QUANTIZATION EFFECTS IN DIGITAL IMAGING SYSTEMS

*Robert Poe** & *James Gordon**

Colors and tones in CEPS and other digital imaging systems do not vary continuously, but in small, discrete steps, because of the limited precision with which the relevant quantities are represented internally. The resulting quantization effects, such as "banding", "contouring", and "noise", can be objectionable to the viewer under certain circumstances. A conceptual analysis of these effects is developed, within the context of non-linear color-reproduction processes and human visual response. The advantage of employing perceptually uniform variables is made explicit. The analysis results in general techniques for minimizing objectionable quantization effects and determining the digital precision required at various points within a system. Typical applications are presented for soft-copy (i.e., monitor display) and hard-copy proofing.

Introduction

Quantization effects are the (generally undesirable) consequences of limited precision in the representation of a continuous quantity. These effects are quite common in digital imaging systems, such as those designed for the graphic

---

*EIKONIX Corporation (a Kodak company)

arts. The high spatial resolution required for good image quality can be achieved only by very fine sampling of the image plane, and the enormous demands placed on a system by the need to store, manage, and process the resulting millions of pixels enforce severe economies in the representation of pixel intensities. Thus, grey-scale images are rarely digitized to a precision higher than 12 bits; 8-bit, or even 6-bit, data are the norm. Even "full-color" systems typically manipulate images with only 8 bits per channel, a total of 24 (or 32) bits per picture element.

Furthermore, the precision available to represent channel intensities may be different in different parts of the system (e.g., disk storage, computer memory, frame store, hardware look-up tables, output peripherals). If insufficient precision is available at any stage, image quality can be degraded by quantization effects. When n bits per channel are used to represent each pixel, the corresponding channel intensity (which, in theory, is a continuous variable) is quantized to 2^n discrete levels. In general, this means that there will be a discrepancy between the true channel intensity at each pixel and the nearest discrete level in the digital representation. If these discrepancies are distributed at random, they are a source of statistical error, or "noise", in the image.

Under certain circumstances, the distribution is *not* random and can give rise to geometric artifacts. Tones and colors vary discontinuously in the picture, stepping abruptly from one discrete level to another. In regions of low intensity gradient, these steps occur along distinct contours, delineating relatively broad bands of constant intensity, and can be highly visible. This phenomenon, which is variously referred to as *staircasing*, *banding*, and *contouring*, is a serious defect in image quality. Higher precision (i.e., increasing n) results in closer spacing of the quantization levels and, therefore, to a decrease in the noticeability of noise and staircasing.

System designers often have to make difficult decisions concerning the precision with which image data will be represented at different stages of processing, with potentially serious impact on system cost, system throughput, and image quality. A clear understanding of quantization effects and their propagation through a system is a prerequisite for intelligent design. Unfortunately, the analysis required to determine the trade-offs is complicated by the overall complexity of the processing hardware and software and by the presence of non-linear components in the system.

In this paper we present a simple conceptual approach to the analysis of quantization effects in digital imaging systems, based on a model of the system as a sequence of non-linear transformations. The model is extended to include the human visual response as a necessary component.

It then becomes possible to analyze the precision requirements that are imposed by image-quality goals and to answer such questions as "How many bits of precision are needed to gamma-correct a monitor adequately?" The utility of perceptually uniform variables in the implementation (as well as the design) of a system becomes evident in the analysis.

The discussion begins with the detailed study of a relatively simple, but non-trivial, example: a monochrome display system. The intuitions gained from this study are then generalized and applied to color-reproduction systems, both soft-copy and hard-copy. The discussion concludes with a brief summary.

A Case Study: Monochrome Display

We consider, for our initial investigation, a system in which picture data have been digitized in a single channel and are to be displayed as grey-scale imagery on a monochrome monitor.

The grey-scale image

For the purposes of this study we specialize to a photometrically calibrated display, one in which the energy of visible light emanating from points on the monitor screen can be accurately and intentionally controlled. Therefore, image intensity (or grey scale) will be identified with the energy of visible light, or *luminance*. Luminance is, by definition, the visible component of light energy, i.e., the integral of spectral radiance weighted by the human photopic response, and can be measured by a photometer at the monitor screen for the purpose of calibration.

