

# Course Content

---

## Image Feature Extraction / Object Representation

### Learning discriminant functions

- From models of feature generation  
*Bayesian decision theory*
- From direct modelling of decision boundary  
*Linear perceptrons; Neural networks; Support vector machines;  
Nearest neighbour*

### Finding structure / Unsupervised learning

- Mixture models
- Clustering

### Combining classifiers

# Your Workload

---

**Homeworks** *Programming & Pen and Paper* exercises (**1.5 hp**)

**Lab Project** Implement a face detector (**1.5 hp**)

**Final Exam** Answer theory and problem questions about the material covered in the course (**3 hp**)

# Task and its Assessment

---

## Homeworks: Pass / Fail

At the end of each lecture, I will assign a relatively simple *programming task* and/or a *pen and pencil exercise*. You will:

1. attempt/finish these tasks,
2. bring your solutions to the next lecture,
3. have it corrected by a fellow student

# Task and its Assessment

---

## **Homeworks: Pass / Fail**

Successful completion of a homework assignment is then defined as:

1. coming to the lecture with the assignment attempted/completed (determined by your peers)

**AND**

2. correcting a fellow student's assignment (according to a provided sample solution)

There will be 11 such assignments and to pass and be eligible to take the exam you must successfully complete 9 of the tasks (see course website for **What if's**).

# Task and its Assessment

---

**Lab Project:** Pass/Fail but bonus points available for exam

In groups of 1 to 3 you will implement a face detector using the boosting algorithm.

This part of the course will be assessed as follows:

**Part I** Complete a basic implementation. (Mandatory)

**Part II** Extend the basic implementation and testing performed for **Part I** and describe these extensions in a *Poster Presentation* to be graded by me and TA.

(Optional, bonus points available “exam mark” +  $\leq 10\%$ )

# Task and its Assessment

---

## **Final Exam:** final grade

Your final grade for the course will be determined by your performance in the final exam.

The exam will have a compulsory part - **Part I** - covering the core topics of the course. This part will have to be answered at a high level to pass the exam.

To obtain higher grades other questions, other topics and **Part I** material at a higher level of understanding, will have to be answered in **Part II**.

Nearer the exam, I will be very explicit about the content and the level of difficulty of **Part I** and **Part II** of the exam.

# Summary

---

To **pass** the course, (Grade E), the following criteria have to be fulfilled:

- Completion of 9 homework assignments
- Completion of **Lab Part I**.
- Score at least a 70% on **Part I** on the final exam.

If you achieve the above then you must perform the following to get:

Grade D

- excellent performance in **Part I** or
- mediocre performance in **Part II** or
- bonus marks from a good poster presentation

## Grade C

- excellent performance in **Part I** + bonus marks **or**
- mediocre performance in **Part II** + bonus marks **or**
- average performance in **Part II**

## Grade B

- average performance in **Part II** + bonus marks **or**
- good performance in **Part II**

## Grade A

- good performance in **Part II** + bonus marks **or**
- excellent performance in **Part II**

Nearer exam time I will be much more specific about my performance demands for each grade. I just want to indicate that a good lab project could bump up your overall grade for the course.



# Register for the course

---

I have set up a small on-line web-form (link from the News part of the course web-site) at

[http://www.csc.kth.se/~sullivan/cgi-bin/pre\\_register\\_DD2427.html](http://www.csc.kth.se/~sullivan/cgi-bin/pre_register_DD2427.html)

You will be asked to fill in some personal details so that I can

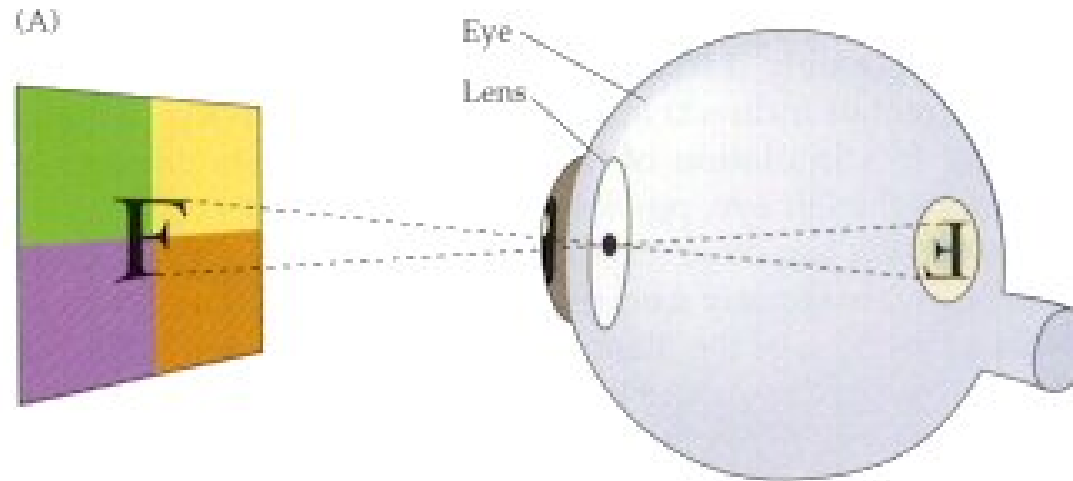
- register you for the course
- add you to the course e-mail list.

