Erasmus+

COSI Computational Colour and Spectral Imaging

Master in Computational Colour and Spectral Imaging (COSI)

UNIVERSITY OF EASTERN FINLAND · UNIVERSIDAD DE GRANADA · UNIVERSITÉ DE LYON · UNIVERSITÉ JEAN MONNET SAINT-ÉTIENNE · NTNU Norwegian University of Science and Technology

High Dynamic range imaging for dual-RGB system for cultural heritage

Master Thesis Report

Presented by

Ali Raza Syed

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. Philippe Colantoni, Jean Monnet University, France
2. Dr. Dmitry Semenov, University of Eastern Finland, Finland

Submission of the thesis: 10th August 2023
Day of the oral defense: 4th September 2023

Abstract

This thesis work explores the usage of a dual-RGB multispectral imaging setup for enhancing the colour accuracy of photographic films capture. A dual-RGB multispectral imaging system captures two images from a trichromatic camera using two different light sources. The combination of trichromatic camera with two light sources give rise to a six-channels multispectral imaging system. One of the applications for such a system is in cultural heritage for digitally preserving the appearance of artworks. The thesis work is mainly divided into six parts. The first part relates to the characterization of the different components of the imaging pipepline including the camera, LEDs light source, and colour targets transmittances. In the second part, dual-RGB simulations are performed on the 630 different combinations of light source pairs to find an optimal light source pair minimizing the ΔE_{00} error. The optimal light sources outperformed a simulated three-channel capture in colour accuracy. The third part of the thesis improves the colour accuracy of dual-RGB capture simulations. This is done by applying optimizations on the LEDs intensities. Two optimization methods of Particle Swarm Optimization (PSO) and genetic algorithm are used. Both algorithms are shown to converge on the same solution. The optimized light sources are referred to being part of the dual-RGB-PSO. In the fourth part of the thesis, real capturing sessions are conducted for three-channel, dual-RGB, dual-RGB-PSO optimized light sources. Dual-RGB and dual-RGB-PSO are shown to outperform the colour accuracy measured from three-channel captures. In real captures, both dual-RGB and dual-RGB-PSO performs equally well in contrast to the simulations where dual-RGB-PSO performed better than dual-RGB. The fifth part of the thesis is related to high dynamic range (HDR) imaging in the workflow of dual-RGB. An HDR pipeline is followed for the dual-RGB and dual-RGB-PSO captures. The results show a promising direction that can be undertaken by implementing HDR imaging techniques in dual-RGB context. The sixth part of the thesis relates to the application of the dual-RGB workflow, which was implemented on photographic films, for reflective objects. The LED light source which was originally built only for photographic films, was adapted to be operable for reflective objects with minimal changes and easy to move back and forth between film and reflective capturing mode. The simulations in the reflective mode converged to similar optimal light sources as were obtained for photographic films. Also the light sources optimized for photographic films showed good results for reflective objects. The results hint at the possibility of the development of a combined multispectral imaging setup where both the films and reflective objects can be captured, and possibly, with minimal changes both in the physical setup and light sources.

Acknowledgment

First of all, I would like to thank *Allah* for giving me the one-of-a-kind chance to take part in the COSI program. Each semester, we went to a different country, which was a very unique part of the program. The area of colour science was also something I had never heard of before. I am glad I found out about it in time.

Thank you, Giorgio, for guiding me through my summer internship and my thesis. Interactions with you was one of the main reasons I became more interested in colour science. So, if you think that one of the primary roles of a teacher is to get students interested in his field, you can say that you have done your job for at least one of your students.

Thank you Mekides for your feedbacks on the thesis work and your general appreciative and understanding nature. You played a pivotal role in arranging the collaborative work between RIT and NTNU for this thesis work.

Thank you Susan for your feedbacks on the thesis work. I always found your comments on the work to be very concise and right to the point. I unfortunately could not have more interactions with you and Mekides as much as I would have liked because of the visa delay issue. But you both have been very understanding and appreciative of the efforts that I put into the work.

Thank you Jana for doing the amazing work of managing the laboratory resources at the Colourlab. It is difficult to find people who do their daily work with all their heart and a big smile, and I consider you one of them.

Thank you Peter for being such a lively soul around the Colourlab. It is always refreshing to talk with you even just of very routine things. Also thank you for arranging the projector and the colourchecker.

Thank you Aqsa, Akash, and Ehsan for your suggestions on the work and for just being there. The thesis work would have become overly burdensome if you were not around. All work and no laughs (read Monopoly) would have made this thesis an unachievable assignment.

Contents

1 Introduction

The journey of a thousand miles begins with one step.

Unknown

This thesis focuses on the implementation of a dual-RGB multispectral imaging system for photographic films and explores its possible enhancements for the application in cultural heritage. The enhancements include the expansion of the solution space with the inclusion of LED intensity variations in the light source and the utilisation of multi-exposure imaging techniques for high dynamic range (HDR) imaging. The setup is also utilized for reflective object captures showing the potential of an interoperable system where both films and reflective objects can be captured. This chapter provides a concise introduction to the ideas of spectral imaging, with a specific focus on one variant known as dual-RGB imaging. Additionally, it explores the application of LEDs as a light source and the utilisation of high dynamic range imaging techniques.

1.1 Spectral imaging

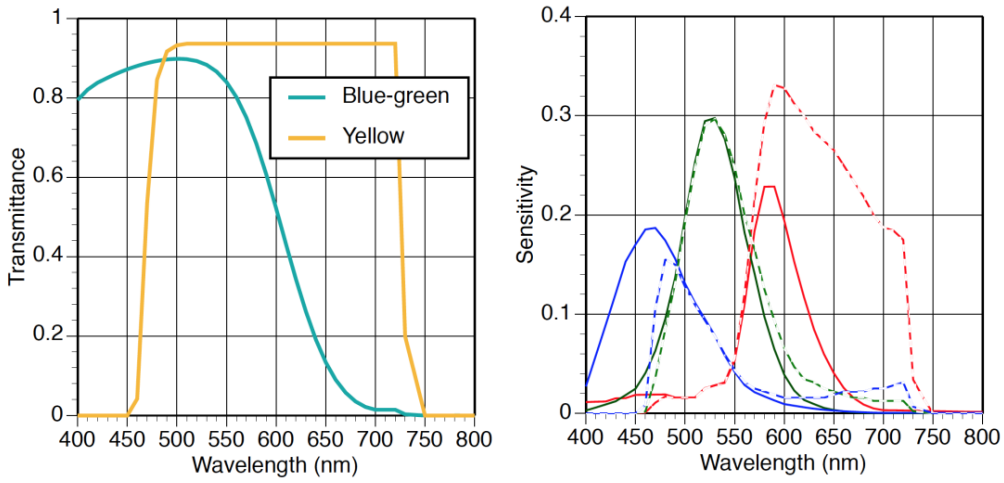
The technique of spectral imaging integrates the disciplines of spectroscopy and imaging. In a spectral image, both spatial and spectral information are retained. Following its effective implementation in remote sensing, spectral imaging has also demonstrated its utility in the domain of cultural heritage. The VASARI project, as documented by Saunders and Cupitt (1993), represents one of the earliest examples of spectral imaging implementation in the field of cultural heritage. This project involved the utilisation of a multispectral imaging equipment. With advancements

in technology, the usage of hyperspectral imaging equipment in the field of cultural heritage has become increasingly prevalent (Casini et al., 2005)(Delaney et al., 2010). Both multispectral and hyperspectral imaging techniques are encompassed under the field of spectral imaging. The primary distinction involves the quantity of spectral channels that are acquired for the object. Although there are no clear distinctions between the number of bands in multispectral and hyperspectral systems, we typically refer to a system as multispectral if it contains between five and a few dozens bands, and as hyperspectral if it contains more than that. Multispectral systems typically sample the electromagnetic spectrum using various filters and light sources. Hyperspectral systems typically utilise either a prism-based or diffraction grating-based configuration to disperse the incident light into discrete spectral bands and capture it.

1.2 Dual-RGB imaging

Despite the availability of spectral imaging as a non-invasive means of inspecting and documenting artworks, its adoption in heritage institutions remains limited. The cause of the limitation is the installation of a complex system. Furthermore, the operation of these devices need individuals with a specialised understanding of colour science and spectral imaging. This motivated researchers to develop approaches that not only have lower setup and maintenance costs, but also require a lower level of scientific expertise. One of the techniques employed in this context is referred to as dual-RGB imaging (Kuzio, 2023).

Dual-RGB imaging system, conceptualised by Berns et al. (2005), refers to the use of a color-filter array camera and two transmissive filters to generate a six-channel multispectral image. The utilisation of dual-RGB imaging has been demonstrated to result in improved colour fidelity when compared to alternative, more complex methodologies including the use of monochrome cameras coupled with liquid crystal tunable filters (LCTFs) or transmissive filters (Berns et al., 2005). The selection of the final filters was facilitated by conducting simulations on a total of 2380 theoretical filters, which were generated by pairing 41 Schott glass filters. The criterion for selection was minimising the root-mean-square error (RMSE) of the predicted reflectances of the colour targets (Berns et al., 2004). Figure.1.1 shows the transmittances of the filters and the corresponding sensitivities of the six channels.



(a) (b)

Figure 1.1: *Dual-RGB imaging setup proposed by Chen and Berns (2012)(Kuzio, 2023). (a) shows the transmittances of the optimal filters for the dual-RGB multispectral imaging and (b) shows the sensitivities of the resulting channels.*

1.3 Dual-RGB with LEDs light source

A simpler kind of dual-RGB imaging has emerged as a result of developments in LED light sources. The filters-based dual-RGB system employs a broadband light source along with selective transmissive filters on the camera to achieve the multispectral characteristics. While this configuration may be preferred by photographers due to its naturalness, it doesn't come without its drawbacks. Broadband light sources are commonly avoided in the context of cultural heritage due to their potential to cause damage to artwork (France et al., 2010) (Christens-Barry et al., 2009). Misregistrations between images can occur as a result of using various filters, which can be caused by physical movements during the installation of the filters or variations in their optical properties. Additionally, the filters tend to become dirty with use (Kuzio, 2023). The usage of LED light sources eliminates the need for filters and broadband light sources. These benefits are in addition to better colour accuracy in LED-based dual-RGB imaging when compared with optical-filter-based dual-RGB imaging (Kuzio and Farnand, 2022a).

1.4 Optimization of LED intensities

LED light sources have the capability to be programmatically tunable, allowing for the adjustment of the intensity of each LED. This functionality is analogous

to the dynamic construction of different optical filters in filter-based dual-RGB setup. Tunable LED-based imaging systems have been used in medical applications. Hyttinen et al. (2018) employed the Particle Swarm Optimisation (PSO) optimisation method to improve contrast in dental lesions images. A similar optimisation method is additionally used in the detection of diabetic retinopathy in retinal images (Fält et al., 2011) (Bartczak et al., 2017). The optimisation of a tunable LED-based system has also been used in the context of cultural heritage. Durmus et al. (2018) optimises a seven-LED system to reduce the light absorption and energy consumption of light sources with minimal perceptual changes to paintings. Like PSO, another evolutionary multi-objective genetic algorithm is used for the optimization.

Berns (2011) used theoretical LED spectra that were simulated using Gaussian functions. Two different optimisations were done using the Generalised Reduced Gradient (GRG) Nonlinear method. The objective of the first optimisation was to achieve a match with the CIE D65 chromaticity, while the second optimisation aimed to either maximise or minimise the chroma. CIE D65 was matched due to its widespread use as a daylight illuminant and its rendering-related benefits, such as in chromatic adaptation transformations when computing colour differences. The use of light sources that increase chroma can serve as a means to compensate for the diminished colorfulness of light-sensitive materials when they are exposed to only low levels of illumination, resulting in a reduction of their original colorfulness. Likewise, certain paintings are produced in conditions of low illumination, such as the interiors of churches, resulting in a reduction in colourfulness in the depicted objects. Artists used high chromatic colours in order to counterbalance the diminished colourfulness within the interior spaces. Light sources decreasing the chroma can be used to compensate for the increased chroma of such artwork.

Schanda et al. (2016) utilised a combination of four red, green, blue, and warm white LEDs in order to maximise colour fidelity and minimise damage. The optimised light source was deemed the most acceptable solution based on visual evaluations conducted by curators. Tuzikas et al. (2014) utilised four red, green, blue, and amber LEDs to optimise the damage factor and colour appearance while maintaining constant irradiance.

1.5 High dynamic range imaging

The data for a spectral image is usually recorded on a charge-coupled device (CCD) or Complementary Metal-Oxide-Semiconductor (CMOS) sensors. The electron well-depth of these sensors is fixed. The electron well-depth is indicative of the capacity for storing information in a single exposure. This constrains the dynamic range of the scene which can be recorded in a capture. This presents a challenge for

both standard three-channel and spectral images. This limitation is also highlighted during the capture of a high dynamic range artwork in cultural heritage. The spectral image of the artwork may exhibit areas that are either overexposed or underexposed (Daniel et al., 2016). Other situations of over- or under-exposed areas may occur due to the constraints in the controllability of uniform illumination on the artwork. This is typically the case when the size of the artwork is very large or when the artwork's condition prevents it from being placed on a hyperspectral linear stage, ultimately leading to non-uniform illumination (Martínez et al., 2019).

Martínez et al. (2019) discusses extreme examples of dynamic range by examining two hypothetical sections of a painting. One section comprises a material with extremely low reflectance, on which a very low level of illumination falls, at a wavelength to which the camera sensor has minimal sensitivity. In this region, an extremely long exposure time would be required to detect any signal above the noise floor. Now consider the other region of the painting, which contains a material with a very high reflectance and is illuminated by a very bright light at a wavelength where the sensitivity of the camera sensor is very high. A very low exposure time would be required for such a region in order to avoid saturating the sensor. Typical hyperspectral cameras lack the capability to adjust exposure time for different regions of object or wavelengths. This restricts the application of hyperspectral cameras to scenes that fall within the dynamic range of the camera sensor.

High dynamic range (HDR) imaging provides a method for recovering the scene's radiance up to any theoretical limit. Practically, this is restricted by the sensor noise floor and the maximum amount of exposure before it is damaged (Mann et al., 2010) (Lapray et al., 2017). In a single exposure, a high exposure image can provide information about areas with low illumination, while a low exposure image can provide information about areas with high illumination. One method for producing HDR images is to convert a set of low dynamic range (LDR) images pixel values into a quantity proportional to scene radiance. This is achieved by dividing the image pixel values by the exposure time. This operation increases low exposure time values while decreasing high exposure time values. As a result, all of the values are converted to a similar scale that is proportional to the scene radiance. In practise, this operation should only occur after compensating for the non-linear processing on the camera's digital values. This nonlinearity is generally known as a camera response function (CRF). The non-parametric method proposed by Debevec is one of the most prevalent approaches to estimating this function Debevec and Malik (1997). This method predicts the radiance values up to a scale factor using the principle of reciprocity. Reciprocity is the principle that the total exposure of a sensor is determined by the product of incoming irradiance and exposure time. It implies that altering either the irradiance by one-half of it or the

exposure time by doubling it does not result in any change in the overall exposure.

1.6 Thesis contribution

The main contributions of this thesis work is summarized below:

- Implementation of dual-RGB multispectral imaging workflow for photographic films.
- LEDs intensities optimizations in dual-RGB for improved colour accuracy.
- Defining an HDR workflow for dual-RGB imaging.
- Adaptation of photographic films capturing setup for reflective objects.
- Analysis of the interoperability between dual-RGB optimized for films and reflective objects.
- Study of anomalous behaviour in the camera response functions of mechanical shutter cameras.

1.7 Thesis outline

The thesis outline is listed below:

- Chapter 2 discusses provides a background on spectral imaging and cultural heritage.
- Chapter 3 provides a theory of camera simulation.
- Chapter 4 discusses the data collection and methods utilized.
- Chapter 5 discusses the methods and results for the simulations of three-channel and dual-RGB imaging pipeline.
- Chapter 6 provides details and results about the exploration of change of intensities of LED light source with multiple optimization techniques.
- Chapter 7 discusses the real captures done against the simulations.
- Chapter 8 provides a background on HDR imaging and how it can be implemented in dual-RGB setup.

- Chapter 9 discusses how LED light source was modified to convert to reflective capture mode. The chapter also discusses the simulation and real captures results for dual-RGB imaging for reflective objects.
- Appendix A provides a study on a characteristic non-linear behavior in a supposedly linear region of CRF of a camera with mechanical shutter. A method to characterize this behavior is proposed and a comparison of cameras with mechanical shutters are drawn with electronic shutter cameras.
- Appendix B discusses a background on colorimetry. This is written for familiarizing readers with colorimetry who do not have a background in this field.
- Appendix C discusses different methodologies that were undertaken to calculate the photographic films transmittances.

Usage of tools: For the work on this thesis, QuillBot was used for grammatical errors checking and ChatGPT was used minimally for things such as generating BibTeX against a reference, transposing a LaTeX table, etc.

2 Spectral imaging and cultural heritage

A wider view reveals a larger canvas of opportunities.

Unknown

Spectral imaging is an advanced imaging technique that seeks to capture the spectral information of light. This chapter provides a concise overview of imaging, its expansion into spectral imaging, and its application in the preservation and processing of cultural heritage through the creation of digital copies.

2.1 Spectral imaging

Spectral imaging originates from two distinct fields, namely spectroscopy and imaging, both of which have undergone significant advancements and continue to be effectively employed in various applications. In the next sections, a concise overview of both fields is provided, followed by a discussion on their integration to create spectral imaging.

2.1.1 Imaging

According to Garini et al. (2006), imaging refers to the methodologies employed for capturing light signals in both spatial and temporal dimensions, resulting in the generation of either monochromatic or trichromatic signals. A simplified depiction is presented in Fig.2.1. The imaging workflow is outlined below:

1. **Light source:** The light source emits electromagnetic radiations outwards into space.

2. **Reflectance:** The object absorbs light in some wavelengths and reflects back in other wavelengths. The amount of reflected light depends on its reflectance.

3. **Lens:** Within the camera devices, the incident light undergoes transmission through the lens, which serves the purpose of directing and concentrating the light onto the plane of the sensors.

4. **Aperture:** Subsequently, the light traverses an aperture, having the dual purpose of regulating the quantity of light entering the camera and determining the depth of field. The term of depth of field refers to the extent of distance within a scene that appears in sharp focus, meaning that each point in the scene corresponds to a single point on the sensor plane.

5. **Bayer filter:** Next, there is an array including three different kinds of filters: red, green, and blue. The configuration of these filters is commonly referred to as Bayer filters. Take note of the additional green filters. This is done in order to collect more samples in the green region, which has the highest overlap with colour luminance as described by the photopic luminosity function. The HVS exhibits more sensitivity to variations in luminance compared to changes in chromaticity, hence requiring a greater number of green filters. This is also useful for estimating more precise measurements while interpolating pixel values.

6. **Sensors:** Once the incident light has passed through the Bayer filter, it gets recorded by the sensors. It should be noted that these sensors lack the ability to discern between various colours, making them "color-blind." The sensors record values based on their spectral sensitivities. The ability to record colours is only possible because the Bayer filter separates light into three regions of electromagnetic spectrum.

7. **Demosaicking:** For each channel, there are missing pixel values due to the Bayer filter pattern. Interpolation is done for the missing pixel values for each channel to predict the missing values and generate three channels image.

As stated before, this is a basic approach for image capturing. There exist numerous additional parameters that affect the quality of images captured by a camera, some of which are given below:

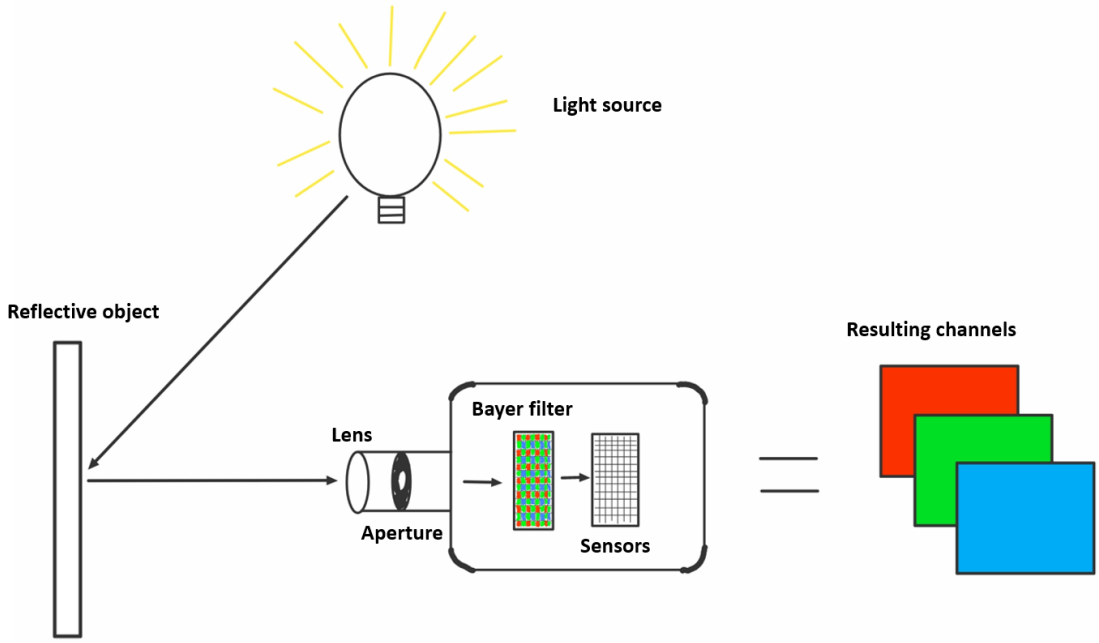


Figure 2.1: *Workflow of trichromatic capturing. 1. Light source emits electromagnetic radiations. 2. Object reflects light in particular wavelengths according to its reflectance. 3. Lens focuses light on sensors. 4. Aperture controls amount of light and depth of field. 5. Bayer pattern filters light into three different regions of electromagnetic spectrum. 6. Sensors record light signals. 7. Missing channel pixel values are interpolated (demosaicking) to produce a complete three-channel image.*

1. **Spatial resolution:** Spatial resolution of the images determine the how much details of the scene can be resolved through the image. The more the spatial resolution is, the more the details can be extracted.

2. **Spectral sensitivity:** Spectral sensitivity of the sensors determine which wavelengths can be recorded and with how much efficiency. This efficiency is also given by quantum efficiency which is the number of electrons generated per photon of a wavelength light. The more the sensor is sensitive to a wavelength of light, the better the signal can be extracted for that wavelength.

3. **Dark noise:** Dark noise is the measurements which is generated even if no light signal is input to the imaging system. This occurs due to the thermal excitations of the electrons in the sensors. This becomes more relevant in the applications where availability of incoming photons are limited in applications such as live-cell imaging. The lesser the dark noise there is, the better the signal can be extracted.

4. **Well depth:** Pixel well depth determines the number of electrons that can be stored in the sensor for valid readings before being saturated (and potentially being overflown corrupting the neighboring sensor readings). This determines the dynamic range capabilities of the sensor. The more the pixel well depth, the more the dynamic range of the scene can be captured in a single shot.

The quality of the imaging setup is influenced by many other variables, such as the numerical aperture of the objective lens, the fill area of the sensor (photo-sensitive area of the sensor), the range of integration time, the resolution of the analog-to-digital conversion, and additional factors.

2.1.2 Spectroscopy

In the 17th century, Newton made the groundbreaking discovery that white light is composed of a range of different component lights. This discovery led to the development of spectrographs in the 19th century, which facilitated the measurement and analysis of these component lights called as electromangetic spectrum, so enabling the widespread use of spectroscopy in the study of matter. Spectroscopy is directly related to the energy-band structure of atoms. Electrons in atoms are excited from their ground-state energy level to an excited state by light absorption. When they discharge, they emit a 'signature' spectrum of light, which aids in the atomic study. Spectroscopy is concerned with obtaining and recording the spectrum of light from a single spatial point in space, making spectroscopic measurements only a function of wavelength. Similar to imaging, spectrometry quality is determined by a number of factors, some of which are discussed below:

1. **Spectral resolution:** Spectral resolution that determines the narrowest wavelength difference that can be detected
2. **Spectral range:** Spectral range that determines the minimum and the maximum recordable wavelengths
3. **Dark noise:** Same as in imaging.
4. **Well depth:** Same as in imaging.

2.1.3 Combining imaging and spectroscopy

According to (Li et al., 2013), spectral imaging is a combination of spectroscopy and imaging techniques, enabling the acquisition of both spatial and spectral data simultaneously. In the context of imaging, the intensity of an image can be denoted

as $I(x, y)$, while the spectrum collected through spectroscopy can be represented as $I(\lambda)$. A spectral image contains both spatial and spectral information represented by $I(x, y, \lambda)$. Generally speaking, in order to obtain a spectral image that corresponds to n-dimensional spatial information in a scene, at least $(n + 1)$-dimensional sensor array is required. There exist multiple hardware configurations for the acquisition of a spectral image, as depicted in Fig.2.2 and explained below:

1. ***Whiskbroom:*** In this method, each point in the scene its captured one at a time to get is spectrum and form the whole spectral image. Symbolically, this corresponds to taking measurements of (x_1, y_1, λ), (x_1, y_2, λ), ... (x_1, y_n, λ), (x_2, y_1, λ), (x_2, y_2, λ), ... (x_m, y_n, λ).

2. ***Pushbroom:*** This is also known as line-scanning. This method improves the time required for the captures by recording a line of spatial location in a single capture. In other words, in each capture it measures (x_1, y, λ), $(x_2, y, \lambda) \ldots (x_m, y, \lambda)$.

3. ***Band-sequential:*** This is also known as staring. This involves capturing the whole spatial information wavelength-by-wavelength i.e. in each capture it measures (x, y, λ_1), $(x, y, \lambda_1) \ldots (x, y, \lambda_k)$.

4. ***Snapshot:*** Although it is in very early stages but this method captures all the spatial and spectral information in a single capture which means it captures the whole (x, y, λ) simultaneously.

The aforementioned techniques are mainly used for hyperspectral imaging. The aforementioned techniques employ either a diffraction grating or a prism configuration to disperse the incident light signal into its component parts, next recording it. An alternative method exists for capturing each of the individual components of the light signal, but with a reduced level of spectral resolution. These setups are commonly referred to as multispectral imaging. Fig.2.3 illustrates the two methods by which filters can be employed in combination with a conventional 3-channel RGB camera, resulting in the setting up of a multispectral imaging system. The figure shows the configuration of two filters, resulting in a system that is defined as a 6-channel multispectral setup.

It is important to note that in both hyperspectral and multispectral setups, data acquisition is carried out using imaging sensors such as CMOS or CCD, each possessing their own spectral sensitivities. In order to record a quantity which is proportional to the incoming radiance, it is important to remove the effect of the sensor sensitivity from the measurements.

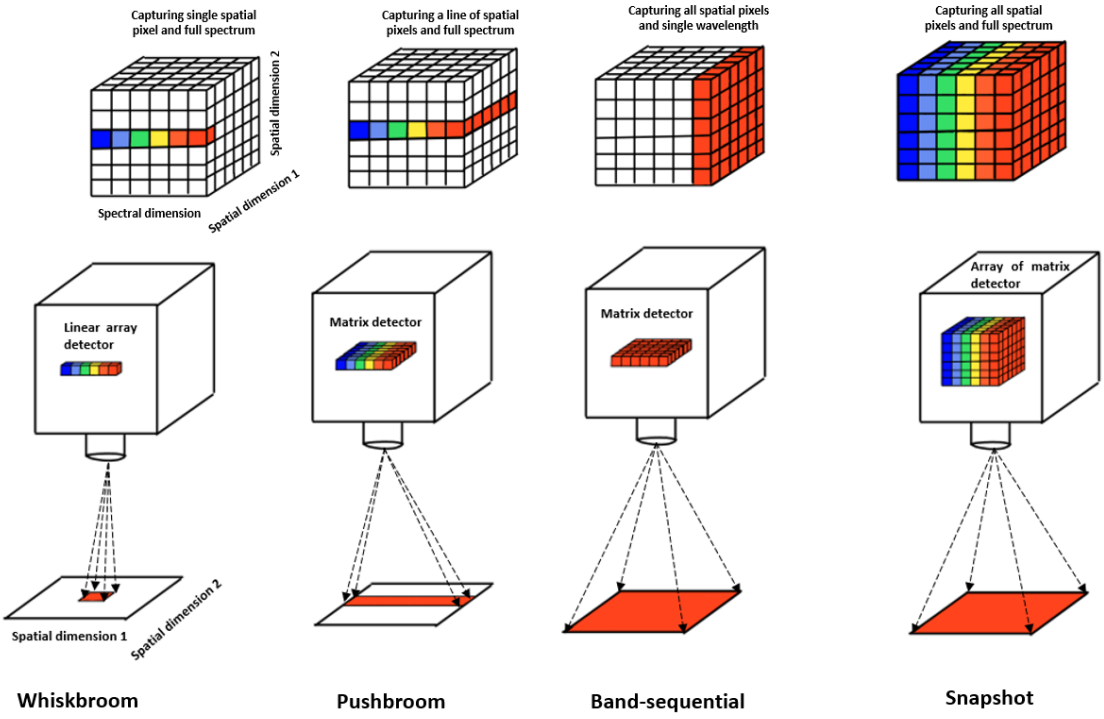
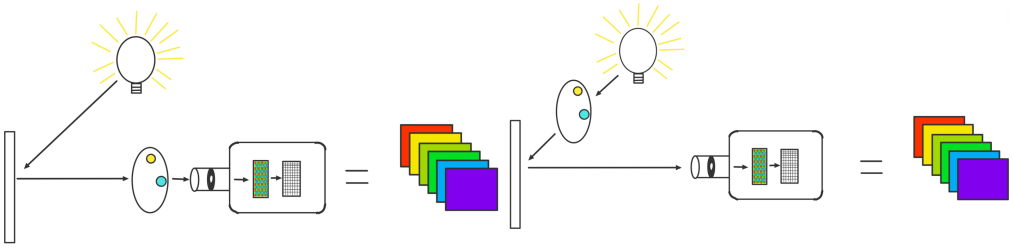


Figure 2.2: *Hyperspectral capturing techniques. Whiskbroom captures a single spatial point, at a time, with a linear array of sensors. Pushbroom captures a line of spatial points with a matrix of sensors. Band-sequential captures all the spatial points against a single wavelength. Snapshot captures all the spatial points against all the wavelengths.*



(a) **(b)**

Figure 2.3: *(a) represents passive mutltispectral imaging setup where the filter array is placed before the camera sensor. (b) represents active multispectral imaging setup where filter array is placed after the light source.*

2.2 Cultural heritage

Brumann (2015) defines cultural heritage as '... the sites, movable and immovable

artifacts, practices, knowledge items, and other things that a group or society has identified as old, important, and therefore worthy of conscious conservation measures, often at the hands of specialized institutions.' Within the scope of our thesis, our primary focus revolves around artworks that mainly exist in a two-dimensional form, such as paintings and photographic films.

The preservation of artwork is essential for gaining insight into how our ancestors lived and how the great minds of their time interpreted their surroundings. When considering art paintings, the optimal condition for preserving the artwork would involve isolating it from its environment and storing it in complete darkness (Durmus et al., 2020). This goal is in direct contradiction with having the best colour appearance in museums for visitors, which requires at least 1000 lux illumination. The ultraviolet and infrared portions of illumination can damage pigment chromophores, resulting in a colour change, and heat up adsorbed moisture, altering the painting's dimensions and structure (Schanda et al., 2016). Thus, the deterioration of artworks is an unavoidable reality.

Making digital copies of artworks has the potential for their preservation. However, the task of digitally capturing artworks in a manner that sufficiently fulfils the requirements of curator interpretation, historical accuracy, human perception, and the limitations of digital reproduction media is very challenging, if not impossible. The use of digital copies of artwork offers a non-invasive method for studying the artwork, eliminating the need for invasive micro-sampling techniques. Both invasive and non-invasive methods possess their own set of advantages and drawbacks (Pinna et al., 2009)(Artioli, 2010). Numerous scientific investigations employ a hierarchical methodology, wherein non-invasive imaging techniques are initially employed for preliminary material screening. Subsequently, if necessary, an invasive procedure is undertaken to extract micro-samples of the object for further analysis. Non-invasive methodologies typically avoid altering the artwork being examined, thereby enabling the implementation of a continuous monitoring system for the artwork. This system can be utilised to track the effectiveness of cleaning and restoration procedures, as well as to monitor the impact of environmental factors such as humidity and temperature (Picollo et al., 2020b).

There exist a variety of non-invasive approaches that are employed based on factors such as the value and dimensions of the artwork, study objectives, and available financial resources. Different methods include ultraviolet-induced fluorescence (UVF), infrared reflectography (IRR), X-ray radiography (XRR), multispectral imaging (MSI), hyperspectral imaging (HSI), Macro X-Ray Flourescence (Ma-XRF), and Time-Domain, Terahertz Imaging (THz-TDI). This study primarily focuses on the application of spectral imaging (MSI and HSI).

For many years, MSI and HSI have been successfully adapted to the application of digitizing art paintings (Cucci et al., 2019)(Delaney and Picollo, 2019). These

works were created for panel paintings that had to cover a maximum of a few square metres. There have been advances in the application of techniques similar to remote sensing to much larger objects, such as frescoes and mural paintings (Liang et al., 2014)(Cucci et al., 2015)(Cucci et al., 2017)(Camaiti et al., 2017). In addition, hyperspectral imaging (HSI) techniques have been employed in recent studies to capture photographic materials such as colour negative and positive films (Picollo et al., 2020a). Compared to the paintings discussed previously, these films are of very small size and finely detailed.

Digital copies of art paintings are also subject to being limited to how the painting appeared at the time of capture and to the criteria for a decent appearance at that time. If the observing conditions or criteria for good appearance have changed, the shoot must be redone. Decoupling the capturing and rendering stages are thus important (Berns and Chen, 2012). This issue can be resolved by incorporating multispectral reflectance data that is independent of observation conditions or aesthetic considerations. This not only benefits preservation efforts, but also provides access to more accurate analyses of art (Faries, 2005).

Due to the variety of requirements and criteria for digital reproduction of artwork, a number of different goals can be defined (Ribés et al., 2003), with emphasis on some or all of them depending on the application context. The following aims are defined as:

- Acquiring high resolution multispectral images.
- Faithfully predicting appearance of the painting in different observing conditions.
- Providing high fidelity color reproduction.
- Reconstructing spectral reflectance across all the pixels of the image.

3 Theory of camera simulation

Prediction is very difficult, especially if it's about the future.

Niels Bohr

Simulations play a crucial role in the design of devices. The simulation can be conducted in order to examine the influence of various parameters on the final objectives of the design. For instance, while examining a printer, the enhancement of toner particle noise, banding, and halftone patterns on a uniform grey patch can be achieved by modifying various design parameters of the printer. The criteria include several factors such as pixels per inch, bits per pixel, and memory efficiency, among others (MacDonald and Luo, 1999). Likewise, in the context of displays, the primary objective may involve generating stimuli that produce optimal colour and luminance accuracy. This objective must be balanced against several design considerations, including resolution, heat dissipation, size, power consumption, and other relevant factors. The complexity increases as many design parameters have interdependence. For example in the case of displays, a brighter output is linked with a higher power consumption. However, increasing power consumption necessitates improved heat dissipation, which in turn would need an increase in the size of the device. With so many conflicting requirements and complexities involved in the design of a device, simulations are an invaluable resource. This chapter focuses on the assumptions used during the simulation of devices and various functionalities, with a specific emphasis on camera simulations. These assumptions play a crucial role in capturing the essential characteristics of the device.

3.1 Underlying assumptions in simulation

Because a simulation represents a very complicated device in the real world, it is critical to consider simplifications that will keep the simulation models below tolerable error limits. These simplifications not only facilitate the implementation of these models but also produce faster outputs. The significance of faster outputs in a simulation model becomes increasingly evident when doing optimisations on various combinations of design parameters. In this section, fundamental assumptions associated with device simulations are discussed.

3.1.1 Device linearity

Devices are assumed to be linear or to have a fixed nonlinearity embedded. CCD sensors are well-known for their ability to provide a linear response across a wide range of irradiance levels (Lomheim and Kalman, 2020). Some devices also have a static non-linearity involved. In the context of displays and cameras, the term 'gamma function' has been used, whereas in the case of printers, it is referred to as the 'tone reproduction curve.'

3.1.2 Principle of superposition

Once linear device assumption is in place or the device's fixed non-linearity has been corrected and made linear, it can be assumed that the principle of superposition applies to the device outputs. Based on this principle, the device's response to a combination of inputs is equal to the sum of the responses produced by each individual input in isolation. The aforementioned assumption, although may appear trivial, effectively reduces the computational burden on the simulator. For instance, in the context of cameras, this assumption indicates that if we know the spectral and spatial response to a point light source, we can compute the response for any given input. In the context of printers and displays, this assumption indicates the possibility of calculating the response to any given input if the response against a single pixel is known.

3.1.3 Shift invariance

The devices' spectral and spatial response are considered to be shift invariant. This implies, for instance, that in the case of displays, the response of each pixel is identical regardless of its position on the screen. In the context of cameras, this

implies that the response to a point light source remains the same irrespective of its location on the sensor array.

3.2 Functions for simulating cameras

Considering a camera as the simulated device, with the assumptions of linearity, superposition, and shift invariance already defined, the camera's complete behavior can be defined with three functions called as camera response function (CRF), camera spectral response function (CSRF), and camera point spread function (CPSF).

3.2.1 Camera response function (CRF)

Within the context of a camera, the radiance of a scene traverses the optical system of the lens and then impinges onto the sensor as irradiance. This results in the generation of a voltage at the sensor, which is subsequently transformed into a digital reading or number. Finally, this converted data is recorded within the camera.

$$\text{Radiance} \Rightarrow \text{Irradiance} \Rightarrow \text{Digital number}$$

The relationship between irradiance and digital number is linear if the image is read in RAW-file format and is non-linear if it is read as a gamma-encoded version such as an image encoded in sRGB. After defining the relationship between irradiance and digital number, the relationship between radiance and irradiance can be formulated. In the simplest of situations, both radiance and irradiance can be considered proportional. Therefore, establishing a relationship between the radiance of the scene and the digital number becomes straightforward if the measurements, up to a scaling factor, are sufficient. The camera response function (CRF) is the mathematical representation that describes the relationship between scene radiance and digital number.

The presence of a non-linear relationship between digital numbers and scene radiance implies that a doubling of the digital number at one pixel is not necessarily indicative of a doubling of the scene radiance compared to the other pixel. This nonlinear relationship between digital numbers and radiance must be reversed for a variety of tasks in many fields, including colour science and computer vision, for algorithms such as creation of high dynamic range images (Mitsunaga and Nayar, 1999) (Debevec and Malik, 1997) (Mann and Picard, 1994a), photometric stereo (Basri and Jacobs, 2001) (Nayar et al., 1990) (Woodham, 1980), shape from shading (Szeliski, 2005), reflectance estimation (Luong et al., 2002), scene

illumination (Luong et al., 2002), and color constancy (Finlayson et al., 2001) (Land and McCann, 1971). Simulating motion blur is also shown to produce more realistic results if done in linear radiance domain rather than the non-linear camera domain (Debevec and Malik, 1997).

Non-linearity in the mapping of radiance to digital numbers can be intentional for the purpose of compressing the high dynamic range of irradiance falling on the sensor to a limited number of digital number levels that can be produced by the sensor. Many different algorithms have been proposed to estimate this non-linearity by assuming a form of gamma-curve (Debevec and Malik, 1997)(Mann, 2000)(Mann and Picard, 1994a)(Mitsunaga and Nayar, 1999)(Tsin et al., 2001). One commonly employed method for determining the gamma parameter involves capturing many images of a given scene, each with varying exposure levels. Other analytical forms have also been proposed (Mann, 2000) and some have also proposed to assume no particular form and applying smoothness constraints (Debevec and Malik, 1997)(Tsin et al., 2001). In Fig.3.1, CRFs for different cameras and films are shown.

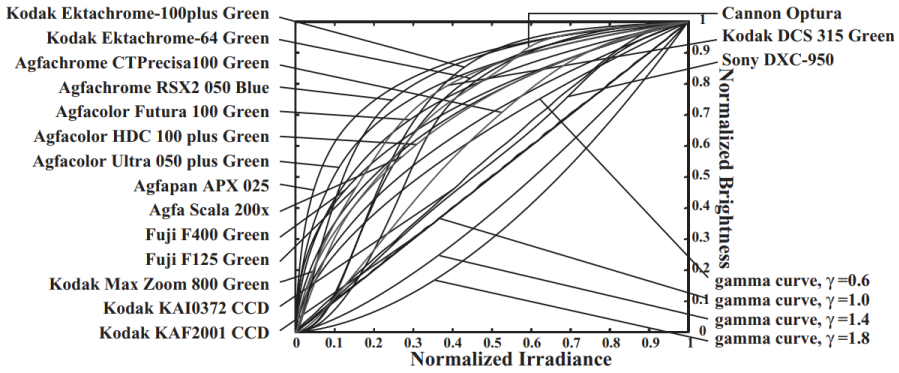


Figure 3.1: *Different CRFs provided in Grossberg and Nayar (2003). The CRFs include both photographic film response and sensor response.*

3.2.2 Camera spectral response function (CSRF)

The camera spectral response function (CSRF) is the function relating camera sensor response to different wavelengths of radiance. In a regular 3-channel Bayer filter camera, CSRF will be a set of three functions which will combine the effect of sensor wavelength-wise response and the Bayer filters transmittances.

The most common way to characterize the CSRF is the use of a monochromator (Poorvi and Farrell, 2018) (Farrell et al., 2008). This is also the recommendation by EMVA standard (EMVA, 2017). Since this method requires both time and

resources which might not be feasible for some applications, there have been many methods devised using the captures of different colorcheckers. Hubel et al. (1994) estimated the CSRF of the camera by taking captures of Macbeth Colorchecker using wide-band tungsten light source combined with 8 broadband and 16 narrow-band filters. Several methods were used to estimate the CSRF like rank-deficient pseudo inverse method (Farrell, 1993)(Sharma et al., 1993)(Trussell and Sharma, 1994) and Wiener estimation method (Lewis and Odell, 1971)(Mancill, 1975)(Pratt and Mancill, 1976). We show an example of CSRF in Fig.3.2.

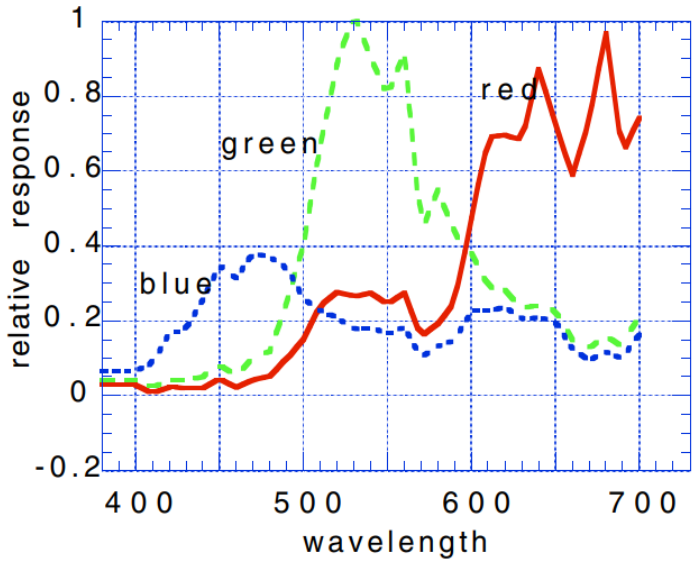


Figure 3.2: *CSRF of Kodak KAF-1600C sensor in DCS-200 camera (Hubel et al., 1994)*

Mathematically, CSRF can be defined as:

$$D_c = \int_{\lambda_{\min}}^{\lambda_{\max}} CSRF_c(\lambda) I(\lambda) R(\lambda) d\lambda \tag{3.1}$$

where D_c is linearized digital count, $CSRF_c$ is the camera spectral response function of channel c, I is the illumination, R is the reflectance of object, and $c \in \{R, G, B\}$. Since in the cameras we get access to the digitized values, we can write the continuous function definition in discrete form as

$$D_c = \sum_{\lambda=\lambda_{\min}}^{\lambda_{\max}} CSRF_{c,\lambda} I_\lambda R_\lambda \tag{3.2}$$

3.2.3 Camera point spread function (CPSF)

In linear and time-invariant (LTI) systems, the system can be completely defined by knowing the impulse response. This means that for an LTI system, we can calculate output of the system to any arbitrary input if we know the response of the system to an impulse. An impulse is generally defined as a Dirac delta function. Dirac delta function's output is non-zero only at the input value of zero. One example of an estimated CPSF in Fig.3.3.

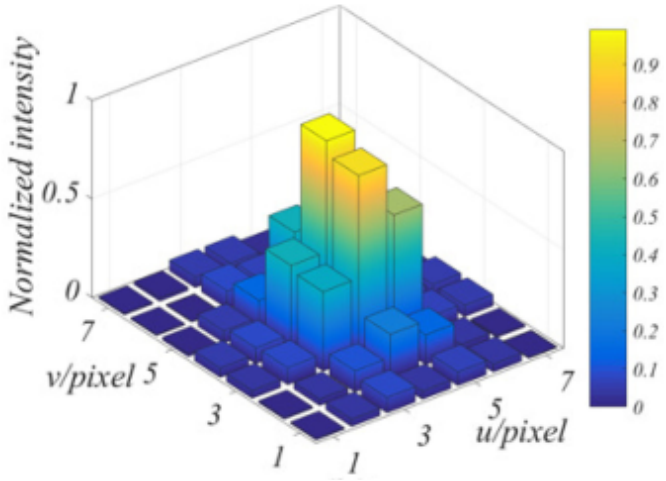


Figure 3.3: *One example of CPSF (Jiang et al., 2018). The graph shows the respone of the camera for a single point light source. The u- and v-axis show spatial position of pixels and the vertical axis shows the response of the pixels. Notice that multiple pixels get activated against a point light source.*

The response of a camera to a 2D Dirac delta function is called as a camera point spread function (CPSF). 2D Dirac delta function can be realized by capturing an image of a very small point light source. The CPFS quantifies the degree of blurring shown by the acquired image of a point light source. Usually this function is different for different wavelengths of light source. Hence CPSF becomes a function of both spatial and spectral dimensions.

