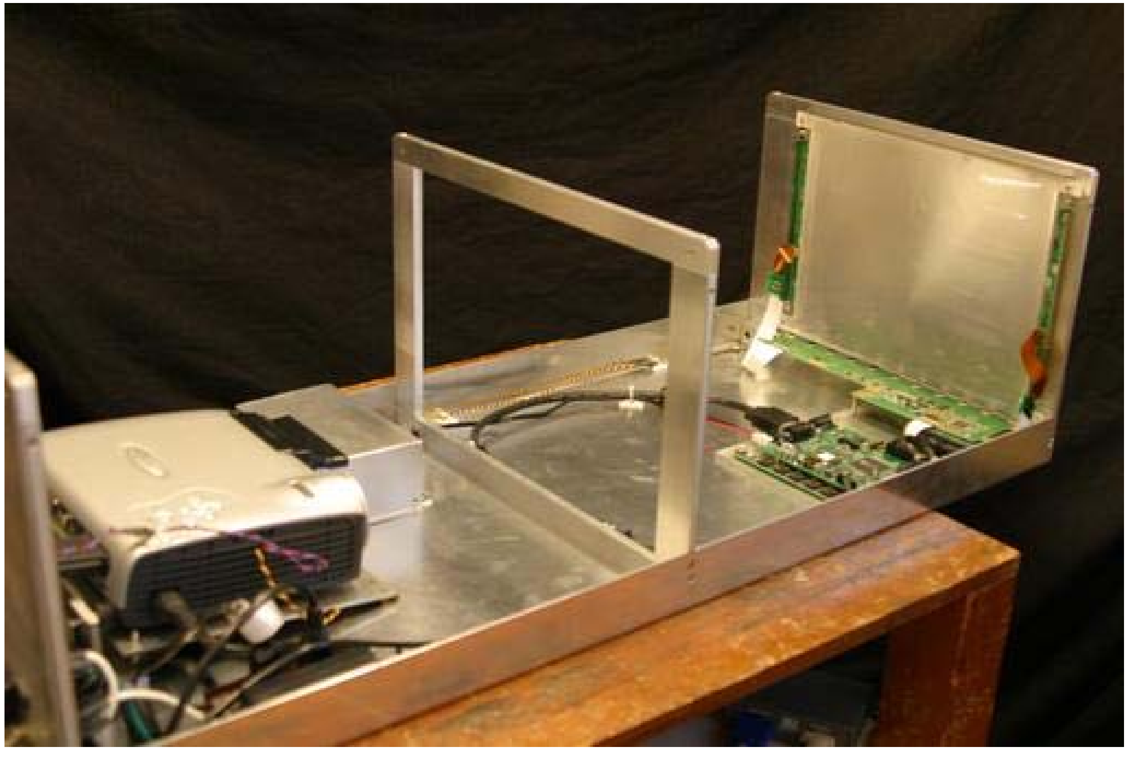
Tone mapping

Digital Visual Effects, Spring 2008
Yung-Yu Chuang
2008/3/4

with slides by Fredo Durand, and Alexei Efros

Display HDR

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

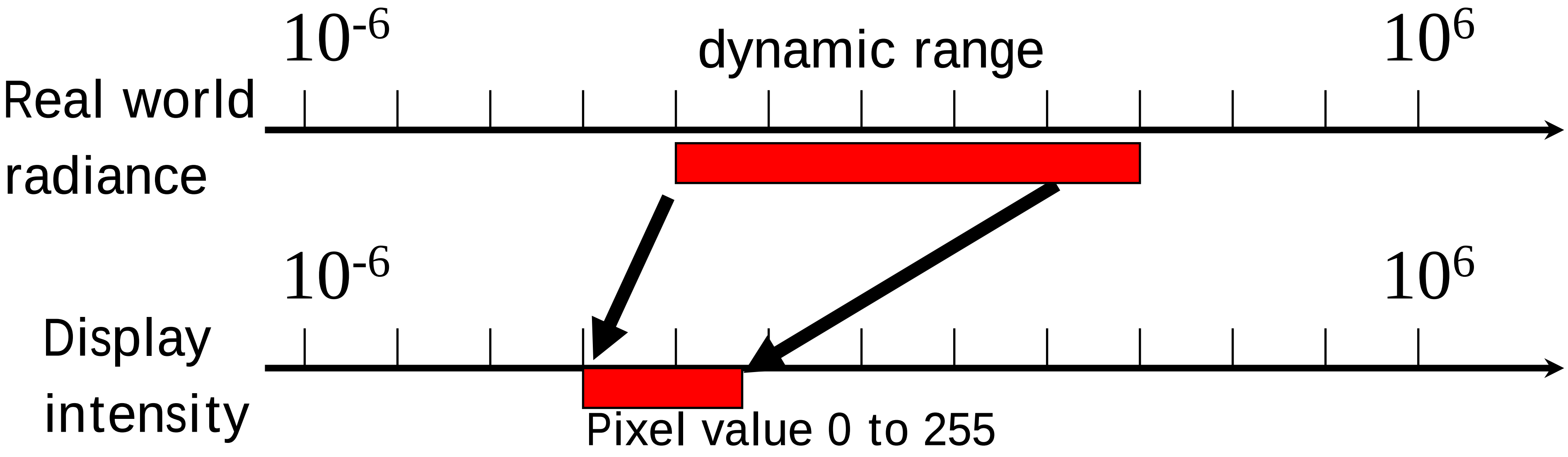


HDR display system, Sunnybrook, 2004

Tone mapping

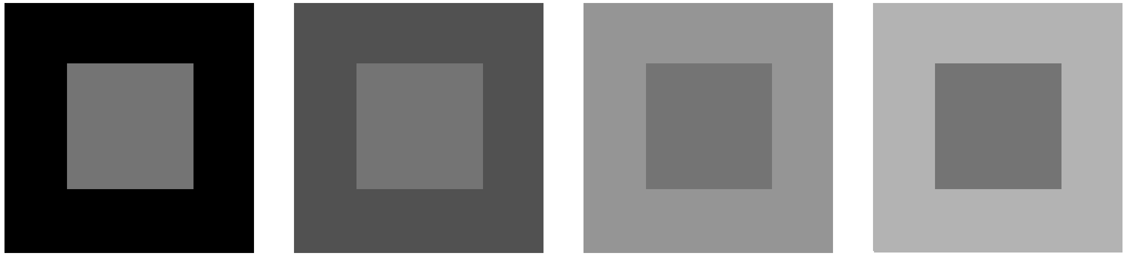
- How should we map scene luminances (up to 1:100,000) to display luminances (only around 1:100) to produce a satisfactory image?

  Linear scaling?, thresholding?



CRT has 300:1 dynamic range

Eye is not a photometer!



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

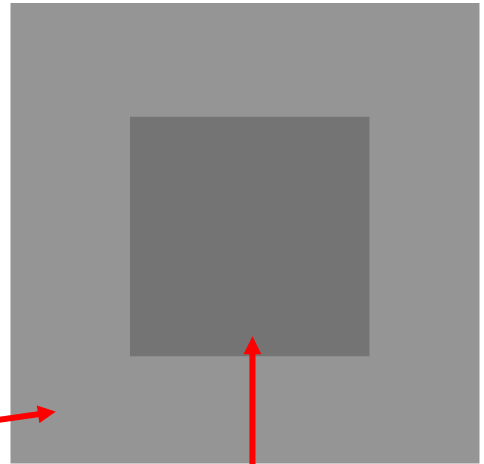
We are more sensitive to contrast

- Weber's law

Just-noticeable Difference (JND)

$$\frac{\Delta I_b}{I_b} \sim 1\%$$

background intensity



flash

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

Preliminaries

- For color images

$$\begin{bmatrix} R_d \\ G_d \\ B_d \end{bmatrix} = \begin{bmatrix} L_d \dfrac{R_w}{L_w} \\ L_d \dfrac{G_w}{L_w} \\ L_d \dfrac{B_w}{L_w} \end{bmatrix}$$

- Log domain is usually preferred.

Tone mapping operators

- 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

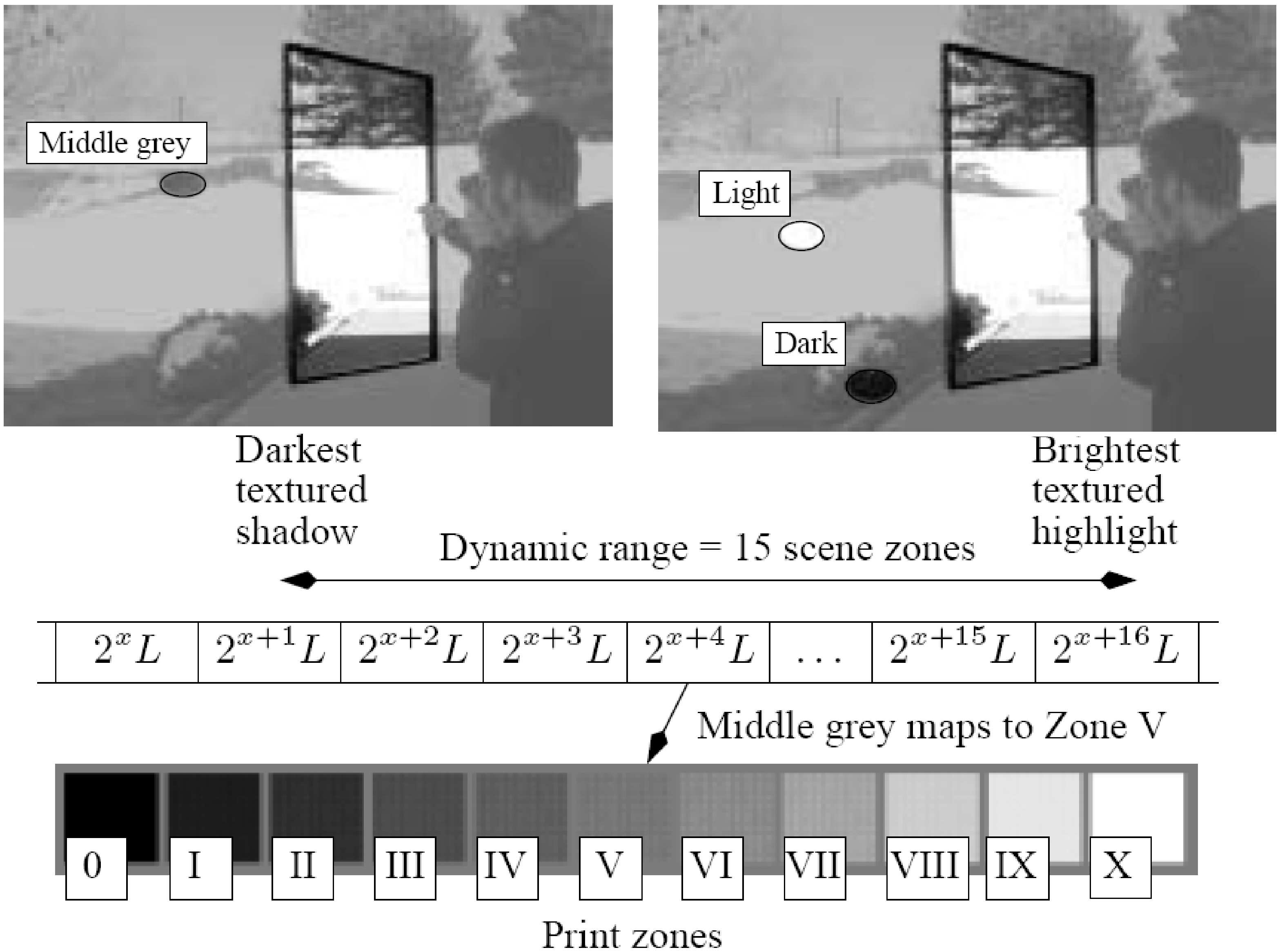
Peter Shirley   Jim Ferwerda

SIGGRAPH 2002

Photographic tone reproduction

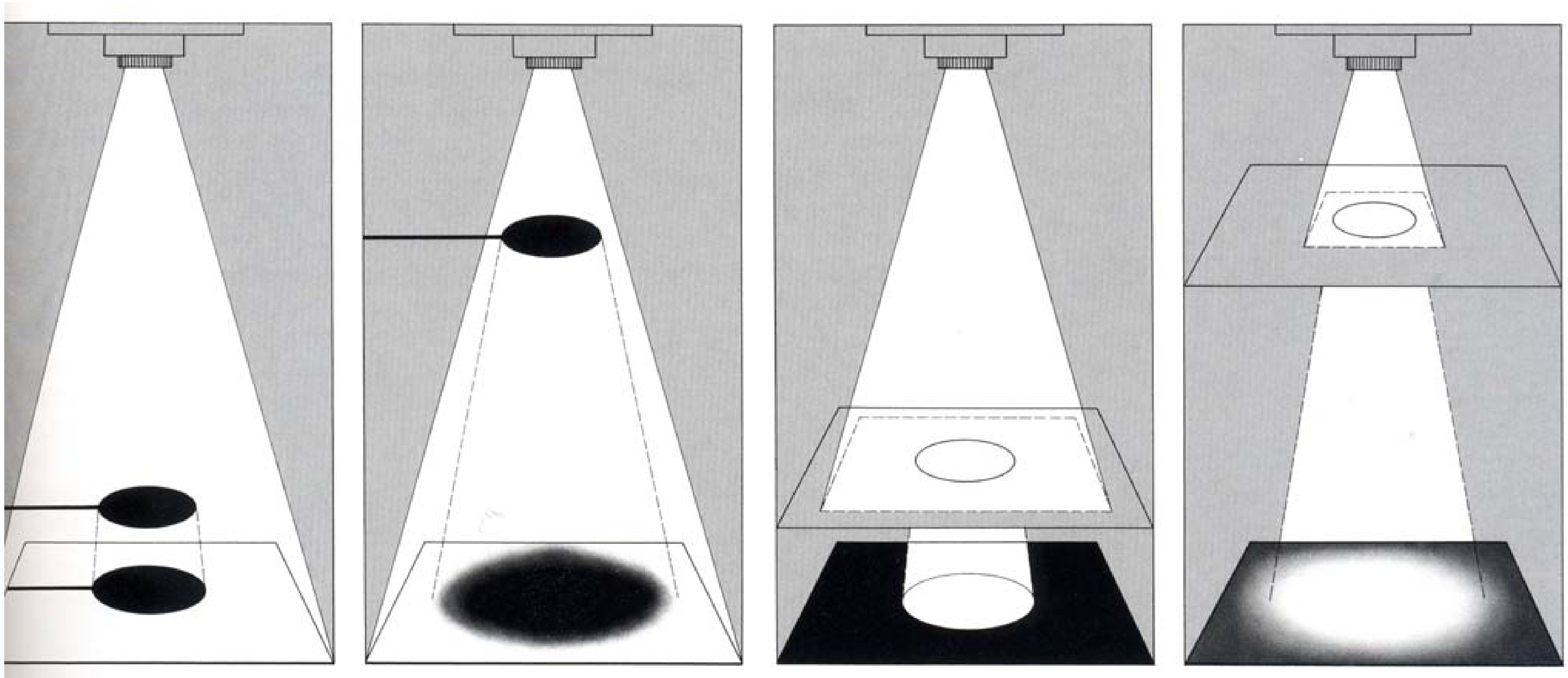
- Proposed by Reinhard et. al. in SIGGRAPH 2002
- Motivated by traditional practice, zone system by Ansel Adams and dodging and burning
- It contains both global and local operators

