High dynamic range imaging and tonemapping

15-463, 15-663, 15-862
Computational Photography
Fall 2017, Lecture 12

http://graphics.cs.cmu.edu/courses/15-463

Course announcements

- Homework 3 is out.
  - Due October 12th.
  - Shorter, but longer bonus component.

- Homework 4 will involve making some substantial use of a camera
  - Sign up to borrow one of the DSLRs we have for class.
  - You can work in teams of two (but each needs to submit their own homework).

- Wednesday we have our first guest lecture
  - Ravi Mullapudi will tell us about high-performance image processing

Overview of today's lecture

- Our devices do not match the world.
- High dynamic range imaging.
- Tonemapping.
- Some notes about HDR and tonemapping.

Slide credits

Many of these slides were inspired or adapted from:

- James Hays (Georgia Tech).
- Fredo Durand (MIT).
- Gordon Wetzstein (Stanford).


HYATT


HYATT


HYATT


THE INTERNATIONAL FUR EXCHANGE


HYATT


HYATT

Our devices do not match the world

The world has a high dynamic range

1

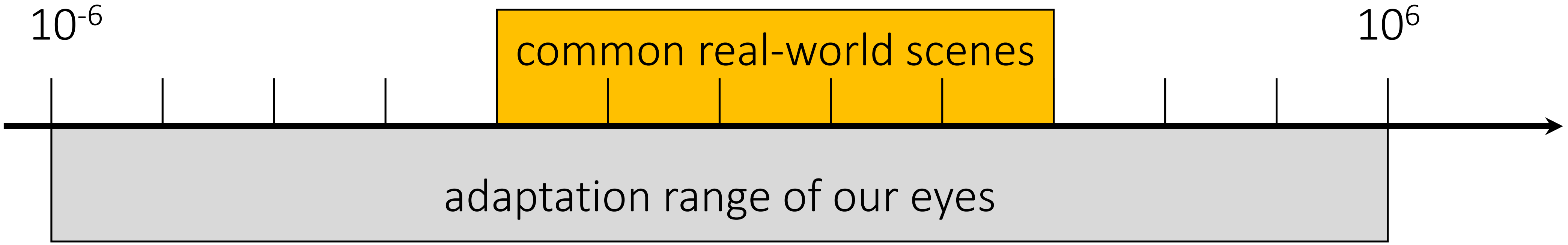
1500

25,000

400,000

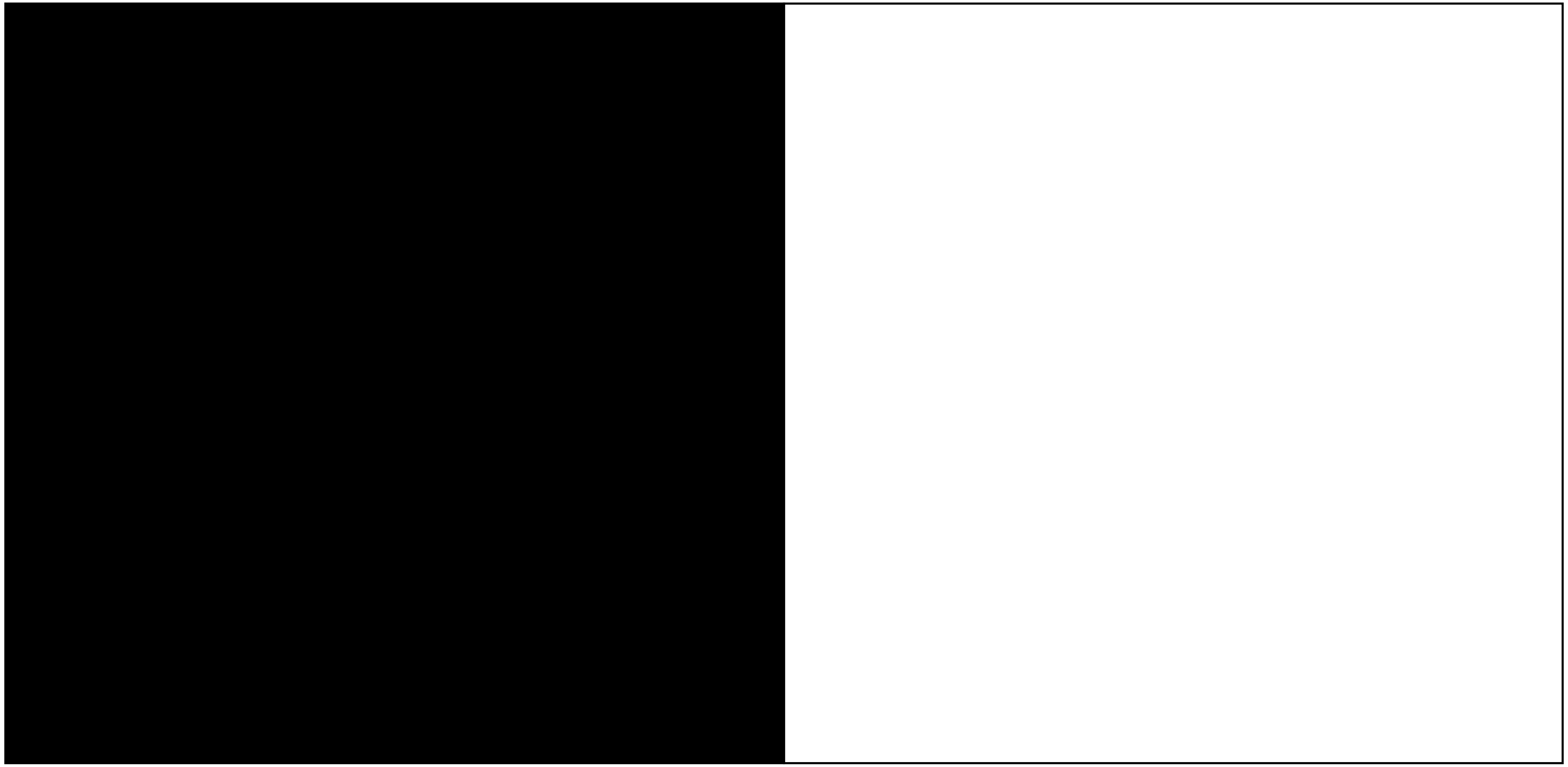
2,000,000,000

The world has a high dynamic range



(Digital) images have a low dynamic range

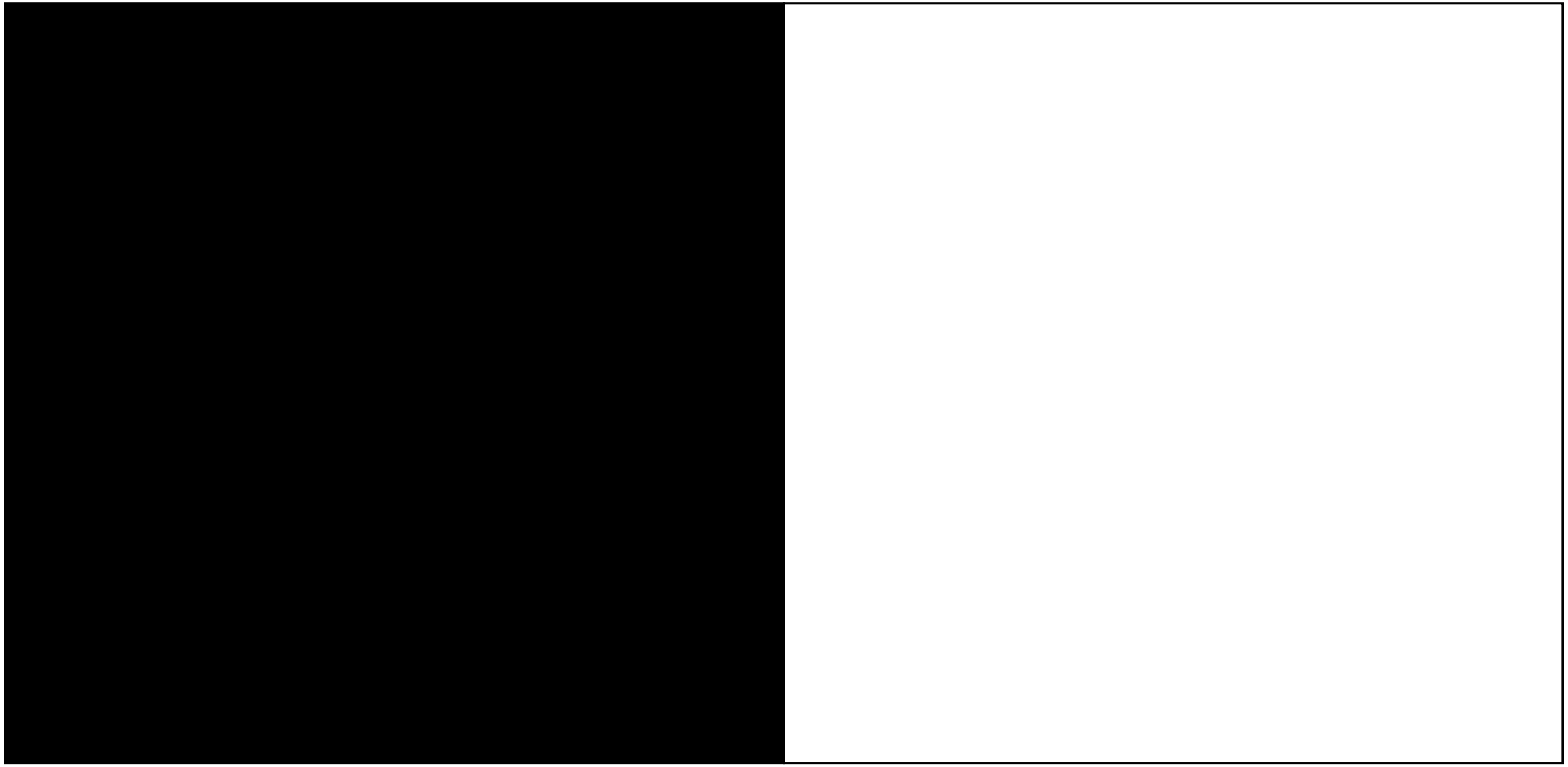
Any guesses about the dynamic range of a standard 0-255 image?



pure black pure white

(Digital) images have a low dynamic range

Any guesses about the dynamic range of a standard 0-255 image?

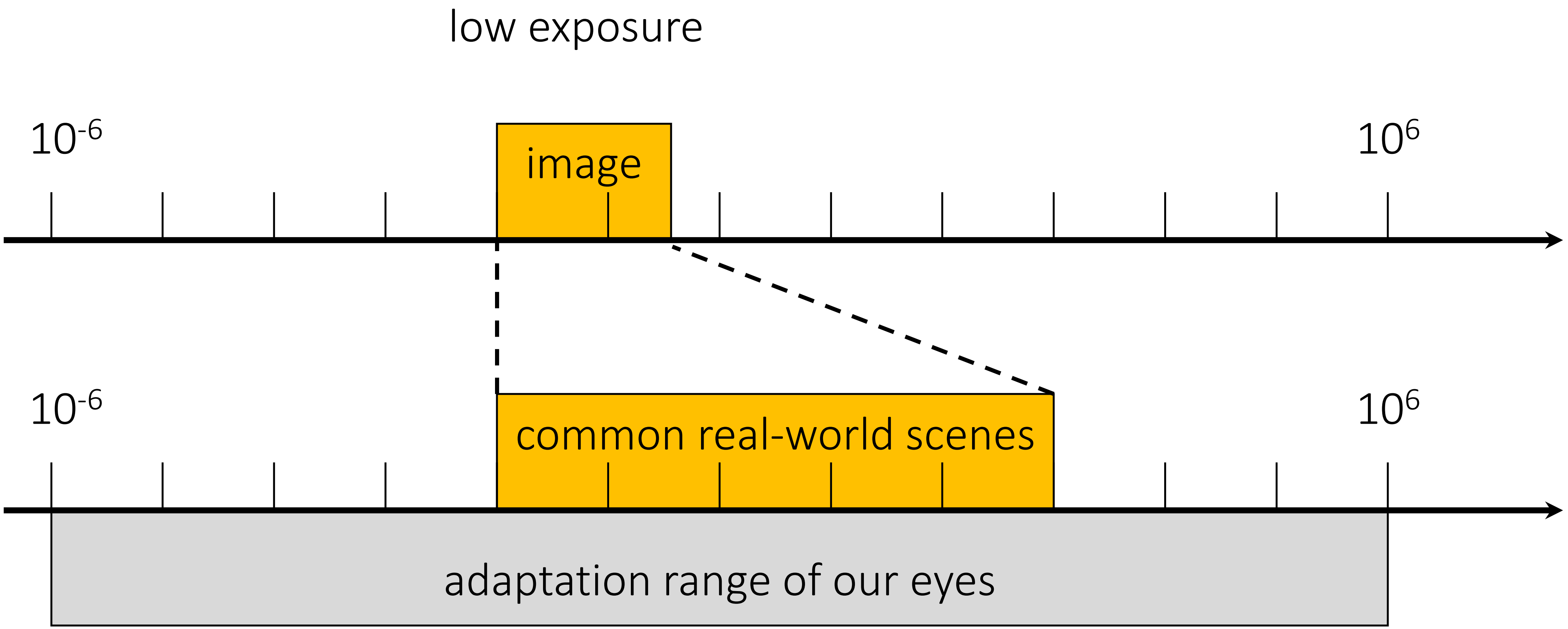


about 50x brighter

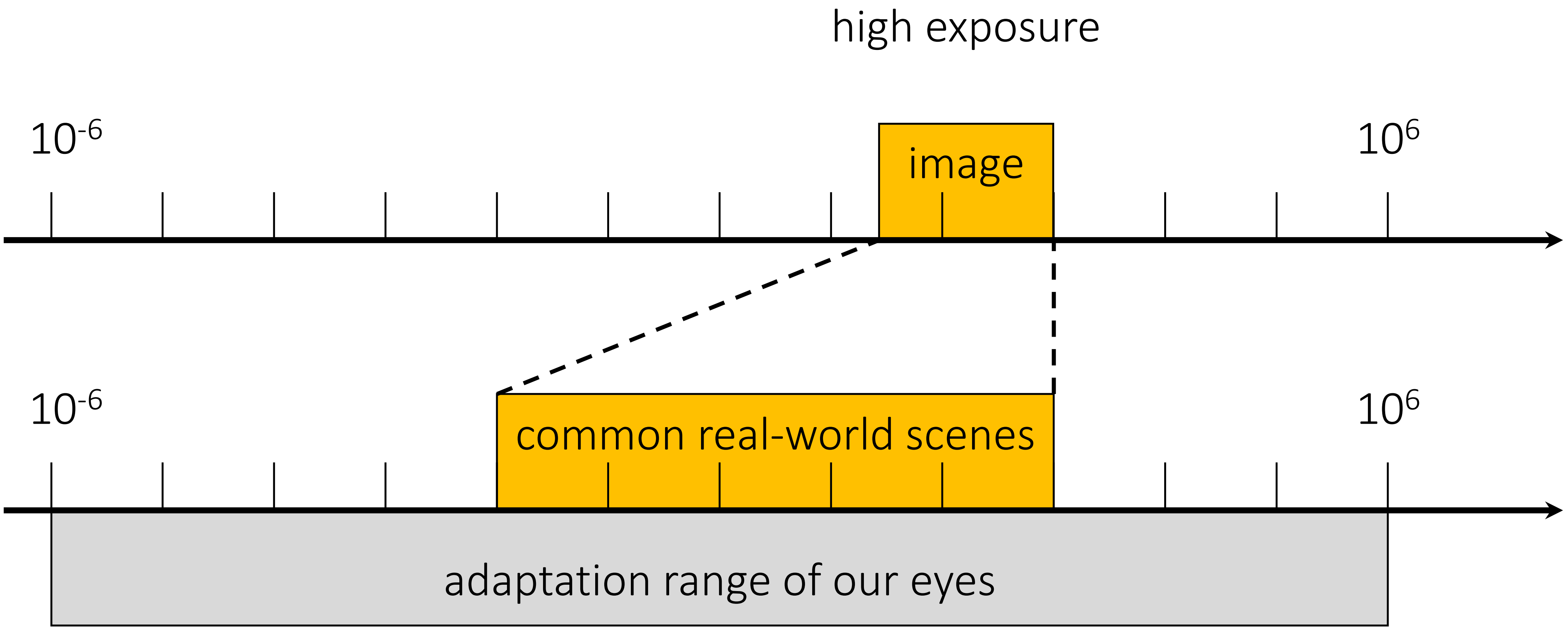
pure black

pure white

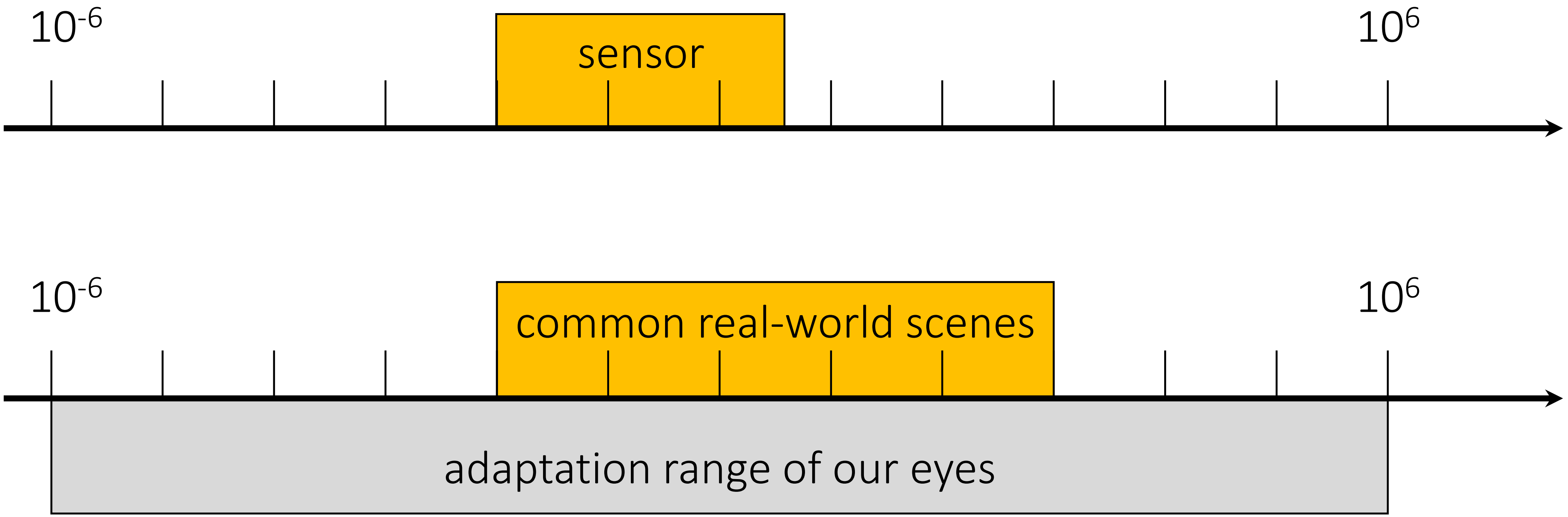
(Digital) images have a low dynamic range



(Digital) images have a low dynamic range



(Digital) sensors also have a low dynamic range



Our devices do not match the real world

• 10:1	photographic print (higher for glossy paper)
• 20:1	artist's paints
• 200:1	slide film
• 500:1	negative film
• 1000:1	LCD display
• 2000:1	digital SLR (at 12 bits)
• 100000:1	real world

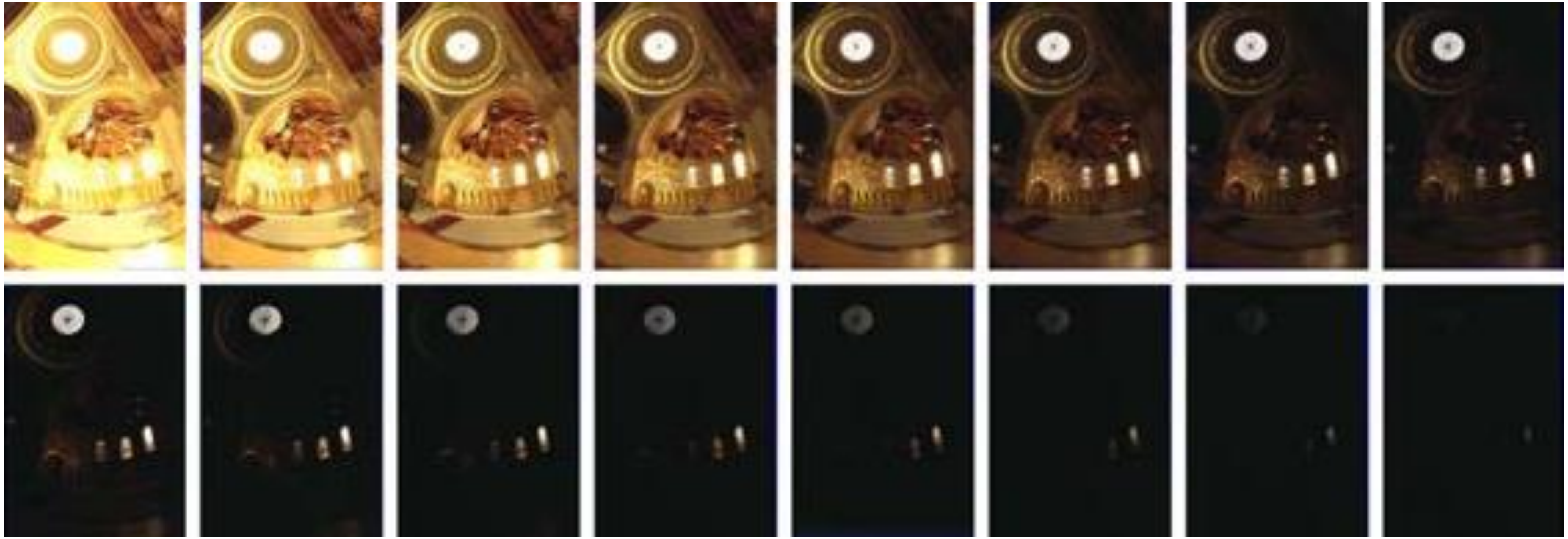
Two challenges:

1. HDR imaging – which parts of the world to include to the 8-12 bits available to our device?
2. Tonemapping – which parts of the world to display in the 4-10 bits available to our device?

