

# A Laser Light Source For Endoscopy

James Larimer, PhD

## Executive Summary

This white paper makes the case for the potential benefits to be gained by replacing today's arc lamp illumination sources with a laser light source for endoscopic surgery. The primary advantage of a laser lamp is that it is many times brighter than an arc lamp illuminator. Bright illumination increases the signal amplitude and absolute luminance differences in spatial details within the image formed on the endoscope camera's sensor. The resulting higher signal contrast results in improved image quality.

Lasers have a lower service-life cost relative to arc lamps. They use less energy to make light than an arc lamp. Arc lamps require frequent and costly lamp replacement, whereas lasers do not require any replacement during the illumination system's lifetime. Additionally, a laser source light can be designed to reveal important surface details that would otherwise be invisible. This adds to the visual information available to the surgeon during the surgery. For example, spectral and temporal manipulation of laser-generated light can reveal the location of blood vessels, thus aiding the surgeon in avoiding accidental severing of an artery during the surgery and the resulting uncontrolled bleeding.

Lasers have been criticized because of their color rendering index (CRI) scores, but this is a simplistic and potentially irrelevant metric for this application. Color rendering as measured by the CIE's CRI is discussed in detail in this white paper. Two physically different lights, i.e., lights with different radiant power density functions within the visual band, can nonetheless appear to be identical to an observer. This is the fundamental property of human vision and it is called metamerism. It enables color TV. Two lights that are physically distinct but that appear to be identical to an observer are called metamers. Metamerism forms equivalence classes of physically distinct lights within the set of all possible lights that appear to be identical to an observer. Metameric equivalence classes are unique to individuals; two lights that appear identical to one person can look different to another. The CIE's Standard Observer describes metamer equivalence classes for an average observer; it does not precisely describe anyone's ability to see color differences. The CRI is based upon this Standard Observer.

Surface color appearance changes with lighting; this is a well-documented phenomenon of human vision. The idea that perceived color is constant under all lighting conditions is simply false. The context of a scene, the familiar objects within it, in conjunction with the nervous system's ability to extract meaning from a visual scene results in the human ability referred to as *appearance constancy*. *Appearance constancy* refers to the human ability to correctly name a surface color or to see the shape of an object independently of most viewing geometries and lighting conditions. Helmholtz attributed this perceptual invariance to unconscious inference. *Appearance constancy* too is imperfect; illusions do occur. The CRI is not a measure of *appearance constancy* and it has been criticized in the technical literature for this reason. Discounting laser illumination sources and solid-state lighting with similar CRI scores ignores task requirements and real appearance filtered by the cognitive abilities of the observer.

Laser lamps utilizing three or more primary bands produce more than adequate color rendering and enable image improvements, for example, revealing otherwise invisible surface features, not possible with arc lamps. The technical literature has recommended new light sources for endoscopy to achieve these goals; lasers greatly increase the potential utility of endoscopic illumination systems.

# A Laser Light Source For Endoscopy

James Larimer, PhD

## INTRODUCTION

The light source for a surgical endoscope serves three essential functions: (i) it provides sufficient light energy to illuminate the work-site and support the camera's requirement for a strong signal, i.e., bright light, to capture high quality images, (ii) it provides a signal with sufficient spectral diversity to support the discrimination of object surfaces based upon the visual band spectral reflection differences unique to materials, and (iii) it provides a light that reveals features in anatomical structures that are essential or useful for the successful completion of the surgeon's task. Finally, given the strong push to reduce medical care costs, the light source must provide these functions in an increasingly cost advantageous fashion in order to be adopted.

Broadband light sources based upon arc lamp and halogen technologies satisfy some but not all of these functions. They lack the endoscopic brightness required to produce high quality images that contain the contrast variations produced by object detail revealed in the tone scale of the image. Captured image contrast by eye or by sensor requires absolute signal differences above the noise floor of the sensing apparatus to be accurately recorded. The signal to noise ratio of the sensor, regardless of whether or not that sensor is a biological system (i.e., the surgeon's eyes) or the camera that collects the data to be reconstructed on a display to be viewed by the surgeon, is the rate limiting step to the optical performance of any endoscopic system. Adequate discrimination performance requires a bright light source to produce large surface reflectivity differences. In a dim light these differences in surface reflection are small. A bright light source produced by lasers eliminates the weak signal problem and produces significantly improved image quality.

The poor performance of filament lamps, arc lamps, and LEDs is the result of their large étendue and their poor intrinsic brightness relative to a laser source. Filament, arc and LED light sources produce incoherent light with orders of magnitude less intrinsic brightness than a laser. The small exit aperture of a laser can be efficiently coupled with minimal losses to a light guide. The inefficient optical coupling of arc or filament lamps to a light guides limits the brightness delivered to the work site and makes integrating these light sources into an illumination system a more challenging and less rewarding task.

The efficiency of arc and filament lamps, i.e., lumens per watt, is reduced because they produce a range of wavelengths outside the useful visual (VIS) band not present in laser light. Some of these wavelengths, particularly those in the far ultra-violet, must be filtered out because they can be harmful. The wavelengths used in a laser source can be individually modulated and crafted to serve several purposes. For example, it is conceivable to use the illumination source to perform pulse oximetry while illuminating the surgical site, or, through a combination of unique spectral bands and signal processing, to reveal surface features that might otherwise not be visible using ordinary lamps. The wavelengths generated by the laser light source can be selected entirely

within the VIS band, so no energy need be produced that will not be used in illuminating the work site. Additionally, useful wavelengths near the VIS band that are visible to the camera could also be added to improve image details that would otherwise be invisible, Silicon based sensors see electromagnetic radiation outside of the VIS band, especially into the near IR. These are wavelengths that interact strongly and non-destructively with biological materials. Sensor bandwidth when coupled with well-chosen laser wavelengths offers new opportunities in medical imaging not available with ordinary broadband sources.

### **SEEING COLOR: Physics & Chemistry**

Color and images are unique attributes of biological vision; they are not part of the physical world or separate from the living beings with visual sensory capabilities. Without biological visual systems there are no images in nature. Sensing our environment through vision, seeing objects in it and experiencing their surface colors are ubiquitous yet unique properties of our conscious experience. It is therefore hard for us to imagine that images are not independent physical attributes of nature. In the purely physical world of light and material objects, all that exists is a light field consisting of wavefronts of light produced by light emission from the sun and spectral shaping of that light as it reflects off of and interacts electrically and chemically with materials. There are terrestrial materials that produce light either by chemical reactions or through excitation induced by radiant and electrical energy. Light is absorbed, reflected, and sometimes – through excitation by radiant energy – materials become emitters. The cooling of rocks at the end of the day is an example of this reemission process, usually in the IR bands. These phenomena are illustrated in Figure 1.

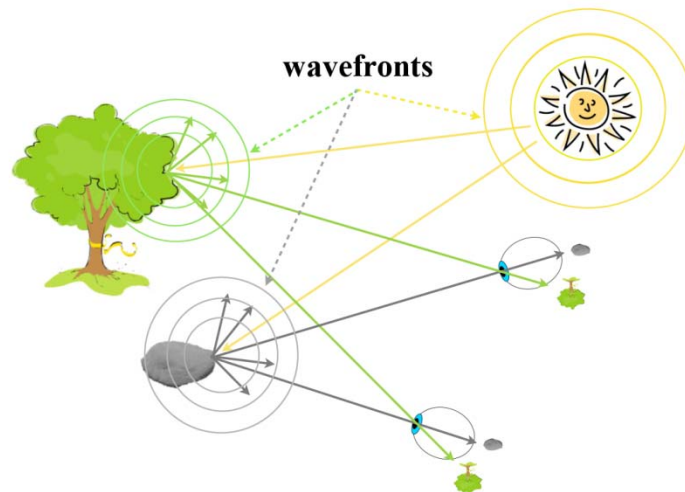


Fig. 1. Light sources illuminate the environment by emitting wavefronts of light that interact by absorption, reflection and luminescence produced by excitation by the radiant energy. The reflected and emitted light at each point on an object's surface produce new wavefronts added to those produced by other surfaces and sources. The collection of all wavefronts within a volume in space at a point in time is called the light field. The aperture of a camera or eye samples the light field moving towards it forming images from the information in the light field on a surface behind the aperture. If the aperture contains a lens, the lens perfects the directional information in the light field forming the image.

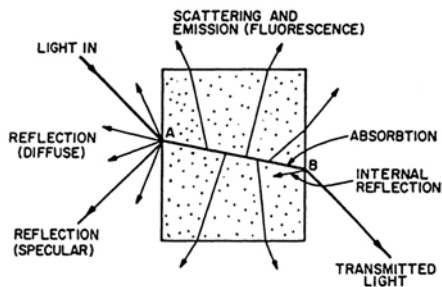


Fig. 2. This figure illustrates the interaction of light with materials.

On a microscopic level the mechanisms of interaction between radiant energy and materials are illustrated in Figure 2. This figure is taken from Kurt Nassau's chapters in a collection of essays entitled **Color for Science, Art and Technology** [1]. Nassau describes the 15 ways in which the spectral power distribution of an illuminant is altered by its interaction with materials. Illuminants are: reflected without change, the specular and diffuse reflections at the air-material interface; they are transmitted into the material where they are differentially

extinguished as a function of wavelength as they pass through or scatter out of the material; and finally they can excite the material to fluoresce creating new light or radiant energy.

