Issues Relating to the Transformation of Sensor Data into Standard Color Spaces

Jack Holm
Hewlett-Packard Laboratories
Palo Alto, California 94304

Background

Since the introduction of the Sony Mavica in the early eighties, people have speculated about the demise of film as the primary means for making photographs. In some cases, such as videotape vs. small format motion picture film, the demise occurred rather rapidly. However, film systems have retained primacy in still imaging, despite numerous predictions to the contrary. The reason for this is that for electronic systems to become dominant, they need to successfully compete with film systems on all fronts. VHS and Beta can seriously compete with Super 8, but as of a few months ago no digital camera could compete with a film camera of even remotely similar cost.

This situation is changing. Until recently, a significant number of technical problems had to be solved to make digital photography practical. These included low cost production of low and medium resolution sensors, image data storage protocols and devices, capture devices, and hardcopy output devices. Now, solutions for many of these problems are entering the marketplace. However, one obstacle remains—the optimization of image processing and the associated problem of the interpretation of image data, processed or unprocessed. This remaining problem is of extreme importance, and could considerably slow the acceptance of digital photography if not addressed correctly and completely.

A number of processing techniques in the areas of spatial reconstruction and color have the potential for proprietary advantage. It is not reasonable to expect competing companies to expend resources developing algorithms to be shared with competitors. However, the work done could prove of little use if the algorithms produced are not implementable in some sort of standardized framework. The responsibility of standards developers is therefore to make absolutely sure that all potentially valuable strategies can be implemented. Fortunately, formal standards protocols are set up to address input from all legitimate sources and require broad consensus. Formal procedures may be relatively slow, but should allow for robust standards if they are followed with reasonable rigor. Also, it is possible to move rapidly through formal procedures if resources are expended to deal with comments and develop consensus through aggressive research and communication.

However, it is essential that the parties involved in standards development recognize that politicization of the formal standards organizations for personal gain or commercial advantage is devastating to the process. All participants, without exception, must work solely to establish a fair, non vendor specific baseline structure. Vendor differentiation results from how well each company can use the structure. If a particular company or group of companies attempts to bias standards to the disadvantage of others, the entire formal structure breaks down. Corporate standards strategies should center on determining where formal standardization is desirable. Standards to be used for leveraging particular technological approaches are most appropriately done by industry consortia, which have structures more suited to this purpose. This leaves the formal structures intact so they can fulfill the purposes to which they are suited.

The goal of formal standards development in relation to the processing of image data is therefore to allow for crystal clear description and communication of the nature of the data in any form that may be required (as opposed to convenient) for a variety of processing algorithms and applications. This can be achieved by specifying the physical characteristics of capture and/or output devices, along with the nature of any encoding used for transmission and storage. First order standards relating to color reproduction should therefore define highly repeatable methods for measuring digital camera and scanner opto-electronic conversion functions (OECF's), and spectral sensitivities and/or spectral products when used in combination with standard or designated illumination sources. Analogous measurement standards are required for output devices, and all of these standards must provide for high enough accuracy so that measurement errors are insignificant when creating pictorial images. Several standards along these lines are under development[1,2] and more are needed.

Some image data formats, such as TIFF/EP,[3] allow OECF and spectral information to be included in image file headers. Few image processing applications make use of such information, primarily because in pictorial imaging it is possible to simplify the data description. This simplification is based on the fact that the human visual system (HVS) does not spectrally analyze light, but has a limited number of spectrally integrating channels. If the sole purpose of the data is to produce an image for viewing, one can mathematically transform the data into a representation based on color matching functions, or a color space. This does not mean that spectral data is unnecessary, as it may be used for determining the transformation, and in spatially reconstructing the image. After this processing is done, however, the description of the resulting image data can be greatly simplified. In many cases it will also be easier for subsequent processing algorithms to use the transformed data.

Discussion

If one assumes that image data will be transformed into a color space, it follows that formal processes may be a good way to determine and establish standard color spaces. The intent is for them to be used by anyone. The danger of this type of work is that it is very important to choose standard spaces which are well suited to all applications, to choose as few spaces as possible to reduce complexity, and to provide an extremely rigorous descriptions of the spaces. Some guidance as to the intended uses of the spaces is also helpful. A number of formal standards organizations have recognized this need, and work is proceeding on several fronts. Some time ago, the CIE defined color spaces in general by standardizing the CIE 2° and 10° observers, and the associated color matching functions.[4] More recently, the ITU standardized some RGB primaries, based on the CIE 2° color matching functions, which are representative of cathode ray tube displays in general.[5] Two obvious choices for standard color spaces are therefore already in existence, CIE XYZ and ITU-R BT.709 based RGB. When expressed in terms of linear radiance, these spaces and the transformations between them are well defined.

