The University Of Sheffield.

EEE6081 (EEE421)
Visual Information Engineering (VIE)

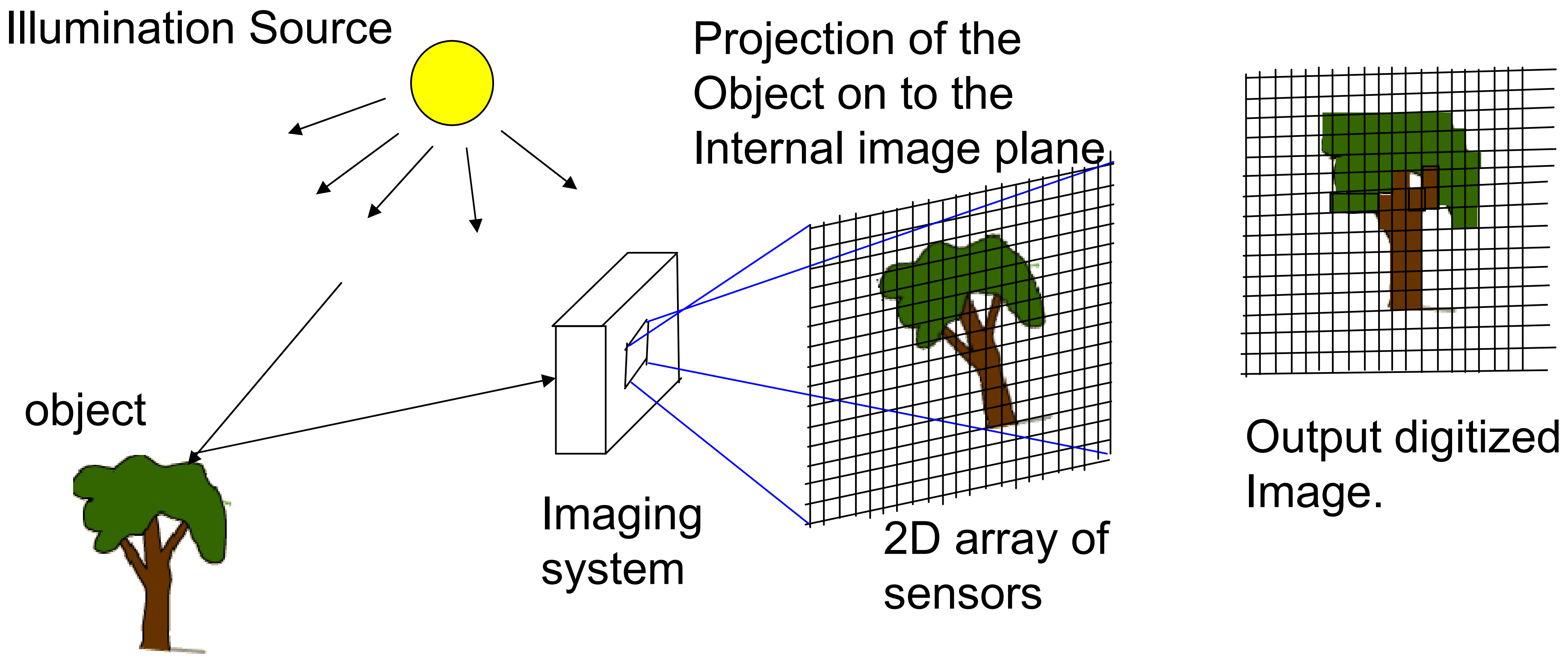
Topic 02: Revision – Background knowledge

- Part 1: Imaging Preliminaries
  - Image acquisition
  - Spatial resolution
  - Color depth resolution
  - Image data rate computation
  - Matlab

•Background reading: Digital Image Processing (Gonzalez / Woods) Chapters 1 and 2.

Dr Charith Abhayaratne
Email: c.abhayaratne@sheffield.ac.uk
Office: F176 Tel: 25893

Image Acquisition



- The image acquisition is performed by a 2D array of sensors. This array is coincident with the focal plane of the imaging system.
- At each sensor the incoming illumination energy is transformed into a voltage.
- The output voltage waveform is the response of sensors.
- This waveform is digitized to obtain the digital image.
- What are the two processes involved in digitizing?

