COLOR MANAGEMENT FOR DIGITAL CINEMA

A Proposed Architecture and Methodology for Creating, Encoding, Storing and Displaying Color Images in Digital Cinema Systems

Edward J. Giorgianni
Submitted to the Science and Technology Council,
Academy of Motion Picture Arts and Sciences
November 25, 2005

Notices

Some material described in this document is protected by one or more U.S. and/or foreign patents.

The following illustrations are reproduced from *Digital Color Management, Encoding Solutions*, E. J. Giorgianni and T. E. Madden, copyright Eastman Kodak Company: Fig. 1.5.1a, Fig. 1.5.1b, Fig. 6.1, Fig. 1.6.2, Fig. 7.6.1, Fig. 7.11.2, Fig. 8.1, Fig. 8.5.1, Fig. 10.1.1, Fig. 10.1.3, and Fig. 10.1.4.

A number of definitions given in the glossary of this document are from the Annual Book of ASTM Standards, copyright American Society for Testing and Materials.

The author has taken care in the preparation of this document but makes no expressed or implied warranty of any kind and assumes no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information contained herein.

Table of Contents

Color Management for Digital Cinema

Edward J. Giorgianni

Introduction

The purpose of this document is to present a comprehensive proposal for implementing color management in an inclusive and extensible digital cinema system. The proposed system supports input from electronic cameras, computer-generated images, and photographic films and provides output to digital projectors, self-luminous displays, and writers for hardcopy media. The proposal addresses color-related problems identified by the File Format Project Committee and suggests solutions to those problems. Included are a conceptual architecture defining how the overall system would work and recommended methodologies for encoding, manipulating, storing, and communicating color throughout the system.

1. Committees, Myths and Misconceptions

It has been my experience that the work of color-management technical committees is often impeded by a number of widely held misconceptions regarding color science, colorimetry, and color imaging. Often, little progress can be made until all participants fully agree on the relevant technical issues and their implications. Disagreements often cause the same issues to be argued over and over, without resolution. Worse yet, misconceptions can lead to compromised and undesirable outcomes that would have been avoided had the issues been properly understood.

It is important, then, that any such misconceptions be expunged before proceeding with this description of the proposed system. Otherwise, the proposal may seem obscure and unnecessarily complex. If, for example, one believes that the colorimetry of a displayed cinema image ideally should equal that of the original scene, the output-rendering transforms described here will seem entirely pointless. Similarly, the appearance transformations included in this proposal will make little sense if one is under the impression that encoding images in terms of recommended CIE color spaces such as CIELAB provides a description of their color appearance.

In the following subsections, a "top-ten list" of issues I have found most likely to cause conceptual and technical problems are discussed. Please note that an extensive glossary is included with this paper to clarify the terminology used here and elsewhere in related color-management literature. The intent of this preliminary discussion is to help the committee avoid pitfalls that have made the work of other groups more difficult and have led to unsatisfactory outcomes.

1.1. CIE Colorimetry

Perhaps the most unrecognized (or overlooked) fact regarding CIE colorimetry is that it was developed for one specific purpose: to predict whether pairs of color stimuli will visually match, according to a set of responsivities defined for a standard human observer. One stipulation is that the stimuli being compared must be in close proximity and viewed simultaneously under identical viewing conditions.

It is critical to understand that this standard colorimetric practice is *not* a predictor of color appearance—nor was it ever intended to be. It is only a determination of whether stimuli match; it is not a numerical specification of what the stimuli involved look like. Misunderstanding of this one point alone has led to the demise of numerous color-imaging systems and color management products.

The appearance of a given color stimulus will depend not only on its CIE tristimulus (*XYZ*) values but also on the viewing conditions under which it is presented to the observer and on the adaptive state of that observer. These additional factors must be taken into account if CIE colorimetry is to be used appropriately in imaging applications. Procedures for doing so are described later in this paper.

1.2. CIE Recommended Color Spaces

A related prevalent notion is that if CIE *XYZ* values do not describe color appearance, conversions of those values to other CIE recommended color spaces such as CIE 1976 $L^*a^*b^*$ (CIELAB) will produce values that do directly relate to color appearance. That would be wonderful if it were true. Unfortunately, it is not.

The real purpose of such conversions is to produce colorimetric values such that observed equal color *differences* among pairs of stimuli are represented in a reasonably uniform way throughout the color space. This is quite useful for specifying colorimetric tolerances because a computed colorimetric difference, such as a given CIELAB ΔE value, will be similarly noticeable regardless of where the involved stimuli are in the color space. But a specification of the location of a stimulus in CIELAB or a similar color space is not a description of its color appearance.

1.3. The CIE Standard Observer

Perhaps due to difficulties correlating CIE colorimetry and color appearance, it is sometimes suggested that there must be something "wrong" with the specified

CIE Standard Observer. (Actually, there are several CIE observer specifications. The CIE 1931 Standard Colorimetric Observer, also called the 2º observer, is appropriately used for most imaging applications.)

My experience suggests that if such observer-related problems exist at all, they are so small as to be insignificant in practical imaging applications. I have encountered situations where color problems were attributed to deficiencies in the CIE Standard Observer specification. However, in every case, I determined that the problems actually resulted from more mundane causes such as incorrect device calibration and mathematical errors. In fact, I have never encountered a single imaging-related application where using the CIE Standard Observer has been a problem. This includes extremely demanding applications, such as diagnostic medical imaging, and in situations involving severe metamerism, such as matching colors on hardcopy media with previews generated on self-luminous displays.

I believe that any colorimetric errors introduced into a digital cinema system by the Standard Observer specification will be far smaller than the colorimetric variability that will occur from variations in illuminants, devices, media, calibrations, mathematical transformations, and numerous other system components. To be robust and practical, a system must be designed to deal with these normal variations; thus, this system should easily handle any slight errors that might be contributed by the specified responsivities of the CIE Standard Observer.

1.4. Scene Colorimetry and Displayed Colorimetry

The conditions under which standard colorimetric methods directly apply were described earlier. One is that compared stimuli must be viewed under identical conditions. However, in motion picture film and digital projection, the displayed image is viewed in conditions very different from those of the original scene. Specifically, the projected images is:

- not viewed simultaneously with the original scene
- two-dimensional, not three-dimensional
- viewed at an absolute luminance level much lower than that of the original scene
- viewed at a different state of chromatic adaptation
- displayed in a darkened environment
- surrounded by an extensive dark field
- affected by stray light present in the projection viewing environment

Because these differences will produce substantial changes in the physical and perceived color of the displayed image, the colorimetry of that image must be altered such that its appearance will be correct in the intended viewing environment. In the jargon of the industry, this alteration is one aspect of what is called *rendering*, and images having colorimetry appropriate for viewing are referred to as *rendered images*.

The colorimetry for all color reproductions must be rendered. That is true for paintings, conventional photographic images, graphic arts prints, self-luminous displays, and projected images. Colleagues whose experience is limited to graphic arts will sometimes argue that this is not true; but it is important to realize that in that field, the "original" is itself a reproduction. It is correct to say that the appearance of a reflection print can be matched by another reflection print having identical CIE colorimetry (assuming they are viewed under identical conditions). Similarly, the appearance of a motion picture print can be matched by another motion picture print having identical colorimetry. But as with all color reproductions, for correct color appearance, the colorimetry of those images must be entirely different from that of an original live scene.

1.5. Scene Luminance Dynamic Range

Standard colorimetric procedures often make use of perfect white references, which are defined as ideal isotropic diffusers with a spectral reflectance or spectral transmittance of 100% at each wavelength of interest. Perhaps because of the words "perfect" and "ideal" there is a common misconception that nothing can have a luminance greater than that of a perfect white. It would seem to follow that there must be nothing to capture and nothing to display beyond that point; but that statement is incorrect on both counts.

Scenes often contain information at luminances above that of the perfect white. These levels can be produced by specular highlights, certain types of diffuse highlights, fluorescent colors, and secondary light sources within a scene. It is quite common for high levels of luminance to be produced in scene areas that are more highly illuminated than the principal subject area. In the example image below (Fig. 1.5.1a) and associated histogram (Fig. 1.5.1b), the clouds are highly illuminated and produce luminance levels higher than those of the whites in the principal subject. As a result, the luminance dynamic range shown in the scene's histogram extends beyond the perfect white.

Figure 1.5.1a: The extended luminance range of this scene results from different levels of illumination in the sky and principal subject areas.

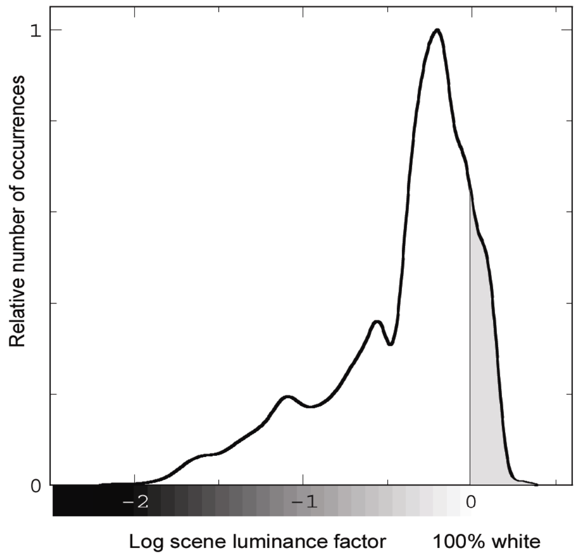

Figure 1.5.1b: A histogram of the image of Fig.1.5 1a.

In addition, many scenes actually are not a single scene; they are made up of multiple "sub-scenes" in which the lighting can be quite different. For example, the exposure histogram in Fig. 1.5.2 below is that of a representative back-illuminated scene. The exposure distribution is bimodal because there effectively are two scenes, one in the principal subject area (foreground) and another in the background. Similar multiple-illumination levels occur in many other circumstances, such as when an indoor scene has multiple areas of localized lighting or includes a window open to a highly illuminated outdoor scene.

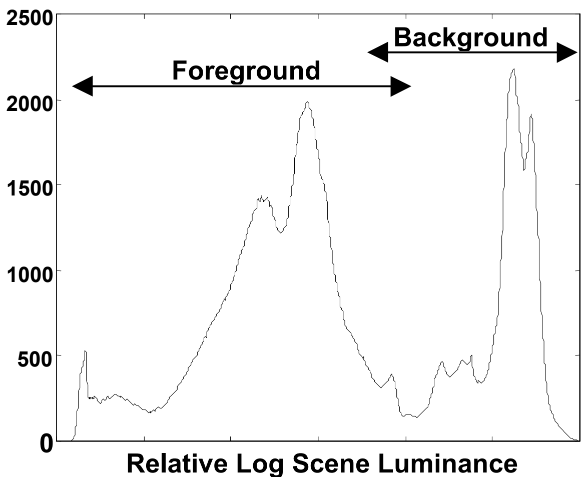

Figure 1.5.2: A representative histogram (occurrences vs. scene luminance) for a back-illuminated scene.

Many original scenes, therefore, contain an extensive luminance dynamic range, and input devices and media are indeed capable of capturing much of that range. It is important, then, that the color encoding of a color-managed system have a dynamic range as extensive as that of its inputs. In addition, imaging systems must deal with under and overexposed images. Thus in the stages prior to exposure correction, the color encoding must be capable of handling an even greater dynamic range of information.

1.6. Rendered-Image Luminance Dynamic Ranges

In discussing luminance dynamic range in the context of rendered images, it is important to recognize that there actually are *two* fundamentally different dynamic ranges involved. The more obvious one is the relative luminance range that can be displayed by the rendering device or medium. This range often is quite large because it is necessary to compensate for viewing flare and various psychophysical and psychological factors involved in image viewing.

The other luminance dynamic range associated with rendered images is less obvious. This range relates to the range of system input information that is retained and can be represented in a rendered image. In high-quality rendered images, this representation must include information corresponding to scene relative luminance values above the perfect white. This fact is somewhat counterintuitive and is often disputed, so the topic warrants further explanation. It is important here because the luminance dynamic range requirements for rendered images will be a critical consideration for the color encoding of this proposal.

Obviously it is not possible for a reflection print to have a diffuse reflection of greater than 100%, nor is it possible for a projected image to reflect more light than the maximum supplied by the projector. However, all high-quality imaging systems are capable of displaying a *representation* of scene high-luminance information with detail that is at least acceptable. This is illustrated below in Fig. 1.6.1, where the overall system grayscales for four types of media are shown.

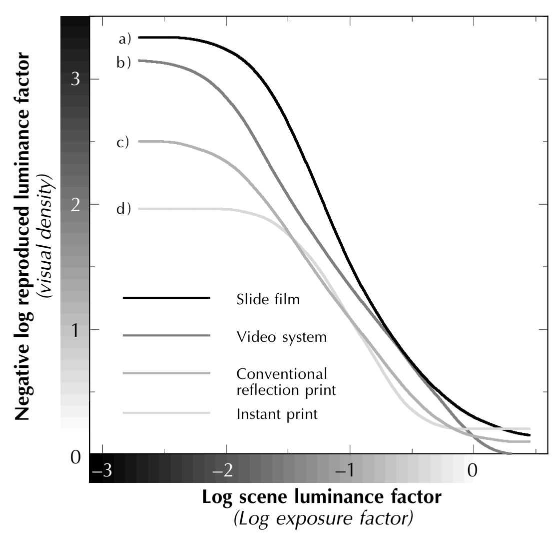

Figure 1.6.1: Display of information from scene luminance levels greater than that of a perfect white.

Notice in Fig. 1.6.1 that, except in the case of the instant print, the system grayscales have gradient at the 100% scene white point (zero on the X-axis) and beyond. This means that the systems can display, to various extents, information from very high-luminance areas of the original scene. It is important, then, for this scene information to be acquired, encoded, and made available to the output devices and media.

In Fig. 1.6.2 below, the same grayscales have been adjusted for equivalent brightness. (More regarding that process will be discussed later.) When this adjustment is made, it becomes evident that projected slide images and projected or self-luminous electronic images are capable of creating the *illusion* of brightness greater than that of a perceived perfect white (i.e., negative visual densities in the figure).

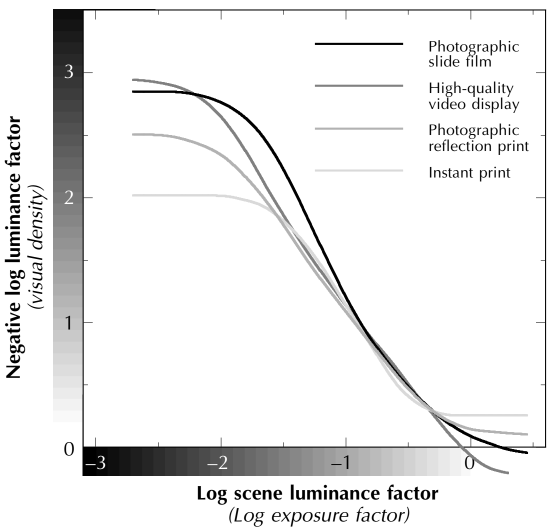

Figure 1.6.2: Brightness-adjusted system grayscales show that displays can create the illusion of brightness greater than that of a perceived perfect white.

Digital and film motion picture projections can create an even greater range of perceived brightness levels. In a darkened theater, the illusion of very bright subjects, such as fiery explosions, can be entirely convincing. Thus the encoding of a high-quality digital cinema system must be capable of representing an extensive luminance dynamic range through to the output and display stages. This encoding method must allow for "whiter-than-white" values such that displayed image areas that will be *perceived* to be brighter than a perfect white have values equivalent to CIE Y values (and L^* values) greater than 100.

This concept of "whiter-than-white" may lead to great debate within the committee. It has, in fact, frequently been argued that there is no point in encoding CIE Y or L^* values greater than 90 or so because reflection print supports and motion picture screens do not reflect more than that. But allowing for higher dynamic range values makes it possible to fully capture and preserve the creative intent during content creation. It is critical, then, that in a digital cinema system, high-luminance information be acquired, retained throughout the imaging system, and displayed appropriately based on the capabilities and limitations of each particular type of output.

1.7. Color Encoding Methods

Discussions regarding digital color encoding generally focus on two issues: image file formats and color spaces. The underlying assumption of such discussions is that once there is agreement on these two issues, the problems of representing and interchanging color images are solved. However, before these issues should even be addressed, a much more basic question needs to be answered: What color-encoding *method* is appropriate for this application?

The color-encoding method determines the actual meaning of the encoded data. For example, the encoding method might be that encoded values represent standard CIE colorimetric measurements of reproduced images. If, for the particular circumstances of a given system, that method is appropriate, then and only then should the discussion move to what color space (CIEXYZ, CIELAB, etc.) and format should be used for that standard colorimetric specification.

Eventually, of course, a decision on color space does have to be made. But it is important to recognize that if the encoding method is inappropriate, the system will fail regardless of what color space is chosen. By analogy, if a problem can be solved only by measuring the mass of an object, and its length is measured instead, the metric selected for expressing that length (centimeters, inches, etc.) will not matter. The problem will still not be solved.

It is also important to recognize that there is no single "best" method for representing color. Determining the appropriate encoding method for a particular imaging system requires an evaluation of what inputs and outputs are to be supported by the system and the principal intended usage of the system. Consequently, the best encoding method for one system may be entirely inappropriate for another.

1.8. Color Encoding Spaces

To start a heated debate among a group of color scientists, just ask them for the "best" color space. There will be many answers because the question alone under-defines the problem. This should lead to other questions: Best for what? Input processing? Color gamut? Quantization? Editing? Compression? Gamut mapping? Output processing?

As was the case for color encoding methods, there is no single best color space. It is not possible for one to be best because there are many different—and often conflicting—criteria. For example, all other things being equal, a color space having a smaller color gamut will have quantization that is less visible than that of a space having a larger color gamut. So which space is "better". Again, there is no answer unless the relevant criteria are specified.

In practice, the signal processing for virtually all modern systems frequently transforms images through a number of different spaces, each appropriate for a different purpose. In a consumer digital camera, for example, different color spaces are likely to be used for pixel interpolation, uniformity correction, color balancing, sharpening, achromatic noise reduction, chromatic noise reduction, color correction, and spatial compression. Different color spaces may be used for these operations because each space is well suited for some operations but not others.

Similarly, on a workstation, some operations are more easily performed in one space than in another. For example, correcting the color balance of an image from a digital camera is straightforward if one is using an RGB color space similar to the camera's own RGB signals. If the image has been transformed into a nonlinear space such as CIELAB, such color balancing becomes far more complicated. Simple shifts in a^* and b^* are likely to provide unequal balance corrections throughout the space. On the other hand, changing only the hue or chroma of a color in an image is very difficult to accomplish using RGB adjustments; but it is straightforward using color spaces where luminance and chrominance information are independent.

In certain cases, there will be some compelling reason to select one color space over another. That tends to happen mostly on low-end systems where computational resources are limited, making it advantageous to use a color space that requires minimal input and/or output signal processing. For digital cinema, I doubt that will be the case. Therefore the choice of color space may not be obvious (and more to the point, probably not that critical).

1.9. Color Encoding Data Metrics

A color space is one aspect of what I refer to as a color encoding data metric. A complete specification for an encoding metric includes the color space and the mathematical properties of the metric (bit depth, variable ranges, etc.) Sometimes there is a single overriding concern that essentially dictates the most appropriate data metric. Again, this situation usually arises on systems having limited computational resources, which is not applicable here.

What will be critical for digital cinema is that any data metrics used must not restrict the capabilities or performance of the system in any way. Care should be taken to ensure that all current and anticipated future inputs and outputs will be fully supported by the encoding and its associated data metric.

1.10. Device Dependent/Independent Encoding

One divisive topic related to data metrics that is very likely to arise in Committee discussions is that of "Device Independent Color" vs. "Device Dependent Color". In the past, this one topic alone has been responsible for countless hours of mostly pointless debate and numerous unsatisfactory outcomes. The opposing positions can be summarized as follows:

- Color encoding should always be in a Device Independent space (which advocates generally define as any of the CIE recommended spaces), because only those spaces encompass the entire gamut of visible colors.
- Color encoding should always be in a Device Dependent space (which advocates define as signal values for a particular type of input or output device or medium), because this minimizes the complexity of the data transformations needed to convert signals to and/or from the space.

Although both positions seem to make sense, what is often missed is that they are not mutually exclusive. It is in fact possible to use what some might call a device dependent color space, yet still maintain the ability to represent all visible colors (or, of course, any subset of these colors). This can be done in a number of ways.

For example, many cyans and some yellows are outside the gamut of a typical CRT, so a color space based on that device normally could not represent those colors. However, out-of-gamut cyans, for example, can be represented in terms of CRT signal values using positive green and blue values but negative red values. I used that principal on the Kodak Photo CD System by allowing for negative signal values in the data metric (which, by the way, can be done without the use of signed integers). Doing so preserved the feature of easy output to CRTs (the negative red signals can either be mapped or simply clipped to zero). However, the encoded values can also be used for output to other devices or media having larger color gamuts. The output signal processing will generate positive signal values such that the full gamut capabilities of the particular output are realized. Similarly, an extended luminance dynamic range can be preserved by allowing signal values to extend above those corresponding to a perfect white. These values can be mapped or clipped when sent to devices or media having smaller dynamic ranges.

Other techniques also can be used to resolve the apparent conflict between having signal values that require little or no signal processing for a specified output device while retaining an extended color gamut and luminance dynamic range. For example, an image file containing data for the reference output device can be accompanied by an auxiliary file or metadata containing additional image data that can be used by other devices and media having greater capabilities.

The point here is that if it becomes necessary to do so, it is possible to have a color encoding data metric that provides easy output to a selected device without sacrificing capabilities that could be utilized by other existing or future types of output devices and media.

Summary of the "Top-Ten" Misconceptions

In order to avoid the pitfalls described in this section, it will be important for the Committee to keep the discussed myths and misconceptions in mind. For convenience, the main issues are summarized below:

1. Standard CIE colorimetry does not—and was not intended to—represent color appearance.
2. CIE XYZ values that have been converted to other recommended color spaces such as CIELAB still do not represent color appearance.
3. Use of the CIE 1931 Standard Colorimetric Observer should not be a concern for most imaging applications.
4. For proper color appearance, the colorimetry of a displayed image always must be altered (rendered) from that of an original live scene.
5. Original scenes routinely have areas of luminance greater than that of a perfect white in the principal subject area.
6. Two distinctly different luminance dynamic ranges must be considered for rendered images. One corresponds to range that can be displayed, the other corresponds to the range of original luminance information the display can represent.
7. Successful color encoding begins with a determination of an appropriate encoding method rather than with a color space or data metric.
8. There is no one "best" color space for digital images because different spaces are best suited for different purposes.
9. The design of a data metric must consider the resources of the particular system for which it is being designed.
10. Device-dependent color encoding methods can be designed such that they are unrestricted by the limitation of actual devices.

2. Color Management System Architectures

The architecture of most color management systems is based on the use of a single color-encoding method with an associated data metric—a combination I refer to as a Color Encoding Specification (CES). In this classic architecture, illustrated in Fig. 2.1, the data path from each of any number of inputs includes an input transform to convert signals to those specified by the CES. Similarly, the data path to each of any number of outputs includes an output transform to convert CES values to those required by the particular output.

When designed and implemented correctly, this classic architecture provides many desirable features. In particular, it allows images from various types of inputs to be brought to a common representation where editing, compositing, and other operations can be performed without requiring differentiation or knowledge of the actual input source. It also allows CES images to be sent to any of the outputs supported by the system. Another important feature is that new inputs and outputs can be added to the system at any time by the use of an appropriate input or output transform for each new addition.

The principal challenge in designing a system based on this architecture is that of determining a CES that can support all the inputs and outputs while also meeting other system requirements. In most cases, however, there is one overriding requirement that essentially will dictate the definition of the CES. For example, the requirement could be that CES values must be very close to the final output code values for the principal output of a resource-limited system.

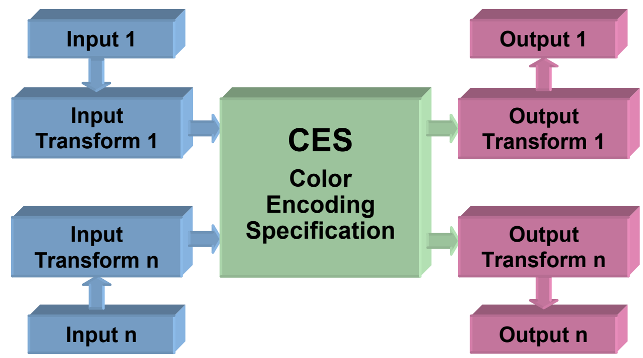

Figure 2.1: Classic use of a Color Encoding Specification (CES) in a multiple input/output imaging system.

One common strategy—one that I have never willingly used—is to base the CES on the lowest quality imaging component of the system. For example, input and output for various types of devices and media can be supported using a CES based on the rendering properties of a reference reflection print medium. Although it provides interoperability within the system, this "lowest common denominator" approach creates a system where images from even the highest quality inputs are irrevocably reduced to the quality of what is perhaps the poorest system component. Similarly, images produced on the best output devices and media can be little or no better than those specified by the reference reflection print of the CES. This approach clearly is not suitable for professional imaging applications, and it would be particularly inappropriate for motion picture systems where extraordinary image quality is expected. It would be equivalent to basing a professional studio audio recording and mastering system on MP3.