Zone system


Middle grey
Light
Dark
Darkest textured shadow
Brightest textured highlight
Dynamic range = 15 scene zones
$2^x L$ | $2^{x+1} L$ | $2^{x+2} L$ | $2^{x+3} L$ | $2^{x+4} L$ | ... | $2^{x+15} L$ | $2^{x+16} L$
Middle grey maps to Zone V
0 I II III IV V VI VII VIII IX X
Print zones

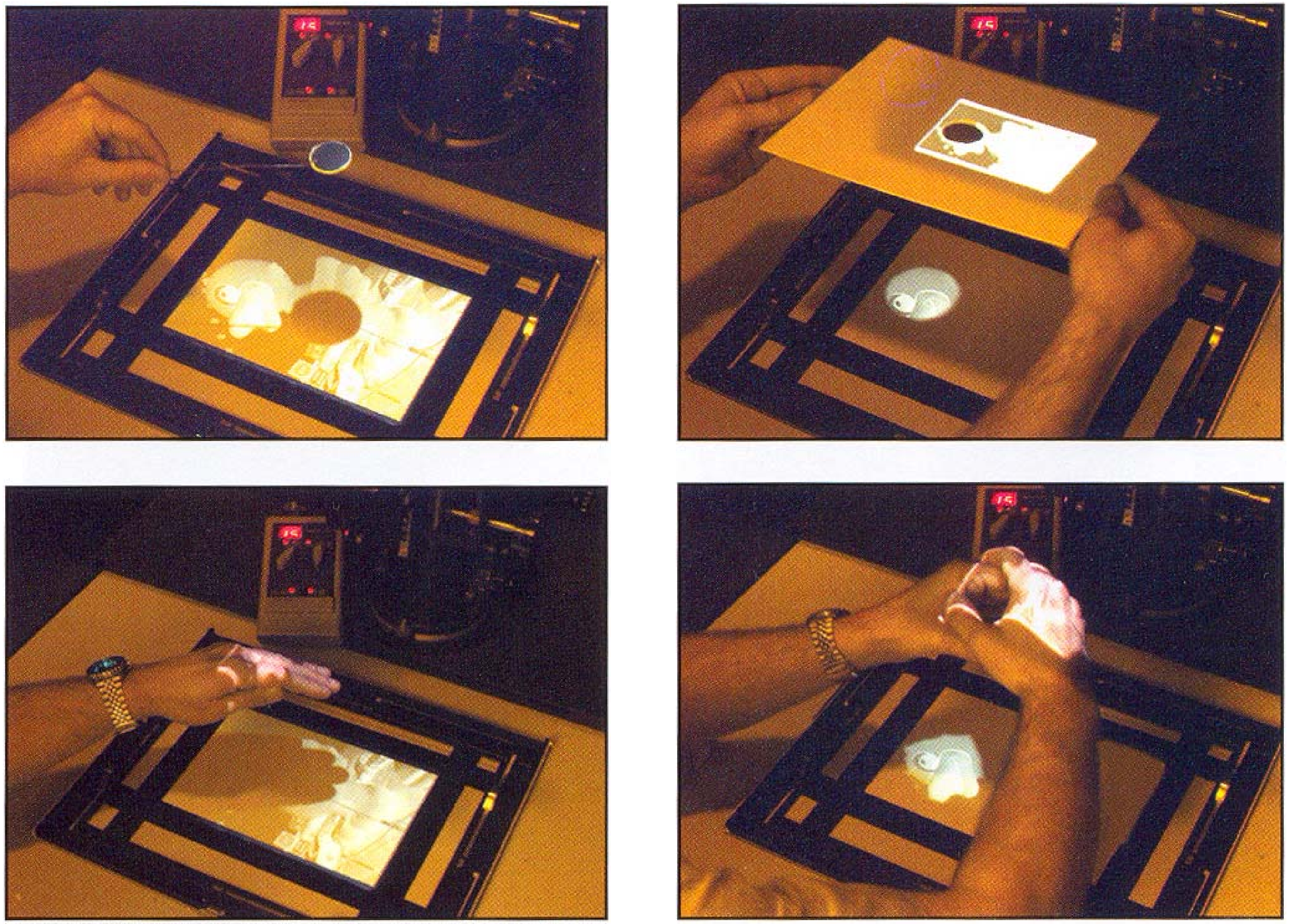
Dodging and burning

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



From The Master Printing Course, Rudman

Dodging and burning

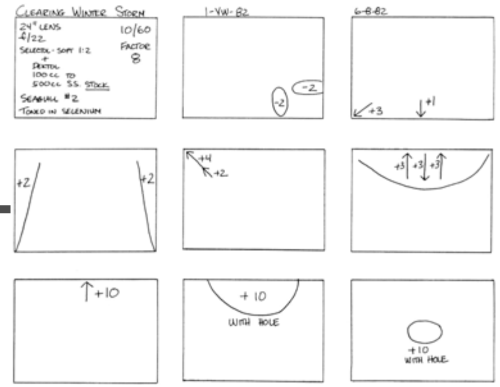


dodging burning

From Photography by London et al.

Dodging and burning

- Must be done for every single print!



Straight print

After dodging and burning

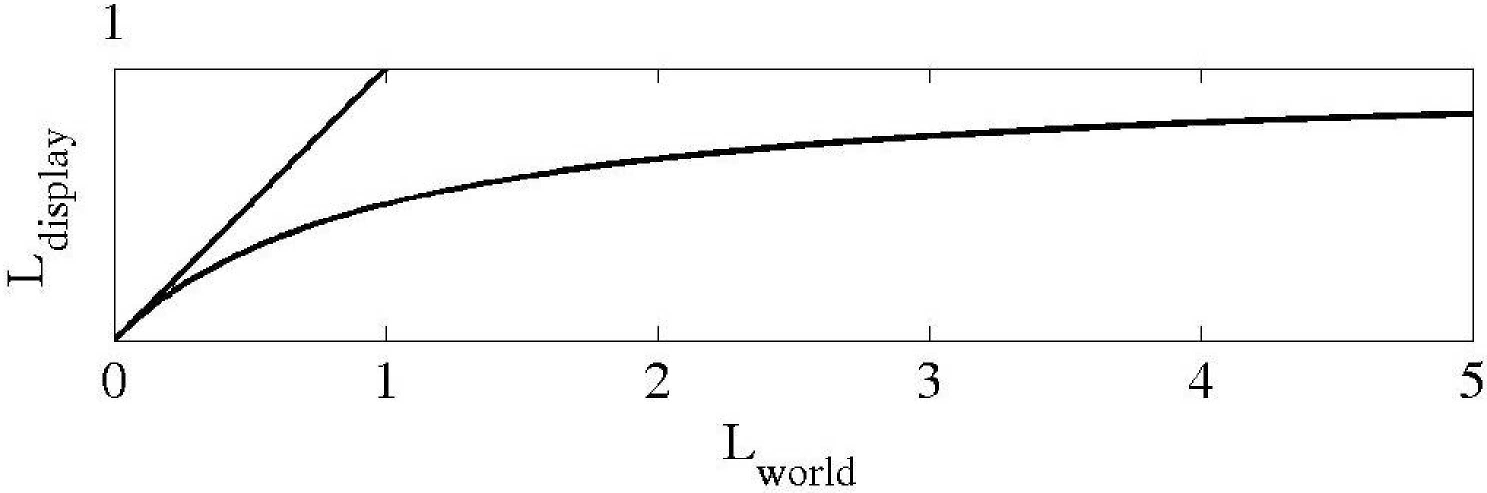
Global operator

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

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

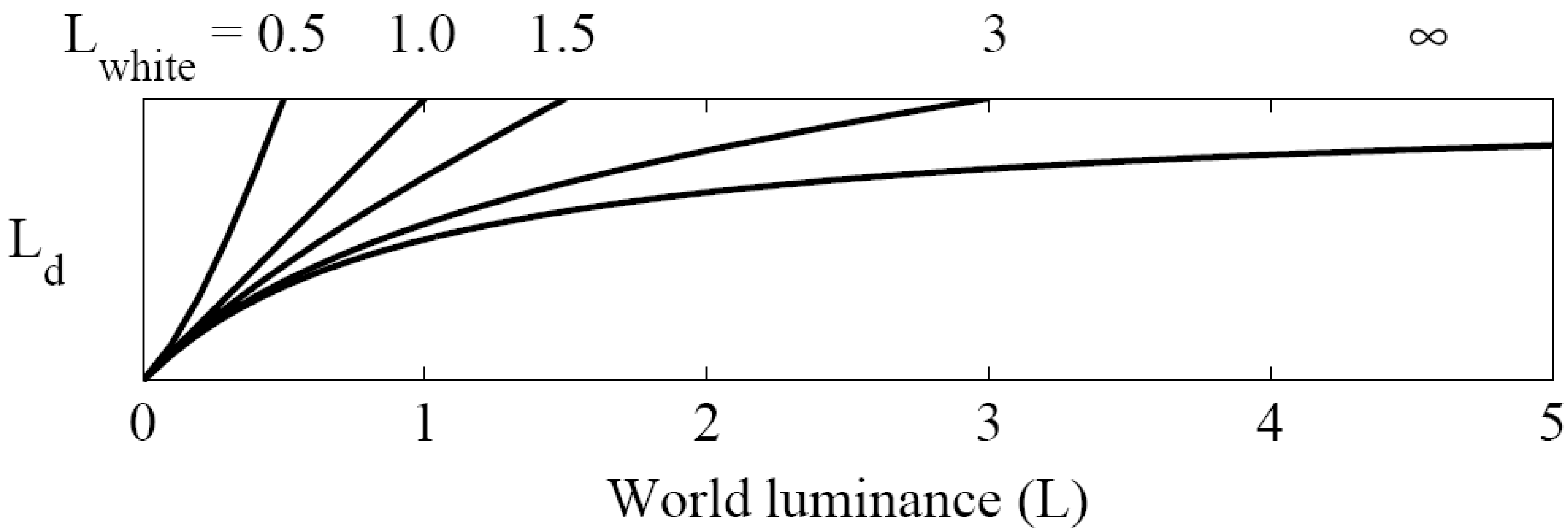
$$L_m(x,y) = \frac{a}{\bar{L}_w} L_w(x,y) \qquad L_d(x,y) = \frac{L_m(x,y)}{1 + L_m(x,y)}$$

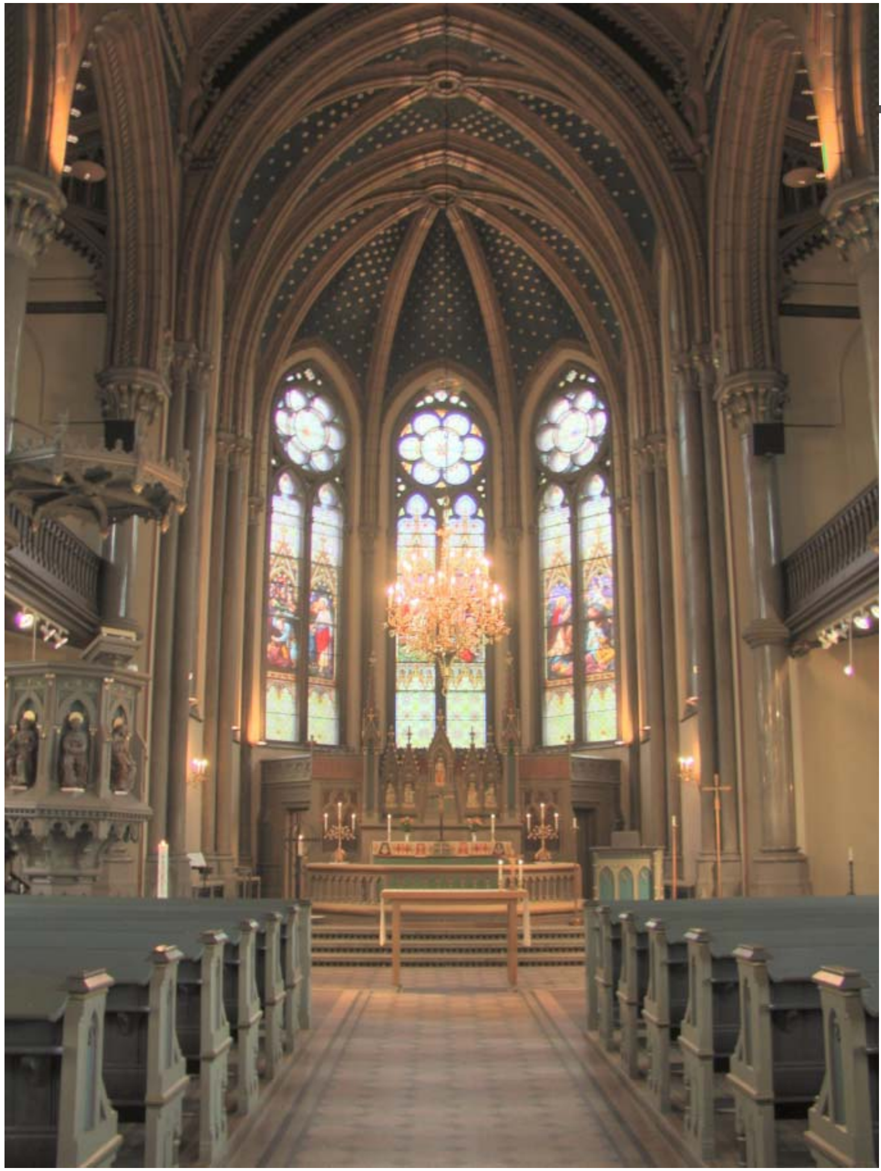


Global operator

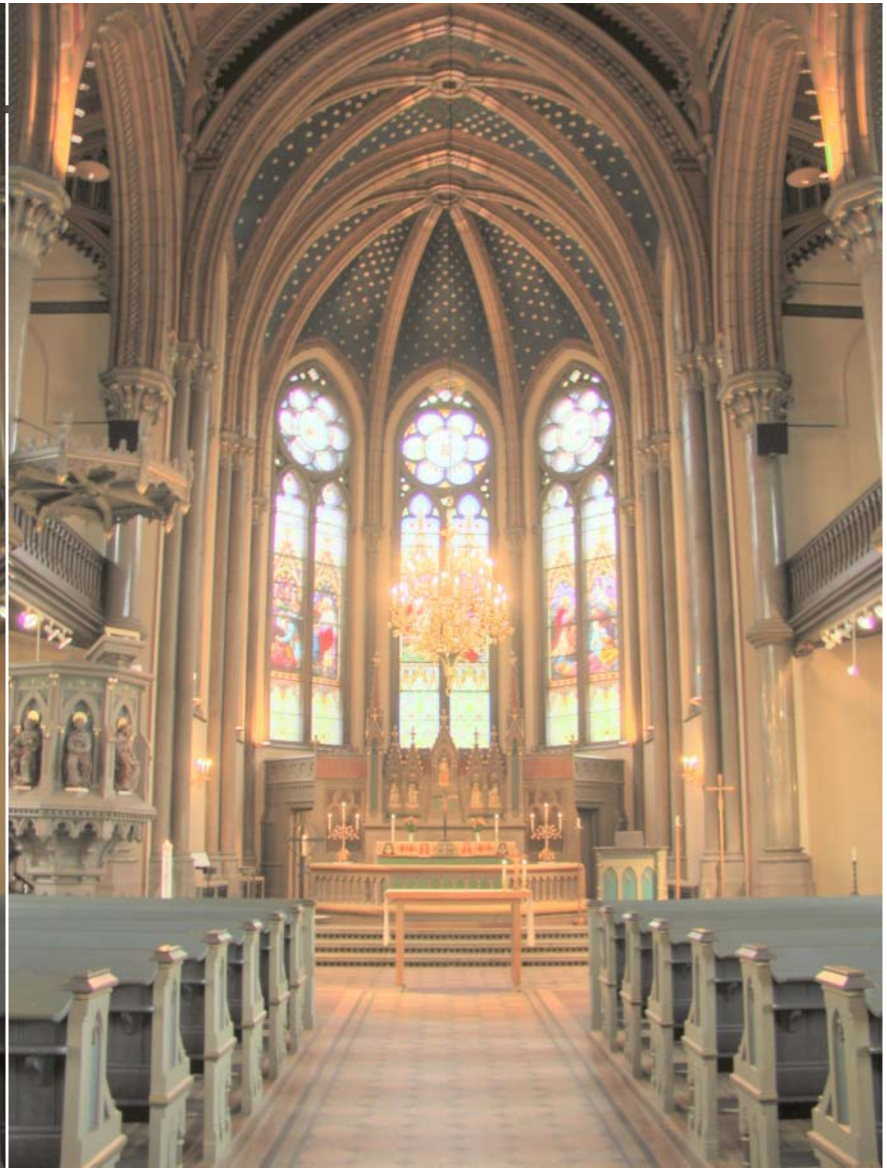
It seldom reaches 1 since the input image does not have infinitely large luminance values.

