Introduction to Computer Graphics
Human visual perception & colour
CGRA 151

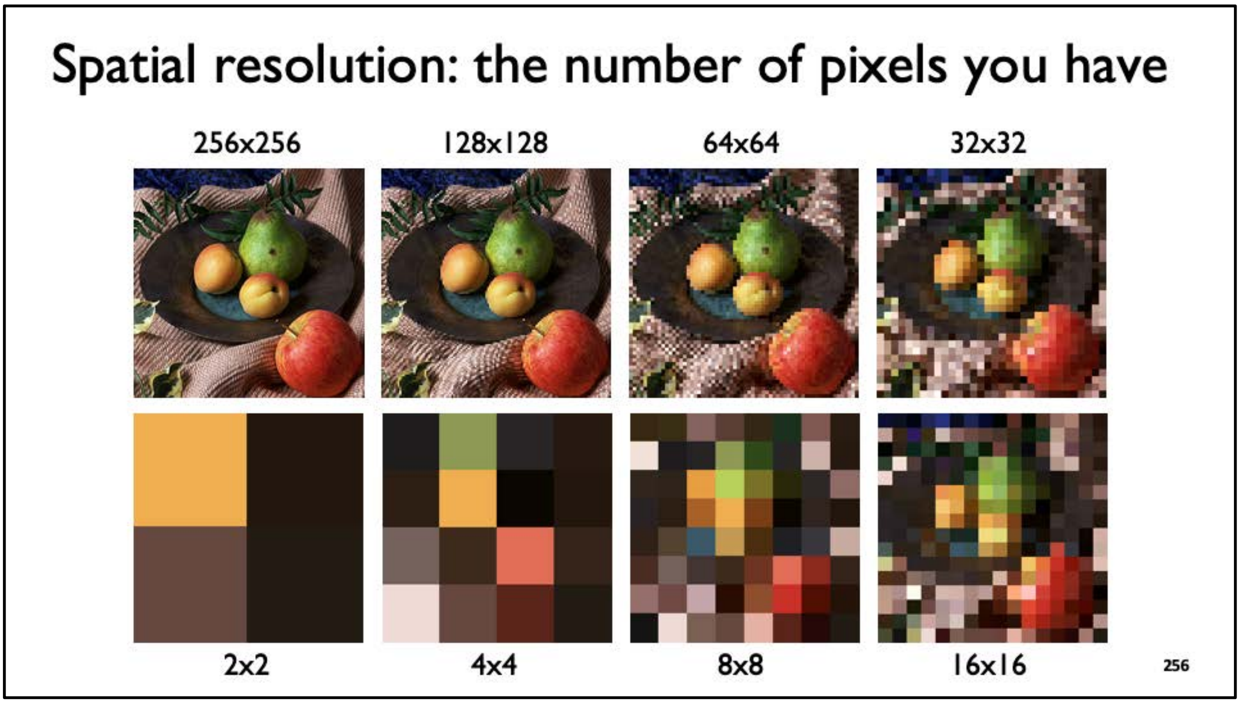
Prof. Neil Dodgson
neil.dodgson@vuw.ac.nz

255

Image from http://pixar.wikia.com/wiki/Inside_Out

The five emotions in the film "Inside Out" represent something about what goes on in our brains. This part of the course considers how the eyes and brain together handle visual perception.

The five emotions are coloured with the five principal colours according to Munsell's theory of colour (from around 1910): red, yellow, green, blue, violet. We will look at the many ways in which colour is represented and why something seemingly so simple turns out to be wonderful and complex.



This slide appeared much earlier in the course. Here it is just reminding us that resolution relates to the number of pixels that you have.

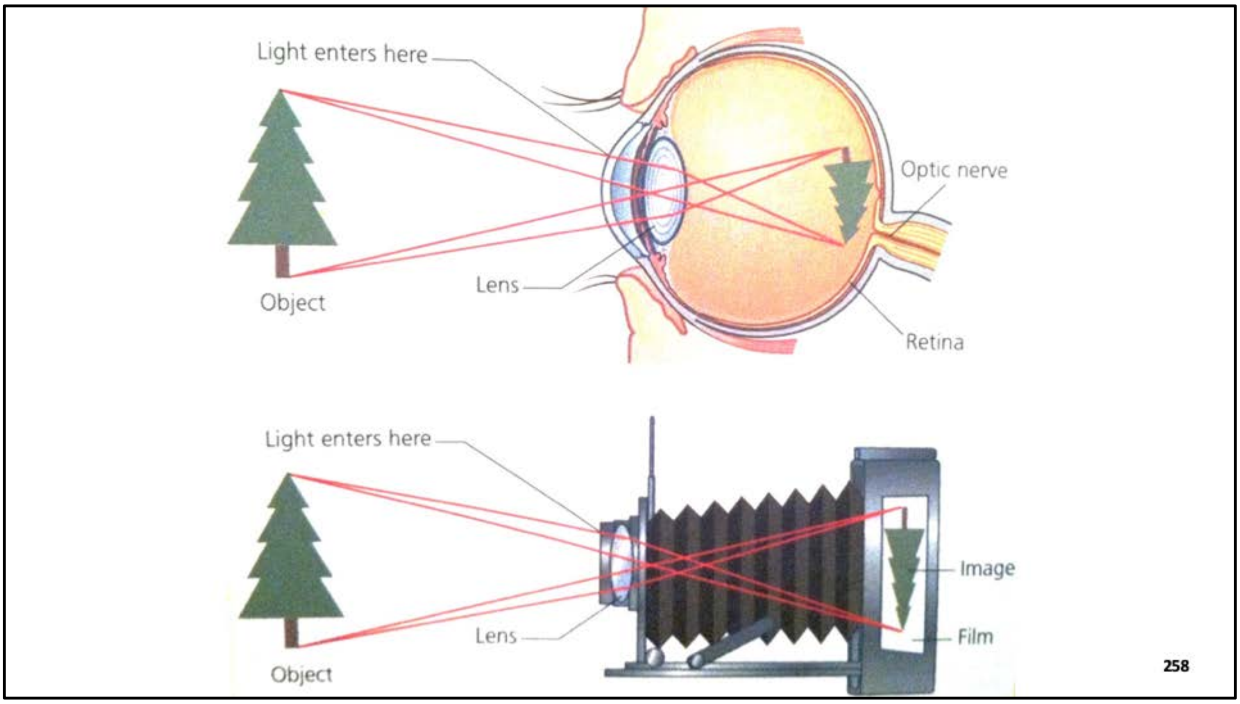
However, what becomes important at this stage is how many pixels there are for a given distance in space.

We measure this in pixels per inch. Displays are usually between 100 and 400 pixels per inch, with each pixel able to display a wide range of colours. Printers tend to be between 600 and 2500 pixels per inch, but each pixel on a printer can only be on or off, either ink is there or it is not. You should ask yourself how you would create a range of shades of colour when you have pixels that can be only on or off. We will look at that later in the course.

As well as pixels per inch, we can also be interested in pixels visible across a given angle. For example, in terms of the angular resolution, a 400 pixel per inch mobile device held 25 centimetres from your eye is as good as a 100 pixel per inch TV viewed from 100 centimetres, because an individual pixel covers the same visual angle.

All visual computer graphics output is for consumption by human eyes. We need to make our display devices sufficiently good that a human eye cannot see the individual pixels, but our displays do not need to be any better than that: because a human being would not be able to tell.

In order to find out what resolution we need (and also to find out how colour vision works), we need to know more about the working of the human eye.



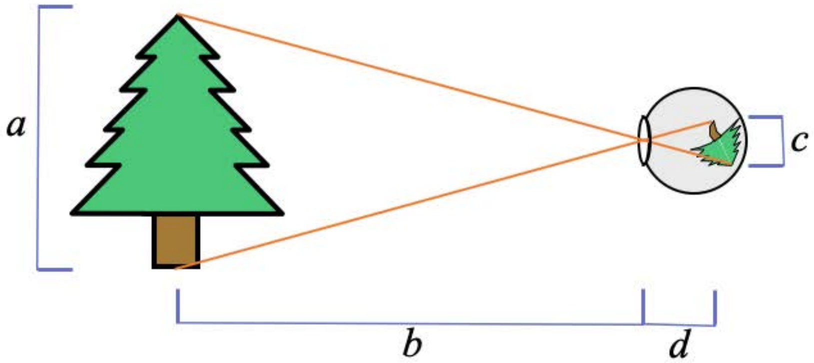
http://discoveryoursolutions.com/images/camera_eye.jpg

The human eye works in a similar way to a camera. A lens focuses the image onto an imaging surface. In an old fashioned film camera, the imaging surface is photographic film. In a modern phone camera, the imaging surface is a light-sensitive silicon chip with millions of light detectors in a rectangular grid. In the human eye, the imaging surface is the curved back of the human eye, called the *retina*.

Do not worry that the image is upside down on the back of the eye. This is the correct way that eyes work. The image goes through a great deal of processing before it is sent off to be interpreted by the brain, as we'll find out.

What is the smallest thing you can detect?

$$\frac{a}{b} = \frac{c}{d}$$



259

If we want to find out how small our pixels need to be out in the real world, then we need to know the size of the smallest light detector in the back of our eyes.

The mathematical equation at the top is based on the similar triangles in the diagram. The two triangles are called *similar* because they have exactly the same angles, which means that corresponding distances are in proportion. Thus the height a is to the distance b in the same proportion as the height c is to the distance d.

Structure of the human eye

- the retina is an array of light detection cells
- the fovea is the high resolution area of the retina
- the optic nerve takes signals from the retina to the visual cortex in the brain

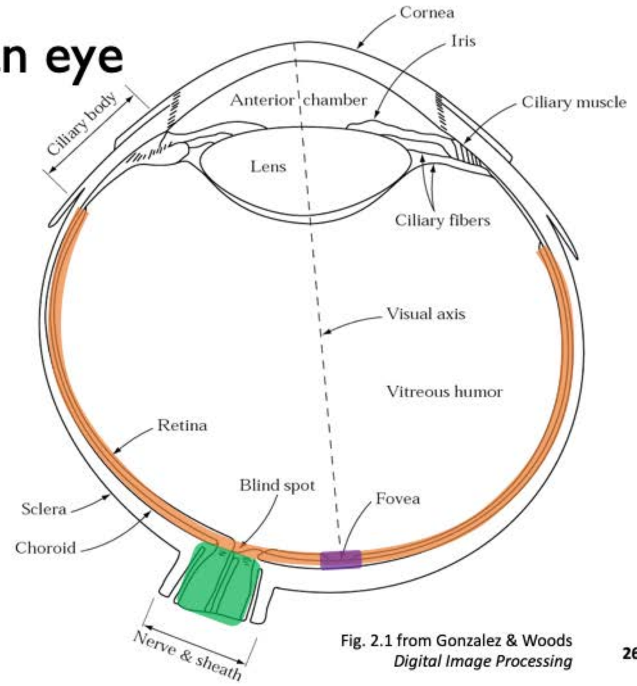


Fig. 2.1 from Gonzalez & Woods *Digital Image Processing*

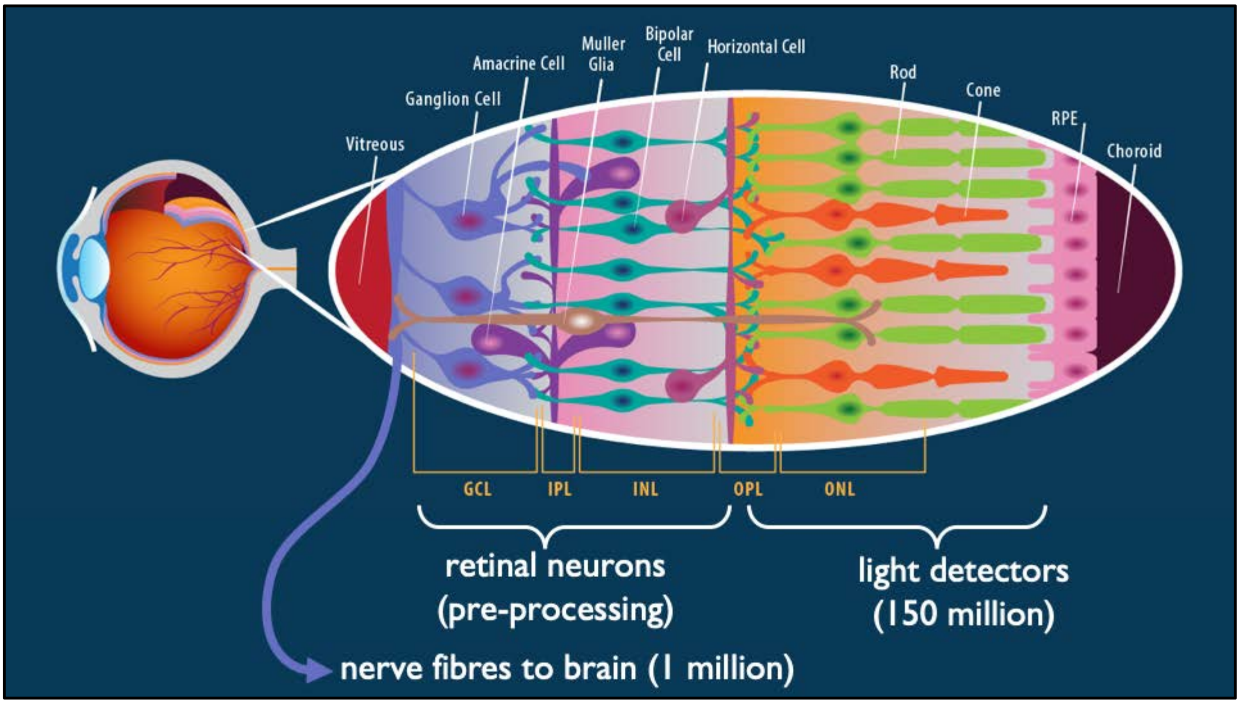
260

Light enters the eye from the top. It passes through the *cornea*, a tough transparent membrane that protects the eye. The *lens* focusses the image onto the back of the eye. The *ciliary fibers* and *muscles* pull on the lens to change its focal length, allowing you to focus on near objects or on far objects or anything in between. The *iris* is the pretty coloured part of the eye that you can see from outside. It can change to make a bigger or smaller hole (the *pupil*) to let more or less light through. It can change in diameter from 2 mm to 8 mm, making a difference in light let through of 1:16.

The image is formed on the retina, which comprises about 150 million light detection cells.

The fovea is the highest resolution part of the retina and is the bit of the eye on which falls the light from the thing that you are currently looking at.

The eye does a lot of pre-processing on the signals from the light detection cells. It is the signals from this pre-processing that are sent down the optic nerve to the brain.

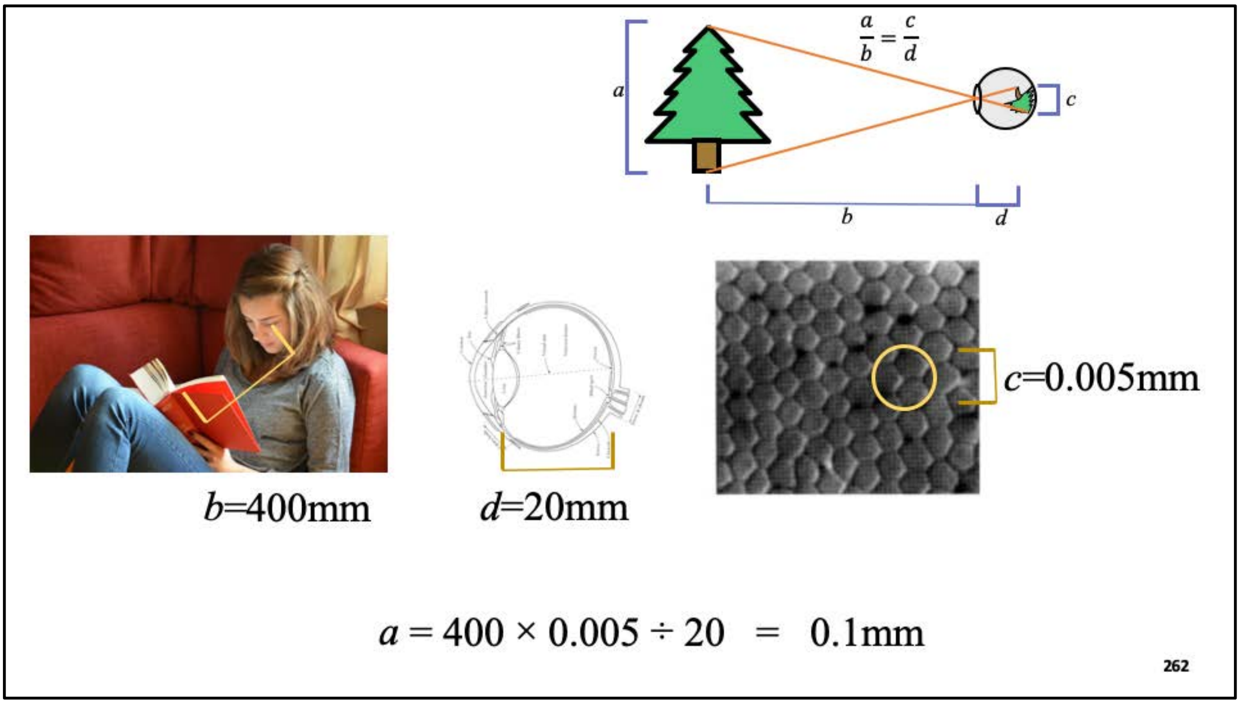


http://media.discovery.lifemapsc.com/pub/uploadedFiles/images/The_Cellular_Structure_of_the_Retina.png

This highly coloured diagram shows the structure of the retina. On the right are the rods and cones, which are the light detection mechanisms. They are long and thin and oriented in the same direction as the light beams. Light enters them from the left, the length of the rods and cones helps to maximise the possibility of a photon of light interacting with the detector.

