Camera and visual veiling glare in HDR images

John J. McCann
Alessandro Rizzi

Abstract — High-dynamic-range (HDR) images are superior to conventional images. The experiments in this paper measure camera and human responses to calibrated HDR test targets. We calibrated a 4.3-log-unit test target, with minimal and maximal glare from a changeable surround. Glare is an uncontrolled spread of an image-dependent fraction of scene luminance in cameras and in the eye. We use this standard test target to measure the range of luminances that can be captured on a camera's image plane. Further, we measure the appearance of these test luminance patches. We discuss why HDR is better than conventional imaging, despite the fact the reproduction of luminance is inaccurate.

Keywords — *HDR imaging, veiling glare, calibration of multiple exposure techniques, spatial algorithms, Retinex, ACE.*

1 Introduction

This paper is the second of two consecutive papers on HDR imaging. The prior paper[1] reviews the long history of HDR imaging from Renaissance paintings to modern digital imaging. This paper measures the effects of veiling glare on camera image capture and the appearance of the same scenes viewed by humans.

Recently, multiple exposure techniques[2] have been combined with LED/LC displays that attempt to accurately reproduce scene luminances.[3] However, veiling glare is a physical limit to HDR image acquisition and display. We performed camera calibration experiments using a single test target with 40 luminance patches covering a luminance range of 18,619:1 (4.3 log units). Veiling glare is a scene-dependent physical limit of the camera and lens.[4–7] Multiple exposures cannot accurately reconstruct scene luminances beyond the veiling-glare limit.

Human observer experiments, using the same targets, show two independent and opposing visual mechanisms. Intraocular veiling glare reduces the luminance range on the retina while physiological simultaneous contrast[8] increases the apparent differences.

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 (HDR) displays, it processes scenes using spatial comparisons.

2 HDR test targets

Although ISO 9358:1994[9] provides a standard to compare different lenses and apertures, we wanted to measure the effects of veiling glare on HDR imaging. We used a single calibrated test target with 40 test luminance sectors (dynamic range = 18,619:1). Nearly 80% of the total target area was an adjustable surround; 20% of the area was luminance test patches. Using opaque masks to cover the surrounding portions of the scene, we photographed three sets of HDR test scenes with different amounts of glare. The experiment compares camera digits with measured scene luminance over a very wide range of luminances and exposure times. This experiment measured the extent that veiling glare distorts camera response in situations common with HDR practice.

In 1939, Kodak patented the Projection Screen Print Scale for making test prints.[10] It is a circular step wedge with 10 pie-shaped wedges, each with a different transmission. The range of transmissions was 20:1. After focusing a negative in an enlarger on the print film plane, darkroom technicians would place this scale on top of the unexposed print film in the dark. The wedges transmitted 82, 61, 46, 33, 25, 17, 14, 9, 8, and 4% of incident light so as to make a quick and accurate test print to select the optimal print exposure. We used these scales to make 10 different test luminance sectors.

The components of our test display are shown in Fig. 1. The display is made of transparent films attached to a high-luminance lightbox. There are four Kodak Print Scale transparencies mounted on top of 0.0 (ScaleA), 1.0 (B), 2.0 (C), and 3.0 (D) N.D. filters. The 40 test sectors are constant for both minimal (*4scaleBlack*) and maximal (*4scaleWhite*) glare so that both targets have the same range of 18,619:1. For minimal glare, we covered all parts of the display except for the pie-shaped projection scales with an opaque black mask (*4scaleBlack*). For maximal glare, the opaque black mask was removed so that the zero-glare surround was replaced with maximal glare (*4scaleWhite*). The diagonal line in *4scaleWhite* is an opaque strip in front of the display. To further reduce glare, we covered the background and scales A, B, and C, leaving only the light coming from scale D (*1scaleBlack*) with a 20:1 range.

Extended revised version of a paper presented at the 14th Color Imaging Conference held November 6–10, 2006 in Scottsdale, Arizona.
John J. McCann is with McCann Imaging, 161 Clafin St., Belmont, MA 02478; telephone/fax 617/484-7865, e-mail: mccanns@tiac.net.
A. Rizzi is with the Universita degli Studi di Milano, Italy.

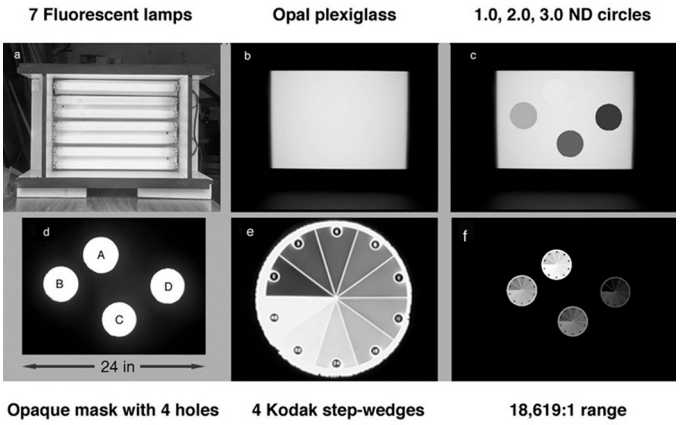

FIGURE 1 — (a) Light source made of seven fluorescent tubes (20 W). (b) An opal-Plexiglas diffuser placed 15 cm in front of the lamps. (c) The addition of three circular neutral density filters attached to the Plexiglas with densities of 1.0, 2.0, and 3.0. (d) An opaque mask that covered the entire lightbox except for four circular holes registered with the N.D. filters. (e) An enlarged view of a single Kodak Projection Print Scale. (f) Assembled *4scaleBlack* target with a dynamic range of 18,619:1 (2049 to 0.11 cd/m^2). Using opaque black masks, the luminance of each sector was measured with a spot luminance meter (Konica-Minolta LS-100C), one wedge at a time in a dark room.

3 Camera veiling-glare limits

We made separate sets of measurements, first with a digital camera then with a 35-mm film camera using both slope 1.0 slide duplication and conventional negative films. We also used a lensless pinhole camera. We used all three HDR calibrated targets to measure the camera response. With the *1scaleBlack* target, we measured the camera response using only the lowest luminances with a 20:1 range. With the *4scaleBlack* target, we measured the camera response using a high display range of 18,619:1 with minimal glare. With the *4scaleWhite* target, we measured the camera response using the same display range with maximal glare.

3.1 Digital camera response

We made photographs using a typical compact, high-quality digital camera (Nikon Coolpix 990) with a manual mid-range aperture (set to *f*/7.3) and exposure time controls. The experiment photographed three sets of records shown in Fig. 2.

