Accurate Reproduction of Real-World Lighting

by

Brian Karr

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Manufacturing Group

2019
## Contents

List of Tables iv  
List of Figures v  
List of Publications viii  
Acknowledgments ix  
Declarations x  
Abstract xi  

### Chapter 1 Introduction 1

1.1 Research Aim and Objectives 2  
1.2 Current State of the Art in HDR 3  
1.3 Evaluating Dynamic Range of Imaging Systems 5  
1.4 Radiometric Calibration 6  
1.5 Thesis Outline 6  

### Chapter 2 Background and Related Work 8

2.1 Background 8  
2.1.1 Radiometry, Photometry, and Luma 8  
2.1.2 Colourimetry 12  
2.1.3 Dynamic Range and Signal to Noise Ratio (SNR) 20  
2.1.4 Data Formats 25  
2.2 Related Work 34  
2.2.1 Analysis of Capture Processing Pipelines 34  
2.2.2 Gaps in Capture Processing 37  
2.2.3 Multiple Exposure HDR Techniques 38  
2.2.4 Noise Based Weighted Average 39
Chapter 3 Methodology

3.1 Objectives .............................................. 42
3.2 Approach ............................................. 44
3.3 Rationale .............................................. 45
3.4 Structure of Research ................................. 46
	3.4.1 Chapter 4: Analysis of a Real World Rocket Launch Environment .... 46
	3.4.2 Chapter 5: Calibrated Measurement of System Dynamic Range ........ 47
	3.4.3 Chapter 6: Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model ........................................... 48
3.5 Summary .............................................. 48

Chapter 4 Analysis of a Real World Rocket Launch Environment .......................... 50

4.1 Background ............................................. 51
4.2 Film Baseline and History ............................... 52
	4.2.1 Dynamic Range of Film ............................... 54
	4.2.2 Auto-Exposure and Contrast Ratio ....................... 55
	4.2.3 Linear and Non-linear Quantization ..................... 56
	4.2.4 Playback and Display ................................... 57
	4.2.5 Summary of Film Use ................................... 58
4.3 HDR Imaging of Spacecraft Field Experiments ............................................. 59
	4.3.1 High Definition Camera Test During STS-129 ................... 59
	4.3.2 RED ONE M Camera Test During STS-131 .................. 61
	4.3.3 HDR Imaging during the Final Space Shuttle Flight STS-135 ........ 63
	4.3.4 RED EPIC HDR Mode during Delta 4 WGS 5 .................. 70
4.4 Review of HDR Field Experiments ............................................. 71
4.5 Summary .............................................. 73

Chapter 5 Calibrated Measurement of System Dynamic Range .............................. 74

5.1 Method and Materials .................................... 75
	5.1.1 Design ............................................ 75
	5.1.2 Materials .......................................... 76
5.2 Test Procedure ........................................ 79
	5.2.1 Data Workflow ..................................... 81
5.3 Results ............................................... 83
5.4 Quality Descriptors .................................... 84
Chapter 6 Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model

6.1 Introduction 88
6.2 Calibrating a Multiple Exposure HDR Workflow 91
   6.2.1 Bayer Raw (BR) Data 92
   6.2.2 Performance Parameters via PTC 93
   6.2.3 Relationships of Noise Parameters 94
   6.2.4 Read and Shot Noise 96
   6.2.5 Per Pixel Gain Factors 97
   6.2.6 Radiometric Scene Calibration 98
   6.2.7 Glare 99

6.3 Implementation and Results 100
   6.3.1 Method 100
   6.3.2 Effect of Glare on the Calibration 105
   6.3.3 Spatial Nature of Glare 108

6.4 Discussion 110
   6.4.1 Variability Due to Sensor and Lens 110
   6.4.2 Variability Due to Image Content 112
   6.4.3 Comparison to Commercially Available Software 113

6.5 Summary 114

Chapter 7 Conclusions and Future Work 115

7.1 Contributions 115
   7.1.1 Analysis of a Real World Rocket Launch Environment 116
   7.1.2 Calibrated Measurement of System Dynamic Range 117
   7.1.3 Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model 119

7.2 Impact 121
7.3 Future Work 122

Appendix A Dynamic Range Conversion Chart 125
List of Tables

2.1 Photometric and corresponding radiometric quantities ............... 11
2.2 Capture and Processing Pipelines ........................................ 38

4.1 Summary of HDR Imaging Spacecraft Field Experiments ............. 60

5.1 Xyla-21 Density Calibration Data ........................................... 78
5.2 Post-Processing Workflow Summary ....................................... 83
5.3 Dynamic Range Measurement of Digital Imagers .......................... 85

6.1 Summary of PTC Determined Values ..................................... 103
6.2 Canon 5DM3 #1 with Sigma 24-70 mm #1 Compared to Similar
               Equipment ............................................................. 112
List of Figures

1.1 Example exposures, varied via post-processing, illustrating overall dynamic range of an oxygen tank over-pressurization event. Source: Advanced Imaging Lab, Kennedy Space Center. 2
1.2 Night launch Space Shuttle Discovery illustrating wide range of light 4

2.1 CIE photopic luminous efficiency curve \( V(\lambda) \) 10
2.2 Spectral sensitivities of the long (L), medium (M), and short (S) cones [Stockman et al., 1993] 13
2.3 CIE 1931 RGB colour matching functions [Schanda, 2015] 14
2.4 Chromaticity plane \( xy \) projected in \( XYZ \) space [Fairman et al., 1997] 15
2.5 CIE 1931 rg (xy) chromaticity diagram [Fairman et al., 1997] 16
2.6 Colour matching functions of the CIE 1931 standard (2° : full lines) and CIE 1964 standard (10° : dotted lines) [Schanda, 2015] 18
2.7 Image and Noise Transfer Model 22
2.8 Arbitrary Quality Descriptions and SNR in Imatest 25
2.9 sRGB Encoding Luminance vs. Code Value 28
2.10 IEEE Float vs Encoded Value 30
2.11 OpenEXR Half Float vs Encoded Value 31
2.12 HDR LogLuv36 TIFF Encoding Luminance vs. Code Value 33

3.1 Research strategy 43
3.2 HDR Workflow 44
3.3 Night-time launch image of Delta IV 4 rocket taken with a RED EPIC camera operating in HDR+6 mode. 45

4.1 Engine debris during night mission 51
4.2 Characteristic curve for a negative emulsion 55
4.3 Film to digital transfer with display on 8 bit monitor 57
4.4 HD1100 Cam A (1/60 exposure) 61
6.10 Scene 1 multiple exposure single images (L:R, T:B, 1/8000, 1/2000, 1/500, 1/125, 1/30, 1/8, 2, 8, 32 secs) ........................................ 105
6.11 Scene 1 multiple exposure recombined false colour image of Canon 5DM3 with Zeiss lens ....................................................... 106
6.12 Luminance percent difference for scene 1 grey scale patches, Canon 5DM3 with Sigma 24-70 mm lens at 24 mm ................................. 107
6.13 Luminance percent difference for scene 1 grey scale patches with Zeiss 25 mm lens .......................................................... 108
6.14 Luminance percent difference for scene 1 with diffuse lighting using Zeiss 25 mm lens .......................................................... 109
6.15 Scene 2 multiple exposure single images (L:R, T:B, 1/8000, 1/2000, 1/500, 1/125, 1/30, 1/8, 2, 8, 32 secs) ................................. 110
6.16 Scene 2 with dark regions both near and far from high luminance levels ................................................................. 111
6.17 Luminance percent difference for scene 2 with dark regions near and far from high luminance levels ........................................ 111
6.18 Canon 5DM3 Camera Body #1: Difference Plot with Sigma 24-70 mm Lens #1 versus Sigma 24-70 mm Lens #2 ....................... 112
6.19 (L) Our calibrated recombination method (R) Photosphere method . 113
7.1 Example radiometrically calibrated scene measurement of a spacecraft launch at night, tone-mapped for display .......................... 123
7.2 Example radiometrically calibrated scene measurement of a spacecraft launch at night, false-colour version ............................... 124
A.1 Dynamic Range Conversion Chart ........................................ 126
List of Publications

Publications


Publications under preparation

I would like to express my sincere appreciation to my advisors, Alan Chalmers and Kurt Debattista, whom have skilfully guided me through this endeavour. I set several initial goals for myself, to learn to think as a researcher as much as an engineer, to understand the research process and value of collaboration and peer review, and finally to dive deeper into a world for which I am passionate about, all of which were achieved as a result of your leadership and efforts. The best thanks I can give, is to say that I am completing this process with the same motivation and excitement as when I began.

Thanks to the many new colleagues and friends I have met at conferences and meetings. I believe a true measure of an individual’s knowledge and expertise is the ability, and willingness, to share it with others. I am honoured to have met, discussed, and learned from so many fellow researchers with whom I have crossed paths.

Finally, I would like to thank my family, including my wife Tammy, and son Adam. Your sacrifices of missed time, and your never-ending support, is appreciated beyond words.
Declarations

The work in this thesis is original and no portion of this work has been submitted in support of an application for another degree or qualification at this university or at another university or institute of learning.
Abstract

The dynamic range of a scene is defined as the difference in lighting between the brightest and darkest areas of that scene. The goal of High Dynamic Range (HDR) imaging is to capture all of the detail in a scene, no matter what the dynamic range. HDR is essential when capturing safety critical environments, for example during a rocket launch. Here it is crucial to be able to see simultaneously all the detail in the rocket plume, on the rocket itself and within the launch environment. Although the capture of HDR images and video has been well studied, little is known about the physical accuracy of the captured image compared to the scene itself. Just how truly representative is the captured HDR image of the real world lighting? In this thesis, the image capture pipeline is analysed in detail, while not assuming linearity, in order to maintain physical accuracy throughout the workflow. Rather than approximating the camera response function, calibrated measurements are employed to characterize both the image sensor and the associated optics.

As a first step in this research, existing HDR capture technology was investigated. A number of latest generation digital imaging systems were selected for testing and evaluation, and several limitations are identified in the evaluation of existing HDR capture technology. First, in most cases, no standard is utilized when stating imager dynamic range capabilities. Second, in-camera processing of data is described generically, such as with a vendor specific log curve applied, but only in some cases is the processing described, or the reverse transform provided in order to linearize the data. Third, there is lack of discussion, or available data, related to acceptable noise level or the noise level used in the determination of signal to noise based sensor characterizations. From the imager test and evaluation, it was concluded that the processing pipelines differ significantly from manufacturer to manufacturer, and that there is no clear definition as to how imaging system dynamic range is either measured or specified.

To address these limitations, a new method for characterizing HDR response from a “system” perspective was developed, including the combined effect of the optics, the sensor hardware, and the camera processing. The results indicated that the accuracy of manufacturer stated capabilities are dependent upon the chosen maximum RMS noise defining the minimum luminance step used to determine dynamic range. Different manufacturers appear to use different metrics, making comparison difficult.

The model was further improved to treat the optical and sensor elements separately, so that the contributions from each are known. To provide maximum accuracy, the model in this thesis categorizes error via methodical calibration; including determination of sensor noise attributes, system gain response, and the effects of lens glare. Additionally, using reference measurements from a calibrated luminance meter, absolute radiometric calibration was achieved, albeit with a degree of error. The results highlighted that the effects of glare limit the ability to accurately capture and measure the lower luminance portion of a scene. To capture an HDR image most accurately, glare effects must be reduced by using higher quality optics, or by reducing sources of glare in the scene.
Chapter 1

Introduction

There is a wide range of light and colour in the real world. While the human eye is capable, through adaption, of seeing scenes from starlight ($10^{-5}$ cd/m$^2$) to bright sunshine ($10^6$ cd/m$^2$) traditional imaging technology is not. Such Low Dynamic Range (LDR), frequently also called Standard Dynamic Range (SDR), technology is only able to capture light in a scene where the difference between the darkest and brightest regions, the so called dynamic range, is 256:1. This can result in over- or under-exposed areas of the captured image where substantial information is missing.

As an example, figure 6.1 shows different exposures of a frame from a test flight of an experimental NASA launch and landing test vehicle developed under the project name Morpheus [Olansen, 2014]. While the Morpheus vehicle ultimately resulted in 13 successful free flights, during an initial test flight the vehicle crashed, resulting in the oxygen tank over-pressurization and explosion shown in figure 6.1. Had the capture been made using an LDR process, shadow and highlight detail, important to the root cause crash investigation, may have been lost during the capture process. High Dynamic Range (HDR) imaging provides a potential solution to overcome this problem, and in this example the capture device, an ARRI Alexa, is a professional level digital camera with reported dynamic range capability of 14 stops. The greater dynamic range allows for additional capture detail in the shadows and highlights. Figure 6.1(a) shows a single exposure post-processed for display on a typical LDR monitor (for visualization purposes, as the full range would not otherwise be visible on print or on a traditional display). With the exposure optimised per the pre-explosion average luminance level of the scene, the oxygen cloud is saturated with mostly white pixels. The raw data however contains additional luminance detail in the oxygen cloud, and when re-processed, shifting the displayed exposure range towards the highlights, the oxygen cloud detail is
then observed as shown in figure 6.1(b). In figure 6.1(c), the exposure range is further shifted, such that detail in the brightest area of the oxygen cloud can be seen. This ability to capture a wide range of light provides the opportunity to view and work with additional detail in the image, and in some cases with physically accurate representations of real world lighting. Although the images serve as a good example and includes bright highlights, the total luminance range of the scene remains unknown. Further, being a daylight scene, the total dynamic range is not expected to be as wide as would be encountered with a night scene that includes a similar or brighter flame. A goal of this thesis is to investigate and calibrate those types of scenes containing extremely wide dynamic ranges, in the ranges from starlight ($10^{-5}$ cd/$m^2$) to direct sunlight ($10^6$ cd/$m^2$).

(a) Frame with missing (saturated) detail in highlights  (b) Frame detail recovered in oxygen cloud  (c) Frame detail recovered in highlights

Figure 1.1: Example exposures, varied via post-processing, illustrating overall dynamic range of an oxygen tank over-pressurization event. Source: Advanced Imaging Lab, Kennedy Space Center.

1.1 Research Aim and Objectives

HDR capture forms a fundamental first step in the HDR pipeline. As of yet there is little literature on how best to capture physically accurate linear HDR data without losing vital information. Although there is a wealth of publications related to the HDR pipeline including: capture, processing, tone-mapping, and display, the dynamic range of the test datasets is rarely taken to be physically accurate or truly representative of real world lighting. This presents a problem when working to characterise scene(s), when specifying the capabilities of pipeline equipment, and when comparing manufacturer specifications and/or processing techniques.

The main aim of this thesis is to evaluate and calibrate for real world luminance values at every stage of the capture pipeline. This thesis proposes a novel method for the calibrated capture of HDR imagery, develops a solution for the collection of physically accurate radiance maps, and investigates limitations due to the image sensor, camera processing, and optics.
The main objectives of this work are:

1. A comprehensive literature review of HDR, in order to obtain a detailed knowledge of previous research in the field.

2. An analysis of HDR capture systems, challenges in implementing these systems, and differences between workflows. The launch environment, as described below, is utilized as a test bed in reviewing the state of the art.

3. A new calibrated method to accurately measure and objectively compare the dynamic range capture capability of modern imaging systems using off the shelf equipment.

4. The proposal of a new complete HDR capture workflow with known and calibrated dynamic range and a physically accurate radiance map.

5. To clearly demonstrate the new system’s capability to reproduce real world lighting via a challenging test scenario, and to discuss the efficacy and limitations of this new method.

1.2 Current State of the Art in HDR

The range of light encountered in the real world can be huge. Using another spacecraft example, figure 1.2 shows a night Space Shuttle launch, where detail is equally important in shadow regions for detached external tank foam identification, as well as in bright regions, where debris can pass in front of the flame plume. Accurate imaging is crucial during the launch to enable ground controllers to determine if there have been any anomalies, such as an object becoming detached from, or striking, the vehicle. Traditional film, which has a dynamic range of approximately 15 stops, has been used as the primary method to record Space Shuttle launches. The use of film precludes the images being studied in real-time (in fact the footage typically is only available to be viewed, after it has been developed and printed, several days after the launch). HDR imaging offers the potential for investigating all the detail on the vehicle in real time during the launch or shortly afterwards.

The application of HDR to spacecraft launches is a completely uncontrolled real-world problem that requires HDR to be accurate and the slightest discrepancies to be understood in order to be able to take decisions with confidence from the resulting footage. It therefore represents a highly challenging scenario that is unachievable without a robust and quantifiable HDR workflow. A transition from film cameras to digital imagers is anticipated in the post-Space Shuttle era at Kennedy
Space Center, Florida, USA. Analysis of digital imaging systems and comparison to the film baseline is important for determining the transitional roadmap. Resolution, dynamic range, uniformity, colour response, sharpness, timing, etc. are examples of imaging metrics for consideration. The focus of the research in this thesis is dynamic range. As previously discussed, traditional imaging methods are unable to capture the wide range of luminance present in a natural scene. The ability to capture engineering quality imagery with a wide degree of dynamic range during rocket launches is critical for post launch processing and analysis, especially if a problem occurs [Columbia and Gehman, 2003; Challenger, 1986].

Looking past rocket launch examples, other market segments potentially benefiting from HDR include building lighting modelling, gas plume analysis, image processing, and most recently, the introduction of autonomous vehicles driving in highly variable lighting environments. To help address these challenging scenarios, HDR technology is currently witnessing significant and rapid advancement. HDR capture systems, processing techniques, and display devices are evolving, with new products increasingly coming to market. Examples include popular camera systems such as those available from Red Inc. [RED, 2015a]. and ARRI Inc. [ARRI, 2015] that at a minimum meet, and can exceed, the Dynamic Range (DR) capability of film [Karr et al., 2017]. Technologies now available include processing and display
software from goHDR Ltd. [goHDR, 2014], on-camera HDR monitoring support from RED [RED, 2015b], increased dynamic range professional displays from Sim2 [Sim2, 2014], Dolby [Dolby, 2015], and Sony [Sony, 2015], as well as consumer HDR displays from LG [LG, 2015] and Panasonic [Panasonic, 2015]. In order to fully realize the potential benefits afforded by these new technologies, primarily images that more accurately represent real-world lighting and an improved viewing experience, a complete HDR production workflow is required.

1.3 Evaluating Dynamic Range of Imaging Systems

The problem currently encountered in the capture portion of the HDR pipeline is the lack of a clear definition of the measurement process. HDR data is often being presented without a common reference, making comparisons difficult [Bantterle et al., 2011]. As interest in HDR grows, especially with its inclusion in the UHDTV definition (ITU-R Recommendation BT.2020) [ITU-R, 2012b] a number of imaging vendors are now beginning to offer capture systems which they claim to be HDR [Chalmers and Debattista, 2017]. To validate such claims, new methods are needed that evaluate system dynamic range across a wider luminance range. Dynamic range, as it relates to noise, can be viewed as one aspect of image quality, the evaluation of which can be broadly categorized as either subjective or objective. Subjective evaluation generally utilizes a large sample of human observers and a carefully designed experiment, often based on a standardized test procedure [Pappas et al., 2000; ITU-R, 2012a]. Subjective methods are considered to be the most reliable when considering the Human Vision System (HVS), however they can be costly, time consuming, and demanding when considering the knowledge and effort required to obtain meaningful results [Reinhard et al., 2010]. The complex behaviour of the HVS response and perception of brightness, including content-dependency, to chromatic and achromatic images is known as Perceived Dynamic Range (PDR) [Hulusic et al., 2017]. Objective evaluation has the goal of accurately predicting standard subjective quality ratings, and when selected and designed for a specific application, can also produce reasonable cost-efficient results [Narwaria et al., 2015]. Objective evaluations make use of computational methods to assess the quality of images, thus human factors such as mood, past experience, and viewing environment are not considered. Tests can be practically performed for large test samples, and repeatable results can be obtained utilizing identical test conditions. In the case of our proposed dynamic range calculation, a no-reference metric is based on the evaluation of noise, using a single frame of a 21 density step HDR test chart as an input.
to determine the ratio of the maximum unclipped input luminance to the minimum acceptable input luminance level meeting specified noise criteria. Measured results are used to compare the increased dynamic range capabilities of modern image capture systems, and to examine the arbitrary assignments of perceived quality terms to the results.

1.4 Radiometric Calibration

HDR data may be created via a variety of methods including multiple exposure techniques, computer graphic generation, beam splitting using multiple sensors, and extended range single sensor systems [Banterle et al., 2011]. Computer vision algorithms, tone-mappers, and other HDR data processing systems usually require the recorded luminance values to be linearly related to scene radiance. Furthermore, if calibration between measured data and actual luminance values is desired, linear data is required before a luminance factor can be applied as a constant multiplication factor relating the relative and actual luminance. A number of studies have examined specific use cases requiring absolute calibration of the resulting HDR luminance map [Mitsunaga and Nayar, 1999; Debevec and Malik, 2008; Inanici and Galvin, 2004; Coutelier and Oumortier, 2003], but none of these has provided a detailed analysis of the complete HDR capture workflow, and in particular the impact of each of the stages on the overall resultant accuracy. As part of the radiometric calibration, this thesis will examine in detail the linearization of raw data from a sample camera, implement the recombination of a HDR scene from full resolution multiple exposures based on a rigorous noise model, and examine the Per Pixel Gain Response (PPGR) and optical effects on the results. Camera noise measurements are determined via generation of the Photon Transfer Curve. A purposely designed controlled test scene is used to challenge the calibrated reconstruction efforts, including low luminance levels, glare, and spatial inclusion of lens vignette over the full imaged area. Results demonstrate that even with careful processing and recombination of the LDR data, the total system dynamic range is limited in the shadow region as a result of optical glare. A significant improvement of the new method over standard practice using a commercially available HDR recombination software is also demonstrated.

1.5 Thesis Outline

This thesis consists of seven chapters arranged as follows:
Chapter 2 - Background and Related Work: Provides an overview of relevant HDR terms, discuss key concepts such as dynamic range and signal to noise ratio, and examines data formats used to process and store HDR data. Additionally, a literature review of modern methods addressing the HDR capture pipeline will be presented, along with identified gaps in existing pipelines.

Chapter 3 Methodology: Provides details as to how the new calibrated HDR capture system will be designed, implemented, and thoroughly evaluated.

Chapter 4 - Analysis of a Real World Rocket Launch Environment: As a compelling example of the need for accurate HDR imaging, this chapter provides an analysis of the real-world HDR problem; the capture of a rocket launch during day-light and night-time conditions.

Chapter 5 - Calibrated Measurement of System Dynamic Range: Presents the techniques developed to accurately measure and objectively compare imaging systems dynamic range, and examines the arbitrary assignment of perceived quality terms to the results.

Chapter 6 - Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model: In order to accurately determine the optical effects on the HDR workflow, we examine in detail the effects of glare and vignette on the linearization of raw data from a sample camera, implementing the recombination of an HDR scene from full resolution multiple exposures based on a rigorous noise model.

Chapter 7 - Conclusions and Future Work: This concludes the work and describes directions for future research.
Chapter 2

Background and Related Work

This chapter will present background and related work. In particular, background terms and definitions used to describe light are first presented, followed by a review of dynamic range and signal to noise ratio, as well as the noise model as it relates to image capture. Data storage formats for captured imagery are then discussed. These are necessary for the proper processing of HDR data. Finally, an analysis of current processing pipelines is given, along with the existing gaps in these workflows as they relate to HDR processing.

2.1 Background

2.1.1 Radiometry, Photometry, and Luma

Radiometry  Radiometry is the measurement of electromagnetic radiation energy that when considering ultraviolet, visible, and infrared light, ranges spectrally in wavelength from 100nm to 1mm. It can be used to describe the transfer of energy or power from a source to a detector, and may be integrated over time, space, or angle. Radiometric quantities are based on optical power, termed the radiant energy:

\[
radiant\text{ }energy = Q_e = Joules
\]  \hspace{1cm} (2.1)

If the radiant energy is considered with respect to time \(t\), the time derivative is defined as the radiant flux (or radiant power) emitted, reflected, or transmitted per unit time as:

\[
radiant\text{ }flux\text{ }\text{(or power)} = \Phi_e = \frac{\delta Q_e}{\delta t} = Joules\text{ }second = Watts
\]  \hspace{1cm} (2.2)
In radiometry, a fundamental quantity is known as the radiant sterance, or radiance. It can be used to calculate the radiant flux onto a system, or more formally, the amount of radiant flux (watts) radiated into a cone of incremental solid angle $dw$ (steradians), from a source whose area is measured in squared meters $A$ [Holst and Lomheim, 2007].

$$\text{radiance} = L_e(x, \omega) = \frac{\delta \Phi_e(x, \omega)}{\delta \omega \cdot \delta A} = \frac{\text{Watts}}{\text{steradian} \cdot \text{meter}^2} \quad (2.3)$$

In picture formation, light that strikes a camera sensor pixel focused via a lens, from a particular direction, is the measurement of radiance, and thus important as radiance is proportional to the pixel voltage generated. For Lambertian surfaces, having matte properties and adhering to Lambert’s cosine law, radiance emitted by a surface per unit area is defined as the radiant exitance (or radiant emittance):

$$\text{radiant exitance} = M_e(x) = \frac{\delta \Phi_e^{\text{emitted}}}{\delta A} = \frac{\text{Watts}}{\text{meter}^2} \quad (2.4)$$

Similarly, the radiant flux received onto a target, a real or imaginary unit surface area, arriving from all directions over a hemisphere, is defined as the radiant incidence or irradiance (or radiant flux density):

$$\text{irradiance} = E_e(x) = \frac{\delta \Phi_e^{\text{arriving}}}{\delta A} = \frac{\text{Watts}}{\text{meter}^2} \quad (2.5)$$

The radiant flux specified in a particular direction from an infinitesimally small point light source, as a function of unit solid angle, is defined as the radiant intensity:
radiant intensity \( I_e = \frac{\delta \Phi_e}{\delta \omega} \) \( \text{Watts steradian} \) (2.6)

**Photometry**  Photometry is analogous to radiometry, with the differentiation of integration over the visible wavelengths and spectral weighting to the human visual response. Formally, the unit of radiation is normalized to the spectral sensitivity of the human eye as described by the 1924 CIE luminous efficiency function \( V(\lambda) \). The human eye is sensitive to wavelengths approximately between 380 and 830 nm, however human sensitivity to visible light is not uniform across this spectral band. An approximation to the spectral response for the average viewer (the CIE Standard Observer) is known as the photopic luminous efficiency curve as shown in figure 2.1 (correlation with visual colour matching of fields subtending between about 1° and about 4° at the eye of the observer, data shown based on 5nm intervals).

