# High Dynamic Range Imaging of Natural Scenes

*Feng Xiao, Jeffrey M. DiCarlo, Peter B. Catrysse and Brian A. Wandell*
*Stanford University*
*Stanford, California*

## Abstract

The ability to capture and render high dynamic range scenes limits the quality of current consumer and professional digital cameras. The absence of a well-calibrated high dynamic range color image database of natural scenes is an impediment to developing such rendering algorithms for digital photography. This paper describes our efforts to create such a database. First, we discuss how the image dynamic range is affected by three main components in the imaging pipeline: the optics, the sensor and the color transformation. Second, we describe a calibrated, portable high dynamic range imaging system. Third, we discuss the general properties of seventy calibrated high dynamic range images of natural scenes in the database (http://pdc.stanford.edu/hdri/). We recorded the calibrated RGB values and the spectral power distribution of illumination at different locations for each scene. The scene luminance ranges span two to six orders of magnitude. Within any scene, both the absolute level and the spectral composition of the illumination vary considerably. This suggests that future high dynamic range rendering algorithms need to account jointly for local color adaptation and local illumination level.

## Introduction

The management of data obtained from a scene with a large range of luminance is an important issue in image acquisition, analysis, and display [1, 2]. Difficulties in acquiring and rendering scenes that span a large luminance range are usually described by the informal phrase "system dynamic range" limitations.

The distribution of natural scene luminance often spans a range of 10,000 (4 $\log_{10}$ units) [3]. For example, in a typical indoor scene with a window, the luminance of the sky seen in the window is on the order of 10,000 cd/m2; the luminance of a face within the room may be 50 cd/m2; a dark surface in the room may be 1 cd/m2. This type of scene presents a challenge to the most recent consumer and even professional digital cameras. We have no trouble appreciating details in both the bright and dark regions because our visual system adjusts its sensitivity across fixation points [4, 5].

While it has been difficult to capture accurately the full luminance range in such scenes, recent advances in sensor circuit design have produced new imaging devices that can capture this luminance range [6, 7]. Output devices, however, cannot reproduce these large luminance ranges. They are the main limiting step in the imaging pipeline and likely to remain the main bottleneck for some time [1, 2]. Hence, it is important to develop image-rendering algorithms that transform scene images to display images by matching the scene appearance as closely as possible.

Research on high dynamic range rendering algorithms has mainly used synthetic scenes [3, 8, 9]; only a few high dynamic range images of natural scenes are available [10, 11]. Important scene and system information, such as the responsivity functions of the camera and the spectral power distribution of the illuminant, is not recorded with these images. People, the main subject of photography, are absent from most of these images. The presence of people and other familiar subjects will add more challenge to rendering algorithms [12]. Hence, we propose to develop an open database of calibrated, high dynamic range, color images that represent the typical scenes in digital photographs. The availability of these images should help the development and evaluation of rendering methods; the database also can be used to evaluate limitations imposed by image sensors and lens configurations.

In this paper, we describe our initial methodology for creating such a database. In developing these methods, we characterized three elements of the digital imaging pipeline that affect the dynamic range of captured image: the optics, the sensor and the color transformation. Then we created a portable high dynamic range imaging system and used it to take more than seventy high dynamic range scenes that are representative of challenging visual scenes. The spectral power distribution of illumination at different locations in each scene was also recorded. Finally, we summarize the properties of these scenes.

## Background

Theoretically, scene dynamic range is defined as:

$$DR_{scene} = \frac{Y_{\max}}{Y_{\min}} \qquad (1)$$

$Y_{max}$: Maximum scene luminance
$Y_{min}$: Minimum scene luminance

There are three important locations within the imaging pipeline that affect the estimation of scene dynamic range:

the optics, the sensor and the color transformation. The optics transforms the scene into an optical image, just prior to the sensor. The sensor transforms the optical image into electron charges that are quantized into a digital sensor image, and the color transformation converts the sensor image into a tristimulus image, which is rendered for display. We consider each of these separately.