There are various ways to measure or estimate the CPSF of camera. It can majorly be divided into three categories of blind CPSF estimation (Sun et al., 2013)(Cho et al., 2011)(Joshi et al., 2008), non-blind CPSF estimation (Joshi et al., 2008)(Mosleh et al., 2015)(Kee et al., 2011)(Brauers et al., 2010), and direct measurement (Du and Voss, 2004)(Lehr et al., 1998)(Jemec et al., 2017)(Navas-Moya et al., 2013).

The estimation of CPSF is an ill-posed problem, indicating that several combinations of the input sharp image and CPSFs can result in the same blurred image.

In blind CPSF estimation, constraints can be put either on the blurring kernel or the sharp image. Joshi et al. (2008) adopts a methodology that imposes constraints on the sharp edges of the image, assuming them to be step edges. By using this methodology, it is possible to estimate both uniform and spatially varying CPSFs if the image contains a sufficient number of edges with different orientations. Sun et al. (2013) adopts an optimisation strategy to estimate a statistical prior by using a database of natural images. However, this methodology is ineffective when applied to spatially varying CPSFs, as it operates under the assumption that the CPSF remains constant across the whole image.

In non-blind estimation methods, fixed calibration images are used. The type of calibration images used varies from checkerboard with circles inside the blocks (Kee et al., 2011), arc-shaped checkerboard (Joshi et al., 2008), and random noise target (Brauers et al., 2010). The CPSF estimation methods also vary from least-square minimization (Kee et al., 2011) and Bayesian method with maximum a posteriori (Joshi et al., 2008).

The third type of CPSF estimation is directly measuring it. These methods are useful for covering the spatial-variance of CPSF. Navas-Moya et al. (2013) and Jiang et al. (2018) both uses LCD screen with illuminated pixels as point light sources. Navas-Moya et al. (2013) method is affected by dark noise because of large required exposure time. Jiang et al. (2018) proposes to utilize the principles of single-pixel imaging for measurment of CPSFs which is less affected by noise. Other than LCD screens, lasers have also been used as a point light source achieving high SNR (Du and Voss, 2004) and sub-pixel accuracy (Jemec et al., 2017). Though lasers can provide high SNR but it is difficult to measure spatially varying CPSFs with them.

4 Simulation data

Predicting rain doesn't count. Building arks does

Warren Buffet

The preceding chapters have provided an overview of imaging and the mechanics involved in simulating these processes. The present chapter provides a discussion of the data necessary for the simulations and the methods employed to gather such data for the context of this thesis.

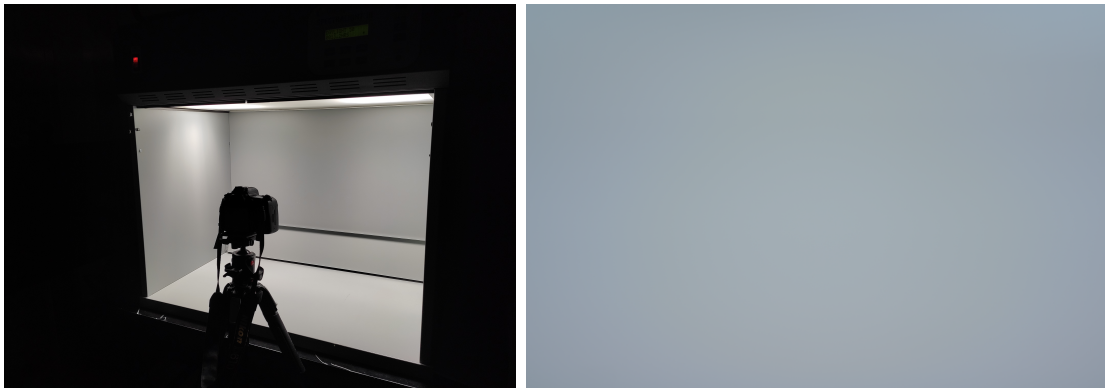
The Nikon D610 DSLR camera is used for our imaging workflow. This camera was chosen based on its availability in our lab and the fact that it was one of the most up-to-date models we owned. Given that simulations will be performed using this camera, it is important to thoroughly characterise its behaviour. The two most significant functions for us are its CRF and CSRF.

4.1 Camera response function (CRF)

4.1.1 Experimental setup

Fig.4.1 shows the experimental setup for measuring CRF as well as one example capture. The X-Rite Spectrallight III lightbooth is utilised as the light source. The light simulation is configured to use the CIE D50 illuminant. The camera's lens is detached during the process of capturing images. It enables a uniform irradiance across the camera's entire sensor plane. The camera is set to a minimum ISO of 50.

For reading the images, the RAW file format is used. As discussed in Chapter 3, this RAW file format has a linear relationship between the incoming radiance on



(a) (b)

Figure 4.1: *Experimental setup for measuring CSF. (a) shows Nikon D610 camera on a tripod pointing towards X-Rite Spectrallight III lightbooth. The camera has its lens removed for having a uniform irradiance over its sensor plane. (b) shows one of the captures at* $\frac{1}{160}$ *seconds integration time.*

the camera and recorded digital numbers. The RAW file format gives access to the least altered information that the sensor of the camera received without any compression or post-processing applied. This format gives access to the mosaicked image. For measuring CRF, average values from any of the three channels can be taken. In this study, the green channel is selected so more pixels data can be averaged as discussed previously that there are more green channels in Bayer filter. The bit-depth of the data is 14-bits which means the digital numbers range from 0 to 16383.

4.1.2 Results

The resulting camera response function is shown in Fig.4.2. It is clear that the camera is linear till the digital count of 11,000. This is approximately one-third of the bit-depth range. It is worth noting that the relationship plotted is between integration time and digital number, as opposed to radiance or irradiance v.s digital number, as discussed in Chapter 3. This is done since the light booth does not have an adjustable illumination level. However, raising the radiance or the integration time by the same factor has the same effect on the camera sensor reading, which is known as the principle of duality. As a result, the technique of modifying the camera integration time was chosen for this study.

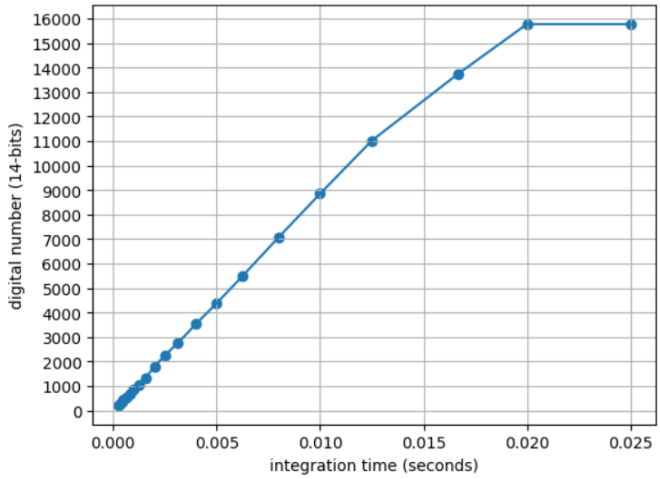


Figure 4.2: *An example of demosaicked capture for measuring CRF at the exposure time of 1/160 seconds. The digital numbers are in 14-bits. Note that after 11,000 digital number, the sensor shows a non-linear behavior.*

4.2 Camera spectral response function (CSRF)

4.2.1 Experimental setup

To characterise the CSRF, direct measurement with a monohcromator is adopted as the method of choice. The experimental setup for the measurement of CSRF is shown in Fig.4.3. A halogen lamp is used as the light source. This light source is fed through a Bentham monochromator, which converts it into monochromatic light whose wavelength is adjusted by Bentham software. The monochromatic light output from the monochromator is diffused by an integrating sphere. The camera is positioned perpendicular to the monochromatic light and is aimed directly at the integrating sphere. The telespectroradiometer (TSR, Konika Minolta CS2000) is used to take measurements in the same location as the camera. DigiCamControl software (digiCamControl, 2023) is used to control the captures from the computer so that the camera does not move throughout the capturing sessions. The camera's ISO is set to its lowest setting of 50. The integration time is adjusted so that the highest response obtained across all monochromatic light and camera channels is considerably below the non-linear region.

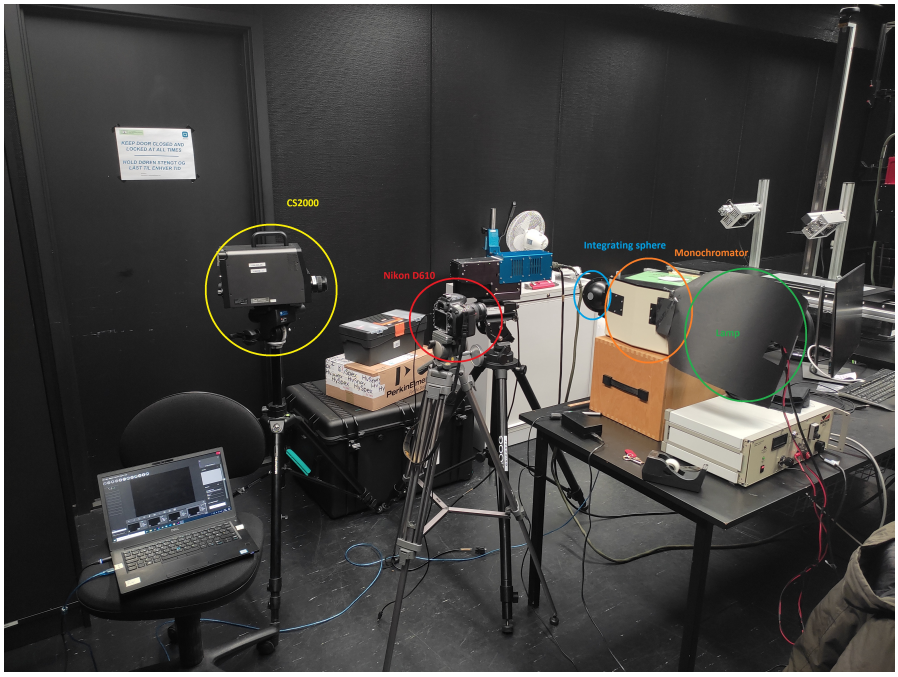


Figure 4.3: *Camera spectral response function measurement setup. The green circle shows the halogen lamp covered with a black covering. The lamp directs its light inside the Bentham monochromator highlighted in orange. The monochromator output is diffused by an integrating sphere shown in blue. The Nikon D610 camera, highlighted in red, is fixed to take measurements perpendicularly to the integrating sphere. The telespectroradiometer (TSR) Konica Minolta CS2000 is shown in yellow. Note that while capturing from TSR, it is placed in place of D610.*

4.2.2 Results

The monochromatic lights utilised in the measurements have a peak wavelength range of 380nm to 800nm with a step size of 10nm. Fig.4.4 shows monochromatic light measurements from TSR and the light source used in the monochromator.

Fig.4.5 shows the captures of the monochromatic lights at 480nm, 540nm, and 600nm. It shows the mosaicked image, demosaicked imaged, and a four pixels in the central region of the measurement area. Note the mosaicked image is greyscale since Bayer pattern sensor arrangement is a 2D-matrix of sensors. The color code of the four pixels also emphasize the Bayer pattern which is $\begin{bmatrix} R & G \\ G & B \end{bmatrix}$. The spatial resolution of each of the captured images is 4028x6080 (rows x columns) and the averaging region defined is of 250 x 250 dimension.

The wavelengths across the different captures where the three filters display

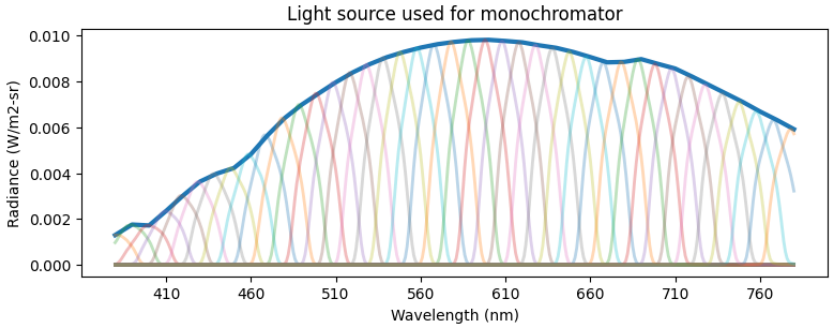


Figure 4.4: *Light source used for monochromator. The transparent spectra shows the monochromatic light measurements. The solid line indicates the peak of each of monochromatic light spectra which defines the light source spectrum.*

highest values are shown in Table 4.1. It can be seen that the captures are much below the 11,000 digital value, after which the CRF shifts to the non-linear region.

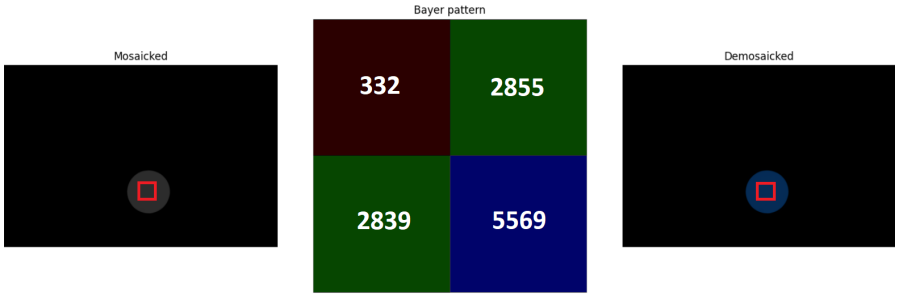
Table 4.1: *Table showing the peak monochromatic wavelengths across which the three channels showed maximum values. For example, when the monochromatic light is 480nm, the peak blue pixel value obtained in all samples from 380nm to 800nm is 5670. All the values are less than 11,000 which marks the end of the linear region of the CRF. The data is in 14-bits.*

Filter	Monochromatic light	Digital number peak value
Blue	480	5670
Green	540	8786
Red	600	7546

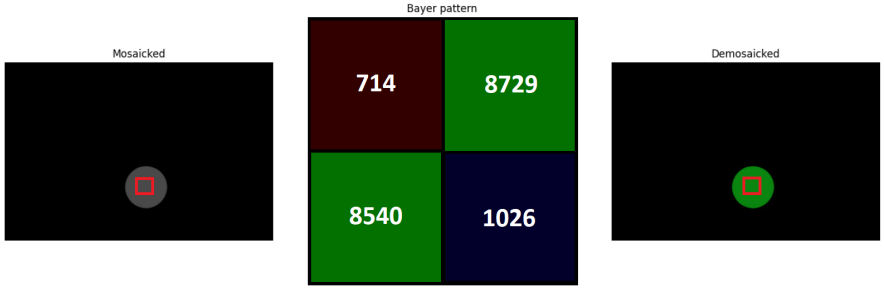
The CSRF at each wavelength is calculated as:

$$CSRF(\lambda) = \frac{DN(\lambda)}{L(\lambda)} \tag{4.1}$$

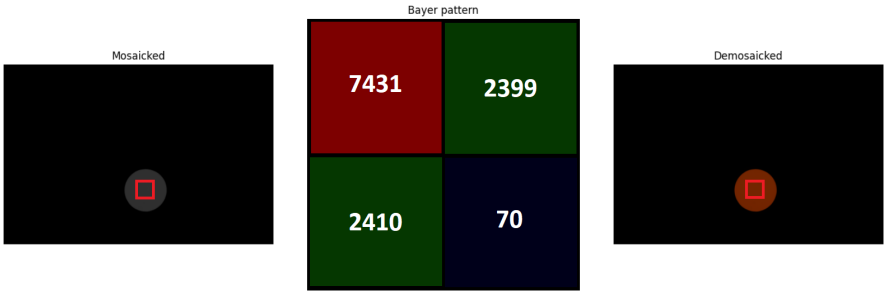
where $CSRF(\lambda)$ is the camera spectral response function when the peak monochromatic light at wavelength λ is captured, $DN(\lambda)$ is the mean digital number that we get from each square region (see Fig.4.5) defined in capture, and $L(\lambda)$ is the peak value of the light source radiance captured at λ. Also note that the mean values of the channels are taken after demosaicking. Results of CSRF before demosaicking are also checked but no difference is found. The resulting CSRF is shown in Fig.4.6.



(a) *Capture of 480nm monochromatic light*



(b) *Capture of 540nm monochromatic light*



(c) *Capture of 600nm monochromatic light*

Figure 4.5: *The first column shows the Bayer-patterned mosaicked image. The red squares represent the area of the images where the mean digital numbers of the images were taken for each channel as the response of camera at the monochromatic light. The second column is one of four pixels in square region where the color code shows the Bayer pattern and the numbers show the digital numbers at those pixels in 14-bits. The third column is the demosaicked image.*

4.3 LED light source

4.3.1 Equipment

Fig.4.7 and 4.8 show the 10 LEDs-based light source and its spectra. The light source is originally developed for capturing photographic films (Trumpy et al.,

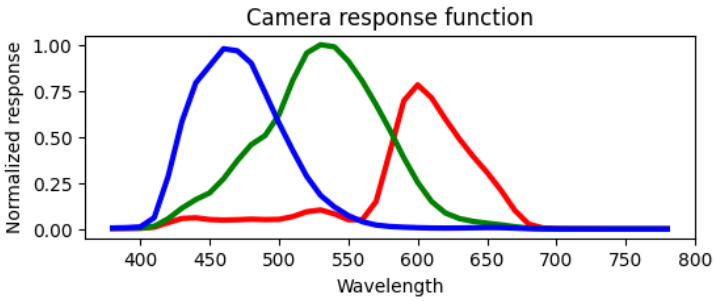


Figure 4.6: *CSRF for three channels of the camera. The color coding represents the measurements against the red, green, and blue channels.*

2021). Notice that the relative intensity of 6th and 7th LEDs are quite low in comparison with other LEDs. Table 4.2 shows the LEDs relative and absolute intensities along with at which wavelengths the peaks occur.

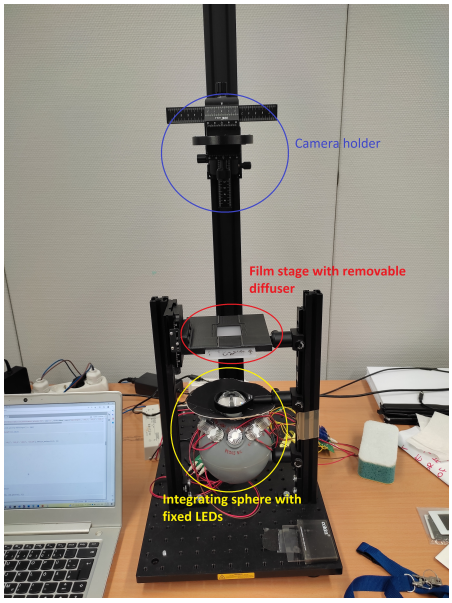
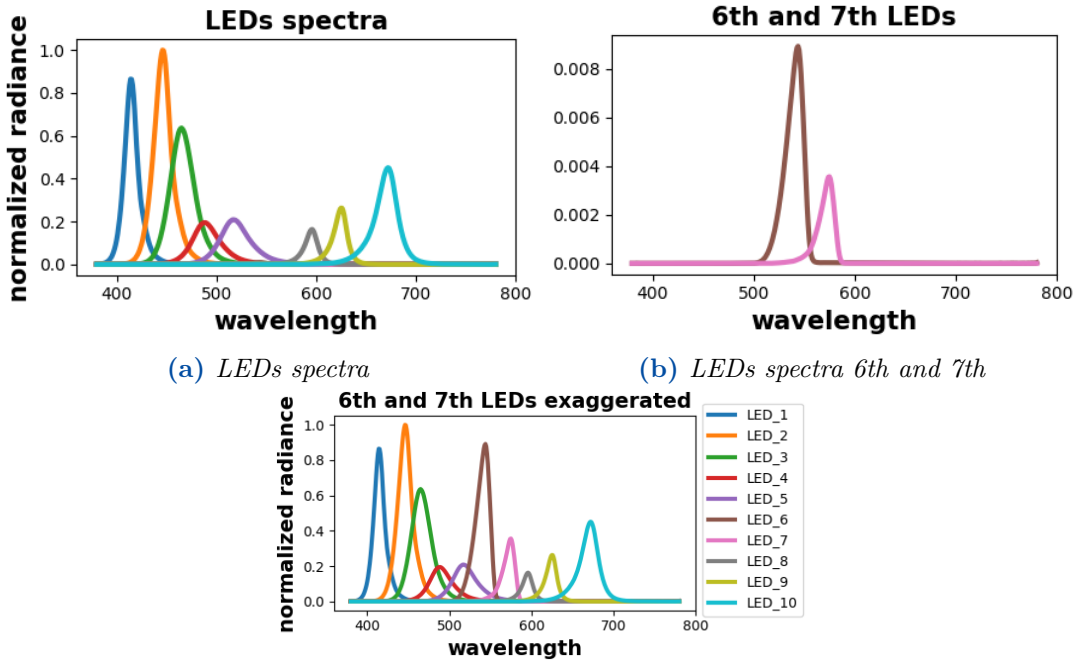


Figure 4.7: *LED imaging device. The yellow circle highlights the integrating sphere with 10 LEDs fixed across its radius. At the top of integrating sphere it has a black covering to stop the unwanted light from the integrating sphere leaking into the captures. At the center of black covering it has a collimating lens. In the red circle it has the stage where the photgraphic films can be fixed. It also has the option to slide in a diffuser. In the blue circle, the camera holder is shown.*



(a) *LEDs spectra*

(b) *LEDs spectra 6th and 7th*

(c) *LEDs spectra 6th and 7th exaggerated*

Figure 4.8: *(a) shows the 10 LEDs spectra of the light source. (b) shows the 6th and 7th LEDs seperately. (c) shows all the 10 LEDs spectra again but 6th and 7th LEDs scaled up 50 times.*

Table 4.2: *Table shows the wavelengths at which the LEDs peaks occur in the 10-LEDs light source. It also shows their relative and absolute intensities.*

LED	Peak wavelength (nm)	Relative peak value	Peak value ($Wsr^{-1}m^{-2}$)
1	416	0.88	0.0244621
2	448	1.00	0.0275599
3	467	0.663	0.0182769
4	491	0.183	0.00504516
5	520	0.198	0.0054696
6	545	0.014	0.00037581
7	572	0.003	0.0000795212
8	597	0.165	0.00455702
9	626	0.300	0.00827383
10	680	0.565	0.0155592

4.3.2 LEDs stability

In order to check the stability of LEDs over time, measurements are taken for each LED for 1 hour at a step size of 10 seconds. The results are plotted in Fig.4.9.

In all of the measurements, the wavelengths shift from lower to higher (in nm). Notice that the measurements for LED 10 nearly died down towards the end to an irrecoverable state. This LED was not replaced with a new one further on, because as it will be seen later on, LED 10 does not end up being used in the simulations and real captures because it does not contribute in minimizing the errors.

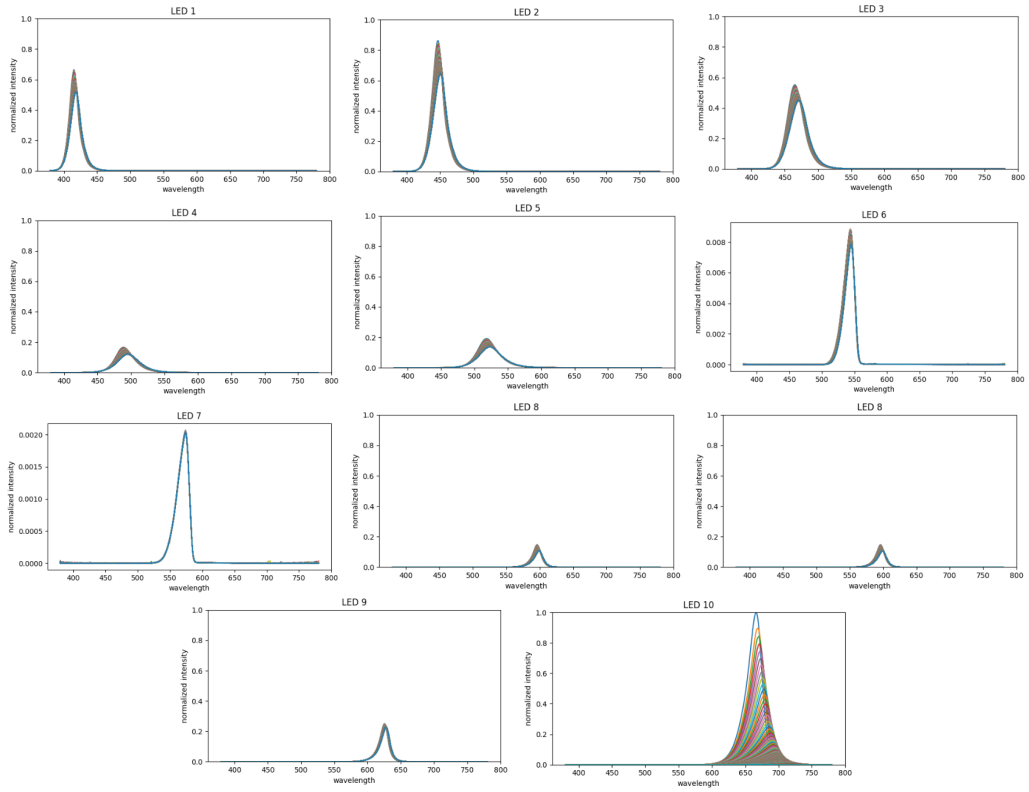


Figure 4.9: *10 LEDs spectra measured for 1 hour at an interval of 10 seconds. All LEDs shift their spectra towards higher wavelengths (in nm). LED 10 malfunctioned by the end of the measurements.*

The shift in peak wavelength for each LED with respect to the first peak wavelength when the LED is turned on is shown versus time in Fig.4.10. LED4 has the greatest shift of $7nm$. There is no change in peak wavelength for LED7. Also, for the first 10 seconds, there is no shift in wavelength peaks across all LEDs. This finding, along with the fact that LED 10 malfunctioned after a longer period of operation, encouraged the usage of LEDs within the first 10 seconds of their on-time. This is also consistent with how the original developers of the LED light source used it, where Trumpy et al. (2021) used the average value of the first twenty seconds measurements of LED spectra for the calculations.

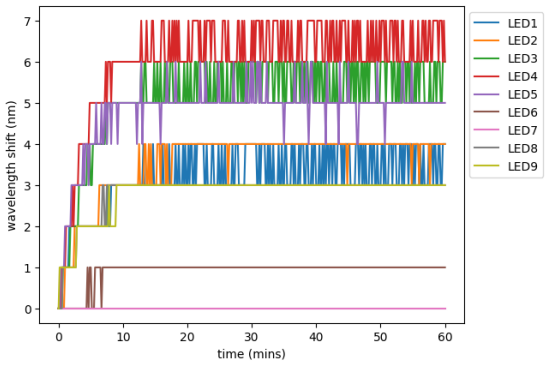
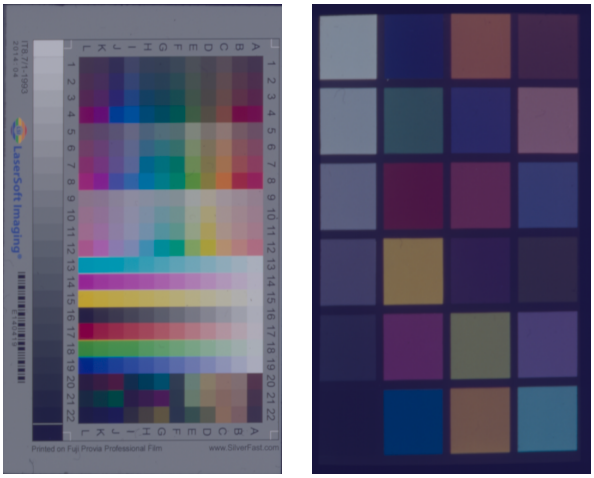


Figure 4.10: *Shift in peak wavelength of 10 LEDs w.r.t to the first peak wavelength measured when the LED is switched on plotted against time. The minimum difference that can be detected is of 1nm because the spectral resolution of the used TSR is 1nm.*

4.4 Training and testing

For training and testing photographic films of IT8 Colourchecker and Colourchecker24 are used respectively as shown in Fig.4.11. The transmittances of the colorcheckers have been calculated through various different setups which have been described in the appendix. This section describes the setup whose measurements were finalized for usage in the experiments and simulations. It is to be noted that the photographic films are 35mm films having the dimensions of 24mm x 36mm. The IT8 colourchecker has 288 patches in the small mentioned area. For measuring the transmittances using the conventional means of TSR is not possible. A setup is devised utilizing a overhead projector with halogen lamp. This setup allows to enlarge the colourchecker image making it possible to use TSR for the transmittance measurements.

The overall setup for transmittance calculation is shown in Fig.4.12 and the workflow for measurement is depicted in Fig.4.13. An overhead projector with a halogen lamp is used for projecting the photographic films on a neutral gray pattern tile. This setup allows for the enlargement of the film image by which we can use the TSR for the spectral measurements. Although the lamp was made to warm up for an hour, light source measurement for each of the patches measurements are done. This was deemed necessary to be extra sure about the correctness of the measured transmittances in case the light source spectrum changes significantly over time. In Fig.4.14, the measured transmittances of the IT8 colorchecker and colorchecker24 are shown.



(a) (b)

Figure 4.11: *(a) IT8 colourchecker used for training data. (b) Colourchecker24 used for testing data.*

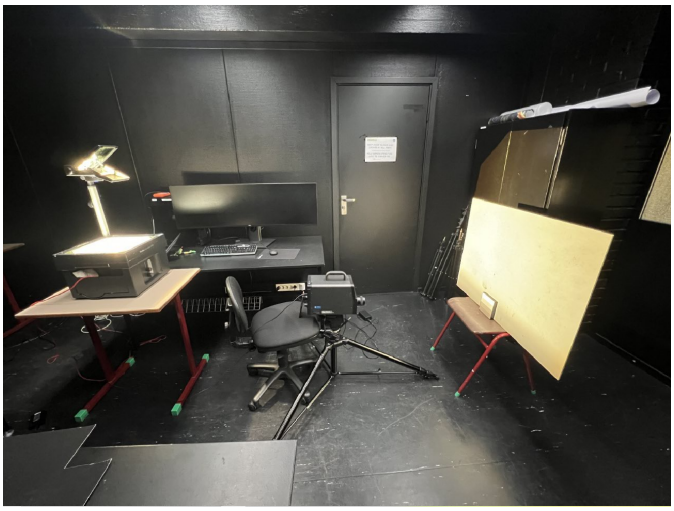
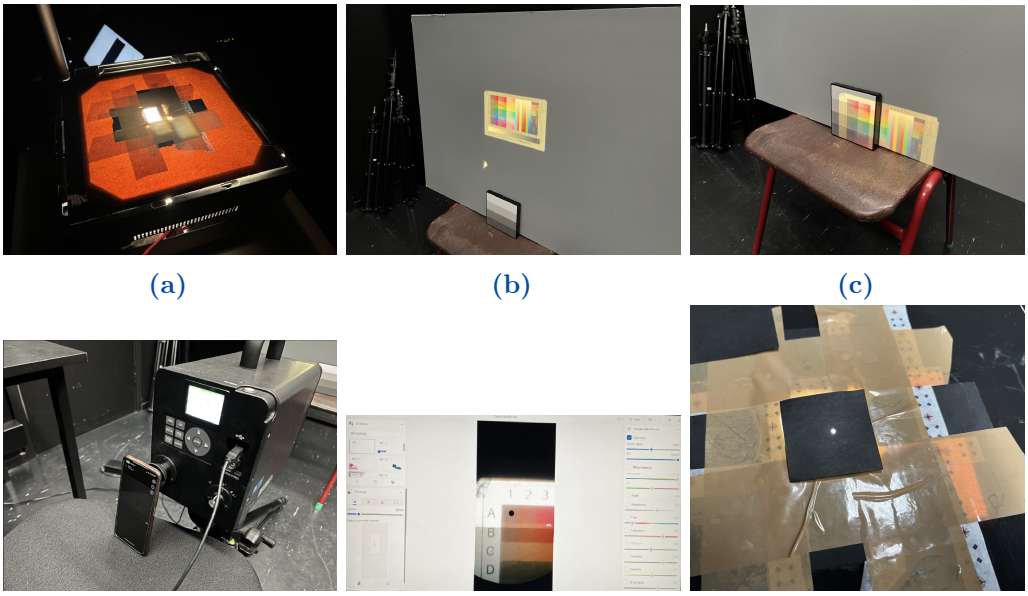


Figure 4.12: *Setup for measuring transmittances of photographic films. The overhead projector utilizing a halogen lamp illuminates a neutral gray pattern tile. The background of the tile is covered with a gray cardboard so the light does not reflect off the metal cabins at the back. The TSR is placed to take reading perpendicularly from the gray pattern tile.*

4.5 Function interpolations

The data described till now are all functions of wavelengths, but these are defined over a different range and increments of wavelengths. Table 4.3 summarizes the



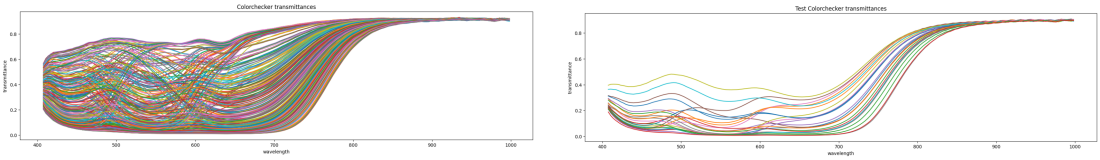
(a) (b) (c)

(d) (e) (f)

Figure 4.13: *(a) shows how the film is projected by placing it over an overhead projector. The film is fixed in a black paper cutting fit for the film size. This stops the redundant light other than the one passing from the film. A small opening is made beside the film cutting on the paper. This serves as an opening for light source measurements. The paper cuttings are made slide-able in order to take measurements of both the film and light source. (b) shows the projected colourchecker on the back cardboard for the sake of demonstration. Also notice the small opening of light below the colourchecker projection. This is the light source opening that was mentioned earlier. (c) shows the projected colourchecker on the gray pattern tile. (d) shows the setup of TSR where a phone's camera is set up to see through the eye piece. The phone's camera is connected to the computer over WiFi. This allows for the TSR measurement area to be visible while handling the film. (e) shows the interface where the TSR eye-piece images are forwarded to computer. The software used is Camo Studio (Studio, 2023). This software also allows to change image attributes like exposure time, hue, saturation. too. This helps in distinguishing the boundaries of the patches which are very similar. (f) shows another black paper cutting of the size of a single patch on colourchecker. This opening is used on top of the colourchecker patch to stop the redundant light from other patches.*

domains of the functions.

In order to work with these functions together, the domain of wavelength needs to be the same. The starting wavelength for all the functions is selected to be 410nm, ending wavelength to be 780nm, and at a step-size of 10nm. For this



(a) *IT8 colorchecker transmittances* (b) *Colorchecker24 transmittances*

Figure 4.14: *Transmittances of colorcheckers used for training and testing*

Table 4.3: *Table showing the region of wavelengths over which the CSRF, LEDs spectra, and colourchecker transmittances are defined.*

Function	From (nm)	To (nm)	Stepsize (nm)
CSRF	380.00	780.00	10.00
LEDs spectra	380.00	780.00	1.00
Colorchecker trasmittances	407.47	998.11	3.19

purpose linear interpolation is carried out. The most significant impact is on CSRF as shown in Fig.4.15, where it can be seen that the spectra of the first LED is cut in half.

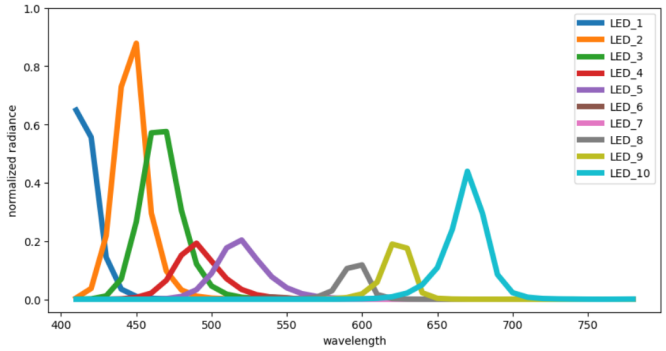


Figure 4.15: *CSRF after skipping values of 380, 390, and 400nm because of interpolation.*

4.6 Calculation of ground-truth tristimulus values

The ground-truth tristimulus values are calculated as:

$$X = \frac{100 \sum_{410}^{780} L(\lambda)T(\lambda)\bar{x}(\lambda)}{\sum_{410}^{780} L(\lambda)\bar{y}(\lambda)}, Y = \frac{100 \sum_{410}^{780} L(\lambda)T(\lambda)\bar{y}(\lambda)}{\sum_{410}^{780} L(\lambda)\bar{y}(\lambda)}, Z = \frac{100 \sum_{410}^{780} L(\lambda)T(\lambda)\bar{z}(\lambda)}{\sum_{410}^{780} L(\lambda)\bar{y}(\lambda)} \tag{4.2}$$

where X, Y, and Z are tristimulus values, L is the CIE D65 illuminant, $\bar{x}$, $\bar{y}$, $\bar{z}$ are 2° CMFs, and T is the transmittance.

5 Simulation

The best way to predict the future is to create it

Abraham Lincoln

The preceding chapter detailed what and how data for the simulation was acquired. The discussion of the simulation itself continues in this chapter. The simulation is performed for a pair of captures acquired by a 3-channel camera under two distinct light sources in order to minimise the ΔE_{00} error on the training data. Instead of directly simulating the final setup, this chapter builds up from the basic principles of imaging in order to see the different problems encountered during the simulation and appreciate their solutions.

5.1 3-channel imaging workflow

Firstly, the capturing workflow for a regular 3-channel camera is shown in Fig.5.1. Considering a light source emitting a radiance of $L(\lambda)$, onto an object with tristimulus values of X_1, Y_1, Z_1 (taken under CIE D65 by CIE 2° observer), and a reflectance of $R(\lambda)$. Also consider this reflected light is captured by a 3-channel camera with optical path trasmittance of $O(\lambda)$, a Bayer-pattern three filters $B(\lambda)$, sensor spectral response $S(\lambda)$, and giving rise to a triplet of R_1, G_1, B_1 values. The final values of R_1, G_1, B_1 generated can be calculated as follows:

$$c = \int_{\lambda_{min}}^{\lambda_{max}} L(\lambda)R(\lambda)O(\lambda)B(\lambda)S(\lambda)d\lambda \tag{5.1}$$

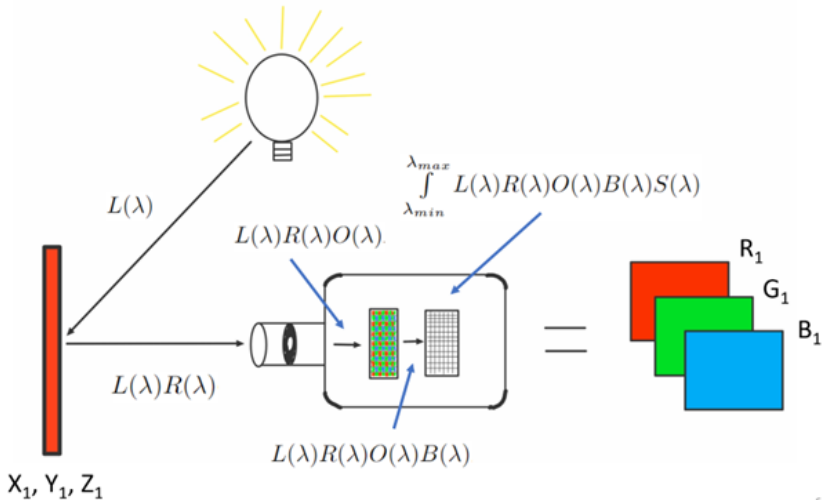


Figure 5.1: *3-channel imaging workflow. The light starts its journey from the light source towards a reflective object. The object reflects light according to its reflectance. The reflected light enters the camera through its lens and aperture. The light is filtered into three channels by the Bayer pattern filter and the transmitted light is recorded by the imaging sensor. Finally the missing pixel values are demosaicked which results in final three channels of the image.*

where $c \in \{R, G, B\}$.

If the same procedure is followed for N different uniformly colored patches, a mapping from RGB to XYZ values are formed for the patches. This allows to set up an equation where B is the tristimulus values, C is the RGB values, and A is the transformation matrix:

$$B = CA \tag{5.2}$$

$$\begin{bmatrix} X_1 & X_2 & \dots & X_n \\ Y_1 & Y_2 & \dots & Y_n \\ Z_1 & Z_2 & \dots & Z_n \end{bmatrix}^T = \begin{bmatrix} R_1 & R_2 & \dots & R_n \\ G_1 & G_2 & \dots & G_n \\ B_1 & B_2 & \dots & B_n \end{bmatrix}^T \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \tag{5.3}$$

Finding the transformation matrix A would have been straightforward by calculating $C^{-1}B$, but this requires C to be a square matrix, which is not the case in our situation. Nonetheless, the transformation matrix A' can be estimated which minimizes the squared difference between the predicted and ground-truth through Moore-Penrose pseudoinverse as (Penrose, 1956):

$$A' = (C^T C)^{-1} C^T B \tag{5.4}$$

Then using the estimation, the tristimulus values can be predicted as:

$$B' = CA' \tag{5.5}$$

$$\begin{bmatrix} X'_1 & X'_2 & \dots & X'_n \\ Y'_1 & Y'_2 & \dots & Y'_n \\ Z'_1 & Z'_2 & \dots & Z'_n \end{bmatrix}^T = \begin{bmatrix} R_1 & R_2 & \dots & R_n \\ G_1 & G_2 & \dots & G_n \\ B_1 & B_2 & \dots & B_n \end{bmatrix}^T \begin{bmatrix} a'_{1,1} & a'_{1,2} & a'_{1,3} \\ a'_{2,1} & a'_{2,2} & a'_{2,3} \\ a'_{3,1} & a'_{3,2} & a'_{3,3} \end{bmatrix} \tag{5.6}$$

Once the predictions are in place, the mean error ME can be calculated as:

$$ME = \frac{1}{n} \sum_{i=1}^{n} \Delta E_{00}(B_i, B'_i) \tag{5.7}$$

5.2 Extension to 6-channel multispectral imaging setup (dual-RGB imaging)

Though the spectral imaging provides a non-invasive technique for the inspection and recording of artworks, the conservators and museum curators have been able to utilize such setups in normal practices to a very limited cases. One reason for it is the installation and operational costs of these setups. Additionally, an expertise in the field of colour science and spectral imaging are often required for the operation of such setups. One method to lower down the barrier of spectral imaging utilization in cultural heritage context is called as dual-RGB imaging (Kuzio, 2023). In this section, the extension of the 3-channel imaging workflow to 6-channel multispectral imaging is discussed. Six channels are chosen because they have been shown to provide a good trade-off between additional channels and colour accuracy, where a 6-channel multispectral imaging setup provided comparable accuracy as a 10-channels setup in the context of cultural heritage (Kuzio and Farnand, 2021).

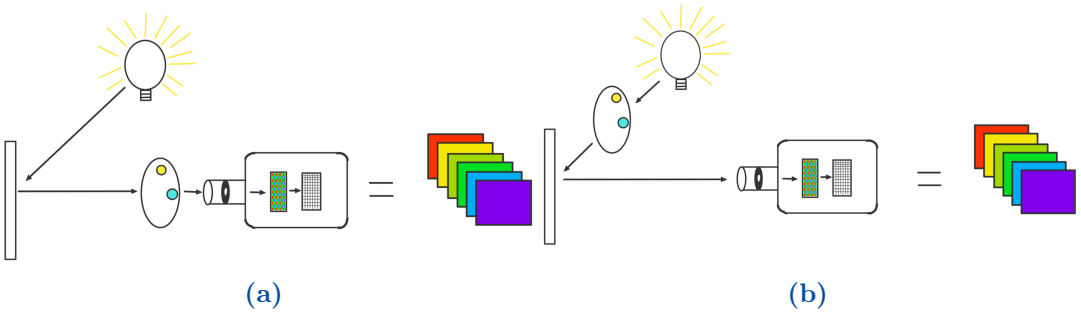
In order to extend the 3-channel imaging setup, discussed in last section, to 6-channel multispectral imaging setup, two filters can be introduced just before the light enters the camera. These two filters will be used to capture two different captures. This is called the passive multispectral imaging setup. This approach of filtering the incoming light into the camera through the usage of filters had been conceptualized by Berns et al. (2005) in the context of cultural heritage (Berns, 2005). Such a setup has been termed as dual-RGB imaging. Notice now the equation for the pixel value becomes

$$c = \int_{\lambda_{min}}^{\lambda_{max}} L(\lambda)R(\lambda)F(\lambda)O(\lambda)B(\lambda)S(\lambda)d\lambda \tag{5.8}$$

where $F(\lambda)$ is the filter transmittance.

