Towards Colour-Accurate Documentation of Anonymous Expressions

Adolfo Molada-Tebar [1,*], Geert J. Verhoeven [2]
[1] Department of Cartographic and Land Engineering, Higher Polytechnic School of Avila, University of Salamanca, Hornos Caleros 50, 05003, Ávila; E-Mail: admote@usal.es
[2] Ludwig Boltzmann Gesellschaft, LBI ArchPro, 1190 Vienna, Austria; E-Mail: geert@projectindigo.eu
* Corresponding author

Abstract
Colour is a powerful communication element in most forms of cultural heritage. This importance of colour notwithstanding, the documentation of cultural heritage typically captures the geometrical aspects and seldom the spectral dimensions of an artefact. This is partly because the science of colour (called colorimetry) is non-trivial. In addition, capturing accurate colour data with digital cameras remains challenging due to the operating principle of standard imaging sensors and the need for a stable and well-characterised illumination source. Despite these limitations, the heritage science project INDIGO made it one of its central aims to generate colour-accurate photos from graffiti captured with standard digital cameras in varying outdoor illumination conditions. This paper first discusses the importance of colour accuracy in graffiti documentation. Afterwards, the text details (in a non-mathematical manner) essential colorimetric and camera principles that underlie the generation of colour images from raw image sensor data. This in-depth coverage supports clarifying the main hurdles to accurate photo colours. Finally, the paper introduces the open-source COOLPI software resulting from this research. We are confident that COOLPI will benefit any other heritage documentation project, or any application where digital cameras play a fundamental role in acquiring correct colour values.

Keywords
camera characterisation; CIE colour spaces; colorimetry; colour transformation; COOLPI; graffiti; image processing; Python; RAW photo

Acronyms

ADC	Analogue-to-Digital Converter
AW	Adopted White
CAT	Chromatic Adaptation Transform
CCM	Camera Characterisation Matrix
CCT	Correlated Colour Temperature
CFA	Colour Filter Array
CIE	International Commission on Illumination
CMF	Colour-Matching Function
CMYK	Cyan-Magenta-Yellow-blacK/Key
CRS	Coordinate Reference System
CS	Coordinate System
DN	Digital/Data Number
EV	Exposure Value
Exif	Exchangeable image file format
FoV	Field of View
HVS	Human Vision/Visual System
JP(E)G	Joint Photographic Experts Group
OSM	Output Space Matrix
RGB	Red-Green-Blue
SPD	Spectral Power Distribution
TIFF	Tag(ged) Image File Format

1. Introduction

1.1. Colour and Cultural Heritage

Even though colour is not a physical attribute of an object but a perceived human physiological sensation, colour should be a significant focus point in cultural heritage documentation (Molada-Tebar, Marqués-Mateu, & Lerma, 2019b). As a descriptive attribute, colour is indispensable for proper object recognition and cataloguing, but also for tasks like damage detection or restoration, to name but a few (Boochs et al., 2014). However, colour is so much more; it is an effective communication tool that evokes emotions and can profoundly affect viewers (Chen et al., 2020; Hanada, 2018). That is why colour is crucial in creating and studying graffiti. Graffitists often aim to communicate social, cultural or political ideas so that they remain engraved in the minds of passers-by. In this sense, graffiti can be considered a powerful form of visual communication free from conventional restrictions (Velikonja, 2020). To reach the intended impact of their visual message, graffiti creators often rely on a varied, vivid and striking colour palette (Feitosa-Santana et al., 2020).

Colour is thus one of the absolute distinguishing features of a graffito, often even more than its geometrical aspects. Given this significance, one of graffiti project INDIGO's primary research aims is to obtain digital colour values for each part of a graffito. These values should be as close to reality as possible (i.e. colorimetric) to digitally preserve the spectral characteristics of the work and the essence of the message that the creators intend to convey. However, obtaining colour-accurate data remains a significant challenge in heritage documentation (Korytkowski & Olejnik-Krugly, 2017). Using colorimetric instruments directly on cultural assets might be forbidden or can be cost-ineffective. For example, it would take very long to acquire sufficient samples with a portable spectrophotometer from all the differently coloured regions of a painting, like a mural. That is why INDIGO relies on digital photographs. Given the large number of samples in a digital photo (i.e. each of the many million image pixels has at least a red, green and blue spectral value), photographs might be considered a good and fast approach to obtaining dense colour data over a large spatial extent. In addition, photographing is a cost-effective, non-contact and physically non-invasive method. When many overlapping photos are available, it also becomes possible to extract realistic and accurate 3D models that digitally encode the surface geometry (Verhoeven et al., 2022) or support various analytical tasks like change detection (Palomar-Vazquez et al., 2017). This explains why many traditional heritage documentation endeavours and some less-conventional projects on contemporary graffiti recording, like INDIGO and others (Rodriguez-Navarro et al., 2020), consider photographs the primary source to obtain spatial and spectral data about the study object(s).

However, whereas the acquisition and processing workflows for photo-based 3D surface modelling have achieved a certain maturity and consensus, this is not true for photo-based colour extraction. It turns out that generating colour-accurate photos is still relatively complex, as it requires a thorough understanding of colorimetry (i.e. the science of colour), a camera's hardware plus processing pipeline, and the illumination source used while photographing. In addition, various colorimetric limitations related to surface reflectance and digital cameras must be considered (Kirchner et al., 2021). In summary: one cannot rely on a standard digital photo camera for rigorous and objective colour determination without a robust photo acquisition and colour management procedure (Molada-Tebar et al., 2018). The topic of this paper is to provide the necessary background for understanding this problem and to offer an approach for retrieving colorimetric data from digital photos of graffiti.

1.2. Colour Within Project INDIGO

Project INDIGO aims to build the basis to systematically document and digitally disseminate almost 13 km of uninterrupted graffiti along the *Donaukanal* (Eng. Danube Canal), the central waterway of Vienna, Austria (Verhoeven et al., 2022). The project hopes to digitally preserve this unique form of volatile cultural heritage and open new analytical pathways for such large graffiti-scapes. To that end, the photo-based documentation aims to include the geometrical (i.e. shape and dimensions), spectral (i.e. colour), geographical (i.e. location), temporal (i.e. time of creation and lifespan) and contentual (i.e. subject matter and meaning) aspects of every graffito.

Striving for colour-accurate graffiti photographs was prompted since colour is essential in graffiti's visual appearance and study. INDIGO wants to ensure that the result of a red spray can is digitally archived as red pixels and not as orange ones. Colour-accurate photo records not only keep the intended visual impact of the graffiti creator but can also open new windows for analysis: what are the dominant colours in this graffiti-scape? Can we link certain brands of spray cans to specific colour values? Do certain creators consistently use the same colour palette? In addition, project INDIGO wants to create an online platform where everybody can query and visualise graffiti records. A part of this platform should provide an extensive 3D surface model of the entire research zone. This 3D model will include overlapping, multi-temporal textures (i.e. a texture patch for every new graffito). Enforcing colour accuracy for every texture patch should help minimise tonal differences across these textures' seams, thereby contributing to a smooth viewing experience.

Finally, all of INDIGO's data will become freely available at the project's end. Since it is not unimaginable that machine learning engineers will use this vast photo collection to automate the classification of graffiti by creator or style, consistency of the digital records is vital.

1.3. Colour and Graffiti: an Example

Before diving into the realm of light, colour and cameras, it is beneficial to start with an example to get some basic feeling for the parameters that influence the final colour in a photograph. On top of several camera-dependent factors, one must keep many camera-agnostic variables in check to generate

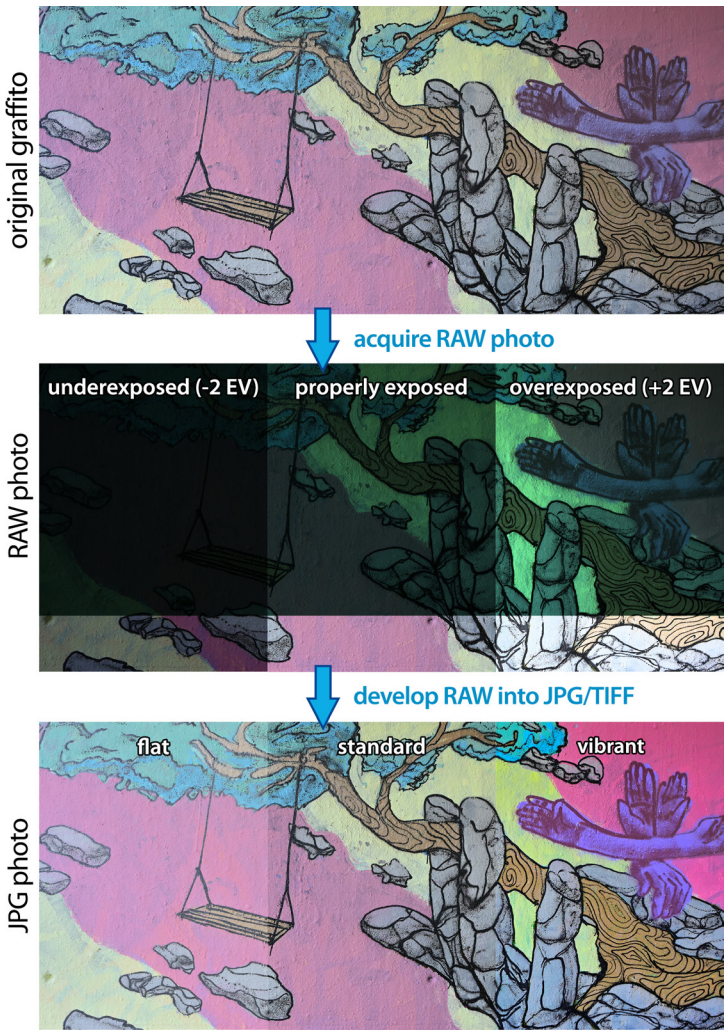


Figure 1. Digitally photographing a graffito (upper row) generates a RAW image (central row), of which the exposure parameters largely influence the pixel values. However, even an adequately exposed RAW image can result in different colours (lower row) depending upon the RAW development settings used inside the camera or by the computer software.

image pixels with approximately correct colour values. Two of those are the exposure values and the development settings—used internally by the camera or externally by dedicated computer software—to convert the raw sensor data into the final pleasing JPG/JPEG or TIFF image. Figure 1 illustrates both. A graffito (Figure 1 top row) created on a wall below Vienna's Rossauer Bridge is photographed with a Nikon Z7ii and a Nikon NIKKOR Z 20mm f/1.8 S lens. The ambient incident light levels were measured with a Sekonic L-358 FLASH MASTER light meter, yielding the following exposure values: ISO 100, *f*/5.6 and 1/13s. These values were dialled into the camera to generate the RAW image in the middle of the central row. The lens aperture and sensor sensitivity were kept invariant at *f*/5.6 and ISO 100 for the photos on its left and right. However, an underexposure of two photographic stops or Exposure Values (EVs) was achieved via a shutter speed of 1/50 s, while a shutter speed of 1/3 s yielded the 2 EVs overexposed photo on the right. Because the values of the photographs are unprocessed or raw (thus constituting a RAW image), two things can be noticed:

- The image looks very green since a camera's imaging sensor is most responsive in the green spectrum. This apparent colour imbalance must be considered when rendering the final JPG or TIFF file.
- The photos look very dark, even the correctly exposed and overexposed ones. The pixel values of a RAW image are ideally varying in a perfectly linear fashion with the incoming light. Like the human visual system, more collected light means higher values. However, computer monitors apply a non-linear tonal transformation when displaying images. This transformation is 'undone' when rendering the final JPG or TIFF file from the RAW image. Since they do not consider this non-linear transformation by the monitor, RAW camera images look very dark.

To understand how these exposures would yield different colours in the final photo, the lowest portion of the RAW images depicts a processed version into a neutrally rendered JPG image. However, the last parts of the entire rendering process also largely influence the resulting colours (Karaimer & Brown, 2019). The lowest row of Figure 1 depicts three different renderings of the properly exposed RAW image. Most cameras and RAW conversion software feature rendering settings like "flat", "standard", and "vibrant". Each setting applies one or more algorithms to 'lift' the shadows or 'suppress' the highlights, to sharpen details and boost or restrain colour vibrancy. All these sub-settings—which solely aim at turning the raw, unprocessed data into a pleasing photo with a specific style—influence the finally rendered pixel values, thus increasing the potential of ending up with inaccurate colours (Ramanath et al., 2005).

In addition, RAW-to-JPG operations are complex, diverse, typically manufacturer-specific and thus proprietary (Chakrabarti et al., 2009), making it hard to understand all rendering operations fully. If they are executed within the camera (i.e. when the camera is set to deliver a pleasing, fully-rendered output image), the process can usually only be controlled via a few basic settings like sharpness, vibrancy and shadow recovery. These, plus many other factors, prevent digital photos from being accurate colorimetric records. If rigorous colour determination based on image pixels is the aim, a colour-aware procedure with control over each parameter is essential (in addition to auxiliary data). Our paper presents the workflow and software to achieve this within project INDIGO. Since such a description necessitates much particular terminology, we decided to let clarity rule over textual conciseness. This makes it possible to provide the reader with a rather in-depth but non-mathematical overview of all critical colour imaging principles and their supporting colorimetric ideas.

The rest of the article is structured as follows. First, we introduce the basics of creating, measuring and communicating colour (Section 2). Second, the basics of creating pixels with a digital imaging sensor are tackled (Section 3). Based on these colorimetry and camera principles, Sections 4 and 5 explain why standard photographs are not colorimetric and which data acquisition and processing steps INDIGO has put in place—in the form of a free and open-source software package COOLPI—to largely alleviate that shortcoming. Finally, Section 6 provides some workflow reflections and covers the imaging situations COOLPI will not be able to handle. All our work is guided by colorimetric standards, particularly those published by the *Commission Internationale de l'Éclairage* (CIE; Eng. International Commission on Illumination).

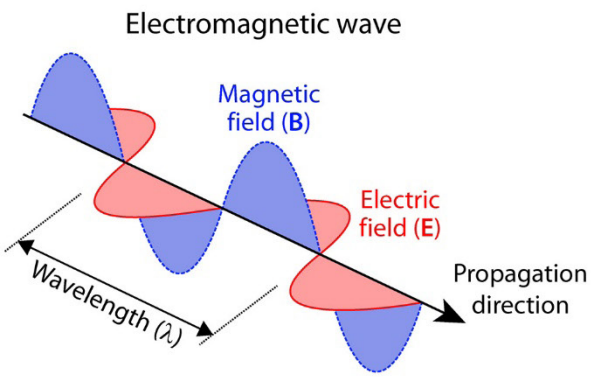
Although written in British English, this text uses the non-British variant "colorimetric" because "colour" and "colorimetric" are official CIE terminology (CIE, 2018).

2. Understanding Colour

Photographing an object works because the light created by at least one source reaches that object; the object then reflects a part of that light towards the camera. However, many hard- and software parts are needed inside the camera to turn the captured signal into a photo that closely resembles the object we witnessed with our eyes. This entire photo acquisition pipeline leverages many colorimetric concepts. The following five sub-sections—based mainly on Verhoeven (2016)—provide a basic introduction to the concepts essential to understanding the proposed image processing workflow.

2.1. Light

Electricity and magnetism are intimately related. Moving a magnet around an electric wire creates an electric current. Still, a moving electric field will also produce a magnetic field. Since both fields create each other, they oscillate together and create a so-called electromagnetic wave (see Figure 2 on the left). Being a wave-like phenomenon, electromagnetic radiation can be distinguished by the length of its waves, called the wavelength (λ). Electromagnetic radiation with a wavelength between 400 nm (400×10^{-9} m) to 700 nm (700×10^{-9} m) is called visible light or simply light (although colorimetric applications often use 380 nm as cut-on and 780 nm as cut-off wavelength). Light is thus only a very narrow spectral band out of all possible electromagnetic radiation (see Figure 2 on the right and Figure 3) and the only wavelengths to which human eyes respond with a visual sensation. To both sides of the visible band resides radiation that does not produce a visual sensation: gamma rays, X-rays and ultraviolet radiation with shorter-than-visible wavelengths, while the long-wavelength region encompasses infrared radiation, microwaves and radiowaves (Figure 3). The part of the electromagnetic spectrum that includes the complete ultraviolet to infrared bandwidth, comprising radiation with wavelengths between 10 nm (0.01 µm) to 1 mm (1000 µm), is called the optical electromagnetic spectrum (Ohno, 2010) (see the classification in Figure 2).



Optical electromagnetic radiation

Division	Subdivision	Abbreviation	Cut-on (nm)	Cut-off (nm)
UltraViolet (UV)	Vacuum UV	VUV / UV-D*	10	200
	Far UV	FUV / UV-C*	200	280
	Middle-UV	MUV / UV-B	280	315
	Near-UV	NUV / UV-A	315	400
Visible (Vis)	Blue	B	400	500
	Green	G	500	600
	Red	R	600	700
InfraRed (IR)	Near-IR	NIR	700	1 100
	Short Wavelength IR	SWIR	1 100	3 000
	Mid Wavelength IR	MWIR	3 000	6 000
	Long Wavelength IR	LWIR	6 000	15 000
	Far/Extreme-IR	FIR	15 000	1 000 000

Figure 2. On the left an electromagnetic wave consisting of electric and magnetic oscillating fields. The oscillating electric field vectors are indicated in red, while the blue lines represent the magnetic field vectors. On the right are the divisions of optical electromagnetic radiation.

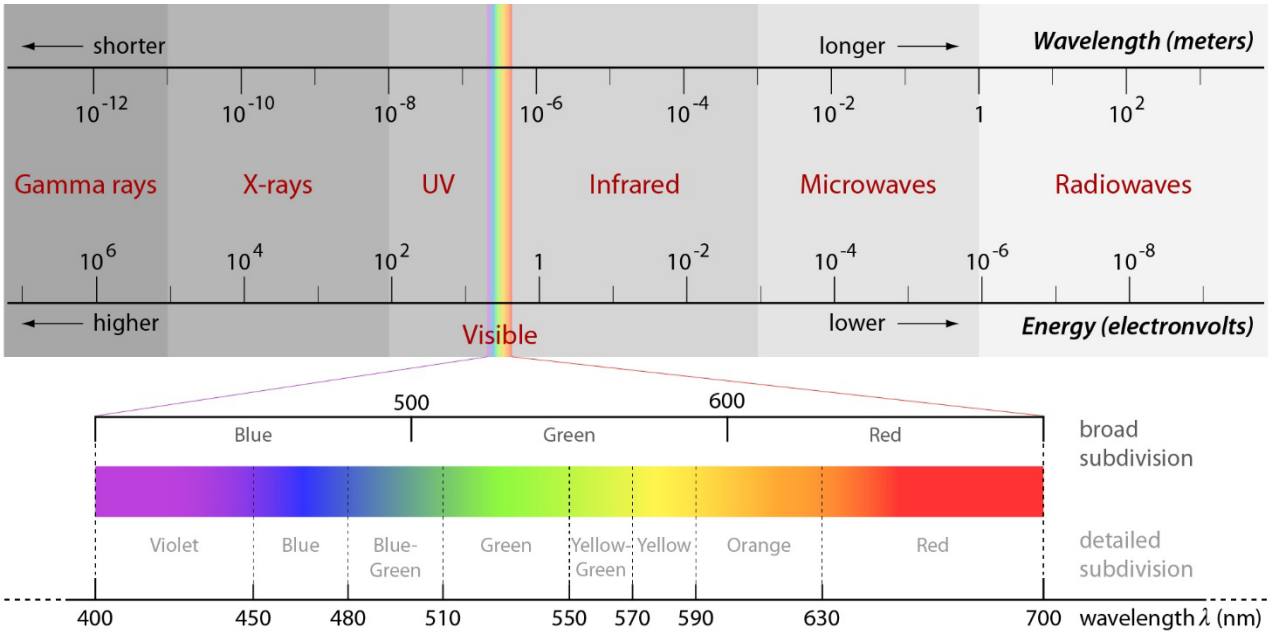


Figure 3. The complete electromagnetic spectrum with the spectral subdivisions of the visible waveband. The spectral hue names are after Hunt (2004).

In the visible wavelength range, each wavelength of light correlates with a sensory impression of a particular colour (or, more technically correct, 'hue'). Even though colour is thus not a physical property of electromagnetic radiation, the light spectrum may be roughly divided according to it, as indicated in Figure 3. The latter shows that the light spectrum contains all hues visible in a rainbow: varying from red on the long-wavelength side over orange, yellow, green and blue to violet on the short-wavelength side (sometimes the hue indigo is defined between blue and violet). For the sake of simplicity, the visible spectrum is usually considered to consist of only three bands: Blue (400 nm – 500 nm), Green (500 nm – 600 nm) and Red (600 nm – 700 nm). Although a coarse approximation, many image-related devices, such as digital cameras and monitors, base their physical working principles on this subdivision.

In addition to the wave properties mentioned above, electromagnetic radiant energy exhibits particle-like behaviour. These indivisible particles are called photons, discrete energy packets with energy levels that differ according to the wavelength. Due to this quantisation, a visible photon with a wavelength of 650 nm will always have 1.9 eV of energy, while photons with quantum energies of 3.6 eV characterise 345 nm ultraviolet radiation. These numbers show that shorter wavelengths have higher radiative energies (see Figure 3). This also explains why highly energetic ultraviolet radiation causes sunburns.

None of the wave-like and particle-like descriptions of electromagnetic radiation is complete by itself. Still, each offers a valid description of some aspects of electromagnetic radiation's behaviour. However, one could as well forget about this wave-particle duality if in need of absolute physical accuracy. Essentially, there are no waves and particles, just quantised fields with discrete excitations. That is also why quantum field theory is the theoretical framework behind the standard model of particle physics (Mandl & Shaw, 2010). However, to understand how light contributes to the creation of colour, this naïve interpretation of light as 'waves' and a stream of 'particles' is satisfying enough.

2.2. Creating Colour

Human colour perception results from the visible electromagnetic radiation received by the eye's photoreceptors and how the brain subsequently interprets these signals (B. B. Lee, 2004). In brief, colour 'happens' inside the human head, and this combined eye-brain processing is known as the Human Vision/Visual System (HVS). Although the general principles of the HVS are known and basically identical amongst all humans, smaller aspects can differ. For instance, some people might be more or less sensitive to Red light. This leads to varying forms of colour deficiencies in male and female populations but also explains why two human observers might perceive colour in more or less different ways. In addition, the emotions that colour evokes are subjective, dependent on the observer's experience and culturally determined (Jonauskaite et al., 2019).

These differences notwithstanding, there is always an interaction between three elements required to generate or see colour: a light source, an object and an observer. At the origin of this imaging chain lies the interaction of light with the scene or object. This interaction determines which portion and quantity of light the HVS (or digital imaging sensor) will detect and process. Figure 4 details this: visible electromagnetic radiation falls onto an object; depending on its physical and chemical properties, the object reflects, absorbs and transmits a fraction of all the incident light. Finally, a digital camera or a human observer picks up the diffusely reflected portion. This signal emanating from the object is called the spectral stimulus signal, the spectral stimulus or simply stimulus. The stimulus is thus always a function of the light source and the object's spectral reflectance. If one of those two components is altered, so will the stimulus. For example, healthy grass is green because it mainly reflects incident solar light between 500 nm and 600 nm, a range perceived by the HVS as green. If grass gets solely illuminated by a light source devoid of wavelengths between 500 nm and 600 nm (e.g. a blue disco light), this part of the spectrum cannot be reflected, making it impossible for the HVS or a digital camera to render grass green.