**The form is password protected. The password is....**

# **Imaging Basics: Images and Pixels**

# Image Formation

---



Projection from a surface to a 2D sensor. The **eye** and the **camera** are examples of such devices.

# Analog Image

---



A black and white image on film is a two-dimensional function  $f : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}^+$  which maps a spatial location  $(x, y)$  to a gray value/brightness,  $f(x, y)$ , at that point. However, **cannot store** such a **continuous image** on a digital computer.

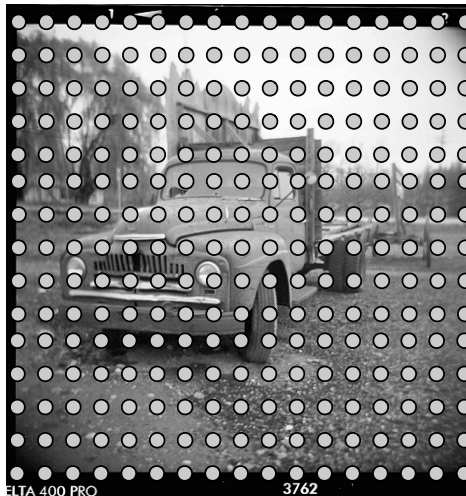


# Digital Image

---

Spatial locations and grayscale values are discrete.

Effectively, at a discrete set of points, a regularly spaced grid, the average value of  $f(x, y)$  in a region around each point is computed and quantized into one of 256 values - 8 bits.



(have  $3 \times 8$  bits per point for color images)

# Image Digitization

---

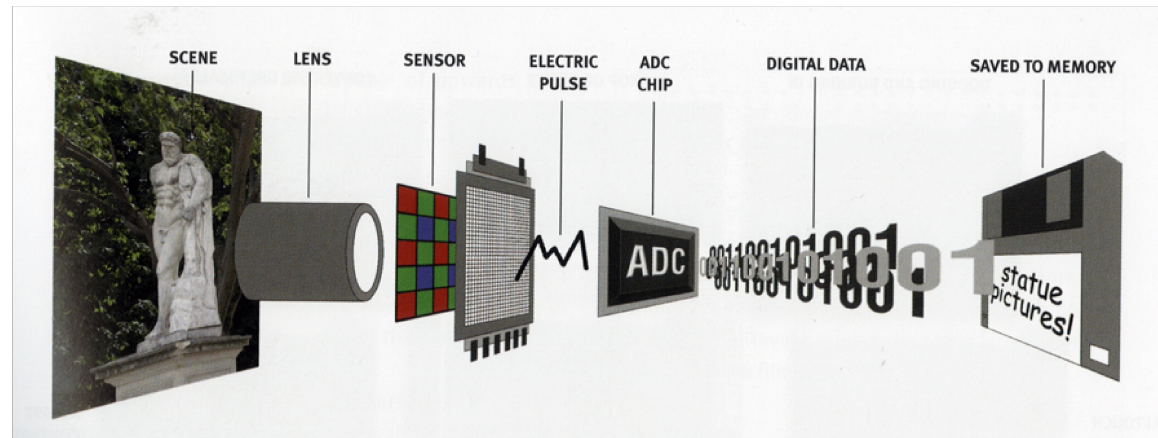
Each pixel represents a small region in the continuous image being digitized.

The spacing and number of pixels determine the size of image features that can be seen.

- Fewer pixels  $\Rightarrow$  inferior quality and smaller file size
- More pixels  $\Rightarrow$  superior quality and larger file size

# Digital Camera

---



- Light travels from world's surfaces into the lens of the camera.
- Photons hit the CCD (charge-coupled-device) array - the silicon sensor.
- Each CCD produces an electronic charge - the more light, the bigger the charge. (For colour images - colour filters placed over the pixel sensors of an image sensor to capture colour information.)
- The electronic pulses are collated, converted to digital information (via an ADC chip) and stored according to their position on the CCD.



- This information is stored on the camera's memory card.

### **Press Release from the Nobel Prize in Physics 2009:**

A large share of the traffic is made up of digital images, which constitute the second part of the award. In 1969 **Willard S. Boyle** and **George E. Smith** invented the first successful imaging technology using a digital sensor, a CCD (Charge-Coupled Device). The CCD technology makes use of the photoelectric effect, as theorized by Albert Einstein and for which he was awarded the 1921 year's Nobel Prize. By this effect, light is transformed into electric signals. The challenge when designing an image sensor was to gather and read out the signals in a large number of image points, pixels, in a short time.

