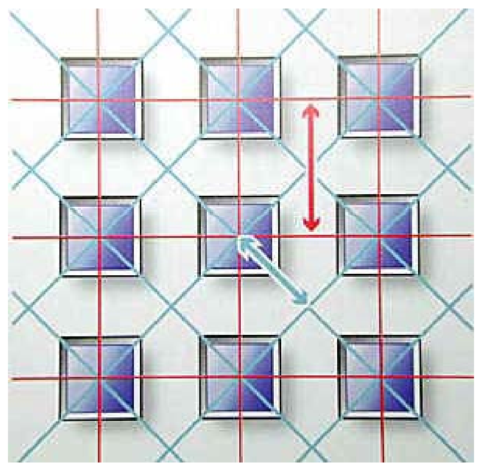
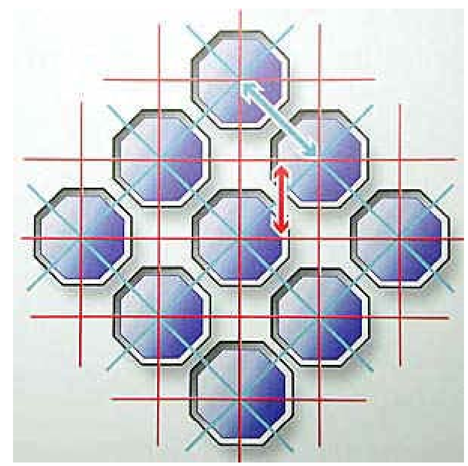
Camera sensors

SINA – 08/09

Camera Sensors

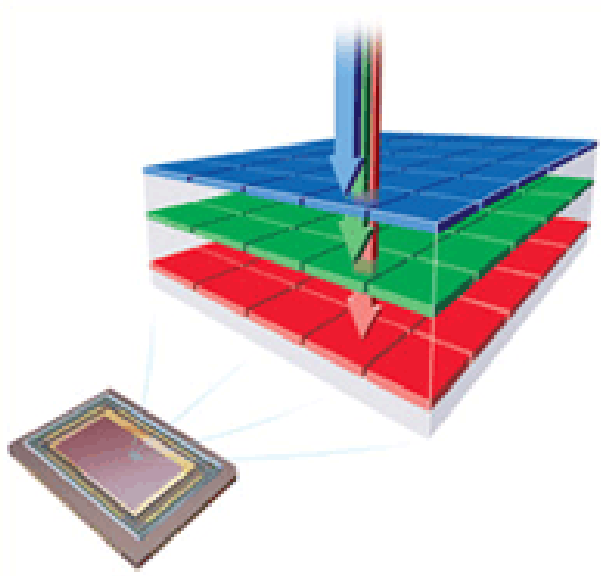
- A digital image is made up of tiny elements called *pixels*
- Photosites on the sensor capture the *brightness* of a single pixel
- The typical layout is a rectangular grid





Technologies

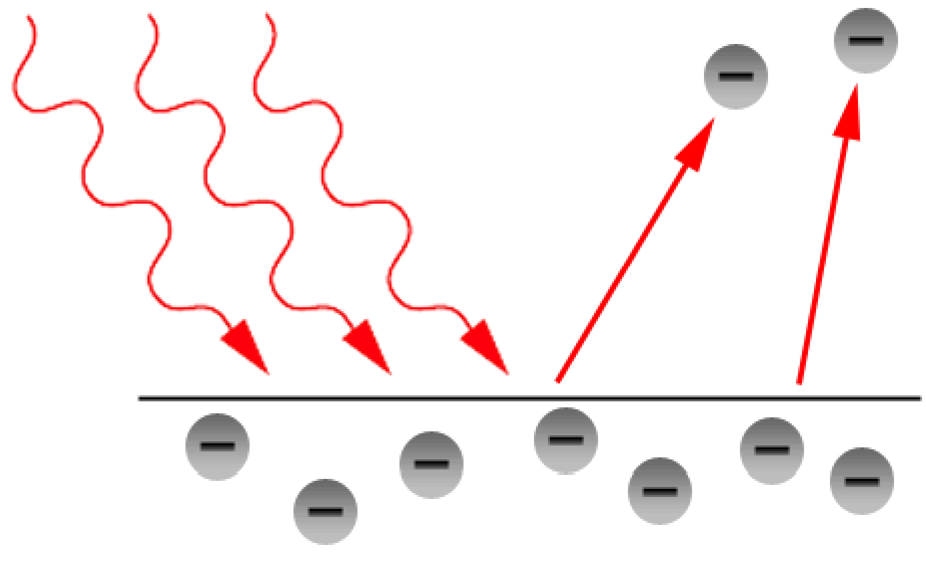
- CCD consists in photosensitive cells able to store charge produced by the light-to-electron conversion; in addition, the charge can be transferred to an interconnected, adjacent cell. In this case charges are shifted out of the sensor (bigger sensors, better quality, but additional circuitry)
- CMOS, transistors within the photosite perform charge-to-voltage conversion and allow the pixels to be read individually (higher integration, less power consumption, but less sensitive, higher noise)
- In both CMOS and CCD all photosites are sensitive to visible light, detect only brightness, not color
- Foveon: three layers of CMOS



CCD technology

- Analog shift register that enables charge to be transported through stages (capacitors) under the control of a clock signal
- BTW: and in fact was invented for totally different reasons (memory, delay lines)
- CCD refers to how the read out is performed
- Photons are converted into electrons by a special monocrystalline layer of silicon – photoelectric effect
- The photoactive region can be seen as an array of capacitors

Photoelectric effect



The effect on the semiconductor is to kick electrons from the valence to the conducting band (still inside the material)

Types of CCD

- Three architectures
  - Full frame: requires mechanical shutter, no further electronic circuitry
  - Frame transfer
  - Interline transfer
- Different approaches to the problem of shuttering
- Negative effect: smearing

CCD, Frame Transfer

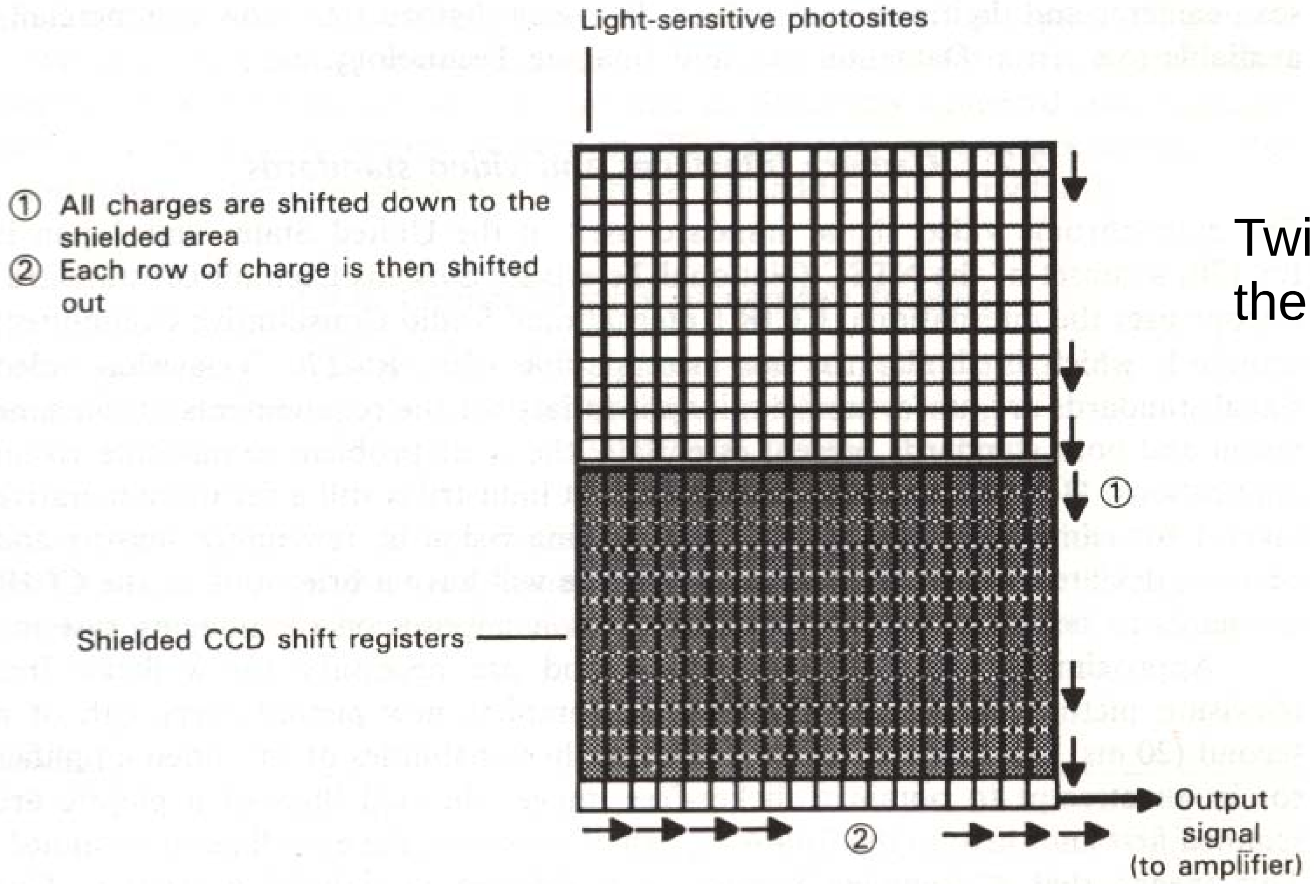


Figure 2.3 Frame transfer of charge in CCD sensors.