Image Acquisition

- We denote images by 2D functions:
  - $f(x,y)$ - This means amplitude of f at coordinates x and y.
    - (what coordinate system?)
- When an image is generated from a physical process, its values are proportional to energy radiated by the physical source.
- Therefore we can say $f(x,y)$ must be non-zero and finite.
  - $0 < f(x,y) < \text{infinity}$
- $f(x,y)$ is characterized by two components:
  - 1. The amount of source illumination incident on the object – illumination $i(x,y)$
  - 2. The amount of illumination reflected by the object – reflectance $r(x,y)$
- Now we can define $f(x,y)=i(x,y)r(x,y)$
- Usually, $0 < i(x,y) < \text{infinity}$ (theoretical values)
- $0 <= r(x,y) <=1$ 0 is total absorption 1 is total reflectance. $r(x,y)$ is a characteristic of the imaged object.
- E.g., black velvet $r=0.01$ stainless steel $r=0.65$ snow $r=0.93$

Monochrome images

- The intensity of a monochrome image at any coordinate (x.y) is called the gray level (l)
  - we can write $l=f(x,y)$
- We can define the lower and upper bounds for l.
  - $L_{min} <= l <= L_{max}$
  - where $L_{min}=i_{min}r_{min}$ and $L_{max}=i_{max}r_{max}$
  - Typical values L_{min}~10 and L_{max}~1000
- The interval $[L_{min} , L_{max}]$ is called the gray scale.
- It is common to re-map this scale to [0, L-1]
- 0 corresponds to black
- L-1 corresponds to white
- What about all intermediate values in the scale?

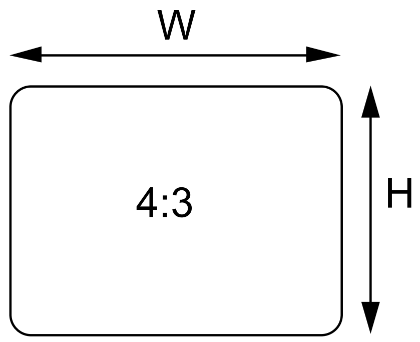
Image Representation

- To generate digital images from the sensed data (voltage waveform output)
  - We need to convert the continuous signal into digital form.
- Remember digitisation of data?
- Two processes:
  - 1. sampling
  - 2. quantisation
- In a 2D array of sensors, the spatial sampling is determined by the arrangement of sensors. (we will discuss this later)
- If there is no motion, the output of each sensor is quantised and gray level values are mapped into discrete levels.
- That means we can represent a digital image as a matrix. (An example of image using M x N array of sensors.

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}$$

What is image resolution?

- Describes how detail the image is
- Resolution types in video:
  - Spatial resolution
    - Usually expressed in terms of "pixels"
    - Pixels = Picture-elements (pels)
      - A tiny *dot* that makes up the representation of a picture in a computer's memory.
      - Expressed as W x H
      - Aspect ratio W : H = 4:3 (for standard display) 16:9 for wide screen display
  - Colour depth resolution
    - Defines how many different colours per pixel (C)
    - Usually expressed as how many bits per pixel (N)
  - Temporal resolution
    - Frame rate (frames per second)



Spatial Resolution

An example

The spatial resolution of an image is determined by sampling.

Spatial resolution defines the smallest details that can be discernible.

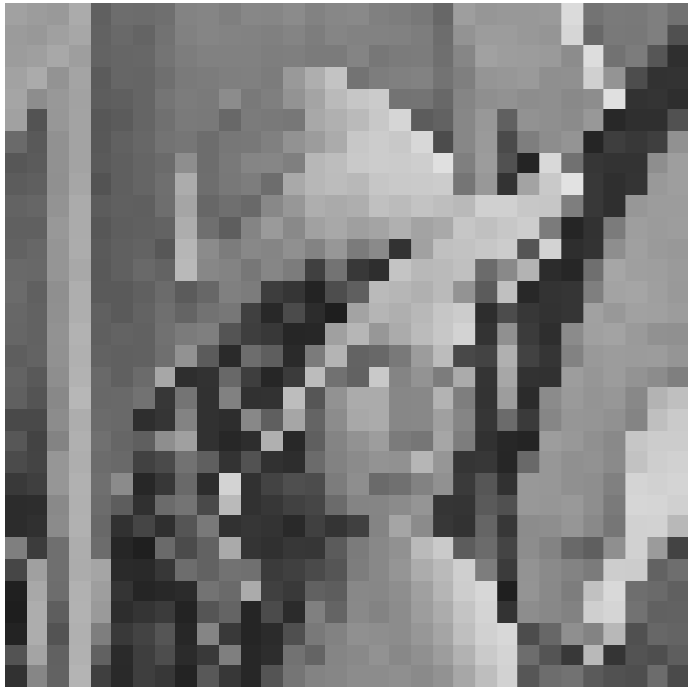
Reduction of spatial resolution (although data size can be reduced) results in loss of details.