The important interactions of light with materials occur within the VIS band of frequencies, illustrated in Figure 3. Electromagnetic radiation within this band of frequencies is called light because we can see it. Light interacts with the outer electrons of materials causing reactions that incorporate radiant energy into the material. This loss of energy in the illuminant changes the spectral composition of the radiation as it passes through the material. Fluorescence adds energy at unique frequencies to the reflected radiation. Frequencies below the VIS band, the IR and lower, generally warm up materials by causing vibrations within the material lattice. If we could see into the IR band we would see our own body heat; we are IR emitters. Higher frequencies above the VIS band, i.e., radiation in the UV band, ionize materials. Ionizing radiation destroys or is harmful to living tissue. Evolution selected the VIS band to develop the sense of sight because of its unique non-destructive interactions with materials. These interactions impart a unique material signature onto the light reflected from object surfaces. Spectral shaping through reflective interactions provides useful information to the observer.

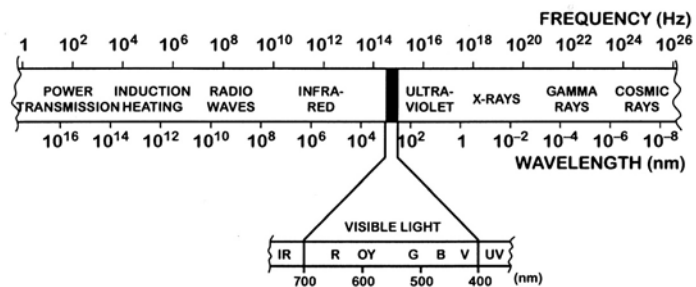


Fig. 3. The VIS band (light) within the spectrum of all electromagnetic radiation.

### SEEING COLOR: Psychology & Color Matching

Colors, like images, are ubiquitous attributes of our conscious experience but they are not a feature of physical phenomena apart from living beings. Colors result from the way biological vision systems code and interact with the light imaged onto the retina by our eyes. Newton is responsible for this non-intuitive discovery. In Newton's words:

*"And if at any time I speak of Light and Rays as coloured or endued with Colours, I would be understood to speak not philosophically and properly, but grossly, and accordingly to such*

*Conceptions as vulgar People in seeing all these Experiments would be apt to frame. For the Rays to speak properly are not coloured. In them is nothing else than a certain Power and Disposition to stir up a Sensation of this or that Colour.” [2]*

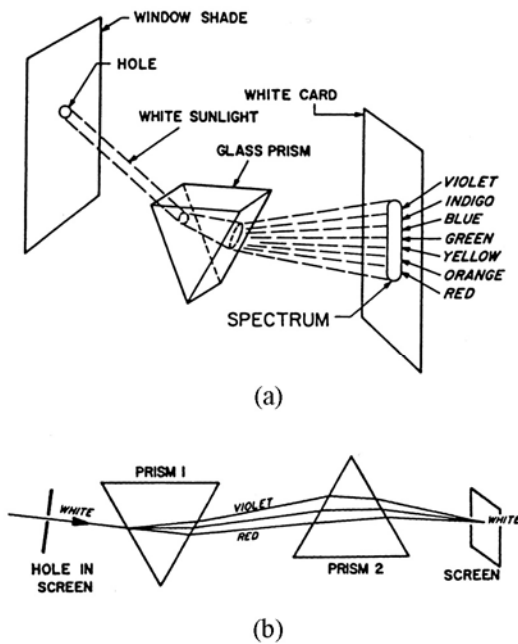


Fig. 4. Newton's 1666 experiments with light conducted in his rooms at Cambridge.

Newton's experiments with prisms are illustrated in Figure 4. Newton obtained prisms at a county fair that he later experimented with in his rooms at Cambridge. He shuttered his rooms leaving only a narrow slit for the light to enter. Into the beam of light he placed the prisms. Newton found that a prism broke sunlight into a spectrum or rainbow of colors. He concluded from this that white sunlight was actually made up of a mix of particles that he referred to as corpuscles that differed in their apparent colors. This is illustrated in (a) in Figure 4. He was also able to reverse the separation by using a second prism (and later mirrors to superimpose the individual bands back onto a single spot) thus recreating white light. He also experimented with mixing various bands of the rainbow and discovered that a mixture of the red and

green bands could always be adjusted in relative intensity to appear exactly the same to his eye as the yellow band of the rainbow. These identical appearing yellows, now called **metamers**, could be separated once again by a prism into red and green whereas the rainbow yellow when passed through a second prism remained yellow. From this observation Newton concluded that color was a sensory phenomena, an experience unique to vision, and not a property of light. Color is a property of seeing light.

The definition of a metamer is any two lights that appear to be identical in side-by-side comparison but that are physically distinct. Later through the insight of Thomas Young, Clerk Maxwell, and Hermann Grassmann a trichromatic theory was created to explain color-matching phenomena. A set of three primaries can always be found such that an additive mixture of their light in some proportion will match any arbitrary light. There is one caveat to this rule: to match an arbitrary color experience it may be necessary to add by superposition one of the primaries to the light to be matched. The primary desaturates the color to be matched in this case and the trichromatic ratio of primaries for this match contains a negative value. The reason color matches behave this way is that the eye contains three photoreceptors each with a unique broadband spectral sensitivity within the VIS band. Light isomerizes pigments within these detector elements and that process is linear. The eye uses the neural signals produced by these three photoreceptors when stimulated by light on the retina to encode a three dimensional representation of the light reflected from object surfaces.

### **SEEING COLOR: The dimensionality of color experience.**

Before discussing how color matching works and is quantified by the CIE, it is worth noting why only three detector mechanisms are required. The cone photoreceptors, responsible for vision in bright illumination, are spatially distributed within the retinal surface and act as punctuate sampling elements to encode the retinal image. The cones acquire information about the spectral quality of the light imaged onto the retinal surface, which is shaped by the interaction of light with the materials that reflect it, and they provide a signal to the visual nervous system to be analyzed, yielding information about the location, range, motion and type of objects within the immediate vicinity of the viewer.

The development of sensory systems occurred during the early phase of the Cambrian period of global evolution when life forms first developed the ability to move about with purpose. From a Darwinian perspective, the visual sensory system evolved to find prey and to avoid being prey. The evolutionary driver or forcing function of the visual system is to provide information about the environment immediately surrounding the individual. Heading and range to objects in our immediate environment is critical information for survival. Information contained in the retinal image allows us to identify objects, thus discovering opportunities and threats within our immediate vicinity.

Spatial visual capabilities developed to recognize objects, just as temporal visual capabilities evolved to inform flight or fight responses as objects move or we move toward objects. The design goals for visual systems exist in three distinct domains: temporal change, spatial detail, and object recognition through surface properties such as color. The visual system trades off these three objectives as it gathers image information. All three depend upon the amplitude of the retinal signal, so visual systems evolved to measure and encode intensity differences in space, time and chroma (color). Gathering visual information needed to identify objects within our immediate vicinity by their material interactions with light, i.e., color, is no more important than the heading, range and spatial details required to inform and motivate actions.

The retinal surface is limited in extent as are the dimensions of the cone photoreceptors. The entrance aperture of a cone photoreceptor for the average person is approximately  $3\mu\text{m}$  in diameter within the fovea of the retina in the human eye. This puts a limit on the spatial resolution of the eye, matched closely to the diffraction limit of the eye's optics. The spectral tuning of the cones varies with the three cone photoreceptor types. A uniform field of light on the retina will generate a pattern of three distinct response amplitudes due to differences in the spectral tuning or sensitivities of the three cone types. This creates a fixed pattern of signal differences within a region in the retinal image illuminated by homogenous light. This pattern of amplitude differences required to identify the color of surfaces limits the spatial resolution of the eye. In a camera sensor these amplitude differences produced by spectral tuning are called fixed pattern noise. This trade-off is discussed in more detail in Appendix A.

To see spatial detail in the retinal image, amplitude differences in neighboring cone signals provides spatial information about the objects imaged onto the retina. The cone photoreceptors near an edge will signal the location of the edge by the spatially contiguous amplitude variations in their signals. One side of an edge may be a bright red in color; the average amplitude and cone triad pattern of amplitudes will all be different from the opposite side of the edge. If the edge was created by a shadow falling across the surface of the object, the cone signals within the shadow will all be of a lower amplitude and due to less intense illuminant in the shadow, i.e., all scattered light. The triad amplitude signals will also vary in the shadow as scattered light has a different spectral composition than direct illumination.

The spatial ability of the visual system is limited by its ability to detect the spectral quality of the light reflected from surfaces because spectral sensitivity differences within neighboring cones on each side of the edge or near it add a spatial variation similar to shot noise. This is discussed with an example in Appendix A. At an edge boundary this cone signal spatial variation caused by spectral tuning will limit the visual system's ability to precisely extract the edge location, thus limiting the spatial resolution of the eye. The spectral signature of a surface is contained in the cone signal differences within spatially homogenous regions in the image, i.e., the image of a object's interior, and in conjunction with spatial details, i.e., edges and textures, provide useful information to the individual.

