HDR Imaging

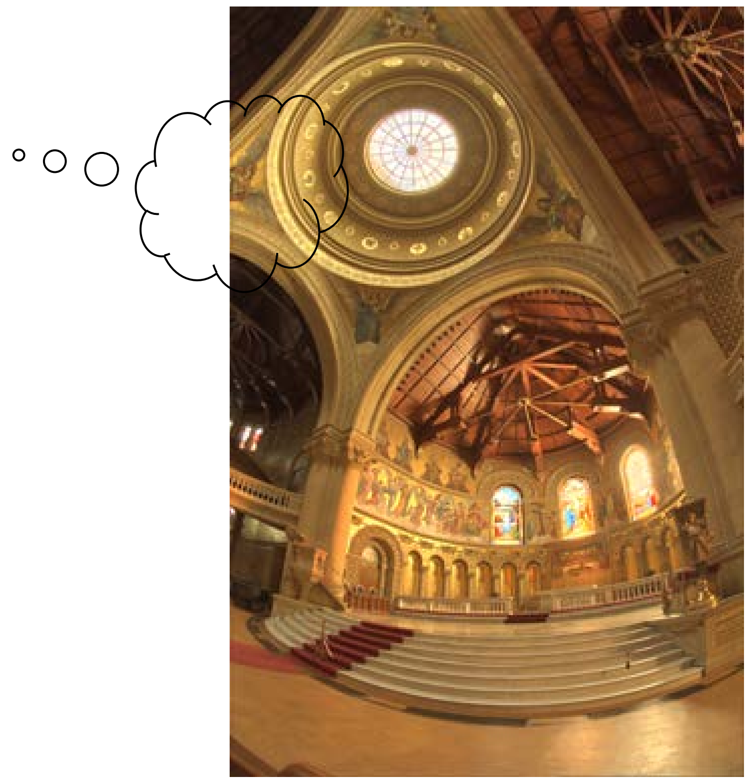
CVFX 2015

Today's Plan

- *Recovering High Dynamic Range Radiance Maps from Photographs*
  - Debevec and Malik
  - SIGGRAPH 1997
- *What Is the Space of Camera Response Functions?*
  - Grossberg and Nayar
  - CVPR 2003

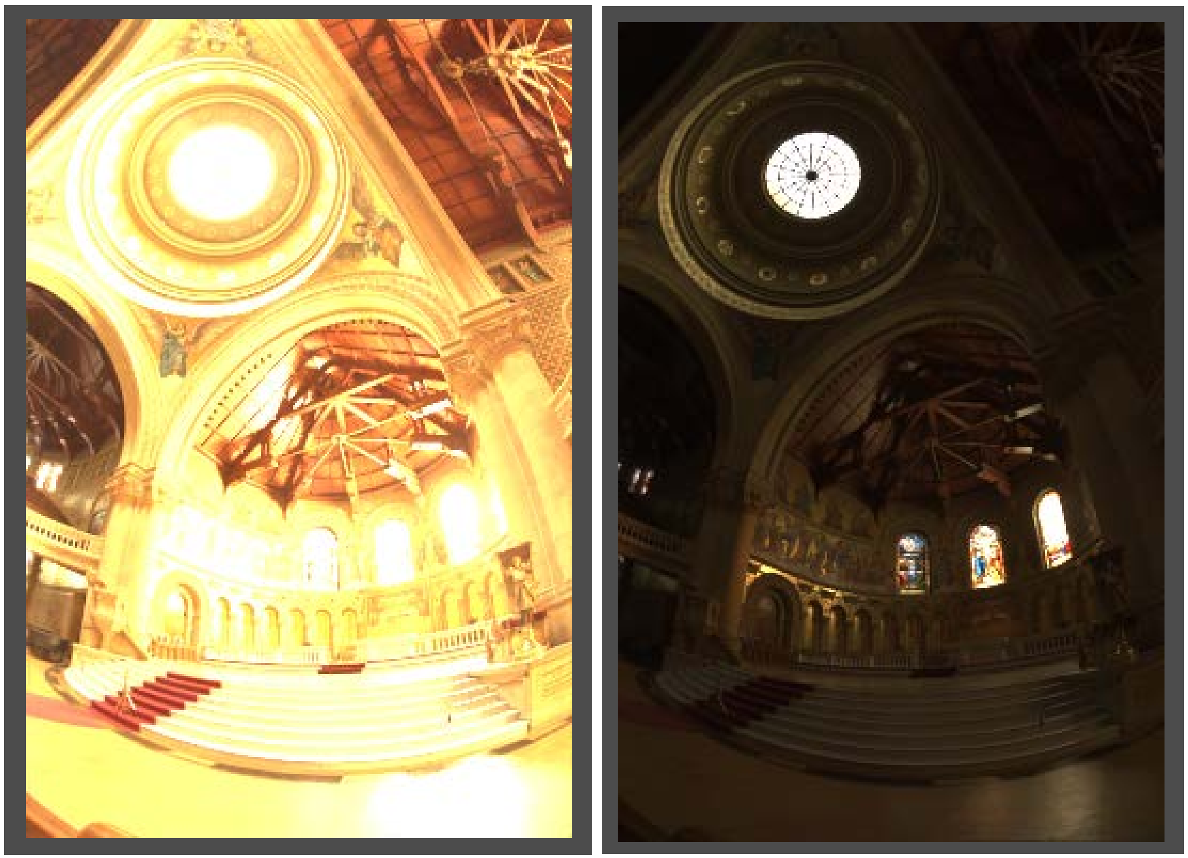
In the Old Days...

- You took some pictures on your trip



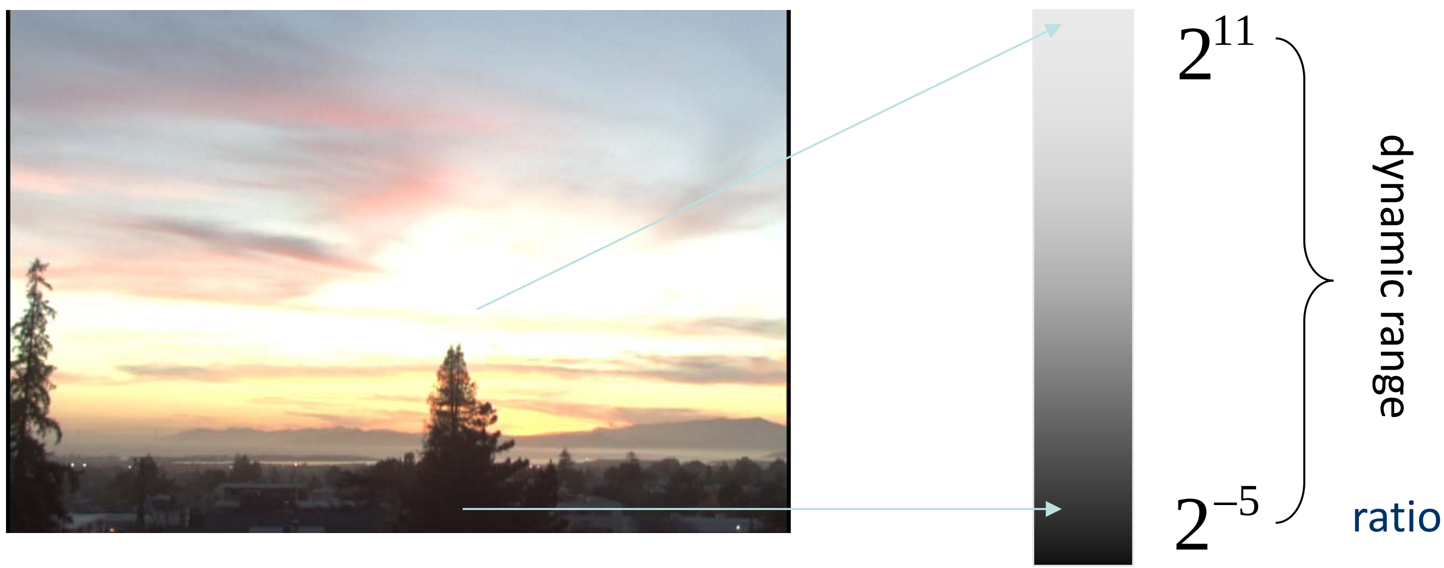
They Might Turn out like …

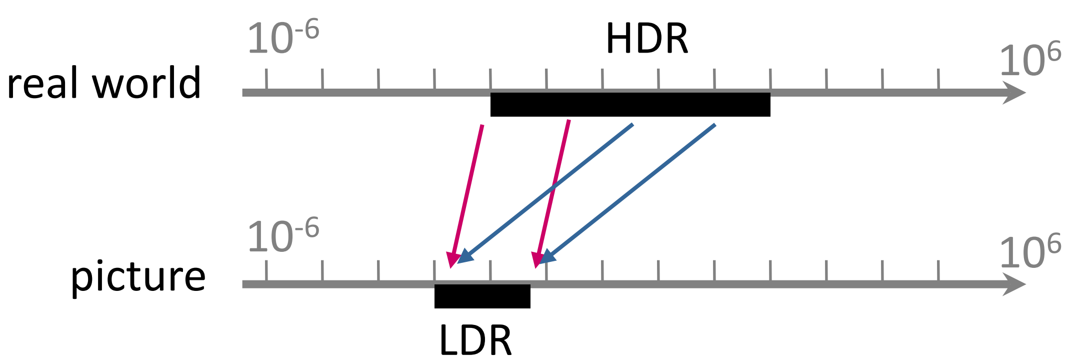
- What's wrong?

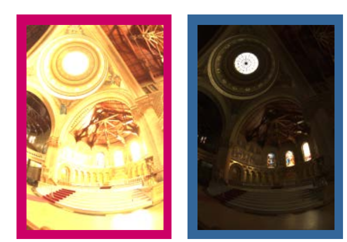


Dynamic Range

- Dynamic range: *contrast in the scene*
- The real world is of high dynamic range (HDR)

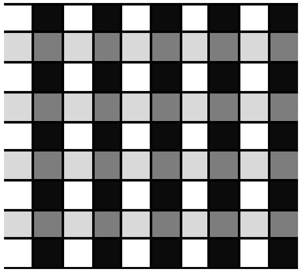






Two Issues

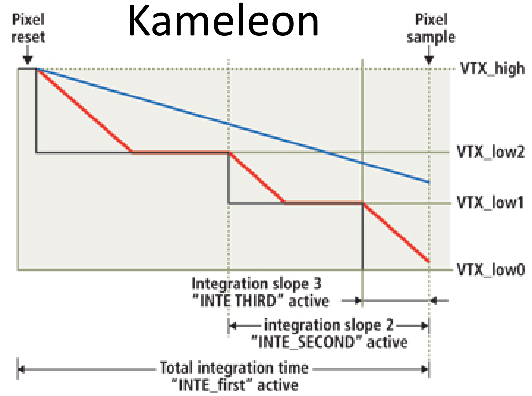
- How to capture
  - Debevec and Malik …
    - » Radiance map in RGBE format
  - Nayar and Mitsunaga
  - Aggarwal and Ahuja

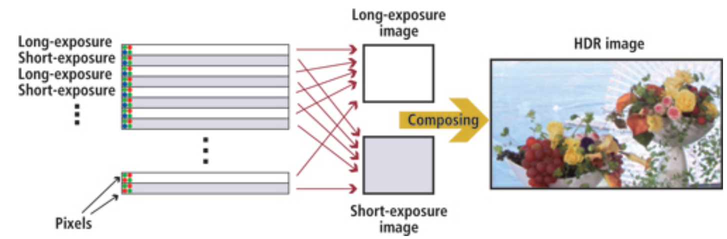


- How to display
  - LCD contrast – 3000:1
  - Photo paper – 100:1

How to Capture

Kameleon

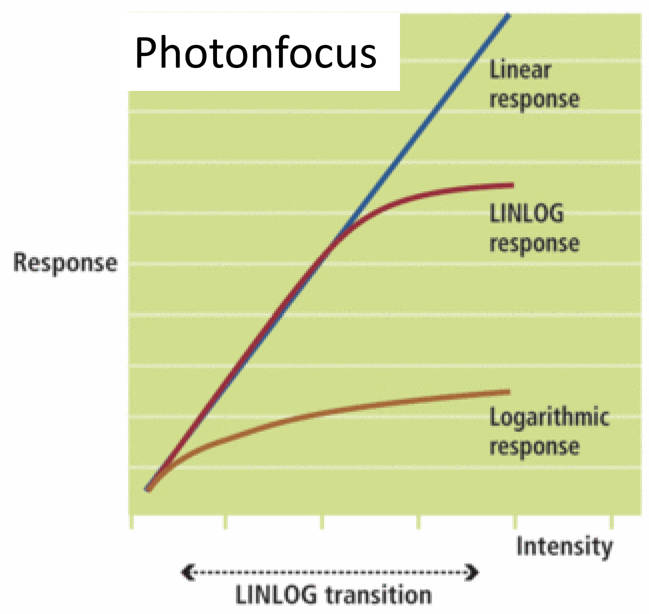


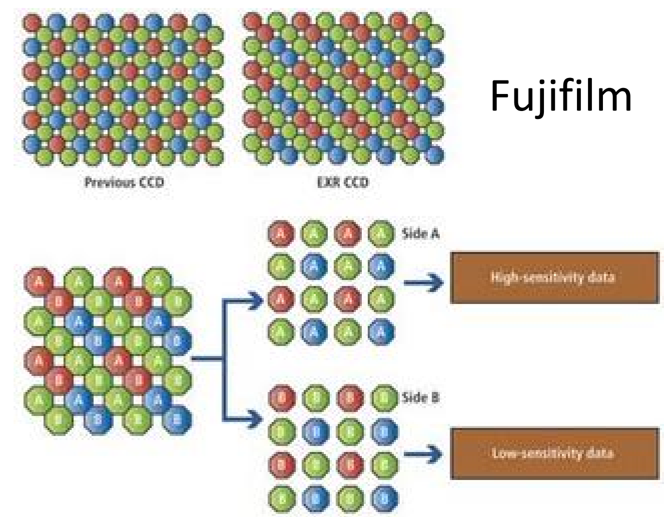


Toshiba

exposure time

Photonfocus





Fujifilm

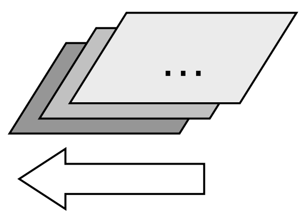
spatially varying

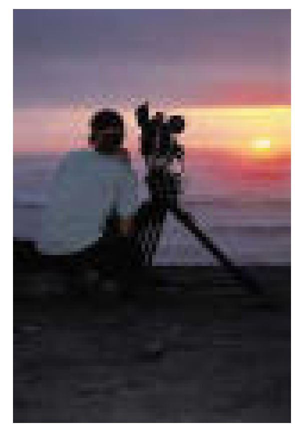
images from vision-systems.com

Radiance Map

- Combine multiple images taken under different exposure settings
  - Debevec and Malik