It can be seen that the position of the factors of $R(\lambda)$ and $F(\lambda)$ can be interchanged without changing the equation which would convert it into an active multispectral setup. Both the setups of active and passive multispectral imaging are shown in Fig.5.2. The active setup can be written mathematically as:

$$c = \int_{\lambda_{min}}^{\lambda_{max}} L(\lambda)F(\lambda)R(\lambda)O(\lambda)B(\lambda)S(\lambda)d\lambda \tag{5.9}$$



(a) (b)

Figure 5.2: *6-channel imaging workflow. The six channels are achieved by the introduction of a filter wheel either after the light source or the reflective object. The filter wheel filters the light either directly from the light source or the reflected light. After which the light enters the camera through its lens and aperture upon which it is filtered into three channels by the Bayer pattern filter and then recorded by the imaging sensor. Finally the missing pixel values are demosaicked which results in three channels of the image. When this process is repeated for the other filter in filter wheel and the resulting three channels are stacked with the previously got three channels, a 6-channel image is formed.*

The approach of the replacement of optical filters with LEDs light sources has been shown to provide better colour accuracy than optical filters utilization in the context of cultural heritage (Kuzio and Farnand, 2022a). Moreover a dual-RGB based approach utilizing LEDs light source is also being employed in the context of cultural heritage into the open-source software termed Beyond RGB (Kuzio and Farnand, 2022b).

The filter wheel and the light source combinedly work to an equivalent setup where two different light sources are used. This motivates the replacement of both light source and filter wheel with a N-LED system (N=10 in this case) to provide two different light sources as shown in Fig.5.3. The figure shows that one capture is done using 6 LEDs as the light source and the second capture is done with 4 LEDs. The resulting camera response can be given as:

$$c_i = \int_{\lambda_{min}}^{\lambda_{max}} L_i(\lambda)R(\lambda)O(\lambda)B(\lambda)S(\lambda)d\lambda \tag{5.10}$$

where $c \in \{R, G, B\}$, and $i \in \{1, 2\}$.

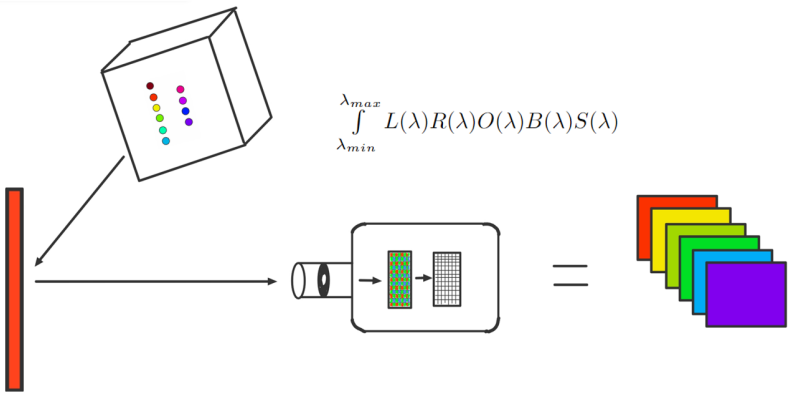


Figure 5.3: *6-channel imaging workflow with LEDs. The light source uses six LEDs to emit light towards a reflective object. The object reflects light according to its reflectance. The reflected light enters the camera through its lens and aperture upon which it is filtered into three channels by the Bayer pattern filter and then recorded by the imaging sensor. Finally the missing pixel values are demosaicked which results in three channels of the image. When this process is repeated with the light source using the other four-LEDs as a light source and the resulting three channels in camera are stacked with the previously got three channels, a 6-channel image is formed.*

The pair of the RGB images which are captured with two different LED combinations are called as a dual-RGB capture. Notice that in all of these setups a mapping from 6-channel RGB images to tristimulus values are formed which can be written mathematically as:

$$B = CA \tag{5.11}$$

$$\begin{bmatrix} X_1 & X_2 & \dots & X_n \\ Y_1 & Y_2 & \dots & Y_n \\ Z_1 & Z_2 & \dots & Z_n \end{bmatrix}^T = \begin{bmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,n} \\ G_{1,1} & G_{1,2} & \dots & G_{1,n} \\ B_{1,1} & B_{1,2} & \dots & B_{1,n} \\ R_{2,1} & R_{2,2} & \dots & R_{2,n} \\ G_{2,1} & G_{2,2} & \dots & G_{2,n} \\ B_{2,1} & B_{2,2} & \dots & B_{2,n} \end{bmatrix}^T \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ \vdots & \vdots & \vdots \\ a_{6,1} & a_{6,2} & a_{6,3} \end{bmatrix} \tag{5.12}$$

Estimating the transformation using the pseudo inverse similar to last section:

$$B = CA \tag{5.13}$$

$$\begin{bmatrix} X'_1 & X'_2 & \dots & X'_n \\ Y'_1 & Y'_2 & \dots & Y'_n \\ Z'_1 & Z'_2 & \dots & Z'_n \end{bmatrix}^T = \begin{bmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,n} \\ G_{1,1} & G_{1,2} & \dots & G_{1,n} \\ B_{1,1} & B_{1,2} & \dots & B_{1,n} \\ R_{2,1} & R_{2,2} & \dots & R_{2,n} \\ G_{2,1} & G_{2,2} & \dots & G_{2,n} \\ B_{2,1} & B_{2,2} & \dots & B_{2,n} \end{bmatrix}^T \begin{bmatrix} a'_{1,1} & a'_{,12} & a'_{1,3} \\ a'_{2,1} & a'_{2,2} & a'_{2,3} \\ \vdots & \vdots & \vdots \\ a'_{6,1} & a'_{6,2} & a'_{6,3} \end{bmatrix} \tag{5.14}$$

Now the mean error can be calculated as in Eq.5.7.

5.3 Dual-RGB imaging for photographic films

In the context of this thesis work, the setup of dual-RGB imaging works on transmissive photographic films. So the actual working is a slight modified version as shown in Fig.5.4. The difference is the replacement of reflectance $R(\lambda)$ with transmittance $T(\lambda)$.

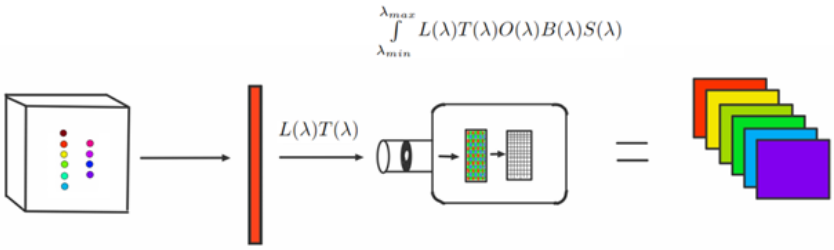


Figure 5.4: *Dual-RGB imaging workflow for transmissive objects. The only difference between this and Fig.5.3 is the replacement of reflective object with transmissive film. The light source uses six LEDs to emit light towards a transmissive object. The object transmits light according to its transmittance. The reflected light enters the camera through its lens and aperture upon which it is filtered into three channels by the Bayer pattern filter and then recorded by the imaging sensor. Finally the missing pixel values are demosaicked which results in three channels of the image. When this process is repeated with the light source using the other four-LEDs as a light source and the resulting three channels in camera are stacked with the previously got three channels, a 6-channel image is formed.*

5.3.1 Which LED combination to choose?

5.3.1.1 Iterating over each LED combination

The very first question to answer for moving forward is to decide on which LED light source pair one should choose. The following criterion is defined for selecting the best LED light source pair:

Criterion: *Choose the light source pair which minimizes the mean ΔE_{00} error on the training set.*

A straightforward solution will be to iterate over each LED light source pair possible and check the mean ΔE_{00} error and select the one which gives the minimum error. Running a crude approximation for the total number of LED pairs: There are 10 LEDs in a light source which makes it 2^{10} different ways to define a single light source. Since two light sources need to be defined, this makes $2^{10} \cdot 2^{10} \approx 1$ million different combinations of light sources. This poses a practical challenge for running the simulations. Hence a better strategy needs to be defined to choose the LED pair for evaluation.

5.3.1.2 Reducing the number of LED combinations to evaluate

The efficacy of a multispectral imaging setup is highly dependent upon the amount of new information that each channel carries. This gives a hint on what combinations of light source pairs are more probable to provide reduced errors. The effective sensitivity of the imaging system is defined as:

$$S_{eff} = L(\lambda)O(\lambda)B(\lambda)S(\lambda) \tag{5.15}$$

The LED spectra are divided into three bins having 3, 3, and 4 LEDs from shortest to longest wavelength as shown in Fig.5.5. A single light source is defined by an LED triplet where each of the LEDs are taken from each of the bin defined: one bin, one LED. The motivation for doing this is if two LEDs are chosen from a single bin, it does not add much new information that is expected from a new separate channel in the imaging system. This approach is similar to the one employed by Kuzio and Farnand (2022b) for the reduction of LEDs light sources combinations in a 10-LEDs based system.

Using this LED pair combination reduction scheme, A single light source can be defined by 3 LEDs which can be chosen in $3 \cdot 3 \cdot 4 = 36$ different ways. Two light sources can be defined by $36 \cdot 35 = 1260$ different ways. Here 35 is used because the pair used as first light source should not be used as the second light source i.e. two light sources should not be same. Since order of light sources does not matter, the total number of light source combinations are $1260/2 = 630$. Hence the total

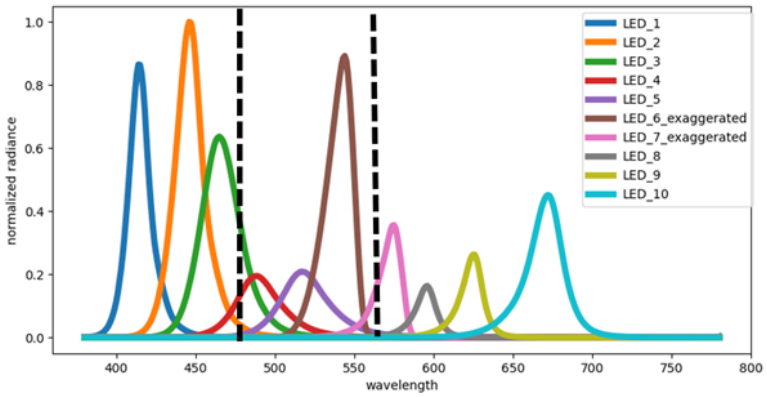


Figure 5.5: *Spectra of the 10-LEDs system (see Chapter 4) is shown with LED 6 and 7's spectra scaled up 50 times. The dotted black vertical lines represent the boundaries of the bins selected. Each bin have 3, 3, and 4 LEDs. A single light source can be defined by taking one LED from each of the bin.*

number of possible light source pairs combinations reduced from 1 million to only 630. This allows to change the previous light source pair selection criterion to:

Criterion: *Choose the pair of LED triplets from the three defined bins which minimizes the mean ΔE_{00} error on the training set.*

The above described workflow is summarized in Fig.5.6.

5.4 Simulation results

5.4.1 Optimal light sources

After simulating the errors from all 630 pairs of light sources, ΔE_{00} is calculated for each of them. The best three combinations are shown in Table.5.1 and the corresponding light source plots in Fig.5.7.

Table 5.1: *The table shows the best pair of light sources defined by two LED triplets. The table also shows the training error calculated by using each light source.*

Light source 1	Light source 2	Error
(1, 6, 8)	(3, 5, 9)	0.31
(1, 6, 8)	(3, 5, 10)	0.35
(1, 6, 9)	(3, 4, 8)	0.37

The minimum error is 0.31 from light sources pair formed from LEDs (1, 6, 7)

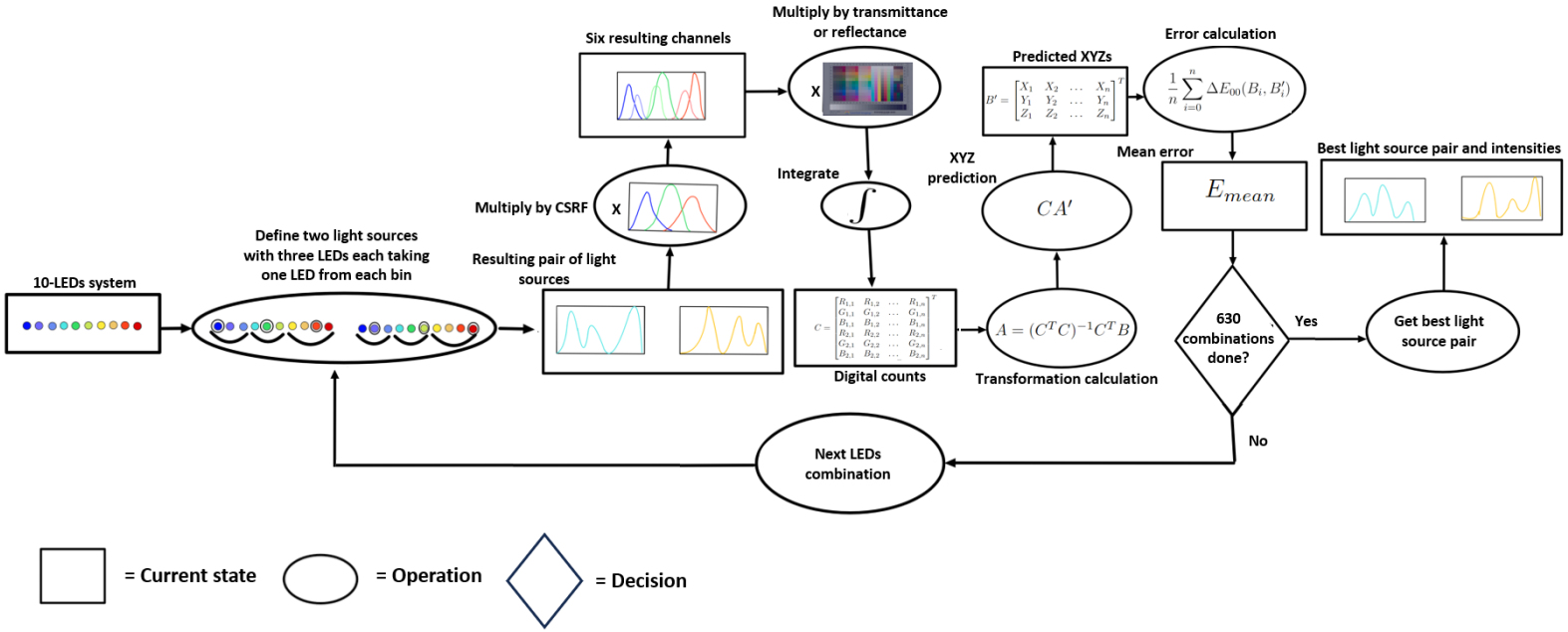
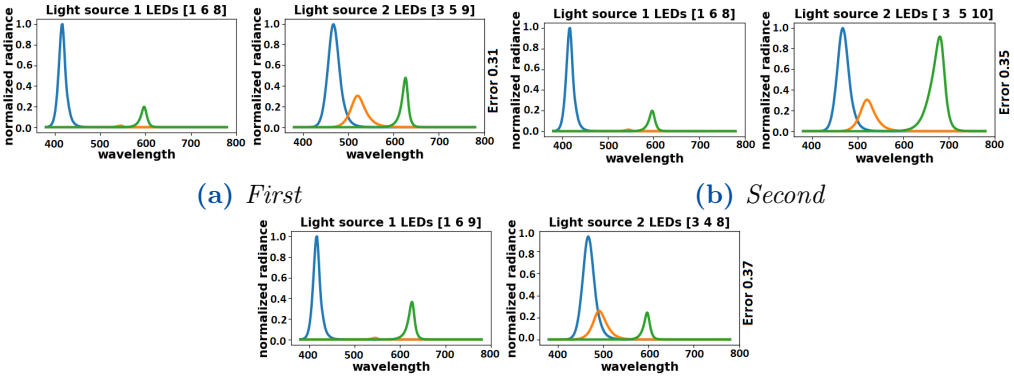


Figure 5.6: *Dual-rgb imaging workflow. Two light sources are defined from the 10-LEDs light source in such a way that each light source comprises of three LEDs each of which are taken from each bin. The defined light sources are mutliplied by the CSRF to get six channels of the imaging setup. These channels are multiplied with the trasmittances of each of the patches of the colourchecker and integrated to simulate the digital numbers captured by the camera. The transformation matrix is estimated by using the pseudo-inverse. This transformation matrix is then used to predict the tristimulus values of the patches. The ΔE_{00} difference can be calculated for each of the patches using the ground-truth tristimulus values and can be averaged to provide mean error. This process is done for all the 630 combinations of light source pairs. Finally the pair of light source is selected which gives the least mean error.*

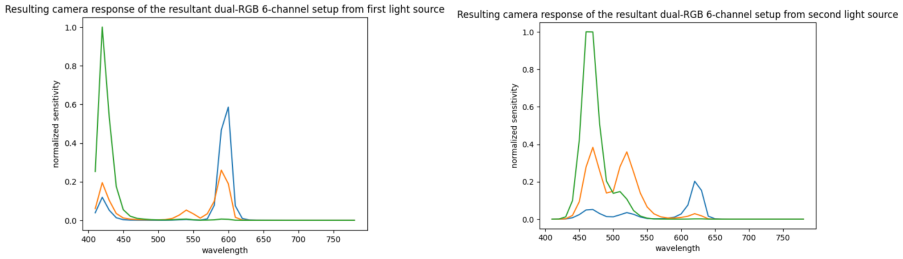
and (3, 5, 9). The test error for this pair of light sources on Colourchecker24 is 0.66. The final channels of the imaging system is the product of the pair of light sources with CSRF which is shown in 5.8.



(a) *First* (b) *Second*

(c) *Third*

Figure 5.7: *The best pair of light sources defined by two LED triplets. The light source combinations are mentioned in the title of each plot. The training errors calculated by using each of the light source is also mentioned on the right side of each plot pair.*



(a) *Camera response light source 1* (b) *Camera response light source 2*

Figure 5.8: *The resulting channel sensitivities for the best pair of light source.*

5.4.2 Error on patches

For the visualization of errors in this report, the colour patches are converted to sRGB colour space. The sRGB colour space is a widely used colour space across electronic devices, making it important to know the chromaticities of the patches with respect to the gamut of sRGB (W3C, 2022). Note that a gamut of the colour space is a three-dimensional construct, containing both the information of chromaticity and luminance. But in this thesis work, the colour specifications

are normalized by the luminance of the white patch of the training colourchecker. Hence only the chromaticities are compared for the gamut comparison. Fig.5.9 shows the chromaticities of both training and testing data as well as the gamut of sRGB. The figure shows that there are 42 data points in training and one in testing which are out of gamut for this sRGB visualization. Especially for these out-of-gamut points, it is better to only refer to the numerical errors.

Moreover, in order to have a comparison with the dual-RGB imaging setup, simulations are conducted for a three-channel imaging setup under the CIE D65 illuminant. The three-channel simulations can be processed as discussed in Section 5.1 and Eqs.5.2-5.7.

The Fig.5.10 and Fig.5.14 shows the ΔE_{00} for each of the training and testing patches as well as mean ΔE_{00} values along with the results of the three-channel imaging workflow. The advantage of dual-RGB imaging over three-channel imaging is evident where the training mean ΔE_{00} error reduced from 1.25 to 0.31 (75% reduction) and the testing mean ΔE_{00} error from 1.53 to 0.66 (57% reduction). This is also visible by the colour coding visualization of errors where the colour code varies from brown, red, orange, yellow, to green for highest to lowest errors. It can be observed that the dual-RGB has more green codes and three-channel has more red and yellow codes assigned. For the neutral patches, a trend is visible where the error generally decreases as the lightness of the patch increases.

The Fig.5.11 and Fig.5.15 shows a reproduction of ground-truth, dual-RGB predicted, and three-channel predicted colours in sRGB colour space for training and testing data. Each patch consists of four sub-blocks in the form $\left(\begin{smallmatrix} GT & TC \\ GT & DR \end{smallmatrix}\right)$, where GT is for ground truth, TC for three-channel, and DR for dual-RGB sub-block. The yellow boundaries represent the out-of-gamut colours for the sRGB colour space. As mentioned before, the difference in patches cannot be visualized for these out-of-gamut colours, limiting the observations to only their ΔE_{00} values for these patches.

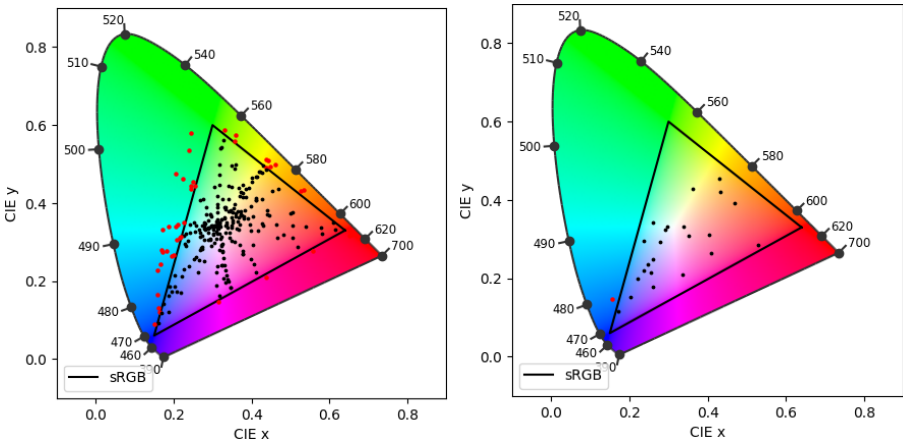
Moreover, a generally good correlation is observed between the ΔE_{00} errors and perceivable differences in colour patches. The Fig.5.13 shows five example patches taken from the Fig.5.11. The ΔE_{00} errors are mentioned below each block. The errors are taken between the ground-truth and three-channel predictions. The patches are selected such that the errors are in approximate increments of one ΔE_{00}. This figure highlights a good correlation between the perceivable difference and the ΔE_{00} errors. This signals to the correctness of the parameters used for the ΔE_{00} in this thesis work.

The Fig.5.12 shows plot for Lightness vs. Chroma and a^* vs. b^* in CIELAB colour space for training data, where the positions of dots represent the ground-truth data. The plots show both colour codes for errors as well as colour code representing the sRGB reproduction of colour patches. Considering the Lightness

vs. Chroma plot elaborates more clearly on the trend that was mentioned before that the errors generally decrease as the lightness increase. It can be seen that as the plot is traversed from low lightness to high lightness values, the errors generally decrease. This plot highlights another factor into this observation, that the errors also decrease with decreasing chroma. This means that the most light and the least chromatic colours perform the best in terms of ΔE_{00} errors. These observations are true for both three-channel and dual-RGB setup.

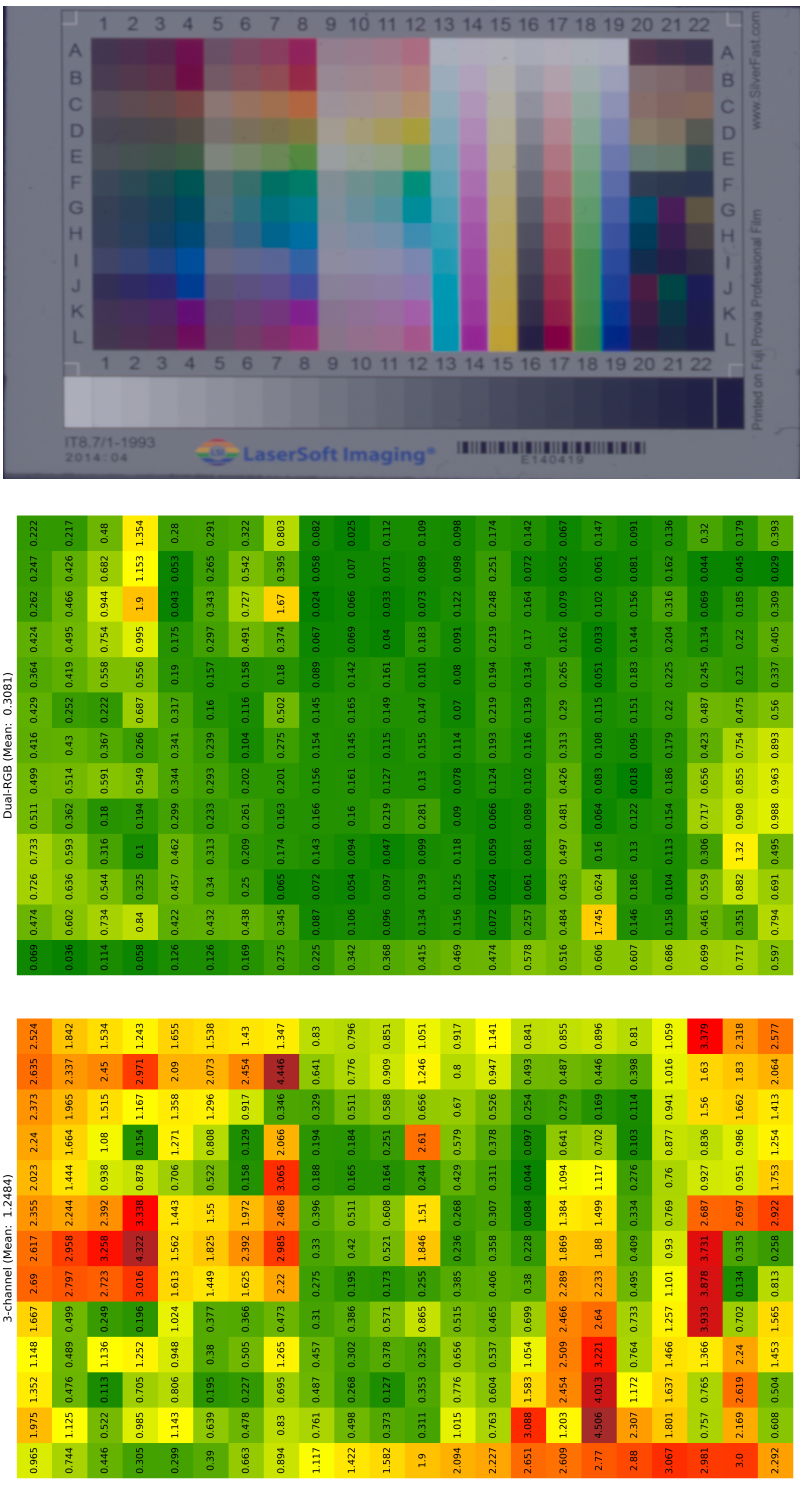
Looking at the a^* vs. b^* plots for training, more elaboration on the previous observations are found. Previously it was seen that increasing the chroma also generally increased the colours. The a^* vs. b^* plots show that these errors tend to fall more on the side of bluish and reddish colours. It can also be seen that errors generally increase moving out from the center which is also consistent with increasing errors with increasing chroma observation.

The Fig.5.16 show similar plots for Lightness vs. Chroma and a^* vs. b^* for testing data. For testing data no obvious patterns are visible which are as much visible as it was for training data. This might be because of the low number of colour patches in testing data.



(a) (b)

Figure 5.9: *Data points on chromaticity diagram with a comparison to sRGB colour space. (a) shows the chromaticity points for training data and (b) shows the chromaticity points for testing data. The black points correspond to in-gamut and red points correspond to out-of-gamut chromaticities.*



(a) *Train error 3-channels* (b) *Train error dual-RGB* (c) *Train colorchecker*

Figure 5.10: *Comparing training errors for each patch in training for 3-channels and dual-RGB. The color code goes from brown → red → orange → yellow → green. The mean errors are shown above each plot. For reference the original colourchecker image is also shown.*

Figure 5.11: *Train data reproduction from groundtruth data and predictions from dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\begin{pmatrix} GT & TC \\ GT & DR \end{pmatrix}$*, where GT is for ground truth, TC for three-channel, and DR for dual-RGB block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. The yellow boundary represents out-of-gamut-colour for sRGB colour space*

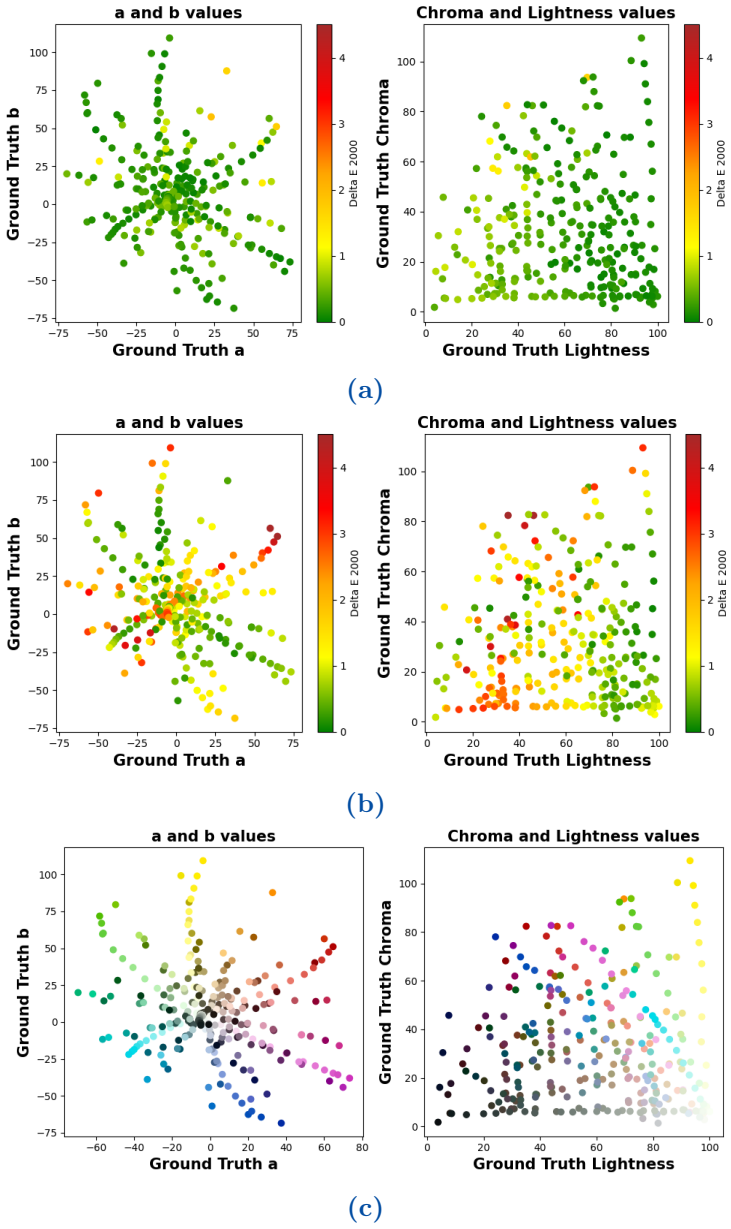


Figure 5.12: *Plots for dual-RGB and three-channel training patches. (a) shows a^* vs. b^* and Lightness vs. Chroma plots for dual-RGB.(b) represents the same set of plots for three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a^* vs. b^* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

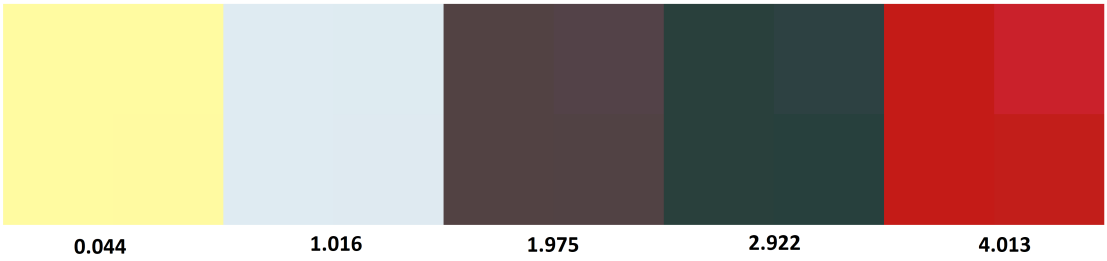


Figure 5.13: *The figure shows five example patches from Fig.5.11, where each patch block is made up of four sub-blocks in the form* $\left(\begin{smallmatrix} GT & TC \\ GT & DR \end{smallmatrix}\right)$*, where GT is for ground truth, TC for three-channel, and DR for dual-RGB block. The* ΔE_{00} *errors are mentioned below each block. The errors are taken between the ground-truth and three-channel prediction. The patches are chosen such that the errors are in approximate increments of* ΔE_{00}*. This highlights the correlation between the* ΔE_{00} *errors and perceptible differences.*

5.5 Discussion

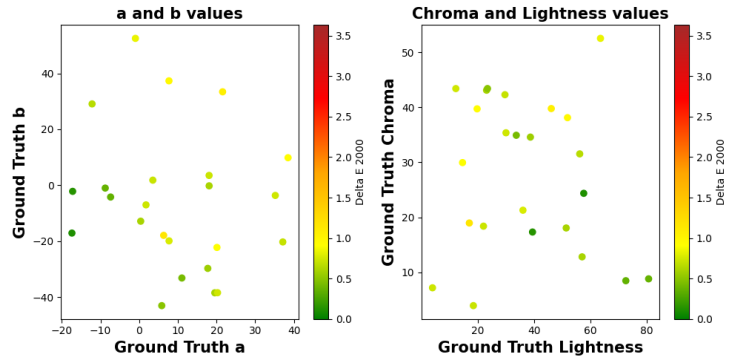
This chapter discussed the workflow for a three-channel imaging and 6-channel multispectral imaging (dual-RGB). The approach for predicting a ground-truth tristimulus values from the simulated digital numbers from the camera was elaborated. Different approaches for dual-RGB were discussed including active and passive multispectral imaging using filters and multispectral imaging with LEDs-based light source. The approach of LEDs-based light source was further elaborated which was used with transmissive photographic films. It was discussed how the different combinations of the 10-LEDs could generate around a million different pair of light sources for dual-RGB captures. These combinations were reduced to 630, by dividing the LEDs into three bins of 3, 3, and, 4 LEDs. This limited the number of combinations of light source pairs because now one light source could only be created by a total of three LEDs where each LED is selected from different bin.

All the different 630 combinations of light sources were simulated and the one giving the least ΔE_{00} mean error was selected as the final light sources pair. The mean error for training was 0.31 and for testing it was 0.66. These errors were compared with the mean error from a three-channel imaging which gave 1.25 for training and 1.53 for testing. This showed that the dual-RGB approach reduced the errors by 75% for training and 57% for testing. Looking at the individual errors of patches in Fig.5.10 and Fig.5.14, a general trend was seen for neutral patches that as the lightness increased, the errors were reduced. This is sensible since more lightness would correspond to more light being transmitted from the film increasing the signal to noise ratio for the capture.

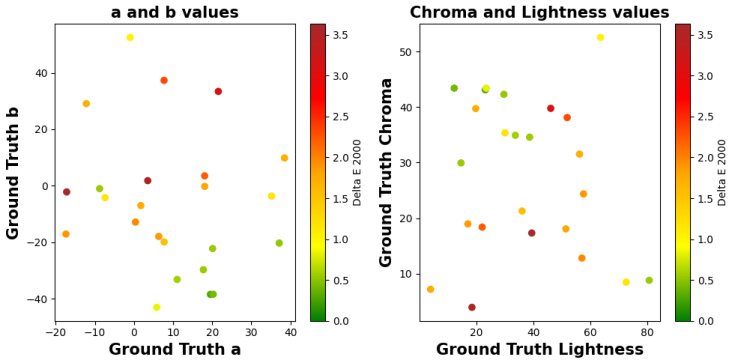
(a) *Test colorchecker* (b) *Test error 3-channels* (c) *Test error dual-RGB*

Figure 5.14: *Comparing testing errors for each patch in training for 3-channels and dual-RGB. The color code goes from brown → red → orange → yellow → green. The mean errors are shown above each plot. For reference the original colourchecker image is also shown.*

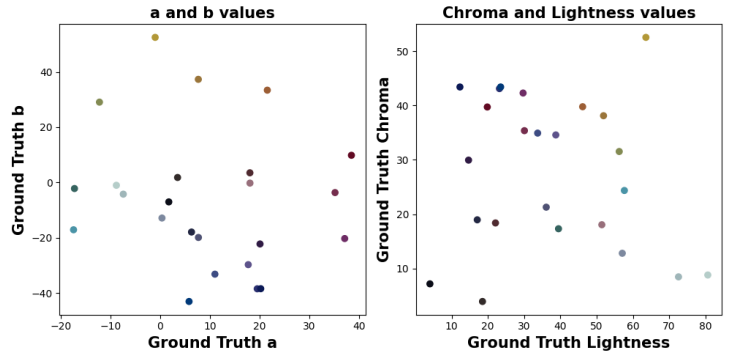
Figure 5.15: *Test data reproduction from groundtruth data and predictions from dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\begin{pmatrix} GT & TC \\ GT & DR \end{pmatrix}$*, where GT is for ground truth, TC for three-channel, and DR for dual-RGB block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. The yellow boundary represents out-of-gamut-colour for sRGB colour space*



(a)



(b)



(c)

Figure 5.16: *Plots for dual-RGB and three-channel testing patches. (a) shows a^* vs. b^* and Lightness vs. Chroma plots for dual-RGB.(b) represents the same set of plots for three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a^* vs. b^* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

The sRGB visualizations of the ground-truth and the predicted patches are also shown in Fig.5.11 and Fig.5.15 along with the highlighting of colours which were out-of-gamut for sRGB colour space. These figures showed the perceivablity of the difference in three-channel and dual-RGB imaging. The three-channel predictions gives a perceivable different colour especially for reddish and bluish patches. The dual-RGB approach generally gives imperceivable results from the ground-truth even for reddish and bluish regions.

The Fig.5.12 and Fig.5.16 showed the CIELAB a^* vs. b^* and Lightness vs. Chroma plots for the training and testing data. These plots further confirmed the obeservation of decreasing errors as the lightness increased. These plots also showed another pattern of decreasing error with decreasing chroma. This can be attributed to the fact that for less chromatic colours, more channels participate in providing the signal hence increasing the signal to noise ratio. These trends are not obvious for testing data, one reason might be because of the low number of colour patches for testing.

With all these different plots and visualizations, it is clear that in simulations, dual-RGB performs better than a regular three-channel imaging workflow in terms of colour accuracy.

6 | Exploring change of intensities

To improve is to change; to be perfect is to change often.

Winston Churchill

In the previous approach for selecting the pair of LED triplets, the LEDs were considered at their full intensities. But a single pair of LED triplets can have a varied intensity, giving very different results on the colour difference. As an example, it is shown here how much of a difference changing the intensity of one LED can make on the error. As it was found out previously, the best pair of LED triplets are (1, 6, 8) and (3, 5, 9). As an example, the intensity of LED 5 can be varied from 0 (no intensity) to 1 (full intensity) at a step-size of 0.01 as shown in Fig6.1. Fig.6.2 shows that the optimal intensity for LED 5 (keeping the intensity of other LEDs at full) is 0.57, which gives an error of 0.29 in comparison to 0.31 at full intensity.

Although the error might not have been reduced to a significant level for this particular LED and its combination, this definitely gives a hint of the potential of exploring changing intensities of LEDs for the reduction of errors. For this reason, for each of the light source pair combinations, the varying intensities for each LED are also explored to search for the optimal LED intensities that give the minimum error. For finding these optimal intensities, Particle Swarm Optimization (PSO) has been implemented. The PSO is selected because of its simplicity and because it gives the best success rate and solution quality and the second best processing time in continuous and discrete optimization tests competing against the four other evolutionary algorithms, i.e., the genetic algorithm, memetic algorithm, ant-colony systems, and shuffled frog leaping (Elbeltagi et al., 2005).

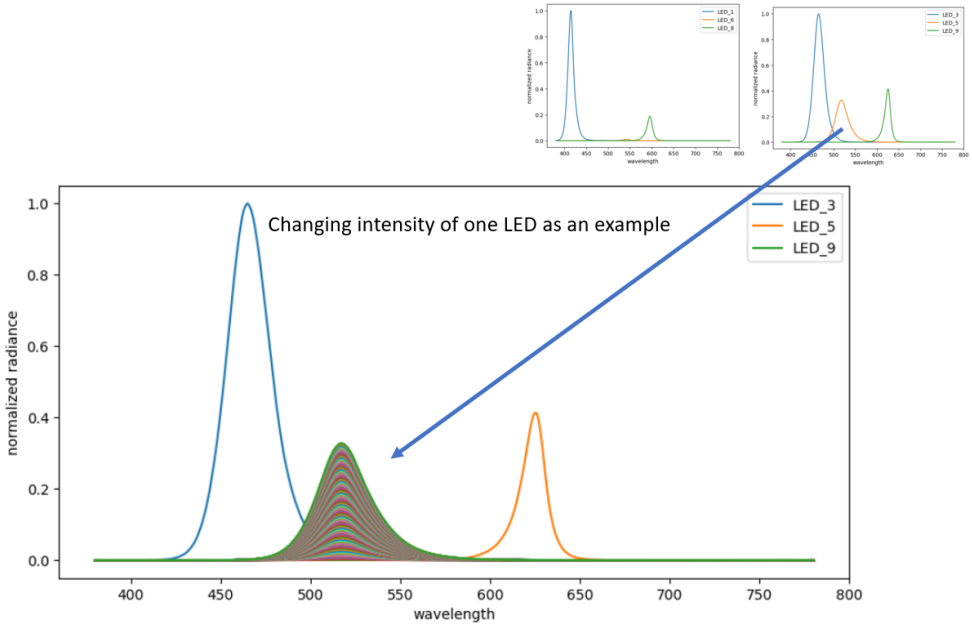


Figure 6.1: *From the best pair of light sources selected in the Chapter 5, the intensity of LED5 is changed from 0 (off) to 1 (full intensity) at a step size of 0.01 while keeping the rest same.*

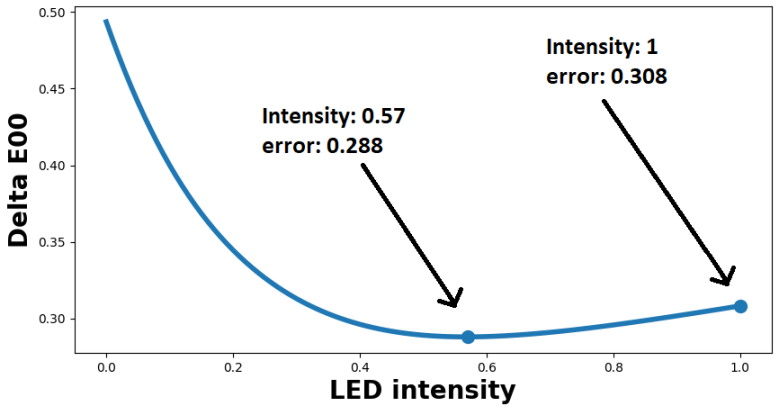


Figure 6.2: *The plot shows the mean ΔE_{00} error found for the best light source pair found in Chapter 5 by changing the intensity of LED 5 from 0 (off) to 1 (full intensity) while keeping the rest same. The error is found out to be 0.29 when the intensity is 0.57 in comparison to 0.31 when it was full.*

6.1 Particle Swarm Optimization (PSO)

PSO is part of a larger class of algorithms called as evolutionary algorithms. Evolutionary algorithms are stochastic searching methods that are inspired by either the biological evolution of species or their social behaviour, or both. This is analogous to ants finding the shortest path to food or birds finding the destination in migration (Lovbjerg, 2002). This behaviour for finding the target is analogous to optimization in computational problems. This way of optimization is simulated computationally in the category of evolutionary algorithms. In literature, genetic algorithms were the first evolutionary algorithms proposed based on the principles of survival of the fittest and reproduction. Genetic algorithms have demonstrated their ability to solve a large set of problems in both science and engineering. An example of it is the 'evolved antenna' called X-band in NASA's ST5 Mission, whose design is far from being proposed by an experienced antenna designer, giving maximum data throughput with lesser power requirements (Hornby et al., 2006). With the benefits also comes the time complexity that genetic algorithms have. This leads to the creation of other evolutionary algorithms that can be faster and even provide better solutions; PSO is such an algorithm (Elbeltagi et al., 2005).

Kennedy and Eberhart (1995) proposed PSO as being inspired by the social behaviour of birds migrating to an unknown destination. In PSO, each solution is considered a bird and called a particle. The whole flock is called a swarm. Each particle has its own intelligence, which allows it to search its local space for solutions. The whole swarm has a combined intelligence that allows each of its particles to direct themselves towards the best solution found by the swarm. Each bird has its own location, direction, and speed of flight. After getting information about the best solution found by the flock, the particle updates its location which gets affected by the best solution this particle has found, its speed and direction, and the best solution found by the whole swarm.

6.1.1 Algorithm

The PSO is summarized in Algorithm 1 and the parameters are explained below:

1. C_0 refers to the inertia of particle. $C_0 < 1$ implies deceleration of particle and $C_0 > 1$ implies accleration.

2. C_1 refers to the weightage given to the cognition of the individual particle.

3. C_2 refers to the weightage given to the global cognition of the swarm.

4. L_{max} is maximum number of iterations for optimization.

5. N is total number of particles in swarm for optimization.

Data: $L_{\max}, N, C_0, C_1, C_2$
Result: $x = [\mu_1, \mu_2, ..., \mu_M]$
for n *in* $[1, \ldots, L_{max}]$ **do**
  for i *in* $[1, \ldots, N]$ **do**
    Randomize $\mathbf{r}_1$ and $\mathbf{r}_2$
    Update velocity: $\mathbf{v}_i \leftarrow$
    $C_0\mathbf{v}_i + C_1\mathbf{r}_1 \otimes [\mathbf{x}_{i,\text{ best}} - \mathbf{x}_i] + C_2\mathbf{r}_2 \otimes [\mathbf{x}_{\text{best}} - \mathbf{x}_i]$
    Update position: $\mathbf{x}_i \leftarrow \mathbf{x}_i + \mathbf{v}_i$
    Enforce constraints for $\mathbf{x}_i$
    Calculate value of cost function: $c_i = \text{COST}(\mathbf{x}_i)$
    if $c_i < c_{i,\ best}$ **then**
      $c_{i,\text{ best}} \leftarrow c_i$
      $\mathbf{x}_{i,\text{ best}} \leftarrow \mathbf{x}_i$
    end
    if $c_i < c_{\ best}$ **then**
      $c_{\text{ best}} \leftarrow c_i$
      $\mathbf{x}_{\text{ best}} \leftarrow \mathbf{x}_i$
    end
  end
end

Algorithm 1: Algorithm for Particle Swarm Optimization

6.2 PSO for LED intensities optimization

In the dual-RGB setup for this thesis, there are 630 different combinations of light sources that can be created by utilizing the three bins of LEDs. In each of these 630 combinations, there is a pair of LED triplets that define the light source pair. These 630 combinations of light source pairs are created by considering each LEDs at their peak illumination. A big part of the solution space is skipped for the minimization of ΔE_{00} by not considering the change of the illumination level for each of the LED involved in the LEDs triplet pair. In order to optimize the LEDs illumination level in a pair of LED triplets, the problem needs to be formulated in such a way that it can be used in the PSO algorithm. The algorithm is detailed in Fig.6.3 and the problem formulation is defined in subsequent subsections.