We will, therefore, assume that a *luminance image* has been prepared in some digital representation within the system. The task, then, is to display this image accurately on the monitor without objectionable quantization effects.

We will use the symbol Y for the luminance measured at the monitor and the symbol Y_d for the internal (digital) representation of the luminance image. The desired goal, then, is

$$Y = Y_d \ .$$

(This goal can be achieved only up to a certain tolerance.)

It will be convenient to consider these quantities, and all other "intensity" variables in this discussion, as being normalized to the range [0.0, 1.0]; digital representations, such as Y_d, will then consist of fractional bits only.

"Gamma" response and "gamma correction"

Typical cathode-ray-tube (CRT) monitors respond in a highly non-linear fashion to the signals impressed upon them. That is, the luminances measured at the screen are not proportional to the voltages applied to the electron guns. Traditionally, the CRT response curve has been described as a power law, and (by analogy with the exposure response of photographic film) the power is called *gamma*. In fact, the power law is not actually a very good fit [Cowan 1983]. In addition, the digital-to-analog converter (DAC) may not respond with perfect linearity. Therefore, we will speak more generally of the *gamma response* of a display (including the effects of the DAC and the CRT), written $G(s_d)$ (where s_d represents the digital display signal input to the DAC), without assuming a power law. Thus, we write

$$Y = G(s_d) .$$

A representative gamma-response curve is shown in Figure 1; typically, the display is relatively

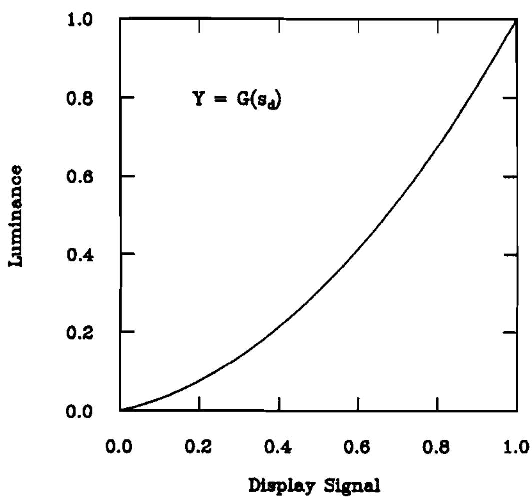

Figure 1: Typical gamma response

insensitive at low signal levels (i.e., the shadows and other dark regions of a picture) and increases its sensitivity dramatically as the signal level increases.

Obviously, if we wish to achieve a photometrically accurate display ($Y = Y_d$), we must introduce some compensation for the gamma response. This mechanism is called *gamma correction* and is applied to the digital luminance data (for instance, in a look-up table) prior to analog conversion. We write

$$s_d = G^{-1}(Y_d) ,$$

which emphasizes that the gamma-correction transformation, G^{-1}, must be the inverse of the gamma-response transformation. A typical gamma-correction curve is shown in Figure 2; it

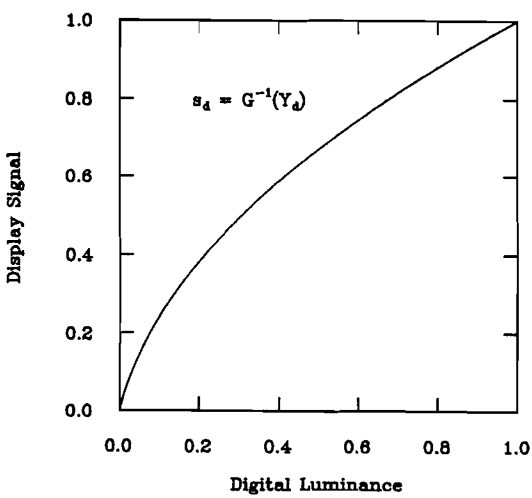

Figure 2: Gamma correction

must stretch the contrast in the shadows (at the expense of the highlights) in order to compensate for the response of the CRT.

