Introduction to Computer Vision

Computer Vision

Taught by

- prof. Luc Van Gool
- Prof. Ender Konukoglu

- Guest starring by prof
- Orcun Goksel

The course comes with a course text that covers most – but not all ! – material.
Slide decks for all lectures will be made available on eDoz or similar

We got questions about which course to take

Computer Vision (D-INFK), or
Image Analysis and Computer vision (this course)

IN ANY CASE, DO NOT TAKE BOTH !

If you took the introductory course on CV at D-INFK,
then best take Computer Vision

If you did not take that course,
then best take Image Analysis and Computer Vision

We got questions about which course to take

Computer Vision (D-INFK), or
Image Analysis and Computer vision (this course)

IN ANY CASE, DO NOT TAKE BOTH !

If you took the introductory course on CV at D-INFK,
then best take *Computer Vision*

If you did not take that course,
then best take *Image Analysis and Computer Vision*

… it is crucial …

The central take-home message:

For people vision is their most crucial sense, for good reason

Vision is important

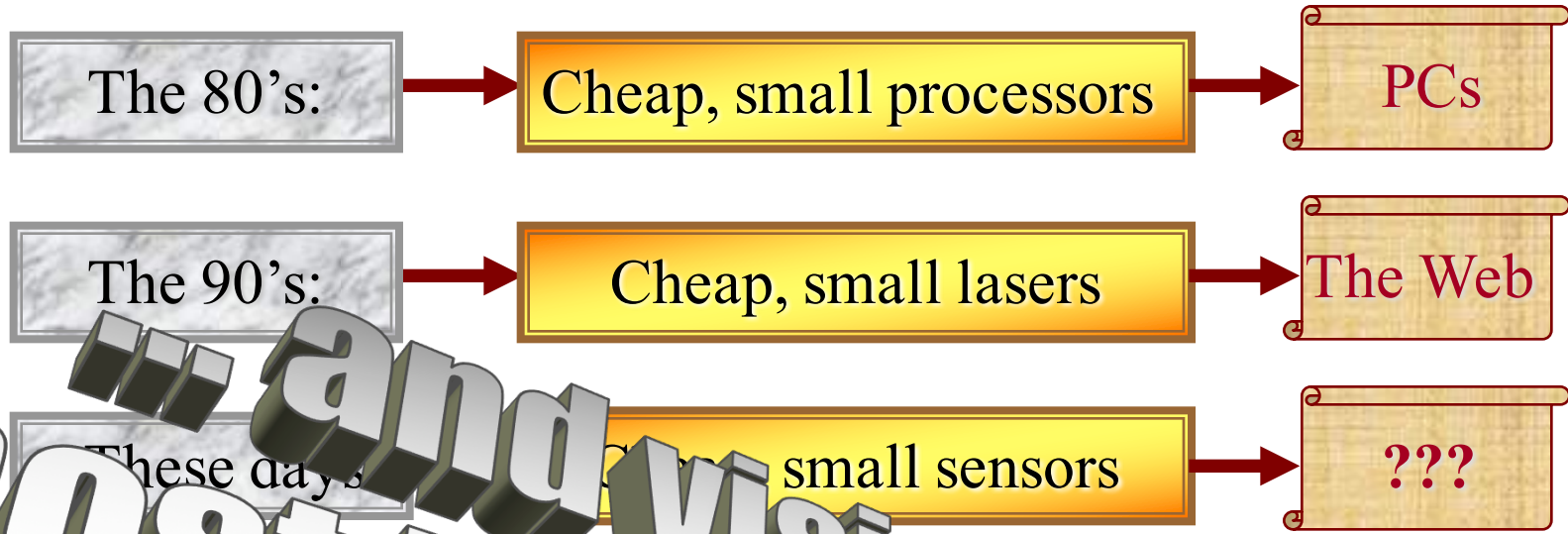
- half our brain is devoted to it
- developed many times during evolution
- it is non-contact
- it can be implemented with high resolution
- works with ambient E-M waves
- yields colour, texture, depth, motion, shape

… it is hot …

The central take-home message:

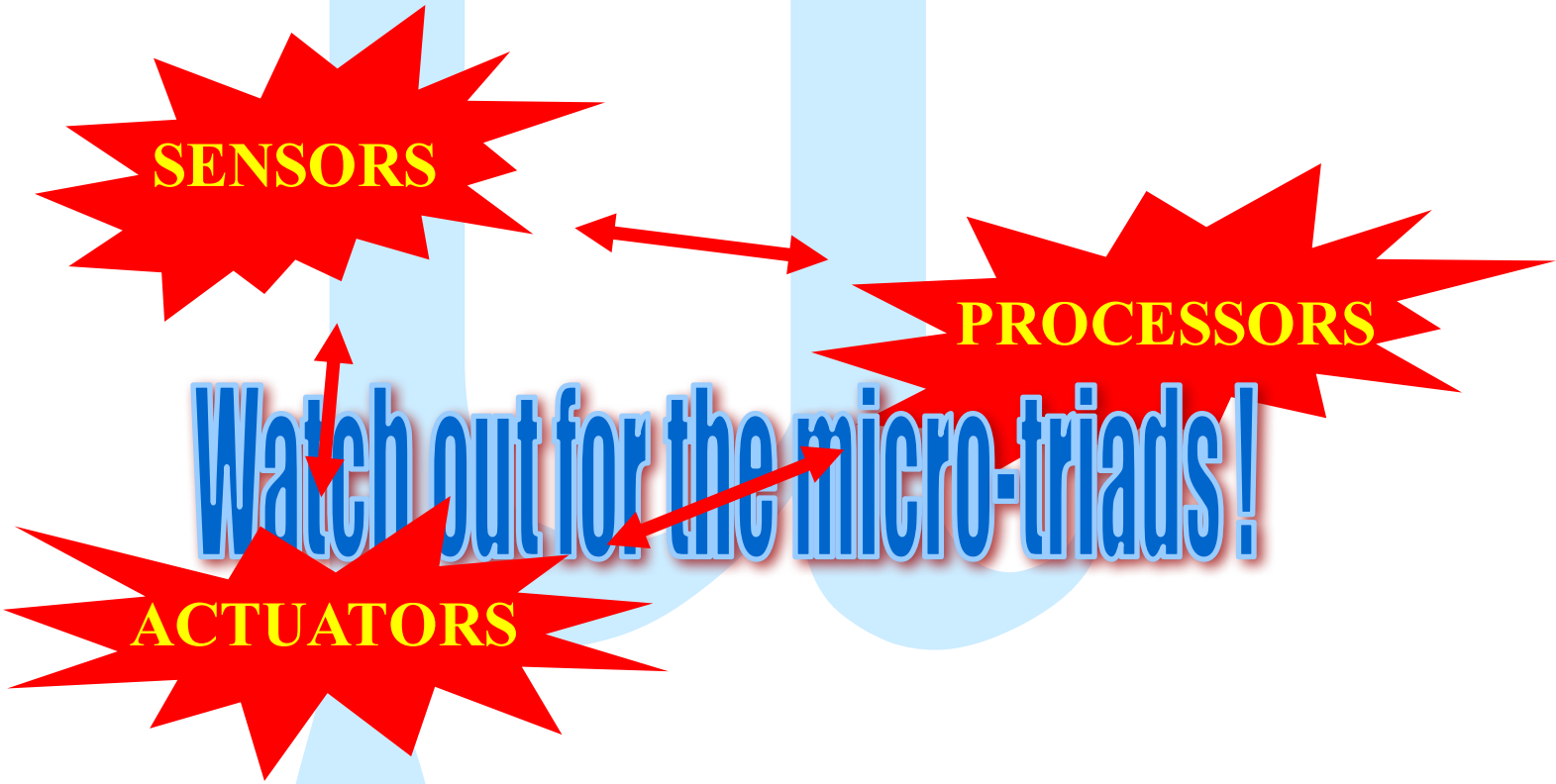
It is feasible now to let most things see their environment

CV at the forefront of technology



"... and vision is our most important sense!"

Watch out for the micro-triads

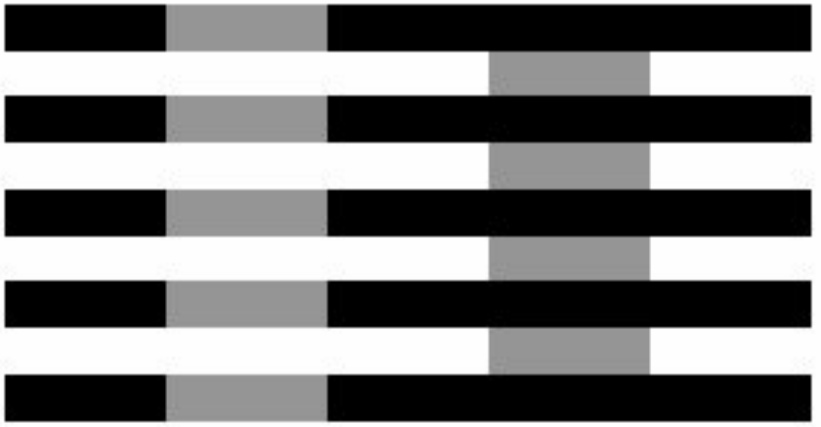


… it is intriguing …

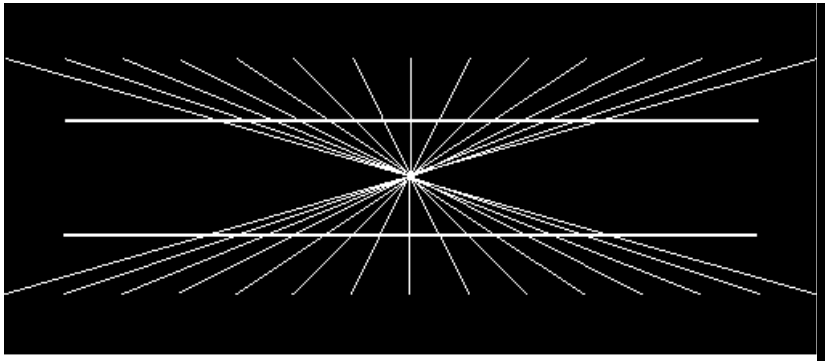
The central take-home message:

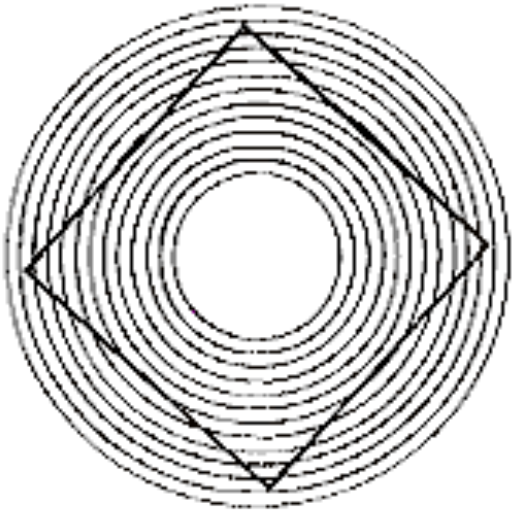
Effective vision needs more than sheer filtering and measuring

