Image Processing Fundamentals

SCC0251/5830 – Image Processing

Prof. Moacir A. Ponti

Instituto de Ciências Matemáticas e de Computação – USP

2019/1

Agenda

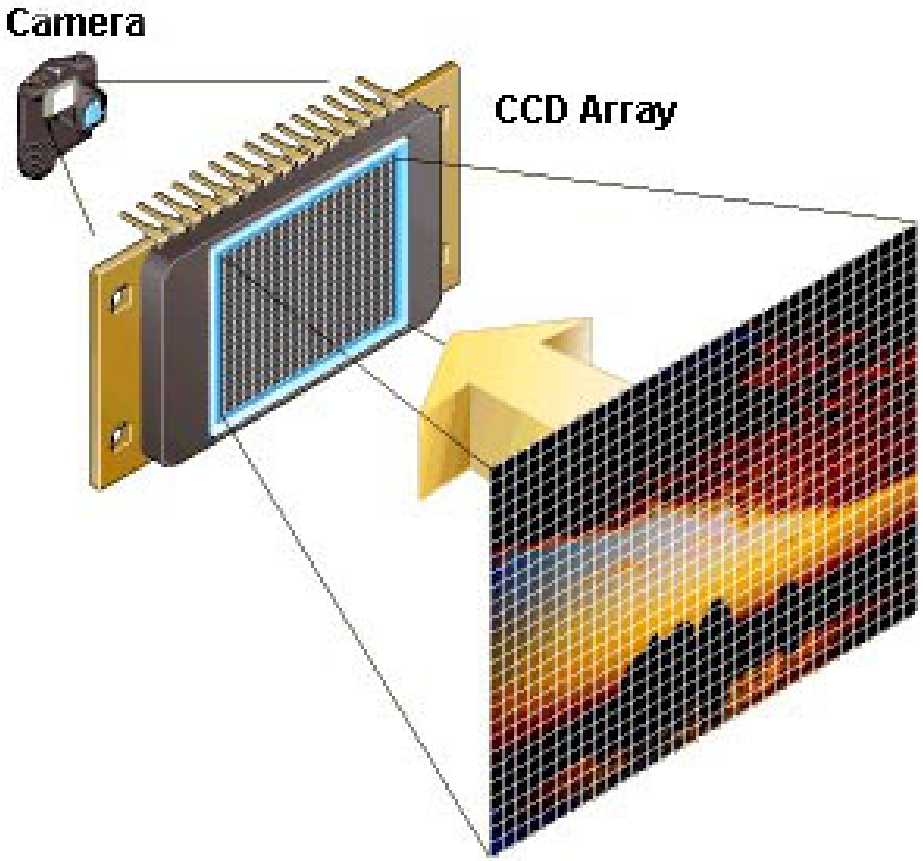
1. Image and Digital Image
2. Natural/biologic vision
3. Samping and Quantisation
4. History and typical image sources

Image

- Bidimensional function (2-d) of intensity of light $f(x, y)$:
  - x and y are spatial coordinates
  - f at (x, y) represents the intensity of light or color in the given coordinate
  - in practice, we define those functions as rectangular regions
- Continuous in space
- Coutinuous in amplitude

Acquisition

Acquisition



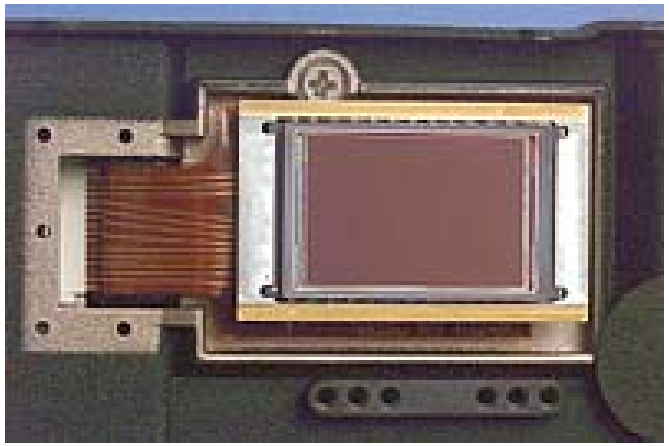
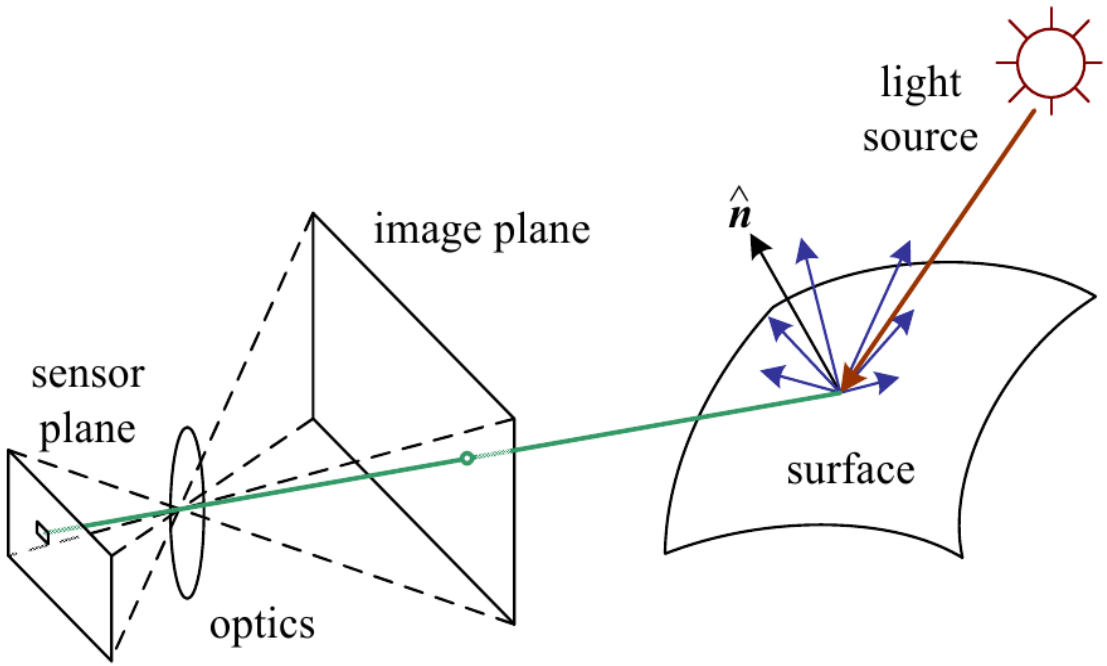
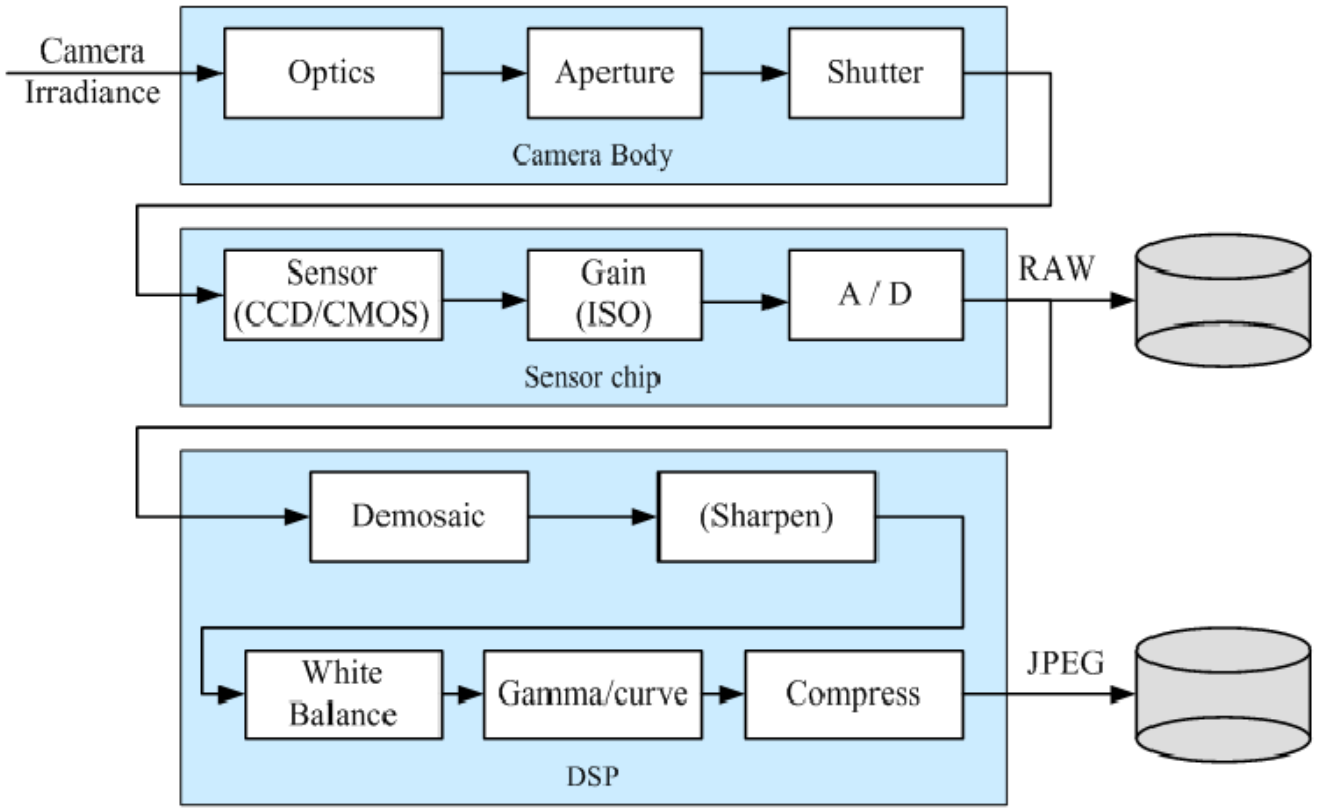


