

Technical Note

Standard Colour Spaces

Richard Kirk

Document ref.	FL-TL-TN-0417-StdColourSpaces
Creation date	January 9, 2004
Last modified	30 November 2010
Version no.	4.0

This document has been superseded.
Visit the following link to access
the latest version:
www.filmlight.ltd.uk/support/documents/other/legacy_tl.php

Summary

In 1931, the Commission Internationale d'Éclairage (CIE) recommended a system for colour measurement. This system allowed the specification of colour matches using the CIE XYZ tristimulus values. In 1976, the CIE recommended the CIE LAB and CIE LUV colour spaces for the measurement of colour differences, and colour tolerances. These colour spaces, and their more modern variants are the basic tools for modern colorimetry.

You can use Truelight without knowing all about CIE colour spaces. However, if you wonder why the XYZ and the L*a*b* calibration for the same monitor look different, you may find your explanation here.

Whites are dealt with in a separate section. Most of us know what white paper and white paint is. We might think we know what white light is too. A full discussion of what is and what is not 'white' is much too big a task for this small note, but we introduce a few basic issues.

Scanners and printers usually communicate in their own device-dependent RGB. Video has its own colour standard. Densitometers use Status A and Status M colour spaces. We describe all the spaces that Truelight uses to build up a colour transform.

Finally we describe a number of effects that are not covered by the standard colour models. These may prevent us getting a perfect visual match. There is not much we can do about these other than to be aware of them, and to try and avoid their more extreme effects.

Contents

1	Measurement and Units.....	3
	1.1 Intensity of a point light source.....	3
	1.2 Brightness of a broad light source.....	5
	1.3 Wavelength.....	6
	1.4 Screen gain.....	7
	1.5 Illumination.....	7
2	Matching a light source.....	8
	2.1 Matching a colour.....	8
	2.2 Matching an out of gamut colour.....	9
3	CIE XYZ Tristimulus.....	10
	3.1 Measuring tristimulus values.....	11
	3.2 XYZ units.....	12
	3.3 The human eye LMS primaries.....	12
4	Luminance-chrominance co-ordinates.....	13
	4.1 Yxy co-ordinates.....	13
	4.2 Yu'v' co-ordinates.....	13
5	Perceptually uniform colour spaces.....	14
	5.1 CIE L*a*b*.....	14
	5.2 CIE L*u*v*.....	16
	5.3 Delta E.....	16
	5.4 Colour appearance modelling.....	17
6	The Luminous White Point.....	18
	6.1 Colour temperature.....	18
	6.2 Calculating Colour Temperatures.....	20
	6.3 Thermal spectra and CIE daylight illuminants.....	21
	6.4 Psychophysical phenomena.....	22
	6.5 Colour Rendering Index.....	24
7	Densitometry.....	25
	7.1 Status A colour densitometry.....	26
	7.2 Status M colour densitometry.....	27
8	Cineon Exposure Space.....	28
	8.1 Reference White and Reference Black.....	29
	8.2 Aim Gammas.....	30
	8.3 Calibrating a Recorder.....	30
	8.4 Exposure Units.....	31
9	Video.....	32
	9.1 Digital to Voltage conversion.....	32
	9.2 Voltage to Luminance conversion.....	33
	9.3 RGB colours.....	35
	9.4 Luminance-Chrominance Coding.....	36
	9.5 Setting the Contrast.....	37
	9.6 Digital Cinema Interface.....	37
10	Real vision and simple models.....	38
	10.1 Variations in cone pigments.....	39
	10.2 Previsual variations.....	40
	10.3 Photopigment masking.....	41
	10.4 Colour opponent coding.....	42
	10.5 The influence of the rods.....	44
	10.6 The influence of melanopsin ganglion cells.....	45
	10.7 The effect of polarization.....	45
	10.8 Variation with retinal position.....	46
	10.9 The special case of the white point.....	47
	10.10 Conclusion.....	47

1 Measurement and Units

The SI unit of power is the *watt*.

We can measure the power in a beam of light using a bolometer. A bolometer is a black object with an electrical heater and a thermometer of some kind. We shine light on the bolometer with the heater off, and we measure the rise in temperature. We turn off the light and adjust the current through the heater to get the same rise in temperature. This gives us the power in the beam, or the energy in a flash of light.

Most light detectors are more sensitive to some wavelengths than others. A bolometer is equally sensitive to any radiation that can be absorbed, so it will give you the power of a beam of any radiation in watts whatever the wavelength is. Bolometers are usually not sensitive enough for our needs, but they are used to calibrate more sensitive devices.

1.1 Intensity of a point light source

The number of watts in a beam of light does not tell us how bright it looks unless we know what wavelengths are present. The unit of brightness was originally based on standard candles, then on a variety of standard lamps with a fixed whiteish spectrum. The current standard candle is effectively defined as a uniform green light source emitting $4\pi/683$ W of light with a wavelength of 555nm.

The 1979 definition of the SI standard unit of luminous intensity the *candela* is...

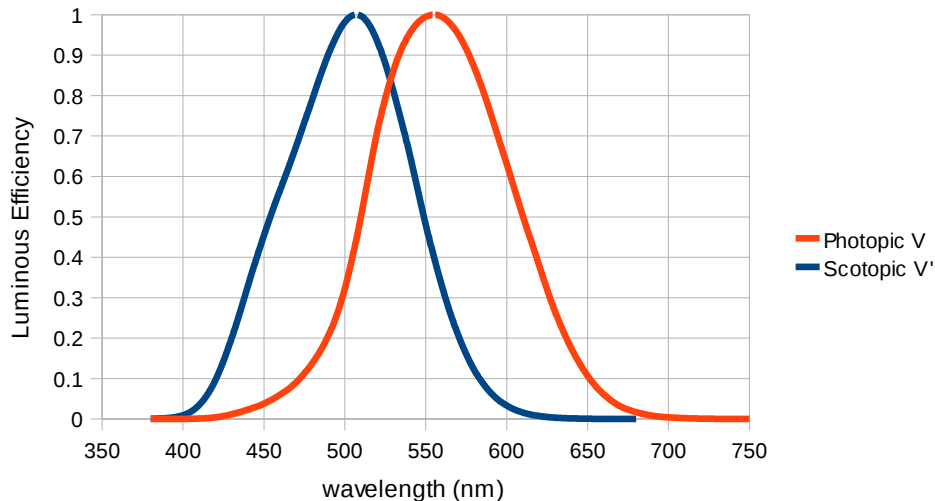
The candela is the luminous intensity in a given direction, of a source emitting monochromatic radiant energy of frequency 540×10^{12} Hz and whose radiant intensity in that direction is 1/683 watt per steradian.

The light is defined by frequency, not by wavelength, but it is still 555nm to all practical purposes. The output is quoted in watts per unit solid angle, not as total power so it does not need the factor of 4π . The 'magic number' of 683 was chosen to match the new units to the previous lamps.

The *lumen* is the SI unit for the total amount of light, equal to the intensity in candelas multiplied by the solid angle in steradians. A small digital projector may output a few thousand lumens.

You cannot meaningfully express the intensity of a parallel beam of light in candelas, but you can give the amount of light in lumens. Optics can change the solid angle of a beam of light, and so change the intensity in candelas. Unless light is absorbed, the amount of light in lumens stays the same.

If we want to measure the luminous intensity of wavelengths other than 555 nm, we need to weight them according to their visibility. There are two standard photometric observers – the photopic standard (CIE 1924) for normal colour vision, and the scotopic standard (CIE 1954) for lowlight monochrome vision.



These functions are usually scaled so the peak luminous efficiency is 1.0.

In this graph 555nm is the peak of the photopic luminance efficiency. We are more sensitive to seeing 1 watt of 555nm light than 1 watt of any other wavelength with our normal colour vision. If we based our luminous efficiency on photons per second instead of watts, as research papers often do, the graph would peak at a lower wavelength.

The new definition of the candela works with scotopic as well as photopic measurements. Scotopic vision peaks at 507nm, so we have to scale up our efficiencies by $1/V'_{555}$, or about 2.489. If we came up with other standard observers, the candela would be defined for them too.

In section 3.2 we extend the definition of the candela to include colour coordinates.

1.2 Brightness of a broad light source

If you are looking at a luminous monitor or a reflective cinema screen, then the brightness of the surface will be measured in candelas per square metre in SI units. As you move away, the intensity of the source in cd/m^2 should remain the same, but the amount of light reaching your eye will drop as the source subtends a smaller solid angle. If you look at the screen at a different angle, the brightness may change. If the brightness is the same in all directions, then the source is said to be perfectly diffuse or *Lambertian*.

The *lambert* is a similar unit in the old CGS system. A one-lambert surface emits or reflects $1/\pi$ candelas per square cm.

The *foot-lambert* is a strange Imperial cousin of the *lambert*, used in the US motion picture industry. A one *foot-lambert* surface emits or reflects $1/\pi$ candelas per square foot instead of per square cm, making one ft-L equal to 3.426259 cd/m^2 . The spelling and the abbreviation of this unit may vary.

The *nit* is an alternative name for the cd/m^2 . It is not approved by the SI but is used by the CIE, and is becoming a common term in the US.

If you read experimental papers, you may come across the troland (Td). This is the amount of light reaching the retina, being the illumination reaching the eye in cd/m^2 multiplied by the pupil area in mm^2 . This makes it dimensionally the same as a micro candela, but it is only used for light levels within the eye.

Here are some typical values:

Condition	Foot-Lamberts	Candelas/sq.m
Scene in bright sunlight	1500	5000
Optimum for human eye	600	2000
White paper under typical reading light	150	500
SMPTE video standard white	30	100
SMPTE cinema standard white	16	55
Mesopic threshold (rod cells saturating)	0.3	1
Photopic threshold (limit of colour vision)	0.03	0.1

1.3 Wavelength

A spectral band of light may be defined by wavelength, frequency, or energy per photon. At the beginning of this section the luminance unit was defined with respect to light of *frequency* 540×10^{12} Hz. It is rare to define light as a frequency: it is much more common to define light in terms of its wavelength in vacuum. The refractive index of air will mean that wavelengths measured in air are typically slightly smaller. The refractive index of air depends on the humidity and the pressure, so it is hard to quote a typical figure, but it should be less than 1.001. For most colour work you can ignore the difference between wavelengths in air and wavelengths in vacuum.

In this note I shall describe wavelengths in units of nanometers (nm). The visible range is usually given as 380 nm (violet) to 680 nm (red)...

<i>Colour</i>	<i>Wavelength</i>
Extreme red	680 nm
Red	648 nm
Orange	607 nm
Yellow	678 nm
Green	535 nm
Azure	497 nm
Blue	470 nm
Violet	442 nm
Extreme violet	425 nm

These are not standard names.

'Red' is the most saturated looking red. 'Extreme red' is the longest wavelength I could see with an incandescent source. The extreme reds may look slightly more orange than the 648 peak red.

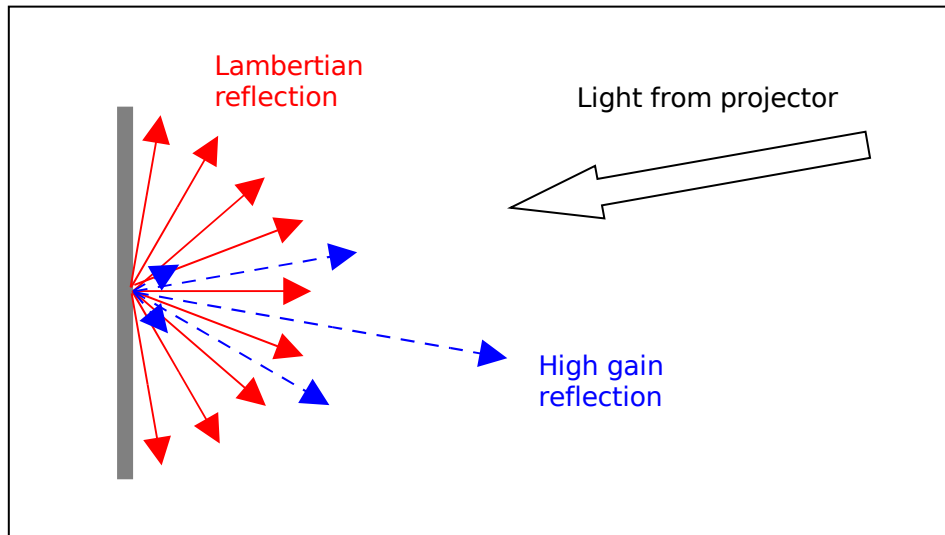
There is no obvious 'orange' wavelength. I have picked the middle of the range of colours.

Yellow could be picked to about 1nm. It is the most precise name in terms of wavelength.

I use 'azure' to describe a blue-green colour. This is sometimes described as 'peacock blue', or 'cyan', or even lumped in with 'blue'. Newton may have been describing this colour when he used the term 'blue', which is why he needed 'indigo' in his spectrum.

1.4 Screen gain

A Lambertian cinema screen would reflect some of the projector light at high angles, spilling it onto the ceiling and walls rather than at the audience.



The *screen gain* is the dimensionless ratio between the peak brightness of the screen and the corresponding brightness of a Lambertian reflector when both are illuminated in a given cinema-like geometry.