![Figure 2.1: CIE photopic luminous efficiency curve \( V(\lambda) \)](image)

Radiometric terms weighted by the photopic luminous efficiency curve result
in their photometric counterparts as listed in table 2.1. Luminous flux is defined as 1/683 Watts of radiant power at a frequency of $540 \times 10^{12}$ Hz, a wavelength of approximately 555 nm corresponding to the peak luminance efficiency as seen in Figure 3. Luminance is photometrically weighted radiance and can be used as an approximate measure of how bright a surface appears [Reinhard et al., 2010]. Luminous flux leaving a surface is luminous exitance, while luminous flux arriving at a surface is illuminance. Illuminance is the familiar quantity measured by a hemisphere light meter, and taking the illuminance times the surface reflectance (non-emissive surface) results in the luminous exitance. Finally, illuminous flux density in a specified direction is luminous intensity, which is independent of distance.

Table 2.1: Photometric and corresponding radiometric quantities

<table>
<thead>
<tr>
<th>Photometric Quantity</th>
<th>Photometric Units</th>
<th>Corresponding Radiometric Quantity</th>
<th>Differentiate Radiant Flux with</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminous Energy ($Q_v$)</td>
<td>$lm \cdot s = Talbot$</td>
<td>Radiant Energy</td>
<td>-</td>
</tr>
<tr>
<td>Luminous Flux (Power) ($P_v$)</td>
<td>$lm = lumen$</td>
<td>Radiant Flux (Power)</td>
<td>-</td>
</tr>
<tr>
<td>Luminance ($L_v$)</td>
<td>$cd/m^2 = nits$</td>
<td>Radiance</td>
<td>Area and Solid Angle</td>
</tr>
<tr>
<td>Luminous Exitance ($M_v$)</td>
<td>$lm/m^2$</td>
<td>Radiant Exitance</td>
<td>Area</td>
</tr>
<tr>
<td>Illuminance ($E_v$)</td>
<td>$lm/m^2$</td>
<td>Irradiance</td>
<td>Area</td>
</tr>
<tr>
<td>Luminous Intensity ($I_v$)</td>
<td>$lm/sr = candela$</td>
<td>Radiant Intensity</td>
<td>Solid Angle</td>
</tr>
</tbody>
</table>

**Luma**  Most imaging systems do not measure luminance directly, as they have Optical to Electrical Transfer Functions (OETF) relating the sensor response due to the input luminance to the actual digital codes recorded in the digital processing circuitry. These OETFs are often non-linear in nature as they have been perceptually linearized, and additionally can be S-curved shaped, providing visually pleasing luminance roll-off at the boundaries of the imager sensor bright and dark response. In LDR video, luma generally represents the traditional gamma function based on a power function [Poynton, 2003]. Gamma correction however is limited in its ability
to match the HVS to luminance values beyond approximately $1000 \text{ cd/m}^2$. In HDR, there are a number of ongoing efforts to define luma [Reinhard et al., 2015; Brooks, 2015; Nilsson, 2015; Rafa et al., 2014].

2.1.2 Colourimetry

**Perception**  In the study of the human visual pathway, image contrast plays a more significant role in how neurons represent light information than absolute light levels. This salient characteristic of human vision, called visual adaptation, is our innate ability to process and respond to variations and changes in image contrast, the relative intensities that are observed across the visual field. While individual neurons have sensitivities of only two to three orders of magnitude, neurons in the peripheral are used to signal local contrast in the image, conveying light information across wide variations of absolute levels of light. The range of contrast generally remains constant as the ambient illumination level changes, typically spanning no more than two orders of magnitude [Wandell, 1995].

**Visual Photoreceptors**  The human eye contains two fundamentally different types of photoreceptors, rods and cones. Each eye contains approximately 100 million rods and 5 million cones [Wandell, 1995]. Rods provide vision predominately under low, or scotopic, light levels. Although the large number of rods allow for fine sampling of the retinal image with good sensitivity, the signals from groups of rods converge onto single neurons within the retina, resulting in a loss of fine spatial detail. Cones provide vision predominately under higher, or photopic, light levels. Each cone signal is interpreted by several neurons, resulting in greater spatial sampling. The central fovea of the eye provides the highest visual acuity in the retina, an area with no rods but the highest concentration of cones ($\approx 50K$). Cones are less sensitive to light at most wavelengths than the more sensitive rods. The intensity where both rods and cones contribute to vision is referred to as the mesopic light level. Colour vision is trichromatic in nature as there are three different types of cone photoreceptors within the retina with varying sensitivity to wavelength. The three cone types are the L-cones, M-cones, and S-cones, referring to long, middle, and short peak wavelength sensitivity respectively. The spectral sensitivities for the three cones is shown in figure 2.2.

**Device Independent Colour Matching Functions**  Following the trichromatic nature of colour vision, a test colour stimulus $Q_\lambda$ reaching the eye of an observer can be matched colourimetrically in chromaticity and brightness by additive mixture of
Figure 2.2: Spectral sensitivities of the long (L), medium (M), and short (S) cones [Stockman et al., 1993]

three primary tri-stimuli \( r(\lambda) \), \( g(\lambda) \), and \( b(\lambda) \) along with the scalar multipliers \( R \), \( G \), and \( B \) as:

\[
Q_{\lambda} \triangleq R\hat{r}(\lambda) + G\hat{g}(\lambda) + B\hat{b}(\lambda) \tag{2.7}
\]

In the 1920s John Guild at the NPL and David Wright at Imperial College performed independent investigations of colour matching, yielding similar results [Fairman et al., 1997]. At the CIE meeting of 1931, colour matching functions for a standard observer were defined based on primaries centred at wavelengths of 700 nm, 546.1 nm, and 435.8 nm, known as the CIE 1931 RGB data as shown in figure 2.3.

The ratio of the luminance factors of \( R \), \( G \), and \( B \) at the primary wavelengths were determined from a least-square best fit of the colour matching functions to the 1924 photopic luminous efficiency curve \( V(\lambda) \) and are given as:
One of the formulating principles agreed upon at the 1931 meeting was that the coordinates of all real stimuli were to be positive. The CIE defined then a new $XYZ$ system characterized by primaries $X$, $Y$, and $Z$ comprised of mixtures of the $RGB$ primaries:

$$X \triangleq a_{11}R\hat{r}\bar{a}_{21}G\hat{g}+a_{31}B\hat{b}$$

$$Y \triangleq a_{12}R\hat{r}\bar{a}_{22}G\hat{g}+a_{32}B\hat{b}$$

$$Z \triangleq a_{13}R\hat{r}\bar{a}_{23}G\hat{g}+a_{33}B\hat{b}$$

In terms of a test colour stimulus $Q_\lambda$ the three primary tri-stimuli $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ along with the scalar multipliers $X$, $Y$, and $Z$ can be expressed as:

$$Q_\lambda \triangleq X\bar{x}(\lambda)+Y\bar{y}(\lambda)+Z\bar{z}(\lambda)$$

To determine the scaler multipliers in $XYZ$ space, three points in $RGB$ space must be defined for which both $RGB$ and $XYZ$ tri-stimulus values are known. The projection of $XYZ$ space is made to the Maxwell triangle whose vertices are unit
vectors in space, or the plane $X + Y + Z = 1$. To identify three common points in $RGB$ and $XYZ$ space, a standard CIE chromaticity diagram is defined that is a parallel projection to the $Z$ axis, projecting points to the $XY$ plane. The $XY$ axis are relabelled as $xy$ to signify that they now represent chromaticity as shown in figure 2.4.

![Figure 2.4: Chromaticity plane $xy$ projected in $XYZ$ space [Fairman et al., 1997]](image)

The $XYZ$ tri-stimulus values and the projected $xy$ chromaticity coordinates can be expressed algebraically as:

$$X = x \quad (2.13)$$

$$Y = y \quad (2.14)$$

$$Z = 1 - x - y \quad (2.15)$$

The new CIE colour matching functions include additional desirable properties including scale factors such that the $R, G, B = 1, 1, 1$ maps to the point $X, Y, Z = 1, 1, 1$ and the definition that $X = Y = Z = 1$ for a constant white point ($x = y = \frac{1}{3}$).
Furthermore, $\gamma(\lambda)$ is purposely defined as the primary to carry the luminance function, and is equal to the photopic luminous efficiency curve $V(\lambda)$. The CIE selected $xy$ chromaticity axes projected in $XYZ$ space are shown in figure 2.5.

![CIE 1931 rg (xy) chromaticity diagram](Fairman et al., 1997)

The resulting RGB to CIE 1931 $XYZ$ conversion matrix based on the $xy$ chromaticity axes is given by:

$$
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
= \frac{1}{0.17697}
\begin{bmatrix}
0.49 & 0.31 & 0.20 \\
0.17697 & 0.81240 & 0.01063 \\
0.00 & 0.01 & 0.99
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
$$

$$
(2.16)
$$
where the scale factor results in the luminance function $\gamma(\lambda)$ equivalence to the luminous efficiency curve $V(\lambda)$. The luminance-chromaticity perceptually non-uniform colour space is called the CIE $xyY$ colour space, and is defined as:

$$x = \frac{X}{X + Y + Z} \quad (2.17)$$

$$y = \frac{Y}{X + Y + Z} \quad (2.18)$$

$$Y = \text{luminance of the colour} \quad (2.19)$$

The 1924 CIE luminous efficiency function $V(\lambda)$ dataset used in formulation of the CIE 1931 XYZ primaries was determined later to be too low in the blue part of the spectrum, transferring into the colour matching functions [Schanda, 2015]. Furthermore, the $2^\circ$ field of vision of the 1931 standard, being within the angle subtended by the fovea, was considered inadequate in that it did not consider enough peripheral vision of the observer. Later experiments by Stiles and Burch in the 1950s led to new colour matching functions known as the CIE 1964 standard colourimetric system [Stiles and Burch, 1959]. The 1964 standard additionally widened the observers field of vision to $10^\circ$, where the eye has a slightly different response. A comparison chart between the CIE 1931 and CIE 1964 standards is shown in Figure 8 and is documented in the CIE technical report on colourimetry [CIE, 2004].

As a result of decades of use and widespread adoption, the CIE 1931 standard continues to be applied as a device independent system for colourimetry.

**Perceptually Uniform Colour Space** The definition of luminance was previously given as radiance integrated over the visible spectrum, weighted by the spectral sensitivity function, the luminous efficiency curve $V(\lambda)$ (figure 2.1). While luminance is not a perceptually uniform quantity, human perception to luminance is nonlinear in nature. This nonlinear response is referred to as lightness $L^*$, the perceived value of human vision to relative luminance. Lightness has been modelled by polynomials, power functions, and logarithms [Wyszecki and Stiles, 2000], however in 1976 the CIE standardized the lightness function as a power function with the substitution of a linear portion near black as:

$$L^* = \begin{cases} 
903.3 \frac{Y}{Y_n}; & \frac{Y}{Y_n} \leq 0.008856 \\
116(\frac{Y}{Y_n})^{\frac{1}{3}} - 16; & 0.008856 < \frac{Y}{Y_n} 
\end{cases} \quad (2.20)$$
where $L^*$ has a range of 0 to 100, $Y$ is CIE luminance, and $Y_n$ is the luminance of reference white. The calculation of relative luminance ($\frac{Y}{Y_n}$) is not required if luminance is normalized to unity [Poynton, 2003]. If the luminance value is approximately 18% of $Y_n$ or $\frac{Y}{Y_n} = 0.18$, the correlate for lightness is approximately 50, commonly referred to as middle grey and is the basis for 18% grey cards often used for calibration and reference. An approximation to the lightness curve can also be determined as a power function with an exponent of 0.4.

Applying the concept of perceptual uniformity to colour spaces, the CIE standardized two colour spaces in 1976 having not agreed on a single system, known as CIE $L^*u^*v^*$ and CIE $L^*a^*b^*$. Both systems transform $X$, $Y$, $Z$ tri-stimulus values into lightness of a 0 to 100 range, and to dual colour components of a ±100 approximate range. One unit of Euclidean distance in CIELUV and CIELAB corresponds with approximately one just noticeable difference (JND) in colour [Poynton, 2003].

**CIELUV** The conversion of CIE 1931 tri-stimulus values $X$, $Y$, $Z$, lit by a known illuminant (reference white) $X_n$, $Y_n$, $Z_n$, to CIELUV is given by:

---

Figure 2.6: Colour matching functions of the CIE 1931 standard ($2^\circ$: full lines) and CIE 1964 standard ($10^\circ$: dotted lines) [Schanda, 2015]
\[ L^* = 116 \left( \frac{Y}{Y_n} \right)^{1/3} - 16 \quad \frac{Y}{Y_n} > 0.008856 \]  

(2.21)

\[ u^* = 13L^*(\dot{u} - \dot{u}_n) \]  

(2.22)

\[ v^* = 13L^*(\dot{v} - \dot{v}_n) \]  

(2.23)

Notice that \( L^* \) above represents lightness in terms of the power function. When the linear portion of the lightness function is calculated, \( L^* \) is given as:

\[ L^* = 903.3 \frac{Y}{Y_n} \quad \frac{Y}{Y_n} < 0.008856 \]  

(2.24)

The intermediate quantities are defined as:

\[ \dot{u} = \frac{4X}{X + 15Y + 3Z} \quad \text{and} \quad \dot{u}_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n} \]  

(2.25)

\[ \dot{v} = \frac{9X}{X + 15Y + 3Z} \quad \text{and} \quad \dot{v}_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n} \]  

(2.26)

**CIELAB** For CIELAB the CIE 1931 tri-stimulus values \( X, Y, Z \), lit by a known illuminant (reference white) \( X_n, Y_n, Z_n \), are computed in a similar manner and are given by:

\[
\begin{align*}
L^* &= 116 \left( \frac{Y}{Y_n} \right)^{1/3} - 16 \\
a^* &= 500 \left[ \left( \frac{X}{X_n} \right)^{1/3} - \left( \frac{Y}{Y_n} \right)^{1/3} \right] \quad \frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} > 0.008856 \\
b^* &= 200 \left[ \left( \frac{Y}{Y_n} \right)^{1/3} - \left( \frac{Z}{Z_n} \right)^{1/3} \right]
\end{align*}
\]  

(2.27)

For the linear portion of the transfer function, the calculations are given by:

\[
\begin{align*}
L_m^* &= 903.3 \frac{Y}{Y_n} \\
a_m^* &= \left[ f\left( \frac{X}{X_n} \right) - f\left( \frac{Y}{Y_n} \right) \right] \quad \frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} \leq 0.008856 \\
b_m^* &= \left[ f\left( \frac{Y}{Y_n} \right) - f\left( \frac{Z}{Z_n} \right) \right]
\end{align*}
\]  

(2.28)

where the function \( f(r) \), with \( r \) representing a \( X, Y, \) or \( Z \) ratio, is given by:
\[ f(r) = \begin{cases} \frac{r}{r}^{\frac{1}{2}}, & \text{for } r > 0.008856 \\ 7.787r + \frac{16}{116}, & \text{for } r \leq 0.008856 \end{cases} \quad (2.29) \]

2.1.3 Dynamic Range and Signal to Noise Ratio (SNR)

DR is defined as the ratio of luminance of the lightest and darkest elements of a scene or image. In an absolute sense, this ratio can be considered the brightest to darkest pixel, but more commonly it is dened as the ratio of the largest non-saturating signal to one standard deviation above the camera noise under dark conditions [Holst and Lomheim, 2007], [ISO, 2013]. An equation for \( DR_{SNR} \) can be described as the ratio of maximum to minimum input luminance level:

\[ DR_{SNR} = \frac{L_{sat}}{L_{min}}, \quad (2.30) \]

where \( L_{sat} \) is the maximum unclipped input luminance level defined as 98% of its maximum value (250 of 255 for 24 bit colour), and \( L_{min} \) is the minimum input luminance level for a specified SNR. The ratio can be specified as instantaneous for a single instant of time, or temporal, where the luminance values change over time, i.e. light and dark elements over a series of images. Poynton [Poynton, 2003] uses the definitions of simultaneous for the instantaneous case, and sequential for the temporal case.

In an attempt to provide a definition of terms to describe dynamic range, the standards body MPEG stated in their recent “Requirements and Use Cases for HDR and WCG Content Distribution” document [Luthra et al., 2014] (following input from COST Action IC1005 in October 2013) as: “Overall, the dynamic range of a scene can be described as the ratio of the maximum light intensity to the minimum light intensity” [ITU-R, 2012b]. In digital cameras, the most commonly used unit for measuring dynamic range is in terms of f-stop, which describes total light range by powers of 2. The current ad hoc use of the term f-stop, refers to the following dynamic ranges:

- 10 f-stops = a difference of \( 2^{10} = 1024:1 \) contrast ratio.
- 14 f-stops = a difference of \( 2^{14} = 16,384:1 \) contrast ratio.
- 16 f-stops = a difference of \( 2^{16} = 65,536:1 \) contrast ratio.
- 20 f-stops = a difference of \( 2^{20} = 1,048,576:1 \) contrast ratio.
It is indeed possible with HDR technology to capture more than the human eye can see at any level of adaption, a range of more than 20 f-stops. New classes of imaging sensors are being developed that are capable of directly capturing larger dynamic range on the scale of 5 orders of magnitude (5 log_{10}, 16.6 stops) [Ma et al., 2015] [Cade, 2013]. For comparison, human vision generally is agreed upon to have a simultaneous dynamic range capability of approximately four orders of magnitude (4 log_{10}, 13.3 stops) [Banterle et al., 2011] [Kunkel and Reinhard, 2010], and a sequential dynamic range via adaption of approximately ten orders of magnitude (10 log_{10}, 33.2 stops) [Ferwerda, 2001] [Wandell, 1995].

In order to objectively calculate the capture system’s DR, an understanding of the imaging and noise transfer function is required. An image sensor measures the radiant power (flux) $\Phi$ at each pixel $j$ for each image $i$. The amount of photons $n_{pj}$ that are absorbed at each pixel area $A_j$ over time period $t_i$ is given by:

$$n_{pj} = \frac{A_j u_j t_i}{E_p},$$  \hspace{1cm} (2.31)

where $j$ is the mean incident spectral irradiance at pixel $j$. $E_p$ is the photon energy defined by:

$$E_p = \frac{hc}{\lambda},$$ (2.32)

with the wavelength of the photon given by $\lambda$, the speed of light $c = 2.9979 \cdot 10^8$ m/s, and Planck’s constant $h = 6.626 \cdot 10^{-34}$ Js. Photons incident on the pixel over the exposure time generate an average number of electrons $n_e$ based on the quantum efficiency of the sensor $\eta(\lambda)$:

$$\eta(\lambda) = \frac{n_p}{n_e},$$ (2.33)

producing a photocurrent $i_{ph}$ and photon noise $n_p$ at each pixel site. Thermally dependent leakage current in the electronics (free electrons in the silicon) known as dark current $i_d$ contributes to the total charge as shown in figure 2.7. In conventional current mode, the photocurrent produces a charge $Q_j$ on a capacitive integrator generating a voltage $V_j$ amplified with gain $g$. The voltage gain product $V_j g$ includes readout noise $n_r$ and amplifier noise $n_a$. Typically, a buffer circuit follows the gain amplifier contributing an additional buffer amplifier noise $n_b$, leading to an analogue to digital converter (ADC) with quantization noise $n_q$, resulting in the measured pixel Digital Number (DN) $d_i j$. In linear systems, the DN is proportional to the number of electrons collected with an overall system gain $G$ of units of $DN/electrons$.

Given a linear transfer model, the noise sources can be summed and we can
write an equation for the total system noise at the output of the ADC as:

\[ n_{\text{total}}(j) = n_p(j) + n_d(j) + n_r(j) + n_a(j) + n_b(j) + n_q(j). \]  
(2.34)

When specifying noise units, it is understood that the noise magnitude is the RMS value of the random process producing the noise [Holst and Lomheim, 2007]. If \( n_{\text{postgain}} \) is defined as the sum of the post-gain amplifier noise sources \( n_r, n_a, n_b, \) and \( n_q \), the variance is given as:

\[ \sigma^2_{n_{\text{total}}} = \sigma^2_{n_p} + \sigma^2_{n_d} + \sigma^2_{n_{\text{postgain}}}. \]  
(2.35)

The variance represents the random noise power of the system, and the standard deviation characterizes the system RMS noise [Reibel et al., 2003]. Finally, Reibel et al [Reibel et al., 2003] noted that no system is strictly linear, and introduce a non-linear contribution \( C_{NL} \) to describe the signal variance that is modeled by a dependence on the total electrons transferred to the sense node capacitor (photons-electrons and dark current electrons) and the overall system gain \( G \) resulting in:

\[ \sigma^2_{n_{\text{total}}} = \sigma^2_{n_p} + \sigma^2_{n_d} + \sigma^2_{n_{\text{postgain}}} + C_{NL}. \]  
(2.36)

A linear transfer model not only permits the summation of noise sources, it also ensures the relative brightness of the scene is maintained. To achieve linearity over the entire sensor, and to minimize fixed pattern noise (including streaking and banding), a radiometric camera calibration is generally required for each pixel resulting in a total camera response function. The calibrated camera response function can be combined with optical characterization and other camera post-processing functions such as gamma, tonal response, noise reduction, and sharpening to produce an overall camera opto-electronic conversion function (OECF) [Fiete, 2010] [Healey and Kondepudy, 1994].

Noise is meaningful in relation to a signal (i.e. Signal to Noise Ratio, SNR)
and can be defined differently depending on how the measurement is performed. If the measurement is concerned with the sensor itself, isolated from the camera post-processing functions described above, noise can be referenced to the original scene. This can be accomplished by linearizing the data using the inverse OECF, and defining the signal as a pixel difference corresponding to a specified scene density range. Alternatively, if raw data is unavailable or the OECF is unknown (or beyond the scope or measurement capability), or if the focus is on complete system capability (optics, sensor, and camera performance) rather than sensor specific, the signal can be defined with respect to an individual patch pixel level. The evaluation of dynamic range is based on the latter, with a goal of describing a straightforward method that can be used to produce an ever increasing database of comparable camera system dynamic range measurements. If the variance $\sigma_{signal}^2$ of the signal is known and signal is zero mean, the $SNR$ for the minimum input luminance can be expressed as:

$$SNR_{L_{min}} = \frac{\sigma_{signal}^2}{\sigma_{n_{total}}^2}$$

(2.37)

where $\sigma_{n_{total}}^2$ is given by equation (2.35). Alternatively, equation (2.37) can be expressed as the ratio of the mean pixel value $\mu_{mean}$ to the standard deviation $\sigma_{StdDev}$, and equivalently to the total RMS noise $\eta_{totalRMS}(j)$ as:

$$SNR_{L_{min}} = \frac{\mu_{mean}}{\sigma_{StdDev}} = \frac{\mu_{mean}(j)}{\eta_{totalRMS}(j)}.$$  

(2.38)

Minimum input luminance level is often determined from the density step having a SNR of 1 or greater [ISO, 2013] [EMVA, 2010]. The processing software utilized in this study is Imatest [Koren, 2014], which performs automatic calculations for $SNR_{L_{min}}$ values of 1, 2, 4, and 10. It can be convenient to re-evaluate equation (2.38) in terms of stops. Stops is a common photography term where one stop is equivalent to a halving or doubling of light, corresponding more closely to the relative luminance response of human vision [Mantiuk et al., 2007]. We can express $SNR_{L_{min}}$ in terms of noise in stops by:

$$n_{stops_{L_{min}}} = \frac{1}{SNR_{L_{min}}}.$$  

(2.39)

where a $SNR_{L_{min}}$ of 1 is equivalent to 1 stop of noise. Figure 2.8 illustrates the assignment of the arbitrary quality descriptions (high, medium-high, medium, and low) for 0.1, 0.25, 0.5 and 1 stop of noise as calculated and defined in Imatest, as well as the corresponding $SNR$ values in terms of a luminance ratio as well as in
decibels (dB). Finally, $DR_{SNR}$ can also be expressed in terms of relative luminance or stops as:

$$DR_{stops} = \frac{\log_{10}L_{sat} - \log_{10}L_{min,SNR}}{\log_{10}2}.$$  (2.40)

The question that remains is: How are the arbitrary quality descriptors chosen and why? We can start to answer this question by reviewing early 1940s era work in human signal detection by Albert Rose, known as the Rose model. The Rose model provided a good approximation of a Bayesian ideal observer, albeit for carefully and narrowly defined conditions [Rose, 1948]. Rose defined a constant $k$, to be determined experimentally, as the threshold $SNR$ for reliable detection of a signal. The value of $k$ was experimentally estimated by Rose to be in the region of 3-7 based on observations of photographic film and television pictures. Additional experiments using a light spot scanner arrangement resulted in an estimate of $k$ of approximately 5, known as the Rose criterion, stating that a $SNR$ of at least 5 (13.98 dB) is required to distinguish features at 100% certainty. The early empirical determinations of $k$ had several issues. Dependency on a specification of percent chance of detection (i.e. 90%, 50%), included limitations of human vision such as spatial integration and contrast sensitivity, and included dependence on viewing distance amongst other factors. The development of Signal Detection Theory (SDT) and objective experimental techniques worked to overcome these limitations [Green and Swets, 1966] [Nevin, 1969]. For comparison, a binary (yes/no) Signal Known Exactly (SKE) detection experiment, utilizing 50% probability of signal detection and an ideal observer false alarm rate of 1%, resulted in $SNR$ values of 2.3 (7.23 dB) and 3.6 (11.13) for 50% and 90% true-positive rates respectively [Burgess, 1999].

