433-380 Graphics and Computation

Department of Computer Science and Software Engineering, The University of Melbourne

slides by Adrian Pearce,

includes some slides adapted from Les Kitchen

and figures from Foley and Rowe texts.

Contents

1   Colour

Aim of this lecture: understand the properties of light and colour

Reading: Foley Sections 11.2 Chromatic colour and 11.3 Colour models for raster graphics

1.1   Properties of (visible) light

Light is electromagnetic radiation in the wavelength range of 400nm (blue) to 700nm (red).

- Almost all light is a mixture of wavelengths, e.g. the rainbow spectrum of white light from the sun.
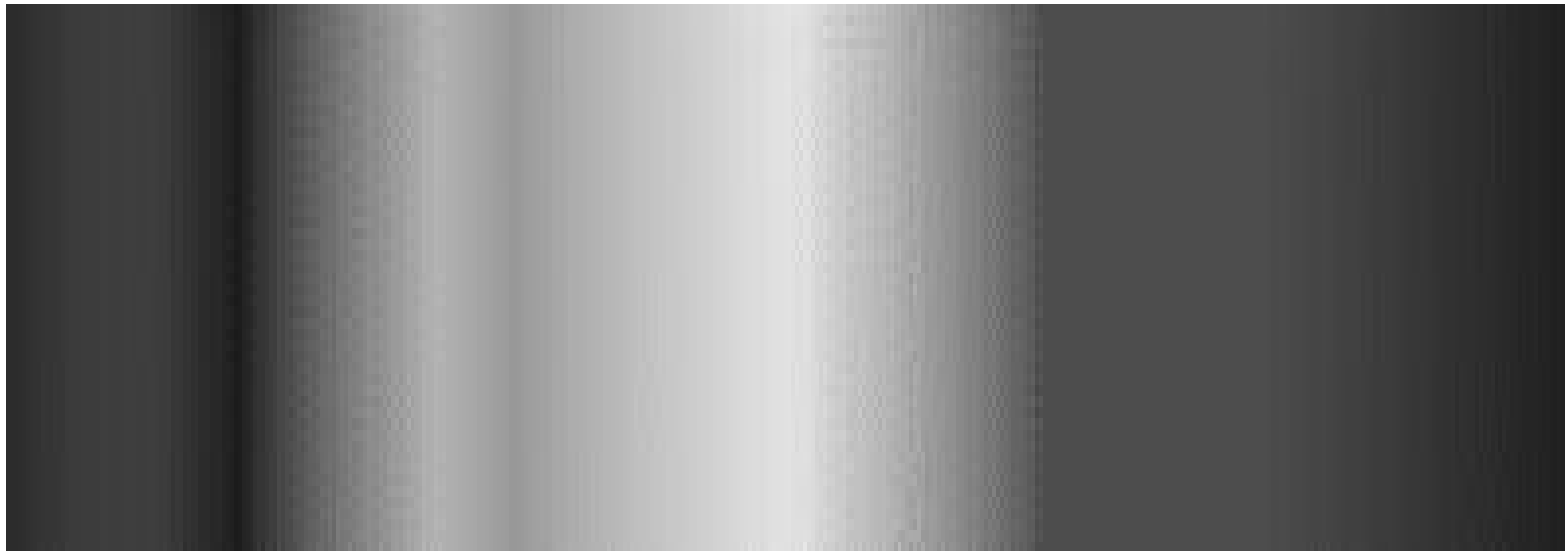
Colour vision is based on the tri-stimulus theory of colour perception where three kinds of cones are sensitive to red, green and blue light.

- Use of tri-stimulus values is known as the RGB model. Another value system is hue, saturation, brightness (HSB) model, closely related to human preprocessing of light (both RGB and HSB are typically supported in most graphics API's, including Java).

