1 Introduction

High-dynamic-range (HDR) scenes are the result of nonuniform illumination falling on reflective material surfaces. In fact, material surface reflections have a very low dynamic range (LDR). Light falling on any surface does one of three things:

- A fraction reflects off the front surface of the object, i.e., specular reflection.
- A fraction reflects from the subsurface material the object is made of, i.e., diffuse reflection
- A fraction is absorbed by the material, i.e., absorption

Specular reflections from front surfaces are about 3% to 5% of the incident light. Diffuse reflection happens when illumination is scattered within the material itself and back out of the material. We see objects by sensing the light reflected to our eyes. Many diffuse reflections send light over a very wide angle. The amount of pigment, dye, or natural colorant in a material controls how much of the illumination is diffusely reflected from an object. White, gray, and black papers with matte surfaces are a good example of the range of material reflections. White paper reflects nearly 100% of the incident light. Black papers reflect about 3%. Adding more black pigment to a black paper does not lower the percent reflectance. The 3% limit is caused by front surface reflections. Adding a flocked surface, such as black velvet, will lower the amount of reflected light but only by a small amount. The dynamic range of objects in real scenes is thus quite small, about 32:1 or 2^5 in binary notation (Fig. 1).

Although front surface reflections limit the range of light from reflected objects, illumination can have extremely large dynamic ranges. It can be the ratio of a very bright light source itself to no emitted light. The HDR imagery discussed in this book refers to the scenes found in our natural environment with particular interest in photographic scenes, their capture, and reproduction.

Figure 2 is a photograph of a bunkhouse; its surface is uniformly coated with white paint. Although its surface reflectance is perfectly uniform, the pattern of reflected light is far from uniform. The side of the bunkhouse on the left shows the glossy appearance of specular reflected light from other parts of the scene. Brewster’s angle surface reflections, called *surface glare*, are polarized light. They can be removed by linear polarizers crossed to the orientation of the surface glare.

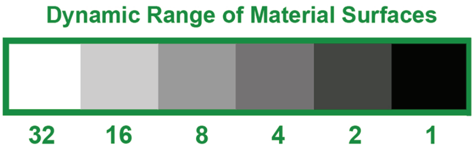

Figure 1 The step gradient of 32:1 illustrates the range of diffuse surface reflectances.

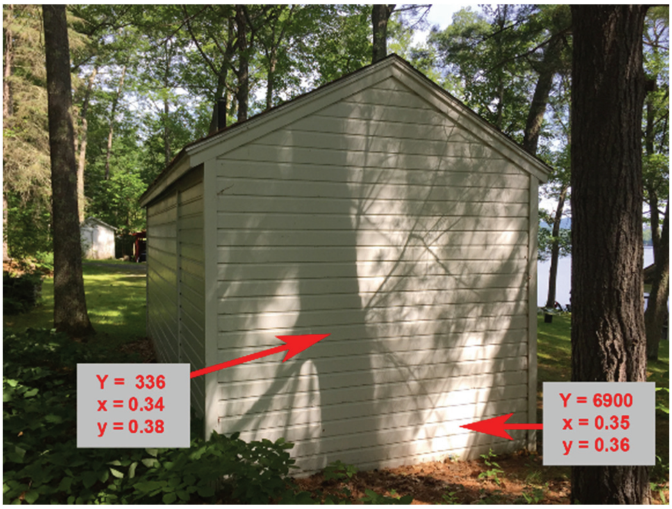

Figure 2 A photograph of a bunkhouse.

Although this is a typical natural scene, we see that nonuniform illumination increases the scene's dynamic range. The sunlight *luminance* of the white paint is 6900 cd/m^2, whereas the same paint in the shadow of the tree is 336 times that value.