There are 150 million light detectors in each eye. However, there are only about one million nerve fibres going to the brain, so the 150 million signals from the light detectors have to be reduced to one million signals.

In front of the light detectors are up to four layers of neurons, which do pre-processing on the signals. They combine the signals from several light detection cells to reduce the number of signals that need to be sent to the brain. The number of detectors combined into a single signal increases dramatically as you move away from the fovea. Most light detectors have input to multiple retinal neurons.



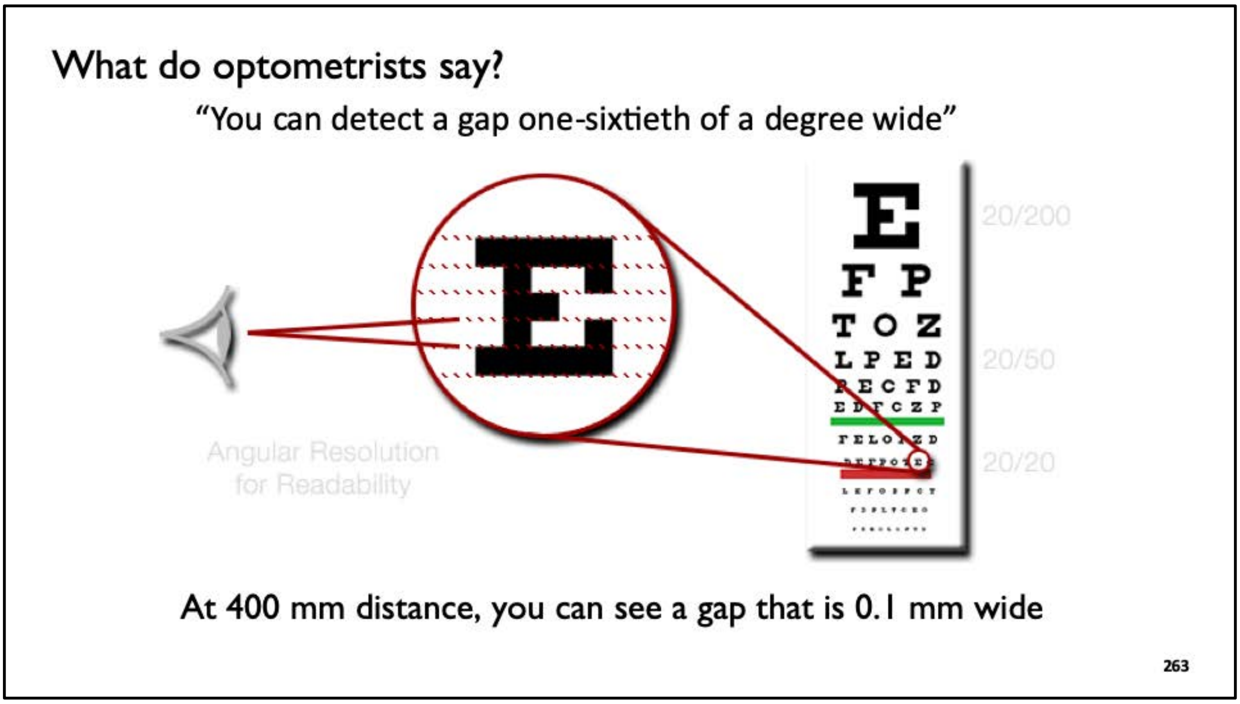
In your fovea, your retinal neurons will average together the signals from three or four cone light detectors to create a single signal going to the brain. This is the smallest thing that you can detect.

There are 150,000 cones per square millimetre in the fovea. That means that each cone is about 2.6 micrometres across*. Combining three or four together makes an area about 5 micrometres across, or 0.005 mm.

The eyeball is about the size of a ping pong ball. Both are about 20 mm across.

If you are reading text at a distance of 400 mm, then the calculation tells us that the smallest feature that you can distinguish is about 0.1 mm across.

* How did we calculate this? We are told that there are 150,000 cones in a single square millimetre of the fovea. This means that, if we put cones side-by-side, there would be $\sqrt{150{,}000} = 387.3$ cones across one millimetre. The size of one cone is therefore $1/387.3 = 0.00258$ millimetres, which is 2.582 micrometres, which (as this is just an approximation) we can happily round to 2.6 micrometres.



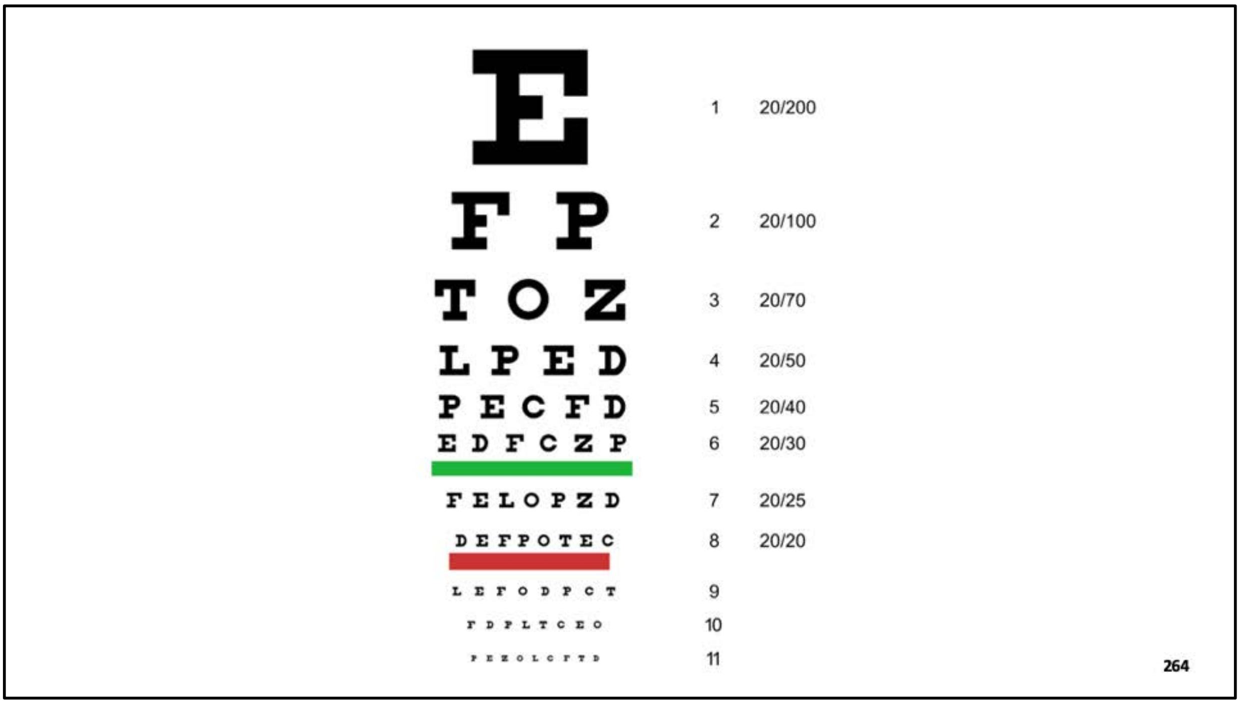
https://commons.wikimedia.org/wiki/File:Snellen_chart.svg

Rather than doing all that calculation, we could simply have asked an optometrist.

Scientific observation of human beings’ vision shows that the smallest resolvable white gap between two black lines is one-sixtieth of a degree wide (0.0003 radians).

The eye chart is designed with letters that have stems (the black parts) and gaps that are all the same width, so that you can read the characters if they are bigger than a certain size, but not if they are smaller. Someone with perfect vision is able to distinguish characters where that width is one-sixtieth of a degree wide, which is 0.1 mm at a distance of 400 cm, or 1mm wide at a distance of 4 metres.

Degenerative diseases, short- or long-sightedness reduces the visual acuity of the eye, generally by blurring the image that is presented to the retina. That means that people with poor vision do not have such good resolving power.



https://commons.wikimedia.org/wiki/File:Snellen_chart.svg by Jeff Dahl, based on the public domain document.

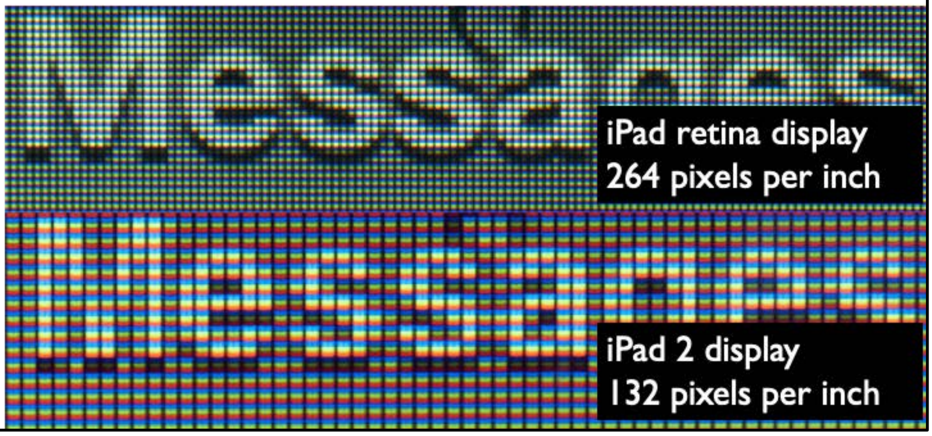
This is a typical Snellen chart. Originally developed by Dutch ophthalmologist Herman Snellen in 1862, to estimate visual acuity.

To use the chart, you need to position it at a distance away from you so that the text on line 8 subtends a visual angle of 5 minutes (i.e., 5 sixtieths of a degree = one twelfth of a degree).

Optometrists use a version of this chart that has line 8 being 0.35 inch high (8.9 millimetres), positioned at a distance of 20 feet (6 metres) from the viewer. The acuity ratings at right are all of the form 20/X, such as 20/20 or 20/200. The optometrist works out which is the lowest line that the viewer can read accurately and that gives their visual acuity. Snellen defined a "standard viewer" who is someone able to read line 8 from 20 feet. A visual acuity of 20/X means that a person can read, at 20 feet, text that a standard viewer could read from X feet, so 20/40 means that, from a distance of 20 feet, the viewer can only read text (line 5) that could be read by a standard viewer from twice that distance. In the USA 20/200 is the boundary to be considered legally blind.

Monitor resolution in pixels per inch

- At 400 mm distance, you can see a gap that is 0.1mm wide:
  ⇒ you need at least 10 pixels per millimetre
  ⇒ about 250 pixels per inch
- Retina displays, about 300 pixels per inch
- Conventional displays, about 100 pixels per inch

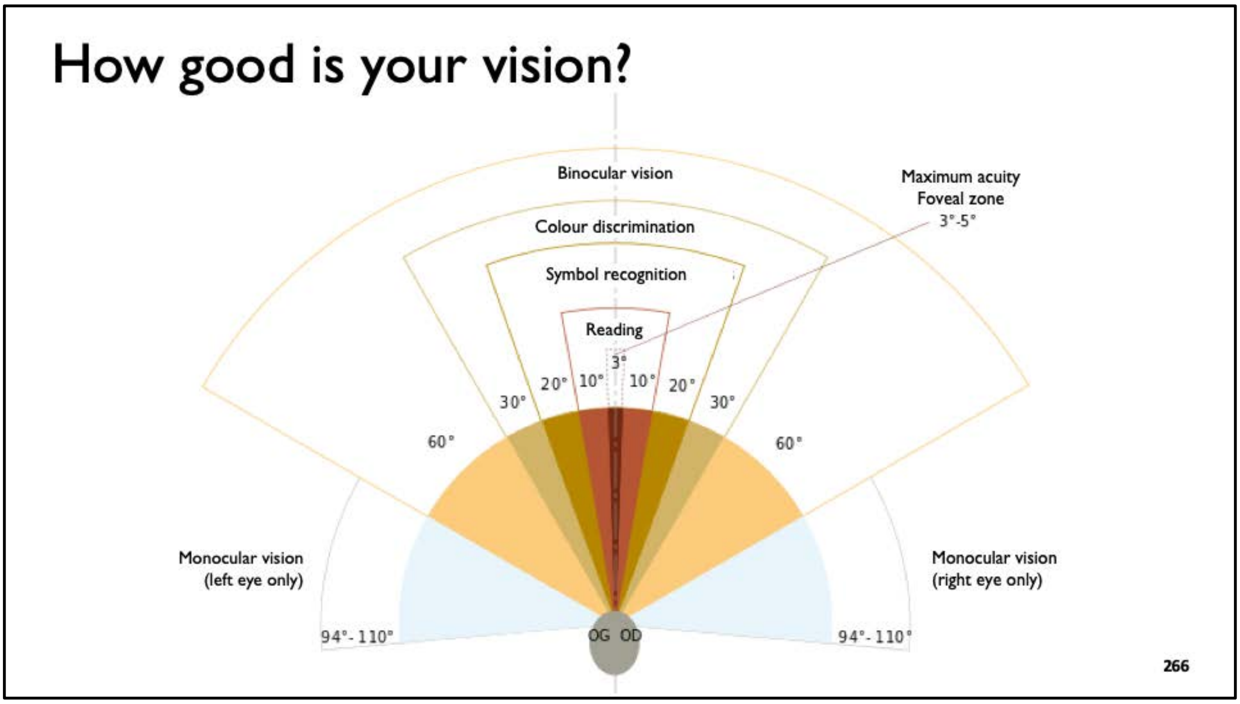


Until 2012, computer monitors had resolutions around 100 pixels per inch.

The early Macintosh (1984) had a resolution of 72 pixels per inch. The IBM PC had a similar resolution. For thirty years only small progress was made in resolution, so that by 2011 we had 132 pixels per inch on the state-of-the-art iPad 2.

A step change came with the introduction of the "iPad retina", which had a "retina display": one that matched the resolving power of the human retina. The difference was considered so profound that it was the main selling point of the new iPad, even getting included in the name of the display.

Today, there are displays with up to about 400 pixels per inch. The limitations on display resolution are in the physics of how the devices are constructed and operated.



http://www.lemonteam.com/wp-content/uploads/2014/02/blog_hmd_pic1.jpg

You get highest resolution only in the very centre of your vision.

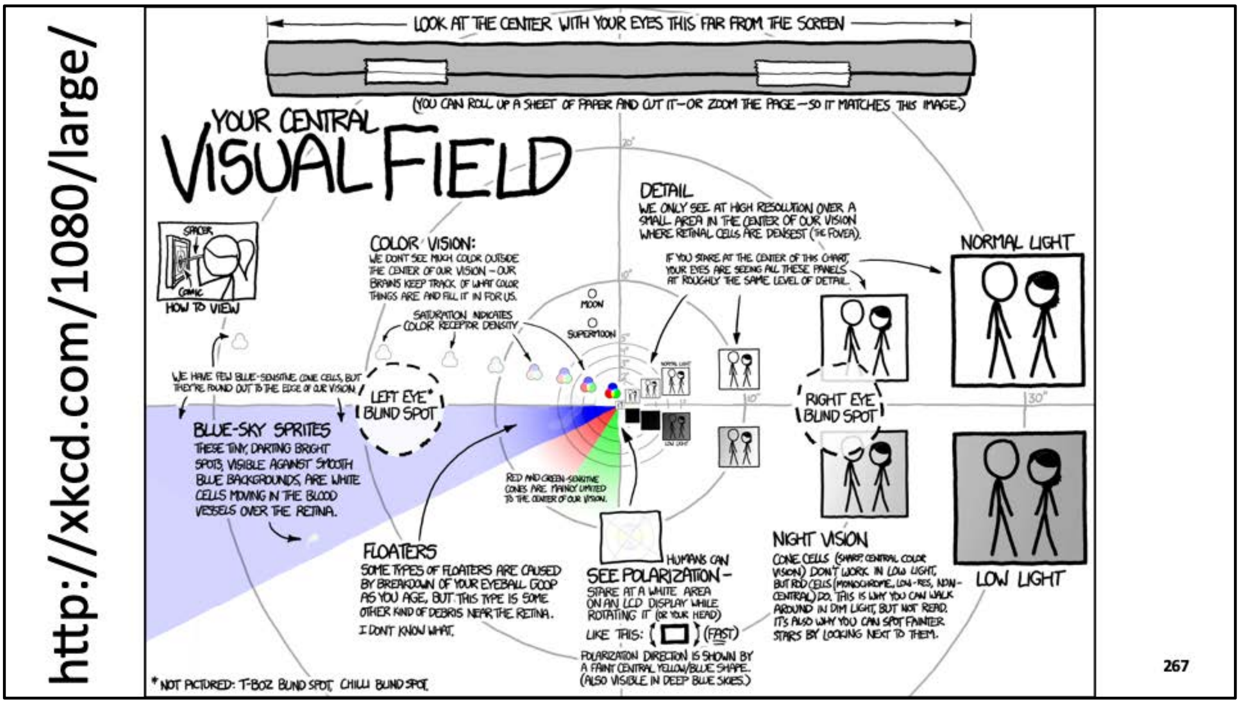
There is a slightly broader zone in which you can read normal sized text.

Symbols can be distinguished in a wider zone still.

Good colour discrimination goes out to 30 degrees either side of where you are looking.

Binocular vision (3D vision), which needs you to see the same thing with both eyes covers a range of 120 degrees: 60 degrees either side of centre.

