Title
Perceptually Based Tone Mapping for Low-Light Conditions

Author
Kirk, Adam

Publication Date
2011

Peer reviewed|Thesis/dissertation

Perceptually Based Tone Mapping for Low-Light Conditions

by

Adam Garnet Kirk

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

James F. O'Brien, Chair
Maneesh Agrawala
Martin S. Banks

Fall 2011

Perceptually Based Tone Mapping for Low-Light Conditions

Abstract

Perceptually Based Tone Mapping for Low-Light Conditions

by

Adam Garnet Kirk

Doctor of Philosophy in Computer Science

University of California, Berkeley

James F. O'Brien, Chair

We present a perceptually based algorithm for modeling the color shift that occurs for human viewers in low-light scenes. Known as the Purkinje effect, this color shift occurs as the eye transitions from photopic, cone-mediated vision in well-lit scenes to scotopic, rod-mediated vision in dark scenes. At intermediate light levels vision is mesopic with both the rods and cones active. Although the rods have a spectral response distinct from the cones, they still share the same neural pathways. As light levels decrease and the rods become increasingly active they cause a perceived shift in color. We model this process so that we can compute perceived colors for mesopic and scotopic scenes from spectral data. While our tone mapping operator works with spectral data from any source, we show how to produce spectral data of static scenes using multiple images and a camera with known spectral sensitivity. If the spectral sensitivity isn't provided by the camera manufacturer, we describe a one-time calibration procedure to estimate the sensitivity. Should obtaining spectral data of a scene be infeasible, we also describe how the effect can be approximated from high dynamic range RGB images by learning a mapping from RGB to rod and cone responses. Once we have determined rod and cone responses either directly or through our approximation process, we map them to RGB values that can be displayed on a standard monitor to elicit the same color perception when viewed photopically. Our tone mapping method focuses on computing the color shift associated with low-light conditions and leverages current HDR techniques to control the image dynamic range. We include results generated from both spectral and RGB images as well as experimental data.

To my parents, Christopher and Kathleen Kirk.

Contents

List of Figures

Acknowledgments

There are a great many people I have to thank for their contributions to my academic successes, and for most this is better done over a stiff drink. But there are a select few for whom I would like to publicly express my gratitude. These are all people who have fundamentally changed the course of my academic and professional life, and who made my time at Berkeley far more interesting, productive, creative, introspective, and just plain fun.

First of all, I would like to thank my graphics and vision colleagues. I was fortunate enough to be surrounded by a group of some of the finest engineers and researchers I have ever met, who all started as academic peers and ended as friends. In particular, I wish to thank: Adam Bargteil, Andrea Frome, Ashley Eden, Bryan Feldman, Bryan Klingner, Chen Shen, Deva Ramanan, Floraine Berthouzoz, Hayley Iben, Jason Sanders, Lillian Chu, Leslie Ikemoto, Mira Dontcheva, Pushkar Joshi, Ryan White, Tolga Goktekin, and Wilmot Li for their support and inspiration.

A few of my graphics colleagues warrant special mention. Anthony Lobay has been a great friend through many phases of my time at Berkeley, and in particular co-developed our annual camping trip Pinecrest which has been an outstanding success. Okan Arikan has been not only a great friend but also a great mentor, and is one of the finest researchers I know.

I also was lucky enough to meet some outstanding colleagues from outside graphics and vision, and wish to thank: Alejandro De La Fuente, Alic Chen, Christine Ho, David Garmire, Jon Burns, Lisa Fowler, Scott Lederer, and Steven Volkman. All of these people were kind enough to share the graduate school experience, introduce me to new subjects, and I count them all as friends.

I had several faculty mentors at Berkeley who fundamentally influenced my research ideas. I'd like to thank Maneesh Agrawala, Marty Banks, and David Forsyth for all of their insights and patience. And I would especially like to thank my advisor, James O'Brien, for all of his help and support.

I would also like to thank Kaccie Li and Susan Kim for their help with this dissertation. Kaccie was kind enough to help with the camera calibration work, and Susan (who is one of the finest artists I've ever met) donated her artwork for my examples.

In addition to the people I met academically, I also was fortunate to make some great friends during my time in the bay area. Genevieve Breed and Tynia Yang have been incredibly supportive through this process and made my time in graduate school much more fun. And I especially want to thank Arti Desai for all her kindness, support, and many adventures over the last few years.

I am grateful for the unending encouragement from those who have known me the longest. While a great many friends from back home have supported me for a great many years, I especially want to thank: Jaron Danioth, Roy Rowland, Ryan Geros, and Trina Paquette.

Last, but certainly not least, I wish to thank my entire family for all their love, support, and inspiration. In particular, I want to thank my sister Hillary Dorksen, who I know is

secretly proud of me but endeavors to make sure I never let it go to my head, and her husband Bryan Dorksen for all their support. And, while they're a bit young to understand the significance of my dissertation, I'm sure their children Mason and Caden would have encouraged me throughout this process. More than that, though, I'm sure my sister will let me hear it if I don't mention my nephews.

I was lucky enough to finally live in the same place as my cousin Heather Caughron, and I wish to thank her for all her kindness and support over the last few years.

I also want to thank Jamie Springberg. Aunt Jamie was always one of my biggest supporters, and her eternal love and support are immeasurable.

My grandfather George Martin has had a profound influence on my life. Starting at a very young age, we had numerous conversations about various engineering disciplines, many of which while out fishing. He passed his life-long interest in science and engineering on to me, and encouraged me every step of the way.

I especially want to thank my grandmother Bobbi Kirk. Through our many trips to various museums, she instilled in me an appreciation of art and culture at a very young age. She has been a major benefactor for the last few years, and her love and support have been critical in the completion of my degree.

Finally and most importantly, I want to thank my parents Christopher and Kathleen Kirk. They were always my best teachers, and continue to guide me through life. They instilled in me a love of art and science, and a desire to make the world a better place. They are the best parents anyone could hope for, and I can't thank them enough.

Chapter 1

Introduction

Reproducing the perception of low-light scenes presents challenges due to changes in how the human visual system responds at different light levels. In well-lit scenes, the eye behaves photopically with light perception mediated by the short, medium, and long cone cells. The three types of cone cells have distinct spectral response functions and they allow perception of a three-dimensional color space. In near-dark scenes, the eye functions scotopically, with only the rod cells active. The rod cells have a spectral response function that is distinct from the cones, and when only the rods are active color discrimination is dominated by a single perceptual axis, leading to primarily monochromatic vision. In between the photopic and scotopic regimens are low-light scenes, such as the one shown in Figures 1.1,1.2, and 1.3, where the eye functions mesopically. In mesopic vision all four types of receptors are active and contribute to color perception.

Each of these four types of photoreceptor has a spectral response that is linearly independent from the others. (See Figure 2.1.) This independence implies that the eye measures a four-dimensional colorspace. However, there are only three distinct types of signaling pathways to the brain from the retina. These pathways are dominated by the cones in photopic vision and by the rods in scotopic vision, but in mesopic vision the signals from the rods and cones are combined. Further, the relative weighting in this combination varies across the retina as local illumination varies within the mesopic regime. Therefore, although the eye senses in four-dimensions, communication of color to the brain still occurs within a three-dimensional colorspace. The compression from four to three dimensions is nonlinear and depends locally on the color and intensity of the scene. Perceptually, the change in the combination of signals manifests as a shift in hue that is commonly known as the Purkinje effect.

Unfortunately, the Purkinje effect is not replicated by current cameras or displays. Because humans photopically perceive a three-dimensional colorspace, commercially available camera and display technologies also operate in three dimensions, for example using three color sensors or three color phosphors. Color space standards, such as ISO RGB or CIE XYZ, explicitly assume that three dimensions are sufficient to model human color percep-

Figure 1.1: A high dynamic range (HDR) image of UC Berkeley's South Hall captured at night without perceptual tone mapping.

Figure 1.2: The result of our perceptual tone mapping algorithm for low-light conditions when applied to the image in Figure 1.1 using a simulated eight second exposure.

Figure 1.3: The result of our perceptual tone mapping algorithm for low-light conditions when applied to the image in Figure 1.1 using a simulated one second exposure.

tion. Monitors and printers that include more than three primaries typically do so to expand the gamut within this three-dimensional space, not to add extra dimensions of color. These systems assume photopic viewing and cameras/displays have no way to capture/stimulate rod response independent from cones. As a result, low-light images captured by standard cameras either look underexposed or they look very similar to well-lit images. Color variations that do occur between photographs of dimly and brightly lit scenes are due to different colored illuminants, and do not account for the perceptual color changes due to the Purkinje effect.

Color shifts associated with mesopic vision are an important lighting cue. Accordingly, photographers and cinematographers have developed ad hoc techniques to mimic color effects of vision under low-light conditions. Known as day-for-night techniques, they include adjusting the color balance, shooting with color filters, or using short exposures. Other approaches filter the color of environmental lighting [16]. While these techniques can be effective, they require a high degree of expertise to effectively simulate mesopic vision, and for optimal effect each scene must be treated independently. We propose a scene-independent method that is based on the biology of the eye. Our method does not require the user to modify the environment and is accesbile to novice users. In addition to artistic application, perceptually motivated models of scotopic and mesopic vision also play an important role in designing spaces and making lighting choices, for both aesthetic and safety reasons. One such example is highway design, where achromatic simulations of mesopic vision are used to measure the effects of lighting conditions on driving performance [2].

We present a computational model of the changes to perceived color that occur for low-light scenes, and a perceptually based tone mapping algorithm that generates standard RGB images that appear closer to what the standard viewer would experience in low-light conditions. This algorithm uses only a single point of control, uniform scaling of scene luminance, to determine a color change corresponding to the Purkinje shift. We use the term exposure to refer to this scaling as the apparent effect of uniform doubling/halving of scene luminance would correspond to doubling/halving the length of exposure when taking a photograph. Our algorithm captures low-light perceptual effects by approximating the psychophysical process through which color is perceived. It computes a standard observer's rod and cone receptor responses to a scene at a user-specified exposure, and determines how those responses would be combined to produce the output signals from the eye. Our method then computes the photopic RGB display image that most closely elicits the corresponding cone response for the standard observer viewing a given display.

In order to determine how the human visual system responds to a given scene, our tone mapping algorithm ideally works on spectral input images. While standard images approximate the spectral density function at a pixel using only three channels, a spectral image approximates the continuous distribution of energy at every pixel using a higher dimensional representation. This representation allows the cone and rod responses to be computed directly by integrating their response functions against the pixel's spectral density function. While our method works on spectral data from any source, we present a method

to determine a spectral image for a given static scene using multiple images under different color filters. This method assumes a camera with known spectral sensitivity. If the spectral sensitivity isn't provided by the camera manufacturer, we describe a one-time calibration procedure that estimates the sensitivity.

