# Introduction to Computer Vision

Computer Vision

Taught by

- prof. Luc Van Gool
- Prof. Ender Konukoglu

- Guest starring by prof Orçun Göksel

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*

… it is crucial …

# 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

The central take-home message:

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

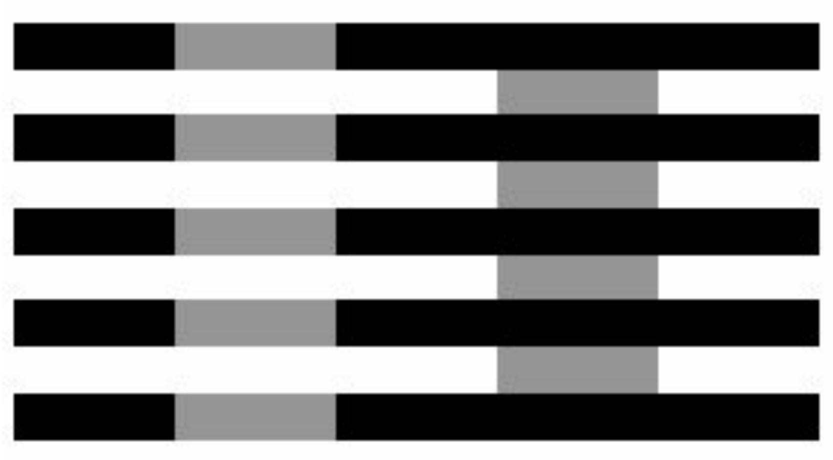
# … it is intriguing …

Computer Vision

INTRO

**perception**
applications
light

# The perception of intensity

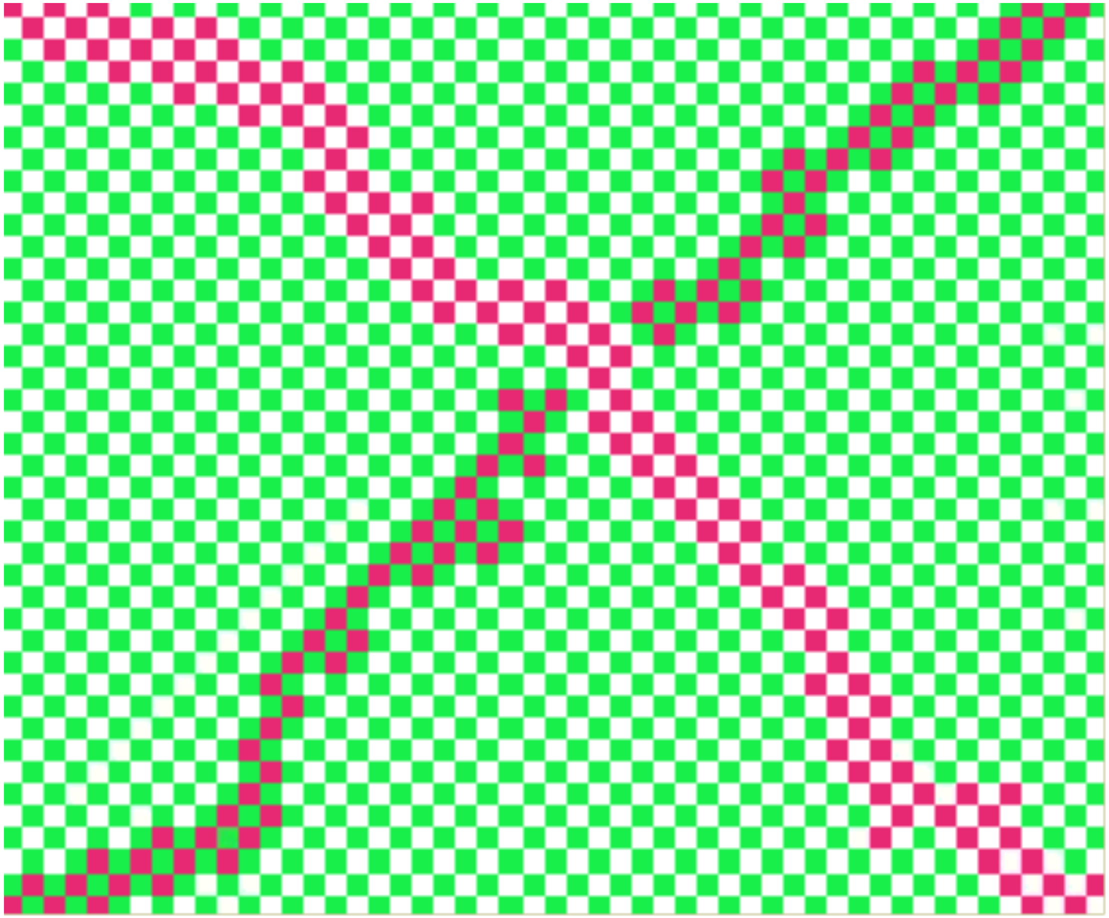


Computer Vision

INTRO

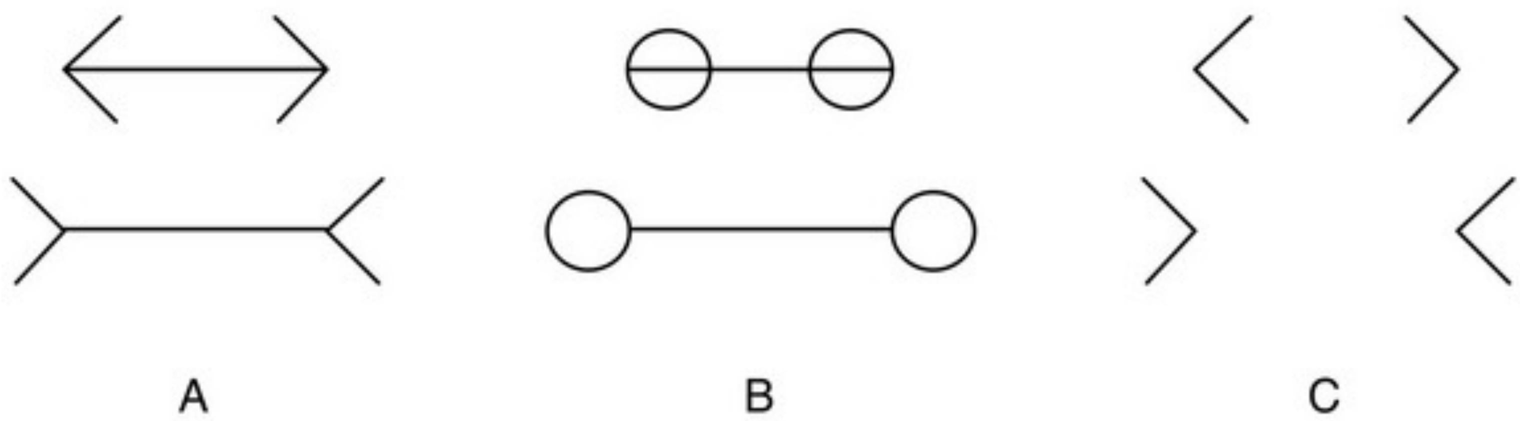
**perception**
applications
light

# The perception of color

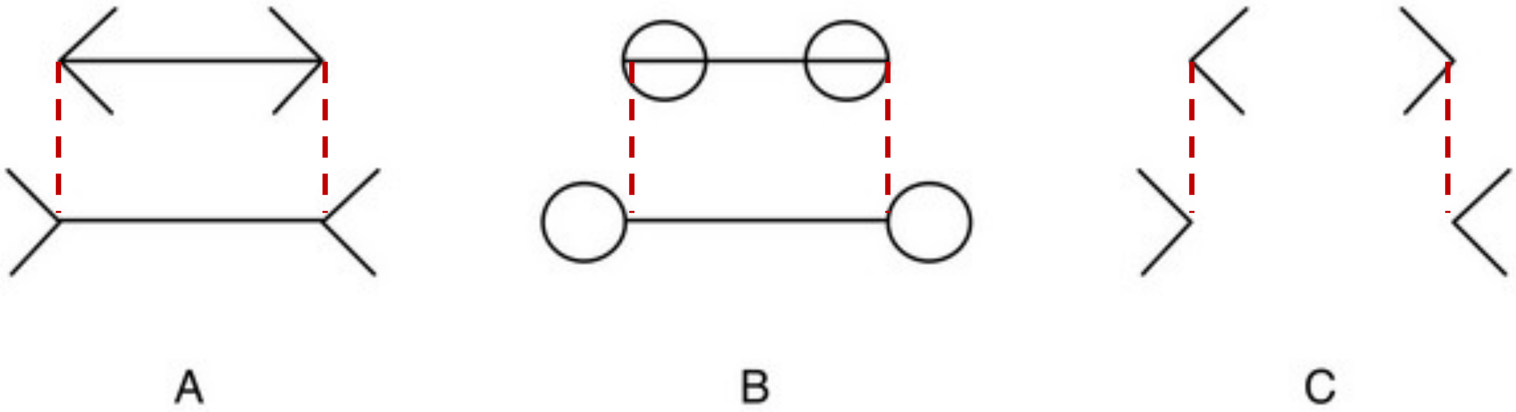


The red squares have equal color…

# The perception of length



# The perception of length



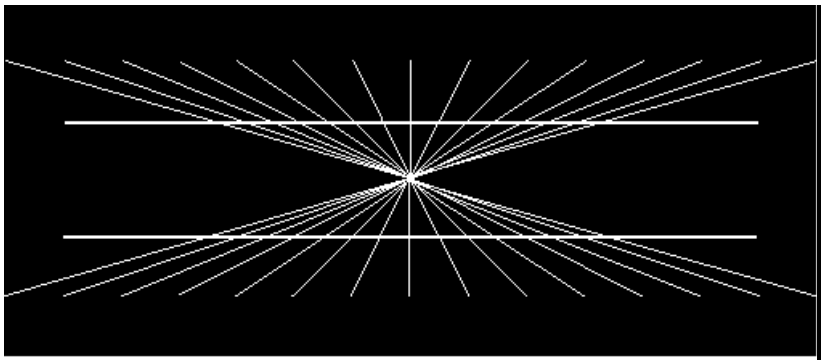

The horizontal lines are equally long…

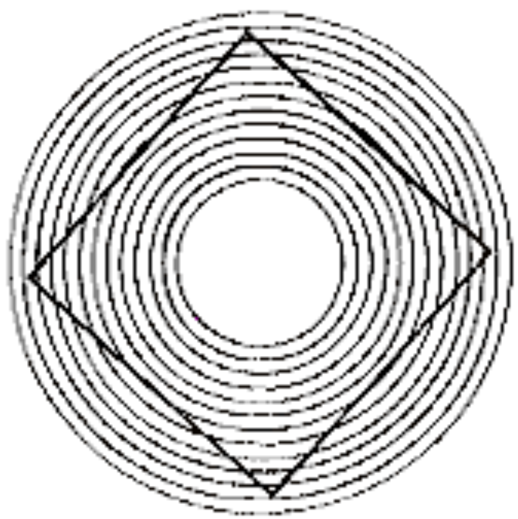
Computer Vision

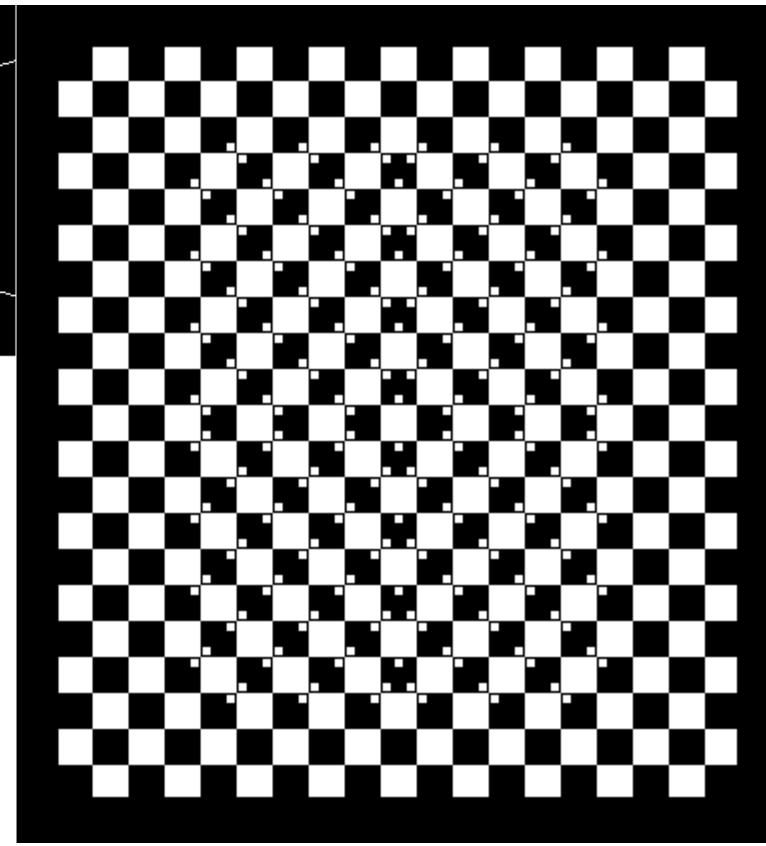
INTRO

**perception**
applications
light

# The perception of lines being straight





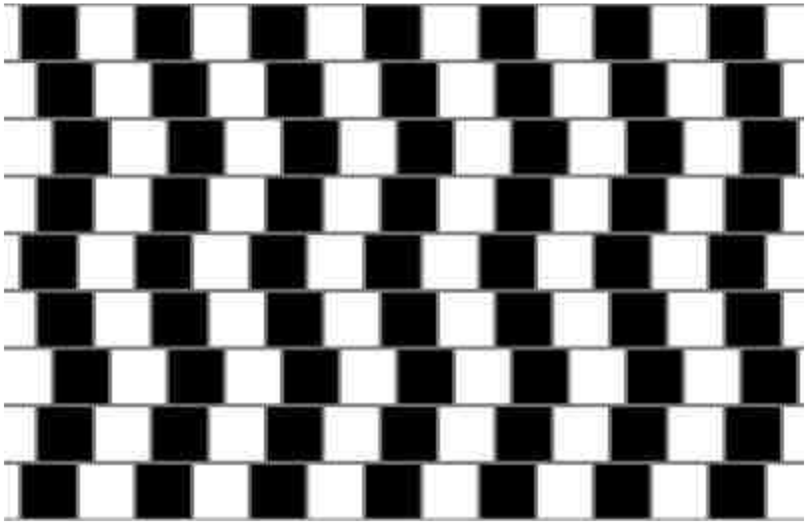


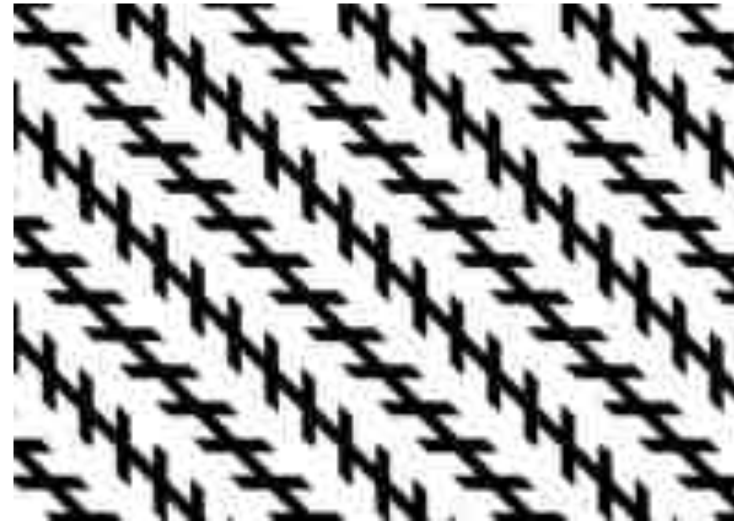
Computer Vision

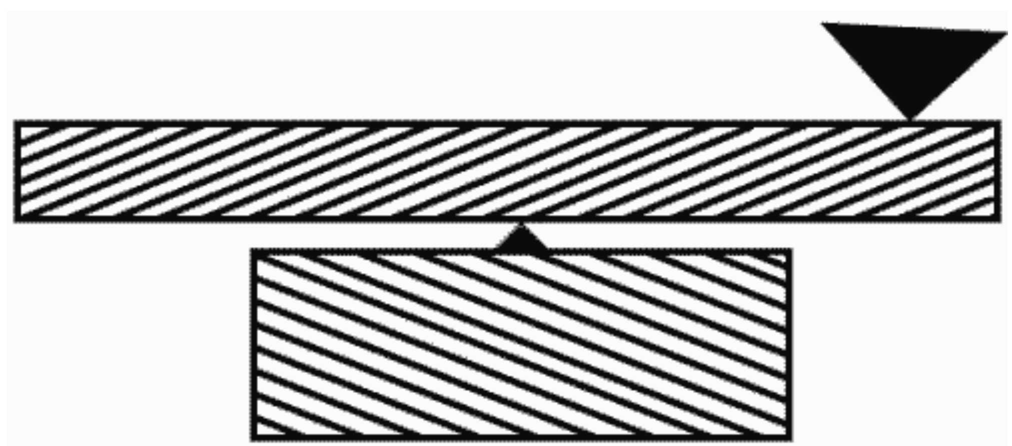
INTRO

**perception**
applications
light

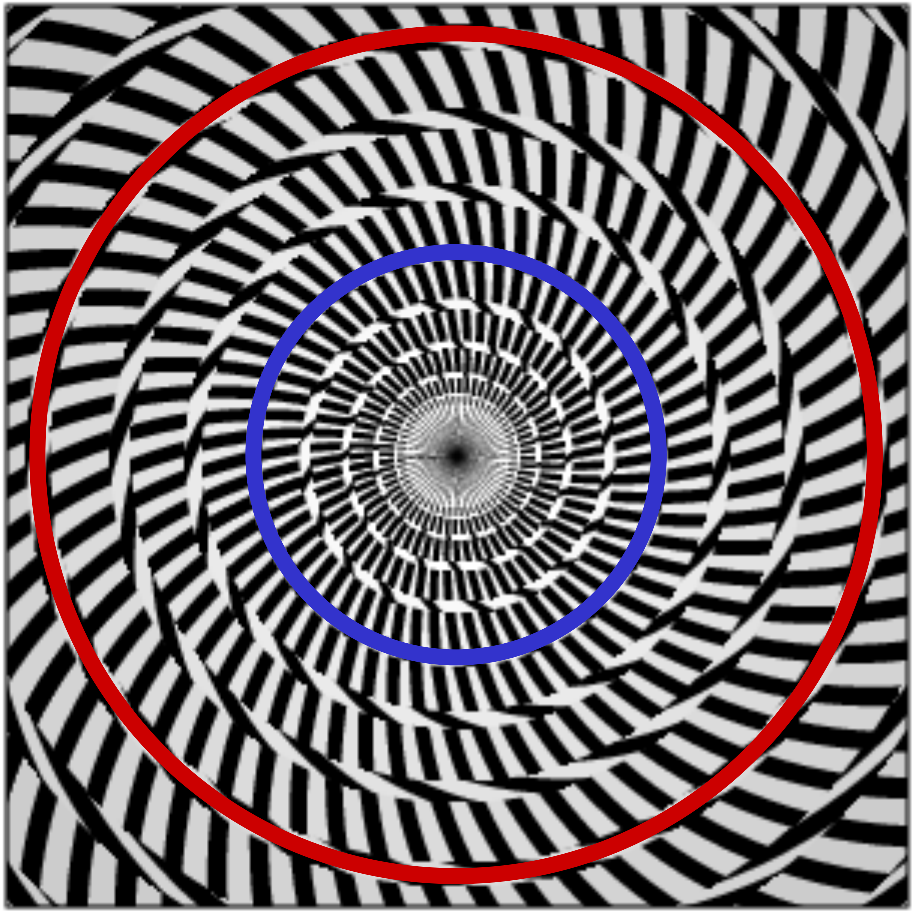
# The perception of parallelism







# The perception of curvatures



Illusions : interference of differently oriented patterns via adaptation

Computer Vision

INTRO

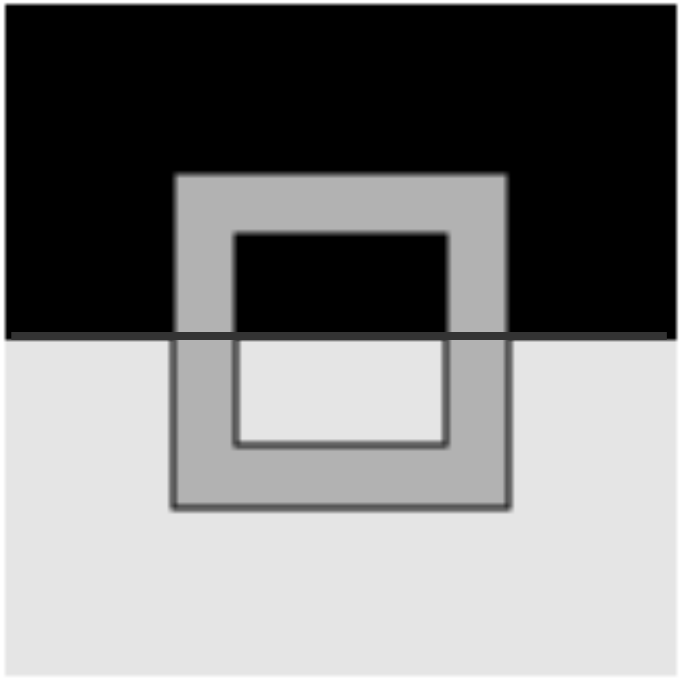
**perception**
applications
light

# The perception of motion

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

# It's not that more context solves it all…

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

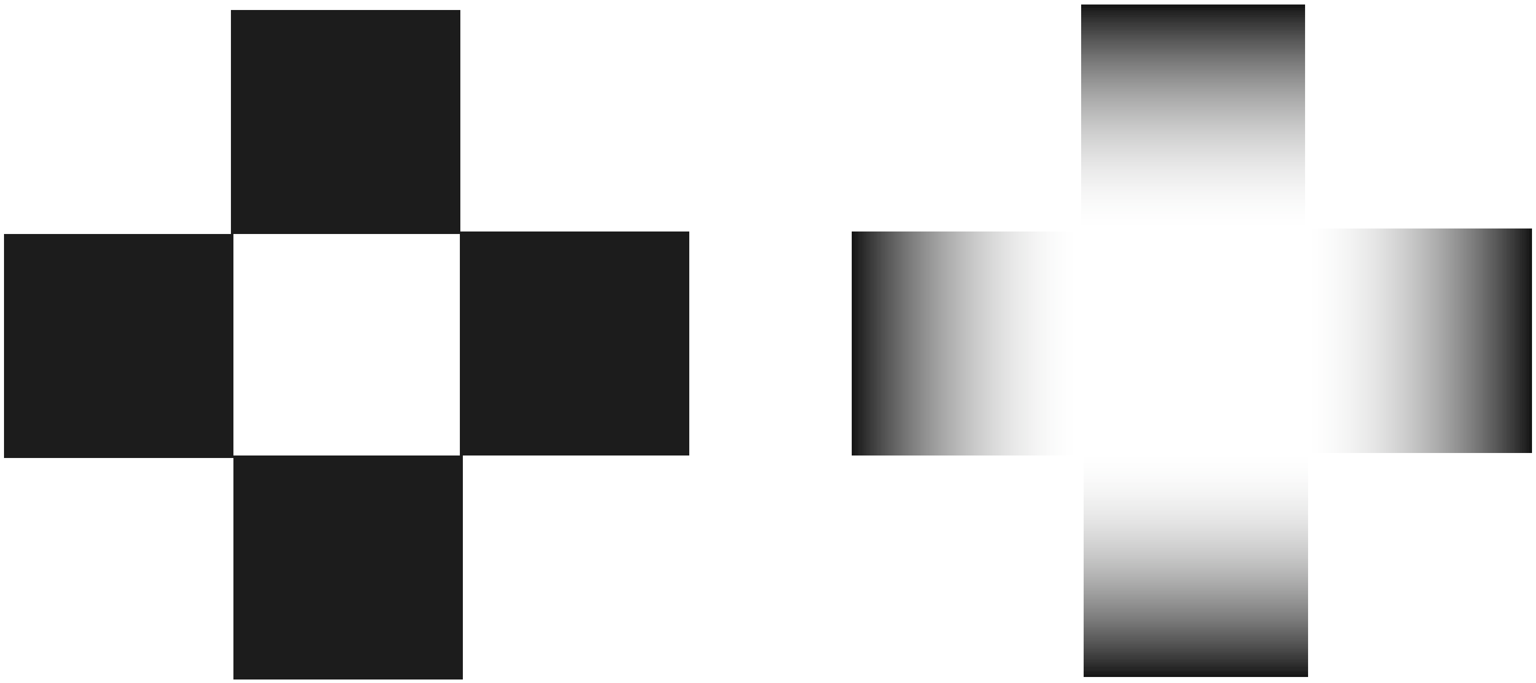


Computer Vision

INTRO

**perception**
applications
light

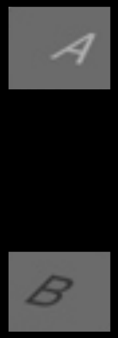
# The perception of intensity



Computer Vision

INTRO

**perception**
applications
light

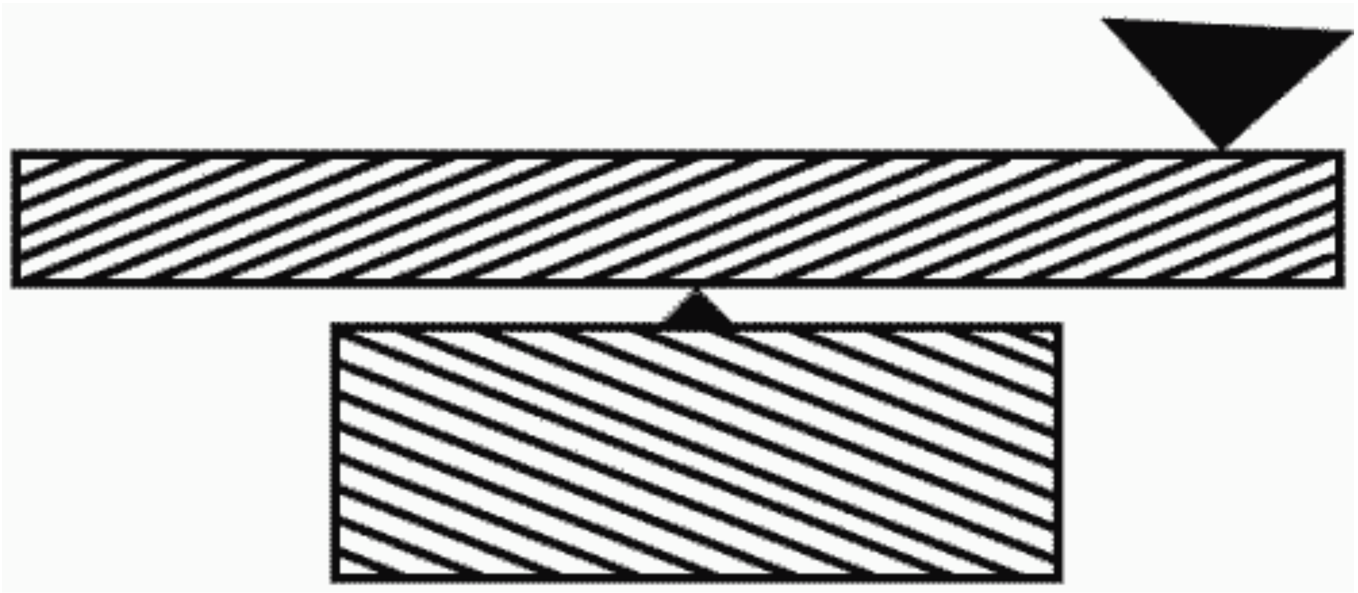


Computer Vision

INTRO

**perception**
applications
light

# Parallelism again…



Computer Vision

INTRO

**perception**
applications
light

# Kanisza illusion


Fill-in : averaging of perceived contrast at edges over regions possibly obtained via extrapolation of the edges… in any case *such illusion seems to help people to detect patterns in the world.*

Computer Vision

INTRO

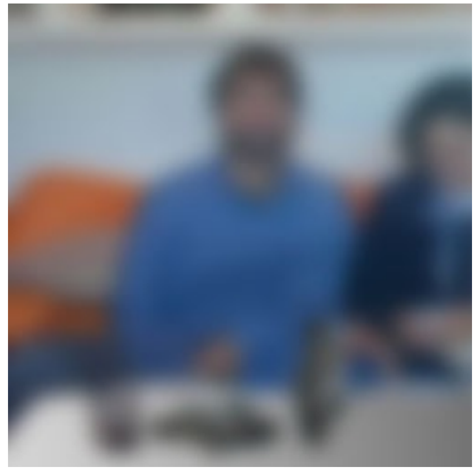
**perception**
applications
light

Computer Vision

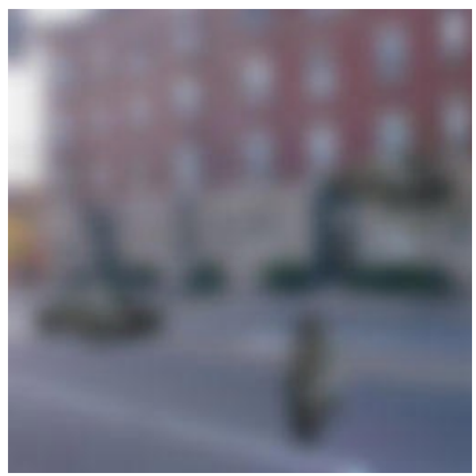
INTRO

**perception**
applications
light

# The role of context




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



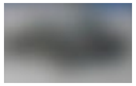

Computer Vision

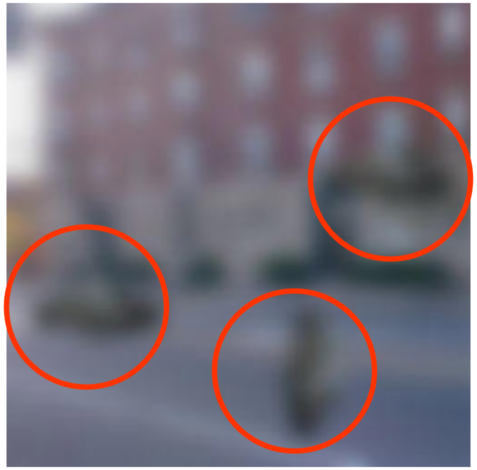
INTRO

**perception**
applications
light

# The role of context

All encircled patterns are identical:





Computer Vision

# The role of context

INTRO

**perception**
applications
light

Person?

Computer Vision

INTRO

**perception**
applications
light

# The role of context

Person?

# The role of context

Person?

Computer Vision

INTRO

**perception**
applications
light

# The role of context

Person?

Computer Vision

INTRO

**perception**
applications
light

Car?

