CPSC 4040/6040
Computer Graphics
Images

Joshua Levine
levinej@clemson.edu
Lecture 15
Tone Mapping
Oct. 8, 2015

Slide Credits:
Frédo Durand
Alexei Efros
Agenda

- PA04 questions?
- Final Project Proposals — coming up after Fall Break
Last Time
The World is a High Dynamic Range (HDR)
Multiple exposure photography

- Sequentially measure all segments of the range

Real world

\[10^{-6} \quad \text{High dynamic range} \quad 10^6\]

Picture

\[10^{-6} \quad \text{Low contrast} \quad 10^6\]
Response Curve

- Radiance is unknown, fit to find a smooth curve

Assuming unit radiance for each pixel

After adjusting radiances to obtain a smooth response curve

Slide stolen from Alyosha Efros who stole it from Paul Debevec
HDR File Formats
Overview of HDR Encoding

• RGBE/XYZE
  • 24 bits/pixels as usual, plus 8 bit of common exponent
  • Introduced by Greg Ward for Radiance (light simulation)
  • Enormous dynamic range

• OpenEXR
  • By Industrial Light + Magic, also standard in graphics hardware
  • 16bit per channel (48 bits per pixel) 10 mantissa, sign, 5 exponent
  • Fine quantization (because 10 bit mantissa), only 9.6 orders of magnitude

• Others: 48-/96-bit TIFF, 32-bit LogLuv TIFF, PPM Extension (32-bit/channel), JPEG200 Extension (can be lossy)

Radiance RGBE and XYZE

- Simple format with free source code
- 8 bits each for 3 mantissas and 1 exponent
- 76 orders of magnitude in 1% steps
- Run-length encoding (20% avg. compression)
- RGBE format does not cover visible gamut
- Dynamic range at expense of accuracy
- Color quantization not perceptually uniform
Radiance RGBE Format (.pic, .hdr)

32 bits/pixel

<table>
<thead>
<tr>
<th>Red</th>
<th>Green</th>
<th>Blue</th>
<th>Exponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>145</td>
<td>215</td>
<td>87</td>
<td>149</td>
</tr>
<tr>
<td>145</td>
<td>215</td>
<td>87</td>
<td>103</td>
</tr>
</tbody>
</table>

(145, 215, 87, 149) = (145, 215, 87, 103) =

(145, 215, 87) * 2^(149-128) = (145, 215, 87) * 2^(103-128) =

1190000 1760000 713000 0.00000432 0.00000641 0.00000259

ILM OpenEXR Format

• 16-bit/primary floating point (sign, 1; exponent, 5; mantissa, 10)

• 9.6 orders of magnitude in 0.1% steps

• Additional order of magnitude near black

• Wavelet compression of about 40%

• Negative colors and full gamut RGB

• Alpha and multichannel support

• Open Source I/O library released Fall 2002

• Slower to read and write
ILM’s OpenEXR (.exr)

- 6 bytes per pixel, 2 for each channel, compressed

  ![Sign Exponent Mantissa Diagram]

- Several lossless compression options, 2:1 typical
- Compatible with the “half” datatype in NVidia's Cg
- Supported natively on GeForce FX and Quadro FX

- Available at [http://www.openexr.net/](http://www.openexr.net/)
HDR File Performance Compared

See http://www.anyhere.com/gward/hdrenc/hdr_encodings.html
Tone Mapping Overview
The Radiance Map
The Radiance Map

Linear mapping to display is insufficient!

Linearly scaled to display device
Chapter 4: Tonemapping effects. Everything was simulated correctly according to the books, but it didn't really look so convincing until game makers started to master tonemapping.

In this real-time context, the tonemapping operator can be considered the virtual camera the game is shot with. It has to dynamically meter the exposure as the player travels across varying lighting conditions and constantly readjust exposure accordingly. But that is only half the story. A simple exposure adjustment doesn't make it cinematic enough, and that is where most early games failed. Nowadays tonemapping operators also take on the role of virtual film stock. These algorithms use complex rule sets and custom tone curves to keep some indication of detail visible in shadows and highlights. The way the tones in between are distributed and how that virtual film stock responds to colors are nowadays crucial elements to define the look of a game. That can be the gritty bleached look of a militaristic ego shooter or the award-winning picturesque look of the *Uncharted* adventure series—in any case, the tonemapping style is meticulously crafted to express a certain mood and support the overall art direction. High-end games even contain a data track, where the game artists can animate exposure and other tonemapping parameters along with the camera motion, so it becomes a truly handcrafted experience.

