Master in Computational Colour and Spectral Imaging (COSI)

High Dynamic Hyperspectral Imaging:
An empirical investigation of tone mapping on spectral radiance for HDR scene

Master Thesis Report

Presented by

Aqsa Hassan

and defended at the

Norwegian University of Science and Technology

September 2023

Academic Supervisor(s): Dr. Giorgio Trumpy
Host Supervisor: Dr. Mekides Assefa Abebe and Dr. Susan Farnand
Jury Committee:

1. Dr. Dmitry Semenov, University of Eastern Finland, Finland.
2. Dr. Arash Mirhashemi, University of Eastern Finland, Finland.

Submission of the thesis: 10th August 2023
Day of the oral defense: 5th September 2023

Abstract

High dynamic range (HDR) images require tone-mapping to scale the large range of luminance information that exists in the real world so that it can be displayed on a device that is capable of only a limited dynamic range in luminance, to adequately reproduce their perceptual qualities. Here the main objective is to preserve the original HDR scene appearance, including contrast, sharpness, and color with the focus on overcoming the limitation of the output media to achieve the best match given the limited gamut and dynamic range. To the best of our knowledge, these HDR rendering algorithms have been proposed only for traditional three-channel (RGB) images. It is well known that having multispectral data can produce better color-accurate images, but whether the same thing can be said for tone mapping too or not, has not been explored yet. In addition, limited research work regarding tone mapping for HDR hyperspectral images is itself a challenge whereas there is no such publicly available database for HDR hyperspectral images.

To address these issues, this master's thesis aims to investigate the effective utilization of spectral radiance for improved color fidelity and tone-accurate reproduction of HDR images. To achieve this, an HDR hyperspectral radiance image has been captured using an approach similar to the multiple exposures HDR technique which helps to significantly recover the details of a high dynamic range (HDR) scene both in the dark and bright areas, overcoming the problem of capturing underexposed and overexposed data. The HDR hyperspectral absolute radiance image is further improved by applying the proposed linearity correction method while hyperspectral interpolation has been performed to verify and account for the missing wavelengths due to pixel saturation. It also introduces a spectral image color appearance model titled SiCAM which is the first of its kind, designed for tone mapping a HDR hyperspectral radiance image to a three-channel Low Dynamic Range (LDR) image. It is to be noted that SiCAM is inspired by the iCAM06 image color appearance model where we adapt the iCAM06 for hyperspectral input by embedding a spectral adaptation method rather than chromatic adaptation as in the case of iCAM06 which requires only three-channel RGB input.

Additionally, a psychophysical experiment has been conducted for perceptual evaluation of the proposed HDR rendering method to determine the effectiveness of having more spectral data for the tone mapping of HDR images in comparison to the performance of iCAM06 and the gamma operator. Besides, the objective image quality assessment (IQA) of these reproduced LDR images has been performed. The results from both subjective and objective evaluation indicate that SiCAM outperformed the other two HDR rendering methods in terms of both accurate color appearance and pleasantness. Finally, we presented a dataset containing four HDR hyperspectral radiance cubes and their respective three-channel HDR images as our contribution to future research in this domain.

Acknowledgment

I am eternally grateful to the Almighty Allah for guiding me through this journey and providing me with the wisdom and knowledge to complete my master's thesis.

I would also like to express my sincere gratitude to my dedicated supervisors Giorgio Trumpy, Mekides Assefa Abebe, and Susan Farnand. Your invaluable advice, consistent support, and expert insights have shaped the trajectory of this novel and challenging research. Thank you for being an integral part of my academic endeavor and contributing in completing my thesis.

I am indebted to many people who provided help and support during this research. I would like to thank Ivar Farup for highlighting that we need to convert camera spectral sensitivity to quantum efficiency to get a fair comparison. I am grateful to Fereshteh Mirjalili for providing me monochromator to perform camera characterization and also for reviewing my findings. I also like to acknowledge Federico Grillini for providing helpful suggestions regarding the hyperspectral camera and also providing valuable feedback on the methodology of data acquisition. Furthermore, I am grateful to Miguel Angel for his invaluable guidance and support during the initial phases of this research topic. My appreciation extends to Jana Blahova and Gregory High for providing access to the instrument and other necessary resources to conduct the research work.

Special thanks to Mark Fairchild for reviewing the proposed SiCAM model and for providing useful theoretical insights that were extremely helpful in proposing this model. I extend my heartfelt gratitude to Ph.D. student Saeedeh Abasi for sharing the HDR display. A special thanks goes to Stephanie Livingston-Heywood who facilitated me through my research at MCSL. I would also like to express my gratitude to the entire research team at the Munsell Color Science Lab(MCSL) for creating an inclusive research environment.

I wish to express my gratitude to all esteemed professors within the COSI program who introduced me to the fascinating subject of color science, enriched my understanding of the field, and solidified my passion for color science. I am also deeply grateful to all my colleagues and friends of COSI, especially to Ali Raza, Akash Harijan, and Ehsan Ullah for their thoughtful and constructive discussions on refining my research methodology. Your input has been invaluable in enhancing the rigor and depth of my work.

And most of all, I would like to thank my parents and family for their encouragement, belief in my abilities, and continuous motivation. Your love and support have been my driving force.

Acronyms

CCD	Charged Coupled Device
CAM	Color Appearance Model
CFA	Camera Filter Array
CLUT	Color Look-Up Table
CMF	Color Matching Function
CMOS	Complementary Metal-Oxide Semiconductor
CMOS	Complementary-metal-oxide Semiconductor
CRF	Camera Response Function
CIE	Commission internationale de l'éclairage
CSF	Contrast Sensitivity Function
CRT	Cathode Ray Tube
EOTF	Electro-optical Transfer Function
FOV	Field of View
HDR	High Dynamic Range
HDRI	High Dynamic Range Imaging
HSI	Hyperspectral Image
HVS	Human Visual System
IC	Integrated Circuit
ICC	International Color Consortium
ISP	Image Signal Processor
iCAM	Image Color Appearance Model
IQM	Image Quality Metric
IQA	Image Quality Assessment
LDR	Low Dynamic Range
PCS	Profile Connection Space
SDR	Standard Dynamic Range
SPD	Spectral Power Distribution
SiCAM	Spectral Image Color Appearance Model
TRC	Tone Response Curve
TMOs	Tone Mapping Operators

Contents

1 Introduction

The One who illuminates and reveals the divine light of all the worlds...

Almighty

High dynamic range imaging (HDRI) is a set of techniques that computationally enhance the conventional or standard dynamic range of an image and/or sequence of images, which has recently become quite prevalent by being more produced, distributed, and consumed in different applications such as photogrammetry, entertainment, computer vision, computer graphics, animation, etc. The ultimate goal of HDRI is to mimic the perception of the human visual system (HVS) which has already revolutionized the way digital visual content is processed. Human vision is capable of adapting to lighting conditions of nearly 10 orders of magnitude (Hoefflinger, 2007; Ferwerda et al., 1996). Nevertheless, the HVS can operate only within a limited portion of this wide range at any given instance. The static dynamic range, which is observed when the HVS is fully adapted, is commonly accepted to be approximately 10,000:1 as given in Figure 1.1. Not only are real-world scenes brighter and more vibrant than their digital reproductions, but they also have much higher contrast, both locally between adjacent objects and globally between distant ones. The HVS has evolved to handle such high contrasts that are present in nature which evokes essential perceptual cues. Unlike HDRI, conventional imaging is unable to reproduce such high-contrast scenes.

The dynamic range of a display represents the ratio of luminance between the brightest white and the darkest black that it can reproduce. A standard dynamic range (SDR) display employs a 2.2 gamma electro-optical transfer function (EOTF) often with 8 bits of color depth. This restricts the dynamic range of the traditional SDR display to two orders of magnitude while a high dynamic range (HDR) display

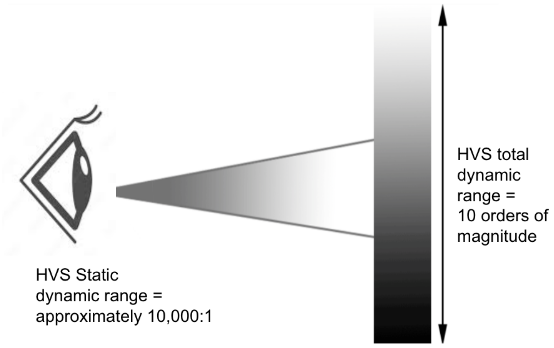


Figure 1.1: *The HVS exhibits the remarkable ability to adapt to a wide range of lighting conditions up to 10 orders of magnitude. However, it should be noted that the HVS has the capability to function within a restricted segment of its wide at any given moment. The usually accepted static dynamic range, observed under the full adaptation of the human visual system (HVS), is roughly 10,000:1.*

employs an EOTF curve that is extended on both ends (i.e., PQ curve), with a minimum color depth of 10 bits. In the case of HDR display, it is relatively achievable to exceed the maximum luminance from 1000nits to 4000nits and beyond (Reinhard et al., 2010). Color depth refers to how precisely we can encode color information. The color channel of the SDR display can be encoded up to 256 levels (8-bit) for each of the three primary colors (red, green, and blue) that are used to synthesize the entire color gamut of these additive displays. Using this 8 bits encoding range, one can specify up to 16,777,216 colors in conventional SDR display monitors. In contrast, most HDR standards enhance the precision with which colors can be specified from 8-bit per channel to 10-bit, resulting in a total of 1,073,741,824 colors. These ranges demonstrate the advancements in HDR displays intending to replicate the wide luminance ranges perceived by the HVS, surpassing the limitations of traditional SDR displays(Eilertsen et al., 2017b; Mantiuk et al., 2007). The HDR displays allow the perception of a larger range of luminances and a higher bit depth compared to conventional standard dynamic range displays. These features make human vision to have more plausible and realistic representations due to the clear perception of details in both the darker and brighter areas at the same time. The HDR images are considered to be scene-referred, which means that an HDR image stores the radiometric quantity i.e., the radiance of the actual HDR scene that is only scaled by the sensor's sensitivity. This is to ensure that the dynamic range compression and other adjustments are carried out just during the display phase, and only if a device is incapable of accurately reproducing the information. With its greater contrast ratio, HDR demonstrates more details in extremely bright and dark scenes, as well as highly saturated and vibrant colors.

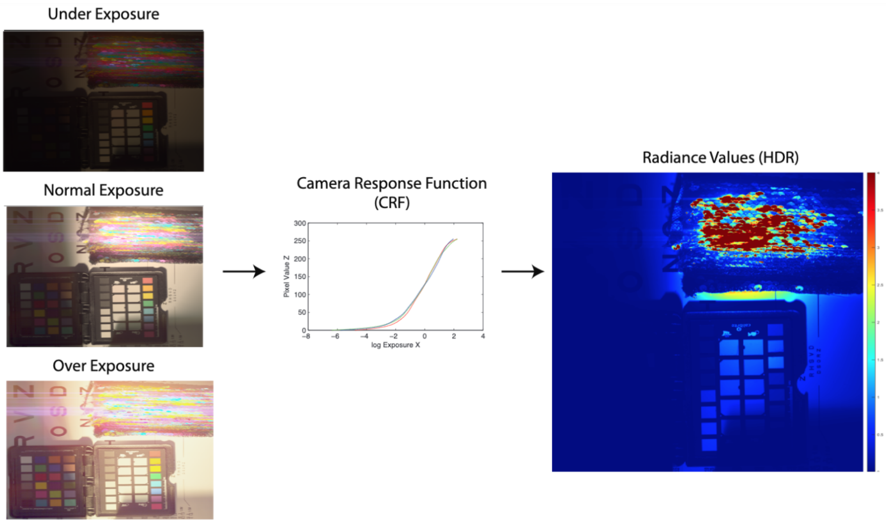


Figure 1.2: *The multiple exposure HDR capture utilizes a set of LDR images taken at different exposure settings such that all the details of the HDR scene are captured. Later, these LDR images are used to recover the camera response function which is employed for linearity correction. Once the camera response function is obtained, the LDR images are transformed into linear representations. These linear LDR images are then combined or merged together to create an estimation of the radiance map of the actual HDR scene.*

Consequently, HDR images appear more realistic and stand out more noticeably. As depicted in Figure 1.2, the multiple exposure HDR capture is the most widely used technique of HDR imaging where multiple shots of an HDR scene are captured sequentially at varying exposures (Reinhard et al., 2010; Burt and Kolczynski, 1993; Tomaszewska and Mantiuk, 2007)

On the other hand, hyperspectral imaging (HSI) is one of the advanced imaging technologies which is beyond the limitations of HVS. While HVS is limited to only a narrow range of the electromagnetic spectrum (Malacara, 2003), HSI extends beyond this limitation by capturing a broader range of spectral information, including wavelengths that are outside the visible range. It provides significantly more information i.e., a wider spectrum per data point of the scene than RGB images captured by a normal color filter array-based camera. The normal cameras only operate on acquiring three channels of RGB image corresponding to red, green, and blue regions of the visible spectrum. Because each material has a unique spectral signature that can be used as its identification, HSI greatly improves the capability to classify objects based on their spectral properties. With the advent of HSI, it has been widely used in the field of remote sensing in recent years (Shukla and Kot, 2016) while it has become an emerging technology in medical diagnosis and image-guided surgery (Glaubitz et al., 2014). The field of cultural heritage has also involved HSI systems for the analysis and restoration of artworks (Pouyet et al., 2018; Kubik, 2007; Linhares et al., 2008). Additionally, deep learning has already demonstrated its efficacy with hyperspectral data and produced state-of-the-art results on hyperspectral classification (Audebert et al., 2019; Polak et al., 2017). As a result of recent developments in sensor technologies and efficient computation capabilities, hyperspectral imaging devices have evolved from relatively slow and unreliable research prototypes to reliable, accurate, and precise analytical instruments.

In past research, limited studies have been found integrating hyperspectral data to improve the quality and color accuracy of HDR image reconstruction and vise-versa. In an attempt to utilize hyperspectral information for enhancing HDR reproduction, Martínez et al. (2019) proposed a framework for capturing spectral reflectance through HDR hyperspectral imaging and multi-focus stacking which was applied for the analysis of artwork. They used three different focus settings while for each focus, three different exposure settings had been used to obtain the spectral bands ranging from 400nm to 1000nm. This work was focused on retrieving spectral reflectance data of artwork such as paintings, documents, posters, photographs, etc. Lapray et al. (2017) also presented an interesting work that merged HDR into spectral imaging by proposing an HDR spectral imaging pipeline for a multispectral filter array camera which was intended to estimate relative multi-spectral radiance of HDR scenes. The authors took multispectral captures at three integration settings and merged them by using a hat weighting function i.e., giving more weightage to middle-range intensity pixels. However, it is important to note that the final output images were not the absolute radiometric radiance of an HDR scene. A novel hyperspectral visualization approach based on high dynamic range imaging was presented by Ertürk et al. (2014) which aimed to retain visual detail and provide a superior result in terms of visual quality when converting hyperspectral images to three-channel color images. Anand Swamy et al. (2022) also proposed lossless compression of hyperspectral images by employing decorrelation and a multiscale HDR technique. In addition, Wang et al. (2018) propose an HDR 3D measurement method based on spectral modulation and hyperspectral imaging.

Besides, the concept of spectral adaptation was first introduced by Fairchild (2007b) in the literature where the author applied it to the full spectral reflectances, while Khan et al. (2017, 2018) attempted to generalize it to the sensor measurements. Also, note that the HDR was not incorporated in any of these studies. Khan, H. A et al. proposed the idea of multispectral constancy, which allowed for the acquisition of multispectral images independent of illumination. A spectral adaptation transform (SAT) (Khan et al., 2018) was utilized to change data representation from the actual sensor domain to a canonical one to achieve multispectral constancy. The simulation results demonstrated that a diagonal SAT was helpful to estimate similar reflectance reconstruction as when the samples are acquired under the calibrated illumination. By achieving multispectral constancy, multispectral imaging can be used with any unknown illuminant without the need for re-calibration when the imaging environment is changed. However, when the SAT was evaluated based on an estimate of illumination, errors in illuminant estimation significantly reduced the performance of the proposed approach. Furthermore, it is unknown what level of accuracy is required for illuminant estimation for this concept to be useful for computer vision applications.

1.1 Motivation

HDR images require tone mapping to adjust the wide range of luminance found in real-world HDR scenes. It involves compressing these wide ranges of luminance which is necessary to display the HDR image on devices with a limited dynamic range, while still preserving its original appearance, including contrast, sharpness, and color. The primary goal is to overcome the limitations of the output media and achieve the best possible match, considering the restricted gamut and dynamic range of the display device. To the best of our knowledge, existing HDR rendering methods (Reinhard et al., 2010; Banterle et al., 2017; Eilertsen et al., 2017b; Mantiuk et al., 2007; Salih et al., 2012) have been proposed only for traditional three-channel (RGB) images. However, it is well known that having spectral data can produce better color-accurate images, but whether the same thing can be said for tone mapping or not has not been explored yet. There are also certain problems with hyperspectral data acquisition such as bands being unevenly affected by noise at different exposures while sensitivities of different wavelengths are also not uniform. Therefore, significant postprocessing of the raw hyperspectral cubes is required to reduce the effects of noise and to obtain accurate radiances. Other relevant issues include the influence of chromatic aberration and glare effect especially for HDR scene captures (McCann and Rizzi, 2007). In addition, limited research work regarding tone mapping for HDR hyperspectral is itself a challenge whereas there is no such publicly available database for HDR hyperspectral images.

In this context, the main research objective of our study is to investigate how the visual quality perception and realism in HDR-rendered images are affected by using HDR hyperspectral data instead of traditional three-channel HDR images. It is also motivated by the fact that although trichromatic vision serves as the first stage of color vision processing, it might be possible that the higher levels of the visual system have some access to spectral information which may complement trichromatic mechanisms. This research endeavor has the potential to advance the field of high dynamic range imaging and contribute to the development of more sophisticated rendering techniques across various applications of imaging technologies.

1.2 Research questions

To commence our research, we first highlighted our research questions which are as follows:

- **Research question 1:** By leveraging the advantages of spectral radiance in the context of tone mapping, can we achieve superior results when dealing with high dynamic range content?

- **Research question 2:** What methodology to use for capturing HDR hyperspectral data? To address research question 1, it is important to develop the methodology for capturing HDR hyperspectral data considering factors such as hardware requirements, calibration techniques, and data processing methodologies. This involves acquiring spectral bands covering the entire visible spectrum, enabling us to gather detailed spectral information for each pixel in the HDR scene. Because no HDR hyperspectral data is publicly available for our application, developing the methodology for capturing such data is crucial to support our research objective.

- **Research question 3:** How to effectively utilize hyperspectral radiance data for better tone-accurate HDR image rendering with the intent to adequately reproduce their perceptual impression of physical HDR scenes? In other words, how to integrate the HDR and hyperspectral data for improving HDR reproduction?

- **Research question 4:** How to render these HDR hyperspectral images on HDR and SDR display? This involves addressing the challenges posed by the varying dynamic ranges and color gamuts of HDR and SDR displays.

1.3 Major contributions

Given the mentioned issues and objectives, the major contributions of this research are as follows:

- The proposed method for HDR hyperspectral image acquisition and post-processing workflow.

- A novel hyperspectral image color appearance model titled SiCAM has been introduced which is the first image color appearance model designed for HDR hyperspectral radiance images. It is to be noted that the proposed model is inspired by the three-channel based image color appearance model iCAM06 (Kuang et al., 2007a).
- Another main contribution of this work is the utilization of hyperspectral adaptation which has been incorporated into SiCAM for generating tone-accurate LDR images from HDR hyperspectral radiance cube as an input.
- Both subjective and objective image quality assessment of the reproduced LDR images has been performed to determine the effectiveness of having more spectral data for tone mapping.
- Considering the lack of a widely accessible standard pipeline for rendering HDR content of arbitrary luminance ranges, we proposed a display-referred HDR rendering workflow that effectively translates HDR images into the specific range of luminance supported by the HDR display.
- As a contribution to extended research in this domain, a dataset containing four HDR hyperspectral radiances cubes and their respective three-channel HDR images has been proposed where each of them consisted of a different HDR scene captured in an indoor setting.

1.4 Thesis outline

This master's thesis report is organized as follows: In Chapter 1, the theoretical background of color science is discussed which is necessary to understand the objectives and methodology of this study. Chapter 2 addresses the fundamentals of image formation along with an overview of HDR and hyperspectral imaging which are an integral part of this research study. The details of HDR hyperspectral data acquisition and post-processing are given in Chapter 3. The proposed HDR imaging rendering pipeline for HDR displays and its comparison with the HDR standards are also discussed in this chapter. Chapter 5 provides a concise overview of HDR tone mapping followed by the proposed SiCAM model for HDR rendering. The HDR display calibration and characterization to ensure faithful HDR reproduction and to avoid potentially adverse impacts during the perpetual evaluation of the proposed model is summarized in Chapter 6. The details of the psychophysical experiment and its methodology for perceptual-based evaluation of SiCAM with that of iCAM06 and gamma tone mapping are given in Chapter 7. Additionally, Chapter 8 presents a summary of HDR image quality assessment (IQA), and also discussed the metrics employed for evaluating the tone-mapped images involved in psychophysical study. Lastly, the results obtained from the data analysis of the rating scores given by the observers along with the results of objective assessment are presented in Chapter 9. This chapter also includes a discussion of the findings and highlights the limitation of this research study. The poster presented at MCSL open house session during the Color Impact Conference 2023 is attached in the appendix. The observer consent form used during the psychophysical experiment and the CITI program completion certificate is also attached to the appendix.

Acknowledgement: I would like to thank Dr. Mekides Assefa for providing the initial code for display calibration and the template for the GUI of the psychophysical experiment. These codes were further modified and tested according to the use case of this thesis work. Besides, I would like to acknowledge the use of AI-powered tools for grammar checks such as Quillbot, Grammarly, etc.

2 | Fundamentals of color science

Like beauty, color is in the eye of the beholder

Yogi Bera

The ultimate goal of electronic vision is to mimic the capabilities of human vision and beyond in some aspects, as highlighted in Chapter 1. Since human vision as a whole is so versatile and powerful, any approach for extracting the perceptual cues of human vision with electronics and computer vision technologies requires an understanding of how to model some of the higher and lower-level processing involved in human vision to collect the most comprehensive information about our physical world. Likewise, the emerging field of high dynamic range (HDR) imaging is intrinsically such an interdisciplinary field that it involves not just imaging formation concepts but also revolves around radiometry, photometry, colorimetry, color appearance, etc., to deal with certain characteristics of light and human perception. Hence, this chapter summarizes a brief overview of these fundamental concepts of color science, which are an integral part of the research involved in this master's thesis. It is important to note that this chapter offers a concise introduction to the realm of color science and its associated concepts. For a more comprehensive understanding, please refer to (Wyszecki and Stiles, 2000; Hunt and Pointer, 2011; Fairchild, 2013; Schanda, 2007).

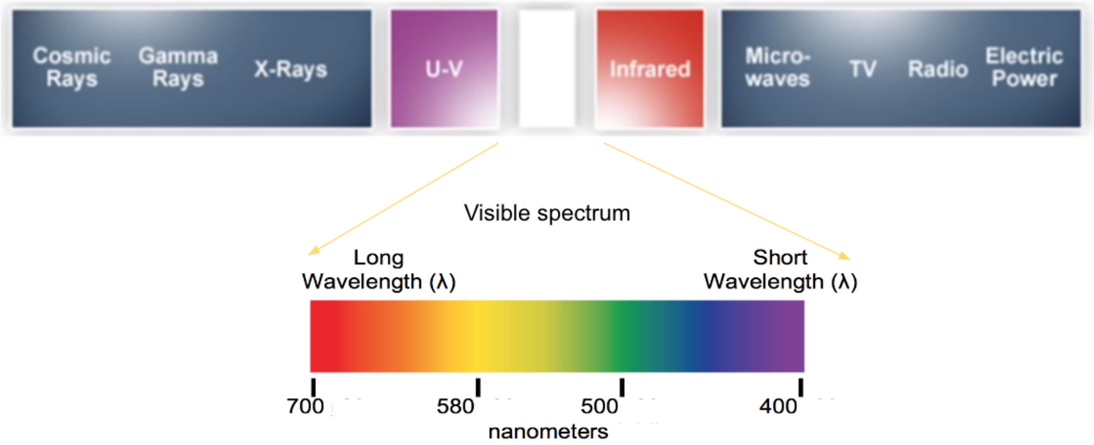


Figure 2.1: *The full electromagnetic spectrum of light consists of ultraviolet, visible, and infrared regions where the visible spectrum is within the range wavelengths from 380nm up to 780nm.*

2.1 Theory of light: Radiometry and Photometry

Light is electromagnetic radiation that propagates in space and interacts with different materials in our environment. This interaction of light with the materials results in several optical phenomena where it can be absorbed, refracted, reflected, and transmitted depending on the properties of the material and the characteristics of the light. The full electromagnetic spectrum of light involves ultraviolet, visible, and infrared regions, where the visible spectrum of light lies within the range of 380nm to 780 nm, as illustrated in Figure 2.1. Radiometry and photometry are two different categories of light measurements, which are further discussed in the following subsection.

2.1.1 Radiometry

Radiometry deals with the measurement and quantification of electromagnetic radiation across the entire electromagnetic spectrum, including visible light, ultraviolet light, infrared light, and beyond. It is concerned with the absolute measurement of light energy in terms of power, radiant flux, radiant intensity, and other radiometric quantities as given in the table 2.1. Typically, these measurements are expressed in watts (W) or joules per second (J/s).

Table 2.1: *The summary of radiometric quantities along with their photometric counterparts (Wyszecki and Stiles, 2000).*

Radiometric Quantities			**Photometric Quantities**		
Quantity	*Symbol*	*Units*	*Quanity*	*Symbol*	*Units*
Radiant Power	Φ_e	W	Luminous Flux	ϕ_v	$lumens(lm)$
Radiant Intensity	I_e	W/sr	Luminous Intensity	I_v	lm/sr
Irradiance	E_e	W/m^2	illuminance	E_v	lm/m^2
Radiance	L_e	$W/m^2/sr$	Luminance	L_v	$lm/m^2/sr$

2.1.2 Photometry

Photometry is a subfield of radiometry that focuses on the measurement of light as perceived by human vision, where the radiometric quantities are scaled by the spectral response of the human eye. It takes into account that HVS has different sensitivity at different wavelengths of visible light as defined by the Commission Internationale de l'Eclairage (CIE) as given in Figure 2.2. This curve is known as the photopic luminous efficiency curve (V(λ)) where it is found to be highly sensitive at a wavelength of about 555nm. Hence, all the radiometry qualities defined in the previous table are scaled by the V(λ) to convert into photometric quantities, which are essentially the spectrally weighted radiometric quantities by taking into account the human vision sensitivity of visible light.

It is important to note that photometric measurements do not tell us how bright a stimulus will appear because the appearance of the brightness (i.e., the perceived luminance) is also dependent on viewing conditions, whereas in any one set of viewing conditions, the brightness and photometric measures are not proportional (Hunt and Pointer, 2011). This means that increasing the luminance will not result in increased brightness, and vice versa. Additionally, stimuli of the same luminance appear brighter by increasing their colorfulness, while the surroundings of the stimuli also impact their appearance. These examples indicate that beyond radiometric and photometric measurements of light, there are other more complex phenomena involved in the human vision system (HVS), which are briefly discussed in the following section.

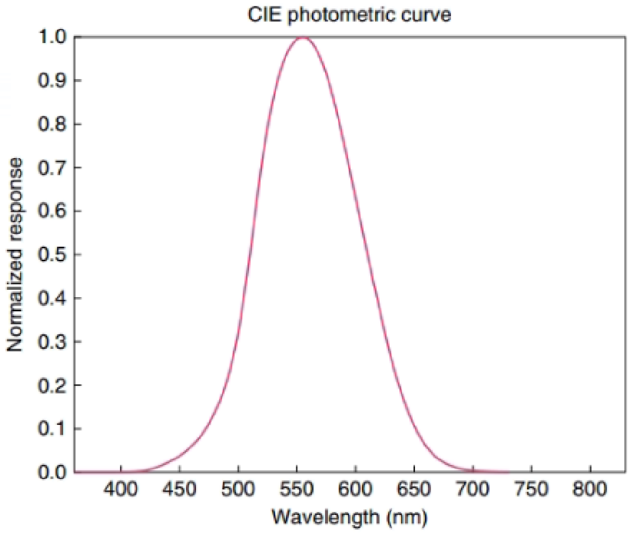


Figure 2.2: *The figure illustrates the photopic luminous efficiency curve ($V(\lambda)$) defined by the Commission Internationale de l'Eclairage (CIE) which is the spectral sensitivity curve of human vision (Reinhard et al., 2010).*

2.2 Turning light into color: Human visual response to light

In order to develop a better understanding of HDR imaging, it is essential to understand how our visual system works. Illuminants (light sources), which are identifiable by their unique spectral power distributions (SPDs), objects, which are identifiable by their own material properties, and the human eye are typically the three main stages involved in human visual perception. When the light from some illuminant hits the object, based on the material's properties, the incoming light interacts with the object and is then reflected back in the surrounding area. The reflected light then finally reaches the human eye, as depicted in Figure2.3.

When light from the environment enters the human eye, the HVS is able to interpret various characteristics of that light. The most fundamental characteristic is the intensity of the light (power or luminance), which we perceive as brightness. The HVS is also sensitive to variations in the intensity of light. The spatial variations in the intensity of light produce contrast, which enables us to perceive lines, contours, and, finally, the shapes of objects in the environment, while the spectral variations in the intensity of light allow us to discern the object's colors.

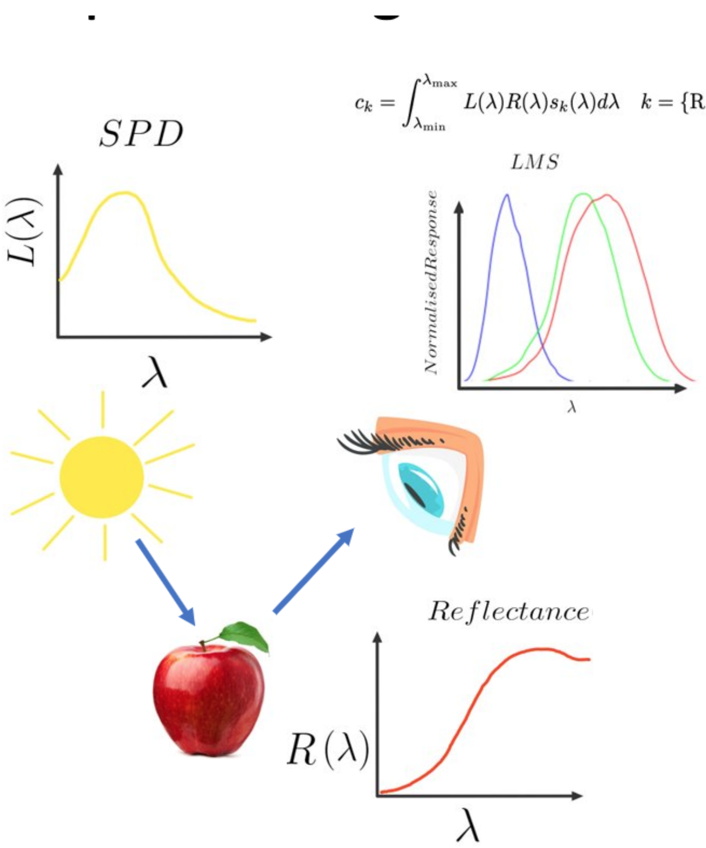


Figure 2.3: *The figure illustrates the three fundamental components of color vision perception which includes: illuminant, object and human eye.*

2.2.1 The human visual system:

Figure 2.4 illustrates the schematic diagram of the human eye along with the labeling of key structures. The human eye accumulates incoming light on photoreceptors, which then convert this light into biological signals. These signals are transmitted via the optical nerve to the visual cortex, a region of the brain that processes these signals to produce the perception of the physical scene. Light entering the eye is first transmitted through the cornea, a transparent membrane. The light then enters the pupil, which is an aperture modified by the iris, a muscular diaphragm. The lens then refracts light, which strikes photoreceptors in the retina. There are two liquids inside the eye: the vitreous and aqueous humors (Banterle et al., 2017). The former occupies the eye, maintaining its form and the retina's contact with the inner wall. Between the cornea and the lens, the aqueous humor maintains intraocular pressure.

Photoreceptors are specialized cells in the retina that detect light and convert it

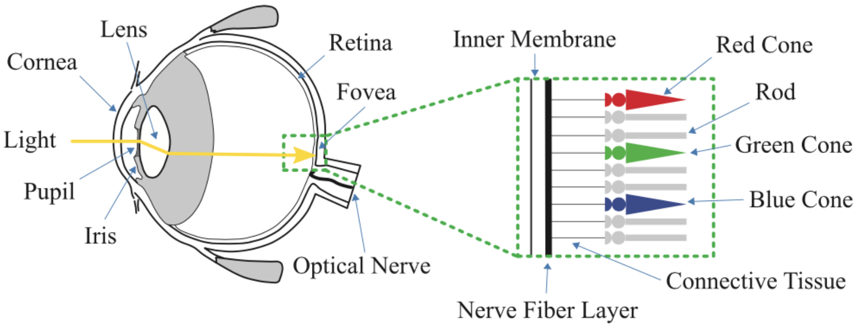


Figure 2.4: *The human eye. (Banterle et al., 2017).*

into signals, which are conveyed to the brain via the optic nerve. The cones and rods are the two categories of photoreceptors, where rods are sensitive at low luminances, producing significantly low visual acuity. This vision, when the rods are active, is known as scotopic vision. The cone photoreceptors are sensitive at higher luminance ranges, producing significantly high visual acuity. This vision, where cones are activated, is known as photopic vision, which allows color perception. Mesopic vision is where both types of photoreceptors contribute to vision at intermediate luminance levels. There are far more rods than cones; however, approximately 6–7 million cones are primarily located in the fovea, where we perceive the best spatial and color vision, and the fovea is the point on the retina where the projection of objects on which we fixate our gaze falls.Furthermore, color vision is the result of signals generated by three types of cones: short-wavelength cones (S), which are sensitive to wavelengths around 435 nm; middle-wavelength cones (M), which are sensitive to wavelengths around 530 nm; and long-wavelength cones (L), which are sensitive to wavelengths around 580 nm. These signals generated by photoreceptors are subsequently processed by the brain to create a visual impression of the world around us and are primarily responsible for color vision perception.

2.2.2 Basics of color vision

When cones are activated in the presence of light, the photoreceptors in the retina generate the signals that lead to color vision. However, there are several theories that attempt to explain transformations of these signals which serve as useful insights to understand the overall function of human color vision. Some of the most relevant mechanisms to our discussion are briefly described as follows:

2.2.2.1 Trichromatic vision theory

As stated above, cones are the photoreceptors that are primarily responsible for color vision and serve as the first stage of visual processing for color perception. Scientists discovered that there are three types of photoreceptors, namely long (L), middle (M), and short (S) cones, which are sensitive to the red, green, and blue regions of the visible spectrum, respectively. The trichromatic theory of color vision is also referred to as the Young-Helmholtz theory (Hunt and Pointer, 2011).

The trichromatic nature of color vision differs significantly from person to person. In addition to color vision deficiencies due to the lack of a particular cone photopigment and anomalous trichromacy, large individual variability also exists in normal color vision (Asano et al., 2016). Thus, a pair of two different spectral stimuli may create the same trichromatic response for one observer but a mismatch for another observer. This phenomenon is known as observer metamerism. Typically, observer variability would not be a significant issue in many observation settings, owing to the fact that natural spectra are broad-band and do not require side-by-side comparison. However, variability in color vision and observer metamerism, on the other hand, could be a severe concern in color-critical applications such as cross-media color matching (Long and Fairchild, 2014). Due to the narrow band primaries of modern wide color gamut and HDR display technologies, observer metamerism has become an even more serious problem, and normal observers may perceive different colors for the same color values shown on the calibrated displays, which may appear mismatched (Bergquist, 2008; Hung, 2019).

2.2.2.2 Opponent colors theory

The trichromatic vision theory is based on the assumption that the three cone photoreceptors generate three separate images of the physical scene and transmit them to the brain for further processing. However, the assumption of sending three images to the brain is not sufficient to explain other dominant phenomena. Hence, the opponent color theory explains that the three separate color images generated by LMS cones are not transmitted directly to the brain, but there is a further transformation of these three types of signals into opponent signals (i.e., red-green and yellow-blue). These opponent signals consist of an achromatic response generated by the sum of three individual signals from L, M, and S cones, while two chromatic responses are generated by taking the difference of these cone responses, i.e., red-green is formed by $L - M + S$ whereas yellow-blue is formed by $L + M - S$ as shown in Figure 2.5.

The encoding of signals from LMS to the achromatic and chromatic signals serves to decorrelate the color information carried in LMS responses, which allows for more efficient signal transmission and reduces difficulties in their efficient

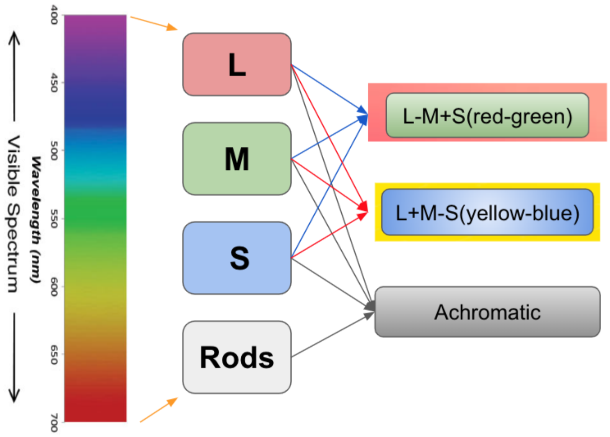


Figure 2.5: *The opponent colors theory argues that the human visual system analyzes color information by processing opponent signals from a combination of responses generated by cone and rod photoreceptors as illustrated in this diagram.*

processing. Based on this phenomenon of decorrelation, there are several image encoding algorithms that are inspired by opponent color theory and incorporate the principle of decorrelation for efficient data compression and transmission.

2.2.2.3 Adaptation Mechanisms

In addition to the transformation from trichromatic vision to two opponent color signals and an achromatic signal, the dynamic mechanism of human visual adaptation by adjusting its sensitivity to the level of illumination to further optimize the visual response to certain viewing conditions. These adaption mechanisms are discussed as follows:

- **Dark adaptation** It refers to the adjustment of visual sensitivity that occurs when the prevailing intensity of illumination is reduced. For instance, when walking into a darkened room on a sunny afternoon, at first everything appears dark, but eventually, we recover our sight and become able to see the objects around us. This is because the visual system takes time to adapt by becoming more sensitive in low-light conditions after being exposed to a brighter environment.

- **Light adaptation** The inverse of dark adaptation is light adaptation where the visual system becomes less sensitive due to the presence of high illumination in order to produce useful perception.

- **Chromatic adaptation** In addition to the illumination adaptation, human

> vision also adapts to color of the object or its surround. Chromatic adaptation is a tendency of the human visual system to disregard the color of the illumination and maintain the appearance of an object to a reasonable degree. In other words, the magnitude of change in the appearance of objects is far lesser than the magnitude of the change in its radiometry or colorimetry. Examining a white object under different forms of illumination, such as daylight and incandescent, reveals chromatic adaptation. The visual system achieves color constancy by discounting scene illumination. As an approximation, we can say that the perceived chromaticity of the scene is normalized by removing the perceived illumination. Although, it is safe to assume that human vision has a natural tendency to correct for the effects of light source color but the mechanism underlying this ability is not fully understood yet.

The adaptation to intensity is analogous to exposure setting in electronic cameras and chromatic adaptation can be thought of as analogous to the automatic white balancing.

2.2.2.4 Spatial and temporal properties of color vision

Contrast sensitivity is one of the most essential visual measurements. For an object to be distinguishable from its background, the object's luminance or color must differ from that of the background. This distinction is referred to as contrast. The contrast in the field of visual perception is determined by the difference in color and/or luminance between an object and its background or other objects within the same field of view.

Contrast sensitivity is a measurement of the ability to distinguish between different levels of luminance in a static image. The contrast Sensitivity Function (CSF) is defined by the threshold response to a function of spatial or temporal frequency. The contrast sensitivity function that is obtained by measuring contrast thresholds over a range of spatial frequencies is illustrated in 2.6 , where it can be observed that the spatial CSF for luminance is bandpass in nature. It also demonstrates that human vision is most sensitive to an intermediate range of spatial frequencies (i.e., between 2 and 8 cycles per degree). For comparison, luminance and chromatic CSFs for both spatial and temporal frequencies in the logarithmic domain are given side by side in Figure 2.7. It can be observed in this figure that the chromatic response is lowpass in nature for both temporal and spatial frequencies.

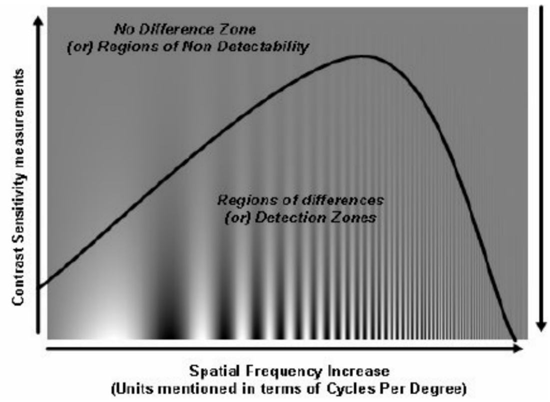


Figure 2.6: *The figure shows the spatial contrast sensitivity for the luminance of human vision which is bandpass in nature where the human vision is most sensitive to an intermediate range of spatial frequencies.(Image credits: Lecture slides of color imaging course(NTNU).)*

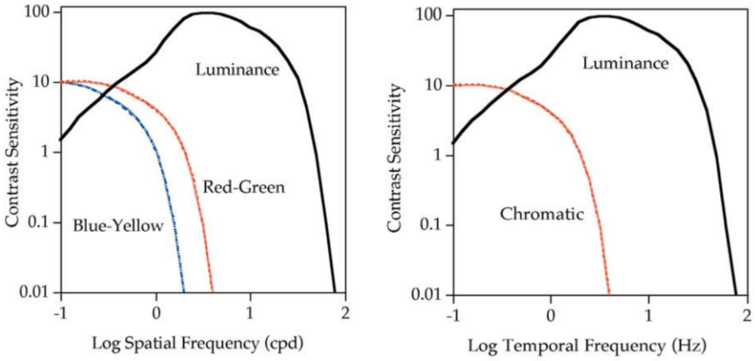


Figure 2.7: *The figure represents the spatial luminance and chromatic CSFs on the left side whereas the temporal luminance and chromatic CSFs are given on the right sideFairchild (2013).*

2.2.2.5 Important perceptual attributes of color

For developing an understanding of the fundamental scientific concepts, it is crucial to learn the definitions of the commonly used terms in that field. Hence, the definitions of basic perceptual attributes of color are given as follows:

- **Brightness vs Lightness:** The attribute of a visual sensation according to which an area exhibited more or less light is known as brightness whereas the brightness of an area judged in proportional to the brightness of a similarly illuminated area is known as lightness.

$$\text{Lightness} \;=\; \frac{\text{Brightness}}{\text{Brightness (Reference\_white)}} \tag{2.1}$$

- **Colorfulness vs Chroma:** The attribute of a visual sensation according to which an area exhibited more or less of its hue whereas the chroma is the colorfulness of an area judged relative to the brightness of a similarly illuminated area.

$$\text{Chroma} \;=\; \frac{\text{Colorfulness}}{\text{Brightness (Reference\_white)}} \tag{2.2}$$

- **Hue** An area appears to be similar to one of the four unitary hues: red, yellow, green, and blue, or a combination of two of them

- **Saturation** The colorfulness in proportion to its brightness is known as saturation.

$$\begin{aligned}
\text{Saturation} &= \frac{\text{Chroma}}{\text{Lightness}} \\
\text{Saturation} &= \frac{\text{Colorfulness}}{\text{Brightness (Reference\_white)}} \times \frac{\text{Brightness (Reference\_white)}}{\text{Brightness}} \\
\text{Saturation} &= \frac{\text{Colofulness}}{\text{Brightness}}
\end{aligned} \tag{2.3}$$

2.3 Theory of Colorimetry

In the midst of all these complex mechanisms involved in the perception of color by human vision, it is a challenging task to effectively communicate and measure colors. As discussed in the previous sections, color is determined by the underlying biology of the human visual system, particularly the spectral characteristics of the three types of light-sensitive cone photoreceptors in the eye. The trichromatic nature of color and color matches means they can be specified by three linearly related variables. The primary objective of colorimetry is to be able to specify these variables for all spectral compositions of light. (Schanda, 2007). Because color is a perceptual phenomenon, engineering measurements cannot be used to quantify colors. Consequently, colorimetry is concerned with the quantitative description of the color of visual stimuli.

2.3.1 Color matching functions

Trichromacy implies that the color-matching behavior of an individual can be examined by determining the intensities of three independent fixed-wavelength "primary" lights to match a target stimulus where independent means that no two primaries will match the third. By performing a series of color matches for the entire spectrum of visible light, one can estimate the color-matching function (CMF) for an observer. The maximum saturation method and Maxwell's method are the two principal methods that attempt to estimate a standard CMF. As a result of these color-matching experiments, a system of colorimetry was introduced based on the principles of Grassmann's laws that color matching must be linear and additive (Wyszecki and Stiles, 2000). The main idea of this system is that the color matches can be specified in terms of the intensities of three additive primaries required to visually match the color of a stimulus as given by the equation 2.4 where C is the unknown stimulus and a, b and c indicates the amount of the RGB primaries required to match C.

$$[\mathrm{C}] \equiv a[\mathrm{R}] + b[\mathrm{G}] + c[\ \mathrm{B}] \tag{2.4}$$

The researchers performed this color-matching experiment where they projected a monochrome light from the visible range on the reference field and asked the observers to match the color of this reference field by modifying the test field that is usually made up of an additive mixture of three independent primaries as illustrated in the Figure 2.8.

