Robotics: Principles and Practice

Module 5: Robot Vision

Lecture 1: Computer vision; optics and sensors; image acquisition; image representation

David Vernon
Carnegie Mellon University Africa

www.vernon.eu

Computer Vision

Computer Vision is concerned with the

- content
- organization, and
- behaviour

of a 3D world

by the automatic analysis of images of that world

Computer Vision

IMAGE → KNOWLEDGE

Extract descriptions of the world from images

Descriptions of what kind? qualitative vs. quantitative

- Geometric: shape and position of object or distances in 3D world
- Semantic: what objects do I see?
- Dynamic: scene changes, object velocities, human actions, …

Credit: Francesca Odone, University of Genova

Computer Vision

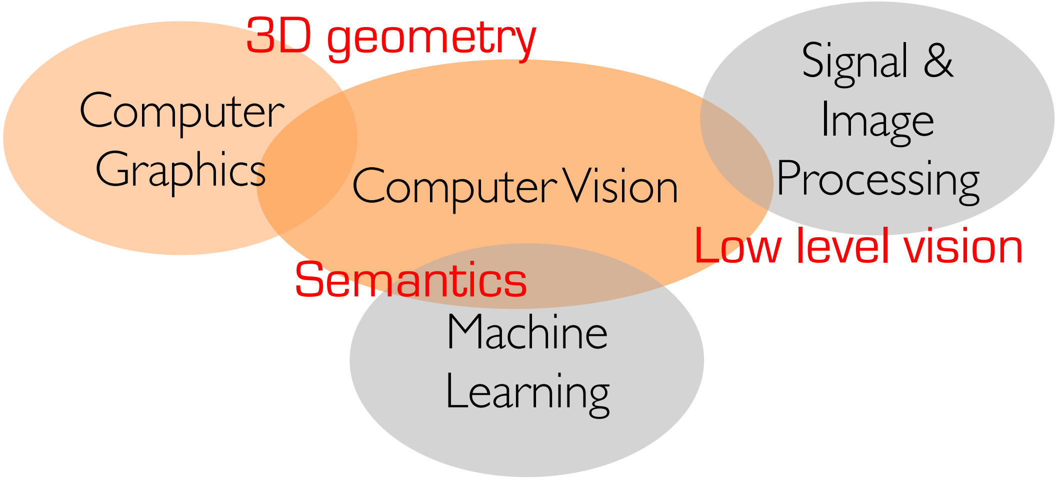
Recognizing objects from images . . .

. . . may be difficult!

151	138	153	138	137	119	111	111	110	74	74	62	69	74	62	78
137	144	145	137	128	137	119	111	129	86	69	69	74	62	69	69
134	137	133	129	128	128	126	119	119	74	69	62	64	78	62	64
79	79	74	69	74	79	69	74	62	69	69	74	74	72	64	64
62	62	64	64	57	57	55	57	62	69	69	72	69	74	62	69
57	64	74	57	64	74	64	74	74	74	69	74	69	78	62	64
86	84	91	84	86	86	84	86	86	84	69	74	69	69	74	69
94	86	94	94	86	94	91	86	91	79	78	64	69	78	62	69
94	99	102	94	94	103	94	99	103	86	69	64	69	74	69	69
128	138	133	144	138	138	145	144	142	100	78	74	69	72	69	64
176	176	183	170	176	170	183	170	168	126	78	74	78	78	69	69
174	176	176	189	170	176	170	189	170	135	74	69	74	72	69	69
183	181	183	174	181	181	189	181	170	135	78	69	69	69	69	64
181	181	183	174	181	189	170	181	172	146	86	69	74	72	62	69
176	174	181	181	183	174	183	181	174	148	84	69	69	74	74	69
181	176	183	183	174	181	181	181	176	141	99	69	69	72	69	69
183	183	183	181	181	189	183	174	181	151	91	69	69	69	64	69
181	181	189	174	181	181	174	189	183	161	91	69	69	78	62	69
176	181	183	181	174	189	176	181	174	165	91	78	74	72	69	69
181	176	189	181	176	176	183	174	181	163	91	78	69	78	62	69
176	174	183	189	176	181	174	170	176	168	91	74	69	74	69	69

Credit: Alessandro Saffiotti, University of Örebro

Computer Vision



An incomplete view of related disciplines

Credit: Francesca Odone, University of Genova

Computer Vision

- The image is two-dimensional
- We lose information in the projection process, i.e., in passing from a 3D world to a 2D image
- The images are digital images
  - they are a discrete representation (i.e. they have distinct values at regularly sampled points)
  - they are a quantised representation (i.e. each value is an integer value)

Computer Vision

Image formation system

- part illumination
- sensing element
- associated optics

is critical to the successful deployment of industrial systems

Computer Vision

The task of the image acquisition and processing sub-system is

- to convert this signal into a digital image

- to manipulate the resultant image to facilitate the subsequent extraction of information

Computer Vision

The image analysis phase is concerned with the extraction of explicit information regarding the contents of the image (*e.g.* object category, identity, position, size, orientation)