This information gathering trade-off is a classic engineering design problem; it was resolved in biological visual systems by a Darwinian forcing function. Increasing the capability to discriminate the spectral composition of the light imaged onto the retina necessarily requires giving up more spatial visual capabilities. The question remains why are three cone types sufficient for the eye to properly classify a surface? The answer again lies in the way radiant energy in the VIS band interacts with materials.

The type of material and light interactions in the VIS band that impart an object's signature onto the light those objects reflect involves mostly outer electron orbits within

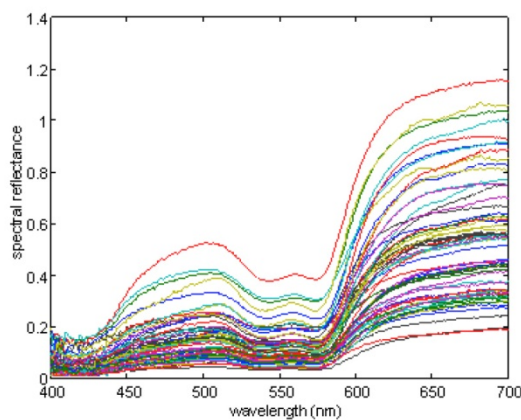


Fig. 5. Examples of measured reflectance spectra from approximately 50 rectal membrane tissue samples.

the material. The physical phenomena evoked by light that result in changing the spectral composition of the reflected light are highly constrained as many potential band gaps within the materials become involved in the interactions that originated mostly in the outer electron orbits. There is a resulting high degree of correlated activity within the spectral frequency domain contained in the VIS band.

The exception to this rule is interference, or equivalently the diffractive interactions of light with physical structures on a scale of the wavelength of light (such as

Newton's rings). These interactions are sharply tuned in the frequency domain because they depend upon geometry. With this exception, light and material interactions turn out to be of low dimensionality. Consider the samples of broadband white light reflected from approximately 50 samples of human rectal membrane tissue shown in Figure 5. These samples are characteristic of light's interaction with virtually all materials in nature.

The spectral samples shown in Figure 5 were subjected to a principle components analysis to find the *eigen* values and *eigen vectors* forming a linear basis set that spans these spectra [3]. The results of that analysis are shown graphically in Figure 6.

It is apparent from Figure 6 that 95% of the variance in the spectra is accounted for based upon a three-dimensional vector space approximation. That is to say, this family of functions, i.e., the reflections from tissue, can each be described accurately with three parameters. A linear combination of just three spectral functions will almost perfectly fit any of these reflectance functions. This means that the spectral reflectance functions are inherently of low dimension and increasing the number of spectral samples will do little to improve the sensor's ability to recover the unique signature of the surface.

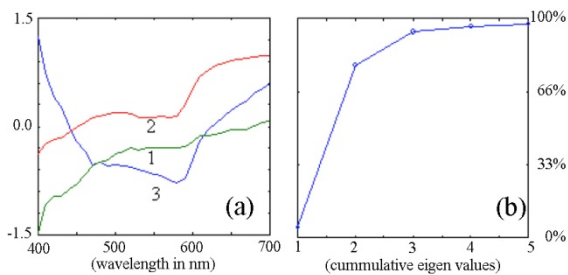


Fig. 6. Eigen vectors (a) and cumulative variance accounted for, i.e., goodness of fit, in spectra by number of eigen vectors employed in basis set (b).

The use of principle components analysis to characterize the dimensionality of reflected light has a long and rich history; for recent developments and discussion see [14]. The general rule is that three basis functions will span most spectral reflection functions within the VIS band encountered in natural settings regardless of the light source, and with a high degree of accuracy using only three basis functions to describe the spectra.

Evolution apparently discovered this excellent physical approximation and used it to refine the trade-off of spectral resolution for spatial resolution in our visual system. We have a three dimensional visual system in the amplitude domain of the image formed on the retina because adding more sensor types would only make a small improvement in our ability to recognize the unique surface reflection signatures of objects. Adding more cone types would sacrifice our ability to analyze spatial details in the image.

That a three band light source is fully adequate to illuminate objects and to obtain good color rendering was pointed out in the 1960s by Thornton in his papers recommending modifications to the more efficient florescent lamp [5]. This idea has already been proposed for use in endoscopes [6]. A laser light source is the ideal embodiment of Thornton's design recommendations for efficient lamps with fully adequate color rendering properties.

### SEEING COLOR: The CIE Standard Observer.

Color matching is described as a three-dimensional linear vector space, because the sensory behavior of color matching obeys what is known as Grassmann's Laws. These laws define a three-dimensional linear vector space. The physiology behind these laws is the manner by which light is encoded by the visual nervous system. Light imaged onto pigments in the photoreceptors isomerizes the pigments, i.e., bleaches them, and this process is linear. The three cones have different action spectra determined by a pigment molecule unique within each cone type that bleaches in light. Because the system is linear, color matches can be described and captured by any basis set defined over the VIS band that is a linear transformation of the action spectra of the cones.

The description of color matches and their characterization would be precise for all people if everyone had identical cone action spectra; we almost do have this condition. For a large set of the human population, easily 95%, our cone pigments are identical. What varies across people and during our lifetimes is what can be described as a personal chromatic filter through which each of us views the world. This unique filter means that my color matches will not be identical to yours; patches of light that are metamers for me, i.e., that is they are physically different spectra but I am unable to tell them apart by looking at them with my unaided eye, may appear to be different to your eye. It is the action spectra of the cone photoreceptors in combination with these personal filters that form the basis set for each person's unique color matching functions.

The CIE has resolved this problem by not attempting to describe individual color matching behavior but instead to develop a normative standard, the CIE Standard Observer, that represents an average observer. All CIE color standards use this representation to characterize dyes, paints, fabrics, light sources, illuminants, etc. -

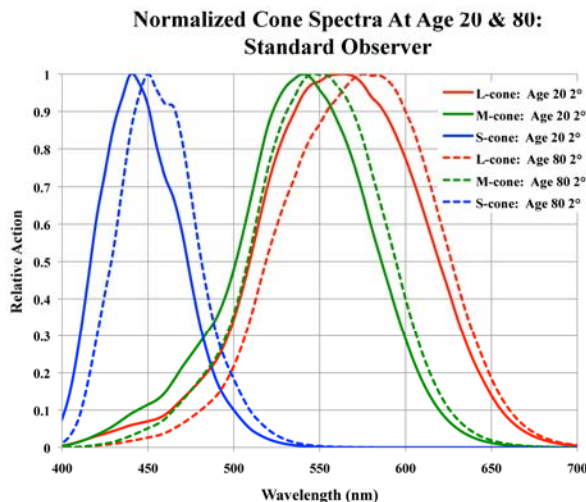


Fig. 7. The solid spectra represent the normalized cone action spectra for a 20 year old norm group viewing a 2° patch foveally and the dashed spectra represent the same condition for a 80 year old norm group.

anything that impacts a color match made by a Standard Observer. The CIE Standard Observer has a built in uniqueness filter that represents the average person's personal filter. A complete description of the CIE Standard Observer is available online at [7].

Recently the CIE has adopted a new standard that incorporates individual differences stratified by age cohort groups. This new standard uses cone-fundamentals and an age and image size dependent

filter to describe color-matching performance in age related norm groups. The standard is meant to be a research tool rather than an engineering tool, but it is useful in illustrating

the differences between age groups. Figure 7 illustrates the differences in the cone action spectra for 20 year olds compared to a norm group of age 80. These differences are due primarily to yellowing in the lens, cornea and media of the eye and a loss of sensitivity due to age related entopic light scatter in the eye [8].

The CIE Standard Observer tristimulus values that form the basis for all CIE standards referring to light and color appearance are defined as:

$$X = \int_{380}^{780} I(\lambda)\bar{x}(\lambda)d\lambda,$$

$$Y = \int_{380}^{780} I(\lambda)\bar{y}(\lambda)d\lambda, \text{ and}$$

$$Z = \int_{380}^{780} I(\lambda)\bar{z}(\lambda)d\lambda,$$

where  $I(\lambda)$  is the spectral power distribution for a homogenous patch of light imaged onto the retina and measured at the cornea, and  $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$ , and  $\bar{z}(\lambda)$  are the Standard Observer's color matching functions. X, Y and Z, the tristimulus values, represent the putative signals captured in the each of these three transformed color mechanisms. A 3x3 matrix linear transform relates these tristimulus values to the signals generated in the three types of cone photoreceptors. Whenever the tristimulus values are the same for two physically different lights  $a$  and  $b$  with respective power density functions  $I_a(\lambda) \neq I_b(\lambda)$  those two lights will appear to be identical to the Standard Observer. Even though two lights,  $a$  and  $b$ , are a color match for the standard observer, they are likely to be slightly different in appearance to real observers due to individual differences produced by the each person's unique personal filter.

### **SEEING COLOR: Color Constancy.**

The color of surfaces provides important information with survival value to the observer. The color of food crops and grasses reveals a need for water and the health of the crop; the color the sky often indicates the likelihood of a storm; the color of food can reveal when food has spoiled even when it does not smell spoiled; and flesh tones in another person's complexion can be a good indicator of that person's health. A cyanotic or jaundiced hue in a person's cheeks can be a harbinger of a serious and contagious illness. For all of these reasons recognizing surface colors is an important sensory ability.

