The general solution to HDR rendering

John McCann
McCann Imaging, 161 Claflin St., Belmont, MA 02474, USA

ABSTRACT

Our High-Dynamic-Range (HDR) world is the result of nonuniform illumination. We like to believe that 21st century technology makes it possible to accurately reproduce any scene. On further study, we find that scene rendition remains a best compromise. Despite all the remarkable accomplishments in digital imaging, we cannot capture and reproduce the light in the world exactly. With still further study, we find that accurate reproduction is not necessary. We need an interdisciplinary study of image making - painting, photography and image processing - to find the general solution. HDR imaging would be very confusing, without two observations that resolve many paradoxes. First, optical veiling glare, that depends on the scene content, severely limits the range of light on cameras' sensors, and on retinas. Second, the neural spatial image processing in human vision counteracts glare with variable scene dependent responses. The counter actions of these optical and neural processes shape the goals of HDR imaging. Successful HDR increases the apparent contrast of details lost in the shadows and highlights of conventional images. They change the spatial relationships by altering the local contrast of edges and gradients. The goal of HDR imaging is displaying calculated appearance, rather than accurate light reproduction. By using this strategy we can develop universal algorithms that process all images, LDR and HDR, achromatic and color, by mimicking human vision. The study of the general solution for HDR imaging incorporates painting photography, vision research, color constancy and digital image processing.

Keywords: High-Dynamic-Range (HDR) imaging, scene rendition, appearance, spatial interactions, contrast, assimilation, color constancy, image processing, veiling glare.

1. INTRODUCTION

This paper reviews many of the disciplines that contribute to High-Dynamic-Range (HDR) imaging. Painters have rendered HDR scenes for centuries. In the early days of photography, the range of outdoor scenes was a considerable challenge. Nevertheless, photography can successfully render the HDR world. Similarly, the range of early digital cameras was limited. Now we use digital techniques for better rendition of the details in whites and blacks. The limited range of reflectances in photographic prints, and luminances in electronic displays limit the possibility of accurate HDR scene reproduction. HDR renditions are the best compromise. We describe the hypothesis that the best rendition of the HDR world is calculated appearance written to display media. The HDR world is best rendered in LDR media, just as artists have done for centuries. Figure 1 illustrates the different disciplines that influence our understanding of HDR imaging. This paper reviews how each discipline contributes to HDR scene rendition.

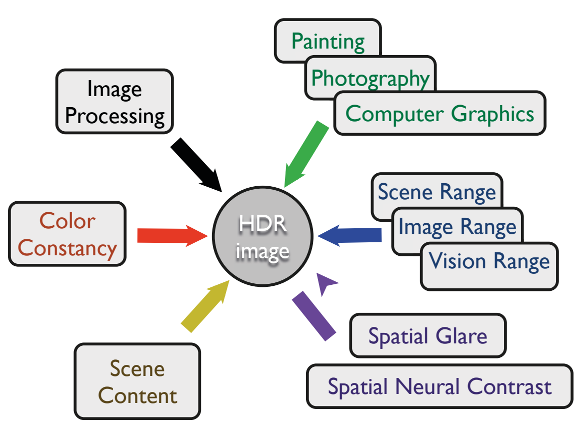


Figure 1. Photograph of raft.

Human Vision and Electronic Imaging XVII, edited by Bernice E. Rogowitz, Thrasyvoulos N. Pappas, Huib de Ridder, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8291, 829102 • © 2012 SPIE-IS&T
CCC code: 0277-786X/12/$18 • doi: 10.1117/12.914965

This paper puts forward the hypothesis that the general solution for rendering all types of scenes is to mimic human vision. The ideal rendition of every scene should be its calculated appearance (*sensation*). [1]

Figure 2. Photograph of raft.

The psychophysics of color constancy has been studied for nearly 150 years. In fact, there are a number of distinct scientific problems incorporated in the field. These studies ask observers distinctly different questions, and get answers that superficially seem to be contradictory. The computational models of color constancy for colorimetry, sensation, and perception are good examples. The Optical Society of America used a pair of definitions for sensation and perception that followed along the ideas of the Scottish philosopher Thomas Reid. Sensation is "Mode of mental functioning that is directly associated with the stimulation of the organism".[2] Perception is more complex, and involves past experience. Perception includes recognition of the object. It is helpful to compare and contrast these terms in a single image to establish our vocabulary as we progress from 18th century philosophy to 21st century spatial image processing.

Figure 2 is a photograph of a raft, -- a swimming float -- in the middle of Mascoma Lake, NH. [3,4] The photograph was taken in early morning: the sunlight fell on one face of the raft, while the skylight illuminated the other face. The sunlit side reflected about 10 times more 3000°K light than the 20,000°K skylight side. The two faces had very different radiances, and hence very different colorimetric values.

For sensations, observers selecting the colors they see from a lexicon of color samples, such as the Munsell Book or the catalog of paint samples from a hardware store. The question for observers was to find the paint sample that a fine-arts painter would use to make a realistic rendition of the scene. Observers said that a bright white with a touch of yellow looked like the sunlit side of the raft, and a light gray with a touch of blue looked like the skylit side. The answer to the sensation question was that the two faces were similar, but different.

For the perceptions, observers selected the colors from the same catalog of paint samples, but with a different question. The perception question was to find the paint sample that a house painter would use to repaint the raft the same color. Observers selected white paint. They recognized that the paint on the two sides of the raft is the same despite different illuminations. The perception question rendered the two faces identical. In summary, the raft faces are very different, or similar, or identical depending on whether the experimenter is measuring colorimetry, or sensation, or perception. Following discussion of the raft picture, Arend & Goldstein's subsequent experiments asked the same question, using a slightly different vocabulary.[5] They found the same result. Namely, observer's responses depended on the observers' psychophysical task.

We need completely different kinds of image processing in order to model these three questions. David Wright argued that colorimetry models predict only the receptor quanta catch[1:Chapter 2]; sensation models predict the color appearance that goes beyond light absorption; and perception models predict the observer's estimate of the object's surface using cognitive processes.

2. RENDERING HDR SCENES IN PAINTING, PHOTOGRAPHY, AND DIGITAL IMAGING

There is a long list of HDR imaging over 5 centuries.[6, 1:Chapter 4] The best way to categorize these examples is to organize them by rendering intent. Painters from daVinci to modern photorealists rendered HDR images so that the scene's illumination was as important as its people and objects. The rendering was a combination of both aesthetic and range compression intent. Robinson's early examples of multiple exposure silver-halide photography had the same intent as painters: Render a scene of extended luminance range to a limited one with aesthetic design.

