This is a draft of an invited paper submitted to the Journal of the Society of Information Display, 2007
It is the first of a pair of our papers in that issue.

McCann, J. J. (2007) 'Art Science and Appearance in HDR images', *J. Soc. Information Display*, vol. 15(9), in press.
McCann, J. J. and Rizzi, A (2007) 'Camera and visual veiling glare in HDR images', *J. Soc. Information Display,* vol. 15(9), in press.

ART, SCIENCE AND APPEARANCE IN HDR

John J. McCann
McCann Imaging, Belmont, MA 02478, USA

Keywords: HDR imaging, veiling glare, calibration of multiple exposure techniques, spatial algorithms, Retinex, ACE

ABSTRACT

High Dynamic Range (HDR) image capture and display has become an important engineering topic. The discipline of reproducing scenes with a high range of luminances has a five -century history that includes painting, photography, electronic imaging and image processing. HDR images are superior to conventional images. There are two fundamental scientific issues that control HDR image capture and reproduction. The first is the range of information that can be measured using different techniques. The second is the range of image information that can be utilized by humans. Optical veiling glare severely limits the range of luminance that can be captured and seen. It is the improved quantization of digital data and the preservation of the scene's spatial information that causes the improvement in quality in HDR reproductions.

Keywords: HDR Imaging, veiling glare, tone-scale maps, human dynamic range, spatial algorithms, Retinex, ACE.

1. INRODUCTION

This paper is the first of a pair of our papers on HDR imaging in this issue. This paper reviews the long history of HDR scene capture from Renaissance paintings to modern digital imaging. The second paper[1] measures the effects of veiling glare on camera image capture and the magnitude estimates between white and black of scenes viewed by humans.

This paper defines HDR imaging as the reproduction of scenes with a high range of luminances. More specifically, these scenes have a greater range of luminances than reflected or emitted from the reproduction media. This definition includes a great many scenes, because print and emissive display minima are controlled by their ambient surface reflections. Even if the display does not emit light at a pixel, the room light reflected from the surface of the display is that pixels minimal luminance. A black in a print may have an optical density of 2.2, measured with a 0/45 densitometer, but a density of 1.5 with a spot photometer. Also, this definition has relevance to human vision. Although retinal receptors have a dynamic range of more than 10 log units, optic nerves that transmit the retinal response to the cortex has a range of only 2 log units. Vision makes a low dynamic range representation of HDR scenes.

Examples of HDR images in painting date back to the Renaissance. Examples in photography date back 160 years. Examples of electronic image processing date back to analog imaging. This paper reviews the evolution of many approaches, including multiple exposures, tone-scale functions and spatial image processing in rendering HDR scenes. As well, this paper reviews HDR issues concerning scene capture and human viewing of reproductions. Cameras can capture scene luminances. Optical veiling glare limits the dynamic range of luminances on the camera image plane. Human intraocular scatter transforms scene luminances to a much lower range of retinal luminances. Both capture and viewing are major issues in understanding HDR reproduction.[1]

There must be reasons, other than accurate luminance, that explain the improvement in HDR images. The multiple exposure technique significantly improves digital quantization. The improved quantization allows displays to present better spatial information to humans. When human vision looks at high-dynamic range displays, it processes scenes using spatial comparisons.[1]

2. HISTORY OF HDR IMAGING

For many centuries the arts portrayed people and objects in a style that rendered the subject without the need to accurately reproduce the surrounding scene and lighting environment. That changed in the Renaissance, first with Brunelleschi's perspective for rendering the geometry of the spatial environment, and later with chiaroscuro for rendering the high-dynamic-range illumination environment.

2.1 Painters' Rendering

Pre-renaissance paintings render people and scenes in uniform illumination. Leonardo da Vinci is credited with the introduction of chiaroscuro, the painting of light and dark.[2] His portraits of Benois Madonna, (1478) and Lady with an Ermine, (1483-1490) capture the illumination as well as the figures. One sees that the illumination comes from a particular direction and that there are highlights and shadows. Caravaggio's paintings, such as *The Musicians*, (1595-6), portray people and illumination with equal importance. In turn Caravaggio influenced several Dutch painter, among them Gerrit van Honthorst (Figure 1).

Figure 1 shows van Honthorst's 1620 painting "The Childhood of Christ". The boy holding the candle has the lightest face. The father, further from the light, is darker. The other children, progressively further from the light are slightly darker.

Rembrandt's, *Night Watch*, (1642) is an almost life size painting (363x437 cm) of a military company receiving orders to march. It is known for its effective use of light and shadow, and perceived motion.

In all these paintings the depicted light range is greater than the physical dynamic range rendered by the oil painting technique. They are just a few examples of many examples of HDR scenes that were rendering by painters in the low-dynamic-range of reflective paints.

2.2 Photographic Rendering

With the growth of photography in the mid 19th century HDR scenes presented a severe problem for films available at that time. Multiple exposure techniques for rendering HDR scenes go back to the earliest days of negative–positive photography. H.P. Robinson's (1858) composite print "Fading Away" was made using five differently exposed negatives.[3] This dramatic still life was staged using actors.

Figure 2 shows H.P. Robinson's 1858 photographic print "Fading Away" made from 5 combined negatives.

Robinson's techniques were empirical. Later in the 1870's and 1880's Hurter and Driffield established the field of photographic sensitometry.[4] They measured the sensitivity function of silver halide films. C. K Mees repeated this work in his thesis at University College London.[5]

Figure 3 shows Mees's combined enlargement from two negatives.

Mees's "The Fundamentals of Photography" (1920), shows an example of a print made with multiple negatives with different exposures (Figure 3).[6]

Over the years the science of silver-halide imaging improved rapidly. Mees established standards for high-dynamic range image capture on the negative, and high-slope rendering on prints.[7] Negatives are designed to capture all the information in any scene. The negative response function changes very slowly with change in exposure (low-slope film). This property translates into the ability to capture a wide range of scene luminances. Further, it is relaxes the requirements for cameras to make accurate film exposures. Once the scene is captured and the negative is developed, the final print can be made under optimal conditions at the photofinishing facility. The print paper has a rapid S-shaped nonlinear response to light (high slope). The resulting positive print is higher in contrast than the original scene. The loss of scene detail occurs in this high-contrast print rendering.