The earlier discussion of the physics and chemistry of surface colors described how light reflected from a normally non-emitting surface in a particular direction is the result of a complex interaction of the illuminant with the material and geometry of the surface. Even the orientation of the surface normal relative to the location of the illuminant and the observer's position impact the quality and quantity of reflected light. There are specular and diffuse reflections of the illuminant at the air-material interface that redirect some of the illuminant spectrally unaltered towards the observer. There are transmitted and scattered components of the illuminant, spectrally altered by unique extinction events created by interactions of the illuminant and the material, and finally there is the

possibility that components of the illuminant either from inside or outside the VIS band have excited florescent events within the material causing VIS band emissions in the reflected wavefronts moving towards the observer. These phenomena can be characterized coarsely by the expression:

$$I(\lambda) = \alpha \cdot e(\lambda) + (1 - \alpha) \cdot e(\lambda) \cdot \rho(\lambda) + \phi(\lambda),$$

where  $\alpha$  represents the diffuse and specular reflections at the air-material interface reflected towards the observer's eye, with  $0 < \alpha < 1$ ;  $e(\lambda) \cdot \rho(\lambda)$  represents the transmitted and/or scattered light produced by the illuminant interacting with the material and headed towards the eye;  $e(\lambda)$  is the spectral power distribution of the illuminant at the surface;  $\rho(\lambda)$  is the spectral reflectivity (or spectral albedo) of the surface, and  $\phi(\lambda)$  is any light emitted by the material as a result of florescence excited by the illuminant.

A change in the position of the observer, or the location of the illuminant, or the spectra power distribution of the illuminant typical of sunlight during the diurnal cycle – any of these changes – will lead to a change in the spectral power distribution,  $I(\lambda)$ , arriving at the cornea of the observer. Any change in  $I(\lambda)$  is likely to produce a change in the corresponding *XYZ* tristimulus values. In the vernacular, colors change with the lighting; this is something familiar to everyone.

Drapes that match the wall color in sunlight no longer match when a room is illuminated with florescent light. Colors appear to be more or less saturated at dusk than they do in high noon illumination. In the color science literature these phenomena are referred to as metamerism and refer specifically to surfaces that appear to be identical under one illuminant changing color appearance when viewed with a different illuminant.

Because color is an important attribute in terms of identifying objects, biological vision systems have evolved mechanisms that attempt to maintain a constant color experience under changing illumination conditions. For example, in reddish firelight white surfaces reflect more red light than the same surface viewed under a neutral 'whitish' illuminant because firelight contains more long wavelength photons. The visual system has evolved two types of mechanisms that can be thought of as restoring the white balance in the visual scene to maintain a constant color experience [9]. One of these mechanisms operates like an automatic gain control in a radio. As one cone signal tends to dominate in a region of the image the gain in that cone pathway is reduced. This is referred to as von Kries adaptation and it supports making neutral white or grayish surfaces appear neutral even when illuminated by nonwhite illuminants. But this mechanism is imperfect, for example, a bright yellow surface in a dark surround can appear to be greenish when illuminated by firelight. Observations like this underscore the failure of perfect color constancy.

A second mechanism of color constancy is a restoring force that offsets the coloration in the illuminant by producing an internal signal in the color pathway of the visual system that is complementary to the coloration of the illuminant. This signal acts like a desaturant added to the illuminant. Colored afterimages often last longer than the time

required for a von kries gain change indicating a separate time course for this second mechanism designed to “white balance” our color experience. The point is that neither of these mechanisms works perfectly, color constancy is at best approximate. More often than not our experience of the color of a surface changes with changes in the illuminant that we understand are natural events. We see the colors change, but our cognitive processes are able to attribute this to changes in the illuminant and not to changes in the surface. This topic will be discussed again below in the context the color rendering index and what has been called appearance constancy in the technical literature. Helmholtz labeled these phenomena *unconscious inference* implying that they are the result of cognitive processing occurring through a subconscious neural mechanism involving prior experience.

The eye operates primarily by sensing differences rather than absolute quantities of the light imaged on the retina. As many photographers understand, the eye is an unreliable estimator of light levels. For example, a sunrise in a movie may appear to be blinding bright when the light on the cinema screen is actually quite dim in absolute measure. The human visual system is said to be ac-coupled, meaning that it senses differences and not absolute quantities. There is plenty of evidence in the vision research literature for this statement with the most striking evidence being that if an image is completely stabilized on the retina it vanishes from our conscious experience despite no change in the energy on the retinal encoding apparatus [10]. Bright light sources produce greater absolute differences in reflectivity and this turns out to be related to the eye’s ability to see relative differences.

### **SEEING COLOR: The CIE’s Color Rendering Index (CRI).**

When significant surface colors change in appearance due to the illuminant, observers often notice these changes. Some florescent lamps produce a cyanotic cast to flesh tones that can be highly objectionable. The previously described loss of a color harmony between draperies and wall coverings under some illuminants has motivated a great deal of effort in the illumination manufacturing industry to produce lamps that minimize these effects. As discussed above, color constancy is imperfect; everyone sees surface colors changing with illumination. Maintaining color matches and minimizing color mismatches as the illumination changes remains an important consideration, for example, when an automobile is repaired after an accident and new paint must be applied. In these circumstances, preserving color matches under as many illumination conditions as possible is a reasonable goal. To aid this effort, the CIE created a measure called the color-rendering index (CRI) to quantify color match preservation [11]. It is important to note at this point that exploiting the differences in surface appearance created by changing the illuminant can also be an advantage by revealing details that might otherwise not be apparent. Dermatologists use special lamps to detect tone and color variations in skin lesions that would be difficult to see with ordinary light.

The CIE Standard Observer defines a three-dimensional vector space, *XYZ* tristimulus values, where these vectors correspond to the signals produced when the Standard Observer *sees* a surface. Two different colors will produce two distinct vectors in this

color space. The distance between these vectors is related to how different these lights appear to the Standard Observer.

Name	Appr. Munsell	Appearance under daylight	Swatch
TCS01	7.5 R 6/4	Light greyish red	
TCS02	5 Y 6/4	Dark greyish yellow	
TCS03	5 GY 6/8	Strong yellow green	
TCS04	2.5 G 6/6	Moderate yellowish green	
TCS05	10 BG 6/4	Light bluish green	
TCS06	5 PB 6/8	Light blue	
TCS07	2.5 P 6/8	Light violet	
TCS08	10 P 6/8	Light reddish purple	
TCS09	4.5 R 4/13	Strong red	
TCS10	5 Y 8/10	Strong yellow	
TCS11	4.5 G 5/8	Strong green	
TCS12	3 PB 3/11	Strong blue	
TCS13	5 YR 8/4	Light yellowish pink (skin)	
TCS14	5 GY 4/4	Moderate olive green (leaf)	
TCS15	1 YR 6/4	Asian skin	

Fig. 8. The 15 test samples specified in the most recent version of the CIE color rendering index, see [11].

The  $XYZ$  axes based upon the  $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$ , and  $\bar{z}(\lambda)$  color matching functions define the Standard Observer. These axes were selected in 1931 to make the computation of the tristimulus values easy to accomplish on a mechanical adding machine. Additionally, the  $\bar{y}(\lambda)$  color matching function was made to correspond to  $V_\lambda$ , the photopic luminosity function. The luminance of a surface is also determined along with its tristimulus values in the color metrology procedure. The choices made to structure this procedure were conventional choices made by the CIE in 1931 to make the standard useful tool for industry.

$V_\lambda$  is a measure of the eye's ability to see an intensity difference across an edge.  $V_\lambda$  was originally determined by flicker photometry but it applies equally to edge visibility. To see spatial or temporal differences in intensity, i.e., an edge, the visual system *computes* the signal difference between neighboring cones or the temporal difference in a single cone within a short interval of time. For this reason it is not surprising that edge distinctiveness should be closely related to color matching. The color matching functions of the Standard Observer are related by a linear transformation, i.e., a 3x3 matrix, to the cone action spectra of the Standard Observer. An arbitrary linear transformation, in this case for computational convenience, does not preserve metric distances. The computed metric distance between two points in XYZ tristimulus space is not necessarily related to the discriminability of the two corresponding patches of light.

In 1965 [12], the CIE adopted a methodology to measure the CRI. The method included the adoption of a color transform space, in this case a non-linear transformation of the  $XYZ$  tristimulus space, where the metric distances in this transformed space were believed to be more closely related to perceived color differences, i.e., the discriminability, of colors. The discriminability of two surfaces depends upon their hue difference and brightness differences. The space adopted by the CIE in 1965 attempts to take both into account.

Quantifying how well color matches are maintained when the illumination changes is the primary goal of the CRI. All color matches, however, are not equally important to observers. Part of the CRI standards definition effort was therefore to select a set of surfaces which are believed to be representative of the most

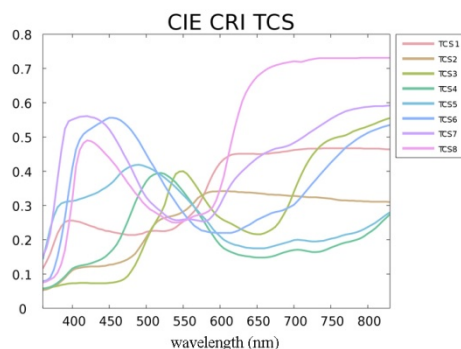


Fig. 9. The spectral reflectance functions of the first 8 CRI test surfaces, see [11].

important matches to be preserved when the lighting changes. Eight surface reflectances were selected and later seven more were added making a total of 15 surfaces [11]. The approximate color of these surfaces and their Munsell values are shown in Figure 8. The spectral reflectance functions for the first 8 are shown in Figure 9.