The basic approach of using a reference rendering medium can be used to create a much more useful system if the reference rendering medium of the CES is modified to have properties beyond those of any real medium. This is the basis for the "Universal Paradigm System" proposed in my textbook. In the illustrative example used in that book, the luminance dynamic range of the reference medium extends to an L^* of about 120 (versus a maximum L^* of only about 90 for a reflection print having a minimum visual density of 0.05). This additional "headroom" in the CES of the Universal Paradigm allows the encoding of much higher quality images, such as those produced by digital projection, photographic slides, and motion picture print films. The L^* upper limit of 120 could, of course, be extended as needed.

It might seem, then, that using the approach of a CES based on an idealized reference rendering medium would be ideal for digital cinema applications. But I do not think so, for the following reason:

Ordinarily, color-imaging systems fall into one of two categories: they are focused primarily either on the inputs or on the outputs. The Kodak Photo CD System, for example, was primarily driven by the desire to produce output TV images using CDRs and a specialized player. Accordingly, the system's color encoding was optimized for that output-driven purpose. The principal focus of the Universal Paradigm also was on output-driven applications such as office imaging and desktop publishing; so again, the color encoding is indicative of that output focus. On the other hand, the emphasis of the Cineon Digital Film System was the inputs. The objective was to scan motion picture negatives in a way that retained all the image information for editing and subsequent output to intermediate negative films. To support that function and to provide for the digital archiving of negatives, an input-based CES of photographic printing-density values was used.

What sets the digital cinema system apart from these prior systems is that there are very demanding requirements that must be met for both the inputs *and* outputs. Moreover, the input requirements are quite different from those of the output. As a consequence, a CES that meets one set of requirements will not meet the other, and I believe a compromise CES that attempts to meet both will be satisfactory for neither.

What is needed, then, is an alternative architecture that allows the encoding-related objectives for the system to be achieved. In the following sections, those objectives are listed, the proposed architecture is described, and the associated color encoding details are discussed.

3. Proposed Objectives for Digital Cinema

The color-encoding architecture and methodologies that will be described are intended to provide a foundation for a Digital Cinema System that meets the following objectives:

- The system will support all currently available image acquisition devices, techniques, and media including electronic cameras, computer generated images, and photographic films.
- The system will be capable of incorporating and supporting future means of image acquisition, including those having imaging capabilities that exceed those of current electronic cameras.
- The system will incorporate a method and means for digitally representing input color information in a way that allows images from disparate sources to be adjusted and edited without regard to the actual sources of images.
- The system will support both the retention and the removal of unique input color characteristics associated with individual input sources.
- The system will provide for images to be archived such that all acquired color information is retained.
- The system will support the conversion of images to and from other color representations for editing and other forms of image manipulation in existing or future user-defined workspaces.
- The system will provide a means for the artistic and creative intent of images to be unambiguously specified, stored, communicated, and retained throughout the system to the final display.
- The system will support all relevant existing means of image output and display.
- The system will be capable of incorporating and supporting future means of image output and display, including those having capabilities that exceed those of current devices and media.

4. Proposed System Architecture

As suggested earlier, it does not appear possible to meet these objectives using the classic color-management architecture of Fig. 2.1. In particular, I see no way the full dynamic range of acquired input information can be retained in a CES that is also appropriate for rendered output. However, it is essential for the system to include some means for conveying the intended output appearance to the final display. Therefore, I suggest the Committee consider the alternative architecture shown in Fig. 4.1 below.

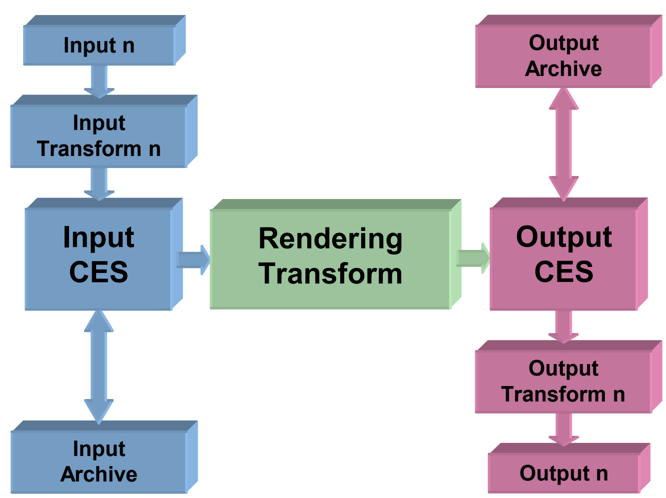

Figure 4.1: Proposed architecture for a Digital Cinema System, incorporating two Color Encoding Specifications.

In this multiple-input multiple-output arrangement, the inputs and outputs have different Color Encoding Specifications, each of which is optimized to meet a different set of criteria. The Input CES is designed to encode all information acquired by existing and future input devices and media, including those having the largest anticipated image-capture capabilities. The Output CES is designed to encode the rendered color appearance of a hypothetical display device having both a dynamic range and color gamut exceeding those of present or anticipated display devices and media.

Details regarding each CES, the associated input and output transforms, and the connecting rendering transform are discussed in the following sections.

5. Input CES Encoding Method

The color-encoding method recommended for the input CES of the system is based on Scene-Space Color Encoding. In the most basic form of this encoding method, original-scene color stimuli are represented by their original colorimetric values rather than by values corresponding to those of intermediate signals or rendered reproductions of those stimuli produced by any device or medium. There are many advantages to this method of encoding color information.

When scene-space encoding is used, input image information from all sources is brought into a common, input-independent representation. As a result, scene-space encoded images from various sources can be edited together to form seamless composite images. All editing and other image modifications can be done with one suite of tools because all images are encoded in a consistent way.

Although these basic objectives also may be approached by certain other encoding methods, scene-space encoding is unique in that it simultaneously provides both of the following features:

- It supports input from any and all types of media and devices, regardless of their disparity.
- It allows information to be encoded in a way that places no limits on luminance dynamic range or color gamut. Any color that can be seen by a human observer can be represented.

The inherent unrestricted nature of scene space makes it ideal for input encoding in this application. (In some cases, it may be appropriate to limit the encoded space in the *data metric*. But that is a separate design issue that will be discussed later. Scene space itself is unlimited). Any of the alternative encoding methods based on rendered-image spaces will, by definition, have restricted luminance dynamic ranges and possibly color gamut boundaries. As a result, image color information captured in acquisition will be lost in an encoding process based on rendered-image spaces. The extent of that loss would depend on exactly how the space is defined. If the input CES were based on the rendering properties of a real display device or medium, the loss might be considerable. A CES based on the rendering properties of a reflection print, for example, would result in the loss of important highlight information and would make that approach unacceptable for applications where high quality input and outputs must be supported.

Referring again to Fig. 4.1, it is important to note that as in any imaging system, the flow of information is “downhill” from the first block to the last. The most color information available is that which is originally provided by the input. If the input transformation is done correctly and the CES is designed well, no information will be lost in the process of Input CES encoding. Rendering to the Output CES loses information (even if that process is based on an idealized output); and still more is lost in the transformation for any real output.

In this proposal, then, the strategy is to retain as much information as possible for as long as possible. This means that the color information should stay in the input CES space until all editing and other creative work is complete and the image is ready for rendering to output. If archiving of input images is desired, that too should be done from the CES. By analogy, high-quality digital audio is recorded and stored at bit depths and sampling frequencies several times greater than those of any present digital audio output medium or device. Editing is done at full resolution, and signal processing is performed using real numbers. Only after this work is complete and stored are the resulting files reduced to the lower digital standards of conventional CDs, MP3s, etc. A similar strategy is appropriate for digital cinema.

6. Color in Scene-Space Encoding

Earlier it was stated that, at its most basic, scene-space encoding represents original-scene color stimuli in terms of their actual colorimetric values. That unique capability is ideal for *some* imaging applications. For example, an accurate representation of original-scene colorimetry is a necessary first step when colorimetric accuracy is the principal objective of the entire imaging system. Scene-space encoding therefore is appropriate (and has been used) for many scientific, medical, law enforcement, military, and other technical applications. It is also appropriate for use in more conventional applications where accurate color reproduction is important, such as in the imaging of art objects for high quality printing.

The most likely objection to the use of scene-space encoding is that there are many applications where the color objective is *not* one of accuracy. Instead, the desire might be to deliberately alter colors. For example, a cinematographer might select a particular film or electronic camera because it “distorts” scene colors in a way that helps create a desirable look. Similarly, a computer-generated image might be created in which entirely unrealistic scene colors are deliberately used to produce a desired effect.

For this proposal to be accepted by the industry, then, it is very important that the following fact is fully understood: ***The use of scene space is not limited to the encoding of actual scene colors.***

I realize that statement might sound contradictory, but consider this example: Imagine an outdoor scene on a heavily overcast day. Now imagine the same scene, but on a clear sunny day. The colors would be much more saturated. So what, then, are the "original-scene colors"? The answer is that they can be whatever someone wants them to be. Since it would be perfectly valid to accurately record the scene colorimetry under the conditions of either day, it is equally valid to digitally alter one day's image to look like the other. In fact, it is valid in scene space to alter the color of any image for any reason. It is also valid to use computer generated image techniques to create virtual scenes with any desired color properties, accurately realistic or completely imaginary. Scene-space, then, can incorporate scene colors that actually existed, colors that were altered by the characteristics of some input device or medium, colors that were modified in editing, and colors that were created entirely from scratch using computer generated imaging.

What sets scene-space apart from other color representations is that its colors, regardless of their actual origin, are understood to exist in an unrendered state. In this state, images can be combined, edited, adjusted, and manipulated in any number of ways, essentially without restriction. It is only after that work is complete that images are rendered for output, where restrictions necessarily must be imposed.

7. Getting Into Scene Space

The two next most likely objections to the use of scene space encoding are that it is difficult to get into and that it requires input transformations that are obscure and complex. That can be true for film but, as we will see, certainly not for other types of input. Even film input can be dealt with using a "black-box" approach that is easily implemented. Moreover, many of the complications involved (e.g., camera flare) must be dealt with regardless of what type of input encoding method is used. In many cases, such complications actually can be handled more straightforwardly in the process of transforming to scene space than to other forms of color representation.

In the following subsections, the process of transforming images from a number of different types of input devices and media into scene space will be discussed. The examples will start with the simplest form of input and progress to the most complex. For that reason, it is suggested they be read in order. Each example builds on the previous ones, and each contributes to the complete explanation.

Issues regarding data metrics will be dealt with later in this paper. For now, it is only necessary to state that the encoding method of the Input CES will be based on scene-space CIE colorimetric values. Various color primaries and color spaces can be used to express that colorimetry; but again, such decisions are data-metric issues and not encoding-method issues. For simplicity, then, the objective of the input transforms in these examples will be to produce CIE XYZ tristimulus values from color data provided by disparate types of input device and media.

7.1. Input from an Ideal Digital Camera

The most direct input device for scene-space encoding would be an ideal reference digital camera having the following properties:

- The spectral responsivities would exactly match those defined for the CIE 1931 Standard Observer.
- The optical characteristics would be perfect, i.e., no spatial nonuniformity and no camera flare.
- The relationship between sensor exposure levels and digital code values output by the camera would be perfectly linear.

As defined by these properties, the reference camera is an imaging colorimeter. Figure 7.1.1 below illustrates a reference camera being used to capture the color stimuli of an original scene. Note that no input transform is required for this reference camera because the code values it produces directly correspond to the CIE XYZ values required for the Input CES.

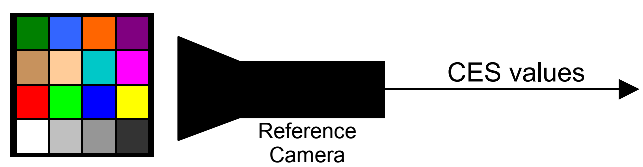

Figure 7.1.1: Input CES values are generated directly from original color stimuli by this reference camera.

This first example is straightforward, but it is also very important because it illustrates the fundamental concept of scene-space color encoding. The examples that follow will discuss inputs that depart in one or more ways from this idealized reference input. In each case, an input transform will be needed to compensate for such departures. When these transformations are designed and implemented correctly, the CES values derived from any of the inputs will match, as closely as possible, those that would have been directly generated by the reference camera, had it captured the same original color stimuli.

7.2. Input from a Colorimetric Digital Camera with Nonlinear Signal Processing.

In this next example the input again is a digital camera that is inherently colorimetric, with responsivities matching those of the CIE 1931 Standard Observer, and having optics that are ideal. However, in this case, code values output from the camera are not linearly related to the sensor exposure levels.

Nonlinearity is not unusual in actual cameras and can arise from several causes. Although virtually all electronic sensors are inherently "photon counters" that produce linear analog electronic signals, a camera's analog-to-digital converters may not be perfectly linear. In addition, the manufacturer may choose to introduce nonlinearity in the digital signal processing. The nonlinearity can be subtle or it might be significant if, for example, the code values are meant to correspond to HDTV digital video standards. For critical work, it is useful to determine if the signal processing is linear. If it is not, compensation must be included in the input transform. This can be implemented using 1-dimensional lookup tables (1D LUTs). In some cases, the departures from linearity might be somewhat different for each color channel. If so, the respective LUTs will also have to be different.

Compensating for nonlinear signal processing by using an appropriate input transform is simple if 1) the nonlinearity is known, and 2) it is mathematically reversible. Ideally, manufacturers would provide data

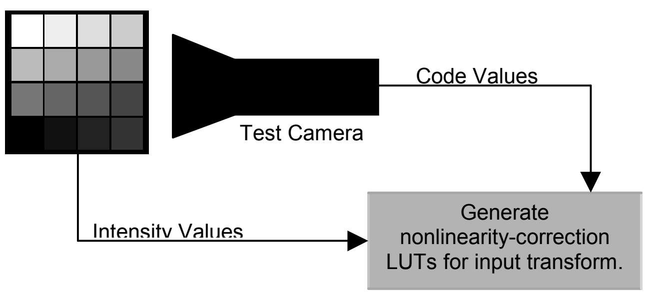

Figure 7.2.1: Generating input signal processing LUTs to compensate for camera nonlinear signal processing.

regarding any nonlinear processing included in their cameras; but they seldom will do so. An alternative is to characterize the processing experimentally. (My experience is that such experiments are worth doing even when a manufacturer has provided data. Discrepancies and inconsistencies are not uncommon.)

In the example being discussed, a linearity test is straightforward because there are no confounding effects due to optics. As illustrated above in Fig. 7.2.1, a set of neutral test stimuli of known relative intensities is first imaged by the camera. The corresponding camera code values for each area are then determined. These results are then used to generate the LUTs of an input transform for converting camera code values to relative linear intensity values. Transformed values would correspond directly to Input CES XYZ values due to the camera responsivities and ideal optics specified in this example.

7.3. Input from a Colorimetric Digital Camera with Optical and/or Sensor Nonuniformity.

In this example, the input again is a digital camera having spectral responsivities that match those of the CIE 1931 Standard Observer. We will assume that any nonlinearities in the camera signal processing have already been compensated for by an input transform. Unlike the previous examples, however, the camera is no longer quite so ideal. In particular, when an area of uniform light is imaged, light at the plane of the camera's image sensor is not uniform.

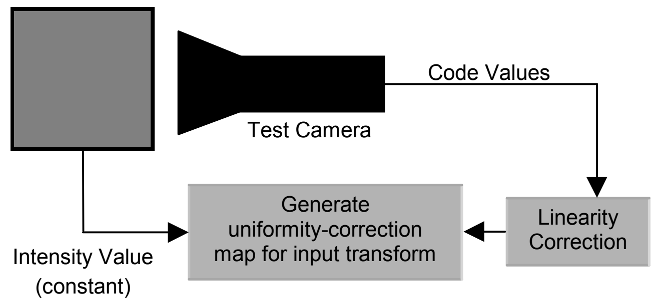

Figure 7.3.1: Generating a uniformity correction map for use in an input signal processing transform.

Due to lens falloff and other optical factors, such variations in light intensity at a camera's imaging sensor are likely. Also, a sensor may not produce the same signal value at each location when uniformly illuminated. The level of such nonuniformities in most professional cameras generally is not of great concern for most imaging purposes. For some applications, however, the colorimetric errors created as a function of spatial position may be unacceptable.

Generating a compensating transform for the effects of nonuniformity is quite straightforward. It can be done using experimental data generated using the setup shown above (Fig. 7.3.1), in which the test target is a perfectly uniform field. An ideal camera image of this field should, of course, produce code values that are identical at all locations in the image. That is highly unlikely. A correction map can be computed and built into the input transform as follows:

1. The image of the uniform field is processed through the nonlinearity-correction transform described previously. The processed image will then be in terms of linear intensity values.
2. The processed image provides a spatial map of the effective intensity nonuniformity of the camera, which includes optical and sensor nonuniformities.
3. Since the nonuniformity map is in linear exposure space, a compensating map is essentially its mathematical reciprocal.

The compensation map would then be included as part of the input transform, following the LUTs used to provide linearity. Images passing through the complete input transform would have Input CES values that are linear and independent of spatial location. Again, the ultimate goal is to have input from all actual sources match that from the theoretical reference camera of Example 1, and that is what has been accomplished in this more realistic example.

An alert reader may have noticed the somewhat Catch 22 situation represented by the two previous examples. The uniformity-correction transform was built using camera values that had first been processed through a nonlinearity-correction transform. However, that transform was built using a test chart like the one shown in Fig. 7.2.1. If the camera has uniformity problems, code values measured from an image of the chart will be incorrect. It would seem, then, that camera nonlinearity cannot be measured without first correcting for nonuniformity; but nonuniformity cannot be measured without first determining nonlinearity! Fortunately, there are ways around this apparent dilemma. One such way, which has been used in traditional photography for more than 60 years, is based on the use of a compensating gray target, such as that illustrated in Fig. 7.3.2 below.

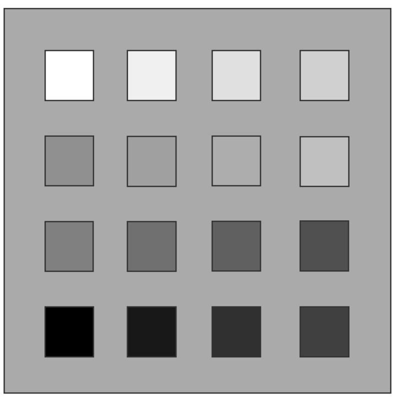

Figure 7.3.2: Compensating gray charts such as this can be used to correct image measurements in situations where camera and/or lighting uniformity is a problem.

The chart in this figure has the same test patches as those of the chart in Fig. 7.2.1. The difference is that in the compensating gray chart, each patch is surrounded by areas of uniform gray. When a test patch in a chart image is measured, measurements also are made of the surrounding gray areas. These measurements then are used to normalize the patch measurement. Applying this normalization process to each patch effectively corrects the measurements for camera (and lighting) nonuniformity. Based on much personal experience, the use of such charts is highly recommended.

7.4. Input from a Colorimetric Digital Camera with Camera Flare.

In this example, the input once again is a digital camera having spectral responsivities that match those of the CIE 1931 Standard Observer. We will assume that any camera nonlinearities and nonuniformities have been compensated for by the use of appropriate input transforms. However, the camera has another optical problem in that is produces flare.

Camera flare results from light scattered by the camera's lens and other components. It is present in all cameras, although to various degrees. The amount of flare can be expressed as a percentage of the light at the image plane from a perfect white in a photographed scene. Professional cameras can have as little as 0.25% flare, although it can be considerably higher depending on the lens type, cleanliness, and f-stop used. A value of 1% is reasonable for this discussion.

Camera flare is a non-imagewise redistribution of light in which light scattered from brighter image areas is added to whatever light is present in darker areas. Although this addition of light is linear, the perceived effect of that addition is quite different. This is illustrated in Fig. 7.4.1 below.

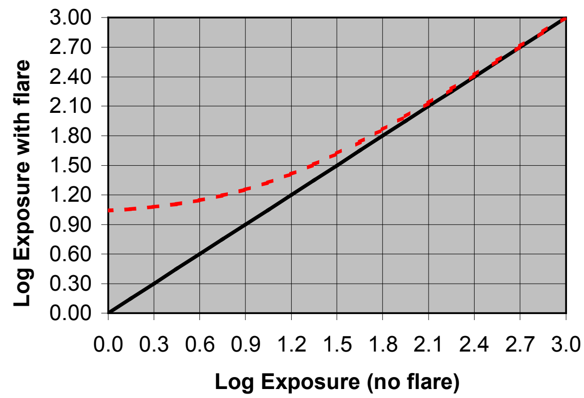

Figure 7.4.1: The effect of camera flare. The red (dashed) line corresponds to 1.0 % flare.

In the figure, the black (solid) line represents zero camera flare and the red (dashed) line represents the effect of 1.0 % camera flare. The axes are logarithmic to better illustrate the perceived effect of flare at low levels of exposure. In the final display, even small amounts of camera flare will cause darker areas to appear lighter or "smoky", and darker colors will appear noticeably less saturated. Both effects result from the initial addition of (neutral) flare light within the captured image.

Because the effects of camera flare are so apparent, the errors it introduces in the measurement of scene colorimetry should be compensated for if possible. Some camera flare, such as the creation of a halo around a candle flame, is local and cannot be removed using an input transform that will be applied to all images. However, much of the effect is global and can be compensated for to a significant degree.

When camera flare is present, it will confound a camera signal-processing measurement performed as described in the previous examples. That may not necessarily be undesirable. The measured grayscale will include the net effect of flare and signal processing. An input transform based on that measured grayscale will correct the total nonlinearity present. Thus it will include an approximate compensation for camera flare, which may be satisfactory for most applications. This essentially is the strategy used in photographic films. Their grayscales are designed to compensate for an anticipated amount of flare.

An alternative is to measure camera flare separately using techniques known in the trade. One method involves photographing a test target consisting of a "black hole" light trap and a white surround. The resulting image is then measured to determine the percentage of (flare) light that reached the black area in the camera image.

Once the camera flare is quantified, its value can be used to correct colorimetric values measured by the camera. This is quite simple to do in a linear scene space, as shown later. First, though, we will pause for some reflection on what has been covered to this point and for some philosophical discussion.

7.5. Encoding Philosophies.

In the previous examples, it was shown that input transforms can be used to compensate for camera signal-processing nonlinearities, optical and/or sensor nonuniformity, and camera flare. The construction of transforms requires some effort, however, which raises the question of whether the corrections they impart are really necessary.

Certainly if the Input CES is based on linear values, any fundamental camera signal processing nonlinearity (such as a video camera "gamma" curve) must be accounted for in the input transform. Aside from that, however, the question is more application-specific and philosophical.

If the objective is to make the camera serve as a colorimetrically-accurate imager for an application such as medical diagnostics, then yes, every effort should be made to compensate for factors that would reduce the accuracy of the device. For more conventional applications, however, that may not be necessary; and it may not even be desirable.

For example, two camera models may be virtually identical and have signal processing that is substantially linear. But each may depart slightly from perfect linearity in different ways. As a result, each camera will have a different "look" or "personality". Use of camera-specific input transforms would eliminate the subtle departures from linearity, effectively making images from the cameras identical. That generally would be good for projects that involve forming composite images from the two cameras. However, eliminating the distinction between the cameras might not be acceptable to a cinematographer who has a preference for one camera's particular "look" over that of another camera.

Similarly, the vignette effect produced by lens falloff may be aesthetically pleasing in some types of images. Even camera flare can produce effects that might be desirable in some circumstances because it can produce a softer, lower contrast, less colorful rendition that might be desirable for a given scene. My point, then, is that the question of whether or not to use *all* the capabilities that can be provided by input transforms is as much about preferences as technology. In the transformation to scene space, individual "personality traits" of the input devices and media can be removed to make all inputs essentially identical or they can be left in place to preserve individual looks. That means there are decisions that must be made, and the question of how they are made is something that will be discussed later. For now, let me summarize the options regarding scene-space input transformation components for electronic cameras, as discussed so far.

If a camera's code values are substantially linear to sensor exposure (the case for most camera RAW files):

- No linearity correction is required for normal imaging applications. Subtle departures in linearity normally produced by the camera will be retained in the encoding and passed through to the final displayed image.
- A transform can be used temporarily, if needed, to perfect the linearity when making measurements of camera nonuniformity or flare.
- A transform can be used permanently, if needed, to perfect the linearity when the camera is used in objective colorimetric imaging applications.

If there are optical and/or sensor nonuniformities in the camera:

- No correction is required for normal imaging applications. Encoded image values will include the effects of any nonuniformity, and those effects will appear as they normally would in the displayed image.
- A transform can be used temporarily, if needed, to remove the effects of any nonuniformity when making measurements of camera flare.
- A transform can be used permanently, if needed, to perfect the uniformity when the camera is used in objective colorimetric imaging applications.

If there is camera flare:

- No correction is required for normal imaging applications. Encoded image values will include the effects of the camera flare, and those effects will appear as they normally would in the displayed image.
- A transform can be used temporarily, if needed, to remove the effects of flare when making measurements for colorimetric transformations (discussed in the next sections).
- A transform can be used permanently, if needed, to compensate for flare when the camera is used in objective colorimetric imaging applications.

7.6. Input from a Colorimetric Digital Camera with Non-Standard Responsivities.

In the previous examples, each digital camera was an inherently colorimetric device. That is because each had spectral responsivities corresponding to those of the CIE 1931 Standard Observer.