While spectral images are useful, they are also difficult to acquire. In cases when only a standard RGB image is available, we describe a learning process that enables estimation of a plausible rod and cone response to the unknown spectral distribution. Although this preprocessing allows us to apply our algorithm to RGB input images and produces reasonable results, working with true spectral data is preferred as it captures effects that are lost by premature transition into a three-dimensional color space. In particular, photopic metamers may appear distinct when viewed mesopically or scotopically due to independent rod contribution. Premature projection into a three-dimensional color space precludes estimating a rod response independent from cone responses.

We include results for several images to demonstrate the effectiveness of our algorithms. We also provide supplemental data including measured filter transmission curves and camera sensor responses that may be of use to others working on similar problems.

Chapter 2

Background

2.1 Vision Science

The human eye perceives light using four types of photoreceptor: long, medium, and short cone cells, and rod cells. These types of receptor each detect a specific distribution of wavelengths, and their sensitivity functions are linearly independent. The spectral sensitivities are shown in Figure 2.1. In addition to spectral sensitivity, the receptors are also characterized by a range of intensities over which they are active. Cones operate in photopic, or well-lit, conditions and rods operate under dark, or scotopic, conditions. Photopic and scotopic illumination levels overlap in a mesopic range, where both rods and cones contribute to color perception.

While we have four types of photoreceptor, we only perceive a three-dimensional colorspace. This limitation exists because the rod cells use the same pathways as the cones for conveying information to the brain. This communication of color to the brain is believed to be characterized by the three-dimensional opponent color model. This model states that there are three perception channels that each operate in an antagonistic space. The first is a combination of the long and medium cone cells, which corresponds approximately to the red versus green quality of color. The second is a combination of all three cone cells, which generally describes the blue versus yellow qualities. The final channel is another combination of the long and medium cone cells, and gives an approximation to overall intensity. Experiments show the rod cells also convey information to the brain using all three of these channels [1].

The comparative contribution of rod and cone receptors in the opponent color model determines how subjects perceive color. In photopic vision rod contribution is dominated by the cone responses, so much so that vision is accurately modeled using only these three receptors. In this case, there is a linear mapping from receptor signals to opponent color model. The large amount of cone contribution in photopic vision leads to the perception of many different color tones. In scotopic vision cones are inactive and intensity is perceived

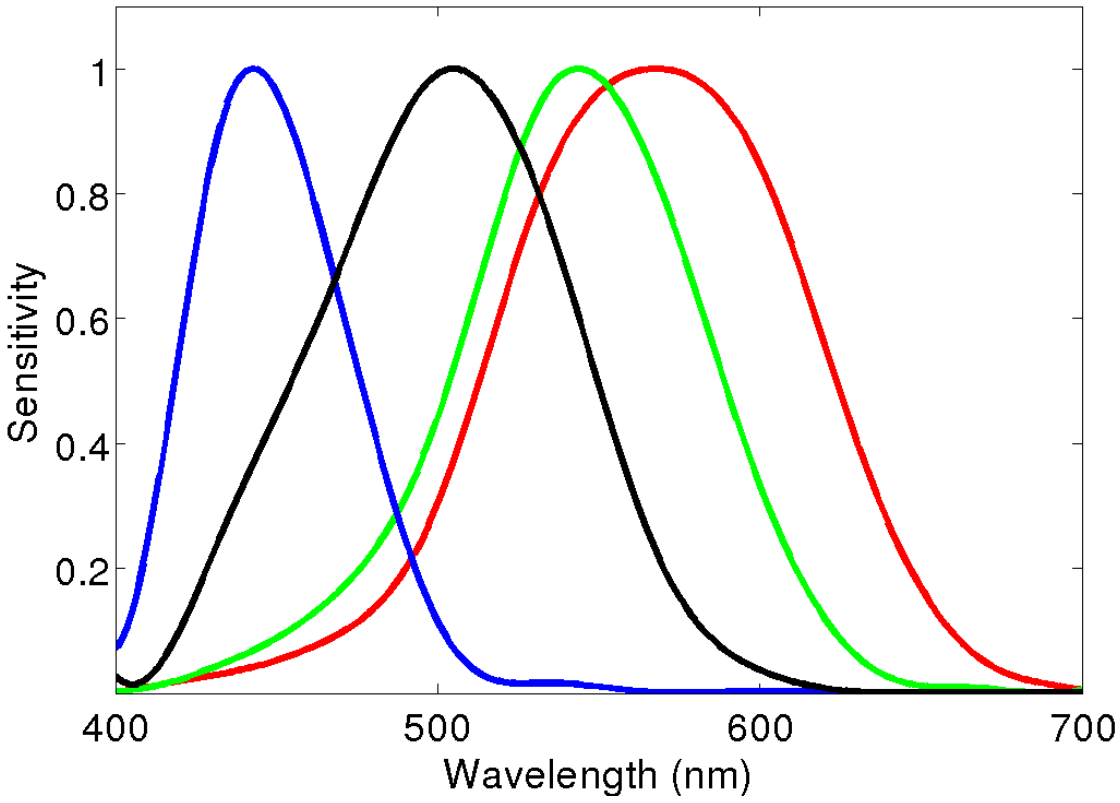

Figure 2.1: Plot of the normalized cone and rod spectral sensitivities. Long cone sensitivities are shown in red, medium cone sensitivites in green, short cone sensitivites in blue, and rod sensitivites in black [3, 27]. Note that these curves are independent, and no combination of the cone responses is equal to the rod sensitivity.

using primarily rod receptors. Here again the mapping from receptor signals to opponent color model is straight-forward, because cone contribution is negligible. Note that while scotopic vision is one-dimensional, subjects do not perceive gray scale. Instead, subjects report the perception of monochromatic scenes. This is due to the proportion in which the rods contribute to the different channels in the opponent color model. In mesopic vision both rods and cones are active and contribute to the three opponent color pathways. This mapping from four to three dimensions is nonlinear and varies with intensity. As the intensity changes, the opponent color model transitions non-linearly from a full three-dimensional colorspace down to a monochromic space. During this transition, subjects experience three-dimensional color perception, albeit with a reduced gamut and color shift [26]. This particular biological configuration explains how the rod cells can create the sensation of color, and it also gives a sense of why the Purkinje phenomenon is difficult to capture with a simple model [5].

In the vision science literature, researchers have done experiments to characterize the Purkinje effect and proposed theories to explain the underlying mechanism. Shin and colleagues [25] have proposed a partial model based on data from user studies. Their data-driven model is based in LMS space (LMS space refers to the three axes of the space: long, medium, and short cone response). They collected data from several subjects on how color perception shifts at a set of illumination levels that are evenly separated perceptually. They then built a multilinear model based on LMS values as well as illumination level. Their model fits their data well, however it is difficult to extrapolate their results to lighting conditions different from those in their experimental setup. Cao and colleagues [5] also fit a model based on user studies, however their model is based more heavily on human biology. As a result, their model is more effective at a wide rof illumination. Our tone mapping algorithm is based on the Cao model.

2.2 Spectral Images

Our method models the responses of the cone and rod receptors to light in a scene and then uses them to compute perceived color. These responses can be computed directly from a spectral image, which gives an approximate distribution of energy in the visible wavelengths for each pixel. Integrating the per pixel spectral distribution against the rod and cone spectral sensitivities gives the eye's response for each of the four receptors to every pixel in the image. This information lets us leverage current results from vision science and human perception discussed above to estimate the perceptual effect of the scene.

There are several existing methods for acquiring spectral images. Burns and Berns [4] use a set of band pass filters. They sampled the scene under these filters and solved for the image spectra. The benefit of this approach is that accuracy is easy to analyze and increases with the number and diversity of filters, however in the limit it requires many filters. Furthermore, bandpass filters are most accurate with a small angle of incidence which may cause errors towards the edge of the image. Mohan and colleagues [20] built a system to dynamically select the spectra best associated with their device's input channels. While they get limited data in a single photograph, such a system could be used to construct a spectral image with multiple images. The benefit of their system is that the spectral bands used in reconstruction can be selected dynamically. Schechner and Nayar [24] show that it is possible to spectrally image a scene using video taken through a spatially varying filter. They attach a horizontally varying bandpass filter to the front of a video camera and use it to build a spectral image. By panning across the scene, they capture each scene point under the entire range of their filter. They then construct a spectral image using image correspondence.

Our spectral image acquisition process, discussed in Chapter 4, is mathematically similar to that of Park and colleagues [22]. They showed that one can reconstruct the reflectance spectra of objects in a scene by imaging the scene under known multiplexed lighting conditions. This process requires a calibrated light source with a controller to gather the data and

a camera with known spectral sensitivity. Their model of pixel generation is a linear function of incident lighting, material reflectance, and camera sensitivity. Based on this model and the known camera sensitivity, they show that the material reflectance can be solved for by capturing multiple images each under a unique known set of lights. With known reflectances at each pixel and known lighting conditions they can determine the spectral image of a scene.

The method of Park and colleagues [22] discussed above relies on a known camera spectral sensitivity. They also show these sensitivity curves can be determined experimentally if the information is not published by the camera manufacturer. To do this, they simply image a scene with known material reflectances. Because they know the incident lighting and reflectances, they can determine the spectral sensitivity of the camera. They show the mathematics are similar to scene acquisition. We adopt the same approach and adapt our scene aquisition process to calibrate the camera.

Of particular note, recent work shows minor modifications to existing camera sensors can enable the capture of more spectral information in a single picture. Yasuma and colleagues [29] simulated a camera designed to capture more than three unique frequency bands. They proposed a modification to the sensor's Bayer pattern to introduce additional filters at a lower resolution. With such a system, a traditional RGB image can still be reconstructed, and at the same time one can obtain lower-resolution spectral information in a single picture. Similarly, Kumar and colleagues [15] proposed a panchromatic sensor, which adds a clear or white filter to the Bayer pattern. They motivate their work with reduced noise in low-light images. However, they don't address the perceptual qualities of such images, and the resulting images appear as if they were taken under higher intensity lighting. Having four independent spectral channels in a single image makes it possible to map directly to the four independent photoreceptor responses. These approaches, which both modify the physical sensing hardware, have the strong advantage of being applicable to scenes that contain moving objects and where the lighting cannot be controlled. The availability of commercial imaging devices that could easily acquire spectral images could potentially have many applications including our low-light tone mapping method.