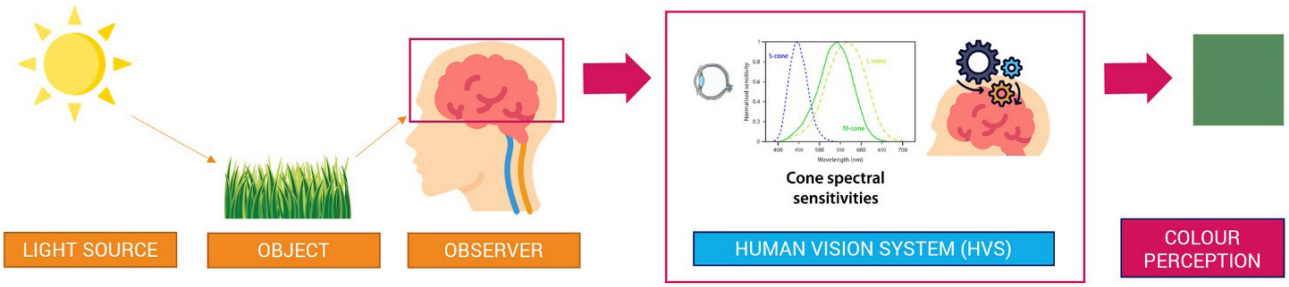


Figure 4. Three elements are needed to perceive colour: a light source, an object and an observer (HVS).

The spectral stimulus which arrives at a digital camera gets integrated over millions of small photodetectors, all of which feature one of three spectral response curves (approximately situated in the Red, Green and Blue parts of the spectrum—see Figure 5A). A human observer integrates the stimulus over the five to six million cones in the retina at the back of the eye (Wandell, 1995). Similar to the camera's imaging sensor, the centre of the retina is densely packed with three cone variants with a specific response to visible wavelengths: cones sensitive to Short, Middle and Long wavelengths (S, M and L), named according to the part of the visible spectrum to which they are most sensitive (Stockman & Sharpe, 2000). Figure 5B depicts the linear normalised spectral sensitivity curves of the S-, M- and L-cones. These curves are known as the cone fundamentals.

Since a digital photo camera and the HVS sample the incoming signal in three different spectral regions, they are so-called trichromatic devices. Their respective spectral response/responsivity curves (also known as their spectral sensitivities) determine the likelihood that a photon of a specific wavelength will contribute to the imaging process

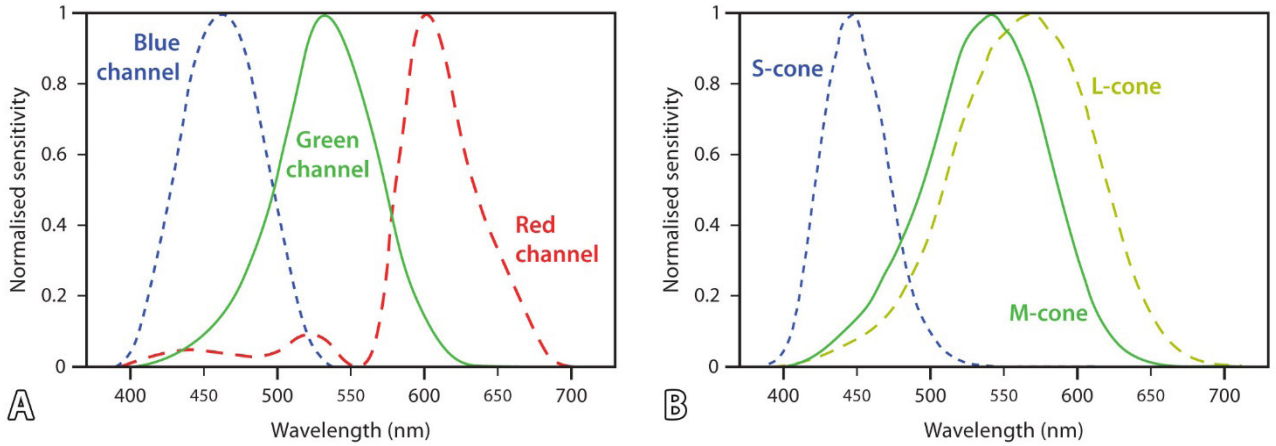


Figure 5. The normalised spectral response curves of a Nikon D200 (A) and the cones in the retina of a human eye (B). The latter curves are also called cone fundamentals, and the graph uses the 10° data from Stockman and Sharpe (2000).

(Cornsweet, 1970). They can thus be interpreted as probability functions. In short: when photons of a stimulus arrive at the eye or digital camera, they get weighted by the three sensitivity functions. Each of the three resulting photon collections is added up to produce a signal, and the three signals are proportional to the area under the three curves.

The raw trichromatic signals generated by the camera are then processed inside or outside the camera (see Sections 3 and 4) to yield Red-Green-Blue (RGB) colour values that—when combined—ideally depict the colour as perceived in the real world by a human observer. Inside the HVS, the brain's post-retinal structures process the trichromatic retinal signals to yield a colour perception. Although both processing chains are vastly different, the relative proportion of the photons absorbed by the three cones/spectral response curves determines which colour gets perceived in both cases. For completeness, it is interesting to know that the brain encodes the retinal signals into three opponent signals: two chrominance signals and one achromatic channel representing luminance information (Valberg, 2005). Later in this paper, we will see why splitting colour into its chromaticity and luminance components is helpful.

2.3. Standard Observers and Colour-Matching Functions

Because it was impossible in the early 20th century to directly measure the response functions of the S-, M- and L-cones inside the human eye, visual colour-matching

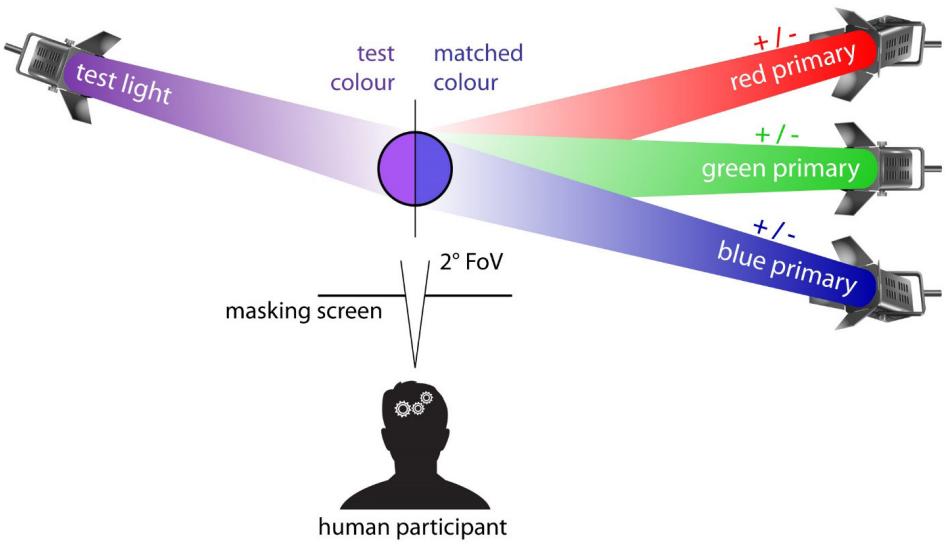


Figure 6. The setup of a colour-matching experiment.

experiments were conducted to derive a so-called CIE standard observer (H.-C. Lee, 2005). This CIE standard observer is a mathematical representation of the average colour vision of humans in the 380 nm to 780 nm range. Characterised by three curves called the Colour-Matching Functions or CMFs, the standard observer is a central pillar of present-day digital imaging and colour science.

As a matter of fact, two CIE Standard Observers exist: one initially determined in 1931 and another defined in 1964. Both were established with the help of healthy and young human participants asked to look at a white screen through a small aperture. Different test colours from the visible spectrum were projected onto that screen, thereby filling half of the screen (see Figure 6). At the same time, these participants could work with one Red, Green and Blue monochromatic light (i.e. a light source producing one wavelength and technically called a primary). The participants adjusted the amounts of these three lights/primaries on the other half of the screen until the additive mix of these three lights matched the projected test colour (see Figure 6).

This colour-matching process was repeated until most colours across the visible spectrum were covered. For some test colours, no match could be obtained by additively mixing the three lights. The only way to make a match was to remove one of the lights and add it to the test colour. When this happened, the primary was given a negative value. In 1931, the CIE published the three curves that resulted from averaging two such colour-matching experiments: one by David Wright (1928-29) and one by John Guild (1931). These curves became known as the CIE 1931 2° standard observer. The "2°" in the name comes from the fact that the participants had to look through a hole that allowed them to have a 2° Field of View (FoV) during the colour-matching experiment (see Figure 6). This 2° FoV was used because it was believed all cones were located in a small area of the retina called the fovea. Scientists later found out that the cones covered a larger area, spreading beyond the fovea.

That is why the CIE used the colour-matching experiments by Stiles, Burch and Speranskaya (Speranskaya, 1959; Stiles & Burch, 1959) to propose a second, 10° standard observer in 1964. Besides a larger FoV and more observers (49 versus the initial 17), the wavelengths of the Red, Green and Blue primaries also differed for these new tests: [645.16, 526.32, 444.44] nm versus [700.0, 546.1, 435.8] nm for the first experiments in the 1920s. So today, two standard observers exist: the CIE 1931 2 degrees observer and the CIE 1964 10 degrees observer (Figure 7A). Their CMFs are technically denoted CIE 1931 2° RGB CMFs $\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)$ and CIE 1964 10° RGB CMFs $\bar{r}_{10}(\lambda), \bar{g}_{10}(\lambda), \bar{b}_{10}(\lambda)$, respectively (Ohno, 2000). Slight adaptations to these CMFs have been proposed in the past two decades (CIE, 2006, 2015), but many engineers and scientists still favour the original ones.

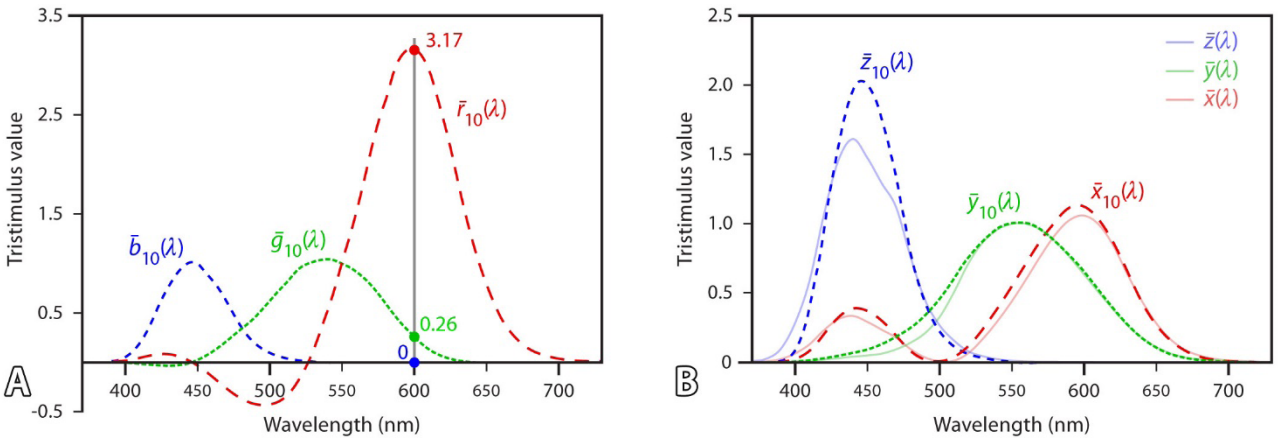


Figure 7. A) The CIE 1964 10° RGB CMFs; the tristimulus values for an orange hue of 600 nm are indicated. B) The CIE 1931 2° XYZ CMFs (modified by Judd (1951)) and CIE 1964 10° XYZ CMFs with primaries *X*, *Y* and *Z*. All datasets from Colour and Vision Research Laboratory (2021).

Both sets of RGB CMFs allow determining the relative quantity of the corresponding standardised primaries (i.e. the specific Red, Green, and Blue lights) needed to match a monochromatic test light of a particular wavelength. As an example, Figure 7 displays the 10° RGB CMFs and indicates that a combination of 3.17 parts of a 645.16 nm light, 0.26 parts of a 526.32 nm light and 0 parts of a 444.44 nm light source will produce a perfect match for the colour sensation of a 600 nm monochromatic test light. As such, most perceivable colours can be described by a known set of *RGB* primaries and three numbers corresponding to the amount of each primary needed to create that specific colour. These numbers are called the tristimulus values for that colour, and colour televisions, computer monitors, scanners and digital photo cameras exploit this trichromatic additive colour-mixing phenomenon.

However, Figure 7 also shows these RGB CMFs to have negative lobes. This results from the previously mentioned negative primary. The observer of the colour-matching experiment had to remove one primary from the mixture and add it to the monochromatic test light to achieve a match with the remaining two primaries (Wandell, 1995). Even though this behaviour is due to the similar spectral responses of the L- and M-cones (see Figure 5B), these CMFs are inconvenient for colorimetric applications because they create a physically non-achievable solution for RGB devices. For example, a digital camera should be negatively sensitive to particular wavelengths, meaning it should emit them.

The CIE performed a linear mathematical transformation of the initially calculated RGB CMFs to avoid negative values at all wavelengths. This yielded a new set of imaginary *XYZ* primaries and corresponding XYZ CMFs: CIE 1931 2° XYZ CMFs $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ and CIE 1964 10° XYZ CMFs $\bar{x}_{10}(\lambda), \bar{y}_{10}(\lambda), \bar{z}_{10}(\lambda)$, for a 2° and a 10° FoV (CIE, 2018). Figure 7B shows that both CMFs are slightly different, as the CIE 1964 10° XYZ CMFs also involve regions of the retina that are less densely packed with cones (Malacara, 2011). Note that the *X*, *Y*, and *Z* primaries are physically unattainable; they are purely mathematical constructs made up from the real CIE *R*, *G*, and *B* primaries using a matrix transformation. Even though no single natural person is probably exactly like any of the CIE standard observers (Nimeroff et al., 1961), and notwithstanding the drawback that the CIE XYZ CMFs use imaginary primaries because no physical matching lights can obtain these functions (Hung, 2006), these CMFs are still essential in all aspects of colorimetry and a vital element for understanding colour perception.

Finally, it is necessary to understand that the 10° cone fundamentals depicted in Figure 5B are an exact linear transform of the 10° XYZ CMFs (which also counts for the 2° curves). Via multiplication with a 3x3 matrix containing nine elements, one can quickly get one set of curves from the other and vice versa (Berns, 2019). The CMFs and cone fundamentals are thus duals of one another (Horn, 1984). The following section will explore how these CIE XYZ CMFs allow calculating the fundamental *XYZ* tristimulus values for any perceivable colour.

2.4. Expressing Colour: Colour Models and Spaces

In the digital world, colour is represented using global colour models and more specific colour spaces. Colour is often mathematically defined as a three-dimensional (3D) property. However, a point in a 3D space can be determined using many different Coordinate Systems (CSs). In cartography, a CS is determined by its dimensionality (i.e. the number of coordinate axes) and the attributes of these axes: their name, abbreviation, units, direction and order (Iliffe & Lott, 2008). As an example: a 3D Cartesian XYZ system using the meter. If any of these attributes changes, the CS changes.

The same goes more or less for colour (Verhoeven, 2016). Many CSs for colour exist, typically built upon three or four coordinate axes with a specific name, direction and order. Two well-known examples are the RGB and CYMK colour models. Whereas the RGB colour model is used in digital cameras and monitors to describe colour values with an *RGB* triplet, a 4D system like CMYK is used in printing to describe colour via specified amounts of a Cyan, Magenta, Yellow and blacK/Key primary. However, a specific *RGB* triplet such as *R*:130 – *G*:110 – *B*:255 does not define a particular colour but only indicates the ratio of the three components used. Although 255 is the maximum value for each *R*, *G* and *B* channel in an 8-bit image, a colour model does not specify how 'vibrantly blue' this maximum should

be. In other words: a colour model is a mathematical system without describing how these values should be interpreted in terms of real-world, quantified colours. Colour models such as RGB, CMYK, HSV (Hue, Saturation, Value), and HSL (Hue, Saturation, Lightness) are, therefore, said to be relative.

To accomplish an unambiguous description of colours, one needs colour spaces. And again, cartography functions as a handy comparison. Cartographers use the term Coordinate Reference System (CRS) when a CS is fixed to a specific object. For example, a 3D Cartesian XYZ coordinate system with units in meters related to the Earth by fixing the CS' scale, orientation and the position of its origin. Similarly, the CIE XYZ colour space is associated with the operating principles of the HVS, and all other colour spaces are related to the XYZ colour space. Once coordinates are given in any colour space, one knows unambiguously which colour these coordinates represent. To map from relative *RGB* values to absolute CIE *XYZ* values, RGB colour model-based colour spaces are defined by three primaries, a specific CIE illuminant and a gamma value.

Defining a specific primary for each axis is similar to choosing the three monochromatic lights in the colour-matching experiments. From a conceptual point of view, establishing a primary determines the greenest green, the bluest blue and the reddest red, represented by the highest value on the *R*, *G*, and *B* axes. Since these primaries can be freely defined, each colour model has almost unlimited colour spaces. In other words: a particular colour space is just one possible, absolutely defined instance of the more general colour model.

The best-known RGB colour spaces for digital photographs are sRGB, Adobe RGB (1998) and ProPhoto RGB (also known as ROMM RGB), all with three unique primaries. Figure 8 provides the data for these three colour spaces and compares all three in the CIE *xy* chromaticity diagram. The chromaticity coordinates (*x*, *y*) are defined by:

$$x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z} \qquad (1)$$

so that a colour's chromaticity can be described irrespective of its luminance (Hunt, 2004). The three points that form a colour space triangle correspond to the chromaticity coordinates of that colour space's primaries (Figure 8 also gives these values). The area in the triangle is known as the colour gamut and encompasses all the colour values one can create by mixing those three primaries. From Figure 8, one can infer that the sRGB space has a relatively limited gamut (which corresponds to those of most monitors). In contrast, Adobe RGB (1998) and ProPhoto RGB have much wider gamuts, even though all three are based on the RGB colour model. In other words, the ProPhoto RGB colour space can represent more colours than the Adobe RGB (1998) space, which can store a broader range of colours than the sRGB colour space. Since the horseshoe-shaped chromaticity diagram represents all chromaticities visible to a standard human observer, the gamut of human vision is clearly not a triangle. The outer curved diagram boundary is known as the spectral (or monochromatic) locus, because monochromatic lights can generate these chromaticities. The diagram includes the wavelengths of those monochromatic lights for completeness.

Combining three primaries with a particular CIE illuminant and gamma value (concepts detailed in Sections 2.5 and 4.6, respectively) thus effectively maps *RGB* values to the CIE XYZ reference colour space. As mentioned, *X*, *Y*, and *Z* primaries are purely mathematical entities, but they allow us to uniquely create and describe all possible colours that an average human perceives. The values mentioned above (*R*:130 – *G*:110 – *B*:255) expressed in the sRGB colour space correspond to *X*: 32.8 – *Y*: 23.1 – *Z*: 97.3, describing a purple-ish tone (see Figure 8 on the right). A more blue-ish tone results when the same *RGB* values are expressed in the ProPhoto colour space. As one would expect, its objective description in CIE 1931 2° XZY values changes to *X*: 29.8 – *Y*: 24.2 – *Z*:82.5. However, it is the reverse conversion which makes the XYZ colour space so powerful: given a colour expressed in terms of CIE *XYZ* tristimulus values, its corresponding value in any possible colour space like sRGB or Adobe RGB (1998) can be computed. And because the CIE XYZ colour space is device independent, it functions as connection space between input and output devices. In other words, devices like scanners and digital cameras have

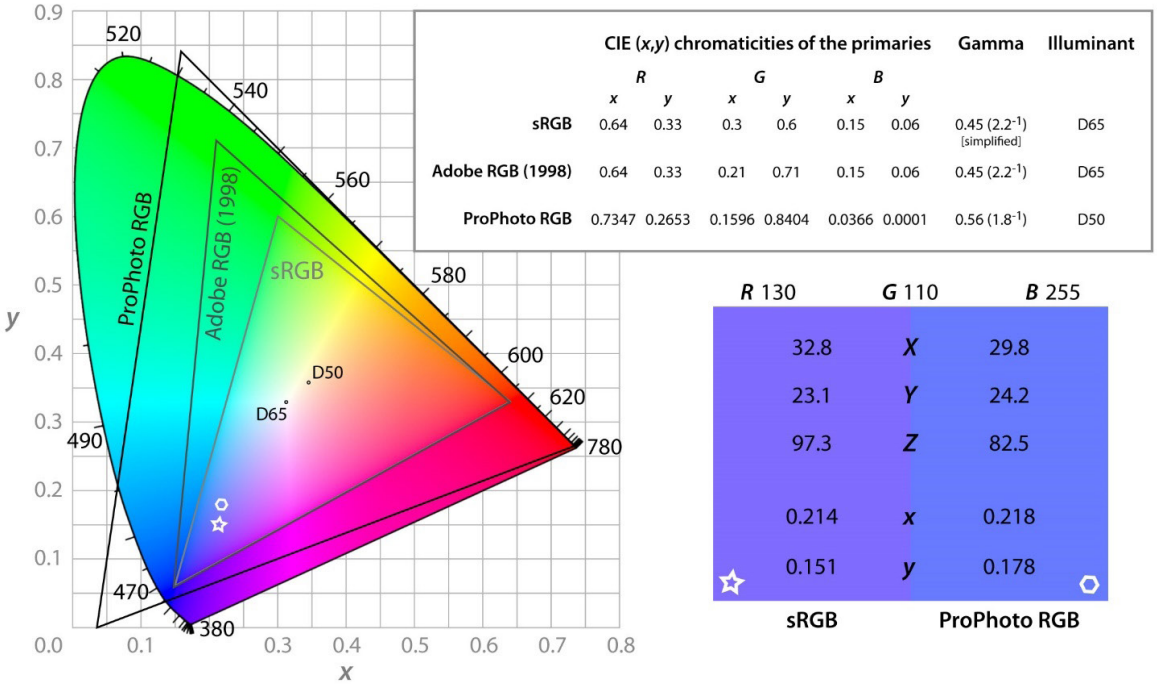


Figure 8. On the left, the primaries and bounds of three RGB colour spaces are drawn on top of the *xy* chromaticity horseshoe diagram that represents the limits of human colour vision. The top right inset provides all colour space-specific data. The bottom right illustrates how identical *RGB* values (130, 110, 255) result in different colours (and CIE 1931 *XYZ* or 1931 *xy* values) when expressed in the sRGB or ProPhotoRGB colour space. Both colours are indicated in the chromaticity diagram.

to find a way to map their device-specific *RGB* values to corresponding *XYZ* values. As we will see in Section 5, this is one of the most critical steps in producing adequately accurate colour values with a digital camera.