It should be noted in passing that non-linear transformations, such as gamma correction and gamma response, have a strong influence on precision requirements, since they modify the relative spacing of quantization levels. For instance, if Y_d is quantized to a number of equally spaced levels, these will be mapped to *unequally* spaced levels by gamma correction; some of those levels, particularly in the highlights, may be so closely spaced that they are indistinguishable in the digital representation of s_d; if so, some sets of neighboring levels will collapse together into single levels.

Therefore, although we can write

$$Y = G(s_d) = G(G^{-1}(Y_d)) = Y_d$$

as an overall description of the display transformation, this equation may not hold precisely in every case.

Psychometric response

Before we can draw conclusions regarding the precision required for Y_d and s_d, we must take into account the response of the human visual system to the luminances in the displayed image. It is well known that this response is non-linear and is much more sensitive in the shadows than in the highlights.

Human visual response to luminance variations is a complicated phenomenon, depending on the state of adaptation of the observer, the spatial content of the scene, and other factors. For our purposes, the phenomenon can be treated by defining a psychometric variable, called *lightness*, which is a monotonically increasing function of luminance and which represents an approximately uniform perceptual scale. The more accurate this function is, and the more specifically it is adjusted to the actual viewing conditions appropriate to the imaging system, the better the results will be; however, even a simple, approximate model will be of great help in analyzing the precision issues involved.

We will therefore assume that the human visual response is represented by a transformation, H, from luminance, Y, to lightness, L:

$$L = H(Y) \ .$$

For illustrative purposes, we will model the transformation H on the variable L^* introduced by the *Commission Internationale de l'Eclairage* (CIE) under the designation *CIE 1976 psychometric lightness* [CIE 1978]. To be consistent with our normalization convention, we will put

$$H(Y) = L^*/100 \ .$$

This choice of lightness variable has the virtue of having a simple mathematical expression:

$$H(Y) = \begin{cases} 9.033Y, & 0 \leq Y \leq 0.008856, \\ 1.16Y^{1/3} - 0.16, & 0.008856 < Y \leq 1. \end{cases}$$

(where we have implicitly assumed that the observer is adapted to a white reference with $Y = 1.0$), as well as the virtue of being a widely-used international standard. Figure 3 shows the form of $H(Y)$.

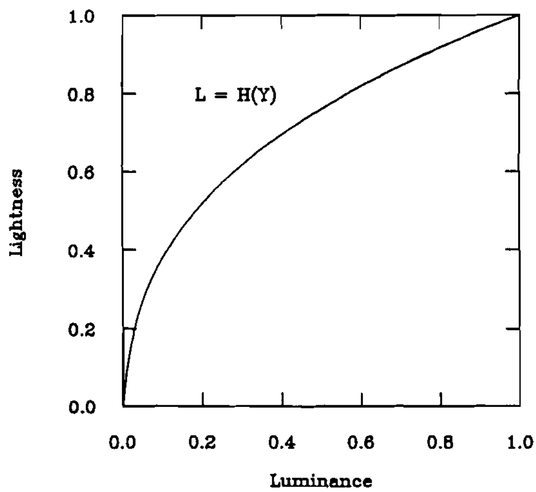

Figure 3: Model of visual response

The display system can thus be regarded as a sequence of non-linear transformations, from the digital representation of luminance, to a digital gamma-corrected signal (and its analog equivalent), to measured luminances at the screen, to psychometric lightness as perceived by the observer. This conceptual view is represented in

Figure 4.

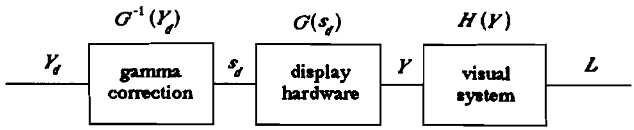

Figure 4: Monochrome display system

We are now in a position to establish precision requirements for the system.

Reverse system analysis

The best way to proceed with the analysis is to fix a goal for image quality in terms of perceived lightness and to work backwards through the system to determine the precision requirements.

Since lightness is measured on a uniform perceptual scale, the image-quality goal can be expressed simply as an upper bound (or tolerance), h, on the spacing between consecutive lightness levels:

$$dL < h \ .$$

The assumption then is that, if this goal is met, quantization effects will be held down to an acceptable level.