The intent of multiple exposures in HDR imaging is to calculate a new image with a significantly greater dynamic range record of the scene.[2] The idea is simply to assume that scene flux (cd/m^2-sec) generates a unique camera digit. Longer exposures collect more scene photons and can move a dark portion of the scene up onto a higher region of the camera response function. This higher exposure means that the darker portions of the scene have better digital segmentation, more digits for better quantization of luminances. Digital HDR multiple-exposure techniques,[2,11–14] claim to extend the cameras range by calibrating flux *vs.* camera digit. Debevec and Malik make the specific argument that

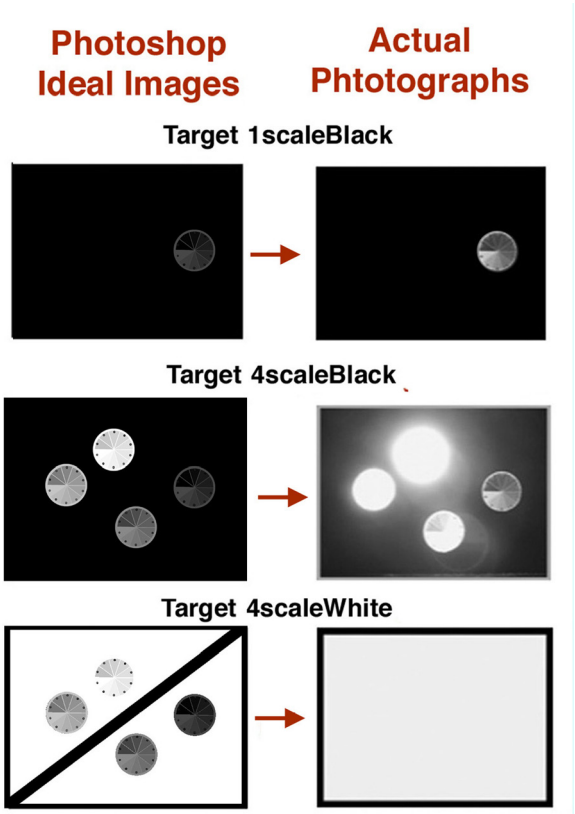

FIGURE 2 — Ideal synthetic Photoshop images of the three test targets (left) and the actual 16-sec exposures of the test target acquired with the digital camera (right). The 16-sec exposure is optimal for recording the luminances of the lowest luminance scale D. The punctual luminance values at each wedge sector remain unchanged in the three scenes. A 16-sec exposure of the *1scaleBlack* target shows a typical camera response with digits from 37 to 201. Veiling glare has a small but significant incremental effect on camera response to *4scalesBlack*. Glare from the other test sectors has increased camera digits. The darkest sector digit increased from digit 37 to 98. Veiling glare overwhelms the camera response to *4scalesWhite*. All pixels have the same saturated maximum value (242).

calculations using multiple exposure data derive high-dynamic-range scene luminance. For reference, this technique will be described below as multiple exposure to scene luminance (ME2SL). This is to distinguish this use of multiple exposures from others described in the previous paper.[1]

Given our calibration measurements of the test target scene, we can paste together in Photoshop the desired ideal image. By dividing the 2049 to 0.11 luminance range into 256 levels, we obtain an 8 cd/m^2 increment per gray level. These synthetic ideal images are shown in the left column of Fig. 2. The right column shows actual photographs. The differences between ideal and actual images are due to unwanted veiling glare. The results in Fig. 2 show that glare can be substantial.

We measured the veiling glare's influence, with 16 shots taken with variable exposure times and the same *f*/7.3 aperture (Fig. 3). We selected this camera because it has manual controls for both aperture and time of exposure. The *1scaleBlack* photographs have the lowest veiling glare and provide an accurate measure of the camera sensor

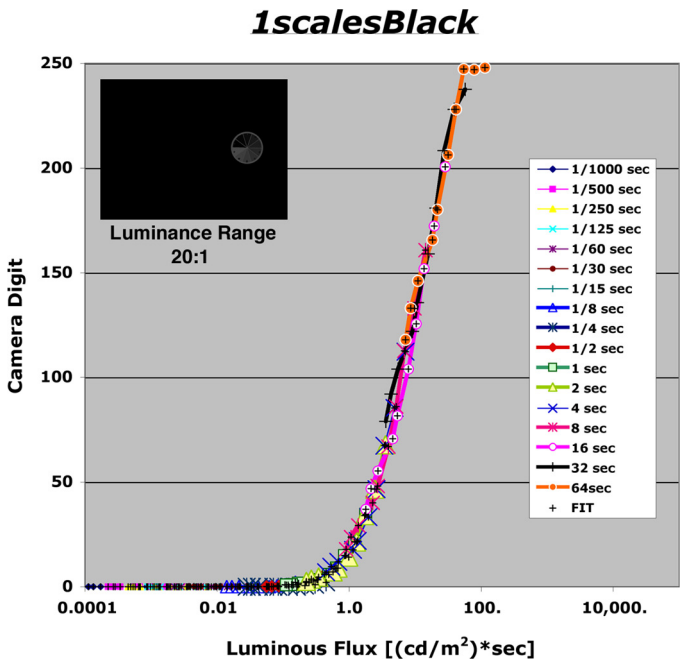

FIGURE 3 — *1scaleBlack* (range 20:1) camera digits. It shows the desired coincidence of camera digits and flux. The curve provides us with an accurate camera response function. The sensor digits saturate at 247 with a flux of 78.4 sec-cd/m^2; at digit 1 the flux is 0.107 sec-cd/m^2. The camera dynamic range is 731:1, or 2.9 log units. The black + symbols plot lookup table data determined from the average measured digital data. This lookup table converts the camera digit scene flux.

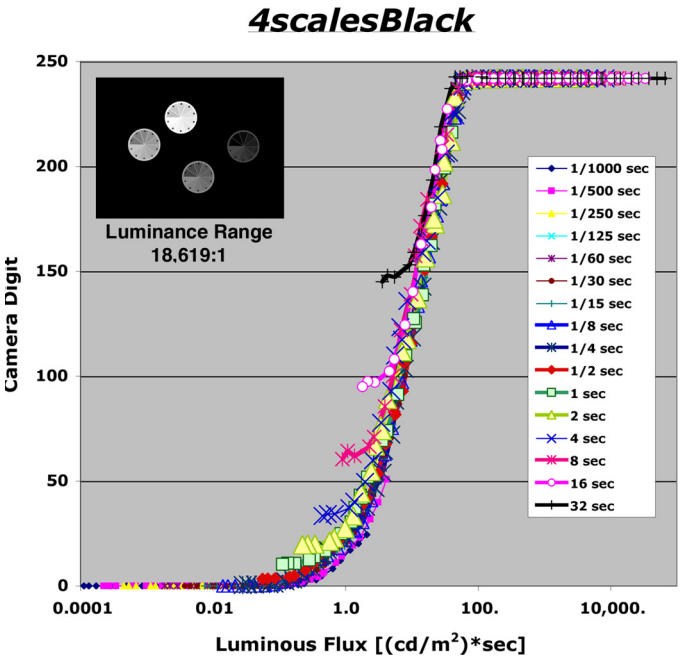

FIGURE 4 — *4scaleBlack* (range 18,619:1) camera digits. It shows the minimal effects of glare for this range and configuration using a black surround between test scales. For an optimum exposure (1/2 sec), the sensor digits saturate at 242 with a flux of 119 sec-cd/m^2; at digit 11 (departure from camera-response curve) with a flux of 0.84 sec-cd/m^2. The glare-limited dynamic range is 141:1, or 2.2 log units for this exposure. The effects of glare are seen as departures from a single camera function in low-luminance sectors.

response function. The only sources of glare are the test patches themselves (range 20:1). The camera response is the average digital value (from 491 pixels) calculated in Photoshop for a circular area falling inside the pie-shaped luminance sector.

