Aesthetic Considerations in Tone and Color Management

Robert F. Poe
Color Management Systems, Eastman Kodak Company
Billerica, Massachusetts

Abstract

Color-management engineers have concentrated their attention on scientific issues (colorimetry, adaptation) in their efforts to provide consistent color reproduction on a variety of media and devices. Aesthetic issues have mostly been left to the user. However, some of these issues are related to the characteristics of media and are amenable to analysis and automation. A study of tone management exemplifies this approach.

Introduction

Contemporary color-management systems have arisen in an effort to facilitate some measure of color consistency among the various peripheral devices and media accessible on a computer system or network. Prior to the emergence of these systems, it was the usual case that hard-copy output of the same image on various devices was highly variable in appearance and generally differed from a soft-copy display of the image on the system monitor and, in the case of scanned photographs, from the rendering on the original medium. Furthermore (and this is a point that has been insufficiently stressed), the output reproductions were frequently unacceptable in quality, when judged on their own terms, as well as being inconsistent with one another. Indeed, for many applications the inconsistency among various renderings might have been quite tolerable if the individual renderings had been aesthetically pleasing.

Nonetheless, the focus of the color-management effort has been on consistency in color reproduction, rather than on aesthetic considerations. In the prevailing architectural model, the latter are felt to lie in the domain of the user, interacting with the application software, where aesthetic choices can be made on an image-by-image basis. Color-management, on the other hand, is a system-level function intended to handle the properties and idiosyncrasies of devices and media, so that the application level can be independent of these concerns. Color consistency is a logical goal, therefore, for color-management: the user makes aesthetic color choices, in a device-independent manner, at the application level; the system then attempts to obtain the selected colors on any desired device or medium.

Unfortunately, this concept, as attractive as it may seem, does not work out so neatly in practice. For one thing, there has been some confusion over the meaning of “color consistency”. Initially, color-management relied on the concepts and methods of *CIE colorimetry*[1] to achieve “device-independent color”. This approach has been quite successful in improving color consistency and user satisfaction, but as color management is extended to a wider range of devices and media it has become clear that something is lacking. Colorimetry is a technique for quantifying visual *color stimuli*, but it does not describe the response of the human visual system to these stimuli. The visual response, or *color appearance*, is a complex neuro-psychological phenomenon that is heavily influenced by the state of adaptation of the observer receiving the color stimuli, which, in turn, is affected by the viewing environment and conditions. Thus, the appearance of a color stimulus on a hard-copy print viewed in reflection under typical office lighting will differ from that of the same stimulus presented on a CRT in a semi-darkened room and from that of the same stimulus projected by a tungsten lamp from a 35 mm slide onto a screen in a fully-darkened room.

Today the goal of color-management systems has become that of consistent color appearance, rather than consistent color stimuli, and the better systems attempt to attain this goal by adjusting or correcting the colorimetry for various adaptive effects in renderings on media intended for different viewing conditions. These efforts were sanctioned last year by the International Color Consortium (ICC) in its specification[2] for a universal profile format to be used by any color-management system on any computer platform; adaptive corrections were explicitly mandated as responsibilities of the ICC profile. (In the ICC system, each color device has a “profile” defining the transformation between its native coordinates and a device-independent *Profile Connection Space*, or PCS.) There is still some uncertainty concerning the nature, magnitude, and implementation of these corrections, but progress can be expected over the next few years.

But color appearance, while a clarifying concept, may ultimately prove insufficient as a system goal. Once these systems have evolved to a point where they really are capable of producing consistent color appearance across a broad range of media, the long-ignored aesthetic issues will finally have to be addressed. When the users’ demand for consistency has been satisfied, many of them will find that they still are not obtaining the quality of color reproduction that they expect on various media.

