Color Vision responds to Natural Scenes: Roles of Glare, Receptor Quanta Catch, and Neural Spatial Comparisons

John McCann [1]

[1] McCann Imaging; mccanns@tiac.net

Abstract

The goal of human color vision research is to understand how we see Color. Our vision has evolved to guide us through the world's Natural Scenes. Appearances made by our color vision are the result of: optics of the eye, molecular quanta catch of receptors, and spatial image processing of neurons.

The scene in front of the lens is the first critical variable in modeling vision. The spatial distribution of radiances coming to the eye initiates the first major visual event, namely intraocular glare. Imaging the retinal image redistributes the very large dynamic range of light in complex Natural Scenes.

The second event is the receptors' response, namely, Light/Matter reactions in atoms and molecules. Color results from the different spectral sensitivities of rods and cones.

The third event is that receptors initiate the network of neural spatial comparisons that lead to Appearances. This network is stimulated by the output of all receptors in the retina simultaneously.

This talk is about the effects of scenes, glare, quanta catches, neural image processing, and Appearances. This talk introduces 8 different studies of vision that trace the light from the scene, through measurements of Appearances. It discusses Theoretical Color experiments, and practical Color technologies that respond to complex Natural Scenes.

Keywords: *Intraocular glare, Rod/Lcone Color, Appearance in Complex Scenes, Calculate Appearances, Neural Spatial Comparisons, Algorithms that mimic vision.*

INTRODUCTION

Studies of Color Vision and Light/Matter interactions are 24 centuries old. Plato and Aristotle speculated about color vision. Ancient Greeks reported light-sensitive matter. When certain substances were exposed to light, the substances changed, indicating a chemical reaction. These Light/Matter interactions are the result of photons transferring energy to atoms and molecules. This energy transfer is called quanta catch, or matter's spectral response to light. When light exposes silver-halide salt, each quanta catch converts a silver ion to a stable silver atom. All Light/Matter interactions take place in the spatial dimensions of angstroms.

Human vision captures information over a solid angle of >140°. It begins by optically imaging light on receptors in the retina. After that it sends the quanta catch information from 100,000,000 receptors to the brain, using only 1,000,000 optic nerves. From neural junctions every receptor sends its light responses to neurons. These neurons make Spatial Comparisons to make visual Appearances.

Traditional photography, movies, and analog videos are records of sensors' responses to quanta catch. Today, we talk of picture elements, or *pixels*. Cameras capture light, and record it as digital pixel values. Imaging technologies transmit the array of all those digit values to different media: print, television, and digital displays. The vast majority of imaging technologies perform this task one pixel at a time. Namely the capture, transmission, and presentation to observers works the same as a landline telephone wire connecting the sensor response with the display output. These imaging pipelines, from photon capture to display, are constant for all local image segments. Every segment is processed separately - independent of all the signals from all the other segments.

In 1961, as a Freshman at Harvard, I began a part-time Lab Tech job for Edwin Land, and Nigel Daw of Polaroid in its Vision Research Laboratory. Land's study of Red and White (Two-Color) Photography made him realize that vision is fundamentally different from photography. Light/Matter interactions are limited to a single pixel's response to light. Silver halide film responds to individual microscopic scene segments, independent of the rest of the scene. However, Land realized that visual

Appearances respond to the spatial content of the entire scene, after the receptors' responses. In the mid-1960's Land and I spent time discussing vision's response to the Natural Scene. One of the themes we discussed was "How human spatial vision worked". Could understanding human spatial vision lead to better photographs? Land was equally fascinated with both Color challenges: Theory and Practice. How could technology "mimic human vision"? Over the decades that challenge was refined to become: accurately capture light from the entire scene, calculate appearances of all scene segments (using multi-resolution spatial comparisons), and write all appearances on film, now media.

Figure 1 illustrates three stages of photography. The three photos were taken out of the same latticed window in Fox Talbot's home, Lacock Abbey, Wiltshire, UK. Figure 1A is a reproduction of Talbot's 1835 negative photo. It is the oldest surviving silver halide negative image. Figure 1B is a digital rendition of the same scene using three fixed (R,G,B) tone-scale reproduction responses to spectral light. This photo mimics the combined positive and negative responses of Kodacolor and Fujicolor prints. In other words, it mimics the pipeline process that made most color photo prints from 1940 to 2000. Each input pixel determined the value of the output pixel. All pixels used the same pipeline (Tone Scale Function). Figure 1B is a pixel-by-pixel rendition of the films' quanta catch.

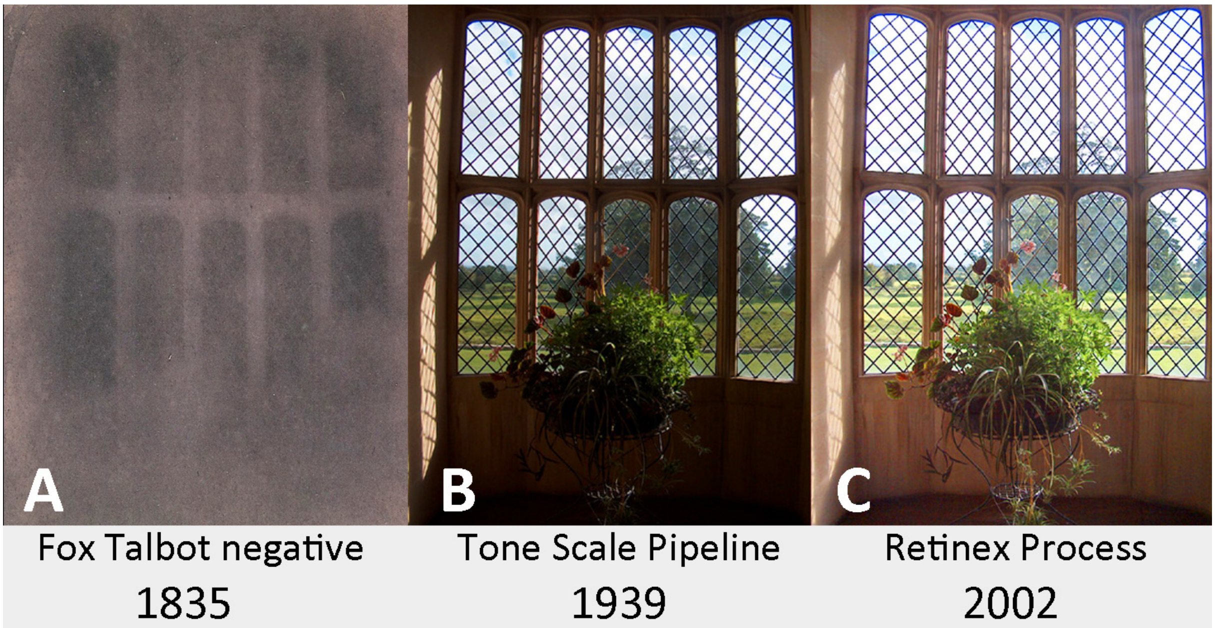

Figure 1. Three photos of Fox Talbot's window. 1A is a reproduction of Talbot's surviving negative. 1B illustrates color prints made using 3 (RGB) fixed Tone Scale Curves (eg. Kodacolor). 1C Digital digital camera image made with spatial comparisons to compress the scene's high range of light to fit the print's low range.

