**Tone mapping**

Digital Visual Effects  
*Yung-Yu Chuang*

with slides by Fredo Durand, and Alexei Efros

---

**Tone mapping**

- How should we map scene luminances (up to $1:100,000$) to display luminances (only around $1:100$) to produce a satisfactory image?

  Linear scaling?, thresholding?

<table>
<thead>
<tr>
<th>Real world radiance</th>
<th>Dynamic range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-6}$</td>
<td>$10^6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Display intensity</th>
<th>Pixel value 0 to 255</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-6}$</td>
<td>$10^6$</td>
</tr>
</tbody>
</table>

CRT has 300:1 dynamic range

---

**The ultimate goal is a visual match**

- Dynamic range along the visual pathway is only around $32:1$.
- The key is adaptation

---

**Eye is not a photometer!**

- We do not need to reproduce the true radiance as long as it gives us a visual match.
Eye is not a photometer!

Are the headlights different in two images? Physically, they are the same, but perceptually different.

How humans deal with dynamic range

- We’re more sensitive to contrast (multiplicative)
  - A ratio of 1:2 is perceived as the same contrast as a ratio of 100 to 200
  - Makes sense because illumination has a multiplicative effect
  - Use the log domain as much as possible
- Dynamic adaptation (very local in retina)
  - Pupil (not so important)
  - Neural
  - Chemical
- Different sensitivity to spatial frequencies

We are more sensitive to contrast

- Weber’s law

Just-noticeable Difference (JND)

\[
\frac{\Delta I_b}{I_b} \approx 1\%
\]

Preliminaries

- For color images

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

- Log domain is usually preferred.
**HDR Display**

- Once we have HDR images (either captured or synthesized), how can we display them on normal displays?

**Sunnybrook HDR display**

- Use Bright Source + Two 8-bit Modulators
  - Transmission multiplies together
  - Over 10,000:1 dynamic range possible

**How it works**

- LED Backlight × LCD Screen = Combined Result

**Brightside HDR display**

- 37”
- 2000000:1
- Acquired by [Dolby](https://www.dolby.com)
Tone mapping operators

- Spatial (global/local)
- Frequency domain
- Gradient domain

- 3 papers from SIGGRAPH 2002
  - Photographic Tone Reproduction for Digital Images
  - Fast Bilateral Filtering for the Display of High-Dynamic-Range Images
  - Gradient Domain High Dynamic Range Compression

Photographic Tone Reproduction for Digital Images

Erik Reinhard  Mike Stark
Peter Shirley  Jim Ferwerda
SIGGRAPH 2002

Photographic tone reproduction

- Proposed by Reinhard et. al. in SIGGRAPH 2002
- Motivated by traditional practice, zone system by Ansel Adams and dodging and burning
- It contains both global and local operators

Zone system

- Darkest textured shadow
- Dynamic range = 15 scene zones
- Brightest textured highlight
- Middle grey maps to Zone V
- Print zones
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

Source: Ansel Adams

The Zones

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

Dodging and burning

- During the print
- Hide part of the print during exposure
  - Makes it brighter

From The Master Printing Course, Rudman
**Dodging and burning**

- Must be done for every single print!

**Global operator**

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

\[
L_m(x, y) = \frac{a}{L_w} L_w(x, y)
\]

\[
L_d(x, y) = \frac{L_m(x, y)}{1 + L_m(x, y)}
\]

**Transfer function to compress high luminances**

- \(L_w = \exp\left(\frac{1}{N} \sum_{x,y} \log(\delta + L_w(x, y))\right)\)

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

- \(L_{white}\) 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

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

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

\[ s_{\max} : |V_{s_{\max}}(x, y)| < \varepsilon \]

Dodging and burning (local operators)

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

- 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
Frequency domain
• First proposed by Oppenheim in 1968!
• Under simplified assumptions,

\[
\text{image} = \text{illuminance} \times \text{reflectance}
\]

- low-frequency attenuate more
- high-frequency attenuate less

Oppenheim
• Taking the logarithm to form density image
• Perform FFT on the density image
• Apply frequency-dependent attenuation filter

\[
s(f) = (1 - c) + c \frac{kf}{1 + kf}
\]
• Perform inverse FFT
• Take exponential to form the final image