Image formation



Digital image generation pipeline



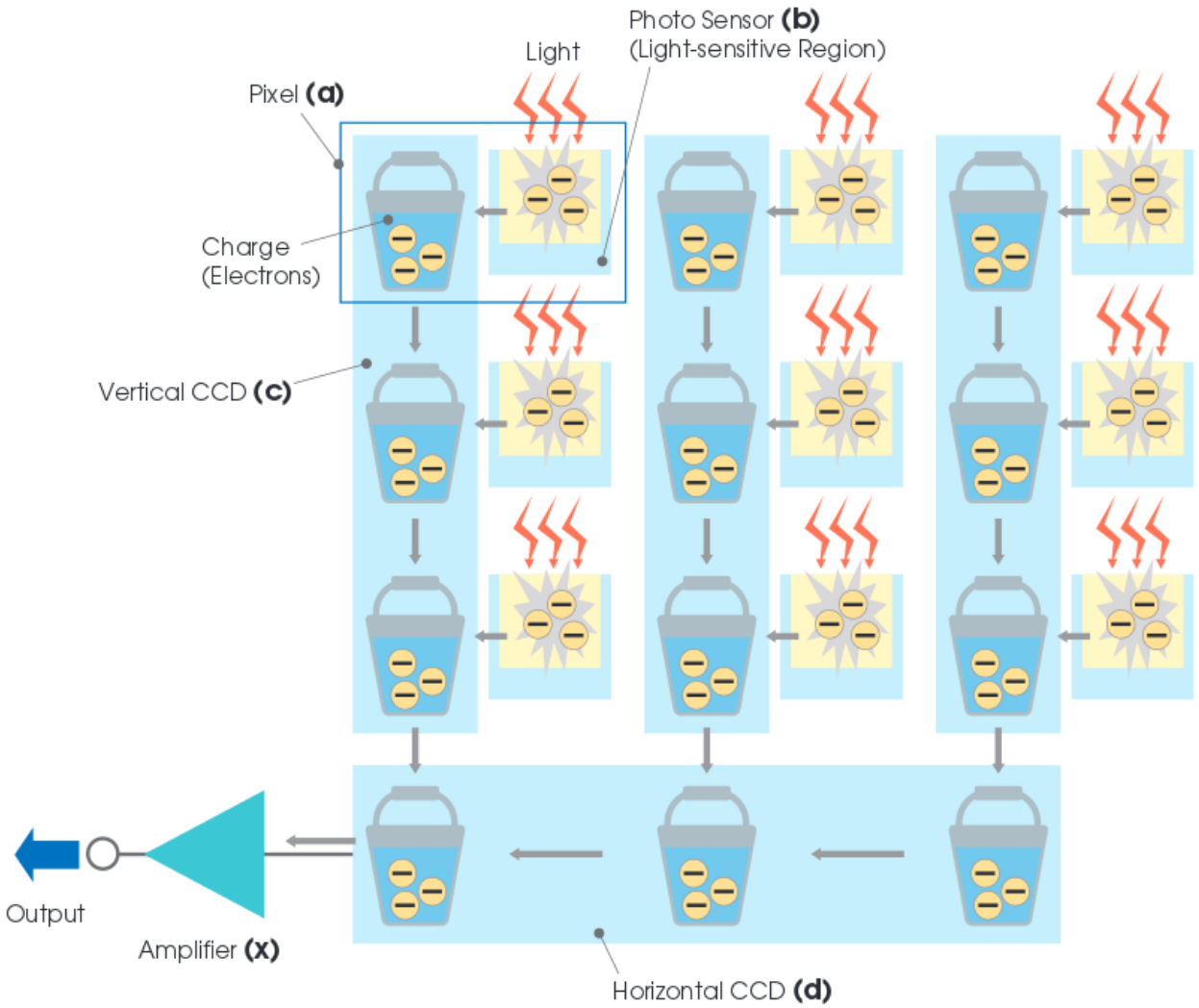
Digital image

- When acquiring images, the continuous function is **sampled**, its amplitude is **quantised**.
- As a result, a **digital image** is an image representation in a *2-d array of discrete samples*.
- Each element of the matrix is called **pixel**.