Beyond that you can only see with one eye, but you need that peripheral vision to help provide context for the central high resolution part, and to warn you that Ug, the Caveman from the neighbouring tribe, is creeping up on you to hit you over the head with a rock.



http://xkcd.com/1080/large/

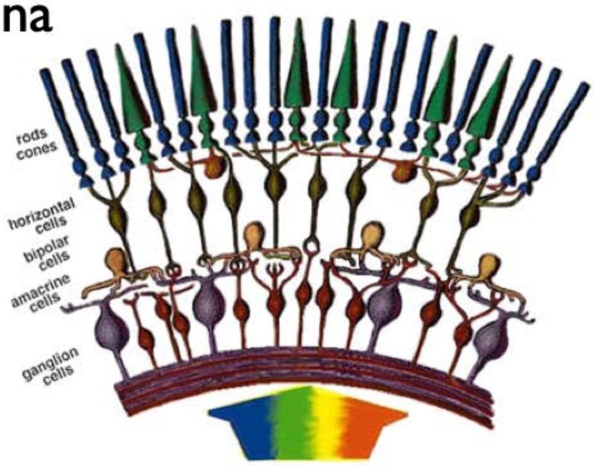
This xkcd comic illustrates several features of the human eye.

The most important are:

- Our ability to see detail decreases rapidly as we move out from the centre.
- Colour perception decreases significantly away from the centre.
- The rods are more sensitive to very low light level than the cones. There are only cones in the centre of vision so, counter-intuitively, in very dark conditions you see better away from the centre of vision.

Light detectors in the retina

- two classes
  - rods
  - cones
- cones come in three types
  - sensitive to short, medium and long wavelengths
  - allow you to see in colour
- the cones are concentrated in the macula, at the centre of the retina
- the fovea is a densely packed region in the centre of the macula
  - contains the highest density of cones
  - provides the highest resolution vision



268

Notice that the cones are described as being sensitive to short, medium and long wavelengths. They are not "blue, green and red" cones. In fact, the medium and long wavelength cones have very similar responses to light. More on this in the lectures on colour.

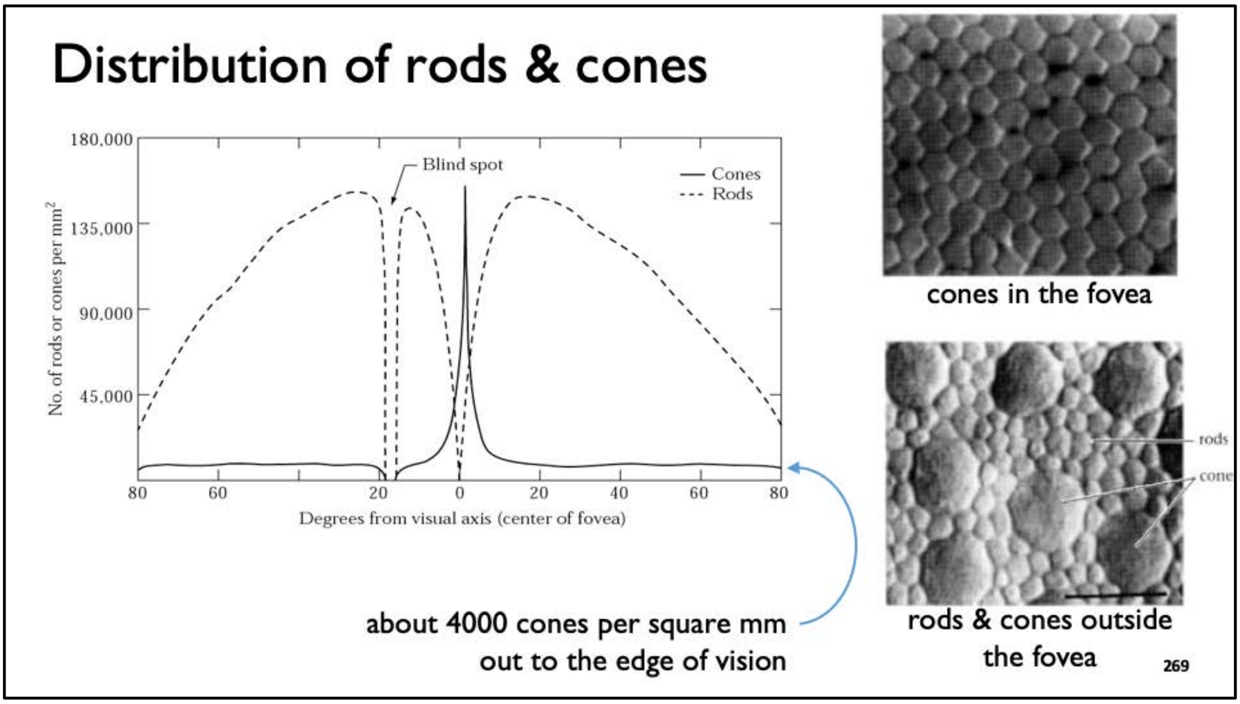


Fig. 2.2 from Gonzalez & Woods
www.cis.rit.edu/people/faculty/montag/vandplite/pages/chap_9/ch9p1.html

(1) cones in the fovea are squished together more tightly than outside the fovea: giving higher resolution vision;
(2) as the density of cones drops the gaps between them are filled with rods;
(3) cones come in three types, that allow colour vision (see later);
(4) rods come in one type, and it responds to light across the whole visible spectrum;
(5) rods work only in low light levels but they keep working in darker situations than cones;
(6) as we move away from the centre, more and more light receptors get grouped together into a single signal to the brain;
(7) colour vision does not disappear as you move to the edges, it just gets very poor as more light detectors get clumped together into a single signal.

Some of the processing in the eye

- discrimination
  - discriminates between different intensities and colours
- adaptation
  - adapts to changes in illumination level and colour
  - can see about 1:100 contrast at any given time
  - but can adapt to see light over a range of 10^{10}
- persistence
  - integrates light over a period of about 1/30 second
- edge detection and edge enhancement
  - visible in e.g. Mach banding effects

270

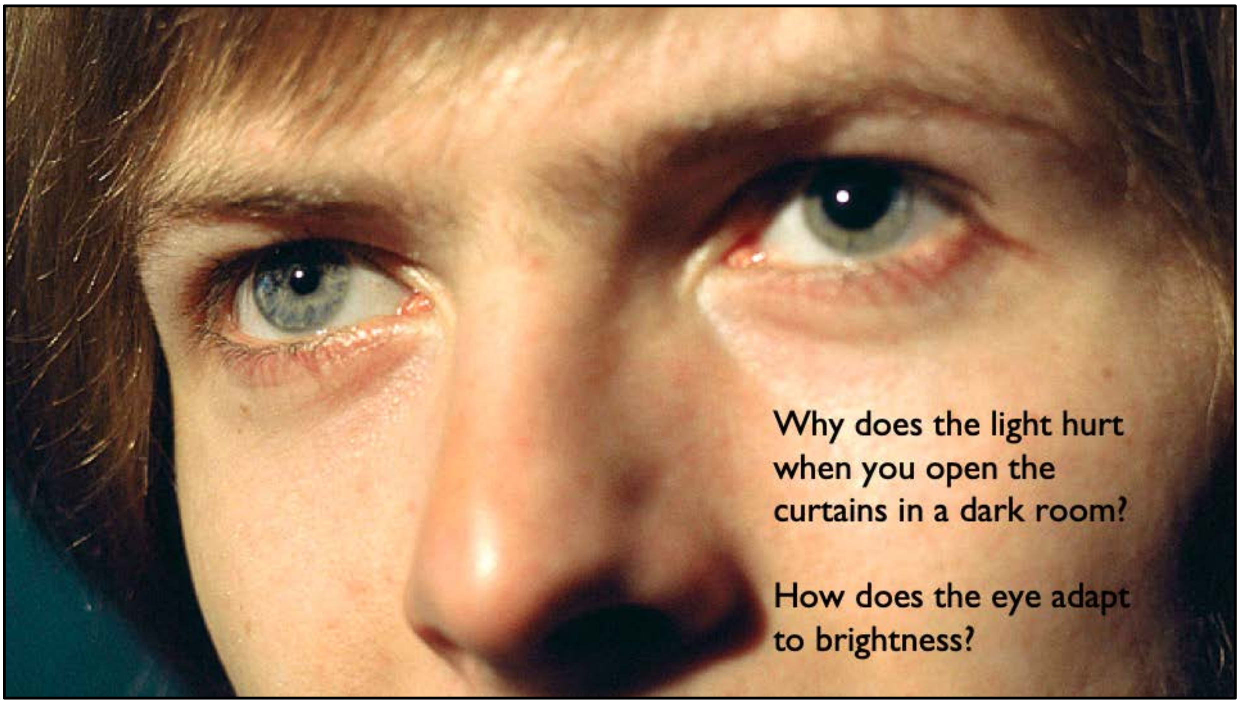
Intensity adaptation

271

https://upload.wikimedia.org/wikipedia/commons/b/b3/HDRI-Example.jpg

When you look at the stained glass window, your eyes adapt to the brighter light and you see the glorious colours of the light shining through the window.

When you look into the corner of the roof, your eyes adapt to the lower light and you see the structure of the wooden framing.



Your pupil can have a diameter between 2 mm and 8 mm, this gives you a 1:16 ratio of light coming into your eye, instantly adjustable by the pupil. This is what the eye uses for instant adjustments as you move from looking at the bright stained glass windows to the dark corner of the room.

However, 1:16 is nowhere near enough to cope with the vast differences between a dark bedroom and the bright sunlight beyond the curtains. There are chemical processes that handle those big changes but they take a few minutes to adjust.

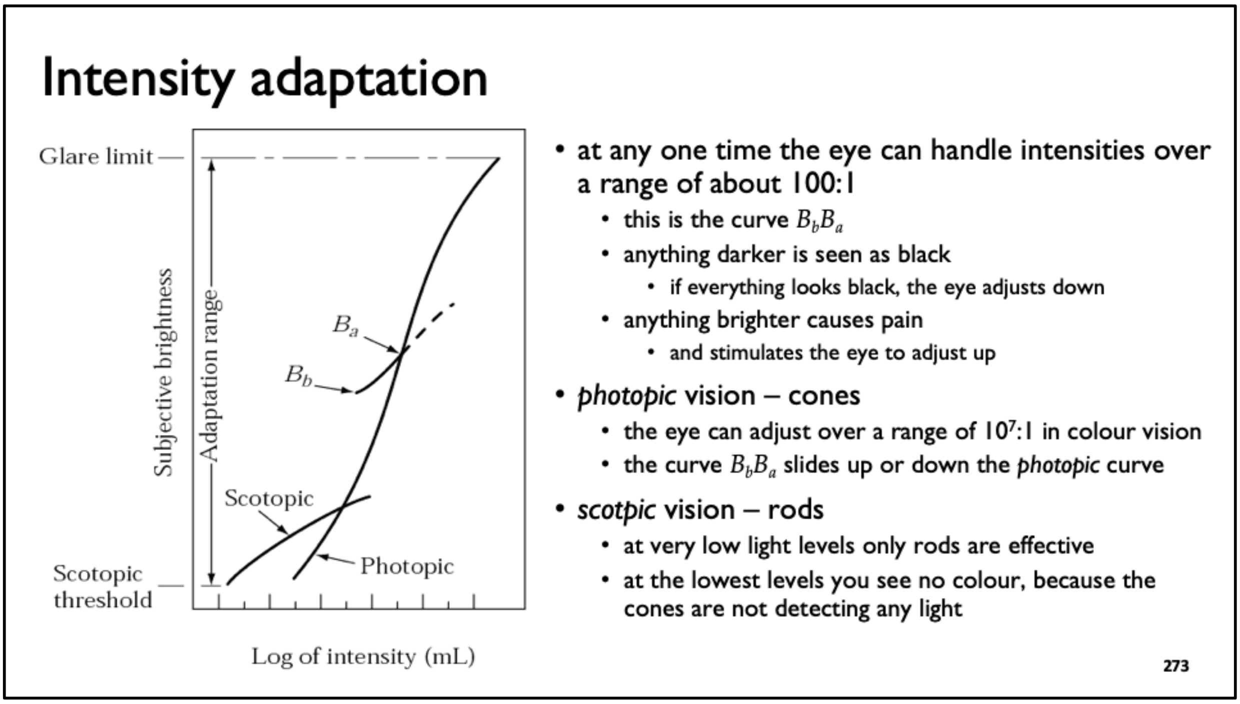


Fig. 2.4 from Gonzalez & Woods

At any one time your eye can handle a range of intensity of about 100 to 1. The brightest thing that you can see will be 100 times brighter than the darkest. Anything darker than that will look black.

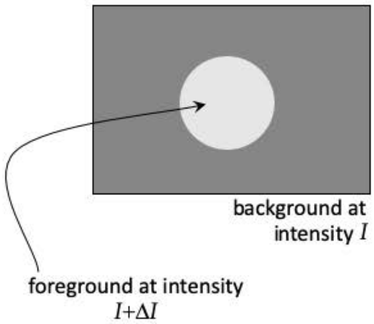
The eye handles bigger changes in brightness by adjusting the chemical balances in the back of the eye. If you look at something that is too bright, and your pupil is already as small as it can get, you get a sensation of pain and the chemicals in your eye start to adjust to handle the brighter light. This takes several minutes.

There is a limit to how bright you can handle. Your eyes cannot handle looking directly at the sun, and doing so will physically damage your eyes.

Equally, when you turn out the lights, your eye takes several minutes to adjust the chemicals to allow you to see in the dark. There is a limit to how dark you can handle, but it is incredibly low.

Intensity differentiation

- the eye differentiates between different colours and different intensities
- Weber's Law describes how well the eye can distinguish different intensities using *just noticeable differences*



for a range of values of I

- start with $\Delta I=0$
  increase ΔI until human observer can just see a difference
- start with ΔI large
  decrease ΔI until human observer can just *not* see a difference

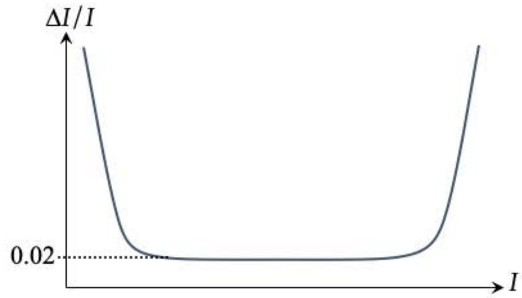
274

You can easily tell the difference between blue and yellow, between dark grey and light grey, but how fine a difference can you distinguish?

The experiment described on the slide allows you to find the just noticeable difference: the minimum difference that a human is able to distinguish

Intensity differentiation

- results for a "normal" viewer
  - a human can distinguish about a 2% change in intensity for much of the range of intensities
  - discrimination becomes rapidly worse as you get close to the darkest or brightest intensities that you can currently see



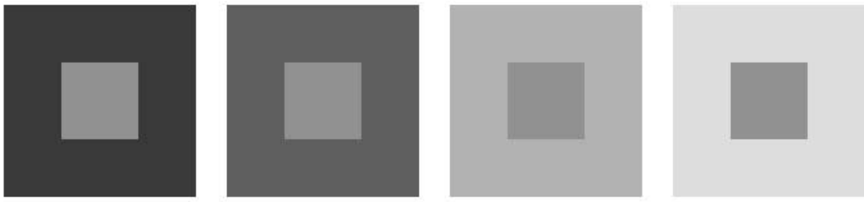
275

Once the eye has adapted to the current range of intensities, it has good differentiation in the middle of that range (from dark grey to bright grey), where it can distinguish a 2% change in intensity between two patches of grey.

However, at the dark end, where things look black, it has a hard time distinguishing different shades of black. Equally, at the bright end, it is challenging to distinguish between two slightly different whites.

Simultaneous contrast

- the eye performs a range of non-linear operations
- for example, as well as responding to changes in overall light, the eye responds to local changes



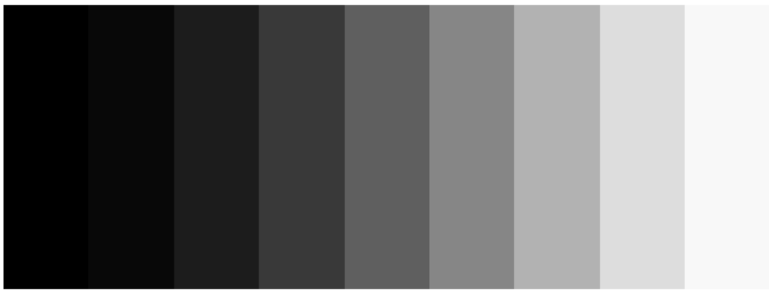
The centre square is the same intensity in all four cases but does not appear to be because your visual system is taking the local contrast into account

276

When preparing this slide, I checked in Powerpoint and can confirm that the centre square truly is the same intensity in all four cases. That's not what your brain tells you.

Mach bands

- show the effect of edge enhancement in the retina's pre-processing

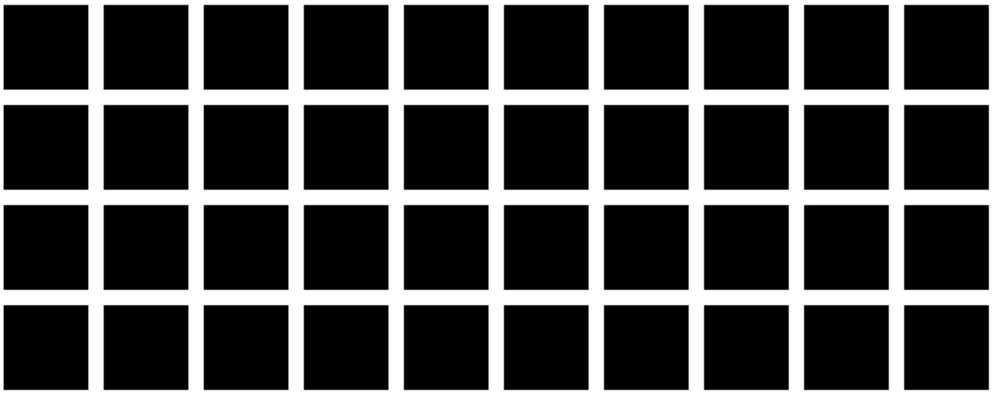


Each of the nine rectangles is a constant colour but you will see each rectangle being slightly brighter at the end which is near a darker rectangle and slightly darker at the end which is near a lighter rectangle

277

Ghost squares

- another effect caused by retinal pre-processing
  - the edge detectors *outside* the fovea cause you to see grey squares at the corners where four black squares join
  - the fovea has sufficient resolution to avoid this "error"



278

The retinal neurons combine the signals from the rods and cones in various ways, taking larger and larger groups as you move away from the fovea.