To compute the CRI for a particular illuminant, the black body radiator with the closest correlated color temperature to this illuminant is determined. The light reflected from each of the eight primary test surfaces when they are illuminated by the test illuminant and a second ideal illuminant of an equivalent black body temperature are determined. This produces 16 power density functions (PDF) representing the reflected light, two each for each of the eight test surfaces. Next the XYZ tristimulus values are computed for each PDF and transformed to the CIE CRI color space, see [11]. The triads for the 8 test patches are transformed to the same von kries adaptation state equivalent to the adaptation state that would be produced by the ideal black body radiator. Then for each test surface pair the metric distance is computed. Finally, an average color distance is computed for the 8 pairs and the formula for the CRI is:

$$R_a = 100 - 4.6\Delta\bar{E}_{UVW},$$

where  $\Delta\bar{E}_{UVW}$  is the average color distance between the paired eight surfaces illuminated by the two illuminants. A block diagram showing the steps required to compute the CRI is shown in Figure 10. An example of this procedure using 18 MacBeth Color Checker Samples is provided in Appendix B.

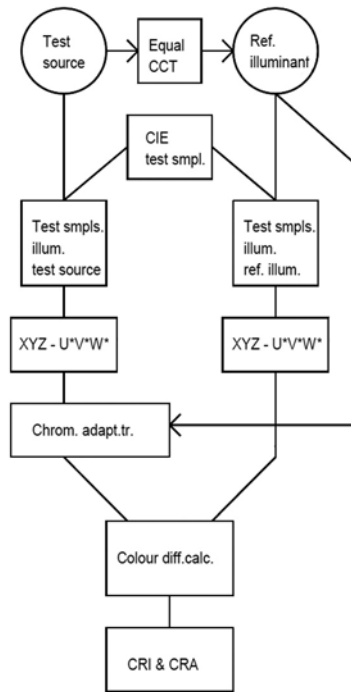


Fig. 10. A block diagram of the steps required calculating the CIE color-rendering index CRI.

### **SEEING COLOR: A Critique of the CRI and Comments.**

Complaints about the CIE CRI arose in the technical literature soon after the first version of the standard had been adopted by the CIE [13]. It has been revised several times since its original publication. Some of the early complaints focused on the choice of test surfaces and whether or not they were sufficiently representative of the surfaces that are encountered in everyday usage. A more challenging critique is based upon the observation that color appearance is not constant, and, moreover, a color space that more closely represents the Standard Observer's ability to see that two surfaces are different or the same will necessarily depend upon parameters determined by the viewing context [14]. Color appearance depends upon interactions with neighboring colors as well as the spectrum of the light imaged on the retina.

Colors will often shift and change with brightness levels. Dimming the illuminant, but not changing its spectrum, can produce an apparent color shift. Most of the time the context in which surfaces are viewed provides sufficient contextual information to the observer for the apparent color to remain unchanged subjectively despite the observer noticing that the surface is now darker or illuminated by a colored light source. A shadow falling across a homogenous surface does not appear to change the color of the surface despite making it darker nor does transparency phenomena where the viewer sees a surface below and through a transparent surface. The subjective ability of observers to correctly see the surface color independent of the illuminant is referred to as *appearance constancy* to distinguish it from color constancy [14]. Models of color experience that attempt to account for appearance constancy move out of the purely sensory psychology of vision into the realm of cognitive psychology. Some models are now being proposed to bridge this gap [15].

Whether or not this level of complexity has practical significance depends clearly upon the task requirements. For example, the manufacture of automobile paints may only care about maintaining color matches for touch up paint within a set of illuminants determined by the nature's changing seasonal illumination plus how the paint maintains color matches in the automobile showroom. For this task an application of the current CIE CRI restricted to a few notable light sources and perhaps some task related surfaces might be completely sufficient.

There is good evidence that three narrow bands, the so-called RGB primary colors, when combined in an illumination source and properly white balanced actually do an exceedingly good job, i.e., at least as good as appearance changes that occur naturally, as traditional broadband illuminants based upon filaments and arcs [5]. Thornton [5] compared lamps not only in terms of their CRI but their power efficiency, eta denoted by the Greek letter  $\eta$ . Eta is defined as:

$$\eta = \frac{\int_{380}^{780} V(\lambda)I(\lambda)d\lambda}{\int_{380}^{780} I(\lambda)d\lambda},$$

and its units are lumens per watt. Lumens per watt are a critically important measure of a light source characterized in the task environment. In highway lighting, for example, it is the number of lumens on the street surface per watt of lamp power that is critical. Eta does not include the power required to produce  $I(\lambda)$  that also must be taken into account when the power budget is the issue. Low-pressure sodium lamps have a poor CRI and a very high  $\eta$  making them the light source of choice for streetlights. Thornton proved that by employing three carefully chosen narrow spectral bands both CRI and  $\eta$  can be optimized. The exact optimization is task dependent.

In the endoscopic surgical task, the ability to reveal critical surface details and changing surface characteristics are as important – or even more important – than exactly matching the color appearance of surfaces when seen in natural light. Seeing surface detail is a spatial task of the human visual system. To perform it well requires bright illumination. Today because of the large exit aperture of arc lamps and resulting étendue it is difficult to optically couple these lamps with the endoscopic illuminator resulting in lower than desired light levels at the surgical site.

A laser lamp does not have this defect and it can provide a much brighter light with both adequate measures of CRI and  $\eta$ . CRI is not directly related to image quality in part because no lamps have CRI scores of 100 and changing color appearance of surfaces by changing the illuminant is highly task specific. Making flesh tones appear bluish maybe bad for restaurant lighting and exactly what is needed in the dermatologist’s examining room where revealed lesion details may have diagnostic relevance.

### SEEING SURFACE DETAIL

The ability to see surface detail is a spatial vision task. The human visual system, as noted previously, is thought to be ac-coupled. It is designed to detect differences. A useful metaphor for how spatial vision works is to imagine that the visual system computes derivatives as it analyzes images formed on the retina. Texture and structure are spatial differences in light intensity in an image. There are several measures of contrast that have been used to describe images. The most familiar is simple contrast defined as

$C = L_{\max}/L_{\min}$ <sup>1</sup>. Displays are often characterized by this measure where  $L_{\min}$  is the measured luminance of the full screen black state and  $L_{\max}$  is the luminance of the full screen white state. Vision research has used Michelson Contrast to characterize the contrast in an image. Michelson Contrast is defined as

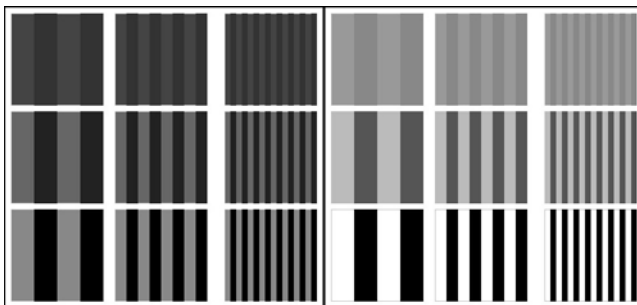


Fig. 11. The type of grating targets used to characterize spatial vision are shown in this figure. Typically the gratings used are sine wave gratings not bar gratings. The set of 9 gratings on the left represent a low average or mean luminance in 3 columns of different spatial frequencies, increasing left to right. The 9 gratings on the right where constructed at the same spatial frequencies and contrasts as the 9 gratings on the the left but at a higher mean luminance.

11cd/m<sup>2</sup>, many flat panel TVs deliver 500

$C_m = (L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$ . If in a small region of the image half the area is at the  $L_{\max}$  luminance and the other half is at  $L_{\min}$  luminance, then the Michelson Contrast is essentially  $C_m = 2AC/DC$ .

Sine wave gratings similar to the bar gratings shown Figure 11 were used by van Nes and Bouman [16] to measure the contrast tuning characteristic of the eye, called the contrast sensitivity function (*csf*) in the technical literature. For each of several mean luminance levels a sine wave frequency was selected and the contrast varied to determine the minimal contrast at which the grating was still visible. This was repeated at a series of spatial frequencies and at several mean luminance levels. The results are shown in Figure 12.

Each curve in Figure 12 represents a different mean luminance, measured in Trolands, a luminance measure of the intensity of the light falling on the retina<sup>2</sup>. As the mean luminance increases the sensitivity to contrast at most spatial frequencies except the lowest ones also increases, i.e., less contrast is required to see the grating. The eye goes from acting as a low pass spatial filter at low luminance levels to being band pass at high luminance levels.

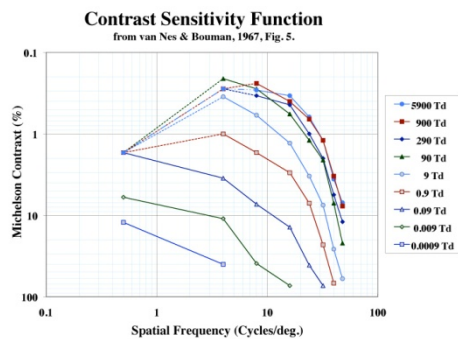


Fig. 12. The contrast sensitivity function.