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





low key (0.18)



high key (0.5)

Dodging and burning (local operators)

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

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

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

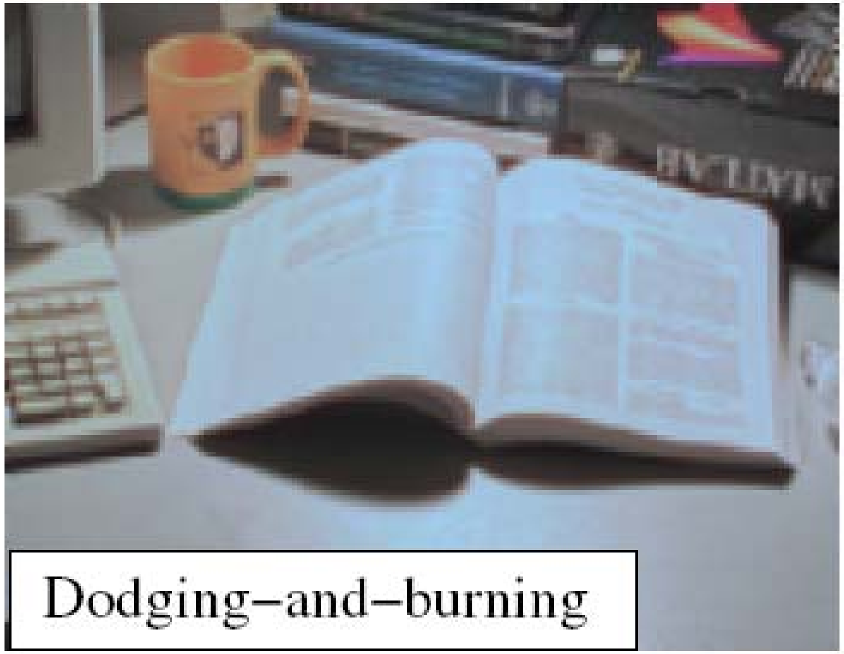
$$s_{\max} : \left| V_{s_{\max}}(\mathrm{x},\mathrm{y}) \right| < \varepsilon$$

Dodging and burning (local operators)

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

- 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

Dodging and burning



Frequency domain

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

image = illuminance * 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

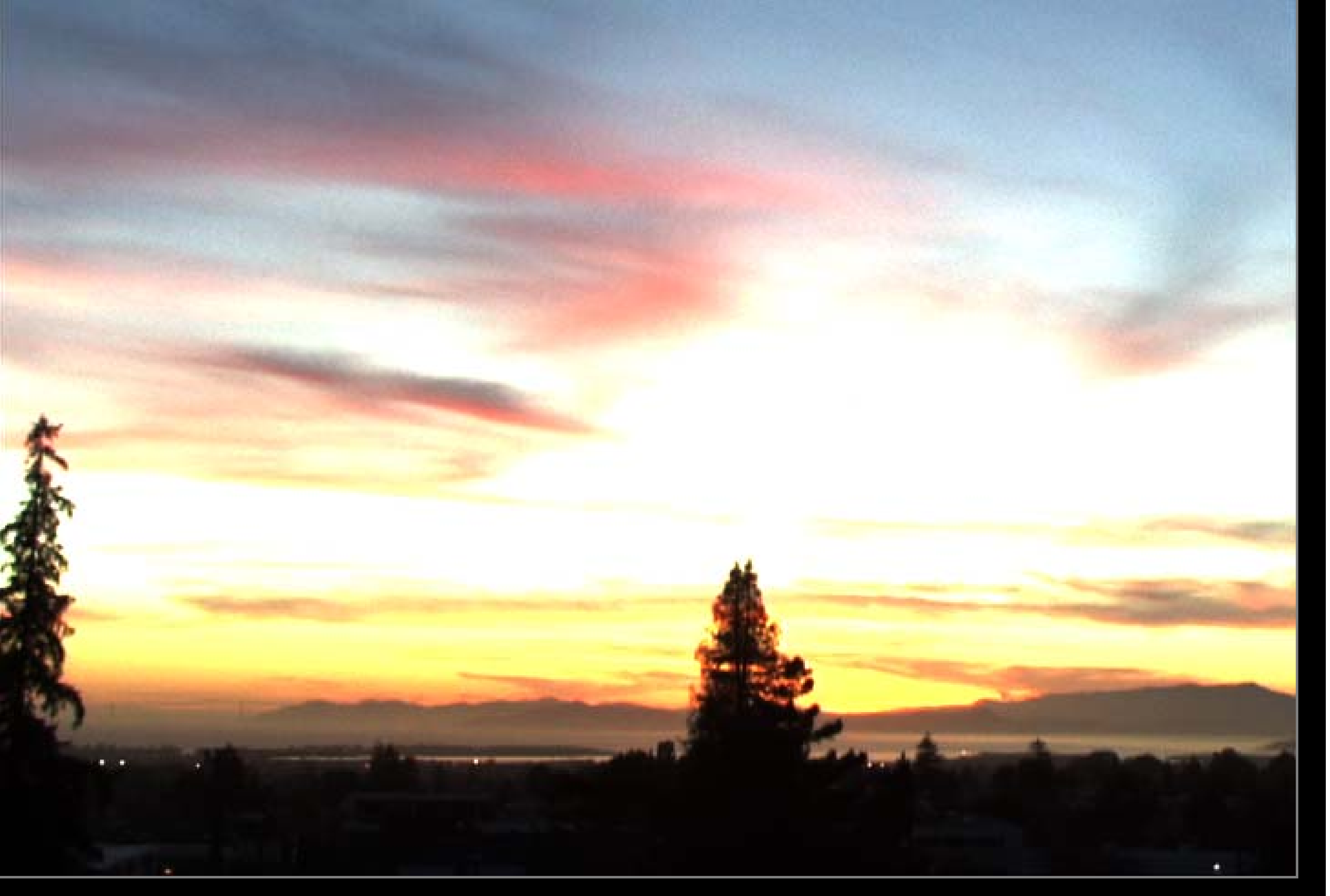
Fast Bilateral Filtering for the Display of High-Dynamic-Range Images

Frédo Durand & Julie Dorsey

SIGGRAPH 2002

A typical photo

- Sun is overexposed
- Foreground is underexposed

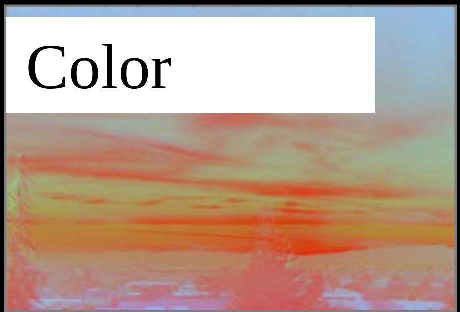


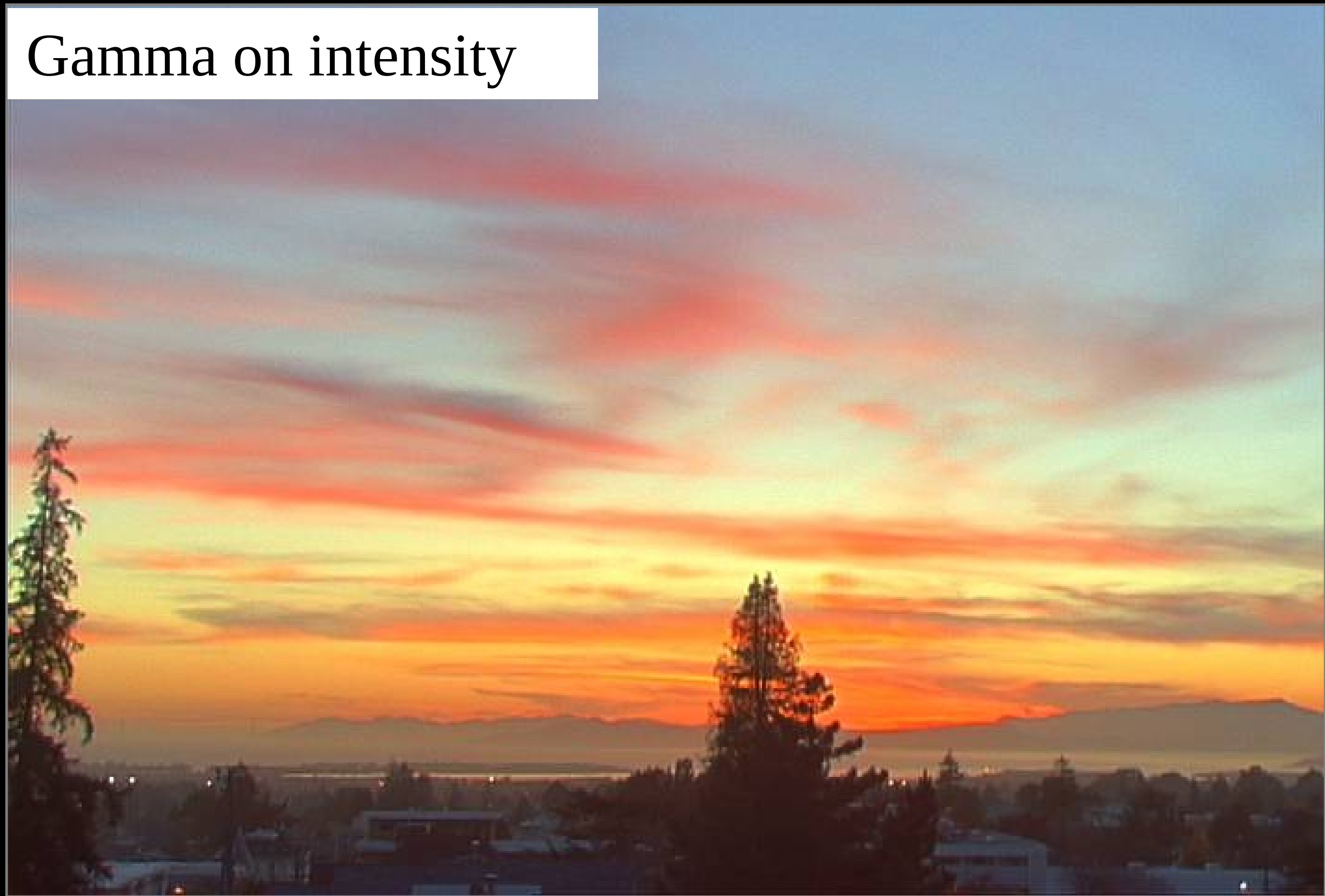
Gamma compression

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

Gamma compression on intensity

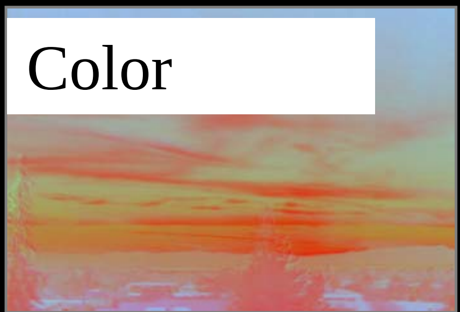
- Colors are OK, but details (intensity high-frequency) are blurred





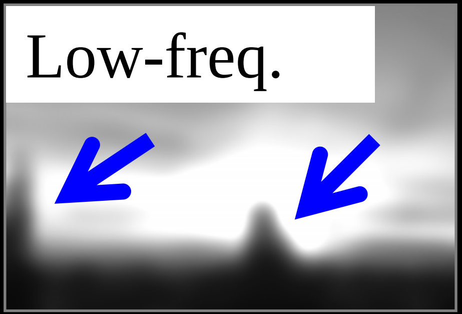
Chiu et al. 1993

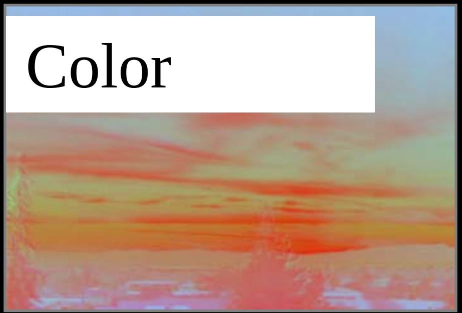
- Reduce contrast of low-frequencies
- Keep high frequencies

