An Incomplete and Short Introduction to a Few Usually Forgotten Foundamentals of Computer Vision

(a.k.a. the eazyFUFF of Vision)

Julio Martín-Herrero*

1 Elements of a vision system

Computer vision is a complex subject involving many scientific disciplines and technological fields. This is a direct consequence of the different elements involved in a vision system:

Reality

Vision is the sense by which some complex systems can obtain information about targets without any physical contact. Targets exist in the physical world, where they configure *scenes*. Only targets capable of interaction with the electromagnetic field can be sensed by vision. Therefore, ghosts and other celestial nonbeings are excluded. Everything else is a potential target for a vision system. Electromagnetic radiation (the phenomenon allowing transfer of energy between different locations by altering the electromagnetic field) and wave-matter interaction are the physical-world mechanisms that make vision possible, portraying scenes into images.

Radiation source

Only radiant objects emit radiation. Non radiant objects have to be illuminated by one or several sources of radiation. Some systems are based on natural sources which are not controlled by the designer or by the user (eg. the Sun). Some systems are exclusively based on artificial sources (eg.

---

*Dept. of Signal Theory and Communications, University of Vigo (Spain).

a lamp inside a dark chamber). Many systems are based on one or several artificial sources, but are also subjected to one or several additional artificial and natural sources (interfering sources). Hopefully, in artificial systems the designed sources are under the control of the user or the engineer. Usually neither the artificial nor the natural interfering sources are controllable by the system. Sometimes even the designed sources are not under the control of the engineer.

Camera

The camera collects the radiation received from the target and keeps the spatial relationships in the target, i.e. there should be a coherent mapping between the different emitting/reflecting areas in the target and the image plane in the camera. The image plane is a plane in the camera where *images* are formed. An image is an spatially coherent representation of the level of incoming energy within a given bandwidth coming from the different regions in a scene. Cameras can be as simple as a black box with a tiny hole, or as complex as an X-ray tomograph. Lenses usually have a few things to say in this issue.

Sensor

The sensor produces an electrical signal suitable for processing which is proportional to the incoming radiation (received radiative flux density). The image plane can be covered by a 2D array of sensors (matrix array, or matrix sensor), by a 1D array, if a 1D scanning system is added to the system (linear array, or linear sensor), or by a single sensor, if a 2D scanning system is added to the system. Electronics is one of the main keywords here.

Processor

The processor digests the signal fed by the sensor, generates higher level descriptions of the image, stores images (either in raw format, as higher-level descriptions, or a combination of both) in a special place called *memory*, by a process usually refered as *memorization*, establishes relationships between different features at different levels of description and from different sources, including the memory and other sensing devices, and produces outputs for acting devices in order to alter the reality causing the scene.

Actuators

These are the acting devices which receive orders from the processor. Sometimes they are considered a part of the vision system, as in a motion tracking system, or in a navigation system, and sometimes not, as the legs and arms in a tennis player.

Yes, a tennis player. The elements enumerated above are common to artificial and biological vision systems. An eye is nothing but a camera, a retina is "just" a sensor, a brain (and attached gadgets) is quite a good distributed parallel processor, and reality... reality is the place some struggle to escape and where some others would like to act.

The possibility of establishing a direct relationship between artificial and biological systems allows us to learn from living organisms in order to develop better engineering solutions, and allows cognitive scientists and neuroscientists to enhance their understanding of the human vision system and the human mind. However, care has to be taken in order to not be misled either by biomimic engineering solutions far less than optimal, nor by pretended models of living systems which may just be oversimplifications of very complex phenomena, structures, and processes. Don't worry, ghosts are still out of scope.

Knowing how we see and how we process visual information is a must for any researcher involved in image, no matter if it is computer vision, computer graphics, or display systems design. You should get a solid background on the human vision system, and it is always good having a reference on the way that evolution has solved (in the extent that we may already know it) a given task involved in a given engineering problem.

2 Light and radiation

Light is the electromagnetic radiation within a narrow portion of the spectrum (380 to 780 nm). Every portion of the electromagnetic spectrum obeys the same fundamental physical laws. Radiation of different wavelenghts, however, has different interactions with matter. We have a special name for the visible part of the spectrum because it is the band our visual system has evolved to be sensitive to.