Typically, games use global tonemappers for performance reasons. Local enhancements will surely come up in the future, but they are not urgently necessary because game makers have total control over lighting and textures anyway. They can simply tweak their world to have a lower contrast to begin with. However, these real-time tonemappers are still capable of adapting dynamically to changing lighting conditions and applying some very controlled looks. That makes them work well for batch processing, time lapse, and HDR video.

**Figure 4-23:** Tonemapping plays an essential role in real-time rendering. It defines the look and enhances the cinematic appearance of high-end games like *Uncharted 3* by Naughty Dog.
Problem: How should we map scene radiances (up to 1:100,000) to display radiances (only around 1:100) to produce a satisfactory image?

Goal: match limited contrast of the display medium while preserving details.
Visual Matching

• We do not need to reproduce the true radiance as long as it gives us a visual match.
Recall: Eyes and Dynamic Range

• We're sensitive to contrast (multiplicative)
  • A ratio of 1:2 is perceived as the same contrast as a ratio of 100 to 200
  • Use the log domain as much as possible
• Eyes are not photometers
  • Dynamic adaptation (very local in retina)
  • Different sensitivity to spatial frequencies

Headlights are ON in both photos
Weber’s Law

Just Noticeable Difference (JND) \[
\frac{\Delta I}{I} = K_w
\]

Weber fraction (constant!)

• We can detect \(~0.5\%\) changes in intensity
  • The perceived minimum increment is the just noticeable difference

• True for luminance as well as other things (length, weight, sound, etc.)

• Takeaway: The amount of difference that we can detect is relative to the context between the two things, not an absolute quantity
  • If you change the context, you change the increment
Chapter 4: Tone Mapping

Haze removal

Landscape shots often have a lot of atmospheric haze, especially when a vista stretches far into the distance. Sometimes that may be a desired effect, but more often you would rather see clearly all the way to the horizon. It's not exactly what they're designed for, but local detail enhancement algorithms just happen to be well suited for enhancing clarity and punching through this haze. And that works even though the overall scene contrast typically doesn't justify a real HDR shoot—a single RAW is normally all you need for this application.

Just like natural toning, this is actually a reconstruction job. It's very tempting to over-shoot, especially when you start to discover more things than you could actually see with your naked eye. But ultimately, you want this to be a natural image that looks like it was shot on a clearer day. It helps to mentally set this goal up front.

Since you're boosting contrast in the horizon area quite a bit, this process needs very clean and noise-free image data to begin with. Otherwise the image will break up and look manipulated. Shooting RAW is essential, at the highest possible bit depth and the lowest possible ISO. Browse back to section 3.2.9 for some hints on pseudo-HDR conversion. Even if the hazy horizon will fit in a single RAW, it doesn't hurt to combine multiple exposures just for the sake of averaging out the last bit of noise.

Texture extraction

When people think about using HDRI for computer graphics and VFX, they often overlook the applications of tone mapping for preparing texture images. The importance of good reference before rendering can be directly linked to the process of tone mapping for preparing textures.

Figure 4-21: Atmospheric haze is easily wiped away with tonemapping techniques. Recovering all these details from the murky shadows is exactly where local TMOs excel.
Common Artifacts to Avoid

• Tone Reversal: large image areas have swapped brightness
• Noise boost: noise emphasized in regions which were smooth
• Oversaturation: increasing color beyond what is natural
• Halos: bright glows around dark edges
• Flattening: over compressing the dynamic range into something less than the display can do
• Webbing: Division of the image into cells based on natural lines
Global tonemappers are immune to halos, and most local tonemappers nowadays have a dedicated halo reduction built in. In some rare cases, halos can also benefit an image by adding a strong separation between foreground and background.