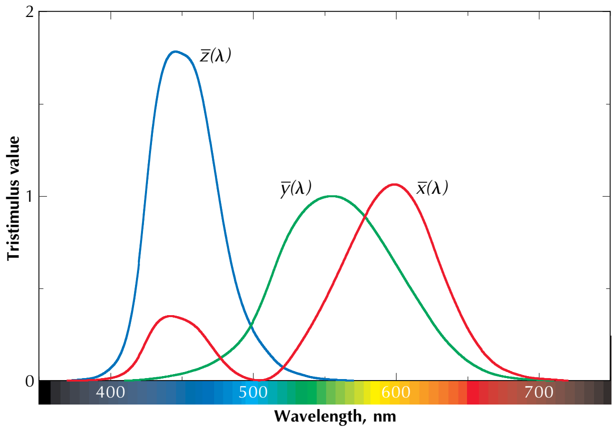

Figure 7.6.1: Spectral responsivities for the CIE 1931 Standard Colorimetric Observer.

The spectral responsivities specified for the CIE 1931 Standard Observer are shown in Fig. 7.6.1 above. Because these responsivities are all-positive (no imaginary negative lobes), it would be possible to build a camera having equivalent responsivities. That feature would simplify the task of generating CIE XYZ values for the Input CES. To the best of my knowledge, however, no such digital motion picture camera exists. There are engineering reasons, primarily related to the significant overlap of the "red" and "green" responsivities, why building a camera with these responsivities is problematic.

Nevertheless, a colorimetric camera can still be practical if it is based on some other set of CIE color-matching functions where the responsivities have less overlap. This is a tactic frequently used by camera manufacturers. In this example, it will be assumed that the camera responsivities exactly correspond to a set of CIE color-matching functions other than those of the Standard Observer.

A fundamental principle of color science is that all sets of color-matching functions are simply linear combinations of all other sets. That means that any set can be transformed into any other set using an appropriate 3x3 matrix. Moreover, it means that a 3x3 matrix can transform the tristimulus values corresponding to one set of color-matching functions to those corresponding to another set.

All that is required, then, to allow the camera of this example to be used for input to the system is the inclusion of an appropriate 3x3 matrix in the input transform. It must operate in linear space, so it must be placed after the linearization operation.

Determining the matrix coefficients is a textbook procedure. The color-matching functions of the camera will be associated with a unique set of RGB primaries. When the primaries are known, the derivation of a transformation matrix to CIE XYZ values is a simple algebraic exercise.

7.7. Input from a Non-Colorimetric Camera.

In this example, the digital camera that will be used is no longer inherently colorimetric. This means its spectral responsivities do not correspond to the color-matching functions of the CIE 1931 Standard Observer, nor do they correspond to any other set of visual color-matching functions.

This example is entirely realistic because camera spectral responsivities always differ, at least to some extent, from actual color-matching functions. There is a long list of reasons why manufacturers make deliberate departures from color-matching functions. (It is *not* simply a lack of understanding of the color science involved.) Regardless of the reasons, the result is that for virtually all real cameras, the encoding problem becomes one of deriving colorimetric information from a device that is not inherently colorimetric. Doing so is, of course, impossible. Nevertheless, in practice it has to be done anyway.

Before going any further, an important side note is that this problem is not *caused* by the use of a colorimetric scene-space Input CES. Transforming into some other type of CES would not change or eliminate the basic problem in any way. Use of this type of CES just makes the same inherent problem more explicit and more directly quantifiable.

How can non-colorimetric camera RGB values be transformed into colorimetric values? The methods used by manufacturers are complex and proprietary. When done well, the consequences of colorimetric capture errors are minimized. Creating an appropriate transform, in the form of a 3x3 matrix, involves at least as much art as science. For example, it might seem reasonable to derive a matrix simply by running a regression between camera (linear) RGB values and corresponding CIE XYZ values. If the camera spectral responsivities are known, camera RGB values (and corresponding XYZ values) can simply be calculated for an array of color stimuli. The linear regression would generate a matrix that minimizes the mathematical errors, but it is very unlikely the matrix will produce the most pleasing images. That is where art must be used together with science.

Manufacturers have mathematical tools, image simulation capabilities, and experienced people who can determine matrices that minimize the perception of color shifts resulting from input colorimetric errors. My suggestion, then, would be to have manufacturers furnish the RGB-to-XYZ transformation matrices required for input. It would be to their advantage to provide the best possible transforms for their products, and they already generate such matrices anyway. Virtually all cameras include a data path that includes a transform to ITU-709 RGB primaries (for video, sRGB, HDTV, etc.). The conversion from 709 to XYZ is a defined matrix, and the two matrices can be combined to produce the Input CES transform matrix.

7.8. Input from a Camera Having Unknown Spectral Responsivities.

In the previous example, it was suggested that the preferred method of obtaining a matrix for transforming camera RGB values to CIE XYZ values was to obtain it from the camera manufacture. Also described was an alternative method based on computations of camera RGB values. These computations require knowledge of the camera's spectral responsivities. That data may or not be available from the manufacturer. Even when it is, its accuracy should be considered somewhat suspect.

There are many factors that contribute to a camera's net spectral responsivities. A partial list includes the basic spectral sensitivity of the photodetector and the spectral characteristics of the color filter dyes, the IR and other filters, and the lens and other optical components. These individual components generally are produced in batches, often by a variety of suppliers, and the variability of every component contributes to the total variability in effective responsivities. Therefore, while published data can be considered representative for the camera model, it is unlikely to correspond to the exact responsivities of any individual camera.

For very critical work, then, or in situations where the manufacturer can supply neither a transformation matrix nor the camera spectral responsivities, the responsivities can be measured experimentally. In concept, this is straightforward for electronic cameras; but the equipment involved (optical benches, calibrated monochrometers, etc.) is not commonly available. The measurement procedures are also somewhat time-consuming and can be quite tedious, although they are not unreasonable if relatively few cameras are involved.

There are two basic approaches that can be used in measuring spectral responsivities. In one, the camera is sequentially exposed to a series of monochromatic (or narrow band) light, and camera red, green, and blue exposures (linear code values) are measured at each wavelength band in the series. These measurements, when corrected for any wavelength-band intensity variations in the light source itself, define the camera responsivities.

I prefer to use a variation on this method. Again in the first method, the basic concept is that a constant intensity of light will be presented at each wavelength, and the varying camera signals will indicate the camera's responsivities. That means the results will be influenced by any nonlinearities that might exist in the camera. Since even a 1% measurement error can result is significant color errors, my preference is to use a technique that essentially reverses the experiment. In this method, at each wavelength in the series, the light source intensity is adjusted until some predefined constant signal value is produced in the camera. The light intensity needed to produce that signal is then recorded. For a color camera, of course, three different intensities would be required at each wavelength setting. (This experiment can be *very* time consuming unless the signal values can be read essentially in real time). When the measurements are complete, the results represent the reciprocal of the responsivities. (At wavelengths where responsivity is high, the amount of energy needed to produce the signal value is low, and vice versa). Taking the reciprocal of the intensity measurements, then, yields the responsivities. My experience is that this method provides very reliable results. Neither of the described methods is simple, however. Perhaps if the industry's need for this work is sufficient, a specialized service should be established to provide the required measurements.

An alternative method for developing an input transformation matrix when spectral responsivities are unknown is to avoid the need for that information by determining the matrix based on a *characterization* of the camera. Characterizing a device essentially means treating it as a "black box". For an input device, this is done by exposing the device to a known set of stimuli and measuring the resulting signals. When a sufficient number of stimuli and corresponding measurements have been collected, a correlation that describes the behavior of the device can be established.

A digital camera can be characterized by photographing an array of illuminated reflection color patches, measuring the camera RGB exposures (linear code values) for each patch, and relating those values to the corresponding colorimetric values.

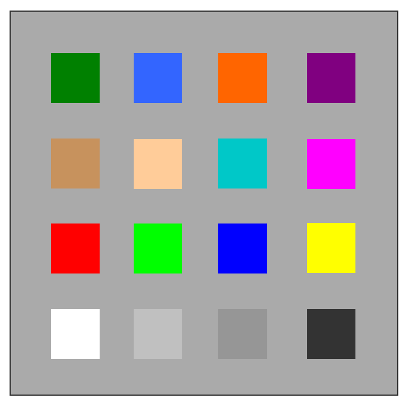

Figure 7.8.1: A color chart that allows measured camera values to be corrected for camera nonuniformity.

Figure 7.8.1 above illustrates a small array of test colors arranged on a uniform gray background. This arrangement allows measured camera exposure values to be corrected for camera nonuniformity. An alternative is to photograph one test color (placed in the same location) at a time, but that is a very tedious procedure. Whether lighting uniformity is a concern or not depends on how the colorimetry for the color patches is determined. If it is computed from patch reflectances and light source power distribution, lighting nonuniformity will have to be accounted for in the experiment. This and other complications can be avoided if instead the colorimetry of the patches is measured at the same time the photography is done. I would recommend use of a telespectroradiometer, such as a PhotoResearch PR-705 Spectrascan, placed next to the camera. Lighting uniformity and other possible variations are not a problem in this arrangement because the camera and the instrument are simultaneously looking at the same stimuli.

A critical element in this experiment is the selection of color patches. There are several factors that must be considered; in particular, the number of patches and their spectral distributions. If too few patches are used, the results may not be indicative of the camera's response to real world colors. Of course that concern must be balanced against using an unnecessarily large number of patches and making the test more difficult than necessary. Test colors should be selected to include all areas of color space of interest. Moreover, the spectral characteristics of the colors must be considered.

A common mistake is to use test colors from a single family, such a set of Munsell colors. Although hundreds of Munsell colors are available, most are created from a relatively limited set of colorants. So a large array of such colors may effectively represent only eight or ten unique spectral vectors. Very often the outcome of using such color sets is a transform that works well for those colors, but not for anything else. The test set, then, should contain color patches having not just similar colorimetry but also similar spectral characteristics to real-life colors important to the particular application.

It is not possible to recommend a specific number of test colors or to specify their spectral distributions. Those factors will vary depending on the application and the nature of the camera responsivities. If the responsivities exactly correspond to a set of color-matching functions, there will be one and only one transform from camera RGB to target colorimetric values. Any set of colors, then, will generate the same result. The more the responsivities differ from a set of color-matching functions, the more the generated transform will depend on what test colors are used. It can be useful to test that dependence. If it is determined that the transform is highly dependent on the color set, the color set will have to be carefully selected for the specific application. For critical work involving only a subset of colors, it might even be advisable to use a transform derived specifically for those colors alone. In deriving a transform for portraiture, for example, the color set should contain numerous patches with spectral properties representative of skintones and hair colors. Once again, there are decisions and judgements to be made here that involve as much art as science.

7.9. Input from Multiple Digital Cameras

In each of the examples to this point, components for an input transform to Input CES values have been developed based on modeling and characterizations of individual devices. If nothing else, this discussion has shown that transform development requires a fair amount of effort. It is advantageous, then, to minimize the number of transforms needed in a system. Different transforms are required, of course, for different types of devices. However, it is not necessary (or good practice, in my experience) to construct unique transforms for each individual device.

In many systems, the number of input devices can be substantial, but many may be nominally the same. For example, a system may have twenty individual scanners, but of only two different models. In such situations, it makes sense to divide the input transform into two fundamental components, characterization and calibration, which are defined as follows:

Characterization is a procedure for defining the color characteristics of a representative model of a given type of device. *Calibration* is the procedure of correcting for any deviations of a particular device from the characterized representative model.

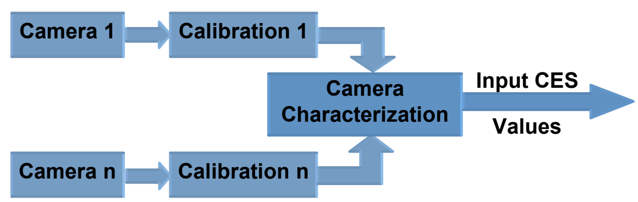

Figure 7.9.1: Use of individual calibration transforms and a common characterization transform for input processing.

Figure 7.9.1 above illustrates the use of calibration transforms for individual cameras of the same type, together with a characterization transform for that camera type. Although it might seem that all this has done is add to the number of transforms, there actually is an advantage to this arrangement. As has been discussed, characterization transforms can be difficult and time consuming to construct. Calibration, on the other hand, is generally much simpler. It should be just a "tweak" of the code values of a device to make it correspond to those of the characterized type. As such, calibration transforms can be built much more quickly and easily (often automatically), using far fewer test colors (usually just gray or color scales).

It should be pointed out that camera calibration is no more or less critical for scene-space encoding than it would be for any other type of color encoding. Whether it is used or not depends once again on the intended application. Calibration can help ensure that images from a group of cameras are interchangeable. If that is not important, the characterization transform can be used alone, without the calibration transforms.

7.10. Input from Black-and-White Negatives

The process of transforming images into scene space is fundamentally the same regardless of the type of devices or media involved. The basic steps can be summarized as follows:

1. A transform relating image signal values to original exposure values is determined by modeling or characterization.
2. Image signal values are acquired.
3. Image signal values are processed through the transform to determine original exposure values.

The examples discussed to this point were for digital cameras because they directly provide image signal values in digital form. Photographic films must, of course, be scanned to provide such values. The scanned values are those of images formed by chemical signal processing that is often quite complex. As a result, it can be considerably more difficult to determine a transform relating scanned values to original exposure values.

In the case of black-and-white films, however, the relationship between scanned values and exposure values is straightforward. That makes it a good place to start the technical discussion of film input, and it provides an opportunity to further discuss some of the philosophical issues involved in scene-space encoding.

A film scanner is simply a scanning densitometer. Black-and-white film contains only one channel of information that is represented in terms of a silver image that is nearly spectrally flat. That makes the measurement of image signals very straightforward. Like any densitometer, a scanner requires periodic calibration (adjustments of electronic gain and offset). Such calibration, which is often automatic, is standard in virtually all scanners. For very critical work, some field uniformity correction might be useful, but it is not necessary for normal applications. Some scanners automatically provide this uniformity correction.

The relationship of scanned density values and exposure values for a particular black-and-white film can be characterized by scanning an image of a grayscale target of known exposure values. The results can be used to construct a characteristic curve relating Scanned Density and Relative Log Exposure, such as that shown in Fig. 7.10.1. This characteristic curve is the transform required for encoding.

Grayscale-characterization images can be exposed using a sensitometric instrument specifically designed for that purpose, or they can be created simply by photographing a chart of neutral test patches. A compensating gray chart, discussed earlier, would be an excellent choice. It would be difficult, however, to build a chart in which the dynamic range of the test patches is sufficient to cover the extensive exposure dynamic range of a photographic negative film. For example, on a Macbeth ColorChecker chart, the reflectance of the white patch is about 89%, and for the black patch it is about 3%. That corresponds to a range of only about 1.47 in log exposure. The exposure dynamic range of a typical photographic negative film is several times that large. However, that problem can be overcome using a simple trick: The chart can be photographed in an exposure series. Each exposure will cover a different portion of the film's exposure range. Depending on the dynamic range of the test chart, covering the entire range of the negative may require an exposure series of four or more images. It is good practice to bracket the exposure series such that there is overlap. This produces useful redundancy in the experiment. For example, the same film density should result from the 20% gray patch of a normal exposure, the 40% gray of a one-stop underexposure, and the 10% of a one-stop overexposure.

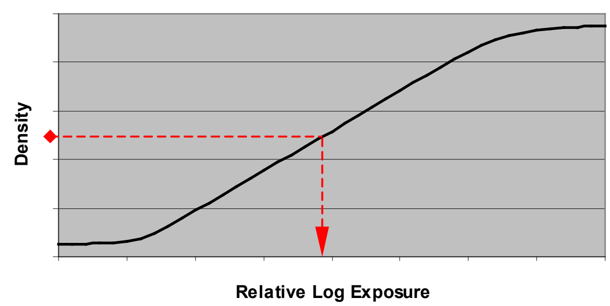

Figure 7.10.1: Example graysclae characteristic relating scanned values to relative exposure for a B&W film.

It should be pointed out that if the grayscale is determined from camera images, the effects of camera flare would be included in the measured scale. If the grayscale is left uncorrected for flare, its use in the input transform will (approximately) remove the contribution of camera flare. If instead it is desired to retain the effect of camera flare in images to be encoded, a flareless grayscale such as that produced by contact exposure in a sensitometer can be used. Alternatively, the amount of camera flare can be determined independently and used with the measured grayscale to compute a flareless grayscale.

Once the test-chart exposures have been made, the film is processed, and the images are scanned to measure their density values. Alternatively, the images could be measured on an ordinary densitometer. However, the silver of black-and-white images scatters light, so density readings can be influenced by measurement geometry and other optical factors. The consequences of any such densitometric differences are canceled if the same device that will be used to scan normal images is used for measuring the characterization images. Once the density values have been determined, they are used to construct the grayscale characteristic curve.

The red (dashed) line in Fig. 7.10.1 illustrates the transformation of a scanned density value to relative log exposure values. That essentially is all there is to the scene-space input encoding method for black-and-white films. The remaining steps of the encoding only involve data metric conversions. This might include a normalization to ensure that a properly exposed reference patch (usually a reference white patch) will have the correct encoded value, an exponentiation to linear exposure values, and finally a conversion to Input CES data metric values.

If the CES is encoded in terms of normalized RGB exposure values, the exposure values of all three channels would be set equal to the normalized exposure value of the film. If the CES is encoded in terms of XYZ values, the Y tristimulus value would be set equal to the film's normalized exposure value, and the corresponding X and Z values computed so as to produce a specified chromaticity. The chromaticity aim could be that of an achromatic neutral; or, if desired, it could correspond to some desired color tint. The following equations can be used to determine the X and Z values for a pixel having a given Y value and chromaticity aim values x and y:

$$X = x\,(Y/y), \text{ and}$$

$$Z = z\,(Y/y), \text{ where}$$

$$z = 1 - x - y.$$

The degree of accuracy of this procedure for a given image depends in large part on how well the grayscale characteristic used in the transform represents the actual grayscale inherent in the image. For ultimate accuracy, the characterization grayscale chart would be included within the image itself. That would ensure the grayscale and the rest of the image stay together through subsequent latent-image keeping, chemical processing, and scanning. Although this technique has been used for very critical image-simulation work, it is unlikely to be required in almost any other application.

In practice, there is a wide range of acceptability that will determine how closely the characterization grayscale must correspond to an image to be encoded. The grayscale could, for example, be placed on a frame adjacent to the image, on an end of the same piece of film, on another roll of film from the same batch, or on a different roll of the same film product. The characterization grayscale could, in fact, be a generic curve, representative only of the basic type of film.

A decision as to how closely the characterization must correspond to the images to be encoded obviously depends on the degree of accuracy required in the extraction of film exposure values. Less obvious is that the use of a more generic characterization actually might be *preferred* in some applications.

Consider the two characterization curves in Fig. 7.10.2. Assume the black (solid) curve represents an average film grayscale. The blue (dashed) curve then represents a film of lower-than-average contrast. If the same scene were photographed with both films, a darker area of the scene corresponding to the exposure labeled "1" would produce essentially the same density on each film. A brighter area of the scene, corresponding to the exposure labeled "3", would produce the density value labeled "DH" on the average film and "DL" on the lower contrast film.

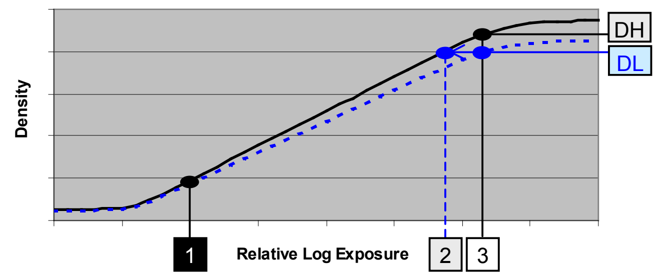

Figure 7.10.2: The effects of using two different gray-scale characterization transforms.

If the encoding of the average film image uses a transform based on the grayscale of that film, and the encoding of the lower-contrast image uses a different transform based on the grayscale of its film, the respective "DH" and "DL" densities will produce the same exposure value of "3". The result of using these film-specific transforms, then, is that the encoded images from the two films will be identical. That accurately reflects the fact that both films photographed the same original scene. For some applications, that outcome is ideal. For example, it allows images from the two films to be merged seamlessly in forming composite images. However, a photographer who had deliberately chosen the lower contrast film for its look would not be pleased with this film-independent outcome.

As discussed earlier, the individual "personalities" of input media can be retained if images are processed through "generic" or "universal" transforms corresponding only to the basic type of device or medium. In this example, that transform would be based on the grayscale for the average film. As shown in Fig. 7.10.2, when the density value "DL" of the lower-contrast film is transformed through the average-film grayscale instead of its own, the resulting exposure value will be "2" instead of "3". Use of the average-film transform, then, will have the effect of compressing the encoded exposure values, thus retaining the lower-contrast look of the film.

This dual nature of scene-space encoding can be confusing, but it addresses an important issue that will arise again and again. I would go so far as to say that the acceptance of the proposed system will depend in large part on how well the versatility of scene-space encoding is understood and conveyed to the industry. The ability to select the encoding paradigm appropriate for a particular application must be part of the final system, and the fact that selection options are available must be emphasized.

7.11. Input from Color Negative Films

In all the examples discussed to this point, the input devices and films inherently have had a characteristic that I refer to as *channel independence*. Independence means that an imaging channel produces a signal that is detectable and separate from signals produced by other imaging channels.

Black-and-white films are "channel independent" by definition because they have only one channel. They capture a single channel of exposure information and produce a single and measurable signal in the form of a silver image. A monochrome electronic camera is similarly channel independent. Color electronic cameras also are channel independent devices because each of the red, green, and blue image-capture channels produces its own detectable electronic signal at the output of the image sensor(s). That independence simplifies the process of transforming image signal values to exposure values.

If the color channels of photographic film were similarly independent, the process of determining scene exposure values from scanned density values would be essentially the same as that for black-and-white films. The only difference would be that each pixel would have three density values and three corresponding exposure values, which would be determined using three respective characteristic curves of a film grayscale exposure. That process is illustrated in Fig. 7.11.1 below.

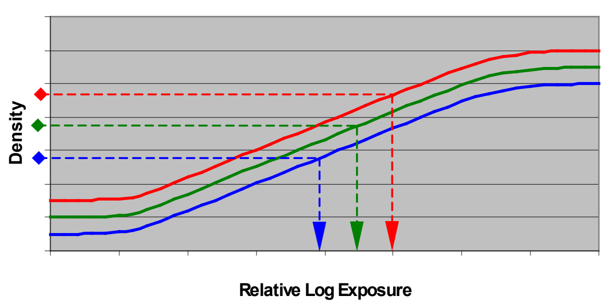

Figure 7.11.1: Transformation of a set of RGB density values to corresponsing RGB exposure values using respective characterization transforms.

The illustration suggests that red exposure values can be determined from red density values, green exposure values from green density values, and blue exposure values from blue density values. That is valid only if the measured density values for the three color channels are independent. Unfortunately, such independence is not inherent in color photographic films. Therefore measured red density values, while primarily indicative of red exposures, are also related to exposures that may have been captured by the green and/or blue image layers of the film. For the mapping process shown in the above figure to be valid, any density interdependence present in both the image and characterization scales first must be overcome.

There is only one point in the process of forming a color image when the color channels are independent. That is when the film has been exposed but not yet chemically processed. At that point the exposure recorded in the red-sensitive layer is, by definition, the red exposure. That red exposure signal is exactly what the scene-space encoding process is intended to determine. Similarly, the exposure signals for green and blue are recorded in their respective layers. These recorded signals, although independent, are in the form of a latent image that essentially is undetectable. Chemical processing is therefore needed in order to amplify these microscopic signals and form measurable color images. It is in that chemical process and in subsequent optical measurements that color channel interactions are formed.

In the image-forming stage of chemical processing, exposed silver is developed, and an associated dye image is produced. Numerous byproducts of the chemical reactions involved (some deliberate, others unavoidable) can influence the silver development and/or dye formation taking place in other layers. So, for example, development of the latent image of the green-sensitive layer can affect the amount of silver and image dye formed in both the red-sensitive and blue-sensitive layers. As a result, some fraction of the green exposure information will have crossed into the dye-image signals of the red and blue channels. Similarly, red and blue exposure information will cross to the dye-image signals of the other two channels. Thus even if a film were measured in terms of its analytical dye amounts, the measured signal values would not be channel independent.

Once an image is formed, it must be optically measured. That process also creates color-channel interactions, even in the absence of prior chemical interactions. Assume, for a moment, there were no chemical interactions. Exposure in the red-sensitive layer of the film then would result in the formation of cyan dye and no magenta or yellow dye. The cyan dye image itself would be independent of the other channels, but scans or other optical measurement of that dye would create another form of color channel cross-talk. This is illustrated in Fig. 7.11.2.

The upper graph of the figure shows spectral transmission density curves for the cyan, magenta, and yellow image-forming dyes of a representative color negative film. The lower graph in the figure shows the spectral responsivities for a representative film scanner. Note that the cyan dye has optical density not only in the region measured by the red responsivity of the scanner, but also (to a lesser extent) in the regions measured by the scanner's green and blue responsivities. Similarly, the magenta dye has optical density not only to the scanner's green responsivity but also to its red and blue, and the yellow dye has optical density to the scanner's green responsivity in addition to its primary density to its blue responsivity.