Digital Image: CCD sensor

- The light is captured via the lens system and integrated during the exposure time (usually expressed in fractions of a second),
- CCD: charge-coupled device,
  - each cell in the array is photo-sensitive, acting like a bucket
  - electrons are accumulated in each cell, in proportion with the amount of captured photons,
  - after exposure, the electrons run from cell to cell to an amplifier for charge-to-voltage conversion.

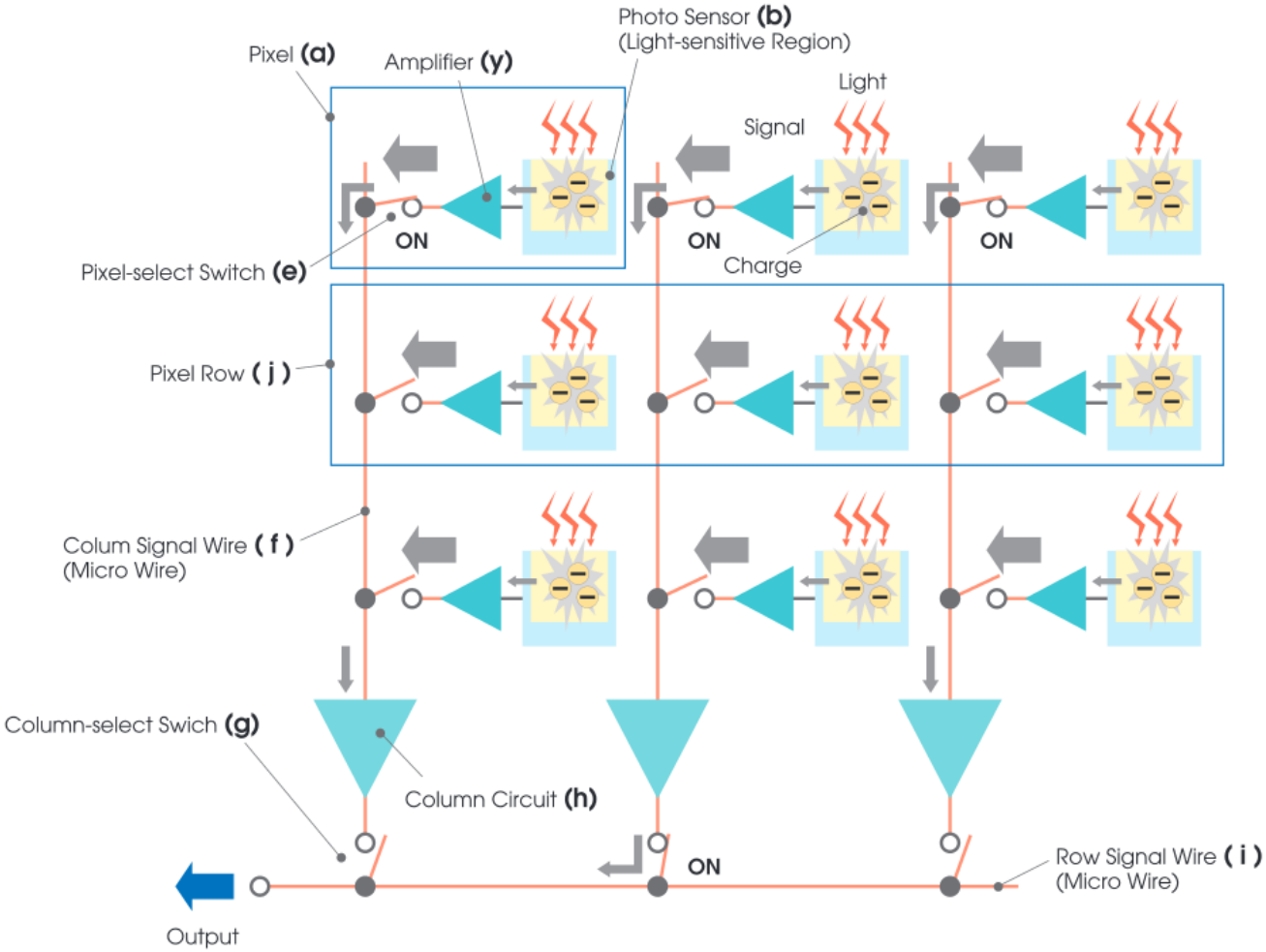
Digital Image: CCD sensor



Digital Image: CMOS

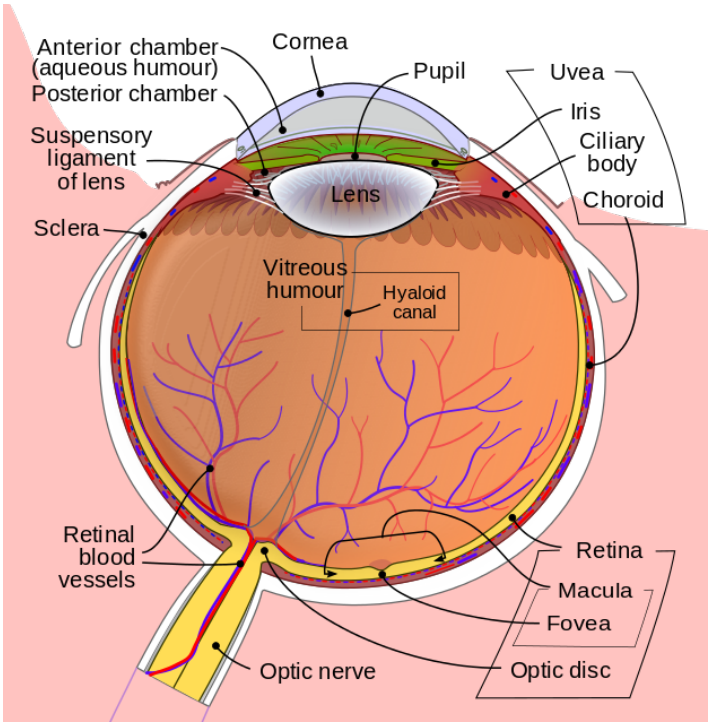
- CMOS: complementary metal oxide on silicon.
  - (*Active Pixel Sensor*): photons affect directly the conductivity of each cell, which is locally amplified,
  - is prone to more energy dispersion among nearby cells then CCDs,
  - transmission via microwires that already have the voltage.