In 1939 Ansel Adams first described the zone system for photographic exposure, development and printing. It described three sets of procedures: first, for measuring scene radiances; second, for controlling negative exposure to capture the entire scene range, and third, spatial control of exposure to render the high-range negative into the low-range print.[8]

Adams used a spot photometer to measure the luminances of image segments and assigned them to zones in the scale from white to black in the final photographic print. The zone system imposed the discipline of visualizing the final image by assigning image segments to different final print zones prior to exposing the negative. Adams was a professional performing pianist. He often described the negative as the analog of the musical score and the print as the performance. It was essential that the negative recorded all the information in the scene and that the printing process rendered this information in the print.

Photographic contrast is the rate of change of density vs. exposure. In the negative, the low-zone values are controlled by exposure, and the high-zone values are controlled by development and exposure. The zone system provided the necessary information to select appropriate exposure and processing for each scene's dynamic range.

The final stage was to control the amount of exposure for each local part of the image (dodge and burn) to render all the desired information from a high dynamic range scene into a low-dynamic range print. This process starts with a preliminary test print using uniform exposure. Examination of the print identifies the areas with overexposed whites and underexposed blacks that have lost spatial detail. Dodging refers to holding back exposure from areas that are too dark. Less exposure lightens this local region of the negative-acting print paper and gives better rendition in the blacks. Burning refers to locally increasing the exposure to make an area darker. This is a local spatial manipulation of the image. Not only can these techniques preserve detail in high and low exposures, they can be used to assign a desired tone value to any scene element.

Adams described the dodging and burning process in detail for many of his most famous images.[9] He executed remarkable control in being able to reproducibly manipulate his printing exposures so that the final print was a record of his visualization of his desired image, not a simple record of the radiances from the scene. In fact, Ansel's Zone System process was the 1940's equivalent of an *all-chemical* Photoshop™.

The Jones and Condit study of 128 outdoor scenes provided two important benchmarks in HDR imaging.[10] First, it compared photographic images (measured camera luminances) and spot photometer (Luckiesh and Taylor)[11] measurements from scenes. It characterized these images into 3 homogeneous groups of scene: Illumination characteristics (sunlit, haze, light clouds, heavy clouds), Viewing aspects (front, cross and back lighting) and Spatial distributions (distant, remote, near-by, close-up). In all there are 17 classifications. The minimum dynamic range ratio was 27:1; the maximum was 750:1; and the average was 160:1. These measurements did not include specular highlights.

Second, and more relevant here, Jones and Condit devoted a significant portion of the paper to the careful analysis of flare in the image falling on the camera's image plane. They calculated a flare factor for each image. Intentionally, they used a Bausch and Lomb Vila Protar 8-element lens cemented into two-components in a Speed Graphic camera with a 5x7 inch image size to reduce flare. They explained that small image sizes in typical cameras (by 1940's standards) had a much larger glare problem. By comparison, using 1/4 inch electronic sensors (cell phones), and 9 element zoom lenses (digital cameras) generates even more flare than measured in the Jones and Condit study.

Jones and Condit also measured the film's limit of response to light. At some maximal exposure the film stops getting darker with increases in exposure. As well, at some minimal exposure the film stops getting less dense with decreases in exposure. The limit of film response to no light exposure is called the fog level of the film. This is the equivalent of the various noise limits in blacks in CCD and CMOS sensors. Jones and Condit showed that the fog limit was significantly lower than the camera flare limit. Although many papers discuss digital camera noise limits, few discuss flare limits. Flare, not sensor signal-to-noise ratios of noise limits, sets the usable dynamic range of cameras,

In 1961, Charles Wyckoff [12,13] developed a multi-layered color film with different ASA panchromatic sensitivities (speeds) in each set of color forming layers. The film was used to photograph the detonation of nuclear explosions. The bottom, yellow color separation, layer had ASA=2 and the top, cyan layer ASA = 600. This made a single exposure pseudocolor film able to measure dynamic range of 8 log units.

In 1978, A. Adams describes a pre-exposure technique that adds low-level uniform exposure to a negative to increase the total exposure of darkest area above response threshold.[14] The intentional addition of fog increases visual discrimination between blacks in the negative and the print. Although adding fog may seem counterintuitive, it dramatically improves image rendition by providing spatial details. This technique decreases the actual dynamic range to increase the apparent range.

2.3 Spatial Vision

Over the past century psychophysical and physiological experiments have provide overwhelming evidence that vision is a result of spatial processing of receptor information. Hecht and others showed that threshold detection mechanism use pools of retinal receptors.[15] Rod and cone receptors respond to light over a dynamic range of over 10 billion : 1. That is the range of radiances from snow on a mountaintop to the half-dozen photons needed for a dark-adapted observer to say he saw the light. In 1953 Kuffler[16] and Barlow[17] showed that the signal traveling down the optic nerve has spatial-opponent signal processing. In one example, the center of the cell's field of view is excited by light (more spikes per second). The receptors in the surround of the cell's field of view are inhibited by light (fewer spikes per second). The net result is the cell does not respond to uniform light across its field of view and is highly stimulated by edges. It has the greatest response to a white spot in a black surround. Hartline and Ratliff (1958) found spatial processing in the compound eye of Limulus Polyphemus.[18] Dowling[19] showed pre- and post-synaptic behavior of the retina establishing post-receptor spatial interactions in mammals.

In 1963 Land proposed his Retinex theory[20], asserting that these cone types act as sets, where the response was determined by their spatial interactions. The phenomenon of color constancy is best explained by independent long-, middle-, and short-wave spatial interactions. Zeki found color constant cells in V4 with predicted spatial properties.[21] Hubel and Wiesel studied the organization of the primary visual cortex's response to stimuli projected on a screen in front of the animal.[22] In each small region of the cortex they found a three-dimensional array of different representations of the visual field. Each segment of the visual field has columns of cortical cells that report on the left-eye image next to a column for the right-eye image. The cells perpendicular to the left/right eye columns respond to bars of different orientations. The third dimension has cells with different retinal size segments of the field of view. Campbell and colleagues showed that there are independent spatial-frequency channels corresponding to bar detectors of different visual angle.[23] J. J. Gibson, the noted Cornell psychologist, described the importance of bottom-up spatial image processing.[24]

The 1960's provided a decade of new evidence that human vision had mechanisms using spatial comparisons. These physiological and psychophysical experiments provided a background for making algorithms that mimic human vision.

