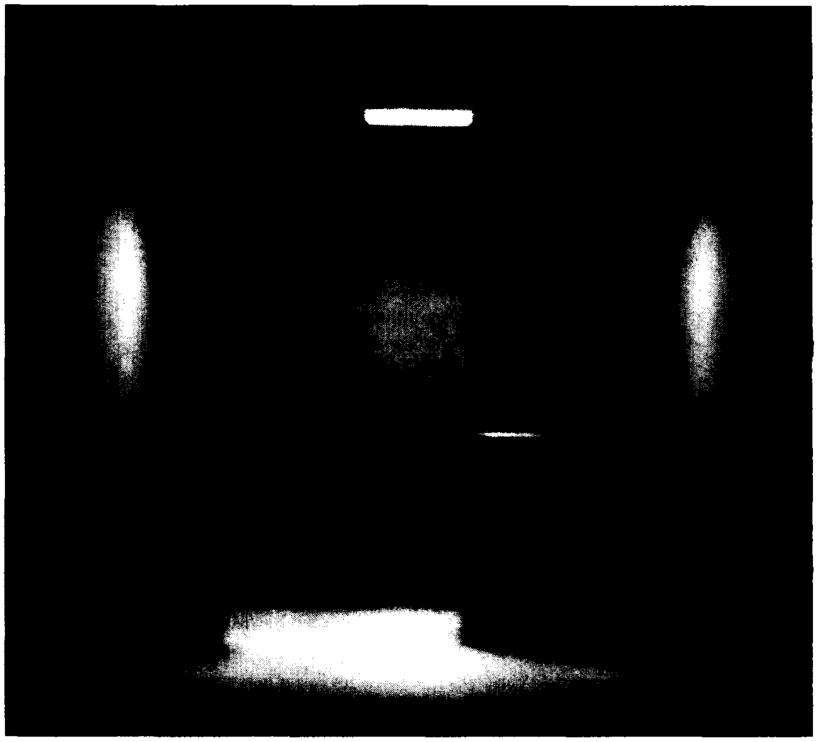
Tone Reproduction for Realistic Images

Jack Tumblin and Holly Rushmeier
Georgia Institute of Technology



Display methods should compensate for all light-dependent changes in the way we see. Better tone reproduction for computer graphics will help solve display range problems.

Radiosity and other global illumination methods for image synthesis calculate the "real world" radiance values of a scene instead of the display radiance values that will represent them. Though radiosity and ray tracing methods can compute extremely accurate and wide-ranging scene radiances, modern display devices emit light only in a tiny fixed range. The radiances must be converted, but ad-hoc conversions cause serious errors and give little assurance that the evoked visual sensations are truly equivalent. Common methods of conversion can fail spectacularly for extreme lighting conditions because they ignore light-dependent changes in the way we see.

These conversions deserve attention because better ones are easily implemented by computer. Sensation-preserving conversions for display are already known in photography, printing, and television as *tone reproduction* methods. Computer graphics workers can apply the existing photographic methods, but may also extend them to include more complex and subtle effects of human vision using the published findings of vision researchers. We will demonstrate how to construct a sensation-preserving display converter, or *tone reproduction operator*, for monochrome images. See the sidebar for an explanation of why tone reproduction matters.

Background

Radiance values appear to be exactly as we see them, but humans are very poor judges of absolute radiance.[3] The eye's sensitivity to light varies with wavelength, even within the visible band. More importantly, the visual system is far better at sensing spatial and temporal changes in radiance than its absolute value. Humans apparently form perceptions of light strength from judgments of these changes. Good tone reproduction needs at least three different measures of light: the absolute magnitude of light energy in the visible band, the magnitude of light sensible to the human eye, and the perceived strength of the light. We use radiance, luminance, and brightness respectively for these three measures.

Radiance is the quantity of electromagnetic energy per unit projected area, time, and solid angle. The spectral radiance is the radiance per unit wavelength at a particular wavelength. In general, global illumination methods such as radiosity or path tracing compute spectral radiance. Radiance can be measured with a radiometer.

Luminance is the physical measure of light sensible by a "standard" human eye. The luminance is obtained by integrating the spectral radiance weighted by the eye's luminous efficiency

0272-17-16/93/1100-0042$03.00 © 1993 IEEE

Why tone reproduction matters

Many computer graphics image synthesis algorithms can't tell the difference between night and day—differences that are obvious to any human eye. Figure A depicts a room of uniformly diffuse surfaces lit by a single extended light source with adjustable emitted power. Suppose we choose two settings, one at an emitted power equal to that of two squashed fireflies and another with the power of an anti-aircraft searchlight. The global illumination solution is linear in source radiance; that is, results for any two light source strengths are directly proportional. Accordingly, the image of the room with firefly lighting is identical to the image of the room lit by a searchlight, except for a scale factor of about 10^{11}.