# The role of context

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

The central take-home message:

**Effective vision needs more than sheer filtering and measuring**

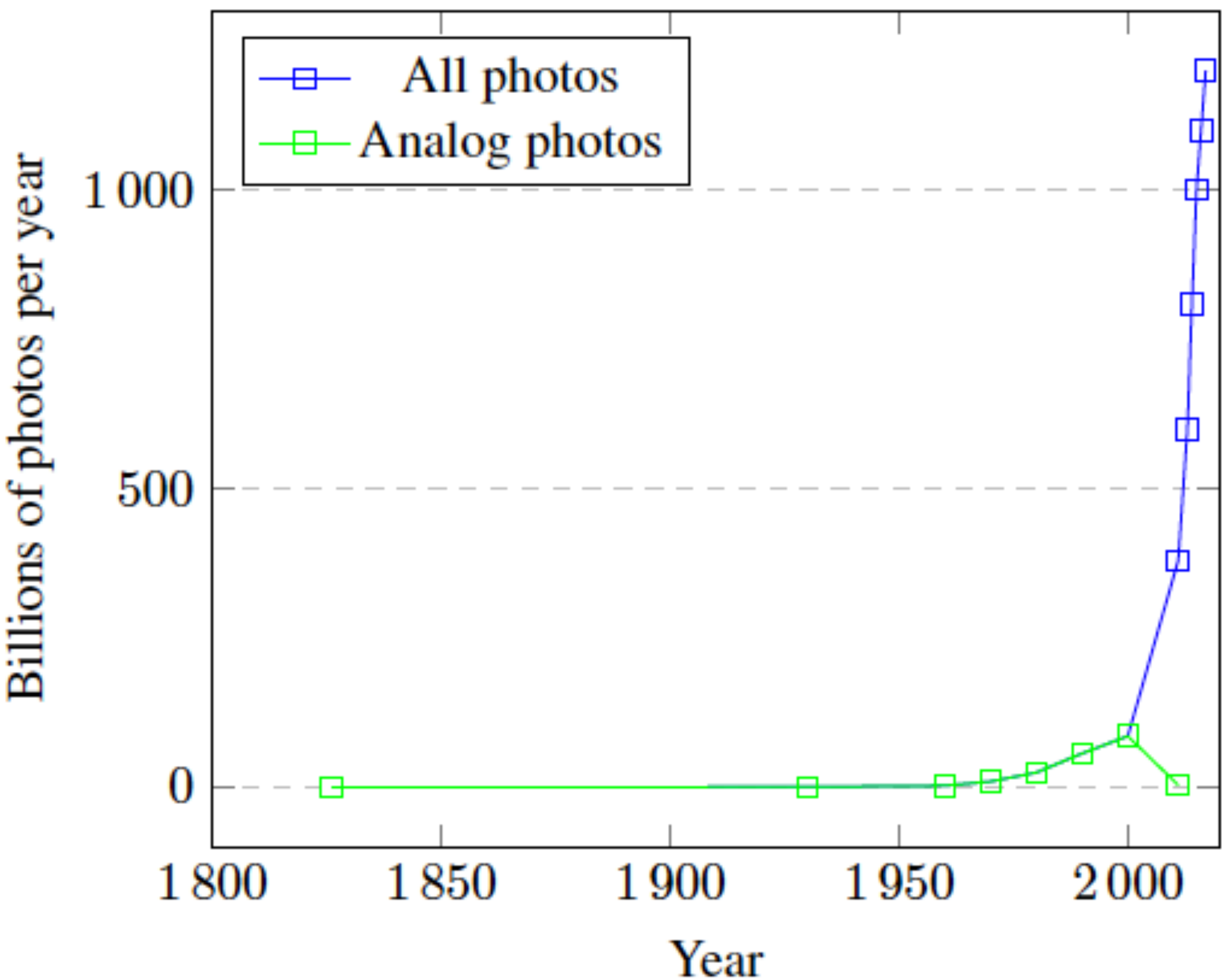
… it is hot …

Computer Vision

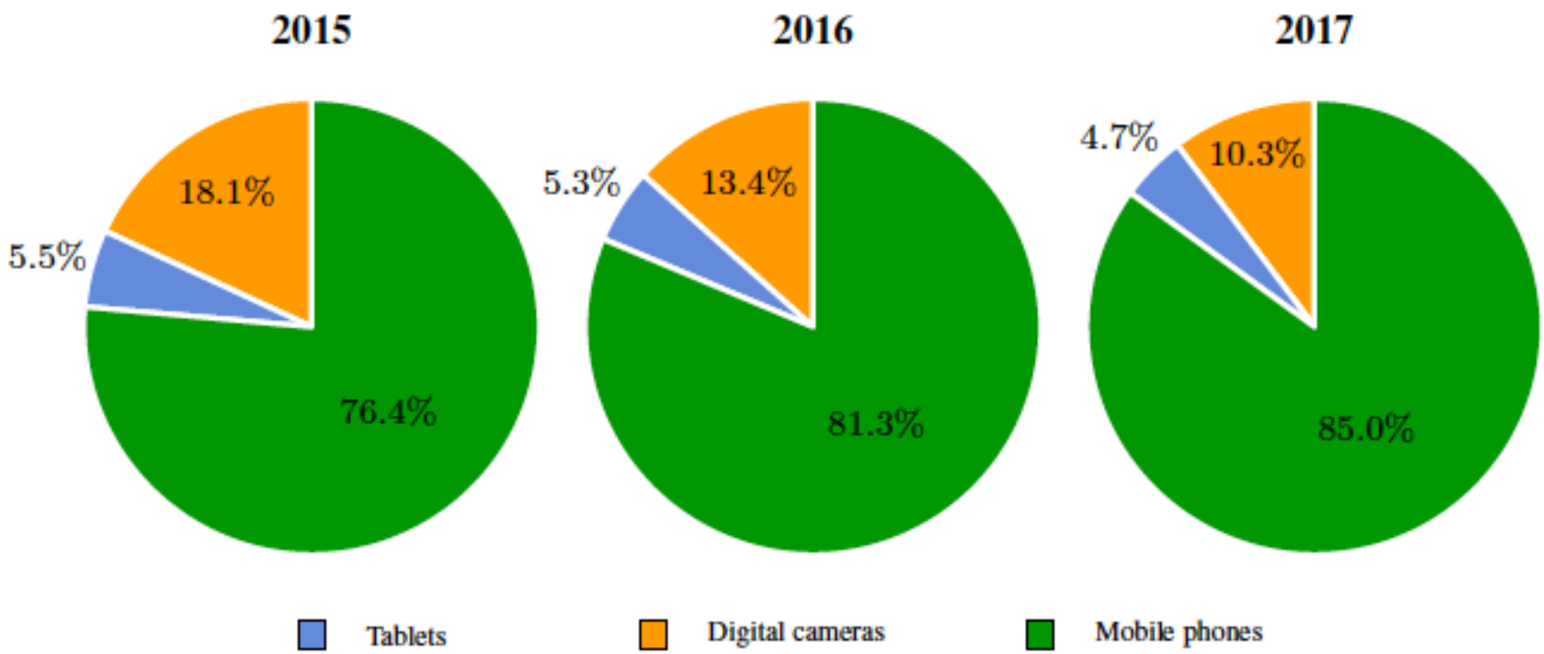
INTRO

perception
**applications**
light

# The explosion of photography



# The explosion of photography




Easier than ever to take a photo
The cost is extremely low (cheap memory)
Most people carry a camera most of the time

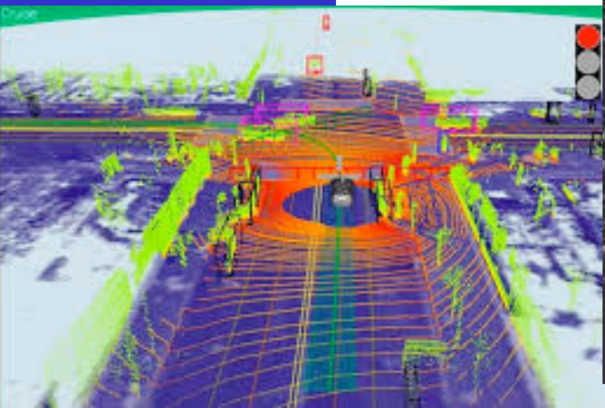
# The development of computer vision apps

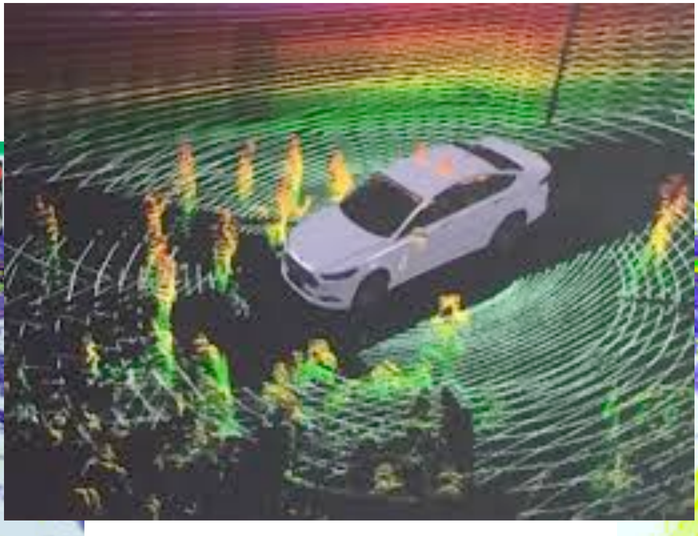
Most early applications where found in production environments, as these *allow for controlled conditions* and *have little uncertainty*

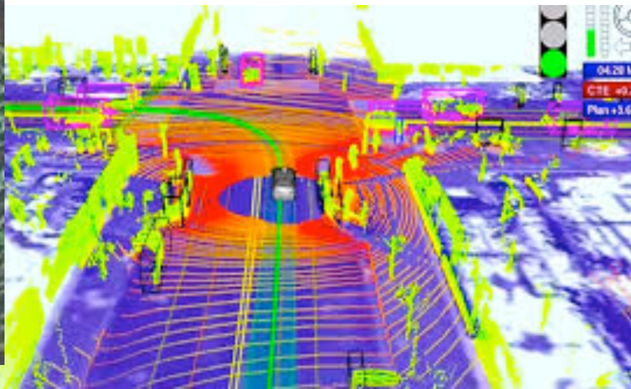
some areas do not allow for much control: medical IP, remote sensing, surveillance, etc.

currently CV is conquering the less controllable areas by storm

# Ex App: autonomous vehicles





Computer Vision

INTRO

perception
**applications**
light

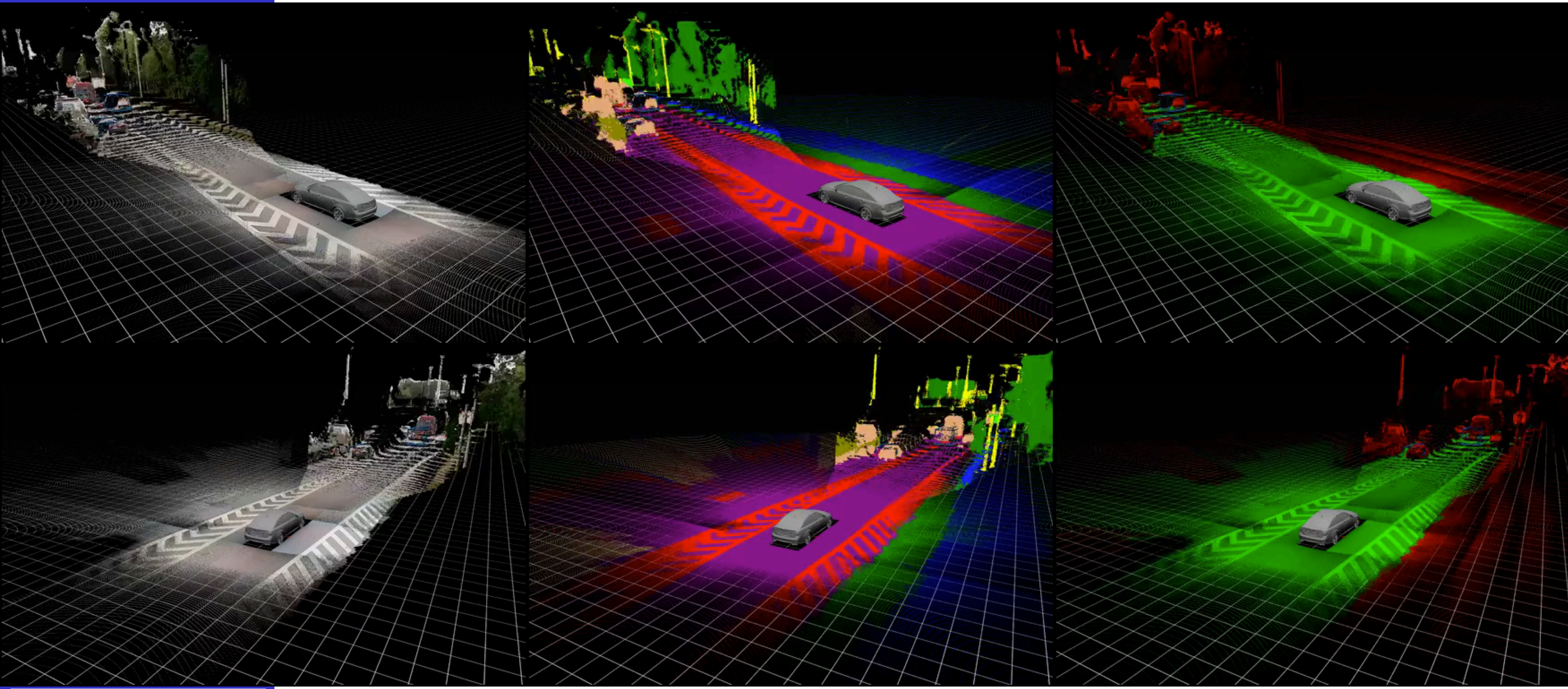
# Ex App: autonomous vehicles

car detection:

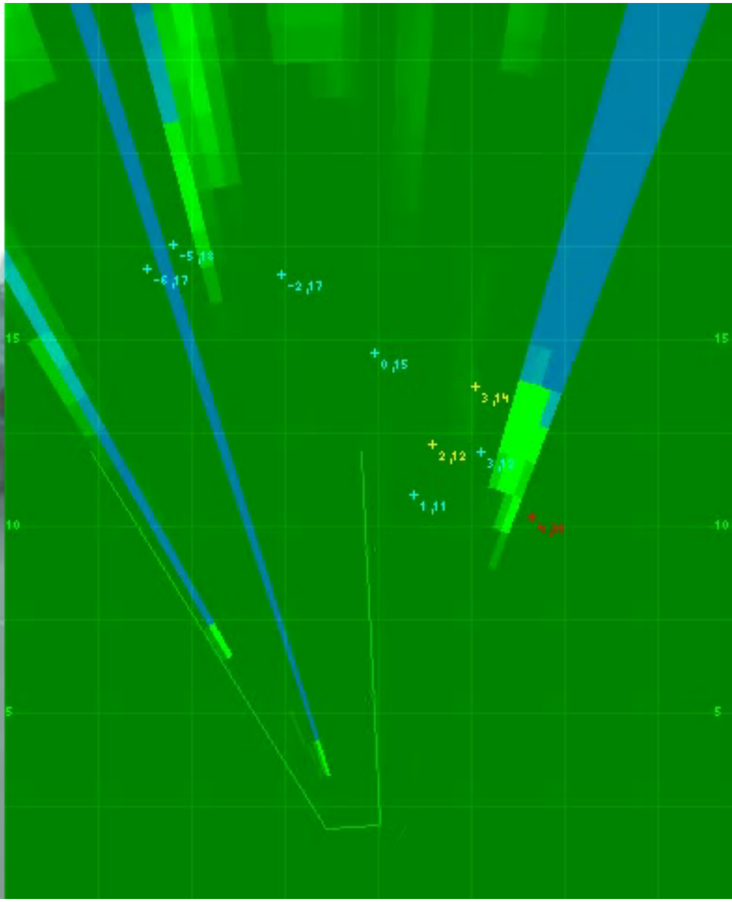
Computer Vision

# Ex App: autonomous vehicles

putting vision modalities together:



# Ex: autonomous mobile platform



Computer Vision

# Ex App: image retrieval, captioning, …

## Describes without errors


A person riding a motorcycle on a dirt road.


A group of young people playing a game of frisbee.


A herd of elephants walking across a dry grass field.

## Describes with minor errors


Two dogs play in the grass.


Two hockey players are fighting over the puck.


A close up of a cat laying on a couch.

## Somewhat related to the image


A skateboarder does a trick on a ramp.


A little girl in a pink hat is blowing bubbles.


A red motorcycle parked on the side of the road.

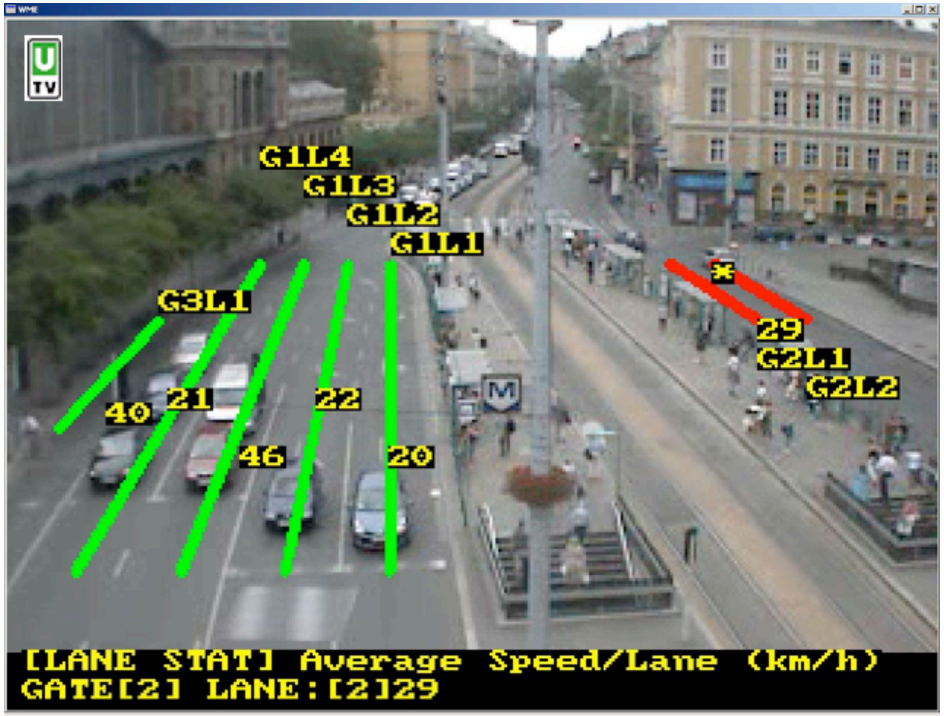
## Unrelated to the image


A dog is jumping to catch a frisbee.


A refrigerator filled with lots of food and drinks.

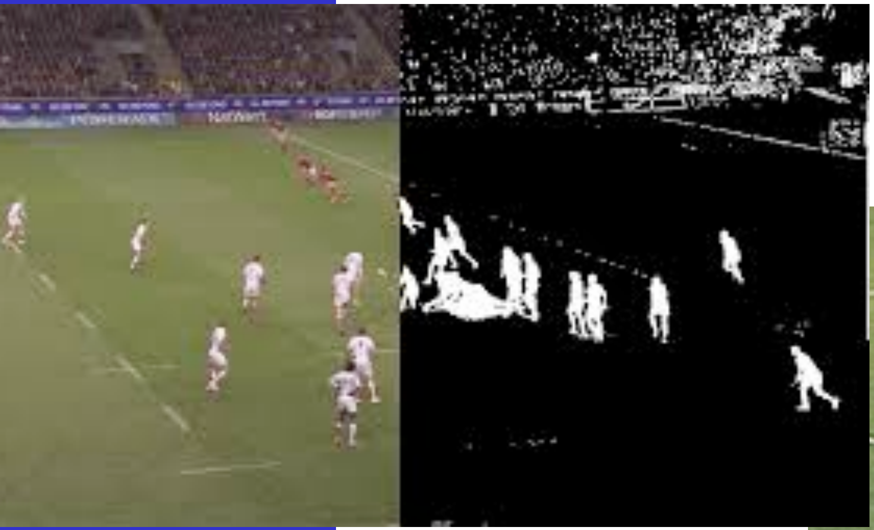

A yellow school bus parked in a parking lot.

# Ex App: visual surveillance



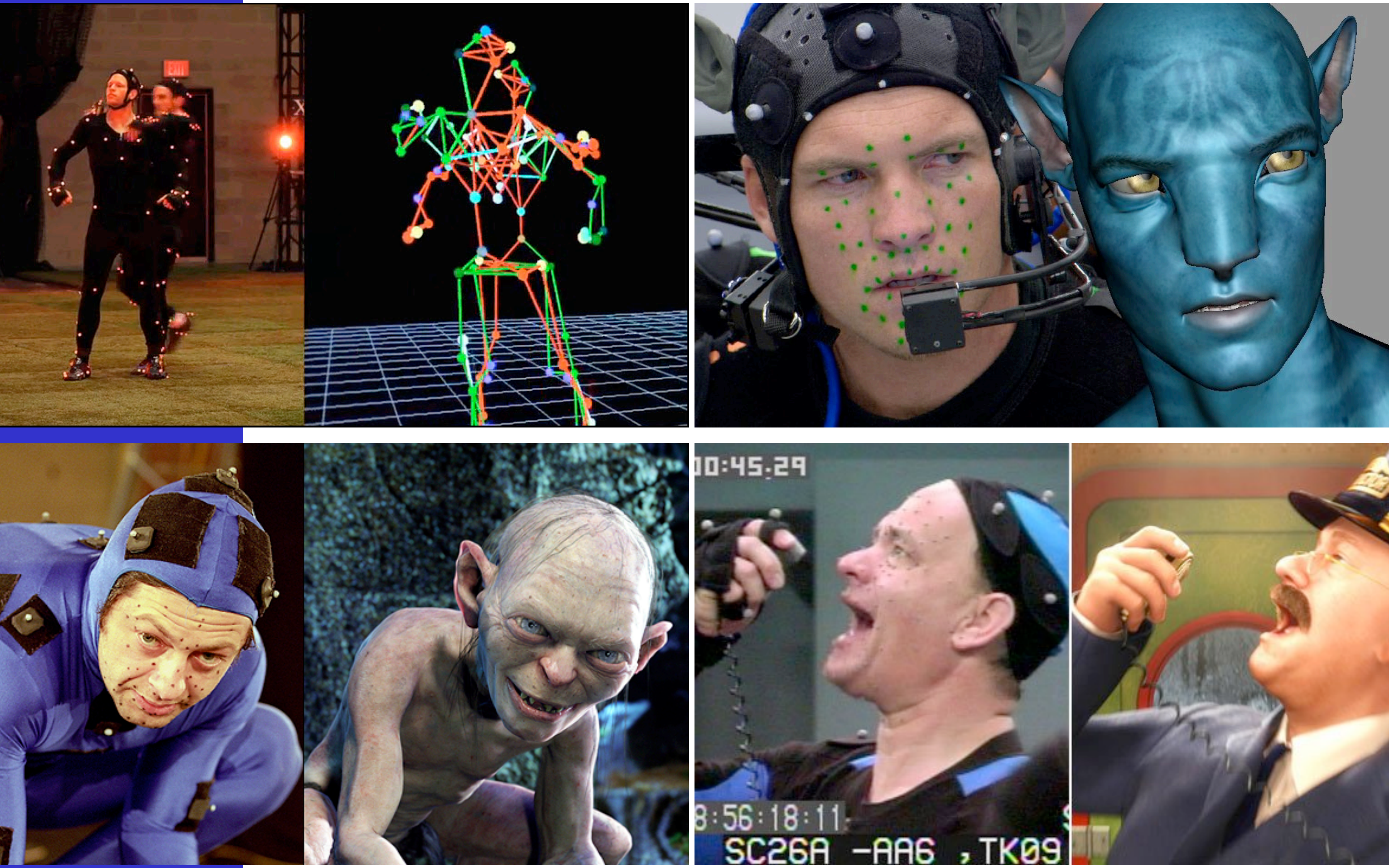
Computer Vision

# Ex App: Augm. Reality, eg sports

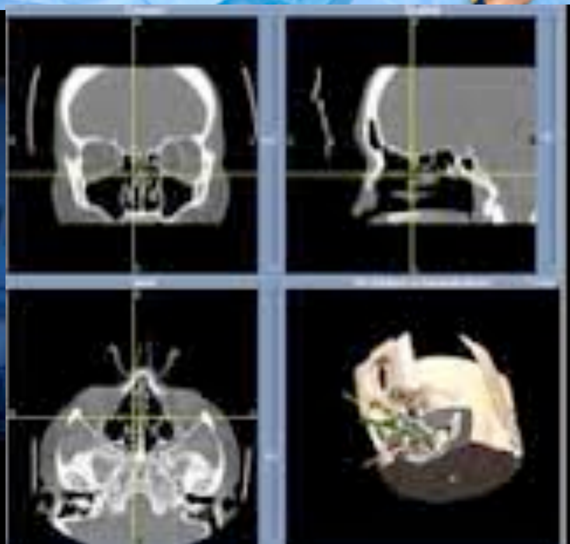


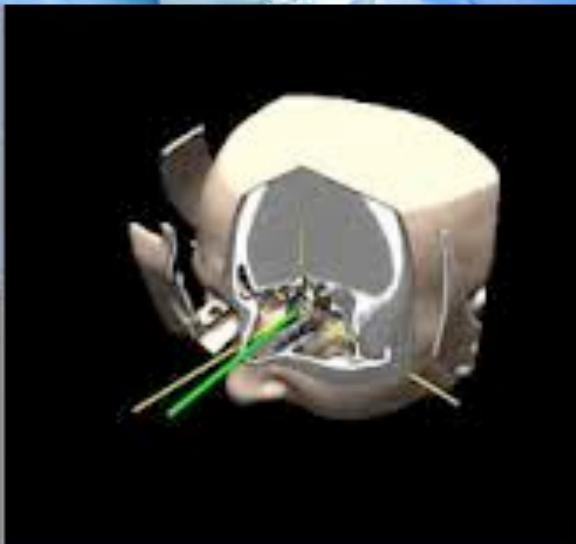


# Ex App: motion capture for movies/games



# Ex App: computer-assisted surgery





# Mobile mapping

The central take-home message:

**It is feasible now to let most things see and interprete their environment**

# … it needs light …

Computer Vision

INTRO

perception
applications
**light**

# And then there was Light…

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

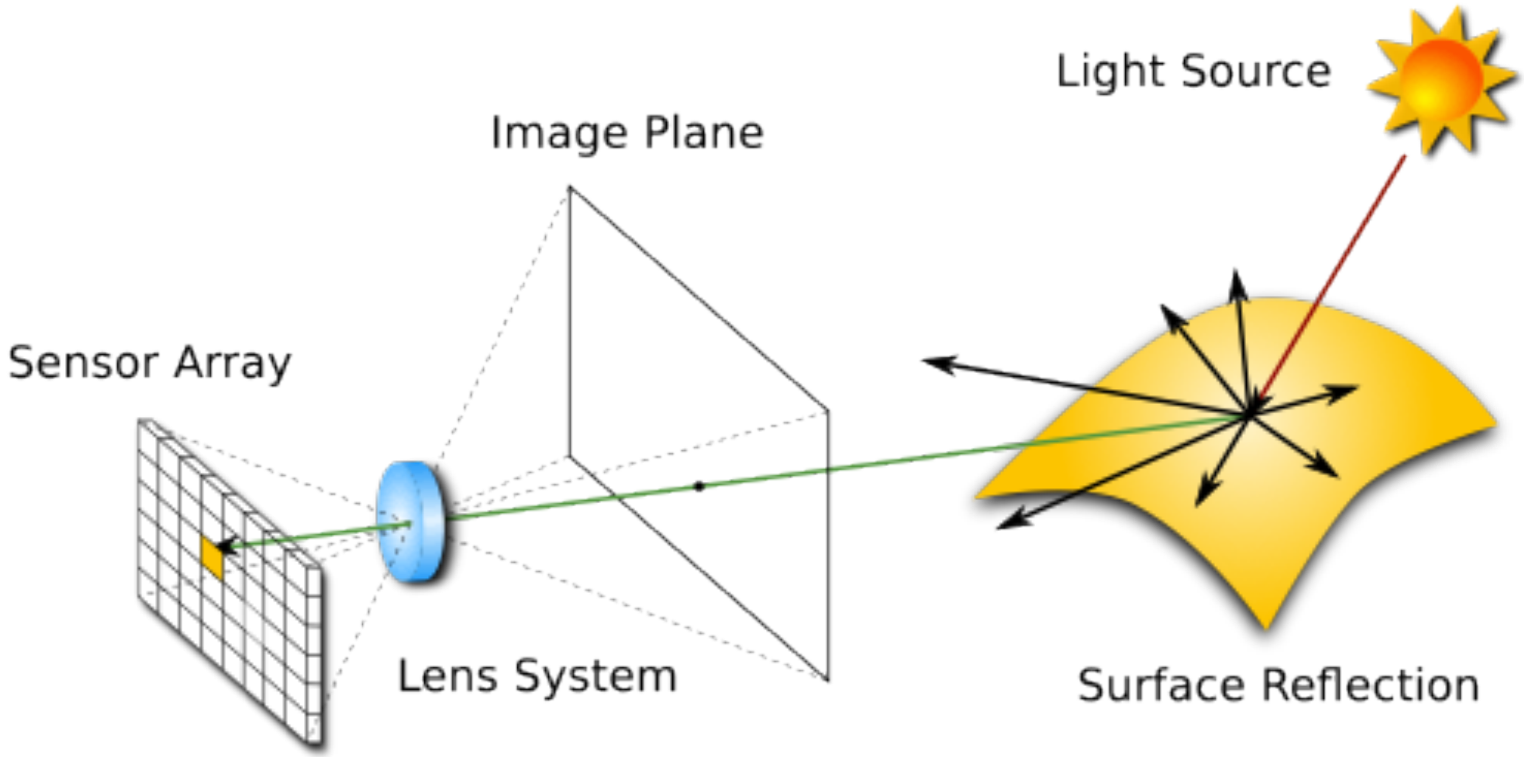

"What the...?"