The degree of overlap of multiple exposure responses with flux is possible because of the accuracy of the camera's exposure-time mechanism and the level of veiling glare found in the 20:1 test target. The multiple exposure to scene luminance (ME2SL) technique works well in these conditions. The data in Fig. 3 is plotted as flux (luminance-time) because that is related to the total number of photons falling on the camera's CCD sensor. We cannot derive actual photon counts without detailed knowledge of the cameras spectral sensitivities, and anti-blooming, noise reduction, and tone-scale circuits. The results in Fig. 3 provide a consistent measure of camera response and a lookup table that allows us to convert camera digit to estimated scene luminance. This estimate is accurate as long as the ME2SL technique is error free, as shown in Fig. 3.

In *4scaleBlack*, the camera's digit responses to four 10-step scales attempt to capture a combined scene dynamic range of 18,619:1 (Fig. 4). This target measures the minimum glare for a scene with this range, because it has an opaque black surround. The only source of glare is the test patches that vary from 2094 to 0.11 cd/m^2.

The data from Fig. 4 shows that camera digit does not predict scene flux because the data for scale D fails to fall on the single-camera-response function measured in Fig. 3. The same scene luminance generates different (exposure-dependent) digits. This is important because this display was intended to measure the minimal glare for an 18,619:1 scene. Despite the fact that 80% of the target area is glare free, we measure a problem with the ME2SL technique.

When we removed the black mask covering the lightbox in the background, we go to the situation with maximal veiling glare (*4scaleWhite*). Nearly 80% of the pixels are making the highest possible contribution to veiling glare (Fig. 5). Figure 5 shows that the influence of glare is dramatic. For scales C and D, camera digits are controlled as much by glare as by scene flux.

The data from all three sets of photographs are different. Data from *1scaleBlack* (Fig. 3) provides a single-camera sensor response function. Camera digit correlates with scene flux. Data from *4scaleBlack* shows a lack of correlation for low luminances at some exposures. Data from *4scaleWhite* shows that glare corrupts camera digit correlation with scene flux. This is a major problem for the ME2SL technique. It works well only when veiling glare is low. Veiling glare is scene-dependent.

Camera digits from multiple exposures cannot provide a trustable means of measuring HDR scene flux. Camera digits cannot accurately record HDR scene flux because of glare. Veiling glare is scene dependent. We also have performed tests using different cameras and various changeable lenses, and we obtained similar results.

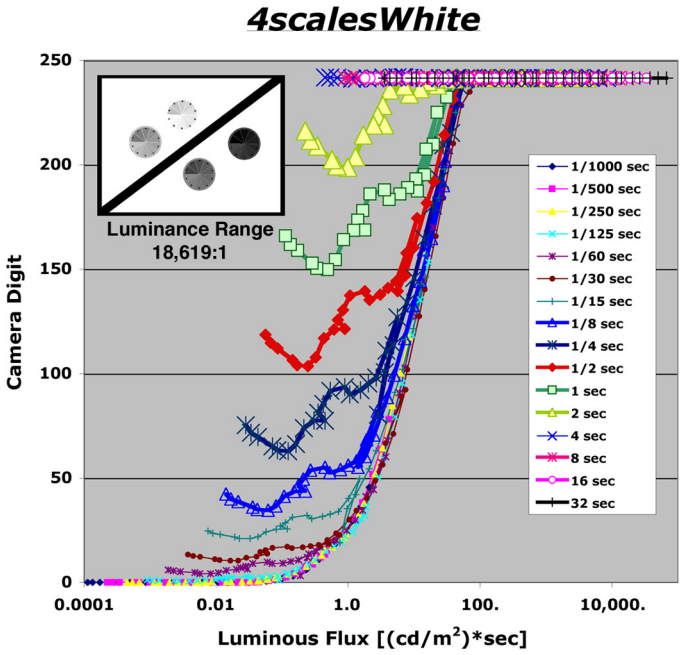

FIGURE 5 — The camera digits from 16 different exposure times for the *4scaleWhite*, the high-glare target. The many large departures from a single line are due to scene-dependent glare. All departures from the camera response function (Fig. 3) are errors in the ME2SL technique.

We took the data of scale D from *1scaleBlack* to generate a lookup table that describes camera digit as a function of flux and its inverse (see FIT in Fig. 3). We then used this camera response lookup table to convert the camera digits from *4scalesBlack* and *4scalesWhite* to the calculated flux. We took the ratio of camera-estimated flux to actual measured flux. This ratio is a measure of ME2SL error of each target sector. If the camera digit accurately predicted scene flux, then this ratio equals 1.0. Ratios greater than 1.0 measure the magnitude of errors introduced by glare. Figure 6 plots these ratio values *vs.* scene luminance.

The *4scaleBlack* target has no glare from 77% of the scene area, yet shows worst-case errors as large as 300% distortions. The *4scaleWhite* target (maximal glare) shows 10,000% errors. If we hypothesize a variety of different surrounds to substitute for the all white, or all black surround, all possible luminance backgrounds, around scales A, B, C, and D, will fall in between the white and the black data sets. Substituting all possible surrounds for the white, and the black, will generate veiling-glare luminance estimate errors between 300 and 10,000% for this scene.

The data from *4scaleWhite* in Fig. 6 shows a series of parallel lines deviating from the slope 0.0. They show different, large glare distortions for the same luminance, depending on exposure. This adds *exposure* to the list of physical attributes controlling glare. The others are: scene, camera body, lens, and aperture size. The ME2SL technique is subject to glare limits that are difficult to estimate for any scene, camera, lens, aperture, and exposure.

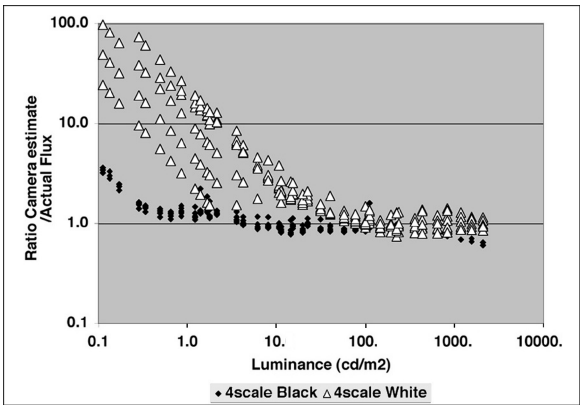

FIGURE 6 — Ratio of camera-estimated flux to actual flux for *4scaleBlack* and *4scaleWhite*. If the camera digit accurately measures scene luminance (ME2SF), then all the data must fall on a horizontal line (ratio = 1.0). The results support that hypothesis from 50 to 2048 cd/m^2 (range 40:1). Below 50 cd/m^2, the *4scaleWhite* data show that veiling glare distorts the ratios, and hence the luminance estimates. The same is true for *4scaleBlack* below 3 cd/m^2. The target has a range of 4.3 log units. Camera estimates of luminance are accurate on average over 2.8 log units with black surround (minimal glare) and 1.6 log units with white surround (maximal glare).

