A Tone Mapping Algorithm for High Contrast Images

Michael Ashikhmin

Department of Computer Science, SUNY at Stony Brook

Abstract

A new method is presented that takes as an input a high dynamic range image and maps it into a limited range of luminance values reproducible by a display device. There is significant evidence that a similar operation is performed by early stages of human visual system (HVS). Our approach follows functionality of HVS without attempting to construct its sophisticated model. The operation is performed in three steps. First, we estimate local adaptation luminance at each point in the image. Then, a simple function is applied to these values to compress them into the required display range. Since important image details can be lost during this process, we then re-introduce details in the final pass over the image.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/image generation; Display algorithms

1. Introduction

Many current state of the art image synthesis algorithms model the physics of light propagation through a scene. If provided with an accurate description of the environment, these methods are capable of computing luminance values at each pixel which closely match those measured at corresponding points in the real scene [28]. As a result, the range of these values can be vast. The problem of sensibly mapping the results of a physics-based image synthesis algorithm to the limited range of displayable values is commonly referred to as tone mapping. With recent active developments in the field of high dynamic range photography [4], the importance of the problem has risen significantly since even systems which are not based on accurate lighting simulations are now potentially capable of producing full range of real world luminance values.

Arguably, to construct a good tone mapping operator, we need to consider how humans see the world around them. The problem is therefore complicated by our lack of complete understanding of human visual system (HVS). However, enough information has been accumulated over years of research to develop computational models of initial stages of the HVS [15] which have proven useful in computer graphics, including construction of several good tone mapping operators. Unfortunately, such operators are usually rather complex. In the other extreme, there exist several very simple operators but their results commonly fail to reproduce some important features of visual experience for the corresponding scene.

In this paper we present a new tone mapping operator which mirrors two relevant functions of HVS: signalling absolute brightness and local contrast. Most previous approaches have either addressed only one of these issues or attempted to deal with both parts within a single complex computational framework. In contrast, we address these two problems separately. We use local adaptation level at a particular image position to signal absolute brightness and non-linearly compress these adaptation luminances into displayable range using a simple mathematical function we derive. We also provide a novel way to estimate local adaptation luminance needed for this part of the algorithm. The final stage of the algorithm ensures that fine image details are present in the tone mapped image. We believe that the main advantage of our approach compared with current state of the art methods is that it produces an algorithm which is significantly simpler without compromising the quality of the results. Moreover, clearly separating phases responsible for different tasks allows for easy replacement of constituent algorithms as better methods become available.

The rest of the paper is organized as follows. We review relevant work on the tone mapping problem in section 2. Our assumptions and guiding principles are given in Section 3.

The main stages of the algorithm are developed in Section 4. We present our results in Section 5. Section 6 concludes the paper.

2. Related Work

Advances in understanding of the human visual system have been significant [14]. For a comprehensive review of relevant literature from the computer graphics perspective, see [15, 6]. We state here only a few results which are important for better understanding of our approach.

It has been discovered that light plays a rather peculiar role in visual perception. It is necessary to project information onto the retina but there are both physiological and psychophysical evidences [20, 26] that absolute brightness information is of secondary importance to HVS and tends to be largely discarded on very early stages of visual processing through mechanisms of brightness constancy. Local contrasts are used instead to convey the wealth of information about the outside world. For a tone mapping algorithm it is therefore very important to preserve this contrast information. Second important fact is that human sensitivity to luminance changes is not constant throughout the range of luminances, but is rather given by threshold vs. intensity (TVI) function which gives the value of just noticeable luminance difference for given adaptation level L_a. This function has been experimentally measured (see Figure 1 below and discussion in [15] on relevant techniques). Moreover, it has also been shown that L_a can vary across the image, which is known as local adaptation [12, 3]. TVI function and local adaptation play a key role in our tone mapping procedure.

Researchers and practitioners in several areas encounter problems similar to that of tone mapping. In image processing it is often desirable to increase image contrast without introducing artifacts. A full range of enhancement techniques has been developed [8]. It includes simple methods such as histogram equalization along with some recently developed algorithms [9] which provide excellent results by incorporating multiscale processing based on retinex theory of HVS [10]. Some image processing techniques served as a starting point for tone mapping procedures [11], but as these authors note, there are two fundamental differences between image enhancement and tone mapping. First, maintaining perceptual fidelity is usually of no importance to an image enhancement technique. Second, tone mapping procedure deals with undistorted world luminances while image processing algorithm is already presented with a limited dynamic range input. Since artifacts tend to grow with dynamic range, direct application of image processing techniques may be not appropriate.

In photography, dodging and burning techniques are popular to address the problem of obtaining a good looking image of a high contrast scene and these ideas were exploited in computer graphics [2, 13]. In particular, our algorithm was developed in parallel to and is very similar in some respects to that of Reinhard et.al. [13] who provide an alternative justification for the procedure presented in section 4.2 as well as a wealth of references to photographic literature. The main difference of this work from [13] is that we attempt to develop a method for specific case of high contrast images which is based mostly on insights from human perception and uses little user intervention while Reinhard et al. allow for more detailed user control based on direct adaptation of several techniques from photographic practice.

In computer graphics, the need for mapping world luminances into display range was recognized since the development of first physics based rendering algorithms. Simple linear scaling is one of the most popular practical approaches. Ward [27] proposed to choose the linear scaling factor based on the ratio of world and display just noticeable differences in luminance. This approach was later used in [7] with more accurate expressions for the TVI function. As discussed below, there are some serious inherent limitations to linear mappings which encouraged the search for nonlinear techniques.

