9

Color Measurement and Calibration in Medical Photography

Elizabeth Allen

Contents

9.1 The Physics of Color

In everyday language humans ascribe certain colors to objects or light sources, for example, “the sky is blue.” However, color is not an object attribute but rather an attribute of visual sensation; it cannot exist without an observer which has a means of both detecting and interpreting radiation in the form of color, whether the human visual system or an image capture device.

As described in Chap. 11, the human perception of color arises from the response and interpretation by the human visual system (HVS) to radiation in the visible part of the electromagnetic (EM) spectrum, which ranges between wavelengths of 380 and 720 nm. The EM spectrum extends far beyond that range, and various medical imaging modalities make use of this, including ultraviolet and x-ray radiation at the short wavelength, high-frequency end of the spectrum, and infrared radiation which occupies the region just beyond the long wavelength and low-frequency portion of the visible spectrum. In

E. Allen (✉)
Computational Vision and Imaging Technologies
Research Group, University of Westminster,
London, UK
e-mail: E.allen01@westminster.ac.uk

© Springer Nature Switzerland AG 2020

P. Pasquali (ed.), *Photography in Clinical Medicine*, https://doi.org/10.1007/978-3-030-24544-3_9

the human observer, the retinal photoreceptors detect the physical stimulus (electromagnetic radiation), and then the neural connections in the visual pathway, and the cognitive system, process and interpret the signals produced by the stimulus [1]. The perception of color is as a result of the integrated response of the three different retinal cone receptors (ρ, γ, β) [2].

The observer is the final component of what is termed the *triangle of color* [3], which describes what happens when emitted light falls upon a surface and is then observed. The first component is a source of visible electromagnetic energy, which will have its own spectral signature. The second component is an object, the chemical and physical properties of which modulate the energy from the source. The final color forms from the combination of all three components of the triangle.

Image capture within an imaging system may be modelled in a similar manner to an observer, replacing the receptors of the visual system with those of the imaging system. Aside from the variability of the colored surface itself, variability in the illuminating light and the color responses of the receptors have an impact on the final color seen or imaged. The color of the object itself therefore cannot be considered in isolation. To further complicate matters, the viewing geometry also affects color appearance.

Visualizing the color triangle as a series of spectral distributions can be useful in understanding how the appearance of a final color is produced. The *spectral power distribution* (SPD) is the relative spectral power output measured at regular increments across the visible spectrum The SPDs of a series of standard light sources are shown in (Fig. 9.1).

The color produced when a surface is illuminated, occurs as a result of the material properties of the surface, selectively absorbing and reflecting (or transmitting) the wavelengths present by different amounts. Because the color produced is a combination of the spectral properties of the illuminant and the surface, and the spectral responsivities of the receptors, it is possible for two different colors to appear to match under some illuminants (see Fig. 9.2). This phenomenon is known as *metamerism*, and it creates complexity in color reproduction and color management but is also the basis for trichromatic matching which allows the appearance of colors to be matched under diverse viewing conditions and with differing white points.

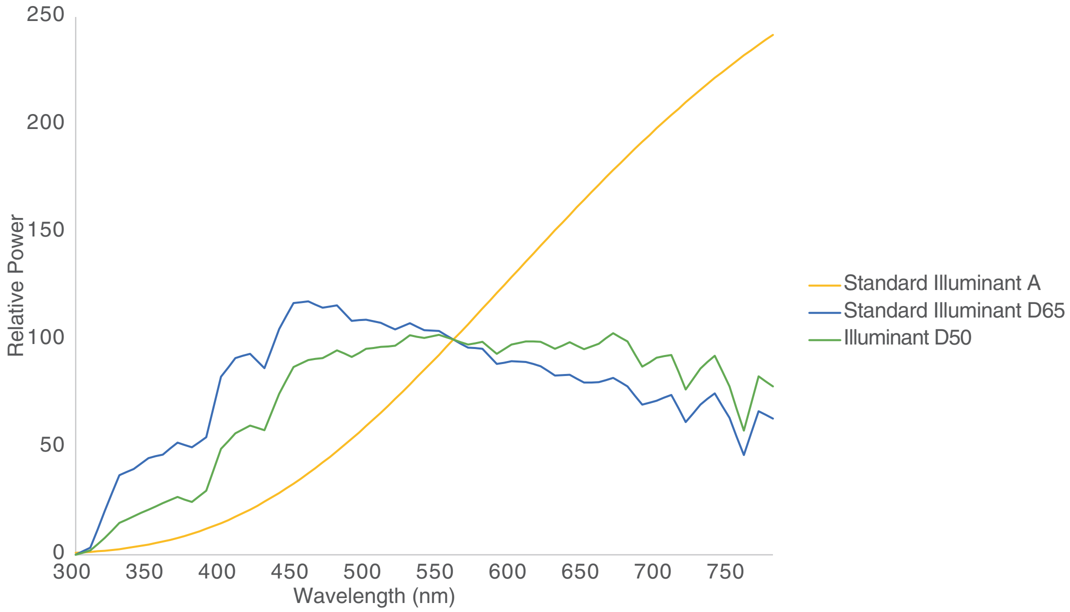

Fig. 9.1 Spectral Power Distributions of Different Illuminants. These SPDs are computed from the CIE data for Standard Illuminant A (to simulate incandescent light sources), Standard Illuminant D65 (simulating mean noon daylight) and Illuminant D50. *Data sourced from CIE 15:2004 Colorimetry standard published by the Commission Internationale de l'Eclairage*

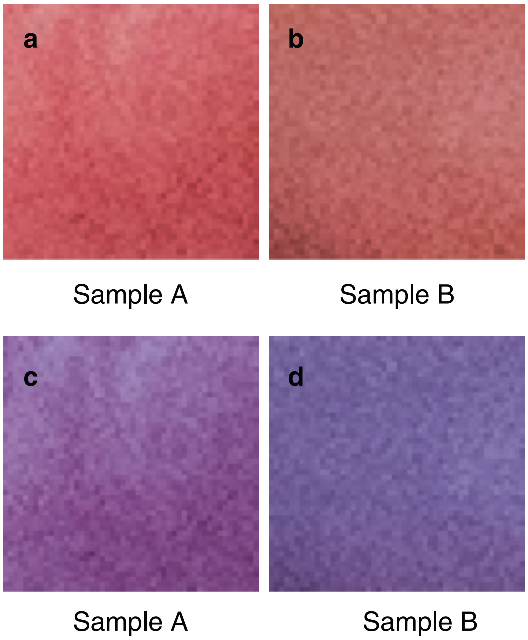

Fig. 9.2 Metamerism. Two equally sized areas from the same image simulating the effects of image capture under an incandescent light source and daylight. The color of the patches is similar under the first illuminant (top row) but shows a more significant hue shift under the second illuminant (bottom row)

9.2 The Appearance of Colors

The color triangle explains what is happening in terms of spectral properties during color capture or viewing, but there are numerous other factors that influence how a color appears, and awareness of these is important in understanding the need for standardization wherever possible, in medical photography and other color critical applications.