2.3 RGB Approximation

Our work uses spectral images to perform perceptually based tone mapping. However, others have addressed the issue using regular RGB image data. Tumblin and Rushmeier originally introduced the problem [28]. Ferwerda and colleagues [9] model several characteristics of the visual system, including: threshold visibility, color appearance, acuity, and time sensitivity. While this work addresses color change due to decreases in light level, at scotopic levels their model produces a greyscale image, which is not what viewers report. Durand and Dorsey [7] address this problem by blending a particular color of blue into scenes as light is decreased. Khan and Pattanaik [12] use a similar approach to Durand and Dorsey. In their work, they hypothesize that the blue shift in moonlit scenes is caused by rod receptors contributing color

as if they were short cones. They estimate a cone response from the initial scene and, as the light is reduced, blend an attenuated rod response into the short cone response. Finally, they convert to RGB space. While both Khan and Pattanaik and Durand and Dorsey introduce a color shift, blending a linear amount of a single blue color is inconsistent with psychophysical data.

Pattanaik and colleagues [23] examined a variety of important perceptual effects related to high-dynamic-range scenes, including effects that occur at low-light levels. They also explicitly consider rod contribution to perception. Our approach to modeling the Purkinje effect goes beyond what they present with a more comprehensive model of the color effects. Kuang and colleagues [14] presented the iCAM06 model for rendering HDR images. Their framework accounts for rod contribution in low-light conditions, but with a less detailed model than ours that operates directly in RGB space. Both of these methods examined a host of related issues, such as increased noise and loss of spatial acuity at low light levels, that would complement our work, which focuses only on color effects.

In addition to these automated methods, it is possible for an artist to tone map an image for low-light conditions using standard image editing tools. Photo editing tools in the hands of a skilled artist can create stylized effects that can be quite attractive and convey the general impression of low light. However, manual tone mapping requires the artist to make arbitrary decisions on how the colors should be modified and is generally not perceptually correct.

Chapter 3

Spectral Calibration

Our tone mapping method ideally operates on spectral images, but obtaining these images is difficult. In Chapter 4 we show how to capture a spectral image using multiple images each under a different color filter, but that technique requires a camera with known spectral sensitivity. We show how to calibrate a camera's spectral sensitivity using a known source, colored filters, and multiple exposures. This calibration need only be performed once for a camera sensor. The mathmatics of our camera calibration are similar to spectral image acquisition and many of the ideas transfer to the following chapter.

Digital cameras capture images by integrating the light's spectral distribution against the camera sensor's spectral sensitivity. To calibrate the spectral sensitivity of the camera, we image a known source through a series of known transmissive filters. Each image provides a sample that maps from a specific spectral distribution of light to an RGB color. Given enough samples we show how to solve for the camera's spectral response curves. A brute force approach would be to sample a known source with a series of bandpass filters, for example at each nanometer, and build a lookup table that gives the sensitivity across the spectrum. Instead, we use a set of linearly independent color filters and solve for a smoothed set of sensitivity curves.

We calibrate a camera's spectral sensitivity by sampling the mapping from a light's spectral distribution to an RGB color. The naive experimental setup would be to image a known source modulated by a color filter directly onto the camera sensor. There are two ways to build the naive experiment. First, the object plane could be co-located with the image plane at a reasonable distance from the source. The further from the source, however, the lower the intensity and smaller the size of the illumination field. The alternative naive experiment is to place the source directly at the object plane. However, this setup will generate an undesirable image of the filament in the image. This is similar to a classic problem in microscopy, where the image of the filament would be visible on the sample. Placing the object plane Koehler illumination has become the standard technique to provide an evenly illuminated field at the object plane [10].

Our experimental setup is shown in Figure 3.1. In this setup the spectral distribution

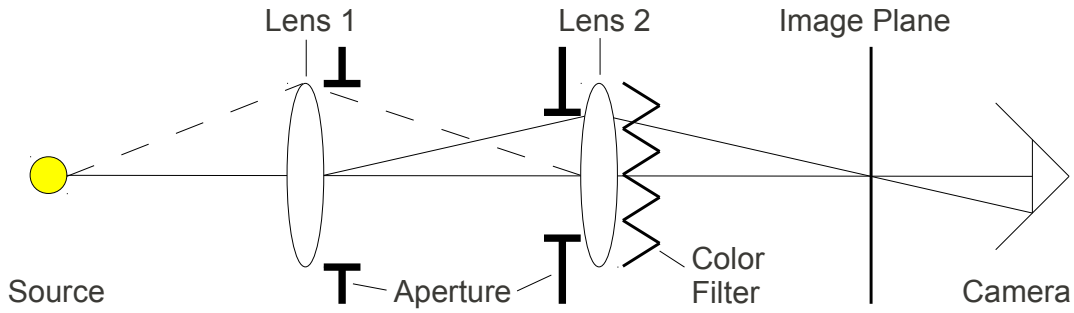

Figure 3.1: Diagram of a Koehler Illumination system used to calibrate the spectral sensitivity of a camera. This diagram is not to scale.

of the source is known, as is the spectral transmission of the color filters. The object plane is located at lens 1, where the source is reasonably decorrelated. The location of the image plane is determined by lens 2, such that the decorrelated object plane is focused onto the image plane. The camera is placed so the camera sensor plane is conjugate to the image plane. In practice, precise location of the sensor plane with respect to the image plane is obtained using the focus of the camera lens system. Colored filters are inserted before the image plane. The source is highly decorrelated here, which limits the effects of any dust or irregularites in the filter. Note that lens 2 is not needed if the source can be imaged directly onto the entrance pupil of the camera lens. However, the entrance pupil is inside the camera lens enclosure, which makes it difficult to locate. Adding lens 2 avoids this problem and improves the quality of the data.

Our calibration setup uses lenses with effective focal lengths of 10 and 20 cm for lens 1 and lens 2 respectively. Both lenses are achromatic doublets so that chromatic aberration is minimized [11]. We also selected relatively large lenses (2 inch diameter) so more light is coupled into the illumination system. Both apertures in the setup are adjustable irises. Aperture 1 limits the field and aperture 2 adjusts illumination coherence. Koehler illumination is achieved by placing the lenses so that the source is imaged onto the entrance pupil of the imaging system (front surface of lens 2). The object plane is chosen to be at the position of aperture 1, so the image plane is easily located by observing where the image of aperture 1 is in focus. Our source is a SoLux Q50MR16/CG/47/36 4700 Kelvin Halogen Bulb driven by a regulated DC power source, chosen for its broad-spectrum properties. We use a set of Roscolux polycarbonate filters whose transmission properties can be seen in Figure 3.2. All images were captured using a Canon EOS 300D in RAW format and without gamma correction so that the data was in a linear energy space.

Because we know the spectral properties of the source light as well as the transmission of

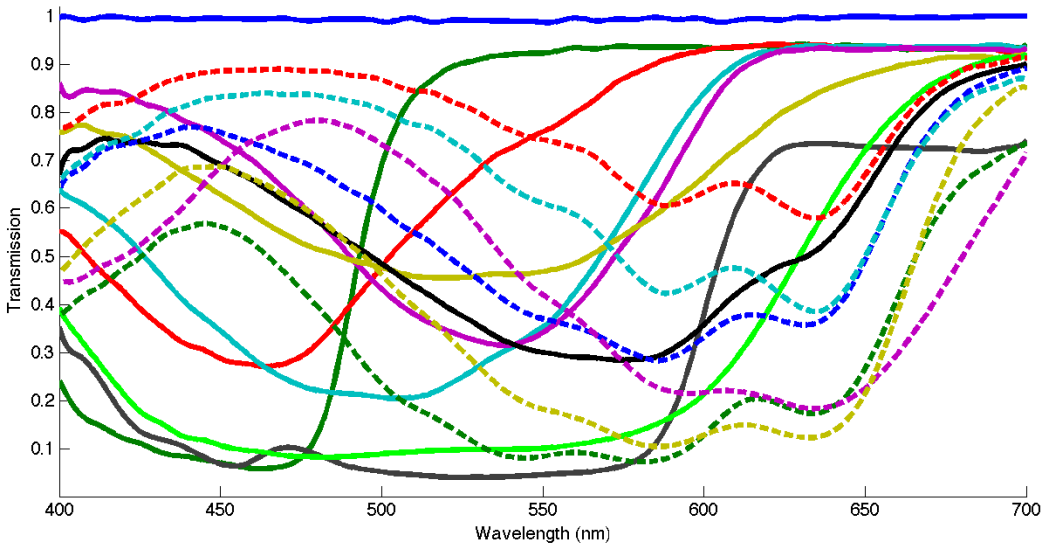

Figure 3.2: Transmission characteristics of various Roscolux filters.

the colored filter, we can determine the spectra of the incident light. Taking a picture with that incident light will give us a corresponding spectra and RGB pair, where the RGB value is determined by the camera's spectral sensitivity. This imaging setup creates images such as those seen in Figure 3.3. While the image of the filtered source illuminant is not entirely uniform, it is smoothly varying and relatively stable at the center of the projection area. In Figure 3.4 we plot a subset of the row in the image that passes through the center of the bright spot. Note that there is a significant window of pixels in which the illumination is noisy but reasonably flat. To determine the value of a pixel under our known illumination, for each color channel we average a small pixel neighborhood from the flat intensity profile that describes the brightest area in the calibration image. We assume that each intensity value (red, green, and blue) is representative of what any element of the camera sensor would record under the same illumination.

Each image with a different color filter gives us a unique spectral distribution of light and its corresponding RGB color, for our camera. We use this data to solve for the camera spectral sensitivity. Our model for pixel formation follows. For a given pixel p, let p_c be the value recorded by a camera in color channel $c \in \{\text{red}, \text{green}, \text{blue}\}$. If the camera sensor's sensitivity function is $\mathcal{S}_c(\lambda)$, then we can write

$$p_c = \int_{\Omega} \mathcal{S}_c(\lambda)\, \mathcal{R}(\lambda)\, \mathrm{d}\lambda \tag{3.1}$$

where $\mathcal{R}(\lambda)$ is the spectral distribution of light incident at pixel p, and $\Omega = [400\text{nm}\ 700\text{nm}]$.