32-bit RGBE

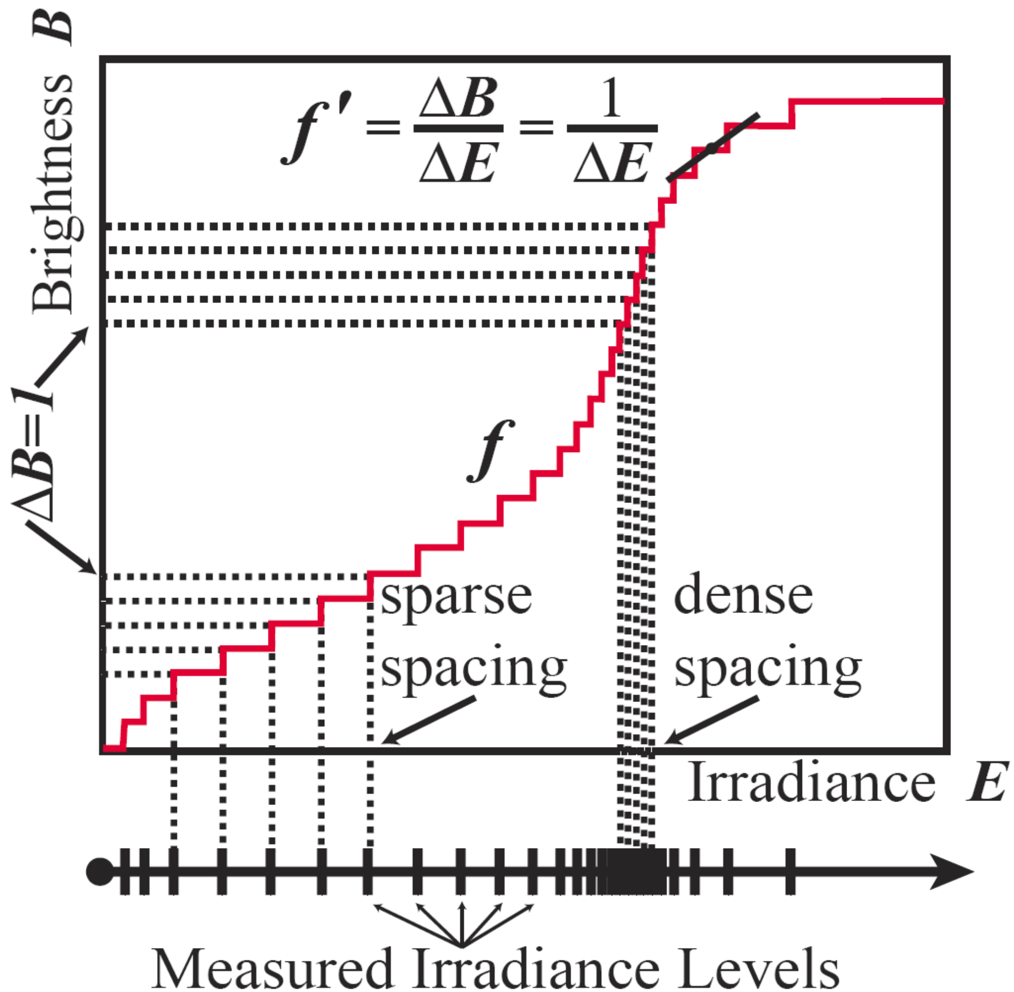




R,G,B + shared Exponent

Matlab hdrread, hdrwrite, makehdr

Recovering the Response Function

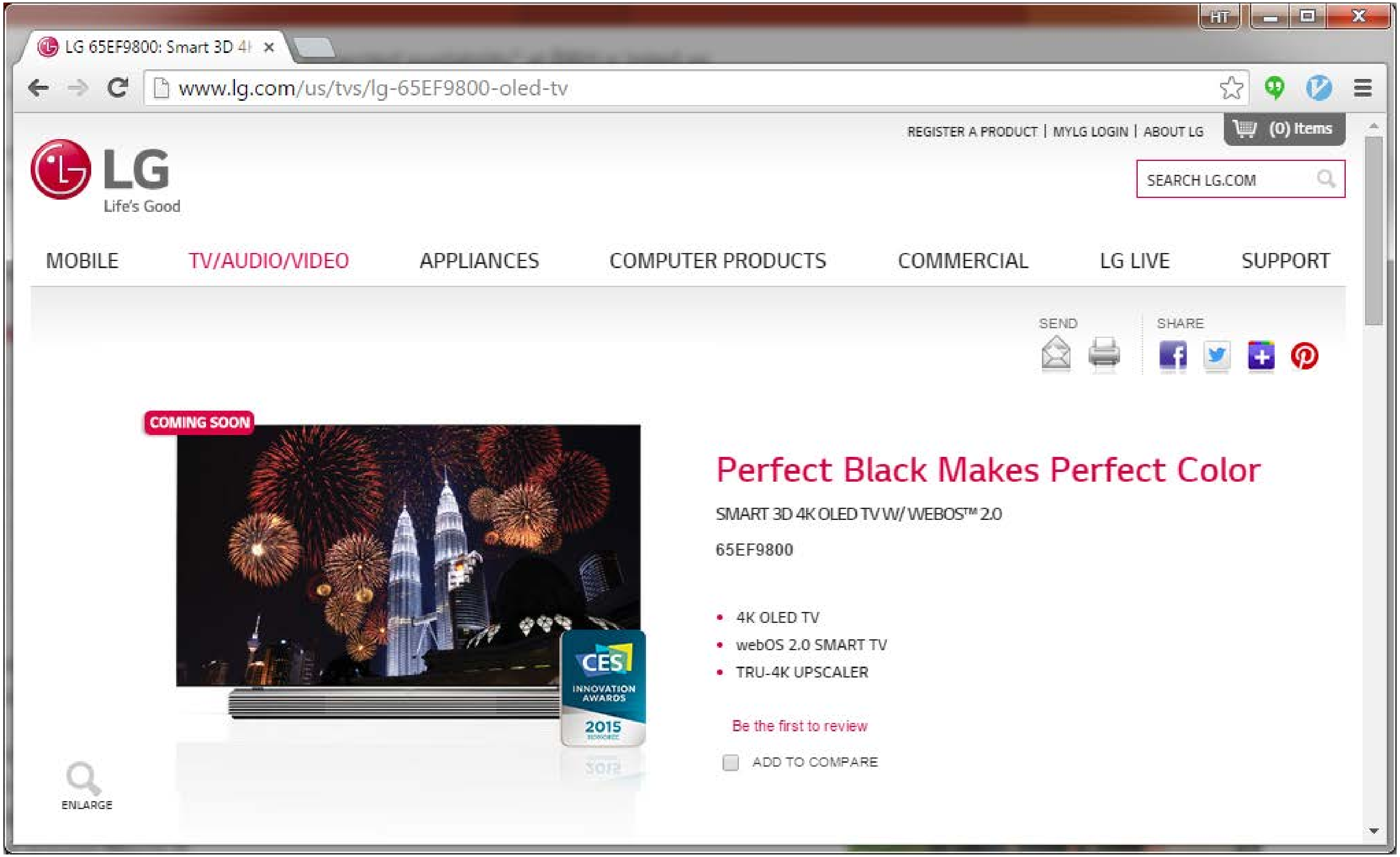


How to Display a Radiance Map

- Hardware

- Software

High Dynamic Range Display



$9999

image from LG's website

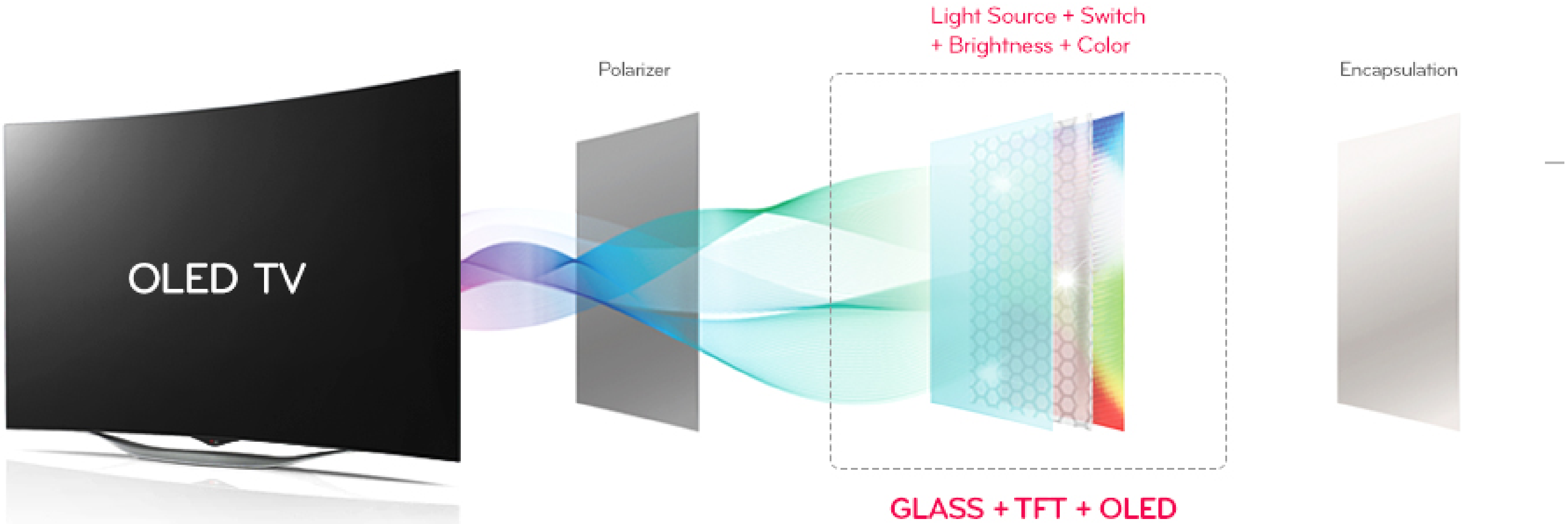
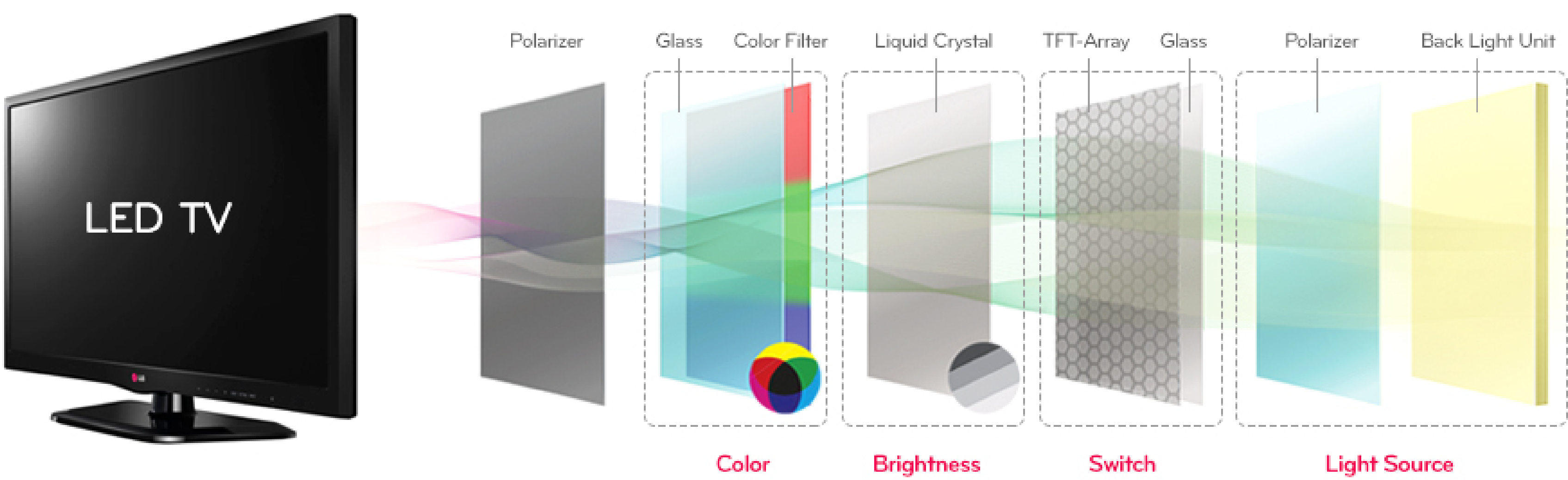
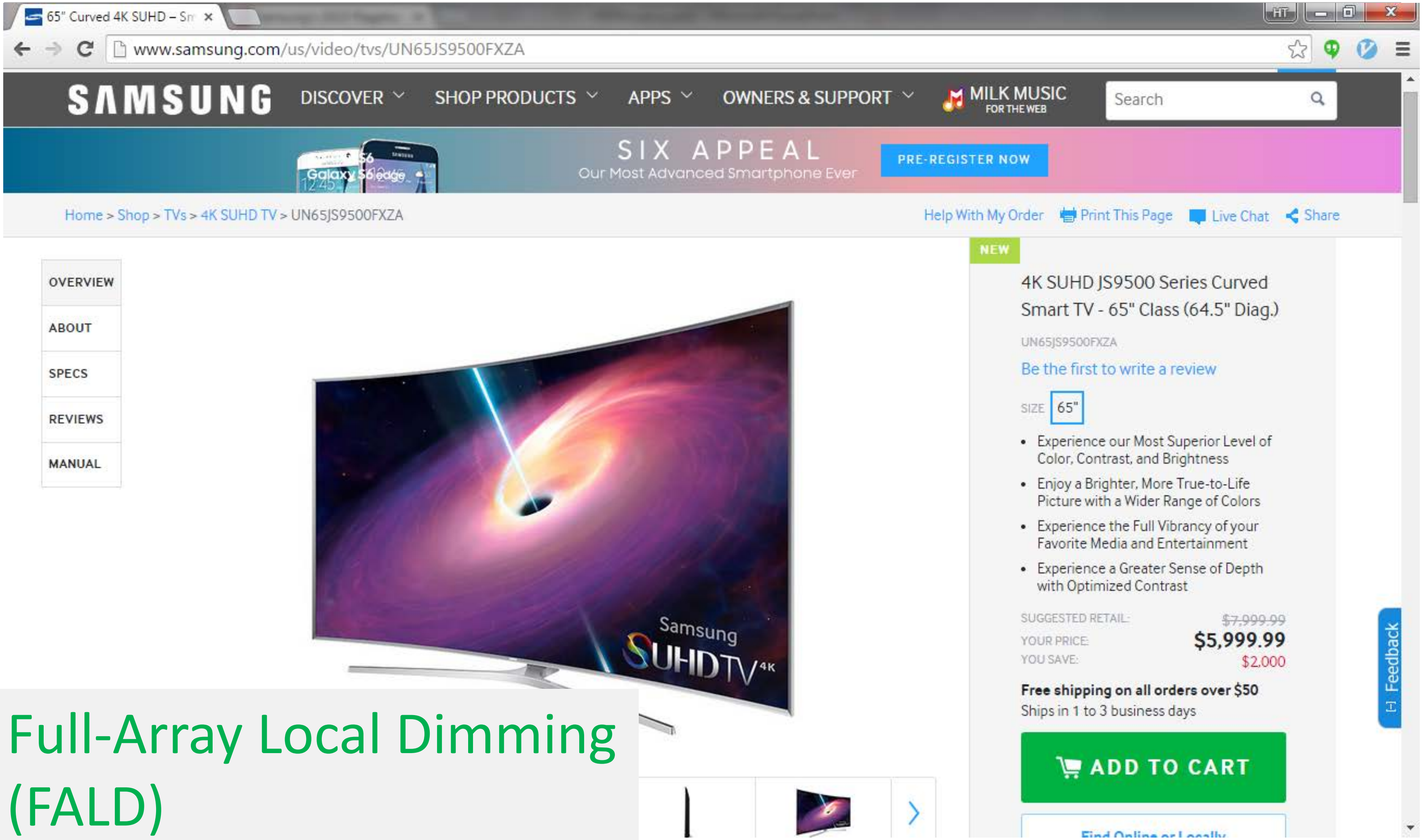