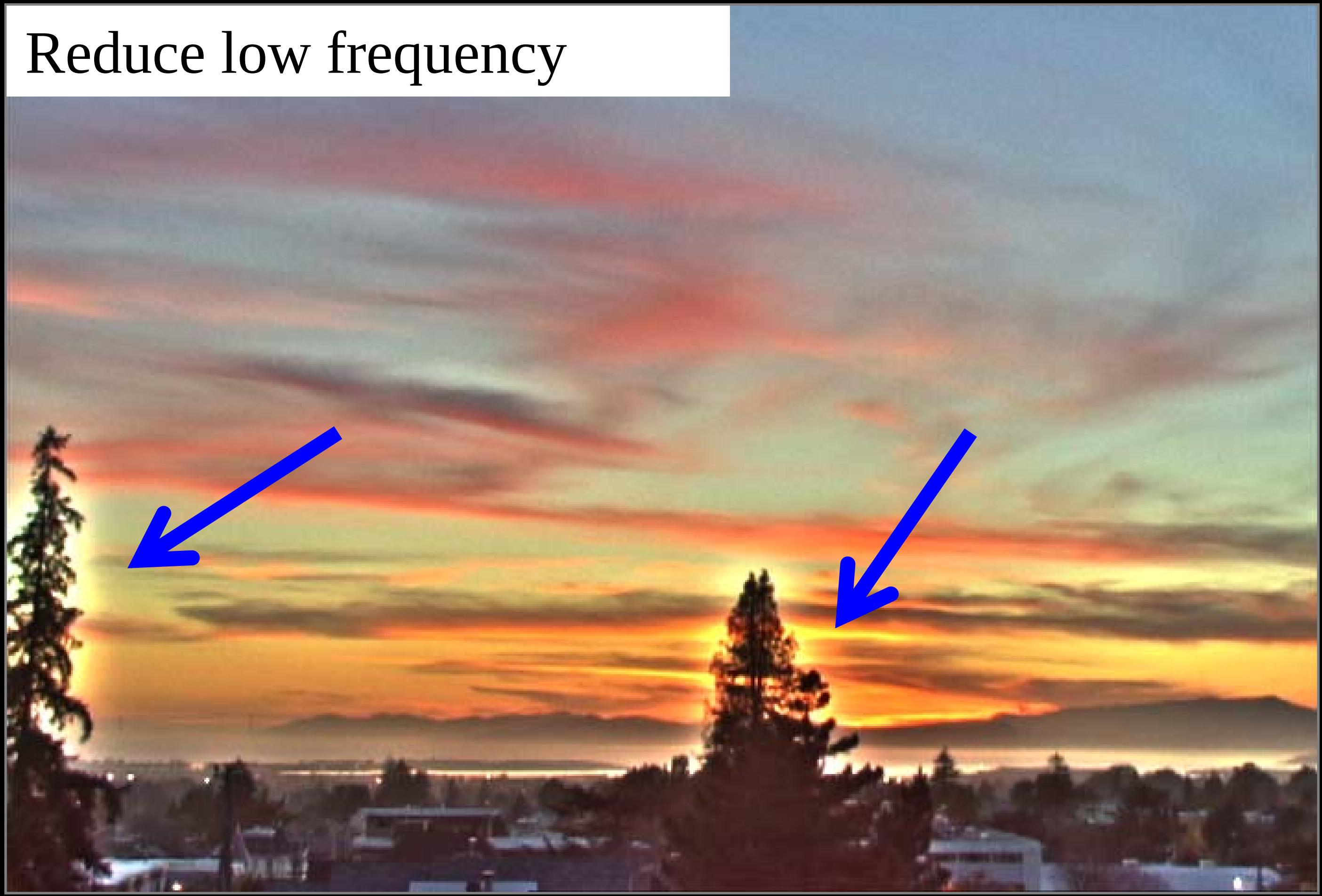


The halo nightmare

- For strong edges
- Because they contain high frequency

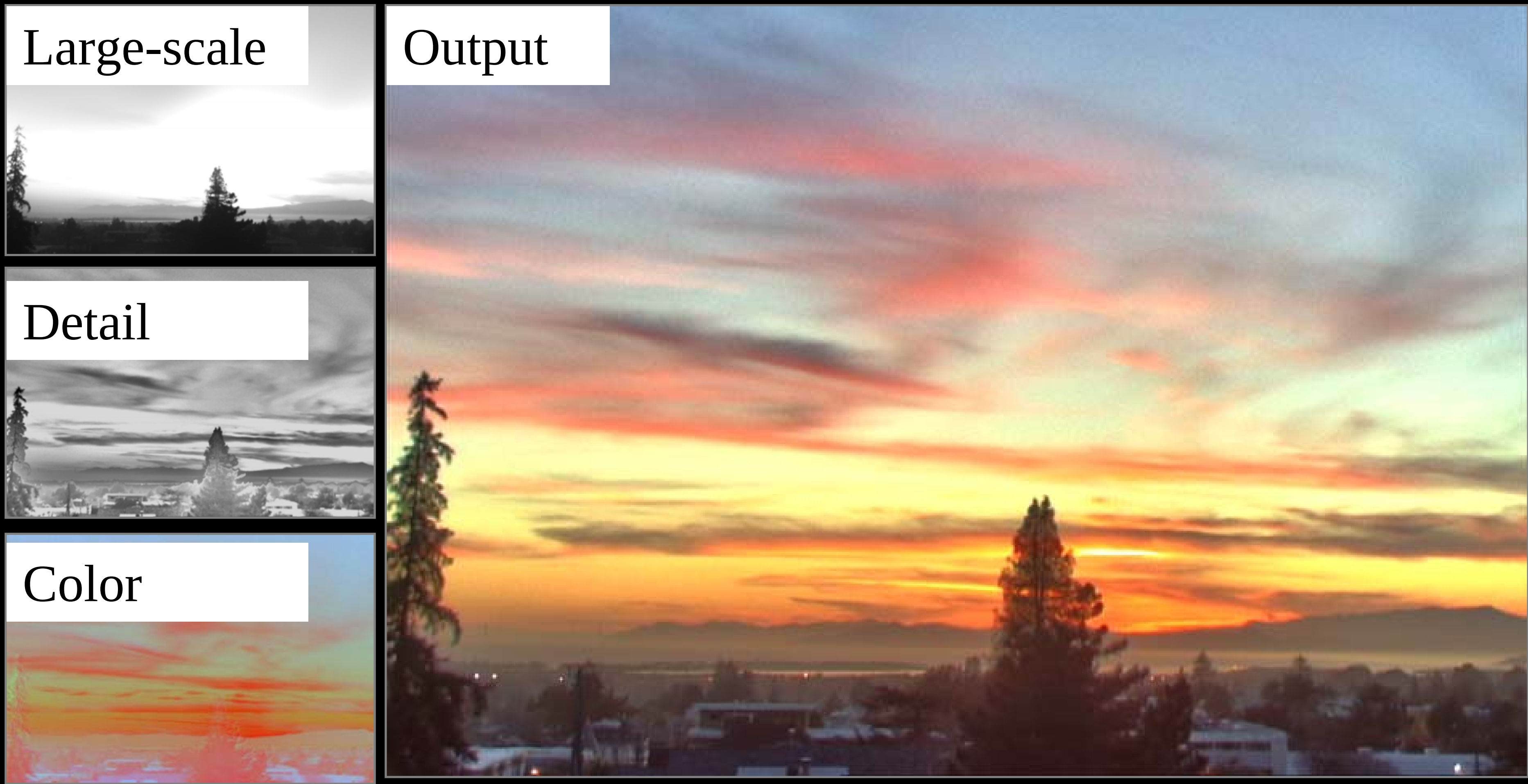






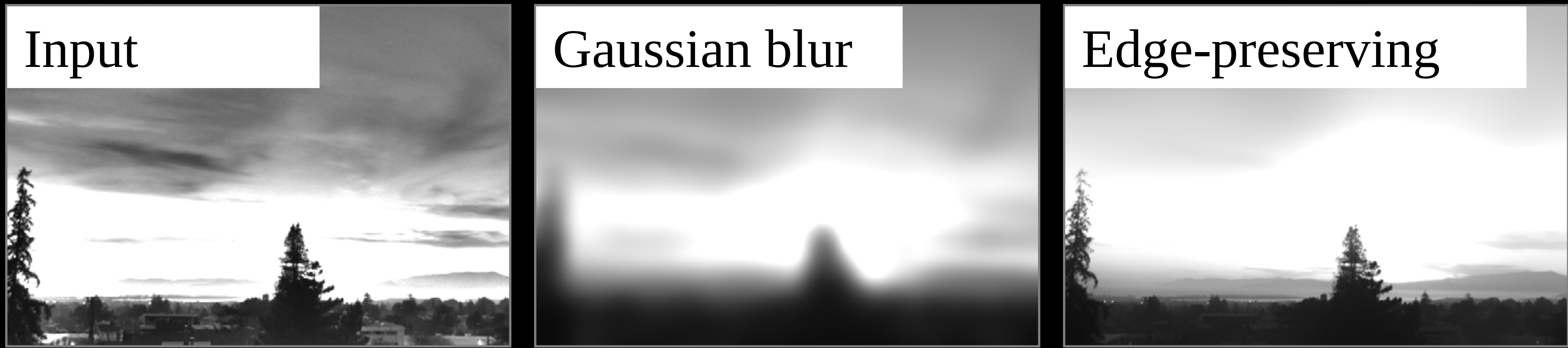
Durand and Dorsey

- Do not blur across edges
- Non-linear filtering



Edge-preserving filtering

- Blur, but not across edges

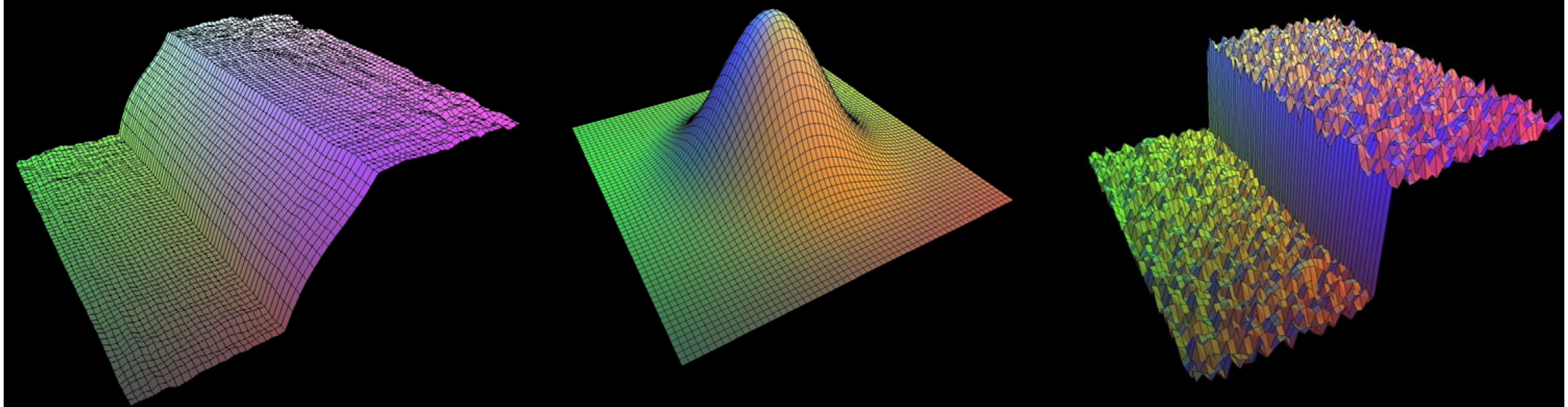


- Anisotropic diffusion [Perona & Malik 90]
  - Blurring as heat flow
  - LCIS [Tumblin & Turk]
- Bilateral filtering [Tomasi & Manduci, 98]