An early HDR experiment observed that one can find situations in which the range of illumination in "sun and shade" can equal that of the "range of objects' reflectances." On a sunny day in Yosemite Valley, California, spot meter measurements showed that the shadows were 5 stops or 32 times darker than the sunlight. That meant that the white paper in the shade had the same light-meter reading as the black paper in the sunlight (Fig. 3).[2]

The HDR scene *John at Yosemite, 1981* is a problem for standard photography. The best exposure for the white paper in shade (Fig. 3(a)) renders the entire ColorChecker® as detail-free white. Furthermore, the best exposure for the black ColorChecker square (Fig. 3(b)) renders the entire shaded area much too dark. Ordinarily, we would use tone scale adjustments to attempt to improve the reproduction rendition of this scene. The special challenge that this scene presents is that both white in sunlight and black in shade papers have the same luminance and thus identical camera output digit values. A tone scale manipulation can improve the rendition of the white paper in the shade, but the sunlit portion becomes more overexposed. Alternatively, a very different tone scale can improve the rendition of the black in the sun, but the shaded portion becomes more underexposed. It is impossible to find a tone scale function for all identical pixel values that improves the rendition of both white and black papers in this scene. This HDR scene requires spatial image processing analogous to that found in human vision (Fig. 3(c)).

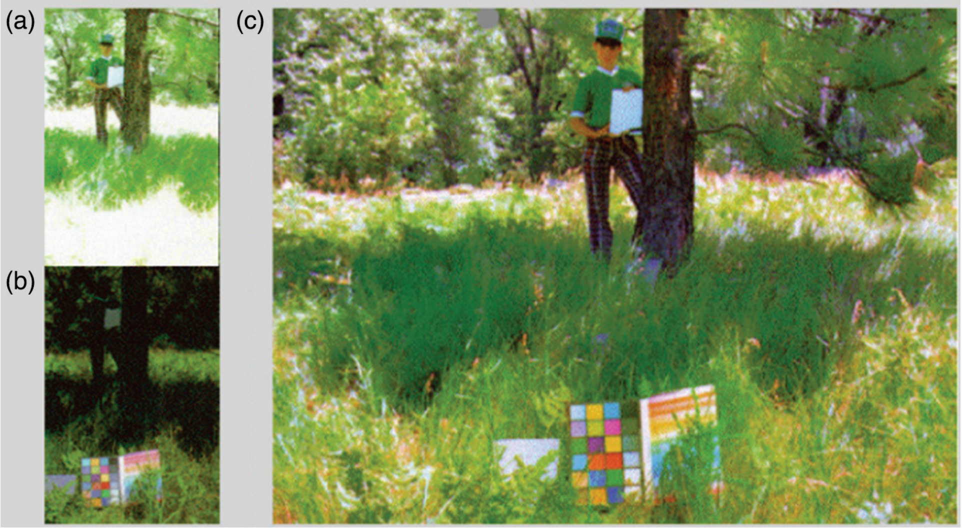

Figure 3 (a and b) Conventional prints made with high and low exposures. Meter readings of the scene showed equal radiances from the white paper under the tree and the black paper in the ColorChecker® in sunlight. (c) Retinex HDR spatial image processing. While conventional photography reproduces parts of the HDR scene, spatial processing can render the appearance of the entire scene using LDR media.

1.1 Goals

The field of HDR imaging has three major branches:

- Branch 1: Paint an image using human visual feedback.

  Since the Renaissance, painters have put down on canvas the combination of oil paints that render the appearance of HDR scenes. Their vision provided the spatial image-processing feedback that determined the equivalent radiance information. Such reproductions did not reproduce scene radiances, just the appearances generated by the artist's visual system. The artist rendered each scene individually using a unique transformation of scene radiance to reproduction output. Human visual feedback with eye/hand coordination is the underlying mechanism. The painter uses unique *Visual Inspections* to control the reproduction of each part of each image.

- Branch 2: Calculate appearances from camera data: Write sensations on LDR media.

  Using the best practices of Adams and his zone system, we can capture the entire range of information available on the camera's image plane. Using a *Vision-based Model* that incorporates the neural spatial processing that takes place after the rods and cones in the retina, one can calculate the sensations that humans see when observing the original HDR scene. These calculated sensations can render the scene in LDR media.

- Branch 3: Capture and accurately reproduce scene radiance.