2.4 Analog Electronic Rendering

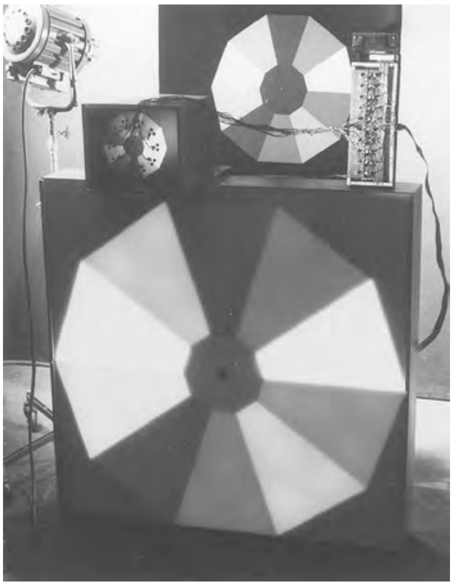

Figure 4 shows Land's Retinex analog image processing demonstration, using spatial comparisons.

In 1967, Land demonstrated the first electronic (analog) HDR rendering in his Ives Medal Address to the Optical Society of America. (Figure 4)[25,26] Here, the intent was to render HDR images using spatial comparisons that mimic human vision. This paper took the ideas of Hans Wallach[27] that suggested that lightness correlated with spatial ratios and expanded it beyond the restraints of uniform illumination. The idea was that what we see was synthesized from the ratio at an edge multiplied by the ratio at all other edges. This process synthesized an image based on the relationship of all edges in the scene, independent of the luminances of each.[26] The history of the development of this idea is found in Land's Friday Evening Discourse to the Royal Institution, London.[28]

Land's Ives Medal Address included the Black and White Mondrian. This experiment showed that two areas with identical luminances generated the white and black lightnesses in the same scene, at the same time. [26] As discussed below, this is an important test image for HDR rendering.

The pair of Land and McCann patents[29,30] described analog embodiments of calculating matches from arrays of radiances. The second patent introduced the idea of non-linear reset to the maxima, that is critical in distinguishing Retinex processing from subsequent spatial filtering. Human vision normalizes to local maxima, rendering them as white or near white.[31] This property of vision is modeled by the reset to maximum in Retinex algorithms.

Tom Stockham saw one of Land's frequent lectures demonstrating the Black and White Mondrian experiment at MIT. He became interested in the application of Fergus Campbell's and Arthur Ginsburg's[32] spatial frequency approach to imaging. In 1972 Stockham wrote a paper on rendering high-dynamic-range scenes using a low-spatial frequency filter to compress the image.[33] This, along with Hugh Wilson's spatial-frequency models[34] and Marr[35] & Horn's[36] work on gradients, became the foundation of substantial interest in spatial frequency techniques of rending images. All of the models using low-spatial frequency filters assign a specific set of filters to emulate vision. These filters are then applied to all images. Human vision is different. Human vision uses the equivalent of scene-dependent spatial filtering. The reset in Retinex has been shown to generate the equivalent of a scene-dependent spatial frequency filter. Human vision and the reset-Retinex algorithm with fixed model parameters both generate scene-dependent rendering.[37]

2.5 Digital Electronic Rendering

The actual practical embodiment of the principles articulated by Land and McCann needed two technological developments: first, the digital image processing hardware, and second, an efficient algorithmic concept that reduced the enormous number of pixel to pixel comparisons to a practical few, enabling rapid image synthesis. The hardware became commercially available in the early 1970's for the display of digital satellite and medical images. The efficient image processing began with the Frankle and McCann's patent[38] using I^2S image processing hardware with multiresolution software. The explanation of this work and its relation to other multiresolution and pyramid processing is found in the literature.[39]

Figure 5 shows an example of an HDR scene processed with spatial comparisons. (1978 Frankle and McCann patent application). The illumination on the white card in the shadow is 1/32th that on the black square in the sun. Both the white card in shade and black square in sun have the same luminance. The spatial processing converted equal input digits (~log luminance) into very different output digits, thus rendering the HDR scene into the small range of the reflective print shown here.

Figure 5 shows an example of a very efficient digital, multi-resolution HDR algorithm, using spatial-comparisons first shown in the Annual Meeting of Society of Photographic Scientists and Engineers in 1984. Here, spot photometer readings show that the illumination in the sunlit foreground is 32 times brighter than in the shade under the tree. That means that the sunlit black square has the same scene luminance as the white card in the shade. Prints cannot reproduce 32:1 in sun, plus 32:1 in shade, (dynamic range $32^{\wedge 2}$) because the entire print range is only 32:1 in ambient light. Using the spatial comparison algorithms described in detail by Frankle and McCann[38], it is possible to synthesize a new 32:1 image that is a close estimate of what we see.

In 1984 & 85 McCann described HDR image capture[40] using low-slope film, a graphic-arts scanner and digital image processing in Siggraph courses. These results agreed with those of Jones and Condit. Typical sun and shade images had a range of roughly 3.0 log units. New in this study was the effect of changes in spectral composition of illumination. Tungsten light, without shadows showed a 3.0 log unit range because the long-wave light was significantly greater than the short–wave light. Conversely, in skylight the short-wave light was significantly greater than long-wave light. Both showed a 3.0 log unit range, but this was due to spectral shifts in illumination composition, not due to sun and shade.

In 1985 Ochi et al.[41] of Sony patented a multiple-exposure CCD system using one imaging lens, a beam splitter, and two CCD image regions. An object of the invention was to "produce a still image of excellent quality and which prevents deterioration of the image quality which results from smearing and blooming."

In 1987, Alston et al.[42] of Polaroid patented an electronic imaging camera for substantially expanded dynamic exposure range by combining two successive exposures with different durations.

In 1993, S. Mann of MIT investigated a series of different digital image fusion techniques.[43] One of these techniques was to merge different exposures to capture a wider dynamic range. This work was expanded in a further paper with R. W. Picard[44]. They called the collection of multiple exposures a "*Wyckoff set*" in honor of Charlie. Their concept was to use multiple digital frames to make "undigital" images with floating point precision. They wrote: "Double precision (64bit) floating point number is close to analog in spirit and intent". For them the problem was the loss of image quality from each successive image-processing step, leading to gaps in histograms.