However, computer vision does not depend on visible radiation. There are imaging detectors for other frequencies of electromagnetic radiation (from X rays to radiowaves) and even other types of radiation (sound, for instance). Remote knowledge about objects is obtained not only from the spatial distribution of emitted and reflected radiation, but also from the portion of radiation that is emitted, scattered, absorbed, or reflected. Thus, penetration depth and surface reflectivity, the interaction of waves and matter, are esential processes. All of these processes depend on the wavelength, and thus the spectral distribution of the detected radiation is also a useful source of information.

Electromagnetic radiation

Electromagnetic radiation is a perturbation of the electromagnetic (EM) field (which is just a property of space) in the shape of waves propagating energy through the space. These waves are characterized by their frequency and wavelength, inversely related by the speed of light c, $\lambda = c/f$. The speed of light depends on the wavelength and the medium. Non-monochromatic light is composed of different wavelenghts. This is the reason for *refraction* (change of direction at interfaces between different media) and *dispersion* (separation of different wavelenghts).

EM radiation can be treated as a flow of photons travelling with speed c. Photons are discrete particles with quantized energy $e_{\mathrm{p}} = \mathrm{h}f$, where $\mathrm{h} = 6.626 \cdot 10^{-34}$ J·s is Planck's constant. Thus, the total energy is given by the number of photons. Photon energy is usually expressed in electron volts (eV): the energy of an electron accelerated by a potential difference of one volt. Photons do not have electric charge, but they are usually detected by their interaction with charged particles in sensors. For a given sensor material, photons with energy above the bandgap energy of the sensor's semiconductor can move electrons from the valence band to the conduction band. This sets a fundamental limit to the sensitivity range of the different materials. Thus the bandgap energy of InSb is as low as 0.18 eV, and therefore it is sensitive to wavelenghts up to 6.9 μm, while Si has a bandgap of 1.1 eV, and thus only reaches 1.1 μm. Therefore, InSb is suitable for infrarred cameras in the 3-5 μm range, while Si is suitable only for VIS and VNIR applications. Note that Si exceeds the visible range (goes up to 1 100 nm), and an IR filter can be necessary in some applications. The probability of a photon at a given wavelength moving an electron accross the bandgap is the *quantum efficiency* of the sensor at that wavelength, an importat factor describing the sensitivity of the sensor, together with the *fill factor*, the percentage of the area of the sensor effectively sensitive to the impinging photons.

Polarization

An EM wave can also be described as oscillating coupled electric and magnetic fields. Both are given by the solution of a set of linear differencial equations, Maxwell's equations. Linearity implies superposition, and therefore the sum of two solutions to Maxwell's equations is also a solution. In general, the orientation of the electric field vector $\vec{E}$ changes with time. A mixture of waves shows in general random distribution of the orientation of $\vec{E}$. In the special case that the orientation of $\vec{E}$ is confined to a plane, the wave is said to be *linearly polarized*. When two linearly polarized waves travel in the same direction, the resulting $\vec{E}$ vector is the sum of the individual $\vec{E}$ vectors. Depending on their phase shift, the resulting $\vec{E}$ may remain confined to a plane (linearly polarized, 0^o phase) or rotate around the propagation direction. In this case, if the phase is 90^o the wave is said to be *circularly polarized*, otherwise it is called *elliptical polarization*.

Coherence

Mixtures of waves emitted from conventional sources do not show any spatial or temporal correlation. Such radiation is called *incoherent*. Special types of sources, such as lasers, emit waves with a given relationship between the phases of their $\vec{E}$ vectors. This is *coherent* radiation, which can be added in constructive and destructive interference, and thus can reach very high amplitudes and cause damage.

Radiometry

Radiometry quantifies radiated energy, and radiometric quantities are the basis for any light- or radiation-based imaging system. At the root of any output of such a system it is always a radiometric quantity.

Radiant energy and flux

Radiant energy Q [J] quantifies the total energy emitted by a source or received by a detector. Radiant flux Φ [W] is radiant energy passing through a surface per unit time, $\Phi = dQ/dt$, also called radiant power.

Photon derived quantities are based in number of photons instead of Joules, and are identified by the subscript p. The number of photons given the radiant energy is

$$N_{\mathrm{p}} = \frac{Q}{e_{\mathrm{p}}} = \frac{Q\,\lambda}{\mathrm{h\,c}}. \tag{1}$$

Photon flux Φ_{p} is the number of photons per unit time,

$$\Phi_{\mathrm{p}} = \frac{dN_{\mathrm{p}}}{dt} = \frac{\Phi\,\lambda}{\mathrm{h\,c}}. \tag{2}$$

All other photon derived quantities can be obtained from the corresponding energy counterparts by dividing by the energy of a photon, e_{p}.