More recently, the ISO 12232-2006 standard defines two “noise based speeds” for low noise exposures that is based on objective correlation to subjective judgements of the acceptability of various noise levels in exposure series images [ISO, 2006]. The two noise based speeds are defined as $S_{noise_{40}} = 40$ (32.04 dB) providing “excellent” quality images, and $S_{noise_{10}} = 10$ (20 dB) providing “acceptable” quality images. Examining figure 2.8, a commonality point between the ISO 12232-2006 standard and Imatest occurs at the 20 dB point, pertaining to “acceptable” quality in the ISO standard and “high” quality in Imatest. We note that the ISO standard is based on prints at approximately 70 pixels/cm, at a standard viewing distance of 25 cm. For subjective models utilizing sensor pixels per centimetre values $P$ differing from the ISO standard, an $SNR$ scaling ratio of $\frac{70}{P}$ is recommended.

Previously reported results for objective calibrated dynamic range test data
have yielded limited results and the test methodology used is often unclear. The Cinema5D test website [Wober, 2014] notes “We have at one point in 2014 updated our dynamic range evaluation scale to better represent usable dynamic range among all tested cameras. This does not affect the relation of usable dynamic range between cameras”. However the author does not clearly state what evaluation scale is currently, or was previously, used. In another example, published on the DVInfo website [Wilt, 2014], the author describes the minimum input luminance level as “typically, where you consider the noise amplitude to be significant compared to the amplitude of the darkest stop visible above the noise”, but does not further define the term “significant”. Results published on the DxOMark website [DxOMark, 2015] primarily include still camera and mobile phone tests, implementing a self-designed dynamic range test chart that is limited to 4 density steps (13.3 stops).

2.1.4 Data Formats

One of the most fundamental yet important decisions in creating an HDR workflow is the selection of the data format(s). A particular data format can have a ripple effect throughout the pipeline, effecting choice in hardware, processing power, storage, communication protocol, and of course, the ability to maintain physical accuracy. Data formats primarily are either integer or float types, and may be stored either
as linear representations or with logarithmic encoding.

**Integer versus Floating Point**  Integer based data storage most commonly consists of 8 bits per the red, green, and blue representation of each pixel, resulting in a total 24 bits per pixel. As hardware and software capability evolves, 10, 12 and even 16 bit integer workflows may be a possibility, resulting in 30, 36, and 48 total bits per pixel respectively.

Floating point numbers provide an alternative data type, available in several typical flavours:

- **Half**: 2 bytes, 16 bits × 3 colour channels = 48 bits
- **Float (Single)**: 4 bytes, 32 bits × 3 colour channels = 96 bits
- **Double**: 8 bytes, 64 bits × 3 colour channels = 192 bits
- **Long Double (Quad)**: 10, 12, or 16 bytes, 80, 96 or 128 bits × 3 colour channels = 240, 288, or 384 bits respectively

The minimum number of significant digits, used to represent the exponent and mantissa, as well as minimum and maximum exponent range, is dependent upon the system implemented and can often be found in the float header file or documentation. Typical values include:

- **Half**: Minimum 3 significant digits, exponent 5 bits, mantissa 11 bits (10+sign bit)
- **Float**: Minimum 6 significant digits, exponent 8 bits, mantissa 24 bits (23+sign bit)
- **Double**: Minimum 15 significant digits, exponent 11 bits, mantissa 53 bits (52+sign bit)
- **Long Double**: Minimum 18 significant digits, exponent 15 bits, mantissa 64 bits (63+sign bit)

On the surface, it would appear that floating point data has the capability for achieving greater dynamic range and more accurate pixel data, at the cost of increased size and processing requirements. Before considering the trade-offs of implementing one data type over the other however, one must first consider the type of encoding to be implemented, either linear or non-linear.
Linear versus Non-Linear Encoding  We begin the discussion by describing linear encoding using the integer storage type. Using an 8 bit integer as an example, coding values range from 0 to 255 resulting in a theoretical contrast ratio of 255:1 or 8 stops of dynamic range, but an issue arises with respect to human perception. The issue at hand, as described by Poynton, is the “code 100” problem [Poynton, 2003]. For pixel values below code 100, luminance difference between codes become increasingly perceivable, resulting in artefacts such as contouring or banding. For pixel values above 100, codes eventually fall beyond the 1% luminance difference threshold of the HVS (also referred to as contrast sensitivity) resulting in unnecessary codes. Increasing the number of bits anywhere from 8 to 16 bits increases the overall contrast ratio, but does not solve the “code 100” problem below 100, and results in further wasted codes above 100. To put this in perspective, to cover 8 stops of dynamic range (a contrast ratio of 255:1) with 1% luminance difference, the number of linearly encoded values required is:

\[
\frac{255}{0.01} = 25,500 \text{ codes or 15 linearly encoded bits.} \tag{2.41}
\]

As an alternative, implementing the near logarithmic response of the HVS results in codes spaced at the ratio of 1.01, the 1% perceivable luminance step. Implementing the nonlinear 1.01 ratio, the number of codes required is:

\[
\frac{\log(255)}{\log(1.01)} = 557 \text{ codes or 10 non-linearly encoded bits.} \tag{2.42}
\]

As one can see, the more efficient method is to pick code values based on a non-linear scale that is perceptually uniform, so as to (as a minimum) satisfy the 1% contrast sensitively characteristic of the HVS. In order to better understand the various integer and floating point data types available, as well as the trade-offs of linear versus non-linear encoding, several popular HDR data choices are presented. We begin with formats stored as RGB triplets, followed by formats using log luminance and chromaticity coordinate colour space.

sRGB: Non-Linear Integer  sRGB, a popular “LDR” 8 bit format, is actually a display-referenced colour space that includes a primarily non-linear luminance correction similar, but not exact, to CRT “gamma” (and the Rec-709 Standard). The format is referred to as display-referenced, where the primary goal is to mimic the output capabilities of the display device, not necessarily to capture a complete radiance range such as in an HDR type scene. For the luminance correction, a linear transfer function is implemented for normalized [0, 1] RGB values below 3.1308 ×
For greater values, a non-linear transformation is used. The transformation from relative scene tri-stimulus values $L$ to sRGB coded values $\hat{V}$ is summarized as:

$$
\hat{V}_{sRGB} = \begin{cases} 
12.92L & \text{for } L < 0.0031308 \\
1.055L^{\frac{1}{2.4}} - 0.055 & \text{for } L \geq 0.0031308
\end{cases}
$$

and from coded values back to tri-stimulus values:

$$
L = \begin{cases} 
\frac{1}{12.92}\hat{V}_{sRGB} & \text{for } 0 \leq \hat{V}_{sRGB} < 0.03928 \\
\left(\frac{\hat{V}_{sRGB}+0.055}{1.055}\right)^{2.4} & \text{for } 0.03928 < \hat{V}_{sRGB} \leq 1.0
\end{cases}
$$

sRGB is an 8 bit integer format, where coded values are converted into full 8 bit form from 0 to 255. No headroom or footroom is applied, as may be the case in other 8 bit display-referenced formats (Rec709). sRGB relative luminance as a function of code value is shown in figure 2.9.

Figure 2.9: sRGB Encoding Luminance vs. Code Value

**Cineon and Pixar: Log Integer** The Cineon [Kennel and Snider, 1995] and Pixar formats implement 10 and 11 bits respectively per RGB channel for a total
of 30 and 33 bits per pixel. Each channel is a logarithmically encoded integer value with the intent of covering the typical full dynamic range of film stock. Cineon, along with the related format DPX, were developed for use in the film scanning process to capture the 2.048 density range of film in 0.002 steps. The Pixar format covers 3.8 orders of magnitude in 0.4% steps. It is a good example of the use of higher integer bit depths with non-linear encoding to increase the dynamic range capability beyond sRGB. The limitation of dynamic range to a maximum 3.8 orders of magnitude however (Pixar), is insufficient for HDR applications, so it is not covered in detail here.

**Radiance HDR: Floating Point**  
Radiance HDR is the storage format employed in the Radiance rendering program [Larson and Shakespeare, 1998]. There are two options, a 4 byte RGBE format and a XYZE CIE variant. The XYZE variant has associated quantization errors however, whose choice is not recommended [Ward, 2001]. In the RGBE format, RGB floating point values are represented by three 8 bit RGB mantissas, utilizing a common 8 bit exponent, for a total of 4 bytes. In most floating point formats, the mantissa is normalized to values between 0.1 and 1 [Ward, 1991]. RGBE uses the same exponent for the RGB mantissas, and therefore only the largest value may be normalized. The largest primary value is favoured, but the same largest value also dominates the displayed pixel colour.

**RGBE** provides over 77 orders of magnitude of dynamic range from $10^{-38}$ to $10^{38}$, allowing HDR image processing without losing significant accuracy. RGBE is however limited to positive values and therefore cannot represent all visible chromaticities. Overall, the RGBE format includes a large total dynamic range at the expense of accuracy, which is approximately 1% perceptual steps.

**TIFF Floating Point Formats**

**IEEE Float TIFF**  
The IEEE format, or Float TIFF, is a 32 bit per component floating point representation for a total of 96 bits with RGB colour space. This is a storage intensive format that is primarily used in lossless intermediate compositing applications or when accurate reference image representation is required for evaluation and comparison.

Each 4 byte component consists of a sign bit, 8 exponent bits, and 23 mantissa bits, noting that the last 12 bits or more will contain random noise as there are not image sources that output this type of accuracy [Reinhard et al., 2010]. The range of values extends from the smallest value of $1.17549E^{-38}$ to $3.40282E^{+38}$ as shown
in figure 2.10; a dynamic range of $2.89480E^{+76}$. The relative error is small at 0.00003%. The floating point values can be expressed by:

$$\text{floating point value} = (-1^{\text{sign}}) \times (1 + m) \times 2^{\text{exp}-127}$$ \hspace{1cm} (2.45)$$

where $$\text{sign}$$ is the sign bit, $$m$$ is the decimal mantissa, and $$\text{exp}$$ is the decimal exponent value between 1 and 254. Although the total range is similar to $RGE$, IEEE $RGB$ has smaller relative error at the expense of large data sets.

**Half Float TIFF**  A half float encoded version is available in the TIFF library consisting of 16 bits per component for a total of 48 bits per pixel with $RGB$ colour space. Half float with linear encoding suffers reduced dynamic range of 5.4 orders of magnitude when limiting the perceivable luminance steps to less than 1%. Additionally, half float represents only positive primary values and therefore does not cover the full colour gamut. Perhaps a more widely used and supported implementation of the half float is the OpenEXR format.

**OpenEXR: Floating Point**  OpenEXR (Extended Range) was developed by Industrial Light and Magic (ILM) and is similar to the half float format in that each
RGB pixel utilizes a total of 48 bits [Bogart et al., 2004]. Each channel consists of 5 exponent bits, 11 mantissa bits, and a sign bit, the same as the half float. The floating point values can be expressed as:

\[
\text{floating point value} = \begin{cases} 
(-1)^{\text{sign}} 2^{\text{exp}-15}(1 + \frac{m}{1024}) & \text{for } 1 \leq \text{exp} < 30 \\
(-1)^{\text{sign}} 2^{-14}(\frac{m}{1024}) & \text{for } \text{exp} = 30 
\end{cases}
\]  

(2.46)

where \(\text{sign}\) is the sign bit, \(m\) is the decimal mantissa, and \(\text{exp}\) is the decimal exponent value between 1 and 30. The range of values extends from the smallest value of \(6.10352E^{-5}\) to \(6.55040E^{+4}\) as shown in figure 2.11, a dynamic range of \(10.7322E^{+10}\). The relative error is small at 0.1%, not nearly as small as in the case of the IEEE Float, but an order of magnitude better than Radiance RGBE.

![OpenEXR Half Float Value vs Encoded Value](image)

Figure 2.11: OpenEXR Half Float vs Encoded Value

OpenEXR has several “usability” features that make it an attractive option. The format matches NVidias 16 bit half format and is directly supported. Standard C++ classes in the OpenEXR programming library provide a general interface for reading and writing of files, a specialized interface for RGBA, and an interface for programs written in C. Files may be stored as uncompressed for fast read and write
cycles when paired with a fast file system such as a RAID configuration, or in a compressed format for reduced file sizes. There is support for three different lossless data compression methods (PIZ, ZIP, and RLE), offering trade-offs in speed versus compression ratio. PIZ compression is wavelet based giving the best compression ratio for photographic images. ZIP compression is based on differences in horizontal pixels with faster decompression than PIZ, but slower compression. RLE compression uses run-length encoding on differences between horizontal pixels, a generally fast method albeit with larger compressed files sizes.

**LogLuv TIFF** The LogLuv format is an encoding of RGB images into log luminance values and CIE \((u, v)\) chromaticity coordinates. A driving factor in its development is to store HDR data using a reduced amount of storage space as opposed to formats such as the full float IEEE RGB. LogLuv is stated to provide the following benefits [Larson, 1998]:

- Covers the full visible colour gamut
- Covers the full range of perceivable luminance
- Uses imperceptible step sizes in a perceptually uniform space
- May be calibrated to absolute luminance and colour
- Enables optimal visual fidelity on any output device.

Additional stated goals of the format are that it be compact, compressible, easily converted to \(XYZ\) and \(RGB\) formats, and that it be incorporated into the TIFF standard [Larson, 1998]. There are two versions of the LogLuv encoding, a 24 bit and 32 bit option as follows:

- 24 bit: 10 bit log luminance + 14 bit indexed uv mapping
- 32 bit: 16 bit log luminance + 8 bits each for \(u\) and \(v\)

At the expense of one extra byte per pixel, the 32 bit version provides greater dynamic range, precision, and compression capability. The 32 bit implementation is considered the standard, and is what will be described herein. The Most Significant Bit (MSB) is used as a sign bit, leaving 15 bits for encoding of real luminance \(Y\) to log luminance \(L_c\) as follows:

\[
L_c = 256(\log_2 Y + 64),
\]  

(2.47)
and from log luminance back to real luminance as:

\[ Y = \exp_2 \left[ \frac{L_e + 0.5}{256} - 64 \right]. \tag{2.48} \]

Log luminance as a function of code value is shown in figure 2.12. The TIFF library is readily available at no cost, and is widely utilized and tested, making it a strong potential choice as an HDR format.

![Figure 2.12: HDR LogLuv36 TIFF Encoding Luminance vs. Code Value](image)

**Lossy Formats**  In addition to the lossless file formats discussed, there are several lossy formats available, or under development, including JPEG HDR-XT, HDR-JPEG2000, MPEG-HDR, and H.264-HDR [Banterle et al., 2011]. Lossy formats provide practicality in file size, processing, and distribution, while including several benefits such as backward compatibility and incorporation of tone mapping operators to reduce dynamic range per display capability. The trade-off is generally a loss of physical accuracy, and the ability to maintain and display real-world scenes, a focus of this thesis. For this reason, the selection of file formats for use in the evaluations will be of the lossless types.
**File Format Discussion**  From the file formats discussed above, the format with the highest quality standard for comparison of data sets is the 96 bit IEEE Float TIFF. IEEE float would be selected for the research concern where file size or processing limitations is secondary to overall accuracy. Alternatively, if the benefits of industry standardization and the implementation of lossless wavelet compression is of value, OpenEXR is a potential choice. OpenEXR has a small relative error of 0.1% while supporting 10 orders of magnitude (although the maximum value is limited to 65,504). Another more commonly implemented format is LogLuv TIFF, with relative error of 0.3% while supporting 38 orders of magnitude. LogLuv has the benefit of supporting the full range of perceivable luminance. For our application, the Radiance HDR float point format provided the best fit for two reasons. First, there are existing compatible Matlab image processing functions available for the import and export of Radiance HDR data, as well as other standard processing functions. Second, we intend to compare data produced by our methodology with current HDR processing techniques, such as via the software program Photosphere. Using the same data format simplifies comparisons. Photosphere combines multiple exposures into an HDR image using the Radiance HDR format.

A summary of dynamic range representations, as well as the number of sampled bits required to meet 1% linear, 1% log, and 0.27% LogLuv relative spacing is provided in Appendix A.

## 2.2 Related Work

### 2.2.1 Analysis of Capture Processing Pipelines

**LDR** Historically, in the commercial and entertainment markets, the LDR processing pipeline has in general been implemented to meet specifications such as the ITU Rec. BT.709, using colour spaces such as sRGB to model CRT “gamma” as described in section 2.1.4. Poynton [2003] provides an excellent overview of gamma and the standard definition processing chain in Chapter 23 of his text, including the nonlinear relationship between the physical quantity luminance and the perceptual quantity of lightness. As HDR applications develop, new processing pipelines are being formulated from both standards committees as well as industry.

**MPEG** The Moving Picture Experts Group (MPEG) has put forth a best practices document for conversion and coding practices as part of the ISO/IEC 23008 MPEG-H standard titled “High Efficiency Coding and Media Delivery in Heterogeneous Environments” [Rec, 2013]. A sub-document, ISO/IEC TR 23008-14, “Con
version and coding practices for high-dynamic-range and wide-colour-gamut video” includes a description of processing steps for converting linear light, RGB, 4:4:4 video into ST 2084, Y’CbCr, 4:2:0 video before encoding [Samuelsson et al., 2017]. It is stated in the document, however, that the content preparation step is out of scope; further stating that although content preparation may include filtering and image enhancement processing such as denoising, colour correction, sharpening filtering etc., they are not specifically described. The sub-document assumes that the input to the HDR System is linear light, RGB, 4:4:4 video.

Mantiuk Mantiuk et al. [2007] discuss HDR image acquisition as part of a paper on the HDR imaging pipeline and its relationship to perception motivated representation of visual content. They note that image processing associated with each stage of the pipeline usually aims at delivering the best looking image, such as adding colour and sharpness for effect, while operating on vaguely defined data that is difficult to standardize. It is stated that one reason that actual imaging pipelines do an acceptable job is a result of the human eye having a huge tolerance for the imperfections of images as compared to the real scenes. It is concluded that the goal of most imaging systems is to reproduce the appearance of a scene, but not to deliver the same linear light physical stimuli to our eyes. Mantiuk et al. note that for HDR image acquisition from multiple exposures, the values of luminance recovered are not meant to be exact measurements, since they can be severely distorted by camera optics, and are only approximately proportional to the luminance of a real scene via a camera response function. Further, typical cameras are not linear in most cases, and only with a measured camera’s response function defining the relation between pixel values and the intensity of light reaching the camera’s sensor, can the data be considered approximately linear. Mantiuk et al. references Reinhard et al. for deriving of the camera response function in order to approximately linearize the data [Reinhard et al., 2010]. Generally, camera response functions are not provided by manufacturers as they are often considered proprietary. Reinhard et al. describe techniques for the deriving the camera response function from a series of aligned exposures, including those fromDebevec and Malik [2008] and Mitsunaga and Nayar [1999]. An alternative method is described from Grossberg and Nayar [2003, 2004] that avoids pixel correspondence and works instead on image histograms. Grossberg makes the assumption that the response curve is the same for each point in the image.
Debevec and Malik  Debevec and Malik extend earlier pioneering work by Mann and Picard [1995] for deriving the camera response function from a series of aligned exposures. Different image pixels, taken at different exposures, provide different portions of the response curve. Using linear optimization, a smooth curve then is determined from the individual portions by minimizing mean square error. When determining the optimal values, a weighting hat function $w(Z_{ij})$ is utilized for image position $i$ at exposure $j$ as:

$$w(z) = \begin{cases} 
    z - Z_{\text{min}}, & \text{for } z \leq \frac{1}{2}(Z_{\text{min}} + Z_{\text{max}}) \\
    Z_{\text{max}} - z, & \text{for } z > \frac{1}{2}(Z_{\text{min}} + Z_{\text{max}}) 
\end{cases} \quad (2.49)$$

Mitsunaga and Nayar  Mitsunaga and Nayar use an $N$ dimensional polynomial approximation to the response function, while additionally determining the exact exposure ratios. The exposure ratios are important if the exact aperture and shutter speeds are not known. The response function is defined by $N + 1$ coefficients of the polynomial, or $[c_0, c_1, ... c_N]$, as given by:

$$f(M) = \sum_{n=0}^{N} c_n M^n \quad (2.50)$$

The coefficients are determined by minimizing an error function for a given candidate exposure ratio. In selecting the polynomial degree $N$, the authors recommend solving for every degree polynomial up to a maximum exponent, selecting the solution with the smallest error.

Grossberg  Grossberg uses intensity mapping between images, determined solely from the image histograms. Conducting intensity mapping in this fashion avoids the issue of pixel correspondence, and the need for registration, for scene radiances that remain almost constant between images. An important benefit is that some motion of both camera and scene objects is allowable, as scene motion will not change the histogram of scene radiances significantly.

Akyuz  Akyuz and Gentav [2013] provides an evaluation study of the various camera response algorithms, finding that Grossberg and Nayar’s algorithm is most accurate, Mitsunaga and Nayar’s algorithm is the most consistent, and Debevec and Malik’s algorithm is the most resistant to noise. It was also found that, statistically, none of the algorithms outperformed the others in terms of accuracy. Like Grossberg, Akyuz makes the assumption that the optical and sensor elements are typically
linear, with Akyuz indicating nonlinearity is introduced during analogue to digital conversion and remapping. With regard to the accuracy metric, Akyuz states two reasons for not directly comparing the recovered camera responses with the actual responses of the camera. First, it is stated that recovering the actual response of a camera is an immensely challenging task due to the effects of optical glare. Second, an observation is made that one does not strictly need the ground truth curve to measure accuracy if the camera response is the aggregate effect of nonlinearities in an imaging system. This assumes, however, that the camera response curve is the same for all lighting conditions and that principle of reciprocity holds; meaning that measured pixels are proportional to the product of the sensor irradiance and the image exposure.

Hasinoff  Hasinoff et al. [2016] present an HDR imaging pipeline that captures, aligns, and merges a burst of frames, taken with constant exposures, as opposed to bracketed exposures. The method operates on Bayer raw frames rather than demosaiced RGB frames, providing more bits per pixel. The pipeline assumes camera raw data is linear. By aligning and merging multiple frames at the same exposure per frame, an image is produced which is reported as having higher bit depth, higher dynamic range, and reduced noise as compared to the input frames. The method has a clearly stated goal of producing natural-looking scenes, and that HDR tone-mapping is employed such as boosting of the shadow regions, brightening corners of the images to compensate for lens vignetting, dynamic range compression, dehazing, global tone adjustment, chromatic aberration correction, sharpening, hue specific colour adjustments, as well as dithering.

Ward  Perhaps the most useable pipeline for the processing and calibration of HDR datasets is via the freely available software program Photosphere, developed by Ward [2010]. Photosphere allows for the input of multiple exposure bracketed images, determines the camera response function as part of the HDR merge and image creation process, and allows for the selection and input of an absolute luminance reference measurement. As this is one of the more popular and easier methods to perform, it is the method we shall use to compare with our new calibrated method.

2.2.2 Gaps in Capture Processing

The image processing pipelines presented in the previous section are summarized in Table 2.2. The methods all share a common gap in that linearity is either assumed, approximated, or of less importance than perceived realism. In the event that
absolute calibration is conducted, without varied spatial and brightness reference measurements throughout the full scene, only limited radiometric accuracy can be determined. Therefore, in all previous pipelines, the absolute dynamic range of the scene is never truly determined, and identification of limitations of specific aspects of the capture system are not identified.

Table 2.2: Capture and Processing Pipelines

<table>
<thead>
<tr>
<th>Source</th>
<th>Dynamic Range</th>
<th>Linearize Data</th>
<th>Absolute Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poyton</td>
<td>SDR</td>
<td>No</td>
<td>Not discussed</td>
</tr>
<tr>
<td>ISO/IEC TR 23008-14</td>
<td>HDR</td>
<td>Assumed</td>
<td>Not discussed</td>
</tr>
<tr>
<td>Mantiuk et al.</td>
<td>HDR</td>
<td>Approximate using CRF</td>
<td>Reference Measurement</td>
</tr>
<tr>
<td>Akyuz et al.</td>
<td>HDR</td>
<td>Approximate using CRF</td>
<td>Not discussed</td>
</tr>
<tr>
<td>Hasinoff et al.</td>
<td>HDR</td>
<td>No Tone-mapped</td>
<td>Not discussed</td>
</tr>
<tr>
<td>Ward (Photosphere)</td>
<td>HDR</td>
<td>Approximate using CRF</td>
<td>Reference Measurement</td>
</tr>
</tbody>
</table>

2.2.3 Multiple Exposure HDR Techniques

In this section we introduce self-calibrating procedures that have been developed for the purpose of combining multiple exposures into HDR radiance maps [Mitsunaga and Nayar, 1999; Debevec and Malik, 2008; Inanici and Galvin, 2004; Coutelier and Oumortier, 2003], as well as the concept of weighted averages. Generally, the film or digital camera response curve is first recovered, and the HDR radiance map is then determined from a weighted average of scaled pixel values from multiple exposures. One of the earliest proposed weighted averages by Mann and Picard [Mann and Picard, 1995] fitted the exposure data to a power curve for use as the weighting function. Debevec and Malik [Debevec and Malik, 2008] introduced a self-calibrating procedure utilizing a physical property of imaging systems called reciprocity. Reciprocity specifies that in the ideal case, the exposure $X$ is a product only of the irradiance $E$ and the exposure time $\Delta T$. Their hat based function assigns higher
weights to middle exposure values [Debevec and Malik, 2008]. Additional methods for implementing weighting functions have since been developed by Robertson et al. [Robertson et al., 1999] (similar to Gaussian curve), Mitsunaga et al. [Mitsunaga and Nayar, 1999] (first order approx. to signal to noise ratio), and Ward et al. [Reinhard et al., 2010] and Akyüz et al. [Akyüz and Reinhard, 2007] (modification of Mitsunaga implementing broad-hat model). Noise based models for weighting functions have been proposed by Tsin et al. [Tsin et al., 2001] (statistical characterization), Kirk et al. [Kirk and Andersen, 2006], Granados et al. [Granados et al., 2010] and Hasinoff et al. [Hasinoff et al., 2010]. Noise based models consider a more in-depth analysis of the acquisition process, including both temporal and spatial noise sources. The noise based approach is increasingly important if the final HDR image is to be used for physically accurate irradiance representation in addition to visualization. Use of the software package Photosphere [Ward, 2010] to determine the camera response function and fuse multiple images was carried out by Inanici et al. [Inanici and Galvin, 2004], including determination of the vignette effect and Point Spread Function (PSF). Popadić et al. [Popadić and Todorović, 2016] present a comparison of four image fusion algorithms where the images are treated from a global approach as opposed to a traditional pixel level method. Finally, Griffiths et al. [Griffiths and Wicks, 2016] develop an unbiased test methodology for determining the aggregated error while using weighted averages as a maximum likelihood estimator.