Figure 1C is an icon of photography that mimics human vision. It is a digital photograph using HP's *Digital Flash* image processing that incorporated Frankle and McCann's (1983) Retinex algorithm. The image used spatial comparisons to transform the scene's high-dynamic-range into a low-dynamic-range image for printing. By emphasizing edges and controlling gradients the algorithm compressed the scene's range of light to render the Appearance of the bright garden and dim interior.

Vision's response to the Natural Scene is to build Appearances out of multi-resolution spatial comparisons. These spatial comparisons tend to cancel the effects of optical glare by transforming quanta catch into a new unique rendition of the scene. Optical glare's scene transformation responds to the distribution of light in the scene. Vision counteracts glare by neural spatial processing that responds to the distribution of light on the retina**.** (McCann & Rizzi, 2012). This talk compares and contrasts vision's spatial processing with the single pixel pipelines used in photography from 1835 to 1981. Until digital image processing, photography lacked the technology to compare quanta catches from every part of the scene. As well, this talk discusses how digital photography is currently moving from Tone Scales to Spatial Processes to improve its rendition of all scenes. This talk describes 8 studies in which the receptor quanta catches after intaocular glare, and neural spatial comparisons interact with the Natural Scene, and other complex images. These studies emphasize that vision's spatial comparison network vastly expands our ability to do much more than just count quanta. Further, studies of Natural Scenes "opens our eyes" to incorrect hidden assumptions associated with single-pixel pipelines.

1. Uniform Appearance Color Space

Newton, Goethe, and painters describe many different hue circles. Munsell was a fine-art painter who taught at Mass College of Art, Boston. He also invented a uniform color space. Munsell did not use theoretical principles to define his space. Instead, he gave observers the practical task of selecting a sample that appears equidistant from two endpoints in a Natural Scene. The Munsell space we use today is a work of the first 1975 Judd Medalist, Dorothy Nickerson. Her work defined the volume of the uniformly spaced papers in today's Munsell Re-notation (Newhall, Nickerson, & Judd, 1943).

Figure 2A shows Munsell's early painting of a color space in 1900. It is an idealized color sphere with white as north pole, and black as south pole. Different hues are arranged around the equator. Nickerson's model (Figure 2B) shows the actual color space built up using equal increments in hue, lightness, and chroma. The white/black polar axis is a cube-root function of reflectance. The hue circumference is a much more complex shape. It is the result of the spectra of illumination, the atomic orbitals of dye and pigment molecules, the quanta catch of human receptors, and their post-receptor neural processing. The Color Space that has uniform spacing is far from spherical. Each hue angle has a unique lightness/ chroma population of colorants.

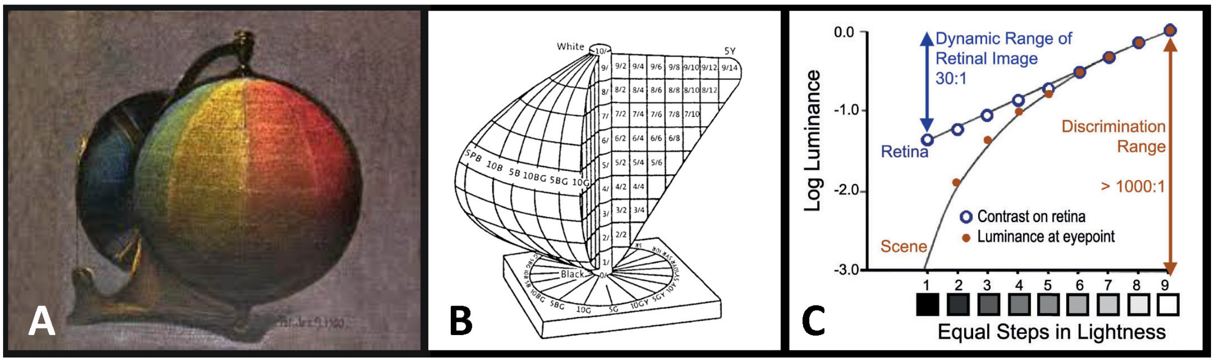

Figure 2. Three representations of Munsell's Uniform Color Space. 2A Munsell's painting of a color sphere in 1900. 2B Nickerson's plot of 1943 Re-notation. 2C Stiehl's 1983 plots of 9 HDR equal steps in Lightness (scene & retina).

The 1943 Munsell-OSA space has traded the globe's elegant simplicity for the most valuable property in digital imaging, namely a Uniform Color Space. Distances in color appearance have to be calculated in a uniform 3-D space. Digital images live in the domain of numbers, and their manipulation. The molecular physics of light sensors (silver halides and silicon chips) counts photons to make a linear scale. The average of 2 linear quanta catches is the midpoint between them. All subsequent pipeline image-processing manipulations are non-linear in radiance. In non-linear spaces ordinary arithmetic averages act differently. Nonlinear digit values need definitions and calibrations.

In the reflectance of papers, the white is ~100%, and black is ~4%; average is ~52%. Munsell asked observers to tell him the midpoint between white and black, and extended the principle to make an entire uniform color space. Munsell's middle-gray N/5.0 reflects 20% of incident photons not 52%. Photographers recognize middle-gray as the antiquated Kodak 18% Gray Card, used to set camera exposures. Photographers used 18% Gray Cards as an object at the center of the Natural Scene.

Why is middle gray 20%, not 52% reflectance?

