A 360° OMNIDIRECTIONAL PHOTOMETER USING A RICOH THETA Z1

IAN MACPHERSON

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

GRADUATE PROGRAM IN COMPUTER SCIENCE
YORK UNIVERSITY
TORONTO, ONTARIO

JULY, 2023

© Ian MacPherson, 2023

Abstract

Spot photometers measure the luminance emitted or reflected from a small surface area in a physical environment. Because the measurement is limited to a "spot," capturing dense luminance readings for an entire environment is impractical. This thesis demonstrates the potential of using an off-the-shelf commercial camera to operate as a 360° luminance meter. The method uses the Ricoh Theta Z1 camera, which provides a full 360° omnidirectional field of view and an API to access the camera's minimally processed RAW images. Working from the RAW images, this thesis describes a calibration method to map the RAW images under different exposures and ISO settings to luminance values. By combining the calibrated sensor with multi-exposure high-dynamic-range imaging, a cost-effective mechanism for capturing dense luminance maps of environments is provided. The results show that the Ricoh Theta calibrated as a luminance meter performs well when validated against a significantly more expensive spot photometer.

Acknowledgements

I would like to first and foremost thank my parents Scott, Vicky, and brother John for their unrelenting support in allowing me to choose my path in life. Their openness in letting me try as well as quit almost every extra-curricular activity in my childhood until I found my passion instilled a mindset that embraced trial without the fear of error. A skill which - as it turns out - is quite practical when applied to research and innovation.

My grandparents Louie and Helen taught me the value of hard work, determination, and how to build a life from no more than an idea and a stubborn belief in your abilities. Their mindset of not knowing how to do something, doing it anyways, and becoming the best at it, is a marvel of the entrepreneurial and immigrant spirit. Something I try to carry with me in every aspect of my life.

I would like to further thank my Supervisor, Dr. Michael S. Brown, for building a nurturing lab culture that embraces collaboration, open discussion, and personal growth. Michael shows nothing but the utmost patience and kindness when discussing new research directions and does so by striking the perfect balance between a positive outlook and a realistic approach. His lessons shared in personal and group discussions set an incredible example for developing quality research and living a well-rounded life.

I thank those in the York University Computational Vision and Imaging (CVIL) lab who helped me directly and indirectly with this thesis through various high-level discussions. Those include but are not limited to: Jason Yu, Sai Tedla, Harrish Thesarathan, Hoang Minle, Beixuan Yang, Matthew Kowal, Trevor Canham, and Javier Vasquez-Corral.

I am thankful for the continued mentorship of Dr. Kevin Swersky, and Dr. Konstantinos G. Derpanis. Thank you for challenging my notions of what is possible and encouraging me to go beyond my limits and continue to strive to be the best that I can be.

A special note is reserved for my lovely partner Josette who went out of her way to ask questions and understand my thesis while coming from a non-technical background. Her help with early testing of the method led to excellent locations and beautiful captures leading directly to the diagrams in this thesis and publication. Thank you for being supportive of my ambitions, career pursuits, and the long hours that necessarily come with them. I am lucky to have you in my life, and wouldn't trade it for the world.

Contents

Abstract ii

Acknowledgements iii

Contents iv

List of Tables vi

List of Figures vii

1 Introduction 1

2 Background and Preliminaries 6

2.1 CMOS 6

2.2 In-camera processing pipeline 8

2.2.1 Demosaicing 9

2.2.2 Lens shading correction 10

2.2.3 Sensor noise and denoising 13

2.3 Exposure, ISO, and HDR imaging 14

2.3.1 Exposure 14

2.3.2 ISO 15

2.3.3 High dynamic range imaging 15

2.4 Omnidirectional luminance estimation 19

2.5 Ricoh Theta Z1 20

2.6 Summary 22

3 Converting the Ricoh Theta to a 360-degree photometer 23

3.1 Monitor calibration 24

3.2 Ricoh Theta calibration 25

3.3 Summary 30

4 Testing the Ricoh Theta as a 360-degree photometer 31

4.1 Testing procedure 31

4.1.1 Image capture 33

4.1.2 High dynamic range imaging 33

4.1.3 Ground truth luminance measurement 35

4.1.4 Luminance measurement to luminance image comparison 35

4.2 Results 38

4.3 Summary 42

5 Conclusions and Future Work 43

5.1 Discussion 43

5.2 Future Work 45

5.3 Conclusion 45

Bibliography 47

List of Tables

4.1 The reported results of the method across the test scenes. 38

List of Figures

1.1 The output of a single omnidirectional 360° JPEG image from the Ricoh Theta and the resulting luminance map from RAW images using multi-exposure fusion. 3

2.1 Illustration of a CMOS sensor. This diagram shows a close-up view of three pixels in an RGB CMOS sensor. 7
2.2 The CFA pattern on the Ricoh Theta Z1 is shown in the illustration. The pattern is defined by the top-leftmost region, highlighted by the purple rectangle, which repeats across the sensor in an RGGB arrangement. . . 10
2.3 An example pipeline for correcting the radial falloff (i.e., lens shading) in the intensity of light from the centre of a camera sensor. 12
2.4 An illustration of how the modification of exposure and ISO settings allow a camera to capture different portions of the dynamic range of a scene. . . 17
2.5 Photopic luminous efficiency function which approximates the human visual system's experience of light intensity between 450nm-770nm. This response is based on an observer with normal vision. 20
2.6 Overview of the Ricoh Theta architecture and RAW dual fish-eye output. 21

3.1 An overview of the Ricoh Theta calibration procedure. 24
3.2 This diagram shows the linearity in response of the Ricoh Theta camera sensor to increasing luminance values. Each linear plot shows the response for a single exposure and fixed ISO. 26

3.3 The camSPECS device for colour calibration. Diagram from Image Engineering's product page [9]. 28
3.4 The spectral sensitivity of the Ricoh Theta across colour filters in comparison to the photopic luminance efficiency function. 29

4.1 RAW dual fish-eye image captured in one of the test scenes. 32
4.2 RAW dual fish-eye image annotated in label studio. 36
4.3 Results of the method, respective HDR fusion map, false-coloured luminance image, and measurements made from the SpectraScan and Ricoh Theta. 39
4.4 Results of the method, respective HDR fusion map, false-coloured luminance image, and measurements made from the SpectraScan and Ricoh Theta. 40
4.5 Example of producing perspective images from a single 360° image. . . . 41

Chapter 1

Introduction

Lighting conditions are an essential property of indoor and outdoor environments, particularly for visual perception. Almost all retinal images are highly ambiguous, and human vision relies on stable lighting properties to overcome this ambiguity and infer the properties of objects and surfaces [17]. Therefore, having a reliable statistical model of an environment's lighting from omnidirectional measurements is useful for research in this area.

Specialized spot photometers such as the Jadak SpectraScan PR-655 [10] or the Konica Minolta LS-160 [13] are required to measure perceived brightness accurately in terms of luminance (i.e., candela/meter^2 or nit). These devices operate in a point-and-shoot manner and measure the light emitted or reflected from a small surface patch in the scene. It can take up to six seconds to measure a "spot" in the scene, depending on the settings of the device and the luminance of the scene. The auto-exposure behaviour of the photometer increases or decreases integration time according to the intensity of the measured

region, resulting in longer capture times for darker regions than brighter ones. Furthermore, multiple sparse measurements only provide an understanding of a fraction of the light present in a scene, while the environment may be changing between each photometer measurement. Additionally, spot photometers can be expensive, with state-of-the-art devices costing several thousand dollars.

Unlike luminance spot meters that are limited to a single spot in a scene, consumer cameras provide a convenient, cost-effective, and intuitive way to record large portions of light in an environment. Human observers can quickly interpret camera images to reveal relative brightness in an environment falling within the camera's field of view and dynamic range. Recent advances in lens technology, sensor size, and optical path folding allow cameras like the Ricoh Theta [20] to capture a full 360° omnidirectional field of view in azimuth and altitude, as shown in Fig 1.1. However, digital cameras do not aim to reproduce an accurate recording of scene luminance; instead, they strive to produce visually pleasing photographs in a display-referred colour space, such as sRGB [11]. As a result, onboard camera hardware performs proprietary image processing routines to enhance the image's contrast and colours such that the final pixel values have a complex and indirect relationship with the real environment. Recovering such relationships is difficult unless the nonlinear operations applied are known, and this is commonly proprietary information that differs between device manufacturers. Additionally, most digital camera sensors have a limited dynamic range compared to photometers, and the final output JPEG image is further reduced to 8-bit values and compressed for compatibility with display technology, storage, and faster file transfer.

Unprocessed RAW images allow the closest access to the direct response of a cam-

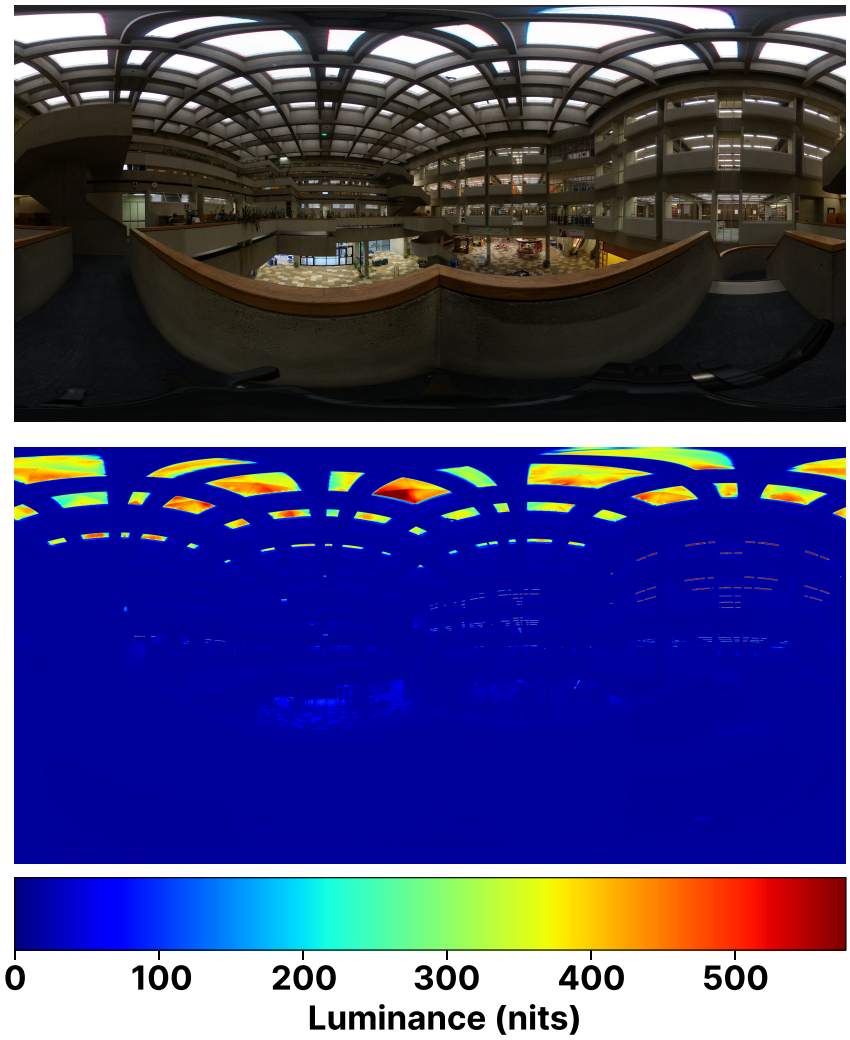

Figure 1.1: (Top) The output of a single omnidirectional 360° JPEG image from a Ricoh Theta Z1 camera. (Bottom) The corresponding omnidirectional luminance map computed from calibrated RAW images using multi-exposure fusion. This image has been falsely coloured to reveal measured luminance (nits) per pixel.

era sensor. It is important to denote these images as minimally processed, as there may be some low-level processing steps applied to the image to account for hardware-specific defects. A minimally processed RAW image from this definition (i.e., no additional processing steps) has a direct linear relationship between the physical world and per pixel value. The values of RAW images are not in the range 0-255 as is common in sRGB images. Instead, RAW image are encoded in digital values that are within the range of sensors bit depth. Most sensors encoded digita values in the range of 10-bit (i.e., 0-1023), 12-bit (i.e., 0-4095), or 14-bit (i.e., 0-16353).