**Tone reversal**
This term means that large image areas have swapped overall brightness levels; for example, the sky may suddenly appear much darker than the foreground. Some people call it contrast reversal, and it can also happen on a smaller scale. An obnoxious example is a dark spot in the middle of a lightbulb. Tone reversal is often accompanied by halos as the result of excessive noodling with the image. Global tonemappers are immune, and when using local TMOs, thinking in zones is the best way to keep a thumb on tone reversal. Used with caution in the context of artistic toning, tone reversal can lead to very dramatic results and may even be a desired effect.

**Noise boost**
Local tonemappers are like watchdogs, trained to snoop out details everywhere. Especially when applied without smoothing or with a small radius value, they sometimes confuse residual noise with valuable details worth exaggerating. Smooth gradients like clouds, clear skies, and flat walls are often affected and start looking grainy. One way to prevent this is to start off with a cleaner HDR image. Shoot at the lowest ISO and consider shooting more exposures with smaller EV steps so the noise will be averaged away during the merging process. Another way is to treat the affected areas afterward with a noise removal filter. Noiseware, Topaz DeNoise, Neat Image, and Noise Ninja are all great solutions for that.

---

**Figure 4-33:** Compilation of typical tonemapping artifacts, each one driven to the extreme.
Simple Approaches to Tone Mapping
Can we just scale? Maybe!

- For a color image, try to convert the input (world) luminance \(L_w\) to a target display luminance \(L_d\):

  \[
  \begin{bmatrix}
  R_d \\
  G_d \\
  B_d
  \end{bmatrix}
  =
  \begin{bmatrix}
  L_d \frac{R_w}{L_w} \\
  L_d \frac{G_w}{L_w} \\
  L_d \frac{B_w}{L_w}
  \end{bmatrix}
  \]

- This type of scaling works (sometimes). In particular, it works best in the log and/or exponential domains (because of Weber's law)
Recall: Gain, Bias, and Gamma

\[ C_{out} = (\alpha C_{in} + \beta)\gamma \]

- \(\alpha\) is known as gain (exposure)
- \(\beta\) is known as bias (offset)
- \(\gamma\) maps us to a non-linear curve (gamma correction)
Normalization

- Typical scene has 1:10,000 contrast, display has 1:100
- Simplest reduction destroys midtones
Gamma Compression

• \( C_{\text{out}} = C_{\text{in}}^{\gamma} \), where \( 0 < \gamma < 1 \) applied to each R,G,B channel

• Colors are washed out, why?
Gamma Compression on Intensity

• Colors ok, but details in intensity are blurry
Histogram Equalization

- Questions to answer: how many bins?
- Better to equalize the log of intensity

**Figure 3.7.** An example of the Histogram adjustment by Larson et al. [110] to the IDL HDR image. (a) The histogram of the HDR image. (b) The tone mapped image.
Using Local Operators
Abstract

A classic photographic task is the mapping of the potentially high dynamic range of real world luminances to the low dynamic range of the photographic print. This tone reproduction problem is also faced by computer graphics practitioners who map digital images to a low dynamic range print or screen. The work presented in this paper leverages the time-tested techniques of photographic practice to develop a new tone reproduction operator. In particular, we use and extend the techniques developed by Ansel Adams to deal with digital images. The resulting algorithm is simple and produces good results for a wide variety of images.

CR Categories: I.4.10 [Computing Methodologies]: Image Processing and Computer Vision—Image Representation

Keywords: Tone reproduction, dynamic range, Zone System.

1 Introduction

The range of light we experience in the real world is vast, spanning approximately ten orders of absolute range from star-lit scenes to sun-lit snow, and over four orders of dynamic range from shadows to highlights in a single scene. However, the range of light we can reproduce on our print and screen display devices spans at best about two orders of absolute dynamic range. This discrepancy leads to the tone reproduction problem: how should we map measured/simulated scene luminances to display luminances and produce a satisfactory image?
Ansel Adams
The Zone System