The Mees (1920) example of multiple exposures was not so much an artistic technique, as it was a demonstration of an improvement in image quality. Mees, as director of Research at Kodak for half a century, led the development of negative films that can capture the entire range of light possible using glare limited lenses. These films captured all the possible range in single exposures.[7] Such film designs were the result of extensive photographic research, with Jones and Condit[1:Chapter 5], as an example. This work led to a single tone-scale reproduction function used in all manufacturers' color films for the second half of the twentieth century. Innumerable experiments in measuring users' print preferences led the massive amateur color print market to use a single tone-scale-system response. Even digital camera/printer systems mimicked this function.[8] It is important to note that this tone-scale function is not slope 1.0. It does not accurately reproduce the scene. It compresses the differences in luminances in both whites and blacks, enhances the mid-tones (increasing color saturation) and renders only light skin tones, and near-blacks accurately. Photographs of the identical objects (a bowl of lemons with identical chromaticities) have variable chromaticities in their standard digital and film color photographs.[9]

The Ansel Adams Zone System combined the chemical achievements of capturing wide ranges of luminances in the negative with dodging and burning to synthesize his aesthetic intent. Controlling exposure and development capture all the desired scene information. Spatial manipulations (dodging and burning) fit the captured range into the limited print range.[1:Chapter 6]

Land and McCann's 1971 Retinex algorithm [10], starting with analog electronics and quickly expanding to digital imaging, used a new approach. It incorporated the initial stage of the Mees and Adams HDR scene capture as its first step. Instead of using the Adams aesthetic rendering, it adopted the goal that image processing should mimic human visual processing. The Retinex process writes calculated visual sensations onto print film, rather than recording scene luminances.[1:Chapter 7] To this aim, Retinex renders scenes using spatial computations that take into account ratios among areas. These renditions required extremely efficient algorithms (Frankle & McCann, [11]). In computing these spatial relationships the reset step is essential to mimicking vision. It is a powerful non-linear operator that applies the equivalent of a scene dependent spatial frequency filter.[12]

Rizzi and colleagues introduced a new family of advanced of Automatic Color Equalization (ACE) that used information from the entire image so that the output values depended on scene content.[1:Chapter 33]

Stockham's spatial filtering of low-spatial frequency image content intended to combine Land's Mondrian experiments with Fergus Campbell's multi-channel spatial frequency model of vision. This concept was the basis of a great many image processing experiments and algorithms. It differs from the original Retinex algorithm because it lacks the non-linear reset, which locally normalizes images to maxima.[1:Chapter 32]

Subsequent digital processes provided methods to increased digital camera sensors range to approach that possible in negative films. McCann's scans of slide duplicating film (slope 1.0) and Mann's "Wyckoff Set" of multiple exposures had the rendering intent of better digital segmentation between max and min.[1:Chapter 7]

In 1997, Debevec and Malik [13] had a new and different rendering intent for the very old multiple-exposure technique. Using the simple, but efficient, single pixel processes they attempted to make accurate records of scene luminances. This process assumed that one can calculate scene luminance from digital camera values. This led to proposals for digital image files covering extended dynamic ranges up to 76 log units. [16: Chapter 3] It also led to the development of HDR display technology [14] with a modulated DLP projector illuminating an LCD display. This approach uses a simple tautology: namely, a display that accurately reproduces all scene radiances, must look like that scene. The weakness in the argument is that it is impractical and nearly impossible to reproduce every single pixel in the entire color space accurately. The measurements of scenes and captured scenes below demonstrate the difficulties of accurate reproduction of every scene pixel.

Table 1 summarizes the dates and rendering intents for 6 important HDR examples.

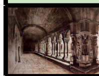
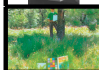
Image	Year	Author	Rendering Intent
	1851	Edouard Baldus 10 negatives	Scene Rendition
	1939	Adams Adams Zone System Dodge & Burn (spatial)	Make the Visualized Image
	1968	Land & McCann Analog Product Spatial Comparisons	Display Calculated Appearances
	1978	Frankle & McCann Digital Multi-resolution Processing	Efficient Spatial Computation
	1984	McCann Scan HDR Positive Film Single Exposure	Scan HDR Scene Capture on Film
	1997	Debevec & Malik Multiple Exposures	Capture Scene Radiances

Table 1. Important experiments in photographic and electronic HDR imaging.

2.1 Pixel vs. Spatial Image Processing

A central theme of this paper is the distinction between *pixel-based* tone-scale approaches and *spatial synthesis* of what we see.

- The *pixel-based* tone-scale approach is indifferent to the content of the image. Every pixel with an input luminance capture value of *I* has the same final output value *O*.
- The *spatial synthesis* approach is indifferent to the pixel's input capture value. Every pixel with a capture value *I* can have any output value.

Recalling the Black and White Mondrian experiment [10] is very helpful here. The display was an array of rectangular achromatic papers with a single bright light on the floor. More light fell on the papers at the bottom. Using a spot photometer, the experimental procedure was to find a black paper at the bottom with the same luminance as a white paper at the top. This was easy to do by adjusting the distances between the light and the white and black papers. The experiment asked observers about the appearances of the equal luminances. They reported that one was white, and the other was black. The fact that they have equal luminances is inconsequential to the experiment that measures what we see. The Black and White Mondrian experiment is central to understanding how HDR images can improve the reproduction of HDR scenes.

- First, it disposes of the idea that a single tone-scale function can be helpful in rendering all HDR images. Tone scale adjustments designed to improve the rendering of the black do so at the expense of the white in less illumination. As well, improvements to white make the blacks in the bright light worse. When two Mondrian areas have the same luminance, tone-scale manipulation cannot improve the rendering of both white and black.
- Second, Land & McCann, [10,15] made the case that spatial algorithms can automatically perform spatial rendering, doing what dodging and burning did to compress HDR scenes into the limited range of prints.

In order to mimic human vision, output response must be independent of the sensor's quanta catch. Early HDR algorithms [1:Chapter 7] never attempted to determine actual scene radiance, since radiance is almost irrelevant to what we see [1:Sections C, D & E]. Instead, these spatial algorithms mimicked vision by synthesizing LDR visual renditions of HDR scenes using spatial comparisons. After all, the high-dynamic range of the rods and cones feeds into the very low dynamic range of the cells in the optic nerve. The intent of Land and McCann's electronic HDR imaging was to render high-dynamic range scenes in a manner similar to human vision. They made the case for spatial comparisons as the basis of HDR rendering in the B&W Mondrian experiments.[15] The advent of electronic imaging made possible spatial manipulation of entire images. Such extensive spatial processing is not possible in silver halide chemical photography. Quanta catch at a pixel determines the system response, namely density of the image. Digital imaging processing, or its equivalent, had to be developed in order for each pixel to be able to influence each other pixel. Digital image processing unchained imaging from being bound to universally responsive pixels. Spatial interactions became technologically possible. Ironically, recent HDR tone-scale processes impose pixel-value-dependent global restrictions on digital systems. Global tone-scale functions re-chain Prometheus unchained. Lloyd Jones's work in the 1920's showed that the accuracy of their luminance record is unimportant. The spatial relationships of objects in shadows are preserved in multiple exposures. Spatial-comparison image processing has been shown to generate successful rendering of HDR scenes. Such processes make use of the improved differentiation of the scene information. Since 1968, there have been many different examples of spatial algorithms used to synthesize improved images from captured image plane luminances. [1:Section F] Efficient digital spatial algorithms, such as Frankle & McCann [11], have been used to display high-range scenes with low-range media. It provides an extensive discussion of HDR image processing. HDR imaging is successful because it preserves local spatial details. HDR has shown considerable success in experimental algorithms, and in commercial products. [1:Chapter 9]

2.2 Rendering Intent

There are four different rendering intents used extensively in HDR. They are:

- 1. An artist's aesthetic intent
- 2. Improved image quality
- 3. Calculated sensations written to film and display
- 4. Making accurate records and reproductions of scene radiances at every pixel

Calculating aesthetics is beyond our scope, but calculating an optimal HDR reproduction is not. The improvement in HDR images, as compared to conventional photography, does not correlate with accurate luminance capture and display. Accurate reproduction has never been the goal of amateur photography. The improvement in HDR images is due to better digital quantization, and the preservation of relative spatial information. Successful HDR image processing algorithms mimic processes developed by human vision, chiaroscuro painters, and early photographers. They render HDR scenes in low-range outputs accessible to human vision.