However, not just the values generated by a digital camera must, at a certain point, be expressed in *XYZ* coordinates. The XYZ colour space is omnipresent in colorimetry, and everything—from illumination sources to stimuli—can be described in terms of *XYZ* coordinates. This is the topic of the following two sections.

2.5. Characterising Illumination

Any light source can be expressed by its Spectral Power Distribution (SPD): a series of numbers quantifying the amount of each visible wavelength (between 400 nm and 700 nm) included in the light source. Blue disco light emits mainly wavelengths between 400 nm and 500 nm and has almost no output from 500 nm to 700 nm. On the other hand, a camera flashlight contains all visible wavelengths in varying proportions. Plotting these numbers yields a graph like the ones in Figure 9A. Both SPDs were measured by the INDIGO staff with a Sekonic C-7000 SPECTROMASTER portable handheld spectrometer. The graph shows two different illuminations of a graffito just before photographing it. These curves always combine direct solar radiation and diffuse skylight. The former might be zero on very cloudy days or when a graffito is in the shadow.

The SPD of any light source can thus be measured with a spectrometer. However, in colorimetric applications, one usually does not work with just any light source but with so-called CIE standard illuminants: quantified and standardised illumination sources. The CIE has described several standard illuminants (International Organization for Standardization, 2022) based on physically obtainable

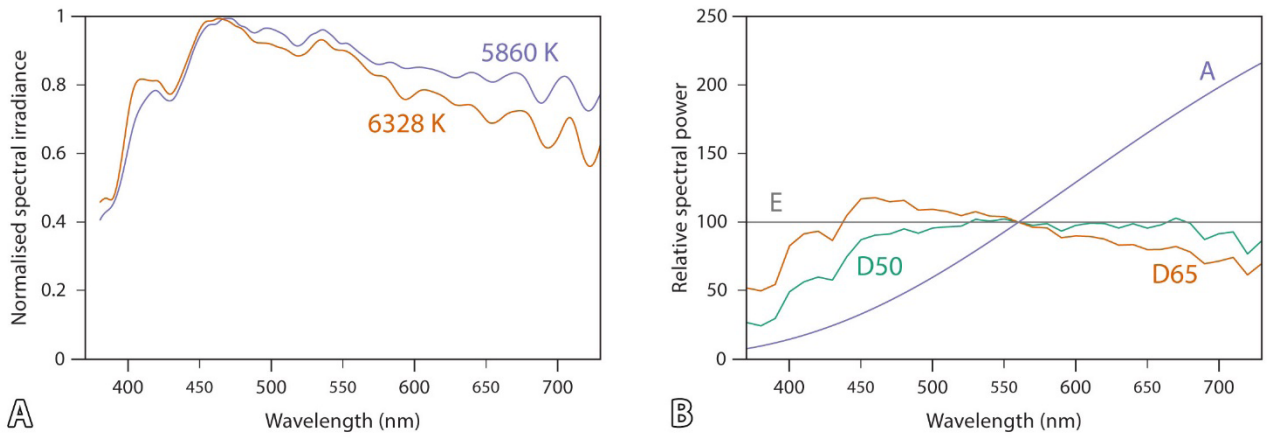


Figure 9. Inset A depicts two spectral power distributions of the outdoor illumination just before photographing a graffito. Inset B represents the SPDs of the CIE D50, D65, A and E standard illuminants. The SPDs of the latter four come from the ISO/CIE 11664-2:2022 standard (International Organization for Standardization, 2022).

light sources or their statistical representations (see Figure 9B). Of those, the CIE D65 is the most widespread. The "D" part of the name indicates that D65 is a Daylight illuminant, part of a family that is representative of the various phases of daylight (Judd et al., 1964). Its SPD (see Figure 9B) aims to roughly correspond to the average western European midday light (comprising both direct sunlight and diffuse skylight). The "65" part refers to this illuminant's correlated colour temperature of approximately 6500 degrees kelvin (CIE, 2018).

Summarising metrics like the Correlated Colour Temperature (CCT) exist since it is not always possible nor practical to consider the illumination's entire SPD. The CCT characterises the illumination's dominant colour with a temperature reading on the kelvin scale. The kelvin value (symbolised by K) refers to the temperature at which a theoretical object (called a blackbody) must be heated so that its SPD generates the same colour experience as the SPD of the illuminant. A blackbody is a theoretical object that absorbs all incident electromagnetic radiation; it is black. Any electromagnetic radiation originating from a blackbody is emitted solely as a function of its temperature, as described by Planck's law (Walker, 2004). Figure 10 depicts the temperature-dependent output of a blackbody, and although iron is not a blackbody, it pays off to visualise heated iron to understand the curves in Figure 10 better. When heated to 700 K (i.e. 700 K – 273.15 = 426.85 degrees Celsius or 427 °C), the piece of iron is characterised by a deep red glow. The iron shall radiate brighter, reddish-orange light when raising the temperature to about 1000 K (727 °C). Increasing the temperature to its boiling point at 3134 K (2861 °C) yields a brilliant orange-yellowish light. In other words: the SPD of the emitted light is only a function of the iron's absolute surface temperature, hence the term Colour Temperature (CT). Since a blackbody is an idealised object and illumination sources are not ideal blackbody radiators, the entire SPD of these sources cannot be described solely as a function of temperature. However, their colour or, more accurately, their chromaticity can. This gave birth to Correlated CT or CCT: the kelvin temperature at which a blackbody SPD yields the same chromaticity experience as the illumination source under consideration (Borbély et al., 2001).

Being a one-dimensional metric, CCT is a convenient but limited way to characterise and summarise many illumination sources. Although an in-depth discussion of its (dis)advantages is beyond the scope of this paper, one of the problems with CCT is that different SPDs can

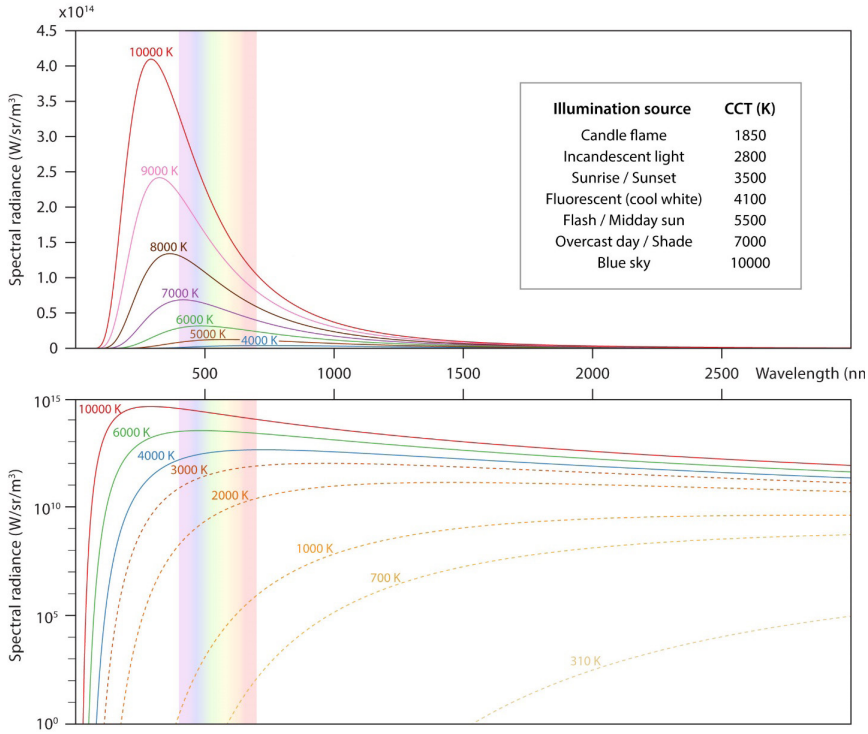
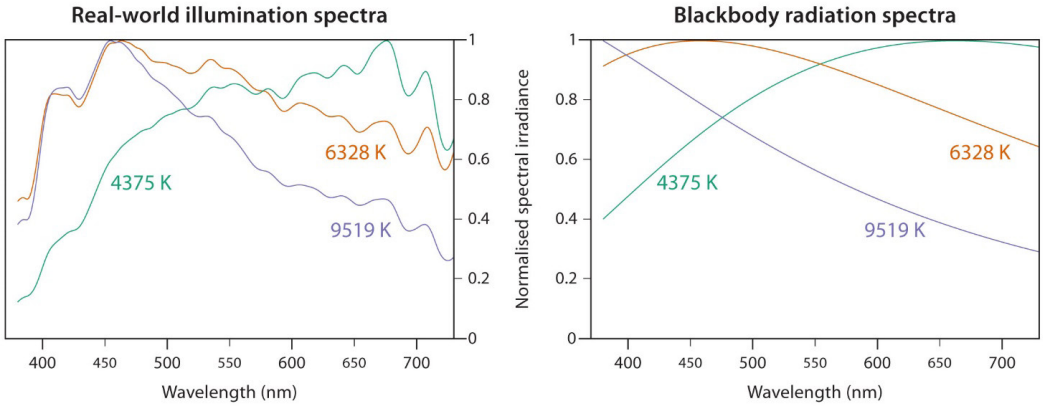


Figure 10. The blackbody radiation output for different kelvin temperatures (K). The lower graph is a semi-log plot with the spectral radiance displayed on a logarithmic scale to better visualise the radiation emission at lower temperatures. In this way, it is possible to show that a blackbody with the temperature of an average human body at 37 °C (or 310 K) does not emit any visible light. Only around 700 K (or 427 °C) gets visible red light emitted. Dashed lines in the lower, semi-log plot indicate blackbody radiation at temperatures not depicted in the upper plot. Blackbody radiance curves of a specific temperature share the same colour in both plots. The inset provides the CCT values of common illumination sources.

result in an identical CCT value. Figure 11 illustrates this. The left side displays three real-world daylight SPDs and their CCTs. These SPDs—measured with a Sekonic C-7000 SPECTROMASTER spectrometer—are normalised to their maximum value to ease comparison. Although the Sun is close to an ideal blackbody, these real-world daylight SPDs are pretty different from the SPDs of a blackbody heated at those temperatures (Figure 11, right side), mainly because solar illumination is altered by various absorptions via atmospheric gasses.



0.9711	1.0000	0.6706	0.3676	0.3785	4375 K	0.9960	1.0000	0.7305	0.3653	0.3668
0.9533	1.0000	1.0661	0.3157	0.3312	6328 K	0.9692	1.0000	1.0949	0.3163	0.3264
0.9380	1.0000	1.4105	0.2801	0.2986	9519 K	0.9719	1.0000	1.4565	0.2835	0.2917
X	*Y*	*Z*	*x*	*y*		*X*	*Y*	*Z*	*x*	*y*

Figure 11. The SPDs and CCTs of three outdoor illumination scenarios are displayed on the left, while the right side depicts the blackbody spectra at the same kelvin temperatures. The tables below both plots provide the CIE 1931 2° *XYZ* and *xy* coordinates of each SPD.

Instead of CCT, illumination SPDs are also characterised by their *XYZ* values (see Figure 11). To obtain these *XYZ* values, the illumination's SPD is multiplied by the CMFs of a CIE standard observer. Then, all the visible wavelength values are summed for each of the three CMFs to obtain the CIE *XYZ* tristimulus values for that incoming stimulus (this multiplication and summation combination is known as integration; see also Figure 13, but without the reflectance.). The resulting *XYZ* coordinates are said to represent the white point of the SPD. In other words, expressing the colour of an illumination source as an *XYZ* coordinate triplet tells us how much of the imaginary *X*, *Y* and *Z* primary is needed to produce that colour.

Despite being purely mathematical, the CIE XYZ colour space features a few unique properties. For instance, the *Y* value represents the luminance data generated by the HVS, which is approximately the physical counterpart of the perceptual quantity brightness (CIE, 2018). That is why the white point's *XYZ* coordinates are usually divided by *Y*, so *Y* always equals 1. Because the *Y* value represents the luminance of a colour, the resulting *X* and *Z* coordinates characterise the SPD's white point without considering its luminance. Having two coordinates (*X*, *Z*) instead of one (like CCT) offers advantages. For instance: CCT as a principle relies on glowing hot objects. However, there is nothing that glows green or magenta. So if an SPD of a given illumination source would be perceived as a bit greener than the blackbody radiation at that temperature, this can be expressed by the *XZ* white point coordinates but not by the CCT. Table 1 provides some white point coordinates of standard CIE illuminants. Via equation (1), these white points can also be expressed as chromaticity coordinates (*x*, *y*).

Four CIE standard illuminants

Illuminant	White point CIE 1931 (*x*, *y*)	White point CIE 1931 (*X*, Z), *Y* = 1	CCT (K)	Represents
A	0.4476, 0.4074	1.0985, 0.3558	2855.5	tungsten filament lamp
D50	0.3457, 0.3585	0.9642, 0.8250	5001	average daylight around 5000 K
D65	0.3127, 0.3291	0.9504, 1.0886	6503	average daylight around 6500 K
E	0.3333, 0.3333	1, 1	5455	equi-energy radiator

Table 1. Some summarising data on four CIE standard illuminants. The CCT values for A, D50 and D65 are from the ISO/CIE 11664-2:2022 standard (International Organization for Standardization, 2022). The white points were computed with the CIE 1931 2° XYZ CMFs, and the standard illuminants' SPD tabulated in the same ISO standard.

Finally, it is essential to know that the 2° and 10° XYZ CMFs (see Figure 7B) have arbitrary units because they were normalised to yield identical *XYZ* values for a spectrally flat stimulus (i.e. an equi-energy stimulus). That is why the equi-energy CIE illuminant E has [1, 1, 1] as *XYZ* white point (see Table 1). Illuminant E is, therefore, called the reference white of the CIE XYZ colour space. It was already mentioned above and indicated in Figure 8 that any RGB colour space has a reference or native white (like D65 for sRGB). In each case, the SPD of this illumination results in a maximum identical value for each axis (i.e. [1, 1, 1]), thus representing white. For example, the stimulus in Figure 19C equals the reference white SPD for a Nikon D700.

2.6. Measuring Colour

Colorimetry, or the science of colour measurement, intends to express colours quantitively based on colorimetric standards (Hunt & Pointer, 2011). Since 1931, the CIE has provided several systems for that purpose, like the CIE XYZ colour space mentioned before. Since the text already explained the concepts of SPD, standard illuminant and their characterisation via CIE *XYZ* coordinates, it becomes now straightforward to explain colour measurement.

To express any object's colour as *XYZ* coordinate triplet, one needs three elements again (see Figure 12): the light source or illuminant, the object or sample and an observing instrument. The whole principle is identical to the one explained in Figure 4, with the constraints that the

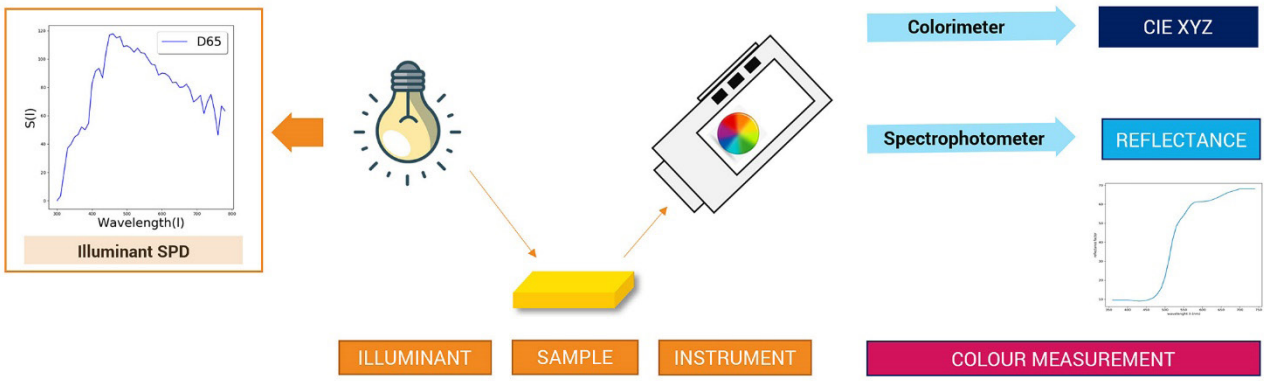


Figure 12. Colour measurement: an interplay between illuminant, sample and instrument.

light source is usually a CIE standard illuminant and the instrument a CIE standard observer. Once the light of the (standard) illuminant hits the object, a fraction of the entire SPD gets reflected—which, as said before, depends on the object's physical and chemical structure. Like the output of a light source, the spectral reflectance of an object can be measured and graphed. This graph is called the spectral reflectance curve; it quantifies which wavelengths get reflected and by how much (see Figure 12 on the right or Figure 13 in the middle).

Finally, a colorimeter will produce *XYZ* tristimulus values. Colorimeters have internal filters and processing algorithms, so they can mimic the CIE 1931 2° or CIE 1964 10° XYZ CMFs and integrate the stimulus over them. Colorimeters are great devices, but they are not as accurate as spectrophotometers. In addition, the operating principles of colorimeters and spectrophotometers are dissimilar (Berger-Schunn, 1994). A colorimeter solely observes the incoming stimulus and provides the *XYZ* coordinates to characterise the colour of that stimulus. As such, they are often used to characterise computer monitors. In contrast, a spectrophotometer has a standard light source integrated, usually a simulation of the D65 illuminant (even though they may include additional standard illuminants). This light source illuminates the sample, which partly reflects a portion back into the spectrophotometer. Since the instrument knows the exact SPD of its integrated light source, it can compute the object's reflectance from the stimulus it receives. In that way, spectrophotometers can be placed on any part of an object to accurately measure that sample's spectral reflectance. This spectral reflectance curve can be considered a unique spectral fingerprint of that object part. At that stage, the instrument has all the pieces to generate the *XYZ* values of the sample: it knows the SPD of the illuminant and the sample's spectral reflectance. Both can now be multiplied into a stimulus; the latter gets integrated over a CIE standard observer to yield 2° or 10° CIE *XYZ* tristimulus values (see Figure 13) (CIE, 2018; Hunt & Pointer, 2011).

Figure 12 depicts these two different pathways: either a colorimeter computes CIE *XYZ* values directly, or the latter are obtained mathematically after first calculating the exact spectral reflectance of the sample. The problem with a colorimeter is that the illuminant can remain entirely unknown. And since the stimulus is a function of the illuminant, a sample will feature different *XYZ* values according to the light source. That is why *XYZ* values for colour reference targets (like the well-known ColorChecker series by X-Rite, now produced by Calibrite) are always defined for a particular illuminant. And that is why well-defined standard illuminants are essential. If a non-standard illuminant were used to express the *XYZ* values of an object, these values would be worthless without considering the entire illumination SPD. Agreeing on some standardised illuminants makes communication of colour values thus more straightforward. Moreover, using a spectrophotometer to directly measure the object's spectral reflectance, the *XYZ* colour values can be computed

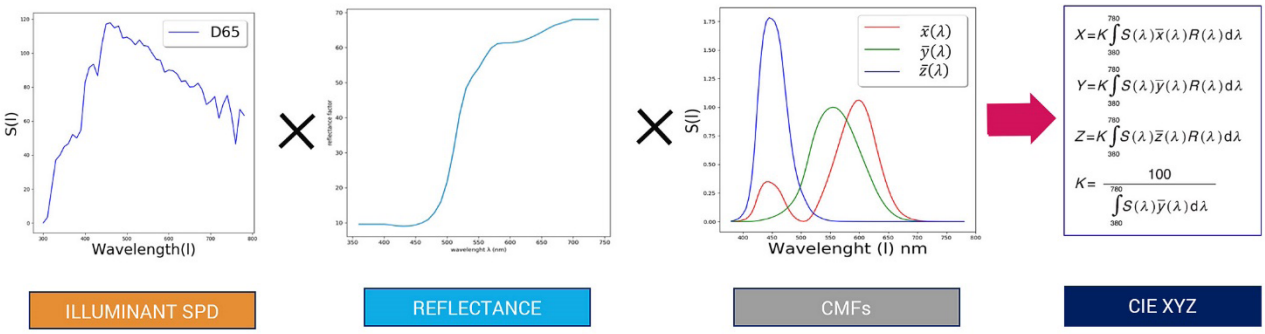


Figure 13. Scheme for the computation of CIE *XYZ* values from reflectance data.

for any known SPD. That is why spectrophotometers are recommended, even though they are more expensive than colorimeters.

At this point, it is essential to address if one can rely on a conventional digital camera for colour measurement, thus effectively simulating a colorimeter or spectrophotometer. Using the reasoning established before, one can state that every pixel of a digital graffito photo results from an illumination's SPD interacting with the graffito's reflectance at a specific point, thus yielding a spectral stimulus which gets integrated over the camera's spectral sensitivity curves. In this chain, it is only straightforward to measure or estimate the SPD of the illuminant with a spectrometer (see the left graph of Figure 13). So how can one get accurate colour data from the raw pixel values registered by the camera? Sections 3 and 4 will cover this topic.

3. The Birth and Storage of Image Pixels

3.1. Image Sensors: A Collection of Photosites

Analogue signals are continuous and exist in the tangible world as functions of space and time. Digital signals are found inside computers and are merely a collection of discrete states. Once an analogue signal is digitised, perfect clones become possible. The digital product is, however, always an approximation of the analogue reality because this numerical translation is accomplished by the processes of sampling and quantisation (both explained below). The analogue signal digitally captured by photography or any other form of optical imaging is the continuously varying electromagnetic radiation reflected or emitted by the scene. In the case of traditional photography, a digital image is generated by converting the visible portion of radiant energy into an electrical output signal which is then digitised.

All cameras comprise optical elements such as lenses and filters that gather electromagnetic radiation and focus it onto their imaging sensor. For applications like conventional photography, these imaging sensors consist of a two-dimensional array of individual photon-sensing sites or photosites (Figure 14). The photodetector is the key component of such a photosite, as it collects light during the exposure time. Nearly all imaging sensors have one photodetector per photosite, although Foveon's X3 image sensor features three detectors per site (Lyon & Hubel, 2002). Depending on the sensor design, the individual photosites may contain more or less circuitry, and the photon-receiving surface area of the photodetector may be smaller or larger. Throughout the years, academia and industry have proposed diverse photosite arrangements and photodetector designs to achieve specific performances (e.g. increase the image's spatial resolution or optimise the spectral sensitivity of the imaging sensor).