A high gain screen - sometimes called a 'silver' or a 'grey' screen' because of its appearance - reflects more light at the audience, making the projector appear brighter. Unfortunately, the screen gain must fall off rapidly with viewing angle, so the viewed image may be dull at the edges, and have a noticeable bright spot – often called a 'hot spot' – centred on the specular reflection of the projector. Real cinema screens are a compromise between brightness and flatness.

1.5 Illumination

The SI unit of illumination is the *lux*. One *lux* is defined as an illumination of one lumen per square meter from any angle. If the surface were perfectly white and Lambertian, it would reflect a brightness of $1/\pi$ cd/m^2 .

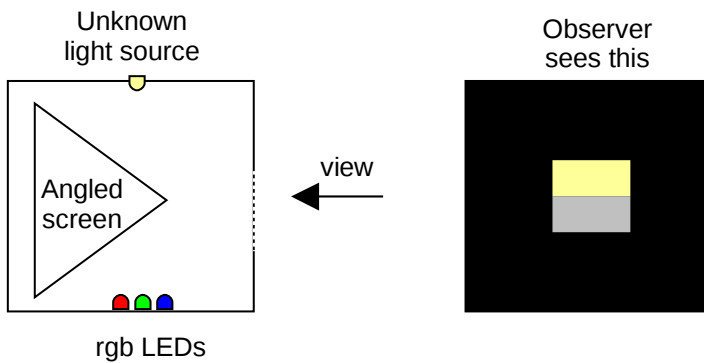
An instrument that measures in lux will typically have a white diffuser so it collects light from all angles over a known area. An instrument that measures in cd/m^2 will typically have optics or collimation to control the solid angle that it is measuring.

Typical scenes reflect about 20% of the light. An 18% grey card used to be used to set camera exposure levels.

One *lux* corresponds to typical auditorium or telecine suite low lighting.

2 Matching a light source

Suppose we have three ideal LEDs, which give out red, green, and blue light, where the LED output intensity is proportional to the current, but the spectrum is fixed. We could use these three LEDs to estimate the colour of an unknown light source using an apparatus like the diagram below.



We adjust the current through the LEDs until the two sides of the screen appear the same colour. We then measure the current through the three LEDs in milliamps. This gives us three rgb numbers that will identify the unknown light source. Of course, our rgb figures would be meaningless to someone who did not have our apparatus.

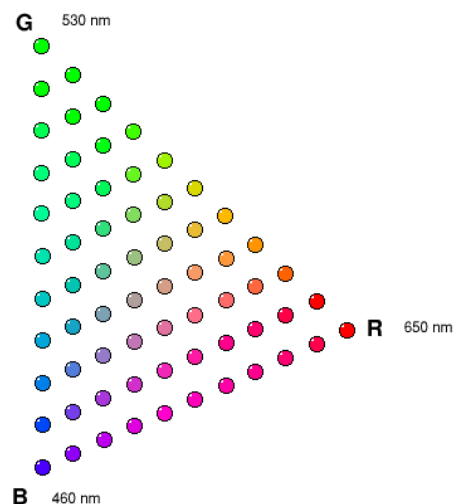
2.1 Matching a colour

Suppose we halved the signal going through our three LEDs. Our LEDs would be outputting the same ratio of red, green, and blue light, so the LED illuminated screen would appear the same colour as our unknown light source, but it would be half as bright.

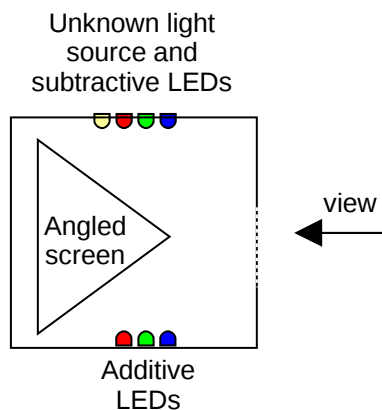
Let us make a new **RGB** = $rgb/(r+g+b)$.

Our RGB colour space has three dimensions. Our new **RGB** colour space has only two independent dimensions, because $R + G + B = 1$. Our scaling has taken out the brightness factor. This does not mean all the colours will appear the same brightness to us. That will depend on our original choice of RGB primaries. But it will mean every colour is represented with a unique brightness.

The diagram opposite shows the triangle of colours we get from a particular set of **RGB** primaries. If you plot colours like this, the colours you get by mixing any two colours will always lie on the straight line joining those two colours.



2.2 Matching an out of gamut colour

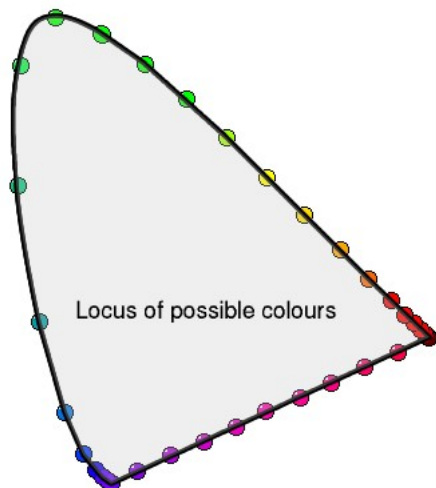
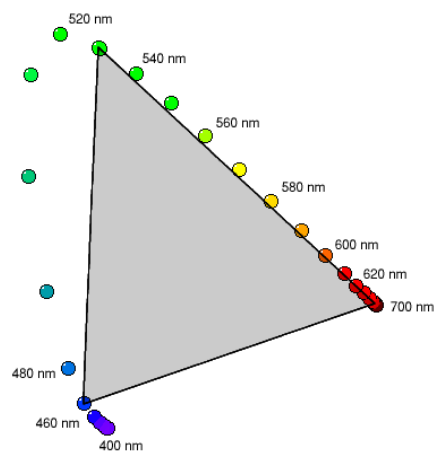


Our **RGB** colour space in section 2.1 has a triangular gamut. There may be colours outside this triangle, but we cannot get a match for them without using negative RGB values. We cannot get negative RGB illuminations from our LEDs, but we can modify the apparatus in section 2.1.

We cannot add negative amounts of blue light, but we can now add light to the other side of the balance, which has the same effect. Now we can have negative RGB values, we can extrapolate outside the **RGB** triangle.

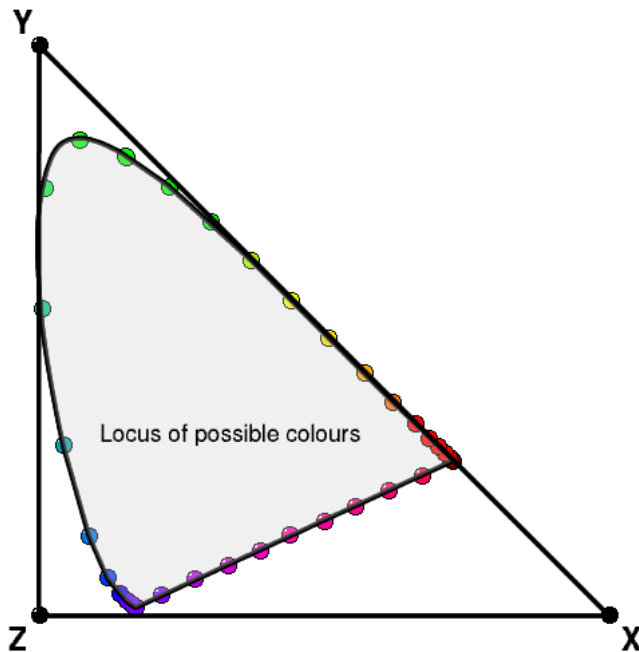
In the diagram opposite, we have plotted out the spectrum at 10nm intervals to the same scale as the diagram in section 2.1. The grey triangle represents the gamut of positive RGB values.

In the diagram in section 2.1 we had primaries of 650nm (red), 530nm (green) and 460nm (blue). The original CIE researchers for the 1931 standard used white lights with coloured filters. These experiments were repeated when suitable lasers came available.



This locus of all the possible colours is the grey region shown in the diagram opposite. The convex curved edge is the locus of all spectral colours. The straight line closing the bottom of the spectral locus is the locus of all colours made by mixing the shortest (violet) and the longest (red) visible wavelengths.

3 CIE XYZ Tristimulus



We have taken the diagram from section 2.2, and added three new points, labelled X, Y, and Z. These three colours are the CIE XYZ primaries, normalised in the manner of RGB in section 2.2.

We have not chosen our XYZ points to form a nice, right-angled triangle. Rather, we arranged our earlier plots, so things would turn out this way. The CIE reasons for choosing their XYZ primaries are more cunning.

The 1931 CIE researchers used primaries of 700nm (red), 546.1nm (green) and 435.8nm (blue) to measure their spectral locus. These primaries did not have equal brightness, so they could choose an extrapolated X and Z with zero visual brightness. They could do this because XYZ are not real colours.

Y is the only primary with brightness. It is the Y brightness value in colour spaces such as Yxy, and Yu'v'.

NB: Y is not the Y in video colour spaces such as YIQ and YCrCb.

The XY line was chosen to be a tangent to the spectral locus.

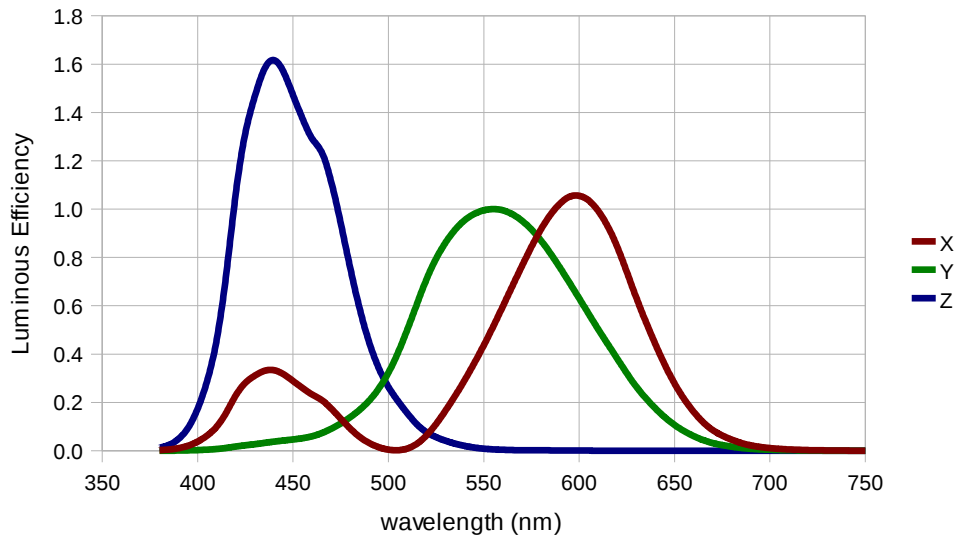
The ends of the spectral locus are the limits of human vision. These would seem to be unique points in the visual colour space. However, it is difficult to locate these points reliably, because the eye has very little sensitivity to these wavelengths. The ends of the spectral locus do not have any significant values in XYZ. The violet point does not lie on the ZX line.

The point X=Y=Z is a whiteish colour. It has no particular significance.

3.1 Measuring tristimulus values

We can calculate the relative spectral contributions from the diagram of section 3.

The CIE published the original standard observer colour matching functions in 1931. There have been several revisions since then as better monochromatic light sources have become available. This graph shows the CIE 1931 observer as modified by Vos (1978).



The main changes have been in the blue-violet region (350-450 nm), increasing the bend at the violet end of the spectral locus.

This can be important: not all CIE XYZ measurements are quite the same. You can get slightly different values from the same spectrum, depending on the standard observer functions. Once you have converted to XYZ, there is no way to correct for the differences in standard observers because you cannot reconstruct the original spectra from XYZ. There is a lot of published work based on the original 1931 observer, which is why 1931 XYZ is still used, even though more accurate measurements exist.

Actually, for most practical work, there is little difference between the different CIE colour matching standards. Some phosphors have a violet emission line, but most artificial light contains little violet, and the eye is not particularly sensitive to violet anyway. Often, the differences between different colour matching functions are smaller than the experimental or calibration errors you get with real colour measuring instruments. Truelight does not export or import XYZ images. If you use the same instrument for all your XYZ measurements, any systematic errors should cancel themselves out.

3.2 XYZ units

XYZ measurements are usually given in foot-lamberts (ft-L) or candelas/sq.m (cd/m²).

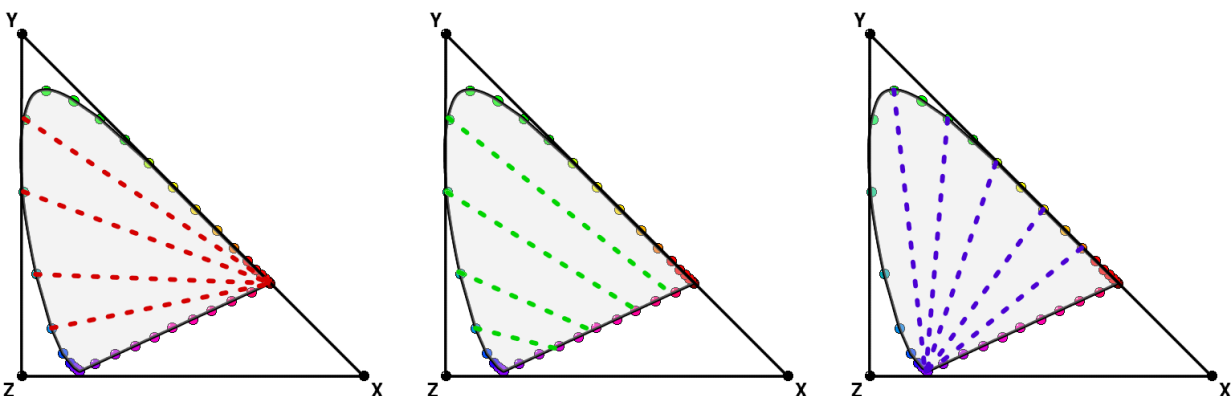
The Y values in the CIE XYZ weights are the same as the CIE 1924 standard photopic observer values plotted in section 1.1. Clearly we can express Y in cd/m². But what about the other two channels? Do we scale the weights as we did earlier to get the same value at 555 nm?