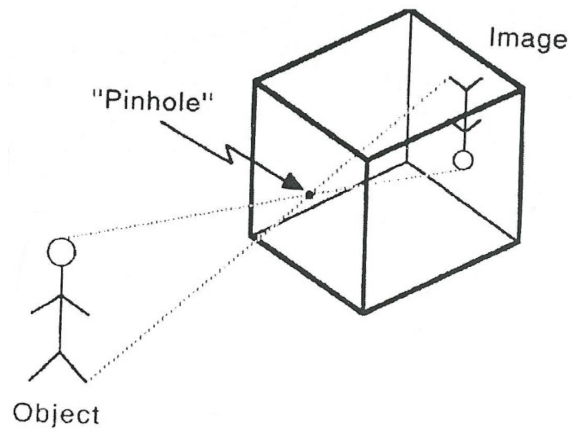
There is a fundamental difference between image processing and image analysis

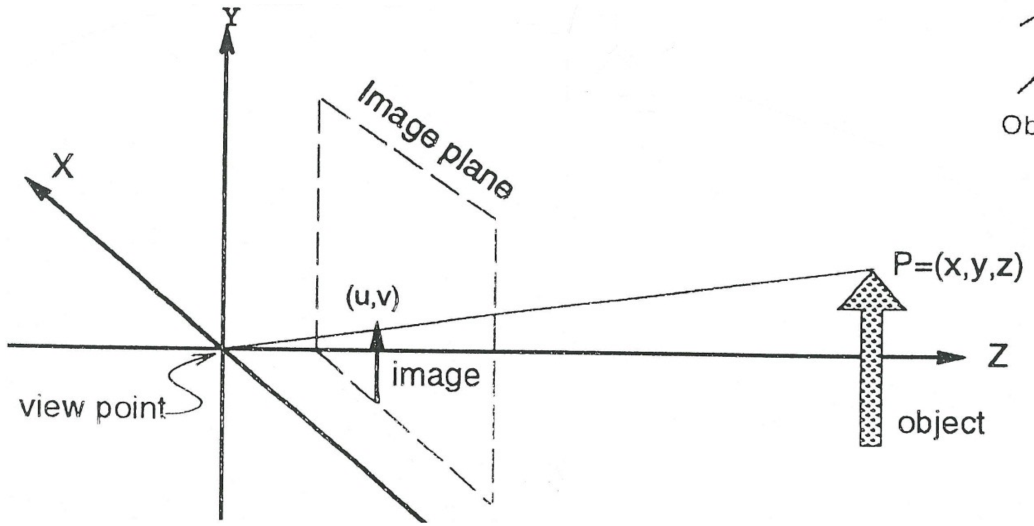
- Image Processing facilitates transformations of images to (hopefully, more useful) images
- Image Analysis facilitates the transformation from an image to explicit (symbolic) information

Optics and Sensors

Optics

- Perspective Projection
- Pin-hole model of a camera





Basic model of imaging process

Optics

Lenses are required to focus part of the visual environment onto the image sensor

Lenses are defined by:

- their Focal Length (quoted in millimetres)
- their Aperture (the f number)

These parameters determine the performance of the lens in terms of light gathering power and magnification, and it often has a bearing on its physical size

Optics

The focal length of a lens is a guide to the magnification it effects and its field of view

Selecting the focal length which is appropriate to a particular application is simply a matter of applying the basic lens equation

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

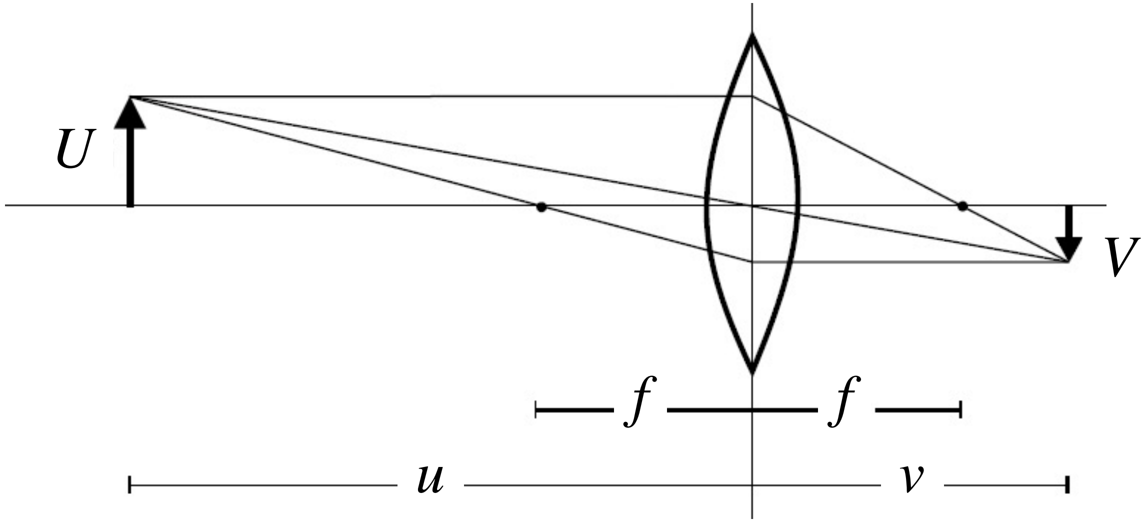
where

- v is the distance from the lens to the image
- u is the distance from the lens to the object
- f is the focal length