One way in which the signals are combined is in simple edge detection, detecting where there is a boundary between something bright and something dark.

Look at one of the junctions in the image. You can see that it is white. But the junctions away from your central vision have ghostly grey squares that disappear when you look directly at them.

The fovea's edge detectors are high resolution, so you see no grey squares at the junction you look directly at. The edge detectors further away from central vision are low resolution and they combine in the brain to indicate that there is something going on where all of those edges meet. It is probably significant that this sort of pattern does not occur in nature.

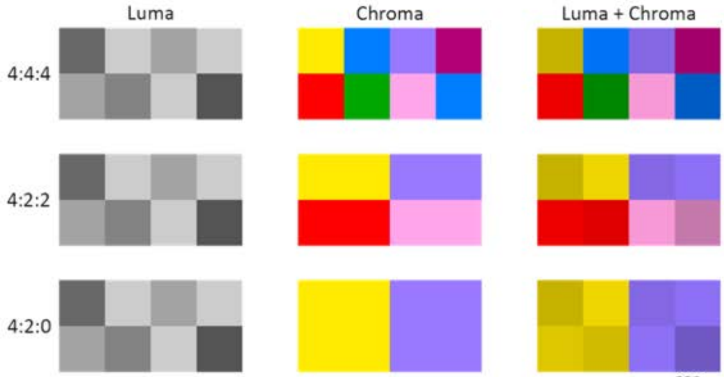
Summary of what human eyes do…

- sample the image that is projected onto the retina
- adapt to changing conditions
- perform non-linear pre-processing
  - makes it hard to model and predict behaviour
- combine a large number of basic inputs into a much smaller set of signals
  - which encode more complex data
    - e.g. some of those signals encode the presence of an edge at a particular location with a particular orientation; others encode simpler things such as the intensity at a particular location
- pass pre-processed information to the visual cortex
  - which performs extremely complex processing

279

Implications of vision on resolution

- the acuity of the eye is measured as the ability to see a white gap, 1 minute wide, between two black lines
  - about 250 pixels per inch at a viewing distance of 40 cm
  - this corresponds to about 2 cone widths on the fovea
- resolution decreases as contrast decreases
- colour resolution is much worse than intensity resolution
  - this is exploited in encoding video — allocate fewer bits to the colour information (see examples at right)



280

What does this mean for computer displays?

Remember: 1 minute is 1/60th of a degree

A phone is viewed from a comfortable distance of about 30 cm, so does not need much more than about 300 pixels per inch.

A desktop monitor is viewed from a comfortable distance of about 75 cm, so does not need much more than about 120 pixels per inch, hence why we were happy with monitors of about 100 pixels per inch in the years before tablets and smart phones.

A television is viewed from a comfortable distance of about 3 metres. That means it does not need to be much more than about 30 pixels per inch, hence why we were happy with TVs that had a resolution of only 640×480 pixels, so long as they were smaller than a 26-inch diagonal.

What happens when we want a bigger TV? An HDTV has a resolution of 1920×1080 pixels. From 3 metres away, this is going to be good enough for any size TV up to about 73 inches.

What about 4K Ultra HD? That's 3840×2160 pixels. That's good up to about 150 inches, but is overkill for anything below about 70 inches.

Implications of vision on quantisation

- humans can distinguish, at best, about a 2% change in intensity
  - not so good at distinguishing colour differences
- we need to know what the brightest white and darkest black are
  - this determines the contrast ratio: the brightness of the brightest white divided by the brightness of the darkest black
- ⇒ 12–16 bits of intensity information
  - assuming intensities are distributed linearly
    - this allows for easy computation
  - 8 bits are often acceptable, except in the dark regions

Display type	Contrast ratio
Cathode Ray Tube (CRT)	200:1
LCD digital projector	200:1
DLP digital projector	500:1
Movie film	1000:1
LCD monitor	1000:1
Modern TV	4000:1

281

TV used to have dreadful contrast. The TV screen would look grey when off and the only reason that you thought it could show black was because the brightest white it could show was about 200 times brighter than the grey background, so the grey looked black when compared to the white.

Movie film was better. It could do 500:1 to 1000:1 contrast. Watching a movie in a darkened theatre helped too because the screen reflected any ambient light from the room, thus reducing the contrast ratio. Remember: a movie screen is actually white, so when you see black on the screen, it is only black when compared to the brighter things being projected around it.

DLP digital projectors (discussed later in the course) can now produce images nearly as good as film. Digital projectors started to be adopted in cinemas from about 2008, and now over half the world's cinemas use digital.

Modern TVs, using such technologies as OLEDs and quantum dots, are able to achieve stunningly good contrast ratios. An OLED TV can be described as an infinity:1 contrast ratio because when the OLED is turned off it is truly black. However, in any real world environment, there will be at least some external light (or reflected light) splashing onto the screen which will mean that black is not true black.

Implications of vision on refresh rate

- The human eye integrates light over a period of some fraction of a second
- Displays and films work by producing successive images at a rate faster than the eye can see

- Flicker fusion threshold
  - 24Hz — traditional number of frames per second for film
  - 50 or 60Hz — used by television for decades
  - 90Hz — needed for virtual reality applications

282

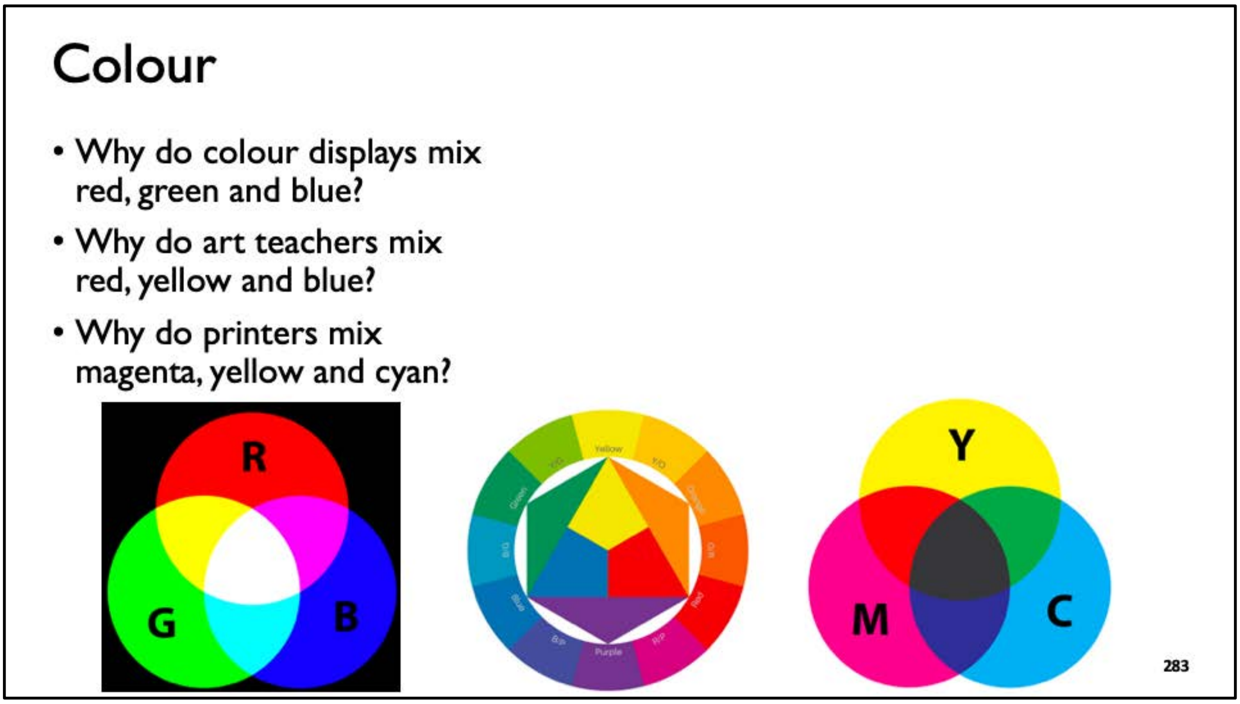
http://www.artvalue.fr/auctionresult--marey-etienne-jules-1830-1904-negative-film-strip-of-a-horse-1889377.htm
The image at right is of a sequence of photos of a horse taken in the 1890s, one of the earliest examples of a sequence of frames that could be projected quickly to give the illusion of motion.

The flicker fusion threshold is the frequency at which an intermittent light stimulus appears to be completely steady to the average human observer.

Many things contribute to the ability to see flicker. Your rods integrate more quickly than your cones, so you see flicker better outside your zone of central vision.

Traditional movie film caught images at 24Hz. This is why movies were called "the flicks". To remove the annoying flicker, movie projectors were rigged to display each image three times, so flicking at 72Hz, but only changing the image at 24Hz. This gives very little flicker and the illusion of continuous motion.

Televisions refreshed their screens at 50Hz (UK and NZ) or 60Hz (USA) but only refreshed the full image at 25Hz or 30Hz.



These questions are designed to make you think about why we do things the way we do – they are answered later in these lectures.

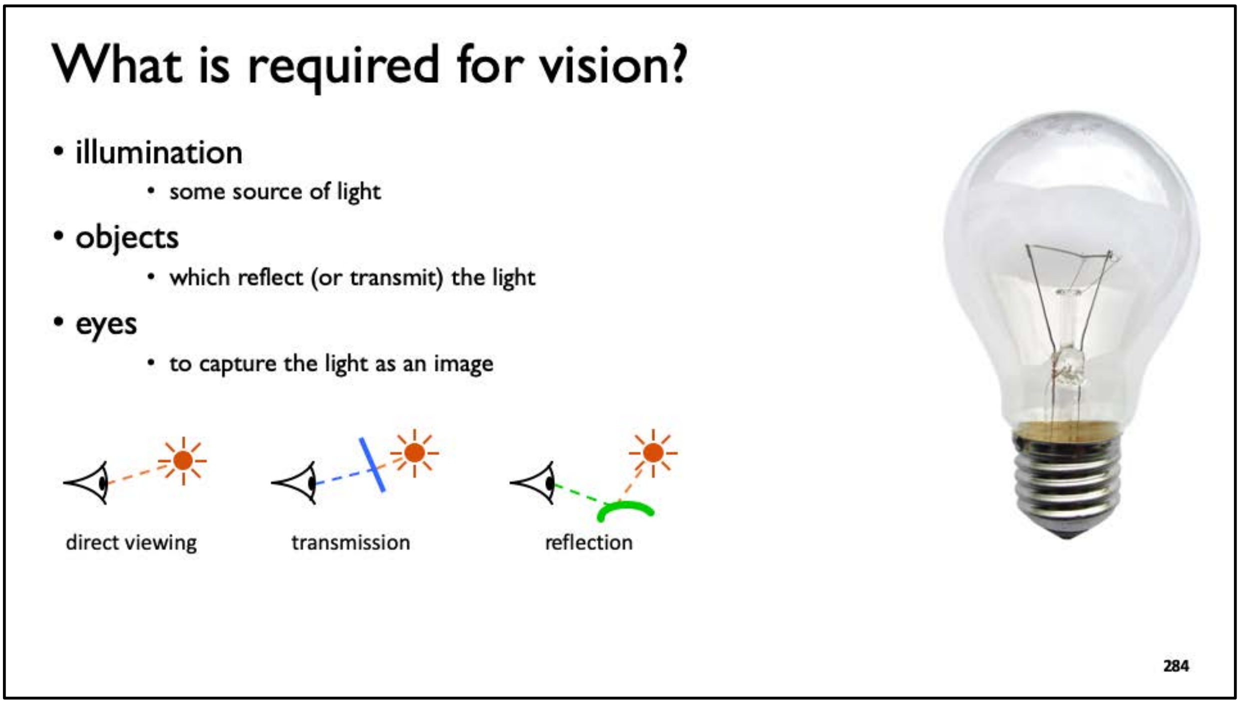
You should also ask yourself why we only need three colours to mix to make all possible colours.

Here are very short answers…

Red-Green-Blue are additive colours – used in mixing lights

Magneta-Yellow-Cyan are subtractive colours – used in mixing inks

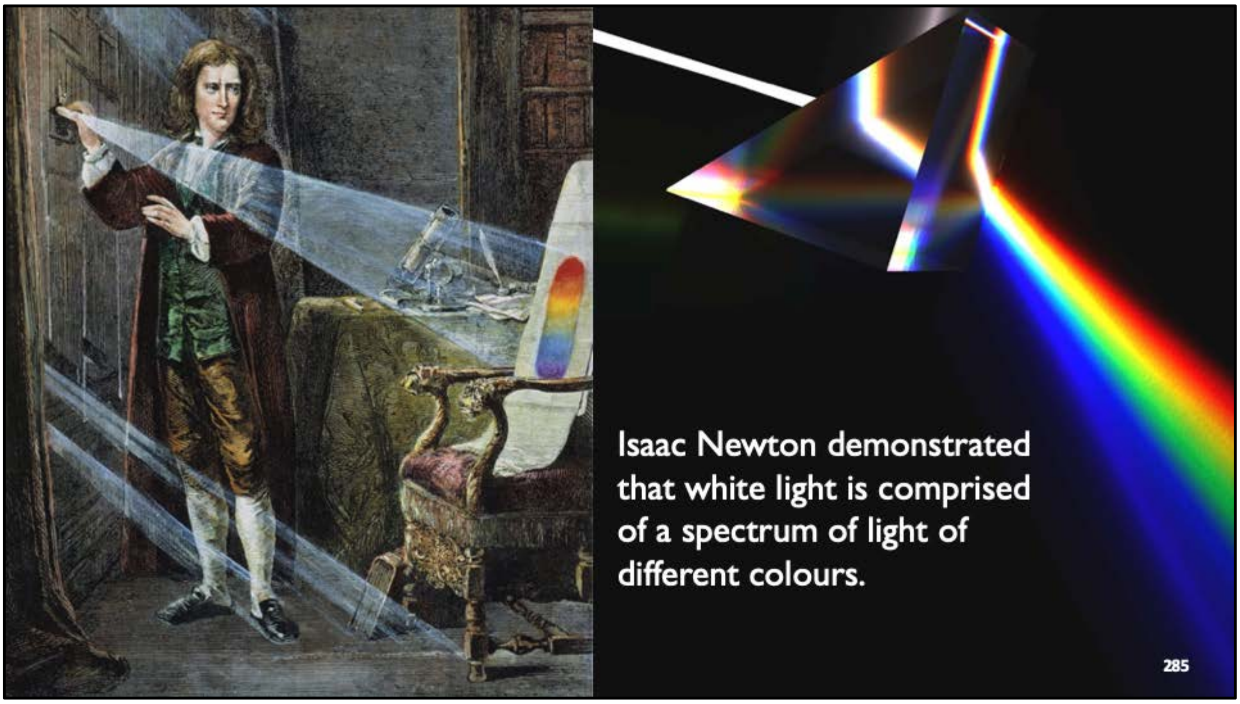
Red-Yellow-Blue are a simple approximation to Magenta-Yellow-Cyan, suitable for primary school children because red, yellow and blue are basic colour categories in English (so children know them) whereas magenta and cyan are not



You cannot see anything if there is no light

You cannot see anything if there is nothing there to see

You cannot see anything if you have no eyes to see with



http://vignette3.wikia.nocookie.net/warehouse-13-artifact-database/images/1/12/Newton's_prism_eperiment.jpg/revision/latest?cb=20150514171746

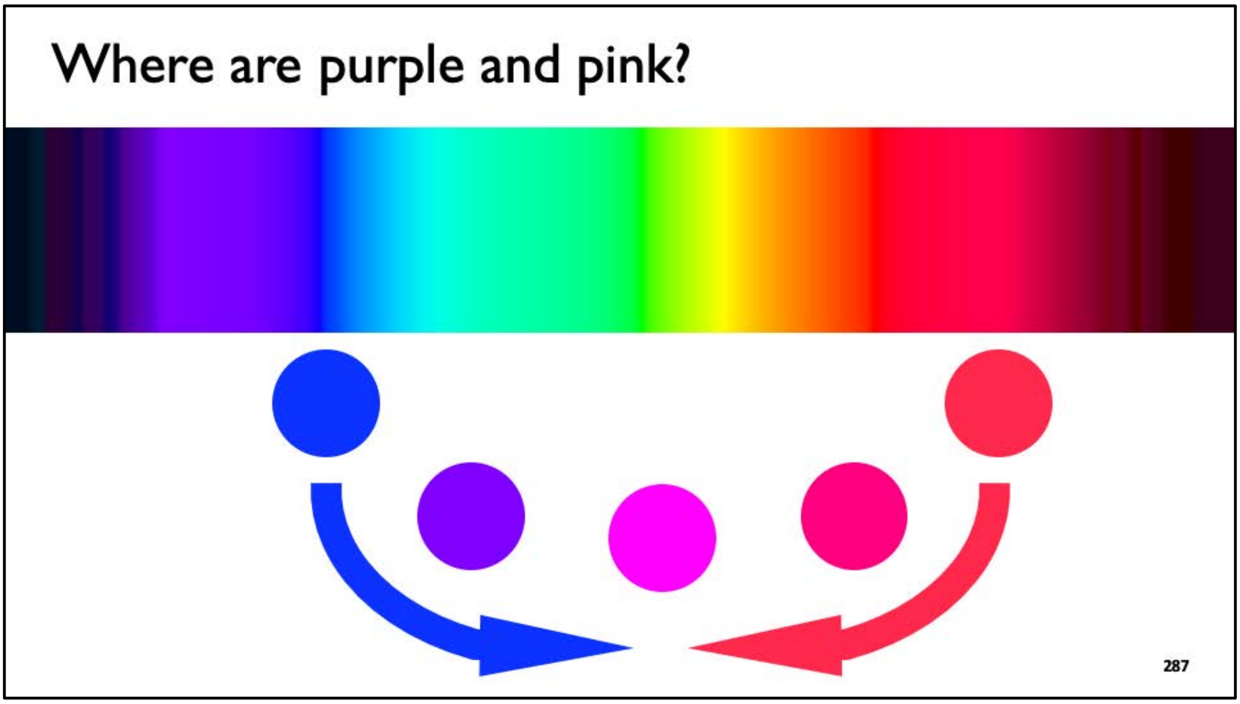
Isaac Newton (1666–1726) undertook his work on optics when in his late 20's at Trinity College in the University of Cambridge.
He was the first person to do serious work on how white light is actually composed of a whole range of different colours, which today we know as the spectrum of different wavelengths of visible light.
He put a prism near his window (notice in the painting that he had wooden shutters on the inside of the windows, which were more common than curtains at the time). The prism refracts the light, with shorter wavelengths (blues and violets) being refracted more than longer wavelengths.

https://commons.wikimedia.org/wiki/File:Spectrum-sRGB.svg

All children in English-speaking countries are taught the seven colours of the rainbow: ROYGBIV: red, orange, yellow, green, blue, indigo, violet.
But there are several problems with this:

- The rainbow is continuous. There are no distinct colours.
- Newton originally identified *five* principal colours: red, yellow, green, blue, violet.
- He revised this to seven to fit with the mystical properties of the number seven, including that this was the same number of notes as in the standard musical scale (doh, re, mi, fa, sol, la, te).
- The colour that he identified as ”blue” we would today call “blue-green”, “turquoise”, or “cyan”.
- The colour that he identified as “indigo” we would today call “blue”.