This inequality can be immediately translated into a tolerance in the spacing of luminance levels; it is merely necessary to introduce the first derivative H' of the transformation H:

$$dY = \frac{dY}{dL}\, dL = \frac{dL}{H'(Y)} < \frac{h}{H'(Y)} .$$

This is a *non-uniform* tolerance and is strictest when H' is a maximum, viz., in the shadows; there, the level spacing is constrained by:

$$dY < h/H'(0) .$$

The tolerance on the spacing of s_d levels can be determined by defining the combined transformation F:

$$\begin{aligned} L = H(Y) &= H(G(s_d)) \\ &= F(s_d) ; \end{aligned}$$

with this definition, we can follow a similar line of reasoning to deduce

$$ds_d < h/F'(s_d) .$$

This is also a non-uniform tolerance. Although the gamma response G partially compensates for the psychometric response H, the combined transformation F still reaches its steepest slope in the shadows. Therefore, the closest tolerance is:

$$ds_d < h/F'(0) = \frac{h}{G'(0)H'(0)} .$$

Now, this expression establishes a requirement on the precision of the gamma-corrected input to the

DAC: since the inherent precision of a digital representation is uniform over the range of values represented, it must be adequate to resolve the closest levels required. Thus, the number of bits, n, used to represent s_d must satisfy the condition:

$$1/2^n < \frac{h}{G'(0)H'(0)} .$$

Similarly, the non-uniform tolerance on the spacing of levels in Y_d can be determined by working backwards one more stage, through the gamma-correction transformation:

$$\begin{aligned} dY_d &< (h/F')/(G^{-1})' = hG'/F' \\ &< h/H'(Y_d), \end{aligned}$$

or, in the shadows,

$$dY_d < h/H'(0) .$$

(This, of course, is the same constraint obtained above for dY: if gamma correction is accurate, this is just what should be expected.) As in the case of s_d, this constraint on the spacing of levels in the shadows then establishes the precision required in the digital representation of luminance.

It is instructive to insert some actual numbers into these expressions. Suppose that it has been decided that a level spacing $dL < 0.01$

provides acceptable image quality. We thus set $h = 0.01$ as the system goal. Since $H'(0) = 9.033$, we find that we must control luminance, at least in the shadows, to a tolerance:

$$dY < 0.01/9.033 = 0.0011 \; .$$

Suppose further that our monitor calibration has determined that $G'(0) = 0.23$. We then find

$$ds_d < 0.01/(0.23 \times 9.033) = 0.0048 \; .$$

Since 8-bit precision gives us a level spacing of $1/256 = 0.0039$, our gamma-correction table need be only one byte wide. (However, seven bits would not suffice, since 1/128 exceeds 0.0048.)

On the other hand, we find that the constraint on the luminance image ($dY_d < 0.0011$) requires at least 10 bits of precision in the representation of Y_d. If, indeed, a 10-bit representation is employed, there will be 1024 quantization levels available. The gamma-correction table (10 bits in, 8 bits out) can then be accommodated in 1 kilobyte of storage.

Image encoding

At this stage of the analysis, it becomes clear that considerable waste is involved in the storage of the luminance image. In practice, a 16-bit word (for each pixel) will probably be needed to provide the 10 bits of required precision. Furthermore, by the time the 1024 available levels are gamma-corrected, displayed, and observed, most of them will have been collapsed together and compressed into the upper part of the lightness range, with a level spacing far below the established tolerance of 0.01. Only in the shadows is the 10-bit quantization really required.

Now, in theory, 101 quantization levels, *equally spaced in lightness*, would satisfy the given tolerance ($h = 0.01$). This suggests that, for storage purposes, it is advantageous to encode a digital *lightness image*, rather than the luminance image. We propose, therefore, to introduce another transformation into the system, for digital encoding. Let L_d symbolize the digital representation of lightness. Then the encoding transformation is

$$L_d = H(Y_d) \ ,$$

and

$$Y_d = H^{-1}(L_d)$$

describes the decoding of this representation to obtain (digital) luminance. Regarded from one end to the other, the system now acquires a symmetrical form:

$$\begin{aligned} L &= H(G(G^{-1}(H^{-1}(L_d)))) \\ &= H(H^{-1}(L_d)) \\ &= L_d \ . \end{aligned}$$

(See Figure 5.)