Twice the area of the sensitive part

Example: vertical smear

CCD, Interline Transfer

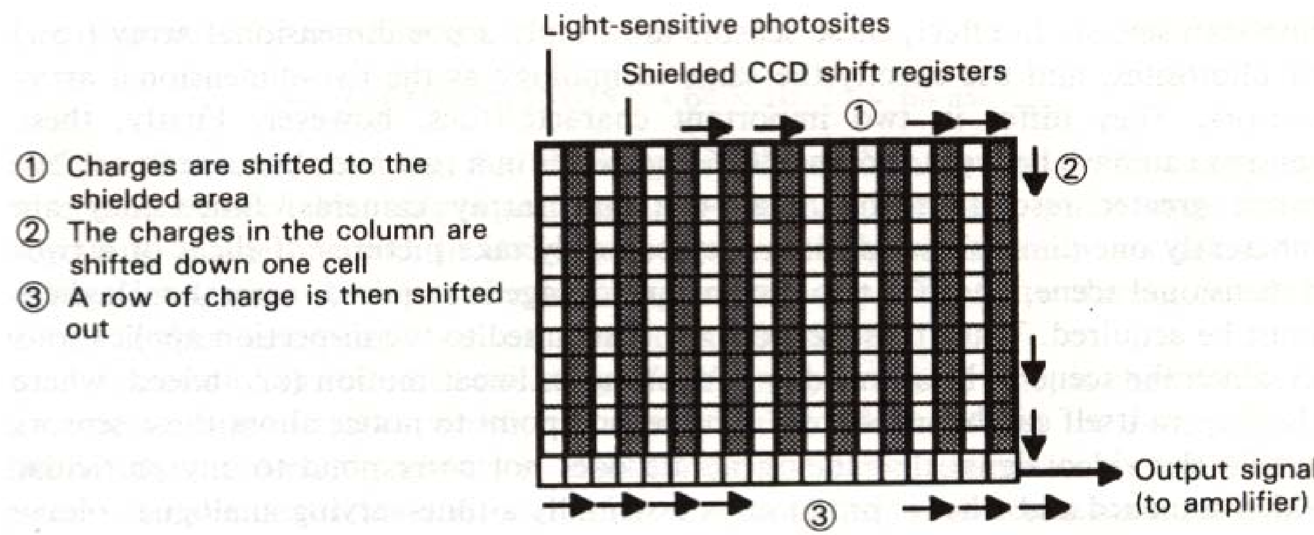


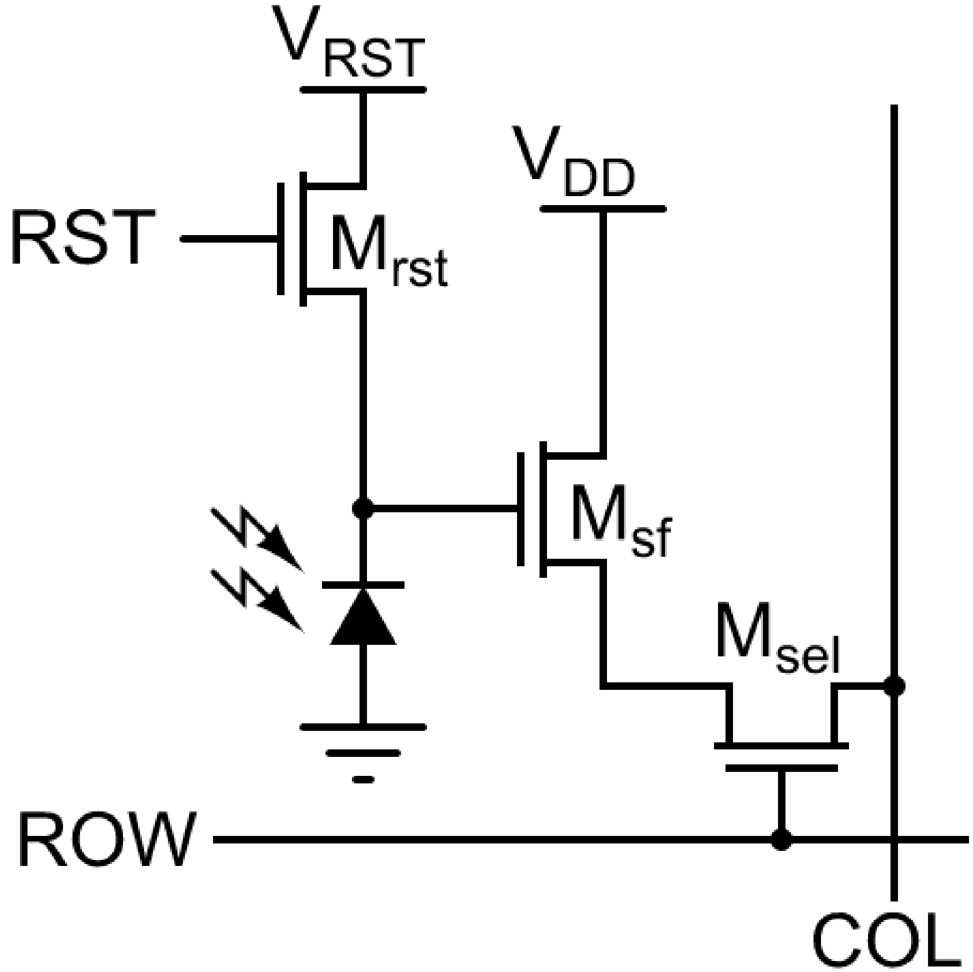
Figure 2.2 Interline transfer of charge in CCD sensors.

Even faster, 1 pixel transfer (< 1us), fill factor 50%
Modern design, uses microlenses to increase the fill factor (90%)

Common parameters

- Quantum efficiency 70% -> film only 2% (of course it depends on the film sensitivity, ISO): photons to electrons ratio (depends further on the channel/color)
- Fill factor: up to 90-100% depending on the arrangement (full frame preferred for example in astronomy, minimum smear)
- Smear: < few percentage points
- Dark current (no light): thermal, should be small

CMOS technology



Meaning

- Photodiode, the converter
- RST, reset signal, discharges the diode
- Msf is the read out transistor (source follower), reads the voltage, very small current, preserves charge
- Msel, selector (addressed from ROW)
- COL, output

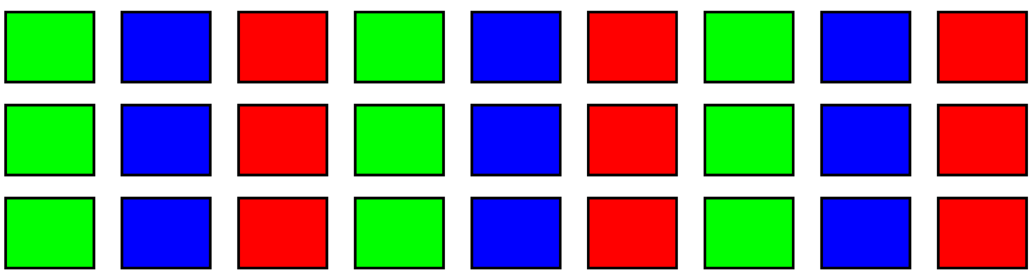
CMOS parameters

- Smallest number of transistors
- Greater active area

- Tradeoff: reset mechanism vs. image lag
  - Better reset means more complicated electronics (affecting the fill factor), better reset means lower noise
  - Better reset takes time, larger lags