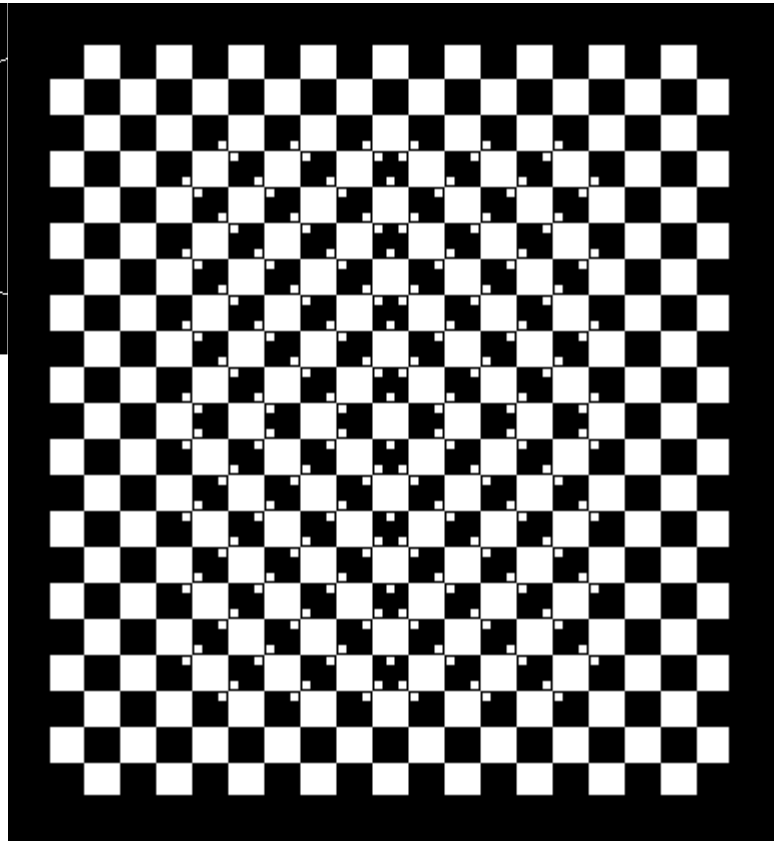
The perception of intensity



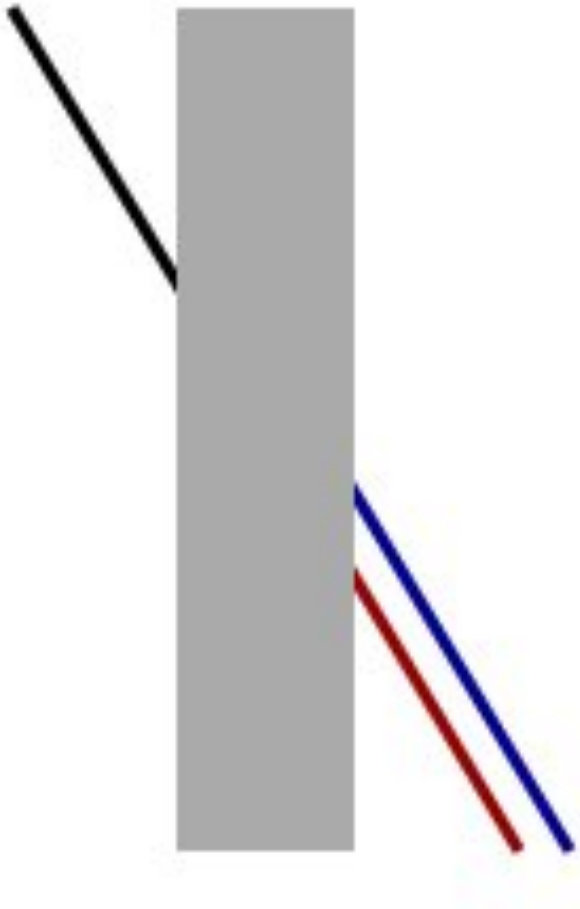
The perception of lines being straight



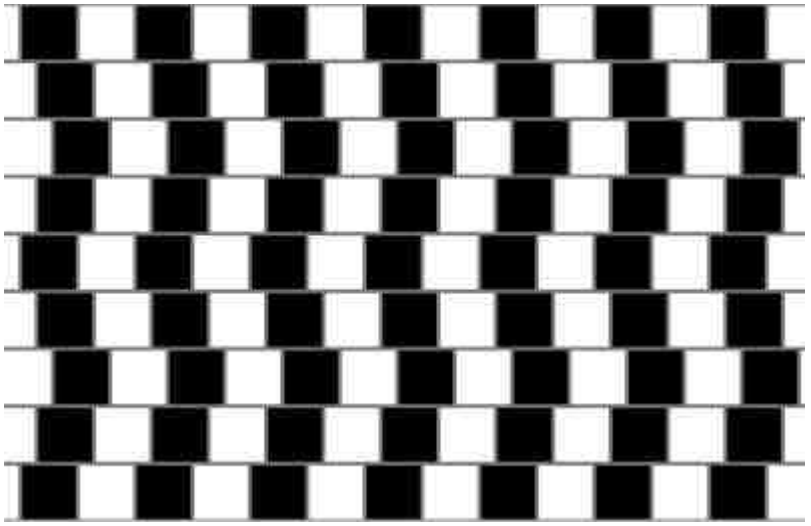


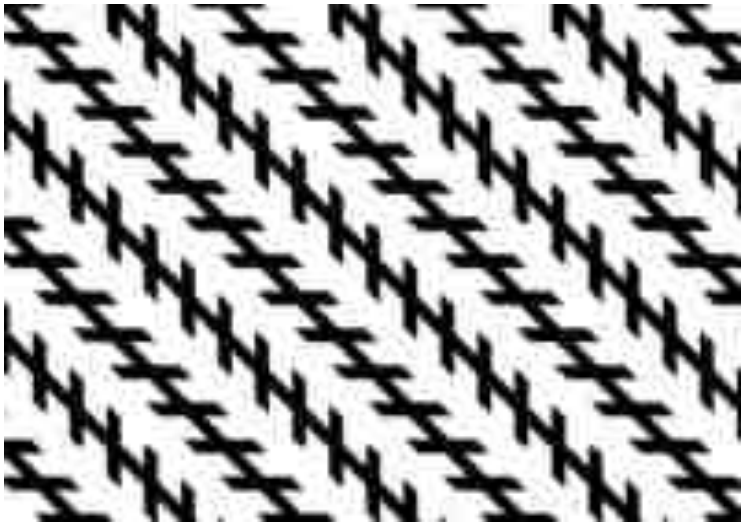


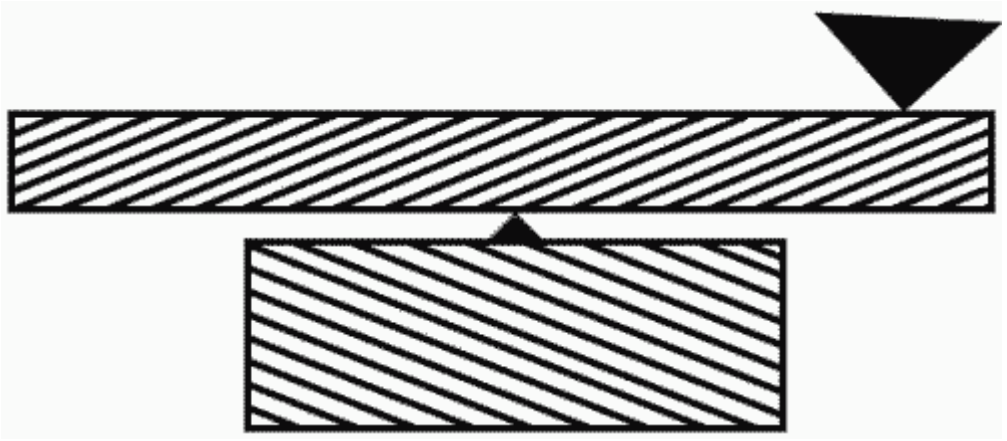
The perception of colinearity



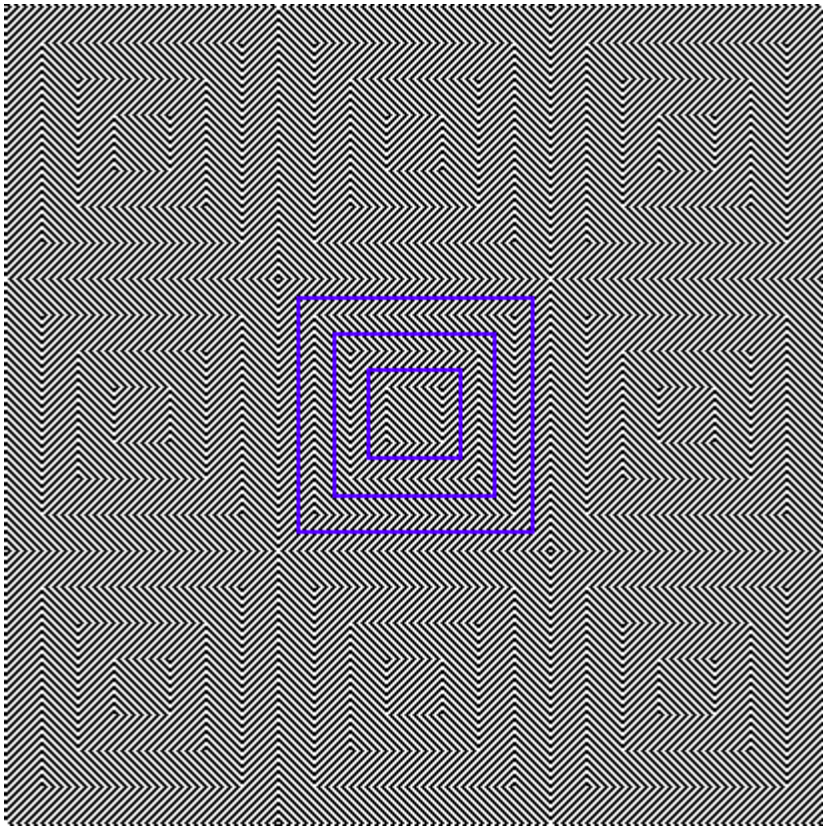
The perception of parallelism



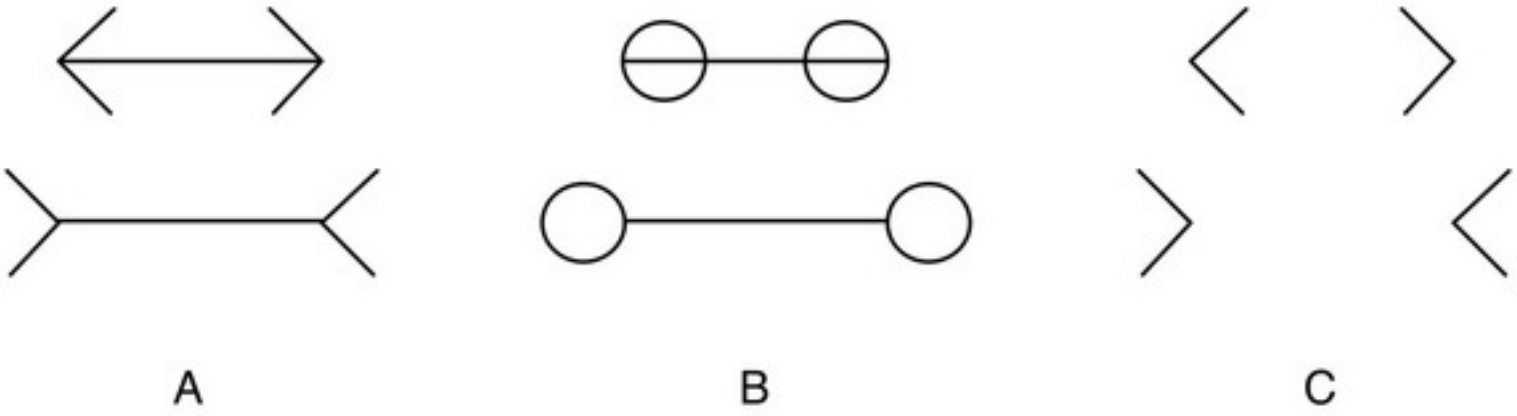




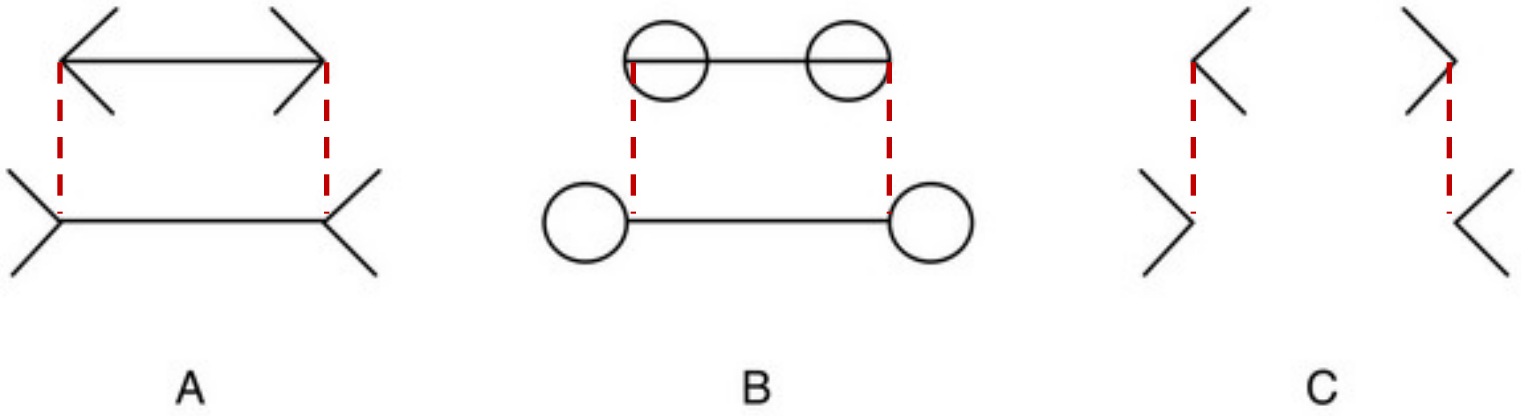
The perception of squares



The perception of length

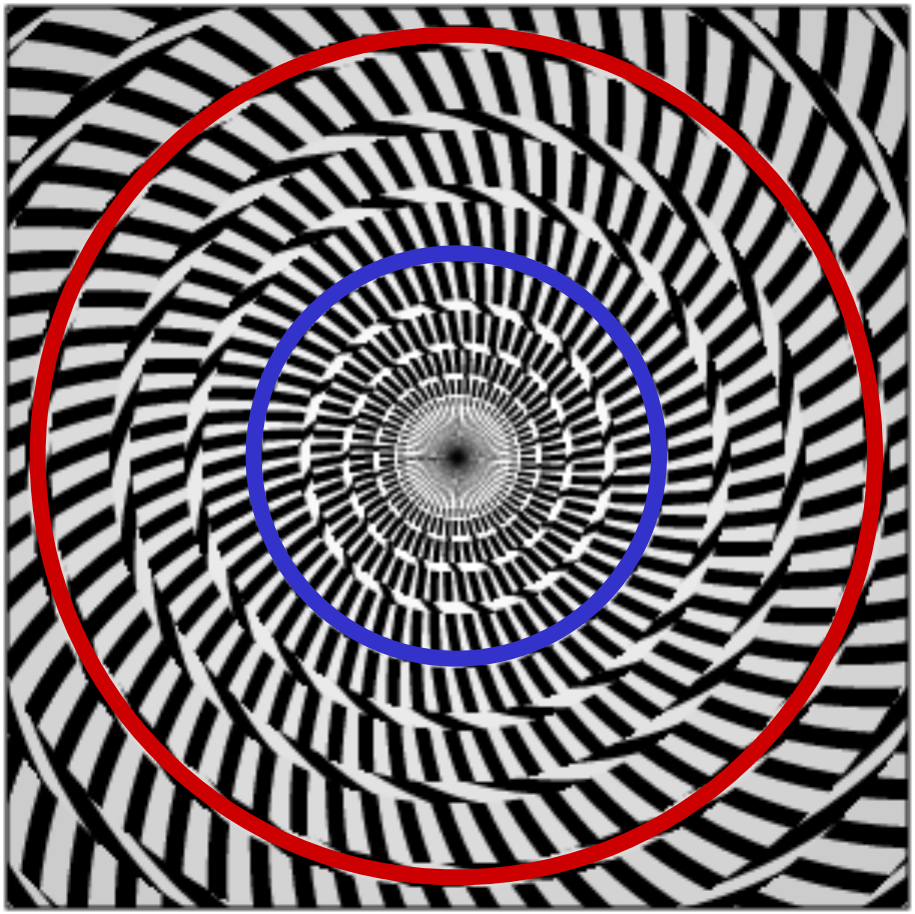


The perception of length



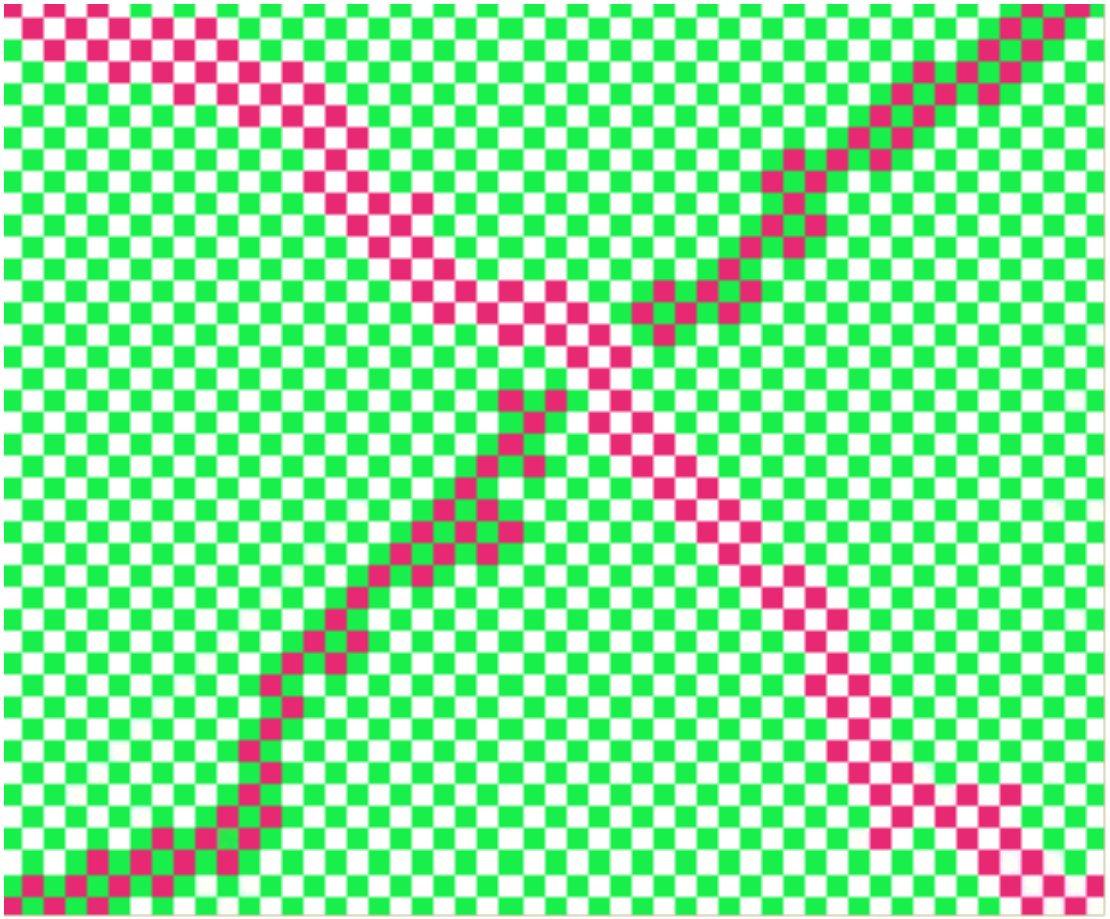
The horizontal lines are equally long…

The perception of curvatures



Illusions : interference of differently oriented patterns via adaptation

The perception of color

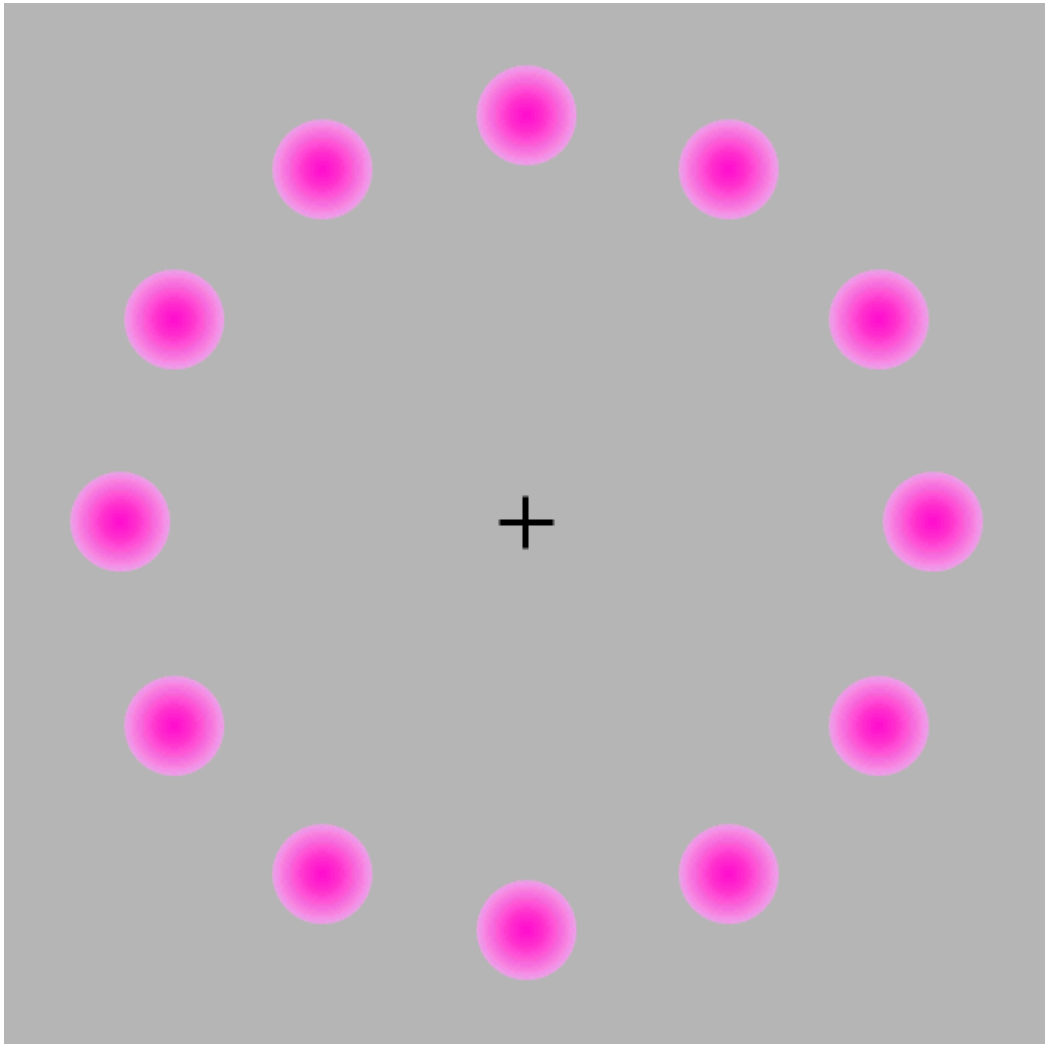


The red squares have equal color…

The perception of color

The dogs are identical…

The perception of color



Focus on the cross and you see a virtual green dot circling…

The perception of motion

The `barber pole' rotates about the vertical, it does not translate vertically…

… it is useful …

Applications of computer vision

The development of computer vision apps

Most early applications where found in production environments, as these *allow for controlled conditions* and *have little uncertainty*

some areas do not allow for much control: medical IP, remote sensing, surveillance, etc.

currently CV is conquering the less controllable areas by storm

Rationales

- increase productivity
- increase reliability
- increase flexibility
- decrease costs
- assist with quantitative aspects of a job
- guarantee constant vigilance/assistance
- realise intelligent man/machine interfaces
- automate complex processes
- generate more complete 3D models

Rationales

- increase productivity
- increase reliability
- increase flexibility
- decrease costs

Classical apps

- assist with quantitative aspects of a job
- guarantee constant vigilance/assistance
- realise intelligent man/machine interfaces
- automate complex processes
- generate more complete 3D models

New apps

Rationales

- increase productivity
- increase reliability
- increase flexibility
- decrease costs

Classical apps

- assist with quantitative aspects of a job
- guarantee constant vigilance/assistance
- realise intelligent man/machine interfaces
- automate complex processes
- generate more complete 3D models

Rationales behind computer vision for traditional visual inspection

a well engineered vision system can:

- increase productivity
- increase reliability
- increase flexibility
- decrease costs

Computer Vision

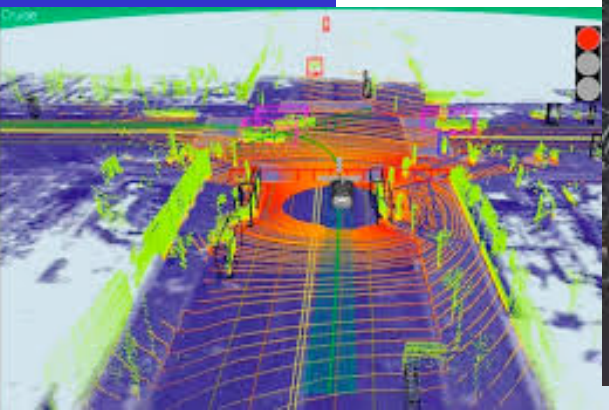
Rationales

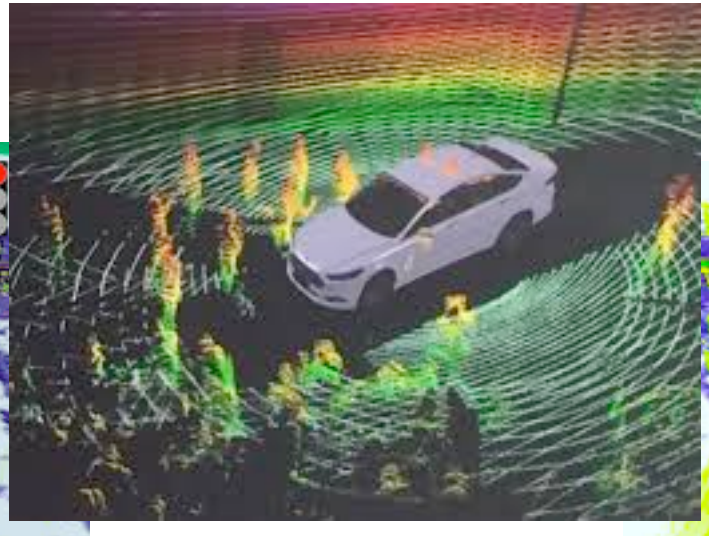
- increase productivity
- increase reliability
- increase flexibility
- decrease costs
- assist with quantitative aspects of a job
- guarantee constant vigilance/assistance
- realise intelligent man/machine interfaces
- automate complex processes
- generate more complete 3D models

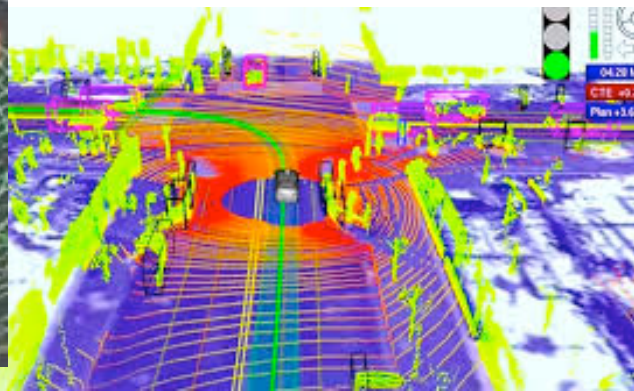
New apps

Computer Vision

Ex App: autonomous vehicles

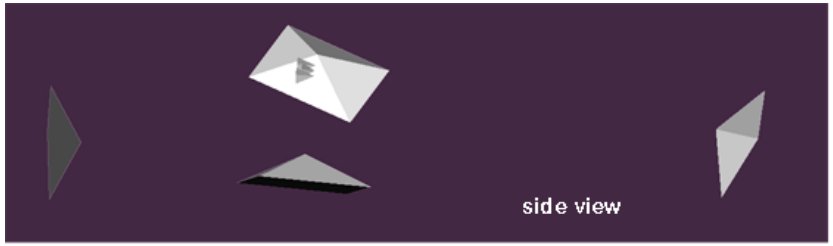


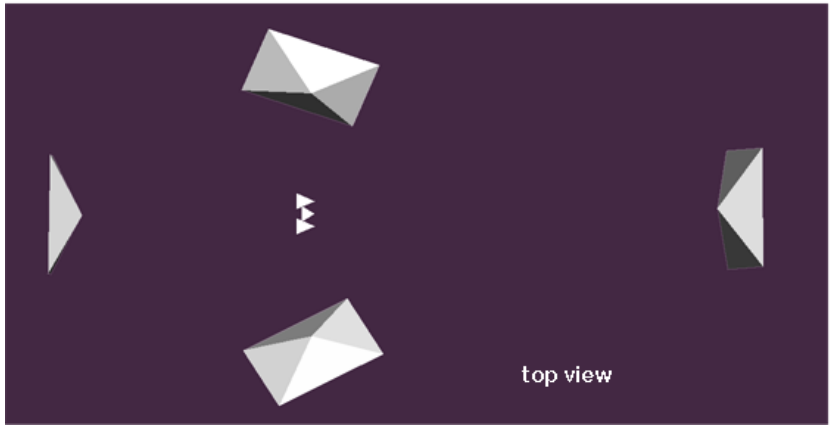


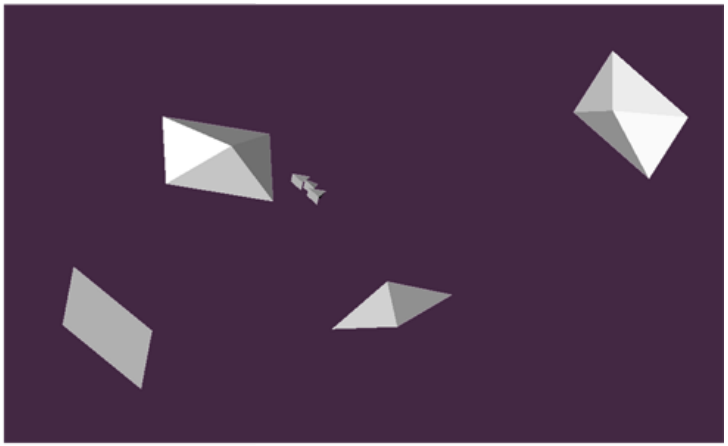


Ex App: autonomous vehicles

- 3 forward normal (1 stereo pair + 1 wide angle) + 4 fish eyes for 360 vision
- Images below show the result of calibration
- We also take into account if cams are behind glass







Computer Vision

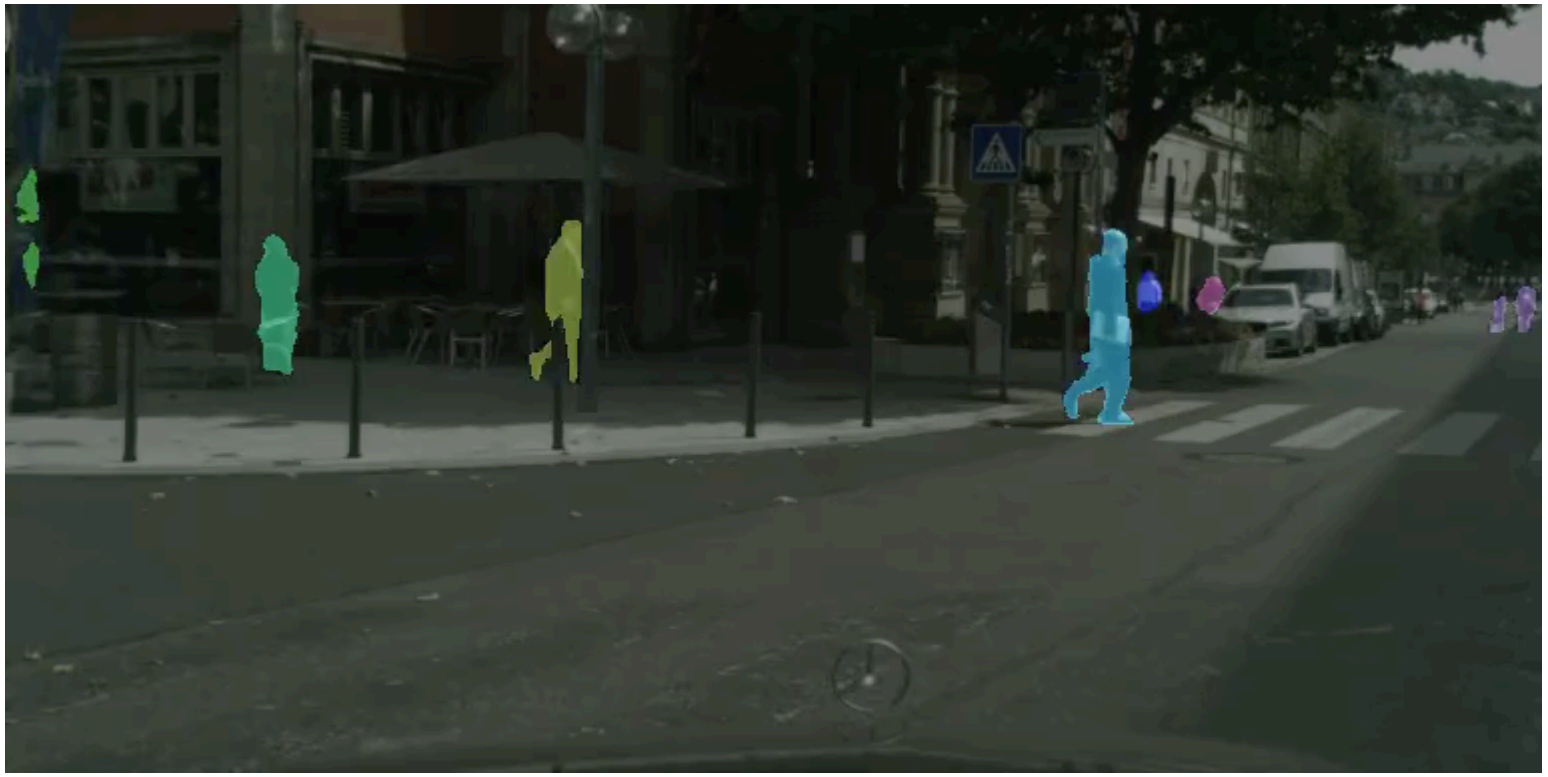
Ex App: autonomous vehicles

car detection:

Computer Vision

Ex App: autonomous vehicles

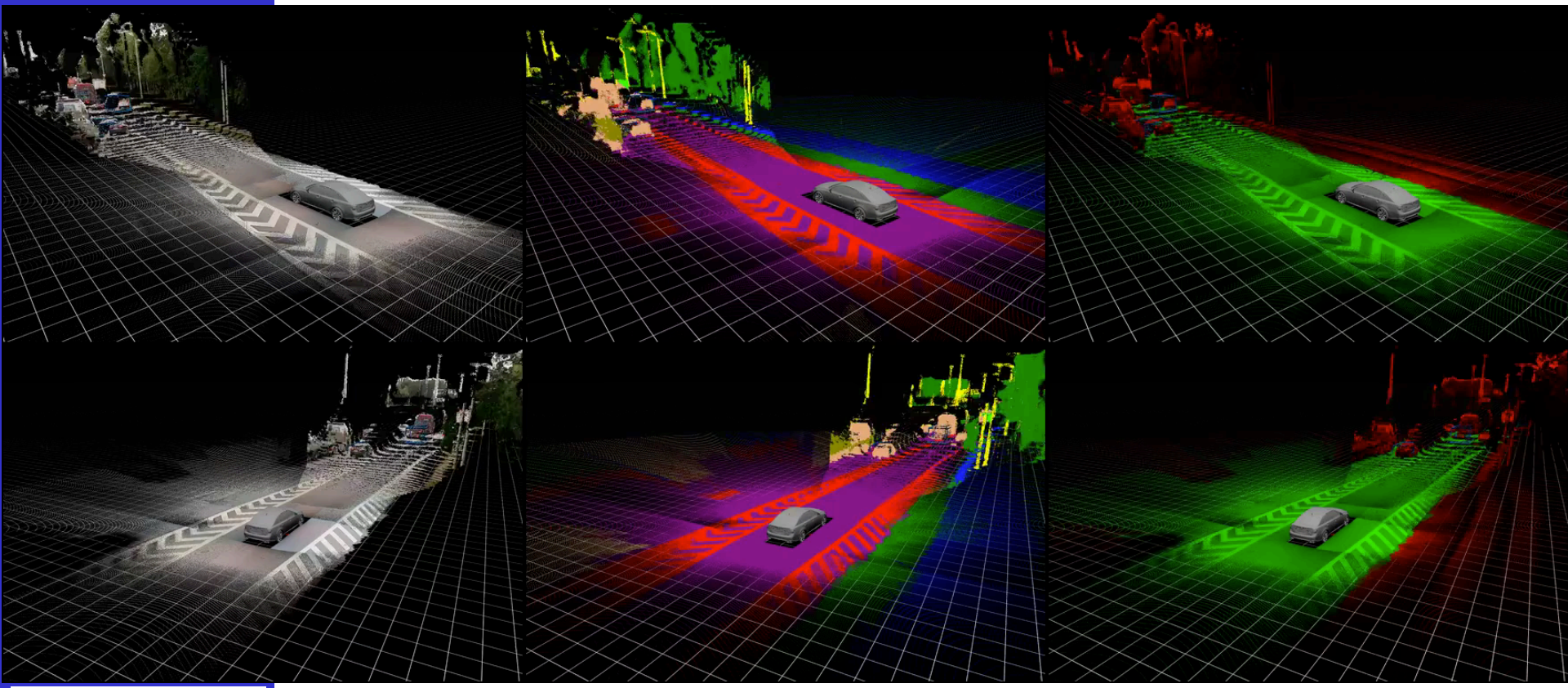
pedestrian detection:



Computer Vision

Ex App: autonomous vehicles

putting vision modalities together:

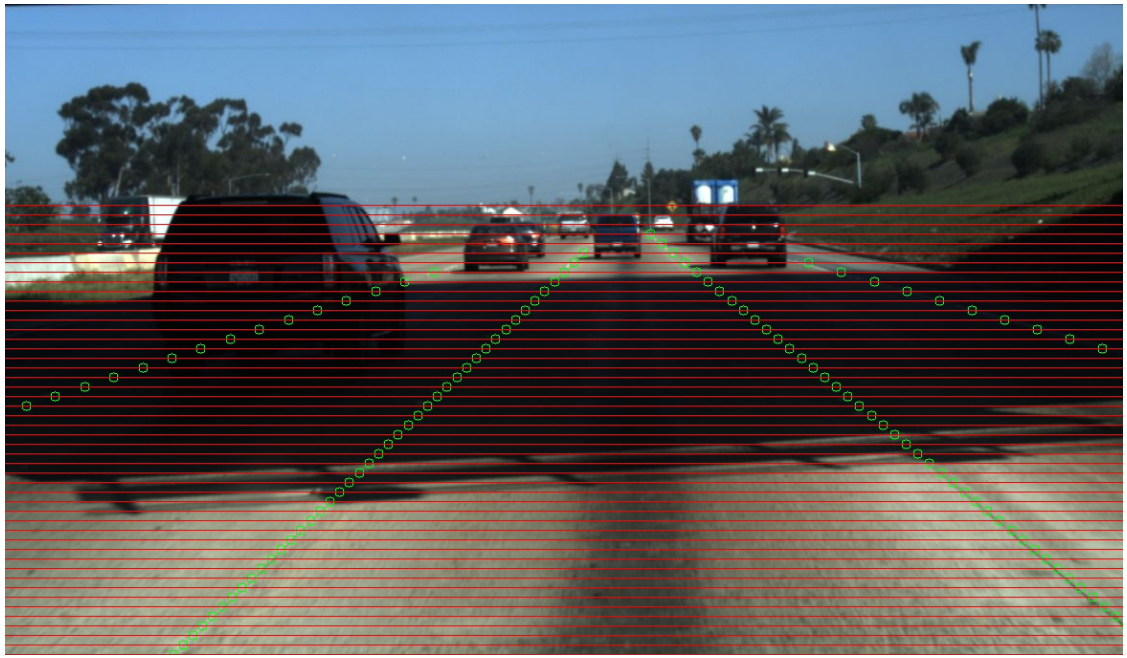


Computer Vision

Ex App: autonomous vehicles

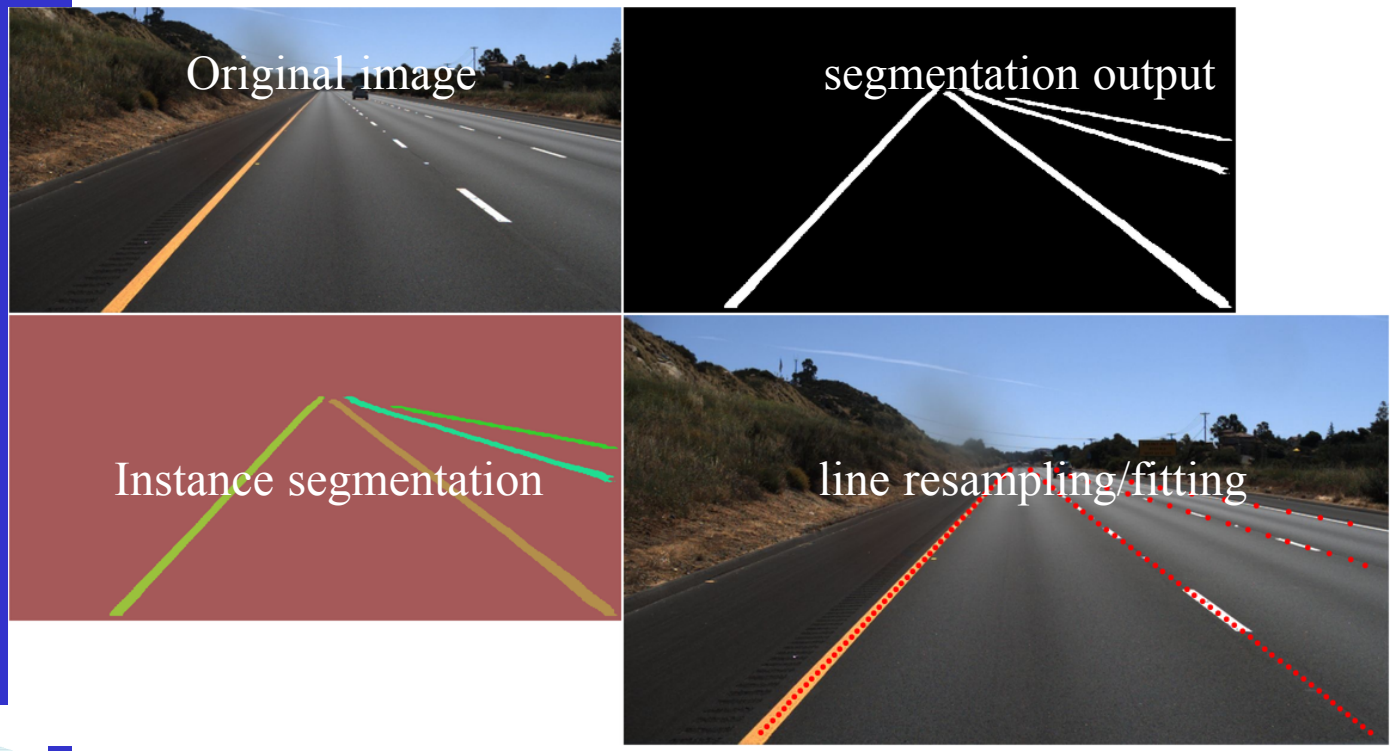
Latest addition: lane marking detection

Annotations based on new benchmark : http://benchmark.tusimple.ai/#/t/1
Markings are annotated as series of connected points on equidistant rows (see example picture below)
Note: lanes are also annotated when markings are not present, not visible, or even occluded by cars



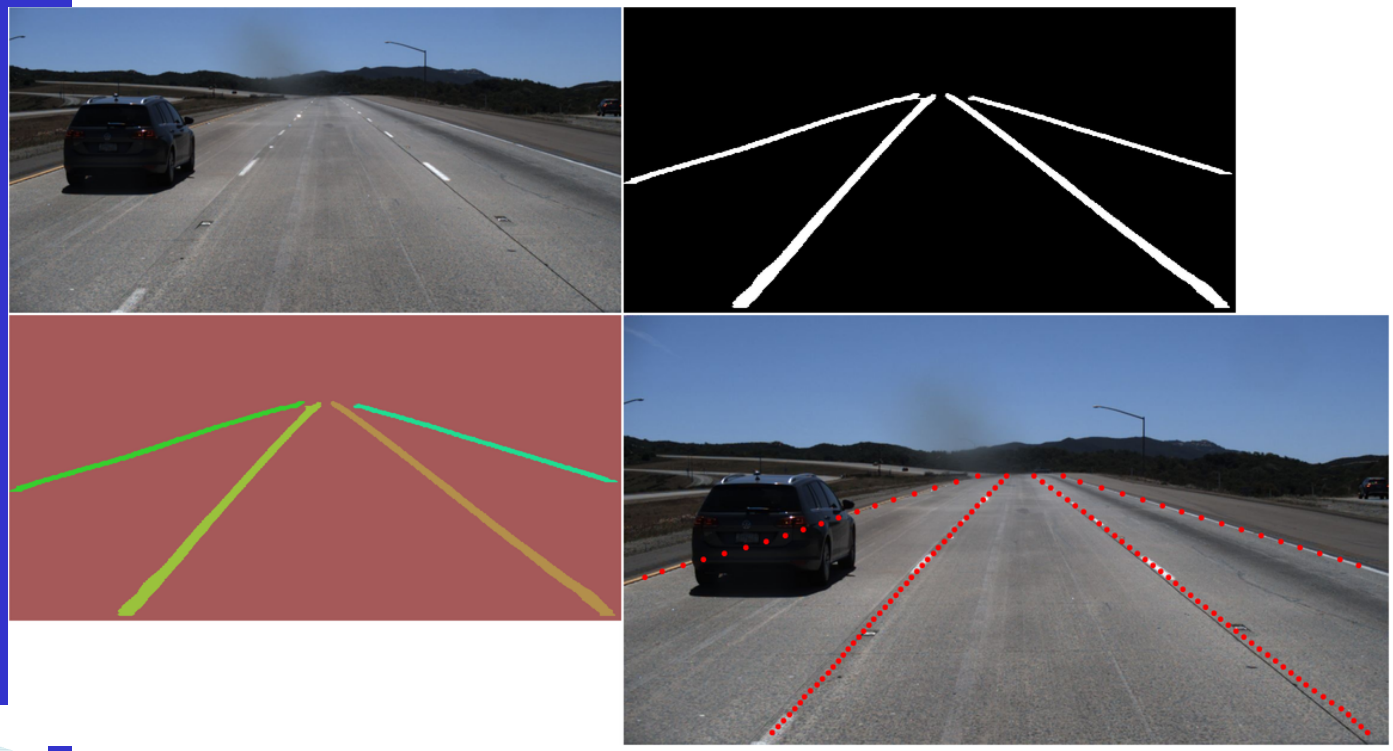
Ex App: autonomous vehicles

Latest addition: lane detection



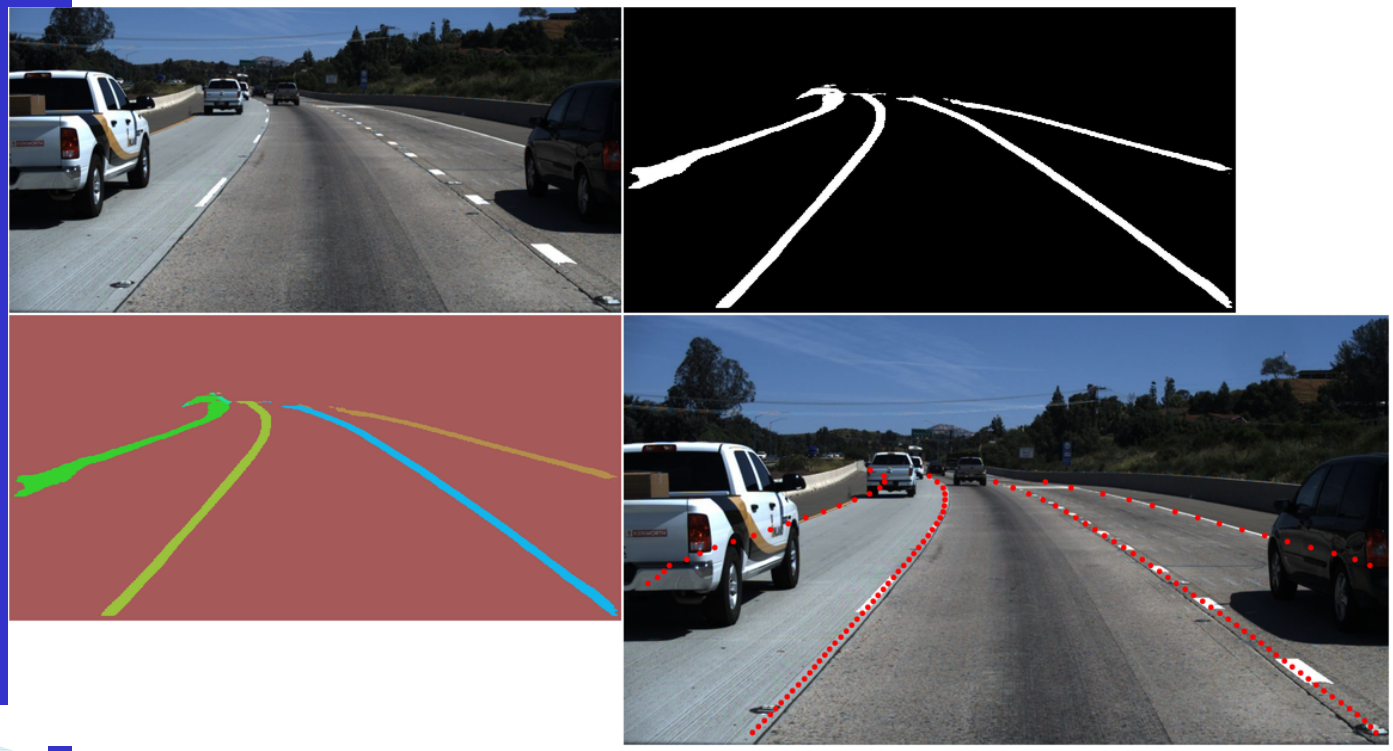
Ex App: autonomous vehicles

Latest addition: lane detection



Ex App: autonomous vehicles

Latest addition: lane detection



Computer Vision

Ex: autonomous mobile platform

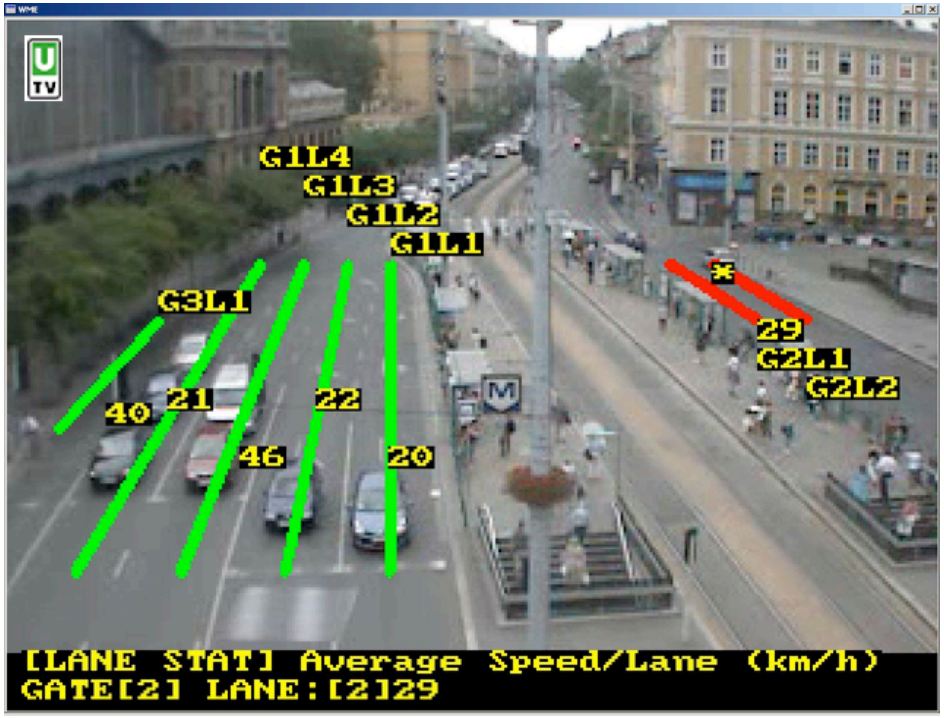
Computer Vision

Ex App: image retrieval, captioning, …

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image
 A person riding a motorcycle on a dirt road.	 Two dogs play in the grass.	 A skateboarder does a trick on a ramp.	 A dog is jumping to catch a frisbee.
 A group of young people playing a game of frisbee.	 Two hockey players are fighting over the puck.	 A little girl in a pink hat is blowing bubbles.	 A refrigerator filled with lots of food and drinks.
 A herd of elephants walking across a dry grass field.	 A close up of a cat laying on a couch.	 A red motorcorcycle parked on the side of the road.	 A yellow school bus parked in a parking lot.