image from LG's website



High Dynamic Range Display



Full-Array Local Dimming (FALD)

image from Samsung's website

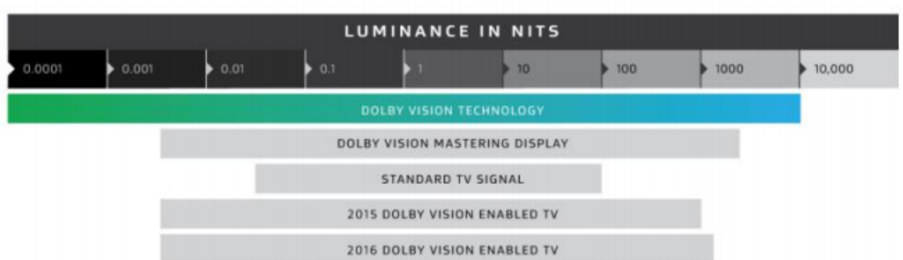
High Dynamic Range Solution

www.dolby.com/us/en/technologies/dolby-vision/dolby-vision-white-paper.pdf

DOLBY VISION™ WHITE

HOW MUCH DYNAMIC RANGE DO WE NEED?

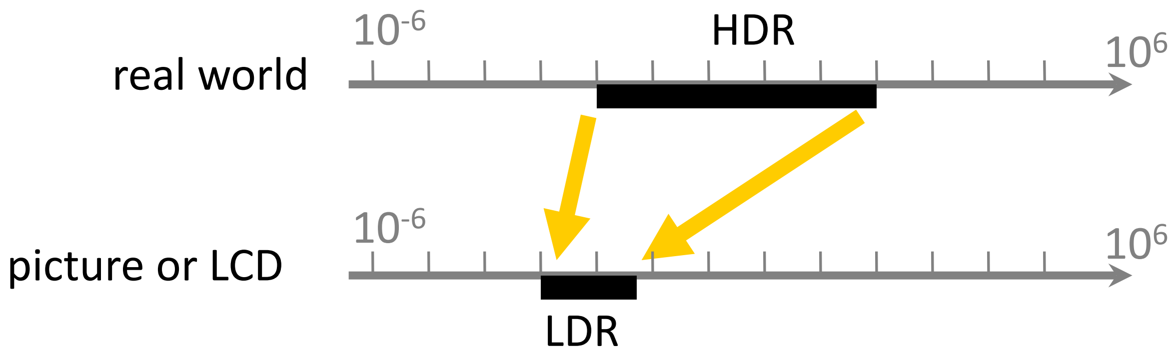
The Dolby image research team ran a set of experiments with ordinary viewers to answer this question. The researchers tested what viewers preferred for black level, diffuse white level, and highlight level. They determined that a system that could reproduce a range of 0 to 10,000 nits satisfied 90 percent of viewers asked to pick an ideal range.



How to Display a Radiance Map

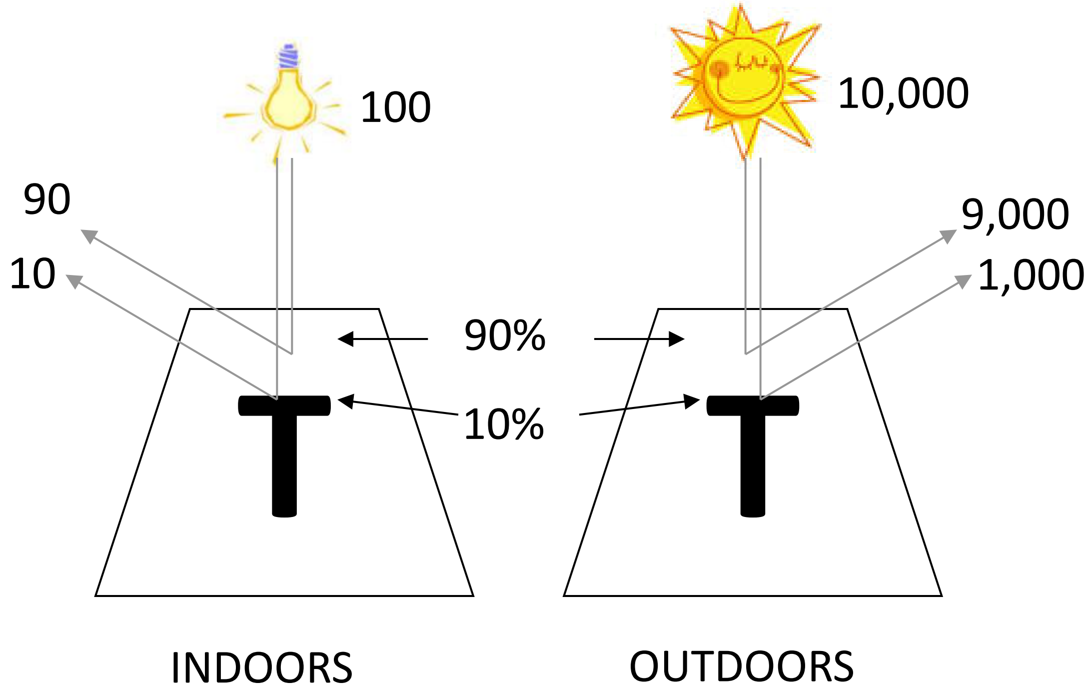
- Human eyes perceive higher dynamic ranges than those reproduced on LCD or photo paper

- Tone reproduction problem
  - How do we map perceived scene luminance to display luminance and produce a satisfactory image?