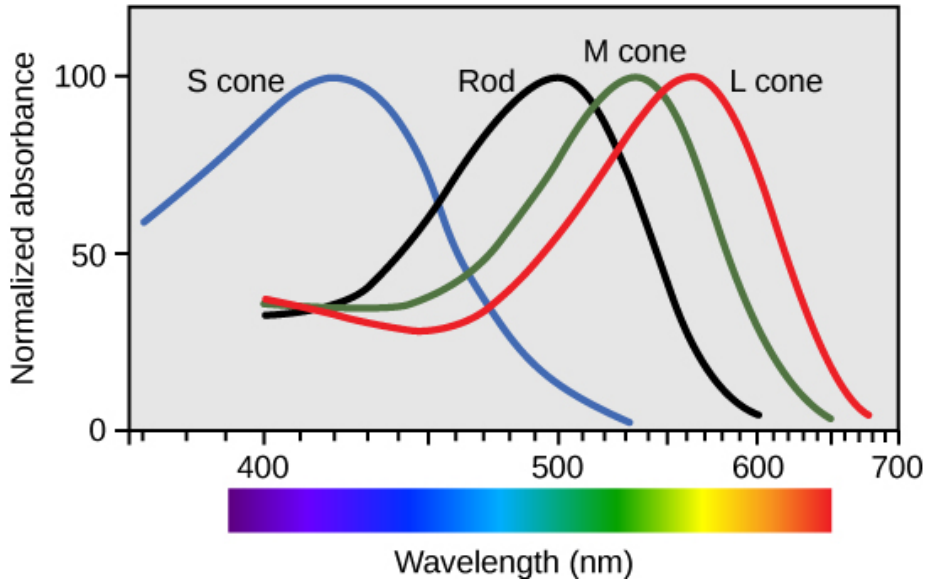
Digital Image: CMOS sensor



Natural/biologic vision

- Cornea, pupil, lens, retina and optical nerve play the most important roles in vision
  - Cornea, pupil and lens control the amount of light and how it refracts,
  - Retina/fovea contains the photo-sensitive cells: rods and cones,
  - Optic nerve is the neuron that carries the activations to the brain.





Sampling and resolution

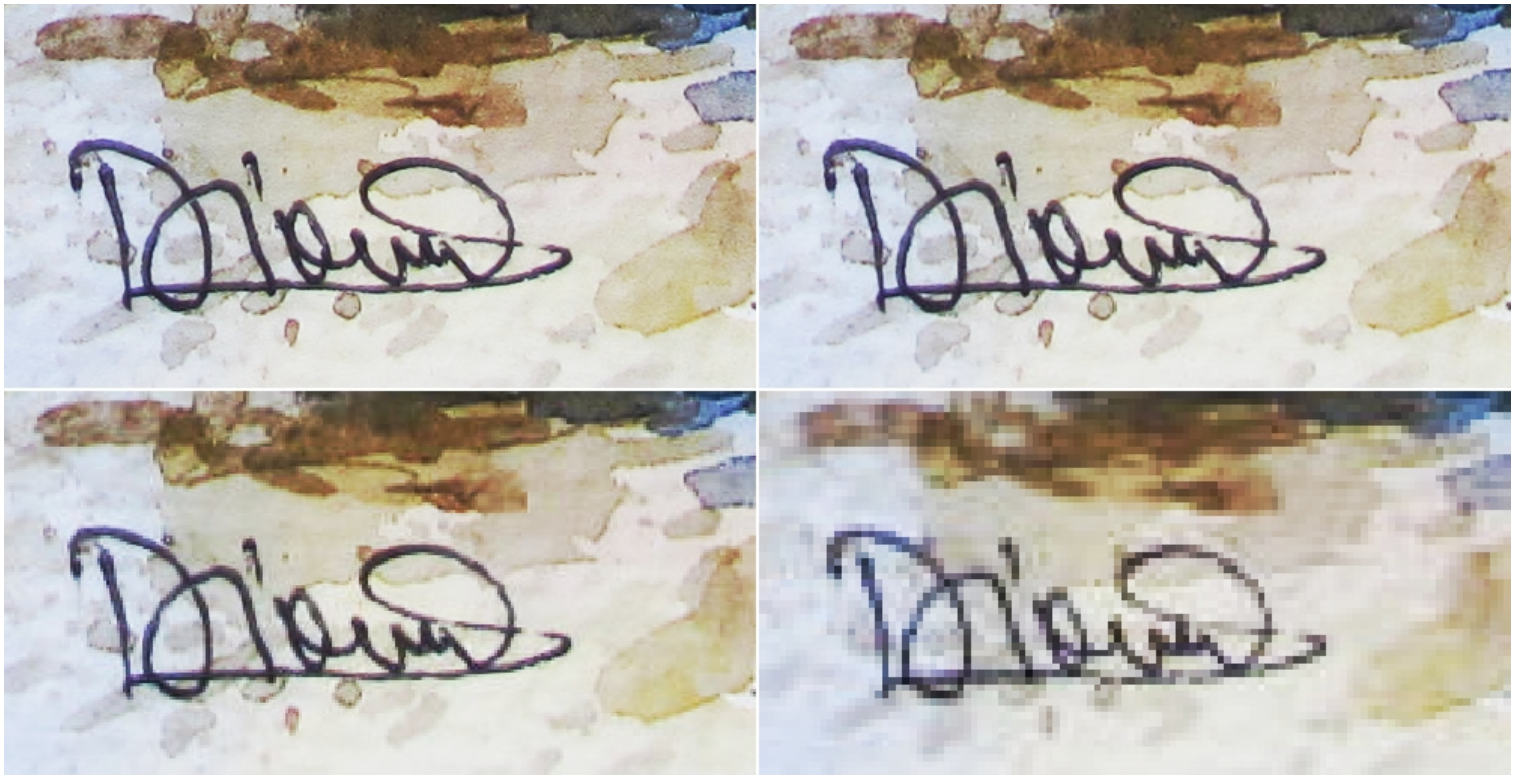
- The size of the sensor defines most of the image quality, followed by the analog gain (can be simulated via ISO) and the sensor noise.
- More megapixels not always mean more quality.

Typical sensor sizes:

	1/3"	1/2.7"	**1/2.5"**	**1/2"**	**1/1.8"**	1/1.7"	2/3"	1"
Width	4.8	5.37	5.76	6.4	7.18	7.6	8.8	12.8
Height	3.6	4.04	4.29	4.8	5.32	5.7	6.6	9.6
Size	17.3	21.7	24.7	30.7	38.2	43.3	58.1	123

Sampling and resolution

Images obtained using the same sensor but different sampling parameters:



Sampling and resolution

- Spatial resolution of human vision: how many points can be distinguished?
- Visual field of around 24000×24000 points, or 576 megapixels.
- However, our brains cannot see all in the same way, a smaller region is often processed at once.
- Displays are often adaptations of visual human vision, trying to provide something close to a continuous image.