However much remains to be done, because in a digital world it is extremely inefficient to represent image data in terms of linear radiance. HVS perception is strongly nonlinear, so in maintaining the necessary accuracy in dark image areas, much more accuracy than is necessary is maintained in bright areas. Rendering is also an issue. Rarely is it desirable to for a reproduction to have the same colorimetric description as that of a scene. The white points may be different, the dynamic range of the reproduction medium may be different from that of the scene, the viewing conditions and states of adaptation may be different, and viewers often prefer reproductions in which tones and colors has been altered for æsthetic reasons.

Perceptually compact representations, white points, dynamic range differences, viewing condition differences, and reproduction preferences are all separate issues. Unfortunately, in the past they have frequently been confused to the point of causing serious problems and a mistrust of computational color reproduction in some areas. In many cases this confusion has resulted from oversimplification and a lack of understanding. It may be possible to lump the above considerations together to produce a reproduction model for a particular situation, but when this model is applied in a different situation it no longer functions correctly because the different considerations interact differently.[†] An explicit understanding of the nature of each consideration is necessary for the development of generic approaches. It is also important to note that several considerations are not related to appearance; a perfect appearance model, if and when one is developed, will still not deal with every consideration relevant to digital photography.

The most pathological situation in digital photography is the capture of natural scenes. Transformations to standard color spaces are indeterminate because the spectral correlation statistics of the scene radiances are unknown and frequently variable across the scene. A large variety of white points, states of adaptation, and viewing conditions are possible. Dynamic ranges are frequently anywhere between 10:1 and 2000:1. The capture of natural scenes is therefore the most general problem to be solved in color reproduction. Once this problem has been solved, the same generic philosophy can be applied to all other pictorial imaging systems. However, the solutions for specific applications may appear to be different: film scanners may be able to take advantage of known film spectral correlation statistics, copying systems do not need to repeat preferred reproduction (unless it was not done initially), most reflection media and monitors have similar dynamic range capabilities, etc.

In defining color spaces for digital photography, it is therefore necessary to explicitly define non-linearities, considerations relating to appearance, and considerations relating to preferred reproduction. The considerations relating to appearance also need to be distinguished from each other. At present, we have fairly good ideas about which factors affect appearance, but are considerably more in the dark about exactly *how* and *why* these factor affect appearance. The safest approach is to specify the factors and leave the treatment to the user of the standard, which is the same way preferred reproduction is handled. This means that while digital photography standards may be based to some extent on color spaces, they must be based on physical metrics as opposed to appearance measures.

The preceding discussion points to the necessity of applying rendering processing to the data captured by digital cameras. This processing should take into account appearance as well as preferred reproduction issues. In designing the algorithms these issues may be separated, but the processing itself can be viewed as a black box. Formal standards which support this approach therefore need to specify standard color spaces which apply to image data before it is rendered (data that colorimetrically describes the scene), and after it is rendered (data that colorimetrically describes the reproduction).

Given these distinctions, it is possible to envision six scenarios for the transformation of data into standard color spaces. A seventh scenario which represents the video paradigm is also described. These scenarios are outlined because it is important to be aware of the exact purpose of a transformation when it is determined.

Output Rendering

In this scenario, the raw sensor data is rendered for reproduction on a particular output device by a single program. The image data appropriate for the designated output device is then saved. Output rendering programs fold the transformation of the sensor data into a standard color space, the appearance and preferred reproduction considerations, and the output device characteristics together. Color spaces for output rendering *must* describe a physically realizable output medium. An example of an output rendering color space is the proposed ISO display RGB.[6] It would also be possible to base an output rendering space

---

[†] A specific illustration of such a situation is as follows: One notices that if the media white point relative CIE L*a*b* measurements of an image displayed on a 6500K monitor in a dim room and a reflection print viewed using 500 lux tungsten illumination are made to be equal, that the print and monitor representations will appear to be similar in successive viewing with adaptation, particularly if a white surround is used. However, if the monitor is then placed under the tungsten illumination used to view the print, the monitor will appear to be too blue because of the partial adaptation to the tungsten illumination, and too dark because of the reduction of dynamic range due to veiling glare. It is interesting to note that the monitor image will appear to be too dark, even though the L* values, if re-measured, will have increased to be lighter than those of the print.

on metrics using a different transformation of CIE XYZ, such as CIE L*a*b*. However, in specifying an output space, it is essential that physically measurable values are used, and that all considerations which might affect appearance or preferred reproduction (such as white points, dynamic range, and viewing conditions) are clearly delineated.