2.2.4 Noise Based Weighted Average

A benefit of a noise based model for determining weighting averages is the consideration of individual noise sources of the sensor acquisition process. Early weighting methods for linear sensors provide variance estimates based on a simplified camera noise model, transferring the measurement uncertainty into the weighting function. Granados et al. [Granados et al., 2010] establishes an optimal weighting function implementing a more rigorous noise model, one that addresses both the temporal and spatial noise sources. This is the method we have used as the foundation of our new approach, building on the method by performing a radiometric calibration. Granados’ method implements an iterative optimization for the maximum likelihood estimate of the irradiance and its uncertainty, such that the weighting function is not perturbed by measurement noise. The predicted uncertainty can also be used to optimally denoise the resulting irradiance map. The noise based optimal weighting
function $w_{\text{opt}}$ for HDR reconstruction is given as:

$$w_{\text{opt}}(v_i) = \frac{1}{\sigma^2_{X_i}} = \frac{t_i^2 g^2 a_j^2}{g^2 t_i (a_j \mu_x + 2 \mu_D) + 2 \sigma^2_R}.$$  (2.51)

where $v_i$ is the non-saturated mean radiance of the $i$th exposure, $\sigma^2_{X_i}$ the uncertainty of irradiance $X_i$, $t_i$ the exposure time, $g$ the overall camera gain factor, $a_j$ the per-pixel gain factor for pixel $j$, $\mu_x$ the mean irradiance, $\mu_D$ the mean dark current, and $\sigma^2_R$ the spatial variance of the readout noise. $X_i$ and $\sigma^2_{X_i}$ are valid for $v < v_{\text{sat}}$, where $v_{\text{sat}}$ is the saturation limit. As part of the optimal recovery of $\mu_x$, estimates for the temporal readout noise parameters including spatial mean $\mu_R$ and spatial variance $\sigma^2_R$, the saturation limit $v_{\text{sat}}$, and the overall camera gain $g$, are generated by calibrating the camera sensor using the Photon Transfer Curve (PTC) method. The per-pixel gain factors $a_j$, including spatial noise components Photo-Response Non-Uniformity (PRNU) and Dark Current Non-Uniformity (DCNU), are estimated via flat field measurement with background subtraction.

With the PTC derived values, the per-pixel gain factors, and the LDR multiple exposure images $v_i$ and dark frames $b_i$, the optimal HDR recombination based on a rigorous noise model is performed using iterative estimates of the irradiance and its uncertainty. The goal is to reconstruct the mean radiance having the lowest variance from the set of multiple exposures to obtain the maximum likelihood estimate per Granados (Eq. 2.51) [Granados et al., 2010].

### 2.3 Summary

In this chapter, background terms and definitions used to describe light are presented including the topics of radiometry, photometry, luma, and colourimetry. Fundamental concepts of dynamic range and signal to noise ratio are then discussed, as they relate to image capture, along with a discussion on the use of arbitrary quality descriptors such as noise based speeds. Data formats are discussed next, including the comparison of integer and float point values, linear versus non-linear encoding, and the various HDR capable file formats available for use.

Related work follows, beginning with an analysis of image capture processing pipelines. Pipelines implemented for LDR, by standards committee’s such as MPEG, and by the research community, are presented. Gaps in capture processing are identified and summarised, creating a basis for this thesis and the need for a more robust method for determining linearization and calibration of HDR images. Finally, techniques for creating multiple exposure HDR images are discussed, along with a
specific review of the noise based weighted average implemented in this thesis.
Chapter 3

Methodology

As discussed in previous chapters, there is very little academic literature on how best to capture physically accurate linear HDR data without losing vital information. Although there is a wealth of publications related to the HDR pipeline including: capture, processing, tone-mapping, and display, the dynamic range of the test datasets is rarely taken to be physically accurate, or truly representative of real world lighting. In Table 2.1, we identified the main gaps in the current capture and processing pipelines; that linearity is either assumed, approximated, or of less importance than perceived realism. The goal of this thesis is to investigate methods for the capture and calibration of HDR imagery with the ultimate goal of developing solutions, and identifying limitations, in the collection of physically accurate radiance maps. Figure 3.1 illustrates the three part strategy for the research:

1. Investigating current methods and workflows for conducting HDR capture in a real world HDR environment,

2. Determining the “system” dynamic range, including both camera and optics, for comparison of dynamic range capability, and

3. Developing a methodology for creating physically accurate HDR radiance maps while identifying limitations.

3.1 Objectives

At a high level, HDR imaging workflows are structured similarly to that of traditional LDR imaging. Pipeline steps include capture, in-camera processing, storage, post processing, transmission, and display. Within each pipeline step however, HDR diverges from LDR, as increased capability in captured luminance range leads to a
Figure 3.1: Illustration of the research strategy: an investigation of existing HDR capture techniques, developing a system level dynamic range calibration, and finally the development and evaluation of a radiometric calibration based on a rigorous noise model. The methods, while having some manufacturer specific data workflow variations, can be applied to a variety of camera systems.

myriad of possibilities in the handling of data (figure 3.2). With content display, consider that SMPTE RP 71 standardized LDR studio monitors to a reference white luminance of $103 \text{ cd/m}^2$. Given a sample black level of $1 \text{ cd/m}^2$, the resulting contrast ratio of roughly 100:1 can be adequately encoded in a perceptually uniform relationship via 8 bits (i.e. Rec709) [Poynton, 2003]. With HDR capable monitors, both black and white capabilities are increased, and the ripple effect of this increase is perpetuated throughout the processing pipeline. Working backwards from display, new methods are required for storage and transmission, gamma and/or other non-linear electro-optical transfer functions, optimized data storage formats, treatment of non-linear camera response, and inclusion of advanced meta-data. The field of HDR has experienced significant development in all of these areas as demonstrated in the literature, yet the image datasets that are utilized in their development are, for the most part uncalibrated, usually with few details given as to how they were produced. Is the luminance data stored within the image physically accurate? If so, throughout the entire range, or only a portion of the range? In solving these questions, is it known whether the data is linear, or non-linear, or both? How do noise and saturation affect the range? And finally, what methods are available in the capture of the HDR data? These questions lay the foundation of this dissertation, and are inherently important to the HDR workflow, as without a complete understanding of the HDR image data itself, efforts in the resulting pipeline may be
imprecise, unnecessary, or lacking.

To answer these questions, the main objectives of this work are organized as follows:

1. In chapter 4, an analysis of HDR capture systems, challenges in implementing these systems, and differences between workflows is presented. The launch environment, as described below, is utilized as a test bed in reviewing the state of the art.

2. In chapter 5, a new calibrated method is developed to accurately measure and objectively compare the dynamic range capture capability of modern imaging systems using off the shelf equipment.

3. In chapter 6, a new complete HDR capture workflow is proposed with known and calibrated dynamic range and a physically accurate radiance map.

Figure 3.2: HDR pipeline from scene luminance to display

3.2 Approach

The approach used in this research will be to begin with an in-depth investigation into existing HDR capture technologies, followed by an examination of the types of vendor equipment available. In order to quantify the performance we have selected, an initial attempt to collect HDR data will be performed in a challenging real-world scenario: a rocket launch. Figure 3.3 shows a sample night-time rocket launch image taken with a dual exposure HDR camera. Based on the results of the investigation and the knowledge gained, a methodology will be developed to perform capture calibration utilizing off the shelf hardware and software. Finally, with the lessons learned, a detailed radiometric calibration of a sample capture system will be undertaken, and the implications to the HDR workflow will be addressed.
Figure 3.3: Night-time launch image of Delta IV 4 rocket taken with a RED EPIC camera operating in HDR+6 mode. HDR+6 mode takes two successive frames with different exposures times. An oval compositing mask is used to illustrate the exposures of the darker exposure in the flame area as opposed to the lighter exposure in the rocket area. Source: Advanced Imaging Lab, Kennedy Space Center

3.3 Rationale

Performing research into existing technologies provides an initial background into not only what HDR capabilities are currently available, but how manufacturers chose to characterize their capabilities, providing a baseline for comparison. Additionally, by utilizing existing hardware and capturing HDR data, an initial HDR capture workflow is established that will become the processing baseline. We will see that this processing baseline is similar in some circumstances, and varied in others, based on differences in both technology and methodology. We expect to identify gaps in our understanding of the HDR capture workflow, and with these gaps, provide the motivation for our final step to develop as closely as possible a radiometrically correct linear calibration of an HDR capture system.

As previously stated, HDR capture forms a fundamental first step in the HDR pipeline, and as of yet there is little literature on how best to capture physically accurate linear HDR data without losing vital information. This thesis proposes a
novel method for the calibrated capture of HDR imagery, develops a solution for the collection of physically accurate radiance maps, and investigates limitations due to the image sensor, camera processing, and optics.

3.4 Structure of Research

Having reviewed the current state of the art in HDR capture systems, the research now investigates challenges in implementing these systems, and identifies differences between workflows. A new calibrated method is proposed that accurately measures and objectively compares the dynamic range capture capability of modern imaging systems using off the shelf equipment. This enables a complete HDR capture workflow with known and calibrated dynamic range and a physically accurate radiance map to be developed. The efficacy and limitation of the new method are then comprehensively evaluated.

3.4.1 Chapter 4: Analysis of a Real World Rocket Launch Environment

To investigate existing HDR capture technology, a number of latest generation digital imaging systems were selected for testing and evaluation. Our knowledge of existing capabilities was initially limited, therefore a review of current manufacturer technology and documentation serves as a valuable resource for gaining initial insight into the HDR capture process. For instance, do manufacturers merely make image data available in a raw type format, or are post-processing tools made available? Is high bit depth data considered to be HDR data, and what makes it HDR data? What file formats and data storage methods are employed? These questions will be investigated in detail in Chapter 4. A review of published vendor specifications is also conducted, including capabilities described as “HDR” specific or capable. Each capture system is thoroughly examined for setup, operation, and post-processing, per vendor information availability. A discussion on the NASA requirements for rocket launch imagery in the Space Shuttle era is included, providing insight into real world imaging requirements. Rocket launches, especially those at night, present a challenging HDR environment due to the brightness of the plume and the desire to see detail of the launch vehicle masked in the night sky. Seeing detail in both the highlight and shadow regions of launch footage provides valuable information in determining flight safety. A review of film based capture capabilities is included, both for historical significance, and due to film being the Space Shuttle launch capture technology baseline. Film based imaging systems are not evaluated during our
testing, as film is being phased out due to unavailability, reduction in processing facilities, and overall operational costs. Work performed in evaluating existing HDR capture technology is organized around specific rocket launches, as they provide the targets of opportunity for testing. Our method for evaluating captured image data at this stage is to utilize vendor provided and/or third party software tools, based on the assumption that each vendor has created tools specific to their camera data and workflow. Data obtained from the existing capture technology investigation will be leveraged towards the development of more robust capture and calibration methods.

3.4.2 Chapter 5: Calibrated Measurement of System Dynamic Range

Several limitations are identified in the evaluation of existing HDR capture technology. First, in most cases, no standard is utilized when stating imager dynamic range capabilities by the various vendors. The issue is similar to “contrast ratio” in commercial televisions, where manufacturers state product contrast values independent of a common specification, making apples to apples comparisons meaningless. Second, in-camera processing of data may be described generically, such as with a vendor specific log curve applied, but only in some cases is the processing described, or the reverse transform provided in order to linearize the data. Third, there is lack of discussion, or available data, related to acceptable noise level or the noise level used in the determination of Signal to Noise (SNR) based sensor characterizations. While addressing the limitations at this stage of the research, a new method characterizing HDR response from a “system” perspective is proposed, including both sensor and optical effects together. The method requires only a small set of commercially available test equipment, primarily an HDR test chart (Xyla-21) and processing software (Imatest). This method is investigated first, as the equipment is readily available, and there is publically available test data from a commercially based image test website\(^1\) using a similar technique. In developing our version of the method, the arbitrary assignment of perceived quality terms to the results and their relationship to noise were also examined. This is an important aspect that is not addressed on the image test data website. To provide context for the measurement, a review of image sensor theory, the associated noise model, and calculation theory is included. The test method is described, including data workflow and analysis procedures. Results are presented for a number of modern imaging systems, including some stating HDR capture ability, along with manufacturer published capabilities. Results are discussed as they relate to perceived dynamic range quality, and the

\(^1\)www.DXOMark.com
limitations of the method are identified.

3.4.3 Chapter 6: Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model

A comparison of dynamic range data from our system based calibrated measurement, manufacturer stated data, and commercially published test website data, is hampered by the fact that the results cannot be compared directly without a more stringent definition of the test parameters, and specifically the definition of an acceptable noise level. Further, there is not a good basis for determining if our measurements, or any of the measurements used for comparison, are indeed accurate, as there are no calibrated independent luminance measurements for comparison. We also do not know the effects of the optical system versus the sensor, and if contributions from either are effecting the measurement. Contemplation of these unknowns leads to the realization that if we are to truly understand the basis of HDR capture, from which an appropriate HDR workflow can be developed, a detailed radiometric calibration needs to be performed. The calibration must treat the optical and sensor elements separately, so that the contributions from each are known. Further, processing steps taken at each stage of the capture pipeline must be considered in detail, ensuring that the resulting optical electrical transfer function is determined correctly. The method presented categorizes error via methodical calibration; including determination of sensor noise attributes, system gain response, and the effects of lens glare, such that the dynamic range of the imaging system can be characterized to a greater extent than previous methods. Additionally, using reference measurements from a calibrated luminance meter, absolute radiometric calibration is achieved, albeit with a degree of error. Results obtained from this method are used to improve the calibrated measurement technique based on commercially available equipment. The radiometric calibration leads to a greater understanding of the limits of HDR capture, which in turn will effects designs of entire HDR workflows.

3.5 Summary

This chapter has presented the methodology that will be used throughout the rest of the thesis. In Table 2.1, gaps in the current capture and processing pipelines are identified; mainly that linearity is either assumed, approximated, or of less importance than perceived realism. Additionally, absolute calibration is not verified versus a separate reference standard. In other words, HDR data is produced via the referenced pipelines, but whether the HDR data is a true representation of the
actual scene dynamic range, and what the scene total dynamic range truly is, is unknown.

The next chapter starts the investigation by performing an analysis of a real world rocket launch environment. By reviewing the current manufacturer technology and documentation, initial insight is acquired on the HDR capture process. Further, in evaluating captured image data, we utilize vendor provided and/or third party software tools, based on the assumption that each vendor has created tools specific to their camera data and workflow. In the chapter after this, we propose a new method for characterizing HDR response from a “system” perspective, including both sensor and optical effects together utilizing a small set of commercially available test equipment, primarily an HDR test chart (Xyla-21) and processing software (Imatest). Using this new method, we compare the stated manufacturer dynamic range versus the obtained results, and identity any shortcomings. Finally, we present a chapter on a new method for radiometrically calibrating an imaging system while evaluating the optical effects on HDR measurement. The processing steps taken at each stage of the capture pipeline are considered in detail, ensuring that the resulting optical electrical transfer function is determined correctly. The method presented categorizes error via methodical calibration; including determination of sensor noise attributes, system gain response, and the effects of lens glare, such that the dynamic range of the imaging system can be characterized to a greater extent than previous methods.
Chapter 4

Analysis of a Real World Rocket Launch Environment

The application of HDR to spacecraft launches is a completely uncontrolled real-world problem that requires HDR to be accurate and the slightest discrepancies to be understood in order to be able to take decisions with confidence from the resulting footage. It therefore represents a highly challenging scenario that is unachievable without a robust and quantifiable HDR workflow. Accurate imaging is crucial during the launch to enable ground controllers to determine if there have been any anomalies during the launch, such as an object becoming detached from, or striking, the vehicle. As an example, night rocket launches present a two-fold problem, capturing detail of the vehicle and scene that is masked by darkness, while also capturing detail in the engine plume. Figure 4.1 contains two frames of a debris event as captured using traditional non-HDR imaging. The images originated from an 8 bit, 4:2:2, Rec 709 colour space data file during a spacecraft night mission. Figure 4.1 (left) illustrates debris to the left and above the vehicle plume in the night sky. The camera exposure was set to capture detail in the flame, while the limitation in dynamic range results in little to no observable detail in the vehicle area. Figure 4.1 (right) includes debris in front of the rocket engine plume several frames later. While the debris can be observed in and near the plume area, there is a loss of useful visual information of the rocket that may show the origination point of the debris from the vehicle, how the debris was generated, and the remaining state of the vehicle. Detail of the debris itself is also reduced as a result of the limited dynamic range, amongst other factors such as image resolution and optical system sharpness.

The focus of this chapter is the development and testing of digital imaging systems with respect to the challenging scenario of HDR imaging of spacecraft. A
review of the dynamic range of film, the current engineering imaging baseline, is presented along with relevant characteristics. Past work related to the imaging of the Space Shuttle is included. Field experiments are then described, with the inclusion of sample image frames.

4.1 Background

The NASA goal for engineering quality imagery, including processing and compression, is being visually lossless, i.e. no loss of perceived quality once compression is applied, as compared to the original captured file. The ability to capture engineering quality imagery with a wide degree of dynamic range during rocket launches is critical for post launch processing and analysis [Columbia and Gehman, 2003]. Rocket launches often present an extreme range of lighting during both day and night launches. During the Space Shuttle Program, engineering imagery was collected at Kennedy Space Center (KSC) primarily through the use of 16mm and 35mm Photosonics high speed film cameras. The film cameras include automatic exposure control, shifting the available dynamic range in response to scene luminance. HDR images may be captured through a single sensor using multiple exposure settings, dual gain sensors, or multiple sensors in combination with beam-splitters and neutral density filters [Mann and Picard, 1995; ARRI, 2013a]. In addition, new
classes of imaging sensors are being developed that are capable of directly capturing a larger dynamic range, for example those proposed by Fuji and Panasonic [Cade, 2013]. Workflows in typically lossless HDR formats, as discussed in Section 2.1.4 Data Formats, such as Radiance RGBE [Ward, 1991] allowing for half-float, as well as OpenEXR [Bogart et al., 2004] and TIFF allowing for half-float and full-float representations, are required for HDR data [Ward, 1991]. Lossy compression formats such as goHDR and HDR MPEG [Mantiuk et al., 2006a] may also be utilized for bandwidth reduction and smaller file size. The ability to display HDR imagery directly on HDR displays is also evolving rapidly with the advent of LED modulated LCD HDR displays. These HDR capture, processing, and display workflows aim to reproduce the simultaneous contrast ratio capability of the Human Vision System (HVS), and indeed even more dynamic range than the human eye can see [Seetzen et al., 2004]. The entire capture, processing, and display workflow must be considered when evaluating dynamic range, as each subsystem can potentially enhance or degrade the overall system capability.

4.2 Film Baseline and History

16 and 35 mm film was utilized as the engineering imagery baseline during the Space Shuttle Program. Investigation is now ongoing to consider replacing or at least complementing film with advanced digital technology. Digital cinema has a specific objective in striving to replicate the look of film, including colour and tonal response. Film look is a cinematic term describing a common Hollywood style of shooting. Shallow depth of field, 24 frame per second framerate, and 180° shutter are primary characteristics, along with lighting, cinematography, post-production, and even distribution method as additional contributors to style. Film studios additionally view long term protection of the digital negative as paramount, necessitating the need for a robust digital archival system as seen with the Academy of Motion Picture Arts and Sciences (AMPAS) introduction of the Academy colour Encoding Specification (ACES) [SMPTE, 2012]. In the broadcast market segment, as tone and colour characteristics of digital displays improve, the broadcast industry is evolving with new specifications such as BT.2020 to expand tonal and colour capability beyond current broadcast standards that are based on CRT phosphor approximation. The scientific and engineering community alternatively is quite often looking to push boundaries by developing and verifying new ideas and concepts, utilizing RAW or nearly RAW imaging data in the analysis of their work [Arens et al., 2014; Karr et al., 2016; Ashok and Wray, 1995; Madzsar et al., 1992; Larson and Shake-
speare, 1998; Reinhart and Herkel, 2000; Mantiuk et al., 2006b; Akyz and Reinhard, 2007]. The purchasing power of the scientific and engineering industry however, as compared to cinema and broadcast, has generally not been at the forefront of the field and thus unable to significantly influence the migration path of digital video development, except perhaps in niche high end examples such as high speed camera systems. While each market faces its unique challenges, they collaboratively require standardization throughout the HDR pipeline including the capture process, the production workflow, and finally in the rendering of the final images to the display [Moir and Chalmers, 2016].

When considering technical merits, manufacturers of digital technologies strive to demonstrate the extent of their latest achievements versus the film baseline. Recent technological advancements have been fast paced, and key metrics of frame rate, resolution, colour gamut, and dynamic range have progressed with varied stages of implementation. With regard to resolution, film scanning resolution, as determined by the system Modulation Transfer Function (MTF), improves dramatically from 1000 to 3000 samples per width with reduced improvement to 4000 samples per width, and negligible improvement beyond [Kennel, 1994]. For this reason, film scanning systems such as by Cineon are designed around 4K resolution, a resolution that is commonly available and even exceeded in digital capture imaging systems today. Therefore, many digital capture resolutions are on par with 4K film scanning resolutions. The second metric, colour gamut, now has increasing capability after relatively little progression for decades. Previous slow development is primarily due to the fact that colour gamut encoding for video has been standardized since 1990 with the adaption of ITU Rec. BT.709, a moderate colour gamut with primaries chosen to approximate CRT phosphors [Poynton et al., 2015]. BT.709 can render an approximate 33% of the visible colour gamut, as opposed to a representative film gamut that covers approximately 66% [Miller et al., 2012]. Increased colour gamut is achievable today via the SMPTE standardized reference projector gamut known as P3, but this is primarily used in digital cinema as part of the Digital Cinema Initiative (DCI). More recently, the concept of Wide colour Gamut (WCG) has been introduced to capitalize on advancements made through the use of technologies such as LEDs and laser light, enabling display gamuts exceeding that of BT.709 and P3. New proposals are included in the ITU-R BT.2020 Recommendation on Ultra High Definition Television (UHDTV) [ITU-R, 2012b], and the Academy colour Encoding System (ACES) [ACES, 2015], that should invoke further discussion as they are implemented. It is the third technical merit, dynamic range, that is experiencing a windfall of new development, and whose capability is still actively debated and
compared to film [Chalmers and Debattista, 2017]. Modern imaging systems however, particularly those involving HDR, are surpassing the dynamic range of film as well as the simultaneous contrast capabilities of human vision. In recent years, digital imaging technologies have challenged film as the gold standard in the capture of motion images.

4.2.1 Dynamic Range of Film

The dynamic range of “film” from a system perspective should include the negative, processing, and printing/display combination. Modern motion picture original colour negative stocks such as Kodak Vision3 capture images with red, green, and blue records representing up to a 13 stop scene [Kodak, 2010]. Kodak estimates that the usable digital scan range is 10-13 stops depending on the scanning and encoding scheme [Kodak, 2009]. One method for retaining the full density range is through the use of adjusted density mapping. Film negatives however capture more than can be reproduced on the print, and follow a characteristic “film” curve for Relative Log Exposure. The classical approach to densitometry (scientific analysis of exposure) was devised by Hurter and Driffield and is called the H&D curve [Davis and Walters, 1922]. The H&D curve (also called the \(D \log E\) curve, \(\log E\) curve, or Relative Log Exposure) plots the amount of exposure \(E\) in logarithmic units along the horizontal axis, and the amount of density change \(D\) along the vertical axis. A sample \(D \log E\) curve is shown in figure 4.2 [Kodak, 2006].

In practice the exposure curve indicates that film response is not linear throughout the exposure range. A theoretical perfectly linear response would have a 45° slope. The slope of a relative exposure response in practice is a measure of its contrast, while the slope of the straight-line portion of the curve (essentially ignoring the shoulder and toe) is referred to as the Gamma. The upper and lower end of the film emulation response differs than that of the Gamma section. At the low end, the film does not initially respond as it “sees” radiant energy and then responds sluggishly for initial increases. At the upper end, the response is compressed as the emulsion becomes overloaded. These areas are indicated by the toe and shoulder regions in the H&D curve. The logarithmic data could be linearized using a nominal film transform such as a Linearized Printing Display (LPD). Use of a linear scene-referred representation would require detailed characterization of each film emulsion, which film manufacturers often consider confidential. Additionally, the inverse linearization of the film H&D curve would produce very high gain in the non-linear toe and shoulder regions, potentially producing high gain artifacts [Kennel, 2014]. The film lab timer sets the exposure range of the negative that the
print will be exposed to in a development process using reducing chemicals. The process is stopped either by washing or using a stop bath. This is a relative process where the exposure is determined by the timer. Once exposed, printed, and digitally scanned, the process will need to be repeated if an adjustment in overall exposure is required. This process is both time intensive and costly. For Space Shuttle engineering film processing, the film was flown from Kennedy Space Center, Florida, to a processing facility near Atlanta, Georgia, processed and scanned, with a goal to return the digital transfers within 24 hours. As film processing labs have reduced in number, processing cost continues to increase.

### 4.2.2 Auto-Exposure and Contrast Ratio

Film cameras implemented auto-exposure systems during Space Shuttle launch imaging. At times there was some confusion amongst analysts with regard to the available dynamic range with the use of auto-exposure. Contrast ratio, the ratio of luminance of the lightest and darkest elements of a scene or image, can be used to describe different aspects of auto-exposure. When auto-exposure is implemented, the lighting information captured often changes with time. Therefore two measures of contrast ratio may be used (Poynton, 2003). Simultaneous contrast ratio (static) is the contrast ratio at one instant of time. Sequential contrast ratio (dynamic) is
the contrast ratio separated in time, i.e. light and dark elements over a series of images. Human vision may have a simultaneous contrast ratio of close to four orders of magnitude (3.7 density, 12.3 stops) (Kunkel and Reinhard, 2010), while human sequential contrast ratio via adaptation can be as great as ten orders of magnitude or approximately 33 stops (Ferwerda, 2001). Film and digital imagers both can be defined in terms of simultaneous and sequential contrast ratio.