Computer Vision

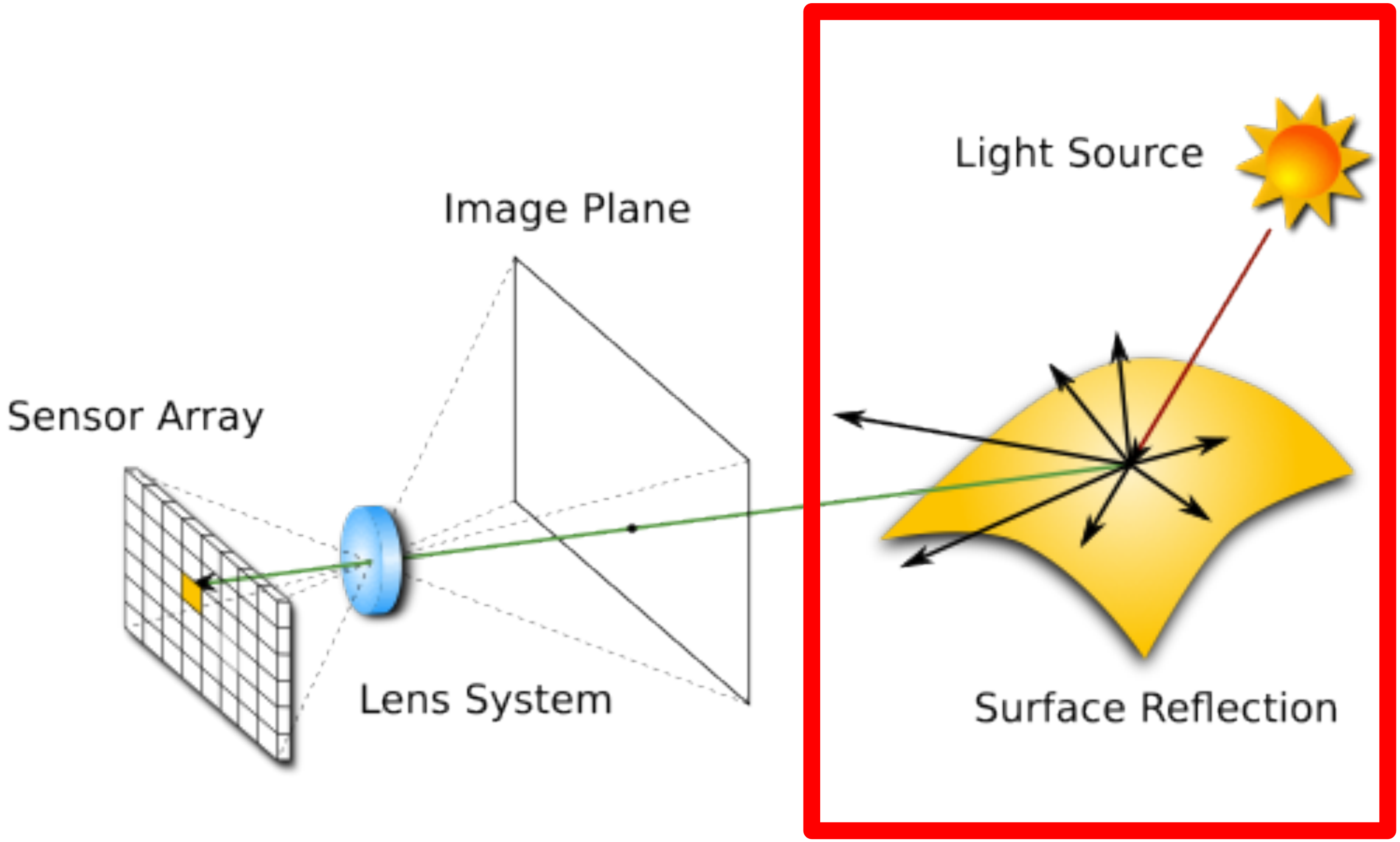
INTRO

perception
applications
**light**

# 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

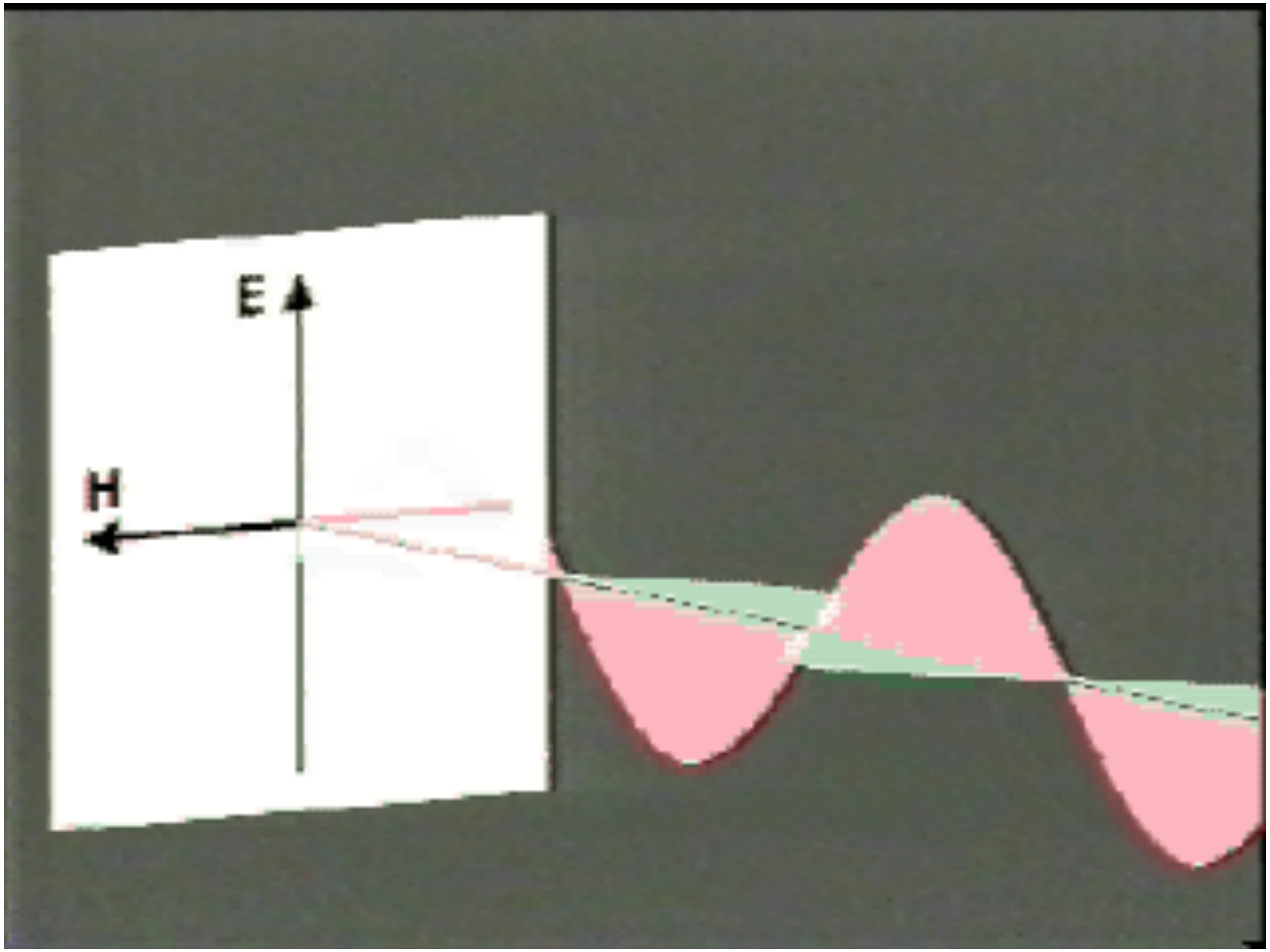
➔ **wave character**

Computer Vision

INTRO

perception
applications
**light**

# Light as electromagnetic waves



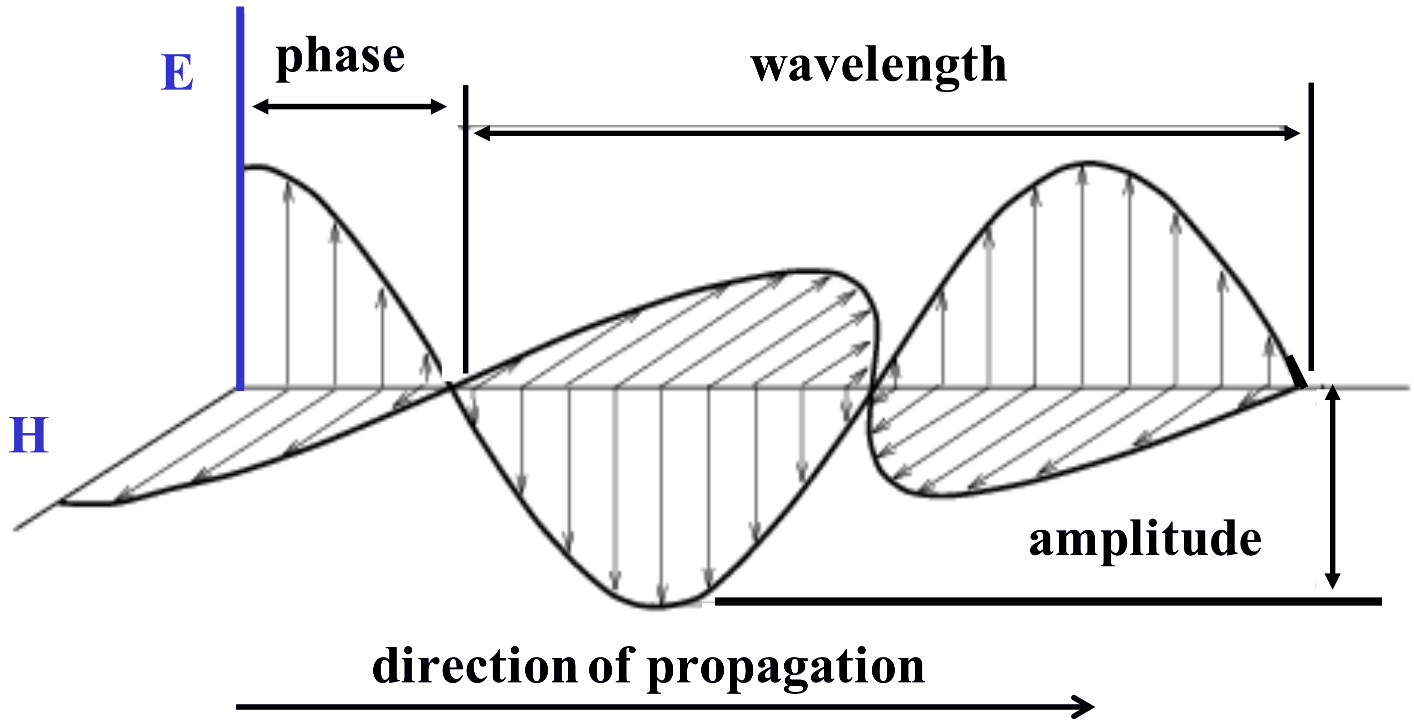
Computer Vision

INTRO

perception
applications
**light**

# 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

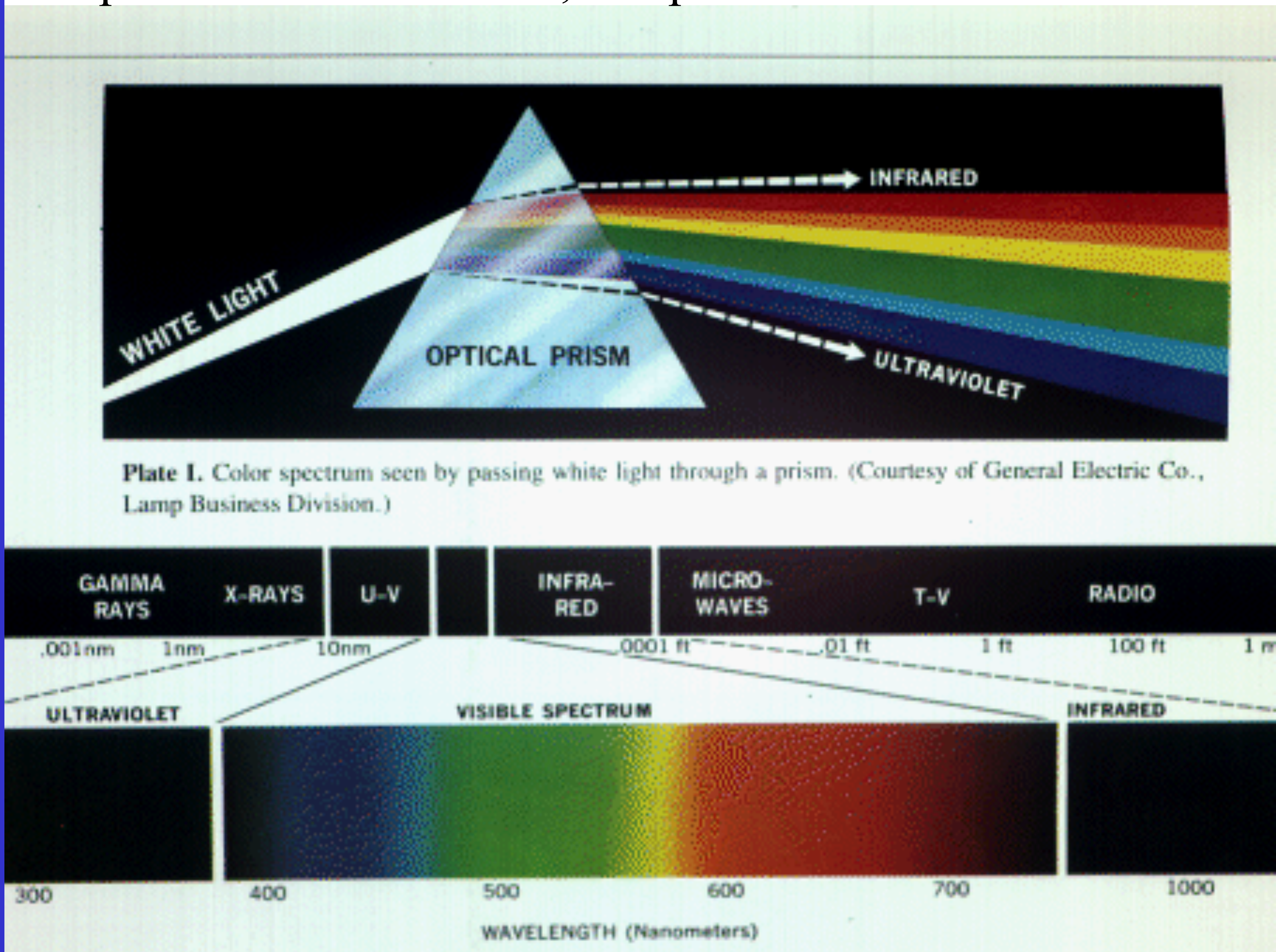



Plate 1. Color spectrum seen by passing white light through a prism. (Courtesy of General Electric Co., Lamp Business Division.)

Computer Vision

INTRO

perception
applications
**light**

# 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 |

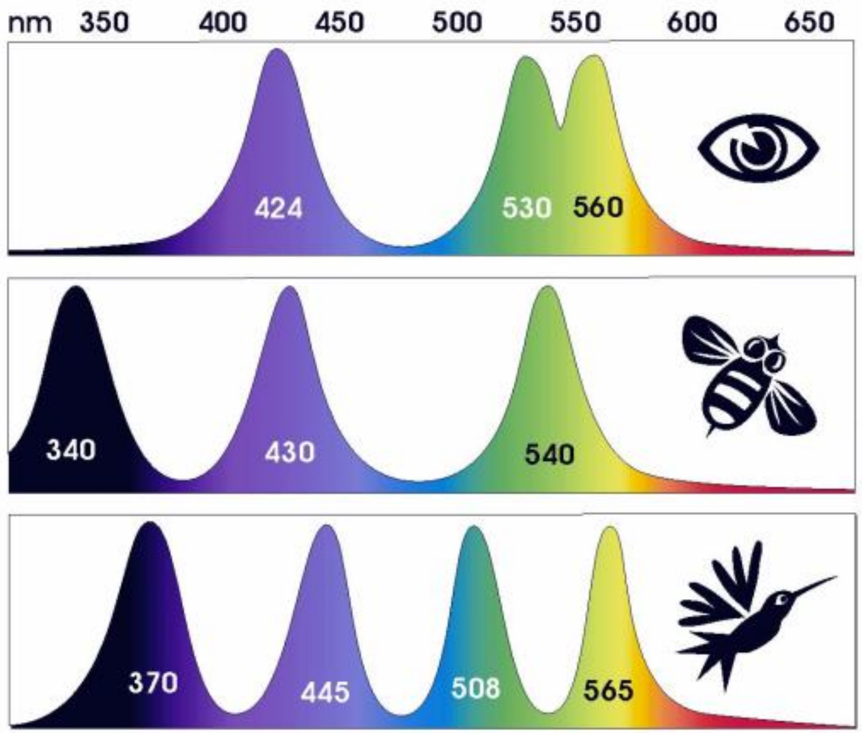
**NOTE** : Cameras may have different spectral sensitivities (i.e. also different from human vision)

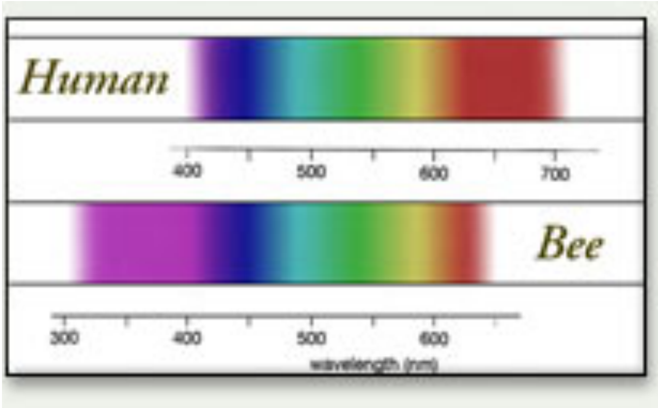
Computer Vision

INTRO

perception
applications
**light**

# The visible range



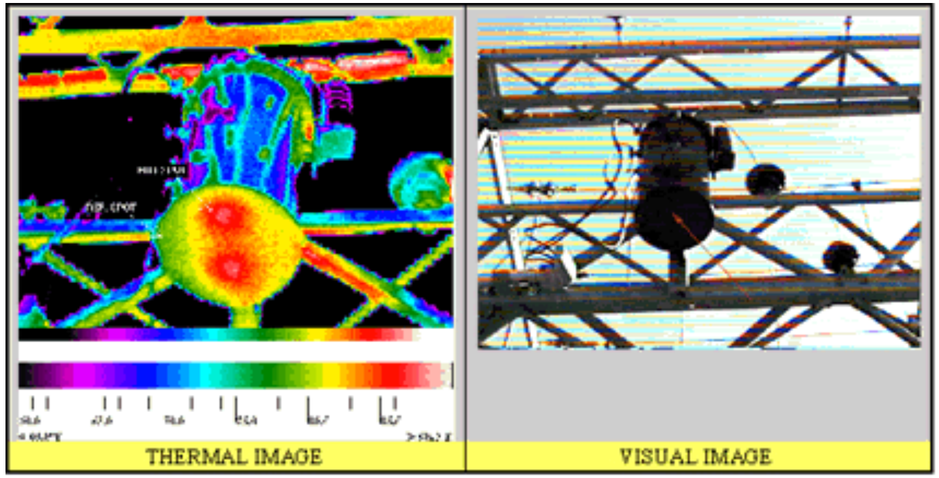

**NOTE** : animals may have different spectral sensitivities (i.e. different from human vision), and may also have a Different number of cone types, like 4 in most birds.

Computer Vision

INTRO

perception
applications
**light**

# 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:

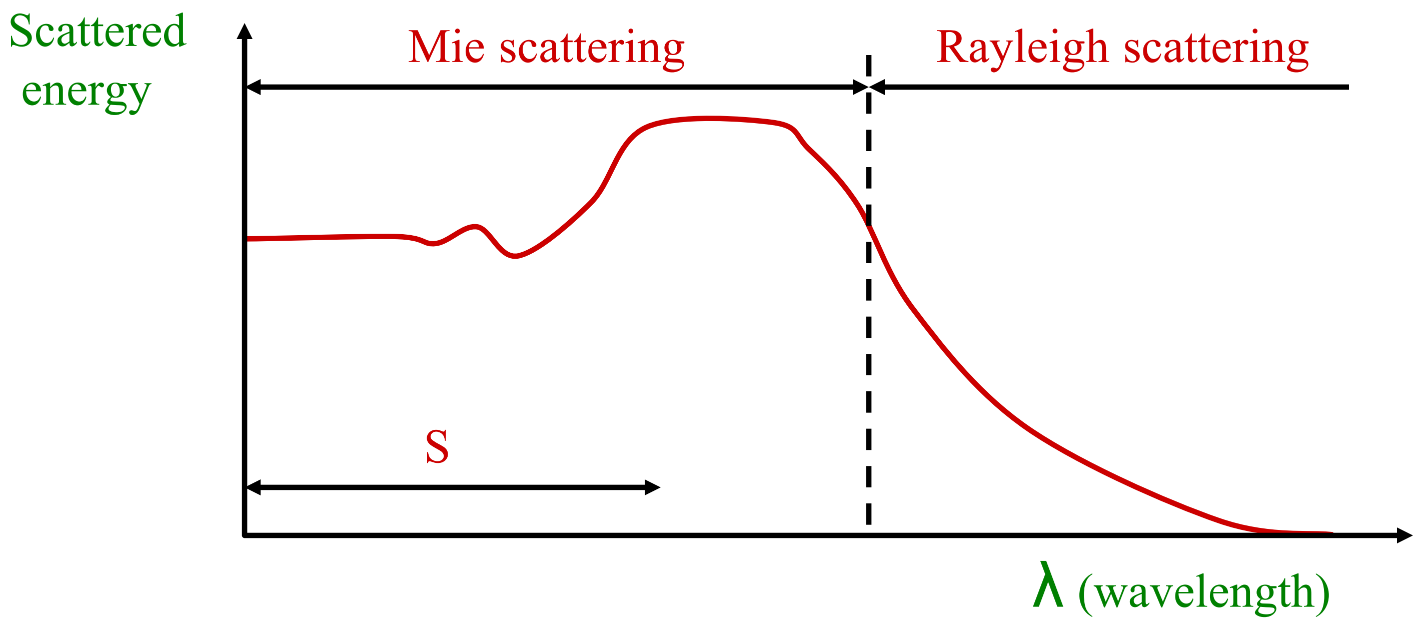
1. small particles: *Rayleigh* (strongly wavelength dependent)
2. comparable sizes: *Mie* (weakly wavelength dependent)
3. Large particles: *non-selective* (wavelength independent)

Computer Vision

INTRO

perception
applications
**light**

# 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

Computer Vision

INTRO

perception
applications
**light**

# Atmospheric showcase

*Rayleigh:*
Tyndall effect (blue sky)
Red, setting sun

*Non-selective:*
Grey clouds

*Mie:*
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 |

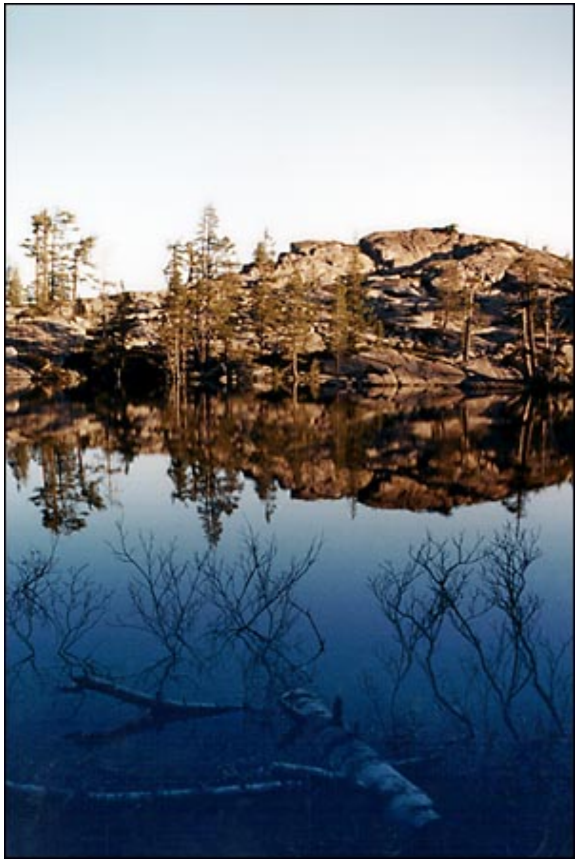
+ diffraction

Computer Vision

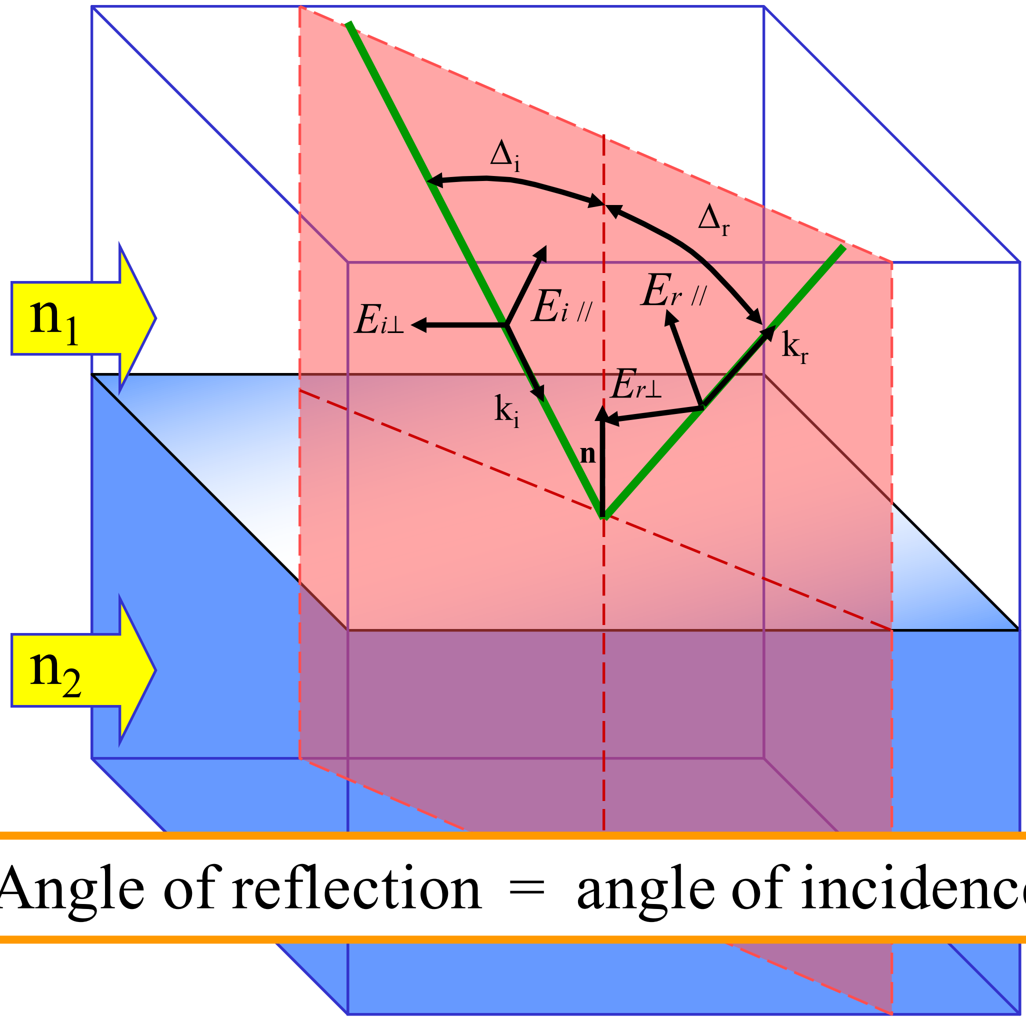
INTRO

perception
applications
**light**

# Mirror reflection



# Mirror reflection



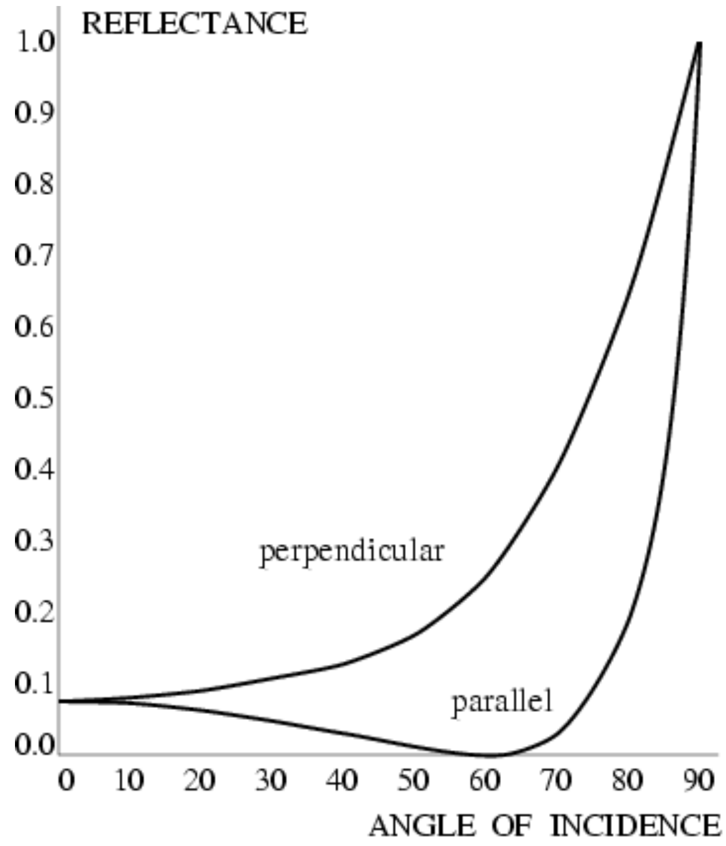

Angle of reflection = angle of incidence

Computer Vision

INTRO

perception
applications
**light**

# Mirror reflection : dielectric




Polarizer at *Brewster angle*

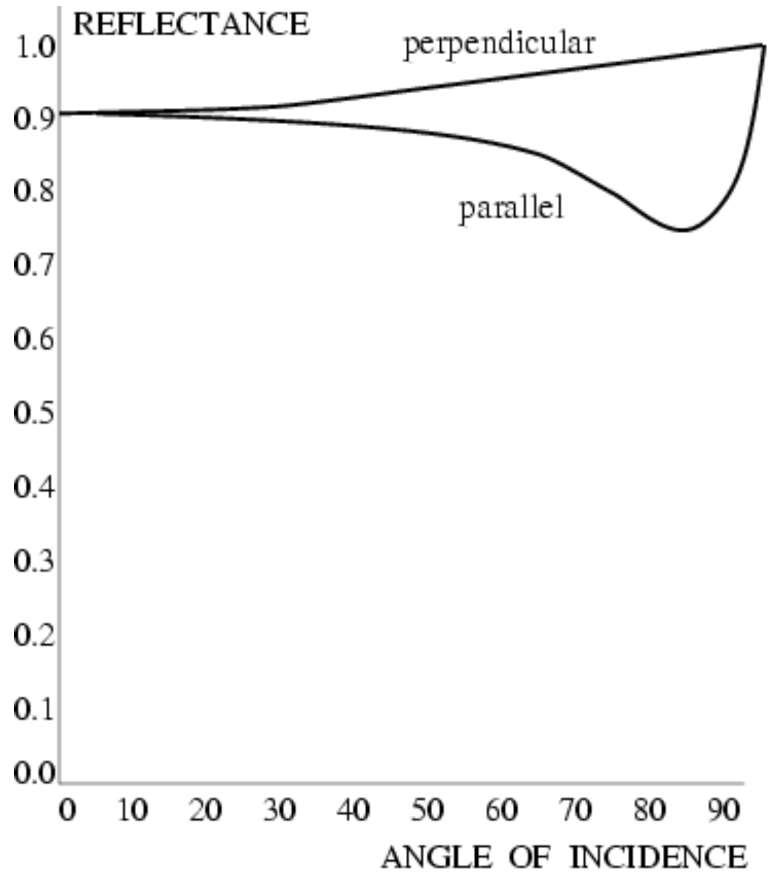
Full reflection at grazing angles

Computer Vision

INTRO

perception
applications
**light**

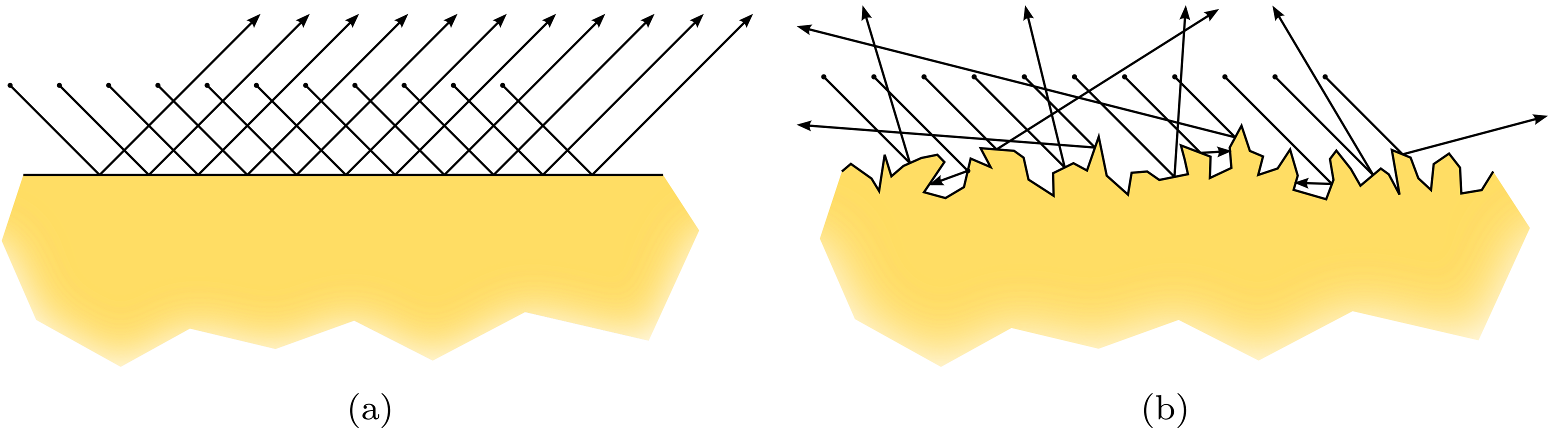
# Mirror reflection : conductor




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

# Roughness of surfaces leads to `diffuse' reflection

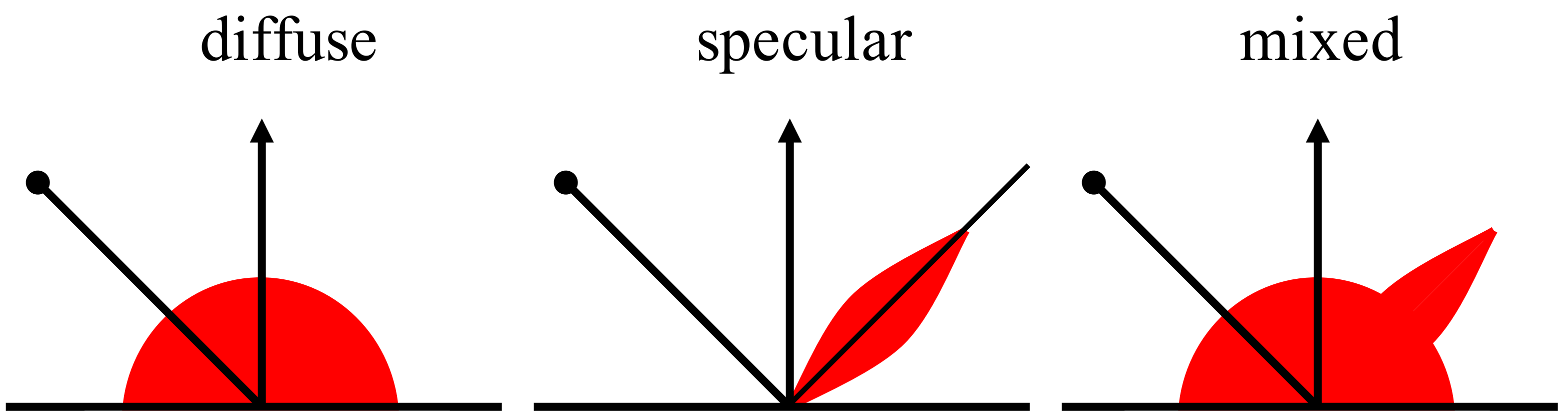





(a) Mirror or `specular' reflection, (b) diffuse reflection

# … and to mixed reflection for most real surfaces

three types of reflection :



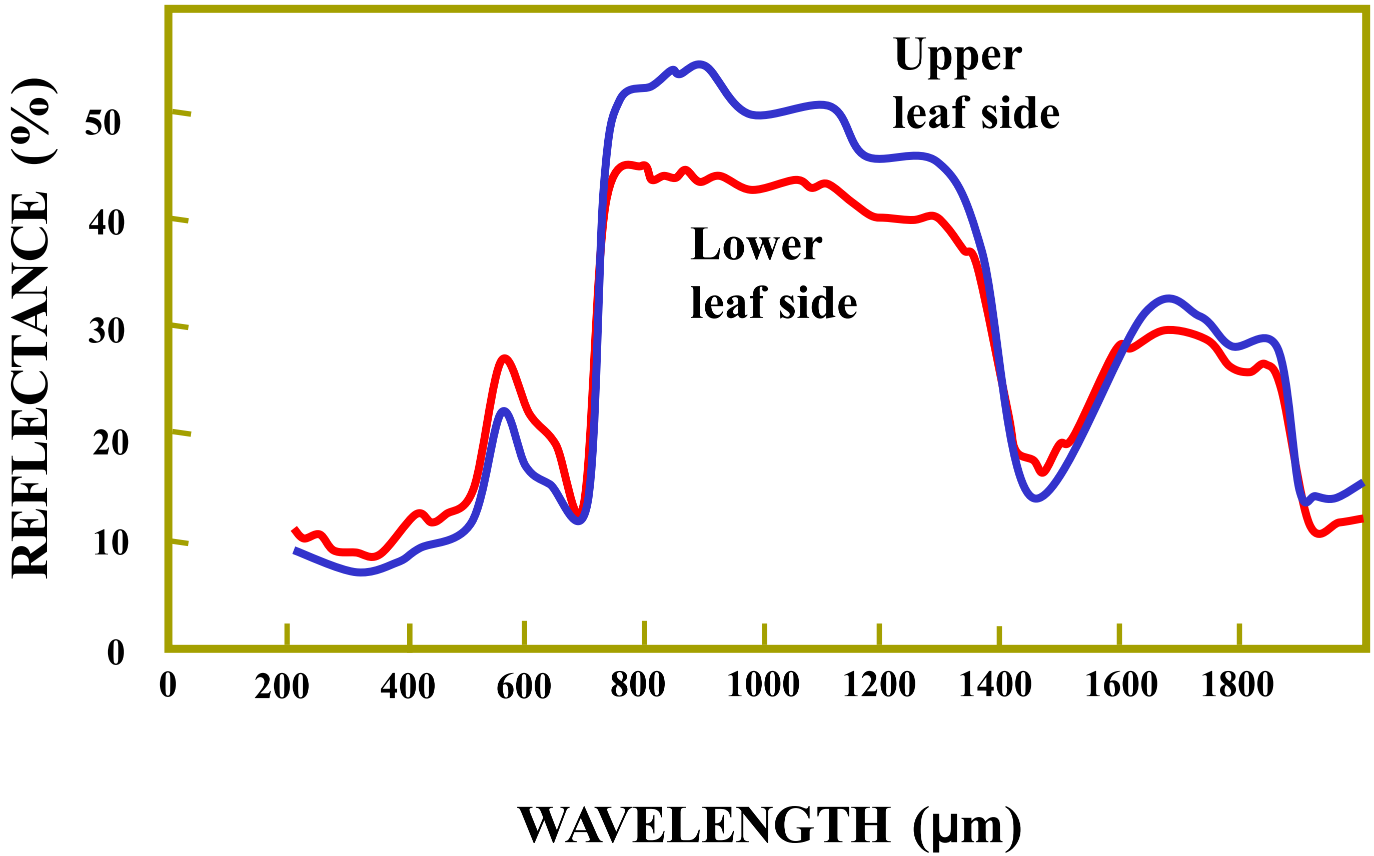

Note : Lambertian example of diffuse reflection

Computer Vision

INTRO

perception
applications
**light**

# Spectral reflectance
# e.g. vegetation



Computer Vision

INTRO

perception
applications
**light**

# Ideally: spectral BRDF at all points known

BRDF = bidirectional reflectance distribution function

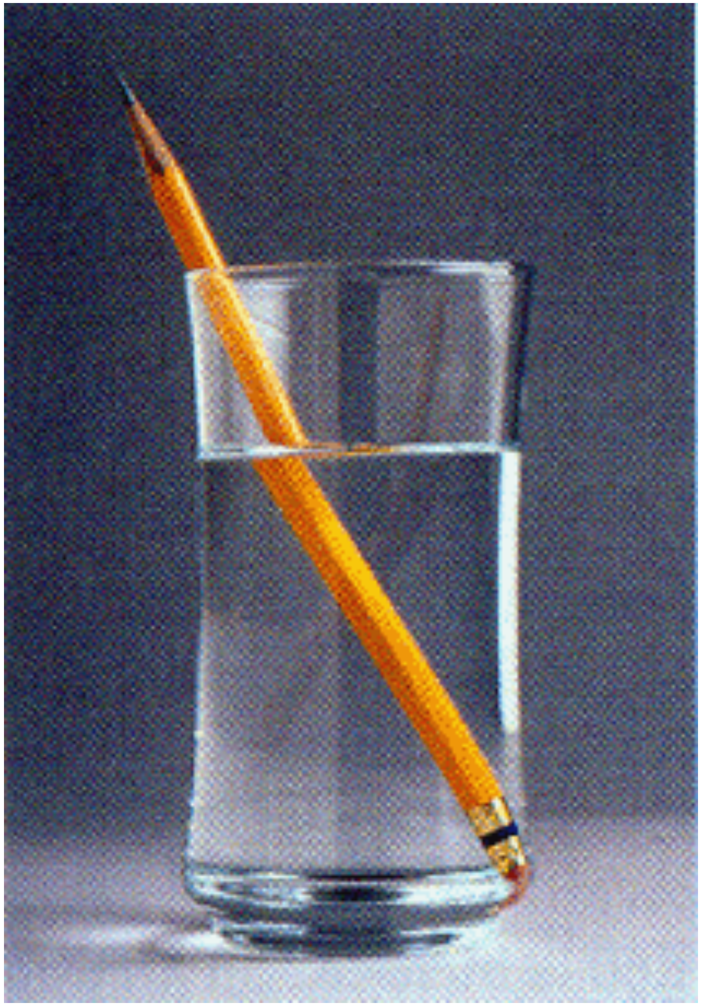
# Interactions with matter

four types :