Most photo cameras feature an imaging sensor in which one photodetector contributes one effective pixel to the image. For instance, a 24-megapixel digital camera has an imaging sensor built-up of at least 24 million photosites/photodetectors distributed in rows and columns (for example, 6000 columns by 4000 rows). 'At least' is essential here since additional photosites handle tasks like dark

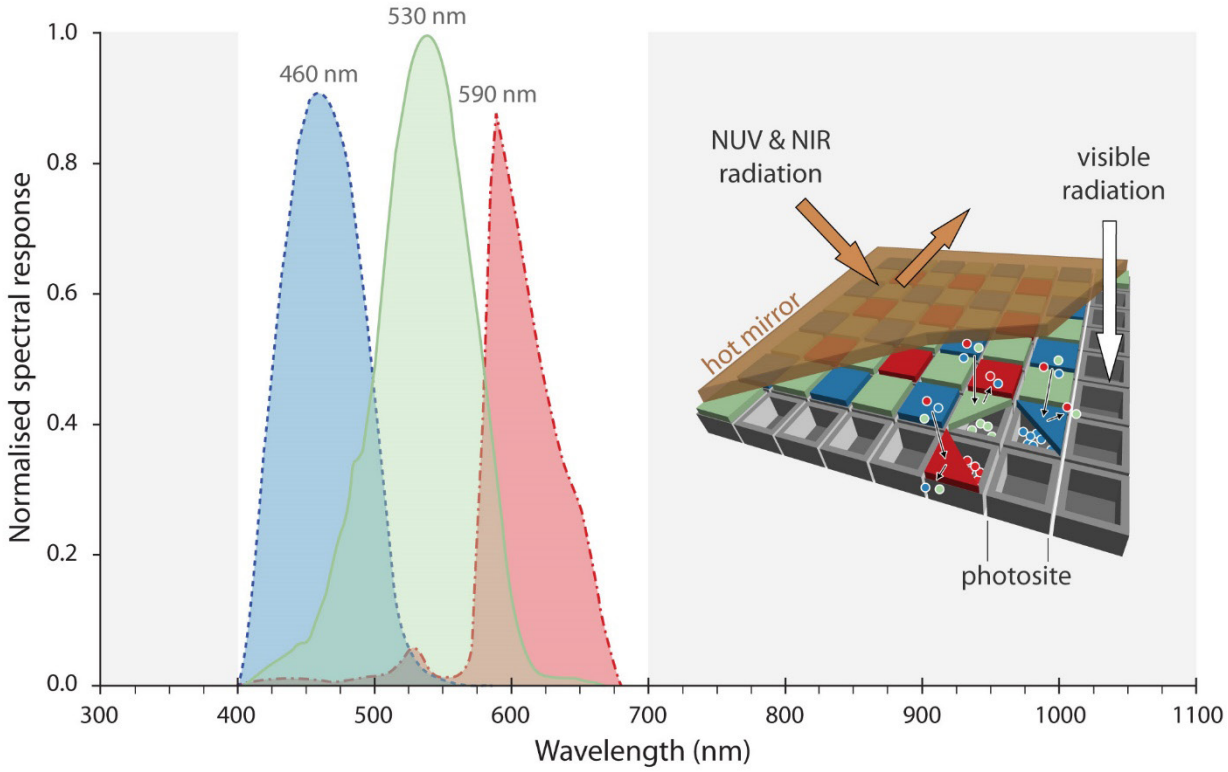


Figure 14. The relative spectral response curves (also known as the spectral sensitivity functions), sensor element layout and working principle of a typical off-the-shelf digital photo camera with a Bayer-filtered imaging sensor. NUV and NIR mean Near-UltraViolet and Near-InfraRed, respectively.

signal correction (see Section 4.1) and white balancing (see Section 4.2).

3.2. Sampling: Collecting Photons

The fundamental building blocks of any digital image are called pixels (Billingsley, 1965) or pels (Schreiber, 1967), coined terms for picture elements. To create these pixels, an imaging sensor inside a conventional photo or video camera collects incoming photons over the area of every photosite. However, not all photons are collected, only from the incident visible electromagnetic radiation (i.e. light). To achieve this, the imaging sensor features a filter on top (the so-called hot mirror; Figure 14) that blocks non-visible electromagnetic radiation. Without this filter, the sensor would also detect near-ultraviolet and near-infrared radiation, making it impossible to render the scene colours as humans perceive them.

As Figure 14 depicts, none of the photosites captures the entire visible spectrum. A mosaic of thin, coloured filters ensures that every photosite only gathers one particular part of the incident light. This Colour Filter Array (CFA) comes in various designs, but digital cameras mainly use a so-called Bayer pattern which features twice as many green filters as blue or red ones (Bayer, 1975). So per photosite, photons are solely gathered in one of the three 100-nm-wide spectral bands, being the Blue waveband (with wavelengths from 400 nm to 500 nm), the Green (500 nm to 600 nm) and Red (600 nm to 700 nm) spectrum (Figure 14).

This process lasts as long as the exposure lasts (e.g. 2 s or 1/250 s). All the absorbed photons generate an electrical charge in every photodetector. After the exposure, every single detector's charge—which is linearly proportional to

the amount of incoming radiation—is read out. It represents a sample of the visible electromagnetic energy originating from the imaged scene (see Figure 15). In other words, photographic pixels are created by sampling the scene in the spatial, spectral and temporal dimensions. A digital photograph is thus never a very accurate reproduction (in absolute terms) since its pixels represent averages in space, spectral range, and time.

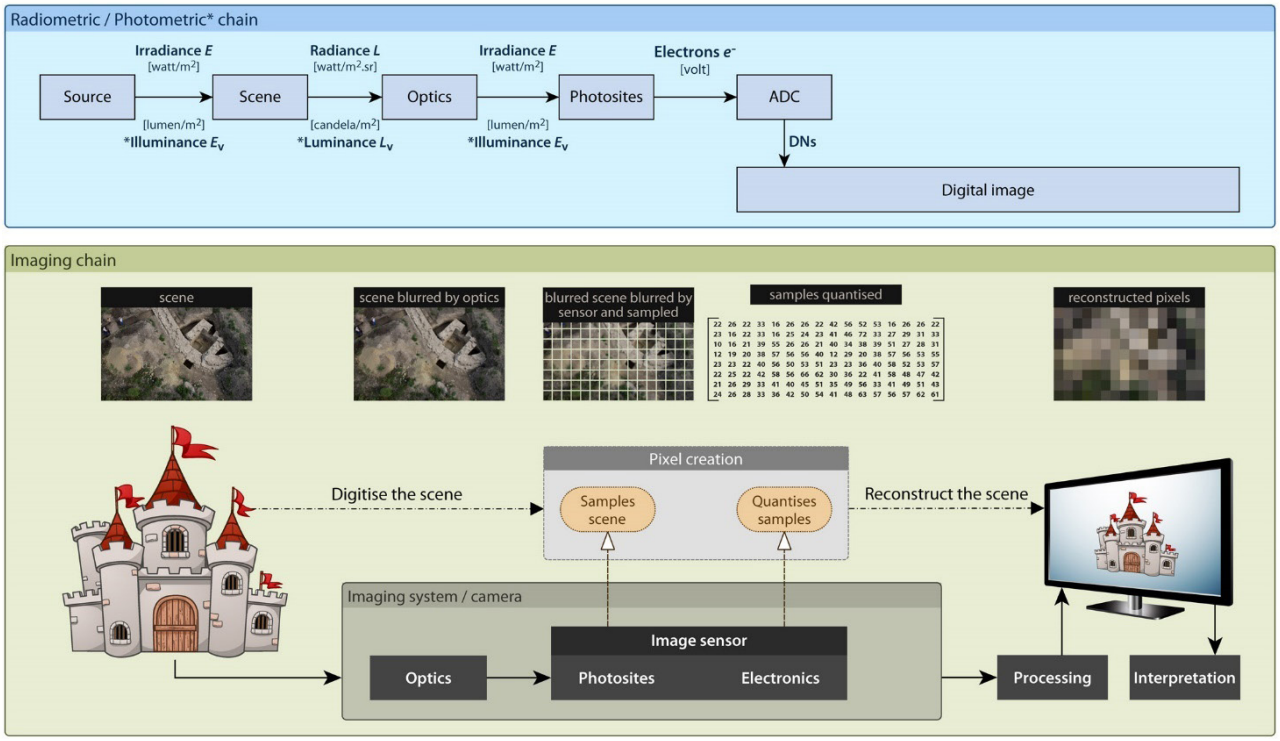


Figure 15. The complete imaging chain: a digital camera digitises an analogous real-world scene which gets reconstructed on a monitor. The lower part of the illustration depicts the overall pipeline (from left to right), including the pixel creation stage. On top, the imaging chain is broken down into its radiometric and photometric components. The latter quantities and units apply when imaging visible electromagnetic radiation.

3.3. Quantisation: Handing Out Digital Numbers

These sample values (i.e. the generated charges) have to be mapped onto a discrete set of numbers by a process called quantisation. In their most raw form, image pixels are thus created by quantising every photosite-specific sample to a discrete Digital/Data Number (DN) by the Analogue-to-Digital Converter (ADC). The total range of tones or quantisation values an ADC can create is called the tonal range. The ADC's bit depth determines the image's tonal range: quantisation with *N* bits rounds all possible charges to these 2^N values. For example, a conventional 12-bit ADC can discriminate 2^{12} or 4096 tones; every pixel gets one of those 4096 possible discrete DNs, 0 for the lowest and 4095 for the highest possible charge.

Raw photo pixels are thus sampled and quantised versions of a continuous analogue spectral signal (Verhoeven, 2018). These samples are determined by a pair of pixel coordinates (*r* and *c*, indicating row and column) and one specific value (the DN). An array of these pixels is called a digital image,

mathematically represented as an $M \times N$ matrix of numbers, M and N indicating the image dimensions in pixels. However, at this stage, every pixel only has one value, one DN. Section 4 will cover all the processing steps that are needed to turn these raw DNs into a colour photograph where every pixel contains one digital number for the Red, one for the Green and one for the Blue spectral band (which explains why photographs in their final, processed form are often called RGB images). So, the processed photos a camera generates from the initially captured raw pixels can be represented by O matrices of $M \times N$ elements, O being the number of spectral bands. O is three for an RGB image and one for a greyscale image.

3.4. RAW Versus JPEG and TIFF

The last paragraph mentioned two types of photographs: photos containing raw pixel values and fully processed RGB output photos. This dichotomy or 'choice' is also reflected in many digital cameras, which can usually produce and store both image types:

1. An image which holds per pixel the single DN captured via its corresponding photosite. This image is typically referred to as the RAW photo or RAW file. RAW is not an acronym. It only signifies raw or minimally processed image sensor data with pixel values that are linearly related to the incoming radiation in the Red, Green or Blue visible spectral band. RAW can thus be considered the only scientifically justifiable file format (Verhoeven, 2010). However, the RAW format is not all roses. Even though most dedicated digital cameras support saving RAW files, they all have a manufacturer-specific structure and extension, like *.NEF for Nikon, *.RAF for Fuji, *CRW or *.CR2 for Canon and *.GPR for GoPro. Adobe also launched its open-source Digital NeGative (or *.DNG) format in 2004, attempting to standardise the RAW file format. However, most manufacturers refrain from implementing it. In addition, RAW data need many processing steps to end up with the second image type: a normal-looking photo.
2. A highly-processed viewable image with pixels nonlinearly related to the captured stimulus. This image is usually expressed in the sRGB colour space and saved as a JPG/JPEG-compressed image or TIFF file. When talking about a photo, this viewable type of image is meant. Even though some dedicated cameras (and smartphones) might not offer the option to save the RAW image, the latter always forms the basis to yield a pleasing output photo.

In addition to the pixel values or DNs that encode the real-world scene, both image types contain Exif (Exchangeable image file format) metadata. These metadata describe image acquisition parameters (such as the serial number and model of the camera, the aperture, focal length, shutter speed, possible flash compensation, and the date plus time of photo acquisition) in mandatory, recommended and optional tags stored in a separate segment of the file (Camera & Imaging Products Association, 2010-2019). Suppose the camera is GNSS (Global Navigation Satellite System)-enabled. In that case, tags can also hold the latitude, longitude and altitude of the camera's geographical location. All these Exif-defined tags are created by the camera and stored simultaneously with the DNs in the image file, making it possible to analyse them afterwards.

4. 'Developing' or Rendering Raw Digital Numbers

Many processing or 'digital development' or 'rendering' steps are involved in producing a normal-looking output photo from a RAW image. These steps either happen inside the digital camera or are executed afterwards via dedicated RAW conversion computer software. Despite the variety of RAW processing pipelines detailed in the scientific literature, there is agreement on the central processing stages (Karaimer & Brown, 2016, 2019; Ramanath et al., 2005; Sumner, 2014). Within the scope of project INDIGO, we have developed the software package COOLPI. COOLPI, or the COlour Operations Library for Processing Images, is open-source and freely available on GitHub: https://github.com/GraffitiProjectINDIGO/coolpi. The software can also be installed directly from the Python Package Index repository or PyPi (https://pypi.org/project/coolpi) using pip (by running "pip install coolpi" on the system shell).

COOLPI relies on rawpy (https://pypi.org/project/rawpy), a Python wrapper for the LibRaw library (https://www.libraw.org), to facilitate many of the well-known RAW processing steps, but with adaptations at several stages to prioritise colour accuracy (see also Molada-Tebar et al. (2018)).

Generally, one can distinguish six main steps to convert a RAW image file into a viewable RGB photo expressed in a colour space like sRGB. Some implementations might change the sequence of these steps, but the order in which they are detailed below is relatively standard. Figure 16 summarises all these steps and related terminology. It can, therefore, function as a visual guide for the rest of this paper.

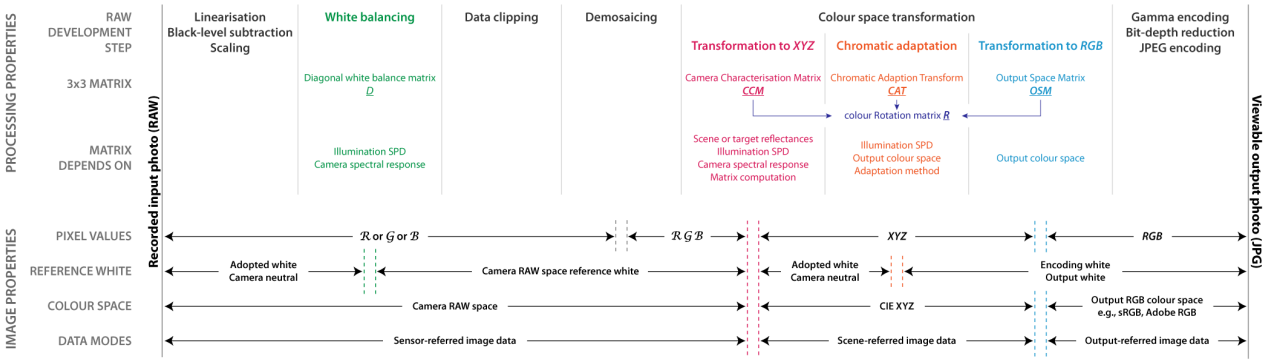


Figure 16. The figure outlines the entire RAW development chain covered in this paper. All matrix operations directly influencing the final colour are indicated in a chromatic tone. The factors on which those matrix operations depend, and the colour space or white point transformations they effectuate, are also indicated.

4.1. Linearisation, Black-level Subtraction and Scaling

Although the sensor's response to the incoming illumination is mainly linear (Chakrabarti et al., 2009), some cameras introduce a non-linear operation to compress the data. For example, many Sony cameras of the NEX and ILCE series feature a 14-bit ADC, but these 14-bit data get compressed to store them as 11-bit data, thereby reducing the file size considerably. Since all subsequent RAW development steps expect linear data, it is necessary to correct any non-linearity from the start (Ramanath et al., 2005). This linearisation step is straightforward: the linearisation curve embedded in the metadata of the RAW file gets applied to the stored data (from now on indicated as $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$), thereby decompressing the RAW data back to its original bit-depth. For example, Figure 17 shows how the Nikon D70's linearisation curve maps the initially stored but compressed data with a limit value of 683 back to the 12-bit maximum value of 4095. In other words: the original RAW data in the D70 image are compressed by the camera into 9.4 bits (i.e. $\log_2 683$). Via the linearisation curve, the compressed data are 'unpacked' to 12 bits (i.e. DNs from 0 to 4095). The Nikon D5600 linearisation curve shown in Figure 17 indicates that the 14-bit RAW data (i.e. DNs from 0 to 16383) are compressed into 12 bits. Figure 18 visualises the effect of this curve. Figure 18A shows the non-linear version of the RAW file. Halfway through the image's width, B depicts the linearised version.

A black-level correction to compensate for dark current accompanies this step. Dark current is a signal generated even when the sensor is not illuminated. In other words: pixels which should be perfectly black still have some value. This value increases with exposure time and is temperature dependent. The warmer an image sensor, the more dark current gets produced. To render black as truly black, the lowest value in the RAW file (or some default value) is considered pure black and gets subtracted from every image pixel (Figure 18C).

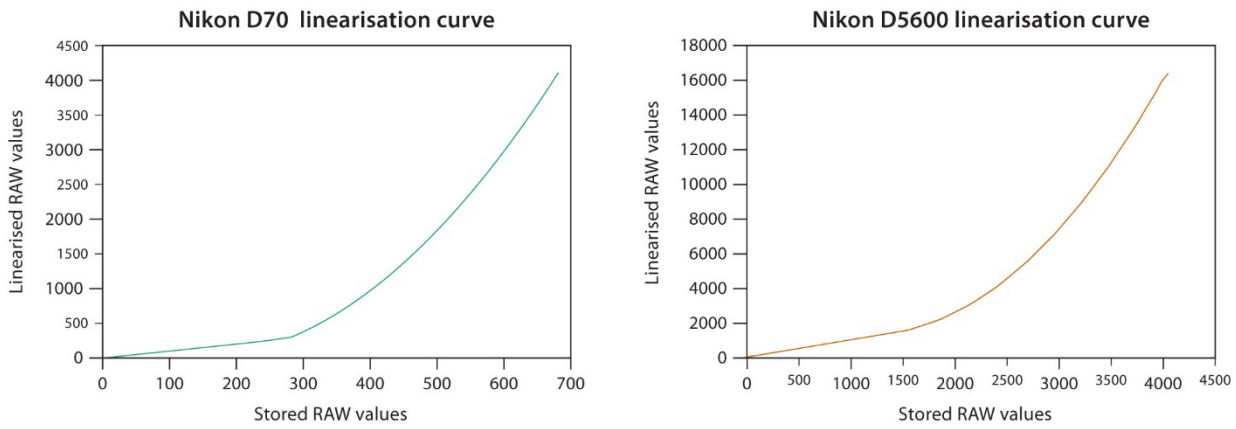


Figure 17. The linearisation curves of the Nikon D70 and Nikon D5600. These curves reside in the Exif metadata of the RAW files.

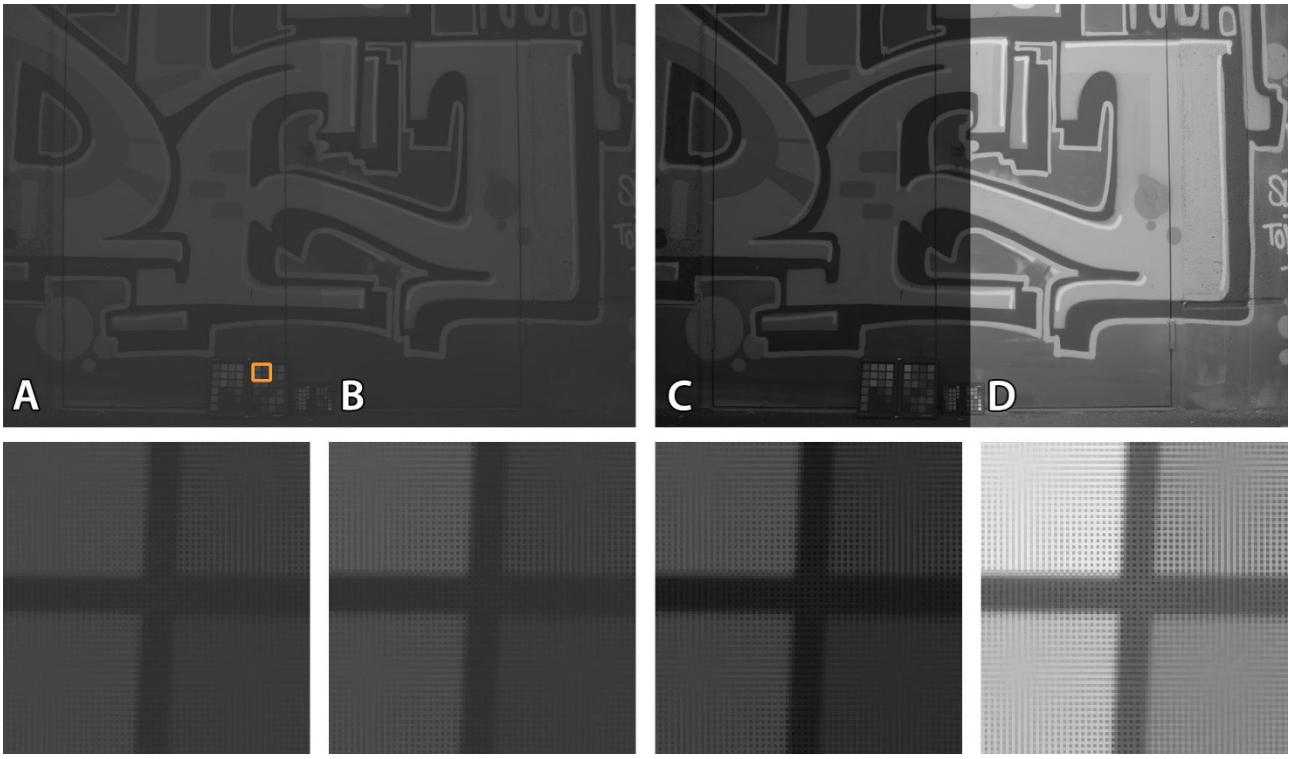


Figure 18. A Nikon D5600 RAW file in its initial processing stages. A) indicates the non-linearised image, while B) shows the linearised version (which is slightly brighter). C) displays the linearised, black-subtracted version, while D) shows the RAW image scaled to the [0.0, 1.0] interval. The CFA patterning is apparent from the zoomed-in parts in the bottom row. All images have been subjected to a 1/2.2 gamma encoding, as they would otherwise be too dark for display.

Some RAW converters then scale the input data (usually 12- or 14-bit) to 16-bit integer values. For a 14-bit file (which can hold 2^{14} or 16,384 different integer values), the most straightforward scaling approach is multiplying all DNs with four so that every pixel features a value between 0 and 65,535 (i.e. 2^{16}-1). Other RAW converters (and COOLPI) rescale the raw $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ DNs to the [0.0, 1.0] floating-point interval (Figure 18D). There are also more adaptive scaling approaches possible. These could ensure that details in the brightest portions of the image (the so-called highlights) never get clipped. The 'Exposure' or 'Exposure compensation' sliders found in RAW converters like Adobe Camera Raw (https://helpx.adobe.com/camera-raw/using/supported-cameras.html) or RawTherapee (https://www.rawtherapee.com) enable interactive scaling to account for over- or underexposure. Although COOLPI does not offer such interactive scaling, the authors are experimenting with automatic image-specific scaling to unify the exposure of all photographs from one particular graffito because exposure values cannot be spot on every time. Finally, some RAW converters also feature sliders that non-linearly scale the DNs, typically to brighten the shadows or compress the highlights in the image. If a colour-accurate rendering of the scene is wanted, one should stay clear of such operations.