In many cases, of course, aesthetic considerations are image-specific and, therefore, lie outside the domain of color management. However, there is a class of aesthetic issues which are medium-specific, rather than image-specific, and which can (and should) be handled by the system. Color-reproduction media can be characterized by their capabilities and limitations—their color gamut and dynamic range. The most satisfactory renderings on different media will, in general, *not* have the same color appearance, because they will be "tuned" to the characteristics of the individual media, adjusting gracefully to their limitations and making effective use of their capabilities. In many cases, this tuning has occurred empirically, evolving over time according to the progress of technology and the tastes of particular markets or applications. Thus, various media acquire different "looks" or "personalities"—low contrast, high contrast, low chroma, high chroma, enhanced yellows, exaggerated greens, subtle pastels, etc. Thus, a 4-color offset press sheet, prepared by separating a photographic transparency, does not necessarily aim at the appearance of a snapshot, a projected slide, or a video image: it has its own intrinsic look.

Therefore, to a certain extent the aesthetic aspects of color reproduction are not device-independent, and the line between the application level and the system level does not coincide with the line between aesthetics and consistency. In particular, a color-management system should offer the option of automatically modifying the personality of an image when cross-rendering from one medium to another. Sometimes, certainly, the goal is to preserve the appearance of an original picture or graph, including the characteristic look of the original medium; but, at other times, the goal is to re-render the original, to bestow upon it a new personality better suited to the output medium.

Aesthetic Issues in Tone Reproduction

In order to analyze these possibilities, it is helpful to restrict the discussion to a simpler related problem: the management of tone (which is one-dimensional), rather than color (which is three-dimensional). Here it is useful to study the tone-reproduction curves of various media in rendering images of real-world scenes. This approach may be unfamiliar to those experienced in computer graphics, in which there is often no real-world scene to serve as a reference, but it is common among those working directly with the design of color-reproduction systems, such as photography, printing, and television, or in certain branches of digital image processing. The treatment here is similar to that in Hunt;[3] the purpose of the present discussion is to broaden the scope of color management to include aesthetic considerations and to encourage research and analysis into the technical issues involved.

We begin with the problem of reproducing a daylight-illuminated outdoor scene in an indoor viewing environment. Initially, at least, our goal is to preserve the color appearances of the scene, in the presence of various adaptive and physical effects. Probably the dominant effect to consider is the much lower level of illumination present in the viewing environment than in the scene—say, 300 lux, rather than 30000 lux. Now, the visual system adapts very well even to such large differences, so that (to a first approximation) it is sensitive primarily to the *relative* luminances of the objects. However, the adaptation is not perfect: it normalizes the overall lightness and darkness, but there will be a loss of apparent contrast in the reproduction. It has been found that an increase in the "gamma" of the reproduction is needed to recreate the original contrast.

It is useful here to imagine a spectrally nonselective, perfectly diffusing surface of 100% reflectance, illuminated in the same way as the other objects; we will call this *diffuse white*, and it will serve as the white reference both in the scene and in the viewing environment. Figure 1 is a logarithmic plot showing the relation between the luminances in the scene and those in the reproduction, both referenced to diffuse white. (The ordinate is actually *visual density*—the negative of the logarithm of relative luminance.) The curve marked "gamma = 1" represents the identity relation: relative luminances in the scene are rendered by the same relative luminances in the viewing environment. The curve marked "gamma boost" indicates the required contrast compensation. A third curve, marked "gamma boost, flare comp.", has been further compensated for a typical level of stray light ("flare") in the viewing environment, which also tends to lower the tonal contrast by physically lightening the shadows.