As soon as a color is viewed in context, surrounded by other colors, for example, within an image, or on a colored background, or illuminated by a source of a particular spectral quality, or on a glossy instead of a matt paper surface, these environmental conditions will affect color appearance. The human visual system is adept at accounting for changes as a result of viewing conditions [1] and has evolved a number of adaptation mechanisms to assist with object identification and interpretation. One of these is *color constancy*, an adaptation that ensures that the perception of colors remains relatively constant in appearance under changing lighting conditions, for example, when illumination levels change, or under light sources with different spectral characteristics. Human perception of color and particularly perceptual differences between colors are highly influenced by their perceived relationship to white or the light source; hence this mechanism has a significant impact upon color perception. This process is sometimes termed *discounting the illuminant* [3] where the visual system automatically incorporates the white point of the illuminant, effectively shifting the perception of all colors in response to this. Color constancy is believed to be a combination of two effects: *chromatic adaptation* and *memory colors* [3]. Chromatic adaptation may be viewed as a form of automatic white balance, whereby the relative sensitivity of the cone receptors in the eye changes in response to the wavelengths within a light source; for example, under a light source containing more blue wavelengths, such as noon daylight, an increase of the L-cone sensitivity and a decrease in the S-cone sensitivity will ensure that white objects will appear white rather than blue. Memory colors are prototypical colors associated with particular highly recognizable objects; these are individual to the observer, but certain trends are common, for example, grass is typically remembered with more of a green hue than it has [4].

There are other color appearance phenomena that may affect images when viewed separately or on different systems. *Simultaneous contrast* describes the shift in color appearance as a result of the color or tone of the background [5]. Apparent shifts are based upon the opponent theory of color vision, whereby the signals from the three cone receptors are combined to form three opponent signals, red-green, blue-yellow, and light-dark [6]. Simultaneous contrast results in images on a lighter background appearing to be darker; those on a dark background appear lighter, red colors make the colors next to them appear greener and vice versa, and blue colors make colors next to them appear yellower and vice versa. This phenomenon highlights the necessity to control viewing conditions at output; whether viewing on a display or on a print, the background color and lightness will affect the image appearance in terms of color and contrast.

Initial object recognition is generally driven by the spatial and tonal properties of images, fundamentally related to global and local contrast characteristics. The visual system has light and dark adaptation mechanisms to respond to differing illumination levels and is far more sensitive to tonal differences than color differences; once object recognition has occurred, color adaptation mechanisms, as described above, affect color appearance. These mechanisms are important as color evaluation almost always involves comparison of colors, either against another colored object or against a memory object.

The various adaptation mechanisms work well for the human visual system, but without their effects being built into color reproduction of images, they can produce unpredictable results when viewing images produced or viewed under different conditions. Where images are to be reproduced on the same medium for comparison, and the viewing conditions and the white point of the illuminant are the same, color appearance phenomena are less likely to influence the evaluation. But as soon as there is a change in medium (e.g., from display, which is emissive, to print, which is reflective), the contrast characteristics and the white point may potentially change. Additionally, printed images are perceived as illuminated objects, and therefore the visual system discounts the illuminant, whereas this mechanism will not occur with the display, because it is self-luminous, and no known illuminant is present [3]. Essentially different color appearance mechanisms are at work when images are displayed upon different media.

9.3 Color Reproduction in the Digital Imaging Chain

The imaging chain encompasses the devices and encoding of image data from input to output. As described by Triantaphillidou [7], a color space encoding is the digital encoding of a color space, which is an *n*-dimensional coordinate system to specify the position of a color within the space. The color space encoding describes the digital encoding of the underlying continuous color space, for example, the bit depth and quantization, which determines the range of values that may be taken within each color channel. Color spaces may be classified according to the method of color reproduction (additive such as RGB or subtractive such as CMY), whether they are specific to a device and whether they have a defined relationship with human color perception.

In a typical imaging chain, image data may be transformed through a number of different color encodings, as illustrated in (Fig. 9.3), the nature of which are related to the image format, the characteristics of the device, and color management decisions made based upon the color reproduction objectives of the application.

Digital color reproduction requires the transformation of image data values between color models and color spaces in a manner that repro-

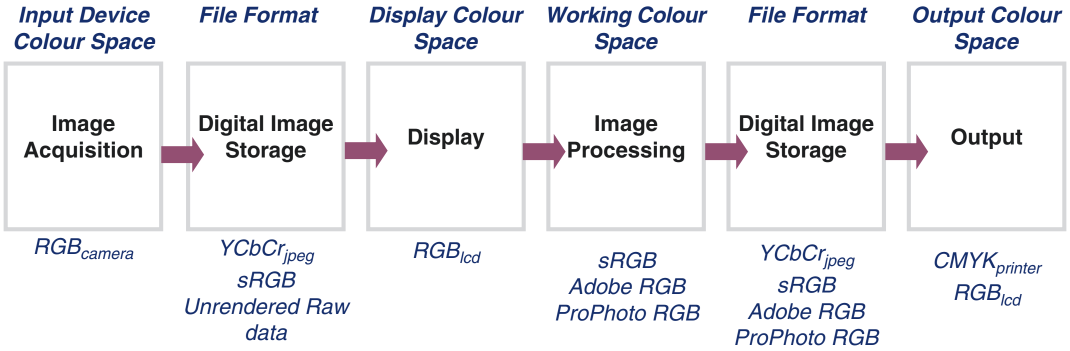

Fig. 9.3 Color through the imaging chain. A typical digital imaging chain, from capture to output, with examples of the possible color spaces that the image may be encoded in to at different stages

duces the appearance of the colors as accurately as possible. The color spaces native to input and output devices are *device-dependent* color spaces; the coordinates used to represent the colors are defined by the nature of the color space, and the color gamut is based upon the characteristics of the device, determined by the following factors:

9.3.1 Spectral Characteristics of the Device Primaries

All digital color devices have a set of primaries, which are the set of colors, which, when mixed in different proportions, will produce other colors. These determine the color channels used to create color and are chosen to produce the widest possible color gamut in a given imaging system. In input and display devices, the primaries are additive RGB, meaning that white is created as an additive combination of the three red, green, and blue primaries at their maximum. In print devices, the color space is subtractive, meaning that the color channels subtract from the reflected illuminant spectrum; white is produced when the color channels are at a minimum.

The primaries in a digital camera are defined by the spectral responsivities of the filtered CCD or CMOS pixels, so they are a result of the sensor sensitivity and the spectral transmissivity of the overlaid RGB filters. In a scanning device, and in other illuminated optical devices, such as dermatoscopes or endoscopes, the primaries are a result of the RGB-filtered sensors and the spectral characteristics of the illuminant. In the latter case, color rendition has also been found to be affected by illumination intensity [8]. Display device primaries are defined by each of the three color channels when at a maximum (e.g., when a pure red pixel is displayed, which in a 24 bit color system will have a value of [255, 0, 0]); in LCD devices, they are defined by the spectral transmission characteristics of the filters overlaying the backlight.