In some instances, a match cannot be made, necessitating the addition of light to the reference field, resulting in negative values for the color-matching functions. Although any three independent primaries can be used, monochromatic red (long wavelength), green (medium wavelength), and blue (short wavelength) are typically chosen.

$$[\mathrm{C}] = a \cdot [\mathrm{R}] + b \cdot [\mathrm{G}] + c \cdot [\mathrm{B}]$$

$$a = \int_{380\ \mathrm{nm}}^{780\ \mathrm{nm}} r(\lambda)P(\lambda)\mathrm{d}\lambda \quad b = \int_{380\ \mathrm{nm}}^{780\ \mathrm{nm}} \bar{g}(\lambda)P(\lambda)\mathrm{d}\lambda \quad c = \int_{380\ \mathrm{nm}}^{780\ \mathrm{nm}} \bar{b}(\lambda)P(\lambda)d\lambda \tag{2.5}$$

Repeating the same color-matching process of the whole visible spectrum as given by equation 2.5, resulted in CMF. The spectral response curves of an average human eye is defined by CIE in 1931 based on the independent experiments of William David Wright and John Guild in the late 1920s. The data was averaged and presumably based on observers with normal vision this then became the basis

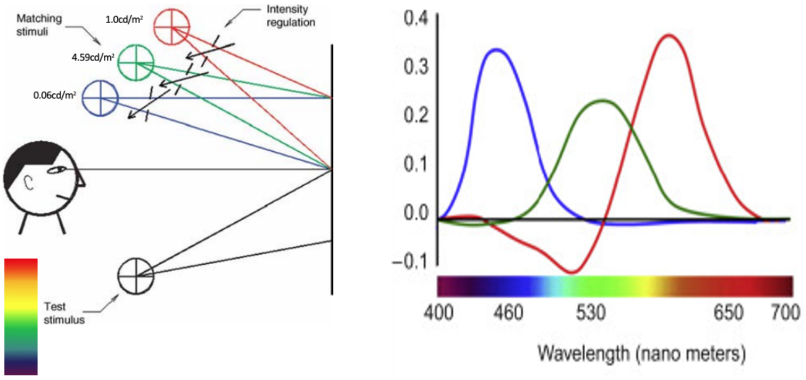


Figure 2.8: *The figure illustrates the color matching experiment where three primary colors are used to match or reproduce the color of a target stimulus. This is often done by adjusting the intensities of these primary lights until their combination creates the same color sensation as the target stimulus. By performing a series of color matches across the entire spectrum of visible light, a color-matching function (CMF) for that particular observer can be determined. (Fairchild, 2013; Schanda, 2007).*

for what we call the standard observer. The negative values in the CMFs made the calculation more difficult and hence CIE decided to transform from the real rbg primaries to CIE XYZ. It is important to note that even in the early 1900's it was understood that the field of view had an impact on color perception and a 2° field of view was selected to help neutralize this variable and that is why the standard observer model is often referred to as a CIE 1931 2° standard observer. Most of the display monitors work on the principle of additive color theory where the linear combination of the display primaries defines entire color gamut of that display.

2.3.2 Color spaces

A color space is a mathematical explanation of how to describe color which is basically a mapping between numeric values and specific colors. It is a three-dimensional model containing all the combinations of the three fundamental independent vectors known as primary colors. Color spaces are classified into two types: device-dependent and device-independent which are as follows.

2.3.2.1 Device-dependent color space

This defines the color information with respect to the hardware and software capabilities of the device for reproducing the color. In the case of a display monitor,

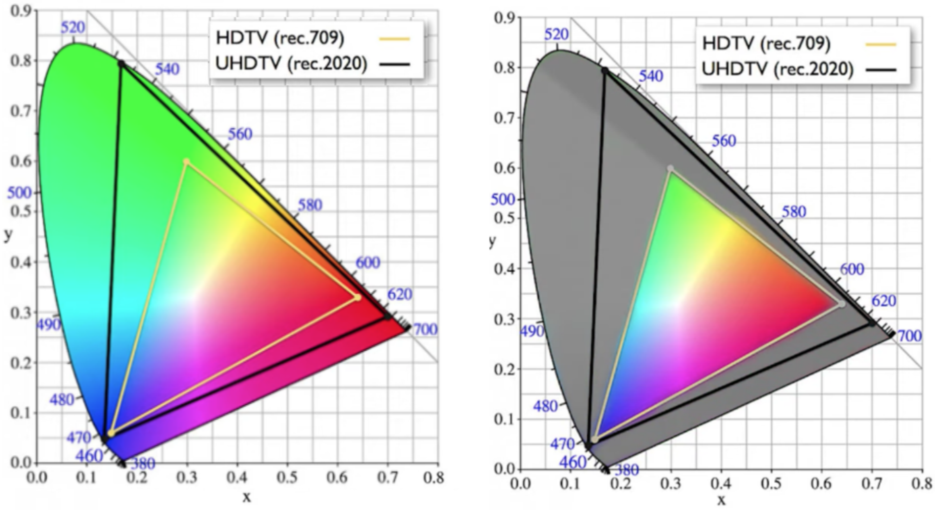


Figure 2.9: *The figure represents CIE chromaticity diagram. All the colors which are in the spectral local(boundaries of the horseshoe) can be perceived by the human eye, where we performed the mapping of a 3D color space into 2D to obtain this CIE XY chromaticity diagrams using 1931 CMF. However, due to technological limitations, the display standards define how much of this visible gamut can be achieved by a device. As shown on the chromaticity diagram given on the right side of the figure, the colors that are within the gamut of REC709 are only reproducible by the display which is designed to operate on this standard colorspace.*

it is determined by the set of primary phosphors, whereas it is determined by the set of primary inks in an ink-jet printer. For example, display monitors usually operate on sRGB color space while prints work on CMYK color space.

2.3.2.2 Device-independent color space

The device-independent color space is the superset of all colors visible to the human eye based on the CIE 1931 2° standard observer and hence used by all color management software for the transfer of color information between different devices in its workflow. Examples of such color spaces are CIE XYZ, CIE Luv, and CIE Lab.

2.3.3 Color difference

Color is a crucial attribute that defines the appearance of objects, and understanding color differences is essential for describing these appearances. In industrial and practical applications, these differences hold significant importance. Color difference models play a key role in quantifying the disparities between visual perceptions of color (ΔV) and the differences calculated using specific formulas (ΔE). The visual color difference is obtained through psychophysical experiments, whereas the computed difference is determined using appropriate formulas. CIELAB color differences are the most used (Hill et al., 1997). They are based on perceptually uniform color space and they optimize the hue linearity. It is similar to a Euclidean distance metric, so it is easy to calculate. Even so, depending on the application, the perceptually uniform CIELAB does not always works because it still contains some issues regarding the blue and purple non-linearity.

Another widely used color difference is the CIE $\Delta E2000$ color difference formula Luo et al. 2001, which includes compensation for the non-linearity of CIELAB. This formula was recommended by the CIE for industrial applications as it has outperformed previously proposed color difference metrics. Although CIE $\Delta E2000$ is considered better than CIELAB it also has its own shortcomings which have to be resolved in the future (Melgosa et al., 2008). For example, (Kirchner et al., 2015) found out that in many cases the CIE $\Delta E2000$ formula predicts the color differences to be equal while the observers can perceive smaller color differences. Moreover, they concluded that new color difference equations are needed to obtain more accurate predictions.

2.3.4 Color appearance model

Color appearance models (Fairchild, 2013) are mathematical models which are designed to provide the prediction about how humans perceive the appearance of colors under specific viewing conditions. These models consider various factors such as illumination, background, viewing angle, and the human visual system's response to different wavelengths of light. The color appearance models take into account at least five attributes of uniform color stimulus which include brightness, lightness, colorfulness, hue, and saturation.

Some models, such as CIELAB (and its lastest version, CIELUV) (Mahy et al., 1994), offer a relatively straightforward method for describing and quantifying color differences and appearances. These models are extensively employed for a variety of practical applications, such as color quality control, color matching, and color space transformations. Complex color appearance models, on the other hand, account for a broader spectrum of perceptual phenomena where color constancy and simultaneous contrast phenomena are incorporated. CIECAM02 (Moroney et al., 2002) and CAM02-UCS (Li et al., 2017) are examples of more comprehensive and complex models of color appearance that attempt to account for a broader range of perceptual effects.

3 The fundamentals of Imaging

In the legacy of 'What you can see, you can capture.'

Erannman Camera

This chapter provides a historical overview of photography, allowing us to grasp the remarkable progress in imaging technology, from its origins to its current trajectory. Subsequently, this chapter presents the fundamental workflow of the digital camera pipeline, highlighting its key stages and processes. Additionally, the factors influencing image quality are discussed, acknowledging the limitations that arise within this context.

In the latter part of the chapter, we provide concise summaries of two advanced imaging techniques: high dynamic range imaging (HDRI) and hyperspectral imaging(HSI). Both of these imaging techniques are an integral part of our research study. These cutting-edge methods offer exciting possibilities for capturing and processing images with enhanced dynamic range and spectral information, respectively.

3.1 A brief history of Photography

For nearly 200 years, photography has allowed human beings to capture the fascination of a single moment in time. Thanks to the inventive minds that have long been looking for ways to improve or indeed to perfect photography. The fundamental principle of image formation is the mapping of the 3D world around us into the 2D plane. In the early days, painting was the dominant medium in the art world to capture these 2D visual representations. Later, the invention of cameras allowed capturing light with a digital sensor or film, to create an image. The

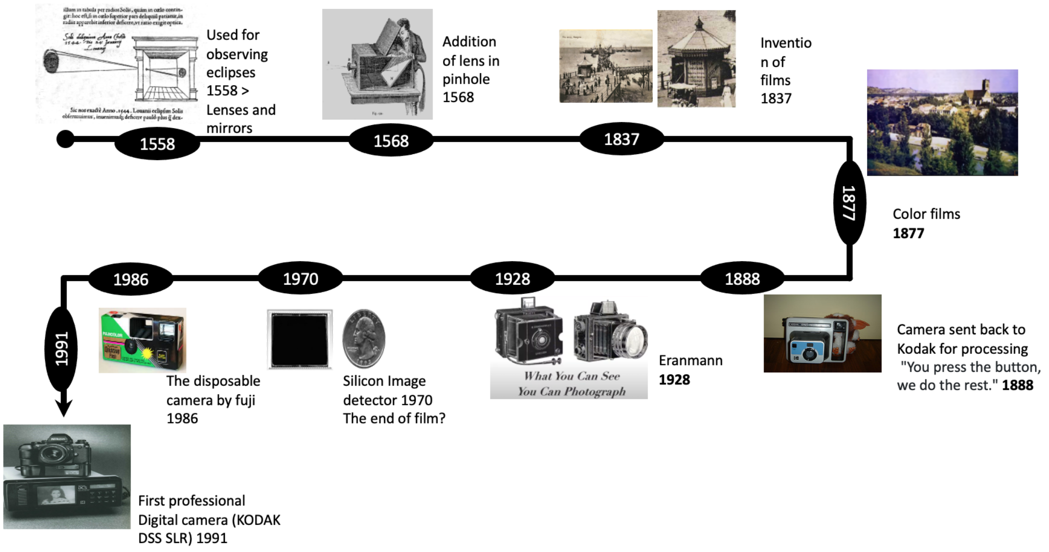


Figure 3.1: *The above figure depicts a concise chronology of major developments in the field of photography.*

evolution of the simplest imaging device is so fascinating at every different point in its history. While our understanding of the camera has evolved significantly, the "camera obscura" is regarded as the ancient concept upon which all subsequent revolutionary developments and inventions in the field of photography were built as illustrated in Figure 3.1. This involves light propagation through a small pinhole that results in an upside-down mapping of a real-world scene on a 2D plane. This was essentially designed to facilitate the artist to draw images based on perspective projection. Sooner it was discovered that by using a smaller pinhole the projection may appear more sharper but also dimmer. The pinhole allows less light to pass and hence requires long exposures to capture a bright image.

Therefore, the pinhole camera was enhanced with the addition of an optical lens. The use of a lens allowed much more light into the camera obscura and, with a mirror placed at 45 degrees to the rear wall, meant that the image could be projected onto an opaque glass screen on the top side of the camera. This enabled the artists to trace the projected image onto paper placed on the screen. Although the advancement in this imaging device brought a dramatic shift in painting through an increase in realism, particularly the more naturalistic treatment of faces. However, one can consider the invention of films as a significant development in the history of photography which gave artists the freedom to capture the real optical look of a 3D scene and made it possible to shoot multiple pictures one after the other. It basically opened the doors of photography for everyone. It is worth noting that capturing an entire photograph with full colors is an extremely challenging task. Photographic technology took an exciting step forward with advances in color film. The consumer's film cameras soon appeared in the market with the advisement of "what you can see, you can capture".

The invention of silicon image detection paved the road to digital photography. The advent of digital imaging sensors such as CCD (Charge-Coupled Device) and CMOS (Complementary Metal-Oxide-Semiconductor) sensors brought the transition from Film photography to digital imaging that broaden the horizons of visual communication.

3.2 Digital camera pipeline

All digital camera involves three main components which include the optics, the imaging sensor, and the processing module. These imaging devices operate on a similar underlying principle where they take in light as an input and convert it into photons which are eventually transformed into electrons through the photoelectric effect of semiconductors. At a certain point, the electrons are converted into voltages, and then these voltages are subsequently converted to digital numbers. The typical pipeline of digital cameras is illustrated in Figure 3.2.

Light enters the camera through the lens which focuses and directs light onto the imaging sensor of the camera. The lens's quality and design characteristics are critical in determining the image's sharpness, clarity, and overall performance. The imaging sensor receives incoming light and converts the optical image created by the lens into an electrical signal. The camera uses either CMOS or CCD sensors depending on the requirements of their application. However, CMOS sensors are widely used in consumer cameras. It is necessary to transform the analog electrical signals produced by the image sensor into digital information. An analog-to-digital converter (ADC) executes this process by quantizing the analog signal into digital values and sampling the analog signal at discrete intervals. The RAW file format is used to store the digital values obtained from the ADC. These RAW files are sensor data that has undergone very little processing, preserving all of the data that was recorded. Compared to processed image formats like JPEG, RAW files give photographers more flexibility during post-processing because they contain more dynamic range and color information.

The Bayer filter array is used in most camera sensors, with each pixel capturing only one of the three primary colors (red, green, or blue). The process of reconstructing full-color information for each pixel by interpolating the missing color information based on the values of neighboring pixels is known as demosaicing Li et al. (2008). After the demosaicing, the camera applies color correction algorithms Gasparini and Schettini (2003) to guarantee accurate color representation and noise reduction algorithms to eliminate electronic noise and sensor artifacts that may

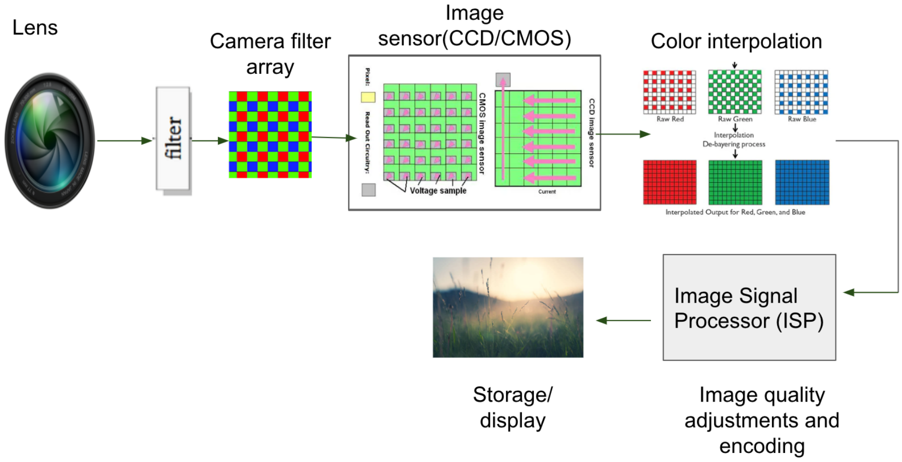


Figure 3.2: *The given pipeline illustrates a generic workflow of digital cameras.*

have been introduced during image capture, particularly in low-light situations. The image signal processing (ISP) stage consists of a number of image processing algorithms that include gamma correction, white balance, sharpness enhancement, and contrast adjustments. These stages refine the appearance of the final image and optimize it for display. Finally, the camera may apply lossy compression algorithms Yang and Bourbakis (2005) to reduce the file size and then store it on the camera's memory card. Depending on the camera settings, this could be a RAW file, JPEG file, or both.

3.2.1 Image quality factors

During the early year of digital imaging, it was the image sensor with its limited number of pixels which was considered as the main obstacle for capturing images with higher resolution. Today, high resolution relies on the high quality of camera optics to produce fine details, especially in the image corners. Due to various optical distortions such as vignetting and chromatic aberration, the image quality decreases toward the periphery compared to the image center.

Additionally, an imaging sensor can generate a wide range of noise artifacts, including thermal noise, readout noise, and shot noise Bigas et al. (2006). These artifacts diminish image clarity and are more apparent in images captured with a high ISO setting or in low-light conditions. Moire patterns also appear when the scene being captured contains repeating textual or lines that are tightly spaced. The pixel pattern of the sensor can interact with these patterns, resulting in visual interference or false colors in the image. Many cameras use a rolling shutter to read data from the sensor line by line. A rolling shutter is used by many cameras to read data from the sensor line by line. When photographing fast-moving individuals or scenes, the rolling shutter effect can generate distortions such as skew or ghosting artifacts because various parts of the scene are captured at different instances. Image sensors have a limited dynamic range, which is the luminance ratio between the darkest and brightest parts that can be captured in a single exposure. The real-world scenes with a dynamic range in luminance can cause the highlights and shadows to be clipped, leading to a loss of details. It's worth noting that advances in sensor technology and image processing techniques are aimed at reducing these limitations of digital imaging.

3.3 Advanced imaging techniques: High dynamic range imaging

The aspiration of electronic vision is to replicate the functionalities of the human eye and potentially surpass it in some domains. The eye and human vision possess remarkable versatility and power, necessitating a targeted approach in the development of electronic and information technology strategies aimed at emulating these capabilities. The objective is to gather the most comprehensive information about our physical world by focusing on certain aspects of the human visual system. Such two advanced fields of imaging science are discussed in the last part of this Chapter.

One of the key technologies of advanced imaging is High dynamic range imaging (HDRI). It is a set of techniques that computationally enhance the standard (conventional) dynamic range of digital images. The difference in the dynamic range of the human visual system (HVS) (Hoefflinger, 2007) compared to standard cameras/displays provides an inherent motivation to develop methods for capturing and displaying HDR images that better simulate the sensation of watching a real HDR scene. HDRI has emerged in a variety of domains, including computational photography, computer graphics, animation, etc. HDR content can be generated in a variety of ways, including merging several lower dynamic range images for HDR reconstruction, simulating or raytracing via computer graphics tools, or using HDR sensors for HDR data capture. In the following subsections, a brief summary of the different aspects of this technology is addressed. However, to develop a better understanding of HDR imaging and further details on HDR imaging pipeline please refer to Reinhard et al. (2010).

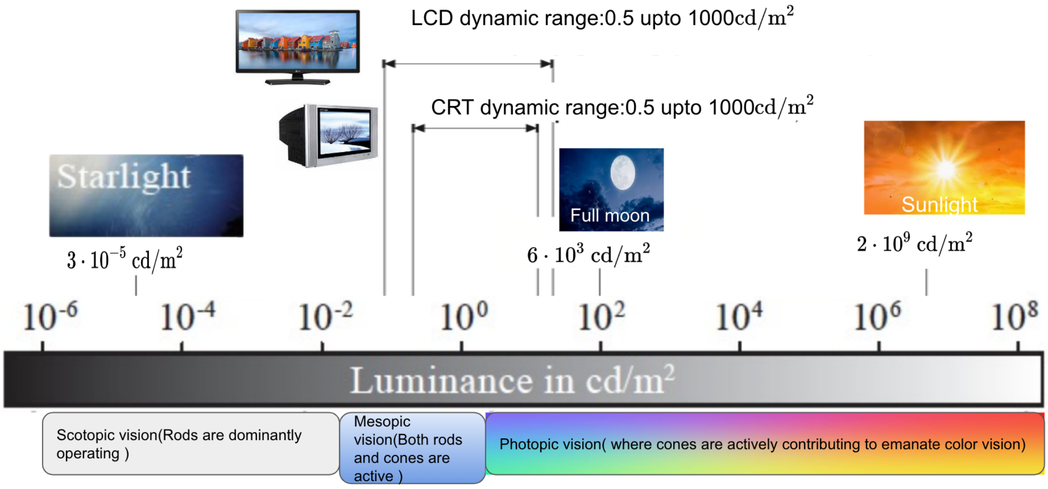


Figure 3.3: *The figure illustrates luminance ranges commonly encounter in real-world scenarios and commonly used display screens.*

3.3.1 Dynamic Range

The definition of the term "dynamic range" requires more clarification as it is a dimensionless quantity that can be used to refer to a variety of distinct physical measurements. Following are the widely encountered scenarios in which the concept of dynamic range differs.

- **Images:** The dynamic range of an image is the ratio between its brightest and darkest pixel. The minimum and maximum pixel values of any image are arguably outliers by definition, so this way of calculating dynamic range is not necessarily robust. However, excluding a certain percentage of the lightest and darkest pixels is a more reliable approach. Commonly researchers use the 99 percentile and 1 percentile respectively to estimate the dynamic range of an image.

- **Display:** The dynamic range of a display is the ratio between the maximum and minimum luminance levels that it can emanate.

- **Camera:** The ratio between the luminance that saturates the sensor and the luminance that shifts the camera response one standard deviation above the noise floor.

The HVS can adapt to illumination conditions that vary by nearly ten orders of magnitude, while simultaneously operating over a range of approximately five orders of magnitude within the scene. Figure 3.3 depicts the luminance ranges commonly encountered in real-world scenarios and widely used display screens

3.3.2 High vs low dynamic range imaging

High dynamic range (HDR) images and videos contain pixels that can represent a significantly greater range of colors and brightness levels than those offered by current dynamic range standards. HDR does not only offer brighter and more colorful images than their SDR reproduction but also contains much higher local and global contrast between neighboring and distant objects respectively (Pouli et al., 2016; Narwaria et al., 2016).

The electro-optical transfer function (EOTF) of a standard dynamic range (SDR) display typically follows a 2.2 gamma curve and employs 8 bits of color depth. As a result, the maximum luminance of the display is approximately 500 nits. High dynamic range (HDR) displays, on the other hand, utilize an extended EOTF curve, such as the Perceptual Quantizer (PQ) curve, and a minimum color depth of 10 bits. This enables HDR displays to achieve significantly higher luminance levels, exceeding 1000 nits and even exceeding 4000 nits (Seetzen et al., 2004).

Color depth refers to the level of precision that can be used to encode color information. The additive displays use three primary colors (red, green, and blue) to synthesize all display gamut colors. In the case of standard dynamic range (SDR) display, each color channel can be encoded with 8 bits, allowing for 256 levels of color shades. By utilizing this 8-bit encoding range, standard SDR displays can show up to 16,777,216 colors. Although this appears to be a vast number, it should be noted that conventional images mostly depict approximately two orders of magnitude of dynamic range, which is insufficient for many real-world scenes. In contrast, HDR standards improve color accuracy by increasing the color depth per channel to 10 bits. This yields a total of 1,073,741,824 range of colors for HDR displays. Hence, HDR color space is considered to be the superset of all the SDR traditional color spaces. This expanded color gamut enables HDR displays to generate more realistic and vibrant images with finer color gradations, thereby contributing to a more immersive visual experience.

3.3.3 The HDR imaging pipeline

The HDR imaging pipeline can be divided into five main stages i.e., capturing HDR content, processing, rendering and finally visualizing it on the screen as shown in Figure 3.4.

The first stage involves capturing HDR content which may be captured in a variety of ways. The hardware-based methods of HDR capture encompass the utilization of specialized sensors or devices explicitly engineered to capture a wider dynamic range of light in comparison to conventional cameras. These techniques facilitate the simultaneous acquisition of different exposures or enhance the sensor's

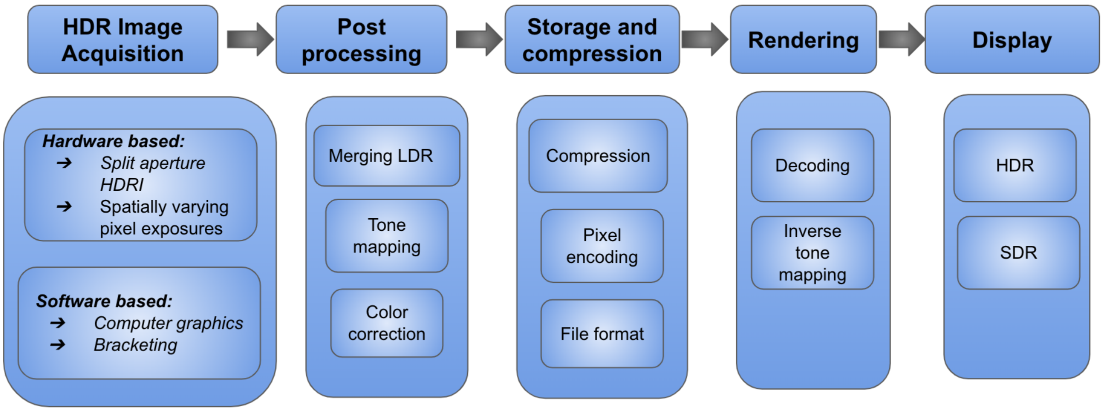


Figure 3.4: *The HDR imaging pipeline consisted of four main stages which involve HDR capture, post-processing of captured HDR content, efficient encoding and storage, and finally rendering on the output display screen.*

capacity to capture real-world HDR scenes. Examples of such two methods are: splitting the aperture and varying pixel exposure sensitivity. The software-based methods include merging LDR images taken at multiple exposures, using computer graphic tools and/or using computer vision techniques for HDR reconstruction from single-exposure images. In the next stage, the captured HDR data may be further posted in a number of ways such as merging the LDR captures for expansion of dynamic range, to account for the artifacts removal or color shift adjustments, etc. Besides, tone mapping can also be employed in this stage to adjust the dynamic range given the limitations of the output display device.

For storage and distribution, specialized HDR data encoding techniques and file formats are employed as the HDR content may exhibit a much greater file size compared to its LDR counterpart, owing to the explicit representation of HDR values. At the display stage, the HDR images are rendered by employing pixel data decoding considering the Electro-optic transfer function of the display device.

3.3.4 Methods of HDR Acquisition

HDR images can be created through various approaches, each of which employs different techniques to capture or reproduce HDR scenes. Some common methods include merging multiple LDR images to reconstruct an HDR image, utilizing learning-based models to enhance dynamic range, simulating HDR scenes via ray-tracing, or employing specially designed HDR sensors to capture real-world HDR scenes which are discussed as follows.

3.3.4.1 Hardware-based single shot HDR capture

A traditional digital camera normally gives a dynamic range of two orders of magnitude i.e., the ratio of intensity between the brightest and darkest pixels. However, many real-world scenes have a wider range of brightness. As a result, some parts of digital camera photos are undersaturated or oversaturated. One way to expand the dynamic range of digital cameras is to use different exposure setting to capture different regions of the HDR scene. To capture multiple exposures simultaneously the researchers have been found using a beam splitter in front (multiple cameras) or behind (single camera) the lens to split the light into multiple beams reaching multiple sensors which results in capturing multiple exposures in a single shot(Wang et al., 2005; Aggarwal and Ahuja, 2001). Another approach is to utilize a sensor with varying pixel sensitivity (Nayar and Mitsunaga, 2000; Nayar and Branzoi, 2003; Brajovic and Kanade, 1996). Although these hardware-based HDR capture techniques result in high quality HDR images with no temporal misalignment and no ghosting artifacts, however, the required HDR acquisition setup and post-processing are usually expensive.

3.3.4.2 Multiple exposure HDR capture

The most common technique of HDR imaging is to capture multiple shots of the scene at varying exposures in a sequential manner (Reinhard et al., 2010; Burt and Kolczynski, 1993; Tomaszewska and Mantiuk, 2007). The LDR images taken at low exposures capture relevant details in bright scene regions, whereas the LDR images taken at high exposures provide helpful data in dark scene regions. As a result, the captured photos can be merged to produce a single HDR image. This approach is taken a step further by Debevec and Malik (2008) where the author attempts to recover the radiometric response of the imaging system from these multiple exposures HDR images in order to estimate the radiance map of the HDR scene as depicted from Figure 3.5.

Ghosting artifacts caused by moving objects or misalignments is a key challenge in high dynamic range (HDR) imaging for dynamic scenes. The methods described above are, of course, only applicable to static scenes; the imaging apparatus, scene objects, and their radiances are all required to be constant during the sequential capture of LDR at different exposures. Therefore, the quality of file HDR image is highly dependent on LDR image alignment, merging, and other post-processing algorithms.

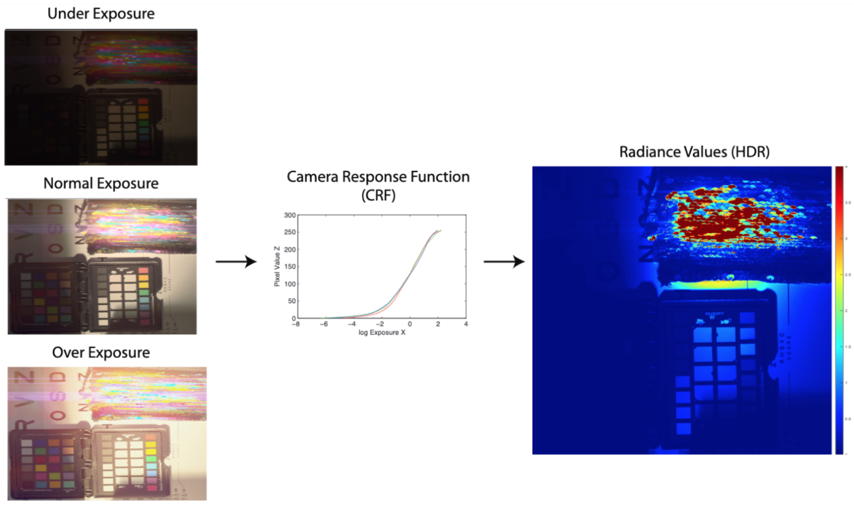


Figure 3.5: *(Reproduced from Figure 1.2 for convenience) The multiple exposure HDR capture utilizes a set of LDR images taken at different exposure settings such that all the details of the HDR scene are captured. Later, these LDR images are used to recover the camera response function which is employed for linearity correction. Once the camera response function is obtained, the LDR images are transformed into linear representations. These linear LDR images are then combined or merged together to create an estimation of the radiance map of the actual HDR scene.*

3.3.4.3 Learning-based HDR reconstruction

In recent years, there is a remarkable advance in HDR reconstruction using deep learning technologies. Learning-based HDR reconstruction is an advanced approach that employs deep learning algorithms to generate high dynamic range (HDR) images from a single or limited number of low dynamic range (LDR) images. Instead of relying on multiple exposures, this method utilizes a trained deep-learning model to predict the missing HDR details and enhance the image's dynamic range. These models can also handle challenging lighting conditions and scenes with moving objects better than traditional image processing-based methods. By training on a large dataset of paired LDR and HDR images, a deep learning model extract the features and use them to comprehend the relationship between the input LDR images and their corresponding HDR versions. This trained model can then be used to analyze any LDR image and estimate the missing HDR data. The result is an enhanced image with a wider range of luminance levels, simulating the visual characteristics of a true HDR image (Eilertsen et al., 2017a; Park et al., 2022; Pan et al., 2020).

Others also attempted to incorporate the domain knowledge of the LDR image formation pipeline to design the learning-based model to estimate the inverse camera response function for restoring the missing details caused by quantization and reducing the visual artifacts in the underexposed regions and over-exposed regions (Liu et al., 2020). These techniques have also been escalated to learn a mapping from a stack of LDR images to generate a ghost-free HDR image where attention-guided networks have also been employed for artifacts removal (Wu et al., 2018).

3.3.5 Post Processing of HDR content

In order to refine and enhance the final HDR reproduction considering the output device, post-processing plays a vital role in the HDR imaging workflow. Several post-processing techniques for optimizing the HDR image's appearance and accounting for the undesired atifacts are proposed which are discussed below.

3.3.5.1 Tone compression

HDR images are considered to be scene-referred which means that an HDR image stores the radiometric quantity i.e., radiance of the actual HDR scene which is further scaled by the sensor's sensitivity. HDR images, unlike output-referred JPEG files produced by traditional cameras, are not pre-processed for rendering on SDR display i.e., they are not gamma encoded pixel values. Since the HDR stores linear pixel values, HDR images are unsuitable for direct display on standard display. The majority of current display technologies are SDR displays which cannot display HDR images because of the substantially narrower dynamic ranges of the display. To visualize HDR images of high luminance on a normal SDR display, tone mapping is generally required to perform tone compression while preserving the maximum details that are present in HDR images. The details of tone mapping are discussed in Chapter 4.

It is also important to mention that not just SDR displays require tone compression but also for HDR display it is required. For instance, HDR images containing absolute luminance or radiance units in HDR images are sometimes higher than the maximum luminance of the available HDR display. Hence it is often needed to be scaled down the absolute luminance of HDR images by a constant factor before displaying on HDR display them as discussed in Chapter 5.

3.3.5.2 Color correction

One simple example of the need for color correction in HDR images is that the colorfulness of an object increases at high levels of illumination (Hunt, 1952). To accurately represent a scene under low levels of illumination, it is necessary to artificially enhance the object's colorfulness in order to get a comparable visual appearance. To ensure that the colors in the final HDR image are accurate and aesthetically pleasing, color correction is an important step in the HDR imaging pipeline. Due to the fact that HDR images are created by combining multiple exposures with varying levels of brightness, color variations can occur due to differences in illumination conditions and camera settings. The objective of color correction techniques is to address these issues and bring uniformity to the image's colors. Additionally, HDR images are often processed in a wide-gamut HDR color space i.e., REC2020 to preserve color accuracy and detail. Color space conversion may be necessary when exporting or displaying the HDR image on standard monitors which usually operate in sRGB color space.

3.3.6 Methods of HDR storage and compression

The fundamental difference between HDR and traditional SDR image format and encoding techniques lies in the fact that HDR consistently utilizes device-independent and high-precision data. This ensures that the dynamic range compression and other adjustments are carried out just during the display phase, and only if a device is incapable of accurately reproducing the information. Following are the commonly used pixel encoding and file formats for recommended HDR content

3.3.6.1 HDR pixel encoding

In general, pixel values in HDR images are linearly related to luminance, the photometric quantity that defines the perceived intensity of light per unit surface area regardless of color. The cameras used to capture HDR images have a different spectral sensitivity from the luminous efficiency function of the human eye (which is used in the definition of luminance). However, HDR pixel values are considered a close approximation of these photometric quantities. When three color channels are taken into account, each color component in an HDR image is sometimes referred to as radiance. The physical definition of radiance implies that the light is integrated over all wavelengths, whereas the spectral characteristics of red, green, and blue HDR pixel values are constrained by the spectral sensitivities of a camera system. Hence the HDR pixel values are assumed as the relative radiance of the HDR scene. However, the most accurate term describing the quantities that are stored in HDR pixels is trichromatic color values. Following are the commonly used HDR pixel encoding approaches that aim to reduce the number of necessary bits while maintaining adequate accuracy and the capacity to encode a wide dynamic range.

- **Half-precision float(fp16) encoding:** In half-precision float i.e., fp16 encoding consists of one bit of sign, 5-bit exponent, and 10-bit mantissa.

- **RGBE encoding:** RGBE is a 32-bit floating point encoding of pixel values that employs 8 bits for exponent and additional 8 bits for mantissa (8E8). RGBE encoding takes advantage of the fact that three is a high correlation among all color channels in the RGB color spaces and have values of at least the same order of magnitude. Consequently, there is no need to record a different exponent for each color channel.

- **Logarithmic encoding:** The logarithmic encoding decorrelates the luminances and chrominance by encoding them separately such as the 32-bit LogLuv encoding utilizes two bytes for luminance and other two bytes for encoding the chrominance. In this way, pixel encoding can be roughly correlated to the perception of HVS.

- **Perceptually uniform encoding:** The pixel values in LDR (Low Dynamic Range) exhibit a desired characteristic wherein their values demonstrate an approximately linear correlation with the perceived brightness of the corresponding pixels. As a result, the utilization of LDR pixel values is quite appropriate for image encoding as the artifacts caused by image compression have the same visual impact across the whole range of image values. In contrast, HDR pixel values do not possess this characteristic. Consequently, when a similar level of distortion is applied to regions of low luminance and high luminance in an image, the artifacts become more discernible in the low-luminance regions. Hence, a number of encodings have been proposed which utilize more precise models of the HVS sensitivity to variations in high luminance such as Mantiuk et al. (2005); Miller et al. (2013).

3.3.6.2 HDR file format

Different file formats have been specifically designed to store HDR images. In the following sections, the two most widely adopted HDR image formats i.e., Radiance HDR and OpenEXR are discussed

- **Radiance HDR:** The Radiance HDR images typically use 32 bits encoding range per channel to provide enough precision to represent a wide color gamut and dynamic range. It is one of the early HDR file formats represented by *.hdr* file extension which contains a short header text followed by run-length encoded RGBE pixels values. The MATLAB *.hdr* is also a radiance HDR file using RGBE encoding which has been utilized in our work.

- **OpenEXR:** OpenEXR is a well-known HDR image format that was developed by Industrial Light and Magic (ILM) and made available in 2003. The format

Table 3.1: *Overview of HDR Standards*

HDR Standard	Color Space	Metadata	Bit Depth	Max. Luminance	Access
HDR10	Rec. 2100	Static	10-bit	1000	Open Source
HDR10+	Rec. 2020	Dynamic	16-bit	10000	Royalty-Free
Dolby Vision	Rec. 2020	Dynamic	12-bit	10000	Proprietary
HLG	Rec. 2020	Static	10-bit	1000	Royalty-Free

was created to satisfy the needs of the visual effects industry and has since been widely adopted by the animation, film, and gaming industries.

3.3.7 HDR standards

It is also worth mentioning that 4K and UHD (which stands for ultra-high definition) refer to the screen's resolution, which impacts the image's sharpness. In contrast, HDR increases the contrast ratio, which is the luminance difference between an image's brightest and deepest regions as discussed above. With its greater contrast ratio, HDR demonstrates more details in extremely bright and dark regions of the high dynamic range scenes, delivering HDR reproductions that appear more realistic and stand out more noticeably.

To support HDR imaging in a variety of industrial applications, a number of HDR standards have been created among which the widely used are briefly discussed in the following subsections while Table 3.1 provides a summary of these HDR standards.

3.3.7.1 HDR10

HDR10 is an open standard and one of the most popular HDR formats for consumer electronics. It employs the ITU-R BT.2020 color space and a static metadata approach, offering 10-bit color depth and a maximum luminance of 1,000 nits. HDR10 format supports the compressed transmission of HDR video content. For display of all levels, they are all required to support the industry standard HDR-10 format in order to properly display HDR content.

3.3.7.2 HDR10+

HDR10+ is an open, royalty-free HDR format developed by Samsung and supported by various content creators and manufacturers. The difference between HDR10 and HDR10+ is that the former employs static metadata, which applies the same parameters (such as maximum luminance and color settings) to the entire video, thereby limiting the scene-specific optimization. HDR10+ dynamically adjusts HDR data scene-by-scene or even frame-by-frame, allowing for more precise and granular control over HDR presentation.

3.3.7.3 Dolby Vision

Dolby Laboratories developed the Dolby Vision proprietary HDR format. Utilizing dynamic metadata similar to HDR10+, it enables content creators to specify scene-by-scene adjustments for contrast and color. Dolby Vision supports a higher color depth, a broader color gamut, and also a higher peak luminance range than HDR10. However, HDR10+ competes directly with Dolby Vision. The adoption of HDR10+ has been steadily increasing, and it provides an attractive option for content creators and consumers who seek a high-quality HDR experience without the need for proprietary technology or licensing fees.

3.3.7.4 Hybrid Log-Gamma(HLG)

HLG (Hybrid Log-Gamma) is an HDR standard developed together by the BBC and the Japan Broadcasting Corporation (NHK). It was established to overcome the issues of distributing HDR content to a wide range of devices, including HDR and SDR (Standard Dynamic Range) displays, without the need for tone mapping based on metadata. One of the key benefits of HLG is its backward compatibility with SDR screens. When an HLG signal is displayed on an SDR display, the content will seem as if it were a conventional broadcast, with no color or brightness aberrations. This is due to the fact that HLG employs a hybrid gamma curve, which includes both standard gamma information for SDR and HDR enhancement information.

While HLG has the advantage of backward compatibility and simplicity, it may lack the precision and scene-by-scene adjustments that dynamic metadata-based formats such as Dolby Vision or HDR10+ offer. HLG, on the other hand, remains a valuable HDR standard, particularly for applications where compatibility with SDR displays is essential.

3.4 Advanced imaging techniques: Hyperspectral imaging

Hyperspectral imaging gathers hundreds of images at various wavelengths for the same spatial. While the human eye is only limited to trichromatic vision, hyper-

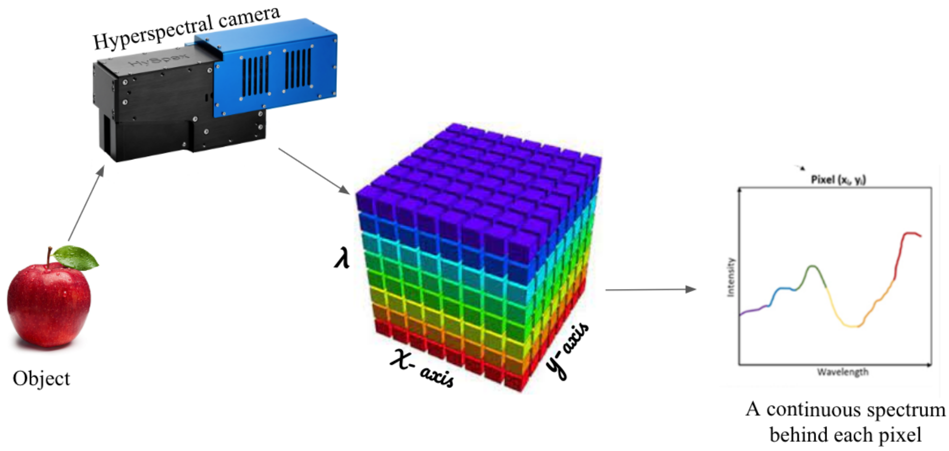


Figure 3.6: *The hyperspectral imaging measures a continuous spectrum (in radiance or reflectance) and can be used to precisely characterize the objects within the scene with great detail*

spectral imaging measures the continuous spectrum of light for each pixel of a scene with fine wavelength resolution in both the visible and near-infrared. The first two dimensions of the hyperspectral cube represent the spatial extent of the scene(i.e., rows and columns), whereas the third represents its spectral composition as shown in Figure 3.6. The capability of digital imaging and spectroscopy is combined in hyperspectral imaging, also known as imaging spectroscopy. A hyperspectral camera measures the light intensity (radiance) for a large number of contiguous spectral bands for each pixel in an image. Every pixel in the image thus has a continuous spectrum (in radiance or reflectance) and can be used to precisely characterize the objects within the scene with great detail.

Clearly, hyperspectral images provide significantly more information about the scene than images captured by a normal RGB camera, which only acquires three spectral channels corresponding to the visual primary colors channel i.e., red, green, and blue. Each material has a unique spectral signature that can be used as its unique identification. Therefore, hyperspectral imaging greatly improves the capability to classify scene objects based on their spectral properties. This technology has numerous applications in various fields, such as agriculture, environmental monitoring, forensics, medical imaging, remote sensing, and food quality control. As a result of recent developments in sensor technology and computing power, hyperspectral imaging has evolved from relatively slow and unreliable research prototypes to reliable, accurate, and precise analytical instruments.

3.4.1 Types of hyperspectral acquisition modes:

Hyperspectral imaging can be acquired using different modes, each with its own advantages and disadvantages.The four primary hyperspectral acquisition modes are:

- **Point scanning (whiskbroom mode):** In this mode, the hyperspectral camera obtains a hyperspectral image by scanning a single spatial point across the scene. Typically, a mirror that oscillates back and forth is used to sweep the scene during the scanning process. This mode is also referred to as whiskbroom mode because it resembles sweeping the scene with a broom. The main advantage is its high spectral and spatial resolution, but the data acquisition rate is limited.

- **Line scanning (pushbroom mode):** In this mode, the hyperspectral camera simultaneously captures a scene line by line in all spectral bands as it travels over the target area. Compared to point scanning mode, this mode provides greater spectral and spatial resolution and quicker data acquisition rates.

- **Plane scanning (area scanning mode):** In this mode, the hyperspectral camera simultaneously captures a two-dimensional scene region across the spectral range. This mode offers a rapid data acquisition rate, but its spatial resolution is inferior to that of the point scanning and line scanning modes.

- **Single shot mode:** In this mode, the hyperspectral camera simultaneously captures an entire scene in a single snap. This mode is helpful for applications requiring instant information capture, such as medical imaging or surveillance. Nevertheless, it typically has inferior spatial and spectral resolution than other modes.

Overall, the choice of hyperspectral imaging mode depends on the specific application requirements and tradeoffs between spatial and spectral resolution, data acquisition rate, and equipment cost.

4 High Dynamic Hyperspectral Data Acquisition and Processing

If we knew what it was we were doing, it would not be called research, would it?

Albert Einstein

The main focus of this chapter is to discuss the process of adequately performing hyperspectral data acquisition and the post-processing of the raw hyperspectral cubes to estimate the radiances in the HDR hyperspectral cube (hyperspectral radiance map) accurate to the actual HDR scene. Due to uncontrolled lighting conditions and limitations of the imaging devices such as a low dynamic range, it becomes difficult to recover the details in hyperspectral images with bright regions and dark shadows at the same time. To overcome this problem, HDR techniques can be used which allow us to significantly recover the radiance of the HDR scene.

The widely used technique of HDR image capture from multiple exposures as discussed in Chapter 3, has been employed for capturing the HDR hyperspectral cube. The main idea was to capture hyperspectral cubes at different integration settings (i.e., exposures) such that all the details of the HDR scene are properly captured avoiding under or overexposed pixel values. During the post-processing stage, the properly exposed region in each of the hyperspectral cubes can be merged together to get the full radiance map(i.e., hyperspectral mean radiance cube) of the HDR scene.

As imaging sensors are linear in the lower and middle regions of their linearity, the sensor's response becomes non-linear in the upper region which results in incorrect radiance values. Hence sensor linearity correction has also been implemented to recover the digital counts beyond the linearity threshold of the sensor. Furthermore, the hyperspectral interpolation was performed to verify and account for the missing wavelengths in the calculated mean radiance cube (i.e., radiance map) where the sensor was saturated across all the integration settings. Finally, the mean radiance cube has been converted to a linear three-channel RBG image and stored in 32-bit hdr format.

In this chapter, the hyperspectral camera is discussed to understand its working and output quantities followed by the characterization. The characterization involved three primary tasks: assessing the camera's linearity, estimating the spectral sensitivity curve, and validating the radiance outputs produced by the camera. Later, we discussed the methodology for data acquisition and proposed the workflow for storing and displaying it. For displaying HDR images on an HDR display, we have proposed a display-referred HDR rendering workflow which has also been compared with the ITU-recommended pipeline.

4.1 HySpexVNIR-1800

For calculating the HDR hyperspectral radiance image, HySpexVNIR1800 hyperspectral camera was used. The range of this hyperspectral camera is from 400 nm to 1000 nm and contains 186 bands with a step size of 3.26 nm. The HySpexVNIR1800 is a line scanner where the scanning speed is synchronized automatically with the integration time. The integration time can manually be adjusted through its software namely HySpex GROUND. In this experiment, a 30 cm cylindrical lens with a field of view of approximately 86 mm was used and it captures 1800 spatial pixels across a line. Because of limited bandwidth and dynamic range, the cameras usually perform compression and convert linear sensor response into gamma-corrected pixel values. However, HySpex VNIR-1800 stores the captured data as RAW digital number images which can be converted to calibrated digital numbers and their corresponding absolute radiances using HySpex RAD software(see section 4.1.2 of details of supporting packages).