When cameras allow access to the minimally processed RAW image, there is an opportunity to calibrate the sensor such that each pixel's digital value can be interpreted in terms of physical or perceptual measurements. This is possible due to the linear relationship between the intensity of light in the physical world and the digital response of the sensor. Spot luminance meters use silicon photodiodes with photopic filters that are similar architecturally to CMOS sensors (i.e., a common imaging sensor architecture) used in consumer cameras. Recently, the manufacturers of the Ricoh Theta provided a software API [20] that allows access to RAW dual fish-eye images on certain Ricoh devices. These RAW captures can be stitched through proprietary software and viewed as 360° omnidirectional images. The availability of off-the-shelf access to this omnidirectional RAW image whose pixel values maintain a relationship to the physical brightness in a scene is the impetus for the work in this thesis.

Contribution: This thesis presents a working proof of concept for a 360° omnidirectional luminance meter based on a Ricoh Theta Z1 camera. Specifically, a calibration procedure is described to estimate the response of the Ricoh sensor with respect to scene luminance at different ISO and exposure settings depending on the requirements of the captured scene. To overcome the limited bit-depth and dynamic range of the CMOS sensor, high dynamic range (HDR) imaging is used to translate pixels from multiple images with different exposure times to their corresponding luminance values (i.e., cd/m^2). The result is a photometric device that can quickly measure the luminance across an entire environment with minimal captures required. When the device is used at test time to capture images within the calibrated exposure and ISO range, the results show good performance compared to measurements made with a more expensive spot photometer.

Chapter 2

Background and Preliminaries

This chapter covers the underlying technical aspects related to the work in this thesis. The primary focus is on the CMOS sensor architecture and the relevant in-camera processing pipeline steps whose understanding is essential for motivating the methods and experiments described. While the discussed pipeline steps may not be used (i.e., the Ricoh images are not passed through a traditional in-camera processing pipeline), the underlying concepts are incorporated into the design of various procedures for achieving the reported results of the method.

2.1 CMOS

The work in this thesis is based on the hardware technology called complementary metal oxide semiconductor (CMOS) [5] which is used in most imaging sensors and provides measurements of light in an environment. The CMOS sensor consists of an array of photodiodes which each convert the intensity of incoming light (i.e., photons) to a correspond-

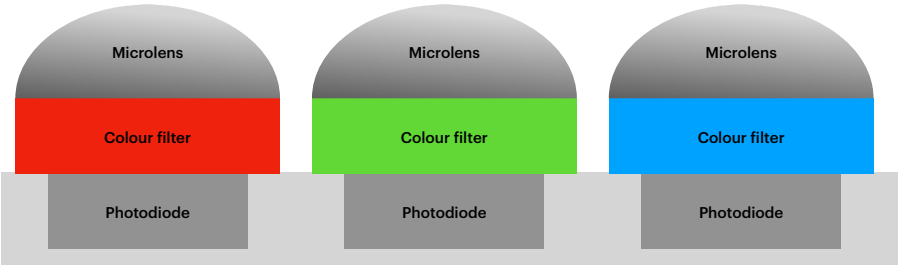

Figure 2.1: Illustration of a CMOS sensor. This diagram shows a close-up view of three pixels in an RGB CMOS sensor.

ing digital value. In this regard, the sensor in its simplest form provides a per-pixel photon count in digital values across the pixel array. The range of the recorded digital values per pixel depends on the bit-depth of the given sensor. For example, a 16-bit sensor will return raw sensor values in the range of 0-65535, while a 10-bit sensor will return values in the range of 0-1023. The values within the bit-depth are quantized responses of the per-pixel photon count, usually with some form of analog gain applied by the sensor at capture time. The number of photons that each pixel can measure, or the bit-depth, is related to the physical size of each of the pixel's photo-sensitive diodes (see Fig. 2.1), and is referred to as the well-capacity of the pixel.

CMOS sensors used in consumer cameras typically consist of three layers, depending on the variation of architecture, to capture colour images. The layers are a photodetector (diode), which is responsible for photon count; a colour filter, which allows transmission of specific spectral information; and a microlens, which directs light onto the single photodetector below it. The photodetector produces an electrical charge as photons are collected in its "well." The colour filters limit the transmission of photons associated with different spectral wavelengths. The microlens is used to direct and increase the amount of light toward a given photodetector in the array.

Typically, the photon count is integrated over a specific exposure time and returned by the sensor in a "rolling shutter" pattern where each row is exposed and read sequentially. This is a hardware-specific readout pattern and can vary depending on the sensor manufacturer. For example, some CMOS sensor architectures allow for a global shutter, where all pixels are integrated, exposed, and read out at the same time instead of in a rolling shutter pattern. More recent architectures can have varying exposures on a per-pixel level [21], offering benefits for high dynamic range image capture.

The RAW digital value output of the CMOS sensor scales linearly with respect to the intensity of light present in the scene. Meaning, a halving or doubling of the intensity of light in the scene will correspond with a halving or doubling in the RAW digital value output by the sensor. The linearity of the raw sensor response to the intensity of light is an important property to retain when using cameras as scientific sensors, as the correlation between the intensity of light in the physical world and digital value is preserved. Device-specific Perceptual processing steps which are common in in-camera processing pipelines will break this linear relationship and as such are avoided by the access to the minimally processed RAW image of the Ricoh Theta. More detail on the importance of RAW image linearity is discussed in the following section.

2.2 In-camera processing pipeline

The *in-camera processing pipeline* refers to the sequence of processing steps applied to the raw photon measurements of a camera sensor to produce a visually pleasing image. The measurements of light produced by the CMOS sensor are returned as digital values,

and operations such as demosaicing, lens shading correction, and denoising are applied to produce the final image suitable for human observation. Delbracio et al. [4] provide an in-depth review of the complex camera pipelines commonly found in consumer mobile phones.

The steps described in the following section are a subset of what can be found in consumer camera pipelines. Many of the operations in consumer pipelines are nonlinear and break the relationship between the resultant image and the physical world. However, this thesis is interested in using the camera as a measurement device instead of a device for creating visually pleasing images. As a result, the work in this thesis will incorporate some of the early-stage pipeline steps to support or enhance the downstream task of 360° luminance measurement and maintain the linear relationship between the incoming light and the final pixel values which represent a per-pixel luminance measurement.

2.2.1 Demosaicing

As discussed in Sec. 2.1, a colour filter array (CFA) is placed over the photodiodes to sample different parts of the visual spectrum. The most common pattern for colour filter arrays is referred to as the Bayer filter, named after its inventor Bryce Bayer from Kodak Labs. This filter organization follows a blue-green-green-red (BGGR) pattern, which describes the top leftmost two-by-two pattern that repeats across the entire array. An example of the filter pattern found on the Ricoh Theta can be seen in Fig. 2.2. This means each pixel on the CMOS will represent only one colour. Ideally, we would like to have three colours per pixel to reconstruct a digital representation that is similar to the stimulus of the human visual system.

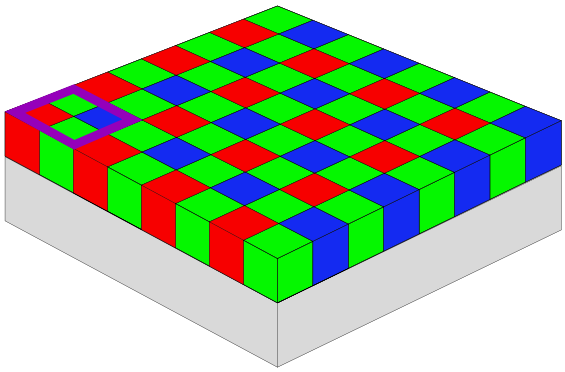

Figure 2.2: The CFA pattern on the Ricoh Theta Z1 is shown in the illustration. The pattern is defined by the top-leftmost region, highlighted by the purple rectangle, which repeats across the sensor in an RGGB arrangement.

To convert the single-channel "mosaic" to a full-resolution, three-channel image, a weighted interpolation is performed to calculate the intensity of the missing colour channels at each spatial location from the sensor. Simple bilinear interpolation based is the most straightforward algorithm for predicting these values, although more complex algorithms exist [7, 8, 22, 26]. This process is commonly referred to as "demosaicing." Since only a single colour is measured at each pixel, two additional values must be calculated at each pixel in the Bayer pattern to form the final image. This means that $\frac{2}{3}$ of the digital image does not exist and must be inferred using neighbouring values. It is common to pair the demosaicing algorithm with other in-camera processing steps such as denoising to reduce the time required to process an image.

2.2.2 Lens shading correction

Camera sensors encounter a lens shading phenomenon, where the light across the camera sensor decreases radially in intensity from the center. This occurs because the light that

reaches the center of the sensor is more intense than the light that reaches the outer edges. This is particularly problematic for digital sensors found in mobile phones due to the size of the sensor and the recessed position of the optical center within the camera's housing. Without lens shading correction, each captured RAW image will have a vignetting artifact present which skews the accuracy of the response to light spatially. Although such an effect may be added in post-processing for aesthetic purposes, vignetting impedes digital cameras from providing reliable measurements for scientific tasks due to the spatial variation of signal across the image. An example of lens shading in a RAW image, and the method for its correction can be seen in Fig. 2.3.

To perform lens shading correction (also referred to as flat-field correction), a correction mask is calibrated and applied to each image during capture. The correction mask applies an inverse gain to the radial falloff, which is specific to each sensor and can be calibrated across different camera settings. To measure the radial falloff, multiple captures of an evenly illuminated scene, such as a white canvas, are taken and averaged. The average normalized response is then inverted to compute the lens shading mask. Pixels near the center of the image are multiplied by a value close to one, while the darker pixels near the edge of the sensor are digitally gained relative to the intensity of light in the middle of the image. Once calibrated, the lens shading correction mask, or gain function, is stored on the device and applied at capture time, similar to a gamma function or a per-pixel gain map. The process of lens shading correction does not impede the linearity between the brightness in the scene and the sensor response and is one of the low-level processing steps included when referring to the minimally processed RAW image from the Ricoh Theta.