Color printers use a subtractive set of primaries, in the form of dye-based or pigment-based inks in the case of inkjets or dye particles in the case of dye sublimation printers. These may be three channels of cyan, magenta, and yellow or have a fourth black "key" channel (the "k" in CMYK) to improve contrast. The primaries are a function of the spectral reflectance of each individual dye or ink layer at its maximum saturation combined with the spectral reflectance of the underlying substrate (the paper color). When the full saturation of all three- or four-color channels is overlaid, they subtract light to produce black.

9.3.2 White Point

The white point has different characteristics, depending on the type of device and whether it incorporates an illuminating device. The white point in a digital camera is related to the scene conditions, i.e., the spectral qualities of the illuminant, but also the white point that the camera is adapted to, which may be achieved by adjusting the analogue gain of the three channel signals [9] or by post-processing the digital signal during image rendering. The white points of scanners, dermatoscopes, and endoscopes are defined by the spectral characteristics of the unfiltered illuminant. In a display incorporating a backlight, the white point is defined by the spectral characteristics of the backlight when the color filters are completely turned off (meaning that all the light is emitted from the display). The white point of a printer is determined by the white point of illuminating light and the spectral reflectance distribution of the paper.

9.3.3 Transfer Functions

The transfer function is the device tone reproduction function, sometimes called the gamma function; it defines the range of tones and the contrast of the device and the relationship between the input and output intensities. Each color channel has its own transfer function, and these are set during the calibration of the system (e.g., by changing the gamma value during display calibration). A change to one or more of the transfer functions will result in significant color shifts as the relationship between the color channels changes and may result in a color cast within neutral greys.

Device-dependent color spaces are not only variable with the make, model, and spectral characteristics of the primaries of the devices; there can be variability between devices of the same model and in reproduction from the same device over time. The variability may be as a result of different hardware or software settings (e.g., through different white point setting or gain adjustment of individual channels) or as a result of aging of device components (e.g., the change in white point as the backlight source ages) or color drift (e.g., in an inkjet printer a change in the amount of different inks being laid down as a result of print head blockages). A reduction in device variability is one of the purposes of calibration, as described later in this chapter.

A further limitation of device-dependent color spaces is the lack of a defined relationship with the response of the HVS and a lack of perceptual uniformity in terms of color differences.

Device dependence is one of the key issues to be addressed through color management. Early color management was achieved between two devices (e.g., from a scanner to a printer) by measuring the color reproduction of a specific set of colors from the two devices and creating a direct color transform that could be applied on the data from the input device to correctly reproduce the colors on the output device. This *closed loop* color management worked well between two devices but required detailed knowledge of the two devices and an experienced operator to monitor color drift and recalibrate regularly [10]. Such a system may be suitable in a medical photography context, for example, if the images are always to be captured in controlled lighting conditions and always viewed on the same display, without the need for transmission or archiving. But in many contexts, this is an unsatisfactory solution, which relies on unchanging technology and does not future proof the images.

As a result of the growth in the use of true color images in medical imaging, image colors need to be translatable and consistent between multiple different devices, different imaging chains, to different output devices, and, in telemedicine, between different imaging locations. The growth in fields such as teledermatology and telepathology allows access for patients to specialists in distant locations, but poor color translation can impede diagnosis [11]. As in consumer digital imaging, closed loop color management is inadequate for situations with multiple devices, rapidly growing in complexity, and therefore a standardized framework for color reproduction and management is required. The International Color Consortium (ICC) architecture provides this through open loop color management, which uses an intermediate device-independent color space into which all device color spaces can be translated (Fig. 9.4).

The work of the ICC Medical Imaging Working Group [12], introduced in Chap. 11, aims to bring together different working practices from disparate specialisms to define standards and guidelines for color reproduction. Generally, it has been found that color within various clinical specialisms tends to be managed in a rather ad hoc manner. Digital microscopy, telemedicine, medical photography (particularly ophthalmic, dental, and dermatology), and display calibration have been identified as priority areas in which improvements in the consistency and standardization of color reproduction would produce significant and tangible benefits [13]. To date the ICC MIWG have produced a number of white papers, reports, and other resources with recommendations for best practice in different areas, including pathology, medical displays, a color eye model, digital photography, and a range of other areas of interest.

9.4 CIE Colorimetry and CIE Color Spaces

The measurement of color using colorimetry is at the heart of color reproduction and ICC color management. It aims to define color measurements and specifications that directly relate to human color perception. The Commission Internationale d'Eclairage (CIE) approaches to colorimetry and CIE color spaces are device independent, providing absolute color specification and are therefore used as intermedi-

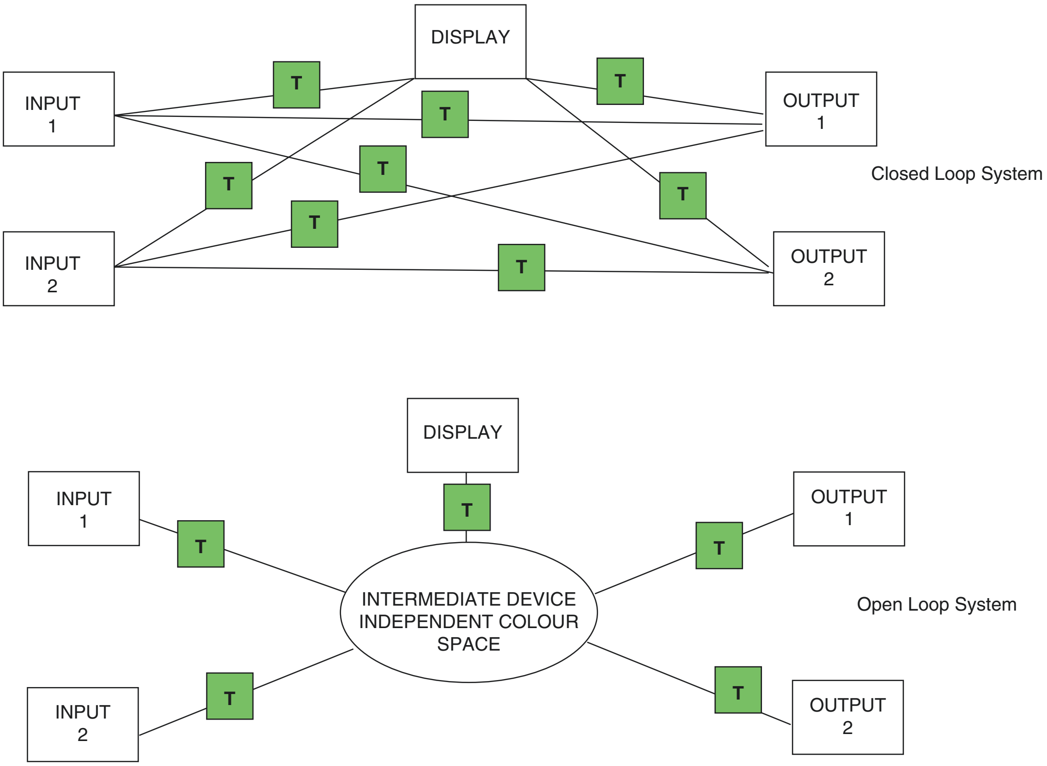

Fig. 9.4 Open loop color management uses a central connecting color space to reduce the required number of color transforms, compared to closed loop color management systems

ate connection spaces in open loop color management.