How should such firefly- and searchlight-strength room images be displayed? One widely used method normalizes all computed scene radiances by the value of the strongest nonemitting surface in the image, then these normalized values drive a gamma-corrected display device. This method, used for Figure A, gives the appearance of pleasant office lighting. This ad-hoc normalization removes the scale factor that distinguishes the firefly- and searchlight-powered room images; thus, Figure A is the displayed result for both of them, absurd as it seems.

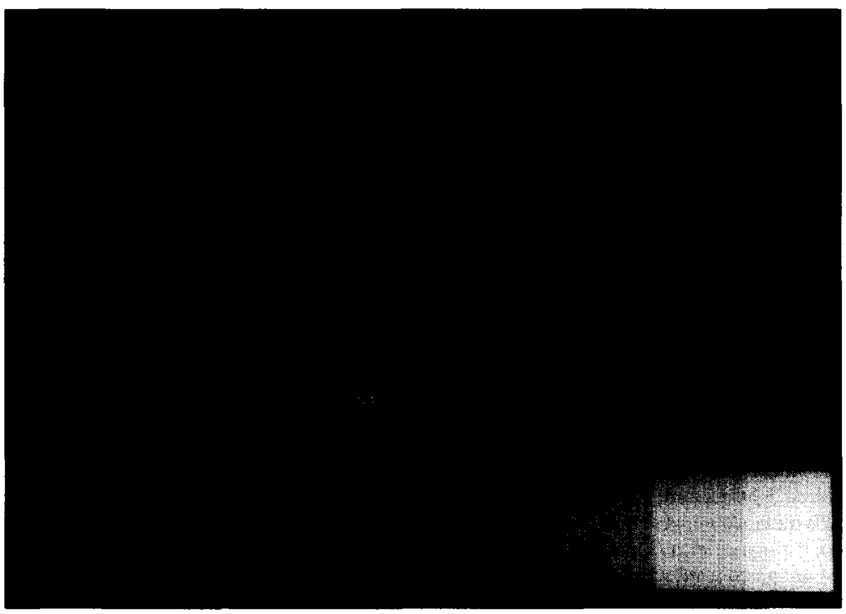
Human observers see these rooms quite differently. To a dark-adapted viewer, the firefly-powered room appears as little more than a very dim area, while an observer of the searchlight-illuminated room might see only harsh shadows against glaring white. Neither observer would recognize Figure A—it is wholly inaccurate. Tone reproduction operators correct these deficiencies.

Figure A is also a dubious result for moderate illumination. As discussed in Meyer et al.,[1] the rendering process has two steps: the calculation of scene radiances, and the use of principles of perception to map these radiances to the display device. To be treated scientifically, the rendering process must be subject to verification at each stage. A tremendous amount of recent research deals with rigorous, first principle methods for the first step. Researchers have studied color transformations in the second step as well.[2] However, the mapping of real world radiances to the display device (tone reproduction) has been generally ignored.

Fine image details in Figure A were very expensive to compute, yet the human observers in the firefly- and searchlight-illuminated rooms cannot see them. A good tone reproduction operator can predict visibility thresholds for errors in each region of the displayed image and avoid these wasted efforts. By setting the precision of the computed radiosity solution, tone reproduction addresses one of the basic questions remaining in global illumination—how accurate does a solution have to be?

Figure A lacks the strongly light-dependent effects characteristic of human vision needed for good tone reproduction. The subjective accuracy of the displayed image can be improved by including models of the complex, dramatic changes to human vision that occur over the firefly-to-searchlight range. People have strong differences in sensitivity, acuity, contrast perception, and color sensitivity; high contrast effects such as glare, dazzle, afterimages, color washout, and diffraction; spatial effects such as Mach banding and hyperacuity; and temporal effects such as changes in adaptation, thermal noise, and motion blur. All of these effects contribute to realistic images, but useful results are demonstrated here using just a few of them. We wish to stress the importance of explicit tone reproduction operators and encourage their use, rather than to champion any particular expression.

Figure A. Display of radiosity solution using ad-hoc scale factor.