Fast Bilateral Filtering for the Display of High-Dynamic-Range Images
Frédéric Durand & Julie Dorsey
SIGGRAPH 2002

A typical photo
• Sun is overexposed
• Foreground is underexposed
Gamma compression

- $X \rightarrow X^\gamma$
- Colors are washed-out

Gamma compression on intensity

- Colors are OK, but details (intensity high-frequency) are blurred

Chiu et al. 1993

- Reduce contrast of low-frequencies
- Keep high frequencies

The halo nightmare

- For strong edges
- Because they contain high frequency
Durand and Dorsey

- Do not blur across edges
- Non-linear filtering

Start with Gaussian filtering

- Here, input is a step function + noise

Edge-preserving filtering

- Blur, but not across edges

Start with Gaussian filtering

- Spatial Gaussian f

Input

Gaussian blur

Edge-preserving

Anisotropic diffusion [Perona & Malik 90]
- Blurring as heat flow
- LCIS [Tumblin & Turk]

Bilateral filtering [Tomasi & Manduci, 98]
Start with Gaussian filtering

- Output is blurred

The problem of edges

- Here, “pollutes” our estimate $J(x)$
- It is too different

Principle of Bilateral filtering

- [Tomasi and Manduchi 1998]
- Penalty $g$ on the intensity difference
Bilateral filtering

- [Tomasi and Manduchi 1998]
- Spatial Gaussian f

$\mathbf{I}(x) = \frac{1}{k(x)} \sum_{\mathbf{y} \in B(x)} w_{x,y} \mathbf{I}(y)$

$w_{x,y} = \min\left( f(|\mathbf{I}(x) - \mathbf{I}(y)|), g(|\mathbf{I}(x) - \mathbf{I}(y)|) \right)$

Normalization factor

- [Tomasi and Manduchi 1998]
- $k(x) = \sum_{\mathbf{y} \in B(x)} w_{x,y}$

Bilateral filtering is non-linear

- [Tomasi and Manduchi 1998]
- The weights are different for each output pixel
Contrast reduction

Input HDR image

Contrast too high!

Contrast reduction

Input HDR image

Intensity

Color

Contrast reduction

Input HDR image

Intensity

Fast Bilateral Filter

Large scale

Detail

Color

Contrast reduction

Input HDR image

Intensity

Fast Bilateral Filter

Large scale

Detail

Color
Bilateral filter is slow!

- Compared to Gaussian filtering, it is much slower because the kernel is not fixed.
- Durand and Dorsey proposed an approximate approach to speed up.
- Paris and Durand proposed an even-faster approach in ECCV 2006. We will cover this one when talking about computational photography.
Log domain

- Logarithm is a crude approximation to the perceived brightness
- Gradients in log domain correspond to ratios (local contrast) in the luminance domain

The method in 1D

1. Log
2. Derivative
3. Exponentiate
4. Integrate
The method in 2D

- Given: a log-luminance image \( H(x,y) \)
- Compute an attenuation map \( \Phi(\|\nabla H\|) \)
- Compute an attenuated gradient field \( G \): 
  \[ G(x, y) = \nabla H(x, y) \cdot \Phi(\|\nabla H\|) \]
- Problem: \( G \) may not be integrable!

Solution

- Look for image \( I \) with gradient closest to \( G \) in the least squares sense.
- \( I \) minimizes the integral: 
  \[ \iint F(\nabla I, G) \, dx \, dy \]
  \[ F(\nabla I, G) = \| \nabla I - G \|^2 = \left( \frac{\partial I}{\partial x} - G_x \right)^2 + \left( \frac{\partial I}{\partial y} - G_y \right)^2 \]

Solving Poisson equation

- No analytical solution
- Multigrid method
- Conjugate gradient method

\[
\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} = \frac{\partial G_x}{\partial x} + \frac{\partial G_y}{\partial y}
\]

\[
G_x(x, y) - G_x(x-1, y) + G_y(x, y) - G_y(x, y-1)
\]

\[
I(x+1, y) + I(x-1, y) + I(x, y+1) + I(x, y-1) - 4I(x, y)
\]
Attenuation

- Any dramatic change in luminance results in large luminance gradient at some scale
- Edges exist in multiple scales. Thus, we have to detect and attenuate them at multiple scales
- Construct a Gaussian pyramid $H_i$