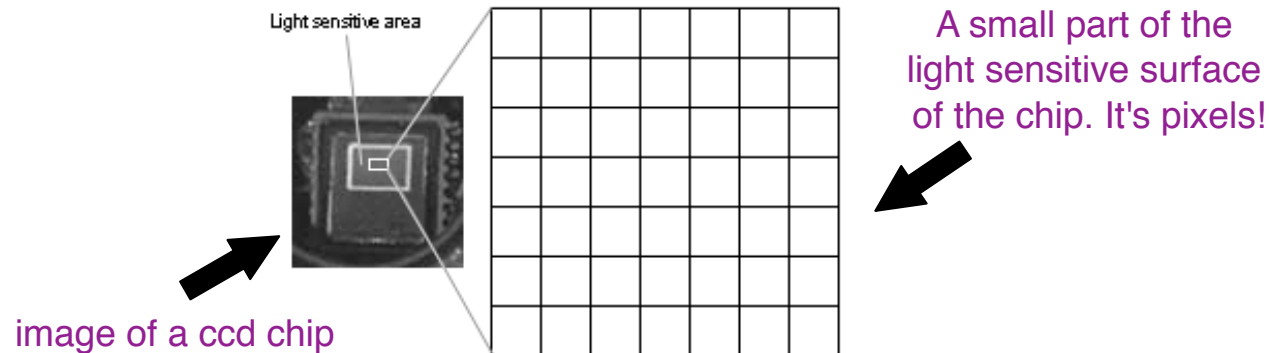
The CCD is the digital camera's electronic eye. It revolutionized photography, as light could now be captured electronically instead of on film. The digital form facilitates the processing and distribution of these images. CCD technology is also used in many medical applications, e.g. imaging the inside of the human body, both for diagnostics and for microsurgery.

Digital photography has become an irreplaceable tool in many fields of research. The CCD has provided new possibilities to visualize the previously unseen. It has given us crystal clear images of distant places in our universe as well as the depths of the oceans.

# What is a CCD?

---

A CCD chip is a light sensitive device, made of silicon. It is an array of light sensitive pixels. When light falls on a pixel, it is converted to a charge. This charge is captured in the pixel. It can't go to other pixels. As more light falls on the pixel, the more charge accumulates in the pixel.



The amount of charge is a measure of the amount of light that fell

on the pixel.

The amount of charge in every pixel makes the image. The computer can read this charge and convert it to an image.

CCD array

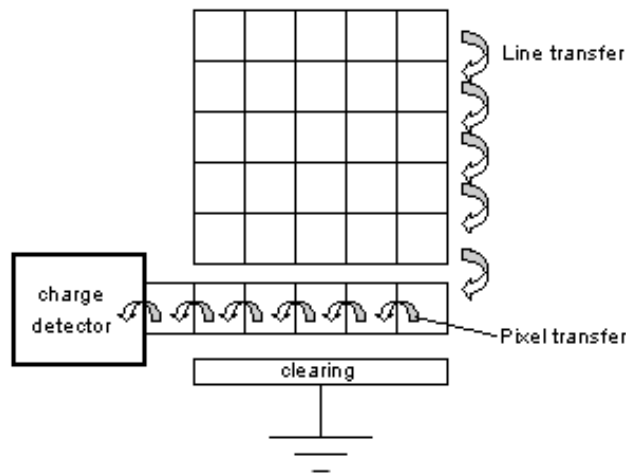
80	92	84	96	104
75	87	54	65	98
83	97	207	159	134
78	79	34	56	21
37	89	65	34	74

The value in every pixel gives the charge,  
which is related to the light intensity.

# From the CCD to an Image

---

- To get the image into your computer, it has to read every pixel's charge.
- The **charge detection node** in the CCD chip is able to measure a pixel's charge.



- Every pixel's charge must be shifted to this device. The lowest line can shift pixels to the left in the direction of the device. These pixels can be measured one-by-one.
- When all the pixels of the lines are measured, all lines are shifted one line down.
- Then again the lowest line can be measured. This is repeated until all lines are measured.
- The computer can read the measurements through the camera electronics.
- Another important task is the conversion of the analog output signal of the charge detection node to a digital format the computer is able to read and it is done by an ADC (Analog to Digital Converter).