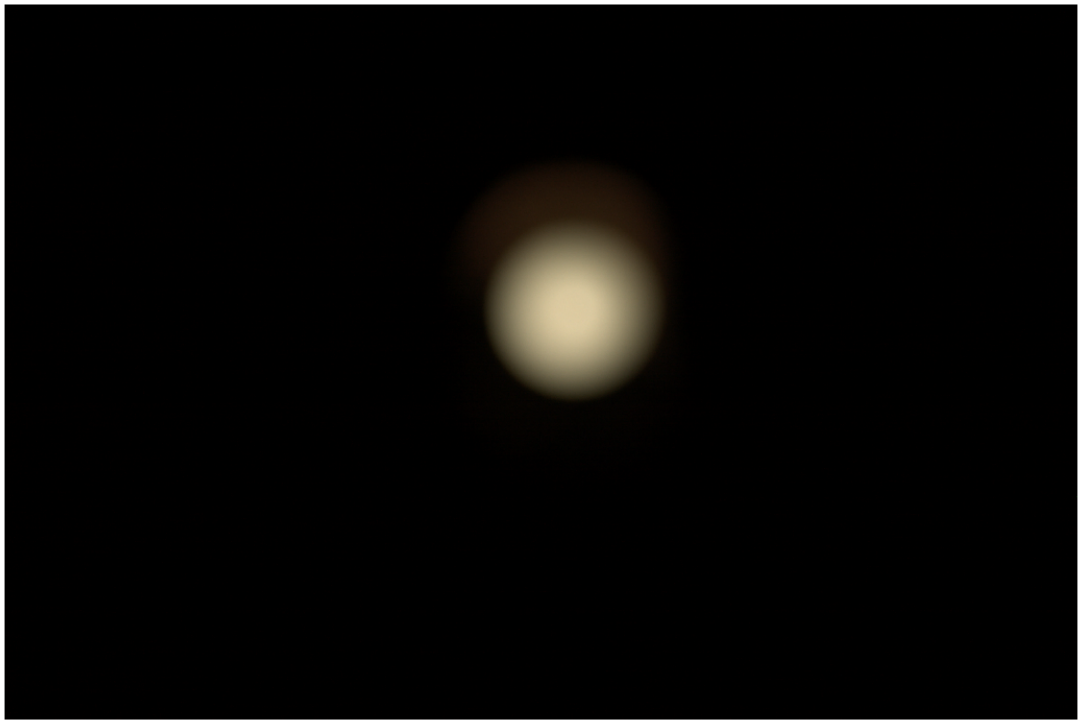

Figure 3.3: Example image from known source used in spectral sensitivity calibration.

To calibrate the camera for spectral photography, it is necessary to know $\mathcal{S}_c(\lambda)$ for all c. By using a set of known illuminations (which gives multiple equations each with a different $\mathcal{R}(\lambda)$), we can determine spectral sensitivity experimentally. First, we parameterize the function. We represent the spectral functions using B-spline basis functions so that, for example, we have

$$\mathcal{S}_c(\lambda) = \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{s}_c \tag{3.2}$$

where $\boldsymbol{\mathcal{B}}(\lambda)$ is a vector of uniform B-spline basis functions over the interval Ω, $\boldsymbol{s}_c$ is the vector of control points for $\mathcal{S}_c(\lambda)$, and $\cdot^{\mathsf{T}}$ denotes transpose. We use nine knot points uniformly distributed over Ω. This number was experimentally determined to be sufficient.

With this discretization we can rewrite (3.1) as

$$p_c = \int_{\Omega} \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{s}_c \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{r} \, \mathrm{d}\lambda \ . \tag{3.3}$$

Given multiple samples of known illuminations $\mathcal{R}(\lambda)_i$ like those shown in Figure 3.3, we can solve Equation 3.3 for the camera sensitivity. In this case, all quantities except $\boldsymbol{s}_c$ are

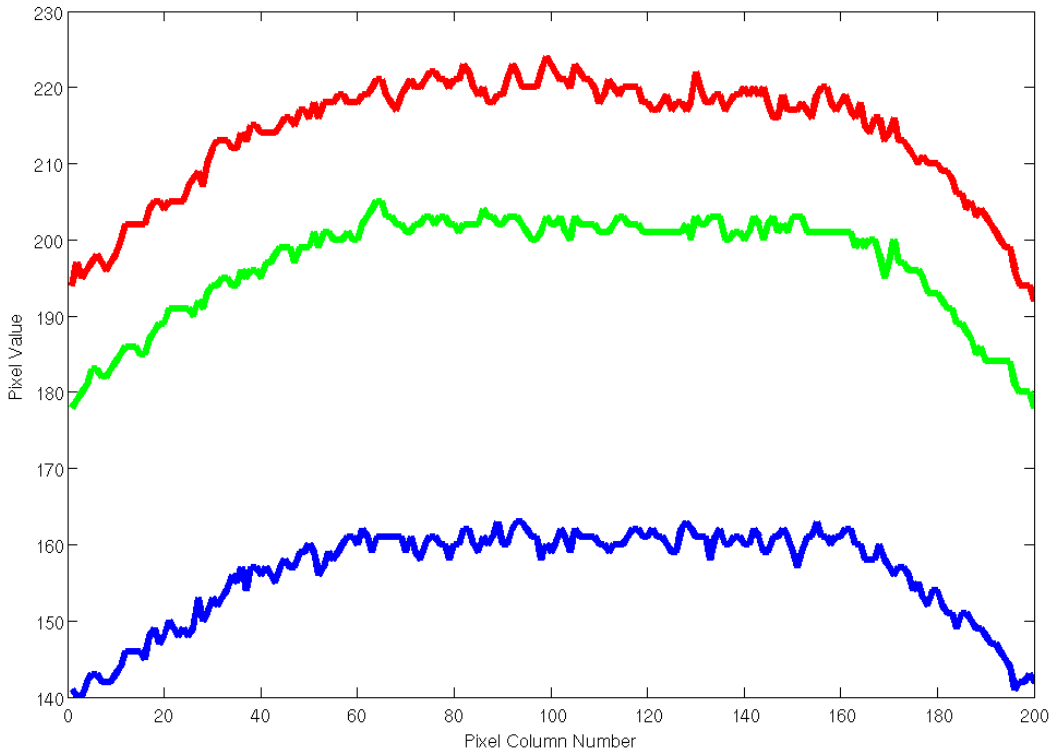

Figure 3.4: Plot of pixel data from a single row taken from the center of the Koehler illumination field.

known, so we move $\boldsymbol{s}_c$ out of the integral and rewrite as

$$\begin{aligned} p_{c,i} &= \left(\int_{\Omega} \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{r}_i \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \, \mathrm{d}\lambda \right) \boldsymbol{r} \\ p_{c,i} &= \boldsymbol{u}_{c,i}^{\mathsf{T}} \boldsymbol{s}_c \end{aligned} \tag{3.4}$$

where $\boldsymbol{u}_{c,i}^{\mathsf{T}}$ is equal to the expression in parentheses.

We collect all of these equations for each channel and all n known illuminations together as

$$\boldsymbol{U}_c = \begin{bmatrix} \boldsymbol{u}_{c,1}^{\mathsf{T}} \\ \vdots \\ \boldsymbol{u}_{c,n}^{\mathsf{T}} \end{bmatrix} \quad \boldsymbol{p}_c = \begin{bmatrix} p_{c,1} \\ \vdots \\ p_{c,n} \end{bmatrix} \quad . \tag{3.5}$$

We then solve for $\boldsymbol{s}_c$ in a least-squares sense. However, to avoid physically impossible solutions we constrain the solution to the positive domain, and set the value and derivative

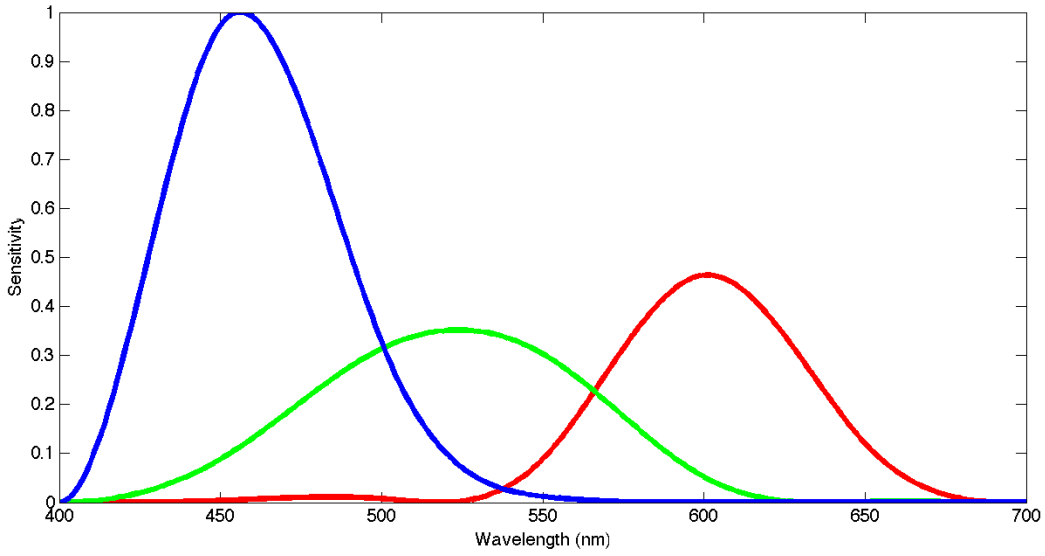

Figure 3.5: Plot of camera sensor sensitivity for a Canon 300D.

of the endpoints to be zero, as such:

$$\begin{aligned} \min_{\boldsymbol{s}_c} \quad & \|\boldsymbol{U}_c\, \boldsymbol{s}_c - \boldsymbol{p}_c\|_2^2 \\ s.t. \quad & \boldsymbol{\mathcal{B}}(\lambda)^\mathsf{T} \boldsymbol{s}_c \geq 0 \\ & \boldsymbol{\mathcal{B}}(400)^\mathsf{T} \boldsymbol{s}_c = 0 \\ & \boldsymbol{\mathcal{B}}(700)^\mathsf{T} \boldsymbol{s}_c = 0 \\ & \frac{d}{d\lambda}\boldsymbol{\mathcal{B}}(400)^\mathsf{T} \boldsymbol{s}_c = 0 \\ & \frac{d}{d\lambda}\boldsymbol{\mathcal{B}}(700)^\mathsf{T} \boldsymbol{s}_c = 0 \quad . \end{aligned} \tag{3.6}$$

We constrain the solution at the endpoints because the signal to noise ratio at those frequencies is poor. Most camera sensors attenuate frequencies towards both ends of the visible light range. The lack of data at these points can lead to overfitting if not constrained. The resulting normalized sensitivity curves are shown in Figure 3.5.

Chapter 4

Acquiring Spectral Images

The first step in our tone mapping procedure is to acquire a spectral image of the scene. We build spectral images using multiple RGB images, each taken through a different color filter. Because we know the sensitivity functions of the camera sensors and transmission functions of the filters, we can solve for the unknown spectral distribution of the light entering the camera at each pixel. This method is analogous to the method of Park and colleagues [22]. In their work, they image a scene under a series of differently colored lights and solve for surface reflectance. Mathematically the methods are nearly identical, but our method solves for the light entering the camera rather than surface properties and does not require controlled illumination.