4.2.3 Linear and Non-linear Quantization

In both scanned film and digital-based imaging of spacecraft, the final result is often a digital data file. In non-typical cases, the film print may be utilized directly in a projection system. In the film case, the processed negative is scanned within the capabilities and bit depth of the scanning system to a digital file generally referred to as a Digital Intermediate (DI). In the digital imager case, the analog sensor voltages are converted to linear digital values within the capabilities and bit depth of the camera sampling system. A primary constraint in both cases is the digital bit depth, or the number of discrete quantization levels (or codes) available to represent a given relative luminance range. There are two primary methods for bit depth encoding that are implemented by digitizer manufacturers, linear and non-linear scales. Non-linear scales are often referred to as log or log-c, even though the functions are often in practice based on power law as it can be a better approximation to human visual response. In general, linear scale quantization is utilized when the sampling resolution is to be kept constant over the entire exposure. Non-linear scale quantization is utilized when the sample resolution follows the logarithmic response similar to human perception of luminance. By implementing non-linear quantization, fewer bits are required to theoretically match the viewing capability of human luminance response; the visual threshold. The tradeoff is that a reduced subset of total exposure information is stored, limiting image analysis capability when detail within a subset of the contrast range is required.

A typical industry scanning metric is SMPTE 268M Digital Moving Picture Exchange (DPX), logarithmic with 10 bit integer calibrated over a 2+ density range (SMPTE, 2003). Density range is defined in the equation:

\[
\text{Density Range} = \log_{10} \left[ \frac{\text{# of photons transmitted through the film}}{\text{# of photons hitting the film}} \right] \quad (4.1)
\]

with the relationship to dynamic range in terms of stops defined in the equation:
In practicality, a Cineon digital film scanner is calibrated for a 2.048 density range [Kennel and Snider, 1995]. HDR imagery, on the other hand, is typically stored utilizing float point values.

### 4.2.4 Playback and Display

The final, yet equally important, piece of the dynamic range workflow for both film and digital technologies is the playback software, hardware and display. Regardless of the capture and quantization methods, the end user will only be able to view the simultaneous contrast ratio within the limitations of the end playback software, hardware, and display system. An example film workflow utilizing a Digital Intermediate and a typical 8 bit display is shown in figure 4.3.

![Figure 4.3: Film to digital transfer with display on 8 bit monitor.](image)

In this example, the full dynamic range benefit is not realized as a result of playback and display limitations. Limitations in dynamic range were present in the review of Space Shuttle imagery, depending upon the equipment available at each participating NASA Center. One example display device, a 10 bit 4K projection system, is limited in dynamic range due to a peak luminance of only 60 cd/m² measured at the projection screen. For redundancy, Image Analysts were available at Johnson Space Center (JSC), Marshall Space Flight Center (MSFC), and Kennedy Space Center (KSC), albeit each with different equipment and capabilities. Although the differences between Centers were a result of several factors including...
areas of expertise, program responsibilities, and budget, the differences do have a benefit of a more comprehensive and overlapping analysis approach.

Figure 4.3 is dependent on the specifics of the DPX scan, as well as the setup and calibration of the monitor. In converting from the full-range 10 bit log digital negative to an 8 bit Rec 709 video representation, the preferred method is to limit the density range that is translated to that of a normal exposure with scene contrast range of 100:1, or $6\frac{2}{3}$ stops. The resulting scene will have reasonable contrast when displayed on a standard graphics display monitor [Kennel and Snider, 1995] where the $D_{\text{min}}$ code value of 95, and the 90% white card code value of 685, are mapped to an output range of 0 to 255 respectively in the 8 bit representation.

4.2.5 Summary of Film Use

During the Space Shuttle program, 16 and 35 mm high speed film cameras provided greater dynamic range capability than digital imagers available at the time. The film assets were only minimally utilized, on average for 4.5 Shuttle flights per year. With an on-going maintenance program, they did not require regular replacement. The development of digital imaging technology, along with increases in dynamic range capability, occurred as the Space Shuttle program reached end of life. As the program ended, an upgrade of film assets to digital technology was cost prohibitive in terms of equipment, infrastructure, training, and operation.

A transition to digital technology for future programs, however, can provide several benefits. Although the film emulsion can capture up to 13 stops, in the majority of cases the film is transferred to a digital intermediate with reduced dynamic range, often displayed via limited contrast ratio display. The process of developing film is time consuming and costly, and is becoming more difficult to procure. Operation and maintenance of film systems requires specialized skills sets. Post launch, the exposed films are collected, transported to a local airport, and flown to a development house often out of state. The film is developed, scanned, and the digital files are transported back to Kennedy Space Center (KSC) after as much as 24 hour delay.

In the case of digital capture, files are available much faster, often with access to the RAW data quickly after launch. HDR data capture surpassing the dynamic range capability of film is becoming more prevalent, HDR file containers are available that support HDR data types, and HDR workflows allow for the processing of HDR data. HDR displays are becoming available capable of expanded contrast ratio. The entire HDR workflow aims to fulfill a primary requirement of image analysis, the ability to process, display, and analyze imagery with increased simultaneous contrast.
4.3 HDR Imaging of Spacecraft Field Experiments

Testing of HDR digital image capture began in 2009 by the Advanced Imaging Lab (AIL), in preparation for an eventual film to digital transition. Experiments were initially conducted based on equipment currently on-hand or available via loan, while preparations were made to obtain the latest generation digital technology. As the Space Shuttle program was coming to a close, field experiments were often planned in advance of lab testing, in order to capture actual launch imagery while the launch opportunities were still available.

The following section summaries the lessons learned during the imaging experiments, presented in the chronological order during which a continuing knowledge base was advanced. The overview of the experiments is included in Table 4.1.

4.3.1 High Definition Camera Test During STS-129

The first HDR experiment utilized two industrial 8 bit HD1100 \( \frac{2}{3} \) inch sensor cameras with similar optics to capture similar fields of view, albeit with different exposures. The experiment was conducted during STS-129 [Clements et al., 2009] with the imaging systems mounted to a common tracking system. The A camera was mounted to the upper side of the tracker right wing and the B camera was mounted to the underside. At the time of the experiment, the HD1100 shutter speed was limited to 1/300 of a second. The difference in exposure from Camera A to Camera B was \( 2\frac{1}{4} \) stops based on available exposure times of \( \frac{1}{60} \) and \( \frac{1}{300} \) of a second. Both cameras were mounted with Meade 8 inch aperture telescopes, focal length of 2032 mm. The \( \frac{2}{3} \) inch sensors (8.8 x 6.6 mm) have an equivalent 3 5mm crop factor of 3.9. The telescope and crop factor result in a tight field of view with the tracker distance of approximately 3 miles from the Shuttle Pad.

Atmospheric effects of ground heating resulted in a reduction of overall image quality as can be observed in figure 4.4. The effects are particularly pronounced with daytime launches during the Florida summer months, where temperatures are commonly in above 30 C. Limitations in shutter speed also contributed to the softness of the imagery during the actual launch. Mechanical errors in the camera alignment stage are evident when comparing figure 4.4 and figure 4.5 complicating efforts to recombine the synchronized frames into single HDR frames. Independent vibration affected each of the two image sequences, primarily a result of the launch
Table 4.1: Summary of HDR Imaging Spacecraft Field Experiments

<table>
<thead>
<tr>
<th>Launch Opportunity</th>
<th>Experiment</th>
<th>Camera Utilized</th>
<th>Key Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Shuttle</td>
<td>Dual cameras with telescopes</td>
<td>HD1100 Industrial Box Cameras</td>
<td>Atmospheric effects noted, mechanical alignment error, independent vibrations</td>
</tr>
<tr>
<td>STS-129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Single sensor camera during night launch</td>
<td>RED ONE M Sensor</td>
<td>Digital capture of lighting conditions during night launch, no alignment issues with single sensor</td>
</tr>
<tr>
<td>STS-131</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Single sensor camera during day launch</td>
<td>RED ONE M Sensor</td>
<td>Highlights saturated based on exposure settings</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Compare “film look” versus “engineering” exposure settings</td>
<td>ARRI Alexa</td>
<td>“Engineering” exposure sharper due to faster exposure time, “Film Look” exposure required additional ND filter, Alexa post-processing workflow uses 3D LUTs</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Evaluate camera with LogFilm contrast curve</td>
<td>RED EPIC 5K</td>
<td>DR similar to RED ONE imagers tested, higher pixel count allows for increased resolution when in details</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Evaluate high frame rate camera</td>
<td>Photron SA2</td>
<td>DR appears similar to other cameras, highlights saturated, high frame and short exposures provide sharp images</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Evaluate high frame rate camera</td>
<td>Cooke DiMax</td>
<td>Banding observed around saturated scene areas</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Shuttle</td>
<td>Compare high frame rate camera with film camera</td>
<td>Vision Research Phantom HD Gold</td>
<td>Film camera auto-exposure demonstrates sequential contrast ratio, Phantom camera highlights saturated with fixed exposure, but do not bloom into surrounding pixels</td>
</tr>
<tr>
<td>STS-135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta 4 WGS 5</td>
<td>Evaluate camera HDR mode</td>
<td>RED EPIC with HDR Mode</td>
<td>Produces dual exposures known as A and X frames, Vendor supplied software tools for “combining” dual frames to single image</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Acoustics on the mechanical support structures. Recombination efforts did not result.
in acceptable video due to the misalignments and vibration affects.

4.3.2 RED ONE M Camera Test During STS-131

As a result of misalignment issues noted with the two HD1100 imagers, market research was conducted to identify the latest single sensor imagers advertised with
increased dynamic range capability. One contender identified was the RED One imager from RED Inc. RED specifies the dynamic range as 11.3 stops (assumed to be MX sensor). The AIL measured 10.4 stops with the M sensor utilizing Log gamma, and 11.5 stops with the MX sensor utilizing LogFilm gamma. Both measurements were made with respect to 0.5 stop RMS noise, as discussed in the section on calibrated measurement of imager dynamic range. M sensor versions of the cameras were available and configured for STS-131 [Clements et al., 2010] to acquire imagery of the launch from five camera sites around the Shuttle pad perimeter as illustrated in figure 4.6.

![Diagram of Space Shuttle pads perimeter camera site locations](image)

**Figure 4.6:** Space Shuttle pads perimeter camera site locations (Courtesy Marshall Space Flight Center Engineering Photographic Analysis).

Deployed in pairs to each site, the cameras provided imagery not only for HDR testing purposes, but also to generate stereoscopic 3D image sets. A diagram of the Space Shuttle pad perimeter camera site locations is shown in figure 4.6. The RED ONE M experiment was successful in illustrating the capability of a single exposure digital imager at the time. Clearly the Solid Rocket Booster (SRB) plume is saturated and outside the available dynamic range. Detail of the vehicle is present, primarily lit from the rocket plume, with the addition of some ground-based spotlights. The five camera locations had similar exposure settings, resulting in similar dynamic range detail for all image sets.
Figure 4.7: RED ONE composite of 5 camera sites.

Figure 4.8 is a full frame from the RED ONE M at Camera Site 3, southeast of the Launchpad, illustrating the available lightness detail of the full plume. Launch debris may be present throughout the field of view, ranging from the Solid Rocket booster (SRB) bright plume to the dark shadow areas of the plume cloud. If the shadow areas are comparable to starlight, and the SRB bright plume is considered to be greater than direct overhead sunlight conditions, the total dynamic range of the launch environment is greater than 8 orders of magnitude. Future effort will be made to quantify the launch environment total dynamic range.

4.3.3 HDR Imaging during the Final Space Shuttle Flight STS-135

Market research was expanded for the final Space Shuttle flight of STS-135 to identify the growing number of available single sensor imagers with increased dynamic range capability. New systems identified included the Vision Research Phantom HD Gold High Speed, RED EPIC with 5K resolution sensor, Cooke DiMax High Speed, Photron SA2 High Speed, and the ARRI Alexa [Lane et al., 2011]. The RED ONE with M sensor was again deployed as a baseline for comparison.

RED ONE M Test

A RED ONE camera with M sensor was deployed to the infield location near Camera Site 4. A 25 mm RED lens was used for the wide angle field of view. The resulting
imagery was similar in dynamic range response to previous tests. A sample frame is shown in figure 4.9. Although the SRB plume is saturated, the overall imagery is high quality, with detail observed throughout the field of view. An additional benefit is that the RED ONE imager is capable of 4K resolution.
ARRI Alexa Test

Two ARRI Alexa cameras were placed within the Shuttle pad perimeter. The A camera was configured to capture a “film look” exposure with 180° shutter at 30 FPS (16.67 mS exposure time) with a 50 mm lens. The B camera was configured to capture an “engineering” exposure with 22.5°shutter at 120 FPS (0.52 mS exposure time) with a 40 mm lens. Whereas the 180° shutter creates the familiar film look motion blur, the engineering 22.5° has the goal of reduced motion blur and sharp individual frames. Film look is a subjective term. It is commonly associated with the human visual experience when viewing a film-capture based motion picture. Typical characteristics of “film look” include 24 FPS frame rate, 180° shutter angle, narrow depth of field, and film grain type noise. An engineering exposure alternatively, has a goal of reproducing as sharp an image as possible in each individual frame, with high frame rate.

Figure 4.10: ARRI Alexa 120 FPS B Cam sample frame.

The 180° A camera required the use of ND filter due to the longer exposure time. For the B camera, this test was one of the first utilizing ARRI’s 120 FPS feature. The ARRI Alexa is marketed as capable of 14 stops of dynamic range. The AIL lab test measured 13.9 stops at 0.5 stop RMS noise in Log mode. Figure 4.10 contains a sample frame from the “engineering” 120 FPS B camera. Although the SRB plume is saturated, the overall imagery is of a high quality. When digitally cropped on a sub-section of the digital image that includes fast moving content, such
as the orange coloured SRB exhaust plume, the 180° shutter A camera image is less sharp than the 22.5° shutter B camera. The Alexa imagery workflow was considered one of the simplest whilst still producing high quality imagery. The straightforward application of a 3D Look Up Table (LUT), generated from the online ARRI LUT Generator, to the log data files resulted in realistic looking colour and lightness without further grading required [ARRI, 2013b]. The LUT, in the case of AIL processing, was generated for the DaVinci Resolve colour grading system.

RED EPIC Test

The RED EPIC without HDR mode was lab tested by the AIL as 11.5 stops at 0.5 stop RMS noise, LogFilm mode. Red LogFilm applies a log curve to the linear sensor data without the addition of a film-like contrast curve. LogFilm replicates the characteristic curve of Cineon film scans [RED, 2013]. For the STS-135 test, HDR mode was available but had not yet been tested, so it was not utilized. The RED EPIC has the benefit of a 5K resolution imager, upgradable to the Dragon 6K sensor. The EPIC was configured to capture at 96 FPS at a 5210×2700 resolution, 100mm lens, and a 1ms shutter period. The imagery captured was similar in dynamic range to the other imagers tested in that the SRB plume was saturated. Overall the EPIC footage was of a high quality and the increased pixel count allowed for higher quality cropping within the imagery. A sample RED EPIC frame is shown in figure 4.11.
**Photron SA2 Test**

A Photron SA2 high speed digital camera with 2K resolution was deployed near Camera Site 3 utilizing a 500 mm lens. The frame rate was set to 500 FPS. Figure 4.12 illustrates a frame from the SA2 showing a debris item that was observed.

The field of view was selected specifically to image the Space Shuttle Main Engine (SSME) hydrogen burn off sparklers. Portions of the SSME flame are saturated, although detail can be seen in other portions as evident in the identification of a debris item passing behind the plume.

![Figure 4.12: Photron SA2 sample frame with debris item inset.](image)

**Cooke DiMax Test**

A Cooke DiMax high speed digital camera was mounted on a tracker near Camera Site 2 with an 85mm lens. The camera resolution was 2016×2016 capturing at 250 FPS. The camera did exhibit some banding immediately surrounding the saturated section of the plume as seen in figure 4.13, but otherwise operated as expected. Upon
reviewing imagery post-launch, the manufacture recommended a darker capture exposure setting.

**Figure 4.13: Cooke DiMax sample frame.**

**Vision Research Phantom HD Gold Test**

A Vision Research Phantom HD Gold high speed digital camera was located on the 275 foot level of the Shuttle Fixed Service Structure (FSS). A 16mm film camera was also located at this level, with similar fields of view configured for comparison. A sample frame from the Phantom Gold is shown in figure 4.14, while sample frame from the 16mm film camera is shown in figure 4.15. Both cameras were operated at 400 FPS. The film was scanned via a film to HD 10 bit log transfer. The Phantom
captured 14 bit linear data with gamma applied during the colour grading process. Both frames have been cropped per their respective aspect ratios.

The darkened film image in figure 4.14 is a result of the automatic exposure control. More detail is observed in the plume area in the film image as compared to the digital image, at the expense of the darkened overall image. The Phantom camera did not utilize auto-exposure and was configured for proper exposure of the vehicle. The SRB plume is observed to saturate in the Phantom image.

The changing exposure level of the film represents sequential contrast ratio, complicating comparison over the entire image sequence. From a single frame perspective as shown, simultaneous contrast ratio can be observed, albeit at different
exposure levels. The dynamic range of the Phantom Gold HSD imagery was better than expected as the Solid Rocket Booster (SRB) plume was saturated but did not bloom into surrounding pixels.

4.3.4 RED EPIC HDR Mode during Delta 4 WGS 5

The RED EPIC capability to capture in HDR mode, allowing the selection of an additional 1 to 6 stops, was tested in the post-Space Shuttle period. The EPIC stores two different frames in HDR mode, the A frame and the X frame. The RAW A frame contains the “standard exposure” portion of the capture, the RAW X frame contains the “highlights” portion of the capture. RAW refers to the high quality encoding format of the RED EPIC camera known as REDCODE which is a proprietary RED un-debayered linear wavelet based compression [RED, 2012].

An example EPIC HDR capture was made during the Delta 4 WGS 5 mission [Karr, 2013]. This HDR+6 mode captured approximately 17 stops. In order to illustrate this range in limited print form, or via a LDR monitor, the imagery must be shown as two separate frames. The A frame is shown in figure 4.16 and the X frame in figure 4.17.

The capability to capture this range of brightness detail is unprecedented for a single imager based system, and is an excellent example of the current progress in HDR image capture.
4.4 Review of HDR Field Experiments

This section discusses field testing of digital imagers, providing insight into the current state of technology and the ever increasing complexity of working with a multitude of digital formats, as well as the benefits and tradeoffs of the different imaging systems.
State of Technology

How captured digital data is provided to the user is manufacturer dependent. For example, ARRI utilizes a Dual Gain Architecture (DGA) with two paths of different amplification for each pixel [ARRI, 2013a]. ARRI processes the dual path data internal to the camera, providing a single output file available either as RAW (with external recorder), LogC, or Rec709. Generation and application of post-processor specific LUTs is straightforward, resulting in a particularly user friendly workflow.

RED provides its HDR digital data as two separate file sequences comprising of A and X frames. Access to both exposure frames has the benefit of increased flexibility in post-processing, at the cost of increased complexity in understanding the tone-mapping process. The RED software CineXPro includes two tone-mapping functions, Simple Blend and Magic Motion, assisting the user in the tone-mapping process. Overall, the RED HDR mode had the highest dynamic range measured, as will be discussed in Chapter 5, Calibrated Measurement of Imager Dynamic Range.

The high speed camera manufacturers such as Vision Research, Photron, and Cooke DiMax all provide the capability to export data from manufacturer formats to either linear or gamma corrected image sequences. In most cases, 16 bit short integer linear TIFF sequences were utilized, as HDR type formats utilizing half floats such as LogLUV TIFF were not available [Larson, 1998]. The accompanying software packages from the high speed vendors generally include basic image processing tools, with the expectation that the majority of image processing will be conducted in a third party application. As the RED and ARRI are considered cinematic cameras encompassing tools and workflows to aid the user in producing cinematic like images, high speed cameras have traditionally been utilized in industrial applications such as vehicle safety crash testing, spacecraft launch environments, and munitions testing.

Digital Formats

A significant challenge in testing a large sampling of digital imagers is developing data workflows. For as many vendors utilized, there are just as many proprietary data formats, encoding algorithms, export formats, software packages, and post-processing guidelines. Metadata varies and is non-standard. Color range and exposure was observed to vary from camera to camera, even within manufacturers, potentially requiring custom LUTs per individual imager for precise characterization of colour and luminance response as required.

In some cases, file wrappers, encoding schemes, gamma, and colour space functions are vendor specific such as with RED Redcode files, but have wide accep-
tance and are importable and manageable either through vendor provided software or popular third party applications. In other cases such as with ARRI, file formats typically use ARRI LUTs for gamma and colour space, but are wrapped in standard ProRes422 or ProRes444 wrappers. Some high speed cameras such as the Vision Research Phantom have file formats that are natively importable to popular third party applications, while others are not and must be exported as an image sequences for import to third party applications.

**Limitations**

From the experiments and evaluation of the state of the art in HDR image capture, we find that the processing pipelines differ significantly from manufacturer to manufacturer, and that there is no clear definition as to how imaging system dynamic range is either measured or specified. Gaps exist between how to accurately evaluate imaging requirements against Vendor described capabilities, what the range of the dynamic scenes truly are, and how best to process the captured data in an HDR processing pipeline.

### 4.5 Summary

The chapter began with a review of film capture, serving as the historical baseline for the capture of scenes containing wide luminance ranges. The inherent characteristic response of film is described, known as the Hurter and Driffield (H&D) curve, along with the post-production digitization process of the film negative. An example playback and display workflow for the digitized film is then presented, along with processing limitations. Next, field experiments conducted with modern imaging systems are presented, using rocket launches as targets of opportunity. Experiments were conducted using dual imagers in a side by side configuration, single imagers with single exposure capture, and single imagers with dual exposure capture. As part of the experimental process, knowledge was obtained as to different manufacturer data availability, file formats, processing software, and in some cases, in-camera data processing. Finally, a review of field experiments is presented, including the state of technology, and a review of the file formats encountered. We conclude from these experiments that the processing pipelines differ significantly from manufacturer to manufacturer, and that there is no clear definition as to how imaging system dynamic range is either measured or specified. This primary finding leads to our next chapter, the calibrated measurement of system dynamic range. In this chapter, we develop a methodology for the testing and comparison of modern imaging systems.
Chapter 5

Calibrated Measurement of System Dynamic Range

In the previous chapter the experiments and evaluation of the state of the art in HDR image capture demonstrated that the processing pipelines differ significantly from manufacturer to manufacturer and that there is no clear definition as to how imaging system dynamic range is either measured or specified. In this chapter we present a method for performing a system calibration of DR where the effects of optics, sensor, and processing are lumped into the system definition. In performing the system calibrations, the processing pipelines for various camera systems are described.

The chapter begins with the overall design of the calibration and the materials required to perform it. Test charts consist of patches with uniform optical density steps, covering a specified range. Many early generation reflective or transmissive test charts are limited in optical density range, and therefore are not suitable for high dynamic range measurements. The method presented includes a latest generation back lit transmissive test chart, the Xyla-21, having 21 patches covering an optical range of 20 stops. The Xyla-21 includes additional design features including uniform backlighting of the patches for increased accuracy, xylophone shaped patch structure from bright to dark patches for reduced glare effect, and factory calibrated density measurement data for each density step. In order to speed up data collection and processing of patch measurements, the test software Imatest is utilized. Imatest image ingest requires the registration of input selection boxes over an approximate 50x50 pixel region of each patch, providing enough patch data to average local noise variation. After all 21 zones are identified, Imatest automatically calculates statistics for each region of interest. Utilizing these new tools, adherence to
existing measurement standards is ensured to provide accurate measurements. Control of ambient light, reflections, test environment temperature, and other factors are factored into the measurement procedure.

In conducting the calibrated measurements of this chapter, a significant contribution is the development of the data workflows necessary in processing data from the many new camera systems, and image capture technologies, marketed with HDR capability. The large sampling of digital camera systems illustrates great variation in processing workflows, including proprietary data formats, encoding algorithms, export formats, software packages, and post-processing procedures.

A final contribution of this chapter includes identifying the many differences in manufacturer quality specifications, descriptive narratives, and use of arbitrary quality descriptors. Value added characterization of quality descriptors to a meaningful metric of RMS noise, and an ISO standard, is presented. A test procedure follows including the data workflows, results, and a discussion on quality descriptors.

5.1 Method and Materials

In this section the methodology for the system level DR calibration is presented, including the test chart materials and processing software required. The cameras selected for testing are then described, where camera selection was based on availability.

5.1.1 Design

Overall, the procedure for measuring system dynamic range involves four primary steps. Step one includes the setup of the test environment so that overall lighting conditions can be controlled while minimizing reflected light. The ability to set, maintain, and read room temperature is also of importance as camera noise may be temperature dependent. The second step is the setup and alignment of the Camera Under Test (CUT) and HDR test chart in order to meet the pixels per patch test requirement. With setup activities complete, the third step includes the collection of sample image frames, followed by post-processing to create compatible input frames for the analysis software. Finally, step four includes processing of the input frames through the analysis software and the review of data.

Three main components in the image capture pipeline that should be considered in measurement of dynamic range include the optics, the sensor, and the camera processing workflow [Xiao et al., 2002]. The optics are important in transforming the scene luminance to the image luminance at the sensor, and can reduce
the dynamic range as a result of diffraction, aberration, defocus and flare [McCann and Rizzi, 2007]. The sensor can have limiting factors to dynamic range as well, including saturation on the intensity high end, and noise limitations on the low end [Reinhard et al., 2010]. The camera data processing workflow begins with the color transformation (for single sensor systems), generally utilizing a Bayer color pattern followed by a luminance estimate as a weighted sum of the RGB values. Once the data is recorded, further processing may be required based on proprietary codecs, log encoding, or other in-camera data manipulation. Care is taken to post-process the collected data into a standard format for analysis. While the effects on dynamic range of each of the three main components in the image capture pipeline can be analyzed in detail, a first step is to perform a system calibration where the effects of optics, sensor, and processing are lumped into the system definition. The measurement and calculation therefore includes the combined effects of the optics, sensor, and camera processing as illustrated in the system level model of luminance conversion in figure 5.1. In the next chapter, the individual contributions of the optics, sensor, and processing blocks will be investigated in detail.