In 1997, Debevec and Malik used multiple exposures and least-square fits to solve for the camera response function and the luminance of each pixel in the scene.[45] Although some people have described its use of multiple exposures as revolutionary, in fact this paper is most significant because it asserted that camera digit can be used to calculate to scene luminance. There are a great many papers based on Debevec and Malik. Reviews are found in the literature.[46,47,48] Many of these papers discuss how to find an optimal tone-scale function for luminance reproduction.

3.0 RENDERING INTENT

This paper, so far, has described a long list of HDR imaging over 5 centuries. The best way to categorize these examples is to organize them by rendering intent. Painters from da Vinci to Rembrandt rendered HDR scenes so that the illumination was as important as the 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 a greater range of luminances than possible on camera image planes[1] for the vast majority of scenes. This film design was the result of extensive photographic research, with Jones and Condit as an example[10]. This work led to a *single tone-scale reproduction function*[49] 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.[50] Even digital camera/printer systems mimicked this function.[51] 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 luminances in both whites and blacks, enhances the mid-tones (increased color saturation) and only renders skin tones accurately.

Ansel Adams Zone System combined the chemical achievements of capturing wide ranges of luminances in the negative with dodging and burning to synthesize Adam's aesthetic intent. Controlling exposure and development capture all the desired scene information. Spatial manipulations (dodging and burning) fit the captured range to the limited print range.

Land and McCann's Retinex, starting with analog electronics and quickly expanding to digital imagery, used a new approach. It assumed the initial stage of Mees's and Adam's wide range information capture for its first stage. Instead of using Adam's aesthetic rendering, it adopted the goal that image processing should mimic human visual processing. The Retinex process writes calculated visual sensations on print film, rather than a record of scene luminances.[52] To this aim Retinex substitute the original pixel luminance values with the results of a spatial computation that take into account ratios among areas. 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. [37]

Stockham's spatial filtering of low-spatial frequency image content intended to combine 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 and image-dependent spatial processing.

Subsequent digital processes (Ochi and Alston) provided methods to increased digital camera sensors range to approach that possible in negative films. Mann's *Wyckoff set* of multiple exposures had the rendering intent of better digital segmentation between max and min.

Debevec and Malik and related papers had a new and different rendering intent. Accurately record the scene luminance. This led to proposals for digital image files covering extended dynamic ranges up to 76 log units.[48] It also led to the development of Brightside technology[53] with a modulated DLP projector illuminating an LCD display. This raised the luminance level of display whites. Raising the luminance of white increases the display's range between white and ambient black. By increasing the range of luminances of the display one can make use of the extended range from HDR capture. There is a simple tautology, namely a display that accurately reproduces all scene radiances must look like a scene.

4.0 OPTICAL LIMITS OF HDR

The problem with the accurate luminance tautology is that cameras cannot capture the luminance information from all scene pixels. As shown by Jones and Condit flare is a more significant limit than fog or S/N limits. Multiple exposures can help to minimize S/N, but they are powerless to reduce veiling glare. A greater problem is that human optics limit range of luminances on the retina more than glass lenses.

4.1 HDR Scene Capture

The best way to determine the dynamic range of a camera is to measure it. The accuracy of a scene luminance calculation needs to be compared with a test target with measured luminance test areas.

The second paper in this pair describes an extended series of measurements of actual and calculated luminance. The differences between actual and camera-based estimates of luminance measures veiling glare. Glare restricts the accuracy calculated scene luminance.[1]

ISO9358:1994 Standard, “ Veiling glare of image forming systems”[54] defines veiling glare and describes two standard techniques for measuring it. It describes how veiling glare is the sum of individual stray light contributions to a single pixel from all other light from the scene, even from light beyond the field of view of the lens. Stray light reflected from lens surfaces, sensor surfaces and camera walls cause veiling glare. The ISO standard defines the glare spread function (GSF), which is a measure of the amount of stray light as a function of angle from a small very bright light source. Veiling glare is measured by ISO9358:1994 as the fraction of a very large white surround scattered into a small opaque central spot. For commercial camera lenses veiling glare values are in the range of 1 to 10 %, depending on the lens and the aperture.

In the second paper in this journal, McCann and Rizzi[1] measured the dynamic range of images on a variety of camera image planes. They measured the range captured vs. the actual scene luminances using typical digital, film and pinhole cameras. The test target had a scene range of 4.3 log units. These measurements were made with a spot photometer in a dark room, one image sector at a time. The rest of the scene was covered with opaque black paper. This procedure insures that veiling glare in the spot photometer optics does not affect calibration.

Measurements using multiple exposures of a scene with 20 to 1 luminance range showed that commercial negative film can capture a 4.1 log unit range of luminance.

McCann and Rizzi used their low-glare test target (See Table 1: top third column) with 4.3 log unit luminance range in an opaque surround to measure the range of their 35mm-camera/negative-film/scanner combination for the *4scaleBlack* target.[1] The results showed an accurate 3.5 log unit rendition of the scene, using a single exposure. The reduction in range from 4.1 to 3.5 was due to glare. The results from the same target showed an accurate range of 2.8 log units using multiple exposures with a digital camera.

In other experiments, McCann and Rizzi used their maximum-glare test target (See Table 1: top fourth column) with 4.3 log unit luminance range in a white surround to measure the range of the *4scaleWhite* target[1]. The results showed an accurate range of 2.4 log units, using a single exposure. The reduction in range from 4.1 for the negative to 2.4 was due to increased glare. The results from the same target showed an accurate range of 1.6 log units using multiple exposures with a digital camera.

The 40 scene luminances (Range 4.3 log units) had camera image plane ranges between 2.4 and 3.5 log units for 35mm camera/film/scanner; camera image plane ranges between 2.8 and 1.6 for a digital camera; and camera image plane ranges between 3.2 and 2.4 for pinhole/ film /scanner (Table 1).