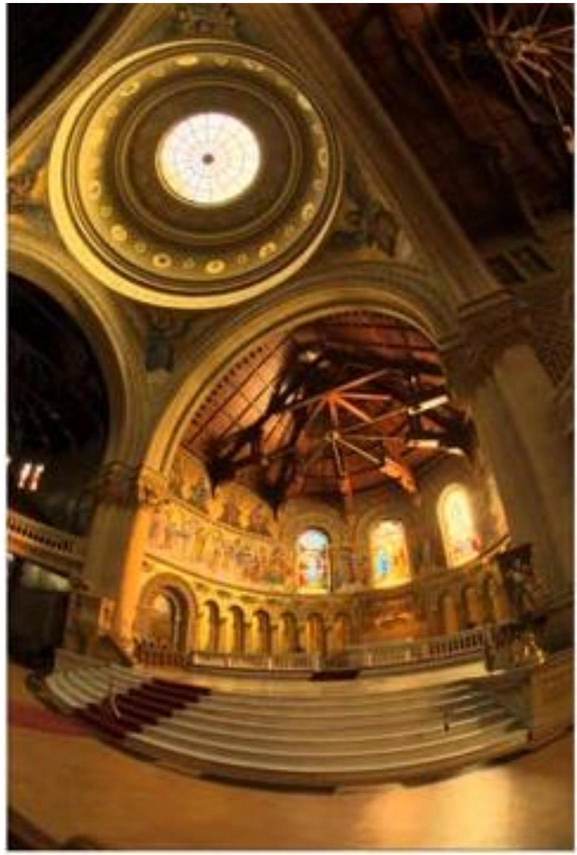
High dynamic range imaging

Key idea

1. Capture multiple LDR images at different exposures

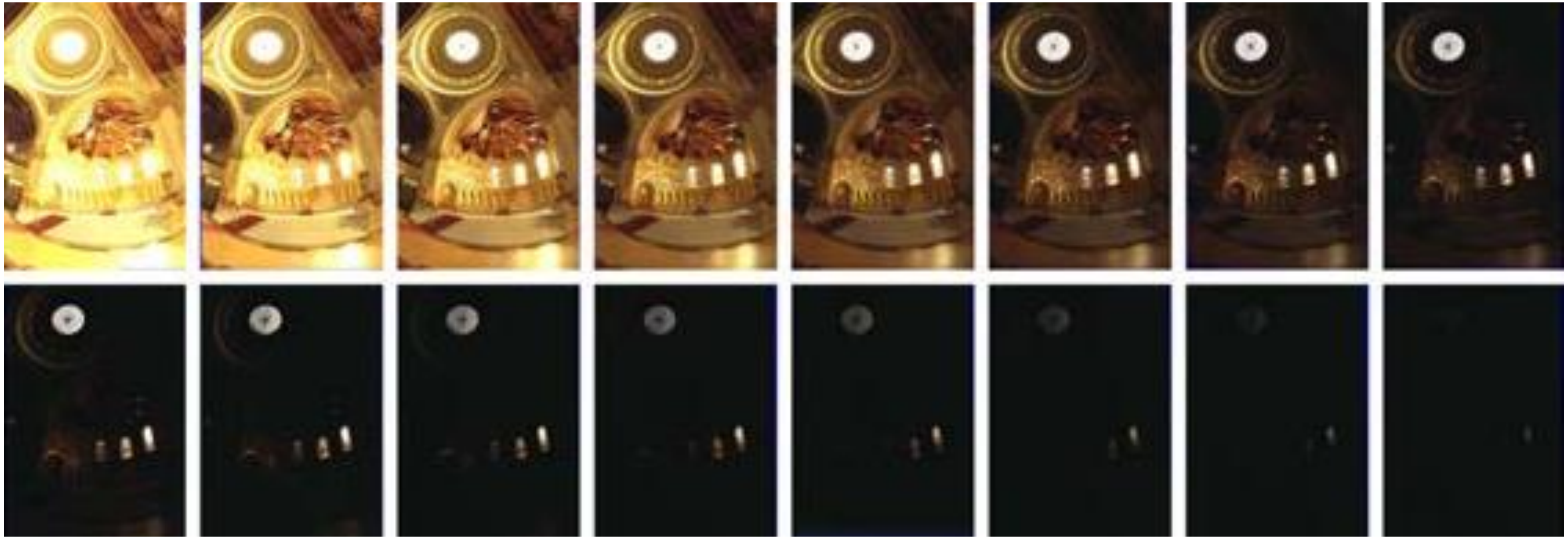


2. Merge them into a single HDR image

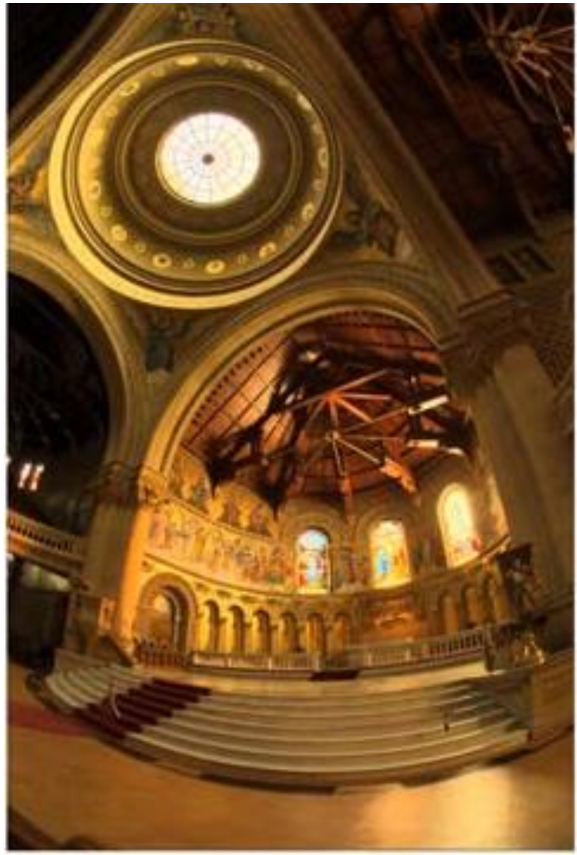


Key idea

1. Capture multiple LDR images at different exposures

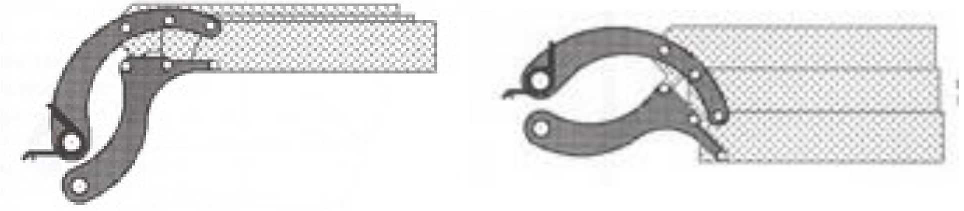


2. Merge them into a single HDR image

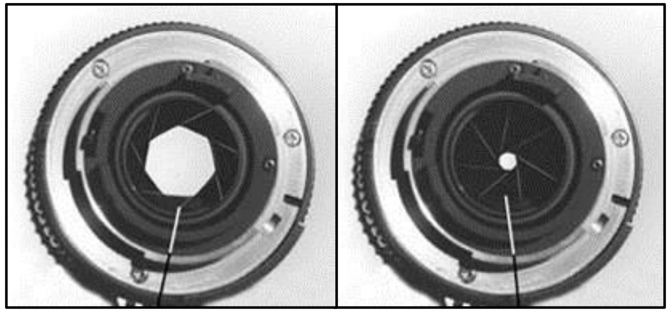


Ways to vary exposure

1. Shutter speed



2. F-stop (aperture, iris)



3. ISO

4. Neutral density (ND) filters

Pros and cons of each?

Ways to vary exposure

1. Shutter speed
   - Range: about 30 sec to 1/4000 sec (6 orders of magnitude)
   - Pros: repeatable, linear
   - Cons: noise and motion blur for long exposure
2. F-stop (aperture, iris)
   - Range: about f/0.98 to f/22 (3 orders of magnitude)
   - Pros: fully optical, no noise
   - Cons: changes depth of field
3. ISO
   - Range: about 100 to 1600 (1.5 orders of magnitude)
   - Pros: no movement at all
   - Cons: noise
3. Neutral density (ND) filters
   - Range: up to 6 densities (6 orders of magnitude)
   - Pros: works with strobe/flash
   - Cons: not perfectly neutral (color shift), extra glass (interreflections, aberrations), need to touch camera (shake)

Shutter speed

Note: shutter times usually obey a power series – each "stop" is a factor of 2

1/4, 1/8, 1/15, 1/30, 1/60, 1/125, 1/250, 1/500, 1/1000 sec

usually really is

1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512, 1/1024 sec

Questions:

1. How many exposures?
2. What exposures?

Shutter speed

Note: shutter times usually obey a power series – each “stop” is a factor of 2

1/4, 1/8, 1/15, 1/30, 1/60, 1/125, 1/250, 1/500, 1/1000 sec
usually really is

1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512, 1/1024 sec

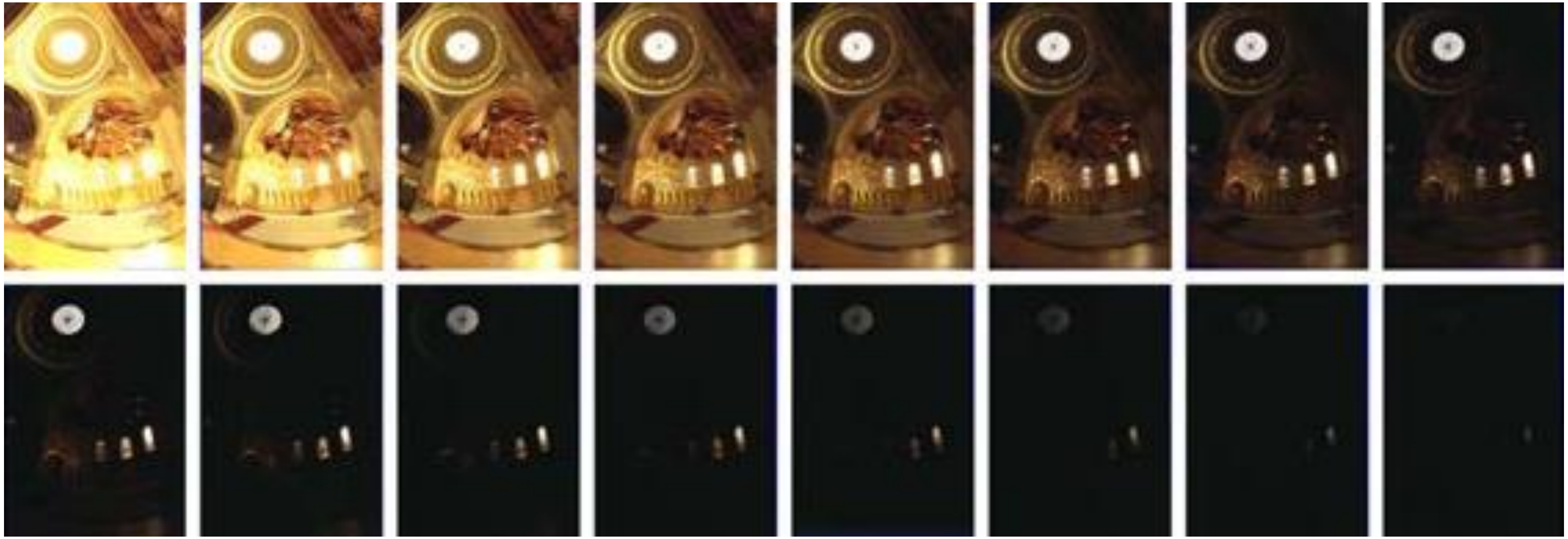
Questions:

1. How many exposures?
2. What exposures?

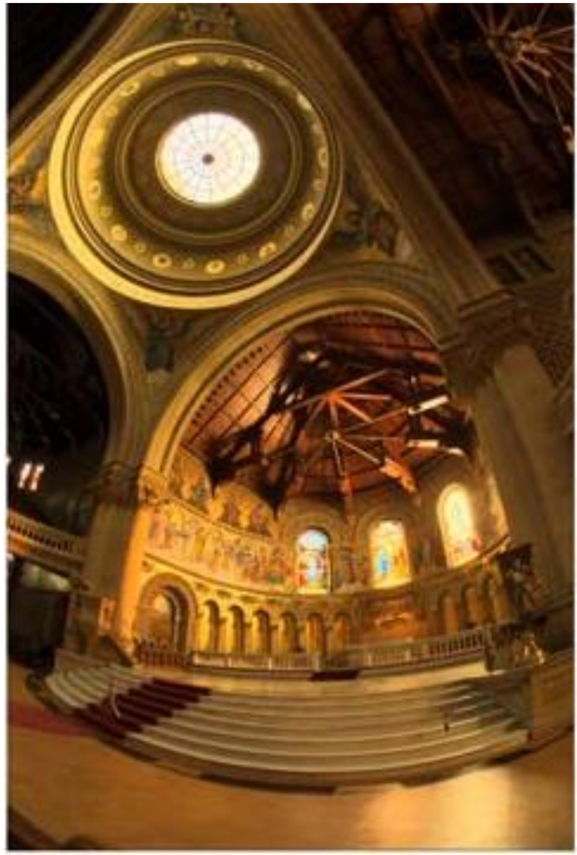
Answer: Depends on the scene, but a good default is 5 exposures, metered exposure and +- 2 stops around that

Key idea

1. Capture multiple LDR images at different exposures

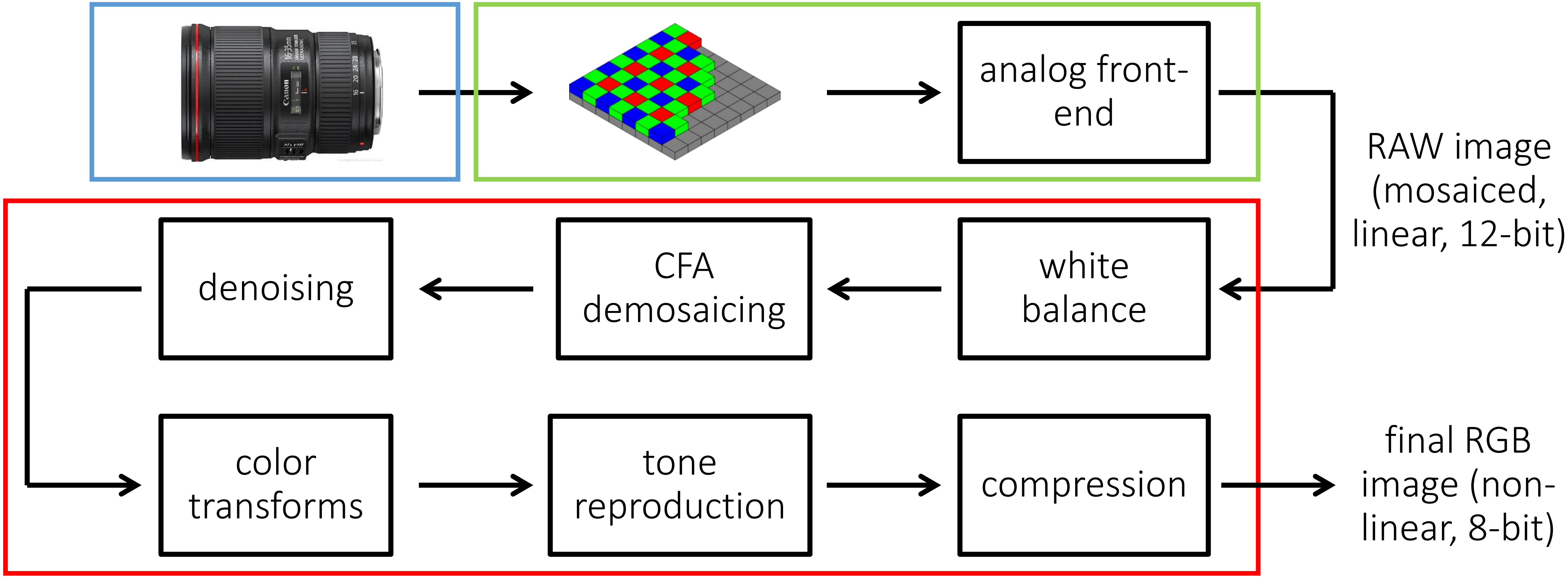


2. Merge them into a single HDR image



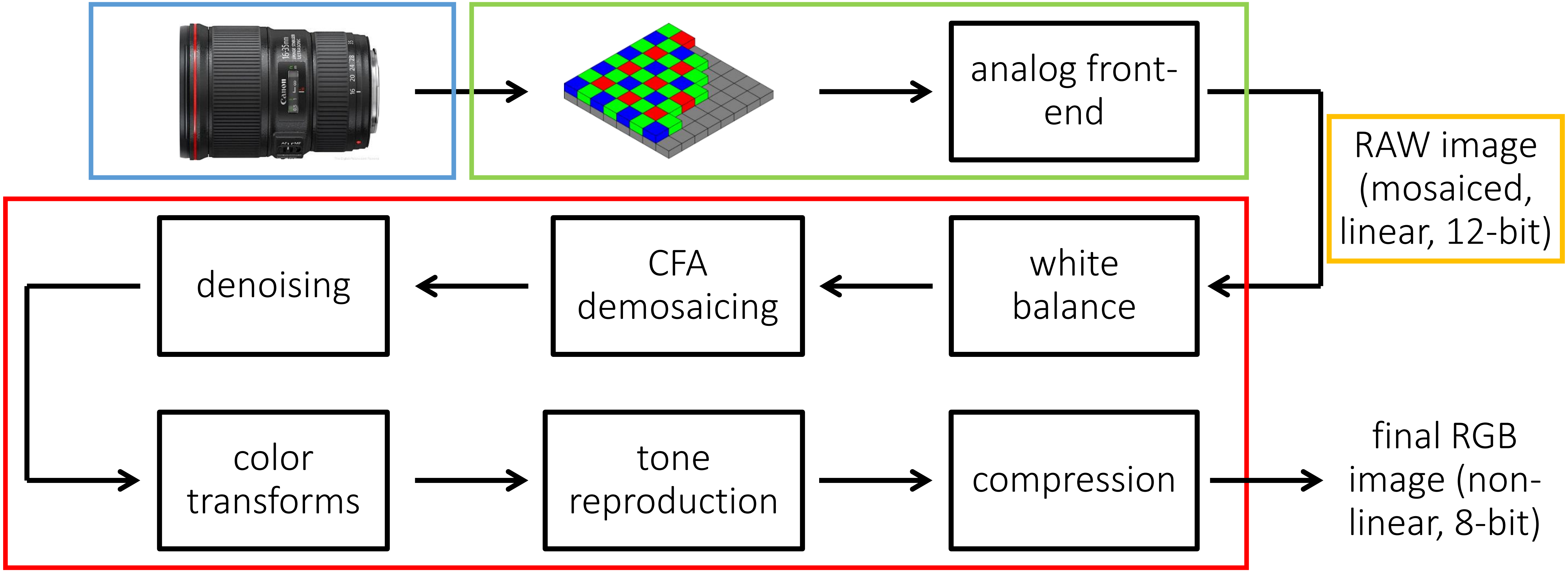
The image processing pipeline

The sequence of image processing operations applied by the camera's image signal processor (ISP) to convert a RAW image into a "conventional" image.



