EUROGRAPHICS ’2000 / M. Gross and F.R.A. Hopgood
(Guest Editors)

Volume 19, (2000), Number 3

Tone Reproduction for Interactive Walkthroughs

A. Scheel, M. Stamminger, H.-P. Seidel

Max-Planck-Institute for Computer Science, Saarbrücken, Germany

---

Abstract

When a rendering algorithm has created a pixel array of radiance values the task of producing an image is not yet completed. In fact, to visualize the result the radiance values still have to be mapped to luminances, which can be reproduced by the used display. This step is performed with the help of tone reproduction operators. These tools have mainly been applied to still images, but of course they are just as necessary for walkthrough applications, in which several images are created per second. In this paper we illuminate the physiological aspects of tone reproduction for interactive applications. It is shown how tone reproduction can also be introduced into interactive radiosity viewers, where the tone reproduction continuously adjusts to the current view of the user. The overall performance is decreased only moderately, still allowing walkthroughs of large scenes.

---

1. Introduction

Physically based rendering can be split into three phases. First a model of the virtual world has to be created. This includes the geometry of the scene, the surface reflectance behavior, and the emission characteristics of light sources. Second, a simulation of light exchange is performed within this model. Typical algorithms for this purpose are ray tracing or radiosity. The third step—creating a faithful representation of the resulting images on the output device—is of equivalent importance, although it is often not considered sufficiently.

Over the past decade several tone reproduction operators were developed for this purpose (e.g. [26, 23, 4, 17, 20, 5, 29, 14, 15, 24, 22]). The goal of tone reproduction operators is to find a representation of the image on the display that closely matches the impression of the real world the solution stands for. The major concern of such operators is to preserve the subjective impression of brightness in the scene. In this context, the term luminance is important. Luminance is the photometric counterpart of radiance, describing the magnitude of light as perceived by a standard human observer. It can be obtained by integrating the spectral radiance weighted by the luminous efficiency function $V(\lambda)$ of a (standard) human eye. Most tone reproduction operators map the luminances encountering the eye in the real world to luminances which can be displayed by the output device. The task is complicated by the limitations of the display device. The dynamic range of real world luminances visible by the human eye (about $10^{-6} - 10^{8}$ cd/m^2 [5]) by far exceeds the dynamic range of the monitor (about $1 - 120$ cd/m^2).

To achieve their goal, most tone reproduction operators mimic the functioning of the visual system (or at least some aspects of it). For example, Ward(1994)[26] tries to maintain perception of contrast, whereas Tumblin and Rushmeier[23] intend to reproduce the perceived brightness.

The functioning of the visual system depends heavily on adaptation. This is the mechanism that allows the eye to function over a wide range of illumination, although its photoreceptors can cope only with a relatively small range of luminances. If for example the illumination changes quickly from bright to dark, the eye will start to adapt to the new luminance level, with the effect that the impression of complete darkness slowly vanishes and contours of objects become visible. Thus, the most important functions of vision like recognition of the surroundings have been restored. Most tone reproduction operators compute the state of adaptation and try to reproduce the image like the adapted eye would see it, thereby exploiting the compression already done by the eye.

In this paper, we want to focus on tone reproduction for global illumination solutions obtained by radiosity methods.[8, 9] These finite element methods are restricted to purely diffuse surfaces. Due to this restriction, global light exchange can be simulated in environments of several hundred thousand patches.[18, 19, 7] Even if this simulation can take hours, the purely diffuse solution can be rendered very quickly by graphics hardware[3, 13] from arbitrary viewpoints. This way, interactive walks through the virtual, globally illuminated world, are possible.

Tone reproduction is necessary for the interactive case for the same reasons that apply to the still image. However, the level of luminance may change with the view point and viewing direction, which has to be accounted for by a change in tone reproduction. This imposes new requirements on a tone reproduction operator in terms of computation speed, application of the operator, and the determination of adaptation. For example, if the observer coming from a bright room enters a dark room, the new view should be mapped brighter than before, because the eye adapts to the new luminance level.

In this paper, we first give a short introduction into adaptation to provide a basis for the further discussion (2) and report on previous work in the field of tone reproduction (3). We then analyze the differences between tone reproduction for still images and tone reproduction for interactive walkthroughs and derive some new demands for the latter (4). In the following section (5) we show a possible way to realize the tone reproduction for interactive walkthroughs. Finally, we will present some results (6) and conclude (7).

2. Adaptation of the Visual System

Everybody knows the phenomena: if we switch off the lights at night, at first we cannot see anything at all. But the eye adapts to the new lighting situation and after one minute we begin to detect first objects and after about twenty minutes most features are visible again (although they lost their color). The same happens if the illumination suddenly changes from dark to bright, but this time the process of adaptation is less noticeable, because it happens much faster—in the first two seconds more than 80 percent of the adaptation is done.[5]

Adaptation is a necessary process to allow the eye to cope with the vast range of luminances in the world. The visual system operates nearly over 14 orders of magnitude, but single photoreceptors only have a dynamic range (i.e. the range where they are able to respond differently) of approximately 3 orders of magnitude.[25] Through adaptation the eye keeps a steady impression of the world. This does not mean that we are not able to perceive different levels of illumination, but they are shifted towards a better "operating" range.

In the following, we will briefly describe the mechanisms of adaptation to provide a basis for the further discussion about adaptation in conjunction with interactive tone reproduction. A very good overview of adaptation is given by Ferwerda et al.[5]

2.1. Adaptation Mechanisms

The range of perception is widened through the presence of two different kinds of photoreceptors in the eye: The rods operate at scotopic lighting conditions, which are encountered at night. They provide only black-and-white vision. The cones function at photopic levels of illumination (e.g. at daylight and in artificial light) and provide color vision. The transitional range between scotopic and photopic vision is called mesopic. However, the coexistence of the two receptor types alone is not sufficient to cope with the high luminance range. Other adaptation mechanisms are required.