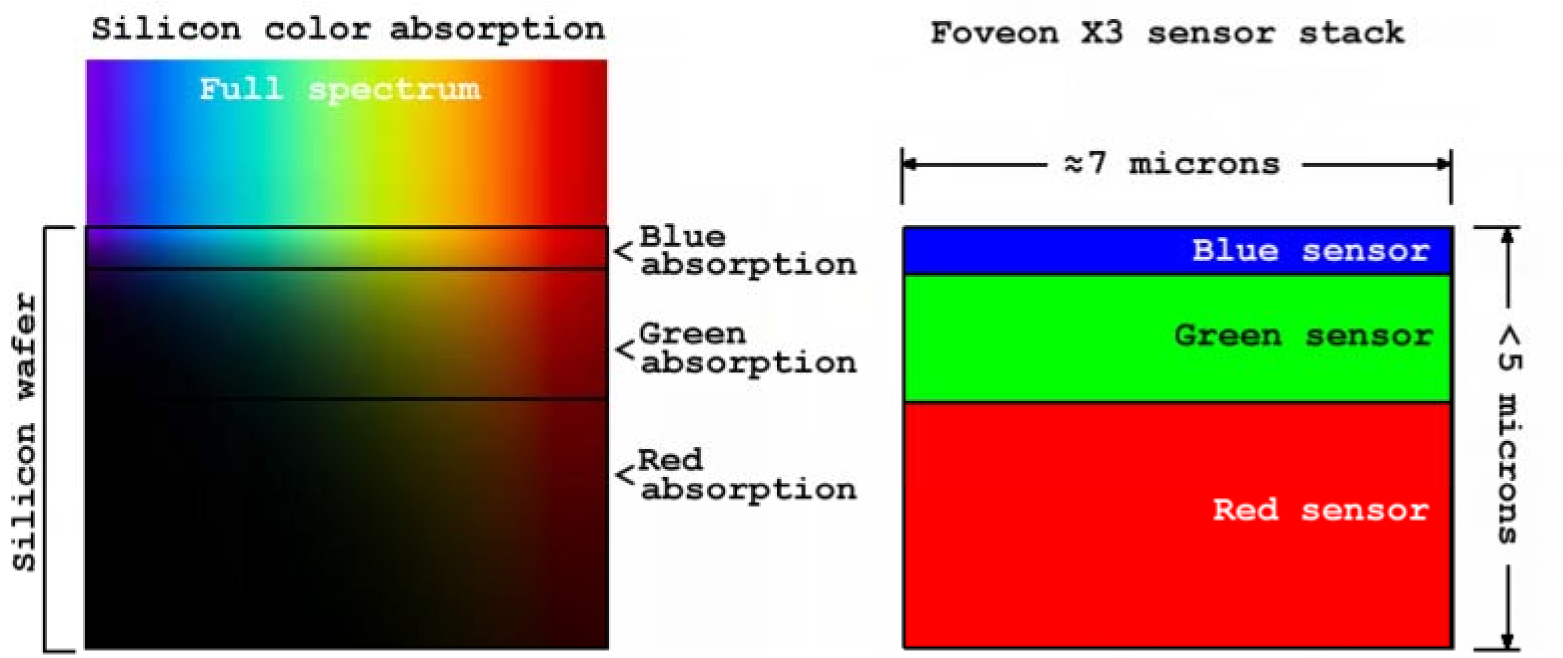
How to sense a color image

- Take 3 shots (temporal multiplexing)
- 3 detectors (e.g. Foveon, photographic film)
- spatial multiplexing
  - human eye
  - sensors are made sensitive to red, green or blue using a filter coating that blocks the complementary light



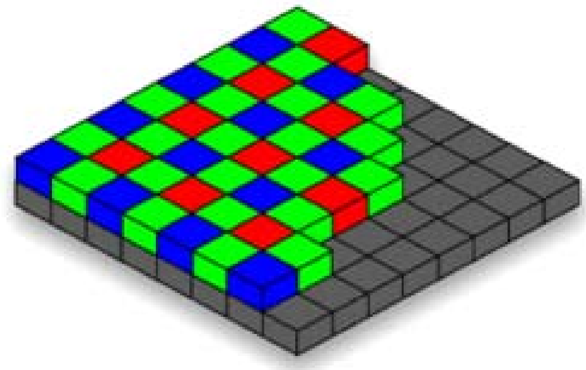
Foveon pixels

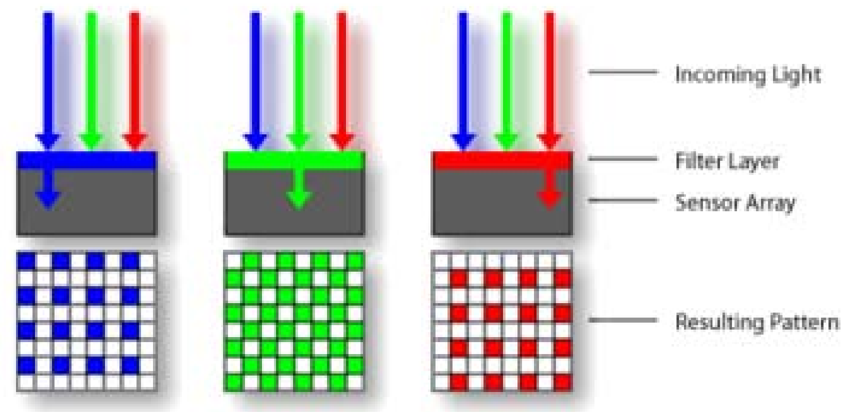
- Complex technology, easier analysis



Bayer Pattern

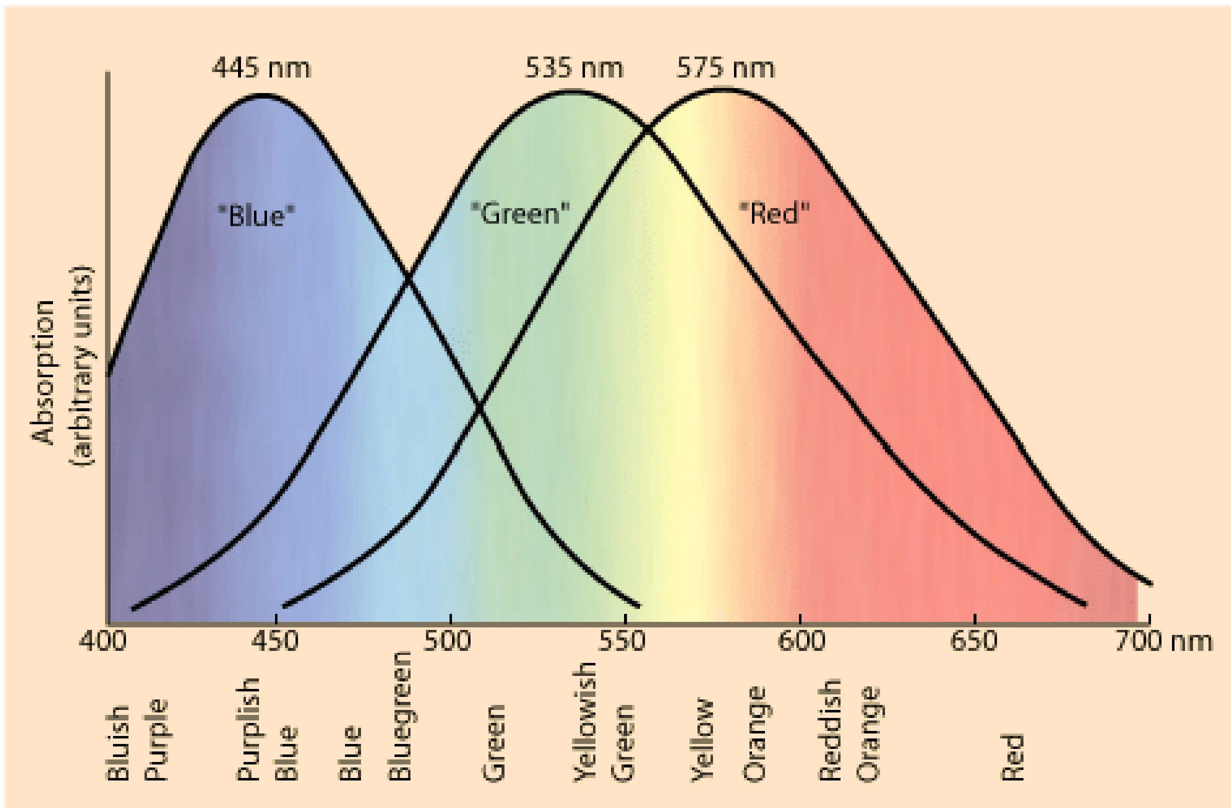
- Invented by E. Bayer at Kodak in 1976, it is a way to arrange RGB filter on a squared grid of photosensors
- 50% green, 25% red, 25% blue
- mimic eye's greater sensitivity to green wavelengths
- need *demosaicing* to interpolate the color information from neighbor units



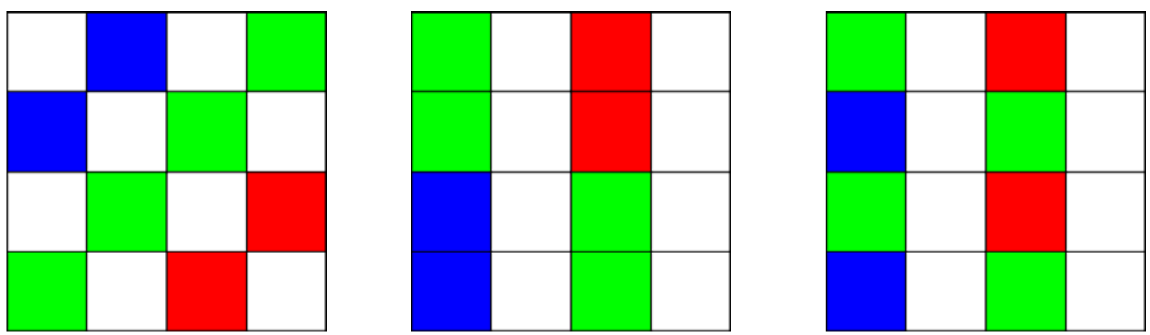


source: http://en.wikipedia.org/wiki/Bayer_filter

Human response to color



- Alternative sensor announced in 2007 by Kodak, add “panchromatic”, that are sensitive to all wavelengths
- Increase sensitivity to light, because panchromatic cells do not filter light