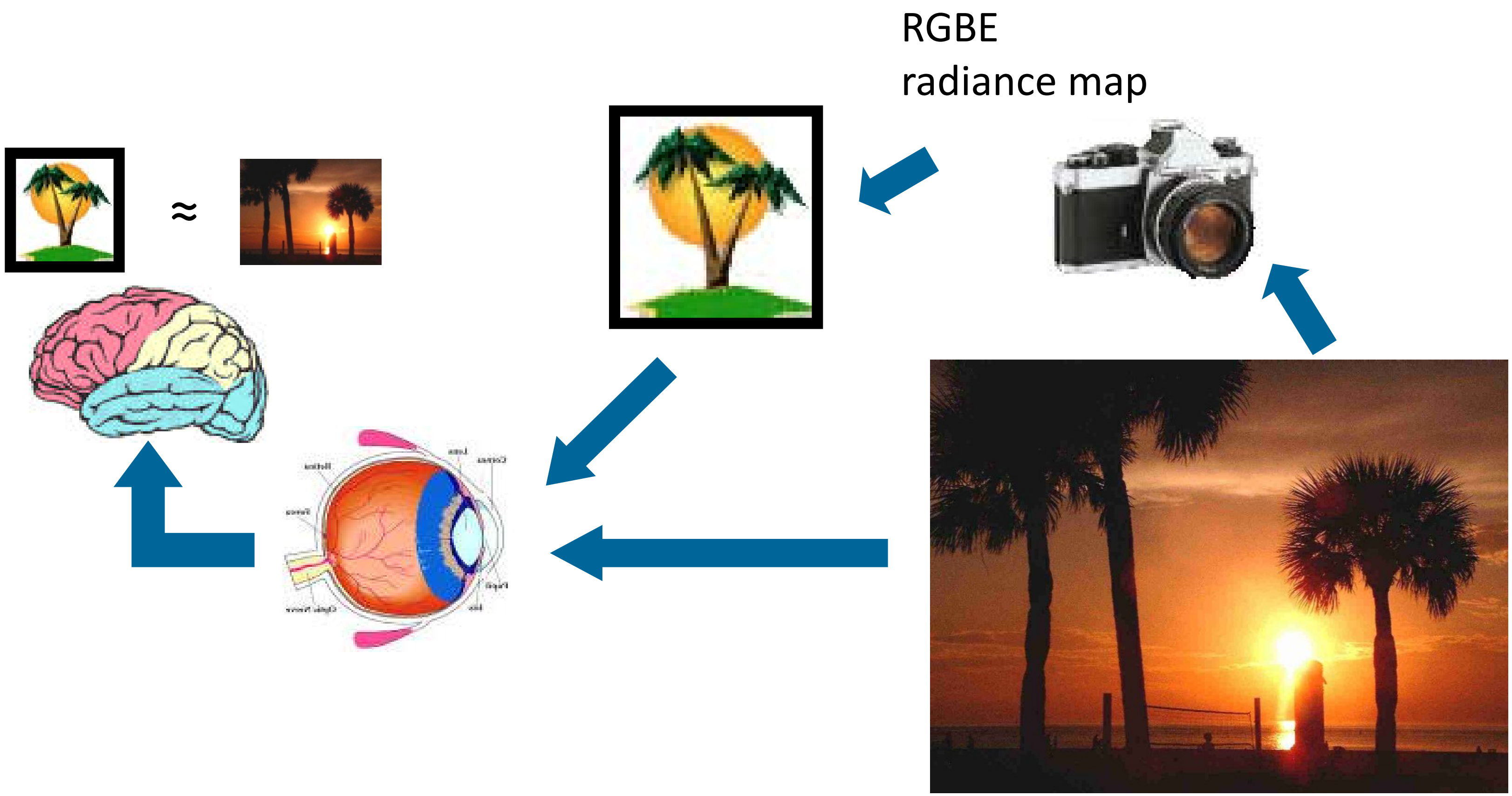
Lightness Constancy

- Illumination and reflectance



Tone Reproduction

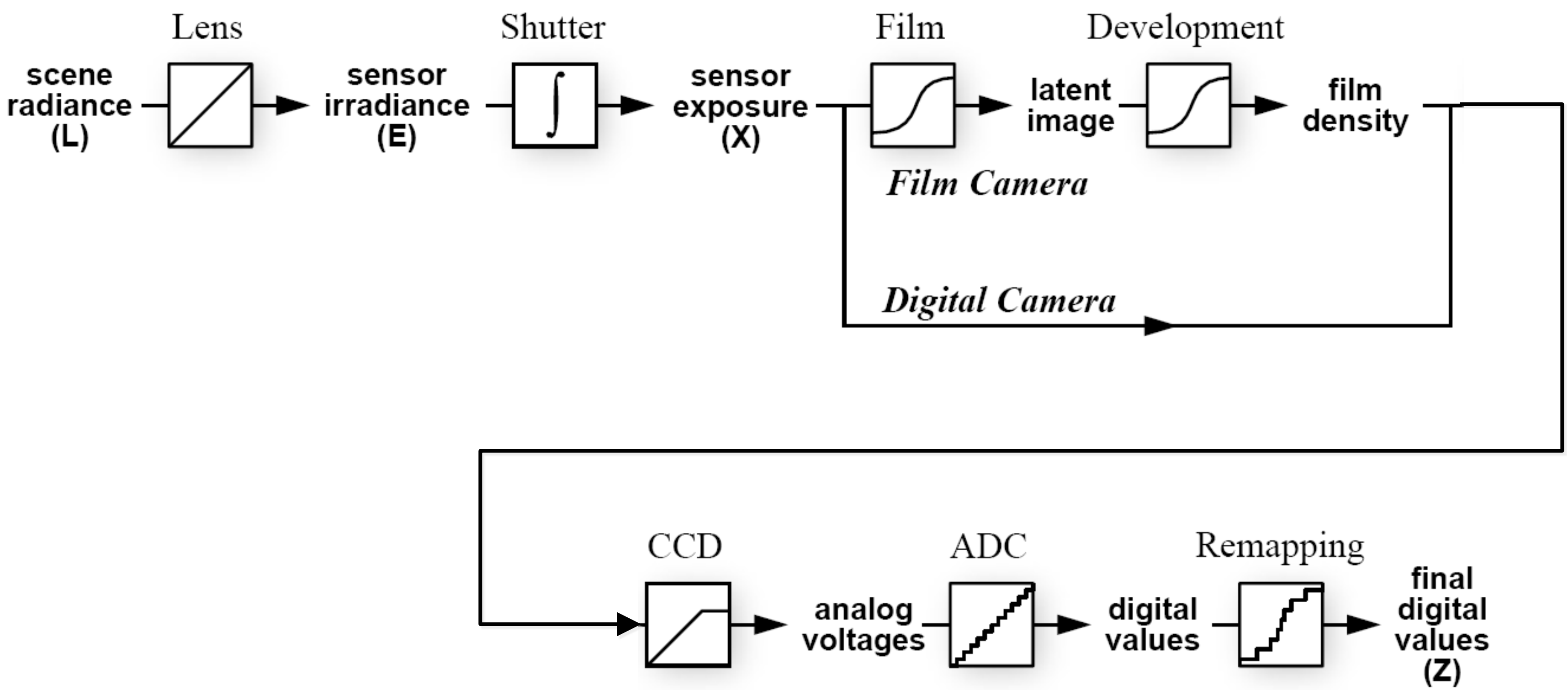
- How to reproduce visual impression



Modeling the Photographic Process

- When we photograph a scene, the acquired brightness values are rarely true measurements of relative radiance in the scene
  - E.g., if one pixel has twice the value of another, it is unlikely that It observed twice the radiance

Image Acquisition Pipeline

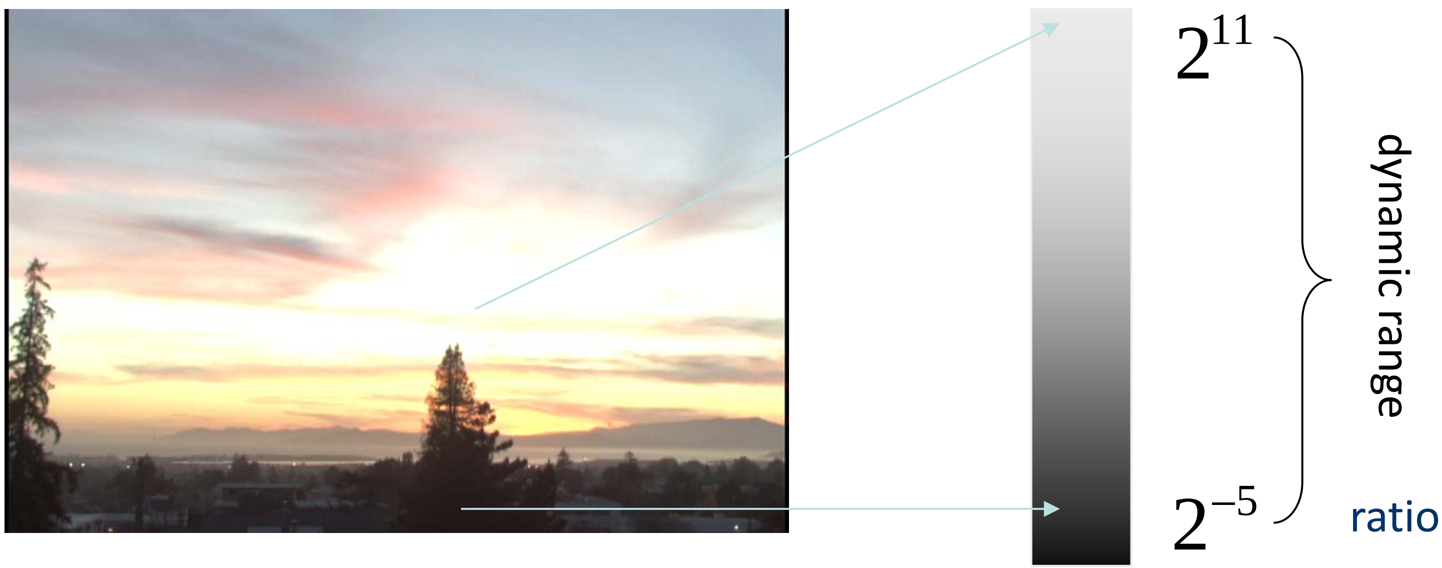


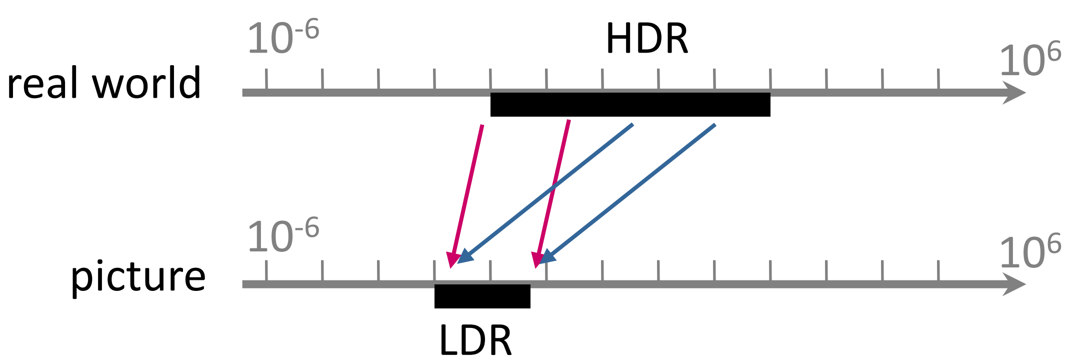
nonlinear response curve

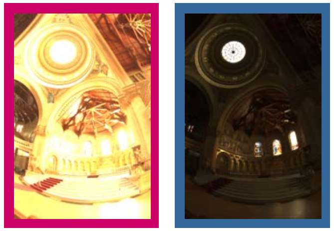
Why Is This a Problem at All?

Limitation on the Dynamic Range

- Dynamic range: *contrast in the scene*
- The real world is of high dynamic range (HDR)

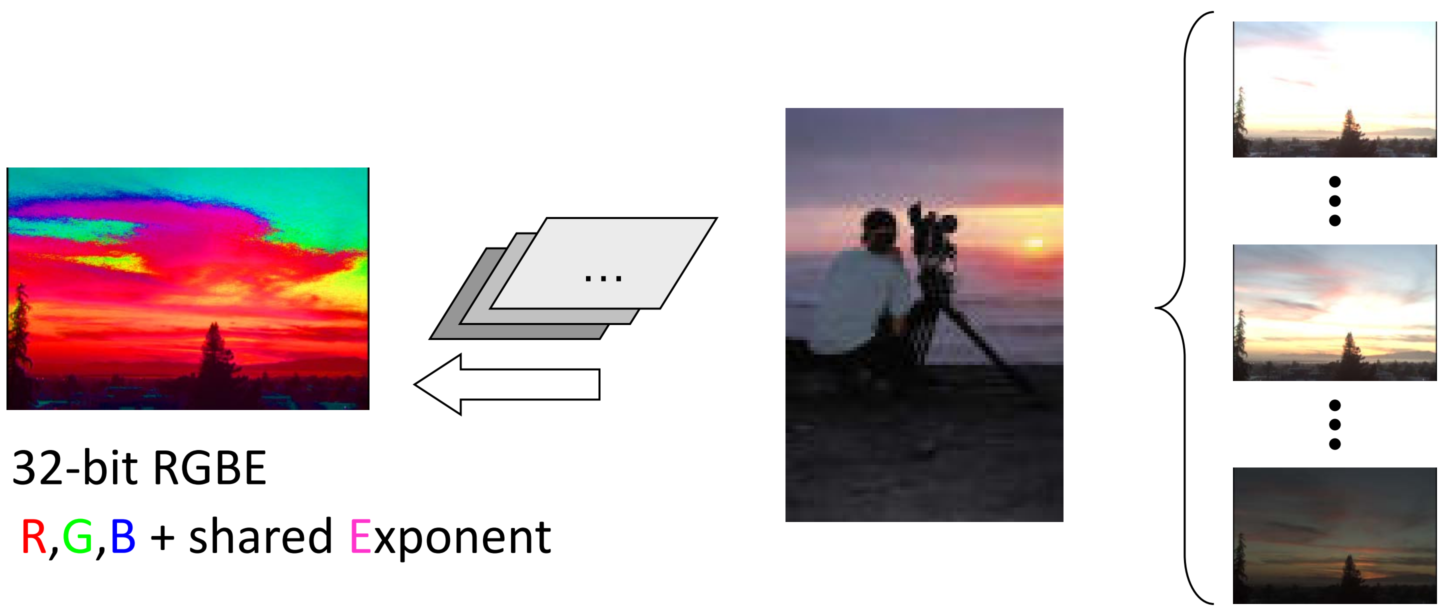






Limitation on the Dynamic Range

- One has to choose the range of radiance values that are of interest and determine the exposure time suitably
- To cover the full dynamic range in a scene, we can take a series of photographs with different exposures



How to Combine Different Exposures?

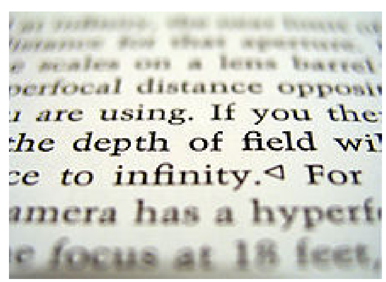
- We need to recover the *response function* so we can use it to estimate the radiance value at each pixel, given the intensity values of that pixel under different exposures

Exposure

- ND (neutral density) filters

- Shutter speed
  - Motion blur
  - Camera shake

- Aperture
  - Depth of field



[Wikipedia]

Varying the Shutter Speed

- The photographs with different exposures are required to be registered
- Assume the scene is stationary
  - There is no motion blur
- Use a tripod
  - Preventing camera shake

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

The Registration Problem

- How to align different exposures efficiently?

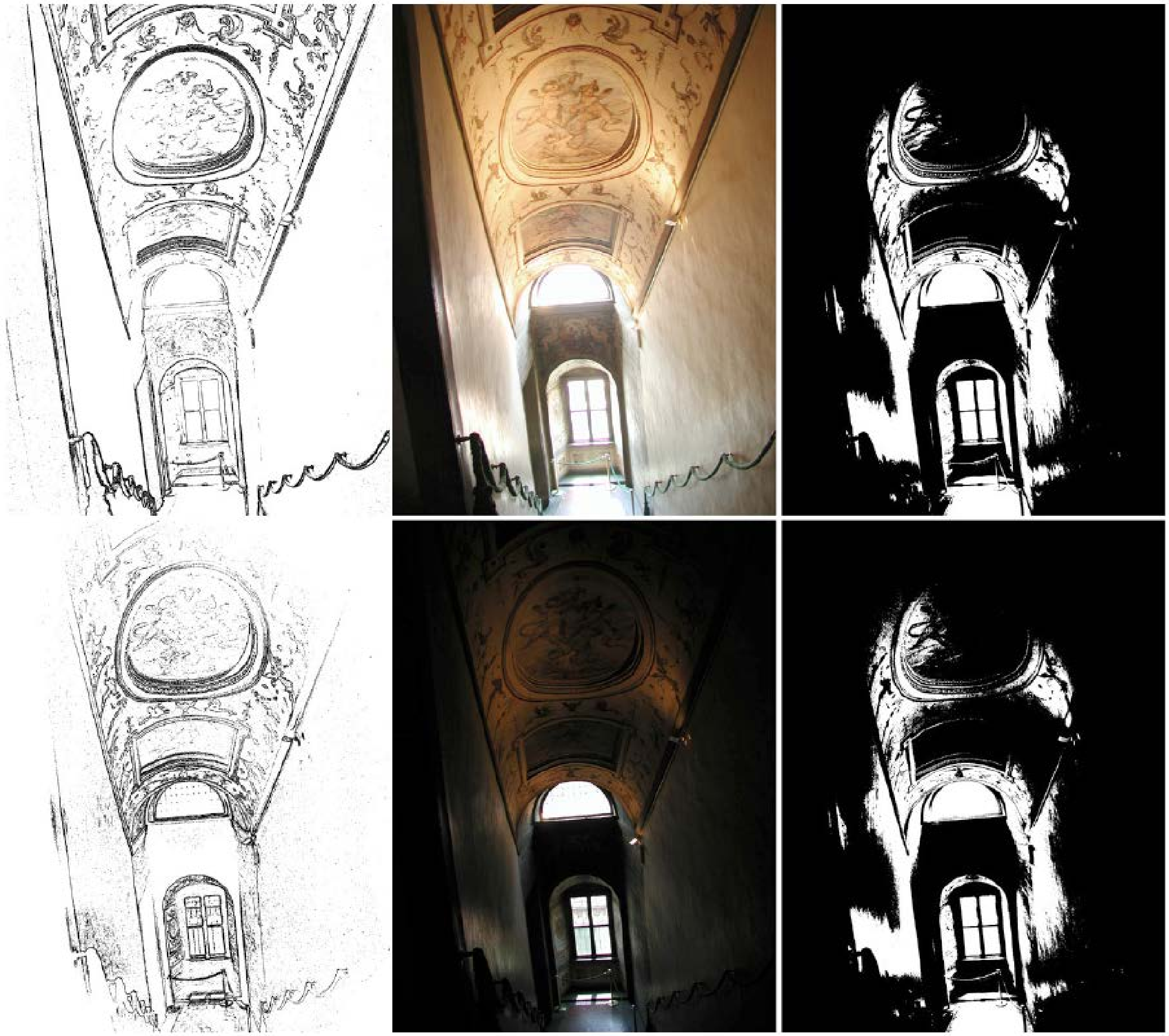
- *"Fast, Robust Image Registration for Compositing High Dynamic Range Photographs from Handheld Exposures"*
  - Greg Ward