Early non-linear methods included applying simple compressive functions, such as cubic root or logarithm to the luminance values. Tumblin and Rushmeier [24] were the first to address the problem of preserving viewer's overall impression of brightness. They proposed an approach based on simple power relation between world luminance and perceived brightness. A single adaptation value was used for the scene and another for the display. In [23] this expression was revised to incorporate a linear scaling factor based on adaptation luminance. The operator is then used to separately compress illumination layers in a computer graphics rendering while preserving surface reflection information. Other methods which rely on availability of some additional information about the scene and are therefore restricted to rendering applications, include recent interactive techniques presented in [5, 18].

All methods mentioned above use a single adaptation value for the entire scene. Attempts to account for local adaptation effects have been done by Chiu et al. [2], Tanaka and Ohnishi [22], and, in a much more consistent manner, by Pattanaik et al. [15] who present the most comprehensive model of human visual system currently used in computer graphics. These techniques are based on applying a linear filtering procedure to the input and using the result to reduce contrast in the tone mapped image either through division by the blurred image [2, 22] or by using a more sophisticated multiscale mechanism [15]. All these methods are susceptible to halo artifacts as discussed in [25, 23] (see also the next section). Schlick [19] concentrated on improving the efficiency of Chiu et al.'s method.

Ward et al. [11] presented a method based on histogram equalization which avoids halo artifacts altogether. It adjusts image histogram so that display contrast does not exceed that of the real scene. However, nothing prevents the displayed

contrast from being too low in some cases. Also, as argued in [25], strictly monotonic mapping might be too restrictive if visually pleasing results similar to those produced by skilled artists are desired. Nevertheless, we believe that this method currently offers the best combination of practical simplicity and uniformly good perceptually advocated results and we chose to compare our algorithm with this technique in section 5.

Tumblin and Turk recently proposed a method based on artist's approach to drawing of high contrast scenes [25]. They separate the image into a hierarchy of detail images through a procedure based on anisotropic diffusion. They then compress the range of each image using different linear scales. This allows to use overlapping intensity ranges in different parts of the scene giving the method greater flexibility than monotonic mapping procedure of [11]. Our approach is similar in spirit, but we compute a single auxiliary image using a much simpler algorithm and use non-linear mappings as well as different detail preservation procedure.

While we concentrate on tone mapping luminance in this paper, other related effects have been addressed in computer graphics literature. They include chromaticity change and acuity loss with darkness [7, 11], blue shift for night scenes [5], chromatic adaptation [15], veiling luminance and other scattering processes in the eye [11, 21], and time dependent adaptation [7, 16]. Many of these methods can be directly incorporated into our framework.

3. Overview

In this section we discuss a simple *functional* model of human luminance perception which can be used to construct a tone mapping operator. Note that we do not attempt to build a full computational HVS model which may account for many more perceptual phenomena but will most likely be too complex for our purposes. In particular, no claim is made that HVS behaves the way we describe, but rather we demonstrate in this paper that a simple model we present successfully deals with the two most important tasks of luminance perception relevant for the tone mapping problem. These two tasks are providing a way of signalling absolute luminance levels and the use of local contrasts for information delivery. In our framework we use local adaptation level to signal absolute brightness and treat local contrast, which we define relative to this adaptation level, as the main information carrying quantity to be preserved in the final image.

We can define local contrast c at a pixel as

$$c(x,y) = L(x,y)/L_a(x,y) - 1 \qquad (1)$$

where L is pixel luminance and L_a is local adaptation level which we take to be just an average luminance over some neighborhood around pixel position (x,y). We present our way to estimate L_a in section 4.2 below. Note that a simple way to preserve contrast is to apply a linear mapping to the input image. This is well understood in computer graphics

where linear mappings are popular due to both their simplicity and contrast preservation property [27, 7]. Unfortunately, a *single* linear mapping applied to the whole image is often not sufficient. While it works for scenes with limited dynamic range, clamping of very low or very high luminance values is necessary for high dynamic range images and contrast preservation is therefore limited to regions in the image which are mapped to a well resolved range of display luminances. In addition, if the linear coefficient applied does not depend on the luminance level, the overall impression of brightness will be completely lost. Humans are not very good in judging the absolute luminances but making a sunlit scene look different from a moonlit one is certainly a desirable quantity of a tone mapping operator.

What we would like to do instead is to is to apply a *locally linear* mapping over much of the image with the coefficient applied depending on the particular image neighborhood so that the mapping would preserve details *throughout* the image. Suppose we compute local adaptation level $L_a(x,y)$ as the average luminance over some local image area (section 4.2). We can then apply a tone mapping function $TM(L)$ to $L_a(x,y)$ which would create a tone mapped adaptation image $TM(L_a(x,y))$ (section 4.1). The main purpose of this function is to compress a high dynamic range image to the display range while trying to convey the overall impression of brightness. Since we are not concerned with a more difficult problem of detail preservation at this stage, we found that the function can be made quite simple. Note that the only argument to $TM(L)$ is a luminance value. In most of our examples it is the adaptation luminance L_a, but in some cases applying $TM(L)$ to the image itself is sufficient to produce good results (see section 4.1).