We use the same model of pixel formation described in Equation 3.1, but in this case wish to solve for $\mathcal{R}(\lambda)$. The three color channel limit of commercial cameras is insufficient for uniquely resolving the function [13]. Ideally we could add additional channels to the camera to further sample the spectra, but because doing so is infeasible we instead take multiple images with each image taken through a different transmissive filter. If $\mathcal{F}_i(\lambda)$ is the transmission function for one of these filters then the pixel value in an image taken through that filter is given by

$$p_{c,i} = \int_{\Omega} \mathcal{S}_c(\lambda)\, \mathcal{F}_i(\lambda)\, \mathcal{R}(\lambda)\, \mathrm{d}\lambda \quad . \tag{4.1}$$

To determine the spectral image we need to compute the spectra $\mathcal{R}(\lambda)$ that best predicts all the observed values $p_{c,i}$.

Similar to calculating the camera's spectral sensitivity, we represent the incoming spectra using B-spline basis functions. In this case

$$\mathcal{R}(\lambda) = \quad \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{r} \tag{4.2}$$

where $\boldsymbol{r}$ is the vector of control points for $\mathcal{R}(\lambda)$. We use 10 knot points uniformly distributed over Ω. This number was selected as sufficient based on spectrometer measurements from a collection of Munsell color squares. Modern illuminants, such as fluorescent or LED lights,

tend to have sharply peaked spectral distributions that would not be well resolved using a smooth B-spline basis. However, we will integrate the distributions against the smooth rod and cone sensitivity functions, so a smoothed representation of sharp peaks will not adversely impact our results.

After discretization we can rewrite (4.1) as

$$p_{c,i} = \int_{\Omega} \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{s}_c \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{f}_i \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{r} \, \mathrm{d}\lambda \ . \tag{4.3}$$

All quantities except $\boldsymbol{r}$ are known, so we move $\boldsymbol{r}$ out of the integral and rewrite as

$$\begin{aligned} p_{c,i} &= \left(\int_{\Omega} \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{s}_c \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{f}_i \, \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \, \mathrm{d}\lambda \right) \boldsymbol{r} \\ p_{c,i} &= \boldsymbol{v}_{c,i}^{\mathsf{T}} \boldsymbol{r} \end{aligned} \tag{4.4}$$

where $\boldsymbol{v}_{c,i}^{\mathsf{T}}$ is equal to the expression in parentheses.

We collect all of these equations for each channel and all n filters together as

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_{red,1}^{\mathsf{T}} \\ \boldsymbol{v}_{green,1}^{\mathsf{T}} \\ \boldsymbol{v}_{blue,1}^{\mathsf{T}} \\ \boldsymbol{v}_{red,2}^{\mathsf{T}} \\ \vdots \\ \boldsymbol{v}_{blue,n}^{\mathsf{T}} \end{bmatrix} \quad \boldsymbol{p} = \begin{bmatrix} p_{red,1} \\ p_{green,1} \\ p_{blue,1} \\ p_{red,2} \\ \vdots \\ p_{blue,n} \end{bmatrix} \ . \tag{4.5}$$

We then solve for $\boldsymbol{r}$ in a least-squares sense while constraining the solution to be positive using

$$\begin{aligned} \min_{\boldsymbol{r}} \quad & \|\boldsymbol{V}\boldsymbol{r} - \boldsymbol{p}\|_2^2 \\ s.t. \quad & \boldsymbol{\mathcal{B}}(\lambda)^{\mathsf{T}} \boldsymbol{r} \geq 0 \ . \end{aligned} \tag{4.6}$$

If a sensor becomes saturated or is too dark it becomes impossible to determine exactly how much energy was incident at that pixel. To capture the spectra in a scene, we use high dynamic range (HDR) input images [6]. In our examples, we captured a static scene with multiple exposure levels under each transmissive filter. We then registered the images from each filter and exposure level to each other and normalized them all to a one second exposure. Finally, for each set of images under a particular color filter we select the best exposure level for each pixel using standard HDR image construction techniques, which gives us an HDR image for each filter.

The spectral image data were captured for static scenes using Cokin P-series filters, specifically filters P001, P002, P003, P004, P005, P006, P020, P036, P047, P050, and P231, shown in Figure 4.1. The goal in selecting these filters was to find a linearly independent set that sufficiently spanned the visible wavelengths. We determined the spectral transmission of

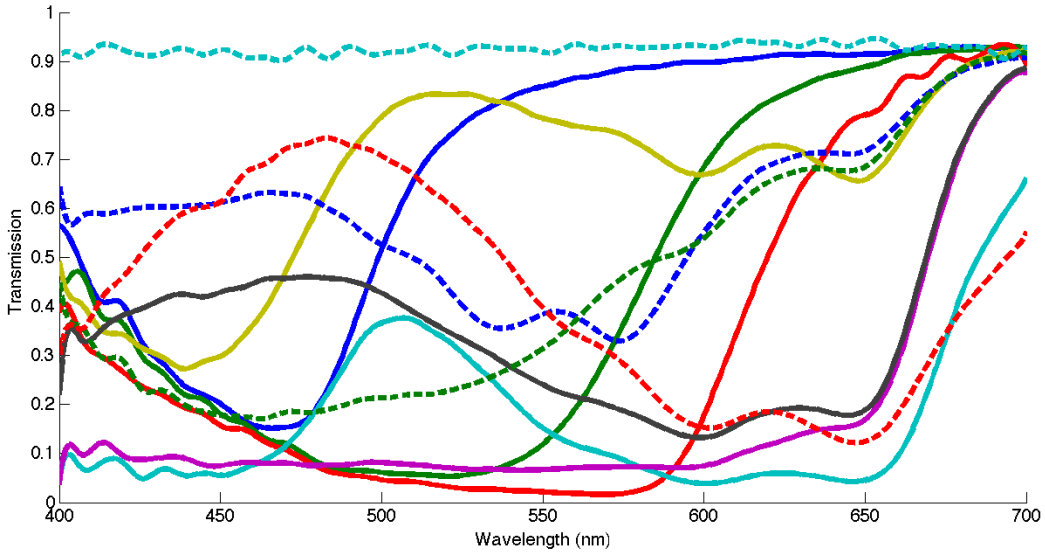

Figure 4.1: Transmission characteristics of various Cokin P-series filters.

each filter as a percentage of incident light using an Edmund Optics Visible Wavelength USB CCD Spectrometer, model BRC111A-USB-VIS, which has a wavelength resolution of 1 nm. A sample input series with an inset plot of percentage of light transmitted per wavelength by the filter can be seen in the left column of Figure 6.1.

With a three channel camera sensor and eleven filters we have 33 independent measurements of the $\mathcal{R}(\lambda)$ which is somewhat excessive for estimating the four values of the rod and cone responses at each pixel. However, these spectral images have uses in other applications and we designed our algorithm to work with this more general data. The camera and filter calibration data and the raw image data is included in the supplemental data.

Chapter 5

Tone Mapping for Low-Light Conditions

5.1 Tone Mapping Spectral Images

Once we have acquired spectral image data, the first step of our tone mapping procedure is to determine the cone and rod response to the acquired data. To convert the spectral image to LMSR (long, medium, and short cones, and rods) space, we integrate each pixel's spectral distribution against the standard rod and cone response curves [3, 27]. If a pixel has spectrum $\mathcal{R}(\lambda)$ and the normalized response curve of the eye for a given receptor is $\mathcal{E}_j(\lambda)$, where $j \in \{\text{Long}, \text{Medium}, \text{Short}, \text{Rod}\}$, then the normalized four-dimensional LMSR response $\boldsymbol{q}$ can be calculated as

$$q_n = \int_\Omega \mathcal{E}_n(\lambda)\, \mathcal{R}(\lambda)\, \mathrm{d}\lambda \tag{5.1}$$

We compute this integral using the same B-spline representation described in the previous section.

At this stage, the LMSR image represents the rod and cone responses that would result from the scene being viewed by a standard observer. We use this information to determine the Purkinje Shift in the opponent color model, which we model as an offset to the long, medium, and short channels such that the result is the perceptually closest photopic signal to the mesopic signal [5]. More specifically, let $\hat{\boldsymbol{q}}$ be a three-dimensional long, medium, and short cone response. We show how to compute $\Delta\boldsymbol{q}$ such that

$$\hat{\boldsymbol{q}} = [q_{\text{Long}}\, q_{\text{Medium}}\, q_{\text{Short}}]^{\mathsf{T}} + \Delta\boldsymbol{q} \tag{5.2}$$

where $\hat{\boldsymbol{q}}$ will be the three-dimensional cone response most perceptually similar to the four-dimensional cone and rod response $\boldsymbol{q}$.

Figure 5.1: HDR still life scene with no tone mapping.

As discussed previously, rods and cones share neural pathways. The signals passing through those pathways are a combination of rod and cone receptors, subject to sensitivity regulation [19]. The regulated signals $\boldsymbol{g}$ are

$$\begin{aligned} g_{\text{Long}} &= 1/(1 + 0.33(\boldsymbol{q}_{\text{Long}} + \kappa_1 \boldsymbol{q}_{\text{Rod}}))^{0.5} \\ g_{\text{Medium}} &= 1/(1 + 0.33(\boldsymbol{q}_{\text{Medium}} + \kappa_1 \boldsymbol{q}_{\text{Rod}}))^{0.5} \\ g_{\text{Short}} &= 1/(1 + 0.33(\boldsymbol{q}_{\text{Short}} + \kappa_2 \boldsymbol{q}_{\text{Rod}}))^{0.5} \end{aligned} \tag{5.3}$$

where κ_1 and κ_2 determine the rod contribution to the neural signals [5]. These mixing terms depend on the mesopic viewing conditions. They are near zero during photopic vision, and reach an approximate value of $\kappa_1 = 0.25$ and $\kappa_2 = 0.4$ during fully adapted scotopic vision [5]. For our results we assume full adaptation and use the above values. We found that modulating κ_1 and κ_2 according to standard mesopic adaptation produced a system

Figure 5.2: The image in Figure 5.1 tone mapped for low-light conditions.

that was too sensitive to exposure and made it difficult to obtain a desired result. We could construct a system with a similar response but higher κ_1 and κ_2 fidelity by giving the user a non-linear exposure control.