For our application, the HySpexVNIR1800 has been used in ground-based operations with a transitional stage to capture the hyperspectral cubes at different exposures in a laboratory setting, as depicted on the left side of Figure 4.1. Other relevant specifications of this hyperspectral camera are summarized in Table 4.1.

4.1.1 Working of HySpex-VNIR1800

The camera's internal operation is depicted schematically on the right side in Figure 4.1. The focusing mirror projects the scene onto a slit through which only

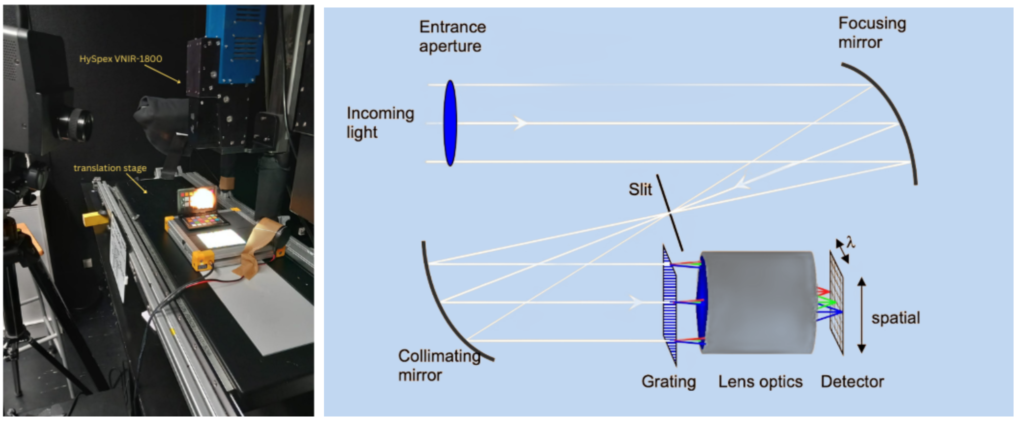


Figure 4.1: *The HySpex VNIR-1800 has been integrated with a translation stage that scans the camera field of view across the scene as shown on the left side while the internal working diagram(taken from manual) of HySpex cameras is given on the right side.*

light from a thin line in the scene passes where each line captures 1800 pixels. After passing from the collimating lens, different wavelengths are separated by a transmission grating which has been used as a dispersive element, and the light is then focused onto a detector array of photodiodes. For each pixel interval along the line specified by the slit, a corresponding spectrum is projected onto a column of detectors on the array as a result of the optics. Therefore, the data read from the array comprises a segment of a hyperspectral image, with spectral information in one direction and spatial (image) information in the other as illustrated in figure 4.2.

4.1.2 HySpex Softwares Overview

Four different software packages are provided with the HySpexVNIR1800 camera which are as follows: HySpex GROUND is the main software package for ground base operations, HySpex AIR is developed for airborne operations, HySpex RAD is intended for the radiometric calibration, and HySpex NAV for resampling the navigation data to the HySpex image data. Only HySpex GROUND and HySpex RAD have been used in this project which are discussed below.

- **HySpex GROUND** is the primary software package for ground-based operations in the laboratory and field, including rotation and translation stage scanning. The hyperspectral images captured by this software are stored in the form of raw digital numbers from the sensor with the "*.hyspex*"

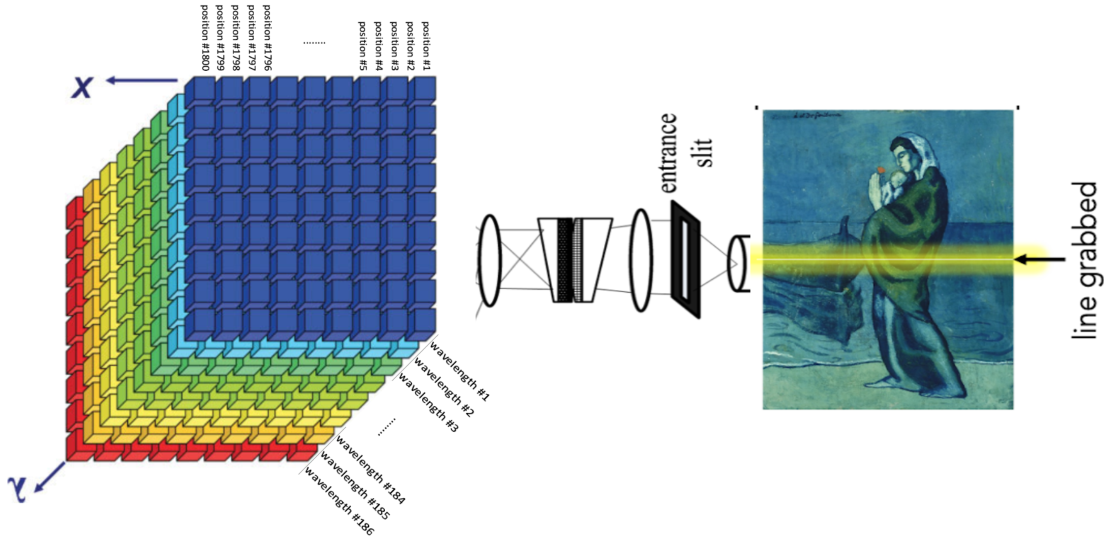
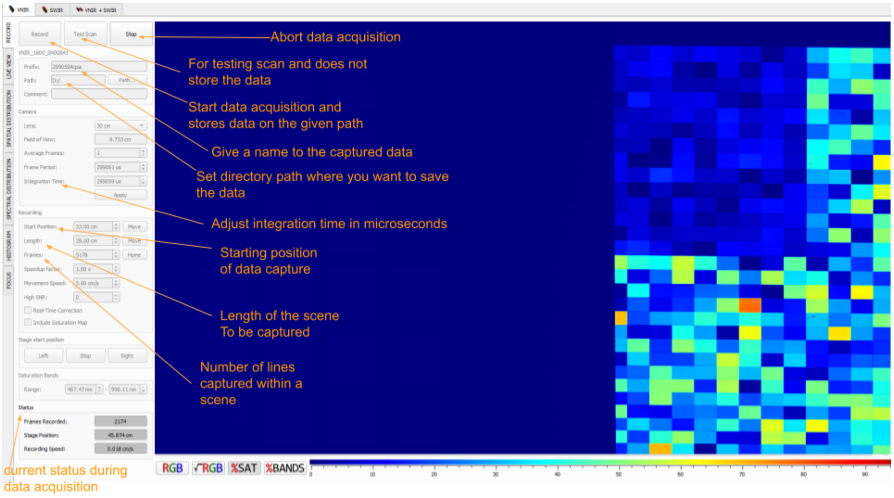


Figure 4.2: *The HySpexVNIR1800 functions as a line scanner as shown in this diagram where it captures 1800 pixels per line and behind each pixel the spectrum spans from 400 nm to 1000 nm containing 186 bands with a step size of 3.26 nm.*

Table 4.1: *The main specifications of HySpex VNIR-1800 camera are given in the table below.*

Specification	Value
Spectral range	400 – 1000 nm
Spatial pixels	1800
Spectral channels	186
Spectral sampling	3.26 nm
FOV*	17°
Pixel FOV across/along*	0.16/0.32 mrad
Bit resolution	16 bit
Noise floor	2.4 e-
Dynamic range	20000
Peak SNR (at full resolution)	>255
Max speed (at full resolution)	260 fps
Power consumption	30 W
Dimensions (l–w–h)	39 – 9.9 – 15 cm
Weight	5.0 kg
Camera Interface	CameraLink

file format. The HySpex GROUND software interface along with the preview



Abort data acquisition

For testing scan and does not store the data

Start data acquisition and stores data on the given path

Give a name to the captured data

Set directory path where you want to save the data

Adjust integration time in microseconds

Starting position of data capture

Length of the scene To be captured

Number of lines captured within a scene

current status during data acquisition

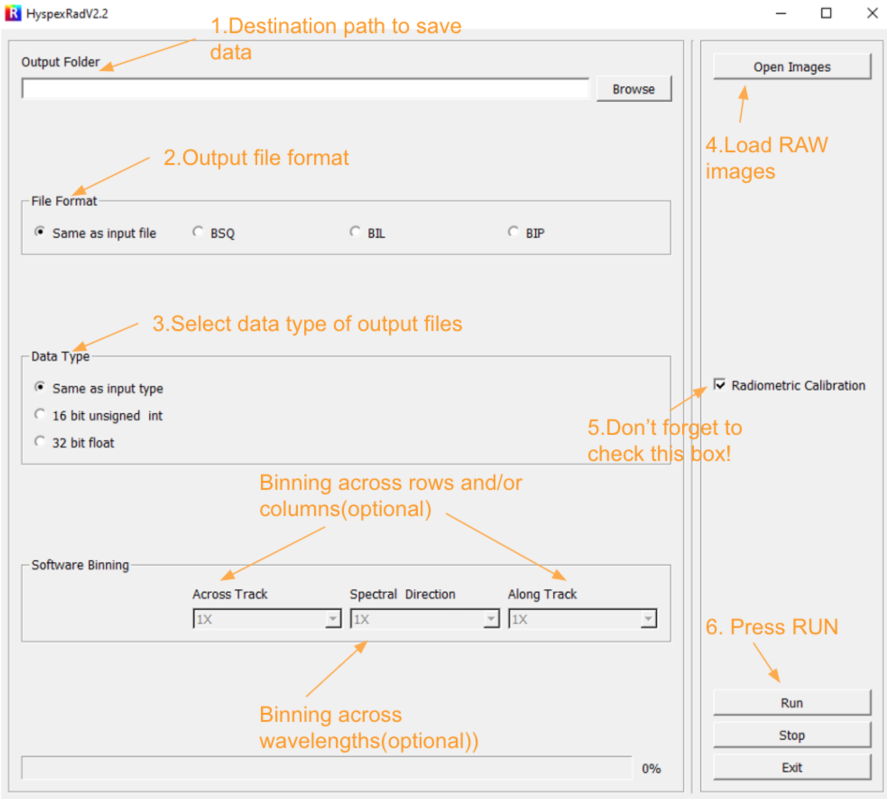
Figure 4.3: *The interface of the HySpexVNIR1800 Ground is illustrated in this figure with the description of how to operate it for data acquisition when the HySpexVNIR1800 camera is set up in Ground mode and integrated with a translation stage.*

window is shown in Figure 4.3. The preview has four different options i.e., $sq\_rootRGB$, RGB, SAT, and $BANDS$ view mode. We have mostly used SAT which shows the saturation map of the capture. If the pixels are closer to red in color, it means that the pixels are about to saturate while it is better to operate in the greenish region to get a better signal-to-noise ratio. Additionally, multiple captures can be taken and averaged to boost the signal, especially in the dark regions of the scene.

- **HySpex RAD** is the software package that can be used for converting RAW images from HySpex GROUND to absolute radiance. The real-time image correction feature handles image data non-uniformity and dark offset correction. If absolute radiance measurements are not required, all subsequent processing can be performed on these real-time corrected data that are stored directly on the given path with the "*raw.hyspex*" file format.

4.1.3 Output Qualities of HySpex VNIR-1800: RAW image vs Radiance image

Image sensors emulate artificially the transduction process of converting one form of signal into another as performed by human vision. As discussed in Chapter 2, rod and cone photoreceptors in the retina work together with ganglion cells to



1.Destination path to save data

2.Output file format

3.Select data type of output files

4.Load RAW images

5.Don't forget to check this box!

6. Press RUN

Binning across rows and/or columns(optional)

Binning across wavelengths(optional))

File Format: ● Same as input file ○ BSQ ○ BIL ○ BIP

Data Type: ● Same as input type ○ 16 bit unsigned int ○ 32 bit float

☑ Radiometric Calibration

Software Binning: Across Track 1X; Spectral Direction 1X; Along Track 1X

Figure 4.4: *The interface of HySpex RAD software is shown in this figure with the labeling of available options for converting RAW data files into corresponding absolute radiances which are stored in 32 bits float format in our study.*

convert photons into an electrochemical signal that the occipital lobe in our brain can then process (see Section 2.4). In the case of imaging sensors, the photoelectric effect is used by both CCD and CMOS sensors to collect a charge in their silicon pixels when light hits the sensor. The amount of charge in each pixel is measured by a capacitor and generated a voltage proportional to the charge which is then Later transferred into digital numbers. These digital numbers can be processed by a computer for image formation as discussed in Section 3.2 of Chapter 3. Likewise, HySpexVNIR1800 also works in a similar manner as depicted in Figure 4.5. RAW images are snapshots of the values registered by the camera's imaging sensor. The raw data output from the camera is digital number (DN) values ranging from 1 to $2N$ minus 1, where N is the digitization bit width. For each readout of the camera, corresponding to one row of pixels in the image, we obtain a digital number matrix $DN[i,j]$, where i is the band number and j is the spatial pixel number as depicted in equation 4.1. Here, $Ni[i,j]$ represents the count of photons received by the jth spatial pixel in the ith band over a period of time t. $QE[i]$ denotes the quantum

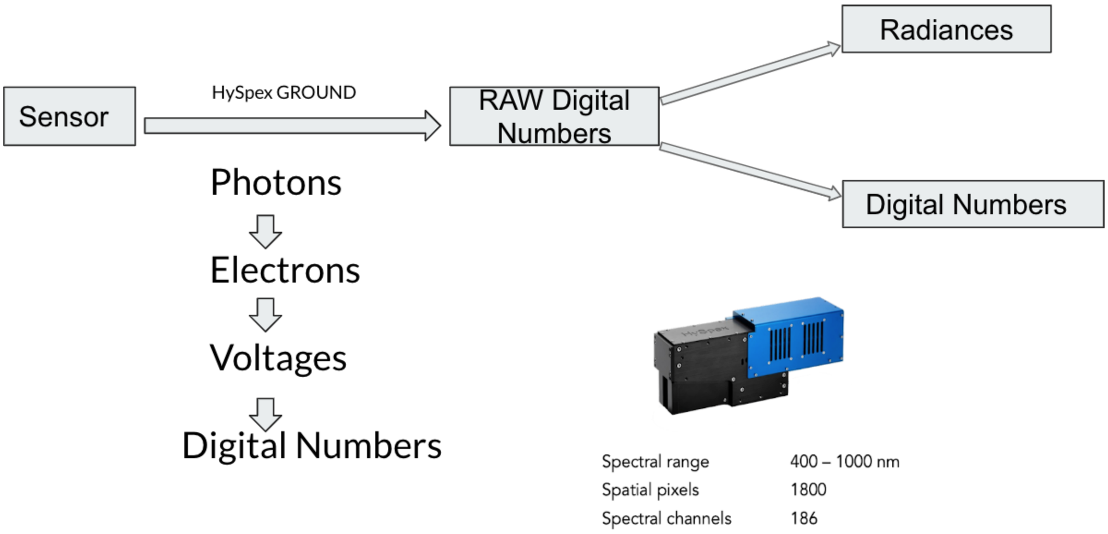


Figure 4.5: *HySpexVNIR1800 converts the photons hitting it's sensor into electrons which are then converted into voltages. At certain stages, these voltages are transformed into RAW digital numbers. Using the software packages as discussed in previous sections, the RAW data can be converted to calibrated digital numbers and corresponding absolute radiance values.*

efficiency of the camera (including optics and sensor) for the *ith* band. Quantum efficiency is generally measured as the ratio of photoelectrons produced per photon. SF is a scaling factor that defines the Digital Number (DN) per photoelectron. RE is a relative responsivity matrix for each detector element, which displays an average value of 1 and indicates the inherent non-uniformity in the sensor's response. Finally, BG denotes the background matrix, which accounts for dark current and offsets.

$$DN[i, j] = Ni[i, j] \cdot QE[i] \cdot SF \cdot RE[i, j] + BG[i, j] \tag{4.1}$$

Because of limited bandwidth and dynamic range, the cameras usually perform compression and convert linear sensor response into gamma-corrected pixel values by applying tone mapping. However, HySpex VNIR-1800 stores the captured data as RAW digital number images which can be converted to calibrated digital numbers and absolute radiances using HySpex RAD software.

To fit the calibrated image into a 16-bit unsigned int format, a scaling factor is used. This scaling factor is included in the header file (.hdr) of each image. The absolute radiance is calculated by dividing the calibrated image by this scaling factor. Converting an image to 16-bit format can occasionally result in information loss. Alternatively, the image can also be converted to a 32-bit format to avoid

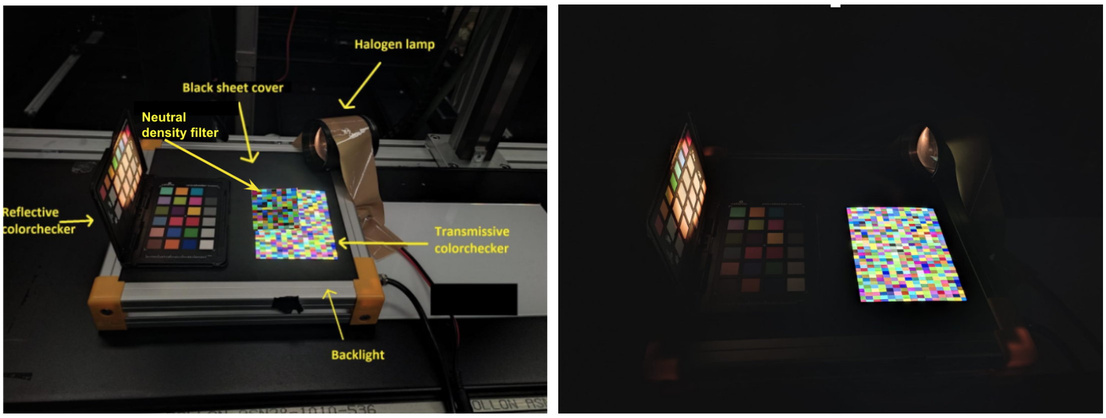


Figure 4.6: *The figure illustrates the data acquisition setup for capturing HDR hyperspectral radiance data. However, it is essential to clarify that the shown images are solely for demonstration purposes. The room was darkened while capturing the data.*

needing a scaling factor to convert it to radiance. However, the disadvantage of this format is that it takes up twice as much storage space on the computer.

The transition from RAW data to absolute radiance data is performed using the parameters that are acquisition-dependent, and variables that are permanently stored in the software as shown in the equation 4.2

$$L(i,j) = \frac{\frac{DN(i,j)-BG(i,j)}{RE(i,j)} \cdot h \cdot c}{QE(i) \cdot SF \cdot t \cdot A \cdot \Omega \cdot \lambda(i) \cdot \Delta\lambda(i)} \tag{4.2}$$

where Digital Number (DN), background matrix (BG), and integration time t are the only parameters that depend on data capture settings while all other variables are permanently stored in the camera's software (details of these parameters are available in the user manual of HySpex VNIR-1800).

4.2 Data acquisition setup

For calculating the HDR hyperspectral image, HySpexVNIR1800 hyperspectral camera was used as discussed above. It is important to recall that the range of this hyperspectral camera is from 400nm to 1000nm and contains 186 bands with a step size of 3.26nm. The HySpexVNIR1800 is a line scanner where the scanning speed is synchronized automatically with the integration time, which can manually be adjusted through its software named HySpex GROUND. In this experiment, a 30 cm cylindrical lens with a field of view of approximately 86mm was used and it captures 1800 spatial pixels across a line. Because of limited bandwidth and dynamic range, the cameras usually perform compression and convert linear sensor response into gamma-corrected pixel values. However, HySpexVNIR-1800 stores the captured data as RAW digital number images which can be converted to calibrated digital numbers and absolute radiances using HySpex RAD software as discussed in the previous sections.

4.2.1 Setting up an HDR scene

For setting up an HDR scene, the backlight as shown in Figure 4.6 has been utilized where we placed a film/transmissive color target and a reflective passport color checker. The rest of the surface of the backlight was covered with a black sheet. We also included some mid-tones in this HDR scene, and for that, we utilized a neutral density filter that was placed on a transmissive target. Another diffused halogen light source was used because of the limitation of the hyperspectral camera to capture the details of the passport colorchecker even at its maximum integration setting. It is important to note that the actual data captures were taken in a dark room, however, the image shown in Figure 4.6 has been taken with dim room lighting for the purpose of demonstration. This section has addressed only one HDR scene while all other HDR scenes in the proposed dataset were also captured using a similar approach.

4.3 Camera characterization

Before capturing the data, the hyperspectral camera was characterized in order to develop a better understanding of its characteristics. The characterization of the hyperspectral camera involved three tasks i.e., 1) estimating the linearity curve of the camera, 2) calculating the spectral sensitivity curve, and 3) finally verifying the measured radiances and comparing them with the values measured with the CS2000A telespectroradiometer. Each of these tasks is discussed in the following subsections:

4.3.1 Linearity of imaging sensor:

A cutting-edge scientific CMOS imaging sensor is used by the HySpexVNIR1800. Because of this, the VNIR-1800 is claimed to be the perfect camera for sophisticated data collection that demands high radiometric accuracy. One important characteristic of the camera's imaging sensor is that the response of the sensor should be linear to the amount of light imping on it. Therefore, 31 captures were taken from the hyperspectral camera using a standard white tile under uniform

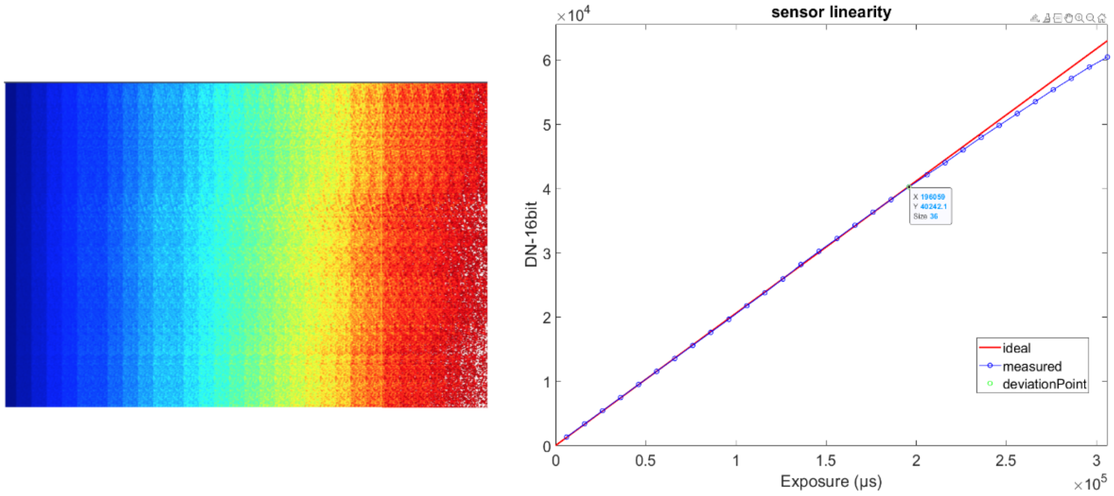


Figure 4.7: *The figure shows the heatmap of 31 measurements of white tile under uniform illumination taken from the HySpexVNIR1800 on the left side. The heatmap indicates that if the pixels are closer to red color, it means that the pixels are about to saturate. The estimated linearity of the HySpexVNIR1800 camera (i.e., digital count vs. exposure time) was found to be linear up to a digital count of nearly 40242 as depicted from the graph on the right side.*

illumination from 6 milliseconds up to 306 milliseconds integration settings with a step-size of 10 milliseconds over the full spectral range (from 400nm to 1000nm) of this camera for characterization. We then selected the spectral band at around 653.3nm as it was highly sensitivity. It has been found that the sensor's response was linear i.e., when the integration time was increased, the digital count also increased proportionally. However, this linearity was limited to nearly 40,000 digital counts as shown in Figure 4.7 where the measured linearity curve deviates from the ideal linear response. The ideal linear response is represented as a red curve which is plotted using linear regression on the data points which were within the linearity range and the green circle shows the linearity deviation point of the sensor from ideal linearity.

Although the graph is represented for digital numbers, it is important to note that the radiance values corresponding to digital numbers within the linearity of the sensor are valid and independent of the exposure setting as shown in Figure 4.8. In this figure, the digital numbers are given on the left side of the graph along with the corresponding radiances on its right side. It is worth noting this graph is calculated from only those captures of white tile which were within the linearity of the HySpexVNIR1800. As long as the digital counts remain within the linearity limit, it is evident from these graphs that the radiances are absolute and independent of the exposure whose unit is cd/m^2, unlike conventional cameras

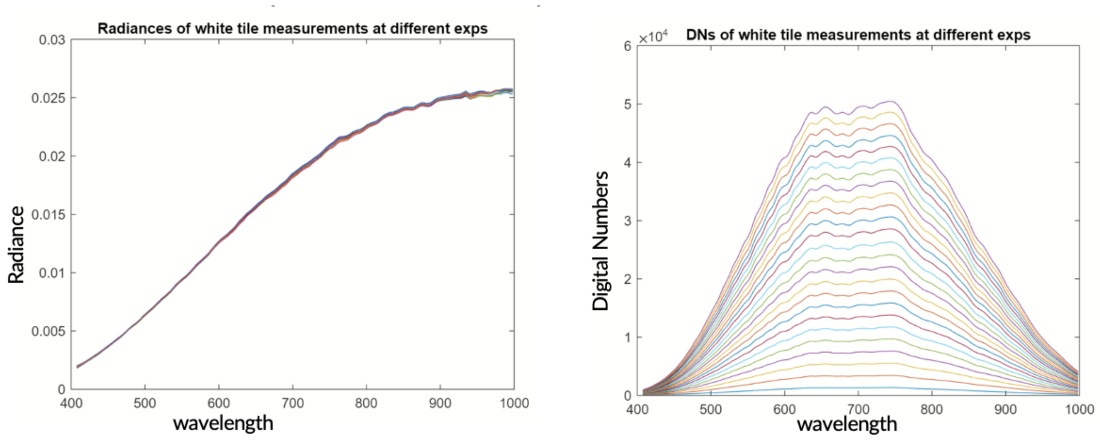


Figure 4.8: *The figure illustrates the digital numbers within the linearity of the imaging sensor of HySpexVNIR1800 which demonstrates that the radiances are absolute and are independent of the exposure as long as the digital count(digital number) is within the linearity limit, unlike conventional cameras where the output quality is gamma encoded or constrained by the sensor's nonlinearity.*

where the output quality is constrained by the sensor's non-linearity.

4.3.1.1 Proposed linearity correction method:

Enhancing the linearity of a sensor involves characterizing the camera's sensor response across its entire range and any non-linear regions can be effectively compensated for by applying the proposed linearity correction using measured data. It can be observed in the graph shown on the right side in 4.7 that the sensor behaves linearly up to a certain threshold (i.e., 40,242 digital counts). Then the measured linearity curve deviates from the ideal linear response which is a red curve plotted using linear regression in the given graph where the green circle shows the point of deviation of the sensor's linearity from ideal linearity.

The aim is to estimate the hypothetical radiance values that would have been obtained if the sensor had exhibited a linear behavior. In order to achieve this, the proposed linearity correction method has been utilized which is based on the lookup table where the objective is to quantify the deviation between the "ideal" response and the measured response. This is achieved by calculating the ratio of the "ideal" response to the measured response. Subsequently, these ratios are stored in the lookup table as the scaling factors which are then applied to the radiance values corresponding to the digital numbers within the non-linear region since digital numbers and radiances are directly proportional to each other. The linearity compensation was applied up to approximately 63,000 digital counts which has been found empirically given the HDR scene.

4.3.2 HySpexVNIR1800 spectral sensitivity verification:

The spectral sensitivity response is the change in the output signal as a function of the wavelength of the input signal. For estimating the spectral sensitivity of the imaging device, the experimental setup involved the utilization of the Bentham MT300 monochromator to direct monochrome light into the integrating sphere. The monochrome wavelength ranged from 440nm to 780nm, with a step size of 20nm while the HyspexVNIR1800 was employed on another end (i.e., opening of integrating sphere) to capture these 21 different monochrome lights impinged into the integrating sphere by the monochromator as depicted in the Figure 4.9. The CS2000A telespectroradiometer was used to measure the spectral power distribution (SPD) of each of these wavelengths and these measurements were treated as the ground truth. It is also important to note the spectral range of this hyperspectral camera is from 400nm upto 1000nam with a stepsize of 3.26nm while CS2000 takes measurements with wavelengths ranging from 380nm up to 780nm with a stepsize of 5nm. Therefore, the measurements taken by the camera were interpolated to the wavelengths at which CS2000A takes reading while both of these measurements were normalized between 0 and 1.

We calculated the spectral sensitivity of the camera, S_{meas} as given by equation 4.3:

$$S_{meas}(\lambda') = max\left(\frac{DN_{\lambda'}(\lambda)}{max(E_{\lambda'}(\lambda))}\right) \tag{4.3}$$

where $DN_{\lambda'}(\lambda)$ is the digital numbers generated from the hyperspectral camera when it captured the monochromatic light of λ' from the integrating sphere. Note that $DN_{\lambda'}(\lambda)$ is itself the function of λ wavelengths since we are using a hyperspectral camera. $E_{\lambda'}(\lambda)$ is the spectral radiance from CS2000A telespectroradiometer and $S_{meas}(\lambda')$ is the measured spectral sensitivity for the camera at λ'. The peaks of these 21 curves were then interpolated and the resultant spectral sensitivity curve was obtained. However, the spectral sensitivity curve was converted to the quantum efficiency (QE) represented by the blue curve as shown in the graph in Figure 4.10 where the red curve represents the ground truth quantum efficiency(QE) curve of HySpexVNIR1800.

Both the estimated curve and provided QE curve are normalized. The reason for converting spectral sensitivity to QE is that the QE curve is the only available data provided by the manufacturer(i.e., stored in the header file of HyspexVNIR1800 output) for verification. We converted the spectral sensitivity to quantum efficiency Q_{meas} using equation 4.4 as given below

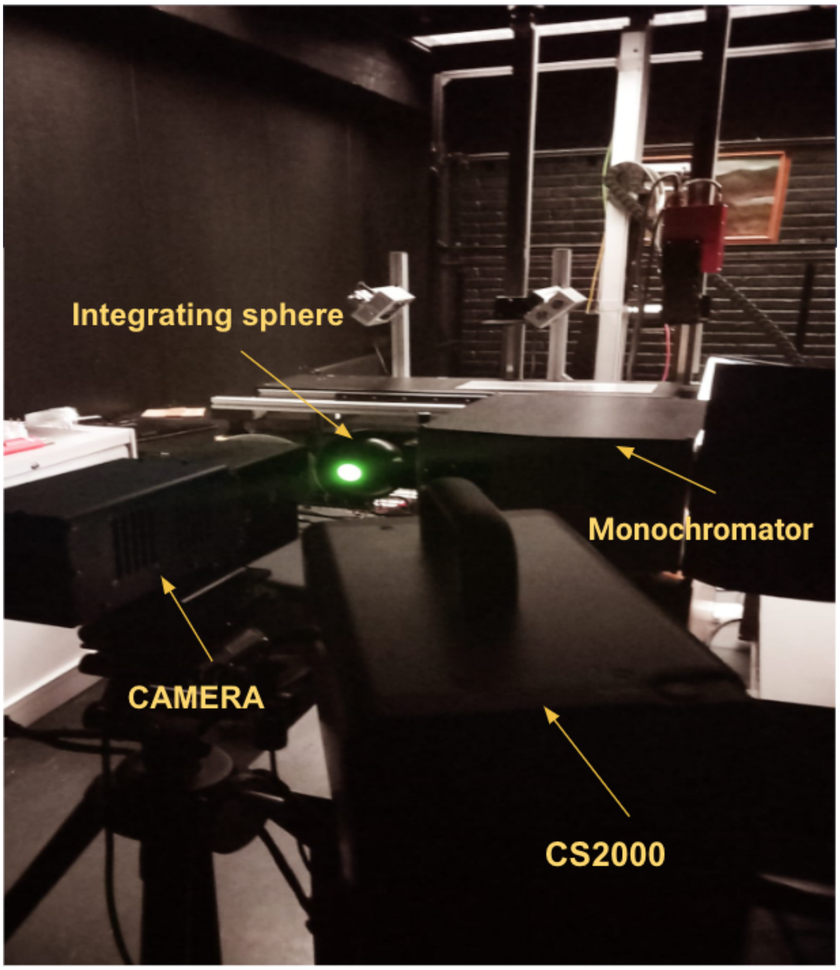


Figure 4.9: *To estimate the spectral sensitivity curve, 21 measurements of monochrome lights were taken using Bentham MT300 monochromator by impinging monochrome light into integrating sphere, and the HyspexVNIR1800 camera was used to capture them. CS2000 spectroradiometer was employed to measure spectral power distribution (SPD) at each of these wavelengths ranging from 440nm to 780nm with a stepsize of 20nm.*

$$Q_{meas}(\lambda') = \frac{S_{meas}(\lambda')}{\lambda'} \tag{4.4}$$

The quantum efficiency (QE) is the ratio of the actual number of photons detected to the number of incident photons. In the imaging sensor such as CCD and CMOS, this QE quantity varies with wavelength. In the range of wavelengths from 300nm up to 900 nm typical this values fall in the range of 0.2–0.75 with maximum efficiency around 500 nm.

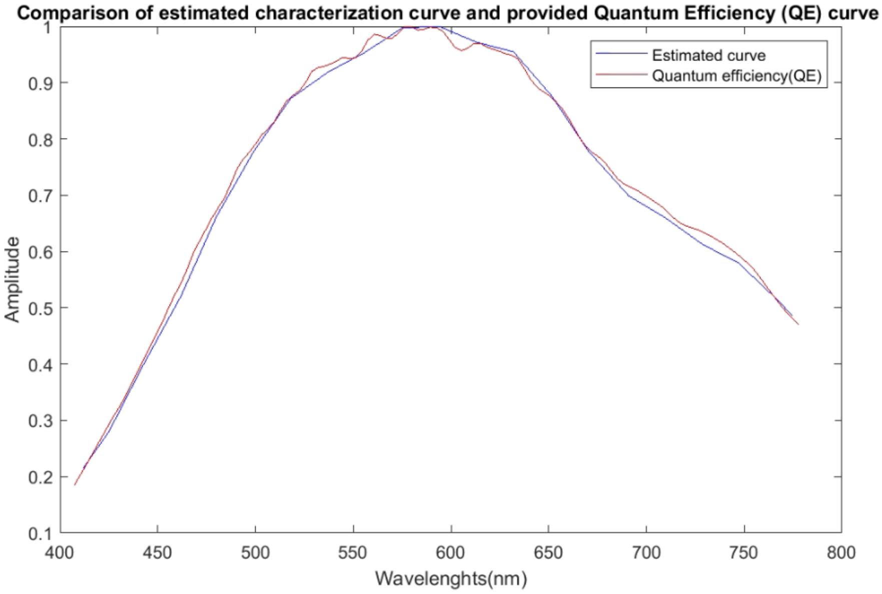
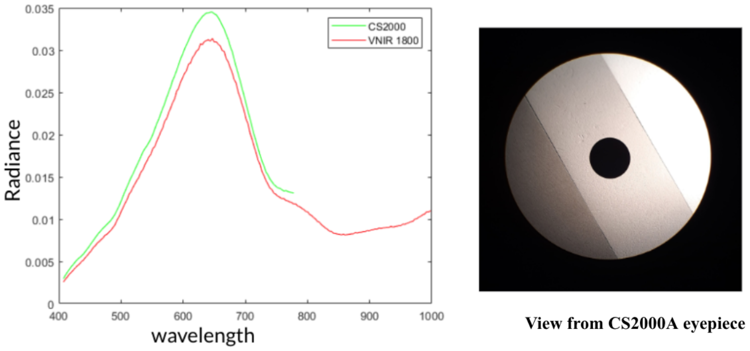


Figure 4.10: *The red curve is a quantum efficiency curve(stored in the header file of the camera's output) and the blue curve is the estimated characterization curve.*



View from CS2000A eyepiece

Figure 4.11: *CS2000A telespectroradiometer vs. HySpexVNIR1800*

4.3.3 HyspexVNIR1800 Radiance verification

In order to verify the accuracy of radiance measurements obtained from the VNIR 1800 hyperspectral, a comparative analysis was conducted by comparing the radiance measurements acquired from a white tile using both the VNIR 1800 and CS2000A radiometric instrument. The results of this comparison are presented in Figure 4.11.

It is worth noting that some data points are missing from the measured CS2000A readings. This is due to the limited wavelength range of the CS2000A telespectroradiometer, which only extends up to 780nm. In order to ensure a fair and consistent comparison, we performed an average calculation over a circular region for both instruments. This was done to account for the fact that CS2000 also averages the radiances over a circular region, thereby allowing a meaningful and equitable evaluation between the two instruments. The average absolute difference between both measurements is $1.8e-3$ and the standard deviation of the absolute differences is $8.54e-4$ which is acceptable for our application.

4.3.4 Noise correction:

Every unprocessed image produced by an image sensor contains at least two types of sensor artifacts. The objective of the noise correction process is to eliminate these artifacts so that the resulting image is a precise representation of the quantity of light that struck the chip sensor. According to fundamentals of optoelectronics (Pollock, 1995), the basic working principle of the imaging sensor is that it is a photon-counting device that counts the photoelectrons produced due to the photoelectric effect. The assumption that these photoelectrons are the only source of the accumulated charge, is challenged by the existence of dark currents. Dark currents occur naturally in semiconductors through the thermal generation of charge carriers. Even in the absence of light, collisions between atoms can excite electrons into the conduction band. The result is electrons residual present in every pixel which refer to as dark current(thermal electrons) as illustrated in equation 4.5. Certain pixels accumulate an excessive quantity of these free electrons which is dependent on the quality of the silicon and fabrication used in the production of the chip. Such pixels are known as hot pixels. This issue is more pronounced in CCD sensors where charge is transferred row by row and due to the delay in pixel readout, thermal motions generate free extra electrons.

$$\text{Measured electrons} = (\text{Thermal electrons}) + (\text{Photo-electrons}) \tag{4.5}$$

The dark current correction can be performed by taking the average of three dark frames taken at the identical settings which were used while capturing the HDR scene (i.e., 20 ms, 40 ms, and 440 ms in our case). Then these averaged dark frames were subtracted from subsequent hyperspectral images. Another issue with the image sensor is the variations in the sensitivity of its pixels. In the scenario when 100 photons impinge onto pixel A and an equal number of photons likewise impinge upon the adjacent pixel B, it is seen that pixel A yields a readout of 55 electrons, whereas pixel B yields a readout of 56 electrons. The observed discrepancy, although typically negligible but might have significant effects in the field of photometry. The hyperspectral camera itself performs dark current correction and cater pixel non-uniformity as discussed in Section 4.1.3.

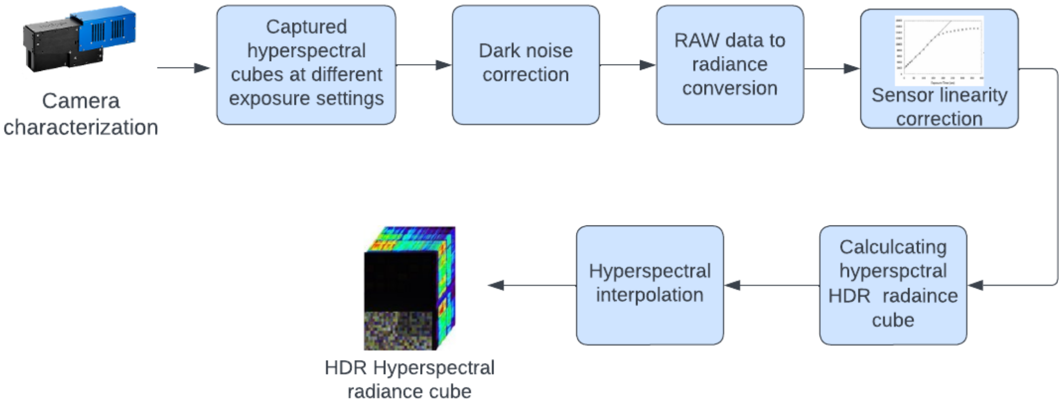


Figure 4.12: *The workflow for processing hyperspectral data to calculate hyperspectral HDR radiance cube.*

4.4 Hyperspectral data acquisition and processing

The workflow of the hyperspectral data processing for capturing and processing the HDR hyperspectral radiance image of our physical HDR scene is given in Figure 4.12 below.

In order to adequately capture the maximum details of this HDR scene, a series of hyperspectral cubes were acquired using the hyperspectral camera, employing various integration settings. We incorporated three different exposure settings as it is the most frequently used in existing literature (Mann et al., 2012; Bouderbane et al., 2016). Also, these individual hyperspectral images were acquired to ensure that the detailed information pertaining in dark shadows, mid-tones, and bright regions of the HDR scene was captured for better estimation of HDR radiances. In particular, two hyperspectral cubes were acquired using integration times of 20 milliseconds and 40 milliseconds, respectively. These short exposure settings were selected to effectively capture the nuanced details present in the brighter regions of the HDR scene. Additionally, an integration time of 440 milliseconds was employed to acquire a hyperspectral cube with longer integration, enabling the capture of details within the darker regions of the HDR scene as shown in Figure 4.13.

HySpexVNIR1800 stores the captured data in the form of raw digital numbers in 16-bit format. However, these raw images can be converted to calibrated digital numbers and absolute radiances using *HySpexRAD* software as discussed in 4.1.2 and verified during the camera characterization as given in 4.3. The size of each

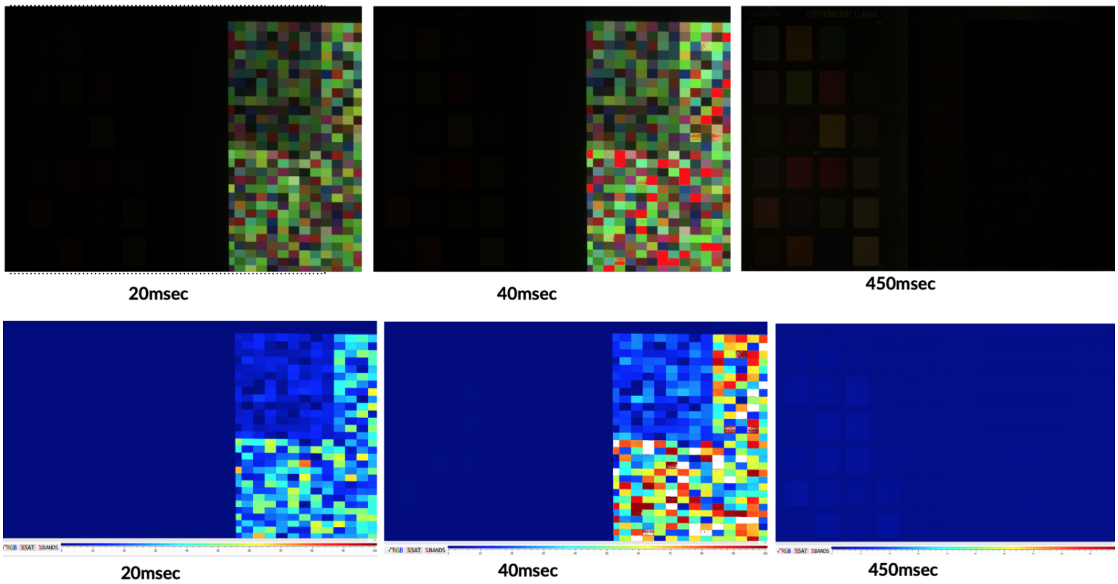


Figure 4.13: *The preview of hyperspectral captures along with their heatmap.*

3D hyperspectral cube is 3976-by-1800-by-186 pixels where 186 is the number of bands that lie within the range of visible and near-infrared wavelengths (i.e., from 407.469nm to 998.107nm), 3976 is the total number of lines captured and 1800 are the number of pixels per line as discussed in the previous section and demonstrated in Figure 4.2. Each of these hyperspectral images was then converted to radiances and was stored in 32-bit float format as demonstrated in Figure 4.4. The transition from raw data to absolute radiance is performed via $HySpexRAD$ software which has embedded the equation 4.6 that contains the parameters that are acquisition-dependent, and variables that are permanently stored in the software.

$$L(i,j) = \frac{\frac{DN(i,j) - BG(i,j)}{RE(i,j)} \cdot h \cdot c}{QE(i) \cdot SF \cdot t \cdot A \cdot \Omega \cdot \lambda(i) \cdot \Delta\lambda(i)} \tag{4.6}$$

where Digital Number (DN), background matrix (BG), and integration time t are the only parameters that depend on data capture settings while all other variables are permanently stored in the camera's software (details of these parameters are available in the user manual of HySpexVNIR1800).

The radiance values corresponding to the digital count less than 40,242 (i.e., the radiances which are calculated from digital numbers within the linearity of the sensor) are valid radiances. In contrast, the radiances corresponding to the digital number greater than 40,242 are corrected up to 63,000 digital counts (which was an empirically defined threshold up to which the linearity correction has been applied as discussed in Section 4.3). The radiances corresponding to the digital number

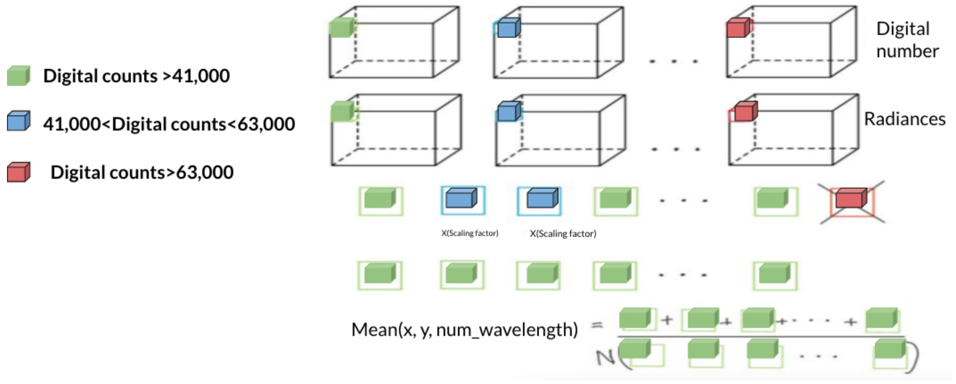
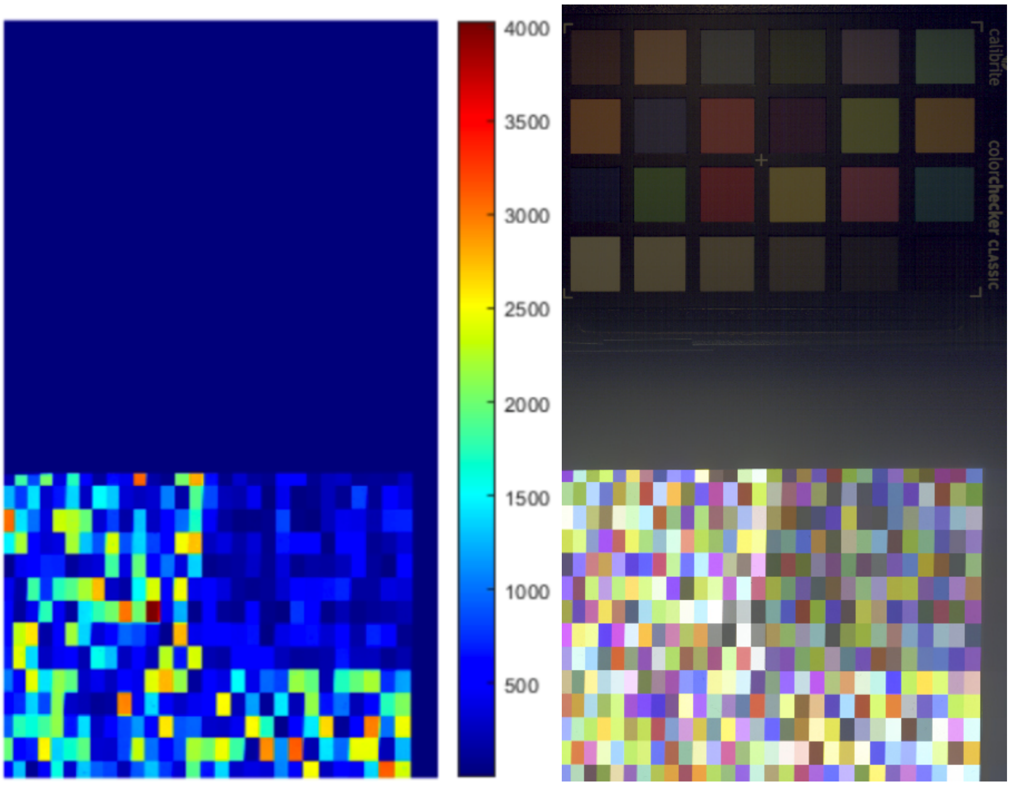


Figure 4.14: *The figure shows radiance values corresponding to the digital count less than 40,242(green boxes) are calculated from digital numbers within the linearity of the sensor and are valid radiances. In contrast, the radiances corresponding to the digital number greater than 40,242 (blue boxes) are corrected up to 63,000 digital counts. The 63,000 was an empirically defined threshold up to which the linearity correction has been applied as discussed in Section 4.3.1.1). The radiances corresponding to the digital number beyond that are invalid and hence they are clipped before merging these hyperspectral cubes by taking an average for estimating an HDR hyperspectral radiance image of our real-world HDR scene.*

beyond that were invalid and hence they were clipped as illustrated in Figure4.14 before merging these hyperspectral radiances cubes.