Sampling and resolution

Displays

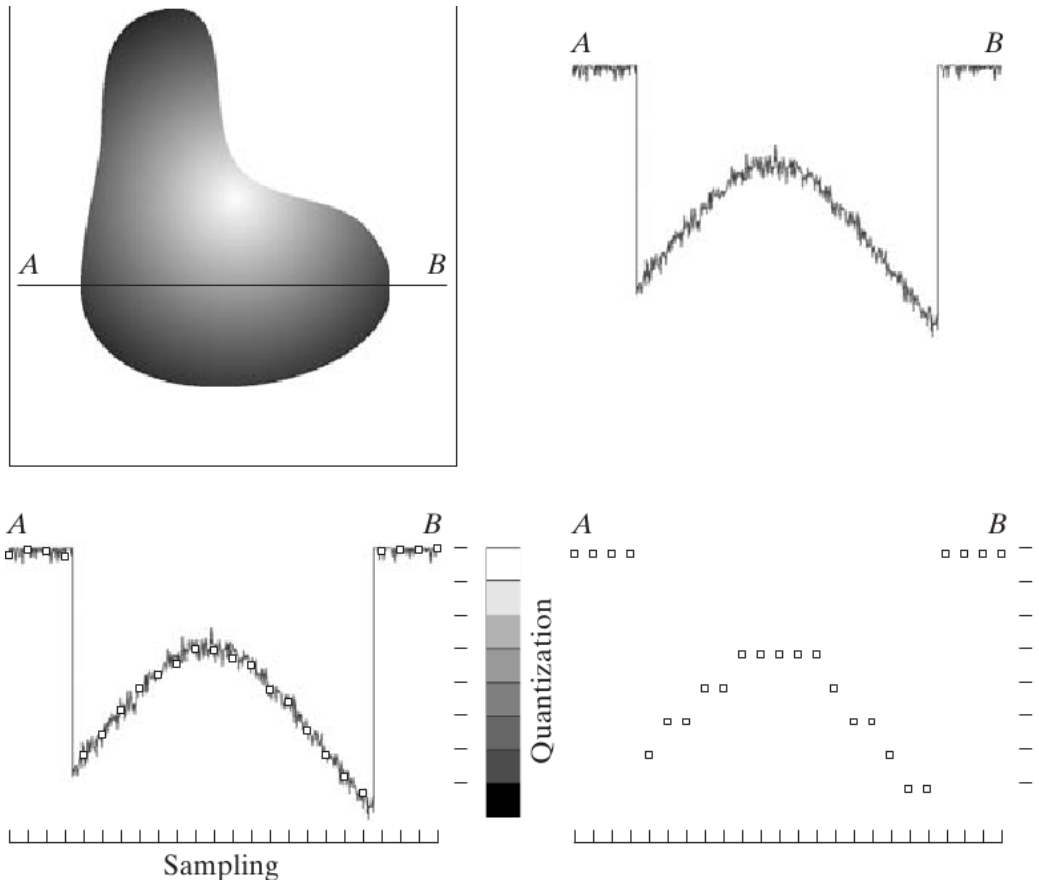
- SD: 512×480 (480 lines)
- HDTV: 1280×720 (720 lines)
- FullHD: 1920×1080 (1080 lines)
- 4K: 4000×2000 (2000 lines)

Aspect:

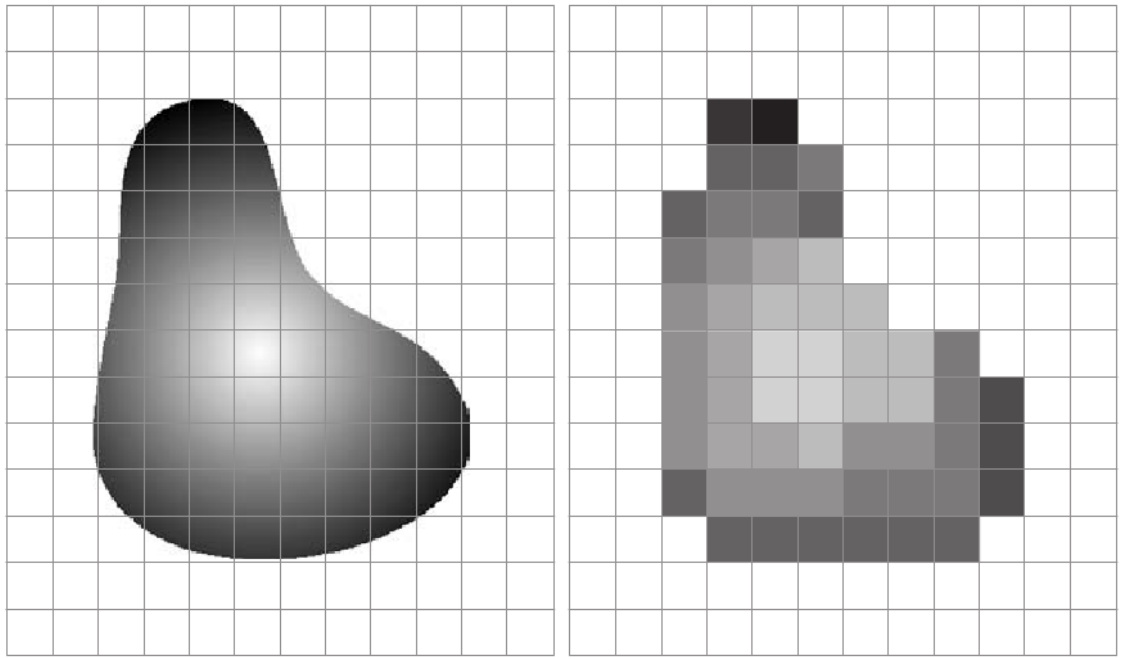
- 4:3
- 16:9 (widescreen)
- 21:9 (ultra widescreen)

Quantisation and number of colors/levels

- After sampling, each "real" observation must be converted into a discrete one, defined by the number of bits



Quantisation and number of colors/levels



Human perception on colour

How many colours can a human distinguish?