Start with Gaussian filtering

- Here, input is a step function + noise

$$J = f \otimes I$$

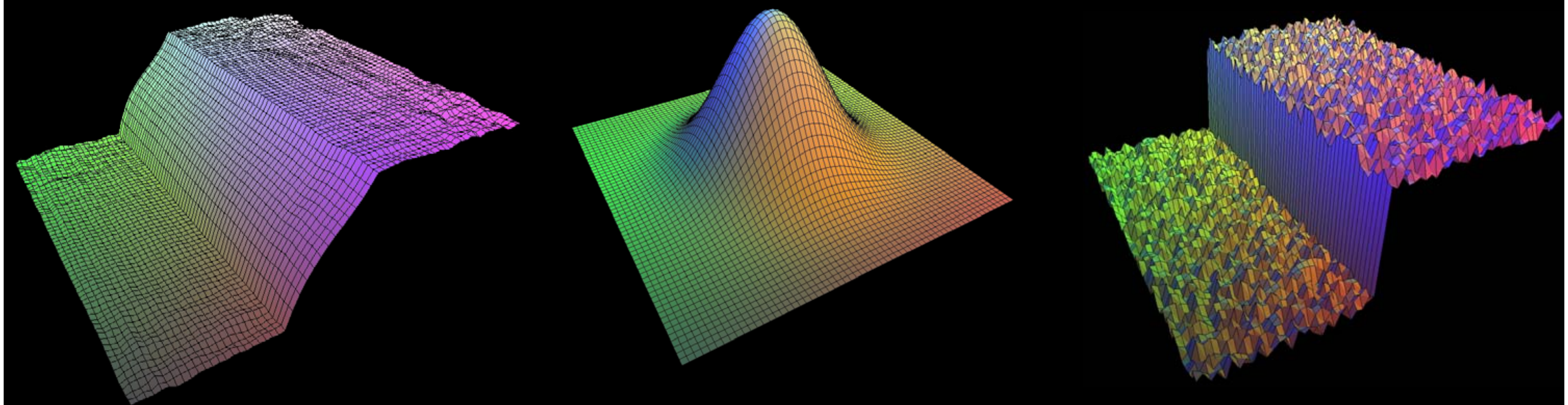


output ← input

Start with Gaussian filtering

- Spatial Gaussian f

$$J = f \otimes I$$

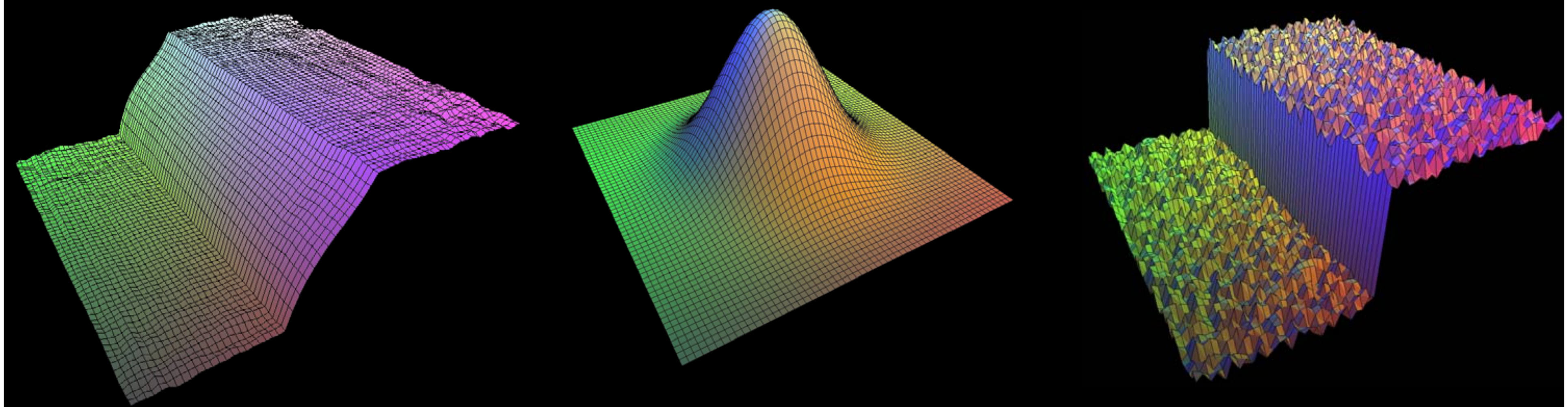


output ⟵ input

Start with Gaussian filtering

- Output is blurred

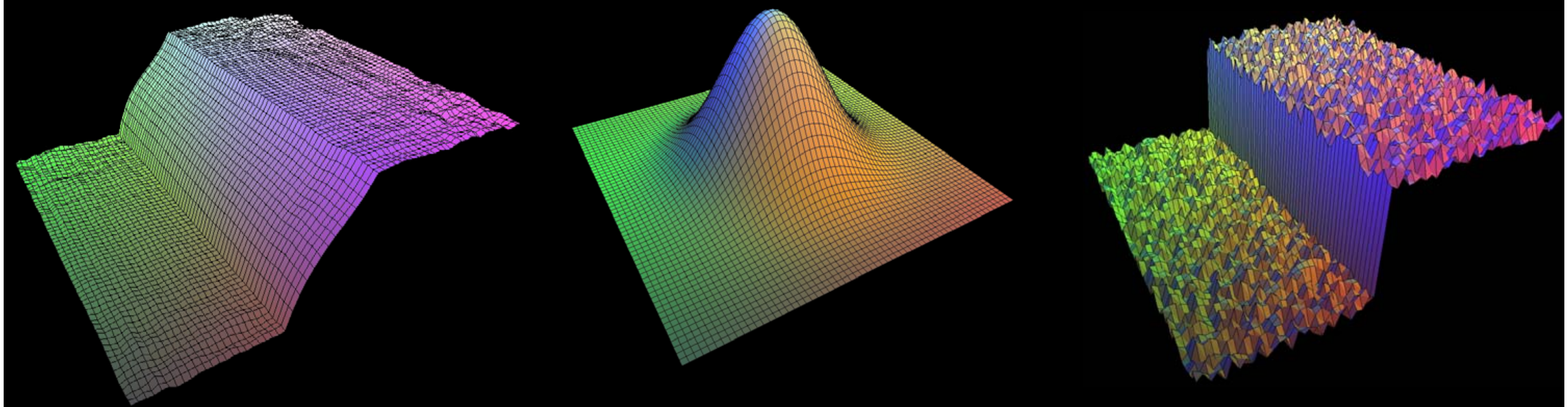
$$J = f \otimes I$$



output ⟵ input

Gaussian filter as weighted average

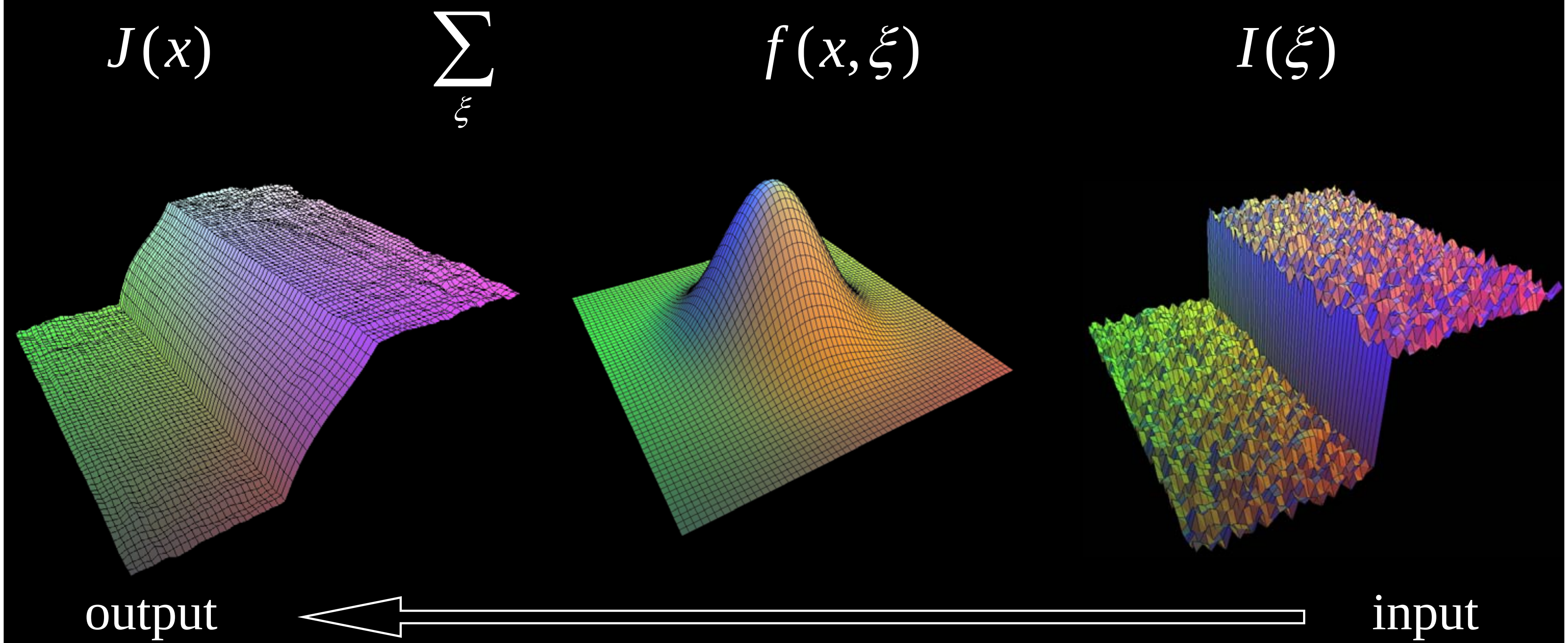
$$J(x) \quad \sum_{\xi} \quad f(x,\xi) \quad I(\xi)$$



output ⟵ input

The problem of edges

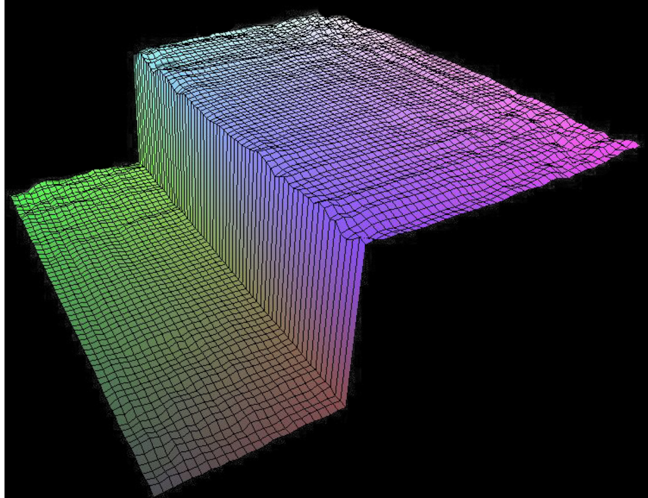
- Here, $I(\xi)$ "pollutes" our estimate J(x)
- It is too different

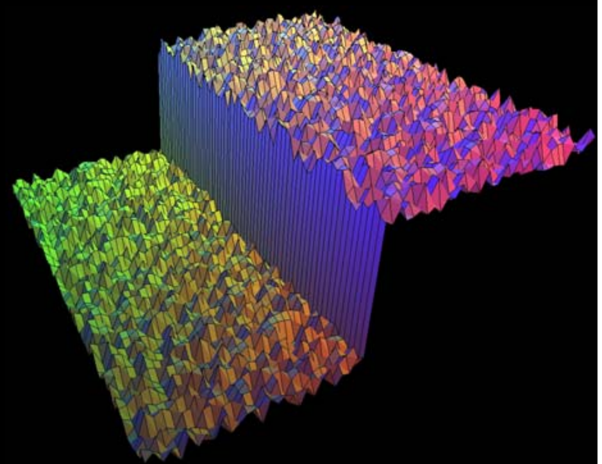


Principle of Bilateral filtering

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

$$J(x) = \frac{1}{k(x)} \sum_{\xi} f(x,\xi) \quad g(I(\xi) - I(x)) \quad I(\xi)$$



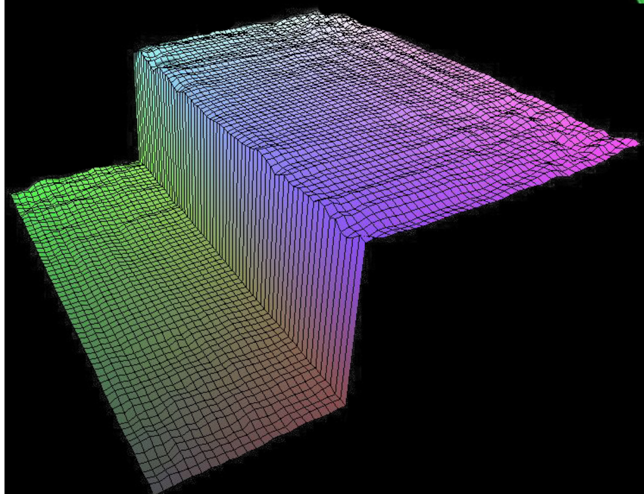


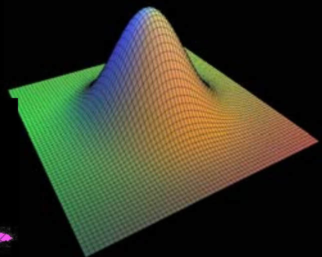
output input

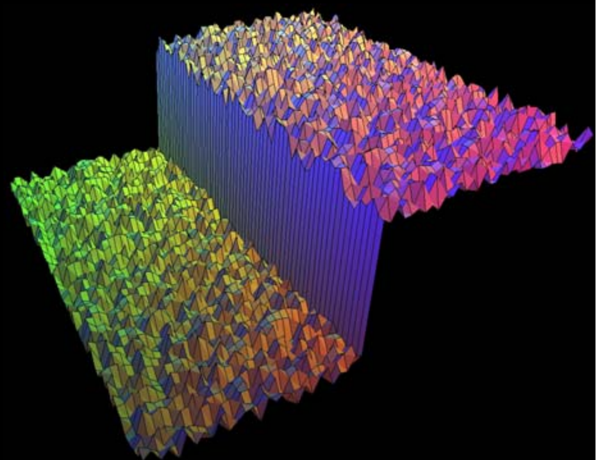
Bilateral filtering

- [Tomasi and Manduchi 1998]
- Spatial Gaussian f

$$J(x) \quad \frac{1}{k(x)} \quad f(x,\xi) \quad g(I(\xi)-I(x)) \quad I(\xi)$$





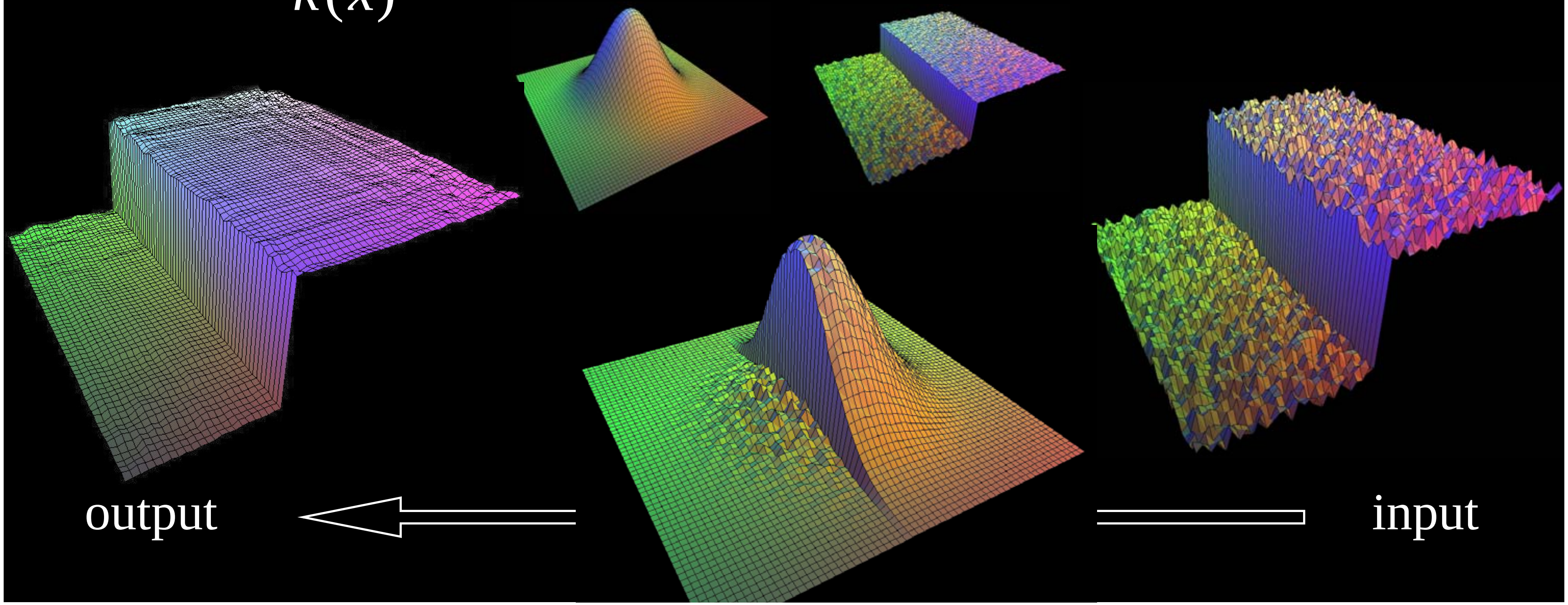


output ⟵ input

Bilateral filtering

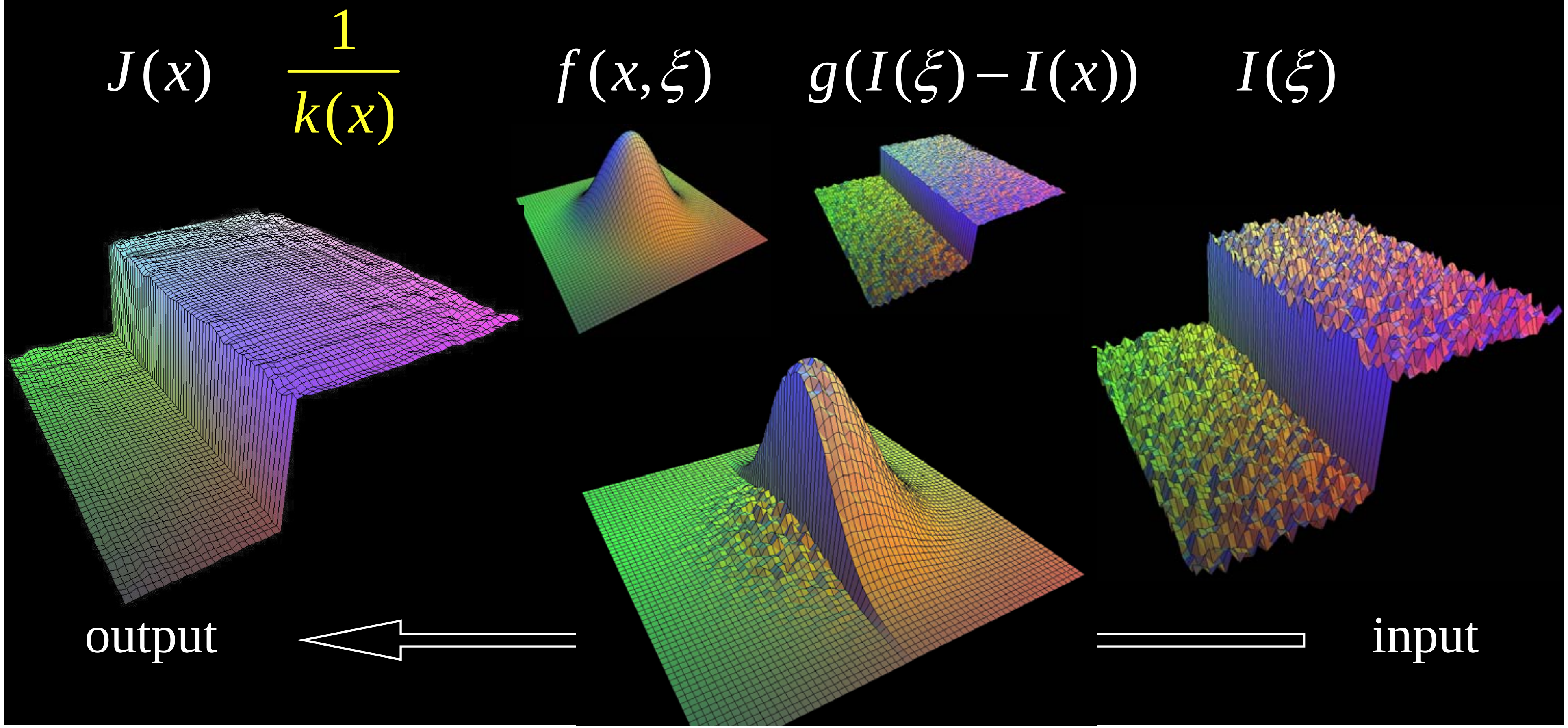
- [Tomasi and Manduchi 1998]
- Spatial Gaussian f
- Gaussian g on the intensity difference