Today we tend to ditch indigo and teach that there are six principal colours: red, orange, yellow, green, blue, violet.



https://commons.wikimedia.org/wiki/File:Spectrum-sRGB.svg

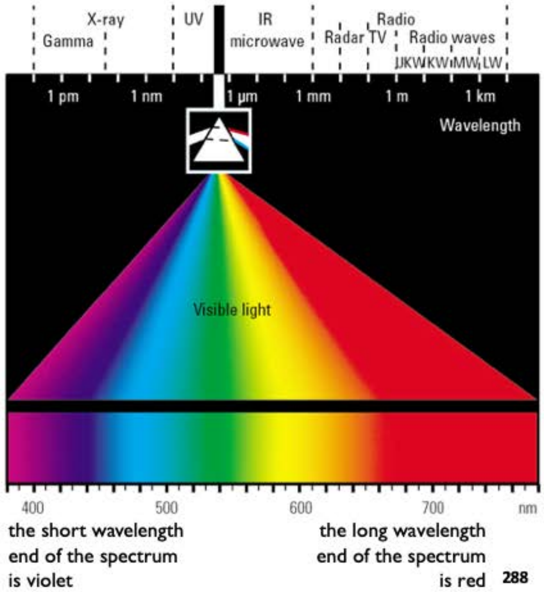
The purples, magentas and pinks are “non-spectral” colours*. They are made by mixing red and blue light together to produce a colour that does not appear anywhere in the spectrum.

These mixtures of blue and red allow us to make a circular “hue” dimension: along the spectrum through red, orange, yellow, green, blue, violet then round through the purples back to red.

*Contrary to the popular children’s song, *I Can Sing a Rainbow*, there is no pink in the rainbow.

Light: wavelengths & spectra

- light is electromagnetic radiation
  - visible light is a tiny part of the electromagnetic spectrum
  - visible light ranges in wavelength from 700nm (red) to 400nm (violet)
- every light has a spectrum of wavelengths that it emits
- every object has a spectrum of wavelengths that it reflects (or transmits)
- the combination of the two gives the spectrum of wavelengths that arrive at the eye



288

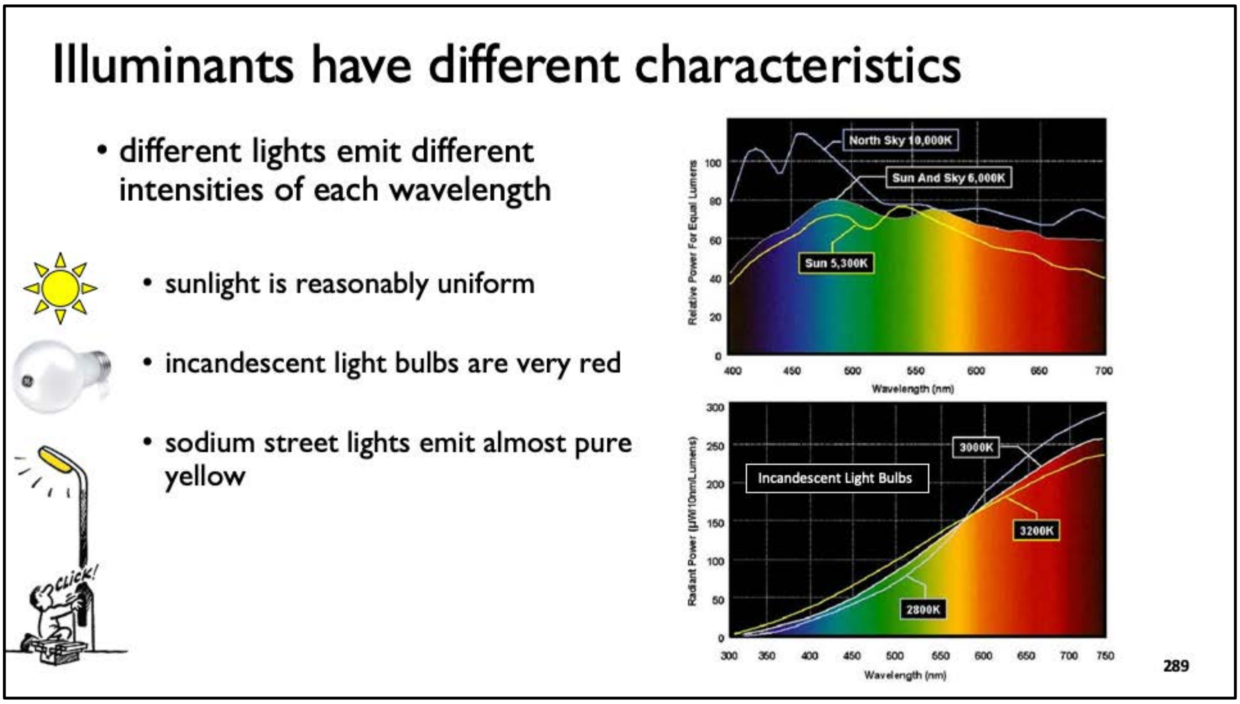
The diagram is only roughly right. Here are wavelengths that map well to what a person with normal vision thinks of as a pure colour:

400nm — violet
475nm — blue
510nm — green
570nm — yellow
590nm — orange
650nm — red

The spectrum is continuous, colours between these number can be described by combinations of the colour names, for example 490nm is greenish-blue (more on colour names later).

Any wavelength longer than 650nm is seen as red, until we hit the limit of what humans can detect, which is around 700nm.

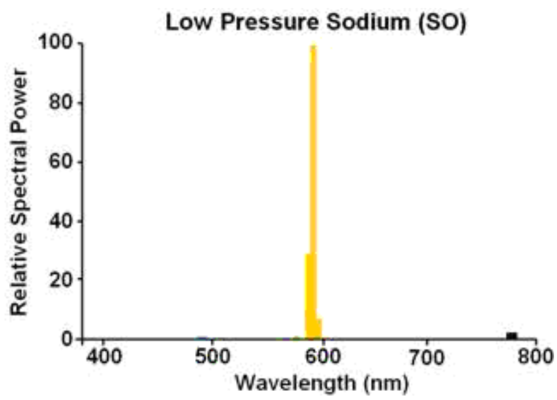
Handy hint: want to remember which end is *violet* and which end is red? *Short* people are often said to have *violent* tempers, so the *short* wavelengths are at the *violet* end .

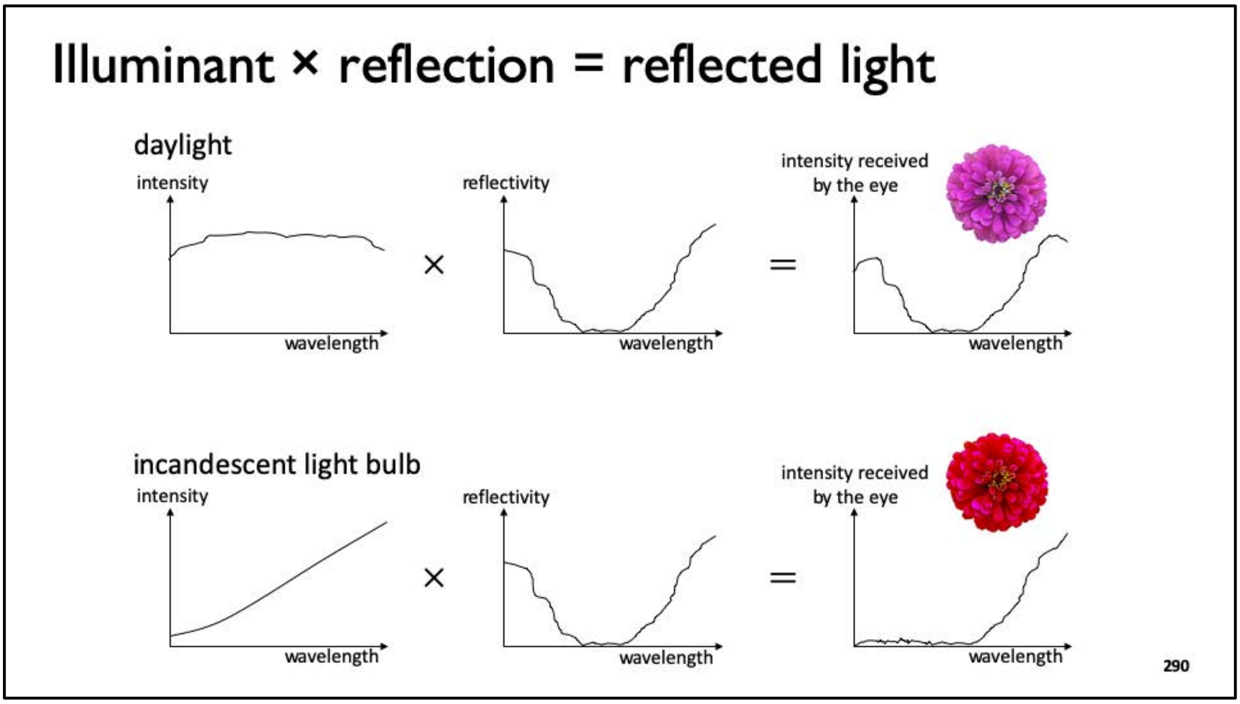


www.gelighting.com/na/business_lighting/education_resources/learn_about_light/

Incandescent light bulbs have a dominantly red spectrum. Modern LED light bulbs are described as "warm", "daylight" or "cool" depending on the spectrum they emit. "warm" lights are like incandescent light bulbs, emitting vastly more red light than other colours. "cool" lights are the opposite, they emit more blue light than red light. "daylight" bulbs attempt to mimic the spectrum of the sun, which is roughly uniform across the visible spectrum.

Sodium vapour lamps have a spectrum that is almost pure yellow, as seen at right.





The light received by your eye is the product of the illuminant and the reflectivity of the object.

You multiply each wavelength individually.

At top, we have a magenta flower, which reflects blue and red light, but absorbs green light. When illuminated by sunlight it looks magenta.

At bottom, we have the same flower, but now illuminated by an incandescent bulb. The flower still absorbs the same light, but there isn't much blue light to reflect, so the flower looks red.

Blue sapphire looks fantastic under sunlight but dull under incandescent light owing to the very limited amount of blue in the illuminant's spectrum.

Alexandrite, a rare form of the gemstone chrysoberyl, changes colour dramatically. In sunlight it ranges from greenish-blue to yellowish-green. Under incandescent light, it looks pink or red. This colour change is attributable to the particular wavelengths of light that the gem absorbs. See http://www.chemistry-blog.com/2013/08/22/alexandrite-effect-not-all-white-light-is-created-equal/

incandescent light bulb camera flash bulb

winter sunlight halogen light bulbs (overhead)

Comparison of illuminants

compare these things:

- colour of the monkey's nose and paws: more red under certain lights
- oranges & yellows (similar in all)
- blues & violets (considerably different)

291

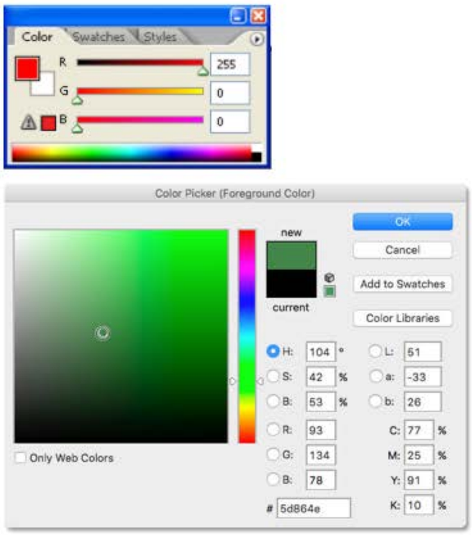
These photos show the differences in the same objects caused by different illuminants.

The camera does not lie: it shows exactly what was seen.

The human visual system does lie: your visual system automatically adjusts to the current illuminant. That means that, if you were looking at these objects in real life, you would see all four scenes in the colours shown in the top right image because your visual system would adjust the colours you perceive based on the overall lighting.

Representing colour

- we need a mechanism which allows us to represent colour in the computer by some set of numbers
  - preferably a small set of numbers which can be quantised to a fairly small number of bits each
- we will discuss:
  - Munsell's *artists'* scheme
    - which classifies colours on a perceptual basis
  - the mechanism of colour vision
    - how colour perception works
  - various *colour spaces*
    - which quantify colour based on either physical or perceptual models of colour

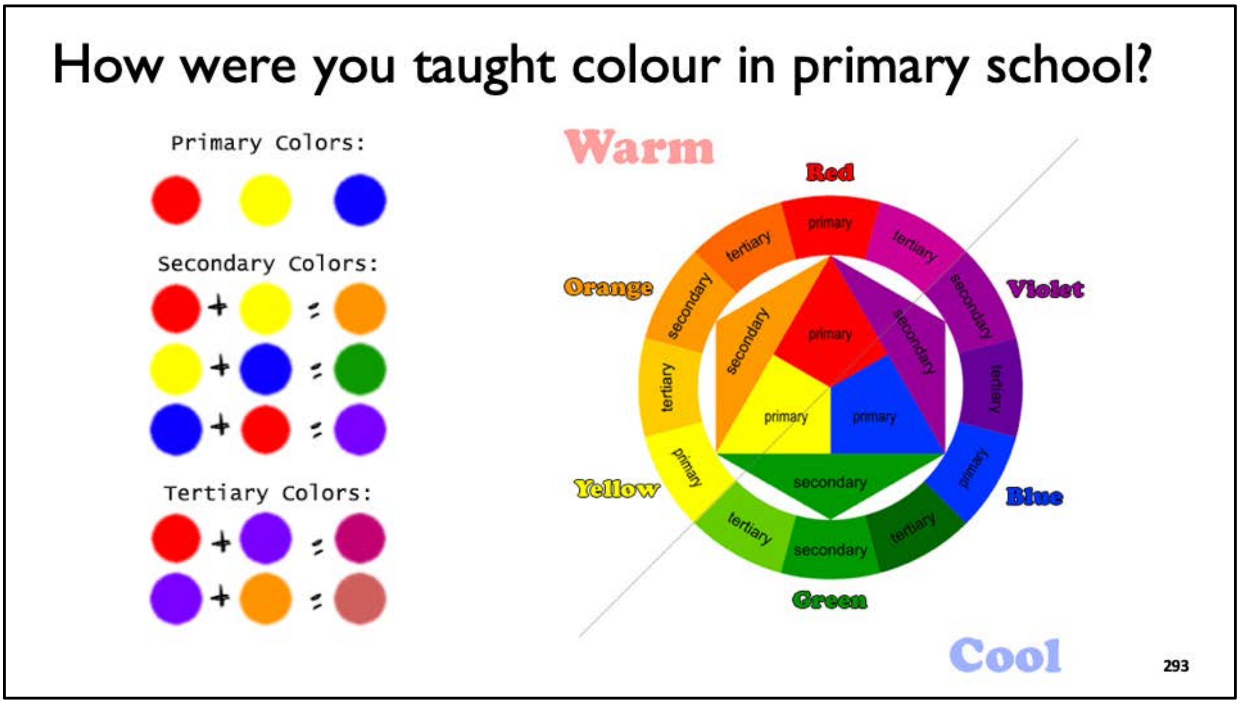


292

The top user interface allows specification of colours using R (red), G (green) and B (blue) sliders.