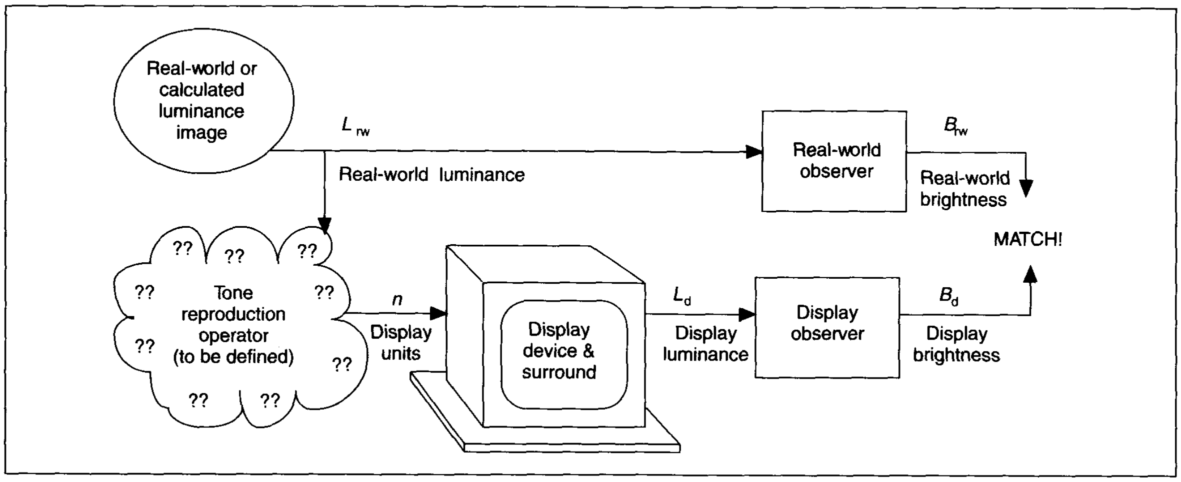


Figure 1. The tone reproduction problem: What operator will cause a close match between real-world and display brightness sensations?

curve over the entire spectrum. Luminance can be measured by a photometer.

Brightness is the magnitude of the subjective sensation produced by light. Brightness cannot be measured by a meter and is known only by indirect psychophysical experiments. Here we use "brils," a linear scale of absolute brightness described by Stevens,[4] where 1 bril equals the sensation induced in a fully dark-adapted eye by a one-second exposure to a five-degree white target of one microlambert (0.0314 cd/m^2) luminance on a completely black background. Some workers disagree with Stevens' generalization of his results, but his extensive suprathreshold vision experiments provide convenient data for a simple tone reproduction operator.

Tone reproduction is only necessary because the eye's input range dwarfs the luminous output range of existing displays. Thus, direct reproduction is almost never possible. For example, most direct-view CRTs have a luminance range between 1 to 100 cd/m^2, but shadows in a starlit forest are just barely visible at about 10^{-5} cd/m^2, while a glinting snow bank might emit 10^5 cd/m^2.

In this article, we consider tone reproduction for gray-scale images only. For a gray environment, the spectral radiance is uniform for all wavelengths, and the luminance is just a constant times the uniform spectral radiance. Each input value to a gray display device causes a unique output luminance, unlike a color display in which two different RGB triplets may have the same luminance. For gray images, then, a method for accurate tone reproduction is completely defined by describing how the simulated real-world luminance is mapped to the luminance on the display device.

A general framework

We have built a general framework to define tone reproduction from the response of two observer models and a display system model, as shown in Figure 1. An observer model is a mathematical model of the human visual system that includes all desired light-dependent visual effects while converting real-world luminance images to perceived brightness images. We denote luminance with L, and brightness with B. "Real world" values are given a subscript "rw," while "display" values are given a subscript "d."

We use two observer models. The *real world observer* views the desired luminance image and corresponds to a human visitor to the room in Figure A, while the *display observer* views the luminance values of the display device. The display model converts display input values n to viewed luminance values L_d. It includes effects of ambient room light and CRT performance. The tone reproduction operator converts real-world luminances L_{rw} to display input values, where n is chosen to closely match the two observer model outputs B_{rw} and B_d.

Expanding Figure 1 as shown in Figure 2 reveals the tone reproduction operator: It is the concatenation of the real-world observer, the inverse of the display observer, and the inverse of the display model. If each is known and robust, then a tone reproduction operator follows easily.