3.2 Duplication film-camera response

We made another set of photographs with a typical high-quality 35-mm film camera (NikonFM with a Nikkor 50 mm 1:2 lens) using Kodak slide duplication film. This follows the single exposure HDR capture technique described by McCann[15] in tutorials at Siggraph conferences in 1984 and 1985. Slide duplication film has a slope of 1.0 on a log exposure *vs.* log luminance plot. In other words, output luminance equals input luminance. Since it is a color film, it can be scanned for color and does not require calibration to remove the color masks found in color negative film. Here,

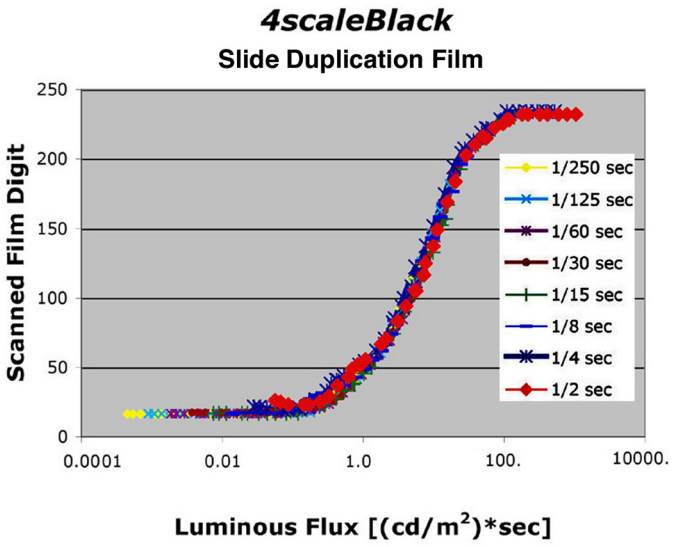

FIGURE 7 — Scanned film digit *vs.* luminous flux for eight exposures ranging from 1/250 to 1/2 sec for *4scaleBlack*. The data from the eight different exposures superimposed to form a single function, except for the very lowest luminance sectors.

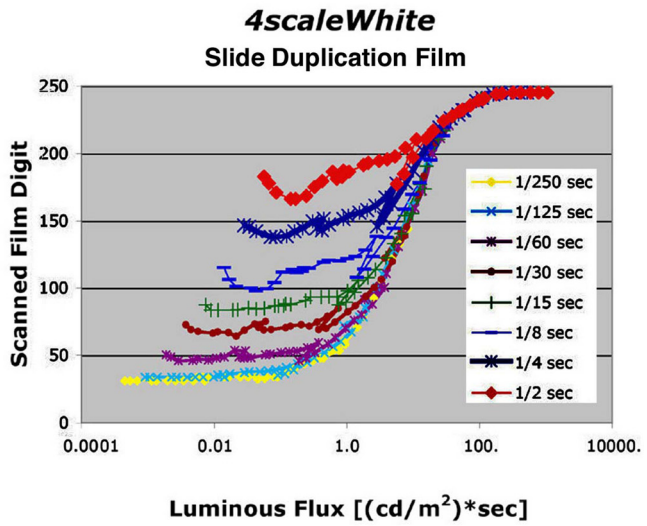

FIGURE 8 — Data for *4scaleWhite*. Here, the white surround adds veiling glare to generate eight different response functions.

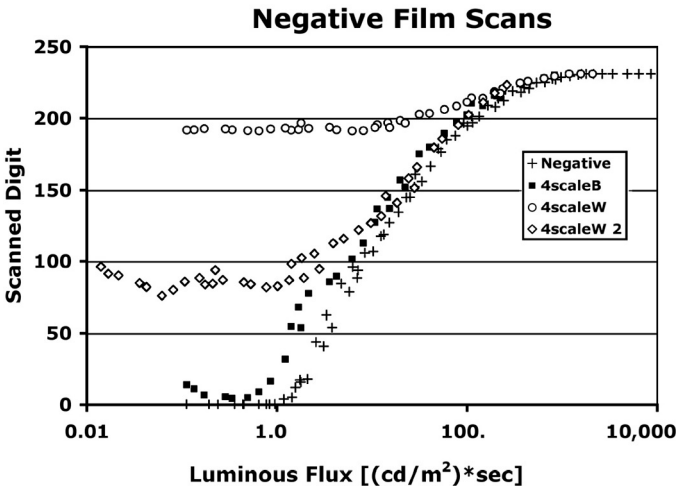

FIGURE 9 — Scanned single negative digits *vs.* log luminous flux for three targets. The *1scaleBlack* data (+ symbol) plots data from seven negatives with different exposures. The + symbols report the response of the camera-film-scanner process with the lowest level of glare (20:1 target). It shows that the negative process can accurately record fluxes from 2639 to 0.24 sec-cd/m^2. The *4scaleBlack* (black squares) and *4scaleWhite* (white circles and diamonds) targets measured negative responses with single exposures. The single exposure curve from *4scaleBlack* saturates at 1181 sec-cd/m^2 and inverts at 0.34 sec-cd/m^2. The inversion is caused by glare that limits usable range. There are two different single exposures for *4scaleWhite* target; one is eight times longer than the other. The data from *4scaleWhite* (white circles) saturates at 1181 cd/m^2 and inverts at 6.17 cd/m^2. The curve from *4scaleWhite2* (white diamonds) has a maximum digit at 2094 cd/m^2 and inverts at 6.17 cd/m^2.

we use multiple exposures to capture both 18,619:1 displays (*4scaleBlack* and *4scaleWhite*). The exposed film was developed with a standard E6 process. All exposures were mounted in a single 35-mm film holder so that all images were scanned at the same time with the same scanner settings. The scanner (Epson Perfection Photo) was calibrated for E6 films. Figures 7 and 8 plot scanned positive film digit *vs.* log flux. Figure 7 data shows that this particular camera-film-scanner system has less veiling glare than the digital camera described in section 3.1.

Although there may be small differences between this data and the digital camera's response in Section 4.1, the same scene-dependent glare dominates both results.

3.3 Negative film-camera response

We made another set of photographs with the same NikonFM camera using Kodak Max 200 negative film. The exposed film was developed with a standard C41 process. Again, all exposures were mounted for a single scan at the same time with the same scanner settings. The scanner was an Imacon Flextight Precision with calibration for C41 films.

First, we used seven different exposures to measure the camera-film-scanner process using the low-glare 20:1 single scale (*1scale Black*). The gray triangles in Fig. 9 plot the combined response of film, camera, development process, and film scanner (called Negative).

We then photographed *4scaleBlack* and *4scaleWhite*, using the same single exposures to capture the 18,619:1 displays. In the case of the white surround, the gray scales had high digital values, covering less than 50 digits. We made an additional single exposure called *4scaleWhite2* negative with shorter exposure time. This scanned film had a digit range of 150, allowing better quantization of the scene.