**Optics**

The imaging optics transforms scene luminance into image illuminance at the sensor. Optical properties, including diffraction, aberrations, defocus and stray light or flare, generally compress the scene luminance range into a smaller illuminance range at the sensor. Diffraction is typically introduced by the finite size of the lens aperture and is characterized by the f-number. Even in well-corrected optics there can be some additional blur due to aberrations or due to defocus. These effects reduce the dynamic range by taking away photons from point like sources with a large luminance and spatially spread the photons. This can be modeled as a (position dependent) point spread function [13]. Two types of lens flare can also significantly reduce the range of the image [14, 15]. When reflections occur in the rear group of elements, a flare effect known as aperture ghosting results and copies of the aperture appear in the optical image. Veiling flare usually originates from reflections in the front group of elements and spreads out over a large area. The result is a haze that adds light to large parts of the optical image. Lens flare reduces the dynamic range of the image by introducing unwanted photons into dark regions. Dynamic range of the optical image can be defined as:

$$DR_{optic} = \frac{Y'_{\max}}{Y'_{\min}} = \frac{Y_{\max} - \Delta Y_{\max}}{Y_{\min} + \Delta Y_{\min}} \quad (2)$$

$Y'_{max}$: Maximum luminance in the optical image
$Y'_{min}$: Minimum luminance in the optical image
$\Delta Y_{max}$: Reduction of maximum luminance
$\Delta Y_{min}$: Addition of minimum luminance

The maximum scene luminance is typically reduced by point spread like effects, while the minimum luminance is increased by stray light. Understanding these effects and characterizing them is particularly important for capturing and rendering images with a high dynamic range. Our calibration efforts focus on lens flare because it affects the dynamic range more than the other lens effects.

**Sensor**

Sensor limitations influence the dynamic range of the captured image in two ways: saturation puts a cap on the maximum detectable intensity while sensor noise limits the minimum detectable intensity. Therefore, the dynamic range of sensor image becomes[1]:

$$DR_{sensor} = \frac{Min(Y'_{\max}, Y''_{\max})}{Max(Y'_{\min}, Y''_{\min})} \quad (3)$$

$Y''_{max}$: Input referred luminance level that fills up sensor well capacity
$Y''_{min}$: Input referred luminance level that equals the sensor noise floor level

From Equation 3, the dynamic range of the sensor image equals that of the optical image if and only if the maximum and minimum intensity levels of the optical image are within the range of sensor's capability. A simple formula for the dynamic range capability of a conventional sensor at exposure duration of $t_{int}$ is [6]:

$$DR_c(t_{\text{int}}) = \frac{Y''_{\max}}{Y''_{\min}} = \frac{i_{\max}}{i_{\min}} = \frac{q_{well} - i_d t_{\text{int}}}{\sqrt{q i_d t_{\text{int}} + \sigma_r^2}} \quad (4)$$

$q_{well}$: Well capacity in electrons
$i_d$: Dark current
$\sigma_r$: Read noise (reset and quantization noise)
$q$: Electron charge

For a quantization noise dominated sensor, Equation 4 can be simplified [16]:

$$DR'_c(t_{\text{int}}) = \frac{Y''_{\max}}{Y''_{\min}} \approx \frac{2^{bits}}{\sqrt{1/12}} \quad (5)$$

bits: Number of ADC bits

For sensors capable of integrating multiple captures with multiple exposure durations [5], the effective dynamic range with exposure durations of $t_{short}$ and $t_{long}$ is:

$$DR_m(t_{short}, t_{long}) = \frac{(q_{well} - i_d t_{short})}{\sqrt{q i_d t_{long} + \sigma_r^2}} \times \frac{t_{long}}{t_{short}} \quad (6)$$

When quantization noise dominates, Equation 6 becomes:

$$DR'_m(t_{short}, t_{long}) \approx DR'_c \times \frac{t_{long}}{t_{short}} \quad (7)$$

As the ratio between the longest and shortest exposure duration grows, multiple captures can increase system dynamic range dramatically. Hence, for an imaging system with a low-noise sensor capable of multiple captures, the main dynamic range limitation is the optics.