Film and television methods

Television and film systems have roughly sigmoid responses to light when plotted on log/log axes, as in Figure 3, and both use similar equations and nomenclature. At the camera, each translates real-world luminances L_{rw} to display luminances L_d in two steps: First the real world luminances are encoded (as transparency T for film and voltage V for television), then decoded to display luminances L_d. Combined encoding and decoding equations describe each system's tone reproduction. The center region of Figure 3 is approximated as a line segment with three parameters: "latitude" is the segment's width, "i" is its horizontal offset or film speed, and "gamma" (γ) is its slope. These three parameters and "$D \log E$" or "H-D" plots (named for Hurter and Driffield, who devised them in 1890), similar to Figure 3, are widely used to describe film performance because all are chemically or optically controlled.[5]

Film encodes real-world luminances within its latitude as transparency values according to

$$T = a_1 L_{rw}^{\gamma} \qquad \textbf{(1)}$$

where

T = film transparency (0 for perfect opacity, 1 for perfect transparency)

a_1 = a constant factor depending on the film speed "i" and on choice of exposure time, lens, and camera aperture

Suppose the film is viewed as a transparency placed on a light

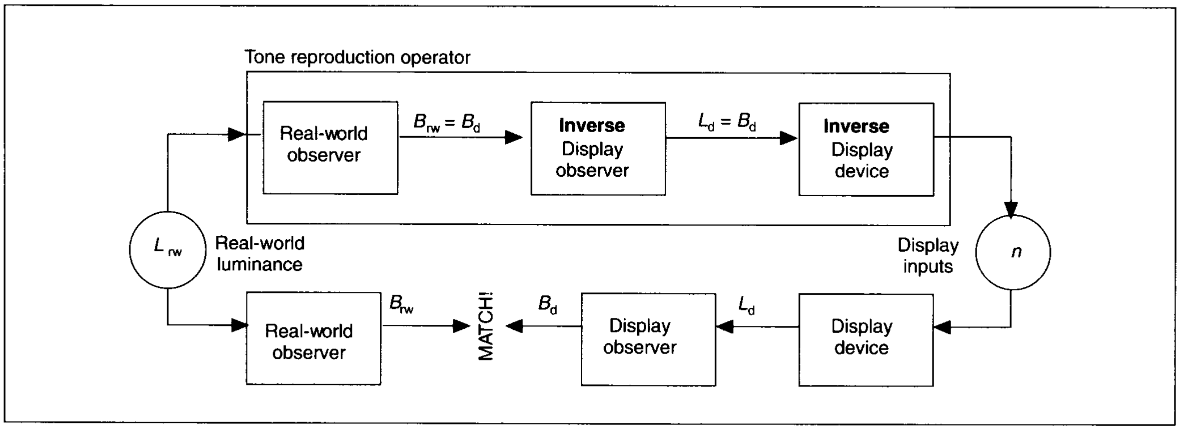


Figure 2. A simple solution: Make a tone reproduction operator by concatenating a real-world observer model, an inverse display observer model, and an inverse display device model.

table that has luminance L_s. The luminance is decoded from the film according to

$$L_d = L_s T \tag{2}$$

Thus film converts real-world luminance L_{rw} to displayed luminance L_d as

$$L_d = a_1 L_s L_{rw}^{\gamma} \tag{3}$$

Since a_1 and γ are set by the photographer's choice of film, lenses, exposure time, aperture, and darkroom processes, human judgments will strongly affect the resulting image.

Television system response also resembles Figure 3. Over the camera's latitude, real-world luminances L_{rw} are encoded to V, then decoded to display luminances L_d as

$$V = (a_2 L_{rw})^{\gamma_{cam}} \tag{4}$$

$$L_d = a_3 V^{\gamma_{crt}} \tag{5}$$

where

- V = normalized video signal, with $0 < V < 1$.
- a_2 = constant set by camera sensitivity, lens, and aperture.
- γ_{cam} = "gamma correction" to compensate for excess receiver CRT gamma and to reduce visibility of transmission noise at the receiver. Typically equal to 1/2.2 (standardized for television CRTs).
- γ_{crt} = the standardized television CRT gamma, in the range 2.8 to 3.0.
- a_3 = maximum CRT luminance; typically $48 \leq a_3 \leq 127$ cd/m^2.[6]