As discussed earlier, our tone mapping procedure is based on rod contribution causing a shift in the opponent color model space. The opponent color model expresses color as three different combinations of the cone sensors. For a pixel in LMS space, $\hat{\boldsymbol{q}}$, the corresponding pixel in opponent color space $\boldsymbol{o}$ is computed as:

$$\begin{aligned} o_{\text{Red/Green}} &= \hat{\boldsymbol{q}}_{\text{Medium}} - \hat{\boldsymbol{q}}_{\text{Long}} \\ o_{\text{Blue/Yellow}} &= \hat{\boldsymbol{q}}_{\text{Short}} - (\hat{\boldsymbol{q}}_{\text{Long}} + \hat{\boldsymbol{q}}_{\text{Medium}}) \\ o_{\text{Luminance}} &= \hat{\boldsymbol{q}}_{\text{Long}} + \hat{\boldsymbol{q}}_{\text{Medium}} \end{aligned} \tag{5.4}$$

where the first component of $\boldsymbol{o}$ roughly measures depth of red versus depth of green, the

Figure 5.3: An HDR night scene featuring UC Berkeley's Sather Tower.

Figure 5.4: The image in Figure 5.3 tone mapped for low-light conditions. Artifacts on the clock face occur because the hand moved during image acquisition.

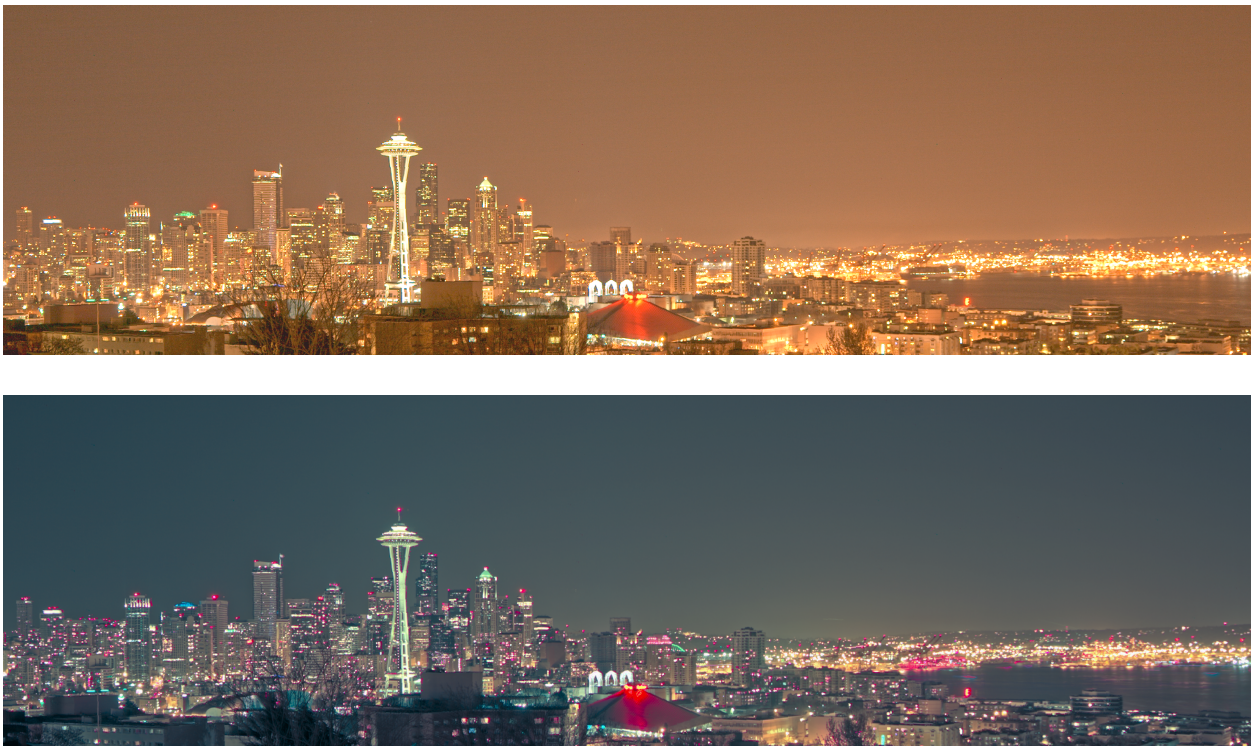

Figure 5.5: Comparison of a nighttime skyline scene. Top: An HDR image of the Seattle skyline as seen from Kerry Park at night with no tone mapping. Bottom: The image is tone mapped for low-light conditions. Some differences are due to lights being switched on and off in the scene during acquisition.

second channel measures depth of blue versus depth of yellowness, and the third channel measures luminance. Note that this transformation can be expressed as an invertible matrix such that $\boldsymbol{o} = \boldsymbol{A}\hat{\boldsymbol{q}}$.

To determine the amount of color shift, each term in the opponent color model has a weighted amount of the rod response added. The weight is a function of the sensitivity regulated receptor response $\boldsymbol{g}$. Specifically, the color shift in the opponent color model $\Delta\boldsymbol{o}$ is:

$$\begin{aligned}
\Delta o_{\text{Red/Green}} &= x\,\kappa_1 \left(\rho_1 \frac{g_{\text{Medium}}}{m_{\max}} - \rho_2 \frac{g_{\text{Long}}}{l_{\max}}\right) q_{\text{Rod}} \\
\Delta o_{\text{Blue/Yellow}} &= y \left(\rho_3 \frac{g_{\text{Short}}}{s_{\max}} - \rho_4 \left(\alpha \frac{g_{\text{Long}}}{l_{\max}} + (1-\alpha)\frac{g_{\text{Medium}}}{m_{\max}}\right)\right) q_{\text{Rod}} \\
\Delta o_{\text{Luminance}} &= z \left(\alpha \frac{g_{\text{Long}}}{l_{\max}} + (1-\alpha)\frac{g_{\text{Medium}}}{m_{\max}}\right) q_{\text{Rod}}
\end{aligned} \tag{5.5}$$

Figure 5.6: An HDR image of the Fremont Troll with no tone mapping.

Where $l_{\max} = 0.637$, $m_{\max} = .392$, and $s_{\max} = 1.606$ are the maximum values of the cone fundamentals [27], and psychophysical data [5] is used to fit the ρ and α as: $\rho_1 = 1.111$, $\rho_2 = 0.939$, $\rho_3 = .4$, $\rho_4 = 0.15$, and $\alpha = .619$. The above leaves free parameters x, y, and z. These parameters control the scale of the shift for each of the channels in the opponent color model. Psychophysical data suggests that x and y are the same, but z is independent [5]. For all the examples in this paper, we used $x = y = 15$ and $z = 5$. Returning to Equation 5.2, we have

$$\hat{\boldsymbol{q}} = [q_{\text{Long}}\, q_{\text{Medium}}\, q_{\text{Short}}]^{\mathsf{T}} + A^{-1}\Delta\boldsymbol{o} \tag{5.6}$$

giving the photopic response $\hat{\boldsymbol{q}}$ most perceptually similar to the original mesopic response.

The final step is to convert from LMS space into RGB space for display. At this stage, we know the desired excitation of cones in the retina and we seek to find the display settings that produce that excitation. This process is essentially a color matching problem where we need to know the mapping from the display's primaries, $\mathcal{M}_{\text{Red,Green,Blue}}(\lambda)$, to viewer stimulus.

Figure 5.7: The image in Figure 5.6 has been tone mapped for low-light conditions.

These spectra are device dependent and one can map values for one display onto another using standard color matching functions. Further, cone responses depend on adaptation to the ambient viewing conditions [18]. Our results were computed using the measured spectral output of an Apple Cinema HD Display with the default display profile, and we assume viewing in a dark room. Common experience shows that viewing RGB images calibrated for one display on different displays typically produces reasonable results, but applications demanding precision should use color matching functions appropriate to the actually display device, display driver color profile, and viewing conditions.

Let $\boldsymbol{M}$ be the matrix that maps the display's spectral emission onto the cone fundamentals

$$\boldsymbol{M} = \begin{bmatrix} m_{\text{Long,Red}} & m_{\text{Long,Green}} & m_{\text{Long,Blue}} \\ m_{\text{Medium,Red}} & m_{\text{Medium,Green}} & m_{\text{Medium,Blue}} \\ m_{\text{Short,Red}} & m_{\text{Short,Green}} & m_{\text{Short,Blue}} \end{bmatrix}, \tag{5.7}$$

Figure 5.8: The result of a standard grayscale image conversion applied to the image in Figure 5.3, scaled to fit the range [0 255] to highlight differences with Figure 5.9.

Figure 5.9: Image showing our measure of the mesopic value of a pixel in Figure 5.4, where lighter indicates a more photopic pixel and darker indicates a more scotopic pixel. This image has been scaled to fit the range [0 255] to highlight differences with Figure 5.8.

with the components of $\boldsymbol{M}$ given by $m_{x,y} = \int_\Omega \mathcal{E}_x(\lambda)\,\mathcal{M}_y(\lambda)\,\mathrm{d}\lambda$. We solve for the final RGB settings $\boldsymbol{p}^*$ that maps to our desired photopic response $\hat{\boldsymbol{q}}$ with

$$\begin{aligned} \min_{\boldsymbol{p}^*} \quad & \|\boldsymbol{M}\boldsymbol{p}^* - \hat{\boldsymbol{q}}\|_2^2 \\ s.t. \quad & \boldsymbol{p}^* \geq 0 \end{aligned} \tag{5.8}$$

When solving for a given pixel, we first try the unconstrained solution, $\boldsymbol{p}^* = \boldsymbol{M}^{-1}\hat{\boldsymbol{q}}$. In most cases, the result of the unconstrained problem satisfies the constraint. In the event that it doesn't, a solution is found by solving a small quadratic programming problem.

The above optimization converts pixels from LMS space into RGB space, but the values still need to be compressed to the appropriate dynamic range. For photopic images, the target dynamic range would be the full range of the display. However, low-light scenes are generally perceived as dim because illumination that falls in the mesopic or scotopic range doesn't approach the saturation level of cone receptors. As a result, images tone mapped for low light levels look odd when displayed on a bright monitor using the device's full range. We account for this effect by mapping to a reduced dynamic range. This method is similar to standard cinema approaches which use neutral density filters or shortened exposures in filming what should appear to be night scenes [16].

