Tone mapping

Digital Visual Effects
Yung-Yu Chuang

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?

The ultimate goal is a visual match

The eye is not a photometer!

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

Real world radiance

Display intensity

CRT has 300:1 dynamic range

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)
  - \( \frac{\Delta I_b}{I_b} \sim 1\% \)
  - background intensity
  - flash

Preliminaries

- For color images
  \[
  \begin{bmatrix}
  R_d \\
  G_d \\
  B_d
  \end{bmatrix}
  =
  \begin{bmatrix}
  R_w \\
  G_w \\
  B_w
  \end{bmatrix}
  \]
- Log domain is usually preferred.
**HDR Display**

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

  - Theoretically, 240,000:1.
  - Due to imperfect optical depth, 54,000:1 measured.

**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”
- 200,000:1
- World’s First Extreme Dynamic Range Display
- Acquired by Dolby
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

Global v.s. local

Example: Gamma Compression

Example: Adaptive Histogram Equalization

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
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

0 Solid black; the same as the film rebate

I Nearly black; just different from Zone 0

II The first hint of texture

III Textured shadow; the first recognizable shadow detail

IV Average shadow value on Caucasian skin, foliage and buildings

V Middle grey; the pivot value; light foliage, dark skin

VI Caucasian skin, textured light grey; shadow on snow

VII Light skin; bright areas with texture, such as snow in low sunlight

VIII Highest zone with any texture

IX Pure untextured white

Source: Ansel Adams
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!

From Photography by London et al.

Global operator

$$L_u = \exp\left(\frac{1}{N} \sum_{x, y} \log(\delta + L_m(x, y))\right)$$

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

$$L_m(x, y) = \frac{d}{L_w} L_w(x, y)$$

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

Transfer function to compress high luminances
**Global operator**

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

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

![Graph showing L_d and L_white](image)

**Dodging and burning (local operators)**

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

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

\[
V_s(x, y) = \frac{L_s^{blur}(x, y) - L_{s+1}^{blur}(x, y)}{2^d \alpha \beta^2 + L_s^{blur}}
\]

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

**Dodging and burning (local operators)**

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

- 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
**Dodging and burning**

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

\[
\text{image} = \text{illuminance low-frequency attenuate more} \ast \text{reflectance high-frequency attenuate less}
\]

**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**

Frédo Durand & Julie Dorsey

SIGGRAPH 2002
A typical photo
- Sun is overexposed
- Foreground is underexposed

Gamma compression
- $X \rightarrow X'$
- 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

Low-freq.  Reduce low frequency

High-freq.

Color

Durand and Dorsey
- Do not blur across edges
- Non-linear filtering

Large-scale  Output

Detail

Color

Edge-preserving filtering
- Blur, but not across edges

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
- Here, input is a step function + noise

output  input
Start with Gaussian filtering

- Spatial Gaussian $f$

Output is blurred

Gaussian filter as weighted average

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

$$I(x) = \sum_{\xi} \xi f(I)$$

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

\[ \sum_{x} (f(x) \cdot g(I(x) - I(f(x)))) \]

Bilateral filtering

- [Tomasi and Manduchi 1998]
- Spatial Gaussian $f$

Normalization factor

- [Tomasi and Manduchi 1998]
- $k(x) =$

\[ \sum_{x} f(x) \cdot g(I(x) - I(f(x))) \]
Bilateral filtering is non-linear

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

\[ \text{output} \rightarrow \text{input} \]

Contrast reduction

Contrast too high!

Contrast reduction

Intensity

Color

Large scale

Fast Bilateral Filter
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.

Gradient Domain High Dynamic Range Compression

Raanan Fattal   Dani Lischinski   Michael Werman

SIGGRAPH 2002
The method in 1D

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: $\int\int F(\nabla I, G)\,dxdy$

$$F(\nabla I, G) = \|\nabla I - G\|^2 = \left(\frac{\partial I}{\partial x} - G_x\right)^2 + \left(\frac{\partial I}{\partial y} - G_y\right)^2$$

- Solve

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

$$G_x(x, y) - G_x(x-1, y) + G_y(x, y) - G_y(x, y-1)$$

$$I(x+1, y) + I(x-1, y) + I(x, y+1) + I(x, y-1) - 4I(x, y)$$

$$\begin{bmatrix} \ldots & 1 & \ldots & -4 & 1 & \ldots \end{bmatrix} \begin{bmatrix} I \end{bmatrix} = \begin{bmatrix} \end{bmatrix}$$
**Solving Poisson equation**

- No analytical solution
- Multigrid method
- Conjugate gradient method

**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 $H_i$

\[
\varphi_k(x, y) = \left( \frac{\|
abla H_k(x, y)\|}{\alpha} \right)^{\beta-1}
\]

\[
\beta \sim 0.8 \\
\alpha = 0.1 \nabla H
\]

**Multiscale gradient attenuation**

- Log(Luminance)
- Gradient magnitude
- Attenuation map
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

<table>
<thead>
<tr>
<th>Gradient domain</th>
<th>Bilateral</th>
<th>Photographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Fattal et al.]</td>
<td>[Durand et al.]</td>
<td>[Reinhard et al.]</td>
</tr>
</tbody>
</table>

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

- HDR display
- Tonemapping result

### Preference matrix

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

#### Preference matrix

<table>
<thead>
<tr>
<th></th>
<th>$t_{mo_1}$</th>
<th>$t_{mo_2}$</th>
<th>$t_{mo_3}$</th>
<th>$t_{mo_4}$</th>
<th>$t_{mo_5}$</th>
<th>$t_{mo_6}$</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{mo_1}$</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>$t_{mo_2}$</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$t_{mo_3}$</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>$t_{mo_4}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$t_{mo_5}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$t_{mo_6}$</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 ($t_{mo_2}$ vs $t_{mo_4}$, $t_{mo_2}$ is better than $t_{mo_4}$)

### 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

#### 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>Sum</td>
<td>3712</td>
<td>3402</td>
<td>2994</td>
<td>2852</td>
<td>1902</td>
<td>1696</td>
</tr>
</tbody>
</table>

**Bright Detail**

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>A</th>
<th>P</th>
<th>II</th>
<th>B</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum</td>
<td>823</td>
<td>688</td>
<td>569</td>
<td>549</td>
<td>474</td>
<td>347</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>Sum</td>
<td>815</td>
<td>793</td>
<td>583</td>
<td>491</td>
<td>485</td>
<td>283</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