Physiologists since the 1930's have measured retinal receptor's response to light. Neural response is proportional to log of the outer segments quanta catch (Oyster, 1999). But, Appearances in a Natural Scene do not fit a logarithmic function. Munsell's Lightness Values, CIE L*, and Savoy's HDR experiment (bisecting Lightness) fit a cube root function. Figure 2C plots log luminance in blue circles, and cube-root luminance in dark red dots from 9 equally space Lightnesses. Stiehl, Savoy, and McCann (1983) used Munsell's bisection technique to make an HDR uniform lightness scale using transparencies on a light box. Figure 2C plots the telephotometer luminance measurements of the display as dark-red dots. The data fits the cube root of scene luminance. The authors also calculated the light on the retina after intaocular glare. They used the work of the 1991 Judd Medalists Hans Vos, namely his Glare Spread Function (GSF) to calculate the receptors' actual stimulus, plotted as blue circles. The GSF converted scene luminances to retinal luminances:

- Equally-spaced Lightnesses fit log Retinal Luminance
- Retinal Receptor response fits log Retinal Luminance
- Lightness is proportional to Retinal Receptor response
- Mid-point in Receptor response = 20%; Mid-point in quanta catch = 52%

Middle-gray, Lightness step 5 in Figure 1C is 20% of the maximum retinal luminance. Above middle gray, the log and cube root functions superimpose. As scene luminances decrease below 20%, optical veiling glare adds scattered light from the rest of the scene to the darker steps. Equal Lightness steps required observers to select darker and darker transparencies in order to overcome scattered light.

[Summary: Lightness is proportional to Retinal Receptor response (after glare)]

Contrast and Assimilation

Visual illusions are a popular pursuit of color scientists. Null experiments are a favorite. How does the rest of the scene influence the appearance of identical stimuli? In Figure 3A we have 8 identical *Gray* luminances (4 circles-top and 4 crosses-bottom). On the left side *Grays* are on a uniform background, and all appear the same *Gray* lightness. On the right the four *Grays* have different backgrounds. On top-right we see *Gray* circles on the traditional *Contrast* backgrounds of black (lighter appearance), and white (darker appearance). Below the Contrast illusion, we have Assimilation by Todorović, (1997). It is scaled to fit the Contrast illusion above it. In Assimilation we see *Gray* circles behind slits in white and black foregrounds. In this spatial arrangement, the mostly white ground makes the *Gray* appear lighter, the mostly black ground makes the *Gray* appear darker.

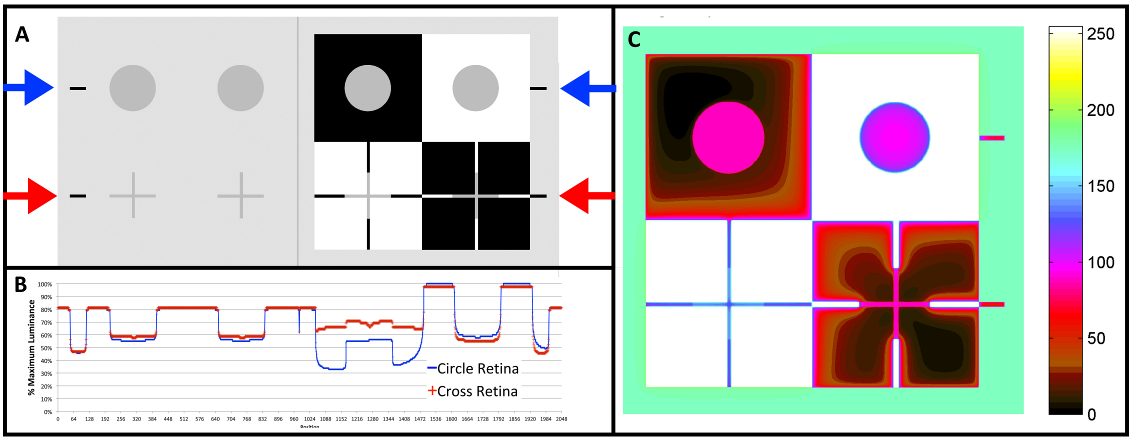

Figure 3. Contrast and Assimilation analysis. 3A Reproduction of the illusion. 3B Two horizontal plots of calculated retinal luminance (blue=Contrast / red=Assimilation). 3C Pseudocolor visualization of 3A with color map.

As scientists, how do we sort out our observations that the top Contrast-white ground makes darker *Gray* appearance; and bottom Assimilation-white makes lighter *Gray*. How does the top Contrast-black ground make lighter *Gray*; and the bottom Assimilation-black make darker gray?

The answer is that there are two different, independent spatial transformations of the light coming to our eyes. The first transformation is pre-retinal, namely intraocular glare. The pair of Contrast/Assimilation illusions has 4 identical *Grays* at the cornea (Scene luminance). They are spatially transformed by glare. The 4 *Grays* become 4 markedly different Retinal stimuli. While all *Grays* are identical at the cornea, the receptor responses are very different because of glare light.

We used the McCann and Vonikakis (2018) Matlab program to calculate the pattern of light on the retina caused by Figure 3A image (2048 x 1024 pixels; 8 bit). Figure 3A was viewed on an iMac computer screen at 30 inches, subtending 15° by 7.5°. The *digit to screen luminance* calibration was made using a Konica Minolta CS100 colorimetric telephotometer. Measurements were made using an opaque mask blocking all other light from the screen in a dark room.

The blue arrows in Figure 3A indicate the location of a horizontal digital scan of the calculated retinal image (Contrast Illusion). The blue plot in Figure 3B is the average retinal luminance of a 3 pixel high scan across the middle of the 4 gray circles in the <u>Contrast illusion</u>. The red arrows in Figure 3A indicate the location of a second horizontal digital scan. The red plot in Figure 3B is the average retinal luminance scan across the middle of the 4 *Gray* crosses in the <u>Assimilation illusion</u>.

The hidden assumption in both Figure 3A Contrast and Assimilation illusions is that there is zero intraocular glare. Comparisons of Contrast and Assimilation assume that equal digits sent to the display device causes equal retinal luminances. However, the red and blue scans of the gray scene components are different. In Figure 3B-left, *Grays* on a uniform light gray background, the crosses scan (red) have more background scatter than the circles (blue). In Figure 3B-right, Assimilation's white foreground is adjacent to the *Gray* cross, and it adds still more glare light. As well, Contrast's Black circle caused the least amount of glare in all 8 *Gray* segments. These plots shows that glare is the first transformation that distorts scene luminances. The neural phenomena that create Contrast and Assimilation begin with unequal retinal luminances. This takes the Null experiment (Equal Stimuli) argument "off the table".

The illusions in Figure 3A do not make a constant neural input for all *Grays* in the scene. Figure 3A is an abstract simulation of the Natural Scene. It is made up of uniform patches of light with perfect square-wave edges. There are no gradients in this digital image. The retinal image of 3A is different. All the sharp edges become a wide variety of different slope gradients. Figure 3C is a pseudocolor rendition of the calculated retinal digits proportional to log retinal luminance. The digital map on the right side identifies the hues used to visualize the gradients in the retinal image. Digit 255 is rendered as white, Digit 127 is gray blue green, and digit 0 is black. By rendering the retinal image in 64 discrete colors, pseudocolor breaks gradients into visible bands that allow us to visualize the shapes of the gradients. Gradients in pseudocolor are much more visible that in an achromatic rendition.