Response Compression: The photoreceptors of the retina can become saturated; e.g. they have a maximum response amplitude. Stimulation by higher luminances beyond this point cannot increase the neuron's answer further.[25] The effect is, that with rising levels of adaptation luminances, the neuron's possible response range is decreased, such that the eye would be "blinded". Therefore, processes are needed, which mitigate the effects of response compression. These processes are either of multiplicative or subtractive nature.[2]

Multiplicative Adaptation: Multiplicative mechanisms scale down the input to a neuron, thus preventing that the maximum response is reached too early (this was called "dark glasses" effect by MacLeod[11]).

- Pupil size: As it is well known, the pupil becomes smaller with rising illumination, such that less light reaches the retina. But the pupil's size can only cause luminance changes in the range of one log unit, hence other processes are needed for adaptation.[25]
- Depletion of photo-pigments: At photopic luminance levels photo-pigments are bleached faster than they can be recovered. Therefore less photons can be absorbed and the neuron becomes less stimulus. In fact, for the cones pigment depletion stops further increases in light absorption already when they have reached half of their maximum response, such that they can operate up to the light-damage limit.[25]
- Postreceptoral gain chances: Also in cells which process the output of the photoreceptors—the horizontal, bipolar and ganglion cells—multiplicative mechanisms can be found.

Subtractive Adaptation: Another possibility for gain control is to subtract a fixed percentage of the signal from the adapting field. One example are the horizontal cells which get input from many cones. If a substantial amount of cones is stimulated so is the horizontal cell. This leads in turn to an inhibition of all cones connected to the horizontal cell.[25] Therefore, the whole input from this field is reduced by a certain amount.

2.2. Spatial Dependency of Adaptation Processes

Only the change of the pupil size can be called a global adaptation process because it depends on the overall illumination.

All other mechanisms work rather local: the photoreceptors can change their sensitivity individually.[25] However, their responses are linked together in the following network, so the adaptive states of a group of receptors influence the adaptive state of later cells.

Moon and Spencer[12] made a psychophysical experiment based on earlier research by Holladay[10], where they tested the effects of a non-uniform surround on adaptation. They presented two results. First, over 90 percent of the total adaptation depend on the small portion imaged on the fovea (e.g. the point of best vision having the by far highest density of cones) which has an angular diameter of approximately 1.5°. This means that about 1.5° of the visual field around the fixation point influence 90 percent of the adaptation. The second outcome was, that the influence of objects being further away from the center than 0.75° falls off with the square of the angle, whereas nearer objects do not depend on the angle.

The adaptation state is described by the so called adaptation luminance L_a. Following the results from Moon and Spencer L_a can be approximated by the following formula (simplified for a circular view):

$$L_a = 0.923L_b + 0.0192\int_{0.0131}^{1}\frac{L_s(\theta)}{\theta^2}\sin(\theta)\cos(\theta)\,d\theta, \quad (1)$$

where L_b is the luminance of the background (e.g. the portion imaged on the fovea; see Figure 1) and $L_s(\theta)$ is a luminance in the surround with angle θ between the focus and the point with luminance L_s. The integration domain is from 0.0131 radian (0.75°) to 1 radian ($\sim 57°$) (the influence of the exact choice of the upper border was found to be quite low).

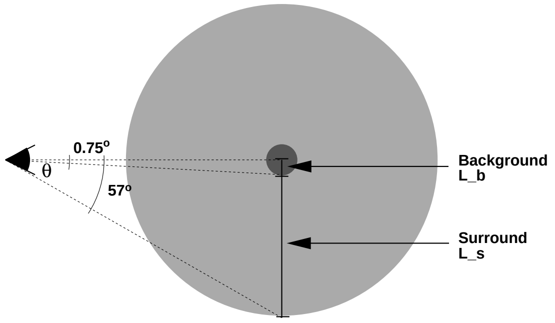

Figure 1: *The experimental setting of* [12]

This experiment was only done for cones. For the rods one would expect a less punctual adaptation. First of all, in contrast to the cones the rods are not centered at the fovea but more evenly distributed over the retina and are even absent in the fovea. For example, at night it is not possible to see a star, when trying to fixate it. Only if the eye is moved slightly such that the image of the star falls onto regions different from the fovea, it can be seen. In addition, the adaptation state of one rod influences directly the adaptation state of the surrounding rods.[25]

3. Previous work

In this section, we will summarize tone reproduction techniques. We will concentrate not only on tone reproduction in computer graphics but as well on tone reproduction for photography as both of these research areas influenced our work.

3.1. Tone Reproduction in Computer Graphics

The task of a tone reproduction operator is to present the global illumination solution for a scene on the display (monitor, print, etc.) such that it closely matches the impression of an observer of the corresponding real world's scene.

The tone reproduction operator has to face two major problems when fulfilling this task:

- It must mimic how the eye would see the real world and the displayed image and find a match between the two impressions.
- It must compress the high dynamic range of the real world to the small dynamic range of the display (e.g. for monitors 1 - 120 cd/m^2).

Fortunately, the eye does some "compression" itself by the process of adaptation, as was explained in Subsection 2.1. So most tone reproduction operators determine a world adaptation and mimic the eye's behavior for this adaptation level. There can be found two classes of such operators: single and multi adaptation level operators, working with one or multiple adaptation levels, respectively.