Optics

Gaussian lens equations: $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$ and $\frac{V}{U} = \frac{v}{u}$

Focal length: $f = \frac{uV}{U + V}$



Optics

Noting the Magnification Factor M is

$$M = \frac{image\_size}{object\_size} \quad \frac{V}{U}$$

Optics

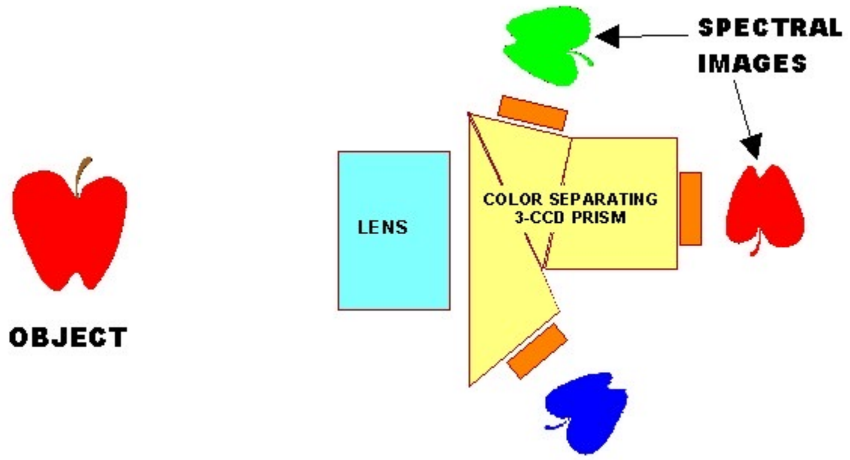
Thus

$$f = \frac{uM}{M+1}$$

Hence if we know the required magnification factor and the distance from the object to the lens, we can compute the required focal length.

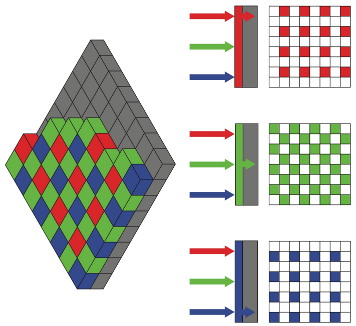
Sensors

3-Chip Colour Camera



1-Chip Colour Camera

(Bayer filter)



C. Bartneck, T. Belpaeme, F. Eyssel, T. Kanda, M. Keijsers, S. Šabanović, Human-Robot Interaction – An Introduction, Cambridge University Press, 2020

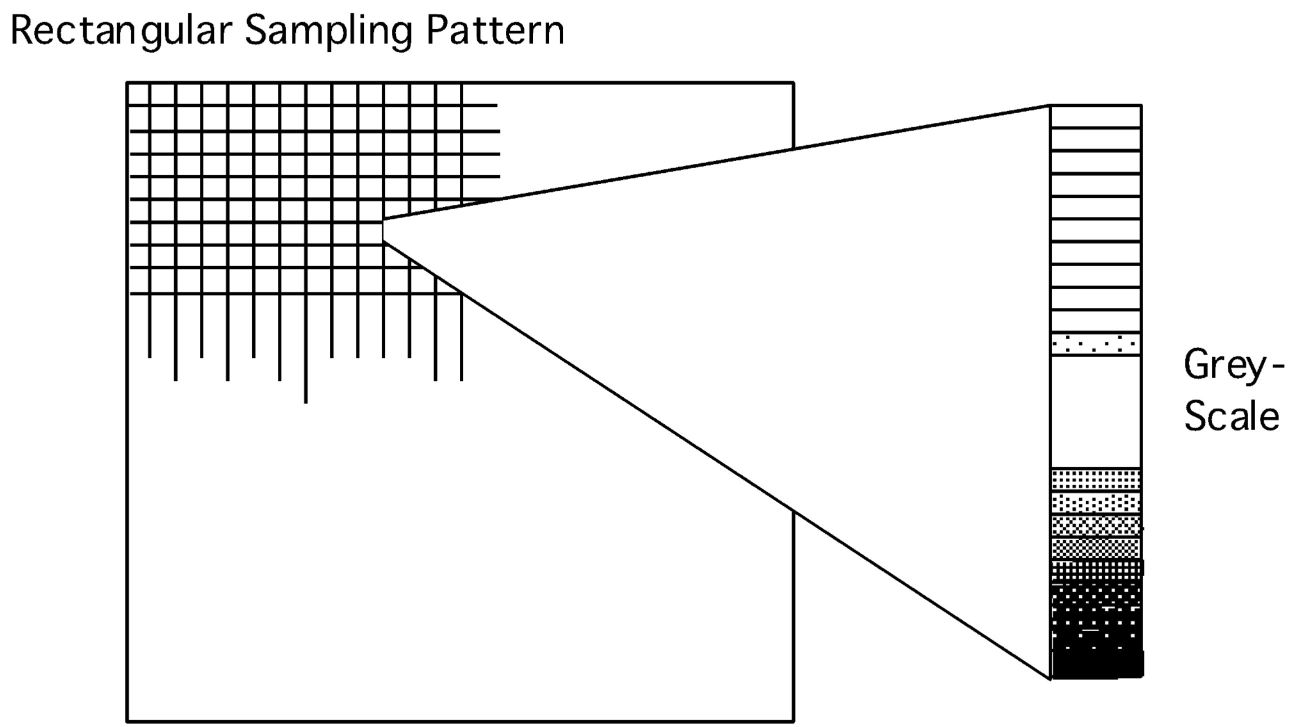
Image Acquisition
and
Image Representation