Figure 9 shows that the negative-camera-scanner process can accurately record fluxes from 2639 to 0.24 sec-cd/m^2 (dynamic range of 11,100:1, or 4.05 log units) (see Table 1). The single exposure data from *4scaleBlack* show a small effect of glare from the addition of 30 higher luminance test pie-shaped areas. This glare reduces the dynamic range of the image in the camera to 3.5 log units. The glare from the white surround in *4scaleWhite* and *4scaleWhite2* reduces the dynamic range of the image measurements to 2.3 and 2.5 log units.

The fact that the dynamic ranges for the two exposures of the *4scaleWhite* target are almost the same is important. Their response curves in Fig. 9 are very different. The *4scaleWhite* scanned digits have a max of 231 and a min of 191. The *4scaleWhite2* scanned digits have a max of 223 and a min of 94. The range of digits representing the scene is only of secondary importance. The range of digits describes the number of quantized levels used to represent the image. It controls discrimination, but does not control the dynamic range of the image.

There is an interesting characteristic of scanned negatives. The lowest density film (highest transmission) response measures the lowest scene luminance levels. These areas send the most light to the scanner and have the best signal-to-noise scanner response. As shown in Fig. 9, there is a long "toe" region that stretches from 100 to 10,000 sec-cd/m^2. Over that 2-log-units scene range, the scanned digits are all above 200. Although the digital quantization is poor, the

TABLE 1 — Table 1 shows the range limits of the negative film plotted in Fig. 9. It lists the maximum flux below system saturation (max flux), the minimum flux above digit reversal from glare (min flux), the dynamic range equal to max flux/min flux ratio (Range), and log ratio (Log Range).

Target	Max Flux	Min Flux	Range	Log Range
Negative	**2639.0**	**0.24**	**11,100**	**4.05**
4scaleB	**1181.4**	**0.34**	**6,345**	**3.54**
4scaleW	**1181.4**	**6.17**	**192**	**2.28**
4scaleW 2	**2094.2**	**6.17**	**340**	**2.53**

sampling (average digital value) discrimination is excellent because of the strong signal read by the scanner.

Conventional negative film can capture a greater range of luminances than falls on the camera image plane from these targets. The dynamic range of a single exposure negative-film-scanner process exceeds the glare-limited *4scaleBlack* scene by 0.5 log units and glare limited *4scaleWhite* scene by 1.8 log units (Table 1). Multiple exposures with negative films serve no purpose. The glare-limited ranges of the camera and these HDR scenes are smaller than the film system range.

L. A. Jones and H. R. Condit[16] measured the luminance range in 128 typical outdoor photographic scenes. They reported a minimum range or 27:1; the maximum was 750:1; and the average was 160:1. Nevertheless, one can increase the scene range by including the light source and specular reflections. The above data from Section 3 suggests that in high, and in average glare scenes, the glare-limited image dynamic range on the film/CCD image plane will be less than 3.0 log units. Only in special cases, very low-glare scenes, will the image plane's dynamic range exceed 3.0. The data here show how well the designers of negative films did in optimizing the process. They selected the size distribution of silver halide grains to make the negative have a specific dynamic range, around 4.0 log units. Thus, single-exposure negatives capture the entire range possible in cameras, with low glare scenes. For most scenes, this image capture range provides a substantial exposure latitude, or margin of exposure error. After reading the papers of C. K. Mees, L. A. Jones, and H. R. Condit, it is easy to believe that this fact is not a coincidence.[16–18]

3.4 Pinhole-camera response

Sowerby, in his *Dictionary of Photography*[19] discusses the reflection of light in lenses as the diversion of an appreciable portion of the incident light from its intended path. The small percentage of light reflected from each air–glass surface is called a parasitic image. Parasitic images that are completely out of focus give rise to a general fog that limits the dynamic range of the image falling on the film plane. The actual *4scaleB* image (Fig. 2) shows a magnified inverted in-focus parasitic image, as well as the out-of-focus fog from other parasitic images. Multiple exposures improve the digital quantization and thus the sensor's performance. Nevertheless, multiple exposures have no effect on the dynamic range of the image falling on the sensor. The digital camera used in Section 3 has nine elements and 153 parasitic images.[20] The film camera used in Sections 3.2 and 3.3 has seven elements and 91 parasitic images.

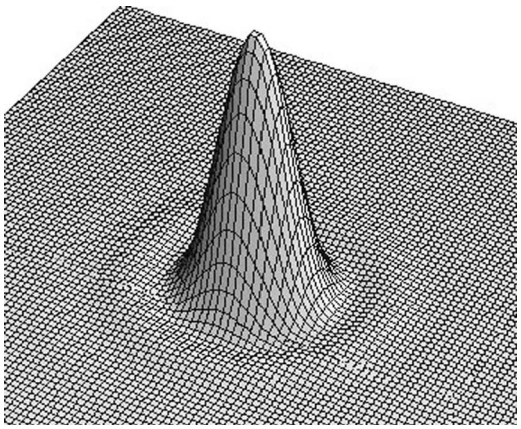

FIGURE 10 — Airy pattern of diffracted light from a point source imaged by a pinhole. The central lobe, called the Airy disc, totals 83.7% of the light falling on the image plane. All other light outside the first minimal ring contributes to diffraction fog.

An interesting problem is to measure the dynamic range of images made with a lensless pinhole camera.[21] We made a pinhole out of soft black plastic, counter-bored with a 1-cm drill and pierced with a needle. The pinhole was slightly elliptical. The average diameter was 376 μm (392 × 362). It was placed 50 mm from the film plane. Each point source in the scene is diffracted by the pinhole. Figure 10 shows the Airy pattern formed by diffraction.[22] The major lobe (peak to first minimum) called Airy's disc, is 83.7% of

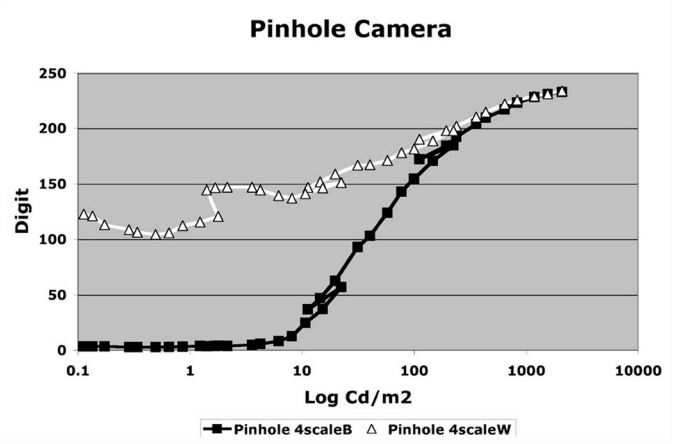

FIGURE 11 — Pinhole camera digits scanned from 800 ASA negative film images using 180-sec exposures.