- *"High Dynamic Range Image Reconstruction from Hand-held Cameras"*
  - Pei-Ying Lu, Tz-Huan Huang, Meng-Sung Wu, Yi-Ting Cheng and Yung-Yu Chuang
  - CVPR 2009

Registration for Compositing HDR Photographs

- Edge detectors are dependent on image exposure, therefore edge-matching or interest-point-matching algorithms are ill-suited to the exposure alignment problem
  - Taking the difference of two edge maps would not give a good indication of where the edges are misaligned.
- Median Threshold Bitmap
  - Consider only integer pixel offsets (no rotation, enough for 90% cases)
  - The input is a series of N grayscale images.
    - Use the green channel approximately or convert into grayscale by $Y=(54R+183G+19B)/256$

edge maps | unaligned exposures | MTBs

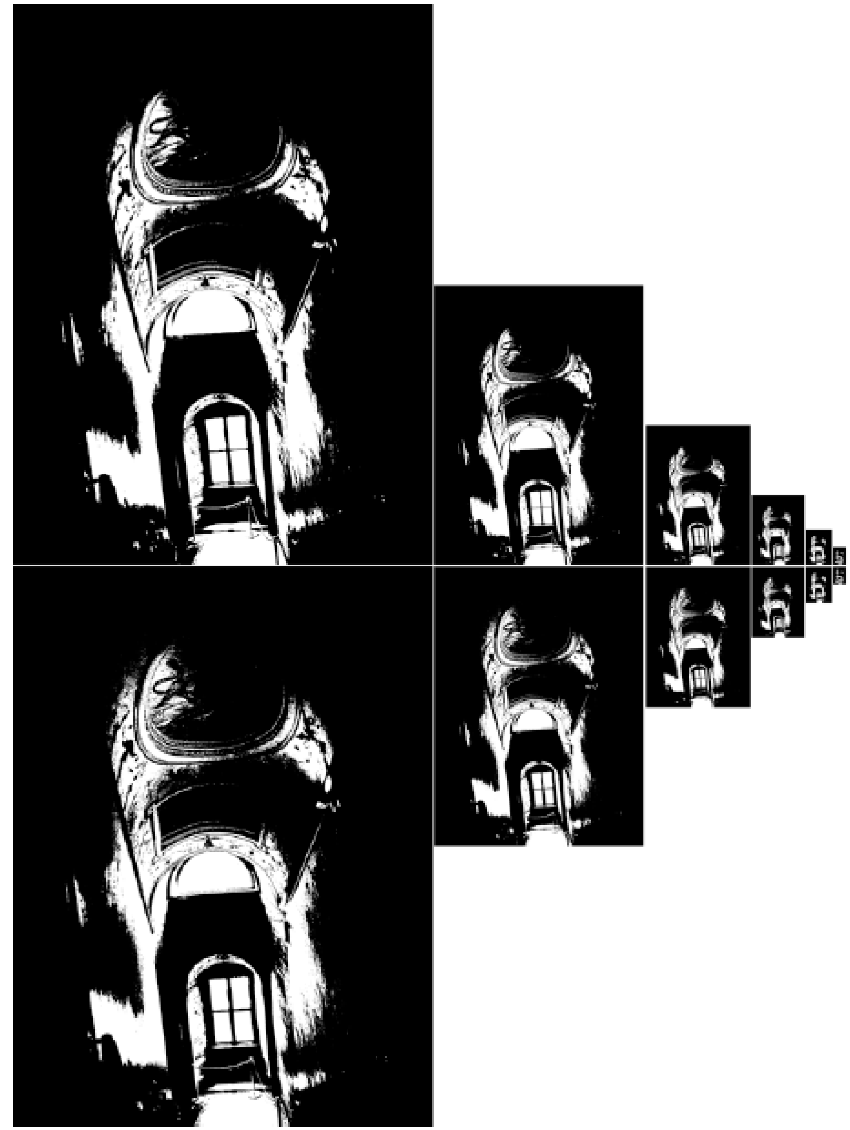


use the median grayscale value as a threshold

partition the pixels into two equal populations

Search for the Optimal Offset

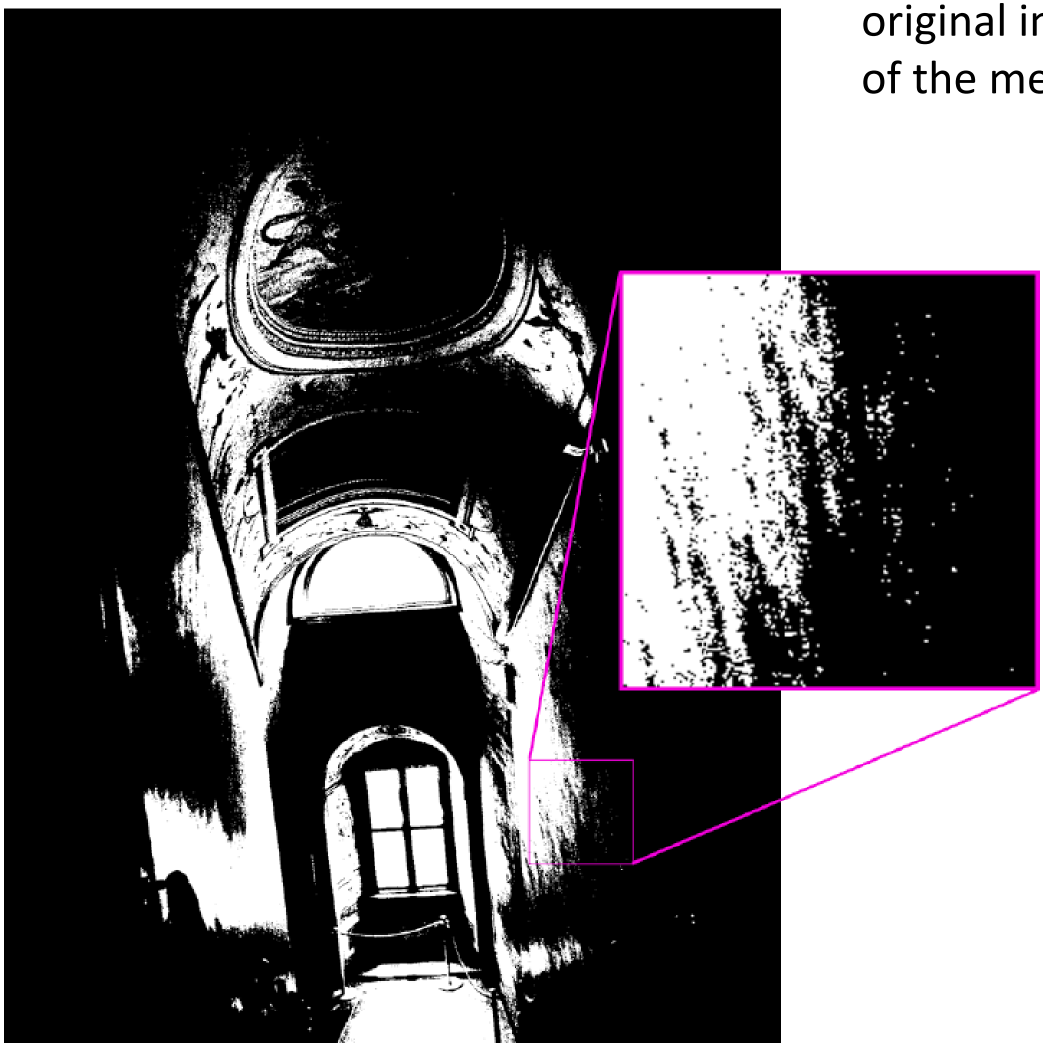
- Multi-scale

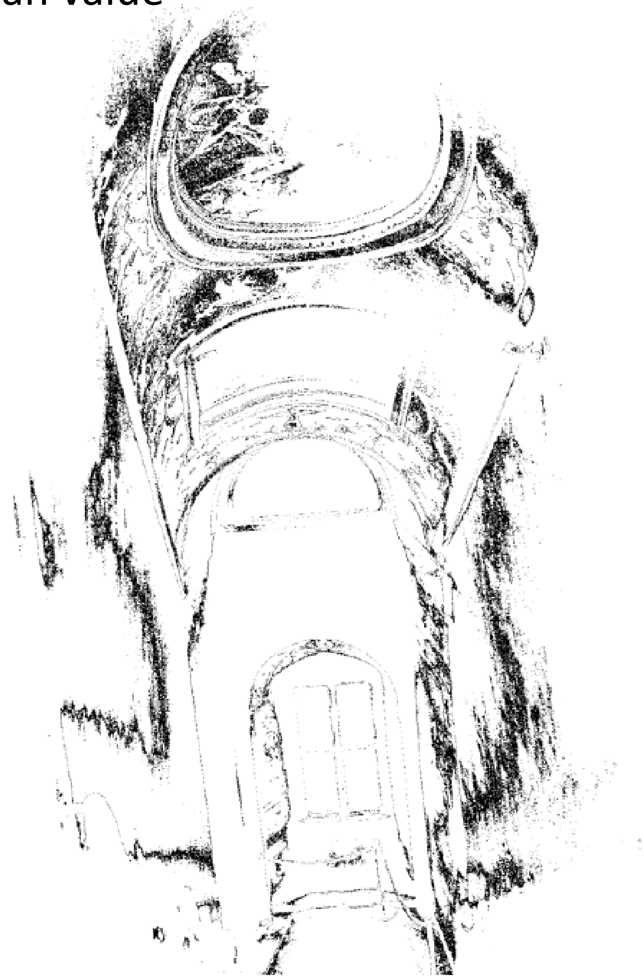


- log(max_offset) levels
- Within a range of -1~+1 pixel in each dimension at the lowest level
- Multiply the offset by 2 at the next level and then find the minimum XOR difference offset within a -1~+1 pixel

Threshold Noise

exclusion map: ignore wherever pixels in the original image are within the noise tolerance of the median value





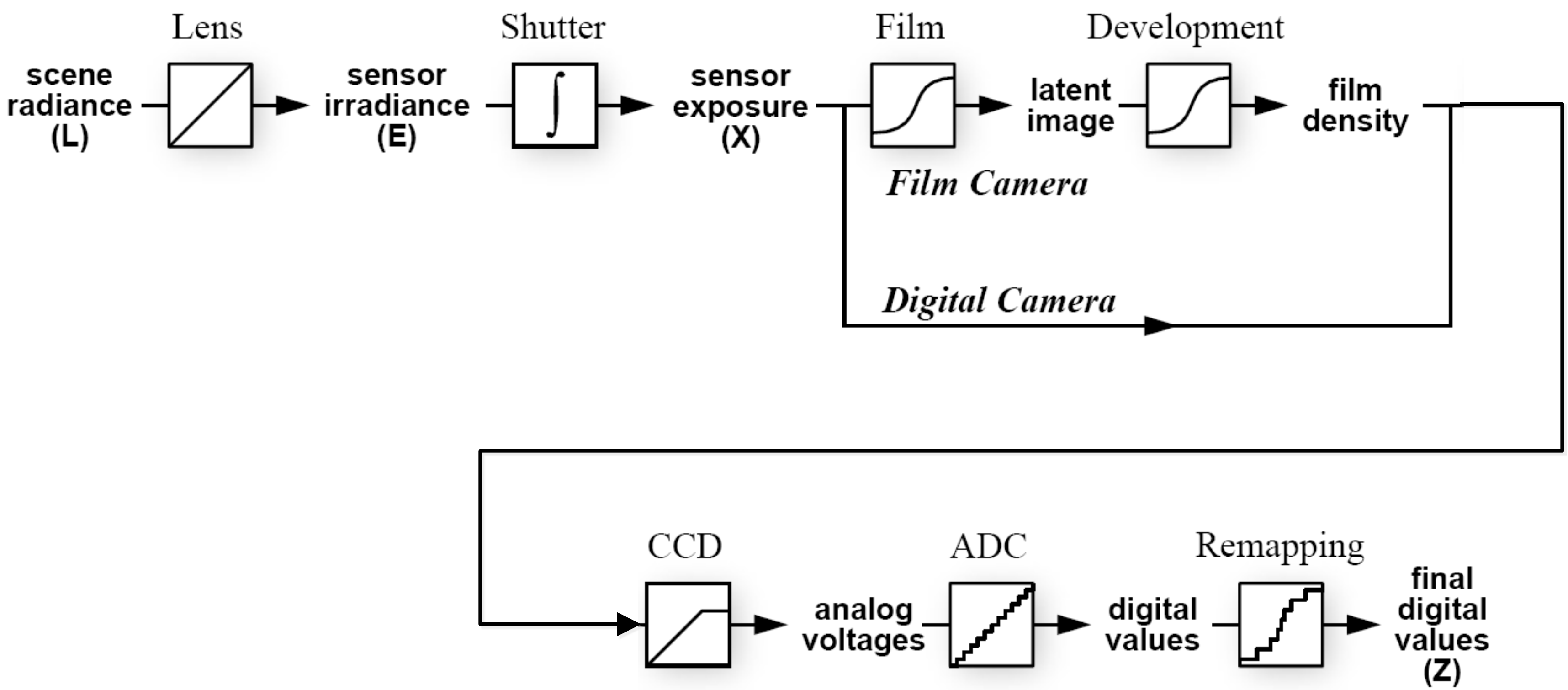
Efficiency Considerations

- Bitmap operations
  - XOR for differences
  - AND for exclusion maps
  - SHIFT for offsets
  - Pre-computed table for bit-counting
    - Number of 1 bits in the binary representations from 0 to 255 (i.e., {0, 1, 1, 2, 1, 2, 2, 3, 1, …, 8})
    - Add together the corresponding bit counts byte-by-byte through the whole bitmap

Result

Success rate is about 84% with about 10% failure due to image rotation, 3% due to excessive motion, and 3% caused by too much high-frequency content.