In general, radiative flux is made up of waves at different frequencies. Correspondingly, all radiometric quantities have spectral distributions over a given band. Let Q [$\circ$] be any radiometric quantity. Then Q_λ is the corresponding spectral quantity concentrated at a specific wavelength within an infinitesimal interval $d\lambda$,

$$Q_\lambda = \frac{dQ}{d\lambda} = \lim_{\Delta\lambda\to 0} \frac{\Delta Q}{\Delta\lambda}. \tag{3}$$

where the units of Q_λ are [$\circ$/m]. Sometimes it is writen in [$\circ/\mu$m] or [$\circ$/nm]. Integrated quantities over a given spectral range $[\lambda_1, \lambda_2]$ are obtained as

$$Q_{\lambda_1}^{\lambda_2} = \int_{\lambda_1}^{\lambda_2} Q_\lambda \, d\lambda. \tag{4}$$

Radiant exitance and irradiance

Radiant exitance M $[\mathrm{W}\cdot\mathrm{m}^{-2}]$ is radiative flux emitted per surface area, $M(\vec{x}) = d\Phi/dS$. Exitance is in general a function of position on the emitting surface, unless Φ is uniformly distributed over the area S, when $M = \Phi/S$.

Irradiance E $[\mathrm{W}\cdot\mathrm{m}^{-2}]$ is radiative flux incident per surface area, $E(\vec{x}) = d\Phi/dS$. Exitance characterizes an active radiating surface, while irradiance characterizes a passive receiver surface.

Radiant intensity

Radiant intensity I $[\mathrm{W}\cdot\mathrm{sr}^{-1}]$ describes the angular distribution of radiation emerging from a point, as radiant flux per solid angle, $I = d\Phi/d\Omega$. It is a function of the direction of radiation, usually specified by spherical coordinates θ and ϕ. It is generally used to specify radiation emitted from point sources, much smaller than the distance to the detector. It is often confused with irradiance and illuminance, but it should not.

Radiance

Radiance L $[\mathrm{W}\cdot\mathrm{m}^{-2}\cdot\mathrm{sr}^{-1}]$ is the amount of radiant flux per solid angle per projected surface area of the emitting source,

$$L(\vec{x},\theta,\phi) = \frac{d^2\Phi}{d\Omega}dS_\perp = \frac{d^2\Phi}{d\Omega}dS\,\cos\theta, \tag{5}$$

where $dS_\perp = dS\,\cos\theta$ defines a surface element perpendicular to the direction of the radiated beam.

Radiance combines exitance and intensity, relating intensity in a certain direction to the area of the emitting surface. It is used to characterize an extended source, with an area comparable to the square of the viewing distance, and is a function of both position on the emitting surface and direction. All other radiometric quantities can be derived from the radiance, by integrating over solid angles or surface areas.

Inverse Square Law

If $I = d\Phi/d\Omega$, then $d\Phi = I\cdot d\Omega$, and the irradiance of a surface can be written as $E = d\Phi/dS = I\cdot d\Omega/dS$. For small surface elements perpendicular to the line from the point source to the surface, at a distance r from the source,

$d\Omega = dS/r^2$. Thus the irradiance at a distance r from a point source with intensity I is $E = I/r^2$.

Lambert's Cosine Law

Radiant intensity from extended sources is usually unevenly distributed with the angle. A surface is called Lambertian if its radiance is independent of the viewing angle, i.e. $L(\vec{x}, \theta, \phi) = L(\vec{x})$. Then the angular distribution of intensity is

$$I(\theta) = \cos\theta \int_S L(\vec{x}) dS = I_0 \cos\theta, \tag{6}$$

which is independent of ϕ, and varies with the cosine of the angle of incidence θ.

The exitance of a planar Lambertian surface is

$$M(\vec{x}) = L(\vec{x}) \int_0^{2\pi} \int_0^{\pi/2} \cos\theta \, \sin\theta \; d\theta \; d\phi = \pi \, L(\vec{x}). \tag{7}$$

The proportionality factor π shows that it yields half the exitance expected for a surface radiating into 2π stereoradians. For a uniformly radiating point source, the proportionality factor would be 2π. Non-Lambertian surfaces have proportionality constants smaller than π.