3. MEASURED DYNAMIC RANGES

There are two fundamental scientific issues in the practice of HDR image capture and reproduction. First, we need to record as wide a range of light information as possible. Second, we need to display that information in the best way for humans.

One obvious approach to accurately reproducing high-dynamic range scenes is to improve the imaging technology. Namely, if we can improve the range of the light sensors, increase the number of digital quantization levels, and improve the range of light emissions of displays, we could make accurate reproductions over a greater range of light.

Before we can assume that multiple exposures solve the problem of accurately recording HDR scenes we need to document the dynamic range over which it actually works.

3.1 Accurate scene capture by cameras

Regardless of the type of camera, digital sensor, film, and lens, HDR images have substantial optical limits. The range of light falling on the sensors is limited by veiling glare from parasitic images in glass and plastic lenses, and diffracted fog in pinhole images. Although the glare is formed by reflections in one case, and diffraction in the other, they both show limited dynamic range.[1:Section B] Accurate capture of very high-dynamic-range scenes from camera images is impossible to achieve. McCann and Rizzi [7] were unable to make accurate camera estimates of scene luminance for two 4.3 log dynamic range scenes with different surrounds. The comparison of white and black surrounds shows dramatic scene dependence. Multiple-exposure HDR is limited by veiling glare that is scene-, lens-, aperture-, and camera-body-dependent. The accuracy of scene-luminance estimates varies with all these parameters.[1:Chapter 11]

Some HDR algorithms attempt to correct for glare.[16, 1:Chapter 10] Given the characteristics of the camera, they calculate the luminances in the scene. The glare spread functions of commercial lenses fall off very rapidly with distance to a very small value. We might think that such small glare values cannot affect distant pixels. However, there are millions of pixels that contribute glare to all other pixels. The light at each pixel is the sum of scene luminance plus scattered light from all other pixels. The sum of a very large number of small contributions is a large number. Sorting out these millions of scene-dependent contributions would be required to precisely correct for glare. ISO 9358:1994 Standard [17] states unequivocally that: "the reverse (deriving luminance from camera response) calculation is not possible."

Claims are made that recent multiple-exposure HDR algorithms capture wider scene luminances, or colors than previously possible [16]. These claims are severely limited by scene and camera veiling glare. The designers of negative films selected a 4.1 log response range. That range exceeds the camera glare limit for almost all cameras and scenes. [7]

The improvement in HDR images, compared to conventional photography, does not correlate with accurate luminance capture and display. Accurate capture in a camera is not possible, and accurate rendition is not essential. Multiple exposures capture edges in the shadows. The improvement in multiple-exposure HDR images is due to better preservation of relative spatial information. Details in highlights and shadows are not lost.

3.2 Accurate scene capture by humans

McCann and Rizzi [1: Chapter 12] used the same targets for measuring cameras to measure the range of lightnesses in human vision. These targets have exactly the same scene luminance in each of the 40 segments. The 4scalesWhite target is created by removing the opaque black mask from the 4scalesBlack target. The transparent pie-shaped segments send the same light to the eye. The only physical difference at the target is the surround. In the white surround, the sector's appearances are darker in lower luminance areas than in the black surround. This is a paradox because the effect of glare from the very large white surround is to increase the retinal luminance of every sector. Despite the fact that observers have much more light on each sector on the retina, they report darker appearances.

The important point here is that we have two different visual mechanisms working against each other. They are both spatial mechanisms that involve all the pixels in the scene. One is veiling glare that increases the amount of light on the retina. The other is neural contrast from spatial interactions that makes *"more light"* look darker. Neural contrast is a collective name for post-retinal spatial interactions responsible for simultaneous contrast, assimilation, white's effect, and mach bands, color constancy, and other spatial phenomena. [1: Sections C,D,E]

4. SEPARATING GLARE FROM CONTRAST: VISION HAS TWO SCENE-DEPENDENT SPATIAL MECHANISMS

Building on the measurements of the dynamic ranges that cameras can capture, and what humans can see, we studied the scene-dependent properties of HDR imaging. The first scene-dependent property is the scatter of light in the ocular media. It reduces the contrast of edges depending upon the content of the scene. It is the direct result of the optical properties of the intraocular media, and the variable reflectances of the retinal surface. The second scene-dependent property is spatial comparison, or neural image processing. It increases the apparent contrast of edges depending upon scene content. Neural contrast is a general name for all of post-retinal image processing. Although these two spatial mechanisms have entirely different physiological causes, and they have very different spatial properties, they tend to cancel each other. The lower the physical contrast of the scene on the retina, due to scatter, the higher the apparent contrast to the observer. These two counteracting spatial mechanisms, *glare* and *neural contras*t play important roles in human vision.

Stiehl et al. [18] studied the role of scattered light in uniformly spaced appearances. They made a test target that had equally spaced steps in lightness. The asked observers to select the middle-gray transparency on a light box that was half-way between white and opaque black. Six more bisections gave nine equally spaced lightnesses. Then, they measured the luminances of these 9 steps. They found the same results as others, namely that lightness was fit by the cube root of luminance. Using the Vos and Walraven glare spread function they calculated that luminance on the retina. They found that lightness was proportional to log retinal luminance. On the retina, equal log luminance increments generate equal steps in appearance. The reason for the cube-root function is intraocular glare. Observers selected darker samples to compensate for glare. The target in front of the eye has to be cube-root to make log-luminance steps on the retina after intraocular scatter.

4.1 Measuring the Range of HDR Appearances

Rizzi, Pezzetti and McCann [31] described a set of magnitude estimation experiments for a series of pairs of test targets. In the first pair, the surround around the 40 different gray patches had a half-white / half-black surround. Observers measured the appearances of Single-Density and Double-Density targets. Their dynamic ranges were markedly different, but the half-white / half-black surround had minimal effect on glare, and did not introduce changes from Simultaneous Contrast. The measured range of useable light information for the 50% white surround target was 2.3 log units. In addition, other pairs of displays varied the surround and found different ranges of useable information. [1: Chapter 15]

4.2 Calculated Retinal Image

Rizzi and McCann[32] calculated the retinal image of *Double-Density* test targets with dynamic ranges just below 6 log units. They measured the target luminance at each pixel, and calculated the retinal contrast array using the 1999 CIE Standard glare spread function. Calculated intraocular veiling glare reduced the 6 log unit range of the test target to as little as 1.5 log units on the retina. Intraocular glare reduced the high-dynamic-range target luminances to much smaller ranges. Below middle gray, glare reduces the retinal contrast variable amounts, depending on scene content. However, the interaction of neural mechanisms with scene content tends to counteract glare. Lower actual contrast retinal images have higher apparent contrast. Their data agreed with Stiehl's, namely, that lightness is a logarithmic function of the amount of light on the retina. The slope of that function depends on the amount of white in the scene. [1: Chapter 16]

4.3 Visualizing the Retinal Image

McCann and Rizzi described pseudocolor visualizations of the HDR scenes and retinal images. They showed that veiling glare reduces a 250,000:1 scene to be very similar to a 500:1 scene. For this pair with constant glare, but very different dynamic ranges, we find nearly the same image on the retina. Almost all of the increase in scene dynamic range was removed by glare.