256x256

128x128

64x64

32x32

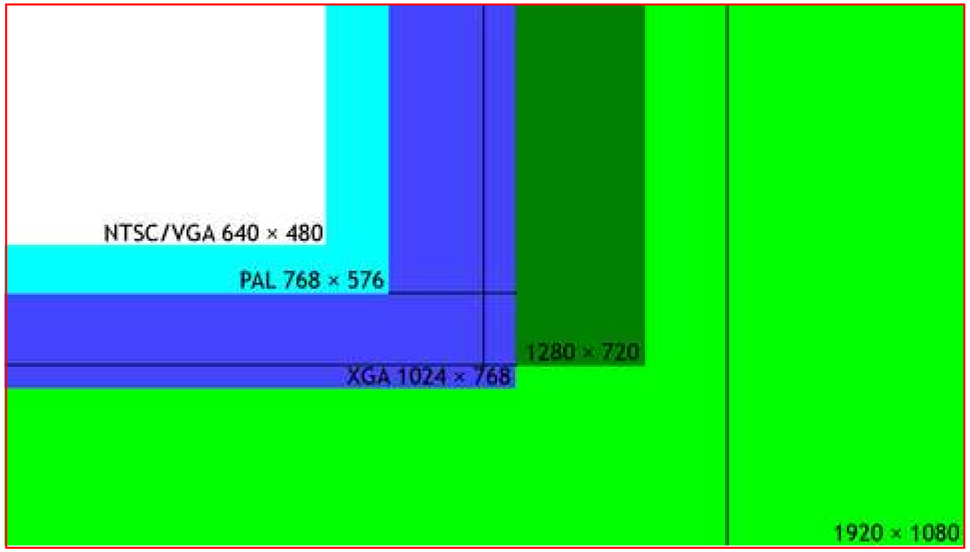


Current Spatial-temporal Resolution Examples

QQCIF	90 x 72	Thumbnail videos for browsing
QCIF	180 x 144	Mobile phone displays
CIF	360 x 288	VCD resolution
4CIF	720 x 576	DVD resolution
SDTV	576 x788 (576i50 (PAL, SÉCAM), 576p25)	
EDTV	576p50, 720i50, 720p24, 720p25	
HDTV	720p50, 1080p24, 1080p25, 1080i50	
UHDTV	4 x HDTV	

Compute Aspect Ratio

Are they 4:3 or 16:9?



p=progressive
(non-interlace)
i=interlace

Colour Depth Resolution

- We will first consider gray level resolution:
- Gray-level resolution = the smallest discernible change in gray level.
- We usually consider the number of discrete gray levels as an integer power of 2.

- Using N bits we can represent 2^N gray levels.
- We have gray levels ranging from 0 to L-1 where L-1 ~ 1000.
- Usually we use 8 to 10 bits per pixel for image representation.
- (Think why 8 bits are sufficient. 2^{10}=1024)

- How can you compute the total number of bits required to represent an N-bit image of resolution P x Q.?
-

- What happens when colour bit-depth resolution is reduced?
-

Gray-level Resolution example
Lena and Mandrill images

256 levels
8 bits

16 levels
4 bits

4 levels
2 bits

2 levels
1 bit

Gray-level Resolution example Lena and Mandrill images

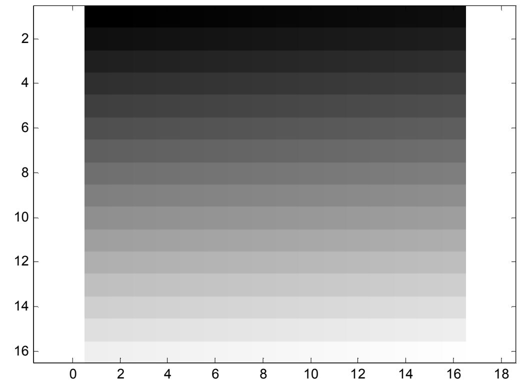
- Consider the 8-bit representation:
  - Which regions contain low spatial frequency components?
  - Which regions contain high spatial frequency components?
- Now consider low bit representations:
  - How are the high spatial frequency regions affected?
  - How are the low spatial frequency regions affected?

Colour Depth Resolution

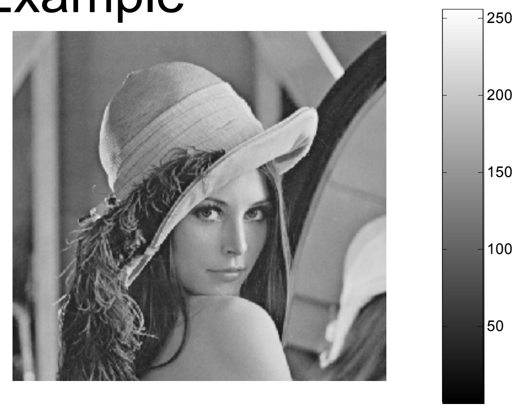
For gray scale: 1 byte per pixel = 8 bits per pixel (bpp)
This can represent $2^8 = 256$ different gray levels.