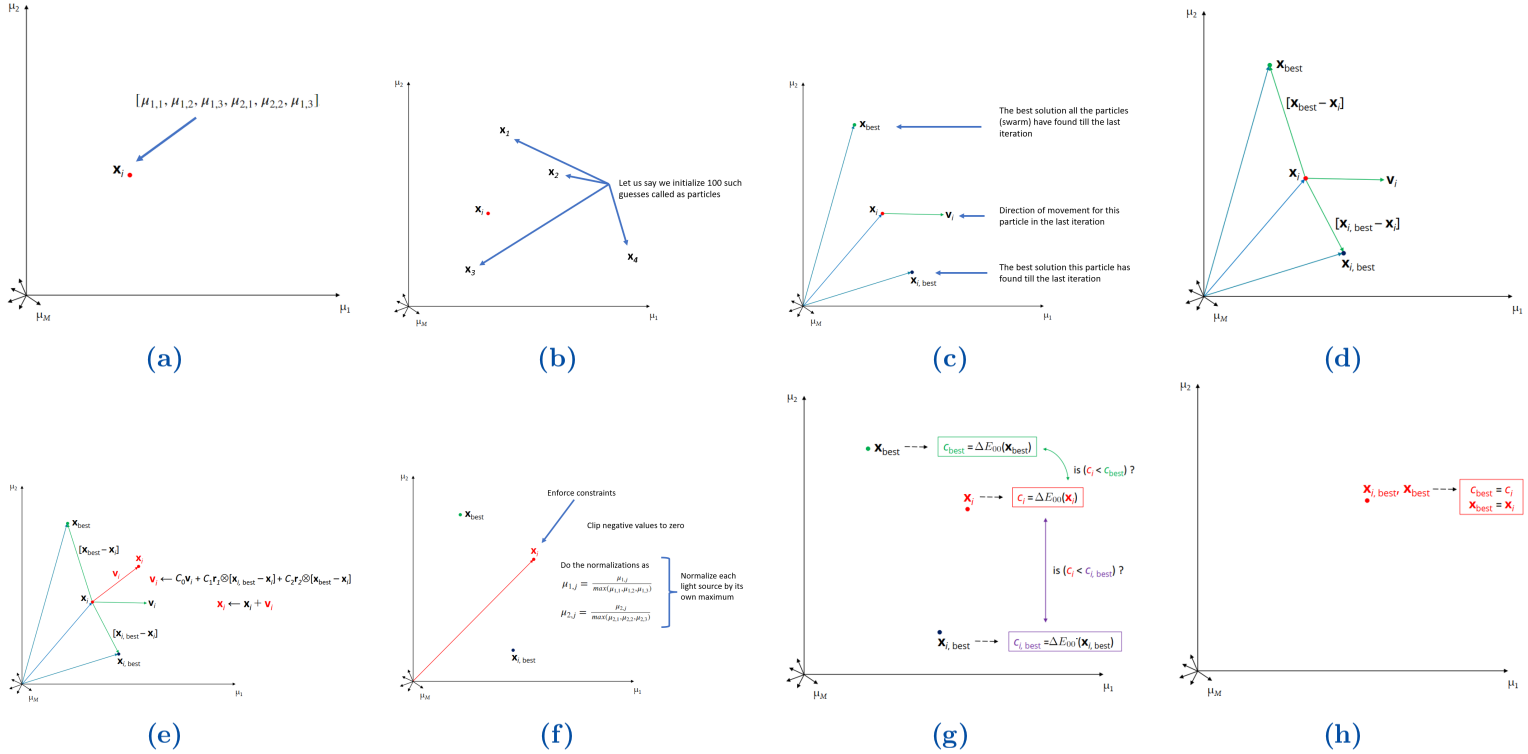


Figure 6.3: *Detail of Particle Swarm Optimization (PSO) algorithm. (a) Parameterize into six parameters to be optimized. (b) Initialize randomly multiple solutions across the solution space. (c) Define the velocity, particle best, and swarm best. (d) Calculate difference vectors. (e) Update particle's position. (f) Enforce constraints. (g) Compare particle's new cost with particle's and swarm's best costs. (h) Update particle's and swarm's best costs. The figures are modified for this thesis work from Falt (2022).*

6.2.1 Parameters

The intensity of an LED is defined as a value in the range 0 to 1 where 0 corresponds to no illumination of the LED and 1 corresponds to maximum illumination. The optimization problem is parameterized by considering the intensities of the LEDs as the parameters to be optimized i.e. six dimensions of solution space are considered, where each dimension corresponds to each of the LED intensity. The solution is defined as:

$$x = [\mu_{1,1}, \mu_{1,2}, \mu_{1,3}, \mu_{2,1}, \mu_{2,2}, \mu_{1,3}] \tag{6.1}$$

where $\mu_{i,j}$ is the intensity for i^{th} light source having j^{th} LED.

6.2.2 Constraints

For the optimization to give results that are physically realizable and to make the modelling closer to real captures, two types of constraints are defined.

6.2.2.1 Clipping of negative intensities

The negative LED intensity values are clipped because they correspond to negative illumination, which implies that an LED reduces the illumination at the wavelengths it is defined. Achieving this would technically be possible if the LEDs had high overlapping regions in wavelengths because a light source is defined by additively mixing three LEDs. In this case, a negative contribution to illumination from a single LED can still lead to net positive illumination across all wavelengths because of other LEDs contributions. This would make the light source physically realizable. But the LEDs wavelengths do not overlap substantially, so any negative intensity is clipped to zero during the optimization steps. Mathematically, it can be written as:

$$\mu_{i,j_{clipped}} = max(0, \mu_{i,j}) \tag{6.2}$$

6.2.2.2 Normalization of intensities

For the simulations, the relative intensities of the LEDs are important hence LEDs spectra are normalized. Since there are two light sources each defined by a LEDs triplet, each LEDs triplet is normalized with the the maximum intensity among them. Mathematically, it can be written as:

$$\mu_{1,j} = \frac{\mu_{1,j}}{max(\mu_{1,1}, \mu_{1,2}, \mu_{1,3})} \tag{6.3}$$

$$\mu_{2,j} = \frac{\mu_{2,j}}{max(\mu_{2,1}, \mu_{2,2}, \mu_{2,3})} \tag{6.4}$$

This type of normalization also reflects the real captures setting more closely because for each of the two captures in dual-RGB imaging, the exposure time is set individually for each capture. This also allows for compensating for the very low illumination of the LEDs 6 and 7 since the exposure time can be increased for the captures involving the light sources with these LEDs to get more signal in the image.

6.2.3 Cost function

The cost function is ΔE_{00} with the parameters k_L, k_C, and k_H defined to be 1.

6.2.4 Tuning parameters

- C_0 is set as the average of 1.4 and 0.5 values range proposed in Shi and Eberhart (1998).
- Both C_1 and C_2 are set as 2 from the recommended values of Kennedy and Eberhart (1995).
- L_{max} is set as 10 and N as 200. These are empirically set by author considering the trade-off between computational resources, time, and errors being achieved.

6.3 Results

The PSO is applied on each of the 630 combinations of light source pairs to optimize the intensities of the LEDs involved. In this way, the pair of light sources and their corresponding LEDs intensities which minimize the ΔE_{00} error are found. The workflow is summarized in Fig.6.4.

The Fig.6.5 shows the final optimized LED light source pair. Notice that the light source pair has changed from (1, 6, 8) and (3, 5, 9) selected at full intensity to (1, 6, 7) and (2, 4, 9). Since the effective channels of the dual-RGB imaging setup is the combined effect of LED illumination with the CSRF, the channels of the overall system are shown in Fig.6.6. This highlights the fact that optimizing the LEDs

at varying intensities expands the solution space. The results are summarized in Table.6.1.

The Fig.6.7 and Fig.6.10 show the patch-wise ΔE_{00} error for training and testing data similar to previous chapter. The training mean error comes out to be 0.17 and testing error of 0.28. The training error has reduced further by 86% and testing by 82% of what had been calculated for three-channels imaging setup in Chapter 5. Note that dual-RGB had reduced the errors by 75% for training and 57% for testing. Dual-RGB-PSO provides further reduction of 11% for training and 25% for testing. Similar to the observation in neutral grays of decreasing errors with increasing lightness, a similar trend is shown in these figures.

The Fig.6.8 and Fig.6.11, like in last chapter, shows the colour reproduction of ground-truth and predicted patches. This time each block has four sub-blocks in the form $\left(\begin{smallmatrix} GT & TC \\ DRP & DR \end{smallmatrix}\right)$, where GT is for ground truth, TC for three-channel, DR for dual-RGB, and DRP for dual-RGB-PSO sub-block.

The Fig.6.9 shows the plot for Lightness vs. Chroma and a^* vs. b^* in CIELAB for training data. like as shown last chapter. For Lightness vs. Chroma, dual-RGB-PSO plot also shows a similar trend of having least errors at the bottom right of the plot which represents highest lightness and minimum chroma. As the lighthness is reduced or chroma is increased, the errors are generally increasing. The a^* vs. b^* plot shows, again, the same trend as dual-RGB i.e. having least error at the center of the plot representing neutral colours and increasing errors as the point move outwards. The Fig.6.12 shows Lightness vs. Chroma and a^* vs. b^* plots for testing data with no obvious patterns to observe.

Table 6.1: *Details and comparison of PSO optimized intensity with full intensities and 3-channel system.*

Condition	3-channel	Full intensity	PSO intensity
Source 1 LEDs	D65	(1, 6, 8)	(1, 6, 7)
Source 2 LEDs	N/A	(3, 5, 9)	(2, 4, 9)
Source 1 intensities	N/A	(1.0, 1.0, 1.0)	(0.07, 0.58, 1.0)
Source 2 intensities	N/A	(1.0, 1.0, 1.0)	(0.12, 0.56, 1.0)
Train error	1.25	0.31	0.17
Test error	1.53	0.66	0.28

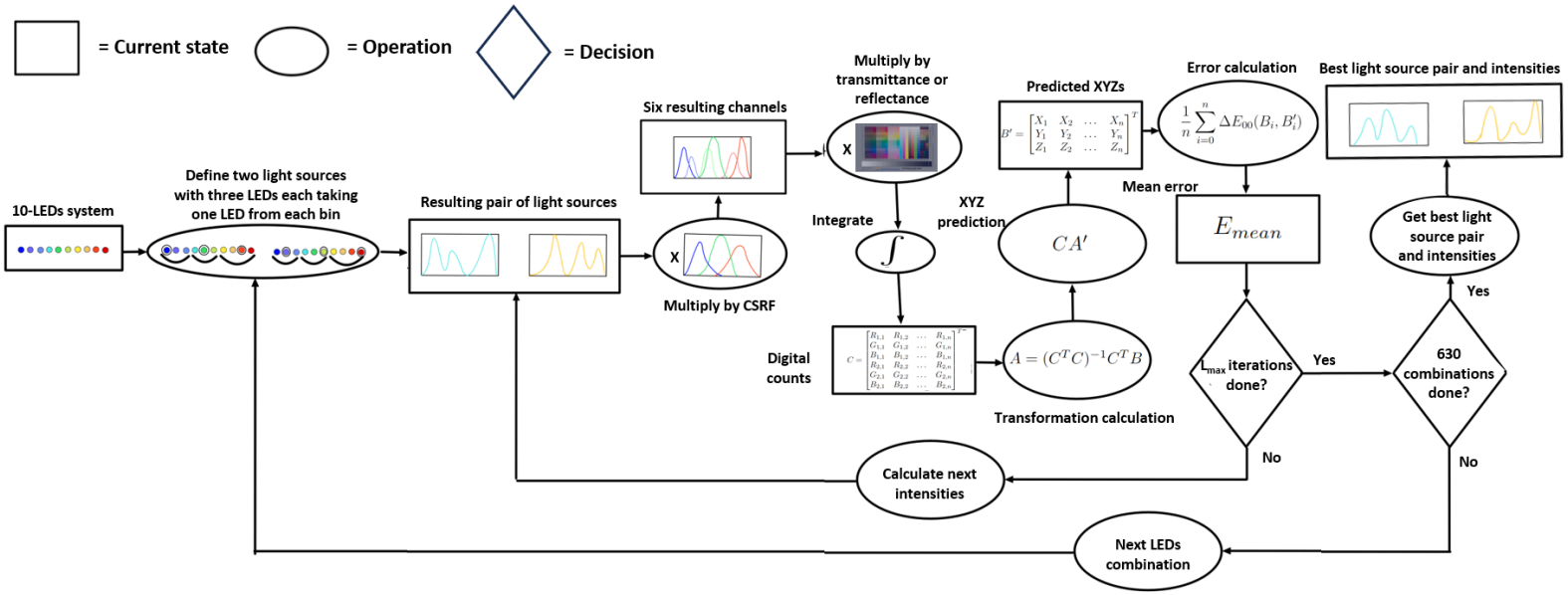


Figure 6.4: *Dual-rgb imaging workflow with LED intensity optimization. Two light sources are defined from the 10-LEDs light source in such a way that each light source comprises of three LEDs each of which are taken from each bin. The defined light sources are mutliplied by the CSRF to get six channels of the imaging setup. These channels are multiplied with the trasmittances of each of the patches of the colourchecker and integrated to simulate the digital numbers captured by the camera. The transformation matrix is estimated by using the pseudo-inverse. This transformation matrix is then used to predict the tristimulus values of the patches. The ΔE_{00} difference can be calculated for each of the patches using the ground-truth tristimulus values and can be averaged to provide mean error. The optimization method suggests the next set of LED intensities to use in the current light source pair untill maximum number of iterations are completed. This process is done for all the 630 combinations of light source pairs. Finally the pair of light source and its corresponding LED intensities are selected which gives the least mean error.*

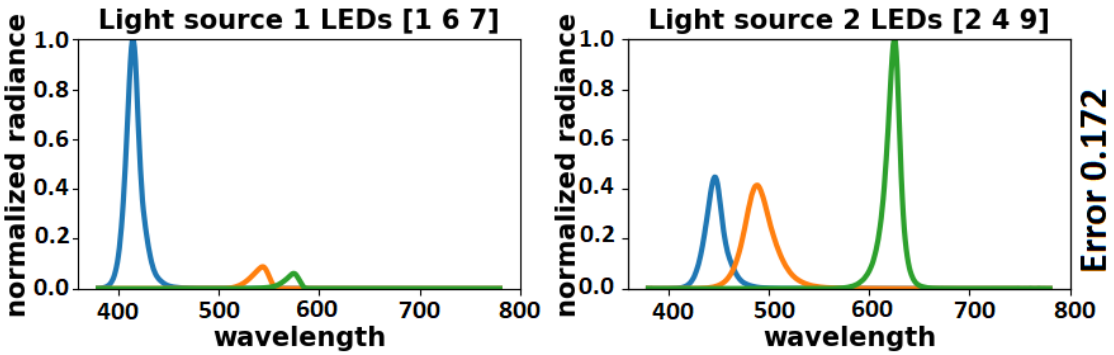


Figure 6.5: *Final PSO-optimized light sources among all the 630 different combinations of light sources pairs.*

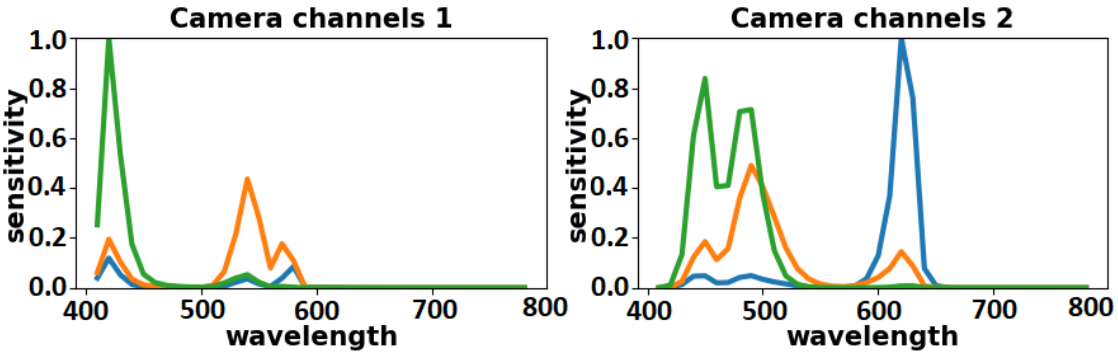
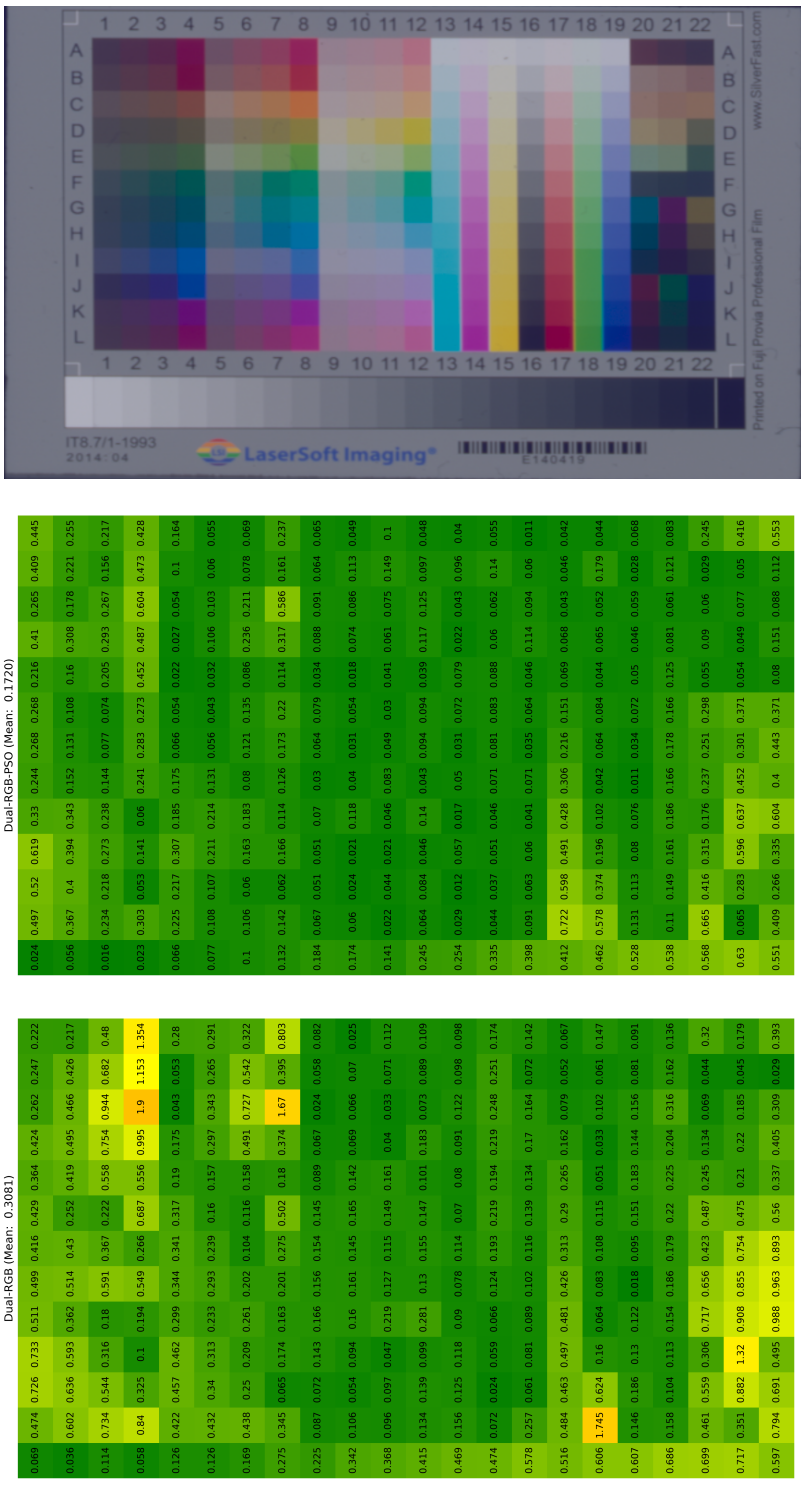


Figure 6.6: *Resulting channels of the overall multispectral system got after multiplication of light source pairs with CSRF.*

6.4 Further exploration on intensities: Genetic algorithm

The solution obtained from the PSO might be a result of some local minimum in the solution space or, in the worst scenario, a by-product of some algorithm-specific details. This gives motivation to get a 'second opinion' from other optimizing algorithms too. For this purpose, genetic algorithm is selected because it has been utilized in literature for LEDs optimization in art paintings context (Durmus et al., 2020) and also because of its wide usage in science and engineering problems (Elbeltagi et al., 2005).

First a population of μ randomly generated solutions are generated. In each iteration λ offsprings are generated by either crossover operation with a probability of p_c or mutation. If the offspring is different from the parent then the fitness function is evaluated, otherwise the same fitness value as parent is assigned to the



(a) *Train error dual-RGB* (b) *Train error dual-RGB-PSO* (c) *Train colorchecker*

Figure 6.7: *Comparing errors for each patch in training for dual-RGB and dual-RGB-PSO.*

Figure 6.8: *Train data reproduction from groundtruth data and predictions from dual-RGB-PSO, dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\begin{pmatrix} GT & TC \\ DRP & DR \end{pmatrix}$*, where GT is for ground truth, TC for three-channel, DR for dual-RGB, and DRP for dual-RGB-PSO block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits.*

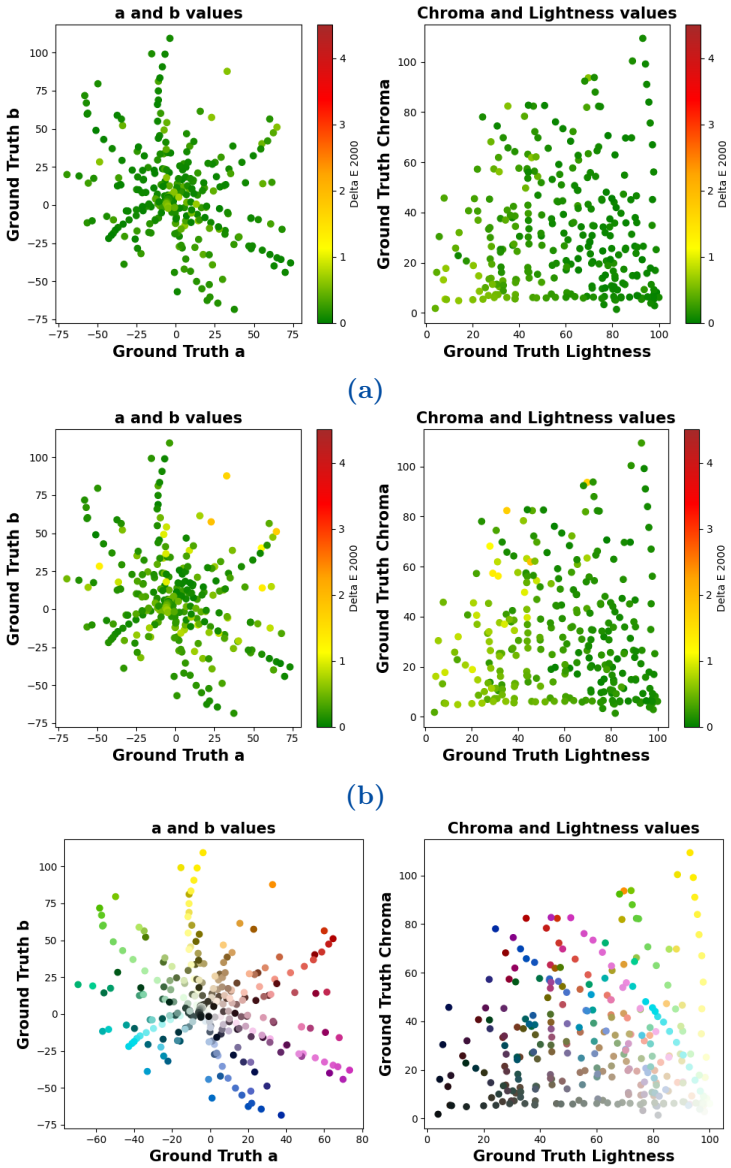


Figure 6.9: *Plots for dual-RGB-PSO and dual-RGB training patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for dual-RGB-PSO. (b) is reproduced from Fig.5.12a for convenience. It represents a* vs. b* and Lightness vs. Chroma plots for dual-RGB. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) is reproduced from Fig.(c) for convenience. It shows similar a* vs. b* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

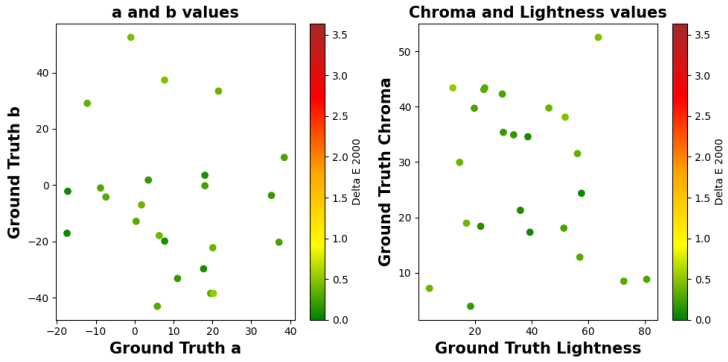
(a) *Test colorchecker*

(b) *Test error dual-RGB*

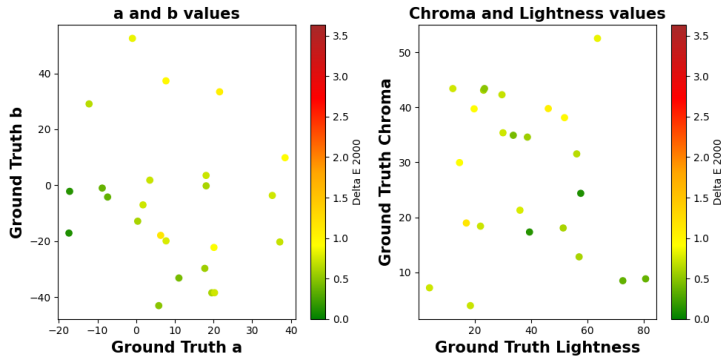
(c) *Test error dual-RGB-PSO*

Figure 6.10: *Comparing errors for each patch in training for dual-RGB and dual-RGB-PSO.*

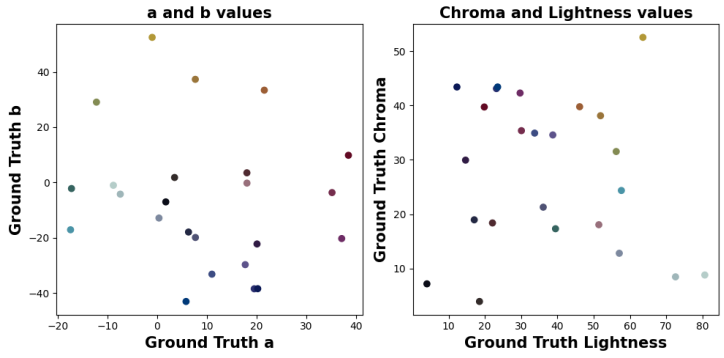
Figure 6.11: *Test data reproduction from groundtruth data and predictions from dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & TC \\ GT & DR \end{smallmatrix}\right)$*, where GT is for ground truth, TC for three-channel, and DR for dual-RGB block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits.*



(a)



(b)



(c)

Figure 6.12: *Plots for dual-RGB and three-channel testing patches. (a) shows a^* vs. b^* and Lightness vs. Chroma plots for dual-RGB-PSO.(b) is reproduced from Fig.5.16a which represents a^* vs. b^* and Lightness vs. Chroma plots for dual-RGB. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) is reproduced from Fig.c, which shows similar a^* vs. b^* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

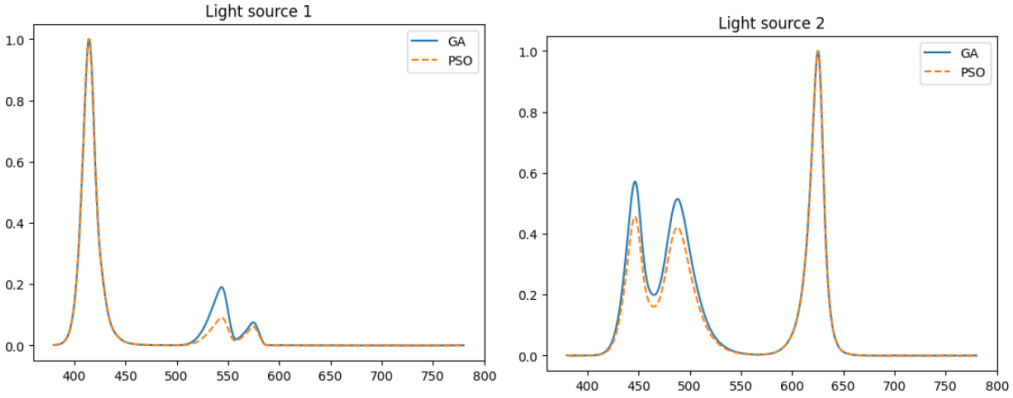
offspring. The best μ solutions from both parents and off-springs are selected to move to the next generation (iteration). The crossover is a function which generates an offspring by taking parts of the parameters from parents and concatenating them to generate an offspring solution. The mutation is another of way generating an offspring which is to take a single parent and randomly change one of its parameters. The fitness function evaluates the error for a solution. The algorithm is summarized in Algorithm.2.

Data: Population sizes μ, λ, crossover probability p_c, mutation rate p;
Result: $P = \{x^{(1)}, x^{(1)}, \ldots, x^{(\mu)}\}$
Set $P = \{x^{(1)}, x^{(1)}, \ldots, x^{(\mu)}\}$ randomly across the solution space
for t *in* $[1, \ldots, L_{max}]$ **do**
  $P' \leftarrow \emptyset$
  for i *in* $[1, \ldots, \lambda]$ **do**
    Sample $r \in [0, 1]$;
    if $r \leq p_c$ **then**
      select two individuals x, y from P;
      $z^{(i)} \leftarrow$ Crossover(x, y);
      if $z^{(i)} \notin \{x, y\}$ **then**
        evaluate $f(z^{(i)})$;
      else
        infer $f(z^{(i)})$ from parent;
      end
    else
      select an individual x from P;;
      $z^{(i)} \leftarrow$ Mutation(x;)
      if $z^{(i)} \neq x$ **then**
        evaluate $f(z^{(i)})$
      else
        infer $f(z^{(i)})$ from parent;
      end
    end
    $P' \leftarrow P' \cup \{z^{(i)}\}$;
  end
  P is updated by the best μ points in $P \cup P'$
end

Algorithm 2: Genetic algorithm $(\mu + \lambda)$

6.4.1 Results for genetic algorithm and comparison with PSO

As shown in Fig.6.13, Fig.6.14, and Table.6.2, genetic algorithm optimizes to not just the same LED pairs but also very closely to the relative intensities of the LED pairs too. This gives more confidence for the LED optimization solution by PSO for being more closer to global minimum rather than some local one and also not being a result of some algorithm-specific result.



(a) *Light source 1* **(b)** *Light source 2*

Figure 6.13: *Comparison of light sources optimized from PSO and genetic algorithm.*

Table 6.2: *Details and comparison of PSO optimized intensity with Genetic algorithm intensities*

Condition	PSO intensity	Genetic intensity
Source 1 LEDs	(1, 6, 7)	(1, 6, 7)
Source 2 LEDs	(2, 4, 9)	(2, 4, 9)
Source 1 intensities	(0.07, 0.58, 1.0)	(0.06, 1.0, 1.0)
Source 2 intensities	(0.12, 0.56, 1.0)	(0.15, 0.68, 1.0)
Train error	0.17	0.18
Test error	0.28	0.28

6.5 Discussion

This chapter discussed the limitation of the solution space in the dual-RGB workflow elaborated in Chapter 5. In dual-RGB workflow, the LEDs were only considered

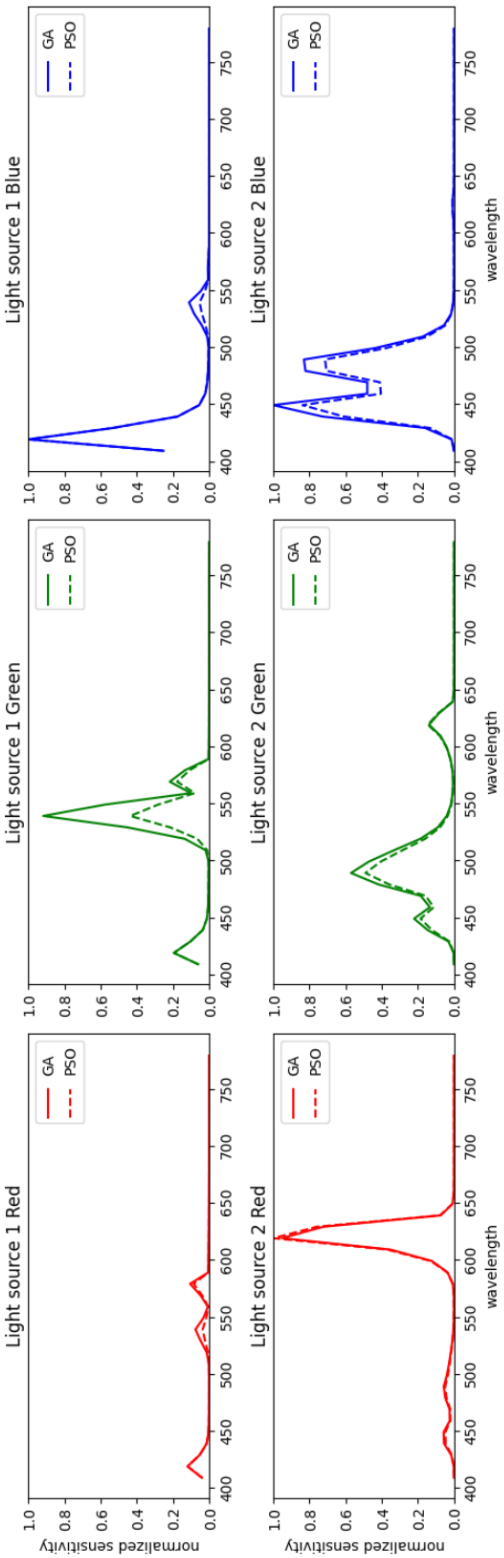


Figure 6.14: *Comparison of multispectral channels formed by PSO and genetic algorithm optimized light sources.*

at their full intensities for defining a pair of light sources. It was shown that how the optimal light source pair, in terms of mean ΔE_{00} error, from the dual-RGB simulations can further be optimized by changing the intensities of LEDs defining the light source. As an example, the Fig.6.2 showed that by changing only one LED keeping the rest at full, the mean error was reduced from 0.31 to 0.29.

For the LEDs optimizations, an evolutionary algorithm called Particle Swarm Optimization (PSO) was chosen and adapted for dual-RGB (hence called dual-RGB-PSO). After parameterizing the problem of LEDs intensity optimization, PSO was applied to all the 630 combinations of light source pair which could be defined in dual-RGB. Those LEDs and their corresponding intensities were selected as the final light source which gave the minimum mean ΔE_{00} error. The final pair of light sources was shown in Fig.6.5 and their corresponding spectral channels in Fig.6.6. The mean error for dual-RGB-PSO was 0.17 for training and 0.28 for testing. This shows a reduction of 86% and 82% for training and testing in comparison to the three-channel imaging errors. In comparison, the dual-RGB had reduced the errors to 75% and 57%. Moreover, the dual-RGB-PSO selects different LEDs than the dual-RGB which highlights the fact that dual-RGB was exploring a limited solution space for the minimization of the errors.

The sRGB colour reproductions of ground-truth, three-channels, dual-rgb, and dual-RGB-PSO colour patches for training and testing data were shown in Fig.6.8 and Fig.6.11. Like in Chapter 5, three-channel shows a perceivable difference from the ground-truth. For the dual-RGB and dual-RGB-PSO, generally there is no perceivable difference. This can be attributed to the fact that the errors for both the dual-RGB and dual-RGB-PSO are below the just noticeable difference. So the visualizations for both of the system would generally appear the same.

The Fig.6.9 and Fig.6.12 showed the CIELAB a^* vs. b^* and Lightness vs. Chroma plots for the training and testing data. For dual-RGB-PSO, we observe similar trends of decreasing error with increasing Lightness and decreasing Chroma as was found for three-channels and dual-RGB in Chapter 5. For testing data, these trends are not visible by inspection which, again, can be attributed to low number of sample points.

In order to check the robustness of the optimal solution provided by PSO, Genetic algorithm was used to optimize the LEDs intensities similar to what PSO had done. It was shown in Fig.6.13 and Fig.6.14, that the genetic algorithm not only selects the same LEDs as optimal light sources, but also a very similar corresponding intensities. This gives more confidence in the optimal solution provided by the PSO for the best light sources and their LED intensities.

Going through the comparative errors and plots for dual-RGB-PSO, it can be seen that exploring intensities for defining light sources for dual-RGB the solutions can be further improved in comparison to the three-channel imaging. And for the

purpose of optimization of the LEDs intensities, both PSO and genetic algorithm can be utilized to give optimal solution.

7 | Real captures

The unexpected always happens when you least expect it.

Lola Perkins

This chapter discusses the real captures that are taken for the optimized dual-RGB, dual-RGB-PSO, and three-channel imaging setup simulated in previous chapters. The Fig.7.1 shows the setup while taking a dual-RGB capture. Each of the image captured is corrected for dark noise and non-uniform illumination is handled through flat-fielding. The Eq.7.1 is used for dark noise correction and flatfielding.

$$img_f = \frac{img_r - img_d}{img_w - img_d} \tag{7.1}$$

where img_r is the raw image, img_d is the dark noise image, and img_w is the white image.

Dark noise image, img_d, is taken as the average of the 10 captures taken with lens cap on the camera. The white image, img_w, is taken as the open-gate image of the light source.

The white, train, and test colorchecker captures for light sources optimized dual-RGB and dual-RGB-PSO are shown in Figs.7.2 and 7.3. For all the captures the ISO was set to 50 (minimum possible in the camera). The integration time was set using the open-gate images of both light sources i.e. one integration time for each light sources. The details of the captures are mentioned in Table.7.1:

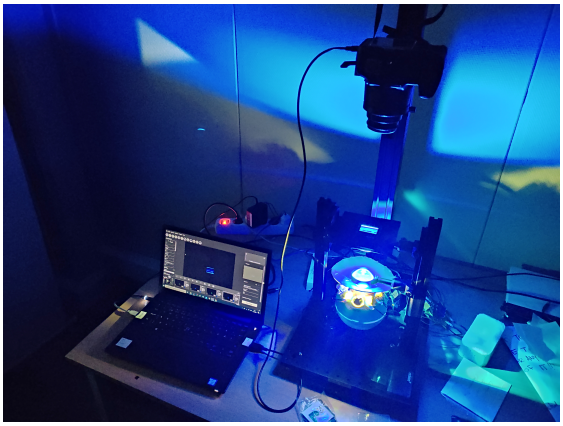


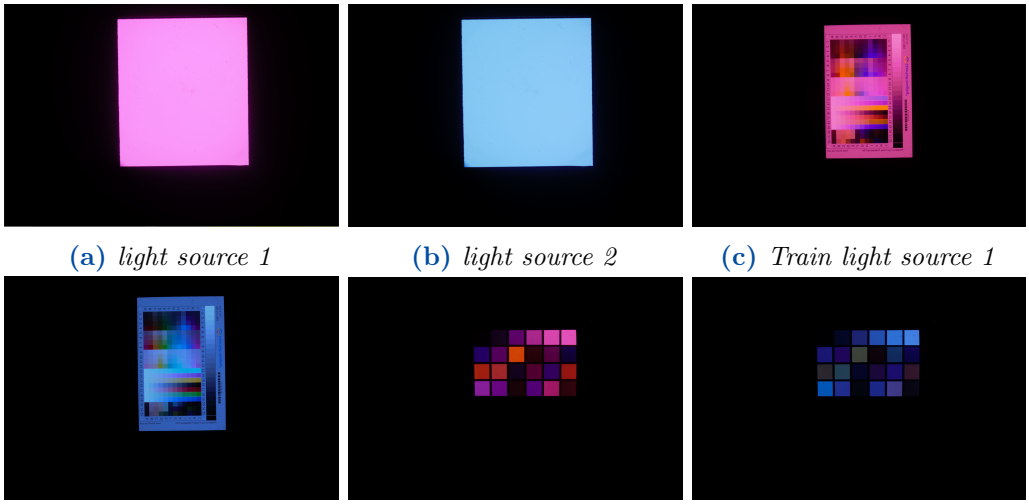
Figure 7.1: *Capturing the dual-RGB image*

Table 7.1: *Table showing the integration times for the captures.*

Light source	Integration time (s)
Full intensity Light source 1	1/60
Full intensity Light source 2	1/500
PSO intensity Light source 1	1/2
PSO intensity Light source 2	1/40

7.1 Real captures Vs. simulated captures validation

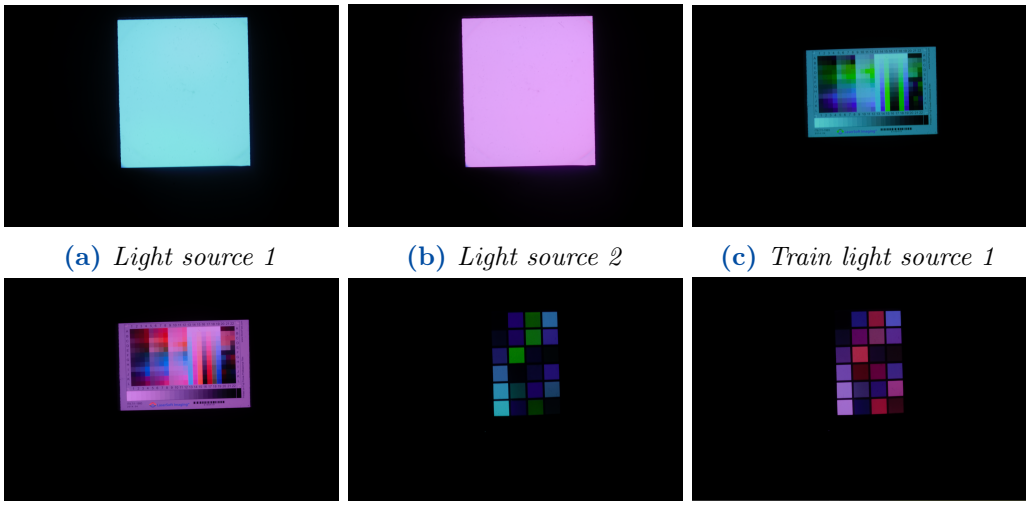
In order to compare the real captures and the simulated captures calculated in previous sections, the intensity values of the simulated pixel intensities against the real intensities are plotted in Fig.7.4. Each point corresponds to the simulated and real capture values of each patch in the training data. The black line is the line found through linear regression on all the points. The camera response are shown for each of the individual channels of the dual-RGB system separately. The R^2 error for the regressed line is shown above each plot. It can be clearly seen that the real and simulated captures lie well on the regressed line for each of the channel. All of the R^2 errors are also above 0.98 which makes the fit very well. There are two outliers found for the training data which are marked with red colours. The results of the comparison after removing the outliers are shown in Fig.7.5. From the improvement in R^2, it can be seen the regressed line better fits the data proving a linear relationship between both actual and simulated captures.