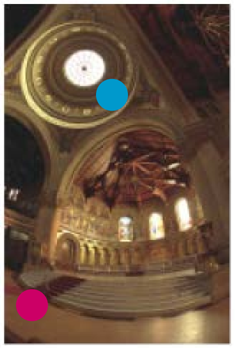
Image Acquisition Pipeline



nonlinear response curve

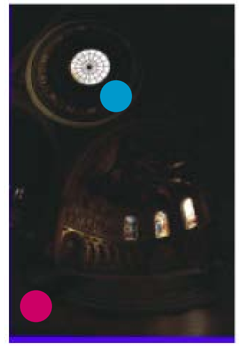
exposure $X = E\Delta t$

Recovering the Response Curve

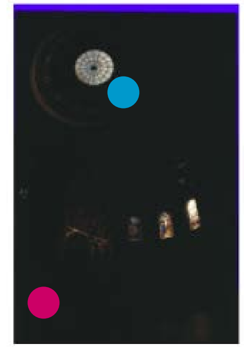


$\Delta t =$ 2 sec

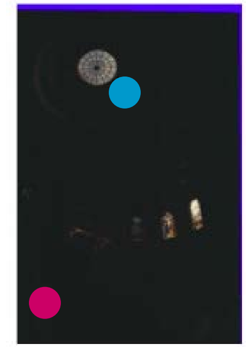
$\Delta t =$ 1/2 sec



$\Delta t =$ 1/8 sec



$\Delta t =$ 1/64 sec



$\Delta t =$ 1/256 sec

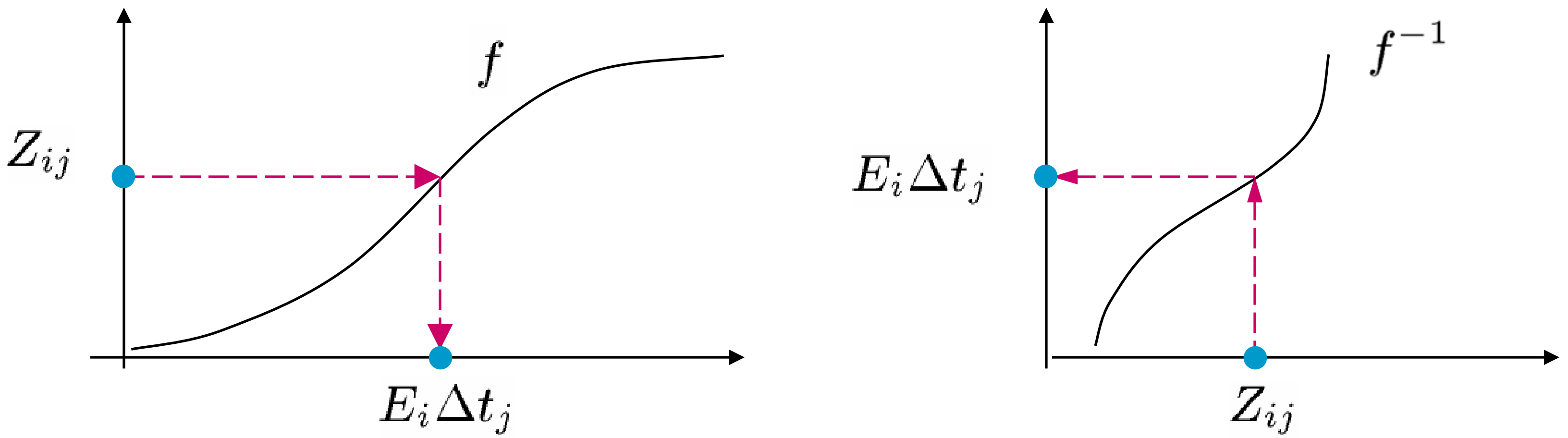
$$Z_{ij} = f(E_i \Delta t_j)$$

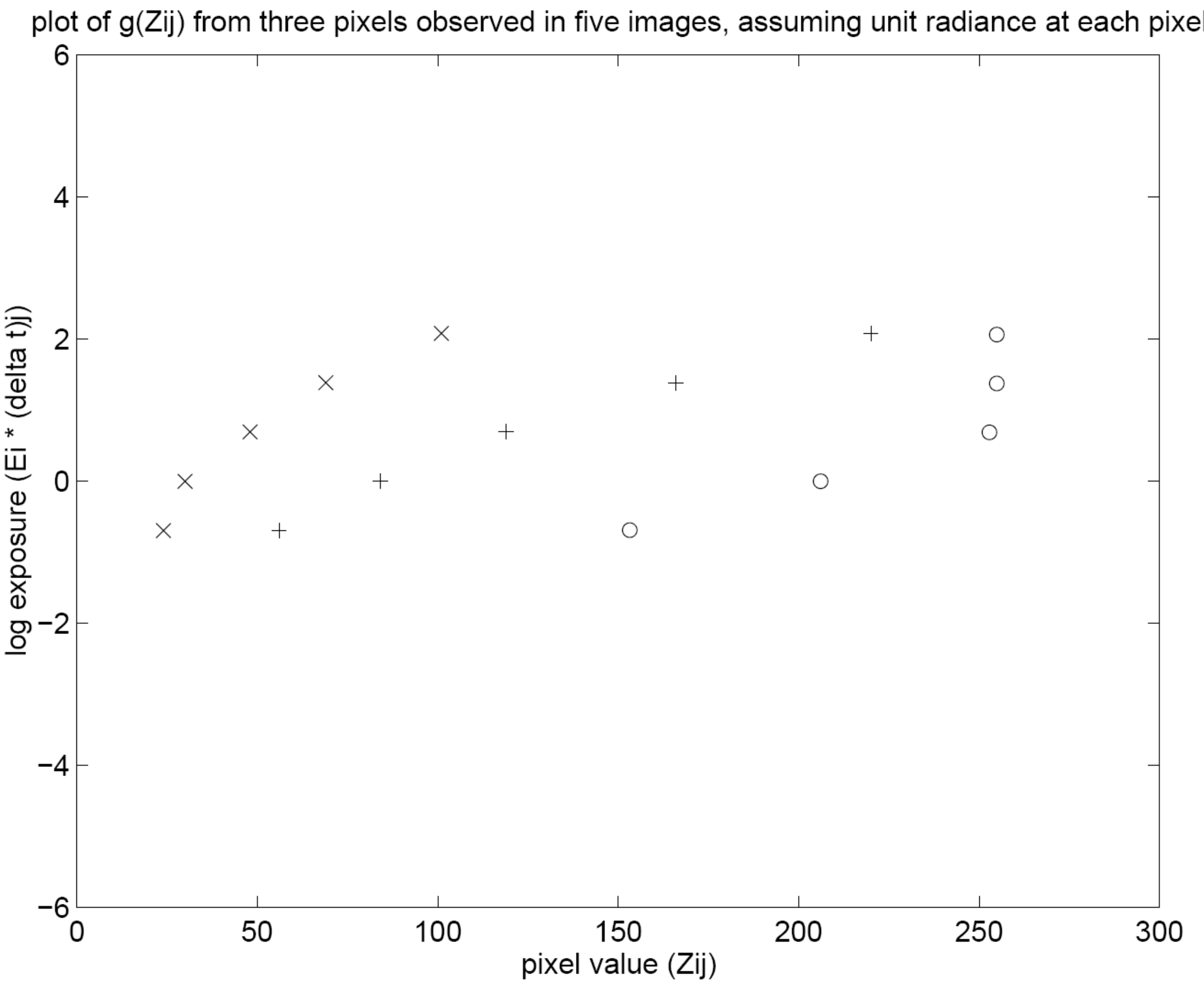
i: index for sampled pixel locations

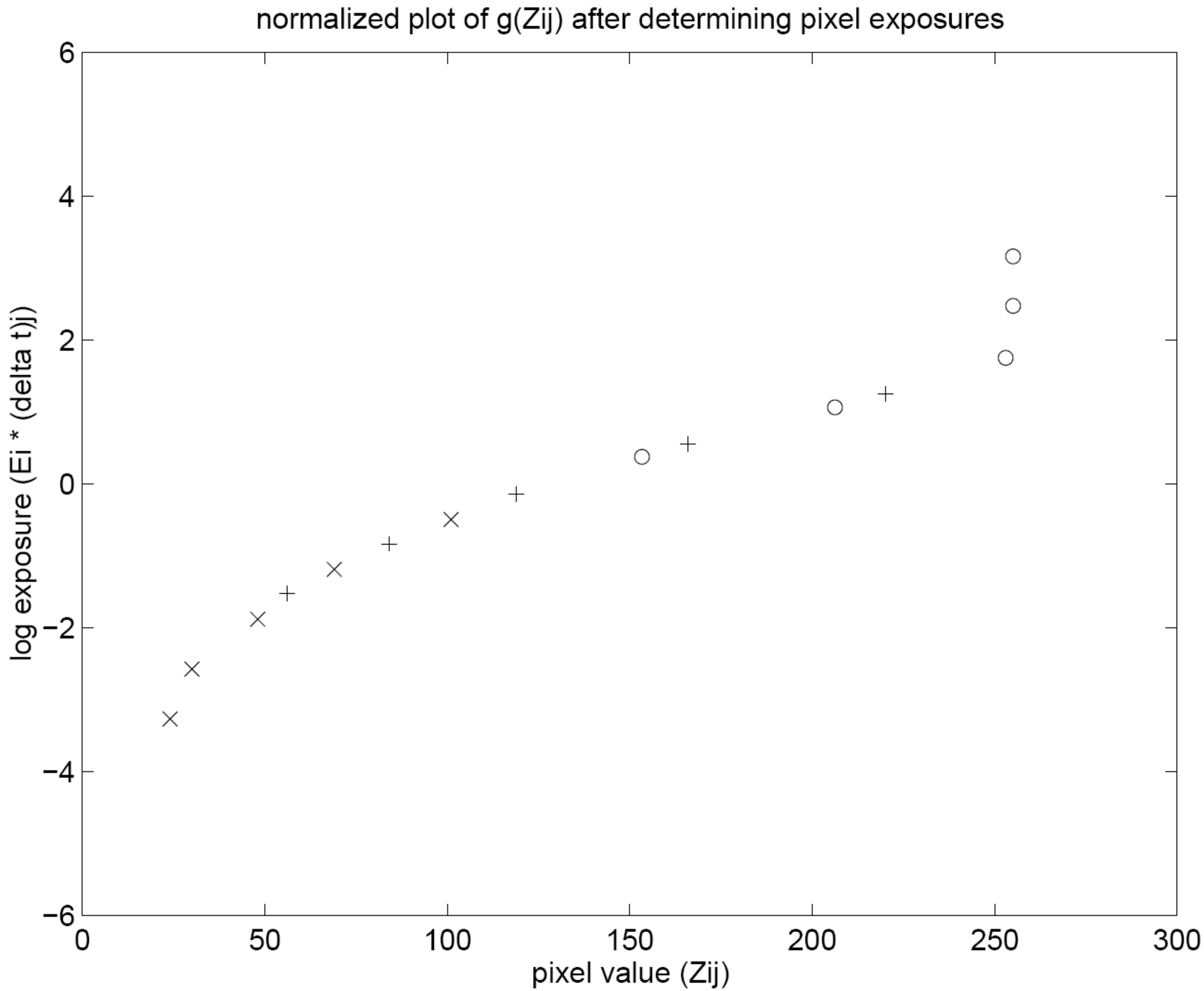
j: index for exposures

Recovering the Response Curve

- Finding the inverse of the response curve




plot of g(Zij) from three pixels observed in five images, assuming unit radiance at each pixel
log exposure (Ei * (delta t)j)
pixel value (Zij)
6
4
2
0
−2
−4
−6
0
50
100
150
200
250
300


normalized plot of g(Zij) after determining pixel exposures
log exposure (Ei * (delta t)j)
pixel value (Zij)
6
4
2
0
−2
−4
−6
0
50
100
150
200
250
300

$$Z_{ij} = f(E_i \Delta t_j)$$

known: Z_{ij} Δt_j
unknown: g E_i

assume f is monotonic, it is invertible

$$\ln f^{-1}(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

define function $g = \ln f^{-1}$

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

The Problem

- Recovering g only requires recovering the finite number of values that $g(z)$ can take since the domain of Z is finite ($0 \sim 255$)
- Let N be the number of pixel locations and P be the number of photographs, we formulate the problem as one of finding the $(Z_{max} - Z_{min} + 1)$ values of $g(Z)$ and the N values of $\ln E_i$ that minimize

$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} g''(z)^2$$

$$g''(z) = g(z-1) - 2g(z) + g(z+1)$$

- The solution can be only up to a scale
  - Introduce the additional constraint

$$g(Z_{mid}) = 0, \text{ where } Z_{mid} = \tfrac{1}{2}(Z_{min} + Z_{max})$$

- A simple hat weighting function

$$w(z) = \begin{cases} z - Z_{min} & \text{for } z \le \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z & \text{for } z > \frac{1}{2}(Z_{min} + Z_{max}) \end{cases}$$

$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} \{w(Z_{ij})\,[g(Z_{ij}) - \ln E_i - \ln \Delta t_j]\}^2 +$$

$$\lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} [w(z) g''(z)]^2$$

$P=11$, $N=50$, $(Z_{max}-Z_{min})=255 \rightarrow N(P-1)>(Z_{max}-Z_{min})$

- Sampling
  - The pixel locations should be chosen so that they have a reasonably even distribution of pixel values from Z_{min} to Z_{max}, and so that they are spatially well distributed in the image
  - The pixels are best sampled from regions of the image with low intensity variance so that radiance can be assumed to be constant across the area of the pixel, and the effect of optical blur of the imaging system is minimized