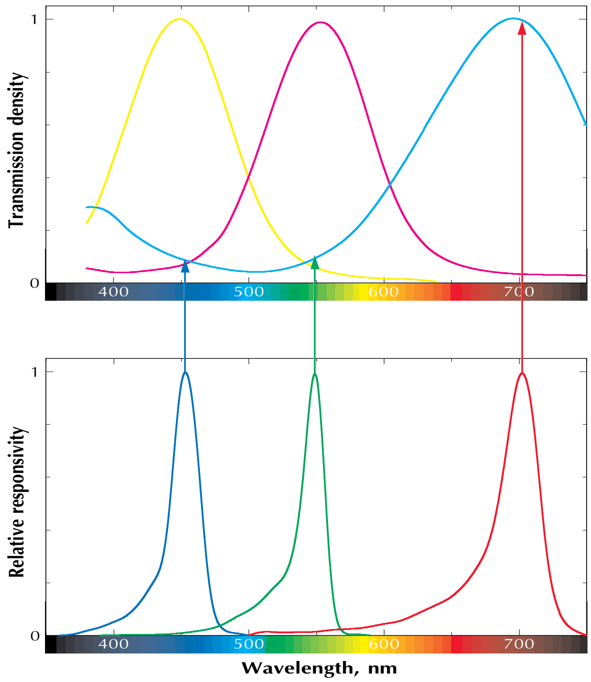

Figure 7.11.2: Color channel interdependence resulting from optical measurements of image-forming dyes.

The problem, then, is that the use of grayscale characterization curves for transforming color density values to color exposure values requires the density-to-exposure relationships to be 1-dimensional. As has just been shown, however, interactions among the color channels instead cause those relationships to be 3-dimensional. To solve this problem, a transform can be used to computationally remove the net effect of the chemical and optical cross-talk interactions from scanner measurements of grayscale-characterization images. This generates a set of channel-independent grayscale characterization curves appropriate for use in the film Input Transform. By applying the *same* transform to scanned image values, they too are transformed to channel-independent density (CID) values that can be used with the channel-independent 1D characterization scales to determine film exposure values. The process is illustrated in Fig. 7.11.3 below.

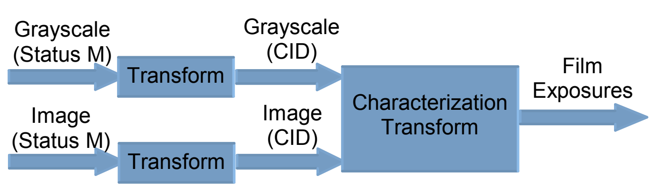

Figure 7.11.3: Use of a transform to generate channel-indendent grayscale-characterization and image values.

Successful implementation of this technique depends in large part on how well the transform removes color channel interdependencies. Depending on the particular film it is derived form, the transform can be quite difficult to derive. I would suggest, then, that film manufacturers assume the responsibility for providing product transforms. This would ensure that the transforms are generated by those having the most experience with their creation. It would be unreasonable, however, to expect manufacturers to provide transforms not just for each film but for every combination of scanner and film. The number of transforms required would be entirely impractical.

Therefore, I further suggest that the industry adopt standards to define the densitometric properties of motion picture film scanners. Specifically, I suggest the use of the ISO Status M standard for film scanners used with color negative films. Status M was, in fact, originally developed for densitometry of motion picture color negative films, and it is still well suited for that application. Moreover, Status M densitometers are widely available, relatively inexpensive, yet very reliable, precise, and accurate. Accordingly, a densitometer provides an excellent means of verifying scanner calibration: On a properly calibrated scanner, scanned values for a set of film test colors should match those read on the densitometer.

If this densitometric standard were adopted for all scanners, only one transform (at most) would be required for each film product. As shown previously in Fig. 7.11.3, the transform would convert Status M RGB values to channel-independent RGB values. Far fewer transforms would be necessary if the use of a "generic" transform is acceptable (or preferred, as discussed earlier) for a related group of film products. I would suggest that manufacturers provide generic transforms for their products and at least one product-specific transform that can be used in applications where the highest degree of accuracy in determining original scene colorimetry is required.

For a practical implementation of scanner calibration, I would urge the industry to produce film calibration targets in a format suitable for scanning. The targets should contain a full grayscale and a variety of test colors. Certified ISO Status M values for each color patch should also be provided with the target. Such targets would be relatively easy to generate, and automated programmable densitometers could be used to measure the densitometric values.

Use of a film target for scanner calibration would be quite straightforward. The first step would be to follow the scanner manufacturer's procedures for routine calibration. The calibration target then would be scanned, and the scanner RGB values for each patch would be determined. A simple regression would then be run to determine a transform relating measured scanner RGB values to reference Status M RGB values provided with the target. In most cases, a simple 3x3 or 3x4 matrix should suffice (See Appendix 1). The regression should not be restricted because differences in scanner spectral responsivities from those of Status

M produce color contrast differences that require an unnormalized matrix for correction. Also, it would be advisable to examine the resulting matrix. If the coefficient values are large, it indicates that the scanner's responsivities are significantly different from those specified for Status M. It might be advisable to determine if filters or other components could be changed to correct that difference.

Scanner calibration eliminates one major source of variability in the encoding process. Another potential source of variability is film (chemical) processing. Process variations are unlikely to have a significant effect on the 3-dimensional color characteristics of the processed film because the typical variations do not alter chemical interactions or image dye hues. It is the grayscale characteristic that is most likely to be affected. In particular, grayscale minimum density (D_{min}) levels can vary, and this will cause problems in determining film exposure values.

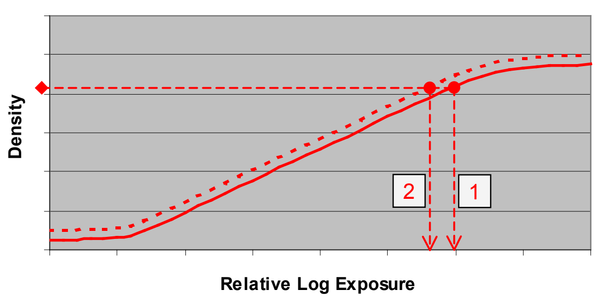

Figure 7.11.4: Uncorrected, shifts in film minimum densities result in errors of determined exposure values.

Figure 7.11.4 above illustrates the effect of a D_{min} shift in one color channel. The solid line represents the red characterization curve of the transform, and the dashed line represents the curve of the actual scanned film. For the example red density value shown, the characterization transform would yield the exposure value labeled "1". The correct value would have been the exposure labeled "2". Thus a density shift resulting from a D_{min} variation has been misinterpreted as a different film exposure value. In the linear portion of the curve, exposure values will simply be shifted higher or lower. While such shifts are not particularly desirable, they can be corrected by subsequent adjustments of overall exposure and/or color balance. However, exposure values derived from density values in the nonlinear toe and shoulder portions of the characteristic curve will be distorted, and some values will be clipped at the ends of the curve.

The effects of process variations can be avoided, of course, if new characterization curves were built for each processed batch of film. However, my experience is that it should not be necessary to go to that extreme, since most variations are primarily in D_{min} levels. I suggest, then, that all characterization transforms be built in terms of *D_{min}-subtracted* Status M density values. Likewise, scanned image values should be provided to the transform as D_{min}-subtracted Status M density values. This can be accomplished easily by scanning a minimum-density area of the film as the images are scanned. The D_{min} values can then be subtracted as part of the process of converting scanner values to D_{min}-subtracted Status M values. This method has the added benefit of nullifying any drifts in the zeroing of the scanner electronics.

Referring again to Fig. 7.11.3, the procedures that have been discussed will transform film scanned density values to film exposure values. To complete the encoding to an Input CES based on scene-space CIE colorimetry, film exposure values must be transformed to standard colorimetric values. Because the spectral sensitivities of photographic films do not correspond to a set of visual color-matching functions, the transformation of film exposure values to CIE colorimetric values is not straightforward.

This topic was discussed earlier in Section 7.7 with regard to non-colorimetric digital cameras. I suggested that because camera manufacturers have the necessary expertise, they would be in the best position to provide the required transforms. That suggestion applies to film manufacturers as well.

As an alternative to providing both a film characterization transform (Status M RGB to film RGB exposures) and an exposure transform (film RGB exposures to XYZ values), a film manufacturer could instead provide a single transform in which these functions are combined. Doing so would provide a somewhat less explicit description of a film's behavior, which might be preferred by its manufacturer. However, as will be discussed later, it is useful to have direct access to film RGB exposure values at some point in the system. This requires separate transforms.

If manufacturers are reluctant to provide any transforms, another alternative would be for them to provide film targets and accompanying data specifying aim Status M density values, aim scene-space colorimetric values for each patch, and (preferably) film RGB exposure values. As was just discussed, such targets could be used for scanner densitometric calibration. In addition, the combination of film targets and accompanying data would provide information sufficient for constructing input encoding transforms.

7.12. Input from Print Films, Other Output Media

The primary inputs to the proposed system are digital cameras, computer-generated images, and negative photographic films. However, it is quite likely that on occasion there will be a need to input images from other media, including media that normally function as outputs. For example, it might be necessary to input images that only exist in the form of a motion picture prints. The challenge of meeting that requirement provides a good opportunity to demonstrate the flexibility and inherent inclusiveness of the system.

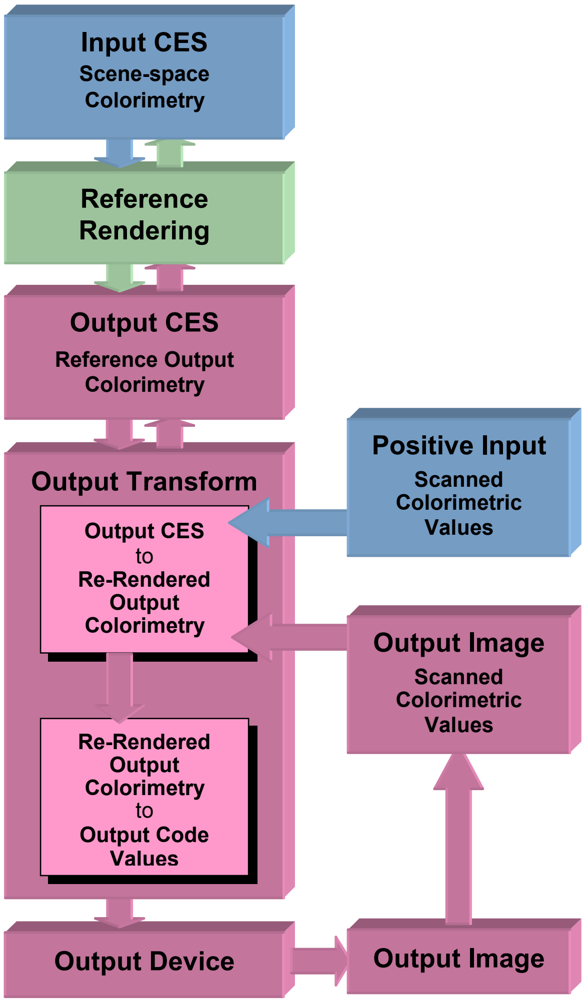

Figure 7.12.1: A method for transforming images from positive media to Input CES values.

The left-side arrows of Fig. 7.12.1 above illustrate the normal flow of color information in the system. Input CES values are rendered for a reference output medium to Output CES values, which are then transformed in a device-specific output transform to code values for the particular output. Two basic functions, described in more detail later, take place within an output transform. In the first, colorimetric values specified for the reference output are transformed to colorimetric values for the particular output device, medium, and viewing conditions with which the transform is associated. The second function determines the code values required to produce that colorimetry on the output. If the output is properly calibrated, an output image having the specified colorimetry will be produced.

Imagine now that scanned colorimetry of that output image were measured, and the values fed *back* into the system at the location shown in the diagram. If these values were then processed "backwards" through the signal-processing chain (right-side arrows), they ultimately would produce the same input CES values that produced the output image. In other words, a closed-loop path would have been followed from Input CES values to output-image measured colorimetry back to Input CES values. (That is always an interesting real-life exercise. It certainly reveals any calibration problems or other flaws in a system.)

Just as the described process can determine Input CES values for an image generated by the system, it also can determine Input CES values for other images on the same type of output medium. So for example, an output transform for a film writer and motion picture print film can be used to transform colorimetric values measured from any image on motion picture print film to Input CES values. Note that since the measurements are colorimetric, not densitometric, the particular dye set of the print film is not an issue.

Do Input CES values determined by this process represent *the* original scene originally photographed for that print? There is no way to know for certain without having access to complete information regarding the print film, the printer, the camera film, any intermediates involved, etc. I would argue, however, that it also does not matter.

The process will have created an image in terms of Input CES values based on *this* system, with its known data path, and with a known output medium. That means two things: First, if the Input CES image is processed through the system, it will generate a new image that is a visual match to the original print-film image. That in itself is an important capability. Second, and perhaps more importantly, derived Input CES values represent original-scene colorimetry that currently would generate the same appearance as that print film image. As such, those values are compatible and interchangeable with all other Input CES values. They can be treated the same in editing, adjusting, merging, and other input-image processing, and they can be processed by the system for output to any type of output device or medium included in the system.

Images from motion picture print films were used in this example because it seems reasonable that they are likely to be used for input. However, the same basic procedure can be used for any hardcopy or softcopy image for which colorimetry can be determined. The procedure is particularly easy to implement for images from devices or media comparable to those already included in the system. In such cases, the required signal-processing path can be created simply by inverting the first output transform and the reference rendering transform.

7.13. Input of Computer Generated Images

Computer-generated images can be brought into the system in a number of ways. I would suggest they be brought into the Input CES, where they then would be interchangeable with images input from all other sources. The process for doing that is described next.

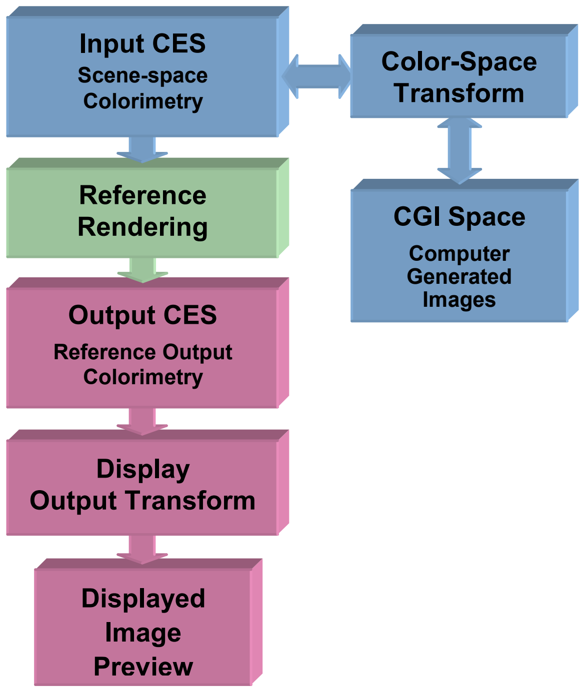

Figure 7.13.1: Input of computer-generated image data.

Computer-generated images could be created and manipulated directly in terms of Input CES values. It is more likely, however, that practitioners will prefer to work in terms of other color spaces that are familiar and perhaps more appropriate for the particular tasks being performed. For that reason, Fig. 7.13.1 above includes a transform relating the Input CES and a user-specified CGI color space. The two-way arrows indicate that images may originate in the CGI space and then be transformed to the Input CES and that other Input CES images can be transformed to the CGI space for inclusion in work being done there.

Because the Input CES directly specifies input-image scene-space colorimetry, transformations to and from commonly used CGI spaces are straightforward. If we assume, for example, that Input CES values are expressed in terms of (linear) CIE XYZ, then transformation to various linear RGB and YCC (luminance/chrominance) spaces require only a 3x3 matrix. Transformation to video code values would only require a matrix and a 1-D LUT. Such transformations are simple, fast, and reversible.

Also included in Fig. 7.13.1 is a preview display and its associated signal-processing from Input CES. The display would provide a real-time preview of final image appearance. This would be of great value in image creation. Alternatively, the Display Output Transform could include additional processing to assist image-creation work. For example, a transform that can temporarily boost the chroma of lower-chroma colors greatly simplifies tasks such as color-balancing and color-matching pastels and other subtle colors.

7.14. Connecting to Other Workspaces

The final consideration related to the encoding method of the Input CES is its link to other color–image workspaces, where various types of image manipulations would be performed. Depending on the data metric chosen for the Input CES, some types of image manipulations might be performed directly in the Input CES color space. As was the case for computer image generation, practitioners likely will prefer to work in various existing color spaces that are familiar and well suited for particular types of imaging operations. Spaces based on perceptual attributes such as lightness, hue, chroma, and saturation, for example, are useful for adjusting colors of individual objects. Other spaces are better suited for merging multiple images into seamless composite images.

Figure 7.14.1 below illustrates a data path that includes the use of color-space transforms from the Input CES to other color spaces, which could be selected by the user. User image manipulations then would operate on the image in that selected space. Upon completion, a modified image would be processed through an inverse transform to return it to Input CES space. From there, the image would continue as normal through the system.

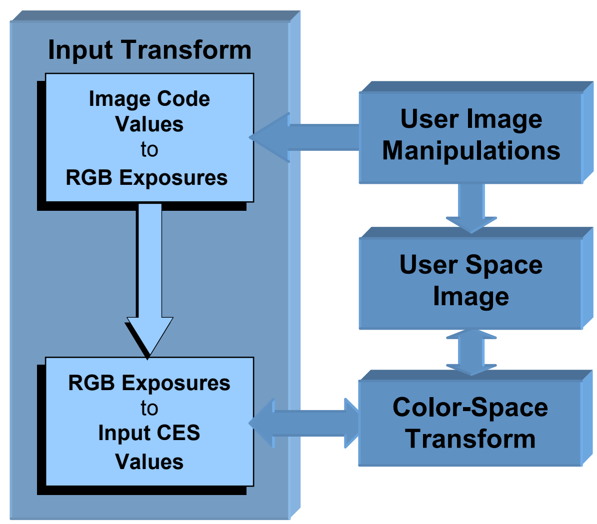

Figure 7.14.1: A data path allowing image manipulations to be applied to RGB exposure or Input CES values.

Some types of image adjustments, however, are best made before the input device/medium code values are fully transformed to Input CES values. In particular, I would suggest providing direct access to input device/medium RGB exposure values before their conversion to CIE colorimetric values. There are several reasons for this preference.

First, it is very likely there will be errors in the derived RGB exposure values. Errors can result of under- or over-exposure in the original photography, color balance errors due to imperfect white-point adjustment of an electronic camera or illuminant/film color-temperature mismatch, differences among the devices and/or media that are not fully accounted for by input calibration, and possible errors in the transformation process itself. Because these errors will have occurred in the exposure space of the particular input device or medium, they are likely to be most easily fixed in that same space, rather than later in the Input CES space. How much more difficult they would be to fix in the CES will depend on its data metric.

Assume, for example, that due to an incorrect white-point adjustment, the overall red exposure of a digital camera image is high relative to the green and blue. This is easy to identify and fix in the camera's own RGB exposure space. All that is required is a scaling (down) of all red linear-exposure values. If the uncorrected image is instead transformed to a CES based on another linear RGB space or CIE XYZ space, the high red exposure values would affect all three color channels as a result of the matrix involved in the transformation. So a simple one-channel color balance problem becomes a three-channel problem. The image still may not be too difficult to color balance, but some scaling of all three channels would be required.

If the image were converted through a nonlinear transformation to a space such as CIELAB, the high red exposure values would affect all L^*, a^*, and b^* values. If the $L^*a^*b^*$ values for the color-balanced and unbalanced images were plotted, the impression would be that something very complex has occurred. Most likely, the images could not be matched simply by making overall shifts of the $L^*a^*b^*$ values. Shifts that perfectly correct some colors may not correct others. So now a simple one-dimensional shift will have been turned into a complex three-dimensional problem that can be very difficult to resolve.

My experience suggests that it generally is best to fix problems at the time and in the same space where they occur. In this case, that means fixing device and media exposure-space errors prior to transformation to the Input CES. In Fig. 7.14.1, access to the RGB exposure values is provided by forming the input transform in two parts. A user (or algorithm) could then manipulate images at the exposure level, as shown in the figure. For this and other image manipulations, the rest of the system can be operating so that a real-time preview can be viewed as changes are made.

Adjusting exposure values prior to transformation to the Input CES values is particularly important if the transformation is to be implemented in the form of a 3D LUT. These LUTs can be quite sensitive to the correspondence of the input data to the data used in the construction of the table. Very often, shifting input data such that they no longer are perfectly "aligned" with the table can cause unexpected and problematic results. The possibility of this happening depends on the complexity of the table, which in turn will depend on the definition of the color space and data metric of the Input CES. That topic will be discussed later.

8. Rendering for Output

Input CES image values represent scene-space colorimetry. For the system to produce high-quality images, Input CES values must be rendered to output colorimetric values appropriate for projection and other forms of display.

In the subsections that follow, the function of rendering, the factors involved in the rendering process, and the implementation of rendering in the proposed Digital Cinema System are discussed.

Figure 8.1: The upper image represents scene colorimetry, the lower image represents rendered colorimetry.

8.1. Why is Rendering Needed?

Figure 8.1 above illustrates the effects of display-image rendering. The upper image represents original scene colorimetry, and the lower image represents the results of rendering that colorimetry for output.

The images demonstrate that although scene-space images may be colorimetrically accurate, when displayed directly they are perceived as "flat" and "lifeless". The fundamental reason rendering is needed, then, is to translate original-scene colorimetric values to output colorimetric values that produce images having a preferred color appearance.

8.2. Rendering and Color Appearance Models

In recent years, considerable progress has been made in the development of what are called "color appearance models". The applicability of such models to imaging applications is a subject of much debate. The underlying issues need to be discussed here, before proceeding with a description of image rendering, because one position being promoted, and which may need to be addressed by the Committee, is that the use of a color appearance model is all that is needed for determining output-image color.

Color appearance models are intended to expand the applicability of standard colorimetry. In particular, they are intended to eliminate the restriction that color stimuli being evaluated for matching must be viewed under identical conditions. A color-appearance model allows, for example, the valid comparison of stimuli pairs viewed at different levels of illumination. That and other capabilities offered by color appearance models can be useful in image applications. However, it is important to recognize that the basic objective of such models is still the same as that of standard colorimetry: to determine if color stimuli visually match. As will be discussed in the following subsection, the objective of rendering is quite different.

In addition, generating rendered images requires dealing with the practical limitations of output devices and media. For example, a color-appearance model can be used to compute the colorimetry of a projected image of a color patch required to match the appearance of an original color patch viewed outdoors. The computation would determine that, due to the substantial difference in illumination levels, color-appearance matching requires that the chroma of the projected color must be several times greater than that of the original. That is true. But what is also true is that it is not possible to produce such chroma levels with any known display technology. Moreover, images judged to have optimum color quality have chroma levels much lower than those predicted by models based on stimuli matching alone.

Although appropriate image-rendering techniques definitely do include elements of color-appearance modeling, the use of output colorimetry based solely on objective color-appearance matching produces results that are neither practical nor optimum for color imaging. That is because objective color matching is not the intent of image rendering.

8.3. Image Rendering Intent

The principal intent of image rendering is to produce displayed images judged to be optimum according to subjective—rather than objective—standards. Subjective image assessments and color-appearance matching assessments are influenced by psychophysical factors, such as image luminance level and surround. However, image assessments also involve psychological factors, such as color memory and color preference. While these factors generally are not a consideration in color-appearance matching, they must be considered in the process of image rendering.

Moreover, rendering is intended to produce images that are excellent *reproductions*, not *re-creations*, of original scenes. The assessment of reproductions is influenced by many factors other than appearance matching, including aesthetic expectations developed throughout human cultural history. Centuries of visual art, more than a century of traditional photography, and decades of electronic imaging have contributed to accepted conventions for image reproductions. For example, certain colorimetric modifications of luminance and chroma levels in a reproduction are used to *suggest* and *represent* (not duplicate) a brightly illuminated outdoor scene. This is done only by relatively subtle shifts, not by color matching.

The intent of the rendering process, then, is to produce image displays that are *consistent* with the influences, expectations, and conventions that contribute to the interpretation and assessment of all forms of image reproductions. The technical factors involved in that process are discussed next.

8.4. Image-Rendering Factors

If one were to make a list of all the factors that might be considered in determining an optimum relationship of reproduced colors to scene-space colors, the task of developing a rendering transformation would seem hopeless. For example, there are dozens of known psychophysical effects that, if considered, would greatly complicate the relationship.

My experience, however, is that under the conditions relevant to typical imaging applications, many psychophysical effects are not significant. Thus only a manageably small number of factors need to be considered. I would suggest, then, that practical rendering transforms can be built for digital cinema if the key factors discussed below are accounted for appropriately. (Techniques for doing that are described in some detail in my textbook.)

- **Viewing Flare**. Flare light in the viewing environment physically lowers the luminance contrast of the display, especially in shadow areas, and it also desaturates colors. To compensate, rendering must include appropriate adjustments of the image grayscale and chroma levels.
- **Image Luminance**: In nearly all situations, the (absolute) luminance of the displayed image will be considerably lower than that of the original. This lowers the perceived luminance contrast and colorfulness of the display. To compensate, rendering again must include adjustments of the image grayscale and chroma levels.
- **Observer Chromatic Adaptation**: The perception of color is strongly affected by the observer's state of chromatic adaptation. Rendering must modify the chromaticities of scene-space colorimetry to be

consistent with the observer's state of chromatic adaptation in the display viewing environment.