The bottom row of Figure 18 illustrates that, at this stage, the image looks like a patterned greyscale image. The greyscale nature relates to the fact that there is still only one DN per pixel (either $\mathcal{R}$, $\mathcal{G}$ or $\mathcal{B}$); the patterning originates from the CFA. It is easy to discern because most objects reflect differently in the Red, Green and Blue spectral ranges. Neighbouring photosites capture thus very different DNs, even under uniform illumination.

4.2. White Balancing

White balancing is the most crucial RAW development step to guarantee that the output photo represents colours faithfully. In short, white balancing ensures that white looks white and grey looks grey in the output image. A perfectly white pixel should have identical *RGB* values, like *R*: 255 – *G*: 255 – *B*: 255 for an 8-bit output image. However, this does not only hold for white pixels. Any spectrally neutral or achromatic object (i.e. reflecting any wavelength of light with identical magnitude) that varies from black over grey to white should feature equal RGB values in the final output image. So, when a light source with a flat SPD over the visible spectrum (like illuminant E) gets diffusely reflected from a light grey and spectrally neutral card, one expects the camera to digitise the resulting equi-energy stimulus with identical $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ pixel values. Figure 19A shows this not to be the case since the spectral response curves of the camera's sensor have different widths and heights. This usually results in a high $\mathcal{G}$ value and lower $\mathcal{B}$ and $\mathcal{R}$ values. The camera counteracts this by normalising the $\mathcal{R}$ and $\mathcal{B}$ values to the $\mathcal{G}$ DN via multiplication with a scaling factor or so-called multiplier value, thus ensuring identical $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ values for all pixels of the achromatic card (see Figure 19A).

However, the illumination source also influences these multipliers. Suppose illuminant A shines on the same spectrally neutral card. The subsequent stimulus will now be most substantial in the Red spectrum, which results in high $\mathcal{R}$ DNs (sometimes even higher than the $\mathcal{G}$ ones) and low $\mathcal{B}$ values (see Figure 19B). In this case, the camera has to slightly reduce the $\mathcal{R}$ pixel values (with a multiplier smaller than 1) and boost the $\mathcal{B}$ DNs to ensure the achromatic card looks light grey and not orange-red.

To determine these multipliers accurately, the camera must estimate (or be told) the scene illumination when acquiring the photo. Since the camera cannot consider an entire SPD, the illumination's dominant colour is characterised by the Correlated Colour Temperature (CCT), a concept introduced in Section 2.5. Human eyes constantly adjust to such CCT changes and can tell a wall is white, irrespective of the illumination conditions (Hung, 2006; Livingstone, 2002). Digital sensors and film are unable to do so. In the analogue era, one had to change the type of film or use appropriate corrective filters to avoid colour casts. In digital photography, the camera needs to know the predominant colour of the illumination to calculate the correct multipliers (E. Y. Lam & Fung, 2009). That is why the camera automatically estimates the illumination's CCT when acquiring a photo; or the photographer manually determines this number, often via the camera's white balance presets like "cloudy", "fluorescent", "shade", or "direct sunlight" (see Figure 10 for common values). With an established CCT, the camera computes the illumination's white point, from which the correct multipliers get derived. Finally, those multipliers

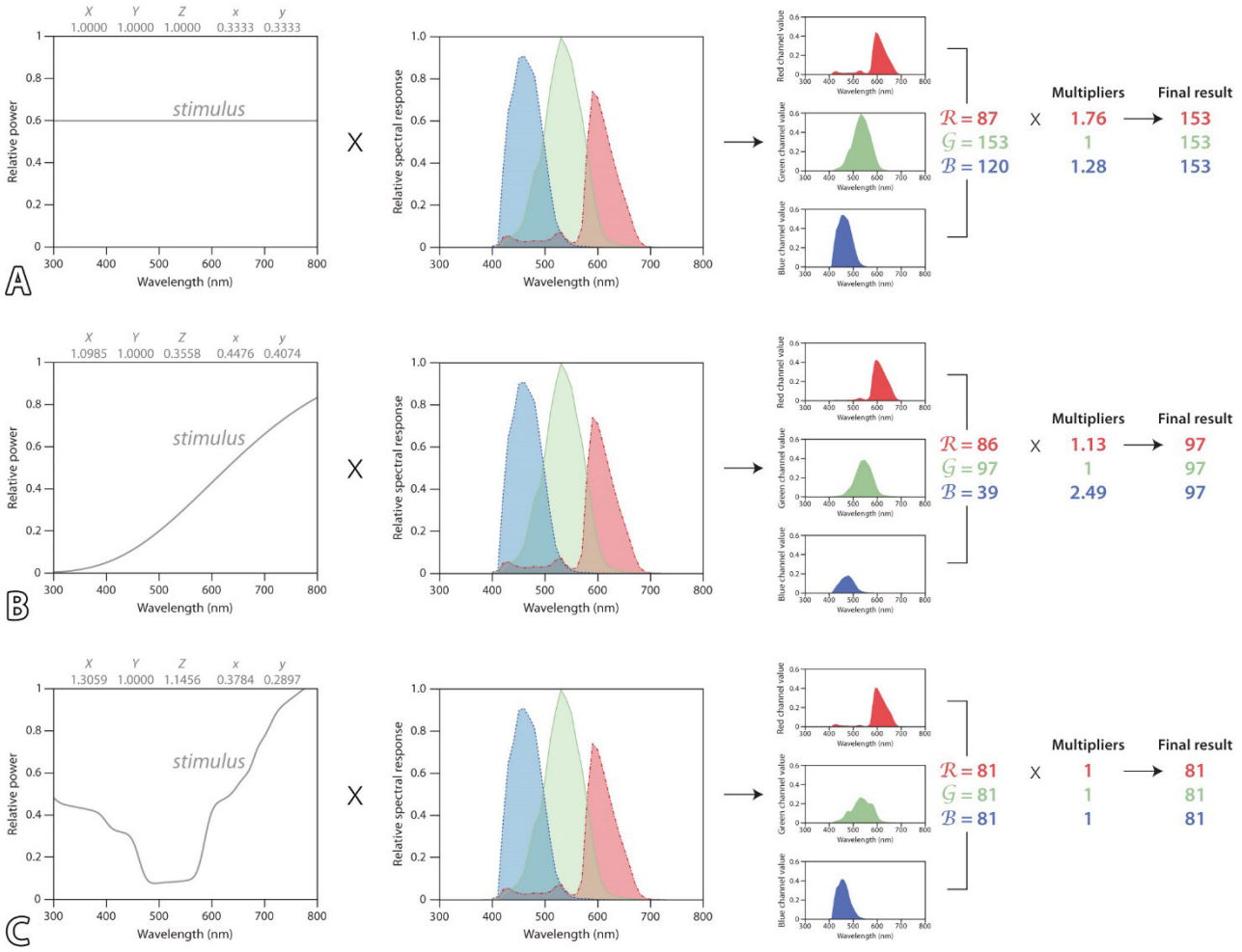


Figure 19. A) depicts how the interaction of a spectrally flat stimulus with the spectral responsivity curves of a digital camera yields unequal $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ values. The multipliers can be computed knowing that the stimulus is supposed to generate three equal values. Row B depicts a similar situation with standard illuminant A diffusely reflected off a spectrally neutral reference card. This stimulus mainly contains red light, so the multipliers are adjusted accordingly. Row C displays an experimentally defined stimulus, resulting in maximum RAW values [1, 1, 1] for the Nikon D700. The illumination's SPD (which looks like the stimulus since the illumination got diffusely reflected from a spectrally neutral card) can be summarised via its *XYZ* tristimulus values [1,3059, 1, 1.1456] or (*x*, *y*) chromaticity values (0.3784, 0.2897). These values constitute the reference white of the Nikon D700 RAW space. Data on the D700's spectral response curves come from the camspec database (https://www.gujinwei.org/research/camspec/camspec_database.txt) published by Jiang et al. (2013).

get saved in the Exif image metadata. Figure 19 mentions the white point values above the stimuli. Irrespective of its determining procedure, the estimated scene illumination white point is known as the adopted white (International Organization for Standardization, 2012) or camera neutral (Adobe, 2021).

When developing the RAW file (either by the camera or—as in the case of INDIGO with COOLPI—afterwards on a computer) into a JPG or TIFF, these multipliers are used to recalculate the linearised, black-subtracted and scaled raw pixel values of all channels to make sure that the scene's spectrally neutral zones—and by extension also all the other

colours in the digital image—appear without major colour casts in the final output image, irrespective of illumination condition. Technically, a diagonal matrix $\underline{D}$ with the channel multipliers takes care of this (see Figure 20).

Figure 16 indicates that white balancing chromatically adapts the camera's adopted white to the so-called camera RAW space reference white (Rowlands, 2020b). As mentioned in Section 2.5, the reference white of the camera RAW space equals the SPD or chromaticity coordinates of the illumination that generates identical $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ values when diffusely reflected off a spectrally neutral target. In other words, only this illumination will yield the equal RAW values one expects for a spectrally neutral target. Row C of Figure 19 shows the reference white SPD and chromaticity coordinates for a Nikon D700 camera. When locating these chromaticity coordinates in the diagram of Figure 8, it is clear that this camera's reference white is more of a reddish, almost magenta colour rather than a neutral one. In the colour space transformation step (Section 4.5), this RAW space reference white will be chromatically adapted to the output colour space's D50 or D65 reference white (called the encoding white).

Although digital cameras are usually reasonably good at estimating the CCT value/the white point for outdoor illumination conditions, more reliable values can be computed from a photographed spectrally neutral card (a so-called white balance card) or the illumination's SPD (see Figure 20, and also Section 6.2 for a comparison). That is why the INDIGO staff uses a Sekonic C-7000 SPECTROMASTER portable handheld spectrometer to measure the graffito's spectral illumination directly before starting the photographic documentation. The latter begins with a photo from an X-Rite ColorChecker Passport Photo 2 reference target. Since this target features a few spectrally neutral patches, one can use their RAW DNs to extract the channel multipliers directly. Thus, this initial photo serves as a backup solution to the spectrometer data (see Verhoeven, Wogrin et al. in this volume for more details about the data acquisition pipeline).

The image is still greyscale at this point (see Figure 20), but

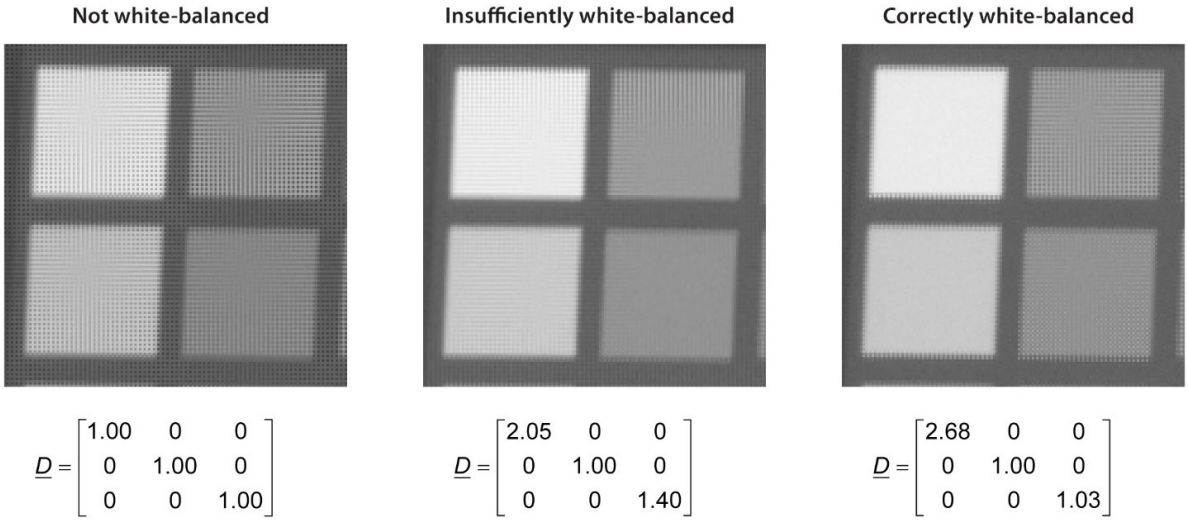


Figure 20. Three crops of the Datacolor Spyder Checkr reference chart, visible beside the smaller X-Rite ColorChecker Passport Photo 2 below a graffito in Figures 18 and 21. Both reference targets contain spectrally neutral patches (the outer left ones in the crop). On the left is a crop of a RAW image that is not white-balanced, which is conceptually identical to multiplying it with a white balance matrix $\underline{D}$ containing only 1's on its diagonal. The middle crop represents a white-balanced RAW image using the multipliers determined by the camera's automatic illumination estimation and stored in the image's Exif metadata. These multipliers are not sufficiently accurate; the neutral patches still display some CFA patterning, indicating that their pixels feature dissimilar $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ values. The image on the right eliminates this because it is adequately white-balanced by multiplying every $\mathcal{R}$ DN by 2.68 and every $\mathcal{B}$ DN by 1.03. These numbers were determined from the RAW pixel values of the lower neutral patch.

spectrally neutral objects (like dedicated white balance cards or neutral patches of a colour reference chart) no longer display CFA-related patterning. In other words, their linear $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ values will all be very similar (ignoring image noise and illumination differences). However, the mosaic pattern is still apparent for chromatic objects (like the two patches on the right of the chart crop in Figure 20) since they have different reflectances in the visible spectrum's Red, Green and Blue parts.

4.3. Data Clipping

White balancing commonly leaves the $\mathcal{G}$ DNs generated by the green-filtered photosites alone. After scaling to either a [0.0, 1.0] floating-point range or a [0, 65535] integer range, no $\mathcal{G}$ DN will surpass either 1.0 or 65,535. However, that is not the case for the $\mathcal{B}$ and $\mathcal{R}$ DNs. Since they have all been multiplied with a constant factor in the white balancing step, the $\mathcal{R}$ and $\mathcal{B}$ values of highlight pixels could easily go beyond those floating-point and integer maxima. If left untreated, this will result in highlights with odd colours.

Imagine a 12-bit RAW image of a graffito with a white outline. The entire graffito lies in the shadow, apart from a tiny sunlit part of the outline. With a camera exposure suitable for the shaded region, the pixels of the white outline will feature high $\mathcal{G}$ DNs like 3800 (see Table 2). The photosites that sample the white sunlit part will overflow with photons, yielding saturated pixels with the highest possible 12-bit DN: 4095. Since digital cameras are usually less sensitive in the Red and Blue spectral range, the $\mathcal{R}$ and $\mathcal{B}$ DNs that make up the white outline are substantially lower; even the sunlit pixels are not saturated (see Table 2). After rescaling to the [0.0, 1.0] range, the $\mathcal{G}$ DNs of the outline in the shadow are close to 1. Still, those of the saturated white pixels will be 1 (again, the maximum possible value). At this stage, the $\mathcal{R}$ and $\mathcal{B}$ DNs are still far below their maximum value because the same constant value was used to scale all DNs linearly. However, white balancing changes this picture.

Even though white-balanced pixels of the white outline feature identical high values (i.e. 0.93), a problem occurs with the $\mathcal{R}$ and $\mathcal{B}$ DNs of the pixels representing the sunlit white outline. Using a $\mathcal{R}$ multiplier of 1.4 and a $\mathcal{B}$ multiplier of 1.6, both DNs surpass 1.0 with a value of 1.21. At the same time, the maximum $\mathcal{G}$ DNs remain 1.0 because they stay unaltered during white balancing. Because the $\mathcal{G}$ DNs are too low with respect to the $\mathcal{B}$ and $\mathcal{R}$ ones, and because green and magenta are complementary colours, such green-deficient pixels look pinkish or magentish. A processing step that clips all channels to the lowest value of the maximum $\mathcal{R}$, $\mathcal{G}$ and $\mathcal{B}$ DNs (i.e. 1.0 here) prevents such false highlights colours. Some advanced RAW converters typically offer the user a 'highlight recovery' slider, which estimates the underrepresented highlight component(s) rather than clipping it (note that this approach also includes a highlight compression to bring all DNs back in the [0.0, 1.0] range). And yes, RAW converters must take similar precautions to avoid false colours in the shadow areas, which could occur with white balance multipliers smaller than 1.0.

	DN range	Majority of the white outline			Sunlit white outline		
		$\mathcal{R}$	$\mathcal{G}$	$\mathcal{B}$	$\mathcal{R}$	$\mathcal{G}$	$\mathcal{B}$
12-bit RAW	[0, 4095]	2714	3800	2375	3529	4095	3088
Black-subtracted, scaled RAW	[0, 1]	0.66	0.93	0.58	0.86	1.00	0.75
	[0, 65535]	43444	60815	38015	56472	65535	49415
Black-subtracted, scaled, white-balanced RAW	[0, 1]	0.93	0.93	0.93	1.21	1.00	1.21
	[0, 65535]	60821	60815	60824	79061	65535	79064

Table 2. Clipping in the $\mathcal{R}$ and $\mathcal{B}$ channels is often necessary so their maximum values do not surpass those of the $\mathcal{G}$ DNs after scaling and white balancing.

4.4. Demosaicing

Even though the linearisation, black-level subtraction, scaling and white balancing have changed the initial $\mathcal{R}$, $\mathcal{G}$, and $\mathcal{B}$ DNs, the RAW file is still a single-channel image in which each pixel has only one value, representing either the collected Red, Green or Blue incoming light (Figure 21A). To

end up with an image where every pixel features an $\mathcal{RGB}$ triplet, the remaining two values are interpolated from the surrounding pixels in a process called CFA interpolation, de-Bayering or demosaic(k)ing. An immense range of demosaicing algorithms has been developed to tackle this interpolation process and the various artefacts it can introduce. Most commercial RAW converters only offer one undisclosed algorithm. At the same time, free and open-source software like RawTherapee and COOLPI provide various demosaicing options. The demosaiced result is an $\mathcal{RGB}$ image, a three-channel or three-band image featuring a 'channel' or 'band' for all the $\mathcal{R}$ values, one for all the $\mathcal{G}$ and one with all the $\mathcal{B}$ values (see Figure 21B).

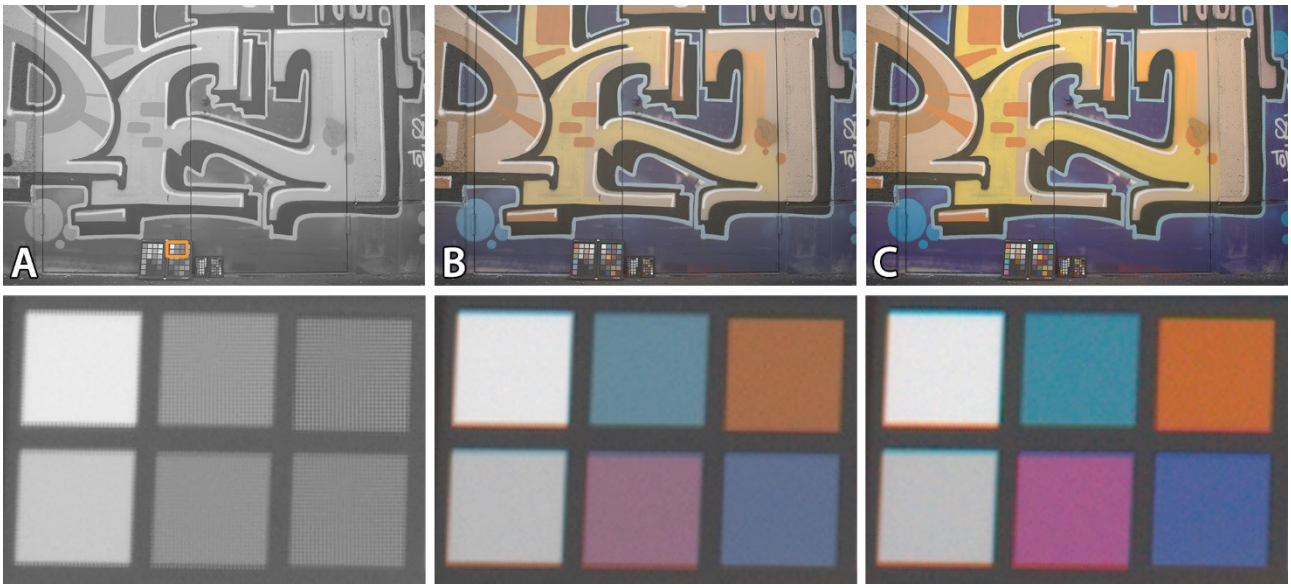


Figure 21. A mosaiced image (A) is a one-band, greyscale image that still shows CFA patterning in chromatic areas. After demosaicing (here with the algorithm by Malvar et al. (2004)), a three-band colour image is obtained for the first time in the RAW development process. The image colours seem realistic but desaturated. The insets of column C result from transforming B via a colour rotation matrix $\underline{R}$ into the linear sRGB output colour space, followed by a gamma encoding. Although nonlinear gamma encoding occurs only at the end of the RAW development chain (i.e. Figure 21C), it was also applied to Figures 21A and B because they would otherwise appear too dark and be useless as illustrations.

COOLPI, and many other RAW converters, also provide the option to create minimally/half-size/2x2 demosaiced images. The result is not obtained via an actual demosaicing algorithm because no values are interpolated. However, this approach bins each quartet of 2x2 pixels ($\mathcal{R}$, $\mathcal{G}$, $\mathcal{G}$, $\mathcal{B}$) into a single $\mathcal{RGB}$ pixel. The $\mathcal{B}$ and $\mathcal{R}$ values are taken as-is, while the $\mathcal{G}$ DN is the average of the two $\mathcal{G}$ values. The result is an image with only a quarter of the megapixels of a fully demosaiced RAW image.

At this point, we can also explain Figure 1 better. The middle row shows RAW colour images with a strong green colour cast. These demosaiced images were not white-balanced so that they could visualise the dominance of the green spectral sensitivity curve. The images are also very dark because they lack gamma encoding. The lack of this non-linear encoding is expected at this stage of the RAW processing pipeline as it only happens at the very end (see Section 4.6). However, since this makes it hard to visualise what is happening, a 1/2.2 gamma encoding was applied to Figures 18, 20, 21A and 21B for visualisation purposes.

4.5. Colour Space Transformation

Although the RAW processing pipeline finally yielded a colour image, the pixels' $\mathcal{RGB}$ values are still expressed in the camera-specific RAW space. To ensure that imaging

hard- and software know how to interpret these numbers, pixel values must be converted into a standard RGB output colour space like sRGB or Adobe RGB (1998). This step is likely the most complex, and together with white balancing, critical in pursuing accurate photo colours. Without going into detail yet, the RAW converter uses multiplications with one or more matrices to express the black-subtracted, scaled, white-balanced, clipped, and demosaiced pixel values into a standard colour space. Various approaches exist to obtain these matrices (Rowlands, 2020a), all with varying accuracy and applicability.