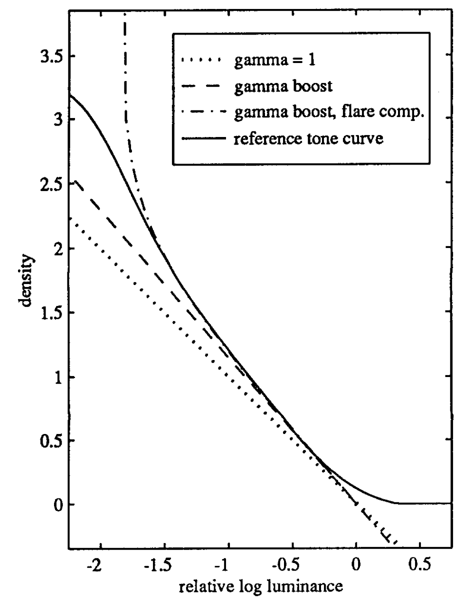

Figure 1. Constructing the reference tone curve

The resulting curve represents a theoretical aim for indoor tone reproduction in a "normal surround" (i.e., where the illumination of the environment is similar to that of the image), corresponding to the *reference viewing environment* of the ICC profile specification. Thus, this curve may be regarded as one that preserves the (tonal) appearance of the scene, within the limitations of the viewing conditions. (The assumed flare puts a lower limit on reproducible dark tones.) In line with the ICC specification, we may hypothesize a *reference medium* having a substrate of 100% reflectance (zero density): in order to reproduce tones brighter than diffuse white in the scene (such as specular reflections, self-luminous or fluorescent objects, secondary sources of illumination, etc.), we will want to reshape the "toe" of the tone curve so as to compress these highlights smoothly into the range of the medium. We will also be well advised to reshape the "shoulder" of the curve so as to compress the dark shadows of the scene into the flare-limited range of the viewing environment. Thus, we arrive at the solid line of Figure 1, labeled "reference tone curve"; it describes an idealized color-reproduction system which is capable of recreating the appearance of a real-world scene on the ICC reference medium in the ICC reference viewing environment. It has only a minimal personality, due to the reshaping of the toe and shoulder.

It is interesting, at this point, to observe that certain actual systems have tone-reproduction curves that closely resemble this reference curve. Figure 2 compares the reference tone curve with the tone curves of broadcast television and an experimental negative-positive system that was found to produce excellent reflection prints. Within the dynamic-range limitations of these media, the agreement is striking; we are tempted, therefore, to describe these media as, by and large, preserving the (tonal) appearance of the scene and recreating it in the viewing environment.

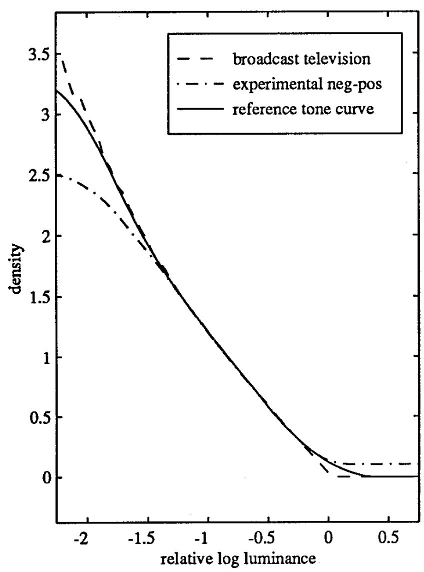

Figure 2. Some real media and the reference tone curve

On the other hand, there are other common media having tone curves of distinctly different shapes, which might be regarded as altering the appearance of the scene—perhaps enhancing it—and lending a distinctive personality to it. Figure 3 contains examples of these, in comparison with the reference curve. One of these is for an instant-film product, which can be compared directly with the negative-positive system of Figure 2: it employs its limited dynamic range to reproduce a relatively small part of the scene's luminance range with heightened contrast. The curve labeled "reversal film" has been adjusted in accordance with a model of the adaptive mechanisms present in the viewing environment of a darkened projection room, in which there is no white reference outside of the projected image itself. The curve marked "idealized video aim" has undergone similar adjustments; it corresponds to a proposed rendering of high-quality images on a CRT display in a dim-surround environment. Note that these last two curves extend into the negative-density region, indicating that these media, in their intended viewing environments, can create the illusion of tones brighter than diffuse white—brilliant, sparkling reflections from shiny surfaces, for instance—, albeit compressed. (Unfortunately, the ICC reference medium is limited to zero density.) These tone curves also maintain higher three-quarter-tone and shadow contrast than the reference curve, which should tend to impart a particular personality to the renderings.

