Realistic Image Synthesis

- HDR Capture & Tone Mapping -

Philipp Slusallek
Karol Myszkowski
Gurprit Singh
## LDR vs HDR – Comparison

<table>
<thead>
<tr>
<th>Standard Dynamic Range</th>
<th>High Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QUALITY OF CONTRAST &amp; COLOR</strong></td>
<td></td>
</tr>
<tr>
<td>50 dB</td>
<td>Camera Dynamic Range</td>
</tr>
<tr>
<td>1:200</td>
<td>Display Contrast</td>
</tr>
<tr>
<td>limited</td>
<td>Color Gamut</td>
</tr>
<tr>
<td>display-referred</td>
<td>Image Representation</td>
</tr>
<tr>
<td>display limited</td>
<td>Fidelity</td>
</tr>
</tbody>
</table>
Various Dynamic Ranges (1)

Luminance [cd/m²]

10⁻⁴  10⁻²  10⁰  10²  10⁴  10⁶  10⁸

Realistic Image Synthesis SS19 – HDR Image Capture & Tone Mapping
Various Dynamic Ranges (2)

Luminance [cd/m²]

Contrast
1:1000
1:1500
1:30
High Dynamic Range

- HDR Image
- Usual (LDR) Image

Realistic Image Synthesis SS19 – HDR Image Capture & Tone Mapping
# Measures of Dynamic Range

<table>
<thead>
<tr>
<th>Measure</th>
<th>Formula</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast ratio</td>
<td>( CR = 1 : (Y_{peak}/Y_{noise}) )</td>
<td>displays (1:500)</td>
</tr>
<tr>
<td>Orders of magnitude</td>
<td>( M = \log_{10}(Y_{peak})-\log_{10}(Y_{noise}) )</td>
<td>HDR imaging (2.7 orders)</td>
</tr>
<tr>
<td>Exposure latitude (f-stops)</td>
<td>( L = \log_{2}(Y_{peak})-\log_{2}(Y_{noise}) )</td>
<td>photography (9 f-stops)</td>
</tr>
<tr>
<td>Signal to noise ratio (SNR)</td>
<td>( SNR = 20\log_{10}(A_{peak}/A_{noise}) )</td>
<td>digital cameras (53 [dB])</td>
</tr>
</tbody>
</table>
HDR Pipeline
Lecture Overview

- Capture of HDR images and video
  - Multi-exposure techniques
  - Photometric calibration

- Tone Mapping of HDR images and video
  - Early ideas for reducing contrast range
  - Image processing – fixing problems
  - Alternative approaches
  - Perceptual effects in tone mapping

- Summary
HDR: a normal camera can’t…

- linearity of the CCD sensor
- bound to 8-14bit processors
- saved in an 8bit gamma corrected image
HDR Sensors

- logarithmic response
- locally auto-adaptive
- hybrid sensors (linear-logarithmic)
HDR with a normal camera

Dynamic range of a typical CCD $1:1000$

Exposure variation ($1/60 : 1/6000$) $1:100$

Aperture variation ($f/2.0 : f/22.0$) ~$1:100$

Sensitivity variation (ISO 50 : 800) ~$1:10$

Total operational range $1:100,000,000$ High Dynamic Range!

Dynamic range of a single capture only $1:1000$. 
Multi-exposure Technique (1)

Realistic Image Synthesis SS19 – HDR Image Capture & Tone Mapping
Multi-exposure Technique (2)

- **Input**
  - images captured with varying exposure
    - change exposure time, sensitivity (ISO), ND filters
    - same aperture!
    - exactly the same scene!