In the illusions white areas contribute more glare than they receive, They show little change. Glare transforms uniform scene blacks into a wide assortment of gradients on the retina. Figure 3C shows many different local spatial transformations of the "equal grays" in the scene. Intraocular glare upsets the null experiment; it redistributes Scene's light in the retinal image.

<u>*[Summary: Retinal Luminances ≠ Scene Luminances]*</u>

<u>*Intraocular Glare is the first Spatial Transformation of scene information]*</u>

2. Light/Matter interactions of molecular physics → Colorimetry

Color is our response to light from a scene. Colorimetry's Matches are the result of the Light/Matter interactions of molecular physics. Colorimetry's input is the spectral radiances of a spot of light. As David Wright, the 1977 Judd Medalist, pointed out "colorimetry ends once the light has been absorbed by the colour receptors in the retina and that appearance science begins as the signals from the receptors start their journey to the visual cortex" (Wright, 1987).

David Wright splits Color into two topics: Colorimetry, and Appearance. Colorimetry is a triumph of psychophysics. The spectral sensitivities of the receptors are the input to vision (Figure 4A). Light/ matter interactions lead to the color matches that David Wright contributed to Colorimetry. These matches are the result of the individual quanta catch of independent light receptors. Equal receptor quanta catches make matches that take place in the retinal rod and cone outer segments. The red ellipse at the top of the left side identifies the only site of light/matter cone interactions. The only scene radiance measurements allowed in all CIE Colorimetry calculations are the X, Y, Z values of a single small spot of light on these receptors. Colorimetry predicts matches from the spectral radiances of single *spots of light* on a no-light background. If there is no glare from surrounding scene segments, then the only glare comes from the spot itself. The red ellipse in Figure 4A illustrates Wright's *stop sign* for Colorimetry, namely the independent quanta catch in receptor outer segments (McCann, 2020).

[Summary: Colorimetry calculates equal quanta catches to predict color matches. Applying CIE X,Y,Z to Natural Scenes requires an additional calculation of the effects of intraocular glare]

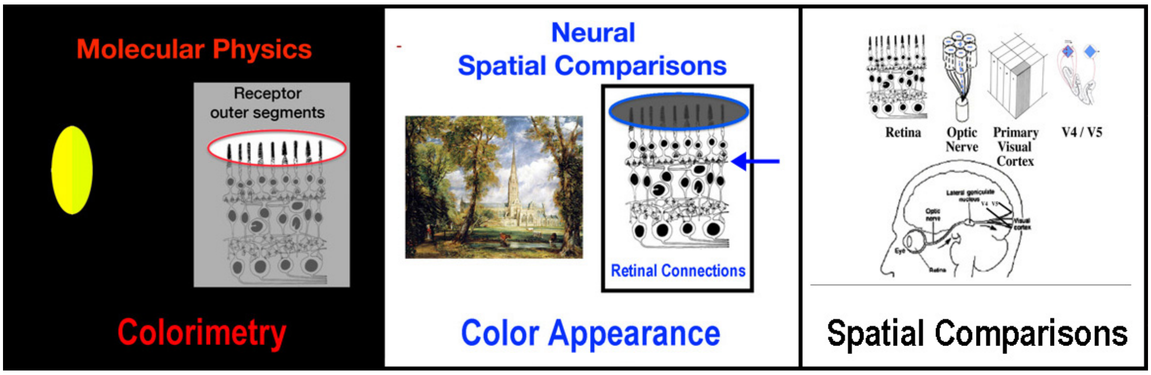

Figure 4 Wright's Colorimetry/ Appearance distinction. 4A Wright's *Colorimetry* is limited to receptors' quanta catch. 4B Color Appearance begins to renders HDR scenes using a network of neurons that make Spatial Comparisons in the retina. 4C Spatial Comparisons continue at every stage in the visual pathway.

3. Appearance: Neural Spatial Comparisons → rendition of natural HDR scenes

Figure 4B uses a scene icon, namely, John Constable's 1825 HDR painting. It illustrates the Bishop, his cathedral in sunlight, and his garden in shade. It is a record of what Constable saw. His Appearances of the scene were recorded by his painting. Unlike Colorimetry, real HDR Natural Scenes have abundant optical glare in both vision and cameras (McCann and Rizzi, 2012; McCann, Vonikakis & Rizzi, 2018). Glare transforms the HDR scene radiances into a substantially different image on the retina.(McCann and Vonikakis, 2018) While scene radiance of a single spot is the appropriate input to Colorimetry's Matches, we cannot use the single pixel radiance as the input to neural spatial processing. A spot of scene radiance does not include the substantial scattered light from all other scene segments. Neural input is the complete array of <u>all</u> receptors' quanta catch after glare, not the <u>scene radiance</u> of a single pixel.

In Figure 4B at the neural junctions (blue arrow), on the lower end of the rod and cone receptors, the study of Appearance begins. Color Appearance models use all receptors' response as input to the retina's complex spatial processing. Figure 4B is based on John Dowling's map of retinal connections. It is the start of the cascade of neural interactions that travel down the optic nerve and throughout the brain. Figure 4C shows the many different types of neural spatial comparisons, and their location along the visual pathway. It illustrates the work of Dowling, Kuffler, Barlow, Daw, Hubel, Wiesel, Devalois, Zeki, and Conway. The visual pathway is a cascade of spatial comparisons starting with receptor's synapses and continuing at every neural stage (McCann, 2020).

[Summary: <u>Neurons build Appearances using spatial comparisons]</u>

<u>[Neurons make the second Spatial Transformation of scene information (quanta catch)]</u>

4. Neural Spatial Comparisons - Edges not Light control Color Appearance

Figure 5A is a simple illustration of edges. On a white page we see a reproduction of a smooth digital gradient from 255 to 0. On top of that gradient are 9 identical rectangles, each with uniform digit 146. (Use your computer's Digital Color Meter to verify that all these rectangles are identical.) The Appearance of the 9 rectangles varies from light gray to dark gray. Scene luminance does not correlate with appearance. The uniform gray luminances have distinctly different edge ratios with the surrounding gradient. Edges, not uniform luminance, control Appearance.

Figure 5B (left) illustrates a section of Land's Black&White Mondrian - complex array of achromatic matte papers.(Land and McCann, 1971) The illumination was a smooth luminance gradient (low at top / high at bottom). Two middle gray papers are in the red rectangle. There are two points on this pair that measured 160. The upper 160 is the combination of a higher reflectance paper (right) in slightly lower illumination. The lower 160 is the combination of a lower reflectance (left) in slightly higher illumination. The luminance of the right paper is 200 in the higher illumination.

The lower change from luminance 160 to 200 is an edge caused by an edge in luminance. Edges causes substantial changes in Appearance.