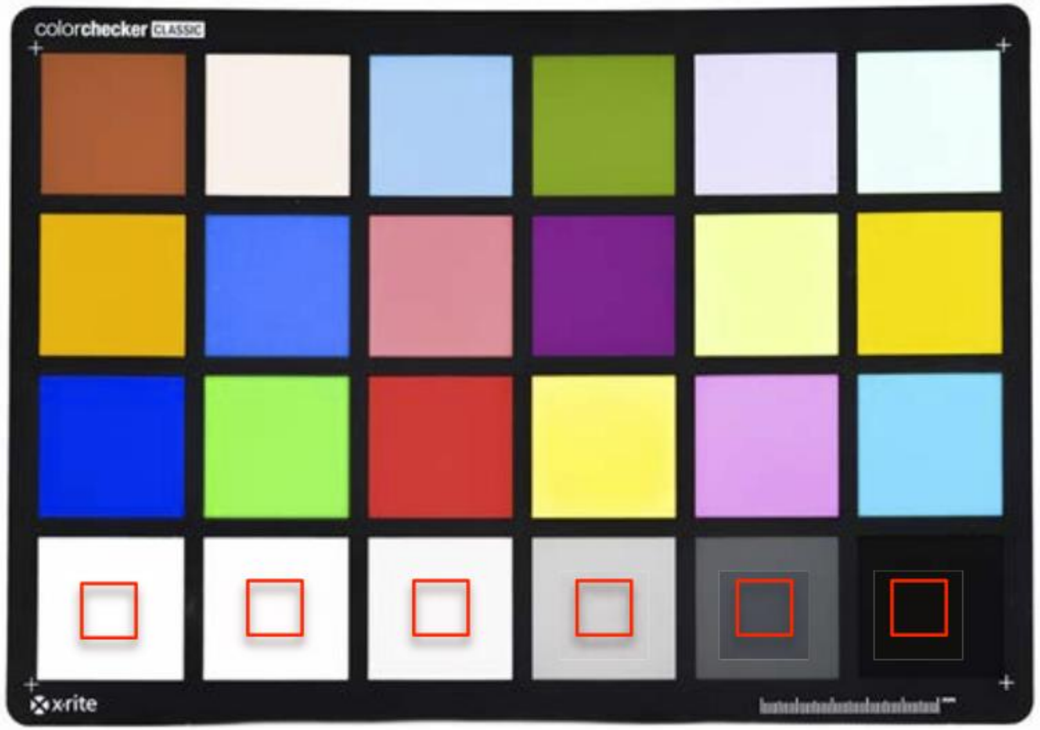
The image processing pipeline

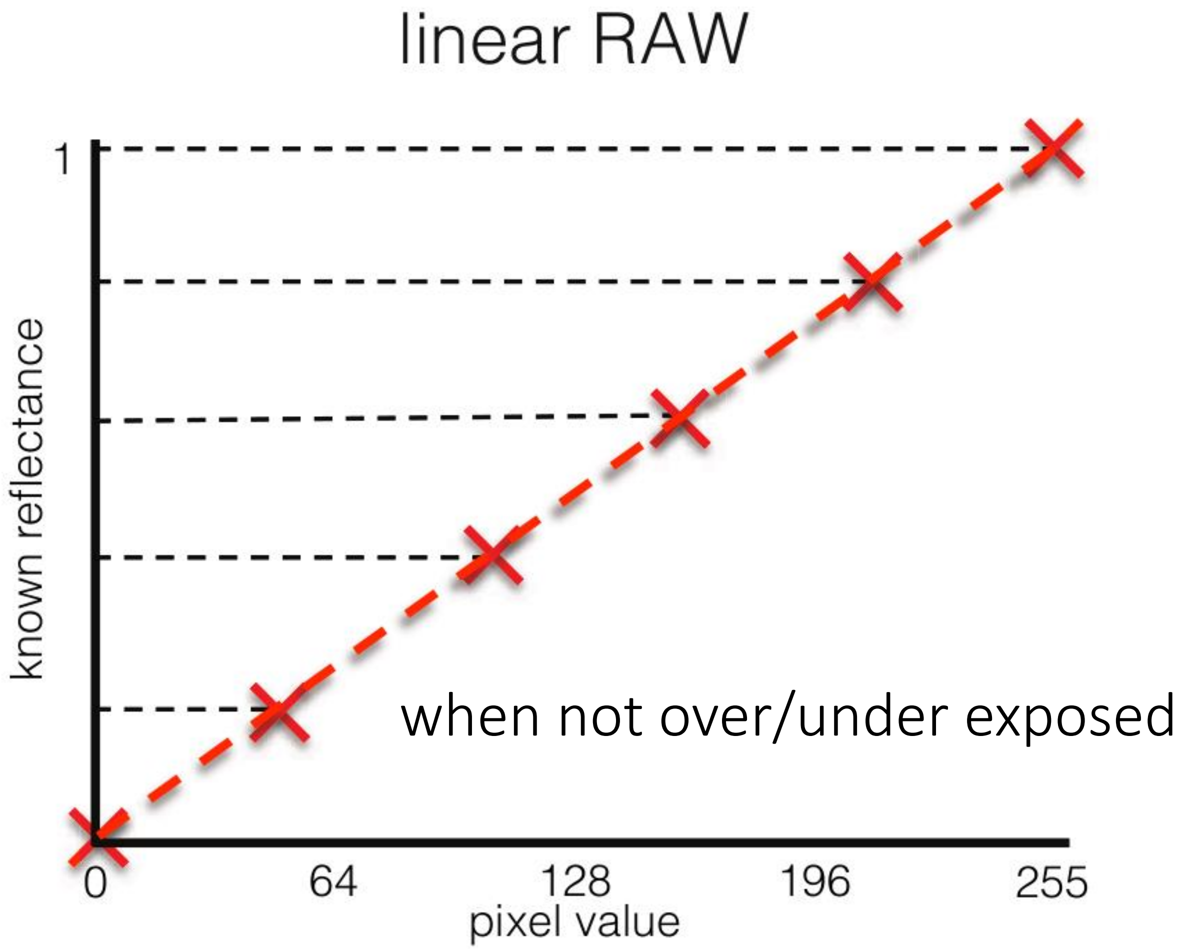
The sequence of image processing operations applied by the camera's image signal processor (ISP) to convert a RAW image into a "conventional" image.



RAW images have a linear response curve

<u>No need for calibration in this case</u>

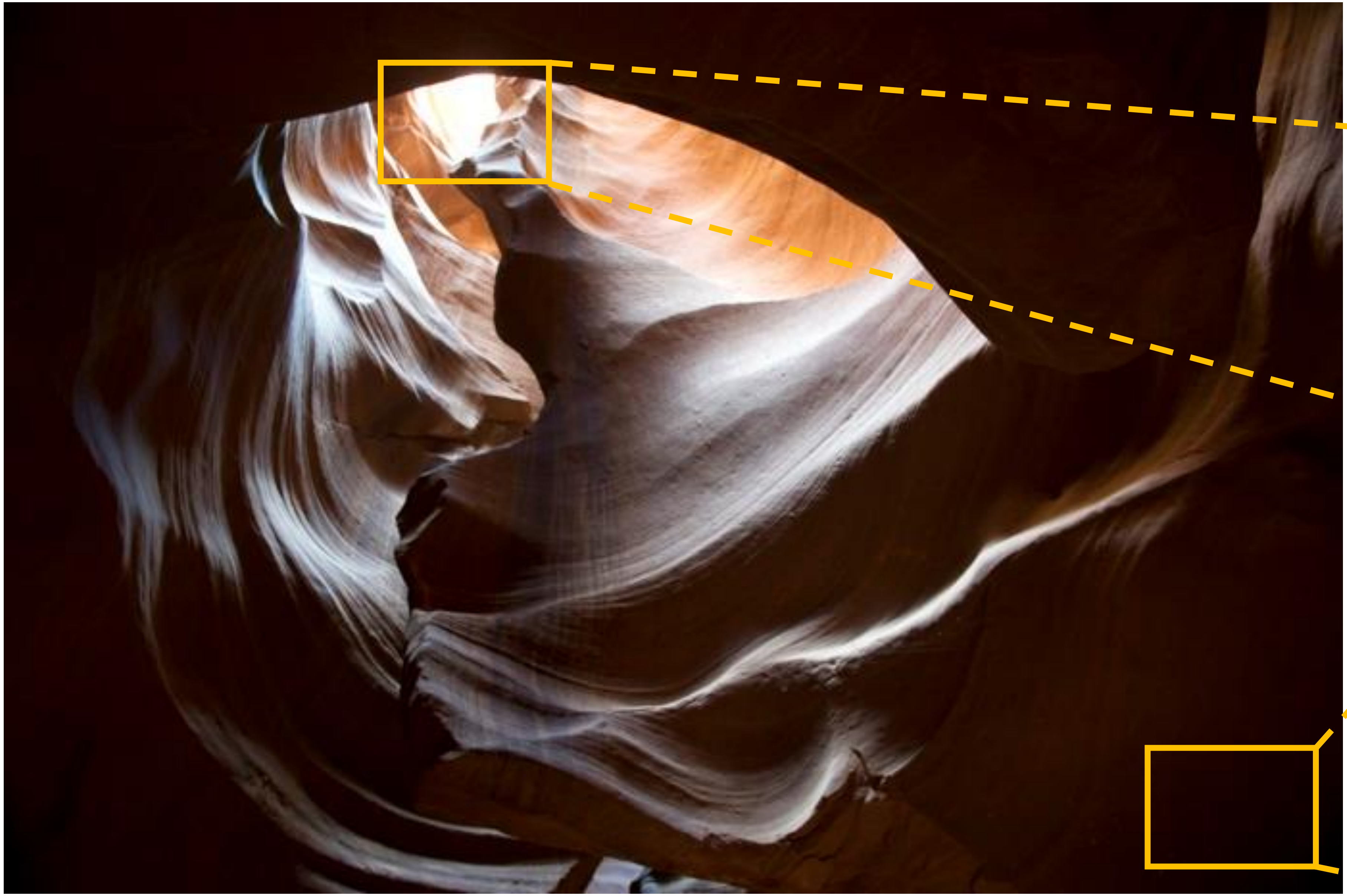




Calibration chart can be used for:

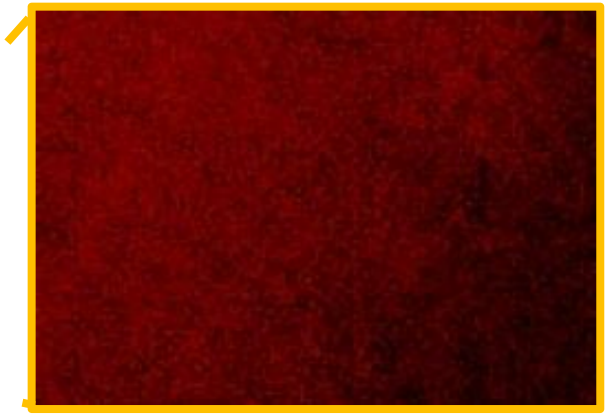
1. color calibration
2. radiometric calibration (i.e., response curve) using the bottom row

Over/under exposure



in highlights we are limited by clipping

in shadows we are limited by noise



RAW (linear) image formation model

Real scene radiance for image pixel (x,y): $L(x, y)$

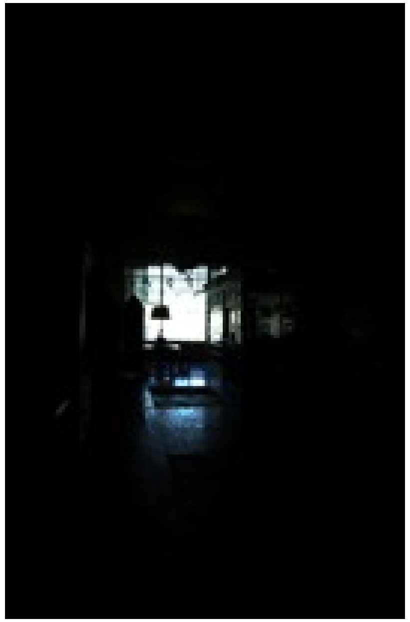
Exposure time:

t_5

t_4

t_3

t_2



t_1

What is an expression for the image I(x,y) as a function of L(x,y)?

RAW (linear) image formation model

Real scene radiance for image pixel (x,y): $L(x, y)$

Exposure time:

t_5 t_4 t_3 t_2 t_1

What is an expression for the image $I_{linear}(x,y)$ as a function of $L(x,y)$?

$$I_{linear}(x,y) = \text{clip}[\, t_i \cdot L(x,y) + \text{noise} \,]$$

How would you merge these images into an HDR one?

Merging RAW (linear) exposure stacks

For each pixel:

1. Find “valid” images

2. Weight valid pixel values appropriately

3. Form a new pixel value as the weighted average of valid pixel values

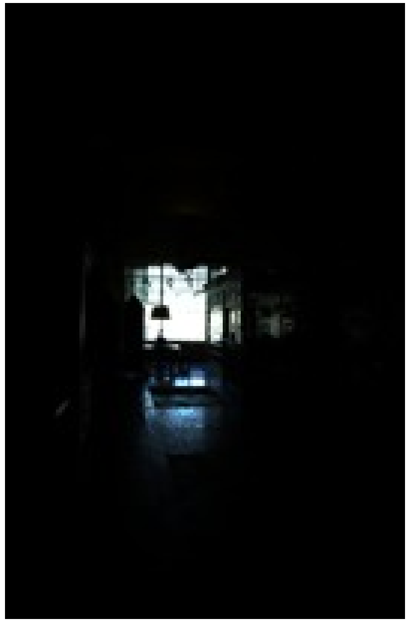
How would you implement steps 1-2?

t_5

t_4

t_3

t_2



t_1

Merging RAW (linear) exposure stacks

For each pixel:

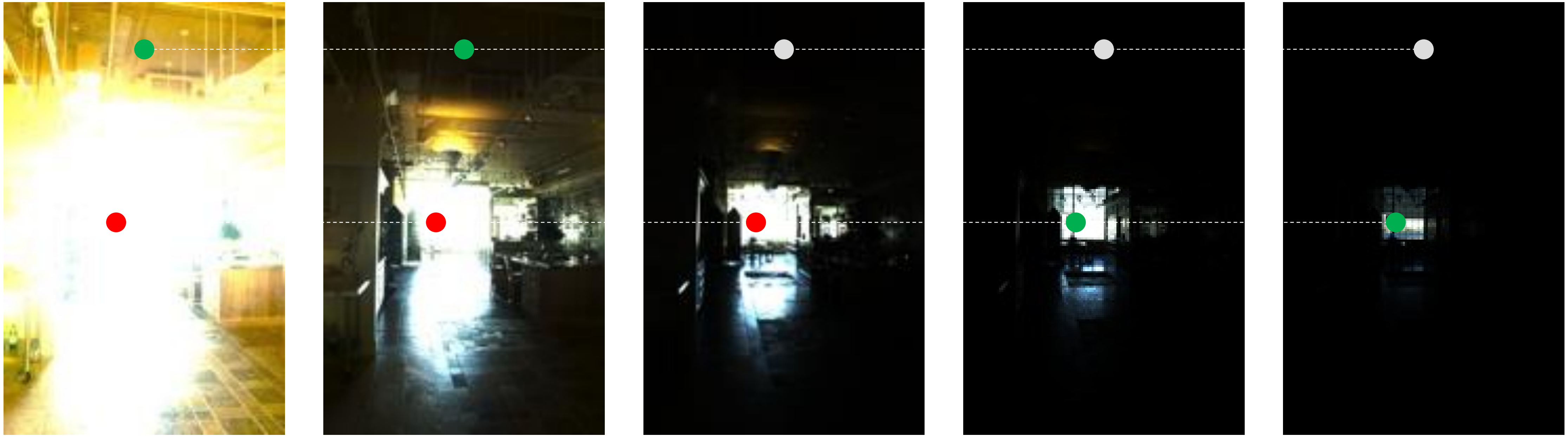
1. Find “valid” images ← (noise) 0.05 < pixel < 0.95 (clipping)

2. Weight valid pixel values appropriately

3. Form a new pixel value as the weighted average of valid pixel values

noise
valid
clipped

t_5 t_4 t_3 t_2 t_1



Merging RAW (linear) exposure stacks

For each pixel:

1. Find "valid" images ⟵ (noise) $0.05 < \text{pixel} < 0.95$ (clipping)

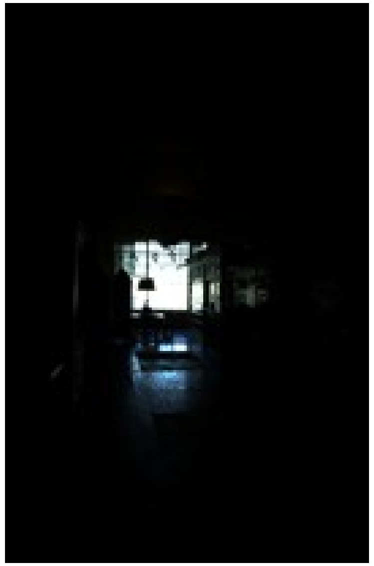
2. Weight valid pixel values appropriately ⟵ (pixel value) / t_i