**Color Transformation from RGB to luminance**

Once we have sensor image (RGB values), it is a common practice to estimate luminance (Y tristimulus value) by a weighted sum of sensor RGB values [17-19]:

$$Y''' = w_1 \times R + w_2 \times G + w_3 \times B \quad (8)$$

In practice, these three weights can be negative. Coupled with noise, the estimated minimum luminance value can be negative for a significant proportion of scenes. Image

[1] To simplify following discussion, we assume a monochrome sensor with Y spectral response.

*Figure 1. Portable high dynamic range imaging system*

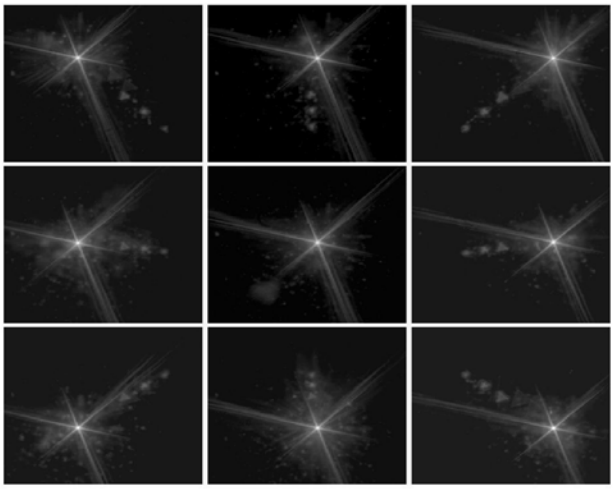

*Figure 2. Images of point light source at different locations on a log scale.*

dynamic range makes no sense for these cases. To avoid negative luminance value, these three weights are constrained to be nonnegative.

## Calibration Methods

To measure natural scenes, we created a portable camera system based on a QImaging Retiga 1300 scientific color CCD camera (Figure 1). The camera is equipped with a Computar manual-focus, manual-iris lens. A laptop computer controls the camera. A large battery on wheels powers the entire apparatus. The lens has 8.5mm focal length and F-number from 1.3 to 16. Angle of view is 57.4° horizontal and 43.8° vertical. The image sensor has a 10-bit linear output, 1024x1280 Bayer primary color pattern. The whole capture and analysis process is controlled through a Matlab interface.

**Lens Calibration**

To explore how aperture diffraction and lens flare influence system dynamic range, we measured the image of a point source at different locations (F-number ~ 16). As Figure 2 shows, there are three components to the image: (a) a star-like PSF, (b) triangular ghost images, and (c) haze. The star-like PSF corresponds to the diffraction from a triangular aperture. It spreads extensively across the image but the amplitude drops to 0.1% of peak value at 1.5 degree away from the star's center. The triangular aperture is visible as a series of ghost images. These images are created by three different lens surfaces from the group of lens elements behind the aperture. The amplitude of the ghost image at 20 degree away from its peak location is still 0.01% of the peak value. The veiling glare is small but spread throughout the image. As aperture size increases, the amplitudes of these three components decline slowly. In general, for scenes with large bright areas or very bright illuminants (sun for example), the optical components set a limit on the effective dynamic range of the capture system. Hence, the dynamic range is scene-dependent.

**Sensor Calibration**

We verified the linearity of the sensor response and measured its spectral responsivity (Figure 3). Table 1 shows the measured dark current and noise property of the CCD sensor at room temperature (25°). Dark current is small for exposure durations shorter than 10 second.