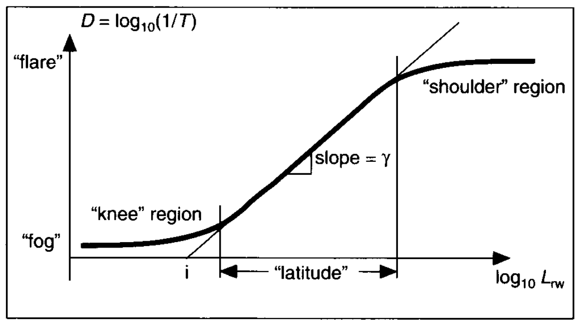
Combined encoding and decoding yields Equation 6, which is identical to Equation 3 if constants a_2, a_3 and γ_{cam}, γ_{crt} are combined:

$$L_d = a_3 (a_2 L_{rw})^{\gamma_{cam}\gamma_{crt}} \tag{6}$$

Human observer effects in film and television

Linear reproduction, where a's are nominal and γ's equal 1.0, are poor choices for good tone reproduction. The best choices depend heavily on the strength of the displays and their surrounding room lighting: They vary according to light-dependent changes in the way we see. Subjective image quality experiments for film[5] and television[7] show ideal γ or $\gamma_{cam}\gamma_{crt}$ values fall between about 1.1 and 1.5. Television viewers prefer decreasing $\gamma_{cam}\gamma_{crt}$ as surroundings brighten, with 1.5 preferred for darkness, 1.2 for dim light (14 cd/m^2), and approaching 1.0 for bright light. Similarly, reflection prints (film mounted on white paper backing) with $\gamma = 1.0$ are overwhelmingly rejected as too dark and lacking in contrast. However, when viewed as a strongly backlit transparency in a dimly lit room, the same film is preferred over any other reflection print.

Figure 3. Hurter and Driffield's "characteristic curve" for photographic film.



Components of a tone-reproduction operator

Simply applying film tone-reproduction operators to synthetic images might be inadequate. Film and television reproduction is primarily concerned with the light-dependent effects on the display observer alone; real-world observer effects are left to the judgment of the camera operator, who may change exposure settings or even adjust scene radiances to improve the displayed result. For example, the practice of "day for night shooting" uses optical filters and high-contrast, low-angle lighting to achieve the appearance of night while filming motion pictures in daylight. However, a practical operator for accurate

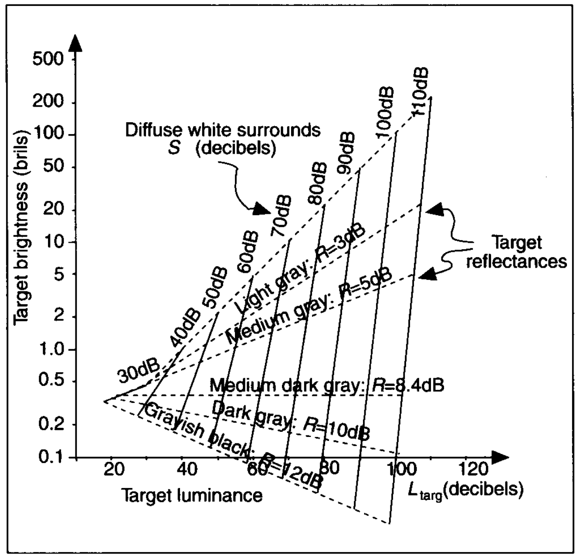


Figure 4. Stevens' 1960 experiment: Measured brightness versus luminance at various adaptation levels.

computer graphics must be judgment-free, since images often represent real-world scenes never viewed before. We build such an operator by finding robust display and observer models.

Observer model

One way to build an observer model is to approximate the eye's performance with the same sort of power-law relations used for film. However, the eye adjusts to its surroundings, so the sensitivity and contrast compression constants a_1 and γ must be made to change with the viewed luminances L_{in}; name these new functions $10^{\beta(L_{in})}$ and $\alpha(L_{in})$, respectively. The power-law relationship between brightness B and L_{in} is then

$$B = 10^{\beta(L_{in})} L_{in}^{\alpha(L_{in})} \quad (7)$$

Equation 7 applies to both the real-world observer ($L_{in} = L_{rw}$) and the display observer ($L_{in} = L_d$). This simple observer model is not new: It is used implicitly in Jones's graphical method for photographic tone reproduction, as discussed by James,[5] and agrees with the "power law" models of human neural response advocated by Stevens.[8] We will derive $10^{\beta(L_{in})}$ and $\alpha(L_{in})$ from their data.