Image Acquisition

Digital images represent the reflectance function of a scene but they do so in a sampled and quantised form

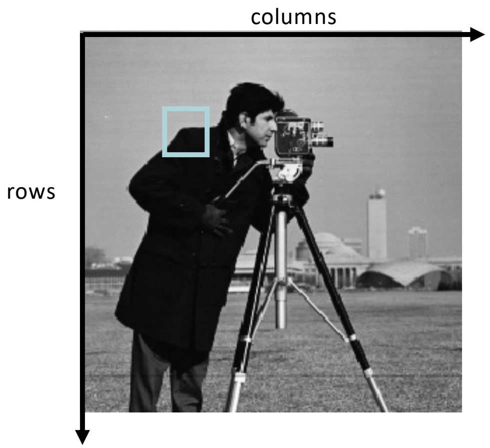
Each quantized integer value is known as a pixel and is the smallest discrete accessible sub-section of a digital image.

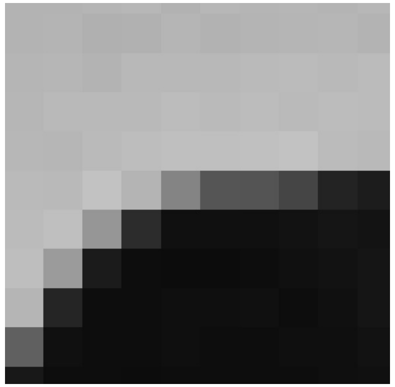
Image Acquisition



Sampling and Quantisation

Image Acquisition

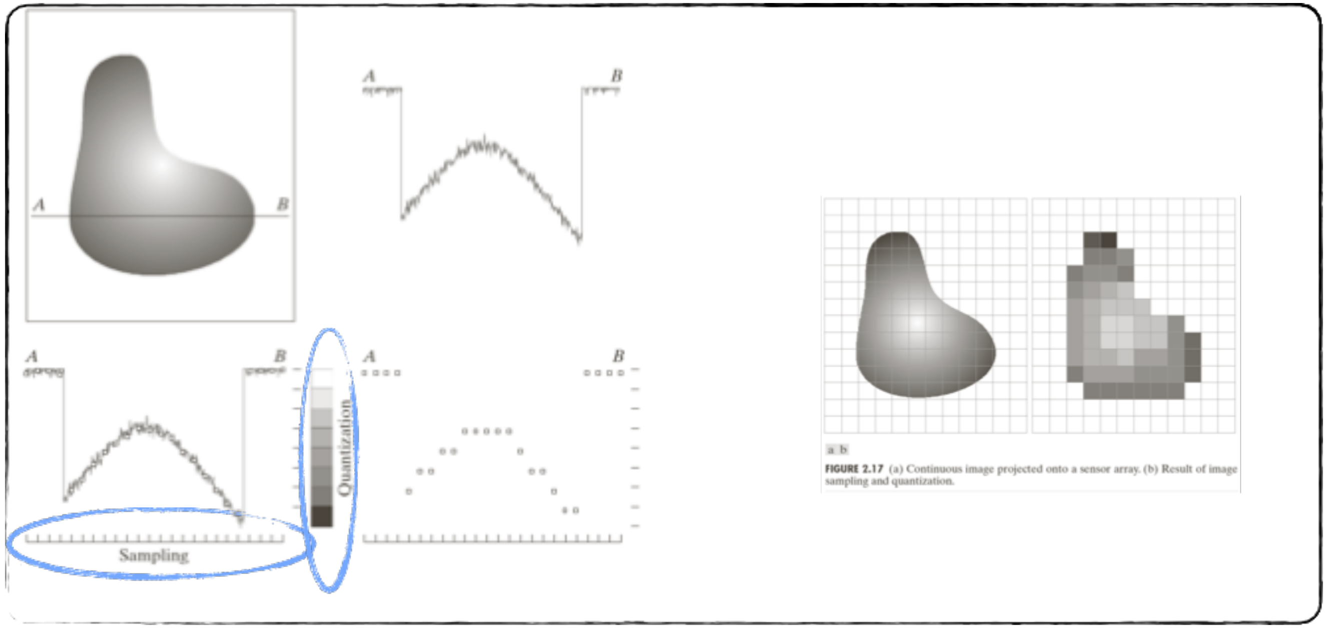




180	183	181	181	179	181	180	180	182	181	183	181	183	184	184	184	184	181	184	183	180
178	178	178	177	182	182	180	177	182	180	183	179	181	181	183	180	181	180	179	180	180
179	178	178	180	177	178	179	179	181	183	178	183	181	182	180	183	181	181	180	181	182
179	179	177	179	177	171	179	180	176	177	181	178	180	181	182	179	183	181	179	177	179
180	179	180	181	180	181	181	182	179	184	184	184	186	187	185	187	182	183	187	186	182
178	179	181	182	180	182	182	185	185	185	188	186	188	186	188	187	190	191	188	186	186
181	178	181	178	186	180	183	182	186	189	191	191	192	194	187	185	178	178	180	184	188
179	177	180	181	180	183	186	185	194	180	132	85	84	69	35	28	25	25	27	29	41
176	177	181	180	183	181	187	191	150	44	16	16	17	18	21	19	19	18	20	19	21
181	181	180	181	181	183	190	155	27	13	12	12	13	16	18	21	18	19	19	19	17
179	179	181	179	186	185	181	37	14	14	15	15	16	14	16	21	19	18	20	19	19
182	180	184	182	183	195	96	16	14	14	15	14	14	15	15	18	20	17	19	18	18
180	180	183	186	191	159	24	13	13	12	13	14	14	14	15	16	20	17	19	19	17