Sensory systems follow a rule that is called Weber's Law. Consider this example; you have a thimble and it is half full of water. How much water do you need to add to the thimble to notice that the water level has risen? The answer is not much. Now suppose instead that you have a bathtub half filled with water, would the same amount of water that you had previously added to the thimble now noticeably raise the level in the bathtub? It is easy to understand that it would not. This is an example of Weber's Law.

When Weber's Law is applied to seeing increases in light levels, we find that once a critical minimal level as been obtained, this rule applies:  $\Delta L / L_{\text{background}} = k$ . The amount of light that you must add to the background amount of light that is already present is a constant proportion of the background. The contrast at the threshold of seeing the grating remains constant. What this means in words is that it is easy to see a candle burning at a great distance at night when the level difference between the dark night sky and the candle is large enough, but during the day even a search light, with many orders of magnitude more luminance relative to the candle light, might go unnoticed at the same distance.

<sup>2</sup> The amount of light entering the eye depends upon the pupil diameter. The Troland takes into account this factor. For example a light of 100 cd/m<sup>2</sup> would be four times more intense on the retina when the pupil is 4mm than when it is 2 mm. The correction factor converting the cd/m<sup>2</sup> is the proportionality factor determined by the area of the pupil.

Weber's Law is quite similar to the way Michelson Contrast is computed. Weber's Law means that once a certain mean luminance or DC level is obtained, the contrast required to see the grating reaches a maximum of sensitivity and does not improve with further increases in the DC level. This level is reached very quickly as brightness increases for low spatial frequencies, which is why the curves all meet up on the left side of the graph in Figure 12. The van Nes and Bouman data shown in Figure 12 are replotted as a function of mean luminance in Figure 13.

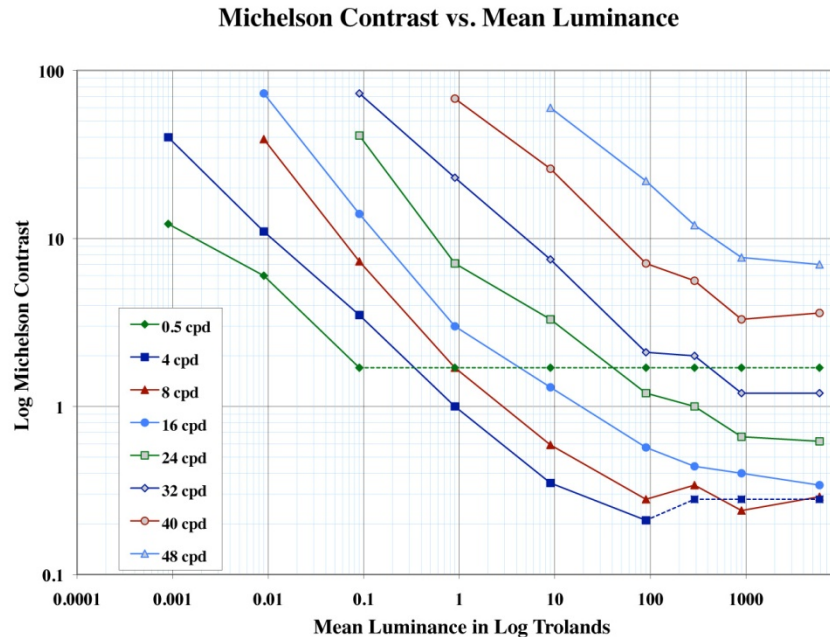


Fig. 13. The data of van Nes & Bouman (1967) plotted as a function of mean luminance level. Weber's Law applies when the curves flatten and become parallel to the abscissa. At 1000 Trolands all frequencies obey Weber's Law.

The Weber fraction  $k$  depends upon spatial frequency and is optimal around 4 cycles per degree visual angle where the human eye has its best spatial sensitivity. As light levels increase from darkness to bright illumination the eye's sensitivity to spatial details continuously improve until they reach a maximum where Weber's Law prevails. This point occurs at the foot of the negative sloped portion of each spatial frequency curve in Figure 13 starting from a high contrast on the left side of the graph. By 1000 Trolands which correspond to approximately  $140 \text{ cd/m}^2$  for a 3 mm pupil, typical of adults in moderately bright illumination, the optimal spatial sensitivity is achieved for every spatial frequency. The ability to see contrast at all spatial frequencies within the human visual system's window of visibility is critical to visually guided tasks such as surgery. This is why surgical theaters are so brightly illuminated to provide the best viewing conditions for the surgeon. The ability to see surface details, blood vessels and color and brightness changes in organs and structures at the site of the surgery is necessary for successful outcomes.

### **SEEING SURFACE DETAIL: Endoscopic Image Capture and Display.**

Cameras using CMOS and CCD sensor arrays share many properties with human vision. In a sensor, a quantum well, a capacitor, is charged by light. Over much of the dynamic range of the sensor the charge level is linear with the light flux per unit area and time. However, the electronics are noise limited, so the light to charge a level transfer curve characterizing the sensor pixel has a knee and shoulder similar to the dynamic transfer function of photoreceptor in the eye or film. The transfer curve relating light level to signal is s-shaped with floor, i.e., where light has little impact, and ceiling, where the process saturates. The region between the floor and ceiling show linear increases in signal with light levels.

Unlike the eye, which has a regional gain function so that different regions of the retinal surface can operate at different gain levels, CMOS and CCD sensors tend to have only a global gain mechanism. Both cameras and eyeball optics are limited by diffraction and light scattered from lens surfaces, the interior surfaces of the camera or eye, and in the case of the eye by scattering in the transparent humors that fill the outer and inner chambers of the eye. Scatter in the eye is called disability glare and it limits the dynamic range of intensity differences possible in the image on the retina to about 60 db. The very best cameras also have a glare limited dynamic range of about 60 db (3 log units). The tone scale or differences in levels must be extracted from this optimal dynamic range.

When the illuminant lighting the surgical work site is dim, the camera will be operating near the noise floor of the sensor and in a constrained dynamic range. This means that the number of level differences that can be captured will be limited and below the optimal potential of the camera. A bright illuminant is required to mitigate this situation. When the number of captured levels is non-optimal, no amount of signal processing or dynamic range stretching on the display device can make up for the loss in dynamic resolution produced by low light levels at the surgical site. This is why a bright endoscopic illumination system is essential to provide peak performance of the endoscopic vision system.

Additionally, a laser illumination system, because it can select narrow spectral bands designed to enhance the contrast of blood vessels and other anatomical structures, can actually provide a brightly illuminated work site with the potential to reveal features that might be difficult to visualize with a less wavelength selective illumination source, such as a Xenon arc-lamp. Lasers can be temporally modulated at a high frequency well above the flicker fusion frequency of the human eye and wavelength difference information can be used to enhance the image, revealing, for example, blood vessels that would otherwise not be visible.

A laser light source will render colors in a manner that is useful and natural for the surgeon's task. It is even possible to enhance the image, revealing important surface details that might otherwise be obscured by a broadband light source. By adding brightness to the work site, the tone scale rendering of the captured image will be improved making surface details more visible to the surgeon's eye viewing the work site remotely on a display. The most important quality of a laser light source is the great improvement in light levels it provides at the critical surgical work site. Laser light can

also enable new features when coupled with signal processing that are not possible with arc lamps. The long service life of laser lamps will reduce system operating costs thus providing a better endoscopic solution with a lower cost of ownership.

### References

- [1] Nassau, K., (1998) Fundamentals of color science, &, The fifteen causes of color. In (Nassau, K ed.) *Color for Science, Art and Technology*. Elsevier:Amsterdam, p. 1-30 & p. 123-168, ISBN 444-898468.
- [2] Newton, I. (1730) *Opticks*, 4<sup>th</sup> edition (W. Innys, London; reprinted: Dover, New York, 1952) pp 155.
- [3] Miyake, K., et el., (2005) Development of New Electronic Endoscopes Using the Spectral Images of an Internal Organ. In *13th Color Imaging Conference Final Program and Proceedings*, IS&T/SID, p. 261-263.
- [4] Maloney, L. T., (1986) Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA*, 3/10, p. 1673-1683; Maloney, L. T. (1999), Physics-based approaches to modeling surface color perception. In Gegenfurtner, K. R., & Sharpe, L. T. [Eds], *Color Vision: From Genes to Perception*. Cambridge, UK:Cambridge University Press, pp. 387-422.
- [5] Thornton, W.A., (1971) Luminosity and Color-Rendering Capability of White Light. *JOSA*, 61/9, p. 1155-1163; Brill, M.H., Finlayson, G.D., Hubel, P.M., Thornton, W.A., (1998) Prime colors and color imaging. In *6th Color Imaging Conference Final Program and Proceedings*, IS&T/SID, p. 33-42.
- [6] Shiobara, T., Zhou, S., Haneishi, H., Tsumura, N., Miyake, Y., (1996) Improved color reproduction of electronic endoscopes. *JIST*, 40/6, p. 496-501.
- [7] [http://en.wikipedia.org/wiki/Standard\\_colorimetric\\_observer#The\\_CIE\\_standard\\_observer](http://en.wikipedia.org/wiki/Standard_colorimetric_observer#The_CIE_standard_observer)
- [8] CIE 2007 functions at <http://www.cvrl.org/>; <http://www.cie.co.at/publ/abst/170-06.html>; Sharpe, L. T., Stockman, A., Jägle, H., & Nathans, J. (1999). Opsin genes, cone photopigments, color vision and colorblindness. In K. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From Genes to Perception* (pp. 3-51) Cambridge: Cambridge University Press.; Stockman, A., & Sharpe, L. T. (1999). Cone spectral sensitivities and color matching. In K. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From Genes to Perception* (pp. 53-87) Cambridge: Cambridge University Press.; Stockman, A., Sharpe, L. T., & Fach, C. C. (1999). The spectral sensitivity of the human short-wavelength cones. *Vision Research*, 39, 2901-2927; Stockman, A. and Sharpe, L.T. (2000) Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40, 1711-1737.; Stockman, A. and Sharpe, L.T. (2000) Tritanopic color matches and the middle- and long-wavelength-sensitive cone spectral sensitivities. *Vision Research*, 40, 1739-1750.; Stockman, A., Sharpe, L.T.,