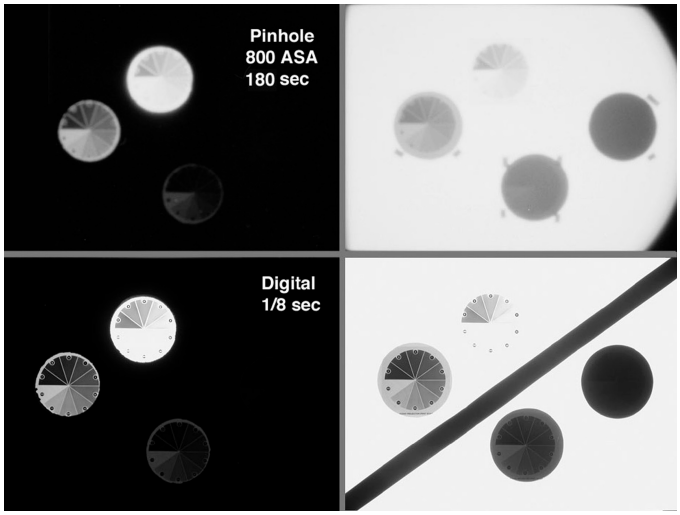

FIGURE 12 — Pairs of images made with the pinhole (without diagonal bar) and the digital camera. The pinhole images are less sharp and the *4scaleW* is lower in luminance range than the digital images.

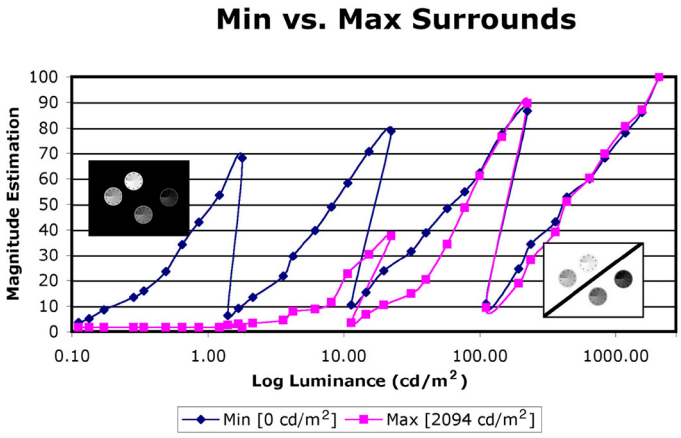

FIGURE 13 — Magnitude estimation of appearance *vs.* calibrated luminance for the 40 sectors in *4scalesBlack* and *4scalesWhite* test targets. Although the luminances are exactly equal, the appearances are not. With a black surround, observers can discriminate all 10 sectors in all four displays. With a white surround, observers cannot discriminate below 2 cd/m^2.

the light falling on the image plane.[23] The diameter of that lobe in 550-nm light in this camera is 0.178 mm. The remaining 16.3% of the light is diffracted outside the disc to form a diffracted fog that limits dynamic range.

Figure 11 plots the scanned digits from negatives taken in the pinhole camera with 180-sec exposures of the *4scaleW* and *4scaleB* targets.

The diffraction fog from the white surround increases responses of the gray sectors compared to the same luminances in the black surround. Figure 12 shows pinhole camera images compared to one of the digital pairs from Section 3.1. We selected the 180-sec time with the pinhole camera to optimize exposures for both targets. The average digit for the highest-luminance sector in scale B (*4scaleBlack*–Pinhole) was 185. We selected the 1/8-sec digital image pair because (*4scaleBlack*–Digital) had an average digit of 185 for the same sector.

Summarizing, regardless of the type of camera, film, and lens, HDR images have strong optical limits. The range of light falling on the sensors is limited by veiling glare from parasitic images in glass 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. That range depends on both the lens/camera and the scene. Scene dependence is a substantial problem for any ME2SL HDR algorithm[24] attempting to measure scene luminance.

4 Visual response to HDR display

The second effect of veiling glare on HDR imaging is intraocular scatter that controls the dynamic range of luminances on the retina. In Section 3.0, we saw that camera glare limits the range of luminances falling on the camera sensor plane. Human glare is caused by Tyndall scattering by macromolecules in the intraocular media, as well as the layers of neurons between the lens and the sensors. Scatter limits the eye's dynamic range more than glass lenses limit cameras. Here we will describe the range of discrimination and the corresponding range of retinal luminances. In addition, we will measure the observed appearance for both *4scales Black* and *4scales White* test targets.

4.1 Visual appearance of HDR displays

We asked observers to evaluate the appearance of the *4scaleBlack* and *4scaleWhite* displays using magnitude estimation. Observers sat 1.9 m from the 61-cm-wide display. The radius of each sector was 5.1 cm; subtending 2.4°. Three observers were asked to assign 100 to the "whitest" area in the field of view, and 1 to the "blackest." We then instructed them to find a sector that appeared middle gray and assign it the estimate 50. We then asked them to find sectors having 25 and 75 estimates. Using this as a framework, the observers assigned estimates to all 40 sectors. The data from each observer (ages 31, 64, and 68) was analyzed separately. No difference between observers was found.

The average magnitude estimate results (Fig. 13) show very dramatically the role of spatial comparison and scattered light in vision. The *4scalesBlack* and *4scalesWhite* appearance estimates overlap for only the top five luminances. Below that, simultaneous contrast makes the luminances in the white surround darker. The white surround makes the local maxima in scales C and D darker than in the zero-luminance surround. Scattered light from the white surround severely limits all discrimination below 2 cd/m^2.

The *4scalesBlack* estimates are very different from those in *4scalesWhite*. In *4scalesBlack*, the pie-shaped sectors with the highest luminance in each scale all appear light (Estimates: A = 100, B = 90, C = 80, D = 69). As shown in other experiments, the local maxima generate appearances that change slowly with luminance.[25] Nearby areas, with

less luminance, change more quickly (physiological simultaneous contrast).

One can think of the *4scalesBlack* experiment as a four-scale version of Land's Black and White Mondrian.[26] We have four different identical targets in four different illuminations. When the targets are isolated in the opaque surround, we have three different examples of whites and blacks generated by the same luminances [(For 147 cd/m^2, magnitude estimate = 17 in A and 87 in B), [(For 15 cd/m^2, magnitude estimate = 16 in B and 71 in C) [(For 1.8 cd/m^2, magnitude estimate = 10 in C and 68 in D)].

Also, we have four different luminances [1.06, 8.4, 63.5, and 414] that generated the same appearance [ME = 50]. The same holds for all magnitude estimates except for the lightest and darkest. As argued in the original work,[26] there is no correlation of a pixel's luminance with appearance. Models of appearance require a strong spatial component.

The opposite roles of physiological contrast and scattered light are seen in the difference appearances in white and black surrounds (Fig. 13).[27] In the presence of white surrounds, glare increases and appearances gets darker. These two targets change the amount of veiling glare, but do not measure the effect of glare because simultaneous contrast interferes. Different targets are necessary to understand the role of glare alone. We need pairs of targets that have constant contrast, but different dynamic ranges. Rizzi, Pezzetti, and McCann have studied single- and double-density targets with half-white and half-black surrounds.[28] The double-density transparency target squares the dynamic range, yet has only a small effect on the magnitude estimates of appearance. Out of a possible range of 6.0 log units of display density, observers estimate that blacks are 3.0 log units darker than white. Additional range serves no purpose. Further, the 3.0 log range of luminances equals the range of conventional transparency film.