181	183	185	190	185	49	15	16	16	13	12	12	13	14	14	14	19	17	17	18	19
184	184	187	192	104	22	20	17	16	15	12	14	13	15	16	16	18	19	18	18	18
186	189	193	133	24	17	16	16	15	15	16	13	13	15	16	16	16	17	17	16	18
188	192	137	31	16	16	16	17	18	18	15	14	13	14	15	16	17	16	17	16	16
193	126	26	17	17	16	16	16	17	16	16	14	14	13	14	16	16	15	15	16	16
131	24	16	15	16	16	16	16	16	16	17	15	14	12	13	14	15	15	14	16	16
24	17	17	15	16	16	16	17	16	16	17	16	15	15	14	15	15	17	15	15	16
15	15	13	13	15	17	16	15	15	15	14	13	13	13	13	14	14	14	14	15	14

Credit: Francesca Odone, University of Genova

Image Acquisition



Credit: Francesca Odone, University of Genova

Image Acquisition

Dynamic range

Total number of distinctive values occurring in the image

- it is limited by the number of bit per pixel we may want to use
- it is also limited by the physical dynamic range of the sensor

this is related to the quantization process...

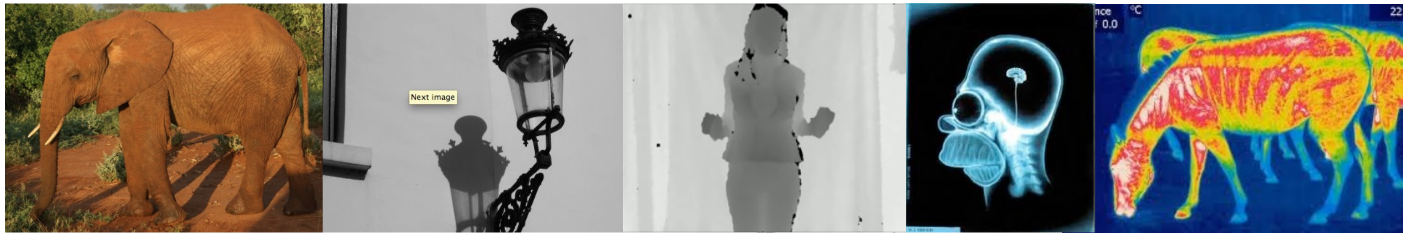
- To represent black-white images 1 bit is sufficient
- Grey level images: usually associate a byte to a pixel 2^8=256 gray levels
- Color images: usually 1 byte per channel ("millions of color")

Credit: Francesca Odone, University of Genova

Image Representation

Pixel content depends on the image type

- Gray level pictorial digital images ("black and white photos"): intensity
- Color pictorial digital images: color (modeled as triplets, eg RGB)
- Range images: depth information
- Medical images: radiation absorbance level
- Thermal images: heat