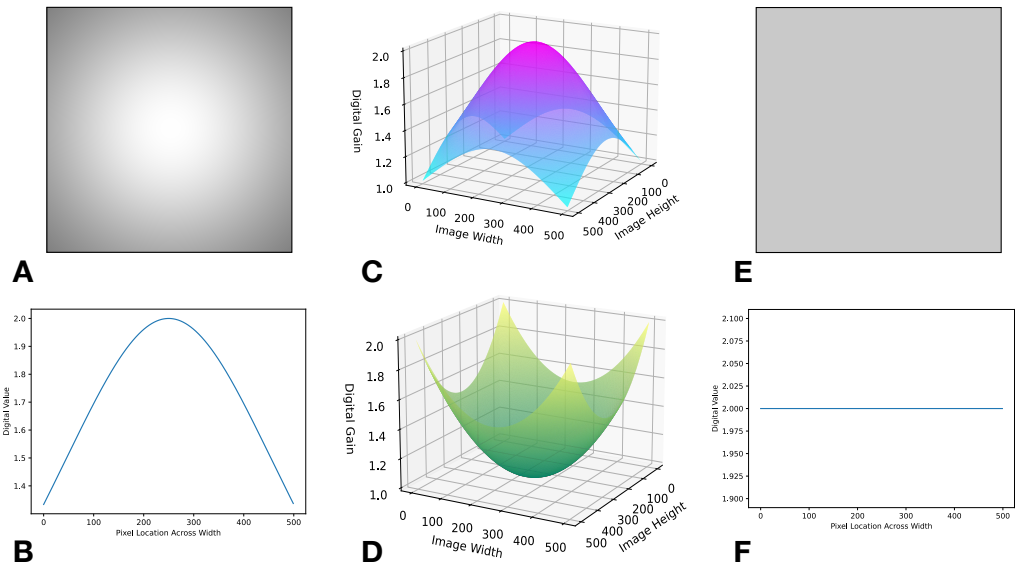

Figure 2.3: An example pipeline for correcting the radial falloff (i.e., lens shading) in the intensity of light from the centre of a camera sensor. In (A), an example flat field image is shown, where the radial falloff (i.e., lens shading) is visible. A (central) vertical or (central) horizontal cross-section of pixels across the middle of the image is plotted as a normalized histogram in (B). In (C), a 3D normalized view of the image's radial falloff is shown. In (D), the inverted flat field correction for the given sensor is shown. The resultant image after correction is shown in (E), along with the same central cross-section of pixels plotted in (F). It can be seen that this process fixes the issue of nonuniformity across the image due to lens shading.

2.2.3 Sensor noise and denoising

The digital values produced by the sensor's photodiodes are susceptible to noise, which can be influenced by camera settings such as the ISO and exposure, as well as the operating temperature of the CMOS. Various post-processing techniques have been developed to address sensor noise, including classical algorithms such as the bilateral filter [24] and non-local means denoising [2], as well as more recent neural network-based methods [23, 27]. However, in this thesis, no explicit denoising is applied using an explicit post-processing algorithm on a per-image basis. Instead, the issue of sensor noise is addressed by averaging multiple images during the calibration process.

The pipeline steps discussed in this section are particularly relevant to the task of luminance measurement using a consumer-grade camera sensor. However, certain additional steps commonly found in the in-camera processing pipeline, such as defective pixel correction, black level subtraction, auto-focus, and auto-white balance, are not necessary for achieving luminance measurement and hence are not discussed. It is important to understand the Bayer pattern and the impact of lens shading and noise to comprehend and execute the proposed method in this thesis. The next section will delve into camera parameters and their relationship with sensor response, which is crucial for developing an understanding of the motivation behind the process of capturing a high dynamic range image.

2.3 Exposure, ISO, and HDR imaging

In this thesis, the primary objective is to measure the luminance of a scene accurately. Thus, it is crucial to have a comprehensive understanding of exposure, ISO (i.e., a parameter named after the International Organization for Standardization, where the name is not expressive of the behaviour of this camera setting), and their effects on image capture when used in combination. Both exposure and ISO are intrinsic camera parameters that directly impact the hardware at capture time and are determined relative to the subjects present in the scene and the dynamic range of the lighting conditions. By adjusting these settings, cameras can capture brighter or darker images depending on the desired aesthetic or downstream task. These camera parameters are commonly used in conjunction, as ISO has the side-effect of boosting noise, and longer exposures will reduce this noise but suffer from motion artifacts. An intentional tuning of both parameters together is necessary to mitigate the shortcomings of each.

2.3.1 Exposure

Exposure is determined by two factors: the aperture of the camera, which controls the size of the hole through which light enters, and the exposure time, which regulates the amount of time that light is integrated by the sensor. For the experiments in this thesis, the aperture of the Theta device is fixed and all further mentions of exposure can assume the use of a consistent aperture. A shorter exposure time means that fewer photons reach the sensor during the given period, resulting in an overall darker image (i.e., low exposure). Conversely, when a longer exposure time is used, more photons reach the sensor, leading to a

brighter image (i.e., long exposure). However, long exposures may also average motion in the scene, causing motion-blur or ghosting artifacts in regions where motion is at a higher frequency than the current exposure time. To address these artifacts, the best approach is to limit the camera's motion during capture by using a tripod or camera stabilizer and capturing a static scene. The type of CMOS sensor used (i.e., rolling or global shutter) may affect the appearance of long exposure artifacts.

2.3.2 ISO

ISO is a multiplicative gain factor that scales the digital values and produces a brighter image. In some scenarios, a short exposure time may be needed to account for motion in a dark environment, but it may not provide sufficient light for proper exposure. In such cases, increasing the ISO setting will brighten the image without the risk of capturing ghosting artifacts, since it gains the low exposure image. However, the gain factor also increases the noise, creating a trade-off between ISO and exposure time. The two parameters can be intentionally combined to balance this trade-off between capturing high noise or exposure artifacts.

2.3.3 High dynamic range imaging

The dynamic range of a scene is the ratio between the darkest and brightest intensities of light that are present. Typically, this ratio is greater than what can be captured with a camera sensor at a specific exposure and ISO setting. As a result, the image will not accurately represent the scene as perceived by the human visual system, which can adjust to this situation naturally. Therefore, methods for extending the dynamic range of the

sensor must be introduced, as it is unlikely that every imaging scenario will fall within the sensor's dynamic range. Although exposure and ISO settings can be adjusted, the sensor architecture and manufacturing specifications provide an upper limit on the dynamic range that can be captured with a single image. This means that each image, regardless of the changes in exposure and ISO settings, will have a similar dynamic range. Since each combination of settings captures a different portion of the dynamic range of the scene, a weighted fusing of images can be calculated to produce an image which represents a higher dynamic range than any of the individual captures with their respective settings.

When the dynamic range of a scene exceeds the dynamic range of the sensor, some pixels become overexposed or underexposed, resulting in their digital values being close to or at the upper and lower bounds of the sensor's bit-depth. As a consequence, the reliability of the pixel values in this range as a correspondence to the light in the scene is uncertain. For instance, a pixel that provides a maximum reading of 4095 on a 12-bit sensor (i.e., the highest value representable with a 12-bit sensor) is considered overexposed.

Pixel overexposure does not imply that the physical value cannot be captured by the sensor. Instead, the pixel has attained its maximum photon measurement under the current ISO and exposure settings, limited by the dynamic range of the sensor. Similarly, the measurements near the boundaries of the sensor bit-depth are unreliable for imaging tasks, reflecting the lower end of the dynamic range of the sensor. The limitations of camera sensor hardware to provide accurate readings near the edges of its bit-depth, along with limited capture of dynamic range motivate the need for a method which can extend the dynamic range that can be captured. This method is known as high dynamic range (HDR) imaging.

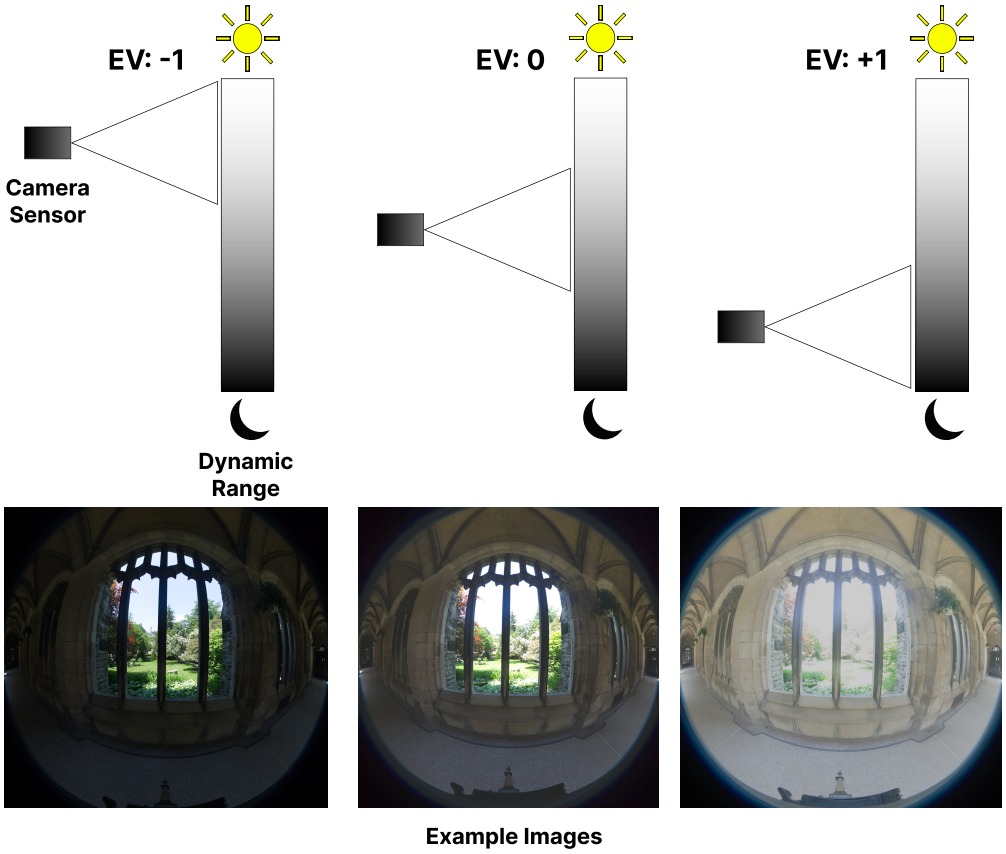

Figure 2.4: This illustration demonstrates how adjusting camera settings such as exposure and ISO can help capture a wider range of luminance values in a scene. However, it is important to note that the camera may not capture the entire dynamic range of the scene in a single capture. To capture a higher dynamic range than any individual capture, multiple images can be combined by varying exposure, ISO, or exposure and ISO settings between each. It is important to understand that longer exposures can capture darker regions more effectively, while shorter exposures can capture brighter regions more effectively. Having this understanding is crucial for the development of high dynamic range imaging techniques.

High dynamic range imaging [19] can help overcome the limited dynamic range of a camera sensor by fusing images captured with different exposure and ISO settings into a single image with a dynamic range beyond any of the individual captures. Each combination of exposure and ISO captures a different portion of the physical scene's dynamic range, but the fusion of a stack of images allows for each pixel to be algorithmically combined to produce an image with a higher dynamic range than the sensor. An illustration of this concept is shown in Fig. 2.4.