Numbers 0 – 255 are mapped to different gray shades 0= black 255=white

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111
112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255



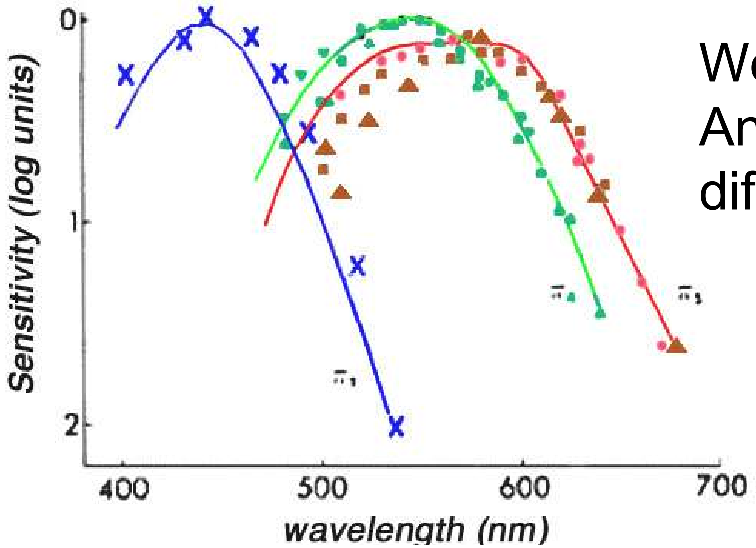
An Example



Digital images are represented as an matrix.
Manipulating numbers → image processing

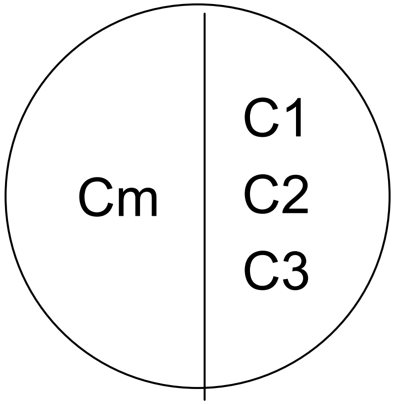
Colour depth Resolution

Tri-stimulus Theory of Colour



We possess three different colour-senstive cones. Any colour is a combination of the output of these different cones

A colour(Cm) = C1(blue) + C2(green) + C3(red)



Grassman's Law

Match to Cm by varying C1, etc.