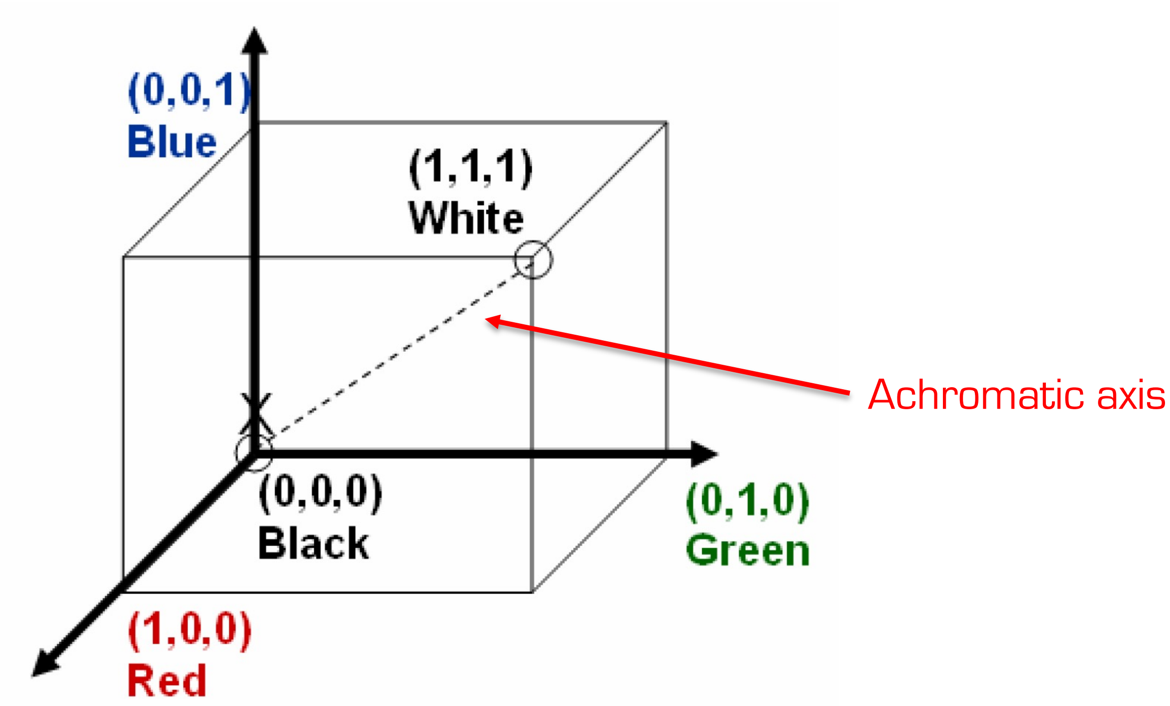
Credit: Francesca Odone, University of Genova

Colour Spaces

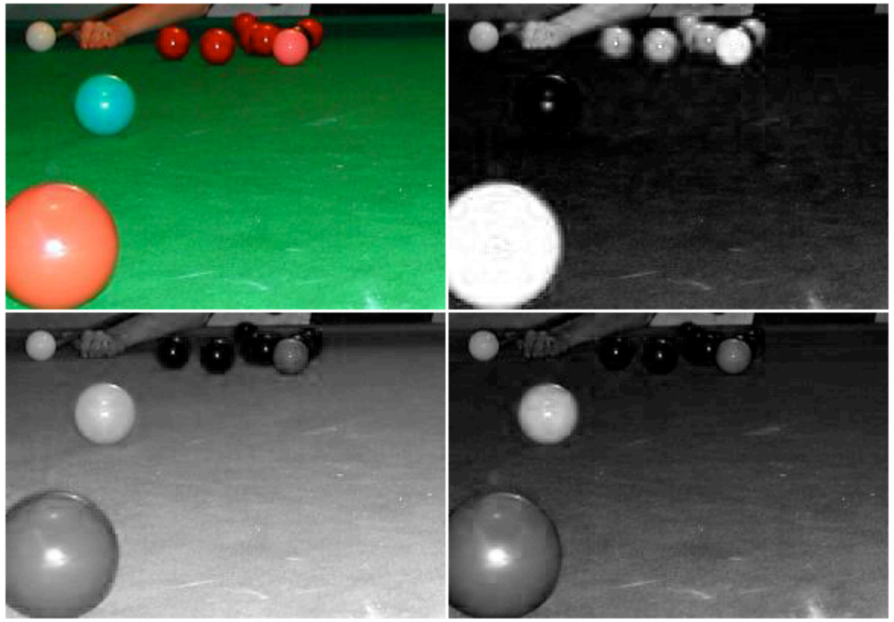
There are many colour representations

RGB	Red, Green, Blue
CMY	Cyan, Magenta, Yellow
YUV	Luminance (Y), Blue minus Luminance (U), Red minus Luminance (V)
YCrCb	Scaled version of YUV
CIE XYZ	Standard reference colour space based on the response of human eye
CIE L*u*V*	Perceptually uniform colour space
CIE L*a*b*	Device independent colour space (all colours perceived by humans)
HSV	Hue, Saturation, Value
HLS	Hue, Luminance, Saturation
HSI	Hue, Saturation, Intensity