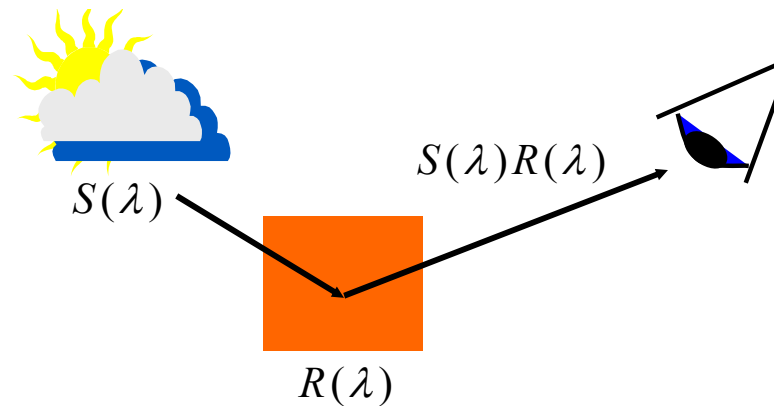
**What about colour ?**

# The physics of colour

---

**Objects** have the **colour** of the **light leaving their surfaces** which depends on

Spectrum of the incident illumination;  
Reflectance properties of the surface;  
Angles of illumination and viewing

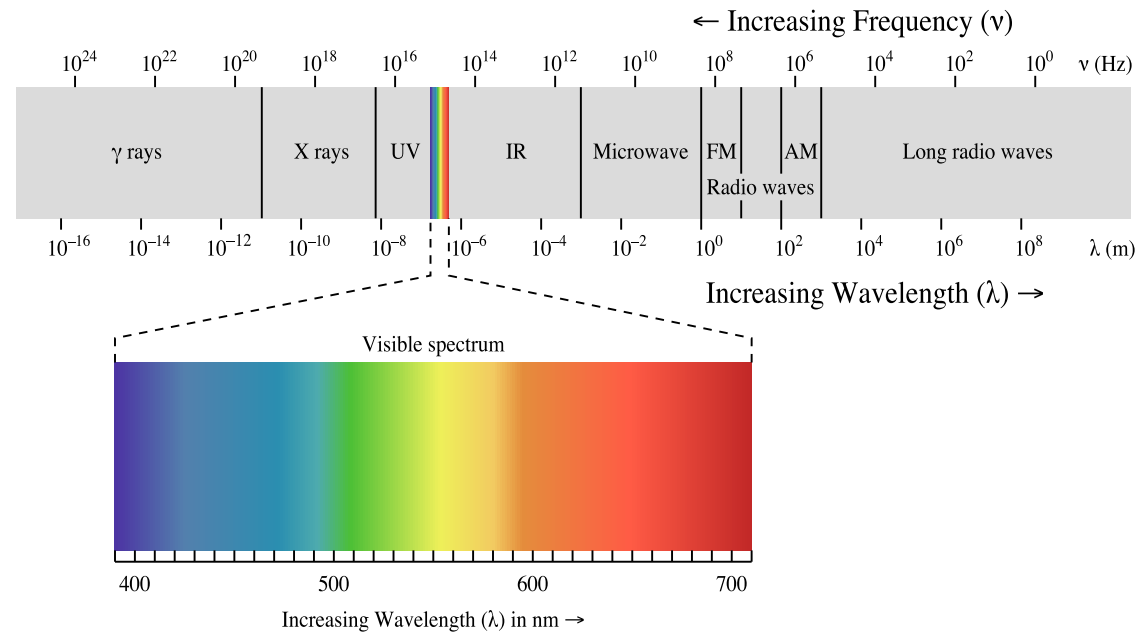


While the **perceived colour** of an object depends on:

light leaving the object's surface **AND** eye and brain observing it

# Human perception of colour

The **electromagnetic spectrum** with visible light highlighted





# Human perception of colour

---

## To recap

The sensation of color is caused by the brain in response to the presence/absence of light at various wavelengths.

For humans, visible light spans the electromagnetic spectrum from 400 to 700 nm.