			
Camera	Sensor Range	Min Flare	Max Flare
Negative	**4.1**	**3.5**	**2.4**
Digital	**2.9**	**2.8**	**1.6**
Pinhole	**4.1**	**3.2**	**2.4**

Table 1 lists the measured dynamic ranges of three camera systems for two test targets. Both targets have the same 40 calibrated luminance areas covering 20% of the target. Min Glare has 80% opaque area (Top row 3rd column). Max Glare has 80% maximum-luminance area (Top row 4th column). In all cases: maximum flare reduces the dynamic range on the image plane to 2.4 log units, or less; minimum glare for this scene is 3.5 log units, or less. All possible surrounds around the 20% area covered with luminance patches will fall between these values.

These are measurements of specific cameras, lenses and scenes. They cannot be generalized to other cameras, or other scenes. Nevertheless, the results demonstrate that multiple exposures cannot measure accurate scene luminance beyond these glare limits, described by Jones and Condit and ISO ISO9358. Image plane luminance is scene, lens, aperture, exposure and camera-size dependent.[1]

4.2 Human Vision and HDR

The second fundamental HDR issue concerns the range of display that is detectable by human vision. Vision adjusts itself to respond to sunlit snow scenes at mountain top, and to half a dozen photons in prolonged darkness (a range greater than $10^{\wedge 10}$). That remarkable range is possible because of physiologically different receptors with adaptive sensitivities, and adaptive pupil size. However, the range of visual responses within a scene involves these mechanisms and spatial contrast and intraocular veiling glare. In other words, human vision has the same ocular flare issues and sensor range issues as cameras. In addition, it has physiological spatial-image-processing contrast mechanisms that change the matches of a particular luminance depending on other luminances in the field of view.

Lightnesses in complex scenes are the result of two powerful mechanisms. The first is physical, namely, optical glare. The second is physiological, namely, spatial contrast. The universal example of simultaneous contrast is a white square next to a black square with two identical gray patches in the center of each. The gray appears darker surrounded by white. The gray has less intraocular glare when surrounded by black, and hence lower retinal luminance. Physiological contrast makes the higher luminance gray in a white surround look darker.[1]

Starting with a dark night sky with bright stars we find a minimum of both intraocular glare and physiological contrast. As we change the scene's composition towards gray spots in a completely white surround, we increase the amount of glare (lowering retinal contrast) and we observe an increase in physiological contrast (increasing apparent differences). Spatial contrast acts as an antidote to glare.[1]

In 1983, Stiehl et al.[55] described an algorithm that calculated the luminance of each pixel in an image on the human retina after intraocular scatter. It calculated the luminance at each pixel on the retina based on that display pixel's measured luminance and the calculated scattered light from all other scene pixels. The calculation used Vos and Walraven's measurements of the human point spread function. Stiehl et al. measured the actual display luminance and calculated the retinal luminance for the display (Figure 6).

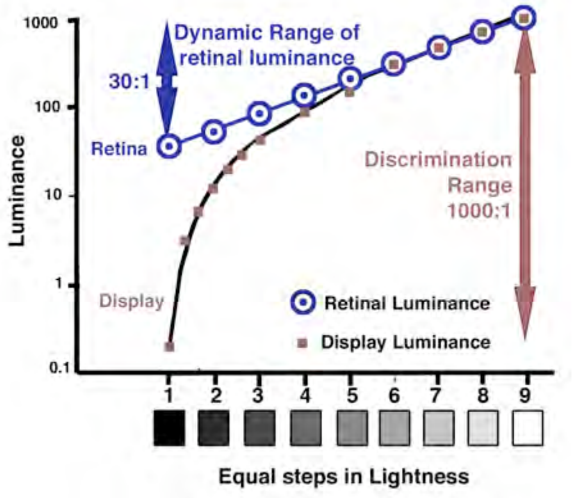

Figure 6 shows dots plotting the display luminance for the HDR transparent lightness scale. Stiehl's observers selected these values of transparency to be equal steps in lightness. The circled dots plot the calculated luminance on the retina, after intraocular scatter. Observers can discriminate display luminances over of 3 $\log_{10}$ units. This discrimination is made using calculated retinal luminances of only 1.5 $\log_{10}$ units.

These results show that measures of discrimination are distinct from measurements of dynamic range. Humans continue to discriminate between display blacks that are 1/1000th the white luminance, although the stimulus on the retina is limited by scatter to only 1/30th the white. Discrimination has to do with spatial comparisons. Cornsweet and Teller measured discrimination thresholds with different lightnesses. Using a pair of joined semi-circular test areas, with an adjustable surround, they measured observers' ability to detect edges. By changing the surround luminance around the test semicircles, they changed the test areas from light to dark. They showed that discrimination depends on the local, glare corrected, stimulus on the retina.[56] The observers' discrimination was the same regardless of whether it was white or black.

In the following paper in this journal[1] we asked observers to measure the appearance, between white and black, of the test patches used for camera measurements. In both targets, 20% of the scene area contained 40 calibrated luminance pie-shaped sectors in 4 circular patterns. The remaining surround (80% of the area) was either opaque, or maximum luminance. Image dependent intraocular scatter can transform identical display luminances into completely different sets of retinal luminances.

Observers were asked to assign 100 to the whitest area and 1 to the blackest, for both the white and black surrounds. The data is plotted in Figure 7 (bottom). The results in Figure 7 show the competing roles of glare and contrast. The squares plot the magnitude estimates for the high-glare white surround. Human contrast mechanisms make the gray sectors appear darker than a white than in a black surround. Glare limits discrimination within the lowest luminance circle.

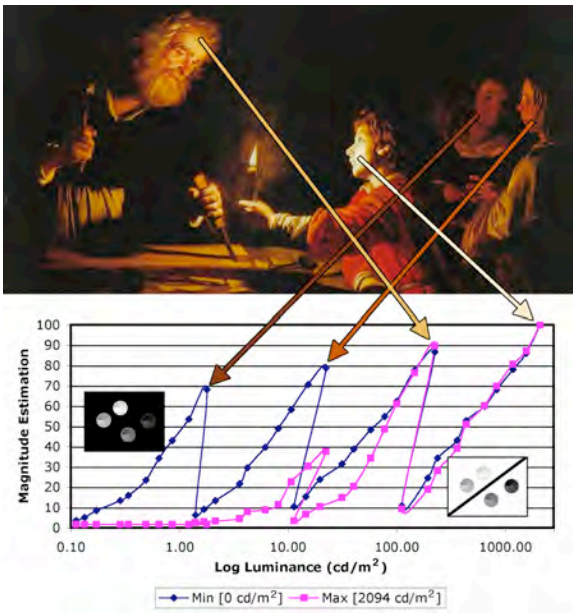

Figure 7 plots magnitude estimations between white and black vs. calibrated luminance for the 40 sectors in *4scalesBlack* and *4scalesWhite* test targets. Although the test luminances are exactly equal, the magnitude estimates are not. With a black surround observers can discriminate all 10 sectors in all four displays. With a white surround observers cannot discriminate below 2 cd/m^2. The *4scalesBlack* data is analogous to van Honthorst's rendering of figures in different illumination.

