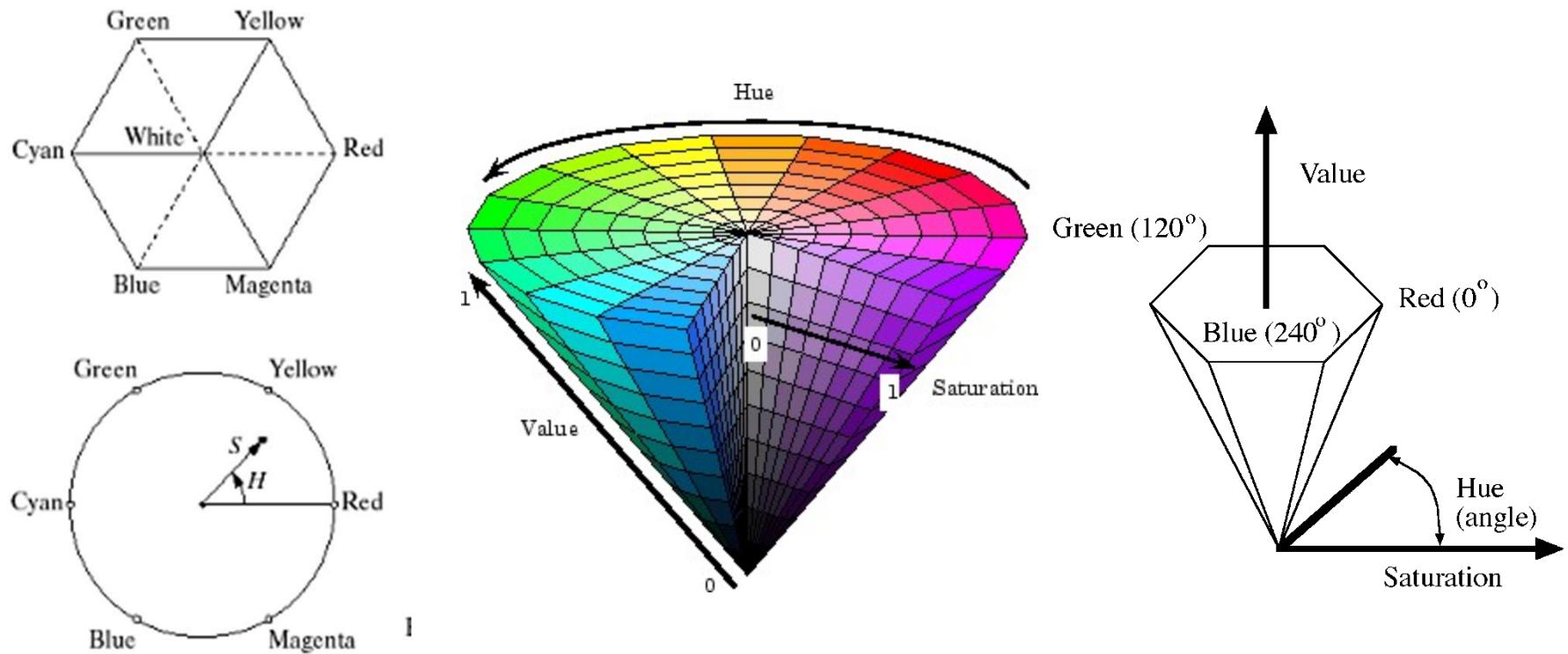


- $p_1 = 645.2 \text{ nm}$
- $p_2 = 525.3 \text{ nm}$
- $p_3 = 444.4 \text{ nm}$

Mapping functions for RGB



Model HSV



Coordinates choosed for human reception:
Hue (Тон), Saturation(Насыщенность), Value
(Intensity) (Интенсивность)

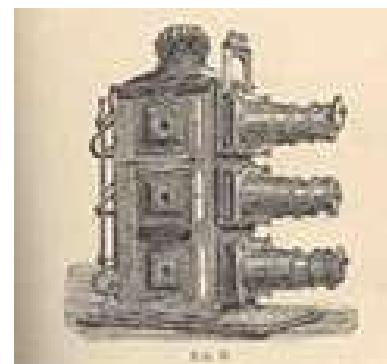
Why linear?

- Grassmann's Law:
 - There are 2 sources of lights, с цветом (R_1, G_1, B_1) и (R_2, G_2, B_2)
 - Then merge has the following color $(R_{1+2}, G_{1+2}, B_{1+2})$

First colored photos

Sergei Prokudin-Gorsky (1863-1944)

Photos of the Russian Empire (1909-1916)