- It is an over-determined system of linear equations and can be solved by finding the least squares solution

Optimization

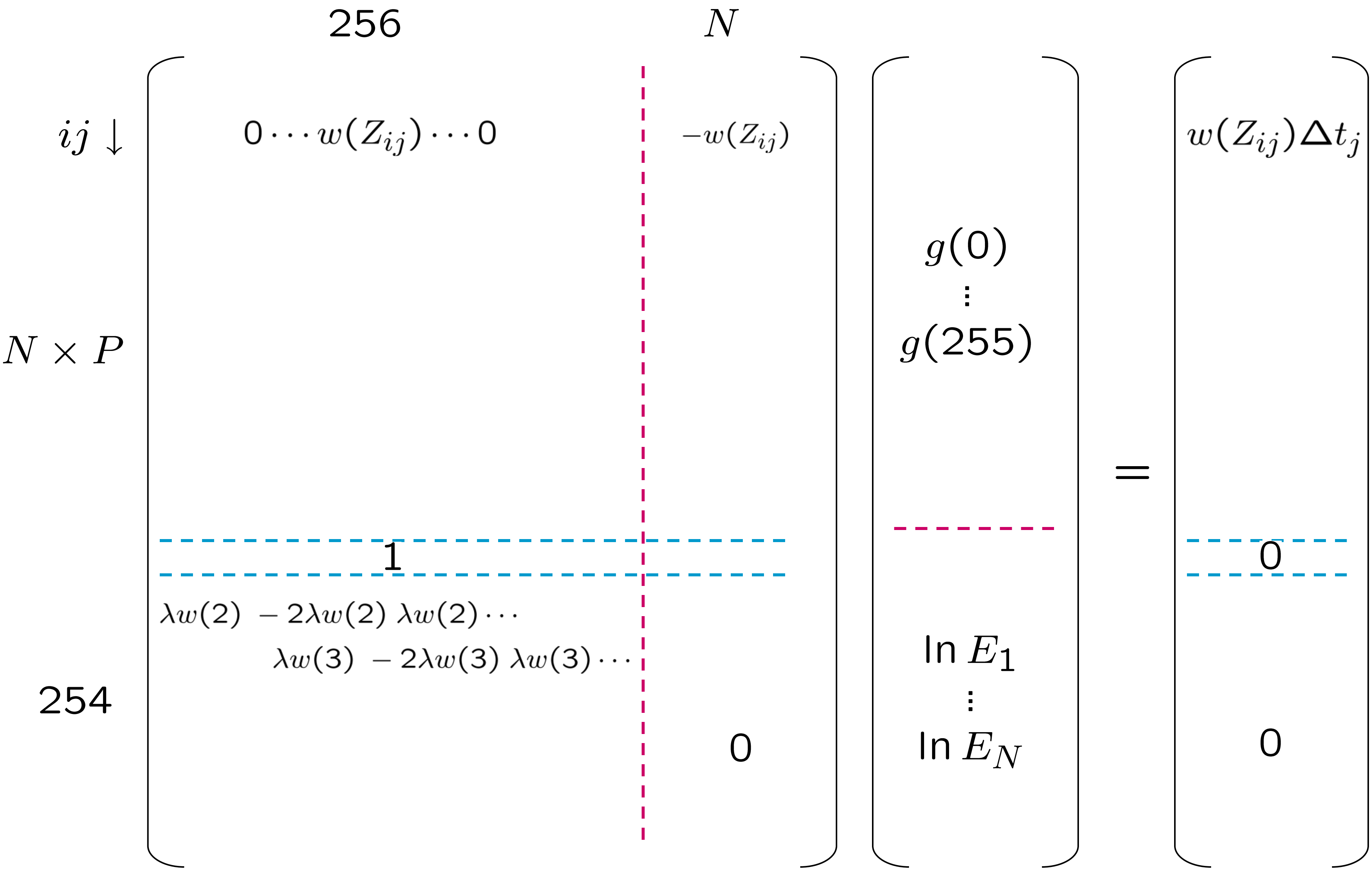
$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} \{w(Z_{ij}) \left[g(Z_{ij}) - \ln E_i - \ln \Delta t_j\right]\}^2 +$$

$$\lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} \left[w(z) g''(z)\right]^2$$

$$g(Z_{mid}) = 0, \text{ where } Z_{mid} = \tfrac{1}{2}(Z_{min} + Z_{max})$$

least squares solution to $Ax = b$

A Sparse Linear System



Least Squares Solution to an Over-determined Linear System

$$Ax = b$$

We are given *m* equations in *n* unknowns, with *m* > *n*

The vector x that minimizes $\|Ax - b\|^2$ is the solution to the normal equations

$$A^T A x = A^T b$$

This vector $x = (A^T A)^{-1} A^T b$ is the least squares solution to $Ax = b$

Proof of Least Squares Solution

MATLAB code

```
%
% gsolve.m - Solve for imaging system response function
%
% Given a set of pixel values observed for several pixels in several
% images with different exposure times, this function returns the
% imaging system's response function g as well as the log film irradiance
% values for the observed pixels.
%
% Assumes:
%
%  Zmin = 0
%  Zmax = 255
%
% Arguments:
%
%  Z(i,j)  is the pixel values of pixel location number i in image j
%  B(j)    is the log delta t, or log shutter speed, for image j
%  l       is lamdba, the constant that determines the amount of smoothness
%  w(z)    is the weighting function value for pixel value z
%
% Returns:
%
%  g(z)    is the log exposure corresponding to pixel value z
%  lE(i)   is the log film irradiance at pixel location i
%
```

MATLAB code

```
function [g,lE]=gsolve(Z,B,l,w)

n = 256;

A = zeros(size(Z,1)*size(Z,2)+n+1,n+size(Z,1));
b = zeros(size(A,1),1);

%% Include the data-fitting equations

k = 1;
for i=1:size(Z,1)
  for j=1:size(Z,2)
    wij = w(Z(i,j)+1);
    A(k,Z(i,j)+1) = wij;  A(k,n+i) = -wij;        b(k,1) = wij * B(i,j);
    k=k+1;
  end
end
```

$$\sum_{i=1}^{N} \sum_{j=1}^{P} \left\{ w(Z_{ij}) \left[g(Z_{ij}) - \ln E_i - \ln \Delta t_j \right] \right\}^2$$

MATLAB code

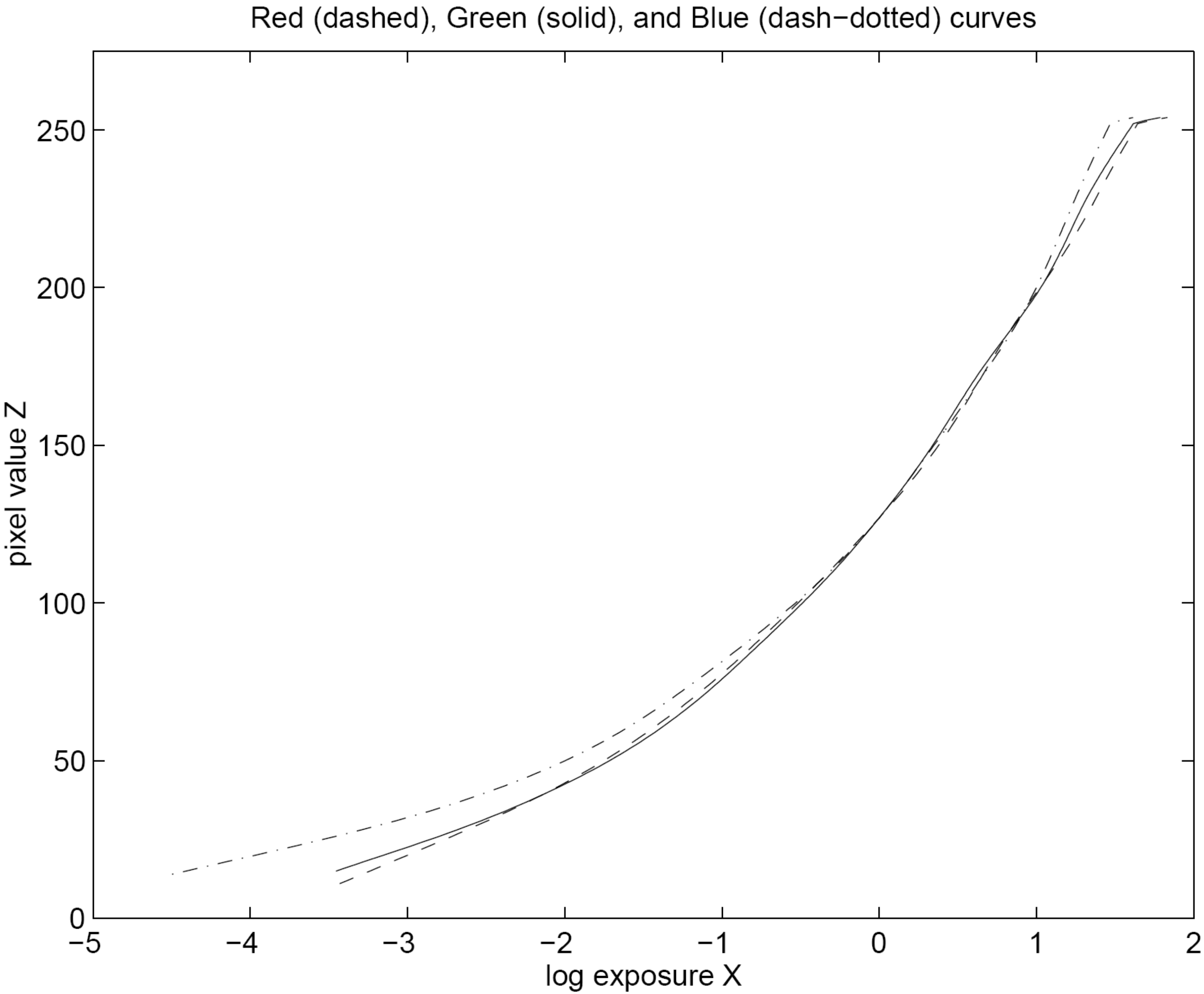
```
%% Fix the curve by setting its middle value to 0

A(k,129) = 1;
k=k+1;
```

$$g(Z_{mid}) = 0, \text{ where } Z_{mid} = \frac{1}{2}(Z_{min} + Z_{max})$$

```
%% Include the smoothness equations

for i=1:n-2
  A(k,i)=l*w(i+1);        A(k,i+1)=-2*l*w(i+1);  A(k,i+2)=l*w(i+1);
  k=k+1;
end

%% Solve the system using SVD

x = A\b;

g = x(1:n);
lE = x(n+1:size(x,1));
```

$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} \{w(Z_{ij}) [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]\}^2 +$$

$$\lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} [w(z) g''(z)]^2$$


Red (dashed), Green (solid), and Blue (dash−dotted) curves
pixel value Z
250
200
150
100
50
0
−5
−4
−3
−2
−1
0
1
2
log exposure X

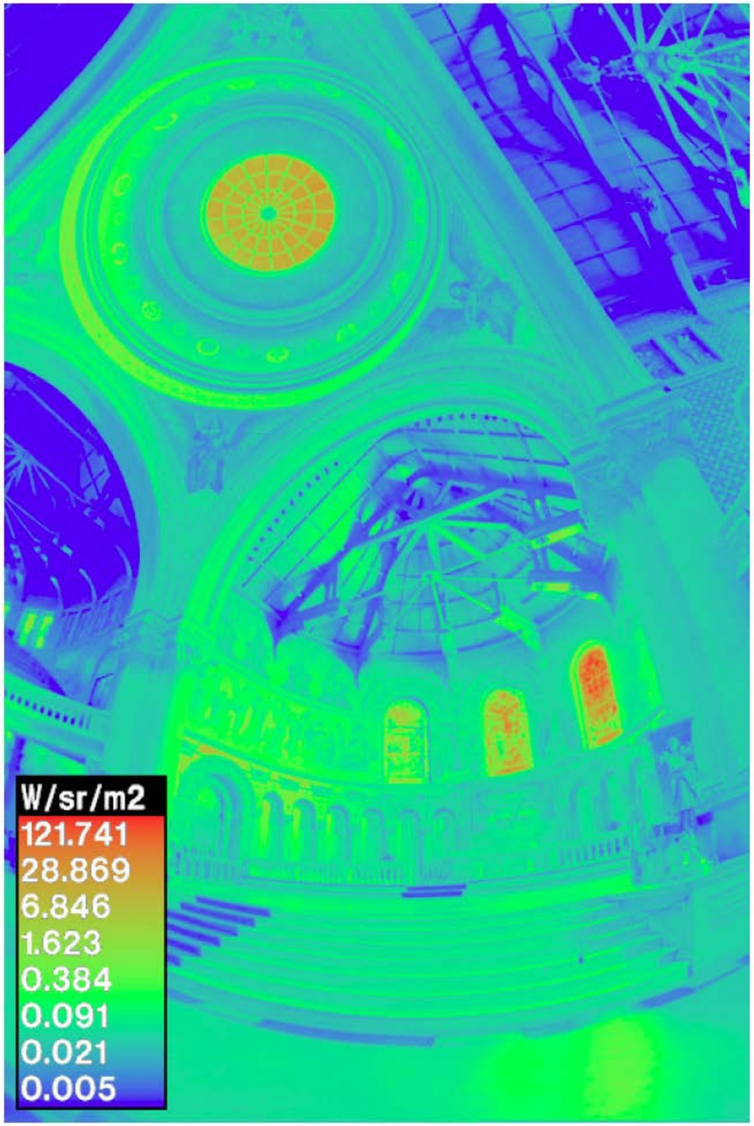
Constructing the HDR Radiance Map

$$\ln E_i = g(Z_{ij}) - \ln \Delta t_j$$

We should use all the available exposures for a particular pixel to compute its radiance:

$$\ln E_i = \frac{\sum_{j=1}^{P} w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^{P} w(Z_{ij})}$$