6 Discussion

Veiling glare limits HDR imaging in two distinct ways. First, camera glare limits the luminance range that can be accurately measured (Section 3). Multiple exposures improve the quantization of digital records, but fail to accurately record scene luminance. Second, intraocular scatter limits the range of scene luminances falling on the retina (Section 4).

Accurate portrayal of scene luminances from camera images is both impossible to achieve and inessential to the visual process.

We were unable to make accurate camera estimates of scene luminance for the 4.3 log dynamic range scenes studied here. The comparison of white and black surrounds shows dramatic scene dependence. In addition, the camera flux estimates, when compared with actual flux, show a different error with each exposure. It may be tempting to look for some type of average-flux curve that represents data with smaller errors, but that idea is in conflict with the fundamental aim of the process; namely, recording accurate scene luminance. Multiple-exposure HDR is limited by veiling glare that is scene-, exposure-, lens-, aperture-, and camera-dependent. The accuracy of ME2SL scene-luminance estimates varies with all these parameters.

Some HDR algorithms attempt to correct for glare.[29–31] 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. 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 states unequivocally that "the reverse (deriving luminance from camera response) calculation is not possible."[9]

Claims are made that recent multiple-exposure HDR algorithms capture wider scene luminances or colors than previously possible.[32] These claims are severely limited by scene and camera veiling glare. As shown above, the designers of negative films selected a 4.1 log response range. That range exceeds the camera glare limit for almost all scenes.

Veiling glare for human vision is much worse than for cameras. Nevertheless, human vision has a much greater apparent dynamic range than camera systems. Humans can see details in highlights and shadows much better than conventional films and conventional electronic cameras can record. Although the rods and cones in the retina respond to more than a 10 log range, the ganglion cells that transmit the retinal response to the brain have only a 2 log range. There is no simple correlation between retinal quanta catch *at a pixel* and appearance.

The interplay between glare and physiological contrast is very complex. They act in opposition to each other, with physiological contrast tending toward canceling glare.[27] Since their mechanisms are so different, there is no actual image-wise cancellation, as seen when a negative image is combined with a positive one, so that they make a uniform image. Intraocular glare, because it is the sum of contributions from all other pixels, adds a scene-dependent low-spatial-frequency mask to the scene. The effect of that mask is seen in the *4scalesWhite* appearances. Discrimination is lost below 2 cd/m^2. By comparison, in *4scalesBlack* with minimal glare observers can discriminate the entire 4.3 log range. Physiological contrast is a high-spatial-frequency edge-based mechanism.[33–35] The human visual system synthesizes images from edge information, using local maxima as a reference.[26] The local maxima in *4scalesBlack* have estimates ranging over 100–69. Appearance below these maxima decrease rapidly with luminance. The physiological contrast-based HVS image synthesis renders the four-level targets almost the same, despite large changes in lumi-

nances. Physiological contrast builds significant visual differences from small luminance differences on the retina. It is easy to understand that if the surround around the 40 test sectors increases, then the glare increases, the magnitude of the luminance ratios at edges decreases, and physiological contrast makes the appearance of the smaller value darker. This counteraction of contrast and glare is fundamental to how we see. If the visual environment were only stars at night, then these low-glare scenes are easy to interpret because appearance tracks luminance.[25,36,37] As scenes change from the starry night to grays in a white surround, glare decreases the edge ratios on the retina and physiological contrast increases relative differences in appearance.

Often we see discussions of HDR image capture in which dynamic range is equated to digital bit depth. Luminance and bit depth are equivalent only in a very special case that rarely happens. This case requires that the digits resulting from camera-image-plane luminances fall on a slope 1.0 plot with luminance. In the history of silver halide film, the only example of slope 1.0 is *slide duplication* film. Consumers do not like accurate (slope 1.0) pictures. Linear, slope 1.0, response functions are rare in digital cameras. The fundamental mechanism of CCDs tells us that the number of quanta caught is proportional to the charge in the pixel well. However, charge in a pixel is not linearly proportional to the output digit. There are anti-blooming (high luminances) and noise reduction (low luminances), and tone-scale functions built into most digital cameras. Digital output value cannot be assumed to be proportional to image-plane luminance. A camera's dynamic range is not equal to pixel bit depth. As shown above, the camera's dynamic range has to be measured. The number of bits determines the number of quantization levels between the max and min. The number of quantization levels determines whether two slightly different luminances are reproduced as different or the same, which determines whether spatial detail in the scene is preserved or lost in the reproduction.

Improved digital quantization, which allows discrimination of adjacent objects, can be used in spatial comparison algorithms. Unlike analog film density responses, digital imaging, particularly in the 1960s and 70s was limited to the number of bits available in electronic-imaging devices. Although appropriately spaced, 24 bits of data per pixel is close to being able to record the entire range detectable to human vision, it lacks the digital resolution to handle computations, such as de-mosaicing, image processing, printing, and displaying a satisfactory picture. The number of bits, and more importantly, the luminance spacing between each bit is critical to having an artifact-free image.[14] If digital cameras had photon well sizes and pixel digitizer circuits with 4.1 log range comparable to silver halide film, then multiple exposures would not be necessary. Such devices are too expensive for modern camera markets and have not been developed. Multiple exposures provide the mechanism to effectively increase the number of quantization levels that improves spatial discrimination of camera luminance data. Multiple exposures improve local information – pixel A is darker than pixel B. That information is essential for synthesizing visual appearances of HDR scenes.

Summarizing, the significant role of intraocular glare has always prevented retinal receptors from seeing actual scene luminances. Physiological mechanisms, described as simultaneous contrast, work to reduce the adverse effects of glare. It follows that computational approaches to render HDR scenes for humans should use spatial comparisons as the essential tool in synthesizing the optimal display.[1,38–40] The best approach to HDR is to follow the lead of artists' spatial rendering techniques, but with computational rather than subliminal mechanisms.

7 Conclusions

This paper measures how much veiling glare limits HDR imaging in image capture and display. 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. We used 4.3 log dynamic range test targets and a variety of digital and film cameras. In each case, the camera response to constant luminances varied considerably with changes in the surrounding pixels. HDR image capture cannot accurately record the luminances in these targets. Second, we measured the appearance of the same targets. Appearance did not correlate with luminance at a pixel; it depended on physical intraocular glare and physiological contrast.

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. The improvement in HDR images is due to better preservation of relative spatial information that comes from improved digital quantization. Spatial differences in highlights and shadows are not lost. Spatial HDR image-processing algorithms mimic processes developed by human vision, by chiaroscuro painters, and by early photographers that renders HDR scenes in low-range outputs.[1]

Acknowledgments

The authors wish to thank S. Fantone, D. Orband, and M. McCann for their very helpful discussions. The authors also want to thank the guest editors and referees for their assistance.

References

1 J J McCann, "Art, science, and appearance in HDR images," *J Soc Info Display* **15**/9, 709–719 (2007).
2 P E Debevec and J Malik, "Recovering high dynamic range radiance maps from photographs," *ACM SIGGRAPH '97*, 369–378 (1997).
3 H Seetzen, W Heidrich, W Stuerzlinger, G Ward, L Whitehead, M Trentacoste, A Ghosh, and A Vorozcovs, "High dynamic range display systems," *ACM Trans Graphics* **23**(3), 760–768 (2004).