1.2   Response of human cone photo-receptors (relative)
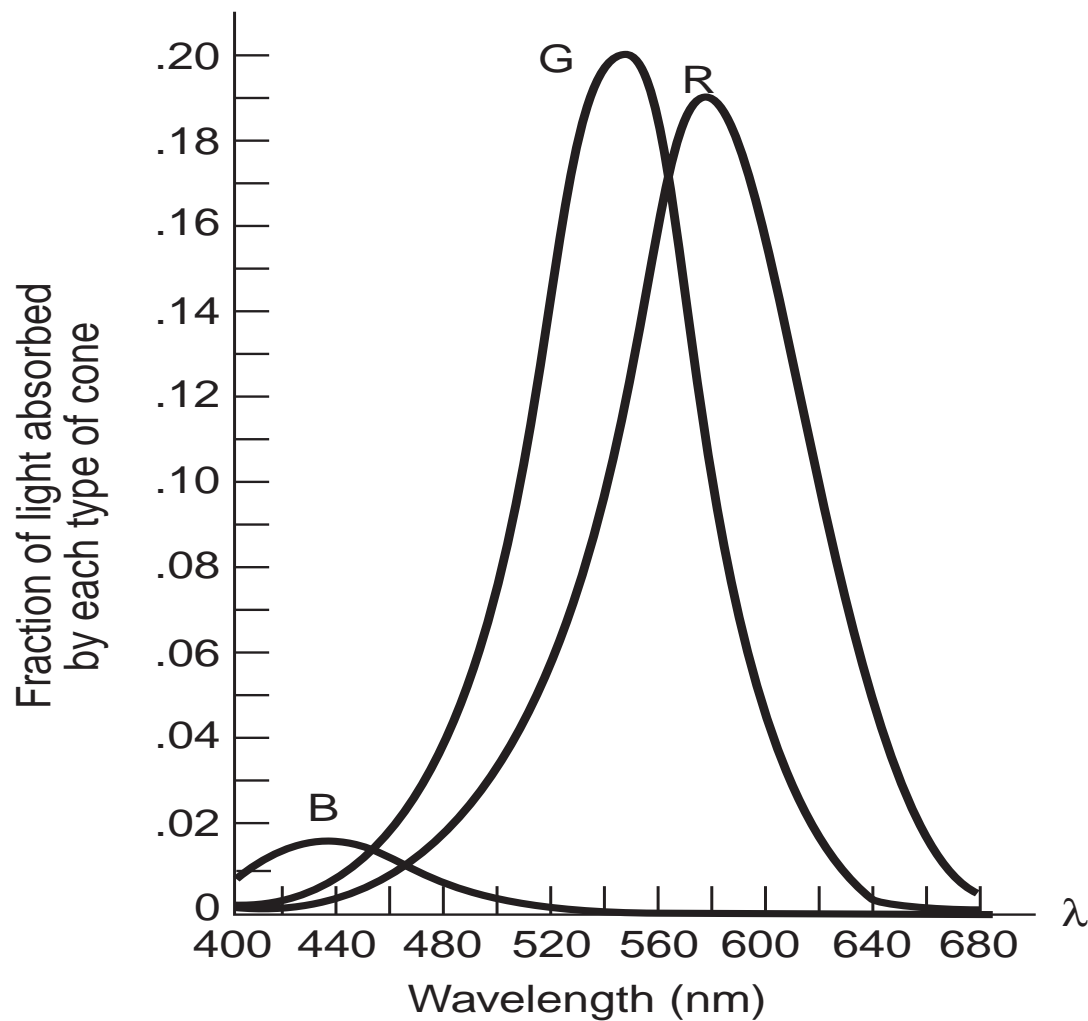
Luminous efficiency of human eye
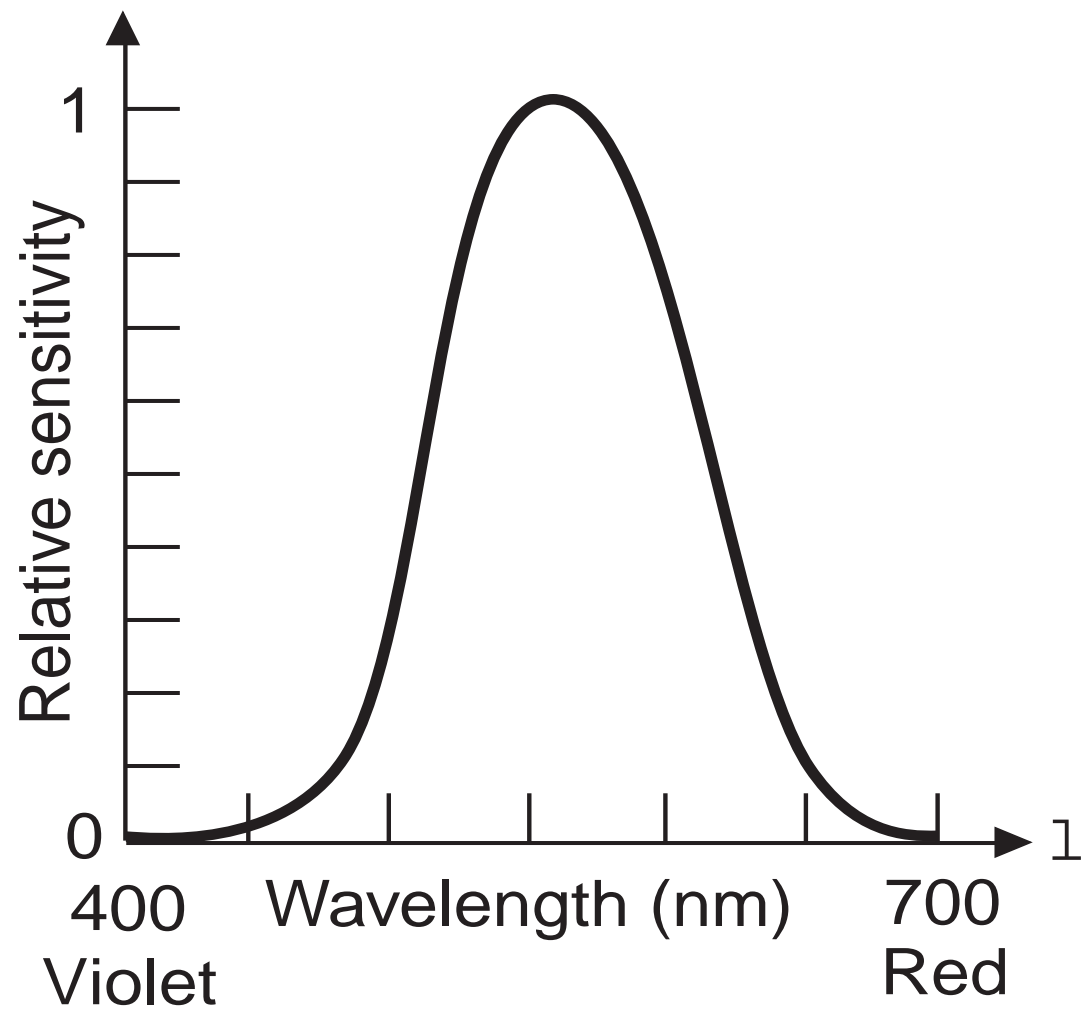
1.3   Human colour vision characteristics

