CSCI435 – Computer Vision

Lecture 3

Image Formation

Lecturer: Dr. Igor Kharitonenko
Room 3.108
ph: 4221 4825
igor@uow.edu.au

Machine Vision Concept (review)

- Machine Vision is the process where a digital computer automatically controls the process of image capturing and image analysis to understand the content of the image

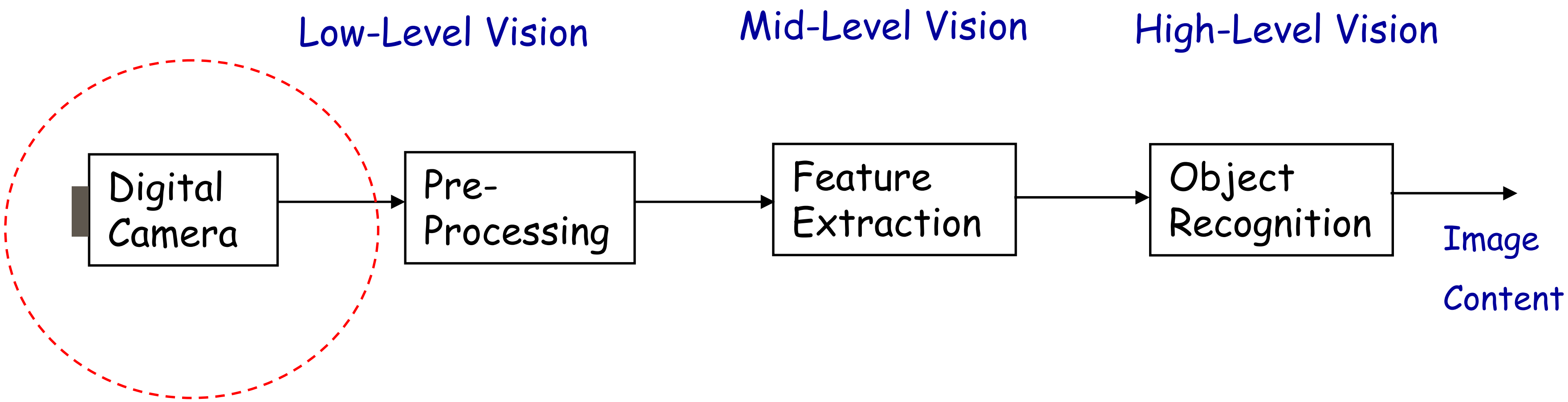
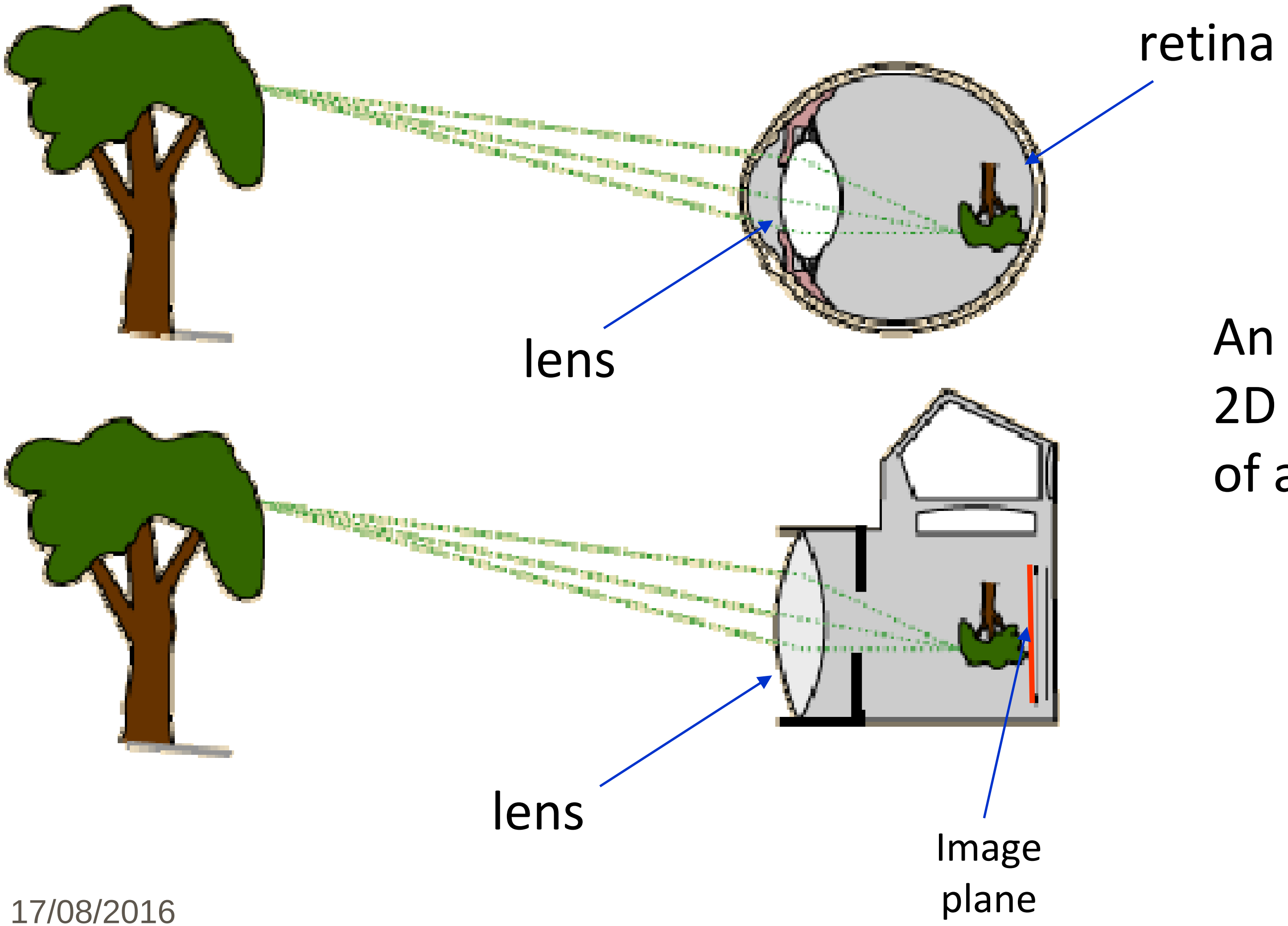
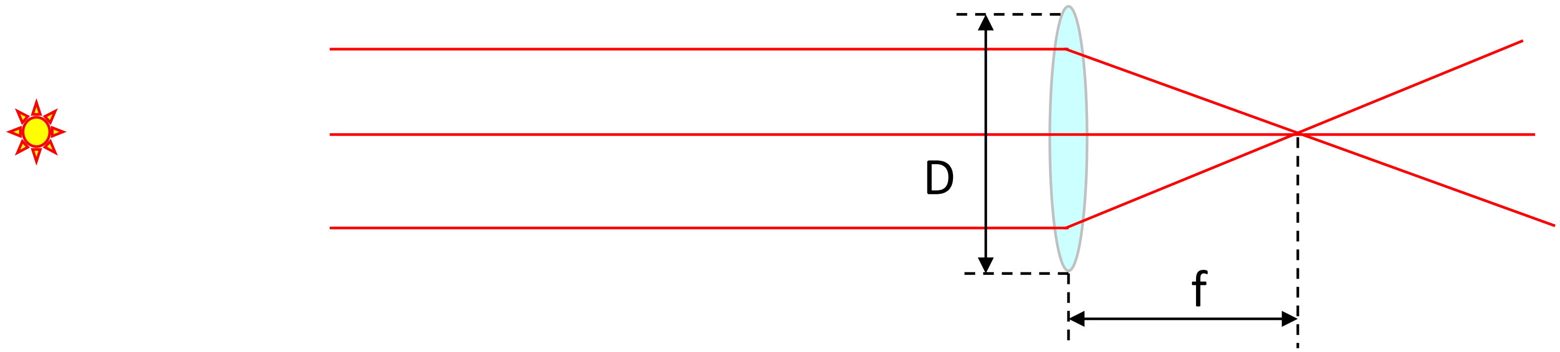


Image Formation



An image is a
2D projection
of a 3D scene

Lenses



Major parameters:

`f` – focal length

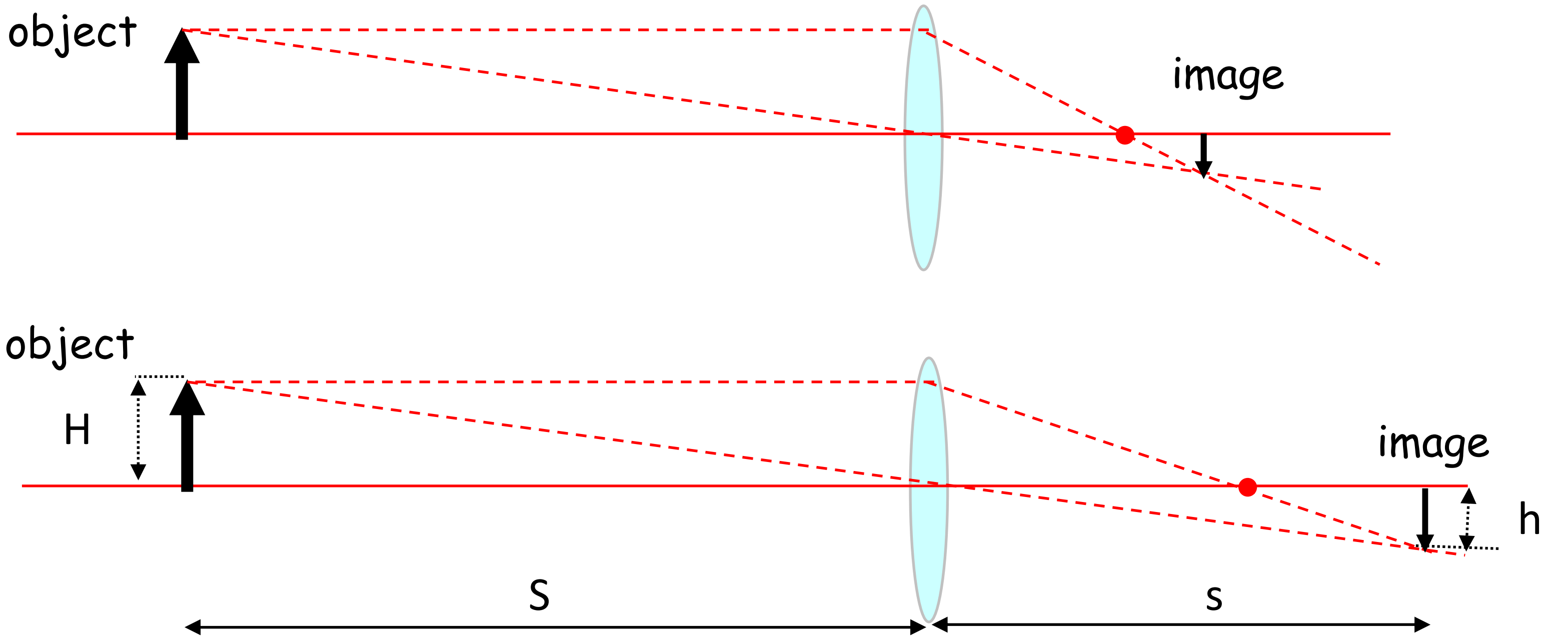
D – diameter

F – F-number

Focal length `f` is the distance required to focus incoming parallel rays to a single point. It indicates magnification property of a lenes

`F = f/D` indicates light collection property. The smaller F, the more light is collected

Image Formation Geomety (thin lense)