Lantern
projector

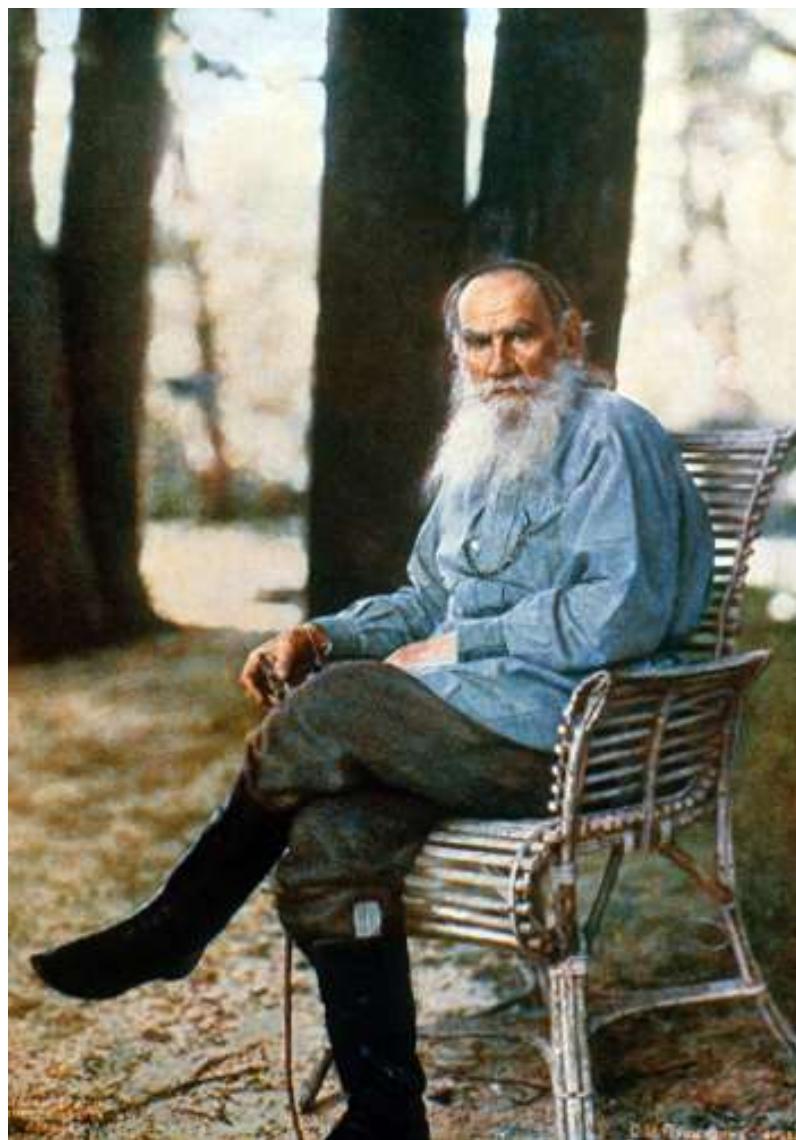


http://en.wikipedia.org/wiki/Sergei_Mikhailovich_Prokudin-Gorskii

<http://www.loc.gov/exhibits/empire/>

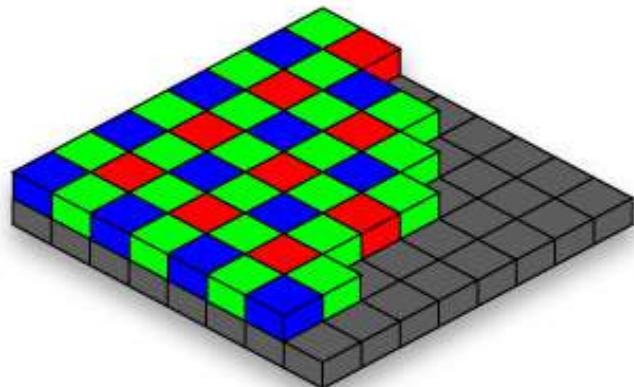
Slide by A.Efros

Лев Толстой

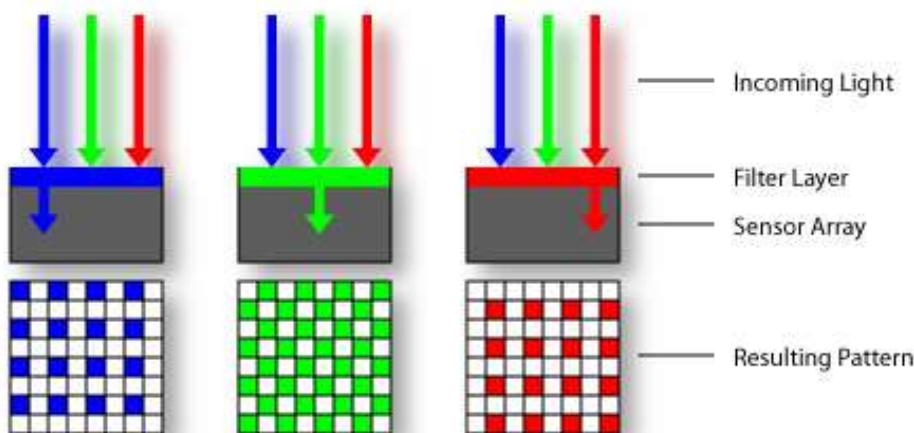
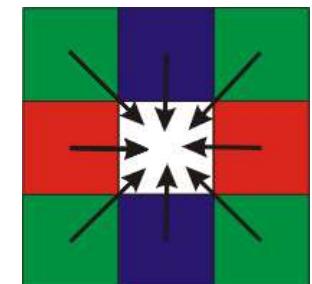


Color digital image

Bayer's template



Democating
(estimation of missing
color values)