(a) *light source 1* (b) *light source 2* (c) *Train light source 1*

(d) *Train light source 2* (e) *Test light source 1* (f) *Test light source 2*

Figure 7.2: *Real captures of open-gate, train, and test for full intensity light sources.*



(a) *Light source 1* (b) *Light source 2* (c) *Train light source 1*

(d) *Train light source 2* (e) *Test light source 1* (f) *Test light source 2*

Figure 7.3: *Real captures of open-gate, train, and test for PSO intensity light sources.*

7.2 Results

In this section, the errors for the predictions using dual-RGB, dual-RGB-PSO, and three-channel imaging are discussed. The Table 7.2, Fig.7.6, and Fig.7.9 shows the overall and patch-wise ΔE_{00} error. The mean calculation is done after removing

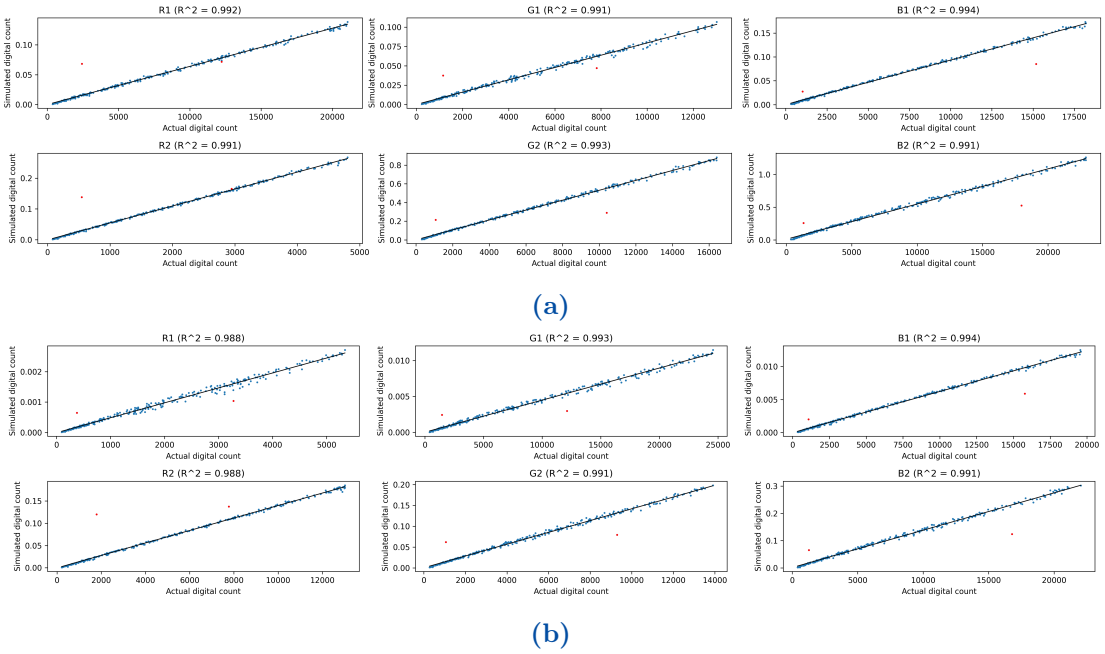


Figure 7.4: *Correlation plots for real and simulated digital numbers. (a) shows the plot for dual-RGB and (b) shows the plot for dual-RGB-PSO. Notice that there are two patches which are outliers.*

the outliers mentioned before. The mean error in the three-channel imaging is 1.89 which is reduced to 1.06 in dual-RGB and 1.11 in dual-RGB-PSO. These represent a reduction of 43% and 41% for dual-RGB and dual-RGB-PSO respectively. For testing, the errors reduced from 2.52 in for three-channels to 1.61 for dual-RGB and to 1.87 for dual-RGB-PSO. These represent a reduction of 36% and 26% for dual-RGB and dual-RGB-PSO respectively. Unlike in the simulations, where the dual-RGB-PSO performed better than dual-RGB, in real captures the performance of dual-RGB and dual-RGB-PSO are comparable.

The Fig.7.7 and Fig.7.10 shows the colour reproduction of ground-truth and predicted patches similar to Fig.6.8. each block has four sub-blocks in the form $\left(\begin{smallmatrix} GT & TC \\ DRP & DR \end{smallmatrix}\right)$, where GT is for ground truth, TC for three-channel, DR for dual-RGB, and DRP for dual-RGB-PSO sub-block. This time another set of patches are highlighted as red which are the outliers detected previously.

The Fig.7.8 shows the plot for Lightness vs. Chroma and a^* vs. b^* in CIELAB for training data. For Lightness and Chroma, as expected, the trend of decrease in error with increasing Lightness and decreasing Chroma is observed. For a^* vs. b^*, it can be seen that, unlike the simulations, no clear trend of higher errors are observed along any specific direction. No particular trend also extends to Fig.7.11 which shows the Lightness vs. Chroma and a^* vs. b^* plots for testing data.

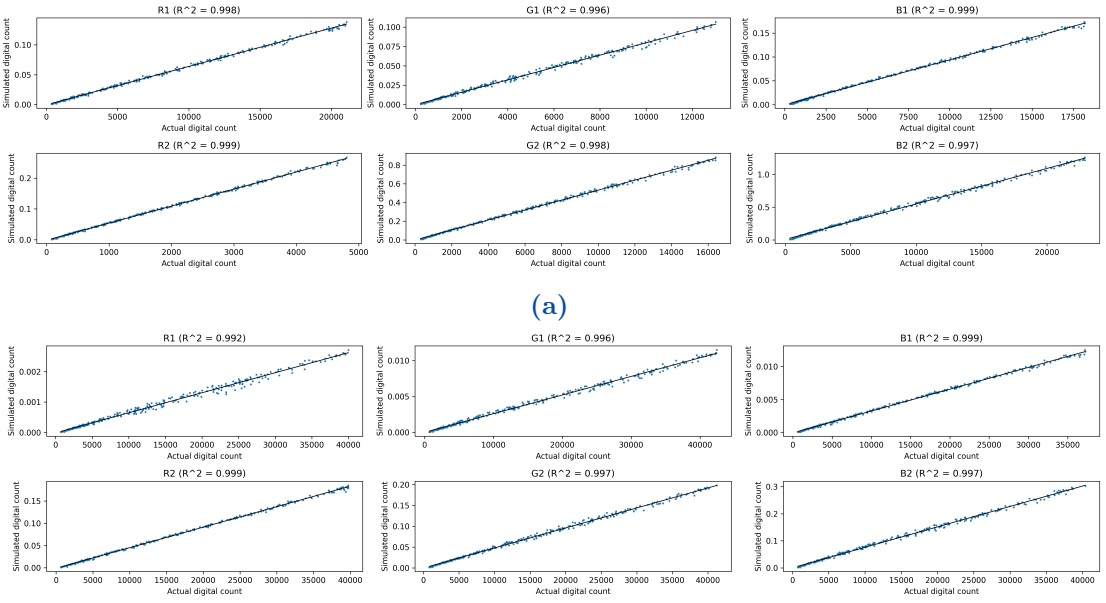


Figure 7.5: *Correlation plots for real and simulated digital numbers. (a) shows the plot for dual-RGB and (b) shows the plot for dual-RGB-PSO. The outliers are removed from the calculations and plots*

Table 7.2: *Table showing the errors for the real captures*

Metric	3-channel	Full intensity	PSO intensity
Mean train Error	1.89	1.06	1.11
Median train Error	1.36	0.62	0.72
Mean test Error	2.52	1.61	1.87
Median test Error	2.64	1.72	1.83

7.3 Discussion

This chapter described the processing and results for real captures corresponding to the optimal light sources described in simulations in Chapter 5 and 6. For the real captures, RAW images are used which are linear to the irradiance on the sensor. In order to validate the real captures, a comparison with the simulations was discussed. In Fig.7.4 a plot was shown of the digital counts of real captures against simulated captures. In order to validate the linear relationship between these two quantities, linear regression is performed to fit all the data points. It can be seen that the line fits the data quite well with all the R^2 errors above 0.98. In these plots, two outliers are detected which correspond to the same patch in training. The Fig.7.5

(a) *Train colorchecker* (b) *Train error 3-channel* (c) *Train error dual-RGB* (d) *Train error dual-RGB-PSO*

Figure 7.6: *Comparing errors for each patch in training for 3-channel, dual-RGB, and dual-RGB-PSO.*

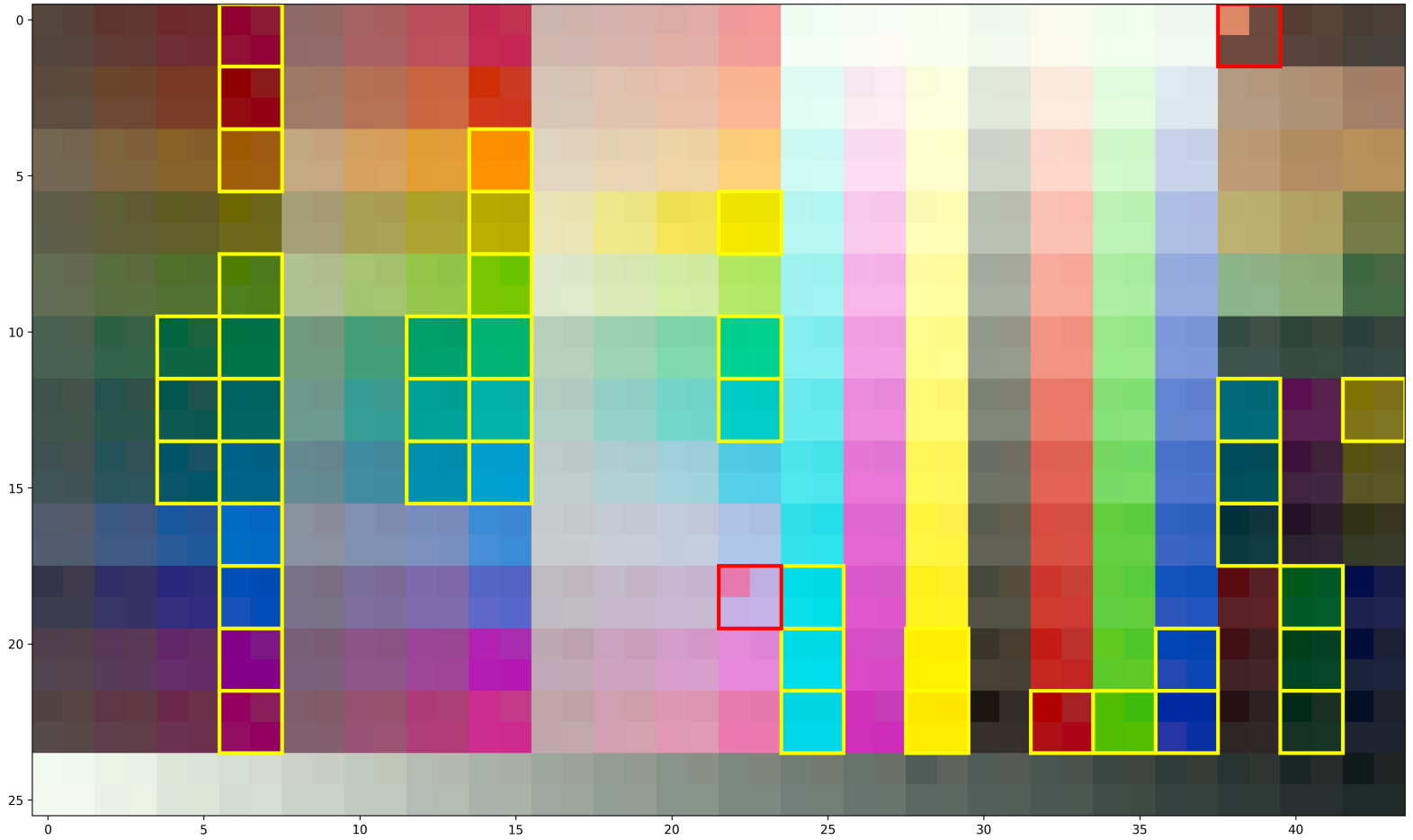
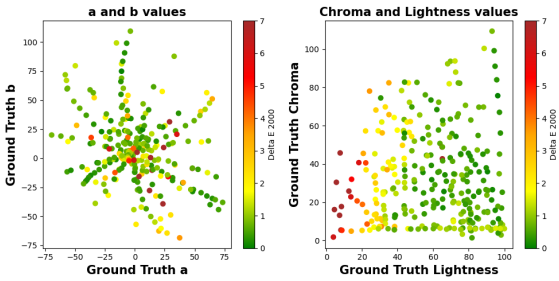
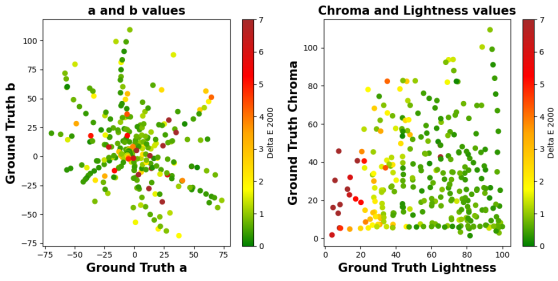


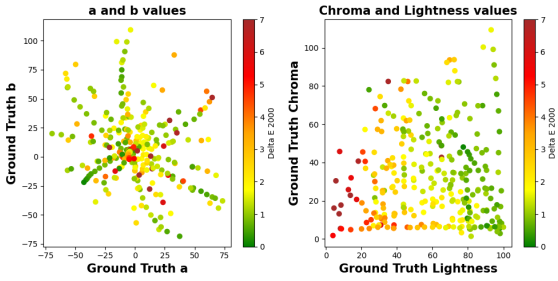
Figure 7.7: *Train data reproduction from ground-truth data and predictions from dual-RGB-PSO, dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & TC \\ DRP & DR \end{smallmatrix}\right)$*, where GT is for ground truth, TC for three-channel, DR for dual-RGB, and DRP for dual-RGB-PSO block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. Patches with yellow boundaries are out-of-gamut chromaticities for sRGB colour space limiting the actual demonstration. Patches with red boundaries are outliers which are discarded from the mean calculations.*



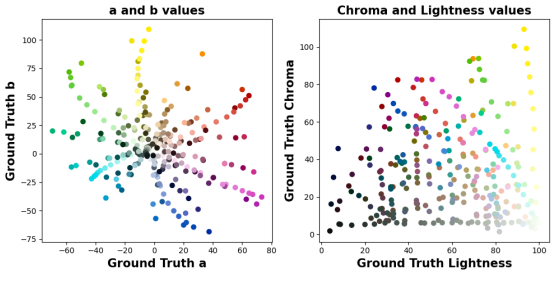
(a)



(b)



(c)



(d)

Figure 7.8: *Real captures plots for dual-RGB-PSO, dual-RGB and three-channel training patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for dual-RGB-PSO.(b) and (c) represent the same set of plots for dual-RGB and three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. For (d) the colour codes are the sRGB representation of patches*

(a) *Test colorchecker* **(b)** *Test error 3-channel* **(c)** *Test error dual-RGB* **(d)** *Test error dual-RGB-PSO*

Figure 7.9: *Comparing errors for each patch in testing for 3-channel, dual-RGB, and dual-RGB-PSO.*

Figure 7.10: *Test data reproduction from ground-truth data and predictions from dual-RGB and three-channels imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & TC \\ GT & DR \end{smallmatrix}\right)$*, where GT is for ground truth, TC for three-channel, and DR for dual-RGB block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. Patches with yellow boundaries are out-of-gamut chromaticities for sRGB colour space limiting the actual demonstration. The patches with red boundaries are the outliers which are detected in the real and simulated digital numbers correlation plot.*

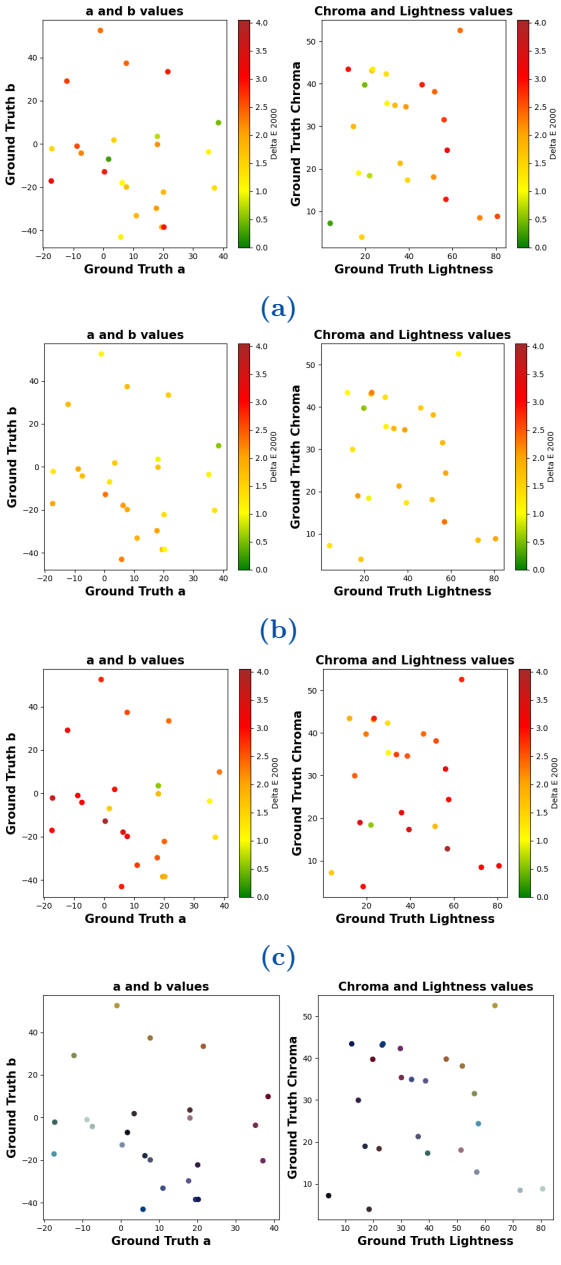


Figure 7.11: *Real capture plots for dual-RGB-PSO, dual-RGB and three-channel testing patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for dual-RGB-PSO.(b) and (c) represent the same set of plots for dual-RGB and three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. For (d) the colour codes are the sRGB representation of patches.*

shows the same plots with outliers removed. A further improvement is observed in the R^2 error. The two outliers might be attributed to some inaccuracy while measuring the transmittances through the overhead projector setup as discussed in Chapter 4. In further calculations, these outliers were removed.

Fig.7.6 and Fig.7.9 showed the three-channel, dual-RGB, and dual-RGB-PSO imaging mean ΔE_{00} error. For three-channel imaging, the mean error was 1.89 for training and 2.52 for testing. dual-RGB reduces the training error to 1.06 and testing error to 1.11. This was a 43% and 41% reduction from three-channel imaging error. For dual-RGB-PSO, the errors reduced to 1.11 for training and 1.87 for testing. This was a 41% reduction in error for training and 26% reduction in testing.

Similar to what was shown earlier in Chapters 5 and 6 in simulations, real captures also prove that having extra channels in the dual-RGB and dual-RGB-PSO provide a better colour accuracy in comparison to three-channel imaging both for training and testing data.

One apparent misalignment is between the errors of dual-RGB and dual-RGB-PSO. In simulations, dual-RGB-PSO clearly outperformed the dual-RGB training and testing by having approximately half the errors as dual-RGB. In real captures, both perform equally well in comparison to three-channel image. One possible explanation for it might be the usage of percentage while seeing the improvements between dual-RGB and dual-RGB-PSO. For example, in the simulations the training errors for dual-RGB and dual-RGB-PSO were 0.31 and 0.17. If the errors are computed between the predicted patches of dual-RGB and dual-RGB-PSO (instead of calculating errors of dual-RGB vs. three-channel and dual-RGB-PSO vs. three-channel), the mean error is 0.28 which would be imperceivable by HVS. In other words, the difference between predicted dual-RGB and dual-RGB-PSO colours are imperceivable. If the same errors between dual-RGB and dual-RGB-PSO are calculated for the real captures, the mean error is 0.52 which also would be imperceivable to HVS. In this regard, both the simulations and real captures agree on the imperceivability of the difference between dual-RGB and dual-RGB-PSO. One possible solution for exploring the effectiveness of intensity optimizations in dual-RGB workflow would be to use a different training and testing data which offer a larger variety of colour patches to be trained on and to be predicted. This would allow both dual-RGB and dual-RGB-PSO to work on larger scales of colour differences and possibly giving optimized solutions which are much different from each other.

The sRGB colour reproductions of ground-truth, three-channels, dual-rgb, and dual-RGB-PSO colour patches for training and testing data were shown in Fig.7.7 and Fig.7.10. The reproductions highlight a generally better colour accuracy of the dual-RGB and dual-RGB-PSO over the three-channel predictions. Also

the previously discussed observation can be seen here about the imperceivability between the dual-RGB and dual-RGB-PSO predictions.

The Fig.7.8 and Fig.7.11 showed the CIELAB a^* vs. b^* and Lightness vs. Chroma plots for the training and testing data. Similar trends as found in simulations in Chapters 5 and 6 are found in real captures too. For training data, a decrease in errors are observed with the increasing Lightness and decreasing Chroma. For testing, no obvious trends can be deduced by inspection.

Exploring and analyzing the results from all the different plots and visualizations, it can be easily seen that having more channels not only give better performance in simulations but also in real capturing workflow too. For the dual-RGB and dual-RGB-PSO solutions, no perceivable difference is found for the real captures. One reason could be the scale of differences between dual-RGB and dual-RGB-PSO, giving a hint to possible future directions one of which is to use different set of training and testing data having a larger variety of samples.

8 High Dynamic Range Imaging

Nature always wears the colors of the spirit.

Ralph Waldo Emerson

Uptill now, the workflow on dual-RGB imaging has been implemented both in simulation and in real captures. It was also seen how changing the intensities of the LEDs involved in the light source has the potential to improve the colour accuracy. In this chapter, a workflow is explored on how high dynamic range (HDR) imaging techniques can be utilized to potentially improve the results of dual-RGB and dual-RGB-PSO imaging.

8.1 Background

One of the ways to define dynamic range is the ratio of the maximum and minimum input signal which is in the range that a change in signal is detectable. It is also useful to define a quantity called dynamage range as ratio of the maximum input signal which does not damage the sensor/device and minimum input signal which is in the range that a change in signal is detectable (Mann et al., 2010)(Mann et al., 2012). HDR imaging can be done either by the combination of several low dynamic range images or by the acquisition of HDR sensors. First known form of combining images to improve dynamic range has been by Mann (1993) (Robertson et al., 2003).

During a camera capture, the light falling on the sensor, q, can not be called as radiometric because the sensor sensitivity is different for different wavelengths. It can also not be called as photometric because the sensor sensitivity also does not

follow the spectral response of the human eye or the color matching function and luminous efficiency function defined by CIE (Reinhard et al., 2010)(Debevec and Malik, 1997).

If the quantity q is considered as a spatially varying quantity across the sensor array $q(x, y)$ then two classes of improvements can be defined that can be done in recording this quantity. The first class acting on the domain of (x, y) and the second class acting on the range of q. Then define the improvements can be defined as either increasing the resolution or extent of these quantities as

1. Increasing pixel density (domain resolution)
2. Increasing number of pixels (domain extent)
3. Increasing bit-depth (range resolution)
4. Increasing dynamic range (range extent)

Multiple images can be taken with shifts to increase the spatial resolution of images to subpixel accuracy (domain resolution) (Irani and Peleg, 1991)(Mann and Picard, 1994b). Image stitching methods mainly work by having multiple images captured by rotating the camera along the same point. Image stitching increases the number of pixels in an image extending its domain (Mann, 1993)(Mann and Picard, 1997)(Szeliski, 1996).

Mann and Ali (2016) models $q(x, y)$ as:

$$q(x, y) = \int_0^\infty q_{\text{ss}}(x, y, \lambda) s(\lambda) d\lambda \tag{8.1}$$

where q_{ss} is light falling on the sensor at continuous position (x, y) on the sensor plane consisting of wavelength λ, and $s(\lambda)$ is the spectral sensitivity of the sensor assumed to be same for each sensor across the sensor array. As can be seen from Eq.8.1, at a single location (x, y) the spectral information of $q_{\text{ss}}(x, y, \lambda)$ is lost. Thus under different noises such as sampling, aliasing and blurring which act on the domain (x, y) and other noises such as quantization and non-linear response acting on the range q, only the quantity $q_s(x, y)$ can be estimated.

There is a difference between what a sensor reads through the photo-electric conversion at the very beginning and what quantity is finally accessible for reading. The historical context for this is quite interesting where it was found that the television displays did not have a linear relationship between the voltage supplied to the cathode ray tube screens and the output luminance. The relationship was approximately modelled as a gamma function as

$$L = V^{2.5} \tag{8.2}$$

where L is the screen luminance and V is the voltage supplied. One solution was to cater for this non-linearity at the television level which would be cumbersome as it was already a commercialized product with many users. The other solution was to compensate for this non-linearity at the camera level which not many people had yet. The later solution was used and an 'opposite' non-linearity was applied to the captured images or videos. This is shown in Fig.8.1.

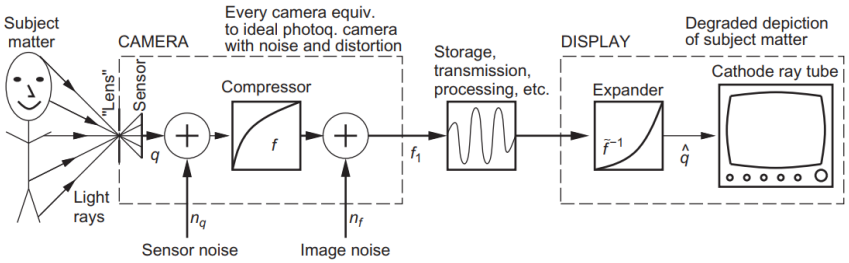


Figure 8.1: *Encoding and decoding at capturing and display side (Mann and Ali, 2016).*

The 'opposite' non-linearity came to be called as encoding defined as:

$$V = L^{\frac{1}{2.5}} \tag{8.3}$$

Now with the help of this encoding, the display could decode this through Eq.8.2. During capturing side, this not only allowed for linearizing the capture-display workflow but it also provided a way to encode the tones of the images with more perceptual uniformity. This happened because, it turned out later on, that HVS also perceive brightness in a non-linear logarithmic fashion which closely aligns with gamma encoding done during capturing.

In practice, the full gamma factor is not compensated for during encoding. Say a gamma of $\frac{1}{2.2}$ is encoded while decoding happens for 2.5 giving rise to an overall gamma of $\frac{2.5}{2.2}$ or 1.14. This is done in order to compensate for the reduction in contrast caused due to the dim surround in which the television screens were normally viewed (Bartleson and Breneman, 1967)(Pitt and Winter, 1974). This also results in the increase in purity of colors and shift in their dominant wavelength. Increase in purity helps for compensating the lost saturation but the shift in dominant wavelength shifts hue.

The digital sensor in cameras are very much limited in capturing the full dynamic range of the scenes commonly occuring in nature. So in order to estimate $q(x, y)$, the method to combine several exposure images to a single HDR image becomes the definition of HDR reconstruction.

It can be formulated as:

$$f_i(x, y) = f\left(c_i q(x, y)\right) \tag{8.4}$$

where c_i controls the exposure of the image and f is the function which maps incoming light on the sensor to the final accessible reading from the sensor which can be both linear or non-linear or just an identity mapping. In order to estimate $q(x, y)$ from $f_i(x, y)$:

$$q(x, y) = \frac{1}{c_i} f^{-1}\left(f_i(x, y)\right) \tag{8.5}$$

A more sophisticated formulation and much closer to reality will be to include the noise terms as:

$$f_i(x, y) = f\left(c_i q(x, y) +\right) + n_{\mathrm{f}_i} \tag{8.6}$$

In the presence of noise, getting images with different exposure helps us because in the underexposed images, those regions of the image are under the noise levels which would have been covered up by saturation noise at the regular exposure. Also in the overexposed images, those regions of the image are above the noise levels which would have been flooded by dark level noise.

Now by reversing the Eq.8.6:

$$q(x, y) = \frac{1}{c_i} f^{-1}\left(f_i(x, y) - n_{\mathrm{f}_i}\right) \tag{8.7}$$

The estimation of $q(x, y)$, $\hat{q}_i(x, y)$, can be written as:

$$\hat{q}_i(x, y) = \frac{1}{\hat{c}_i} \hat{f}^{-1}\left(f_i(x, y)\right) \tag{8.8}$$

where the estimations of f^{-1} has been introduced as $\hat{f}^{-1}$ but also for c_i as $\hat{c}_i$ because in some scenarios no information related to exposure is known.

In Fig.8.2, the overall process for the estimation of $q(x, y)$ is shown. Three captures of the same scene are taken with different exposures manifested by c_i. The function f is applied which can be linear or non-linear. Different instances of the same image noise is added n_{f_i}. In each of the images, inverse function f^{-1} is estimated as $\hat{f}^{-1}$ and apply it to get the linearized version of the same scene. The exposure difference can be undone by dividing by the estimated exposure constant $\hat{c}_i$, if estimated, or c_i if known. This step would bring back the differently exposed pixel values of the same scene to the same scale. Since there will be different noises and estimations, the pixel values can be combined with a weighting function w_i to get the estimated $\hat{q}_i(x, y)$. And then for illustration, estimated $\hat{c}_i$ capture can be simulated by multiplication with exposure. This highlights one of main benefits

of HDR, that since the HDR image tries to approximates the scene radiance, any given exposure that is required by the application can be simulated.

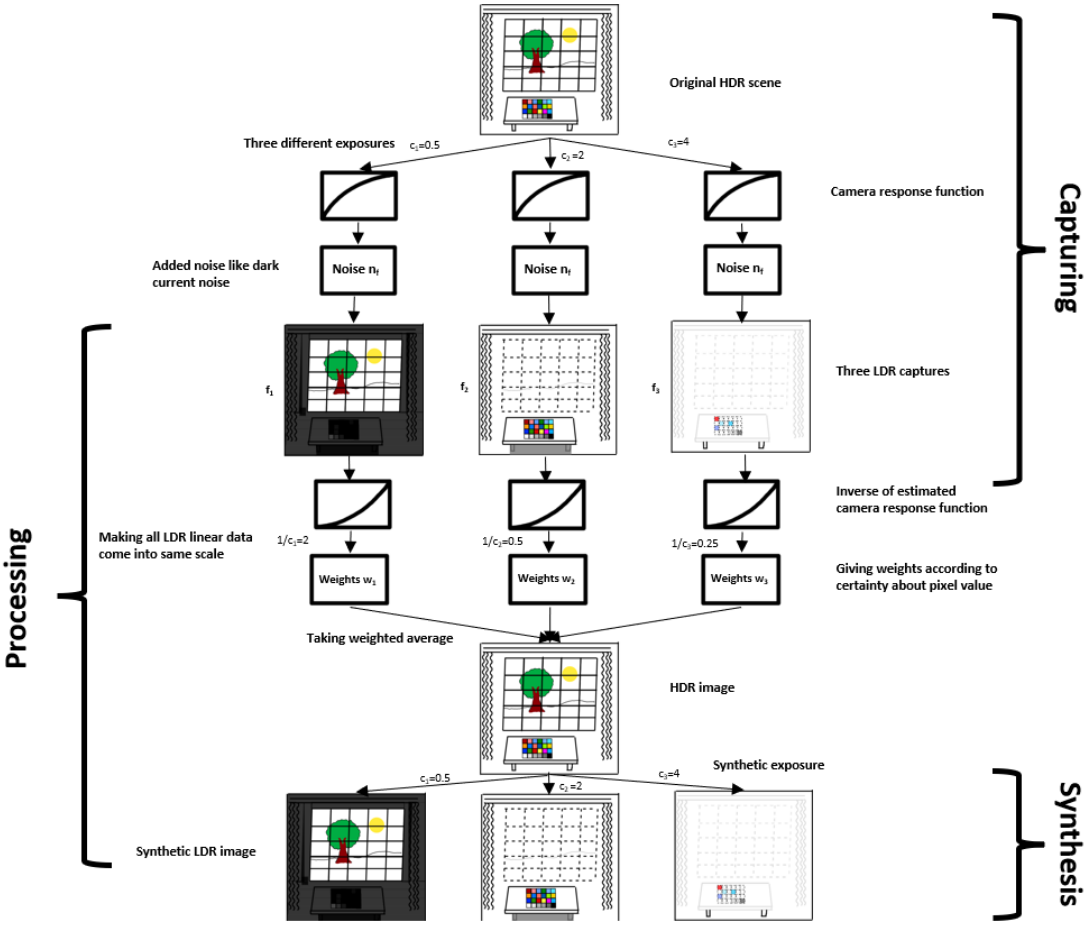


Figure 8.2: *Process of multi-exposure captures*

8.2 HDR imaging and dual-RGB

This section discusses the utilization of HDR imaging techniques into the workflow of dual-RGB. The final HDR images are termed as HDR-Dual-RGB and HDR-dual-RGB-PSO (when using PSO optimized light sources). The subsequent procedure is mentioned for HDR-Dual-RGB, but the same procedure can be carried out for HDR-Dual-RGB-PSO too just by utilizing the PSO optimized light sources.

Table 8.1: *Table showing the integration times for the captures of HDR-Dual-RGB setup*

Setup	Integration time (s)
Light source 1 exposure 1	1/1250
Light source 1 exposure 2	1/8
Light source 2 exposure 1	1/2500
Light source 2 exposure 2	1/20

Table 8.2: *Table showing the integration times for the captures of HDR-Dual-RGB-PSO setup*

Setup	Integration time (s)
Light source 1 exposure 1	1/60
Light source 1 exposure 2	2
Light source 2 exposure 1	1/250
Light source 2 exposure 2	1/1.6

8.2.1 Dark noise correction

For dark noise correction, 10 captures are taken with lens cap on at each of the exposure time that is being used. And then an average is taken at each of the pixel to create an average dark noise image for each capture. This is shown in Fig.8.3.

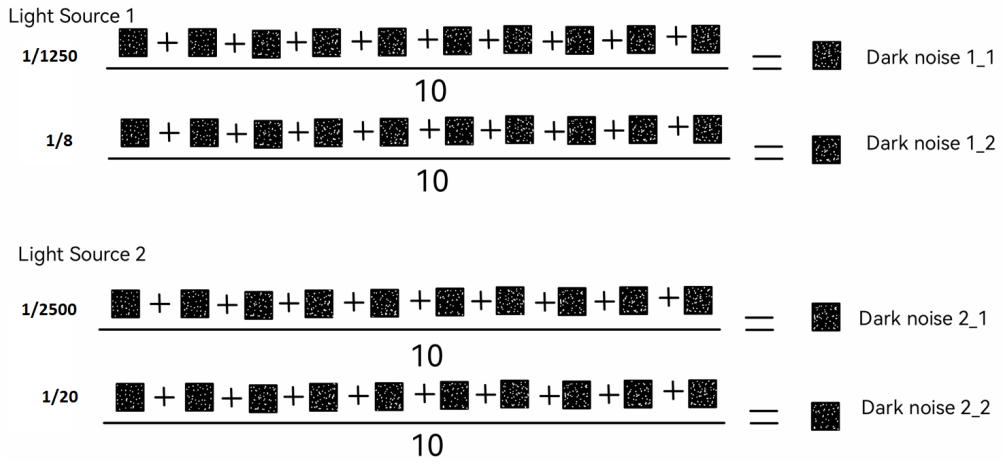


Figure 8.3: *Dark noise correction for HDR-Dual-RGB imaging. 10 dark captures are taken for each of the exposure times involved and are averaged. Same workflow is carried out for HDR-Dual-RGB-PSO with PSO optimized light sources.*

8.2.2 White flat-fielding image calculation

The procedure for calculating the white image for flat-fielding is shown in Fig.8.4. Two white images are needed to be calculated for two of the light sources involved. For each of the light sources, two captures are taken with different exposure times which covers the scene dynamic range (that will be captured next) well. Each of the images is corrected for dark noise with their respective average dark noise images. From both of these images, a weighting function shown in Fig.8.5 is applied. This function gives weights in the range of 0 to 16383 (which also could have been from 0 to 1 equivalently) according to the following equation:

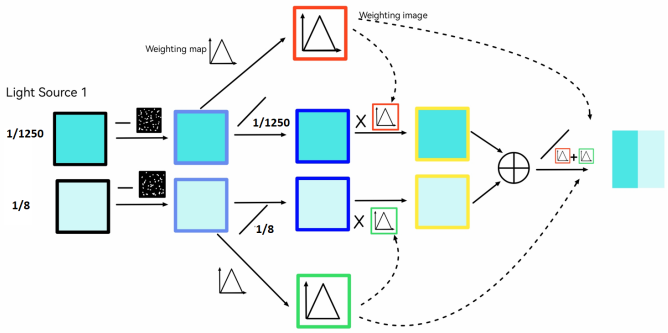
$$f(x) = \begin{cases} x, \text{for } x < 8192 \\ 16383 - x, \text{for } x \geq 8192 \end{cases} \tag{8.9}$$

where x is the dark noise corrected white low-dynamic range image pixel value. This function is used to get a weighting image.

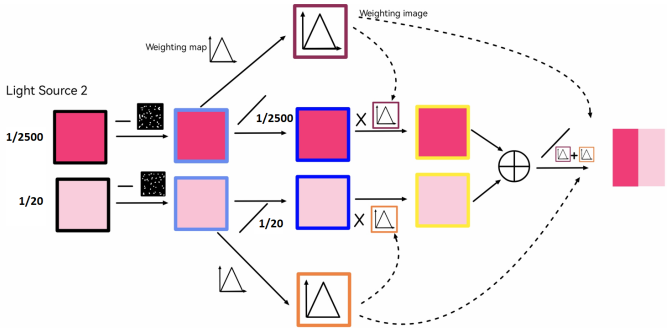
The weighting image is kept aside the low dynamic range white images are divided by their respective exposure time. This allows both of these images to remove the difference in exposure and come on the same scale of irradiance. Then a weighted average of both of these images are taken according to the weighting image that was created previously. This is done by multiplying pixel-wise the images with their respective weighting images and dividing by the sum of the weighting images pixel-wise. After doing this for both of the light sources images, two HDR white images are generated.

8.2.3 HDR dual-RGB image calculation

Once the dark noise is corrected and white images are calculated for each of the light sources, the HDR images of the scene can be calculated as shown in Fig.8.6. A similar procedure is followed as was followed for the calculation of the white image. For each of the light sources low dynamic range scene images, dark noise is corrected by the average dark noise captures at their corresponding exposure times. The weighting images are calculated similar to the white image calculation and kept aside.The low dynamic range images are converted to a common scale of irradiance by dividing them by their corresponding exposure time. The white images calculated in the previous step is divided by the images to flat-field them. Then the images are combined with a weighted average according to their corresponding weighting images. These two images then serve as the two images for our HDR-Dual-RGB pipeline. The same procedure as mentioned for the two images in Chapter 7 for dual-RGB and dual-RGB-PSO can be carried out for these two HDR-Dual-RGB images.



(a) *White calculation for light source 1*



(b) *White calculation for light source 2*

Figure 8.4: *White calculation for HDR-Dual-RGB imaging. Two open-gate white images are taken for each of the light sources. Afterwards, dark noise is subtracted. Weighting image is calculated according to the weighting function and the input pixel value. The images are divided by their corresponding exposure times. The images are combined together by taking a weighted mean for each pixel according to the weighting image calculated. Same workflow is carried out for HDR-Dual-RGB-PSO with PSO optimized light sources.*

8.3 Results

For increasing the dynamic range of the colourcheckers that have been used till now, a neutral density (ND) filter of optical density 2 is used covering half the amount of patches on both training and testing colourcheckers. The Fig.8.7 shows the chromaticities of the training and testing data as well as the gamut of sRGB colour space. There are 42 patches which lie outside the gamut of sRGB hence limiting its reproduction visualization for this report. The Fig.8.8 and Fig.8.12 show the patch-wise ΔE_{00} error for training and testing data similar to previous chapter. For dual-RGB, the training mean error comes out to be 2.30 and testing error of 2.87. For dual-RGB-PSO, the training mean error is 2.35 and testing error of 3.09.

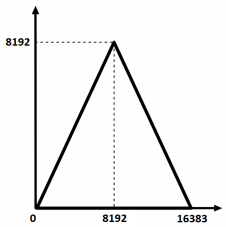


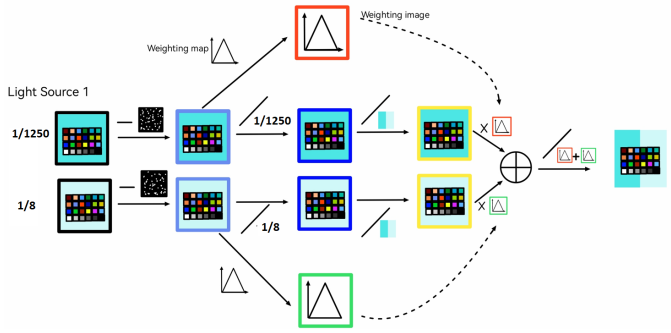
Figure 8.5: *Weighting function used for HDR-Dual-RGB and HDR-Dual-RGB-PSO imaging.*

The Fig.8.9 and Fig.8.13 shows a reproduction of ground-truth, dual-RGB predicted, and dual-RGB-PSO predicted colours in sRGB colour space for training and testing data. Each patch consists of four sub-blocks in the form $\left(\begin{smallmatrix} GT & DRP \\ GT & DR \end{smallmatrix}\right)$, where GT is for ground truth, DRP for dual-RGB-PSO, and DR for dual-RGB sub-block. Since the visualizations are done in sRGB colour space, visualizing the side with ND filter is difficult.

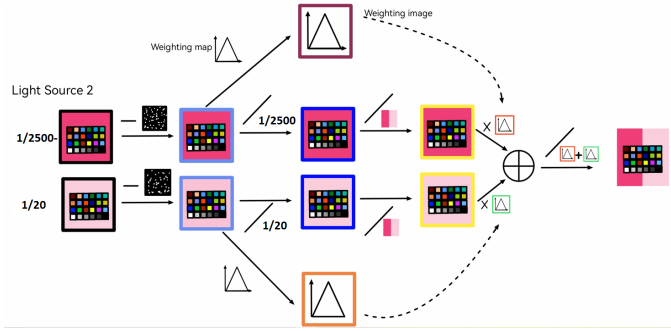
The Fig.8.10 shows the plot for Lightness vs. Chroma and a^* vs. b^* in CIELAB for training data for both dual-RGB and dual-RGB-PSO. Notice the clustering of points near low Lightness region. These are the points towards the side of ND filter. All the points in this cluster have very low errors. A trend of having reduced errors for higher Lightness values is observed on the side without ND filter for both dual-RGB and dual-RGB-PSO. For the plots of a^* vs. b^*, as the points move away from the center, a trend of increasing error is observed. In order to have a better look at these points, Fig.8.11 shows a zoomed in version for these same plots. In these plots, a similar trend of increasing error is observed as the point moves away from the center of the a^* vs. b^* plot. However in this region, as the Lightness increases, the errors are gradually increasing. The Fig.8.14 and Fig.8.15 shows the plot for Lightness vs. Chroma, a^* vs. b^* for testing data, and their zoomed in version. The same trend of low errors for the ND-filters side are observed as well as a higher errors for higher chromatic values are seen.

8.4 Discussion

This chapter discussed the background on HDR imaging and how it allows to extend the limitation of the dynamic range of the sensor. One method for HDR imaging discussed was through the capture of multiple LDR images and by combining them according to some weighting function. This is done after removing any embedded non-linearity in sensor and by dividing the images by their exposure time. A workflow on embedding the HDR techniques in to the dual-RGB working was discussed. For this purpose, several LDR images were captured under both



(a) *HDR creation for light source 1*

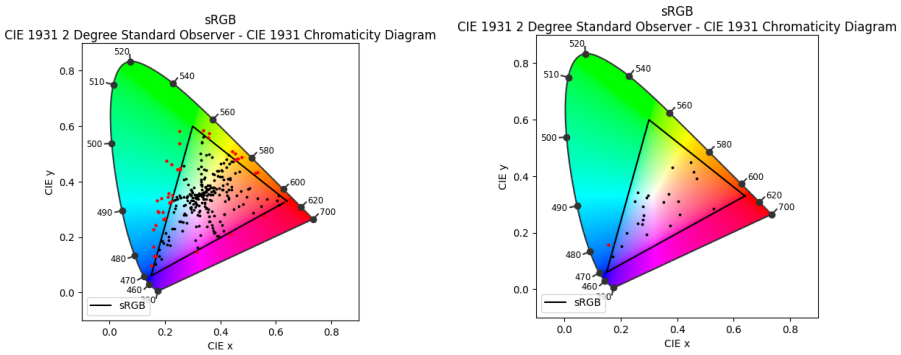


(b) *HDR creation for light source 2*

Figure 8.6: *Final HDR image creation for HDR-Dual-RGB imaging. Two images are taken with different exposures for each of the light sources. They are corrected for dark noise. A weighting image is calculated according to the weighting function and the pixel values. The images are divided both by the exposure time and the white image. A weighted mean is calculated according to the weighting image. Same workflow is carried out for HDR-Dual-RGB-PSO with PSO optimized light sources.*

dual-RGB and dual-RGB-PSO optimized light sources. The capturing scene was created by using an ND filter on half of the patches on the colourcheckers used. Since the images were in RAW format, there was no gamma function needed to estimated. The workflow was described through the help of Fig.8.3 - 8.6. The subsequent HDR images were called as HDR-Dual-RGB and HDR-Dual-RGB-PSO images.

The Fig.8.8 and Fig.8.12 showed the patch-wise error for training and testing data. It can easily be seen that the patches under the ND-filter give the least errors. This was also highlighted in Fig.8.10 and Fig.8.14 where the ND filter data was clustered towards the low Lightness region. Compare this with Fig.7.8, where the low Lightness values had the maximum error. The HDR gets an upper edge because it takes separate captures, targeting different regions of luminance levels.