Merbs, S. and Nathans, J. (2000) Spectral sensitivities of human cone visual pigments determined in vivo and in vitro. In: K. Palczewski (Ed), Vertebrate phototransduction and the visual cycle, Part B. *Methods in Enzymology*, Vol. 316, pp. 626-650. Vol. 316. Academic Press, New York.

- [9] Pugh, E.N. Jr. and Larimer, J. (1980) Test of the identity of the site of blue/yellow hue cancellation and the site of chromatic antagonism in the P-1 pathway. *Vision Research*, 20, 779-788; Larimer, J. (1981) Red/Green opponent colors equilibria measured on chromatic adapting fields: evidence for gain changes and restoring forces. *Vision Research*, 21, 501-512; Nick, J., Larimer, J. (1983) Yellow/blue cancellation on yellow fields: its relevance to the two process theory. In **Colour Vision Physiology and Psychophysics**. (Edited by Mollon, J.D. and Sharpe, L.T.), Academic Press, London; Shevell, S.K., Nick, J., Larimer, J. (1984) Color perception under chromatic adaptation: Super sensitivity with dim backgrounds. *Vision Research*, 24, 491-495; Piantanida, T., Larimer, J. (1989) The Impact of Boundaries on Color: Stabilized Image Studies, *JIST*, 15:2, 58-63.
- [10] Yarbus, A.L. (1967) *Eye Movements and Vision*. Plenum:New York.
- [11] [http://en.wikipedia.org/wiki/Color\\_rendering\\_index](http://en.wikipedia.org/wiki/Color_rendering_index); CIE (1995) *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*, Publication 13.3, Vienna: Commission Internationale de l'Eclairage, ISBN 978-3900734572
- [12] Nickerson, D., Jerome, C. W. (1965) Color rendering of light sources: CIE method of specification and its application. *Illuminating Engineering* (IESNA) 60/4, p. 262-271; Nickerson, D., (1960) Light sources and color rendering. *JOSA*, 50/1, p. 57-69.
- [13] Nayatani, Y., Takahama, K., (1972) Adequateness of using 12 metameric gray object colors in appraising the color-matching property of lamps. *JOSA*, 60/1, p. 140-143.
- [14] van Trigt, C. (1999) Color Rendering, a Reassessment. *Color Research and Application*, 24/3, p 197-206.
- [15] Schanda, J., Sándor, N., (2003) Colour rendering: past, present, future. In 'International Lighting and Colour Conference', Color Research & Application (Cape Town: Wiley Interscience) 28: 397.
- [16] van Nes, F.L., Bouman, M.A., (1967) Spatial modulation transfer in the human eye., *JOSA*, 57/3, p. 401-406.

## Appendix A: Color and Space Trade-offs in Focal Plane Sensors

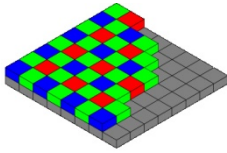


Fig. A.1. The Bayer Pattern used in single sensor camera.

Single Sensor CMOS and CCD cameras employ a Bayer Pattern, an array of color filters, that is placed over the sensor's pixels to differentially sample the spectral power density function of the light imaged onto the sensor. This is similar in concept to the eye's mosaic of cone photoreceptors which samples a single weighted function of the light at each separate retinal location.

The eye, however, moves and creates a perceptual representation of the image on the retina based upon a brief interval of sampling the visual environment. A camera, on the other hand, typically reconstructs the image based upon a single camera frame and does not attempt to analyze the image into components like chairs and people. The visual system's function is much more complex than a camera's function which is designed only to reconstruct the projected image on the sensor. The Bayer Pattern is shown in Figure A.1.

There are many color filter array patterns that are used today. The specific pattern that is optimal depends upon the modulation transfer function of the camera's optics, the size and density of sensor pixels, the pixel tessellation geometry, the number of full image frames that will be used to reconstruct the image, and the reconstruction algorithm. The most apparent issue is that every location in the image requires three values to reconstruct the color signal. This problem is referred to as up-sampling the raw data or interpolating the missing data in the raw captured image frame.

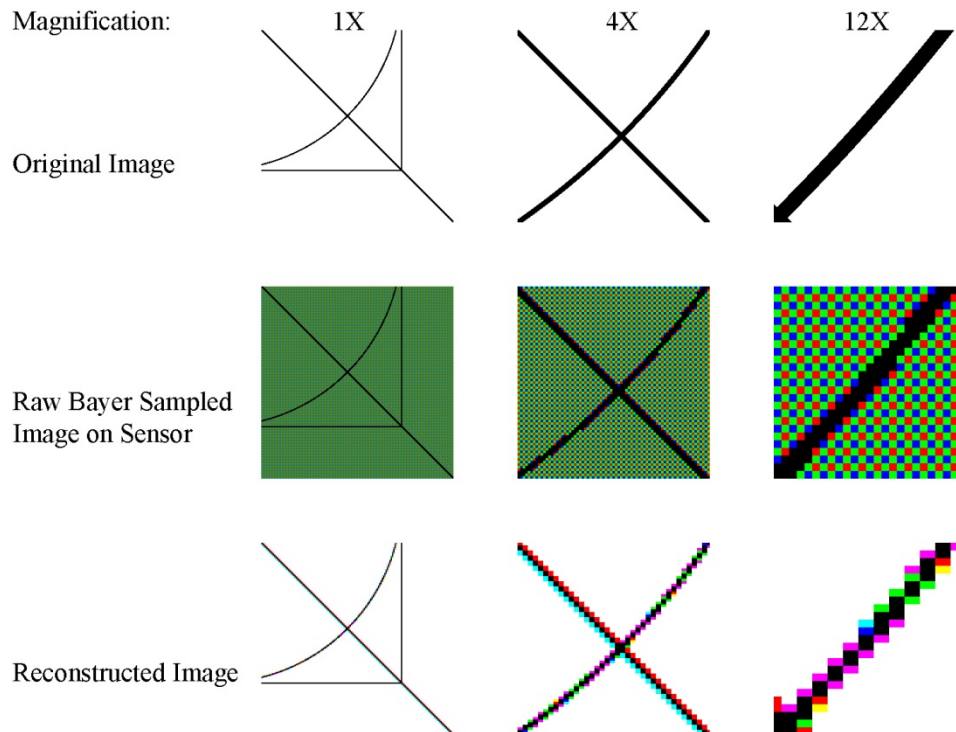


Fig. A.2. An illustration of the Bayer Pattern sampling method is shown above at three levels of magnification. The top row is the raw image. The middle row is the raw Bayer signal on the sensor. The bottom row is the image reconstruction with sampling artifacts.

Figure A.2 illustrates the sampling artifacts created by the trade-off of spatial resolution and chromatic resolution implicit in Bayer Pattern sampling, and similar to the trade-off made in biological vision systems. Unique and different colored filters, in this case red, green, and blue filters, sample each location in the image spectrally. The middle row in Figure A.2 shows what the image samples on the sensor. The sensor only knows charge levels so this row would more appropriately be displayed as grayscale corresponding to charge, but the colors were added to identify the locations of the red, green, and blue filters respectively. In this case a relatively *naïve* 4 tap up-sampling filter was used to reconstruct the missing data in each chromatic sample represented by the red, green and blue samples.

In the chromatic domain, the green spectral band is sampled at one octave below full spatial resolution, i.e., exactly half the pixels, and the red and blue spectral bands are sampled at two octaves below full resolution, i.e., one out of every four pixels. The green spectral band is near the peak of  $V_\lambda$ , the photopic luminosity function. It is the spectral region that most strongly produces signals in the eye. The eye and many cameras have a short-wavelength chromatic aberration – or blur – which effectively spatially pre-filters the signal captured by the blue filtered pixels. No spatial pre-filter is applied to the red and green filtered samples, so, as is expected, these sub-sampling losses are expressed as spatial artifacts in the reconstructed image. This is apparent in the non-smooth reconstruction of the edges shown in the bottom row of Figure A.2. Chromatic artifacts are also introduced by this sampling scheme. The reconstruction filter used in this example was purposely naïve for this illustration. It did not use either spatial correlation or correlations across spectral bands to reconstruct the full signal.