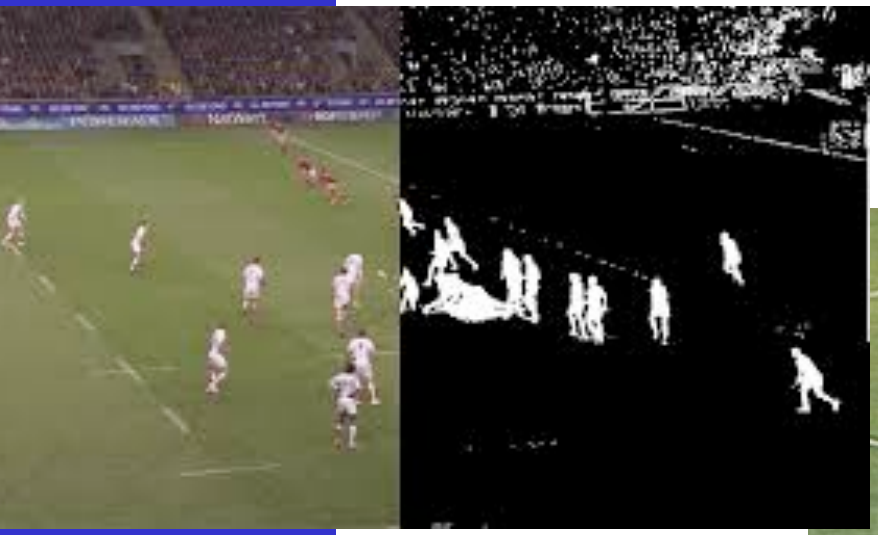
Computer Vision

Ex App: visual surveillance



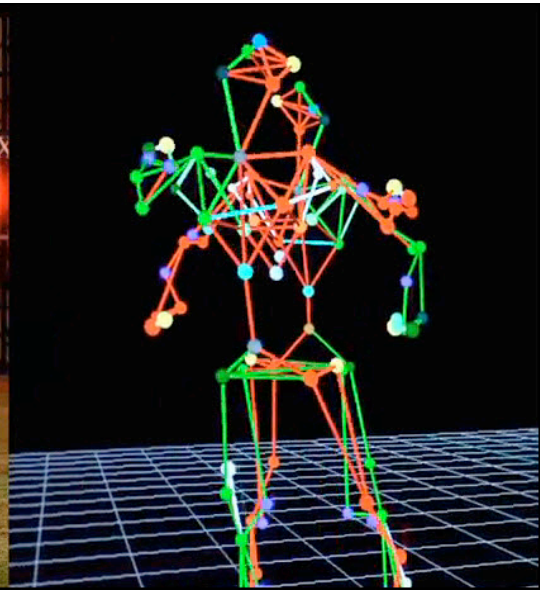
Computer Vision

Ex App: Augm. Reality, eg sports



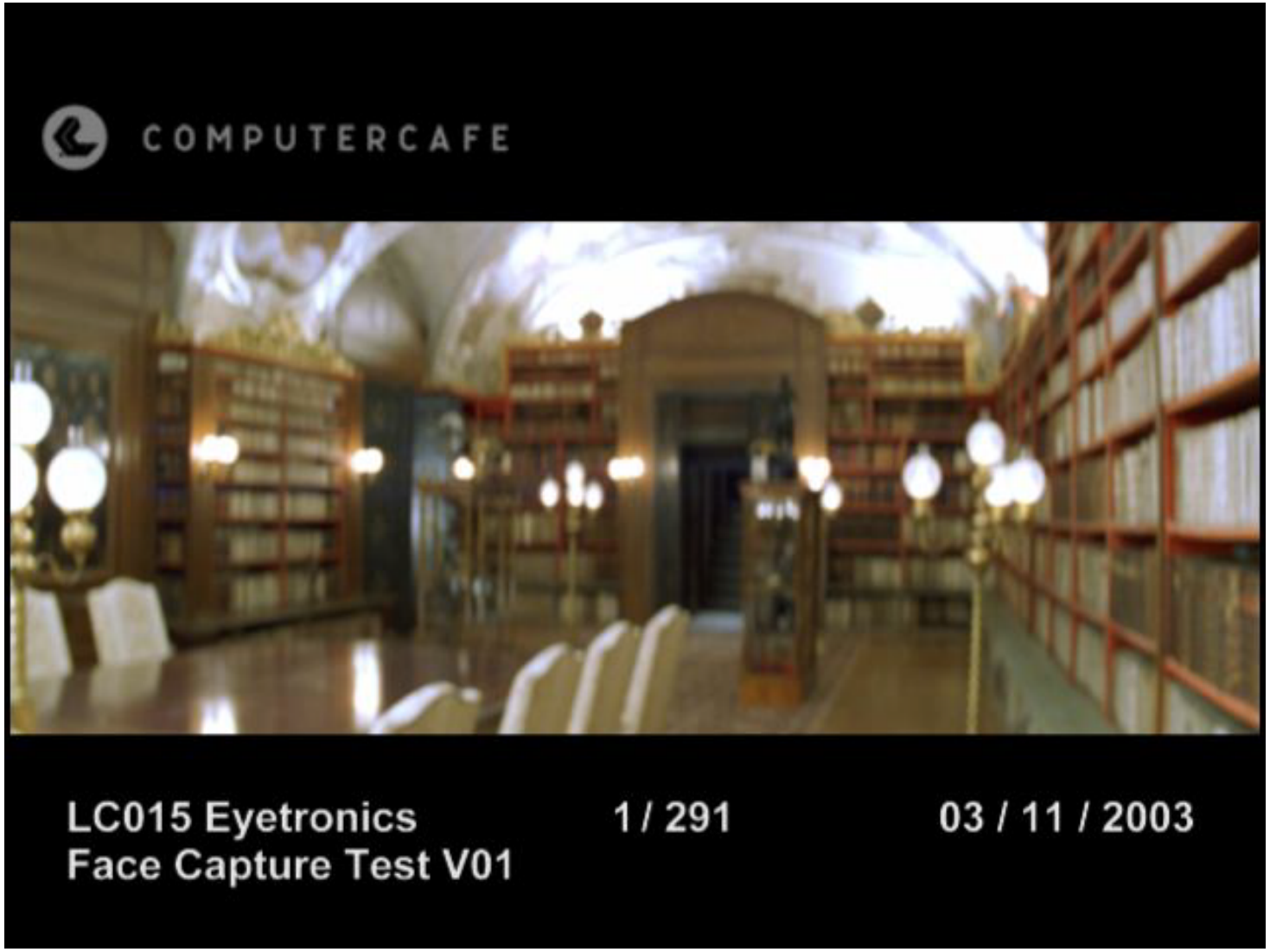
Computer Vision

Ex App: motion capture for movies/games



Ex App: facial performance capture



Ex App: advert replacement

Ex app: Avatar guidance via 3D facial capture

Face/Off: Live Facial Puppetry

PaperID 102

Developed by our spin-off FaceShift, acquired by Apple. The basis for their user-animated emoji's in new iPhone

Computer Vision

Ex App: computer-assisted surgery

Computer Vision

Mobile mapping

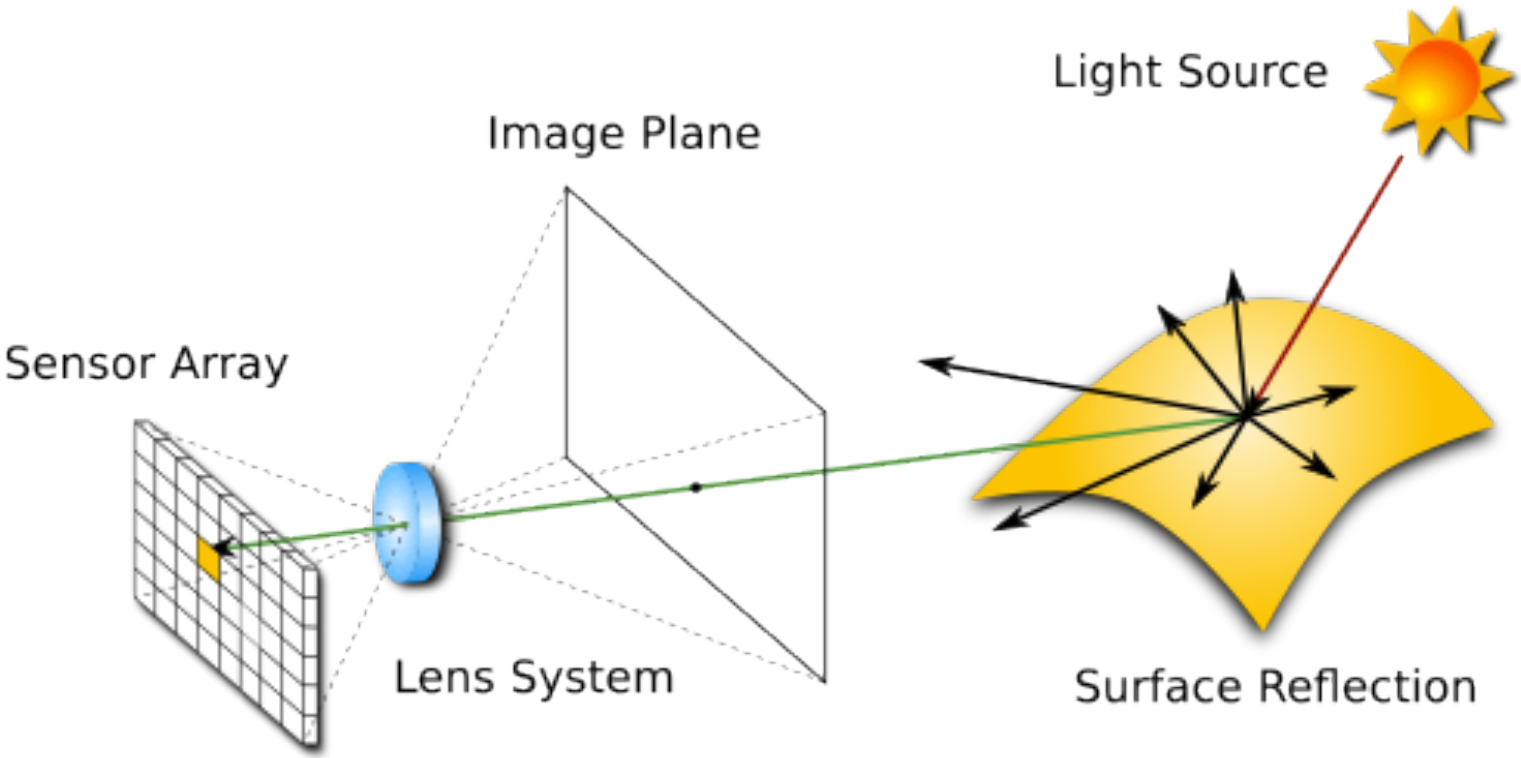
… it needs light …

And then there was Light…

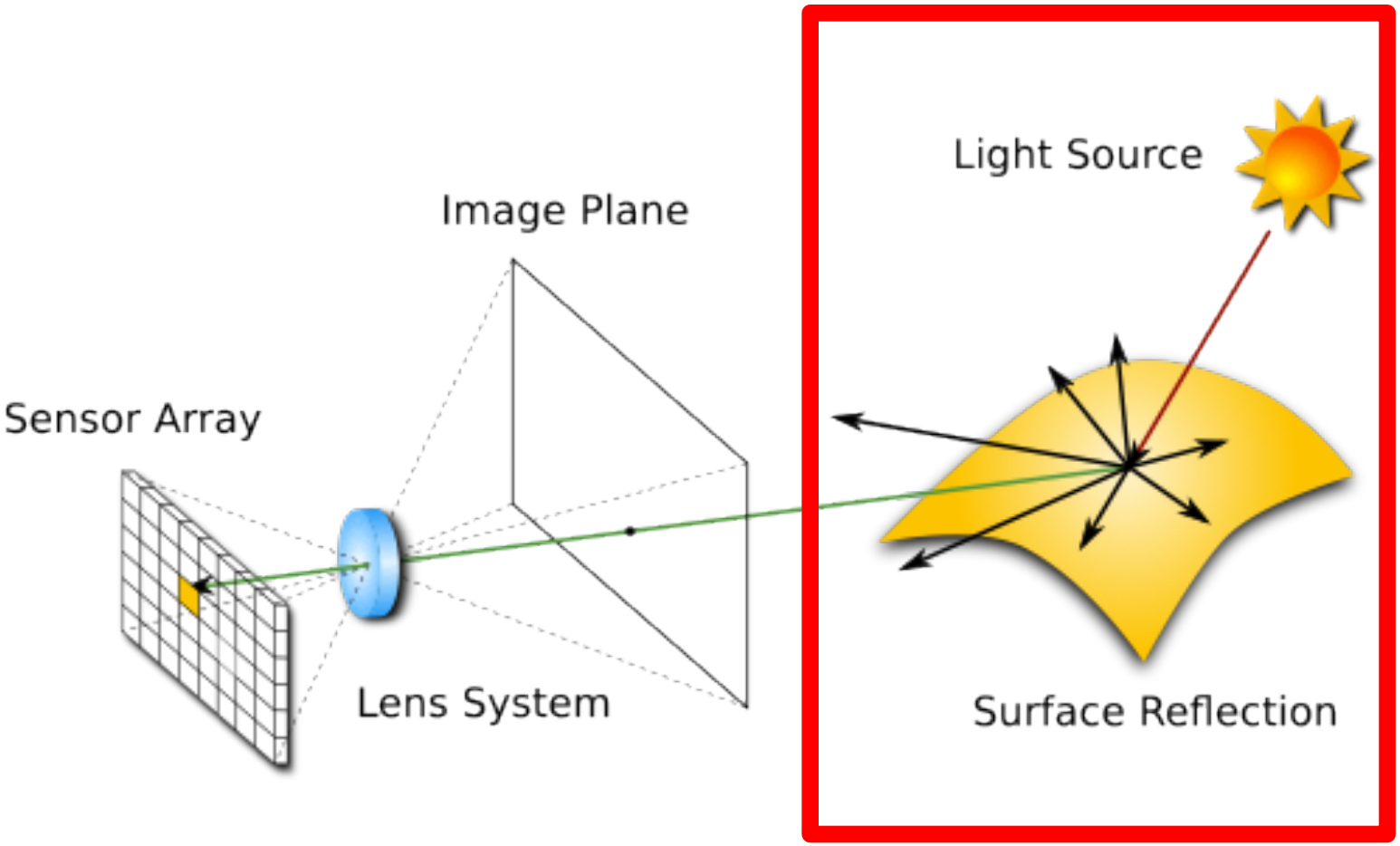
- no vision without light…
- … because it is influenced by objects

"What the...?"

Kickoff: the light, surface, lens & cam



Kickoff: the light, surface, lens & cam



topics

- the nature of light
- interactions with matter

An option on optics

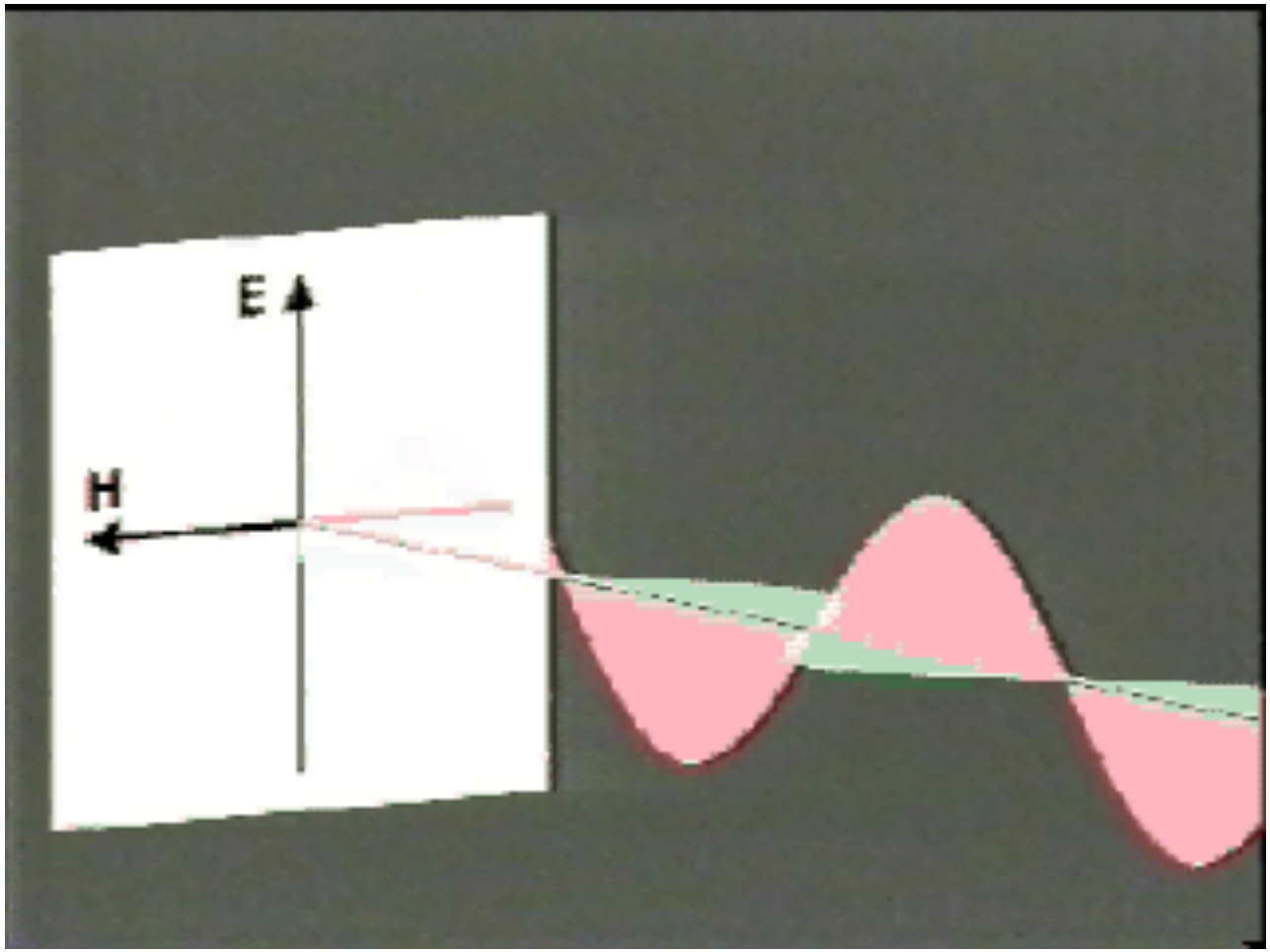
1. Geometrical optics

2. Physical optics, or

3. Quantum-mechanical optics

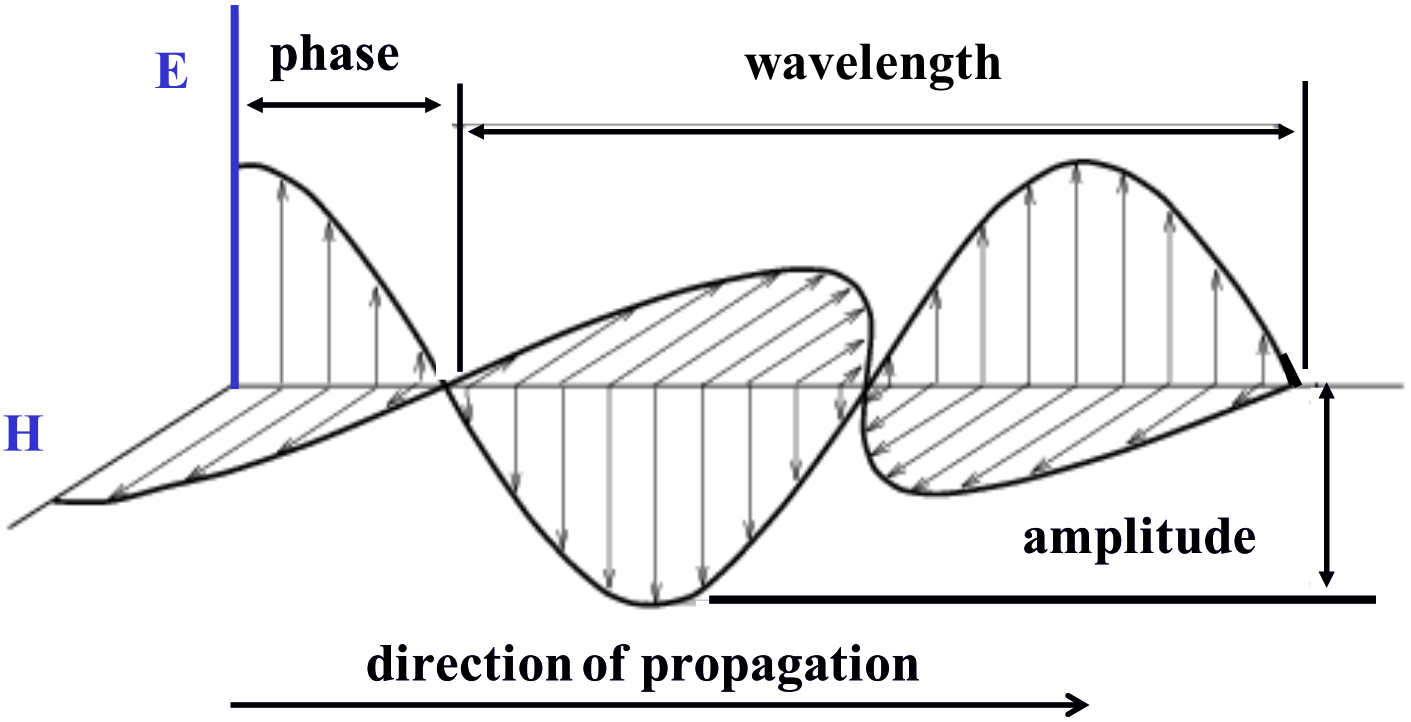
➔ **wave character**

Light as electromagnetic waves



Light as electromagnetic waves

Self-sustaining exchange of electric and magnetic fields



1. wavelength
2. direction
3. amplitude *E*
4. phase
5. direction of polarisation

The spectrum

Normal ambient light is a mixture of wavelengths, polarisation directions, and phases

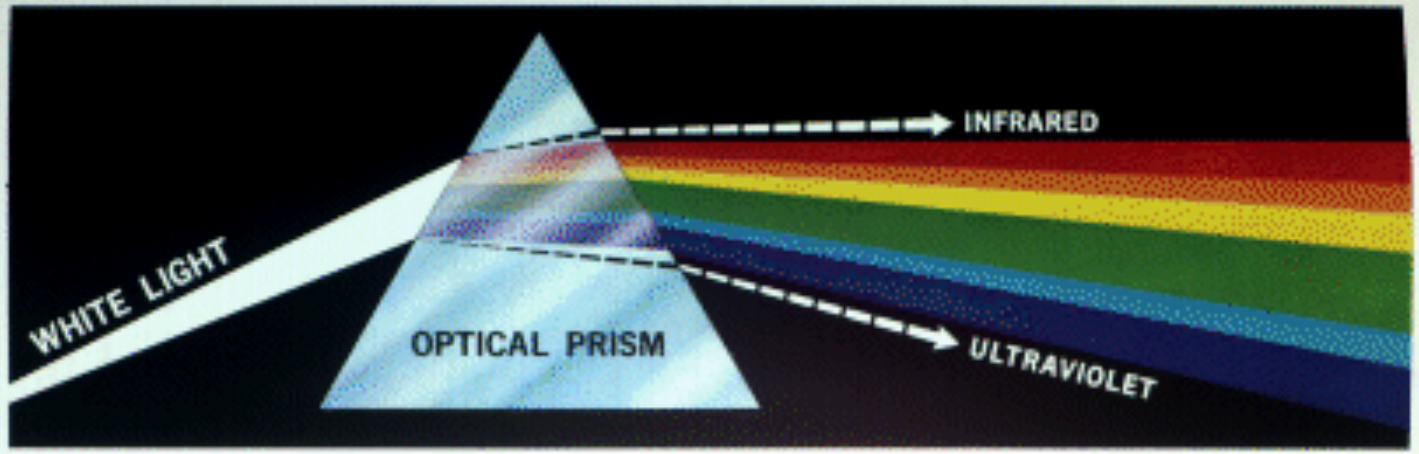
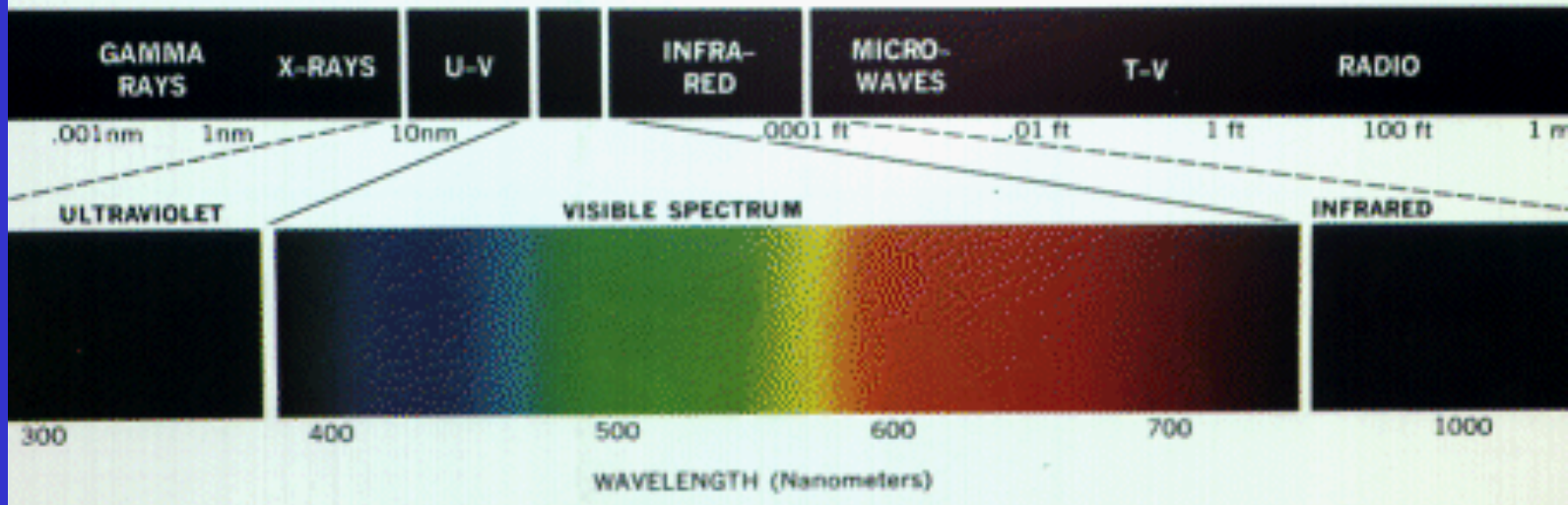


Plate I. Color spectrum seen by passing white light through a prism. (Courtesy of General Electric Co., Lamp Business Division.)