In this discussion, it is probably worth mentioning that the referenced ISO standard monitor RGB white paper[6] has not been formally proposed as a new work item as of the time this manuscript was written. This is because it is very similar to another proposal initiated in the IEC for a standard color space designated as sRGB.[7] The ISO committee felt that it would be better to see how the IEC proposal developed before proceeding with another work item. For purposes of discussion, the standard monitor RGB descriptions provided in the ISO white paper are used.

The advantage of output rendering is that the image data is immediately available for reproduction on the designated output device. The disadvantages are that the image data will frequently be substantially changed from that captured by the sensor, the relation of the image data to the original scene may not be known (this is undesirable for archiving), and if the rendered data does not produce a pleasing image it may be difficult or impossible to re-render it effectively. The latter disadvantage is particularly significant if the user wishes to re-render the image for output on another device with substantially different characteristics.

Output rendering is currently the most likely candidate for consumer digital photography, where immediate accessibility is important, image quality must only exceed that of consumer photofinishing, and the vast majority of output will be monitor display and reflection hardcopy (similar dynamic range).

Embedded Transform Output Rendering

This approach to output rendering involves embedding the output rendering transform in the image file without actually transforming the data. The advantage of this modification is that it is much more acceptable for archiving, since the original data is saved and can be re-rendered. This re-rendering may be substantially easier, and the value of the archived raw data further enhanced, if the camera OECF and spectral information is saved along with the embedded output rendering transform.

Source Rendering

With source rendering, the image data is transformed into a standard color space, but the color values are estimates of the scene colorimetry. It is also frequently desirable to deal with one appearance issue in source rendering—the white point. If no white point is specified, the source rendering must be into high bit depth linear CIE XYZ (or a linear combination thereof). If a white point is specified, it becomes possible to render into more perceptually compact spaces such as the proposed ISO source RGB or CIE L*a*b*. Color spaces for source rendering should be unbounded in dynamic range, and therefore cannot exactly represent real output media.

The advantage of source rendering is that it produces image data well suited to archiving, and can be fed into generic appearance/preferred reproduction algorithms. The disadvantage is that the data is not ready for display. There is also some risk if a white point based source rendering space is used, in that the appearance decision about the white point will have already been made. An incorrect decision will cascade through the rest of the imaging chain to produce poor results.

Embedded Transform Source Rendering

It is also possible to embed the transform in the image file with source rendering. An additional advantage of doing so in this case is that it is possible to concatenate the source transform with some rendering transform and thereby allow for the precision of direct rendering while deferring the decision about the exact nature of the rendering transform. As with embedded transform output rendering, the value of the archived raw data is enhanced if the camera OECF and spectral information is saved along with the embedded source rendering transform.

Source Rendering with Embedded Output Transform

Another variation which is quite useful for image data supplied by digital stock agencies is source rendered data with an embedded output rendering transform. This allows potential purchasers to view an output rendered image, while retaining access to the source rendered data which can be re-rendered by generic programs for a particular output medium and/or artistic intent.

Embedded Transforms for Source and Output Rendering

This scenario is probably the most appropriate for archival stock agencies, where it is desirable to keep the raw data and also provide both source and output rendered images. As with the other embedded transform approaches, the value of the archived raw data is enhanced if the camera OECF and spectral information is saved along with the embedded source rendering transform.

Video Rendering

Video rendering is a special case where the image data is source rendered into ITU-R BT.709 RGB, with the associated gamma function. This source rendered data is then just assumed to be equivalent to output rendered ISO display RGB. In effect, the appearance/preferred rendering transform is the difference between ITU-R BT.709 RGB and ISO display RGB. This first order approximation works reasonably well for commercial video, where scene dynamic ranges are controlled, and in consumer video, where image quality expectations are minimal. It does not produce very good results for pictorial still imaging, except in situations where the dynamic range is relatively fixed at around 50:1 (such as with some types of studio photography).