The eye is clever and uses dynamic extended duration sampling scheme wherein the eyes move during signal capture interval. A more sophisticated analysis is used to reconstruct the imagery. However, it must be noted that we experience scenes not as pixels on our retina or as images similar to a photograph. We perceive the space around us filled with objects that we recognize. We see chairs, people and objects within our visual environment and not the cone signals on our retinas or spatial combinations of these signals. Perception is a cognitive process in that we see objects that we recognize not images that require identification and recognition. Nonetheless, the Bayer Pattern example illustrates the trade-off inherent to sampling both chromatic and luminance differences in projected images regardless of whether or not that process is physical and algorithmic based on a camera or biological within a living visual sensory system. The class of artifacts, their causes and solutions are similar in both cases.

## Appendix B: A Comparison of Color Rendering with an Arc Lamp and Laser Lamp.

### Metrology Test Pattern: The MacBeth ColorChecker & Color Meter.

A Xenon arc lamp endoscopic illumination system and an RGB laser prototype endoscopic illumination system were used to illuminate a MacBeth Color Checker. A MacBeth Color Checker is a calibrated test chart, with a calibration traceable to NIST standards. It is provided to test the performance of lamps, cameras and displays with regard to producing known specific colors. The chart is constructed from a series of pigmented and/or dyed color chips with broadband reflectivities. When the chart is illuminated with a known standard light source the 1931 CIE Standard Observer  $xy$  chromaticities are determined exactly. The values for a D65 illuminant, the standard light source for television and computer monitor applications, are provided with the chart.



Fig. B.1. The calibrated MacBeth ColorChecker Chart is shown in this figure.

Display monitors used in endoscopic systems often are calibrated against the IEC sRGB Standard that is defined in terms of D65 illumination. The MacBeth Chart, therefore, provides a test pattern to measure and assess the performance of a lamp, camera, and display in terms of the chromatic distortion they introduce by illuminating, capturing or reconstructing an image of

the ColorChecker chart. The ideal values for each stage of this multi-stage process are provided with the chart. Figure B.1 shows a likeness of the MacBeth ColorChecker.

The 1931 CIE  $xy$  chromaticities for the eighteen color chips in the chart were determined by measurement with a Konica-Minolta CS-100A Luminance and Color Meter (see <http://www.konicaminolta.com/sensingusa/products/display/luminance-color-meters/cs100a/index.html>). This meter is calibrated to a NIST standard and provides  $Yxy$  values. The rendering performance of both lamps was assessed using this meter in a controlled lighting environment wherein only the illumination from one or the other source could illuminate the test chart.

### Test Procedure

The laser lamp is many times brighter than the arc lamp. This is a known advantage of a laser illumination system compared to an arc lamp system. Both lights are presumably providing D65 illumination to the chart, although no adjustment was possible with the arc lamp. The arc lamp provided a narrow beam of light that was used to illuminate the chart in the interior of a light integrator. The laser light was optically coupled to the integrator and provided an equally intense diffuse illumination adjusted to have a D65 equivalent white balance based upon the laser wavelengths and power output. The color test chips were measured through a viewing port in the side of the integrator.

The CIE  $XYZ$  tristimulus values determined by measurement with the Konica-Minolta meter were equated for luminance so that only the measured chromaticities across illumination sources could vary. The MacBeth chart ideal D65 chromaticities were converted to  $XYZ$  tristimulus values by assigning the same chip luminance values used to

compare the empirical chromaticity measurements of each color chip under the two light sources. Now only the chromaticities varied across chips when comparing the ideal target values to the actual measured values. These 1931 CIE xy chromaticities are shown in Figure B.2.

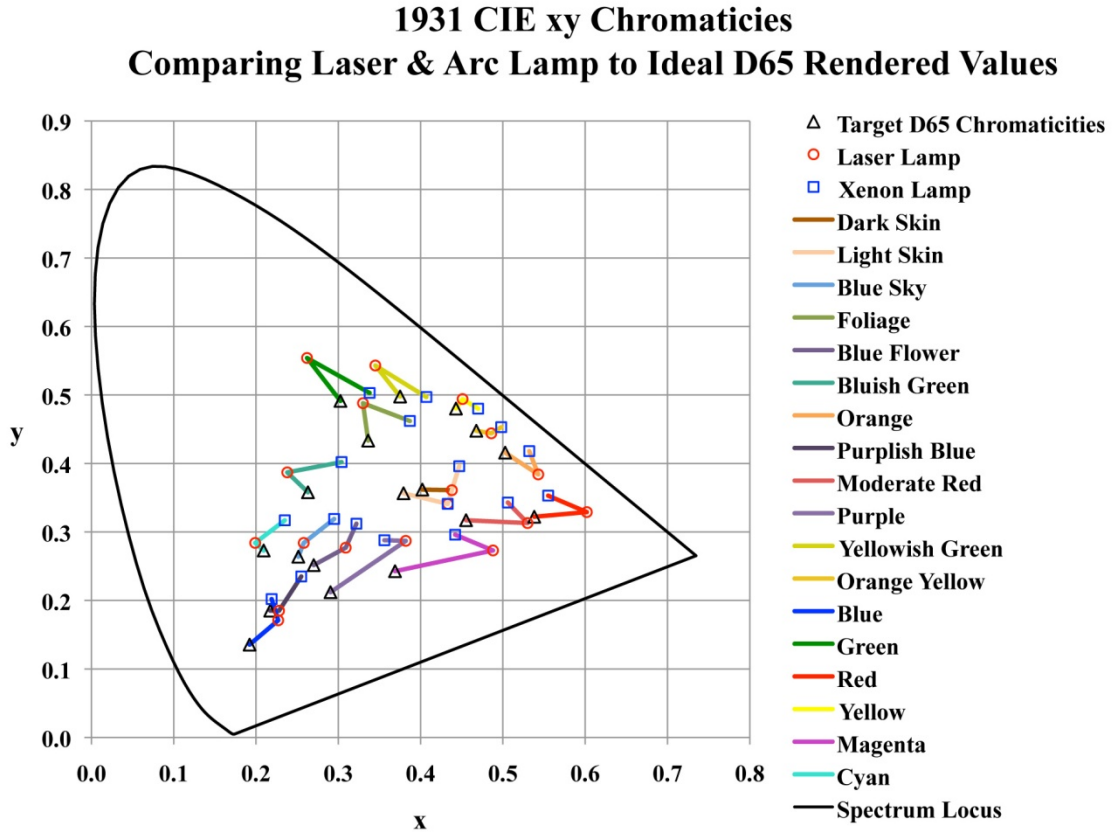


Fig. B.2. The chromaticities of the eighteen MacBeth color chips are compared in this chart. The triangles are the location of the ideal chromaticities for the color samples when illuminated by an ideal D65 illuminant. The circles are the same chips illuminated by the laser light source set to a D65 white balance. The squares are the chips illuminated by the Xenon arc lamp that is set to approximately D65 (typically 6000°K is achievable). The lines of different colors connect the measurements of the same chip under the two lamps and the ideal illuminant.

The data plotted in Figure B.2 show that neither lamp is rendering the chips according to the standard and that both lamps however are creating a broad spread of different colorations in the test chips not far in 1931 CIE xy chromaticity space from the ideal loci. As noted above, however, the XYZ tristimulus space and its projection onto the xy chromaticity plane were not designed to represent visually salient color differences. Other CIE color spaces were designed to provide a perceptually uniform color space. One of these spaces is CIE L\*a\*b\*.

The data shown in Figure B.2 were all transformed to CIE L\*a\*b\* space. This color space defines a color discriminability metric that is virtually the same measure used in the CRI. Since the measurements for each color chip under each illuminant were equilibrated in luminance, only the a\*b\* vector differences were required to assess the visibility of

rendered color differences. The CIE L\*a\*b\* discriminability measure is highly sensitive to luminance differences, so it would detect a small difference in brightness between lamps which is not the object of this assessment as the superior brightness of the laser lamp is not the issue.

The discriminability comparison made here eliminated any brightness difference leaving only color difference as a measure of the lamp performance. The a\*b\* measured value for each lamp on each of the eighteen color chips was compared directly to its corresponding ideal value and an average discriminability was computed using this method. This is similar to the scoring method used to calculate the CRI, but it directly compares the lamps to the ideal rendered values. The results of this comparison are shown in Table B.1.

Arc Lamp		
	Direct Light	Monitor
Ideal	21.22	26.89
Laser Lamp		
	Direct Light	Monitor
Ideal	20.06	18.15

The average distance or discriminability of the color chips when illuminated by the laser lamp is actually slightly better than the arc lamp. This difference is small and in this case a prototype lamp diffusely illuminating the test chips is compared to a light delivered through an endoscope instrument, so precise conclusions should not be drawn from these measures. However, this data clearly shows that the laser lamp is no worse and perhaps even more accurate in color rendering than the arc lamp endoscope. The criticism that lasers do a poor job of rendering colors is not sustainable based upon these data.

In addition to measuring the performance directly with the Konica-Minolta meter, an endoscope camera captured the colors and displayed them on a computer monitor that was also measured. The right most column in Table B.1 shows the computed average discriminability scores comparing the colors rendered on the monitor to the ideal values. Here the laser clearly out performed the Xenon lamp.