Table 1: Noise property of sensor

| Well capacity | Readout Noise | Dark Current |
|---|---|---|
| 1024 DV | < 0.3 DV | <0.03 DV/s/pixel |
| 36000 e- | < 10.5 e- | <1e-/s/pixel |

**Color Transformation Calibration**

We calculated the color transformation weights using method similar to that in [17] but with nonnegative constraints. Scene is assumed to be Macbeth Color Chart under four reference illuminants (D65, D50, Cool white Fluorescent and Tungsten A). The optimal weights are (0.2552, 1.0834, 0).

**Data Acquisition: Multiple Capture Imaging**

For each scene, we acquired multiple pictures at different exposure durations spaced by multiples of four (Figure 4). The shortest exposure duration was chosen to make the brightest pixels close to the saturation level or at least 0.1 ms; the longest exposure duration was chosen to make the darkest pixels at least 10 digital counts or at most 6 seconds. From equation 3, we obtain a very high effective

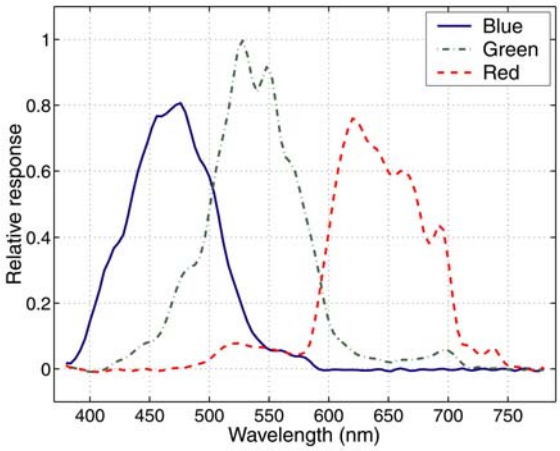


*Figure 3. Camera spectral responsivity*

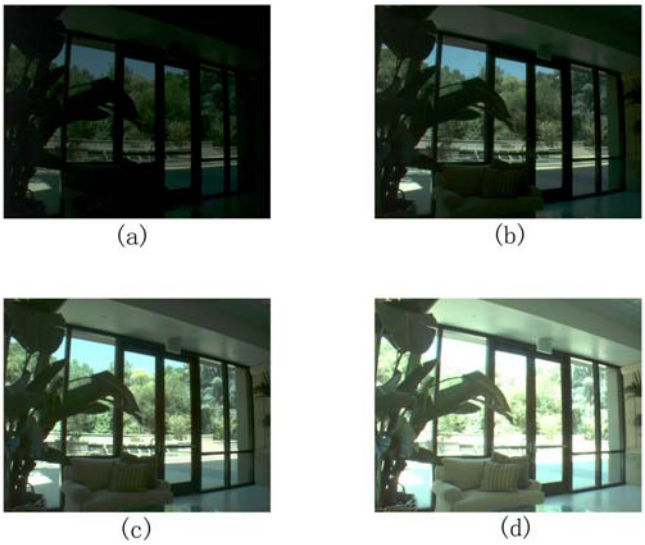


*Figure 4. Multiple pictures taken at different exposure durations (a) 4ms. (b) 16ms. (c) 64ms. (d) 256 ms*

sensor dynamic range using the multiple capture architecture. The minimum detectable input luminance of this system is less than 0.001cd/m$^2$ (F-Number = 1.3 and exposure duration= 6 seconds) and the saturation luminance intensity is greater than 1,000,000 cd/m$^2$ (F-Number ~= 16 and exposure duration = 0.1 ms). This range covers the whole luminance range for the scenes we measure. This multiple capture architecture differs from the conventional use of sensors and bypasses the usual bottleneck at the sensor itself.

The data from each pixel includes multiple samples at different durations. Apart from noise and saturation, these samples should fall on a line whose slope defines the scene intensity. However, saturated samples and motion distortions cause some of the data to deviate from a line. We use a simple motion detection algorithm to exclude moving samples (people, rustling leaves). Similarly, we remove all samples after saturation. Linear regression is used to find the optimal RGB values for the remaining valid pixels. The maximum recording error for most pixels (excluding darkest pixels) is less than 0.4% (10bit, exposure interval of four).