- **Lateral-Brightness Adaptation**: The perception of image luminance contrast is affected by the relative luminance of the area surrounding a displayed image. This is of particular importance to digital cinema because the dark surround of theater projection significantly lowers perceived image luminance contrast. To compensate, rendering must include appropriate adjustments of the image grayscale.
- **General-Brightness Adaptation**: The perception of image brightness is also affected by the relative luminance of areas surrounding the displayed image. Again, this is of particular importance to digital cinema due to its darkened viewing environment. A well-designed rendering would take advantage of this phenomenon by altering the grayscale characteristic to effectively increase the dynamic range of the highlight region. This can create the illusion in displayed luminance levels above those of a perfect white.
- **Local-Brightness Adaptation**: In live scenes, an observer can sequentially focus on and locally adapt to various regions within a scene. This allows details to be seen in areas of deep shadows and areas of bright highlights. Little or no such adaptation takes place in the viewing of reproductions. The effect is emulated in paintings by local adjustments of tone reproduction. The effect can be simulated, at least to some extent, in rendering by global adjustments of the grayscale in the highlight and shadow regions.
- **Color Memory and Color Preference**: Optimum image rendering should account for the psychological influences of color memory and color preference. Again, the goal is to produce reproductions that are judged according to subjective standards. In many cases, that involves further changes in color to produce displays that, although not accurate, conform to an observer's recollections and preferences for color.
- **Output Luminance Dynamic Range and Color Gamut**: The final consideration is that the rendering process must specify colors that are within the luminance and color-gamut limits of the actual output device or medium. Typically, ideal colorimetric values would first be determined, based on scene-space colors modified according to the factors described above. Appropriate gamut mapping would then be applied to transform those colors as needed for real outputs.

8.5. Image-Rendering Effects

Figure 8.5.1 illustrates the effects of rendering a scene-space grayscale (the line of unity slope in the figure) for three different types of output. All the output grayscales have some degree of nonlinearity (in log space). This results from the compensation for viewing flare. In each case, the output grayscale has a slope of at least 1.15. This minimum slope increase is one result of the compensation for a lower image-viewing luminance level. The slide-film grayscale has a further increase in slope to compensate for lateral-brightness adaptation effects in dark-surround projection. The video system curve is for normal-surround viewing. Its slope also would need to be higher if the intended output were a dark-surround display, such as a home theater. As discussed earlier and shown in Fig. 1.6.2, correcting the visual density values for local-brightness adaptation yields negative visual-density values in the highlight regions of the video-system and slide-film grayscales. This corresponds to CIE Y and L^* values greater than 100, which correctly describes the appearance of highlights displayed by these media. The classic S-shape in all the grayscales results from the compression of highlight and shadow information in rendering to simulate local-brightness adaptation. Where the compression begins and ends is a function of the available dynamic range of each medium.

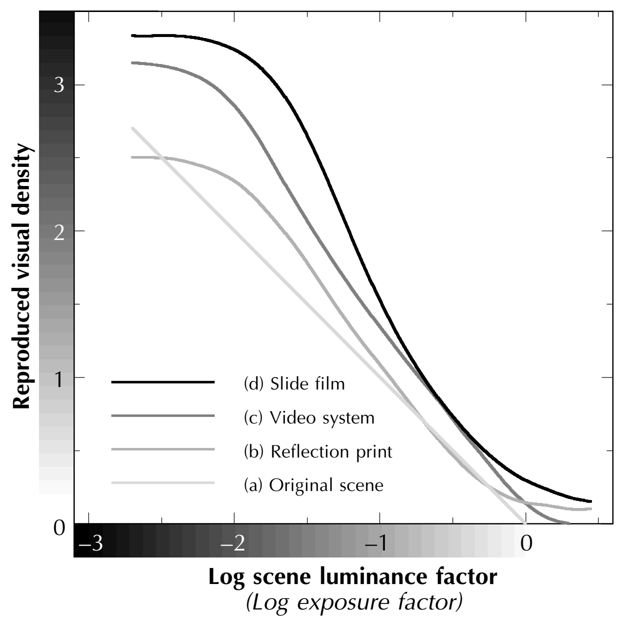

Figure 8.5.1: Grayscales rendered for three types of output.

It is very important to note that the slope increases and flare compensation necessary for output significantly increase the dynamic range required for output. As a result, the original-scene dynamic range that can be displayed will always be limited by the output. The process of rendering therefore produces a loss of information originally captured at input. That is why I continue to stress that all captured scene information should be retained until all adjustments, editing, and other image manipulations are complete. Only then should rendering be applied. Again, it is for that reason that I have recommended against the use of a single CES based on any form of rendered image.

8.6. Rendering Implementation

From the preceding discussion, it is clear that not just one but *multiple* output renderings ultimately would be needed for the proposed system to support various forms of output. This is necessary because the viewing conditions for outputs such as theater projection, home video, and reflection prints (for posters, magazine advertisements, etc.) are significantly different. The imaging capabilities and limitations of the output devices and media involved also are very different.

This might imply that the proposed architecture, which includes both an Input CES and an Output CES, is unnecessary. If the rendering from scene space to each different type of output must be unique, why not just use only the Input CES alone in the classic color-management system architecture (Fig. 2.1) discussed earlier? If that were done, the digital cinema system architecture would be that shown in Fig. 8.6.1 below. The reasons why this approach is *not* recommended, and why the alternative dual-CES approach instead is proposed, are discussed in Section 9 that follows.

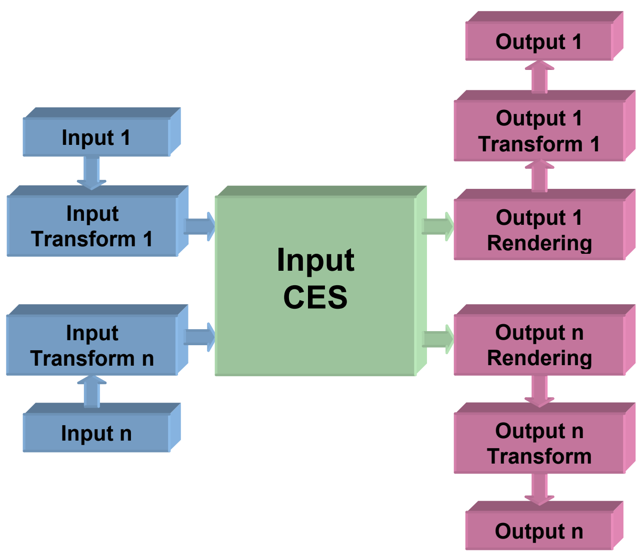

Figure 8.6.1: A digital cinema system based on a classic color-management architeture and the Input CES alone. Note: This approach is NOT recommended!

9. Reference Rendering

The recommended output architecture for the digital cinema system is shown in Fig. 9.1 (next page). The architecture includes the step of Reference Rendering, which transforms Input CES scene-space colorimetry to Output CES colorimetry for a defined reference output device or medium. As the figure illustrates, the signal processing for each basic type of output includes a further transformation from the Reference Rendering to a rendering specific for that output type.

The mathematical results from this process are not different from those of the system in Fig. 8.6.1. The net effect of the signal processing from Input CES values to output-device code values will be the same whether the processing is done by a single rendering or by a process that first renders for a reference output and renders further for a particular output.

If the output results will be no different, why bother with an intermediate rendering to some (possibly hypothetical) reference device? The reason for including the step of reference rendering is this: It provides an Output CES that is *complementary* to the scene-space encoding used for the Input CES.

As I have stated in other publications, to be of value, a CES must satisfy two criteria: First, it must be *unambiguous*: The meaning of a set of CES values must be unique and fully defined by the values alone. If other information is necessary in order to interpret the meaning of CES values, the specification is not unambiguous. Second, the CES must be *unrestricted*: System devices, media, and functions must not be constrained by limitations of the CES itself.

These two criteria are somewhat conflicting. It is easy to design an encoding that is just one or the other. For example, one could simply dictate that all images must be encoded in terms of code values for a defined CRT-based monitor in a particular location. While that certainly is unambiguous, it is also restricted in that the encoding capability is subject to the limitations of the monitor. It is also easy to design an encoding that is unrestricted, such as by simply specifying that colors will be represented in terms of standard CIE colorimetric values. That would be ambiguous, however, because a given colorimetric specification can be associated with almost any color appearance.

In the proposed system, the use of two Color Encoding Specifications allows these conflicting attributes to be balanced appropriately, and somewhat differently, for input and output. The Input CES is unrestricted in that it supports input from all devices and media, and it does so in a way that retains the full capabilities of each input. It is also unambiguous in that it specifically defines colors in a real or virtual original scene. However, that is an unambiguous specification of what a color *is*; it is not necessarily an unambiguous specification of the artistic intent of how the color should *look*. The function of the Reference Rendering, then, is to provide a complementary means for specifying how the color ideally would look if it were rendered with little or no restriction.

It is useful to recall the debates that occurred when the technology for "colorizing" black-and-white motion pictures was introduced. The arguments centered on the concept of original intent. Would the director have used color if it had been available at the time? If so, what color "look" would have been used? Strong, vibrant colors; or muted, subdued colors? A warm, neutral, or cold overall look? Those arguments could not be resolved because there was no means to convey the original artistic intent. In the proposed system, Reference Rendering provides that means.

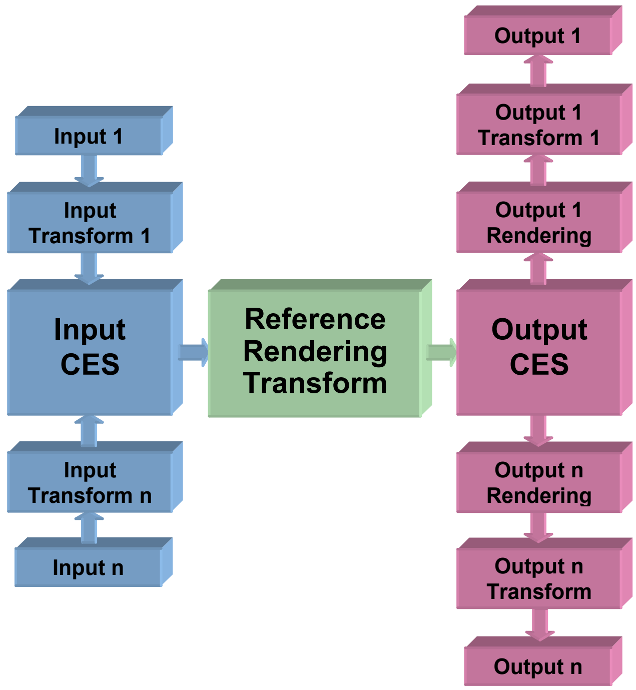

Figure 9.1: The recommended system architecture includes a Reference Rendering process to transform Input CES values to reference Output CES values.

This discussion has emphasized the importance of recognizing that in a digital cinema system, color representation requirements change as images move from origination through to production, distribution, and display. To address the different requirements involved in this process, there are three stages of color representation on the proposed system:

- In the Input CES, color is represented in the context of real or imaginary scenes.
- In the Output CES, color is represented in the context of ideal renderings of those scenes.
- After each Output Transform, color is represented as true to the Reference Rendering as possible, within the capabilities of the particular output.

9.1. Reference Rendering Approaches

There are three fundamentally different approaches that could be taken in the development and application of the Reference Rendering transform. I would suggest that a clear decision should be made, and that the decision should be explicitly defined. If not, history suggests that numerous rendering approaches and implementations will evolve, and the meaning and usefulness of Reference Rendering will be lost.

In one approach, multiple Reference Renderings could be used as a means for adding various "looks" to Input CES images. I do not support this approach, and I bring it up only because it is likely to be proposed or simply used by default if an alternative is not clearly specified. The idea of adding or altering the look of images certainly is appealing. However, having "multiple references" is essentially the same as having "no reference". I would propose, then, that alterations to image looks should not be made in the Reference Rendering but made instead in the Input CES. As discussed earlier, the system supports translation to Input CES values to and from other workspaces, where such alterations can be performed. Use of a Reference Rendering transform and a high quality display can provide a real-time preview of the altered Input CES images. Any look, even the look of a particular negative film and print film, can be emulated this way, without losing acquired information and without creating a confusing array of "Reference" Renderings.

A second approach would be to specify a single Reference Rendering transform that renders Input CES values and also imparts a certain look, such as that of an existing film, electronic, or hybrid system. This is a much better alternative in that it is based on a single transform. Use of a single rendering removes any ambiguity regarding the intended look of Input CES images. However, I am not certain that any particular look can be agreed upon within the industry, and I am also not sure doing so is best for the long term.

I would suggest, then, use of a single Reference Rendering, and one that is as "neutral" as possible. Its function would be to modify Input CES colorimetric values only as necessary to generate Output CES colorimetric values for an image that, when displayed, is true to the Input CES color. The role of the transform, then, becomes one of *delivering* color in the intended viewing environment, not one of *creating* new color. As stated earlier, creative color intent is best represented in Input CES values. The combination of an Input CES image file and a Reference Rendering transform unambiguously communicates that intent.

9.2. Reference Rendering Transform Development

Development of an appropriate Reference Rendering transform will require experimentation, but it should not be difficult. I would suggest basing the transform on a defined Reference Display Device. For simplicity, the device should be an additive color projector, which can be specified in terms of its electro-optic transfer function, its RGB primaries, and its white point. The following six-step approach can be used to develop and evaluate that specification:

1) The viewing conditions associated with images from the Reference Display Device must be specified. The specification would include factors such as viewing flare, image luminance level, and surround type that influence the appearance of the displayed images. These conditions are defined in the Output CES, which will be discussed later.

2) A luminance dynamic range must be specified for the Reference Display Device. This can be based on existing devices or media, with perhaps some speculation on possible future improvements.

3) A grayscale curve, consistent with the specified luminance dynamic range can then be drawn based on the flare and viewing-condition factors discussed previously. The system grayscales of current film and electronic systems can provide useful guidance. I would suggest using aspects of both types of systems.

Grayscales for electronic systems generally are less compressed in the highlight region, and that would be a useful feature in the reference grayscale. When adjusted for brightness adaptation, the grayscale should extend to an $L*$ of about 130. This would provide good retention of highlight detail. A motion picture film system grayscale, which has a greater overall luminance dynamic range, may provide a better model for the shadow regions.

4) It would be very useful at that point to create monochrome (black-and-white) images based on the proposed grayscale. This would allow images to be evaluated for tone reproduction alone, without the complications of color. A matrix can be used to create monochrome image files from color images in the Input CES. Following the matrix, all three channels would be identical and equivalent to CIE Y values.

5) Color images would then be produced and evaluated. Again, the objective here is not to create new color, but only to render Input CES colors such that they will appear appropriate in the reference viewing conditions associated with the Output CES and the Reference Rendering.

A good starting point in developing this process would be to first transform images from the Input CES primaries to the primaries defined for the Extended Reference Input Medium Metric RGB (ERIMM RGB) color space (Fig. 9.2.1a). The transformed values can then be mapped through the Reference Rendering grayscale curve to produce modified ERIMM RGB values that then can be converted to CIE XYZ or other values according to the data metric of the Output CES. This process effectively translates scene-space colorimetry to rendered-space colorimetry.

Although the Reference Rendering grayscale could be applied to image color values expressed in other primaries, the results may not be satisfactory. The ERIMM RGB primaries were selected specifically for the application of nonlinear operations, and for grayscale mapping in particular. Their use minimizes hue rotations, especially those of more sensitive colors, that can be caused by nonlinear transformations. One change in these primaries should be considered: When they were derived, there was a Photoshop-related constraint that no chromaticity coordinate value could be zero or negative. That constraint no longer exists. A blue of about $x = 0.1$, $y = -0.1$ should be investigated.

6) Depending on the design of the grayscale, the above processing may cause too little or (more likely) too great an increase in rendered-image chroma levels. If so, a simple matrix operation can be included in the process to adjust the chroma levels as needed.

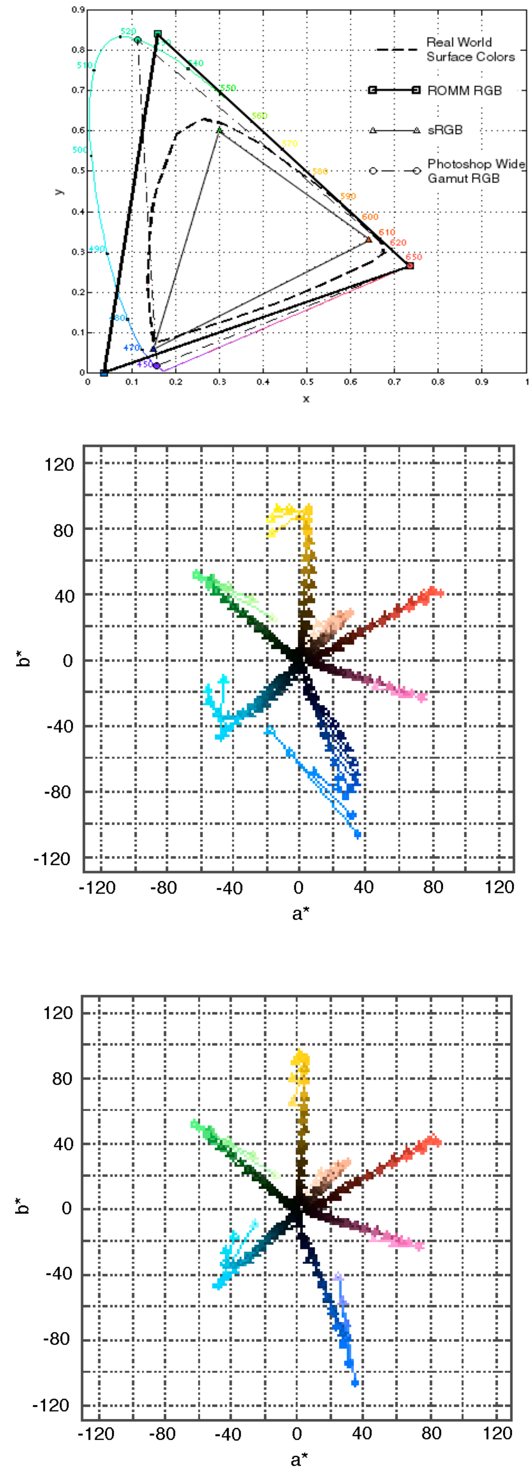

Figure 9.2.1: a) Chromaticity coordinates of ERIMM RGB and two other set of primaries; b) hue shifts (from shadow-to-highlight) resulting from the application of a nonlinear transformation in Photoshop Wide Gamut RGB space; c) reduced hue shifts resulting from the application of the same nonlinear transformation in ERIMM RGB space.

10. Output from the Output CES

In the preceding section, a process of Reference Rendering was used to transform Input CES scene-space colorimetry to Output CES colorimetry for a defined reference output. In the final stages of the system, Output CES are processed through a series of transformations that ultimately produce digital code values appropriate for each specific output device. The sequence, shown in Fig. 10.1 below, is discussed in the following subsections.

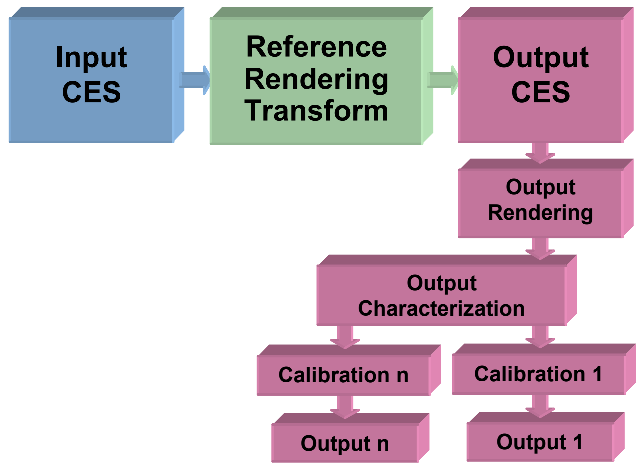

Figure 10.1: The sequence of output signal processing from Output CES values to output device code values.

10.1. Output Rendering

As Fig. 10.1 above illustrates, output signal processing begins with a second rendering (or what might be called "re-rendering") operation. This transformation adjusts Reference Rendered Output CES values to rendered values appropriate for a particular type of output. Once again, the intent is not to create anything new; it is to deliver the color specified in the Output CES as faithfully as possible, within the limits of the given output. Of course different re-renderings would be required for outputs having different luminance dynamic ranges, color gamuts, and/or viewing environments.

The degree of re-rendering performed in the Output Rendering operation will depend on how the characteristics of the actual output device and viewing environment correspond to those of the reference. In retrospect, then, it becomes clear that the specified Reference Rendering characteristics should be fairly realistic in order to minimize the complexity of the Output Rendering transformations. However, this needs to be balanced with the Output CES objective of retaining information that someday might be used by future types of display devices.

Figure 10.1.1 is an example illustration of the relationship between a reference grayscale and a grayscale of a real system on which the output rendering will be based. The example is for a reflection print system, rather than a motion picture system, but it can be used to discuss the basic concepts involved. The relationship of the hypothetical grayscale shown in light gray to the actual grayscale corresponds most closely to the relationship proposed here between the Reference Rendering grayscale and that of an actual projection output device. Note that the curves essentially are the same over much of their range, but there is significantly more gradient in the highlight and shadow regions of the hypothetical grayscale. Basing the Reference Rendering and the Output CES encoding on that type of grayscale makes possible the encoding of significantly more highlight and shadow information in the Output CES.

The process of Output Rendering generally is straightforward. In most cases, it involves nothing more than a mapping from the grayscale of the Reference Rendering to that of the actual output. This is a simple one-dimensional transformation, performed on each of the color channels, and it can be executed in practice by a set of 1D LUTs.

The arrows in Fig. 10.1.1 indicate the mapping of two example data points from their values on the reference curve to their corresponding values on the actual system curve. The figure helps illustrate once again the basic strategy of this proposal: In this system, each stage from input to final output is capable of encoding image information sufficient for all following stages. The Input CES encodes the greatest range, because that is the origination space of the system and where images may need to be adjusted and manipulated. When that process is complete, less information is required to produce rendered images; but enough information must be retained for rendering to any current or anticipated future output. That is the function of the Output CES. Only when a particular output has been identified is the information re-rendered down to the actual limits of that output.

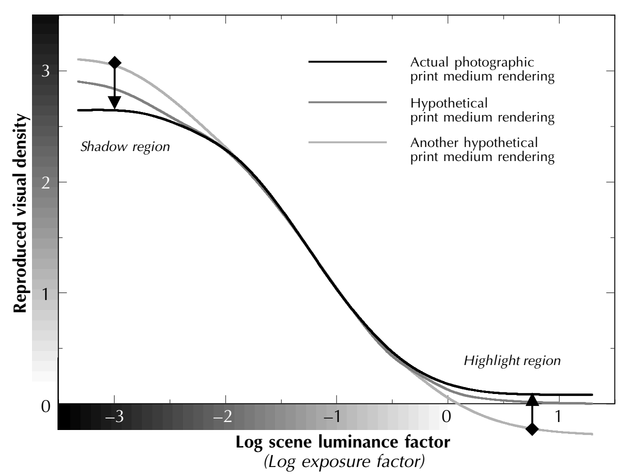

Figure 10.1.1: Grayscales for a reflection-print imaging system and two hypothetical reference systems.

The preceding discussion emphasizes why the proposed Reference Rendering curve is not an actual product grayscale. In particular, it is not the grayscale of a product, such as a reflection print, having limited luminance dynamic range. Fig. 10.1.1 clearly shows, especially in the highlight region, that once information has been rendered to a grayscale region having little or no gradient, that information becomes irrecoverable. The transfer of information from input to output is an inherently "downhill-flowing" process. It is important, then, that the flow is controlled such that sufficient information is retained at each stage to fully support the requirements of all subsequent stages.

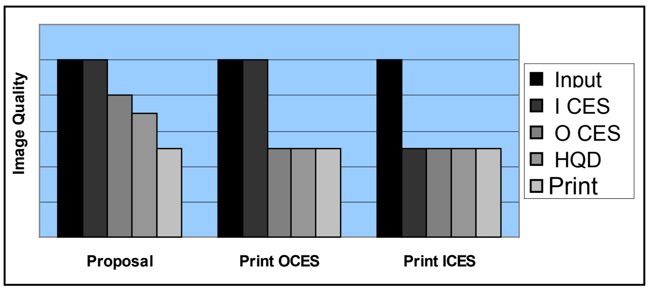

Figure 10.1.2: Comparison of information retention in the proposed system vs. two reflection print-based alternatives.

Figure 10.1.2 above is a conceptual illustration of available image information, and thus potential image quality, at various stages in a color-managed system. In the proposed system (left block in the figure), all original input information is retained in the Input CES, and it is gradually reduced as it is Reference Rendered to the Output CES and then re-rendered to a high quality display device or a reflection print. If the Reference Rendering were based on the limited grayscale of a reflection print system (center), the Output CES would no longer be capable of producing high quality images on the high quality display device. The same would be true if the Input CES (right), or the single CES of a classic color management architecture, were based on the rendered grayscale of a reflection print having limited dynamic range.

For the grayscale mapping from reference to a given output to be correct, the respective grayscales must be aligned in two ways. First, they must be "speed balanced", i.e., they must be aligned properly along the exposure axis. This is done by aligning the curves based on the exposures of known reference test patches. In Fig. 10.1.3, the grayscale curves of four systems have been aligned along the relative log exposure axis, based a knowledge of the aim visual density each should produce from a normally exposed neutral test patch of 20% reflectance (corresponding to –0.7 log luminance factor, the dotted line).