Single Adaptation Level The simpler tone reproduction operators like [26, 23, 5] assume a single world adaptation level. This level can be for example the average or the log average of the scene luminances. It is also useful to exclude light sources from the average computation if they are directly visible and small. The tone reproduction operator of Tumblin and Rushmeier[23] is based on a study of [21], who tested the brightness impression depending on the adaptation. The operator of Ward[26] tries to preserve contrast. For this purpose a so called threshold versus intensity (t.v.i.) function is used which determines the minimum discernible luminance for a given adaptation luminance. A linear tone mapping factor is then chosen such that differences visible by an observer being adapted to the world are also visible on the display with respect to the adaptation state of the display observer. Ferwerda et al.[5] enhanced Ward's operator with a t.v.i. function for the rods, such that it can also be used for scotopic luminance levels. Additionally, the operator accounts for reduced visual acuity at low light levels by removing frequencies from the image which would not be detectable for the world observer. Single adaptation level operators work best for scenes with low deviation from the average luminance. For scenes with high dynamic range problems arise: Consider a scene which shows the view into a dim and into a bright room. The average scene luminance would be somewhere in between these two luminance levels with the consequence that the dark room will be underexposed whereas the bright room will be over-exposed. The compressive power is reduced for the single-adaptation-level operators, but on the other hand they are fast to compute and work well for scenes with medium dynamic ranges.

Multi Adaptation Level In the past years new tone reproduction operators were developed which more closely mimic the visual system and enable larger compression by using multiple adaptation levels. Ward(1997)[29] filters the image into 1° blocks to find the different adaptation levels (based on the work by [12], see Section 2.2) The world luminances are then mapped such that differences visible at the corresponding adaptation level in real world should also be visible on the display (using the human t.v.i function). A highly perception based approach was taken by Pattanaik et al.[15] To account for the way the information gathered by the photoreceptors is processed by the neural network in the retina and in the cortex, a multi-scale representation of pattern, luminance, and color is used. In a sense also Schlick[16] works with different adaptation levels and produces a spatially non-uniform mapping. Schlick did not base his work on exact physiological examinations, but focused more on the preservation of details in high dynamic images.

A completely different approach for viewing still images was developed by Tumblin et al.[22]. In their "foveal display program", the viewer can explore an image interactively. The user tells the viewer its current fixation point through the mouse pointer. The adaptation is then recomputed by averaging the image luminances in a small region around the cursor position. If another point is selected the adaptation and thus the tone reproduction mapping is changed. As tone reproduction operator, a modification of the operator from [23] is used. Their approach is interactive, but it is a method for precomputed still images. The user can interactively change his fixation point, but not interactively walk through the scene. The method also borrows ideas from exposure metering in photography, which will be the topic of the next section.

3.2. Photography and Film

Photo and video cameras have to cope with the same problems as tone reproduction operators for synthetic images: they must somehow compress the world's dynamic range to the dynamic range of the film thereby maintaining the impression of reality. But they have only limited possibilities to achieve this goal. Cameras can just vary the exposure, that is the amount of light falling onto the film. The exposure depends in turn on the shutter speed (opening time) and the lens aperture (opening size).

In order to determine the exposure the (automatic) camera measures a weighted average luminance. In the classical center-weighted measurement sensitivity is concentrated to the center region of the view. Typically 75% of the measured result depend on a spot in the center that covers 15% of the image. Some cameras increase the importance of the lower half of the image, because in outdoor scenes usually one wants to recognize objects on the ground and not the clouds. For more difficult situations, e.g. a back-lit person, spot meters are favored, which solely measure light in a small spot at the center. Professionals average over different important parts of the image manually with such spot meters.[1]

Because the camera determines a single luminance level to guide the exposure, this principle is related to single adaptation level tone reproduction operators. The camera has the advantage that the method how to determine this level can be chosen by the user. Tone reproduction operators on the other hand should be fully automatic.

4. Demands on Tone Reproduction for Interactive Walkthroughs

Usually, tone reproduction operators are developed as an image post-processing step for high quality still renderings. The input images are generated by ray tracing or other, costly global illumination methods. Since the computation of such images takes minutes, hours or even days, a post-processing step requiring seconds is acceptable. When viewing a globally illuminated scene interactively, the luminances in the field of view change quickly. Therefore, ideally for each frame a new tone reproduction operator has to be computed. For this reason, the computation of the operator as well as its application to the image have to be fast.

Speed is not the only issue for interactive tone reproduction. The way the adaptation is determined which is needed to compute the tone reproduction operator must also be reconsidered:

The goal of tone reproduction operators for still images is to reproduce visibility in all areas of the image. This corresponds to a viewer, whose eyes wander over the image. The adaptation depends on the region of the image the viewer looks at and is thus not constant. The adaptation is usually determined as some average over all luminances of the image. In the "foveal display program" of Tumblin et al.[22] the user can give hints about his current fixation point through the mouse cursor. The displayed image is then optimal for looking at the chosen point. Multi-adaptation-level tone reproduction operators try to produce mappings which are reasonable for all fixation points over the entire image.

In interactive walkthrough applications, the situation is different. Users explore a scene they are moving through less by moving over the image with the eyes. Instead, the camera is panned towards interesting objects, in order to have a "closer look". As a result, objects of interest tend to be in the middle of the viewing area. So, the user mainly focuses to the center of the image.

Following the psychophysical experiments of Moon/Spencer[12] described in Section 2.2 the eye adapts mainly to a field of about 0.75° around the focus direction. In our type of application we just know that the user probably focuses to some point near the center of the image. Therefore, the observer's state of adaptation will be influenced by a field wider than 0.75° and the influence will decrease with distance to the center. Consequently, to compute the adaptation state of the world observer the luminances of the center should be weighted strongest and the weight should decrease towards the periphery.

Ideally, the weighting function is the convolution of the user's focus probability with Moon/Spencer's adaptation weighting function. Since we do not have a good model for the first one, we will employ a center weighted averaging—a technique borrowed from film technology.[1]

5. Realization of Tone Reproduction for Interactive Walkthroughs

In this chapter we will introduce a method how tone reproduction could be used for interactive walkthroughs in a radiosity solution. The integration of tone reproduction requires two steps: First the mapping curve has to be determined from the computed image of world radiance values. Then this curve is applied to the computed world luminances to create an image in the framebuffer or in a file.