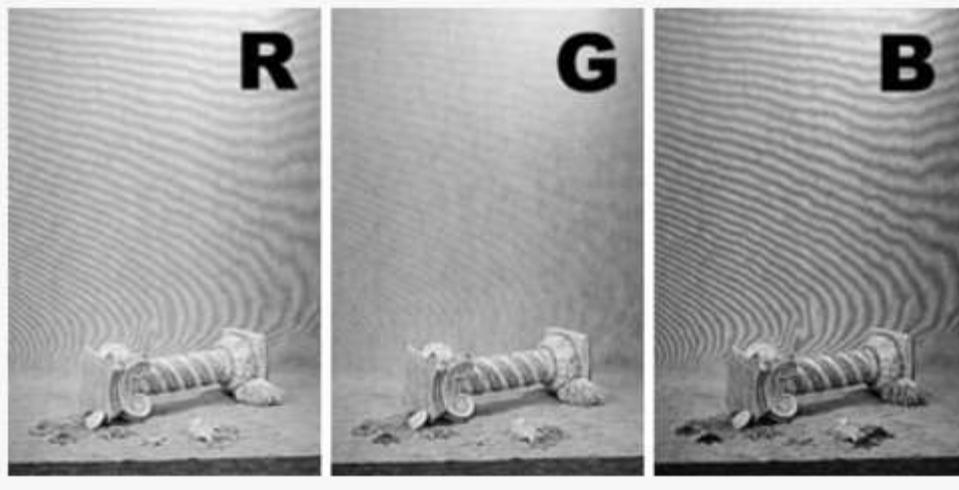
Source: Steve Seitz

Erase demo errors

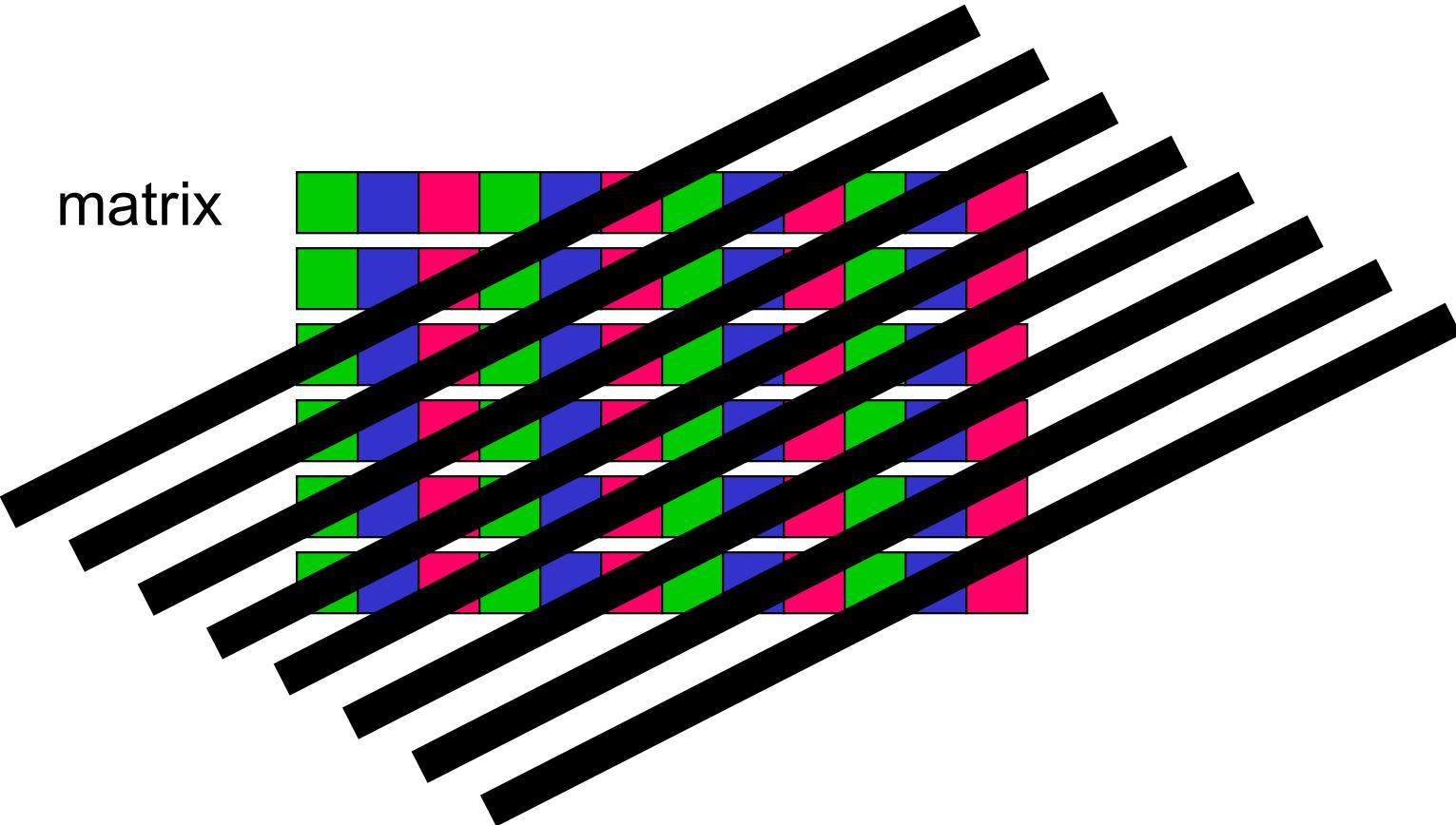


Рис. 1. Изображение с цветовым
муаром

Color moire



The reason



Thin black and white details
interpreted as color changes

Color reception

Constancy of color and illumination

- The ability of the human visual system to evaluate its own reflective properties of surfaces, regardless of the conditions of illumination