Second, the grayscales must be aligned along the visual-density axis, such that they are matched for perceived brightness. Some publications have stated that this can be accomplished simply by matching the curves at their point of maximum brightness (minimum

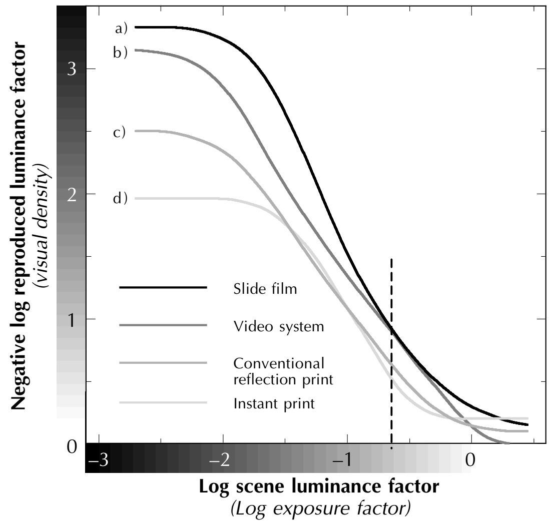

Figure 10.1.3: Four grayscales aligned for relative log exposure (unadjusted for general-brightness adaptation).

visual density). The idea is based on the mistaken assumption that the brightest areas within the visual field determine the observer's state of general brightness adaptation. Numerous psychophysical experiments—and decades of practical imaging experience—have proven that assumption to be entirely untrue. If that concept, which has been referred to as "relative colorimetry" or "media-relative colorimetry", is applied to the four grayscales being discussed, images based on the resulting adjustments actually will be more poorly matched for brightness than images based on the unadjusted grayscale curves.

In Fig. 10.1.4 below, the four (exposure matched) system grayscales have been properly aligned for general brightness adaptation. In a linear space, this would be done by applying appropriate scaling factors, which is equivalent to shifting the curves along the

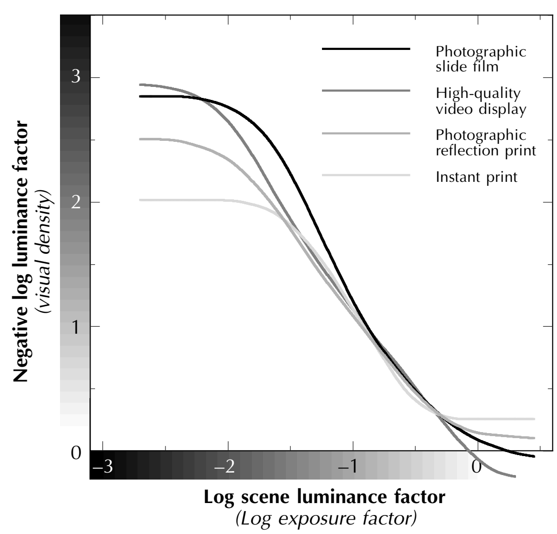

Figure 10.1.4: Four grayscales aligned for relative log exposure and adjusted for general-brightness adaptation.

visual-density axis used in the figures being discussed. This alignment can be determined experimentally by visually matching images for brightness. The results shown in Fig. 10.1.4, for example, were determined by generating monochrome images for each of the grayscales. The images were displayed *simultaneously* on a high quality monitor, and software was used to adjust the images until they all appeared to match for overall brightness. Note that because the grayscales are very different, the resulting images did not look identical. That was not the intent. The purpose of the experiment was to match the images as closely as possible using brightness as the only variable.

The results shown in Fig. 10.1.4 make intuitive sense. Note that the curves are closely aligned in the region corresponding to exposures from 20-40% reflectors (about -0.7 to –0.4 log exposure factor). That is in agreement with the results of psychophysical experiments and is reasonable given that objects having such reflectances are visually predominant in most scenes. Also notice that the minimum visual densities of the adjusted curves are very different. That agrees entirely with the visual impression of the images. For example, when matched for overall brightness to the other images, the video grayscale produces images having much brighter and more detailed highlights.

With some experience, it usually is possible to align grayscales for equivalent brightness without performing any elaborate experiments. In practice, any alignment errors will show up soon enough once images are produced. If everything seems to be working satisfactorily except that images from one type of output are consistently too dark or too light, it is a simple matter to adjust the associated grayscale accordingly. Again, the easiest way to isolate such problems is with monochrome images.

Once the proper exposure and brightness relationship has been established between the Reference Rendering grayscale and a given output grayscale, the mapping table to transform one to the other can be developed. I would suggest that the same color-space primaries used in the Reference Rendering process also be used for this output re-rendering process. A note of caution: Grayscale transformations should *not* be done by using a YCC space and transforming only the Y channel. Treating luminance and chrominance information so differently produces unrealistic images of very poor quality. The process should be performed in an RGB space in which visual neutrals are normalized to have equal RGB values. When this is done, an identical LUT can be used in all three color channels.

The last issue related to rendering for output is color gamut. To function as a true reference, the Reference Rendering and its associated Output CES must have a gamut that supports all current and anticipated forms of output. It would be expected, then, that some form of gamut mapping will be required in going from the Output CES to any real output device or medium. I would suggest, however, that if the grayscale mapping just discussed is performed in ERIMM RGB or another space based on similar primaries, that process alone may also provide color gamut mapping that is entirely satisfactory. I cannot guarantee that will happen in every situation, but my experience it that it has never failed to produce results that are robust and visually pleasing. I would certainly suggest that the results be evaluated before concluding that some type of complex gamut mapping is required.

I would also suggest that if it is determined that some additional gamut mapping is required, it should be applied after the grayscale re-rendering operation. This would retain the intent of the output process. That process is executed in two steps—rendering “up” from the Input CES to the Reference Output CES and back “down” to the actual output; but it should be identical to rendering directly from the input to the output. Interposing gamut mapping within the two-step process is likely to cause unwanted color distortions.

One final comment regarding gamut mapping: The Committee may hear it stated that in order to develop a gamut mapping table, the gamut limits of both the input and the output must be specified and restricted. If this were the case, it would lead to the conclusion that a device or medium defined as the basis for reference rendering also must be specified to have a fully defined and restricted color gamut. This position is commonly held, but it is not correct.

A point to consider is that if it were true that the gamuts of both the input and the output have to be known, it would be impossible to build a photographic film or any other type of imaging system that records original scenes. The color gamuts and dynamic ranges of real-world scenes are widely variable and essentially unlimited. Yet a conventional slide film, with no “knowledge” of the gamut limits of any given scene, can still be used to take pictures of whatever a photographer chooses to shoot. If that is possible, why then would it not be possible to develop a gamut-mapping strategy for a film writer, using the very same slide film, unless the gamut of the input images is both known and restricted?

The fact is that practitioners with appropriate skill and experience routinely construct gamut-mapping transforms based only on a knowledge of the limits of the specific output involved. Knowledge of the input limits sometimes can be useful if it is available, but it is not necessary. The transforms are built in such a way that they can handle the entire color space of image values supported in the CES that provides input to the transform. In this proposal, then, there is no need to specify or restrict the gamut boundaries of the Reference Rendering device, nor is it necessary to limit the Output CES based on concerns regarding output gamut mapping.

10.2. Output Characterization

To summarize the processing to this point: Input CES values have been transformed through the Reference Rendering Transform to Output CES values. The Output CES values then were re-rendered and gamut mapped, if necessary, in a device-specific Output Rendering transformation to form aim colorimetric values for the specified output device/medium and viewing conditions. A complete system would include multiple Output Rendering transforms, each associated with a particular type of output device or medium.

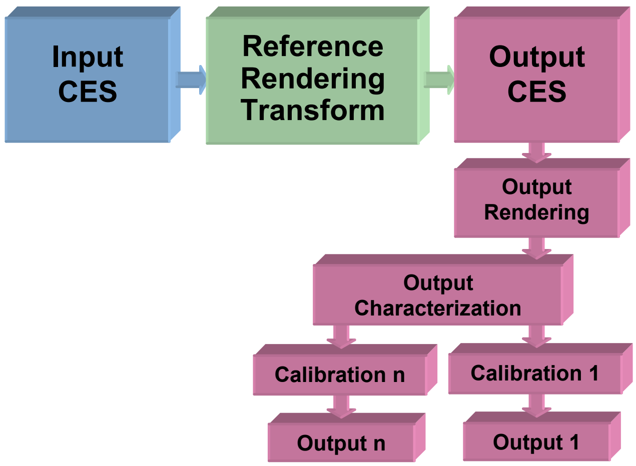

Figure 10.2.1: Output Characterization follows Output Rendering in the sequence of output signal processing from Output CES values to output device code values.

As shown in Fig. 10.2.1 above, the next step in the output signal processing is Output Characterization. Output characterization is a procedure for defining the colorimetric characteristics of a single device that is representative of a group of actual devices used for output on the imaging system. A colorimetric output-characterization transform can be developed from empirical data obtained by measuring the colorimetric values of an appropriate number of color patches produced on the representative output device from an array of output-device code values (Fig 10.2.2). The output characterization process essentially mirrors that of input characterization. However, it is somewhat less problematic because the input consists of digital image values rather than physical test targets.

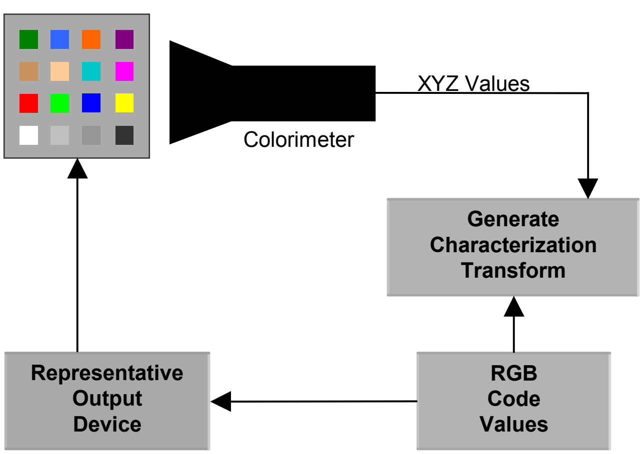

Figure 10.2.2: Construction of an output characterization transformation from an array of RGB code values.

For the most part, output characterization is straightforward. Nevertheless, there can be problems. For example, some computer monitors essentially cannot be characterized because the relationships between code values and light output are not predictable. With CRT-based monitors, this may result from an insufficient power supply. The relationship between a given set of code values and light output becomes dependent on the average signal level of the displayed image. Displays being considered for image preview and other critical work should be screened for this problem (referred to as a *clamping failure*). My experience is that failures are inherent in model designs and not in individual units; hence, testing a single representative unit generally is sufficient. I have found no correlation to price or manufacturer. each different model is suspect until proven otherwise.

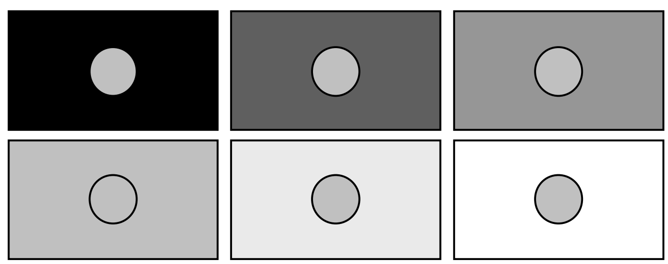

Figure 10.2.3: Measurements of the test areas of images such as these can detect display device problems.

One simple test consists of sequentially displaying a series of images, like those shown in Fig. 10.2 above, and measuring the center test area. The code values of that area are identical in each image and are chosen to produce a mid-value gray. The surrounding area differs in each image in the series, ranging from full black to full white. The center area should, of course, measure identically for all images. On a good monitor or other type of display, that indeed will be the case. This indicates there is a consistent and predictable relationship between code values and light output, which is an essential criterion for successful characterization. On some displays, however, the center intensity can be affected by the average image signal level, sometimes by as much as a factor of four! When this happens on CRT displays, the center area darkens as the surround intensity increases, due to increased draw on the power supply. Obviously, such a device cannot be characterized.

There may be other problems as well, including variations in display uniformity, lack of color purity, sensitivity to image orientation, internal flare, etc., that can complicate or prevent the characterization of an output writer or display device. Such problems may not be evident in a routine characterization. It is good practice to test the robustness of a characterization developed for any type of output device with an independent set of test images in which different test colors, different geometric layouts, and different backgrounds are used. Often a simple test can reveal problems that might long go undetected with normal pictorial images. For example, an image of opposing but otherwise identical grayscales, like those shown in Fig. 10.2.4 below, often exposes problems in scanning-type output devices where the output level in a given area can be affected by the levels of preceding areas.

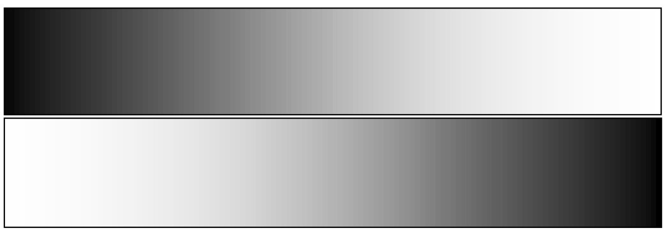

Figure 10.2.4: Measurements of a displayed image of opposing grayscles can detect display device problems.

Some outputs (especially additive-color devices) can be easily modeled; so a mathematical model can serve as the representative device for characterization. This has many obvious advantages. Needless to say, any such model should be verified experimentally to confirm its ability to predict the colorimetric characteristics of an actual device.

10.3. Output Calibration

In some situations, the use of output characterization transforms alone is feasible and practical. For example, it might be reasonable to build characterization transforms for each output device in an operation where there are relatively few devices, where the devices tend to be stable, and where the procedure for building the transforms is fast and economical. In most circumstances, however, it is far more practical to build a single characterization transform representative of all devices of the same type, as described above, and to then provide unique calibration transforms for each individual device. This arrangement was shown previously in Fig. 10.2.1.

A calibration transform corrects any deviation of a particular device from the representative device on which the characterization transform was based. This combined characterization/calibration approach has a number of advantages. In particular, since the bulk of the transformation is performed in the characterization transform, calibration transforms generally can be quite simple. In most cases, they can be derived using a relatively small set of test colors. For most three-color systems, calibration of the grayscale characteristic alone is sufficient. For most four-color systems, calibration can be based simply on four individual color scales. This generally makes the calibration procedure very fast and inexpensive to perform.

Other important advantages of using a combination of calibration and characterization, rather than characterization alone, involve issues of expense, efficiency, workflow and image distribution. Building and testing a characterization transform can be expensive and time consuming. In some situations, devices can drift faster than new characterization transforms can be built. Moreover, many devices have provisions for calibration but not characterization. For example, many printers have raster image processors (RIPS) that include a set of ID LUTs, which makes them well suited for calibration. Having calibration resident on each individual output device entirely changes the workflow and distribution of images.

Consider one real-life example: Several years ago, I was part of a small team sent to install a color management system at a large manufacturing plant that had more the 150 workstations and 50 output devices. All the workstations were the same model, as were the outputs. The intended color-management software was based entirely on characterization and included no provision for separate calibration. You can imagine the customer's reaction to being told that 200 transforms, each costing $800, would be required. Moreover, since the outputs were electrophotographic printers, which are notorious for drifting, each would have to be re-characterized approximately 3 times per day! Worse yet, it took about 8 hours to build each characterization transform. Since the plant ran three shifts per day, it would work out "perfectly" that all transforms would be obsolete just as soon as they were completed. In addition, every workstation would have to be equipped with all 50 printer-characterization tables, and the operators would have to know which printer was going to be used for a particular job so that the appropriate characterization transform could be applied. Obviously, an alternative solution was needed, and calibration provided that alternative.

The team developed simple calibration procedures for the monitors and printers. This was easy because each monitor could be calibrated electronically, each printer had a RIP with programmable 1D LUTs, and the customer already had a programmable densitometer on which calibration scales could be read and uploaded quickly and easily. This approach changed everything. Only two characterization transforms, instead of two *hundred*, were required, and they never had to be changed. With calibration physically installed on each printer, it no longer mattered which printer was used on any particular job. From the perspective of the workstation, all printers became identical.

The example illustrates several basic concepts related to device calibration that have implications important in digital cinema applications. The following benefits, which apply to input as well as output, would be realized if calibration were incorporated on each device in the digital cinema system:

- Relatively few characterization transforms would be required in the system. This would allow each to be built carefully and tested thoroughly before use. This in turn would enhance the system's reliability and stability.
- Because calibrations can be performed quickly and easily, it would be practical to calibrate at whatever frequency is necessary to maintain device consistency. This, too, would enhance the system's reliability and stability.
- Transforms for individual devices would not be needed at the workstation level, which would simplify the color-management software.
- Information regarding which specific input device supplied image data or which specific output device will be receiving the data would not be needed at the workstation level. This would greatly simplify the color-management software and its use by an operator.

11. Digital Color Encoding

Attention to this point has focused on overall system design and architecture. The remaining major topic of this paper—the attributes of the digital color encoding proposed for representing color throughout the system—are discussed in this section.

11.1. Color Encoding Method

The method of encoding color in the proposed system is described in detail in my textbook, where it is the basis for a universal color-managed imaging system. Its use is appropriate for digital cinema because the overall system requirements, including support of disparate inputs and outputs, are fundamentally the same as those of a universal system.

The encoding method is based on the concept that color can be represented and communicated in a way that is unambiguous and unrestricted if its description includes *both* of the following:

1) A fully defined colorimetric specification.
2) A defined set of reference viewing conditions.

The reason for the colorimetric specification is obvious. The reason for specifying reference viewing conditions is that doing so eliminates the ambiguity inherent in the colorimetric specification alone. As previously discussed, a given colorimetrically-specified stimulus can have very different color appearances depending on how it is viewed.

Although many factors could be considered in these reference specifications, my experience is that, for imaging applications, it is only necessary to specify the six described in the following subsections.

11.1.1. Colorimetric Specification

Image colorimetry can be expressed in various ways, using various metrics. The principal specification here is that the colorimetry corresponds to flareless image measurements. This might sound confusing, but the previous statement does not mean there is no flare in the images themselves. An image could be that of an original scene taken such that the image was flooded with lens flare. That is not an issue. The specification applies only to the measurement of the image, not to its content. Flareless measurements are specified because they provide the greatest dynamic range and the least compression of shadow-region image data. Therefore it is better to compute the effects of more flare, when necessary, rather than attempt to go the other way.

11.1.2. Viewing Flare

Flare light in the viewing environment will physically alter the displayed color stimuli of an image. Specifications of the relative amount and chromaticity of flare light, together with the flareless colorimetric measurement of an image, as described above, provide a complete colorimetric description of that image.

11.1.3. Luminance Level

As discussed previously, an observer's perception of image luminance contrast and colorfulness are influenced by the (absolute) luminance level at which the image is displayed. It generally is not necessary to know the level exactly, and specifying a general range is usually adequate. The principal distinction would be between outdoor (daylight) levels and those typical of display environments.

11.1.4. Chromatic Adaptation

The perception of a color stimulus is strongly affected by the observer's state of chromatic adaptation. A meaningful interpretation of a colorimetric specification therefore requires an accurate specification of that state. This is done by specifying the *chromaticity* of a stimulus that would be perceived to be perfectly achromatic by an observer in the specified viewing environment.

I emphasize the word "chromaticity" here because this concept is widely misunderstood. It is often stated that for this approach to be unambiguous, the full *spectral power distribution* of a reference viewing *illuminant* must be specified. That is incorrect on two counts. First, observers may or may not be adapted to the chromaticity of the viewing illuminant. In imaging applications, due to factors such as mixed illumination, partial chromatic adaptation, and media color balance it is likely that observers are adapted to a chromaticity different from that of the viewing illuminant. Moreover, the *intended* colorimetric contribution of the viewing illuminant is already included in the colorimetric specification. So, for example, if an image is deliberately illuminated with purple light to produce a particular effect, the influence of that light is represented in the image's colorimetry. The function of the reference viewing specifications is not to attempt to override that intended appearance. It is just the opposite: The specifications contained in the reference viewing environment are meant to provide a means of preserving and conveying an intended appearance.

It is important for the Committee to be aware of this misunderstanding because it is indicative of the broader lack of understanding of the true function of the reference viewing conditions in this method of color encoding. Contrary to an interpretation often presented, the reference conditions of this method are *not* “A set of viewing conditions under which all your images must be measured and viewed”. What they *are* is “A set of conditions under which an observer, viewing an image having the colorimetry specified, would see colors as you intended”.

11.1.5. Lateral Brightness Adaptation

Although its influence on pictorial images is significantly less than what is described in most of the literature, the relative luminance of the area surrounding an image does influence perceived image luminance contrast enough to warrant its being specified. It is usually adequate to classify the surround as normal/average, dim, or dark.

11.1.6. General Brightness Adaptation

This topic was discussed earlier, and the process of adjusting system grayscales for brightness adaptation was described. The reference specification is simply an assigned luminance factor, or percent luminance factor, for a stimulus that would be perceived to be a perfect white reflector in the context of an image. This value, together with the specified chromaticity for chromatic adaptation, define the *observer adaptive white*. The encoding method can represent highlights above the luminance level of this reference white.

11.2. CES Reference Viewing Conditions

The viewing-environment factors described above must be specified for the Input CES, for the Output CES, and for the actual viewing environments associated with each input and output. When actual input or output viewing conditions differ from those of their respective CES, image colorimetric values must be transformed to compensate for the resulting effect on the specified image colorimetry and on an observer’s perception of that colorimetry.

Again, the key to understanding the fundamental concept of this color-encoding method is to recognize that the purpose of the transformations is not to determine how the colorimetry or its perception would be altered by changes in viewing conditions. The purpose of the transformations is to maintain an intended color appearance even as viewing conditions are changed. That is why the encoding method is not constrained by the reference encoding conditions, why it supports input from any and all image-capture environments, and why it supports output to any and all display viewing environments.

Because the conditions for the CES reference viewing environments do not dictate actual conditions that must be used for input or output, they could be specified essentially arbitrarily. In practice, however, it is more sensible to specify conditions that are realistic and consistent with those of the associated inputs and outputs. Doing so enhances simplicity and minimizes the magnitudes of the colorimetric transformations.

Viewing conditions can be defined by specifying four characteristics: viewing flare, image luminance level, image surround type, and observer adaptive white. Suggestions for the viewing conditions of the Input CES and Output CES are given in the following subsections. They are presented primarily to serve as illustrations. The Committee may want to consider other specifications that might be better aligned with current or anticipated industry practices.

Input CES Reference Viewing Conditions

The objective here is to define a set of viewing conditions representative of most forms of input. To be consistent with the concept of scene-space encoding in the Input CES, the specifications are representative of those of live scenes viewed in average daylight.

Viewing Flare: 0%

This specification might seem incongruous, since many scenes will have a great deal of flare; but that flare is part of the image itself, and it is represented in the image’s colorimetry. The specification of zero flare recognizes that fact, and simply means that no further accounting for flare is required.

Luminance Level: 6000 cd/m^2 or greater

This value corresponds to the luminance of a white reference illuminated by daylight on a cloudy-bright day. The exact level is not critical. The specification is meant to convey a level at least two orders of magnitude greater than that of typical indoor viewing. This indicates that the Input CES colorimetry later will have to be adjusted for luminance-level effects as part of the rendering process to Output CES.

Surround Type: Normal

Scenes typically are “surrounded” by more of the same scene, and what is captured within the frame is what the photographer meant to convey. As with viewing flare, then, the image represents the actual situation, and no colorimetric adjustment is needed for scene-space encoding.

Adaptive White: Y=100, x=0.3324, y=0.3474

These values specify the colorimetry of a stimulus that would be perceived to a perfect achromatic white in the original scene. Note once again that this specification is colorimetric, not spectral. The chromaticity is that of CIE Standard Illuminant D_{55}, an illuminant reasonably representative of average daylight. However, as discussed previously, this specification defines the adaptive state of the observer; it does not specify, or imply, or necessarily correspond to, the actual light source of the captured scene.

An important aspect of the reference white specification is that it unambiguously defines the chromaticity of what will be considered an achromatic neutral in the encoding. Any image or image area intended to be perfectly achromatic thus should have the same chromaticity as the reference white. Consistent with the philosophy underlying the proposal, this is not a mandate that all "neutrals" must have that chromaticity; it is a designation of what is achromatic in the Input CES, what will be treated as such throughout the system, and what will appear achromatic in the final display(s). If the intent is to produce warm or cold color images or toned black-and-white images, such intents are fully supported by appropriate specifications of image colorimetry relative to the encoding reference-white chromaticity.

If these or similar specifications were used for the reference viewing conditions of the Input CES, the encoding of scene colorimetric values determined according to the procedures described in Section 7 would be straightforward. At most, a single transform would be needed to account for any difference in an observer's state of chromatic adaptation in the actual and reference environments.

A chromatic adaptation transformation can be a simple matrix (von Kries) operation. More complex transforms are available, but I do not recommend their use for imaging applications. Of particular concern is that some attempt to account for partial chromatic adaptation. This is a valid concept which recognizes that in many situations, the adaptive white will not be the same as the average chromaticity of the visual field. For example, an observer looking at images or scenes under 2800K illumination may only be adapted such that a chromaticity corresponding to a somewhat higher color temperature, perhaps 3000K, would appear achromatic. As a result, objects being viewed at 2800 K will appear somewhat yellow-orange.