Once the adaptation image and its tone mapped version are produced, we can state the requirement of local contrast preservation as $c_d(x,y) = c(x,y)$ (throughout the paper we will add subscript d to refer to display, i.e. tone mapped, quantities) and using equation 1 obtain the formula for the final mapping

$$L_d(x,y) = \frac{L(x,y)TM(L_a(x,y))}{L_a(x,y)} \qquad (2)$$

This operation "infuses" details back into the scene. Note that mapping 2 presents a locally linear (and, therefore, contrast preserving) transformation as long as $L_a(x,y)$ (and, as a consequence, $TM(L_a(x,y))$) is a slow varying function over much of the image. Since the real scene can have very high local contrasts not reproducible on a display device, this immediately shows that L_a can not be just an average over *constant* size neighborhood around the given pixel. Treating it as such leads to a well known problem of halos (inverse gradients) when, for example, a very bright light source extends a dark aura around it in a tone mapped image. Halos present a serious problem affecting even some of the most sophisticated tone mapping algorithms. We deal with this by adjusting the size of the region used to compute adaptation luminance. In smooth regions, this size is rather large while in

the neighborhood of a high contrast feature it shrinks to that of a pixel. The key observation here is that we do not need a particularly accurate value of adaptation level as long as it experiences a sharp jump at high contrast image features. In section 4.2 we show that this adjustment can be done by applying a simple constraint on the allowed output contrast signal associated with the neighborhood of a given size.

Preserving local contrast in a way suggested by equation 2 produces visually pleasing high contrast images. In many cases, visual satisfaction is the main criteria of success in computer graphics. However, in some applications, it might be more important to preserve *visible* contrast rather than the one given by equation 1. We can define visible contrast c_v as

$$c_v(x,y) = (L(x,y) - L_a(x,y))/TVI(L_a(x,y)) \qquad (3)$$

where $TVI(L_a)$ is threshold vs. intensity function which gives the value of just noticeable intensity difference for given adaptation level L_a. We can modify our procedure to introduce image details according to this new definition in which case the analog of equation 2 will be

$$L_d(x,y) = TM(L_a(x,y)) + \frac{TVI(TM(L_a(x,y)))}{TVI(L_a(x,y))}(L(x,y) - L_a(x,y)) \qquad (4)$$

See section 4.1 for the discussion of the choice of the appropriate TVI functions. Also note similarity between this formula and linear mapping suggested by Ward [27] in that the scaling coefficient is given by the ratio of display and world threshold sensitivities. However, since our framework does not attempt to fully follow the operation of HVS, the level of accuracy obtained might not be sufficient for some applications (a safety panel visibility estimation, for example) and a more complete computational model [15] might be more appropriate in these cases. Our main intention is to use enough relevant information about HVS to create a simple tone mapping procedure for non-critical computer graphics applications where visual satisfaction is of greater importance.

In summary, the following simple steps are performed:

- Separate out luminance channel and construct a Gaussian pyramid to facilitate the next step, see discussion below
- Compute local adaptation luminance $L_a(x,y)$ by contrast-limited neighborhood growing procedure, section 4.2.
- Apply tone mapping function $TM(L_a(x,y))$, equation 9.
- Compute final pixel luminance, equation 2
- Re-assemble color image by applying the scaling obtained for luminance to each of the original RGB channels.
- Perform gamma correction to obtain pixel values for display

Several possible modifications to this basis version of the algorithm are mentioned throughout the paper.

4. Details of Our Method

In the previous section we have separated the complex tone mapping problem into two parts dealing with two independent issues: compressing the dynamic range and preserving

local details. This allowed us to deal with each of the sub-problems using simpler algorithms. We will now present the details of the two key developments of this paper: derivation of the function $TM(L)$ and an algorithm to compute the local adaptation level $L_a(x,y)$. Even though the first operation we perform is the estimation of local adaptation luminance, it is more convenient to start with the description of the tone mapping function.

4.1. The tone mapping function $TM(L)$.

In this section we will deal with the problem of mapping the range of luminances in the input image to display luminances with no regard to preserving feature visibility (i.e. local contrast information) but rather making an attempt to account for deviations of HVS from a pure contrast measurer. This will allow us to be concerned only with given world luminance values and not with their spatial distribution. To construct the tone mapping function $TM(L)$ we first introduce a notion of *perceptual capacity* of a range of luminance values. The intuition is that human sensitivity to luminance changes, which is given by the TVI function, provides a natural scaling factor for a given small range of luminances ΔL. A range of luminances corresponding to, for example, five just noticeable differences can be considered to be of equal perceptual importance no matter where in the range of luminances it lies and we therefore define perceptual capacity as $\Delta L/TVI(L_a)$.

The question of the argument to pass to the TVI function is not a trivial one. Threshold sensitivity depends not on luminance itself but on some adaptation luminance L_a. The simplest way is to assume that there exists a single L_a value for the whole image. This is the assumption used for both world and display luminances by most tone mapping algorithms which introduce the concept of adaptation. However, studies show that adaptation is a highly localized phenomenon [12, 3]. Ignoring this fact by choosing a single L_a value is not a problem for display luminances which cover only a limited range. On the other hand, for a high contrast input image adaptation level can be very different in different parts of the image. Since we take adaptation luminance to be average luminance over some image neighborhood, it is better to take world luminance L itself as approximate measure of adaptation level. We therefore define an auxiliary capacity function C for a single world luminance value L as