For efficiency reasons, we restricted ourselves to operators that map simulated world luminances L_w to display luminance L_d with a mapping function that is invariant of the pixel position. This excludes operations like spatially changing mapping curves [4] or blurring in particularly dark or bright regions [29].

5.1. Determination of the Mapping Curve

5.1.1. Sampling

The first issue to be addressed is the computation of the tone reproduction operator itself. This computation requires knowledge about the distribution of luminance values for the current view. Since the graphics hardware produces the frames, we have no information about which parts of the scene are currently visible. Even by looking into the framebuffer after each frame—which is a very costly operation—we can only get information about the intensity distribution *after* the mapping.

Instead, we cast some rays in between the frames into the scene from the current view. For the hit points, the world luminance values are obtained from the radiosity solution and Equation 2.

5.1.2. Sampling Pattern

As was described in Section 4 the world adaptation should be determined mainly from luminances in the center. The periphery is less important for the adaptation.

To estimate the adaptation we borrowed a measurement method from film technology—center weighted averaging. Figure 2 (left) shows a typical sampling pattern, where the samples are concentrated around the center. This fits well with the results of [12], who state that the periphery contributes little to the adaptation and that the exact representation of the objects in the periphery is not important (see Section 2.2). As mentioned before, the exact focus of the viewer is not known and, unless the user does not stare to one point, the focus will move constantly—probably inside the center region. Therefore, the region of dense sampling is chosen larger then just 0.75°.

In our application an array of sample positions according to the sample distribution function shown in Figure 2 is precomputed before the walkthrough starts. Since the time for this preprocess is not important, we use a simple rejection technique for the sample creation. The precomputed sample positions are then used cyclically for shooting the rays.

5.1.3. Number of Samples

The number of samples which are used to compute the mapping is critical. If it is too small, the stochastic luminance value distribution change is noticeable, even if the view is only changed slightly. As a disturbing result, the overall brightness changes rapidly. Even worse, the operator Ward(1997) does not work properly without a sufficient number of samples (see Section 6).

Our approach is as follows: we store an array with a sufficient number of samples (e.g. 1000). In between single frames a minimum number of rays is shot (e.g. 100) and the oldest samples are replaced. If there is still enough time left further rays are shot, as long as some rendering framerate can be kept. Before a frame is rendered, the tone reproduction operator is recomputed, possibly reusing old samples from previous frames.

The fact that only some of the samples change from frame to frame can also be exploited to speed up the recomputation of the tone reproduction operator. In the case of Ward(1994), we only need the current average luminance, which can quickly be updated everytime a new sample arrives. For Ward(1997), the situation is similar: whenever a new sample is obtained, only the luminance histogram has to be updated accordingly. The costly parts of the computation, namely iterating over the input image, can be omitted this way.

Furthermore, the approach prevents rapid changes in the mapping curve, because only a subset of the samples has been replaced, resulting in delayed adaptation. A similar effect appears with a video camera. If the view is changed

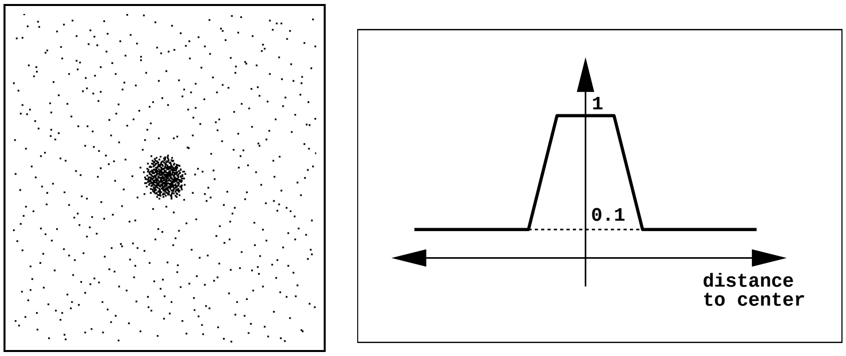

Figure 2: *Left: Typical sampling pattern. Right: Sample distribution function.*

quickly from dark to bright, the adaptation the automatic exposure control of the camera takes a fraction of a second to adapt to the new environment.

5.2. Application of the mapping curve

The tone mapping curve can be determined for each frame as described in the previous section. The next point to be discussed is, how the mapping curve could be applied to the current view.

5.2.1. Discussion

Since graphics hardware does the rendering of radiosity solutions, we have no intermediate images with world luminances. Instead we have a finite element solution of the scene with radiosity values, which first have to be converted to displayable color values, usually integers between 0 and 255. Then the scene is passed to the graphics hardware using these mapped colors.

Note, that usually the scene is passed *once* to the graphics hardware and then rendered from different view points. In OpenGL, this is done by creating a display list which is then rendered with varying view points. By this, the data transfer is dramatically reduced. Unfortunately, this rules out changing the mapping from radiosity to hardware color values. However, we need a method to be able to quickly change the mapping, ideally before each single frame is drawn.

In OpenGL, colormaps seem to be a solution for this problem. At first glance, it is possible to use a colormap for mapping world luminances to display luminances. The mapping could then quickly be changed by loading a new colormap. Our experiments using colormaps for our purpose failed, though. Even after several problems have been eliminated (e.g. we had to put dummy textures onto the objects for the colormaps becoming active at all), we stopped the approach due to accuracy and performance problems. First, colormaps themself are often restricted in size, second the usual framebuffer's 8 bit accuracy for colors is not sufficient. Finally, on our hardware the application of colormaps decreased performance significantly.