• Formalism to talk about exposure, density

• Zone = intensity range, in powers of two

• In the scene, on the negative, on the print

Ansel Adams
## The Zones

<table>
<thead>
<tr>
<th>Zone</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Solid black; the same as the film rebate</td>
<td><img src="image1" alt="Image" /></td>
</tr>
<tr>
<td>I</td>
<td>Nearly black; just different from Zone 0</td>
<td><img src="image2" alt="Image" /></td>
</tr>
<tr>
<td>II</td>
<td>The first hint of texture</td>
<td><img src="image3" alt="Image" /></td>
</tr>
<tr>
<td>III</td>
<td>Textured shadow; the first recognizable shadow detail</td>
<td><img src="image4" alt="Image" /></td>
</tr>
<tr>
<td>IV</td>
<td>Average shadow value on Caucasian skin, foliage and buildings</td>
<td><img src="image5" alt="Image" /></td>
</tr>
<tr>
<td>V</td>
<td>Middle grey: the pivot value; light foliage, dark skin</td>
<td><img src="image6" alt="Image" /></td>
</tr>
<tr>
<td>VI</td>
<td>Caucasian skin, textured light grey; shadow on snow</td>
<td><img src="image7" alt="Image" /></td>
</tr>
<tr>
<td>VII</td>
<td>Light skin; bright areas with texture, such as snow in low sunlight</td>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td>VIII</td>
<td>Highest zone with any texture</td>
<td><img src="image9" alt="Image" /></td>
</tr>
<tr>
<td>IX</td>
<td>Pure untextured white</td>
<td><img src="image10" alt="Image" /></td>
</tr>
</tbody>
</table>
Zone System

Dynamic range = 15 scene zones

<table>
<thead>
<tr>
<th>$2^x L$</th>
<th>$2^{x+1} L$</th>
<th>$2^{x+2} L$</th>
<th>$2^{x+3} L$</th>
<th>$2^{x+4} L$</th>
<th>$2^{x+5} L$</th>
<th>$2^{x+6} L$</th>
<th>$2^{x+7} L$</th>
<th>$2^{x+8} L$</th>
</tr>
</thead>
</table>

Middle grey maps to Zone V
The Zone System

• You decide to put part of the system in a given zone

• Decision: exposure, development, print
Zones Example

Figure 4-32: Zone chart of the assignments for this image.
Dodging and Burning
Limited dynamic range can be good!

- W. Eugene Smith photo of Albert Schweitzer. 5 days to print!
- Things can be related because the intensity is more similar
- Balance, composition
Dodging and burning

• During the print process, selectively hide parts of the print during exposure
• Dodging (left) block light from regions to make darker
• Burning (right) overexpose regions to make brighter

From The Master Printing Course, Rudman
Dodging and Burning

*Dodging* holds back light during the basic printing exposure to lighten an area.

*Burning* adds light after the basic exposure to darken an area.

From Photography by London et al.
Dodging and burning

- Downside: Must be done for every single print

Straight print  After dodging and burning
Dodging & Burning is Difficult

A. The straight work print without additional burning-in.

B. This print shows the result of trying to mask off the foreground by using a moving card. An even more obvious light band will appear in the sky if the card is not kept moving.

C. In order to remove the light band in fig B, the mask has been lowered. This, however, has caused parts of the horizon to become black.

D. The halo effect, here deliberately exaggerated, resulted from dodging the stones during the second exposure while burning-in the sky.

E. It is very difficult to cut a dodging card with precision, especially for a relatively small print like this. As a result, parts of the sky at the horizon are white, although careful spotting can disguise this problem when it is small. But parts of the mid-grey hill tops have gone jet black, which is less easy to rectify.

Source: Rudman
Digital Zoning
Global operator

\[
\bar{L}_w = \exp \left( \frac{1}{N} \sum_{x,y} \log(\delta + L_w(x, y)) \right)
\]

Approximation of scene’s key (how light or dark it is). Map to 18% of display range for average-key scene

User-specified; high key or low key

\[
L_m(x, y) = \frac{a}{\bar{L}_w} L_w(x, y) \quad L_d(x, y) = \frac{L_m(x, y)}{1 + L_m(x, y)}
\]
Global operator

It seldom reaches 1 since the input image does not have infinitely large luminance values.