We do not scale the values. X=Y=Z is a whiteish colour which suggests that the X, Y, and Z values should be similar too. This X=Y=Z colour corresponds to a white where every waveband had an equal energy. This is arbitrary but as good a colour as any other. The CIE decided to stick with this white, and give the X and Z components the same units as Y.

3.3 The human eye LMS primaries

Our eyes generate a red, green, and blue signal. Why not model these signals instead of using XYZ?

We can get an idea of the actual spectral sensitivities of the human eye by investigating people with defective colour vision. A person with normal three-component colour vision would see each point in the locus of possible colours as a unique colour. People who have only two-component colour vision confuse colours on the dotted lines.



The dotted lines in each diagram radiate from the missing primary. The eye primaries are conventionally labelled LMS for Long, Medium and Short wavelength. The L primary is close to the longest red, and the S primary is closest to the shortest violet. The M primary is off the graph with a negative Y. The LMS primary triangle contains no real colours.

LMS primaries are used in colour vision research. Those LMS values are usually based on experimental estimates of cone spectral sensitivities, and not on the XYZ primaries as our construction suggests. Some modern standard colour spaces such as CIECAM02 use a matrix to transform XYZ into approximate eye primaries.

4 Luminance-chrominance co-ordinates

4.1 Yxy co-ordinates

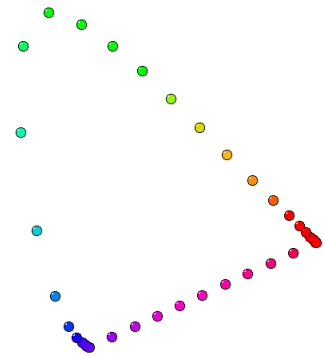
Yxy is the simplest luminance-chrominance space. It is derived from XYZ using...

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

Y is the luminance measured in foot-Lamberts or candelas/sq.m, and the x and y co-ordinates are dimensionless. The inverse transform is...

$$X = \frac{x \cdot Y}{y}, \quad Z = \frac{(1-x-y) \cdot Y}{x}$$

The xy colour space is simple, but it is uneven. A small difference in x and y in the blue-violet corner is more visible than the same difference in the green corner.



4.2 Yu'v' co-ordinates

Yu'v' is a uniform luminance-chrominance space. It is derived from XYZ using...

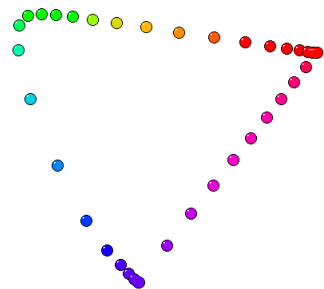
$$u' = \frac{4X}{X+15Y+3Z}, \quad v' = \frac{9Y}{X+15Y+3Z}$$

Y is the luminance measured in foot-Lamberts or candelas/sq.m, and the x and y co-ordinates are dimensionless. The inverse transform is...

$$X = \frac{9Y u'}{4v'}, \quad Z = \frac{3Y(4-u')}{4v'} - 5Y$$

The blue-violet corner has been stretched out and the green corner has been compressed. It is not perfect, but it is much better than Yxy. Additive mixtures still lie on straight lines.

Yuv was an earlier attempt at a uniform colour space that was replaced by Yu'v' in 1975. There is no reason to use Yuv today.



5 Perceptually uniform colour spaces

Our eyes adapt to different illumination conditions. An object will seem to have the same colour to us under very different lighting conditions. Our adaption is not perfect, but most of the time and over a range of lighting conditions, we see colour as a property of an object, and not a property of the illumination.

This suggests that any 2:1 luminance contrast should look the same to us, irrespective of the absolute luminance. The contrast between 20 and 10 ft-lamberts ought to look the same to us as the contrast between 200 and 100 ft-lamberts, or 2.0 and 1.0 ft-lamberts.

There are limits to this. The table in section 3.2 suggests we may get dazzled if our illumination goes much over 2000 ft-lamberts, and we will not see colour at 0.2 ft-lamberts. We would not see a contrast between 2.0 and 1.0 ft-lamberts in an average luminance of 100 ft-lamberts, because of stray light within the eye. However, there should be a luminance range where equal contrast ratios gave equal eye stimuli.

Suppose we have two grey patches with luminance of 90 ft-lamberts and 10 ft-lamberts. A patch of 50 ft-lamberts would be halfway between the two in luminance units. A patch of 30 ft-lamberts should appear to be halfway between the two, because it has a 3:1 contrast ratio either way.

5.1 CIE L*a*b*

The CIE L*a*b* is defined in terms of the stimulus XYZ and the reference white X_n, Y_n, Z_n using a function f as follows:

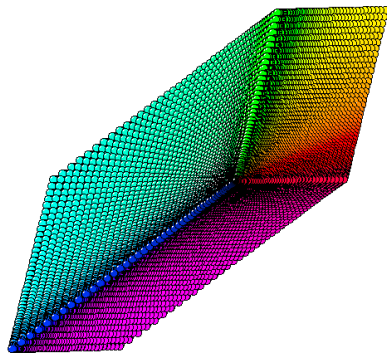
$$f(W) = \begin{cases} W^{1/3} & W \geq 0.008856 \\ 7.787 \cdot W + 0.1379 & W \leq 0.008856 \end{cases}$$

$$L^* = 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16$$

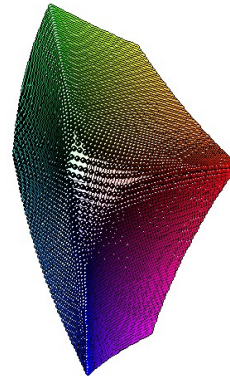
$$a^* = 500 \cdot \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right]$$

$$b^* = 200 \cdot \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right]$$

The f function has a cube root region that approximates to the constant contrast response we anticipated in section 5, and a linear region for very deep shadows where we expect the constant contrast model to break down. The function should be smooth and continuous about the value 0.008856.



A video RGB colour gamut plotted in CIE XYZ



A video RGB colour gamut plotted in CIE LAB

For these diagrams, we made a list of all the RGB points with values 0,5...255 on the surface of the ITU rec709 video gamut, calculated the tristimulus values, then plotted them out in XYZ and $L^*a^*b^*$. In both views, we are looking at the black corner of the gamut. The colours should be correct when viewed on a monitor with a gamma of 2.0.

The XYZ gamut has straight edges, because the red, green, and blue phosphors are independent. The light points are spaced out, but the dark points are all crammed together. This is a result of the ITU monitor gamma, rather than a property of the XYZ space. The $L^*a^*b^*$ gamut contains the same colours as the XYZ gamut, but we have expanded the dark regions where our colour resolution is good, and compressed the cyan region where our colour resolution is poor.

The luminance co-ordinate is L^* . L^* values can go from zero (black) to 100 (white). Negative L^* values are errors. L^* values beyond 100 should not happen if the reference white is chosen properly.

The chrominance co-ordinates are a^* and b^* . The reference white has zero a^* and b^* . You can get values of a^* and b^* greater than +100 and less than -100 for very saturated colours. The exact gamut depends on the reference white. The a^* value is positive for reds, and negative for greens. The b^* value is positive for yellows and negative for blues.

5.2 CIE L*u*v*

The CIE have another colour space called L*u*v*. It uses the same luminance value L as in L*a*b* (section 5.1) and the u*v* chrominance co-ordinates come from u' and v' are defined in section 4.2...

$$u^* = 13L(u' - u'_{\text{white}}), \quad v^* = 13L(v' - v'_{\text{white}})$$

L*u*v* is marginally easier to calculate, but it is less good at predicting colour shifts than L*a*b*, and it can predict colours outside the spectral locus.

There are no good reasons for using L*u*v* these days.

5.3 Delta E

ΔE^*_{ab} is the difference between two L*a*b* colours as given by...

$$\Delta E^*_{ab} = \sqrt{(L^*_0 - L^*_1)^2 + (a^*_0 - a^*_1)^2 + (b^*_0 - b^*_1)^2}$$

CIE L*a*b* space is based on measurements of just perceptible differences between colours. If any two colours are separated by a ΔE^* of 1.0, then you should just see the contrast between them under certain idealised viewing conditions. This is not wholly accurate, but it is close enough to be a useful measure of perceived contrast.

Because ΔE^*_{ab} is based on just perceptible differences, it does not necessarily follow that it is a reliable measure of large differences. Two colour pairs of colours with a ΔE^*_{ab} of 30 may not show the same contrast.

ΔE^*_{uv} is the difference between two L*u*v* colours as given by...

$$\Delta E^*_{uv} = \sqrt{(L^*_0 - L^*_1)^2 + (u^*_0 - u^*_1)^2 + (v^*_0 - v^*_1)^2}$$

This is not the same as ΔE^*_{ab} , but it is similar. A ΔE^*_{uv} contrast of 1.0 should just be visible. ΔE^*_{uv} is occasionally useful. When comparing two white points with $L^* = 100$ it reduces to...

$$\Delta E^*_{uv} = \sqrt{(u^*_0 - u^*_1)^2 + (v^*_0 - v^*_1)^2}$$

5.4 Colour appearance modelling

We have described some of the 'classical' CIE colour spaces. These give us the ability to measure the differences between colours, and, to a limited extent, to predict the effects of changing the white point or the overall luminance. This is enough to allow us to simulate the appearance of a cinema image on a monitor in a darkened room with a similar brightness and surround. They cannot really be used to compare, say, the image we see in a cinema against the original daylight scene, which may have a thousand times the intensity, and no clear surround.

There are more general colour appearance models that cover these conditions. The most recent CIE colour appearance model at the time this was written is called CIECAM02. This came out in 2002, replacing CIECAM97. It is widely recognised that there are still significant aspects of colour appearance phenomena that are not described well. The colour differences are not always accurate, and spatial and temporal effects are not included. People have proposed iCam, or Image Colour Appearance Modelling as an even more complex spatial and temporal colour standard.

The general colour appearance models are very complex, and they are still evolving. There are calls for the adoption of simpler standards such as ZLAB, that are less general but easier to use.

Truelight is used for comparing displays under similar viewing conditions. In the most critical grading conditions, the display and the reference image should be viewed in a darkened room, and both images should have similar brightness and white point, and the images are bright enough for normal colour vision. If we restrict ourselves to these sorts of viewing conditions, then Truelight has no real need to use colour appearance models. Truelight has a simple flare correction for less critical pre-grading work under office lighting conditions, but this is not intended for use when high colour accuracy is important.

In the future, we may use Truelight to match the appearance of the film image to the original scene. If we do this, we may have to cope with dramatic differences in brightness when rendering daylight scenes.

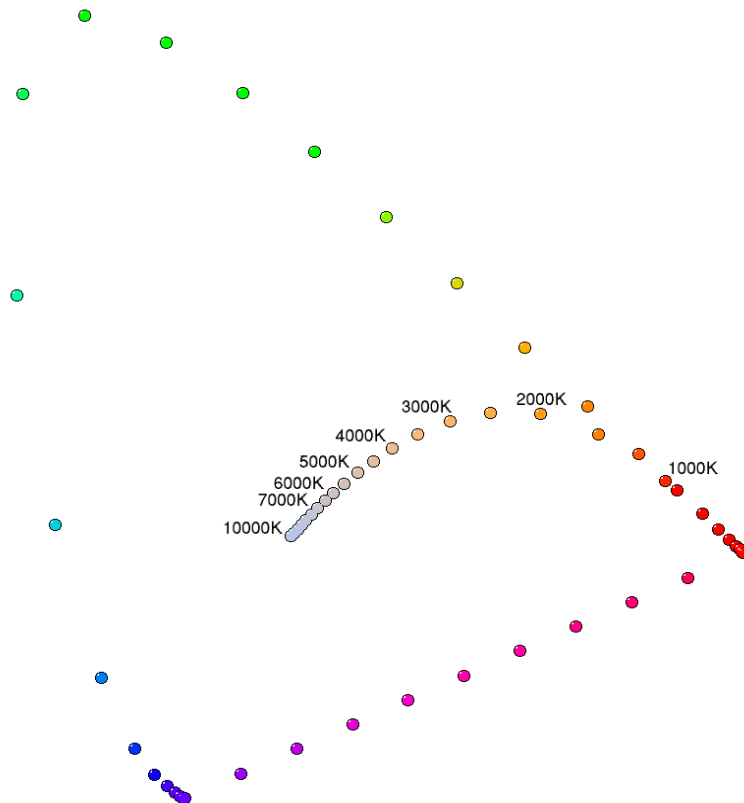
6 The Luminous White Point

The reflection colour white can be defined. An idealised white reflects all visible wavelengths. Practical whites reflect a similar fraction of all visible wavelengths.