The bottom user interface offers four different colour spaces at the same time:
HSB (hue, saturation, brightness)
RGB (red, green, blue)
Lab (lightness, a=red–green, b=blue–yellow)
CMYK (cyan, magenta, yellow, black)

In the screen shot of the bottom user interface, the rainbow slider represents the **hue** axis of the HSB colour space. The large square represents the **saturation** (horizontal axis from zero saturation at left to full saturation at right) and **brightness** (vertical axis from zero brightness at bottom to full brightness at top) axes of the HSB colour space.

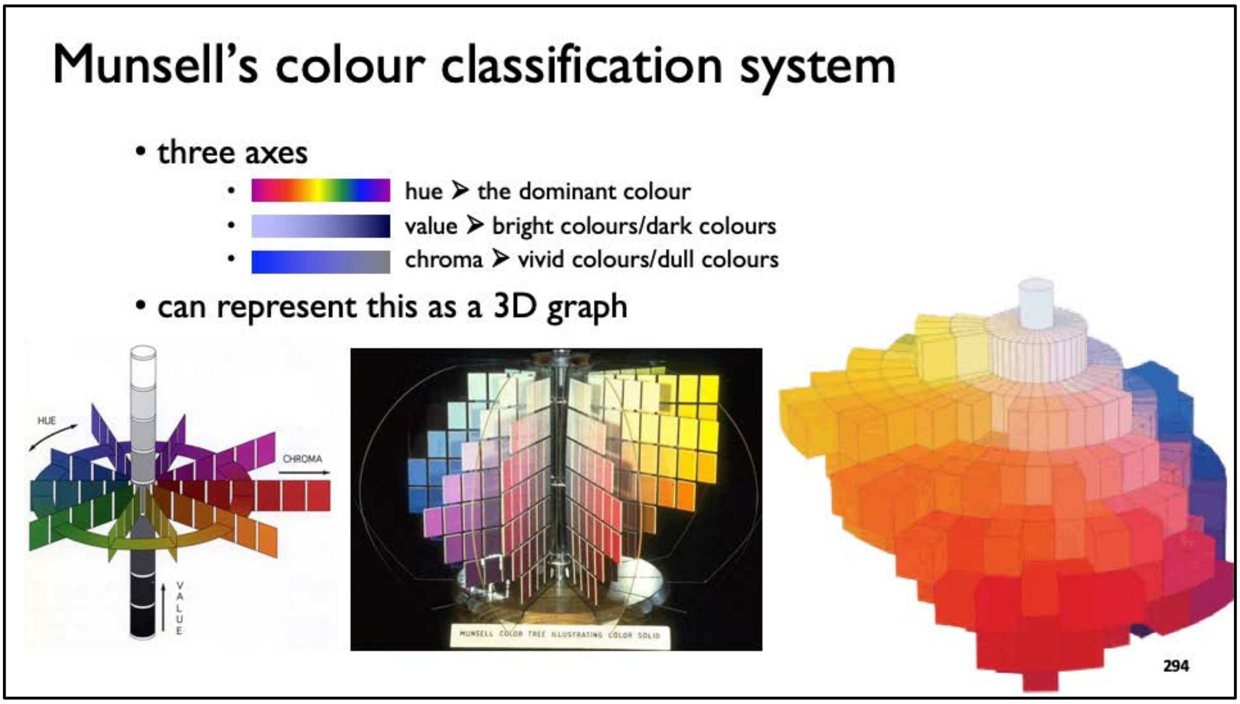


In primary school and in most art classes you are taught the following, which is simple to understand but is only approximately correct. Reality is so much more complicated, as we will see.

The primary school story is this:

- There are three primary colours: red, yellow and blue. These colours cannot be made from any other colours.
- There are three secondary colours made by equal mixes of the primaries: orange, green and violet (= purple).
- There are six tertiary colours made by mixing a secondary with an equal amount of an adjacent primary.

The primaries and secondaries all have well-defined names in English, which indicates something of the importance of these colours. Of the six tertiaries, there are no fundamental names, but I can easily think of simple names for four: blue-violet is called indigo, red-violet is crimson, blue-green is turquoise, yellow-green is lime. Yellow-orange and red-orange don't have easily-recalled simple names, which is likely because orange is a recent (about 500 years ago) addition to the fundamental colour terms in English, whereas red, yellow, green, blue and violet/purple are much older fundamental colour terms.



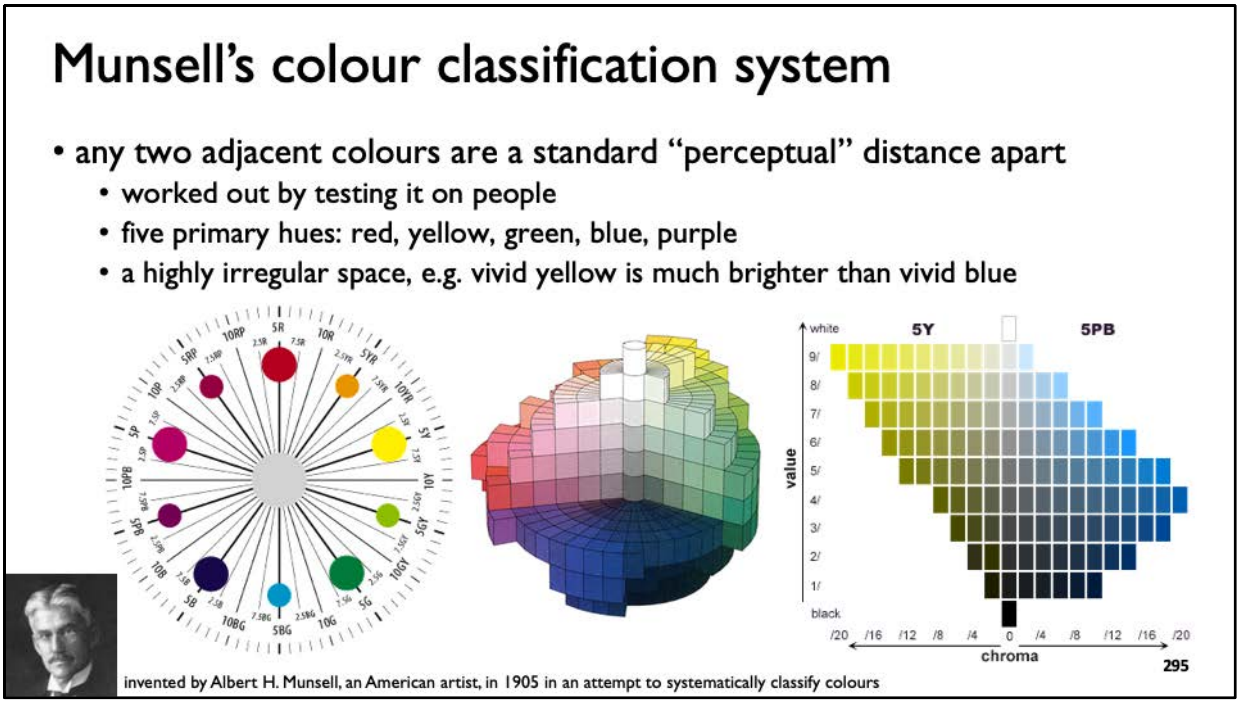
As the twentieth century begin, Albert Munsell, an American artist and professor, wanted to put colour theory onto a firmer footing. He conducted a series of experiments to produce a colour system based on rigorous scientific measurement of human vision.

Munsell's system is three-dimensional:

Hue is the dominant colour, the circular dimension described on the previous slide.

Value is brightness or lightness or intensity: a range from very dark to very bright.

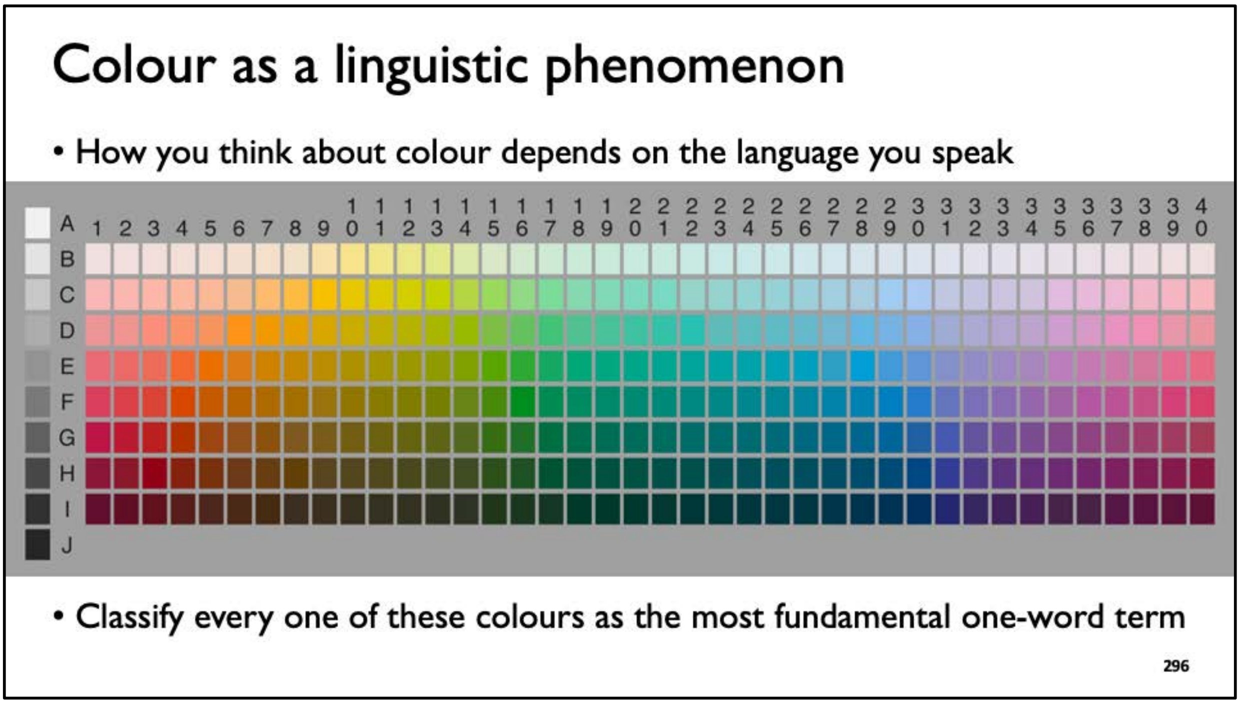
Chroma is Munsell's way of saying how saturated a colour is, it ranges from no colour (grey) to the most saturated colour possible.



In each dimension, Munsell colors are as close to perceptually uniform as he could make them, which makes the resulting shape quite irregular. Munsell explains:

"Desire to fit a chosen contour, such as the pyramid, cone, cylinder or cube, coupled with a lack of proper tests, has led to many distorted statements of colour relations, and it becomes evident, when physical measurement of pigment values and chromas is studied, that no regular contour will serve."

This is owing to the nature of human vision. Our eyes respond more strongly to yellow light than to blue light, so a fully saturated yellow appears to us to be much brighter than a fully saturated blue.



https://evanwarfel.files.wordpress.com/2012/07/wcs23.png

How you think of colour depends on the language you speak.

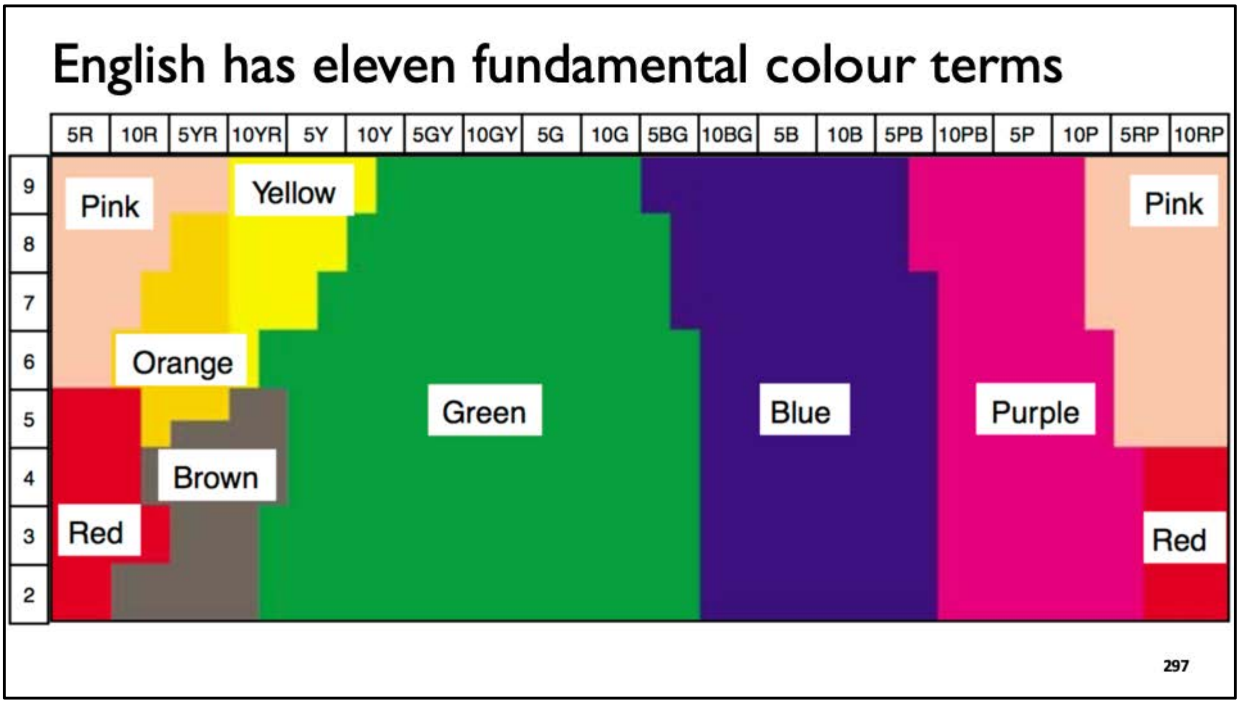
If you classified each of the colours in the chart by the most fundamental colour term you can think of, how many distinct colour terms would you need?

Do you think of "pink" as a fundamental colour in its own right, or is it just a word for "light red"?

Do you think of "brown" as a fundamental colour? Or is brown just a "dark orange"?

Do you think of "lime" as a fundamental colour? Or is it a shade of green?

Is "turquoise" a fundamental colour? Or is it a blue-green?

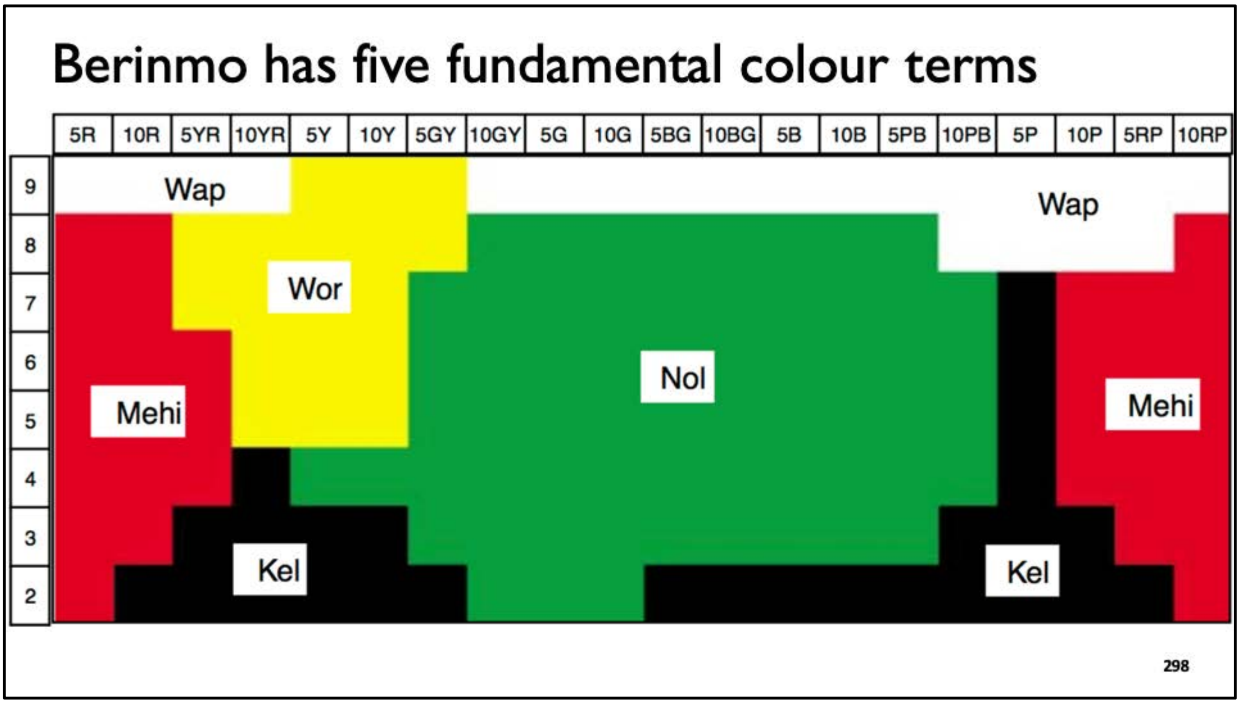


This and the diagram on the next slide from "Language, thought and color: recent developments", Paul Kay and Terry Regier, TRENDS in Cognitive Sciences, Vol.10 No.2, February 2006, pages 51–53.

English has eleven fundamental colour terms:
Red, Orange, Yellow, Green, Blue, Purple
Pink, Brown
White, Grey, Black

Russian has twelve fundamental colour terms. In Russian you distinguish between light blue (**голубо́й,** gah-loo-BOY) and dark blue (**си́ний,** SEE-neey), in the same way that in English you distinguish between light red (pink) and dark red (red).