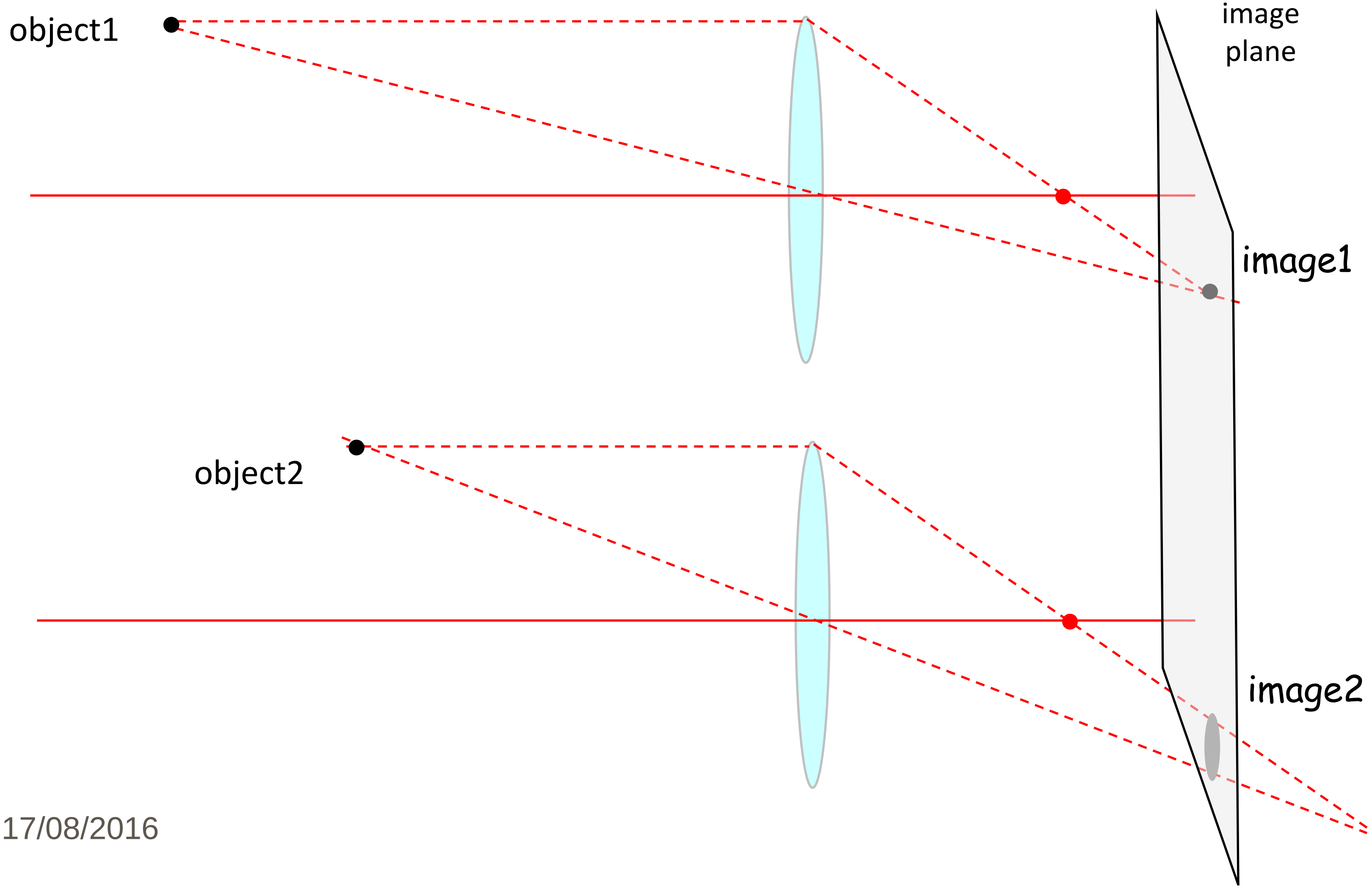


The longer the focal length, the larger the image

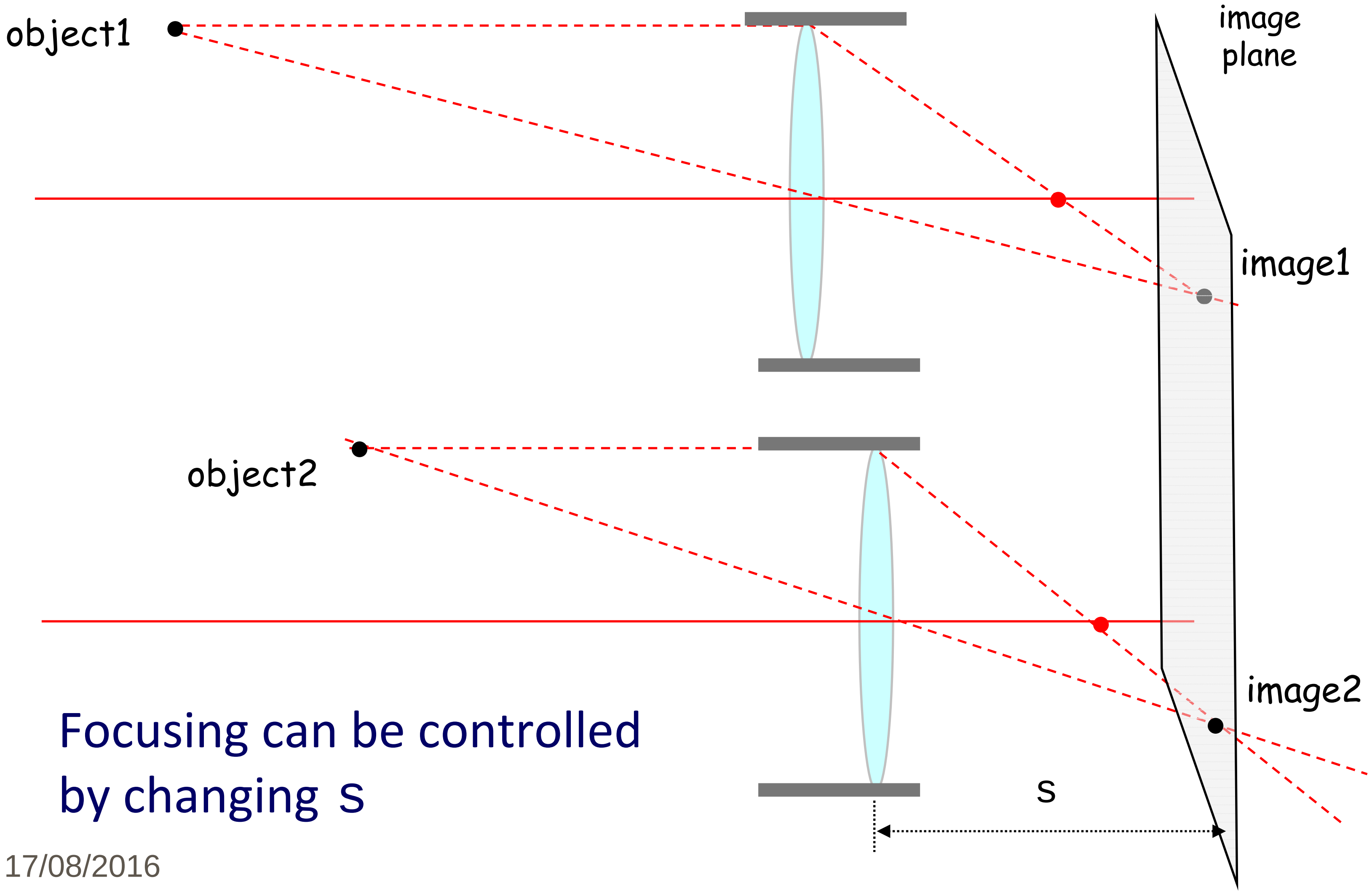
Lens Equation: `1/S + 1/s = 1/f`

Lens Magnification: `m = h/H = s/S = f/(S-f) = s/f - 1`

The Concept of Focusing

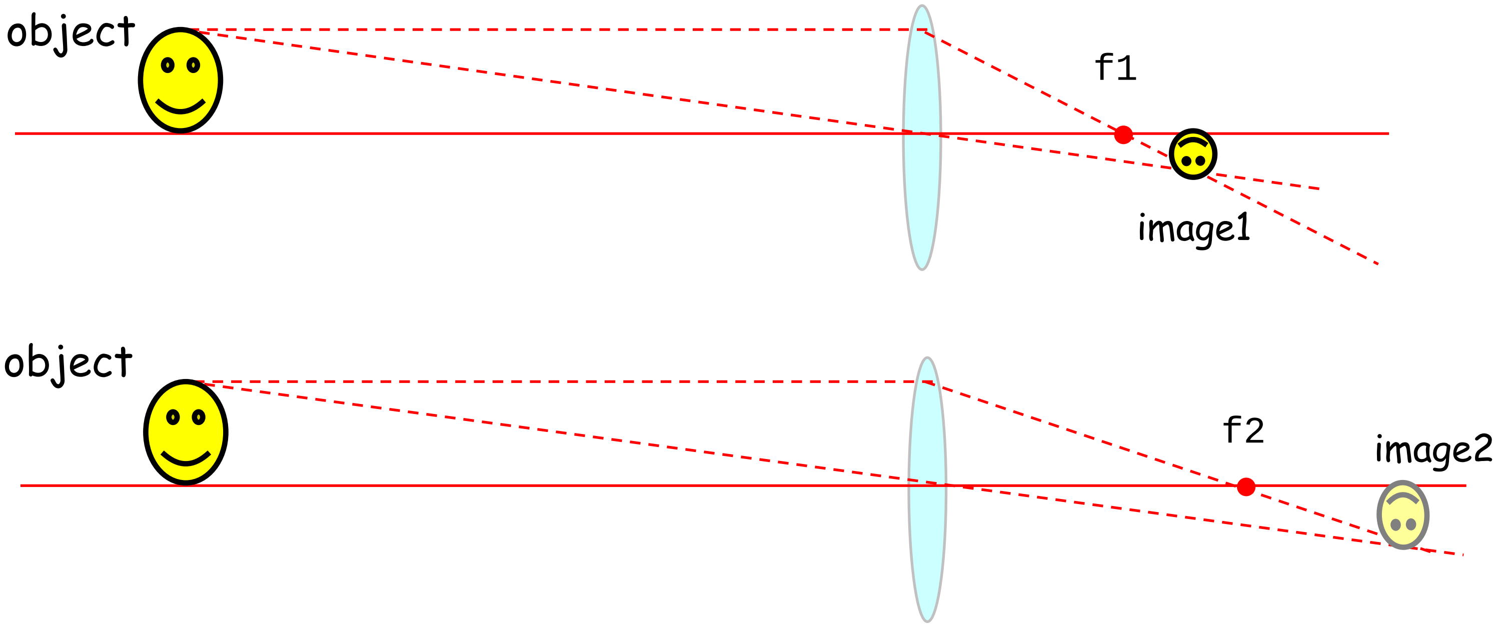


The Concept of Focusing



Focusing can be controlled by changing s

Quiz



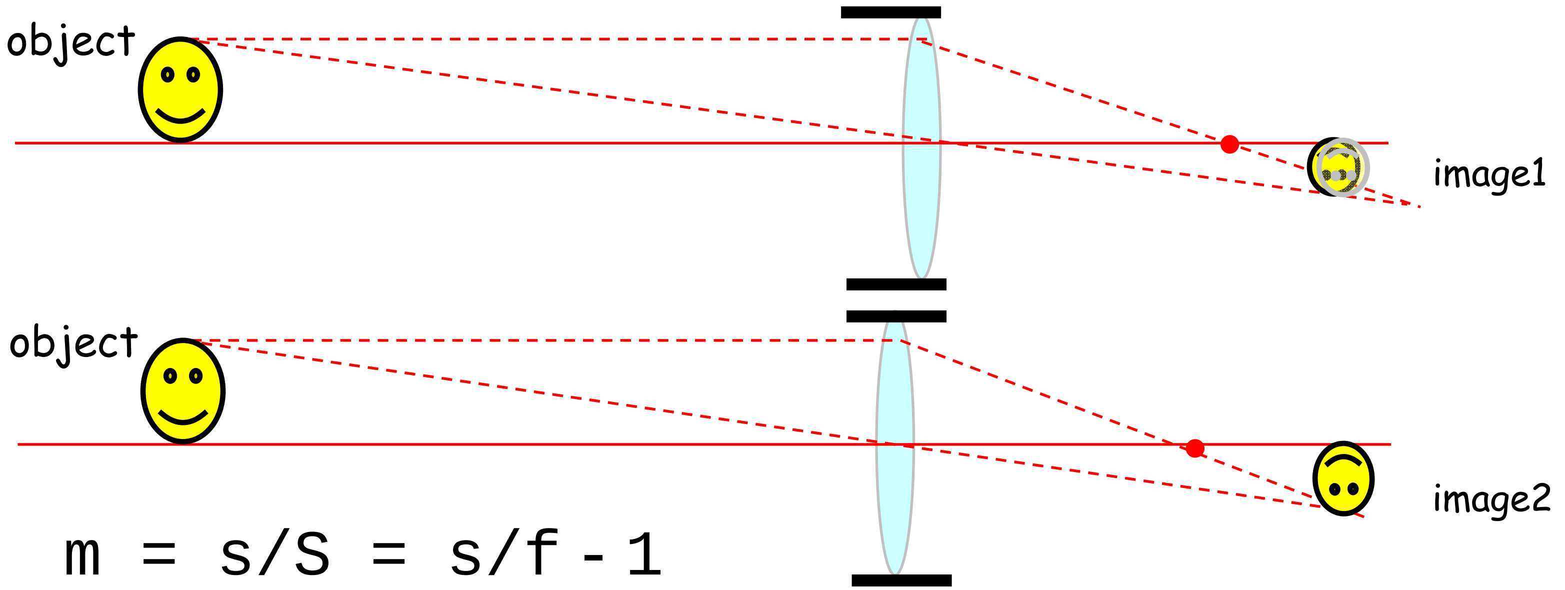
Considering D = const and f2>f1,

1. E(image1) = E(image2)
2. E(image1) > E(image2)
3. E(image1) < E(image2)

Prove your answer using the definition of Illuminance, or F-number

Quiz

$$1/S + 1/s = 1/f$$

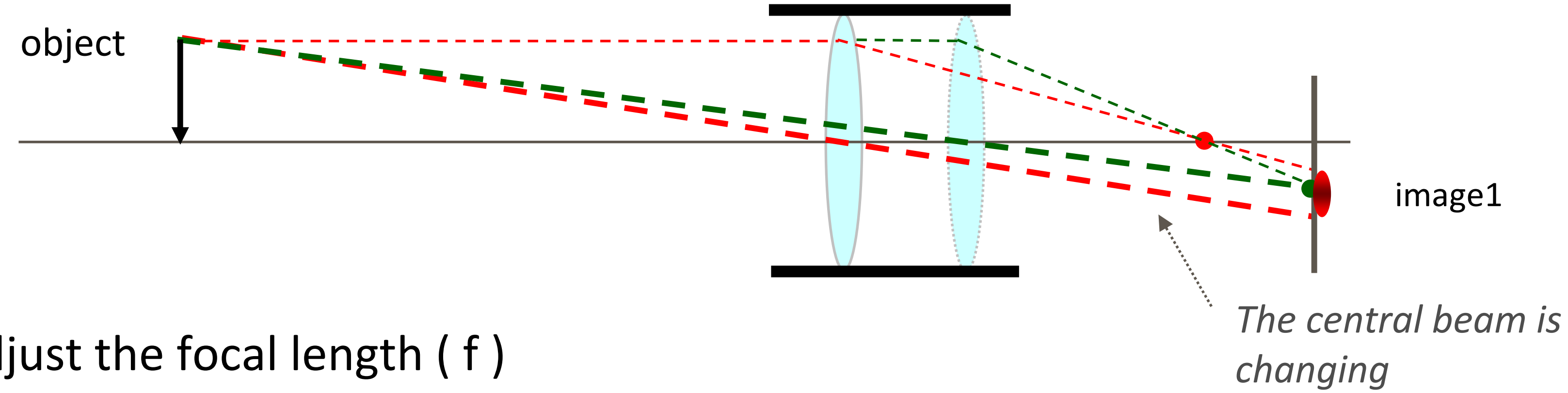


$$m = s/S = s/f - 1$$

- Focusing can be controlled by moving the lens. What is a side effect of this simple approach?
- What alternative method could you propose?

Two ways of changing focus

1. Adjust the lens position (s)



2. Adjust the focal length (f)

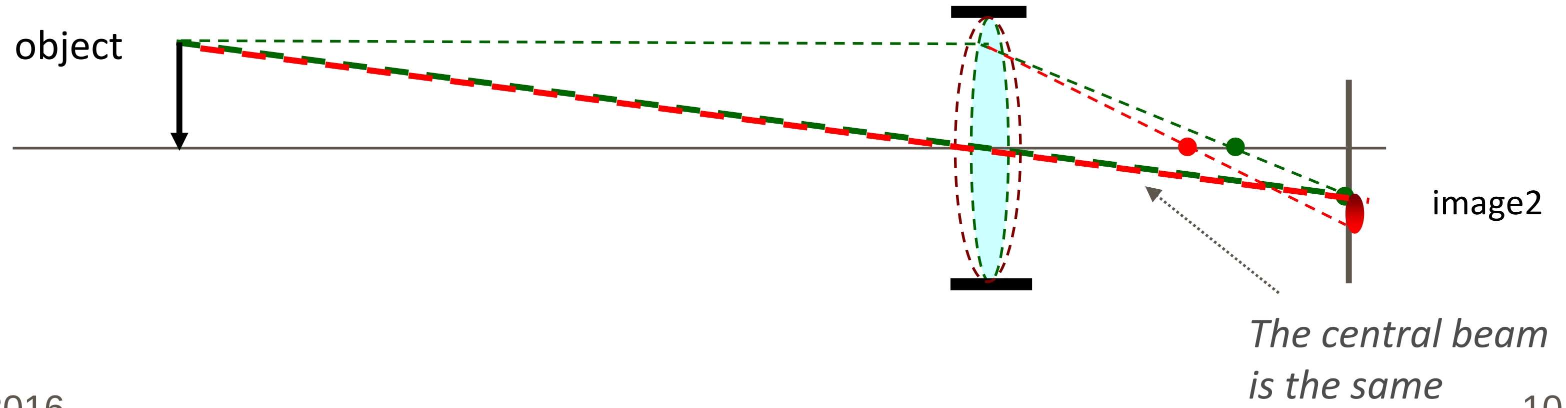
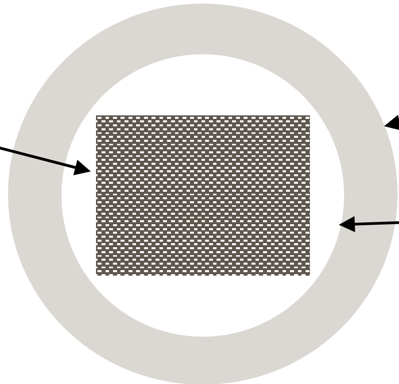


Image Circle



- Image quality produced by a lens is the best at the centre of the image plane with gradual degradation towards the outer image circle

  - geometric distortions
  - blur
  - chromatic aberrations (refractive index $n = \lambda_i / \lambda_{rf}$ is wavelength dependent)

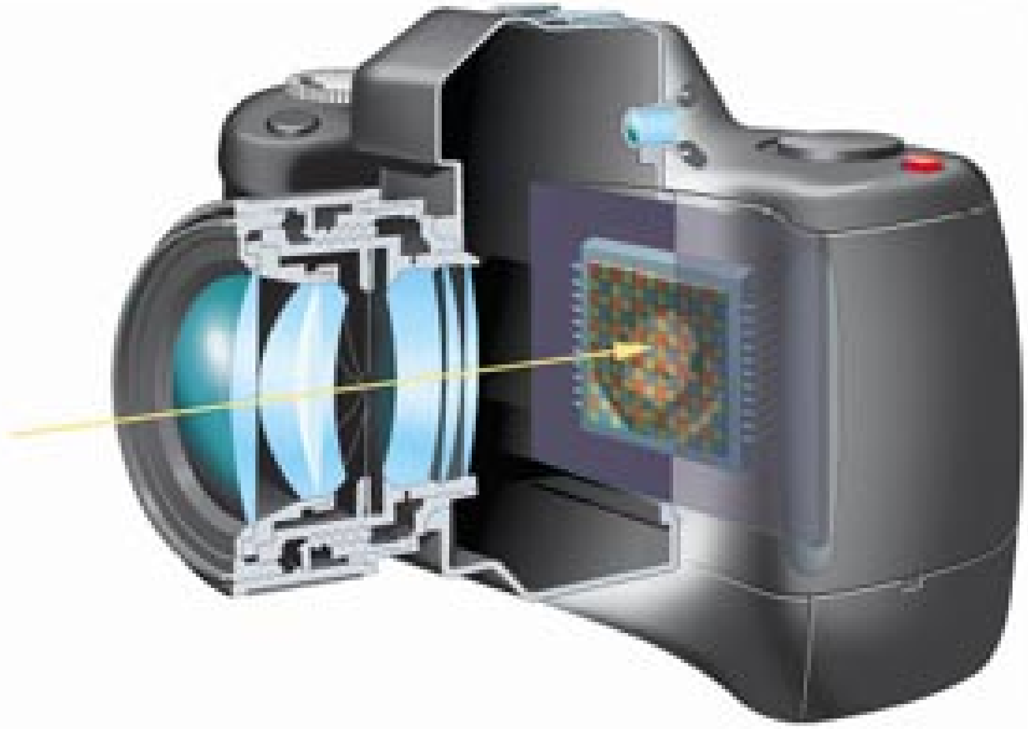
- Outside Inner Image Circle image quality is not guarantied