Since the gamma functions of ITU-R BT.709 and the ISO display RGB are different, video rendering results in a system gamma somewhat greater than unity, and a corresponding boost in luminance contrast and color saturation. This is consistent with preferred reproduction. However in copying applications, preferred reproduction is not desirable - the goal is an appearance match. If one wishes to use a video rendering type approach for copying, it would be better to source render the data into ISO source RGB, and then consider the result to be ISO display RGB.

Tools for Color Management

If agreement is reached on a perspective for color reproduction, formal standards processes provide the opportunity to create tools for color management. Several tools have already been created as described previously. A large amount of new work is also in progress, with one new item related to digital cameras discussed below.

Proposed New ISO Work Item

Last year, a joint ISO TC42 (Photography) and TC130 (Graphic Technology) Task Force was established to propose the development of formal standards in the areas discussed above. In particular, it was felt that formal standards designating methods for determining source rendering transforms for digital cameras would be useful. ISO TC130 had already developed a standard defining targets for determining source rendering transforms for transparency and reflection print scanners.[8,9] An initial proposal was then developed, and is outlined below:

Test Objects and Procedures for the Colour Characterization of Electronic Still Cameras.

Scope

This international standard shall specify test objects, metrology, and procedures for the colour characterization of electronic still cameras.

Purpose and Justification

The spectral response of electronic still cameras does not, in general, match that of a typical human observer, such as that defined by the CIE standard colorimetric observer. Neither do they match each other. Thus, it is necessary to take account of the camera sensitivities, scene illumination, and reference color space.

This standard will address this problem by defining test images, metrology, and procedures for various situations. It will address the problem of such cameras in their most general application; where metameric colours and a range of illuminants may be encountered. However, it will recommend procedures for more closely defined situations in which the illuminant and colorants being imaged are better known.

General Information

The prescribed methods determine transformations for transforming sensor data into standard color spaces.

The goal of the transformations produced is to describe the scene or original using the destination color space, so the purpose of the transformations is source rendering.

The standard does not specify transformations for output rendering, and therefore does not consider appearance/preferred reproduction issues, with the exception of the white point.

Currently, von Kries transformations are used for white point changes, however this should probably be revised once a single more up-to-date white point transformation method, such as the Bradford transformation method, becomes generally accepted.

The default scene illumination sources are as defined in ISO 7589[10] and ISO 14524.[1]

Transformations are defined into CIE XYZ, CIE L*a*b*, and ISO source RGB.

The reference white used to normalize the XYZ values in the calculation of CIE L*a*b* and source RGB values has the same spectral characteristics as a perfectly diffuse reflecting or transmitting white illuminated by the illumination source used, except where the camera is used to capture real (three dimensional) scenes. In this case, the white point luminance is increased by a factor of 1.414 to accommodate (to some extent) the specular reflections that occur in real scenes.

All transformations to L*a*b* or source RGB are white point preserving.

The linearization of the data is accomplished using inverse OECF's as measured according to ISO 14524.[1]

Method A

Method A is applicable under all conditions and is based on camera spectral sensitivity measurements. With method A, the transformation matrix T is determined through matrix multiplication of three matrices and their transposes according to the equation:

$$T = O^t C^t M [M^t C^t M]^{-1}$$

where M is a matrix containing the camera spectral sensitivities, O is a matrix containing the output color space color matching functions, and C is the spectral correlation matrix. A more complete description of these calculations is provided in the paper "White-Point Preserving Color Correction," which can be found in these proceedings.[11]

The spectral correlation is assumed to be one of the following:

1. White point constrained maximum ignorance.
2. Some standard set of surface reflection statistics combined with some known illuminant spectral power distribution (in which case the reflectance statistics and illuminant used should be specified).
3. The actual spectral radiance correlation statistics of the scene or original, if known.

If the camera spectral sensitivities are color matching functions, method A reduces to a linear transformation between color spaces.

Method B

Method B is a target based general method for use when camera spectral sensitivity measurements are not available, and the colorants used in the scene or original are unknown, or are not spanned by the camera analysis channels. The steps involved are as follows:

1. Measure the standard method B test target (the exact nature of which is to be determined) under the desired illumination source, and calculate values for the patches of the target as expressed in the destination color space.
2. Capture image data of the target.
3. Determine the constant coefficient transformation matrix that produces the minimum mean square error between the linearized image data and the XYZ or linearized RGB measured target values, with the transformation matrix constrained to preserve neutrals. (L*a*b* values are calculated from the XYZ values.)