In such models, the adaptation point would be entered as the chromaticity of 2800K, and the model would attempt to determine the actual adaptive chromaticity. In this proposal, however, the adaptive white means just what it says: it is an explicit specification of the chromaticity perceived as achromatic. In this example, then, the adaptive white would be specified as the chromaticity of 3000K. If that value were entered into a model that attempts to account for partial adaptation, the value would be inappropriately altered. In other words, there would be double accounting for the same effect. A von Kries transformation matrix is derived from the exact adaptive chromaticities specified, which makes it appropriate for use in this application.

11.2.1. Output CES Reference Viewing Conditions

The objective here is to define a set of viewing conditions primarily representative of the principal output of the system, digital cinema projection. The following is an initial estimate of those conditions:

Viewing Flare: 0.02% to 0.10%

This specification is meant to indicate that the system grayscale for the Reference Rendering to Output CES need only compensate for the relatively low level of viewing flare expected in motion picture theaters. A range of values is specified because a single value would imply an unintended degree of precision. If the flare were specified as 0.50%, for example, it might imply that re-rendering is required for output to an environment with only 0.40%. The specified range instead indicates that nothing further needs to be done in re-rendering for outputs viewed in environments reasonably similar to those of most theaters. However, viewing flare in home-video and reflection-print environments can be considerably higher, perhaps 0.5% to 1.0%. That degree of difference is significant and warrants additional compensation in re-rendering.

Luminance Level: 25 to 150 cd/m^2

This range is intended to encompass average white-reference luminance levels associated with viewing electronic and film theater projection, home video, and reflection prints. It basically is meant to convey a level at least two orders of magnitude lower than that of typical outdoor viewing. That degree of difference indicates that Input CES colorimetry has to be adjusted for luminance-level effects as part of the rendering process to Output CES.

Surround Type: Dark

This is self-explanatory.

Adaptive White: Y=100, x=0.3140, y=0.3510

The chromaticity coordinate values shown are from the most recent draft of the document "Proposed SMPTE Recommended Practice for Digital Cinema Reference Projector and Environment For Display of DCDM in Review Rooms and Theatres". This document is a work in progress, so the values may change. Although it is certainly worthwhile to specify an adaptive white chromaticity that is reasonable, what is most important is that it be defined explicitly and precisely. This specificity is needed to provide a rigorous definition of a visual neutral chromaticity in the Output CES. A reference neutral chromaticity is required to properly derive output transforms and chromatic adaptation transforms for output environments in which the adaptive white differs from that of the Output CES.

11.3. Color Encoding Data Metrics

At this point in the design of a color-managed system, the need to balance system requirements of signal-processing efficiency, digital quantization, color gamut, compression, etc. usually constrains the choices of possible data-metrics. In this system, however, no such constraints are apparent. In particular, it is my understanding that 48-bit encoding (16 bits per channel) will be used for encoding. As a result, there are many possibilities for the input and output data metrics, as discussed in the following subsections.

11.3.1. Input CES Data Metric

In the previous discussion regarding input signal processing, Input CES values were described in terms of XYZ values. One possibility would be to base the data metric directly on these values. Although there would be nothing wrong with that, I would suggest the use of RGB spaces also should be considered.

No signal-processing operations, at least that I am aware of, are best performed directly in XYZ space; so transformations are always required from XYZ to appropriate RGB spaces. Also, RGB spaces are somewhat more intuitive (high red values mean red colors, etc.). In a normalized RGB space, neutrals have equal RGB values. This makes neutrals easy to identify, and it makes them easy to track through a system for verification and troubleshooting. Moreover, equal RGB neutrals pass through normalized matrices, polynomials, and 3D tables untouched, which provides added simplicity and robustness to the system.

One downside of RGB spaces compared to CIE XYZ space is that the primaries would have to be selected. The Committee may decide it is not worth the trouble of working through that decision. If, however, an RGB space is considered, I would suggest starting from the ERIMM primaries. As discussed before, some adjustment should be considered for the blue primary. It should be possible to increase the chromaticity boundaries of the color gamut while maintaining (or even further reducing) the small hue rotations created when the space is used for Reference Rendering.

Given that 48-bit encoding is available, I would not anticipate problems using linear encoding. If quantization visibility does arise in testing, a nonlinear function such as that of ERIMM could be used. That function is essentially logarithmic, with a linear portion at low values for mathematical reversibility. The visual impact of using this or similar power-law functions roughly corresponds to adding another two bits per channel. That should be more than adequate to allow the encoded exposure dynamic range to equal that of a motion picture camera negative film.

11.3.2. Output CES Data Metric

As is the case with the Input CES, there does not appear to be anything that would strongly suggest the use of any particular data metric for the Output CES. However, since Input CES values will have to be converted to an ERIMM-like set of RGB primaries for Reference Rendering, it might make sense to simply stay with those primaries for the Output CES.

The dynamic range selected for the Reference Rendering grayscale should be used to set the limits for the Output CES data metric. Again, the range must represent luminance-factor values greater than those of the adaptive white. I would suggest that the highlight limit should correspond to an L^* of at least 130 (CIE Y of about 200). My understanding is that 48-bit encoding also will be used in the Output CES, so again quantization is not likely to be a problem.

11.3.3. Multiple Data Metrics

Because there are no technical criteria mandating the use of any particular data metrics, I would anticipate a suggestion will be made that multiple data metrics be allowed within each CES. This of course would require image files to be tagged or otherwise identified as to their metric. I cannot make a technical argument against this idea; but my experience is that, in practice, the use of multiple data metrics is undesirable.

An inclusive system like the one described in this paper can have a positive influence in unifying an industry. That benefit is lost, or at least diminished, when different groups continue to use “their own” metrics rather than one in common. It would be easy to dismiss this as inconsequential; but while subtle, the effect is real and ultimately detrimental.

11.3.4. Multiple Encoding Methods

A related proposal that might arise would allow inclusion of image files that differ not only in the data metric but also from the encoding method of the respective CES. I would argue strongly against that proposal. What it would create is not a single system at all; instead it would be a collection of functionally separate systems running not together but in parallel. Images would share little more than a common file format. There would be no meaningful unification, and the features discussed in this paper would not be realized. My experience is that when such proposals are made, it reveals a lack of understanding of the fundamental distinction between encoding methods and encoding data metrics, and a lack of recognition of the benefits that are realized in a truly unified system.

However, I would also caution against a quite opposite position that also can harm an industry. Often there is an attempt to mandate that *everything* must be done according to some new standard or recommended practice. Yet in every industry there are established practices, with skilled practitioners and highly developed methodologies, which work well and meet particular needs. My opinion is that there is no reason to change such practices. Moreover, history suggests change cannot be dictated. If necessary, practitioners will find ways of getting around new standards that interfere with what they need to do.

I would suggest, then, that the proposed system should be promoted not as a replacement for existing methodologies but as a means of unifying them. Successful practices need not be made obsolete. Instead, methods should be provided so that, at the appropriate place in the imaging path and at the appropriate point in the imaging workflow, images can be brought into the proposed unified system. This concept was discussed earlier in regard to input from computer generated images. Similarly, transforms from other workspaces, such as printing-density space for negative films, should be made available to allow current practices to continue as needed while providing for their incorporation into this larger, unified system.

12. Summary and Recommendations

This paper has described a proposal for implementing color management in a comprehensive, inclusive and extensible Digital Cinema System. The principal features and recommendations of the proposal are summarized below:

1) The recommended color-management architecture of the system is based on an Input Color Encoding Specification (CES), a separate Output CES, and a Reference Rendering transformation that links the two.

2) Each CES is optimized for its specific function.

3) In the Input CES, colors are represented in terms of colorimetric values measured in the absence of flare and associated with a defined scene-space reference viewing environment.

4) Input CES values can represent accurate scene colorimetry, adjusted scene colorimetry, manipulated scene colorimetry, scene colorimetry incorporating the looks of various input or output media, or imaginary scene colorimetry created by computer-generated imaging techniques.

5) Input CES values can be derived from electronic camera raw or processed image files, from computer generated image files, and from scans of negative and positive photographic media through the use of appropriate input signal-processing transformations.

6) It is recommended that manufacturers provide input transformations and/or characterization aids, and device calibration aids related to their products.

7) Input CES values can be transformed to and from other workspaces for image editing and manipulation.

8) Various data metrics could be used for Input CES images, including metrics based on linear CIEXYZ, linear RGB, or nonlinear RGB. However, the use of multiple Input CES data metrics is not recommended.

9) The data metric of the Input CES includes a luminance dynamic range equal to the exposure dynamic range of current color negative photographic motion picture films.

10) Input CES values are transformed to Output CES values by a Reference Rendering transformation.

11) The resulting rendered colors are represented in the Output CES in terms of colorimetric values measured in the absence of flare and associated with a defined reference viewing environment consistent with that of motion picture theaters.

12) The Reference Rendering transformation is based on the properties of a defined additive-color display device having a luminance dynamic range and color gamut that meet or exceed the capabilities of current or anticipated display devices and media.

13) It is recommended that the luminance dynamic range the Reference Rendering extends to an L^* of at least 130 (Y of approximately 200).

14) The Reference Rendering transform includes a conversion of Input CES values to a normalized RGB space having primaries selected to minimize undesirable hue rotations that might otherwise result from the transformation.

15) The Reference Rendering transform also includes nonlinear transformations and chroma adjustments to create colors appropriate for reference viewing.

16) Output Rendering transforms are used to gamut-map reference Output CES colorimetric values to colorimetric values attainable by various types of output display devices and media.

17) Output Characterization transforms are used to determine device code values for representative output devices and media.

18) It is recommended that Output Calibration transforms, used to determine device-specific code values, be separate from characterization transforms.

19) The use of device-resident calibration throughout the system is strongly recommended.

13. Acknowledgements

I would like to thank Andrew Maltz and Richard Patterson of the Science and Technology Council for extending the invitation to develop this proposal and for their support, encouragement and input during its preparation. I also would like to thank Thomas Maier and Gabriel Fielding of Eastman Kodak Company for their valuable input and reviews of preliminary drafts.

14. References

1. E. J. Giorgianni and T. E. Madden, Digital Color Management, Addison-Wesley, Reading MA, 1998.
2. "Basic parameter values for the HDTV standard for the studio and for international programme exchange," Recommendation ITU-R BT.709 (formerly CCIR Recommendation 709).
3. C. M. Daniels, E. J. Giorgianni, and M. D. Fairchild, The effect of surround on perceived contrast of pictorial images, *IS&T/SID 5th Color Imaging Conference, Scottsdale, 12-16 (1997)*
4. K. E. Spaulding, G. J. Woolfe and E. J. Giorgianni, "Reference Input/Output Medium Metric RGB color encodings (*RIMM/ROMM RGB*)," in *IS&T's 2000 PICS Conference*, pp. 155-163 (2000).
5. K. E. Spaulding, G. J. Woolfe and E. J. Giorgianni, "Image States and Standard Color Encodings *(RIMM/ROMM RGB),*" in *Eighth Color Imaging Conference: Color Science and Engineering: Systems, Technologies, Applications*, pp. 288-294 (2000).
6. "*KODAK Photo CD* System—A Planning Guide for Developers," Eastman Kodak Company, Rochester, NY (1991).
7. "Multimedia Systems and Equipment – Colour Measurement and Management – Part 2-1: Colour Management – Default RGB Colour Space – sRGB," IEC 61966-2-1 (1999).
8. "Proposed SMPTE Recommended Practice for Digital Cinema, Reference Projector and Environment for Display of DCDM in Review Rooms and Theatres", SMPTE, Rev. 2.7 (6/1/2005)

15. Appendix 1: Mathematical Transforms

The transformations described in this proposal can be implemented using a number of different mathematical techniques. The best method for a given application will depend, of course, on the hardware, software, and computational resources available. Nevertheless, different types of mathematical operations generally are appropriate for particular types of transformations, and the comments in the following subsections regarding each are applicable in most circumstances.

15.1. One-dimensional LUTs

One-dimensional LUTs are an obvious choice for transforming a set of channel-independent values to a new set of channel-independent values. Typical applications include reshaping grayscales for rendering, applying color balance and exposure shifts, applying or removing camera nonlinearity in digital video and JPEG images, input and output device calibration, etc.

In systems with limited resources, one-dimensional LUTs can be particularly useful. For example, in an 8-bits-per-channel system having 8-bit 1D LUTs and a 3D LUT limited in size to 64 cubed, it would be advantageous to perform any inherently 1-dimensional transformations using the 1D LUTs. Each would have 256 explicit input and output points, whereas the 3D LUT would have only 64 points (at the node locations along a one-dimensional slice of the table). Thus interpolation would be required in processing 8-bit data. My experience is that this often leads to problems, particularly hue shifts in neutrals and near-neutral colors.

15.2. Normalized 3x3 Matrices

In a normalized (or restricted) matrix, the coefficients of each row add to the same value. There are six independent variables—the off-diagonal coefficients—in a restricted matrix. The three diagonal coefficients are used to adjust the sums of their respective rows, thus they are dependent variables.

In imaging applications, this restriction is particularly valuable where matrices are used for color correction and for color-space transformations. When used with a normalized color space (i.e., a space in which neutrals are represented by equal values in all three channels), a normalized matrix will have no effect on neutral colors. For example, if a particular neutral has RGB code values all equal to 128, passing these values through a matrix that is row-sum normalized to 1.00 will leave all three values unchanged at 128. I highly recommend the use of normalized color spaces and normalized matrices, exactly for that reason. In particular, use of normalized color spaces and 1D-LUTs together with normalized matrices creates a simple and robust system in which neutral and color signal processing are separable. Normalized matrices can be removed entirely from the system without affecting the grayscale. This makes it simple and convenient to troubleshoot 1D problems, free of complications created by color interactions.

When applied in normalized linear spaces, normalized matrices can be used to convert from one set of RGB primaries to another. Similarly, conversions can be made to spaces in which luminance and chrominance are separate (YCC spaces, which are really just a special type of RGB). Conversion to a YCC space is often done for image compression, where the degree of spatial subsampling applied to the chrominance signals generally is greater than that applied to the luminance signal.

A normalized matrix in linear space will primarily affect high-chroma colors. That is because in a linear space, the ratios among the color channels are large for such colors. In a nonlinear space, such as a logarithmic space, a normalized matrix will tend to affect most non-neutral colors somewhat independently of their chroma level. In a system that includes matrices in both spaces, this distinction provides some ability to adjust colors differently as a function of their chroma level. That can be useful to adjust, for example, the hue of skin tones without affecting the hue of higher-chroma reds.

15.3. Unnormalized 3x3 Matrices

In an unnormalized matrix, the row sums are not restricted to have equal value. This means the matrix can produce different gains in each row, in addition to the interactions produced by the off-diagonal terms. An unnormalized matrix combines a 1x3 matrix, which produces the gains, with a normalized 3x3.

In linear exposure space, changing the gains equally in each color channel is equivalent to changing the overall exposure level. Changing the gains unequally in the channels is equivalent to changing the color balance (or white point). An unrestricted matrix can be used, then, to transform between spaces where the white points are different. One example of this usage would be for what is commonly called a phosphor matrix. Such matrices are used to transform RGB signal values to monitor CIE XYZ values. Phosphor matrices are derived based on the chromaticity of each of the three phosphors and the chromaticity of the monitor white.

A likely use of an unnormalized matrix in the proposed system would be for scanner calibration. Transforming from scanner RGB density values to ISO Status M values will almost certainly require individual red, green, and blue gain adjustments in addition to any cross-talk adjustments. Of course this could be done instead using 1D LUTs and restricted matrices. But that would be necessary only if it is determined that linearity differences, and not just simple differences in gains, are involved.

15.4. Three-by-Four Matrices

A 3x4 matrix is a 3x3 matrix (restricted or unrestricted) with constants included in each row. In linear exposure space, increasing a constant of positive value corresponds to an addition of exposure or light, which is like adding flare light. Using a negative value, then, provides flare compensation. This is a convenient way to compensate for varying amounts of flare, without having to alter any system 1D or 3D tables.

In log-exposure space, changing the constants equally is equivalent to a change in overall exposure (and equivalent to a gain in linear exposure space). Changing the constants unequally shifts the color balance and is similar to using a colored filter over the camera lens.

One appropriate use of a 3x4 matrix in the system is for scanner calibration. Transforming from scanner RGB density values to ISO Status M values is likely to involve differences in D_{min} values. Use of 3x4 matrix allows any such differences to be accounted for by the constants. Similarly, the constants can be used to provide densitometric measurements in terms of D_{min}–subtracted values, which I have recommended be used in the system.

When running a regression that should require only a 3x3 matrix, it is advisable to make an initial run using 3x4 equations. If everything is as it should be, the determined constants will be very nearly zero. If so, the regression for the 3x3 matrix can be run with confidence. This is a useful way to check that there are no unexpected offsets "hidden" in the data.

15.5. Polynomial Equations

Polynomial equations can be used in color-imaging applications where simple matrices are insufficient to create or compensate for complex color interactions. In the proposed digital cinema system, the most likely need for such equations would be in the transforms used to convert film integral density (Status M or Printing Density) values to channel-independent density (CID) values, as required in determining film RGB exposure values.

In addition to an included 3x3 matrix, a set of polynomial equations also would have cross-product terms (e.g., red x red, red x green, red x blue) and possibly terms of higher than second order. (My "personal best" was a set of 3x36 equations used in modeling a photographic slide film.)

Some caution should be exercised here. In fitting measured data, the use of increasingly complex equations generally will yield better statistical fits. However, when there is noise in the data (which is always the case for densitometry), a very complex set of equations may simply be providing a better fit to that noise. A set of less complex equations will tend to smooth over measurement noise, and therefore may provide a more useful transform.

15.6. Three-dimensional LUTs

Three-dimensional LUTs are commonly used in image signal processing because they provide fast execution of complex transformations. Depending on their size, three-dimensional LUTs also can provide an almost unlimited number of degrees of freedom in relating input and output values. That is both their strength and their weakness. It is an obvious strength in that it allows very complex relationships to be characterized. However, that strength also provides opportunities for inappropriate use.

As with complex polynomials, 3D LUTs can be "overzealous" in fitting every last wrinkle of noise in a set of measured data. Therefore, it is highly recommended to first fit raw data using a set of polynomials or other equations of no more complexity than necessary for a satisfactory fit. The resulting equations then can be used to compute the data required to construct the 3D-LUT. This procedure not only generates a smoother set of data, it also provides a method of generating values that exactly correspond to the nodes of the table.

There is another danger in using such a powerful tool in that it can encourage less thinking and less understanding of the actual mechanisms occurring in the imaging process. A 3D-LUT can simply be used to fit input and output values with no regard as to what the values actually mean. Perhaps the best example of this involves the derivation of transforms used for converting RGB exposure values from non-colorimetric films and digital cameras to CIE XYZ values. It can take some experience and thought to come to the realization that an appropriately derived 3x3 matrix performs this "impossible" transformation as well as it can be done. The extra degrees of freedom provided by 3D LUTs are of no value, and can do great harm, in applications such as this where the relationships between the input and output values are neither systematic nor totally predictable. The most appropriate use of such tables is in applications where interactions that are complex, but also systematic, are to be characterized.

16. Appendix 2: Calibration Tools

A number of calibration methods and tools were discussed in the body of this paper. They are described in somewhat greater detail here.

16.1. Compensating Gray Charts

Compensating gray charts consist of a series of neutral test patches mounted in randomized order on a uniform gray background. They are used to measure device or media grayscale characteristics when measurements might otherwise be confounded by nonuniformities in lighting, optics, sensors, or other components. Patches and backgrounds generally consist of Munsell neutrals or other spectrally-nonselective materials.

16.2. Flare Targets

A compensating grayscale target together with an independent measurement of a flare-free grayscale can be used to measure camera flare. Camera flare can be determined by comparing a film grayscale image of a compensating gray target to a grayscale made by contact exposure on the film. The same grayscale-comparison technique can be used for electronic cameras by characterizing the sensor opto-electronic response in a manner that is independent of camera optics.

16.3. Uniformity Cards

Test charts for measuring camera uniformity generally consist of nothing more than a uniform gray card. However in using such cards, care must be taken to ensure that the illumination is perfectly uniform. Alternatively, a telephotometer can be used to measure the illuminated card from the camera position. Any measured nonuniformity then can be factored into subsequent camera measurements. Also, care should be taken to ensure that the chart is oriented such that it is parallel to the image plane of the camera.

16.4. Compensating Color Charts

Compensating color charts consist of an array of color patches mounted on a uniform gray background. They are used to measure device or media color-capturing characteristics when measurements might otherwise be confounded by nonuniformities in lighting, optics, sensors, or other components. Neutral test patches and backgrounds generally are constructed from Munsell papers or other spectrally-nonselective neutrals. Color patches should have spectral reflectance characteristics representative of actual colors of principal interest. The selection of colors and spectral characteristics becomes increasingly important as the correspondence of the spectral responsivities of the taking device or medium to a set of color-matching functions decreases.

16.5. Negative Film Targets for Calibration Only

Negative film targets for scanner calibration must be on the same film stock that is to be scanned and must be in a format that is appropriate for scanning and subsequent measurement. Multiple targets can be used if necessary. If so, each should have at least one patch in common to provide a means of detecting and correcting any image-to-image variations in scanning.

Test areas should include a D_{min}, a grayscale, color scales, and an assortment of colors. The exact color set is not critical because all colors will have been formed from combinations of a limited number of film image-forming dyes. Certified ISO Status M values should be provided for all target colors.

16.6. Negative Film Targets for Scanner Calibration and Input Transformations

Negative film targets for use in developing input transformations and for use in scanner calibration must be on the same film stock that is to be scanned and must be in a format that is appropriate for scanning and subsequent measurement. Multiple targets can be used if necessary. If so, each should have at least one patch in common to provide a means of detecting and correcting any image-to-image variations in scanning.

Test areas should include a Dmin, a grayscale, color scales, and an assortment of colors. The color set used here is critical because it will influence the transformation from film RGB exposure values to CIE colorimetric values. Certified ISO Status M values, film RGB exposure values, and Input CES values should be provided for all target colors.

16.7. Positive Film Targets for Scanner Calibration and Input Transformation Development

Positive test targets on print-film stock for use in developing input transformations and for use in scanner calibration must be on the same film stock that is to be scanned and must be in a format that is appropriate for scanning and subsequent measurement. Multiple targets could be used if necessary. If so, each should have at least one patch in common to provide a means of detecting and correcting any image-to-image variations in scanning.

Test areas should include a Dmin, a grayscale, color scales, and an assortment of colors. The exact color set used is not critical because all colors will have been formed from combinations of a limited number of film image-forming dyes. Certified ISO Status A values, Input CES values, and (optionally) Output CES values should be provided for each individual target sample.

16.8. Reference Input CES Digital Image Files

An array of digital image files, certified to be in Input CES space, should be provided to users of the proposed system. Such images serve as checks to which images from other sources can be compared. Certified input images also provide a means to examine and troubleshoot the entire imaging chain.

Images should include uniform areas, grayscales, color scales, and a factorial array of color patches. Scales or images of objects having highlight-to-shadow series (varying luminance, constant hue and saturation) at various hue angles should be included. An assortment of pictorial images also should be included. Ideally, images would be created using computer generated imaging techniques to ensure that all image values are ideal Input CES values.

16.9. Reference Output CES Digital Image Files

Digital images having certified Reference Output CES values also should be provided to system users. Such images would allow system inputs and outputs to be tested and verified independently. In particular, they would provide a means of visually confirming setups of monitors and other display devices. They would also provide a basis of comparison for Output CES images generated by other means. Certified Reference Output CES images can be generated readily using the above Reference Input CES digital images together with a certified Reference Rendering transform.

17. Glossary

a* b* diagram
A plot of *a** and *b** values of the 1976 CIE *L*a*b** (CIELAB) color space.

absorption
The transformation of radiant energy to a different form of energy by interaction with matter; retention of light without reflection or transmission.

achromatic
Perceived as having no hue; white, gray, or black.

adaptation
The process by which the visual mechanism adjusts to the conditions under which the eyes are exposed to radiant energy.

adaptive white
A color stimulus that an observer, adapted to a set of viewing conditions, would judge to be perfectly achromatic and to have a luminance factor of unity.

additive color
Color formed by the mixture of light from a set of primary light sources, generally red, green, and blue.

advanced colorimetry
Colorimetric measurement and numerical methods that include colorimetric adjustments for certain physical and perceptual factors determined according to perceptual experiments and/or models of the human visual system.

AgX
Silver halide; a light–sensitive crystalline compound used in conventional photographic materials.

average surround
An area, surrounding an image being viewed, that has a luminance factor of about 0.20 and chromaticity equal to that of the observer adaptive white; also called a normal surround.

bit
Contraction of *binary digit*; the smallest unit of information that a computer can store and process.

block dye
A theoretical dye having equal absorption of light at each wavelength within a given range of wavelengths and no absorption at all other wavelengths of interest.

brightness
An attribute of a visual sensation according to which an area appears to exhibit more or less light.

brightness adaptation (general)
The process by which the visual mechanism adjusts in response to the overall luminance level of the radiant energy to which the eyes are exposed.

brightness adaptation (lateral)
A perceptual phenomenon wherein a stimulus appears more of less bright depending on the relative brightness of adjacent stimuli.

CCD
Abbreviation for *charge coupled device*; a solid state sensor often used in digital still cameras and scanners to convert light into an electrical signal.