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

+ diffraction

# Refraction

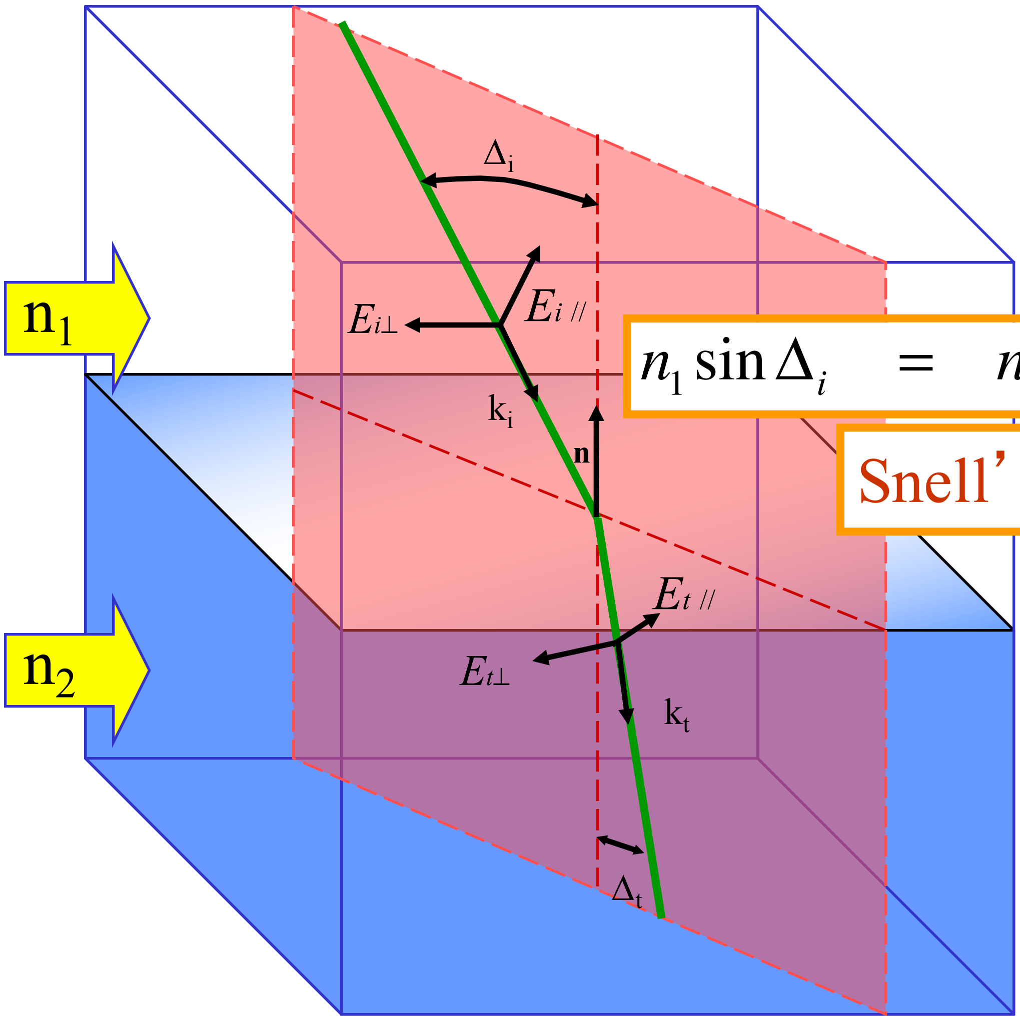


Computer Vision

INTRO

perception
applications
**light**

# Refraction




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

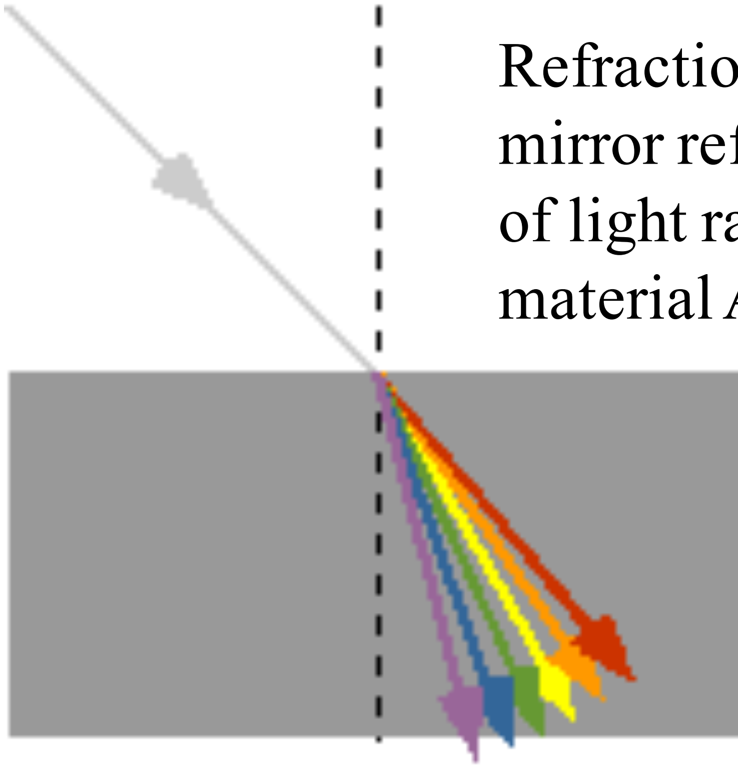
Snell’s law

Computer Vision

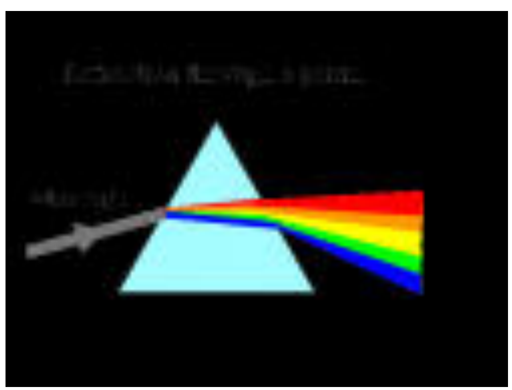
INTRO

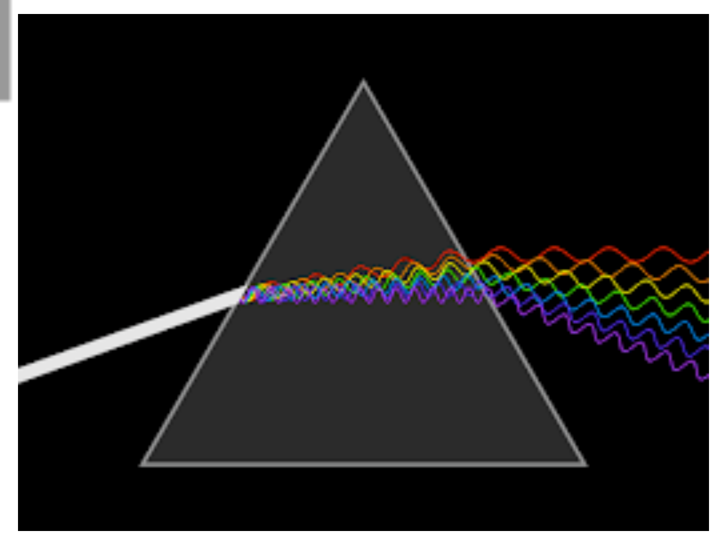
perception
applications
**light**

# Dispersion



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





# Interactions with matter

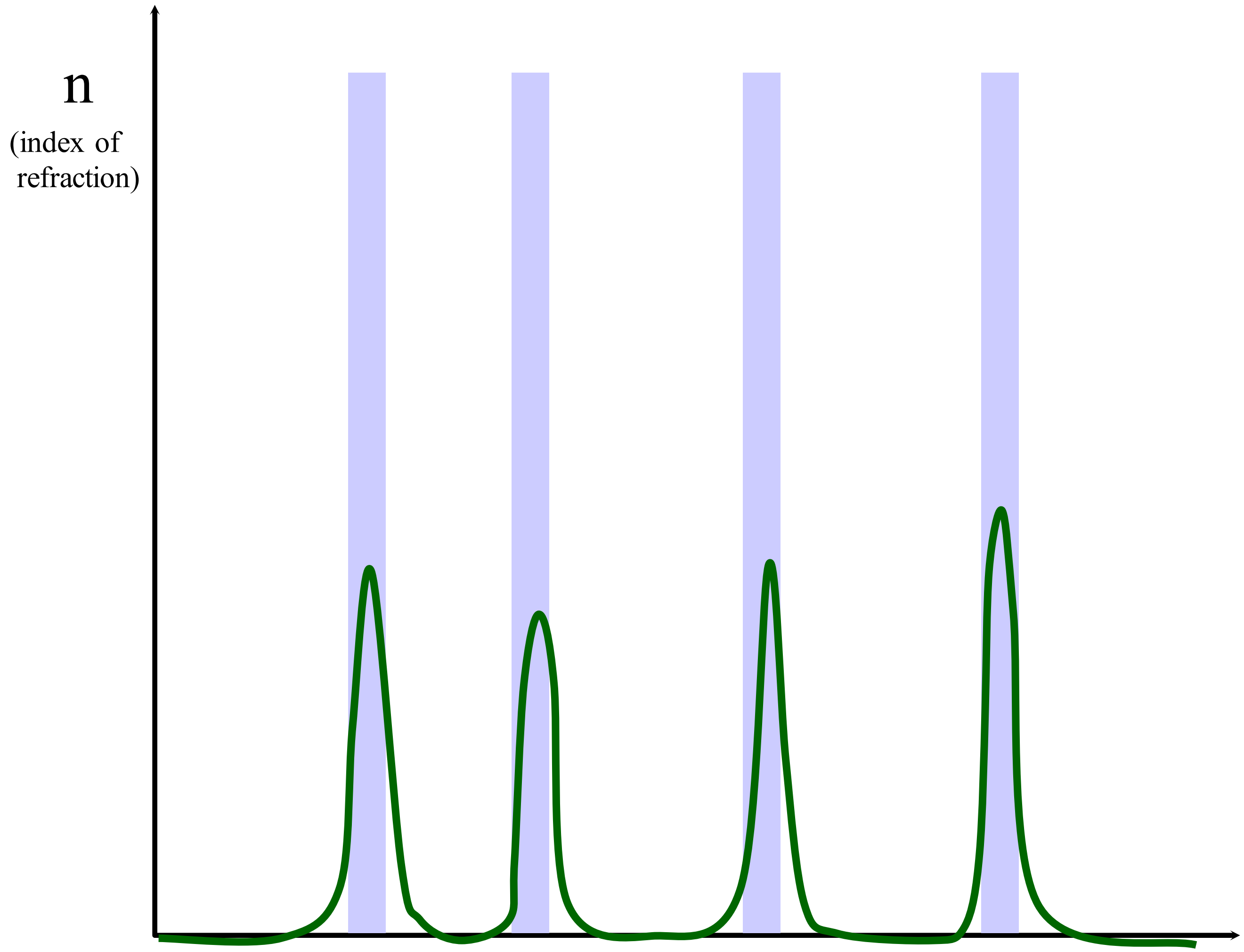
four types :

| 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

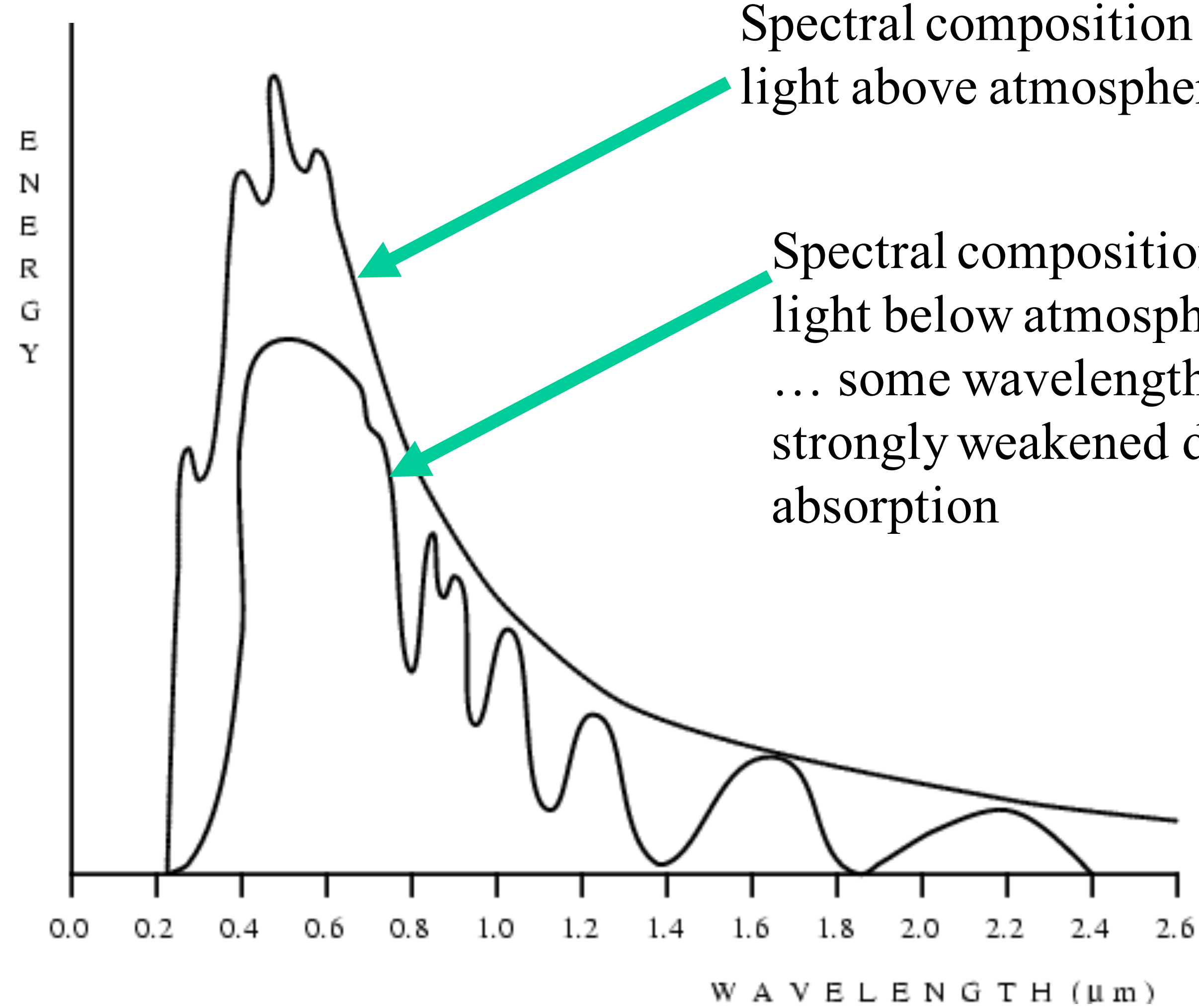



Based on resonance frequencies of molecules -> peaks

Holes in sky light spectrum observed by Fraunhofer

# The solar spectrum

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




Spectral composition of light above atmosphere

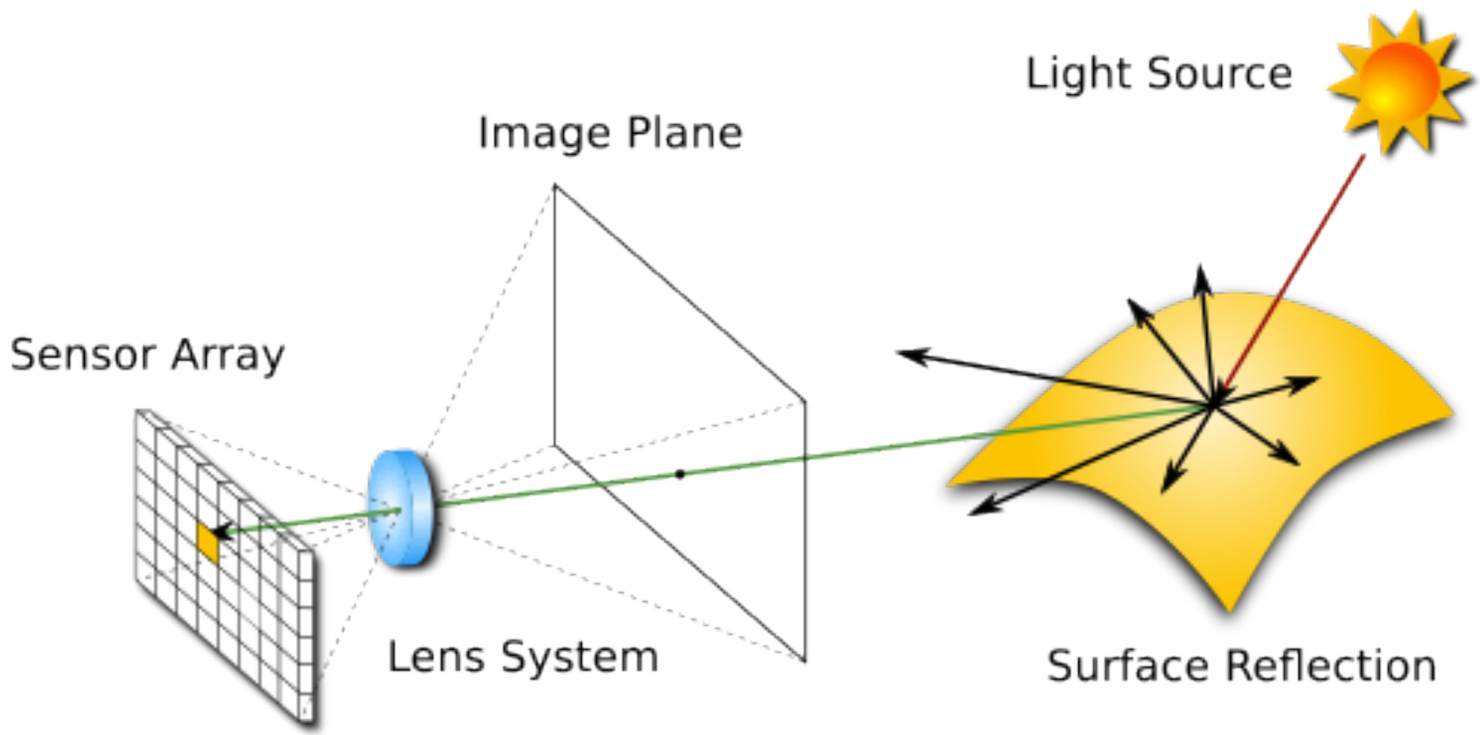
Spectral composition of light below atmosphere … some wavelengths get strongly weakened due to absorption

# Acquisition of Images

# Acquisition of images

We focus on :

1. illumination
2. cameras



Computer Vision

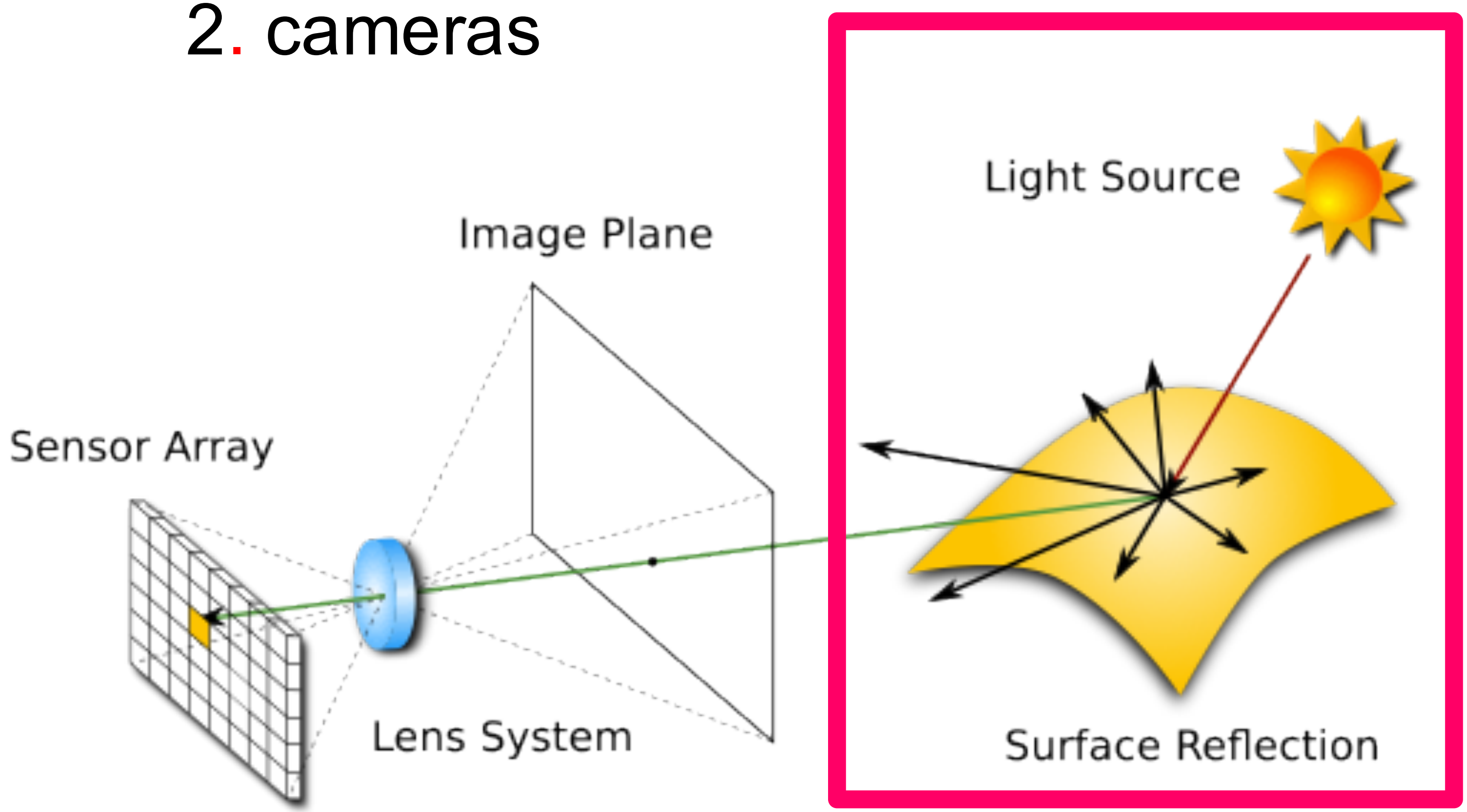
ACQUIS.

illumination
cameras

# Acquisition of images

We focus on :

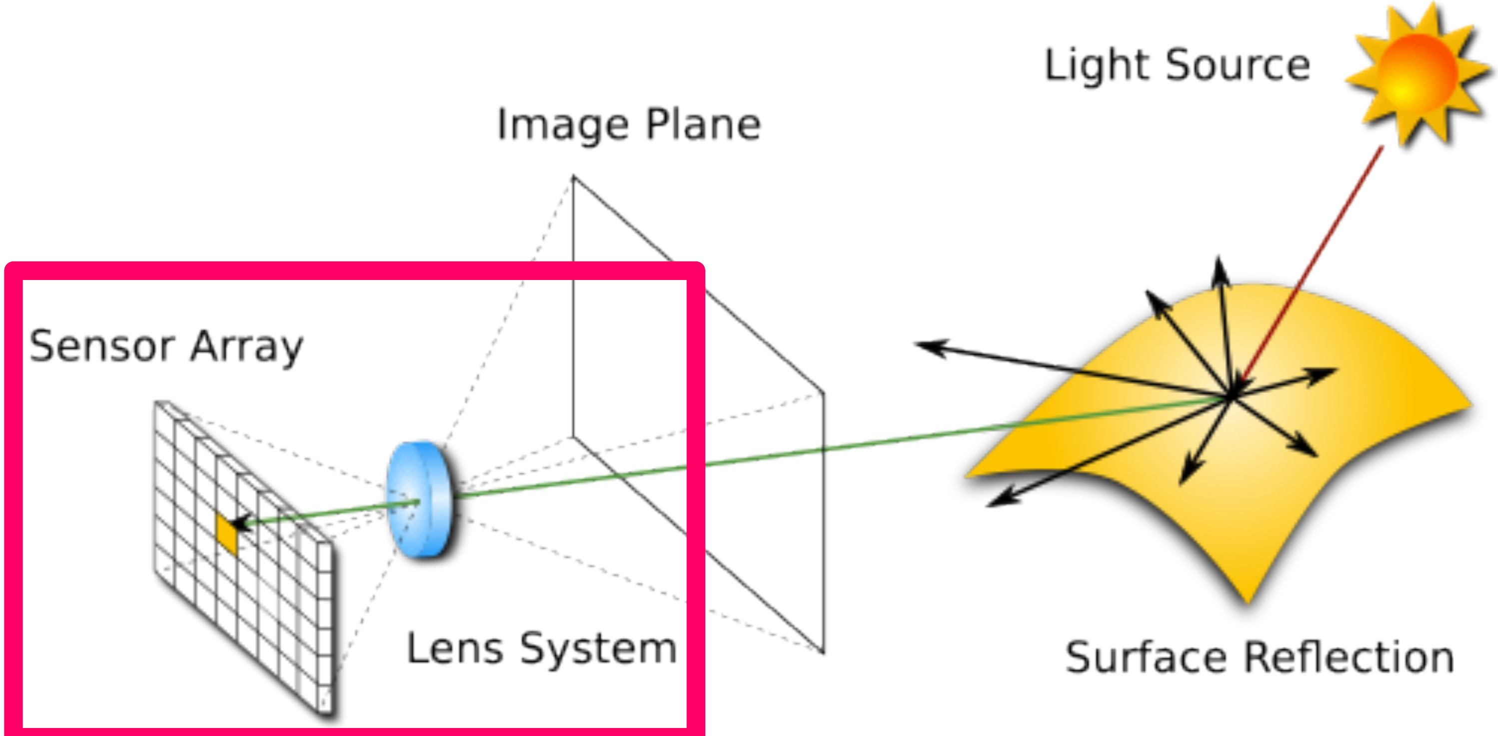
1. illumination
2. cameras



# Acquisition of images

We focus on :

1. illumination
2. cameras



# illumination

Computer Vision

ACQUIS.

**illumination**
cameras

# Illumination

Well-designed illumination often is key in visual inspection


*The light was good, but the hot wax was a problem...*

# Illumination techniques

Simplify the image processing by controlling the environment

## An overview of illumination techniques:

1. back-lighting
2. directional-lighting
3. diffuse-lighting
4. polarized-lighting
5. coloured-lighting
6. structured-lighting
7. stroboscopic lighting

# Back-lighting

lamps placed behind a transmitting diffuser plate, light source behind the object

generates high-contrast silhouette images, easy to handle with *binary vision*

often used in inspection

Computer Vision

ACQUIS.

**illumination**
cameras

# Example backlighting

➔

# Directional and diffuse lighting

## Directional-lighting

generate sharp shadows
generation of specular reflection
(e.g. crack detection)

shadows and shading yield information about shape

## Diffuse-lighting

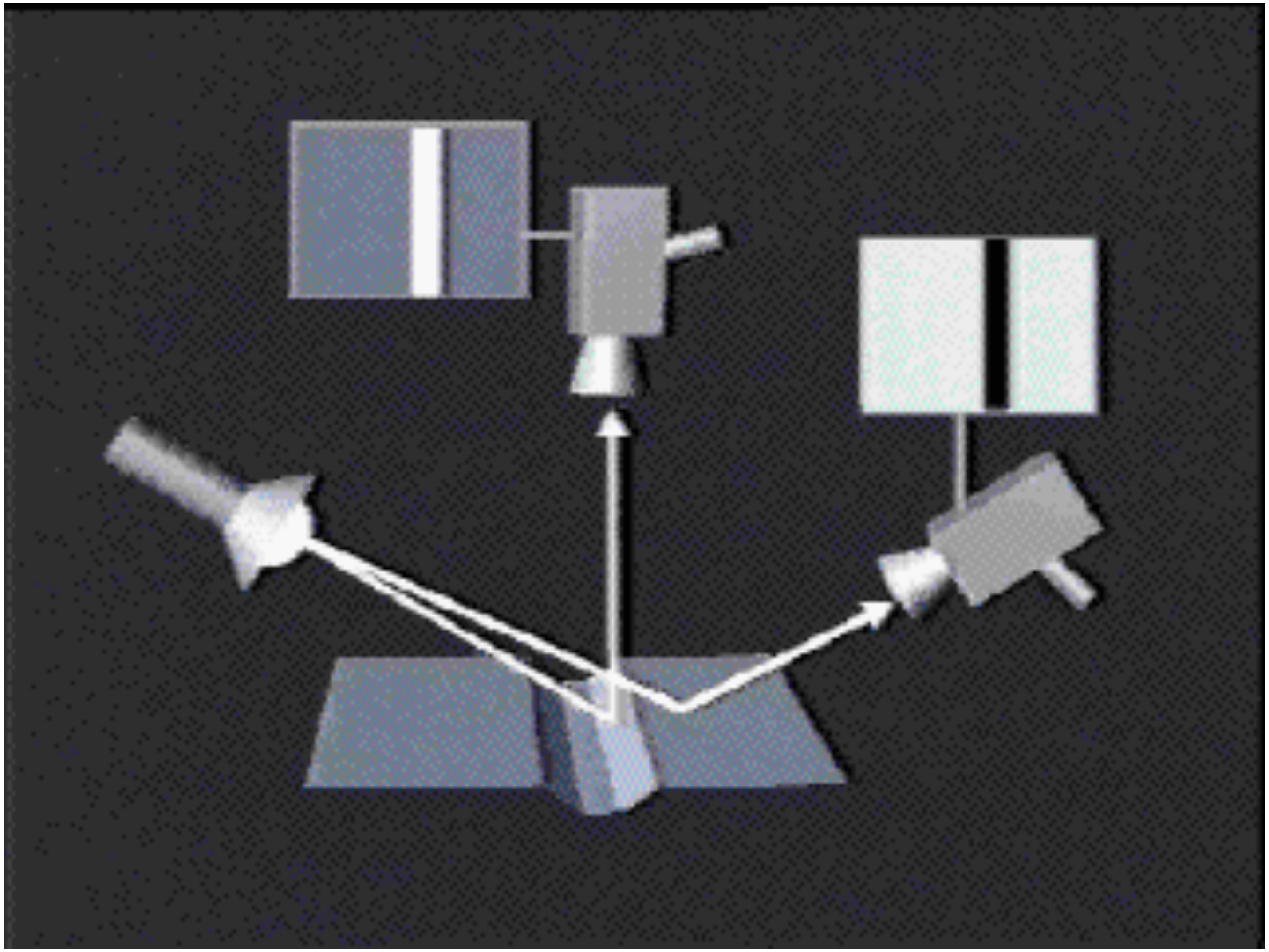
illuminates uniformly from all directions
prevents sharp shadows and large intensity variations over glossy surfaces:
all directions contribute extra diffuse reflection, but contributions to the specular peak arise from directions close to the mirror one only

Computer Vision

ACQUIS.