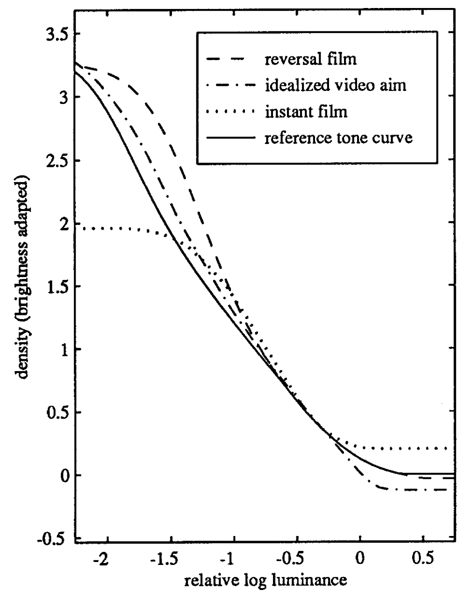

Figure 3. Ref. tone curve and media with different personalities

Tone Management

The existence of a variety of tone-reproduction curves, some of which may have evolved along with the media and others which may have been consciously designed, poses some questions for color-management systems. It seems possible, to some extent, to separate those aspects of these empirical curves which seem to compensate for adaptive effects from those which are based on aesthetic considerations. Should the aesthetic aspects—the personality of the medium—be preserved, or should they be transformed by these systems when cross-rendering from one medium to another?

There is no simple answer to this question. In some applications the overriding purpose is to produce the most aesthetically pleasing rendering possible on any output medium selected by the user. In others it is important to preserve the look of the original—to use the output as a simulation of the input.

Thus, we anticipate the need for a variety of transforms for different purposes. In the ICC context, this means a number of profiles for each device, selected according to need. We propose that the preservation of appearance be regarded as the base-level capability for a profile, as it is more amenable to objective specification (based on colorimetry and appearance-modelling) and is generally useful for a number of applications. Other profiles, designed to modify the apparent tone and personality of images, may be optimized for those applications requiring them, such as high-quality printing and publishing or presentation graphics.

Basic (Appearance-Preserving) Profiles

The straightforward way to achieve the base-level capability is to provide input profiles that produce PCS coordinates that accurately record the tonal appearance of the original and output profiles that accurately recreate the tonal appearances specified in the PCS. This goal can be achieved through colorimetry, device modelling, and modelling of the adaptive corrections required for specific viewing conditions. However, note that: (a) since the reference medium cannot achieve negative densities, those media and devices which *can* will need to have some highlight tonal compression in their input profiles and decompression in their output profiles; (b) the procedure recommended above is contrary to Appendix D of the ICC spec, which states that "the D_{min} of the input medium must be mapped to the $(XYZ)_{D50}$ point of the PCS"—i.e., diffuse white; and (c) this procedure is based on "absolute", not "relative", colorimetry.

To elaborate on these last points: luminance (more properly, *luminance factor*) in CIE colorimetry is commonly quantified with respect to a diffuse-white reference. This procedure has come to be called "absolute colorimetry" in the color-management community to distinguish it from the practice of quantifying luminance with respect to the minimum density (D_{min}) of the medium, which is now called "relative". In this sense, "relative colorimetry" is an oxymoron and without scientific basis: human observers do not find that papers of different reflectances constitute a colorimetric match. The practice of "relative colorimetry" has evolved because of the desire, in some applications, to maintain similar highlight detail on media of different reflectances—a desire belonging to the domain of "special" profiles and better satisfied through more sophisticated means (see below). For preserving appearance, standard (i.e., "absolute") colorimetry should be used to define the color stimuli required to produce the desired color appearance in the reference viewing environment; adaptive corrections can then be modelled and computed to adjust the colorimetry for other environments, as required, but always in reference to diffuse white.