There is no corresponding definition for the whiteness of a light source. We can see reflection colours by daylight, incandescent light, fluorescent lamps, or arc lamps. All of these light sources may seem 'white' to us when viewed in isolation, but may seem coloured when compared to other lights. Our eye-brain system somehow identifies the dominant light source, and adjusts for it.

We can see many colours as white given the right conditions. However, most illuminants we normally regard as white lie close to the set of colours we get from incandescent sources, even though the physics behind the light sources are very different. The D55 and D65 daylight standards have very different spectra to the 5500K and 6500K Planck's law for thermal spectra, but their colour values are very similar.

Here is the thermal locus plotted relative to the spectral locus we saw in 2.2. In this diagram, we plot the colours with 6500K as neutral. Any point down to 2000K can look white to us.



6.1 Colour temperature

Because most 'white' useful light sources lie close to the blackbody curve, it is common to quote light source colours in terms of the temperature of the nearest thermal point. Here are some typical values...

Light source	Temp
Old monitor white	9300K
Old monitor white	7200K
CIE D65 daylight standard	6504K
Video ITU rec. 709 white	6500K
German cinema standard	5500K
Kodak slide viewer recommendation	5500K
Kodak Cineon view recommendation	5400K
US Cinema standard	5300K
European cinema standard	5200K
ICC print viewing D50	5000K
Emerging art museum viewing standard	3700K
Steenbeck (incandescent)	3000K
CIE illuminant A	2856K

People often quote colour temperatures. This reduces our two colour co-ordinates to a single easily remembered physical value. Unfortunately, it does this by ignoring the offset in the green-pink direction perpendicular to the thermal locus, so it is only meaningful when this offset ΔE is less than one (see section 5.3).

Colour temperature is also not a perceptually uniform measure. At the beginning of this section we plotted out points in 500K intervals. The 2000K point is well separated from its neighbours, while the 9500K point is overlapped by its neighbours. An infinite thermal temperature has a sensible blue-white colour according to Planck's law.

Sometimes colour temperatures are measured in mireds, or micro-reciprocal-degrees. A colour temperature of 2000K is 500 mireds. Measurements in mireds are more perceptually uniform. You can also get colour correcting filters with positive or negative mired values, which will, when added to the mired value of any thermal source, give the mired value of the filtered light. For example, a -100 mired filter will change our 2000K 500 mired source to a 2500K 400 mired source.

6.2 Calculating Colour Temperatures

If our colour is on the thermal locus then we will have an exact figure for the colour temperature. As we go away from the thermal locus, our temperature figure becomes slightly ambiguous. It would be logical to choose the temperature of the nearest point on the thermal locus. However, our choice of 'nearest point' will depend on our choice of colour space.

Unfortunately, there is a tradition of calculating colour temperatures using the 1960 uv space. This is similar to the u'v' space of section 4.2, but the scaling of the v axis is different...

$$u = u' \quad v = \frac{2}{3} \cdot v'$$

Back in 1976, this seemed like a good thing to do – it allowed new colour temperature measurements to be compared to told ones. In fact, the backward compatibility only goes back as far as 1968, when the current value of the Planck constant was adopted.

Unfortunately, this change in values can make a significant difference to the reported temperatures. The DCI white point ($x=0.314$, $y=0.351$), for example, is closest to 6000K in u'v' space, but to 6350K in uv space. The ΔE values in both cases are very similar (about 19.0).

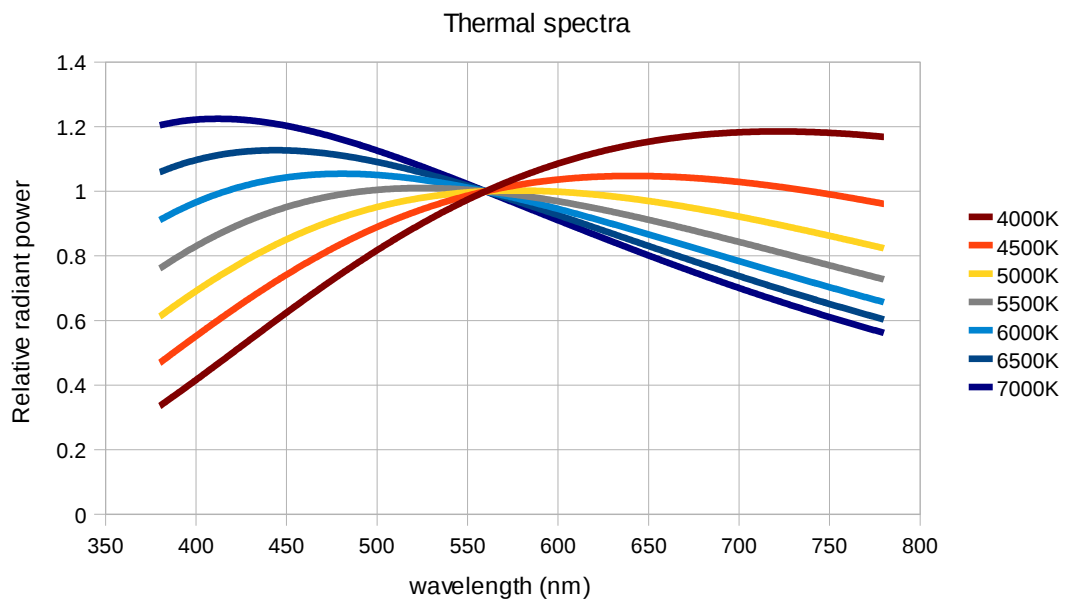
There is no great physical significance to these differences in colour temperature. The greenish-yellow colour of the DCI white point does not look particularly like 6000K or 6350K. If you go further away from the thermal locus, colours such as bright green do not have meaningful colour temperatures at all. Nevertheless, 6000K and 6350K are very different temperature values, so it is important to be consistent.

The Truelight tools calculate colour temperatures in u'v' space by default. This is more perceptually uniform than uv space, so the visual match should be better. When the ΔE values is signed, positive ΔE is for offsets on the green side of the thermal locus, and negative ΔE is for offsets on the pink side. There may be an option for calculating traditional uv colour temperatures.

Most documents and colour measuring instruments still give colour temperatures calculated in uv space.

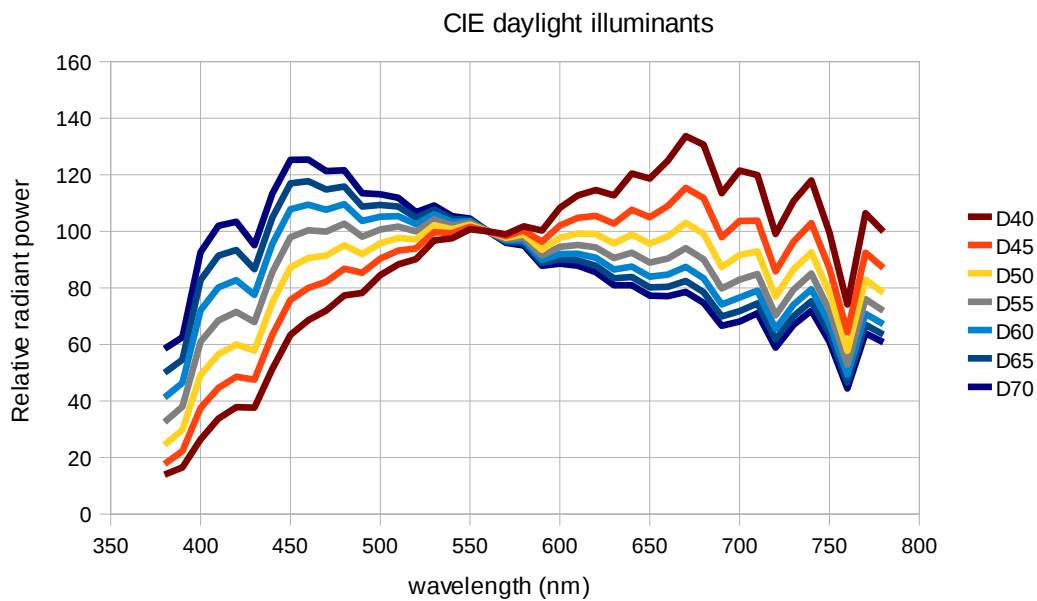
6.3 Thermal spectra and CIE daylight illuminants

A hot body will give off radiation. The spectral power distribution of this radiation will normally depend on the shape of the body, and the material that the body is made of, as well as the body's temperature. However, if we have a large cavity at a uniform temperature, and we look at the radiation through a small hole so we do not disturb the radiation field significantly, then the spectrum will depend on the body temperature alone. These spectra are variously called blackbody spectra, thermal spectra, incandescent spectra, cavity spectra, or Planckian spectra...



The CIE standard illuminant A is the thermal spectra of a 2856K body. Tungsten has a melting point above 3650K, so you cannot get any of the spectra in the graphs above directly from a solid body, though you can get good approximations using the mired filters described in the last section.

The CIE standard daylight illuminants are based on measurements of daylight spectra. The daylight spectrum was measured under a variety of conditions, and reduced to a mean spectrum, and two eigenvectors, which can then be used to calculate a whole family of spectra. These spectra are usually named according to the approximate colour temperature. D65, for example, is a Daylight illuminant with an approximate colour temperature of 6500K.



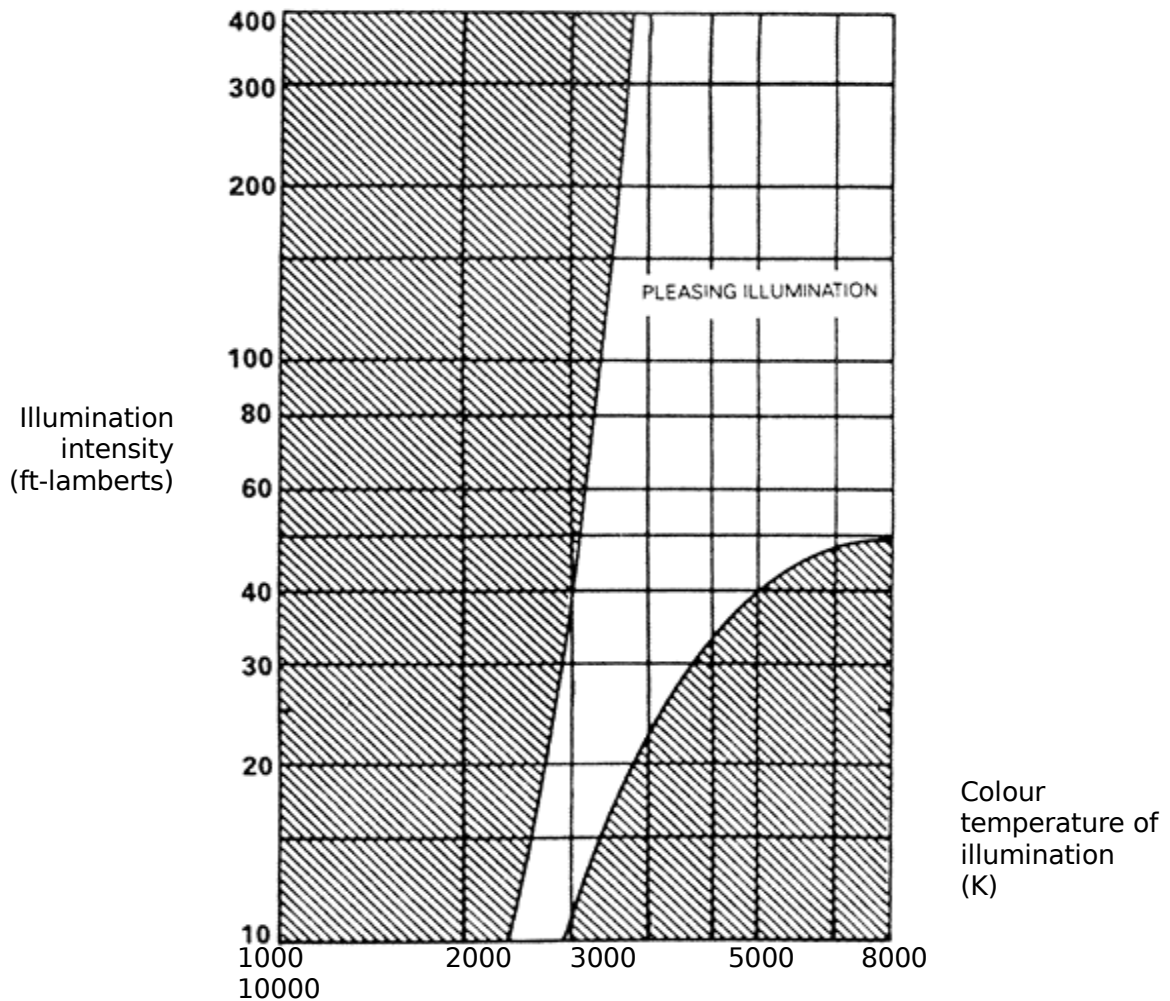
You cannot get a standard light source with any of these spectra: the complex and jagged spectral distribution is hard to reproduce in an artificial lamp. These CIE daylight simulation illuminants only exist as a theoretical ideal for calculations such as the colour rendering index described in section 6.5. When people describe a fluorescent tube as being D65, they probably mean the colour is matched to D65.