In Japanese, the word for blue (青 *ao*) traditionally referred to both blue and green. A separate term for green (緑 *midori*) was introduced around the 10th century but for centuries it was a shade of *ao*, rather than a fundamental colour. Even today some green things (grass, leaves, apples) will be called *ao* rather than *midori*. In Japanese, traffic lights are red, yellow, and **blue**.



Wap is white and pale versions of all other colours except yellow.
Kei is black, dark versions of brown, yellow-green, blue-green, and blue. It also incorporates certain shades of brighter blue-purple.
Mehi covers red, dark pink, and reddish purples.
Wor covers yellow and yellowish shades of green, orange and pink.
Nol covers greens and blues.

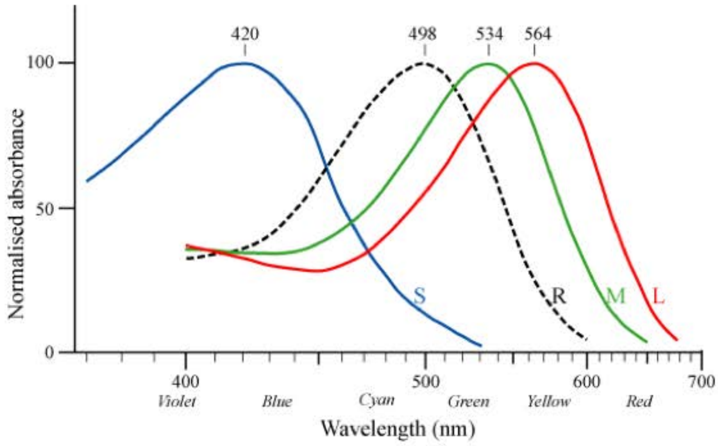
Berlin and Kay, in their 1969 book "Basic Color Terms", theorise that every human language has between two and twelve fundamental colour terms. Furthermore, that the most fundamental of these fundamental colour terms evolve into language always in the same order:

1,2. Light-warm *and* dark-cool
(white-pale-red-yellow *and* black-dark-green-blue)
3. Red.
4. Yellow *or* Green.
5. Yellow *and* Green.
6. Blue.
7. Brown.
8. Purple *or* pink *or* orange *or* grey

Berinmo is a language that matches the model.
Later work has loosened this strict ordering by demonstrating that there are languages where it is not correct.

Colour vision

- there are three types of cone
- each responds to a different spectrum
  - *very* roughly long, medium, and short wavelengths
- different numbers of the different types
  - far fewer of the short wavelength receptors
  - so cannot see fine detail in blue

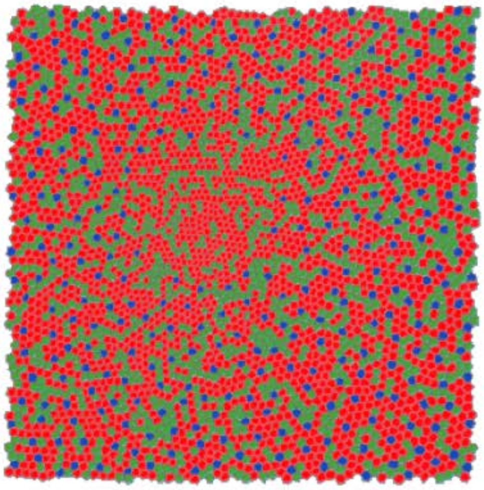


299

This is the tri-stimulus or tri-chromatic theory of colour vision. Tri-stimulus means that there are three different stimuli that affect how we see in colour. Tri-chromatic means that there are three different types of detector, each responsive to a different colour spectrum.

The dashed line is the response curve of the rods. These only work in low light conditions.

Distribution of different cone types



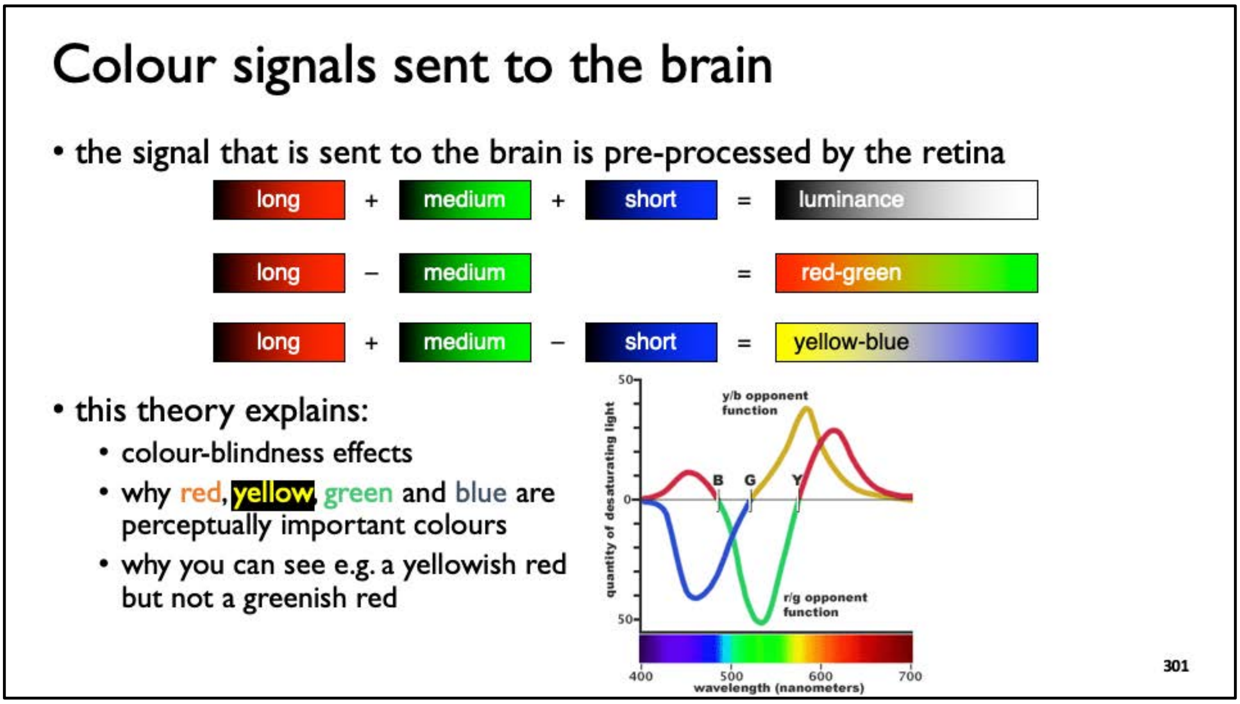
simulated cone distribution at the centre of the fovea

- this image is about 1° of visual angle (about 120 cones across)
- distribution is:
  - 7% short
  - 37% medium
  - 56% long
- short wavelength receptors
  - regularly distributed
  - not in the central $^1/_3$°
  - outside the fovea, only 1% of cones are short
- long & medium wavelength receptors
  - about 3:2 ratio long:medium

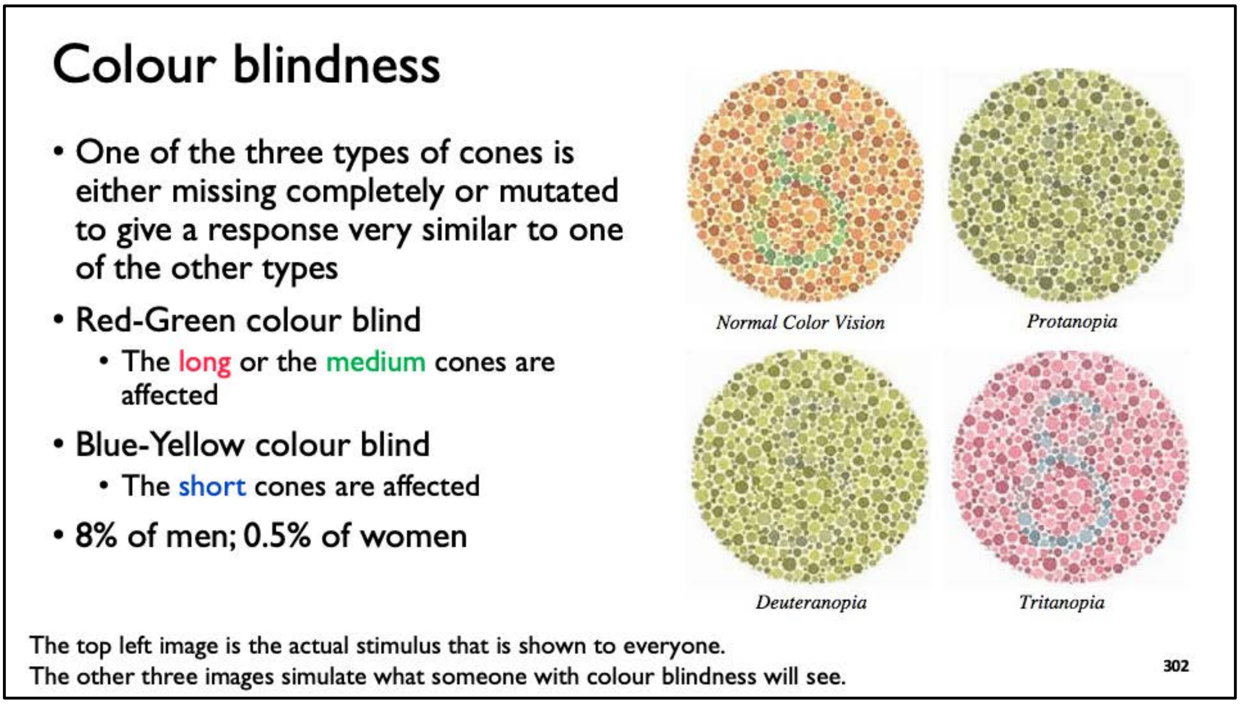
300

www.cis.rit.edu/people/faculty/montag/vandplite/pages/chap_9/ch9p1.html

This demonstrates that you cannot see fine detail in the short wavelengths, because there are not enough detectors available to allow you to see fine detail.



This is the colour-opponent theory of colour vision. It says that, there are three types of signal going to the brain:
Luminance is made by adding up the signals from groups of cones, and this is the highest resolution channel going to the brain, with signals coming from groups of just three or four cones.
Red-green is lower-resolution, with larger groups of cones being combined by subtracting the medium signals from the long signals.
Yellow-blue is even lower-resolution, made from combining the signal from large groups of cones.



Protanopia: missing the long cones.
Deuteranopia: missing the medium cones.
Tritanopia: missing the long cones.

In both protanopia and deuteranopia, the person is red-green colour blind. What is perfectly distinguishable to someone with normal colour vision (the numeral 8 in green on an orange background) is completely indistinguishable to someone who is red-green colour blind.

Someone with tritanopia is blue-yellow colour blind. They see a different effect to the person with normal colour vision, but they can distinguish the numeral 8 against the background. There are other colour schemes used to test for tritanopia.

Red-green colour blindness is linked to the sex chromosomes. The codes for long and medium cones are on the X-chromosome. Men have only one X-chromosome so, if it is incorrect, they are red-green colour blind. Women have two X-chromosomes, and both need to be incorrect for a woman to be red-green colour blind. The short cones are encoded on a different chromosome, so men and women have an equal (and very small) chance of being blue-yellow colour blind.

Chromatic metamerism

- many different spectra will induce the same response in our cones
  - the values of the three perceived values can be calculated as:
    - $l = k \int P(\lambda)\, l(\lambda)\, d\lambda$
    - $m = k \int P(\lambda)\, m(\lambda)\, d\lambda$
    - $s = k \int P(\lambda)\, s(\lambda)\, d\lambda$
  - k is some constant, $P(\lambda)$ is the spectrum of the light incident on the retina
  - two different spectra (e.g. $P_1(\lambda)$ and $P_2(\lambda)$) can give the same values of l, m, s
  - we can thus fool the eye into seeing (almost) any colour by mixing correct proportions of some small number of lights

303

Chromatic = colour
Metamerism = two different things that look the same

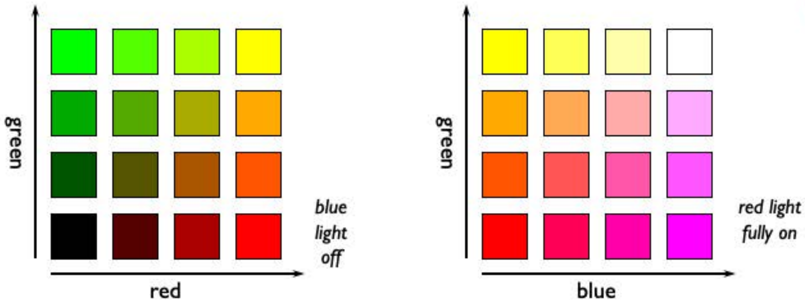
So: two different spectra can produce exactly the same response in our cones.

The value that a cone produces depends on its response function, e.g. $m(\lambda)$ for the medium cones, and on the spectrum of light that hits the cone, $P(\lambda)$. You multiply the two together for every wavelength and then add them up for all possible wavelengths (the integral sign means do a sum of all wavelengths).

For example, a mix of red and green light will look exactly like a yellow light to someone with normal colour vision.

Mixing coloured lights

- by mixing different amounts of red, green, and blue lights we can generate a wide range of responses in the human eye

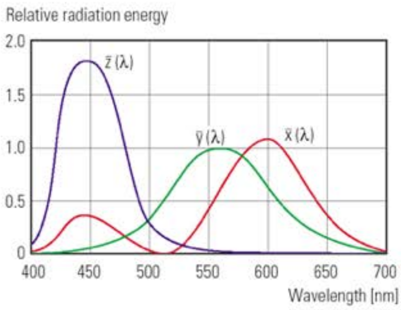


- not all colours can be created in this way

304

XYZ colour space

- not every wavelength can be represented as a mix of red, green, and blue lights
- but matching & defining coloured light with a mixture of three fixed primaries is desirable
- CIE define three standard primaries: X, Y, Z



Y matches the human eye's response to light of a constant intensity at each wavelength (***luminous-efficiency function of the eye***)

X, Y, and Z are not themselves colours, they are used for defining colours – you cannot make a light that emits one of these primaries

XYZ colour space was defined in 1931 by the *Commission Internationale de l' Éclairage* (CIE)

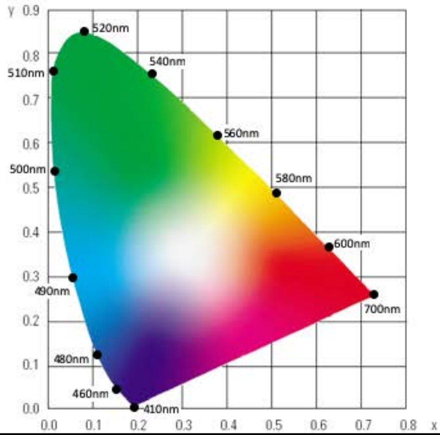
305

CIE chromaticity diagram

- *chromaticity* values are defined in terms of x, y, z

$$x = \frac{X}{X+Y+Z}, \quad y = \frac{Y}{X+Y+Z}, \quad z = \frac{Z}{X+Y+Z} \quad \therefore \quad x+y+z=1$$

  - ignores luminance
  - can be plotted as a 2D function
- pure colours (single wavelength) lie along the outer curve
- all other colours are a mix of pure colours and hence lie inside the curve
- points outside the curve do not exist as colours



306

Colour spaces

- CIE *XYZ*, *Yxy*
- Uniform
  - equal steps in any direction make equal perceptual differences
  - CIE $L^*u^*v^*$, CIE $L^*a^*b^*$
- Pragmatic
  - used because they relate directly to the way that the hardware works
  - *RGB*, *CMY*, *CMYK*
- Munsell-like
  - used in user-interfaces
  - considered to be easier to use for specifying colour than are the pragmatic colour spaces
  - map easily to the pragmatic colour spaces
  - *HSV*, *HLS*

307

XYZ is not perceptually uniform

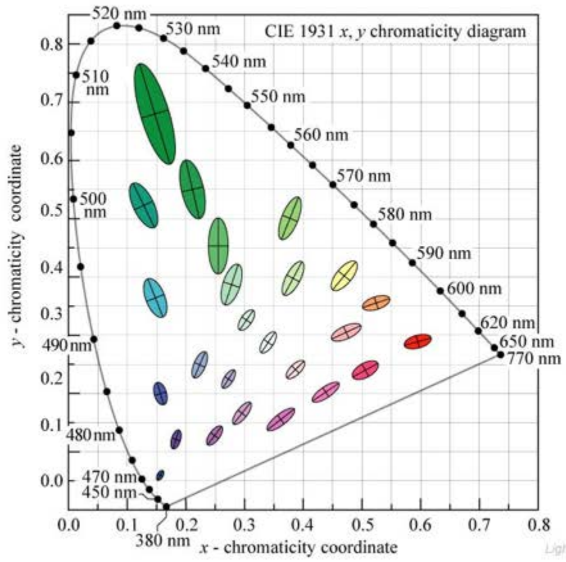


Fig. 17.5. MacAdam ellipses plotted in the CIE 1931 (*x*, *y*) chromaticity diagram. The axes of the ellipses are ten times their actual lengths (after MacAdam, 1943; Wright, 1943; MacAdam, 1993).