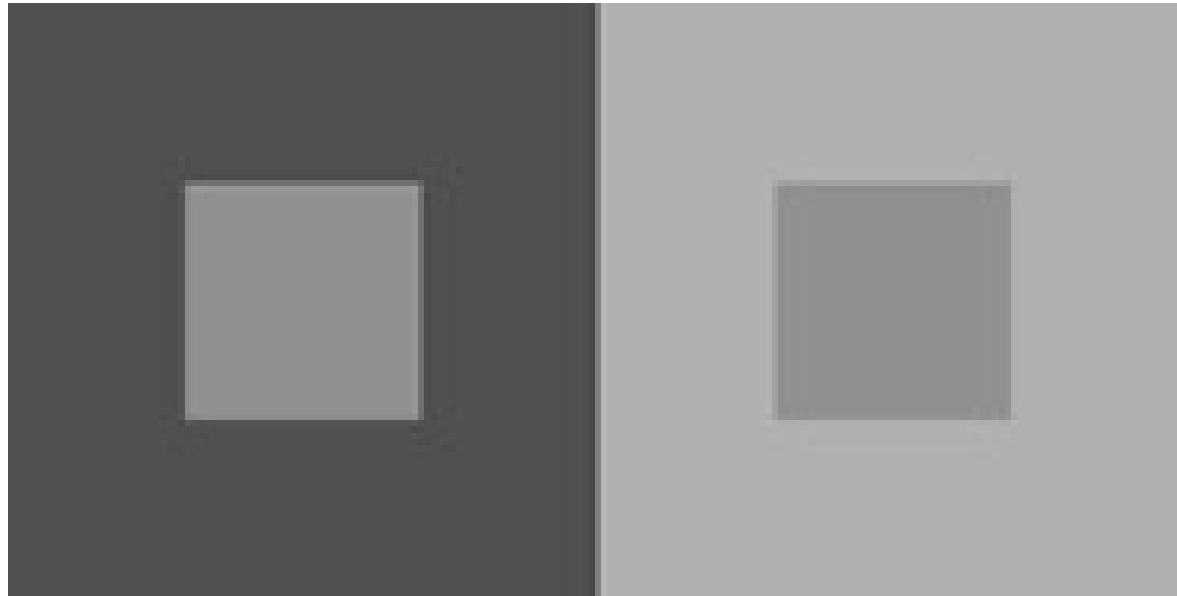
## Note

The perception of color can also result from:

Pressure on the eyelids; Dreaming, hallucinations; Context

# Human perception of colour

---



The perception of color can result from: **Context**

Context

# Description of visible colours

---

The characteristics used to distinguish one color from another are **brightness, hue** and **saturation**. *Chromaticity* is the term used to refer to:

**Hue** dominant wavelength in a mixture of light waves

**Saturation** determined by a combination of light intensity and how much it is distributed across the spectrum of different wavelengths.

*Bright* colours have *high* saturation.

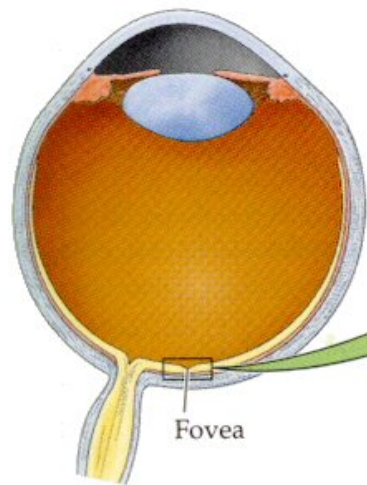
For instance pink is less saturated than a vibrant red.

# Human eye sensing of colour

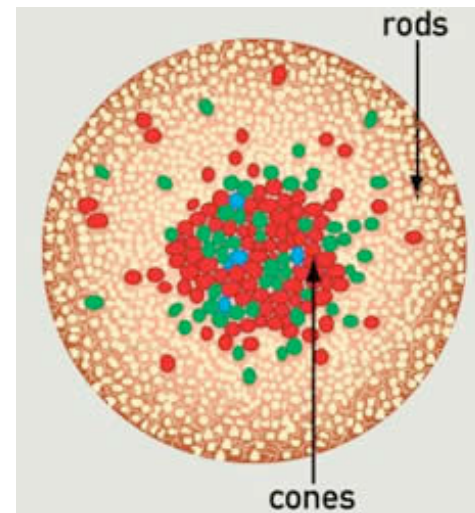
---

At the rear of the eye is the **retina**. This is a light-sensitive panel of cells onto which the eye projects light. On the retina there are two types of light receptors:

## Cones and Rods.

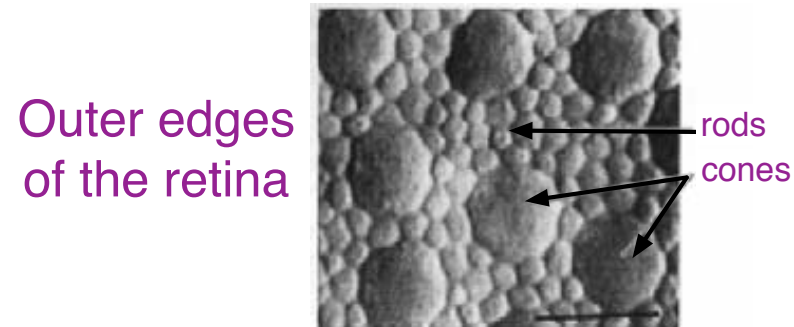


Eye

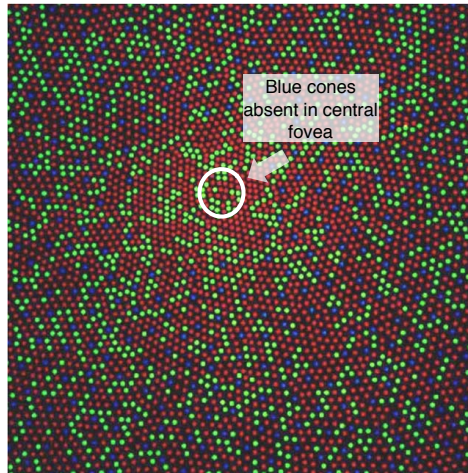


Retina

**Rods** are responsible for low-light monochrome (black-and-white) vision. They cannot distinguish colours. But they are extremely sensitive to light. They are concentrated at the outer edges of the retina and are used in peripheral and night vision. There are roughly 120 million rods in the human retina.