Stevens and Stevens[9] attempted to measure the entire gamut of the human brightness versus luminance response. In their experiments, an observer's eye was allowed to adapt thoroughly to a uniform white background luminance, L_w. Then they briefly presented a small gray target with luminance L_{targ} against this background and assessed brightness. They measured sensations from "black" to "white" over most of the usable range of the eye. (Brightness values beyond this measured range might be analogous to the "knee" and "shoulder" regions of Figure 3. For simplicity's sake, we ignored these regions, even though this caused "clipping" in the results shown in Figure 5.) Figure 4 presents their results. Luminances were expressed in decibel units (dB), where 0 dB is defined as 10^{-10} lamberts (3.18×10^{-7} cd/m^2). Stevens' background luminance S (in dB) is converted to L_w in lamberts by

$$S = 100 + 10 \log_{10}(L_w) \quad (8a)$$

Stevens denoted the ratio between the target luminance L_{targ} in lamberts and the background, adapting luminance L_w as target reflectance R in dB:

$$R = 10 \log_{10}(L_w / L_{targ}) \quad (8b)$$

Using S and R, Stevens summarized the data in Figure 4 by Equation 8c, in a form that reveals the "brightness constancy" contour at $R = 8.4$ dB, where target brightness appears constant over a wide range of background luminance:

$$\log_{10}(B) = 0.004[(S - 27)(8.4 - R) - 108] \quad (8c)$$

Since Equation 8 describes line segments on log/log axes, it can also be written in the form of Equation 7, with luminance expressed in lamberts:

$$\alpha(L_{in}) = 0.4 \log_{10}(L_w) + 2.92 \quad (9)$$

$$\beta(L_{in}) = -0.4(\log_{10}(L_w))^2 + (-2.584 \log_{10}(L_w)) + 2.0208 \quad (10)$$

To complete the observer model, we need a workable value of L_w from the input luminances L_{in}: What uniform white background luminance would cause an amount of adaptation equivalent to that of the complicated image L_{in}? James[5] noted that several workers used average image luminance as an adaptation measure. We chose instead to use the assumption that the eye adapts in an attempt to keep most brightnesses near the "brightness constancy" contour of 8.4 dB below L_w in Figure 4, so

$$\log_{10}(L_w) = E\{\log_{10}(L_{in})\} + 0.84 \quad (11)$$

where $E\{\log_{10}(L_{in})\}$ is the expected value of $\log_{10}(L_{in})$.

Weakness of observer models

The eye's response to light is still not well understood. It is difficult to quantify because vision blends smoothly with higher brain functions and because the eye's behavior is strongly dependent on the content of the viewed image. Brightness response is usually described by several processes, including at least adaptation, simultaneous contrast, brightness- and color-constancy, memory, and cognitive processes. Many of these are interdependent, self-adjusting, and difficult to measure separately. Each tends to obscure the others, so brightness rules inferred from simple tests often fail when applied to more complex images.[10] Since brightness can only be measured indirectly, some researchers such as Schrieber[11] suggest that no quantitative model is above suspicion. We chose Stevens' model for its mathematical convenience.

A simple inverse CRT display model

We can build a display model from the television decoding equation (Equation 5) with an additive term for the ambient

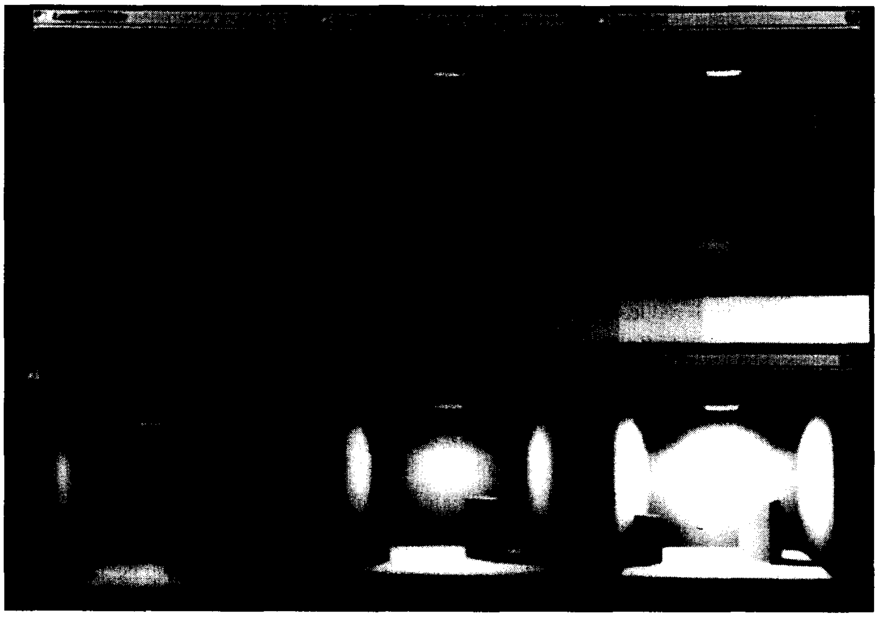


Figure 5. New tone-reproduction operator applied to radiosity solutions of Figure A.