The vertical change from luminance 160 to 200 is caused by a smooth gradient in luminance. Gradients cause almost invisible changes in Appearance.

[Summary: Edges, not light, control Appearance.

In Natural Scenes Edges and Gradients are critical variables, not Reflectance and Illumination.]

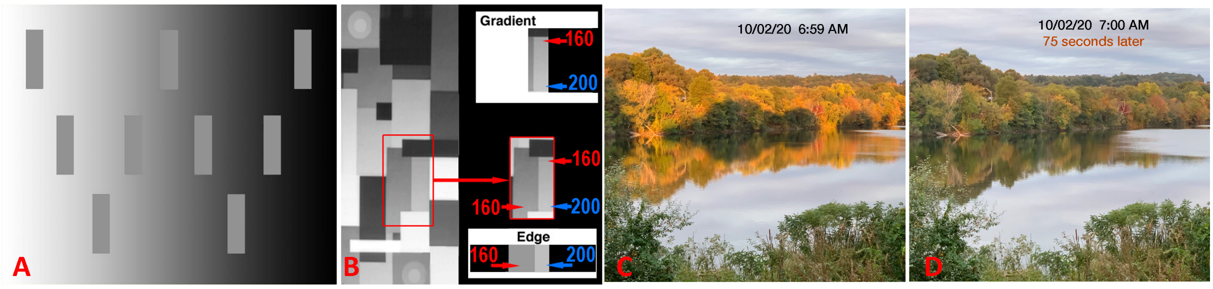

Figure 5 Edges and Gradients. 5A Nine identical gray luminances have 9 different appearances. 5B Edges cause large changes in appearances, while gradients are almost invisible. 5C Photo with edge in sunlight illumination. Vision does not discount edges in illumination. 5D Photo without illumination edge taken 75 seconds later.

Edges in Illumination

Figure 5C,D is a pair of photographs made 75 seconds apart. The photos were taken just after sunrise in the fall of 2020. In 5C the distant trees were in sunlight. A large clump of closer trees were in the shadow of a low cloud. That cloud created a sharp edge in illumination. In 75 seconds a cloud moved to block all sunlight (5D). The edge in illumination was gone. The trees in the photographs have markedly different appearances. Edges in illumination cause major changes in Appearance.

Most people pay attention to objects in the Natural Scene. Photographers study illumination. They observe the interaction of objects and illumination. We are good at recognizing its approximate luminance; sunny vs. cloudy day; sunrise, dusk, sunset. Illumination contributes to our emotional state. The two photographs in Figure 5C,D have different emotional impact, even though two-thirds of their pixels are identical. There is strong evidence that Helmholtz's *Unconscious Inference* exists. The important question, however, is whether there is any actual evidence that *Inferences* affect Appearances, rather than just being a consequence of them. In Figure 5C, observers do not discount illumination.

[Summary: Edges in Illumination generate the same visual response as Edges in Reflectance.]

5. Capturing the Appearance of multi-colored Natural Scenes

J.C. Maxwell's invention of color photography used three different R, G, B spectral photo records of the light from the scene. This invention is a most helpful process in studying Color in the Natural Scene because it allows us to experiment independently with the scenes' color information, and the spectral transfer to receptors. Experimenting directly by changing illumination on objects changes both scene information and receptor responses at the same time.

Figure 6A shows the R,G,B color separation photographic records. It lists three wavelengths used to make an additive color projection (656, 546, 450 nm illumination).

Note the six color squares added to the photograph (RGBYMC). Also note their grayscale rendition in each of the three achromatic separations. Red square is the color associated with white square in 656nm, black in 556nm, black in 450 nm. The Y square is black in the B record, and white in both G and R. Full color reproductions are possible with RGB spectral information of the entire scene, and different wavelengths of light to transmit the color information to the L-, M-, S-cones.(Fig 6A) These photo color separations are a constant record in all wavelengths. Full color photographs require different spectral information (R,G,B separation records) and different sensory color channels (L, M, S, rods). If the separation records are identical, the appearance is monochromatic. L, M, and S cones work best in roughly equal energy illumination, such as daylight.

Two- color natural scenes

We can use *spatial records*, and *channel wavelengths* to analyze natural complex colored scenes. Maxwell's color separation photographs on black & white film has frozen the spectral information from each scene segment. Full color reproductions requires three separations and three transmission wavelengths. However, two spectral separation photos, and two illumination spectra make multicolor scenes (Figure 6B). Here we have the same three wavelengths illuminating the records. However, we have used the G records in both the 556nm and 456nm light. The result is a two-color rendition of the scene that has many different hues, lightnesses, and chromas, but not all of them.

[Color Separations freeze scene's spatial information that Spectral light sends to them to LMS cones]

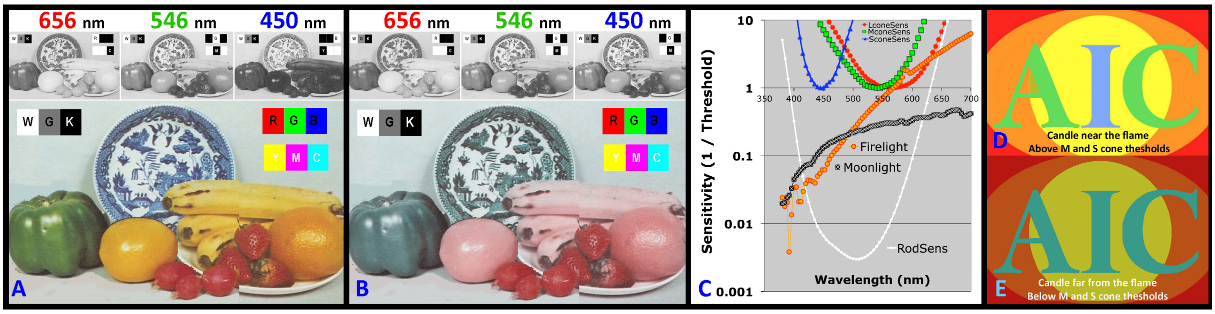

Figure 6. Three Color, two Color channels, and Color from rods and L-cones. 6A Photo mimicking Maxwell's 3 color photos using 3 records in 3 wavelengths. 6B Two channel color. 6C Spectra of receptors, moonlight and firelight. 6D Appearance in bright firelight (> M&S cone threshold). 6E Appearance in dim firelight (< M&S cone threshold).