There is no reason why the CIE daylight illuminants should fall on the thermal locus. The points come close, but there are differences...

Thermal	x	y	CIE	x	y
4000K	0.3820	0.3792	D40	0.3823	0.3838
4500K	0.3620	0.3656	D45	0.3621	0.3709
5000K	0.3460	0.3532	D50	0.3457	0.3587
5500K	0.3330	0.3421	D55	0.3325	0.3476
6000K	0.3224	0.3324	D60	0.3217	0.3378
6500K	0.3137	0.3239	D65	0.3128	0.3292
7000K	0.3063	0.3164	D70	0.3054	0.3216

6.4 Psychophysical phenomena

In 1941, A.A.Kruithof published a graph (plotted below), which summarised the relationship between colour temperature, intensity, and the 'pleasant' quality of a light source. For example, a space will appear white when illuminated with 6000K light at 100 ft-lamberts, but will appear grey and gloomy when illuminated at 20 ft-lamberts. This is the sort of effect that makes a dimmed fluorescent tube look 'grey', even though 'grey light' should not be possible.



A cinema screen will typically have a colour temperature of about 6000K, and an open gate intensity of 16 ft-lamberts. If we assume the open gate white corresponds to the colour of a matt white object in the original scene (not necessarily true, but probably not too far out for many scenes on real films) then our view in a cinema would be like looking at the scene lit to 16 ft-lamberts. The Kruithof curve should predict that a typical cinema scene looks really gloomy and unpleasant.

A typical cinema scene does not look gloomy and unpleasant. Our eye-brain system may associate certain colour temperatures with certain illumination levels, but it is sophisticated enough to learn to reject these associations under special circumstances. The Kruithof curve is not wrong, but we have learned to override it when we are in a cinema.

There is a similar effect when viewing CRTs under office lighting. You could match a monitor to typical 3500K office lighting using instruments, but it would appear as orange as the 3500K dot in the diagram. Even standard 5000K lighting looks strange. We seem to know that CRTs have their own white - about D65 - and we do not expect them to fit the ambient white. I have not found this effect described in the literature, but it was at one time well known in printing, where it was common to view images on monitors and compare them to printed images under standard lighting. As monitors are increasingly set up to the video D65 white, we might expect this effect to become stronger in the future.

There is not much we can do about these effects. They are hard to measure objectively, and they may be different from one person to another. We cannot consistently correct for the degree to which we 'know' that a CRT does not use ambient light, or 'know' that a cinema image is not a real scene. Like optical illusions, the best we can do is to know they exist, and to keep alert.

6.5 Colour Rendering Index

The D65 lighting standard was designed to match noonday ambient light. Typical D65 tubes will use five phosphors to give a similar, though not identical, power distribution from red to violet. A cheaper D65 light might only use three narrow-band phosphors. 6500K is beyond the melting point of tungsten, but we could match D65 using a negative mired filter. We could even match the white of D65 using just two wavelengths - for example red and blue-green. All of these illuminants would look the same colour; whites and greys would look the same; but colours would look very different. In the most extreme case - the two-wavelength white - a full colour image would appear to have only two primaries.

The CIE have defined a number, called the Colour Rendering Index. This is a measure of how well the light source will match an ideal incandescent light source. A perfect match (another incandescent source) would have a colour-rendering index of 100. A good illuminant such as our five-phosphor tube or a high pressure Xenon arc should have a colour-rendering index of 90 or more. A poor illuminant such as a high-pressure sodium lamp would have a colour-rendering index of 25 or less.

The colour-rendering index was designed to test the ability of the light source to render arbitrary colours. It is not suited to choosing a bulb for a projector that will pass light through three fixed film dyes. A lamp with a high colour-rendering index will probably behave a bit like a Xenon arc lamp, which also has a high colour-rendering index. However light sources with lower colour rendering indexes can actually make better projector lamps. A lamp with dark bands between blue and green, and between green and red will give brighter, more saturated colours with typical film dyes. Nevertheless, if you want to match the current industry standard, then a high colour-rendering index is probably a good thing.

7 Densitometry

A densitometer consists of a light source and a detector. A transmission densitometer will have the light source on the opposite side of the film from the detector. The instrument measures the transmittance of the film in density units:

$$\text{density} = -\log_{10}(\text{transmittance})$$

A clear film has zero density; a film that transmits 10% of the light has unit density, and so on. Measure with no film to set the densitometer zero.

Typical films scatter light, as well as absorbing it. In an ideal world, we would choose our densitometer optics to match our application. If we wanted to match the appearance of a typical f/2.0 projection system, we would want an f/2.0 light source and detector. If we wanted to measure the appearance of a film on a light box, we would want a diffuse light source and a collimated detector. In practice, many densitometers use a 90-degree (f/0.5) input and exit cones. This design makes the densitometer insensitive to light scattering artefacts such as fingerprints, but it does mean the densitometer will be picking up some scattered light that would not be seen in a projected image.. Fortunately, most colour film stocks do not scatter much light, so this does not affect the measurement much. Film greys made from silver grains with a hard high-contrast edge will scatter much more light, so you may have to take extra precautions when measuring black and white film.

The first film densitometers measured black and white material. Film greys made from silver deposits will generally have the same transmittance for all wavelengths, so we would expect different light sources to give similar density readings. However, the scattering can give yellowish tints to silver images. If you wanted to measure the density of the image to estimate how it would print, then you would use a bluish light, because orthochromatic print stock is more sensitive to blue light.

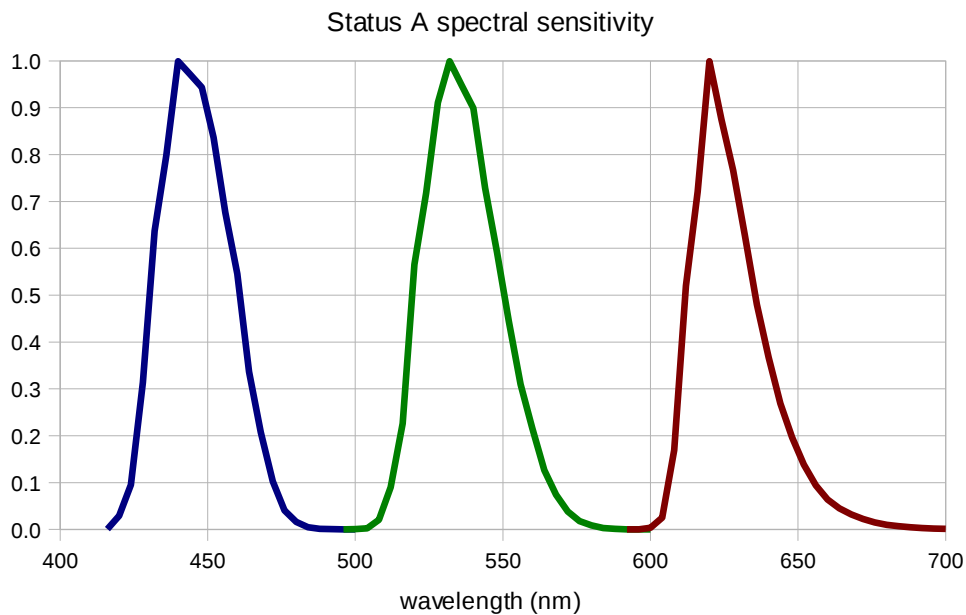
Colour film can absorb different amounts of red, green, and blue light. We can measure the density to red, green, and blue light using three monochromatic densitometers with different light sources. Usually, it is more convenient to use a single instrument with one light source and three filtered detectors.

Before making colour measurements, the densitometer should be zeroed. If we are taking absolute density measurements, the densitometer is zeroed without film. If we are taking measurements relative to the film base, then the densitometer is zeroed on a bit of clear film base.

There are three ISO standard sets of filters, known as Status A, Status M, and Status T.

7.1 Status A colour densitometry

Colour print film has cyan, magenta, and yellow dyes. A densitometer with status A (Analytic) filters is used for measuring how much of these dyes are present. They are filters with well-separated sharp peaks...



Status A measurements are designed for measuring combinations of cyan, magenta and yellow dyes. If you measure other materials, you may get misleading results. A material that absorbed 500nm (blue-green) or 600nm (orange) would give the same status A measurements as clear film, but might look very different. A neutral bit of colour film might have the same status A measurements as a piece of black and white film, but it would look lighter because combinations of cyan, magenta, and yellow dyes have transmission peaks at about 500 and 600nm.

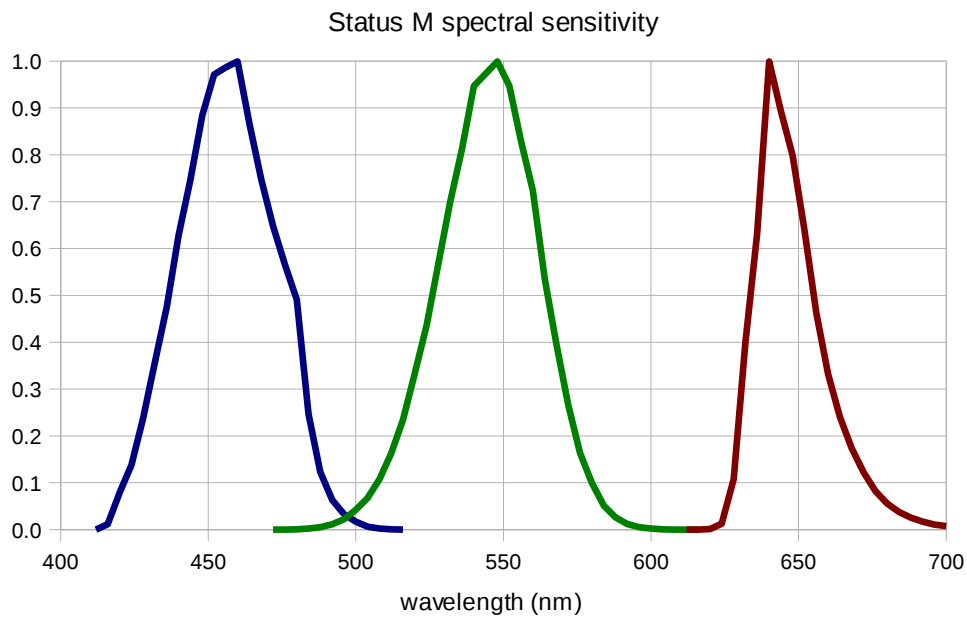
Status A measurements can be made with a heat-absorbing filter. This makes the spectral sensitivity cut off faster as you approach the infrared. Under most circumstances this should make little difference to your measurements.

Status T filters are used to measure positive print densities for colour separations. The peaks are narrower, and more separated.

7.2 Status M colour densitometry

Status M measurements are useful for predicting how negatives will print. Modern colour film printing is done using an additive lamp house. An additive lamp house will have three light sources matched to the red, green, and blue sensitivities of the colour print stock. A densitometer that is matched to print conditions should have similar spectral peaks to the additive lamp house.

The status M spectral sensitivity curves are like this...



The ISO status M standard is used for most negative work, including motion pictures, still pictures, and reversal prints. The SMPTE RP 180 recommended practice paper describes a similar densitometry standard that is aimed specifically at the motion-picture printing industry. Unfortunately, there is not yet any equivalent to the tried and tested Xrite TR310 that supports RP 180.

8 Cineon Exposure Space

'Cineon' is a Kodak trademark for an image processing package, and an image file format. The image file format has a standard way of representing negative densities in 10-bit data that is been widely used in the film industry. For the rest of this section, we shall use the term 'Cineon' for this colour space.

You may think it very strange to define an image in terms of negative densities. There are many ways of printing an image. You can completely change the appearance of the image by changing the printing stock or the exposure. You can get different coloured densities on the negative when you use a different film recorder, as the Cineon calibration only is done along the neutral axis. However, Cineon space has useful properties for calibration, and it is consistent enough for some people to exchange images.

The most popular Cineon file format packs one pixel's RGB values into 10 bits of data. In this note we shall describe Cineon values as 10-bit integers from 0 to 1023.

The graph shows the relationship between Cineon value and negative density for an aim gamma of 1.0.

The graph is a straight line. Each Cineon unit corresponds to a density shift of $0.002 * \text{aim_gamma}$. Low Cineon values are light on the negative and dark on the print. Cineon RGB images viewed without colour correction on a monitor look like a washed-out version on the print.

Kodak defines absolute RGB 'Dmin' densities for each of their intermediate stocks. In the Kodak documentation this is described as the blackest black that can be recorded onto that stock. It is approximately the black you get when you take a picture of a 1% black card. They assign the Cineon reference black value 95 to this Dmin density.