**illumination**
cameras

# Crack detection



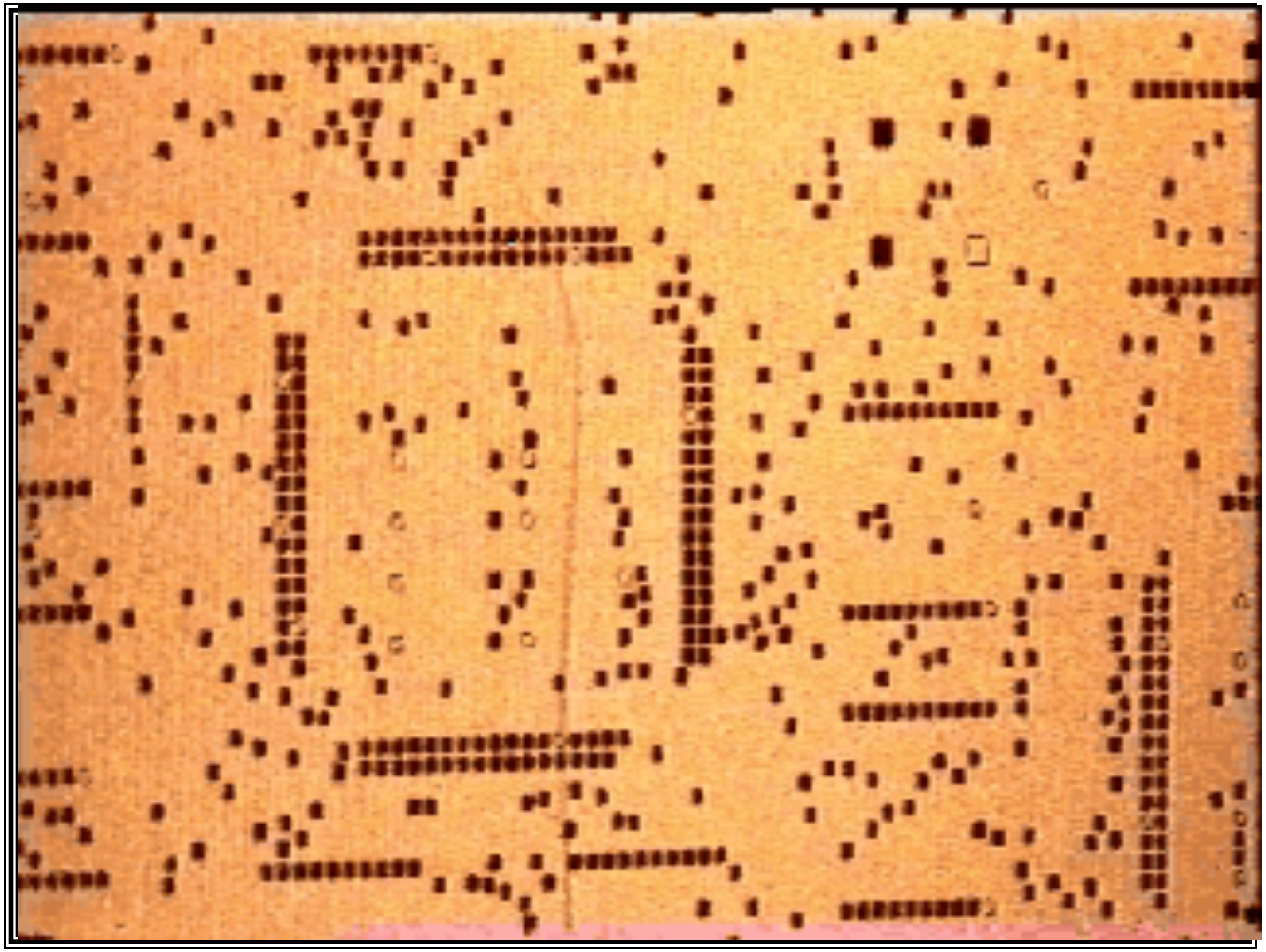
Computer Vision

ACQUIS.

**illumination**

cameras

# Example directional lighting

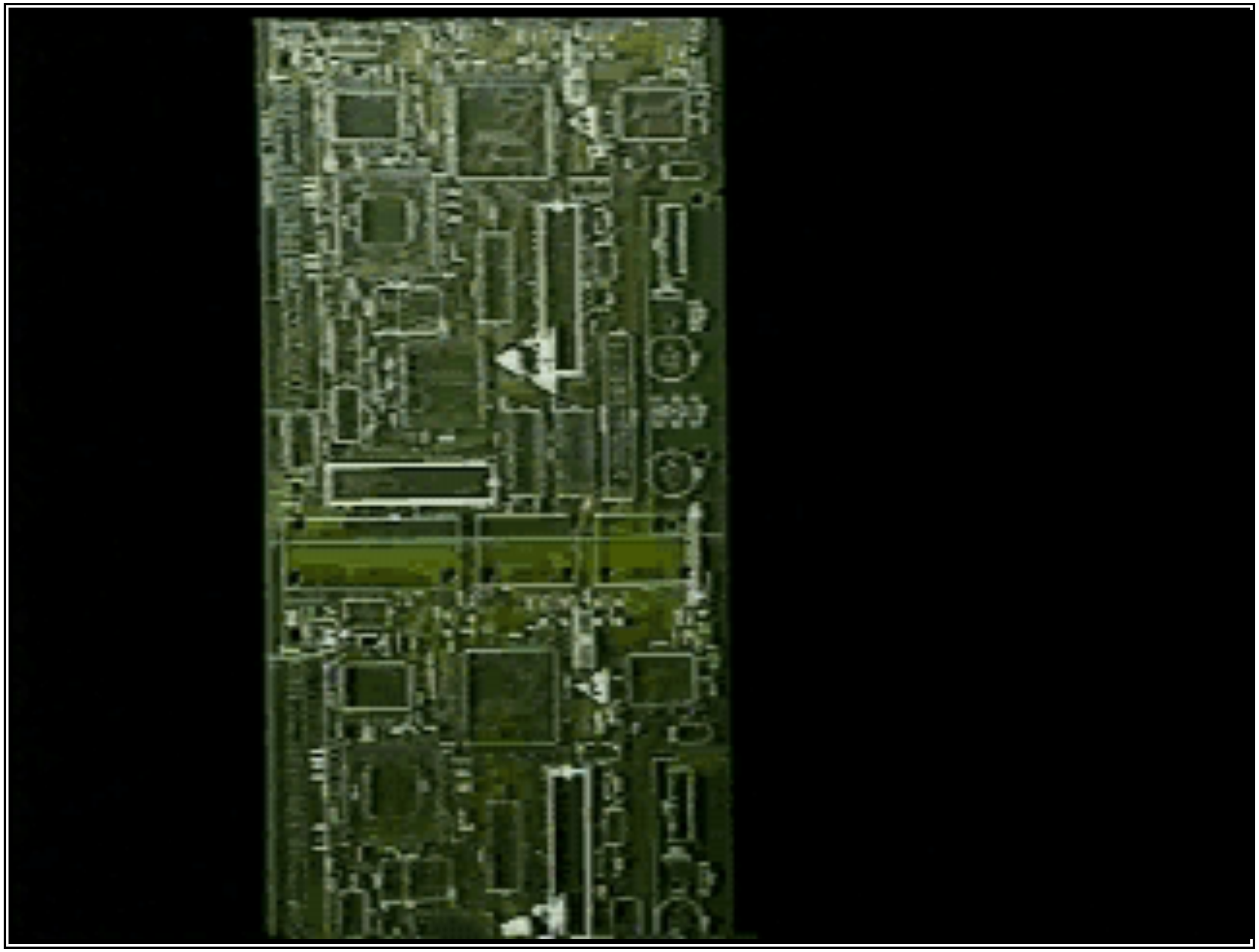


Computer Vision

ACQUIS.

**illumination**
cameras

# Example diffuse lighting



# Polarized lighting

2 uses:

1. to improve contrast between Lambertian and specular reflections

2. to improve contrasts between dielectrics and metals

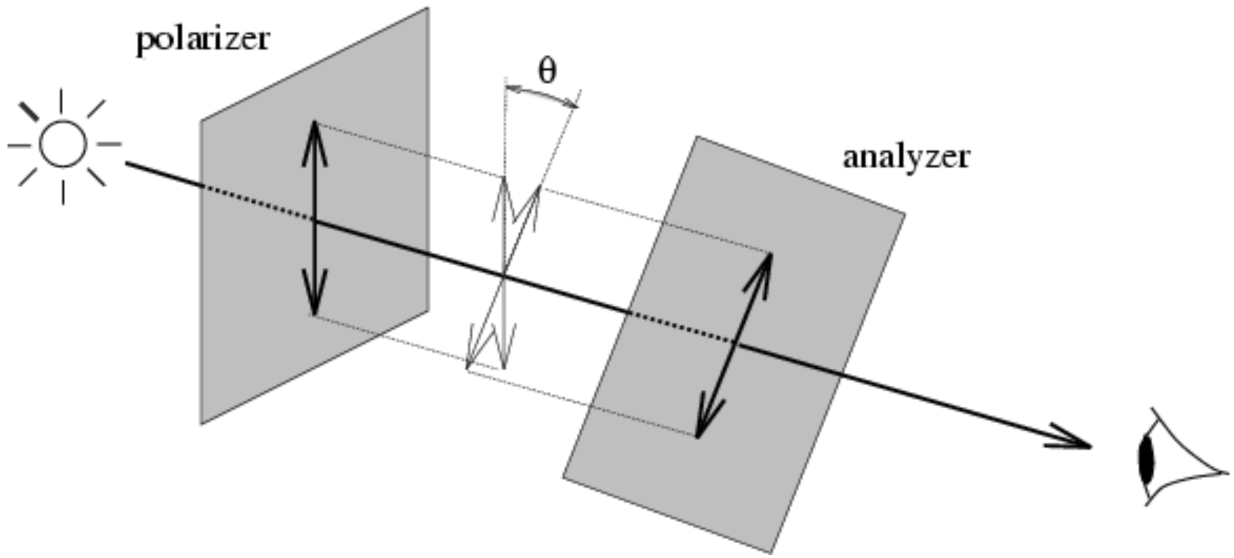
Computer Vision

ACQUIS.

**illumination**
cameras

# Polarised lighting

polarizer/analyzer configurations




law of Malus :

$$I(\theta) = I(0)\cos^2\theta$$

# Polarized lighting

2 uses:

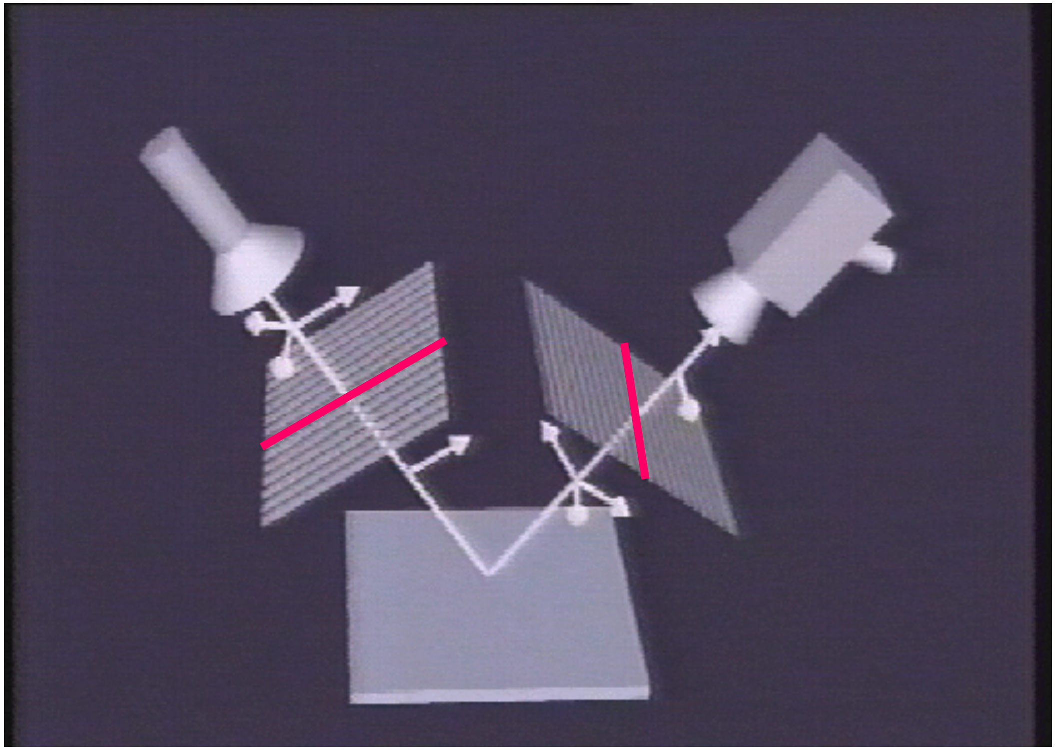
1. to improve contrast between Lambertian and specular reflections

2. to improve contrasts between dielectrics and metals

# Polarized lighting

specular reflection keeps polarisation :
diffuse reflection depolarises

suppression of specular reflection :



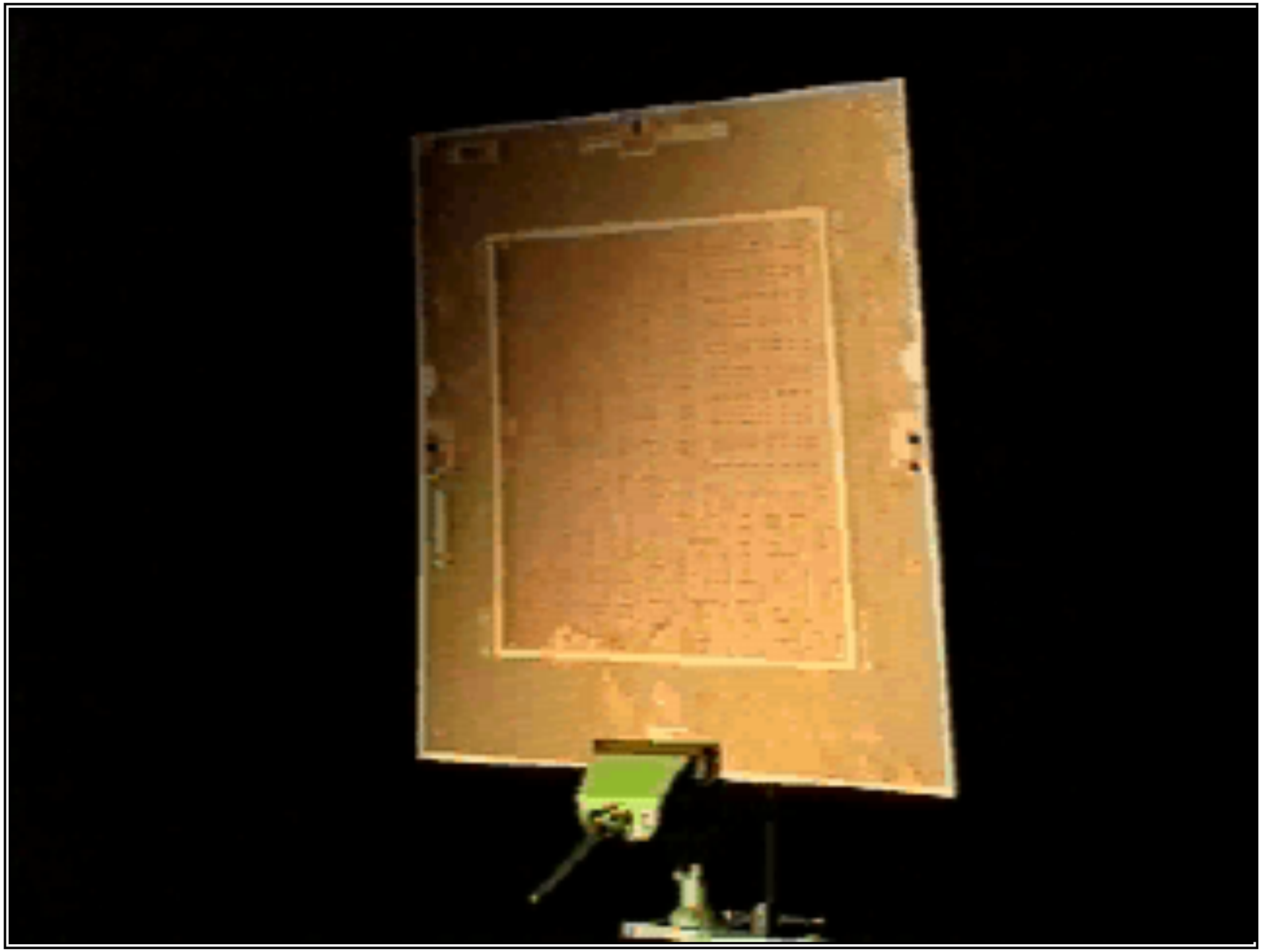
polarizer/analyzer crossed
prevents the large dynamic range caused by glare

Computer Vision

ACQUIS.

**illumination**
cameras

# Example pol. lighting (pol./an.crossed)



Computer Vision

ACQUIS.

**illumination**
cameras

# Polarized lighting

2 uses:

1. to improve contrast between Lambertian and specular reflections

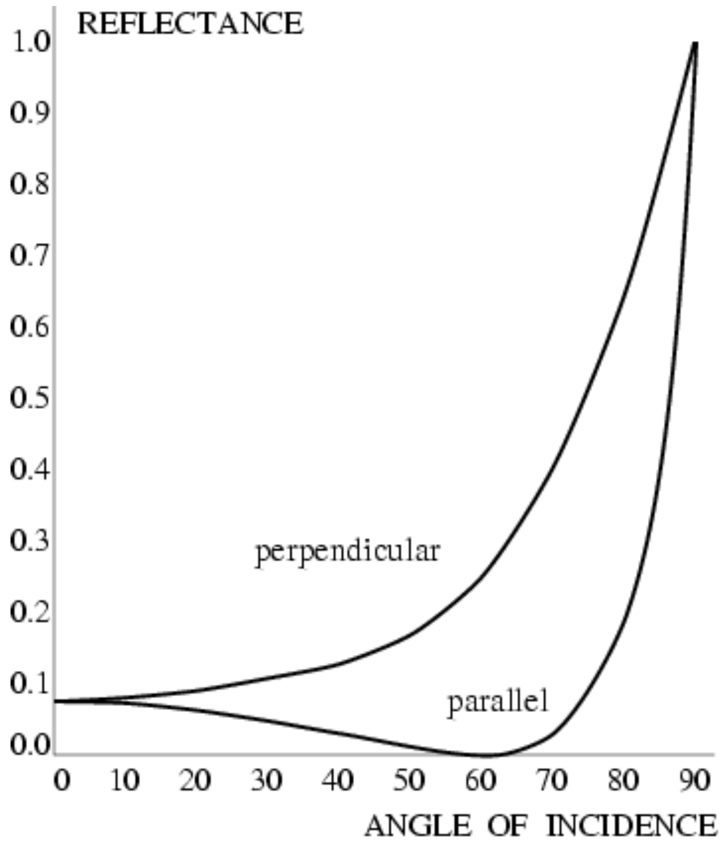
2. to improve contrasts between dielectrics and metals

Computer Vision

ACQUIS.

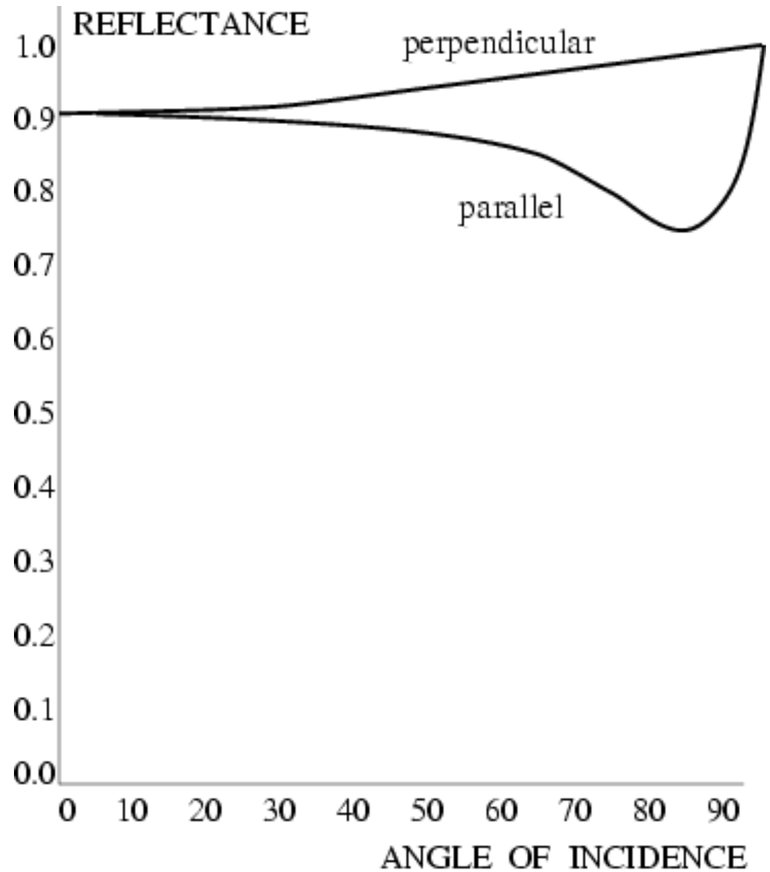
**illumination**
cameras

# Reflection : dielectric




Polarizer at *Brewster angle*

# Reflection : conductor

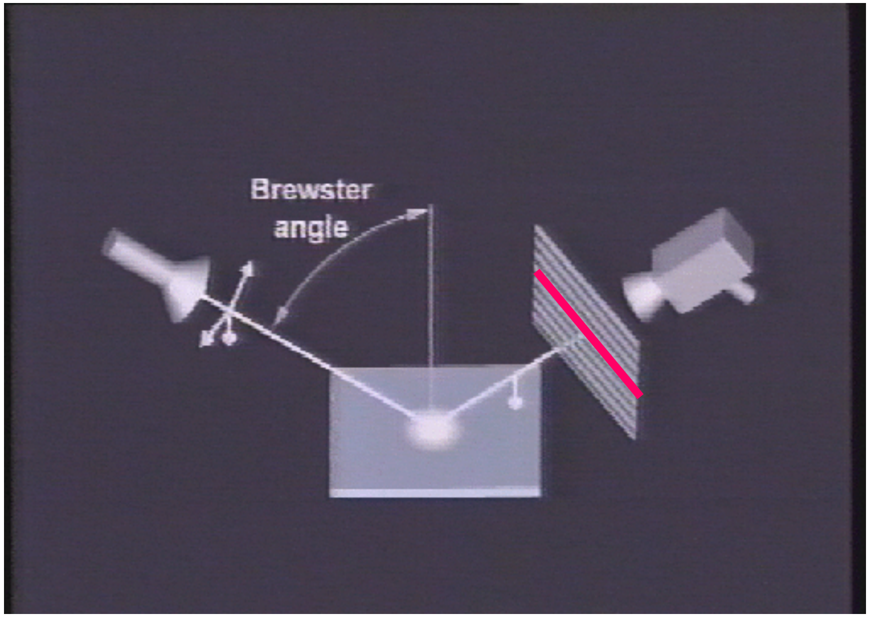



strong reflectors
more or less preserve polarization

# Polarised lighting

distinction between specular reflection from dielectrics and metals;
works under the Brewster angle for the dielectric
dielectric has no parallel comp. ; metal does

suppression of specular reflection from dielectrics :



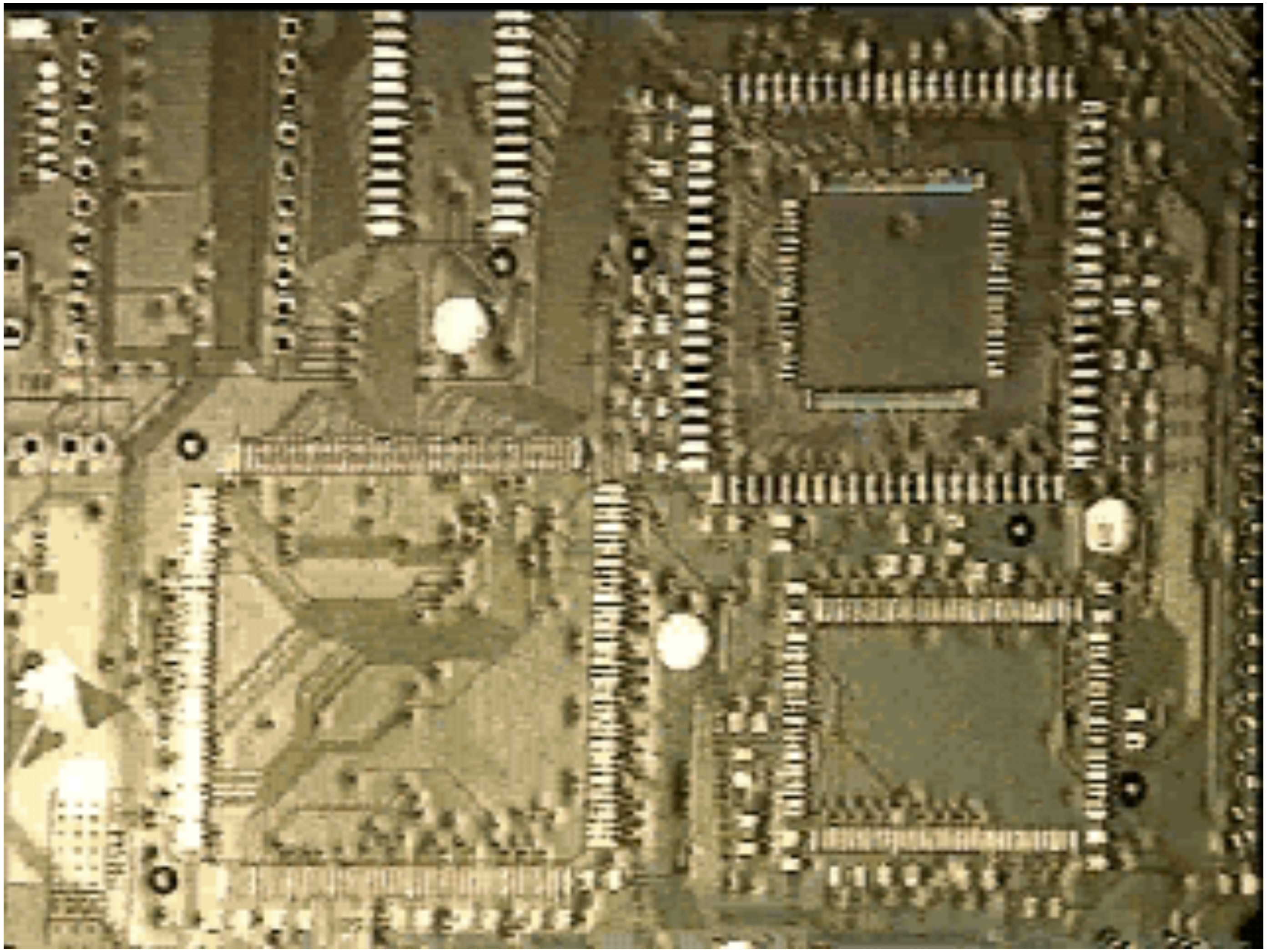

polarizer/analyzer aligned
distinguished metals and dielectrics

Computer Vision

ACQUIS.

**illumination**
cameras

# Example pol. lighting (pol./an. aligned)



➔

# Coloured lighting

highlight regions of a similar colour

with band-pass filter: only light from projected pattern (e.g. monochromatic light from a laser)

differentiation between specular and diffuse reflection

comparing colours ⇨ same spectral composition of sources!

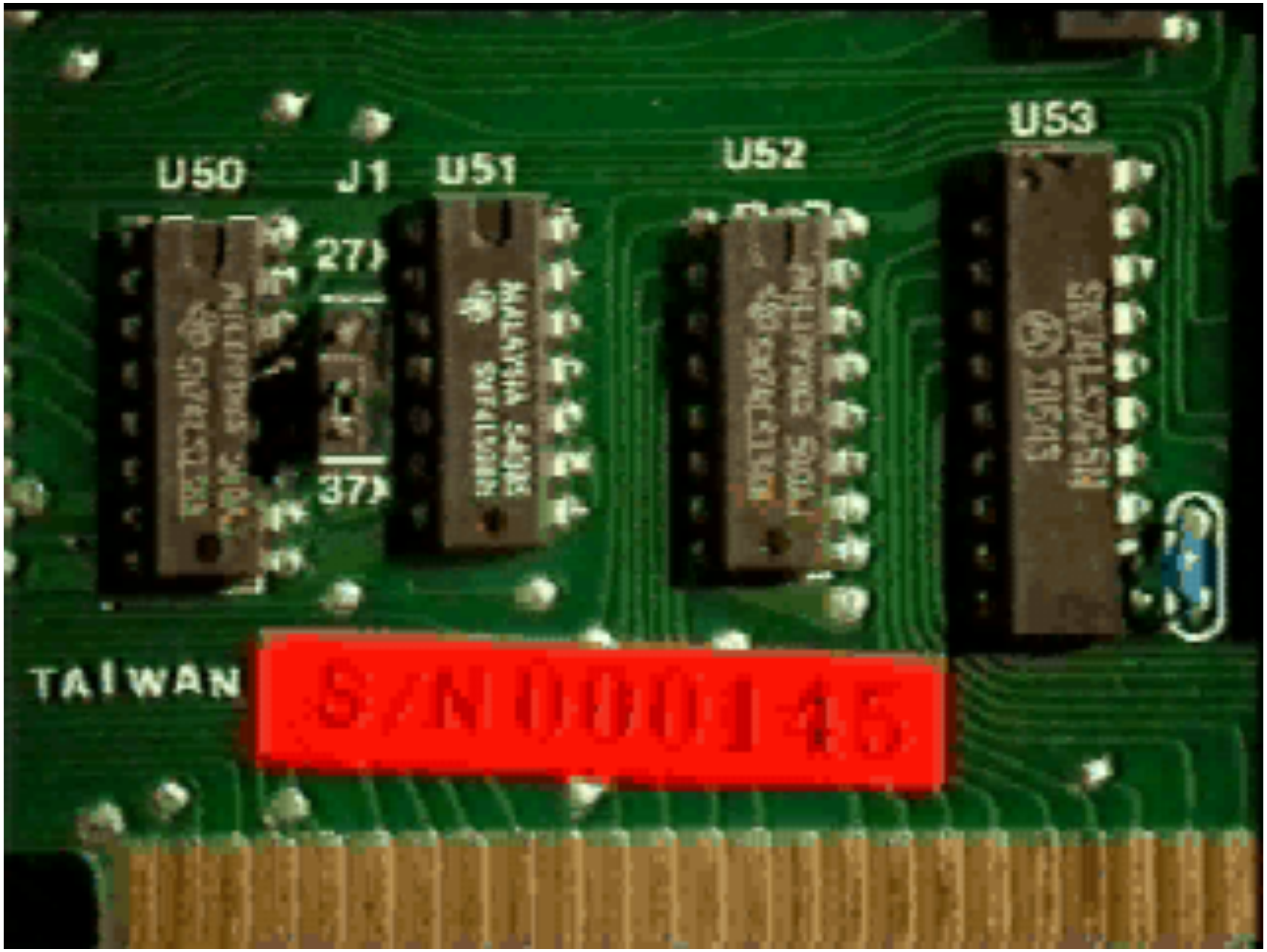
spectral sensitivity function of the sensors!

Computer Vision

# Example coloured lighting

ACQUIS.