- 100K different colours (*disputable*)

24 bits

08 bits

04 bits

03 bits

Color components

Red (**R**)

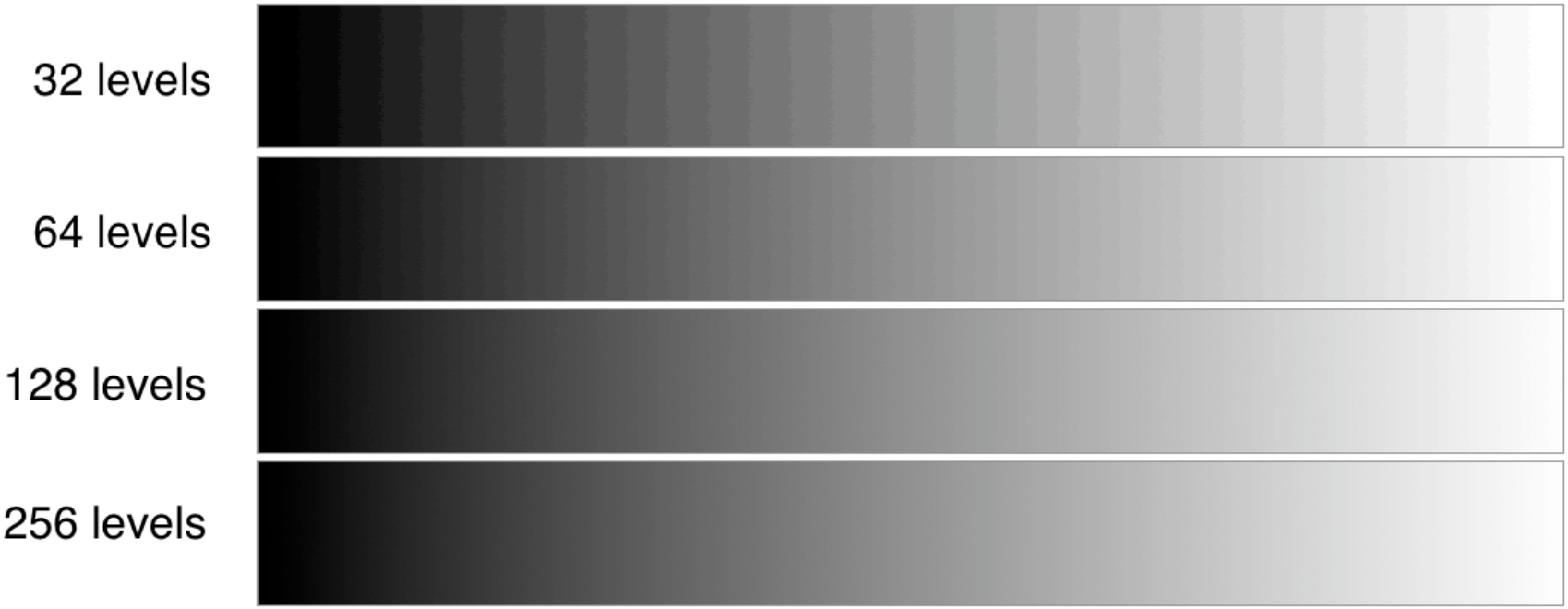
Green (**G**)

Blue (**B**)

24 bits (8 + 8 + 8)

Grey levels

Different parameters of quantisation, ordered can show false contours.

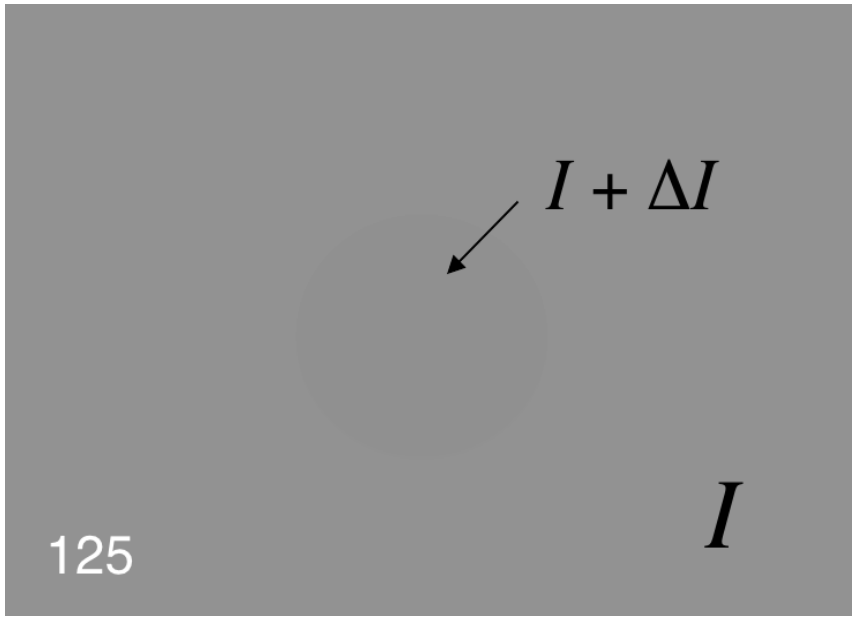


Quantisation

(Domício Pinheiro / Agência Estado)

Binary image (0-1)

Grey levels: bright discrimination



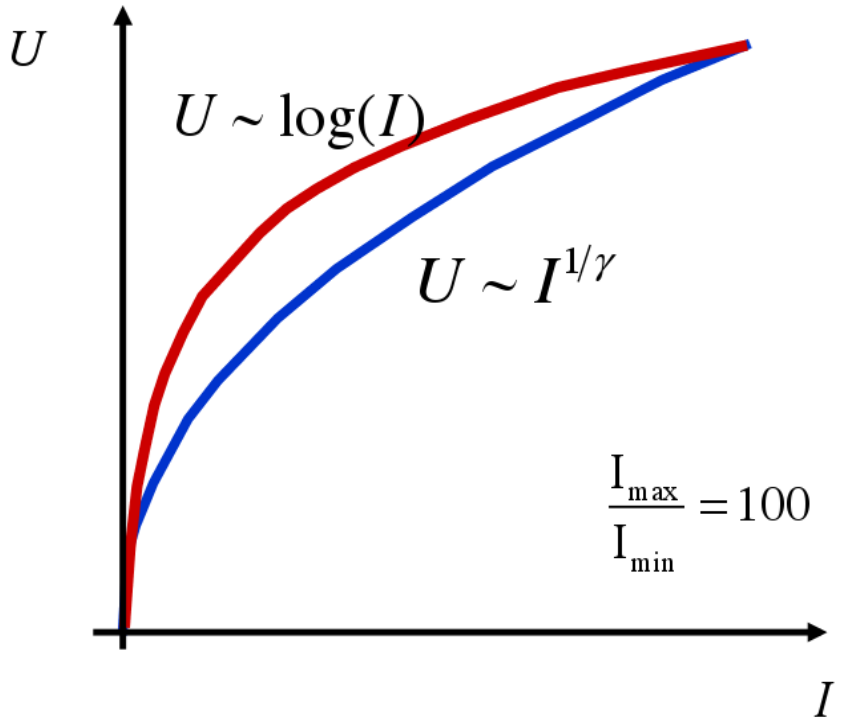
- A threshold was experimentally found by Weber:

$$\Delta I / I \approx K_{\text{Weber}} \approx 1..2\%,$$

law of Fechner-Weber. Human brightness perception is uniform in $\log(I)$

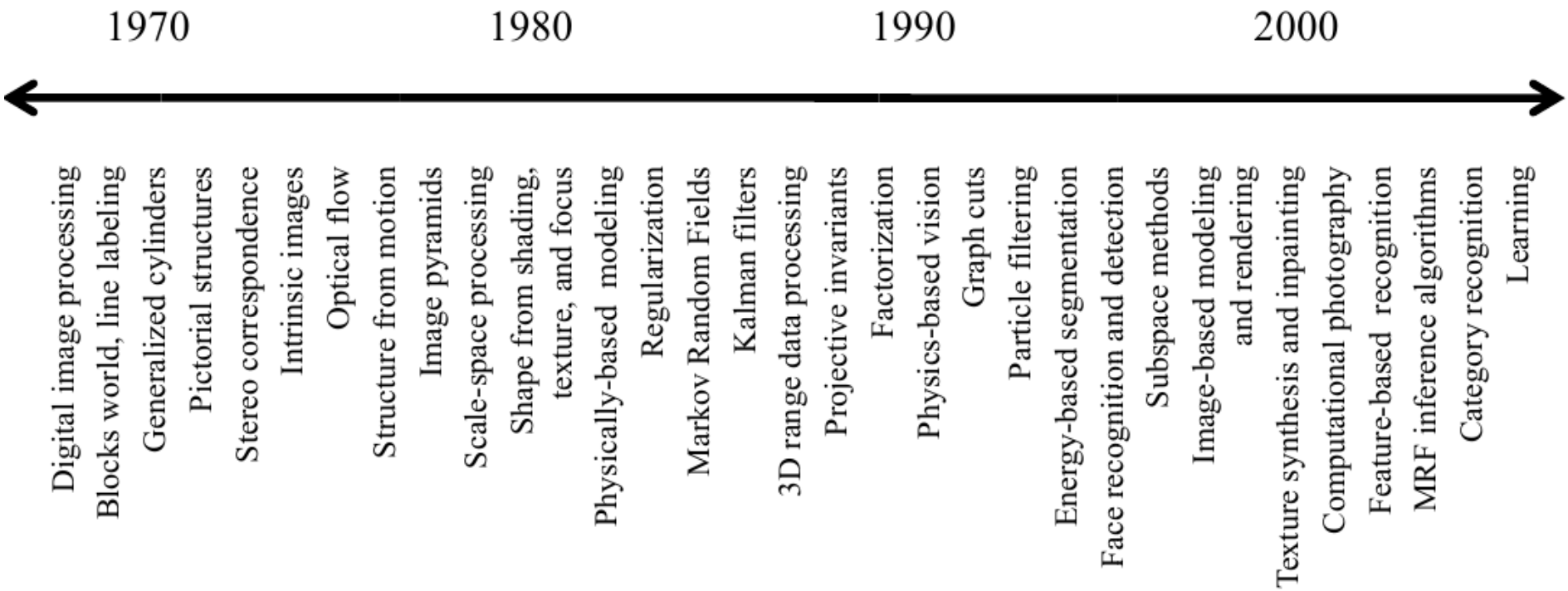
log vs γ-predistortion

displays/cameras have y-predistortion, while our perception is in $\log(I)$

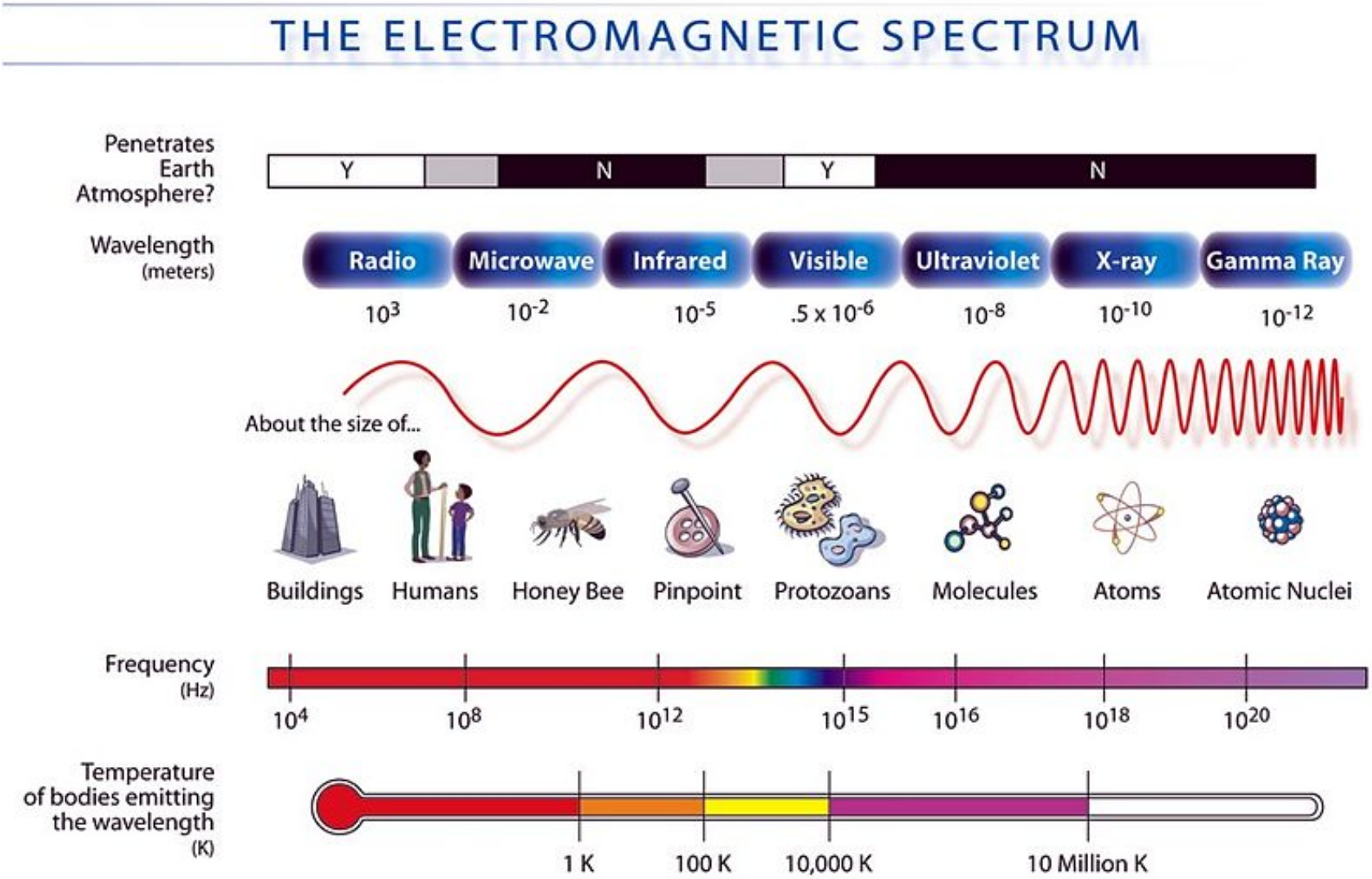


Agenda

History

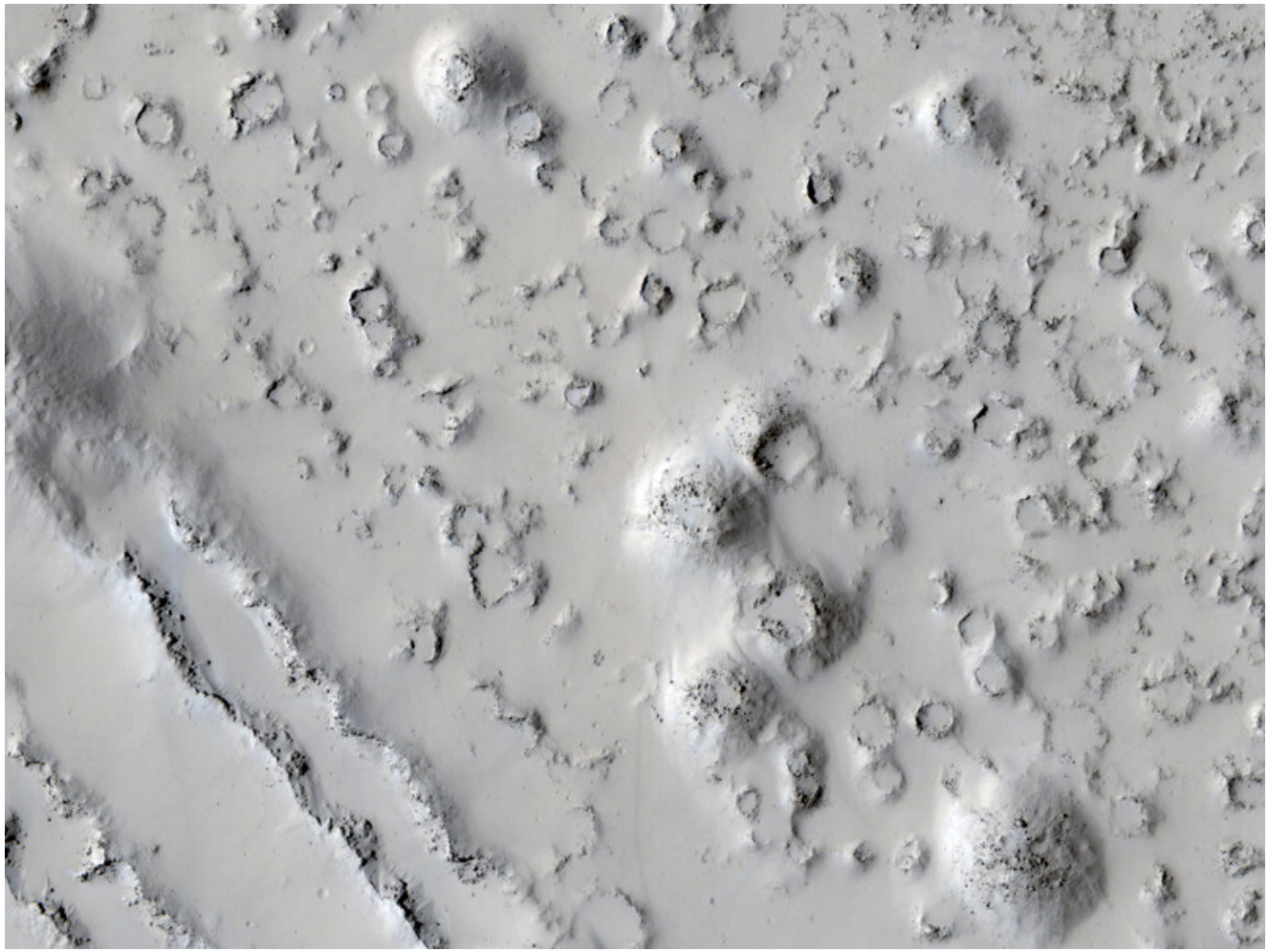

1970
1980
1990
2000
Digital image processing
Blocks world, line labeling
Generalized cylinders
Pictorial structures
Stereo correspondence
Intrinsic images
Optical flow
Structure from motion
Image pyramids
Scale-space processing