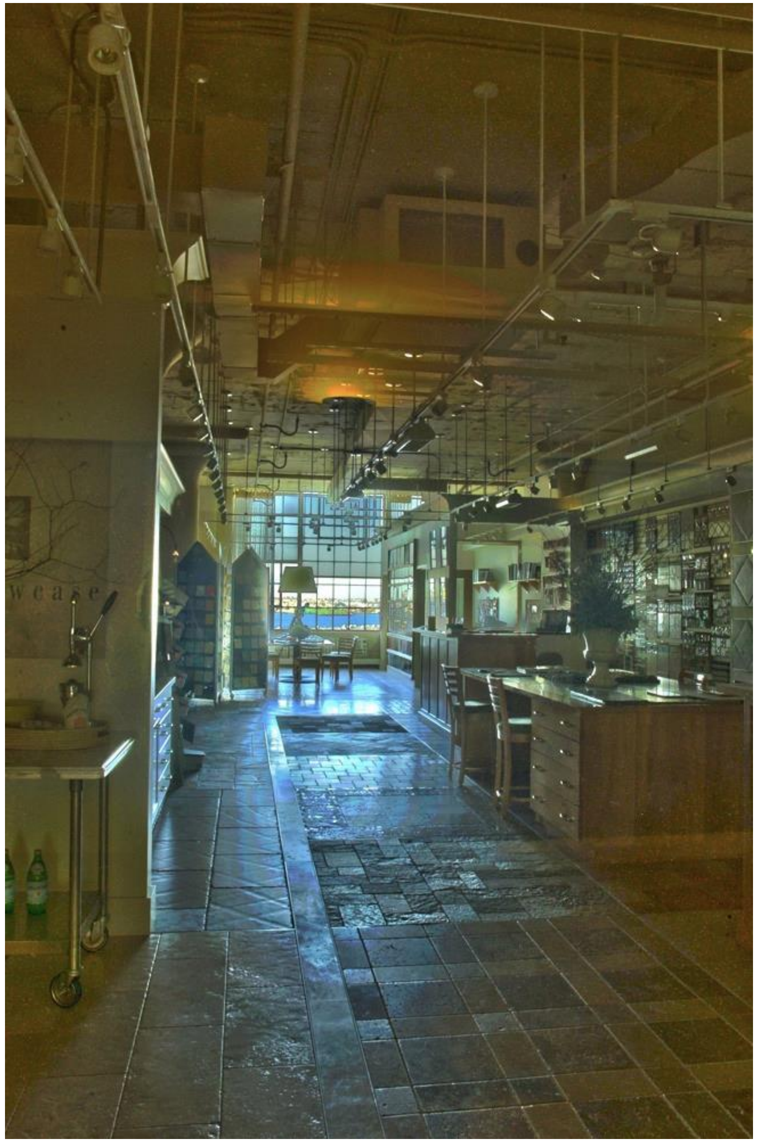
3. Form a new pixel value as the weighted average of valid pixel values

t_5 t_4 t_3 t_2 t_1

Merging result (after tonemapping)

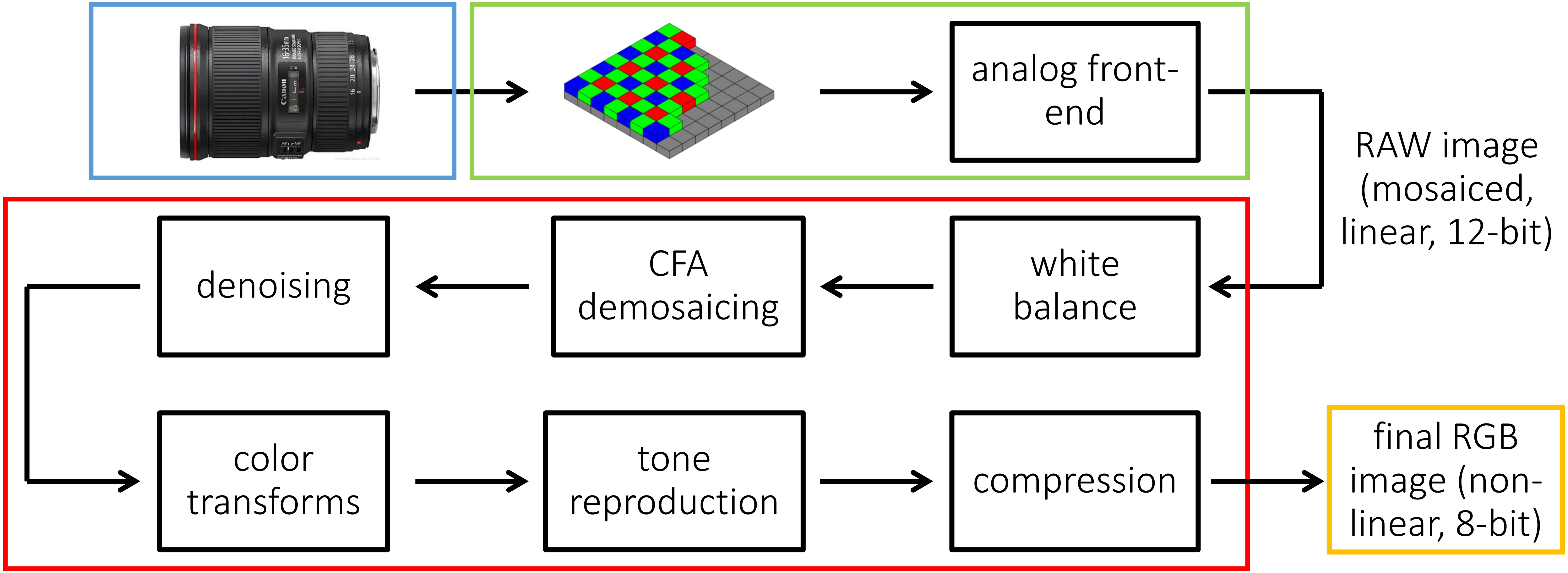




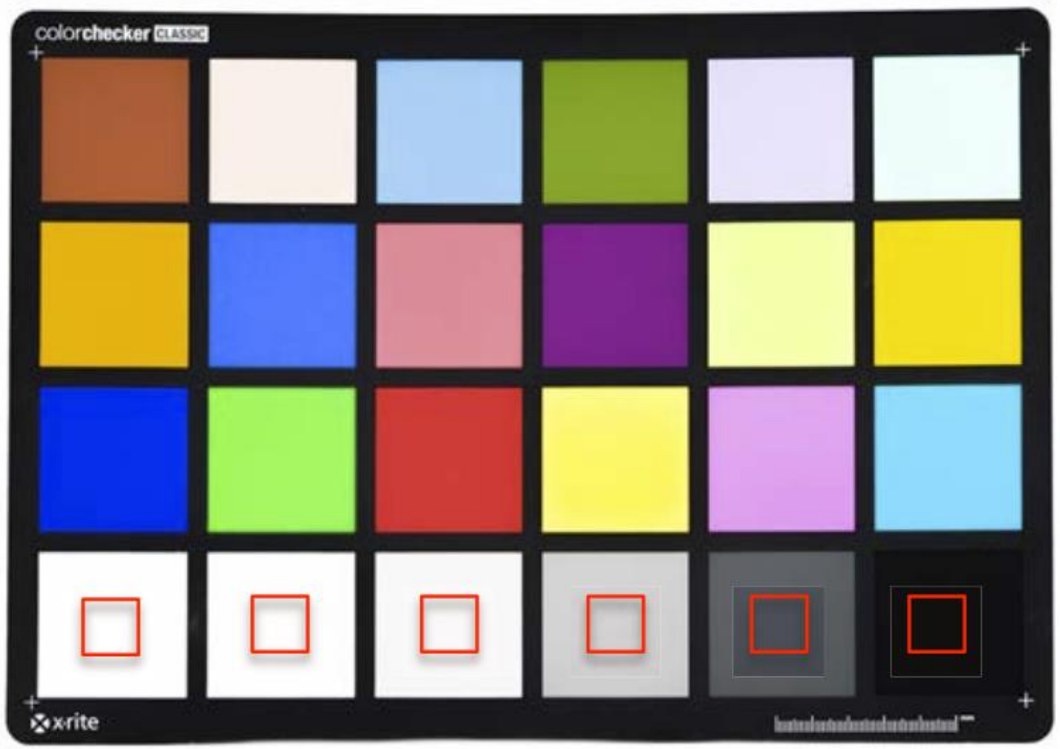
The image processing pipeline

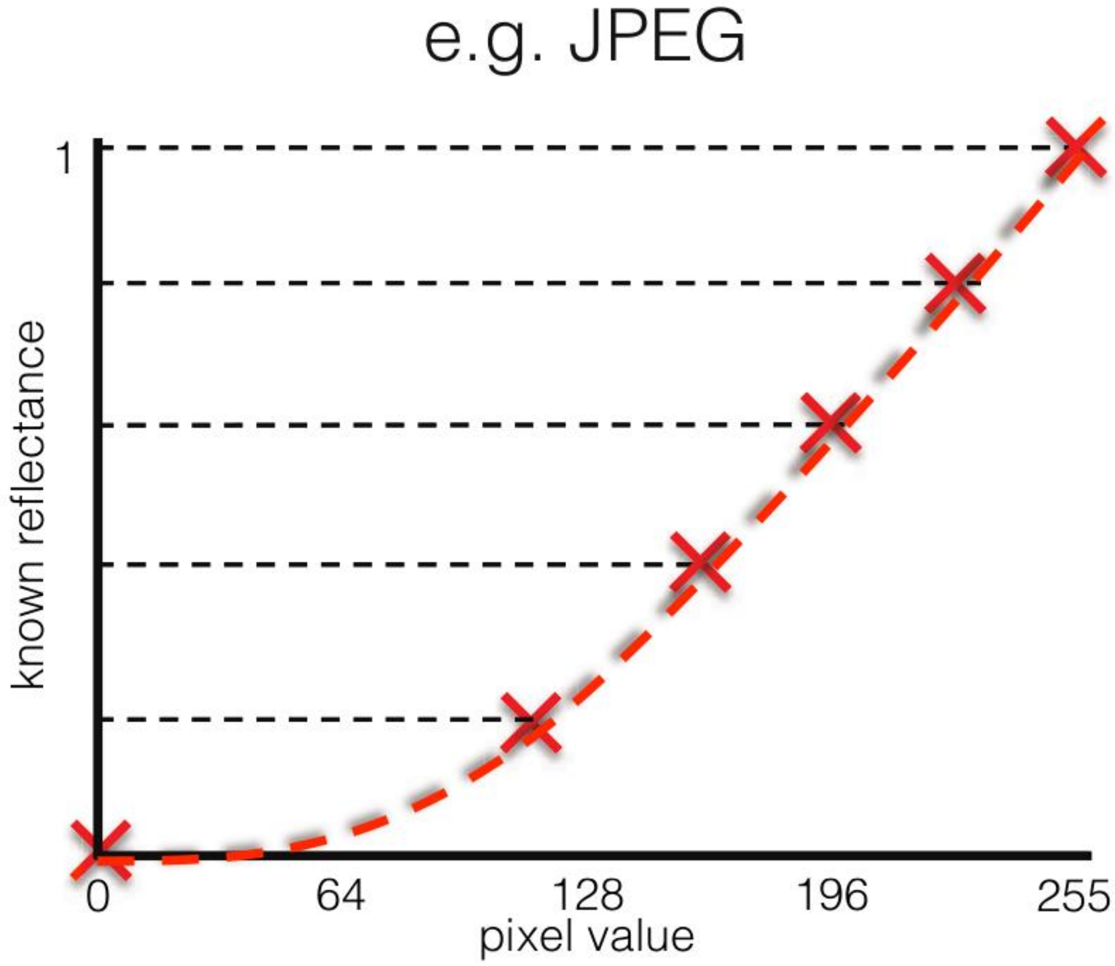
The sequence of image processing operations applied by the camera's image signal processor (ISP) to convert a RAW image into a "conventional" image.



Processed images have a non-linear response curve

We must calibrate the response curve



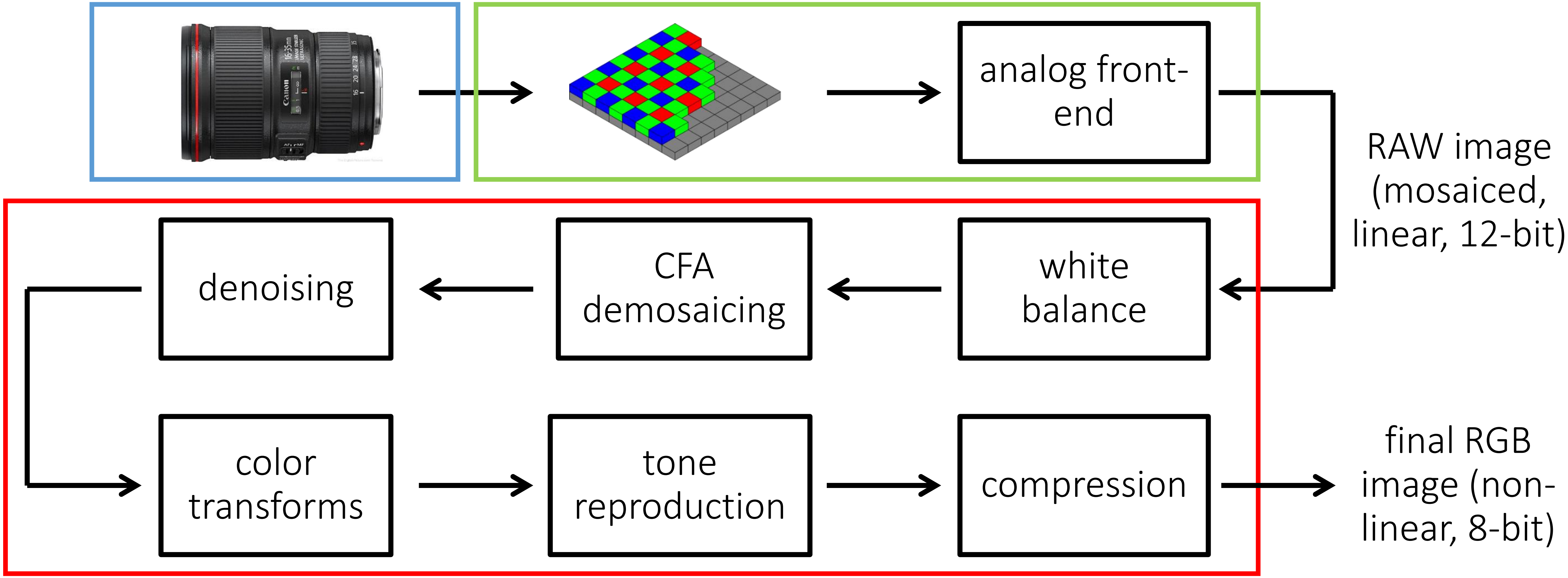


Calibration chart can be used for:

1. color calibration
2. radiometric calibration (i.e., response curve) using the bottom row

The image processing pipeline

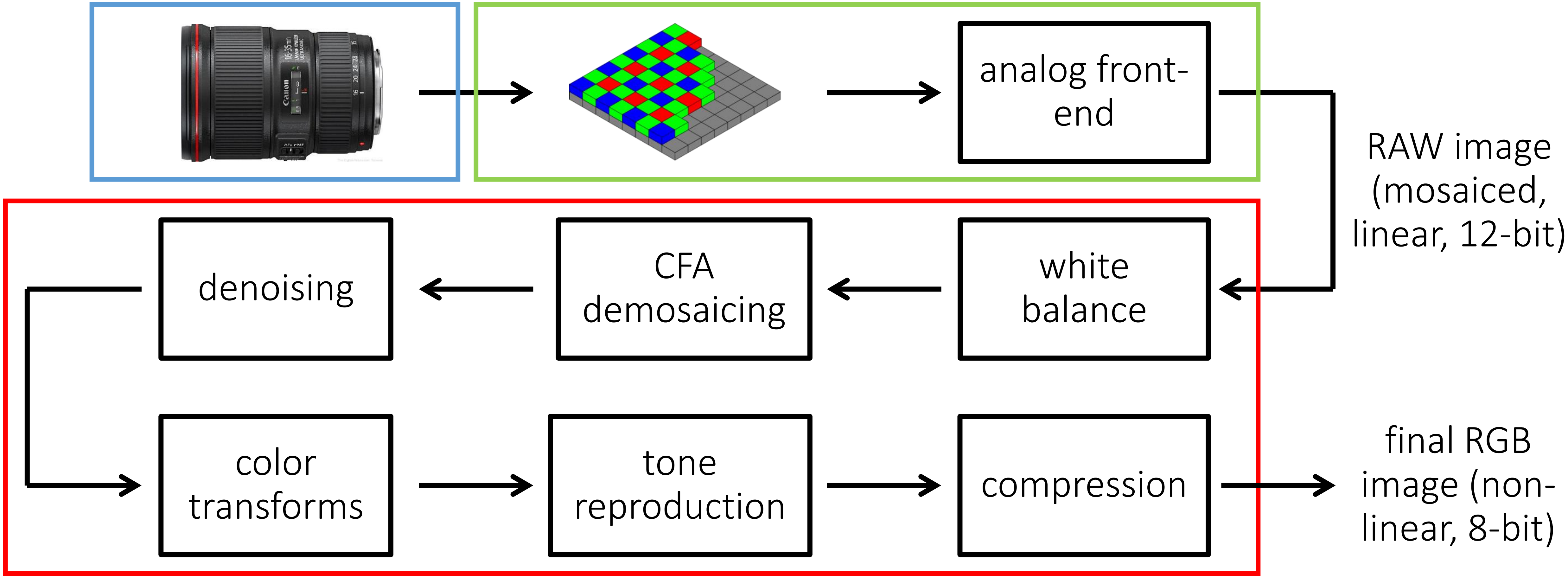
Which part of the pipeline does the non-linear response curve correspond to?



The image processing pipeline

Which part of the pipeline does the non-linear response curve correspond to?

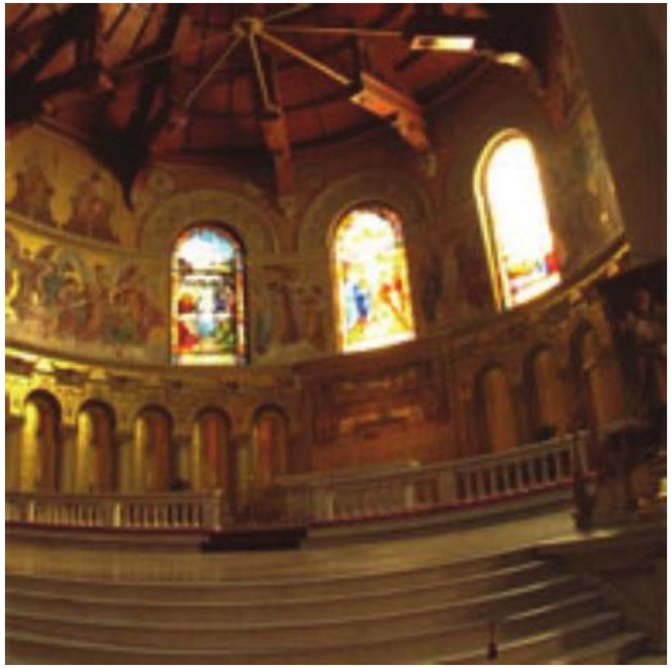
- The tone reproduction (mostly).