**illumination**
cameras



# Structured and stroboscopic lighting

spatially or temporally modulated light pattern

## Structured lighting

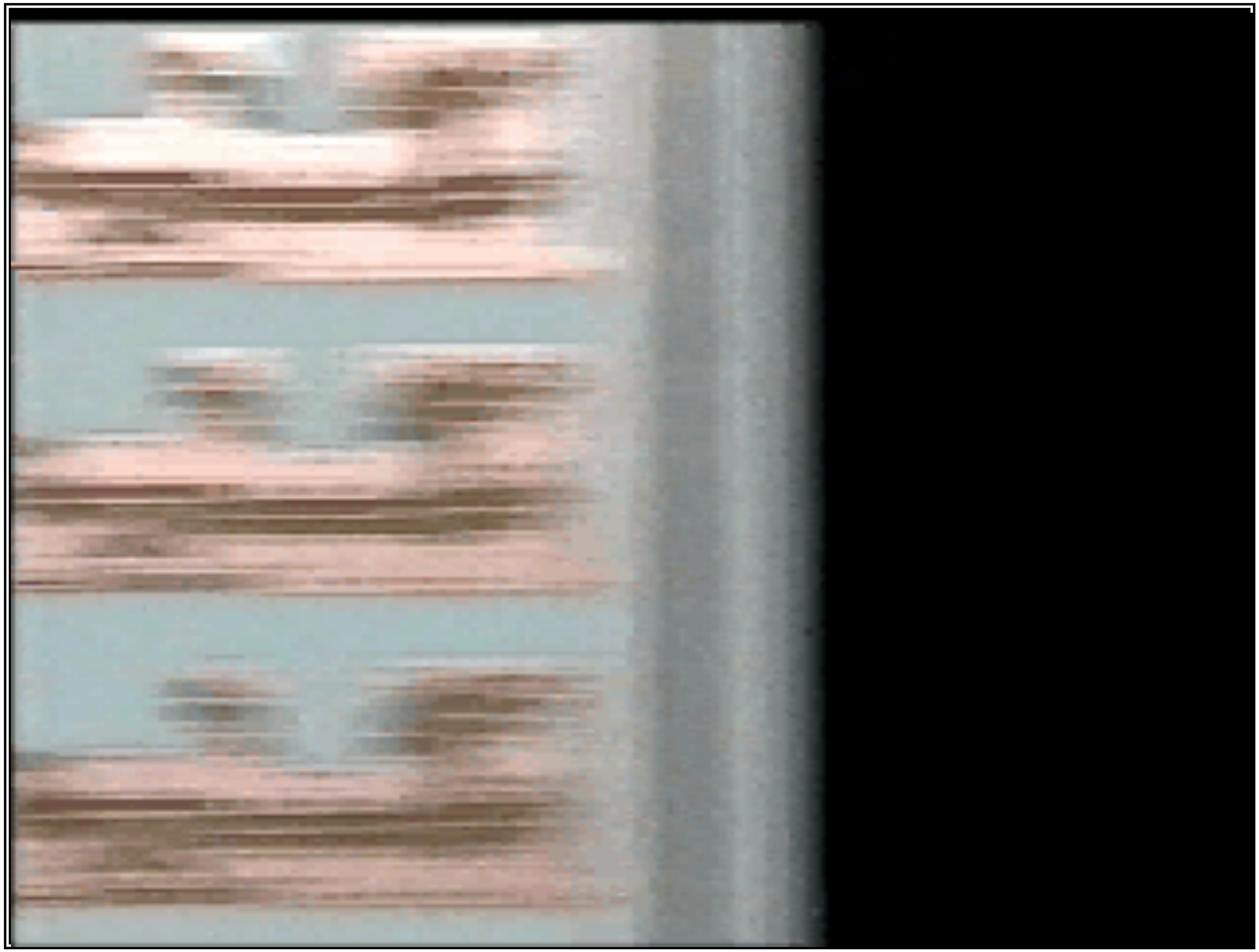
e.g. : 3D shape : objects distort the projected pattern
(more on this later)

## Stroboscopic lighting

high intensity light flash

to eliminate motion blur

# Stroboscopic lighting

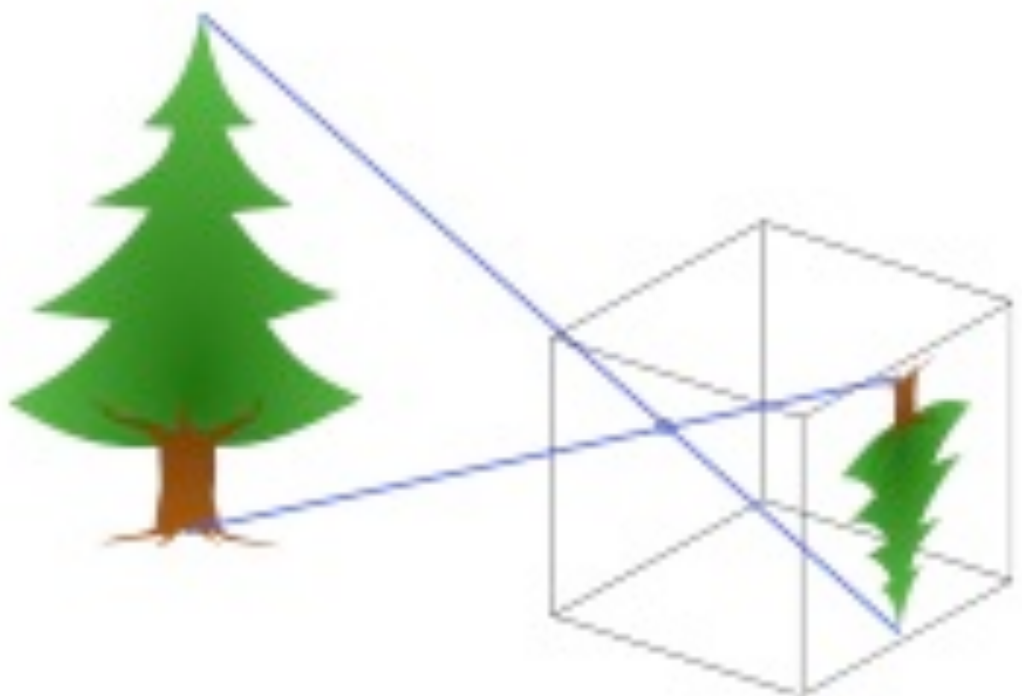


# App: vegetable inspection (colored light + polarization)

# cameras

# Optics for image formation

the pinhole model :



Computer Vision

ACQUIS.

illumination
**cameras**

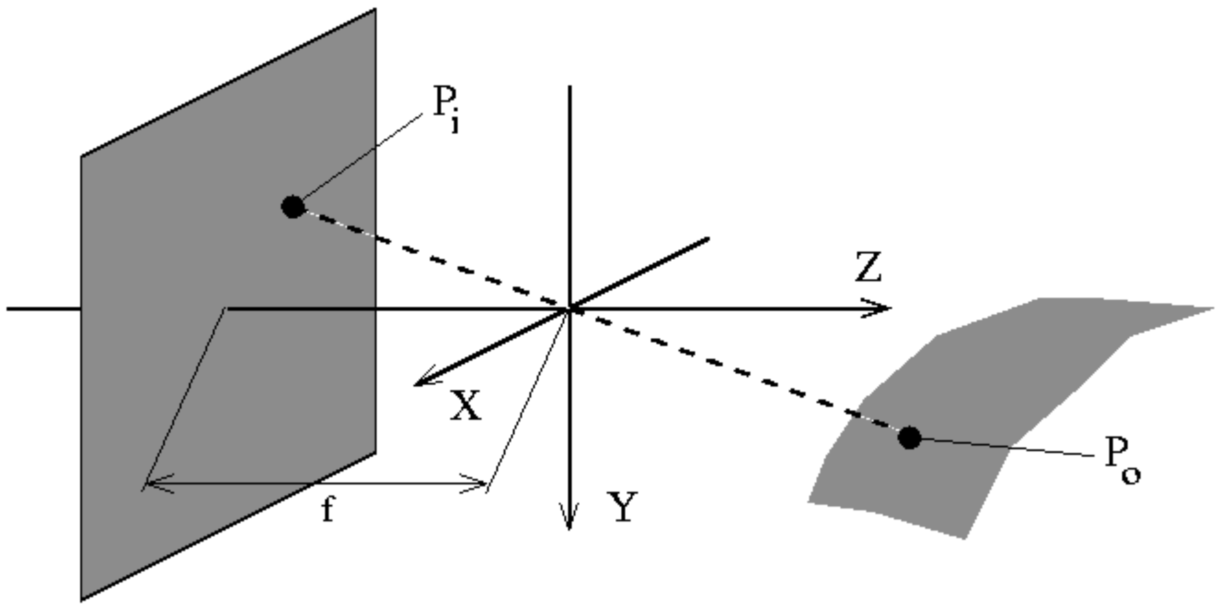
# Optics for image formation

the pinhole model :

hence the name:
CAMERA
obscura

# Optics for image formation

the pinhole model :




$$\frac{X_i}{X_o} = \frac{Y_i}{Y_o} = \frac{f}{-Z_o} = -m$$

*(m = linear magnification)*

# Camera obscura + lens

Computer Vision

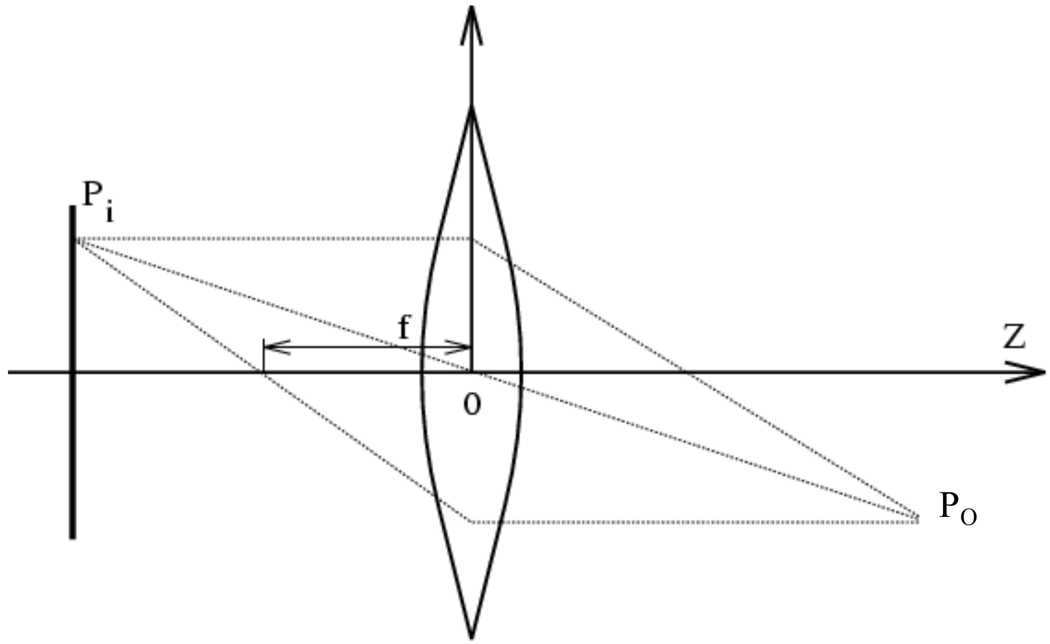
ACQUIS.

illumination
**cameras**

# The thin-lens equation

lens to capture enough light :

$$\frac{1}{Z_O} - \frac{1}{Z_i} = \frac{1}{f}$$




assuming

- spherical lens surfaces
- incoming light ± parallel to axis
- thickness << radii
- same refractive index on both sides

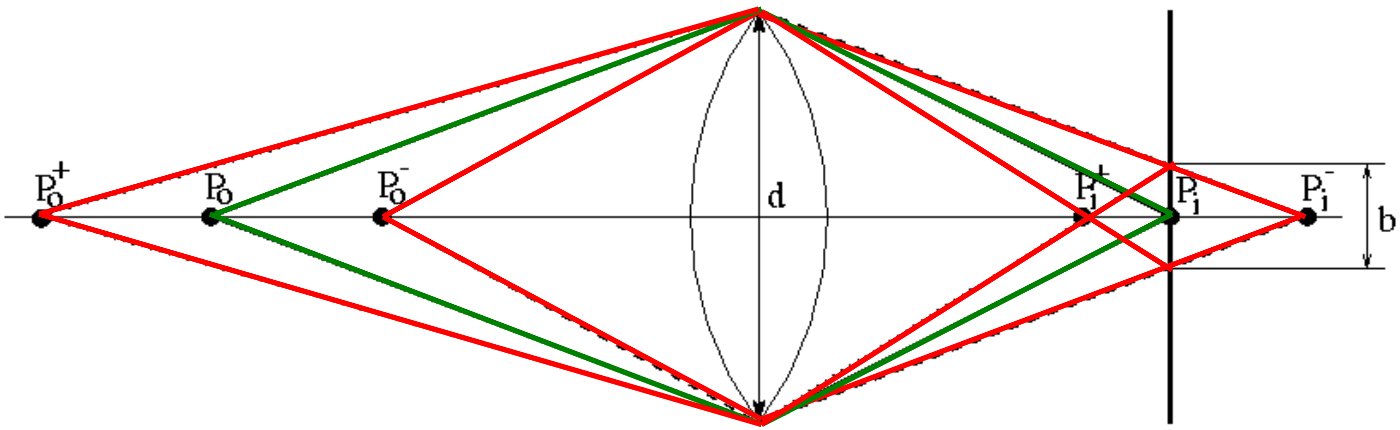
Computer Vision

ACQUIS.

illumination
**cameras**

# The depth-of-field

Only reasonable sharpness in Z-interval




$$\Delta Z_0^- = Z_0 - Z_0^- = \frac{Z_0(Z_0 - f)}{Z_0 + f\, d/b - f}$$

decreases with $d$, increases with $Z_0$

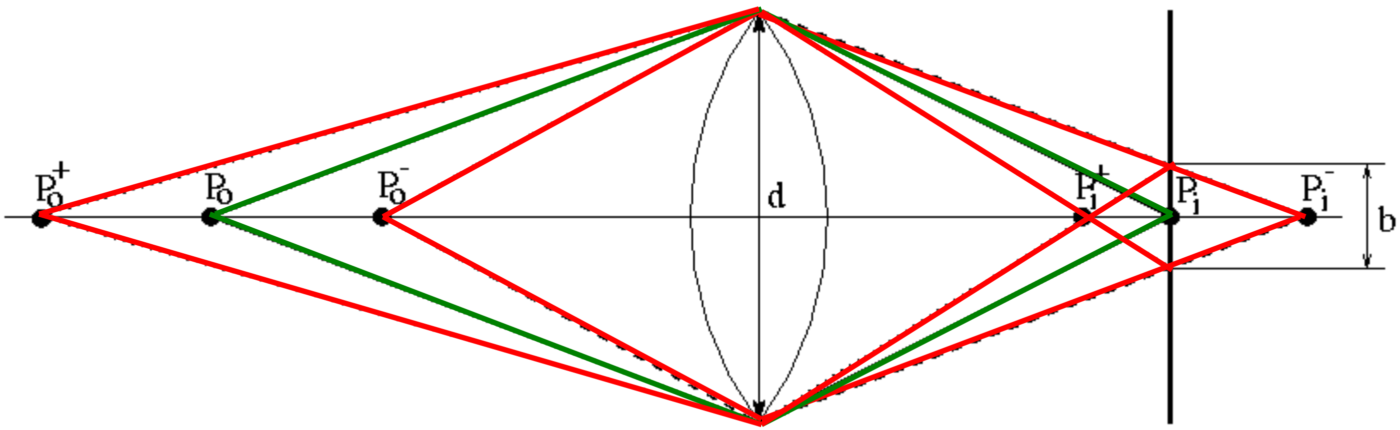
strike a balance between incoming light ($d$) and large depth-of-field (usable depth range)

Computer Vision

ACQUIS.

illumination
**cameras**

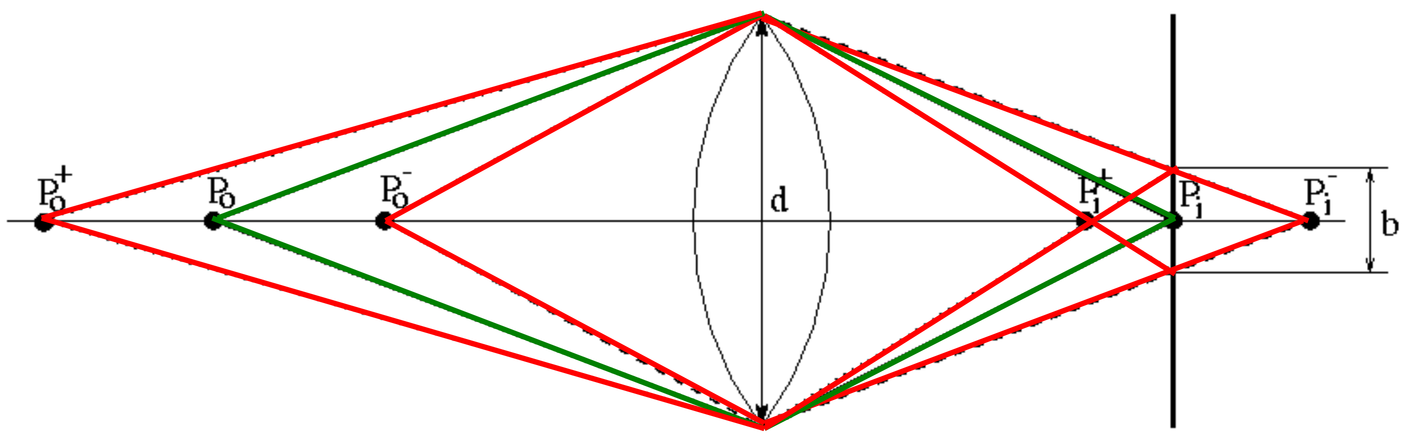
# The depth-of-field




$$\Delta Z_0^- = Z_0 - Z_0^- = \frac{Z_0(Z_0 - f)}{Z_0 + f\, d/b - f}$$

Similar expression for $Z_O^+ - Z_O$

# The depth-of-field




$$\Delta Z_0^- = Z_0 - Z_0^- = \frac{Z_0(Z_0 - f)}{Z_0 + f\, d / b - f}$$

Ex 1: microscopes -> small DoF

Ex 2: special effects -> flood miniature scene with light

# Deviations from the lens model

3 assumptions :

1. all rays from a point are focused onto 1 image point

2. all image points in a single plane

3. magnification is constant

deviations from this ideal are *aberrations*

# Aberrations

2 types :

1. geometrical

2. chromatic

*geometrical :* small for paraxial rays

*chromatic :* refractive index function of wavelength (Snell's law !!)

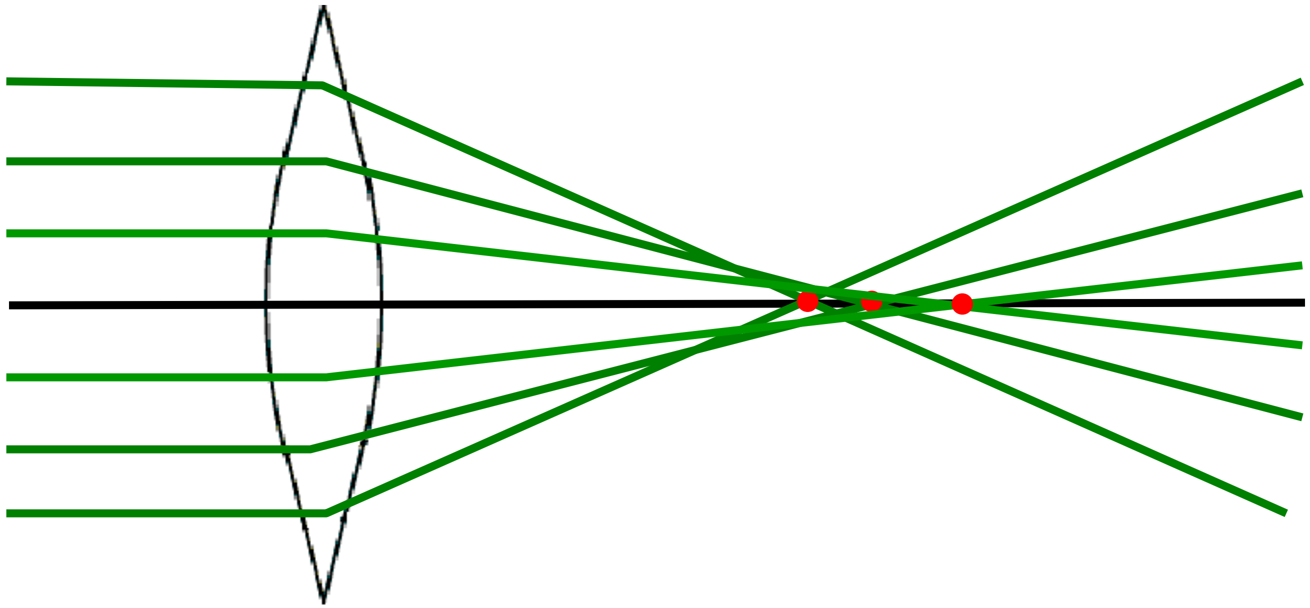
# Geometrical aberrations

- spherical aberration
- astigmatism
- radial distortion ← the most important type
- coma

# Spherical aberration

rays parallel to the axis do not converge

outer portions of the lens yield smaller focal lenghts



Computer
Vision

ACQUIS.

illumination
**cameras**

# Radial Distortion

magnification different for different angles of inclination

*barrel*

*none*

*pincushion*

Computer Vision

ACQUIS.

illumination
**cameras**

# Radial Distortion

magnification different for different angles of inclination

*barrel*

*none*

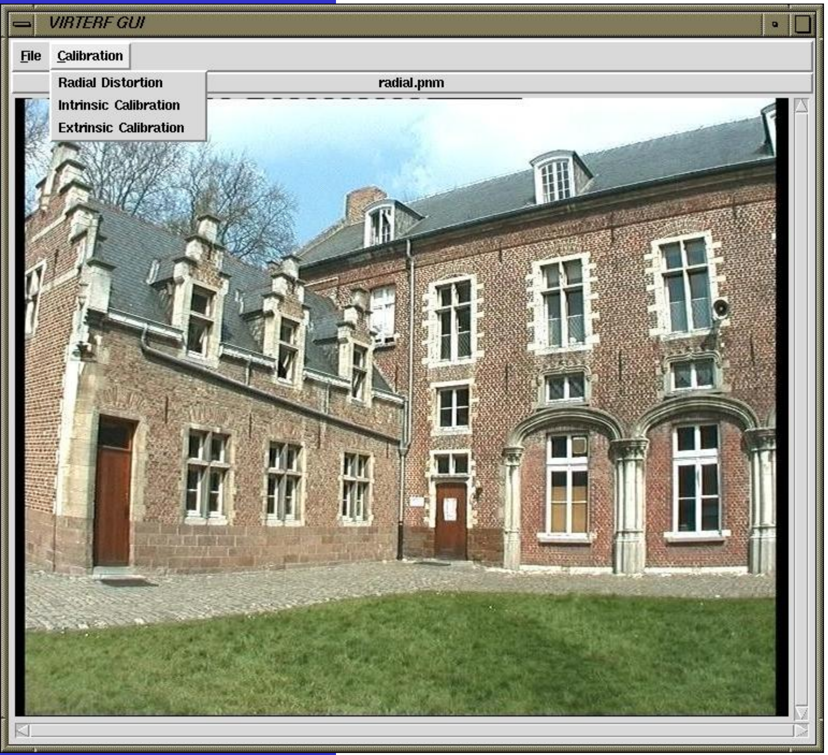
*pincushion*

The result is pixels moving along lines through the center of the distortion – typically close to the image center – over a distance $d$, depending on the pixels' distance $r$ to the center

$$d = (1 + \kappa_1 r^2 + \kappa_2 r^4 + \ldots)$$

# Radial Distortion

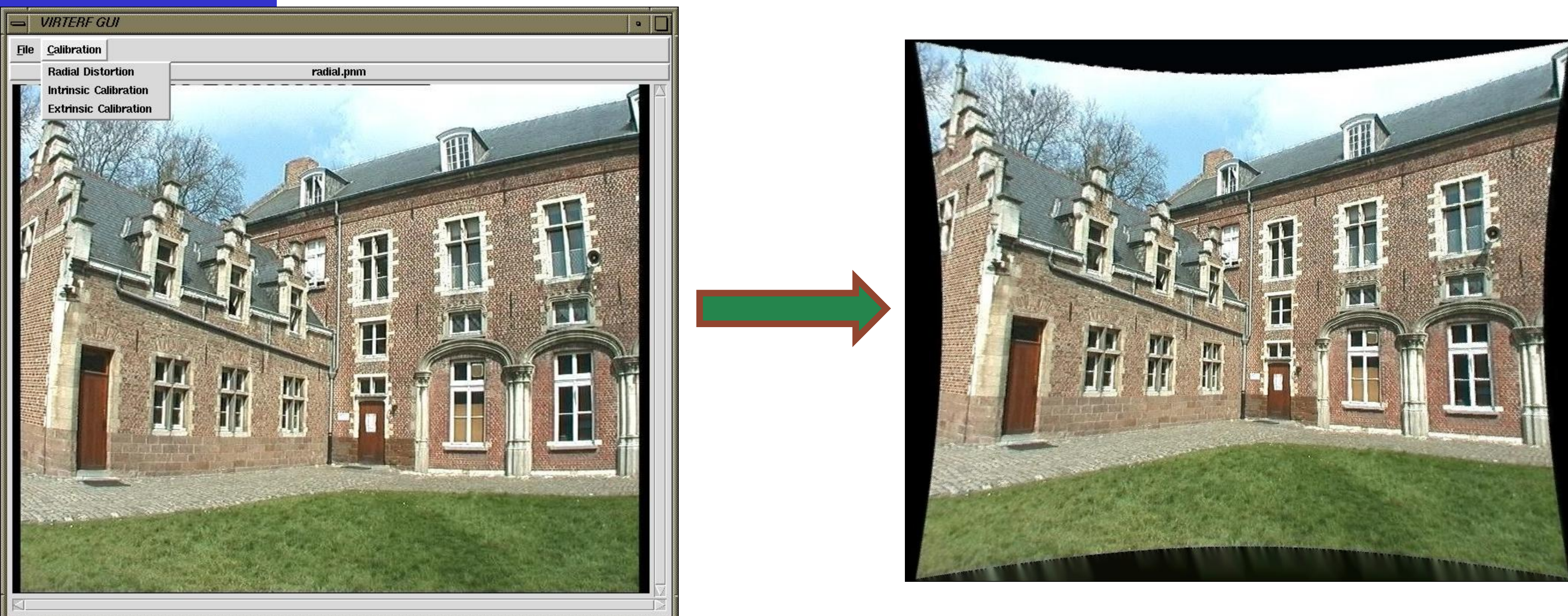
magnification different for different angles of inclination




This aberration type can be corrected by software if the parameters ( $\kappa_1$ , $\kappa_2$ , … ) are known

# Radial Distortion

magnification different for different angles of inclination




Some methods do this by looking how straight lines curve instead of being straight

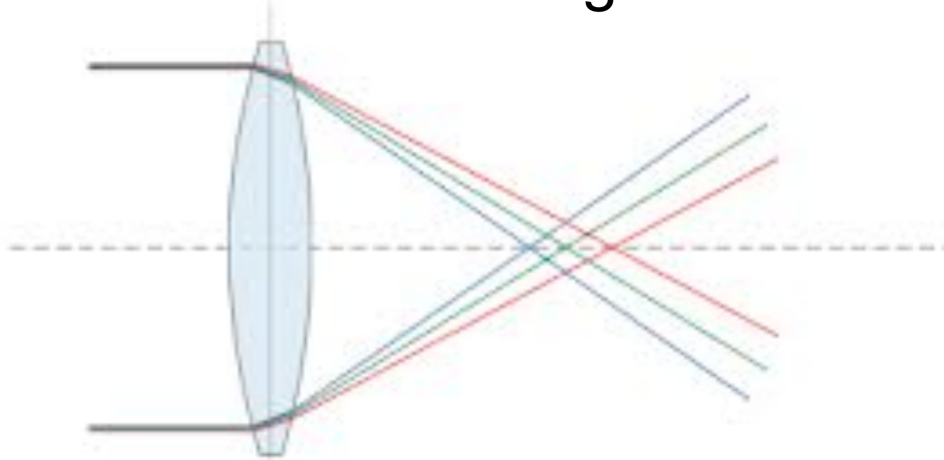
Computer Vision

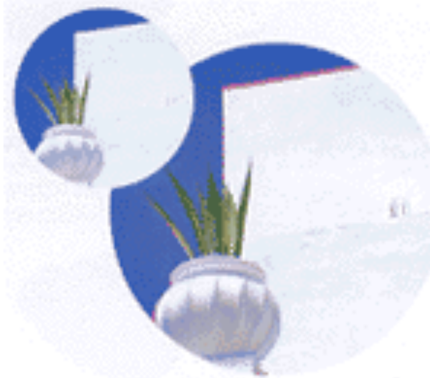
ACQUIS.

illumination
**cameras**

# Chromatic aberration

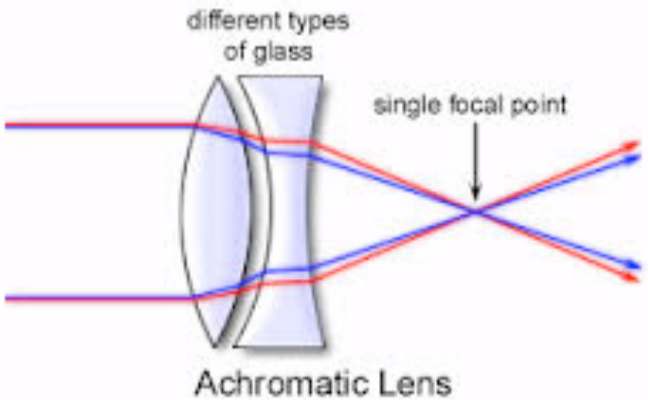
rays of different wavelengths focused in different planes





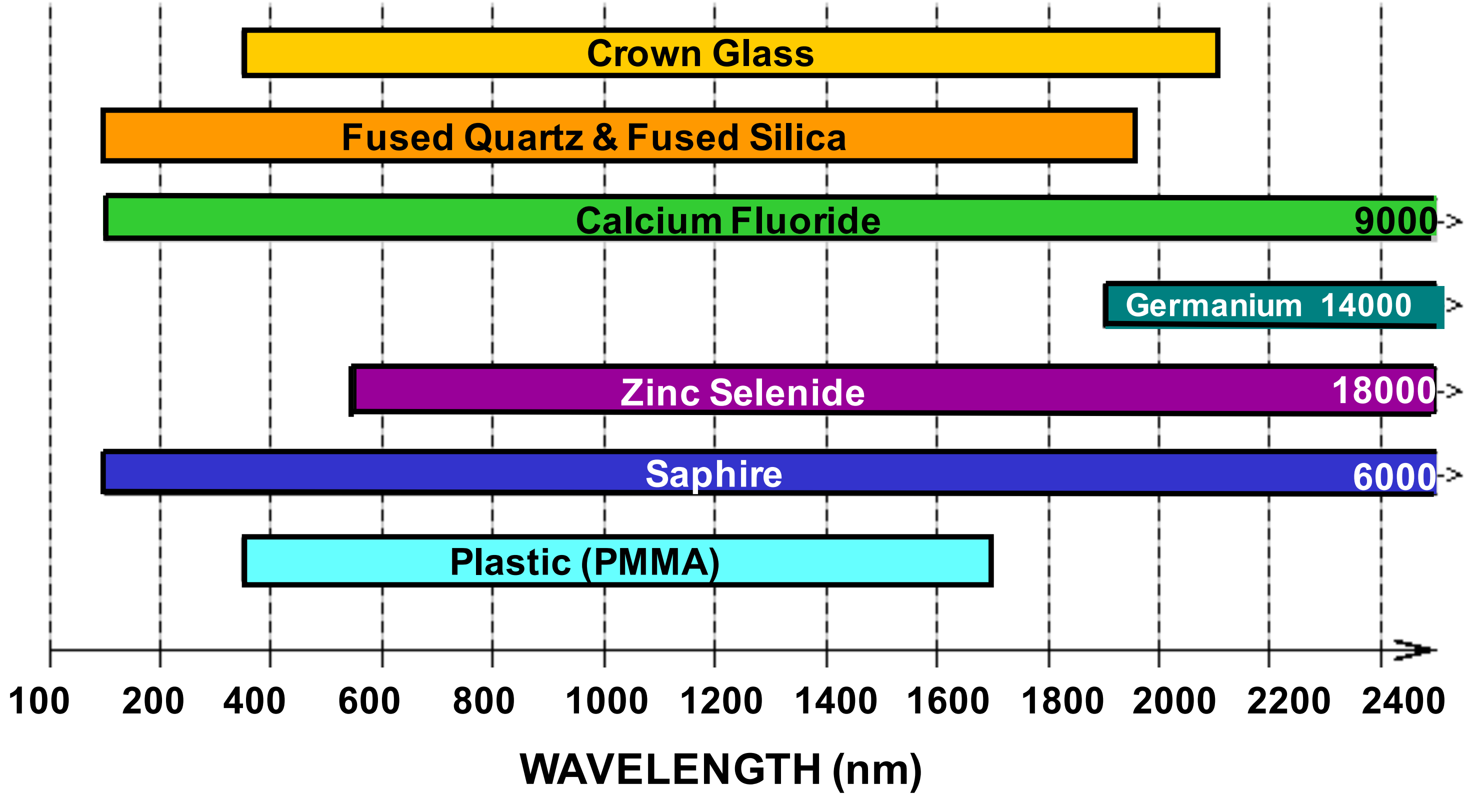
The image is blurred and appears colored at the fringe.

cannot be removed completely
but *achromatization* can be achieved at some well chosen wavelength pair, by combining lenses made of different glasses




sometimes *achromatization* is achieved for more than 2 wavelengths

# Lens materials




the figure shows wavelengths that materials let pass

additional considerations :

humidity and temperature resistance, weight, price,...