For most of the high dynamic range scenes, typically two illuminants are present. We recorded the spectral power distribution of these two illuminants by measuring a white patch put into bright and dark regions respectively.

To map the camera measurements to absolute luminance levels, in the lab we placed a white surface under a tungsten light source (color temperature 3000K). Then we measured the luminance of the patch with a PR-650 spectroradiometer. We repeated these measurements for a series of aperture settings that spanned those used in acquiring natural scenes. From these lab measurements, we derive the correspondence between luminance levels and camera RGB values for different aperture settings and exposure durations.

## Results

**Scenes**

We measured more than 70 high dynamic range images of indoor and outdoor natural scenes around Stanford University campus at different times of day. People are the main subjects for one third of the scenes. The luminance level ranges from 0.001 cd/m2 to 400,000 cd/m2 (Figure 5). The highest luminance is from sunlight glare and the lowest is from night scenes. When we compare the luminance ranges measured across individual pixels, the image dynamic ranges span two to six orders of magnitude.

We also measured the dynamic range of local spatial neighborhoods within the images in the database. The dynamic range within single blocks is generally smaller than the dynamic range of the entire image (Figure 6). The local dynamic range increases as the size of the image block increases. This observation suggests that it will be possible to devise rendering algorithms for high dynamic range images that preserve the local spatial contrast while compressing the overall scene dynamic range.

Figures 7 and 8 show scenes which users find challenging with respect to camera dynamic range. Figure 7 is a typical indoor scene with strong backlighting. The intensity of outdoor daylight is approximately 180 times higher than the intensity of indoor fluorescent light. Figure 8 is a typical outdoor scene. The luminance in the bright sunlight is approximately 20 times more than that in the Shadow. The dynamic ranges of these scenes are moderately high (2000:1 for Figure 7 and 400:1 for Figure 8), but rendering images containing multiple illuminants that combine different mean levels and spectral composition is a significant challenge.

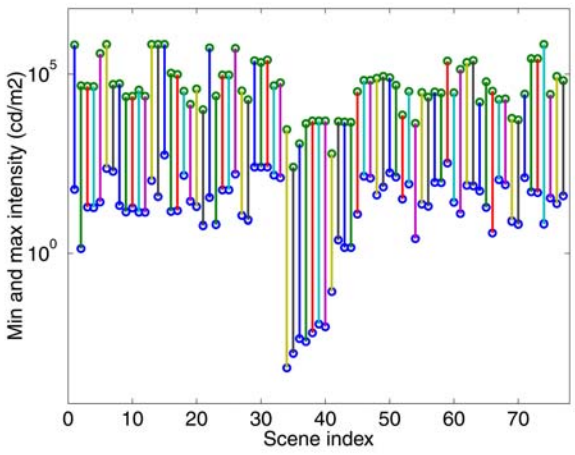


*Figure 5. Maximum and minimum pixel luminance of each scene.*

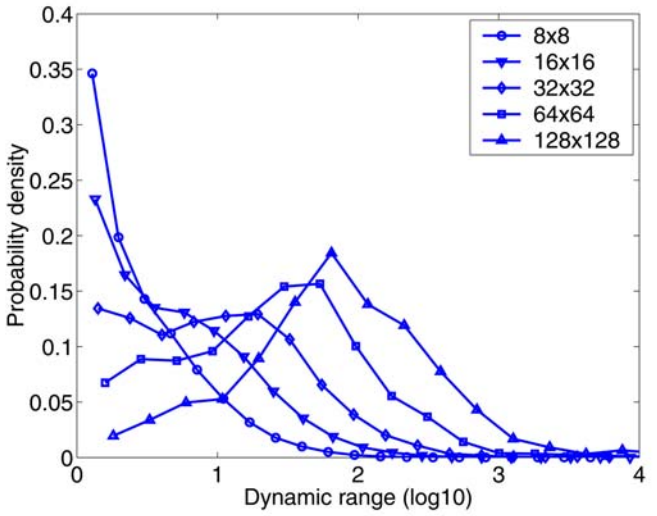


*Figure 6. Histogram of local dynamic range when the images are divided into regions of different size.*