This study of the retinal image also shows, as one would expect, that glare increased the values of the black pixels. It also shows that small white areas have lower values because of a decrease of scattered light from nearby white pixels. Uniform patches in the scene become complex gradients on the retina. Their discussion pointed out the need for different size-dependent spatial contrast mechanisms.[1: Chapter 17]

4.4 Understanding the Separate Roles of Glare and Contrast

Both glare and neural contrast are spatial processes. The changes they generate depend on the contents of the entire scene. In order to measure the glare limits of vision, we designed a new set experiments that isolated the effects of glare from those of neural contrast. We needed to measure the effect of pre-retinal glare independent of post-retinal processing.

Intraocular glare helps to solve an important visual paradox. Uniform Color Spaces do not correlate with cone responses. Lightness correlates with the image on the retina after scatter. Equal steps in log retinal contrast generate equal steps in appearance. [1: Chapter 18, 19]

The contents of the scene determine the unique amount of glare added to each pixel. Glare is scene dependent. Neural contrast is also scene dependent, but tends to cancel glare. In the following section we measure the effects of post-retinal image processing. Spatial comparisons, instead of pixel response, are the basis of apparent lightness.

5. SCENE CONTENT CONTROLS APPEARANCE:

Traditional 19th century psychology segmented appearance into different Aperture and Object Modes, because constant stimuli generated variable responses. Vision is not limited to a single appearance from a single set of receptor quanta catches. The 19th century approach was that observers inferred a Mode from the stimuli and the Mode mechanism generated appearances appropriate for the scene content. Later, 20th century thinking showed that top-down Modes can be replaced with bottom-up spatial processing. The same data that stimulated Aperture and Object Modes can be used to measure the spatial properties of the visual process.

5.1 Appearance of Maxima

McCann measured the effects of changing the illumination on a group of gray test areas in three different surrounds. The scene maxima behaved the same regardless of the surround. The maxima matches became slowly darker with large changes in illumination. The data was consistent with observations of stellar magnitude and magnitude estimates of brightness. As well, the appearance of shadows in 3-D displays was consistent. We call this slow change of appearance of the maxima the "Hipparchus line" after the Greek astronomer who first documented stellar magnitude. [1: Chapter 21]

5.2 Contrast with Maxima

The appearances of other scene segments follow different rules. In all complex images, segments with less luminance than the maxima get darker quickly with small changes in light. The rate of getting darker varies with scene content. For the same decrement in luminance, lightness decreases most rapidly in a white surround, less rapidly in a gray surround, still less with grays on a black surround, and finally fall on the Hipparchus line with single spots of light, or local maxima.

The responses to complex images are evidence for post-receptor processing. The neural contrast mechanisms respond to scene content to calculate lightness appearances. [1: Chapter 21]

5.3 Influence of the Maxima

Lightness constancy is impressive, but not perfectly constant. As well, objects do not change in lightness as they move through real scenes, or in complex Mondrian displays. Many of the everyday changes of appearance are so small that they are essentially unnoticed, or regarded as unimportant.

5.3.1 Take away the Maximum

To make a dramatic visual demonstration we need to introduce a new maximum, or remove an old one. McCann and Rizzi [1: Section D] described a number of different ways to modify and remove maxima from complex scenes. We made modifications of the maximum by:

- varying the separation between *maximum white* and *test gray*
- varying the enclosure of *test gray* by *maximum white*
- varying the size and distribution of the *maximum white*
- varying the relative size o*f test gray* and *maximum white*
- varying the adjacency of *maximum white*

All of these parameters have significant influence on the appearance of the *Test Gray* areas. These studies reveal a great deal about the influences of the *neural contrast* mechanisms that generate our spatial vision.

5.3.2 Phenomena Names vs. Scene Content

Many of the above test targets have been studied in previous experiments, but have been described by a wide range of different names, implying that they are different phenomena, with different neural mechanisms. We have Simultaneous Contrast and Assimilation, Aperture and Object Modes, lateral inhibition and global frameworks.

McCann and Rizzi's goal was to concentrate on the data. We also wanted to present the data as simply as possible. We avoided the extreme hypotheses, such as over-simplified lateral inhibition, and over-complex perceptual frameworks. Neither the simple bottom-up, nor the perceptual top-down mechanisms can model this collection of data. Simple bottom-up hypotheses, (using local lateral inhibition, or global averages) cannot account for separation and enclosure experiments. Perceptual frameworks cannot account for asymmetric assimilation, and separation. Currently there is no single model, that can predict this collection of data and those in the literature. However, we can make a number of conclusions:

- Scene maxima have their own independent function of appearance vs. luminance.
- The smaller the extent of maxima in images, the slower the darkening of appearance with decrease in luminance.
- Maxima with small extent, large separations, and mismatched relative sizes have less influence on test grays.
- Image statistics, global averages, maximum values, Grayworld, and local averages cannot account for the set of appearance measurements.

The data continues to support the hypothesis that appearances of both *Aperture* and *Object Modes* can be translated to *Hipparchus line* behavior of maxima (spots of light) and *neural contrast* response to image content below the maxima.

There is a critical need for future research to integrate all of the above data into a comprehensive model of lightness appearance. Given a complex scene, with all sizes, shapes, and retinal contrasts we need an analytical description that predicts appearance. Such a model, that can reliably calculate appearance from *retinal radiance*, will be of great help in identifying the best image processing algorithms, as discussed below.

6. COLOR HDR

McCann and RIzzi reviewed Color Constancy in Mondrians and real scenes. Color Mondrians provide a large amount of quantitative data on spatial interactions. Many different experiments led to the following conclusions:

- Color is a spatial process, not a pixel-based process such as film.
- There are three channels that build three independent lightness images from long-, middle-, and short-wave scene content (Retinex).
- As with achromatic lightness, the maxima play a special role in each channel.
- Areas darker than the maxima show rapid changes in appearance with small radiance changes.

Experiments failed to find a significant influence of adaptation on Mondrians' appearances.

They extended the search for the influence of visual pigment adaptation in complex color images. Despite many attempts to measure it, they failed to find evidence that visual pigment adaptation plays a role in the color constancy of complex scenes. [1:Chapter 27]

6.1 Constancy's On/Off Switch

Identical stimuli everywhere in the field of view have to look the same. McCann and Rizzi [1:Chapter 28] describes experiments, proposed by Vadim Maximov, in which two sets of carefully selected papers in two different illuminations combine to make two identical visual stimuli. Although these pairs of papers are difficult to make, it is possible, if one can manufacture exactly the colored papers required. The results showed that color constancy in complex scenes can be shut off. The experiments also showed that additions to the scene, new maxima in any of the L, M, S channels, restored color constancy. Again, supporting the idea of independent channels.