One possible approach is to use the matrices that research labs or imaging companies have derived. For instance, the Paris-based company DXOMARK (https://www.dxomark.com) scientifically assesses cameras and lenses. Amongst a plethora of quantitative data, their website provides the matrices for most digital cameras to convert the white-balanced RAW $\mathcal{RGB}$ data to sRGB for the D50 and A illuminants. Figure 22 depicts both illuminant-specific matrices for a Nikon Z7ii camera. Imagine a scene photographed using D50 illumination. One would merely need to multiply the matrix of Figure 22 by the white-balanced, demosaiced $\mathcal{RGB}$ values obtained after step 4.4 to end up with perfect colours. To ensure that neutral tones in the white-balanced image (for example, [0.3, 0.3, 0.3]) get mapped to neutral tones in the final sRGB image, the sum of all row coefficients always equals [1, 1, 1] for these matrices.

CIE-D50

		R_{sRGB}	G_{sRGB}	B_{sRGB}
Color matrix as defined in ISO standard 17321	R_{raw}	1.81	-0.72	-0.09
	G_{raw}	-0.14	1.44	-0.31
	B_{raw}	0.03	-0.46	1.43

CIE-A

		R_{sRGB}	G_{sRGB}	B_{sRGB}
Color matrix as defined in ISO standard 17321	R_{raw}	1.8	-0.72	-0.09
	G_{raw}	-0.18	1.39	-0.22
	B_{raw}	0.03	-0.65	1.61

Figure 22. These matrices from https://www.dxomark.com/Cameras/Nikon/Z7II---Measurements transform white-balanced, demosaiced Nikon Z7ii RAW image data to the sRGB colour space. A 3x3 matrix is provided for two illumination conditions: the standard illuminants D50 and A. Some authors call such matrices colour rotation matrices *R* (Rowlands, 2020a).

However, illumination never has the same SPD as the theoretical D50 illuminant (although natural daylight could come close). To deal with these and other complexities, the colour transformation step usually comprises three steps instead of just one matrix. First, white-balanced and demosaiced $\mathcal{RGB}$ values are transformed into the CIE XYZ colour space. Afterwards, another matrix transforms the *XYZ* values into the linear form of a specific *RGB* output colour space like sRGB or Adobe RGB (1998). Between both transformations, a chromatic adaptation accounts for the fact that these output colour spaces are determined for an illuminant which likely differs from the illumination used while photographing. Because of its importance and complexity, Section 5 further details this step.

4.6. Gamma Encoding, Bit-depth Reduction and JPEG Encoding

At this stage, the entire RAW conversion process has been linear. However, because the first generations of computer monitors displayed pixel values in a non-linear fashion, colour spaces like sRGB and Adobe RGB (1998) also have a gamma value defined to cancel out this non-linear behaviour. This value is fixed for every colour space but often around 1/2.2 (see Figure 8). Multiplying every pixel with this gamma value is the only non-linear tonal transformation in the entire RAW development pipeline.

RAW converters often combine this colour space-related gamma encoding with tonal mapping. The latter can produce photos with more contrast and punchier colours than real life (see Figure 1 at the lower right), as most people prefer that (Parulski & Spaulding, 2003). To keep colour accurate, COOLPI does not apply such tonal mapping. If tone mapping would really be needed—for instance, to fit the high contrast ratio of a digital photo into the often lower contrast ratio supported by monitors or photo paper—it should ideally occur in the colour space transformation step (Torger, 2018). Finally, this step includes a bit-depth

reduction and JPEG encoding since most photographs are saved as 8-bit JPG files.

A possible seventh step—occurring before the bit depth reduction—to achieve the final rendered photo is via additional (global or local) pixel editing. Operations like sharpening, correcting lens distortion or chromatic aberration, denoising and fixing red eyes all aim to remove imperfections and make the image more pleasing. Although many RAW converters combine the conversion steps 1 to 6 with such editing capabilities (e.g. Adobe's Lightroom, Phase One's Capture One, RawTherapee), pure RAW converters like dcraw (Coffin, 2008) and LibRaw (LibRaw LLC, 2021)—and right now also COOLPI—do not offer this functionality. The absence of such image processing capabilities in COOLPI explains why the lower right blue patch in Figure 21C still contains image noise, while the neutral patches' boundaries still suffer from 'fringes' of colour due to chromatic aberration of the lens.

5. The Complex Marriage Between Digital Cameras and Colour

5.1. Luther and Ives

In an ideal world, a camera's sensor spectral response curves would mimic the cone spectral sensitivity curves of the human eye (or be a linear transformation thereof, like the 2° or 10° CMFs) (Sharma, 2003). Only in that way would a camera reproduce accurate colours congruent with the colour experience of a human observer. This is the Luther–Ives condition (Ives, 1915; Luther, 1927). However, Figure 23 illustrates that a camera's spectral response curves might seriously deviate from the human cone responses or any existing set of positive CMFs (Parulski & Spaulding, 2003). Figure 23 displays the Samsung Galaxy S8 smartphone camera's spectral response curves and their area-normalised version. The second plot means that the area under each sensitivity curve is identical, so they would produce equal $\mathcal{RGB}$ values if equi-energy illuminant E were the stimulus. This plot looks quite different from CIE 1931 2° CMFs (right part of Figure 23), which are by default area-normalised.

What does this difference mean? Imagine a given spectral stimulus. This smartphone would digitise this stimulus into raw $\mathcal{RGB}$ values. At the same time, the human eye and a colorimeter would generate *LMS* and *XYZ* tristimulus values. But because the smartphone's digital camera does not satisfy the Luther–Ives condition, it is impossible to linearly transform the camera's raw $\mathcal{RGB}$ DNs to exactly end up with these ideal *LMS* or *XYZ* coordinates. However, this is not solely so for smartphone cameras. It is safe to state that no standard digital photo camera meets the Luther–Ives

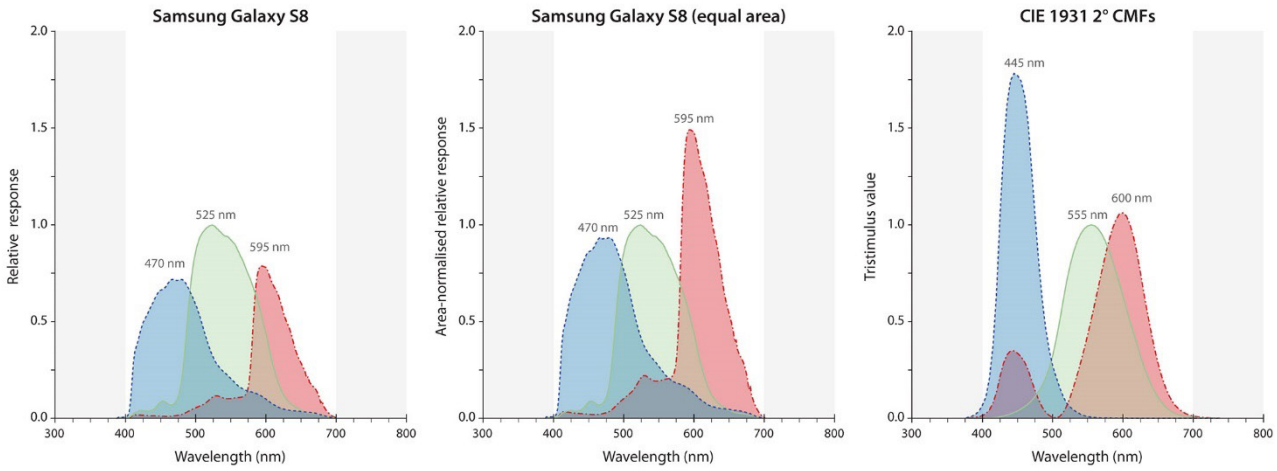


Figure 23. The relative spectral sensitivity curves of a Samsung Galaxy S8 smartphone camera on the left (data from Burggraaff et al. (2019)) with its area-normalised version displayed in the middle. This plot can be compared to the CIE 1931 2° CMFs on the right (CMF data from Colour and Vision Research Laboratory (2021)). The onset of the invisible near-ultraviolet and near-infrared regions is always indicated in grey.

condition (see also Holm (2006) and Jiang et al. (2013)). In other words: digital cameras are far from perfect colour-capturing instruments because colour accuracy is only one out of several, often mutually exclusive image criteria (such as image noise, light gathering efficiency, resolving power, and manufacturing costs) that are taken into account when designing an imaging sensor (Berns, 2001; Imai et al., 2001).

5.2. Camera-Specific Transformations

Although a camera's sensor spectral response curves would ideally mimic the cone spectral sensitivity curves of the human eye or any of the CMFs, Figure 23 shows these curves are different in shape and seem slightly shifted with respect to one another. One must thus find a way to minimise the differences between the two sets of curves or develop a good transformation from the $\mathcal{RGB}$ values into *XYZ* values. A long look-up table with all corresponding values (e.g. $\mathcal{RGB}$ values [768, 3200, 2304] correspond to *XYZ* values [5, 78, 19]) would be possible but impractical. However, with linear algebra, one can find a 3x3 matrix to do that.

Suppose a perfect 3x3 matrix to transform the camera's initial $\mathcal{RGB}$ values to *XYZ* values would exist. Even then, that transformation would be camera-dependent as the imaging sensors used by different camera brands and models feature minor to more considerable variations in spectral responsivity (see Figure 24). In other words: after demosaicing, all RAW $\mathcal{RGB}$ pixel values are expressed using the RGB colour model, but because different camera sensors will generate different $\mathcal{RGB}$ values when they simultaneously observe the same scene using the same lens under the same lighting conditions, the camera RAW space is device-dependent (Punnappurath & Brown, 2020). And to add a final touch of complexity: this 3x3 transformation is also illumination- and (often) scene-dependent because they both determine the unlimited amount of spectral stimuli a camera can sample. The following three sections cover how one can solve all this complexity.

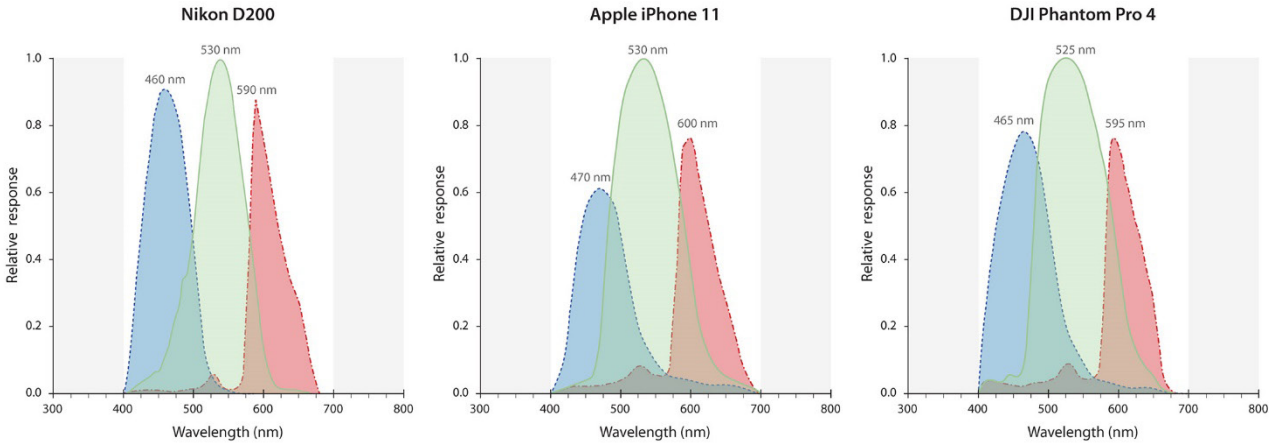


Figure 24. The three spectral responsivity curves of the imaging sensor inside a Nikon D200 semi-professional digital camera, an Apple iPhone 11 mobile phone and the DJI Phantom 4 Pro quadcopter camera. The spectral data for the last two imaging sensors came from Tominaga et al. (2021) and Burggraaff et al. (2019), respectively. The onset of the invisible near-ultraviolet and near-infrared regions is indicated in grey. Note that these response curves are the combined result of the transmittance by the camera lens, hot mirror and CFA elements, plus the quantum efficiency of the silicon sensor.

5.3. Colour Space Transformation: Three Matrices…

Like a coordinate transformation changes coordinates from one coordinate reference system to another (Iliffe & Lott, 2008), colour transformations map colour data from one colour space to another. As explained above, a camera-, scene- and illumination-dependent linear 3x3 matrix can transform the white-balanced and demosaiced $\mathcal{RGB}$ pixel values from the device-dependent RAW space to the device-independent CIE XZY space. Nevertheless, this matrix is only one part of the entire colour space transformation

step mentioned in Section 4.5. Even though that section summarised all processing by one colour rotation matrix *R*, the entire colour space transformation step (and thus matrix *R*) is often split into three 3x3 matrices. In this way, each sub-step is easier to standardise and control (consider Figure 16 for a visual representation of these steps):

1. **Transformation to *XYZ* | $\mathcal{RGB}_{\text{RAWReferenceWhite}}$ to $XYZ_{\text{AdoptedWhite}}$**: the first step transforms the linearised, black-subtracted, scaled, white-balanced, clipped and demosaiced linear $\mathcal{RGB}$ data into *XYZ* values. This is the linear 3x3 matrix mentioned in Section 5.2. In the academic literature, this matrix typically goes by the name Colour Correction Matrix, Colour Conversion Matrix, Camera Characterisation Matrix or Compromise Colour Matrix (all four using the acronym CCM). The *CCM* consists of nine elements (see Equation 2); these depend on the scene illumination, the imaging hardware (i.e. the camera's spectral response plus the lens' spectral transmission) and the scene values used to estimate it (typically 24 patches of a ColorChecker reference chart; see Section 5.5).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{AdoptedWhite}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{ReferenceWhite}} \quad (2)$$

2. **Chromatic adaptation | $XYZ_{\text{AdoptedWhite}}$ to $XYZ_{\text{OutputWhite}}$**: although outdoor illumination can, to a certain extent, approximate the D50 or D65 SPD, it usually differs quite a bit. However, output colour spaces like sRGB, Adobe RGB (1995) or ProPhoto RGB include a standard illuminant in their definition. This standard illuminant is the space's reference white, defining how cool or warm a perfect 'white' pixel looks in that colour space (see Figure 8). Specifying colours to a new reference white is technically called chromatic adaptation. Thus, a so-called Chromatic Adaptation Transform (CAT) is needed; another 3x3 matrix *CAT* to transform the capture's illumination-related *XYZ* values into new *XYZ* values, expressed with respect to the white reference of the output colour space (typically XYZ_{D50} or XYZ_{D65}). Converting pixel values between white points is usually—and thus also by COOLPI—performed with a Bradford CAT (K. M. Lam, 1985).
3. **Transformation to *RGB* | $XYZ_{\text{OutputWhite}}$ to RGB_{linear}**: finally, the XYZ_{D50} or XYZ_{D65} coordinates are transformed into a common output colour space like sRGB or Adobe RGB (1998). COOLPI uses the former by default. Although these colour spaces are not linear, this third 3x3 Output Space Matrix *OSM* converts to a linear version of the colour space. Afterwards, a separate colour space-specific (see Figure 8) gamma encoding follows to end up with the final *RGB* coordinates that are no longer linearly related to the initial $\mathcal{RGB}$ data.

5.4. …But Many Possible Combinations

Every *OSM* to compute linear *RGB* values is standardised and known. For instance, they are freely available on Bruce Lindbloom's webpage (http://www.brucelindbloom.com/index.html?Eqn_RGB_XYZ_Matrix.html). The same website (http://www.brucelindbloom.com/index.html?Eqn_ChromAdapt.html) also provides common Bradford *CAT*s. The main difficulty thus lies in finding the nine unknown elements of the first matrix. Although Section 5.5 details how to determine the *CCM*, the subsequent paragraphs first address a few variations on the workflow sketched in Section 5.3.

- The colour transformation step consists of three 3x3 matrices. Combined with the channel multipliers stored in the diagonal matrix *D*, that makes four matrices (each indicated with a specific colour in Figure 16). However, the scientific literature covers various ways to unite these components (for an overview, see Rowlands (2020a)). In addition, some matrices have slightly different functions despite bearing the same name across the literature. For example, the *CCM* spectrally characterises the digital camera by taking values expressed in the camera-dependent RAW space and transforming them into the device-independent CIE XYZ colour space. Some authors and this paper (see next section) compute

this matrix based on white-balanced $\mathcal{RGB}$ data. In contrast, others derive it from the $\mathcal{RGB}$ values before white balancing. In both cases, the procedure is called camera characterisation, and *CCM* is the resulting matrix. A *CCM* might thus include channel multipliers (indicated with ① in Figure 25), but the one discussed here does not (② in Figure 25). This approach is for a particular reason. Decoupling white balancing from the $\mathcal{RGB}$-to-*XYZ* matrix operation allows for easy finetuning of the channel multipliers afterwards in RAW conversion software like COOLPI (using the illumination's CCT or neutral patches in the image; see Section 4.2 and Figure 20). In addition, some authors suggest that scaling the $\mathcal{R}$, $\mathcal{G}$, and $\mathcal{B}$ DNs before demosaicing might yield a better quality output image (Rowlands, 2020a). Finally, it also makes sense intuitively. By multiplying every $\mathcal{B}$ and $\mathcal{R}$ value with a multiplier, white balancing essentially normalises the spectral sensitivity curves by area (cf. the central graph of Figure 23). Because the CIE XYZ CMFs are area-normalised by definition, both sets of curves thus become instinctively comparable.

- Sometimes, the *CCM* and a kind of *CAT* are combined. Adobe calls this combination the Forward matrix *F*, and it will transform white-balanced and demosaiced $\mathcal{RGB}$ data directly into *XYZ* data for a D50 illuminant (Adobe, 2021). A standard *OSM* from Lindbloom's webpage (together with a *CAT* if the output space's reference white differs from D50) takes these data into the final linear version of the output colour space. Adobe also makes it easy for users. Upon converting any proprietary RAW file to Adobe's open-source DNG file format with their free Digital Negative Converter (https://helpx.adobe.com/uk/camera-raw/using/adobe-dng-converter.html), two forward matrices for that particular camera get written into the DNG metadata: for the standard illuminants A and D65. Adobe has determined both matrices for most digital cameras. An enormous but precious effort as manufacturers typically do not disclose any characterisation information about their cameras. Providing a matrix *F* for two highly different illuminants enables DNG-aware software to interpolate the final forward matrix for the CCT of the actual scene illumination.
- Combining all three matrices yields the colour rotation matrix *R* provided via the DXOMARK webpage (see Section 4.5) or stored in the metadata of certain Olympus and Sony RAW files. As Adobe, DXOMARK provides two matrices—for illuminants A and D50 (see Figure 22)—to interpolate a CCT-specific matrix from. The camera-embedded matrices typically cover a handful of specified CCT ranges. The CCT estimated by the camera or determined in the RAW development software is again used to select (or interpolate) the most suitable matrix *R* from all pre-computed ones.

5.5. Getting to Know Nine Unknowns: Camera Characterisation

Determining the nine unknowns of the *CCM* is a procedure known as camera (spectral) characterisation. Since we have to transform the white-balanced and demosaiced raw data into the CIE XYZ space, having three sets of raw $\mathcal{RGB}$ values generated by a camera under illumination with a known SPD suffices. For example, one could illuminate a graffito with an artificial light mimicking a D50 illuminant and photograph the graffito. After white-balancing and demosaicing the RAW image, the $\mathcal{RGB}$ values of three pixels—each from a graffito patch with a different colour—are written down. Measuring those three actual graffito patches with a spectrophotometer yields the XYZ_{D50} counterparts of these pixel DNs. Using Equation (2), nine equations arise from the two sets of nine values (three sets of three values). Because the *X*, *Y*, *Z*, and the $\mathcal{R}$, $\mathcal{G}$, $\mathcal{B}$ values are known, these nine equations enable computing all unknown elements of the colour correction matrix *CCM*.

However, the resulting matrix would only be effective for that specific camera (plus lens), lighting, and graffito. In other words, altering any of those three would necessitate the creation of a new *CCM*. Since the constant need for new *in-situ* colour measurements would make this procedure practically unfeasible, solutions were developed to simplify the spectral characterisation of cameras. The following

paragraphs detail the chart-based characterisation method since it is ubiquitous and currently the only method COOLPI offers. To aid understanding, the upper part of Figure 25 can function as a visual guideline.

Companies like X-rite (now Calibrite) and Datacolor are known for their photographic colour reference targets. One of the most known and widely used charts is the 24-patch ColorChecker chart. The chart comes in different sizes, ranging from the A4-sized ColorChecker Classic (https://calibrite.com/us/product/colorchecker-classic) to a credit card-sized ColorChecker Classic Mini (https://calibrite.com/us/product/colorchecker-classic-mini). Besides six achromatic patches, every ColorChecker target features 18 patches that should represent everyday colours found in foliage, the sky and human skin (see Figure 25). Although one could measure the spectral reflectance curves of each patch to derive its *XYZ* values for a particular illuminant, XYZ_{D50} or $L^*a^*b^*_{D50}$ values can be found online on sites like BabelColor (https://babelcolor.com/colorchecker-2.htm) or sometimes on a sheet enclosed with the target (note that CIE $L^*a^*b^*$ values can be retrieved via a transformation of the *XYZ* values). Instead of measuring three differently coloured parts of an object (like a graffito) every time it gets photographed, it is much more convenient to include the ColorChecker target in the photo and use the D50 *XYZ* values of its 24 patches to compute the CCM.
Because 24 patches yield 72 values on the $\mathcal{RGB}$ and *XYZ* sides (i.e. 24 * 3), the CCM can be calculated more robustly by solving these 72 equations. Moreover, the matrix becomes

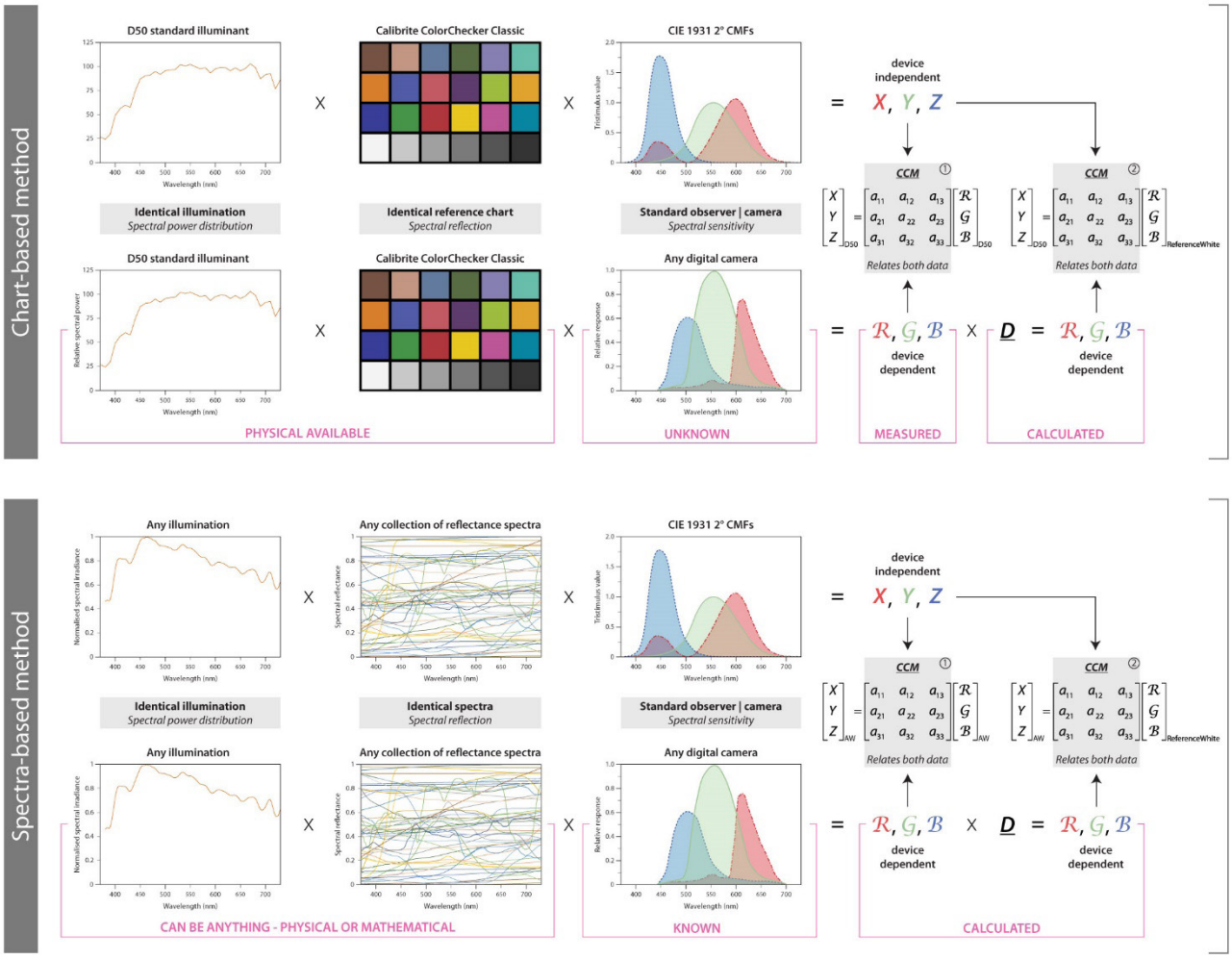


Figure 25. The chart-based versus the spectra-based approach for camera characterisation. In the spectra-based approach, AW means Adopted White. For the chart-based approach, the AW is usually D50 or D65.

applicable for a broader range of photographic scenes as more colours are involved in its computation. For instance, imagine a scene with many blue paints and a reference chart without a blue patch to build the *CCM*. In that case, colour accuracy might suffer for the blue tones. On the other hand, one could extend the number of patches to include typical spray paint colours, thus making the *CCM* more accurate for photographing graffiti. These potential positive effects of the chart-based approach notwithstanding, a few critical observations must be made.