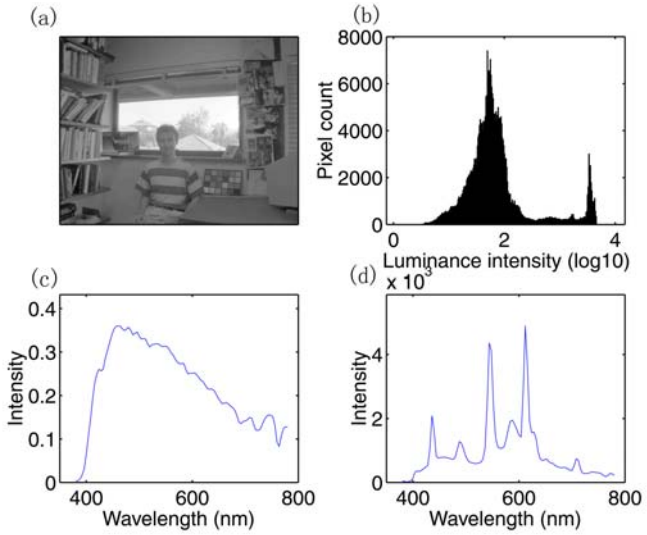


*Figure 7. Indoor scene with strong backlighting. (a) Luminance intensity map (log10). (b) Histogram of luminance intensity (log10). (c) Spectral power distribution of a white surface under bright sunlight. (d) Spectral power distribution of the same surface under fluorescent light.*

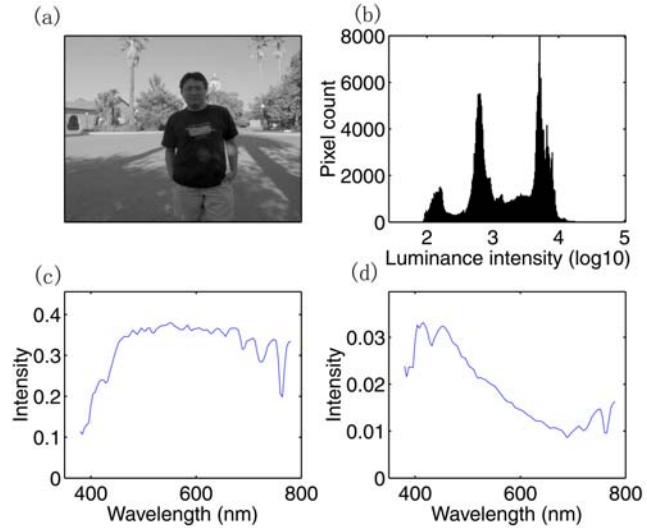


*Figure 8. Outdoor scene with shadow. (a) Luminance intensity map (log10). (b) Histogram of luminance intensity (log10). (c) Spectral power distribution of a white surface under bright sunlight. (d) Spectral power distribution of the same surface in shadow*

**Data Format**

There are several widely used file format for high dynamic range images: Pixar's 33-bit/pixel Log Format, Radiance's 32-bit/pixel RGBE Format and SGI's 32-bit/pixel LogLuv Format [20]. We save our data in two formats: one in the Radiance RGBE Format because of its simplicity and one in our own 56-bit/pixel Format for higher accuracy. Extra information about illuminants and sensor is saved as extra tags. All these images can be accessed at http://pdc.stanford.edu/hdri/.

## Summary

We have described our efforts to create an open database of calibrated high dynamic range images of natural scenes.

Across all images, the absolute luminance varies more than eight orders of magnitude. Within individual images, the dynamic range spans two to six orders of magnitude. These data describe typical spatial variations in illumination, both in luminance and spectral composition. The data suggest that it will be important to combine both luminance compression and color adaptation to render accurately these high dynamic range scenes.