The CIE model of colorimetry originally derived from a color matching experiment implemented in 1931, in which observers were required to match various colored patches by mixing and changing the amounts of three projected superimposed illuminants of short, medium, and long wavelengths. The experimental data were used to obtain the CIE standard observer color-matching functions, which are the theoretical chromatic responses of the average observer across the visible spectrum. Initially expressed as the combination of responses of three monochromatic lights of wavelengths R = 700 nm, G = 546.1 nm, and B = 435.8 nm, these were mathematically transformed to obtain a set of imaginary primaries *X*, *Y*, and *Z*, which could be combined in different proportions to match all possible perceived colors. The color matching functions for the CIE 1931 2° standard observer [14] are illustrated in (Fig. 9.5). Colorimetry defines colors in terms of the proportion of the three primaries, the *XYZ* tristimulus values [15], and may be measured using a spectrophotometer or more commonly a colorimeter.

The original tristimulus values defined from the 1931 experiment are still used in International Color Consortium (ICC) compliant color management systems. Further experiments confirmed the validity of the data and were extended to obtain a dataset for a 10° viewing angle, the CIE 1964 10° standard colorimetric observer.

The CIE *XYZ* color space is a three-dimensional coordinate system in which all colors may be expressed as a combination of the three coordinates; it also relates to the human visual response, whereby each of the *XYZ* color-matching functions may be considered as a linear combination of the cone responsivities. At the heart of color management is the process of transforming colors in the real world and in

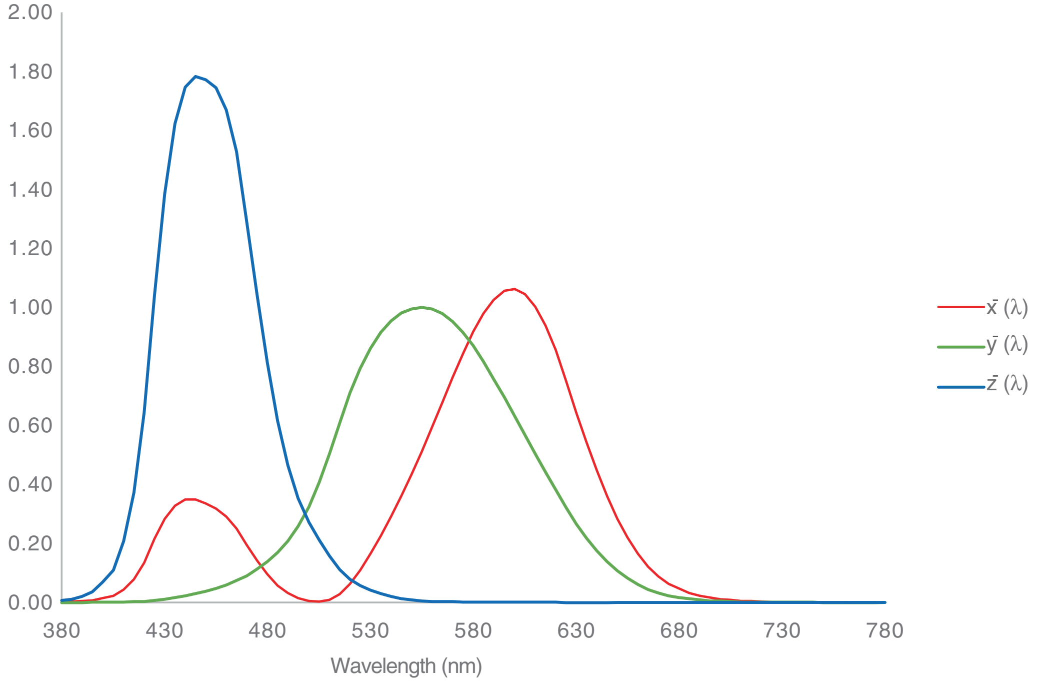

Fig. 9.5 CIE 1931 2° standard observer color matching functions Data sourced from Colorimetry *Part 1: CIE standard colorimetric observers (ISO/CIE 11664-1:2019) published by the Commission Internationale de l'Eclairage*

device-dependent color spaces into colorimetric values. Expressing colors in terms of colorimetric values provides an absolute system of measurement in which colored objects and the white points of illuminants may be mapped, compared, matched, and discriminated in a meaningful manner independently of the devices or media upon which they are reproduced. By doing so, it is possible to transform and reproduce colors on different devices and under different viewing conditions, which will be perceptually as close to the original as possible, within certain defined limits.

The transformations to obtain tristimulus values are designed such that the Y tristimulus value represents the luminance of the color. The XYZ values may be further transformed into chromaticity coordinates as follows [14]:

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = \frac{Z}{X+Y+Z}$$

As x, y, and z are normalized values, only two of them are required to represent a color, and they can therefore be represented on a two-dimensional diagram, which is somewhat easier to understand than a 3D space. An x, y chromaticity diagram provides color information but no luminance information.

The CIE x, y diagram is often used to represent color gamuts of imaging systems for comparison; however this is not without problems, because it is not perceptually uniform, meaning that equal distances between colors in different areas of the chromaticity diagram are unequal in terms of perceived difference. A further transformation of X, Y, Z values in 1976 produced the CIELUV [16] and CIELAB [17] color spaces. These three-dimensional spaces represent colors in terms of lightness and two chromaticity coordinates. The CIE 1976 uniform chromaticity scales diagram, derived from the CIELUV space, represents colors in terms of the u', v' chromaticity coordinates in which the differences are nearly perceptually uniform.

The CIELAB space is used widely in color imaging applications (6,6) and is defined from the XYZ values, incorporating their relationship with the XYZ values of the illuminant white point.