The consequence of Lambert's Law is that Lambertian surfaces appear to have the same brightness from all angles. This is not in contradiction with the cosine dependence of intensity, as can be seen by investigating the radiant power transfer from an extended source to a detector with finite area. Lambert's Law is only valid for perfect radiators and diffusers. In spite of being frequently used for practical purposes, it is not valid for real radiators in general. However, at small angles of incidence, Lambert's Law holds for most surfaces.

Photometry

Photometry studies the response of the human eye in terms of radiometric quantities. Therefore, it deals only with the visible part of the spectrum. Light is concentrated in the retina by the preretinal optics of the eye and perceived by two different types of sensors: rods and cones. Cones are used at high irradiances to detect light and perceive colour (*photopic vision*), while rods are mainly used for low illumination levels (*scotopic vision*). Mesopic vision happens at intermediate levels of irradiance, when both rods and cones are involved. The scotopic and photopic spectral luminous efficiency functions $V(\lambda)$ and $V'(\lambda)$ describe the eye response as a function of the wavelength. Their shape is similar, with a slight downshift of the scotopic curve to 507 nm compared to the 555 nm peak of the photopic curve. The relative curves are normalized to peak value unity, whereas the absolute curves, which are used to convert radiometric quantities (independent of the eye response) to photometric quantities (measuring the eye response), peak at $P = 1754 \text{ lm} \cdot \text{W}^{-1}$ and $P' = 683 \text{ lm} \cdot \text{W}^{-1}$. Any energy-derived radiometric quantity Q_e can be converted to photometric by

$$Q_\nu = P \int_{380\,\text{nm}}^{780\,\text{nm}} Q_\text{e} V(\lambda) d\lambda, \tag{8}$$

with primes for scotopic vision. Evidently the spectral distribution of the radiometric counterpart is needed to obtain any photometric quantity.

Luminous energy and flux

Luminous energy [lm] is the fraction of radiant energy that causes a visual sensation in the eye. Luminous flux [lm · s] is the total luminous energy emitted or received per unit time, i.e. power.

Luminous exitance and illuminance

Luminous exitance and illuminance are the counterparts of radiative exitance and irradiance, measured in $\text{lm} \cdot \text{m}^{-2}$, also called lux, the flux per unit surface area leaving or illuminating a surface. They are integrated over the angular distribution of light.

Luminous intensity

Luminous intensity is the total flux emitted in a solid angle in a given direction. It is used for point sources and light rays, and measured in $\text{lm} \cdot \text{sr}^{-1}$ or candelas [cd]. One candela is the luminous intensity of a monochromatic source at 555 nm with radiant intensity $1/683 \text{ W} \cdot \text{sr}^{-1}$.

Luminance

Luminance is the counterpart of radiance. It describes the *perception* of total brightness as luminous flux per solid angle per surface area perpendicular to a given direction, and is measured in $\mathrm{cd}\cdot\mathrm{m}^{-2}$.

Luminous efficacy

Efficiency (η) measures the ratio of output power to input power in a given system. Luminous *efficacy* measures the ratio of luminous energy output (radiation able to stimulate the human visual system) with respect to power comsuption. The ratio of a photometric quantity to its radiometric counterpart integrated over the entire spectrum is the radiation luminous efficacy, K_r. The efficacy of a light source K_s is the ratio of luminous flux to total power supplied to the source,

$$K_\mathrm{s} = \frac{\Phi_\nu}{P_\mathrm{e}} = \frac{\Phi_\nu}{\Phi_\mathrm{e}}\frac{\Phi_\mathrm{e}}{P_\mathrm{e}} = K_\mathrm{r}\cdot\eta_\mathrm{e}\ [\mathrm{lm}\cdot\mathrm{W}^{-1}]. \tag{9}$$

The efficiency of a source is always less than 1, and therefore the efficay of a light source is always smaller than the radiation efficacy.