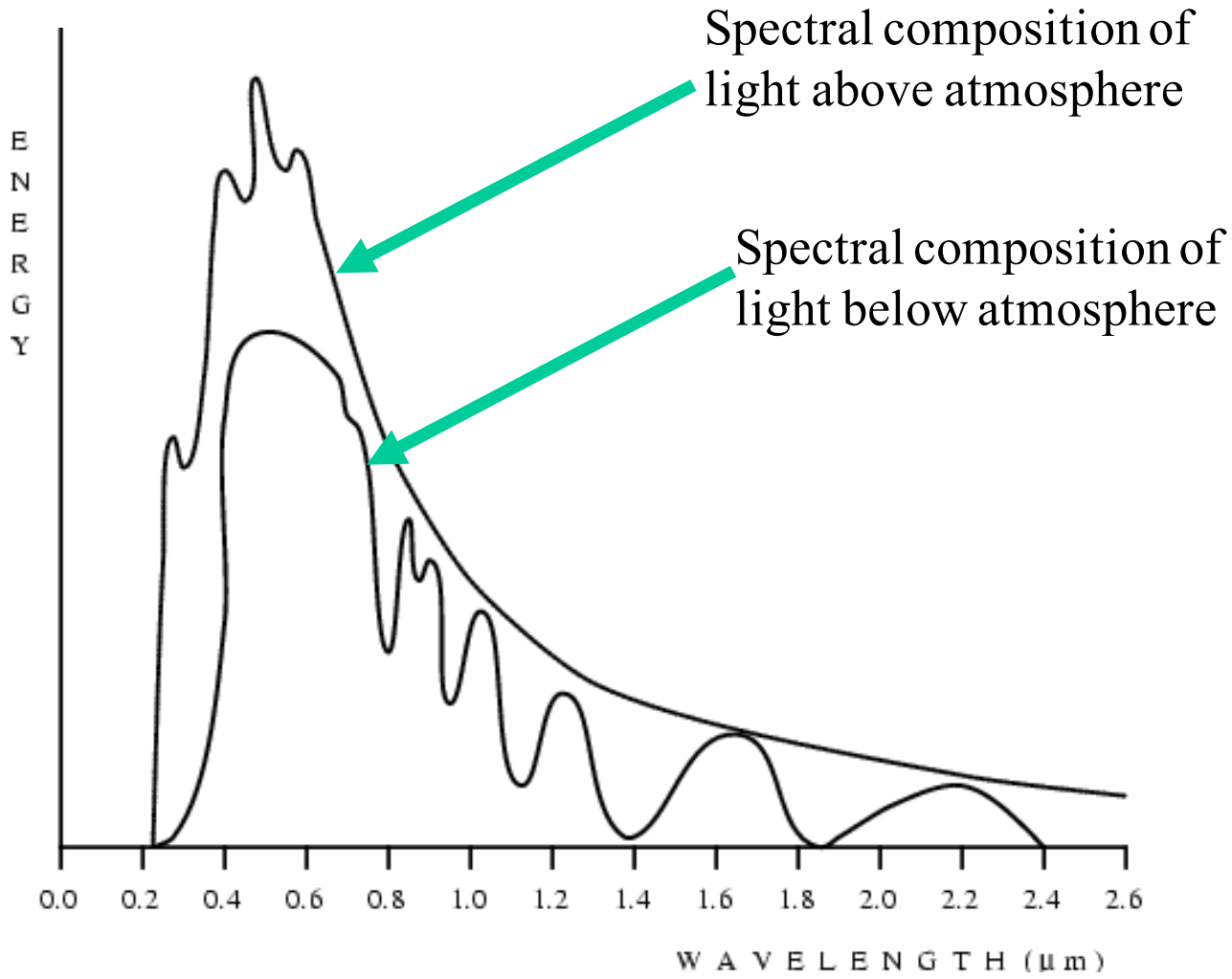
The visible range

Wavelength (in *nm*)	Colour
380 - 450	violet
450 - 490	blue
490 - 560	green
560 - 590	yellow
590 - 630	orange
630 - 760	red

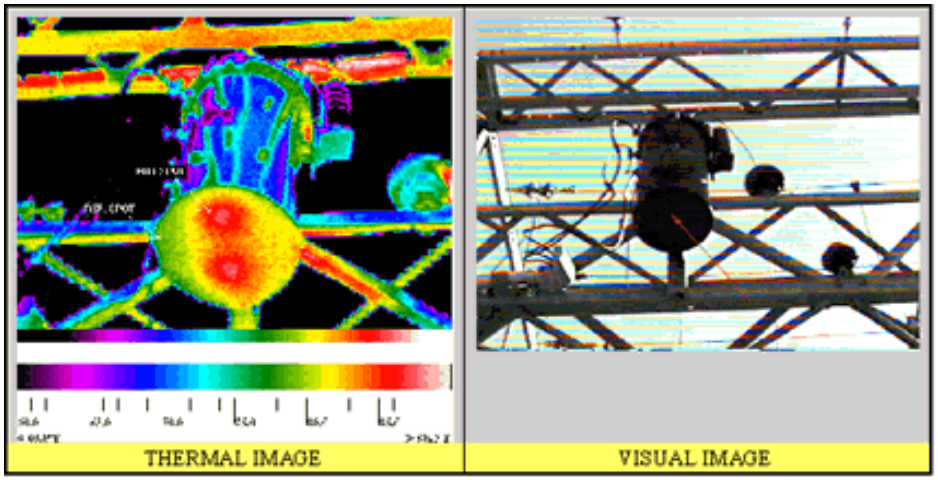
NOTE : Cameras may have different spectral sensitivities (i.e. also different from human vision)

The solar spectrum

Peaks around 500nm, hence human sensitivity for that part of the spectrum



Also cams for non-visible `light', e.g. infrared



Overheating of transformer coils, with far IR

Near infra-red (NIR) space image

NRG -> RGB for visualization (notice the strong reflection in the NIR for vegetation)

Interactions with matter

four types :

phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

+ diffraction

Interactions with matter

four types :

phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

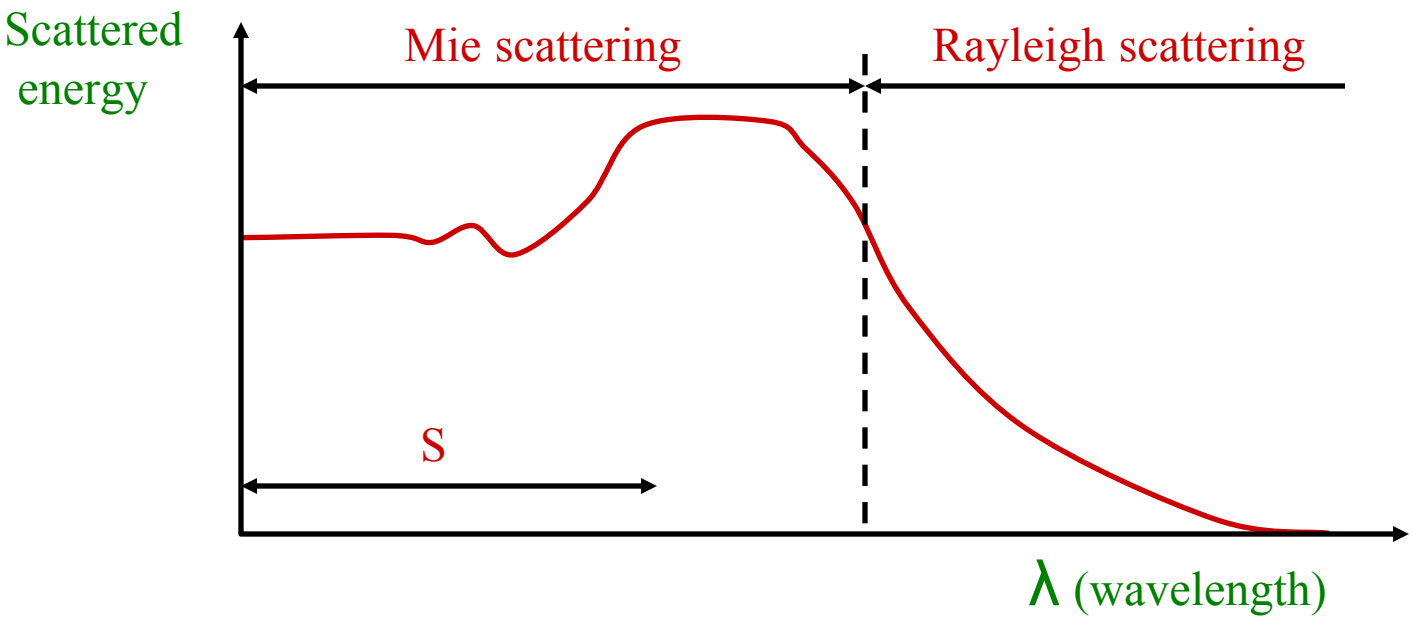
+ diffraction

Scattering

3 types depending on relative sizes of particles and wavelengths:

1. small particles: *Rayleigh* (strongly wavelength dependent)
2. comparable sizes: *Mie* (weakly wavelength dependent)
3. Large particles: *non-selective* (wavelength independent)

Wavelength dependence



Less haze in the infrared (long wavelengths -> little scatter)
Looking through clouds by radar (even longer wavelengths)
NOTE: without scatter we would wander mainly in the dark

Atmospheric showcase

Tyndall effect (blue sky)
Red, setting sun
Grey clouds

Coloured cloud from volcanic eruption

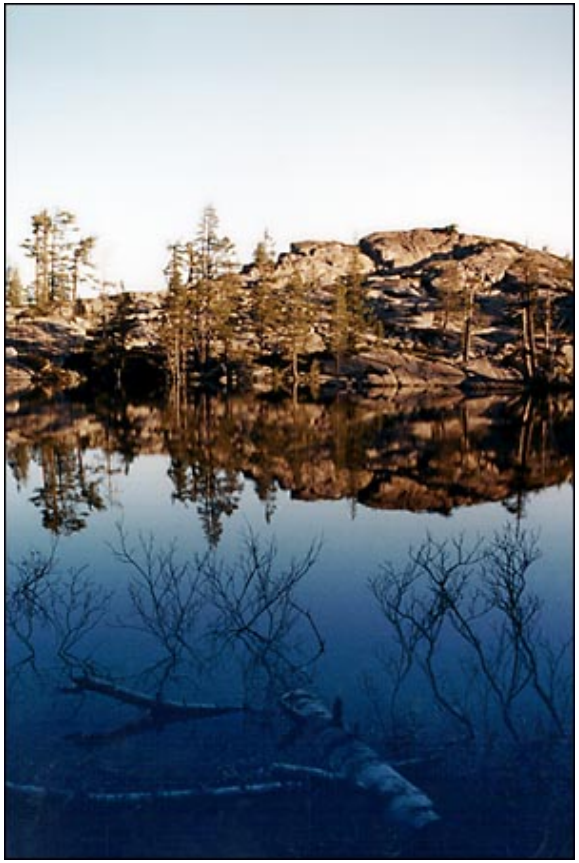
Interactions with matter

four types :

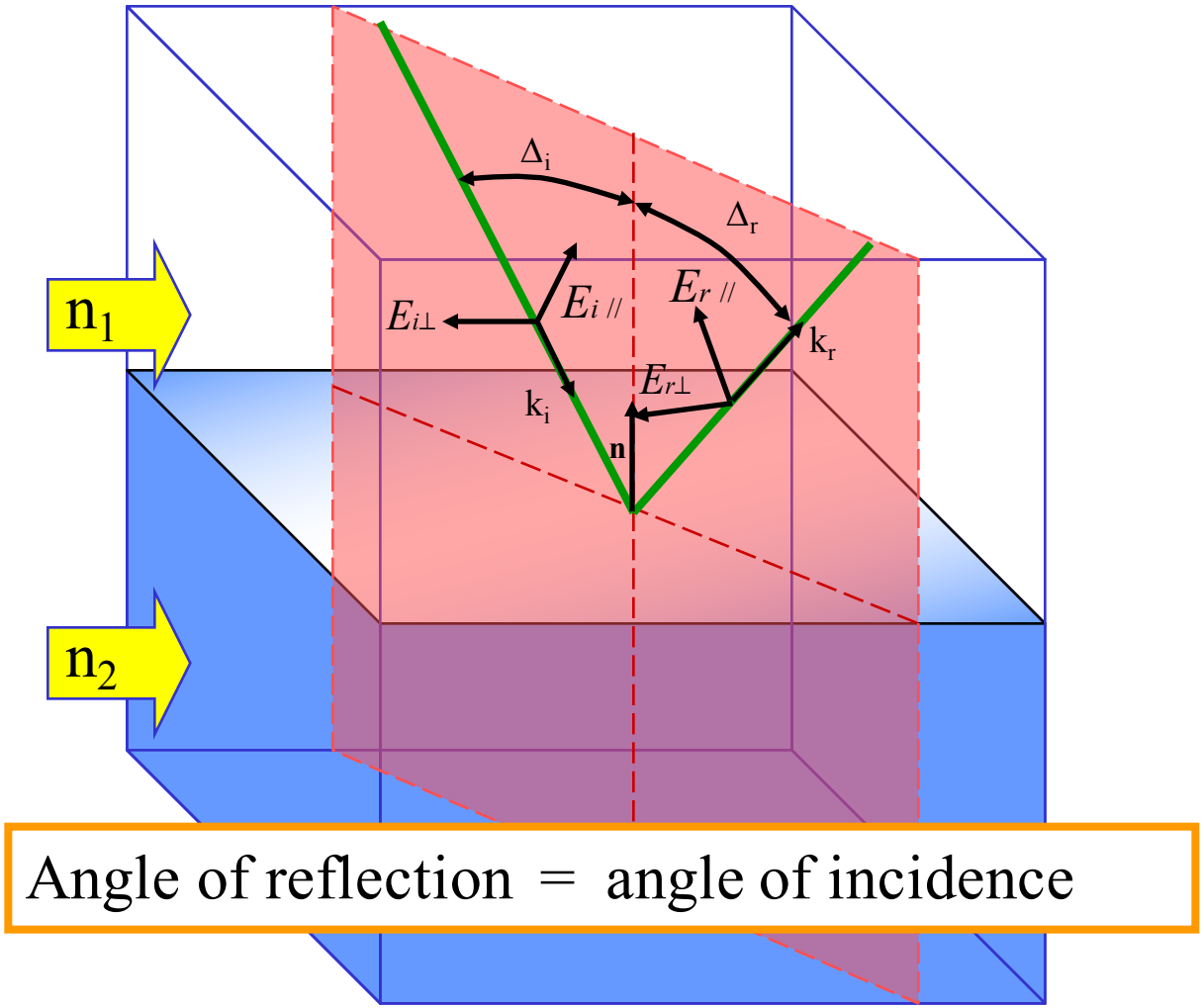
phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

+ diffraction

(Mirror) reflection

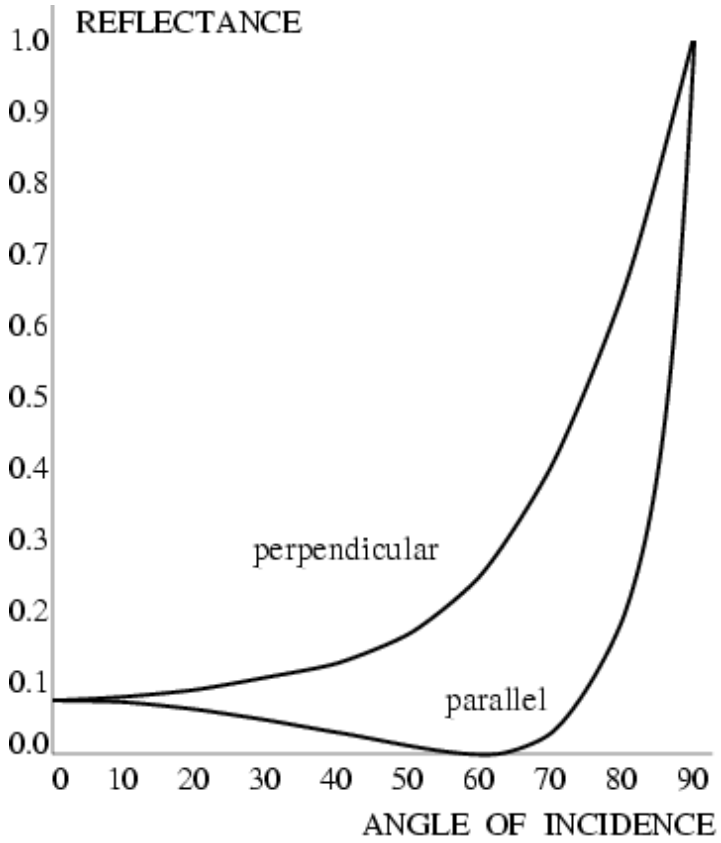


Reflection



Angle of reflection = angle of incidence

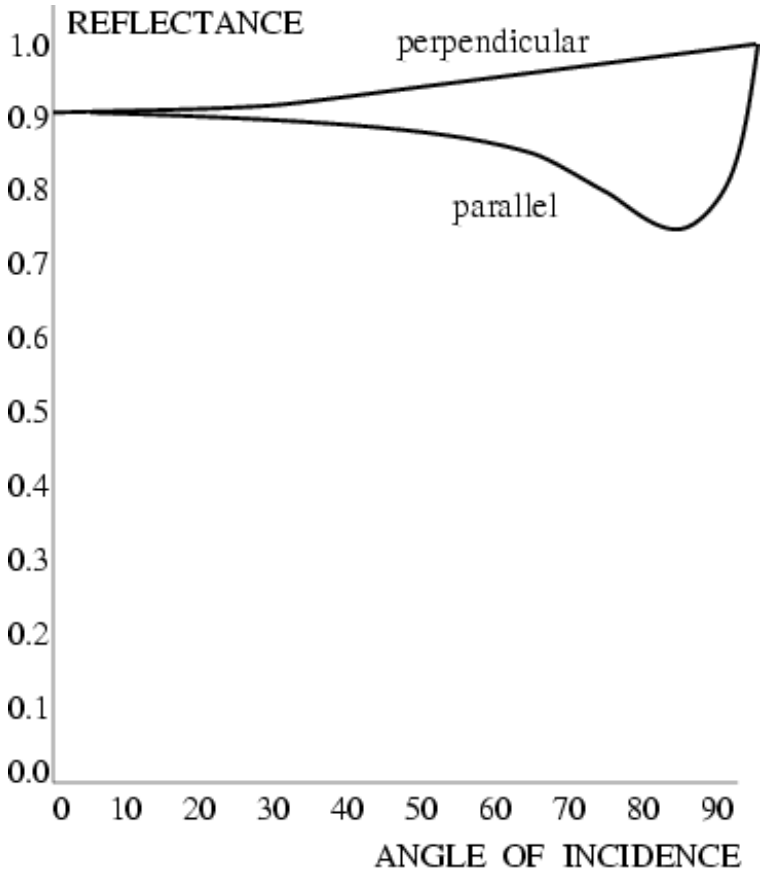
Reflection : dielectric



Polarizer at *Brewster angle*

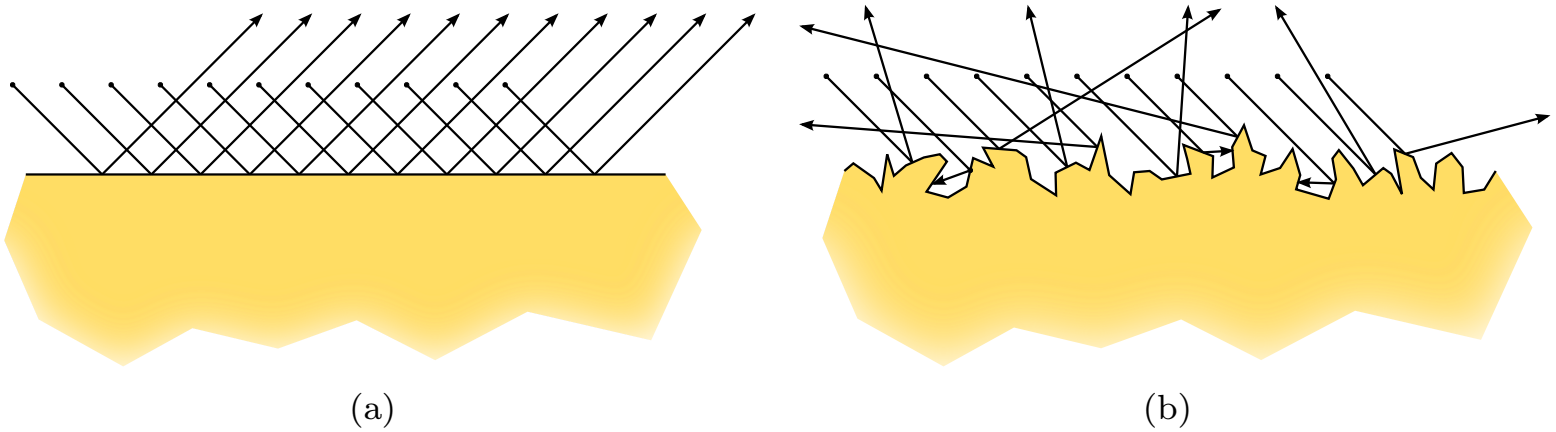
Full reflection at grazing angles

Reflection : conductor



strong reflectors (under all angles)
more or less preserve polarization

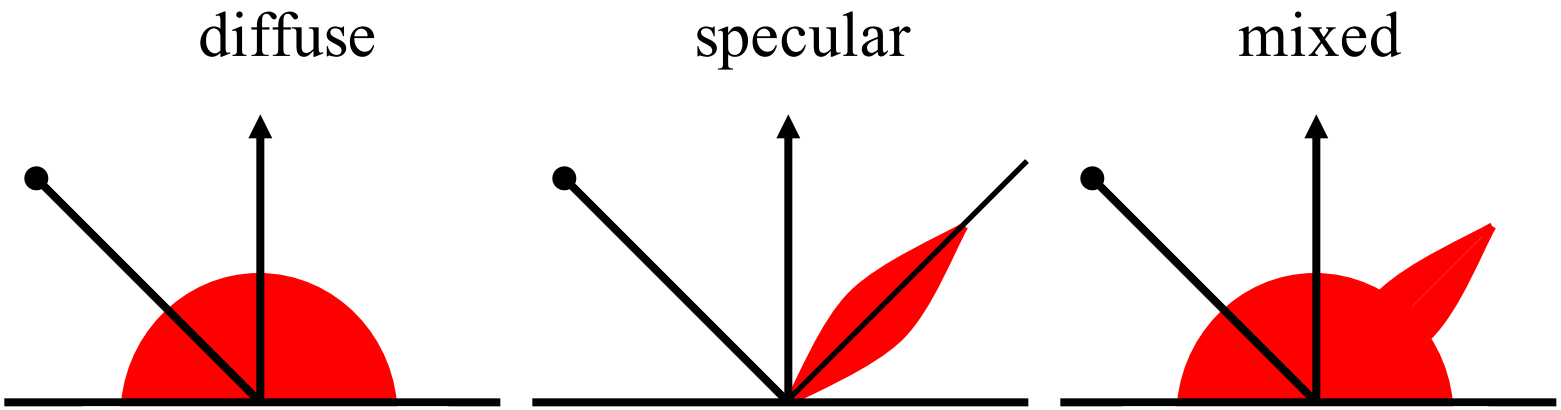
Roughness of surfaces leads to `diffuse' reflection