CCIR (Comité Consultatif Internationale des Radiocommunications)
Abbreviation for the *International Radio Consultive Committee*, an international television standardization organization, now ITU–R.

CCIR Recommendation 601
A document of recommended specifications for digital component video, now referred to as Recommendation ITU-R BT.601, or more informally as Rec. 601.

CCIR Recommendation 709
A document of recommended specifications for high–definition television signals, now referred to as Recommendation ITU-R BT.709, or more informally as Rec. 709.

CD–ROM
Abbreviation for *compact disk read–only memory*; a compact disc used for storing digital data for computer applications.

calibration
Procedure of correcting for any deviation from a standard.

CES
Abbreviation for *color encoding specification*.

channel independent
An imaging channel that produces a signal that is detectable and separate from signals produced by other imaging channels.

characterization
Procedure of defining the color characteristics for a representative operating model of an input or output device.

CCD (charge coupled device)
A solid state sensor often used in digital still cameras and scanners to convert light into an electrical signal.

chroma
1) The colorfulness of an area judged in proportion to the brightness of a similarly illuminated area that appears to be white; degree of departure of a color from a gray of the same lightness.
2) A color component of a color video signal.

chroma subsampling
A technique for compressing image information, generally for storage or transmission, in which luma (achromatic) information is retained at full spatial resolution while chroma (non–achromatic) information is reduced.

chromatic adaptation
The process by which the visual mechanism adjusts in response to the average chromaticity of the radiant energy to which the eyes are exposed; changes in the visual system's sensitivities due to changes in the spectral quality of illuminating and viewing conditions.

chromaticity
The property of a color stimulus defined by its chromaticity coordinates, such as its CIE *x, y, z* values.

chromaticity coordinates
The ratio of each of a set of tristimulus values to their sum.

chromaticity diagram
A plane diagram in which points specified by chromaticity coordinates represent the chromaticities of color stimuli.

chrominance
The properties of a color other than its luminance.

CIE (Commission Internationale de l'Eclairage)
The *International Commission on Illumination*; the body responsible for international recommendations for photometry and colorimetry

CIE colorimetry
Measurement of color stimuli according to the spectral responsivities of a CIE Standard Observer.

CIE 1931 Standard Colorimetric Observer
An ideal colorimetric observer with color–matching functions corresponding to a field of view subtending a 2° angle on the retina.

CIE tristimulus values
The values *X*, *Y*, and *Z*, determined according to the color–matching properties of the CIE 1931 Standard Colorimetric Observer.

CIELAB color space
A color space, defined in terms of L^*, a^*, and b^* coordinates, in which equal distances in the space represent approximately equal color differences.

CIELUV color space
A color space, defined in terms of L^*, u^*, and v^* coordinates, in which equal distances in the space represent approximately equal color differences.

CIEXYZ color space
A color space defined in terms of tristimulus values *X*, *Y*, and *Z*, which are determined according to the color–matching properties of the CIE Standard Colorimetric Observer.

CIS
Abbreviation for *color interchange standard*.

CMY/CMYK
Abbreviations for *cyan* (C), *magenta* (M), *yellow* (Y), and *black* (K), dyes or inks used in subtractive color imaging.

clipping
Condition where variation of an input signal produces no further variation of an output signal.

code value
A digital value produced by, or being provided to, an imaging device.

color encoding
The numerical specification of color information.

color–encoding data metric
The numerical units in which encoded color data are expressed.

color–encoding method
Measurement methods and signal–processing transformations that determine the meaning of encoded color values.

color encoding specification (CES)
A fully specified color encoding scheme, defined by a color–encoding method and a color–encoding data metric, used for encoding color on an individual system. A complete CES also may include specifications for other factors, such as data compression method and data file format.

colorfulness
Attribute of a visual sensation according to which an area appears to exhibit more or less of its hue.

color gamut
The limits of the array of colors that can be captured by an image–capturing device, represented by a color–encoding data metric, or physically realized by an output device or medium.

colorant
A dye, pigment, ink or other agent used to impart a color to a material.

colorimeter
Instrument that measures color stimuli in terms of tristimulus values according to responsivities prescribed for a standard observer.

colorimetry
A branch of color science concerned with the measurement and specification of color stimuli; the science of color measurement.

colorimetric characteristics
Referring to characteristics, such as the color–reproduction characteristics of a device, medium, or system, as measured according to standard colorimetric techniques.

colorimetry (standard)
In this book, refers to colorimetric values determined according to current CIE recommended practices.

color interchange specification (CIS)
A fully specified color interchange scheme that includes a complete colorimetric specification, a defined data metric, and defined a set of reference viewing conditions. A complete CIS also may include specifications for other factors, such as data compression method and data file format.

colorist
An operator who adjusts the electronic signal processing in the transfer of photographic images to video.

color management
The use of appropriate hardware, software, and methodology to control and adjust color in an imaging system.

color–matching functions
The tristimulus values of a sequence of visible monochromatic stimuli of equal radiant power.

color primaries (additive)
Independent light sources of different color (usually red, green, and blue) which may be combined to form various colors.

color primaries (subtractive)
Colorants, each of which selectively absorbs light of one of the additive primaries. A cyan colorant absorbs red light, a magenta colorant absorbs green light, and a yellow colorant absorbs blue light.

color stimulus
Radiant energy such as that produced by an illuminant, by the reflection of light from a reflective object, or by the transmission of light through a transmissive object.

composite (transform)
A single signal–processing transform formed by the concatenation of a sequence of two or more individual transforms.

compositing
Merging portions of various images to form a single image.

compression
A process used to reduce the size of data files, generally for storage or transmission.

concatenation
Process of combining a sequence of two or more individual signal–processing transforms to form a single equivalent transform.

cones
Photoreceptors in the retina that initiate the process of color vision.

contrast (objective)
The degree of dissimilarity of a measured quantity, such as luminance, of two areas, expressed as a number computed by a specified formula.

contrast (subjective)
The degree of dissimilarity in appearance of two parts of a field of view seen simultaneously or successively.

control voltage (CRT)
Voltage signal used to modulate beam current, and thus light output, of a CRT.

corresponding colorimetric values
Colorimetric values for corresponding stimuli (see below).

corresponding stimuli
Pairs of color stimuli that look alike when one is viewed in one set of adaptation conditions, and the other is viewed in a different set.

coupler
An organic compound, used in photographic media, which reacts with an oxidized developing agent to form a dye.

coupler (colored)
A coupler (see above) that is itself colored.

cross–talk
Transfer of information from one color channel to another.

CRT
Abbreviation for *cathode ray tube*.

cyan
One of the subtractive primaries; a cyan colorant absorbs red light and reflects or transmits green and blue light.

DAC
Abbreviation for *digital–to–analog converter*.

dark surround
An area, surrounding an image being viewed, having a luminance much lower than that of the image itself.

dark–surround effect
A manifestation of lateral–brightness adaptation; an observer will perceive an image to have lower luminance contrast if that image is viewed in darker–surround conditions.

data metric
The numerical units in which a given set of data are expressed.

daylight
A mixture of skylight and direct sunlight.

daylight illuminant
An illuminant having the same, or nearly the same, relative spectral power distribution as a phase of daylight.

densitometer
A device for directly measuring transmission or reflection optical densities. For meaningful color measurements, the RGB spectral responses of the densitometer must be specified.

densitometry
The measurement of optical density.

density (optical)
The negative logarithm (base 10) of the reflectance factor or transmittance factor.

device–independent color
As defined by the author, refers to techniques for numerically specifying and encoding color information in a way that is not restricted to either the luminance dynamic range or the color gamut achievable by physically realizable devices.

diffuse
Referring to light that is scattered, widely spread, not concentrated.

digital color encoding
The representation of color information in the form of digital values.

digital quantization
Conversion of continuous quantities to discrete digital values; the number of discrete values is determined by the number of bits that are used.

digitize
Convert analog signals or other continuous quantities to digital values.

display
An image presented to an observer; the process of presenting that image.

duplicate
A reproduction that is a one–to–one physical copy of an original. The spectral properties of the colorants of a duplicate are identical to those of the original.

dyes
Organic colorants used in silver–halide–based photographic media and in other imaging technologies.

dynamic range
Extent of minimum and maximum operational characteristics.

encoder and decoder circuits
Used in video systems to combine various signals into a composite signal and to subsequently extract the individual signals from the composite.

exposure
The quantity of radiant energy received per unit area; the quantity of radiant energy that is captured by a detector or that forms a detectable signal.

exposure factor
Ratio of exposure to that from a perfect diffuser that is illuminated identically.

field
That portion of the surface of a specimen that is illuminated by the illuminator or viewed by the receiver.

film terms
Input signal–processing transforms used on Kodak Photo CD System scanners to convert scanned values to PhotoYCC Space values.

film writer
An output device, used in hybrid color–imaging systems, which produces an image on a photographic film.

flare
Stray light; a non–imagewise addition or redistribution of light.

fluorescence
Process whereby incident radiant power at one wavelength is absorbed and immediately re–emitted at another (usually longer) wavelength.

gamma (photographic)
The slope of the straight–line portion of a characteristic curve relating optical density to relative log exposure.

gamma (CRT)
1) Exponent of a power–law equation relating CRT luminance to control–signal voltage
2) The slope of the straight–line portion of a CRT characteristic curve relating log luminance to log voltage.

gamma correction
The use of signal processing in a video camera to complement the characteristics of a video display device such as a CRT.

gamut (color)
The limits for a set of colors.

gamut boundary
Outermost surface of a color space defined by a particular color gamut.

gamut adjustment (or gamut mapping)
A method for replacing colorimetric values corresponding to colors that are not physically realizable by a considered output device or medium with substitute values that are attainable by that output. In some methods, values within the attainable gamut also are altered.

grayscale
A progression of achromatic colors from blacks to grays to white.

HDTV
An abbreviation for *high–definition television* system, a system having greater spatial resolution than that of current broadcast television systems.

hardcopy
General term referring to solid media such as paper or film base.

hue
Attribute of a visual sensation according to which an area appears to be similar to one, or to proportions of two, of the perceived colors, red, yellow, green, and blue.

hybrid (color–imaging) system
A system which incorporates photographic and electronic imaging technologies.

ICC
Abbreviation for International Color Consortium, an industry group formed in 1993 to promote interoperability among color–imaging systems.

illuminant
A light, which may or may not be physically realizable as a source, defined in terms of its spectral power distribution.

illuminant sensitivity
Propensity for colors formed by a set of colorants to change in appearance as the spectral power distribution of the viewing illuminant is changed.

image dyes (image–forming dyes)
Dyes, usually CMY or CMYK, that make up a displayable image.

independent (primaries)
Sets of light sources in which the chromaticity of each source can not be matched by any mixture of the remaining sources.

ink
A color liquid or paste used in printing.

input
General term referring to imaging media, signals, or data to be put into a color–imaging system.

input compatibility
Expression used by the author to describe the result of color encoding images such that encoded values completely and unambiguously specify the color of each pixel on a common basis, regardless of the disparity of the sources of the image data.

intensity
Flux per unit solid angle; used in this and other texts as a general term to indicate the amount of light.

interlayer effects
Chemical reactions that take place among the various layers of a photographic medium. These interactions are used for color signal processing.

ISO
Abbreviation for *International Standards Organization*.

isotropic
Independent of direction.

ITU
Abbreviation for *International Telecommunications Union*. The United Nations regulatory body covering all forms of communication. ITU–R (previously CCIR) deals with radio spectrum management issues and regulation.

JPEG
Abbreviation for *Joint Photographic Experts Group*; a set of standards developed by this group for compressing and decompressing digitized images.

latent image
An image consisting of a small cluster (a few atoms) of metallic silver within a silver halide crystal, formed by exposure of the crystal to light. During chemical signal processing, crystals with latent–image sites are developed to metallic silver, while those without latent–image sites are not.

lateral brightness adaptation
A perceptual phenomenon wherein a stimulus appears more of less bright depending on the relative brightness of adjacent stimuli.

light
1) electromagnetic radiant energy that is visually detectable by the normal human observer, radiant energy having wavelengths from about 380 nm to about 780 nm.
2) adjective denoting high lightness

light source
A physically realizable emitter of visually detectable electromagnetic radiation, defined in terms of its spectral power distribution.

lightness
The brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting.

look–up table (LUT)
A computer memory device in which input values act as the address to the memory, which subsequently generates output values according to the data stored at the addressed locations.

luma
The achromatic component of a video signal.

luminance
A measure, of a luminous surface, that is an approximate correlate to the perception of brightness.

luminance contrast
Apparent rate of change from lighter to darker areas of an image. Luminance contrast approximately corresponds to grayscale photographic gamma.

luminance dynamic range
Extent of maximum and minimum luminance values, often expressed as a ratio, e.g., 1000:1, or as a logarithmic range, e.g., 3.0 log luminance.

luminance factor
Ratio of the luminance of a specimen to that of a perfect diffuser that is illuminated identically.

magenta
One of the subtractive primaries; a magenta colorant absorbs green light and reflects or transmits red and blue light.

metameric color stimuli
Spectrally different color stimuli that have the same tristimulus values.

metameric pair
Two spectrally different color stimuli that have the same tristimulus values.

metamerism (visual)
Property of two specimens that match under a specified illuminator and to a specified observer and whose spectral reflectances or transmittances differ in the visible wavelengths.

metamerism (degree of)
Reference to the extent to which matching stimuli are spectrally different. A pair of stimuli that match but have very different spectral characteristics are referred to as being highly metameric.

metamerism (instrument)
Property of two specimens that measure identically according to the spectral responsivities of an instrument and whose spectral reflectances or transmittances differ in the wavelengths of those responsivities.

monitor white
Color stimulus produced by a monitor when maximum red, green, and blue code values are applied; measured values for that stimulus.

monochromatic
Of or producing electromagnetic radiation of one wavelength or of a very small range of wavelengths.

nanometer (nm)
Unit of length equal to 10^{-9} meter, commonly used for identifying wavelengths of the electromagnetic spectrum.

negative
A photographic medium, usually intended to be printed onto a second negative–working photographic medium, that forms a reversed image, i.e., higher exposure levels result in the formation of greater optical density.

neutral
Achromatic, without hue.

normal surround
An area, surrounding an image being viewed, that has a luminance factor of about 0.20 and chromaticity equal to that of the observer adaptive white; also called an average surround.

nm
Abbreviation for *nanometer*.

observer metamerism
The property of specimens having different spectral characteristics and having the same color when viewed by one observer, but different colors when viewed by a different observer under the same conditions.

opto–electronic transfer characteristic
Characteristic defining the relationship between exposure and output signal voltage for a video camera.

output
General term referring to images, signals, or data produced by color–imaging systems.

PCS
Abbreviation for *profile connection space*, a fully–defined color space used for linking and/or concatenating a series of profiles.

perfect white
An ideal isotropic diffuser with a spectral reflectance factor or spectral transmittance factor equal to unity at each wavelength of interest.

phosphors
Materials, deposited on the screen of a cathode ray tube, which emit light when irradiated by the electron beam(s) of the tube.

Photo CD Player
A device, similar to an audio compact disc player, which is used to display images from Photo CD Discs on conventional television receivers and monitors.

Photo CD System
A hybrid color–imaging system, developed by Eastman Kodak Company, which produces compact discs of images by scanning and digitally encoding images from photographic media.

PhotoYCC Color Interchange Space
The data metric of the Kodak Photo CD System, in which color data are encoded in terms of a luma value, Y, and two chroma values, C_1 and C_2.

photographic image–forming dyes
The cyan, magenta, and yellow dyes which are formed by the chemical processing of a photographic medium after exposure of that medium to light.

photon
A quantum of light or of other electromagnetic radiation.

pigment
Finely ground insoluble particles that, when dispersed in a liquid vehicle, give color to paints, printing inks, and other materials by reflecting and absorbing light.

pixel
Contraction of *picture element*; a single point sample of an image.

positive
A photographic medium, usually intended for direct viewing, in which higher levels of exposure result in the formation of less optical density.

power
Energy per unit time.

prepress
Term used to describe the process, or components of the process, of preparing information for printing after the writing and design concepts stages.

primaries
Basic colors used to make other colors by addition or subtraction.

printing density
Optical densities measured according to effective spectral responsivities defined by the spectral power distribution of a printer light source and the spectral sensitivities of a print medium.

principal subject area
The area of a scene that is metered or otherwise used in the determination of camera exposure.

product–specific film terms
Input signal–processing transforms used in *Photo CD* Imaging Workstations to convert scanned values to *PhotoYCC* Space values. A product–specific film–term transform is based on the characteristics of the particular film being scanned. When product–specific film terms are used, differences among scanned films are minimized in the color encoding.

profile
A digital signal–processing transform, or collection of transforms, plus additional information concerning the transform(s), device, and data.

profile (abstract)
A profile providing the information necessary to modify color values expressed in a profile connection space (PCS).

profile (destination, or output)
A profile providing the information necessary to convert color values expressed in a profile connection space (PCS) to output device values.

profile (source, or input)
A profile providing the information necessary to convert input device values to color values expressed in a profile connection space (PCS).

profile connection space (PCS)
A fully–defined color space used for linking and/or concatenating a series of profiles.

psychological (signal processing)
Modifier used in this book to refer to visual signal processing that includes higher order mental and cognitive (interpretive) processes.

psychophysical (signal processing)
Modifier used in this book to refer to visual signal processing that includes both physiological and mental processes.

purple boundary
On a CIE chromaticity diagram, the straight line connecting the red and blue ends of the spectrum locus.

quantization
Conversion of continuous quantities to discrete digital values; the number of discrete values is determined by the number of bits that are used.

raw (or RAW) file
Digital camera image files containing unprocessed data from the camera's image sensor, usually in terms of code values that are proportional to exposure.

Rec. 601
Informal name for *Recommendation ITU-R BT.601*, formerly CCIR Recommendation 601, a document containing recommended specifications for digital component video.

Rec. 709
Informal name for *Recommendation ITU-R BT.709*, formerly CCIR Recommendation 709, a document containing recommended specifications for high–definition television signals.

reference image–capturing device
A hypothetical device, associated with the color encoding of the *Kodak Photo CD* System, defined in terms of spectral responsivities and opto–electronic transfer characteristics.

reflectance
Ratio of the reflected radiant or luminous flux to the incident flux under specified conditions of irradiation.

reflectance factor
The amount of radiation reflected by a medium relative to that reflected by a perfect diffuser.

rendered image
An image having attributes that make it appropriate for display and viewing.

rendering
The process of converting scene colorimetric values to values appropriate for image display.

retina
Layer on the back interior of the eyeball, containing various types of photoreceptive cells that are directly connected to the brain by means of the optic nerve.

relative colorimetry
Colorimetric values expressed relative to those of a reference white. In standard CIE calculations, the reference white is defined to be a *perfect* white. In "media–relative" colorimetry, the support of the particular medium being measured is defined as the reference white.

RGB
Abbreviation for *red, green, and blue*.

SBA
Abbreviation for *scene balance algorithm*, an algorithm that automatically adjusts the overall lightness and color balance of images.

SMPTE
Abbreviation for *Society of Motion Picture and Television Engineers*.

saturation
The colorfulness of an area judged in proportion to its brightness.

scanner
A device for forming image–bearing signals from two–dimensional images.

scene balance algorithm (SBA)
An algorithm that automatically adjusts the overall lightness and color balance of images.

sensitivity
Property of a detector which makes it responsive to radiant power.

signal processing
Chemical, electronic, or digital operations, such as linear and nonlinear amplification, by which original signals are altered and/or combined with other signals.

silver halide
A light–sensitive crystalline compound used in conventional photographic materials.

simulation
The use of one medium or system to imitate the appearance of another.

softcopy
Jargon for electronic displays such as CRTs.

source
1. A physically realizable light, the spectral power distribution of which can be experimentally determined.
2. An imaging–system term for *origin*.

spatial compression
A technique for reducing image information, generally for storage or transmission.

spectral
Adjective denoting that monochromatic concepts are being considered.

spectral power distribution
Power, or relative power, of electromagnetic radiation as a function of wavelength.

spectral reflectance
The fraction of the incident power reflected as a function of wavelength.

spectral reflection density
Reflection density as a function of wavelength; the negative logarithm of spectral reflectance.

spectral responsivity
The response of a detection system, such as a scanner or a densitometer, as a function of wavelength. Spectral responsivity is influenced by the spectral power distribution of the illuminant, the spectral

filtration effects of various optical components, and the spectral sensitivity of the detector.

spectral sensitivity
The response of a detector to monochromatic stimuli of equal radiant power.

spectral transmittance
The fraction of the incident power transmitted as a function of wavelength.

spectral transmission density
Transmission density as a function of wavelength; the negative logarithm of spectral transmittance.

spectrum locus
On a chromaticity diagram, a line connecting the points representing the chromaticities of the spectrum colors.

specular
Referring to light that is reflected or transmitted with little or no scattering.

Standard Illuminants
Relative spectral power distributions defining illuminants for use in colorimetric computations.

Standard Colorimetric Observer
An ideal observer having visual response described according to a specified set of color–matching functions.

speed
Term used in photography to describe sensitivity to light. Higher speed means greater sensitivity to light, lower speed means lesser sensitivity to light.

Status A densitometer
Densitometer having spectral responsivities corresponding to those specified by the ISO for Status A densitometers. Status A densitometers are used for measurements of photographic and other types of hardcopy media that are meant to be viewed directly by an observer.

Status M densitometer
Densitometer having spectral responsivities corresponding to those specified by the ISO for Status M densitometers. Status M densitometers are used for measurements of photographic negative media.

stimulus (color)
A spectral power distribution, such as that produced by an illuminant, by the reflection of light from a reflective object, or by the transmission of light through a transmissive object.

subsampling
Sampling within samples; a technique employed to compress digital image files.

subtractive color
Color formed by the subtraction of light by absorption, such as by cyan, magenta, and yellow (CMY) photographic dyes or by cyan, magenta, yellow, and black (CMYK) printing inks.

surface color
Color perceived as belonging to the surface of a specimen, without the specimen appearing to be self–luminous.

surround
The area surrounding an image being viewed.

surround effect
A manifestation of lateral–brightness adaptation; an observer will perceive an image as having lower or higher luminance contrast depending upon the average luminance of the surround relative to that of the image.

tags
In an image file or profile, descriptors of the underlying data.

telecine
An imaging system used to scan motion picture films to produce video signals for taping and television broadcast.

test target
A collection of color samples used in the evaluation of color–imaging systems, generally comprised of spectrally nonselective neutral samples and samples of various colors.

thermal printer
An output device which uses heat to transfer dyes to produce images on reflection or transmission media.

transform
One or more signal processing operations, used in color–imaging systems incorporating digital signal processing.

transmittance
Ratio of the transmitted radiant or luminous flux to the incident flux under specified conditions of irradiation.

transmittance factor
The amount of radiation transmitted by a medium relative to that transmitted by a perfect transmitting diffuser.

transparency
An image formed on a clear or translucent base by means of a photographic, printing, or other process, which is viewed by transmitting light through the image.

trichromatic
Three color.

trichromatic system
A system for specifying color stimuli in terms of tristimulus values based on matching colors by additive mixture of three suitably chosen reference color stimuli.

tristimulus values
The amounts of three matching stimuli, in a given trichromatic system, required to match a particular color stimulus.

tungsten lamp
An electric lamp having filaments of tungsten.

tungsten–halogen lamp
Lamp in which tungsten filaments operate in an atmosphere of low–pressure iodine (or other halogen) vapor.

uniform color space
Color space in which equal distances approximately represent equal color differences for stimuli having the same luminance.

universal film terms
Input signal–processing transforms used on Photo CD Imaging Workstations to convert scanned values to PhotoYCC Space values. A universal film–term transform is based on the characteristics of a reference film of the same basic type as that being scanned. When universal terms are used, differences of each scanned film from the reference film are reflected in the color encoding.

unwanted (spectral) absorption
Spectral absorptions of a colorant in portions of the spectrum where ideally there should be 100% transmission or reflection.

u', v' ***diagram***
Uniform chromaticity diagram, introduced by the CIE in 1976, in which u' and v' chromaticity coordinates are used.

viewing conditions
Description of the characteristics of a viewing environment that physically alter a color stimulus or that affect an observer's perception of the stimulus.

viewing flare
Stray (non–imagewise) light that is present in an environment in which an image is viewed. The amount of viewing flare usually is expressed in terms of an amount relative to the amount of light reflected from, transmitted through, or produced by a white in the image.

visual density
Density measured according to a responsivity corresponding to the CIEXYZ $\bar{y}$(l) function.

visual neutral
A metameric match to a spectrally nonselective neutral viewed under identical conditions.

von Kries transformation
A chromatic adaptation transformation by which changes in chromatic adaptation are represented as adjustments of the sensitivities of the three cone systems.

wavelength
In a periodic wave, the distance between two points of corresponding phase in consecutive cycles.

white balance
The process of adjusting the RGB signals of a video camera such that equal signals are produced from an illuminated white object.

writer
General term for output devices that use photographic films or papers.

x, y ***diagram***
A chromaticity diagram in which the x and y chromaticity coordinates of the CIE XYZ system are used.

yellow
One of the subtractive primaries; a yellow colorant absorbs blue light and reflects or transmits red and green light.

zeroing
Adjustment of an instrument such that a zero signal value would be obtained when an ideal reference specimen is measured. For example, reflection densitometers generally are adjusted such that a zero–density reading would be obtained if a perfect white diffuser were measured.