MORE References:
CIE Chromaticity Diagram

A demo at http://www.cs.rit.edu/~ncs/color/a_chroma.html

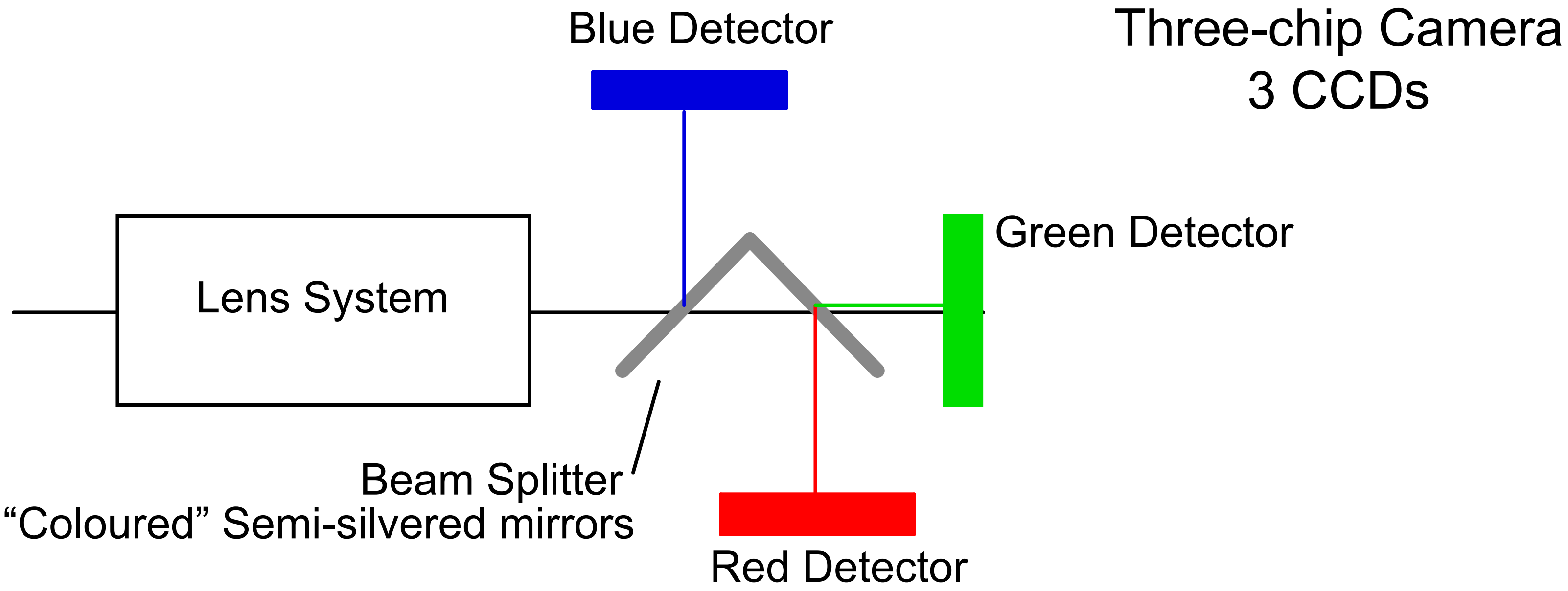
Original

Red

Green

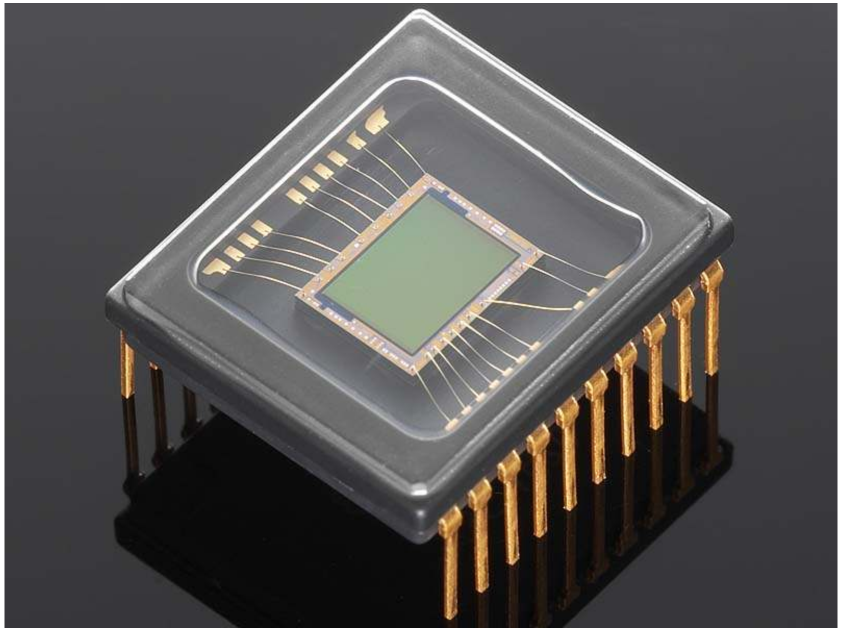
Blue

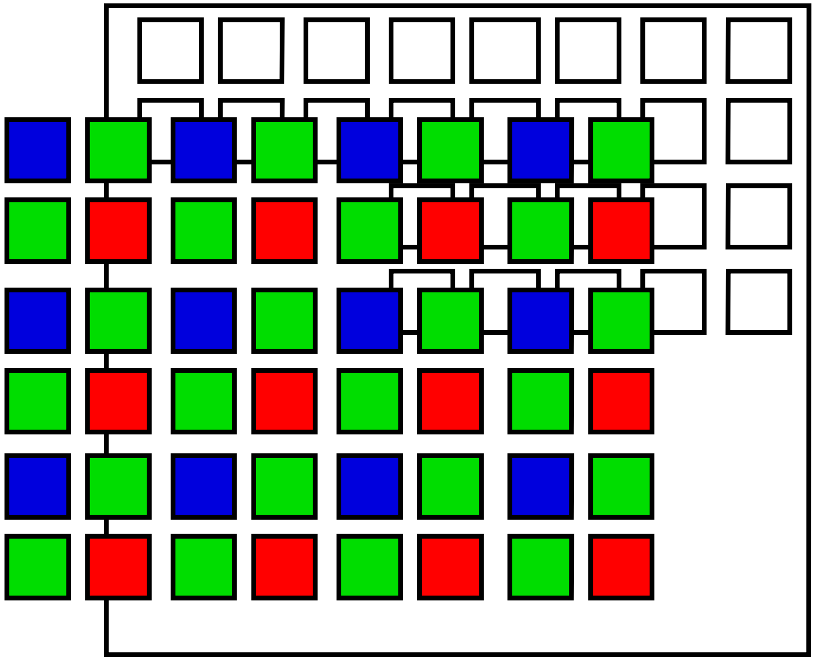
Elements of Camera



Charge-coupled device (CCD) - array of photo-sensitive pixels
Converts incident photons into detected electrons –> readout to produce varying voltage