$$C(L) = \int_0^L \frac{dl}{TVI(l)} \qquad (5)$$

which represents perceptual capacity of the world luminance range from 0 to L. With this definition, perceptual capacity of a (not necessarily small) world luminance range from L_1 to L_2 is simply $C(L_2) - C(L_1)$. Even though we estimate local L_a value as described in the next section, at present we do not see a clean way to incorporate this more detailed information into this expression since C ceases to be a simple "global" function of luminance in this case. For the display image we use a common approximation of single adaptation value L_{ad}:
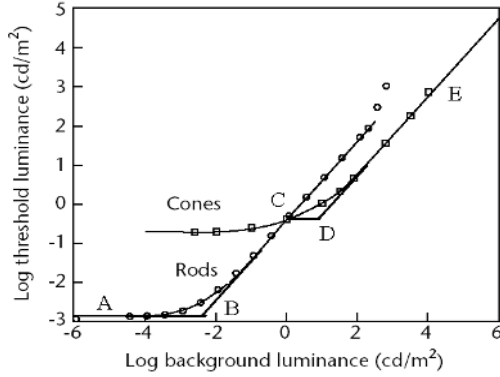
$$C_d(L) = \frac{L}{TVI(L_{ad})} \qquad (6)$$

Figure 1: The TVI function (adapted from [6]). To integrate equation 5 analytically, we use approximation by linear segments AB, BC, CD, and DE.

In addition to the simplicity of this formula, another argument in its favor is that we do not have complete information about viewing conditions for the tone mapped image which can significantly affect adaptation independently from the displayed image. So, equation 6 is probably the best we can do. In fact, our experiments with an analog of equation 5 for display luminances usually produced images which are too dark for typical viewing conditions when extra ambient illumination increases viewer's adaptation level. In a dark viewing environment using an analog of equation 5 might be a valid approach. Extending our framework for this case is straightforward and we skip the details.

We will also use different TVI functions for display and world perceptual capacities. In particular, TVI function for cone vision is appropriate for display luminances, since this is the system active at display intensities. For world luminances, either cone or rod TVI's may be more appropriate. A log-log space plot of these functions is shown on Figure 1. We follow Ward et.al. [11] who simply use a single function combined at the crossing point C. The mathematical form of this function is still too complex to perform analytic integration of equation 5 and some further adjustments to it are required to maintain simplicity. We found that a simplified TVI shape given by four linear (in log-log space) segments AB, BC, CD, and DE is sufficient for our purposes, but a more accurate approximation using more terms or functions which fit TVI curve shape better is certainly possible. With our approximation, world capacity function becomes

$$C(L) = \begin{cases} L/0.0014 & \text{if } L < 0.0034 \\ 2.4483 + log(L/0.0034)/0.4027 & \text{if } 0.0034 \leq L < 1 \\ 16.5630 + (L-1)/0.4027 & \text{if } 1 \leq L < 7.2444 \\ 32.0693 + log(L/7.2444)/0.0556 & \text{otherwise} \end{cases}$$

(7)

where L is expressed in cd/m^2.

The main principle we use to perform the mapping operation is that since we consider perceptual scales to be uniform,

world luminances are mapped into display ones according to their relative positions in corresponding perceptual scales. In other words, if world luminance L "used up" some fraction of total world perceptual range for a given image, the corresponding display luminance will be at the similar position in the total display perceptual range. This condition immediately leads to

$$\frac{C_d(L_d) - C_d(L_{mind})}{C_d(L_{maxd}) - C_d(L_{mind})} = \frac{C(L) - C(L_{min})}{C(L_{max}) - C(L_{min})}$$

(8)

We are interested in high contrast images which calls for using the full display range, so we set $L_{mind} = 0, L_{maxd} = LDMAX$ where $LDMAX$ is the maximum displayable luminance. A simple adjustment in cases when the total perceptual capacity of the input image is less than that of the full display range, is to find L_{maxd} (or L_{mind}) such that perceptual capacities of the input image and new display range are equal. To prevent a single potentially noisy pixel with very large or very low value from affecting this global operator, we can use max and min values of a blurred version of an image as L_{max} and L_{min}.

For any world luminance L Equation 8 uniquely defines corresponding display luminance (assuming invertible $C_d(L_d)$). With linear capacity for display defined by equation 6 and full display range used the tone mapping function becomes

$$L_d = TM(L) = LDMAX \frac{C(L) - C(L_{min})}{C(L_{max}) - C(L_{min})}$$

(9)

Note that TVI value for display adaptation luminance cancels out. Moreover, it is usually sufficient to know *relative* luminances $L_d/LDMAX$ with respect to display's maximum. This conveniently eliminates the question about particular values of L_{da} and $LDMAX$, getting exact answer for which is usually rather tricky. For versions which do require these values (for example, to estimate total perceptual capacity of the display luminances) we can use common values $LDMAX = 100cd/m^2$ and $L_{da} = 0.5LDMAX$. Straightforward adjustments for non-zero display luminance at zero pixel value ($L_{dmin} \neq 0$) are also possible if more detailed information about the display is available.

It is interesting to trace the relationship of our framework to earlier simple tone mapping functions. We can consider our approach to be a generalization of those since assumptions different from ours about world and display adaptation levels give a variety of popular functions. For example, choosing single adaptation values for both world and display luminances gives linear mapping similar to Ward's [27] while world TVI in the form of strict Weber's law ($TVI(l) \sim l$) gives either logarithmic mapping (with linear display $C(L)$) or an image-specific power function mapping (with Weber's law display TVI) similar to that of Tumblin and colleagues [24, 25]. It is further interesting that for mapping of many scenes with directly visible light sources the power given by the ratio of display to world total luminance capacities comes out to be around 1/2 to 1/3 in agreement with simple practice of taking cubic root of luminance as a tone

Figure 2: Tone mapping function $TM(L)$ applied directly to pixel luminances. Left: bathroom scene from [11], right: dimmed to 1/10000 of the original level.

mapping operation. Our function was constructed with no consideration of preserving image details. On some images, however, applying just this function directly to pixel luminances produces satisfactory results. An example is shown on Figure 2 where we present bathroom scene used in [11] under two different illumination levels to demonstrate the operation of $TM(L)$ alone.