Demosaicing

- Reproduce the original image
- Avoid artifacts
- Often must be efficient

1 Simple nearest neighbor, take the missing colors from the nearest pixel

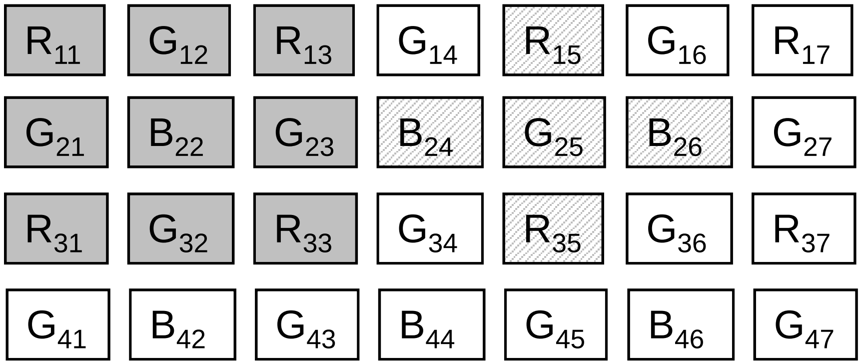
R_{11}	G_{12}	R_{13}	G_{14}	R_{15}	G_{16}	R_{17}
G_{21}	B_{22}	G_{23}	B_{24}	G_{25}	B_{26}	G_{27}
R_{31}	G_{32}	R_{33}	G_{34}	R_{35}	G_{36}	R_{37}

$R_{11} = R_{11}$ $R_{12} = R_{13}$

$G_{11} = G_{12}$ $G_{12} = G_{12}$

$B_{11} = B_{22}$ $B_{12} = B_{22}$ ...etc

2 Bilinear interpolation



$$R_{22} = 0.25 \cdot (R_{11} + R_{13} + R_{31} + R_{33})$$

$$G_{22} = 0.25 \cdot (G_{12} + G_{21} + G_{23} + G_{32})$$

$$B_{22} = B_{22}$$

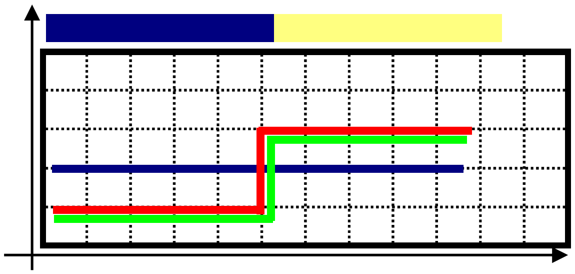
$$R_{25} = 0.5 \cdot (R_{15} + R_{35})$$

$$G_{25} = G_{25}$$

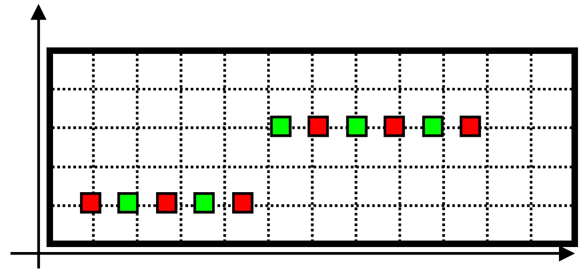
$$B_{25} = 0.5 \cdot (B_{24} + B_{25})$$

3 More sophisticated methods to reduce artifacts (but computationally more expensive)...

Artifacts: color “fringes”

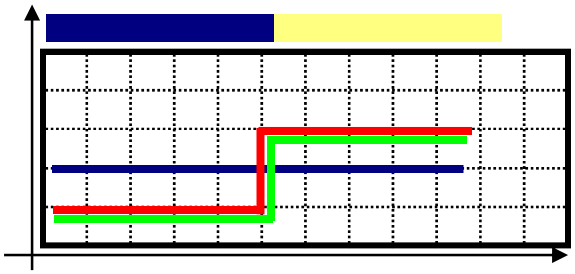


Original signal, an edge from blue (0,0,128) to yellow (255,255,128). Only one scanline shown.

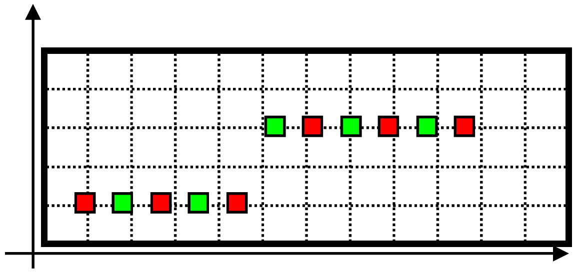


Subsampled Bayer pattern (red scanline)

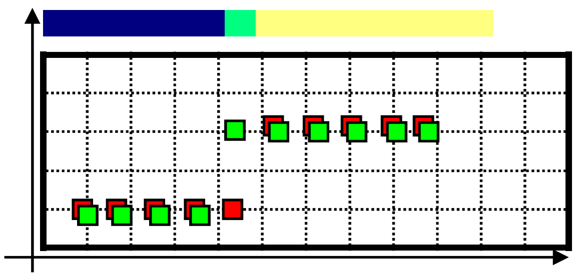
Artifacts: color “fringes”



Original signal, an edge from blue (0,0,128) to yellow (255,255,128). Only one scanline shown.

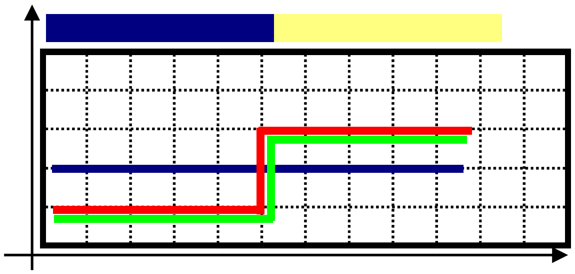


Subsampled Bayer pattern (red scanline)

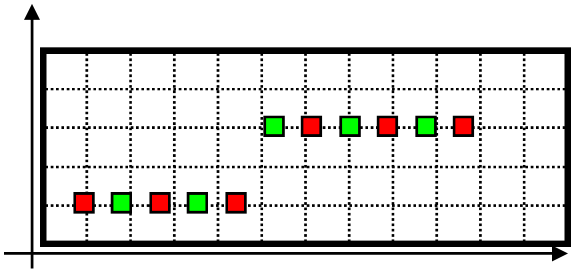


Nearest neighbor color reconstruction.

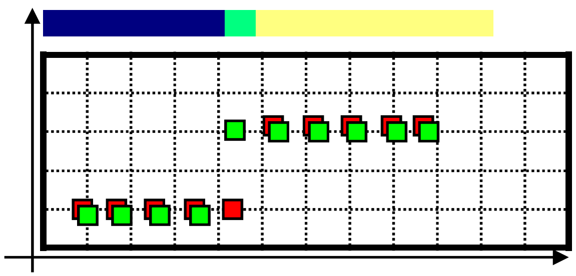
Artifacts: color “fringes”



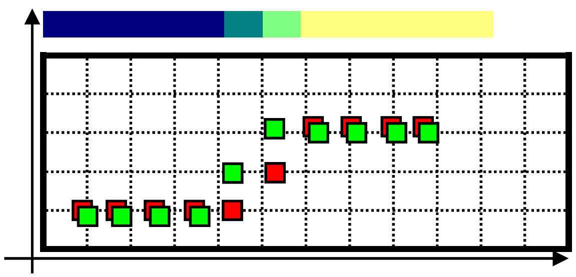
Original signal, an edge from blue (0,0,128) to yellow (255,255,128). Only one scanline shown.



Subsampled Bayer pattern (red scanline)

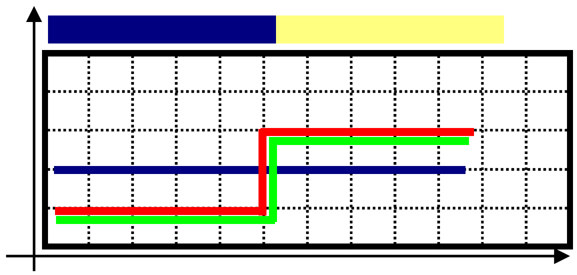


Nearest neighbor color reconstruction.