  Using *Multiple Exposures* to measure the camera response function (CRF), one can attempt to remove the camera's transformations of scene radiances. Using HDR displays, one can attempt to accurately reproduce the actual scene radiances. It goes without saying that if we make an exact scene reproduction, it must generate all of the radiances in the entire scene. That system must then reproduce the scene's appearance. This tautological argument holds only for that special case in which every pixel in the scene is accurately reproduced. Any errors in reproduction invalidate it.

1.1.1 Paint an image using visual feedback

The first HDR imaging goal was achieved with the development of chiaroscuro painting in the Renaissance. Painters added the appearance of illumination, as well as the appearance of objects, to their images. Leonardo da Vinci, Caravaggio, Rembrandt, van Honthorst, Constable, and Martin, for example, synthesized HDR scenes in LDR media, i.e., oil on canvas.[3]

Figure 4 shows Gerritt van Honthorst's HDR chiaroscuro painting. The dynamic range of the image is the low-reflection range of paints. Nevertheless, the LDR painting renders the appearance of the HDR scene very successfully.

This technique simply adjusts the paint on the canvas until the scene has the desired appearance. There is no attempt to capture and reproduce scene radiances.

Figure 4 Gerritt van Honthorst's *The Childhood of Christ, 1620*, Hermitage. Photograph©.

The spatial manipulation of the LDR media is designed to partner with the viewer's visual system. These paintings do not reproduce scene radiances; rather, they reproduce the scene's sensations, as observed by the artist.

The first applications of Multiple Exposures in AgX photography are found in the early 1850s. Edouard Baldus, a trained painter, was an early member of Benito de Monfort's Societe de Heliographique in France. He extended Talbot's calotype process by replacing wax with gelatin in the negative. He made the print shown in Fig. 5 in 1853 using 10 negatives.[4]

During the mid-19th century, homemade AgX emulsions had limited dynamic range. Taking a series of photographs with different exposures recorded different light ranges of the scene. Combining these exposures resulted in images of the entire scene's range. There were many examples of multiple-exposure techniques until the mid-19th century.

George Eastman, founder of Kodak, persuaded C. E. K. Mees to move from London to Rochester in 1912 to become Kodak's Director of Research. Mees described the use of Multiple Exposures to extend range in his first edition of the *Fundamentals of Photography*, published by Kodak in 1920.[5] Also, in this book, Mees introduces the term "tone scale" as the amateur-photographer-friendly substitute term for CRF. For achromatic images, it was today's equivalent of a one-dimensional lookup table (1-D LUT).

Mees and his colleagues measured scene radiances, camera optics, and film response functions. Using that data, they designed AgX emulsions that extended a film's dynamic range to exceed the range of light falling on the film in cameras.

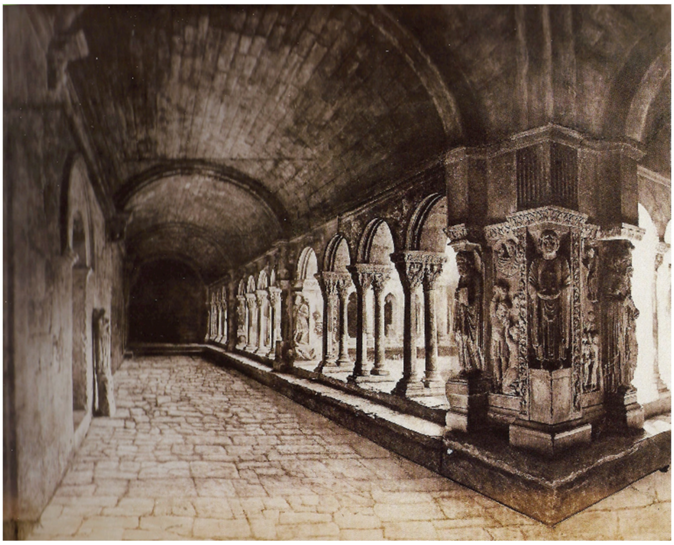

Figure 5 Baldus's *Cloisters of the Church of St. Trophime, Arles*, 1853, U. Texas, Austin, made from 10 paper negatives.

Negative films manufactured in the second half of the 20th century had greater range than most scenes' optical images on film. Single exposures that captured all possible image information were very important for making photography more convenient.