Instant effects

- Simultaneous contrast: background colors affect the perceived color of the object
- Mach bands

Step-by-Step Effects

- Adaptation to light / dark
- Chromatic adaptation
- Residual images (afterimages)

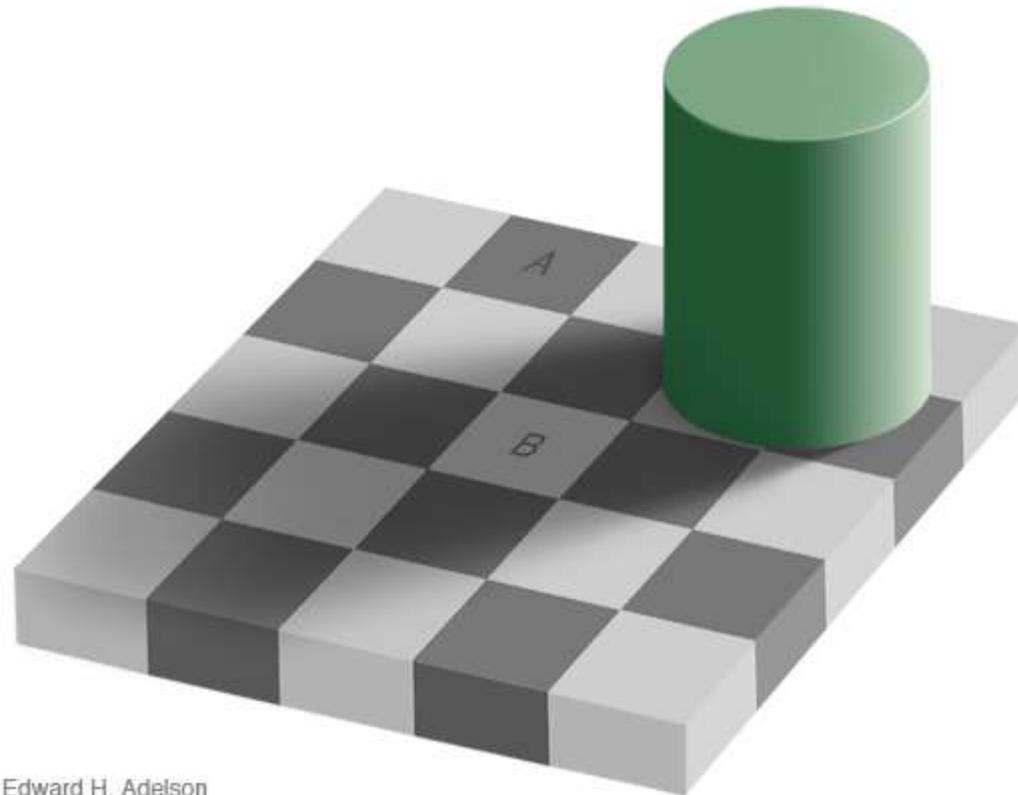
Constancy of illumination



J. S. Sargent, The Daughters of Edward D. Boit, 1882

White color
in shadow
and light

Constancy of intensity

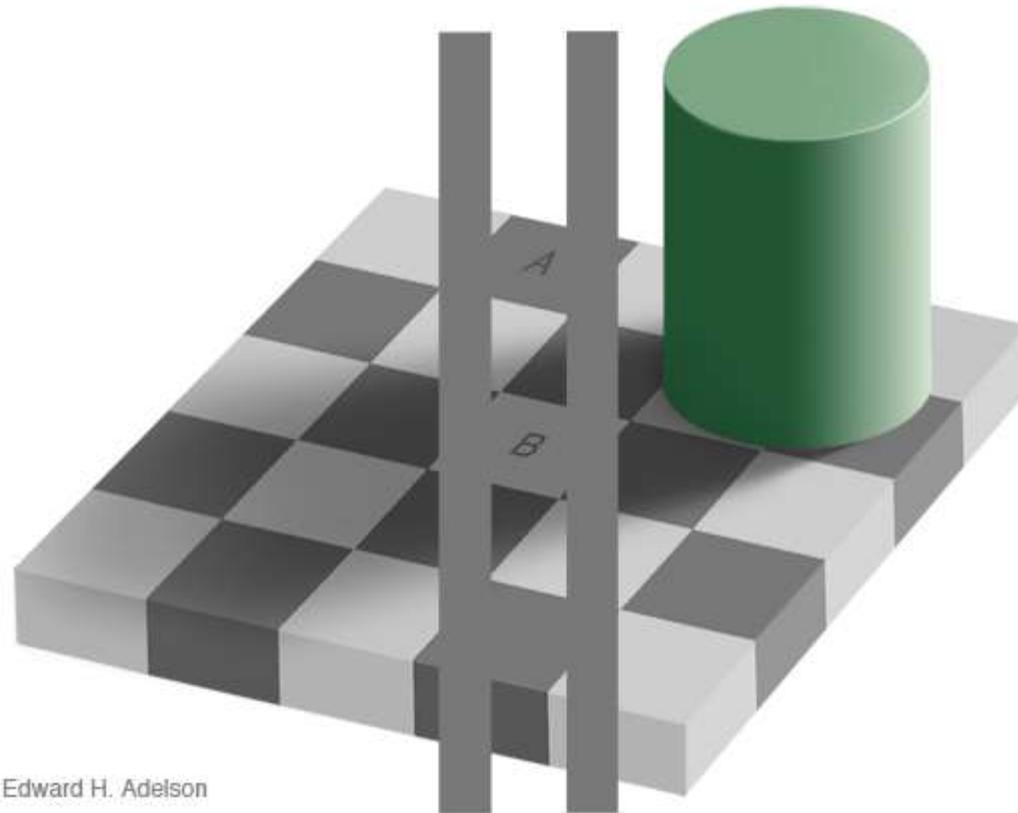


Edward H. Adelson

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

Slide by S. Lazebnik

Constancy of intensity

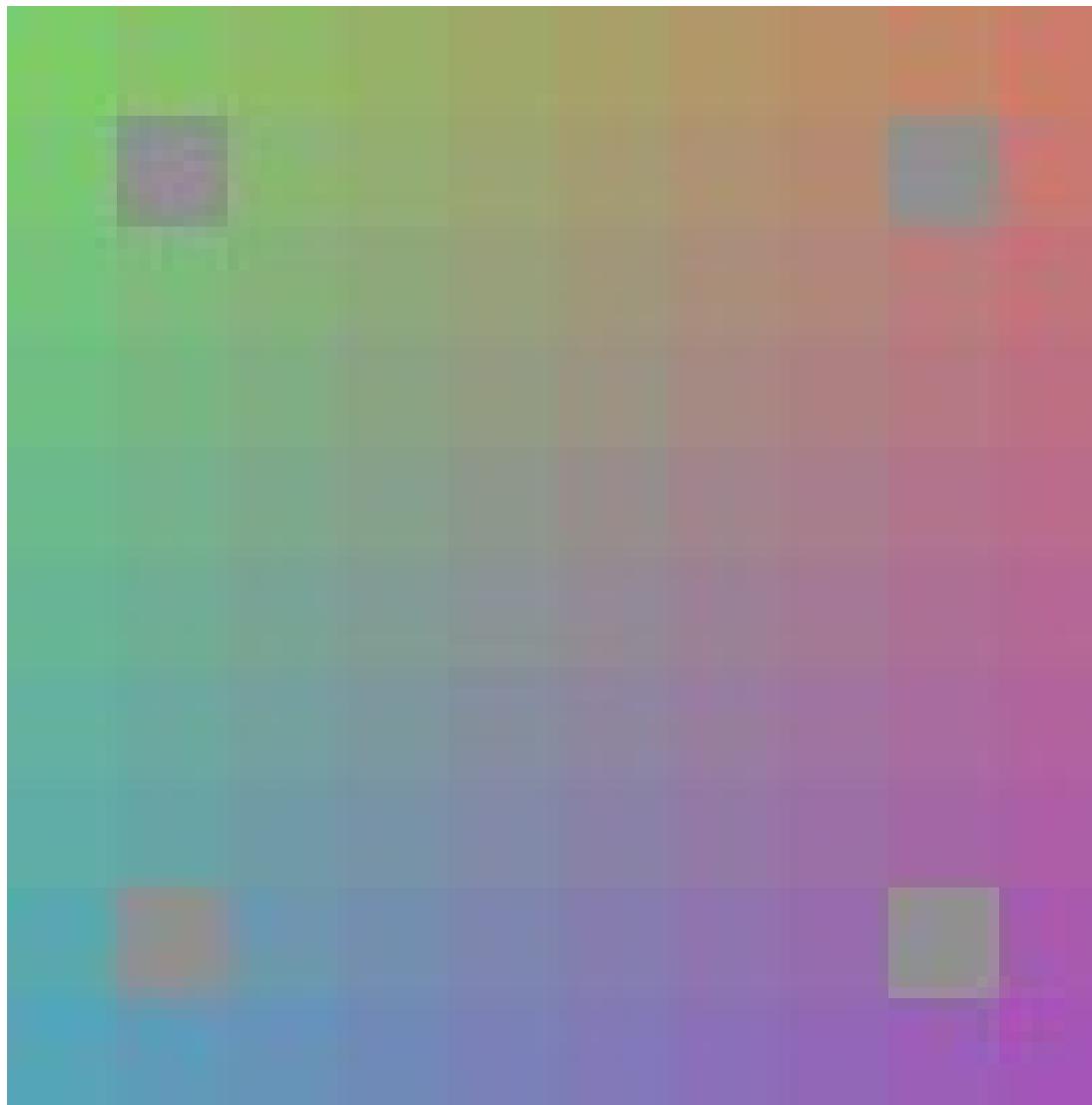


Edward H. Adelson

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

Slide by S. Lazebnik

Mach bands



Source: D. Forsyth

Edwin H. Land



- One of the founders of Polaroid
- Retinex = Retina + Cortex (1971)
- The author of "Color constancy"

Chromatic adaptation

The sensitivity of the visual system varies depending on the dominant illumination of the observed scene

The mechanism is poorly understood

Adaptation to different levels of illumination

The size of the pupil regulates the volume of light falling on the retina

The size changes dramatically at the entrance to the building with a brightly sunlit street

Color adaptation

Retinal cells change their sensitivity

Example: if the proportion of red in the lighting rises, the sensitivity of the cells responsible for red decreases, until the scene looks normal

We are better able to adapt in bright light, when illuminated by a candle everything remains in yellow tones

Balance of white

When we look at a photo or a monitor, the eyes adapt to the lighting in the room, and not to the lighting of the scene in the photo

If the white balance is inaccurate, the colors of the photo appear unnatural

incorrect white balance



correct white balance



<http://www.cambridgeincolour.com/tutorials/white-balance.htm>

Slide by S. Lazebnik

Balance of white

Film cameras:

Different types of film and light filters are used for different scenes?