This experimental data for the black surround (diamonds) shows great similarity to Gerrit van Honthorst's faces in the painting "Childhood of Christ". Each Scale is the analog for the four figures in the painting. As the distance between the candle and the faces grew, the tones rendering the faces got slightly darker. Each person is rendered slightly darker, but the spatial contrasts for each are very similar. Note the correlation with observers' assignment of almost the same tone-scale magnitude estimates to Scales A, B, C, and D. The only differences was that each started and ended a few percent lower in magnitude estimates, despite the substantial decreases in luminance. Figure 7 just assigns numbers to 16th century observations. Chiaroscuro painters did not render luminances; rather they rendered what they saw.

5. DISCUSSION

We have seen four different rendering intents used extensively in HDR. They are: artist's aesthetic intent, improved image quality, calculated sensations written to film and display, and accurate capture and reproduction of scene luminance at each pixel. Calculating aesthetics is beyond the scope of this paper, but calculating an optimal HDR reproduction is not. It is very helpful to begin this discussion with the distinction between pixel based tone-scale approaches and spatial synthesis of what we see. The 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 input capture value. Every pixel with a capture value of ***i***, can have any output value.

Recalling Land's Black and White Mondrian experiment[26] is very helpful here. The display was an array of rectangular achromatic papers illuminated with a single bright light on the floor. More light fell on the papers at the bottom. Using a spot photometer meter, 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 magnitude estimates of the equiluminant papers. They reported that one was white and the other was black. The fact that they had equal luminances was inconsequential to magnitude estimates of what we see.

Pixel-based global tone scale functions cannot improve the rendition of both the black and the white areas in the Mondrian with the same luminance. 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. Land and McCann[26] made the case that spatial algorithms can automatically perform spatial rendering, doing what Adams did to compress HDR scenes into the limited range of prints. Such spatial rendering is not possible with tone-scale manipulations. By their design, global tone–scale functions have the same effect on all pixels with the same digital input value.

The two curves in Figure 7 are central to understanding how HDR images can improve the reproduction of scenes. First, it disposes the idea that a single tone-scale function can be helpful in rendering a 4.3 log unit image. On average an optimal tone-scale function could be a straight line, or other monotonic function, from magnitude estimate 100 at max luminance to 1 at minimum luminance. For *4scalesBlack* that tone scale coincides with visual estimates at only four luminances. All *4scalesWhite* magnitude estimates are darker than this function. Since areas with the different luminances resulted in the same magnitude estimates, there is no single function that can effectively render luminance inputs for either the *4scalesBlack* data or the *4scalesWhite* data. The global tone-scale approach requires that it attempt to do both, namely attempt to optimize all scenes.

In order to mimic human vision, output density must be independent of quanta catch. Early HDR algorithms[39] never attempted to determine actual scene luminance, since luminance is almost irrelevant to what we see.[26,57] Instead, these spatial algorithms mimicked vision by synthesizing HDR visual renditions of scenes using spatial comparisons. 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.[26] There, a white paper in dim illumination had the same luminance as a black paper in high illumination, but their lightnesses were strongly different. Figure 5 shows a real life embodiment.

As discussed in our second paper[1], the *4scaleBlack* target is a four-scale version of the Black and White Mondrian. The same magnitude estimates have four different luminances. Human spatial image processing is controlling these measurements. The advent of electronic imaging made possible spatial manipulation of images. Such spatial processing is not possible[58] in silver halide 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 rechain *Prometheus unchained*.

Details in the shadows are necessary to render objects in shade to humans. Clearly, 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. Therefore, one can make the case that improved quantization is key to successful image processing.[1] Although this argument is sufficient to explain the results of HDR imaging, such logic does not speak to whether other properties of vision could contribute to HDR image quality.

A favorite visual hypothesis is *adaptation*. Unfortunately, the term has many conflicting definitions: dark adaptation, namely the slow chemical migration of retinene; light adaptation, namely the fast neural responses;[19] and von Kries, namely the phenomenon driven model). One, or all, of these adaptations could play a role. Land repeated his Black and White Mondrian experiment in a 1 msec flash of non-uniform illumination. Observers could not report on the lightnesses of all areas from a single flash. A series of flashes were necessary to see the equiluminant white at the top and black at the bottom. Although both areas had the same flux (luminance*sec), one appeared white, and the other black. Viewing in 1 msec flashes affected the visibility of non-foveal details, but not the lightness of these areas. When seen, the lightness was the same as in continuous illumination. This experiment does not demonstrate that adaptation cannot play a role in what we see in a complex image. However, it demonstrates that such adaptation processes must happen in 1 msec illumination; that is, with minimal time and photon count.

Summarizing, high-dynamic-range image capture, with negatives or multiple exposures, provide better quantization of information to be used in spatial comparison algorithms. Spatial comparison images correlate with what we see.[26] By preserving the original scene's edge information, observers can see details in the shadows that are lost in conventional imaging. Spatial techniques have been used by painters since the Renaissance. Photographers have used multiple exposures and dodging / burning for 160 years.

Since 1978[25], there have been many different examples of spatial algorithms used to synthesize improved images from captured image plane luminances. Digital spatial algorithms, such as Frankle and McCann, have been used to display high-range scenes with low-range media. Kiesel and Wray's zoom processes improved processing efficiency.[59,60] McCann's color spatial gamut process used spatial information to increase apparent gamut.[61] Rizzi et al's Automatic Color Enhancement (ACE)[62,63,64] starts from the idea that pixels affect the values of neighbors according to the visual content of the scene. It is a nonlinear effect that decreases with distance. A wide variety of spatial imaging algorithms were compiled in the special session at Electronic Imaging, "Retinex at 40"[65]. Reinhard et al.[48] describes a number of Stockham-like spatial filtering techniques.