4 J J McCann and A Rizzi, "Optical veiling glare limitations to in-camera scene radiance measurements," in *ECVP 2006 Abstracts, Perception* **35**, Supplement, 51 (2006).
5 J J McCann and A Rizzi, "Spatial comparisons: The antidote to veiling glare limitations in HDR images," *Proc ADEAC/SID&VESA*, 155–158 (2006).
6 J J McCann and A Rizzi, "Veiling glare: the dynamic range limit of HDR images," *Human Vision and Electronic Imaging XII*, eds., B. Rogowitz, T. Pappas, and S. Daly, *Proc SPIE* **6492-41** (2007).
7 J J McCann and A Rizzi, "Spatial comparisons: The antidote to veiling glare limitations in image capture and display," *Proc IMQA* (2007).
8 The term *contrast* has different definitions in photography and vision. In photography, it refers to the rate of change in reproduction luminance *vs.* scene luminance. It is the slope of the tone-scale function. In vision, it is the spatial mechanism that enhances differences in appearance. A gray patch in a white surround is darker because of the physiology in the visual system, referred to as simultaneous contrast.
9 ISO 9358:1994 Standard, "Optics and optical instruments. Veiling glare of image forming systems. Definitions and methods of measurement" (ISO, 1994).
10 J W Gillon, U.S. Patent 2,226,167 (December 4, 1940).
11 S Ochi and S Yamanaka, U.S. Patent 4,541,016 (filed Dec. 29, 1982, issued Sept. 10, 1985).
12 L E Alston, D S Levinstone, and W T Plummer, "Exposure control system for an electronic imaging camera having increased dynamic range," U.S. Patent 4,647,975 (filing date: Oct 30, 1985, issue date: Mar 3, 1987).
13 S Mann, "Compositing multiple pictures of the same scene," *Proc IS&T Annual Meeting* **46**, 50–52 (1993).
14 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* **48**, 442–448 (1995).
15 J J McCann, "Calculated color sensations applied to color image reproduction," in *Image Processing Analysis Measurement and Quality, Proc SPIE* **901**, 205–214 (1988).
16 L A Jones and H R Condit, "The brightness scale of exterior scenes and the computation of correct photographic exposure," *J Opt Soc Am* **31**, 651–678 (1941).
17 C E K Mees, *Photography* (The MacMillan Company, New York, 1937).
18 C E K Mees and T H James, *The Theory of the Photographic Process*, 3rd ed. (The MacMillan Company, New York, 1966).
19 A L M Sowerby, *Dictionary of Photography*, 18th edn. (Philosophical Library, New York, 1956), p. 568.
20 The formula to count parasitic images (pi) is pi = $2n^2 - n$, where n is the number of lenses.[19]
21 This question was raised by Jim Larimer (personal communication).
22 P. Padley, "Diffraction from a circular aperture," ttp://cnx.org/content/m13097/latest/.
23 R W Ditchburn, *Light* (Dover, New York, 1991), p. 165.
24 E Reinhard, G Ward, S Pattanaik, and P Debevec, *High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting* (Elsevier, Morgan Kaufmann, Amsterdam, 2006), Chaps. 6 and 7.
25 J J McCann, "Aperture and object mode appearances in images," in *Human Vision and Electronic Imaging XII*, eds. B. Rogowitz, T. Pappas, and S. Daly, *Proc SPIE* **6292-26** (2007).
26 E Land and J J McCann, "Lightness and Retinex theory," *J Opt Soc Am* **61**, 1–11 (1971).
27 J J McCann, "Spatial contrast and scatter: Opposing partners in sensations," in *Human Vision and Electronic Imaging IV*, B. Rogowitz and T. Pappas, eds., *SPIE Proc* **3644**, 97–104 (1999).
28 A Rizzi, M Pezzetti, and J J McCann, "Intraocular glare controls the visible range of high dynamic range images (HDRI)," *Proc IS&T/SID Color Conference* (2007).
29 F Xiao, J M DiCarlo, P B Catrysse, and B A Wandell, "High dynamic range imaging of natural scenes," *Proc IS&T/SID Color* **10**, 337–342 (2002).
30 E Reinhard, G Ward, S Pattanaik, and P Debevec, *High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting* (Elsevier, Morgan Kaufmann, Amsterdam, 2006), Chap. 4.
31 B Bitlis, P A Jansson, and J P Allebach, "Parametric point spread function modeling and reduction of stray light effects," in *Computational Imaging V*, eds. C. Bouman, E. Miller, and I. Pollak, *Proc SPIE* **6498-27** (2007).
32 ttp://www.cs.nott.ac.uk/%7Eqiu/jvci/Call-for-Papers.html.
33 V O'Brien, "Contrast by contour-enhancement," *Am J Psychol* **72**, 299–300 (1959).
34 T Cornsweet, *Visual Perception* (Academic Press, New York, 1970).
35 E H Land, "Smitty Stevens' test of Retinex theory," in *Sensation and Measurement Papers in Honor of S. S. Stevens* (Applied Science Publishers, Ltd., London, 1974), pp. 363–368.
36 J J McCann, "Rendering high-dynamic range images: Algorithms that mimic human vision," *Proc AMOS Technical Conference*, 19–28 (2005).
37 H Takahashi, H Yaguchi, and S Shiori, "Estimation of brightness and lightness in all adaptation levels," *J Light & Vis Env* **23**, 38–48 (1999).
38 J J McCann, "Capturing a black cat in shade: The past and present of Retinex color appearance models," *J Electronic Img* **13**, 36–47 (2004).
39 A Rizzi, C Gatta, and D Marini, "A new algorithm for unsupervised global and local color correction," *Pattern Recognition Lett* **24**, No. 11, 1663–1677 (July 2003).
40 J J McCann, "Retinex at forty," *J Electronic Img* **13**, 1–145 (2004).

John McCann received his B.A. degree in biology from Harvard University in 1964. He managed the Vision Research Laboratory at Polaroid from 1961 to 1996. He has studied human color vision, digital image processing, large-format instant photography, and the reproduction of fine art. He is a Fellow of IS&T (1984). He is a Past-President of IS&T and the Artists Foundation, Boston. He is currently consulting and continuing his research on color vision. He received a Certificate of Commendation, Society for Information Display, 1996. He is the IS&T/OSA 2002 Edwin H. Land Medalist and IS&T 2005 Honorary Member.

Alessandro Rizzi received his B.S. degree in computer science from the University of Milano and received his Ph.D. in information engineering at the University of Brescia (Italy). He taught information systems and computer graphics at the University of Brescia and at Politecnico di Milano. He is currently an assistant professor, teaching multimedia and human-computer interaction, and a senior research fellow at the Department of Information Technologies at the University of Milano. Since 1990, he has been performing research in the field of digital imaging and vision. His main research topic is the use of color information in digital images with particular attention to color-perception mechanisms. He is the coordinator of the Italian Color Group.