6.2 Color of 3-D Mondrians - LDR/HDR Illumination

Until recently, the study of Color Mondrians has been limited to flat displays in uniform illumination. Recent experiments used two identical 3-D Mondrians made with only 11 paints. One Mondrian is in nearly-uniform (LDR illumination), the other in highly directional, mixed spectral light (HDR illumination). They are viewed in the same room at the same time. Observers measured the appearances of constant reflectance in variable illumination. They were viewed in the same room at the same time. All flat facet objects have been painted with one out of a set of 11 paints.[19.20]

We used two different techniques to measure the appearances to observers of these constant reflectance paints. First, we recorded observer magnitude estimates of change in Munsell Notation; second, we measured a Carinna Parraman's watercolor rendition of both scenes. The artist recorded the appearances she say in the water color painting. Then, she measured the reflectances of each area, thus quantifying the appearances as spectral reflectances. Both magnitude estimates and watercolor reflectances showed the same results. There is no general rule, based on illumination and reflectance, to describe the observations. Rather, we measured a great many individual departures from perfect constancy. In nearly uniform illumination more samples appear the same as *ground truth* than in complex HDR illumination. Even small departures from perfectly uniform illumination generate departures in appearances from physical reflectance. If an image-processing algorithm *discounted the illumination*, and succeeded in accurately calculating objects' *physical reflectances*, then that algorithm would not predict appearances of real-life scenes with complex non-uniform illumination. Edges and gradients in illumination behave the same as edges and gradients in reflectance.[1: Chapter 29]

6.2.1 Retinex

Retinex theory uses the entire scene as input and calculates color appearances by making spatial comparisons across the entire scene. The McCann, McKee, and Taylor [21] data show that flat Mondrian appearances correlated with Scaled Integrated Reflectances. This was measured with human cone sensitivities, in uniform illumination, and calculated integrated reflectances by spatial comparisons. The intent was to show that it was possible to calculate those integrated reflectances without ever finding the illumination. Integrated reflectances are different from physical reflectances because of crosstalk.

If one applies a spatial model to these 3-D Mondrians we would not calculate the paints' physical reflectances. Instead we would get a rendition of the scene that treated edges in illumination the same as edges in reflectance. The Retinex spatial model shows correlation with scaled integrated reflectance sometimes, (in flat Mondrians), but not all the time (in 3-D Mondrians). Retinex models of human vision calculate appearance and can be applied to all images. Incorporating a model for vision in a reproduction system provides the much needed dynamic range compression of HDR scenes into LDR media.[1: Chapter 29]

6.2.2 Colorimetry

CIELAB and CIECAM models predict appearances of single pixels by discounting the illumination. It requires measurements of the light coming from the scene, and, more important, the light falling on the scene. These models measure illumination, rather than deriving it from scene radiances. These models calculate the reflectance of the object and scales it's relation to white. CIELAB/CIECAM models measure the X, Y, Z reflectances of individual pixels and transform them into a new color space. There is nothing in the calculation that can generate different outputs from identical reflectance inputs, as frequently observed in color appearances in 3-D scenes [1: Chapter 29] and in different spectral illumination.[1: Chapter 27] These models predict the same color appearance for all blocks with the same reflectance, in every illumination. While useful in analyzing appearances of flat scenes such as printed test targets, it does not predict appearance with different spectral illumination, shadows and multiple reflections. CIE models calculate scaled reflectance and can be used to predict appearance only in one uniform illumination. Incorporating CIE models in image reproduction systems can reduce metamerism problems, and other issues regarding the response of the cones. However, they cannot predict appearances in complex HDR scenes with nonuniform, or variable spectral illumination.

6.2.3 Computer Vision

Computer Vision models work to remove the illumination measurement limitation found in CIE colorimetric standards by calculating illumination from scene data. The image processing community has adopted this approach to derive the illumination from the array of all radiances coming to the camera. Computer Vision has the specific goal of calculating the object's reflectance. The question here is whether such material recognition models have relevance to vision. If a computer vision algorithm correctly calculated cone reflectances of the flat Mondrians, then one might argue that such processes could happen in human vision(Ebner, [22]). He cites McCann, McKee and Taylor as the basis of his appearance equals reflectance assumption used throughout the book. Ebner says "Human color perception correlates with integrated reflectance (McCann et al. 1976)"[22; page 2]. The problem is that his interpretation of our work did not include spatial comparisons and crosstalk. Ebner's and colleague's *physical* integrated reflectance, is markedly different from McCann et al.'s *spatial* integrated reflectance.[1: Chapters 27]

However, the 3-D Mondrians described above, and other experiments, show that illumination affects the observers' responses[1: Chapters 27, 29]) If that same computer vision algorithm correctly calculated 3-D Mondrian physical reflectances, then these calculations are not modeling appearances. Computer Vision is a discipline distinct from human vision, with very different objectives. Computer vision calculates physical reflectances. Incorporating Computer Vision models in image reproduction systems can reduce dynamic range, but, if successful, they remove all traces of illumination. Most photographers believe that the illumination is the most important component of aesthetic rendering. By comparison, Retinex processing reduces the effects of non-uniform illumination in a scene-dependent manner, so as to have variable influence. That preserves the sensations of the illumination, rather than removing it.

6.2.4 Surface Perception

Surface Perception has the specific goal of calculating the observers' estimate of the object's reflectance. We did not ask the observer to guess the reflectance of the facets in 3-D Color Mondrian experiments. We told observers that the blocks had only 11 paints, identified in the color wheel. We asked them to estimate the appearances of the facets.

If asked, observers are likely to get very high correlations of appearance with physical reflectance in the LDR because the 11 paint samples were so different from each other. However, in the HDR illumination, we would expect that there would be more confusion, as shown by the measured appearances.

Modeling surface perception is a distinct field from measuring the appearances (sensations) in complex 3-D scenes. Since observers give different responses to the sensation and perception questions, surface perception models must have different properties from sensation (appearance) models (McCann & Houston, 1983 [3]; Arend & Goldstein, 1987 [5]). Surface perception calculates human estimates of physical reflectance. A different set of experiments is necessary to measure human estimates of reflectances. Such models are not appropriate for HDR appearance data. We asked the observers and the painter to report on the colors they saw.

Humans exhibit color constancy using scene radiances as input. The appearances they see are influenced by the spatial information in both illumination and reflectance. Measuring, or calculating physical reflectance, is insufficient as a model of visual appearances of real complex scenes. Asking observers to guess the reflectance of surfaces is a different task than asking about their appearances.

6.3 Color Constancy is HDR

McCann and Rizzi [1: Chapter 30] reviewed the literature on the only Retinex that can be completely isolated by controlling the content of the illumination. The rods allow us to observe the appearances produced by a single set of receptors.

Further, the rods are perfectly good color receptors when combined with light appropriate for the long-wave cones near their threshold. We see object constancy in HDR imaging from observations in the dark at night to bright daylight. We see the same constancy in color experiments.[33]

7. HDR IMAGE PROCESSING:

We can analyze HDR algorithms by dividing them into a number of categories. Our analysis has the goal of evaluating an algorithm's ability to render all types of scenes, the search for the general solution. If the process mimics the human visual system, then it must appropriately render LDR, HDR, low-average and high-average scenes using the same algorithm, and with exactly the same algorithm parameters. Using this as a performance standard is very helpful in evaluating the great variety of algorithms. We will review image processing algorithms that use light (quanta catch) from: one pixel, a local group of pixels, all the pixels, and all the pixels in a way that mimics vision.