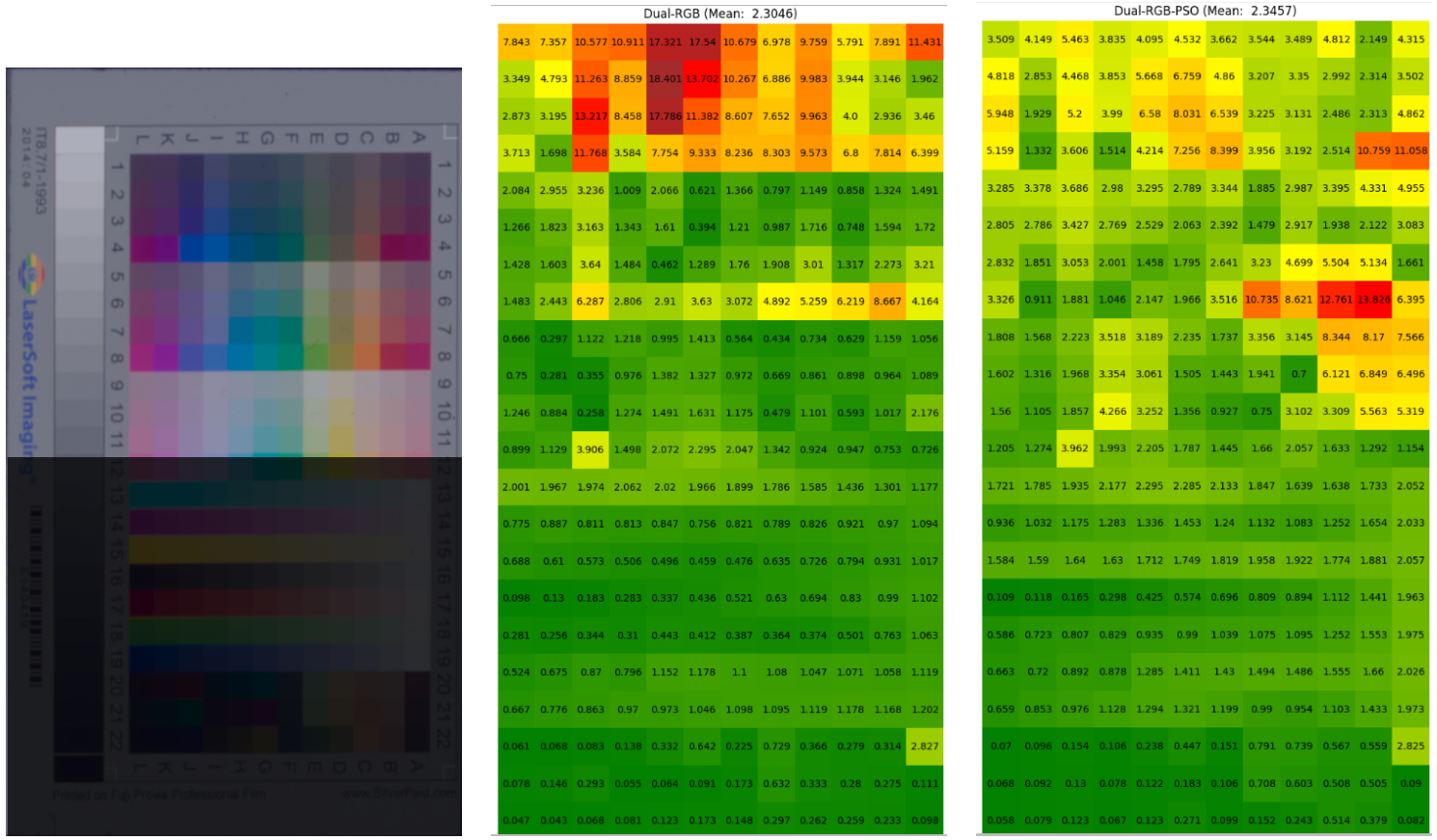
(a) (b)

Figure 8.7: *HDR data points on chromaticity diagram with a comparison to sRGB colour space. (a) shows the chromaticity points for training data and (b) shows the chromaticity points for testing data. The black points correspond to in-gamut and red points correspond to out-of-gamut chromaticities.*

This allows the low Lightness patches to send as much signal to the capturing device as a high Lightness patch can send.

The results for both HDR-Dual-RGB and HDR-Dual-RGB-PSO are comparable, which is consistent with the results of Chapter 7 for dual-RGB and dual-RGB-PSO captures.

Considering the above discussions and visualizations, it can be seen that embedding HDR workflow in dual-RGB technique can have significant advantage for capturing scenes having a high dynamic range.



(a) *Train colorchecker* **(b)** *Train error HDR-Dual-RGB* **(c)** *Train error HDR-Dual-RGB-PSO*

Figure 8.8: *Comparing errors for each patch in training for HDR-Dual-RGB and HDR-Dual-RGB-PSO. Note that the colorchecker image is digitally modified to represent where was the ND filter placed.*

Figure 8.9: *Train data reproduction from ground-truth data and predictions from HDR-Dual-RGB-PSO and HDR-Dual-RGB imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & DRP \\ GT & DR \end{smallmatrix}\right)$*, where GT is for ground truth, DR for HDR-Dual-RGB, and DRP for HDR-Dual-RGB-PSO block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. Patches with yellow boundaries are out-of-gamut chromaticities for sRGB colour space limiting the actual demonstration.*

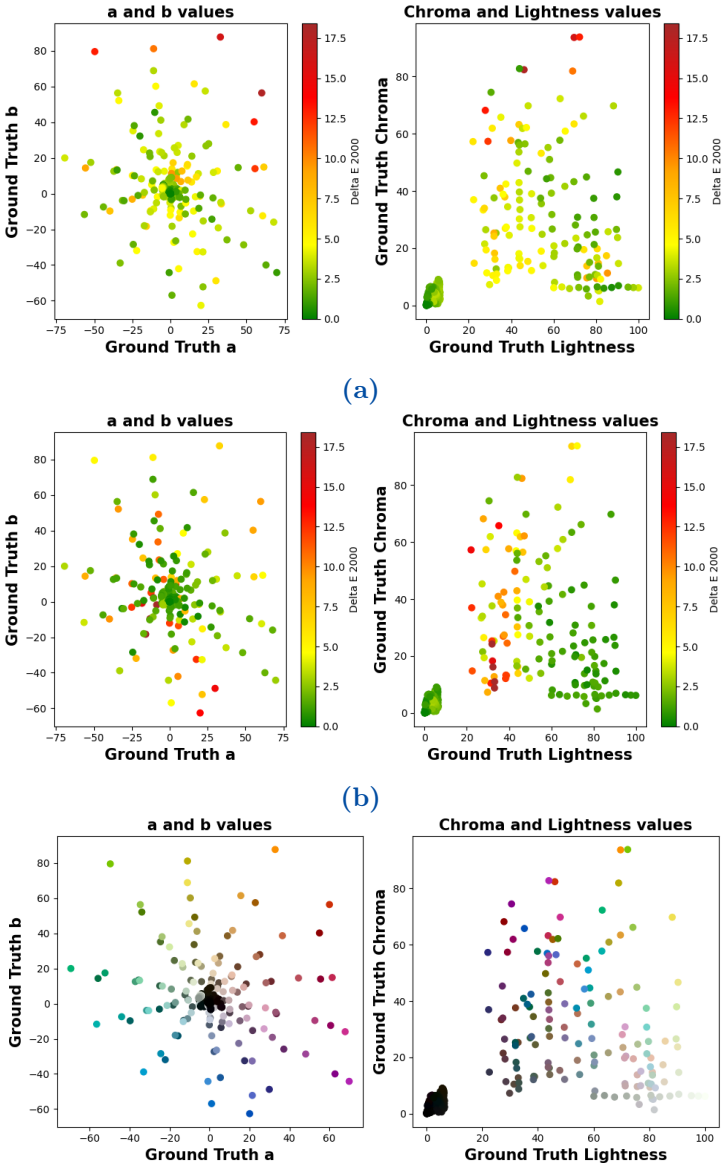


Figure 8.10: *Plots for HDR-Dual-RGB and HDR-Dual-RGB-PSO training patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for HDR-Dual-RGB-PSO.(b) represents the same set of plots for HDR-Dual-RGB captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a* vs. b* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

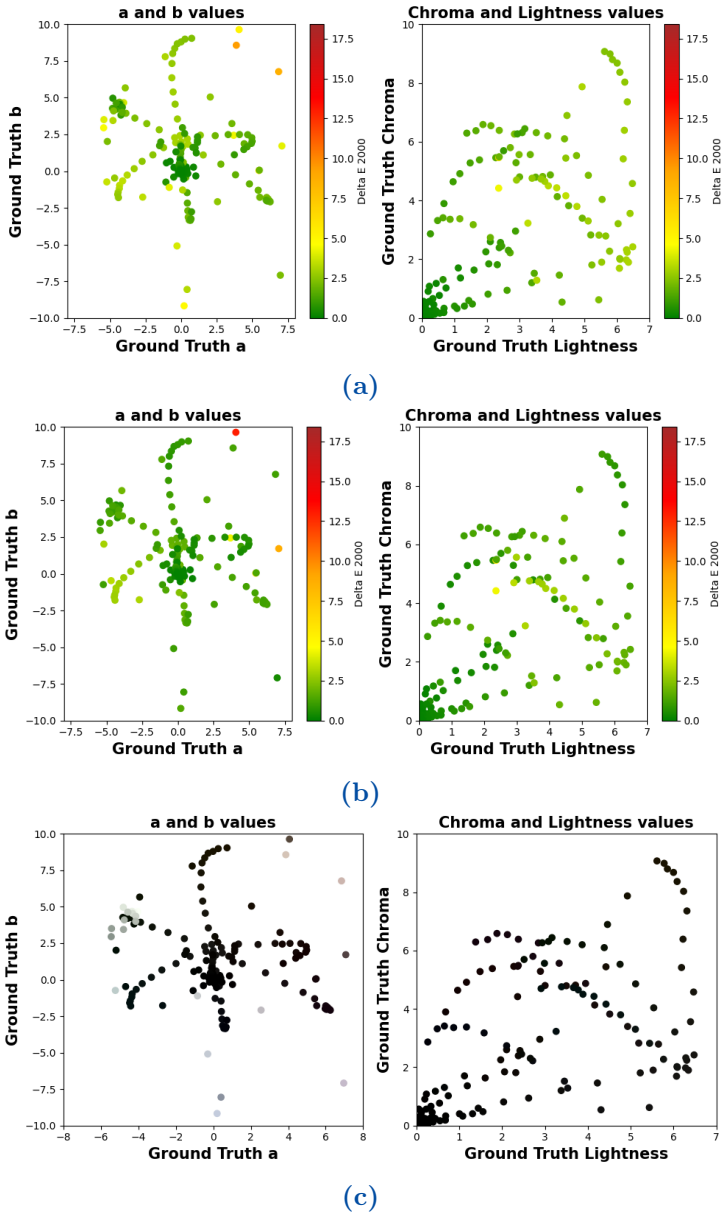


Figure 8.11: *Plots for HDR-Dual-RGB and HDR-Dual-RGB-PSO training patches zoomed in. (a) shows a* vs. b* and Lightness vs. Chroma plots for HDR-Dual-RGB-PSO.(b) represents the same set of plots for HDR-Dual-RGB captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a* vs. b* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

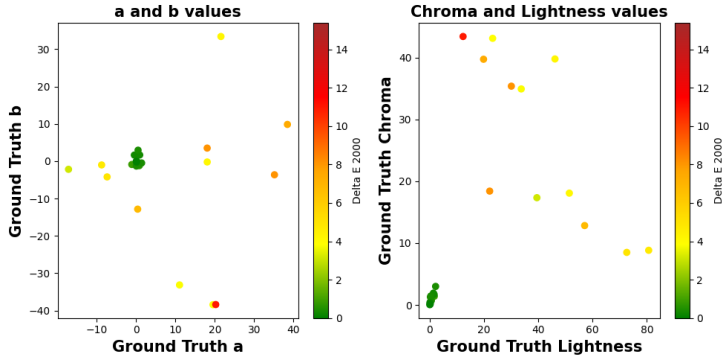
(a) *Test colorchecker*

(b) *Test error HDR-Dual-RGB*

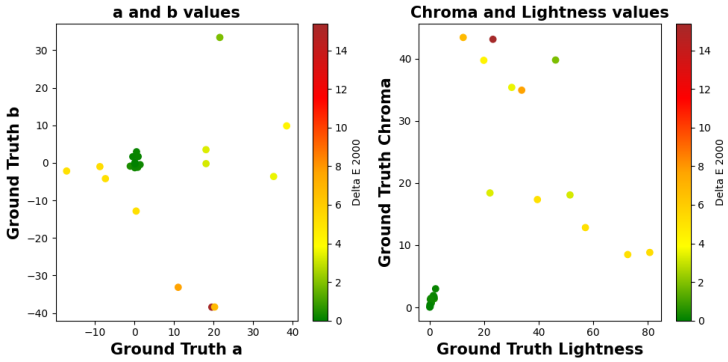
(c) *Test error HDR-Dual-RGB-PSO*

Figure 8.12: *Comparing errors for each patch in testing for HDR-Dual-RGB and HDR-Dual-RGB-PSO. Note that the colorchecker image is digitally modified to represent where was the ND filter placed.*

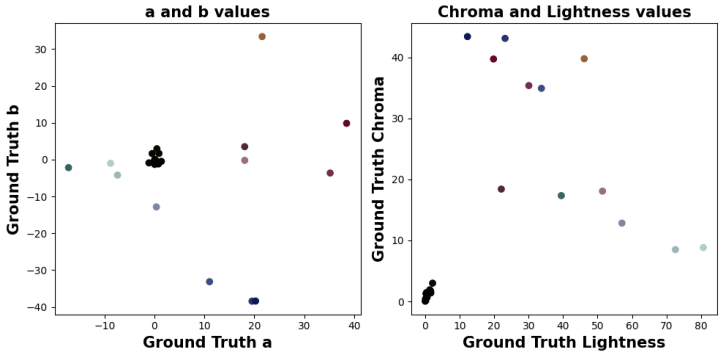
Figure 8.13: *Test data reproduction from groundtruth data and predictions from HDR-Dual-RGB and HDR-Dual-RGB-PSO imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & DRP \\ GT & DR \end{smallmatrix}\right)$*, where GT is for ground truth, DRP for HDR-Dual-RGB-PSO, and DR for HDR-Dual-RGB block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. The yellow boundary represents out-of-gamut-colour for sRGB colour space*



(a)

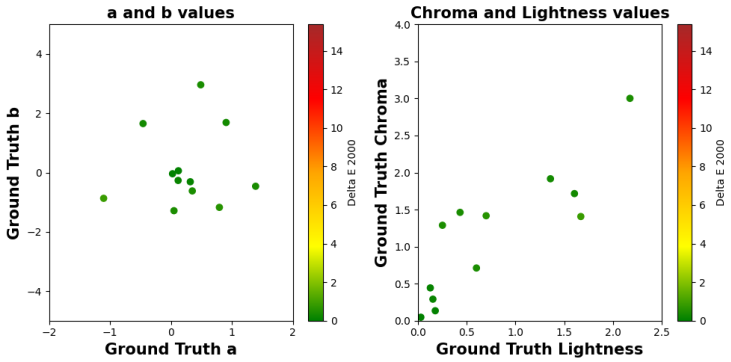


(b)

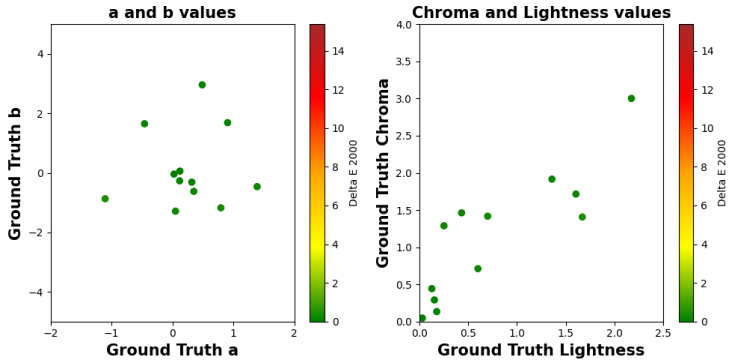


(c)

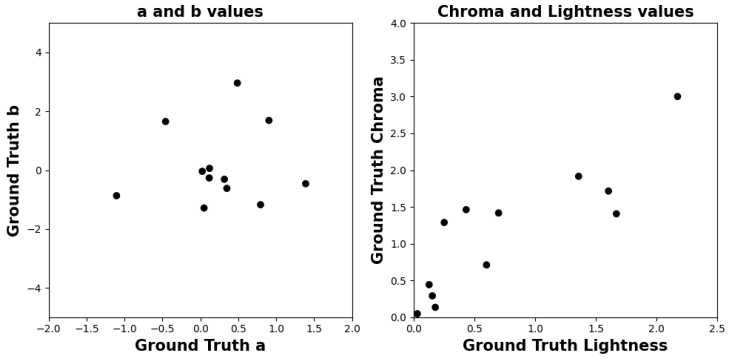
Figure 8.14: *Plots for HDR-Dual-RGB and HDR-Dual-RGB-PSO testing patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for HDR-Dual-RGB.(b) represents the same set of plots for HDR captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a* vs. b* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*



(a)



(b)



(c)

Figure 8.15: *Plots for HDR-Dual-RGB and HDR-Dual-RGB-PSO testing patches zoomed in. (a) shows a* vs. b* and Lightness vs. Chroma plots for HDR-Dual-RGB.(b) represents the same set of plots for HDR captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. (c) shows similar a* vs. b* and Lightness vs. Chroma plots but the colours represent the sRGB colours of the ground-truth patches.*

9 | Reflective setup

You are never too old to set another goal or to dream a new dream.

C.S. Lewis

Since the work discussed till now have been related to photographic films, the dual-RGB multispectral imaging setup is extended to reflective objects with the following objectives:

- To check whether similar results are obtained for reflective objects as were obtained for transmissive objects.
- To extend the setup to reflective objects such that it can easily be converted back and forth between transmissive and reflective mode.

The easiest way to convert the LED setup for reflective object should have been to rotate the integrating sphere. Since the LED system was originally meant for only photographic films, rotating the integrating sphere was not in the original use-case of the LEDs setup. Movement of integrating sphere posed a high probability of changing the LEDs input power and hence the LEDs spectra. This lead to the constraint that for the extension to reflective objects, nothing could be moved on the LEDs device.

This lead to a setup as shown in Fig.9.1. The reflecting mirror from the overhead projector is placed at the location of where the films were used to be placed. This mirror then can be used to project the light outwards towards a reflective objective. The use of mirror allows a very easy way to change the mode of imaging between reflective and transmissive.

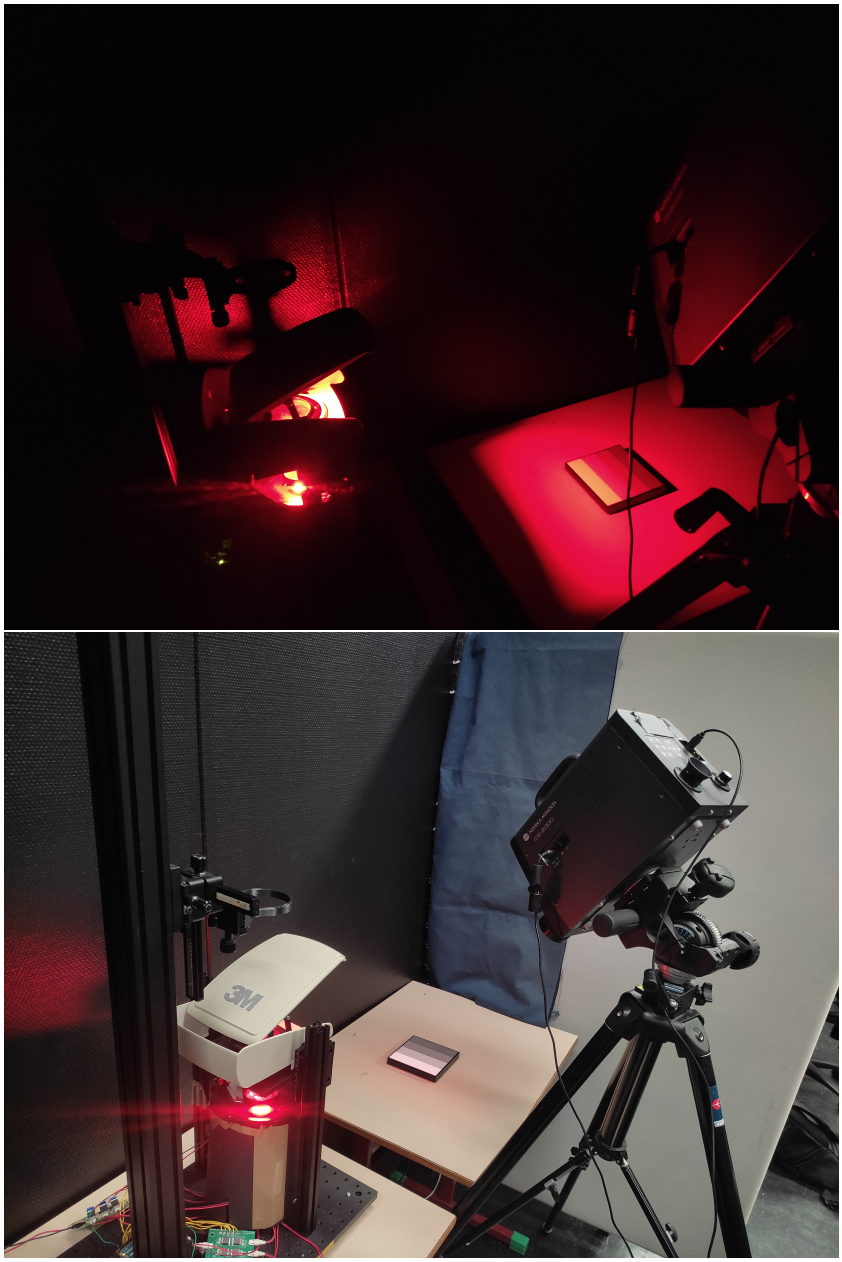


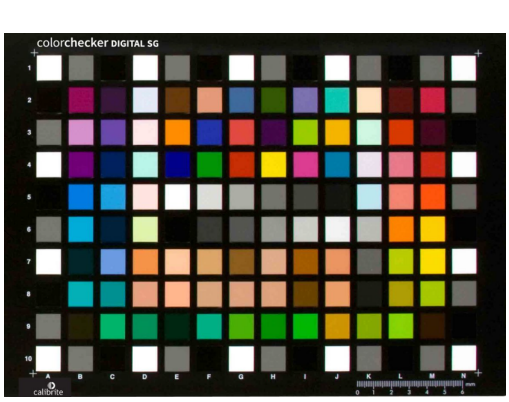
Figure 9.1: *The converted setup of the LEDs light source to be used for reflective objects. This particular image shows an example of the reading of LED spectra impinged on a gray patterned tile with the help of a telespectroradiometer.*

9.1 Simulation

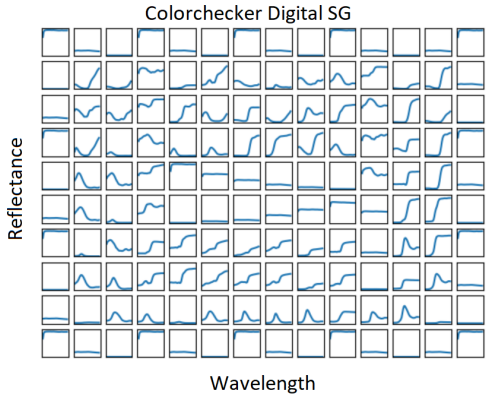
The simulations are run for the reflective setup similar to what was done in Chapters 5 and 6, running the optimizations of LEDs at full intensities and then running PSO for exploring change of intensities.

9.1.1 Data

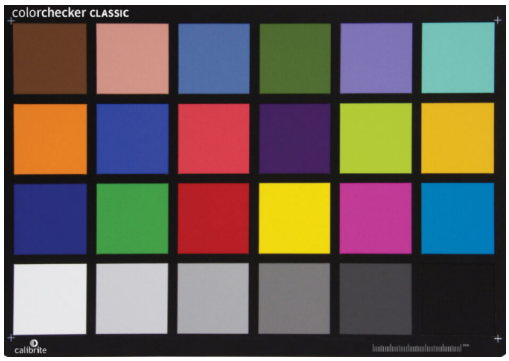
For training the Colorchecker Digital SG and for testing the Colorchecker Classic are used. These are shown in Fig.9.2. For reflectance, the data from BabelColor (2023) is used.



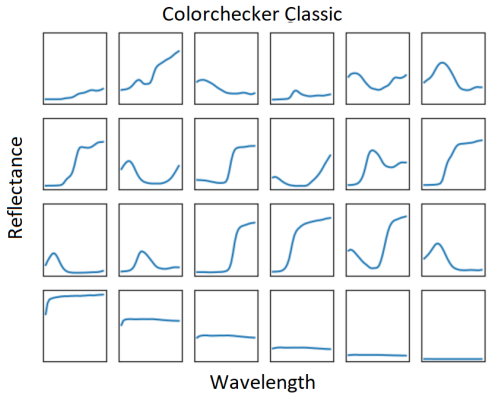
(a) *Colourchecker Digital SG (training)*



(b) *Colourchecker Digital SG reflectances*



(c) *Colourchecker Classic (testing)*



(d) *Colourchecker Classic reflectances.*

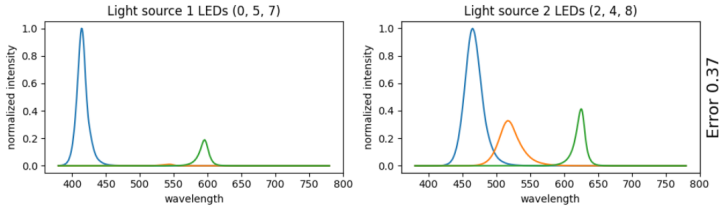
Figure 9.2: *Training and testing data for reflectance setup.*

9.1.2 Full intensities of light sources

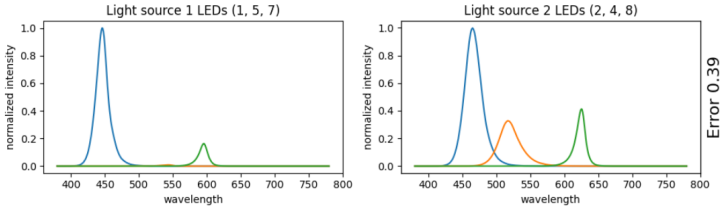
The simulations are run for all the 630 different combinations of light source pairs and mention the top three giving the least errors in Table.9.1 and plotted in Fig.9.3. The best light source pair is exactly the same that was obtained earlier for the transmissive photographic films. The resulting dual-RGB channel sensitivities are plotted in Fig.9.4.

Table 9.1: *Best three combination of light source pairs for dual-RGB*

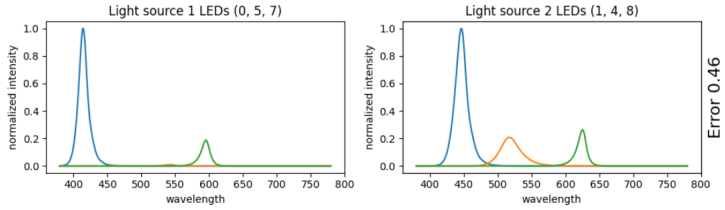
Light source 1	Light source 2	Training error
(1, 6, 8)	(3, 5, 9)	0.37
(2, 6, 8)	(3, 5, 9)	0.39
(1, 6, 8)	(2, 5, 9)	0.46



(a) *First (Best error)*

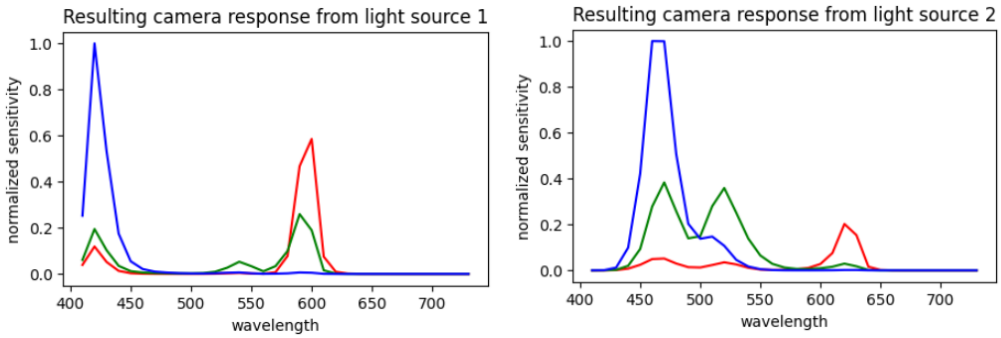


(b) *Second*



(c) *Third*

Figure 9.3: *Top three pair of light sources*



(a) *Camera response from light source 1* (b) *Camera response from light source 2*

Figure 9.4: *Camera spectral response at full intensity of light sources*

9.1.3 Optimization through PSO

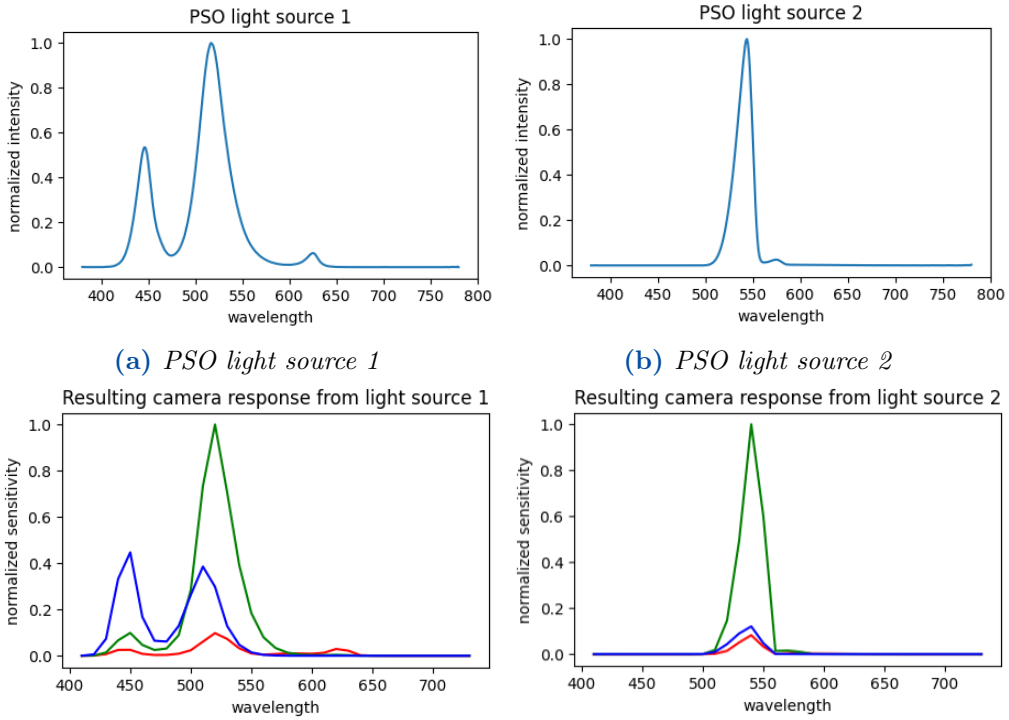
Now similar to what was done in Chapter 6, the intensities of LEDs are optimized in the light source pairs through PSO to find the pair of LED combinations and their corresponding intensities which minimizes the error. The resulting optimized light sources and their corresponding multispectral channels are shown in Fig.9.5.

Table 9.2: *Simulation details and comparison of PSO optimized intensity with full intensities and 3-channel system*

Condition	3-channel	Full intensity	PSO intensity
Source 1 LEDs	D65	(1, 6, 8)	(1, 6, 7)
Source 2 LEDs	N/A	(3, 5, 9)	(2, 4, 9)
Source 1 intensities	N/A	(1.0, 1.0, 1.0)	(0.07, 0.58, 1.0)
Source 2 intensities	N/A	(1.0, 1.0, 1.0)	(0.12, 0.56, 1.0)
$\Delta TrainE_{00}$	1.23	0.37	0.25
$\Delta TestE_{00}$	1.29	0.44	0.28

9.1.4 Reflective vs. transmissive optimized channels

The light sources and the resulting channels for full intensities are exactly the same for both reflective and transmissive. The PSO optimized light sources are compared in Fig.9.6 and the channels are compared in Fig.9.7.



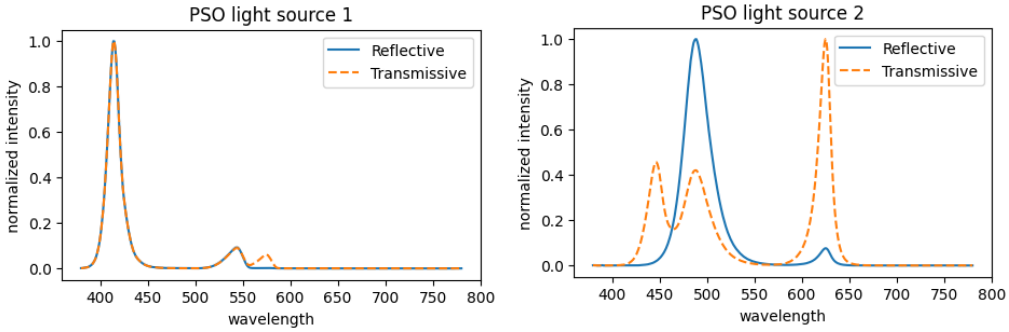
(a) *PSO light source 1*

(b) *PSO light source 2*

(c) *Channels from light source 1*

(d) *Channels from light source 2*

Figure 9.5: *PSO light sources and the corresponding multispectral channels*



(a) *PSO light source 1*

(b) *PSO light source 2*

Figure 9.6: *Reflective vs. transmissive light sources comparison*

9.2 Results for real captures

Real captures were conducted for the light sources optimized for dual-RGB and dual-RGB-PSO for the reflective colourcheckers. The light sources optimized for dual-RGB-PSO for films was also utilized to take another set of captures in order

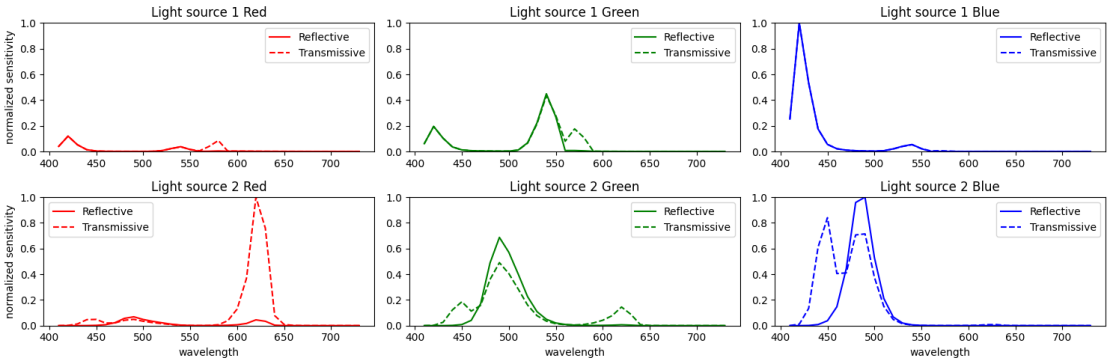


Figure 9.7: *Reflective vs. transmissive PSO optimized channels comparison*

to see whether the same optimized light can be used for both transmissive and reflective objects.

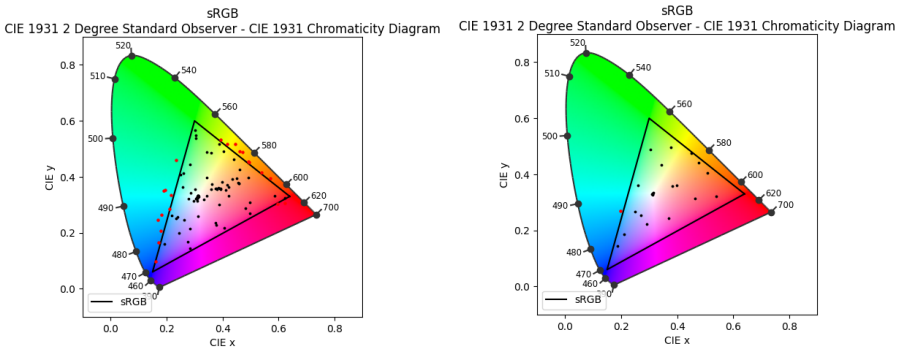
The Fig.9.8 shows the chromaticities of both training and testing data as well as the gamut of sRGB. The figure shows that there are 21 data points in training and one in testing which are out of gamut for this sRGB visualization. The Fig.9.9 and Fig.9.12 shows the ΔE_{00} for each of the training and testing patches as well as mean ΔE_{00} values. The mean training errors are 2.08, 1.84, and 2.18 and the mean testing errors are 2.25, 1.72, and 2.14 for dual-RGB-PSO, dual-RGB, and dual-RGB-PSO (for films).

The Fig.9.10 and Fig.9.13, shows the colour reproduction of ground-truth and predicted patches. Each block has four sub-blocks in the form $\left(\begin{smallmatrix} GT & DRPT \\ DRP & DR \end{smallmatrix}\right)$, where GT is for ground truth, $DRPT$ for dual-RGB-PSO (light sources optimized for films), DR, and DR for dual-RGB-PSO block.

The Fig.9.11 shows the plot for Lightness vs. Chroma and a^* vs. b^* in CIELAB for training data. For Lightness vs. Chroma plots, common observation of having least errors at bottom right is found across all the three setups. In a^* vs. b^* plots, the errors are observed to be equally distributed across all the directions from center. This is same for the test plots in Fig.9.14.

9.3 Discussion

This chapter discussed conversion of the existing dual-RGB imaging setup for photographic films to capture reflective objects. This was motivated by the exploration of whether the workflows undertaken for photographic films can be utilized for reflective objects. Moreover, the easy convertibility of the LEDs light source from photographic mode to reflective mode pave a way for its utilization for future research on reflective objects.



(a) (b)

Figure 9.8: *Data points of reflective data on chromaticity diagram with a comparison to sRGB colour space. (a) shows the chromaticity points for training data and (b) shows the chromaticity points for testing data. The black points correspond to in-gamut and red points correspond to out-of-gamut chromaticities.*

Simulations were conducted similar to Chapter 5 and 6 for the reflective colourcheckers. The simulation of dual-RGB at full intensities converged to the exact same solution as in Chapter 5 for photographic films. For dual-RGB-PSO simulation, the final multispectral channels also converged to, more or less, the same shapes. The setup of dual-RGB and dual-RGB-PSO gave a training error of 0.37 and 0.44 and testing error of 0.25 and 0.28 respectively. This is in comparison to the three-channel imaging simulated for CIE D65 whose training error was 1.23 and testing error was 1.29. This was summarized in Table 9.2. This shows a clear advantage of dual-RGB and dual-RGB-PSO over three-channel imaging.

For the real captures, the three-channel images were not compared due to limitation of time. The captures were conducted for dual-RGB, dual-RGB-PSO light sources optimized for reflective colourcheckers and dual-RGB-PSO light sources optimized for photographic films. The setup of dual-RGB-PSO and dual-RGB optimized on reflective colourcheckers gave a training error of 2.08 and 1.84 and testing error of 2.25 and 1.72 respectively. The dual-RGB-PSO optimized for photographic films gave a training error of 2.18 and testing error of 2.14. Notice that the testing error for dual-RGB has reduced than the training error. This might highlight a limitation in this experimental session that the training colourchecker contains a semi-gloss version of the testing data. Though, due to the semi-gloss, the chromaticity would be impacted but the trained model can still be biased.

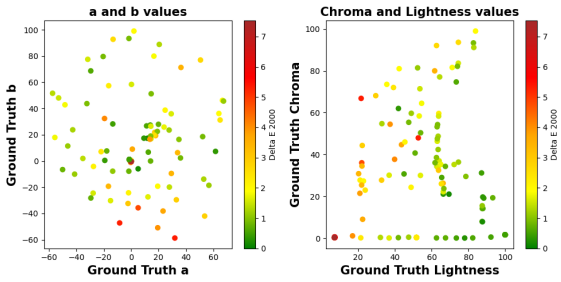
Through the visualizations, similar trends were shown as were found in Chapters 5 and 6. Errors were found to be decreasing as Lightness of the patch increased. Increasing the Chroma also tended to increase the error.

With the above discussion and the visualizations in this chapter, it is seen that

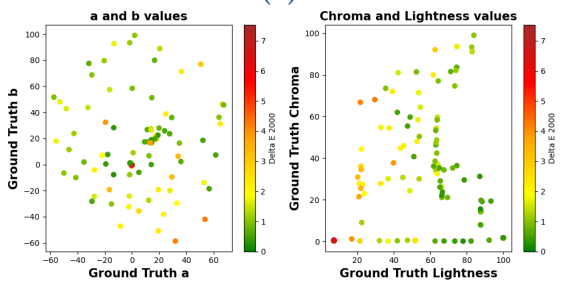
(a) *Train colorchecker* **(b)** *Train error dual-RGB-PSO* **(c)** *Train error dual-RGB* **(d)** *Train error dual-RGB-PSO transmissive*

Figure 9.9: *Comparing real capture errors for each patch in training for dual-RGB-PSO, dual-RGB, and dual-RGB-PSO (optimized for films).*

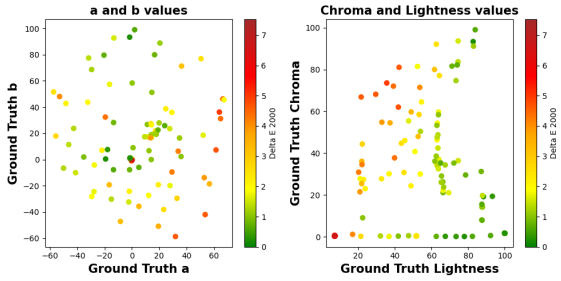
Figure 9.10: *Train data reproduction from ground-truth data and predictions from dual-RGB, dual-RGB-PSO and dual-RGB-PSO (light sources optimized for films) imaging. Each patch consists of four blocks in the form* $\begin{pmatrix} GT & DRPT \\ DRP & DR \end{pmatrix}$*, where GT is for ground truth, DRPT for dual-RGB-PSO (light sources optimized for films), DR for dual-RGB, and DRP for dual-RGB-PSO block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. Patches with yellow boundaries are out-of-gamut chromaticities for sRGB colour space limiting the actual demonstration.*



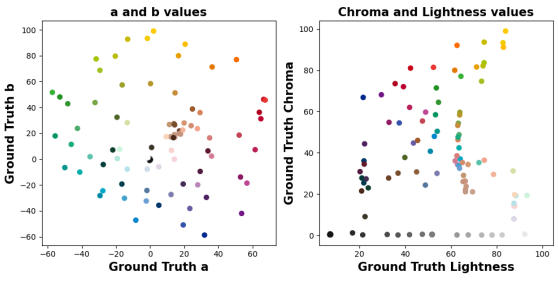
(a)



(b)



(c)



(d)

Figure 9.11: *Real captures plots for dual-RGB-PSO, dual-RGB and dual-RGB-PSO (light source optimized for films) training patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for dual-RGB-PSO.(b) and (c) represent the same set of plots for dual-RGB and three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. For (d) the colour codes are the sRGB representation of patches*

(a) *Test colorchecker*

(b) *Test error dual-RGB-PSO*

(c) *Test error dual-RGB*

(d) *Test error dual-RGB-PSO (optimized for films)*

Figure 9.12: *Comparing errors for each patch in testing for dual-RGB-PSO, dual-RGB, and dual-RGB-PSO (light sources optimized for films).*

Figure 9.13: *Real captures test data reproduction from ground-truth data and predictions from dual-RGB, dual-RGB-PSO, and dual-RGB-PSO- (light sources optimized for films) imaging. Each patch consists of four blocks in the form* $\left(\begin{smallmatrix} GT & DRPT \\ DRP & DR \end{smallmatrix}\right)$*, where GT is for ground truth, DRPT for dual-RGB-PSO (light sources optimized for films), DR, and DR for dual-RGB-PSO block. The tristimulus data is produced by converting to sRGB colourspace, scaling and quantizing the data to 8-bits. Patches with yellow boundaries are out-of-gamut chromaticities for sRGB colour space limiting the actual demonstration.*

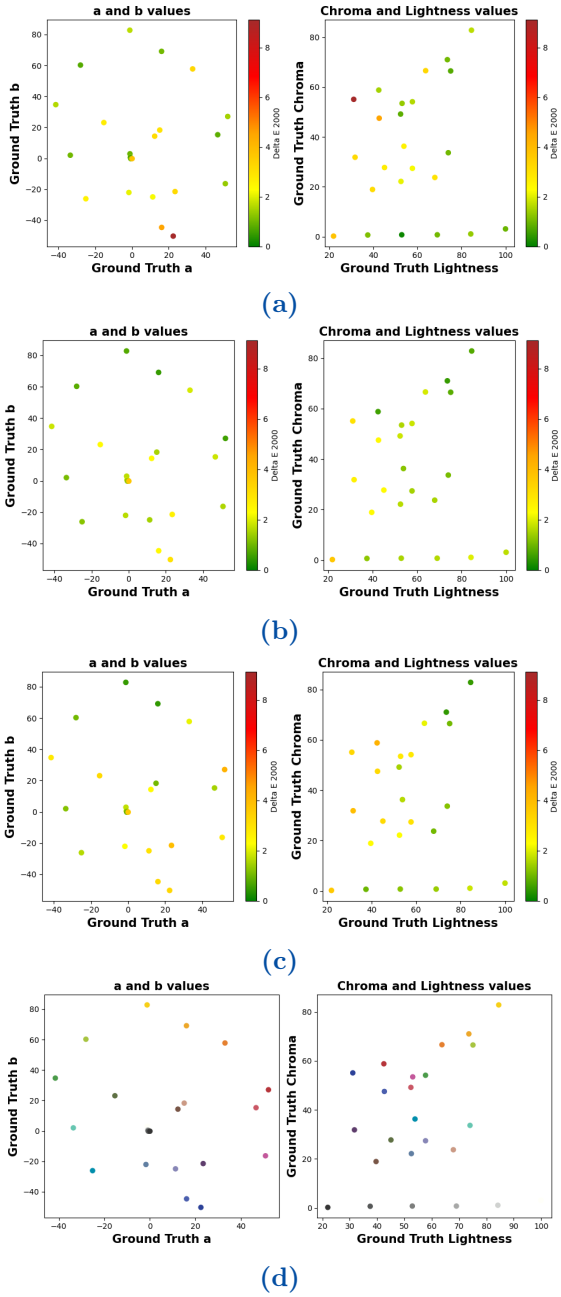


Figure 9.14: *Real capture plots for dual-RGB-PSO, dual-RGB and dual-RGB-PSO (optimized for light sources) testing patches. (a) shows a* vs. b* and Lightness vs. Chroma plots for dual-RGB-PSO.(b) and (c) represent the same set of plots for dual-RGB and three-channel captures. The dots represent the ground-truth data and the color code represents the mean ΔE_{00} error. For (d) the colour codes are the sRGB representation of patches.*

the simulation workflow conducted for the photographic films can be extended towards reflective objects. Also the light sources optimized for photographic films, gave an error on training and testing which can be called similar, depending on the application, to the ones optimized specifically for reflective colourcheckers.