Method C

Method C is applicable when the camera spectral sensitivities are color matching functions. If this is the case, there is a well defined transformation to all of the standard color spaces.

Method D

Method D is applicable when the colorants found in the scene or original are known and spanned by the camera analysis channels. In this case the test target used should be made of the same colorants as are found in the scene or original. The procedure for determining transformation is the same as with method B, except the specialized target is used, and the matrix coefficients are not required to be constants or to preserve white points. Since the colorants used are actually known, more accurate transformations can be obtained by allowing the form of the transformation to be flexible. However, extreme care must be taken to prevent this additional flexibility from allowing errors into the transformations which might result in them producing objectionable results. This method is most appropriate for film and print scanners, in which case the appropriate IT8 targets can be used.[8,9]

Supporting Research

Many of the methods proposed in this standard were evaluated in the research presented in the paper "Matrix Calculations for Digital Photography," which is also included in these proceedings.[12] The results of this research tend to validate the proposed methods, with the exception of the test chart used. The specular reflection characteristics of the Macbeth Color Checker were found to be too variable for it to be generally used to determine repeatable transformations. Other test charts with better surface reflection characteristics, or transmission charts, may give more repeatable results.

Conclusions

A great deal of work remains to be done to achieve consistently excellent color reproduction in digital photography, particularly with images obtained using digital cameras. Formal standards can facilitate the advance and growth of this field by providing a sound framework on which products can be developed. However, there is some risk that standards work could hinder the growth of the industry by establishing structures which are biased toward particular technologies or applications, or do not allow for the implementation of some approaches. It is essential that any formal standards that are developed support a broad and universal view of digital photography. This view need not be segmented according to markets, because if it is truly universal, it will encompass all markets, and can be refined and simplified for particular applications.

An important step forward is the proposed new work item to specify methods for determining source rendering transforms for digital cameras. The general nature of the digital camera color reproduction problem will result in the solution to this problem having broad implications on color management in general. Other important new work relates to the continued formal establishment and acceptance of a few source rendering color spaces, and the development of well defined and useful output rendering color spaces. Future work is also needed with respect to color negative capture. It is the hope of the author that the open and non-competitive spirit which has been embodied in the development of formal photographic standards in the past will continue into this new era.

References

1. ISO/DIS 14524, *Photography - Electronic still picture cameras - Methods for measuring opto-electronic conversion functions (OECF's)*, 28 May 1997.
2. IEC/WD 61966 - 3, *Colour management in multimedia systems - Part 3: Equipment using cathode ray tubes*, ver. 3.12, 1997.
3. ISO/CD 12234-2, Photography - Electronic still picture cameras—Removable memory, Part 2: Image data format—TIFF/EP, 1997.
4. CIE Publication 15.2, *Colorimetry, 2nd edition*, 1986.
5. ITU-R BT.709, *Basic parameter values for the HDTV standard for the studio and for international programme exchange*, 1993.
6. ISO TC42/TC130 white paper, *Graphic Technology and Photography—Standard monitor RGB specifications*, 28 April 1997.
7. IEC/NP 61966 - 2.1, *Colour management in multimedia systems - Part 2.1: Default RGB colour space - sRGB*, 1997.
8. CGATS/ANSI IT8.7/1, *Graphic Technology - Color transmission target for input scanner calibration*, 1993.
9. CGATS/ANSI IT8.7/2, *Graphic Technology - Color reflection target for input scanner calibration*, 1993.
10. ISO 7589, *Photography - Illuminants for sensitometry - Specifications for daylight and incandescent tungsten*, 1984.
11. G.D. Finlayson & M.S. Drew, "White-Point Preserving Color Correction," *IS&T's 5th Color Imaging Conference Proc.,* p. 258 (1997).
12. P. M. Hubel, J. Holm, G.D. Finlayson & M.S. Drew, "Matrix Calculations for Digital Photography," *IS&T's 5th Color Imaging Conference Proc.,* p. 105 (1997).

☆ This paper was previously published in *IS&T's 5th Color Imaging Conference Proc.,* p. 290 (1997).