Figure 5.1: System Level Model of Luminance Conversion

5.1.2 Materials

Test Chart and Processing Software

A DSC Labs Xyla 21 test chart was procured that includes 20 stops of dynamic range via a 21 step rear lit voltage regulated light source. The Xyla chart, shown in figure 5.2, features a stepped xylophone shape, minimizing flare interference from brighter steps to darker steps. The Xyla chart includes an optional shutter system that allows for the isolation of individual steps in order to reduce the effect of stray light. The use of the shutter system comes at the expense of increased measurement and processing time, as each individual step must be imaged and processed as opposed to the chart as a whole. The shutter system was not utilized in these experiments.
due to time constraints, however the evaluation and characterization of the effect of stray light from the chart itself on the noise measurement is important future work. The Xyla chart is preferred to front lit reflective charts as front lit charts are more difficult to relight evenly over time, as test configurations change or the test environment is altered. Rear lit grayscale stand-alone films require the use of an illuminator such as a light box, where special care is required to monitor the light source and voltage. The Xyla 21, being an all enclosed calibrated single unit, simplifies test setup and measurement.

The Xyla-21 chart includes $\log_2$ spaced steps, where each linear density step of 0.3 equates to one stop, a doubling (or halving) of luminance. The calibrated density data for the Xyla-21 chart used is provided in Table 5.1. The calibration is provided by the manufacturer, and the recommended upgrade (re-calibration) date is printed on the chart.

The Stepchart module of the image analysis software Imatest [Imatest, 2014] is utilized to process the captured images of the Xyla-21 test chart. The camera distance to the chart is adjusted to maintain approximately 50 pixels per patch horizontal resolution, required by Stepchart [Koren, 2014]. The Stepchart module includes a Region of Interest (ROI) selection tool that allows for the alignment and selection of ROI for zones as illustrated in figure 5.3. The light grey boxes are the selection region for each individual patch, with software controls allowing
Table 5.1: Xyla-21 Density Calibration Data

<table>
<thead>
<tr>
<th>Step #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0</td>
<td>0.3</td>
<td>0.6</td>
<td>0.9</td>
<td>1.2</td>
<td>1.51</td>
<td>1.81</td>
<td>2.11</td>
<td>2.41</td>
<td>2.71</td>
<td>3.01</td>
</tr>
<tr>
<td>Step #</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td>3.31</td>
<td>3.61</td>
<td>3.91</td>
<td>4.21</td>
<td>4.52</td>
<td>4.82</td>
<td>5.12</td>
<td>5.42</td>
<td>5.72</td>
<td>6.02</td>
<td></td>
</tr>
</tbody>
</table>

adjustment of all patches contained within the red outer box. The right image in figure 5.3 shows the selection tool with zoom and lighten functions turned on, which can be especially helpful when aligning darker zones. A useful alignment hint is to first align the ROIs based on a brightly lit captured frame of the test chart. Imatest will recall this alignment as the default for future images, therefore as long as the camera and chart are not moved, the brightly lit alignment can be used for test frames where the ambient lights have been switched off and darker zones are more difficult to align to. With zone regions selected, Imatest software calculates statistics for each ROI including the average pixel level and variations via a second-order polynomial fit to the pixel levels inside each ROI. The second-order polynomial fit is subtracted from the pixel levels that will be used for calculating noise (standard deviation), removing the effects of non-uniform illumination appearing as gradual variations.

Figure 5.3: Imatest Stepchart Module Region of Interest Selection Screens
Selected Cameras

Camera selection was partially based on availability, while best effort was made to include units from various “market” segments. The RED EPIC and ARRI Alexa represent high end, primarily entertainment industry cameras, both of which claim extended dynamic range. The RED ONE is a very popular entertainment industry camera that continues to find wide use and is available in M and MX models, dependent on the version of the image sensor. The Toshiba and Hitachi cameras represent machine vision cameras that are often used in scientific and engineering applications. The Canon 5D Mark 3 is a latest generation DSLR, and was tested utilizing the manufacturer provided embedded software, as well as open source embedded software known as Magic Lantern [MagicLantern, 2014]. Finally, one of the many new entries into the lower cost “professional” 4K market cameras, the Black Magic Cinema 4K, was also selected.

5.2 Test Procedure

The CUT is mounted on a tripod and placed in a dark test room, kept at an ambient temperature of $23^\circ \pm 2^\circ$C, with no light leakage along with the Xyla test chart. Camera lights and all reflective surfaces are masked to prevent reflections off the chart surface. A 25mm Prime PL mount lens is attached to the CUT utilizing PL type mounts. For large format sensors, in order to maintain the 50 pixels per patch horizontal resolution, a longer focal length lens may be required such as an 85mm for the RED Dragon 6K. Cameras with non-PL mounts (such as C, Canon EF) utilize available compatible lenses.

A one foot straight edge is placed across the front plane of the lens. Measurements are taken at each end of the straight edge to the face of the test chart to ensure the two planes are parallel to an eighth of an inch. The lens aperture is set to full open, and focus is set by temporarily placing a Siemens star on the face of the test chart. Once focus has been verified, the camera is set to the manufacturer recommended native ISO, and data collection begins. The ambient room lights are turned off and the door to the test room is closed so that only the light from the test chart reaches the camera under test. The lens remains at the lowest stop value (wide open) to ensure the luminance of at least the first 2-3 chips will saturate the image sensor at the slowest exposure time (generally $\frac{1}{24}$th or $\frac{1}{30}$th of a second). A few seconds of video are collected, and then the exposure time is reduced by one half (one stop). Again a few seconds of video are collected, and the process is repeated for a total of at least 5 measurements to complete the sample collection process.
Note that only a short video sequence is required, as only a single frame is used for the simultaneous calculation. Imatest does include a sequential (temporal) noise calculation, where additional frames could be utilized in the difference calculation, but the investigation of this will be future work.

Measured frames are processed through Imatest that can be configured to produce output figures relating to exposure, noise, density response, and/or SNR. A graph of normalized exposure pixel level (normalized to 0-255) is one output that is plotted against the step value of the Xyla21 chart. The original image data, whether 8 or 16-bit pixels, is first converted to quad precision (floating point) and normalized to 1 before any processing is performed. The normalization to 255 is only for the purposes of the output graphical display (historically related to 8-bit file depth and the familiar 255 maximum pixel level). Therefore it is important to note that the full precision of the input file data is utilized in the calculations, independent of the normalized output scale. Sample normalized exposure data for the ARRI Alexa is shown in figure 5.4, noting that 7 exposures were collected during this test. The collection of multiple measurements is to ensure an exposure was captured meeting the criteria of the maximum unclipped input luminance level $L_{sat}$ defined in Imatest as 98% of its maximum value.

![Figure 5.4: Normalized Exposure Data for ARRI Alexa](image)

Dynamic range is calculated by Imatest using the minimum luminance step
meeting the maximum RMS noise for each exposure time. This is shown in figure 5.5 for the ARRI Alexa. We can observe that the exposure setting that resulted in the peak dynamic range was the one where the brightest chip was just saturated. In the case of the ARRI Alexa data, this was the exposure time setting of $\frac{1}{190}$ of a second, resulting in 13.9 stops for a maximum 0.5 stop RMS noise utilized in defining the minimum luminance step.

5.2.1 Data Workflow

A significant challenge in testing a large sample of digital imagers is developing data workflows. For as many vendors utilized, there are just as many proprietary
data formats, encoding algorithms, export formats, software packages, and post-
processing guidelines. How captured digital data is provided to the user is also
manufacturer dependent. For example, ARRI utilizes a Dual Gain Architecture
(DGA) with two paths of different amplification for each pixel [ARRI, 2013a]. ARRI
processes the dual path data internal to the camera, providing a single output file
available either as RAW (with external recorder), LogC, or Rec709. ARRI file
formats typically use ARRI LUTs for gamma and color space, but are wrapped in
standard ProRes422 or ProRes444 wrappers. Generation and application of ARRI
specific LUTs for the different outputs is straightforward, resulting in a user friendly
workflow [ARRI, 2013b].

In some cases, file wrappers, encoding schemes, gamma, and color space
functions are vendor specific such as with RED Redcode files, but have wide accep-
tance and are importable and manageable either through vendor provided software
or popular third party applications (RED, 2013). RED provides its HDR digital
data as two separate file sequences comprising of A and X frames. Access to both
exposure frames has the benefit of increased flexibility in post-processing, at the
cost of increased complexity in understanding the tone-mapping process. The RED

Figure 5.6: Normalized Exposure Data for RED EPIC A and X Frame
software CineXPro includes two tone-mapping functions, Simple Blend and Magic Motion, to assist the user in the tone-mapping process [RED, 2014]. Manufacturer specific post-processing workflow and applications used in the creation of sample frames for analysis are summarized in Table 5.2.

Table 5.2: Post-Processing Workflow Summary

<table>
<thead>
<tr>
<th>Imager</th>
<th>Capture Format</th>
<th>Processing Steps</th>
</tr>
</thead>
</table>
| ARRI Alexa       | 12 Bit LogC, ProRes444, 1920×1080      | LogC - Export 16 Bit TIFF from DavR  
|                  |                                       | Rec709 - Apply 3D LUT, export 16 Bit TIFF from DaVR                           |
| RED One M/MX     | 12 Bit Linear, Redcode36, 4096×2034    | Log/LogFilm Gamma - Apply using CineX-Pro, export 16 Bit TIFF                   |
|                  |                                       | Gamma3/Color3 - Apply using CineX-Pro, export 16 Bit TIFF                       |
| RED EPIC/dynamic | 12 Bit Linear, Redcode 8:1, EPIC 4K 16:9, dynamic DRAGON 6K 16:9 | LogFilm Gamma - Apply using CineX-Pro, export 16 Bit TIFF.  
| DRAGON           |                                       | Apply Simple Blend, Magic Motion tone-mapping as required.                      |
| Toshiba HD       | Uncompressed YUV, Blackmagic Hyperdeck Pro Recorder, MOV wrapper | Rec709 gamma applied by camera. Export 16 Bit TIFF from DavR                   |
| Hitachi DK       |                                       |                                                                                   |
| BlackMagic       | Film Mode, Apple ProRes 422HQ, MOV wrapper, 3840×2160 | Film gamma applied by camera. Export 16 Bit TIFF from DavR                     |
| Canon 5DM3       | ML RAW, Full Frame                     | Convert ML RAW using RAW2DNG. Export 16 bit TIFF DavR                          |
| Magic Lantern RAW Still | ML RAW, Full Frame |                                                                                   |
| Canon 5DM3       | ML H.264, 1920×1080                    | Convert H.264 HDR file using AVISynth to HDR JPEG.                              |
| Magic Lantern HDR Movie | ML H.264, 1920×1080 |                                                                                   |

5.3 Results

Results of imagers tested are shown in Table 3, including the native ISO values used and the vendor specification when available. The dynamic range measurements are stated in terms of the minimum luminance step meeting the maximum RMS noise, with high, medium-high, medium, and low quality corresponding to 0.1, 0.25, 0.5 and 1.0 stops of RMS noise respectively. The RED EPIC and dynamic DRAGON total measurements are approximate, as it includes the combination of the two measured A and X frames. Data for the RED tone-mapping functions Simple Blend
and Magic Motion have been included, as well as measurements for the RED EPIC and dynamic DRAGON without utilizing the HDR function.

In the Table 5.3 results, we attempted to select a “most appropriate” single quality term, and accompanying maximum RMS noise defining the minimum luminance step, that correlates with published data. The table has been arranged from highest to lowest dynamic range using the Imatest designated “medium” quality. Sorting of the data based on the medium quality term is made after comparison to available vendor specifications provide where available. The high and medium-high quality results appear to be too stringent to be utilized by most vendors, especially when marketing for typical applications. Medium quality has relatively close agreement when considering high end cameras such as the ARRI Alexa (measured 13.9 stops, specified 14 stops) and industrial cameras such as Hitachi DK-H100 (measured 10.6 stops, specified 10.0+ stops) and DK-Z50 (measured 9.3 stops, specified 9.6 stops). The case could be made for utilizing the low quality term, as measured data does have close agreement to vendor specifications in the case of the RED ONE M Log (measured 11.7 stops, specified 11.3 stops), RED ONE MX Log (measured 12.7 stops, specified 13.0+ stops), and Black Magic 4K camera (measured 10.0 stops, specified 12.0 stops). This option seems less than ideal however, as higher noise levels in dark regions is generally undesirable and the low quality term includes a full stop of RMS noise.

5.4 Quality Descriptors

An evaluation of a no-reference objective quality metric to accurately measure and compare the increased dynamic range capability of modern image capture systems is presented. The method utilizes a 21 step, back lit test chart providing luminance patches over a 20 stop range. Data collection, processing, and analysis steps are described as part of an image capture and post-processing workflow. The workflow requires the use of vendor software, commercially available post-processing software, or in some cases both. Sample data is presented for a number of current generation imaging systems, some of which are extending dynamic range capability beyond traditional digital imaging capabilities.

The image systems tested illustrate the different technologies and implementations available in the market today. In some cases, vendors utilize a single sensor with single output, relying on the intrinsic capability of the sensor for the maximum dynamic range. In other cases, such as the ARRI Alexa, dual gain sensors are implemented with the vendor combining, or tone-mapping the dual data sets.
Table 5.3: Dynamic Range Measurement of Digital Imagers

<table>
<thead>
<tr>
<th>Imager Mode, Gamma</th>
<th>Quality Term</th>
<th>Stops @ 0.1 RMS Noise</th>
<th>Stops @ 0.25 RMS Noise</th>
<th>Stops @ 0.5 RMS Noise</th>
<th>Stops @ 1.0 RMS Noise</th>
<th>Native ISO</th>
<th>vendor Spec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED DRAGON - HDR x6, Log Film, Total</td>
<td>High</td>
<td>7.6 + ~3</td>
<td>10.2 + ~3</td>
<td>11.8 + ~3</td>
<td>13.1 + ~3</td>
<td>250</td>
<td>16.5+</td>
</tr>
<tr>
<td>RED EPIC - HDR x6, Log Film, Total</td>
<td>Medium-High</td>
<td>7.6 + ~5</td>
<td>9.2 + ~5</td>
<td>10.7 + ~5</td>
<td>12.0 + ~5</td>
<td>800</td>
<td>&lt;18</td>
</tr>
<tr>
<td>RED EPIC - HDR x6, Log Film, Simple Blend</td>
<td>Medium</td>
<td>11.6</td>
<td>13.8</td>
<td>15.0</td>
<td>16.1</td>
<td>800</td>
<td>N/A</td>
</tr>
<tr>
<td>RED DRAGON - HDR Off, Log Film</td>
<td>Low</td>
<td>10.2</td>
<td>12.6</td>
<td>13.9</td>
<td>14.8</td>
<td>800</td>
<td>14.0</td>
</tr>
<tr>
<td>RED EPIC - HDR x6, Log Film, Magic Motion</td>
<td>10.0</td>
<td>11.7</td>
<td>13.1</td>
<td>14.2</td>
<td>800</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RED dynamic DRAGON - HDR Off, Log Film</td>
<td>8.4</td>
<td>11.1</td>
<td>12.4</td>
<td>13.6</td>
<td>250</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>ARI Alexa - Log</td>
<td>8.2</td>
<td>10.7</td>
<td>11.8</td>
<td>12.7</td>
<td>800</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RED EPIC - HDR Off, Log Film</td>
<td>8.6</td>
<td>10.6</td>
<td>11.5</td>
<td>12.8</td>
<td>800</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>RED ONE MX - Log Film</td>
<td>8.8</td>
<td>10.4</td>
<td>11.5</td>
<td>12.7</td>
<td>800</td>
<td>13.0+</td>
<td></td>
</tr>
<tr>
<td>RED ONE MX - Gamma 3</td>
<td>8.5</td>
<td>10.4</td>
<td>11.4</td>
<td>12.2</td>
<td>800</td>
<td>13.0+</td>
<td></td>
</tr>
<tr>
<td>Canon 5DM3 - Magic Lantern H.264, HDR DUAL ISO 400/1600</td>
<td>8.5</td>
<td>10.5</td>
<td>11.4</td>
<td>11.9</td>
<td>100</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>Toshiba IK-HR1S HD - Gamma 0</td>
<td>8.5</td>
<td>9.5</td>
<td>11.1</td>
<td>12.0</td>
<td>0dB</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Canon 5DM3 - RAW Still</td>
<td>7.6</td>
<td>8.8</td>
<td>10.6</td>
<td>11.6</td>
<td>100</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Hitachi DK-H100 - Gamma On</td>
<td>3.9</td>
<td>8.3</td>
<td>10.6</td>
<td>11.2</td>
<td>-3**</td>
<td>10.0+</td>
<td></td>
</tr>
<tr>
<td>RED ONE M - Log</td>
<td>7.7</td>
<td>8.9</td>
<td>10.4</td>
<td>11.7</td>
<td>320</td>
<td>11.3</td>
<td></td>
</tr>
<tr>
<td>RED ONE M - Gamma 3</td>
<td>7.8</td>
<td>9.3</td>
<td>9.9</td>
<td>10.6</td>
<td>320</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Hitachi DK-Z50 - Gamma On</td>
<td>5.3</td>
<td>7.2</td>
<td>9.3</td>
<td>10.1</td>
<td>-3**</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>Black Magic Cinema 4K - Film</td>
<td>6.0</td>
<td>7.8</td>
<td>9.0</td>
<td>10.0</td>
<td>200</td>
<td>12.0</td>
<td></td>
</tr>
</tbody>
</table>

**Hitachi cameras do not specify ISO values for digital gain, instead an arbitrary gain of -3, 0, or +3 is available.

into a single output, transparent to the end user. Furthermore, with devices such as the RED EPIC and dynamic DRAGON multiple exposures are taken and stored
as separate data fields, leaving the tone-mapping as an additional step in the post-processing workflow. An observation that can be made from testing these various images systems is that when considering dynamic range, it is equally important to consider the post-processing workflow and how it will be conducted. Manufacturers also offer various gamma functions, which when utilized result in changes in the final dynamic range.

In real world scenes, quality depends on scene content such as edges and gradients as well as target contrast of a signal with respect to the background. Ultimately, we can state that arbitrary quality descriptors are best described as scene dependent. What is deemed “excellent” for one application, or scene, may be “acceptable” for another. Producers of imagery for entertainment may not only allow, but welcome more low light noise to achieve a film like appearance or “look”. Other users such as in the medical, engineering, or scientific fields, may alternatively require increased signal to noise performance. In other cases, the human eye has been shown to detect well-defined targets when the SNR is less than 0.1 [Holst and Lomheim, 2007]. Manufacturers often add to the confusion by publishing dynamic range or minimum signal level without reference to test conditions or SNR, or without stating if the data is presented for the ideal sensor case, making comparisons difficult. We can state the obvious regarding the definition and application of arbitrary descriptive terms, that they are in fact arbitrary, but can have value and can be compared when they are characterized with respect to a meaningful metric such as the SNR of the minimum luminance level used to determine dynamic range. The medium quality level pertaining to a 0.5 RMS noise was chosen as a compromise between results that correlate with vendor supplied data, and a noise level that is undesirable. Medium quality was selected for the general case, however it is open for interpretation for specific use cases. Future work will be conducted in correlating RMS noise based quality evaluation with subjective evaluations.

The results indicate that not all vendor data correlates with a specific maximum RMS noise defining the minimum luminance step. This may be due to differences in test methodologies, to acceptance of different noise levels, or both. Also of note is the higher end ARRI and RED imaging systems resulted in the greatest dynamic range, however, the RED EPIC and dynamic DRAGON results for HDR mode are estimates based on the separate exposure frames. The final realized dynamic range of the RED HDR systems will be a factor of the tone-mapping method employed.
5.5 Summary

As imaging systems continue to develop and evolve, including recent growth in HDR, calibrated measurements of dynamic range that are both transparent and comparable are of increased importance. A first step is to perform a system calibration where the effects of optics, sensor, and processing are lumped as part of the system definition. The system calibration utilizes a commercially available DSC Labs Xyla 21 test chart including 20 stops of dynamic range via a 21 step rear lit voltage regulated light source. The Xyla-21 chart features a stepped xylophone shape to minimize flare interference from brighter steps to darker steps. Additionally, the Stepchart module of the image analysis software Imatest is utilized to process the captured images of the Xyla-21 test chart and determine dynamic range.

Manufacturer specification are often without reference or descriptive narrative as to the computational method. The use of arbitrary quality descriptors may indeed be arbitrary, as what is deemed “excellent” for one application, or scene, may be “acceptable” for another. Value can be added to dynamic range measurement and quality descriptors, and they can be better compared, when characterized with respect to a meaningful metric such as the maximum RMS noise defining the minimum luminance step used to determine dynamic range. A SNR of 20 dB (0.1 stop of RMS noise) was identified as a reference point having commonality to both ISO Standard 12232-2006 and the test software Imatest. Review of measured results for several modern imaging systems indicate that the industry trend in dynamic range reporting correlates closer with the “medium” quality descriptor in Imatest corresponding to a SNR of 6dB (0.5 stop of RMS noise when defining the minimum luminance patch) than with the 20 dB metric termed “acceptable” quality in the ISO Standard 12232-2006 and “high” quality in Imatest.

The imagers tested have been treated as “systems”, with the combined effect of the optics, the sensor hardware, and the camera processing. An assumption is made that lumping the individual effects of the optics, sensor hardware, and camera processing are not introducing systematic error into the method. In the next chapter this assumption is challenged by undertaking a much more detailed approach; performing a radiometric calibration while specifically studying the effects of the optics on dynamic range.
Chapter 6

Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model

Figure 6.1: Example exposures illustrating overall dynamic range of an oxygen tank over-pressurization event

6.1 Introduction

The methods presented in this chapter carefully categorize error via methodical calibration; including determination of sensor noise attributes, system gain response, and optical effects, such that the linear dynamic range of the imaging system can be characterized to a greater extent than previous methods.

For visual computing cases where realism is a must, for example to provide robust visibility in autonomous vehicles, or when making physically accurate light measurement where absolute precision is required, carefully calibrated HDR is fundamentally important. Additional applications for calibrated HDR imagery include...
gas analysis [Arens et al., 2014], rocket plume visualization [Karr et al., 2016; Ashok and Wray, 1995; Madzsar et al., 1992], lighting analysis [Larson and Shakespeare, 1998; Reinhart and Herkel, 2000], and image processing [Mantiuk et al., 2006b; Akyz and Reinhard, 2007]. Absolute calibration and determination of true light levels can also assist in the shaping of data requirements for imaging equipment. As an example, figure 6.1 includes different exposures of a frame containing a dynamic scene from a test flight of an experimental NASA launch and landing test vehicle developed under the project name Morpheus [Olansen, 2014]. While the Morpheus vehicle ultimately resulted in 13 successful free flights, during an initial test flight the vehicle crashed, resulting in the oxygen tank over-pressurization and explosion shown in figure 6.1. The capture device, an ARRI Alexa, is a professional level digital camera with reported dynamic range capability of 14 stops [Karr et al., 2017]. Figure 6.1(a) illustrates the sample Alexa frame, post-processed for display on a typical LDR monitor. With the exposure optimized per the pre-explosion average luminance level of the scene, the oxygen cloud is saturated with mostly white pixels. The raw data however contains additional luminance detail in the oxygen cloud, and when re-processed, shifting the displayed exposure range towards the highlights, the oxygen cloud detail is then observed as shown in figure 6.1(b). In the third representation of the frame, figure 6.1(c), the exposure range is further shifted, such that detail in the brightest area of the oxygen cloud can be seen.

What is not immediately evident from the images is how much of the available dynamic range of the imager is utilized? If the scene were initially darker, would an imager with even greater dynamic range be required? These questions lead to key insights, including for budgetary considerations, such as how much dynamic range is truly required, at what sensitivity, and for what additional cost? We develop a calibrated HDR radiance map to attempt to answer these questions, including methodical linearization of captured image data, while characterizing the limitations due to the effects of optical glare.

In this visual computing application, focus is primarily on a linear workflow, including the removal of non-linear contributions of the imaging system to obtain linear data for further processing. On the contrary, HDR in commercial, broadcast, and cinema markets focuses heavily on log based workflows where HDR imagery is tone-mapped for display for humans within the limits of the available hardware and software [DiCarlo and Wandell, 2000]. The workflows are not mutually exclusive. The methods presented in this paper carefully categorize error in a multiple exposure HDR recombination workflow via methodical calibration; including determination of sensor noise attributes, system gain response, and optical effects, such that the
linear dynamic range of the imaging system can be characterized to a greater extent than previous methods.

The two primary steps of the multiple exposure recombination technique are (1) linearization of the data of each exposure by applying the inverse Camera Response Function (CRF), and (2) the generation of the HDR luminance map from a weighted average of scaled pixel values of the multiple exposures. Weighting schemes include those that are proportional to the Digital Number (DN) [Mann and Picard, 1995], those that are proportional to the CRF [Mitsunaga and Nayar, 1999], and noise based models [Tsin et al., 2001]. Recent reviews of weighting schemes can be found in Granados et al. [Granados et al., 2010] and Griffiths et al. [Griffiths and Wicks, 2016]. Noise based approaches consider the individual noise sources of the acquisition process, giving higher weights to values containing lower variances. One limiting factor to obtaining good dynamic range is noise in the dark areas of the captured raw digital image. This can be optimized both in the selection of the weighting factor [Granados et al., 2010; Griffiths and Wicks, 2016; Hasinoff et al., 2010], as well as in the capture strategy [Hasinoff et al., 2010, 2016]. Properly accounting for the noise sources results in a calibrated linear camera model that minimizes uncertainty while producing more accurate results; an important factor if the final HDR image is to be used as a physically accurate irradiance representation.

Optical effects that influence image capture include vignette, shading, interference fringing, lens flare, and veiling glare. Vignette, shading, and interference fringing are classified as Fixed Pattern Noise (FPN) sources, which can be removed by the process of flat fielding [Janesick, 2007]. Lens flare and veiling glare are scene dependent effects that can reduce contrast visibility due to light scattering, and set a physical limit on the luminance range that can be accurately measured [McCann and Rizzi, 2007]. Lens flare manifests as sharp image artifacts resulting from hard edges, internal to the lens, such as the lens aperture and leaves [Talvala et al., 2007]. At further distances lens flare becomes blurred, and the effects of stray light from lens surfaces, camera walls, and the sensor surface itself combine as veiling glare. Examples of veiling glare include internal multiple reflections between lens surfaces, scatter from lens elements due to imperfections or optical cements, and reflections from lens edges, mounts, and shutter blades [ISO, 2013].