10 Conclusion

Every end is a new beginning.

Unknown

This thesis explored the usage of an LEDs-based multispectral imaging system called as dual-RGB. In dual-RGB multispectral imaging, two images are captured from a trichromatic camera either using two external filters or two different light sources. The combination of trichromatic camera with two external filters or light sources give rise to a six-channels multispectral imaging system. One of the applications for such a system is in cultural heritage for digitally preserving the appearance of artworks.

This thesis work was mainly divided into six different parts. The work on these parts were not done strictly in order. The first part dealt with characterizing different devices that were required for simulations. This included the characterization of the Nikon D610 camera for its response function and spectral response function. Also the 10-LEDs light source that was used for illumination was characterized for its spectra and and its thermal stability. The transmittances of the training and testing colourcheckers were measured by creating a setup utilizing an overhead projector for enlarging the colourchecker image.

The second part dealt with the simulation of the dual-RGB imaging setup utilizing the data characterized in the first part. The LEDs were defined in different bins in order to reduce the combinations of light source pairs. This reduced the number of different combinations from approximately a million to only 630. All of the 630 combinations of light sources combinations were traversed to find the optimal light source pair giving the least mean ΔE_{00} error on the training colourchecker data. The mean error provided by the optimal light source was 0.31

for training and 0.66 for testing. The finalized light source pair was compared against the simulation of a three-channel light source under CIE D65. This showed that for three-channel imaging the training error was 1.25 and testing error was 1.53. This showed a significant improvement of dual-RGB over three-channel imaging.

The third part of the thesis was related to the exploration of how changing the intensities of the LEDs defined in a dual-RGB imaging setup could further improve the colour accuracy of the imaging system. For the purpose of optimization, Particle Swarm Optimization (PSO) was adapted and implemented for the LEDs intensity optimization. The optimizations were done for all the possible 630 combinations of light sources that the dual-RGB was optimized on. PSO explored all the different variations of LEDs intensities giving the least mean ΔE_{00} error on the training data. The final pair of light sources and their intensities were different than what was optimized in dual-RGB optimizations. This new PSO optimized light sources gave a training error of 0.17 and a testing error of 0.28 which was an improvement over dual-RGB light source. This new PSO optimized light source and the method was subsequently referred as dual-RGB-PSO. In this same part, for the purpose of giving more confidence on the optimality of the dual-RGB-PSO solution, genetic algorithm was used to optimize the same LEDs intensity problem. Genetic algorithm converged on the same pair of LEDs as dual-RGB-PSO, as well as approximated the same intensities too. This gave more probability that the dual-RGB-PSO light source was not a local minimum or a side-effect of algorithm specific details.

The fourth part of the thesis concerned with validating the simulated imaging model and conducting the actual captures for dual-RGB and dual-RGB-PSO optimized light sources. The camera simulation was validated through correlation plots of simulated and real captures pixel values. Other than the two outliers found in training data, the rest of the points showed very good linearity. The real captures of dual-RGB gave a training error of 1.06 and testing error of 1.61. The dual-RGB-PSO showed a training error of 1.11 and testing error of 1.87. This was compared with the three-channel captures which gave training error of 1.89 and testing error of 2.52. This data further proved that having more channels in the dual-RGB and dual-RGB-PSO imaging provides a better colour accurate image. One apparent dissimilarity between simulation and real captures were in the comparison of the dual-RGB and dual-RGB-PSO errors. In simulations, dual-RGB-PSO performed better than dual-RGB for both training and testing, reducing the errors approximately by half of dual-RGB errors. One possible explanation for this apparent misalignment is that the scale at which the errors are being optimized in simulation for dual-RGB and dual-RGB-PSO is much less than the perceivability of human visual system. One possible solution for this misalignment can be to use training and testing data which is more difficult to optimize even in simulations.

The fifth part of the thesis was related with high dynamic range (HDR) imaging in the context of dual-RGB and dual-RGB-PSO, subsequently called HDR-Dual-RGB and HDR-Dual-RGB-PSO. Due to limited dynamic range of the colourchecker films being used, a neutral density (ND) filter was used to cover half the number of patches of both training and testing colourcheckers. Several low dynamic range (LDR) images were captured where within both dual-RGB and dual-RGB-PSO, a pair of LDR images were captured for each light source. After dark noise correction and flatfielding, the LDR images were combined with a weighting function to form an HDR image. The HDR-Dual-RGB mean training error was 2.30 and testing error was of 2.87. For HDR-Dual-RGB-PSO the training error came out to be 2.35 and testing error to 3.09. HDR captures also showed similar colour accuracy between the HDR-Dual-RGB and HDR-Dual-RGB-PSO. This also gives more ground to the previous explanation that the errors that the simulations for dual-RGB and dual-RGB-PSO were operating on might be of insignificant magnitude.

The sixth part of the thesis related to the application of the existing dual-RGB and dual-RGB-PSO workflow for reflective objects. The existing LED light source was originally built only for photographic films. The setup for this light source was adapted to be operable for reflective objects with minimal changes and easy to move back and forth between film and reflective mode. Similar procedure was employed for reflective objects as was applied as in the second, third, and fourth parts previously discussed. When the dual-RGB simulation was employed on reflective objects, it optimized to exactly the same light source pair. Applying the dual-RGB-PSO simulation, it ended up with approximately similar imaging channels. Real captures of the reflective colourcheckers were taken with the dual-RGB and dual-RGB-PSO light sources optimized for reflective objects as well as for the light sources of dual-RGB-PSO optimized for the films in the third part of the thesis. The mean training errors were 2.08, 1.84, and 2.18 and the mean testing errors were 2.25, 1.72, and 2.14 for dual-RGB-PSO, dual-RGB, and dual-RGB-PSO (optimized for films). The light sources optimized for the films can be said to be performing comparably, depending on applications, with the reflective optimized light sources on reflective objects. This hints at the interoperability of the film and reflective mode of the dual-RGB and dual-RGB-PSO workflows.

The results of the simulations and real captures done in the thesis work, hints at a very fruitful research direction to explore not just dual-RGB itself but also its extensions toward LEDs intensities optimizations and embedding HDR imaging techniques. The interoperability of the films and reflective mode of the imaging setup, and the similar results of light sources optimized for films when used on reflective objects, point towards a combined multispectral imaging setup operable both on films and reflective objects.

Limitations

There are several improvements and directions that can be explored in future research, some of which are mentioned list below:

- The testing of the implemented methods such as dual-RGB and dual-RGB-PSO can be performed on real scene photographic films.
- Since the accuracy of a dual-RGB workflow is tied with its training data, it is important to consider new set of training data, depending on the application context. One example of it is Artist Paint Target (APT) (Berns, 1976) proposed for the application in cultural heritage. This would allow for a better colour accuracy in the application.
- For dual-RGB-PSO, the parameters utilized in this thesis work had been taken from literature. A hyperparameter tuning can be performed on PSO for getting a set of parameters which work well.
- The interoperability of the light sources optimized for photographic films and reflective objects can be checked more rigorously with more variety of colourcheckers involved in both films and reflective objects.
- The current work captured the reflective objects with the light sources optimized on films. It should be checked whether this observation is also valid for the other way around i.e. whether film captures produce good colour accuracy when the light source optimized for reflective objects are used.
- The advantage of the HDR capturing workflow can be more highlighted if it is compared with an LDR capture of the same HDR scene.

Context of thesis work

The thesis work and directions were impacted due to the delay in visa processing, because of which the author was not able to conduct his research at Munsell Colour Science Laboratory (MCSL) at RIT, USA. The original thesis direction was to improve the dual-RGB using the methods of high dynamic range (HDR) imaging for reflective objects, for which the equipments were in place at MCSL. As an alternative approach of working at the Colourlab at NTNU, the LED imaging system developed for photographic films were utilized to implement the dual-RGB for photographic films from scratch and improving it using the proposed intensities optimizations. Still the direction of HDR workflow and reflective objects were explored to a degree.

A Exposure time linearity characterization

The CRF that we showed in earlier chapter in Fig.4.2 is not the first and only CRF that we extracted for Nikon D610. The first CRF is shown in Fig.A.1. Notice the unusual bumps in the supposedly linear region of the camera. We repeated the experiment for the CRF measurement in a separate light booth to make sure it was not a problem related to unstable light booth, we still found the bumps at the exact same exposure times. This motivated us to investigate this unexpected behavior of the camera rigorously. Our initial impression about these unusual bumps are that these occur because of the defect in mechancial shutter in DSLR cameras. This hypothesis can get further grounds, if we compare mechanical shutter and electronic shutter cameras linearities. In this chapter, we will explain the method that we used to investigate this issue and also compare the results across a wide range of different cameras.

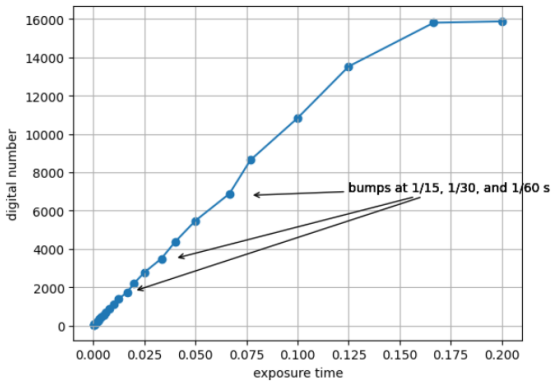


Figure A.1: *The first CRF measured for Nikon D610. The exposure time is in seconds and the digital numbers are in 14-bits*

A.1 Exposure time linearity characterization

In order to investigate the unusual bumps in CRF, we define a type of camera characterization which we propose to call as exposure time linearity characterization.

Exposure time linearity characterization: *Measurements of linear CRF samples across all the possible exposure times for the camera.*

In this characterization we measure samples of CRF of cameras across all the possible exposure time that the camera have. This will allow us to detect unusal deviations from linearity at particular exposure times for the camera. By linear CRF samples we mean points of the CRF where it supposedly should be in linear region.

A.2 Steps to characterize exposure time linearity

A.2.1 Illumination setup

Note that even in regular laboratory conditions, it is not generally easy to have a robust single setup to generate a wide range of illumination conditions to get the CRF points for a camera so that the points lie in the linear region of the sensor. Because of this we loosen up the definition so that the measurements can be taken across different setups for illumination just with the condition that there should be at least three points measured for CRF in the linear region for each of the different illumination. The condition for three points comes from the requirement that we need at least three points in order to detect deviation or bumps from the ideal linear behaviour of the sensor. The illumination setups that we used allowed us to have CRFs with overlapping exposure times which implied that we could observe any unusual deviations across different set of captures. This gave us more confidence that these bumps are not related to sensor linearity but to mechanical shutter.

A.2.2 Dealing with dark noise

For considering signals in which dark noise does not play a significant part we can apply different approaches that we list below:

Dark noise correction: We can apply dark noise correction for each of the captures that we take. For this we would need to take several dark captures with lens cap on for each of the exposure time that we intend to work. We then

average the dark captures for each of the exposure times and subtract it from the linearity captures at corresponding exposure time. This is the most accurate but time-consuming approach.

Maximum exposure dark noise: Another approach would be to take average dark noise capture for the longest exposure time (similar to previous step) only and then take the mean value across all the pixels for this average dark noise capture. Once we have this value, we consider signals from the linearity captures which are well above this dark noise mean value.

Estimating maximum exposure mean dark noise: Instead of taking the dark noise capture at the maximum exposure time, we can also take some estimated value for mean dark noise.

A.2.3 Finding a linear region of sensor

In order to find deviations from linearity in CRF we need to decide upon a linear range for the camera. Otherwise we would not be able to attribute a deviation to either unusual 'bumps' at particular exposure times or to the sensor non-linear behavior itself. There can be two approaches to do this which we list below:

Doing it after the measurements: In this approach we can do the measurements for the CRFs at all the possible exposure times such that we start from the very lowest signal from the sensor (above dark noise) and go to complete saturation. Once we have this data, we can explore the different CRFs that we have measured and select the ones that does not show the bumps in linear region. It might be the case that we end up with no CRFs without bumps but in our experience with the different cameras that we have worked on there was always such a CRF. Note that in this approach we have to take more measurements since in each of the CRFs we need to go from minimum signal to complete saturation.

Estimating the linear region: In this approach we can estimate the linear region to be two-thirds of the total range of digital numbers for the captures. For a camera captures of 14-bits, this would be 10,922. In our experiments with different cameras, this range has turned out to be very good approximation of linear region and also a bit on the safe side i.e. this gives a smaller range than the actual bigger linear range for the cameras. For example, for Nikon D610 through our measurements by the first method, we estimated the linear range to be 11,000 which is higher than the estimated by two-thirds range i.e. 10,922. Note that in this approach we do not need to take the measurments till full saturation of the camera because we have already decided upon the linear part of the sensor. This makes this method faster than the previous one.

A.2.4 Method for describing a deviation:

Once we have CRF samples in linear region across different exposure time ranges, we need to define some way that we mark a CRF point in the linear region as a deviation. There can be two methods for doing this which we list below:

Changepoint detection: The deviations can be calculated through some automated manner. Changepoint detection algorithms like proposed by Bai (1997) and Fryzlewicz (2014) can be used for this.

Manual: We can observe the deviation from linearity by observing the CRFs too. This detection may become subjective.

A.3 Exposure time linearity characterization

In order to study the deviations from linearity and to check our hypothesis that this is caused due to mechanical shutter, we characterize the exposure time linearity of both mechanical shutter and electronic shutter cameras. We have characterized the following cameras:

- Nikon D610 (Mechanical shutter)
- Nikon D850 (Mechanical shutter)
- Nikon D200 (Mechanical shutter)
- Silios CMS-S multispectral camera (Electronic shutter)
- Spectrocam multispectral camera (Electronic shutter)

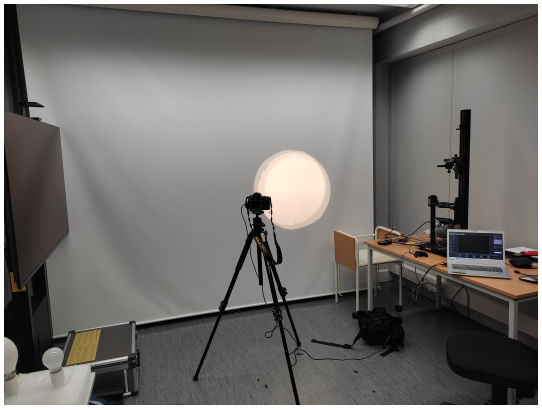
In this section we discuss how we do the exposure time linearity characterization of Nikon D610. We expand upon the setup for Nikon D610, and then provide the results for the other cameras directly in the later section.

A.3.1 Setup for characterization

We show the setup that we used in Fig.A.2. We have set up the camera on a tripod pointing towards a white projector screen. In order to create different levels of illumination we use three halogen lamps, the tunable pair of room lights, and an LCD screen with white background.

With these sources of lights in hand, we create six different illumination levels as described in Table.A.1. In addition to these six illumination conditions we also

(a) *Camera pointing at the white projector screen*



(b) *Projector screen illuminated with three halogen lamps*

(c) *The three halogen lamps*

Figure A.2: *Setup for exposure time linearity characterization for Nikon D610.*

have the data for the light booth for this camera that we had collected earlier. So in total we have seven different illumination conditions for this camera.

For dark noise we go with the approach discussed in *Estimating maximum exposure mean dark noise* and take the values of linearity captures which are above 500 digital counts for 14-bits data (0 to 16383 range). For finding linear region we do what we explained in *Doing it after the measurements* and measure the linear region till 11,000 digital counts, and for changepoint detection we do it manually.

A.3.2 Measurements

In Fig.A.3, we show the range of exposure times which was in the linear range of the camera for each of the illumination setup we mentioned above. The triplets of each color represent the blue, green, and red channel response of the camera. The

Table A.1: *Table detailing the different illumination setup used for exposure time linearity characterization for Nikon D610.*

Illumination setup	Light
1	LCD screen with white background
2	One room light at minimum
3	One room light at medium
4	One room light at full and other at minimum
5	Two room lights at full
6	Two room lights at full and 3 halogen lamps

different colors represent the different illumination setup. So for example, at the bottom right we have the blue channel capture when only LCD screen was on in the room. The blue channel had the measurements from 2.5s to 30s exposure time in the linear range of the camera. The green channel for the same setup was linear in the exposure time from 1.6s to 20s and so on. We show the same thing for all the different illumination setups. The highlighted yellow exposure times are the ones where we found the bumps consistently across the different setups.

For studying the curves, we select the individual channels in different setups to cover the whole exposure time range of Nikon D610 and also to have some overlaps of detected bumps in different illumination setups to be more certain about their existence. We show the selected channels in bold outlines of cells. So for example we should observe bump at 1/60 exposure time in setup 4, 5, and 7 (L, O, and V rows).

In Fig.A.4, we show all the CRFs measured across the different illumination setups across the possible exposure times. We see in the plots, that there are particular exposure times across which the bumps are detected across the different setups.

A.3.3 Exposure time linearity characterization: Nikon D850

We wanted to explore more mechanical shutter cameras available in our laboratory. The next camera that we tested after Nikon D610 was Nikon D850. We show the exposure times used and CRFs in Fig.A.5 and Fig.A.6.

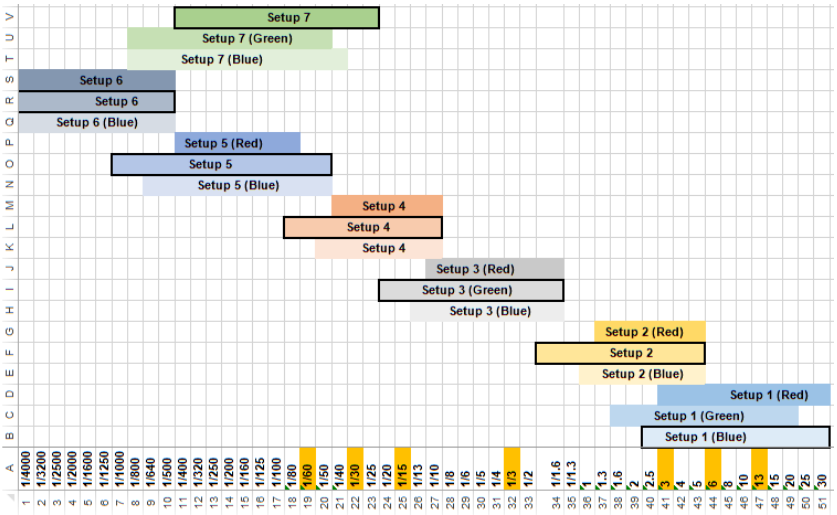


Figure A.3: *Showing which channels were in linear range across the different integration times for Nikon D610*

A.3.4 Exposure time linearity characterization: Nikon D200

The next in line Nikon D200 which is quite old camera. We show the exposure times used and CRFs in Fig.A.7 and Fig.A.8.

A.4 Exposure time linearity characterization: Silios CMS-S

For testing the electronic shutters, we explored multispectral cameras. The first one was Silios CMS-S which has 9-channels Bayer-patterned filters. Since in this camera the exposure time could be set across a wide range of exposure times with a precision of milliseconds. We decided to test a two different randomly selected regions for exposure time. We did not find any unusual bumps in these CRFs. We show the CRFs in Fig.A.8.

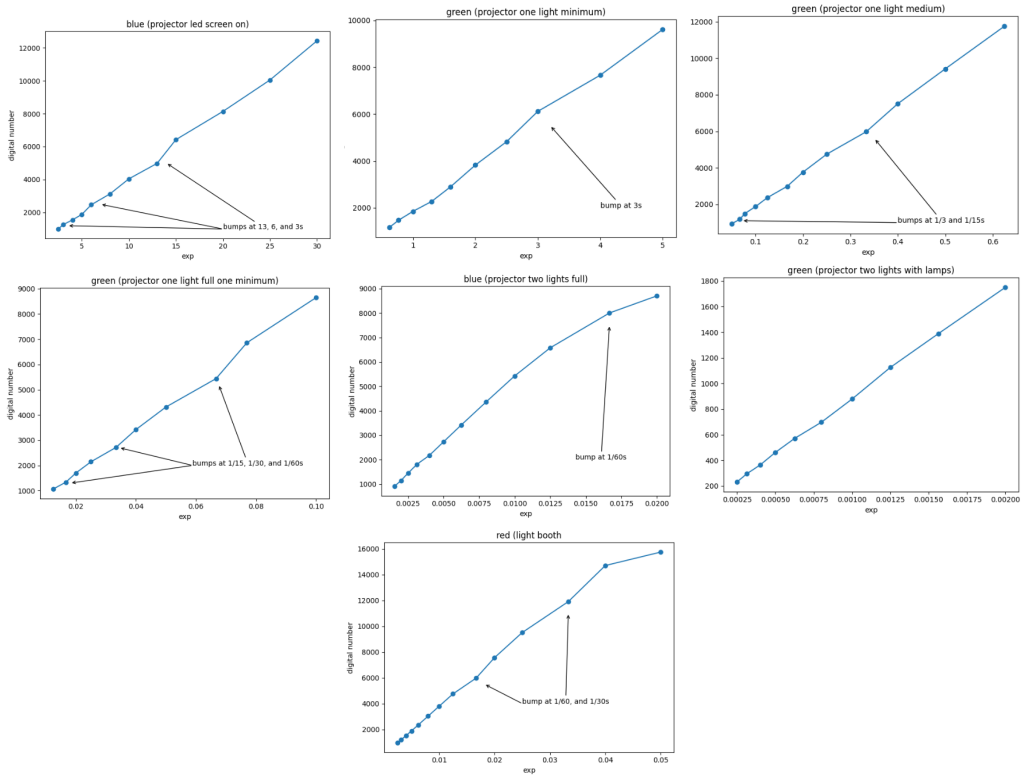


Figure A.4: *Linearity curves showing bumps across different integration times*

A.5 Exposure time linearity characterization: Spectrocam

The next in line for electronic shutter was Spectrocam multispectral camera. The CRFs are shown in Fig.A.10. Similar to previous electronic shutter, we do not find any unusual bumps in these CRFs either. This gives more grounding to our hypothesis that the unusual bumps are because of the mechanical shutter.

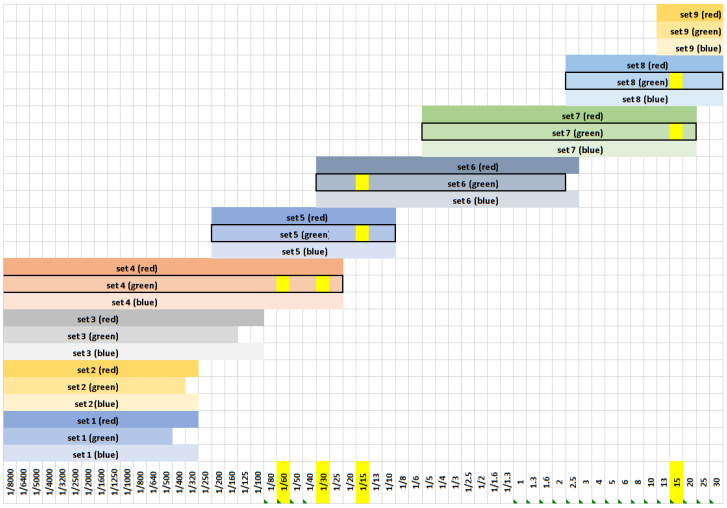


Figure A.5: *Showing which channels were in linear range across the different integration times for Nikon D850*

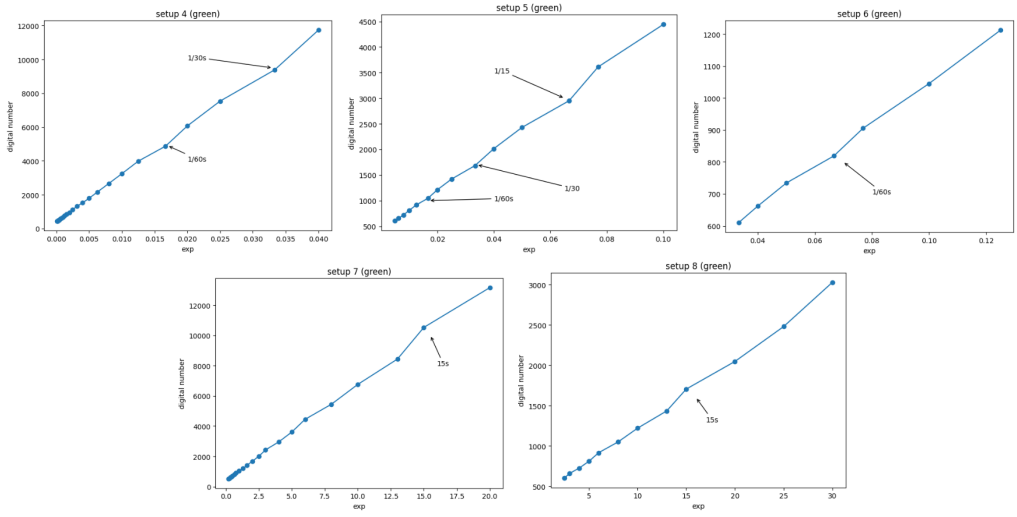


Figure A.6: *Linearity curves showing bumps across different integration times for Nikon D850*

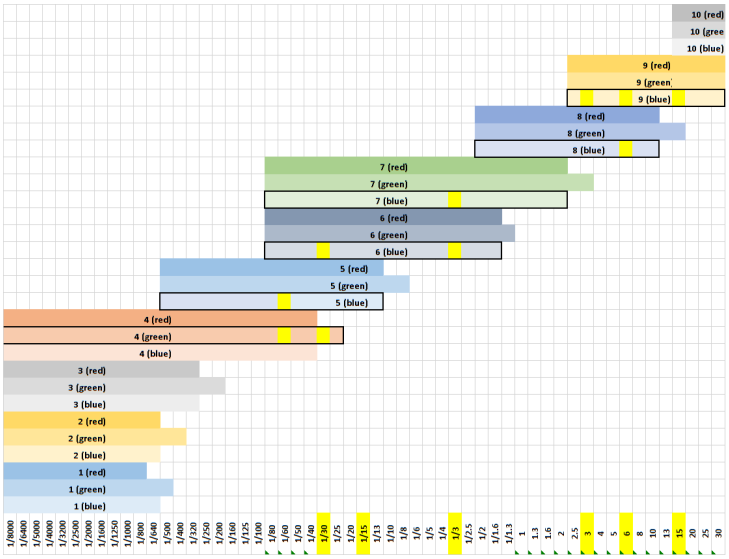


Figure A.7: *Showing which channels were in linear range across the different integration times for Nikon D200*

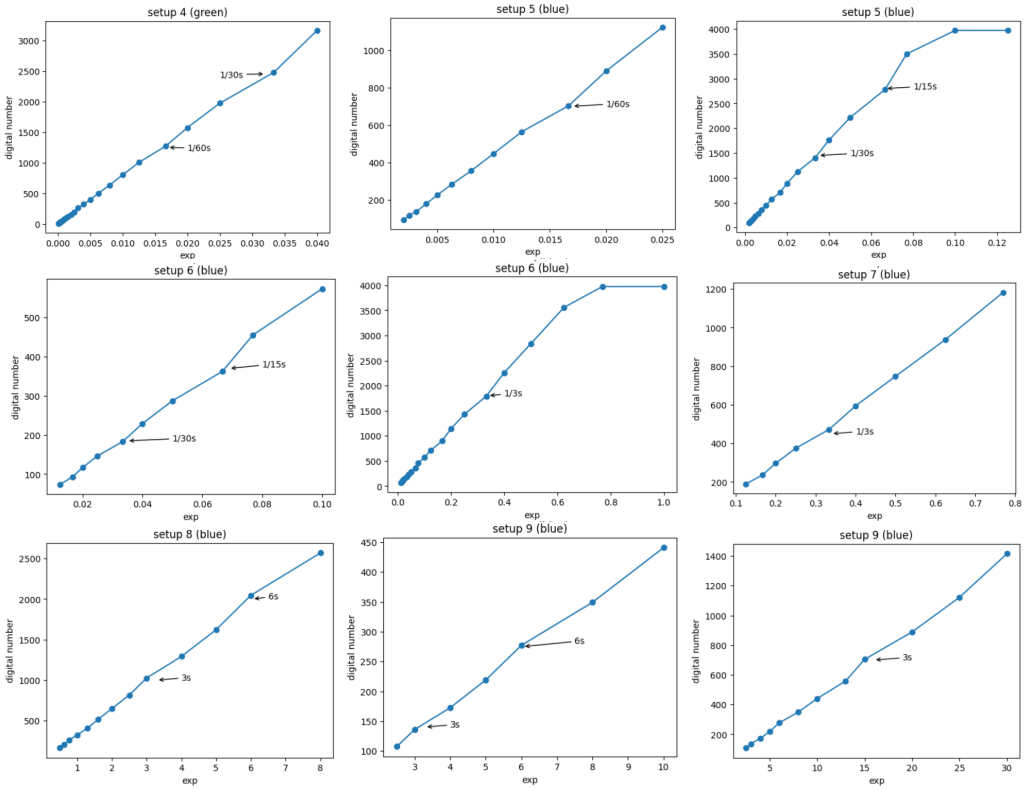
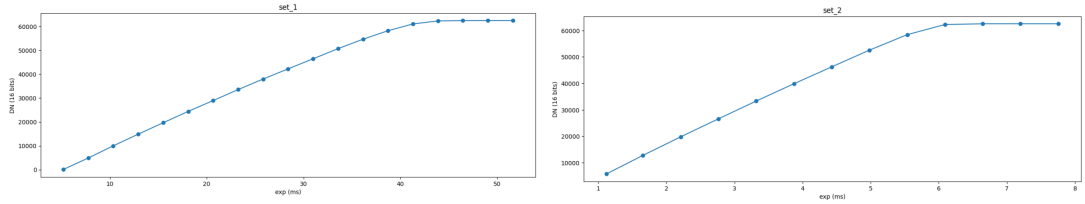
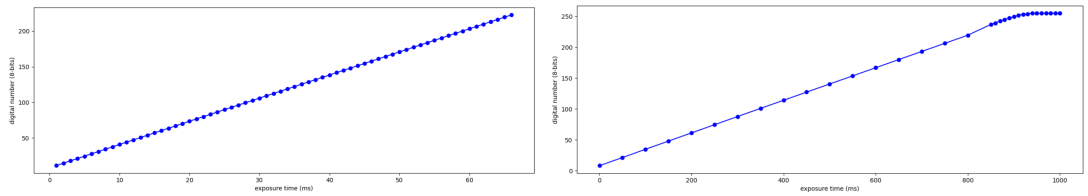


Figure A.8: *Linearity curves showing bumps across different integration times for Nikon D200*

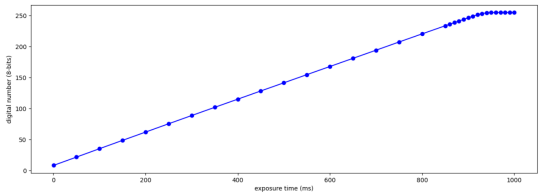


(a) *Set 1* (b) *Set 2*

Figure A.9: *Linearity curves for Silios CMS-S*



(a) *Set 1* (b) *Set 2*



(c) *Set 3*

Figure A.10: *Linearity curves for Spectrocam*

B | Colorimetry

The idea that the range of colours perceivable by humans can be reproduced by the combination of three colours has been evolving throughout the course of several centuries. The earliest documented example of the application can be traced back to the year 1722, when Jakob Christoffel LeBlon employed it in trichromatic printing Weale (1957)Wall (1925)Birren (1981). The subsequent contributions of Thomas Young and James Clerk Maxwell have played a significant role in advancing the understanding of the trichromatic characteristics of the human visual system (HVS) and their occurrence at the retinal level. Demonstrating the trichromatic nature of the HVS and how it operates at the retinal level presents a nontrivial challenge. Dartnall and Lythgoe (1965) observed absorptions in eyes of animals having similar color vision as humans. Researchers also observed that the pupil's reflected light behaved as if red, green, and blue pigments were present and absorbing light in their respective regions (Rushton, 1958)(Rushton, 1957)(Weale, 1959)(Ripps and Weale, 1963)(PK and WALD, 1964)(Brown and Wald, 1964). Postmortem microscopic examinations of the retina have additionally demonstrated the presence of three distinct pigments within specialised structures known as cones (Marks et al., 1964)(Brown and Wald, 1964)(Bowmaker and Dartnall, 1980).

It is very important to highlight the difference between the subjective experience of colour and the external stimuli that trigger such an experience. Color is a perception or subjective experience which makes it impossible to quantify through mechanical devices. However, the external stimuli received by the HVS that result in the formation of such perceptions can be objectively measured and quantified using engineering instruments. The foundation of colorimetry is rooted in the establishment of a relationship between the quantifiable stimulus received as input and the resulting perception it produces.

Colorimetry can be broadly classified into two main categories: basic colorimetry and advanced colorimetry. Basic colorimetry determines whether or not two colour stimuli viewed under identical external conditions are equal (peceptually same). The extent of their differences cannot be determined. Advanced colorimetry goes a step further by attempting to determine the magnitude and direction of difference. Advanced colorimetry also aims to describe physical correlates of HVS perception

such as lightness, chroma, and hue.

Colour stimulus can be produced through the combination of three light sources. Using a display as an example, three sources of red, green, and blue light are positioned so close together that HVS is unable to resolve the gaps between them and interprets it as a single source of light. This lack of resolution is the result of the limited contrast sensitivity of HVS, as defined by the contrast sensitivity function. This method of generating colour stimuli is known as additive mixture. An alternative method for producing colour stimuli is known as subtractive mixing. In this process, three distinct colours or dyes, each with the ability to absorb specific wavelengths of light, are combined in specific proportions. When light is incident upon this mixture, it results in the perception of a particular colour.

The foundation of colorimetry is rooted in the empirical principles established by H. G. Grassmann, which are commonly expressed in contemporary literature as follows (verbatim reproduction): (Hunt, 1998):

1. To specify a color match, three independent variables are necessary and sufficient

2. For an additive mixture of color stimuli, only their tristimulus values are relevant, not their spectral compositions.

3. In additive mixtures of color stimuli, if one or more components of the mixture are gradually changed, the resulting tristimulus values also change gradually.

These laws are valid when the parameters under which the matching is conducted remain consistent, including factors such as stimulus size, retinal illumination, observer's state of adaptation, and the observer himself. While these principles are applicable to a broad range of colorimetric applications and have thus formed the foundation for CIE colorimetry standards, it is known that these laws may not hold under certain circumstances. When exposed to high levels of luminance, the validity of colour matching between monochromatic yellow and a combination of red and green is disturbed (Wright, 1934)(Wright, 1936). Additivity in blue-green matches is also found to break down under a variety of stimuli luminance levels (greater than 15,000 td [1]), however only to a minimal percentage of tristimulus values (Blottiau, 1947)(Trezona, 1953)(Trezona, 1954). Colour difference equations such as CIE94 (CIE, 2023) raises questions about the symmetry property (if A=B, then B=A) of colour match because the errors vary depending on which colour is used as reference and which as test. Furthermore, when considering a colour match that comes within

[1] td stands for Troland. This represents a quantity which is proportional to the illumination on retinal surface. It is calculated as the product of incoming luminance to eye with the area of the limiting pupil. The luminance is considered in candelas per square metre and the pupil area is taken in square millimetres (Schanda, 2007).

the limits of a just-noticeable difference (JND), it is possible for the transitivity property (if A=B and B=C, then A=C) to be disturbed. This implies that if the difference between stimuli A and B falls within the just noticeable difference (JND), and the difference between stimuli B and C also falls within the JND, it does not necessarily ensure that the difference between stimuli A and C will also fall under the JND. In other words, the cumulative effect of two sub-threshold differences may exceed the threshold difference. Thornton (1992) observed a mismatch of colors when different primaries were used to predict matches. This transformation of primaries, which is the backbone of color management, is suggested to breakdown because of the failing of Grassmann's Laws.

To define a colour match, an experimental setup for trichromatic colour matching can be implemented, as depicted in Fig.B.1. An observer sees a bipartite field in which half of the field is illuminated by a monochromatic light while the other half is illuminated by a combination of three lights designated R, G, and B after Red, Green, and Blue. The objective of this experiment is to sequentially examine each monochromatic light within the visible spectrum and determine the respective luminances of the red (R), green (G), and blue (B) lights necessary to achieve a perceptual match with the monochromatic light in question.

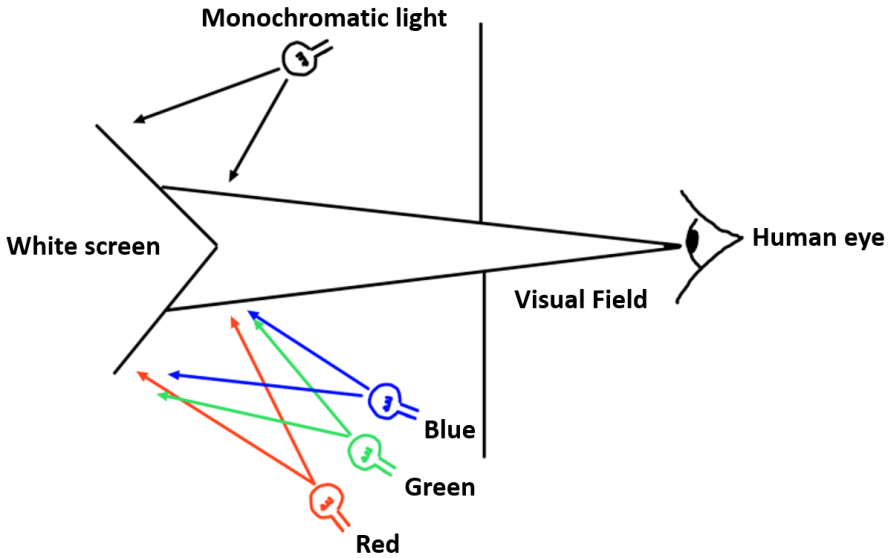


Figure B.1: *Setup for color matching experiment. An observer sees a bipartite field on a white screen. Half of the field is illuminated by a monochromatic light while the other half by a combination of red (R), green (G), and blue (B) primary lights. The objective of the experiment is to match the monochromatic light with the combination of primaries. This has to be done for each monochromatic light in the visible range.*

The colour matching phenomenon, which is determined by the interaction of three light sources, is defined by Grassman's laws and is made feasible by the phenomenon of metamerism. Metamerism refers to the phenomenon wherein the HVS is unable to discern any distinction between two stimuli that generate identical responses in the retinal cones, regardless of the spectral compositions of those stimuli. These pairs of stimuli are referred to as metamers. In the colour matching experiment, some monochromatic colours can be matched in hue and brightness but not in colorfulness, particularly in the blue-green colour region. This obesrvation may appear to contradict Grassman's laws, however it does not. The perception of colorfulness can be achieved by introducing red (R), green (G), and blue (B) colours to monochromatic light, which can be considered as the addition of a 'negative' color.

If the matching primaries R, G, and B are monochromatic lights at 700, 546.1, and 435.8nm, then the quantities of these primaries required to match every visible monochromatic light are defined as shown in Fig.B.2. These three plots are called as color matching functions (CMFs) and each of these plots are symbolized as $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$. It is important to establish the specific quantities of the primaries, R, G, and B, that the observer will be modifying before starting the colour matching experiment. This can be specified in units of power, but is typically defined in arbitrary units so that equal quantities of R, G, and B match the luminance of a predetermined white. In this particular case, the colour white is chosen as the equi-energy source S_e, which has equal power distribution across all wavelengths. The luminance values of 1cd/m^2 of R, 4.5907cd/m^2 of G, and 0.0601cd/m^2 of B matches with S_e. Using a different white would result in the same CMF shapes but with three different ffactors multiplied by the three functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$.

A monochromatic color-match is written as:

$$1.0[\lambda] \equiv \bar{r}(\lambda)[R] + \bar{g}(\lambda)[G] + \bar{b}(\lambda)[B] \tag{B.1}$$

where the equivalent sign means a color-match, the quantities in square brackets represent the pre-selected primaries units, the coefficients (e.g. $\bar{r}(\lambda)$) represent the amount of the stimuli used in the units defined, and the addition sign represents the additive mixture of colors.

According to the Grassmann's Laws, this equivalency holds for proportional quantities too such that for a constant k we have:

$$k[\lambda] \equiv k\bar{r}(\lambda)[R] + k\bar{g}(\lambda)[G] + k\bar{b}(\lambda)[B] \tag{B.2}$$

Grassmann's Laws also dictate the that if we have two stimuli matched as:

$$k_1[\lambda] \equiv k_1\bar{r_1}(\lambda)[R] + k_1\bar{g_1}(\lambda)[G] + k_1\bar{b_1}(\lambda)[B]$$
$$k_2[\lambda] \equiv k_2\bar{r_2}(\lambda)[R] + k_2\bar{g_2}(\lambda)[G] + k_2\bar{b_2}(\lambda)[B]$$

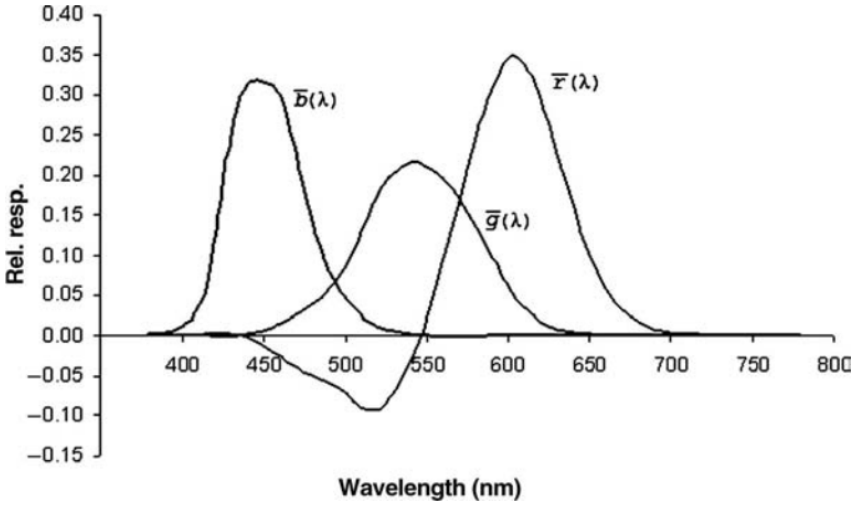


Figure B.2: *Color matching functions (CMFs) $\bar{r}$, $\bar{g}$, and $\bar{b}$ (Schanda, 2007). These curves represent the relative amounts of R, G, and B primaries required for matching a monochromatic light. Read the function values at particular wavelength to get the corresponding amounts of primaries required to match a monochromatic light of that wavelength.*

then the additive property holds:

$$k_1[\lambda]+k_2[\lambda] \equiv (k_1\bar{r_1}+k_2\bar{r_2})(\lambda)[R]+(k_1\bar{g_1}+k_2\bar{g_2})(\lambda)[G]+(k_1\bar{b_1}+k_2\bar{b_2})(\lambda)[B] \quad \text{(B.3)}$$

This implies that if we have any arbitrary stimulus Q for which we know the spectral power distribution $S(\lambda)$, we can calculate the coefficients of primary stimuli which will match the color in appearance (under identical external settings) as

$$[Q] \equiv R_q[R] + G_q[G] + B_q[B] \quad \text{(B.4)}$$

where

$$\begin{aligned} R_c &= S_1\bar{r_1} + S_2\bar{r_2} + ... \\ G_c &= S_1\bar{g_1} + S_2\bar{g_2} + ... \\ B_c &= S_1\bar{b_1} + S_2\bar{b_2} + ... \end{aligned}$$

The values R_c, G_c, and B_c are called as tristimulus values of Q. Note that the luminance of Q can be calculated as

$$L = 1.0000R_c + 4.5907G_c + 0.0601B_c \quad \text{(B.5)}$$

which comes from the fact that we had matched the luminances of the primaries in the same 1.0000 : 4.5907 : 0.0601 proportions beforehand. Most of the time the units used for R, G, and B are in relative luminances making L a relative luminance too.

If different primaries and matching white are utilised, it is possible to linearly transform the CMFs using the proportionality and additive properties stated previously. We only need to know how the new primaries match with the old primaries. We call this as a transformation between color spaces. During the process of standardisation, the presence of negative lobes in the colour matching functions (CMFs) presented challenges in efficiently computing the tristimulus values. As a solution, CIE made the decision to convert the primary colours R, G, and B into a set of imaginary [2] primaries denoted as X, Y, and Z. This transformation was implemented to ensure that the resulting colour matching functions would not exhibit any negative quantities. The transformation is defined as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.768892 & 1.751748 & 1.130160 \\ 1.000000 & 4.590700 & 0.060100 \\ 0.000000 & 0.056508 & 5.594292 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{B.6}$$

Other requirements for this transformation that one of the tristimulus values should correspond to luminance (note that Y is calculated same as luminance), tristimulus values for equi-energy stimulus values should be equal, and the volume of the tetrahedron set by the new primaries should be as small as possible (Schanda, 2007).