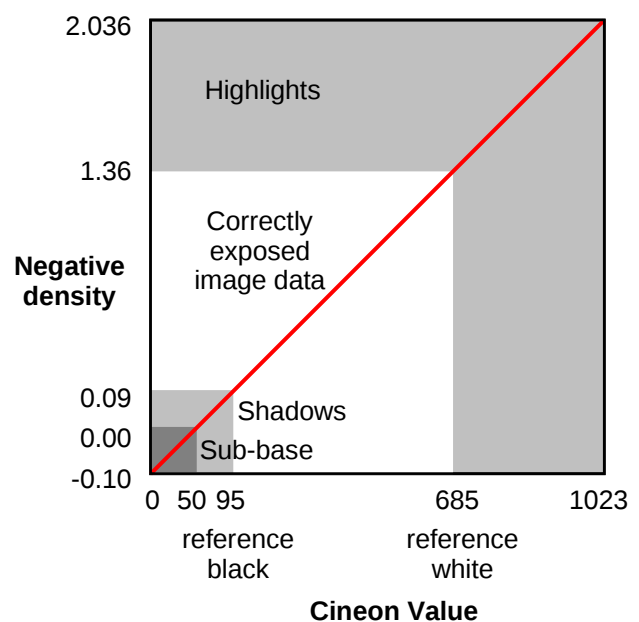
The Kodak formula for the absolute density is...

$$D_{\text{absolute}} - D_{\text{min}} = (\text{value} - 95) * 0.002 * \text{aim_gamma}$$

We show the film base density as having a base density of 50. If you stick to the Kodak definition, then the actual Cineon base value will depend on the actual choice of film stock. The actual Cineon value of 5242 base stock is 42.

In practice, you may not have Dmin values if you are not using Kodak stock, or you may prefer to measure negative densities relative to base. We suggest a nearly equivalent alternative formula...

$$D_{\text{absolute}} - D_{\text{base}} = (\text{value} - 50) * 0.002 * \text{aim_gamma}$$



8.1 Reference White and Reference Black

Cineon space covers all the useful density range for conventional negatives. A Cineon image usually contains more information than you can see in a single print. You can vary the exposure in a Truelight simulation and see more detail in the highlights or shadows.

There is a subset of the Cineon space, with values between 95 and 685 that is commonly used for 'correctly exposed' image data. Kodak documents call 95 the 'reference black', and 685 the 'reference white'. Fixed conversions from Cineon to display RGB used to clip values outside this range. Old tube based film recorders also used to limit their output range to the 'reference white' or some other similar value to preserve tube life.

You are not restricted to the 'correctly exposed' range. Nevertheless, there are reasons for keeping most of your image within these limits...

- The negative exposure curve has its 'toe' beneath the reference black. Down at the toe, accurate interpolation is difficult, and the shape of the tone curve can vary with processing chemistry or the age of the negative film stock. You can avoid both these problems by adding 50 to all your Cineon values, which reduces the transmission of your negative by 20%, then increasing the exposure by 20% when you print.
- Most print stocks have a tone curve with the 'shoulder' above the reference white. This print shoulder has all the same problems as the negative toe. You can avoid these problems by reducing your white to a light grey on the print. The image will look dimmer when viewed on a light box, but will look almost the same in a cinema.

8.2 Aim Gammas

If our Cineon step corresponded to a shift in negative density of 0.002 in all three channels, then a set of patches with equal values in all three channels would give a neutral negative. Unfortunately, a neutral negative wedge will produce a print with the darker colours looking distinctly bluish on Vision stock.

Kodak tabulate a set of aim values in their H-387 document “Kodak Digital LAD test image”. If you plot out these values, they lie pretty close to a straight line. Kodak worked on the RP-180 densitometry standard mentioned briefly at the end of section 7.2, so aim gammas may have been a correction for the differences between RP-180 and status M measurement. Other people have used aim gammas as a correction for the colour shift of the last paragraph.

The Kodak aim gammas for Vision film stock are 0.966, 1.063, 1.087. These are typical values: they are close to 1.0 and the red value is lower than the others.

8.3 Calibrating a Recorder

Usually, film recorders have to be calibrated before taking Cineon data. The recorder is loaded with the appropriate film stock. A set of neutral test patches is output. These test patches should cover at least the range from reference black to reference white. The densities of these patches are measured on a densitometer such as an Xrite TR-310. The measured density values are used to make a look-up table that converts Cineon values to the value to produce the correct density on a neutral patch.

Unfortunately, we need these look-up tables to generate the neutral patches on the test strip. The first time a recorder is calibrated, the usual approach is to generate a test wedge using some supplied set of look-up tables. These are unlikely to give a wholly neutral test strip, but the calibration will be an improvement on the original tables. After a couple of cycles, the test strip should be accurately neutral.

8.4 Exposure Units

Film labs have controls on their printer lamp house that set the RGB exposure. These controls are often marked in units of printer points, with 12 printer points to a stop. Increasing the printer point setting will increase the exposure through the negative, which will make the print darker. Sometimes people quote a figure of 8 printer points to a stop, instead of the usual 12. This is about right when referring to stops on the camera that took the original film, and assuming a typical camera film gamma.

Laboratories work in absolute printer points. The absolute printer point setting to get a given print result on a given stock will be different for each laboratory, and will change with the condition of the chemicals. People requesting prints from a laboratory will often ask for relative exposure shifts in printer points, and leave the absolute settings to the laboratory.

A shift of -1 printer point is equivalent to....
An exposure filter factor of 0.9439
An increase of +0.025 in the negative density.
A shift of +12.5 Cineon values
A shift of about -0.08 in typical mid-tone print densities

9 Video

The original video standards were written for analogue equipment. They defined how analogue luminance and chrominance signals relate to red, green, and blue voltages, and how these voltages relate to light levels in the video image. The new, digital standards retain a surprising amount of this analogue heritage. Digital video engineers use digital equivalents of the traditional analogue tools, such as PLUGE tests and vectorscopes. Digital RGB signals can be converted to luminance-chrominance YCrCb, with luminance sub sampling to save bandwidth, just like the old analogue signals, then converted back again at the other end. In a long video pipeline you may have several such conversions, and the digital errors can accumulate.

9.1 Digital to Voltage conversion

In the standard, the red, green, and blue voltage signals are called R', G', and B'. The voltage is linearly related to the value in the digital standards. R', G' and B' have these ranges:

Limit	signal	V	8-bit value	10-bit value
Black point	0 mV	0	16	64
White point	700 mV	1	235	940

Digital video values can go outside this range. Video that is clipped to this range is known as 'legalized video'. Values outside the range 1-251 (4-1004) can give overflow or underflow errors depending on the hardware, particularly when converting to luminance-chrominance values (see section 9.4). A bright red (255,0,0), for example, may display as a dull green.

Values outside the legal range are not usually allowed on video media, but computer graphics images usually uses the whole 8- or 10-bit range. Some video hardware can compress the range of computer graphics image data when transferring to video, and expand the range of video data when transferring to computer graphics images. This re-ranging is often combined with the luminance-chrominance transformation described in section 9.4. At the time of writing, there are no clear standards on when data is re-ranged and when it is not. If there is room for doubt, you should check the values of a test image in the video and computer graphics formats.

9.2 Voltage to Luminance conversion

The voltage V ranged 0 to 1 is related to the corresponding luminance L for each of the R'G'B' channels using the formula:

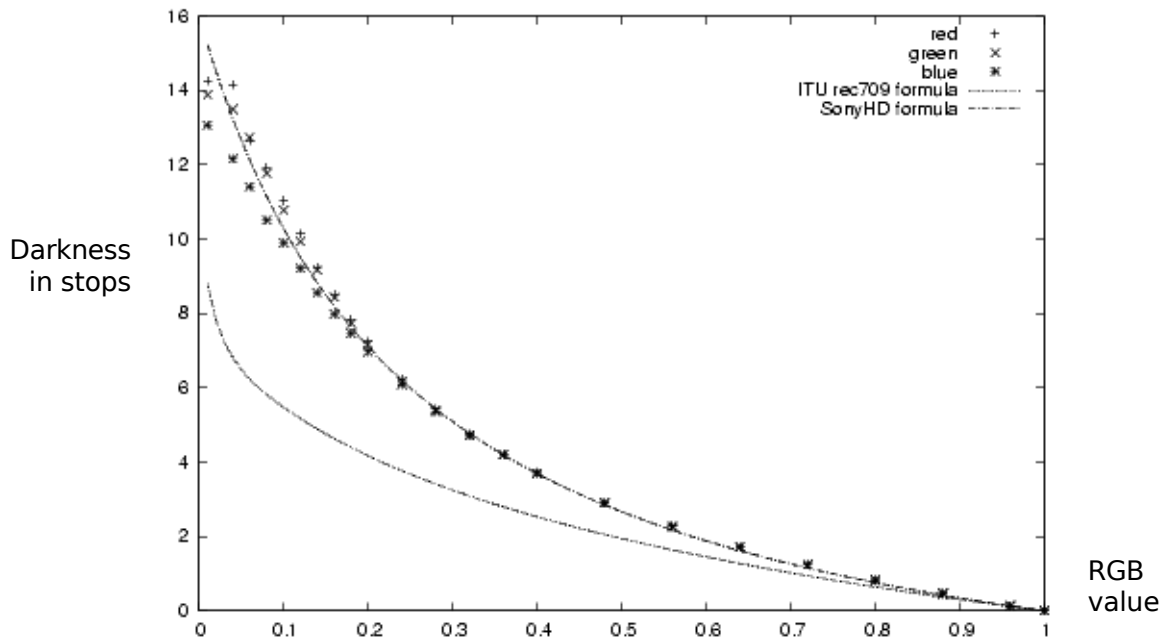
$$\begin{aligned} V &= 1.099 \cdot L^{0.45} - 0.099 & \text{if} & \quad 0.018 \leq L \leq 1 \\ V &= 4.5 \cdot L & \text{if} & \quad 0.018 \geq L \geq 0 \end{aligned}$$

This function is known as the 'opto-electronic transfer characteristic' or the 'gamma correction'.

In the standard, the voltages are called R'G'B', and the luminance values after the gamma correction are called RGB. Outside the standard, the luminance values are rarely referred to, and the voltages are often called 'RGB'.

This formula is simplistic. Over most of the range it approximates to the sort of power law function you get with CRTs, and at the dark end there is a linear section to keep the maths well behaved, a bit like the CIE function in section 5.1). Unfortunately, the linear section at the bottom end does not correspond with real monitor behaviour.

Real monitors are often set up with a PLUGE signal. A pluge signal has bars of -4%, 0%, and +4% grey. A digital 8-bit pluge image might have three bars with neutral values of 7, 16, and 25. The operator adjusts the monitor controls until the 7-16 edge cannot be seen, but the 16-25 edge just can. The 16 value is the video black value. If the monitor had the standard transfer characteristic, then this would set the luminance zero in the right place.



This plot shows experimental data measured off a Sony HD monitor calibrated in a darkened room using the PLUGE test image, the tone curve function taken from the ITU rec709 standard, and a fixed function we use in Truelight to represent Sony HD monitors...

$$L = 2^{\left(3.2 - \frac{4.0}{V + 0.25}\right)}$$

...where V is the RGB value, and L is the relative luminance, both ranged 0-1.

Our measured values are typical of HD monitors. A new HD monitor will give a smooth curve that gets steeper towards the shadows. An older monitor, such as this one, has a background glow that stops our shadow range going beyond about 14 stops.

We might expect a calibrated HD monitor would fit the ITU standard. In fact, we find the ITU standard gives a curve that is very different to the measured one. We could increase the contrast on the monitor, and stretch the ITU curve so it fitted the top half of the experimental data, but the shape at the shadows is still very different. You might not see the difference if you are viewing with ambient lighting, but you will see the difference in a darkened room.

9.3 RGB colours

The standard defines the CIE chromaticity coordinates for the RGB primaries, and the R=G=B white:

Colour	x	y
R	0.640	0.330
G	0.300	0.600
B	0.150	0.060
White	0.3127 (D65)	0.3290 (D65)

The white chromaticity sets the relative brightness of the R, G, and B signals. If we normalize our white tristimulus Y - not video Y or Y' - to 1.0, then XYZ and RGB are related by...

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} +3.240479 \cdot X & -1.537150 \cdot Y & -0.0498535 \cdot Z \\ -0.969256 \cdot X & +1.875992 \cdot Y & -0.0415560 \cdot Z \\ +0.055648 \cdot X & -0.204043 \cdot Y & +1.057311 \cdot Z \end{pmatrix}$$

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} +0.412455 \cdot R & +0.357580 \cdot G & +0.180423 \cdot B \\ +0.212671 \cdot R & +0.715160 \cdot G & +0.072169 \cdot B \\ +0.019334 \cdot R & +0.119193 \cdot G & +0.950227 \cdot B \end{pmatrix}$$

9.4 Luminance-Chrominance Coding

Within the standard, the luminance signal is called Y' and the chrominance signals are called C'_R and C'_B . Outside the standard, these signals are often just called Y , C_R and C_B .

There are two standards for luminance-chrominance coding in common use. The SD (standard definition) standard specified in SMPTE-125M is...

$$\begin{pmatrix} Y' \\ C'_R \\ C'_B \end{pmatrix} = \begin{pmatrix} +0.299 \cdot R' & +0.587 \cdot G' & +0.114 \cdot B' \\ +0.511 \cdot R' & -0.428 \cdot G' & -0.083 \cdot B' \\ -0.172 \cdot R' & -0.339 \cdot G' & +0.511 \cdot B' \end{pmatrix}$$

When the current HD (high definition) standard was defined, these values were replaced for most HD standards by another set defined in SMPTE-274M...

$$\begin{pmatrix} Y' \\ C'_R \\ C'_B \end{pmatrix} = \begin{pmatrix} +0.2126 \cdot R' & +0.7152 \cdot G' & +0.0722 \cdot B' \\ +0.5114 \cdot R' & -0.4646 \cdot G' & -0.0468 \cdot B' \\ -0.1172 \cdot R' & -0.3942 \cdot G' & +0.5114 \cdot B' \end{pmatrix}$$

There is a 1250 line HD standard that uses the old SMPTE-125M equations. Almost all the other HD standards, including all the popular ones, use the SMPTE-274M ones.