Thermal emission and blackbodies

Hot (above absolut zero) matter emits radiation. Thermal radiant exitance increases with temperature. A blackbody is an ideal body that absorbs all incident energy, i.e. a perfect absorber. Perfect absorbers are also perfect emitters. Planck (1900) derived the relationship between the spectral distribution of thermal radiation (spectral radiance, $L_{\mathrm{e},\lambda}$) of a blackbody and its temperature. Though for a perfect emitter, Planck's Law has been shown to describe with enough accuracy thermal radiation of a broad range of objects, from the Sun to water through incandescent lamps.

Wien's Displacement Law gives the wavelength of maximum radiative exitance of a blackbody (by deriving Planck's Law with respect to λ),

$$\lambda_{\mathrm{m,e}}T = 2.891\cdot 10^{-3}\ \mathrm{m}\cdot\mathrm{K}. \tag{10}$$

Exactly 25% of the total exitance is emitted below $\lambda_{\mathrm{m,e}}$. Thus, for instance, typical incandescent lamps with $T = 1\,000$ K have $\lambda_{\mathrm{m,e}} \approx 3\,000$ nm, and therefore they only emit a very low fraction in the visible range; most of the radiation is emitted in the IR. Luminous efficacy of blackbodies peaks at 6 600 K, very close to the temperature of the surface of the Sun, an evidence of adaptation of the human visual system to the solar spectrum.

Wave-matter interaction

Radiation interacts with matter. In general, bodies reflect some of the incident energy, absorb another fraction, transmit yet another fraction, and emit energy in response to incoming energy to maintain the thermodynamic balance. All these processes affect the direction of propagation, the magnitude, the spectral distribution, and the polarization of the incoming waves.

Optical properties of surfaces and objects

The mechanisms governing the interaction of radiation with matter are very complex, but they can be studied using macroscopic properties of objects and surfaces: Reflectivity, also called reflectance, ρ: the ratio of incoming to outcoming radiative flux; Absorptivity, or absorptance, α: the ratio of absorbed to incident radiative flux; Transmissivity, or trasmittance, τ: the ratio of trasmitted to incident radiative flux; and emissivity, or emittance, ϵ: which compares the exitance of a given object to that of a blackbody at a given temperature. Use the suffix -ivity for surface related properties and -ance for volume related properties.

All these quantities have spectral and directional dependencies, i.e. they vary with wavelength and angle. Spectral selective quantities can be obtained by integrating the corresponding optical property over the entire spectrum weighted by a spectral function. Examples of spectral functions used to obtain spectral selective quantities are the scotopic and photopic efficiency functions (to obtain photopic and scotopic luminous transmittance and reflectance), the solar irradiance (for solar transmittance, reflectance, absorptance), and the blackbody irradiance (for emittance).

Refraction index

Maxwell's equations yield the complex index of refraction, $N(\lambda) = n(\lambda) + \mathrm{i}k(\lambda)$. The real part is the index of refraction of geometric optics, while the complex part governs reflection and absorption. Refraction determines the change in the direction of propagation of radiation crossing the interface between two media with different dielectric properties, $n_1 \cdot \sin\theta_1 = n_2 \cdot \sin\theta_2$. This is the scientific basis for the technology of lenses and prisms. Prisms take advantage of the dependency of n with λ to separate radiation of different wavelengths, while lenses generally suffer it in the shape of *chromatic* *aberration.*

Luminiscence

Luminiscence is the emission of light by radiative transition from a excited state to a lower energy state. *Fluorescence* is luminiscence that persists only a few ns after the excitation, i.e. it ceases almost at the same time as the stimulation. *Phosphorescence* is luminiscence that happens between ms and minutes after the excitation, i.e. delayed luminiscence. The sources of estimulation can be light (photoluminiscence), electric charge (electroluminiscence), thermal stimulation (thermoluminiscence), ionizing radiation (radioluminiscence), or chemical reactions (chemoluminiscence, also called bioluminiscence in living organisms).

Coming next

Tired of physics? Well, you will need a little patience yet. You should still deal with optics, the physics of colour, and image sensors. Noman's land between science and technology still for a while. But don't worry, balance on a ridge is unstable, and ultimately you will fall in the depths of the technological well. How did they use to call it? Silicon Valley? Whatever. Chances are rather somewhere in Southeast Asia.

3 The perception of light

Brightness, according to the CIE (Commission Internationale de L'Eclairage), is "the attribute of a visual sensation according to which an area appears to emit more or less light". This is a perceptual quantity which cannot be measured in any objective way. We can only measure physical quantities such as radiant power (radiance), and weight them by the spectral sensitivity functions commented above, to obtain the photometric counterparts, for instance *luminance*, denoted Y.