Creating the scene newly for every frame, which would make it possible to change the mapping continuously, results in a large performance loss. Especially for high-end machines, tools exist (e.g. IRIX Optimizer), which do complex visibility and occlusion culling computations for the rendering of highly complex scenes with the side effect of having to create the scene newly for every frame. This would allow us to change the color mapping. Another possible approach would be to use vertex arrays which are an acceleration method to pass big sets of triangles including color information to OpenGL graphics hardware. By this, the color values remain accessible and can thus be changed in between frames. However, there is still a performance loss compared to a display list, especially on normal desktop computers (in our case an SGI O_2). Furthermore, just the computation of the mapping of all vertex color values of more complex scenes with maybe 100 000 vertices is too expensive for an interactive application.

Our approach is a compromise that makes it possible to continue using display lists, but also to redefine the mapping function quickly. We separately code the luminance of every vertex into texture coordinates. Before rendering the scene, a texture is loaded that maps these luminance coordinates into display luminance values, which modulate the object colors. By redefining the texture in between frames, the luminance mapping can be changed quickly. Unfortunately, with this approach, the normal texturing of objects becomes more difficult. Multipass-rendering is too slow for complex scenes. Instead, multitexturing is meanwhile available widely and can be used for both, texturing the objects and changing the tone reproduction.

Note that our approach only maps luminances, not colors. Operators that change the hue, e.g. by shifting dark saturated regions to grey (e.g. [29]), cannot be represented this way. Spatially non-uniform mappers (e.g. [4]) also cannot be represented by a single mapping curve. However, the majority of the tone reproduction methods only computes uniform luminance mappings and is thus applicable.

In the following we will first describe how to map a computed radiosity value to texture coordinates and a normalized color value. Then it is described how a texture has to be defined that represents a particular mapping curve.

5.2.2. Color Mapping

Let's assume, we want to display a vertex with world radiances (L_r, L_g, L_b). Depending on the underlying RGB color space, this results in a luminance value of

$$L(L_r, L_g, L_b) = E\bar{L} = E(\alpha L_r + \beta L_g + \gamma L_b), \tag{2}$$

where α, β, and γ sum up to one and weight the channels according to their perceived luminance $(V(\lambda))$. E is the luminance of the monitor's white point (1,1,1).

The principle idea is to represent $\bar{L}$ as one texture coordinate and to use the normalized color value $(L_r/\bar{L}, L_g/\bar{L}, L_b/\bar{L})$ as vertex color. Unfortunately, these normalized color values are not in the interval $[0,1]$ yet. Standard OpenGL, however, first clamps these values to the unit interval, thus changing brightness and color hue.

Therefore, we have to extract an additional scaling factor $m = \max\{L_r/\bar{L}, L_g/\bar{L}, L_b/\bar{L}\}$ from the color triple, leaving $(L_r/m\bar{L}, L_g/m\bar{L}, L_b/m\bar{L})$, each component of which is in the interval $[0,1]$. m itself is one for grey color triples and reaches its maximum value $1/\gamma$ for purely blue colors, since γ is always smaller than α and β.

Altogether, the used decomposition is:

$$(L_r, L_g, L_b) = \overbrace{L/E}^{\text{texture } s} \quad \overbrace{m}^{\text{texture } t} \quad \overbrace{(L_r/m\bar{L}, L_g/m\bar{L}, L_b/m\bar{L})}^{\text{color}}$$

The luminance factor L/E is represented as one texture coordinate value as described below. Fortunately, textures can be two-dimensional, so we can use the second dimension for representing m. The remaining color triple is guaranteed to be in $[0,1]^3$ and is given to the hardware as vertex color.

Also L has to be mapped to the unit interval in order to be usable as texture coordinate. Because the possible range of L is very big, we use a logarithmic mapping of L. First minimum and maximum values for L over all vertices are determined. The texture coordinate s is then computed as

$$s = (\log L - \log L_{\min})/(\log L_{\max} - \log L_{\min})$$

Since the logarithm is not defined for zero, we restrict the lower bound to $10^{-4}\mathrm{cd/m}^2$, which is roughly the minimal perceivable luminance.

In our experiments the graphics hardware's precision for texture coordinate computations was sufficient to represent the luminances correctly. Similar luminance mapping techniques are used for storing high dynamic range images, where a luminance precision of 16 bit or less is used.[27] This precision can also be expected for texture coordinates, otherwise it would not be possible to address large textures.

The mapping of m is simpler. For all grey colors ($L_r = L_g = L_b$), m is equal to one. m reaches its maximum for completely saturated blue $(0,0,L_b)$, where $m = 1/\gamma$, which can be as much as 20. So we use $t = 1/m$ as second texture coordinate.

When rendering the scene into a display list, for every vertex with computed radiosity (L_r, L_g, L_b) the above mappings are performed, obtaining a 2D-coordinate (s,t) and a normalized color value (R,G,B). These are then passed as texture coordinates and as normal object color, with all hardware lighting capabilities disabled.

5.2.3. Texture Creation

Let's assume we have a mapping curve $V(L)$, which maps world luminance values L to device intensity values V. Using the above texture coordinate mapping, the texture point (s,t) represents the luminance

$$L(s,t) = \frac{1}{t} L_{\min} \left(\frac{L_{\max}}{L_{\min}} \right)^s,$$

thus a texel at (s,t) must contain the value $V(L(s,t))$.

The tone reproduction is done mainly by the s-coordinate, t is only an artifact of OpenGL's color clamping. Thus we use a much bigger resolution of the texture in the s-dimension than in t. Texture sizes of $128 * 8$ have turned out to be sufficient for all our examples, if linear texture interpolation is used. Nearest neighbor mapping requires rather big textures, otherwise discontinuities in the displayed luminance are clearly visible.

Figure 3 shows three such textures created by our system when viewing a dark region of the scene (left), a part of the scene with high dynamic range (center) and a bright area (right).