Early work in HDR required a calibration step to model the nonlinear photo-finishing applied by the camera [3, 15]. Access to the minimally processed RAW image simplifies HDR significantly; however, calibration is still necessary to model the relationship between different ISO and exposure times and the corresponding scene luminance. HDR is essential for the application in this thesis to expand the range of measurable luminance values, as a single capture can only report a fraction of the dynamic range in a scene. Due to the relationship between luminance and sensor response, capturing only a fraction of the dynamic range with a single capture limits the range of luminance measurements possible when using the Ricoh Theta Z1 for this task.

Motion presents a significant challenge when capturing images with multiple exposures for HDR imaging. The ideal scenario is to capture multiple images of the same content with only the intrinsic camera parameters varying between captures. However, when the camera is handheld, even minor hand tremors can cause misalignment between images, which can be corrected algorithmically but is more easily avoided by employing the use of a tripod. HDR capture may fail when the subject is moving quickly enough to appear in different regions within the field of view between subsequent captures, or during

the exposure integration time. This makes HDR imaging a more challenging problem for applications such as outdoor photography or driverless cars.

2.4 Omnidirectional luminance estimation

Luminance is a measure of the intensity of light emitted by a light source, expressed in candela per square meter (cd/m²), or "nits." It is based on the photopic luminous efficiency function, as shown in Fig. 2.5, which is standardized by the International Commission on Illumination (CIE) and approximates the sensitivity of the human eye to light in the visible spectrum. The function is based on the sensitivity of an observer with normal vision, meaning luminance represents the perceived intensity of light emitted from a source concerning the human visual system's response.

Pioneering work by Tominaga and Tanaka [25] proposed a catadioptric system for omnidirectional imaging, utilizing a conventional camera observing a spherical mirror. Their research laid the foundation for many of the geometric warping techniques required to map between a planar sensor and a spherical world coordinate system. Although they did not directly calibrate the sensor for luminance measurements, their work has undoubtedly influenced later technologies, such as the Ricoh Theta, which uses wide-angle lenses instead of spherical mirrors.

Mury et al. [18] developed a custom device, the "plenopter," for omnidirectional radiance mapping using 12 photocells with a 74° field of view each, capable of capturing the intensity of light in a 360° view from a point in space. This device was purposely developed for omnidirectional luminance measurement, however, since the device employs

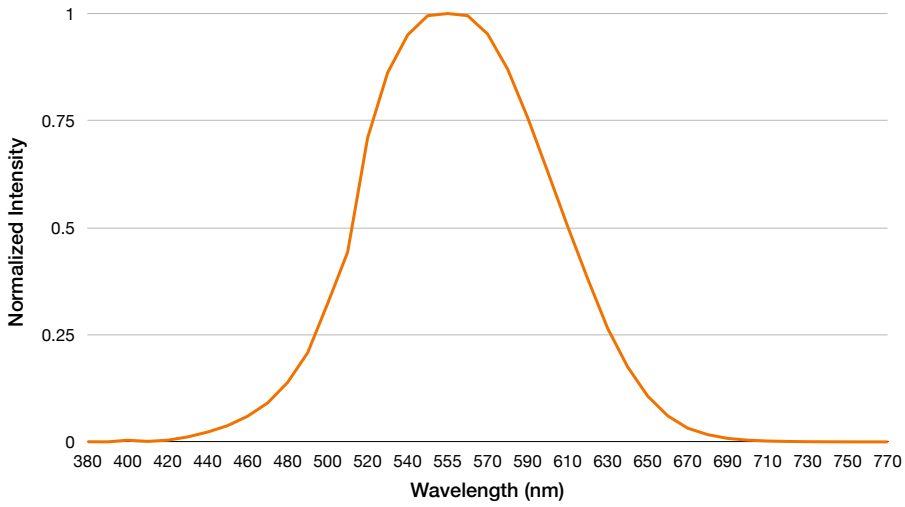

Figure 2.5: Photopic luminous efficiency function which approximates the human visual system's experience of light intensity between 450nm-770nm. This response is based on an observer with normal vision.

only 12 sensors, it does not effectively capture all regions of the scene in a 360° view. To account for this, harmonic interpolation was used to obtain dense omnidirectional measurements from the sparse photocell readings. Morgenstern et al. [16] also used a similar device to study the diffuseness of light based on the assumptions of multiple observers. This work motivates the need for a method which densely captures the intensity of light present in a scene.

2.5 Ricoh Theta Z1

The Ricoh Theta camera uses two wide field-of-view (i.e., "fish-eye") lenses and two CMOS sensors to capture two hemispheres of the environment, as shown in Fig. 2.6. The small overlapping region of the sensor images is stitched together to create a single

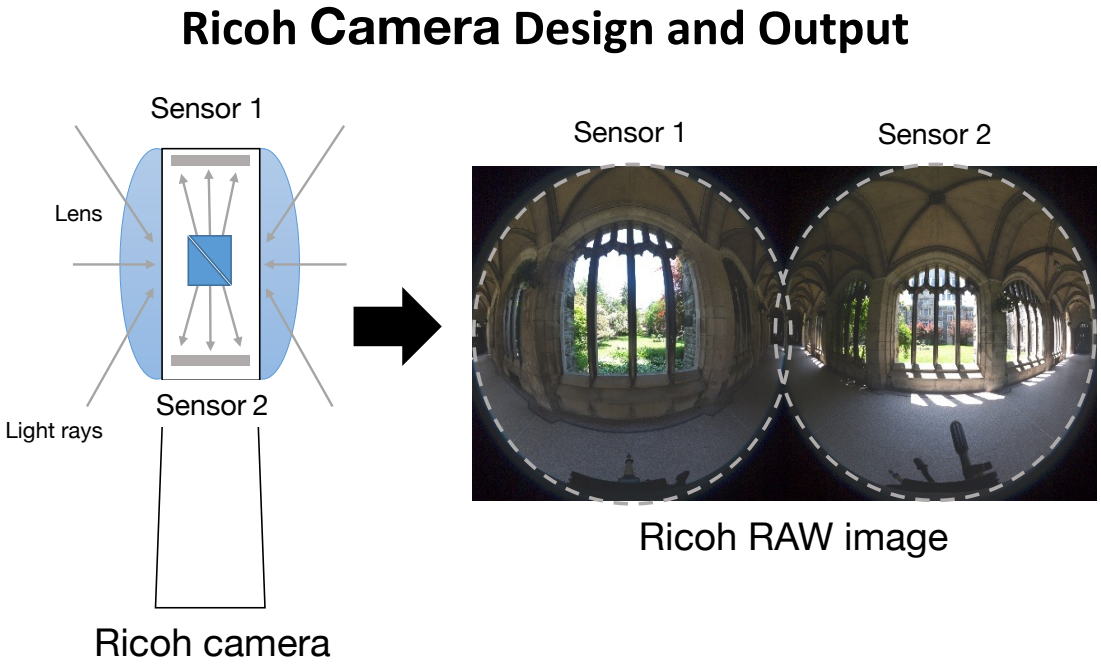

Figure 2.6: The Ricoh Theta Z1 device and its underlying construction are illustrated on the left, with two fish-eye lenses directed towards two separate sensors to capture two hemispheres of the environment. The two fish-eye images are captured and concatenated side by side to produce a single RAW dual fish-eye image.

7296×3648 image. The resulting image can be viewed as a spherical or equirectangular image using the accompanying software provided by Ricoh. The Theta Z1 model is advantageous for scientific measurement due to its ability to capture RAW images with minimal image processing, which is unavailable for most other models in the Ricoh Theta product line. Aghayari et al. [1] provided a geometric calibration of the Ricoh Theta and noted its potential applications in photogrammetry, robotic, and machine vision applications. Their approach uses targets placed across known distances in an enclosed room to determine the device's intrinsic and extrinsic parameters.

2.6 Summary

This chapter provides the necessary background and related work to motivate the use of a consumer-grade camera sensor for luminance measurements and highlights the opportunity that the Ricoh Theta Z1 presents for this task. The Ricoh Theta's RAW image transfer capability provides easy access to the sensor's raw response. The discussion began at the hardware level, with a focus on the CMOS architecture and relevant in-camera processing pipeline steps. The well-depth of the CMOS sensor is a hardware limitation that restricts the dynamic range that can be imaged with a single capture. To overcome this limitation, high dynamic range imaging methods are required to allow the camera sensor to function more like a photometer. The necessary steps of the in-camera processing pipeline to achieve the goal of luminance measurement were discussed and were chosen due to their linearity-preserving properties. The effects of noise, lens shading, and spectral response of the Bayer pattern were discussed in addition to their mitigations. The hardware limitations of the CMOS sensor and the intuition and motivations behind each required step of the in-camera processing pipeline inform the calibration and testing of the proposed method outlined in the following chapters.

Chapter 3

Converting the Ricoh Theta to a 360-degree photometer

This chapter outlines the calibration and testing procedure to turn the Ricoh Theta into a 360° omnidirectional photometer. Initially, an LCD monitor is calibrated using a photometer so it can be programmed to output desired luminance values. Desired luminance values are displayed, and RAW images are captured with the Ricoh Theta to determine the relationship between the luminance output of the monitor, and the RAW sensor response of the Ricoh Theta. This procedure validates the linearity of the sensor concerning scene brightness and serves as calibration (i.e., "training") data for the Ricoh Theta when used as a photometer in real-world scenes. Additional details including using only the green pixel values and the limitations of the method are also discussed. An overview of the procedure to map the RAW sensor response of the Ricoh Theta to luminance can be seen in Fig. 3.1.

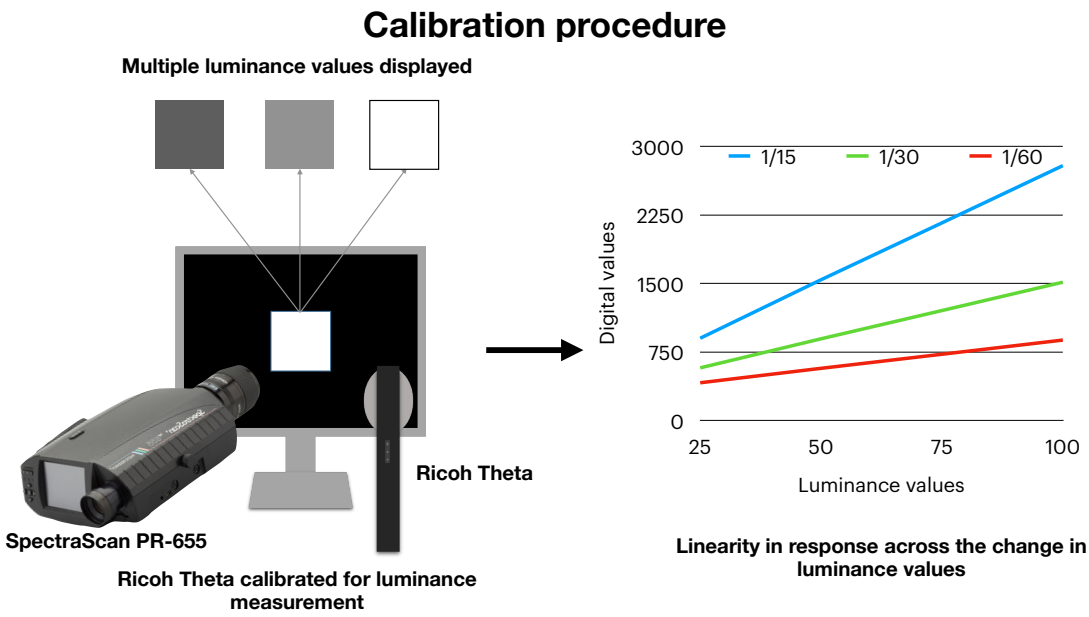

Figure 3.1: An overview of the Ricoh Theta calibration to map luminance to RAW sensor response. Specified luminance values are displayed by an LCD monitor and captured by the Ricoh camera using varying ISO and exposures.