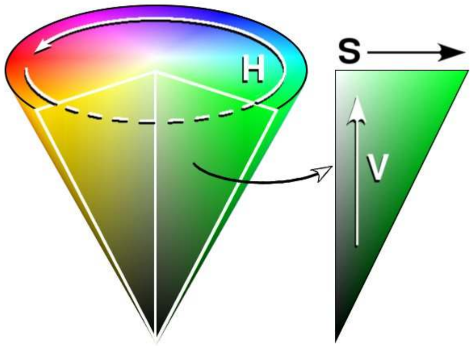
Charge-Coupled Device





Single Chip Camera
Bayer Filter

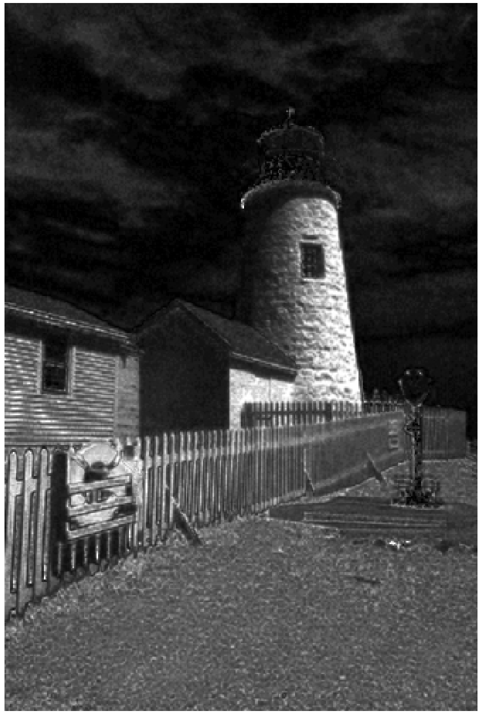
More Colour Representation Schemes



Luminance

Hue

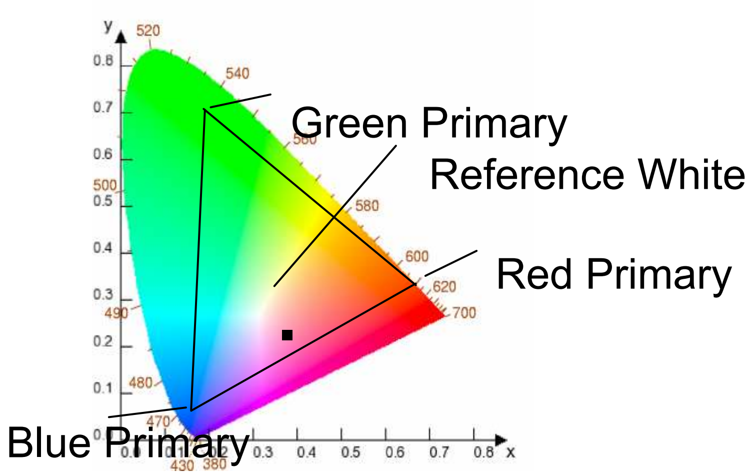
Saturation



Luminance is the intensity of light - independent of colour (black - various shades of grey - white).

Hue is the dominant colour

Saturation is the strength of the colour.



Due to the position of the primary colours and reference white to produce the effect of white light to the viewer need

$$0.30R + 0.59G + 0.11B$$

The luminance signal, Y_s, of an arbitrary colour is given by

$$Y_s = 0.30R_s + 0.59G_s + 0.11B_s$$

Where R_s, G_s and B_s are magnitudes of the three colour component signals (Red, Green and Blue)

Now we have the **luminance** information, need to code the colour information.

Not practical to supply separate colour signals, hence use colour difference or **chrominance** signals

$$C_b = B_s - Y_s$$

$$C_r = R_s - Y_s$$

No need to supply the third signals – (G_s - Y_s)– since:
We only need three equations with three unknowns.

Get back to the original colour signals by:

$$R_s = C_r + Y_s, \text{ etc.}$$

Why transmit luminance and chrominance signals ($Y\ C_b\ C_r$) rather than separate B_s, G_s and R_s signals?