Colour perception depends on colour context (e.g. what colour and intensity resides next to what other colour etc.)

● Blue alone tends to be perceived quite weakly. The greatest colour discrimination is in the green to yellow range.

People vary quite widely, even amongst people with "normal" colour vision (adjust your monitor to suit yourself).

● Forms of colour blindness include red-green, blue-yellow and achromatic, has obvious implications for the display of computer graphics.

1.4   Colour

Light is part of the electromagnetic spectrum, specifically that part which occupies the range of wavelengths λ human eyes are sensitive to, roughly speaking 400 nm (blue end of spectrum) to 700 nm (red end of spectrum).

A full treatment of light has to take this continuum of wavelength into account. A light source will actually have a continuous distribution of *power* over the spectrum of wavelengths, $p(\lambda)$.

The reflectivity of a surface, will also depend on wavelength: $r(\lambda)$.

For a given light source and material, the light intensity reflected will be the product of these two and will also depend on wavelength,

$$p(\lambda)r(\lambda)$$

For example, an intrinsically "red" material would have high reflectivity for longer (red) wavelengths, and low reflectivity for other, shorter wavelengths. Under "white" light (an uniform distribution of power across λ), it would reflect mostly red. Under "blue" light (a distribution with higher power in the blue or short wavelengths, and lower or zero power in the red (long) wavelengths, it would appear dark, since the product $p(\lambda)r(\lambda)$ would be low all the way across.