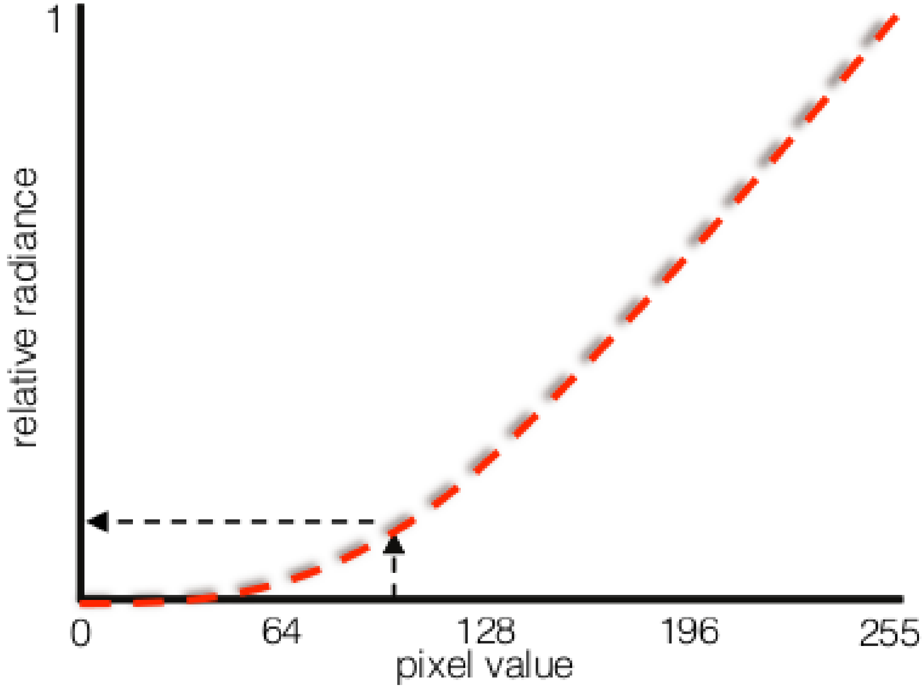


Non-linear image formation model

Real scene radiance for image pixel (x,y): $L(x, y)$

Exposure time: t_i





$$I_{linear}(x,y) = \text{clip}[\, t_i \cdot L(x,y) + \text{noise} \,]$$

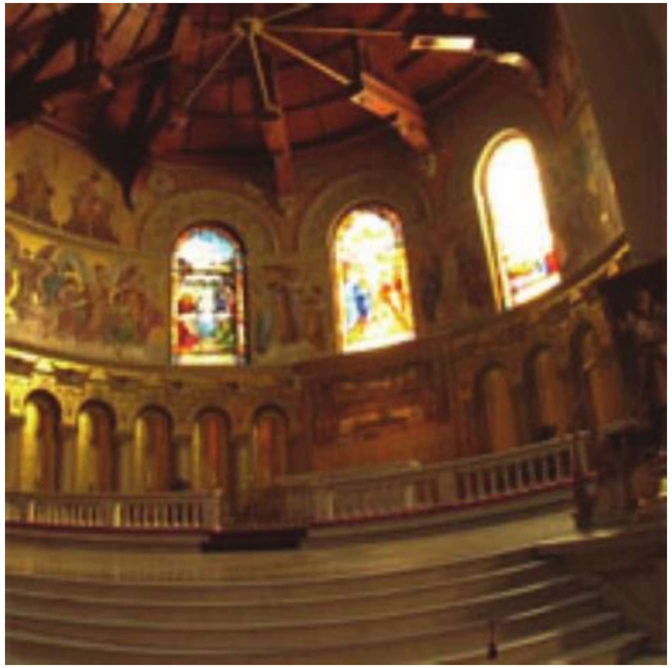
$$I_{non\text{-}linear}(x,y) = f[\, I_{linear}(x,y) \,]$$

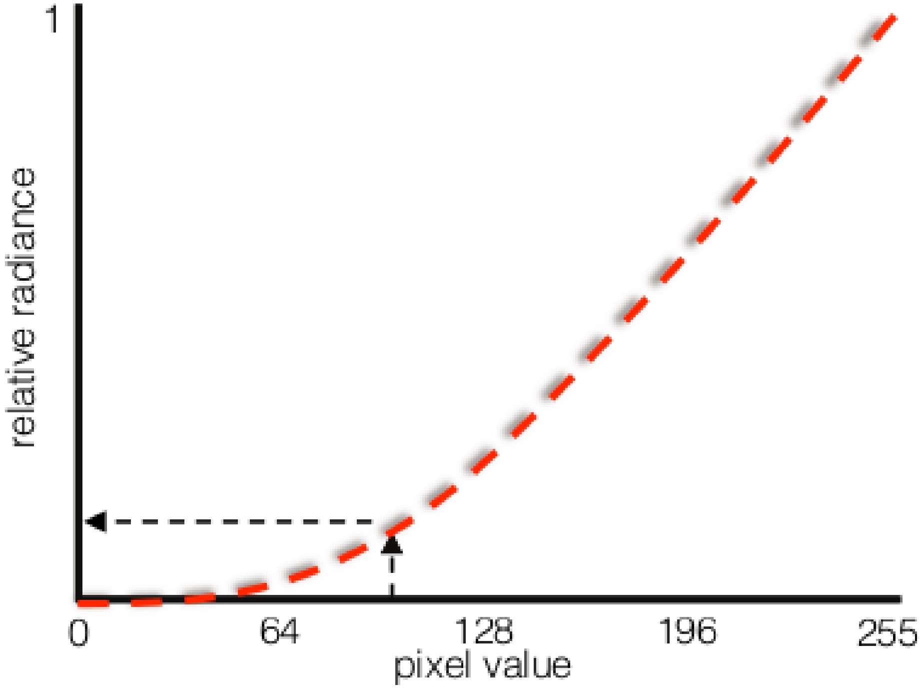
How would you merge the non-linear images into an HDR one?

Non-linear image formation model

Real scene radiance for image pixel (x,y): L(x, y)

Exposure time: t_i





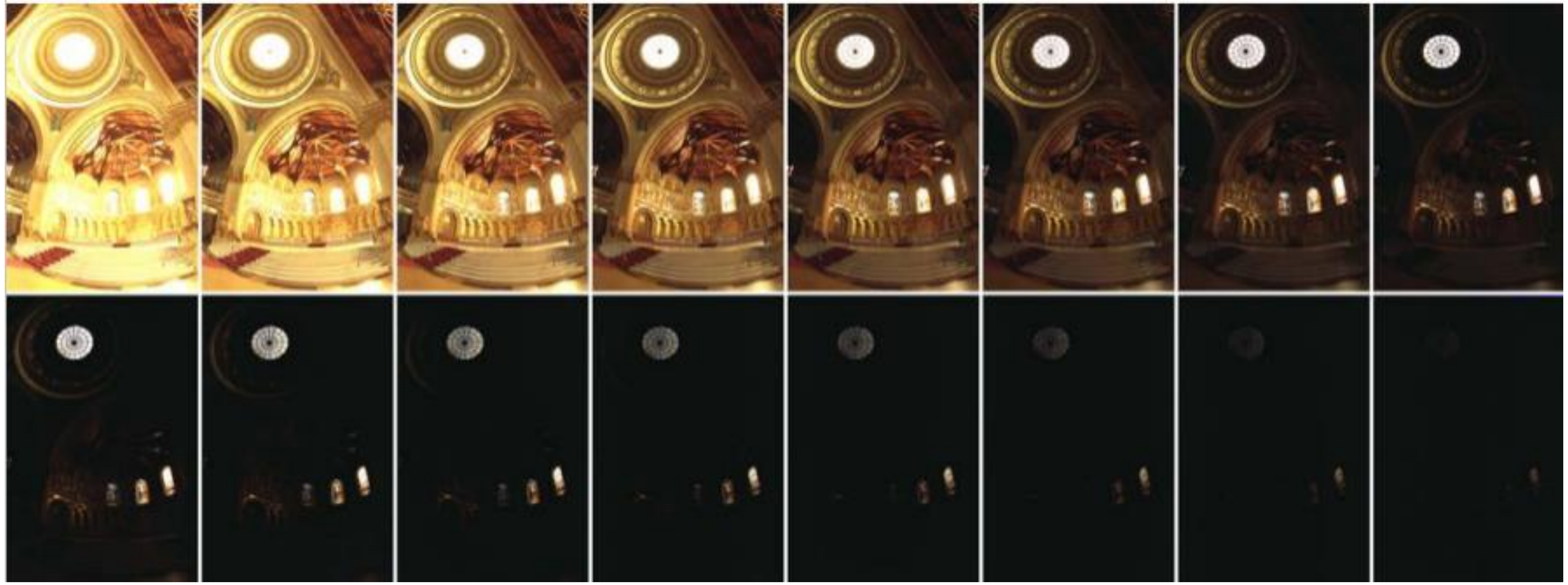
$I_{linear}(x,y) = \text{clip}[\ t_i \cdot L(x,y) + \text{noise}\ ]$

$I_{non\text{-}linear}(x,y) = f[\ I_{linear}(x,y)\ ]$ $I_{est}(x,y) = f^{-1}[\ I_{non\text{-}linear}(x,y)\ ]$

Use inverse transform to estimate linear image, then proceed as before

Linearization

$$I_{non\text{-}linear}(x,y) = f[\ I_{linear}(x,y)\ ]$$



$$I_{est}(x,y) = f^{-1}[\ I_{non\text{-}linear}(x,y)\ ]$$

Merging non-linear exposure stacks

1. Calibrate response curve
2. Linearize images

For each pixel:

3. Find “valid” images ← (noise) $0.05 < \text{pixel} < 0.95$ (clipping)
4. Weight valid pixel values appropriately ← (pixel value) / t_i
5. Form a new pixel value as the weighted average of valid pixel values

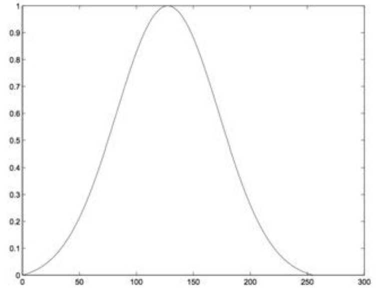
Note: many possible weighting schemes

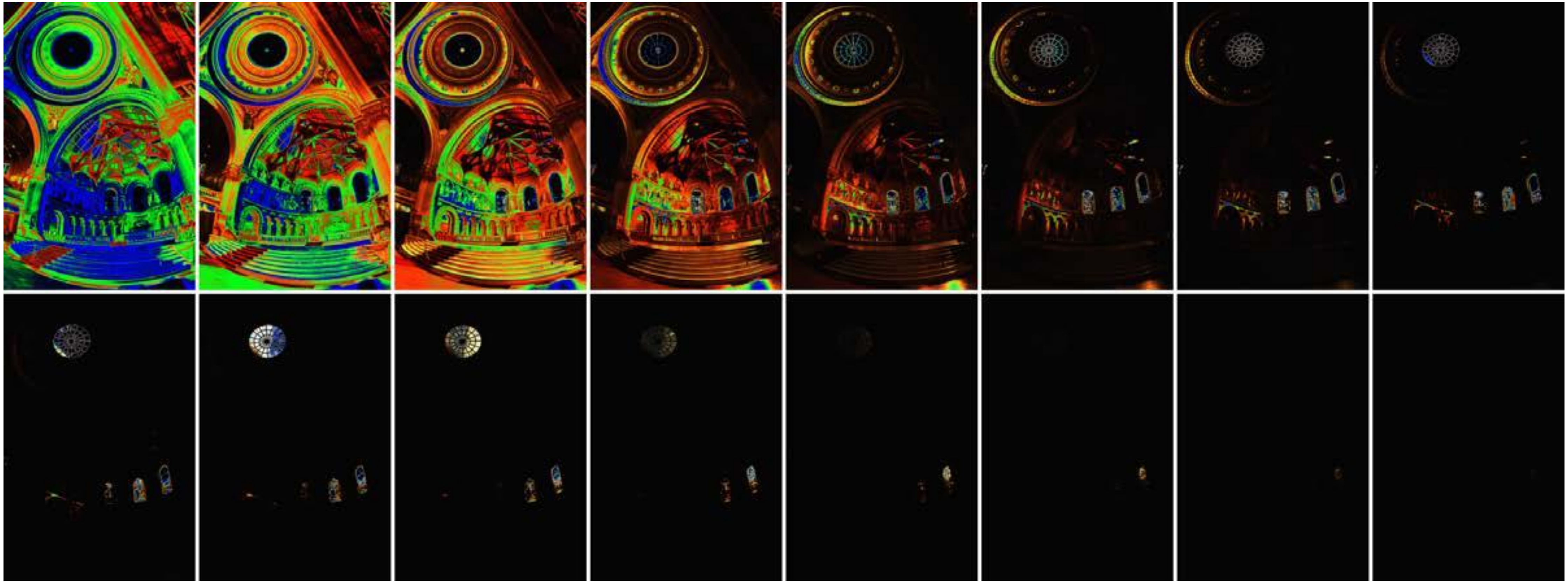
Many possible weighting schemes

You will see more in Homework 4

"Confidence" that pixel is noisy/clipped

$$w_{ij} = \exp\left(-4 \frac{\left(I_{lin_{ij}} - 0.5\right)^2}{0.5^2} \right)$$

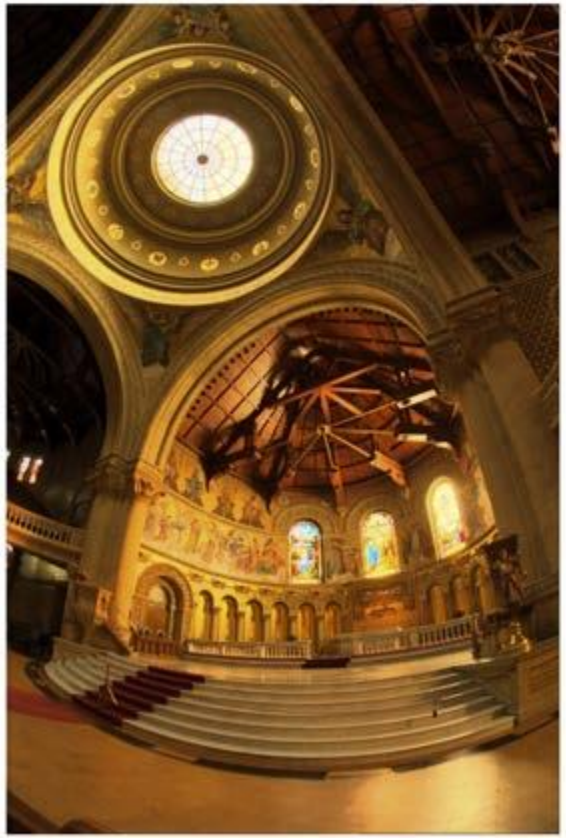




Relative vs absolute radiance

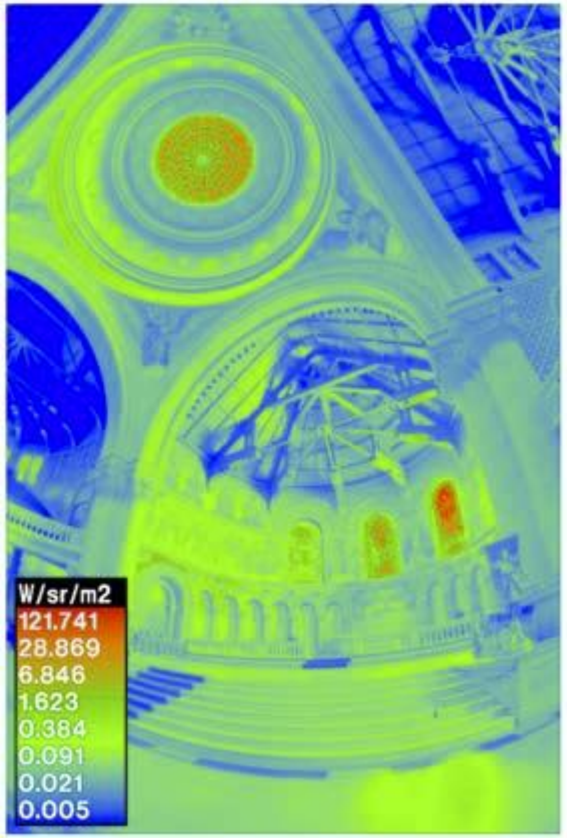
Final fused HDR image gives radiance only up to a global scale

- If we know exact radiance at one point, we can convert relative HDR image to absolute radiance map



HDR image (relative radiance)

spotmeter (absolute radiance at one point)



absolute radiance map

Basic HDR approach

1. Capture multiple LDR images at different exposures

2. Merge them into a single HDR image

Any problems with this approach?

Basic HDR approach

1. Capture multiple LDR images at different exposures

2. Merge them into a single HDR image

Problem: Very sensitive to movement

- Scene must be completely static

- Camera must not move

Most modern automatic HDR solutions include an alignment step before merging exposures

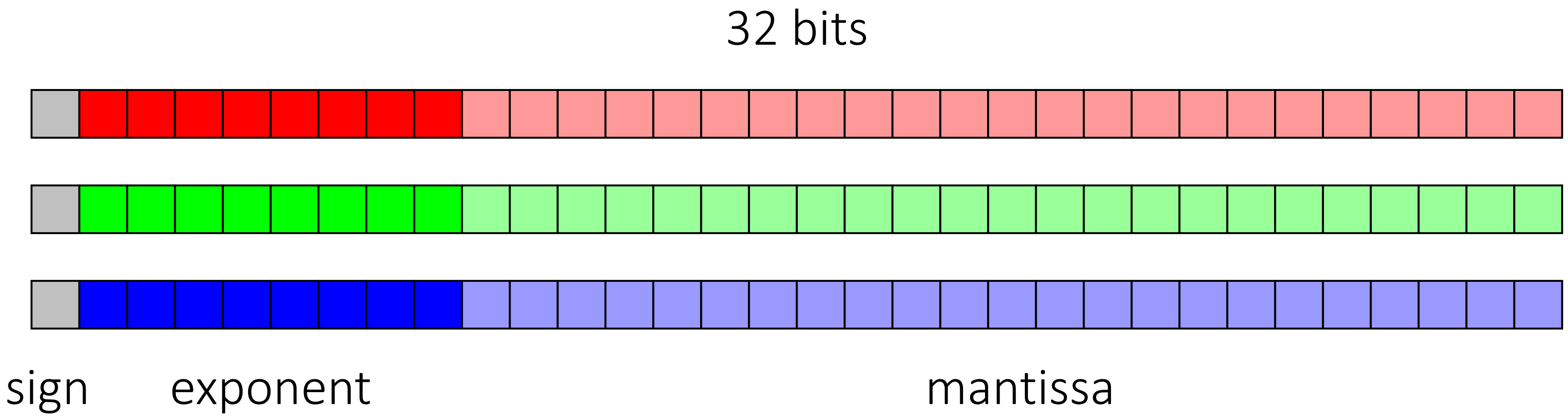
How do we store HDR images?

- Most standard image formats store integer 8-bit images
- Some image formats store integer 12-bit or 16-bit images
- HDR images are floating point 32-bit or 64-bit images

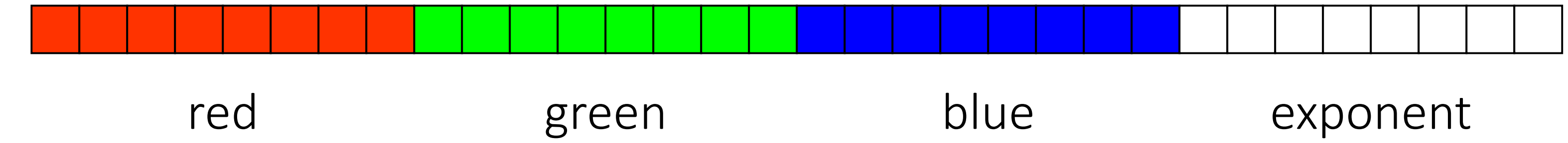
How do we store HDR images?