Note: both standards are calculating the luminance and chrominance from the $R'G'B'$ voltage signals, not from the gamma-corrected RGB luminance signals. This was an engineering compromise in the analogue days that made the separation into luminance and chrominance inaccurate in some parts of the colour space. The C'_R and C'_B values are zero for $R'=G'=B'$ neutrals, but for some colours the chroma will vary slightly with the Y' value.

The digital Y' values are ranged like the $R'G'B'$ values of section 9.1...

Limit	Voltage	Y'	8-bit value	10-bit value
Black point	0 mV	0	16	64
White point	700 mV	1	235	940

The chrominance signals C'_R and C'_B are signed, and have slightly different ranges for the two standards. The digital values are offset and ranged slightly differently...

Value	Voltage	C'_R, C'_B	8-bit value	10-bit value
Neutral	0 mV	0	128	512
Minimum	-350 mV	-0.5114	16	64
Maximum	+350 mV	+0.5114	240	960

9.5 Setting the Contrast

Section 9.2 described the classical way to set up a HD monitor using the PLUGE test chart. This was an ingenious scheme that allowed the user to set their display contrast so they always saw the same amounts of shadow detail without needing instruments to measure the very low light levels at the shadow end, and automatically compensating for the effects of flare and ambient light.

There were drawbacks. It was a skilled process. It was subjective. The settings depended on the ambient lighting. If the room was very dark, and the monitor was good, the shadow levels you set might be well beyond the sensible grading range.

We now have instruments that can measure these low light levels. We suggest you may use these grey and black values to set the contrast.

Patch	16-235 value	0-255 value
Grey point	33	20
White point	235	254
<u>Grey-Black</u> White-Black	0.0776	0.0784

Measure the luminance of the white and the grey patches. Adjust the display controls until you get a grey value that is 1/500 of the white value for each channel. You will need to periodically adjust the white point to keep it on target.

The SonyHD display calibration (see section 9.2) has a 500:1 contrast for RGB = 0.0787, which is pretty close to the values in the table. We do not have monitor controls to adjust the entire shape of the monitor, but if we can get a 500:1 contrast between the white and the grey, our monitor will probably be a close match to other HD monitors and to the SonyHD calibration.

9.6 Digital Cinema Interface

The DCI specification is a recent attempt to extend the gamut of digital images beyond the video gamut. The 12-bit DCI X'Y'Z' values are a simple transform of the CIE XYZ values...

$$X' = 4095 \cdot \left(\frac{X}{52.37 \text{ cd m}^{-2}} \right)^{\frac{1}{2.6}}, \quad Y' = 4095 \cdot \left(\frac{Y}{52.37 \text{ cd m}^{-2}} \right)^{\frac{1}{2.6}}, \quad Z' = 4095 \cdot \left(\frac{Z}{52.37 \text{ cd m}^{-2}} \right)^{\frac{1}{2.6}}$$

X'Y'Z' (usually called "XYZ prime") values can be stored in conventional video formats. X'Y'Z' images will look strange and desaturated when viewed on conventional video equipment, but should look correct when displayed on a DCI projector. For more details, see the DCI standard documents.

10 Real vision and simple models

Sections 2 to 5 described how the standard CIE colour spaces for normal colour vision were derived from some simple assumptions. If you followed this explanation you might believe that...

- We all see each wavelength of light as a unique colour.
- The colour we see does not depend on intensity.
- We see colour using red, green, and blue detectors with three fixed spectral responses.
- We can calculate the response of the eye to two or more wavelengths by summing the individual responses.
- Two colours with the same XYZ values will look the same.

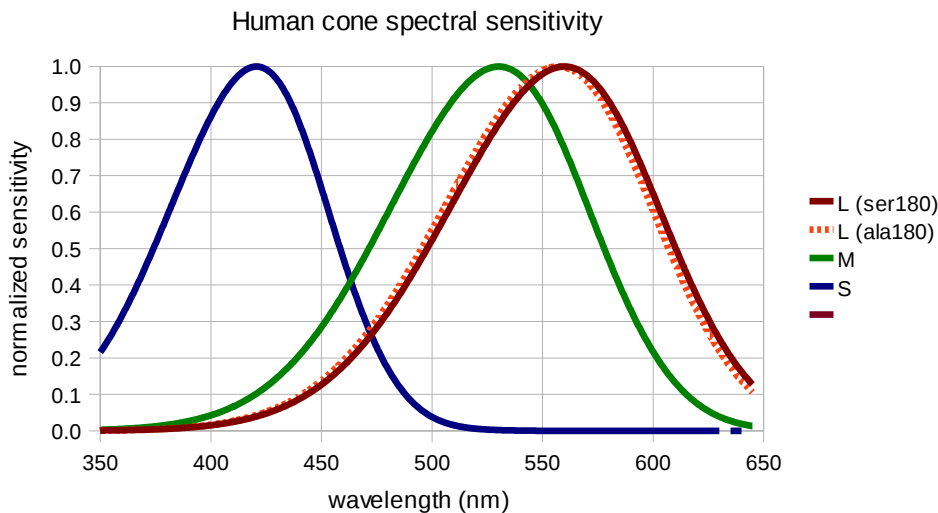
All these statements are wrong. Not completely wrong: they are close enough for the CIE values to be useful, but they are not close enough to predict a colour match under all conditions. We can make two patches with identical XYZ values that will look visibly different, or two patches that look the same but have different XYZ values. Section 10.8 describes a particularly pathological case.

All eyes are not the same. Two observers with good colour vision may still see different colours. You may find your left and right eyes are not quite the same.

This section describes where the errors come from, and why they exist. In most cases, there is not an easy cure.

There are some things we could do to get the CIE model closer to typical cinema conditions. There are other factors we cannot compensate for – factors that differ from one observer to another, or vary with retinal position, or are affected by the adaption of the eye and brain. Such factors are hard to model quantitatively. For example, Truelight contains a flare compensation that attempts to model the effect of a surround on the image. To accurately predict the effect of the surround, we must know the shape of the surround and its exact colour. The effect of the surround would be a lot greater at the edge of the image. It would also depend on where the eye was looking at the time. Even if we could model all these factors, it would be very hard to apply the corrections to a real image to make it look right for two separate observers.

10.1 Variations in cone pigments



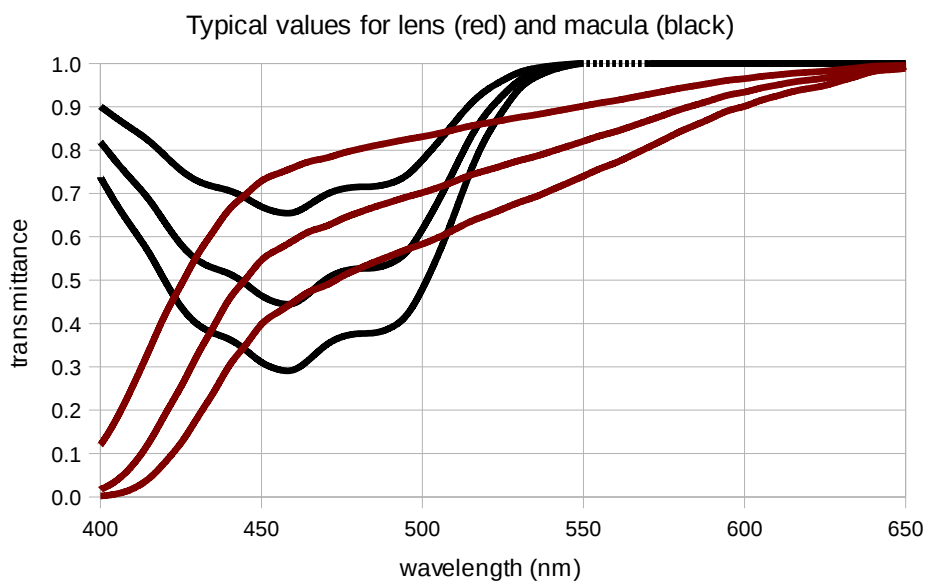
The eye has three types of cone sensor, usually called L, M, and S because their sensitivity peaks in the Long, Medium and Short wavelengths. The L-cone pigment is shown with two variants. Approximately 56% of tested subjects have serine and 44% have alanine at codon 180, which gives a shift of about 2.7nm in the peak sensitivity. There are probably other variants of the M- and L-cones, but this seems to be the most significant variation, and it is the one that has been studied most.

This suggests we need two sets of standard observers. In practice, the difference between these two red sensitivities is much harder to measure than these cleaned up graphs would suggest. The recent Stockman and Sharpe LMS primaries use a weighted average of the two measurements, and this approach is probably good enough for our purposes.

This genetic variation has been widely studied, but it is perhaps the least important variation in this section. It is usually masked by other variations within the eye.

10.2 Previsual variations

Before light reaches the cones on the retina, it has to go through the lens and the macular pigment. The lens contains a yellow pigment that increases in density with age. The macula is a blob of yellow dye that lies over the fovea – the area on the retina where we have our best colour and spatial resolution. The macular density does not vary significantly with age, but the density and distribution of this pigment can vary considerably between observers, and sometimes between eyes in a single observer. These two variations in yellow pigmentation give us most of the variability between observers. The following graph shows spectral transmittance plots for typical observers with high, average, and low pigmentation of either type.



It is hard to measure either of these densities in a living eye. In the past, when observers have two disagreeing sets of colour matching data, they have often assumed different average values of these two pigments to resolve the differences. While this may be part of the explanation, it also can hide other phenomena.

10.3 Photopigment masking

Our rod cells contain a purple pigment called rhodopsin. After we compensate for absorption of the pigments described in section 10.2, the spectral sensitivity of cones is close to the fraction of the light absorbed by the rhodopsin. The early experimenters deduced that the bleaching of the dye was somehow generating the signal that fired the cell. The dye is regenerated steadily, and the balance between the rate of bleaching and the rate of regeneration seemed to allow the cell to be sensitive at low light levels when the dye densities were highest, and less sensitive at high levels when most of the dye is bleached. They could not at first find a corresponding dye in the cones, but they guessed the mechanism was similar.

We now know a lot more about the actual mechanism within the rods and the cones, and we know the early researchers were basically right. There is a dye in rods and cones; light bleaches the dye, and this bleaching can indirectly fire the cell. Rod cells individually adapt to light level, but cone cells pool their signals for adaption. The adaption process controls the photopigment density, but it is not a simple balance between production and bleaching.

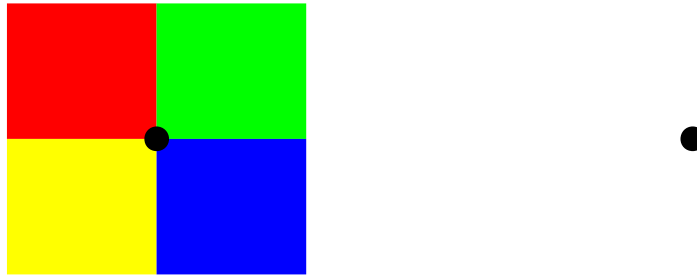
In section 10.2, we described the corrections necessary for the yellow pigments in the lens and the macula. Now we find the detector itself in a pigment. The rod and cone cells are also deep structures, so it is likely the pigment at the front of the cell will be shadowing the back. This shadowing will be greatest at the peak wavelength, which will increase the relative sensitivities away from the peak.

It is possible to view a living retina through stabilized optics, and measure the light transmitted obliquely by an individual cell. When the eye is adapted to bright conditions, the dye is assumed bleached. The difference between the bright adapted transmittance and the dark-adapted transmittance gives an estimate of the dye density. Such measurements give cone dye densities in a living retina of about 0.5; which are backed up by measurements on individual cells in dissected retinas. This suggests that the pigment at the front of a cell may have a considerable effect on the signal at the back. The CIE measurements used high light levels to keep the densities down, and minimise this effect. They may not be appropriate for typical cinema and display light levels. Some researchers believe this effect gives shifts comparable to the errors in estimating the eye's yellow pigments described in section 10.2.

There are problems with measuring individual cells using oblique light. We might guess the cells would be more sensitive to light coming at an angle, and the front no longer shadows the back. In fact, the reverse is true. This suggests the cells somehow act as a waveguide, channelling the normal radiation into the detector, and rejecting some of the transverse component. We do not know enough about this waveguide effect to model it well.

10.4 Colour opponent coding

If you stare at the left-hand dot for 30 seconds, and then switch to the right-hand dot, you should see some afterimage colours.