In order to perform this spectral weighting, we need the spectral distribution of the radiance, also known as the *spectral power distribution* (SPD), of a given light. Luminance is therefore proportional to radiant power, but the same amount of power in different parts of the spectrum will produce different levels of luminance. Luminance is measured in $\mathrm{cd} \cdot \mathrm{m}^{-2}$, but it is usually normalized to 1 or 100 with respect to the luminance of a given reference called *white*. Being proportional to physical power, double the power with the same SPD will produce double the luminance. However, brightness is not linear. A target with one fifth of the luminance will appear half as bright. Therefore we say that luminance is a linear measure of light whereas brightness, and, in general, the perception of light, is not.

Colour

Colour is the perceptual result of the distribution of luminous intensity over the visible spectrum. We perceive colour because we have (at least[1]) three different photopic receptors (cones) with corresponding different spectral sensitivity curves, which act as passband filters. The different combinations of the responses of the three types of cones is what we call colour. Provided that we (at least me, poor simple XY man) have only three types of cones, three numerical components are neccessary and sufficient to describe any perceivable colour. Therefore, colours can be mapped into three dimensional spaces called *colour spaces.*

In 1931[2] CIE determined the *Standard Observer* spectral sensitivity curves. These curves where obtained empirically, by showing to a set of

---

[1] There seem to be tetrachromatic women. See, for instance, Jordan and Mollon (1993).

[2] First time. Since then there have been several revisions. This is the reason for CIE SO curves being always accompanied by a year reference.

people circles of different aperture[3] of different spectral compositions, while asking them to generate the same colour on another screen by actuating on three primary sources.

The description of colour

Colour is usually described in natural language in terms of hue and saturation. We usually say "deep blue", "vivid red", "pale pink", and things like that. Hue is "the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colours: red, yellow, green, and blue, or a combination of two of them" (CIE). It is related to the dominant wavelength in the SPD. Saturation is "the colourfulness of an area judged in proportion to its brightness" (CIE). It goes from *neutral* gray to *pure* colour. It is related to the degree of concentration of the SPD around a dominant wavelength. Grays have a flat SPD, while pure colours are spikes, i.e. monochromatic or single line spectra. A colour can be desaturated preserving its hue by adding the same amount of light at all wavelengths.

Hue and saturation (together with brightness) are the common qualitative mesarues of colour used in natural language to describe different colours, but they cannot be easily quantified. A *colour specification system* is a mapping from the variety (infinite) of possible SPD to a three dimensional space, the colour space, such that each perceivable colour is represented by one and only one point in the space, and two different SPD perceived as the same colour by a Standard Observer are mapped to the same point in the colour space[4]. Colour specification systems are used to accurately reproduce colour, therefore they require high precision, even at the cost of considerable computational overhead. Colour specification systems should not be mistaken with *colour coding systems* such as RGB, CMYK, or HSV.

The CIE XYZ system

The basis for any colour specification system is the CIE XYZ system. Based on the Standard Observer curves, the CIE defined three tristimulus curves, $\bar{x}$, $\bar{y}$, and $\bar{z}$. Weighting of the SPD with the tristumulus curves produces the tristimulus values X, Y, and Z, where Y is luminance, because $\bar{y}$ was

---

[3] This is the reason why CIE Standard Observer curves are also usually accompanied by an aperture reference in degrees, such as the 1931 CIE 2° Standard Observer curves.

[4] Because, yes, two different light spectra may produce the same colour sensation; just take a look a little further on to how the colour coordinates X, Y, and Z –sufficient for colour specification– are obtained from any given spectra. They are called *metamers*.

chosen to match the photopic spectral sensitivity curve. Thus, X and Z are not related to any given primary or any given cone type. Any perceivable colour can be perfectly defined by its XYZ values.