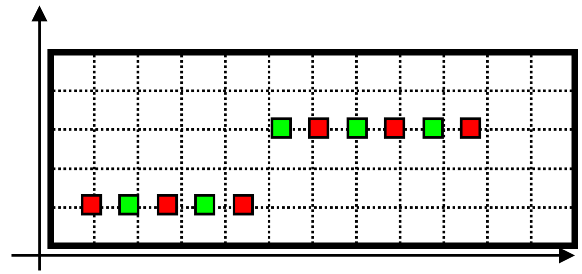


Linear interpolation.

Artifacts: color “fringes”

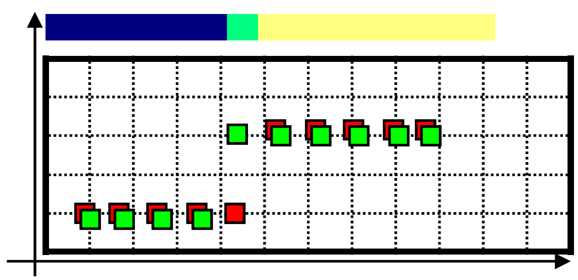


Original signal, an edge from blue (0,0,128) to yellow (255,255,128). Only one scanline shown.

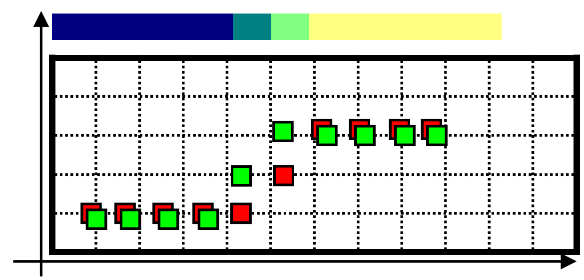


Subsampled Bayer pattern (red scanline)

color fringes



Nearest neighbor color reconstruction.



Linear interpolation.

adapted form: A.Lukin, D.Kubasov, Graphicon 2004

Constant Hue-Based Interpolation

Limit abrupt hue changes (“fringes”) across pixels.

Consider R and B the chrominance values, G is assigned the luminance.
Define hue as: (R/G, B/G)

- First, G values are computed by bilinear interpolation
- Bilinear interpolation of the **hue values** for the R and B channels
- R and B values reconstructed accordingly:

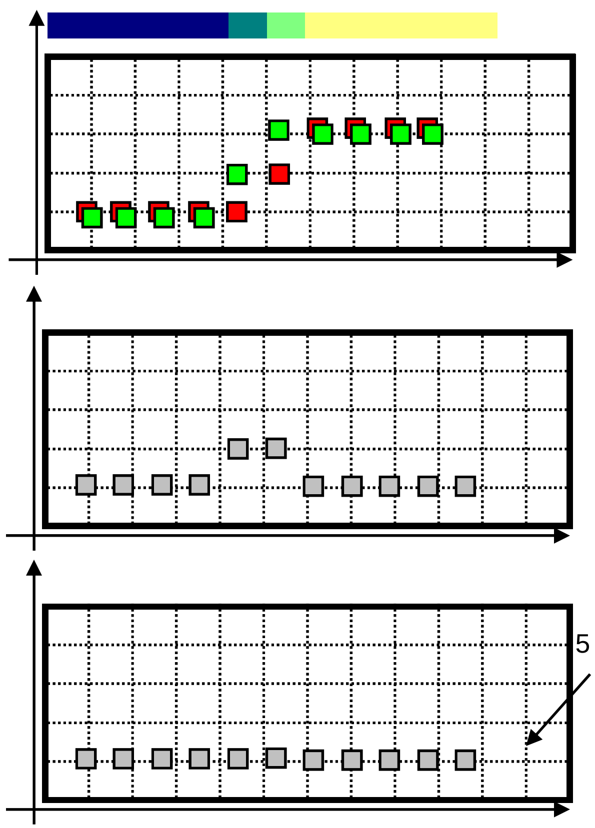
R_{11}	G_{12}	R_{13}	G_{14}	R_{15}
G_{21}	B_{22}	G_{23}	B_{24}	G_{25}
R_{31}	G_{32}	R_{33}	G_{34}	R_{35}
G_{41}	B_{42}	G_{43}	B_{44}	G_{45}

$$R_{22} = G_{22} \cdot 0.25 \cdot (R_{11}/G_{11} + R_{13}/G_{13} + R_{31}/G_{31} + R_{33}/G_{33})$$
$$G_{22} = 0.25 \cdot (G_{12} + G_{21} + G_{23} + G_{32})$$
$$B_{22} = B_{22}$$

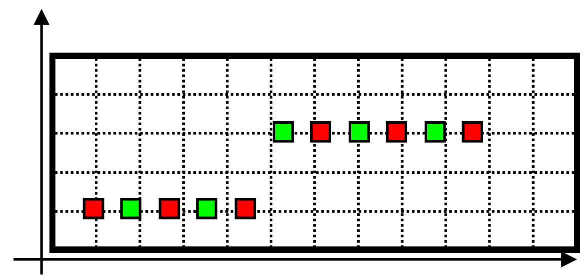
$$R_{33} = R_{33}$$
$$G_{33} = 0.25 \cdot (G_{23} + G_{34} + G_{43} + G_{32})$$
$$B_{33} = G_{33} \cdot 0.25 \cdot (B_{22}/G_{22} + B_{24}/G_{24} + B_{44}/G_{44} + B_{42}/G_{42})$$

Median-Based Interpolation

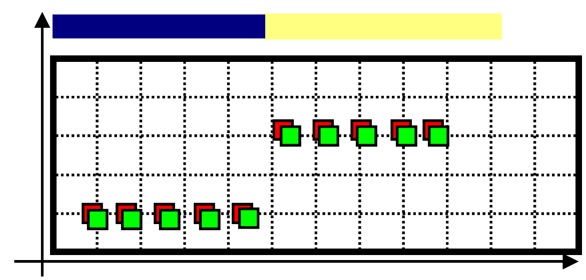
- Perform bilinear interpolation of all channels
- Filter differences R-G and B-G, with a median filter
- Add the median filtered image to the sampled data



5 samples median filter



Original Bayer pattern



Non-uniform sensors

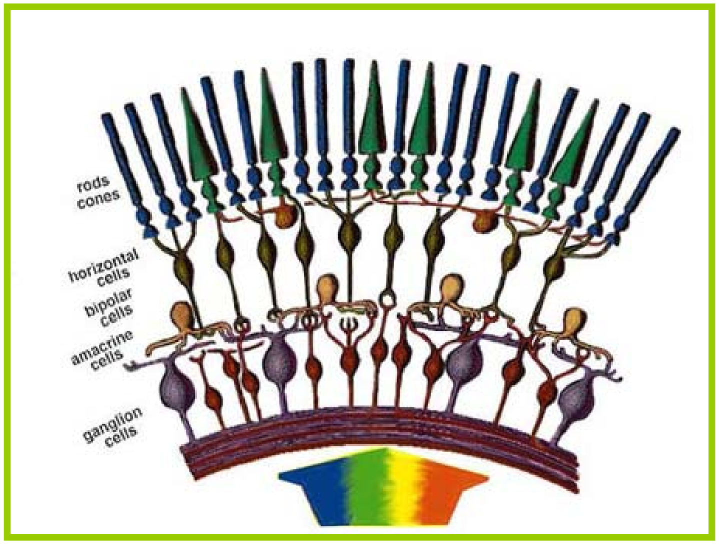
...vision is an information processing task...

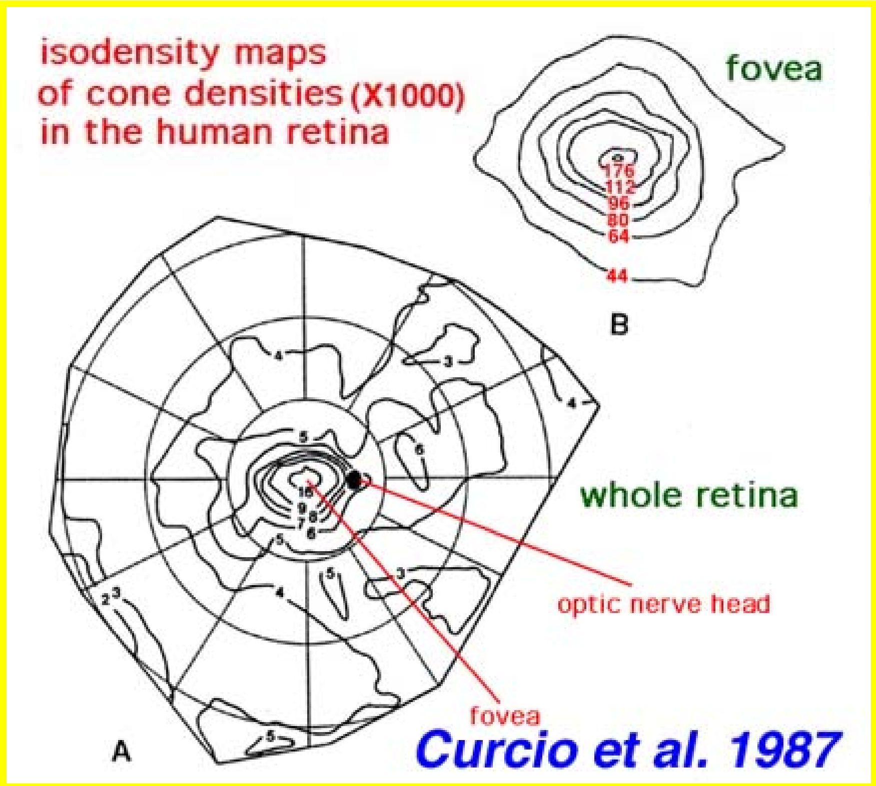
One of the major concerns of evolution has been to shape the sensory systems so that information is reduced as soon as possible (ideally even before it enters the processing device)