Ansel Adams was the unique combination of skilled technician and superb artist. His zone system provides detailed instructions about how to capture scene information. It began with what Adams called the "visualization" of the final print. His 1983 book *Examples: The Making of 40 Photographs*[6] provides many fascinating descriptions of this process.

In visualizing and capturing the scene information, Adams

- used a spot photometer to measure scene luminances,
- mentally assigned scene regions to different tone scale values or zones, and
- selected the specific combination of the exposure and film development procedure for that individual photograph.

By over exposing and under developing the negative, Adams lowered the negative's response slope, thus extending its dynamic range. By under exposing and over developing it, he raised the slope and reduced its dynamic range. Adam's zone system individually tuned his camera's response function for each scene. Adams, a concert pianist in his youth, described the exposed and developed negative as the "score."

In making the print, Adams

- made a test print to find the best baseline exposure,
- spatially manipulated local exposures to render his visualization of the scene,
  - locally dodged (hold back exposure) with an out-of-focus mask to lighten a local region, and
  - locally burned (increase exposure) with a moving baffle to darken a local region,
- developed the print.

Adams described his procedure of making the print as the "performance of the score."

In his scene-capture step, Adams used AgX emulsions that had wide-range logarithmic response functions that accurately captured the relationships of scene radiances. He used large-format cameras with lens shades that minimized optical glare. His score was a highly accurate record of scene information. The dynamic range compression of the score was achieved by the spatial manipulation of the local exposure in the "performance."

Adams used his mastery of technical photography as his tool, his paint brush, to render his aesthetic intent. He never reproduced scenes. He captured their spatial information and rendered his visualization." He used his remarkable photographic skills to synthesize his art.

1.1.2 Calculate appearances from camera data: write sensations on LDR media

Early digital HDR algorithms in the late 1970s (illustrated in Fig. 3) captured a wide range of scene radiances, calculated appearance (using a model of spatial vision), and rendered calculated sensations using standard LDR photographic film. The goal here was to have an algorithm that mimics fine-art painters in making the HDR renditions. The essential element of this approach is a model of vision that can calculate sensations observed in complex, real-life scenes.

A model of vision that successfully calculates appearance for all scenes provides a general solution to HDR imaging based on scene radiances. It removes the need for human, scene-by-scene, Visual Inspection. These spatial image-processing algorithms replaced the role performed by painters and photographers, namely, the rendering of sensations. These algorithms did not attempt to render the aesthetic intent of an artist. Their intent was to provide a computational algorithm to render HDR scenes.[2]

1.1.3 Reproduce radiances

The third HDR-imaging goal used camera Multiple Exposures and computer algorithms to accurately measure scene radiances. In 1997, Debevec and Malik[7] measured the CRF of a camera and calculated its inverse function (−CRF). They used it to digitally remove all of the transformations introduced by the camera. This would allow us to capture accurate scene-radiance information from complex, real-life scenes.

The second half of this approach, i.e., reproduce radiances, created a new kind of display with a much greater range of radiances. A number of HDR displays followed. One idea combined light-emitting diode (LED) illumination and liquid-crystal display (LCD) light absorption. A field of modulated LEDs was used to illuminate a second two-dimensional (2-D) modulated LCD transparent display. The sum of both modulated images synthesized an HDR display. This hybridization of LCD and LED technology produced a variety of HDR displays.[8] The hypothesis was that one could use calibrated cameras to measure actual scene radiance and then use HDR display technologies to reproduce the entire image with actual scene radiances.

This Spotlight does not attempt to describe the field of HDR displays. Display technology and recent commercial products (HDR TV) use high luminances, novel screen materials (organic LED (OLED), quantum dots, and quantum-dot LED (QLED)), and competitive, new tone scale broadcast standards. Each of these display topics is too extensive to include here.

1.2 Different ground truths for different goals

We have just described three very different goals. Each has a distinctly different measure of success. Each goal has a different ground truth.