4.2. Estimating local adaptation level.

Determining appropriate level of local adaptation is a difficult problem. We propose an approach based on balancing two opposing requirements which are faced by HVS: keeping the local contrast signal within reasonable bounds while maintaining enough information about image details. At this stage we will treat HVS as a pure contrast measurer, with its output signal being simply proportional to the local contrast. We assume that deviations from this behavior were satisfactorily taken into account by the construction of tone mapping function $TM(L)$ (through the shape of TVI functions). We compute adaptation level as the average luminance over some pixel neighborhood. Larger neighborhoods allow us to increase the range of pixel contrasts, better preserving image details. For low contrast images, this neighborhood corresponds to about one degree of visual field as determined in physiological experiments [12, 3]. However, making the neighborhood too big for high contrast images would overwhelm the system if it becomes necessary to signal contrasts which are too large. The tradeoff can be stated as the following simple rule: to determine local adaptation, we want to average over the largest neighborhood consistent with some limit on the generated contrast signal, or, equivalently, over the largest neighborhood which is still sufficiently uniform not to generate excessive contrast signals.

The maximum of pixel contrasts given by Equation 1 (or 3) could be used to estimate maximum output signal to place a bound upon, but this way is inefficient since we need to look at every pixel in the neighborhood every time we adjust adaptation level. We would also like the level of local adaptation to vary smoothly far from high contrast features which would be hard to achieve with single pixel sensitive measure suggested by equation 1. We therefore use a measure of neighborhood uniformity instead. The key observation is that for a uniform neighborhood, increasing its size does not affect its average value significantly. To characterize this property mathematically, we will recall that multi-scale contrast processing is a well known feature of HVS and it was suggested that an appropriate way to characterize the effects of local adaptation is to determine the band-limited local contrast lc at a given image location [17]. This quantity is usually calculated by taking the ratio of the difference of two lowpass filtered images (with widths related by a factor of 2) to one of them:

$$lc(s,x,y) = (G_s(L)(x,y) - G_{2s}(L)(x,y))/G_s(L)(x,y) \quad (10)$$

where $G_s(L)$ represent the result of applying gaussian filter of width s to the input image. Choosing G_s instead of G_{2s} in the denominator is somewhat arbitrary and both have been used in the literature [15, 1]. Our choice coincides with the definition of illumination-independent neural activity function from [1]. Note that using illumination dependent form (for example, dividing by $TVI(G(L))$, as in [15]) would not be appropriate since it assumes some knowledge of adaptation level. Fortunately, at this stage in our procedure we need to deal only with contrast measuring abstraction of the HVS.

We will now place a limit on this band-limited contrast associated with an adaptation neighborhood of particular size s as a perceptually advocated uniformity criteria. We will start from the smallest neighborhood (lowpass filter with width $s = 1$ in pixel units) and compute the associated band-limited contrast defined by equation 10. If it is too large, we assume the adaptation value at the particular location to be that of the pixel itself. Note that this corresponds to a pixel forming a high contrast feature in the image. Otherwise, we proceed to a larger neighborhood (increment filter width by one) until we either reach the maximum size allowed by our contrast ceiling or reach some maximum predetermined neighborhood size. This size should be set to about one degree of visual field as dictated by physiological measurements to achieve true resolution independence of the results, but since for many images information about the field of view is not readily available, a reasonable fixed value for the maximum considered filter size $s_{max} = 10$ pixels was found to be adequate. In addition to computational penalty, increasing this maximum size increases the potential of false intensity jumps in the adaptation image. Although our procedure uses discrete steps in s, we assume that $lc(s)$ is a continuous function of s and use linear interpolation to find s more accurately. The exact neighborhood uniformity criteria we

used for all examples in this paper is $|lc(s,x,y)| < 0.5$ and our procedure is effectively just an algorithm to find at every pixel the smallest $s \in [1;s_{max}]$ which solves the equation $|lc(s)| = 0.5$. Once the size s of a locally uniform neighborhood is obtained, L_a is simply taken to be $G_s(L)$ (we linearly interpolate $G_s(L)$ between integer s values). Manually adjusting threshold value (0.5) for particular image (or dividing by $G_{2s}(L)$ instead of $G_s(L)$ in equation 10) improves the results but for more objective evaluation of our method we demonstrate here only results obtained without such user intervention (except on Figure 4 below).

We caution again against interpretation of adaptation levels computed by our algorithm as something which directly corresponds to those present the human visual system. To the best of our knowledge, there is no experimental evidence that HVS chooses adaptation levels through the adaptive neighborhood growing process we describe. However, the results do show that this simple procedure is sufficient to perform our tone mapping operation for a variety of high contrast images.