In order to see the importance of CMFs in color reproduction we take an example of a color stimulus $[C]$ defined as:

$$[C] \equiv [500\text{nm}] + [600\text{nm}] \tag{B.7}$$

From the CMF, we can see that this will be matched to

$$[C] \equiv (-0.1 + 0.8)[R] + (0.2 + 0.2)[G] + (0.2 + 0.0)[B] \tag{B.8}$$

$$\equiv 0.7[R] + 0.4[G] + 0.2[B] \tag{B.9}$$

Note the importance of having negative tristumulus value. Although the stimulus [500nm] is not realizable because we cannot generate a negative luminance while reproducing the color. But when this negative stimulus combines with another stimulus with positive and larger magnitude tristimulus value, this gives rise to a color with net positive tristimulus value.

[2] Note that these imaginary primaries have nothing to do with Complex number.

In Fig.B.3a we assume a uniform patch emitting a combination of 500 and 600nm light. This light falls similarly on the three sensors with filters. Let's assume the combined sensitivity of the sensor and filters are exactly as the CMF that we generated for $[R]$, $[G]$, and $[B]$. The readings generated in the sensors will be proportional to $[C]$:

$$[C_1] \equiv k_1(0.7[R] + 0.4[G] + 0.2[B]) \tag{B.10}$$

The constant k_1 depends upon the exposure of the sensors. This means we have successfully automated the color-matching experiment upto a scaling factor of luminance.

In Fig.B.3b we assume three LEDs with the same tristimulus values as the $[R]$, $[G]$, and $[B]$ primaries used for CMF measurments and also have the same relative luminances. Now if we provide same voltages to these LEDs as the coefficients in Eq.B.10 then we generate a combined output at the screen which is proportional to $[C]$:

$$[C_2] \equiv k_2(0.7[R] + 0.4[G] + 0.2[B]) \tag{B.11}$$

This means that we have successfully reproduced the color which was captured by the sensors up to a scaling factor of luminance. This is a very simplistic explanation for color reproduction workflow with many unrealizable parts, one of those parts being the negative combined sensitivity of the sensor. For more detailed view on the color reproduction subject refer to Hunt (2005).

In many cases we are not interested in the absolute values of tristimulus values but only with the proportions, in which case the representation of the color can be simplified to just two numbers. We calculate the relative proportions as:

$$\begin{aligned} x &= \frac{X}{X+Y+Z} \\ y &= \frac{Y}{X+Y+Z} \\ z &= \frac{Z}{X+Y+Z} \\ &= 1 - x - y \end{aligned}$$

since z can be defined in terms of x and y, so we only need to know x and y. These are called as chromaticity coordinates. The two-dimensional plot showing the x and y dimensions is shown in Fig.B.4.

Once we have formulated a way to describe whether two stimuli are matching perceptually or not, the next immediate question is how much different are they in

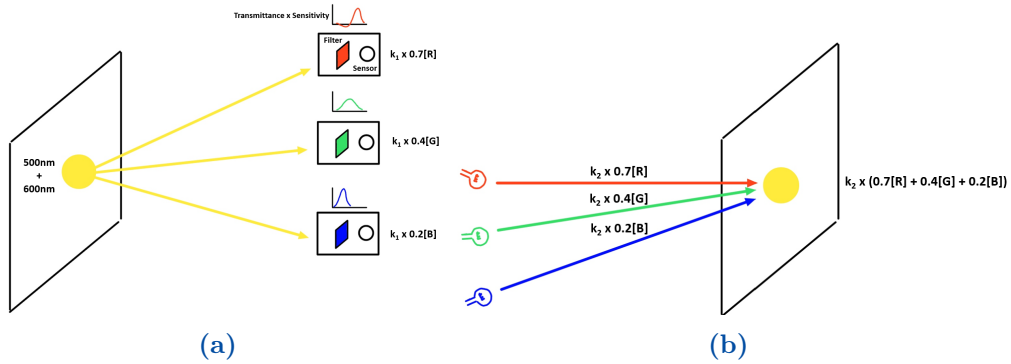


Figure B.3: *A very simplistic workflow for color reproduction. (a) represents a light source emitting a combination of 500nm and 600nm light (yellow). This light passes through three transmissive filters whose transmittances lie in the red, green, and blue regions of spectrum. Then the light signal is recorded by a digital sensor. The combined sensitivity of each of the filter-sensor pair is equal to the three CMFs $\bar{r}$, $\bar{g}$, and $\bar{b}$. The digital numbers recorded by each of the sensor is proportional to the cumulative values of CMFs at 500nm and 600nm. (b) represents three LEDs whose individual colors are matched by the primaries used to derive CMFs. The voltages provided to the LEDs are in the same proportion as the coefficients of the cumulative values of CMFs at 500nm and 600nm. The light from each LED is additively mixed to produce the yellow light. This results in a reproduction of color which only differs from the original yellow color by a single luminance factor.*

the case of a mismatch. The simplest way for doing that is the calculation of a Euclidean distance in the chromaticity diagram as:

$$\Delta E = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{B.12}$$

But it turns out that calculating differences in this way for chromaticities is not suitable for difference measurements because the magnitude of perceived difference is different across the region of the space. This is famously shown by MacAdam (1942) as in Fig.B.5 where the ellipses represent 10 times the just noticeable differences, which means observers do not find any difference in appearance of stimuli if the chromaticities remain within the ellipse. Notice how the regions are very small in blue region and very large in green.

To improve on the situation, researchers have worked on two different approaches. The first approach is to transform the XYZ space to some other space which is more uniform perceptually. The second approach is to use more sophisticated color difference equations which takes into consideration the non-uniformities of the color spaces in which the difference is being taken.

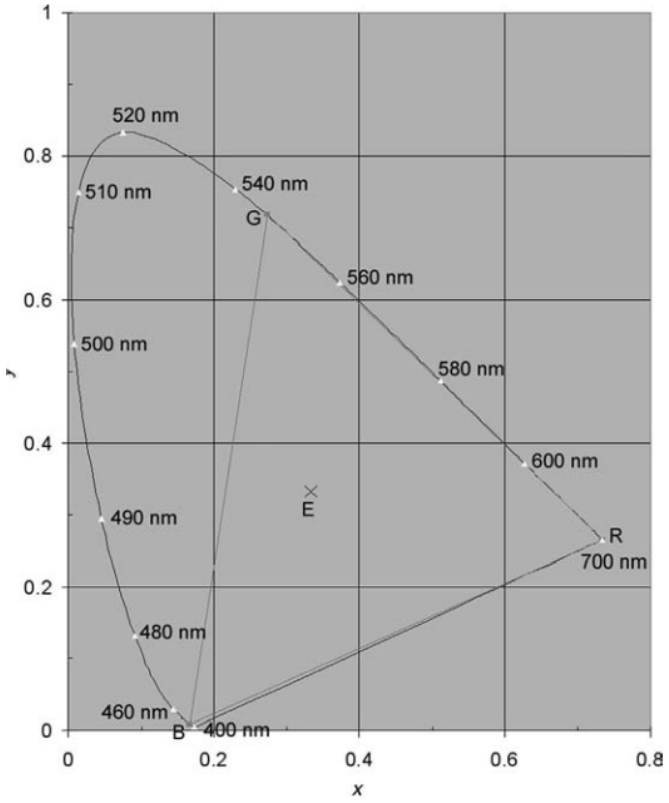


Figure B.4: *Chromaticity diagram showing the chromaticities of the visible range of colors along with the chromaticities of the three selected primaries used for the colour matching experiment (Schanda, 2007).*

Considering the first approach, one of the mostly used color space is CIELAB which is defined as:

$$L^* = 116 f\left(\frac{Y}{Y_\mathrm{n}}\right) - 16$$
$$a^* = 500\left[f\left(\frac{X}{X_\mathrm{n}}\right) - f\left(\frac{Y}{Y_\mathrm{n}}\right)\right]$$
$$b^* = 200\left[f\left(\frac{Y}{Y_\mathrm{n}}\right) - f\left(\frac{Z}{Z_\mathrm{n}}\right)\right]$$

where

$$f(x) = \begin{cases} x^{1/3} & \text{if } x > \left(\frac{24}{116}\right)^3 \\ \left(\frac{841}{108}\right) x + \frac{16}{116} & \text{if } x \leq \left(\frac{24}{116}\right)^3 \end{cases}$$

where X, Y, and Z are the tristimulus values to be represented in the new

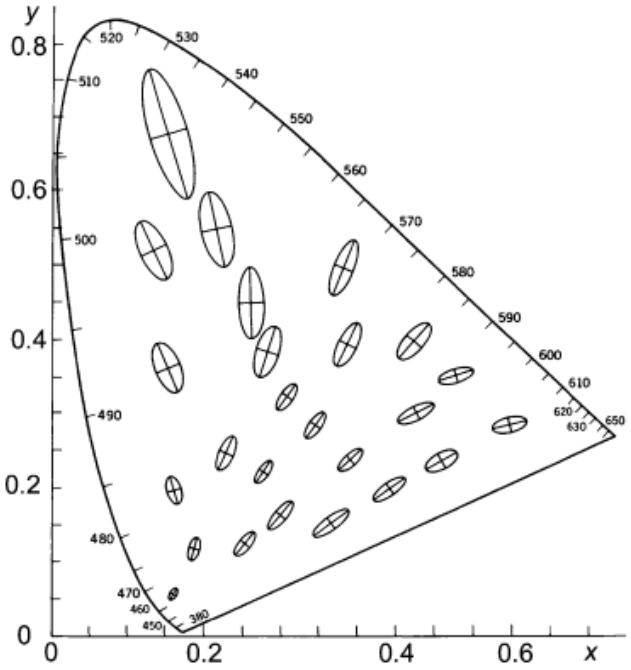


Figure B.5: *MacAdam's ellipsis (Schanda, 2007). The ellipsis represent 10 times the region of the just noticeable difference. This highlight the non-uniformity for the XYZ colour space and why it is unsuitable for taking out the colour differences in this space.*

space and X_n, Y_n, and Z_n are the tristimulus values of the adapting white.

In this space we can define quantities which are correlated with the perceptual attributes of lightness, chroma, and hue as:

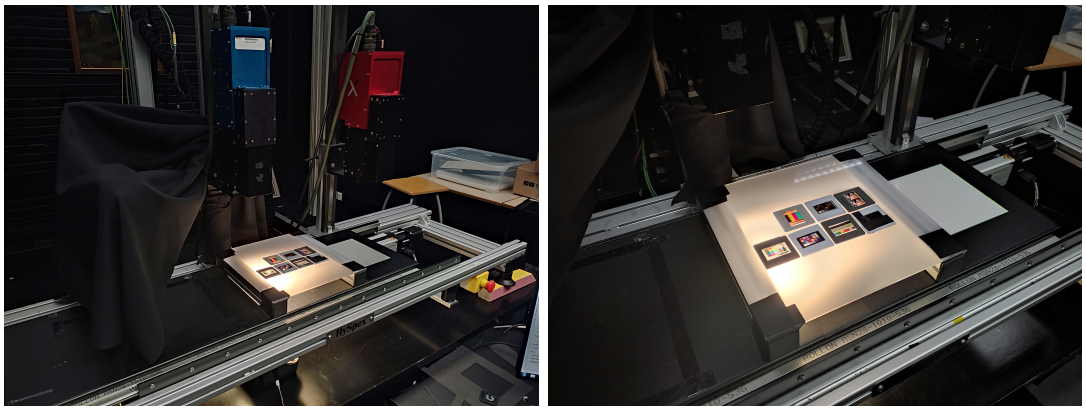
$$\begin{aligned} \text{Lightness:}\ & L^* \\ \text{Chroma:}\quad & C^*_{\text{ab}} = \left(a^{*2} + b^{*2}\right)^{1/2} \\ \text{Hue angle:}\ & h_{\text{ab}} = \arctan\left(b^*/a^*\right) \end{aligned}$$

CIELAB's Euclidean distance, ΔE^*_{ab}, provides a considerably better approximation to perceived difference, i.e. a ΔE^*_{ab} of 1 should result in a similar magnitude of perceived difference regardless of whether it occurs for chromatic colours or neutral grey. The CIELAB colour space has improved the approximation of perceived differences in stimuli to a certain extent. However, it has not completely resolved the issue, as there still exists non-uniformity inside the CIELAB space as well. This leads us to the second strategy for dealing with colour space non-uniformity, which is to increase the sophistication of colour difference formulae. CIEDE2000 has

emerged as a prominent colour difference equation following numerous iterations of such equations. For further explanation of different colour spaces and colour difference equations, it is recommended to consult the work of Fairchild (2013).

C | Transmittances calculation through HySpex

We used HySpex VNIR1800 hyperspectral camera for calculating the transmittances of photographics films that we had. The setup is shown in Fig.C.1. The light source used halogen.



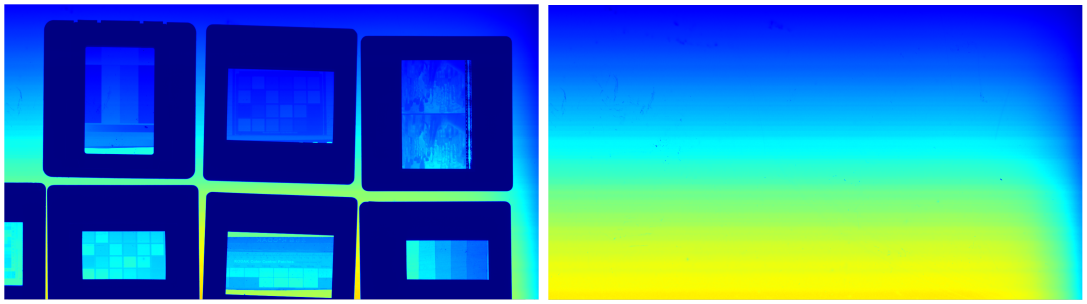
(a) *setup picture 1* **(b)** *setup picture 2*

Figure C.1: *Transmittance calculation*

We show the saturation map from the HySpex software in Fig.C.2

Before using the halogen map, we also had used LED backlight as our light source. The setup is shown in Fig.C.3. The problem that we realized in this setup is the measurement of transmittance values greater than 100%.

The problem is highlighted when we see the saturation map for LED backlight in Fig.C.4. We can see that with films, there is no saturation in the image, but when we take captures with films, we see high saturations around the films. We hypothesized that this might be because when we cover the LED with film frames, the light trying to come out of LED is reflected back and forth within LED case and comes out with more intensity than if there was no film case present. We again took captures of the film with both LED and halogen, with and without frames, to



(a) *Saturation map with films* (b) *Saturation map for light*

Figure C.2: *Saturation maps for the setup*

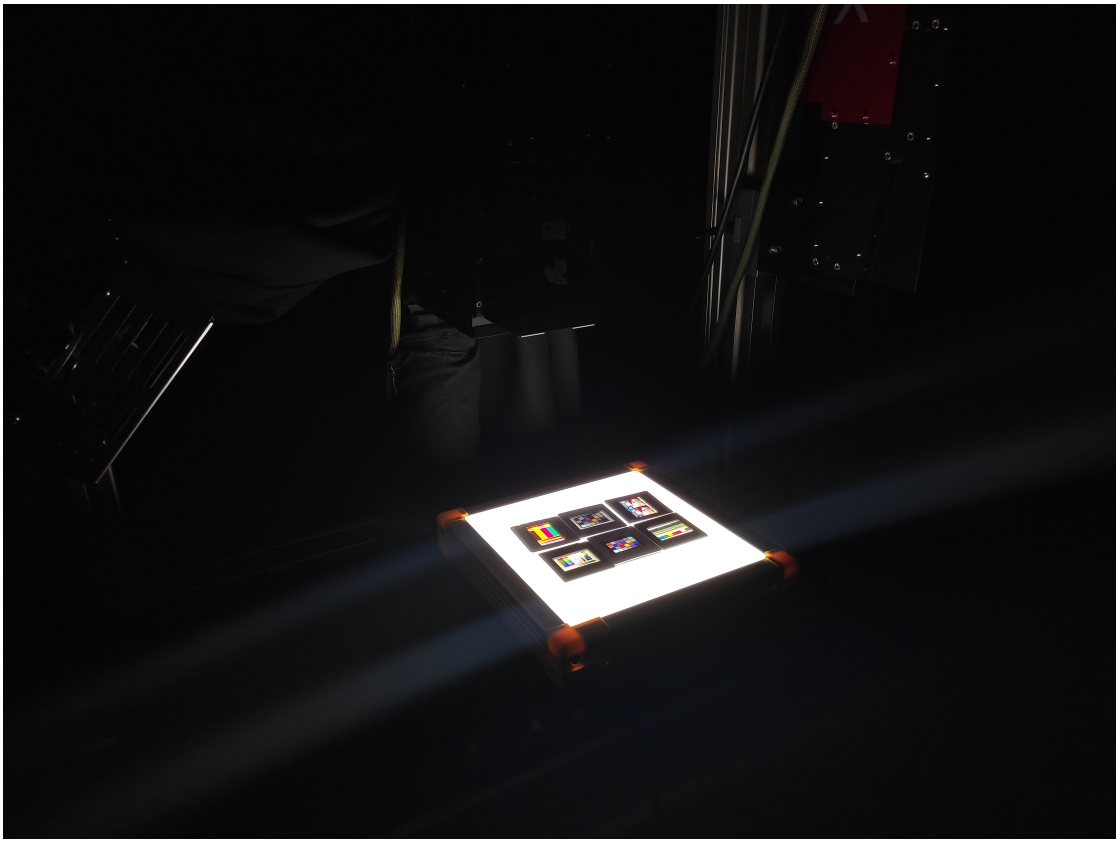
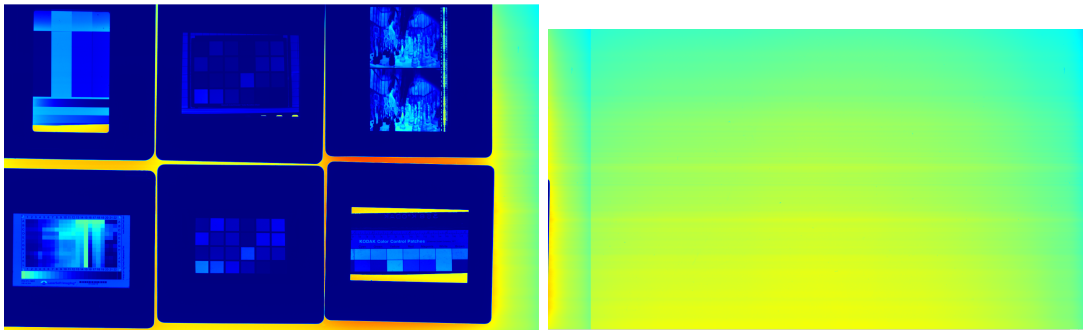


Figure C.3: *Saturation map with films*

check this phenomenon.

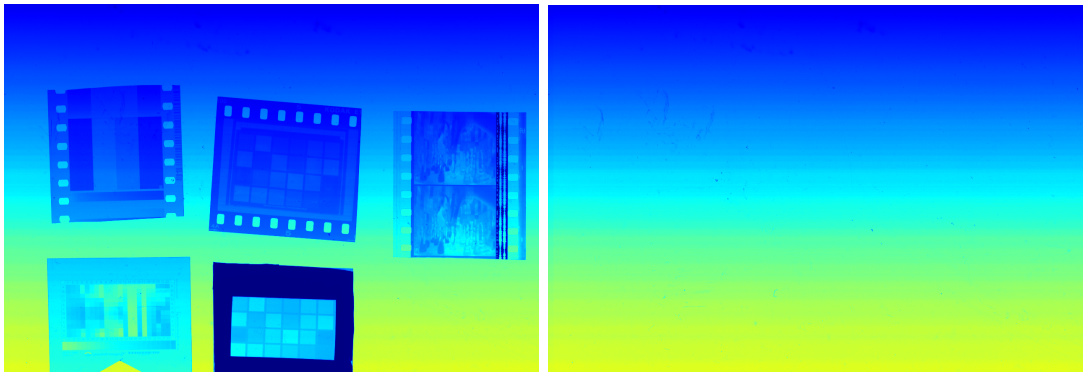
We show the saturation map of the setup with halogen without film frames in Fig.C.5.

We plot the transmittances measured from the four different setups of halogen



(a) *Saturation map with films for LED backlight* **(b)** *Saturation map for light for LED backlight*

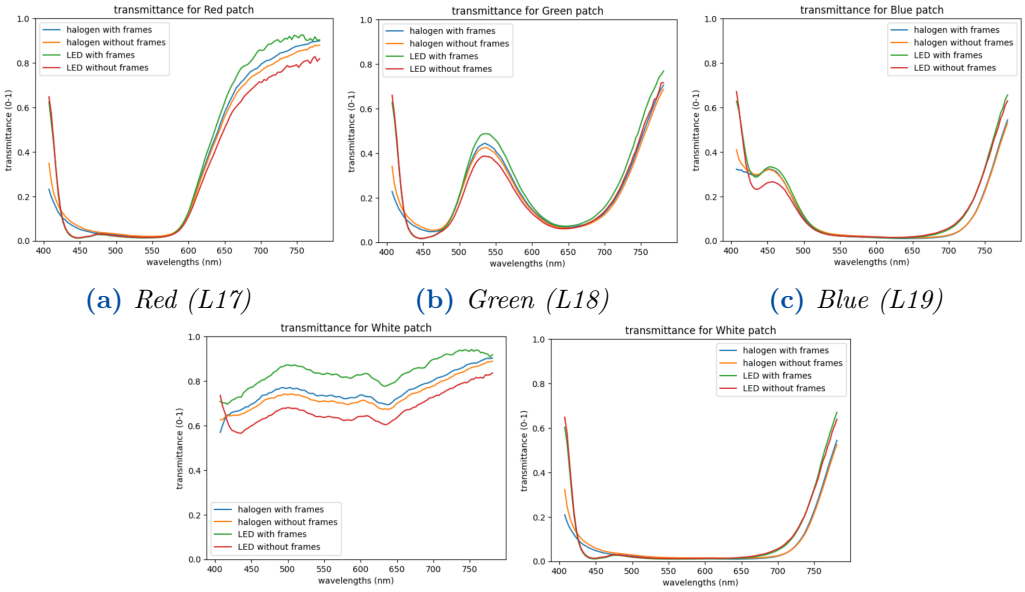
Figure C.4: *Saturation maps for the setup with LED backlight*



(a) *Saturation map with films for halogen without film frames* **(b)** *Saturation map for light for halogen without film frames*

Figure C.5: *Saturation maps for the setup for halogen without film frames*

with/without film frames and LED backlight with/without film frames in Fig.C.6. Due to this difficulty in measurements with the hyperspectral camera. We finally ended up with the setup of the overhead projector as described in Chapter 4.



(a) *Red (L17)* **(b)** *Green (L18)* **(c)** *Blue (L19)*

(d) *White (Top most white)* **(e)** *Black (bottom most black)*

Figure C.6: *Transmittances calculated from different setups.*

Bibliography

Artioli, G. (2010). *Scientific methods and cultural heritage: an introduction to the application of materials science to archaeometry and conservation science*. OUP Oxford. (cited on page 15)

BabelColor (2023). The ColorChecker Pages (Page 1 of 3). [Online; accessed 22. Jul. 2023]. (cited on page 117)

Bai, J. (1997). Estimating Multiple Breaks One at a Time. *Econometric Theory*, 13(3):315–352. (cited on page 138)

Bartczak, P., Fält, P., Penttinen, N., Ylitepsa, P., Laaksonen, L., Lensu, L., Hauta-Kasari, M., and Uusitalo, H. (2017). Spectrally optimal illuminations for diabetic retinopathy detection in retinal imaging. *Optical Review*, 24:105–116. (cited on page 4)

Bartleson, C. J. and Breneman, E. J. (1967). Brightness perception in complex fields. *Josa*, 57(7):953–957. (cited on page 97)

Basri, R. and Jacobs, D. (2001). Photometric stereo with general, unknown lighting. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 2, page II. IEEE. (cited on page 19)

Berns, R., Taplin, L., and Nezamabadi, M. (2004). Modifications of a sinarback 54 digital camera for spectral and high-accuracy colorimetric imaging: simulations and experiments. (cited on page 2)

Berns, R., Taplin, L., and Nezamabadi, M. (2005). Spectral imaging using a commercial colour-filter array digital camera. *The Conservation Committee of the International Council of Museums (ICOM-CC)*. (cited on pages 2 and 41)

Berns, R. S. (1976). Artist paint target (apt): a tool for verifying camera performance. *Journal of Applied Photographic Engineering*, 2:95–99. (cited on page 134)

Berns, R. S. (2005). Color-accurate image archives using spectral imaging. *Scientific Examination of Art: Modern Techniques in Conservation and Analysis*, 12:105–119. (cited on page 41)

Berns, R. S. (2011). Designing white-light led lighting for the display of art: A feasibility study. *Color Research & Application*, 36(5):324–334. (cited on page 4)

Berns, R. S. and Chen, T. (2012). Practical total appearance imaging of paintings. In *Archiving Conference*, volume 9, pages 162–167. Society of Imaging Science and Technology. (cited on page 16)

Birren, F. (1981). Jc le blon, discoverer and developer of the red-yellow-blue principle of color printing and color mixture. *Color Research & Application*, 6(2):85–92. (cited on page 147)

Blottiau, F. (1947). Les defauts d'additivite de la colorimetrie trichromatique. *Revue d'Optique, Theorique et Instrumentale*, 26:193–201. (cited on page 148)

Bowmaker, J. K. and Dartnall, H. (1980). Visual pigments of rods and cones in a human retina. *The Journal of physiology*, 298(1):501–511. (cited on page 147)

Brauers, J., Seiler, C., and Aach, T. (2010). Direct psf estimation using a random noise target. In *Digital Photography VI*, volume 7537, pages 96–105. SPIE. (cited on pages 22 and 23)

Brown, P. K. and Wald, G. (1964). Visual pigments in single rods and cones of the human retina. *Science*, 144(3614):45–52. (cited on page 147)

Brumann, C. (2015). Cultural heritage. In *International encyclopedia of the social and behavioral sciences*, pages 414–419. Elsevier. (cited on page 14)

Camaiti, M., Benvenuti, M., Costagliola, P., Di Benedetto, F., and Moretti, S. (2017). Hyperspectral sensors for the characterization of cultural heritage surfaces. *Sensing the Past: From artifact to historical site*, pages 289–311. (cited on page 16)

Casini, A., Bacci, M., Cucci, C., Lotti, F., Porcinai, S., Picollo, M., Radicati, B., Poggesi, M., and Stefani, L. (2005). Fiber optic reflectance spectroscopy and hyper-spectral image spectroscopy: two integrated techniques for the study of the madonna dei fusi. In *Optical methods for arts and archaeology*, volume 5857, pages 177–184. SPIE. (cited on page 2)

Chen, T. and Berns, R. (2012). Measuring the total appearance of paintings using a linear source, studio strobes, and a dual-rgb camera. Technical report, Technical report, Rochester Institute of Technology, College of Science .... (cited on pages 3 and 175)

Cho, T. S., Paris, S., Horn, B. K., and Freeman, W. T. (2011). Blur kernel estimation using the radon transform. In *CVPR 2011*, pages 241–248. IEEE. (cited on page 22)

Christens-Barry, W. A., Boydston, K., France, F. G., Knox, K. T., Easton Jr, R. L., and Toth, M. B. (2009). Camera system for multispectral imaging of documents. In *Sensors, Cameras, and Systems for Industrial/Scientific Applications X*, volume 7249, pages 32–41. SPIE. (cited on page 3)

CIE (2023). Industrial colour-difference evaluation | CIE. [Online; accessed 25. Jun. 2023]. (cited on page 148)

Cucci, C., Casini, A., Stefani, L., Picollo, M., and Jussila, J. (2017). Bridging research with innovative products: a compact hyperspectral camera for investigating artworks: a feasibility study. In *Optics for Arts, Architecture, and Archaeology VI*, volume 10331, pages 17–29. SPIE. (cited on page 16)

Cucci, C., Picollo, M., Chiarantini, L., and Sereni, B. (2015). Hyperspectral remote sensing techniques applied to the noninvasive investigation of mural paintings: a feasibility study carried out on a wall painting by beato angelico in florence. In *Optics for arts, architecture, and archaeology V*, volume 9527, pages 151–159. SPIE. (cited on page 16)

Cucci, C., Webb, E. K., Casini, A., Ginanni, M., Prandi, E., Stefani, L., Vitorino, T., and Picollo, M. (2019). Short-wave infrared reflectance hyperspectral imaging for painting investigations: A methodological study. *Journal of the American Institute for Conservation*, 58(1-2):16–36. (cited on page 15)

Daniel, F., Mounier, A., Pérez-Arantegui, J., Pardos, C., Prieto-Taboada, N., de Vallejuelo, S. F.-O., and Castro, K. (2016). Hyperspectral imaging applied to the analysis of goya paintings in the museum of zaragoza (spain). *Microchemical Journal*, 126:113–120. (cited on page 5)

Dartnall, H. and Lythgoe, J. (1965). The spectral clustering of visual pigments. *Vision research*, 5(4-5):81–100. (cited on page 147)

Debevec, P. E. and Malik, J. (1997). Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 369–378. ACM Press/Addison-Wesley Publishing Co., USA. (cited on pages 5, 19, 20, and 96)

Delaney, J. K. and Picollo, M. (2019). Jaic special issue on "reflectance hyperspectral imaging to support documentation and conservation of 2d artworks". (cited on page 15)

Delaney, J. K., Zeibel, J. G., Thoury, M., Littleton, R., Palmer, M., Morales, K. M., de La Rie, E. R., and Hoenigswald, A. (2010). Visible and infrared imaging spectroscopy of picasso's harlequin musician: mapping and identification of artist materials in situ. *Applied spectroscopy*, 64(6):584–594. (cited on page 2)

digiCamControl (2023). Easy to use, free solution for complex camera control | digiCamControl. [Online; accessed 10. Aug. 2023]. (cited on page 27)

Du, H. and Voss, K. J. (2004). Effects of point-spread function on calibration and radiometric accuracy of ccd camera. *Applied optics*, 43(3):665–670. (cited on pages 22 and 23)

Durmus, D., Abdalla, D., Duis, A., and Davis, W. (2018). Spectral optimization to minimize light absorbed by artwork. *Leukos*. (cited on page 4)

Durmus, D., Abdalla, D., Duis, A., and Davis, W. (2020). Spectral optimization to minimize light absorbed by artwork. *Leukos*, 16(1):45–54. (cited on pages 15 and 68)

Elbeltagi, E., Hegazy, T., and Grierson, D. (2005). Comparison among five evolutionary-based optimization algorithms. *Advanced engineering informatics*, 19(1):43–53. (cited on pages 59, 61, and 68)

EMVA (2017). [Online; accessed 23. Jun. 2023]. (cited on page 20)

Fairchild, M. D. (2013). *Color appearance models*. John Wiley & Sons. (cited on page 157)

Falt, P. (2022). Color science laboratory. University Lecture. (cited on pages 63 and 181)

Fält, P., Hiltunen, J., Hauta-Kasari, M., Sorri, I., Kalesnykiene, V., Pietilä, J., and Uusitalo, H. (2011). Spectral imaging of the human retina and computationally determined optimal illuminants for diabetic retinopathy lesion detection. *J. Imaging Sci. Technol*, 55(3):030509. (cited on page 4)

Faries, M. (2005). *Scientific examination of art: modern techniques in conservation and analysis*. National Academies Press. (cited on page 16)

Farrell, J., Okincha, M., and Parmar, M. (2008). Sensor calibration and simulation. In *Proceedings Volume 6817, Digital Photography IV*, volume 6817, pages 249–257. SPIE. (cited on page 20)

Farrell, J. E. (1993). Scanner linearity. *Journal of Electronic Imaging*, 2(3):225–230. (cited on page 21)

Finlayson, G. D., Hordley, S. D., and HubeL, P. M. (2001). Color by correlation: a simple, unifying framework for color constancy. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1209–1221. (cited on page 20)

France, F. G., Toth, M. B., Christens-Barry, W. A., and Boydston, K. (2010). Advanced spectral imaging for microanalysis of cultural heritage. *Microscopy and Microanalysis*, 16(S2):728–729. (cited on page 3)

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Stat.*, 42(6):2243–2281. (cited on page 138)

Garini, Y., Young, I. T., and McNamara, G. (2006). Spectral imaging: principles and applications. *Cytometry Part A: The Journal of the International Society for Analytical Cytology*, 69(8):735–747. (cited on page 9)

Grossberg, M. D. and Nayar, S. K. (2003). What is the space of camera response functions? *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.* (cited on pages 20 and 175)

Hornby, G., Globus, A., Linden, D., and Lohn, J. (2006). Automated antenna design with evolutionary algorithms. In *Space 2006*, page 7242. Space 2006. (cited on page 61)

Hubel, P. M., Sherman, D., Farrell, J. E., Sherman, D., and Farrell, J. E. (1994). A Comparison of Methods of Sensor Spectral Sensitivity Estimation. *Color and Imaging Conference*, 2:45–48. (cited on pages 21 and 175)

Hunt, D. (1998). Rwg: Measuring colour. (cited on page 148)

Hunt, R. W. G. (2005). *The reproduction of colour*. John Wiley & Sons. (cited on page 153)

Hyttinen, J., Fält, P., Fauch, L., Riepponen, A., Kullaa, A., and Hauta-Kasari, M. (2018). Contrast enhancement of dental lesions by light source optimisation. In *International Conference on Image and Signal Processing*, pages 499–507. Springer. (cited on page 4)

Irani, M. and Peleg, S. (1991). Improving resolution by image registration. *CVGIP: Graphical models and image processing*, 53(3):231–239. (cited on page 96)

Jemec, J., Pernuš, F., Likar, B., and Bürmen, M. (2017). 2d sub-pixel point spread function measurement using a virtual point-like source. *International journal of computer vision*, 121:391–402. (cited on pages 22 and 23)

Jiang, H., Liu, Y., Li, X., Zhao, H., and Liu, F. (2018). Point spread function measurement based on single-pixel imaging. *IEEE Photonics Journal*, 10(6):1–15. (cited on pages 22, 23, and 176)

Joshi, N., Szeliski, R., and Kriegman, D. J. (2008). Psf estimation using sharp edge prediction. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE. (cited on pages 22 and 23)

Kee, E., Paris, S., Chen, S., and Wang, J. (2011). Modeling and removing spatially-varying optical blur. In *2011 IEEE international conference on computational photography (ICCP)*, pages 1–8. IEEE. (cited on pages 22 and 23)

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, volume 4, pages 1942–1948. IEEE. (cited on pages 61 and 65)

Kuzio, O. and Farnand, S. (2021). Color accuracy-guided data reduction for practical led-based multispectral imaging. In *Archiving Conference*, volume 18, pages 65–70. Society for Imaging Science and Technology. (cited on page 41)

Kuzio, O. and Farnand, S. (2022a). Comparing Practical Spectral Imaging Methods for Cultural Heritage Studio Photography. *J. Comput. Cult. Herit.*, 16(1):1–13. (cited on pages 3 and 42)

Kuzio, O. R. (2023). *Practical LED-based Spectral Imaging for Cultural Heritage Studio Photography*. Rochester Institute of Technology. (cited on pages 2, 3, 41, and 175)

Kuzio, O. R. and Farnand, S. P. (2022b). Toward practical spectral imaging beyond a laboratory context. *Heritage*, 5(4):4140–4160. (cited on pages 42 and 45)

Land, E. H. and McCann, J. J. (1971). Lightness and Retinex Theory. *JOSA*, 61(1):1–11. (cited on page 20)

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 5)

Lehr, J., Sibarita, J.-B., and Chassery, J.-M. (1998). Image restoration in x-ray microscopy: Psf determination and biological applications. *IEEE transactions on image processing*, 7(2):258–263. (cited on page 22)

Lewis, T. O. and Odell, P. L. (1971). Estimation in linear models. [Online; accessed 23. Jun. 2023]. (cited on page 21)

Li, Q., He, X., Wang, Y., Liu, H., Xu, D., and Guo, F. (2013). Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of biomedical optics*, 18(10):100901–100901. (cited on page 12)

Liang, H., Lucian, A., Lange, R., Cheung, C. S., and Su, B. (2014). Remote spectral imaging with simultaneous extraction of 3d topography for historical wall paintings. *ISPRS Journal of Photogrammetry and Remote Sensing*, 95:13–22. (cited on page 16)

Lomheim, T. S. and Kalman, L. S. (2020). Analytical Modeling and Digital Simulation of Scanning Charge-Coupled Device Imaging Systems. In *Electro-Optical Displays*, pages 513–584. CRC Press, Boca Raton, FL, USA. (cited on page 18)

Lovbjerg, M. (2002). Improving particle swarm optimization by hybridization of stochastic search heuristics and self-organized criticality. *Master's thesis, Department of Computer Science, University of Aarhus*. (cited on page 61)

Luong, Q.-T., Fua, P., and Leclerc, Y. G. (2002). The radiometry of multiple images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(1):19–33. (cited on pages 19 and 20)

MacAdam, D. L. (1942). Visual sensitivities to color differences in daylight. *Josa*, 32(5):247–274. (cited on page 154)

MacDonald, L. and Luo, M. R. (1999). *Colour Imaging: Vision and Technology*. Wiley, Hoboken, NJ, USA. (cited on page 17)

Mancill, C. E. (1975). Digital color image reproduction. (cited on page 21)

Mann, S. (1993). Compositing multiple pictures of the same scene. In *Proc. IS&T Annual Meeting, 1993*, pages 50–52. (cited on pages 95 and 96)

Mann, S. (2000). Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. Image Process.*, 9(8):1389–1406. (cited on page 20)

Mann, S. and Ali, M. (2016). The fundamental basis of hdr: Comparametric equations. In *High Dynamic Range Video*, pages 1–59. Elsevier. (cited on pages 96, 97, and 184)

Mann, S., Janzen, R. E., and Hobson, T. (2010). Multisensor broadband high dynamic range sensing: for a highly expressive step-based musical instrument. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, pages 21–24. (cited on pages 5 and 95)

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 95)

Mann, S. and Picard, R. W. (1994a). Being `undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. [Online; accessed 22. Jun. 2023]. (cited on pages 19 and 20)

Mann, S. and Picard, R. W. (1994b). Virtual bellows: Constructing high quality stills from video. In *Proceedings of 1st International Conference on Image Processing*, volume 1, pages 363–367. IEEE. (cited on page 96)

Mann, S. and Picard, R. W. (1997). Video orbits of the projective group a simple approach to featureless estimation of parameters. *IEEE Transactions on Image Processing*, 6(9):1281–1295. (cited on page 96)

Marks, W., Dobelle, W. H., and MacNichol Jr, E. F. (1964). Visual pigments of single primate cones. *Science*, 143(3611):1181–1183. (cited on page 147)

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 5)

Mitsunaga, T. and Nayar, S. (1999). Radiometric self calibration. *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. (cited on pages 19 and 20)

Mosleh, A., Green, P., Onzon, E., Begin, I., and Pierre Langlois, J. (2015). Camera intrinsic blur kernel estimation: A reliable framework. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4961–4968. (cited on page 22)

Navas-Moya, F. A., Nieves, J. L., Valero, E. M., and Garrote, E. (2013). Measurement of the optical transfer function using a white-dot pattern presented on a liquid-crystal display. *J. Eur. Opt. Soc.*, 8:13029. (cited on pages 22 and 23)

Nayar, S. K., Ikeuchi, K., and Kanade, T. (1990). Shape from interreflections. In *[1990] Proceedings Third International Conference on Computer Vision*, pages 2–11. IEEE. (cited on page 19)

Penrose, R. (1956). On best approximate solutions of linear matrix equations. *Math. Proc. Cambridge Philos. Soc.*, 52(1):17–19. (cited on page 40)

Picollo, M., Casini, A., Cherubini, F., Cucci, C., Grazzi, F., and Stefani, L. (2020a). Potenzialità e problematiche dell'uso di tecniche di spettroscopia d'immagine per lo studio di pellicole fotografiche. *Memoria Fotografica. Storia di un Recupero Collettivo; Picollo, M., Cattaneo, B., Vianelli, D., Eds*, pages 43–46. (cited on page 16)

Picollo, M., Cucci, C., Casini, A., and Stefani, L. (2020b). Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios. *Sensors*, 20(10):2843. (cited on page 15)

Pinna, D., Galeotti, M., and Mazzeo, R. (2009). Practical handbook on diagnosis of paintings on movable support. european project artech. *Centro Di: Firenze, Italy*. (cited on page 15)

Pitt, I. and Winter, L. (1974). Effect of surround on perceived saturation. *JOSA*, 64(10):1328–1331. (cited on page 97)

PK, B. and WALD, G. (1964). Visual pigments in single rods and cones of the human retina. direct measurements reveal mechanisms of human night and color vision. *Science (New York, NY)*, 144(3614):45–52. (cited on page 147)

Poorvi, L. and Farrell, J. E. (2018). [Online; accessed 23. Jun. 2023]. (cited on page 20)

Pratt, W. K. and Mancill, C. E. (1976). Spectral estimation techniques for the spectral calibration of a color image scanner. *Appl. Opt.*, 15(1):73–75. (cited on page 21)

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 page 96)

Ribés, A., Brettel, H., Schmitt, F., Liang, H., Cupitt, J., and Saunders, D. (2003). Color and multispectral imaging with the crisatel multispectral system. In *PICS*, pages 215–219. (cited on page 16)

Ripps, H. and Weale, R. (1963). Cone pigments in the normal human fovea. *Vision Research*, 3(11-12):531–543. (cited on page 147)

Robertson, M. A., Borman, S., and Stevenson, R. L. (2003). Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of electronic imaging*, 12(2):219–228. (cited on page 95)

Rushton, W. (1957). Visual problems of colour. In *Proc. NPL Symposium*. (cited on page 147)

Rushton, W. (1958). Visual pigments in the colour blind. *Nature*, 182:690–692. (cited on page 147)

Saunders, D. and Cupitt, J. (1993). Image processing at the national gallery: The vasari project. *National Gallery technical bulletin*, 14(1):72–85. (cited on page 1)

Schanda, J. (2007). *Colorimetry: understanding the CIE system*. John Wiley & Sons. (cited on pages 148, 151, 152, 155, 156, 187, and 188)

Schanda, J., Csuti, P., and Szabo, F. (2016). A new concept of color fidelity for museum lighting: based on an experiment in the sistine chapel. *Leukos*, 12(1-2):71–77. (cited on pages 4 and 15)

Sharma, G., Trussell, H. J., and Trussell, H. J. (1993). Characterization of Scanner Sensitivity. *Color and Imaging Conference*, 1:103–107. (cited on page 21)

Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. In *1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360)*, pages 69–73. IEEE. (cited on page 65)

Studio, C. (2023). Camo - Great video made easy. [Online; accessed 21. Jul. 2023]. (cited on pages 36 and 178)

Sun, L., Cho, S., Wang, J., and Hays, J. (2013). Edge-based blur kernel estimation using patch priors. In *IEEE international conference on computational photography (ICCP)*, pages 1–8. IEEE. (cited on pages 22 and 23)

Szeliski, R. (1996). Video mosaics for virtual environments. *IEEE computer Graphics and Applications*, 16(2):22–30. (cited on page 96)

Szeliski, R. (2005). Fast shape from shading. In *Computer Vision — ECCV 90*, pages 359–368. Springer, Berlin, Germany. (cited on page 19)

Thornton, W. A. (1992). Toward a more accurate and extensible colorimetry. part ii. discussion. *Color research & application*, 17(3):162–186. (cited on page 149)

Trezona, P. (1953). Additivity of colour equations. *Proceedings of the Physical Society. Section B*, 66(7):548. (cited on page 148)

Trezona, P. (1954). Additivity of colour equations: Ii. *Proceedings of the Physical Society. Section B*, 67(7):513. (cited on page 148)

Trumpy, G., Hardeberg, J. Y., George, S., and Flueckiger, B. (2021). A multispectral design for a new generation of film scanners. *Proc. SPIE*. (cited on pages 30 and 33)

Trussell, H. J. and Sharma, G. (1994). Signal processing methods in color calibration. In *Proceedings Volume 2170, Device-Independent Color Imaging*, volume 2170, pages 18–23. SPIE. (cited on page 21)

Tsin, Y., Ramesh, V., and Kanade, T. (2001). Statistical calibration of CCD imaging process. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 480–487vol.1. IEEE. (cited on page 20)

Tuzikas, A., Žukauskas, A., Vaicekauksas, R., Petrulis, A., Vitta, P., and Shur, M. (2014). Artwork visualization using a solid-state lighting engine with controlled photochemical safety. *Optics Express*, 22(14):16802–16818. (cited on page 4)

W3C (2022). CSS Color Module Level 3. [Online; accessed 6. Aug. 2023]. (cited on page 48)

Wall, E. J. (1925). *The history of three-color photography*. American Photographic Publishing Company. (cited on page 147)

Weale, R. A. (1957). Trichromatic Ideas in the Seventeenth and Eighteenth Centuries. *Nature*, 179:648–651. (cited on page 147)

Weale, R. A. (1959). Photo-sensitive reactions in foveae of normal and cone-monochromatic observers. *Optica Acta: International Journal of Optics*, 6(2):158–174. (cited on page 147)

Woodham, R. J. (1980). Photometric Method For Determining Surface Orientation From Multiple Images. In *Optical Engineering, Vol. 19, Issue 1*, volume 19, pages 139–144. SPIE. (cited on page 19)

Wright, W. D. (1934). The measurement and analysis of colour adaptation phenomena. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 115(791):49–87. (cited on page 148)

Wright, W. D. (1936). The breakdown of a colour match with high intensities of adaptation. *The Journal of Physiology*, 87(1):23. (cited on page 148)

BIBLIOGRAPHY

List of Figures

List of Tables