Rods in dim firelight are color receptors (below M-& S-cone threshold)

While life on earth dates back 3.5 billion years, and vision's Opsin chemistry dates back 600 million years, our L-cones mutation is much more recent, namely, 30 to 6 million years.(McCann, 2006a) Most moonlit scenes appear achromatic. Highly colored objects appear as shades of gray without a full moon. Max Schultze (1866) proposed Duplicity Theory; simply stated, rod receptors form an achromatic channel; cones form chromatic channels. Figure 6C plots the relative light needed for absolute threshold of rods and L,M,S cones at each wavelength. Blue curve plots S-cones spectral sensitivity. The green and red curves plot M- and L-cone sensitivities. The white curve plots rod sensitivity. At 500 nm rods are 1000 times more sensitive to light than cones (McCann, 2006a).

A large collection of different experiments have shown that rods and cones interact to generate two-color vision. (For current reviews: see Stabell and Stabell (2006), and McCann, (2021:RodConeColor). These experiments use the many distinct physiological properties of rods and cones to identify whether a particular visual response is from rods, or from cones. These physiological properties include: action spectra, dark adaptation threshold, Stiles-Crawford effect, flicker fusion frequency, gray vs. color appearance of monochromatic stimuli, and apparent sharpness of the image. In these experiments observers reported seeing multicolor images when G record illumination levels were well below M- and S-cone thresholds. Using rods and L-cones makes multicolor scenes. As an example, observers adjusted the G record luminance of a multicolored (two color) image to measure the best color balance - not too warm, nor too cool. The luminance of the 656nm on the R record was constant. The measurement was repeated varying the wavelength of the G record illumination from 420 to 610nm. Observers adjusted each wavelength's luminance for best color balance. The plot of best color balance luminance vs. wavelength measured the action spectra of the color judgement. That action spectra fit the scotopic (rod) sensitivity curve. The action spectra showed that G record adjustments were made using rods alone. (McKee, McCann & Benton, 1977).

Demonstration of Rod/L-cone colors

At low light levels reflected moonlight excites the rods, but does not have enough light to excite the L, M, S cones. The spectrum of firelight is 1700°K. In Fig 6C moonlight and firelight have equal radiances at 500 nm. Firelight has nearly 10 times more radiance at 650 nm than moonlight. Dim firelight is an ideal emission spectra to excite rods and L-cones to making spatial comparisons in two color channels, while having insufficient light to excite M- and S-cones. Observers report 2 channel Color vision at low-illumination intensities in firelight.

Figure 6D illustrates the appearance of an inkjet print illuminated by a single burning candle placed close to the print. Letters **A**, **C** appear green; letter **I** appears blue. These are 3 channel colors.

By simply moving the printed AIC target away from the candle, its illuminance decreases with distance. Figure 6E illustrates the appearance of the same inkjet print illuminated by that candle far from the print. Observers report that letters **A, I** and **C** all appear cyan. When the **AIC** print is far from the candle, the color appearances are the result of 2 channel rod and L-cone color interactions. The **AIC** print in the observer's hand does not change, but the decrease in firelight took it from above L,M,S cone threshold - to below M, S cone threshold. The cyan **AIC** is from rod/L-cone color interactions. Download the **AIC** file, print it , and try it! (McCann, 2021: DemoRod-Lcone).

[Summary: The rods are a perfectly good color receptor in appropriate firelight illumination]

6. What is Vision's Response Function (VRF) to light?

We often talk about the response of the eye to light. Too often we talk about it as if we have a fixed Vision Response Function (VRF) to light, just as silver halide film does. Film catches quanta that convert ions to atoms that absorb light. That response is constant in all parts of all scenes, in all conditions. We all know that human response to light is variable. We know that pupil size is variable, and that dark and light adaptation is variable. Measurements of Human VRF ask the question: "If a spot of light has two times the luminance, how much does appearance change?" This is a trick question. The answer is that there are many distinctly different VRFs. Observers answer specific questions. Each question, such as "Can you detect this light?" gives reproducible answers across many observers. The problem is that different scenes, and different questions give a wide range of reliable answers!

Figure 7A plots the VRF for three scenes.[See inset with illustrations: 100%W (top);50%W(middle); 0%W Bottom)]. It is the plot of Lightness vs. Retinal luminance (after glare) for three different HDR transparency test targets, all with ranges close to 6 $\log_{10}$ units. The VRF of 40 gray squares in the

Black surround (0% Max luminance) is a low slope log plot over the range of 10,000:1 between Appearance White and Appearance Black. In the 100% Max luminance surround, the same range of Appearance White to Appearance Black plots over only 30:1 range of retinal luminances. In 50% Max Luminance the White to Black change plots a 100:1 range of luminances. The content of the retinal image changed the White to Black VRF of retinal luminances by a factor of 300 in range. (McCann & Rizzi, 2012; McCann and Vonikakis, 2018).

Figure 7B plots the VRF for four circular gray scales with different luminances in an opaque surround. These gray-scale transparencies have a range of 20:1. The top circle had maximum luminance; others had neutral filters that reduced luminances by factors of 10, 100, 1000. The overall target has a range of 18,909:1. Figure 7B (blue) plots VRF Lightness vs. log scene luminance. All 4 VRF Appearances have parallel slopes, each normalizes to the local maxima in each circle. The green line plots the Appearances of just the local Maxima.(McCann, 2006b; McCann & Rizzi, 2012: 113-219).

Maxima have the same slope VRF in single spots of light, in complex scenes, and in local regions of a complex scene. Neural processing assigns all maxima to the same (green line) VRF. Other nearby darker scene elements have variable slope VRF that depends on the local content of the scene.

[Summary: The Appearance of Maxima have a fixed slope VRF in all scenes.]

[The VRF Appearance of darker scene segments varies with the content of the scene.]

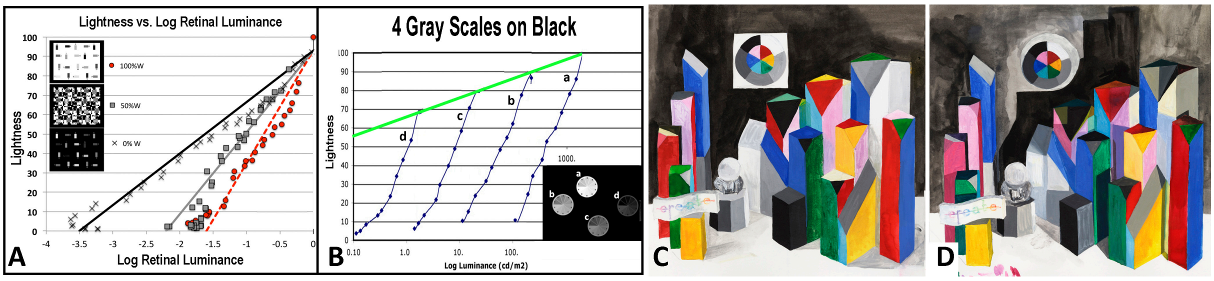

Figure 7. Studies of VRFs and 3-D Mondrians. 7A plots Lightness vs. HDR retinal Luminance in 3 displays with different surrounds. 7B plots Lightness vs. Scene luminance in 4 circles in one display. 7C & 7D shows Carinna Parraman's watercolor paintings of 3-D blocks in LDR (7C) and HDR (7D) illumination.