1.2.1 Paint an image using Visual Inspection

In the first approach, empirical observation is the technique painters and photographers used in learning how to render HDR scenes. It is also the technique we use in Photoshop® to improve an image. We can manipulate lookup tables that convert camera input so as to improve the appearance on the screen. Lookup tables manipulate tone scale values one pixel at a time. Additionally, we can use spatial filters to improve the appearance of the rendering.

The measure of success is simply Visual Inspection. For a painter, the question is: Does the paint look right? For a computer-algorithm designer, the question is the same: Does the screen look right? The best reproduction is the image that gets the highest test scores in observer preference experiments. Do people like the appearance of the result? The ground truth of a Visual Inspection is observer preference.

1.2.2 Calculate appearances from camera data: write sensations on LDR media

In the second approach, an algorithm calculates all of the sensations in the image. It requires a Vision-based Model. We need to verify that the model is accurate in predicting appearance. We need to use psychophysical techniques to measure observer sensations (matches to standards to quantify sensations). This approach's ground truth is the direct comparison of observer matches with model predictions of calculated sensation.

The advantage of calculated sensations is that it is a general solution for all scenes based on their distribution of radiances. A single algorithm, done well, can render both HDR and foggy LDR scenes, using the same algorithm. It can render both sunny and cloudy days. A successful Vision-based Model can render any scene in LDR media. It can do what painters do. However, it does this with computational algorithms instead of manual human manual/visual feedback. Here, the algorithm replaces the painter's brush and all of the skill behind it.

The ground truth of a Vision-based Model is quantitative data acquired by observer matches. Matches are used to measure the model's accuracy over a full range of scenes: HDR, LDR, color constancy, and visual illusions.

1.2.3 Capturing and accurately reproducing scene radiances

The third approach uses a purely physics-based HDR technique. Its goal is the accurate capture and reproduction of scene radiances. It is easiest to verify. It just requires physics-based measurements of the amount of light at specific points in the scene. Then, we need a direct comparison of meter measurements with the calculated radiance values from the camera image. The problem is difficult because real scenes are made up of millions of points of light. Nevertheless, the measurement of scene radiance is strictly a physics-based problem.

The ground truth of a camera's radiance capture calculation is the measured radiance. In order to use a camera as a light-measuring instrument, we need physics-based experiments that verify its accuracy and establish the camera's limits in optics, sensor reciprocity, and system linearity. In the next few sections, we will describe measurements of camera responses in order to understand the limits of physics-based photography.

1.3 Summary

This section describes three very different approaches for making HDR reproductions. Further, it describes the three different ground truths that are needed to measure the success of each approach.

- Painting is usually described as an art rather than a scientific process. How a painter makes an HDR reproduction of a scene is difficult to understand. A painter uses Visual Inspection to validate each part of every image. Ground truth is the observation that it looks right.
- Calculating sensations from captured scene radiances involves both physical and psychophysical disciplines. This approach uses scene radiances as input. It calculates the appearance of the scene and displays its sensations. Ground truth is psychophysical measurements of appearance. Those data describe the properties of a successful computational Vision-based Model. That model is the central core of HDR algorithms. We will return to this combined study later in Part II of this Spotlight book.
- Capturing scene radiance is the foundation of all HDR engineering. The idea is simple. Capture and reproduce the actual scene. Ground truth is radiance measured with a meter. Part I describes the physics-based limitations of accurate scene capture. The problem is more complicated and more interesting than is apparent at first glance.

PART I: THE PHYSICS OF SCENE CAPTURE

2 Multiple Exposures

From childhood, everyone understands that cameras capture light from scenes and reproduce that light. The idea of reproduction is so well-established that any more complicated analysis seems to be contrary to everything we know. Photographic images are everywhere, contributing to everything we do. However, careful examination shows that photos do not accurately reproduce what we see. Usually, the photographs have higher contrast and are more colorful. The photos' shortcomings are that the highest and lowest luminances lack information and discriminable details. Nevertheless, these images are useful and acceptable records of what we see.

There is a related question, namely, do photographs accurately reproduce the radiances from the scene? This question has little practical interest. It has only theoretical importance for those who want to understand the relationship between radiances coming from the scene and their reproduction. In fact, the reproduction is a substantial transformation of scene radiances that is carefully engineered to