Of course, the tonal appearance of the original cannot be duplicated on output to an arbitrary medium or device, since the latter may have a smaller dynamic range than the former. In these cases, there will be a loss of highlight or shadow detail (or both). This loss can be partially alleviated by applying a slight tonal compression at the ends of the tonal range, rather than clipping abruptly, but it must be remembered that the preservation of appearance is a goal that cannot always be attained for all tones (or colors) on all media. Generally, there will be a substantial portion of the tone scale that is common to input and output—say, quarter-tones to three-quarter-tones—where accurate reproduction can be expected.

Special Profiles for Printing and Publishing

As an example of special profiles, we may consider those required for high-quality printing in desktop publishing and graphic arts. In these applications, control over highlight detail is considered essential: users typically demand that the same highlight detail be reproduced on all output media, regardless of the reflectance of the substrate. This goal overrides that of preserving appearance: the same details will be reproduced on a dark paper as on a lighter one, and (naturally) they will *appear darker*, but this is considered preferable to simply losing some of the detail entirely. (It is usually important to preserve shadow detail as well, and these considerations may be more important than the preservation of midscale contrast, for example.)

Now, one way to achieve this goal is simply to rescale the luminances of the image according to the reflectance of the paper. This is the method implicit in "relative colorimetry". However, this means that the image as a whole will appear darker when printed on a darker paper, and the midtones and shadows will then be too dark, as would be quite apparent in a side-by-side comparison; furthermore, the shadows may "plug up" as a result of the rescaling.

For the best results, it will generally be necessary to reshape the tone curve for each output medium, so that the highlights, midtones, and shadows can best be accommodated within the available dynamic range. This has to be done with care and good (aesthetic) judgement concerning trade-offs and compromises.

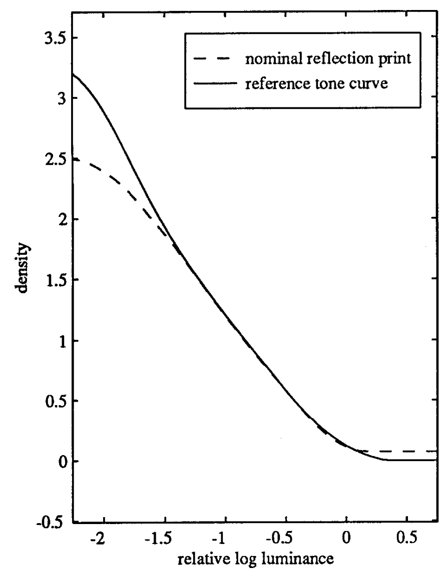

Figure 4. Reference tone curve and a nominal print pattern

In the ICC context, the concept of a "nominal print medium" may be useful here. Input profiles could reshape the tone curve of the original medium for this nominal medium, and output profiles could apply a second reshaping from the nominal medium to the tone curve of the actual output medium. By establishing a fixed convention—for instance, that the nominal medium have a reflectance of 85%—one could guarantee that the areas of blank paper were the same on all media and that all highlight detail would be preserved. Figure 4 shows a possible tone curve for the nominal medium, in comparison with the reference curve.

Monitor Profiles

A lot of the confusion in tone management has to do with image display on a CRT monitor. Part of the problem arises from the variety of viewing environments in which a monitor can be set up: sometimes the ambient room illumination is quite high, creating a "normal" surround and fairly high flare levels; at other times, the room lighting is subdued, and there is relatively little flare and a "dim" surround.

Furthermore, the viewer's state of adaptation may be difficult to analyze in some of these complex environments.