5.3. Choice of Tone Reproduction Operators

We focused our efforts on two operators published by Ward in 94[26] and 97[29]. Both operators were modified in that they measure world luminances like it was pointed out in the previous sections.

As described in Section 3, Ward(1994) uses a single adaptation level, which is determined as the average or logarithmic average luminances of the scene radiances. Depending on the adaptation it then computes a linear mapping factor that tries to preserve contrast in the scene.

$$L_d = m L_w$$

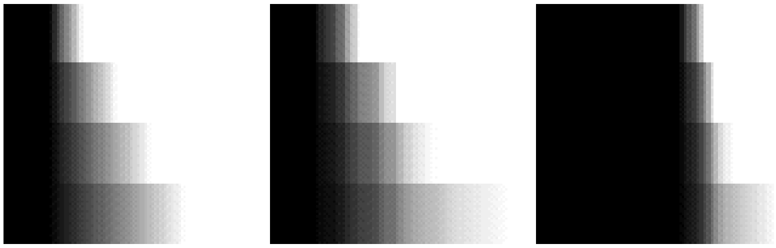

Figure 3: *Three tone reproduction textures for a dark view (left), a view with high dynamic range (right), and a bright view (center).*

A single adaptation level tone reproduction operator shows a behavior like a video camera: the details in the center of the image are presented with an adequate illumination level. In the center, differences visible to a real world observer should be visible to the display observer as well. But it might happen, that features in the periphery are not detectable, because they are under- or overexposed. This behavior can also well be seen in the accompanying video. We used the operator Ward(1994) because of its simplicity. Of course other single-adaptation operators like [23] or [5] would also be applicable.

Multi adaptation level operators offer the advantage that they can better handle uneven lighting conditions in an image (e.g. a dark and a bright region). But on the other hand, they are more complex to compute and thus more time consuming. Tone reproduction operators which map luminances depending also on their spatial position (e.g. [4, 16, 15]) are not suited for our purpose, because for our approach a single mapping curve is needed. Instead, we chose the tone reproduction operator from Ward(1997), because it fits well with our concept. Glare effects, reduction of visual acuity for low luminance levels and a transition to black and white vision at scotopic light levels could not be integrated, though.

Based on a histogram of the image sample luminances, in Ward(1997) a piecewise linear mapping function is determined that compresses low populated radiance intervals and expands dense luminance regions. The entire process is depicted in Figure 4: First, the input image is down-sampled to 1° averages (left) and converted to a histogram of luminance values (center left). The histogram values are cut corresponding to human contrast perception restrictions (center right). Finally, a cumulative histogram is computed which defines the piecewise mapping curve (right). Low sample numbers in a histogram bin result in a flat mapping curve for the corresponding world luminance range, which in turn compresses this luminance range. In contrast, a large number of samples in a bin leads to a steep mapping curve, expanding the display luminance range for the corresponding world luminances.

We expect that the differences between tone mapped views will not be as strong as those produced by Ward(1994), because this operator is able to map all parts of the view into a range of good visibility.

In our experiments, it turned out that Ward(1997) does not work well if the number of samples is too small. We achieved better results by initializing all histogram bins with some constant number of samples. As a result, the variation of the mapping curve is reduced and the curve is moved a bit towards a linear mapping, which both improved visual quality.

6. Results

6.1. Ping-Pong Room

To show the effects of adaptation, we used a scene with a dark and a bright room. An observer standing in the corridor in front of the two rooms alternatively looks either into the dark or into the bright room. In reality, he will be able to recognize the objects in both rooms, although the difference in lighting conditions between the two rooms is perceived clearly. We checked this observation by visual inspection.

Next, we tested how a video camera records this scene (see top row of Figure 5 and part "Real Scene" of the accompanying video). When the camera points to the middle between the two rooms, the dark room is under-exposed (i.e. no features visible) and the bright room is a little overexposed (i.e. looks too bright). So this luminance range is already a difficult situation for the camera. If the camera is panned to the left, it adjusts the brightness of the image and some objects in the dark become visible. If it is panned to the right, the brightness of the image is reduced making the room looking less harsh.

The scene was (approximately) re-built and a hierarchical radiosity solution was computed. Unfortunately, no means to measure the real reflectances and the exact luminance of the light sources were available, so we had to estimate these quantities. For interactive viewing of the radiosity solution the tone reproduction operators of Ward(1994)[26] and Ward(1997)[29] were applied in the way described in the previous sections.

The middle row of Figure 5 (and video part "Ward94")

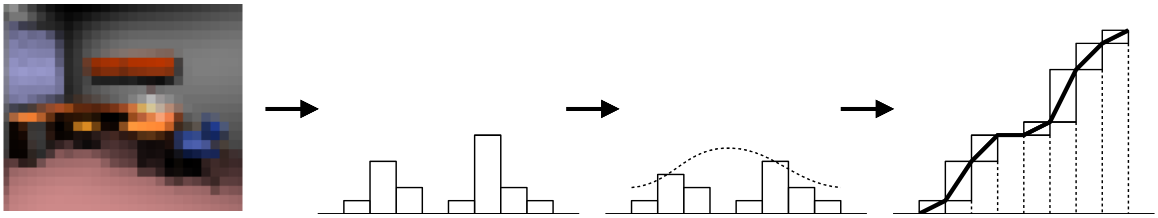

Figure 4: *The tone reproduction of Ward-Larson '97: the input image (left) is converted to a histogram (center left), clamped due to physiological considerations (center right). Finally, the mapping curve is defined by the cumulation of the histogram (right).*

shows the results for Ward(1994). To compute the world adaptation luminance 90 percent of the samples were placed inside the inner circle. Ward(1994) shows a similar behavior as the camera. When the view points to the left room, the scene is mapped brighter such that one can guess that there is a chair in the room. When the view is turned to the right, the brightness of the bright room is shifted to a convenient level.