Any sensor (such as a cone in the human retina) has a response which varies according to wavelength, $h(\lambda)$. So, for any wavelength, the response of the sensor is the product of all three contributions, and the total response of the sensor is the integral over all wavelengths

$$\int h(\lambda)r(\lambda)p(\lambda)d\lambda$$

In human photopic (bright-light) vision there are three kinds of cones, "red", "green", and "blue", each with its own sensitivity, $h_R(\lambda)$, $h_G(\lambda)$, and $h_B(\lambda)$. So the initial raw response of the human visual system to a particular combination of light source and material, can be characterised by three numbers, the total response of these three kinds of cones, R, G, B.

Strictly speaking, it is still necessary to compute $p(\lambda)r(\lambda)$ in terms of continuous (or at least finely sampled) λ. However, for most purposes we can get away with sampling $p(\lambda)$ and $r(\lambda)$ at just three wavelengths, which correspond (more or less) to the peak responses of the three kinds of cones: p_R, p_G, p_B, and r_R, r_G, r_B.

The initial human perception of colour is characterised by three dimensions, the cone responses.

Toward a suitable colour system?

The Munsell and artists pigment-mixing methods involving tints, tones and shades are subjective: they depend on the human observers judgements, the lighting, the size of the object, the surrounding colour and the overall lightness of the environment.
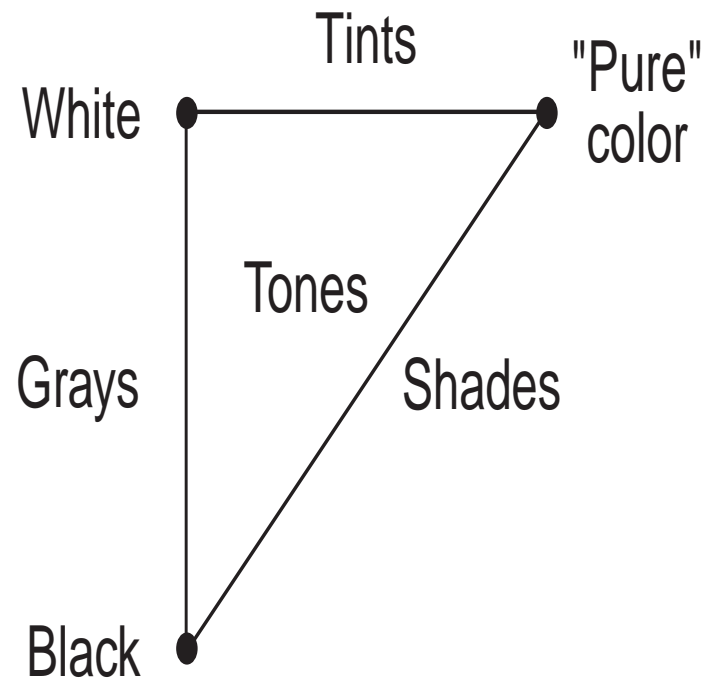
1.5   The CIE colour system

A standard way of representing this 3D colour space is provided by the CIE (Commission International de L'Eclairage). It has three coordinates, X, Y, and Z (not to be confused with the spatial coordinates). These are based on (but not the same as) the human cone responses.

If we take it that the "quality" of a colour is independent of the total light intensity, we can use the *normalised* colours

$$x = \frac{X}{X + Y + Z} \quad \text{etc.}$$

In colour space, this is geometrically the same as projecting colours onto the plane $X + Y + Z = 1$.