Each ellipse shows how far you can stray from the central point before a human being notices a difference in colour

E. F. Schubert
Light-Emitting Diodes (Cambridge Univ. Press)
www.LightEmittingDiodes.org

308

Luv was designed to be more uniform

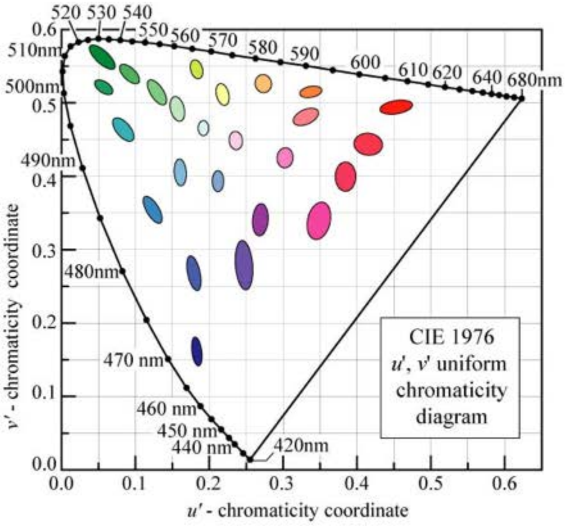
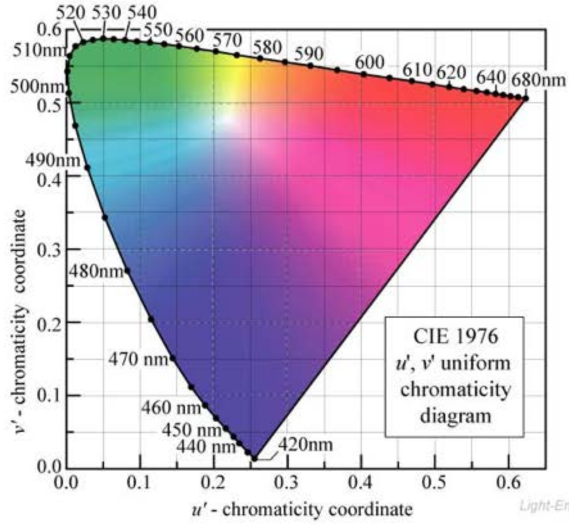


Fig. 17.7. MacAdam ellipses transformed to uniform CIE 1976 (u', v') chromaticity coordinates. For clarity, the axes of the transformed ellipses are ten times their actual lengths. Transformed ellipses are not ellipses in a strict mathematical sense, but their shapes closely resemble those of ellipses. The areas of the transformed ellipses in the (u', v') diagram are much more similar than the MacAdam ellipses in the (x, y) diagram.

E. F. Schubert
Light-Emitting Diodes (Cambridge Univ. Press)
www.LightEmittingDiodes.org

309

Luv colour space



L is luminance and is orthogonal to *u* and *v*, the two colour axes

Fig. 17.6. CIE 1976 (*u*′, *v*′) uniform chromaticity diagram calculated using the CIE 1931 2° standard observer.

E. F. Schubert
Light-Emitting Diodes (Cambridge Univ. Press)
www.LightEmittingDiodes.org

310

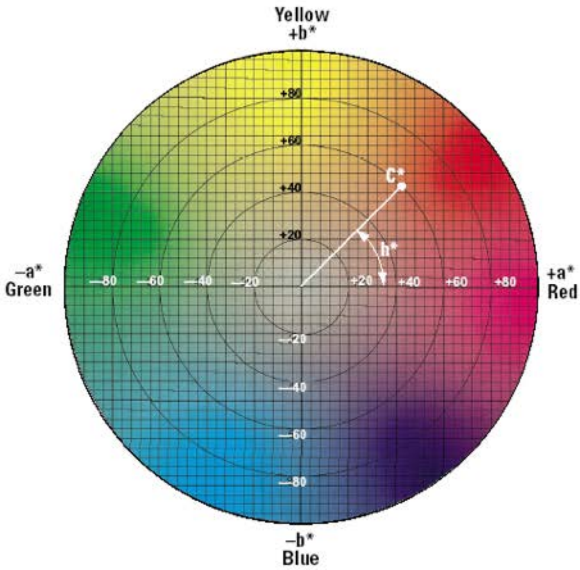
$L^*u^*v^*$ is an official CIE colour space. It is a straightforward distortion of XYZ space.

Lab space

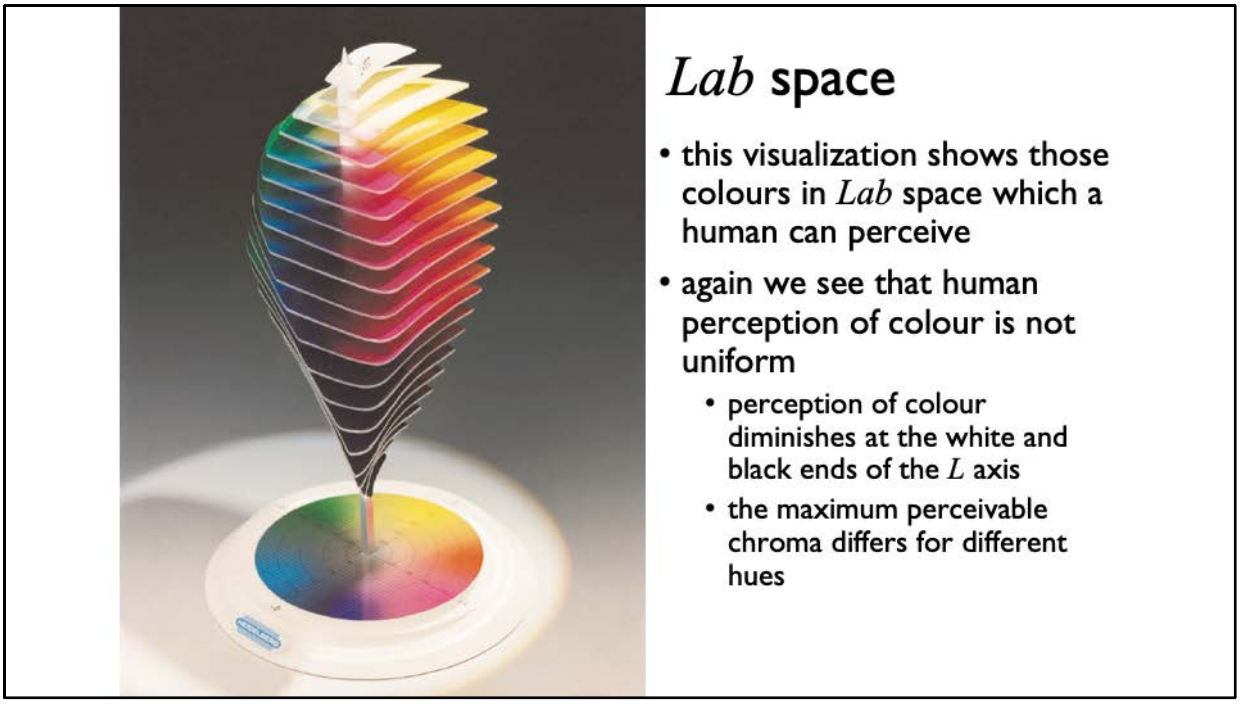
- another CIE colour space
- based on complementary colour theory
- also aims to be perceptually uniform

- L* luminance
- a* red–green
- b* blue–yellow

$L^*=116(Y/Y_n)^{1/3}$
$a^*=500[(X/X_n)^{1/3}-(Y/Y_n)^{1/3}]$
$b^*=200[(Y/Y_n)^{1/3}-(Z/Z_n)^{1/3}]$



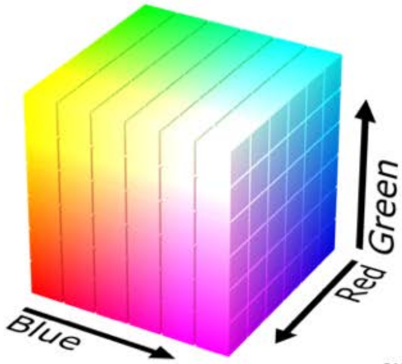
311



This is essentially a more scientific reworking of what Munsell was trying to achieve.

RGB space

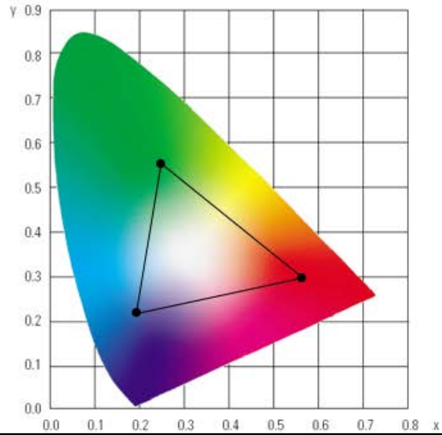
- all display devices which output light mix red, green and blue lights to make colour
  - televisions, video projectors, LCD screens
- nominally, RGB space is a cube
- the device puts physical limitations on:
  - the range of colours which can be displayed
  - the brightest colour which can be displayed
  - the darkest colour which can be displayed



313

RGB in *XYZ* space

- display devices mix red, green, and blue to make all other colours
- the red, green, and blue primaries each map to a point in *XYZ* space
- any colour within the resulting triangle can be displayed
    - any colour outside the triangle cannot be displayed
    - for example: CRTs cannot display very saturated purple, turquoise, or yellow



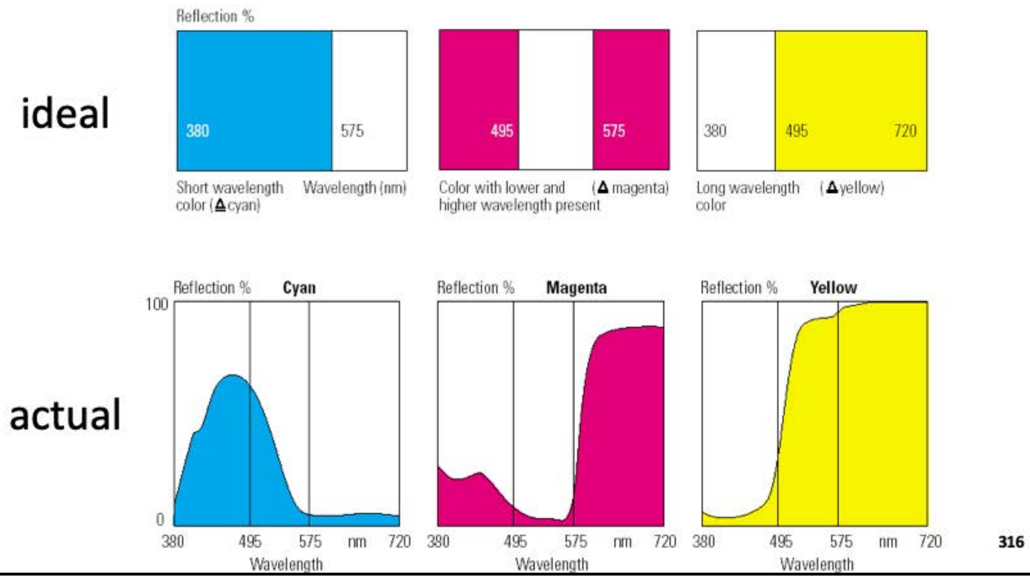
314

CMY space

- printers make colour by mixing coloured inks
- the important difference between inks (*CMY*) and lights (*RGB*) is that, while lights *emit* light, inks *absorb* light
  - cyan absorbs red, reflects blue and green
  - magenta absorbs green, reflects red and blue
  - yellow absorbs blue, reflects green and red
- *CMY* is, at its simplest, the inverse of *RGB*
- *CMY* space is nominally a cube

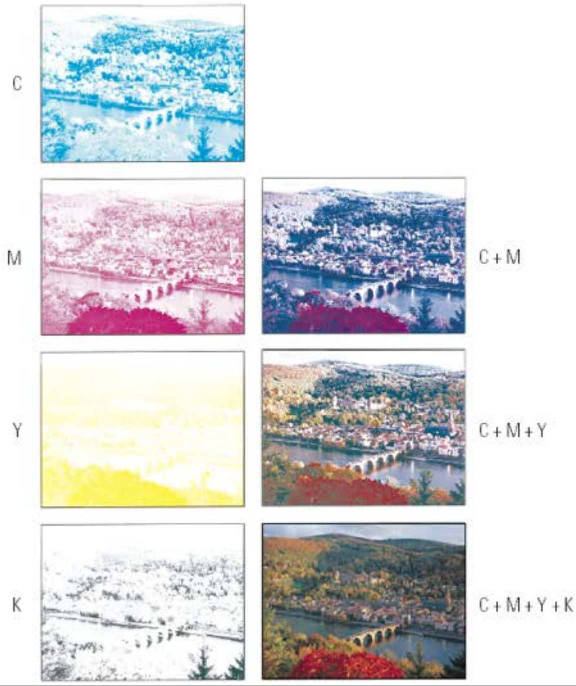
315

Ideal and actual printing ink reflectivities



CMYK space

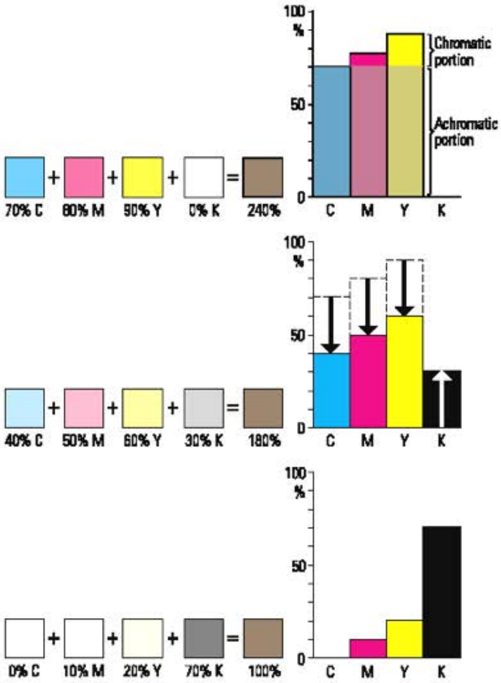
- in real printing we use black (key) as well as *CMY*
- why use black?
  - inks are not perfect absorbers
  - mixing $C + M + Y$ gives a muddy grey, not black
  - lots of text is printed in black: trying to align C, M and Y perfectly for black text would be a nightmare



317

Using K

- if we print using just CMY then we can get up to 300% ink at any point on the paper
- removing the achromatic portion of CMY and replacing with K reduces the maximum possible ink coverage to 200%



318

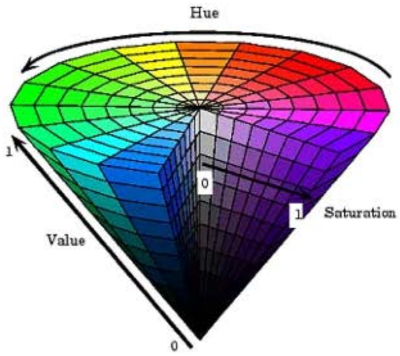
Colour spaces for user-interfaces

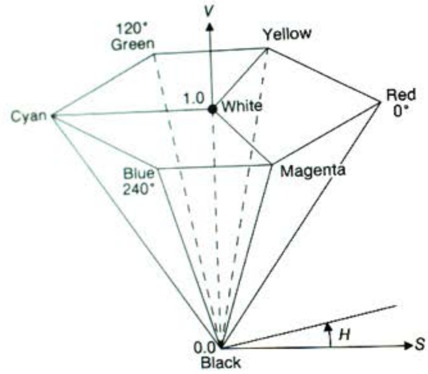
- *RGB* and *CMY* are based on the physical devices which produce the coloured output
- *RGB* and *CMY* are difficult for humans to use for selecting colours
- Munsell's colour system is much more intuitive:
  - hue — what is the principal colour? 
  - value — how light or dark is it? 
  - chroma — how vivid or dull is it? 
- computer interface designers have developed basic transformations of *RGB* which resemble Munsell's human-friendly system

319

HSV: hue saturation value

- three axes, as with Munsell
  - hue and value have same meaning
  - the term “saturation” replaces the term “chroma”



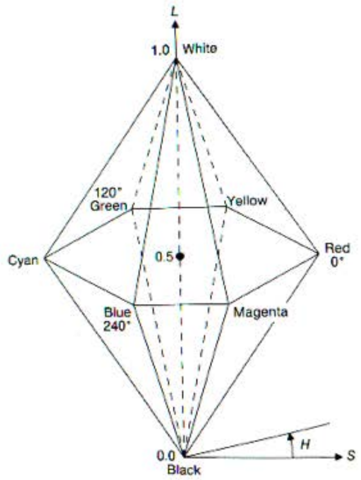


- designed by Alvy Ray Smith in 1978

320

The algorithm to convert *HSV* to *RGB* and back is straightforward because Alvy Ray Smith (the designer of the *HSV* space) wanted something that was intuitive for humans but very simple to convert to *RGB*

HLS: hue lightness saturation



- a simple variation of *HSV*
  - hue and saturation have same meaning
  - the term "lightness" replaces the term "value"
- designed to address the complaint that *HSV* has all pure colours having the same lightness/value as white
  - designed by Metrick in 1979

321

Whenever one colour space is developed, someone always finds a problem with it!

Summary of colour spaces

- the eye has three types of colour receptor
- therefore we can validly use a three-dimensional co-ordinate system to represent colour
- XYZ is one such co-ordinate system
  - Y is the eye's response to intensity (luminance)
  - X and Z are, therefore, the colour co-ordinates
    - same Y, change X or $Z \Rightarrow$ same intensity, different colour
    - same X and Z, change $Y \Rightarrow$ same colour, different intensity
- there are other co-ordinate systems with a luminance axis
  - $L^*a^*b^*$, $L^*u^*v^*$, HSV, HLS
- some other systems use three colour co-ordinates
  - RGB, CMY
  - luminance can then be derived as some function of the three
    - e.g. in RGB: $Y = 0.299\,R + 0.587\,G + 0.114\,B$

322