1. Too often, the values of such reference targets are used irrespective of the photographic illumination conditions. Remember that colour values are generated by integrating a stimulus over a set of XYZ CMFs, and this stimulus is a function of object reflectance and illumination. Although the reflectance of the 24 patches should remain invariant for a couple of years when treating the chart appropriately, illumination conditions continuously vary. The patches' published *XYZ* or *L*a*b** values are thus only correct for a physically not-obtainable D50 illuminant, and they become progressively erroneous (i.e. unusable) with increasingly different illumination conditions. One should ideally compute the patches' *XYZ* values using a relevant illumination SPD.
2. There are different ways to solve the mathematically overdetermined system of 72 equations to obtain the nine elements of the *CCM* (Holm, 2006; Molada-Tebar et al., 2018; Molada-Tebar, Marqués-Mateu, & Lerma, 2019a; Molada-Tebar, Riutort-Mayol, et al., 2019; Westland et al., 2012). One can use the Normal Equation to minimise the sum of the squared differences between the sets of *XYZ* and $\mathcal{RGB}$ values:

$$\underline{CCM} = \left[\left(\mathcal{RGB}^T * \mathcal{RGB} \right)^{-1} * \mathcal{RGB}^T * XYZ \right]^T \quad (3)$$

with *T* meaning the transpose. Or, one can compute a more perceptually relevant solution for *CCM* with an optimisation algorithm that minimises colour differences using ΔE_{00} or CIEDE2000, a colour difference metric based on the working principles of the HVS (CIE, 2018). Although summing the rows of matrix *CCM* should ideally result in the illuminant's white point, that is not the case for the previous solutions. That is why Finlayson and Drew (1997) introduced a White-Preserving Normal Equation, resulting in a neutral *CCM* in which the sum of all row elements equals the illumination's white point expressed as *XYZ*. In other words: any pixel with the maximum possible $\mathcal{RGB}$ values [1, 1, 1] in the white-balanced demosaiced RAW image will get *XYZ* coordinates that match those of the illuminant's white point. Figure 16 reflects this by indicating that the *CCM* maps the RAW space reference white back to the adopted white.
3. Since cameras violate the Luther-Ives condition, there is no perfect linear mapping between the camera's raw $\mathcal{RGB}$ DNs to *XYZ* coordinates. Any 3x3 Colour Correction Matrix *CCM* is thus always a compromise, thus explaining why some authors call it the Compromise Colour Matrix (Kasson, 2015). A *CCM* might work well, but there will always be situations where it yields noticeable errors for specific colours. However, non-linear corrections are needed to deal with these outliers and transform the camera space better. Such adjustments are present in camera profiles that can be created according to guidelines of the International Color Consortium (ICC) or via the DCP (DNG Camera Profile) standard that Adobe promulgates (Adobe, 2021; International Organization for Standardization, 2010). ICC and DCP camera profiles have the potential for excellent colour transformation because they contain large lookup tables with colour values to interpolate. However, a 3x3 matrix approach keeps the data linearity intact, which can be necessary for certain subsequent image processing operations. In addition, ICC and DCP profiles are larger, more complex and thus slower than a *CCM*. That is why COOLPI does not include these table-based profiles, thus rendering them also beyond this paper's scope.
4. The choice of target strongly influences the *CCM* (Cao et al., 2008), so it might often be better to leverage thousands of reflectance spectra for camera characterisation (see the lower part of Figure 25). For this approach to work, one needs to know the spectral sensitivity curves of the camera. The most reliable way to determine them is via RAW photographs of a series of narrowband spectra (e.g. from 650 nm to

655 nm, 655 nm to 660 nm). Ideally, a monochromator generates these small wavebands and steps through all wavelengths of interest (Darrodi et al., 2015; Jiang et al., 2013; Verhoeven et al., 2009). Although one can also apply narrowband interference filters instead of a monochromator (Hubel et al., 1994; Mauer, 2009), both procedures are time-consuming and necessitate expensive laboratory-grade equipment. That is why academia also focuses on mathematical spectral sensitivity recovery methods that use numerical optimisation to estimate the response curves from one (or a small set of) RAW photos that record broadband spectra (Finlayson et al., 2016; Walowit et al., 2017). With the response curves, camera raw $\mathcal{RGB}$ values can be computed for unlimited stimuli. Any possible SPD (even from physically unrealisable illumination) can be multiplied with an extensive collection of reflectance spectra to generate these stimuli (see Figure 25). Since the camera gets characterised via thousands of reflectance spectra instead of the ColorChecker's 24 artificial ones, the resulting *CCM* is more accurate and robust (although tests by Jim Kasson (2022) revealed that the differences might not be as dramatic as is often thought). In a way, the spectra-based *CCM* is no longer scene- or chart-dependent but only specific to the camera and illumination. Knowing a camera's spectral sensitivity curves also optimises white balancing, as one finds optimal channel multipliers upon multiplying the illumination SPD with the camera response curves.

6. Musings and Discussion

6.1. What COOLPI Can, Cannot yet, and Will Never Be Able to Do

INDIGO photographers first take a picture of a ColorChecker Passport Photo 2 target (containing, amongst other charts, the 24 ColorChecker patches) before photographing each new graffito. After that, they measure the graffito's spectral illumination. Multiplying that illumination's SPD with the known spectral reflectance of the ColorChecker patches yields 24 illumination-specific *XYZ* triplets. Combining those with the white-balanced, demosaiced pixel values of the ColorChecker RAW photo should yield an accurate *CCM* for the graffito photographs acquired directly after the spectrometer measurement. In addition, this workflow supports two ways to correctly determine the white balance multipliers: indirectly via the illumination's CCT (see 6.2) or directly from one of the target's spectrally neutral grey patches. In the latter case, the multipliers typically are extracted from the second, third or fourth bottom row grey patch starting from the left (see Figure 25).

Instead of a situation-specific *CCM*, the INDIGO team also experiments with workflows that use a fixed D65-based *CCM*, developed under very controlled lighting conditions in a colour cabin. Finally, project INDIGO plans to determine the spectral sensitivity functions of its cameras because it can bring colour accuracy to the next level. However, any of these approaches necessitate a specific data processing workflow. A dedicated RAW development tool like COOLPI is, therefore, essential because it facilitates and automates the testing and application of these different colour-prioritising RAW development workflows (or any better variant the team might develop).

By making COOLPI open-source (https://github.com/GraffitiProjectINDIGO/coolpi) and bundled with lengthy documentation, the authors hope that many other heritage documentation projects can now pay more attention to accurate colour in their digital records. Although proprietary RAW developing software can be easy-to-use, powerful and yield acceptable colour accuracy, their processing is done 'behind closed doors' and according to a rigid scheme. The free Python toolbox COOLPI allows for very flexible and explicit, well-documented processing pipelines, the latter being a fundamental principle of reproducible science.

However, it is crucial to be aware of issues COOLPI currently cannot solve or might never be able to. For instance, many objects—but certainly spray paints—have an anisotropic reflectance, meaning they look different depending on the angle of observation or illumination. In order to fully describe this angular reflectance behaviour, one needs the Bidirectional Reflectance Distribution Function (BRDF) of every object (Schaepman-Strub et al., 2006). However, it is hard to obtain an object's BRDF and even harder to account for it, so COOLPI cannot deal with BRDF-related colour differences or other non-linear imaging factors such as specular reflection, glare, and flare.

COOLPI can also not solve the problem of metamerism.

Because there are only three types of human cones with rather broad spectral sensitivities, stimuli with different SPDs can produce identical cone responses, thus representing the same colour. Such spectra are called metamers. The perceivable colour match is a metameric match, and the effect is known as metamerism (Hung, 2006; König & Herzog, 1999). Because a camera has three broad spectral response curves, it is also a metameric imager (Sharma, 2003). However, since the spectral responses of the standard observer and a digital camera do not match (i.e. cameras do not satisfy the Luther-Ives condition; see Section 5.1), these metameric matches differ (Fairchild et al., 2001). If the camera sensor generates identical raw values for two stimuli a human discriminates, that information is permanently lost.

Finally, COOLPI will likely never remove lens-related effects like chromatic aberrations or distortion (see Section 4.6). On the other hand, COOLPI's following version could support image denoising and flat fielding, the latter being a procedure to account for sensor dust, nonuniformities in the image sensor and optical vignetting (Berry & Burnell, 2005). Although COOLPI's current version 0.1.18 does not consider the camera's spectral response curves, they should be supported in the RAW image processing as soon as project INDIGO has determined them for the cameras. Finally, the INDIGO team is also trying to optimise the crucial white balancing step and account for uneven illumination. Both topics are highly relevant because knowledge about the illumination conditions during photo acquisition drives many RAW processing steps (see also Figure 16). The following two sections will delve slightly deeper into these topics.

6.2. The Importance of Illumination Estimation

On a summer day with relatively few clouds, the illumination's CCT can stay relatively stable for several hours around noon (see Figure 26). The same can be said for the 'strength' of the illumination, denoted illuminance and expressed in lux. As long as the outdoor illuminance stays invariant, camera exposure should not change. However, Figure 26 illustrates that after 16:00 (at least in Vienna at the end of August), the sun increasingly produces warmer light (i.e. the CCT decreases) while the diffuse skylight becomes bluer, quantifiable though its CCT increase. A decreased illuminance indicates that longer shutter speeds or higher ISO values will be needed compared to the illumination conditions around noon.

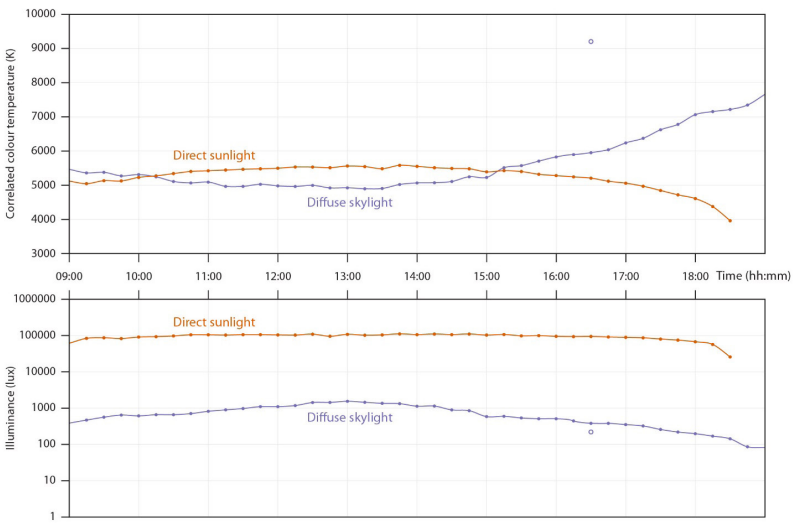


Figure 26. The evolution of the CCT and the illuminance of two outdoor illumination conditions during one sunny, almost cloudless summer day (30-08-2022) in Vienna, Austria. Measurements were performed every 30 minutes with a Sekonic C-7000 SPECTROMASTER portable handheld spectrometer. Diffuse skylight readings occurred in an inner courtyard to avoid any influence of direct sunlight. The Sun's low position prevented direct sunlight measurements after 18:30. The CCT and illuminance graphs of the diffuse skylight indicate an outlier measurement at 16:30. Given the relatively stable illumination conditions, it is considered a measurement artefact. The lower lux value means that the spectrometer's field of view was likely partly blocked by its operator. A data gap at 16:30 was avoided by interpolating a new value from the 16:15 and 16:45 readings.

With the illumination's SPD and CCT measured by the spectrometer, there are various possibilities to get the correct channel multipliers. One can first compute the corresponding (*x*, *y*) chromaticity and then the CIE *XYZ* coordinates for the estimated CCT, or directly obtain the *XYZ* values from the SPD. Multiplying these *XYZ* values with the inverse of the *CCM* (determined without white balancing) yields correct multipliers. However, this approach relies on the correctness of the *CCM*. An easier and maybe more accurate solution is to rely on the colour engineers of the camera manufacturer. Because they know the spectral response of their cameras, ideal CCT-specific multipliers are computed and stored inside each camera, as Figure 28 illustrates for the Nikon Z7ii. One can create such a plot by reading the channel multipliers from a series of RAW images that had their CCT value incrementally changed in the white balance section. Once these multipliers are known, they can be treated as a large look-up table. A third and optimal approach is to integrate the SPD over the spectral sensitivity curves of the camera and normalise the result to the green channel. Finally, one can also read the raw $\mathcal{R}$, $\mathcal{G}$, and $\mathcal{B}$ values of a spectrally neutral object (like a white balance card) that shares the same illumination as the graffito (as Section 4.2 detailed).

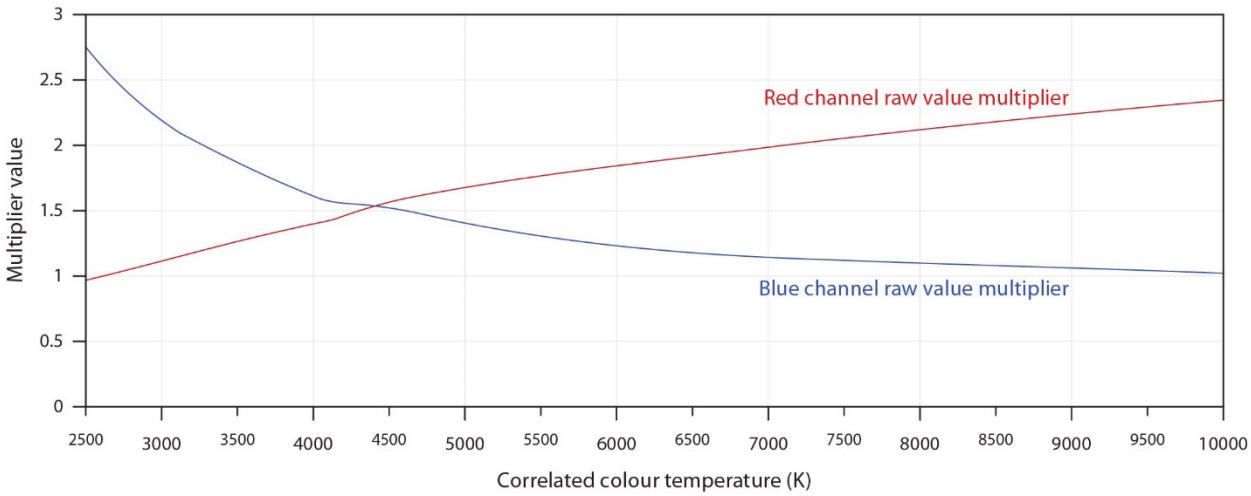


Figure 27. The CCT-specific $\mathcal{R}$ and $\mathcal{B}$ multipliers stored inside a Nikon Z7ii. Since this camera accepts any custom CCT value between 2500 K and 10000 K with a 10 K increment, this graph depicts the 751 channel-specific multipliers. The rate of change in these multiplier values slows down after 6500 K. All values were extracted from the RAW images' Exif metadata via Phil Harvey's free ExifTool (https://exiftool.org).

However, imaging a white balance card or measuring the illumination SPD with a spectrometer just before photographing a graffito might not be all that useful if illumination conditions change quickly. Figure 28 graphs the evolution of the CCT during 90 seconds while the sun is breaking through the clouds. In the first 80 seconds, the CCT drops 700 K, after which it stabilises (see Figure 28, middle row for the detailed view). If photography were to occur during that time, the white balance of the first photos would be very different from the last photographs of that graffito. One could measure the illumination before and after acquiring photographs and interpolate between them to end up with photo-specific CCT values. However, if the CCT change is not gradual throughout image acquisition (for instance, it is stable during a part of the acquisition), this method does not work either. Continuously logging the illumination is also not practical. Not only would it necessitate a spectrometer that can constantly measure and log the data, but one must also ensure that the sampled illumination conditions are always identical to those of the graffito. The only practical solution to deal with illumination changes throughout photo acquisition thus seems to be a software approach, but COOLPI does not contain anything to that end (yet). Or could we rely on the camera to estimate the illumination's CCT?

The lower row of Figure 28 reveals that the illumination estimation by the Nikon Z7ii was very underwhelming. The camera photographed the graffito every second when the sun broke through the clouds; afterwards, its auto-estimated CCT values were retrieved from the Exif metadata. Compared to the reliable spectrometer measurements, the Nikon is, on average, 1500 K wrong, and the CCT trend is not even close to being the opposite of the illuminance curve. As the spectrometer data indicate, one expects the CCT to drop if the proportion of direct solar radiation increases in otherwise cloudy conditions. In summary: for photo colours to be accurate, one needs to rely on a spectrometer measurement or include a white balance card image and hope the illumination's CCT does not change drastically throughout photo acquisition.

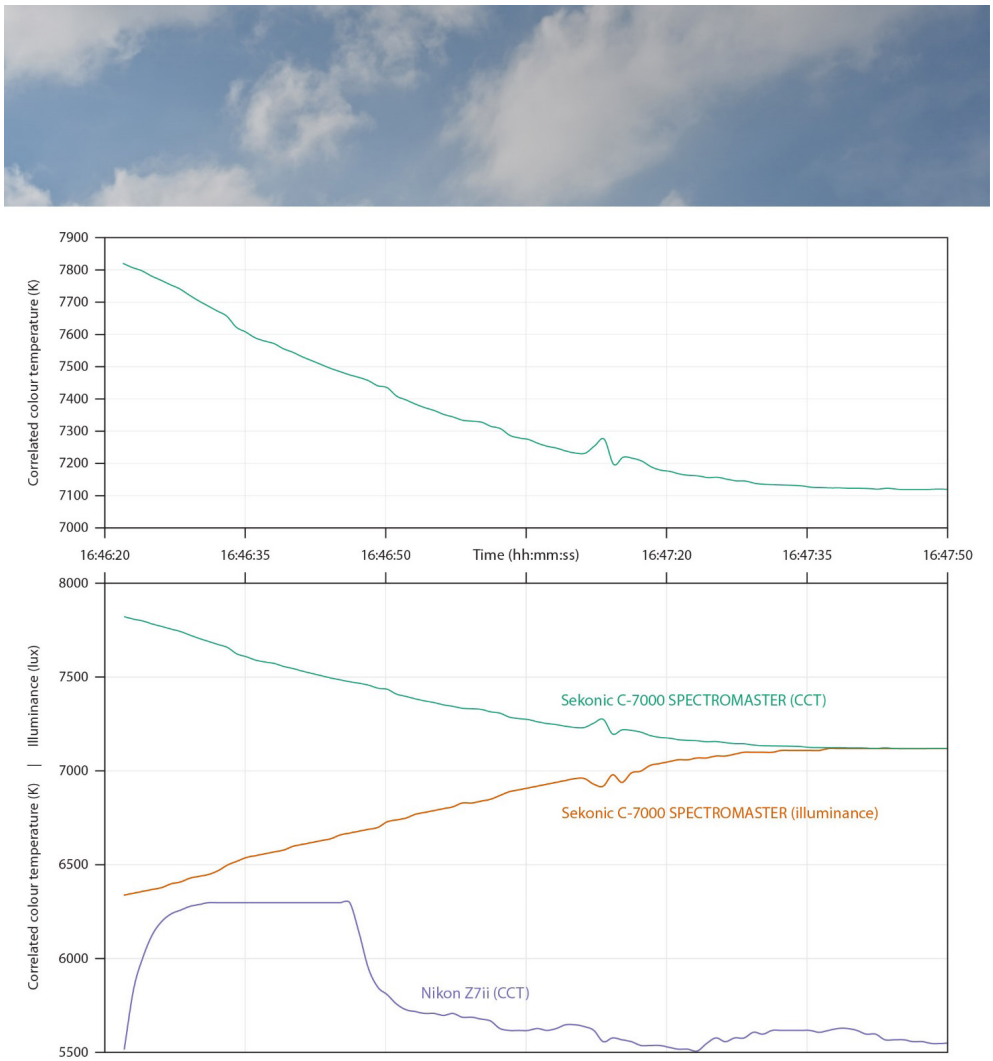


Figure 28. The evolution of the illumination's CCT while the sun breaks through the clouds. The upper part is a photo of the sky conditions. The middle row (i.e. upper graph) depicts in detail the one-second measurements acquired with a Sekonic C-7000 SPECTROMASTER on 05-09-2023 in Vienna (Austria). The spectrometer was mounted on a tripod and pointed in the opposite direction of a graffito. The spectrometer also recorded the illuminance (in lux), which steadily increased while the sun revealed itself more (see the lower graph). By coincidence, the lux values were in the same range as the CCT numbers, which is why both are together on the vertical axis. The CCT estimated by a Nikon Z7ii pointed at the graffito is too low and fails to mimic the real CCT trend.