Digital cameras:

Auto white balance

Preset white balance for typical shooting conditions

Adjustable for reference object

AWB	Auto White Balance
	Custom
K	Kelvin
	Tungsten
	Fluorescent
	Daylight
	Flash
	Cloudy
	Shade

Balance of white

Von Kries adaptation

- We multiply each channel by the transmission coefficient
- In a number of cases, the effect is more complex, corresponding to multiplying by a 3x3 matrix

The simplest method: gray (white) cards

- Photographing a neutral object (white)
- We estimate the weight of each channel
- If the color of the object is written as r_w, g_w, b_w ?
- Then weights $1/r_w, 1/g_w, 1/b_w$



Balance of white

If there are no gray cards, then we need to guess (or estimate) the gain factors

The model of the "Gray World" (Grayworld)

The average level ("gray") for each channel must be the same for all channels

If the color balance is violated, then the "gray" in this channel is larger than the "gray" other channels

Calculate the gain factors so that the average in each

$$\bar{R} = \frac{1}{N} \sum R(x, y); \quad \bar{G} = \frac{1}{N} \sum G(x, y); \quad \bar{B} = \frac{1}{N} \sum B(x, y); \quad Avg = \frac{\bar{R} + \bar{G} + \bar{B}}{3};$$

$$R' = R \cdot \frac{Avg}{\bar{R}}; \quad G' = G \cdot \frac{Avg}{\bar{G}}; \quad B' = B \cdot \frac{Avg}{\bar{B}}$$