7. Measurements of Color Constancy with Edges in illumination

A low-dynamic-range scene can be thought of as an arrangement of papers mounted on a flat surface in perfectly uniform ambient illumination. The range of the scene is limited by the range of surface reflectances of the papers. White to Black matte papers have a range of 30:1. High-dynamic-range is the direct result of nonuniform illumination. Natural scenes, with sun and shadows, often have ranges of >3,000:1. By combining light emitters, reflective surfaces, transparencies, and shadows, one can create almost any scene dynamic range. The practical limits of image dynamic range is imposed by the optics, not scene radiances. How does expanding illumination range affect color constancy?

McCann, Parraman and Rizzi (2014) used a 3-D arrangement of blocks with 104 painted facets. All facets were painted with one of 11 paints (R, Y, G, C, B,M, W, G1, G2, G3, Bk). The experiment used two sets of identical blocks in two different illuminations: LDR illumination using a light-diffusion tent; HDR illumination with directional lights that made sharp shadow edges.

In the first experiment observers used color maps to identify all surfaces painted with one color paint. Observers used magnitude estimation to measure the departures from constancy for all red painted surfaces. Then, observers measured Constancy departures for all paint surfaces in both LDR and HDR scenes. Departures from Constancy were substantially larger in HDR illumination.

In the second experiment Carinna Parraman made watercolor paintings of the LDR (Figure 7C and HDR (Figure 7D) 3-D Mondrians. Her paints matched the appearance of every image segment. She painted them in uniform illumination on the watercolor paper. She then measured the surface reflectance of each segment to quantify their appearances. The reflectances of these matching

Appearances showed the same result. Identical painted surfaces in HDR illumination had much larger departures from perfect constancy.

[*Color Constancy is absent in single spots of light in darkness;*
strongest in flat LDR Mondrians in uniform illumination;
and much weaker in Natural 3-D scenes with edges in HDR illumination.
Prior knowledge of the surface's actual reflectance does not affect Appearances.]

8. HDR renditions

Figure 8 is an illustration of successful HDR rendition algorithms. In Figure 8A Leonardo da Vinci made Chiaroscuro paintings around 1500 that renders both objects and illumination. Figure 8B shows van Honthorst's 1620 painting of four figures at different distances from a candle, in different levels of illumination. The painter used local maxima to compress the range of the scene. Below the painting is a plot of appearances of 4 pie-shaped gray scale transparencies (described above in Figure 7C). The candle and the faces of the four people are local maxima. They darken at a very low rate (Green line). The scenes around them have a very high rate. This recent psychophysical experiment gives quantitative calibration to the Appearances that van Honthorst observed 500 years ago.

Figure 8C shows "John at Yosemite, 1981" digital print of multi-resolution calculated Appearances on film. The HDR photo was taken in Yosemite Valley. The shade of the tree had 30 times (5 stops) less illumination than the sunlight. Spot photometer readings from John's white card (in shade) were equal to ColorChecker®'s black square (in sunlight). The sun-shade scene's dynamic range was 10 stops, or 1,000:1. Color print's range cannot reproduce this scene. However, color negatives have >1,000: 1 sensitivity range. Figure 8C used color negative film to capture the scene; an Itek scanner to digitize it; and Frankle and McCann's (1983) multi-resolution Retinex image processing to calculate appearances to reduce the output range to 30:1 to print on film (McCann and Rizzi, 2012).

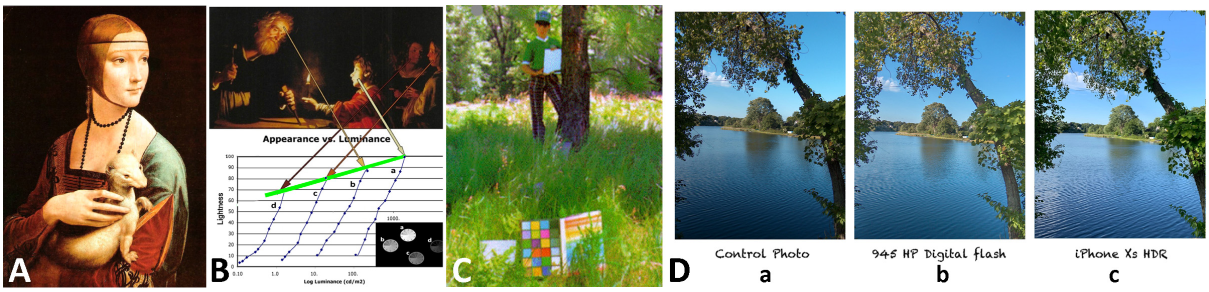

Figure 8. Successful HDR Spatial Processes: (A) daVinci painting; van Honthorst's painting with local normalization; (C) Retinex algorithm; (Da) Control; (Db) HP Retinex Camera; (Dc) Apple iPhone X.

Figure 8D shows three photos of the same scene; (left Da) is a control photo using a fixed RGB tone-scale pipeline for pixels; (Db) is an HP's Digital Flash implementation of Frankle and McCann's Retinex algorithm (1983); (Dc) is a iPhone X synthesis of local regions selected from many different exposures. By selecting the image segments with best local rendition and fusing it with other optimal renditions from different images it synthesizes the best overall rendition. All of these techniques of painting, calculating, and firmware processing are different, but they make renditions that human vision accepts as an Appearance record of the entire original scene. These very different technologies all mimic vision's response to the Natural Scene. They all render the HDR scenes in LDR images. They are all spatial transformations of scene radiances. The most important feature of Figure 8's processes is they work in all scenes, HDR and LDR. Unlike pixel pipelines, vision and multiresolution spatial comparison algorithms successfully process all scene ranges in all types of scenes. Electronic Imaging has begun to learn what painters did 6 centuries ago.