$$J(x) \quad \frac{1}{k(x)} \quad f(x,\xi) \quad g(I(\xi)-I(x)) \quad I(\xi)$$



Normalization factor

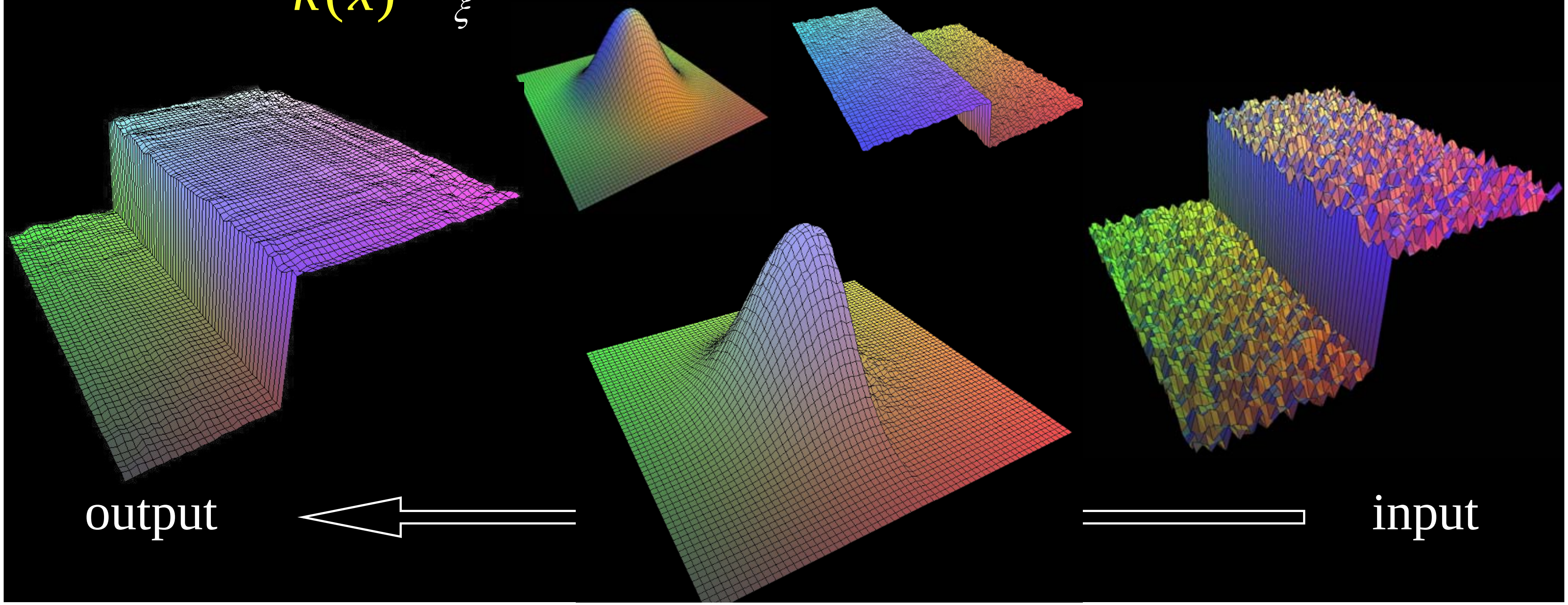
- [Tomasi and Manduchi 1998]
- k(x)= $\sum_{\xi} \boxed{f(x,\xi) \quad g(I(\xi)-I(x))}$



Bilateral filtering is non-linear

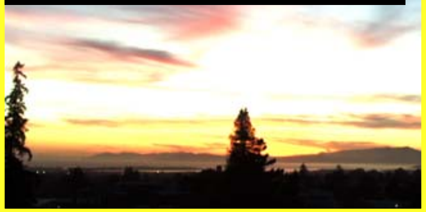
- [Tomasi and Manduchi 1998]
- The weights are different for each output pixel

$$J(x) \quad \frac{1}{k(x)} \sum_{\xi} \quad f(x,\xi) \quad g(I(\xi)-I(x)) \quad I(\xi)$$



Contrast reduction

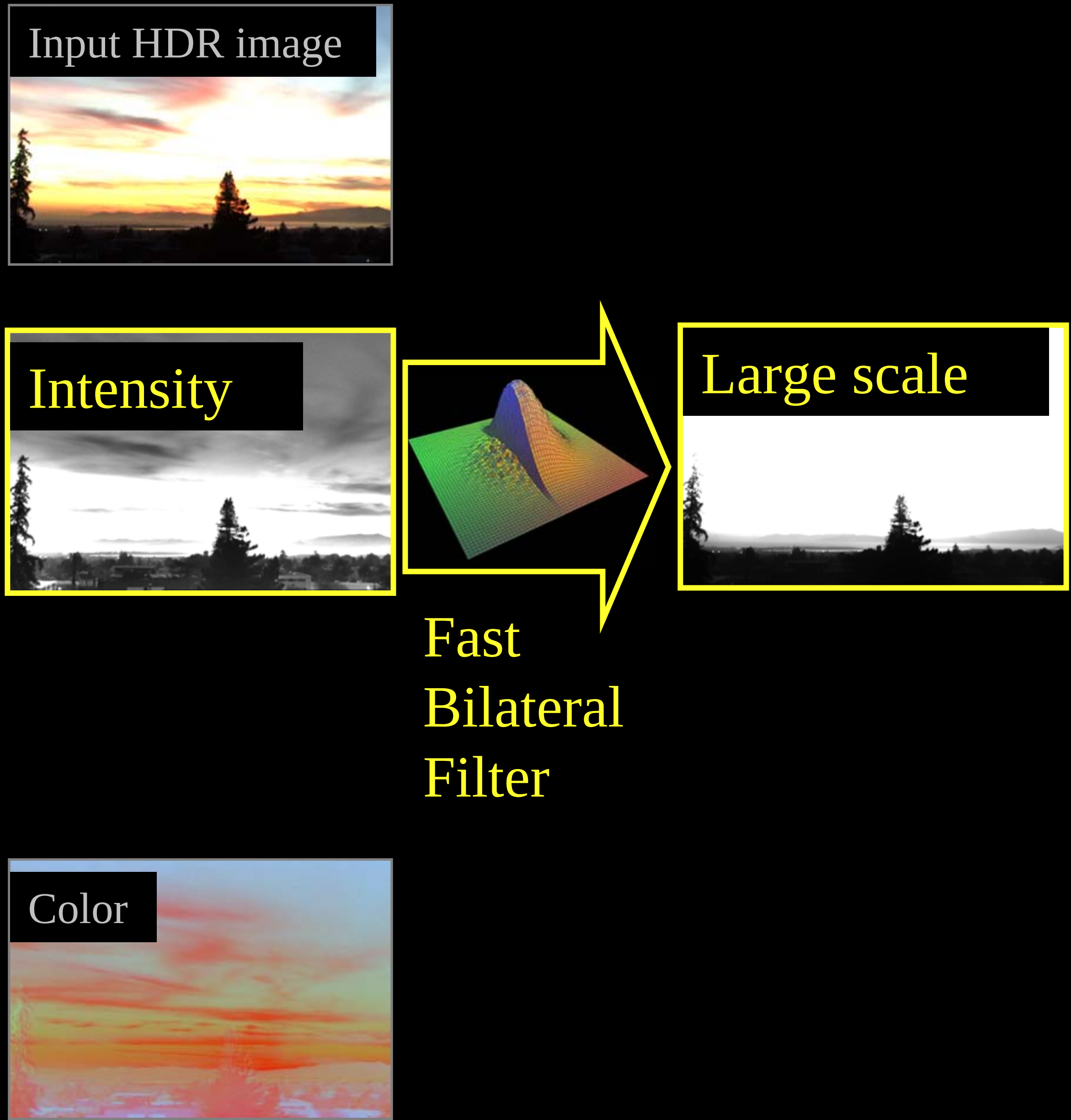
Input HDR image



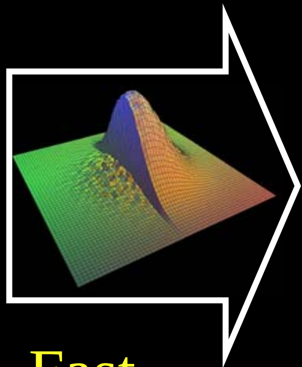
Contrast
too high!