The chromaticity diagram (x, y) is the projection on the XY plane of the cross section of the colour space XYZ by the plane X+Y+Z=1, i.e. x and y are the normalized versions of X and Y: $x = \mathrm{X}/(\mathrm{X}+\mathrm{Y}+\mathrm{Z})$ and $y = \mathrm{Y}/(\mathrm{X}+\mathrm{Y}+\mathrm{Z})$. The perceivable colours in the chromaticity diagram form a curved horseshoe shape with a flat straight base. The curved boundary is the line of pure or *spectral colours*, ranging from pure red to pure blue (both ends of the straight base) across pure green (the top of the horseshoe). All colours in this curve can only be formed by SPD consisting of a single spectral line. You will usually find the wavelength of the corresponding line written along the curved boundary. All colours inside the curve of pure colours are mixtures of pure colours. The flat base is the line of purples, which links pure red with pure blue, at both ends of the visible spectrum. Purples are not pure colours: no single spectral line can produce the colour sensation that we call purple.

Whites

About the center of the horseshoe is the region of whites, produced by similar amounts of the three primaries red, green, and blue. Near the region of whites it passes the *blackbody path*, which is the line described in the chromaticity diagram by the colour emitted by a hot blackbody when it is heated, according to Planck's Law, from a reddish white at about 2 000 K to a bluish white at about 8 000 K and beyond.

The point $x = 0.333, y = 0.333$ is the *Equal Energy Illuminant*, E, the colour of a light produced by exactly equal amounts of energy at all frequencies (null saturation). E is not in the blackbody path. Bluish whites are perceived as brighter whites. That is the reason for the point $x = 0.313, y = 0.329$, known as D_{65}, being often used as the white reference in video systems. White references are often (unaccurately) expressed in degrees Kelvin, even if they are rarely located exactly on the blackbody path, by the temperature corresponding to the closest point in the blackbody path in the chromaticity diagram. Thus, D_{65} is 6 504 K, and E is 5 400 K. Typical tungsten bulbs (incandescent lamps) are at $x = 0.44757, y = 0.40745$, i.e. 2 856 K. That is why light from household incandescent lamps is yellowish. The colour of passive (reflective) targets depends on the SPD of their reflectivity, but also on the SPD of the illuminant. A white paper (flat reflectivity) looks yellowish under tungsten illumination. However, our brain has the capability to adapt to different white references, and thus it is able to correct the effect of the illuminant even for white references relatively apart from E, such as tungsten bulbs. However, this only works in absence of a better white reference. If you look at the same time at a white sheet of paper with one half illuminated by a tungsten bulb and the other half illuminated by direct sunlight, the first half will appear clearly yellowish.

RGB systems

The set of linear combinations of any three points in the chromaticity diagram will cover all colours inside the triangle determined by the location of the three points or *primaries*. The curved shape of the chromaticity diagram implies that there cannot be three perceivable primaries which can cover the whole *gamut* of perceivable colours. RGB systems are colour *coding* systems which use three primaries (red, blue, and green, not necessarily pure) to describe a limited gamut of colours. Different RGB systems have different RGB primaries, and therefore different gamuts. However, there is no RGB system able to describe all perceivable colours. All display systems which produce colour by additive combination of three primary sources have a limited gamut. For a RGB system to be perfectly defined, its primaries have to be defined in terms of their XYZ values, or by their chromaticity coordinates (x, y) and a white reference. Transformation from a given RGB system to CIE XYZ and back is thus achieved by a linear transformation that can be expressed in matrix form.

Colour coding systems do not require such a high degree of accuracy as colour specification systems, and computational cost is much lower. R, G, and B components are usually encoded with integer values (8 bit, 16 bit, 12 bit). RGB colours thus encoded constitute a finite set of discrete samples within a triangle (gamut) inside the chromaticity diagram of all perceivable colours. Perceivable colours *out of gamut* of a given RGB display system are approximated by *in gamut* colours by means of a process called *colour rendering*, according to different criteria of colour fidelity.

Reflective systems (such as colour printers) are often based on the substractive combination of primaries, usually cyan, magenta and yellow, and therefore the overlap of gamuts between RGB displays and CMY(K) printers is not as good as desirable. This is the cause of many disappointments when printing to paper all those beautiful digital images. Colour Management Systems (CMS) are in charge of the translation between the different colour specification systems and colour encoding systems, both device independent and device dependent, by means of the so-called *colour profiles* standardized by the ICC (International Colour Consortium).