7.1 One Pixel - Tone Scale Curves

It is easy and efficient to apply tone scale curves to digital images. While color negative-to-print technology installed a fixed tone scale in the factory, digital photography can modify it using the individual's personal computer. Although the practice is completely different, the principles are the same. Photographic scientists sought the best compromise for optimal rendition of all types of scenes. They measured the limits of veiling glare, and designed negatives that in single exposures captured all the dynamic range possible after glare [1:Section B]. There is no fundamental difference between film and digital tone-scale rendering. Both cannot address the issue raised by Land's Black and White Mondrian, and all other experiments with variable appearances from constant radiances. Nonuniform illumination, sun and shade, puts the same light on the camera's film plane from white areas in the shade, and black areas in the sun. HDR scene rendition that mimics human vision requires spatial image processing.

There is a great irony in using Tone Scales to improve digital imaging. Tone scales were built into the chemistry of film processing. A great deal of practical image-quality research on the scenes people photographed determined the optimal tone scale response for prints and transparencies.[1: Chapter 5] Tone scales for film photography were optimal for pixel based image processing of all scenes. Pixel-based processing had to be the best compromise. It cannot respond differently to scene content. However, digital imaging has made spatial processing possible. Unlike film, that restricted spatial interactions to a few microns, electronic imaging has freed us from the fixed compromise of a global tone scale. How ironic than some use digital image processing to recreate the limitations of film.

7.1.1 One Pixel - Histogram

A very popular tool among digital photographers is histogram equalization. Histograms that plot the number of pixels for all camera quantization levels (e.g., 0-255) are very helpful in evaluating camera exposures. Over- and under-exposure is easily recognized in histogram plots. Engineers have developed programs that group quantization levels into bins, and then redistribute the bins with the goal of having a more uniform distribution of digital values. Histogram equalization programs process the image using the pixel-value population. It assumes that the image will look better if all the bins have an equal count. Histogram equalization programs completely ignore the spatial information in the input, and significantly distort the spatial relationships in the output. Many scenes are harmed by histogram-equalization renditions because they change the spatial relationships of image content.[1: Chapter 32] Humans do not use image statistics in generating appearances; the spatial relationships of the image content control appearances.

7.1.2 One Pixel - LookUp Tables

There is a great advantage to single-pixel processing. It is fast, low-cost, and requires minimal hardware. HDR algorithms frequently use LookUp Tables (LUTs) because they require only a small number of memory locations per color to load the digital values of the output. The operation is to read the memory value out at the appropriate address. This operation is simpler and faster than multiplying two numbers. Single pixel techniques are ideal for manipulating the appearance of individual images which have a particular problem. There are many HDR guide books that can help to solve particular problems for particular scenes using multiple digital exposures, RAW format cameras and LUTs [1: Chapter 31] Although they often use the less powerful three 1-D LUT approach, they improve individual images.

Three-dimensional LUTs are an extremely powerful tool in making accurate reproductions and controlling color profiles Pugsley,1975; Kotera et al, 1978; and Abdulwahab et al.,1989 described 3-D LUT systems.[1: Chapter 31] If we think of color reproduction as the analog of moving from an old house to a new one with different size rooms. The 3-D LUT process lets one use a different strategy in each room. It is much more powerful than three 1-D LUTs. Such 3-D LUTs play a major role in the color calibration processes of digital cameras and printers, displays and color management systems (Green, 2010).

7.1.3 Using 1-D Lookup Tables for Rendering HDR Scene Captures

Kate Devlin [22] and Carlo Gatta [23] wrote extensive reviews of the many tone scale LUT's and their underlying principles. Reinhard et al. [21:Chapter 7] review the digital versions of conventional film tone-scale techniques. Although Reinhard provides many color illustrations of the processes, is difficult to compare these approaches because each of the examples uses a different test image.

Applying the same pixel processing LUT (global operators) to any, and all, scene reproductions is not realistic. Some algorithms use a two step process to incorporate models of human vision. First, they assume that multiple-exposures can measure accurate scene luminance. Second, they apply a global-tone-scale map derived from psychometric functions, measured using isolated spots of light, rather than using scenes.

Many algorithms use some kind of psychophysical data to convert scene radiance to visual appearance. As well, sometimes psychophysical data are used outside of their original context. Similarities with human vision are taken as an inspiration, rather than modeling calibrated scene input with precise implementation of the experimental data. They have been inspired by Steven's data on brightness perception (Tumblin and Rushmeier, [25]), Blackwell contrast sensitivity model (Ward, [26]), models of photopic and scotopic adaptation (Ferwerda, et al. [27]) differential nature of visual sensation (Ward Larson et al. [28]), adaptation and gaze (Scheel et al. [29]). Reinhardt and Devlin [30] used physiological data to modify uncalibrated cameras' scene capture. They used Hood's modification of Naka & Rushton's famous equation describing an electrode's extra cellular post-receptor graded S-potential response of excised fish retinas.

There are many problems with such algorithms that prevent them from accurately predicting human vision. The captured scene data is an inaccurate record of both scene luminance and retinal contrast because of scene-dependent veiling glare. [1:Section B] The psychometric functions do not apply to scenes, they model the appearance of spots of light. [1:Sections C&D]

7.2 Some of the Pixels - Local Processing

As seen above, we need to spatially modify the captured image to make the best rendition of HDR scenes. It is only through spatial processing that we can avoid the need for individual tone scales for each picture. With digital imaging all the pixels in the image can influence the output of other pixels.

The first type of spatial operators uses some of the pixels in the input image. Starting with Leonardo da Vinci, and continuing with 19^{th} century psychologists, we know that the appearance of a test area changes, depending on the surround area around it. We can call this class of models Local Processing mechanisms. This is a non-linear transform, so that all the pixels with the same input value need not have the same output value. Although surround models are popular, too often the proposed mechanisms do not have enough specific information to be able to create an image processing model to verify them. Some framework oriented psychologists (See Gilchrist[34]) describe top-down models using the presence of a surround, rather than the specifics (size, shape, and distances of influence) needed in image processing techniques. Although local spatial operators, using parts of the images, are slower and require greater hardware processing power than LUTs, they are not a challenge to today's computer power.

7.3 All of the Pixels

Following World War II, radar, optical and imaging researchers studied the world in the frequency domain. We can call this class of model spatial frequency mechanisms. The practical use of spatial frequency filtering techniques was helped by the Fast Fourier Transform (FFT) invented by Coole & Tukey in 1965. Without it, the millions of pixels in today's images would make these techniques extremely slow. They are in wide use today with standard computer processing power.

A good example of such a process is: taking an image, converting to its Fourier transform, applying a spatial frequency filter, retransforming the filtered information back to the image domain. The re-transformed image is the output of the model. This, too, is a non-linear transform so that all the pixels with the same input value need not have the same output value. As with tone-scale maps, spatial filtering can substantially improve the appearance of images. However, a given spatial frequency filter has a fixed effect on the image in the Fourier domain. Human vision behavior can be described with a given spatial-frequency filter for a particular image. However, a single filter cannot accurately describe the human appearances of all images. Human vision, in effect, uses scene dependent spatial filters.[12]

7.4 All Pixels & Scene Dependent - The Retinex Extended Family

The fourth type of spatial operations respond to the information in the individual input image. Whereas a spatial frequency filter will apply the same non-linear process to all images, this class of process applies a different modification to the image depending on the image content. We can call this class of model the Retinex Extended Family. It includes image processing algorithms that are modifications and extensions of the original Retinex, and other different algorithms that calculate appearance of an image from the array of all scene radiances. [1:Chapter 32]

The Retinex Extended Family has three unique attributes:

- It uses all the pixels in the image as input
- It is influenced by content of the image
- It attempts to mimic human vision, or make the best rendition for humans

The best HDR imaging is done with processing that render the scene the way that humans do; i.e., mimic human vision. When a scene has gray areas surrounded by white, those areas appear dark. When a scene has minimal white, the same gray areas appear lighter. The Retinex Extended Family alter the input values depending on the content of the scene. This type of spatial operation responds to the information in the individual input image. Whereas a spatial frequency filter will apply the same non-linear process to all images, the Retinex class of process applies a different modification to the image depending on the image content. Retinex Extended Family do not have a predefined behavior in the frequency domain.