Shape from shading, texture, and focus
Physically-based modeling
Regularization
Markov Random Fields
Kalman filters
3D range data processing
Projective invariants
Factorization
Physics-based vision
Graph cuts
Particle filtering
Energy-based segmentation
Face recognition and detection
Subspace methods
Image-based modeling and rendering
Texture synthesis and inpainting
Computational photography
Feature-based recognition
MRF inference algorithms
Category recognition
Learning

Electromagnetic Spectrum

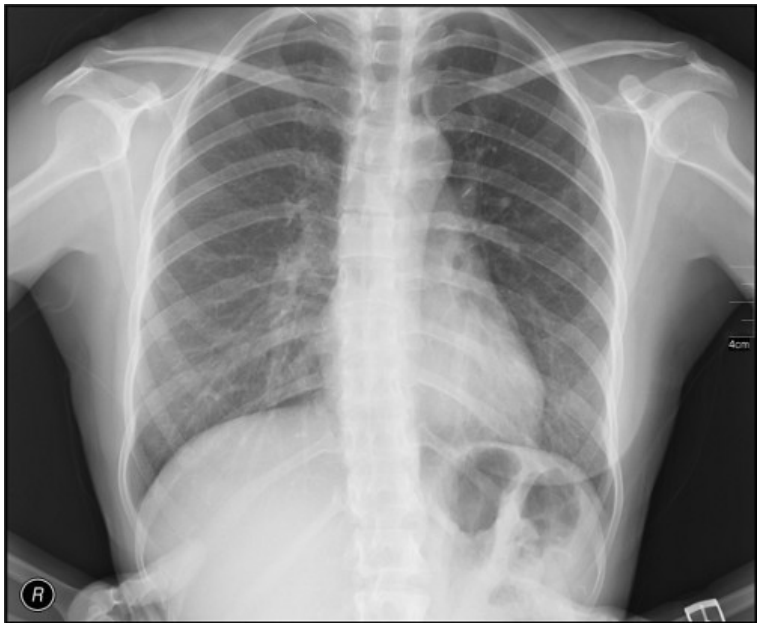


Natural images

Scientific images



Medical imaging



Remote sensing