Finally, these three hyperspectral cubes are merged into an HDR hyperspectral radiance image by simply taking an average of these individual hyperspectral cubes. The average of individual hyperspectral captures also helped to improve the single-to-noise ratio in the dark regions and suppressed the artifacts due to the non-uniformity of pixel sensitivities of the imaging sensor. The hyperspectral interpolation was performed to account for the missing wavelengths in this radiance mean cube due to the saturation of the pixels. Although, this is not the case in this HDR scene but in some of the previous HDR scenes (also given in the proposed dataset), we had speckles that were saturating the pixel. Finally, this gave us the absolute hyperspectral radiance of the physical HDR scene which was stored in a 32-bit float format using MATLAB *.mat* file while the luminance heatmap of this HDR hyperspectral image is shown in Figure 4.15.



(a) (b)

Figure 4.15: *The figure illustrates the luminance heatmap of calculated HDR hyperspectral radiance image of the physical HDR scene for visualization, where the red patch corresponds to the maximum luminance in the scene which is approximately equal to* $4000cd/m^2$.

4.5 Hyperspectral HDR Radiances to three channel HDR image(Scene-referred)

The calculated HDR hyperspectral radiance cube as mentioned above, can be converted to HDR three-channel RGB image using the proposed workflow as shown in Figure 4.16 which maps the visible spectrum of HDR hyperspectral radiance cube to the XYZ color space and from there to the Rec. 2020 color space. To recall, this HDR hyperspectral radiance cube consisted of 186 bands with wavelengths ranging from 400nm up to 1000nm with a step size of 3.26nm and was stored in a 32-bit float *.mat* file. In this workflow, the dot product of HDR hyperspectral radiance cube is taken with 1931 2° CMF(color matching function) to obtain three

Figure 4.16: *The figure illustrates the proposed workflow for converting hyperspectral HDR radiance cube into HDR three-channel RGB image which is finally stored in 32-bit float using .hdr format which is basically a mapping of the visible spectrum of HDR hyperspectral radiance cube to the XYZ color space and from there to the Rec. 2020 color space.*

channel XYZ tristimulus values as given by equation 4.7:

$$\begin{aligned} X(x,y) &= \sum_{\lambda=380}^{780} H(x,y,\lambda)\bar{x}(\lambda) \\ Y(x,y) &= \sum_{\lambda=380}^{780} H(x,y,\lambda)\bar{y}(\lambda) \\ Z(x,y) &= \sum_{\lambda=380}^{780} H(x,y,\lambda)\bar{z}(\lambda) \end{aligned} \tag{4.7}$$

where H is the HDR hyperspectral radiance cube which is a function of spatial pixel coordinates: x and y, and spectral bands λ, and $\bar{x}$, $\bar{y}$, $\bar{z}$ are the 1931 2° CMF. Also, note that the summation limit spans from 380 nm to 780 nm even though the HDR hyperspectral radiance cube is defined up to 1000nm. The wavelength is limited to 780nm because CMFs are defined in the visible region only.

This converts the spectral bands that are within the visible range from the hyperspectral radiance cube to the CIE XYZ color space(See Chapter 2 details of CMF and CIE 1931 color space). Next, these XYZ tristimulus values are converted to ITU-R Recommendation BT.2020 which is commonly known as Rec. 2020 color space.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \boldsymbol{M} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4.8}$$

Hence, the 3x3 transformation matrix M of Rec2020 has been employed to convert these XYZ channels to the linear three-channel RGB image as represented by equation 4.8. This linear RGB image is then normalized with its maximum and referred to as $linRGBNorm$ in the workflow. Finally, $linRGBNorm$ is stored in a 32-bit float *.hdr* format which is the HDR radiance file format. The MATLAB *.hdr* format which is used to store the 3-channel HDR image, uses RGBE pixel encoding where the first three bytes of the RGBE pixel encoding are used to represent the red,

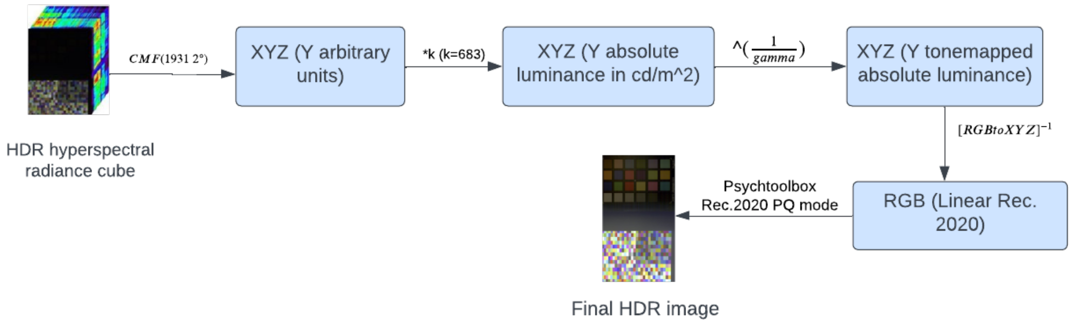


Figure 4.17: *The workflow illustrates the steps involved in the proposed display referred HDR rendering workflow for converting the hyperspectral HDR radiances cube of an HDR scene into three channel HDR image for HDR display.*

green, and blue color channels, while the last byte serves as a common exponent for all channels. RGBE encoding takes advantage of the fact that three is a high correlation among all color channels in the RGB color spaces and have values of at least the same order of magnitude. Consequently, there is no need to record a different exponent for each color channel.

4.6 Proposed display-referred HDR rendering workflow:

The HDR hyperspectral radiance cube contains 186 spectral bands as mentioned above. These hyperspectral bands are required to map to three-channel HDR images that would eventually be rendered on the HDR monitor. The proposed display referred HDR rendering workflow is given in Figure 4.17 which illustrates the required processing steps for rendering HDR images on an HDR display. Firstly, the dot product of the HDR hyperspectral radiance image is taken with 1931 2° CMF to get three channel tristimulus values. Since our hyperspectral HDR cube is a radiometric quantity, the tristimulus values need to be multiplied by 683 (Wyszecki and Stiles, 2000) to obtain absolute luminance where the unit of luminance is cd/m^2.

To display the HDR image, we used a 24-inch HDR Sony display that has a maximum luminance level of 1000 cd/m^2 and was calibrated to Rec. 2020 PQ mode. It is important to note that the maximum luminance of the physical HDR scene was measured to be approximately 4000cd/m^2 using a CS2000A telespectroradiometer and also demonstrated in the luminance heat map given in Figure 4.15 whereas the maximum luminance of our HDR display is 1000cd/m^2. Therefore tone mapping is required to compress the luminance within the capabilities of the available HDR display screen. Hence, a gamma of 1.2 has been utilized to compress the dynamic range of the XYZ tristimulus values so that the maximum luminance would be within the display capabilities. Then, the inverse of the display colorimetric transformation matrix as given by equation 4.9 has been applied for converting XYZ tristimulus values to display linear RGB.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M^{-1}_{RGBtoXYZ} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4.9}$$

The display 3x3 matrix $M_{RGBtoXYZ}$ has been calculated through display calibration as discussed in Section 6.5 of Chapter 6. The display calibration also accounts for the fact that the display sometimes does not strictly follow the targeted color space and deviates from the theoretically perfect system (Berns, 1996). This also makes it important to know the actual colorimetry of the display so that we would be able to correctly transform our XYZ tristimulus data to display $R'G'B'$ pixel values. These negative values of $R'G'B'$ are clipped down to zero as defined in equation 4.10.

$$\begin{aligned} R_{display} &= \max{(0, R')} \\ G_{display} &= \max{(0, G')} \\ B_{display} &= \max{(0, B')} \end{aligned} \tag{4.10}$$

In our case, the EOTF is simulated by PQ curve (Froehlich et al., 2015; Nezamabadi et al., 2014) as the display was calibrated to Rec. 2020 PQ mode. It is to be noted that the final HDR image was displayed using MATLAB Psychtoolbox-3, which assumes a linear Rec. 2020 RGB as input and takes care of PQ encoding itself as shown in Figure 4.17.

4.7 Comparing proposed HDR rendering with ITU recommendation

The workflow illustrated in Figure 4.18 is the conversion of BT.2100 to arbitrary linear color signals for display systems. This conversion process assumes a display referred workflow for both PQ and HLG. The full report is available at `https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-BT.2408-5-2022-PDF-E.pdf`.

It is interesting to compare our proposed HDR rendering as given in the above Section 4.6 with this ITU-recommended HDR rendering workflow as it defines the procedure that can be used to convert linear and normalized BT.2100 RGB

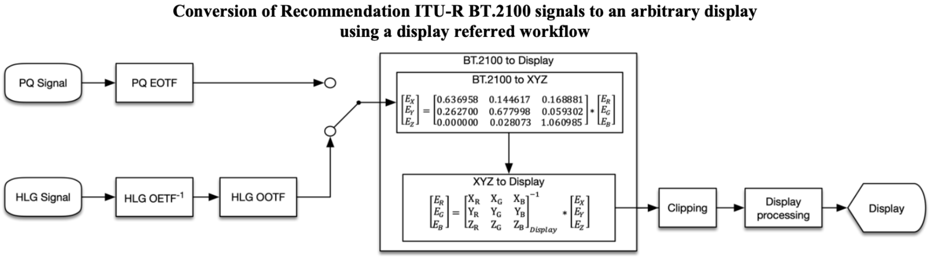


Figure 4.18: *The figure illustrates the steps for conversion of ITU-R BT.2100 HDR signals to an arbitrary display using a display referred workflow*

primaries to the RGB primaries of any display system. It is also important to mention that the input signal is already within the range of 1000 cd/m^2 which is achieved using a display referred camera workflow(refer to report to see details). In this diagram, the input signal is first linearized by applying PQ EOTF and then the equation 4.11 has been utilized for converting the linearized BT.2100 input primaries to display RGB primaries. In this equation, the input BT.2100 RGB is transformed to XYZ using BT.2100 transformation matrix as shown in Figure 4.18. Then the XYZ to display transformation is applied using the inverse of the 3x3 display colorimetric transformation matrix.

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{\text{display}} = \begin{bmatrix} \mathrm{X_R} & \mathrm{X_G} & \mathrm{X_B} \\ \mathrm{Y_R} & \mathrm{Y_G} & \mathrm{Y_B} \\ \mathrm{Z_R} & \mathrm{Z_G} & \mathrm{Z_B} \end{bmatrix}^{-1}_{\text{display}} * \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0.0000 & 0.0281 & 1.0610 \end{bmatrix} * \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix} \tag{4.11}$$

It is likely that not every color from the source color space may be included in the display color gamut. Negative values may be clipped down to zero. Positive values may also be trimmed to the display's capabilities. Although both soft and hard clipping are possible, hard clipping is preferred in many applications, such as when using a reference display as in our case. In this workflow, only negative values are subjected to severe clipping as given by the equation 4.12. These clipped RGB values are then sent to the display for its processing and finally, the output is shown on the screen.

$$\begin{aligned} E_R &= \mathrm{Max}\,(0, E_R) \\ E_G &= \mathrm{Max}\,(0, E_G) \\ E_B &= \mathrm{Max}\,(0, E_B) \end{aligned} \tag{4.12}$$

In our proposed HDR rendering as discussed in Section 4.6, it can be observed that we have almost followed similar transformation steps, and the only additional step in our workflow is applying a tone-mapping or tone compression to the XYZ tristimulus values as illustrated in Figure 4.19. The value of gamma is calculated using the equation 4.13 to map the wide range of HDR values to the limited dynamic range of the available HDR display where n is the maximum luminance of a physical HDR scene and m is the display maximum luminance. In our case, the n is equal to 4000cd/m^2 whereas the value of m is set to 1000cd/m^2.

$$gamma = \frac{\log_{10}(n)}{\log_{10}(m)} \tag{4.13}$$

Furthermore, the display was set to Rec. 2020 PQ mode, and the HDR image was rendered using Psychtoolbox-3. The Psychtoolbox-3 assumes a linear Rec. 2020 RGB as input and takes care of PQ encoding itself. These PQ-encoded RGB values are decoded at the display end while rendering them on the HDR screen.

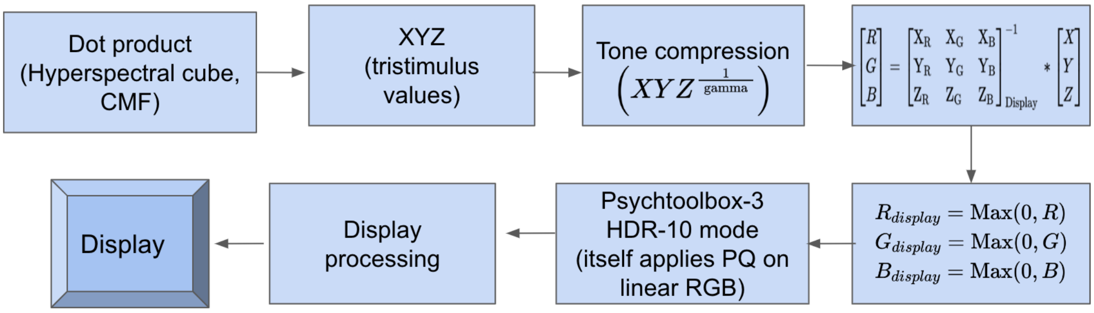
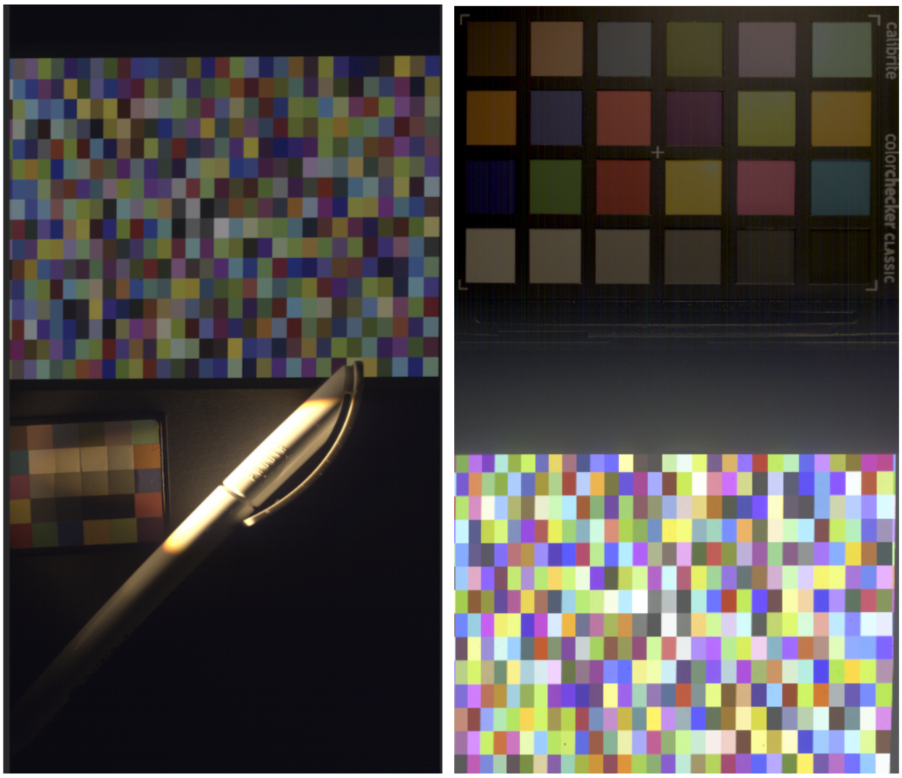


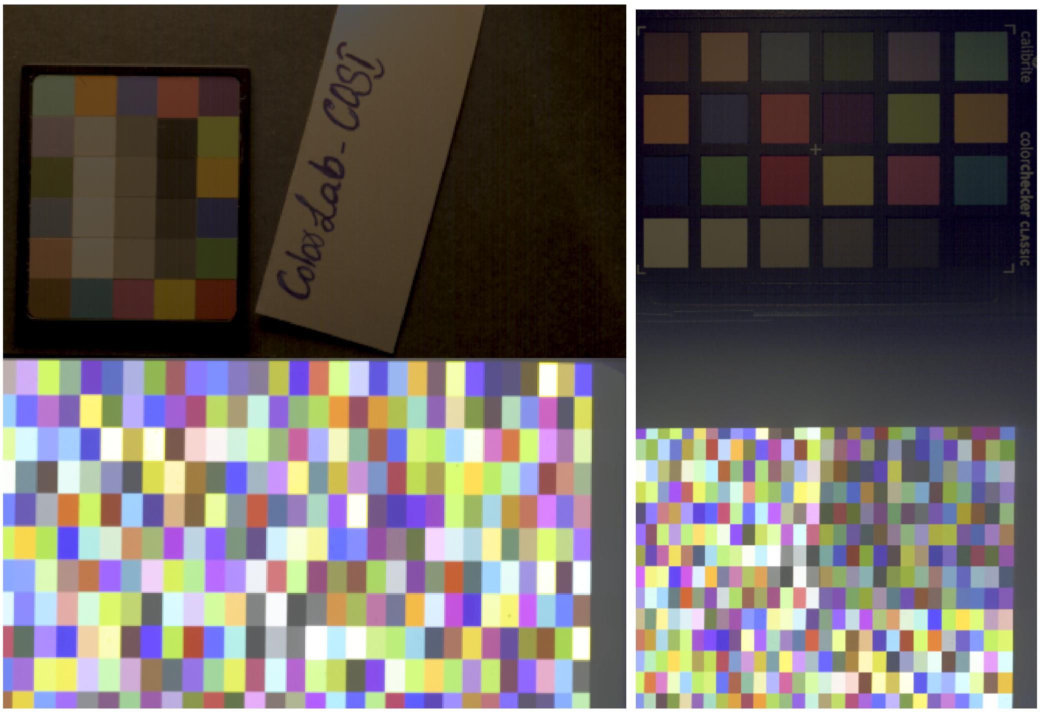
Figure 4.19: *The figure shows the transformation steps involved in our workflow for conversion of HDR hyperspectral radiance cube to display HDR RGB image.*

4.8 Proposed dataset:

As a contribution to extended research in this domain, a dataset containing four HDR hyperspectral radiances cubes and their respective three-channel HDR images has been proposed where each of them consisted of a different HDR scene captured in an indoor setting using HyspexVNIR1800 camera. In Figure 4.20, the tone-mapped images of the four HDR hyperspectral radiance cubes are illustrated whereas the corresponding luminance heatmaps derived from these HDR hyperspectral cubes in our dataset are presented in 4.21 (Please email at ah1333@g.rit.edu to access the dataset).

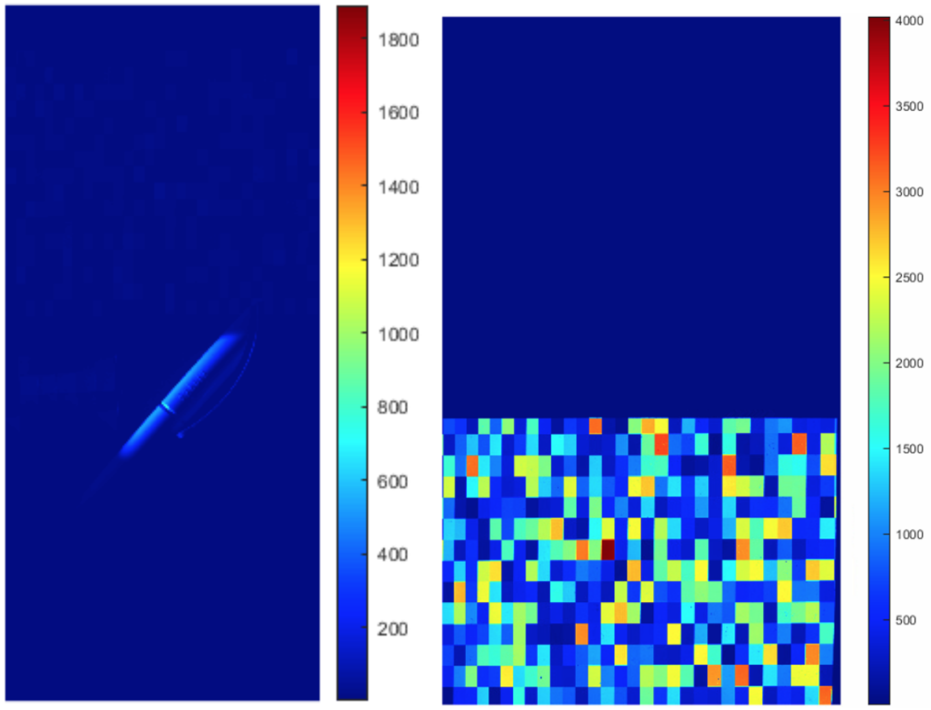


(a) *Pen image* **(b)** *backlight image*

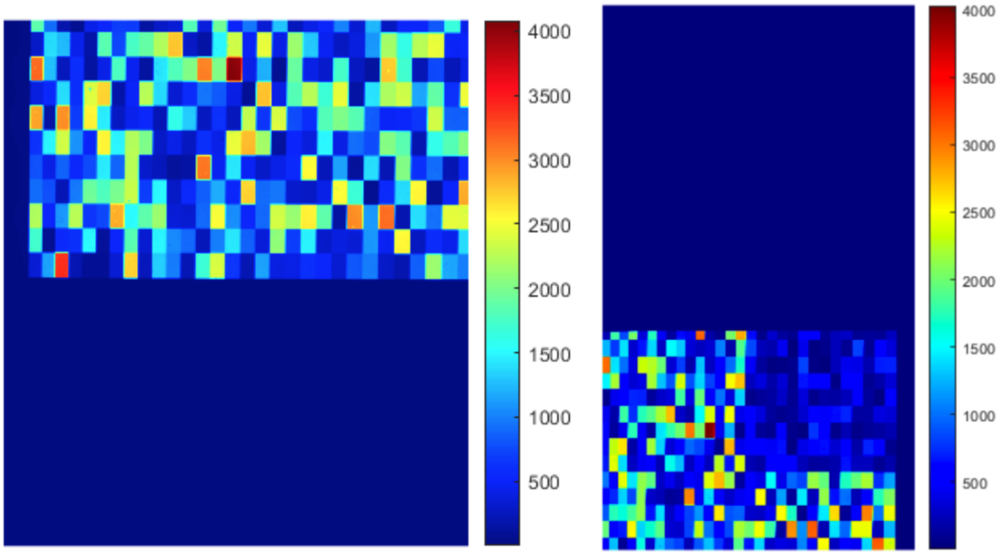


(c) *COSI image* **(d)** *Backlight Midtones image*

Figure 4.20: *The proposed dataset contains four HDR hyperspectral radiance cubes along which their corresponding three-channel HDR images(note that these are tone-mapped images for visualization).*



(a) *Pen image* **(b)** *backlight image*



(c) *COSI image* **(d)** *Backlight Midtones image*

Figure 4.21: *The figure shows the luminance heatmap of the HDR hyperspectral cubes in the proposed dataset.*

5 Tone mapping and Proposed SiCAM model

> *The results of all truly great actions are often accomplished in the minds of those who do them long before the work is done*
>
> **Albert Einstein**

Tone-mapping operators (TMOs), also known as HDR rendering algorithms, are designed to adjust the wide range of luminance information found in real-world scenes so that it can be rendered on display screens with the limitation of producing only lower dynamic range (LDR) images. Tone-mapping objectives might vary substantially depending on the application and discipline. In medical imaging, for example, tone mapping algorithms aim to preserve as much details as possible, which is important for radiologists analyzing medical MR images. The objectives of digital photography may include but are not limited to, producing visually appealing and preferred images, replicating the overall appearance of the original scene on the display, preserving image contrast, and might also include predicting the visibility of artifacts due to tone compression.

Several tone-mapping methods have been proposed in the past while the underlying principles and assumptions of numerous operators in tone-mapping share notable similarities which are discussed briefly in this chapter. Additionally, as image appearance models attempt to predict the perceptual response to spatially complex stimuli, they can provide a unique framework for the prediction of HDR image appearance. Fairchild proposed an image appearance model called iCAM06, specifically designed for rendering High Dynamic Range (HDR) images. This model is built upon the iCAM framework and integrates spatial processing models found in the human visual system. It incorporates contrast enhancement techniques, photoreceptor light adaptation functions that enhance local details in highlights and shadows, as well as functions that predict various color appearance phenomena. Through rigorous evaluation in past research, iCAM06 has consistently demonstrated excellent performance in HDR rendering, both in terms of preference and accuracy. As a result, iCAM06 is a promising choice as a general-purpose tone-mapping operator. Moreover, its potential applications extend to a wide range of image appearance research and practical use. Due to the above mention reasons, an extension of this model namely the spectral image color appearance model (SiCAM) has been proposed in this work, for tone mapping of hyperspectral HDR radiances to generate better tone-accurate LDR tone-mapped images. The main objective of this proposed model is to keep the working principles of iCAM06 (Kuang et al., 2007b) intact and introduce spectral radiances so the appearance pipeline can benefit from the additional information in spectral radiances in comparison to a three-channel image. Additionally, we have proposed a hyperspectral adaptation method in comparison to chromatic adaptation.

This chapter addresses an overview of tone mapping along with the summary of different tone mapping approaches. Refer to Banterle et al. (2017); Eilertsen et al. (2017b); Salih et al. (2012); Mantiuk et al. (2007) for a detailed review of tone-mapping algorithms. Besides other widely used tone mapping algorithms, it also discussed the iCAM06 HDR image rendering model in detail and the proposed SiCAM model which takes a HDR hyperspectral radiance cube as input and generates a three-channel LDR image.

5.1 An overview of tone mapping

Tone mapping is the process of reproducing high-contrast scenes with a potentially wide color gamut on a destination media with limited contrast and color reproduction. It usually entails converting high dynamic range images/videos containing scene radiance or brightness into pixel values that can be shown on a Low Dynamic Range (LDR) display screen. There are two stages in common regardless of whatever category the tone mapping operator (TMO) belongs to. The first step is to extract luminance information from the input HDR image or video frame. This is due to the fact that a TMO normally works on the luminance channel, avoiding color compression. The second step is to restore color information in the compressed image. However, tone-mapping algorithms strive for a wide range of objectives, methodologies, and applications, which are discussed in the following subsections.

5.1.1 Mathematics of tone mapping

A tone mapping operator is a mathematical function that converts HDR scene luminance into the level of luminance range that can be displayed on a limited dynamic range display as depicted in equation 5.1 where LDR is the low dynamic range output image which is in 8 bits(i.e., 0 to 255).

$$f(Image) : \mathbb{HDR}_i^{rows\times columns\times bands} \rightarrow \mathbb{LDR}_o^{rows\times columns\times bands} \tag{5.1}$$

To develop a proper understanding of tone-mapping mechanisms, it is crucial to comprehend how the shape of the tone-mapping function impacts the visual perception of output tone-mapped images. The summary of fundamental mathematical operators used in tone mapping and their effect on the final output is given in Table 6.1. The multiplication of luminance values alters image brightness but has no effect on dynamic range or contrast. As a result, the brightness adjustment parameter B will increase or decrease the overall image brightness as a result of the multiplication. Since the process is similar to adjusting the exposure settings on a camera, it is commonly referred to as "exposure adjustment." The dynamic range of an image can be modified using the Power function. Since reducing the dynamic range has the same effect as reducing the contrast of an image, the two terms are sometimes used interchangeably. The contrast of an image can be adjusted using contrast adjustment parameter C while Ig_{white} is usually assumed to be the scene luminance that is mapped to the peak luminance of a display. The change is relative to the luminance of a reference white point thus the contrast shrinks or expands towards or away from that point. In the literature, this operator is sometimes referred to as gamma correction since the formula is comparable to the display model with the exponent equal to gamma values of 2.2 or 2.4. As a result of the Weber law, adding a constant value to an image has little effect on the bright image areas but has a large effect on the darker image regions. The addition creates a fog or uniform glare impression in an image. The operation will change the contrast and brightness of the image's darkest areas.

Table 5.1: *The given below table presents a comprehensive overview of the fundamental mathematical operators employed in tone mapping and their impact on the final output.*

Operator	***logarithmic domain***	***Visual effect***	***Analogous***
Multiplication	$f(Ig) = Ig + B$	image brightness	exposure adjustment
Power function	$f(Ig) = c(Ig - Ig_{white})$	image contrast	gamma correction
Addition	no representation	both contrast and brightness of darker regions	fog or a uniform glare

5.1.2 Categories of tone mapping

The underlying principles and assumptions of numerous operators in tone mapping share notable similarities. However, in terms of processing, tone mapping operators(TMOs) can be generally categorized into three main categories.

5.1.2.1 Global TMO

The global tone mapping methods are spatially invariant, which means that the processing is identical for all pixel values within an image regardless of the varying pixel values of its neighbors (Banterle et al., 2017; Eilertsen et al., 2016; Mantiuk et al., 2006).

Sometimes the operator will do preliminary processing of the input image to calculate image statistics, which are then utilized for the dynamic range reduction. Maximum brightness, minimum luminance, and logarithmic or arithmetic average values are some common statistics calculated for this type of tone mapping. These statistics are calculated using percentiles to boost robustness and avoid outliers, notably for lowest and maximum values, which may have been impacted by noise during image acquisition. After obtaining the global statistics from the image's luminance, a logarithmic or exponential-like function is used in the pipeline to compute the tone-mapped image.

The fundamental disadvantage of global operators is that, because they use global image statistics, they cannot preserve local contrast and fine details of the original HDR image.

5.1.2.2 Local TMO

Local operators outperform global operators in terms of image quality by attempting to mimic both local and global contrast. This is accomplished by enabling the tone mapping operator to consider local characteristics of the neighborhood surrounding the pixel being tone mapped. However, neighbors must be carefully chosen, i.e., only pixels that are comparable or belong to the same region. Otherwise, local statistics will be skewed, resulting in halo artifacts. These artifacts are most common at boundaries that separate sections with varying intensities and/or color values. It should be noted that halos are often desired when focusing attention on a specific location (Banterle et al., 2017), but if the phenomenon is uncontrolled, it might produce unpleasantness.

5.1.2.3 Segmented TMO

Segmented tone mapping is an approach that divides an image into segments or areas based on criteria such as color similarity or texture coherence. Each segment is then processed independently using a tone mapping operator to change the tone and contrast within that segment. This method enables localized tone mapping modifications that take into consideration the unique qualities and content of distinct sections of the image.

Segmented tone mapping can generate more accurate and visually appealing results by applying distinct tone mapping operators to different segments, especially in scenes with varying levels of brightness and contrast. Another advantage of this method is that gamut modifications are minimized because a linear operator for each segment is often sufficient.

5.1.2.4 Frequency/gradient TMO

The goal of frequency-based/gradient-domain operators is the same as that of local operators: to preserve edges and local contrast. unlike local TMO, dynamic range compression is applied in the frequency/gradient-domain rather than the spatial domain. This method divides the images into low frequencies and high frequencies. Low frequencies are typically tone mapped, while high frequencies (details) are either lightly processed or left unaltered. The higher frequencies usually persevere because they contain fine details of an image. The essential finding for such approaches is that edges and local contrast are preserved only when there is a proper separation between large features and fine details.

5.1.2.5 Others categories

Further categorization might be made based on the TMO's design philosophy or its use as given below:

- **Perception operators** These are tone-mapping methods inspired by Human Visual System(HVS) and aim to model several aspects of it such as chromatic adaptation, temporal adaptation, limited visual acuity at night, degrading of visual due to aging, etc. These perception operators can be global, local, segmentation, or frequency/gradient.

- **Empirical operators:** The main focus of these operators is to generate pleasing outputs inspired by other fields, such as photography, and are usually based on empirical or subjective analysis. Like perception operators, these TMOs can also be of any category such as global, local, segmented, or gradient-based.

5.1.3 Intents of tone mapping

Tone-mapping objectives might vary substantially depending on the application and discipline. Because the range of aims is the source of most uncertainty and misunderstanding about tone-mapping, it is critical to precisely describe these goals. Tone-mapping operators can be generally classified based on their intent.

- **Visual system simulation:** These perception operators imitate the visual system's limitations and properties. A tone mapping operator (TMO), for example, can generate glare, replicate the limitations of human night vision, or lower color and contrast in dark scene sections. Another example is image correction for the difference between real-world scene adaptation conditions and viewing conditions (including chromatic adaptation).

- **Accurate scene reproduction:** When an image is displayed on a device with a reduced color gamut, contrast, and peak luminance, the original scene appearance, including contrast, detail, and colors, is preserved. These operators do not attempt to simulate changes in appearance caused by perceptual effects, such as nighttime loss of acuity and color vision. Instead, they concentrate on overcoming the limitations of the output medium in order to obtain the best possible color gamut and dynamic range match. This is important for maintaining the visual fidelity and realism of the scene.

- **Best subjective quality (BSQ) operators:** Tone mapping operators also consider the subjective visual perception of the image. They are designed to generate the most pleasing images or videos based on subjective preference or artistic objectives. These operators typically include a set of modifiable parameters that can be altered empirically in accordance with artistic goals. Such an example would be photo editing software such as Adobe Photoshop Lightroom. They may incorporate perceptual models to ensure the resulting image is visually pleasing and avoids artifacts such as halo artifacts, noise amplification, or unnatural contrast adjustments.

5.1.4 Basic approaches of tone mapping

In recent years, hundreds of papers on tone mapping have been published, providing an abundance of options. Nonetheless, many of these operators share very similar underlying assumptions and mechanisms. Instead of discussing various tone-mapping algorithms, In this section, a summary of the main approaches is provided along with examples of operators that they rely upon. Refer to Banterle et al. (2017); Eilertsen et al. (2017b); Salih et al. (2012); Mantiuk et al. (2007) for a detailed review of tone-mapping algorithms.

5.1.4.1 Illumination and reflectance separation

The pixel values for most diffuse objects can be thought of as a multiplication of the amount of light hitting the sensor (the irradiance) and the amount of light reflected back from the surface(surface reflectance). This is a simplified model that disregards geometry and more complex reflectance properties, but it is extensively employed in computer vision and related fields. If we know how to separate illumination and reflectance, we can construct a tone mapping operator that only modifies the illumination component without distorting reflectance.

Restricting alterations to the illumination component is particularly advantageous for tone mapping, as illumination is primarily responsible for the wide dynamic range of real-world scenes while reflectance provides information about an object's shape, texture, and color as well as is generally invariant to the viewing conditions. Originally, this approach to dynamic range compression was proposed by Oppenheim et al. (1968).

5.1.4.2 Low-pass filter decomposition

The main challenge is separating illumination from reflectance. Given only an image, it is impossible to arrive at a precise solution for such a problem (Salih et al., 2012). There are, however, various approximations that rely on the mathematical characteristics of light in real-world settings. Unlike reflectance, illumination in a scene typically varies uniformly between pixels. The only strong discontinuities to be expected are at the boundaries of harsh shadows and light sources. Convolving an image with a Gaussian filter is the simplest approach to extracting the gradually changing regions of an image.

5.1.4.3 Bilateral filter decomposition:

The use of Gaussian filtering as the illumination-separation operator has a number of limitations, the most significant of which is that it is unable to detect sudden changes in illumination, such as those that occur at the margins of sharp shadows and light sources. Consequently, Gaussian filtering results in the blurring of the edges in an image. The bilateral filter (Durand and Dorsey, 2002a) is an example of one of these edge-preserving operators where the smoothing by employing the bilateral filter is limited not only in the spatial domain but also in the domain of the intensities of the individual pixels.

5.1.4.4 Retinex algorithms:

In order to provide an explanation for the color constancy phenomena (Land and McCann, 1971), the Retinex algorithm was initially proposed. The goal of the Retinex research was to imitate the ability of the HVS to extract reliable details from the environment that humans experience, despite the variations in illumination(i.e., Human Color Constancy). The technique essentially seeks to isolate reflectance from illumination by suppressing tiny gradients that are associated with illumination and is thus an effective method of tone-mapping images (Kim et al., 2011; Meylan and Susstrunk, 2006).

5.1.5 Gradient and contrast-based methods:

Rather than splitting an image into reflectance and illumination layers, it is feasible to enhance image details (reflectance) before compressing image contrast with a gamma function (or linear scaling in the log domain). Several operators that change visual gradients or local contrast have used this method (Farbman et al., 2008; Fattal et al., 2002; Mantiuk et al., 2006). The primary advantage of operating on gradients rather than pixel values is that it allows for dramatic increases in local contrast without causing unpleasant contrast residuals known as halo artifacts. Local gradient manipulation (Fattal et al., 2002), on the other hand, may result in discrepancies in global image brightness between distant image regions. As a result, modern operators (Farbman et al., 2008; Mantiuk et al., 2006) developed multi-scale structures to retain image contrast at different scales of an image.

5.2 iCAM06: HDR image rendering model

Color appearance models have been used in cross-media color reproduction frameworks, but their application is limited as they are not meant to anticipate the visual appearance of complex spatially variable stimuli (i.e., images or videos). For details of the color appearance model, refer to the book titled "Color Appearance Models" by Fairchild (2013). The image color appearance models (iCAM) proposed by Fairchild and Johnson (2002) extend color appearance models by incorporating spatial and temporal vision features, allowing for the prediction of the appearance of complex stimuli. An image appearance model, given an input of images and viewing conditions, can yield perceptual qualities of each pixel, not just the essential color appearance attributes such as brightness, lightness, saturation, chroma, and hue, but also image attributes like its contrast and sharpness.

As image appearance models attempt to predict the perceptual response to spatially complex stimuli, they can provide a unique framework for the prediction of HDR image appearance. Hence, iCAM has been extended for tone mapping of HDR images without altering its fundamental architecture (Johnson and Fairchild, 2003). Recent tone-mapping operator evaluations demonstrated that iCAM is effective in HDR image rendering (Fairchild and Johnson, 2002); however, iCAM does not outperform alternative operators, such as the bilateral filter (Durand and Dorsey, 2002b). It integrates efficient anisotropic filters and simple tone compression techniques. Even in terms of reproduction accuracy where iCAM claimed to be the most applicable, iCAM renderings were shown to have less local contrast and colorfulness than real scenes. This has motivated further advancement of iCAM which resulted in proposing iCAM06 (Kuang et al., 2007a) for the purpose of an HDR image rendering through the integration of bilateral filtering and photoreceptor response functions to reproduce more accurate and pleasing tone-mapped LDR images.

5.3 Proposed SiCAM model for HDR hyperspectral radiances

The latest version and widely used image color appearance model is iCAM06 (Kuang et al., 2007a) which is specifically designed for rendering High-Dynamic-Range (HDR) images. This model is built upon the iCAM framework (Fairchild and Johnson, 2002) and integrates spatial processing models found in the human visual system as discussed above. It incorporates contrast enhancement techniques, photoreceptor light adaptation functions that enhance local details in highlights and shadows, as well as functions that predict various color appearance phenomena.

Through extensive evaluation, iCAM06 has consistently demonstrated excellent performance in HDR rendering, both in terms of preference and accuracy. As a result, iCAM06 is a promising choice as a general-purpose tone mapping operator. Moreover, its potential applications extend to a wide range of image appearance research and practical use. Due to the above mention reasons, we have extended this model for tone mapping of hyperspectral HDR radiances for generating LDR tone-mapped images and proposed a new model namely the spectral image color appearance model (SiCAM). We kept the working principles of iCAM06 (Kuang et al., 2007b) intact and introduce spectral radiances so the appearance pipeline can benefit from the additional information in spectral radiances in comparison to a three-channel image. Additionally, we have proposed hyperspectral adaptation in comparison to chromatic adaptation. This comes as an extension of applying Von Kries normalization (Chong et al., 2007) in spectral radiances rather than tristimulus values which is a linear Von Kries normalization of hyperspectral bands to the hyperspectral adaptation signal derived from the low-pass adaptation hyperspectral signal at each pixel location. It should be noted that, while a few modules in this model are identical to that in iCAM06 (Kuang et al., 2007a), the SiCAM model is already significantly different since it uses hyperspectral data as input rather than RGB image and chromatic adaptation. The complete workflow of SiCAM is given in Figure 5.1 while the details of each processing stage involved in this model are discussed in the following subsections:

5.3.1 Bilateral filtering for base layer

The initial processing stage in the proposed model is Bilateral filtering. The term filtering in image processing usually implies a function that is dependent on the values of the input image in a small window of the pixel under consideration. The widely used filtering operation is averaging among which Gaussian-weighted averaging is one of the most common approaches. This is a function that estimates the average value for the center pixel by averaging the pixels in the neighborhood such that the weights are assigned based on their distance from the central pixel where the weights fall off as we move farther from the central pixel. This low-pass filtering is based on the assumption that the pixels which are spatially closer together are similar to each other. The averaging operation suppresses the noise in the neighborhood and preserves the signal.

This low-pass filtering breaks down on edges where it also blurs out the edges in an image. Many approaches have been proposed for operations intended to preserve edges (Boult, 1993; Chin and Yeh, 1983; Graham, 1962; Himayat and Kassam, 1993; Huang et al., 1979; Lee, 1980; Nagao, 1979; Narendra, 1981; Edition et al., 2002; Perona and Malik, 1990; Ramponi, 1995; Sapiro and Ringach, 1996; Wang et al., 1981; Yin et al., 1996). One method to perform edge-preserving smoothing is anisotropic filtering, which calculates the local image variations at every pixel (Perona and Malik, 1990; Sapiro and Ringach, 1996). It involves solving partial differential equations in an iterative manner which makes it unsuitable for larger data input or time-critical applications while it also poses issues of stability of solutions. A bilateral filter is a non-iterative edge-preserving filter proposed by Tomasi and Manduchi (1998) and then further developed through linear approximations by Durand and Dorsey (2002a). The bilateral filter works both in the spatial and intensity domain. The filter operation is given in equation 5.2.

$$J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s) g\left(I_p - I_S\right) \quad I_p \tag{5.2}$$

$$k(s) = \sum_{p \in \Omega} f(p-s) g\left(I_p - I_S\right) \tag{5.3}$$

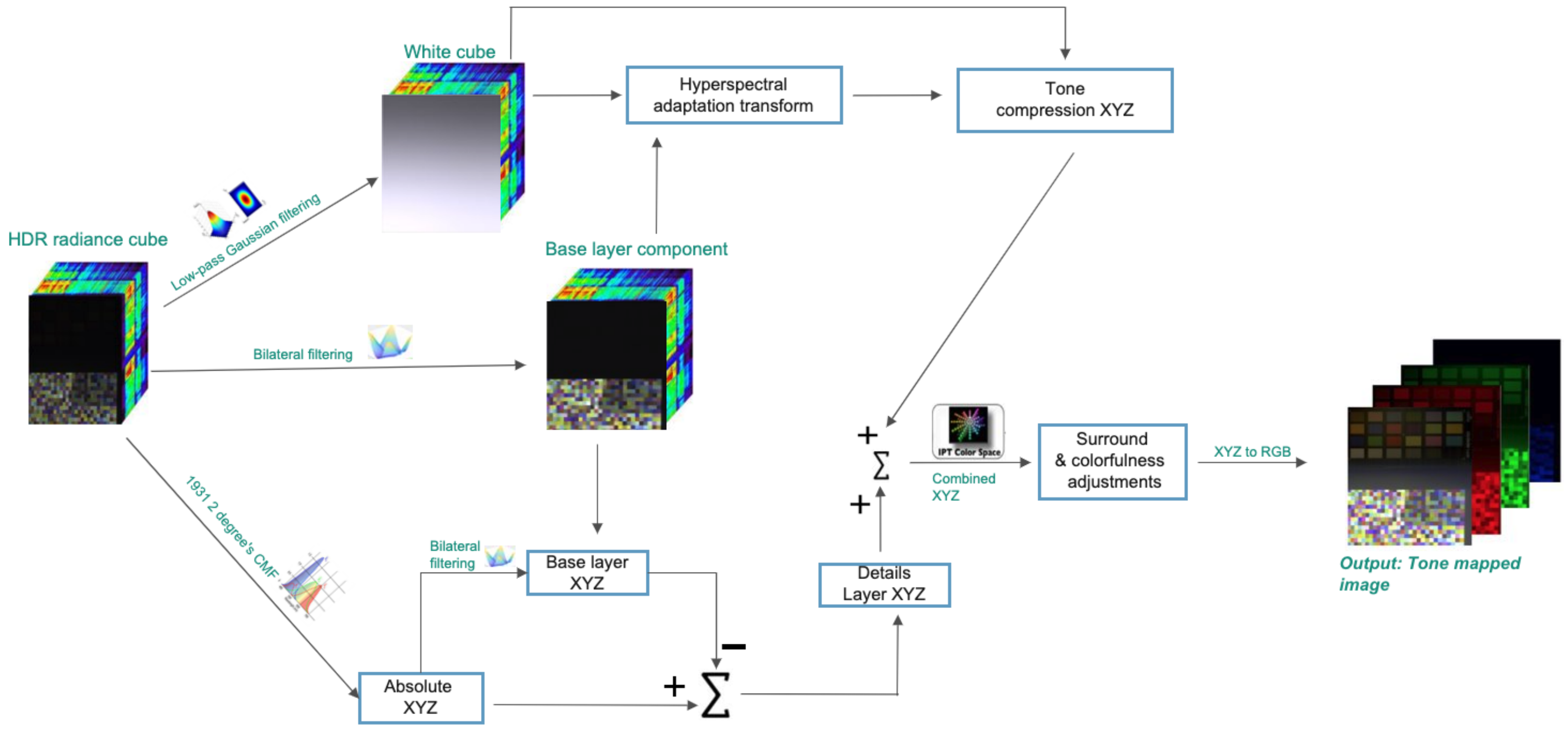


Figure 5.1: *The diagram illustrates the workflow of the proposed SiCAM where input is HDR hyperspectral radiances cube and it generates tone-mapped LDR iamge as output.*

where J_s is the output pixel value at location s, f is the function that operates on spatial domain, g is the function that works on intensity values, Ω is the window around the central pixel, and k is the normalizing function which keeps the sum of the weights equal to 1.