Digital Camera Image Sensors

1. Light goes through lens and hits image sensor plane.

2. Image sensor plane is a checker board pattern of color.

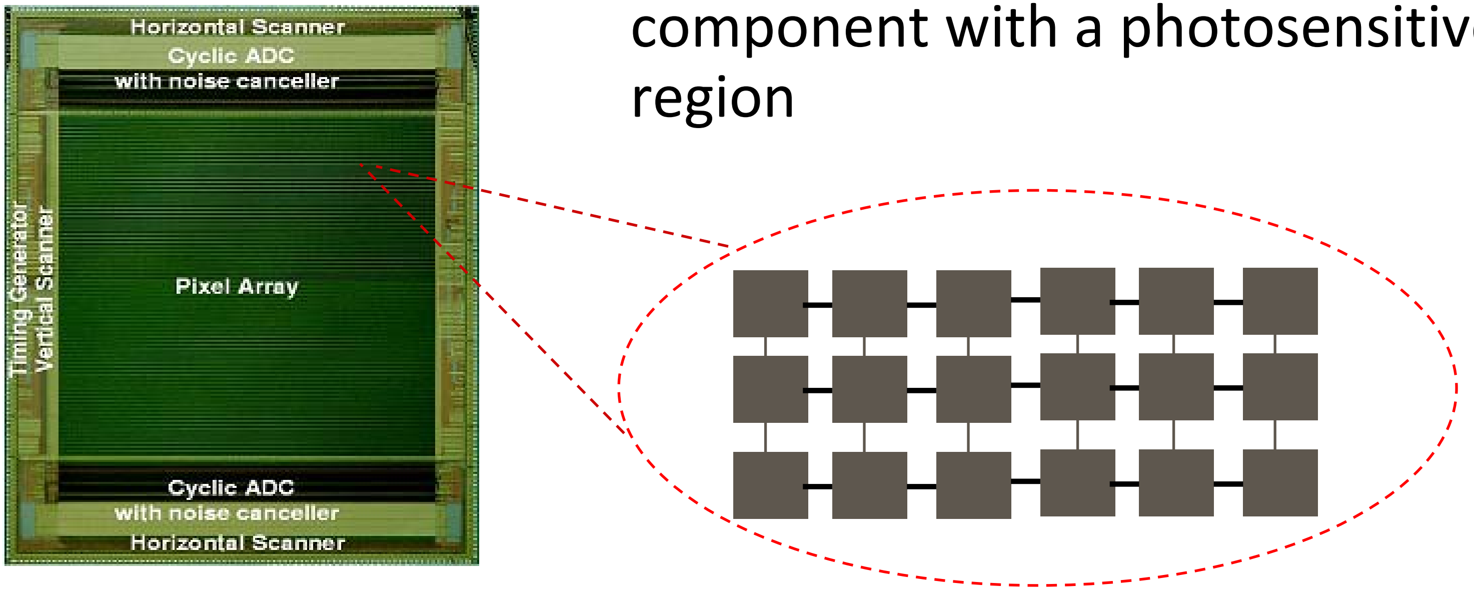
3. Camera estimates image color from checker board pattern.



FOVEON

Image Sensor

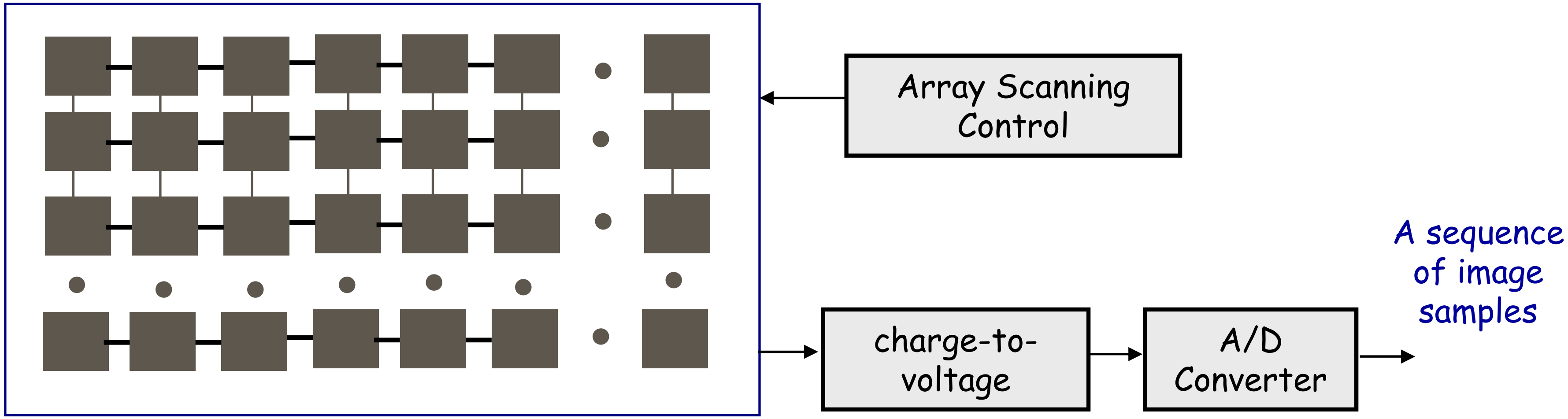
- Image Sensor is semiconductor component with a photosensitive region



- Photosensitive region is a 2D array of photosensitive elements that accumulate electric charges `q(i,j)` proportional to the illuminance `E(i,j)` and exposure time `t`

$$q(i,j) = k*E(i,j)*t$$

Image Sensor



- Accumulated charges q(i,j) are sequentially scanned row-by-row
- Electric charges are converted to voltages v(i,j) and digitised
- The sensor output is a sequence of digitised image samples

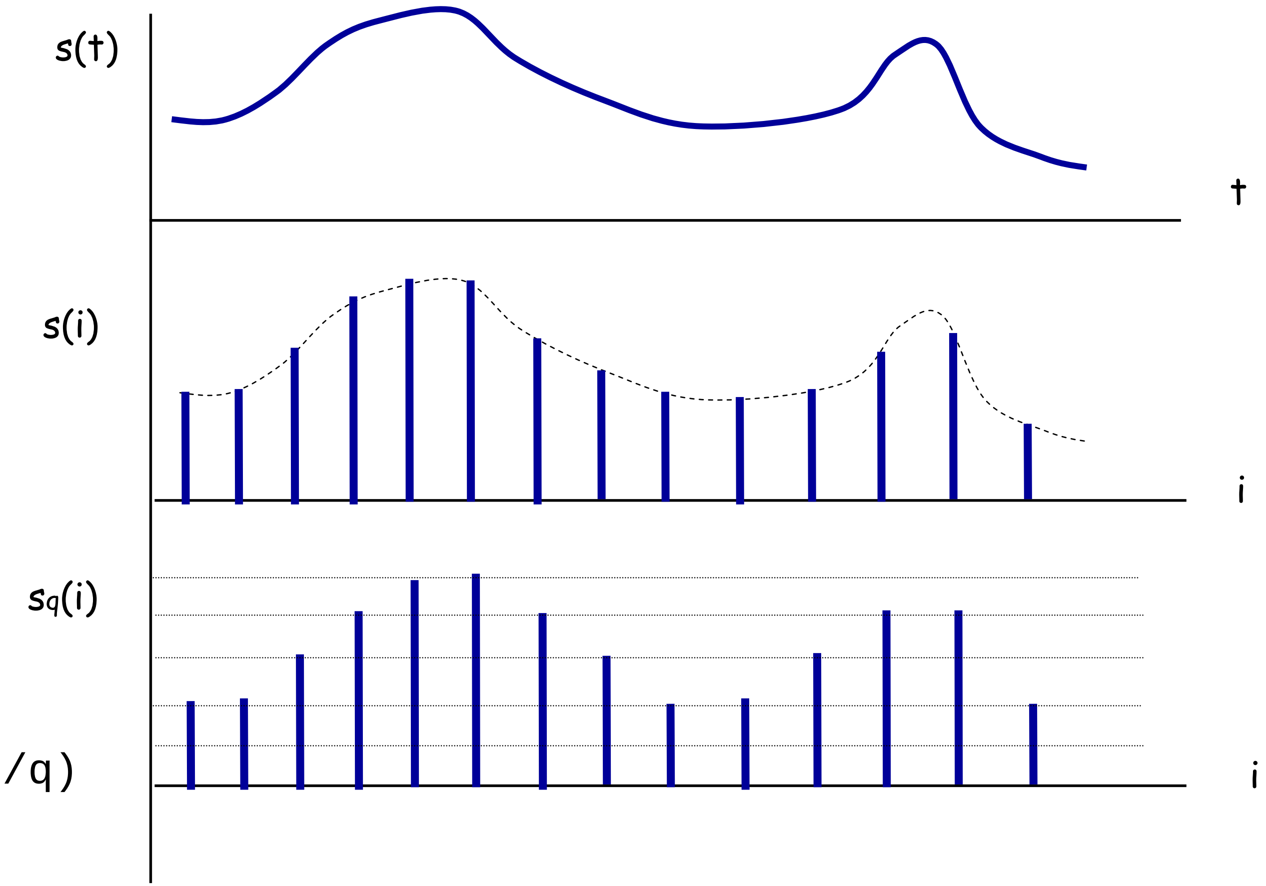
Digitization

Continues signal

Sampled signal

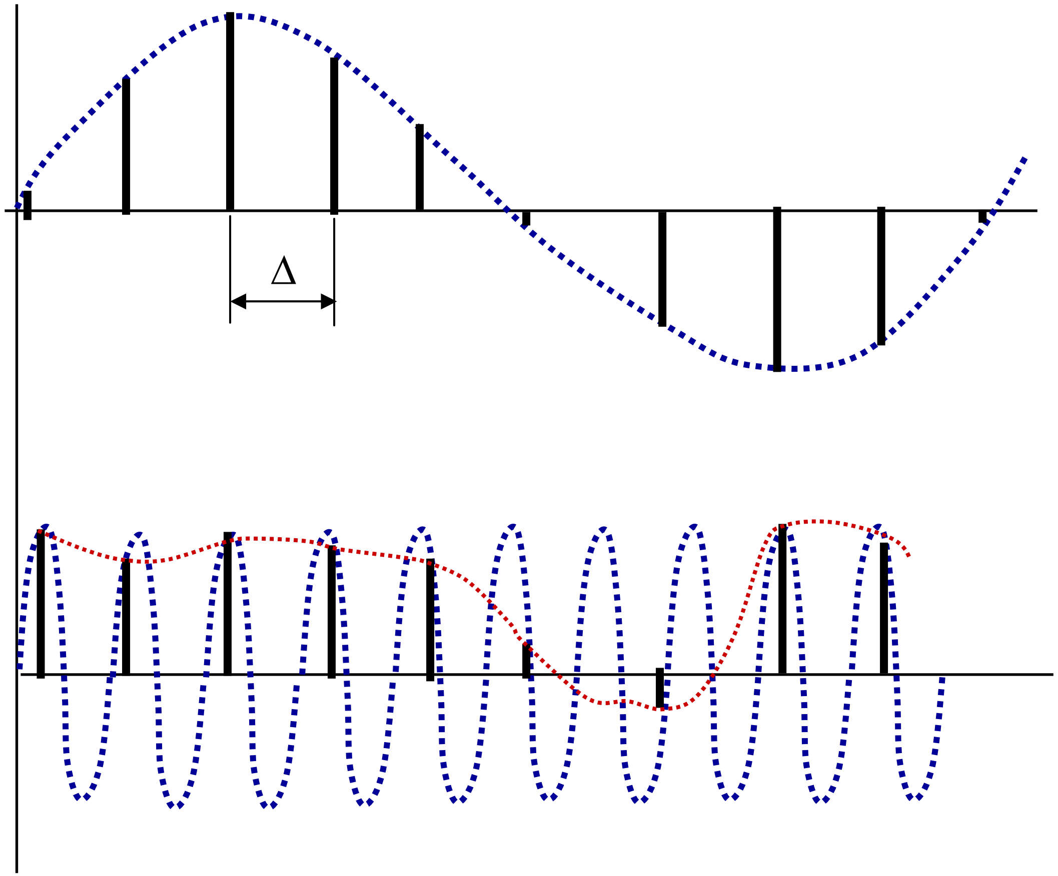
Quantised samples

$s_q(i) = \mathrm{int}(s(i)/q)$



Aliasing

A continues signal can be perfectly reconstructed from its samples, providing the sampling rate is correct



If the sampling interval Δ is not at least half of the period of highest frequency F_{max}, correct reconstruction is not possible

The Sampling Theorem:

$\Delta \leq 1/(2*F_{max})$

Digitization of Images

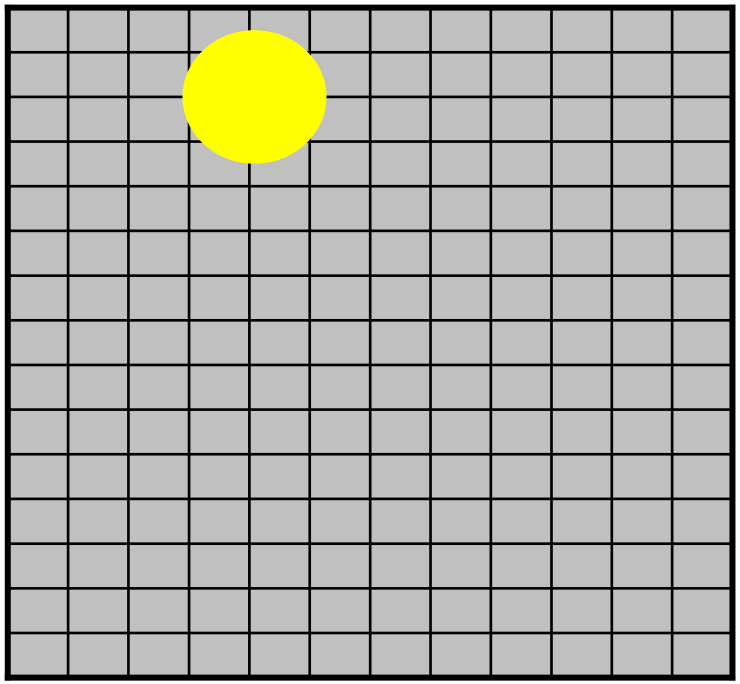
A properly sampled image

An under sampled image

Quiz

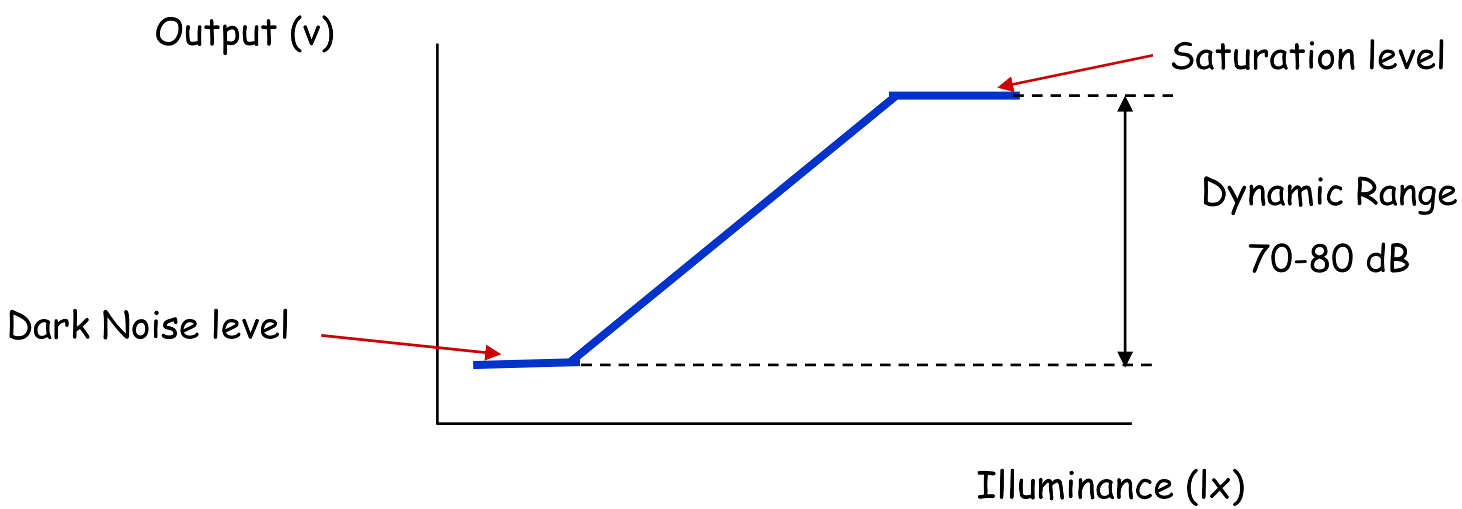
The image sensor is an array of photosensitive elements 5x5μm each. What should be the smallest size of the ray of light produced by a lens to avoid aliasing?