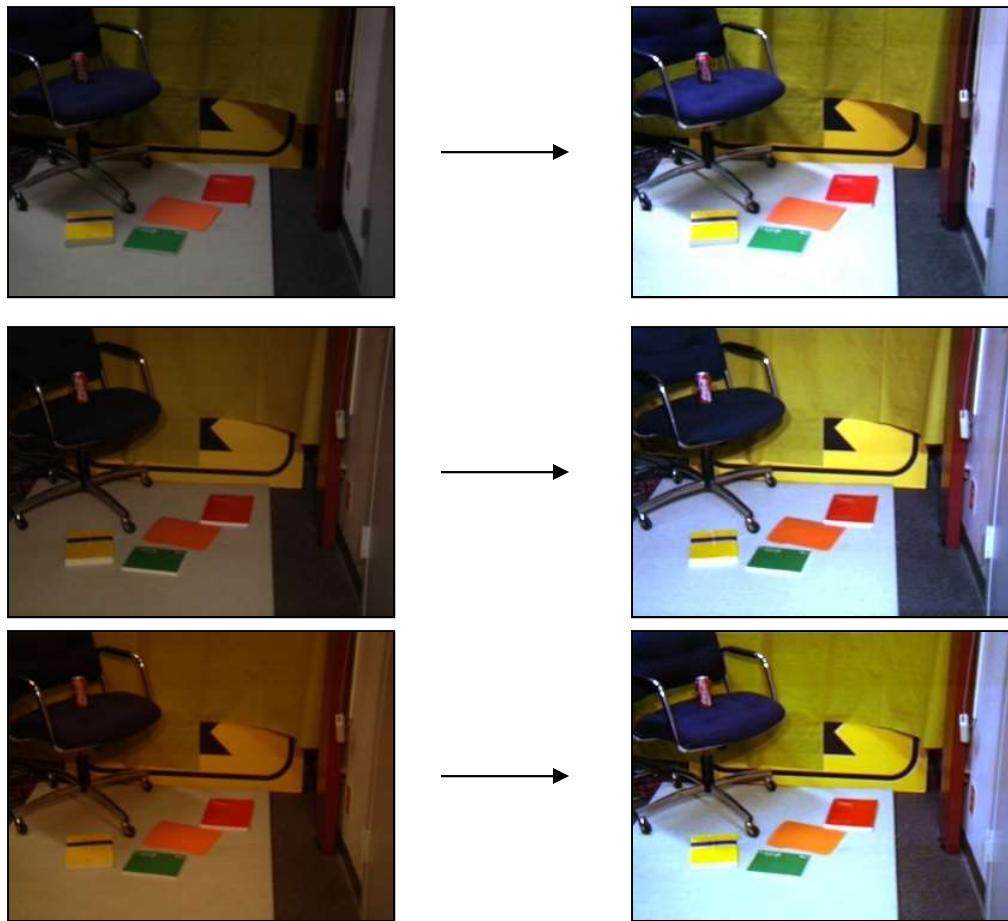
«Gray world» - examples



«Gray world» - example



«Gray world» - example



Other approaches

Light pattern

The colors of the glare usually correspond to the color of the source

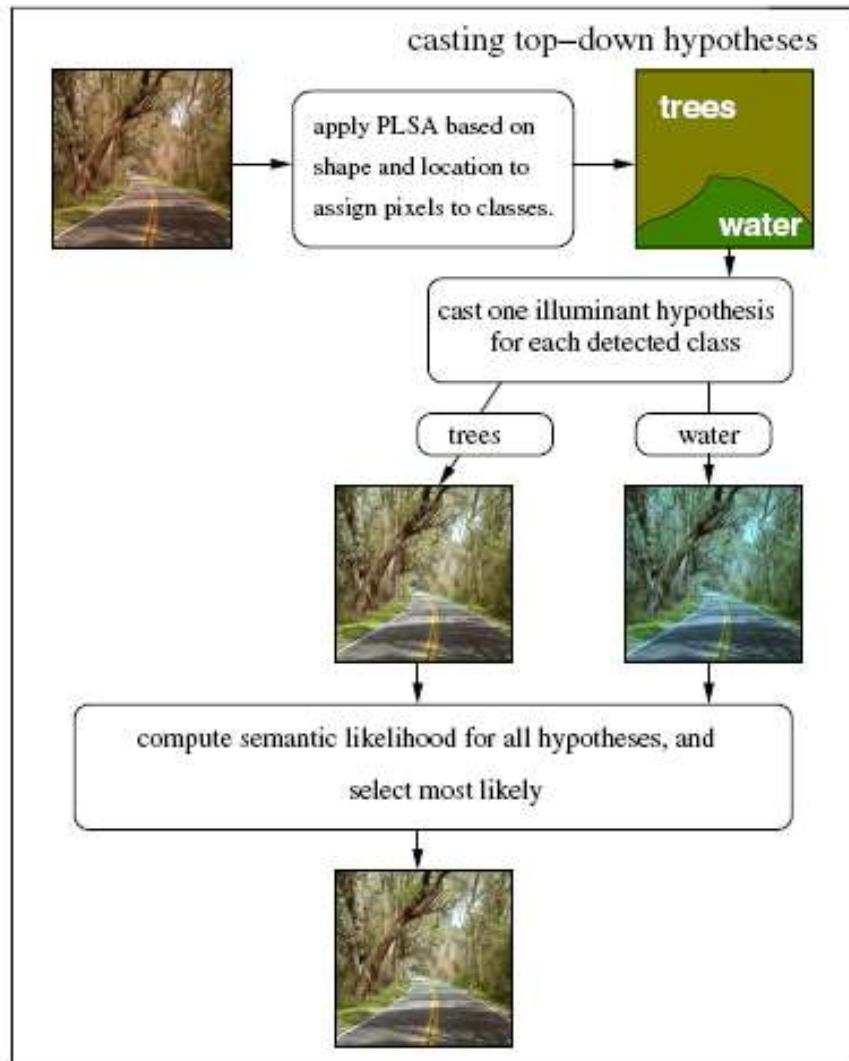
We are looking for the brightest pixel, and we take the weight inversely proportional to its colors

Range Mapping (Gamut)

Gamut: the convex hull of colors of all pixels

Choose a transformation that converts the image range to "standard" in daylight

White balance detection



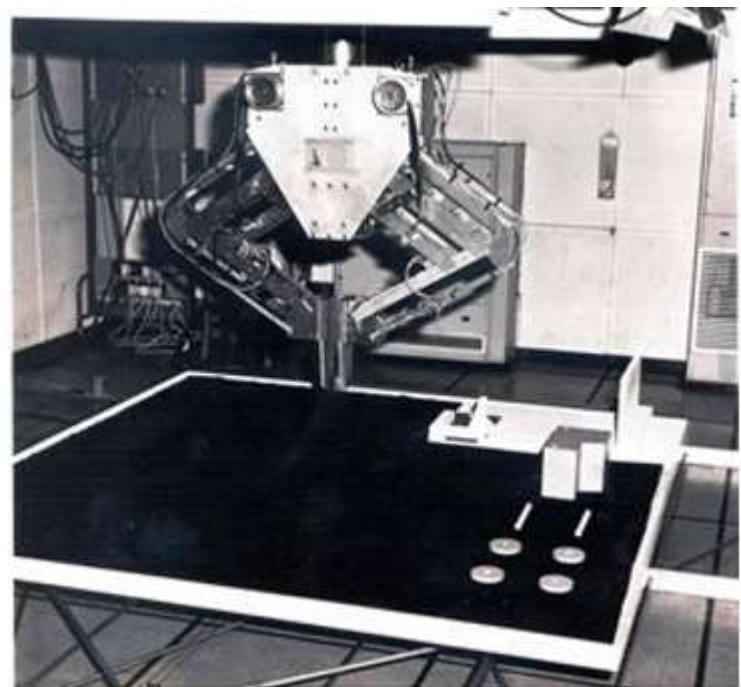
Idea: For each class of objects present in the scene, we calculate the transformation so that the range of colors of the object coincides with the average range of objects of this class on "typical" images

J. Van de Weijer, C. Schmid and J. Verbeek, [Using High-Level Visual Information for Color Constancy](#), ICCV 2007.

Slide by S. Lazebnik

Freddy II 1973

- University of Edinburgh
- One of the first robots with a system of machine vision
- 5 degrees of freedom
- Can collect machines from cubes scattered on the table
- 384KB RAM in the control computer



David Marr (1970)

- "Primal sketch"
 - Low-level image properties: directed edges, segments, etc.
- "2.5D sketch"
 - Arrangement by depth (binocular stereo), texture registration, etc.
- "3D model"
 - Object recognition and representation of the 3-dimensional world

“The Marr Prize”

Viola-Jones (2001)