Contrast reduction

Contrast reduction

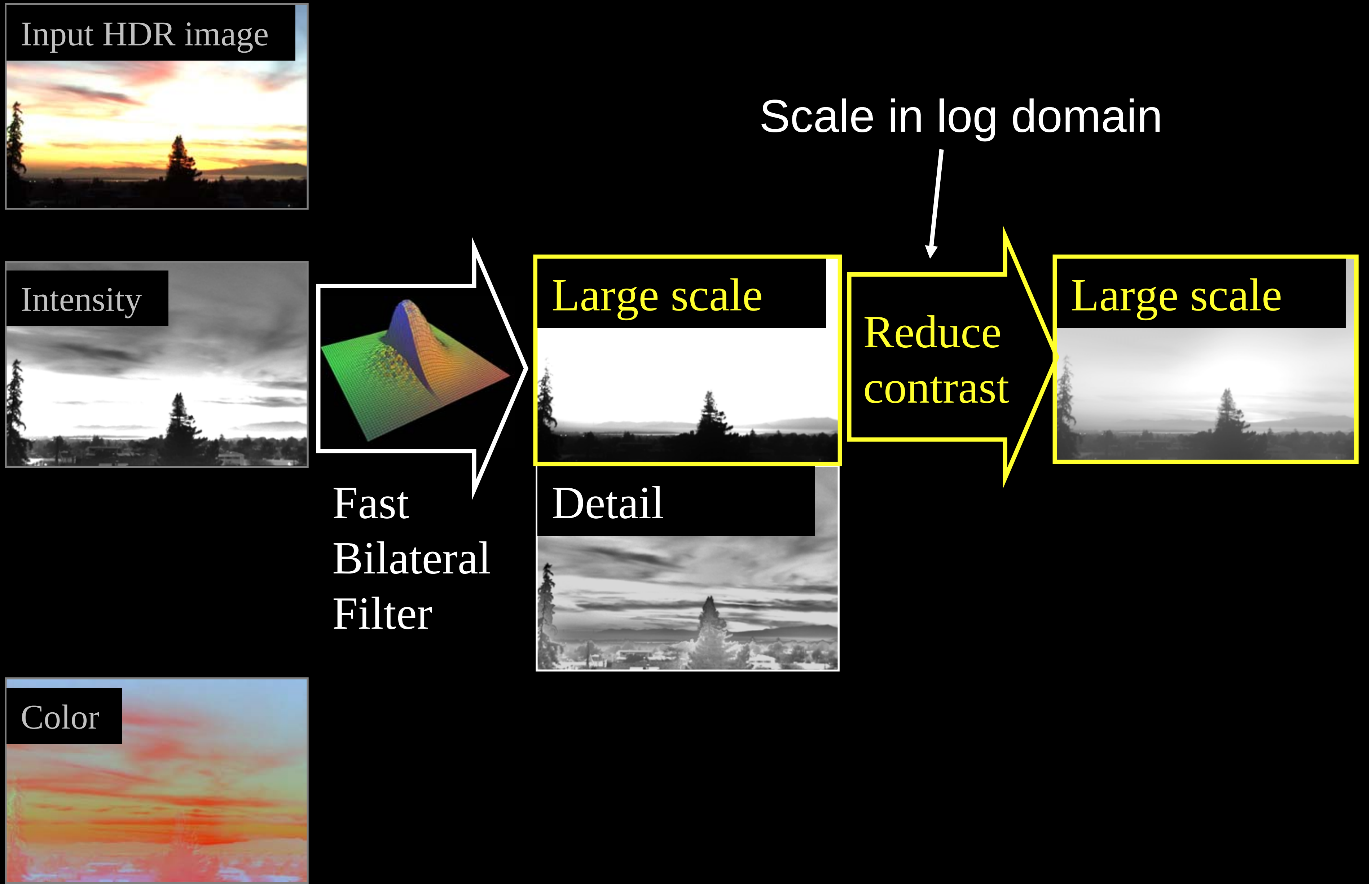


Contrast reduction

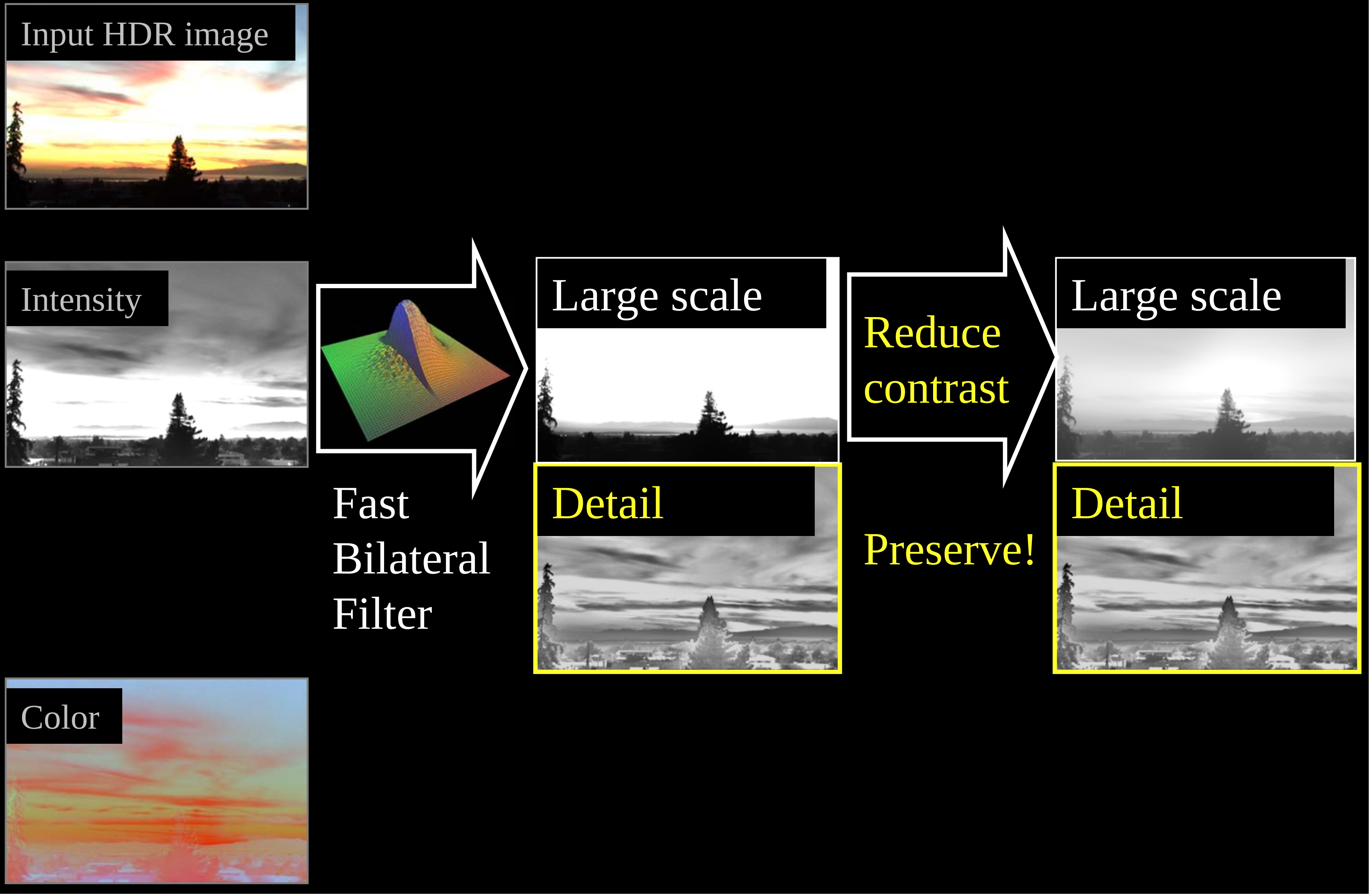


Fast
Bilateral
Filter

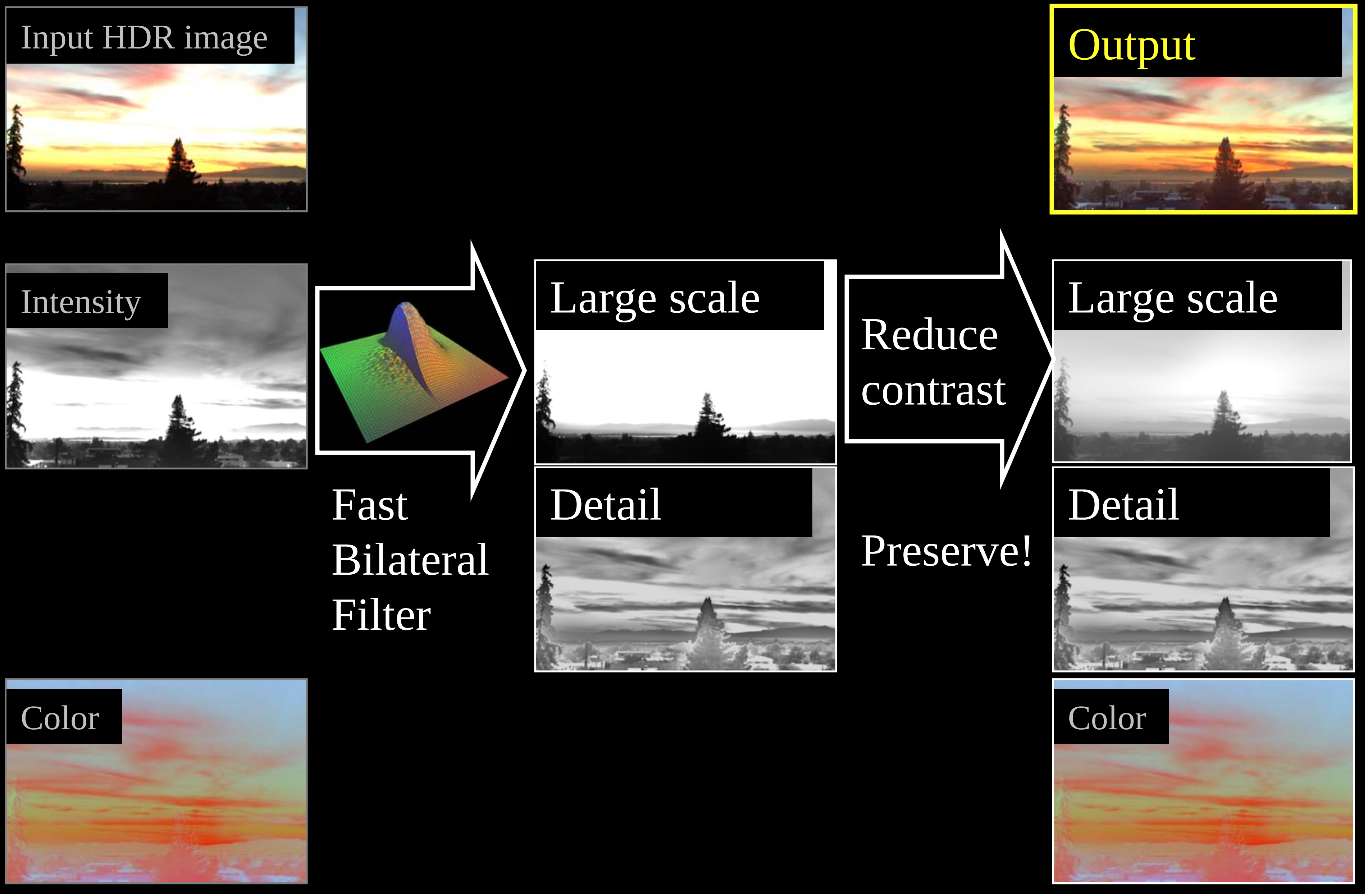
Contrast reduction



Contrast reduction

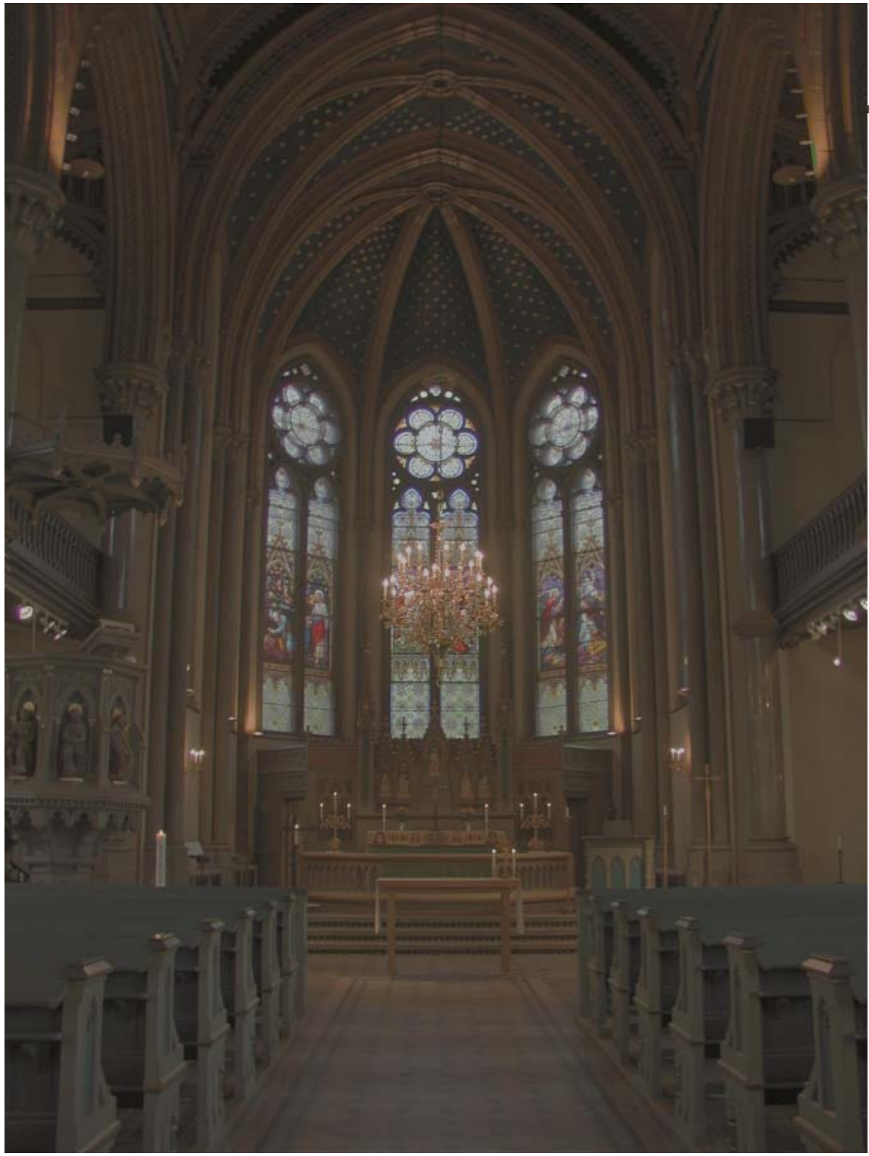


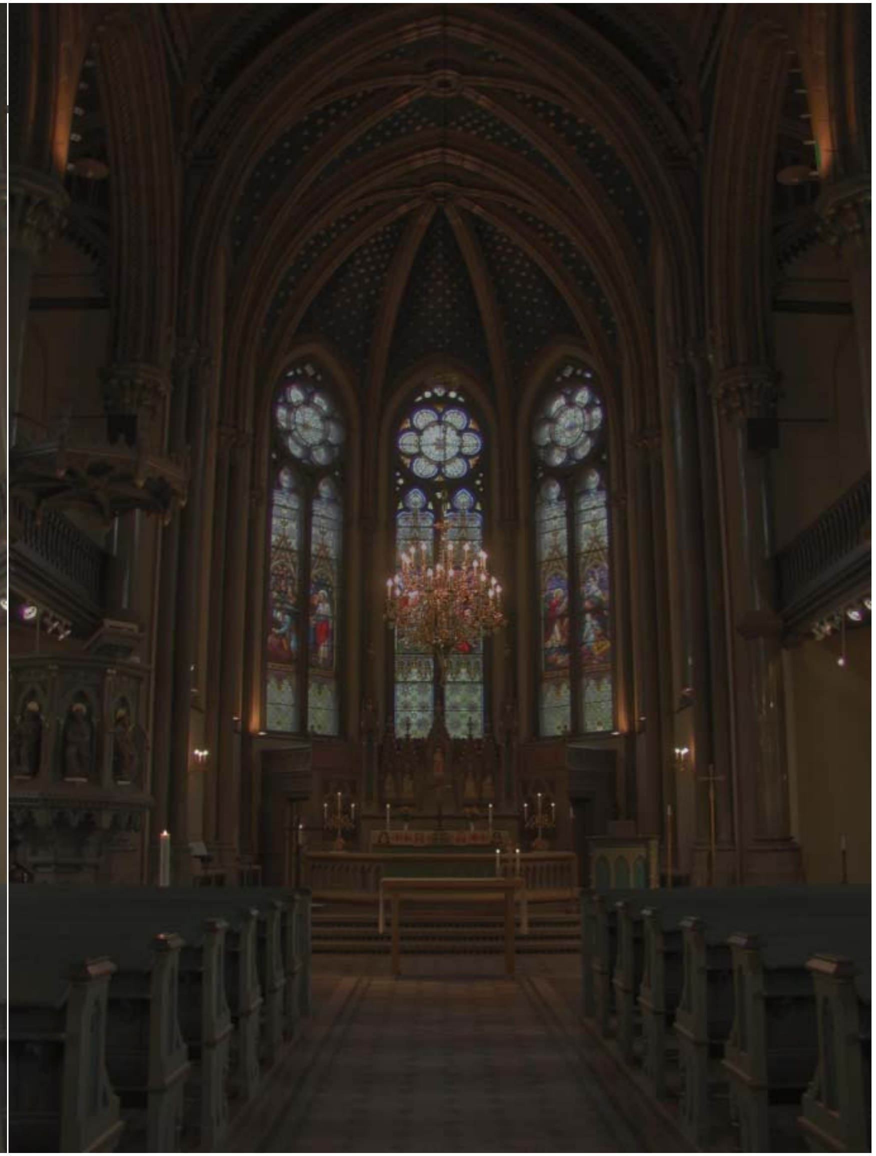
Contrast reduction



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 photogrphy.





Oppenheim | bilateral

Gradient Domain High Dynamic Range Compression

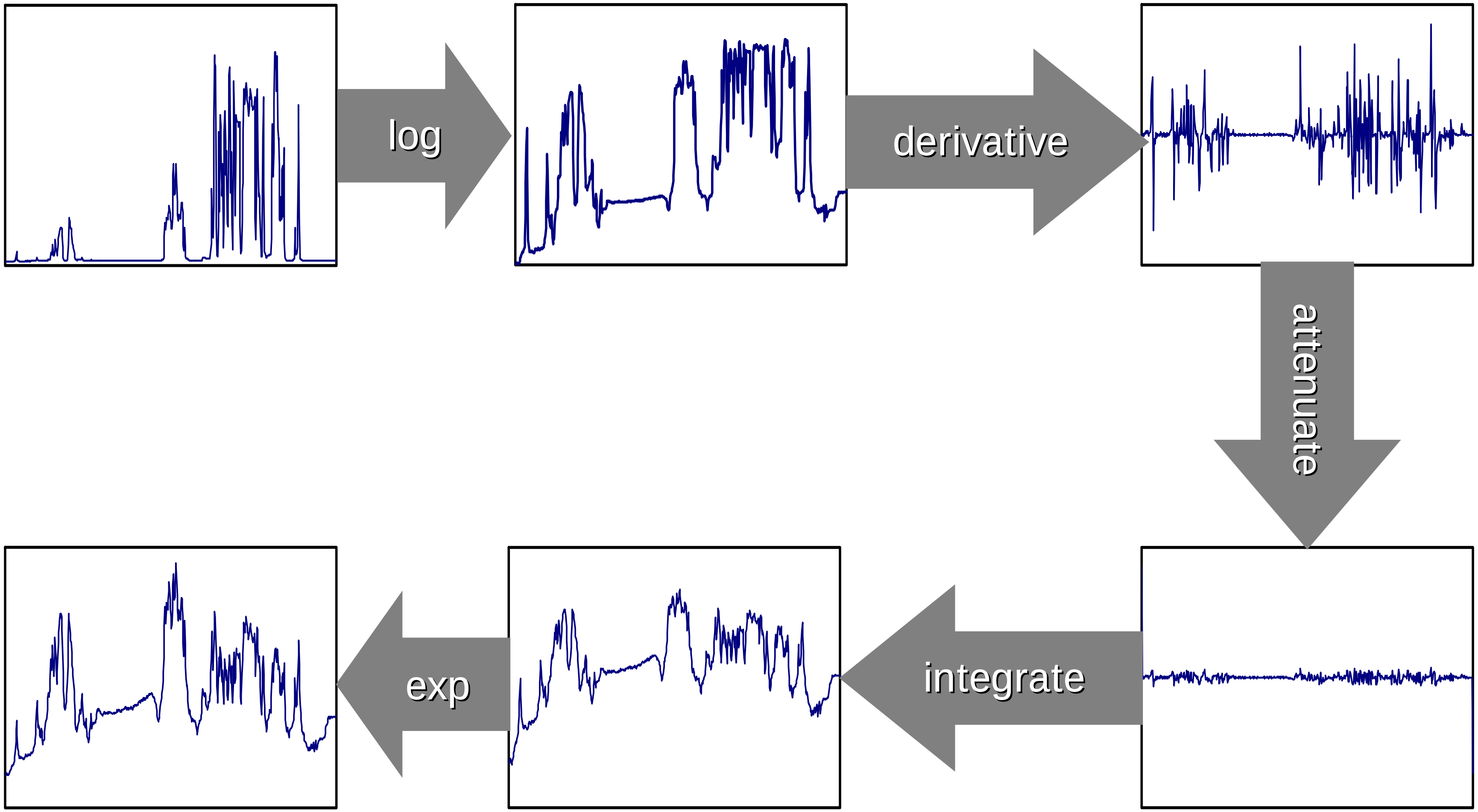
Raanan Fattal   Dani Lischinski   Michael Werman

SIGGRAPH 2002

Log domain

- Logorithm 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



The method in 2D

- Given: a log-luminance image $H(x,y)$
- Compute an *attenuation map* $\Phi\left(\|\nabla H\|\right)$
- Compute an attenuated gradient field $\boldsymbol{G}$:

$$G(x,y) = \nabla H(x,y) \cdot \Phi\left(\|\nabla H\|\right)$$

- Problem: $\boldsymbol{G}$ is not integrable!