\[ L_d(x, y) = \frac{L_m(x, y) \left(1 + \frac{L_m(x, y)}{L_{white}^2(x, y)}\right)}{1 + L_m(x, y)} \]

\( L_{white} = 0.5 \quad 1.0 \quad 1.5 \quad 3 \quad \infty \)

\( L_{white} \) is the smallest luminance to be mapped to 1.
Dodging and burning (local operators)

- Area receiving a different exposure is often bounded by sharp contrast
- Find largest surrounding area without any sharp contrast

\[ L_s^{\text{blur}}(x, y) = L_m(x, y) \otimes G_s(x, y) \]

\[ V_s(x, y) = \frac{L_s^{\text{blur}}(x, y) - L_{s+1}^{\text{blur}}(x, y)}{2^\phi a/s^2 + L_s^{\text{blur}}} \]

\[ s_{\text{max}} : \left| V_{s_{\text{max}}}(x, y) \right| < \varepsilon \]
Dodging and burning (local operators)

\[ L_d(x, y) = \frac{L_m(x, y)}{1 + L_{\text{blur}}^{\text{max}}(x, y)} \]

- A darker pixel (smaller than the blurred average of its surrounding area) is divided by a larger number and become darker (dodging)
- A brighter pixel (larger than the blurred average of its surrounding area) is divided by a smaller number and become brighter (burning)
- Both increase the contrast
The luminance of a dark pixel in a relatively bright region will saturate, while the large, dark rings around luminance steps will form.

This function constitutes our local dodging-and-burning operator.

In summary, by automatically selecting an appropriate neighborhood for each pixel we effectively implement a pixel-by-pixel version of our global tone reproduction operator of Equation 3. This method is insensitive to edge artifacts normally associated with the thresholding operation used for scale selection. The key value setting is determined on a per image basis, while the threshold is set as a function of the pixel’s contrast relative to the surrounding area. For pixels with contrast not close to the luminance of the surrounding area, this operator will decrease the display luminance for very bright pixels, our local operator is not guaranteed to keep these pixels bright. For pixels with contrast close to the luminance of the surrounding area, this operator will darken or brighten certain areas in the final print.

Increasing the threshold \( \epsilon \) forces the appropriate scale \( s \) to be chosen for each pixel. Decreasing \( \epsilon \) slows for large images. Because of the normalization by \( \epsilon \), these convolutions can also be efficiently computed in image operations, and work by manipulating the scale that would be chosen for each pixel.
Dodging and burning

Simple operator

Dodging—and—burning
Bilateral Filtering
(slides from Frédo Durand)
Fast Bilateral Filtering
for the Display of High-Dynamic-Range Images

Frédo Durand and Julie Dorsey
Laboratory for Computer Science, Massachusetts Institute of Technology

Abstract

We present a new technique for the display of high-dynamic-range images, which reduces the contrast while preserving detail. It is based on a two-scale decomposition of the image into a base layer, encoding large-scale variations, and a detail layer. Only the base layer has its contrast reduced, thereby preserving detail. The base layer is obtained using an edge-preserving filter called the bilateral filter. This is a non-linear filter, where the weight of each pixel is computed using a Gaussian in the spatial domain multiplied by an influence function in the intensity domain that decreases the weight of pixels with large intensity differences. We express bilateral filtering in the framework of robust statistics and show how it relates to anisotropic diffusion. We then accelerate bilateral filtering by using a piecewise-linear approximation in the intensity domain and appropriate subsampling. This results in a speed-up of two orders of magnitude. The method is fast and requires no parameter setting.


Keywords: image processing, tone mapping, contrast reduction, edge-preserving filtering, weird maths

1 Introduction

As the availability of high-dynamic-range images grows due to advances in lighting simulation, e.g. [Ward 1994], multiple-exposure photography [Debevec and Malik 1997; Madden 1993] and new sensor technologies [Mitsunaga and Nayar 2000; Schechner and Nayar 2001; Yang et al. 1999], there is a growing demand to be able to display these images on low-dynamic-range media. Our visual system can cope with such high-contrast scenes because most of the adaptation mechanisms are local on the retina.