The main contribution of this chapter is the detailed evaluation and measurement of optical based glare on a calibrated multiple exposure LDR to HDR pipeline, illustrating substantial improvements over non-calibrated workflows of commercially available software. Other contributions of this chapter are:

- An HDR workflow for processing raw linear camera data from a Bayer based
sensor, using floating point data types throughout to maintain precision.

- Absolute calibration of an HDR luminance map.
- Characterization of optical based glare and the resulting impact on dynamic range.

6.2 Calibrating a Multiple Exposure HDR Workflow

In this section we introduce our overall method of generating a calibrated HDR image from multiple exposures (illustrated in figure 6.6), and describe the new steps taken to extend Granados’ noise based recombination method to achieve a calibrated HDR radiance map. The availability of linear data from the captured image file is device dependent, generally requiring characterization of in-camera processing and optical effects to produce a CRF. The inverse CRF can then be applied, linearizing the data. Determination of the CRF, and estimation of the physically accurate scene irradiance, is effectively the calculation of the photometric mean irradiance for each pixel. The nonlinear camera response function $f$ is determined via radiometric calibration that includes the nonlinear contributions of the camera capture system such that $f = g^{-1}$, where $g$ is the inverse function mapping scaled image irradiance to pixel values. Although previous efforts to produce a calibrated radiance map [Mitsunaga and Nayar, 1999; Debevec and Malik, 2008; Inanici and Galvin, 2004; Coutelier and Oumortier, 2003] have included calibrations for many of the system non-linearity’s and temporal/spatial noise, inverse functions have not yet accounted for glare. Further, due to both the scene and lighting specific spatial nature of glare, the effect will constantly change. Our goal is to perform a first step of characterizing the glare effect for a static scene. The calibration steps, to be described below, include:

- Use of the PTC to determine the noise based weighted average performance parameters: read noise, saturation, and camera gain.
- Determination of the per-pixel gain factors by flat field measurement and background subtraction, implementing an integrating sphere for the measurements.
- Radiometric scene calibration, and effect of glare in a static scene, by measurement and comparison to absolute luminance spot meter readings.
6.2.1 Bayer Raw (BR) Data

BR data is defined as the Digital Count (DC) linear greyscale image produced directly from the Analog to Digital Converter (ADC) of a Bayer based imaging sensor. BR captured DCs are typically a linear function of the incoming irradiance throughout at least a portion of the sensor dynamic range. The BR data is usually not directly available from the camera. The “raw” data available from the camera is defined as the Camera Raw (CR) data, and may include a non-linear transformation as well as other processing. The CR file may include pixel counts that are larger than the reported resolution of the camera, as all pixels of the sensor are included, not just those typically reported as the “active” area of the sensor. The Colour Filter Array (CFA) phase for Bayer or similar based imager is first determined (ex. RGGB, see figure 6.2), then the individual RGGB channels are separated into 4 similarly sized arrays containing the individual Bayer RGGB channels. The image data is now at a point where it is considered to be BR, as long as no addition in camera processing has been performed. The data is without dark frame subtraction, demosaicing, white balance, non-linear processing, or compression. We use the term “non-linear processing” in a general sense that minimally may include colour-space conversion, brightness scaling, gamma correction, saturation adjustment, and/or application of additional non-linear transformations such as log or film-like S-response curves. Processing is generally device dependent, potentially applied based on imager functionality.

![Example NxM 'RGGB' Bayer Pattern](image)

Red = R11 + ... + R31 + R33 + ...
Green 1 = G12 + G14 + ... + G32 + G34 + ...
Green 2 = G21 + G23 + ... + G41 + G43 + ...
Blue = B22 + B24 + ... + B42 + B44 + ...

Figure 6.2: Bayer pattern example for RGGB. Each 22 grouping contains twice as many green pixels as red and blue pixels, corresponding with the increased sensitivity to green of human vision.
6.2.2 Performance Parameters via PTC

To characterize error for the benefit of determining the ideal low noise pixels to be used in Granados’s multiple exposure recombination process, we implement photon transfer method as a means of determining sensor performance parameters. PTC’s are useful in the determination of camera system response to a uniform light source without optics. The performance parameters required for our noise model, including read noise mean and variance, charge capacity, and incident photon to ADC sensitivity (camera gain), can be determined from the PTC. The PTC is measured using an integrating sphere setup following the EMVA Standard 1288 [EMVA, 2010], as shown in figure 6.3.

![Camera under test and integrating sphere](image)

Figure 6.3: Camera under test (Canon 5D Mark III) and the integrating sphere. To ensure each pixel receives light from the whole disk-shaped illumination source (the integrating sphere output port), an f-number of 8 is required. By placing the sensor at a distance $D$ of 81.28 cm from the output port having a diameter $d$ of 10.16 cm, an f-number of 8 is achieved, where $f_\# = d/D$.

The PTC contains four identifiable regions dominated by a particular noise type, as described by Janesick [Janesick, 2007], and detailed further in this section. An example for the Canon 5DM3 is shown in figure 6.4. The first region, read noise, is the random noise measured under dark conditions, generally having a slope of 0. Read noise includes a combination of pixel source follower noise, sense node reset noise, thermal dark current noise, ADC quantization noise, offset, and system noise. Read noise is overcome by photon shot noise, which on a log-log plot results in an
approximate slope of $\frac{1}{2}$. Photon shot noise is defined as the standard deviation of the number of photon interactions per pixel, which is a spatially and temporally random phenomenon. Additionally, while photons generate electron-hole pairs via the photoelectric effect, particles other than photons such as high energy electrons, protons, and ions can also produce signal carriers. Following shot noise is the region dominated by Fixed Pattern Noise (FPN). FPN has an approximate slope of 1 indicating that signal and FPN scale together. Finally, the last region is defined as the point where the pixels enter full-well, or saturation. Full-well is defined as a rapid noise deviation from a slope of $\frac{1}{2}$ or 1. From the PTC, required read noise and fixed pattern noise values are obtained.

Figure 6.4: Photon Transfer Curve, illustrating the different noise regions identified as part of the overall total noise curve. Regions are defined by slopes.

6.2.3 Relationships of Noise Parameters

The multiple exposure recombination method to be utilized is based on determining the pixels with the lowest noise. For completeness, this section provides additional detail regarding the relationship between noise performance parameters. Each pixel
has an associated ADC offset that needs to be removed. The offset can be determined as an average of all pixels, or per pixel for increased accuracy.

For each colour channel, the true average signal level \( S_{AVG}(DN) \) or mean for all sub-array pixels \( N_{PIX} \) at each exposure is given by:

\[
S_{AVG}(DN) = \frac{\sum_{i=1}^{N_{PIX}} DN_{ADC,i}}{N_{PIX}} - S_{ADC,OFF}(DN),
\]

where \( DN_{ADC,i} \) is the \( i \)th raw signal value, and \( S_{ADC,OFF}(DN) \) is the ADC offset level. The ADC offset level can be obtained either from the Exif data (metadata), noting that the Exif contains a single value averaged over all pixels, or may be determined per pixel using the average of \( NF \) dark frames. If the Exif data is utilized, the single stored value is used for all Bayer colour assignments, as it is a noise measurement independent of incident photons (i.e. a dark frame). Alternatively, if the offset is experimentally measured per pixel for pixel level accuracy, as was the case in our implementation, the following equation is used:

\[
OFF_i(DN) = \frac{\sum_{j=1}^{NF} off_j(DN)}{NF},
\]

where \( OFF_i(DN) \) is the average offset per pixel \( i \) for \( j \) dark frames \( off_j(DN) \), and \( NF \) is the number of averaged frames. This per pixel offset will be subtracted from the \( i \)th raw signal value, but first temporal effects need to be addressed.

To avoid temporal effects and transient phenomena, two images \( A \) and \( B \) are captured and averaged at each exposure level for both the raw signal value (grey measurement) as well as the ADC offset level (dark measurement). The average signal level for both raw signal level and dark measurement is determined utilizing corresponding \( A \) and \( B \) images as:

\[
\mu_{AB} = \frac{1}{2N_{PIX}} \sum_{i=1}^{N_{PIX}} (A[i] + B[i]).
\]

Therefore, the two grey measurements are averaged as:

\[
DN_{ADC,AB} = \frac{1}{2N_{PIX}} \sum_{i=1}^{N_{PIX}} (DN_{ADC,A}[i] + DN_{ADC,B}[i]),
\]

95
and the two dark measurements averaged as:

\[
OFF_i(DN)_{AB} = \frac{1}{2N_{PIX}} \sum_{i=1}^{N_{PIX}} (OFF_i(DN)_A[i] + OFF_i(DN)_B[i]).
\]  

(6.5)

Now defined are the averaged grey measurements, the averaged dark measurements, and the per pixel offsets. It can now be stated that the signal value of the \(A\) and \(B\) averaged images \(i\)th raw video pixel, without offset, is:

\[
S_i(DN)_{AB} = DN_{i_{ADC,AB}} - OFF_i(DN)_{AB}.
\]  

(6.6)

Finally, the true average signal level, or mean, with temporal averaging and offset removed is:

\[
S_{AVG}(DN)_{AB} = \frac{\sum_{i=1}^{N_{PIX}} S_i(DN)_{AB}}{N_{PIX}}.
\]  

(6.7)

The total noise \(\sigma_{TOTAL}(DN)\) (including read, shot, and FPN) is the standard deviation (RMS) value:

\[
\sigma_{TOTAL}(DN) = \left[ \frac{\sum_{i=1}^{N_{PIX}} [S_i(DN)_{AB} - S_{AVG}(DN)_{AB}]^2}{N_{PIX}} \right]^{\frac{1}{2}}.
\]  

(6.8)

6.2.4 Read and Shot Noise

The multiple recombination method utilizes read noise mean and variance as parameters in determining the ideal low noise pixels. To determine read and shot noise only, as a starting point FPN is removed via pixel by pixel differentiation of two sequential frames (under similar test conditions) at the same exposure level:

\[
\sigma_{READ+SHOT}(DN) = \left[ \frac{\sum_{i=1}^{N_{PIX}} (DN_{1_{i_{ADC}}} - DN_{2_{i_{ADC}}})^2}{2N_{PIX}} \right]^{\frac{1}{2}},
\]  

(6.9)

where \(DN_{1_{i_{ADC}}}\) and \(DN_{2_{i_{ADC}}}\) are the raw signal values of the \(i\)th raw pixel for image 1 and 2, respectively. Note that the variance of the difference of the two values is the sum of the variances, increasing the random noise by a factor of two. Therefore, the equation includes a division factor of 2. To determine the shot noise
\( \sigma_{SHOT}(DN) \) is the read plus shot noise minus the read noise:

\[
\sigma_{SHOT}(DN) = \left[ \sigma_{READ}(DN)^2 \right]^{1/2} \left[ \sigma_{READ}(DN)^2 \right]^{1/2}.
\]

(6.10)

It is desirable to include as many pixels as possible to increase statistical accuracy, while meeting light uniformity of better than 3% across the pixel sub-array in order to limit error in FPN measurements. Measurement accuracy is proportional to the square root of the number of pixels processed (the pixel sub-array), with the major limiting factor in the sub-array size being the uniformity of the illumination reaching the sensor.

### 6.2.5 Per Pixel Gain Factors

There are several spatially based factors that influence the response of the imager that are used in the multiple exposure recombination process, resulting in differences in pixel to pixel sensitivity. First, under identical light intensities, different pixels can consistently produce different DNs as a result of the manufacturing process of the imager and imperfections such as dust particles interfering with light on the sensor. The inherent differences are known as the PRNU pattern. Secondly, as a result of temperature variation, dark current varies between different pixels resulting in a DCNU pattern (or FPN pattern). DCNU can be corrected by subtracting a dark frame from each image frame. The dark frame is captured with the optics receiving no light intensity, and at the same temperature, exposure time, and ISO setting as the image frame. Finally, the vignette effect of the attached lens will result in variation of the light intensity for individual pixels, with the least effect generally along the optical axis and the greatest effect extending radially towards the corners of the image frame. The aperture is held constant during the measurements to minimize variation in veiling glare, which is also dependent on the lens type and manufacturing process [Tomic et al., 2014].

The combined individual per-pixel gain correction accounting for spatially based noise factors PRNU and DCNU, as well as vignette effect, is referred to as the per pixel gain \( a_j \), defined per Granados as:

\[
a_j = \frac{E[ff_j] - E[b_j]}{\frac{1}{N} \sum_j (E[ff_j] - E[b_j])}
\]

(6.11)

Per pixel gain factors are determined using the flat fielding technique [Janesick, 2001], where bias and temporal noise are minimized via correcting for dark current.
by capturing both flat field $E[ff_j]$ (uniform light field) and background frames $E[b_j]$ (lens covered), and then averaging over 36 frames $\Omega$. The number of averaged frames is selected as a balance between noise reduction and data collection and processing time. Our method illuminated the sensor via the attached lens with spatially uniform light from an integrating sphere. SNR is maximized by setting the illumination such that the captured DNs are close to saturation.

### 6.2.6 Radiometric Scene Calibration

To determine the accuracy of a recombined luminance map, luminance values taken from the recombined image of two test scenes are compared with precision measurements made via a commercially calibrated luminance spot meter. Test scene 1 contains an X-Rite ColorChecker Classic as well as 6 self-printed $3 \times 3$ grey patch charts to be used as measurement test targets throughout the scene as shown in figure 6.5. The top left, middle left, and bottom left data correspond to the $3 \times 3$ patch charts on the left side of the scene, masked in darkness. The centre data corresponds with the middle $3 \times 3$ chart, with the ColorChecker chart located below the centre chart. The top right and bottom right data corresponds to the $3 \times 3$ patch charts to the right, and highest luminance, side of the scene. Absolute luminance measurements of the grey patch targets require high accuracy and sensitivity in order to properly characterize optical effects in low signal to noise regions, in this case a Konica/Minolta LS-150 luminance spot meter (accuracy 2%, sensitivity 0.001 $\text{cd/m}^2$) was used. In selecting the spot meter, the meter with the smallest available acceptance angle, 1/3 degree, was chosen in order to limit glare effects in the meter lens. A black sheet of cardboard containing a hole was experimented with to further limit meter lens glare, however positioning and robustly attaching the sheet to the meter was problematic and reduced repeatability, and therefore was not used in the measurements. The use of a cardboard mask with hole is identified as an area for future work. Luminance measurements were taken from the recombined image using the program Photosphere [Ward, 2010], specifically the calibration and selection functions. Photosphere luminance measurements use the radiance to average luminance conversion factor $k$ of 179 lumens/watt, with luminance $L$ defined in terms of $RGB$ primaries as:

$$L = 179 \times (0.265 \times R + 0.760 \times G + 0.065 \times B).$$

(6.12)
Figure 6.5: HDR test scene 1 contains an X-Rite ColorChecker Classic as well as 6 self-printed 3×3 grey patch charts

### 6.2.7 Glare

An ideal point light source, in focus, should illuminate only a single pixel of an image array. In practice, based on lens quality and the distance of the light source from the lens, light will be contributed to other sensor pixels as described by the Glare Spread Function (GSF) [ISO, 1994]. The GSF manifests as lens flare, resulting from sharp edges internal to the lens, including the lens aperture and leaves. Additionally, at further distances lens flare becomes blurred, and the effects of stray light from lens surfaces, camera walls, and the sensor surface itself combine as veiling glare. The total achievable dynamic range of multiple exposure HDR recombination may be limited by veiling glare, potentially effecting the entire image. This can become more evident in lower luminance measurements [McCann and Rizzi, 2007].
6.3 Implementation and Results

Two controlled test scenes are each lit by a quad-LED based light source (Sylvania LED9MR16/DIM/830/FL35/RP, 3000°K) powered by a regulated DC supply (Delta 1145-1071-ND). Scene 1 contains test charts and items of various specular and diffuse reflection, with the light source configured to light much of the scene, while also flagging a darker region. Scene 2 is configured to specifically test the effects of glare in shadow regions, both near and far from the brighter regions of the scene. A multiple exposure set of 10 images are collected using the Canon 5DM3 plug-in MagicLantern (bracket type: exposure, centre exposure 1/8 second, 2EV step) resulting in exposure times as detailed in the figure 6.10 caption. A Sigma 24-70mm lens set at 24mm, with full open IRIS, was attached to the Canon 5DM3 and supported via tripod. The optical axis is aligned approximately with the centre of the scene.

6.3.1 Method

The recombination of multiple exposures includes eight key image processing steps as shown in figure 6.6, where Blocks 1-3 are single capture processes and Blocks 4-8 describe the recombination process. Figure 6.6, Blocks 1 and 2, illustrate the conversion of the Canon CR output .CR2 file to a .DNG file via the Adobe DNG Converter program. The Adobe converter uses dcraw as the low level converter (preference settings: Compatibility Custom, Linear unchecked, Uncompressed checked) [Coffin, 2017]. Dcraw conversion here primarily converts the .CR2 file to a Matlab readable .DNG file. Since Canon CR data does not apply a non-linear transformation, no inverse transformation is required at this step. The active region is cropped using the Exif data, resizing from 3950×5920 to 3840×5760 pixels, so the resulting resolution matches that produced by commercially available software for later comparison. The individual RGGB channels are then separated into 4 similarly sized arrays containing the individual Bayer RGGB channels, i.e. an 1920×2880×4 array as shown in figure 6.6, block 3.

The next step is the recombination of multiple exposures and recreation of the RGGB Bayer pattern as shown in figure 6.6, Blocks 4 and 5. The determination of the performance parameters required for the noise based optimal weighting function, including read noise mean and variance, charge capacity, and incident photon to ADC sensitivity (camera gain), are determined from the PTC. The FPN is determined from the flat field measurement.
Figure 6.6: Processing Blocks illustrating the 8 key processing steps of the recombination procedure. Blocks 1-3 are single capture process and Blocks 4-8 describe the recombination process.

**PTC of the Canon 5DM3**

The PTC is obtaining using a grey image with exposure time of 0.005 seconds (1/200th) resulting in an approximate mid-exposure without Bayer channels nearing saturation. An iterative approach is used to determine the appropriate image sub-region meeting the uniformity requirement of 3% across the pixel sub-array (per EMVA1288), limiting error in FPN measurements. To assist in deciding the range of pixels having spatial non-uniformity of 3%, a two-dimensional finite impulse response (FIR) smoothing filter is applied to the image, smoothing the variations due to pixel noise. The PTC for the Canon 5DM3 is shown in figure 6.7, with linear slope lines for the Read Noise having a slope of 0, the Shot Noise with slope of $\frac{1}{2}$, and the FPN with a slope of 1. The mean Read Noise is determined from the y-axis intercept of a line with slope 0. The shot noise x-intercept of a line with slope $\frac{1}{2}$ provides values for the incident photon to ADC sensitivities $K_{ADC}(e^{-}/DN)$. Full-well (saturation) values are the same for all 4 Bayer channels. Note that in the PTC curve, the black offset (2047) is removed, resulting in a DN value for full-well of 13,236. The Exif full-well value in the captured files is 15,000, whereas the PTC determined value, including black level, is $13,236 + 2,047 = 15,283$. The camera manufacturer specified a lower full-well value so as to introduce all channels into saturation earlier (15,000 versus 15,283), avoiding potential colour shading or casts of the highlights. An example
of colour shading is the case where the green channel saturates, with the blue and red channel just below saturation, resulting in a pink or purple highlight cast. In our processing, the PTC saturation value of 15,283 was used, so as to maintain the linearity of the original values. PTC results are summarized in Table 6.1.

![Photon Transfer Curve (PTC): Read Noise, Shot Noise, Fixed Pattern Noise (FPN)](image)

Figure 6.7: Photon Transfer Curve including individual contributions of read noise, shot noise, and FPN, and linear best fit lines.

**Flat Fielding**

The per pixel gain correction is determined via the flat fielding process, removing PRNU, DCNU, and vignette contributions. The flat field file for the 5DM3 measurement with Sigma 24-70mm lens is shown in figure 6.8; this is tone-mapped for visualization purposes. The per pixel contribution is observed by comparing a scene capture with and without the correction applied. The percent error graph with and without the correction applied is shown in figure 6.9. The centre 3×3 patches are affected the least by vignette, being close to the optical centre of the lens. The ColorChecker chart vertical grey patches percent error increases with...
Table 6.1: Summary of PTC Determined Values

<table>
<thead>
<tr>
<th>Performance Parameter</th>
<th>Symbol (Units)</th>
<th>Red Ch.</th>
<th>Grn Ch.1</th>
<th>Grn Ch.2</th>
<th>Blue Ch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incident Photon to ADC Sensitivity (Camera Gain)</td>
<td>$K_{ADC_{e^-/DN}}$ $(e^-/DN)$</td>
<td>7.7</td>
<td>5.54</td>
<td>5.01</td>
<td>4.24</td>
</tr>
<tr>
<td>Read Noise Mean</td>
<td>$\sigma_{READ_{electrons}}$ $(DN)$</td>
<td>6.84</td>
<td>6.78</td>
<td>6.84</td>
<td>6.75</td>
</tr>
<tr>
<td>Read Noise Variance</td>
<td>$\sigma_{R^2_{electrons}}$ $(DN)$</td>
<td>46.65</td>
<td>45.64</td>
<td>46.78</td>
<td>45.39</td>
</tr>
<tr>
<td>Charge Capacity</td>
<td>$S_{FW}$ $(DN)$</td>
<td>13236</td>
<td>13236</td>
<td>13236</td>
<td>13236</td>
</tr>
</tbody>
</table>

patch distance from the optical centre. The remaining charts further illustrate the increase in percentage difference at greater distances from the optical centre of the lens. We observe that the per pixel gain based error is significant, primarily a result of lens vignette, as high as 70% near the corners of the field of view. The flat fielding calibration is important not just for radiometric measurements, but also when performing test measurements of dynamic range using dynamic range test charts. Results may vary depending on where the test chart is positioned within the scene, if a flat field correction is not applied.

Recombination Result

Upon combination of multiple exposures using the noise based weighted average, the next step is to white balance the resulting data as shown in figure 6.6, Block 6. As a guide, the documentation for Coffin’s image conversion program dcrab describes different white balance behaviors concerning the highlights, known as the -H option. Modes of operation are according to the value range in which the white balance multipliers will be forced to be calculated, with a unity value indicating no correction. Multipliers less than or equal to 1 will maintain the relative proportions of unsaturated pixels, at the risk of colour artefacts in the areas that were originally saturated. In our implementation, multipliers are held to less than or equal to 1 to maintain the relative proportions. The white balance multipliers are read from the camera Exif information. In the case of the 5DM3 under test, values of 0.4756, 1.0, and 0.6611 were obtained for the colour temperature setting of 3050K (selected to
Figure 6.8: Flat field measurement of Per Pixel Gain Factor for the Canon 5DM3 camera with Sigma 24-70mm lens at 24mm, full open IRIS

Figure 6.9: Luminance map percent difference with diffuse lighting, comparing with and without per pixel gain factor correction

match the light source colour temperature. Demosaicing is performed upon white balancing, as shown in figure 6.6, Block 7. For interpolation, we implement Malvars gradient corrected bilinear interpolated approach, while incorporating floating point values in order to maintain precision [Malvar et al., 2004]. The demosaiced image is then stored in Radiance HDR format [Ward, 1991; Larson, 2005] as shown in figure 6.6, Block 8. The original image sequence single exposures are shown in figure 6.10.
6.3.2 Effect of Glare on the Calibration

Figure 6.11 includes the false color representation of the resulting multiple exposure recombined image of test scene 1. Figure 6.12 shows percent error between the factory calibrated spot meter measurements and recombined HDR image measurements for test scene 1 using a Sigma lens attached to the Canon 5DM3 camera. In figure 6.12, we observe a logarithmic percent error increase for luminance levels falling below approximately 0.25 cd/m$^2$, corresponding with the top left, middle left, and bottom left 3×3 patch charts masked in darkness. The spot meter measurements were taken with a high accuracy Konica/Minolta LS-150 luminance meter, and based on this and the carefully performed HDR recombination effort, the low luminance physical accuracy error was attributed to glare. Published efforts to reduce
glare, such as computational methods involving deconvolution, have not performed well as a result of low signal to noise ratio [Talvala et al., 2007]. A currently recommended approach is to physically reduce glare before capturing the image, often through the use of a higher quality lens, however it should be pointed out that glare free lenses do not exist. The use and comparison of a higher quality lens could provide confirmation that the observed low luminance error occurs as a consequence of glare.

Figure 6.11: Scene 1 multiple exposure recombined false colour image of Canon 5DM3 with Zeiss lens, minimum luminance set to 0.01 cd/m$^2$ and maximum luminance set to $1e+05$ cd/m$^2$ illustrating the linear range of the HDR image.

Influence of Lens Quality

The effect of veiling glare can be illustrated via comparison of two different lens types. An additional data set is captured using a Zeiss 25 mm lens marketed by Zeiss as “exhibiting great flare suppression through T* anti-reflex coatings and internal light traps”. The Zeiss lens has a higher price point (currently US$4200) than the less expensive Sigma 24-70 mm lens ($US900). Figure 6.13 shows the percentage error between the spot meter measurements and the HDR image data with the Zeiss lens. The Zeiss lens results in improved glare performance as compared to the
Sigma lens (figure 6.12), particularly in the low luminance range where there is an overall reduction in error. It is observed that the percent error is now consistently under 10% from the maximum luminance patch of 286.8 cd/m² to approximately 0.1 cd/m².