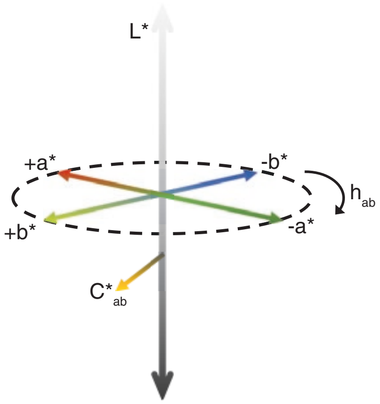

Fig. 9.6 Three-dimensional representation of the CIELAB uniform colour space. The vertical L* axis corresponds to the lightness of a color and the color coordinates a* and b* correspond to redness-greeness and blueness-yellowness respectively. Cylindrical coordinates, C*ab and H*ab are also illustrated; the distance of a color from the central L* axis defines how chromatic it is and the angle around the centre determines the hue

The resulting coordinates L^*, a^*, and b^* correspond to lightness and two color coordinates. The a^* and b^* values define the color on scales which approximate perceptions of red-green (positive a^* to negative a^*) and yellow-blue (positive b^* to negative b^*). From these values, the perceived chroma $C^*_{a,b}$ and perceived hue $H^*_{a,b}$ may be calculated (Fig. 9.6). Defining a color in terms of hue, lightness, and chroma is undoubtedly more intuitive and easier to understand and visualize for the human observer than expressing them as LAB or *XYZ* (or even RGB) values.

9.5 Standard RGB Color Encoding

There are a number of other color spaces that refer to devices, either real or virtual, but with a defined relationship with CIE colorimetric spaces, so that they can be considered to be device independent or device standardized. They are used widely in digital imaging and implemented in imaging software to enable standard color space encoding.

sRGB (standard RGB), introduced in Chap. 11, is one such color space relevant to medical photography. It is an international standard, published by the International Electrotechnical Commission (IEC) [18], and is a type of *output-referred* color space encoding. Output-referred color space encodings are linked to specific real or virtual output devices and viewing conditions [7]. In the case of the sRGB encoding, it is based on typical cathode ray tube (CRT) display primaries and transfer function. Originally developed as a default color encoding for the Internet in 1999, it has found widespread adoption in various imaging industries, including medical photography; if two displays and their viewing conditions are calibrated to the sRGB standard, then the appearance of an image viewed with an sRGB profile on the two systems should be consistent. The standard specifies CIEXYZ tristimulus values for the primaries and the D65 white point, a gamma value of 2.2 and a set of reference viewing conditions in terms of ambient illuminance level, an ambient D50 white point and details also for the background and surround luminance levels.

9.6 Calibration, Characterization, and Color Management

Small changes in color reproduction in medical photography have the potential to significantly impact accurate monitoring of disease and the effectiveness of treatment over time, particularly in color critical applications such as dermatology. This places particular requirements on the system for accuracy and consistency. As stated by Revie and Green [19]: “Many of the current problems in color in medical imaging can be classed as problems of calibration of image capture and display systems” (p. 2). The processes of calibration and characterization are intrinsic to color management. More detail is provided about the characterization of different types of devices later in the chapter; here they are considered more generally.

Calibration is concerned with maintaining consistency and stability within a system, by ensuring that any factors (hardware and software settings, media, viewing environment) that might influence image reproduction remain unchanged and periodically checking and resetting them if they have shifted. Examples include maintaining the system white point in a display or grey balancing a scanner. Selecting the paper and colorant type is part of the calibration of a printer, but color and tone reproduction will vary between batches of both, so printers need recalibrating when a new batch of paper or colorant is installed. The necessary regularity of calibration is dependent upon the device and its susceptibility to drifting away from the calibrated state.

Characterization is a process of measuring the color and tone reproduction of a device and defining the relationship between coordinates in the device-dependent color space and CIE colorimetric coordinates. The output of the characterization process is a device profile; it is produced after the device has been calibrated (although some profiling devices will perform both), and the profile will only be valid and correct for that calibration condition [7], meaning that a change in conditions requires a new profile to be created. Device characterization enables the translation of colors from one color space to another, either directly or via a CIE color space in an ICC framework. An example of the use of profiles in an ICC color managed framework is shown in (Fig. 9.7).

Correctly and regularly calibrated devices, an intermediate colorimetric profile connection color space, and accurate profiles are the building blocks of color management systems. There are two further factors that determine how colors are rendered when transformed from one color space encoding to another in an ICC compliant system.

9.6.1 Gamut Mapping

The gamut of each color space defines the range of possible colors that can be encoded within the space. When transformed into a common colorimetric color space, the color space gamuts occupy different areas and shapes, depending upon the position of their primaries within the space. The three standard RGB color spaces can be illustrated on a chromaticity diagram as shown in (Fig. 9.8).

While there is a significant area of overlap within the center of the gamuts, they have some areas that are unique to each space, and this is a common problem, particularly when transforming from an RGB space of an input or display device to the CMYK space of a printer. There will be some colors in each space that cannot be mapped in an exact way. The process of gamut mapping adjusts the colors of the input image or the input device to fit those of the output device and requires characterized or profiled devices at input and output [7]. In the areas of overlap, colors may be matched, but in areas outside the common area, a decision must be made about what happens with out-of-gamut colors. The approach used within the gamut mapping algorithm depends on the color reproduction objectives. If colorimetric accuracy is paramount, then only the central colors can be correctly reproduced, and the approach is *gamut clipping*, where the colors in the common gamut boundary are unchanged, leaving a few very saturated colors to be clipped to the boundary. In medical imaging applications, this can be a satisfactory approach if these highly saturated colors are not prevalent