Both f and g are defined as Gaussian functions. In Eq.5.2, notice what would happen on an edge. Consider the center pixel to be on the lighter side. For the weights of the lighter side of the window, $I_p - I_S$ would be closer to 0, and $g\left(I_p - I_S\right)$ would be closer to 1. Now consider the center pixel to be on the darker side. For the weights for the darker side of the window, $I_p - I_S$ would be much greater, and $g\left(I_p - I_S\right)$ would be closer to 0. So for the pixels on the lighter side of the edge, the filter gives more weightage to them and very low weightage to the darker side of the window. Hence the edge is preserved. The same reasoning can be carried out for the central pixels to be on the darker side of the edge.

Although the bilateral filters are discussed above as an operator with an intention to suppress the details through its low-pass filtering nature, in iCAM06 its usage is also extended to divide the image into a base layer and detail layer. The base layer is calculated by applying the bilateral-filter channel-wise on the image. After converting the base layer and original image into a logarithmic domain, their difference is taken and converted back to the linear domain to extract the detail layer which corresponds to the fine details in an image. The motivation for the base layer and detail layer comes from the decomposition of the image into its reflectance and illumination. It is also inspired by HVS in the sense that our vision responds strongly to local contrast (detail layer) in comparison to global contrast (base layer).

In SiCAM, the channel-wise application of bilateral filter on XYZ image, as in iCAM06 (Kuang et al., 2007a), is extended to be applied channel-wise on spectral radiance cube as shown in Fig.5.1. We argue for this extension based on our reasoning of how the bilateral filter and contrast sensitivity function (CSF) works. A bilateral filter works similarly as regular Gaussian filtering on surfaces without edges ($g\left(I_p - I_S\right)$ is nearly 1) and for the pixel values near edges, it also takes into account the intensity information of the surrounding pixels as well. We divide our explanation into two parts as follows.

1. Surface without edges: The tristimulus values are the linear combinations of the spectral radiances with CMFs as shown in Eq.5.4.

$$X = r(\lambda_1)\bar{x}(\lambda_1) + r(\lambda_2)\bar{x}(\lambda_2) + \ldots + r(\lambda_n)\bar{x}(\lambda_n) \tag{5.4}$$

where X is one of the tristimulus values, r is the spectral radiance, $\bar{x}$ is the CMF, and λ_n corresponds to the wavelengths withing visible range of the spectrum.

In the region where there are no edges, the bilateral filter approximates a regular Gaussian filter i.e., the weights for the averages are completely defined by the spatial location of the pixels in the window and not their intensity values. Let w, be the weight that is being multiplied by any pixel in the window, then it can be written as:

$$kX = kr(\lambda_1)\bar{x}(\lambda_1) + kr(\lambda_2)\bar{x}(\lambda_2) + \ldots + kr(\lambda_n)\bar{x}(\lambda_n) \tag{5.5}$$

We can see from equation 5.5, that this is equivalent to if we would have calculated the X value by giving the radiances, that made X, the same weight. In the case of a spectral radiance cube, we do know the spectral formation of the radiances, which would be used to calculate X. Hence the same weights can be applied to the spectral radiances.

2. Surface with edges: For the pixels which are on the edges, the weights are calculated by considering both the spatial location of the pixels in the window and also their intensity values. The reasoning from the previous part can not be applicable here, because the weights calculated for X would be different than the weights calculated for individual spectral radiance values. This boils down to our understanding of how luminance contrast sensitivity works for each monochromatic wavelength. We should be more strongly preserving the edges for the wavelengths for which HVS is more sensitive. Such data for HVS sensitivity to wavelength-specific contrast is not available to the best of our knowledge. Hence, we apply the same bilateral filter for all the channels. Although it is a rudimentary approach but it gives satisfactory final results.

After applying the bilateral filter on the spectral radiance cube, we convert both the original spectral cube and the bilateral filter applied cube into XYZ tristimulus values. We extract the details layer by subtracting the base-layer XYZ from the original XYZ in the logarithmic domain. Note that the spectral radiance cube is still retained to be processed in the next stage.

5.3.2 Hyperspectral adaptation transform (HAT)

Chromatic adaptation refers to the ability of the HVS to resist the change in the appearance of objects due to the change in illumination. In other words, the magnitude of change in the appearance of objects is far lesser than the magnitude of the change in its radiometry or colorimetry. This phenomenon is often referred to as color constancy. The term color constancy implies a bit of exaggeration because if constancy had worked perfectly in HVS then colorimetry would reduce to only calculating matrix multiplication between reflectance and CMFs without the need of knowing illumination SPD (Fairchild, 2007b). Very good approximations of chromatic adaptation are achieved with the seemingly rudimentary von Kries scaling. According to this scaling, the response to the neutral patch remains constant in any illumination post-adaptation.

This can be mathematically written as:

$$\begin{bmatrix} L_a \\ M_a \\ S_a \end{bmatrix} = \begin{bmatrix} \frac{1}{L_w} & 0 & 0 \\ 0 & \frac{1}{M_w} & 0 \\ 0 & 0 & \frac{1}{S_w} \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{5.6}$$

where L_a, M_a, and S_a are adapted cone responses, L, M, and S are unadapted cone responses, and L_w, M_w, and S_w are the responses to the scene white. Now in order to find corresponding colors between two different viewing conditions, we can equate the adapted responses of both stimuli to get L_d, M_d, S_d.

$$\begin{bmatrix} \frac{1}{L_{dw}} & 0 & 0 \\ 0 & \frac{1}{M_{dw}} & 0 \\ 0 & 0 & 1/\frac{1}{S_{dw}} \end{bmatrix} \begin{bmatrix} L_d \\ M_d \\ S_d \end{bmatrix} = \begin{bmatrix} \frac{1}{L_{sw}} & 0 & 0 \\ 0 & \frac{1}{M_{sw}} & 0 \\ 0 & 0 & \frac{1}{S_{sw}} \end{bmatrix} \begin{bmatrix} L_s \\ M_s \\ S_s \end{bmatrix} \tag{5.7}$$

$$\begin{bmatrix} L_d \\ M_d \\ S_d \end{bmatrix} = \begin{bmatrix} \frac{L_{dw}}{L_{sw}} & 0 & 0 \\ 0 & \frac{M_{dw}}{M_{sw}} & 0 \\ 0 & 0 & \frac{S_{dw}}{S_{sw}} \end{bmatrix} \begin{bmatrix} L_s \\ M_s \\ S_s \end{bmatrix} \tag{5.8}$$

where L_d, M_d, and S_d are the responses for the corresponding color in destination viewing condition, L_s, M_s, and S_s are the responses for the source viewing condition, L_{dw}, M_{dw}, and S_{dw} are destination scene white, and L_{sw}, M_{sw}, and S_{sw} are source scene white.

As can be inferred from the above discussion on von Kries scaling, the phenomenon of chromatic adaptation is heavily dependent on the trichromatic working of HVS. This approach is widely studied and is also based on well-studied physiological data (Fairchild, 2007b; Khan et al., 2017). It has also been proven effective across different use cases of color appearance phenomena. Over the past few years, there have been very few studies that attempted to explore alternative approaches for adaptation based on the utilization of spectral information(Fairchild, 2007b; Khan et al., 2017; Burns, 2019; Derhak et al., 2020).

(Fairchild, 2007b) has also described several possible physiological reasoning for how the HVS can actually be working on more bands than just trichromatic vision. In multispectral imaging systems, it is a common methodology to include an additional filter on top of the RGB bayer-patterned sensor array, giving rise to a six-channel multispectral imaging system (one capture with filter and one without filter). We can find a correlate of this principle in HVS too. The macula is a yellow-pigmented area from which light passes through for foveal vision which is not present outside the foveal region. This implies that HVS might be able to get multispectral information processing the data from both the foveal and outside the foveal region. Hence, it can be assumed to extract six channels. Also for describing a color fully, we need five dimensions of data i.e., brightness, colorfulness, lightness, chroma, and hue. This also hints at more then three dimensions of signals being processed in HVS. Liu and Fairchild (2006) also showed the need for two CMFs for color matching between the display and surround. Additionally, the existence of both 2° and 10° observers with different CMFs which are not linear transforms of each other also hints at a multi-spectral approach in HVS.

We define the proposed hyperspectral adaptation transform (HAT) as follows:

$$T = C^T(DK_{diag}B + (1-D)B) \tag{5.9}$$

where T is a 3x1 matrix of tristimulus values, C is Mx3 CIE 2° CMF matrix, and B is Mx1 matrix of the spectral radiance of single spatial pixel location, M is the total number of channels in base radiance cube, and D is the partial adaptation factor calculated similarly as in iCAM06 (Kuang et al., 2007a). We define K_{diag} as follows:

$$K_{diag} = \begin{bmatrix} \frac{D65(\lambda_1)}{W(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \frac{D65(\lambda_2)}{W(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{D65(\lambda_m)}{W(\lambda_m)} \end{bmatrix} \tag{5.10}$$

where the W is the white spectral cube taken as an approximation of source white point (discussed below) and $D65$ is the CIE D65 illuminant spectrum. This is similar to von Kries-type scaling applied to spectral radiances rather than tristimulus values. By comparing the approach of Fairchild (2007b), we see that the difference mainly lies in the estimation of source white, where the source illuminant was already known to Fairchild (2007b), whereas in our work we estimate the white cube as shown in Figure 5.1. Additionally, Fairchild (2007b) used uniform colored patches with uniform illumination whereas we are working with complex HDR scene with non-uniform illumination.

White cube: We calculate the white cube by applying a Gaussian filter on the original spectral radiance cube as illustrated in Figure 5.1. This white cube serves as a source white point spectrum for each spatial pixel location. This is similar to the working of iCAM06 extended to spectral radiances. The reasoning for the plausibility of the extension of the Gaussian filter from 3-channels to n-channels is similar to as discussed in the Section 5.3.1. The output of HAT is spectral-adapted tristimulus values to CIE D65 as the later stages of SiCAM operate in IPT uniform opponent color space as depicted in Figure 5.1. The rest of the working modules in SiCAM are almost similar to the iCAM06 (Kuang et al., 2007a) as now we have converted all the spectral information to tristimulus values which are briefly discussed in the following subsections.

5.3.3 Post adaption and tone compression:

Using the CIECAM02 formula given by equation 5.11, the chromatically adapted RGB responses are first converted from the CAT02 space to the Hunt-Pointer-Estevez fundamentals. Even though the CIE XYZ functions are close to a linear transform of the LMS cone response, they are not exactly related, and hence a better approximation is needed. Therefore, a widely used alternative is the Hunt-Pointer-Estevez cone fundamentals which are utilized in color appearance models. This model has been adopted as the cone response prediction model because it has been extensively studied and proven to accurately predict all available visual data.

$$\begin{gathered}\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_{HPE} \begin{bmatrix} X_{HAT} \\ Y_{HAT} \\ Z_{HAT} \end{bmatrix}, \\ M_{HPE} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}\end{gathered} \tag{5.11}$$

The nonlinear tone compression functions used in SiCAM are similar to that of iCAM06 (Kuang et al., 2007a) as given in Equation 5.12 which are initially adopted from CIECAM02; however, the power value has been modified to an adjustable variable p as illustrated in Figure /reftone. This variable p determines the steepness of the tone compression curves whose value can be set between 0.6 and 0.85 while a larger value results in an output image with greater overall contrast. A pilot parameter setting experiment yielded an empirical value of 0.75 as a default value.

$$\begin{aligned} R'_a &= \frac{400\left(F_L R'/Y_W\right)^p}{27.13 + \left(F_L R'/Y_W\right)^p} + 0.1 \\ G'_a &= \frac{400\left(F_L G'/Y_W\right)^p}{27.13 + \left(F_L G'/Y_W\right)^p} + 0.1 \\ B'_a &= \frac{400\left(F_L B'/Y_W\right)^p}{27.13 + \left(F_L B'/Y_W\right)^p} + 0.1 \\ F_L &= 0.2k^4\left(5L_A\right) + 0.1\left(1 - k^4\right)^2\left(5L_A\right)^{1/3} \\ k &= 1/\left(5L_A + 1\right)) \end{aligned} \tag{5.12}$$

In CIECAM02 and earlier models, the FL function is used to predict a variety of luminance-dependent appearance effects. Nevertheless, the computation of the FL factor in iCAM06 differs significantly from that of previous color appearance models, as it is derived from the low-pass adaptation image at each pixel location and Yw represents the luminance of the locally adapted white image by applying a filter of one-third the size of the image which is determined empirically. While the

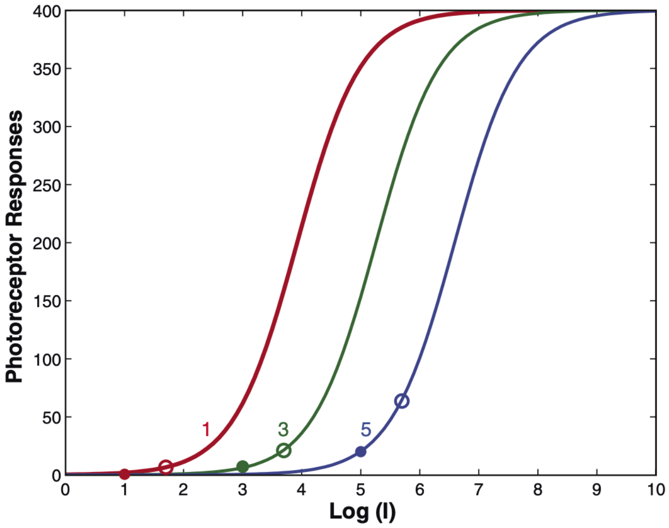


Figure 5.2: *The diagram illustrates the tone compression curves as the response of cones to adaptation plotted against log intensity (log cd/m2) for three levels of adaptation in iCAM06 where open circles represent reference whites while filled circles represent adapting luminances (Kuang et al., 2007b).*

response functions of the rods are adapted from the Hunt Model (Hunt, 2005) and given by equation 5.13

$$\begin{aligned}
&A_s = 3.05B_s\left[\frac{400\left(F_{LS}S/S_w\right)^p}{27.13+\left(F_{LS}S/S_w\right)^p}\right]+0.3\\
&F_{LS} = 3800j^2\left(5L_{AS}/2.26\right)+0.2\left(1-j^2\right)^4\left(5L_{AS}/2.26\right)^{1/6}\\
&L_{AS} = 2.26L_A\\
&j = 0.00001/\left[\left(5L_{AS}/2.26\right)+0.00001\right]\\
&B_S = 0.5/\left\{1+0.3\left[\left(5L_{AS}/2.26\right)\left(S/S_w\right)\right]^{0.3}\right\}\\
&\quad+0.5/\left\{1+5\left[5L_{AS}/2.26\right]\right\}
\end{aligned} \tag{5.13}$$

The final tone compression response is a sum of the cone response and the rod response as given in equation 5.14. Next, the model converts tone-compressed RGB signals back to CIE XYZ images and combines the detail layer. This merged tone-mapped image is then converted into IPT uniform opponent color space and details adjustment is applied to predict the Stevens effect, wherein an increase in luminance level causes an increase in local perceptual contrast which is identical to iCAM06 (Kuang et al., 2007a).

$$RGB_{TC} = RGB'_a + A_S \tag{5.14}$$

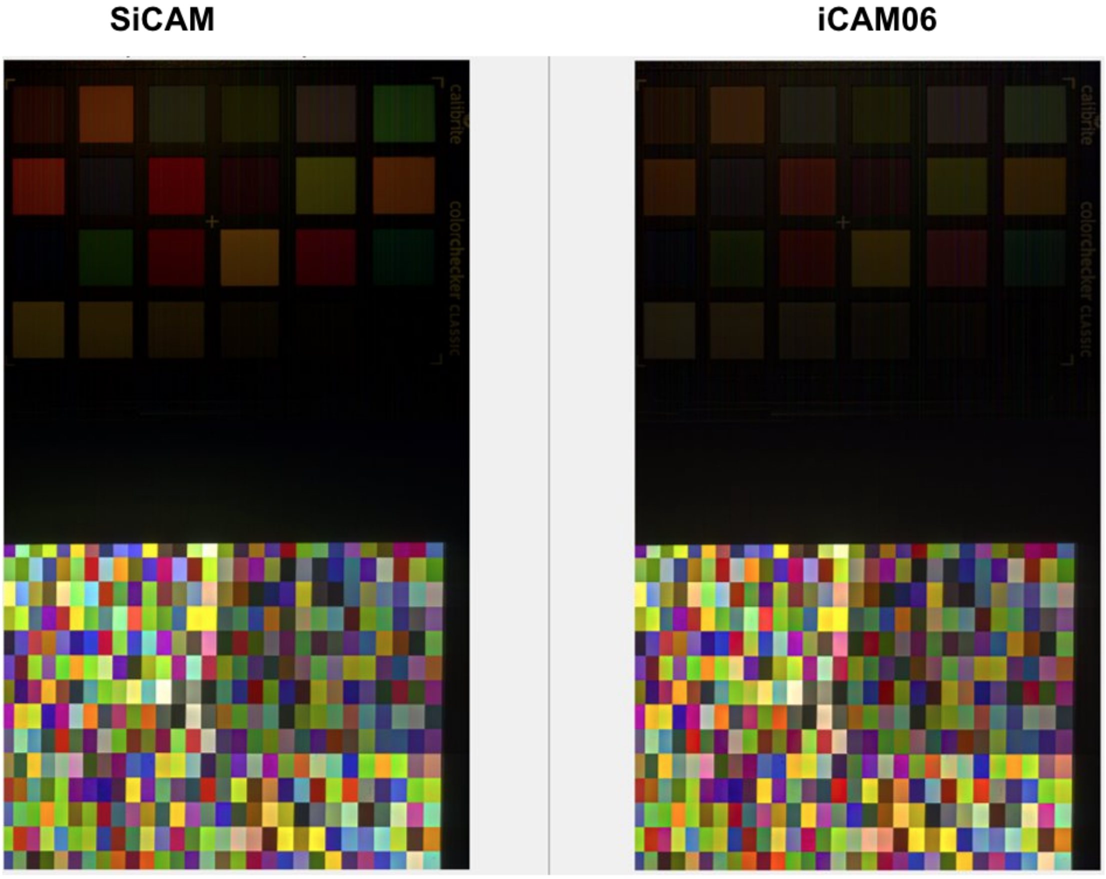


Figure 5.3: *The figure illustrates the outputs generated by SiCAM and iCAM06 (Kuang et al., 2007a).*

5.3.4 SiCAM output:

The final output of both SiCAM and iCAM06 is a three-channel RGB image which is scaled between 0 and 255(i.e., 8 bits encoding). However, SiCAM has been found to generate more tone-accurate reproduction of HDR images as compared to that of iCAM06 (Kuang et al., 2007a) and also sharper LDR images. It can be observed from Figure 5.3 that the output generated by SiCAM using hyperspectral radiances as input resulted in better and more uniform color patches. On the other hand, the iCAM06's output has a slight yellowish color cast and also the patches are non-uniform in terms of color fidelity. Figure 5.4 is the zoomed-in on the patches where the difference clearly visible. We observe that the reflective part in SiCAM is more saturated than the iCAM06. Also, we observe a yellowish white in the transmissive color target(film) of the iCAM06 whereas in SiCAM, the white seems more neutral. Also, we observe that in SiCAM the fixed pattern noise is not visible

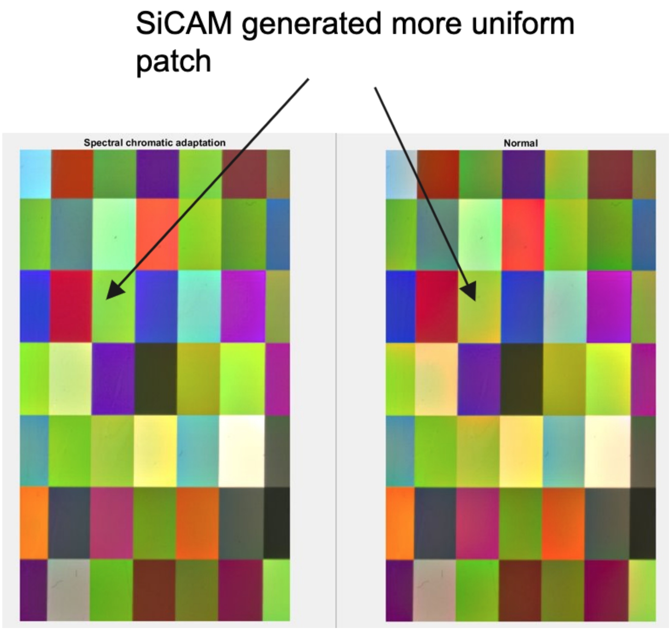


Figure 5.4: *The figure illustrates the zoom-in patches of the outputs generated by SiCAM and iCAM06. It can be observed in these color target patches that the SiCAM's output is more uniform whereas the iCAM06's output has a slight yellowish color cast and also the patches are non-uniform in terms of color fidelity.*

whereas the iCAM06 output has visible noise artifacts.

We have discussed subjective and objective evaluations of SiCAM's outputs in the following chapters, as well as a comparison of those outputs with LDR images created by iCAM06 and the gamma tone mapping operator. The main objective is to assess how the visual quality perception and realism in HDR-rendered images is affected by using HDR hyperspectral data instead of traditional three-channel HDR data.

6 Display calibration and characterization:

In order for the light to shine so brightly, the darkness must be present.

Francis Bacon

To ensure faithful results and to avoid potentially adverse impacts during the perpetual evaluation of the proposed HDR rendering methods, display monitors involved in our psychological experiment have been calibrated and characterized. It has been observed that the color temperature and luminous intensity of the display may alter under continuous use, which can affect the display's performance. Therefore, regular calibration is necessary to restore the display to its original state. Hence, the procedure to achieve this is a two-step process that includes display calibration and characterization. The display is first calibrated to the desired behavior, and it is during the second step that an ICC profile is generated, taking into account the settings that were optimized in the previous step. The fundamental working principle of the display is an additive system, which implies that all the display colors are synthesized using linear combinations of their three RGB primaries. Therefore, it is also important to measure and verify the display model. A display model defines how it generates colors and also reveals the actual color gamut of a display. The display's electro-optical transfer function (OETF) describes the nonlinear relationship between the display's photometric output and digital input while the display's linear 3x3 matrix describes the RGB primary colors. Hence, these two components collectively define a display model.

In this chapter, an overview of color management is discussed to highlight the importance of display profiling, followed by a brief discussion of the necessary concepts involved in display characterization. Later, the method used to characterize both HDR and LDR displays is described to ensure that the display behaves consistently and accurately during the experiment, leading to more reliable results.

6.1 An overview of color management:

Color management is necessary for ensuring consistent color reproduction across various types of devices and multimedia platforms. The advent of the latest color imaging technologies, such as cameras, display monitors, printers, scanners, etc., has made this task more challenging due to the differences in their color reproduction capabilities. The transfer of color information from one device to another necessitate control of how the color is communicated between different devices. For example, when an image is captured with a digital camera, it can be uploaded to a computer for viewing on its screen and then sent to a printer for printing. Managing the translation and communication of color information in such interactions is what we refer to as color management. As long as the device can reproduce the colors, the aim of color management is to ensure that colors appear as close as possible to each other across all of the devices in the color management workflow.

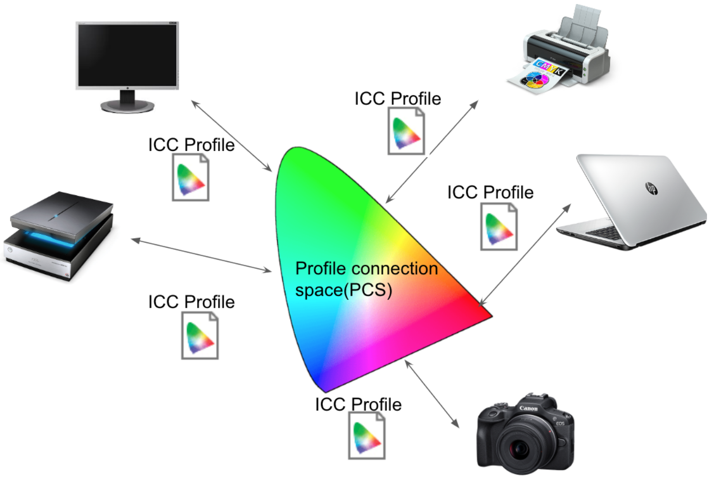


Figure 6.1: *In the color management workflow, individual devices convert their device-dependent color space data to a common device-independent color space called Profile Connection Space (PCS), through which we can efficiently share color content to any other device-dependent space.*

The challenge for color management is that each device has its own color characteristics and therefore, it is not possible to just share the same image across different devices and expect the colors to appear the same throughout. Instead of converting data from between different devices in a one-to-one manner, the solution provided by color management is to convert device-dependent colors to central scale i.e., profile connection space(PCS) using color profiles which is the description of the relationship between device dependent i.e RGB color space to device independent i.e., L*ab or CIE XYZ color space as illustrated in Figure 6.1. The framework for the PCS and the format of such profiles are defined by a regulatory body that governs color management protocols known as the International Color Consortium (ICC).

6.1.1 ICC Profiles

An ICC profile is a data file that describes the color characteristics of an imaging device. The file size usually varies from 4k to 4MB with the extension of *.icc* or *.icm*. There are a total of fourteen profile types provided for in ICC format specification while the most commonly used are input, display, and output profiles. The color transformations defined by ICC profiles are either matrix-based or lookup-table(LUT) based profiles which are discussed below.

6.1.1.1 Matrix/TRC-based profile

A matrix-based profile performs a two-step transformation to convert data from device-dependent color values to device-independent color values. The first is the gamma value to estimate the tone reproduction curve(TRC) which helps to linearize the input data before applying the colorimetric transformation. The second step is the multiplication of a 3x3 transformation matrix that is used to transform the device-dependent color values (e.g., sRGB) to device-independent color values (e.g., CIELAB or CIE XYZ) as shown in Figure 6.2. Before creating a matrix-based profile, it is important to verify that RGB input channels are independent otherwise the LUT-based profile is recommended. To start, a series of XYZ data for different RGB values need to be measured. If the gamma is the same

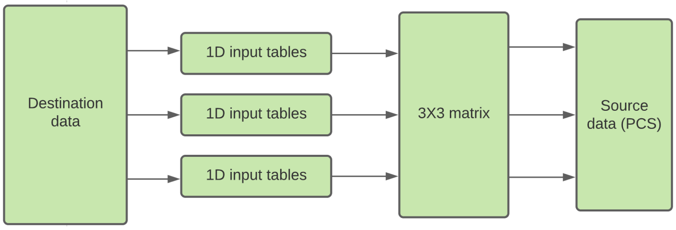


Figure 6.2: *Steps involved in color transformation defined by ICC matrix/TRC-based profile where TRC stands for tone reproduction curve.*

for all the display's primaries, then we can collect neutral RGB color measurements, where the RGB values are the same, at equal step sizes. However, if the value of gamma is not the same for all RGB display primaries, we need to collect equally stepped measurements for individual primaries separately.

Once the data collection is complete, both RGB and XYZ values are converted into the logarithmic domain. The regression is then applied with RGB as the input and Y (luminance) value as the output. By using the logarithmic domain, we calculate the slope of the fitted line which is equivalent to the gamma where the value of gamma is usually 2.2 or 2.4.

Before performing the regression, it is essential to consider the ICC PCS (Profile Connection Space) space, where the illuminant is assumed to be CIE D50. As a result, the XYZ data must be chromatically adapted to CIE D50 illuminant. Additionally, the chromatic adaptation matrix should be included in the ICC profile under the **chad** (chromatic adaptation) tag (see section 6.1.2 for details of profile tags). This inclusion allows the profile to interpret the original white point of the display.

6.1.1.2 CLUT based profile

The colour look-up-table (CLUT) based profiles, allow color management to define a more complex relationship between the different color spaces. Also, note that the LUT-based profiles overcome the limitation of mapping the color spaces with a simple gamma relationship. However, the CLUT does not contain all the different combinations of input domains. This implies that interpolation needs to be done for the intermediate input data values. CLUT can be defined in two ways i.e., input as device RGB (in the case of displays) and output as PCS Lab and vice versa, as illustrated in Figure 6.3.

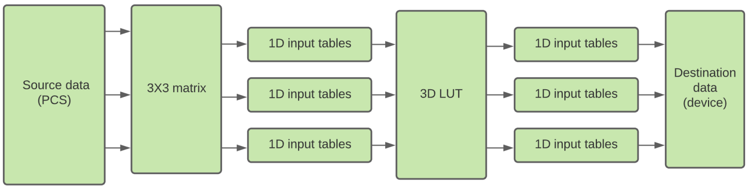


Figure 6.3: *Steps involved in color transformation defined by ICC color lookup table based profile*

6.1.2 Display Profiling

The display profiles are mostly matrix-based and they require only a small number of measurements to determine the phosphor chromaticities and the display intensities. Measurements of color patches on the display screen can be made using any radiometric instrument as discussed in section 6.2 of this chapter. It is essential to determine which measuring instrument would be suitable as a single error in several hundred measurements can seriously weaken the quality of resulting conversions. The selection of the instrument depends on many factors among which spectral responsitivities and integrating time are of great importance. ICC profiles have two main parts i.e., a header that contains a fixed number of items and tags that vary depending on the type of profile. The header reveals if a profile is made in accordance with version 2 or version 4 of the ICC specifications (Green and MacDonald, 2011). Another important part of the header is the class field. This field indicates what sorts of tags to expect in the body of the profile and that is why most processors will look first at the class field. There are seven profile classes: display(mntr), input(scnr), output(prtr), device link(link), colorspace(spac), abstract(abst), and name color (nmcl). There are also three categories of tags in ICC profiles namely basic tags, optional tags, and private tags. The basic tags for the display profile are the Profile Description, Media White Profile tag, Copyright tag, and Chromatic Adaptation tag as shown in Table 6.1.

Table 6.1: *The list of basic tags for display profile*

Tag	**Tag name**	**General Description**
desc	Profile Description	Versions of the profile name for display
wtpt	Media White point	Media XYZ white point
cprt	Copyright	Profile copyright information
chad	Chromatic Adaptation	Method for converting a color from another illuminant to D50

6.1.2.1 Creating Matrix/TRC based Display Profile

The display profiles are generally matrix/TRC-based because they generally fulfill the assumption of additivity where the RGB input channels are independent. One of the scanity check for channel independency to verify the additivity of individual input primaries. For instance, using a spectroradiometer measure the CIE XYZ values of the pure Red channel (255, 0, 0), pure blue channel (0, 255, 0), and pure green channel (0, 0, 255) and add these values together. The sum of these individual

RGB primaries at their full intensities should be approximately equal to the CIE XYZ of white patch measurement(white patch is equal to R=255, G=255, and B= 255 if the input data is 8-bit encoded). Next, the XYZ values of RGB primaries, white point, black point, and a few intermediate ramp colors are determined using radiometric measurement instruments. Then, scale all measured data to the point where the Y of the white point equals 100 and apply chromatic adaptation to the measurement data using the XYZ of the scaled white point as the source illuminant. One way to achieve this is to use ICC $bradford$ function of the MATLAB color engineering toolbox (Green and MacDonald, 2011). The MATLAB command can be used as follows: $data = bradford(xyzs, xyzw, d(50))$ where $xyzs$ is the array of scaled measurement data and $xyzw$ is the scaled CIE XYZ white point. The white point should now correlate to D50's XYZ coordinate system. Divide this value by 100 to ensure that it is within the acceptable range for the ICC profile and use linear regression on this data to estimate the value of gamma. Open an XML file containing a Matrix/TRC profile (for example, MatTRC.icc/MatTRC.icm) or create one using IccToXml.exe using the DemoIccMAX software (Green and MacDonald, 2011). Fill in the appropriate fields in the XML file as given in Table 6.1. For ICC display profiles, the media white point ($wtpt$) is set to the D50 XYZ values. Then utilize IccFromXml.exe to verify all entries, update the header and description sections, and write the profile. To verify the contents of created ICC display profile, the DemoIccMax validation routine (Green and MacDonald, 2011) can be used to test for conformance with the current ICC specification as defined by ICC.

6.1.2.2 Creating CLUT-based Display Profile

Like TRC- based profile, the XYZ values of display primaries, white point, black point, and some intermediate ramp colors are required to measure. Then, this data is scaled to the media relative using the equation 6.1:

$$X_{\mathrm{X_{Mrel}}} = X_{\mathrm{D50}}/X_{\mathrm{W}} \tag{6.1}$$

where X_{Mrel}= is the media-relative X, $XD50$ is the X value of $D50$, and X_W is the X of the media white point. Note this means that the scaling ratios are different for X, Y, and Z, unlike the normalization to $Y = 100$ used for the display colorimetry. To generate a transformation model for the relationship of RGB to XYZ, the $charac3$ function of the MATLAB color engineering toolbox(Green and MacDonald, 2011) can be utilized e.g., $abc = charac3(rgb, xyz)$. Then, generate a uniformly sampled input table using the Matlab command: $lut = create3dlut(n)$ where n is the number of nodes you want in each dimension of the table (n is commonly 9, 17, or 33). Now multiply this by the data encoding maximum

(a) *Densitometer* **(b)** *Colorimeter*

(c) *Spectrophotometer*

Figure 6.4: *The color measuring instruments in color management.*

e.g., $LUT = lut * 255$ for eight-bit data. Convert the LUT to XYZ using the following Matlab command: $XYZ = polyconvert3(LUT, abc)$. Convert this CIE XYZ data to CIELAB. Use encoding code to convert the CIELab to uint8 and then scale it to uint16. Open an XML file for a LUT-based profile or make one from a profile using $IccToXml.exe$, ensuring that the data color space (RGB, CMY etc) matches the data you have. Write the PCSLUT values into the $AToB1 < CLUT >< TableData >$ field. Set the $< GridPoints >$ entry to match the number of nodes chosen. Update the header and description fields and write the profile using $IccFromXml.exe$. Finally, DemoIccMax validation routine (Green and MacDonald, 2011) can be used to verify the profile's conformity.

6.2 Measuring Instruments

The entire process of color management relies on measuring instruments because all calibration and characterization techniques involve taking measurements from the display screen. There are three main categories of color measuring instruments as shown in Figure 6.4 and details about their measurements are given in Table 6.2 which are further discussed in the following subsections:

6.2.1 Densitometers

The densitometers basically work on the principle of photometers that measures the total amount of light reflected from the surface of an object to the incident light falling on the surface of an object from the instrument's lamp. The measured reflectance is then converted to an optical density using $D = -logR$, where D is the density and R represents the reflectance. The density reveals the depth of tone and can also be used to evaluate the thickness of the printing inks. However, it does not provide information about chromaticity and does not relate well to human perception. Therefore their usage is not recommended for taking measurements to perform color management-related tasks. The densitometer is usually preferred to measure the density of photographic materials and printing inks.

6.2.2 Colorimeters

The colorimeter tries to meet Luther's condition by simulating the sensitivities of LMS cones using three or four filters. As the colorimeter has a response approximately equal to CIE standard observer, this can directly measure CIE XYZ values. From XYZ, L*AB and Yxy can easily be determined. The limitation of colorimeters is that they cannot record spectral data. In color management, these devices can be used to make CRT and LCD display profiles because they are portable and inexpensive devices.

6.2.3 Spectrophotometers

Due to the versatility of spectrophotometers, they are becoming the default instrument for color measurement. The spectrophotometer measures the whole spectrum of the sample by splitting the incident light into different wavelengths of light using a diffraction grating. They can also be used as a densitometer and colorimeter. For example, they can be used to measure CIE XYZ and can also be used to measure optical density.

Table 6.2: *Color measuring instruments and what data they can measure.*

	Density	Tristimulus values	Spectral Data
Densitometer	Yes	No	No
Colorimeter	No	Yes	No
Spectrophotometer	Yes	Yes	Yes

Other types of color measuring devices are:**spectrocolorimeter** which only reports the output as CIE XYZ values whereas **telespectroradiometer** which records the ambient light falling on the sample and measures the color at a distance from it. Telespectroradiometer records the full spectrum of incoming light which is usually in the range of 380nm to 780nm.

6.3 Display Calibration

Calibration refers to determining the properties of a display monitor to place them in a predetermined state that can be standardized and repeatable. Display calibration is necessary to be performed on a regular basis since display behavior varies considerably over time. It allows similar display devices at different locations to operate in a consistent and similar way. A good practice is to warm up the display to its normal operating temperature and conditions for at least 30 minutes before calibration and set its resolution to the default screen setting. To ensure that the color reproduction device which is a monitor display in our case operates under optimal conditions, we need to adjust four important parameters of the monitor display which are listed below:

1. **Primary colors:** The types of primary colors (RGB) determine the color gamut of the device system. Display primaries determine every color displayed on a screen that is essentially the additive mixture of these three primary colors in varying intensities. In the realm of display technology, several additive color spaces based on the RGB color model are commonly employed. The two most prevalent of these are the sRGB and Adobe RGB color spaces.

2. **Brightness:** Brightness is the preserved luminance by the human visual system. Display brightness is the intensity of light that is emitted by a monitor which is measured and expressed in Candela per square meters (cd/m2). For example, 160 cd/m2 is a common brightness parameter within the graphic art industry (see ISO 12646 Displays for colour proofing — Characteristics and viewing conditions (ISO, 2023)).

3. **White Point:** The white point is the color of the light emitted by the monitor and measured in Kelvin. The white point levels can be configured by increasing or decreasing the Kelvin degree setting, which corresponds to the screen's color temperature. For example, 5000 Kelvin is a common color temperature parameter within the graphic art industry (ISO, 2023) to use as the white point.

4. **Gamma:** Gamma describes the nonlinear relationship between the pixel levels in the computer and the corresponding luminance of the monitor (the light energy it emits). Gamma is expressed as (Luminance)=Constant x input gamma. While gamma has no effect on black or white, it does affect grey or mid-tones. If a monitor's gamma level is set to high, mid-tones will appear too dark. Likewise, for low gamma values, mid-tones will appear brighter. A gamma of 2.2 is widely used within the graphic art industry (ISO, 2023). The chart given in Figure 6.5 illustrates the relationship for gamma = 1, 1.5, 1.8 and 2.2.

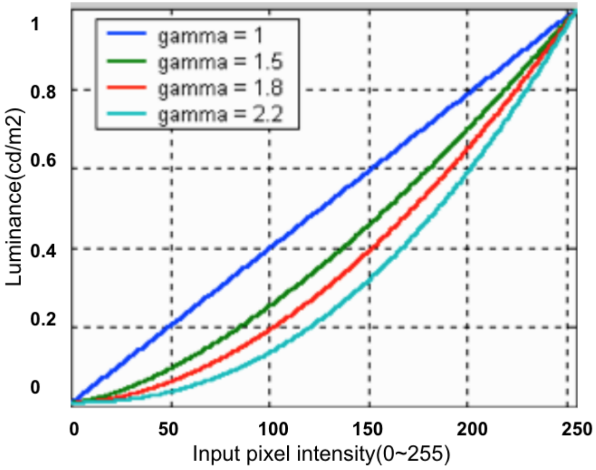


Figure 6.5: *The relationship for different values of gamma = 1, 1.5, 1.8, and 2.2 (Image credits: Lecture slides of cross-media reproduction, NTNU).*

Recommended Monitor calibration according to ISO standard 12646 (2015) are as follows: Primary colors: Adobe RGB, Brightness: 160cd/m2, White Point: 5000 Kelvin and Gamma: 2.2

6.4 Display Characterization:

Calibration means adjusting a display to operate in a predefined working state which is standardized and repeatable. In other words, calibration means adjusting a display's behavior to match a standard such as sRGB. In this case, the display would correctly display an image encoded in sRGB. Characterization means building a model that describes actual display working behavior, thus colorimetric characterization describes how a display is producing the colors on the screen which also reveals the actual gamut of the display as shown in Figure 6.6.

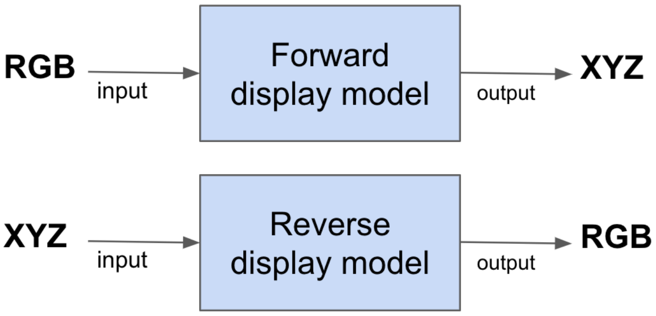


Figure 6.6: *During display characterization we measure and verify the forward and reverse model of the display where the forward model is from display RGB to CIE XYZ while the reverse model is from CIE XYZ to display RGB.*

The display forward model tells the relationship between the input RGB values and the output XYZ tristimulus while the display reverse model is responsible for calculating the display's RGB to accurately generate a desired XYZ output. Characterization means building these display models and performing their verification. There are two steps involved in characterizing a display which are as illustrated in Figure 6.7 i.e., 1) calculating the colorimetric transformation matrix, and 2) the display's Electro-Optical Transfer function (OETF).

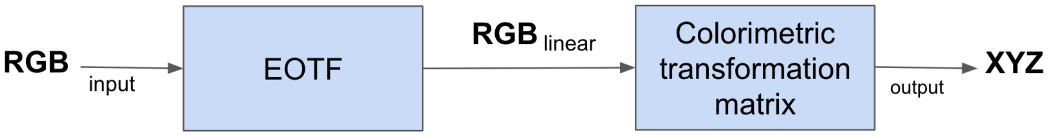


Figure 6.7: *The display model involves the display's Electro-Optical Transfer function (OETF) and the colorimetric transformation matrix.*

6.4.1 Determining the EOFT of display primaries

The Electro-Optical Transfer function (E0TF) defines how the optical or light output relates to the electric or digital input which is a nonlinear relationship and can be approximated by GOG model(Berns, 1996). The EOTF is based on four main assumptions which are as follows:

- **Display primaries**: Three RGB primaries are used to generate the display's colors by their linear combinations.

- **Channel Independence**: The output of one channel does not influence the output of another channel.
- **Separability of nonlinearity**: The chromaticity of the primaries does not change with intensity.
- **Constant offset**: Displays have a black level that is typically not equal to zero due to flare or optical leakage which can be modeled by a constant offset.

This nonlinear EOTF function can be described as a power function that is the input values are nonlinearly related to the physical luminance of the light output as illustrated in figure 6.5 which has also been referred to as a tone reproduction curve (TRC) in ICC color management. Another better approach is to store this relationship in a separate look-up table (LUT) for each channel. The input RGB uses are usually gamma-encoded so we use this function to decode the linear RGB values and then apply a linear transformation matrix to get display independent XYZ values.

6.4.2 Determining the colorimetry of display primaries

To calculate the colorimetric transformation matrix, it is sufficient to have knowledge of display primaries i.e., the CIE XYZ coordinates for a display's maximum red, green, and blue colors. This 3x3 linear matrix is constructed by concatenating the XYZ values of the display primaries as columns as given by equation 6.2 (Green and MacDonald, 2011).

$$\begin{bmatrix} X_{\text{pixel}} \\ Y_{\text{pixel}} \\ Z_{\text{pixel}} \end{bmatrix} = \begin{bmatrix} X_{\text{r,max}} & X_{\text{g,max}} & X_{\text{b,max}} \\ Y_{\text{r,max}} & Y_{\text{g,max}} & Y_{\text{b,max}} \\ Z_{\text{r,max}} & Z_{\text{g,max}} & Z_{\text{b,max}} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.2}$$

Thus, the RGB triplet [1 0 0] corresponds to the XYZ triplet for the red primary, and similarly for the green and blue primaries. For instance, the primary colors of any HDTV display should closely match ITU-R recommendation BT.709 (BT, 2002), and the chroma coordinates are given in Table 6.3 where these chromaticity coordinates can be converted into CIE XYZ values using the formulas given by equation 6.3. Using these CIE XYZ we can obtain the linear transformation matrix for BT.709 as given by the equation 6.4

$$X = \frac{x}{y} \cdot Y, \quad Z = \frac{1 - x - y}{y} \cdot Y \tag{6.3}$$

Table 6.3: *The chromaticity coordinates of ITU-R recommendation BT.709. (BT, 2002)*

Chromaticity	Red	Green	Blue
x	0.6400	0.3000	0.1500
y	0.3300	0.6000	0.0600
Y	0.2126	0.7152	0.0722

$$\begin{bmatrix} X_{\text{red}} & X_{\text{green}} & X_{\text{blue}} \\ Y_{\text{red}} & Y_{\text{green}} & Y_{\text{blue}} \\ Z_{\text{red}} & Z_{\text{green}} & Z_{\text{blue}} \end{bmatrix} = \begin{bmatrix} 0.412424 & 0.357579 & 0.180464 \\ 0.212656 & 0.715158 & 0.072186 \\ 0.019332 & 0.119193 & 0.950444 \end{bmatrix} \tag{6.4}$$

The matrix given by equation 6.4 is a forward transformation i.e., linear RGB pixel values to display XYZ values as shown in 6.2 while the reverse transformation can also be computed by taking the inverse of this matrix. Please note that for further reading on display characterization, refer to the seminal work by Berns (1996) and/or Chapter 6 and Chapter 7 of Color Engineering (Green and MacDonald, 2011).

6.5 Adopted methodology for HDR display characterization

In order to ensure faithful color reproduction and to avoid potentially adverse impacts during the perceptual evaluation of HDR rendering methods both LDR and HDR displays were calibrated and characterized before setting up the experiment. As elaborated earlier, the color characterization of the LDR display was accomplished through ICC profiling. In our psychophysical study, we employed the PVM-X2400 24-inch 4K HDR Sony display with the resolution of 1920-by-1080 pixels, to show an HDR image of the physical HDR scene. The HDR display has been configured to operate in the PQ REC2020 mode using 10-bit pixel encoding. However, for displaying HDR images on this HDR Sony display, we used MATLAB Psychtoolbox-3. It is important to note that MATLAB Psychtoolbox-3 does not support the display ICC profile. Currently, it only supports HDR-10 display mode for rendering HDR images. It requires linear RGB values as input and takes care of encoding these linear RGB pixel values by applying PQ non-linearity curve (i.e., EOTF) itself. Also, note that generally displays sometimes do not strictly follow the targeted color space and deviate from the theoretically perfect system (Berns, 1996). Hence it necessitates estimating the display model by conducting colorimetric characterization of this HDR display. Additionally, in order to convert HDR hyperspectral radiance image to three channel HDR image as illustrated in

Figure 4.17, we need to apply the 3x3 colorimetric transformation metric of this display to ensure accurate color representation instead of using standard REC2020 primaries. This also makes it important to know the actual colorimetry of the display to correctly transform our XYZ tristimulus values to display RGB pixel values.