The CIE $\bar{x}_\lambda$, $\bar{y}_\lambda$ and $\bar{z}_\lambda$ functions
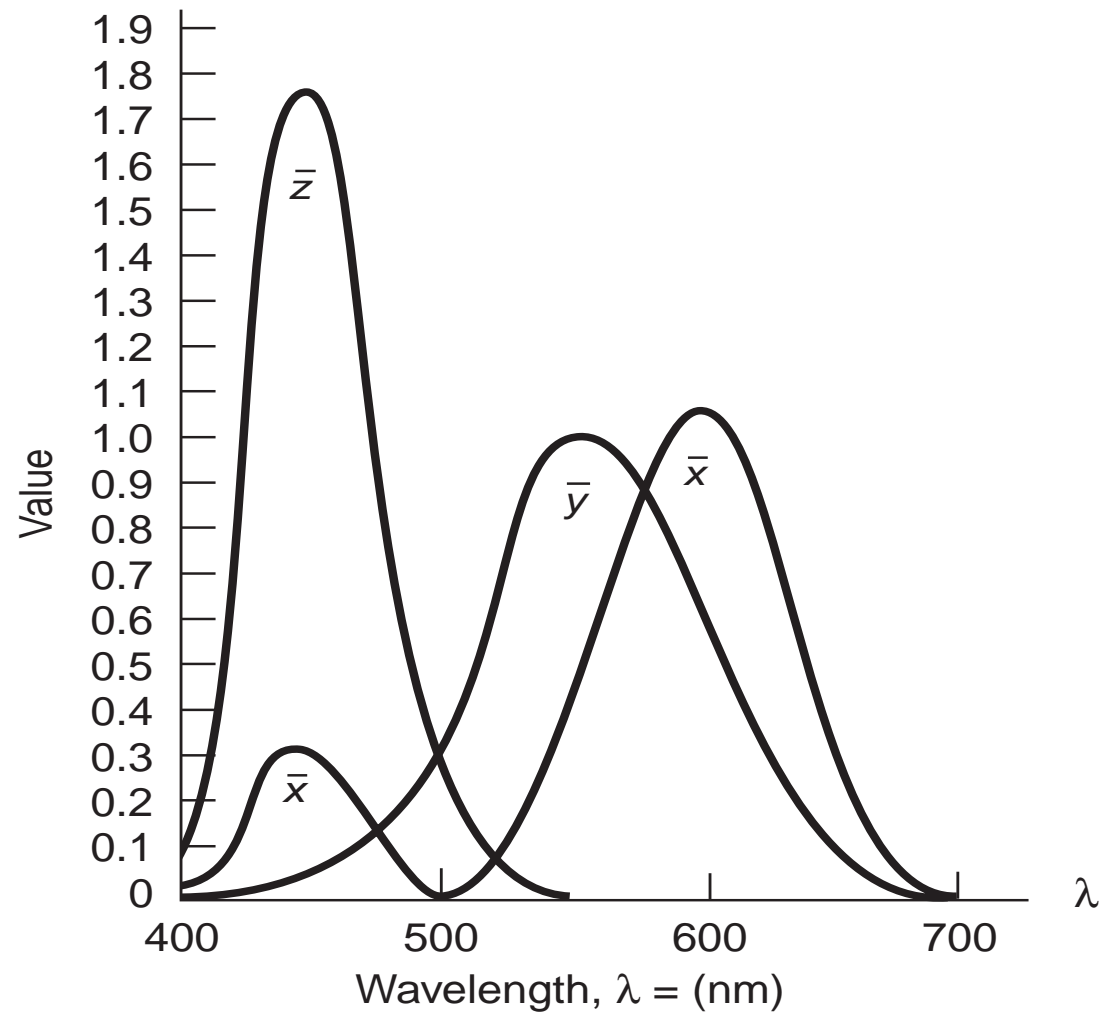
Because of the constraint $x + y + z = 1$, any two of these, conventionally x and y, are sufficient to characterise colour.

In terms of x and y, we can lay out the range of perceivable colours in the CIE chromaticity diagram.

From the theoretical CIE XYZ, we can convert into other colour coordinate systems by a linear transformation (matrix multiply). These include the RGB spaces defined by the colour phosphors used by a particular kind of colour monitor, or by the detectors in some colour video camera, and the colour coordinates used by some T.V. colour standard like NTSC.

1.6   The HSV colour model

Another way of looking at colour is in terms of *hue*, *saturation*, and brightness or intensity *value* (HSV). Value is the total intensity of light, measured by