Use specialized image formats for HDR images

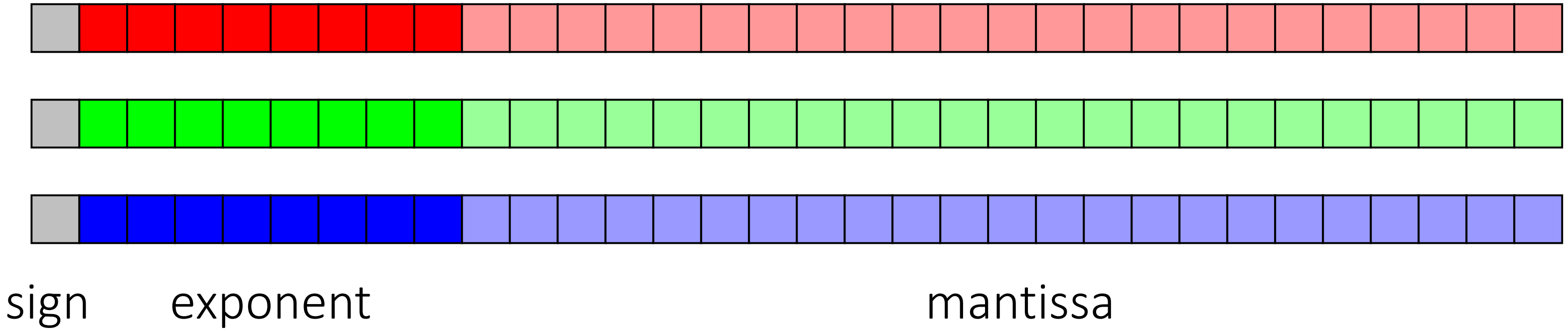
portable float map (.pfm)
- very simple to implement



Radiance format (.hdr)
- supported by Matlab



OpenEXR format (.exr)
- multiple extra features



Another type of HDR images

Light probes: place a chrome sphere in the scene and capture an HDR image

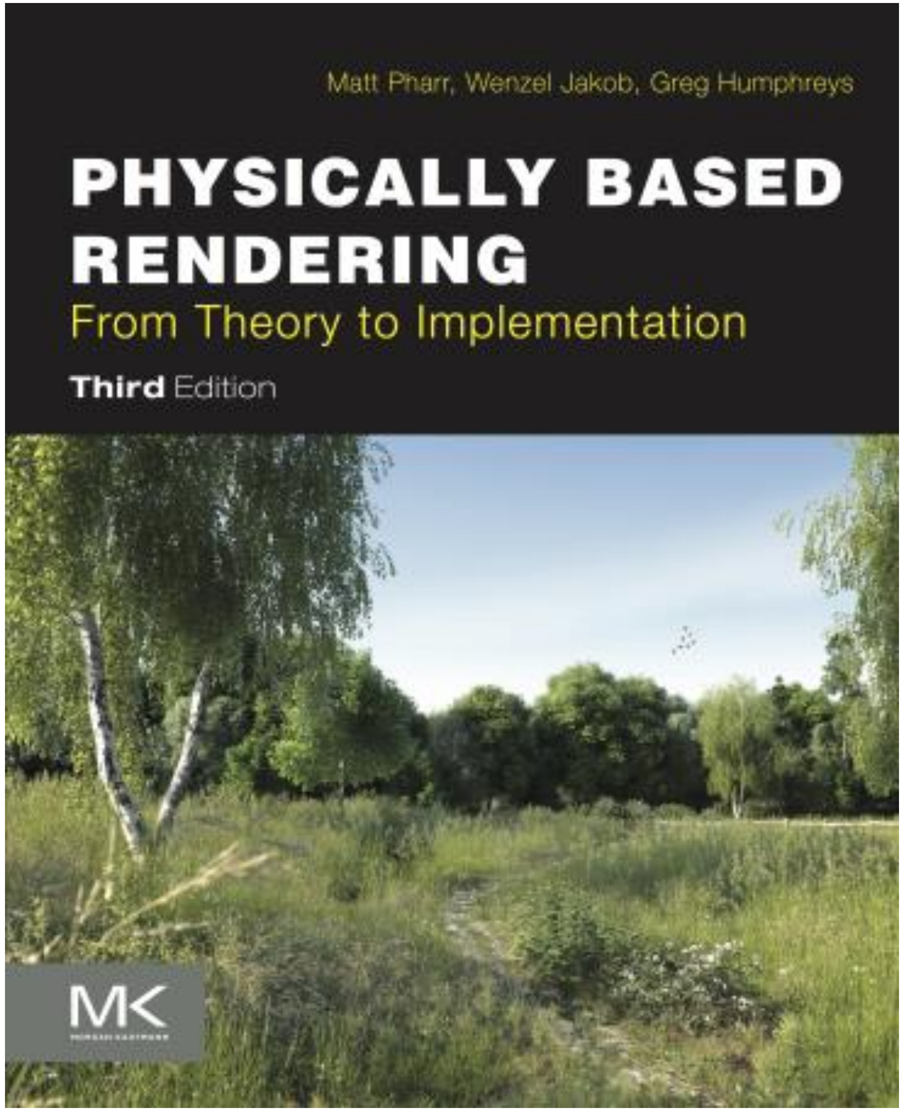
- Used to measure real-world illumination environments ("environment maps")

Application: image-based relighting (later lecture)

Another way to create HDR images

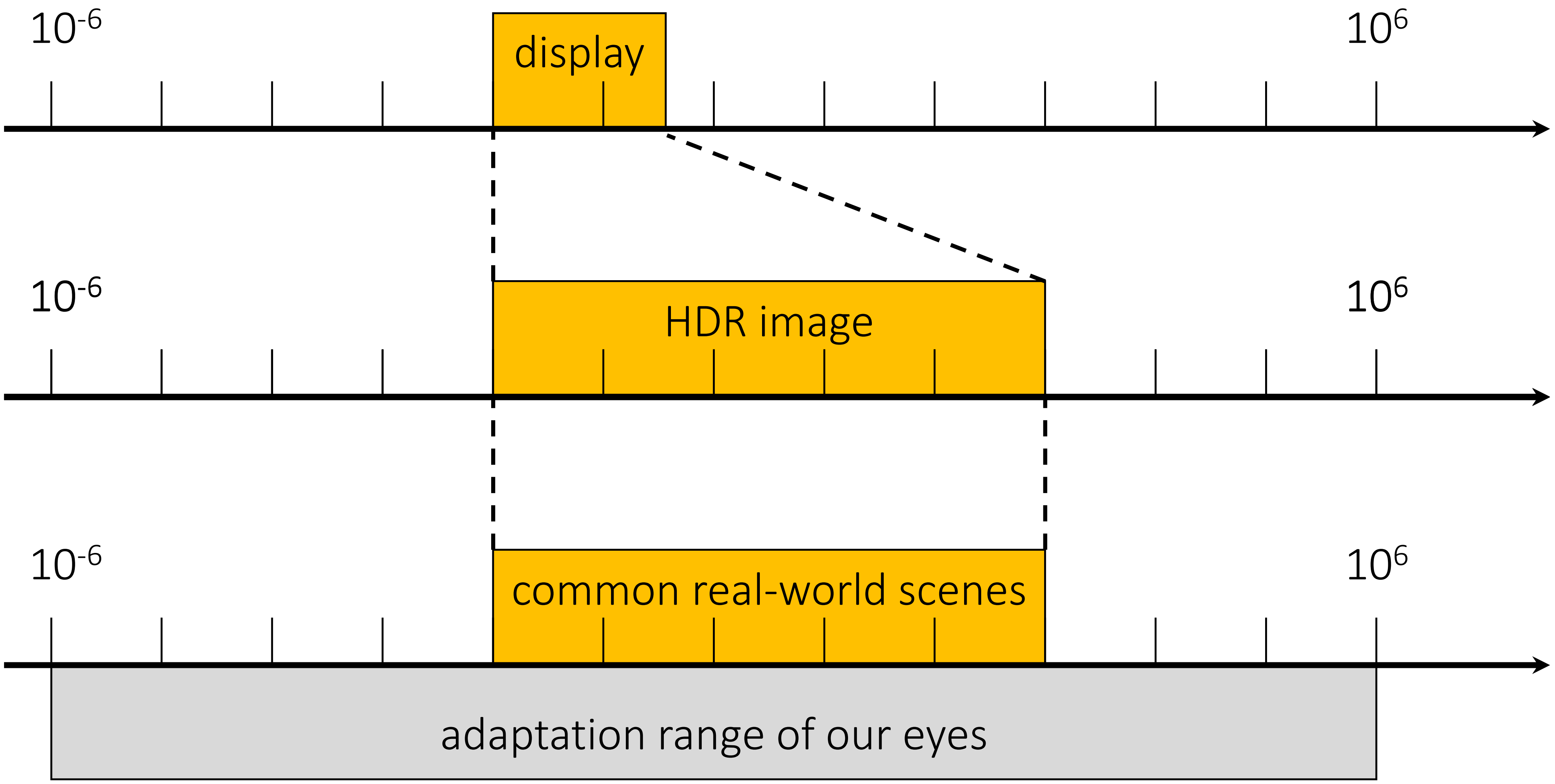
Physics-based renderers simulate radiance maps (relative or absolute)

- Their outputs are very often HDR images



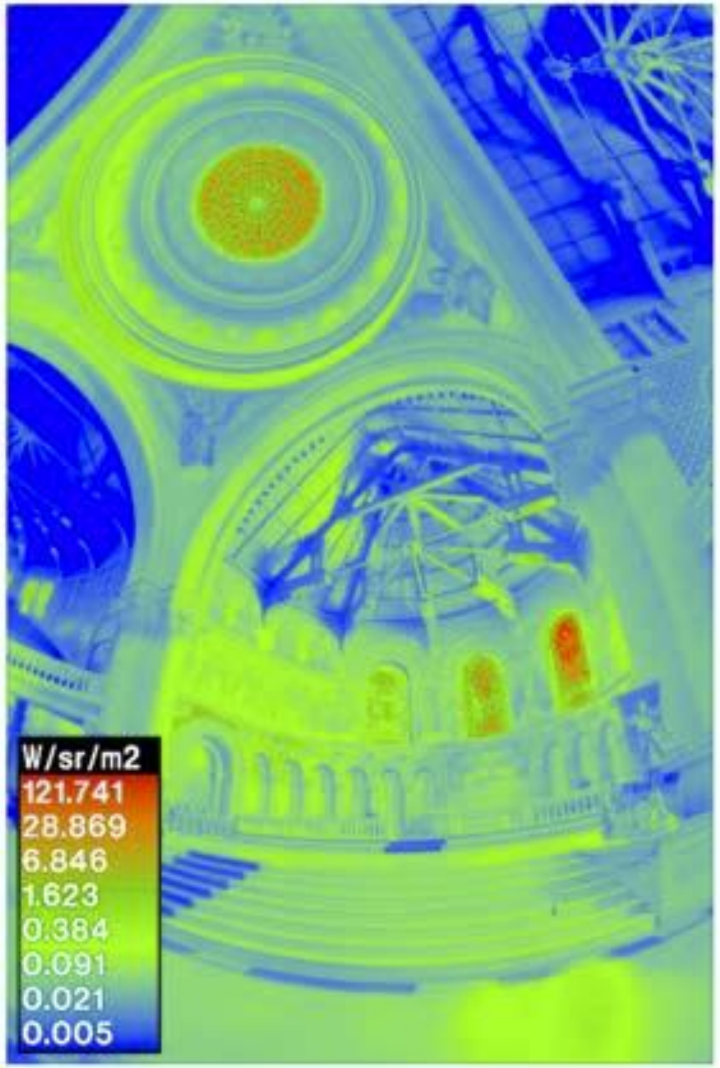
Tonemapping

How do we display our HDR images?



Linear scaling

Scale image so that maximum value equals 1

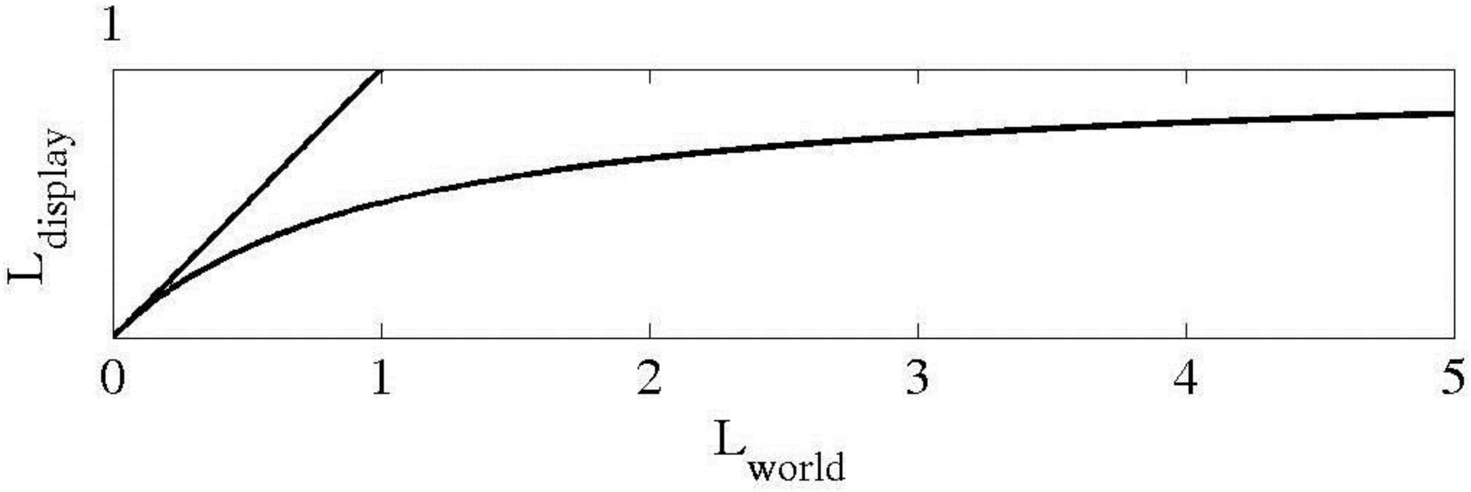


Can you think of something better?

Photographic tonemapping

Apply the same non-linear scaling to all pixels in the image so that:

- Bring everything within range → asymptote to 1
- Leave dark areas alone → slope = 1 near 0



$$I_{display} = \frac{I_{HDR}}{1 + I_{HDR}}$$

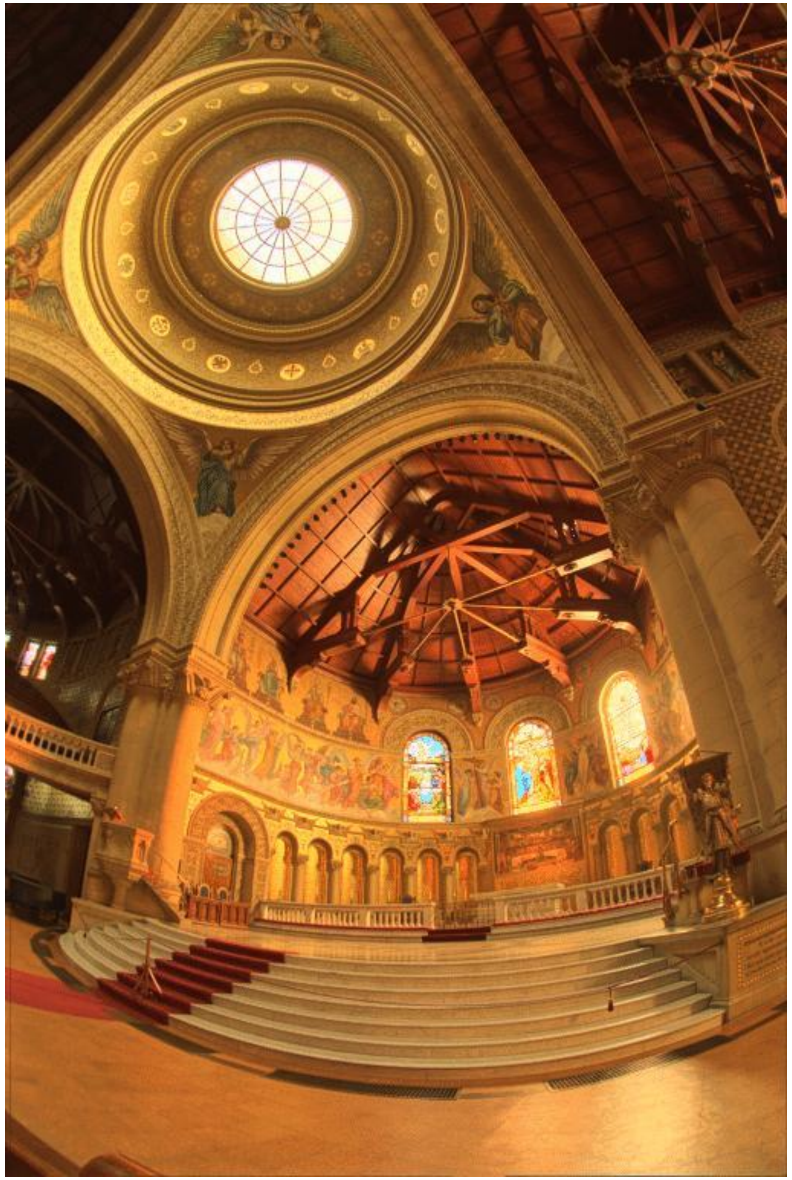
(exact formula more complicated)

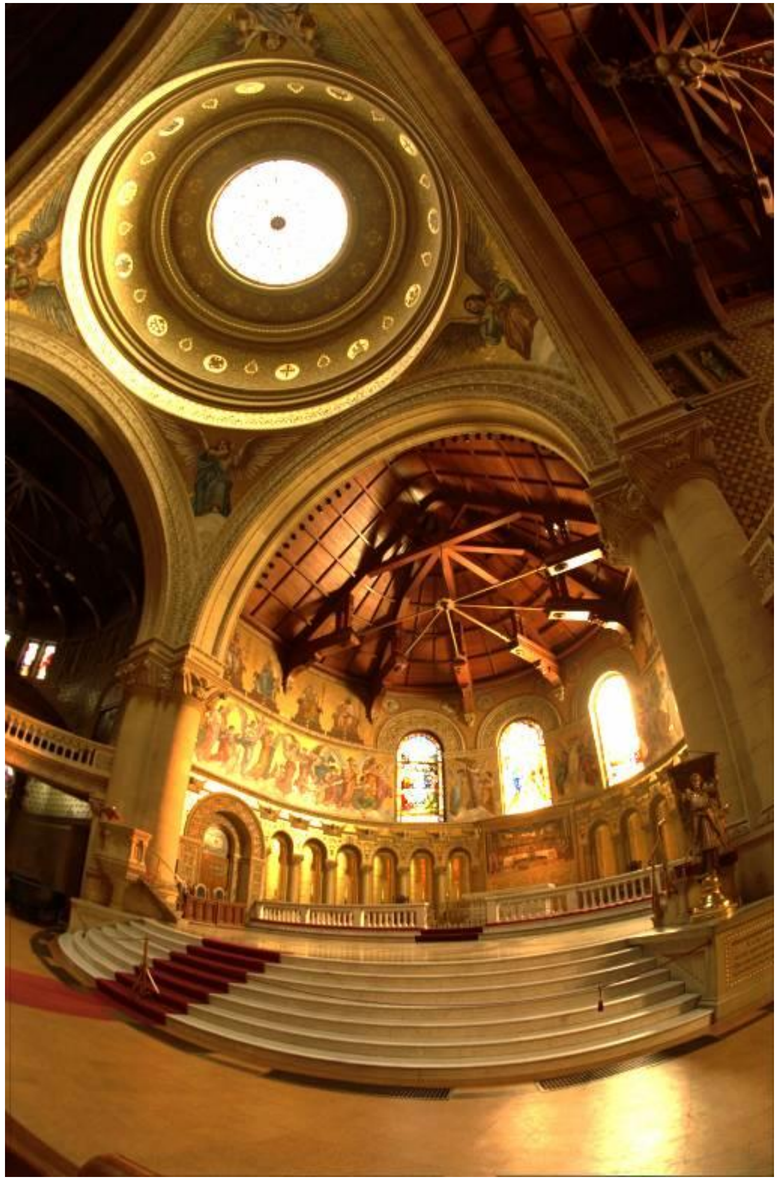
- Photographic because designed to approximate film zone system
- Also perceptually motivated