3.1 Monitor calibration

To map the RAW digital values of the Ricoh Theta to corresponding luminance measurements in cd/m^2, a monitor was first calibrated using a luminance spot photometer. The calibration utility of the Psychtoolbox MATLAB library [12] was used, which interfaces with a wide range of luminance spot photometers. In this work, a SpectraScan PR-655 [10] was the device used for luminance measurements and is connected via USB to a computer controlling the monitor. To calibrate the monitor, a series of intensities were displayed and measured by the SpectraScan PR-655. These intensities are initially programmed into the monitor using arbitrary digital values and do not have a one-to-one relationship with

luminance. For instance, if intensities from the monitor are displayed in an 8-bit range, displaying a value of 255 from the monitor does not mean a luminance of $255cd/m^2$ will be measured with a photometer. The monitor calibration procedure involved displaying a series of luminance intensities from the monitor, with the spot photometer recording the luminance output from each respective intensity. Once completed, a mapping between the arbitrary digital values of the monitor and the resultant output in cd/m^2 is made. The monitor can then be programmed to output desired luminance values directly for calibrating the Ricoh Theta.

3.2 Ricoh Theta calibration

Once the monitor was calibrated to output desired luminance values, the RAW sensor response of the Ricoh Theta could be similarly calibrated to luminance. Except in this case, the luminance output of the monitor in cd/m^2 is mapped to the RAW sensor response of the Ricoh Theta. To do so, the monitor was programmed to display luminance values from approximately 1 cd/m^2 to 400 cd/m^2 (i.e., the maximum output of the monitor) in intervals of approximately $25cd/m^2$. For each luminance value displayed, the Ricoh Theta captured five RAW images across each ISO and exposure setting to mitigate noise and flicker from the refresh rate of the monitor. Once a dataset of RAW images was captured, the response of the Ricoh Theta concerning luminance values was plotted to validate the linearity of the sensor concerning the change in luminance. The result for a single ISO is shown in Fig. 3.2.

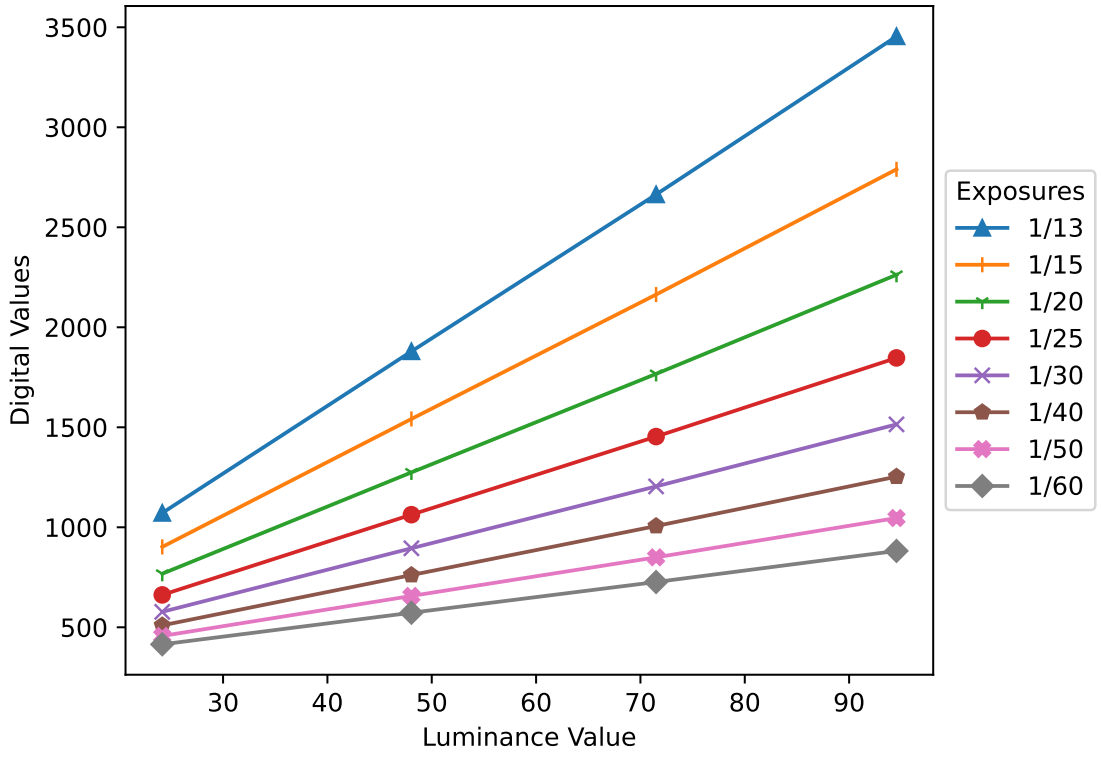

Figure 3.2: This diagram shows the relationship between the RAW image's digital values and luminance for various exposures when calibrated against the LCD monitor, using the lowest ISO of 80 for the camera. The plot demonstrates that the sensor's response is linear concerning incoming scene luminance within this range of luminance values calibrated against the LCD monitor. For this range of exposures, the luminance values are clipped at 100 cd/m^2 due to the oversaturation of the $\frac{1}{13}$ s exposure at greater luminance values. Exposure times are reported in seconds.

The plots in Fig. 3.2 illustrate the results of the calibration procedure. The Ricoh Theta Z1 sensor has 12 bits of dynamic range, resulting in a maximum digital value of approximately 4095. The plot shows that the sensor's response to increasing luminance is linear within the calibration range achievable by the LCD monitor. Ideally, there should be a linear relationship between the increase in exposure and the increase in response. For instance, if the camera outputs a digital value of x when observing a fixed luminance pattern

at an exposure of $\frac{1}{60}$ s, the digital value at an exposure of $\frac{1}{30}$ s is expected to be $2\mathbf{x}$, as the exposure time is twice as long. Unfortunately, this relationship was not observed between exposures, which may be due to the Ricoh API providing only an approximation to the true exposure or the sensor's nonlinear quantum efficiency. Therefore, it was necessary to establish mappings between the sensor's digital values at each exposure and ISO setting.

The Ricoh Theta has an onboard web server that allows for remote control of camera settings and captures, and the software API enables RAW image capture in addition to photo-finished JPEG images. The RAW image includes each fish-eye image side by side before geometric correction is applied to view the image in equirectangular format. The RAW images seem to have lens shading and defective pixel correction applied separately to each fish-eye image before saving the RAW image in Digital Negative (DNG) format. These processing steps are beneficial to the calibration procedure as they ensure a uniform sensor response [11], as discussed in Sec. 2.2.2.

The RAW images are in Bayer format, where each pixel corresponds to a single colour channel according to the CFA pattern of the sensor. The spectral sensitivity of the Ricoh Theta can be measured with an auxiliary device to determine the response of the colour filters across the visible spectrum. The particular device used in this work was the camSPECS device from Image Engineering [9]. The camSPECS device is a camera colour characterization device which is imaged, and, with additional software determines the spectral sensitivity of each colour filter. Since the Ricoh Theta has two sensors, each was calibrated separately and confirmed a low variance between spectral responses. The CamSPECS device can be seen in Fig. 3.3.



Figure 3.3: The camSPECS device for colour calibration. Diagram from Image Engineering's product page [9].

Comparing the spectral sensitivities of the Ricoh Theta to the photopic luminance efficiency function in Fig. 3.4, it can be seen that the green channel is most similar to the response of the human visual system. Since the green channel's spectral sensitivity is similar to the photopic luminous efficiency function used for luminance calculation, it is the only channel used for this method. The range of exposures is limited to a lower bound of $\frac{1}{60}$ s due to the LCD monitor's refresh rate and a maximum of $\frac{1}{2.5}$ s due to rapid pixel over-saturation from the lowest luminance values output by the monitor.

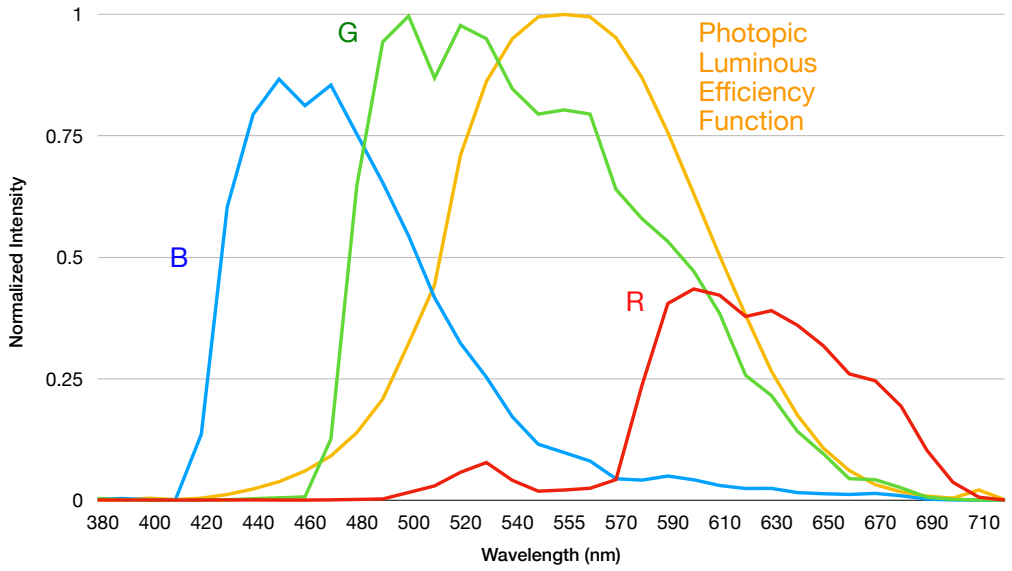

Figure 3.4: The Ricoh Theta's spectral sensitivity versus the photopic luminance efficiency function. The similarity between the response of the green filter (green) to the photopic luminance efficiency function (orange) motivates the use of this filter when approximating the per-pixel luminance from the raw sensor response.

The calibration of the Ricoh Theta against the monitor from Fig. 3.2 can be used as a "training" dataset when the Ricoh Theta is validated against real-world scenes. The linear relationship enables the use of a simple line-fitting function to model the conversion of RAW digital values from the Ricoh Theta sensor to luminance. Doing so allows the RAW image of the Ricoh Theta to be converted to a luminance value on a per-pixel basis. The monitor's refresh rate of 60Hz limited the fastest exposure that could be calibrated to approximately $\geq \frac{1}{60}$ s. Exposure times faster than the refresh rate resulted in black images or images with artifacts due to either a rolling shutter or monitor refresh pattern. Calibration data below this threshold failed to capture the calibration source adequately, making it difficult to correlate the RAW sensor response of the Ricoh Theta with the luminance of

the monitor. The refresh rate of the monitor, therefore, imposed a limitation on the lowest (i.e., fastest) exposure time that could be calibrated to approximately $\geq \frac{1}{60}$ s.