The first HDR computer and electronic implementation (Land & McCann, 1971) made the bold assumption that the model should use all the input pixels and mimic vision.[10] It was a great computational challenge for computers in 1965. But, the goal in Land and McCann was to understand human spatial processes, and build their equivalent in special hardware. Frankle and McCann is a good example. Here, integer summations replaced floating point multiplication. These summations were performed for the whole image in one cycle time of the machine. Special purpose hardware that mimics vision can make the difficult challenge of Retinex Extended Family approaches fast and low cost.

7.4.1 Retinex Algorithms

The judgement of scene rendition is heavily influenced by the calibration and the firmware/hardware imaging chain. The time to process and hardware cost also play a big role in evaluating the merits of the different processes. The Goal of the calculation varies as well. One goal is to model vision. A different, but related goal is to make an imaging device that mimics vision. If the camera acts like vision, it will make a better picture. That was the mixed goal that Land and McCann set out to do in the early 1960's; namely, to study both humans and cameras in parallel and learn things from each that can benefit the other.

7.4.1 Reflectance and Illumination

As well, over the years there has been a dispersion of the goals of the Retinex calculation. Recall that the last sentence in the Land & McCann "Lightness and Retinex Theory" paper reads:

> "Whereas the function of colorimetry is to classify reflectances into categories with similar visual properties, the function of retinex theory is to tell how the eye can ascertain reflectance in a field in which the illumination is unknowable and the reflectance is unknown." [10]

It is always a bit of a shock when a colleague states unequivocally that the goal of Retinex is to use spatial information to find the illumination, so as to be able to discount it, so as to find the objects' physical reflectances. Unfortunately, more than a few have made this modification, or misrepresentation, of the original goal. The original intend of Retinex was to ignore the illumination, just the way that human vision does. The goal of mimicking human vision requires that the model departs from calculating reflectance when human appearances differ from matching the physical measurements of reflectance.

By describing how vision calculates appearances, we can formulate image calculations that render captured scenes mimicking vision. The idea of finding the reflectance of objects goes back to the 19^{th} century, and as we saw in the 3-D Mondrian experiments, reflectance does a poor job of predicting appearance in complex illumination. Further, the act of removing shadows for scenes is a curious anomaly. It does not mimic vision in which shadows are clearly visible, and it makes very dull photographs.

7.4.2 The Collection of Retinex Algorithms

All the Retinex algorithms share common goals, but there are many variations and differences.

They all use spatial comparisons to synthesize a new image from the array of sensor responses to light. The goal of that output image is to mimic vision.

As long as there are three independent spatial comparisons, the process can account for Color Constancy.

The components of the original Land and McCann model for lightness were *Ratio, Threshold, Product, Reset and Average*. The initial image (OldProduct) and the array of input radiances were processed to make the output (NewProduct).

The Ratio Product is the fundamental component that gets beyond the limitations of colorimetry. It assumes that human vision is a spatial mechanism that responds to the entire array of input radiances. Dynamic range compression, color constancy, and most of the phenomena we study in vision, are easy to understand as spatial neural processes.

Very early in our studies we concluded that human spatial processing was neither local, nor global. It had mixed properties of both.

In HDR imaging, Tone Scale processes applies a global mapping that is helpful to many individual images. In the study of spatial HDR processing, the vast majority of work has studied algorithms that make better pictures by preserving edges and manipulating less visible gradients. Compression of gradients can be the result of thresholds, resets, local averages, local sampling and ratio limits. They all work: they just require different parameters for the same amount of dynamic range compression. Some processes are more powerful than others. Table 2 list the names and dates of the major Retinex algorithms, their spatial operators and their initialization and reset properties. See McCann and Rizzi Chapter 32 for a detailed review.[1] These variations in lightness models have developed in the past 40 years.

Some models use computers to render the images, while others use special hardware for increased efficiency. Some labs used carefully calibrated images; others just grabbed stuff off the web. The selection of a favorite Retinex process depends on all of the above.

Retinex	Properties	Initialization
Land & McCann (1971)	Ratio, Threshold, Product, Reset, Average	White
Frankle & McCann (1983)	Multi-resolution (Pixel separation & Zoom pyramid) Ratio Threshold Product Reset Average	White
Land Designator (1986)	Center - Surround (Sampled Pixels) No reset	None
NASA (1999) Kotera (2002)	Multi-scale Designator No reset	None
Gamut Retinex McCann, (1999)	Ratio, Product, Reset, Average	Best Small Gamut
Milano Brownian Path (1993) & Random Spray(2005)	Average ratio (pixel / maximum of selected pixels) Reset	None
Sobol (2002)	Ratio, Clip, Product, Reset, Average	Scene input
ACE & RACE (Ch 33 - not Retinex)	Distance weighted ratio Grayworld scaling	None

Table 2. Major Retinex models along with their properties and initial pixel values.

All of these algorithms sample the input radiances in such a way as to have the locality of a pixel influence it. Locality goes beyond the local surround, and is less than the entire image. This locality sampling is the key feature that makes human vision, and it models, have scene dependent response. We have a vast number algorithms that apply locality influence and better pictures. However, we do not have a clear, data driven, understanding of human spatial processing. We have a great many approaches, but we do not have a way to integrate them into a general solution.

The issue going forward should not be who has the best looking, fastest, lowest cost, or mathematically elegant algorithm for a favorite scene. The important issue is that we understand the general solution to rendering all scenes that include, HDR, normal and LDR inputs. Vision has the unusual property that it does not compress LDR scenes, while compressing the dark regions of HDR scenes. Applying the same Retinex algorithm to both LDR and HDR scenes is a good test of a models effectiveness.

8. SUMMARY

HDR imaging is defined by its physical limits. Whereas the range of light in the natural world is extremely large, all images of that world have a considerably smaller dynamic range. At first glance, this may seem to be a disappointing twist to a popular notion that smart sensors, low-cost electronic image processing, and hybrid display media could expand the dynamic range of imaging. They do not. However, the study of the spatial limitation of optical glare led us to appreciate the magnitude of neural spatial processing counteracting glare. HDR imaging works very well. The reasons have more to do with the synthesis of LDR renditions of HDR scenes. When these renditions use algorithms that mimic the spatial processing of human vision, we find remarkable improvements in photography. So, rather than thinking that HDR research led to a disappointment, in fact, it led to a much more interesting narrative. Modern image processing has successfully begun the development of signal processing that mimics our extremely sophisticated human scene processing. This spatial processing explains the human responses to both LDR and HDR scenes.

The two spatial processes that offset each other are the key to understanding HDR's success. Scene-dependent spatial glare lowers the contrast of retinal images of predominantly white scenes. Spatial neural-contrast image processing renders predominantly white scenes with higher apparent contrast.