Examples

Examples

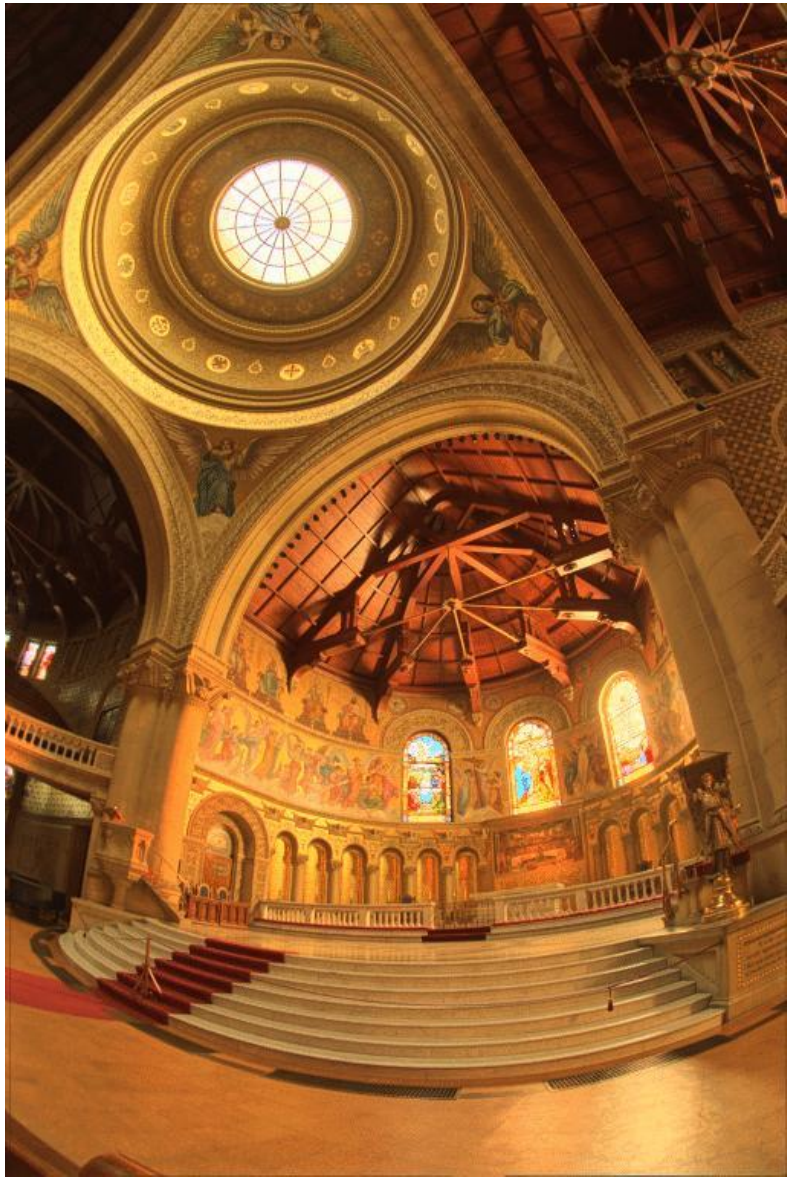
photographic
tonemapping

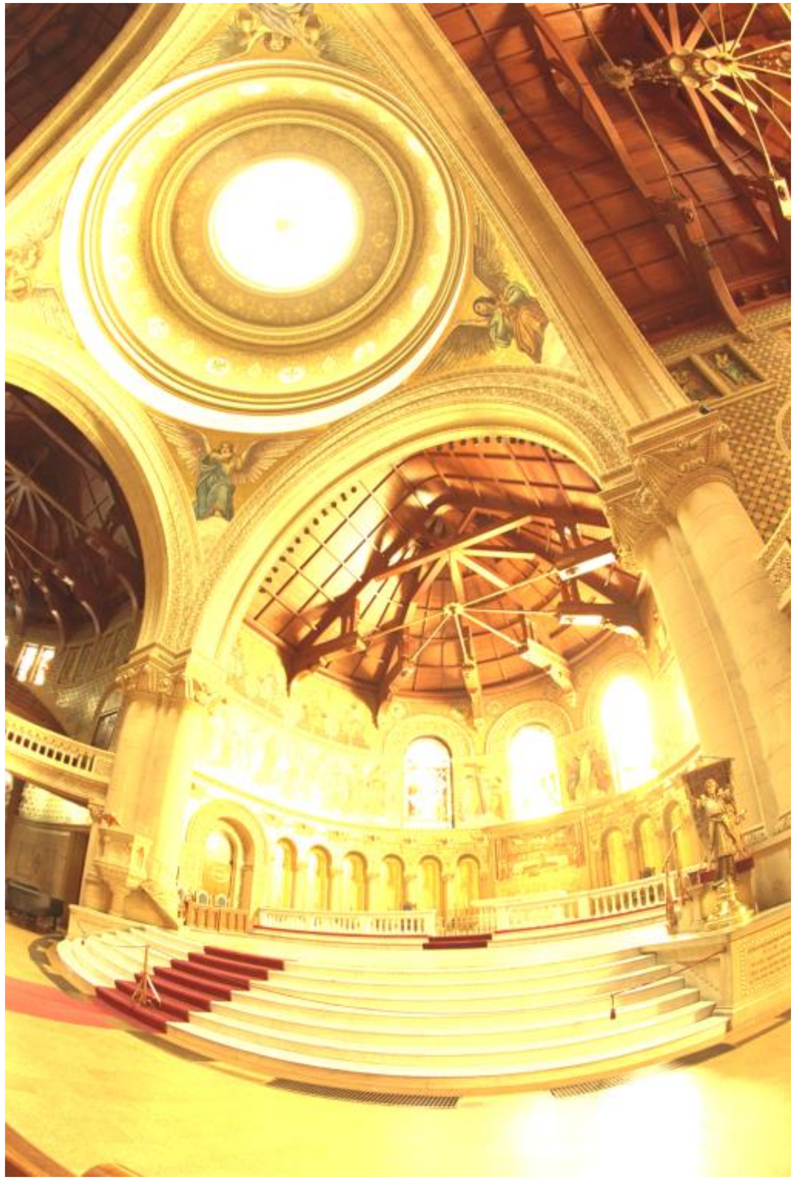


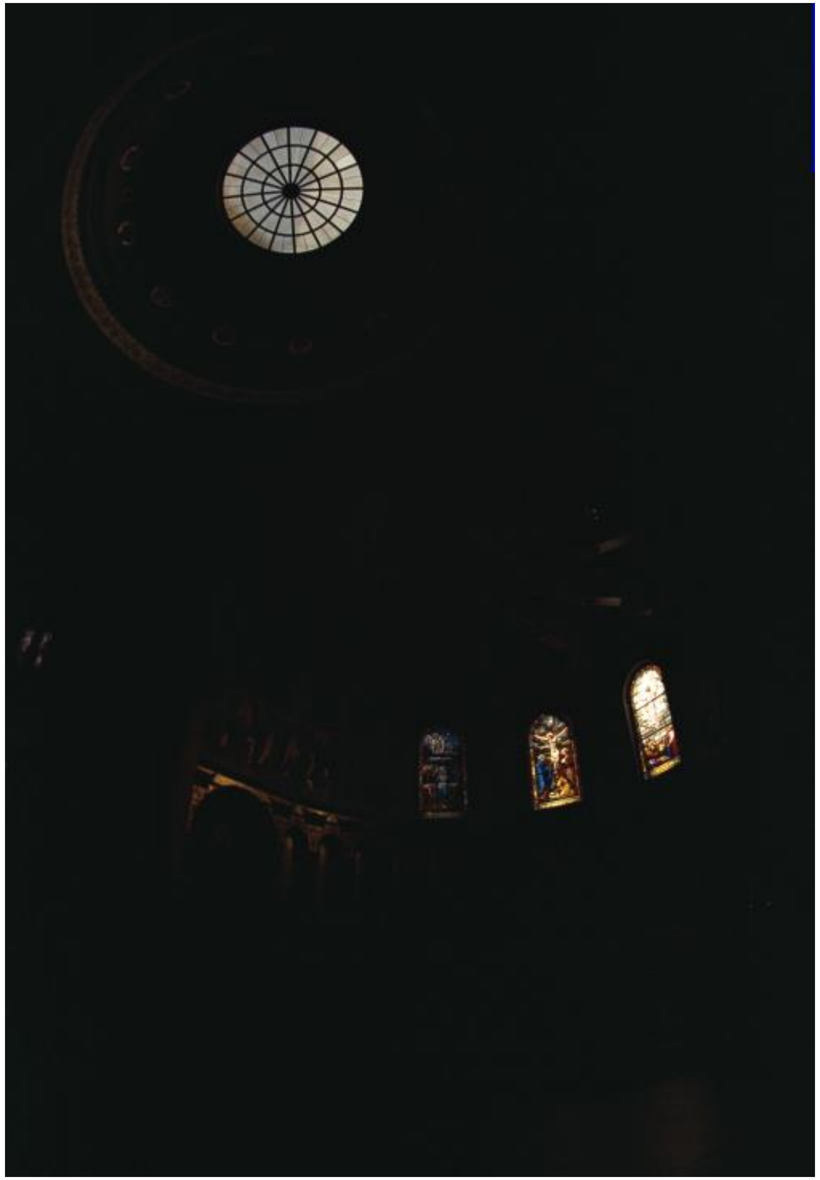


linear scaling
(map 10% to 1)

Compare with LDR images

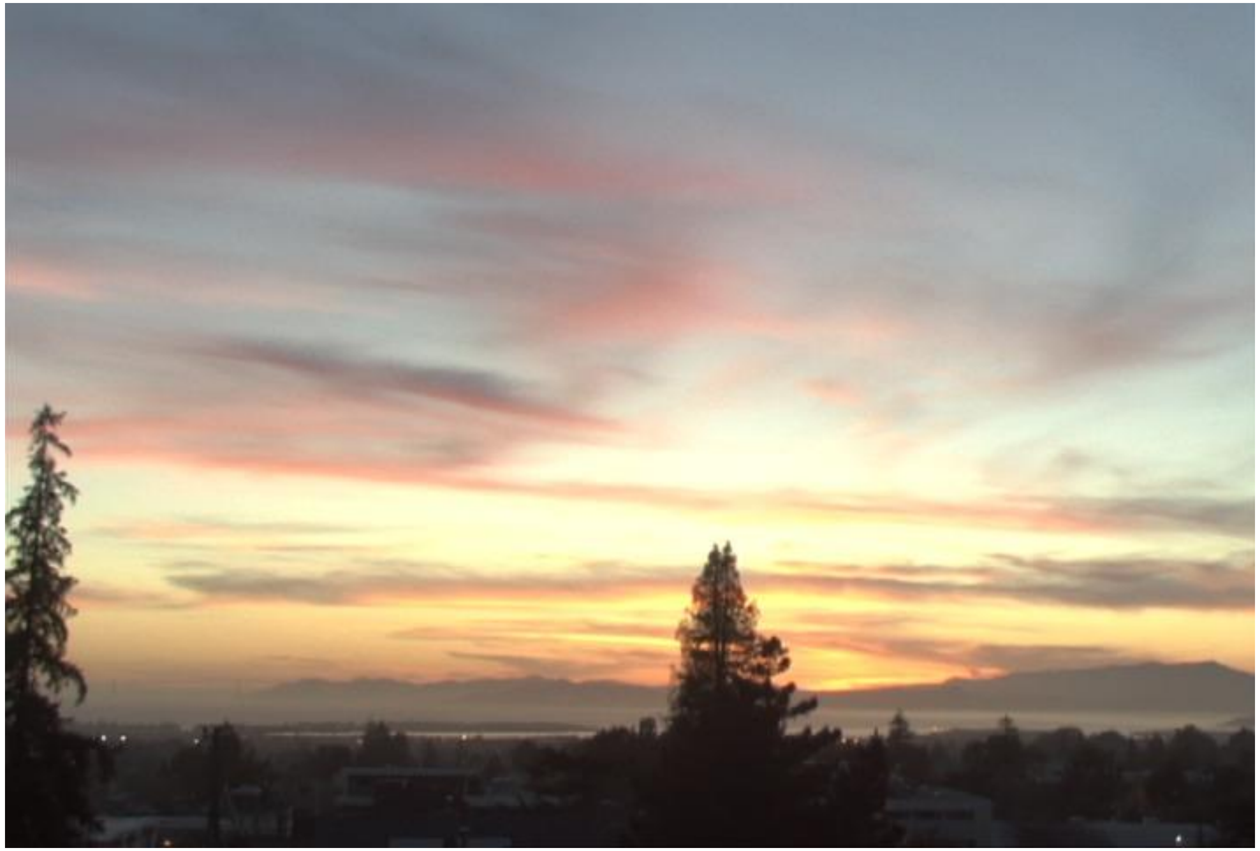






Dealing with color

If we tonemap all channels the same, colors are washed out



Can you think of a way to deal with this?

Intensity-only tonemapping

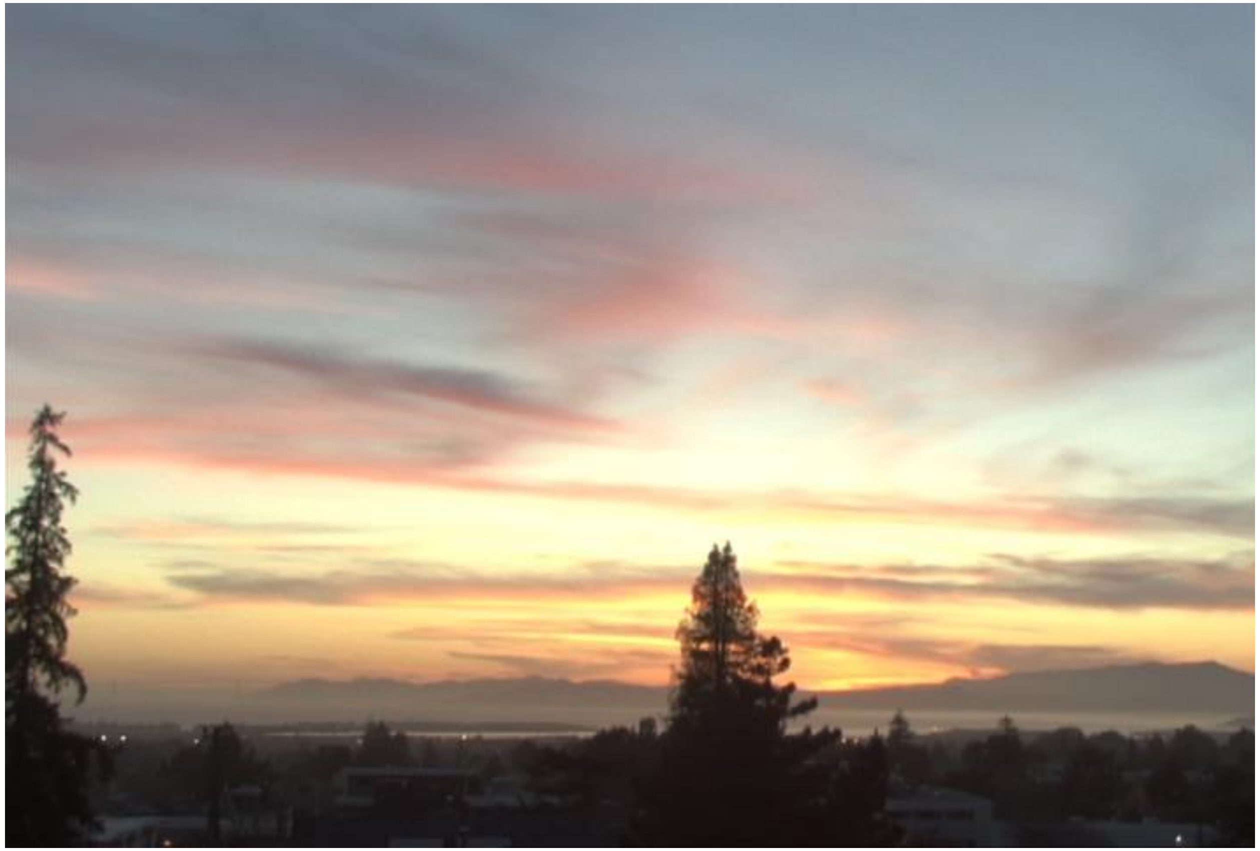
tonemap intensity

leave color the same

How would you implement this?

Comparison

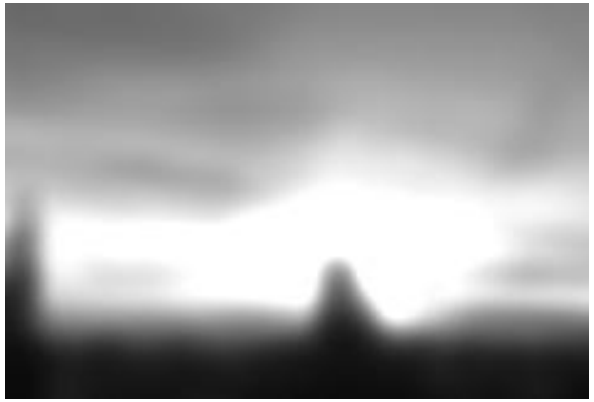
Color now OK, but some details are washed out due to loss of contrast



Can you think of a way to deal with this?

Low-frequency intensity-only tonemapping

tonemap low-frequency intensity component



leave high-frequency intensity component the same

leave color the same

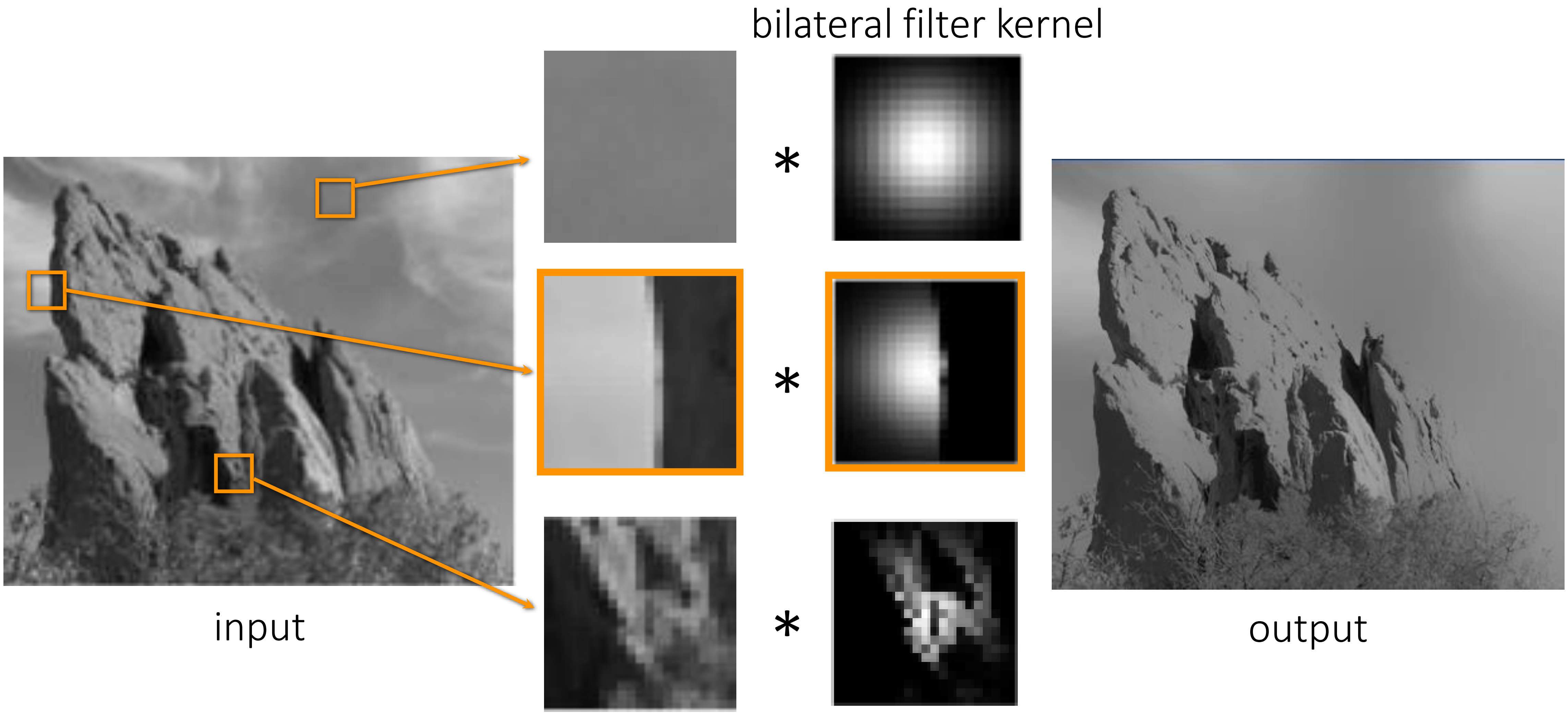
How would you implement this?