- **Unknowns**
  - camera response curve (can be given as input)
  - HDR image

- **Process**
  - recovery of camera response curve (if not given as input)
  - linearization of input images (to account for camera response)
  - normalization by exposure level
  - suppression of noise
  - estimation of HDR image (linear combination of input images)
Algorithm (1/3)

**Camera Response**

\[ y_{ij} = I(x_{ij} \cdot t_i) \]

**Merge to HDR**

- Linearize input images and normalize by exposure time
  \[ x_{ij} = \frac{I^{-1}(y_{ij})}{t_i} \]
  assume \( I \) is correct (initial guess)

- Weighted average of images (weights from certainty model)
  \[ x_j = \frac{\sum_i w_{ij} x_{ij}}{\sum_i w_{ij}} \]

**Optimize Camera Response**

- Camera response
  \[ I^{-1}(y_{ij}) = t_i x_j \]
  assume \( x_j \) is correct

- Refine initial guess on response
  - linear eq. (Gauss-Seidel method)

\[ E_m = \{(i, j) : y_{ij} = m\} \]
\[ I^{-1}(m) = \frac{1}{\text{Card}(E_m)} \sum_{i, j \in E_m} t_i x_j \]

---

\( t_i \) exposure time of image \( i \)
\( y_{ij} \) pixel of input image \( i \) at position \( j \)
\( I \) camera response
\( x_j \) HDR image at position \( j \)
\( w \) weight from certainty model
\( m \) camera output value
Algorithm (2/3)

- Certainty model (for 8bit image)
  - High confidence in middle output range
  - Dequantization uncertainty term
  - Noise level

\[ w(y_{ij}) = \exp \left( -4 \frac{(y_{ij} - 127.5)^2}{127.5^2} \right) \]

- Longer exposures are favored \( t_i^2 \)
  - Less random noise

- Weights

\[ w_{ij} = w(y_{ij}) t_i^2 \]
Algorithm (3/3)

1. Assume initial camera response $I$ (linear)

2. Merge input images to HDR

$$x_j = \frac{\sum_i w(y_{ij})t_i^2 I^{-1}(y_{ij})}{\sum_i w(y_{ij})t_i^2}$$

3. Refine camera response

$$E_m = \{(i, j) : y_{ij} = m\}$$

$$I^{-1}(m) = \frac{1}{\text{Card}(E_m)} \sum_{i, j \in E_m} t_i x_j$$

4. Normalize camera response by middle value: $I^{-1}(m)/I^{-1}(m_{med})$

5. Repeat 2,3,4 until objective function is acceptable

$$O = \sum_{i, j} w(y_{ij})(I^{-1}(y_{ij}) - t_i x_j)^2$$
Other Algorithms

- [Debevec & Malik 1997]
  - in log space
  - assumptions on the camera response
    - monotonic
    - continuous
  - a lot to compute for >8bit

- [Mitsunaga & Nayar 1999]
  - camera response approximated with a polynomial
  - very fast

- Both are more robust but less general
  - not possible to calibrate non-standard sensors
Issues with Multi-exposures

- **How many source images?**
  - First expose for shadows: all output values above 128 (for 8bit imager)
  - 2 f-stops spacing (factor of 4) between images
  - one or two images with 1/3 f-stop increase will improve quantization in HDR image
  - Last exposure: no pixel in image with maximum value

- **Alignment**
  - Shoot from tripod
  - Otherwise use panorama stitching techniques to align images

- **Ghosting**
  - Moving objects between exposures leave “ghosts”
  - Statistical method to prevent such artifacts

- **Practical only for images!**
  - Multi-exposure video projects exist, but require care with subsequent frame registration by means of optical flow
Photometric Calibration

- Converts camera output to luminance
  - requires camera response,
  - and a reference measurement for known exposure settings

- Applications
  - predictive rendering
  - simulation of human vision response to light
  - common output in systems combining different cameras
Calibration (Response Recovery)

- Camera response can be reused
  - for the same camera
  - for the same picture style settings (e.g. contrast)

- Good calibration target
  - Neutral target (e.g. Gray Card)
    - Minimize impact of color processing in camera
  - Smooth illumination
    - Uniform histogram of input values
  - Out-of-focus
    - No interference with edge aliasing and sharpening
Photometric Calibration (cntd.)

acquire target

measure luminance

luminance values

camera output values

camera response
HDR Sensor vs. Multi-exposure

- **HDR camera**
  - Fast acquisition of dynamic scenes at 25fps without motion artifacts
  - Currently lower resolution

- **LDR camera + multi-exposure technique**
  - Slow acquisition (impossible in some conditions)
  - Higher quality and resolution
  - High accuracy of measurements
Lecture Overview