We have assumed that the general solution for rendering all scenes is that the Retinex process should mimic vision. Unfortunately, we still do not know enough about its spatial mechanisms. We have evidence that the maximum radiances normalize the scene, but local maxima play and important role. We know the distance from, and enclosure by, the maxima are important.[1: Section D] We know that the size of the maxima relative to the size of the test are hints at separate spatial frequency, or size based, channels. Experiments using Mondrians and assimilation targets also show importance of maxima in independence of spectral channels. We know the results of many different spatial experiments, but the general theory that unites them all still remains a challenge.

In the original Land & McCann Retinex the iterative calculations starts at white. Every pixel, except the maximum radiance, gets darker with processing. The amount darker depended on the image content, the length of the path, or the number of iterations, i.e, the spatial locality influence. The goal was to find the best description of locality to mimic vision. The idea of iterating to completion was never an option. As we learned that this process simulates vision, we put forth the hypothesis that HDR rendition should calculate sensation, and write that to film.[35] That strategy remains relevant using today's imaging technology. Spatial HDR image processing algorithms mimic processes developed by human vision, by chiaroscuro painters, and by early photographers that render HDR scenes in low-range outputs.

9. ACKNOWLEDGEMENTS

The author appreciates the help of Alessandro Rizzi and Mary McCann in this work.

REFERENCES

[1] McCann J. J., and Rizzi, A., [The Art and Science of HDR Imaging], Wiley, Chichester, (2012).
[2] Committee on Colorimetry, Optical Society of America [The Science of Color], Crowell, New York, (1953).
[3] McCann J. J. & Houston K .L., "Color Sensation, Color Perception and Mathematical Models of Color Vision", in Colour Vision. ed. Mollon & Sharpe, Academic Press, London, 535-44 (1983).
[4] McCann J. J., "Simultaneous Contrast and Color Constancy: Signatures of Human Image Processing," ed. Davis, Oxford University Press, USA, 87-101 (2000).
[5] Arend, L. & Goldstein, R., "Simultaneous Constancy, Lightness and Brightness," J. Opt. Soc. Am., 4, 2281-85 (1987).
[6] McCann J. J., " Art, science, and appearance in HDR images," J. Soc. Info. Display, 709-719 (2007).
[7] McCann J. J., and Rizzi, A., "Camera and Visual Veiling Glare in HDR images," J. Soc. Info. Display, 721-730 (2007).
[8] McCann J. J., "Color imaging systems and color theory: past, present, and future," Proc SPIE, 3299 62-82 (1998).
[9] McCann J. J, and Vonikakis, V., "Accurate Information vs. Looks Good: Scientific vs. Preferred Rendering,"Proc EVCP, Amsterdam, in preparation, (2012).
[10] Land, E. H., and McCann, J. J., "Lightness and Retinex Theory," J. Opt. Soc. Am. 61, 1-11 (1971).
[11] Frankle & McCann,Frankle, J. and McCann J.J.,"Method and Apparatus of Lightness Imaging," U.S. Patent, 4,384,336 filed 7/29/1980 issued 5/17/1983, (1983).
[12] McCann J. J., "High-dynamic-range scene Compression in Humans," in Human Vision and Electronic Imaging XI, Proc.SPIE, 6057, 6057 07-11 (2006).

[13] Debevec P., and Malik P., (1997) "Recovering High Dynamic Range Radiance Maps from Photographs," ACM SIGGRAPH'97, 369-78 (1997).
[14] Seetzen H., Heidrich H., Stuerzlinger W., Ward G., Whitehead L., Trentacoste M., Ghosh A., and Vorozcovs A., "High Dynamic Range Display Systems," ACM Trans Graphics, 760-68 (2004).
[15] McCann J. J., "Capturing a Black Cat in Shade: Past and Present of Retinex Color Appearance Models," J. Electronic Imaging, 13, 36-47 (2004).
[16] Reinhard E., Ward G., Pattanaik S. and Debevec P., [High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting], Elsevier, Morgan Kaufmann, Amsterdam, (2006).
[17] ISO 9358:1994 Standard, "Optics and Optical Instrument: Veiling Glare of Image Forming Systems.Definitions and Methods of Measurement", International Organization for Standardization, Geneva (1994).
[18] Stiehl W. A., McCann J. J., and Savoy R.L., "Influence of Intraocular Scattered Light on Lightness Scaling Experiments," J. opt. Soc. Am. 73, 1143 (1983).
[19] Parraman C. E., McCann, J. J., and Rizzi, A., "Artist's Colour Rendering of HDR Scenes in 3-D Mondrian Colour-Constancy Experiments," Proc. SPIE, San Jose, 7528-1 (2010).
[20] McCann, J. J., Parraman, C. E. and Rizzi, A., "Pixel and spatial mechanisms of color constancy", Proc. SPIE, San Jose, 7528-2 (2010).
[21] McCann J.J., McKee S. P., and Taylor T., "Quantitative Studies in Retinex Theory: A Comparison Between Theoretical Predictions and Observer Responses to 'Color Mondrian' Experiments," Vision Res. 16, 445-58 (1976).
[22] Ebner M., [Color Constancy], Wiley, Chichester, (2007).
[23] Devlin K.," A Review of Tone Reproduction Techniques," Bristol: Technical Report CSTR 02-005, (2002).
[24] Gatta C., "Human Visual System Color Perception Models and Applications to Computer Graphics," PhD thesis, Università degli Studi di Milano, Dottorato di Ricerca in Informatica, XVIII ciclo, (2006).
[25] Tumblin J. and Rushmeier H., "Tone Reproduction for Realistic Images," IEEE Computer Graphics and Applications, 42-48 (1993).
[26] Ward,G., "A Contrast based factor for luminance display", in Graphics Gems II, Academic Press, Boston, 80-3 (1994).
[27] Fewerda J. A., Pattanik S., Ferwerda J., Shirley P. and Greenburg D., "A Model Visual Adaptation for Realistic Image Synthesis," Proc. ACM SIGGRAPH 98, 249-258 (1996).
[28] Larson G., Rushmeier H., and Piatko C., "Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes," IEEE Trans Visualization & Computer Graphics, 291-06 (1997).
[29] Scheel A., Stamminger M., and Seidel H., "Tone reproduction for interactive walkthroughs," Computer Graphics Forum, 301–312 (2000).
[30]Reinhardt, E., and Devlin K., "Dynamic Range Reduction Inspired by Photoreceptor Physiology," IEEE Trans. & Comp. Graphics, 11, 13-24 (2004).
[31] Rizzi, A., Pezzetti M., and McCann J. J., " Glare-limited Appearances in HDR Images," IS&T/SID Color Imaging Conference, 15, 293-298 (2007).
[32] Rizzi A.,and McCann J. J., "Glare-limited appearances in HDR images," J. Soc. Info. Display 17, 3-12 (2009).
[33] McCann, J. J., "Appearance at the low-radiance end of HDR vision: Achromatic & Chromatic", Proc. IS&T/SID Color Imaging Conference, 19, 223-228 (2011).
[34] Gilchrist, A., [Seeing Black and White], Oxford University Press, Oxford, (2006).
[35] McCann, J.J., "Calculated Color Sensations applied to Color Image Reproduction," Proc. SPIE, 901 205-214 (1988).

.