HDR imaging is successful because it preserves local spatial details. This approach has shown considerable success in experimental algorithms,[66] and in commercial products.[67,68] Figure 8 shows the results of spatial image processing from a single exposure using automatic firmware in an amateur camera.

Figure 8 shows images with and without spatial comparisons, both taken with a commercial HP 945 camera. The pictures are hand-held single-exposure images. The image on the right uses Retinex based *Adaptive Lighting/Digital Flash* camera options. The spatial processing removes the over-exposure of the windows while lightening the red rug, white marble altar and pews.

6.0 CONCLUSIONS

This paper describes the five-century history of HDR imaging. Glare is the scene- and camera- dependent scattered light falling on image sensors. First, glare limits the range of luminances that can be accurately measured by a camera, despite multiple exposure techniques. HDR image capture cannot accurately record the luminances in scenes beyond the glare limit. Second, magnitude estimate between white and black do not correlate with luminance: they depend on physical glare and physiological contrast. The improvement in HDR images, compared to conventional photography, does not correlate with accurate luminance capture and display. 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.

ACKNOWLEDGEMENTS

The author wishes to thank Ale Rizzi and Mary McCann for their very helpful discussions. The author also wishes to thank the editor and referees for their helpful advice.

REFERENCES

1 J. J. McCann and A. Rizzi, "Camera and visual veiling glare in HDR images", J. SID Vol. **15(9)** pp. in press, 2007

[2] Clair-obscur (French) and chiaroscuro (Italian) both mean "light-dark". The two terms are used to denote sharp contrasts of light in paintings, and prints.

[3] T. Mulligan & D. Wooters, eds, *Photography from 1839 to today, George Eastman House, Rochester, NY*, (Taschen, Koln, 1999), p. 360

[4] F. Hurter and V. C. Driffield, "*The Photographic Ressearches of Ferdinand Hurter & Vero C. Driffield*", W. B. Ferguson, Ed., (Morgan and Morgan Inc. Dobbs Ferry, 1974)

[5] C. E. K. Mees, *An Address to the Senior Staff of the Kodak Research Laboratories*, (Kodak Research Laboratory, Rochester, 1956)

[6] C. E. K. Mees, *The Fundamentals of Photography* (*Kodak Research Laboratory*, Rochester, 1920), p.82

[7] C. E. K. Mees, *From Dry Plates to Ektachrome Film: a Story of Photographic Research*, (Ziff-Davis Pub. Co, New York, 1961)

[8] A. Adams, *The Negative*, (New York Graphical Society, Little, Brown & Company, Boston, 1981) pp. 47-97

[9] A. Adams, *Examples: The Making of 40 Photographs*, (New York Graphical Society, Little, Brown and Company, Boston, 1984)

[10] L. A. Jones and H. R. Condit, "The Brightness Scale of Exterior Scenes and the Computation of Correct Photographic Exposure", *J. Opt. Soc. Am.*, Vol. **31**. pp. 651-678, 1941

[11] M. Luckiesh and A. H. Taylor, "A brightness meter," J. Opt. Soc. Am., Vol. **27**. p. 132, 1937

[12] C. W. Wyckoff, "An experimental extended response film", Tech Rep. NO. B31-, Edgerton Germeshauser and Grier, Inc., Boston, March 1961

[13] C. W. Wyckoff, "An experimental extended response film", SPIE Newsletter, March 1962

[14] A. Adams, *Polaroid Land Photography*, (New York Graphical Society, Little, Brown & Company, Boston 1978) pp. 146-151

[15] H. Davson, *The Eye: The Visual Process*, Academic Press, New York, 1962) Vol II

[16] S. W. Kuffler, "Discharge Patterns and Functional Organization of the Mammalian Retina", J. Neurophysiol. Vol. **16**, pp 37-68, 1953

[17] H. B. Barlow, "Summation and Inhibition in the Frog's Retina", J. Physiol., **119**, 69-88, 1953

[18] F. Ratliff, *Mach Bands: Quantatative Studies on Neural Networks in the Retina,* (Holden – Day, 1965)

[19] J. E. John E. Dowling, *The Retina: An Approachable Part of the Brain*, (Belknap Press, 1987)

[20] E. H. Land "The Retinex", *Am. Scientist*, **52,** 247-264, 1964.

[21] S. Zeki, "Vision of the Brain", (Blackwell Science Inc., Williston, Vermont, 1993)

[22] D. H. Hubel and T. N. Wiesel, "Brain and Visual Perception", (Oxford University Press, 2005)

[23] C. Blakemore C. and F. W. Campbell, "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images", *J. Physiol., Lond.* 203, pp 237-260, 1969

[24] J. J. Gibson, *The Senses Considered as Perceptual Systems***,** (Allen & Unwin London, 1968)

[25] E. H. Land, "Lightness and Retinex Theory", J. *opt. Soc. Am* 57. p. 1428A, 1967; For citation see J. *opt. Soc. Am,* Vol. **58**. p. 567, 1968

[26] E. Land and J. J. McCann, "Lightness and Retinex Theory", *J. Opt. Soc. Am*., Vol. **61**, pp. 1-11, 1971

[27] H. Wallach, "Brightness constancy and the nature of achromatic colors", *J. Exptl. Psychol,* Vol. **38**. pp. 310-324, 1948

[28] E. H. Land, "The Retinex Theory of Colour Vision", *Proc. Roy. Institution Gr. Britain*, Vol. **47**. pp. 23-58, 1974

[29] E. H. Land and J. J. McCann, "Method and system for reproduction based on significant visual boundaries of original subject", *U.S. Patent* 3,553,360, June 5, 1971

[30] E. H. Land, L. A. Ferrari, S. Kagen and J. J. McCann, "Image Processing system which detects subject by sensing intensity ratios", US Patent 3,651,252, Mar. 21, 1972

[31] J. J. McCann, "Aperture and Object Mode Appearances in Images", in Human Vision and Electronic Imaging XII, eds. B. Rogowitz, T. Pappas, S. Daly, Proc. SPIE, Bellingham WA, Vol. **6292-26**, 2007

[32] A. P. Ginsburg, "Method and apparatus for pattern analysis", US Patent 3,993,976, 1976