## Acknowledgements

We thank Joyce Farrell, Ulrich Tobias Barnhoefer and David Cardinal for their help and comments on the paper. We also thank numerous people for their willing to be pictured. This work was supported by the Programmable Digital Camera project at Stanford University, sponsored by Agilent, Canon, Hewlett-Packard and Kodak.

## References

[1] J. M. DiCarlo and B. A. Wandell, Rendering High Dynamic Range Images, Proc. SPIE: Image Sensors, pg. 392-401 (1999).

[2] G. Ward, High Dynamic Range Imaging, Proc. 9th IS&T/SID Color Imaging Conference, pg. 9-16 (2001).

[3] S. Pattanaik, J. Ferwerda, M. Fairchild, and D. Greenberg, A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display, Proc. SIGGRAPH'98, pg. 287-298 (1998).

[4] B. A. Wandell, Foundations of Vision: Sinauer Associates, 1995.

[5] B. A. Wandell, A. El Gamal, and B. Girod, Common Principles of Image Acquisition Systems and Biological Vision, Proc. IEEE, vol. 90, 5-17, (2002).

[6] D. Yang, B. Fowler, A. El Gamal, and H. Tian, A 640x512 CMOS Image Sensor with Ultrawide Dynamic Range Floating-Point Pixel-Level ADC, IEEE Journal of Solid State Circuits, vol. 34, 1821-1834, (1999).

[7] S. K. Nayar and T. Mitsunaga, High Dynamic Range Imaging: Spatially Varying Pixel Exposures, Proc. IEEE CVPR, (2000).

[8] K. Chiu, M. Herf, P. Shirley, C. Wang, and K. Zimmerman, Spatially Nonuniform Scaling Functions for High Contrast Images, Proc. Graphics Interface'93, pg. 245-253 (1993).

[9] G. W. Larson, H. Rushmeier, and C. Piatko, A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes, IEEE Transactions on Visualization and Computer Graphics, vol. 3, 291-306, (1997).

[10] P. E. Debevec and J. Malik, Recovering High Dynamic Range Radiance Maps from Photographs, Proc. SIGGRAPH'97, pg. 369-378 (1997).

[11] R. Fattal, D. Lishinski, and M. Werman, Gradient Domain High Dynamic Range Compression, Proc. SIGGRAPH'02, (2002).

[12] K. Topfer and R. Cookingham, The Quantitative Aspects of Color Rendering for Memory Colors, PICS 2000: Image Processing, Image Quality, Image Capture, Systems Conference, pg. 94-98 (2000).

[13] W. J. Smith, Modern Optical Engineering: the Design of Optical Systems, 2nd ed: McGraw-Hill, Inc., 1990.

[14] G. Kuwabara, On the Flare of Lenses, Journal of the Optical Society of America, vol. 43, 53-57, (1953).

[15] S. Matsuda and T. Nitoh, Flare as Applied to Photographic Lenses, Applied Optics, vol. 11, 1850-1856, (1972).

[16] G. C. Holst, CCD Arrays Camaeras and Displays, 2nd ed, 1998.

[17] B. A. Wandell and J. E. Farrell, Water into Wine: Converting Scanner RGB to Tristimulus XYZ, Proc. SPIE, (1993).

[18] G. D. Finlayson and M. S. Drew, The Maximum Ignorance Assumption With Positivity, Proc. 4th IS&T/SID Color Imaging Conference, pg. 202-205 (1995).

[19] P. M. Hubel, J. Holm, G. D. Finlayson, and M. S. Drew, Matrix Calculations for Digital Photograpphy, Proc. 5th IS&T/SID Color Imaging Conference, pg. 105-111 (1997).

[20] G. W. Larson, LogLuv Encoding for Full-gamut, High-dynamic Range Images, Journal of Graphics Tools, vol. 3, 15-31, (1998).