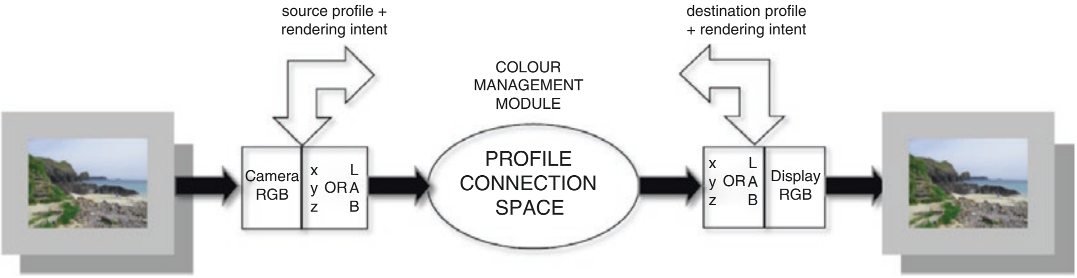

Fig. 9.7 Image capture and viewing in an ICC profiled and color managed system

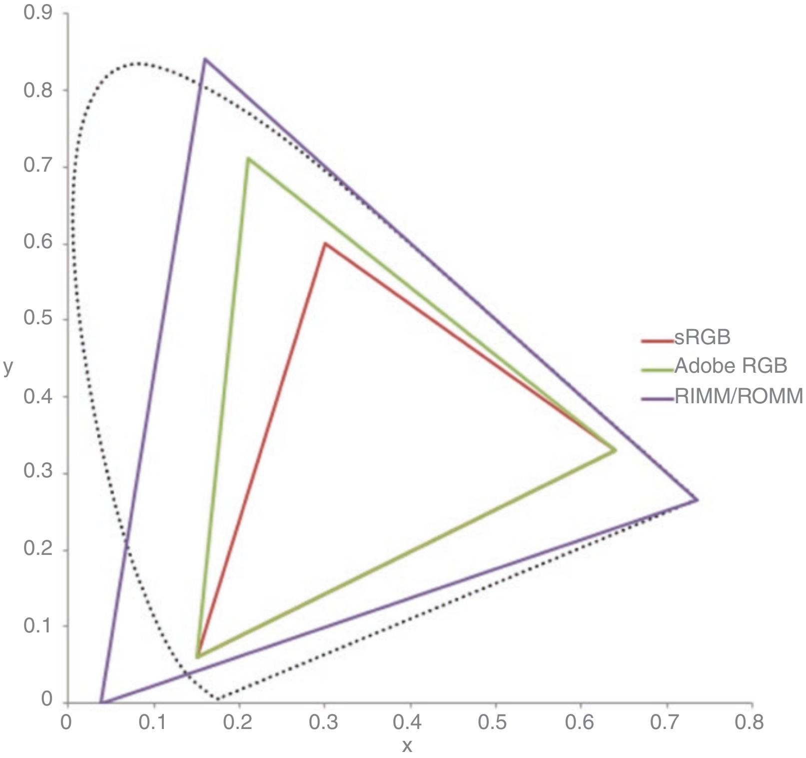

Fig. 9.8 Gamut space comparison of standard RGB color spaces on a CIE x,y chromaticity diagram

so that the color differences in these areas can be ignored. *Gamut compression* mapping algorithms instead compress the entire gamut into a smaller area. In this case, all colors are shifted to become less saturated, but relative color differences are maintained, and the color differences caused by the gamut mismatch are spread across the entire range. In both methods, priority is given to minimizing hue shifts.

9.6.2 Rendering Intents

In ICC color management, the gamut compression method is defined by selecting a *rendering intent* when converting between a source and destination profile. At the time of writing, there are four rendering intents defined within the standard; two of these, the *ICC absolute colorimetric* intent and the *media-relative colorimetric* intent, are suitable for medical photography (the other two are designed to produce pleasing images in more general imaging applications). Both the absolute colorimetric and media-relative intents produce colorimetrically accurate colors, the difference being that in the media-relative intent, the media white point is mapped to the white point of the profile connection space (PCS), whereas in the absolute colorimetric intent it is not. This means that with media-relative rendering, the white point of the image appears white even if it was a yellowish white (this is analogous to chromatic adaptation in the HVS) and all other colors are mapped relative to this white point, meaning that all colors shift. This is more suitable for human visual assessment of images, where images need to be reproduced to be visually accurate. In absolute colorimetric rendering, if the media white point was not the same as the white point of the PCS, it will appear different; a yellowish white point will appear yellowish. In this case, all colors are preserved exactly. This approach is suitable where images are to be evaluated or processed using an image analysis algorithm and absolute colorimetric accuracy is required.

9.7 Camera Calibration and Characterization

As described in Chap. 11, the use of a camera with another device such as a dermatoscope enables the calibration and profiling of the system, in part because the imaging conditions are possible to control, because of the close focusing distance and the use of an illuminating device. In much of medical photography, however, imaging conditions are more variable. The increasing use of mobile devices for image capture, particularly in telemedicine, represents significant challenges in terms of controlling and reducing color errors, which, once introduced at image capture, will be propagated through the entire imaging chain [11]. The quality of cameras ranges from mobile phone cameras, through point and shoot compact cameras, to professional-level DSLRs, and in this range, the level of automation versus the level of manual control is hugely variable; the situation is further compounded by the level of skill required of the camera operator to enable purposeful control of the image in high-end systems. Penczek et al. [11] tested the effect of various imaging conditions upon the color performance of mobile phone, point and shoot, and DSLR cameras and found the most significant errors introduced at capture by position in relation to light source, the type of light source used, and the camera technology. As expected, the overall errors were highest with the mobile phone camera, followed by the point and shoot and then the DSLR. All three cameras performed best under daylight-balanced fluorescent illumination; the DSLR as expected produced the best results, with excellent color rendition of flesh tones under daylight illumination, but performed badly under incandescent and cool white fluorescent illumination. The ICC has developed guidance for improving color in medical photography [20] also highlighting the primary factors that contribute to color error at image capture as lighting uniformity, the spectrum of the lighting, camera technology, and subject color. Poor framing, focus, exposure, and incorrect white balance also have an impact upon color.

Generally, therefore, the best results will be obtained using a professional DSLR, adhering to good practice guidelines for framing, focus, and capturing RAW files to fully utilize the available bit depth and dynamic range of the image sensor (see Chap. 11). If this is not possible, improvements can still be made by optimizing lighting and including a color test chart. Using an illuminant with a spectrum as close as possible to D65 will produce better color rendition regardless of the camera used, and uniform diffuse lighting helps to reduce non-uniformity across the image plane.

To produce calibrated images, an image must be captured of a test target containing a set of known colorimetric values (see examples in Fig. 11.8 of Chap. 11) from which a profile can be constructed that relates the camera RGB values to the colorimetric values of the chart under the same conditions. This can be a challenge in specialisms such as ophthalmology and pathology and is an area being addressed by the ICC MIWG, with the development of a calibration slide for histopathology and a miniaturized color checker chart which can be inserted into a model eye for fundus photography [19].

Images may be captured and rendered in a *scene-referred* or *output-referred* encoding [7, 20]. An output-referred encoding is one that has been optimized for a real or virtual output device, so could be for a printer, or using a color space such as sRGB. The images will be rendered for viewing, and in the case of JPEG files, output rendering is the only option. In such cases some color rendering will be applied automatically to optimize the results for the output color space. This process is proprietary, and so it is difficult to obtain colorimetric values from an output-referred image. Nevertheless, this may be the only option unless using a DSLR. Better results will be obtained if capturing to an uncompressed TIF file, although the images will still be output-referred. In such cases the best option is to capture the image with a color chart or capture the color chart in a separate image under the same lighting conditions, correct any illumination non-uniformity, and gray balance the image using a custom white point correction from the color chart image. A color correction profile or a set of presets for specific conditions can also be created from the color chart image, to be applied to the captured image. It may be possible to retrieve original scene colorimetry if the color rendering method is known.

Scene-referred images are encoded to preserve the original scene colorimetry and may be created within the ICC architecture using two different approaches, either by creating a custom camera profile for the camera illumination or by using a standard scene-referred profile. In both cases the images must be captured as RAW files. The custom camera profile allows direct conversion from camera RGB to colorimetric values in the profile connection space but requires special software to build the profile for the specific camera, settings, lenses, and illumination. Alternatively the images may be converted to a standard scene-referred RGB space; the one recommended by the ICC is the linear_RIMM-RGB_v4.icc profile, which can be downloaded from the ICC website [21]. In general, scene-referred images will produce more consistent results when compared in terms of either colorimetric measurements or viewing over time; it should be noted however that in general, it is not possible to directly view scene-referred image data and therefore the image will finally have to be output referred. This requires that the image is converted to an output-referred encoding; in this case the sRGB standard encoding is recommended, with a media-relative colorimetric rendering intent. It is worth noting however that sRGB has guidelines around viewing conditions which should be adhered to for best results. Adobe RGB 1998 is another standard encoding with a slightly wider gamut, which may be appropriate for soft-proofing images for print.