6.3. One Graffito, Different Illuminations

White balancing can work well if an entire graffito receives one type of illumination. However, problems arise if the subject contains a shaded portion beside one that receives direct solar illumination (see Figure 29). COOLPI can currently not deal with such situations, as one would need to manually or automatically determine the differently illuminated parts and process them separately. However, the authors anticipate such functionality, which is why the INDIGO photographers have been acquiring dual spectrometer readings and ColorChecker Passport Photo 2 images in such situations since the project's start. One can assume that it is not too hard to deal with different illumination conditions like those depicted in Figure 28. However, shadows produced by tree branches lack such evident divisions between the differently illuminated zones. And if those tree branches move in the wind, one can assume it becomes tough to consider these photo-specific shadow patterns.

Figure 29. Different illumination conditions often apply for a more sizeable graffito. In this case, the graffito is located below a bridge which throws part of the creation into the shade.

7. Conclusion

Colour can bring much joy to the world and enlighten one's mood. However, colour is also a tricky phenomenon, and the science of colour is profoundly mathematical. This paper provided an overview of the main concepts needed to understand if and how digital photo cameras can accurately digitise colour. Hopefully, it has become clear that one needs to control and understand every step of the RAW photo processing workflow to ensure a more or less accurate colour recording of heritage objects. To manage that process, the academic graffiti project INDIGO has developed the free and open-source Python toolbox COOLPI. We hope that COOLPI, and the research currently undertaken within the scope of project INDIGO to expand this toolbox, will benefit future actions in graffiti documentation and extend to other scientific fields where recording accurate colour values plays a fundamental role.

Conflict of Interests

The authors declare no conflict of interest.

Acknowledgements

INDIGO is funded by the Heritage Science Austria programme of the Austrian Academy of Sciences (ÖAW).

References

Adobe. (2021). Digital Negative (DNG) Specification: Version 1.6.0.0. Adobe. https://helpx.adobe.com/content/dam/help/en/photoshop/pdf/dng_spec_1_6_0_0.pdf

Bayer, B. E. (1975). Color imaging array (Patent 3971065). USA.

Berger-Schunn, A. (1994). *Practical color measurement: A Primer for the Beginner, A Reminder for the Expert*. *Wiley Series in Pure and Applied Optics*. Wiley.

Berns, R. S. (2001). The Science of Digitizing Paintings for Color-Accurate Image Archives A Review. *The Journal of Imaging Science and Technology*, *45*(4), 305–325.

Berns, R. S. (2019). *Billmeyer and Saltzman's principles of color technology* (Fourth edition). John Wiley & Sons.

Berry, R., & Burnell, J. (2005). *The handbook of astronomical image processing* (Second edition). Willmann-Bell.

Billingsley, F. C. (1965). Digital Video Processing at JPL. In E. B. Turner (Ed.), *Proceedings of SPIE: Vol. 3, Electronic Imaging Techniques I* (p. 15). SPIE. https://doi.org/10.1117/12.970964

Boochs, F., Bentkowska-Kafel, A., Degrigny, C., Karaszewski, M., Karmacharya, A., Kato, Z., Picollo, M., Sitnik, R., Trémeau, A., Tsiafaki, D., & Tamas, L. (2014). Colour and Space in Cultural Heritage: Key Questions in 3D Optical Documentation of Material Culture for Conservation, Study and Preservation. In G. Goos, J. Hartmanis & J. van Leeuwen (Series Eds.) & M. Ioannides, N. Magnenat-Thalmann, E. Fink, R. Žarnić, A.-Y. Yen & E. Quak (Vol. Eds.), *Lecture Notes in Computer Science: Vol. 8740. Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: Proceedings of the 5th International Conference, EuroMed 2014, Limassol, Cyprus, November 3-8, 2014* (pp. 11–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-13695-0_2

Borbély, Á., Sámson, Á., & Schanda, J. (2001). The concept of correlated colour temperature revisited. *Color Research & Application*, *26*(6), 450–457. https://doi.org/10.1002/col.1065

Burggraaff, O., Schmidt, N., Zamorano, J., Pauly, K., Pascual, S., Tapia, C., Spyrakos, E., & Snik, F. (2019). Standardized spectral and radiometric calibration of consumer cameras. *Optics Express*, *27*(14), 19075–19101. https://doi.org/10.1364/OE.27.019075

Camera & Imaging Products Association (2010-2019). *Exchangeable image file format for digital still cameras: Exif Version 2.32* (CIPA DC-X008-2019 / JEITA CP-3451X). Tokyo. CIPA-JEITA. https://www.cipa.jp/std/documents/e/DC-X008-Translation-2019-E.pdf

Cao, F., Guichard, F., & Hornung, H. (2008). Sensor spectral sensitivities, noise measurements, and color sensitivity. In J. M. DiCarlo & B. G. Rodricks (Eds.), *SPIE Proceedings: Vol. 6817, Proceedings of Digital Photography IV* (68170T). SPIE - IS&T. https://doi.org/10.1117/12.766185

Chakrabarti, A., Scharstein, D., & Zickler, T. (2009). An Empirical Camera Model for Internet Color Vision. In A. Cavallaro, S. Prince, & D. Alexander (Eds.), *Proceedings of BMVC 2009* (p. 51). BMVA Press. https://doi.org/10.5244/C.23.51

Chen, Y., Yang, J., Pan, Q., Vazirian, M., & Westland, S. (2020). A method for exploring word-colour associations. *Color Research & Application*, *45*(1), 85–94. https://doi.org/10.1002/col.22434

CIE. (2006). *Fundamental chromaticity diagram with physiological axes - Part 1. CIE technical report: 170-1:2006*. Commission Internationale de l'Eclairage.

CIE. (2015). *Fundamental chromaticity diagram with physiological axes - Part 2: Spectral luminous efficiency functions and chromaticity diagrams. CIE technical report: 170-2:2015*. Commission Internationale de l'Eclairage.

CIE. (2018). *Colorimetry* (Fourth edition). *CIE technical report: 15:2018*. Commission Internationale de l'Eclairage. https://doi.org/10.25039/TR.015.2018

Coffin, D. (2008). *Decoding raw digital photos in Linux*. https://www.dechifro.org/dcraw/

Colour and Vision Research Laboratory. (2021). *Colour & Vision Research laboratory and database*. http://www.cvrl.org/

Cornsweet, T. N. (1970). *Visual perception*. Academic Press.

Darrodi, M. M., Finlayson, G. D., Goodman, T., & Mackiewicz, M. (2015). Reference data set for camera spectral sensitivity estimation. *Journal of the Optical Society of America a*, *32*(3), 381–391. https://doi.org/10.1364/JOSAA.32.000381

Fairchild, M. D., Rosen, M. R., & Johnson, G. M. (08-2001). *Spectral and Metameric Color Imaging* (RIT-MCSL Technical Report). Munsell Color Science Laboratory. http://www.cis.rit.edu/research/mcsl2/research/PDFs/Spec_Met.pdf

Feitosa-Santana, C., Gaddi, C. M., Gomes, A. E., & Nascimento, S. M. C. (2020). Art through the Colors of Graffiti: From the Perspective of the Chromatic Structure. *Sensors*, *20*(9), 2531. https://doi.org/10.3390/s20092531

Finlayson, G. D., Darrodi, M. M., & Mackiewicz, M. (2016). Rank-based camera spectral sensitivity estimation. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *33*(4), 589–599. https://doi.org/10.1364/JOSAA.33.000589

Finlayson, G. D., & Drew, M. S [Mark Samuel] (1997). White-point preserving color correction. In *Proceedings of the Fifth Color Imaging Conference: Color Science, Systems, and Applications: Color science, systems and applications* (pp. 258–261). Society for Imaging Science and Technology.

Guild, J. (1931). The Colorimetric Properties of the Spectrum. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences*, *230*(681-693), 149–187. https://doi.org/10.1098/rsta.1932.0005

Hanada, M. (2018). Correspondence analysis of color–emotion associations. *Color Research & Application*, *43*(2), 224–237. https://doi.org/10.1002/col.22171

Holm, J. (2006). Capture Color Analysis Gamuts. In *Proceedings of the 14th IS&T/SID Color and Imaging Conference (CIC 2006)* (pp. 108–113). IS&T.

Horn, B. K. (1984). Exact reproduction of colored images. *Computer Vision, Graphics, and Image Processing*, *26*(2), 135–167. https://doi.org/10.1016/0734-189X(84)90180-4

Hubel, P. M., Sherman, D., & Farrell, J. E. (1994). A Comparison of Methods of Sensor Spectral Sensitivity Estimation. In *Proceedings of the Second IS&T/SID Color Imaging Conference: Color science, systems, and applications* (pp. 45–48). IS&T.

Hung, P.-C. (2006). Color theory and its applications to digital still cameras. In J. Nakamura (Ed.), *Image sensors and signal processing for digital still cameras* (pp. 205–221). Taylor & Francis.

Hunt, R. W. G. (2004). *The reproduction of colour* (Sixth edition). *The Wiley-IS&T Series in Imaging Science and Technology*. John Wiley & Sons.

Hunt, R. W. G., & Pointer, M. R. (2011). *Measuring colour* (Fourth edition). *Wiley-IS & T series in imaging science and technology*. Wiley.

Iliffe, J., & Lott, R. (2008). *Datums and map projections for remote sensing, GIS, and surveying* (Second edition). Whittles Publishing; CRC Press.

Imai, F. H., Quan, S., Rosen, M. R., & Berns, R. S. (2001). Digital camera filter design for colorimetric and spectral accuracy. In M. Hauta-Kasari, J. Hiltunen, & J. Vanhanen (Chairs), *Proceedings of the Third International Conference on Multispectral Color Science*, University of Joensuu, Finland. http://www.art-si.org/PDFs/Acquisition/MCS01_Imai.pdf

International Organization for Standardization (2010, November 26). *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010* (ISO/TC 130 ISO 15076-1:2010). International Organization for Standardization. https://www.iso.org/standard/54754.html

International Organization for Standardization (2012, October 12). *Graphic technology and photography*

— Colour characterization of digital still cameras (DSCs) — Part 2: Considerations for determining scene analysis transforms (ISO/TC 42 ISO/TR 17321-2:2012). International Organization for Standardization. https://www.iso.org/standard/58003.html

International Organization for Standardization (2022, August 31). *Colorimetry — Part 2: CIE standard illuminants* (CIE ISO/CIE 11664-2:2022). International Organization for Standardization. https://www.iso.org/standard/77215.html

Ives, H. E. (1915). The transformation of color-mixture equations from one system to another. *Journal of the Franklin Institute*, *180*(6), 673–701. https://doi.org/10.1016/S0016-0032(15)90396-4

Jiang, J., Liu, D., Gu, J., & Susstrunk, S. (2013). What is the space of spectral sensitivity functions for digital color cameras? In *Proceedings of the 2013 IEEE Workshop on Applications of Computer Vision (WACV)* (pp. 168–179). IEEE. https://doi.org/10.1109/WACV.2013.6475015

Jonauskaite, D., Wicker, J., Mohr, C., Dael, N., Havelka, J., Papadatou-Pastou, M., Zhang, M., & Oberfeld, D. (2019). A machine learning approach to quantify the specificity of colour-emotion associations and their cultural differences. *Royal Society Open Science*, *6*(9), 190741. https://doi.org/10.1098/rsos.190741

Judd, D. B. (1951). Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight. In *Proceedings of the Twelfth Session of the CIE: Volume 1* (p. 11). Bureau Central de la CIE.

Judd, D. B., MacAdam, D. L., Wyszecki, G., Budde, H. W., Condit, H. R., Henderson, S. T., & Simonds, J. L. (1964). Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature. *Journal of the Optical Society of America*, *54*(8), 1031–1040. https://doi.org/10.1364/JOSA.54.001031

Karaimer, H. C., & Brown, M. S. (2016). A Software Platform for Manipulating the Camera Imaging Pipeline. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *LNCS sublibrary: SL6 - Image processing, computer vision, pattern recognition, and graphics: Vol. 9905. Computer Vision – ECCV 2016: Proceedings of the 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016. Part I* (pp. 429–444). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_26

Karaimer, H. C., & Brown, M. S. (2019). Beyond raw-RGB and sRGB: Advocating Access to a Colorimetric Image State. *Color and Imaging Conference*, *27*(1), 86–90. https://doi.org/10.2352/issn.2169-2629.2019.27.16

Kasson, J. (2015). *Constructing a compromise matrix*. https://blog.kasson.com/the-last-word/constructing-a-compromise-matrix/

Kasson, J. (2022). *Training on the ColorCheckers, testing on natural spectra*. https://blog.kasson.com/the-last-word/training-on-the-colorcheckers-testing-on-natural-spectra/

Kirchner, E., van Wijk, C., van Beek, H., & Koster, T. (2021). Exploring the limits of color accuracy in technical photography. *Heritage Science*, *9*(1), Article 57. https://doi.org/10.1186/s40494-021-00536-x

König, F., & Herzog, P. G. (1999). On the Limitations of Metameric Imaging. In *Pics 1999: Proceedings of the Conference on Image Processing, Image Quality and Image Capture Systems (PICS-99)* (pp. 163–168). Society for Imaging Science and Technology.

Korytkowski, P., & Olejnik-Krugly, A. (2017). Precise capture of colors in cultural heritage digitization. *Color Research & Application*, *42*(3), 333–336. https://doi.org/10.1002/col.22092

Lam, E. Y., & Fung, G. S. (2009). Automatic White Balancing in Digital Photography. In P. A. Laplante (Series Ed.) & R. Lukac (Vol. Ed.), *Image processing series: Vol. 9. Single-sensor imaging: Methods and applications for digital cameras* (pp. 267–294). CRC Press. https://doi.org/10.1201/9781420054538.ch10

Lam, K. M. (1985). *Metamerism and colour constancy* [PhD., University of Bradford, Bradford]. Original source.

Lee, B. B. (2004). Paths to colour in the retina. *Clinical and Experimental Optometry*, *87*(4-5), 239–248. https://doi.org/10.1111/j.1444-0938.2004.tb05054.x

Lee, H.-C. (2005). *Introduction to color imaging science*. Cambridge University Press.

LibRaw LLC. (2021). *About LibRaw*. https://www.libraw.org

Livingstone, M. (2002). *Vision and art: The biology of seeing*. Abrams.

Luther, R. (1927). Aus dem Gebiet der Farbreizmetrik. *Zeitschrift Für Technische Physik*, *8*, 540–558.

Lyon, R. F., & Hubel, P. M. (2002). Eyeing the Camera: Into the Next Century. In *Final Program and Proceedings of the 10th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications* (pp. 349–355). IS & T.

Malacara, D. (2011). *Color vision and colorimetry: Theory and applications* (Second edition). *Press monograph: Vol. 204*. SPIE.

Malvar, H. S., He, L., & Cutler, R. (2004). High-quality linear interpolation for demosaicing of Bayer-patterned color images. In B. Mercer (Ed.), *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing: Volume V* (iii-485-iii-488). IEEE. https://doi.org/10.1109/ICASSP.2004.1326587

Mandl, F., & Shaw, G. (2010). *Quantum field theory* (Second edition). John Wiley & Sons Ltd.

Mauer, C. (2009). *Measurement of the spectral response of digital cameras with a set of interference filters* [Master thesis, University of Applied Sciences Cologne, Cologne, Germany]. Original source.

Molada-Tebar, A., Lerma, J. L., & Marqués-Mateu, Á. (2018). Camera characterization for improving color archaeological documentation. *Color Research & Application*, *43*(1), 47–57. https://doi.org/10.1002/col.22152

Molada-Tebar, A., Marqués-Mateu, Á., & Lerma, J. L. (2019a). Camera Characterisation Based on Skin-Tone Colours for Rock Art Recording. *Proceedings*, *19*(1), 12. https://doi.org/10.3390/proceedings2019019012

Molada-Tebar, A., Marqués-Mateu, Á., & Lerma, J. L. (2019b). Correct use of color for cultural heritage documentation. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W6*, 107–113. https://doi.org/10.5194/isprs-annals-IV-2-W6-107-2019

Molada-Tebar, A., Riutort-Mayol, G., Marqués-Mateu, Á., & Lerma, J. L. (2019). A Gaussian Process Model for Color Camera Characterization: Assessment in Outdoor Levantine Rock Art Scenes. *Sensors*, *19*(21), 4610. https://doi.org/10.3390/s19214610

Nimeroff, I., Rosenblatt, J. R., & Dannemiller, M. C. (1961). Variability of Spectral Tristimulus Values. *Journal of Research of the National Bureau of Standards Section a: Physics and Chemistry*, *65A*(6), 475–483. https://doi.org/10.6028/jres.065A.050

Ohno, Y. (2000). Cie Fundamentals for Color Measurements. In H. Mizes & M. Yuasa (Eds.), *Proceedings of IS&T's NIP16 International Conference on Digital Printing Technologies* (pp. 540–545). The Society for Imaging Science and Technology.

Ohno, Y. (2010). Radiometry and Photometry for Vision Optics. In M. Bass, C. M. DeCusatis, J. M. Enoch, V. Lakshminarayanan, G. Li, C. A. MacDonald, V. N. Mahajan, & E. W. van Stryland (Eds.), *Handbook of Optics: Volume II. Design, Fabrication, and Testing; Sources and Detectors; Radiometry and Photometry* (p. 37). McGraw-Hill.

Palomar-Vazquez, J., Baselga, S., Viñals-Blasco, M.-J., García-Sales, C., & Sancho-Espinós, I. (2017). Application of a combination of digital image processing and 3D visualization of graffiti in heritage conservation. *Journal of Archaeological Science: Reports*, *12*, 32–42. https://doi.org/10.1016/j.jasrep.2017.01.021

Parulski, K. A., & Spaulding, K. (2003). Color Image Processing for Digital Cameras. In G. Sharma (Ed.), *Electrical engineering and applied signal processing series. Digital color imaging handbook* (pp. 727–757). CRC Press. https://doi.org/10.1201/9781420041484.ch12

Punnappurath, A., & Brown, M. S. (2020). Learning Raw Image Reconstruction-Aware Deep Image Compressors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(4), 1013–1019. https://doi.org/10.1109/TPAMI.2019.2903062

Ramanath, R., Snyder, W. E., Yoo, Y., & Drew, M. S. (2005). Color image processing pipeline: A general survey of digital still camera processing. *IEEE Signal Processing Magazine*, *22*(1), 34–43. https://doi.org/10.1109/MSP.2005.1407713

Rowlands, D. A. (2020a). Color conversion matrices in digital cameras: A tutorial. *Optical Engineering*, *59*(11), 110801. https://doi.org/10.1117/1.OE.59.11.110801

Rowlands, D. A. (2020b). *Physics of digital photography* (Second edition). *IOP series in emerging technologies in optics and photonics*. IOP Publishing.

Schaepman-Strub, G., Schaepman, M., Painter, T., Dangel, S., & Martonchik, J. (2006). Reflectance quantities in optical remote sensing—Definitions and case studies. *Remote Sensing of Environment*, *103*(1), 27–42. https://doi.org/10.1016/j.rse.2006.03.002

Schreiber, W. (1967). Picture coding. *Proceedings of the IEEE*, *55*(3), 320–330. https://doi.org/10.1109/PROC.1967.5488

Sharma, G. (2003). Color fundamentals for digital imaging. In G. Sharma (Ed.), *Electrical engineering and applied signal processing series. Digital color imaging handbook* (pp. 1–114). CRC Press. https://doi.org/10.1201/9781420041484.ch1

Speranskaya (1959). Determination of spectrum color co-ordinates for twenty seven normal observers. *Optics and Spectroscopy*, *7*, 424–428.

Stiles, W. S., & Burch, J. M. (1959). N.P.L. Colour-matching Investigation: Final Report (1958). *Optica Acta: International Journal of Optics*, *6*(1), 1–26. https://doi.org/10.1080/713826267

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

Sumner, R. (2014, February 14). *Processing RAW Images in MATLAB*. Santa Cruz. Department of Electrical Engineering, UC Santa Cruz. http://users.soe.ucsc.edu/~rcsumner/rawguide/RAWguide.pdf

Tominaga, S., Nishi, S., & Ohtera, R. (2021). Measurement and Estimation of Spectral Sensitivity Functions for Mobile Phone Cameras. *Sensors*, *21*(15), 4985. https://doi.org/10.3390/s21154985

Torger, A. (2018). *Dcamprof - a digital camera profiling tool*. https://torger.se/anders/dcamprof.html#tone_curves

Valberg, A. (2005). *Light vision color*. John Wiley & Sons.

Velikonja, M. (2020). *Post-Socialist political graffiti in the Balkans and Central Europe. Southeast European Studies*. Routledge.

Verhoeven, G. J. (2010). It's all about the format – unleashing the power of RAW aerial photography. *International Journal of Remote Sensing*, *31*(8), 2009–2042. https://doi.org/10.1080/01431160902929271

Verhoeven, G. J. (2016). Basics of photography for cultural heritage imaging. In E. Stylianidis & F. Remondino (Eds.), *3d Recording, Documentation and Management of Cultural Heritage* (pp. 127–251). Whittles Publishing.

Verhoeven, G. J. (2018). Pixels – So basic but so confusing. *AARGnews*, *56*, 28-33. https://doi.org/10.5281/zenodo.4024319

Verhoeven, G. J., Smet, P. F., Poelman, D., & Vermeulen, F. (2009). Spectral Characterization of a Digital Still Camera's NIR Modification to Enhance Archaeological Observation. *IEEE Transactions*

on Geoscience and Remote Sensing, *47*(10), 3456–3468. https://doi.org/10.1109/TGRS.2009.2021431

Verhoeven, G. J., Wild, B., Schlegel, J., Wieser, M., Pfeifer, N., Wogrin, S., Eysn, L., Carloni, M., Koschi□ek-Krombholz, B., Molada-Tebar, A., Otepka-Schremmer, J., Ressl, C., Trognitz, M., & Watzinger, A. (2022). Project INDIGO – document, disseminate & analyse a graffiti-scape. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLVI-2/W1-2022*, 513–520. https://doi.org/10.5194/isprs-archives-XLVI-2-W1-2022-513-2022

Walker, J. S. (2004). *Physics* (2nd International Edition). Pearson Education.

Walowit, E., Buhr, H., & Wüller, D. (2017). Multidimensional Estimation of Spectral Sensitivities. In *Proceedings of the Twenty-fifth Color and Imaging Conference (CIC 25)* (pp. 1–6). IS&T.

Wandell, B. A. (1995). *Foundations of vision*. Sinauer Associates.

Westland, S., Ripamonti, C., & Cheung, V. (2012). *Computational colour science using MATLAB* (Second edition). *Wiley-IS&T series in imaging science and technology*. John Wiley & Sons.

Wright, W. D. (1928-29). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, *30*(4), 141–164. https://doi.org/10.1088/1475-4878/30/4/301