**Cones** only operate in bright light, they provide us with our sharpest images and enable us to see colors. Most of the 4.5 million cones in each retina are confined to the fovea. Our sharpest and colorful images are limited to a small field of view.

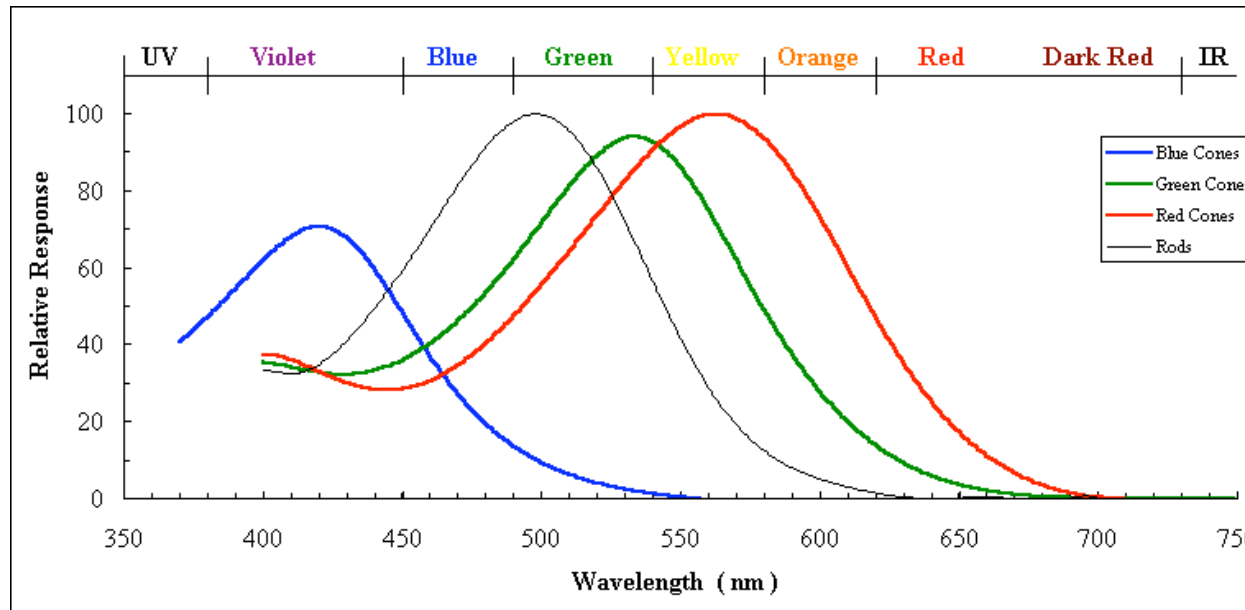


### cones in the fovea

Humans normally have three kinds of cones. Cones are *tuned* to different portions of the visible spectrum:

**Short** (aka Blue); **Medium** (aka Green); **Long** (aka Red)

Approximately 65% of all cones are red, 33% green and 2% blue.



Each class of cone shows an absorption that extends over a range of wavelengths.

# Trichromacy

---

Because of these three types of color-sensing cone cells humans are known as **trichromats**.

Trichromacy is the condition of possessing three independent channels for conveying color information.

Many carnivores, such as dogs, are dichromats.

While birds, fish and insects possess more sophisticated color visual systems than us. They can be tetrachromats (sensitivities to four peak wavelengths) or perhaps even pentachromats (sensitivities to four peak wavelengths).



# Specifying colours numerically

---

Want to specify colours in a standard way so that they can be reproduced on hardware and printing devices.

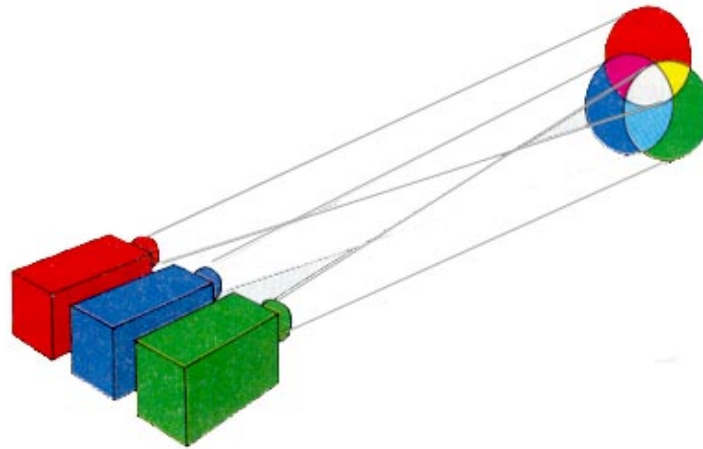
## Note

- Few (about 10) color names are widely recognized by speakers of any language.
- Color reproduction problems are created by the prevalence of digital imaging.
  - Want to ensure that everyone sees the same color

# Additive mixing

---

*James Maxwell* demonstrated that most colours can be matched by superimposing three separate light sources known as primaries; a process known as additive mixing.