**Indirect (Diffuse) Lighting**

Another method for demonstrating the non-linear low luminance response resulting from lens glare is to reduce the scene glare due to the light source. The lighting for test scene 1 is adjusted from direct illumination to reflected illumination by alternatively bouncing the light source off the white ceiling of the test environment, effectively creating a diffuse lighting environment. The diffuse lighting environment has the least effect from lens glare, providing a level of isolation to the per pixel gain factors only. Figure 6.14 shows the percentage error between the spot meter...
Figure 6.13: Luminance percent difference for scene 1 grey scale patches with Zeiss 25 mm lens measurements and the HDR image data for the diffuse lighting environment using the Zeiss lens. It is evident that with the diffuse lighting, the peak luminance patch readings are reduced, measuring no more than 16 cd/m². More importantly, without the direct illumination as in the original scene, the percentage difference resulting from lens flare is significantly reduced, remaining under 20% through the lowest luminance levels.

6.3.3 Spatial Nature of Glare

We next construct a new test scene 2 to characterize the spatial nature of glare throughout the capturing device field of view, with dark regions but near and far from the bright region. The original image sequence single exposures are shown in figure 6.15, containing a region of high luminance on the right side of the field of view, along with a dark region on the left side. A flag curtain is used to separate the
sides, and white reflective cards are used to bounce similar levels of diffuse light onto the 4 dark side 3×3 patch targets. The light side 3×3 patches receive direct light from the lighting source. Figure 6.16 contains a false color version of the multiple image capture recombination of test scene 2.

Spot meter measurements are collected for each of the 3×3 patches in the scene. Percentage difference for scene 2 between calibrated spot meter measurements, and values taken from a Canon 5DM3 image utilizing a Sigma lens at 24mm focal length, are included in figure 6.17. The nonlinear error attributed to glare that was observed in test scene 1 for lower luminance levels (figure 6.12), is again observed in test scene 2, although in this scene the nonlinear error begins at approximately 0.8 cd/m² as opposed to 0.25 cd/m² of test scene 1. It is also noted that glare effects are observed on the bottom centre, bottom left, and top centre targets. Both of these points illustrate the scene and spatial dependant nature of glare.
6.4 Discussion

6.4.1 Variability Due to Sensor and Lens

The process of calibrating a camera sensor and lens, via the procedure described in this paper, is both time intensive and expensive in terms of required laboratory equipment. We consider the reuse of camera and lens calibration data, if similar equipment is employed for image capture. Comparisons of the original Canon 5DM3 and Sigma 24-70 mm lens were made with a same model camera, and several lenses, set at the same focal length of 24 mm. The camera/lens comparison data is included in Table 6.2.

The lowest standard deviation was obtained using the original camera with a rented, identical make and model, Sigma 24-70 mm lens. A difference plot is shown
Figure 6.16: Scene 2 with dark regions both near and far from high luminance levels

Figure 6.17: Luminance percent difference for scene 2 with dark regions near and far from high luminance levels

in figure 6.18. The ring in the centre of the difference plot results from the outer radius of a port cover installed in the rear of the integrating sphere, which is used
Table 6.2: Canon 5DM3 #1 with Sigma 24-70 mm #1 Compared to Similar Equipment

<table>
<thead>
<tr>
<th>Camera Body</th>
<th>Lens (@ 24 mm)</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon 5DM3 #1</td>
<td>Sigma 24-70 mm #2</td>
<td>5.47E-04</td>
<td>0.0431</td>
</tr>
<tr>
<td>Canon 5DM3 #2</td>
<td>Sigma 24-70 mm #1</td>
<td>-3.10E-03</td>
<td>0.1009</td>
</tr>
<tr>
<td>Canon 5DM3 #1</td>
<td>Canon 16-35 mm</td>
<td>-5.20E-03</td>
<td>0.0719</td>
</tr>
<tr>
<td>Canon 5DM3 #1</td>
<td>Canon 24-105 mm</td>
<td>-5.80E-03</td>
<td>0.0662</td>
</tr>
</tbody>
</table>

to align the two images. The borrowed 5DM3 camera (Canon 5DM3 #2), using the original Sigma 24-70 mm lens, resulted in the largest standard deviation. The last two entries are for the original camera with two additional, albeit different model, Canon lenses both set at 24 mm.

![Figure 6.18](image1)

Figure 6.18: Canon 5DM3 Camera Body #1: Difference Plot with Sigma 24-70 mm Lens #1 versus Sigma 24-70 mm Lens #2

6.4.2 Variability Due to Image Content

In scene 2 (figure 6.16), glare related error in the radiance map was determined for the bottom centre, bottom left, and top centre targets. In other words, glare may
effect elements of the scene both close and further away from the high luminance areas. Of these, the bottom centre had the highest error, and when looking at the false color image, glare can be observed shaded in green seeming to “spillover” from the flag curtain towards the bottom centre target. We can deduce that any movement or changes in the light position, flag position, or other elements of the scene will result in some spatial glare variability. Predicting the effect, for variations in lighting and scene content, as well as for different optic types, is a challenging future task.

6.4.3 Comparison to Commercially Available Software

To further assess our method, the multiple exposure CR set of 10 images (direct lighting, Zeiss lens) of scene 1 were converted to tiff images in Photoshop, and recombined into an HDR image via the “Make HDR” option in Photosphere. Figure 6.19 includes false colour representations of the HDR images for our method utilizing per pixel gain factor correction (figure 6.19L), and for the Photosphere method (figure 6.19R). Comparing the two images, increasing differences in luminance are observed at radially greater distances from the optical centre of the lens. This is most evident on the outer edges of the scene, coinciding with the flat field measurement illustrated in figure 6.8. Both data sets suffer from the effects of veiling glare in the low luminance regions, however the Photosphere results indicate that the lack of per pixel gain factor correction results in additional error of nearly 60% in the corner regions, when attempting to create a physically accurate radiance map.

Figure 6.19: (L) Our calibrated recombination method (R) Photosphere method
6.5 Summary

This chapter has presented a novel method of minimizing the optical effects on HDR calibration by a careful multiple exposure noise based workflow. In this new approach, the linearization of raw data from a sample camera is examined for the purpose of recombination, and creation, of the absolutely calibrated HDR scene. In processing the CR data, care is taken to account for the manufacturer defined black level and saturation limit while cropping the full sensor data to the active region. The channels of the Bayer CFA are separated and multiple exposures are combined applying an optimal HDR reconstruction based on a rigorous noise model and a noise based weighted average. White balancing is performed while maintaining the relative proportions of the individual colour channels. The resultant demosaiced HDR image is stored as a Radiance HDR image file. The effects of per pixel gain factor and veiling glare are observed, summarized by stating that per pixel gain factor error has greater significance radially outward from the centre of the lens, and glare error has greater significance in the lower luminance regions.

Traditional methods combining multiple exposures into an HDR image often allow for absolute calibration to a known reference measurement, yet offsetting all image values based on a reference measurement does not provide a measure of the linearity of the data throughout the full captured range. In our work, the calibrated methodology identifies the lower limit of the dynamic range that falls within an acceptable absolute luminance error. The scale in the false colour image in Fig. 6 illustrates that for our captured image of scene 1, the shadow area limit is approximately 0.1 cd/m² based on 10% acceptable error, and the peak luminance for a large enough group of pixels considered to be observable is approximately 1e + 04 cd/m². This results in a physically calibrated linear dynamic range of 16.6 stops.
Chapter 7

Conclusions and Future Work

HDR capture forms a fundamental first step in the HDR pipeline. As of yet there is little literature on how best to capture physically accurate linear HDR data without losing vital information. Although there is a wealth of publications related to the HDR pipeline including: capture, processing, tone-mapping, and display, the dynamic range of the test datasets is rarely taken to be physically accurate or truly representative of real world lighting. This presents a problem when working to characterise scene(s) for applications such as visual computing cases where realism is a must, for example to provide robust visibility in autonomous vehicles, or when making physically accurate light measurement where absolute precision is required. Additional applications for calibrated HDR imagery include gas analysis, rocket plume visualization, lighting analysis, and image processing. The main aim of this thesis is to evaluate and calibrate for real world luminance values at every stage of the capture pipeline.

This thesis proposed a novel method for the calibrated capture of HDR imagery, developed a solution for the collection of physically accurate radiance maps, and investigated limitations due to the image sensor, camera processing, and optics. This chapter concludes the work presented in this thesis and details possible directions for future work.

7.1 Contributions

The overarching contribution of this thesis is the development of a solution for the collection of physically accurate HDR radiance maps. In Table 2.1, the main gaps are identified in the current capture and processing pipelines; that linearity is either assumed, approximated, or of less importance than perceived realism. In our work,
the image capture pipeline is analysed in detail, while not assuming linearity, in order to maintain physical accuracy throughout the workflow. Rather than approximate the camera response function, we employ calibrated measurements to characterize both the image sensor and the associated optics. We investigated in detail the optical effects of vignette and glare on the measurements, and presented strategies for calibrating the associated effects.

7.1.1 Analysis of a Real World Rocket Launch Environment

To investigate existing HDR capture technology, a number of latest generation digital imaging systems were selected for testing and evaluation. Field experiments were conducted using rocket launches as targets of opportunity. Our knowledge of existing capabilities was initially limited, therefore a review of current manufacturer technology and documentation served as a valuable resource for gaining initial insight into the HDR capture process. A discussion on the NASA requirements for rocket launch imagery in the Space Shuttle era is also included, providing insight into real world imaging requirements. A review of published vendor specifications was conducted, including capabilities described as “HDR” specific or capable. Each capture system was thoroughly reviewed for setup, operation, and post-processing, per vendor information availability.

Experiments were conducted using dual imagers in a side by side configuration, single imagers with single exposure capture, and single imagers with dual exposure capture. Of the systems tested, the dual camera system with similar optics provided the least satisfactory results due to the difficulties in achieving accurate mechanical alignment between the systems, the independent vibration of the two systems, and the lack of recombination software containing the capability to remove misalignment and vibration effects. Based on these findings, single sensor camera systems were investigated further for their HDR capabilities. Of the single sensor cameras, most of the imagers experienced sensor saturation in the brightest region of the frame resulting from the rocket plume. One exception was the RED EPIC imager with HDR mode, utilizing a time sequential dual exposure method to capture both a “light” and “dark” frame, by means of two different exposure times. Although the RED EPIC frames utilize the same imager and optics, the time differential between frames will result in some variance in scene motion and luminance. RED EPIC manufacturer software includes a blending utility for combining dual exposure frames via one of two blending methods, however the details of how the blending is performed is not documented in detail in manufacturer literature. The actual dynamic range capability of each of the systems is difficult to compare, as
vendor published data does not specify the test methodology or standard to which capability is measured. In addition, a difficulty encountered amongst all of the imager manufacturers is a lack of detailed description of in-camera data processing, where such information is often considered proprietary. Attempts to process raw data in a third party program such as Matlab is also challenging, as file formats are often not documented or standardized.

From these experiments it is concluded that the processing pipelines differ significantly from manufacturer to manufacturer, and that there is no clear definition as to how imaging system dynamic range is either measured or specified. This primary finding motivated in latter chapters the development of a new methodology for the testing and comparison of modern imaging systems.

List of Contributions

• Definition and review of the film capture baseline.

• An analysis of current HDR processing workflows based on field experiments with modern imaging systems.

• Identification of a lack of a standardized method for measuring dynamic range.

• Identification of challenges in comparing imaging system capabilities, working with imaging system data, and developing processing pipelines.

7.1.2 Calibrated Measurement of System Dynamic Range

Several limitations are identified in the evaluation of existing HDR capture technology. First, in most cases, no standard is utilized when stating imager dynamic range capabilities by the various vendors. Second, in-camera processing of data may be described generically, such as with a vendor specific log curve applied, but only in some cases is the processing described, or the reverse transform provided in order to linearize the data. Third, there is lack of discussion, or available data, related to acceptable noise level or the noise level used in the determination of signal to noise based sensor characterizations. To address these limitations, we presented a new method for characterizing HDR response from a “system” perspective, including both sensor and optical effects together. The method requires only a small set of commercially available test equipment, primarily an HDR test chart (Xyla-21) and processing software (Imatest). To provide context for the measurement, a review of image sensor theory, the associated noise model, and calculation theory is included.
Before presenting results, the arbitrary assignment of perceived quality terms are addressed and their relationship to noise. Manufacturer specification are often without reference or descriptive narrative as to the computational method. The use of arbitrary quality descriptors may indeed be arbitrary, as what is deemed “excellent” for one application, or scene, may be “acceptable” for another. Value can be added to dynamic range measurement and quality descriptors, and they can be better compared, when characterized with respect to a meaningful metric such as the maximum RMS noise defining the minimum luminance step used to determine dynamic range. A SNR of 20 dB (0.1 stop of RMS noise) was identified as a reference point having commonality to both ISO Standard 12232-2006 and the test software Imatest. Review of measured results for several modern imaging systems indicate that the industry trend in dynamic range reporting correlates closer with the “medium” quality descriptor in Imatest corresponding to a SNR of 6dB (0.5 stop of RMS noise when defining the minimum luminance patch) than with the 20 dB metric termed “acceptable” quality in the ISO Standard 12232-2006 and “high” quality in Imatest (0.1 stop of RMS noise).

Of the systems tested, the Black Magic Cinema camera was the most overstated in terms of manufacturer specified DR of 12.0 stops, where the measured value is 9.0 stops for the “medium” quality descriptor of 0.5 stop of RMS noise. The Hitachi DK-Z50 and DK-H100 industrial box cameras measured DR values are 9.3 and 10.6 respectively, of which the DK-Z50 has a manufacturer specified DR of 9.6 stops being fairly close in agreement. The RED ONE and RED EPIC (without HDR mode enabled) cameras measured DR values of 11.4 and 11.5 respectively was less than the published specifications of 13.0 or greater stops. For the RED DRAGON and RED EPIC utilizing the built in HDR dual exposure mode, the results varied based on the post production method used. When considering the dual frames separately and estimating the DR contribution from each, the RED DRAGON and RED EPIC had total measured DR of 14.8 and 15.7 stops respectively as compared to published values of 16.5 and 18.0 stops. The dual capture modes did increase the total DR as compared to single capture modes, however the time sequential nature of the capture results in scene and luminance variance. When using the Simple Blend or Magic Motion software recombination techniques provided by the manufacturer for the RED EPIC, the total DR was reduced to 15.0 and 13.1 stops respectively. Overall, the best performing sensor based on total DR and reduced post-production effort is the ARRI Alexa. The Alexa utilizes single image capture and has 13.9 stops of DR as compared to published specifications of 14.0 stops.
The imagers tested have been treated as “systems”, with the combined effect of the optics, the sensor hardware, and the camera processing. An assumption is made that lumping the individual effects of the optics, sensor hardware, and camera processing is not introducing systematic error into the method. This assumption is investigated further in the next chapter. The results indicate that the accuracy of manufacturer stated capabilities are dependent upon the chosen maximum RMS noise defining the minimum luminance step used to determine dynamic range. Different manufacturers appear to use different metrics, making comparison difficult.

List of Contributions

- A novel method for determining system dynamic range based on commercially available test equipment
- An explanation of the arbitrary assignment of perceived quality terms and their relationship to noise
- A table of system based DR measurement results for latest generation digital imagers.

7.1.3 Radiometric Calibration and Optical Effects Based on a Rigorous Noise Model

A comparison of dynamic range data from our system based calibrated measurement, manufacturer stated data, and commercially published test website data, is hampered by the fact that the results cannot be compared directly without a more stringent definition of the test parameters, and specifically the definition of an acceptable noise level. Further, there is not a good basis for determining if the measurements are indeed accurate, as there is not a calibrated independent luminance measurements for comparison. The effects of the optical system versus the sensor are not known, or if contributions from either are effecting the measurement. Contemplation of these unknowns leads to the realization that if the basis of HDR capture is to be truly understood, from which an appropriate HDR workflow can be developed, a detailed radiometric calibration needs to be performed. The calibration must treat the optical and sensor elements separately, so that the contributions from each are known. Further, processing steps taken at each stage of the capture pipeline must be considered in detail, ensuring that the resulting optical electrical transfer function is determined correctly. The method presented categorizes error via methodical calibration; including determination of sensor noise attributes, system gain response, and the effects of lens glare, such that the dynamic range of
the imaging system can be characterized to a greater extent than previous methods. Additionally, using reference measurements from a calibrated luminance meter, absolute radiometric calibration is achieved, albeit with a degree of error.

A novel method of minimizing the optical effects on HDR calibration by a careful multiple exposure noise based workflow was presented. In this new approach, the linearization of raw data from a sample camera was examined for the purpose of recombination, and creation, of the absolutely calibrated HDR scene. In processing the CR data, care was taken to account for the manufacturer defined black level and saturation limit while cropping the full sensor data to the active region. The channels of the Bayer CFA are separated and multiple exposures are combined applying an optimal HDR reconstruction based on a rigorous noise model and a noise based weighted average. White balancing is performed while maintaining the relative proportions of the individual colour channels. The resultant demosaiced HDR image is stored as a Radiance HDR image file.

Chapter 6, figure 6.12 shows the percent error between spot meter measurements and recombined HDR image measurements for our test scene using a Sigma lens attached to the Canon 5DM3 camera. We observe a logarithmic percent error increase for luminance levels falling below approximately 0.25 cd/m$^2$, corresponding with the top left, middle left, and bottom left 3×3 patch charts masked in darkness. The spot meter measurements were taken with a high accuracy Konica/Minolta LS-150 luminance meter, and based on this and the carefully performed HDR recombination effort, the low luminance physical accuracy error was attributed to glare. Published efforts to remove glare, such as computational methods involving deconvolution, have not performed well as a result of low signal to noise ratio [Talvala et al., 2007]. A currently recommended approach is to physically remove glare before capturing the image, often through the use of a higher quality lens. An additional data set is captured using a Zeiss 25 mm lens marketed by Zeiss as “exhibiting great flare suppression through T* anti-reflex coatings and internal light traps”. The Zeiss lens results in improved glare performance as compared to the Sigma lens, particularly in the low luminance range where there is an overall reduction in error. It is observed that the percent error is now consistently under 10% from the maximum luminance patch of 286.8 cd/m$^2$ to approximately 0.1 cd/m$^2$. The per pixel gain based error observed is significant, primarily a result of lens vignette, as high as 70% near the corners of the field of view.

Traditional methods combining multiple exposures into an HDR image often allow for absolute calibration to a known reference measurement, yet offsetting all image values based on a reference measurement does not provide a measure of the
linearity of the data throughout the full captured range. In our work, the calibrated methodology identified the lower limit of the dynamic range that falls within an acceptable absolute luminance error. In ??, the lower limit corresponds with the logarithmic percent error increase for luminance levels falling below approximately 0.25 $cd/m^2$.

The results indicate that the effects of glare limit the ability to accurately capture and measure the lower luminance portion of a scene. Glare effects can be reduced by using higher quality optics, or by reducing sources of glare in the scene, but cannot be completely eliminated.

List of Contributions:

- A novel method for determining a radiometrically calibrated radiance mapping
- Linearization of raw camera data throughout the image capture pipeline
- Analysis of the optical effects of vignette and glare.

7.2 Impact

The techniques presented in this thesis could impact several areas where HDR imaging is used. Entities performing image system testing can greatly benefit from the explanations of arbitrary defined quality designations, along with the described relationships to noise. Additionally, more accurate comparisons between systems can be made when using commonality of terms, standardizing test methodology, and reporting of similarly tested results. In applications requiring physical accuracy such as launch imaging, industrial lighting and simulation, computer graphics, autonomous driving, etc., the novel method presented for determining radiometrically calibrate data can provide greater insight into the true dynamic range of scenes, leading to improved equipment performance specifications, driving then the requirements of the entire image processing workflow. Calibrated data for machine learning is also very important, especially when replicating real life scenarios. In the case of the Canon 5DM3 imager with Zeiss lens, the shadow area limit is approximately 0.1 $cd/m^2$ based on 10% acceptable error, and the peak luminance for a large enough group of pixels considered to be observable is approximately $cd/m^2$, resulting in a physically calibrated linear dynamic range of 16.6 stops.
7.3 Future Work

With respect to the system based test methodology, utilizing commercially available HDR test charts and processing software to determine the dynamic range of HDR camera and lens systems, future work could include isolating and compensating for the optical effects on the measurement. Specifically, if a PPGR correction is employed as part of the test procedure, determining how the placement of the test chart within an uncalibrated optical field of view could be skewing the results. Other work may include the evaluation of luminance and chromaticity, specifically, using the colour chip section of the ColorChecker chart in our test scene.

In the area of the radiometrically calibrated workflow, the capability to calibrate and create an entire library of HDR images can be leveraged as a useful tool for evaluating the full HDR pipeline; a workflow that normally includes processing, encoding, transmission, and display functions. By evaluating quality at each of these steps in the HDR pipeline with calibrated images, a clearer picture of the effects of the processing can be realized. In some cases, it may be found that the dynamic range capabilities thought to be necessary could be overstated, thereby driving unnecessary hardware, software, and cost. In other cases, with the true dynamic range and luminance levels of a scene known, it may be necessary to invest in new hardware and software, or in some cases may lead to the development of new algorithms to process HDR data within a given quality expectation. Future work should explore this issue in detail. An example of a radiometrically calibrated scene measurement was performed for a rocket launch. A tone-mapped version of the scene is shown in figure 7.1. A false colour version of the scene is shown in figure 7.2. A reference grey card is placed in the scene and lit with a DC voltage based LED light. A luminance measurement of the grey card is taken in parallel with the image capture of the scene. The reference grey card measurement provides the absolute luminance reference for the image, radiometrically calibrated via our presented method. The brightest portion of the image, shown in yellow, is the rocket plume which is approximately $1 \times 10^5 \text{cd/m}^2$. This is an approximate reading, as on inspection of the image data, the pixels containing the plume where found to be saturated and therefore the actual luminance of the plume is most likely greater than measured. To truly determine the maximum luminance of the flame, additional tests will be required on future launches where the exposure is set to prevent saturation. The exposure settings used for the measurement included a shutter speed of $1/8000$ and an aperture of F2. The fastest imager shutter speed of $1/8000$ had already been utilized, and therefore cannot be further decreased to reduce exposure time. The
F2 aperture selection was based on flat fielding data for the imager and lens that had previously been collected and processed. In order to reduce the exposure for future captures to prevent plume saturation, new flat fielding data will be collected at smaller apertures. Another option to reduce exposure would be to introduce neutral density filters, however this will increase reflections between lens and filter, and increase glare. With respect to the minimum luminance of the scene, pixel values of approximately $1 \text{cd/m}^2$ were measured. The flame, however, is most likely causing glare issues, thereby limiting the accuracy of the low luminance readings. In the future we can take an HDR multiple exposure just before the launch, before the rocket flame causes glare issues, to get a more accurate minimum luminance reading. The difficulty in measuring both maximum luminance without saturation, and minimum luminance without glare effects, illustrates why it is difficult to accurately determine the DR of a real scene. The approximate maximum and minimum luminance readings for figure 7.2 of $1 \times 10^5 \text{cd/m}^2$ and $1 \text{cd/m}^2$, respectively, result in a DR of at least 16.6 stops. Future measurements, however, will likely show shadow region luminance is most likely closer to $0.1 \text{cd/m}^2$, with highlights likely $5 \times 10^5 \text{cd/m}^2$, resulting in a total scene DR of approximately 22.3 stops.

![Figure 7.1: Example radiometrically calibrated scene measurement of a spacecraft launch at night, tone-mapped for display](image)

Finally, this thesis has discussed the idea that quality can be defined both ob-
Figure 7.2: Example radiometrically calibrated scene measurement of a spacecraft launch at night, false-colour version

jectively or perceptually. In either case, there is human decision making in defining what quality levels are, and what is deemed “acceptable”, or some other arbitrary descriptor. Future work should consider using a calibrated HDR test dataset to better quantify definitions of quality in terms of known luminance quantities for given scenes and identify to what affect the full HDR pipeline has altered the original scene data.
Appendix A

Dynamic Range Conversion Chart
<table>
<thead>
<tr>
<th>Density</th>
<th>10^0</th>
<th>10^0.15</th>
<th>10^0.2</th>
<th>10^0.3</th>
<th>10^0.4</th>
<th>10^0.6</th>
<th>10^0.8</th>
<th>10^1</th>
<th>10^1.2</th>
<th>10^1.4</th>
<th>10^1.6</th>
<th>10^1.8</th>
<th>10^2</th>
<th>10^2.2</th>
<th>10^2.4</th>
<th>10^2.6</th>
<th>10^2.8</th>
<th>10^3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(100C)</td>
<td>10</td>
<td>100</td>
<td>10,000</td>
<td>100,000</td>
<td>1,000,000</td>
<td>10,000,000</td>
<td>100,000,000</td>
<td>1,000,000,000</td>
<td>10,000,000,000</td>
<td>100,000,000,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decades</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Log(100C)</td>
<td>1.0</td>
<td>10</td>
<td>100</td>
<td>1000</td>
<td>10000</td>
<td>100000</td>
<td>1000000</td>
<td>10000000</td>
<td>100000000</td>
<td>1000000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(100C)</td>
<td>1.0</td>
<td>100</td>
<td>1000</td>
<td>10000</td>
<td>100000</td>
<td>1000000</td>
<td>10000000</td>
<td>100000000</td>
<td>1000000000</td>
<td>10000000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Log</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Log</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Figure A.1: Dynamic Range Conversion Chart
ACES. Academy color encoding system, 2015.


ARRI. Alexa’s sensor, 2013a.

ARRI. Alexa lut generator: Lookup tables, 2013b.

ARRI. Arri alexa camera overview, 2015.


goHDR. Player software overview, 2014.


Samuel W Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T Barron, Florian Kainz, Jiawen Chen, and Marc Levoy. Burst photography for high


Imatest. Step chart module documentation, 2014.


Timo Kunkel and Erik Reinhard. A reassessment of the simultaneous dynamic range of the human visual system, 2010.


LG. 55” class smart curved oled 4k, 2015.


Panasonic. Panasonic 55 class premiere 4k ultra hd smart tv cx800 series tc-55cx800u, 2015.


133


RED. Overview of the redcode file format, 2012.

RED. Understanding red logfilm and redgamma, 2013.

RED. Redcine-x post production software overview, 2014.

RED. Weapon dragon camera overview, 2015a.

RED. Redcast on-camera hdr monitoring, 2015b.


ST SMPTE. Academy color encoding specification (ACES), 2012.


Adam Wilt. Panasonic gh4: Dr and gamma, timelapse, and more, 2014.
Sebastian Wober. Dynamic range - sony a7s vs. the others, 2014.