6.5.1 Experimental setup

The state of the display during measurement would be the state of the display preserved in the display characterization model. Therefore, it is important to calibrate the display to its default setting and also turn off automatic settings such as screen savers and power savers, etc. It is also a good practice to turn on the display for at least 30 minutes warm up and stabilize the display. Due to angular dependence and possible spatial non-uniformity, it is best to measure the display from the observer's perspective viewpoint as illustrated in Figure 6.8.

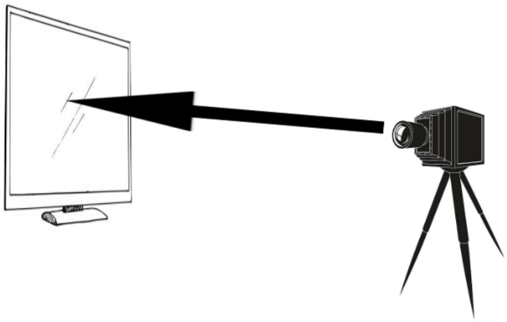


Figure 6.8: *The optimal approach is to take measurements for characterizing a display from the observer's viewing angle due to the angular dependence and possible spatial non-uniformity. Another approach is to take measurements with varying RGB patch sizes.*

For accurate measurements, it is important to avoid any ambient light therefore, the display measurements were taken in the darkroom with black-colored walls and curtains as illustrated in figure 6.9. Considering all the recommendations mentioned above the display was turned on for 30 minutes before taking the measurements and was to its default setting. The Konica Minolta CS2000A telespectroradiometer was utilized for the measurements where the distance between the display screen and the telespectroradiometer was set to 60cm making 2° measuring angle. The room was darkened while taking the measurement.

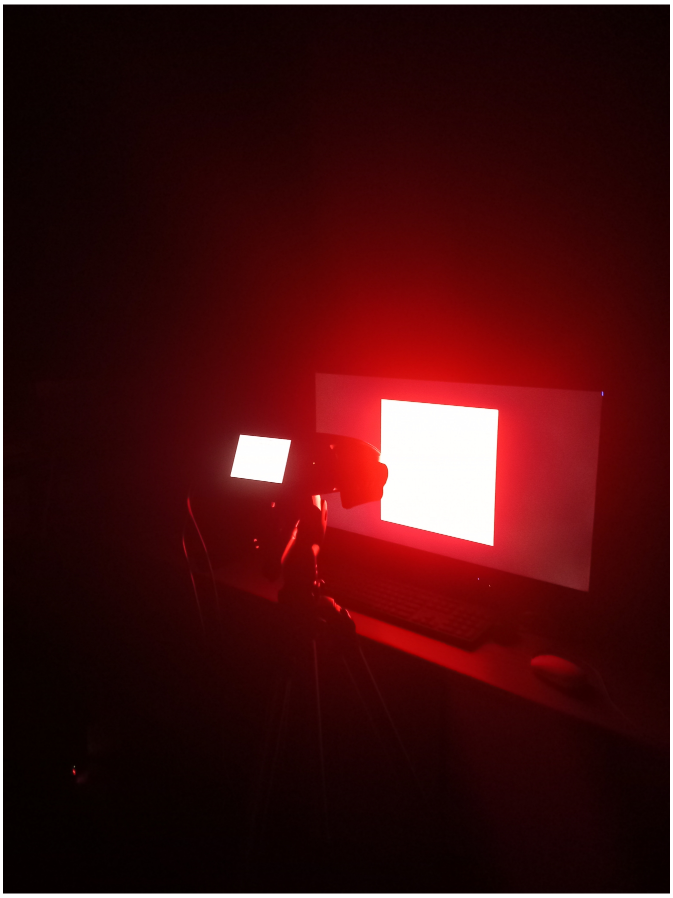


Figure 6.9: *The figure illustrates the setup used for characterizing the HDR display where the Konica Minolta CS2000A tele-spectroradiometer was used to take measurements of different color patches displayed on the HDR screen using Psychtoolbox-3.*

6.5.2 Procedure

The display characterization involved calculating three major components of the display model i.e., the colorimetric transformation matrix M, the XYZ black level or flare values, and the RGB LUTs as illustrated by Figure 6.10. We used MATLAB to generate the RGB patch of 800-by-800 pixels using 10-bit pixel encoding, and Psychtoolbox-3 was used to display those patches.

The colorimetric 3X3 matrix M defines how colors are formed by the display which comes from measurements of CIE XYZ of each of the RGB primaries at full their maximum intensity i.e., (1023, 0, 0), (0, 1023, 0), and (0, 0, 1023) considering 10-bit encoding. The flare vector comes from the measurement of XYZ of black i.e., (0, 0, 0). The LUT comes from measurements of XYZ of a neutral ramp,

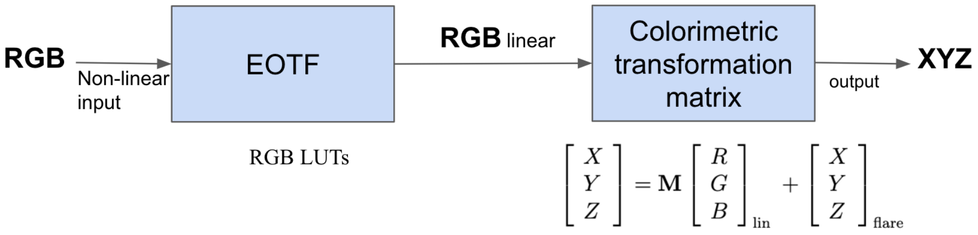


Figure 6.10: *The display characterization involved calculating three major components of the display model namely RGB LUT which defines the EOTF of the display, the 3x3 colorimetric transformation matrix (M), and the black offset/flare as illustrated in this display workflow model.*

where “ramp” simply means a series of increasing or decreasing values whereas neutral ramp means input primaries are all equally incremented (i.e., R=G=B). We created 60 logarithmically spaced 800X800 patches for each RGB primaries and the neural ramp where R=G=B and hence 240 measurements were taken in total. Outliners were first removed from these measurements which exceeded the predefined threshold i.e., their difference from the median chromaticity was greater than 0.2. Then channel independence and spatial independence tests were conducted where if the display exhibits channel independence then the addition of maximum values measured for the individual primaries should equal to the measured peak white. Spatial independence was tested by measuring peak white at varying patch sizes. The errors for both of these tests were below $1\Delta E_{*ab}$ when using the display’s peak white to calculate CIELAB coordinate. Subsequently, the mean average of measurements taken from five black patches (where R=G=B=0) was computed to estimate the flare offset value. It was then subtracted from the RGB patch measurements in order to perform the black correction as shown in equation 6.5

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{r,\max} - X_f & X_{g,\max} - X_f & X_{b,\max} - X_f \\ Y_{r,\max} - Y_f & Y_{g,\max} - Y_f & Y_{b,\max} - Y_f \\ Z_{r,\max} - Z_f & Z_{g,\max} - Z_f & Z_{b,\max} - Z_f \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} X_f \\ Y_f \\ Z_f \end{bmatrix} \tag{6.5}$$

The colorimetric matrix was calculated by concatenating the maximum individual values of RGB primaries into a 3x3 matrix. This matrix was further optimized by employing the Nelder-Mead simplex algorithm (Lagarias et al., 1998) using 1331 training RGB patches where root mean squared error was used as the objective function to minimize the mean CIEDE2000 color difference, calculated by the CIED2000 color difference formula, of test RGB patches sampling the display’s gamut. The final matrix is given in equation 6.6. Finally, the EOTF is simulated by PQ encoding (Nezamabadi et al., 2014; Boitard et al., 2015). As mentioned above the display was set to REC2020 PQ mode and the patches were displayed using Psychtoolbox-3 which assumes a linear RGB input and takes care of PQ encoding itself. These PQ-encoded RGB values are decoded at the display end while rendering them on the HDR screen. Hence, the output is again linearized when displayed on the screen. Therefore the relationship between the input RGB to output CIE XYZ is linear as depicted from the graph 6.11 which is calculated as input neutral ramp (R=G=B) on the x-axis while the output luminance (Y) is given on Y-axis.

$$M_{RGBtoXYZ} = \begin{bmatrix} 0.5818 & 0.1807 & 0.2106 \\ 0.2545 & 0.6671 & 0.0783 \\ -0.0159 & 0.0355 & 1.1380 \end{bmatrix} \tag{6.6}$$

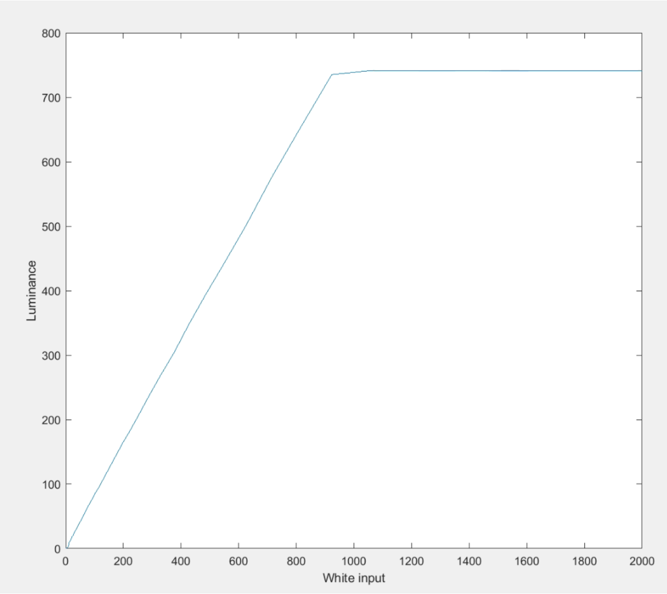


Figure 6.11: *The relationship between the input RGB to output CIE XYZ is linear as depicted from the given graph which is calculated as input neural ramp ($R = G = B$) on the x-axis while the output luminance (Y) is given on the y-axis.*

The measured and estimated peak white luminance and chromaticities are given in Table 6.4. In Figure 6.12, the chromaticity diagram illustrating the color gamut of the HDR display is plotted along with the color gamut of REC2020.

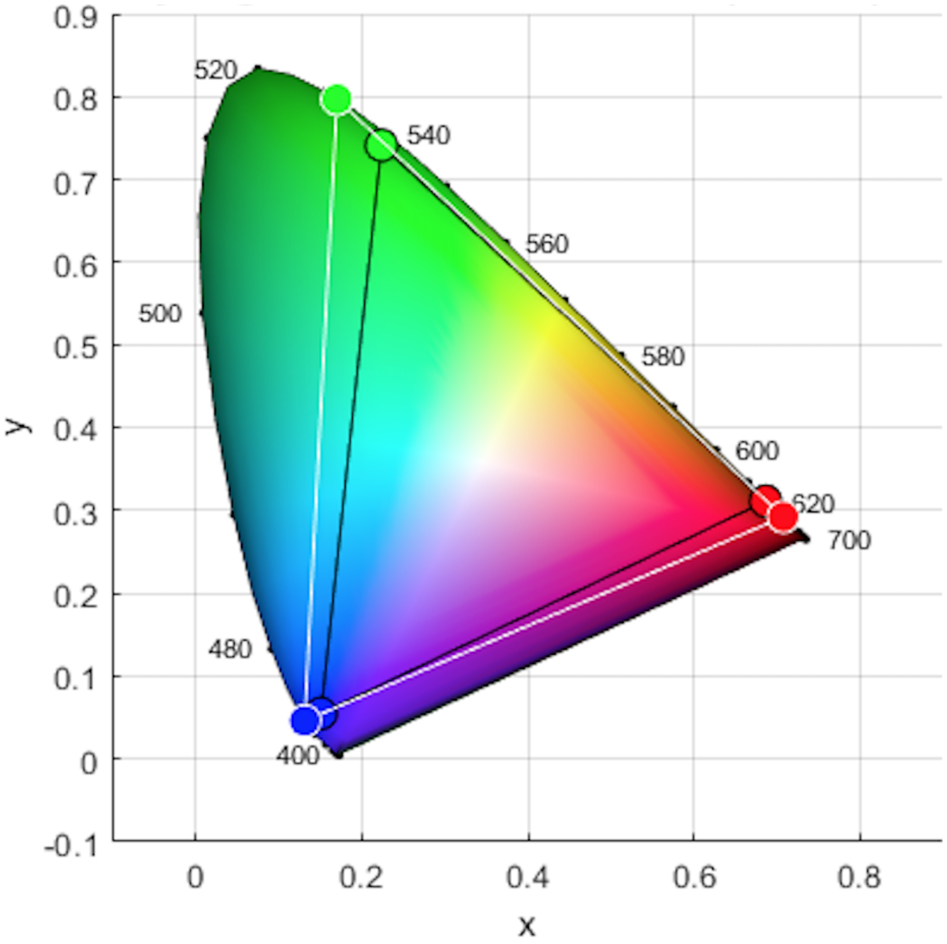


Figure 6.12: *The chromaticity diagram illustrates the color gamut of the Sony HDR display represented by the triangle with black boundaries whereas REC2020 is represented by white boundaries.*

Table 6.4: *Measured and estimated peak white luminance and chromaticities*

Channel	Y (cd/m^2)	x	y
Red	222.0700	0.6889	0.3085
Green	501.400	0.2049	0.7334
Blue	56.868	0.1478	0.0533
Measured peak white	780.338	0.3082	0.319
Calculated peak white	736.102	0.3212	0.3158
Δ	44.2	-0.013	0.003

7 Perceptual evaluation of proposed HDR rendering

Art is the lie that enables us to realize the truth

Pablo Picasso

As human observers are the end users of most image-based applications, the subjective image quality assessment is the most accurate and reliable approach for analyzing the perceptual quality of images. Additionally, there is no such computational model which is designed for perceptual evaluation of HDR rendering with real-world HDR scenes therefore it was necessary to conduct a psychophysical experiment in order to evaluate the performance of the proposed SiCAM model. The subjective IQA involves conducting psychophysical experiments where human observers are asked to judge the quality of images based on a predetermined scale. Category judgment, pairwise comparison, and rank order are the three most common types of psychophysical experiments for subjective IQA. Collecting subjectively annotated data through psychophysical experiments can be a challenging task as these experiments are time-consuming and expensive. They require setting up an appropriately controlled environment that involves many factors including the observer's visual acuity, display calibration, lighting condition, and also the physical and mental state of the observer. In addition, there are many potential flaws that researchers can easily overlook while conducting subjective assessments that could also affect collected subjective data e.g. the recruitment of observers with vision impairment, communicating unclear instructions, flaws in experimental design and outliers in data analysis, etc.

In this chapter, we have given a brief introduction of psychological studies and later our psychological experiment which has been conducted in order to evaluate the performance of the proposed SiCAM model is discussed. The main goal of most tone-mapping algorithms for high dynamic range is to faithfully reproduce the visual appearance of the original scenes while also striving for aesthetically pleasing results. Consequently, this subjective study was specifically designed to evaluate HDR rendering algorithms, considering both preference and accuracy as evaluation criteria.

7.1 An overview of subjective assessment

Subjective assessment entails psychophysical experiments in which human observers are asked to judge a set of images based on a predetermined criterion. Subjective evaluation is a traditional way of comparing the performance of several methods or algorithms that are intended to accomplish a similar objective. In HDR, this is required in a number of applications, e.g., to evaluate different tone mapping approaches, evaluate HDR compression algorithms for images and/or video, evaluate inverse tone mapping techniques, etc. The observers are required to visually compare the outputs generated from various techniques, which are generally referred to as test stimuli, with each other and/or with a ground truth stimulus. The ground truth stimulus is referred to as a reference stimulus. In these psychophysical experiments, usually participants with normal or corrected-to-normal vision are chosen, and at the beginning of each session, the observers are carefully instructed on the task at hand, the grading scale, and/or the sequence and timing of stimuli. The task is generally completed by interacting with a computer interface that displays stimuli and stores the scores given by the observers.

It is to be noted that conducting a psychophysical experiment can be a challenging task that is time-consuming. The quality of results obtained from such studies is highly dependent upon several factors which involve recruiting the observers, experiment methodology, duration of the experiment, viewing conditions, number and types of stimuli, etc.

7.1.1 International standards for subjective assessment:

Several international standards (BT.500-11, 2002; BT.710-4, 1998; P.910, 2008; BT.814-1, 1994; BT.1129-2, 1998; BT.1361, 1998; BT.815-1, 1994) are proposed to perform subjective image quality testing and provide reliable results. Among these ITU-R BT.500-11 (BT.500-11, 2002) proposes various methods for assessing the subjective quality of television images. This is a widely used standard that includes information about general viewing conditions, instructions on how to conduct subjective experiments, test materials, and subjective results presentation. For example, it has been mentioned that at least 15 observers should be used in experiments, participants should be carefully introduced to the method of assessment and should undergo a test session prior to the actual experiment, and details regarding general viewing conditions for subjective assessments in a laboratory environment and selection of test methods. The description of other relevant standards is summarized as below:

- RECOMMENDATION ITU-R BT.500-11 (BT.500-11, 2002): provides several methods for evaluating the subjective quality of television images.
- RECOMMENDATION ITU-T P.910 (P.910, 2008): provides details for assessing digital video quality at transmission rates less than 1.5 Mbits/sec.
- RECOMMENDATION ITU-R BT.814-1 (BT.814-1, 1994): provides recommendations for adjusting the brightness and contrast of the display screens.
- RECOMMENDATION ITU-R BT.1129-2 (BT.1129-2, 1998): provides details for evaluating the quality of the standard definition video sequences.

7.1.2 Methodologies of subjective assessment:

There are several existing methods for carrying out these types of experiments and the choice of a specific methodology impacts the accuracy and reliability of the collected data. Several past studies have compared the strengths and weaknesses of these methods (Pinson and Wolf, 2003; Rouse et al., 2010; Mantiuk et al., 2012). The three most commonly used methods for conducting psychophysical experiments namely category judgment, pair comparison, and rank order are discussed as follows:

7.1.2.1 Pairwise comparison

In pairwise comparison, observers assess quality by comparing image pairs. For instance, observers are tasked to choose the test stimulus with the highest quality or judge the quality difference from the reference image. These experiments can be forced-choice, where the observer must provide an answer, or non-forced-choice, where the observer may judge the two reproductions as equals (tie). Pair comparison experiments do not ask for information on the distance between the images, making them less precise than category judgment but also less complex. The example of a pairwise experiment is shown in Figure 7.1

Please select image with the highest quality

Figure 7.1: *The figure illustrates an example of pairwise comparison where an observer is forced to choose an image of the highest quality .*

7.1.2.2 Category judgment

The category judgment can be based on single and double stimulus as illustrated in Figure 7.2, in which observers rate the quality of a single or a pair of images on a pre-defined scale of five or seven levels which is also known as Absolute Category Rating (ACR). The primary benefit of category judgment is that the rating score of each image is recorded, but the task becomes more challenging for observers than pair comparison (Jones and McManus, 1986). Besides, these experiments are often faster than pair comparisons, with fewer comparisons necessary.

7.1.2.3 Rank order

In rank-order-based experiments, the participant ranks the test stimuli from best to worst according to given criteria as illustrated in Figure 7.3. The results obtained from such experiments are accurate in terms of comparing stimuli. As observers have to take a firm decision on the order of given test stimuli, it again becomes more time-consuming. However, it is faster than a pairwise comparison. These ranking of the test stimuli can also be computed indirectly using pairwise comparison-based experiments.

After collecting the data, we would have a subjective image quality dataset resembling the one shown below in Figure 7.4.

Figure 7.2: *The figure illustrates an example of category judgment where an observer is forced to rate a test image by utilizing the standard five-point Absolute Category Rating (ACR) scale.*

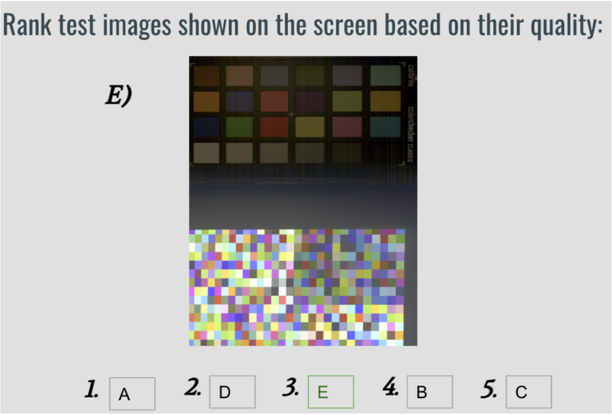


Figure 7.3: *The figure illustrates an example of ranking order where observers are asked to the test stimuli displayed on the screen according to given criteria.*

7.2 Our Experimental Setup:

Subjective assessment is the most reliable method for evaluating the quality of HDR images and their corresponding tone-mapped LDR images (Drago et al., 2003; Ledda et al., 2005). Besides, there is no computational model which is designed for perceptual evaluation of HDR rendering with real-world HDR scenes therefore it was necessary to conduct a psychophysical experiment in order to evaluate the

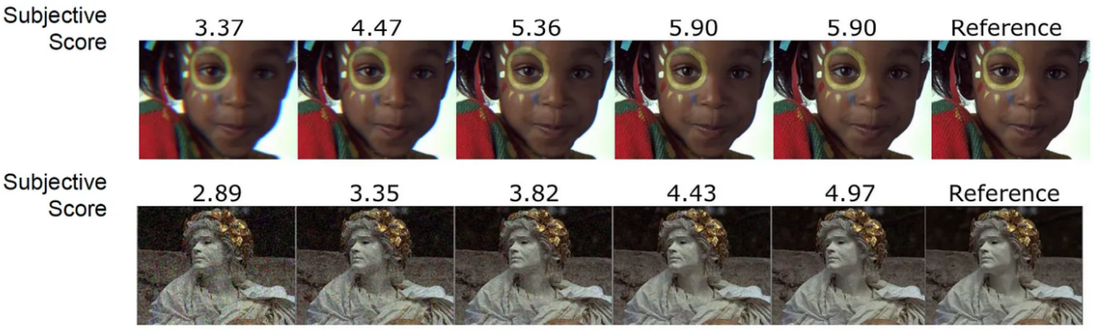
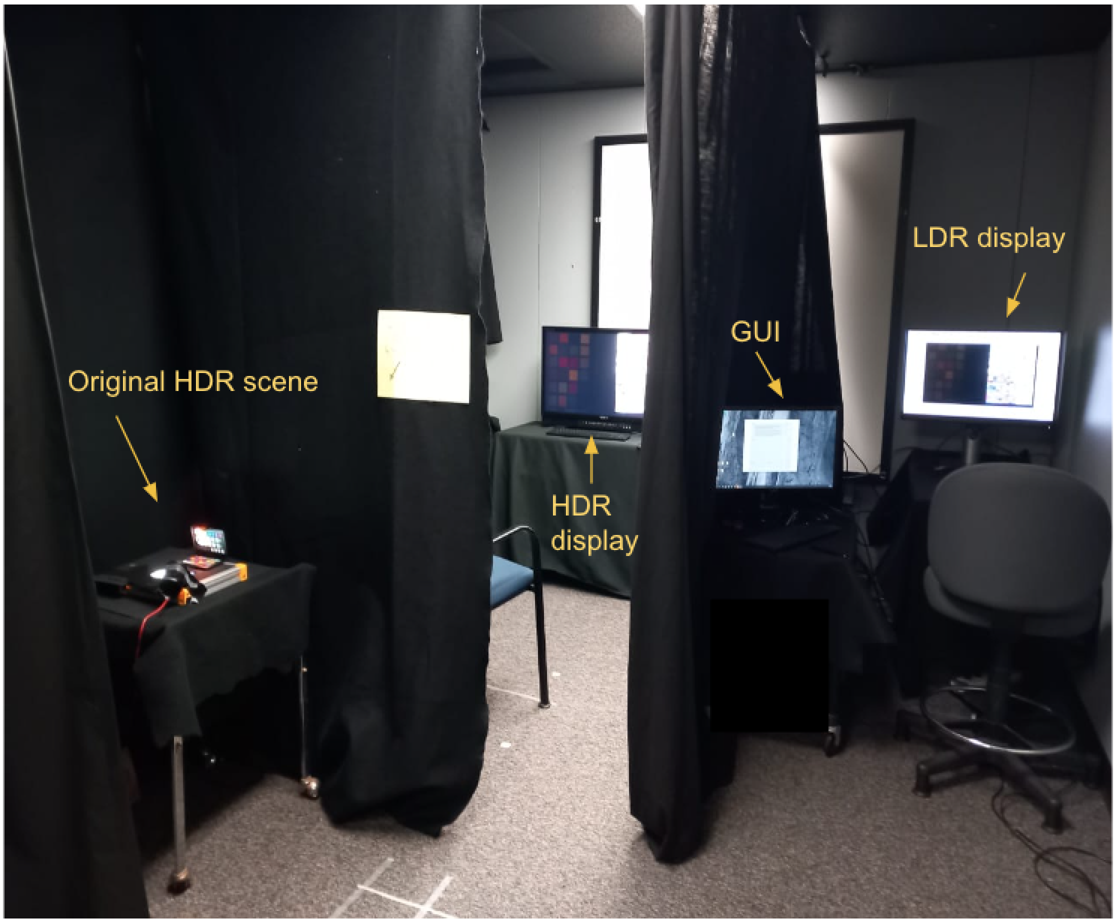


Figure 7.4: *After the subjective experiment, each test image with be associated with a mean opinion score which is calculated using the ratings given by all observers (Ponomarenko et al., 2013).*

performance of the proposed SiCAM model. This experiment was based on rating the HDR rendering on both HDR and LDR displays while the real-world HDR scene was considered as a reference. The primary goal of most tone-mapping algorithms is to faithfully reproduce the visual appearance of the original scenes while also striving for aesthetically pleasing results. Consequently, this subjective study was specifically designed to evaluate HDR rendering algorithms, considering both preference and accuracy as evaluation criteria. The details of the experimental setup are discussed in the following subsections.

7.2.1 Observers:

In order to evaluate the performance of the proposed SiCAM model, we have conducted a psychophysical experiment in which 30 observers participated among which 11 observers were female and 19 observers were male. Among all male observers, one was found to have Colour Vision Deficiency (CVD) and hence eliminated from this study. All the observers who participated were with either normal visual acuity or corrected-to-normal. The average age of the observers was approximately 27 years among which 14 observers were already aware of image quality assessments and most of them had prior experience of participating in such psychophysical studies. The observers were communicated clear instructions on how to participate while a pilot experiment was conducted before starting to perform the actual experiment.



[redacted]

Figure 7.5: *The figures illustrates the psychophysical experimental setup for evaluating the image quality of HDR image rendering by different methods. The experiment was set up in a dark room with three adjoining partitions to avoid interaction for three different types of stimuli during the subjective study which includes an original HDR scene, an HDR image displayed on an HDR monitor, and tone-mapped images which were displayed on an LDR monitor.*

7.2.2 Viewing conditions

The experiment was set up in a dark room with three adjoining partitions to avoid interaction for three different types of stimuli during the subjective study which includes an original HDR scene, an HDR image displayed on an HDR monitor, and tone-mapped images which were displayed on an LDR monitor as shown in Figure 7.5. The HDR image was rendered on an HDR display using HDR Sony display as discussed in Section 4.6 while the calibration method for HDR display is discussed in Section 6.5. To recall, we employed the PVM-X2400 24-inch 4K

HDR Sony display with the resolution of 1920-by-1080 pixels, to show an HDR image of the physical HDR scene. The HDR display has been configured to operate in the PQ Rec. 2020 mode using 10-bit pixel encoding. However, for displaying HDR images on this HDR Sony display, we used MATLAB Psychtoolbox-3. The tone-mapped images were shown on a 24,1 inches, color-calibrated DELL monitor. The LDR display was calibrated to sRGB color space and the pixel resolution was 1920-by-1080 similar to HDR display. The white point was calibrated to 6500K color temperature and luminance was set to 80cd/m2. The observers were positioned at a distance of roughly 100cm from the LDR screen and 160cm from the HDR screen. At these distances, each stimulus was subtended approximately 14.25 degrees of visual angle.

7.2.3 Test stimuli and HDR scene:

In total seven tone-mapped images were used in this experiment which were generated using three different methods as listed in Table 7.1. It is important to note that the input of iCAM06 and gamma tone mapping is three channel HDR image whereas only input of SiCAM is HDR hyperspectral radiance cube. The images generated from the reduced number of hyperspectral radiance bands were included to investigate how many bands are necessary for a better tone-mapped image. The hyperspectral bands are reduced by taking the weighted average of spectral samples by applying a Gaussian averaging as given by equation 7.1 where L is the number of samples in a window, α was set to 1.8 and L was set to 3 and 11 for reducing the number of bands to 39 and 10 respectively. We did not skip the middle wavelengths and opted for an odd number of samples for applying Gaussian weights in order to avoid aliasing.

Table 7.1: *Table.7.1 The details of test images used during the experiment are summarized in the given table. It is important to note that the input of iCAM06 and gamma tone mapping is three channel HDR image whereas only the input of SiCAM model is HDR hyperspectral radiance cube.*

HDR Input	Created from	Tone mapping method
3-channel RGB	119-channel spectral	iCAM06
119-channel spectral	same as input	SiCAM
10-channel spectral	119-channel spectral	SiCAM
39-channel spectral	119-channel spectral	SiCAM
3-channel RGB	119-channel spectral	Gamma
3-channel RGB	10-channel spectral	Gamma
3-channel RGB	39-channel spectral	Gamma

$$w(n) = e^{-\frac{1}{2}\left(\alpha \frac{n}{(L-1)/2}\right)^2} \tag{7.1}$$

7.2.4 Procedure

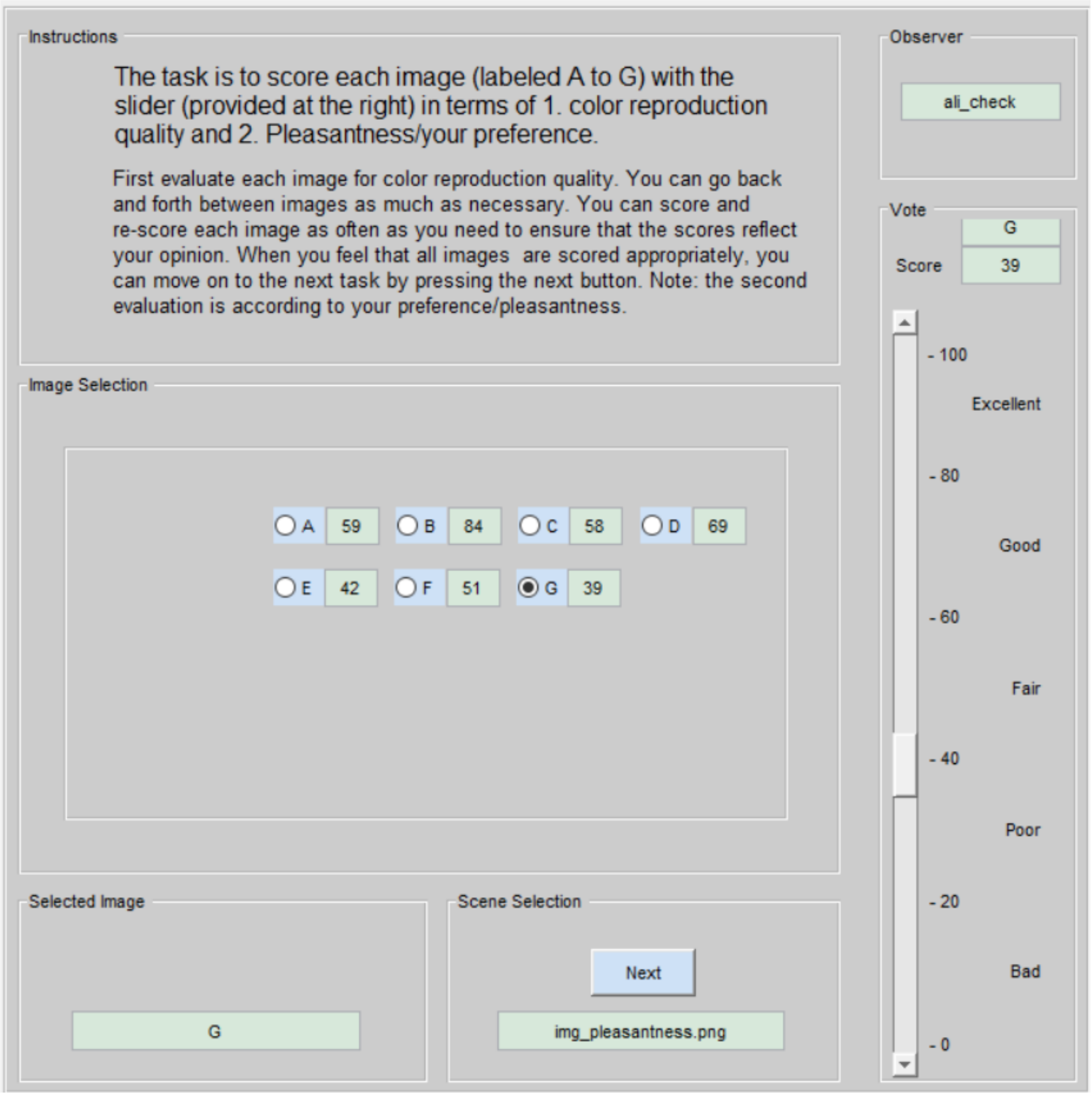
As mentioned earlier, three different stimuli have been utilized in this experiment i.e., original HDR scene, an HDR image displayed on the HDR monitor screen, and seven tone-mapped LDR images which were displayed on the LDR monitor screen. This subjective study was specifically designed to evaluate HDR rendering algorithms based on their respective output images as given in Table 7.1, considering both preference and accuracy as evaluation criteria. The observer's task was to judge the seven LDR reproductions where the original HDR scene was considered as the main reference. These test images were displayed against a gray background with a luminance level equivalent to 20% of the adapting white point while the observer-rated these LDR test stimuli using a user interface as illustrated in Figure 7.6 that was running on a connected screen. The experiment was conducted in two parts as follows:

- In the first part, observers rate the tone-mapped LDR images (total 7 of them) according to accurate reproduction by keeping the original scene as a ground truth. Observers give judgments based on the rating scale of 100 using the slider where a score of 100 indicates excellent reproduction and a score of 0 indicates worst reproduction.
- In the second part, observers rate each of the seven LDR images according to pleasantness/preference by keeping the original scene as ground truth once again on the same rating scale.

It is important to note that the white point and luminance range of the LCD display and the physical HDR scenes differed significantly. As a result, participants were required to undergo a minimum adaptation period of 60 seconds for both the physical HDR scene and the LCD display. The participants were allowed to complete the experiment in a single sitting based on their memory. However, they had the option to revisit the original HDR scene whenever they deemed it necessary. The intention behind enforcing repeated viewing of the original scene was to ensure that participants based their judgments on the accuracy of the rendering, especially during that part of the experiment where the observers were required to rate the LDR reproduction based on accurate reproduction(i.e., the first part of the experiment). The observers performed the experiment in random order where half observers performed part one of the experiment first and then performed the second part while the rest of the observers did the experiment in reverse order. When observers finished both parts of the experiment they were asked an open-ended question to comment about what was the factors or cues that they were considering while making the judgments. Lastly, the observers were also shown an HDR image on an HDR display screen using the proposed display referred HDR rendering workflow as discussed in Section 4.6. They were asked to judge this HDR image by considering the original HDR scene as the reference. Once again the observers were obligated to have a minimum adaptation period of 60 seconds for both the real-world HDR scene and the HDR display while they were allowed to view the HDR scene anytime they felt it necessary while answering the questionnaire. For rating the HDR display after which the observers were asked the following questions:

- Rate the overall impression of image quality of the HDR image by considering the physical HDR scene as a reference on the rating scale of 100 where a score of 100 indicates excellent reproduction and 0 indicates the worst reproduction.
- Observers were also asked to give individual scores for contrast, colorfulness, image sharpness, and overall naturalness of the HDR image on the above-mentioned scale.

we carried out this perceptual evaluation to assess the performance of the SiCAM model compared to iCAM06 and gamma tone mapping. Our goal was to determine how well each algorithm faithfully reproduced the visual appearance of the original HDR scene while also producing aesthetically pleasing results. Thus, this subjective study was designed to evaluate the HDR tone mapping algorithms, considering both preference and accuracy as evaluation criteria. The data analysis and results of the perceptual evaluation are discussed in Chapter 9.



Instructions

The task is to score each image (labeled A to G) with the slider (provided at the right) in terms of 1. color reproduction quality and 2. Pleasantness/your preference.

First evaluate each image for color reproduction quality. You can go back and forth between images as much as necessary. You can score and re-score each image as often as you need to ensure that the scores reflect your opinion. When you feel that all images are scored appropriately, you can move on to the next task by pressing the next button. Note: the second evaluation is according to your preference/pleasantness.

Observer: ali_check

Vote: G

Score: 39

Image Selection:

○ A 59
○ B 84
○ C 58
○ D 69
○ E 42
○ F 51
● G 39

Selected Image: G

Scene Selection: img_pleasantness.png

Figure 7.6: *The figure illustrates the user interface running on a connected screen for rating the test stimuli which were separately displayed to the observers on a calibrated LDR display screen during the psychophysical experiment as shown in 7.5.*

8 Objective assessment of HDR rendering:

To measure is to know

Lord Kelvin

The HDR displays allow the perception of a larger range of luminances and a higher bit depth compared to conventional standard dynamic range displays. These features make human vision to have more plausible and realistic representations due to the clear perception of details in both the darker and brighter areas at the same time. As demonstrated in prior chapters, many techniques have been suggested for various stages within the high dynamic range (HDR) pipeline. Given the considerable number of methodologies available, it is essential to understand the comparative advantages of each. Consequently, a diverse range of HDR image quality assessment (IQA) methods have been proposed to assess and compare the quality of High dynamic range (HDR) images.

In this chapter, we have given a brief introduction to the most popular approaches to image quality assessment which is to use computational algorithms known as image quality metrics (IQMs) to try to predict the quality of natural images. Later, we discussed the IQMs which are fundamentally designed for HDR content since HDR content poses new challenges with respect to conventional Low Dynamic Range (LDR) images. As tone mapping plays a vital role in HDR imaging workflow to visualize HDR images on LDR displays hence a tremendous amount of tone mapping operators (TMOs) have been proposed in past research. In this context, image quality metrics(IQMs) for evaluating the tone-mapped representation of the HDR data have also been addressed. Lastly, the IQMs utilized for evaluating the performance of the proposed HDR rendering methods have been discussed whereas the results of the objective assessment are discussed in Chapter 9.

8.1 An overview of objective image quality assessment

The image quality assessment (IQA) can be classified into subjective IQA and objective IQA. The objective IQA aims to provide quantitative metrics known as image quality metrics (IQMs) that can predict the image quality without the involvement of human observer judgments about the perceived image quality. The ultimate goal of objective IQA is to produce efficient and consistent judgments which should also correlate with human quality judgments. According to the availability of reference (original/distortion-free) images with which the test image (reproduced/distorted) images are to be compared, IQMs can be categorized into three types as follows: full reference, no reference, and reduced reference. The majority of IQA metrics are full reference, which implies that a reference (i.e., original) image is known. However, when the full reference image is not available, making a no-reference or blind quality assessment approach is preferable. In reduced-reference quality assessment, the reference image is only partially provided that uses a collection of extracted features as side data to assess the quality of the test image. These IQMs are often designed by taking into account the knowledge of how the human visual system (HVS) analyzes the image quality and intend to give predictions that are close to the prediction of image quality by an average human observer. These metrics are derived from many ideas and were created for various goals, such as quantifying distortions, producing benchmarks, monitoring quality, optimizing a process, or indicating trouble areas.

Since different metrics are proposed for distinct purposes, it is critical to consider their area of application while evaluating their effectiveness. citechandler2013seven presented a detailed survey of quality assessment where the first half of his review discusses key visual perception attributes, IQA databases, and existing IQA algorithms while the rest of his survey focuses on several outstanding challenges in quality assessment. Nevertheless, his survey focuses more on IQA problems and challenges rather than algorithms. Pedersen et al. (2012) provided a concise summary of over 100 IQMs. Mohammadi et al. (2014) reviewed both subjective and objective IQA methods with a focus on nine classic full reference IQA measures and two emerging techniques: high dynamic range (HDR) and three-dimensional(3D) rendering. While Rousselot et al. (2019) presented HDR IQA based on HDR uniform color spaces and proposed two new databases.

8.1.1 Generic framework of image quality metrics:

As discussed above there are hundreds of proposed image quality metrics however the majority of these metrics usually follow a common framework as illustrated in Figure 8.1. Here we have presented a generic workflow of image quality metrics using full reference as an example. The first step is usually transforming the input RGB values to another more suitable color space where a better representation of the perception of color can be achieved, such as an opponent color space or a perceptually uniform color space is commonly preferred. The second stage involves modeling important aspects of the human visual system (HVS). These models usually simulate low-level features of the HVS, such as contrast sensitivity functions (CSFs) or masking while other high-level features can also be simulated based on the idea that our human visual system is adapted to extract information or structures from the image. Then, the quality calculation is performed based on calculating pixel-wise differences i.e., the most basic approach is to calculate the Euclidean distance. Finally, pooling is performed on these quality scores to reduce these so many numbers and generate a single score indicating overall image quality.

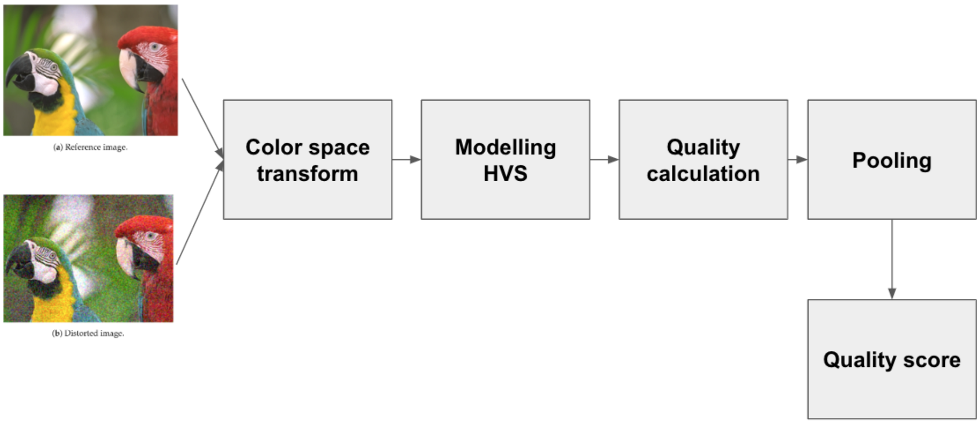


Figure 8.1: *The generic workflow of image quality metrics that involves four main stages which include transforming the input RGB values to another more suitable color space, modeling certain aspects of the human visual system (HVS) such as contrast sensitivity functions (CSFs), estimating quality scores and finally pooling these scores to assign a single quality score value.*

8.1.2 Classification of image quality metrics

The IQMs can be classified into five classes based on their underlying working principle which are listed below. Please refer to the seminal work of Pedersen et al. (2012) for detailed information on each of these groups of metrics.

8.1.2.1 Mathematically based metrics

These IQMs operate only on the intensity of the distortions by calculating the image statistics such as Mean Square Error(MSE), ΔE_{ab} while the metrics measuring color difference also belong to the group of mathematically based metrics. These simple mathematical models are usually not well correlated with perceived image quality.

8.1.2.2 Low-level based metrics

Metrics classified as low-level based metrics simulate the low-level features of the HVS, such as contrast sensitivity functions (CSFs) or masking whiles others also depend on image decorrelation-based methods to separate color and luminance information in an image where the input RGB image is first transformed into CIE XYZ and then further into the opponent color space. The examples of low-level based metrics are S-CIELAB or S-DEE.

8.1.2.3 High-level based metrics

High-level based metrics quantify quality based on the idea that our HVS is adapted to extract information or structures from the image such as SSIM defines the structural information in an image as those attributes that represent the structure of the objects in the scene, independent of the average luminance and contrast.

8.1.2.4 Other metrics

These metrics are either based on other strategies or combine two or more of the above metric types. For example, Visual Signal to Noise Ratio (VSNR) was proposed by [Chandler and Hemami, 2007] which was based on near-threshold and suprathreshold aspects of the HVS i.e., incorporating both low-level features and mid-level features.

8.2 HDR image quality metrics:

Compared to conventional LDR quality assessment, evaluating HDR visual quality poses new challenges (Narwaria et al., 2016). The higher peak brightness and contrast provided by HDR increases the visibility of artifacts, while simultaneously altering how viewers focus their attention in comparison to LDR (Zerman et al., 2017; Hanhart et al., 2015). In addition, color distortion plays a significant role in the overall quality evaluation as a result of the increased luminance (Pouli et al., 2016). Since these and other additional factors interact in a complex manner to determine HDR visual quality. Therefore, LDR image quality metrics can not directly be applied to HDR content. To overcome this performance, the two main approaches are followed to develop HDR image quality metrics which are discussed:

8.2.1 Adapting LDR IQA metrics to HDR images

Because of the lack of IQMs designed specifically for HDR images, various objective LDR metrics are also used to assess the quality of HDR images. However, these LDR metrics assume that pixel values are perceptually linear, i.e., that change in intensity of pixel values corresponds to an equivalent amount of change in perceived luminance whereas the HDR pixel values store linear radiometric value, i.e., pixels are proportional to the scene's physical luminance which is not same as the perceived luminance. Human perception has a more complex behavior where the luminance perception can be approximated by a square root at low luminance levels and is approximately proportional to luminance ratios at higher luminance levels (Kundu and Pal, 1986). Therefore, to get meaningful quality scores using LDR metrics such as from PSNR and SSIM for HDR images, the physical luminance of HDR images needs to be converted to perceptually uniform pixels values by utilizing a logarithmic or perceptually uniform encoding as illustrated in Figure 8.2. The logarithmic encoding is used for a very coarse approximation of the HVS response along the entire visible luminance range and also does not account for the loss of sensitivity for the low light conditions while the drawbacks of the logarithmic encoding can be overcome by applying the PU encoding to generate perceptually uniform pixel values for HDR images (Aydın et al., 2008).