…may be less information is better

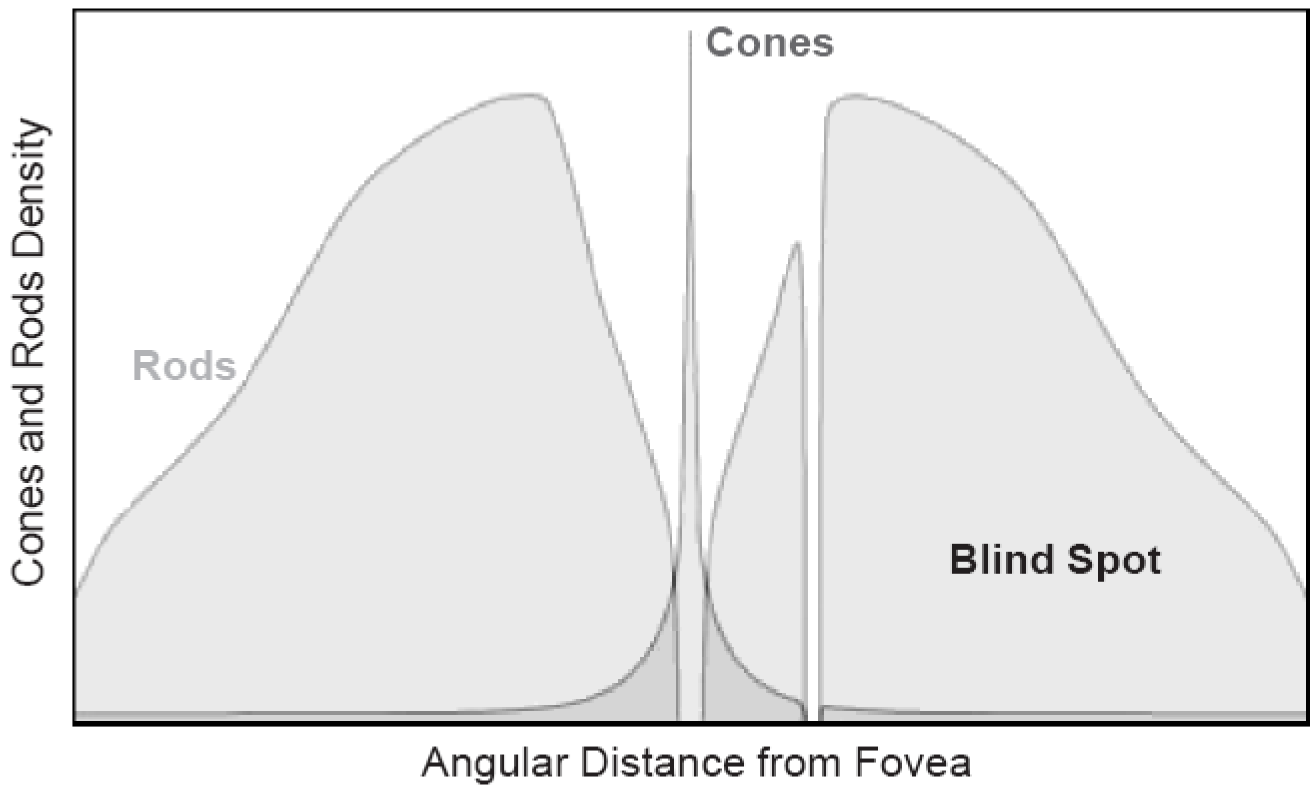
One example is the structure of the human retina

Back to the retina





Distribution of the receptors



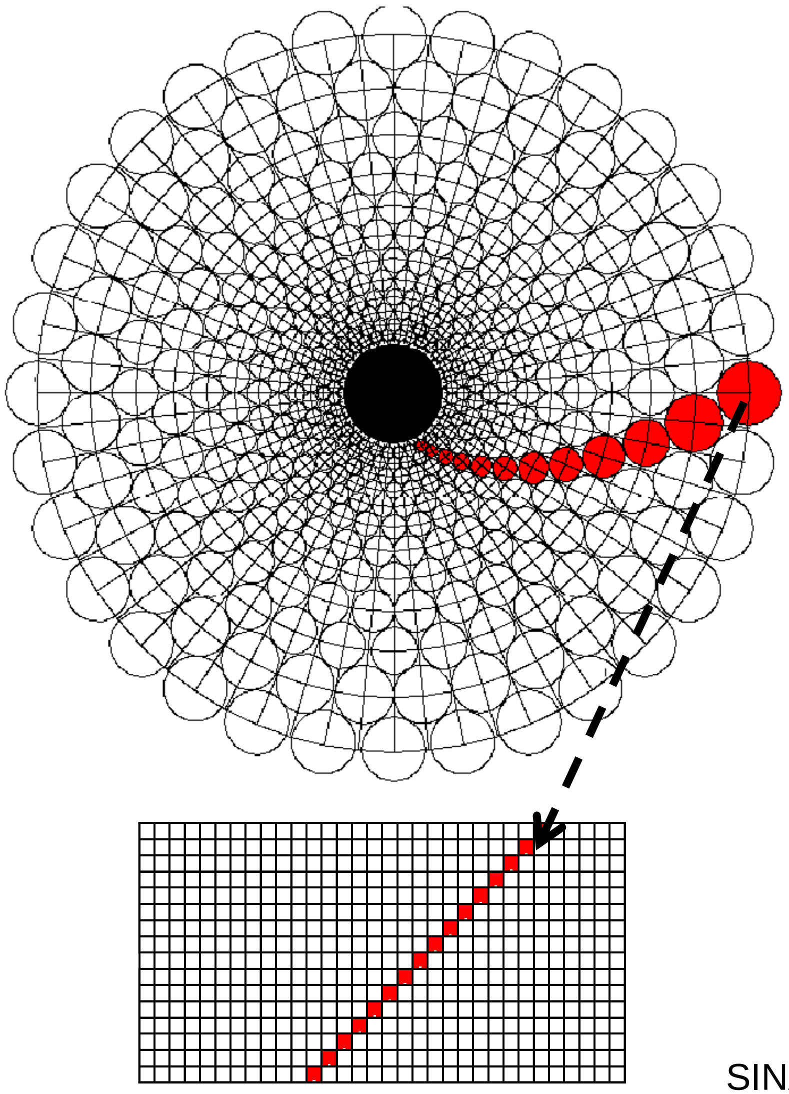
If the retina was designed as a camera

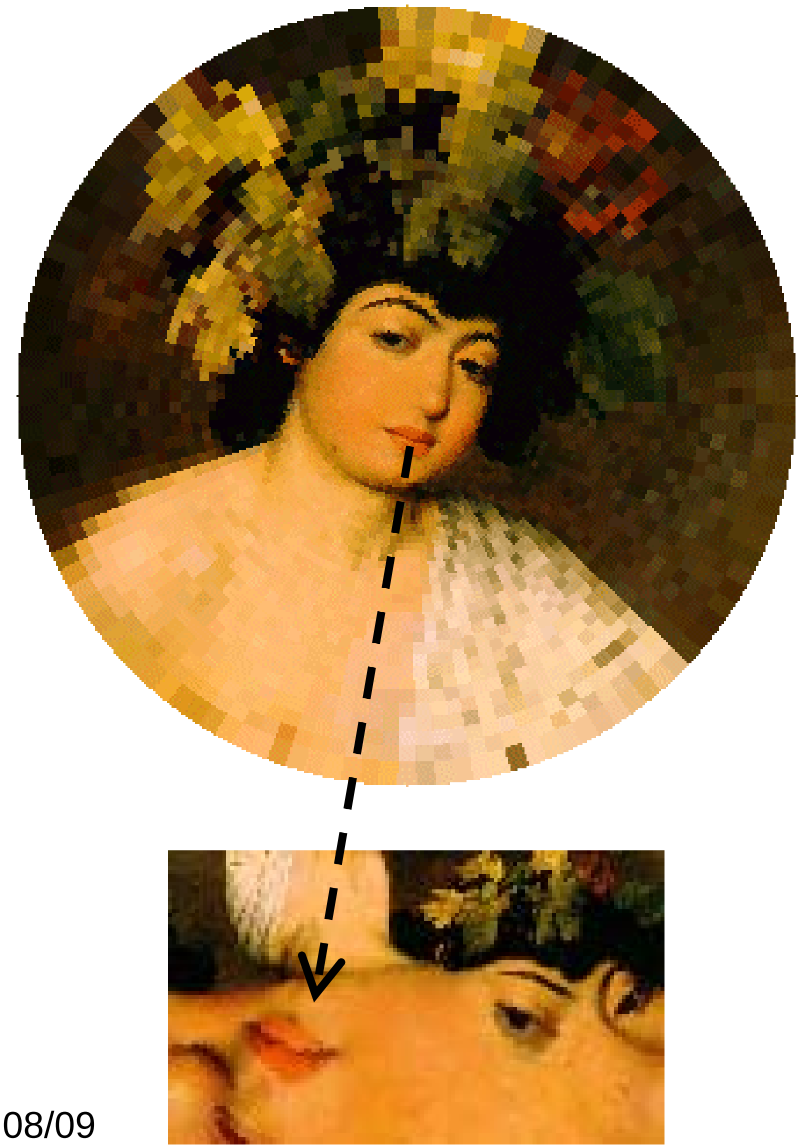
Visual Field: about 160 deg
Maximum Resolution: about 1/60 deg