$$V = R + G + B$$
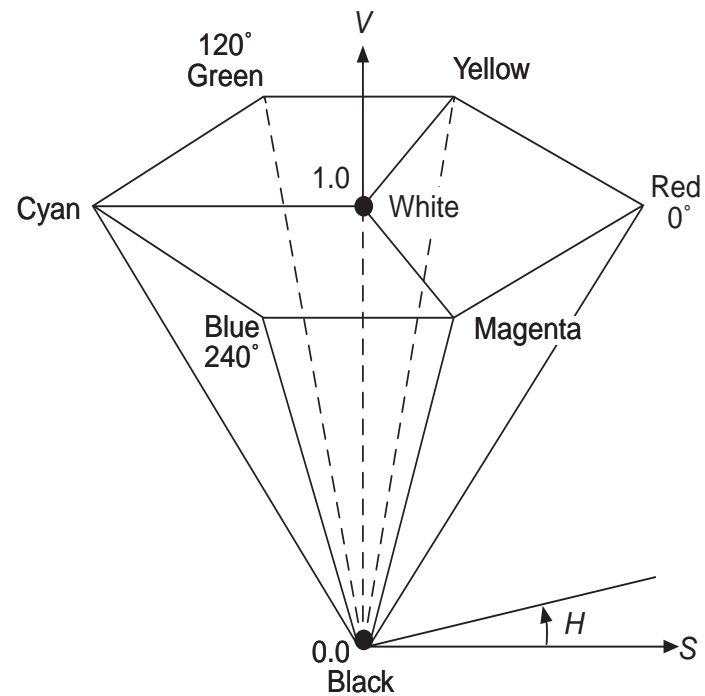
Hue and saturation are essentially polar coordinates in this plane. Saturation measures the distance out from the origin (where there are *achromatic* colours, white, grays, black) towards the periphery (where there are strong "pure" colours).

Hue measures the direction (angle from some conventional direction), and so characterises the nature of the colour independent of intensity and saturation.

This amounts to a conversion from Cartesian RGB coordinates to cylindrical coordinates about the $R = G = B$ axis.

Specification of colours using HSV
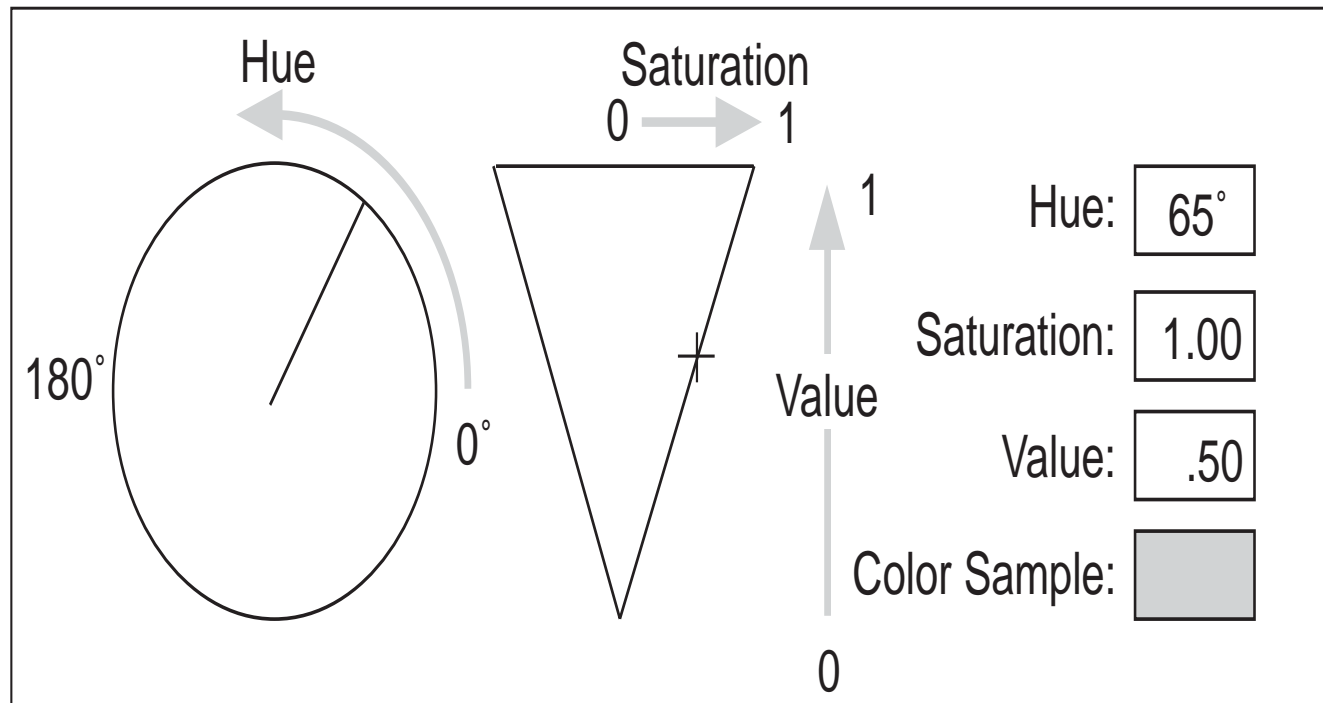
2   Images

Aim of this lecture: understand image representations and formats.