9.8 Output Device Calibration and Characterization

Both displays and printers are classed as output devices, and therefore the accuracy of their calibration and characterization is important at the stage at which the images are to be viewed by clinicians. However carefully an image has been captured, and the capture system has been calibrated and profiled; an uncalibrated or poorly calibrated output device will result in incorrect rendition of colors (and in the case of printers, typically a loss in image saturation and the potential for color shifts because of the often significant gamut mismatches between RGB and CMYK color spaces). As described earlier, the calibration step in each case requires the setting of conditions prior to profiling, followed by regular recalibration to return the device to those conditions and ensure that the profile remains valid.

9.8.1 Display Calibration and Characterization

There are a number of portable devices available for display calibration and profiling, or in high-end workflows self-calibrating monitors, with built-in calibration tools; most methods will perform the calibration immediately before the characterization, which ensures that the profile matches the conditions. It is important to allow some warm-up time to ensure that the display has stabilized to prevent erroneous results.

The steps required for display calibration involve an initial setting of display black and white point luminance to adjust the dynamic range and overall brightness. The calibration involves setting of a target white point color temperature and gamma value to adjust the display transfer function and color balance. If the display is being set up to display sRGB-encoded images, then the target values are according to the sRGB specification [18]. The device may be a colorimeter or a spectrophotometer and measures values from displayed patches with known colorimetry on the screen, either in close contact with or some distance from the screen surface. The device software then creates a profile allowing transformation between the PCS and the device values, which, when applied in the device settings, will be used to adjust values on the video card.

The display viewing conditions have a significant impact upon the color appearance on the screen; hence viewing images on a display in a room illuminated by natural daylight or under different viewing conditions to those used during calibration and profile is unlikely to produce consistent results. As a general rule, the ambient lighting should be at a relatively low level, unchanging, and should not have a color cast. In sRGB the viewing conditions are specified and need to be set up prior to calibration and profiling.

9.8.2 Printer Calibration and Characterization

As described earlier, part of the calibration of a printer is the selection of a paper and ink or dye set, and the process must be performed for each new type of paper and colorant. If a printer is to print on different types of paper, with different white points or surface qualities, a profile will need to be created for each paper type. Because CMYK color spaces are not perceptually linear, a CIELAB linearization stage may also be included as part of the calibration, which involves printing a pre-profiling linearization target which has linear CMYK scales. Once printed, it is measured using a colorimeter or a reflection spectrophotometer. The profiling software then reallocates the CMYK output values to linear CIELAB values. This means that the CMYK values will no longer be linearly spaced but will appear visually equally spaced.

Next, the profiling target should be printed; this tends to have a far larger number of test values than for display profiling, and these are randomized. The measured values are compared to the reference file for the test target and a profile created based on a lookup table (LUT). The aim is for the target to be printed without color correction to test the default uncorrected behavior of the printer. The resulting profile should ideally be selected as the destination profile at the point of printing an image.

9.9 Color Differences, Perceptibility, and Acceptability Thresholds

Perceptually uniform color spaces allow the meaningful calculation of color differences. The $\Delta E^*_{a,b}$ metric is one of the most widely used color difference measures and may be calculated directly either from the differences between the CIELAB values or from the luminance, chroma, and hue difference values [17]:

$$\Delta E^*_{a,b} = \left[\left(\Delta L^*\right)^2 + \left(\Delta a^*\right)^2 + \left(\Delta b^*\right)^2\right]^{1/2} = \left[\left(\Delta L^*\right)^2 + \left(\Delta C^*_{a,b}\right)^2 + \left(\Delta H^*_{a,b}\right)^2\right]^{1/2.}$$

Color differences can be used to benchmark thresholds of perceptibility of various imaging attributes such as color under specific conditions, to identify tolerances in terms of color reproduction for an imaging device, or to explore the effects of changing settings or viewing conditions within an imaging system, thus helping to establish standards appropriate to imaging requirements.

The CIEDE2000 formula [22] is a more recent and complex color difference metric recommended by the CIE, developed to compensate for some non-uniformities in the CIELAB color difference metric. ΔE_{00} is derived from CIELAB values, adjusting the relative weightings of lightness, chroma, and hue components for various illumination conditions. The implementation is rather complex; a step-by-step explanation can be found in [23] and supplementary notes in [24].

Metrics such as $\Delta E^*_{a,b}$ and ΔE_{00} are useful in identifying how much a color has shifted as it moves through an imaging chain. However, to be meaningful, they must be tested and benchmarked using psychophysical experiments with human observers. This process allows values calculated from the metrics to be used to predict the perceptual significance of the differences.

In a typical psychophysical threshold experiment, observers evaluate pairs or triplets of colors to identify the level at which differences between them become perceptible. When a given proportion of observers can notice a difference (commonly either 50% or 75%), the point at which this occurs is known as the just-noticeable difference of perceptibility or JND. For an in-depth explanation of psychophysical experimental methods, refer to Engeldrum [25]. JNDs in this case are a measure of *color fidelity*. When the difference between the two colors is also evaluated with a color difference metric, it is possible to calculate the amount of color difference in $\Delta E^*_{a,b}$ that will produce a just perceptible difference. This allows a prediction of the threshold of color differences that may be tolerated within a given situation, providing a useful benchmark to which devices and systems may be designed and calibrated. The perceptibility of differences, or JNDs, in uniform color patches may be as low as 1.0 $\Delta E^*_{a,b}$ but is generally higher when colors are part of complex scenes.

JNDs of *acceptability* or acceptability thresholds may also be established through psychophysical experiments [26]; in this case they quantify the level at which perceptible differences become unacceptable in a particular imaging application. Such measures are useful in applications where color matching is important, such as prosthetic dentistry [27], where the JND of perceptibility has been found to be approximately 1.0 $\Delta E^*_{a,b}$, similar to that found for uniform color patches, while the JND of acceptability is at around 3.7 $\Delta E^*_{a,b}$. JNDs of perceptibility and acceptability are highly context dependent.

Tolerances for perceptibility and acceptability vary depending upon the requirements of the imaging application, viewing conditions, the subject and image properties, and the experience and expertise of those viewing the images. The thresholds for skin tones in various specialisms are a case in point. As described previously, the HVS is influenced by the memory color effect, which is formed by an observer from familiar objects, provided that they are strongly associated with a typical color [28]. Skin tones can produce significant errors when captured on digital cameras [11] and are also strongly associated with memory colors [29]. Research into color difference perceptibility and acceptability thresholds using artificial skin samples used for maxillofacial prosthetics found that the perceptibility threshold for light skin tones is 1.1 $\Delta E^*_{a,b}$ and the acceptability threshold 3.0 $\Delta E^*_{a,b}$. For darker skin tones the thresholds are 1.6 $\Delta E^*_{a,b}$ for perceptibility and 4.4 $\Delta E^*_{a,b}$ for acceptability.