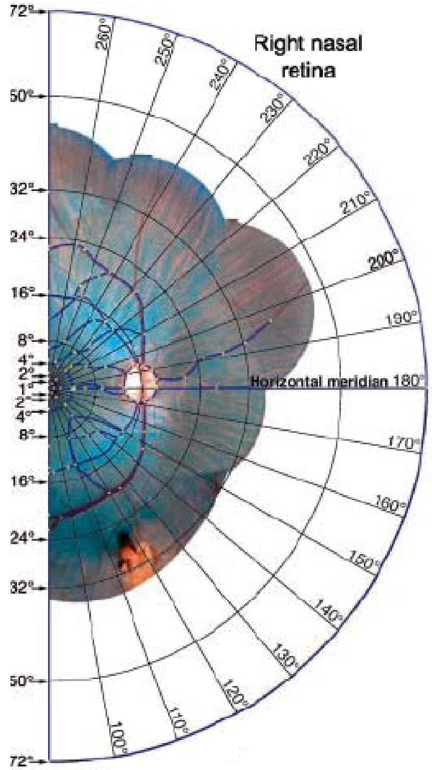
According to K. Nakayama and E. Schwartz the saving is from 5,000 to 30,000 times.

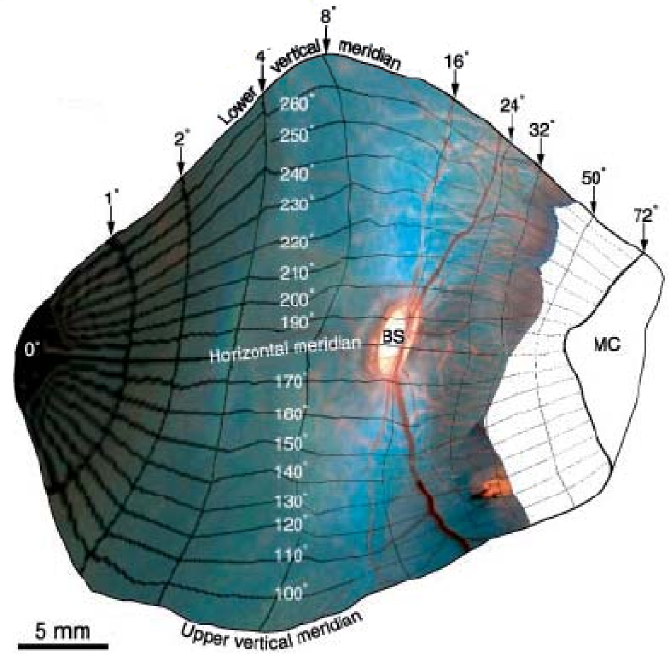
Optic nerve: diameter 4 cm
Brain weight: from about 3 to 20 tons
Amount of food: ?
Processing time: ?

Log-polar images

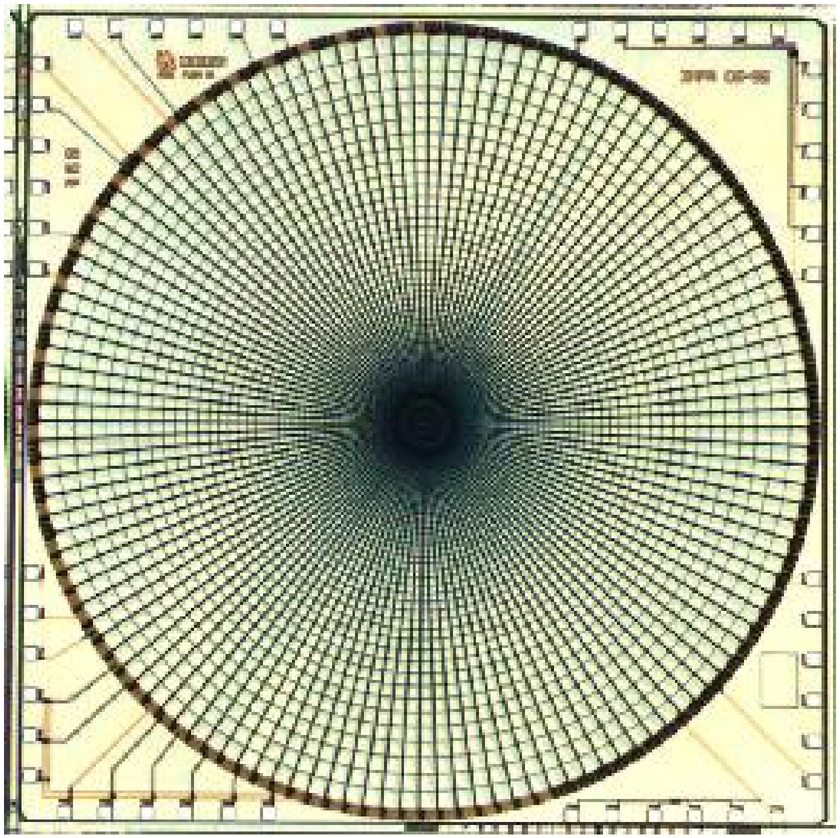




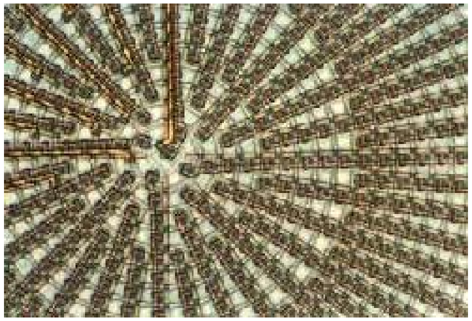

Right nasal
retina
Horizontal meridian 180°
72°
50°
32°
24°
16°
8°
4°
2°
1°
260°
250°
240°
230°
220°
210°
200°
190°
170°
160°
150°
140°
130°
120°
110°
100°


Lower vertical meridian
Upper vertical meridian
Horizontal meridian
BS
MC
0°
1°
2°
4°
8°
16°
24°
32°
50°
72°
5 mm

1996 – CMOS

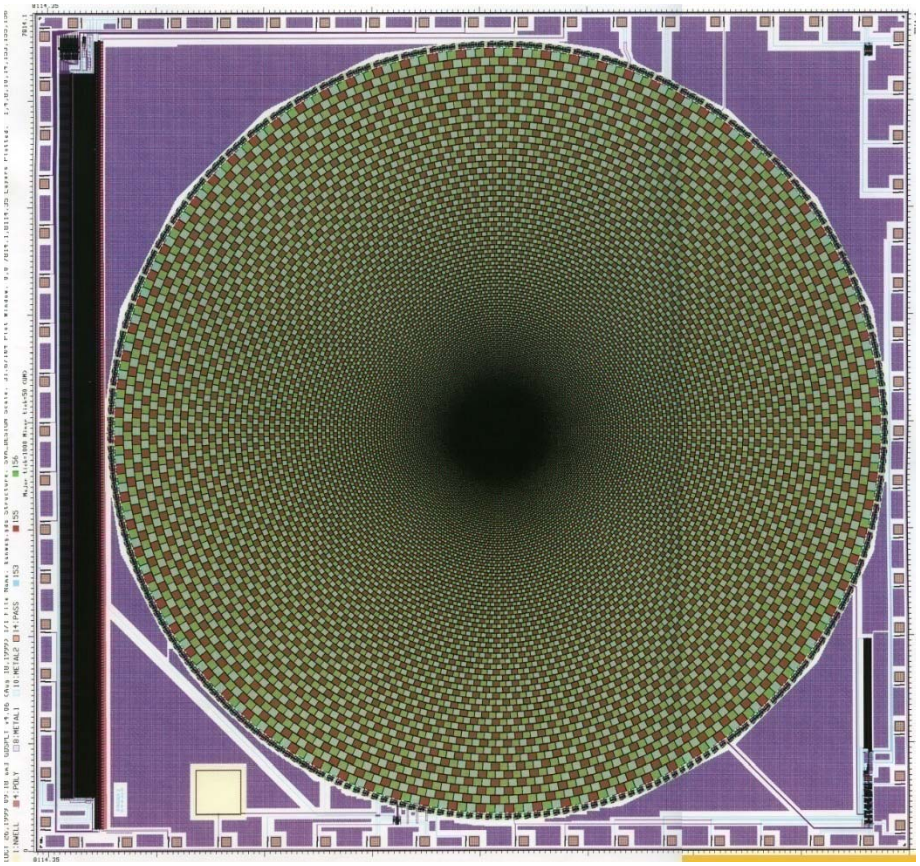


Monochromatic
8000 Pixel (128x56)
Diameter 8.1 mm
Min Pix Size = 14 µm
Max/Min = 14.
Q = 600