All digital cameras have a birefringent crystal plate between a lens and the sensor to restrict the smallest size of the ray and therefore avoid aliasing



Dynamic Range

- Accumulated charge: `q(i,j) = k*E(i,j)*t`



- Dynamic range $20*\log_{10}(V_{max}/V_{min})$

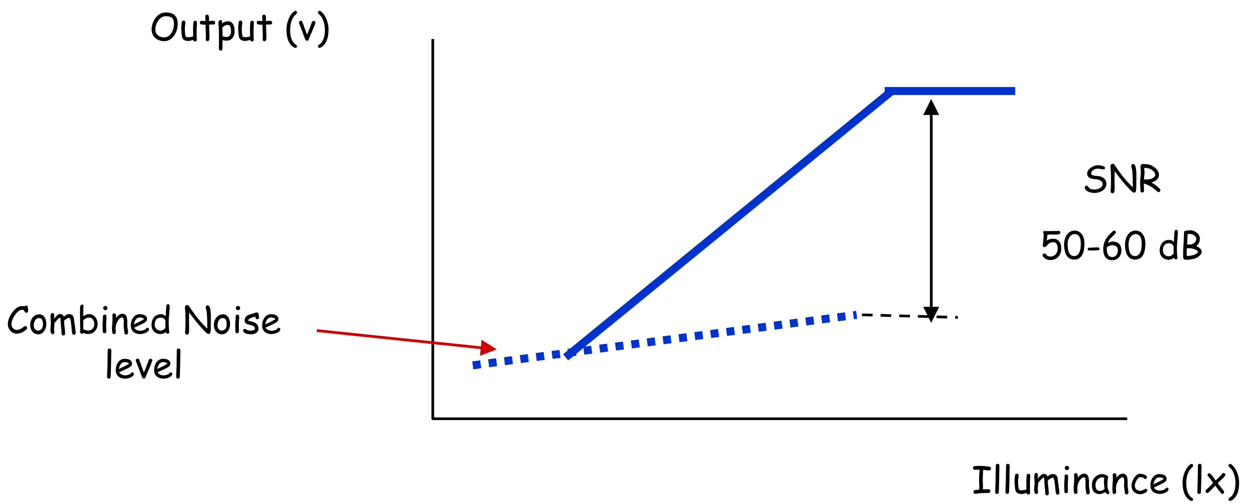
100 times = 40db

1000 times = 60db

10000 times = 80db

Signal-to-Noise Ratio (SNR)

- Besides the dark current noise, there are other sources of noise. Cumulative noise depends on Illuminance



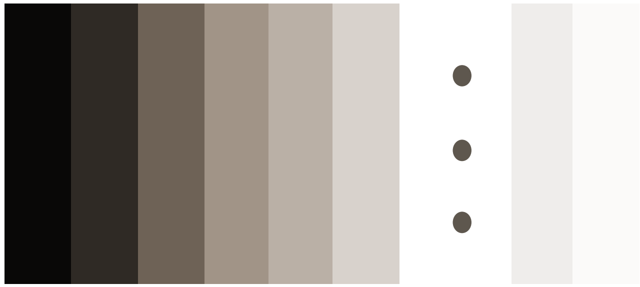
- SNR $20 * \log_{10}(V_{max}/V_{noise})$

What should be the precision of samples if SNR=60db ?

`60db = 1000 times => 10bit/sample`

Gamma Correction

- Each of $2^{10} = 1024$ digital levels represents a corresponding level of illuminance



0 - 1023

- According to Webber's low the noticeable contrast threshold is $k = \Delta I/I \approx 2\%$

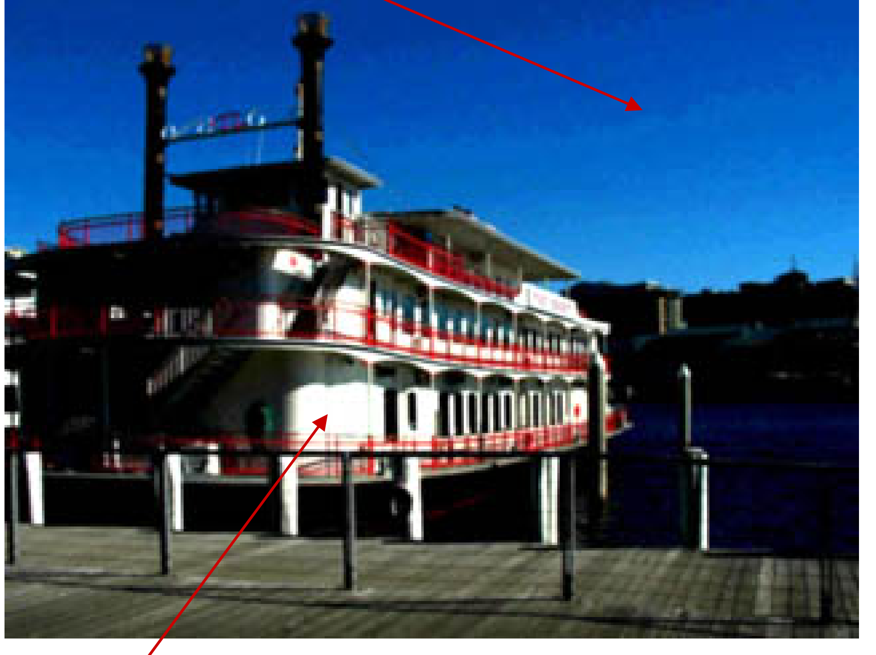
Example: $I_1 = 2$, $I_2 = 3$ => $k = \Delta I/I = 50\%$ over the threshold

$I_1 = 1000$, $I_2 = 1001$ => $k = \Delta I/I = 0.1\%$ below the threshold

1024 quantisation levels are not fully utilised

Gamma Correction

- If the contrast ratio is greater than 2%, smooth-shaded regions may contain false contours



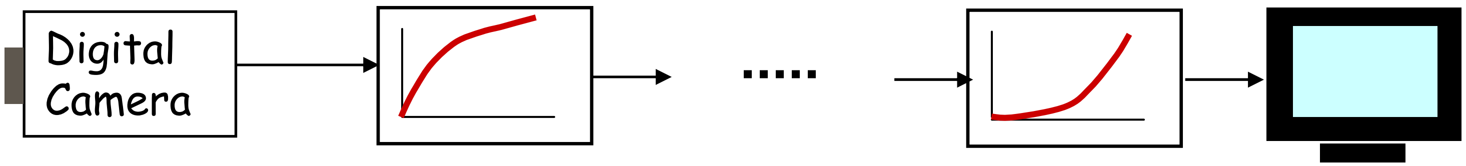
while regions with the contrast ratio less than 2% may contain redundant quantisation levels not distinguishable by human eye

Gamma Correction

- To map quantised samples into the perceptually uniform domain the samples are Gamma corrected

  $s' = s^{\gamma}$ with $\gamma = 0.4 - 0.5$

- Experiments indicate that after gamma correction 8bit/sample precision provides perceptually uniform quantisation



Gamma Correction

$s'=s^{0.45}$

Inverse Correction

$s = s'^{2.2}$

- To restore linearity, an inverse Gamma Correction is applied in all monitors and TVs

Gamma Correction

Example:

- A monochrome gamma corrected image stored with a different number of quantisation levels

2 levels 4 levels 256 levels

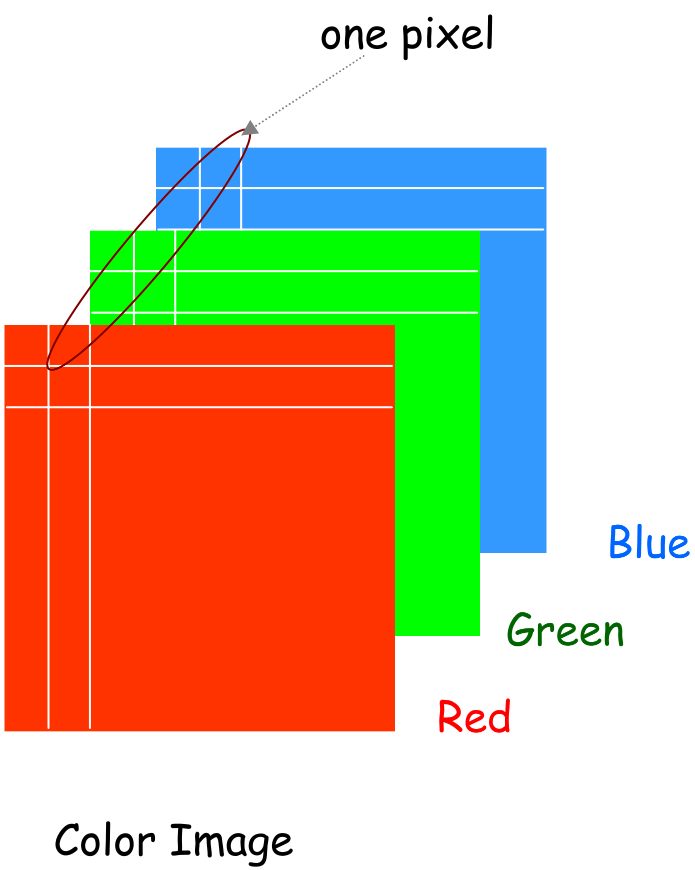
Software Implementation of Gamma Correction

- Real-time implementation of Gamma Correction using the equation $s' = s^{\gamma}$ is not efficient
- A commonly used approach is a Look Up Table

```
/* LUT with precomputed values for every 10-bit sample */
unsigned char LUT[1024] = {0, . . . , 255};
unsigned short sInput;         // 10-bit input sample
unsigned char aGcorrected;     //  8-bit corrected sample
.   .   .   .
   sInput = getNextSample();
   sGcorrected = LUT[sInput];    // gamma correction
   storeSample( sGcorrected );
.   .   .   .
```

Note: Gamma Correction is a processing stage used in all consumed digital cameras, but **it may not be helpful for computer vision applications**

Colour Images

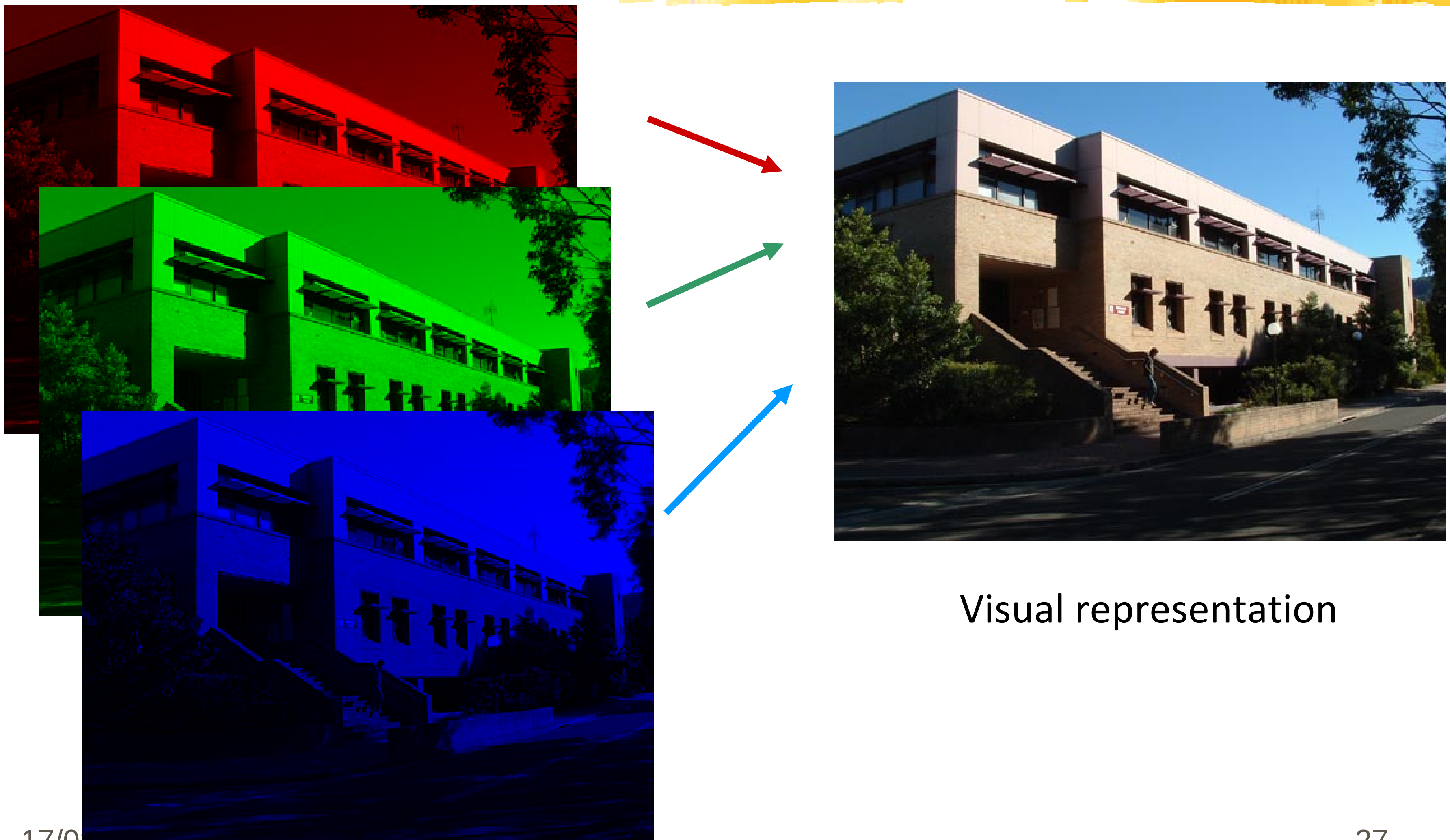


- To preserve information about colours, colour images must contain intensities for three colour channels for every picture element (pixel)

Explain why three colour channels are sufficient to represent real world images

- Each individual channel can be considered as a monochrom image captured in a limited band of spectrum (400-500nm, …, etc)

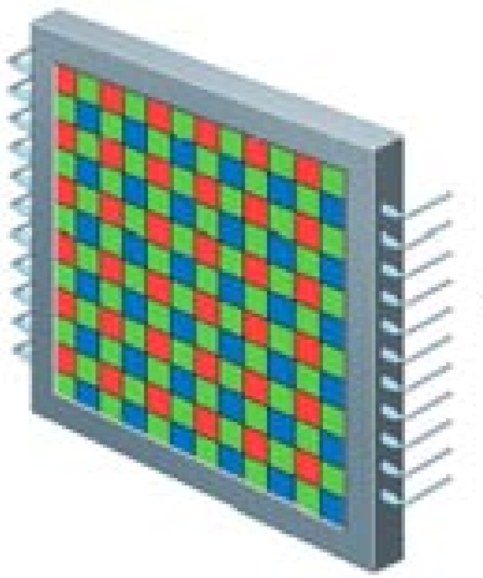
Colour Images



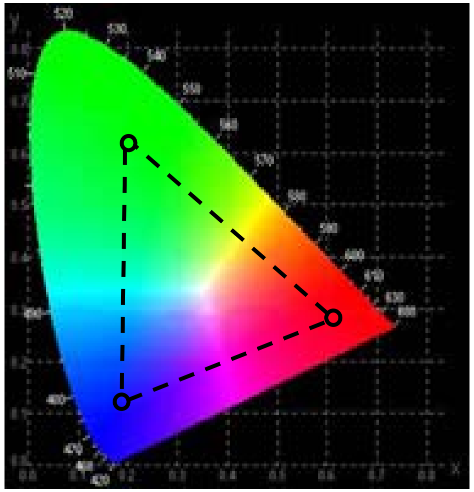
Visual representation

Colour Image Sensors

- Each photosensitive element is covered by a transparent colour filter that passes only one colour band