Psychophysical experiments are a widely accepted approach to test and measure various aspects of the quality of imaging systems, to establish benchmark requirements for imaging standards and to compare the performance of different devices. However, they are intensive and time-consuming and thus often impractical in real imaging scenarios; therefore, having an objective metric such as ΔE, which correlates with human perception and can be derived from colorimetric measurements within an image, is extremely useful.

9.10 Summary

There is little doubt that color accuracy, consistency, and standardization are of key importance in medical photography, to monitor disease and assist diagnosis, and in areas where color matching is important. The increased use of mobile devices and true color images makes it imperative that some of the challenges are met in managing color in medical photography, especially in areas such as telemedicine where images may be transmitted and viewed across different systems and in different geographic locations. Various specialisms have developed individual approaches, but there is currently no common framework. The ICC MIWG has identified a number of key areas of work with the aim of providing a consistent approach and framework for the standardization of color. The reproduction objectives differ across clinical areas, with some necessitating a tailored approach, for example, the development of a color extension to the DICOM grayscale standard display function [30], which is useful for grayscale images with pseudo color overlays, for color visualization of quantitative information. In more general medical imaging, the requirement is for colors in images to be as visually consistent as possible with the original and for this color to be reproduced accurately across different sorts of devices, devices in different locations and in images taken at different times. The ICC color management architecture has the potential to provide a framework in which accurate color reproduction may be implemented. Current work by the ICC MIWG addresses some of the very specific solutions required for best practice guidelines in medical photography.

References

1. Sharma G. Digital color imaging handbook. Boca Raton: CRC Press LLC; 2003.
2. Hunt RWG, Pointer M. Measuring color. Hoboken, NJ: Wiley; 2011.
3. Fairchild MD. Chapter 3: colorimetry. In: Color appearance models. Chichester: Wiley; 2013. p. 59.
4. Bartleson CJ. Memory colors of familiar objects. J Opt Soc Am. 1960;50(1):73–7.
5. Fairchild MD. Chapter 6: color appearance phenomena. In: Color appearance models. Chichester: Wiley. p. 116–21.
6. Triantaphillidou S. Chapter 5: introduction to colour science. In: Triantaphillidou S, Allen EJ, editors. The manual of photography. Focal Press/Elsevier; 2011. p. 77–102.
7. Triantaphillidou S. Chapter 23: digital color reproduction. In: Triantaphillidou S, Allen EJ, editors. The manual of photography. 10th ed: Focal Press/Elsevier; 2011. p. 411–32.
8. Murphy EM, Hegarty F, McMahon BP, Boyle G. Analysis of the colour rendition of flexible endoscopes. In: Opto-Ireland 2002: optics and photonics technologies and applications: proceedings of SPIE, vol. 4876; 2003. p. 26–36.
9. Allen EJ, Bilissi E. Chapter 14: digital cameras and scanners. In: Allen EJ, Triantaphillidou S, editors. The manual of photography. 10th ed: Focal Press/Elsevier; 2011. p. 275–6.
10. Allen EJ. Chapter 26: color management. In: Allen EJ, Triantaphillidou S, editors. The manual of photography. 10th ed: Focal Press/Elsevier; 2011. p. 476–7.
11. Penczek J, Boynton PA, Splett JD. Color error in the digital camera image capture process. J Digit Imaging. 2014;27(2):182–91.
12. International Color Consortium. ICC Medical Imaging Working Group. [Online]. http://color.org/groups/medical/medical_imaging_wg.xalter.
13. Badano A, Revie C, Casertano A, et al. Consistency and standardization of color in medical imaging: a consensus report. J Digit Imaging. 2015;28:41–52.
14. Colorimetry. CIE standard colorimetric observers. Patent BS EN ISO/CIE 11664-1:2019. 2019.
15. Colorimetry. CIE tristimulus values. Patent BS EN ISO/CIE 11664-3:2019. 2019.
16. Colorimetry. CIE 1976 L*u*v* colour space and u′, v′ uniform chromaticity scale diagram. Patent BS EN ISO 11664-5:2016. 2016.
17. Colorimetry Part 4: CIE 1976 L*a*b* colour space. Patent BS EN ISO/CIE 11664-4:2019. 2019 July.
18. Multimedia systems and equipment—Colour measurement and management—Part 2-1: Colour management—Default RGB colour space—sRGB. Patent BS EN 61966-2-1:2000 IEC 61966-2-1:1999. 2003.
19. Revie C, Green P. Current problems and perspectives in colour in medical imaging. In: IS&T international symposium on electronic imaging 2016 color imaging XXI: displaying, processing, hardcopy, and applications; 2016. p. 1–5.
20. ICC White Paper 46: improving color image quality in medical photography. White Paper. International Color Consortium; 2017.
21. Consortium IC. Creating scene-referred images using Photoshop CS3. [Online]. [cited July 2019]. http://www.color.org/scene-referred.xalter.
22. Colorimetry Part 6: CIEDE2000 Colour-difference formula. patent BS EN ISO 11664-6:2016. 2016.
23. Triantaphillidou S. Appendix B. In: Allen EJ, Triantaphillidou S, editors. The manual of photography, 10th ed. Focal Press/Elsevier; 2011. p. 555–6.
24. Sharma G, Wu W, Dalal EN. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. Color Res Appl. 2005;30(1):21–30.
25. Engedrum PG. Psychometric scaling: a toolkit for imaging systems development. Winchester: Imcotek Press; 2000.
26. Allen EJ, Triantaphillidou S, Jacobson REJ. Perceptibility and acceptability of JPEG 2000 compressed images of various scene types. In: Proceedings Volume 9016, image quality and system performance XI, San Francisco; 2014.
27. Ghazal K, Bain PA, Salari S, Dozic A, Kleverlaan CJ, Felzer AJ. Perceptibility and acceptability thresholds for colour differences in dentistry. J Dent. 2014;42(6):637–44.
28. Witzel C, Valkova H, Hansen T, Gegenfurtner KR. Object knowledge modulates colour appearance. i-Perception. 2011;2(1):13–49.
29. Zeng H, Luo R. A new method for skin color enhancement. In: Proceedings volume 8292 IS&T/SPIE electronic imaging | 22–26 January 2012 color imaging XVII: displaying, processing, hardcopy, and applications, Burlingame, California; 2012.
30. Kimpe T, Rostang J, Van Hoey G, Xthona A. Color standard display function: a proposed extension of DICOM GSDF. Med Phys. 2016;43(9):5009–19.