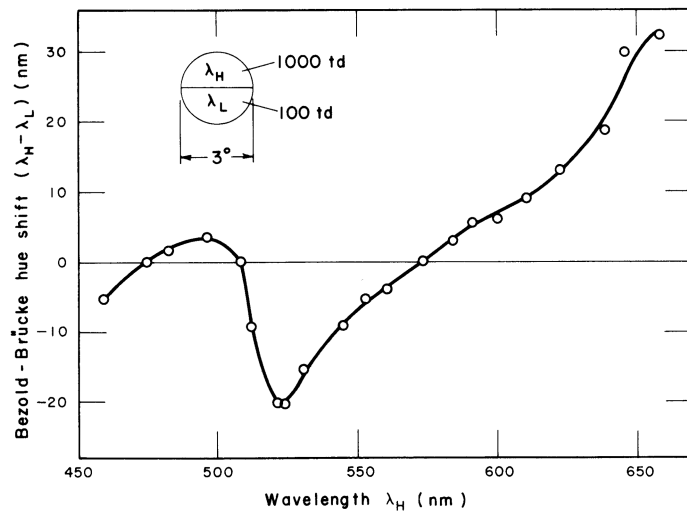


If the afterimage colours were due to adaption of the individual cones, then we might expect red to give cyan, and green to give magenta. Instead, it seems that red gives green, and green gives red.

Behind the retina, there are cells that generate the colour signals. There are red-green opponent cells that are excited by signals from red cones and inhibited by signals from green cones, or vice-versa. There are yellow-blue opponent cells. There are also luminance cells that are stimulated by red, green and blue, that should give the standard photopic observer response graphed in section 1.1.

The red and green signals arrive as pulses. The excitation and the inhibition process have time constants. So, it is likely that a red-green opponent cell output will not depend just on the ratio of red to green light, but will also be influenced by the intensity and any flicker effects.

The graph shows the effect of varying the retinal intensity on the perceived colour of a single wavelength. These shifts are believed to come from the differing response rates of the red-green and blue-yellow opponent cells.



Flicker can also give coloured effects; look up Benham's top if you want to know more. The effect is thought to be small unless the flicker is slow enough for us to be aware of it. The flicker of typical film projectors and displays is unlikely to affect the colours we see.

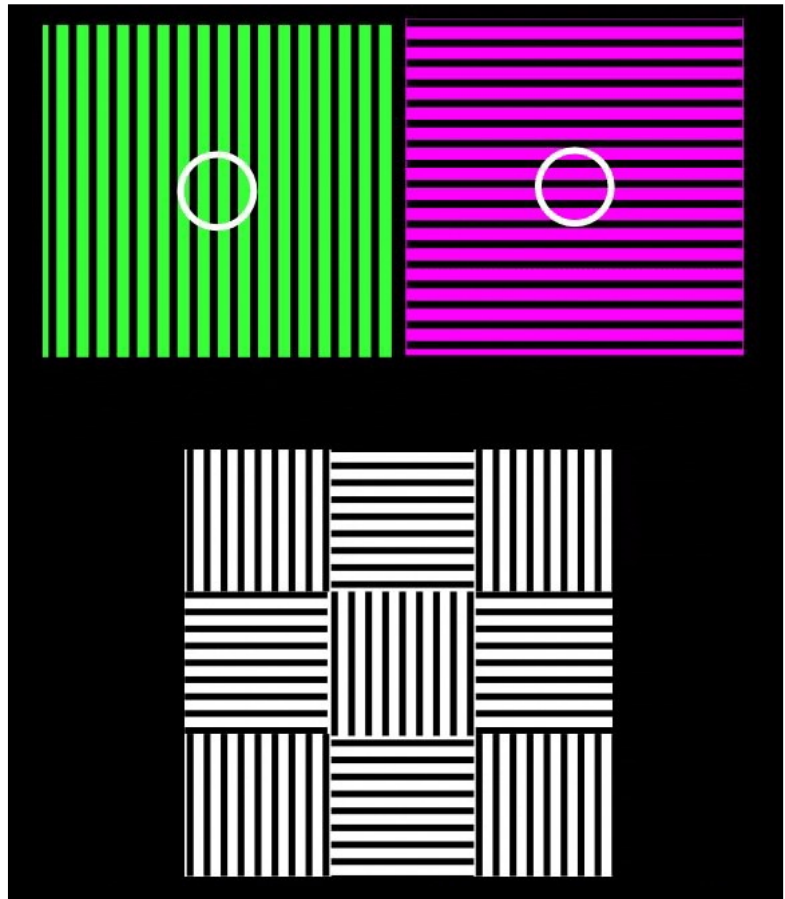
The McCullough effect is another effect that has been linked with the opponent cells. This is a spatial effect, but it is of interest to us because of its surprisingly long lifetime.

Concentrate on the two coloured patches, tracing out the circles slowly with your eyes. Switch from one patch to the other every 15 seconds or so. After about five minutes, look at the neutral patch below. You should see the vertical stripes are tinted pink, and the horizontal stripes are tinted green.

The effect can last for several days. People have even reported seeing the colours two weeks later. It seems likely that the process happens in the cells immediately behind the retina. Drugs that affect inhibition in the visual system modify this effect.

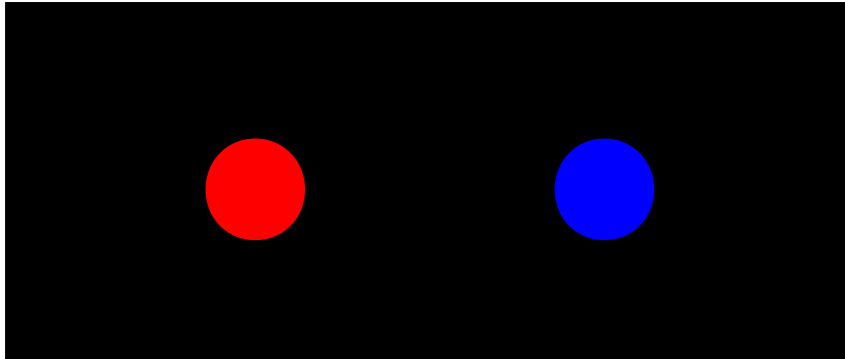
There is a similar long-term effect where the perceived balance of a yellow colour is affected by prolonged exposure (4 hours/day) to red or green light. This, too, has been shown to last for several weeks.

These effects are unlikely to have a severe effect under practical circumstances, but they could make people grade films differently in winter and summer. A change in diet can also affect the density of the macular dye over a similar period.



10.5 The influence of the rods

The rod signal is normally assumed to play no part in colour vision. However, we do know as you go to low light levels, the rods play an increasing part in the perceived luminance signal. The graph in section 1.1 showed that the rod sensitivity peaks at a shorter wavelength than the cones. We would expect a stimulus of 450nm blue to appear ten times as bright to the rods as to the CIE standard cone observer.



Under normal lighting, the red spot will look brighter than the blue one. If you look at a print of this by the light of the full moon, the blue spot should look brighter. This is not because the light from the moon contains more blue (it does not) but because our luminance signal is now dominated by the rods instead of the cones.

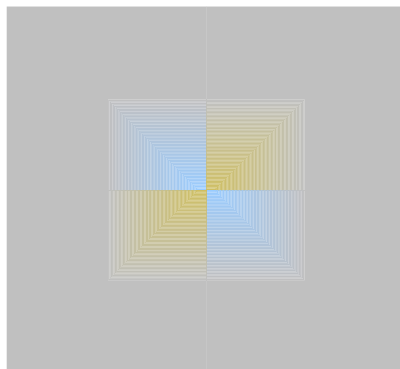
The experience of people with rod-only vision tells us the rods can generate a perfectly useable signal beyond 1000 cd/m^2 , even though they appear totally bleached. They may be contributing to the colours we see on typical displays.

10.6 The influence of melanopsin ganglion cells

The eye has a third type of detector in addition to the rods and cones. As well as rods and cones, there is a rare subtype called intrinsically photosensitive Retinal Ganglion Cells (ipRGC). These cells produce melanopsin with an absorption peak of about 480 nm. These cells have been shown to be sensitive to light, and have a light- and dark-adaptive mechanism similar to the rods and cones. It is unlikely that these cells contribute anything in terms of resolved image detail because they are so rare, but they do cause our brains and bodies to react to daylight, so they could affect how we react to colour stimuli.

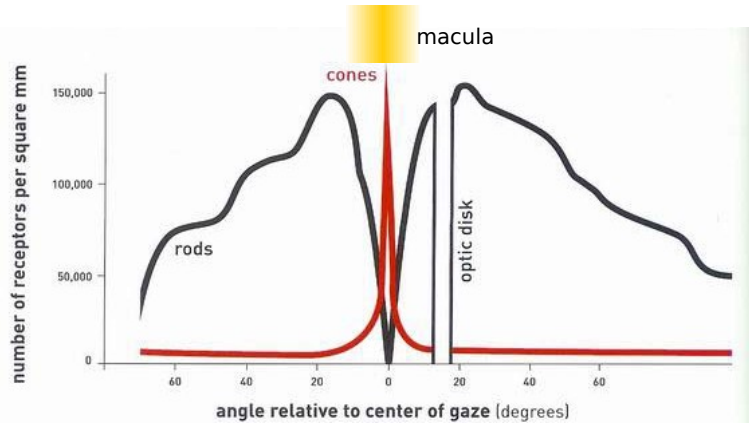
10.7 The effect of polarization

We can detect polarization with the naked eye. If we look at a white region of a LCD display, and tilt our head, we may briefly see a figure like the diagram opposite. This is known as Haidinger's brushes because the yellow regions resembled the head of an artist's brush. It is caused by polarized light from the LCD display interacting with the anisotropic macular pigment. It is not usually a serious problem when colour matching provided you recognize it and avoid it.



10.8 Variation with retinal position

The human eye does not contain a uniform distribution of identical RGB detectors like a digital camera. The very centre 0.25 degrees of vision - the part that we use for most precision vision - contains no blue cones, and no rods. The yellow macular dye of section 10.2 extends to about 3 degrees. As we go away from the fovea, rods replace the cones.



If we look at a plain, white screen, we

might expect our central vision and our peripheral vision to see different colours. Maybe, at the retinal level they do, but our eye-brain system guesses that the central and the peripheral fields are supposed to match, and corrects the signal. You can override this correction by suddenly changing both stimuli. If you look at a dark screen that suddenly turns blue, you will briefly see a dark spot, called the Maxwell spot. If you go on looking, it fades, as the eye and brain compensates.

Suppose you have a cinema screen. On one side you project a white from a digital projector. On the other side you project light from a xenon lamp. If you match the XYZ values of the projector white to the xenon lamp with an instrument, you will probably find the patches do not match to the eye. Typically, the xenon lamp will look duller and yellowish-green and the digital projector will look bluer and cleaner.

Let us match the overall colour of the patches by eye. The patches are much larger than our central vision region, so we are probably doing much of the matching using the peripheral vision. Now, if we now move our eyes around on this blank 'butterfly test' we can see the Maxwell spot. On the xenon field it looks brighter and yellowish-green; on the digital projector side it looks darker and violet. On the boundary, it looks half one, half the other. The eye-brain tries to hide the contrast between the centre and the periphery, but it the colour in the centre is changing depending on the patch we look at. Take away one of the patches and the spot disappears.

If we have two media with different spectra, we can match the colour detail within the Maxwell spot, or we can match the broad impression of colour as seen by the peripheral vision, but we cannot do both at once. This is why we recommend the Truelight white point is set by eye using real image data, rather than using instruments and patches of colour.

10.9 The special case of the white point

The Maxwell spot effect we described in section 10.8 would give us an unpleasant effect if we were doing a 'butterfly' test in a cinema. However, if we were looking at a single image, we would not see the Maxwell spot. We know we cannot simultaneously match central and the peripheral vision. As most cinema images do not have large regions of flat tint in the regions we are likely to be looking, we would normally chose to match the colours at the centre, and ignore the periphery.

The eye is able to extract a sense of the white point, even from images with no white. We have found that the film white - surely the easiest colour to reproduce because it is just the clear film base - shows the same systematic shift in butterfly tests. Though the rest of the image colours may seem like a good match, the whites will look brighter and 'cleaner' (less yellowish-green) on a digital projector. I suspect this is because the eye biases its guess of the white point towards any conspicuously bright bits of the image, such as sunlit clouds, and this in turn gives the bias towards the peripheral signal we might only expect from much larger regions of colour.

This colour shift is not something we can easily correct for. If we edit the digital projector whites to be duller and more yellowish-green, then we may improve the match in some images, but we make it worse in others. And, like our two-coloured Maxwell spot, the colour shift vanishes when we take away one side of the butterfly test. Both images will look the same when seen separately. When seen together, we can match large colour in patches and whites, or colour in small patches. But not both at once.

10.10 Conclusion

Matching colours on different media is a lot harder than you might suppose from the simple theory at the beginning of this note. Indeed, a perfect match is probably impossible. There is no 'right' or 'wrong' here: displays with different spectral characteristics will look different, and any comparison of two images changes the appearance of the images. We should always strive for a better match, but we must also accept that there will be limits to how closely we can match two different media using a colour transform.

We could make colour spaces like CIE XYZ but better matched to typical display conditions. This would be a major task. I am not sure that the benefits would outweigh the disadvantage of being incompatible with the rest of the world. We know we can fix most systematic errors in CIE XYZ by shifting the white point. But, changing the colour space cannot solve the problems of sections 10.8 and 10.9.

We need a robust digital standard to describe image appearance, so we can get repeatable results on different media. Colour science cannot give us a robust predictive model for colour appearance today. Perhaps the best we can do now is to grade our images in some colour space derived from a real display RGB. We can still store the images as XYZ, but this would be a particular XYZ that looked right when viewed on particular display. We should record the primary spectra used when grading if we want future viewers to know the exact look that was agreed at the time.