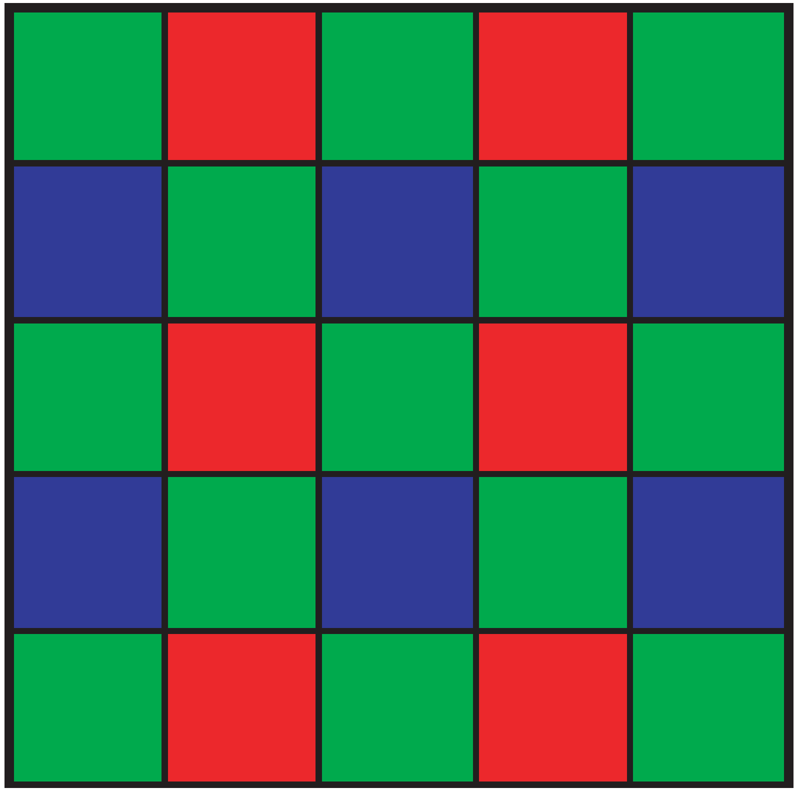


- Colour filter parameters determine the gamut of captured images



- The filters are arranged into a colour filter array (CFA) with the pattern that optimises efficiency

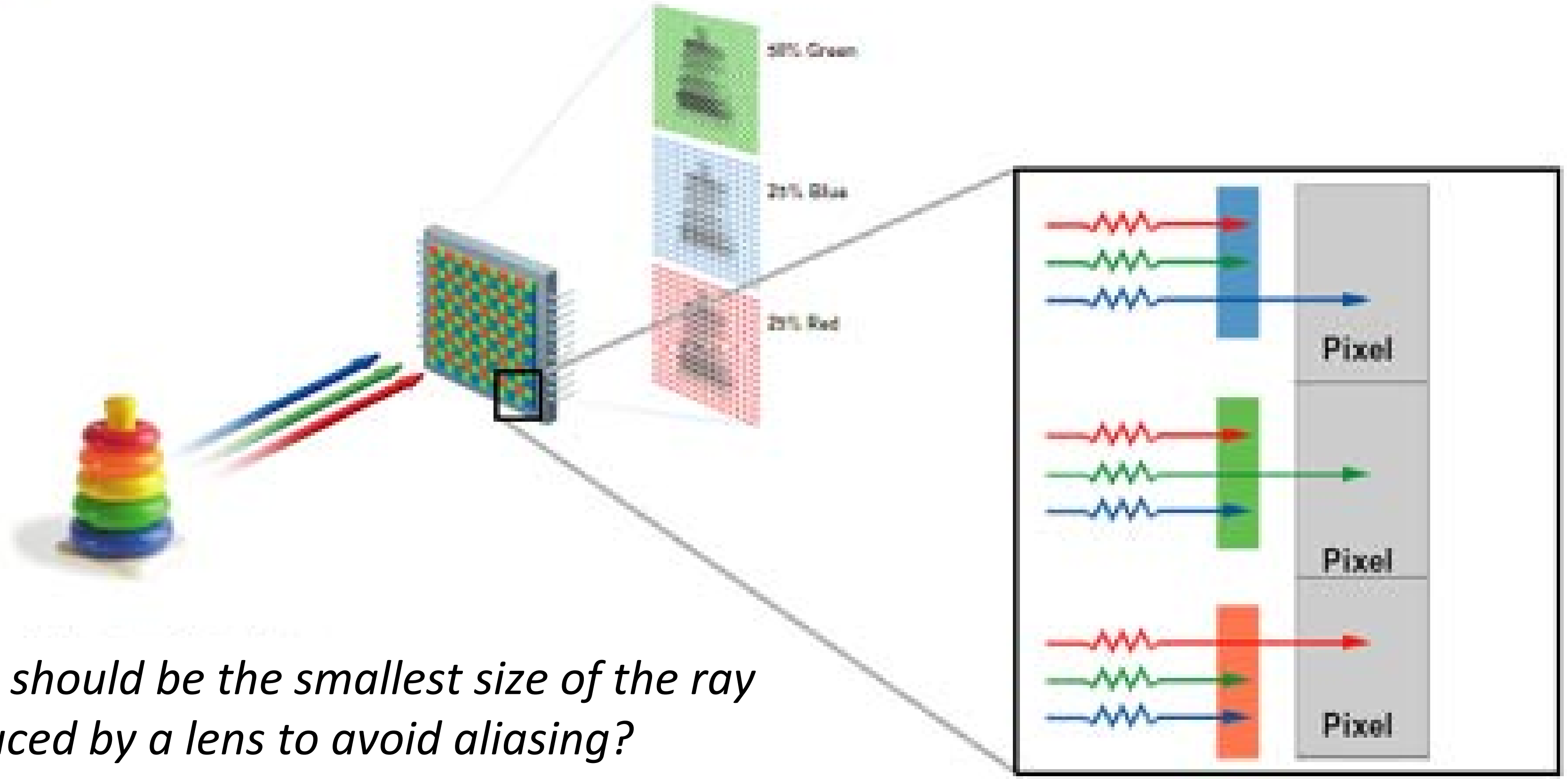
Bayer Pattern CFA



- Each photosensitive element captures only one colour band, thus other two has to be digitally reconstructed
- Filter parameters do not match those of human vision (the colour matching functions), thus colour reproduction has to be corrected
- Colour filters pass infrared that is not visible to humans, thus an additional anti-infrared filter is required
- Colour and anti-infrared filters reflect some light reducing quantum efficiency of sensors

Mosaic Filter: Each Pixel – 1 Color

Captures Only 1/3 of Light



What should be the smallest size of the ray produced by a lens to avoid aliasing?

Is it the same for all colour channels?

FOVEON

CFA Interpolation

- Reconstructs RGB colour samples for each pixel from Bayer colour samples

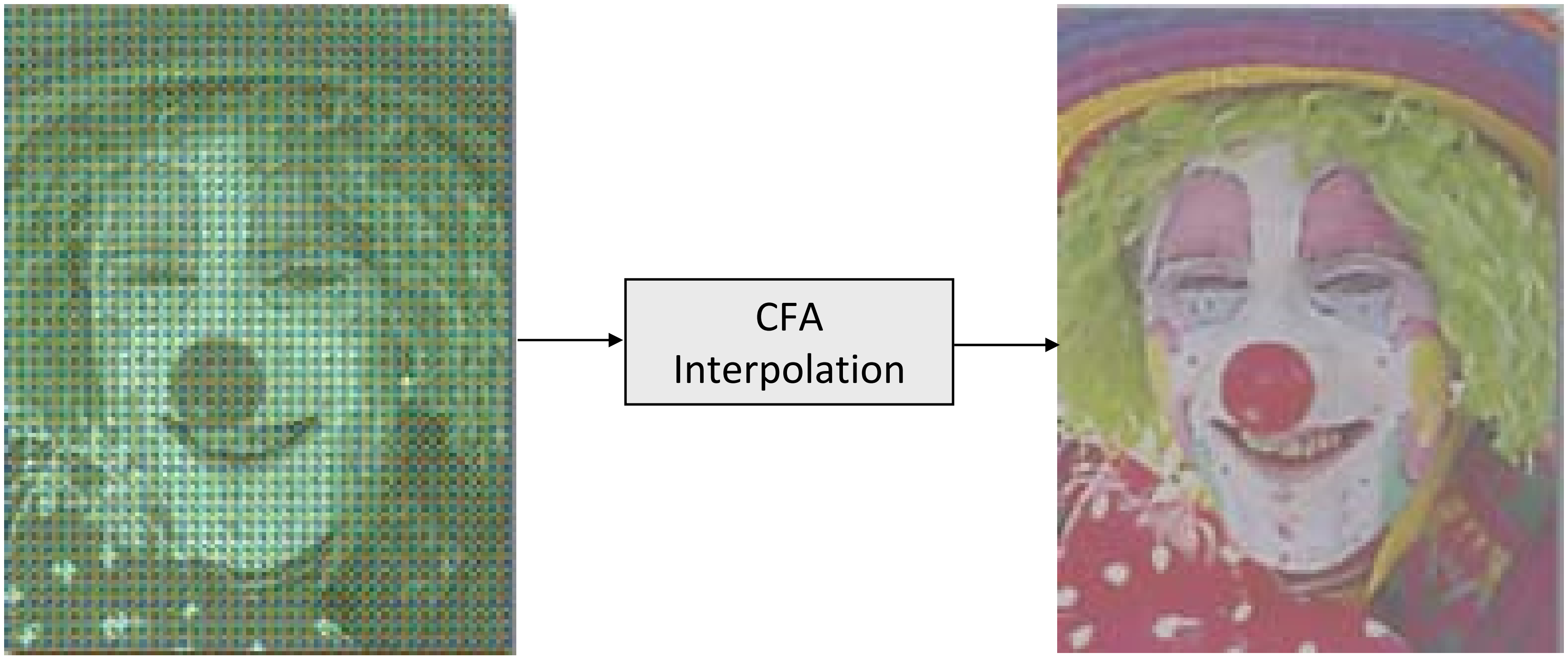


Image captured by a sensor

True colour image

CFA Interpolation

Nearest Neighbor

- Replicate the corresponding values of the nearest neighbor pixel

$R = R_8$

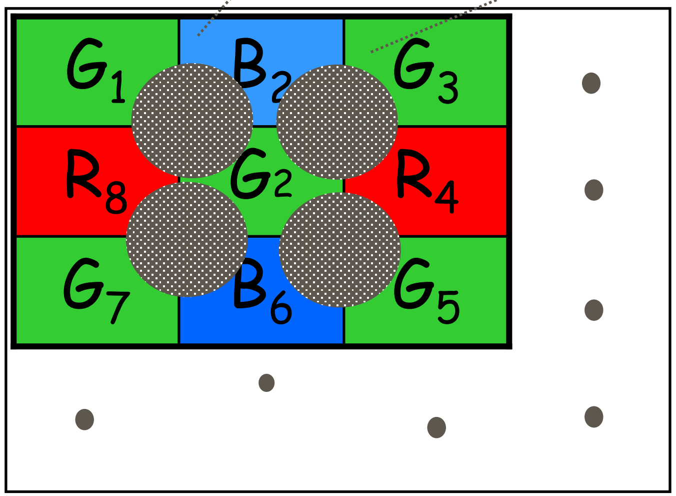
$B = B_2$

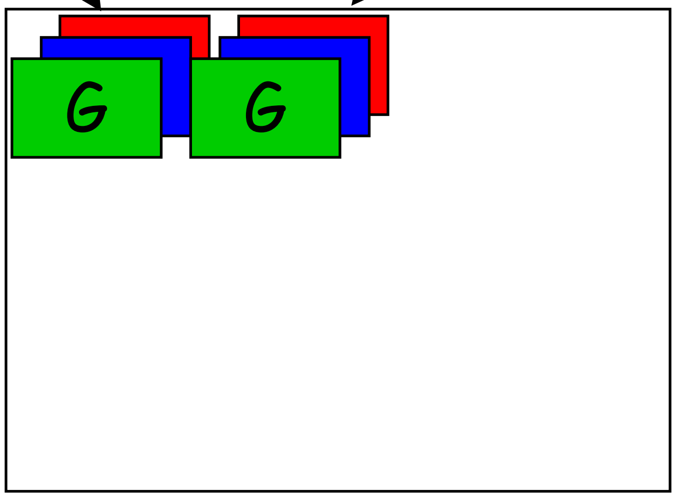
$G = (G_1 + G_2)/2$

$R = R_4$

$B = B_2$

$G = (G_2 + G_3)/2$





CFA Interpolation

Bilinear

- Interpolate missing colour samples as an average of the neighbor pixels

$G = G_2$

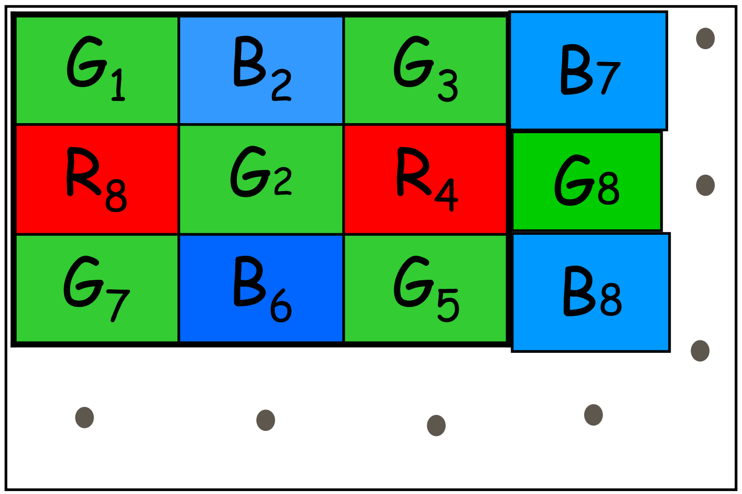
$B = (B_2 + B_6)/2$

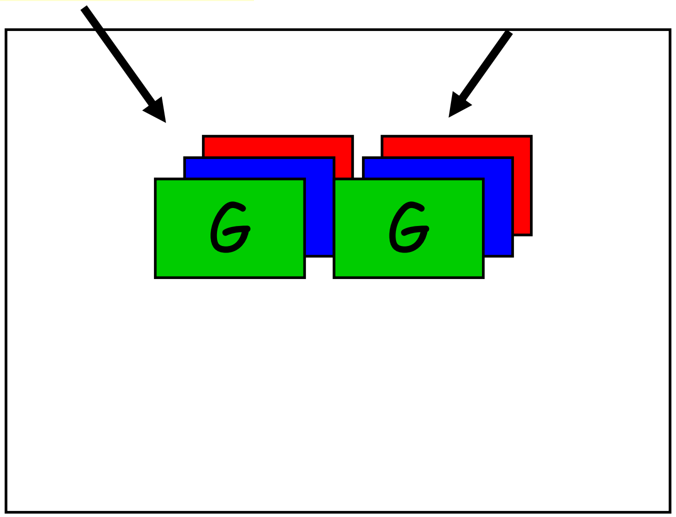
$R = (B_8 + B_4)/2$

$G = (G_2 + G_3 + G_5 + G_8)/4$

$R = R_4$

$B = (B_2 + B_6 + B_7 + B_8)/4$

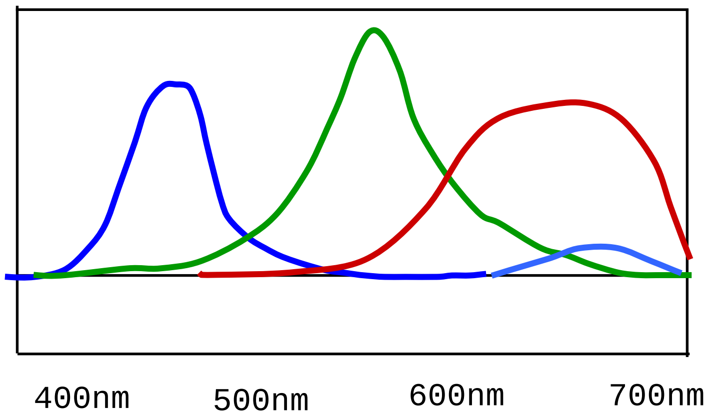




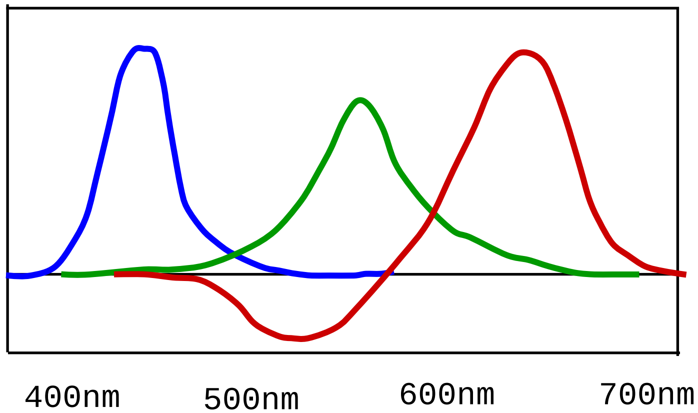
Colour Correction

- Spectral sensitivities of colour filters have significant deviation from CIE Colour Matching Functions