The application of the tone reproduction operator Ward(1997) delivered different results (see bottom row of Figure 5 and video part "Ward97"). The images show a smaller change in brightness, when the view is turned from left to right. But all parts are in a range of good visibility. The compression rate of this tone reproduction operator is much stronger than can be achieved with the linear mapping of Ward(1994) (and without gamma correction). Here, the operator succeeded to map all luminances to appropriate levels. On the other hand, the operator reduces the differences in overall brightness between different views. We therefore suggest to use the operator when it is desired that most detail is visible in a single view.

6.2. Townhouse

Another test scene was the model of a townhouse with three storeys and several rooms, which can be obtained from the RADIANCE website.[28] After turning off the light in some rooms, the scene was a good test environment for our interactive tone reproduction. The accompanying video shows a walk through the model. The snapshots in Figure 6* give an impression of the adaptation of the reproduction operator for a user with constant view point, but different viewing directions.

For the adaptation determination we used a set of 900 samples. At least 100 of them were recomputed in every frame. 900 samples are sufficiently accurate for the scene, and the update of 100 samples per frame reduces the framerate only marginally. For the shown model we measured an average overhead for the interactive tone reproduction computations of less than ten milliseconds per frame for the tone reproduction operator Ward(1994) and about 15 milliseconds for the operator Ward(1997). By decreasing the number of samples to be computed between two frames, the time slot for ray casting and tone reproduction computations can be decreased even more, which has to be bought by slower adaptation of the displayed luminances.

7. Conclusion and Future Work

In this paper we showed that many ideas of tone reproduction can be transfered to interactive viewing of global illumination solutions. Assumptions that are made for the tone reproduction of still images have to be modified for the interactive context, but of course the main ideas remain valid. We showed a way how the mapping of world to display luminances can be represented by a texture, which can be quickly updated in between frames of interactive renderings. If the tone reproduction computation is restricted to coarsely sampled images, the mapping can be adapted from frame to frame corresponding to the current view.

So far, we did not consider the time course of adaptation. The adaptation to brightness is much faster than to darkness. Therefore, the adjustment of the tone reproduction towards dark areas could be delayed. A second idea is to change the tone reproduction if the user stops moving. If the user stops, he obviously wants to explore the current view, so multiple adaptation levels are probably more important.

Figure 5: *The ping-pong scene. Top row: real world images from video camera, middle row: interactive viewing with tone reproduction of Ward91, bottom row: interactive viewing with Ward97.*

References

1. Ansel Adams. *The Camera + The Negative*, volume 1 + 2. Little, Brown and Company, 1980.
2. E. A. Adelson. Saturation and adaptation in the rod system. *Vision Research*, 22:1299–1312, 1982.
3. Daniel R. Baum, Stephen Mann, Kevin P. Smith, and James M. Winget. Making radiosity usable: Automatic preprocessing and meshing techniques for the generation of accurate radiosity solutions. In Thomas W. Sederberg, editor, *Computer Graphics (SIGGRAPH '91 Proceedings)*, volume 25, pages 51–60, July 1991.
4. K. Chui, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface*, pages 245–254, May 1993.
5. James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual adaptation for realistic image synthesis. In Holly Rushmeier, editor, *Computer Graphics (SIGGRAPH '96 Proceedings)*, pages 249–258, August 1996.
6. James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual masking for computer graphics. In Turner Whitted, editor, *Computer Graphics (SIGGRAPH '97 Proceedings)*, pages 143–152, August 1997.
7. S. Gibson and R. J. Hubbold. Efficient hierarchical refinement and clustering for radiosity in complex environments. *Computer Graphics Forum*, 15(5):297–310, 1996. ISSN 0167-7055.
8. Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modelling the interaction of light between diffuse surfaces. In *Computer Graphics (SIGGRAPH '84 Proceedings)*, volume 18, pages 212–22, July 1984.
9. Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. In Thomas W. Sederberg, editor, *Computer Graphics (SIGGRAPH '91 Proceedings)*, volume 25, pages 197–206, July 1991.
10. L. L. Holladay. *Journal of the Optical Society of America*, 12:271, 1926.
11. D. MacLeod. Visual sensitivity. *Annual Review of Psychology*, 29:613–645, 1978.
12. Parry Moon and Domina Eberle Spencer. The visual effect of non-uniform surrounds. *Journal of the Optical Society of America*, 35(3):233–248, mar 1945.
13. Karol Myszkowski and Tosiyasu L. Kunii. Texture mapping as an alternative for meshing during walkthrough animation. In *Fifth Eurographics Workshop on Rendering*, pages 375–388, Darmstadt, Germany, June 1994.
14. Laszlo Neumann, Kresimir Matkovic, Attila Neumann, and Werner Purgathofer. Incident light metering in computer graphics. *Computer Graphics forum*, 17(4):235–247, 1998.
15. Sumanta N. Pattanaik, James A. Ferwerda, Marc D. Fairchild, and Donald P. Greenberg. A multiscale model of adaptation and spatial vision for realisitic image display. In Michael Cohen, editor, *Proceedings of SIGGRAPH '98*, pages pp 287–298. ACM, ACM Press, July 1998.
16. Christophe Schlick. Quantization techniques for visualization of high dynamic range pictures. In G. Sakas, P. Shirley, and S. Mueller, editors, *Photorealistic Rendering Techniques (Proceedings of the 5th Eurographics Workshop on Rendering)*, pages 7–20, June 1994.
17. Christophe Schlick. A survey of shading and reflectance models. *Computer Graphics Forum*, 13(2):121–131, June 1994.
18. François Sillion. Clustering and volume scattering for hierarchical radiosity calculations. In *Fifth Eurographics Workshop on Rendering*, pages 105–117, Darmstadt, Germany, June 1994.
19. Brian Smits, James Arvo, and Donald Greenberg. A clustering algorithm for radiosity in complex environments. In Andrew Glassner, editor, *Proceedings of SIGGRAPH '94 (Orlando, Florida, July 24–29, 1994)*, Computer Graphics Proceedings, Annual Conference Series, pages 435–442. ACM SIGGRAPH, ACM Press, July 1994. ISBN 0-89791-667-0.
20. Greg Spencer, Peter Shirley, Kurt Zimmerman, and Donald P. Greenberg. Physically-based glare effects for digital images. In Robert Cook, editor, *SIGGRAPH 95 Conference Proceedings*, pages 325–334. ACM Press.
21. J. C. Stevens and S. S. Stevens. Brightness function: Effects of adaptation. *Journal of the Optical Society of America*, 53(3):375–385, 1963.
22. Jack Tumblin, Jessica Hodgins, and Brian Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, 1999.
23. Jack Tumblin and Holly Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, 1993.
24. Jack Tumblin and Greg Turk. Lcis: A boundary hierarchy for detail-preserving contrast reduction. In Alyn Rockwood, editor, *Computer Graphics (SIGGRAPH '99 Proceedings)*, pages 83–90, August 1999.
25. Jan Walraven, Christina Enroth-Cugell, Donald C. Hood, Donald I. A. MacLeod, and Julie L. Schnapf. The control of visual sensitivity: Receptoral and postreceptoral processes. In L. Spillmann and J. Werner, editors, *Visual Perception, The Neurophysiological Foundations*, pages 53–102. Academic Press, 1990.
26. Greg Ward. A contrast-based scalefactor for luminance display. In P.S. Heckbert, editor, *Graphics Gems IV*, pages 415–421. Academic Press, 1994.
27. G. Ward-Larson. `http://positron.cs.berkeley.edu/gwlarson/pixformat/tiffluv.html`.
28. G. Ward-Larson. `http://radsite.lbl.gov/radiance/pub/models/index.html`.
29. Greg Ward-Larson, Holly Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, 1997.