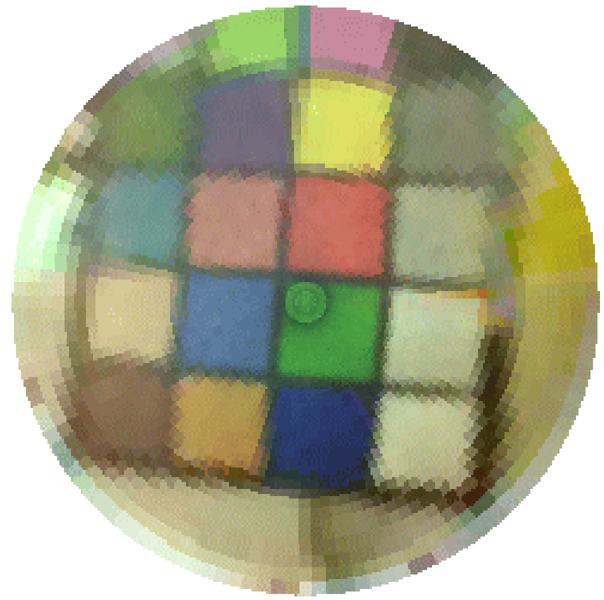


"polar" fovea

1998 – CMOS Color

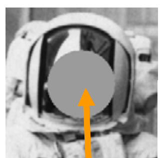


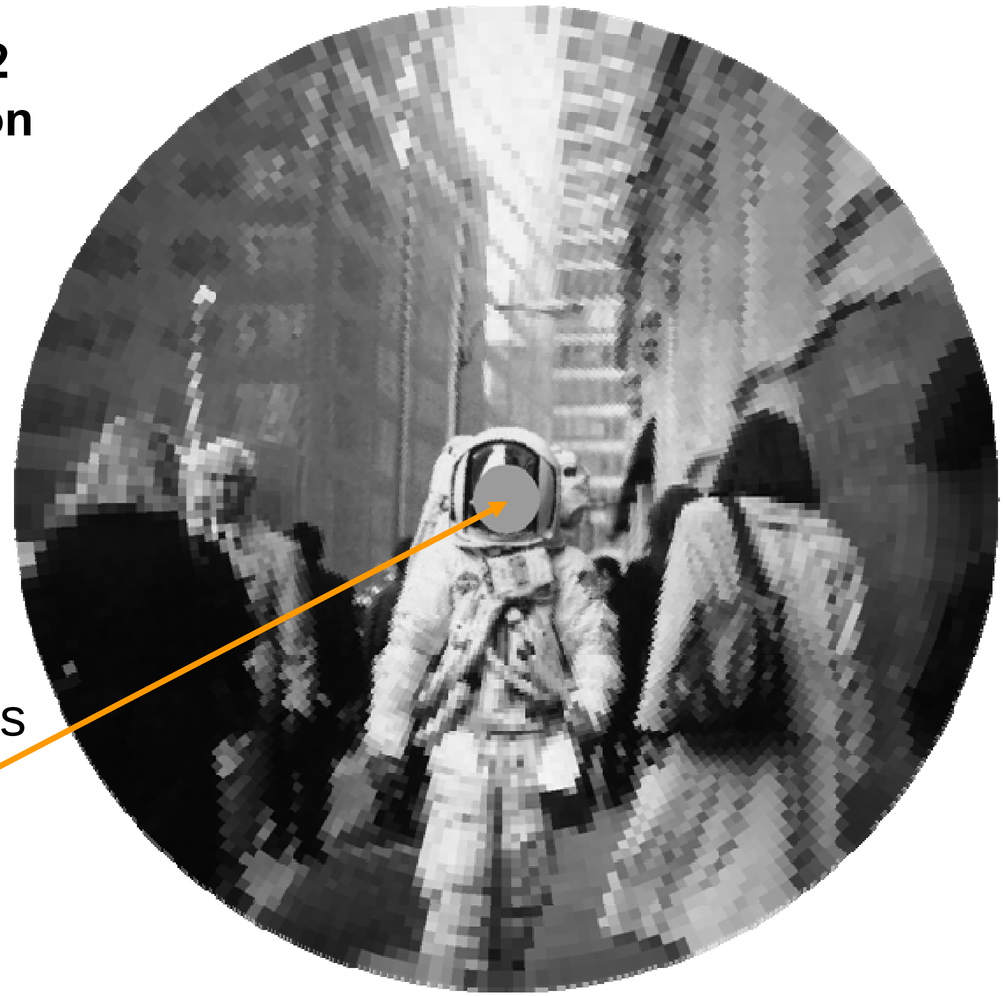
Same Layout



How to compare Giotto-2 with a constant resolution sensor array?

Suppose we start with 5000 pixels at constant resolution





27000 pixels

5000 pixels

The same 5000 pixels
plus 27000 retina-like pixels

33,000 pixels with same field of view (less resolution)

Enlarging the constant resolution image shows the loss in resolution

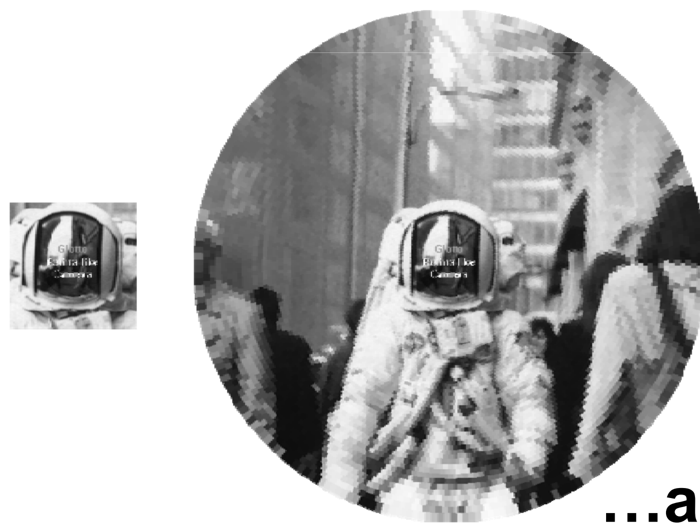
The two images are shown so that the size of the smallest pixel is the same (same maximum resolution)

Another example

6600 Pixels

x 2 Pixels
SubQ-CIF

x 3 Pixels



x 4 Pixels
Q-CIF

…advantages increase

Table 1. A comparison between three generations of log-polar sensors.

Sensor Version	Total Number of Pixels	Pixels in Fovea	Pixels in Periphery	Total Number of Rings	Pixels per Ring
CCD	2022	102	1920	30	64
CMOS 8k	8013	845	7168	76	128
CMOS 33k	33193	5473	27720	152	252

Sensor Version	Rings in Fovea	Rings in Periphery	Pixels in Periphery	Angular Amplitude	Logarithm Base
CCD	—	30	1920	5.413°	1.094
CMOS 8k	20	56	7168	2.812°	1.049
CMOS 33k	42	110	27720	1.428°	1.02337

Sensor Version	Ø of the Sensor	Size of the Smallest Pixel	R	Q	Technology Used	Radius of the Fovea
CCD	9400 μm	30 μm	13.7	300	1.5 μm	317 μm
CMOS 8k	8100 μm	14 μm	14	600	0.7 μm	285 μm
CMOS 33k	7100 μm	6.5 μm	17	1100	0.35 μm	273 μm

Geometrical interpretation

$$\begin{cases} \eta = q\vartheta \\ \xi = \log_a \frac{\rho}{\rho_0} \end{cases}$$

ρ and ϑ are the standard polar coordinates

$$\begin{cases} x = \rho\cos\vartheta \\ y = \rho\sin\vartheta \end{cases}$$

Another view

$$w = f(z) = \log_a(z), \quad w, z \in C$$

$$z = x + iy = r(\cos\varphi + i\sin\varphi)$$

$$w = \rho(z) + i\vartheta(z)$$

$$z = re^{i\varphi}$$

$$\begin{cases} \rho = \log_a r \\ \vartheta = h\varphi \end{cases}$$