Reading: review PNM image format, available from 380 web page links.

Continuous-Domain Images

Images are two-dimensional (2D) patterns of light, that possess intensity and colour properties, and can be considered as a function (bounded by visibility) that maps a plane into some space of measurements, i.e.

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}^+$$

Digitisation

Digitisation involves sampling over a regular spatial grid, typically square or rectangular (rarely hexagonal or triangular, but occasionally in robotics).

Sampling involves quantisation over intensity values, for each of the pixels in an image. Pixels can be either,

- in linear proportion to intensity or in non-linear (adaptive) proportion (as in night vision in the human eye), and either

- continuous in value (grey-scale or colour) or binary in value (bitmap representation of 0's and 1's).

Digital images

Pixels are the elements sampled by the digitisation process. The digital image, or picture captured from each scene is subject to trade-offs, including

- resolution (fineness of sampling or quantisation) versus fidelity, and consequently

- fidelity versus storage and processing costs.

2.1   Image formats

Typically *header* plus *pixel data*

Header may be fixed record structure, or variable (e.g. attribute/value list), or some combination.

Internal versus *external* formats

Depends whether the purpose of the format is for the interchange of images (e.g. streaming of images or video over the Internet) or for optimised, random-access during memory and processor operations.

Image header information

Mandatory information:

- Magic number (that indicates what format image is, generally used in conjunction with file suffix)

- Pixel type such as bit, byte or integer (sometimes unsigned) float, double or complex (imaginary numbers are used by the Fourier transform for smoothing images)

- Image size (in rows and columns)

Optional information:

- Statistics such as minimum and maximum values of intensities or colour values (very useful for display, means don't have to parse entire image first)

- Additional geometrical information such as aspect-ratio or origin offset

- Ownership, program that created image and name of organisation or company

Image pixel data

Images are generally saved in order of raster scanning, like mapping a two-dimensional (2D) array into linear storage (used by computer memory).

- Normally left-to-right then top-to-bottom, however sometimes in other orders (always check before making assumptions, when given images to work with)

- Pixels may be either be stored using (natural machine units) bytes, integers, floats or (packed) bit fields that need to be converted (unpacked) before displaying.

- Often contain padding, to make them an even size for optimised file input and output in many graphics application programming interfaces (API's).

Variations of pixel data

Images vary in the number of dimension according to their sources and can be either multi-dimensional or one-dimensional.

- In Eric Grimson's enhanced-reality medical surgery, both three-dimensional (3D) nuclear magnetic image (NMR) images and motion sequences are involved

Multi-band images include multiple values at each pixel, the most common being the colour red, green, blue (RGB) and hue, saturation and brightness (HSB) formats.

- Multi-spectral remote sensing data is being continuously collected from low-orbiting satellites for geographic applications, such as vegetation or salination information or even spying.

Question: can you think of an example of a one-dimensional image (1D), or even a zero-dimensional (0D) image?

Complex and structured image formats

More complex pixel formats maintain structural information about

image regions or other spatial or temporal (Fourier) qualities.

- include quad-trees, run-length encoding,

- support non-rectangular images of arbitrary shape (however the transparent *gif* format used on web pages cheats by using background overlays), and

- (of course) image compression.

Some image formats use structured picture-description languages, including Postscript, PDF (the format of these lecture slides) and scalable vector graphics (SVG) format frequently used for web pages.

Question: how can a static digital pixel image have (pseudo) temporal qualities?