[*Electronic Imaging is beginning to mimic our color vision's interpretation*
of the Natural Scene]

Summary: Color Vision's responses to Complex Natural Scenes

Human color vision is the result of light captured by receptors, and followed by the neural processes that lead to Appearances. This talk reviews 8 different studies of complex and Natural Scenes. Traditionally vision research emphasizes the quanta catch of molecules as the first step. While necessary, molecular physics is not sufficient because of pre-retinal intraocular glare, and post-retinal neural spatial processing. These 8 studies of complex scenes expand our understanding in many ways. They include Uniform Color Space, Null visual illusions of Contrast and Assimilation, Edges in illumination, Rods as color receptors, Appearance normalization to local maxima, digital image processing, and HDR scene compression technology. Many new ideas and understandings about Color Vision have come from simply using the complex Natural Scene as the experimental stimulus. These ideas and properties of vision cannot be observed in the too-restrictive experiments of spots of light. Furthermore, the integration of multi-resolution spatial comparisons of 3 color channels has led to proven models of Appearance, and practical real-time technology implementations of Color Constancy. These successful spatial-comparison processes do not need unproven inference and illumination detection in order to calculate Appearances. Color is the response of 20th century *edges* and *gradient*, not 19th century *reflectances* and *illumination*. Spatial comparisons make the job of AI computer models of object recognition practical. When spatial processing calculates Appearances, the job of calculating spatial inferences become much, much easier.

Painters have rendered both scene and its illumination since daVinci. They do it by local normalization and controlling edges and gradients. They render the high-dynamic-range world in low-dynamic-range media. These same image processing principles are becoming obvious from the study of the spatial patterns of light on the retina, the variability of visions response to light with the contents of scenes, as well as the neurophysiology of spatial image processing in the visual pathway. These 8 distinct studies show that the contents of Natural Scenes have a major active role in Color Appearances. The contents of each Natural Scene determines the optical glare transformation that becomes the retinal quanta catch. The contents of the retinal image response determines the neural spatial transformations that becomes Appearance. While receptors count quanta, both pre-retinal glare and post-receptor spatial transformation make substantial modifications to scenes' radiances.

Too often we think of Color as a signals from a single spot of retina sent to the brain. In fact, it is the optical imaging of all the scene's light introduces glare that transforms the HDR Natural Scene radiances to a new light distribution on receptors. The retina responds simultaneously to the quanta catch of all receptors. Then, neural image processing make spatial comparisons. Glare is substantial in the Natural Scene, but spatial processing makes it hard to see. Neural vision responds to the content and arrangement of the array of receptor responses. Neurons process the spatial information, not just transmit it to undefined higher cognitive levels. Neurons respond to edges to synthesize Appearances, by compressing dynamic range, and compensating for the effects of optical glare. Unlike the HDR Natural scene, glare is inconsequential in Colorimetry's spot of light. The effects of Natural Scenes on Color Vision cannot be observed in experiments limited to a single spot of light in a dark room. Our vision evolved over the last 30 million years to extract information from all possible Natural Scenes.

Acknowledgements

My interest in the Natural Scene began with my mentors: Edwin Land, his friend Ansel Adams, and Nigel Daw. They taught me an appreciation of vision and photography, and the skills of observing, and measuring, capturing, and rendering scenes. John Dowling taught me the intricate beauty of receptors, retinal neural pathway and its dark- and light-adaptations. I particularly want to appreciate all my colleagues over the years: Jeanne Benton, Jon Frankle, Alan Stiehl, Bill Roberson, Suzanne McKee, Karen Houston, Jay Thornton, Bill Wray, Jim Burkhardt, Vassilios Vonikakis, Carinna Parraman, and Alessandro Rizzi. Their work made this talk possible. I want to thank Mary McCann for her contributions and innumerable conversations about Color.

References

Frankle, J. and J.J. McCann. 1983. Method and apparatus of lightness imaging. US Patent 4,384,336. filed Aug 29,1980; issued May 17, 1983. {"https://patents.google.com/patent/US4384336A/en"}

Land E.H. & J.J. McCann. 1971. Lightness and retinex theory. *J. Opt, Soc. Am.* 61:1-11. {"https://doi.org/10.1364/JOSA.61.000001"}

McCann, J.J. 2006a. Ideal illuminants for rod/L-cone color.Proc.SPIE:SanJose.6058 :605801-605808.<https://www.retinex2.net/Publications/ewExternalFiles/06aEI%206058-1.pdf>

McCann, J.J. 2006b. Measuring Constancy of Contrast Targets in Different Luminances in Complex Scenes. *in IS&T Color Imaging: Scottsdale, Arizona*. 14: 297-303.<https://www.retinex2.net/Publications/ewExternalFiles/06aEI%206058-1.pdf>

McCann, J.J., C. Parraman, & A. Rizzi. 2014. Reflectance, illumination, and appearance in color constancy. *Frontiers in Psychology*, 24,2014, 00005 {"http://dx.doi.org/10.3389/fpsyg.2014.00005"}

McCann, J.J. & A. Rizzi, 2012. *The Art and Science of HDR Imaging. Chichester*: IS&T Wiley.

McCann, J.J. & V. Vonikakis. 2017. Calculating Retinal Contrast from Scene Content:A Program. *Frontiers Psychology*. 2017 {"https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02079/full"}

McCann, J.J, V. Vonikakis, & A. Rizzi. 2018. HDR Scene Capture and Appearance. SPIE Spotlight Tutorial.{"http://spie.org/Publications/Book/2315540?&origin_id=x109925&SSO=1"}

McCann, J.J. 2020. What scene information is needed for Models of Color Appearance in the Natural World?. Coloration Tech,137:5-15.{"https://onlinelibrary.wiley.com/doi/pdf/10.1111/cote.12502"}

McCann, J.J. 2021:DemoRod-Lcone.{"https://www.retinex2.net/Publications/demorod-lcone.html"}

McCann, J.J. 2021:Rod-Lcone. {"https://www.retinex2.net/Publications/rod-lcone.html"}

McKee, S.P., J.J. McCann, & J.L. Benton. 1977. Color Vision from Rod and Long-Wave Cone Interactions. *Vision Research* 17:175-185.{"https://doi.org/10.1016/0042-6989(77)90080-3"}

Newhall, S., D. Nickerson, and D. B. Judd. 1943. Final Report of the OSA. Subcommittee on Spacing of the Munsell Colors, J. Opt. Soc.A. 33: 385-418. {"https://doi.org/10.1364/JOSA.33.000385"}

Oyster, C. W. 1999. *The Human Eye, Structure and Function.* Sunderland, MA: Sinauer Associates.

Schultz, M. 1866. Zur Anatomie und Physiologie der Retina. *Archiv f. mikrosk. Anatomie* **2,** 175–286.

Stabell, B. & U. Stabell. 2009. *Duplicity theory of Vision.* Cambridge: *Cambridge Un. Press.*

Stiehl, W. A., J.J. McCann & R.L. Savoy.198. Influence of intraocular scattered light on lightness-scaling experiments. J. Opt. Soc. Am. 73: 1143-1148. {"https://doi.org/10.1364/JOSA.73.001143"}

Todorović, D. 1997. Lightness and junctions. *Perception* 26: 379–394.

Wright, W. G. 1987. A plea to Edwin Land. *Color Research & Applications* 12(3):119–120