3.3 Summary

This chapter outlines the core methodology for calibrating the Ricoh Theta as a 360° omnidirectional photometer. First, the Ricoh Theta is calibrated against an LCD monitor programmed to output a known luminance. The calibration (i.e., "training") dataset is captured by sweeping the exposure and ISO parameters of the Ricoh Theta to create a linear mapping between the sensor response of the Ricoh Theta and the luminance output of the monitor. Multiple images are captured at each ISO and exposure setting to reduce noise in the calibration dataset.

Chapter 4

Testing the Ricoh Theta as a 360-degree photometer

In this chapter, the performance evaluation of the Ricoh Theta as a 360° omnidirectional photometer is presented. It was found that the device performed exceptionally well, achieving an accuracy of approximately 84% across the test scenes using the calibration data from the LCD monitor. These results provide strong evidence supporting the potential use of the Ricoh Theta as a reliable and efficient tool for 360° luminance measurements.

4.1 Testing procedure

The following sections describe the procedure developed for testing the luminance calibration of the Ricoh Theta in real-world scenes. First, the Ricoh Theta was placed on a tripod and captured a sequence of images in the scene. These images were captured using exposure and ISO values present in the calibration dataset. An overview of this process can

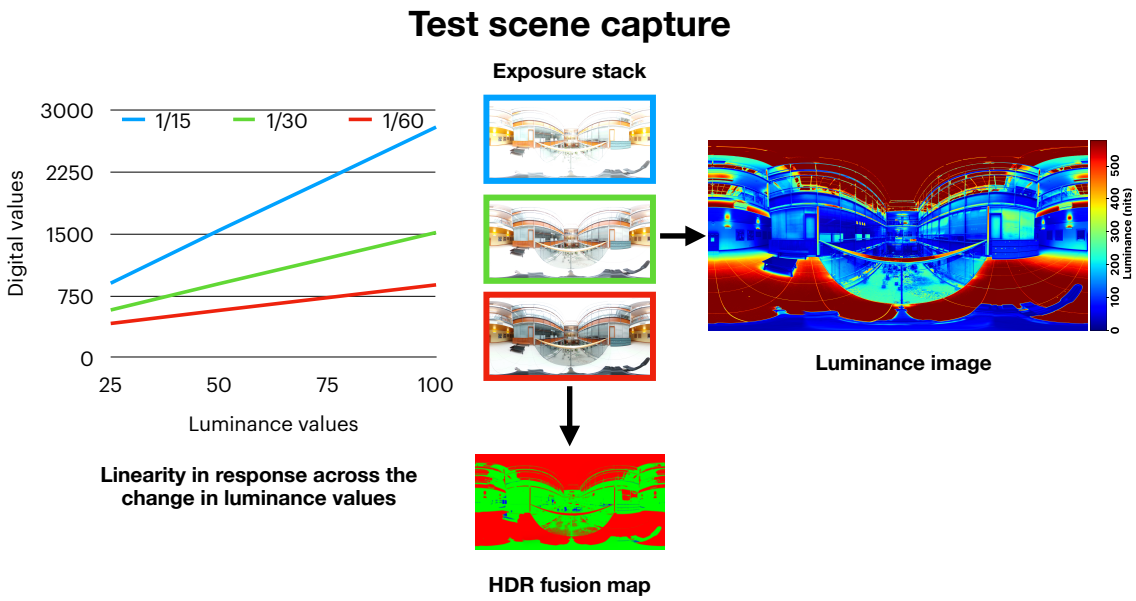

Figure 4.1: The pipeline for capturing images in the test scene, creating the HDR image, and the luminance image.

be seen in Fig. 4.1. Once the images were captured, the SpectraScan PR-655 was placed in the same location as the Ricoh Theta on a separate tripod and 10 measurements across the scene were made as ground truth. The linear functions from the calibration dataset are used to perform a per-pixel conversion of the RAW images of the scene to corresponding luminance measurements. The converted RAW images will be herein referred to as the luminance images. Detailed notes were made as to the location of the SpectraScan PR-655 measurements in the scene and later used to demarcate the locations of the ground truth measurements in the luminance image. Due to changes in lighting between image capture and ground truth measurement, time was limited between Ricoh Theta captures and SpectraScan measurements. This made it necessary to carefully manage potential sources of error, which are addressed in the subsequent chapter.

4.1.1 Image capture

The Ricoh Theta was mounted on a tripod in a scene with limited subject motion. Since the camera captures a 360-degree view in a single image, everything visible from the perspective of the Ricoh Theta could be considered for luminance measurement with the photometer. The position of the tripod holding the Ricoh Theta was recorded, and the height was adjusted to match the height of the Spectrascan PR-655 which was mounted on a separate tripod. The height adjustment was made to closely align the optical centers of Ricoh Theta and the Spectrascan PR-655, such that the rays of light measured from each are approximately the same.

Only the exposure and ISO settings used to capture the calibration dataset can be used for capturing test images with the Ricoh Theta. For every exposure and ISO combination in the calibration dataset, a single linear function is calculated. This function is used to convert the RAW sensor response into a luminance value for test images that are later captured using the same settings. The procedure for capturing the images involved selecting the desired settings and capturing three RAW images using the Ricoh Theta mobile phone application in rapid succession. The RAW images were then saved to the onboard memory of the Ricoh Theta for later transfer to a desktop computer. The procedure for fusing the three RAW images is outlined in the following section.

4.1.2 High dynamic range imaging

To create an HDR image, three RAW images with increasing exposure times are taken consecutively using the Ricoh Theta mounted on a tripod. The scenes selected for HDR capture are limited to containing no motion to achieve optimal conditions. The first image

is taken at an exposure that captures the scene's overall best coverage in terms of well-exposed pixels. This exposure is fixed as the base exposure value (EV0). Once EV0 is established, two more images are taken at EV-1 and EV+1, respectively, which correspond to exposure times that capture the underexposed and overexposed regions of EV0. In cases of underexposure and overexposure, the scene is not captured faithfully and does not reflect the content of the scene. Such pixels are substituted using the information from EV-1 and EV+1.

For image fusion, EV0 is used as the reference to determine the underexposed and overexposed pixels of the image. For any pixels in the EV0 image with a digital value below 500, the luminance reading from EV+1 is used. For pixels in the EV0 image with digital values above 3500, the luminance reading from EV-1 is used. This procedure determines the HDR fusion map in RAW pixel space from EV0 but the images are only fused after each are converted to luminance images. An example of the HDR fusion maps for each scene are in the inset of Fig. 4.4, and Fig. 4.5. Computing the luminance value for each image (EV-1, EV0, EV+1) uses the linear fit calibrated for the specific exposure and ISO settings. The fused luminance image is the final image used for per-pixel luminance measurement.

After combining the luminance image, the auxiliary software provided by Ricoh is used to unwarp the fish-eye views, resulting in a rectilinear output. The luminance composite can be effectively visualized as a heatmap denoting the luminance intensity on a per-pixel basis. Additional software from Ricoh permits the heatmap image to be interacted with in a spherical view, where mouse clicks and drags change the perspective or viewing angle from the optical centre of the Ricoh Theta.

4.1.3 Ground truth luminance measurement

To ensure the reliability and validity of the method, a ground truth metric was required to compare the pixels in the final luminance image with the measurements obtained by the SpectraScan in the scene. The photometer is used to capture 10 measurements at uniformly sampled regions anywhere in the scene from the perspective of the SpectraScan. These measurements were captured to provide a diverse range of luminance measurements that could be used for testing. For each region, detailed notes were taken regarding the measurement location and three to five measurements were averaged to limit noise. The SpectraScan utilized an internal routine similar to auto-exposure to determine the duration required for each measurement. The autoexposure function required a maximum of six seconds for darker regions and less than one second for brighter regions. In the case of scenes with natural lighting, it was crucial to conduct SpectraScan measurements quickly after the images with the Ricoh Theta were captured such that lighting did not drastically change, as this would further introduce errors between the Ricoh Theta images and the ground truth measurements as scene brightness could differ.

4.1.4 Luminance measurement to luminance image comparison

A labelling tool commonly used for object detection was used to compare the pixels in the luminance image with photometer measurements. An example of this can be seen in Fig. 4.2, where a label for each region corresponds to a single luminance measurement from the SpectraScan PR-655. The tool used to achieve this was Labelstudio [6], an open-source data labelling tool. Providing a convenient way to compare the Ricoh Theta luminance images to the photometer measurements, with no further similarities to object detection.

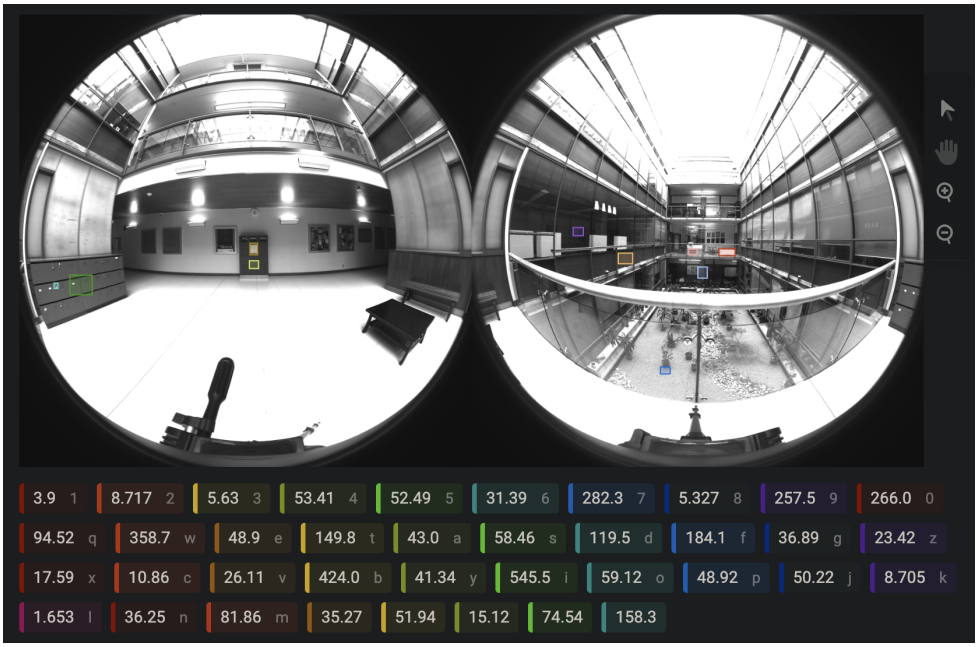

Figure 4.2: An example of the dual fish-eye image annotated in label studio. The detailed notes taken for each measurement with the SpectrScan PR-655 in the scene were translated to approximate regions using a respective bounding box.

For each region in the scene that was measured using the SpectraScan, an approximate area in the image was identified using a bounding box. The location was guided by the recorded position in the scene during the photometer measurement. Instead of the label denoting the object in the region, the ground truth was the luminance measurement from the SpectraScan.