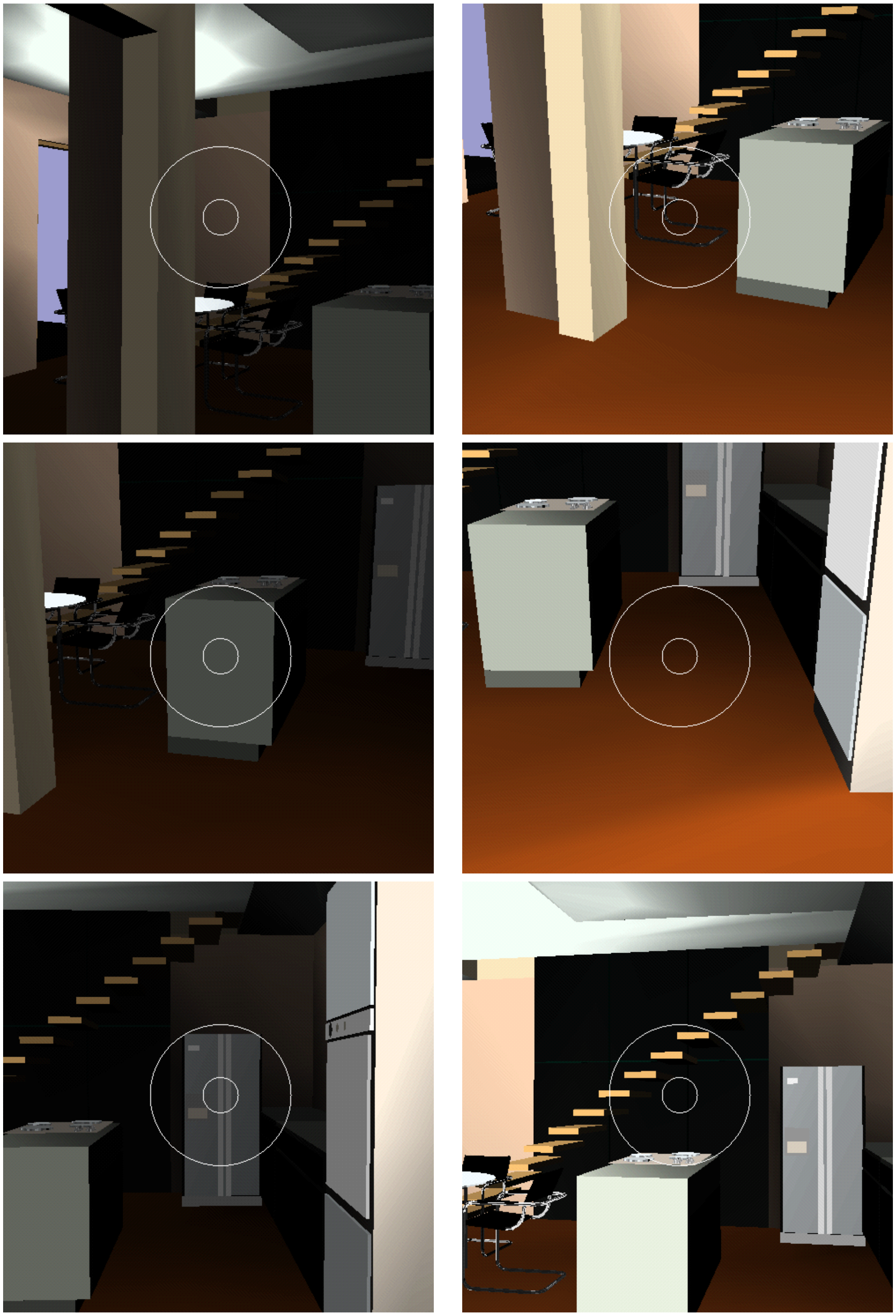

Figure 6: *Townhouse scene viewed interactively with different adaptation levels.*