4.3. The complete procedure.

Once the $L_a(x,y)$ image is computed, we apply the tone mapping function to it to obtain $TM(L_a(x,y))$ and follow equation 2 to compute the final tone mapped value $L_d(x,y)$ at each pixel. This will give a greyscale tone mapped image. The simplest way to deal with color which provides acceptable results in many cases is then to multiply each of the initial RGB components by $L_d(x,y)/L(x,y)$. This is the technique we use for all examples in this paper. Another sufficiently simple way to treat color as well as a way to account for the loss of visual acuity was presented by Ward et al [11]. We implemented this method with no modification in our framework. The only difference is that we can use our computed adaptation levels to judge the degree of color and acuity loss. Not surprisingly, the adjustments to the results are very similar to that shown in [11] and we do not present them here concentrating instead on the newly developed algorithm.

Note that we do not perform any further rescaling of the result, i.e. if our output values are beyond the display range, we simply clamp them. If this occurs at all, it usually happens in high luminance high contrast regions where the true contrast is not reproducible anyway and it is better to clamp the result rather than let other image regions to be affected by applying, for example, another linear mapping. Standard gamma correction procedure is applied as the final step to obtain pixel values for the display. Gamma correction for $\gamma = 2.4$ was applied for all images presented in this paper.

5. Results

We applied our procedure to a number of high contrast images, both rendered and acquired through high dynamic range photography. Rendered images were produced with

Radiance rendering system [28]. Photographs were made with the technique of Debevec and Malik [4]. Original images are publicly available on corresponding web pages. Foggy parking lot and computer desk images are curtesy of J. Tumblin, Northwestern University and S. Pattanaik and J.Ferwerda, Cornell University, respectively. Electronic versions of most of the results included in this paper are available on the author's web page.

To demonstrate non-linear mapping performed by $TM(L)$, we applied this function directly to pixel luminances for the bathroom scene used in [11]. We also apply it to the same image scaled by 1/10000 which approximately corresponds to moonlight illumination. The result is shown on Figure 2. Note that the tone mapped image is darker overall but not nearly in proportion to the scaling applied. Since for this dim image the perceptual capacity of the world luminances is lower than that of complete display range, only a fraction of it is used according to the procedure described in section 4.1.

For this image, details were well preserved by the tone mapping function, but this does not always occur. A more typical example is shown on Figure 3, a and Figure 4, left, where just applying $TM(L)$ to pixel luminances results in an image which is too dull, reminiscent of other simple range-compressing mappings. To correct this problem, we compute adaptation image and apply TM function to it instead of the original image. The result is shown on Figure 3, b. Note that our neighborhood growing procedure successfully captured high contrast features in the image and adaptation level changes sharply at these positions. This prevents halo artifacts from occurring. In the rest of the image it changes smoothly and its value is obtained by averaging over a neighborhood with size close to the maximum allowed one. This corresponds to the best conditions for signalling local details.

Final pass infuses details back into the image by applying equation 2. In effect, we multiply adaptation image by the detail image given by the ratio of pixel luminance to the corresponding local world adaptation level shown on Figure 3,c. At this stage, an alternative definition of contrast 3 can also be used. This is similar to the restriction introduced by Ward et. al. [11] but computes contrast with respect to geometric neighborhood rather than in luminance space. Results of these two alternative procedures are shown on Figure 3, d and e. Contrast is lowered in the last image in dark parts of the scene reflecting reduced contrast sensitivity. Note that even this image has more details than the image obtained with direct application of the tone mapping function, Figure 3, a. This is particularly clear for high luminance areas such as windows.

We found that in some cases the modified procedure emphasizes false edges which can be present in the adaptation image. A simple ad-hoc solution is to use a ceiling on the infused contrast by choosing a minimum of contrasts sug-

Figure 3: Operation of our algorithm. In reading order: a. application of $TM(L)$ to pixel luminances, b. computed adaptation image after tone mapping by TM, c. details image (midlevel grey corresponds to value one), d. final result using equation 2, e. same using visible contrast definition (equations 3 and 4), f. histogram-based tone mapping [11] result.

Figure 4: Tone mapping with different restrictions on local contrast $|lc(s,x,y)|$. Left: $TM(L)$ directly applied to pixel luminances (formally, $|lc| < 0$) Middle: $|lc| < 0.5$ - note some artifacts due to halos and false jumps in adaptation level. Right: $|lc| < 0.1$ - a better overall choice for this image. Corresponding detail views are also shown.

gested by equations 3 and 1. In all other examples we show results obtained with direct application of equation 2 since this is the preferred way of applying our algorithm for visual applications.

Although many algorithms are available, due to limited space we present here a single comparison of our approach with histogram based technique of Ward et. al. [11]. Motivations for this choice are uniformly high quality of the results it produces and availability of standard implementation within the Radiance system. Other recent state of the art methods are either significantly more complex [15, 25], address only part of the problem [24, 7] or have limited domain of applicability [23].

An image tone mapped with the histogram technique is shown on Figures 3, f. To concentrate on comparing algorithmically different parts of the two techniques, for these images we did not perform color correction, blurring, or veiling luminance calculations described in [11]. We did, however, use histogram adjustment based on human sensitivity function (TVI) since both methods attempt to account for corresponding effects but do it in different ways. In very bright areas of images (windows) our method shows more details than histogram mapping. We believe that this is due to the fact that we use a non-monotonic mapping. Compared with other methods, our technique does not create excessive high frequency content which is present in LCIS images [25] and seems to avoids halos better than [2, 15]. A large set of tone mapping results for the image on Figure 3 is presented in [13] showing that loss of fine details in the window region is characteristic to most techniques. A collection of tone mapped high contrast images is shown on Figure 5 (Color plate).