To compare the proposed method against the ground truth, the pixels within the bounding box of the luminance image were averaged, and the average percentage error:

$$err = \frac{1}{N}\sum_{i=1}^{n}\frac{|L_{r_i} - L_{s_i}|}{L_{s_i}}, \tag{4.1}$$

was calculated between the average luminance pixel value and the measurement of the SpectraScan and N, denotes the number of measurement regions across scenes, L_{s_i} is the luminance measurement from the Spectrascan. L_{r_i} is the average of the pixels for the respective annotated region in the luminance image compute as follows:

$$L_{r_i} = \frac{1}{N}\sum_{m=0}^{i}\sum_{n=0}^{j} I_{lum_{mn}}, \tag{4.2}$$

where $i < x$, $j < y$, and x, y are the height and width of luminance image, I_{lum}, in pixels respectively. N in this case denotes the total number of pixels averaged within each luminance image region considered for testing.

Regions in the scene where all the pixels were oversaturated in every image capture of the HDR stack were excluded from the analysis. This is because the response of the sensor in these regions cannot be considered reliable, and including them would not provide an accurate representation of the performance of the method.

Label studio is used to aid in the comparison between the luminance measurement from the SpectraScan and the per-pixel luminance measurement of the Ricoh Theta. Bounding boxes are used to demarcate the location of the real-world luminance measurement onto the luminance image. The results are calculated by averaging the pixels within the demarcated regions and comparing the percentage error against the photometer measurement.

4.2 Results

The method for testing and validation described above yielded the following results for each respective test scene, listed in order of capture. To avoid inaccuracies due to the limitations imposed during calibration, very bright regions of a scene were avoided. To account for this, brighter scenes such as Vari Hall and Lassonde were chosen, but regions known to be saturated in the shortest calibration exposure were avoided with the Spectrascan photometer. Any variations in the results could be due to a combination of sensor noise, photometer measurement noise, and misalignment between optical centers of the devices across scenes. Luminance maps are shown for two scenes with the JPEG image of EV0 converted to its rectilinear format for viewing. Although the processed JPEG image appears bright, the following scenes were relatively low-light, as indicated by the low luminance values in the luminance map.

Test accuracy using LCD calibration data and single green response.			
Scene Name	Average Accuracy (%)	Average Error (%)	Average Sampled Luminance
Lab	85.94	14.06	51.52
Scott Library	80.28	19.72	96.66
Vari Hall	89.11	10.89	135.96
Lassonde	81.23	18.77	111.72
Averages	84.14	15.86	98.97

Table 4.1: The initial results of calibrating the Ricoh Theta against the LCD monitor using only the digital values of the first green channel have been obtained. It is important to note that the average sampled luminance represents the average of the luminance measurements taken by the SpectraScan for specific regions and not the average luminance across the entire scene, as a comprehensive ground truth measurement is not available.

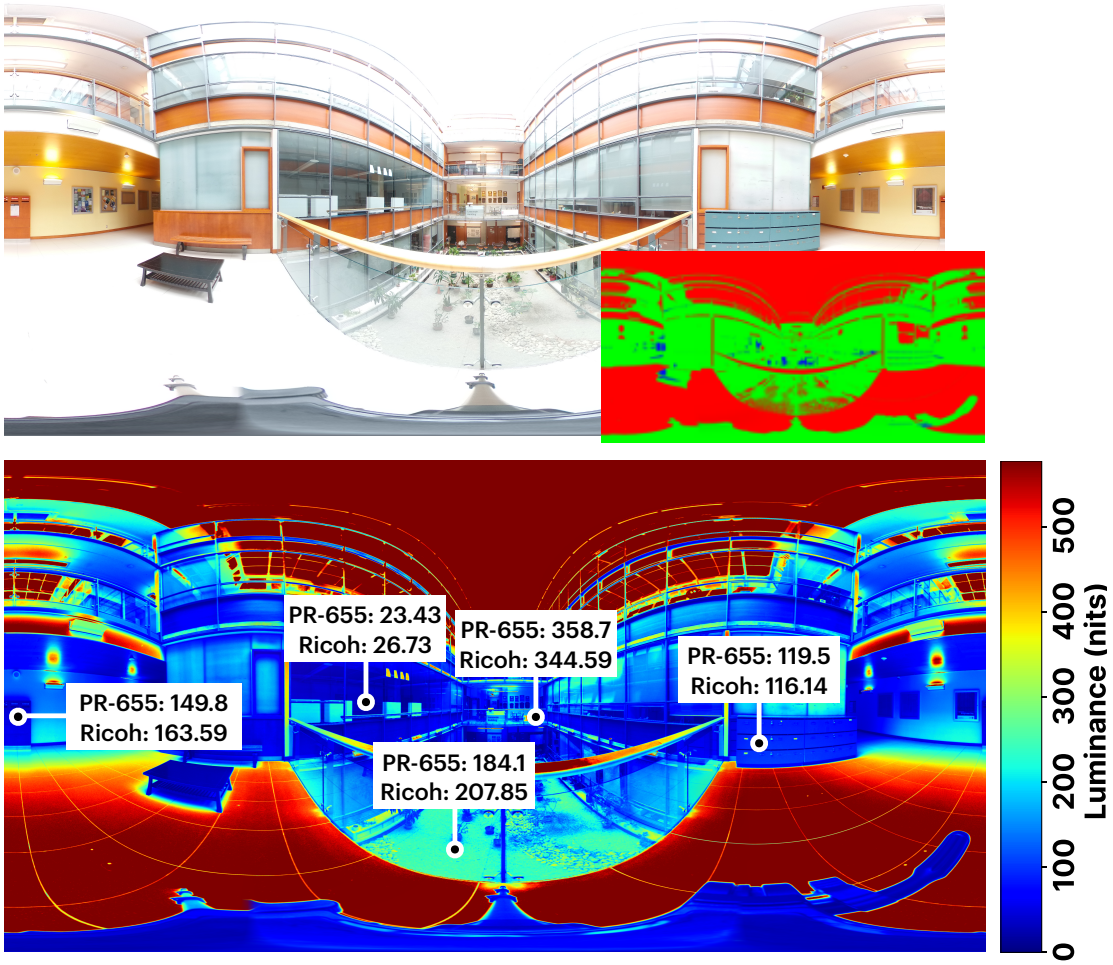

Figure 4.3: Results of the method. The inset image displays the HDR composite map, with green representing EV0, red representing EV-1, and blue representing EV+1, along with the corresponding false-coloured luminance map at the bottom. Luminance readings obtained using the *SpectraScan PR-655* are overlaid on the luminance map at the regions where measurements were made.

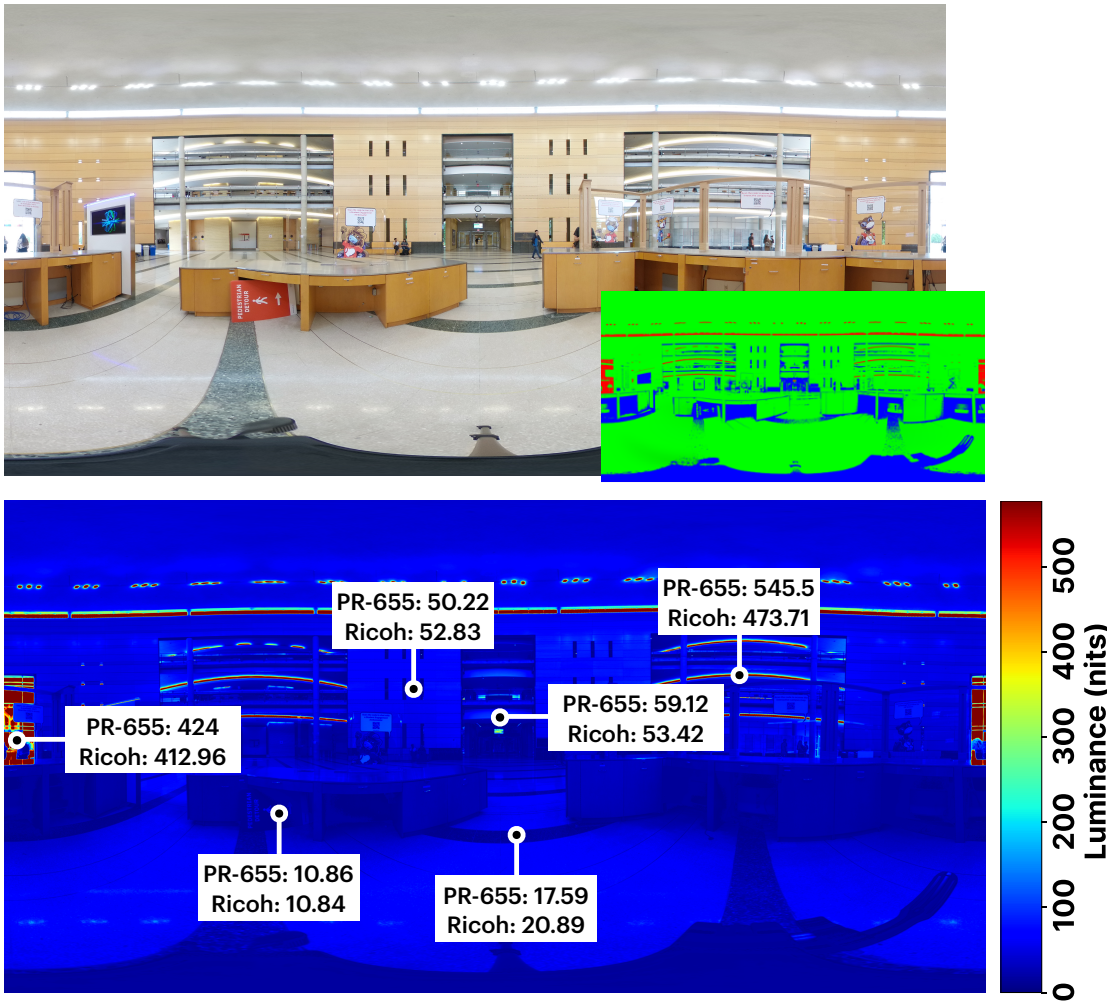

Figure 4.4: Results of the method. The inset image displays the HDR composite map, with green representing EV0, red representing EV$-$1, and blue representing EV+1, along with the corresponding false-coloured luminance map at the bottom. Luminance readings obtained using the *SpectraScan PR-655* are overlaid on the luminance map at the regions where measurements were made.