- Capture of HDR images and video
  - HDR sensors
  - Multi-exposure techniques
  - Photometric calibration

- Tone Mapping of HDR images and video
  - Early ideas for reducing contrast range
  - Image processing – fixing problems
  - Alternative approaches
  - Perceptual effects in tone mapping

- Summary
Objectives of tone mapping
- nice looking images
- perceptual brightness match
- good detail visibility
- equivalent object detection performance
- really application dependent…
General Idea

- Luminance as an input
  - absolute luminance
  - relative luminance (luminance factor)

- Transfer function
  - maps luminance to a certain pixel intensity
  - may be the same for all pixels (global operators)
  - may depend on spatially local neighbors (local operators)
  - dynamic range is reduced to a specified range

- Pixel intensity as output
  - often requires gamma correction

- Colors
  - most algorithms work on luminance
    - use RGB to Yxy color space transform
    - inverse transform using tone mapped luminance
  - otherwise each RGB channel processed independently
General Problems

- Constraints in observation conditions
  - limited contrast
  - quantization
  - different ambient illumination
  - different luminance levels
  - adaptation level often incorrect for the scene
  - narrow field of view

- Appearance may not always be matched
Transfer Functions

- Linear mapping (naïve approach)
  - like taking a usual photo
- Brightness function
- Sigmoid responses
  - simulate our photoreceptors
  - simulate response of photographic film
- Histogram equalization
  - standard image processing
  - requires detection threshold limit to prevent contouring
Adapting Luminance

- Maps luminance on a scale of gray shades
- Task is to match gray levels
  - average luminance in the scene is perceived as a gray shade of medium brightness
  - such luminance is mapped on medium brightness of a display
  - the rest is mapped proportionally
- Practically adjusts brightness
  - sort of like using gray card or auto-exposure in photography
  - goal of adaptation processes in human vision
- Adapting luminance exists in many TM algorithms

$$Y_A = \exp\left(\frac{\sum \log(Y + \varepsilon)}{N} - \varepsilon\right)$$
Logarithmic Tone Mapping

- Logarithm is a crude approximation of brightness
- Change of base for varied contrast mapping in bright and dark areas
  - $\log_{10}$ maps better for bright areas
  - $\log_2$ maps better for dark areas
- Mapping parameter $bias$ in range 0.1:1

\[
Y' = \frac{Y}{Y_A}
\]

\[
L = L_{\text{max}} \cdot \frac{\log_{\text{base}(Y)}(Y'+1)}{\log_{10}(\max(Y') + 1)}
\]

\[
\text{base}(Y') = 2 + 8 \cdot \left( \frac{Y'}{\max(Y')} \right)^{\log_{0.5} bias}
\]
Logarithmic Tone Mapping

These images illustrate how high luminance values are clamped to the maximum displayable values using different bias parameter values.

The scene dynamic range is 1:11,751,307.

$$\left( \frac{Y'}{\max(Y')} \right)^{\log_{0.5} \text{bias}}$$
Sigmoid Response

- Model of photoreceptor

\[ L = \frac{Y}{Y + (f \cdot Y_A)^m} L_{\text{max}} \]

- Brightness parameter \( f \)

- Contrast parameter \( m \)

- Adapting luminance \( Y_A \)
  - average in an image
  - measured pixel (equal to \( Y \))

logarithmic mapping

sigmoid mapping
Histogram Equalization (1)

- Adapts transfer function to distribution of luminance in the image
- Algorithm:
  - compute histogram
  - compute transfer function (cumulative distribution)
  - limit slope of transfer function to prevent contouring
    - contouring – visible difference between 1 quantization step
    - use threshold versus intensity function (TVI)
      TVI gives visible luminance difference for adapting luminance
- Most optimal transfer function
- Not efficient when large uniform areas are present in the image
Histogram Equalization (2)

A linear mapping of the luminances that overexposes the view through the window.  

[Graph showing world to display luminance mapping]

A linear mapping of the luminances that underexposes the view of the interior.