Attenuation

$$\phi_k(x,y) = \left( \frac{\nabla H_k(x,y)}{\alpha} \right)^{\beta-1} \beta \sim 0.8 \quad \alpha = 0.1 \nabla H$$

log(Luminance)  gradient magnitude  attenuation map

Multiscale gradient attenuation

Final gradient attenuation map
Performance

- Measured on 1.8 GHz Pentium 4:
  - 512 x 384: 1.1 sec
  - 1024 x 768: 4.5 sec

- Can be accelerated using processor-optimized libraries.

Informal comparison

Gradient domain [Fattal et al.]  Bilateral [Durand et al.]  Photographic [Reinhard et al.]

Informal comparison

Gradient domain [Fattal et al.]  Bilateral [Durand et al.]  Photographic [Reinhard et al.]
Evaluation of Tone Mapping Operators using a High Dynamic Range Display

Patrick Ledda  Alan Chalmers  
Tom Troscinko  Helge Seetzen

SIGGRAPH 2005

Six operators

- H: histogram adjustment
- B: bilateral filter
- P: photographic reproduction
- I: iCAM
- L: logarithm mapping
- A: local eye adaption

23 scenes

Experiment setting
Preference matrix

- Ranking is easier than rating.
- 15 pairs for each person to compare. A total of 345 pairs per subject.

<table>
<thead>
<tr>
<th></th>
<th>tmo1</th>
<th>tmo2</th>
<th>tmo3</th>
<th>tmo4</th>
<th>tmo5</th>
<th>tmo6</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>tmo1</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>tmo2</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>tmo3</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>tmo4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>tmo5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>tmo6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>

Preference matrix (tmo2->tmo4, tom2 is better than tmo4)

Statistical measurements

- Statistical measurements are used to evaluate:
  - Agreement: whether most agree on the ranking between two tone mapping operators.
  - Consistency: no cycle in ranking. If all are confused in ranking some pairs, it means they are hard to compare. If someone is inconsistent alone, his ranking could be dropped.

Overall similarity

- Scene 8

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>H</th>
<th>B</th>
<th>L</th>
<th>I</th>
<th>A</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>-</td>
<td>24</td>
<td>46</td>
<td>42</td>
<td>10</td>
<td>32</td>
<td>154</td>
</tr>
<tr>
<td>H</td>
<td>24</td>
<td>-</td>
<td>44</td>
<td>32</td>
<td>8</td>
<td>12</td>
<td>120</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>4</td>
<td>-</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>L</td>
<td>6</td>
<td>16</td>
<td>40</td>
<td>-</td>
<td>4</td>
<td>12</td>
<td>78</td>
</tr>
<tr>
<td>I</td>
<td>38</td>
<td>40</td>
<td>46</td>
<td>44</td>
<td>-</td>
<td>38</td>
<td>206</td>
</tr>
<tr>
<td>A</td>
<td>16</td>
<td>36</td>
<td>44</td>
<td>36</td>
<td>10</td>
<td>-</td>
<td>142</td>
</tr>
</tbody>
</table>

Summary

Overall Similarity: Color

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>P</th>
<th>H</th>
<th>A</th>
<th>L</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>3712</td>
<td>3402</td>
<td>2994</td>
<td>2852</td>
<td>1902</td>
<td>1696</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bright Detail

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>A</th>
<th>P</th>
<th>H</th>
<th>B</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>823</td>
<td>688</td>
<td>569</td>
<td>549</td>
<td>474</td>
<td>347</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dark Detail

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>A</th>
<th>I</th>
<th>L</th>
<th>H</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>815</td>
<td>793</td>
<td>583</td>
<td>491</td>
<td>485</td>
<td>283</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Not settled yet!

- Some other experiment said bilateral are better than others.
- For your reference, photographic reproduction performs well in both reports.
- There are parameters to tune and the space could be huge.

References