The choice of the primary light sources affects the gamut of colours obtained by additive mixing.

# Additive matching

---

Exploit these findings and the theory of *trichromacy* and let colours be represented as a mixture of three fixed lights.

$$M = aA + bB + cC$$

Two people **perceive** the same color if they **agree** on a, b, and c.

The range of colours that can be matched with any three primaries is called the **gamut** of those primaries.

No three real primaries can be chosen so that their gamut includes all possible colours. If the primaries are chosen to be red, green, and blue, however, a very large number of colours can be matched.

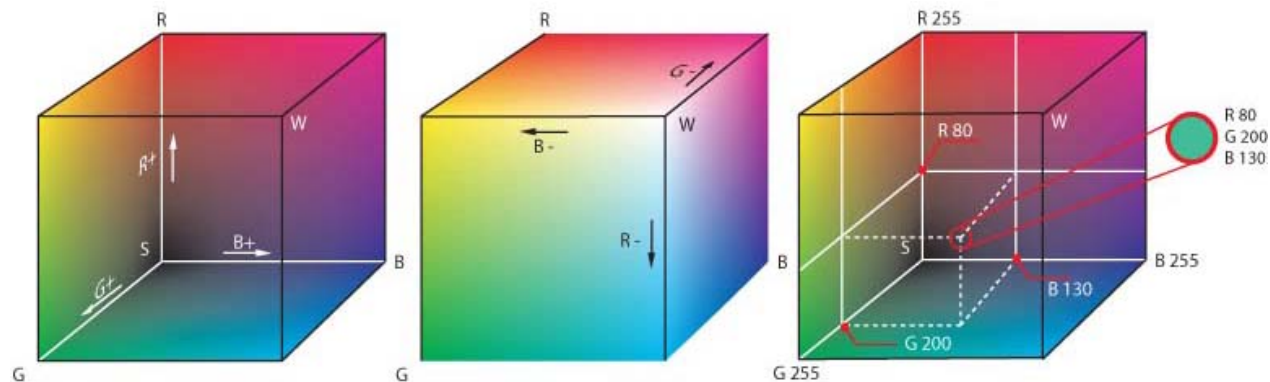
# RGB colour space

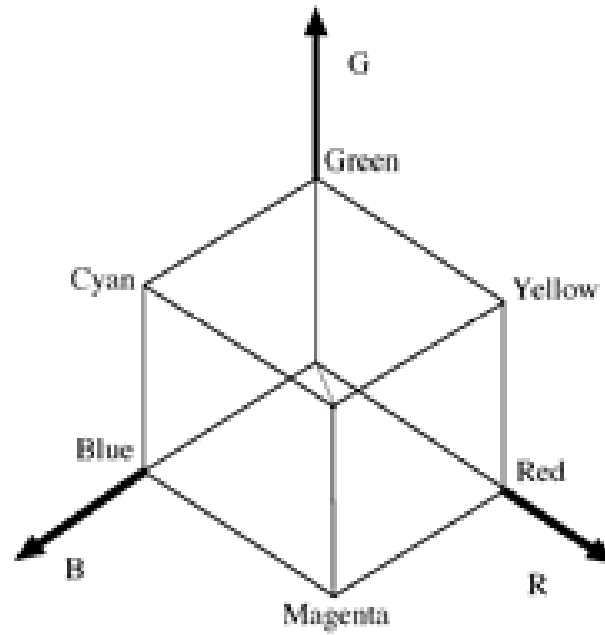
---

Primaries are monochromatic energies around

645.2nm -red light, 526.3nm -green light, 444.4nm -blue light.

R(ed)G(reen)B(lue) model used for many colour monitors and video cameras.



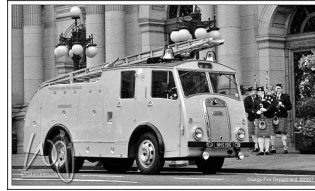


# RGB Example

---



Image



Red channel



Green channel



Blue channel

**Colour Coding of the grayscale images:**

the whiter the pixel  $\Rightarrow$  the higher the  $r$ ,  $g$  or  $b$  value

**Thus:**

The red truck has high  $r$  value and low  $g$  and  $b$  values.

While the white lamps have high  $r$ ,  $g$ , and  $b$  values.