Typical filters of colour sensors



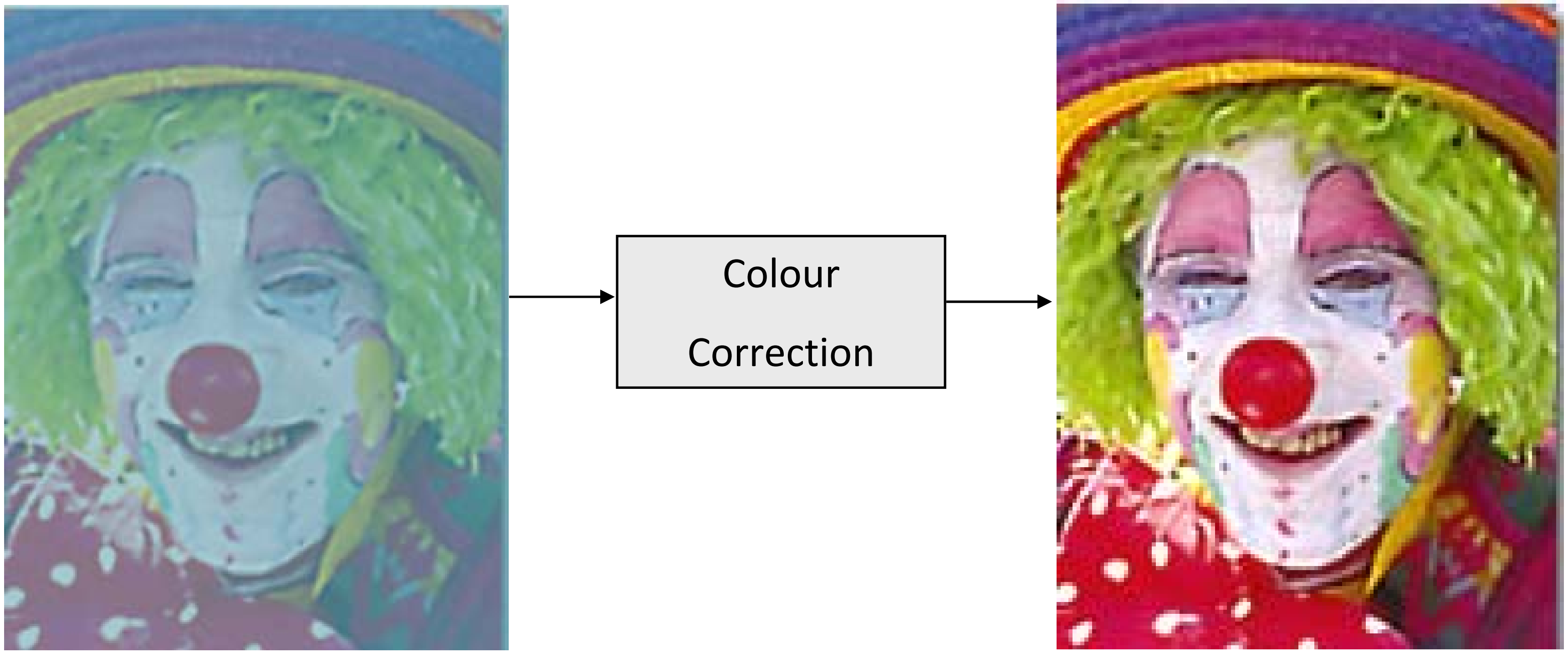
Colour Matching Functions



- Due to mismatch between colour sensitivities, CFA interpolated images have poor colour reproduction

Colour Correction

- Colour Correction digitally compensates deviation of spectral sensitivities of CFA to bring them closer to CIE Colour Matching Functions



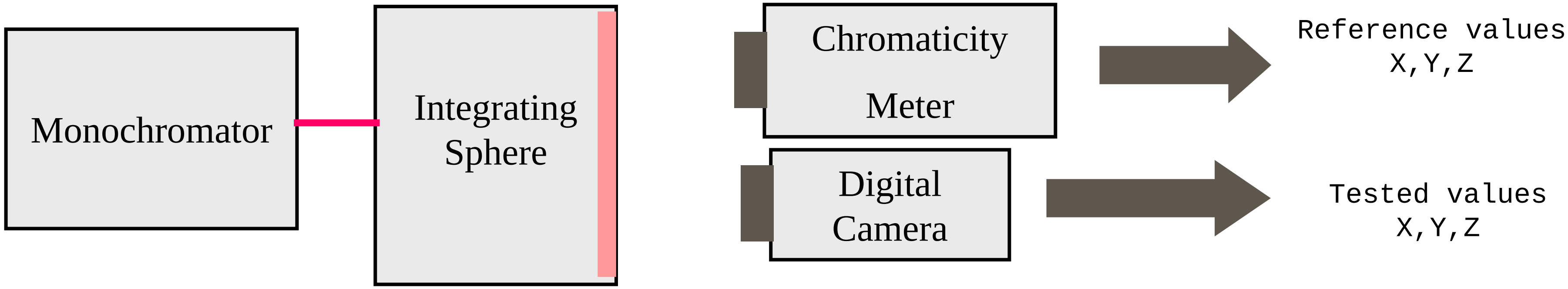
Colour Correction

- Transforms RGB values produced at CFA Interpolation stage into colour corrected RGB values

$$\begin{bmatrix} Rc \\ Gc \\ Bc \end{bmatrix} = T_C \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- T_C is a 3 x 3 color correction matrix, obtained through optimisation to maximise colour fidelity
- Each colour correction matrix is device dependent and may not be optimal for other sensors

Colour Calibration Equipment



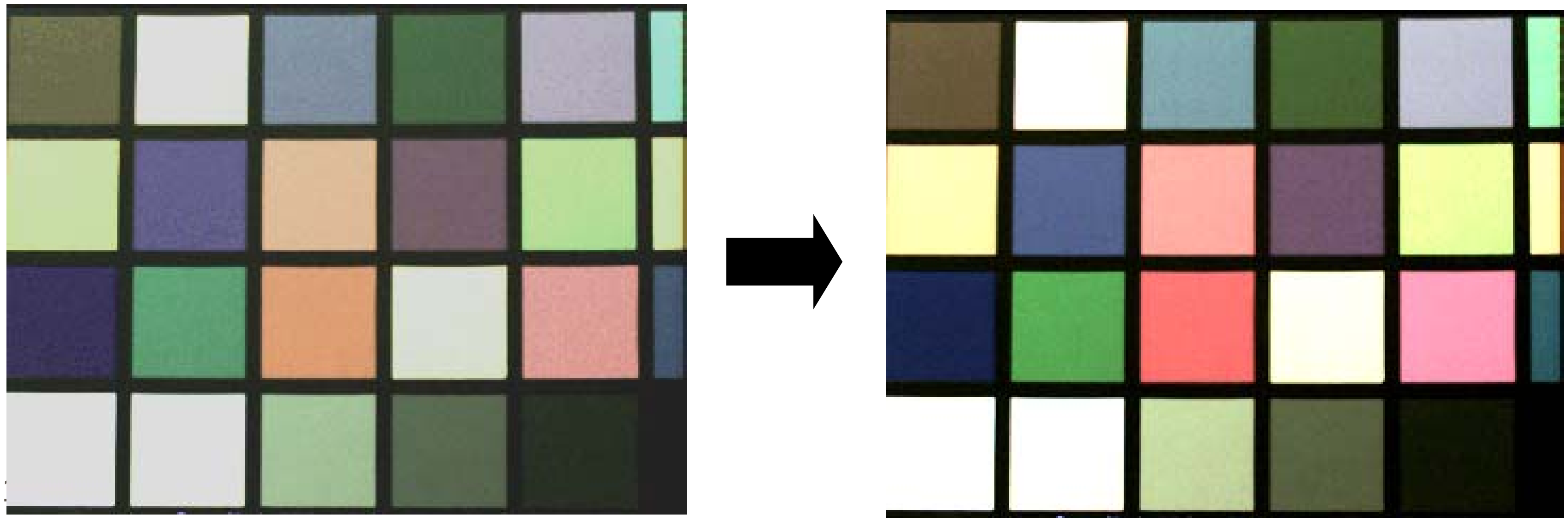
- A monochromator creates a narrow spectrum light ray
- An integrating sphere creates a uniformly illuminated surface from the ray
- A chromaticity meter produces CIE chromaticity values that are used as a reference for camera colour calibration
- The colour correction matrix is optimised to minimise the error between the sets of reference values and tested values

$$Er = \sqrt{\frac{1}{n}\sum\left((X_r - X_t)^2 + (Y_r - Y_t)^2 + (Z_r - Z_t)^2\right)}$$

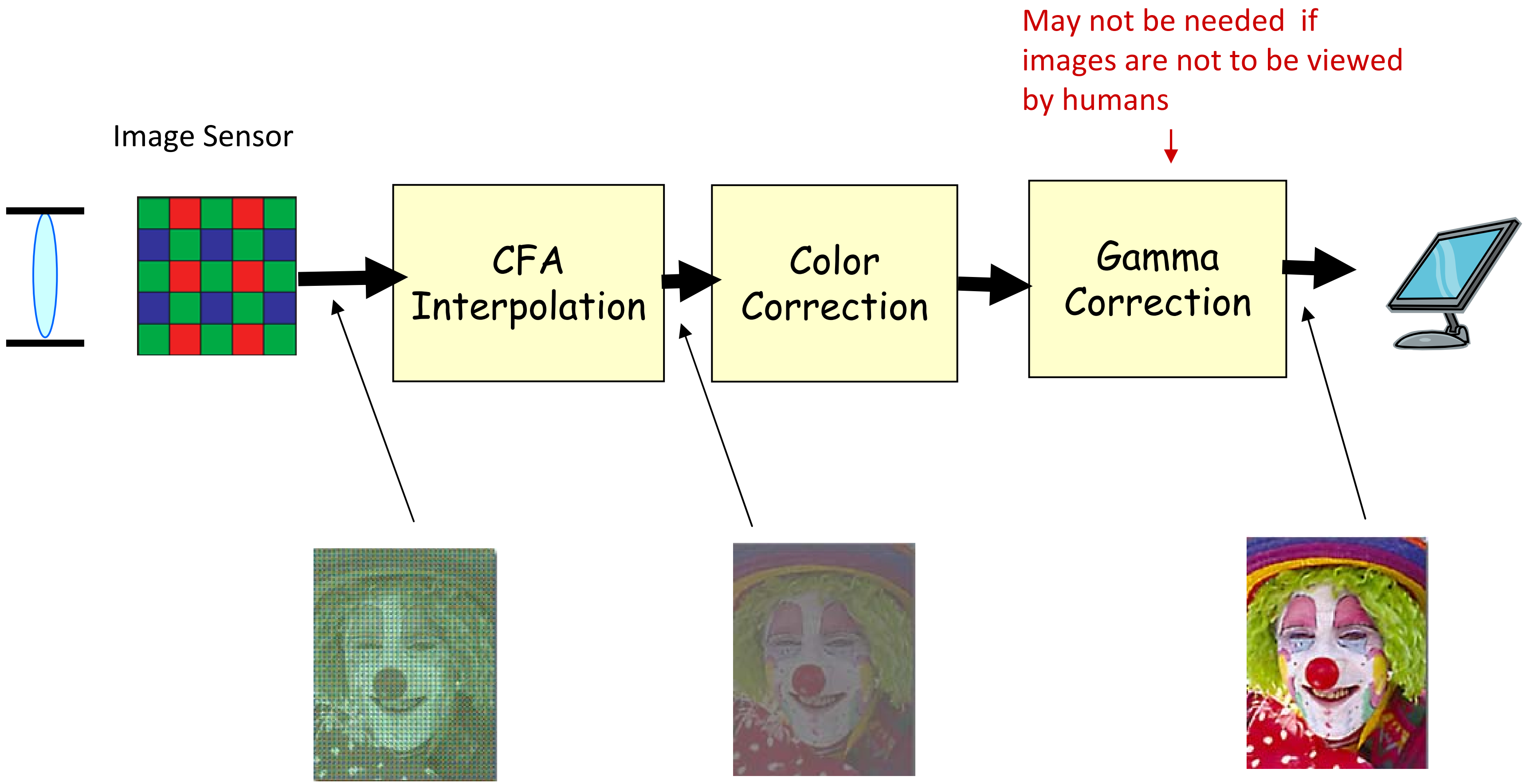
Colour Calibration using Macbeth Colour Checker

- If colour calibration equipment is not available, Macbeth Colour Checker illuminated by D65 standard light source can be utilised
- An example of Colour Correction Matrix optimised for a CMOS sensor using 24 reference colours

$$\begin{bmatrix} Rc \\ Gc \\ Bc \end{bmatrix} = \begin{bmatrix} 1.28 & -0.11 & -0.12 \\ -0.12 & 1.41 & -0.13 \\ -0.14 & -0.15 & 1.36 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

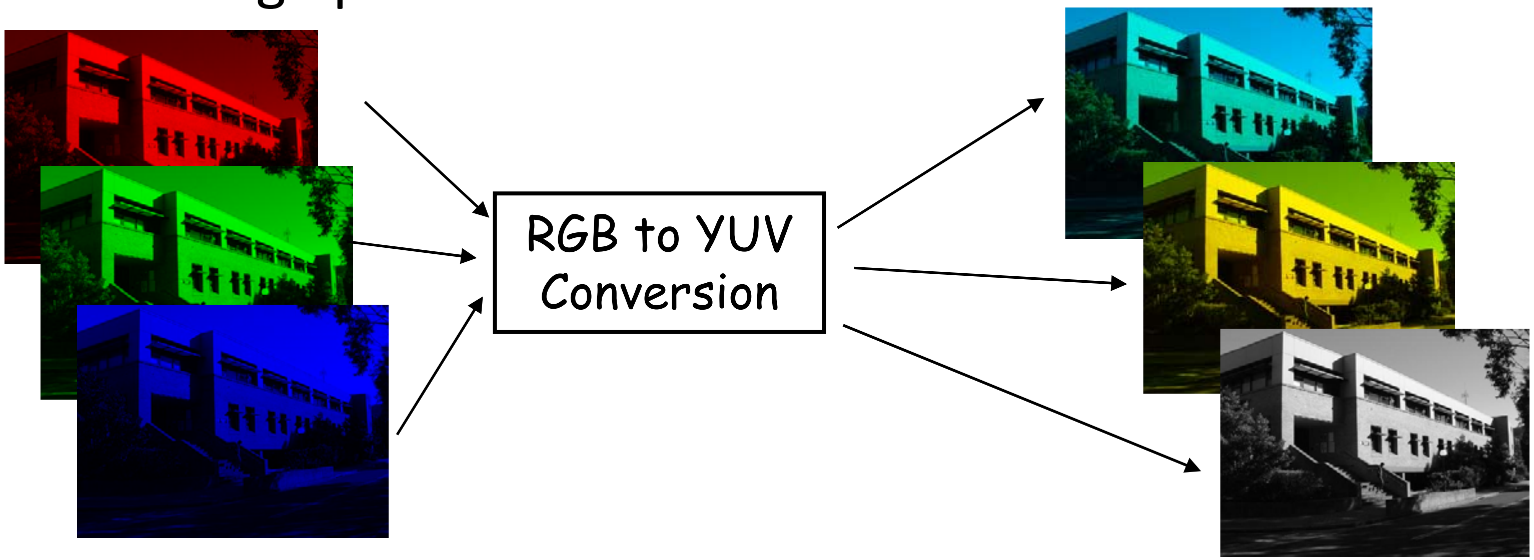


Colour Image Processing Chain



YUV Colour Space

- The human vision is more sensitive to variation in perceived image brightness rather than to variation of any colour component separately
- Some operations can be more efficiently performed in YUV colour space where brightness and colour are stored in different image planes