Separate colour and monochrome information

If the scene is in monochrome, then $B_s = G_s = R_s$

Hence

$C_b = B_s - (0.30B_s + 0.59B_s + 0.11B_s) = 0$, and $C_r=0$

This is called **the Principle of constant luminance**

Also permits monochrome receiver to work satisfactorily on colour signals.

We will learn another advantage soon with respect to reducing the data rate

Colour Representation

Colour space models:

1. RGB (Additive Primaries) – Human eye, image capture in cameras.
2. CMY (Subtractive Primaries) – In printing industry
3. HSB or HSV (Hue-Saturation-Brightness) – in Computer Graphics
4. YC_bC_r (Luma and 2 Chroma – Blue Chroma and Red Chroma)
   - Used in when colour television broadcast was introduced.

RGB -> YCbCr

$$Y = 0.30\,R + 0.59\,G + 0.11B$$
$$Cb = B - Y$$
$$Cr = R - Y$$

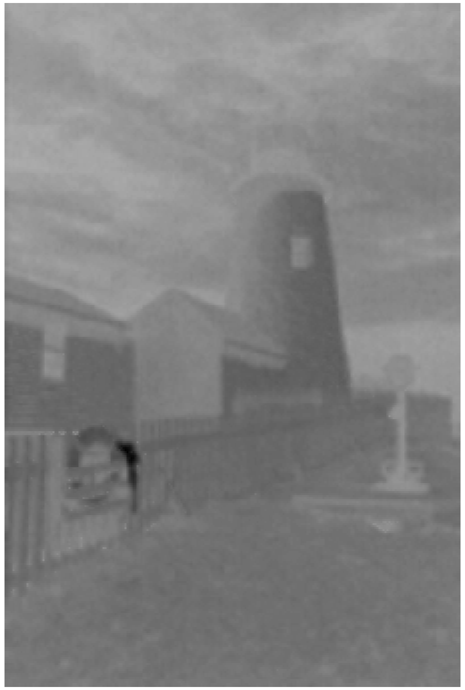
YCbCr -> RGB

$$B = Cb + Y$$
$$R = Cr + Y$$
$$G = (Y - 0.30\,R - 0.11\,B)/0.59$$

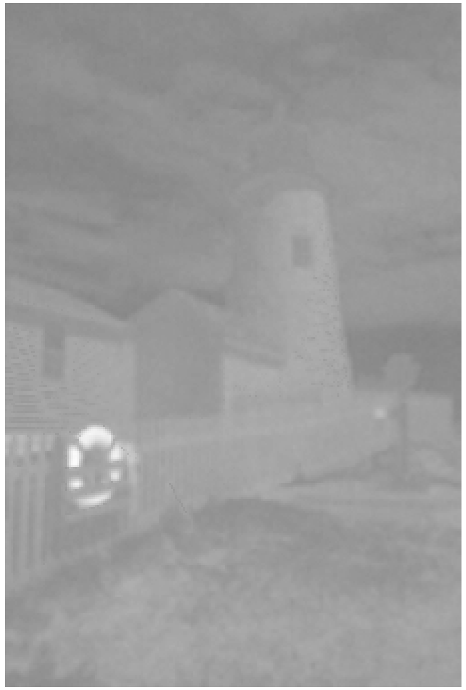
RGB

Y

Cb



Cr



Further reductions of data rate

For analogue – we transmit Luminance (Y) and 2 Chrominance (C_B, C_R) signals.

For digital - could use the separate R, G and B digitised video signals. This requires 3 bytes/pixel memory and then all three signals would have same rate and bits per sample.

But looking at luminance + two chrominance signals, we know that the chrominance signals look smooth, contain fewer finite details and structure is not that visible.

Therefore the human eye is more tolerant to reduced resolution in chrominance channels.

So … down-sample the chrominance channels to reduce data rates.

Chrominance subsampling

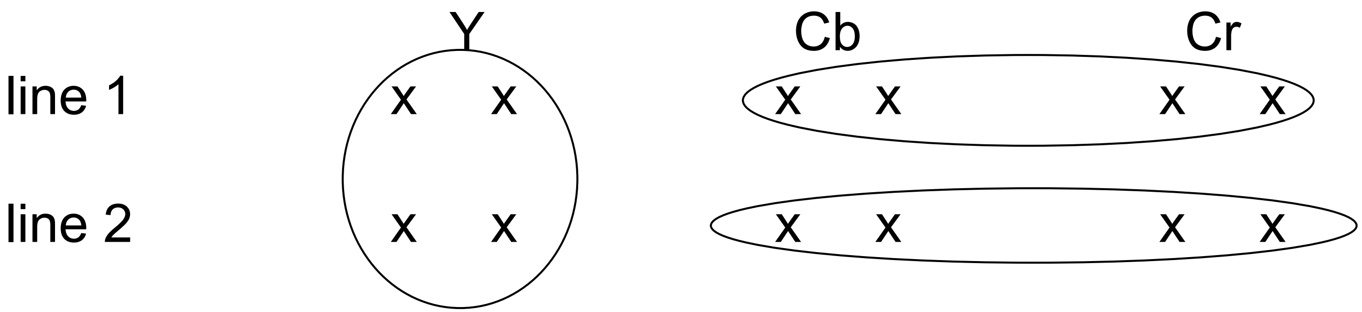
There are three Y Cb Cr formats:

4:4:4 4:2:2 4:2:0

These three numbers A:B:C represent

In a 2x2 group of pixels:

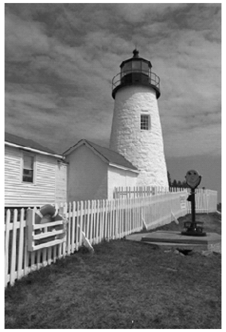
A is the number of Y samples (on both lines).
B is the number of Chrominance (Cb & Cr) samples on line 1
C is the number of Chrominance (Cb & Cr) samples on line 2



W is the Width H is the Height

Y Cb Cr

4:4:4



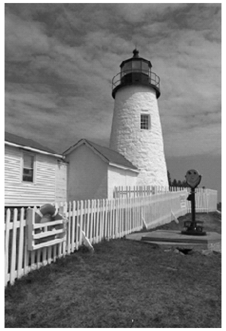
WxH

WxH

WxH

Total bits
=N(WH+WH+WH)
=3NWH

4:2:2



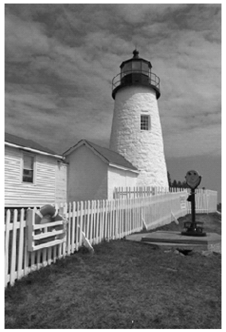
WxH

W/2 x H

W/2 x H

Total bits
=N(WH+WH/2+WH/2)
=2NWH

4:2:0



WxH

W/2 x H/2

W/2 x H/2

Total bits
=N(WH+WH/4+WH/4)
=1.5NWH

N is the bits per colour component

How to find the bit rate?

W	pixels/line
H	Lines/frame

H x W pixels/frame (This is the pixels in Y channel)

S	Chrominance sub sampling factor S=3 for 4:4:4 S=2 for 4:2:2 S=1.5 for 4:2:0
N	bits/pixel (bpp)

S x N x H x W bits/frame

F	frames/sec	(For a video sequence)

S x N x H x W x F bits/sec

Derive the bit rate for 4:2:2 format 4CIF video!

Y channel resolution of 4CIF video = 576 x 720

Considering N= 8 bits per pixel per colour channel

Cb and Cr resolutions using 4:2:2 sampling S= ?

Memory per frame = ?

Memory to store 90 minutes of video (at 50 frames per sec) = ?

How many DVDs are required to record this programme?
(A single layer DVD can store only about 4.5 Gbytes per disk)

The University Of Sheffield.

Another Example

Derive bit rate for video transmission over mobile networks using QCIF 4:2:0 format

Mobile phones display resolution

180 x 144

Typical Frame rate is 6.25 fps

Bits/frame ~

At 6.25 frames/s ~

What is the available mobile phone network bandwidth?

What can we do to further reduce the data rate?

Homework: MATLAB

- MATLAB – Preliminaries
- A good tutorial – MATLAB Primer
  http://www.math.toronto.edu/mpugh/primer.pdf
- Useful commands: help, lookfor, helpdesk, who, whos
- How to read an image into MATLAB (hint: lookfor image and help <Command name>)
- What image file formats can be read in MATLAB?

Homework: MATLAB

- Exercise 1:
  - Download the image "testimage.png" from MOLE into your PC
  - Read testimage.png into MATLAB
  - What are the image dimensions? How many colour components? What "data type" has been used?
  - Display image using MATLAB
  - Now convert the image into its luminance format
  - Save the luminance image and find out the file extension of the saved file
  - Clear the memory space
  - Load back the saved image
  - What is the average luminance value?
  - Reduce the resolution of the image by 2 and display
  - What is the current bit depth resolution? Reduce the bit depth resolution by 2 and display.
  - Display a 100x128 rectangular region starting from the point (200,128) of the original luminance image in a new window.

Homework: DSP

- Next Lecture: Signal processing background
- Revise
    - Discrete time signals
    - Convolution
    - Impulse response
    - Correlation
    - FFT
    - Frequency response
    - Filters (low pass and high pass)