Our range reduction technique starts with standard HDR compression. We use the HDR compression algorithm of Durand and Dorsey [8] implemented using the bilateral filter approach developed by Paris and Durand [21]. From here, our operator further compresses pixels in the low dynamic range (LDR) image based on their mesopic levels. Our measure of mesopic level is based on rod attenuation in the rod-mediated opponent color model. Consider $\Delta o_{\text{Luminance}}$ in Equation 5.5. Rod contribution to $\Delta o_{\text{Luminance}}$ is at a minimum during photopic vision, which corresponds to the rod multiplier on $\Delta o_{\text{Luminance}}$ evaluating to zero. If we pull out this scalar factor as

$$w = \left(\alpha \frac{g_{\text{Long}}}{l_{\max}} + (1-\alpha)\frac{g_{\text{Medium}}}{m_{\max}}\right) \quad , \tag{5.9}$$

then a pixel is fully photopic when $w = 0$. The larger w grows, the further into the scotopic range we can consider the pixel and the more we reduce the dynamic range. By examining the average value of w across an image for a variety of exposures, we can pick a constant β for which we consider the pixel fully scotopic. In our examples, $\beta = 1$. Finally, we choose a parameter $\gamma \in [0\,1]$ which is the maximum ratio by which we reduce the range of a scene. In our examples, γ ranges between 0.25 and 0.5. Given these parameters, the maximum value x for a pixel with mesopic factor w is $x = \max(1 - \frac{w}{\beta}(1-\gamma), \gamma) * \nu$, where ν is the maximum value of the display's dynamic range, normally 255. A pixel with mesopic factor w is then linearly mapped into the range $[0\,x]$. See Figures 5.8 and 5.9 for a comparison of the mesopic factor to a standard grayscale image.

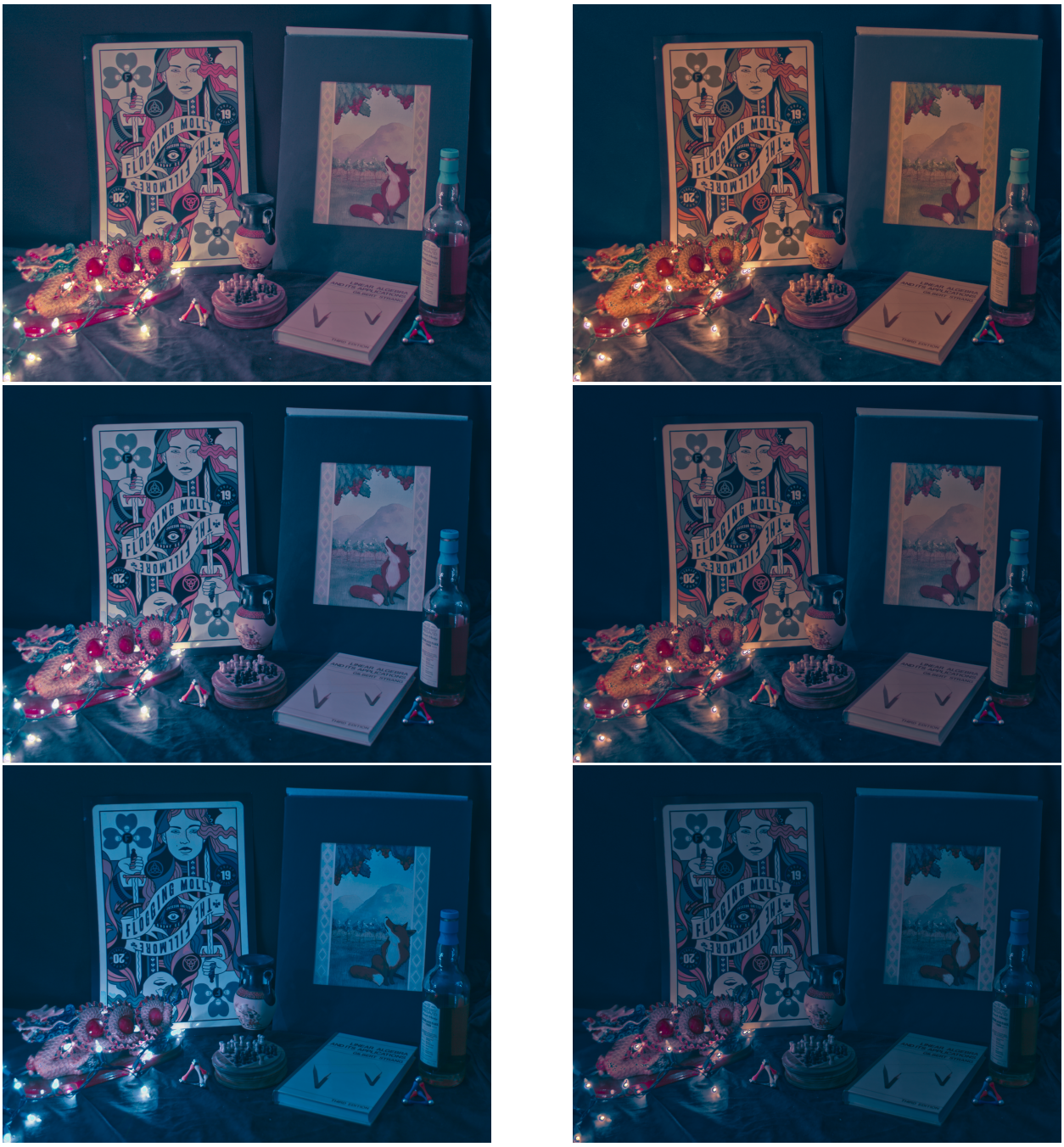

Figure 5.10: Comparison of tone mapping for low-light conditions using spectral and non-spectral HDR images. Left Column: Spectral image tone mapped for low-light conditions. Right Column: Non-spectral HDR image of the same scene mapped into LMSR space before applying our tone mapping procedure. For both columns, exposure decreases from top to bottom. This scene was not used in training our mapping from RGB space to LMSR space.

Original image from the OpenEXR dataset, copyright Industrial Light and Magic.

Figure 5.11: A non-spectral HDR image created by a camera with unknown spectral sensitivity.

5.2 Approximation for Non-Spectral Images

Spectral images contain information necessary to compute rod and cone responses for the tone mapping procedure we have described, however they also require special effort to acquire which may not always be feasible. For cases where spectral information isn't available, we describe a data-driven alternative that, when given an RGB value, infers a plausible LMSR value. This procedure requires lofting from a three-dimensional space to a four-dimensional one, and in the process some extra information must be generated. In particular, while the standard RGB primaries allow a good estimate of cone responses, they don't provide enough information for an independent estimate of rod response. We show that it is reasonable to learn this mapping.

Other researchers have reported that the reflectance spectra for many common scenes can be fairly well approximated using a small number of bases [17]. This result implies that for many scenes the rod response is strongly correlated with the cones. For plain RGB images

Original image from the OpenEXR dataset, copyright Industrial Light and Magic.

Figure 5.12: The result of our tone mapping operator applied to the image in Figure 5.11 after it was mapped into LMSR space.

we take advantage of this correlation by using a linear mapping from RGB space to LMSR space. For the examples in this paper shown in Figures 5.10, 5.11, 5.12, 5.13, and 5.14, we build the mapping using the spectra we acquired for other images. These other datasets give us a per pixel correspondence between RGB images and the LMSR response. Alternatively, one could build the mapping using existing material spectral databases [13] as long as they provide or make it possible to generate similar correspondences.

To build the mapping we solved for the matrix $\boldsymbol{H}$ that best satisfies $\boldsymbol{Q} = \boldsymbol{HP}$, where $\boldsymbol{Q}$ is a matrix of LMSR values from observed spectra and $\boldsymbol{P}$ is a matrix of the corresponding RGB values. For our data, this fit had high residual indicating that rod responses generally are linearly independent from the cones. However, it is not clear how sensitive viewers would be to the types of errors introduced, particularly if they have never viewed the scene firsthand under scotopic or mesopic conditions. Furthermore, mesopic vision is characterized by a reduced gamut which appears to make many errors less noticeable. We found the results to

Original HDR image "Foggy Night," copyright Jack Tumblin.

Figure 5.13: A non-spectral HDR image created by a camera with unknown spectral sensitivity.

Original HDR image "Foggy Night," copyright Jack Tumblin.

Figure 5.14: The result of our tone mapping operator applied to the image in Figure 5.12 after it was mapped into LMSR space.

be generally satisfactory for low light levels, but artifacts appear in brighter scenes. Luckily we have the original RGB values as a photopic reference. Based on this reasoning, we propose the following pipeline for non-spectral images. First, generate LMSR data using the linear mapping described above. Next, tone map the LMSR data to produce an initial mesopic image. Finally, for each pixel blend between the source image and the mesopic image based on the mesopic factor w to produce the final result.

Chapter 6

Results and Discussion

We present several scenes for evaluation throughout this dissertation. None of these images have been subjected to any photo-processing other than the methods described, plus cropping, rotating, and downsampling the final images for publishing. The raw input images and resulting spectral images are included in the supplemental data. We also include LMSR images and standard RGB HDR images. Supplemental data can be found at `http://graphics.berkeley.edu/papers/Kirk-PBT-2011-08/raw_data.zip`.

Figure 6.1 shows a visual comparison of the steps in our method. The left column shows our input sources. Each image in the left column is an HDR image under a different transmissive filter. The second through fifth columns demonstrate different aspects of our methods as exposure varies. Each row in the second through fifth columns has the same exposure. The second column shows the result of our range reduction technique without performing tone mapping. Note that even in the darkest image it is still possible to distinguish color. The third column shows how our mesopic factor changes with exposure. The fourth column shows the result of our tone mapping procedure mapped to the full range of the display. These images clearly demonstrate distinctive hue shifts. The final results of our full algorithm are shown in the fifth column. In the fourth and fifth column, note that similar reds in the apple and mango map to drastically different colors in the shortest exposure.

The majority of the images in this paper demonstrate our tone mapping operator on spectral images. In Figures 5.1 and 5.2, we show how our tone mapping operator varies with proximity to lighting. In the original image, the reddish-brown of the Scotch whisky in the bottle furthest from the light source deepens considerably, while the reds in the dragon remain bright. Figures 5.3 and 5.4 shows an outdoor scene. Neighboring pixels can have extreme lighting discontinuities, and the contrast between the trees and the tower shows our method respects these discontinuities. The well-lit tree between the tower and the building as well as the white walls of the building are good examples of how different colors shift in the transition from photopic to scotopic light conditions. Figures 5.6 and 5.7 show an outdoor statue captured at night. The tone mapped version is not only closer to the viewer's experience in near dark, but also demonstrates a useful artistic effect. Figure 5.5 shows a

night skyline. The original HDR image contains a large amount of light scattering. The tone mapped version demonstrates a reduction in the perceptual effect of this light. Figures 1.1, 1.2, and 1.3 show another night scene. The original image contains little contrast between the doorway and the walls of the building, while there is a strong change in hue in the tone mapped image. Comparing the hue shift for the red brick to the hue change for the bushes and trees demonstrates the difference in relative hue change for different colors. The red brick underneath the street light contains another example of hue change as a function of proximity to light sources.