Source: S. Seitz

Emotions

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



OCR



FineReader, ABBYY, Россия
<http://www.abbyy.ru/finereader/>

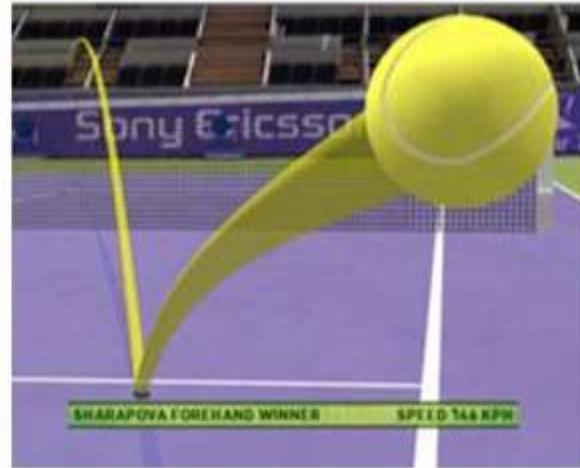


Cognitive Technologies, Россия
<http://cognitiveforms.ru/products/cognitive-forms/>



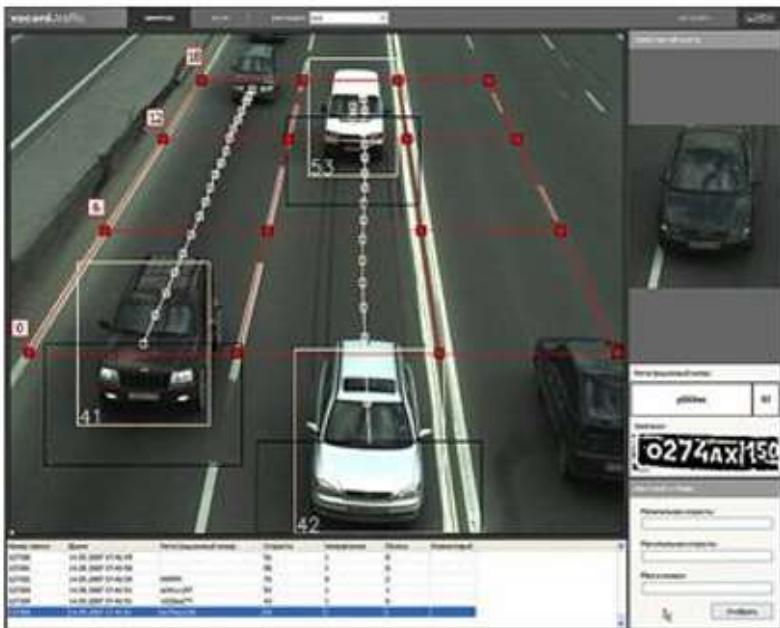
Распознавание текста в
произвольных изображениях пока
работает плохо!

Sport



<http://www.hawkeyeinnovations.co.uk>

Flow

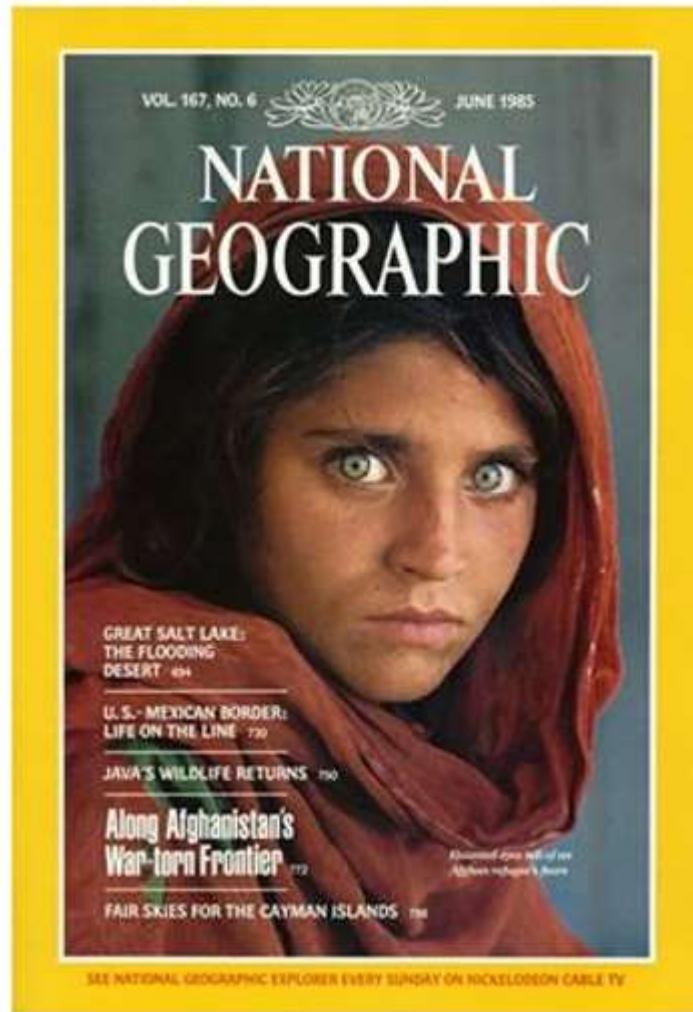


VOCORD Traffic, Vocord, Россия
<http://www.vocord.ru/218/>



Автоматизация видеонаблюдения
за счёт видеоаналитики
(в т.ч. в нашей лаборатории)

Biometrics



Augmented reality



L.A. Noire, Team Bondi/Rockstar, 2011



Kinect