Computational cost of our algorithm is moderate. We found that using a pyramid of filtered images $G_s(L)(x,y)$ with width s doubling on each level and linearly interpolating between them similar to mip-mapping [29] is sufficient for our purposes. Even with this approach, initial filtering is the most computationally expensive part in our current implementation of the algorithm. In our unoptimized C++ implementation we compute image pyramid by direct convolution with gaussian filters of different widths which takes 11 seconds on a 1.2 GHz Pentium IV PC for the Stanford memorial church image (resolution 512x768 pixels). A significantly faster alternative for large images is to apply a Fourier transform and perform the filtering through multiplication in Fourier domain, as done in [13]. Once the pyramid is computed, it takes 0.89 seconds to perform adaptation level computation for the the church image. Direct application of $TM(L)$ to pixel luminances can be performed at several frames per second for this image.

Our method has several important limitations. First of all, the adaptation image computed by our simple procedure can sometimes create jumps in adaptation level away from high contrast image features. This produces artifacts similar to those visible around the light in the middle image on Figure 4. These artifacts are usually mild and can be eliminated by an adjustment of neighborhood uniformity criterion as mentioned in section 4.2 (see figure 4). Second, our procedure is not well suited for images where local contrast is uniform everywhere since it terminates neighborhood growth at locally high contrast positions and it has difficulty correctly finding these places in a uniform image. An example of such an image is the well known Macbeth color checker scene used in [7, 15]. Fortunately, as [7] demonstrated, a simple linear mapping is adequate in these cases. Finally, there is some discussion in the literature on the value of residual halos to accentuate the perception of high contrast image features [2, 25, 22, 23]. Our method seems to avoid halos in most cases but a mild form of them can be visible in some situations (see Figures 3 and 5). We believe that some form of halos is most likely unavoidable in a non-monotonic mapping procedure. Again, adjustments to the neighborhood uniformity criteria can either remove halos completely at the cost of somewhat decreased visual detail or increase their magnitude if desired.

6. Conclusion

In this paper we presented a simple tone mapping algorithm designed for high contrast images. It is based on a multipass technique which first estimates local adaptation level, applies a simple tone mapping function to it and then puts back image details. This reflects two main goals of a tone mapping algorithm: preserving image details and providing enough absolute brightness information in a low dynamic range tone mapped image. In addition to this functional separation, contributions of this paper include a novel technique to estimate local adaptation level and development of a new tone mapping function using the concept of perceptual capacity of a range of luminances. We believe that moderate computational expense and simplicity of implementation combined with good visual results make our algorithm useful for a variety applications.

Our framework is modular and its main parts are independent from one another. This allows to use, for example, a more accurate TVI function with no change to adaptation estimation algorithm. Similarly, if a better way to compute local adaptation luminance is found, it can be substituted for our simple procedure. In particular, we observed that a single global adjustment to the neighborhood uniformity criterion by the user can improve the results. It would be beneficial to investigate more local adjustments and to deal with this issue automatically. Other potential improvements might include incorporating more characteristics of HVS into the framework. For example, the simple color handling procedure we use does not take into account the effects of chromatic adaptation which is a significant effect for the desk image on Figure 5. Adaptation color which would be needed to account for this effect seems to correspond to a neighborhood much larger than that for luminance adaptation we deal with.
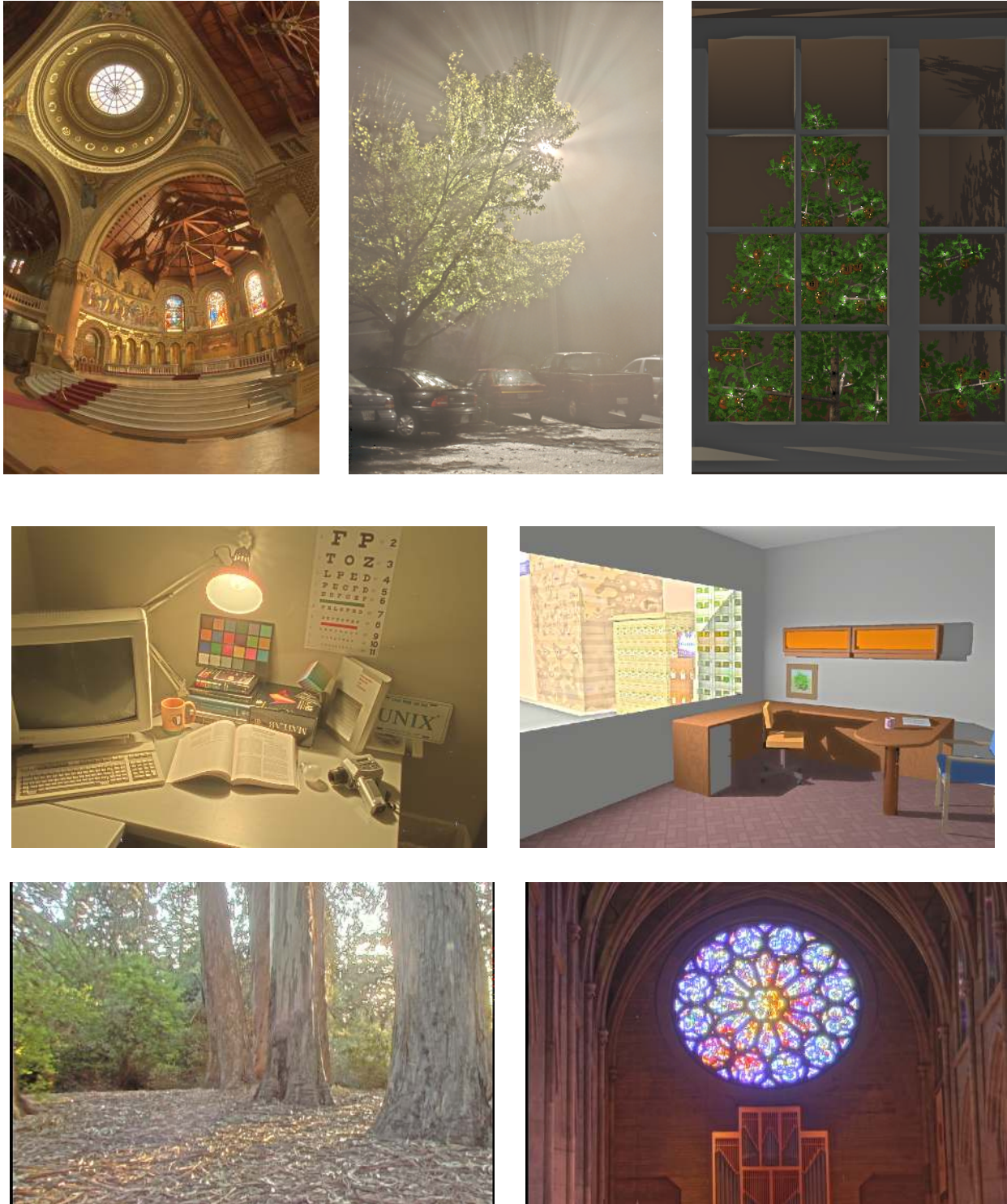
Figure 5: More examples of tone mapped high contrast images.