Finally, we demonstrate our tone mapping process using non-spectral images as input. In Figure 5.10 we compare results from tone mapping spectral data and plain RGB data. Note that while spectral data has a smoother transition as the exposure decreases, the RGB data results in a similar scotopic image and still demonstrates hue change based on proximity to light sources. Figures 5.11, 5.12, 5.13, and 5.14 compare standard HDR images with our tone mapped versions. Note that bright areas in the HDR images and tone mapped images are similar, while the differences in the shadows are more significant.

The human visual system has many adaptive mechanisms, and our work focuses on just one, namely the interplay between rods and cones in the early stages of vision. Subsequent to the processing that occurs in the eye there are other adaptive mechanisms, such as color constancy, that are not incorporated into our perceptual mapping. Furthermore, there are additional scotopic and mesopic effects that we chose not to model. These include well-studied effects such as the loss of acuity and increased noise [9]. If desired, these effects could be combined with our perceptual tone mapping method.

One somewhat arbitrary aspect of our method is that the range of intensities in an image can be scaled however one desires. For example, one could adjust the exposure parameter so that a bright, sunlit scene was rendered with scotopic treatment. Generally doing so will not produce the impression of a nighttime scene without additional processing. Among other issues, the color of the illumination from the sun would be inconsistent with a night scene as would the detail and color in the sky. A less extreme example appears in Figures 1.1, 1.2, and 1.3. The lamp in the center of the scene and other light sources provided a fair amount of illumination, and the middle image is most consistent with our firsthand impression of the scene. Figure 1.3 image takes some artistic license and corresponds to a nonexistent location with much dimmer lighting.

The methods we have described provide a perceptually motivated model of how human viewers perceive low-light scenes. We believe that they will have uses in applications where one wishes to predict actual viewer experiences, convey the correct impression of low-light scene, or generate artistic effects.

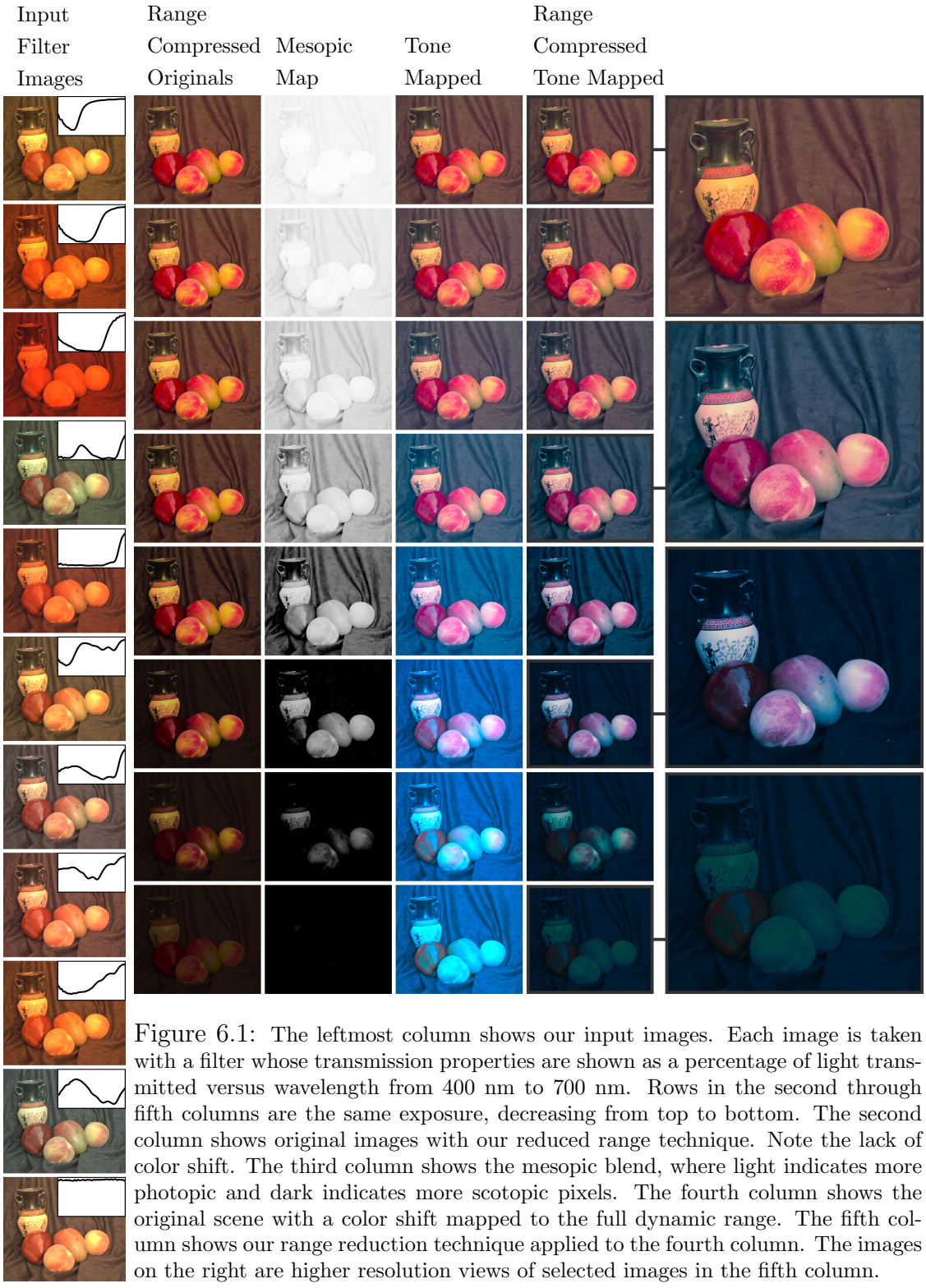

Figure 6.1: The leftmost column shows our input images. Each image is taken with a filter whose transmission properties are shown as a percentage of light transmitted versus wavelength from 400 nm to 700 nm. Rows in the second through fifth columns are the same exposure, decreasing from top to bottom. The second column shows original images with our reduced range technique. Note the lack of color shift. The third column shows the mesopic blend, where light indicates more photopic and dark indicates more scotopic pixels. The fourth column shows the original scene with a color shift mapped to the full dynamic range. The fifth column shows our range reduction technique applied to the fourth column. The images on the right are higher resolution views of selected images in the fifth column.

Bibliography

[1] S.L. Buck, R.F. Knight, and J. Bechtold. Opponent-color models and the influence of rod signals on the loci of unique hues* 1. *Vision research*, 40(24):3333–3344, 2000.

[2] J.D. Bullough and M.S. Rea. Simulated driving performance and peripheral detection at mesopic and low photopic light levels. *Lighting Research and Technology*, 32(4):194, 2000.

[3] Bureau Central de la CIE. *CIE Proceedings (1951)*, volume 3, Paris, 1951.

[4] Peter D. Burns and Roy S. Berns. Analysis of multispectral image capture. In *Proc. Fourth Color Imaging Conference: Color Science, Systems, and Applications*, pages 19–22, 1996.

[5] D. Cao, J. Pokorny, V.C. Smith, and A.J. Zele. Rod contributions to color perception: linear with rod contrast. *Vision research*, 48(26):2586–2592, 2008.

[6] PE Devebec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. SIGGRAPH97*, pages 369–378, 1997.

[7] F. Durand and J. Dorsey. Interactive tone mapping. In *Proceedings of the Eurographics Workshop on Rendering Techniques*, pages 219–230. Citeseer, 2000.

[8] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 257–266. ACM, 2002.

[9] J.A. Ferwerda, S.N. Pattanaik, P. Shirley, and D.P. Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 249–258. ACM, 1996.

[10] Douglas S. Goodman. Basic optical instruments. In Daniel Malacara, editor, *Geometrical and Instrumental Optics*. Academic, 1988.

[11] Eugene Hecht. More on geometrical optics. In *Optics*. Addison Wesley Longman, Inc., 2002.

[12] S.M. Khan and S.N. Pattanaik. Modeling blue shift in moonlit scenes by rod cone interaction. *Journal of VISION*, 4(8), 2004.

[13] Oili Kohonen, Jussi Parkkinen, and Timo Jääskeläinen. Databases for spectral color science. *Color Research and Application*, 31(5):381–390, October 2006.

[14] Jiangtao Kuang, Garrett M. Johnson, and Mark D. Fairchild. icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation*, 18:406–414, 2007.

[15] M. Kumar, E.O. Morales, JE Adams, and W. Hao. New digital camera sensor architecture for low light imaging. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 2681–2684. IEEE, 2009.

[16] Kris Malkiewicz. *Film Lighting: Talks with Hollywood's Cinematographers and Gaffers*. Simon and Schuster, 1992.

[17] Laurence T. Maloney and Brian A. Wandell. Readings in computer vision: issues, problems, principles, and paradigms. chapter Color constancy: a method for recovering surface spectral reflectance, pages 293–297. 1987.

[18] Rafal Mantiuk, Allan G. Rempel, and Wolfgang Heidrich. Display considerations for night and low-illumination viewing. In *Applied Perception in Graphics and Visualization*, pages 53–58, 2009.

[19] E. Miyahara, V.C. Smith, and J. Pokorny. How surrounds affect chromaticity discrimination. *Journal of the Optical Society of America A*, 10(4):545–553, 1993.

[20] Ankit Mohan, Ramesh Raskar, and Jack Tumblin. Agile spectrum imaging: Programmable wavelength modulation for cameras and projectors. *Computer Graphics Forum*, 27(2):709–717, 2008.

[21] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. In *European Conference on Computer Vision*. Springer, 2006.

[22] J. Park, M. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral Imaging Using Multiplexed Illumination. In *IEEE International Conference on Computer Vision (ICCV)*, Oct 2007.

[23] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings SIGGRAPH 98*, pages 287–298, 1998.

[24] Y.Y. Schechner and S.K. Nayar. Generalized Mosaicing: Wide Field of View Multispectral Imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(10):1334–1348, Oct 2002.

[25] J.C. Shin, N. Matsuki, H. Yaguchi, and S. Shioiri. A color appearance model applicable in mesopic vision. *Optical Review*, 11(4):272–278, 2004.

[26] J.C. Shin, H. Yaguchi, and S. Shioiri. Change of color appearance in photopic, mesopic and scotopic vision. *Optical Review*, 11(4):265–271, 2004.

[27] A. Stockman and L.T. Sharpe. The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000.

[28] J. Tumblin and H. E. Rushmeier. Tone reproduction for computer generated images. *IEEE Computer Graphics and Applications*, 13(6):42–48, 1993.

[29] F. Yasuma, T. Mitsunaga, D. Iso, and S.K. Nayar. Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum. Technical report, Nov 2008.