# Cameras

we consider 2 types :

1. CCD

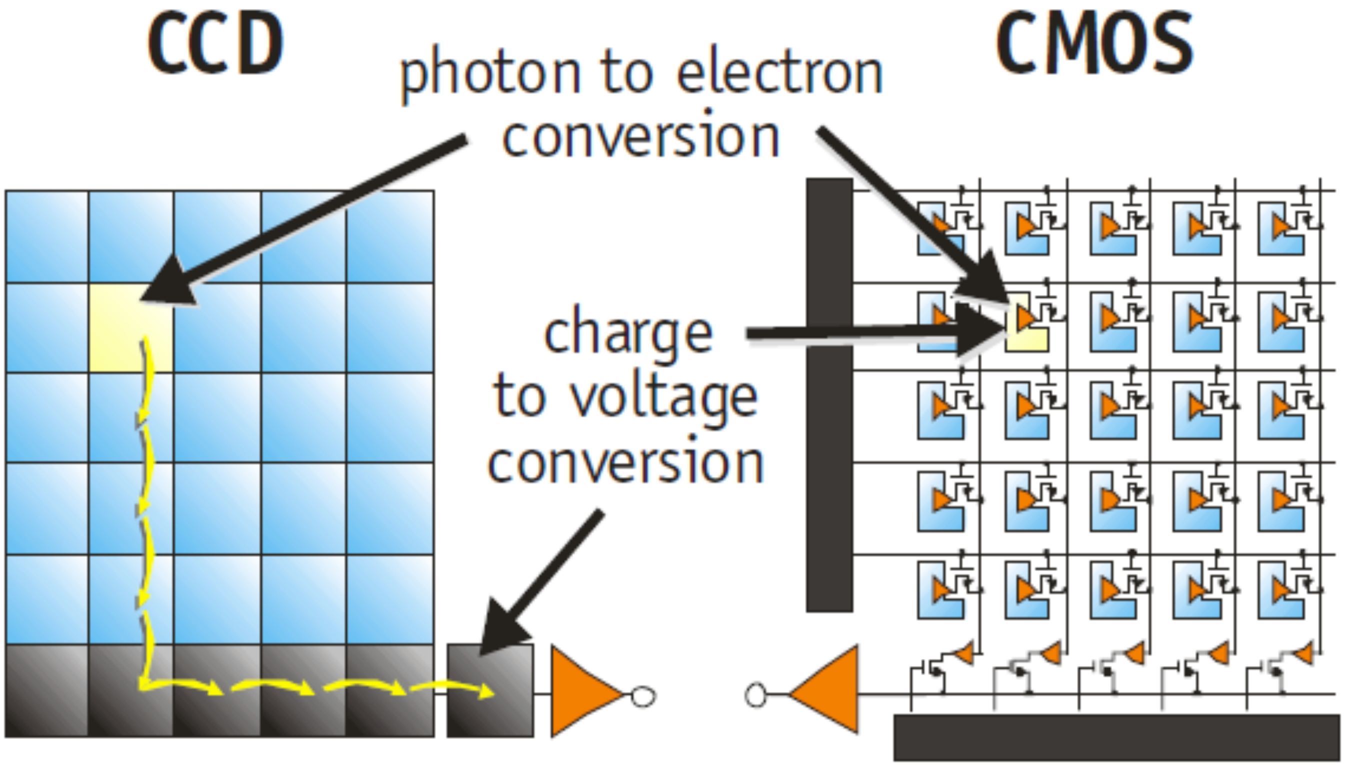
2. CMOS

Computer Vision

ACQUIS.

illumination
**cameras**

# Cameras




CCD = Charge-coupled device
CMOS = Complementary Metal Oxide Semiconductor

Computer Vision

ACQUIS.

illumination

**cameras**

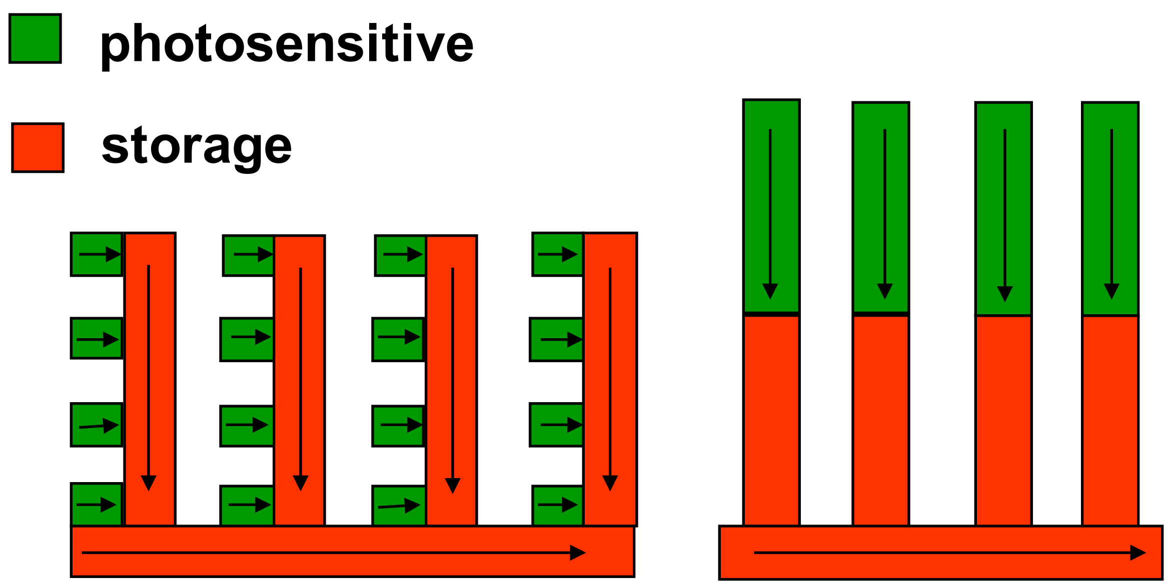
# CCD

separate photo sensor at regular positions
no scanning

charge-coupled devices (CCDs)

area CCDs and linear CCDs
2 area architectures :

*interline transfer* and *frame transfer*

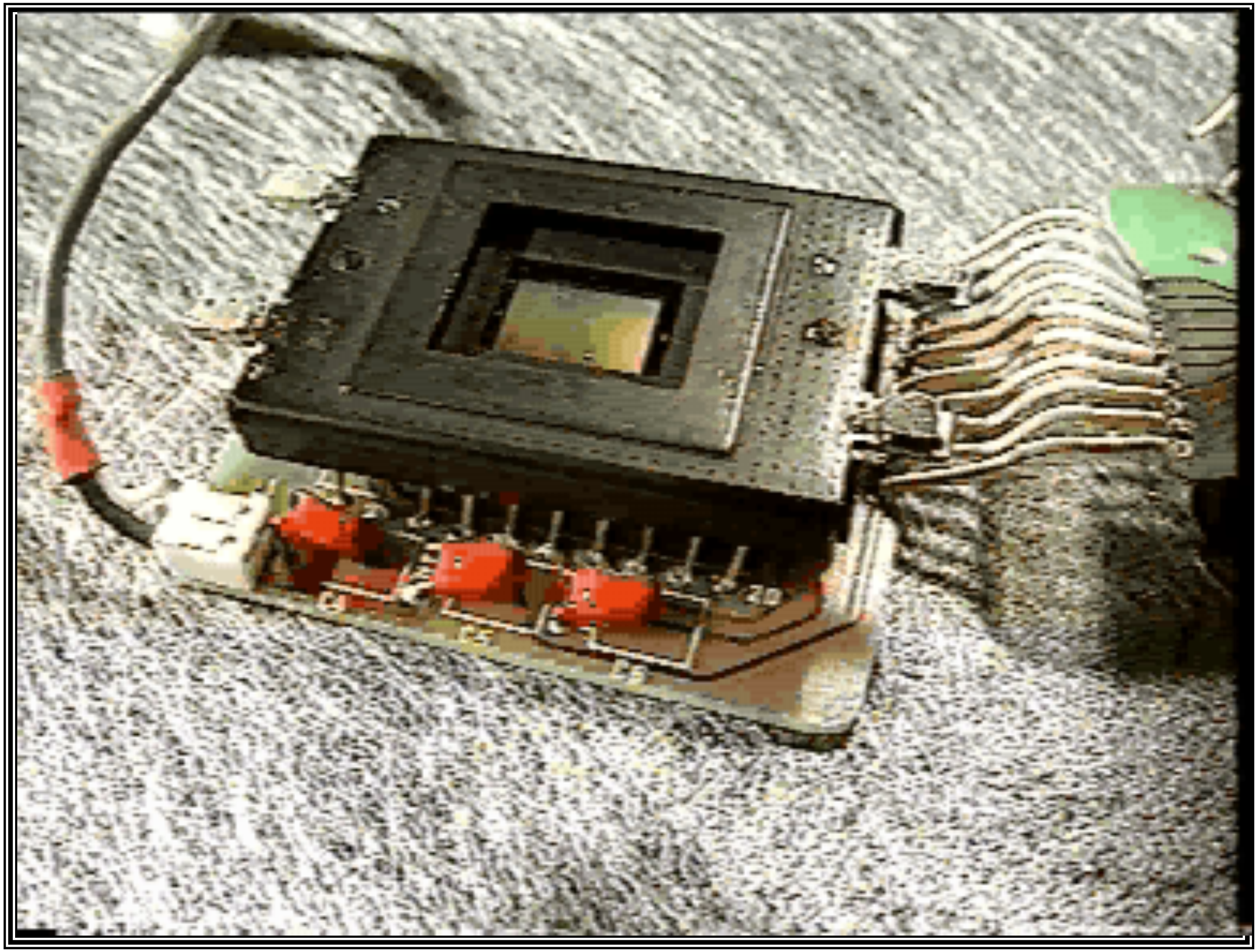


Computer Vision

ACQUIS.

illumination
**cameras**

# The CCD (inter-line) camera



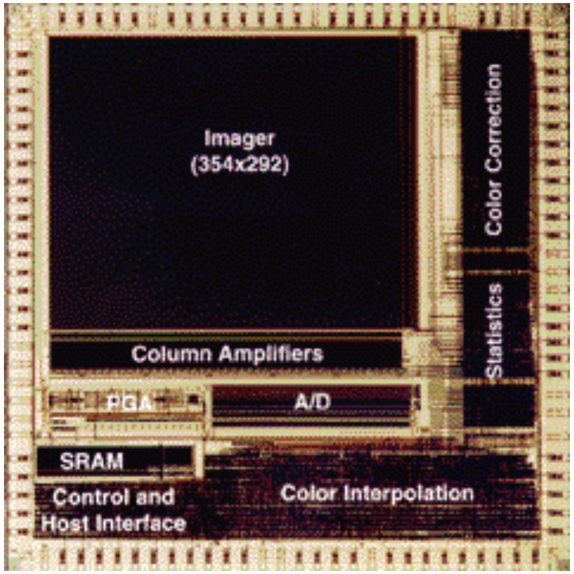
Computer Vision

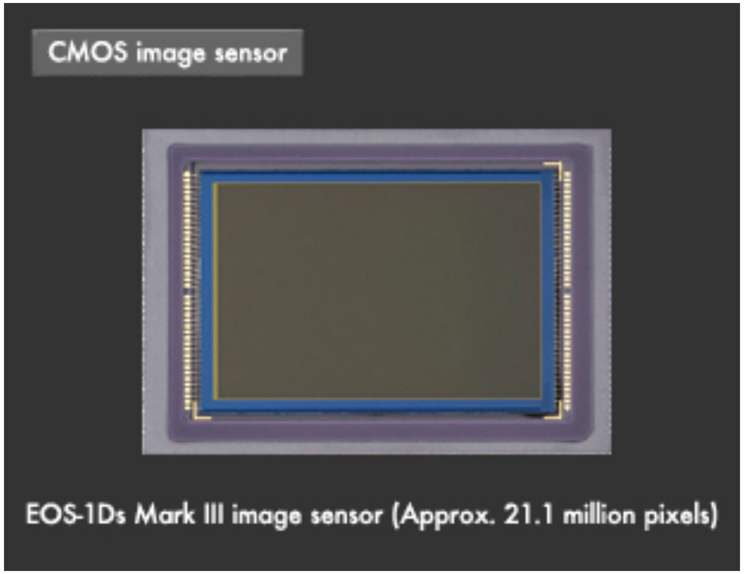
ACQUIS.

illumination
**cameras**

# CMOS

- Same sensor elements as CCD
- Each photo sensor has its own amplifier
  - More noise (reduced by subtracting 'black' image)
  - Lower sensitivity (lower fill rate)
- Uses standard CMOS technology
  - Allows to put other components on chip
  - 'Smart' pixels



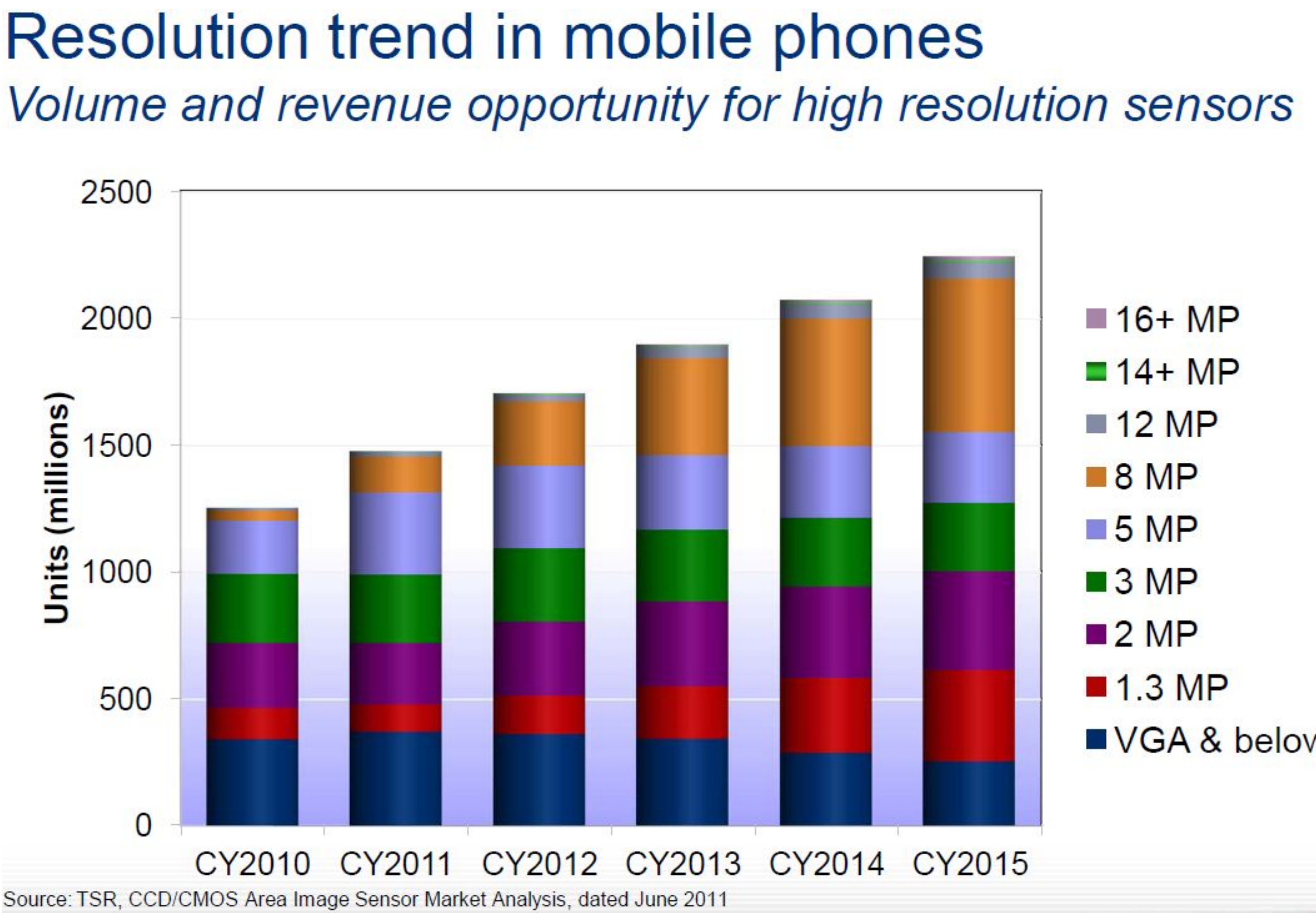



Computer Vision

ACQUIS.

illumination
**cameras**

# CMOS



# CCD vs. CMOS

- Niche applications
- Specific technology
- High production cost
- High power consumption
- Higher fill rate
- Blooming
- Sequential readout

- Consumer cameras
- Standard IC technology
- Cheap
- Low power
- Less sensitive
- Per pixel amplification
- Random pixel access
- Smart pixels
- On chip integration with other components

2006 was year of sales cross-over

Computer Vision

ACQUIS.

illumination
**cameras**

# CCD vs. CMOS

- Niche applications
- Specific technology
- High production cost
- High power consumption
- Higher fill rate
- Blooming
- Sequential readout

- Consumer cameras
- Standard IC technology
- Cheap
- Low power
- Less sensitive
- Per pixel amplification
- Random pixel access
- Smart pixels
- On chip integration with other components

In 2015 Sony said to stop CCD chip production

# Colour cameras

We consider 3 concepts:

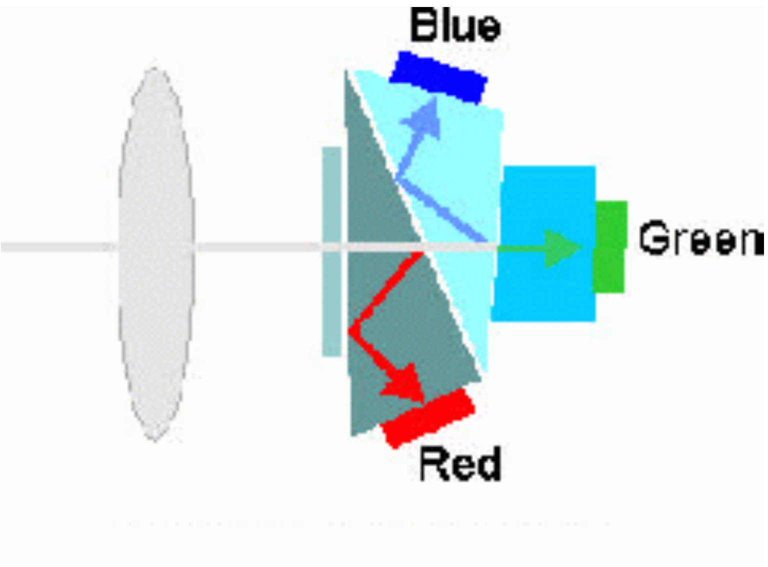
1. Prism (with 3 sensors)
2. Filter mosaic
3. Filter wheel

# Prism colour camera

Separate light in 3 beams using dichroic prism

Requires 3 sensors & precise alignment

Good color separation

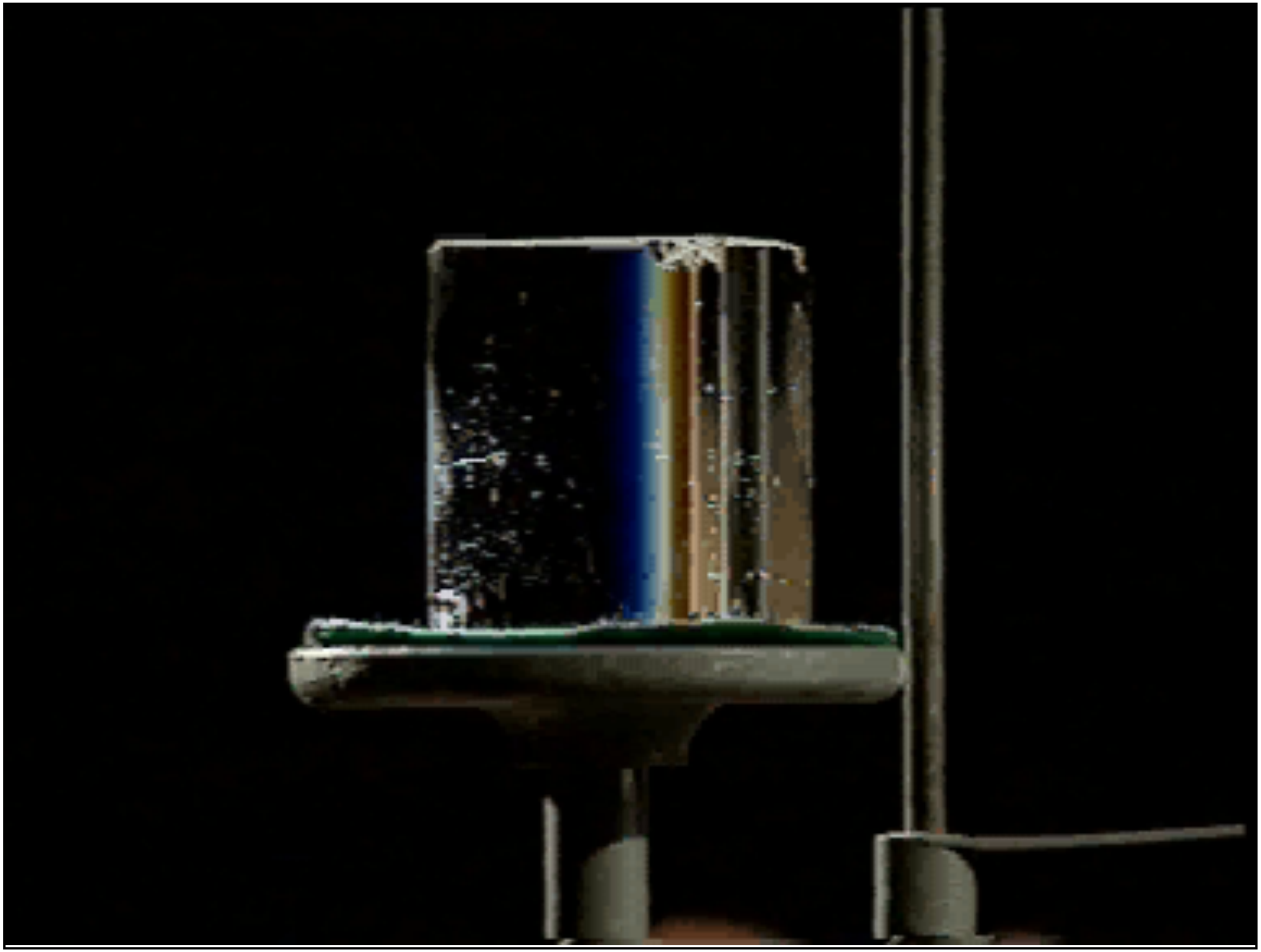


Computer Vision

ACQUIS.

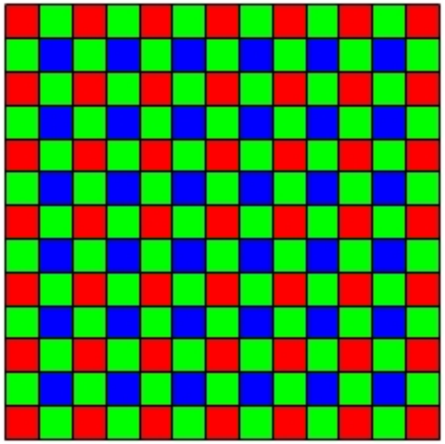
illumination
**cameras**

# Prism colour camera

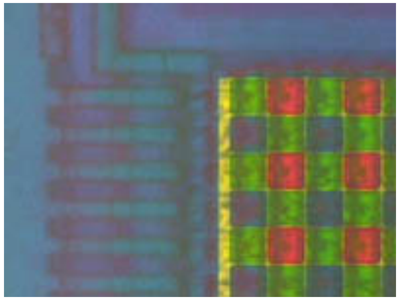


# Filter mosaic

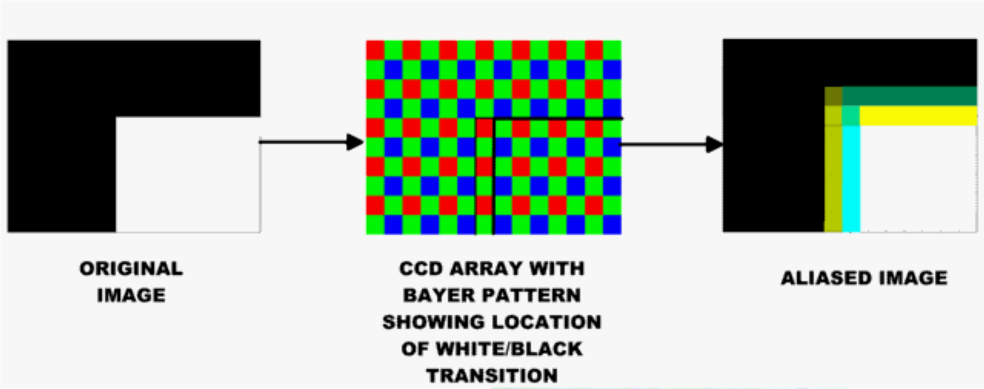
Coat filter directly on sensor



Bayer filter



Demosaicing (obtain full colour & full resolution image)

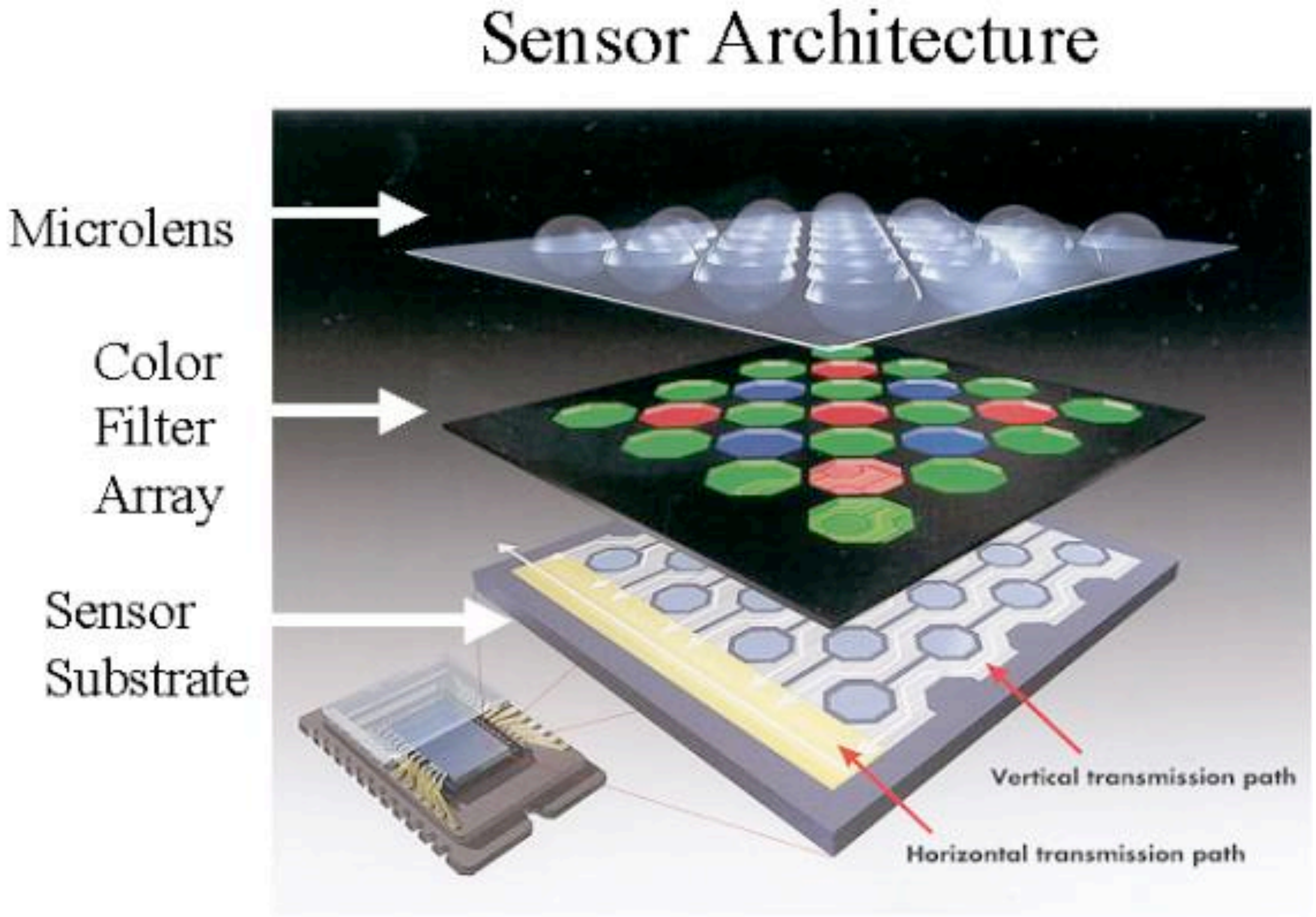


Computer Vision

ACQUIS.

illumination
**cameras**

# Filter mosaic



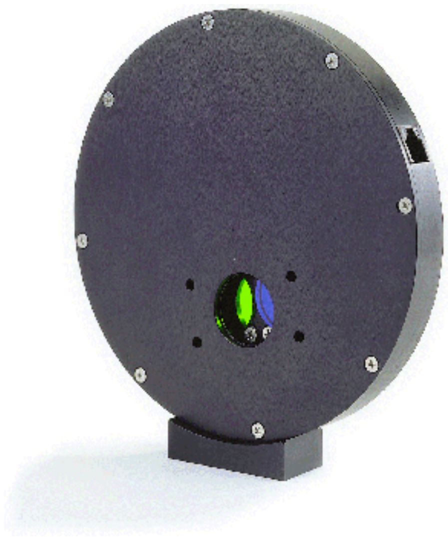

*Fuji Corporation*

Color filters lower the effective resolution, hence microlenses often added to gain more light on the small pixels

# Filter wheel

Rotate multiple filters in front of lens

Allows more than 3 colour bands



Only suitable for static scenes

# Prism vs. mosaic vs. wheel

| approach | Prism | Mosaic | Wheel |
|---|---|---|---|
| # sensors | 3 | 1 | 1 |
| Resolution | High | Average | Good |
| Cost | High | Low | Average |
| Framerate | High | High | Low |
| Artefacts | Low | Aliasing | Motion |
| Bands | 3 | 3 | 3 or more |
| | High-end cameras | Low-end cameras | Scientific applications |

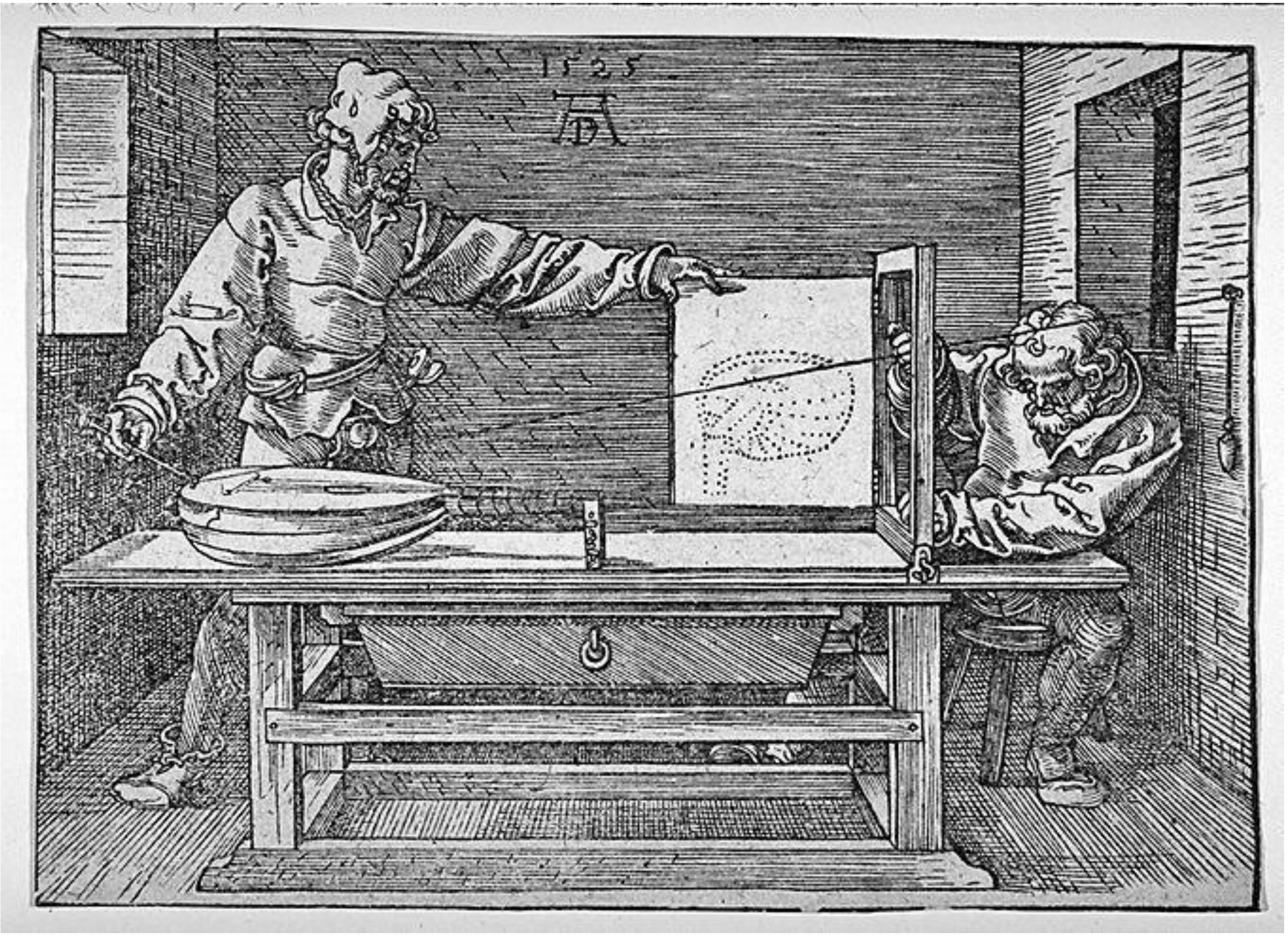
Computer Vision

ACQUIS.

illumination
**cameras**

# Geometric camera model

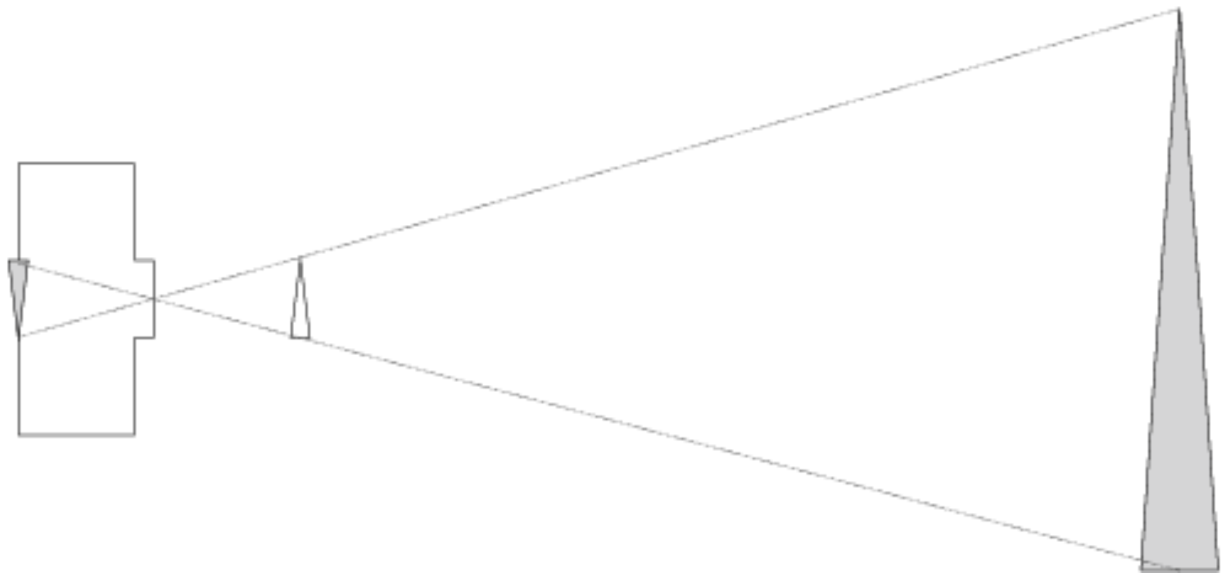
## perspective projection




(Man Drawing a Lute, woodcut, 1525, Albrecht Dürer)

# Models for camera projection

the pinhole model revisited :



center of the lens = center of projection

notice the virtual image plane

this is called *perspective* projection

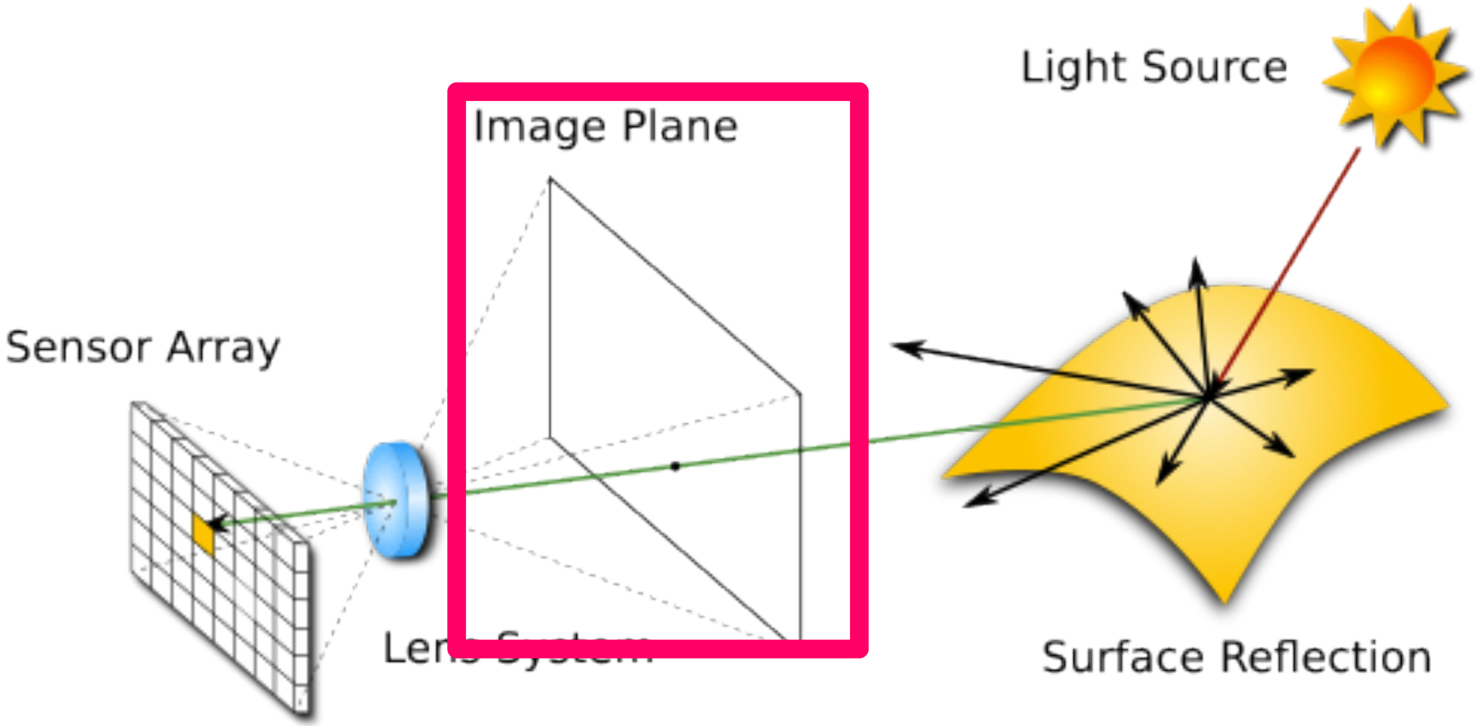
Computer Vision

ACQUIS.