The luminances mapped to preserve the visibility of both indoor and outdoor features.
Transfer Functions Compared

- **Interpretation**
  - steepness of slope is contrast
  - luminance for which output is ~0 and ~1 is not transferred
- **Usually low contrast for dark and bright areas!**
Problem with Details

- Strong compression of contrast puts micro-contrasts (details) below quantization level
Introducing Local Adaptation

- Eye adapts locally to observed area

\[ L = \frac{Y'}{Y' + 1} \quad Y' = \frac{Y}{Y_A} \quad L = \frac{Y'}{Y_L' + 1} \]

Gaussian blur of HDR image, \( \sigma \sim 1 \text{deg of visual angle.} \)
The Halo Artifact

- Scan line example:
  - Gaussian blur under- (over-) estimates local adaptation near a high contrast edge
  - tone mapped image gets too bright (too dark) closer to such an edge
- Smaller blur kernel reduces the artifact (but then no details)
- Larger blur kernel spreads the artifact on larger area
Adjusting Gaussian Blur

- So called: Automatic Dodging and Burning
  - for each pixel, test increasing blur size \( \sigma_i \)
  - choose the largest blur which does not show halo artifact

\[
|Y_L(x, y, \sigma_i) - Y_L(x, y, \sigma_{i+1})| < \varepsilon
\]
Photographic Tone Reproduction

- Map luminance using Zone System
  
  \[ \begin{array}{cccccccc}
  2^x L & 2^{x+1} L & 2^{x+2} L & 2^{x+3} L & 2^{x+4} L & \ldots & 2^{x+15} L & 2^{x+16} L \\
  \end{array} \]
  
  Middle grey maps to Zone V

  Print zones: Zone V 18% reflectance

- Find local adaptation for each pixel
  - appropriate size of Gaussian (automatic dodging & burning)

  \[ |Y_L'(x, y, \sigma_i) - Y_L'(x, y, \sigma_{i+1})| < \varepsilon \]

- Tone map using sigmoid function
  - different blur levels from Gaussian pyramid

  \[ L(x, y) = \frac{Y'(x, y)}{Y_L'(x, y, \sigma_{x, y}) + 1} \]

\[ Y' = \frac{Y}{Y_A}, \quad Y_A = \exp\left(\frac{\sum \log(Y)}{N}\right) \]
Photographic Tone Reproduction

**dodge**  luminance of pixels in bright regions is significantly decreased

**burn**  pixels in dark regions are compressed less, so their relative intensity increases

Automatic dodging-and-burning technique is more effective in preserving local details (notice the print in the book).
Bilateral Filtering

- Edge preserving Gaussian filter to prevent halo
- Conceptually based on intrinsic image models:
  - decoupling of illumination and reflectance layers
    - very simple task in CG
    - complicated for real-world scenes
  - compress range of illumination layer
  - preserve reflectance layer (details)
- Bilateral filter separates:
  - texture details (high frequencies, low amplitudes)
  - illumination (low frequencies, high contrast edges)
Illumination Layer (1)

- Identify low frequencies in the scene
  - Gaussian filtering leads to halo artifacts

\[ J_p = \frac{1}{W_p} \sum_{q \in N(p)} f_{\sigma_s} \left( \| p - q \| \right) \cdot I_q \]

- Spatial kernel with large \( \sigma_s \)

Lost sharp edge
Illumination Layer (2)

- Edge preserving filter – no halo artifacts

\[
J_p = \frac{1}{W_p} \sum_{q \in N(p)} f_{\sigma_s}(\|p - q\|) \cdot g_{\sigma_r}(\|I_p - I_q\|) \cdot I_q
\]

- \( f \) spatial kernel with large \( \sigma_s \)
- \( g \) range kernel with very small \( \sigma_r \)
Tone Mapping Algorithm

Luminance in logarithmic domain.
Illumination & Reflectance
Gradient Compression Algorithm

1. Calculate gradients map of image
2. Calculate attenuation map
3. Attenuate gradients
4. Solve Poisson equation to recover image

\[ L_{\text{d}} = \exp I \]

\[ H = \log L \]

\[ H(x, y) \rightarrow \nabla H(x, y) \rightarrow G(x, y) = \nabla H(x, y) * \Phi(x, y) \rightarrow \nabla^2 I = \text{div} G \]
Attenuation Map