There is a tremendous need for contrast reduction in applications such as image-processing, medical imaging, realistic rendering, and digital photography. Consider photography for example. A major aspect of the art and craft concerns the management of contrast via e.g. exposure, lighting, printing, or local dodging and burning [Adams 1995; Rudman 2001]. In fact, poor management of light – under- or over-exposed areas, light behind the main character, etc. – is the single most-commonly-cited reason for rejecting photographs. This is why camera manufacturers have developed sophisticated exposure-metering systems. Unfortunately, exposure only operates via global contrast management – that is, it recenters the intensity window on the most relevant range. If the range of intensity is too large, the photo will contain under- and over-exposed areas (Fig. 1, rightmost part).

Our work is motivated by the idea that the use of high-dynamic-range cameras and relevant display operators can address these issues. Digital photography has inherited many of the strengths of film photography. However it also has the potential to overcome its limitations. Ideally, the photography process should be decomposed into a measurement phase (with a high-dynamic-range output), and a post-process phase that, among other things, manages the contrast. This post-process could be automatic or user-controlled, as part of the camera or on a computer, but it should take advantage of the wide range of available intensity to perform appropriate contrast reduction.

In this paper, we introduce a fast and robust operator that takes a high-dynamic-range image as input, and compresses the contrast while preserving the details of the original image, as introduced by Tumblin [1999]. Our operator is based on a two-scale decomposition of the image into a base layer (large-scale features) and a detail layer.
Oppenheim 1968, Chiu et al. 1993

• Reduce contrast of low-frequencies
• Keep mid and high frequencies

Low-freq.  Reduce low frequency

High-freq.

Color
The halo nightmare

- For strong edges
- Because they contain high frequency

Low-freq.  
Reduce low frequency

High-freq.

Color
Our approach

• Do not blur across edges
• Non-linear filtering
Start with Gaussian filtering

• Here, input is a step function + noise

\[ J = f \otimes I \]
Gaussian filter as weighted average

- Weight of $\xi$ depends on distance to $x$

$$J(x) = \sum_{\xi} f(x, \xi) I(\xi)$$
The problem of edges

• Here, \( I(\xi) \) “pollutes” our estimate \( J(x) \)
• It is too different

\[
J(x) = \sum_{\xi} f(x, \xi) I(\xi)
\]

output \hspace{1cm} \hspace{1cm} input
Principle of Bilateral filtering

[Tomasi and Manduchi 1998]

• Penalty $g$ on the intensity difference

$$J(x) = \frac{1}{k(x)} \sum_{\xi} f(x, \xi) \ g(I(\xi) - I(x)) \ I(\xi)$$
Bilateral filtering

[Tomasi and Manduchi 1998]

• Spatial Gaussian f

\[
J(x) = \frac{1}{k(x)} \sum_{\xi} f(x,\xi) g(I(\xi) - I(x)) I(\xi)
\]
Bilateral filtering

[Tomasi and Manduchi 1998]

- Spatial Gaussian $f$
- Gaussian $g$ on the intensity difference

$$J(x) = \frac{1}{k(x)} \sum_{\xi} f(x,\xi) g(I(\xi) - I(x))I(\xi)$$
Normalization factor

[Tomasi and Manduchi 1998]

\[ k(x) = \sum_{\xi} f(x,\xi) \quad g(I(\xi) - I(x)) \]

\[ J(x) = \frac{1}{k(x)} \sum_{\xi} f(x,\xi) \quad g(I(\xi) - I(x)) \quad I(\xi) \]
Bilateral filtering is non-linear

[Tomasi and Manduchi 1998]

- The weights are different for each output pixel

\[
J(x) = \frac{1}{k(x)} \sum_{m} f(x, \xi) \cdot g(I(\xi) - I(x)) \cdot I(\xi)
\]
Log domain

- Very important to work in the log domain
- Recall: humans are sensitive to multiplicative contrast
- With log domain, our notion of “strong edge” always corresponds to the same contrast
Recap

Input HDR image

Intensity

Fast Bilateral Filter IN LOG

Color

Large scale

Detail

detail = input log - large scale

Reduce contrast

Preserve!

Output

Large scale

Detail

Color
Lec16 Required Reading
• None. Instead, begin thinking about project proposals.