2.2   Image coordinate systems

Coordinate system used by Java and text (fig1.3)

Increasing x

(0,0)

$y = b$

Increasing y
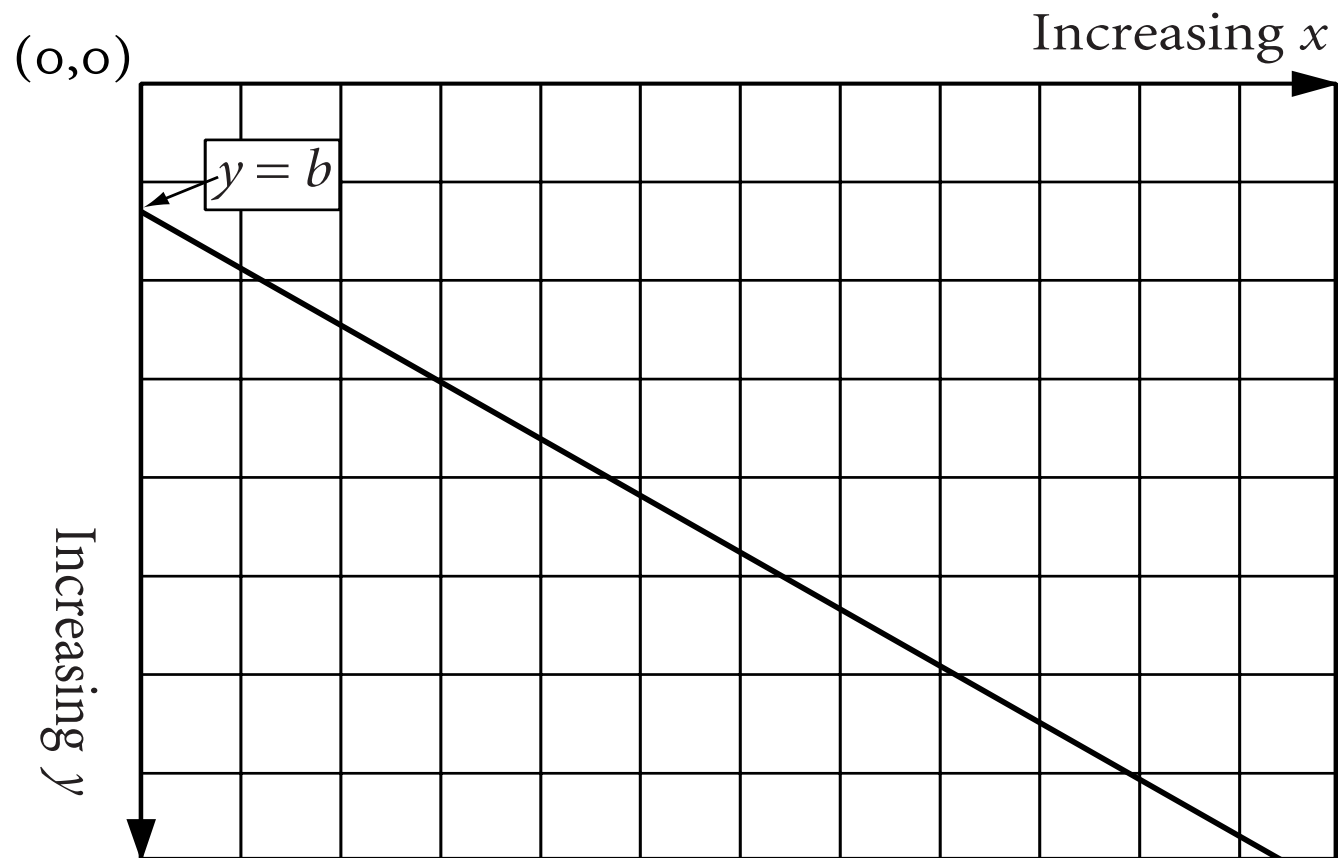
Cartesian coordinates

Beware, Cartesian coordinates differ from those used in Java and many other computer graphics API's. Traditional Cartesian coordinates are as follows,

- (x, y) are integers

- origin is at bottom left

- x increases rightwards

- y increases upwards

- zero-origin

Cartesian coordinates are simple to relate to analytical geometry.

Matrix coordinates

Matrix coordinates are conventional in mathematics and physics.

- (i, j) row i, column j

- origin at top left

- i increases downwards

- j increases rightwards

- zero-origin or one-origin

Matrix coordinates are not as easy to relate to analytical geometry.

Offset coordinates

Offset coordinate systems are often used for specific or internal image format applications, where it is important to set up a relative coordinate system for an image's data.

- involves embedding an image into an absolute coordinate system,

- can measure the offset by various conventions $\pm$

- can be implemented in any of the other coordinate systems