(a) Mirror or `specular' reflection, (b) diffuse reflection

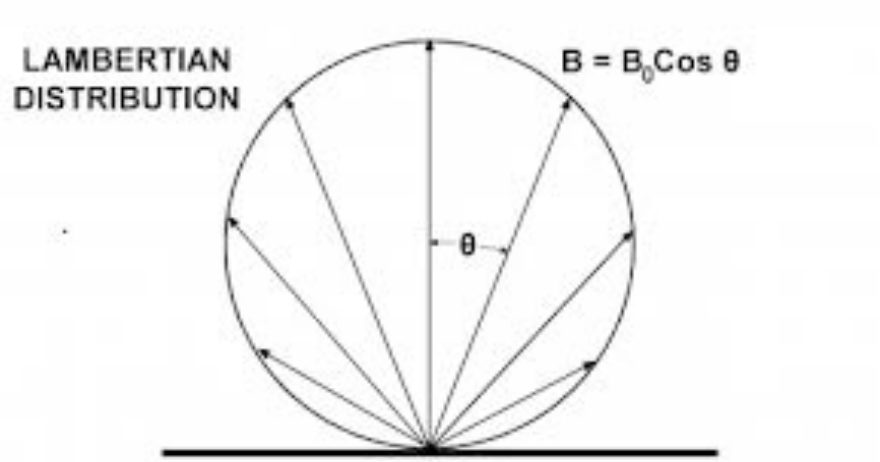
… and to mixed reflection for most real surfaces

three types of reflection :



Note : Lambertian example of diffuse reflection

Lambertian reflection



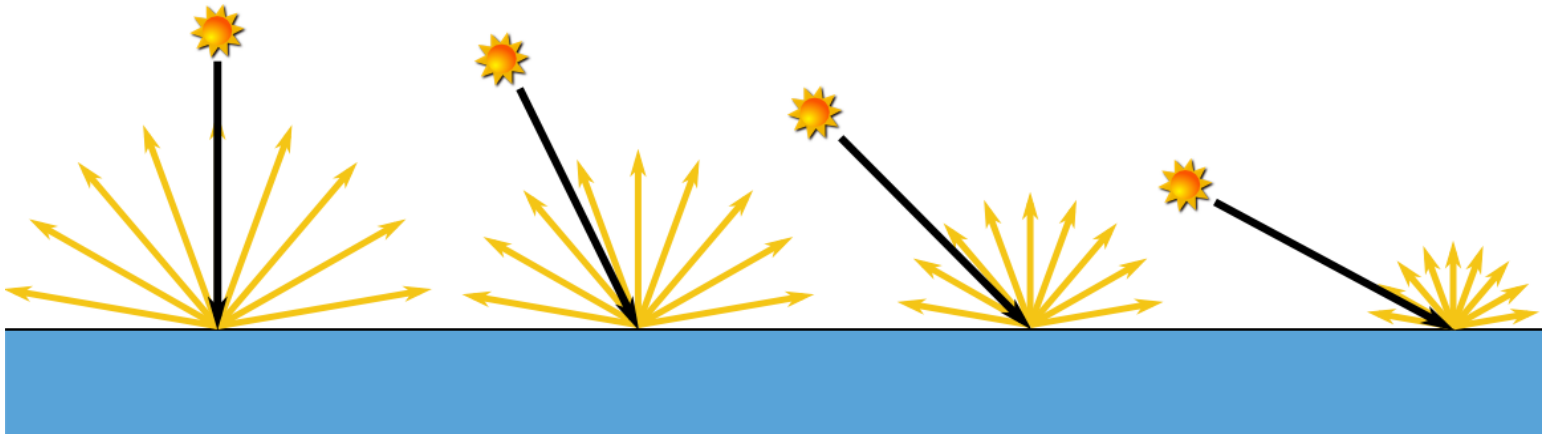
In case of Lambertian reflection the observed brightness only depends on the direction of the incoming light. If that's fixed the source radiates a flux that goes down with the viewing angle (with $\cos\theta$).

On the other hand, the same area on the camera corresponds to a larger surface area as one looks more obliquely ($1/\cos\theta$)

And therefore receives light from a larger surface area.

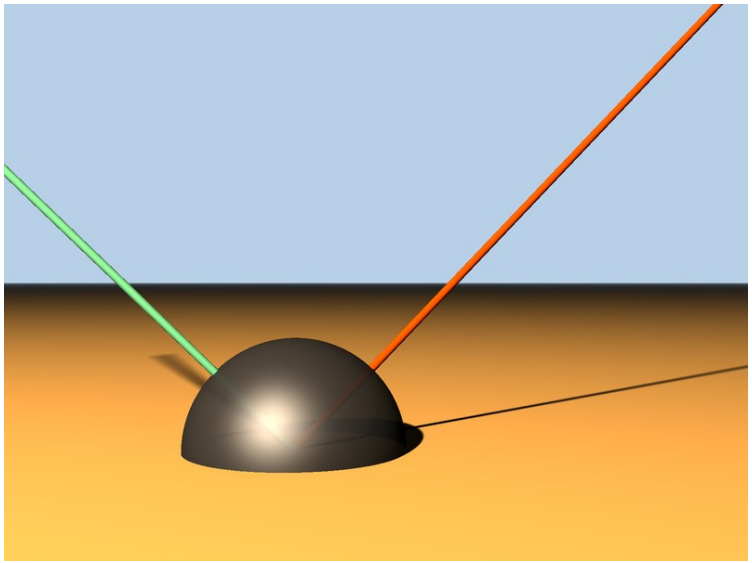
These two factors cancel out, which lets the surface look equally bright from all viewing directions.

Lambertian reflection

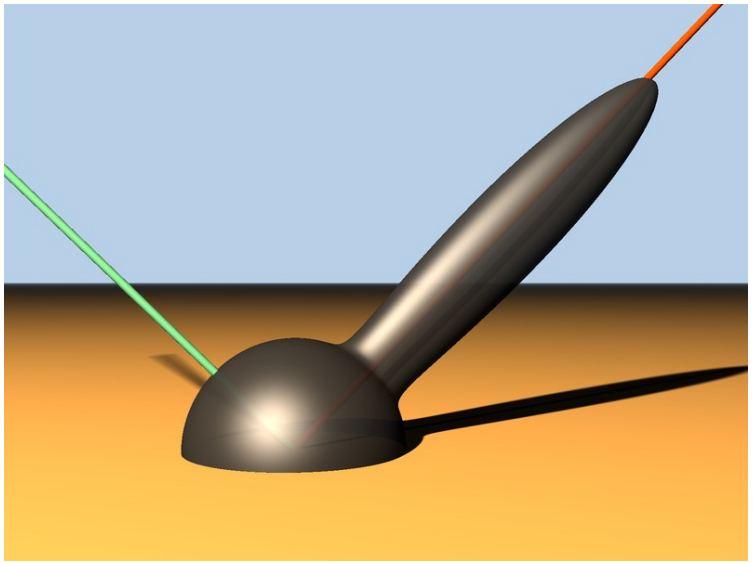


A Lambertian surface does ***not*** look equally bright independent of the lighting direction. Keeping the light intensity fixed but shining more obliquely lowers the amount of light hitting the surface per unit area. Thus there is less light to be reflected and, hence, the surface looks now equally darker from wherever we look.

BRDFs specify reflection



Diffuse example

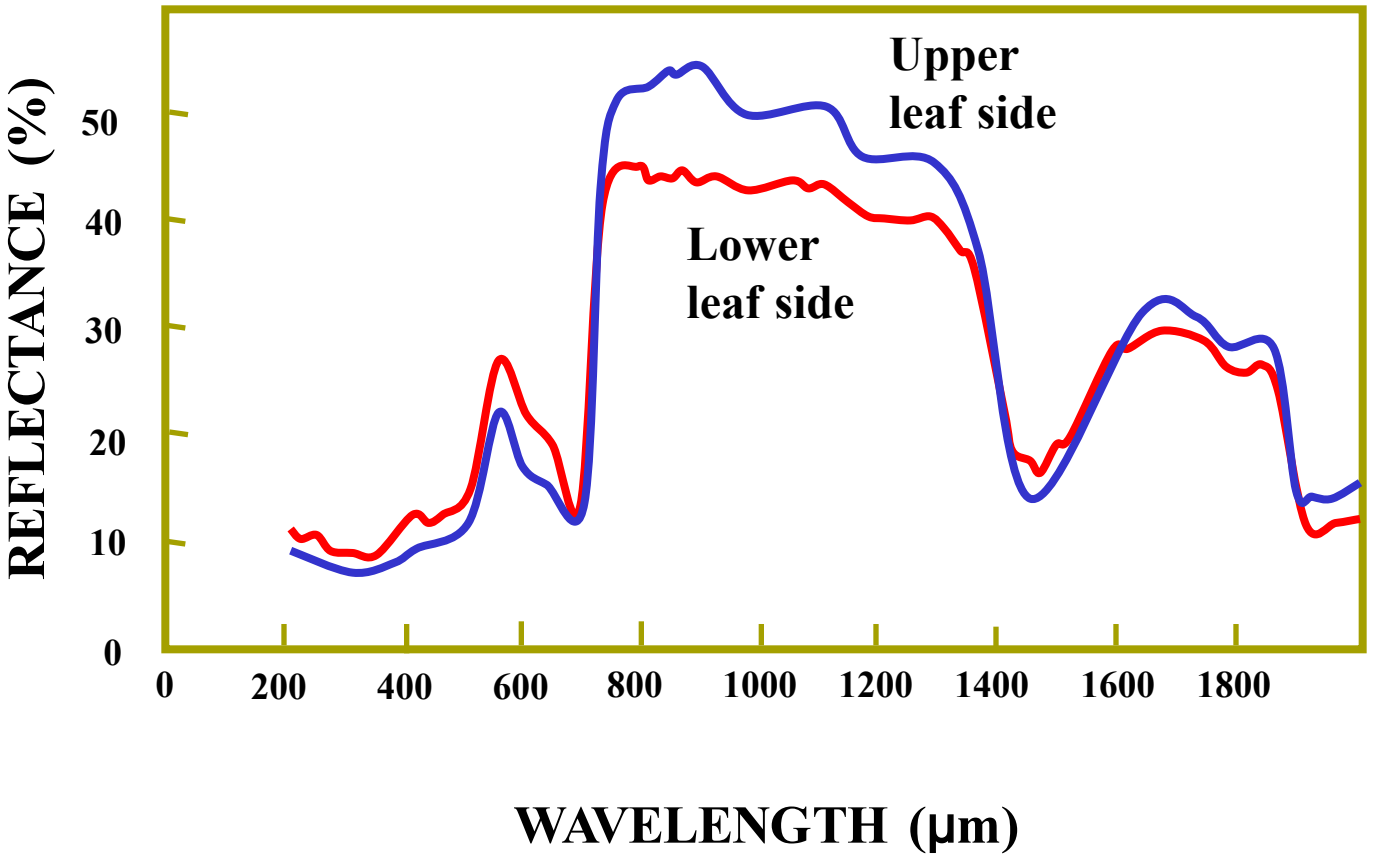


Mixed example

A BRDF is a 4D function, specifying how much light is being reflected in a direction (red lines) for light coming in from a direction (green lines)

BRDFs are still hard to measure in practice

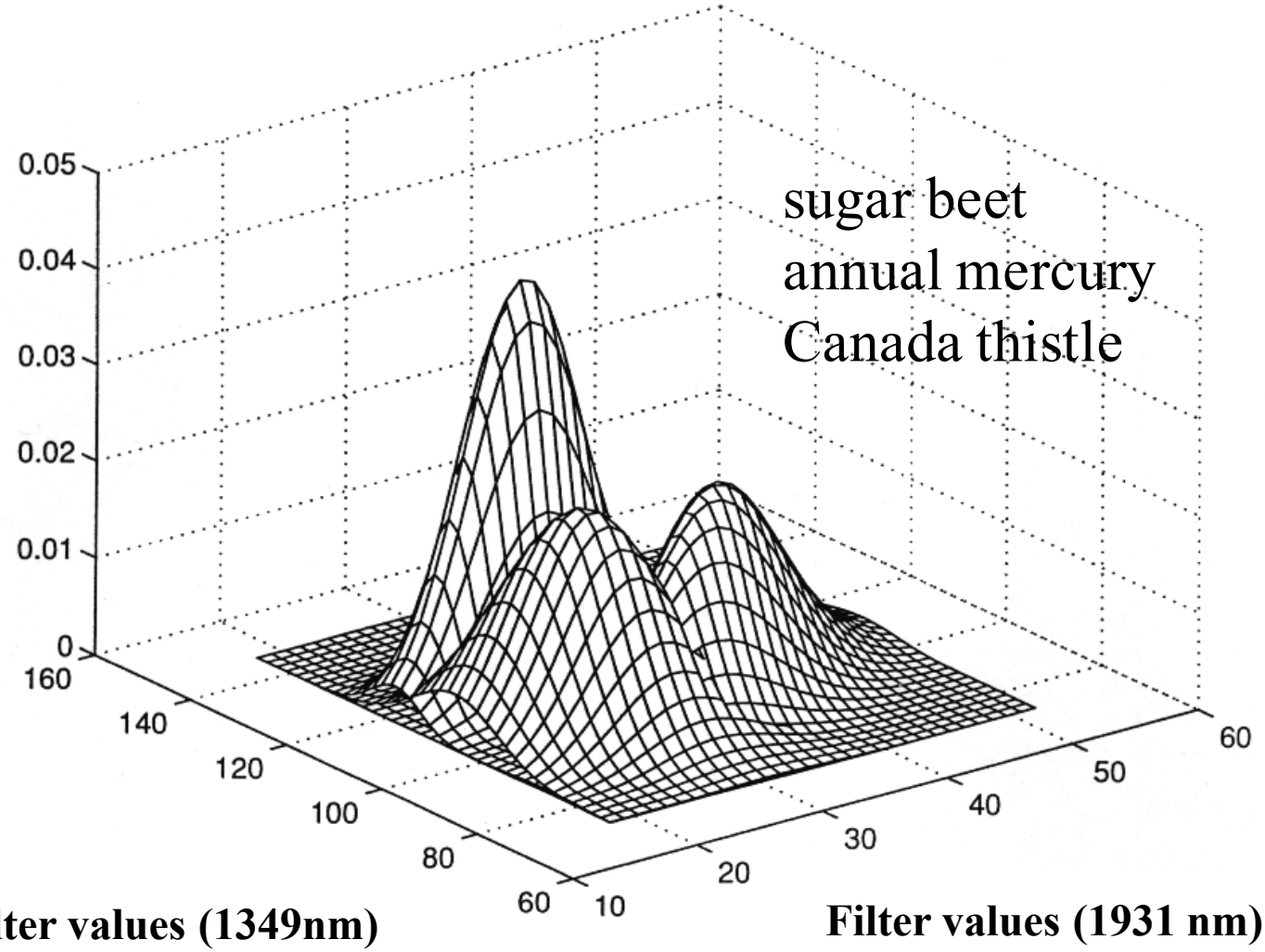
Spectral reflectance e.g. vegetation



Computer Vision

Spectral reflectance: ex. app weed detection for Reduced (selective) herbicide spraying

Gaussian probabilities for different classes



Ideally: spectral BRDF at all points known

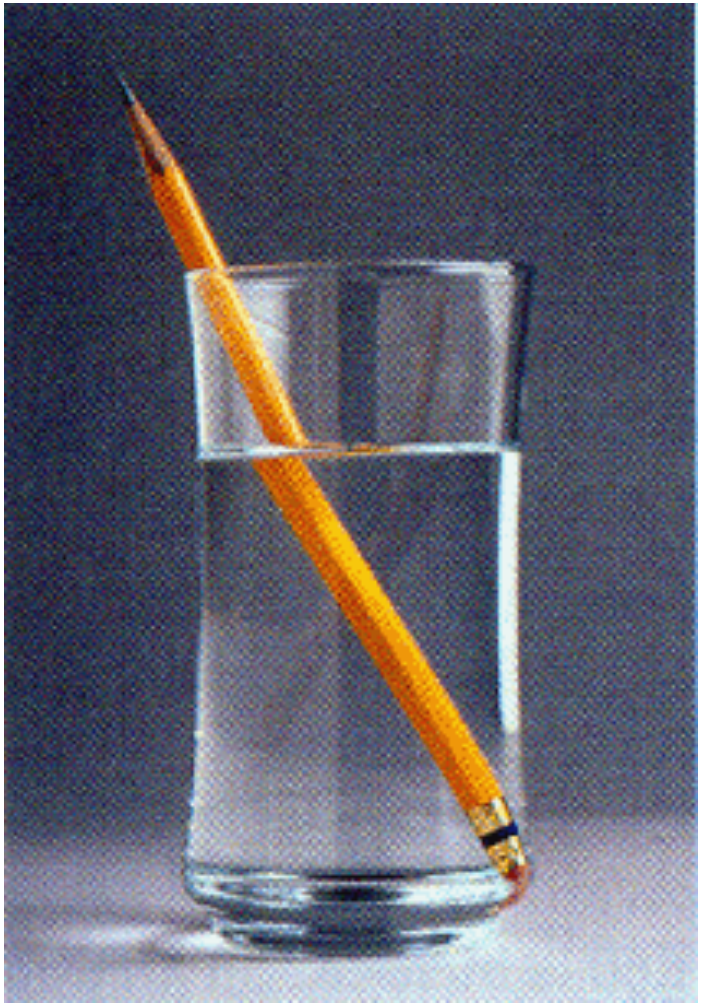
Interactions with matter

four types :

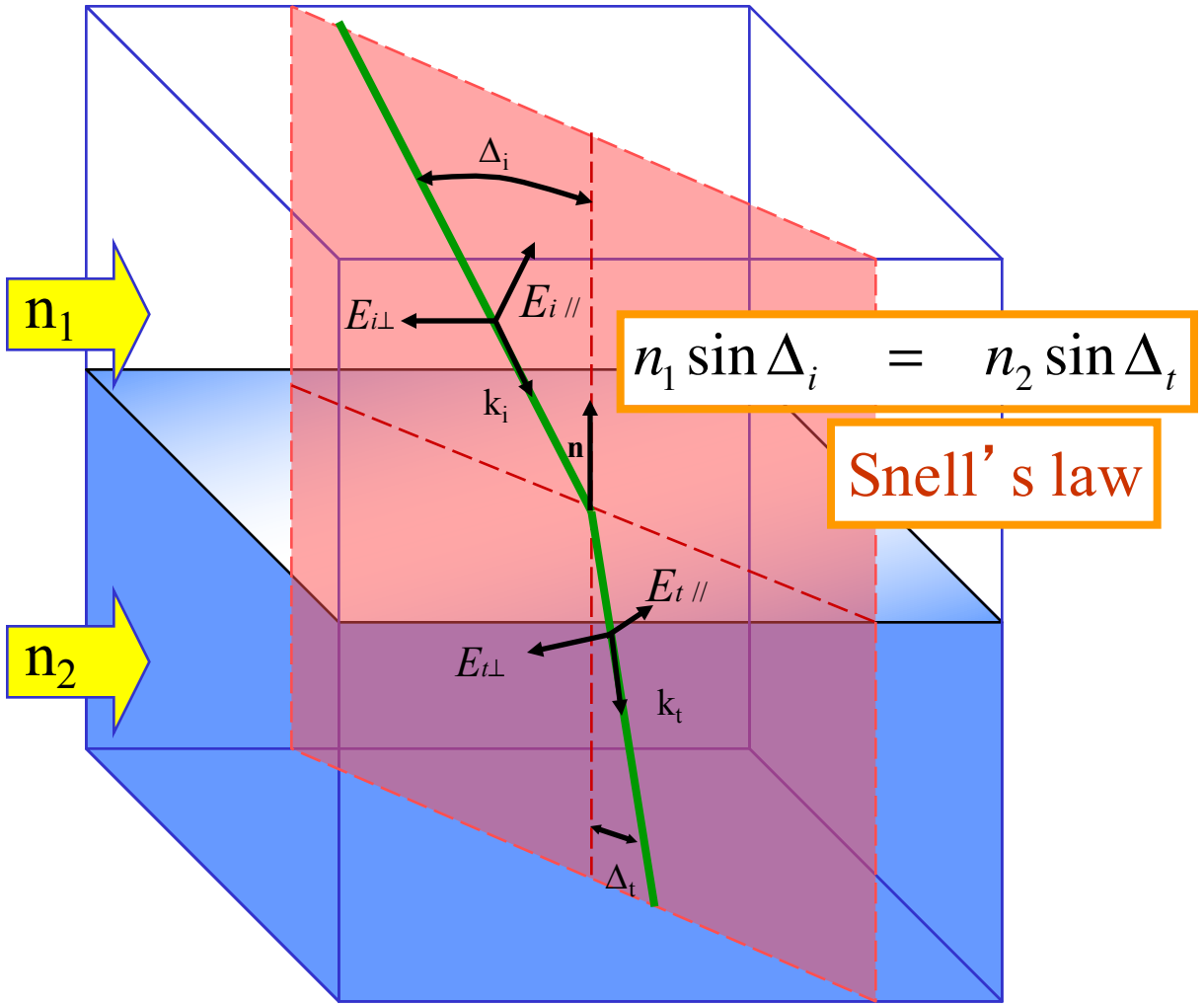
phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