Further complexity arises from the various roles that a CRT monitor can play in a system. Most commonly, a CRT is used for the user interface to the system and application software, and tone reproduction is not critical. Sometimes the CRT is used for "soft-copy proofing" or simulation of a hard-copy reproduction; this may require that the CRT be set up next to a viewing booth for comparison. In other cases, the CRT display is the output medium itself, as for multimedia or video production. In still others, the CRT is used interactively for the creation or editing of graphics to be reproduced ultimately on other media; usually there is no explicit simulation of these media, and the CRT is regarded as a generic graphics medium bearing only a loose resemblance to the eventual output.

In the face of all these uncertainties, we recommend, for the sake of simplicity, that the *basic* display profiles establish a strictly colorimetric relation with the PCS (aside from possible chromatic-adaptation corrections); it would be speculative, pointless, and counterproductive to apply adaptive tonal corrections for an assumed viewing environment in this case. The highest luminance represented in the PCS—that of diffuse white—would be assigned to the brightest displayed white on the monitor. In this way, the monitor display can be regarded as a "window" on the PCS. To the extent that the actual viewing environment for the monitor approximates the reference environment, the monitor will accurately recreate the color appearances recorded in the PCS. And, in other environments, the basic profile may still be useful: certainly it should suffice for non-critical user-interface display, and it can also prove satisfactory for soft proofing, since the state of adaptation (although hard to define) of a viewer looking back and forth between the CRT and a viewing booth is relatively constant, so that a colorimetric match can be visually verified.

If the monitor display is the actual output medium, the appropriate viewing conditions should be established and basic output profiles should be created accordingly. For instance, for video production, the intention may be to create images for display on a calibrated television monitor in a standard environment (dim surround, D65 white point, etc.); adaptive corrections can be modelled for this environment and applied in the profile, just as would be done for any other basic output profile.

For interactive graphics, a special display profile may be required. For applications in printing and publishing, for example, a tonal reshaping may be required in the transformation from the nominal (reflection print) medium to the CRT graphics medium. Here, the brightest white of the nominal medium (corresponding, say, to the 85% reflectance of the nominal substrate) will be displayed at the brightest white of the monitor; graphic artists and other desktop users are accustomed to associating this maximum monitor white with the blank paper of the eventual output, regardless of the actual paper reflectance. A simple rescaling of luminance (from 85% to 100%) is probably sufficient here, at least until a nominal viewing environment for the graphics medium can be defined and analyzed.

Working with the ICC Profile

Ultimately what is required is an overall strategy for providing and communicating options for various aesthetic effects in cross-rendering among color-reproduction media. Preservation of color appearance may be regarded as a base-level capability; alternative effects may be regarded as special capabilities. The ICC profile format can be used to support a variety of these options. Specifically we recommend, for basic profiles: (a) that tonal appearance be preserved to the extent possible through the use of adaptive corrections; (b) that the "relative colorimetry" option *not* be used in basic output profiles; (c) that the (adaptation-corrected) D_{min} *not* be rescaled to diffuse white in basic input profiles; and (d) that basic display profiles (for CRT monitors) preserve relative luminances, associating diffuse white in the PCS with maximum displayed white. We also recommend a variety of special profiles for particular application areas, in which aesthetic tonal remapping ("personality" modification) can be applied for cross-rendering. New tags should be adopted for the profile format to indicate the presence—and perhaps the nature—of such remapping. And, in the future, we propose that the zero-density limitation of the PCS be removed.

References

1. A system of measurement and quantification of color stimuli developed and promoted by the *Commission Internationale de l'Éclairage* (CIE). See CIE Publication No. **15**.2, *Colorimetry*, 2nd ed. (1986).
2. S. Gregory, R. Poe, and D. Walker, "Communicating Color Appearance with the ICC Profile Format", *IS&T/SID Color Imaging Conference Proceedings* **2**, November 1994, pp. 170–174.
3. R. W. G. Hunt, *The Reproduction of Colour*, 4th ed., Fountain Press, 1987, Chap. 6 ("Tone Reproduction"), pp. 49–68.

☆ This paper was previously published in *IS&T's 3rd Color Imaging Conference Proc.,* p. 164 (1995).