Reconstructed Radiance Map



Assignment

- Recovering high dynamic range radiance maps from photographs
  - Taking some photos of a scene in our campus with different exposure levels using your camera obscura

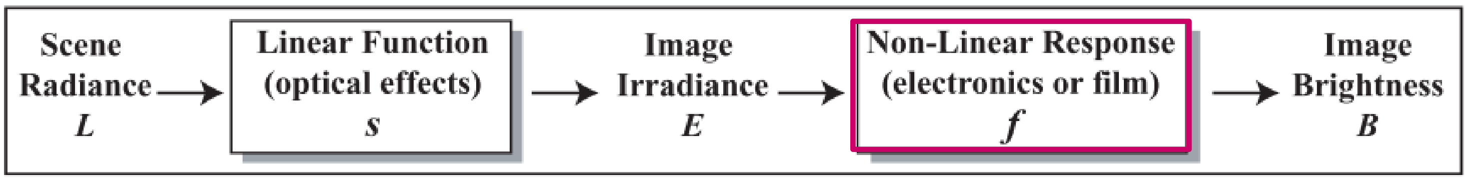
  - Use the MATLAB function 'makehdr' or the code of Debevec & Malik to recover the radiance map
  - Use the MATLAB functions ' hdrwrite' and 'hdrread' to write and read HDR files
  - Use the MATLAB function 'tonemap‘ to show the HDR images

Today's Plan

- *Recovering High Dynamic Range Radiance Maps from Photographs*
  - Debevec and Malik
  - SIGGRAPH 1997
- *What Is the Space of Camera Response Functions?*
  - Grossberg and Nayar
  - CVPR 2003

Camera Response Function

- Mimicking the non-linearity of film
- Spatially uniform

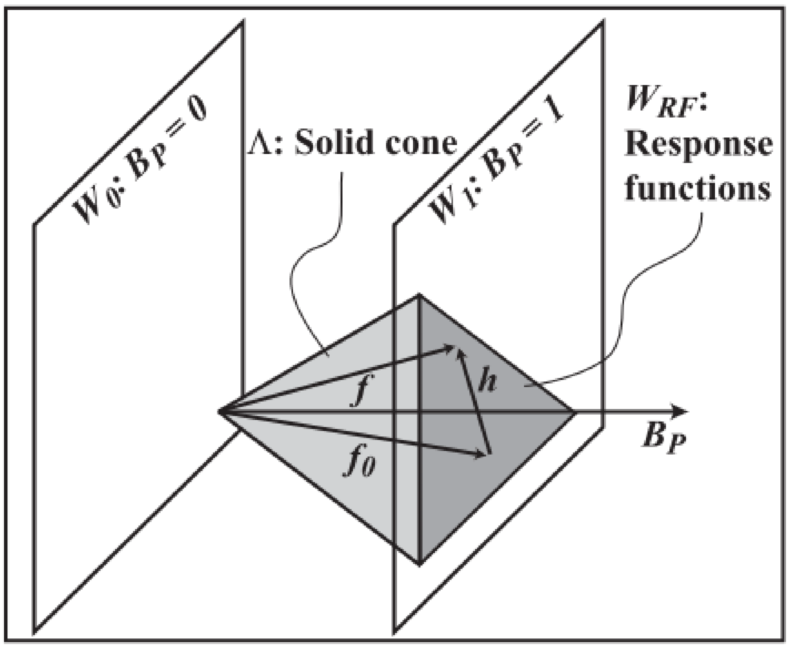


$$f(E) = \alpha + \beta E^{\gamma}$$?

Monotonicity, smoothness constraints

Theoretical Space of Camera Response Functions

$$W_{\mathrm{RF}} := \{f \mid f(0) = 0, f(1) = 1, \text{ and } f \text{ monotonically increasing } \}.$$



positive linear combinations

$$(B_1, \ldots, B_P) = (f(E_1), \ldots, f(E_P))$$

high-dimensional vector

= 1

Approximation Model

$$f_0(E) + \sum_{n=1}^{M} c_n h_n(E)$$

polynomial model

$$f_0(E) := E$$

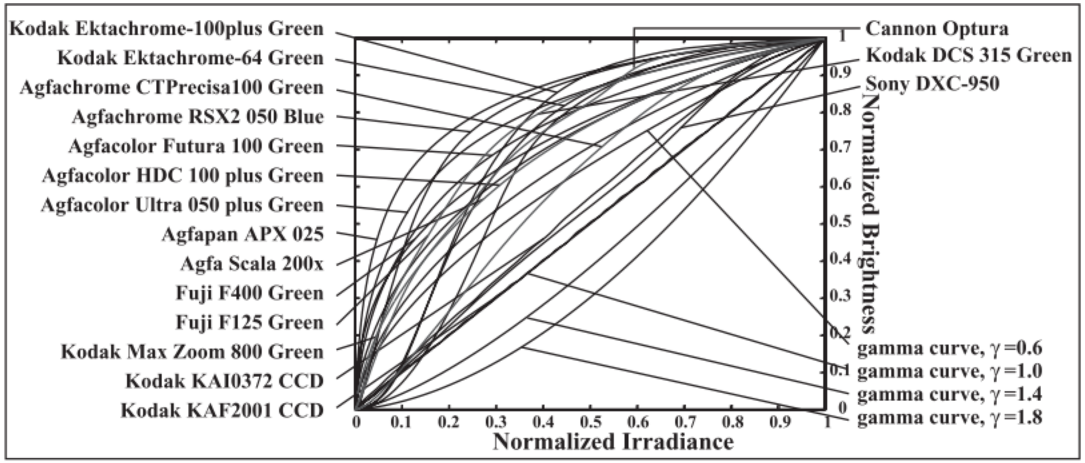
$$h_n(E) := E^{n+1} - E$$

trigonometric approximation model

$$f_0(E) := E$$

$$h_n(E) := \sin(n\pi E)$$

Empirical Model

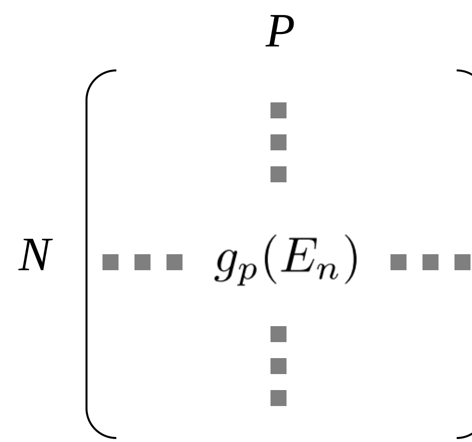


densely sampling at $\{E_1, \ldots, E_P\}$

covariance

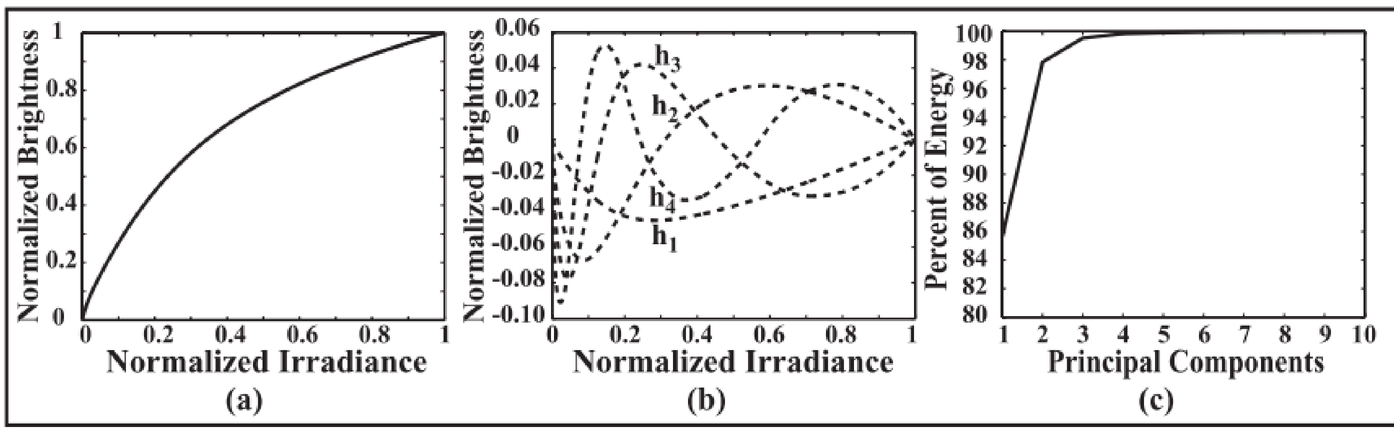
$$C_{m,n} = \sum_{p=1}^{N} (g_p(E_n) - f_0(E_n))(g_p(E_m) - f_0(E_m))$$

mean curve



Principal Component Analysis

- M-dimensional approximation, eigenspaces associated with the largest M eigenvalues of the covariance matrix



$$\tilde{f} = f_0 + Hc \qquad c = H^T(f - f_0)$$

Imposing Monotonicity

$$D\tilde{f}^{mon} \geq 0$$

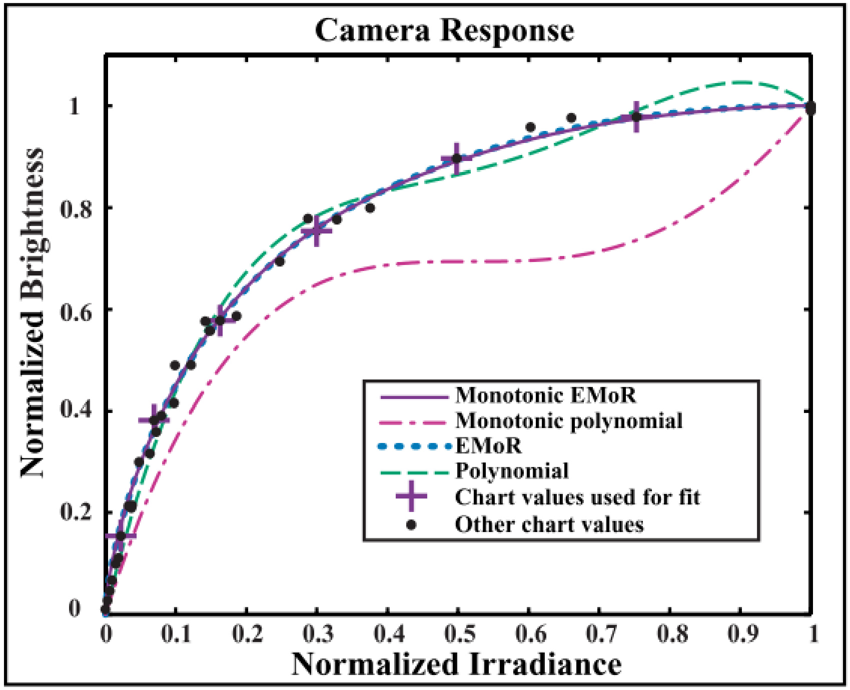
$$\tilde{f}^{mon} = f_0 + H\hat{c}$$

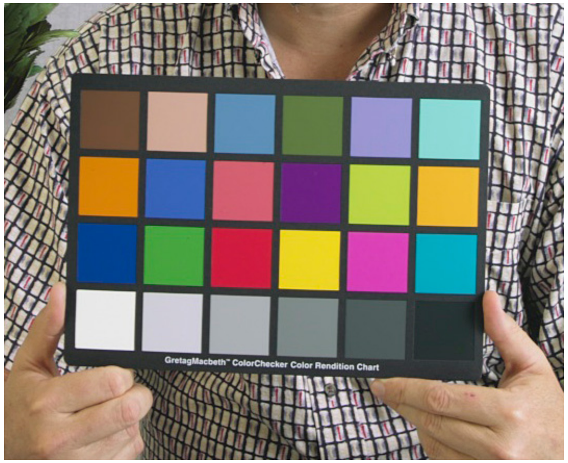
$$\hat{c} = \arg\min_c ||Hc - (f - f_0)||^2,$$

$$DH\hat{c} \geq -Df_0$$

quadratic programming

Interpolation from Sparse Samples





color chart [wikipedia]

six patches with known reflectances

Inverse Response Function from Multiple Images

$$g(B) = g_0(B) + \sum_n^M c_n h_n^{\text{inv}}(B)$$

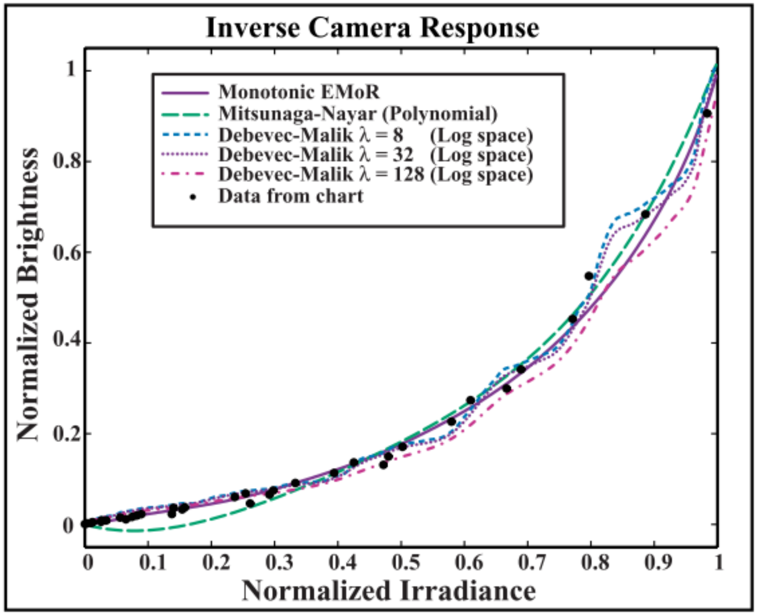
exposure e and $k \cdot e$

$$g(B_a) = kg(B_b)$$

$$g(B_a) - kg(B_b) = 0$$

linear in coefficients c_n

solved by least-squares



Learning by Doing

- Assignments