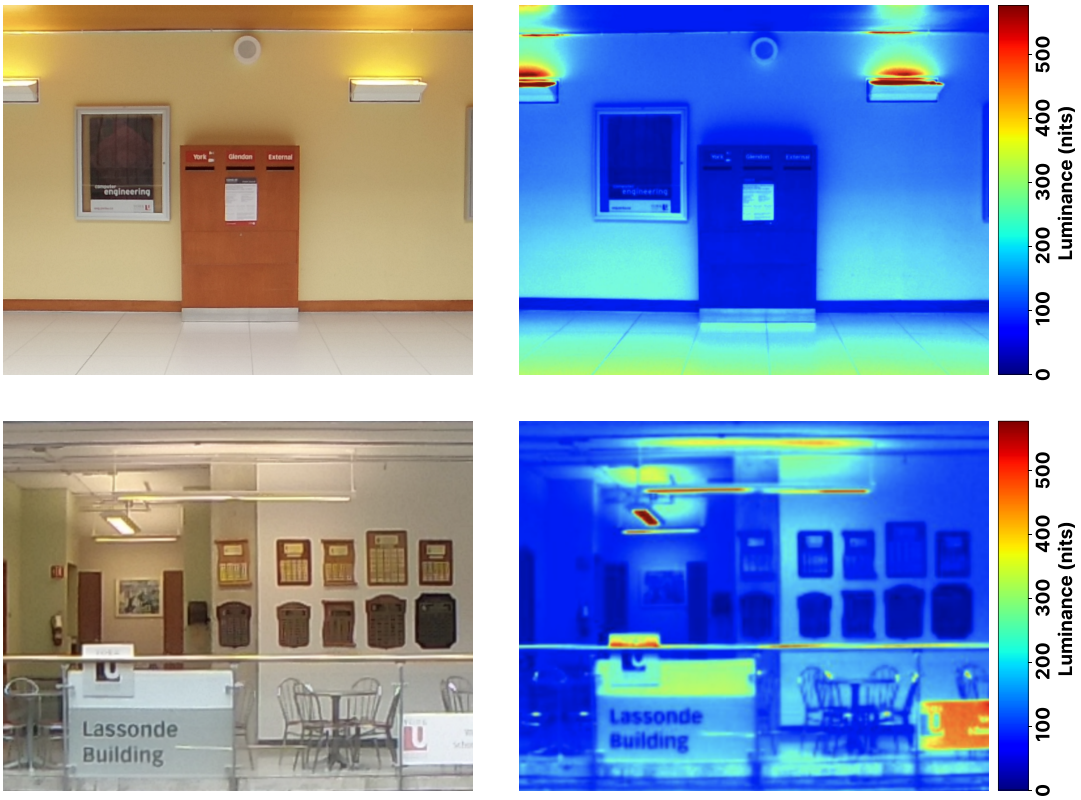

Figure 4.5: Images displayed in the Ricoh Theta auxiliary software and zoomed in to inspect a small portion of the 360° scene. The images on the left are JPEGs produced from Ricoh Theta's camera processing pipeline. On the right, are the same regions of the heatmap image for comparison.

A benefit of using the auxiliary software provided by Ricoh to view the equirectangular image in a 360° view, the luminance image can be viewed next to the image processed through a regular pipeline. An example of comparisons between an image processed using a normal pipeline, versus converted to a luminance image and visualized as a heatmap is shown in Fig. 4.5. This example shows the robustness of the Ricoh Theta device and auxiliary software to provide perspective views of the scene for enhanced visualization.

4.3 Summary

This chapter outlines the process of capturing images and ground truth photometer measurements in real-world scenes using the Ricoh Theta Z1 and Spectrascan PR-655. To accomplish this, both the Ricoh Theta and the Spectrascan PR-655 are placed on tripods in the same location for image capture and ground truth luminance measurement. Careful attention is paid to aligning the Ricoh Theta and the Spectrascan so their optical centers are aligned as closely as possible. When capturing ground truth luminance measurements with the Spectrascan, each location in the scene of the ground truth measurement is noted in fine detail to compare ground truth measurements to the luminance image. This procedure along with its results was published and presented at the 30th Colour and Imaging Conference [14].

Chapter 5

Conclusions and Future Work

This thesis presents the overall success of a low-cost luminance meter in providing dense 360° measurement of a scene using HDR imaging methods. The work provides a proof of concept for this device, but it is important to note that it still lacks colourimetric measurement capabilities. This work focuses only on omnidirectional luminance measurement, and as such the calibration and test dataset focus specifically on optimizing and validating this task. Some additional complexities are introduced for colourimetric measurement and are left for future work. Additional work beyond the scope of this thesis will be discussed in Sec. 5.2, and discuss the need for a new calibration source and what is required to perform colourimetric measurements.

5.1 Discussion

Although the results obtained from this thesis are promising, it is important to note that they serve as a proof of concept for using consumer sensors for luminance measurements.

The calibration range of the LCD monitor had a significant influence on the location of measurements with the spot photometer and reduced the depth of the comparison that could be made during testing. While scenes with higher average luminance measurements were included in the testing data, the average measurements within each scene remained low due to the limitations of the calibration range and difficulties of capturing ground truth data with the photometer.

Moreover, time constraints during ground truth data capture might have led to variations in the luminance between the Ricoh Theta and SpectraScan measurements, particularly in scenes with natural lighting. It could be the case that in some scenes, natural lighting changes the brightness in the scene between when the Ricoh Theta images are captured, and the ground truth photometer measurements are made. This could also be the case between respective SpectraScan measurements due to the internal autoexposure requiring up to six seconds per measurement. Each photometer measurement needed to be captured with the understanding that the Ricoh Theta could be compared at the respective point, which led to the avoidance of bright regions if they were expected to be overexposed in the 360° image. Capturing ground truth measurements more quickly or in controlled environments could improve the accuracy of the results. Calibrating the method against a light source with a higher, or no refresh rate, and increased luminance output will provide a detailed calibration of the sensor, removing the limitation of the LCD monitor.

5.2 Future Work

This section highlights potential areas for future work aimed at addressing the limitations encountered during the development of the proposed method for creating a luminance measurement device using the Ricoh Theta Z1.

The calibration process using an LCD monitor had a limitation in terms of the range of exposures that could be effectively calibrated. To overcome this limitation, a light source with a much faster or no refresh rate will be required for the method to be more robust across scenes. For this, the Philips Hue lightbulbs may be helpful and provide the additional benefit of being controllable via an API.

While colour filters are useful for colour reproduction, they may not be optimal for luminance measurement before correction. The spectral responses of the Ricoh Theta's colour filters do not necessarily align well with the photopic luminous efficiency function beyond the green channel. This raises the question of whether a calibration which solves the system of linear equations between the spectral response of the sensors and the photopic luminance efficiency function would improve the results.

5.3 Conclusion

This thesis introduces a novel approach for measuring luminance in a 360° omnidirectional manner by utilizing RAW image capture from the Ricoh Theta Z1 camera. The calibration process described enables the conversion of the RAW digital values captured by the sensor into per-pixel luminance measurements allowing for dense environmental luminance measurement at a lower cost compared to traditional spot photometers. The calibration

procedure along with the method for capturing ground truth data and correlating it to the image pixels provides a foundation for an omnidirectional photometer using a Ricoh Theta Z1.

Given the challenges in physically aligning the optical centers of the Ricoh Theta and an auxiliary device and the discrepancy between the time of image capture with the Ricoh Theta and subsequent photometer measurements, validating the results remains difficult. Establishing a calibrated environment with controllable luminance is necessary for robust validation. However, despite these challenges, the work in this thesis is expected to inspire further interest in developing 360° omnidirectional luminance meters using existing commercial hardware and expanding the work to include colourimetric measurements.

Bibliography

[1] S. Aghayari, M. Saadatseresht, M. Omidalizarandi, and I. Neumann. "Geometric Calibration of Full Spherical Panoramic Richo-Theta Camera". *Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 237–245, 2017.

[2] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. "Non-Local Means Denoising". *Image Processing On Line*, 1:208–212, 2011.

[3] P. E. Debevec and J. Malik. "Recovering High Dynamic Range Radiance Maps from Photographs". *SIGGRAPH*, 1997.

[4] Mauricio Delbracio, Damien Kelly, Michael S. Brown, and Peyman Milanfar. "Mobile Computational Photography: A Tour". *Annual Review of Vision Science*, 7: 571–604, 2021.

[5] A. El Gamal and H. Eltoukhy. "CMOS image sensors". *IEEE Circuits and Devices Magazine*, 21:6–20, 2005.

[6] Heartex Inc. Open source data labeling, 2023. URL `https://labelstud.io/`. Accessed: 2023-02-20.

[7] K. Hirakawa and T.W. Parks. “Joint demosaicing and denoising”. *IEEE Transactions on Image Processing*, 15:2146–2157, 2006.

[8] Keigo Hirakawa and Thomas W. Parks. “Adaptive homogeneity-directed demosaicing algorithm”. *Transactions on Image Processing*, 14:360–369, 2005.

[9] Image Egineering. camSPECS, 2023. URL `https://www.image-engineering.de/products/equipment/measurement-devices/588-camspecs-express`. Accessed: 2023-02-20.

[10] Jadak SpectraScan PR-655. `https://www.jadaktech.com/products/photo-research/spectrascan-pr-655/`, 2023. Accessed: 2022-06-12.

[11] H. C. Karaimer and M. S. Brown. “A Software Platform for Manipulating the Camera Imaging Pipeline”. In *European Conference on Computer Vision*, 2016.

[12] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard. “What’s new in psychtoolbox-3”. *Perception*, 36:1–16, 2007.

[13] Konica Minolta LS-160. `https://sensing.konicaminolta.us/us/products/ls-160-luminance-meter/`, 2022. Accessed: 2022-06-12.

[14] Ian MacPherson, Richard F. Murray, and Michael S. Brown. A 360° omnidirectional photometer using a ricoh theta z1. *Color and Imaging Conference*, 2022.

[15] S. Mann and R. W. Picard. “On Being ‘Undigital’ With Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures”. In *Proceedings of IS&T*, 1995.

[16] Y. Morgenstern, W. Geisler, and R. Murray. "Human Vision Is Attuned to the Diffuseness of Natural Light". *Journal of Vision*, 14, 2014.

[17] R. F. Murray and W. K. Adams. "Visual perception and natural illumination". *Current Opinion in Behavioral Sciences*, pages 48–54, 2019.

[18] A. A. Mury, S. C. Pont, and J. J. Koenderink. "Representing the light field in finite three-dimensional spaces from sparse discrete samples". *Applied Optics*, 48:450–457, 2009.

[19] E. Reinhard, G. Ward, S. N. Pattanaik, P. E. Debevec, and W. Heidrich. *"High Dynamic Range Imaging - Acquisition, Display, and Image-Based Lighting"*. Academic Press, 2010.

[20] Ricoh Developer Connection. Api reference. `https://api.ricoh/docs/theta-web-api-v2.1/`, 2022. Accessed: 2022-06-12.

[21] Navid Sarhangnejad, Nikola Katic, Zhengfan Xia, Mian Wei, Nikita Gusev, Gairik Dutta, Rahul Gulve, Peter Zhi Xuan Li, Hui Feng Ke, Harel Haim, Manuel Moreno-García, Kiriakos N. Kutulakos, and Roman Genov. "Dual-Tap Computational Photography Image Sensor With Per-Pixel Pipelined Digital Memory for Intra-Frame Coded Multi-Exposure". *IEEE Journal of Solid-State Circuits*, 54:3191–3202, 2019.

[22] Hanlin Tan, Huaxin Xiao, Yu Liu, and Maojun Zhang. "Two-Stage CNN Model for Joint Demosaicing and Denoising of Burst Bayer Images". *Computational Intelligence and Neuroscience*, 2022:6200931, 2022.

[23] Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin. "Deep learning on image denoising: An overview". *Neural Networks*, 131:251–275, 2020.

[24] C Tomasi and R Manduchi. "Bilateral filtering for gray and color images". In *International Conference on Computer Vision*, 1998.

[25] S. Tominaga and N. Tanaka. "Measurement of Omnidirectional Light Distribution by a Mirrored Ball". In *Color and Imaging Conference*, 2001.

[26] Gurjot Kaur Walia and Jagroop Singh Sidhu. "hybridized KNN-Random Forest Algorithm: Image Demosaicing with Reduced Artifacts". *National Academy Science Letters*, 45:517–520, 2022.

[27] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". *IEEE Transactions on Image Processing*, 26:3142–3155, 2017.