lighting falling on the CRT face. This ambient light also places an upper bound on the display's available contrast:

$$\text{contrast} = (L_d/L_{d\,\max}) = n^{\gamma_d} + (BG/L_{d\,\max})$$

where

$L_{d\,\max}$ = maximum possible screen luminance (typically equal to 86 cd/m^2).

γ_d = approximately 2.9 for uncorrected CRTs, or 1.2 for displays with television-standard gamma correction.

n = frame buffer value ($0 < R = G = B = n < 1$) used as display device input.

BG = screen background luminance in cd/m^2, defined as ambient luminance multiplied by screen reflectance plus CRT secondary internal reflections.

L_d = display luminance in cd/m^2.

Let the maximum contrast ratio possible between on-screen luminances be denoted as $C_{\max}$. Now approximate the ratio $(BG/L_{d\,\max})$ as $(1/C_{\max})$. A typical value of $C_{\max}$ for direct-view CRTs is about 35.[6] We can form the inverse display system model that gives the display device input for the required display luminance L_d simply by solving for n:

$$n = [(L_d/L_{d\,\max}) - (1/C_{\max})]^{1/\gamma_d} \tag{12}$$

A complete tone-reproduction operator

Now all the pieces are in place to construct a tone reproduction operator as shown in Figure 2. First, change variable names in Equation 7 to create real-world and display observer equations:

$$B_{rw} = 10^{\beta_{rw}} L_{rw}^{\alpha_{rw}}$$

$$B_d = 10^{\beta_d} L_d^{\alpha_d}$$

The two observers are connected "back-to-back" as in Figure 2 by setting $B_{rw} = B_d$. We can then solve for display luminance in terms of real-world luminance:

$$L_d = L_{rw}^{\left(\frac{\alpha_{rw}}{\alpha_d}\right)} 10^{\left[\frac{(\beta_{rw} - \beta_d)}{\alpha_d}\right]} \tag{13}$$

Note that α_d and β_d are functions of the adapting luminance of the display, $L_{w(d)}$. This cannot be computed exactly from Equation 11 without knowing L_d. To get around this "chicken and egg" problem, we approximate $L_{w(d)}$ as constant. Since the total range of CRT output is small, changes in L_d have little effect on α_d and β_d. Assume display luminances are uniformly distributed on a logarithmic scale, with $\log_{10}(L_{w(d)})$ as its midpoint. Then Equation 13 can be completed using

$$L_{w(d)} = \frac{L_{d\,\max}}{\sqrt{C_{\max}}} \tag{14}$$

Finally, Equation 13 is substituted directly into the inverse display system model of Equation 12 to form a complete tone reproduction operator:

$$n = \left[\left(\frac{L_{rw}^{\left(\frac{\alpha_{rw}}{\alpha_d}\right)}}{L_{d\,\max}}\right) 10^{\left[\frac{(\beta_{rw} - \beta_d)}{\alpha_d}\right]} - \left(\frac{1}{C_{\max}}\right)\right]^{\left(\frac{1}{\gamma_d}\right)} \tag{15}$$

Figure 5 shows the results of this operator when applied to the radiosity solutions of Figure A (repeated in the lower left corner for comparison). Please note that these images were entirely the result of our new tone reproduction operator expressed by Equation 15 — no "tweaks" were used for display. This hands-off, judgment-free property alone justifies tone reproduction operators. Output values n are clipped to 1.0 and drive a CRT display where $\gamma_{crt} = 1.2$, as in Equation 5, and real-world luminances L_{rw} cover the searchlight-to-firefly range. The brightest image results from overhead lamp luminance at 1,000 lamberts (3.18×10^6 cd/m^2), and the lamp luminance is reduced by a factor of 100 with each successive image. The gray-scale strip shows steps of equal n.

The differences in these images show that this simple tone reproduction operator acts as both an exposure control and a contrast compressor. The tone operator is notably lacking in spatial effects. The very dim images should also be blurry, as the eye's resolution fades with decreasing light. The brilliant images should be more harsh, but glare and diffraction effects are also missing. The operator is also flawed for very dim images; for $S < 27$ dB (Equation 11), the slope of the *R-B* line in Figure 4 is negative. We made no attempt to model the limits of latitude in the observer models or display, so clipping occurs in the brighter images. However, even this simple operator appears to be a consistent and plausible solution to the display range problem.