+ diffraction

Refraction



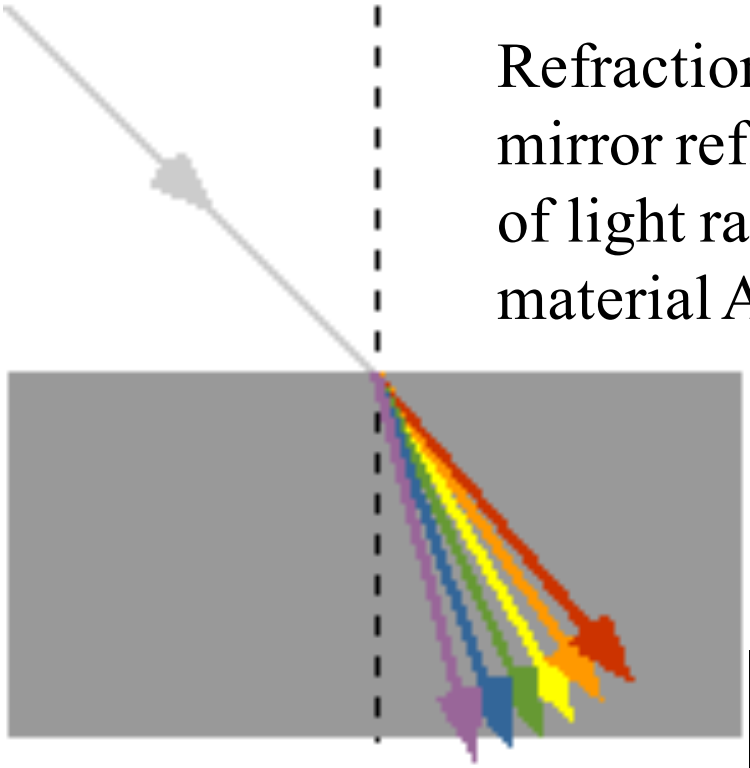
Refraction



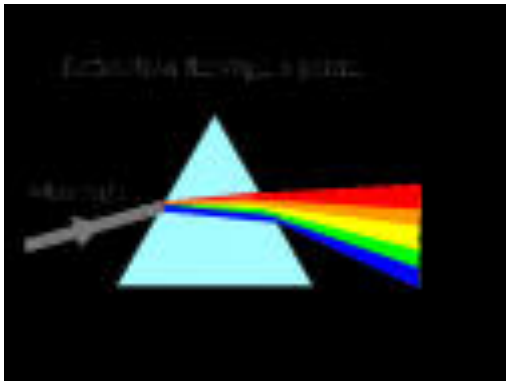
$$n_1 \sin \Delta_i = n_2 \sin \Delta_t$$

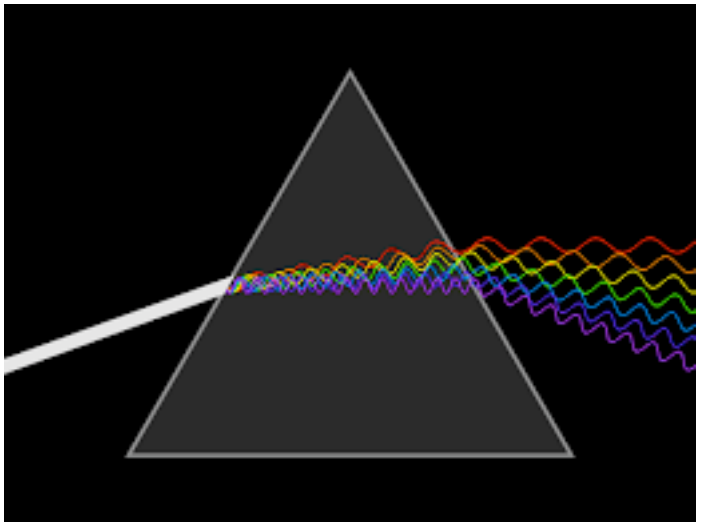
Snell's law

Dispersion



Refraction is more complicated than mirror reflection: the path orientation of light rays is changed depending on material AND wavelength !!!





Interactions with matter

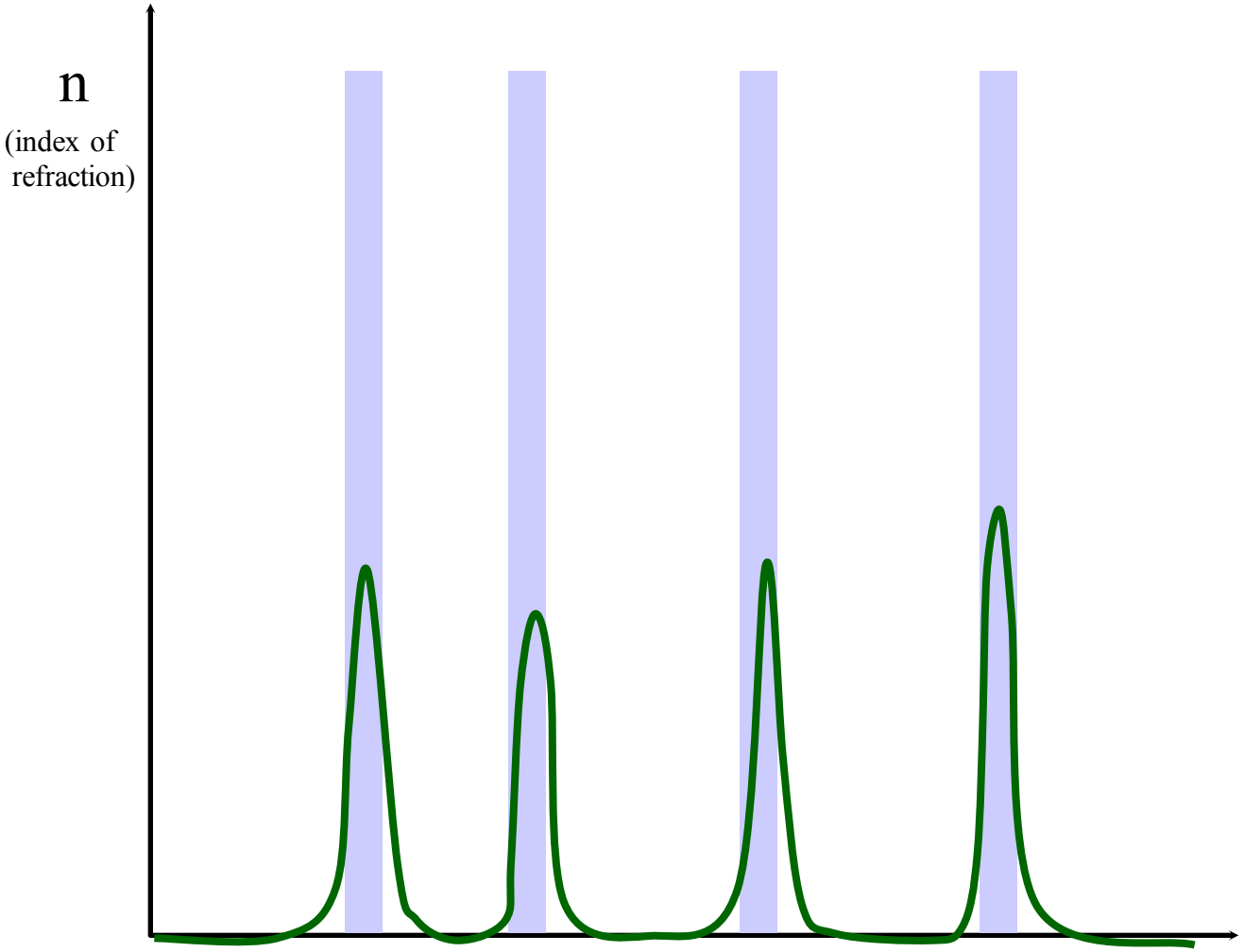
four types :

phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

+ diffraction

Absorption

Dissipation of wavelengths specific for the medium



Based on resonance frequencies of molecules -> peaks

Holes in sky light spectrum observed by Fraunhofer

… it needs brains …

Perception

Knowledge about HVS is important :

- Image quality in interactive systems
- HVS limitations suggest applications
- Might suggest ways to go about

human vision is much more than a bottom-up process of subsequent signal processing steps

Perception

Knowledge about HVS is important :

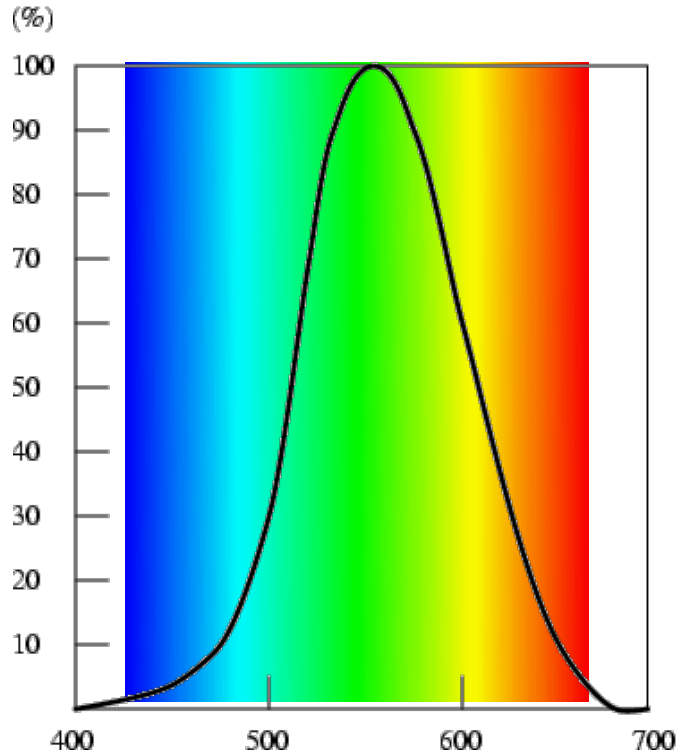
- Image quality in interactive systems
- HVS limitations suggest applications
- Might suggest ways to go about

human vision is much more than a bottom-up process of subsequent signal processing steps

current AI or deep learning techniques bring us closer both methodologically and in terms of performance (but still lacks feedback)

The perception of brightness

- Luminous efficiency function : relates radiometry & photometry



- C.I.E.(Commission Internationale de l'Eclairage) → standards

Link radiometry-photometry (Watt to lumen)

Photometry: subjective impressions
Radiometry: objective, physical measurements

at 555 nm : 1lm = 1/683 W = 1.46 mW

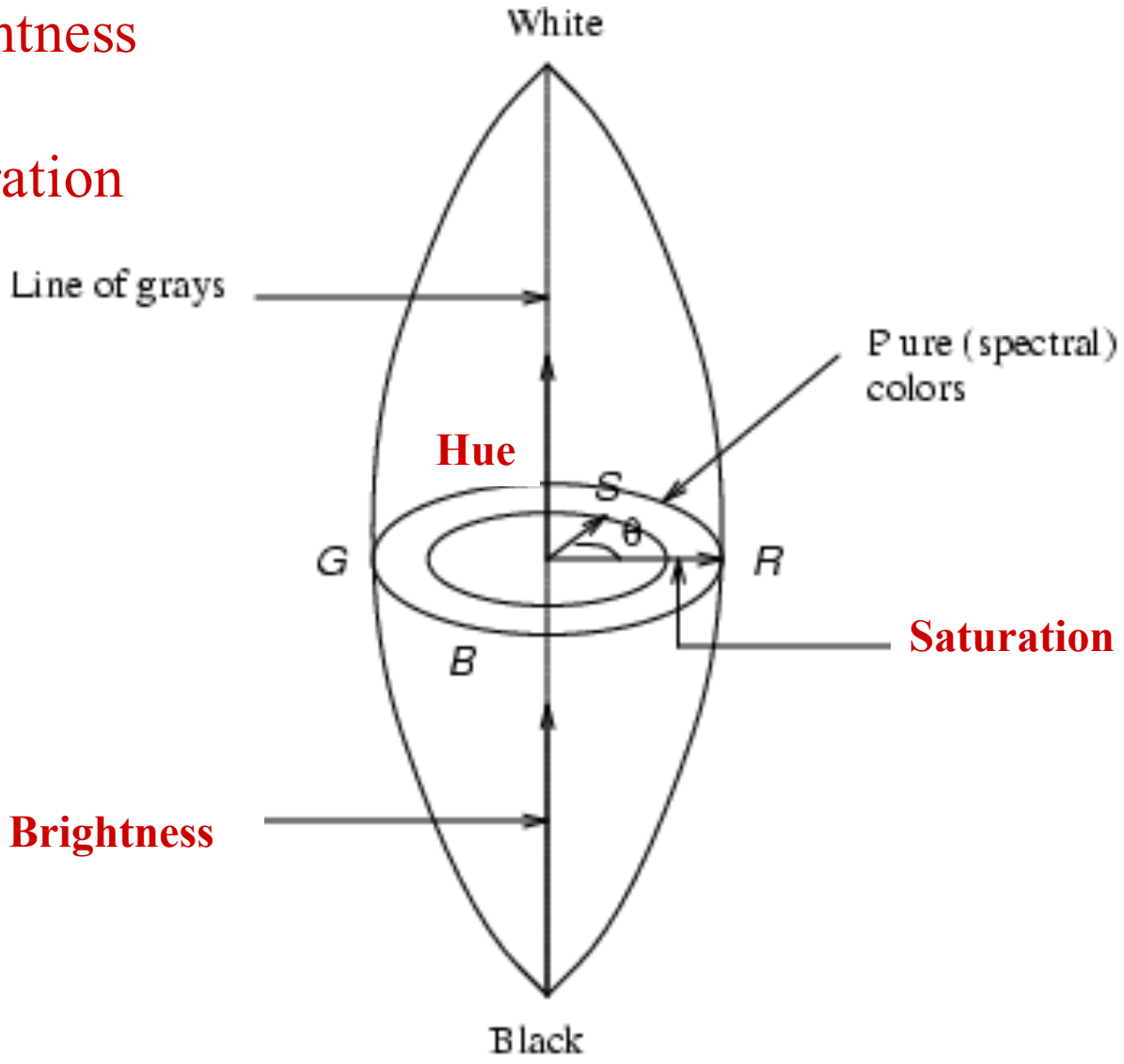
for light with spectral radiant flux $c(\lambda)$

$$l = k\int_{\lambda=0}^{\infty} c(\lambda)v(\lambda)d\lambda$$

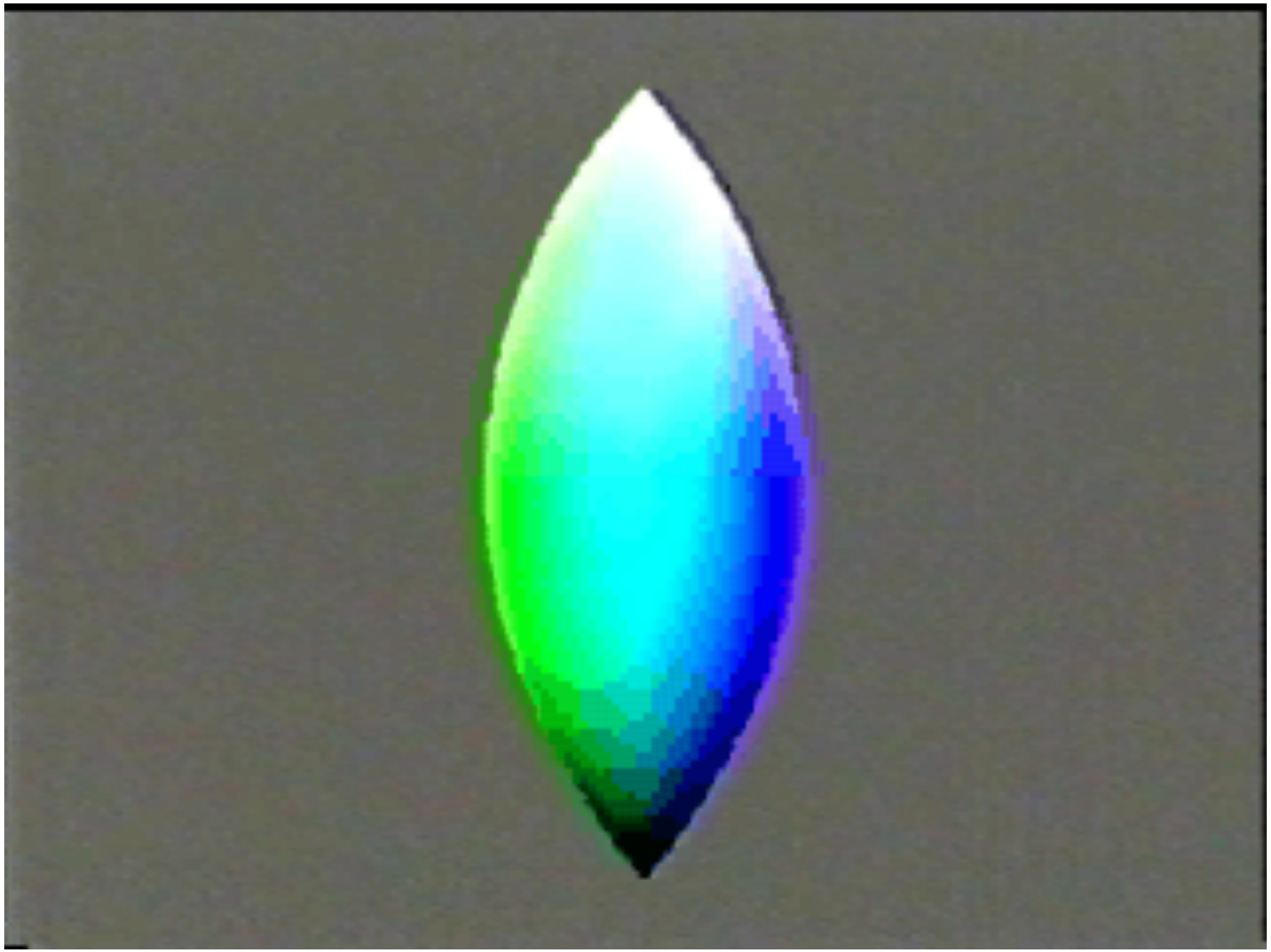
with k is 683 lumens/watt

The perception of colour

1. brightness
2. hue
3. saturation



The perception of colour



The history of colour science

Newton → spectrum

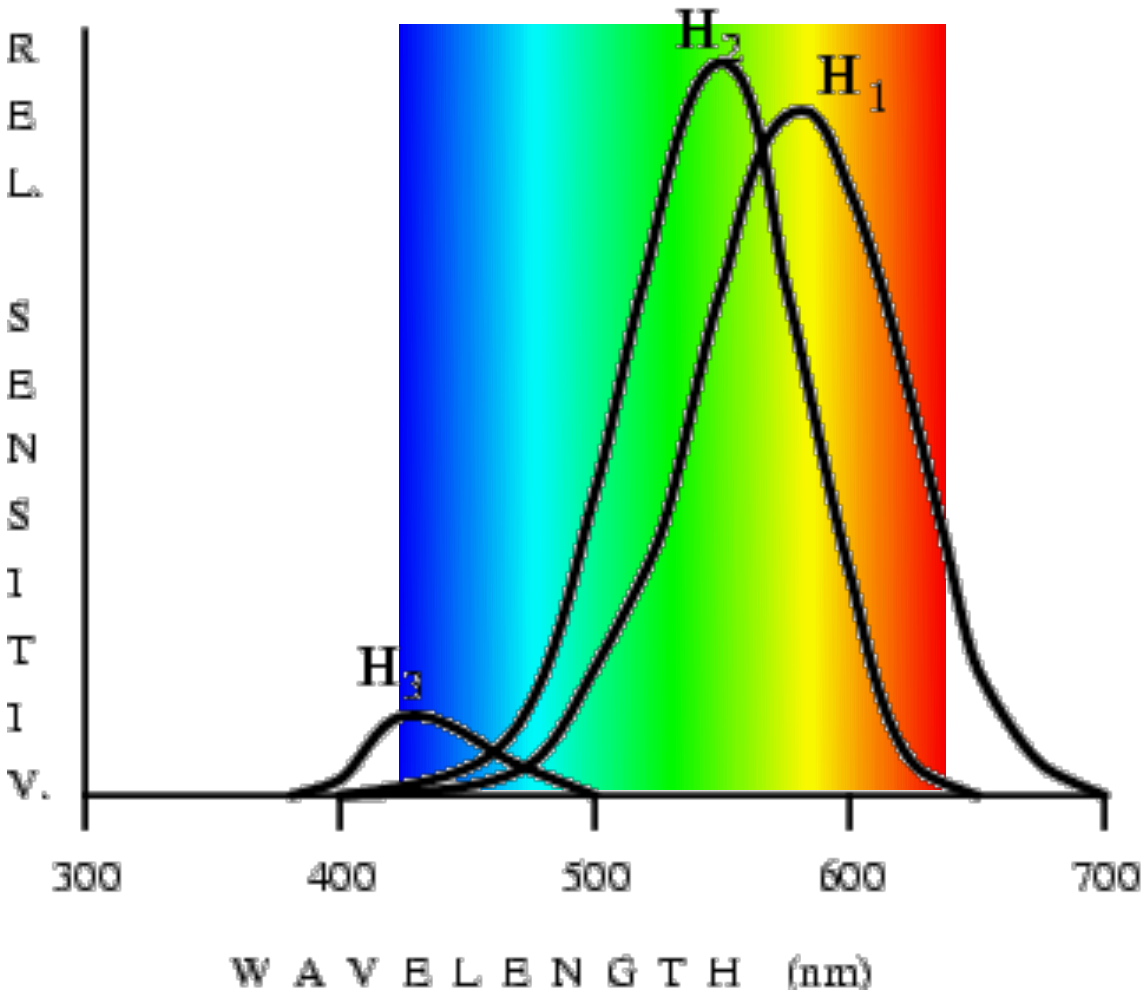
Young → tristimulus model

later : physiological underpinning :