8.2.2 HVS inspired HDR metric

The second approach is to design dedicated HDR IQMs such as HDR VDP2 (Mantiuk et al., 2011) is designed for HDR images and HDR-VQM (Narwaria et al., 2015) metrics is designed for HDR video sequences where both are based upon full reference IQA. These metrics require modeling of the human visual system

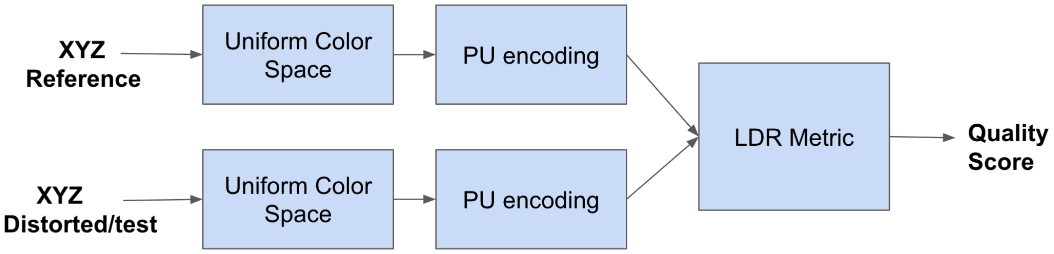


Figure 8.2: *The workflow of extending LDR metrics to HDR content.*

(HVS). For instance, the HDR-VDP metric incorporated the different models of the early stages of HVS, such as intra-ocular scattering, luminance masking, and the achromatic response of the photoreceptors, to accurately predict the visibility and strength of the pixel-wise distortion in HDR images where HDR-VQM embedded a spatio-temporal analysis based model that correlates with the HVS fixation behavior during the video frames.

8.2.3 Relevant HDR datasets:

Table 8.1 provides a concise description of publicly available HDR image databases along with their characteristics and links to access these databases.

Table 8.1: *HDR Image Datasets*

Dataset Name	Number of Images	Observers	Method	Access Link
Valenzise Dataset (HDR_IMT) (Valenzise et al., 2014)	50	15	DSIS	`http://webpages.l2s.centralesupelec.fr/`
IRCCyN-IVC HDR Image Datasets (Narwaria et al., 2013)	160	27	ACR-HR	`http://ivc.univ-nantes.fr/en/databases/JPEG_HDR_Images/`
HDR Photographic Survey (Fairchild, 2007a)	106	-	-	`http://markfairchild.org/HDR.htm`
Rousselot_et_al_EUSIPCO2018 (Rousselot et al., 2019)	104	27	DSIS	`http://www-percept.irisa.fr/software/`
HDRcompressed V2(Zerman et al., 2017)	110	15	DSIS	`http://webpages.l2s.centralesupelec.fr/perso/giuseppe.valenzise/download.htm`
HDdTB HDR dataset (Rousselot et al., 2018)	100	27	DSIS	`https://hevc.hhi.fraunhofer.de/`
Database of HDR images compressed with JPEG XT(Korshunov et al., 2015)	240	24	DSIS	`http://mmspg.epfl.ch/jpegxt-hdr`
HDR Image Dataset(Xiao et al., 2002)	210	29	ACR-HR	-
Tone-mapped Image Database (TMID)(Yeganeh and Wang, 2012)	180	20	ACR-HR	`https://qualinet.github.io/databases/image/tone-mapped-image-quality-database/`
RV-TMO (Krasula et al., 2015)	250	3500	DSIS	`ftp://ftp.polytech.univ-nantes.fr/RV-TMO`

8.3 Objective metrics used for proposed HDR rendering method:

For investigating the effectiveness of utilizing hyperspectral data for HDR rendering and tone mapping, we have mainly relied on perceptual evaluation by conducting a psychophysical experiment. However, we have also utilized a few objective IQA metrics among which two are full reference metrics i.e., Tone-mapped quality index (TMQI) (Yeganeh and Wang, 2012) and Feature similarity index for tone-mapped image (FSITM) (Nafchi et al., 2014). These metrics are designed to evaluate the perceptual quality of tone-mapped high dynamic range images by considering the corresponding HDR image as the reference. Both of these metrics are further discussed in the following subsections.

8.3.1 Tone-mapped quality index (TMQI)

Yeganeh and Wang (2012) introduced the Tone Mapped Quality Index (TMQI) as a comprehensive image quality assessment metric designed to evaluate the quality of tone-mapped high dynamic range (HDR) images. TMQI measures the perceptual similarity between a tone-mapped image and its reference HDR image, taking into account the luminance and contrast adjustments during the tone mapping of HDR images.

The TMQI algorithm operates by quantifying differences between the tone-mapped image and the reference image by considering both global and local image characteristics to provide a comprehensive evaluation of image quality in terms of luminance, contrast, and structural features. It considers the perceptual importance of different frequency components to weight the luminance differences and captures the image's overall contrast changes introduced during tone mapping by utilizing the contrast sensitivity function of HVS. Additionally, it evaluates the structural similarity of textures and edges between the test and reference images by employing a structural similarity index to measure the similarity of local image patterns. Finally, TMQI calculates an image quality score as a combination of these components that represents the quality of the tone-mapped image in comparison to the reference HDR image.

8.3.2 Feature similarity index for tone-mapped image (FSITM)

The Feature Similarity Index for Tone-Mapped Images (FSITM) (Nafchi et al., 2014) is another dedicated metric for assessing the quality of reproduced HDR tone-mapped images. The primary objective of FSITM is to quantify the degree of similarity between the perceptual features of a tone-mapped image and the corresponding HDR reference image. The FSITM algorithm operates by comparing and analyzing the perceptual local similarity maps based on the locally weighted mean phase angle to evaluate the overall quality of tone-mapped images. The authors presented the performance results of FSITM on two datasets and also compared them to the state-of-the-art TMQI where FSITM was found to output the TMQI. It would be interesting to combine the scores of both FSITM and the TMQI to produce a more precise quality evaluation.

8.3.3 Other metrics

Besides, the above-mentioned full reference metrics that are fundamentally designed for evaluating the tone mapping output, other blind metrics have also been considered such as the perception-based Image Quality Evaluator (PIQE) (Venkatanath et al., 2015) where a smaller score indicates better perceptual quality and already demonstrated its efficacy on three widely used image quality datasets. Furthermore, we also estimate the scores of the sharpness of the reproduced LDR images and also calculated their contrast ratio which are two essential factors of image quality. The results of the objective image quality assessment of the tone-mapped LDR images utilized in our psychophysical are discussed in Chapter 9.

9 Results and discussion

The end is where we start from.

T.S. Eliot

The results of both the perceptual evaluation of HDR rendering and the objective image quality assessment with the HDR real-world scene are discussed in this chapter. Additionally, the limitations of this study and future directions have also been addressed. Finally, the conclusion of this master thesis is presented at the end of this chapter.

9.1 Subjective assessment

As human observers are the end users of the majority of image-based applications, the subjective image quality assessment is the most accurate and reliable approach for analyzing the quality of images. In our study, we aimed to assess how the visual quality perception and realism in HDR-rendered images are affected by using HDR hyperspectral data instead of traditional HDR three-channel data. To achieve this, we conducted a psychophysical study, where participants were shown HDR-rendered images generated using two different approaches: one produced using traditional three-channel HDR data and the other using hyperspectral HDR data. In addition, we also reduced the number of spectral bands to generate some test stimuli for investigating the effect of reducing spectral information. Test stimuli are generated using three types of tone mapping methods as listed in 7.1, including Gamma 10-bands, Gamma 39-bands, Gamma Full, iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full. The suffix "Full" indicates images with 119 spectral bands covering the entire visible spectrum, while "n-bands" denotes output generated from reduced spectral bands HDR image as input. It is important to note that SiCAM had a hyperspectral cube as input, making it distinct from the other tone mapping methods which used three-channel HDR data.

The details of how this psychophysical experiment was set up and conducted are discussed in Chapter 7. This experiment was conducted to evaluate the effectiveness of the SiCAM model, discussed in Chapter 5, which leverages hyperspectral HDR information. We compared its performance with that of iCAM06, a model that relies solely on three-channel HDR input. Furthermore, we integrated gamma tone mapping into our test stimuli. Our goal is to determine how faithfully each algorithm reproduced the visual appearance of the original HDR scene while also producing aesthetically pleasing results. Thus, this subjective study was explicitly designed to evaluate the HDR tone mapping algorithms, considering both preference and accuracy as judging criteria.

The distribution of the rating scores given by the observers is depicted using box plots, which offer insights into the central tendency, spread, skewness, and presence of outliers within the collected data during the psychophysical experiment. Additionally, important statistical analyses of these subjective scores are summarized in their respective tables. Lastly, we conducted an ANOVA (Analysis of Variance) (St et al., 1989) on the subjective scores. This involves interpreting the estimated differences in the subjective scores of the different tone-mapping methods to ascertain whether significant differences exist among the subjective scores in our study. This approach aids us in extracting meaningful insights and conclusions from these subjective scores.

The results of this subjective evaluation are discussed in the following subsections, highlighting the differences in perceptual quality and observer's preference between the different types of rendered images as listed in Table 7.1. Through a comprehensive analysis of the rating scores obtained from the psychophysical experiment, we extracted valuable insights regarding the efficacy of employing HDR hyperspectral data for HDR rendering and its impact on the overall visual experience and realism of the HDR images.

9.1.1 Rating scores for tone-mapped LDR images:

In total, seven test stimuli were presented to the observers as listed in Table 7.1 during the experiment. Since it has been observed that the LDR images reproduced by the SiCAM model were preferred by the majority of the participants as discussed in the following sections, it can be concluded that the tone-mapped images of SiCAM have more natural colors and contrast levels, and thus are a more faithful reproduction of the original HDR scene in terms of both accurate reproduction and pleasantness. The individual scores given by the observers for color appearance reproduction and image pleasantness are discussed in the following subsection

9.1.1.1 Color appearance scores

The subjective scores given by the observers to color appearance (i.e., how accurately the tone mapping methods reproduced the appearance of the corresponding real-world HDR scene) are summarized in the box graph given in Figure 9.1. This box plot contains rating scores for seven different tone-mapped reproduced images including iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full(i.e., where Full indicates 119 spectral bands within the visible spectrum). It can be observed that SiCAM has been found to outperform other tone mapping approaches whereas the results of iCAM06 are comparable to that of SiCAM. The box plot given in this Figure 9.1 represents an overall distribution of subjective scores given by the observer during the perceptual evaluation of the test stimuli. The red-colored central line denotes the median value of seven reproduced LDR images which were generated using different tone mapping methods as given in Table 7.1. The width of the blue box in the box plot indicates the interquartile range (IQR), which is the range between the 25th percentile (Q1) and the 75th percentile (Q3) of the data.

It can be observed from this graph that Gamma TMO has obtained significantly lower scores than iCAM06 and SiCAM for accurate color appearance reproduction. Interestingly, the rating scores revealed that SiCAM with 10 bands of hyperspectral input yielded the best results in contrast to our initial hypothesis that using all 119 bands covering the whole range of the visible spectrum would lead to better outputs. The possible explanations for this trend are given in the discussion section. Additionally, the smaller size of the box for SiCAM-10 indicates that the variation in the scores of the observers is also minimum which shows that the rating trend was consistent for this reproduced LDR image.

On the other hand, iCAM06 output also obtained rating scores which are close to that of SiCAM where the input was three channel HDR image. It is worth mentioning that the observers were shown one image at a time during the experiment. It might be possible that the nonuniform target patches(transmissive film) and an overall slightly yellowish color cast on the output of iCAM06 as shown in Figure 9.2, were not that visible to participants affecting the rating of their judgments. For a better comparison of SiCAM and iCAM06, we analyzed color fidelity by estimating the patch uniformity of the target patches in our physical HDR scene as presented in Figure 9.3. To achieve this, we first annotated the patches of film target patches of LDR images that were reproduced by these specific tone mapping methods and estimated their standard deviation. As the uniformity of these patches is inversely proportional to standard deviation, we can observe that the output LDR images generated by SiCAM are more uniform than iCAM06

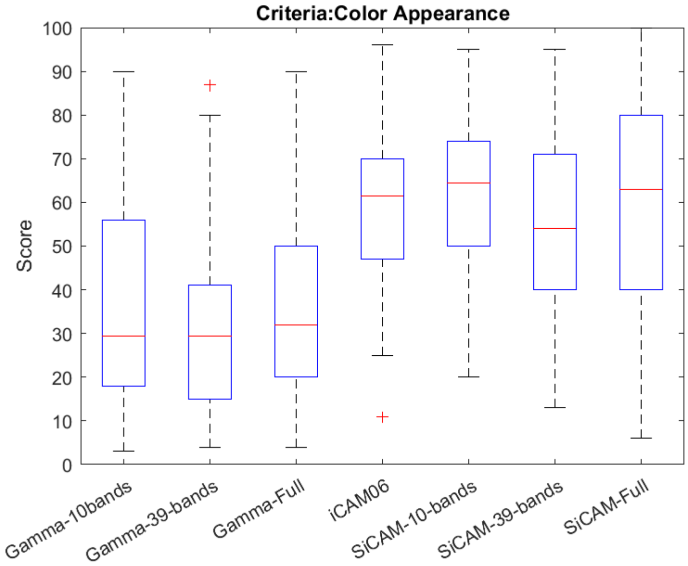


Figure 9.1: *Subjective scores for color appearance.*

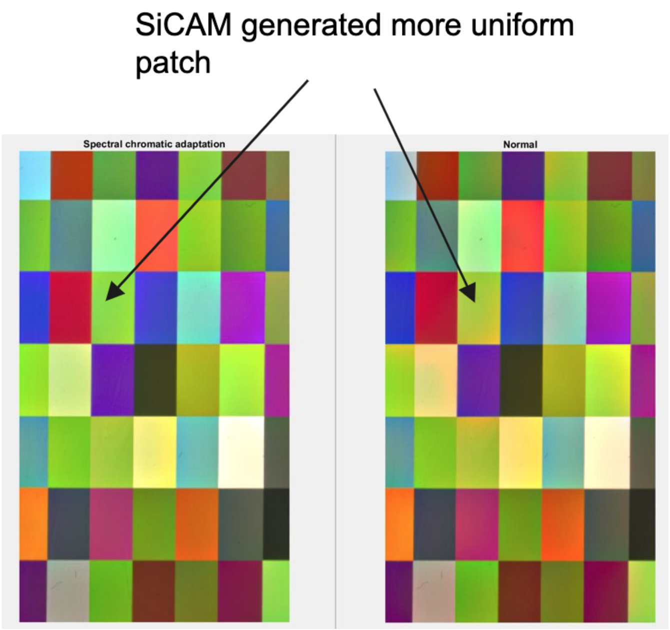
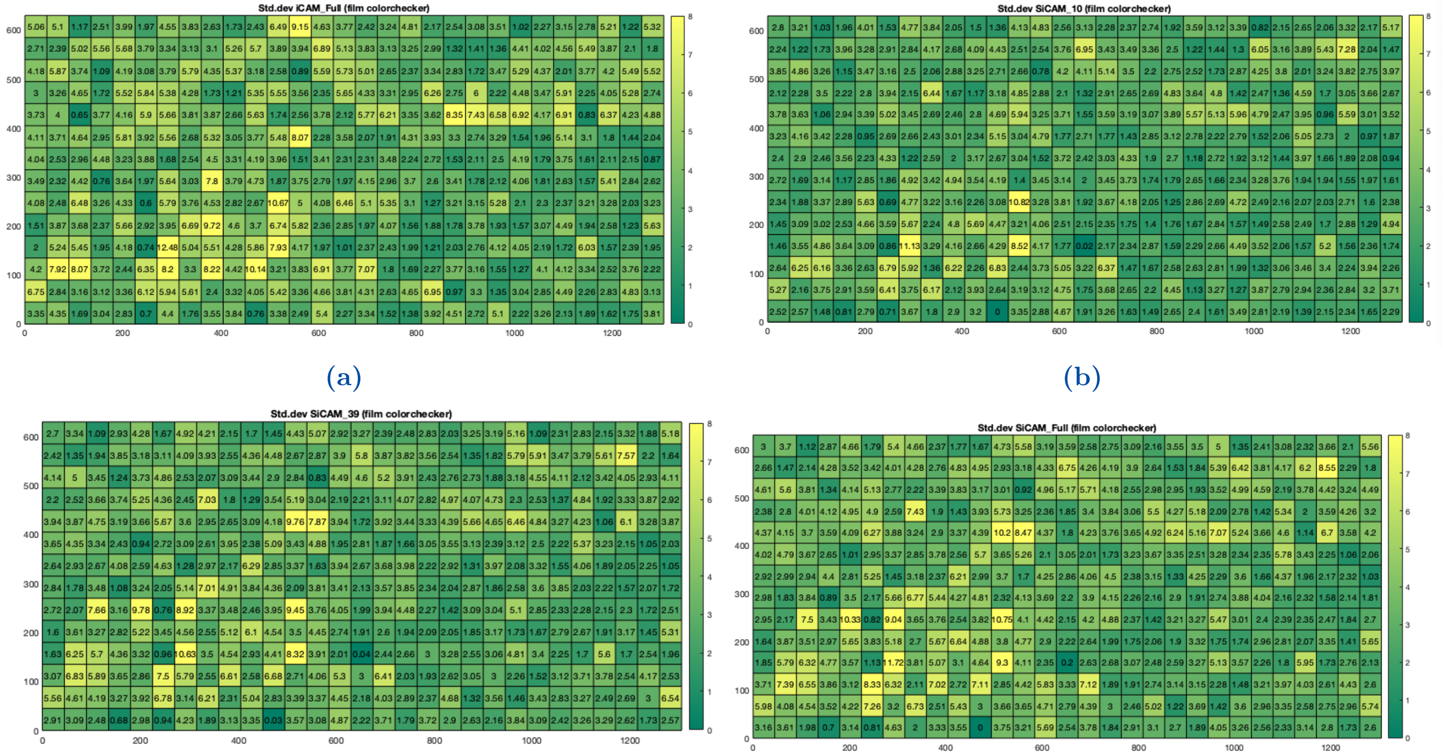


Figure 9.2: *The figure illustrates the zoom-in patches of the outputs generated by SiCAM with 119 spectral bands of hyperspectral input and iCAM06 with three channel HDR image.*

as illustrated in Figure 9.2. Additionally, SiCAM 10-bands have been found to give the best results among these LDR images.



(a)

(b)

(c)

(d)

Figure 9.3: *These given heatmaps illustrate the standard deviation which is inversely proportional to the color uniformity of the color target (film) where the color code varies from green to yellow corresponding to standard deviation low to high. It can be observed that the output LDR images generated by SiCAM are more uniform (i.e., more greenish patches are visible in their heat maps) than iCAM06(i.e., more yellowish patches are visible in its respective heat map) which can also be observed in Figure 9.2.*

Furthermore, as the experiment was conducted during the open house session of Munsell color science Laboratory(MCSL), the participants who had a strong background in color science(i.e., industrial professionals and senior researchers) also took part in this psychophysical experiment. It is interesting to highlight that these participants gave higher scores to SiCAM-Full while there is a discernible difference in the rating scores of iCAM06 and SiCAM-Full as illustrated in Figure 9.4.

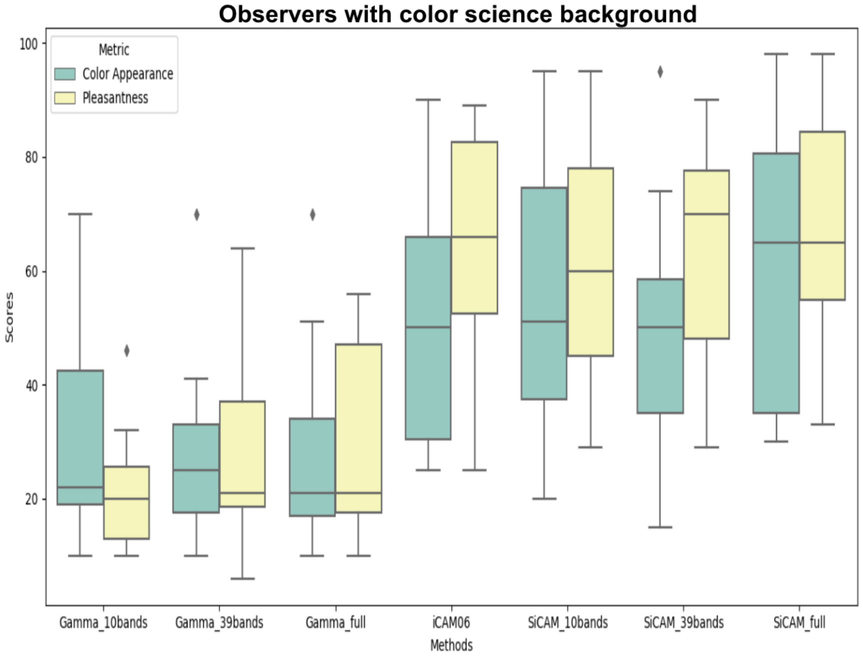


Figure 9.4: *The plot given in this figure shows the rating scores given by the observers with a strong background in color science where there is a discernible difference in the rating scores of iCAM06 and SiCAM-Full.*

Additional statistical analyses to complement the insights gained from the box plot 9.1 are given in Table 9.1 which shows the mean, standard deviation (std.dev), and standard error (std. Error) of the rating scores given by the observer during the experiment where it can be seen that the output images generated by SiCAM 10-bands and SiCAM full have notably higher mean color appearance scores compared to the other methods.

9.1.1.2 Pleasantness scores

The pleasantness scores(i.e., the observer preference irrespective of the accuracy of reproduction) are given in Figure 9.5. When the pleasantness scores are compared,

Table 9.1: *Summary of Color Appearance Scores: The table summarizes the statistical analysis of the observers' score for perceptual evaluation of color appearance for tone-mapped images*

Method	Mean	Std. Dev.	Std. Error
Gamma 10bands	37.458	27.0603	5.523
Gamma 39bands	34.125	25.0847	5.120
Gamma full	36.875	25.9084	5.288
iCAM06	57.125	22.3535	4.562
SiCAM 10bands	**60.291**	**21.0909**	**4.305**
SiCAM 39bands	52.208	24.6488	5.0314
SiCAM full	**62**	**26.929**	**5.497**

SiCAM produced comparable results to that of the iCAM06 model as shown in this boxplot. In the context of reduced spectral bands, SiCAM with the spectral input image of 10 bands gained maximum scores similar to the rating trend of accurate color appearance reproduction of physical HDR scene as depicted from the central red line of the medians of each group as shown in this group. In contrast, it has been noticed that the IQR (represented by box size as mentioned above) is more spread out in this graph presented in Figure 9.5, which tells that the variation in pleasantness scores is higher than the subjective scores for accurate reproduction of color appearance as given in Figure 9.1.1.2. Overall, the rating scores for pleasantness are almost similar for iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full(i.e., 119 spectral bands covering the whole visible spectrum).

The means of the pleasantness scores for the different tone mapping methods in Table 9.2 where Gamma 10bands, Gamma 39bands, and Gamma full exhibit mean pleasantness scores around the range of 30-32, indicating a lower level of pleasantness scores while iCAM06, SiCAM 10bands, SiCAM 39bands, and SiCAM full show much higher mean pleasantness scores, ranging from 57 to 83. These methods seem to yield significantly higher pleasantness ratings compared to the LDR images generated using gamma tone mapping.

The observers with a color science background also gave a similar rating score where the highest scores are given to iCAM06 and SiCAM-Full as illustrated in the boxplot given in Figure 9.4.

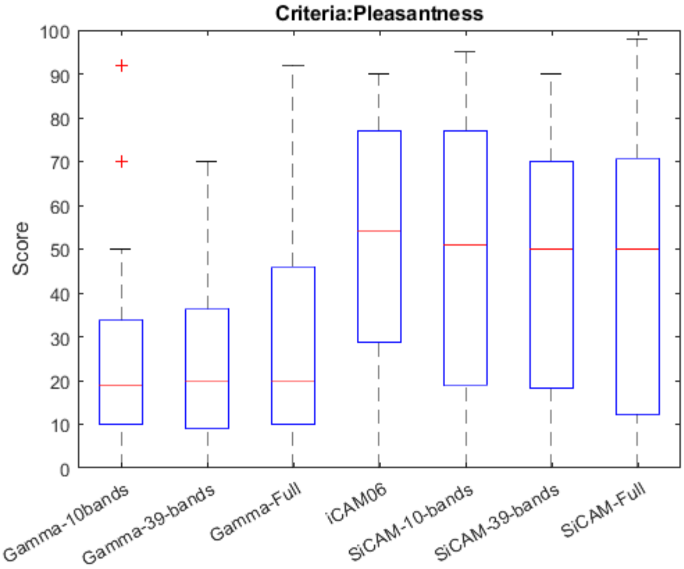


Figure 9.5: *Subjective scores for pleasantness*

Table 9.2: *Summary of Pleasantness Scores*

Method	Mean	Std. Dev	Std. Error
Gamma 10bands	30.16	20.507	4.101
Gamma 39bands	29.76	18.666	3.733
Gamma full	31.583	22.122	4.515
iCAM06	**63.16**	19.349	3.869
SiCAM 10bands	**61.28**	23.055	4.611
SiCAM 39bands	57.20	22.501	4.500
SiCAM full	57.72	**24.354**	4.870

9.1.2 Analysis of Variance (ANOVA)

The ANOVA analysis (St et al., 1989) has been conducted for both the subjective scores of color appearance and pleasantness which are discussed in section 9.1.1 and 9.1.1.2 respectively, to identify and quantify the significance of variability in the subjective scores as given in Table 9.3 and 9.4. In both tables, we conducted variation analyses categorized as "Between Groups" and "Within Groups." Each group corresponds to a specific type of LDR (Low Dynamic Range) image, including Gamma 10-bands, Gamma 39-bands, Gamma Full, iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full. The suffix "Full" indicates images with 119 spectral bands covering the entire visible spectrum, while "n-bands" denotes output generated from reduced spectral bands HDR image as input. It is important to note that iCAM06 and gamma tone mapping had a three-channel HDR image as input while SiCAM operates on the hyperspectral data. The DF (Degrees of Freedom) represents the number of independent pieces of information available for estimating the variability in the data. The SS (Sum of Squares) is a measure of the variability or dispersion of the data whereas the MS (Mean Square) is calculated by dividing the sum of squares by the corresponding degrees of freedom. It provides an estimate of the variance between and within the group. F-Stat (F-Statistic) is the ratio of the variance between groups to the variance within groups. It indicates whether there are significant differences between the group means. Finally, the p-value indicates the probability of observing the F-statistic if there were no significant differences among the groups.

The low p-values i.e., 0.00002 for accurate color appearance reproduction scores and 0.0000 for pleasantness scores indicate highly significant differences between the subjective scores given to these LDR images which were generated using different tone mapping methods. It is to be noted that one-way ANOVA does not identify which group is having the most significant difference. However, it can be deduced from the box plots given in Figure 9.1 and 9.5 as discussed above that a discernible difference is visible between the scores given to LDR images generated using gamma tone mapping and SiCAM whereas the difference between the scores of iCAM06 and SiCAM might not be that significant, especially in the case of pleasantness scores.

Table 9.3: *ANOVA for Color Appearance Rating Scores*

Source	DF	SS	MS	F-Stat	P-Value
Between Groups	6	20954.1141	3492.3523	5.67199	**0.00002**
Within Groups	161	99130.8069	615.7193		

Table 9.4: *ANOVA for Color Pleasantness Rating Scores*

Source	DF	SS	MS	F-Stat	P-Value
Between Groups	63.16	19.349	3.869	13.338	**0.000**
Within Groups	61.28	23.055	4.611		

9.1.3 Rating scores for HDR rendering on HDR display

As mentioned in the methodology of the psychophysical experiment discussed in Chapter 7, the observers were asked to rate the overall image quality of the HDR image based on the real-world HDR screen as the reference. The HDR image was shown on an HDR display using the proposed display referred HDR rendering workflow as explained in Section 4.6. For the HDR image displayed to the observers on an HDR monitor, the rating scores given by the observers are summarized in Figure 9.6. Among these rating scores, the maximum rating is 100% and the minimum rating is 40%. The average of all the rating scores is 68.655% while the standard deviation is 14.125. These statistics reflect that an overall impression of image quality of the HDR image is close to the original HDR scene despite the maximum luminance of the physical HDR scene being approximately 4000nits and the HDR display being limited to only 1000nits of peak luminance.

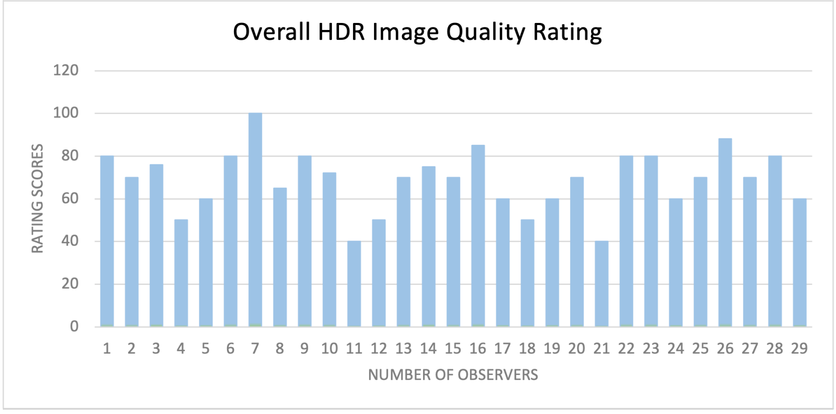


Figure 9.6: *The barplot represents the subjective rating scores given by observers for the overall image quality of the HDR image which has been presented on an HDR display. In this part of the experiment, the original HDR scene has been considered as the reference.*

In addition, the observers were also asked to rate individual attributes of the HDR image displayed on the HDR screen compared to the original HDR scene as a reference. These attributes include colorfulness, sharpness, contrast, and naturalness. The rating scores given by the observers are summarized in the box plot given in Figure 9.7, where the image sharpness obtained the highest rating compared to other attributes while the rating scores for naturalness varied between 65% to 75%. Furthermore, the attribute of colorfulness obtained a slightly higher rating score compared to image contrast, as illustrated in the box plot. Thus, it can be observed that overall perceived the quality of HDR rendering on an HDR display is comparable to that of the original HDR scene.

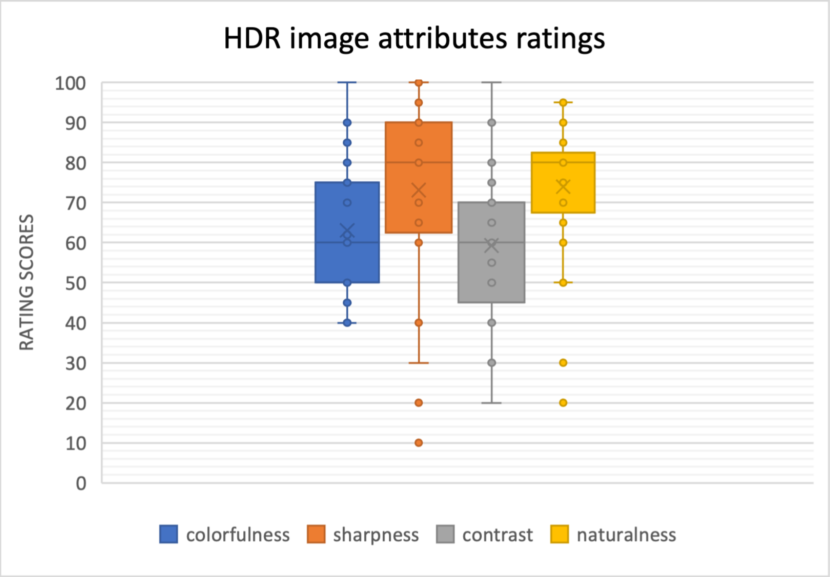


Figure 9.7: *The rating scores revealed that sharpness received the highest ratings compared to other attributes, while ratings for naturalness varied between 65% and 75%. Colorfulness obtained slightly better rating scores than contrast as illustrated in the given plot.*

Lastly, the observers were asked an open-ended question to list down the image quality factors/cues that they were considering while making the judgments. While answering this question, the vocabulary used by the observers is given in Figure 9.8. It can be observed in this plot that almost all observers mentioned brightness/lightness whereas the second most frequently used attribute was color. Among other cues, observers also gave significant importance to saturation, contrast, and perceivable details.

9.2 Objective assessment

The objective image quality assessment (IQA) aims to provide quantitative metrics known as image quality metrics (IQMs) that can predict image quality without the involvement of human observer judgments. The ultimate goal of objective IQA is to produce efficient and consistent judgments which should be correlated with human quality judgments. We have primarily relied on perceptual evaluation by carrying out a psychophysical experiment to investigate the efficacy of using hyperspectral

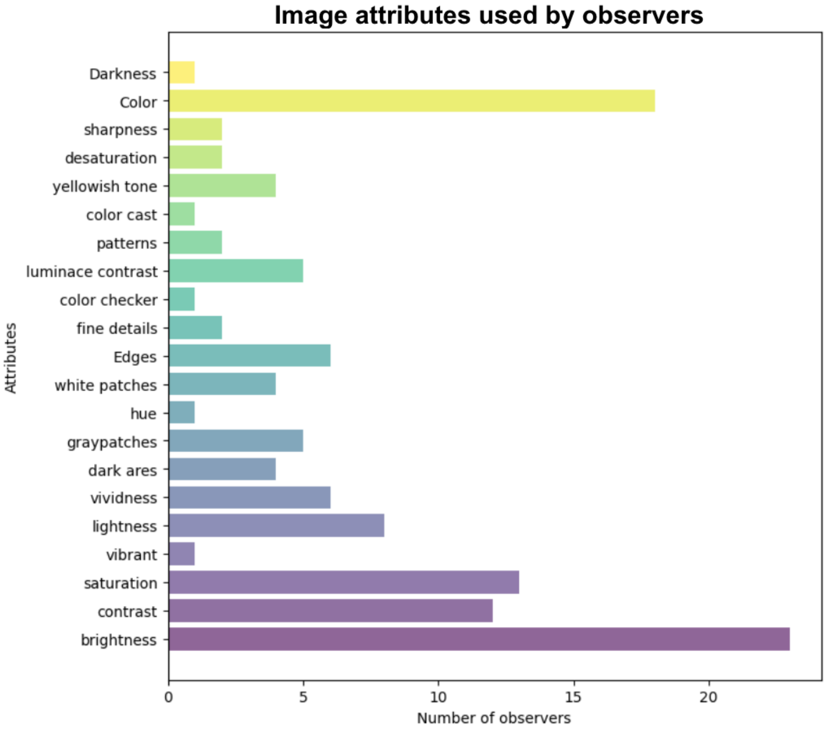
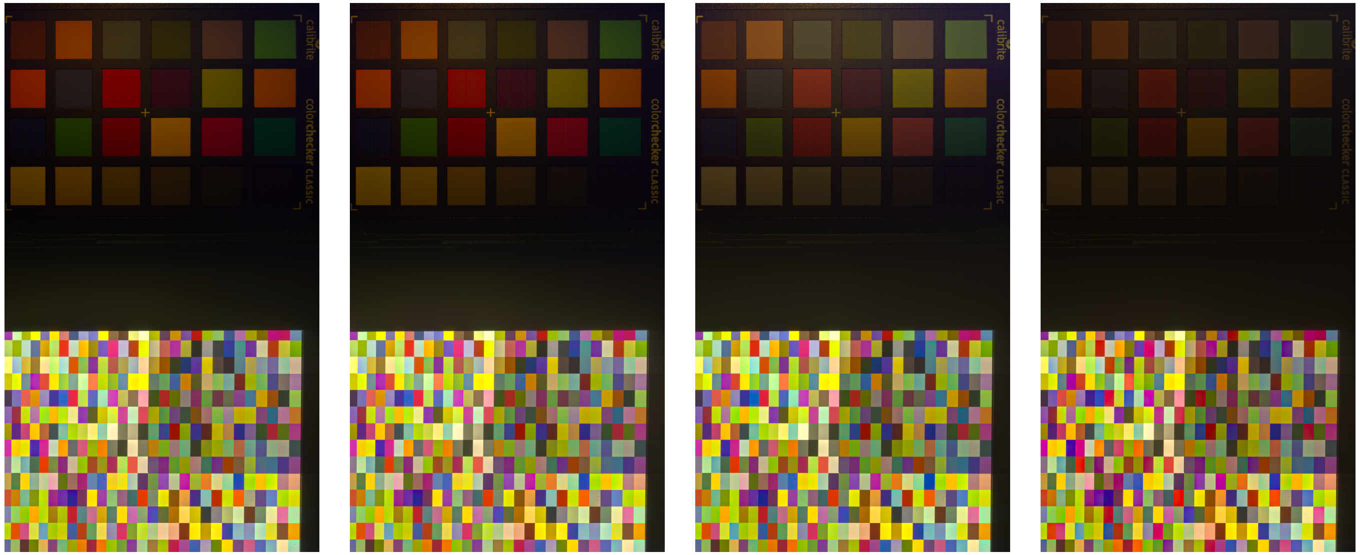
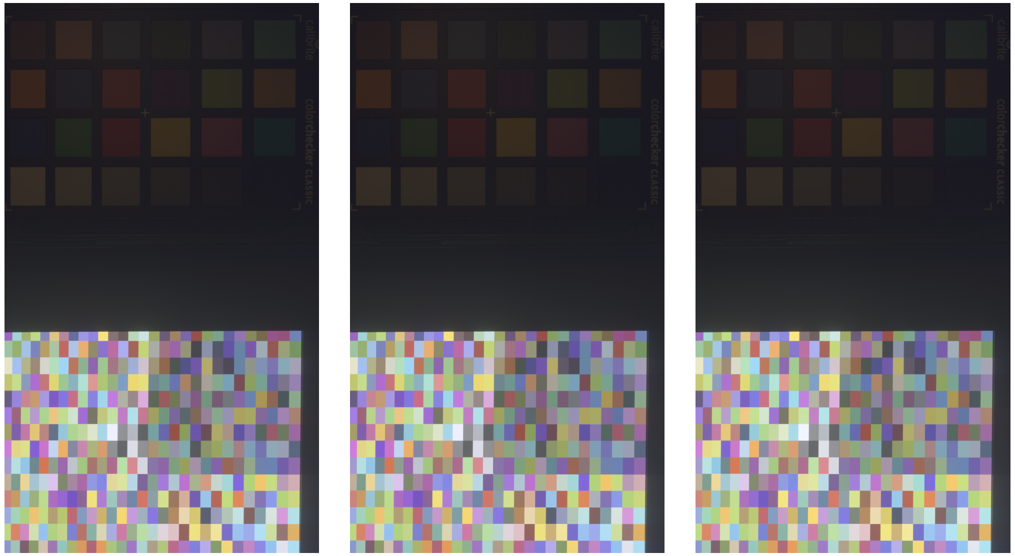


Figure 9.8: *The figure illustrates the image attributes used by observers while making judgments about the LDR tone-mapped images during the psychophysical experiment. Most of the participants said that they considered how bright or light the image was as the most significant attribute while rating the test images. The second most frequently used attribute was the colors of the LDR images. Many participants also talked about other image attributes like how saturated the image was, how much contrast was present, etc.*

data for HDR rendering and tone mapping. Additionally, we have also utilized objective IQA metrics among which two are full reference metrics i.e., Tone-mapped quality index (TMQI) (Yeganeh and Wang, 2012) and Feature similarity index for tone-mapped image (FSITM)(Nafchi et al., 2014). These metrics are designed to evaluate the perceptual quality of tone-mapped high dynamic range images by considering the corresponding HDR image as the reference. Both of these metrics are further discussed in Chapter 5. Besides, the above-mentioned full reference metrics that are fundamentally designed for evaluating the tone mapping algorithms, other blind metrics(i.e., no reference) have also been considered such as the perception-based Image Quality Evaluator (PIQE) (Venkatanath et al., 2015) where a smaller score indicates better perceptual quality. Furthermore, we estimated the scores of the sharpness of the reproduced LDR images and calculated their contrast ratio which are two important factors of image quality.



(a) *SiCAM-full* **(b)** *SiCAM-39bands* **(c)** *SiCAM-10bands* **(d)** *iCAM06*



(e) *Gamma-Full* **(f)** *Gamma-39bands* **(g)** *Gamma-10bands*

Figure 9.9: *Test stimuli are generated using three types of tone mapping methods as listed in 7.1, including Gamma 10-bands, Gamma 39-bands, Gamma Full, iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full. The suffix "Full" indicates images with 119 spectral bands covering the entire visible spectrum, while "n-bands" denotes output generated from reduced spectral bands HDR image as input. It is important to note that iCAM06 had a three-channel HDR image as input, making it distinct from the other methods which used hyperspectral data.*

As mentioned above, FSITM and TMQI have been utilized for evaluating the performance of the proposed SiCAM model with that of iCAM06 and gamma TMO. To calculate these metric scores, the scene referred HDR image has been utilized as a reference (details are given in Section 4.5) whereas the seven reproduced tone-mapped images as given in Table 7.1 were used as test images. The output scores for these metrics are summarized in Figure 9.10. It is clearly visible that the seven tone-mapped images reproduced using different methods have a similar ranking order as we observed in the subjective scores where it has been observed that based on the median of the scores SiCAM-10bands outperformed other tone mapping methods. It is evident from the given graph that these metrics fall short in accurately representing the absolute disparity between the original HDR image and the tone-mapped counterparts. During subjective scores, the gamma tone-mapped images obtained rating scores between 10 to 30 whereas the SiCAM-based LDR images received rating scores between 50 to 87. In contrast, the objective metrics FSITM and TMQI gave almost similar scores to all the test stimuli.

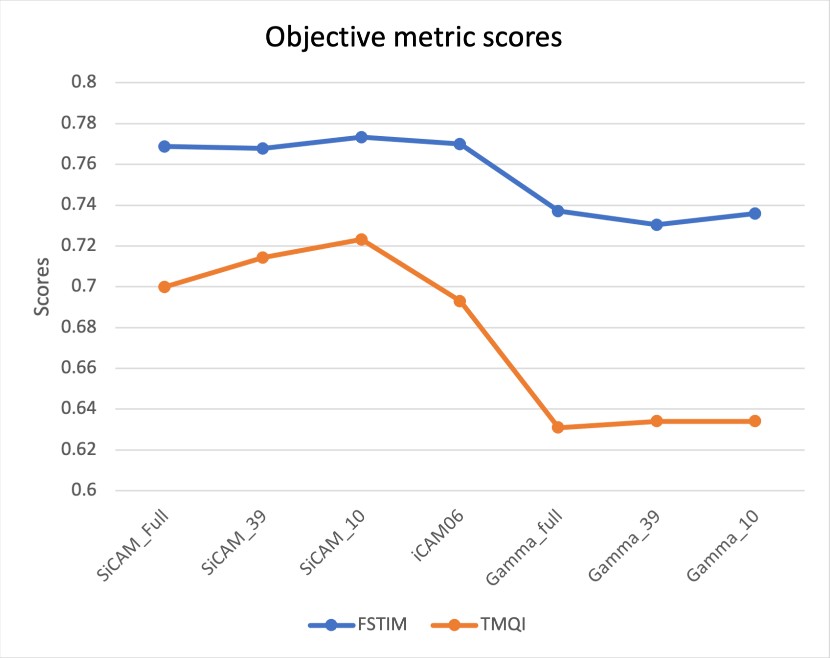


Figure 9.10: *The metric scores of FSITM and TMQI are presented in this graph where HDR image is compared with seven different tone-mapped images reproduced by methods mentioned in Table 7.1 given in Chapter 6.*

We also employed another widely used blind image quality metric named the perception-based Image Quality Evaluator (PIQE) (Venkatanath et al., 2015). Previous research has demonstrated a notable correlation between PIQE scores and

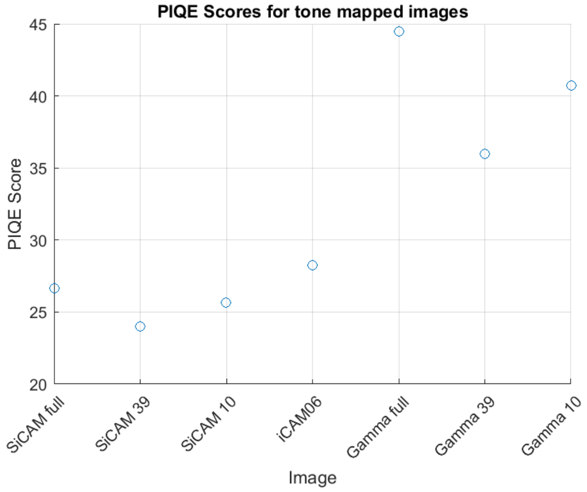


Figure 9.11: *The PIQE scores for tone mapping images are illustrated in the given plot where a smaller score indicates better perceptual quality.*

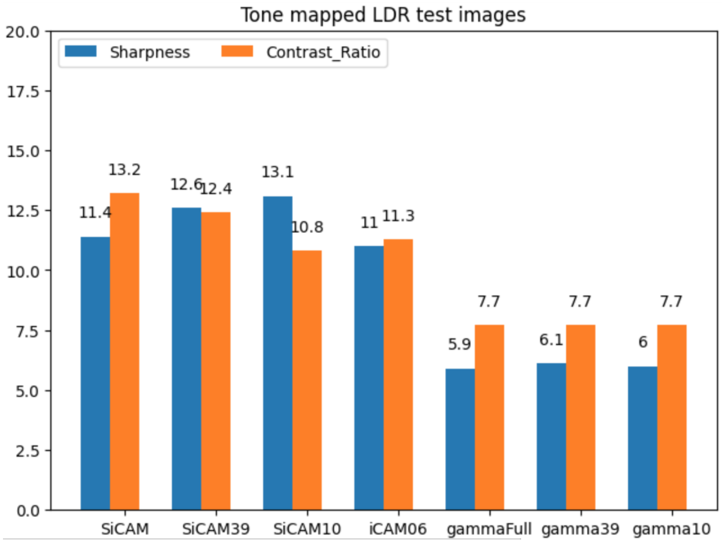


Figure 9.12: *The scores calculated for the sharpness and contrast ratio of tone-mapped LDR images are shown in this graph.*

human subjective scores across diverse databases, including LIVE and TID datasets. Consequently, PIQE proves to be a valuable tool for assessing the perceptual quality of our tone-mapped images. Hence, the PIQE scores are illustrated in Figure 9.11 where a smaller score indicates better perceptual quality. Notably, within this plot, Gamma-10 emerges as the least successful HDR reproduction, while SiCAM-39bands stands out as the image with the highest perceived quality. Besides, sharpness and contrast ratio are also important image quality factors. Hence, the scores were calculated for sharpness and contrast of these LDR reproductions which are given in Figure 9.12. The measurement of image sharpness was performed by quantifying the high-frequency content present in the images using the well-established method based on the normalized gradient magnitudes. By estimating the BT709 luma value for each pixel of LDR images, we quantify the range of luminance variations in an image by calculating the dynamic range by considering the 1 and 99th percentile as lower mean and upper mean values respectively. Lastly as discussed above in Section 9.1.1.1, we analyzed color fidelity by estimating the patch uniformity of the color target (film)in the HDR scene as represented in Figure 9.3 for a better comparison of SiCAM-based LDR images and the output of iCAM06.