YCbCr Colour Space (ITU.BT-601)

- RGB to YCbCr colour space conversion is a matrix operation

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.29 & 0.59 & 0.114 \\ -0.17 & -0.33 & 0.5 \\ 0.5 & -0.42 & -0.08 \end{bmatrix} * \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

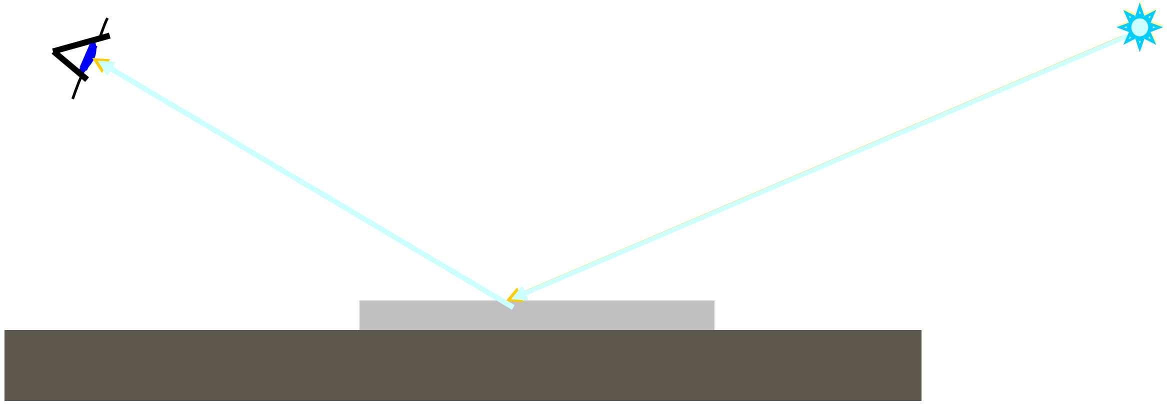
Gamma corrected true colour values

- YCbCr image can be converted back to RGB

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 1.4 \\ 1.0 & -0.34 & -0.71 \\ 1.0 & 1.77 & 0.0 \end{bmatrix} * \begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix}$$

ITU-R Rec. 709 and ITU-R Rec. 2020 define their own conversion matrices

Colour Constancy (review)



$$r_1 = \int S_1(\lambda)\overline{r}(\lambda)d\lambda$$
$$g_1 = \int S_1(\lambda)\overline{g}(\lambda)d\lambda$$
$$b_1 = \int S_1(\lambda)\overline{b}(\lambda)d\lambda$$

$$r_2 = \int S_2(\lambda)\overline{r}(\lambda)d\lambda$$
$$g_2 = \int S_2(\lambda)\overline{g}(\lambda)d\lambda$$
$$b_2 = \int S_2(\lambda)\overline{b}(\lambda)d\lambda$$

According to the formulas, if two sources of light have different spectrums, a white object should change its colour

However, the **perceived** color of objects remains relatively constant under varying illumination conditions

The mechanism of adaptation of the human vision to illumination conditions is still not fully understood

White Balancing

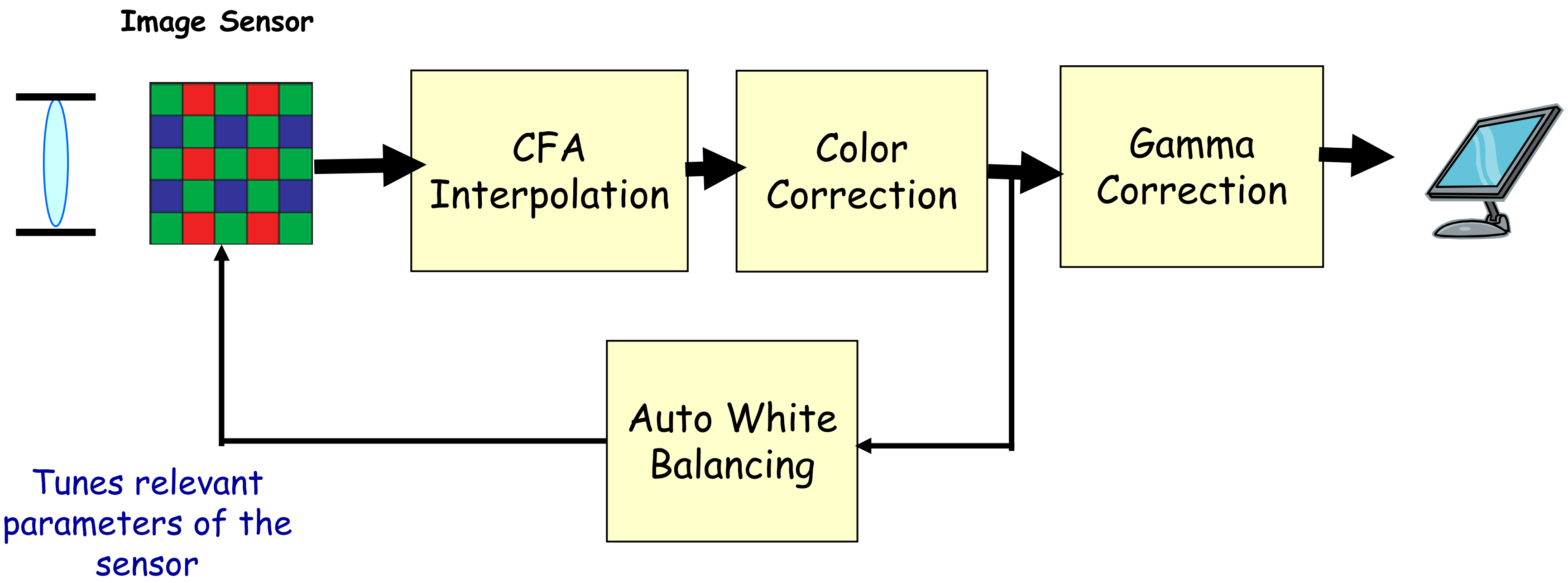
- If the same image is captured at different lighting conditions by a digital camera that does not provide colour constancy, they may have very different gamut

At midday

Before the sunset

Automatic White Balancing

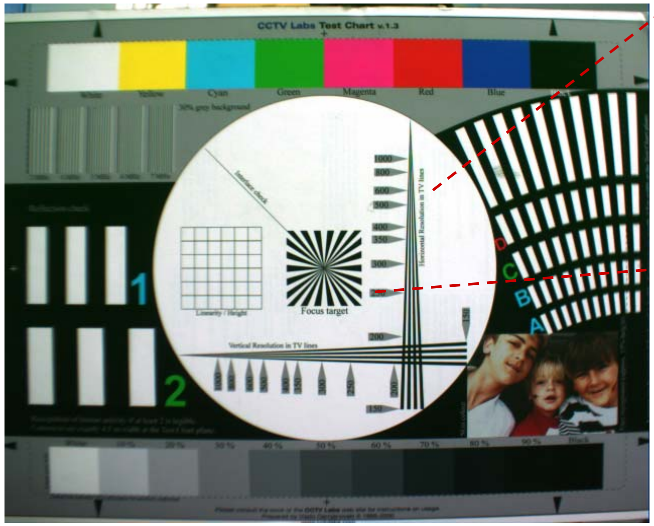
- Every digital camera is equipped with a processing module that provides colour constancy

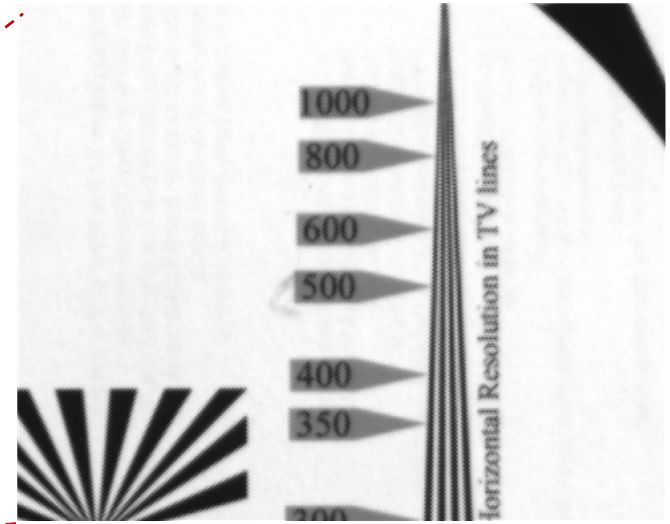


Resolution

- The amount of detail that the camera can capture
- Resolution depends on:

  - sensor resolution
  - optical resolution
  - sharpness of focus
  - CFA interpolation

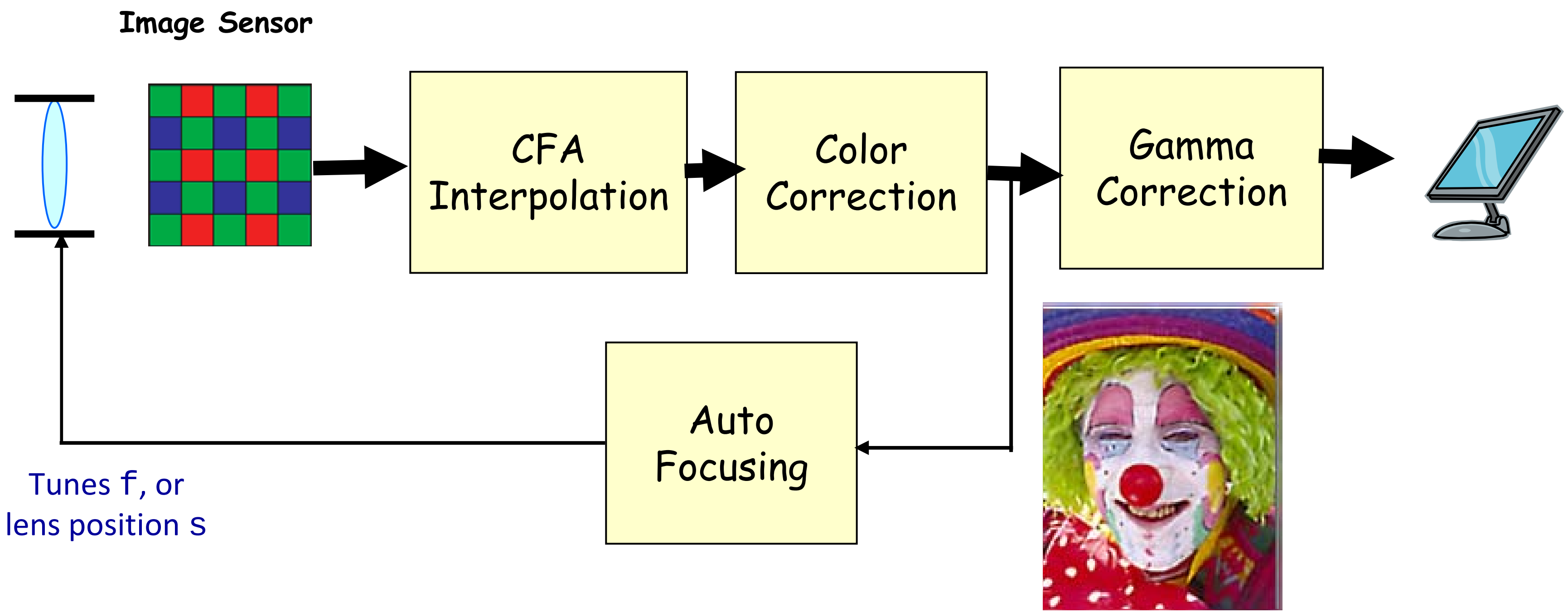




Test
Chart

Automatic Focusing

- Auto focusing tunes relevant parameters of the lens to maximise image sharpness



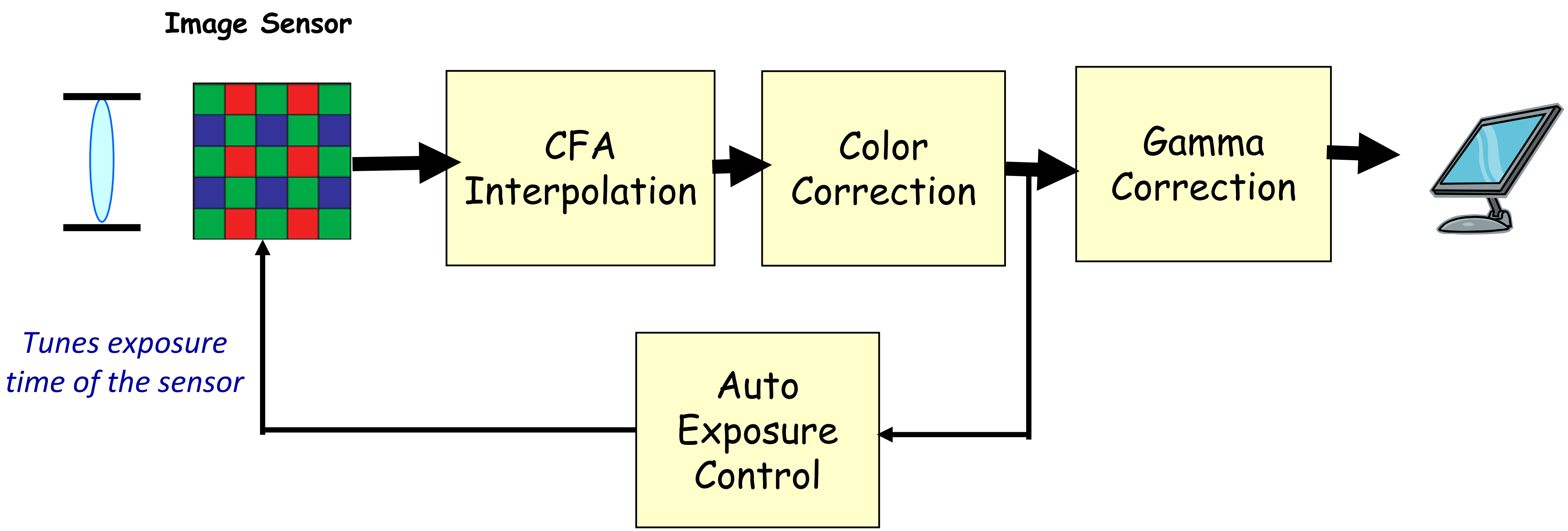
Automatic Exposure Control

- Every digital camera is equipped with a processing module that controls exposure time

  - dynamic range of sensors is 70-80db

  - dynamic range of Illuminance: 160 – 170db

  from 0.001lx (at night) . . . 100 000 lx (direct sun)



Trends of Imaging Technology

- High resolution
- Superior color fidelity
- High SNR
- High Frame Rate
- Low cost

Imaging Technology

DALSA Corporation
33 Megapixel Color CCD
(4992 x 6668, pixel size 7.2x7.2μm)

Traditionally, CCD sensors had better performance than CMOS:

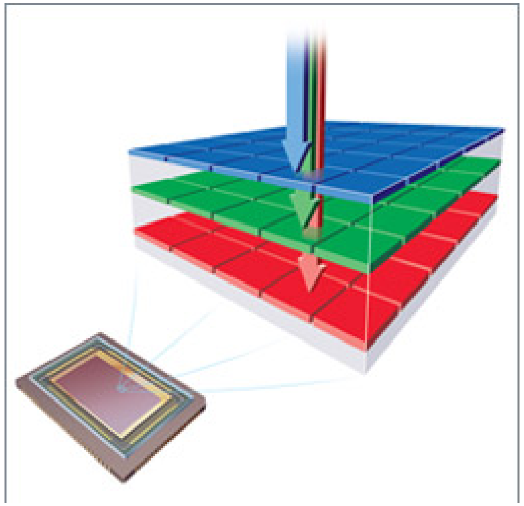
- They are very sensitive due to pixels with a better fill factor (higher dynamic range)
- They use full-frame red out and as a result do not introduce jelly-like geometrical distortions for fast moving objects

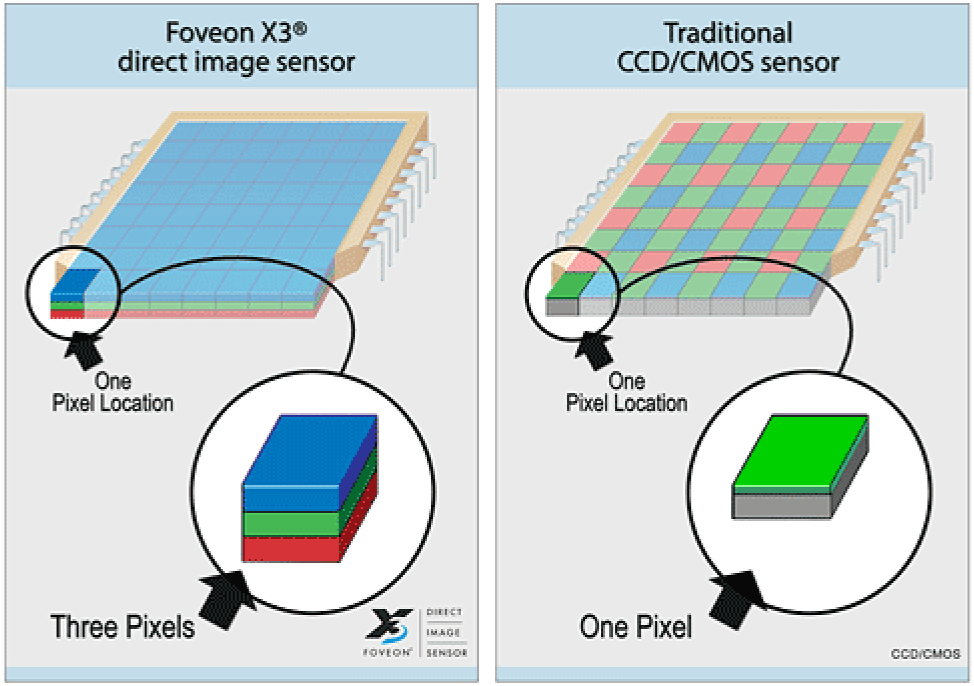
Imaging Technology

Rolling Shutter Distortions produced by CMOS sensors

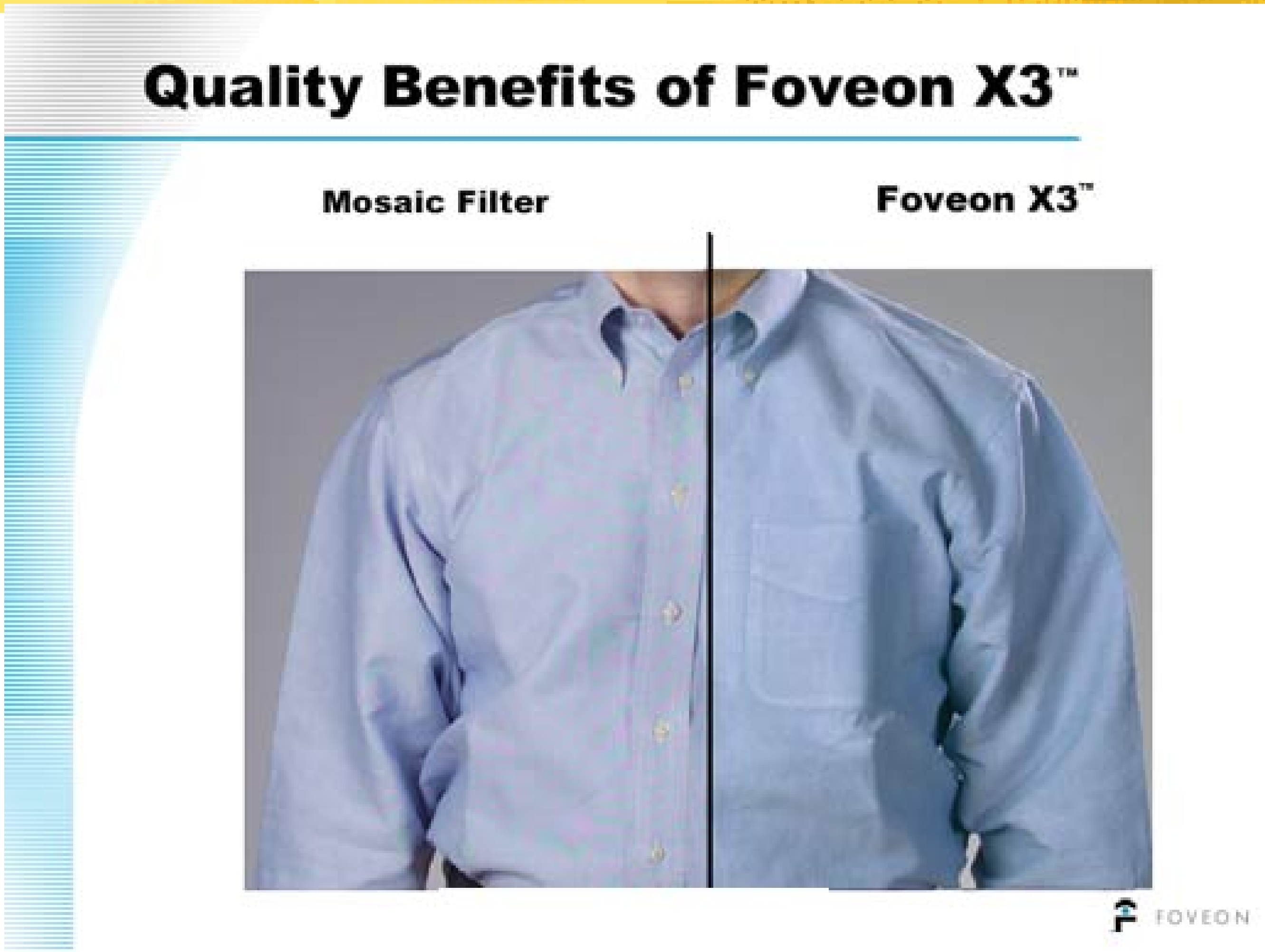
Imaging technology

Foveon X3 CMOS sensors with vertically stacked pixels



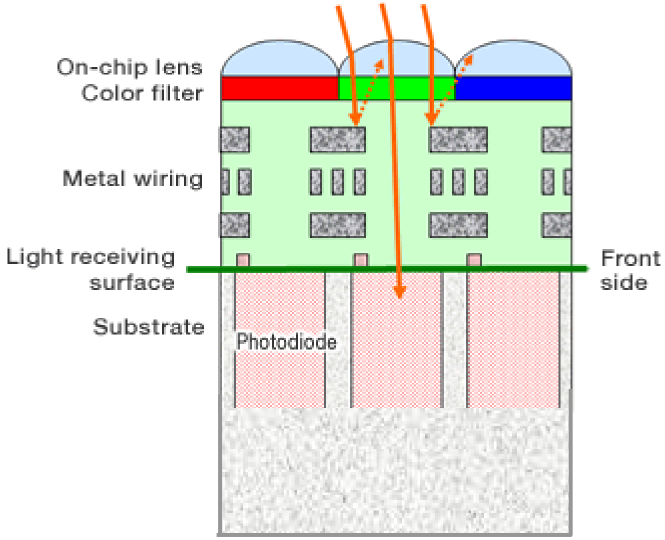


Imaging technology



Imaging technology

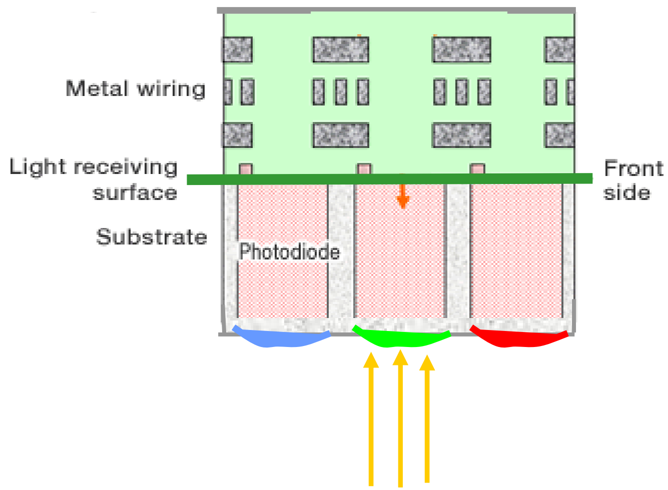
Conventional front illuminated sensors
(fill factor 30-50%)



Back Illuminated CMOS sensors

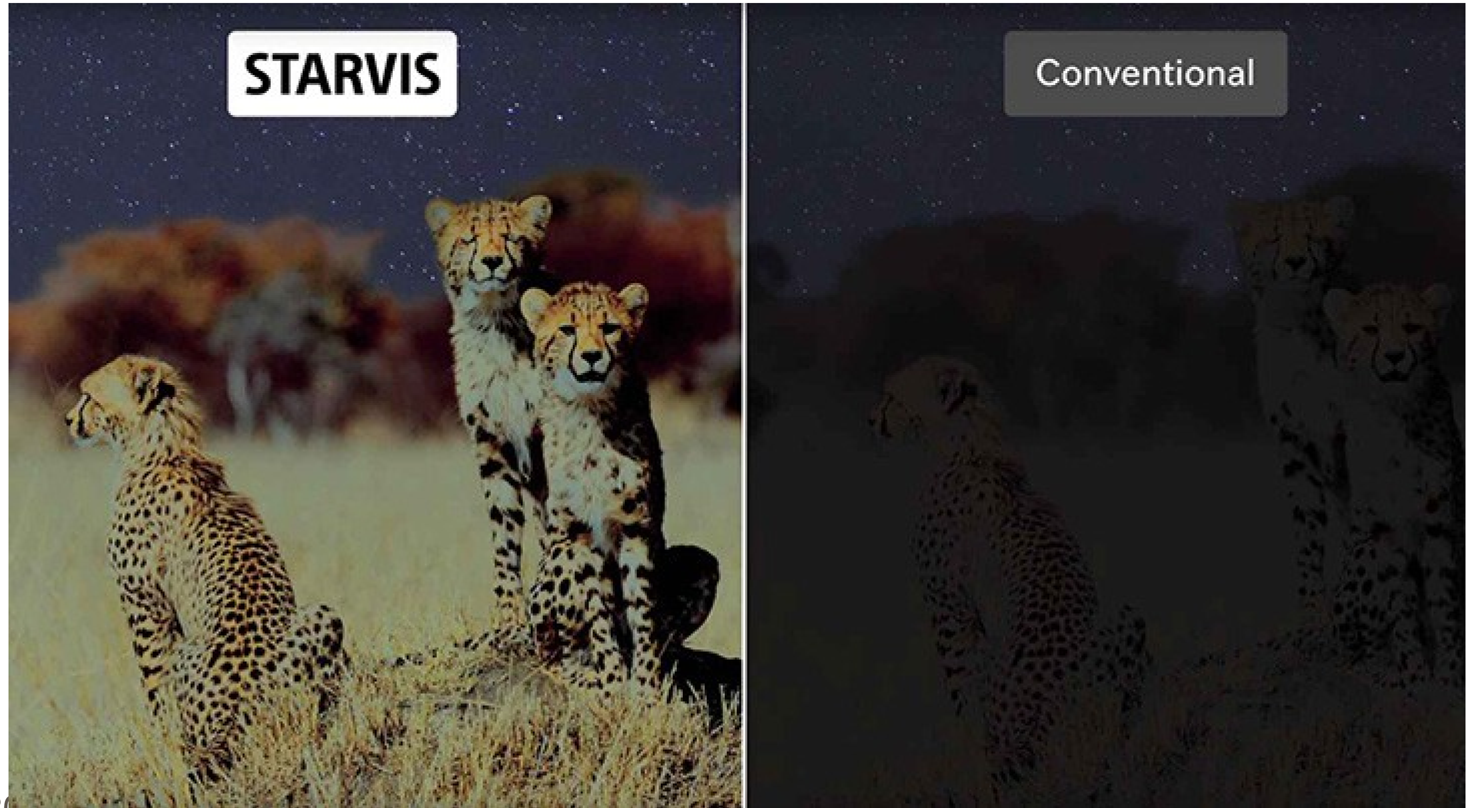
- Sony Exmor R
- Sony STARVIS

(fill factor 80-90%)



Imaging technology

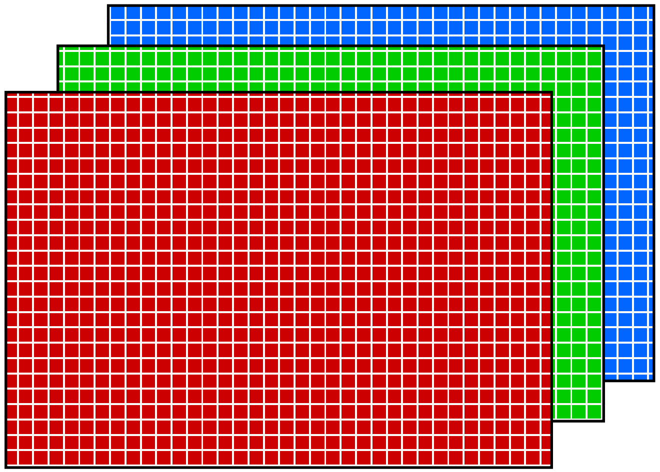
Back Illuminated CMOS sensors



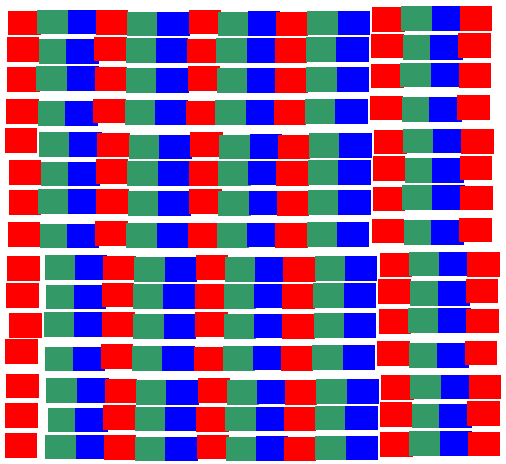
Storage of Images

How to store a 2D array of picture elements into a file that has a linear structure?

- **Picture Element:** Pixel
  - True Colour Image - 3 components/pixel , 8-bit/component
  - Gray-Scale Images - 1 component/pixel , 8-bit/component
  - Binary Images - 1 component/pixel , 1-bit/component



non-interleaved



interleaved

BMP Image File Format

BMP files contain both structured and block data

File Header

Image Header

Colour Table

Image Data

```
struct HEADER
{
  char signature[2];      /* file identifier  'BM' */
  unsigned int size;          /* file size */
  unsigned int unused;
  unsigned int imageDataOffs; /* image data start */
};
```

```
unsigned char blue1, red1, green1;
unsigned char blue2, red2, green2;
unsigned char blue3, red3, green3;
unsigned char blue4, red4, green4;
.   .   .   .   .   .   .   .   .   .   .
```

Standard Raster Scan and BMP Image File Format

Raster scan order

BMP

pixel[0][0]

pixel[height-1][width-1]

Suggested Reading

- D Forsyth, *Computer Vision. A Modern Approach*

  Chapters
  - 3.1 The Physics of Colour
  - 3.2 Human Colour perception
  - 3.3 Representing colours
  - 7.4 Sampling and Aliasing

- G. Bradski, A. Kaehler, *Learning OpenCV*

  Chapter 11
  - Camera Model
  - Basic projective geometry
  - Lens distortions