Comparison

We got nice color and contrast, but now we've run into the halo plague

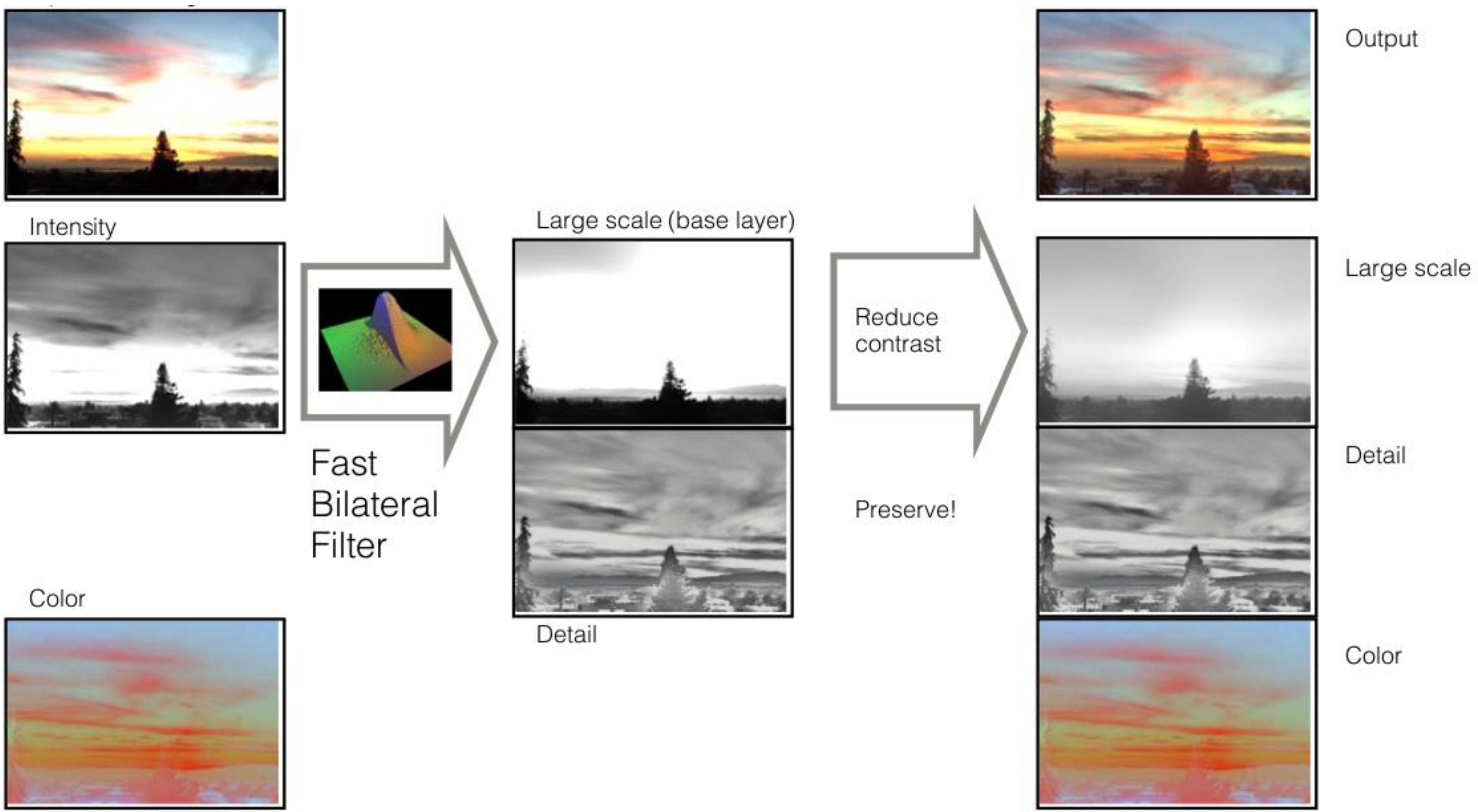
Can you think of a way to deal with this?

The bilateral filtering solution



Do not blur if there is an edge! How does it do that?

Tonemapping with bilateral filtering



Comparison

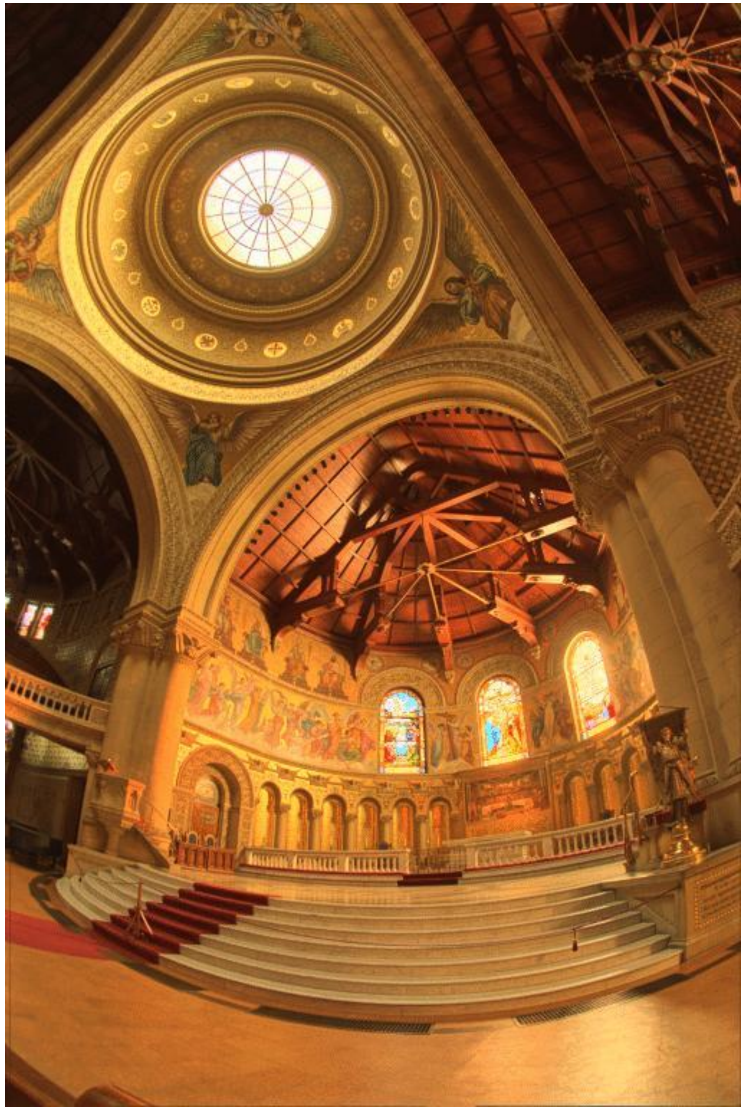
We fixed the halos without losing contrast

Gradient-domain merging and tonemapping

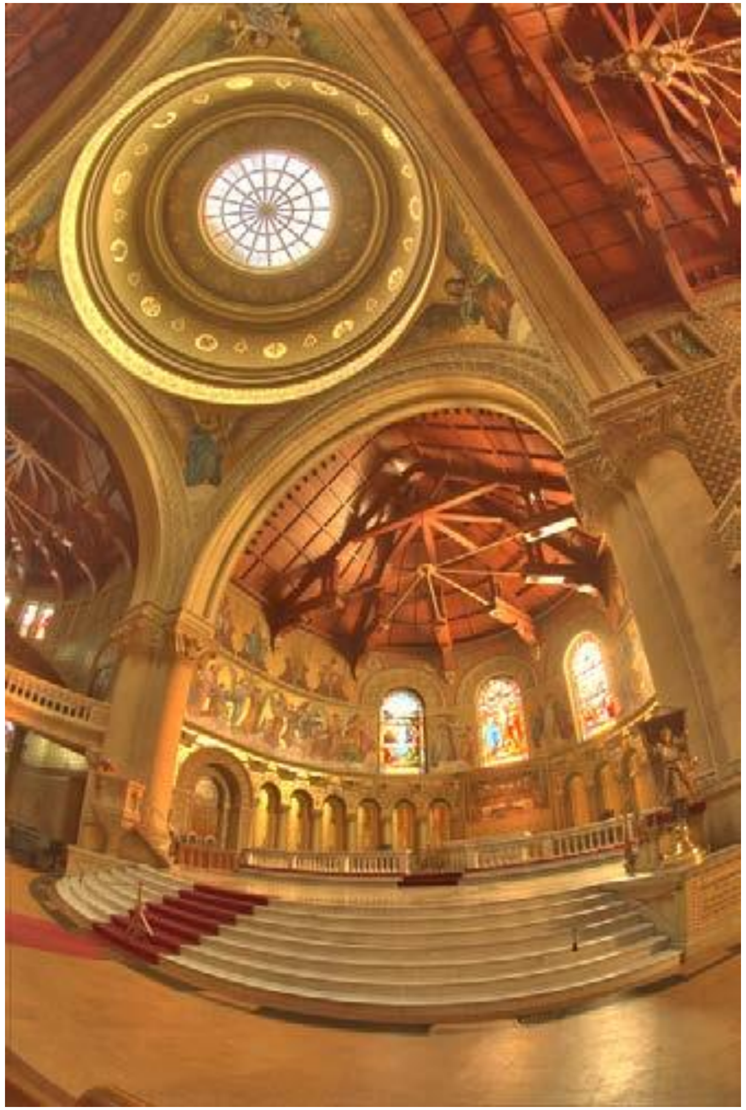
Compute gradients, scale and merge them, then integrate (solve Poisson problem)

Gradient-domain merging and tonemapping

Comparison (which one do you like better?)



photographic



bilateral filtering

gradient-domain

Comparison (which one do you like better?)

bilateral filtering

gradient-domain

Comparison (which one do you like better?)

There is no ground-truth

- which one looks better is entirely subjective

bilateral filtering

gradient-domain

Tonemapping for a single image

Modern DSLR sensors capture about 3 stops of dynamic range

- tonemap single RAW file instead of using camera's default rendering

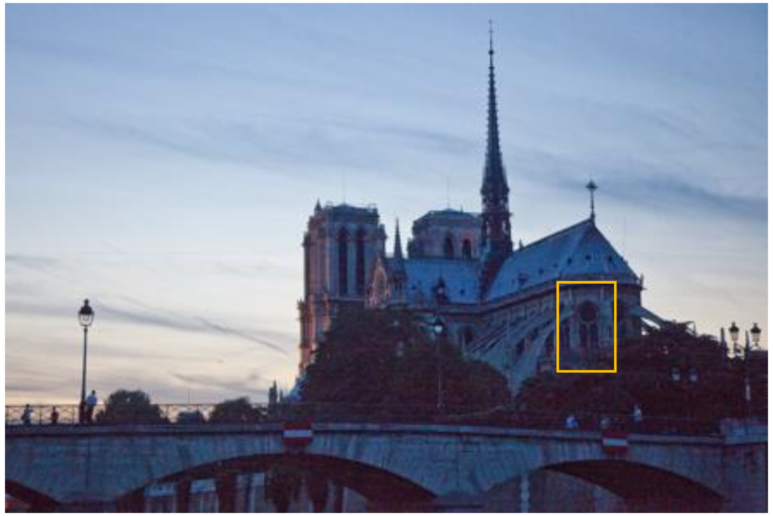
result from image processing pipeline (basic tone reproduction)

tonemapping using bilateral filtering (I think)

Tonemapping for a single image

Modern DSLR sensors capture about 3 stops of dynamic range

- tonemap single RAW file instead of using camera's default rendering



Careful not to "tonemap" noise

- why is this not a problem with multi-exposure HDR?

Some notes about HDR and tonemapping

A note of caution

- HDR photography can produce very visually compelling results

A note of caution

- HDR photography can produce very visually compelling results

- It is also a very routinely abused technique, resulting in awful results

A note of caution

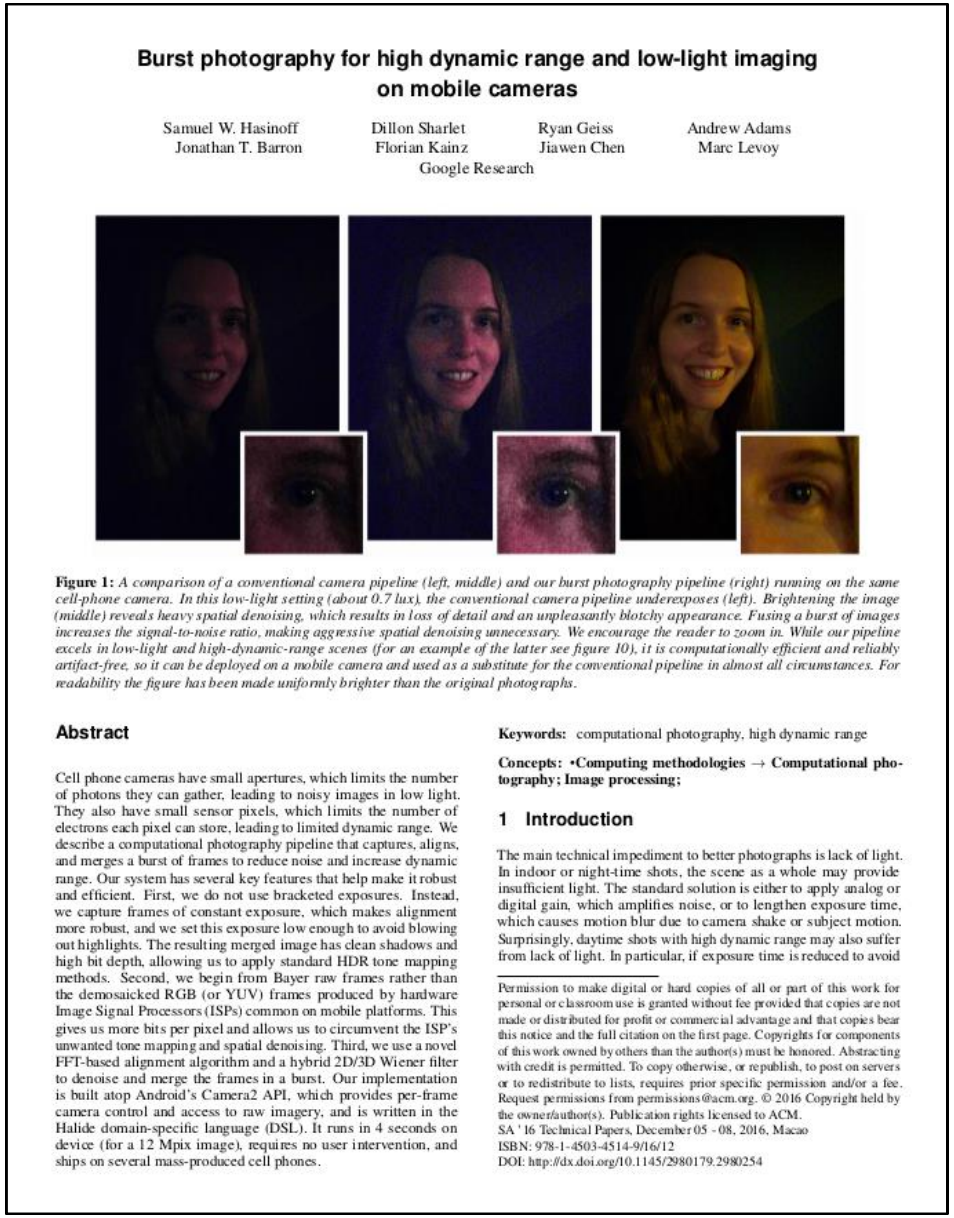
- HDR photography can produce very visually compelling results
- It is also a very routinely abused technique, resulting in awful results
- The problem is tonemapping, not HDR itself

A note about HDR today

- Most cameras (even phone cameras) have automatic HDR modes/apps

- Popular-enough feature that phone manufacturers are actively competing about which one has the best HDR

- The technology behind some of those apps (e.g., Google's HDR+) is published in SIGGRAPH and SIGGRAPH Asia conferences



Burst photography for high dynamic range and low-light imaging on mobile cameras

Samuel W. Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams
Jonathan T. Barron, Florian Kainz, Jiawen Chen, Marc Levoy
Google Research

Figure 1: *A comparison of a conventional camera pipeline (left, middle) and our burst photography pipeline (right) running on the same cell-phone camera. In this low-light setting (about 0.7 lux), the conventional camera pipeline underexposes (left). Brightening the image (middle) reveals heavy spatial denoising, which results in loss of detail and an unpleasantly blotchy appearance. Fusing a burst of images increases the signal-to-noise ratio, making aggressive spatial denoising unnecessary. We encourage the reader to zoom in. While our pipeline excels in low-light and high-dynamic-range scenes (for an example of the latter see figure 10), it is computationally efficient and reliably artifact-free, so it can be deployed on a mobile camera and used as a substitute for the conventional pipeline in almost all circumstances. For readability the figure has been made uniformly brighter than the original photographs.*

Abstract

Cell phone cameras have small apertures, which limits the number of photons they can gather, leading to noisy images in low light. They also have small sensor pixels, which limits the number of electrons each pixel can store, leading to limited dynamic range. We describe a computational photography pipeline that captures, aligns, and merges a burst of frames to reduce noise and increase dynamic range. Our system has several key features that help make it robust and efficient. First, we do not use bracketed exposures. Instead, we capture frames of constant exposure, which makes alignment more robust, and we set this exposure low enough to avoid blowing out highlights. The resulting merged image has clean shadows and high bit depth, allowing us to apply standard HDR tone mapping methods. Second, we begin from Bayer raw frames rather than the demosaicked RGB (or YUV) frames produced by hardware Image Signal Processors (ISPs) common on mobile platforms. This gives us more bits per pixel and allows us to circumvent the ISP's unwanted tone mapping and spatial denoising. Third, we use a novel FFT-based alignment algorithm and a hybrid 2D/3D Wiener filter to denoise and merge the frames in a burst. Our implementation is built atop Android's Camera2 API, which provides per-frame camera control and access to raw imagery, and is written in the Halide domain-specific language (DSL). It runs in 4 seconds on device (for a 12 Mpix image), requires no user intervention, and ships on several mass-produced cell phones.

Keywords: computational photography, high dynamic range

Concepts: **•Computing methodologies → Computational photography; Image processing;**

1 Introduction

The main technical impediment to better photographs is lack of light. In indoor or night-time shots, the scene as a whole may provide insufficient light. The standard solution is either to apply analog or digital gain, which amplifies noise, or to lengthen exposure time, which causes motion blur due to camera shake or subject motion. Surprisingly, daytime shots with high dynamic range may also suffer from lack of light. In particular, if exposure time is reduced to avoid

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. © 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM.
SA '16 Technical Papers, December 05 - 08, 2016, Macao
ISBN: 978-1-4503-4514-9/16/12
DOI: http://dx.doi.org/10.1145/2980179.2980254

References

Basic reading:

- Szeliski textbook, Sections 10.1, 10.2.
- Debevec and Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," SIGGRAPH 1997.
  the paper that more or less started HDR imaging research in computer graphics.
- Reinhard et al., "Photographic Tone Reproduction for Digital Images," SIGGRAPH 2002.
  the photographic tonemapping paper, including a very nice discussion of the zone system for film.
- Durand and Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," SIGGRAPH 2002.
  the paper on tonemapping using bilateral filtering.
- Fattal et al., "Gradient Domain High Dynamic Range Compression," SIGGRAPH 2002.
  the paper on gradient-domain tonemapping.

Additional reading:

- Reinhard et al., "High Dynamic Range Imaging, Second Edition: Acquisition, Display, and Image-Based Lighting," Morgan Kaufmann 2010
  a very comprehensive book about everything relating to HDR imaging and tonemapping.
- Kuang et al., "Evaluating HDR rendering algorithms," TAP 2007.
  one of many, many papers trying to do a perceptual evaluation of different tonemapping algorithms.
- Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography," SIGGRAPH 1998.
  the original HDR light probe paper (we'll see more about this in a later lecture).
- Hasinoff and Kutulakos, "Multiple-Aperture Photography for High Dynamic Range and Post-Capture Refocusing," UofT TR 2009
  a paper on doing HDR by aperture bracketing instead of exposure bracketing.
- Hasinoff et al., "Noise-Optimal Capture for High Dynamic Range Photography," CVPR 2010
  a paper on weighting different exposures based on a very detailed sensor noise model.
- Hasinoff et al., "Burst photography for high dynamic range and low-light imaging on mobile cameras," SIGGRAPH Asia 2016
  the paper describing Google's HDR+.
- Ward, "The radiance lighting simulation and rendering system," SIGGRAPH 1994
  the paper that introduced (among other things) the .hdr image format for HDR images.