Future work

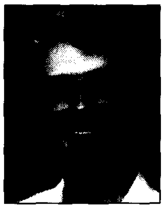
We have shown by example that good tone reproduction does not require ad-hoc methods and subjective judgements. We achieved plausible "hands-off" results with a small amount of computation. More sophisticated observer models should increase the accuracy of the displayed image by including more light-dependent effects of vision, especially extensions to color and spatial filtering. Even without noticeable improvements in brightness reproduction, these methods remove some guesswork from radiosity image display and demonstrate that tone reproduction operators need not be constrained to the same corrections used for film. Because any tone reproduction operator explicitly maps the results of a global illumination method to the display device, it also specifies the accuracy required in the illumination calculations themselves. Since computational costs rise rapidly as error tolerance falls, researchers involved in current efforts to develop efficient, realistic, image-synthesis systems need to explore tone reproduction operators. ❑

Acknowledgments

This work was funded in part by a grant from the National Science Foundation, ECS-9058389.

References

1. G.W. Meyer et al., "An Experimental Evaluation of Computer Graphics Imagery," *ACM Trans. Graphics*, Vol. 5, No. 1, Jan. 1986, pp. 30-50.
2. *Color and the Computer*, H.J. Durrett, ed., Academic Press, Boston, 1987.
3. H. Ripps and R.A. Weale, "Visual Function in Man, Part I: On Human Vision," *The Eye*, Vol. 2A, Ch. 2, "The Visual Stimulus," H. Davson, ed., Academic Press/Harcourt Brace Jovanovich, New York, 1976, pp. 43-99.
4. S.S. Stevens and J.C. Stevens, "Brightness Function: Effects of Adaptation," *J. Optical Society of America*, Vol. 53, No. 3, Mar. 1963, pp. 375-385.
5. T.H. James and C.E.K. Mees, *The Theory of the Photographic Process*, third ed., Macmillan Company, New York, 1966.
6. R.W.G. Hunt, *The Reproduction of Color*, third ed., John Wiley and Sons, New York, 1975.
7. L.E. DeMarsh, "Optimal Telecine Transfer Characteristics," *J. Society of Motion Picture and Television Engineers*, Vol. 81, No. 10, Oct. 1972, pp. 784-787.
8. S.S. Stevens, "To Honor Fechner and Repeal His Law," *Science*, Vol. 133, No. 13, Jan. 1961, pp. 80-86.
9. S.S. Stevens and J.C. Stevens, "Brightness Function: Parametric Effects of Adaptation and Contrast," Program of the 1960 Annual Meeting, *J. Optical Society of America*, Vol. 53, No. 11, Nov. 1960, p. 1,139.
10. C.J. Bartelson and E.J. Breneman, "Brightness Perception In Complex Fields," *J. Optical Society of America*, Vol. 57, No. 7, July 1967, pp. 953-957.
11. W.F. Schreiber, *Fundamentals of Electronic Imaging Systems: Some Aspects of Image Processing*, Springer-Verlag, Berlin, 1986.



Jack Tumblin is pursuing a PhD in computer science. He is researching topics in computer graphics, visual psychophysics, and digital signal processing at the Georgia Institute of Technology. From 1979-1990 he worked in television engineering, then was a founder of Ivex Corp. in Norcross, Georgia, where he designed flight simulator image generators for pilot training. His BSEE (1979) and MS (1990) are also from Georgia Tech.

Readers may contact Tumblin at College of Computing, Graphics, Visualization and Usability Center, Box 33711, Georgia Tech Station, Atlanta, GA 30332, e-mail ccsupjt@prism.gatech.edu.

Holly Rushmeier is currently on the staff of the Computing and Applied Mathematics Laboratory at the National Institute of Standards and Technology. Her research interests include scientific visualization, realistic image synthesis, and radiative heat transfer. Previously, while on the Mechanical Engineering faculty at the Georgia Institute of Technology, she received an NSF Presidential Young Investigator Award. Rushmeier received BS (1977), MS (1986), and PhD (1988) degrees in mechanical engineering from Cornell University.

Readers may contact Rushmeier at Rm. B-146, Bldg. 225, US Dept. of Commerce, National Institute of Standards and Technology, Gaithersburg, MD 20899, e-mail holly@cam.nist.gov.