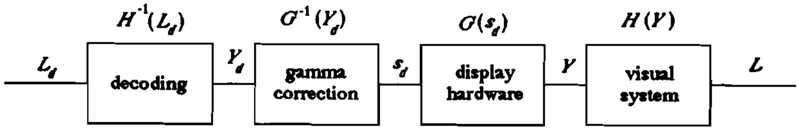

Figure 5: Monochrome display with encoding

Clearly, the level spacing of the encoded image is governed by the condition:

$$dL_d < dY_d/(H^{-1})' = (h/H')/(H^{-1})'$$
$$= hH'/H' = h \ .$$

Therefore, 7 bits will suffice to meet the given tolerance: $1/128 = 0.0078 < 0.01$. In practice, 8 bits would be used, and the total image storage required would be reduced by a factor of 2.

For the purposes of display, it is not necessary to represent the luminance image explicitly in digital form. The decoding transformation can be combined with gamma correction into a single look-up in a table with 256 8-bit entries. This technique allows a factor-of-4 saving in the size of the look-up table.

Of course, whenever luminance must be manipulated directly in the system (as, for instance, in certain kinds of image processing), the 8-bit lightness image must be decoded, with a consequent increase in the precision required (at least 10 bits). However, these manipulations or computations can usually be designed to operate on small segments of the image at a time.

General Analytic Approach

Let us abstract the essential elements of this case study and express them in more general form:

(1) The system is modelled as a sequence of transformations, from the internal digital representation of the stored image, through successive digital and analog representations, to the physical quantities that can be measured for calibration purposes and the psychometric quantities that correlate (approximately) with human visual response.

(2) A tolerance goal is specified in terms of psychometric quantities on a uniform perceptual scale.

(3) This tolerance is propagated back through the successive transformations to derive upper bounds for the spacing of quantization levels in all the representations. The derivatives of the transformation functions are the relevant factors in this calculation. In general, the tolerances will be non-uniform.

(4) In the digital representations, precision requirements are then established by the need to resolve the most closely spaced levels.

(5) The most economical digital representation of the stored image is an encoding that is linear in the same psychometric quantities that were

used to specify the system tolerance goal.

This now constitutes a general approach that can be applied to more complicated problems, such as the precision analysis of a color electronic pre-press system (CEPS).

Color Display

In a CEPS, the image displayed on the workstation monitor can provide invaluable feedback to the operator on the results to be anticipated on press. In order to function as such a "soft-copy proofing" device, the monitor must accurately render the colors and tones of the stored digital image. A necessary condition on the display system (analogous to photometric calibration in the example discussed above), therefore, is *colorimetric calibration*.

One way in which this calibration can be implemented is by colorimetrically determining the tristimulus values at points on the monitor screen; for instance, these can be expressed as the vector $\mathbf{T} = (X, Y, Z)$, where X, Y, and Z are the *CIE 1931 tristimulus values*. A digital *tristimulus image* can be computed or prepared within the system with values $\mathbf{T}_d = (X_d, Y_d, Z_d)$ for each picture element. (In general, the tristimulus image will incorporate corrections for the limited gamut of the press or hard-copy proofing system [Gordon et al., 1987], as well as specific editorial changes invoked by the operator.) The calibration goal, then, is

$$\mathbf{T} = \mathbf{T}_d \ .$$

An RGB monitor is controlled by 3 digital signals, written as the vector $\mathbf{s}_d = (r_d, g_d, b_d)$. Each of these is input to a DAC; the output analog signals are then applied to the red, green, and blue guns of the color CRT. The color response of the display can then be written as

$$\mathbf{T} = \mathbf{C}(\mathbf{s}_d) \ ,$$

where the transformation $\mathbf{C}$ can be further decomposed into the gamma response G of the individual channels and a *linear* transformation $\mathbf{M}$ from the RGB primaries to tristimulus values:

$$\mathbf{C}(\mathbf{s}_d) = \mathbf{M}(G(r_d), G(g_d), G(b_d)) \ .$$

The result of calibration is a determination of G and $\mathbf{M}$ [Cowan 1983] and, therefore, $\mathbf{C}$, so that the gamma-corrected display signals can be computed from the digital tristimulus image by applying $\mathbf{C}^{-1}$:

$$\mathbf{s}_d = \mathbf{C}^{-1}(\mathbf{T}_d) \ .$$

This computation will result in a colorimetrically accurate display--

$$\mathbf{T} = \mathbf{C}(\mathbf{s}_d) = \mathbf{C}(\mathbf{C}^{-1}(\mathbf{T}_d)) = \mathbf{T}_d$$

--, except for quantization effects.

In order to assess the human visual response to the displayed colors, we need to introduce psychometric quantities analogous to the lightness variable employed in the monochrome case. Several attempts have been made, over the years, to define a perceptually "uniform color space" [Tonnquist 1986]. One convenient choice is the CIELUV system [CIE 1978], in which colors are represented by three quantities (symbolized L^*, u^*, and v^*) defined in terms of the tristimulus values $\mathbf{T}$. (The color of the nominal white reference must also be specified.) We will renormalize these quantities by defining the vector

$$\begin{aligned}\mathbf{E} &= 1/100\ (L^*,\ u^*,\ v^*)\\ &= \mathbf{H}(\mathbf{T})\ ,\end{aligned}$$

where the transformation $\mathbf{H}$ is non-linear, but fixed.

In order to suppress contouring and other objectionable quantization effects, a uniform tolerance can now be placed on the spacing of adjacent levels in perceived color:

$$|d\mathbf{E}| < h\ .$$

Since this relation must hold between adjacent levels in any direction in color space, we have:

$$dL^* < 100h\ ,\ du^* < 100h\ ,\ dv^* < 100h\ .$$

These relations can then be propagated backwards through the system to derive (non-uniform) tolerances on **T**, s_d, and T_d. This calculation involves computing the partial derivatives of the transformations **H**, **C**, and C^{-1}, respectively, and combining them appropriately. The analysis is considerably more complex than in the monochrome case, and details will not be given here. However, just as in the single-channel case, constraints can be derived for the digital representations (s_d and T_d) that lead to specifications of their required precision.

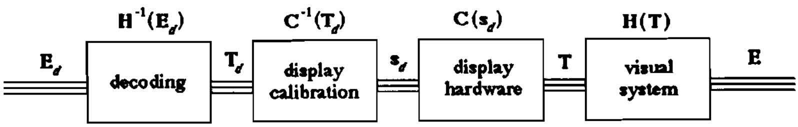

Figure 6: Color display system

Also by analogy with the monochrome example, an economical encoding of the digital image can be developed in terms of the psychometric variables involved. This representation, E_d, is based on the color-space variables employed in setting the tolerances on quantization effects--for example, the CIELUV coordinates. The system is then symmetrical (see Figure 6):

$$E = H(C(C^{-1}(H^{-1}(E_d)))) = H(H^{-1}(E_d))$$
$$= E_d ,$$

and the (uniformly spaced) quantization levels in $\mathbf{E}_d$ are used most efficiently in reducing the (uniformly perceptible) quantization effects in $\mathbf{E}$. For display purposes, $\mathbf{T}_d$ and $\mathbf{s}_d$ can be computed on the fly, as needed.

Color Proofing

A similar approach can be taken for the analysis of hard-copy halftone proofing systems, such as Kodak's Signature, Dupont's Cromalin, 3M's Matchprint, etc. Colorimetric calibration can be achieved by measuring tristimulus values $\mathbf{T}$ on the proof (under a suitable illuminant) in order to establish agreement with the digital tristimulus image $\mathbf{T}_d$. (The latter must be confined to the restricted color gamut reproducible in the proofing process [Gordon et al., 1987].)