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

1. Create Gaussian pyramid
2. Calculate gradients on levels
3. Calculate attenuation on levels - \( \varphi_k \)
4. Propagate levels to full resolution

Realistic Image Synthesis SS19 – HDR Image Capture & Tone Mapping
Transfer Function for Contrasts

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

- Attenuate large gradients
  - presumably illumination
- Amplify small gradients
  - hopefully texture details
  - but also noise
- Equation has a division by zero!
Perceptual Effects in TM

- Simulate effects that do not appear on a screen but are typically observed in real-world scenes
  - veiling glare
  - night vision
  - temporal adaptation to light
- Increase believability of results, because we associate such effects with luminance conditions
Temporal Luminance Adaptation

- Compensates changes in illumination
- Simulated by smoothing adapting luminance in tone mapping equation
- Different speed of adaptation to light and to darkness
Night Vision

- Human Vision operates in three distinct adaptation conditions:

  - **Scotopic**
    - Monochromatic vision
    - Limited visual acuity
    - Luminance: -6, -4, -2
    - Conditions: starlight, moonlight

  - **Mesopic**
    - Good color perception
    - Good visual acuity
    - Luminance: 0, 2, 4
    - Conditions: office illumination

  - **Photopic**
    - Good color perception
    - Good visual acuity
    - Luminance: 6, 8
    - Conditions: sunlight
Visual Acuity

- Perception of spatial details is limited with decreasing illumination level
- Details can be removed using convolution with a Gaussian kernel
- Highest resolvable spatial frequency:

\[ RF(Y) = 17.25 \cdot \arctan(1.4 \log_{10} Y + 0.35) + 25.72 \]
Veiling Luminance (Glare)

- Decrease of contrast and visibility due to light scattering in the optical system of the eye

- Described by the optical transfer function:

$$OTF(\rho, d(\bar{Y})) = \exp \left( -\frac{\rho}{20.9 - 2.1 \cdot d} \cdot \left( 1.3 - 0.07 \cdot d \right) \right)$$

\( \rho \) spatial frequency, \( d \) pupil aperture
HDR Video Player with Perceptual Effects
Papers on Multi-exposure Techniques

- Estimation-Theoretic Approach to Dynamic Range Improvement Using Multiple Exposures
  - M. Robertson, S. Borman, and R. Stevenson

- Recovering High Dynamic Range Radiance Maps from Photographs
  - Paul E. Debevec and Jitendra Malik
  - In: SIGGRAPH 97

- Radiometric Self Calibration
  - T. Mitsunaga and S.K. Nayar

- High Dynamic Range from Multiple Images: Which Exposures to Combine?
  - M.D. Grossberg and S.K. Nayar
Papers about Tone Mapping

- Adaptive Logarithmic Mapping for Displaying High Contrast Scenes
  - F. Drago, K. Myszkowski, T. Annen, and N. Chiba
  - In: Eurographics 2003

- Photographic Tone Reproduction for Digital Images
  - E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)

- Fast Bilateral Filtering for the Display of High-Dynamic-Range Images
  - F. Durand and J. Dorsey
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)

- Gradient Domain High Dynamic Range Compression
  - R. Fattal, D. Lischinski, and M. Werman
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)

- Dynamic Range Reduction Inspired by Photoreceptor Physiology
  - E. Reinhard and K. Devlin
  - In IEEE Transactions on Visualization and Computer Graphics, 2005

- Time-Dependent Visual Adaptation for Realistic Image Display
  - S.N. Pattanaik, J. Tumblin, H. Yee, and D.P. Greenberg
  - In: Proceedings of ACM SIGGRAPH 2000

- Lightness Perception in Tone Reproduction for High Dynamic Range Images
  - G. Krawczyk, K. Myszkowski, H.-P. Seidel
  - In: Eurographics 2005

- Perceptual Effects in Real-time Tone Mapping
  - G. Krawczyk, K. Myszkowski, H.-P. Seidel
  - In: Spring Conference on Computer Graphics, 2005
Acknowledgements

- I would like to thank Grzesiek Krawczyk for making his slides available.