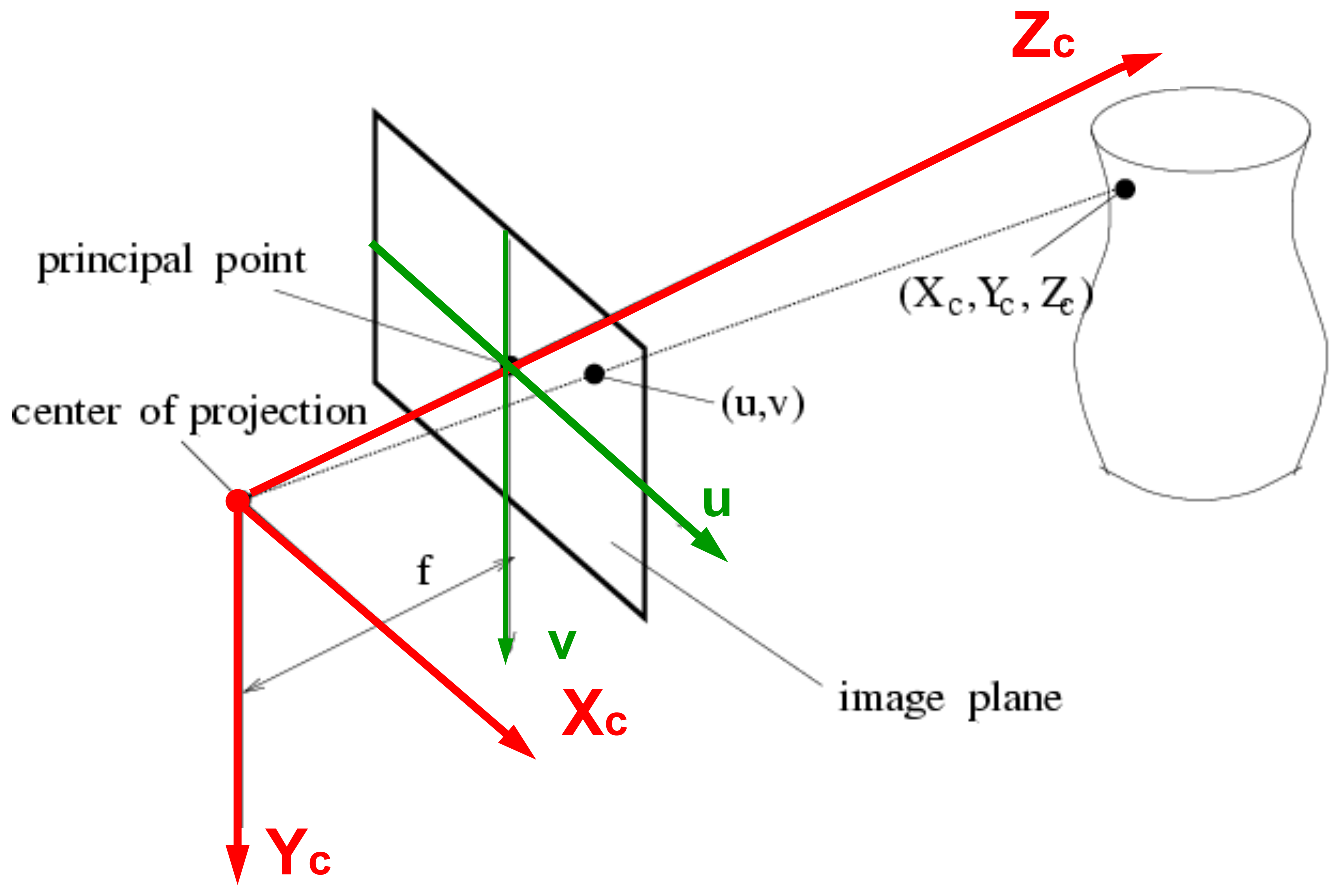
illumination
**cameras**

# Models for camera projection

We had the virtual plane also in the original reference sketch:



# Perspective projection



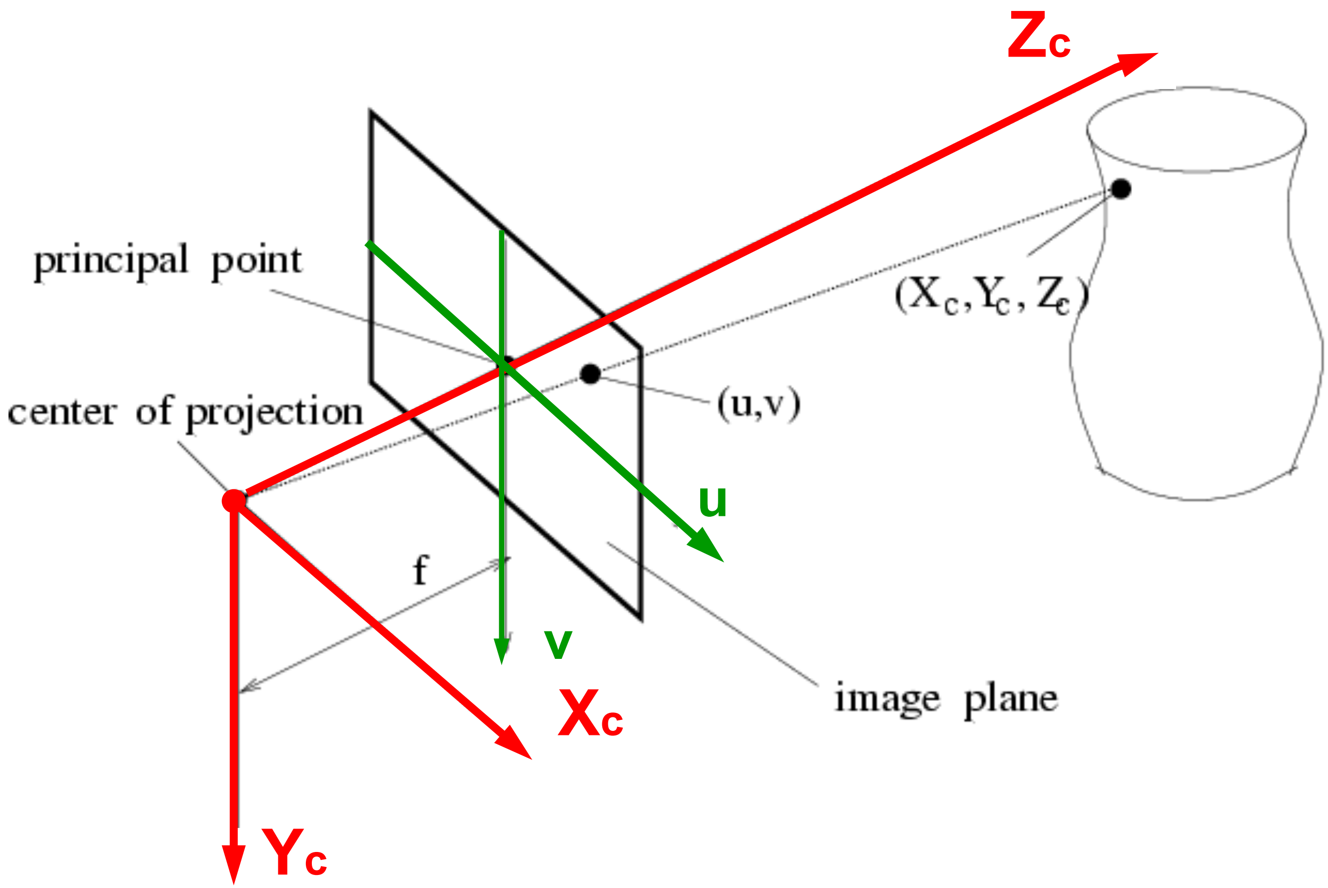

- origin lies at the center of projection
- the $Z_c$ axis coincides with the optical axis
- $X_c$-axis || to image rows, $Y_c$-axis || to columns

Computer Vision

ACQUIS.

illumination
**cameras**

# Perspective projection




$$u = f\frac{X}{Z} \qquad v = f\frac{Y}{Z}$$

# Pseudo-orthographic projection

$$u = f\frac{X}{Z} \qquad v = f\frac{Y}{Z}$$

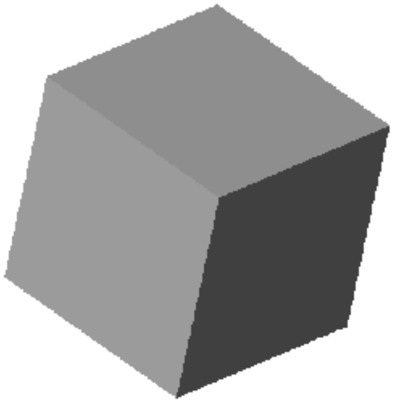
If $Z$ is constant $\Rightarrow x = kX$ and $y = kY$,
where $k = f/Z$

i.e. *orthographic* projection + a scaling

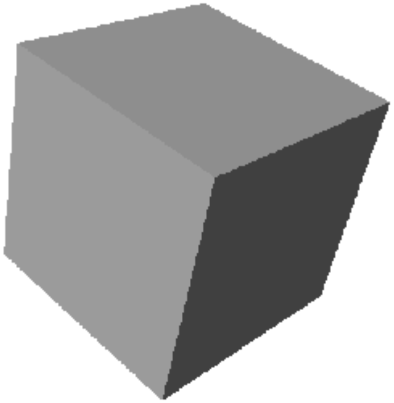
Good approximation if $f/Z \pm$ constant, i.e. if objects are small compared to their distance from the camera

# Pictoral comparison

**Pseudo - orthographic**



**Perspective**



# Projection matrices

the perspective projection model is incomplete : what if :

1. 3D coordinates are specified in a *world coordinate frame*

2. Image coordinates are expressed as *row and column numbers*

We will not consider additional refinements, such as radial distortions,...

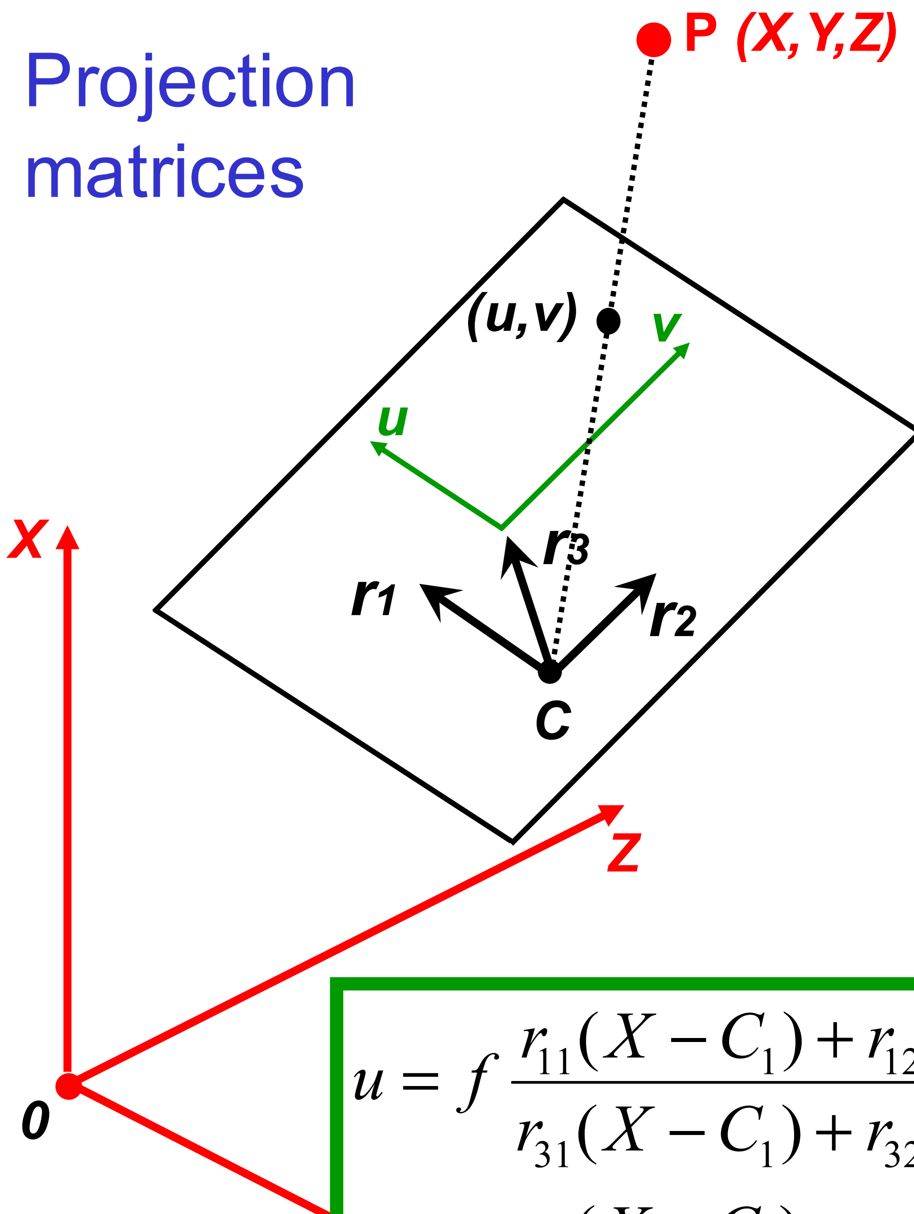
Computer Vision

ACQUIS.

illumination

**cameras**

# Projection matrices




$$u = f\,\frac{\langle r_1, P - C\rangle}{\langle r_3, P - C\rangle}$$

$$v = f\,\frac{\langle r_2, P - C\rangle}{\langle r_3, P - C\rangle}$$

$$u = f\,\frac{r_{11}(X - C_1) + r_{12}(Y - C_2) + r_{13}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)}$$

$$v = f\,\frac{r_{21}(X - C_1) + r_{22}(Y - C_2) + r_{23}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)}$$

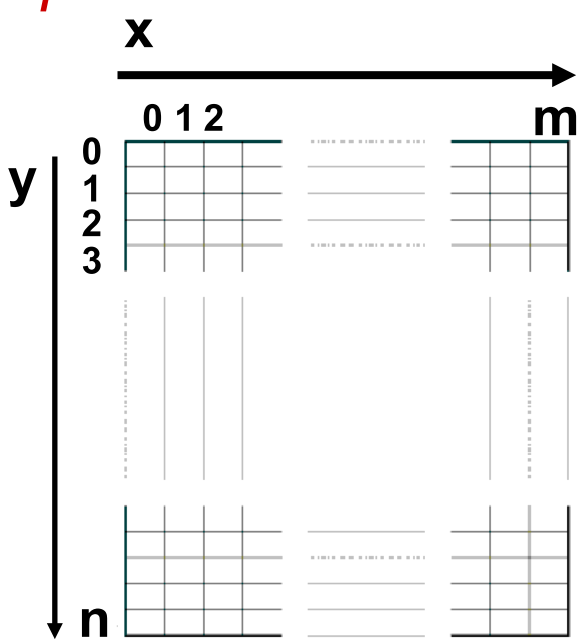
Computer Vision

ACQUIS.

illumination
**cameras**

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*




$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = k_y v + y_0 \end{cases}$$

with :

→ $(x_0, y_0)$ the pixel coordinates of the principal point

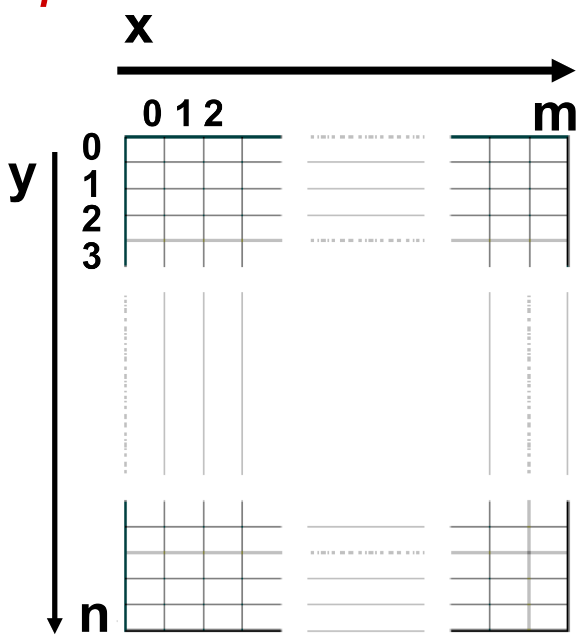
→ $k_x$ the number of pixels per unit length horizontally

→ $k_y$ the number of pixels per unit length vertically

→ $s$ indicates the skew ; typically $s = 0$

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*




$$\begin{cases} x = k_x u + s\, v + x_0 \\ y = \qquad\quad k_y v + y_0 \end{cases}$$

with :

**NB1:** often only integer pixel coordinates matter

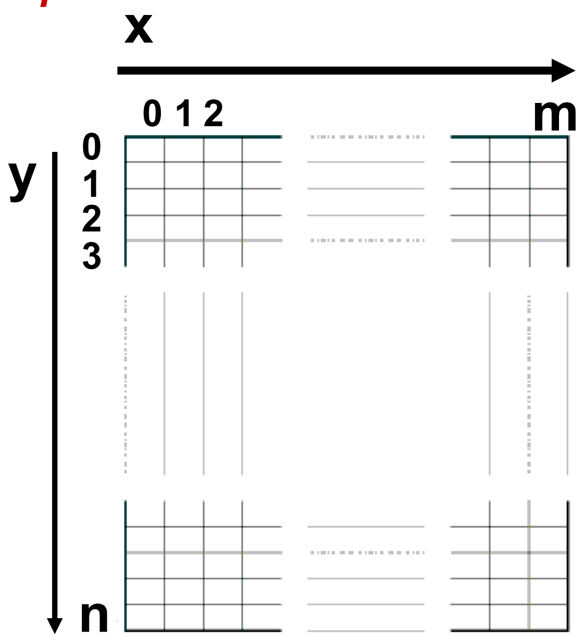
Computer Vision

ACQUIS.

illumination
**cameras**

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*




$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = k_y v + y_0 \end{cases}$$

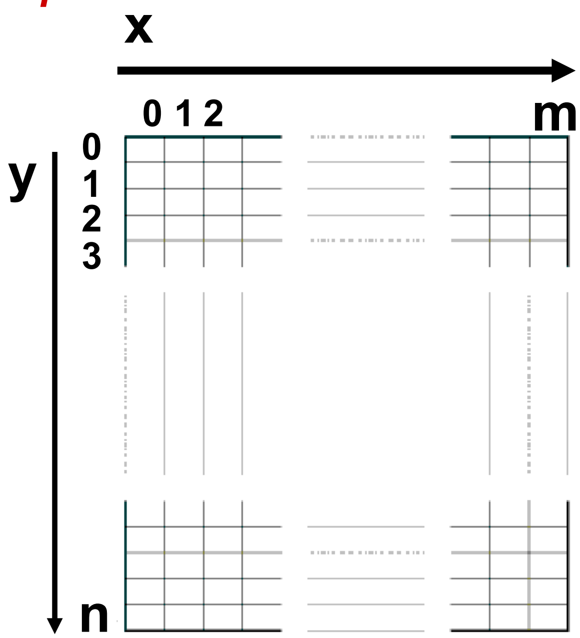
with :

**NB2 :** $k_y/k_x$ is called the *aspect ratio*

➔

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*



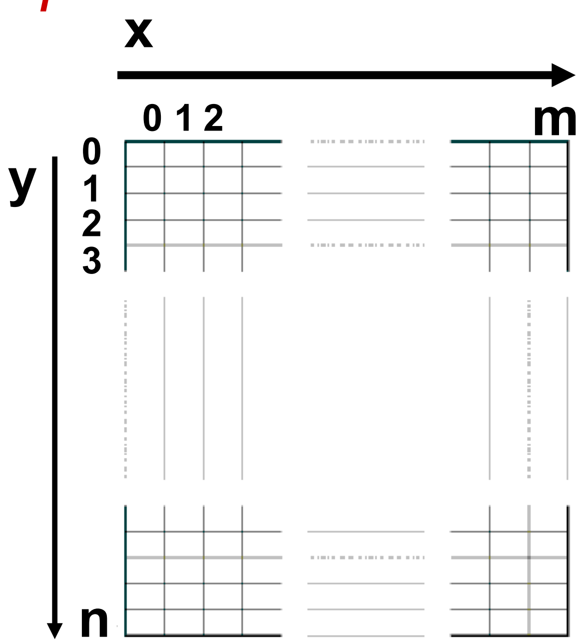

$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = \qquad k_y v + y_0 \end{cases}$$

with :

**NB3 :** $k_x, k_y, s, x_0$ and $y_0$ are called *internal camera parameters*

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*



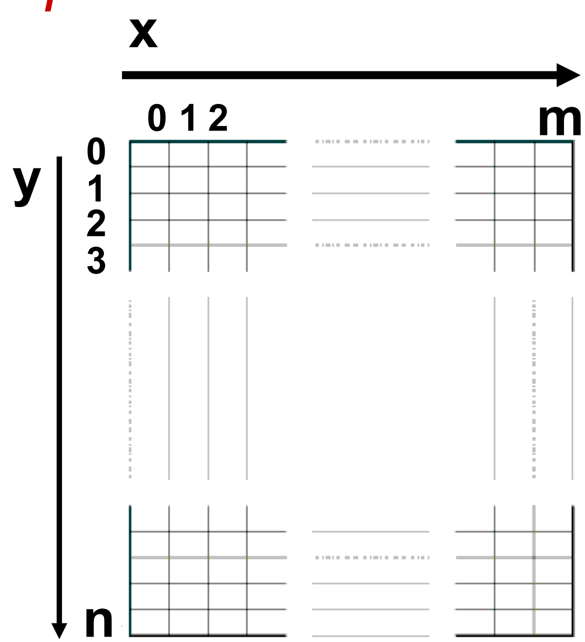

$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = \qquad\quad k_y v + y_0 \end{cases}$$

with :

**NB4 :** when they are known, the camera is *internally calibrated*

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*



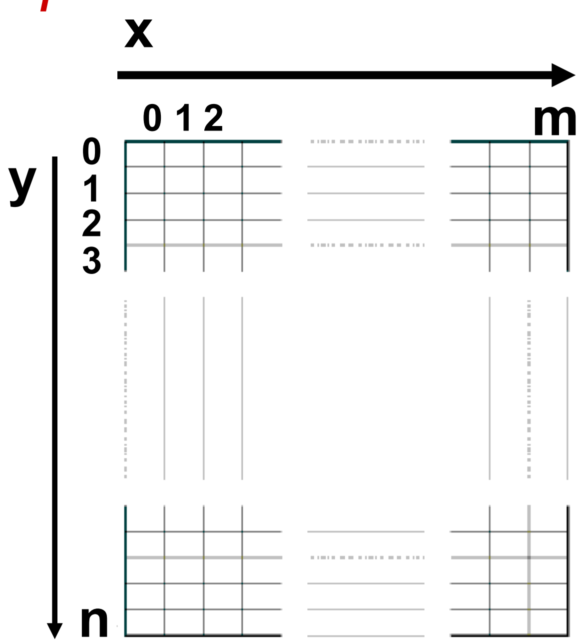

$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = \qquad\quad k_y v + y_0 \end{cases}$$

with :

**NB5 :** vector C and matrix R∈ SO (3) are the *ra* *external camera parameters*

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*




$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = k_y v + y_0 \end{cases}$$

with :

**NB6 :** when these are known, the camera is *ra*
*externally calibrated*

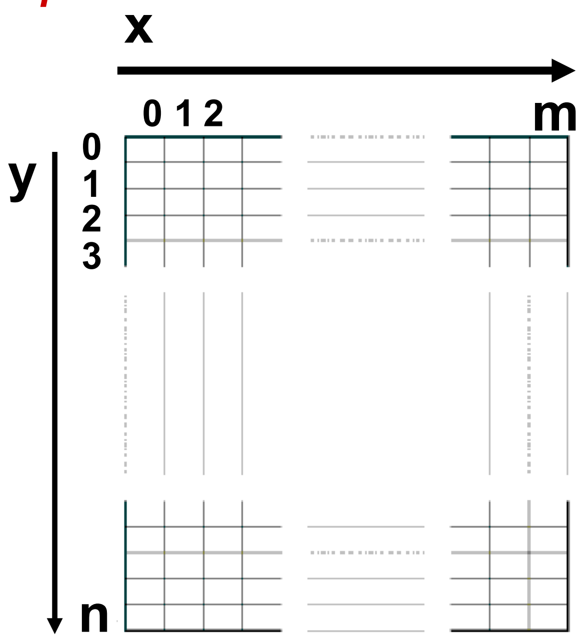
Computer Vision

ACQUIS.

illumination
**cameras**

# Projection matrices

Image coordinates are to be expressed as *pixel coordinates*




$$\begin{cases} x = k_x u + s\, v + x_0 \\ y = k_y v + y_0 \end{cases}$$

with :

**NB7 :** *fully calibrated* means internally and externally calibrated

# Homogeneous coordinates

Often used to linearize non-linear relations

2D
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x/z \\ y/z \end{pmatrix}$$

3D
$$\begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} \rightarrow \begin{pmatrix} X/W \\ Y/W \\ Z/W \end{pmatrix}$$

Homogeneous coordinates are only defined up to a factor

Computer Vision

ACQUIS.

illumination

**cameras**

# Projection matrices

$$u = f\,\frac{r_{11}(X-C_1)+r_{12}(Y-C_2)+r_{13}(Z-C_3)}{r_{31}(X-C_1)+r_{32}(Y-C_2)+r_{33}(Z-C_3)}$$

$$v = f\,\frac{r_{21}(X-C_1)+r_{22}(Y-C_2)+r_{23}(Z-C_3)}{r_{31}(X-C_1)+r_{32}(Y-C_2)+r_{33}(Z-C_3)}$$

Exploiting homogeneous coordinates :

$$\tau\begin{pmatrix}u\\ v\\ 1\end{pmatrix}=\begin{pmatrix}f\,r_{11} & f\,r_{12} & f\,r_{13}\\ f\,r_{21} & f\,r_{22} & f\,r_{23}\\ r_{31} & r_{32} & r_{33}\end{pmatrix}\begin{pmatrix}X-C_1\\ Y-C_2\\ Z-C_3\end{pmatrix}$$

Computer Vision

ACQUIS.

illumination
**cameras**

# Projection matrices

$$\begin{cases} x = k_x u + s\,v + x_0 \\ y = \qquad\qquad k_y v + y_0 \end{cases}$$

Exploiting homogeneous coordinates :

$$\tau \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \tau \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

Computer Vision

ACQUIS.

illumination
**cameras**

# Projection matrices

Thus, we have :

$$\tau\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f\,r_{11} & f\,r_{12} & f\,r_{13} \\ f\,r_{21} & f\,r_{22} & f\,r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}\begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}$$

$$\tau\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}\tau\begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

Computer Vision

ACQUIS.

illumination

**cameras**

# Projection matrices

Concatenating the results :

$$\tau\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} f\,r_{11} & f\,r_{12} & f\,r_{13} \\ f\,r_{21} & f\,r_{22} & f\,r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}\begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}$$

Or, equivalently :

$$\tau\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}\begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}$$

# Projection matrices

Re-combining matrices in the concatenation :

$$\tau\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}\begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}$$

yields the calibration matrix *K*:

$$K = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} f\,k_x & f\,s & x_0 \\ 0 & f\,k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Computer Vision

ACQUIS.

illumination

**cameras**

# Projection matrices

We define $p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}; \quad P = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$

yielding

$$\rho p = KR^t(P - C)$$ for some non-zero $\rho \in \mathbb{R}$

or, $$\rho p = K\left(R^t \mid -R^t C\right)\tilde{P}$$

or, $$\rho p = (M \mid t)\tilde{P}$$ with $\operatorname{rank} M = 3$

➔

# From object radiance to pixel grey levels

After the geometric camera model...

… a **photometric** camera model

2 steps:

1. from object radiance to image irradiance
2. from image irradiance to pixel grey level

# Image irradiance and object radiance

we look at the irradiance that an object patch will cause in the image

assumptions :
radiance *R* assumed known and
object at large distance compared to the focal length

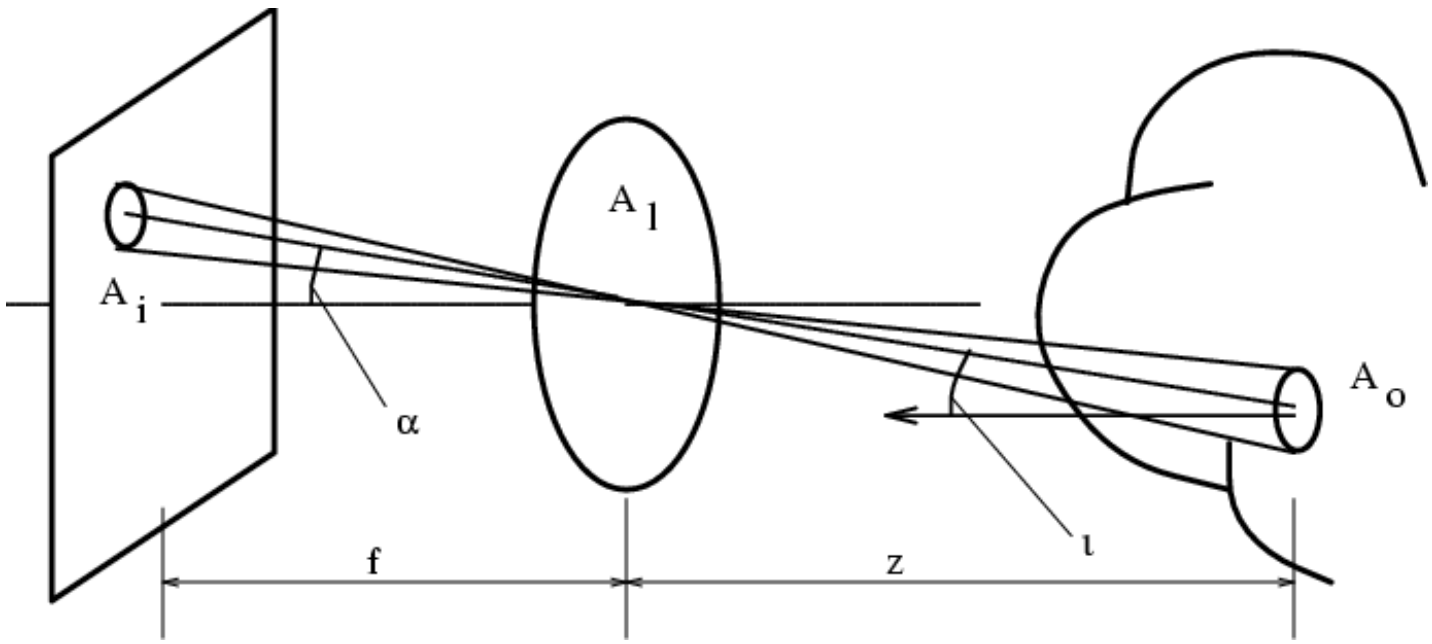
Is image irradiance directly related to the radiance of the image patch?

Computer Vision

ACQUIS.

illumination
**cameras**

# The viewing conditions




$$I = R\frac{A_l}{f^2}\cos^4\alpha$$

the cos$^4$ law

Computer Vision

ACQUIS.

illumination
**cameras**

# The $\cos^4$ law cont' d

Especially strong effects for wide-angle and fisheye lenses

# From irradiance to gray levels

$$f = g\, I^{\gamma} + d$$

Gain

"gamma"

Dark reference

Computer Vision

ACQUIS.

illumination
**cameras**

# From irradiance to gray levels

$$f = g\, I^{\gamma} + d$$

set w. size diaphragm  Gain

close to 1 nowadays  "gamma"

signal w. cam cap on  Dark reference