Colour Spaces



Colour Spaces

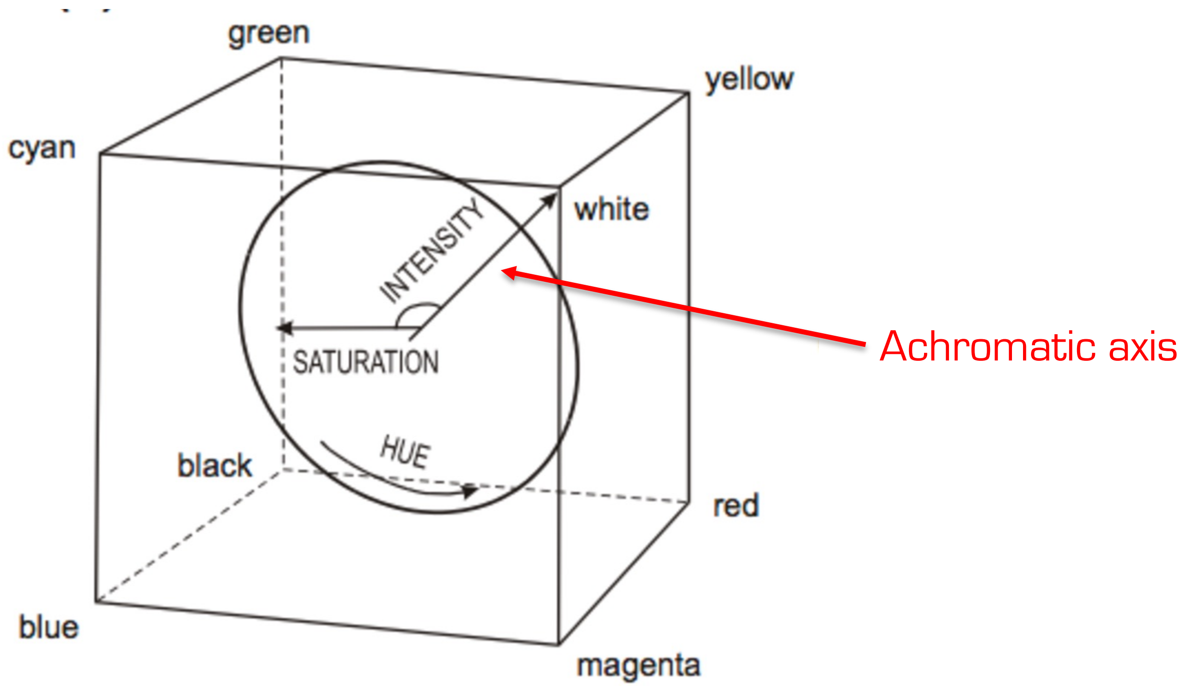


Red Green Blue images

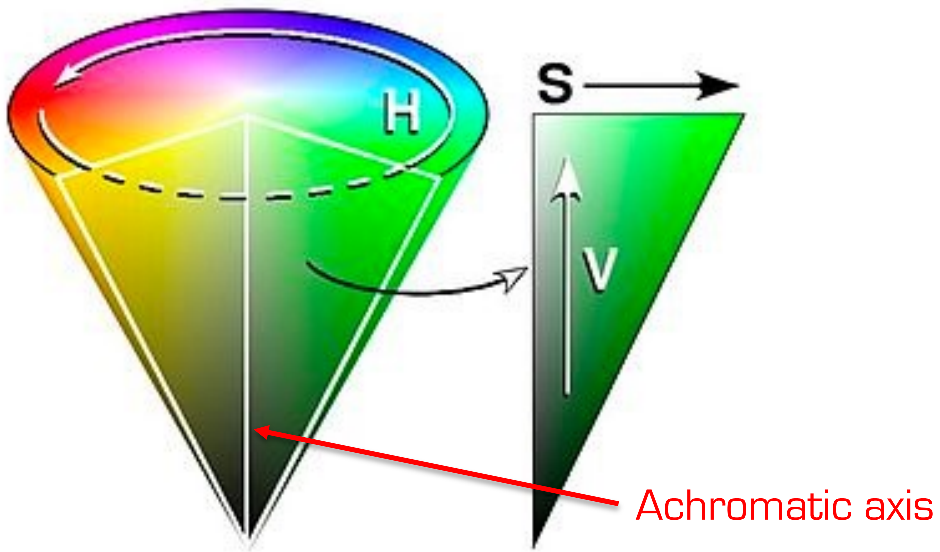
Red (~ 700nm)
Green (~ 546nm)
Blue (~ 436nm)

Credit: Kenneth Dawson-Howe, A Practical Introduction to Computer Vision with OpenCV, © Wiley & Sons Inc. 2014