9.3 Discussion

In our study, we aimed to assess how the visual quality perception and realism in HDR-rendered images are affected by using HDR hyperspectral data instead of traditional three-channel HDR image. To achieve this, we conducted a psychophysical study, where participants were shown tone mapped LDR images generated using two different approaches: one produced using traditional three-channel HDR data and the other using hyperspectral HDR data. In addition, we also reduced the number of spectral bands to generate some test stimuli for investigating the effect of reducing spectral information. The test images include Gamma 10-bands, Gamma 39-bands, Gamma-Full, iCAM06, SiCAM 10-bands, SiCAM 39-bands, and SiCAM Full as shown in 9.9. The suffix "Full" indicates images with 119 spectral bands covering the entire visible spectrum, while "n-bands" denotes output generated from reduced spectral images as input. It is important to note that iCAM06 and gamma tone mapping had a three-channel HDR image as input while SiCAM operates on hyperspectral data. Further details about how these test stimuli were generated and the methodology of conducting the psychophysical experiment is discussed in Section 7.2.4. It is important to recall that the subjective study was specifically designed to evaluate the HDR tone mapping algorithms, considering both preference and accuracy as evaluation criteria.

Based on the evaluation scores obtained from both subjective and objective assessments as discussed in the previous sections 9.1 and 9.2 respectively, it becomes evident that the SiCAM tone mapping method exhibits better performance. As the subjective assessment is the most reliable method for performance evaluation of novel tone mapping algorithms(Drago et al., 2003; Ledda et al., 2005) and considering that there is no such computational model which is designed for perceptual evaluation of HDR rendering with real-world HDR scenes, we primarily relied on the perceptual quality scores given by the observers to investigate the efficacy of using hyperspectral data for HDR rendering and tone mapping.

Hence it can be concluded from the data analysis of the subjective scores given by all the observers to LDR test images that SiCAM-10bands is deemed as the optimal LDR reproduction of the physical HDR scene while the participants with a color science background favored SiCAM-Full as the best LDR image among the test samples as illustrated in Figure 9.4. It can be observed from the box plots given in Figure 9.1 and 9.5 as discussed above that a discernible difference is visible between the scores given to LDR images generated using gamma tone mapping and SiCAM whereas the difference between the scores of iCAM06 and SiCAM is less pronounced, especially in the case of pleasantness scores. It is worth mentioning that the observers were shown one test image at a time during the experiment. It might be possible that the nonuniform target patches(transmissive film) and a slight overall yellowish color cast on the output of iCAM06 as depicted in Figure 9.2, were not that visible to participants affecting the rating of their judgments. Therefore, for a better comparison of SiCAM and iCAM06, we analyzed color fidelity by estimating the patch uniformity of the target patches in our physical HDR scene as where it can be observed that the output LDR images generated by SiCAM are more uniform than iCAM06 as illustrated in Figure 9.2.

These findings suggest that the proposed SiCAM model which utilizes hyperspectral information generated better tone-mapped images. To reiterate, the proposed SiCAM model is an extension of iCAM06 (Kuang et al., 2007b) where we introduced spectral radiances so that the processing pipeline can benefit from the additional information in spectral radiances in comparison to a three-channel HDR image. The detailed explanation of this model is presented in Chapter 5 Section 5.3 and the full workflow of SiCAM is given in Figure 5.1. It should be noted that, while a few modules in this model are identical to that in iCAM06(Kuang et al., 2007b), the SiCAM model is already significantly different since it uses hyperspectral data as input rather than RGB image. Another major difference is that we incorporated the proposed hyperspectral adaptation in the SiCAM model instead of the chromatic adaptation as in the case of iCAM06.

The above-mentioned results support the argument that despite the fact that trichromatic vision serves as the first stage of color vision processing, it might be possible that the higher levels of the visual system have some access to spectral information which may complement trichromatic mechanisms. If that is assumed to be the case then utilizing HDR hyperspectral data instead of traditional three-channel HDR image in tone mapping and other HDR rendering workflow should be deemed to improve visual quality perception and realism in HDR-rendered images. Hence, the answer to our initial research question is "yes" that using HDR hyperspectral data instead of traditional three-channel HDR image positively impact the visual quality perception and realism in HDR-rendered images as the SiCAM outperformed. In the realm of human vision, we can find such supporting shreds of evidence which challenge the underlying working principles of colorimetry that are solely based on the assumption that humans have trichromatic vision (Wyszecki and Stiles, 2000). For instance, the cone photoreceptors detect light that passes through the macula's yellow filter in the foveal vision, whereas there is no macula in the periphery (Fairchild, 2013). Therefore, the human visual system seems to acquire the six spectral channels considering the same scene has been viewed with the foveal and extra-foveal retina which aids spectral adaptation and human vision color constancy. Other examples involve the color appearance models which suggest that at least five dimensions are necessary to consider in order to model the appearance of a uniform color patch (Fairchild and Johnson, 2002) which include brightness, hue, lightness, colorfulness, and chroma. In addition, it has also been found in past research studies that there is a difference in the spectral sensitivity (Abramov and Gordon, 1977) and shapes of the cones which lie within the fovea and those which are in the periphery while the two eyes do not have a similar preretinal absorption. Additionally, the existence of both 2° and 10° observers with different CMFs which are not linear transforms of each other also hints at the spectral approach in HVS (Liu and Fairchild, 2006). All the above-mentioned arguments indicate that human vision utilizes spectral information at a certain level of visual processing and is not completely trichromatic.

Another possible explanation for the SiCAM generating better results is that even though the final output of both HDR and LDR tone-mapped reproduction of real-world HDR scene is the three-channel image, the estimation of these RGB images is more optimal if the spectral information has been utilized in comparison to methods that solely rely on trichromatic information. It has been found that the improvement in the output LDR images of the SiCAM model have emerged by incorporating the hyperspectral adaption module where the fine-grained spectral information enables the SiCAM model to capture more uniform color patches.

It is necessary to acknowledge that the idea of utilizing hyperspectral information for HDR rendering is currently in the infancy stage of research for practical application, necessitating additional investigation and evaluation to ascertain its suitability in advanced imaging algorithms. However, this research serves as the foundational basis for further investigations.

9.4 Limitation and Future Work

The limitations of this research study are highlighted below which can be addressed in future work.

- Due to the limitation of the imaging device, we were confined to only capturing the top view of flat 2D HDR scenes. Additionally, only one HDR real-world scene has been utilized during the perceptual evaluation of the proposed HDR rendering method whereas multiple real-world HDR scenes would have resulted in more reliable analysis.
- In this study, we have compared the performance of the proposed HDR rendering method with iCAM06 and gamma TMO but in future studies, we would be interested in utilizing a large number of state-of-the-art tone mapping algorithms for a more extensive evaluation of the proposed SiCAM.
- As mentioned above, the SiCAM is currently in the infancy stage of research for practical application, therefore, this model can be improved further by incorporating other adjustments such as color saturation compensation, and/or employing a more robust tone mapping curve, etc.
- Additionally, the hyperspectral band selection method can be embedded in the hyperspectral HDR image rendering that would be intended to select a small subset of hyperspectral bands in order to eliminate spectral redundancy and reduce computational costs while preserving the significant spectral information of the HDR scene.
- Since the psychophysical experiment was conducted in just one day due to the limited availability of HDR display, the experiment methodology could be improved by setting up a more detailed experiment with a higher number of observers.
- The objective metrics that are utilized in our work failed to generate quality scores that could correlate with subjective ratings given by the observers. Hence, more sophisticated methods for HDR image quality assessment can be investigated in this context or other LDR-based image quality metrics can be adapted for evaluating tone-mapped images in future work.
- The limited research work in this domain of merging HDR and hyperspectral has been a challenge in determining the formulation of our research methodology.
- The proposed display referred workflow can be further compared with other HDR standards for evaluating its suitability for practical application.
- The proposed methodology for acquiring HDR hyperspectral radiances has the potential to be extended for other hyperspectral cameras by incorporating the method for radiance estimation.

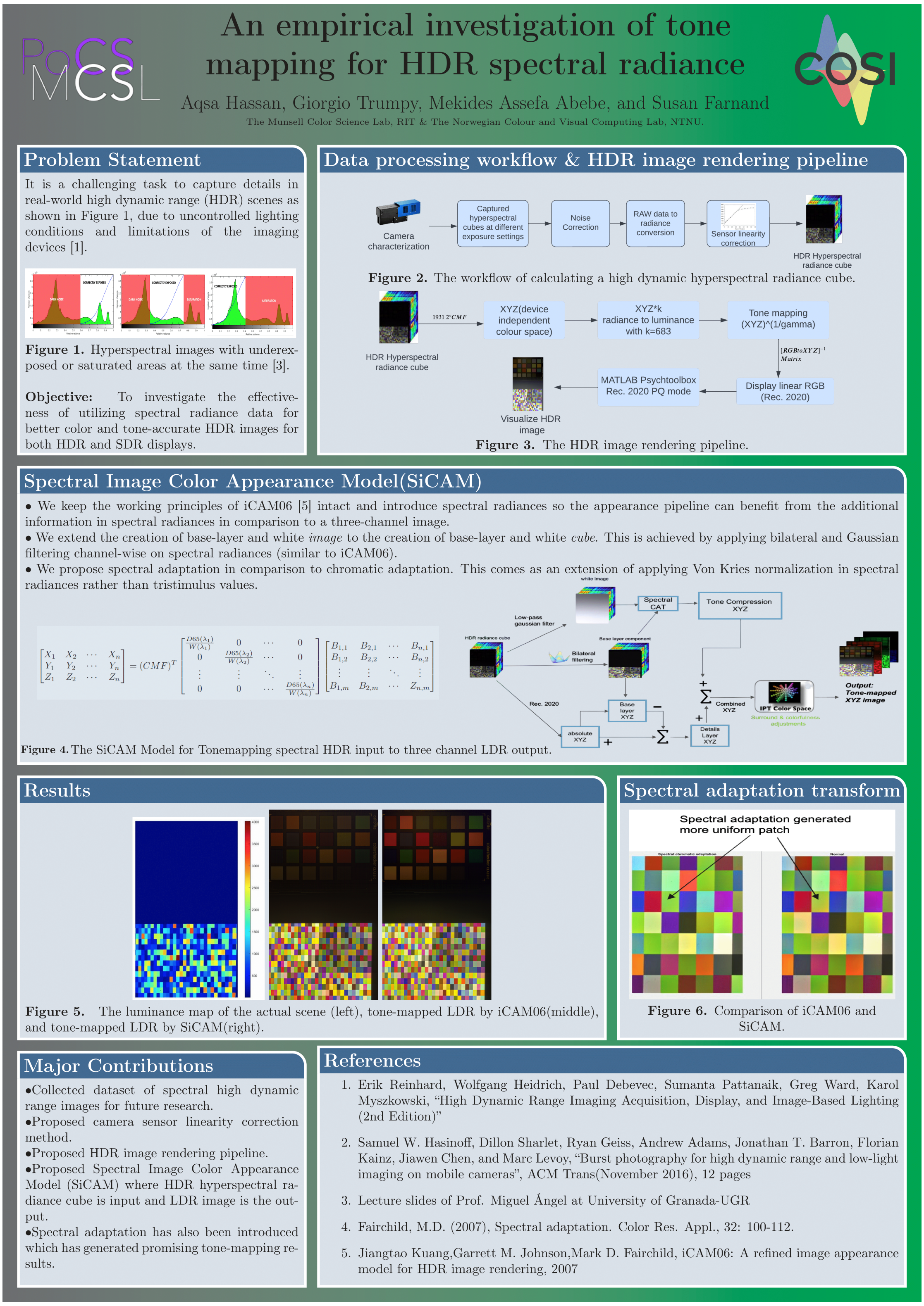
9.5 Conclusion

This work focuses on investigating the effective utilization of spectral radiance for improved color fidelity and tone-accurate reproduction of HDR images. It is difficult to retrieve the details in HDR hyperspectral images where both dark areas and bright areas are present which results in underexposed and overexposed regions. It is due to arbitrary illumination and imaging device constraints, such as limited dynamic range. To address these issues, high dynamic range imaging techniques can be applied which help to significantly recover the details of a high dynamic range (HDR) scene. We commence our study by capturing the HDR hyperspectral radiance cube of a physical HDR scene using an approach similar to the multiple exposures HDR technique. Furthermore, the HDR hyperspectral absolute radiance image was enhanced by employing the proposed linearity correction method and hyperspectral interpolation. The linearity correction was applied for enhancing the linearity of the sensor response to recover the radiances in its upper non-linear region whereas hyperspectral interpolation was employed to check for missing wavelengths due to saturated pixels. Additionally, we proposed the workflow for converting the HDR hyperspectral radiance cube into three channel HDR image. Considering the lack of a widely accessible standard pipeline for rendering HDR content of arbitrary luminance ranges, we proposed a display-referred HDR rendering workflow that effectively translates HDR images into the specific range of luminance supported by the HDR display

A spectral image color appearance model titled SiCAM has been introduced which is the first of its kind, designed for HDR hyperspectral radiance images. Although SiCAM is inspired by the latest iCAM06 HDR image rendering model, it is already significantly different since it uses hyperspectral data as input rather than RGB HDR image. The SiCAM incorporated the hyperspectral adaption transform for generating tone-accurate LDR images from the input hyperspectral HDR radiances image which is another major contribution of this work. We kept the working principles of iCAM06 intact and introduced spectral radiances so the pipeline can benefit from hyperspectral HDR data in comparison to a three-channel HDR image.

Besides, the image quality assessment of the reproduced LDR images using SiCAM has been conducted in comparison to iCAM06 and gamma tone mapping to determine the effectiveness of having hyperspectral data through objective and subjective study. However, we primarily relied on the perceptual quality scores given by the observers to investigate the efficacy of using hyperspectral data for HDR rendering and tone mapping. The findings suggest that the proposed SiCAM model which utilizes hyperspectral information generated better tone-mapped images. It has been observed that SiCAM-10bands generated the optimal LDR reproduction of the physical HDR scene while the participants with a color science background favored SiCAM-Full as the best LDR image among the test samples generated using different tone mapping methods. Consequently, it can be argued that although trichromatic vision serves as the first stage of color vision processing, it might be possible that the higher levels of the visual system have some access to spectral information which may complement trichromatic mechanisms. It is essential to realize that simply having hyperspectral data to improve HDR rendering is not sufficient. Equally important is how well the processing method can leverage this hyperspectral information to produce better tone-mapped images. Given the fact that no HDR hyperspectral radiance data is currently available, a dataset containing four HDR hyperspectral radiances cubes and their respective three-channel HDR images have been proposed where each of them consisting of a different HDR scene captured in an indoor setting.

A Appendix



An empirical investigation of tone mapping for HDR spectral radiance

Aqsa Hassan, Giorgio Trumpy, Mekides Assefa Abebe, and Susan Farnand

The Munsell Color Science Lab, RIT & The Norwegian Colour and Visual Computing Lab, NTNU.

Problem Statement

It is a challenging task to capture details in real-world high dynamic range (HDR) scenes as shown in Figure 1, due to uncontrolled lighting conditions and limitations of the imaging devices [1].

Figure 1. Hyperspectral images with underexposed or saturated areas at the same time [3].

Objective: To investigate the effectiveness of utilizing spectral radiance data for better color and tone-accurate HDR images for both HDR and SDR displays.

Data processing workflow & HDR image rendering pipeline

Figure 2. The workflow of calculating a high dynamic hyperspectral radiance cube.

Figure 3. The HDR image rendering pipeline.

Spectral Image Color Appearance Model(SiCAM)

- We keep the working principles of iCAM06 [5] intact and introduce spectral radiances so the appearance pipeline can benefit from the additional information in spectral radiances in comparison to a three-channel image.
- We extend the creation of base-layer and white *image* to the creation of base-layer and white *cube*. This is achieved by applying bilateral and Gaussian filtering channel-wise on spectral radiances (similar to iCAM06).
- We propose spectral adaptation in comparison to chromatic adaptation. This comes as an extension of applying Von Kries normalization in spectral radiances rather than tristimulus values.

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{bmatrix} = (CMF)^T \begin{bmatrix} \frac{D65(\lambda_1)}{W(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \frac{D65(\lambda_2)}{W(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{D65(\lambda_n)}{W(\lambda_n)} \end{bmatrix} \begin{bmatrix} B_{1,1} & B_{2,1} & \cdots & B_{n,1} \\ B_{1,2} & B_{2,2} & \cdots & B_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1,m} & B_{2,m} & \cdots & Z_{n,m} \end{bmatrix}$$

Figure 4. The SiCAM Model for Tonemapping spectral HDR input to three channel LDR output.

Results

Figure 5. The luminance map of the actual scene (left), tone-mapped LDR by iCAM06(middle), and tone-mapped LDR by SiCAM(right).

Spectral adaptation transform

Figure 6. Comparison of iCAM06 and SiCAM.

Major Contributions

- Collected dataset of spectral high dynamic range images for future research.
- Proposed camera sensor linearity correction method.
- Proposed HDR image rendering pipeline.
- Proposed Spectral Image Color Appearance Model (SiCAM) where HDR hyperspectral radiance cube is input and LDR image is the output.
- Spectral adaptation has also been introduced which has generated promising tone-mapping results.

References

1. Erik Reinhard, Wolfgang Heidrich, Paul Debevec, Sumanta Pattanaik, Greg Ward, Karol Myszkowski, "High Dynamic Range Imaging Acquisition, Display, and Image-Based Lighting (2nd Edition)"
2. Samuel W. Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T. Barron, Florian Kainz, Jiawen Chen, and Marc Levoy, "Burst photography for high dynamic range and low-light imaging on mobile cameras", ACM Trans(November 2016), 12 pages
3. Lecture slides of Prof. Miguel Ángel at University of Granada-UGR
4. Fairchild, M.D. (2007), Spectral adaptation. Color Res. Appl., 32: 100-112.
5. Jiangtao Kuang,Garrett M. Johnson,Mark D. Fairchild, iCAM06: A refined image appearance model for HDR image rendering, 2007

Figure A.1: *The poster presented at MCSL open house session.*

NTNU

RIT | Rochester Institute of Technology of Dubai

An empirical investigation of tonemapping for HDR spectral radiance

INFORMATION SHEET

Researcher(s) Introduction
Aqsa Hassan, Master Student, NTNU

Project Description and Invitation
The project is finding out the benefits that one can get in tonemapping by having more spectral information of the scene.

Participant Identification and Recruitment
Participants must have no visual impairment that cannot be corrected by glasses or contact lenses. Participants must not be colour blind. Participants can participate anonymously.

Project Procedures
Participants will first do an Ishihara test and then they will do the experiment.

Data Management
The collected data will serve research purposes only and will remain strictly confidential.

Participant's Rights
You are under no obligation to accept this invitation. If you decide to participate, you have the right to:

- *decline to answer any particular question;*
- *withdraw from the study (any time).*
- *ask any questions about the study at any time during participation.*
- *provide information on the understanding that your name will not be used unless you give permission to the researcher.*
- *be given access to a summary of the project findings when it is concluded.*

This project has been evaluated by peer review and judged to be ***low risk****. The researchers named in this document are responsible for the ethical conduct of this research. If you have any concerns about the conduct of this research that you want to raise with someone other than the researchers, please contact NTNU's Research Ethics Committee.*

Completion Date 09-Jun-2023
Expiration Date 09-Jun-2026
Record ID 56408405

This is to certify that:

aqsa Hassan

Has completed the following CITI Program course:

Not valid for renewal of certification through CME.

Social & Behavioral Research - Basic/Refresher
(Curriculum Group)
Social & Behavioral Research - Basic/Refresher
(Course Learner Group)
1 - Basic Course
(Stage)

Under requirements set by:

Rochester Institute of Technology

CITI
Collaborative Institutional Training Initiative

101 NE 3rd Avenue, Suite 320
Fort Lauderdale, FL 33301 US
www.citiprogram.org

Verify at www.citiprogram.org/verify/?w314a9f92-37ec-4cae-b9ad-09c5f6eecba3-56408405

Figure A.2: *The CITI program completion certificate was required before conducting a psychophysical study involving human observers.*

Bibliography

Abramov, I. and Gordon, J. (1977). Color vision in the peripheral retina. i. spectral sensitivity. *JOSA*, 67(2):195–202. (cited on page 144)

Aggarwal, M. and Ahuja, N. (2001). High dynamic range panoramic imaging. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 2–9. IEEE. (cited on page 33)

Anand Swamy, A., Mamatha, A., Shylashree, N., and Nath, V. (2022). Lossless compression of hyperspectral imagery by assimilating decorrelation and preprocessing with efficient displaying using multiscale hdr approach. *IETE Journal of Research*, pages 1–12. (cited on page 4)

Asano, Y., Fairchild, M., and Blondé, L. (2016). Individual colorimetric observer model. *PLOS ONE*, 11(2):0145671. (cited on page 15)

Audebert, N., Le Saux, B., and Lefèvre, S. (2019). Deep learning for classification of hyperspectral data: A comparative review. *IEEE geoscience and remote sensing magazine*, 7(2):159–173. (cited on page 4)

Aydın, T. O., Mantiuk, R., and Seidel, H.-P. (2008). Extending quality metrics to full luminance range images. In *Human vision and electronic imaging xiii*, volume 6806, pages 109–118. SPIE. (cited on page 123)

Banterle, F., Artusi, A., Debattista, K., and Chalmers, A. (2017). *Advanced high dynamic range imaging*. CRC press. (cited on pages 5, 13, 14, 70, 72, 74, and 167)

Bergquist, J. (2008). 52.2: Display with arbitrary primary spectra. *Sid Symposium Digest of Technical Papers*, 39. (cited on page 15)

Berns, R. S. (1996). Methods for characterizing crt displays. *Displays*, 16(4):173–182. (cited on pages 64, 99, and 101)

Bigas, M., Cabruja, E., Forest, J., and Salvi, J. (2006). Review of cmos image sensors. *Microelectronics journal*, 37:433–451. (cited on page 28)

Boitard, R., Mantiuk, R. K., and Pouli, T. (2015). Evaluation of color encodings for high dynamic range pixels. In *Human Vision and Electronic Imaging XX*, volume 9394, pages 532–540. SPIE. (cited on page 105)

Bouderbane, M., Lapray, P.-J., Dubois, J., Heyrman, B., and Ginhac, D. (2016). Real-time ghost free hdr video stream generation using weight adaptation based method. In *Proceedings of the 10th International Conference on Distributed Smart Camera*, pages 116–120. (cited on page 58)

Boult, T. E. (1993). G-neighbors. *Vision Geometry*, II. Vol. 2060. (cited on page 78)

Brajovic, V. and Kanade, T. (1996). A sorting image sensor: An example of massively parallel intensity-to-time processing for low-latency computational sensors. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 2, pages 1638–1643. IEEE. (cited on page 33)

BT, I. R. (2002). Parameter values for the hdtv standards for production and international programme exchange. *International Telecommunication Union, Recommendation, May.* (cited on pages 100, 101, and 175)

BT.1129-2, I.-R. (1998). Subjective assessment of standard definition digital television (sdtv) systems. *ITU.* (cited on pages 108 and 109)

BT.1361, I.-R. (1998). Worldwide unified colorimetry and related characteristics of future television and imaging systems. *ITU.* (cited on page 108)

BT.500-11, I.-R. (2002). *Methodology for the subjective assessment of the quality of television pictures.* ITU, Geneva, Switzerland. (cited on pages 108 and 109)

BT.710-4, I.-R. (1998). *Subjective assessment methods for image quality in high-definition television.* ITU, Geneva, Switzerland. (cited on page 108)

BT.814-1, I.-R. (1994). Specification and alignment procedures for setting of brightness and contrast of displays. *ITU.* (cited on pages 108 and 109)

BT.815-1, I.-R. (1994). Specification of a signal for measurement of the contrast ratio of displays. *ITU.* (cited on page 108)

Burns, S. A. (2019). Chromatic adaptation transform by spectral reconstruction. *Color Research and Application*, 44(5):682–693. (cited on page 82)

Burt, P. J. and Kolczynski, R. J. (1993). Enhanced image capture through fusion. In *1993 (4th) international Conference on Computer Vision*, pages 173–182. IEEE. (cited on pages 3 and 33)

Chin, R. T. and Yeh, C.-L. (1983). Quantitative evaluation of some edge-preserving noise-smoothing techniques. *Computer vision, graphics, and image processing*, 23:67–91. (cited on page 78)

Chong, H. Y., Gortler, S. J., and Zickler, T. (2007). The von kries hypothesis and a basis for color constancy. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE. (cited on page 77)

Debevec, P. E. and Malik, J. (2008). Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, pages 1–10. unknow. (cited on page 33)

Derhak, M. W., Luo, E. L., and Green, P. J. (2020). Fast chromatic adaptation transform utilizing wpt (waypoint) based spectral reconstruction. In *London Imaging Meeting 2020: Future Colour Imaging*. Society for Imaging Science and Technology. (cited on page 82)

Drago, F., Martens, W. L., Myszkowski, K., and Seidel, H.-P. (2003). Perceptual evaluation of tone mapping operators. In *ACM SIGGRAPH 2003 Sketches and Applications*, pages 1–1. ACM SIGGRAPH. (cited on pages 111 and 142)

Durand, F. and Dorsey, J. (2002a). Fast bilateral filtering for the display of high-dynamic-range images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 257–266. (cited on pages 75 and 78)

Durand, F. and Dorsey, J. (2002b). Fast bilateral filtering for the display of highdynamic-range image. In *Proceedings of ACM SIGGRAPH 2002, Computer Graphics Proceedings, Annual Conference Proceedings*, page 257–266. (cited on page 77)

Edition, F., Papoulis, A., and Pillai, S. (2002). *Probability, random variables, and stochastic processes.* unkown. (cited on page 78)

Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R. K., and Unger, J. (2017a). Hdr image reconstruction from a single exposure using deep cnns. *ACM transactions on graphics (TOG)*, 36(6):1–15. (cited on page 34)

Eilertsen, G., Mantiuk, R. K., and Unger, J. (2017b). A comparative review of tone-mapping algorithms for high dynamic range video. In *Computer graphics forum*, volume 36, pages 565–592. Wiley Online Library. (cited on pages 2, 5, 70, and 74)

Eilertsen, G., Unger, J., and Mantiuk, R. K. (2016). Evaluation of tone mapping operators for hdr video. In *High dynamic range video*, pages 185–207. Elsevier. (cited on page 72)

Ertürk, S., Süer, S., and Koç, H. (2014). A high-dynamic-range-based approach for the display of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 11(11):2001–2004. (cited on page 4)

Fairchild, M. and Johnson, G. (2002). Meet icam: a next-generation color appearance model. In *IST/SID 10th Color Imaging Conference*, page 33–38. (cited on pages 76, 77, and 144)

Fairchild, M. D. (2007a). The hdr photographic survey. In *Color and imaging conference*, volume 15, pages 233–238. Society of Imaging Science and Technology. (cited on page 124)

Fairchild, M. D. (2007b). Spectral adaptation." color research and application: Endorsed by inter-society color council, the colour group (great britain), canadian society for color, color science association of japan, dutch society for the study of color, the swedish colour centre foundation. *Colour Society of Australia, Centre Français de la Couleur*, 32(2):100–112. (cited on pages 4, 81, 82, and 83)

Fairchild, M. D. (2013). *Color appearance models*. John Wiley and Sons. (cited on pages 9, 18, 21, 23, 76, 144, and 168)

Farbman, Z., Fattal, R., Lischinski, D., and Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM transactions on graphics (TOG)*, 27(3):1–10. (cited on page 76)

Fattal, R., Lischinski, D., and Werman, M. (2002). Gradient domain high dynamic range compression. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 249–256. (cited on page 76)

Ferwerda, J. A., Pattanaik, S. N., Shirley, P., and Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 249–258. (cited on page 1)

Froehlich, J., Kunkel, T., Atkins, R., Pytlarz, J., Daly, S., Schilling, A., and Eberhardt, B. (2015). Encoding color difference signals for high dynamic range and wide gamut imagery. In *Color and Imaging Conference*, volume 2015, pages 240–247. Society for Imaging Science and Technology. (cited on page 64)

Gasparini, F. and Schettini, R. (2003). Color correction for digital photographs. In *12th International Conference on Image Analysis and Processing, 2003. Proceedings.*, pages 646–651. IEEE. (cited on page 27)

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., and Buckler, E. S. (2014). Tassel-gbs: a high capacity genotyping by sequencing analysis pipeline. *PloS one*, 9(2):e90346. (cited on page 4)

Graham, R. (1962). Snow removal–a noise-stripping process for picture signals. *IRE Transactions on Information Theory*, 8(2):129–144. (cited on page 78)

Green, P. and MacDonald, L. (2011). *Colour engineering: achieving device independent colour.* John Wiley and Sons. (cited on pages 93, 94, 95, 100, and 101)

Hanhart, P., Bernardo, M. V., Pereira, M., G Pinheiro, A. M., and Ebrahimi, T. (2015). Benchmarking of objective quality metrics for hdr image quality assessment. *EURASIP Journal on Image and Video Processing*, 2015(1):1–18. (cited on page 123)

Hill, B., Roger, T., and Vorhagen, F. W. (1997). Comparative analysis of the quantization of color spaces on the basis of the cielab color-difference formula. *ACM Transactions on Graphics (TOG)*, 16(2):109–154. (cited on page 23)

Himayat, N. and Kassam, S. A. (1993). Approximate performance analysis of edge preserving filters. *IEEE Transactions on Signal Processing*, 41(9):2764–2777. (cited on page 78)

Hoefflinger, B. (2007). *High-dynamic-range (HDR) vision.* Springer. (cited on pages 1 and 29)

Huang, T., Yang, G., and Tang, G. (1979). A fast two-dimensional median filtering algorithm. *IEEE transactions on acoustics, speech, and signal processing*, 27(1):13–18. (cited on page 78)

Hung, P.-C. (2019). 61-3: Invited paper: Cie activities on wide colour gamut and high dynamic range imaging. *SID Symp Dig Tech Pap*, 50(1):866–869. (cited on page 15)

Hunt, R. W. G. (1952). Light and dark adaptation and the perception of color. *JOSA*, 42(3):190–199. (cited on page 35)

Hunt, R. W. G. (2005). *The reproduction of colour.* John Wiley and Sons. (cited on page 85)

Hunt, R. W. G. and Pointer, M. R. (2011). *Measuring colour.* John Wiley & Sons. (cited on pages 9, 11, and 15)

ISO (2023). ISO 12646:2008. [Online; accessed 12. Jul. 2023]. (cited on pages 97 and 98)

Johnson, G. and Fairchild, M. (2003). Rendering hdr images. In *IST SID 11th Color Imaging Conference*, page 36–41, Scottsdale. (cited on page 77)

Jones, B. L. and McManus, P. R. (1986). Graphic scaling of qualitative terms. *SMPTE Journal*, 95(11):1166–1171. (cited on page 110)

Khan, H. A., Thomas, J. B., and Hardeberg, J. Y. (June 12–14, 2017). Multispectral constancy based on spectral adaptation transform. In *Image Analysis: 20th Scandinavian Conference, SCIA 2017*, Tromsø, Norway. Springer International Publishing. (cited on pages 4 and 82)

Khan, H. A., Thomas, J.-B., Hardeberg, J. Y., and Laligant, O. (2018). Spectral adaptation transform for multispectral constancy. *Journal of Imaging Science and Technology*, 62(2):205041–2050412. (cited on page 4)

Kim, K., Bae, J., and Kim, J. (2011). Natural hdr image tone mapping based on retinex. *IEEE Transactions on Consumer Electronics*, 57(4):1807–1814. (cited on page 76)

Kirchner, E., Dekker, N., Lucassen, M., Njo, L., van der Lans, I., Urban, P., and Huertas, R. (2015). How psychophysical methods influence optimizations of color difference formulas. *JOSA A*, 32(3):357–366. (cited on page 23)

Korshunov, P., Hanhart, P., Richter, T., Artusi, A., Mantiuk, R., and Ebrahimi, T. (2015). Subjective quality assessment database of hdr images compressed with jpeg xt. In *2015 seventh international workshop on quality of multimedia experience (QoMEX)*, pages 1–6. IEEE. (cited on page 124)

Krasula, L., Narwaria, M., Fliegel, K., and Le Callet, P. (2015). Influence of hdr reference on observers preference in tone-mapped images evaluation. In *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE. (cited on page 124)

Kuang, J., Johnson, G. M., and Fairchild, M. D. (2007a). icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation*, 18(5):406–414. Special issue on High Dynamic Range Imaging. (cited on pages 7, 77, 78, 80, 83, 84, 85, 86, and 171)

Kuang, J., Yamaguchi, H., Liu, C., Johnson, G., and Fairchild, M. (2007b). Evaluating hdr rendering algorithms. *ACM Trans. Appl. Percept.* (cited on pages 70, 77, 85, 143, and 171)

Kubik, M. (2007). Hyperspectral imaging: a new technique for the non-invasive study of artworks. In *Physical techniques in the study of art, archaeology and cultural heritage*, volume 2, pages 199–259. Elsevier. (cited on page 4)

Kundu, M. K. and Pal, S. K. (1986). Thresholding for edge detection using human psychovisual phenomena. *Pattern Recognition Letters*, 4(6):433–441. (cited on page 123)

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1):112–147. (cited on page 104)

Land, E. and McCann, J. (1971). Lightness and retinex theory, journal of the optical society of america. *JOlunal of the Optical Society of America*, 61(1). (cited on page 76)

Lapray, P.-J., Thomas, J.-B., and Gouton, P. (2017). High dynamic range spectral imaging pipeline for multispectral filter array cameras. *Sensors*, 17(6):1281. (cited on page 4)

Ledda, P., Chalmers, A., Troscianko, T., and Seetzen, H. (2005). Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics (TOG)*, 24(3):640–648. (cited on pages 111 and 142)

Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, 2:165–168. (cited on page 78)

Li, C., Li, Z., Wang, Z., Xu, Y., Luo, M. R., Cui, G., Melgosa, M., Brill, M. H., and Pointer, M. (2017). Comprehensive color solutions: Cam16, cat16, and cam16-ucs. *Color Research & Application*, 42:703–718. (cited on page 23)

Li, X., Gunturk, B., and Zhang, L. (2008). Image demosaicing: A systematic survey. In *Visual Communications and Image Processing 2008*, volume 6822, pages 489–503. SPIE. (cited on page 27)

Linhares, J. M. M., Pinto, P. D., and Nascimento, S. M. C. (2008). The number of discernible colors in natural scenes. *JOSA A*, 25:2918–2924. (cited on page 4)

Liu, C. and Fairchild, M. D. (2006). The surround color and color matching functions. In *Proc. IST 14th Color and Imaging Conf*, page 203 – 208. (cited on pages 83 and 144)

Liu, Y.-L., Lai, W.-S., Chen, Y.-S., Kao, Y.-L., Yang, M.-H., Chuang, Y.-Y., and Huang, J.-B. (2020). Single-image hdr reconstruction by learning to reverse the camera pipeline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1651–1660. (cited on page 35)

Long, D. L. and Fairchild, M. D. (2014). Modeling observer variability and metamerism failure in electronic color displays. *Journal of Imaging Science and Technology*, pages 30402–1 – 30402–14,. (cited on page 15)

Mahy, M., Van Eycken, L., and Oosterlinck, A. (1994). Evaluation of uniform color spaces developed after the adoption of cielab and cieluv. *Color Research & Application*, 19(2):105–121. (cited on page 23)

Malacara, D. (2003). Color vision and colorimetry: theory and applications. (cited on page 3)

Mann, S., Lo, R. C. H., Ovtcharov, K., Gu, S., Dai, D., Ngan, C., and Ai, T. (2012). Realtime hdr (high dynamic range) video for eyetap wearable computers, fpga-based seeing aids, and glasseyes (eyetaps). In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–6. IEEE. (cited on page 58)

Mantiuk, R., Daly, S. J., Myszkowski, K., and Seidel, H.-P. (2005). Predicting visible differences in high dynamic range images: model and its calibration. In *Human Vision and Electronic Imaging X*, volume 5666, pages 204–214. SPIE. (cited on page 37)

Mantiuk, R., Kim, K. J., Rempel, A. G., and Heidrich, W. (2011). Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on graphics (TOG)*, 30(4):1–14. (cited on page 123)

Mantiuk, R., Krawczyk, G., Mantiuk, R., and Seidel, H.-P. (2007). High-dynamic range imaging pipeline: perception-motivated representation of visual content. In *Human Vision and Electronic Imaging XII*, volume 6492, pages 382–393. SPIE. (cited on pages 2, 5, 70, and 74)

Mantiuk, R., Myszkowski, K., and Seidel, H.-P. (2006). A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)*, 3:286–308. (cited on pages 72 and 76)

Mantiuk, R. K., Tomaszewska, A., and Mantiuk, R. (2012). Comparison of four subjective methods for image quality assessment. In *Computer graphics forum*, volume 31, pages 2478–2491. Wiley Online Library. (cited on page 109)

Martínez, M. Á., Valero, E. M., Nieves, J. L., Blanc, R., Manzano, E., and Vílchez, J. L. (2019). Multifocus hdr vis/nir hyperspectral imaging and its application to works of art. *Optics Express*, 27(8):11323–11338. (cited on page 4)

McCann, J. J. and Rizzi, A. (2007). Veiling glare: the dynamic range limit of hdr images. In *Human Vision and Electronic Imaging XII*, volume 6492, pages 394–403. SPIE. (cited on page 5)

Melgosa, M., Huertas, R., and Berns, R. S. (2008). Performance of recent advanced color-difference formulas using the standardized residual sum of squares index. *JOSA A*, 25(7):1828–1834. (cited on page 23)

Meylan, L. and Susstrunk, S. (2006). High dynamic range image rendering with a retinex-based adaptive filter. *IEEE Transactions on image processing*, 15(9):2820–2830. (cited on page 76)

Miller, S., Nezamabadi, M., and Daly, S. (2013). Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal*, 122(4):52–59. (cited on page 37)

Mohammadi, P., Ebrahimi-Moghadam, A., and Shirani, S. (2014). Subjective and objective quality assessment of image: A survey. *arXiv preprint arXiv:1406.7799*. (cited on page 120)

Moroney, N., Fairchild, M., Hunt, R., and Li, C. (2002). The ciecam02 color appearance model. *Color Research & Application*. (cited on page 23)

Nafchi, H. Z., Shahkolaei, A., Moghaddam, R. F., and Cheriet, M. (2014). Fsitm: A feature similarity index for tone-mapped images. *IEEE Signal Processing Letters*, 22(8):1026–1029. (cited on pages 124, 125, and 138)

Nagao, M. (1979). Edge preserving smoothing. (cited on page 78)

Narendra, P. M. (1981). A separable median filter for image noise smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1:20–29. (cited on page 78)

Narwaria, M., Da Silva, M. P., and Le Callet, P. (2015). Hdr-vqm: An objective quality measure for high dynamic range video. *Signal Processing: Image Communication*, 35:46–60. (cited on page 123)

Narwaria, M., Da Silva, M. P., Le Callet, P., and Pepion, R. (2013). Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality. *Optical Engineering*, 52(10):102008–102008. (cited on page 124)

Narwaria, M., Da Silva, M. P., Le Callet, P., Valenzise, G., De Simone, F., and Dufaux, F. (2016). Quality of experience and hdr: concepts and how to measure it. In *High Dynamic Range Video*, pages 431–454. Elsevier. (cited on pages 31 and 123)

Nayar and Branzoi (2003). Adaptive dynamic range imaging: Optical control of pixel exposures over space and time. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 1168–1175. IEEE. (cited on page 33)

Nayar, S. K. and Mitsunaga, T. (2000). High dynamic range imaging: Spatially varying pixel exposures. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, volume 1, pages 472–479. IEEE. (cited on page 33)

Nezamabadi, M., Miller, S., Daly, S., and Atkins, R. (2014). Color signal encoding for high dynamic range and wide color gamut based on human perception. In *Color Imaging XIX: Displaying, Processing, Hardcopy, and Applications*, volume 9015, pages 98–109. SPIE. (cited on pages 64 and 105)

Oppenheim, A., Schafer, R., and Stockham, T. (1968). Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE*, 56(8):1264–1291. (cited on page 75)

P.910, I.-T. (2008). *Subjective video quality assessment methods for multimedia applications*. ITU, Geneva, Switzerland. (cited on pages 108 and 109)

Pan, Z., Yu, M., Jiang, G., Xu, H., Peng, Z., and Chen, F. (2020). Multi-exposure high dynamic range imaging with informative content enhanced network. *Neurocomputing*, 386:147–164. (cited on page 34)

Park, Y.-I., Song, J. W., and Kang, S.-J. (2022). 65-3: Invited paper: Deep learning-based image enhancement for hdr imaging. In *SID Symposium Digest of Technical Papers*, volume 53, pages 865–868. Wiley Online Library. (cited on page 34)

Pedersen, M., Hardeberg, J. Y., et al. (2012). Full-reference image quality metrics: Classification and evaluation. *Foundations and Trends® in Computer Graphics and Vision*, 7(1):1–80. (cited on pages 120 and 122)

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639. (cited on page 78)

Pinson, M. H. and Wolf, S. (2003). Comparing subjective video quality testing methodologies. In *Visual Communications and Image Processing 2003*, volume 5150, pages 573–582. SPIE. (cited on page 109)

Polak, A., Kelman, T., Murray, P., Marshall, S., Stothard, D. J., Eastaugh, N., and Eastaugh, F. (2017). Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. *Journal of Cultural Heritage*, 26:1–11. (cited on page 4)

Pollock, C. R. (1995). *Fundamentals of optoelectronics.* Irwin. (cited on page 57)

Ponomarenko, N., Ieremeiev, O., Lukin, V., Jin, L., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., et al. (2013). A new color image database tid2013: Innovations and results. In *Advanced Concepts for Intelligent Vision Systems: 15th International Conference, ACIVS 2013, Poznań, Poland, October 28-31, 2013. Proceedings 15*, pages 402–413. Springer. (cited on pages 112 and 173)

Pouli, T., Reinhard, E., Larabi, M.-C., and Abebe, M. (2016). Color management in hdr imaging. In *High Dynamic Range Video*, pages 237–272. Elsevier. (cited on pages 31 and 123)

Pouyet, E., Rohani, N., Katsaggelos, A. K., Cossairt, O., and Walton, M. (2018). Innovative data reduction and visualization strategy for hyperspectral imaging datasets using t-sne approach. *Pure and Applied Chemistry*, 90(3):493–506. (cited on page 4)

Ramponi, G. (1995). A rational edge-preserving smoother. In *Proceedings., International Conference on Image Processing*, volume 1. IEEE. (cited on page 78)

Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., and Myszkowski, K. (2010). *High dynamic range imaging: acquisition, display, and image-based lighting.* Morgan Kaufmann. (cited on pages 2, 3, 5, 12, 29, 33, and 167)

Rouse, D. M., Pépion, R., Le Callet, P., and Hemami, S. S. (2010). Tradeoffs in subjective testing methods for image and video quality assessment. In *Human Vision and Electronic Imaging XV*, volume 7527, pages 108–118. SPIE. (cited on page 109)

Rousselot, M., Auffret, E., Ducloux, X., Le Meur, O., and Cozot, R. (2018). Impacts of viewing conditions on hdr-vdp2. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1442–1446. IEEE. (cited on page 124)

Rousselot, M., Le Meur, O., Cozot, R., and Ducloux, X. (2019). Quality assessment of hdr/wcg images using hdr uniform color spaces. *Journal of Imaging*, 5(1):18. (cited on pages 120 and 124)

Salih, Y., Malik, A. S., Saad, N., et al. (2012). Tone mapping of hdr images: A review. In *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012)*, volume 1, pages 368–373. IEEE. (cited on pages 5, 70, 74, and 75)

Sapiro, G. and Ringach, D. L. (1996). Anisotropic diffusion of color images. *Human Vision and Electronic Imaging*, 2657. (cited on page 78)

Schanda, J. (2007). *Colorimetry: understanding the CIE system.* John Wiley & Sons. (cited on pages 9, 19, 21, and 168)

Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A., and Vorozcovs, A. (2004). High dynamic range display systems. In *ACM SIGGRAPH 2004 Papers*, pages 760–768. ACM SIGGRAPH. (cited on page 31)

Shukla, A. and Kot, R. (2016). An overview of hyperspectral remote sensing and its applications in various disciplines. *IRA Int. J. Appl. Sci*, 5:85. (cited on page 3)

St, L., Wold, S., et al. (1989). Analysis of variance (anova). *Chemometrics and intelligent laboratory systems*, 6(4):259–272. (cited on pages 128 and 134)

Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pages 839–846. IEEE. (cited on page 78)

Tomaszewska, A. and Mantiuk, R. (2007). Image registration for multi-exposure high dynamic range image acquisition. *Skala-UNION Agency.* (cited on pages 3 and 33)

Valenzise, G., De Simone, F., Lauga, P., and Dufaux, F. (2014). Performance evaluation of objective quality metrics for hdr image compression. In *Applications of Digital Image Processing XXXVII*, volume 9217, pages 78–87. SPIE. (cited on page 124)

Venkatanath, N., Praneeth, D., Bh, M. C., Channappayya, S. S., and Medasani, S. S. (2015). Blind image quality evaluation using perception based features. In *2015 twenty first national conference on communications (NCC)*, pages 1–6. IEEE. (cited on pages 126, 138, and 140)

Wang, D. C., Vagnucci, A. H., and Li, C. (1981). Gradient inverse weighted smoothing scheme and the evaluation of its performance. *Computer Graphics and image processing*, 15(2):167–181. (cited on page 78)

Wang, H., Raskar, R., and Ahuja, N. (2005). High dynamic range video using split aperture camera. In *IEEE 6th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, Washington, DC, USA*. Citeseer. (cited on page 33)

Wang, Y., Zhang, J., and Luo, B. (2018). High dynamic range 3d measurement based on spectral modulation and hyperspectral imaging. *Optics Express*, 26:34442–34450. (cited on page 4)

Wu, S., Xu, J., Tai, Y.-W., and Tang, C.-K. (2018). Deep high dynamic range imaging with large foreground motions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 117–132. (cited on page 35)

Wyszecki, G. and Stiles, W. S. (2000). *Color science: concepts and methods, quantitative data and formulae*, volume 40. John wiley & sons. (cited on pages 9, 11, 20, 63, 144, and 175)

Xiao, F., DiCarlo, J. M., Catrysse, P. B., and Wandell, B. A. (2002). High dynamic range imaging of natural scenes. In *Color and imaging conference*, volume 2002, pages 337–342. Society for Imaging Science and Technology. (cited on page 124)

Yang, M. and Bourbakis, N. (2005). An overview of lossless digital image compression techniques. In *48th Midwest Symposium on Circuits and Systems, 2005.*, pages 1099–1102. IEEE. (cited on page 28)

Yeganeh, H. and Wang, Z. (2012). Objective quality assessment of tone-mapped images. *IEEE Transactions on Image processing*, 22(2):657–667. (cited on pages 124, 125, and 138)

Yin, L., Yang, R., Gabbouj, M., and Neuvo, Y. (1996). Weighted median filters: a tutorial. *IEEE Transactions on circuits and systems II: analog and digital signal processing*, 43(3):157–192. (cited on page 78)

Zerman, E., Valenzise, G., and Dufaux, F. (2017). An extensive performance evaluation of full-reference hdr image quality metrics. *Quality and User Experience*, 2:1–16. (cited on pages 123 and 124)

List of Figures

List of Tables