Perceptual uniformity

If you take a look at the CIE chromaticity diagram, you will note that greens cover a considerably greater area than reds or blues. As a matter of fact, the same amount of variation in a component in different parts of the diagram is not perceived as being of the same magnitude. This is because light perception is not linear but logarithmic. To account for this effect, the CIE defined two transformations of the XYZ system which produce two colour spaces approximately perceptually uniform, i.e. where if you move a colour at any location in the diagram a given distance, the Standard Observer will perceive the same amount of variation as if you move the same distance another colour in any other part of the diagram. These are the CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$, also known as CIELAB[5] and CIELUV. In Colour Lab, L^* represents *lightness*, the perceptually uniform counterpart of luminance, and a^* and b^* are two chromaticity components related to the green ($a^* = -128$) – red ($a^* = 127$) axis and the blue ($b^* = -128$) – yellow ($b^* = 127$) axis, respectively. Transformation from CIE XYZ to Colour Lab involves normalization by a white reference, $\frac{1}{3}$ exponents, and some scaling constants to scale the range of L^* to [0, 100] and of a^* and b^* to [-128, 127]. Colour Lab is anything but intuitive, and therefore it is usually used only as a reference colour specification standard, specially in CMS. Photoshop™, for instance, uses Colour Lab to renderize from device dependent or independent RGB to device dependent CMYK colour coordinates, and back.

Hue–Saturation–Whatever systems

Intuition is often used as the excuse to market HSB-like colour systems. HSB stands for Hue–Saturation–Brightness, and it is just one of several similar colour encoding systems (HSV, HSI, HLS,...) that claim to specify colour by its sensation qualities, or in the same terms as natural language does.

---

[5] Yes, the Colour Lab of Photoshop™.

The naked truth is that none of these systems is based on perceptual experimental studies (conversely to CIE XYZ or the systems derived from it) or established colour perception theories. They are rather more or less picturesque formulations (often lacking consistence) that produce colour components which are quite different from what they are supposed to produce. Thus, for instance, the "brightness" component (or "value", "lightness", or however it may be called in the different variations) is mostly defined as the average of the red, green, and blue components, contradicting all empirical studies about the perception of light, which tell us that the contribution of the different wavelengths to the sensation of brightness is not equal. Similar problems arise with the different formulations for saturation and hue, which usually are not correctly decoupled, such that, for instance, a change in the saturation or the hue component usually causes a change in luminance.

However, the greatest underlying problem is that all these "intuitive" colour spaces are just a nonlinear transformation of the RGB cube. This implies three sources of problems, affecting differents domains of application: 1) No RGB system can span the whole gamut of perceivable colours, as we have just seen. 2) There is not such a thing as "the" RGB colour space. There are as many RGB colour spaces as definitions of R, G, and B primaries. Any colour system formulated in terms of unspecified R, G, and B primaries lacks consistency. 3) Nonlinear transformation implies that straight paths across an RGB space are curves in HSB (and the like), and vice versa. This makes colour interpolation (ubiquitous in computer graphics) and colour rendering (for transformation between colour spaces) too complicated, in exchange for nothing.

Thus HSB-like systems should be relegated to human-machine interface (HMI) applications, such as colour pickers, where the typical hue disks with a slider for brightness can be of use for graphic software users not familiarized with colour spaces. However, HSB-like systems can be seen routinely used in vision applications based on hue, such as human skin detection in biometric applications. We have just seen that approach is not sound, and should be avoided unless there is a good reason for not to. If a colour specification system with a clear relationship with the natural description of colour in terms of hue and saturation is needed, it can be obtained by expressing the chromaticity coordinates of CIELAB in polar coordinates, h_{ab} for the angle and s_{ab} for the radius, corresponding to hue (between 0^o and 360^o) and saturation (between 0 and 1), whereas lightness (L*) measures brightness. The CIELAB coordinates can easily been obtained from any RGB camera outputs if the camera uses one of the standard RGB systems (such as the CCIR709), just by passing through the XYZ coordinates.

And what's next?

Well, hopefully now you know a little more on light and why it produces the sensations it produces. It's all about that pretty little thing called vision. Light is a complex source of information, and as such it has many faces, and can be analyzed from many points of view. Did I tell you already that artificial vision was a highly interdisciplinary discipline?

Now it should be the turn to take a look to how do devices designed to "perceive" light work. Those devices are more often than not called *cameras*, and they are the eyes of our seeing machines. They produce the electrical signal which feeds the artificial little "brains" with the necessary information about what is to be "seen". From then on, it is all about digesting that information.

But for now you will have to resort to some other source. There are a lot out there, and some of them really good!