[33] T. G. Stockham, Jr. Image processing in the context of a visual model. Proc. IEEE, Vol. **60(7)**. pp.828—842, 1972

[34] H. R. Wilson and J. R. Bergen, "A four mechanism models for threshold spatial vision", Vision Res., Vol. **26**. pp. 19-32, 1979

[35] D. Marr, "The computation of lightness by the primate retina", *Vision Res.* Vol. **14**. pp. 1377-1388, 1974

[36] B. K. P. Horn, "Determining lightness from an image", Comp. Gr. Img. Proc. Vol. **3**. pp. 277-299, 1974

[37] J. J. McCann, "High-Dynamic-Range Scene Compression in Humans ", in Human Vision and Electronic Imaging, XII, eds. B. Rogowitz, T. Pappas, S. Daly, Proc. SPIE, Bellingham WA, Vol. **6057-47**, 2006

[38] J. Frankle, & J. J. McCann, Method and apparatus of lightness imaging, U.S. Patent, 4384336, May 17, 1983

[39] J. J. McCann, Black Capturing a black cat in shade: past and present of Retinex color appearance models, J. Electronic Imaging, Vol. **13**. pp. 36-47, 2004

[40] J. J. McCann, "Calculated Color Sensations applied to Color Image Reproduction", in *Image Processing Analysis Measurement and Quality*, Proc. SPIE, Bellingham WA, Vol. **901**. pp. 205-214, 1988

[41] S. Ochi and S. Yamanaka, US Patent, 4,541,016, filed Dec 29, 1982, issued Sep 10, 1985

[42] L. E. Alston, D. S. Levinstone, W. T. Plummer, "Exposure control system for an electronic imaging camera having increased dynamic range", US Patent 4,647,975, Filing date: Oct 30, 1985, Issue date: Mar 3, 1987

[43] S. Mann, Compositing Multiple Pictures of the Same Scene", Proc. IS&T Annual Meeting, Vol. **46**. pp. 50,-52, 1993

[44] S. Mann and R. W. Picard, "On Being "Undigital" with Digital Cameras: Extending Dynamic Range by Combining Different Exposed Pictures", Proc. IS&T Annual Meeting, Vol. **48**. pp. 442-448, 1995

---

[45] P. E. Debevec, J. Malik, “Recovering high dynamic range radiance maps from photographs”, *ACM SIGGRAPH’97*, .pp. 369-378, 1997

[46] K. Devlin, “A Review of Tone Reproduction Techniques”, Bristol: Technical Report CSTR 02 005, 2002

[47] C. Gatta “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

[48] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, *High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting*, (Elsevier, Morgan Kaufmann, Amsterdam, chap 4, 2006)

[49] The term *tone-scale-reproduction function* is used in these papers to describe the system response to light in a camera-image-plane pixel. The term applies to systems such as a silver-halide film in which a particular quantum catch uniquely describes that pixels response.

[50] R W. G. Hunt, *The Reproduction of Colour*, Sixth Edition, John Wiley & Sons, Ltd., 2004

[51] J. J. McCann, Color Imaging Systems and Color Theory: Past, Present and Future, Proc. of SPIE , Vol. **3299**. pp. 38-46, 1998

[52] J. J. McCann, “ The Application of Color Vision Models to Color and Tone Reproductions” in Proc. Japan Hardcopy’88 . pp. 196-199, 1988

[53] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, “High dynamic range display systems” ACM Transactions on Graphics, Vol. **23(3)** . pp. 760-768, 2004

[54] ISO 9358:1994 Standard, “Optics and optical instruments. Veiling glare of image forming systems. Definitions and methods of measurement”, (ISO, 1994)

[55] W.A. Stiehl, J. J. McCann, R.L. Savoy, “Influence of intraocular scattered light on lightness-scaling experiments”, *J. opt Soc .Am.*, Vol. **73**. pp. 1143-1148, 1983

[56] T. Cornsweet and D. Teller, “Relation of Increment Thresholds to Brightness and Luminance”, *J. Opt. Soc. Am.* Vol. **55**. pp. 1303-1308, 1965

[57] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods Quantitative Data and Formulae*, *2nd Ed,* (Wiley, New York, chap 6, 1982)

[58] Silver halide development processes using chemical restrainers can affect local departures from quanta-catch proportionality. These local chemical mechanisms have never been demonstrated over a wide enough spatial region to mimic human vision.

[59] K. C. Kiesel and W. R. Wray, “Reconstitution of images”, U.S. Patent 4,649,568, Mar. 10, 1987

[60] W. R. Wray, “Method and Apparatus for Image processing with Field Portions”, U.S. Patent 4,750,211, Jun. 7, 1988

[61] J. J. McCann, “A Spatial Color-Gamut Calculation to Optimize Color Appearance”, in *Colour Image Science: Exploiting Digital Media*, Ed. L. W. MacDonald and M. R. Luo (Wiley & Son Ltd., Chichester, 2002) pp. 213-233

[62] A. Rizzi, C. Gatta, D. Marini, “A New Algorithm for Unsupervised Global and Local Color Correction”, *Pattern Recognition Letters*, Vol. **24**, No. 11, July 2003. pp. 1663-1677

[63] A. Rizzi, C. Gatta, D. Marini, “From Retinex to Automatic Color Equalization: issues in developing a new algorithm for unsupervised color equalization”, *Journal of Electronic Imaging* Vol. **13** pp. 75-84 2004,

[64] E. Provenzi, M. Fierro, A. Rizzi, L. De Carli, D. Gadia, D. Marini, “Random Spray Retinex: a new Retinex implementation to investigate the local properties of the model” *IEEE Transactions on Image Processing*, Vol. **16**, pp. 162-171, 2007.

[65] J.J. McCann, ed., “Retinex at Forty”, *J. Electronic Img*, Vol. **13**. pp. 1-145, 2004

[66] J. J McCann, “Rendering High-Dynamic Range Images: Algorithms that Mimic Human Vision”, in *Proc. AMOS Technical Conference*, Maui, pp. 19-28, 2005

[67] R. Sobol, “Improving the Retinex algorithm for rendering wide dynamic range photographs”, *J. Electronic Img*, Vol. **13**. pp. 65-74, 2004

[68] HP cameras with Digital Flash or Adaptive Lighting (Models 945, 707,717,727) use spatial comparisons to improve image rendering.