# HSV color space

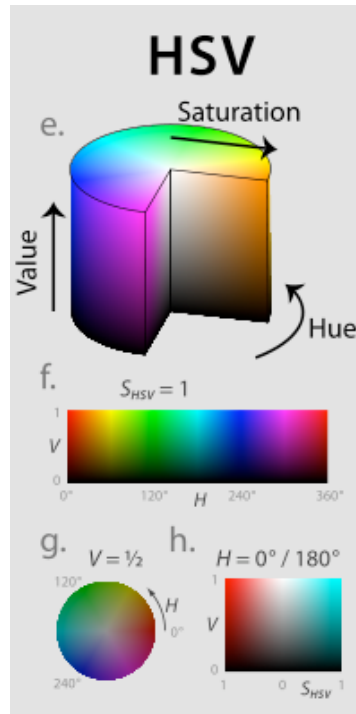
---

HSV is a cylindrical-coordinate representation of points in an RGB color model. It rearranges the geometry of RGB in an attempt to be more perceptually relevant than the Cartesian representation.

It is based on the intuitive appeal of the artist's tint, shade and tone.



Hue and saturation are non-linear functions of RGB because hue relations are naturally expressed on a circle.



$$H \in [0, 360^\circ),$$

$$S \in [0, 1] \text{ and}$$

$$V \in [0, 1].$$

Here are the HSV coordinates are some common colours:

$(-, -, 0) = \text{Black};$

$(-, 0, 1) = \text{White};$

$(0, 1, 1) = \text{Pure Red};$

$(*, 1, 1) = \text{Pure Pigment}$

The  $-$  indicates that the value in this dimension is irrelevant.



# HSV example

---

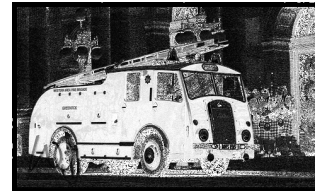
The HSV components of a colour image:



**Image**



**H (colour display)**



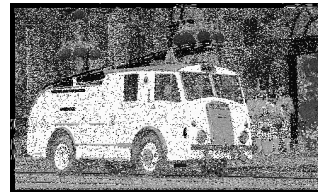
**S**



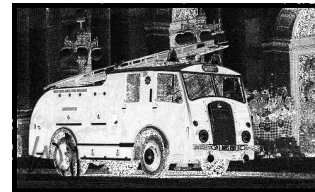
**V**



**Image**



**H (gray display)**



**S**



**V**

Below is the plot of output from Matlab's function `rgb2hsv`.

# HSV example - Hue

---

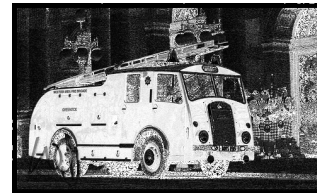
Original Image



Image



H



S



V

Change its Hue



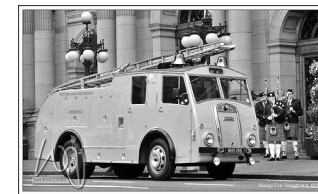
New Image



H + 180°



S



V

# HSV example - Saturation

---

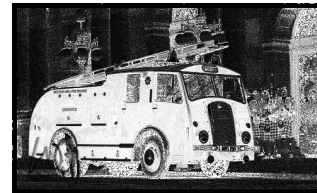
Original Image



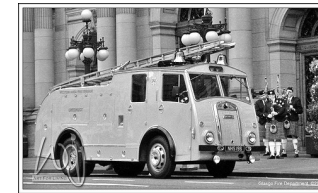
Image



H



S



V

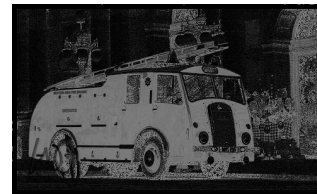
Change its Saturation



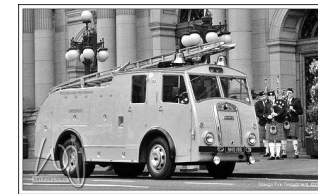
New Image



H



$\frac{1}{2}$  S



V

# HSV example - Value

---

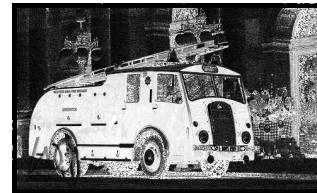
Original Image



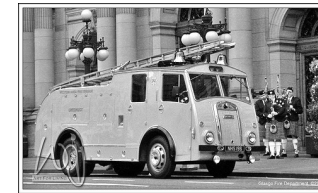
Image



H



S



V

Change its Value



New Image



H



S



$\frac{1}{2}$  V

# Today's programming assignment

# Programming assignment

---

- Details available on the course website.
- Involves running Matlab, loading and displaying images and writing a Matlab function.
- Specifically you will write a function to simulate what a colour image looks like to a person suffering from one of two specific types of colour-blindness. Don't worry; explicit instructions are given as well as background information.
- Mail me with any errors you spot in the *Exercise* notes.
- I will notify the class about errors spotted and corrections via the course website and mailing list.