Proof preparation is controlled through a set of digital inking values for the yellow, magenta, cyan, and black separations. These can be represented as a 4-vector:

$$\mathbf{I}_d = (y_d, m_d, c_d, k_d) \ .$$

These values are the inputs to a dot screener and determine the individual dot percentages on the separation films. When the films are written and assembled, the effects of the screen-generation algorithm, the ink densities and chromaticities, dot gain, interlayer interactions, etc., combine non-linearly in producing the colors that will be measured on proof. This process can be summarized as a transformation, P:

$$T = P(I_d) .$$

Calibration then consists of determining this transformation and inverting it in that part of the system which generates the inking values:

$$I_d = P^{-1}(T_d) ,$$

so that

$$T = P(I_d) = P(P^{-1}(T_d)) = T_d .$$

This problem is considerably more complicated than the corresponding problem of calibrating a color display, but even an approximate solution will suffice for the estimation of quantization effects.

As in the color-display case discussed above, the human visual response can be modelled by transforming to a uniform color space, such as CIELUV. The same transformation, H, applies to this case, and the same tolerance, h, can be used for the spacing of levels in color space. This tolerance can be propagated back through the successive transformations H, P, and P^{-1}, to determine permitted level spacings in T, I_d, and T_d, respectively. For determining the precision requirements for I_d, it is necessary to estimate the partial derivatives of the transformation

$H(P(I_d))$. On the other hand, the precision requirements for T_d will be the same as for color display, since calibration ensures that the perceived colors of the discrete levels are independent of the rendering process:

$$H(P(P^{-1}(T_d))) = H(C(C^{-1}(T_d))) \\ = H(T_d) = E \, ,$$

as long as sufficient precision is available in the intermediate representations (I_d and s_d, respectively).

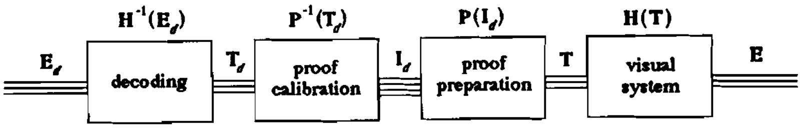

Figure 7: Color proofing system

As in the case of color display, the most efficient and economical digital encoding of the image data is based on color-space coordinates, E_d (see Figure 7). For output to the screener and film writer, I_d can be generated directly from E_d; T_d does not have to be represented explicitly.

Conclusions

The uses of an approximately uniform color space, therefore, go beyond the analysis, design, and evaluation of a system: they can even include aspects of the system implementation. The most efficient way to suppress unwanted quantization effects, within the limited precision available in a digital imaging system, is to represent the image data (at the fundamental level) in psychometric terms. In a color system, this can be achieved by encoding the image in perceptually uniform color coordinates. This encoded image can then be stored in relatively compact form. Other digital quantities (tristimulus values, inkings, etc.) can be derived, as needed, from the color-space encoding; these representations, which generally require higher precision, can usually be generated "on the fly", in the course of data processing, through the use of look-up tables or specific hardware support.

In any case, the analytic method described above provides a consistent and straightforward approach to determining the precision requirements of digital image data, in various representations, throughout a complex system containing non-linear components. Within this approach, a detailed treatment of all the system complexities is not needed. It is only necessary to model the individual transformations to a rough approximation, so that their first derivatives can be estimated. The propagation of quantization levels through the system can then be analyzed, and the effect on the human visual system can be assessed with sufficient accuracy to guide design choices.

Acknowledgements

The authors would like to thank Richard Holub, Gary Newman, and Michael Schwartz (all of EIKONIX Corporation) for helpful conversations.

Literature Cited

CIE (*Commission Internationale de l'Eclairage*)
1978. "Recommendations on Uniform Color Spaces, Color-Difference Equations, Psychometric Color Terms" (*Bureau Central de la CIE*, Paris), Supplement No. 2 to CIE Publication No. 15.

Cowan, William B.
1983. "An Inexpensive Scheme for Calibration of a Colour Monitor in Terms of CIE Standard Coordinates", *Computer Graphics*, Vol. 17, No. 3, pp. 315-321.

Gordon, J., Holub, R., and Poe, R.
1987. "On the Rendition of Unprintable Colors", TAGA *Proceedings* (this volume).

Tonnquist, Gunnar
1986. "Philosophy of Perceptive Color Order Systems", *Color Research and Application*, Vol. 11, No. 1, pp. 51-55.