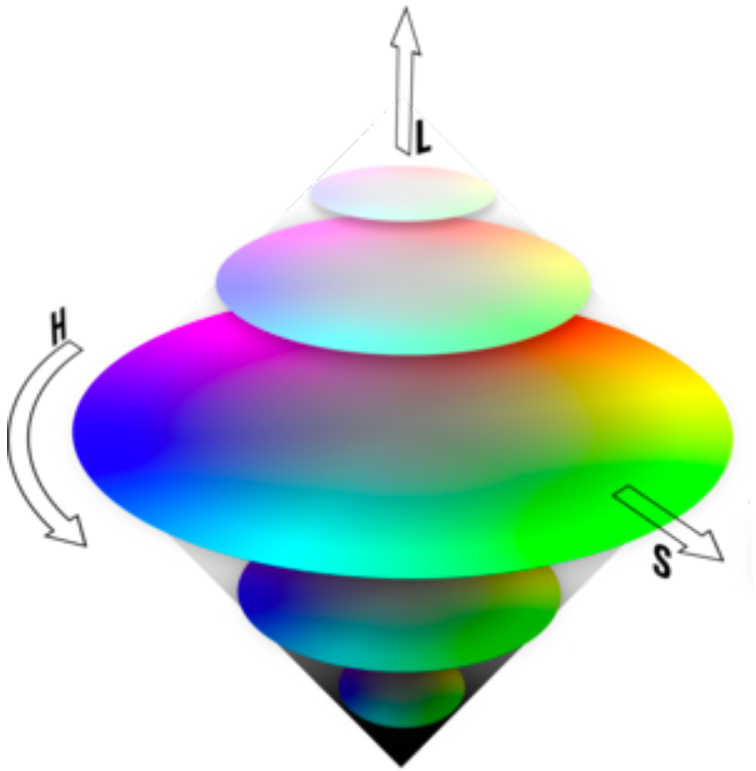
Colour Spaces



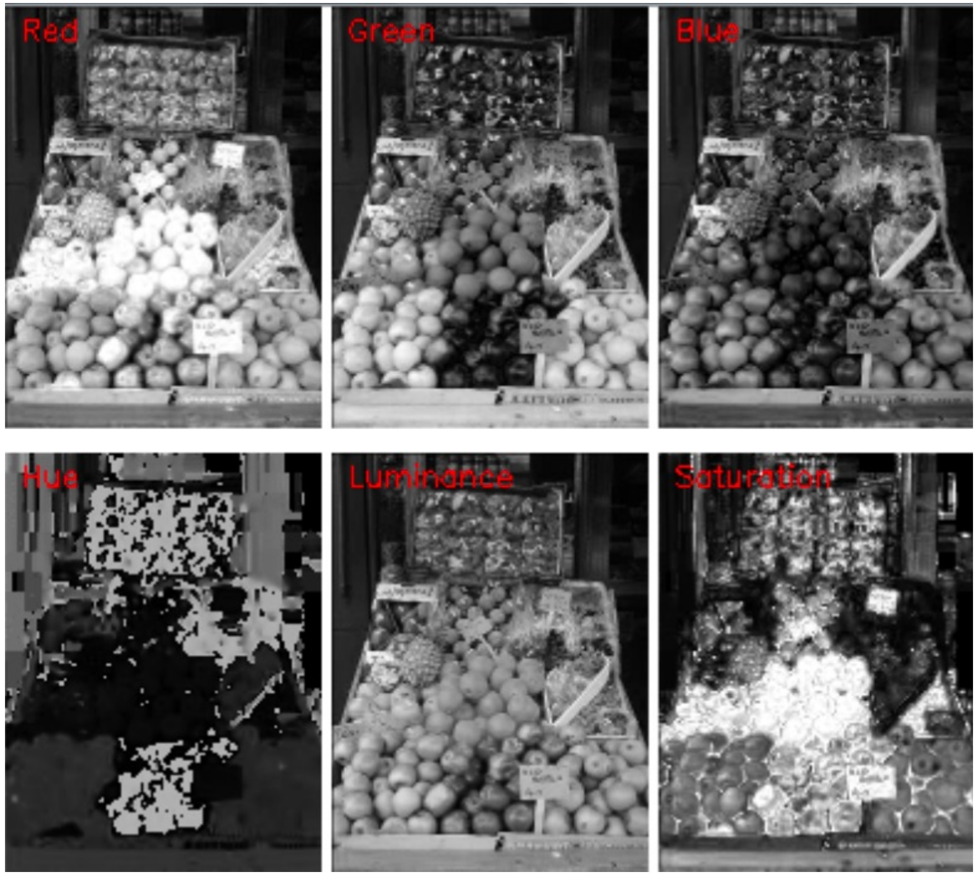
Colour Spaces



Colour Spaces



Colour Spaces



Credit: Kenneth Dawson-Howe, A Practical Introduction to Computer Vision with OpenCV, © Wiley & Sons Inc. 2014

Reading

R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.

- Section 2.2.3 Optics
- Section 2.3 The digital camera
    - Section 2.3.1 Sampling and aliasing
    - Section 2.3.2 Colour

D. Vernon, *Machine Vision*, 1991.

- Section 2.2.1 Image formation: elementary optics
- Section 2.2.2 Camera sensors
- Section 3.1 Sampling and quantization