3 cone types

The retinal cones

3 types : blue, green, yellow-green

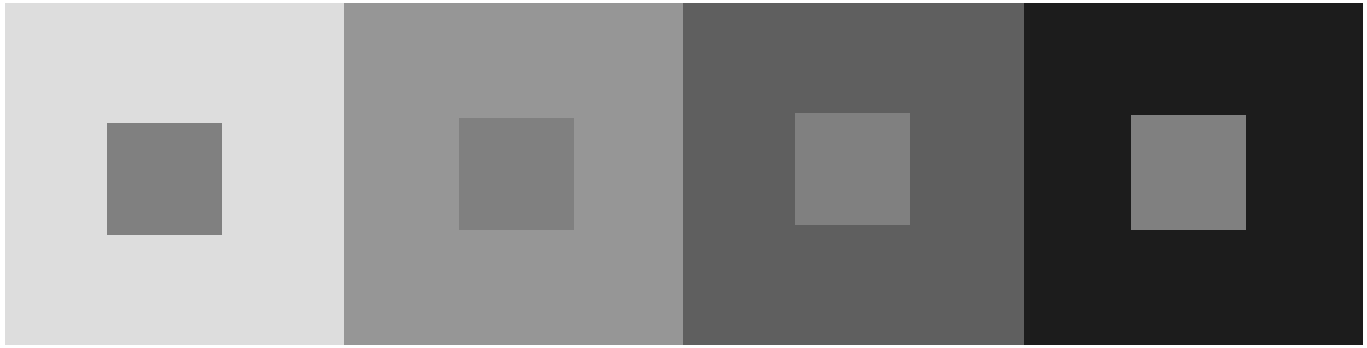


The perception of contrast

relative luminance (context) is the key

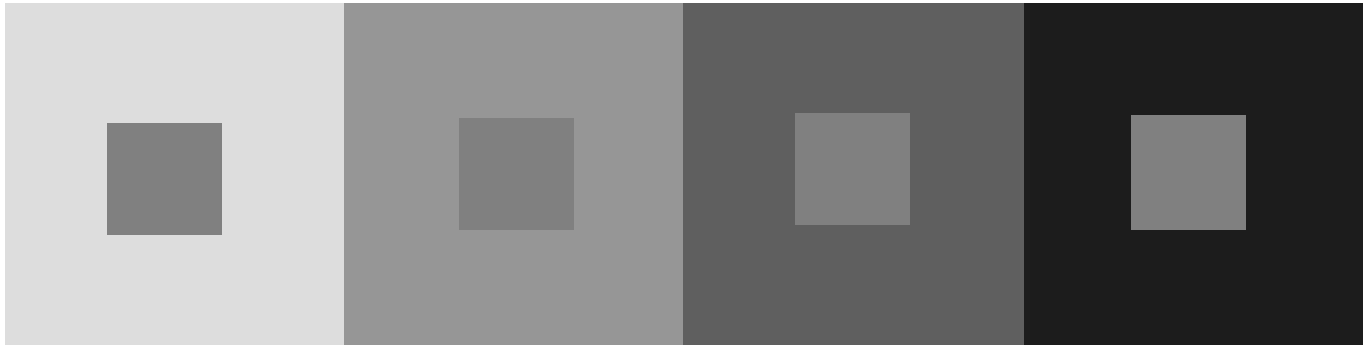
The perception of contrast

relative luminance (context) is the key



The perception of contrast

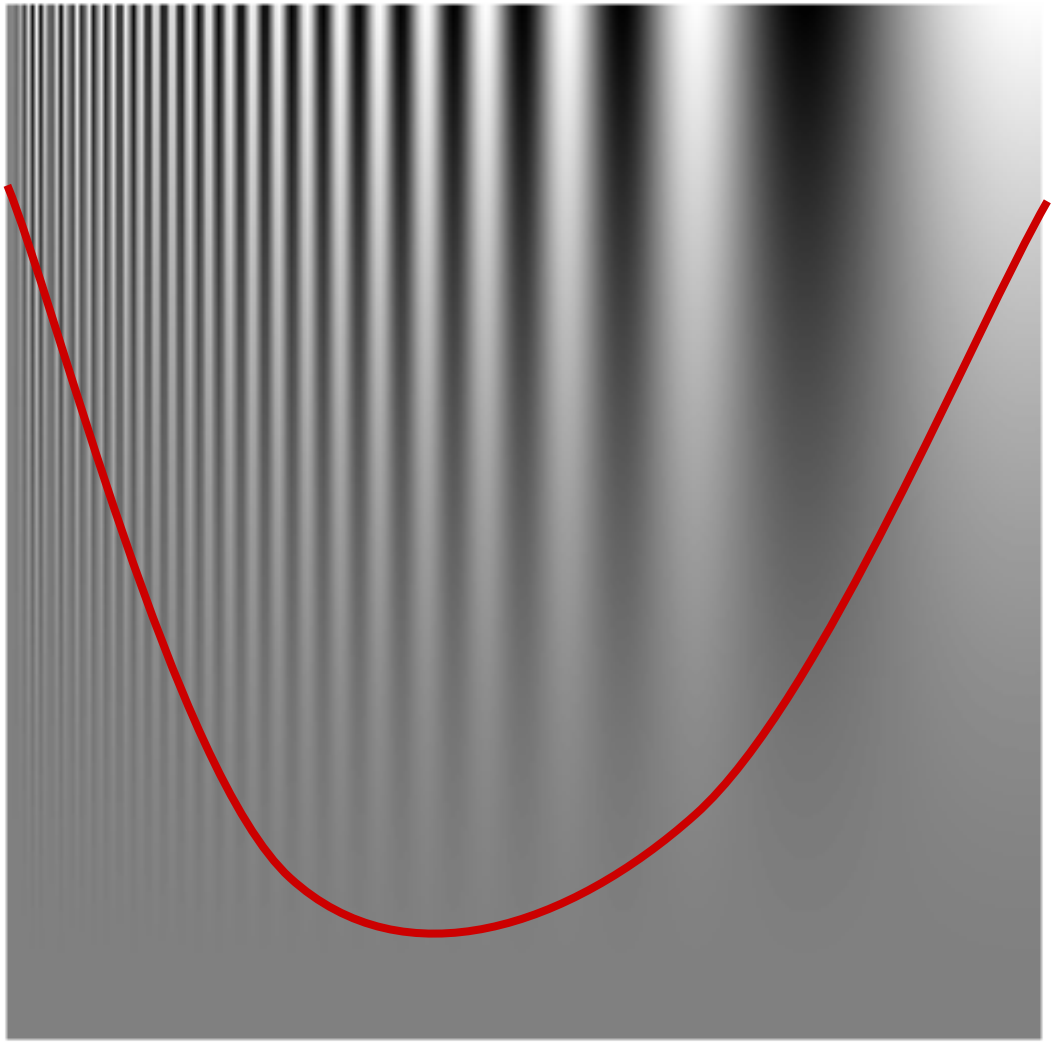
relative luminance (context) is the key



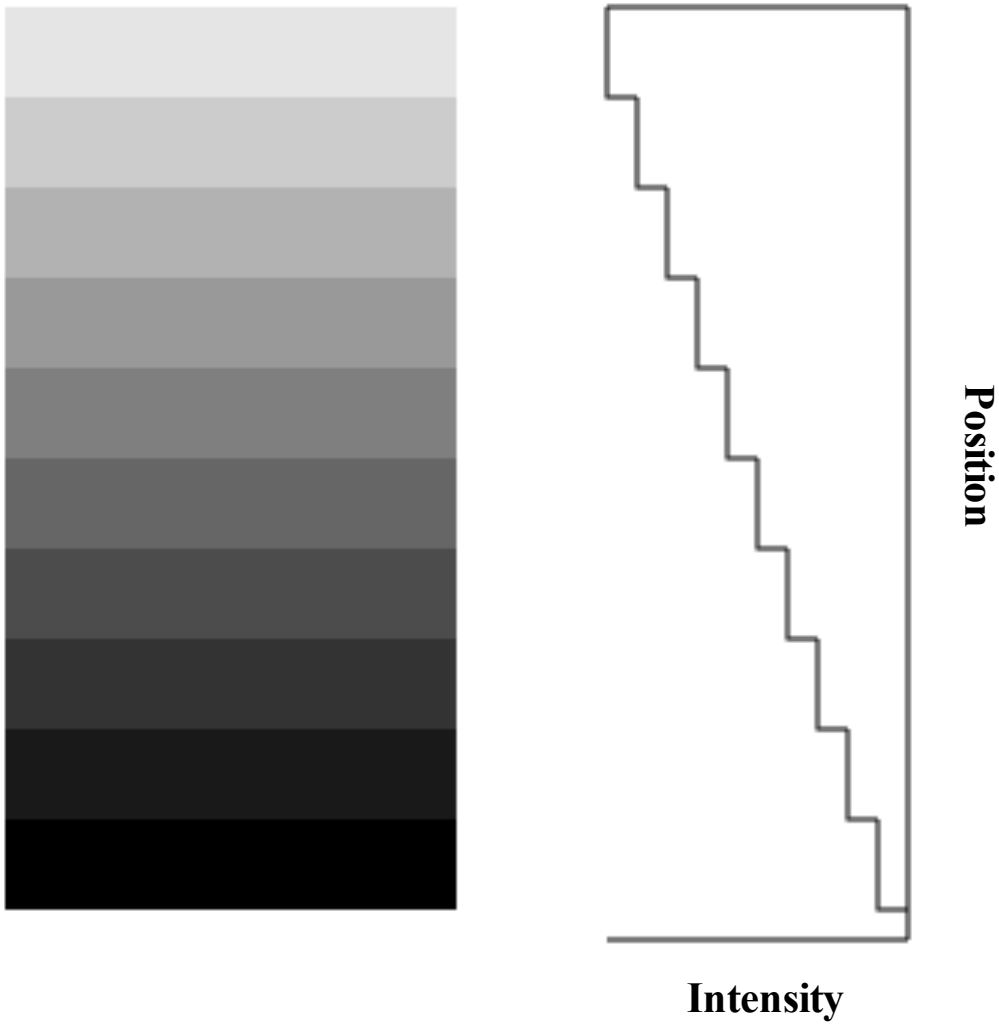
Weber's law : $$\frac{\Delta l}{l_{back}} \approx {}^{c}W$$

Relevant for *just noticeable differences (jnds)* and beyond

The perception of spatial frequency

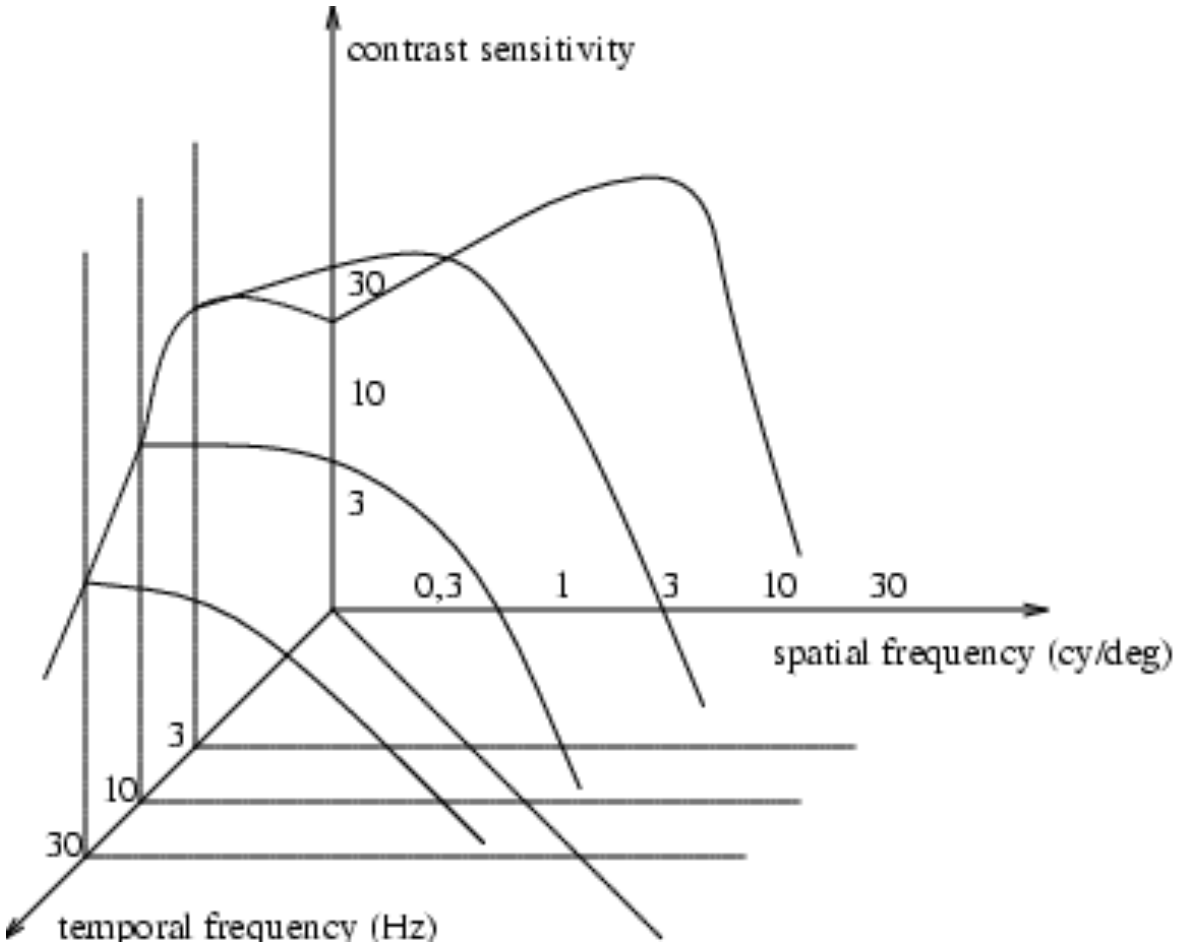


Mach bands



explained by MTF

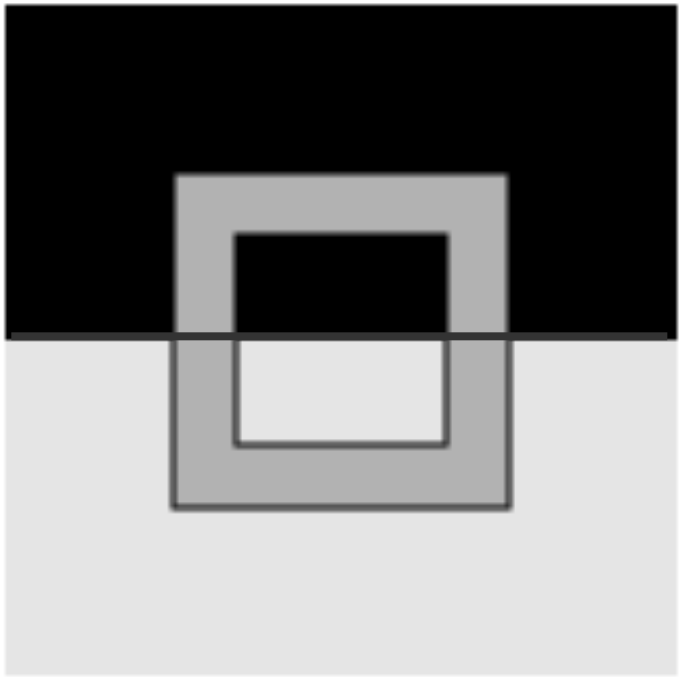
The perception of spatio-temporal frequency



bandpass along both axes

You ain't seen nothing yet

there is literally more than meets the eye,
i.c. a lot of massively parallel processing



Kanisza (filling in)

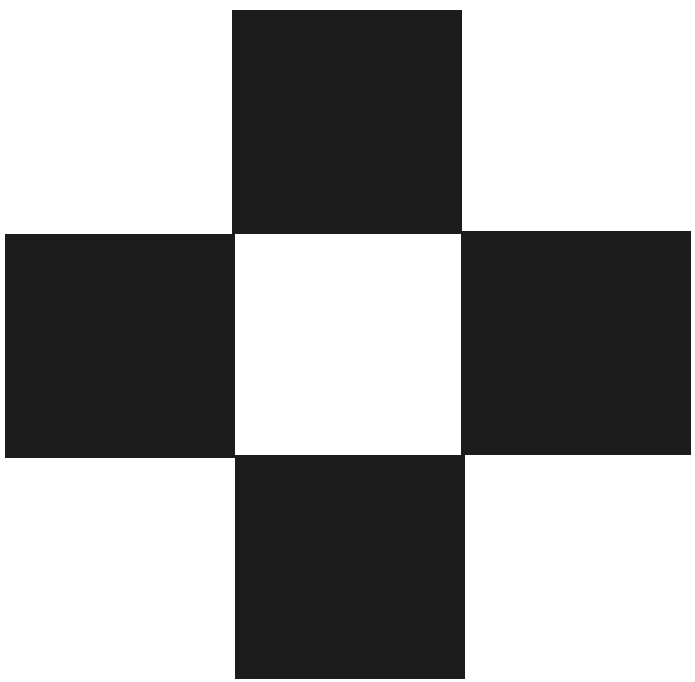
Fill-in : averaging of perceived contrast at edges over regions possibly obtained via extrapolation of the edges

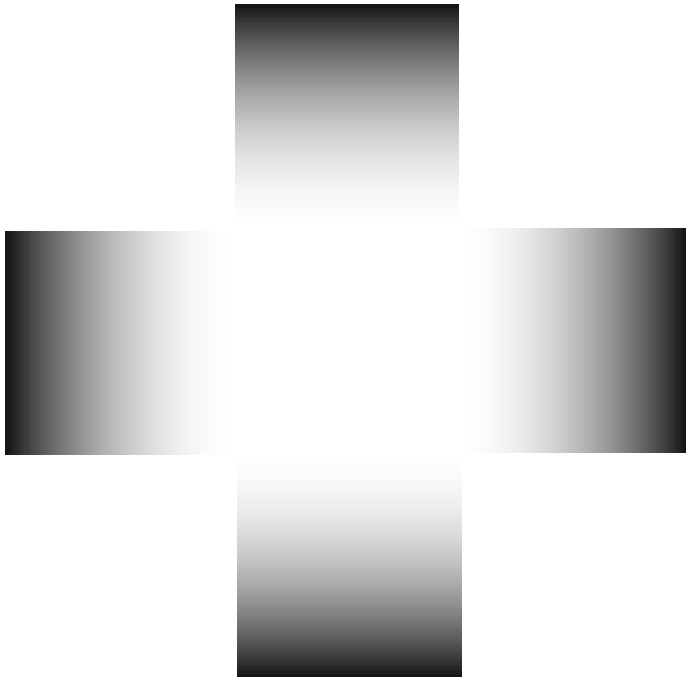
Computer Vision

Computer Vision

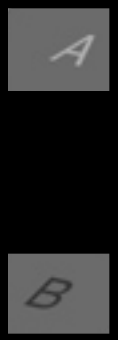
Our visual system eagerly looks for known patte

The perception of contrast

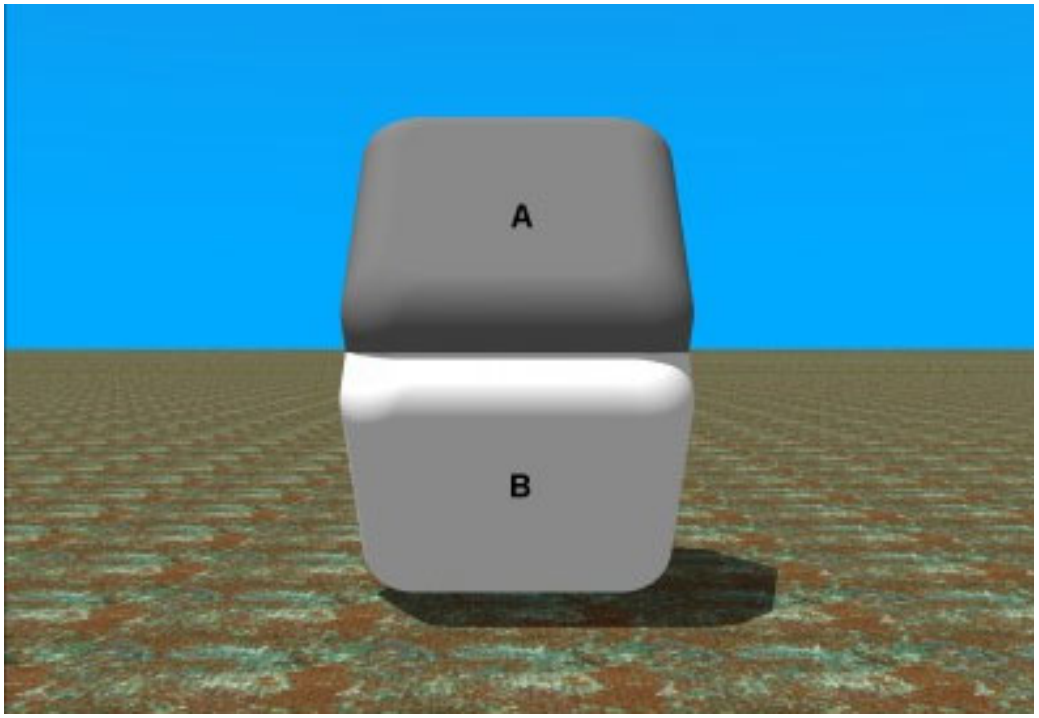




Computer Vision

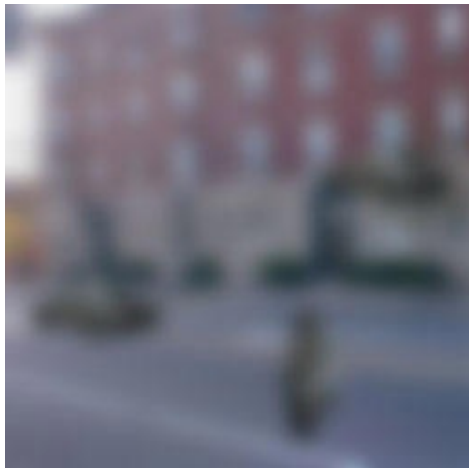


The perception of intensity



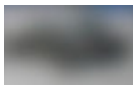
The role of context

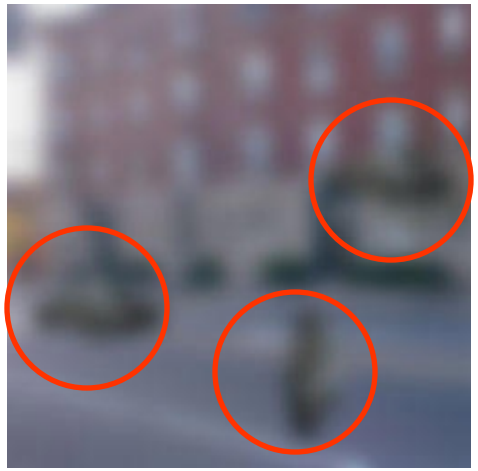
Human vision:
Biederman, Bar & Ullman, Palmer, …



The role of context

All encircled patterns are identical:





The role of context

Person?

The role of context

Person?

The role of context

Person?

The role of context

Person?

Computer Vision

The role of context

human vision is much more than a bottom-up process of subsequent signal processing steps.

Car?

The role of context

A bird, or a giant rabbit flying on a surf board ?

Want to know more?

Foundations and Trends in

Computer Graphics and Vision

Publishes tutorials / reviews

Freely available for ETH students

http://www.nowpublishers.com