Acknowledgements

References

1. Frans J. J. Blommaert and Jean-Bernard Martens. An object-oriented model for brightness perception. *Spatial Vision*, 5(1):15–41, 1990.

2. K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface '93*, pages 245–253, 1993.

3. B. Cleland and A. Freeman. Visual adaptation is highly localized in the cat's retina. *J. Physiology*, 404:591–611, 1988.

4. Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. *Proceedings of SIGGRAPH 97*, pages 369–378, 1997.

5. Fredo Durand and Julie Dorsey. Interactive tone mapping. *11th Eurographics Workshop on Rendering*, pages 219–230, 2000.

6. James Ferwerda. Elements of early vision for computer graphics. *IEEE Computer Graphics and Applications*, 21(5):22–33, 2001.

7. James A. Ferwerda, Sumant N. Pattanaik, Peter Shirley, and Donald P. Greenberg. An adaptation model for realistic image sysnthesis. In *SIGGRAPH '96 Proceedings*, pages 249–258.

8. Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 2002.

9. D. J. Jobson, Z. Rahman, and G. A. Woodell. A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6:965–976, 1997.

10. Edwin H. Land. The retinex theory of color vision. *The Scientific American*, (12):108–128, December 1977.

11. Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, 1997.

12. P. Moon and D. Spencer. The visual effect of non-uniform surrounds. *Journal of the Optical Society of America*, 35:233–248, 1945.

13. E. Reinhard M.Stark, P.Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *Proceedings of SIGGRAPH*, 2002.

14. Stephen E. Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge, MA, 1999.

15. Sumant N. Pattanaik, James A. Ferwerda, Mark Fairchild, and Donald P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *SIGGRAPH '98 Proceedings*, pages 287–298.

16. Sumanta N. Pattanaik, Jack E. Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for realistic image display. *Proceedings of SIGGRAPH*, pages 47–54, 2000.

17. E. Peli. An object-oriented model for brightness perception. *Journal of the Optical Society of America, A*, 7:2032–2040, 1990.

18. A. Scheel, M. Stamminger, and Hans-Peter Seidel. Tone reproduction for interactive walkthroughs. *Proceedings of Eurographics Conference*, pages 301–312, August 2000.

19. C. Schlick. Quantization techniques for visualization of high dynamic range pictures. *5th Eurographics Workshop on Rendering*, pages 7–20, 1995.

20. R. Shapley and C. Enroth-Cugell. Visual adaptation and retinal gain-controls. *Progress in Retinal and Eye Research*, 3:263–346, 1984.

21. Greg Spencer, Peter S. Shirley, Kurt Zimmerman, and Donald P. Greenberg. Physically-based glare effects for digital images. *Proceedings of SIGGRAPH*, pages 325–334, 1995.

22. T. Tanaka and N. Ohnishi. Painting-like image emphasis based on human vision svstems. *Proceedings of Eurographics Conference*, pages 253–260, 1997.

23. Jack Tumblin, Jessica K. Hodgins, and Brian K. Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, 1999.

24. Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, November 1993.

25. Jack Tumblin and Greg Turk. Lcis: A boundary hierarchy for detail-preserving contrast reduction. *Proceedings of SIGGRAPH 99*, pages 83–90.

26. H. Wallah. Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, 38:310–324, 1948.

27. Greg Ward. A contrast-based scalefactor for luminance display. In Paul Heckbert, editor, *Graphics Gems IV*, pages 415–421. Academic Press, Boston, 1994.

28. Gregory J. Ward. The RADIANCE lighting simulation and rendering system. *Computer Graphics*, 28(2):459–472, July 1994. ACM Siggraph '94 Conference Proceedings.

29. Lance Willams. Pyramidal parametrics. *Computer Graphics*, 17(3), July 1983. ACM Siggraph '83 Conference Proceedings.