Solution

- Look for image $\boldsymbol{I}$ with gradient closest to $\boldsymbol{G}$ in the least squares sense.
- $\boldsymbol{I}$ minimizes the integral: $\iint F(\nabla I, G)\,dxdy$

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

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

Poisson equation

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

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

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

$$\begin{bmatrix} \\ ..\ 1 \ldots\ \ 1\ \text{-}4\ 1 \ldots\ \ 1\ .. \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \mathbf{I} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

Solving Poisson equation

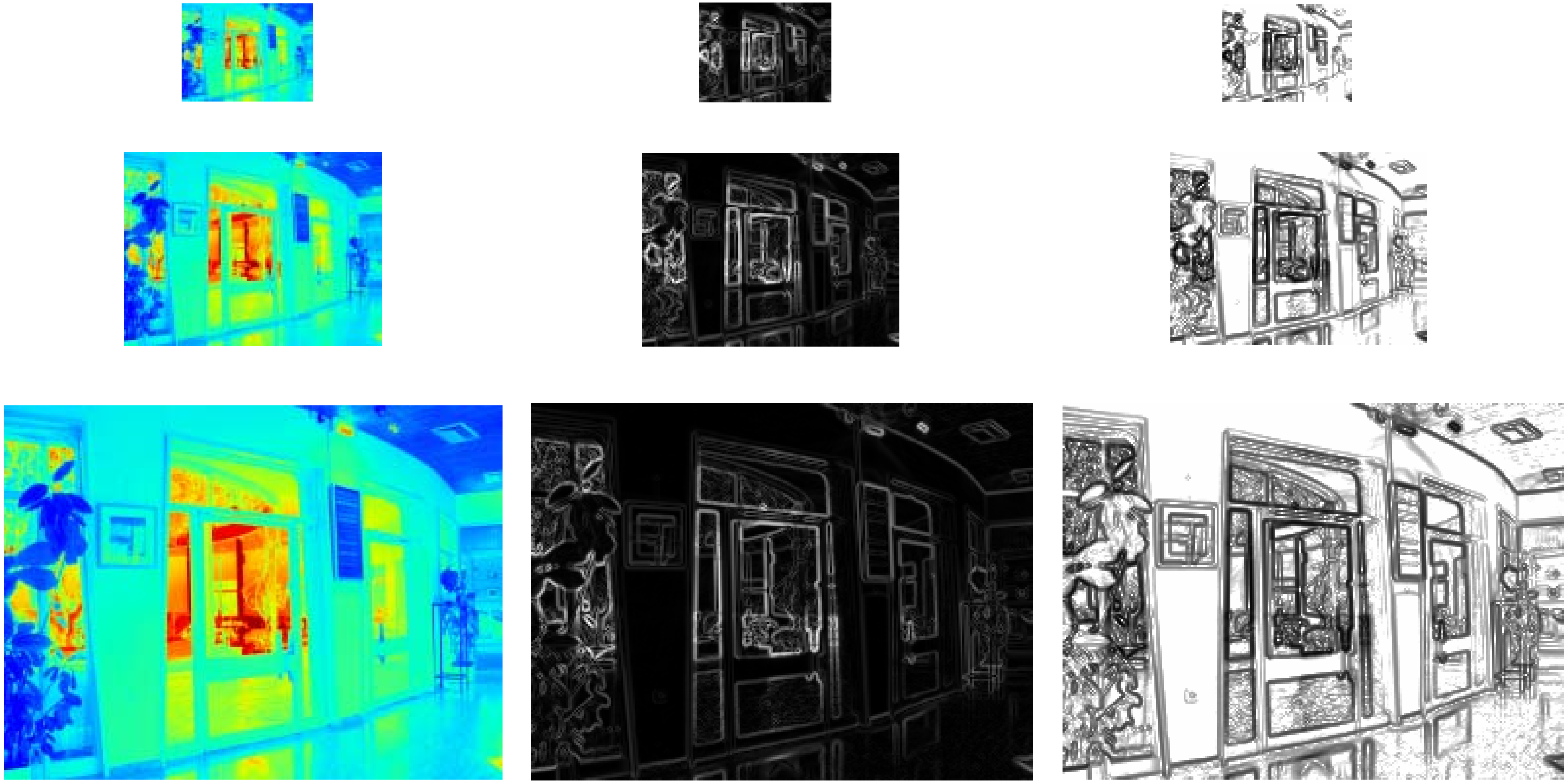
- No analytical solution
- Multigrid method
- Conjugate gradient method

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

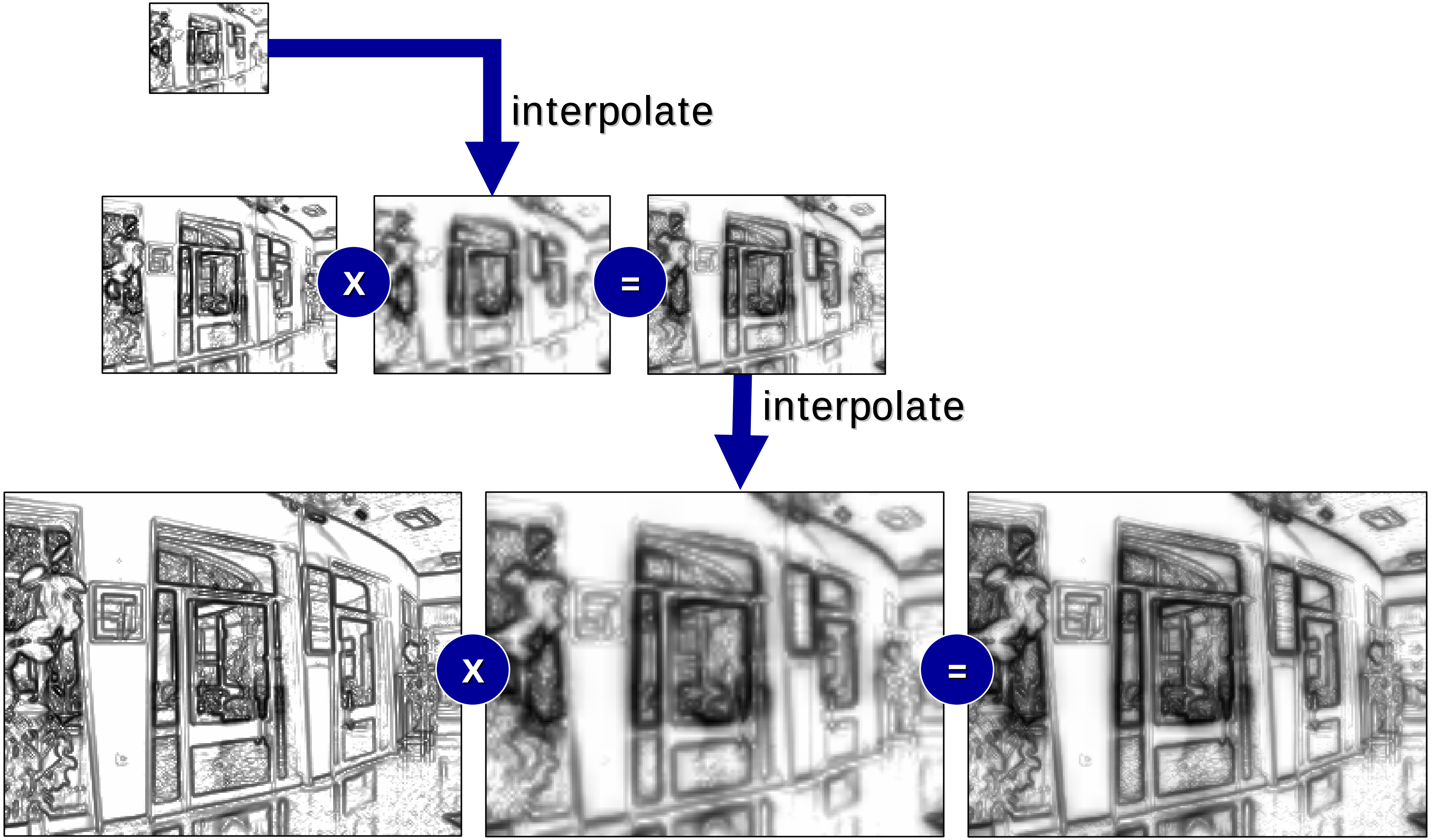
Attenuation

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

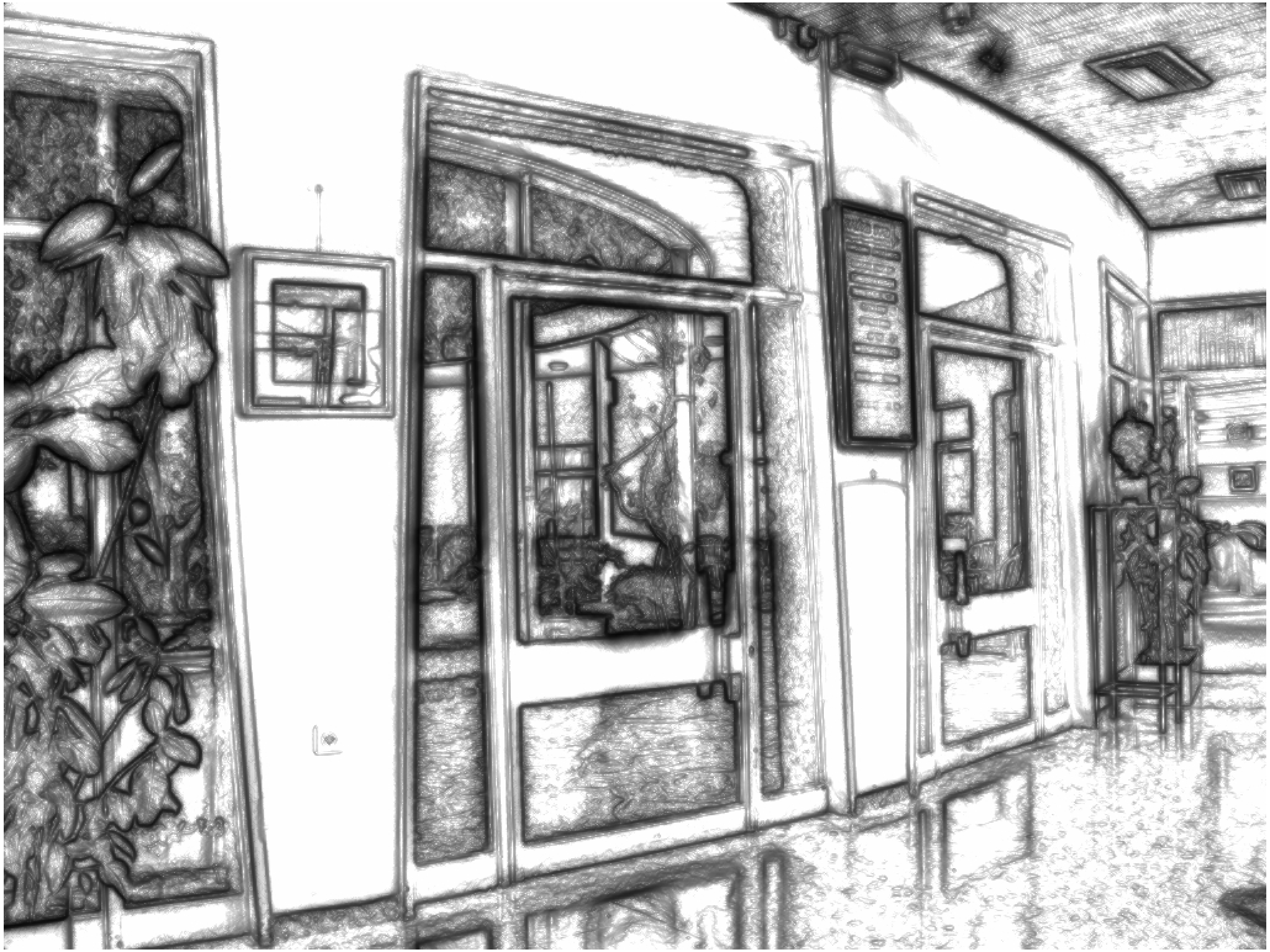


log(Luminance) gradient magnitude attenuation map

Multiscale gradient attenuation

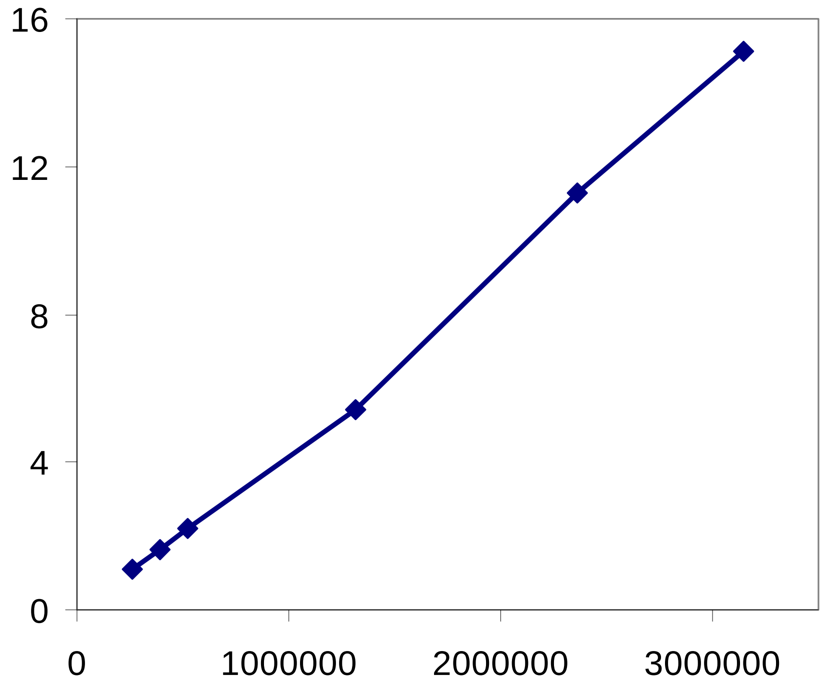


Final gradient attenuation map



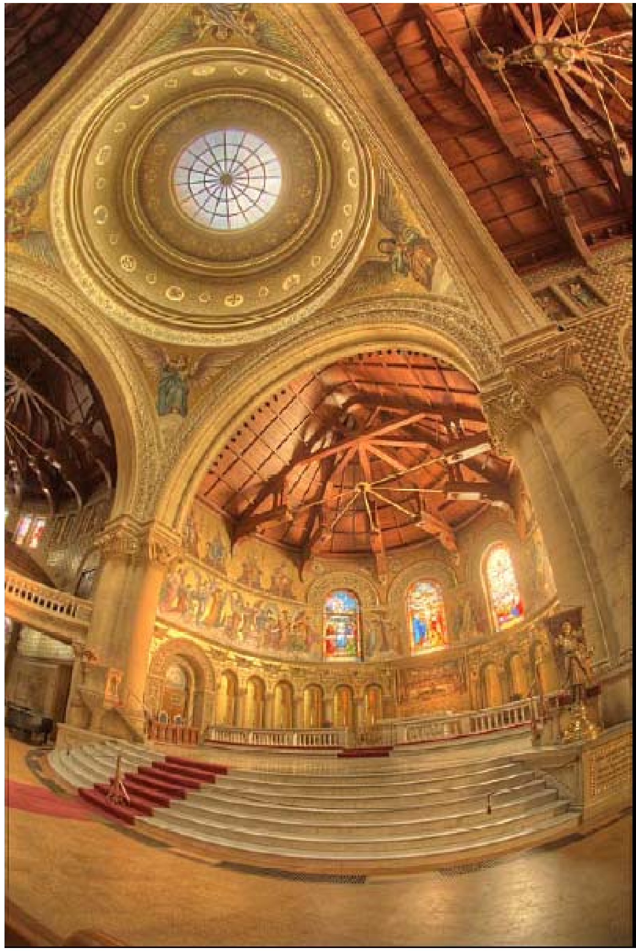
Performance

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



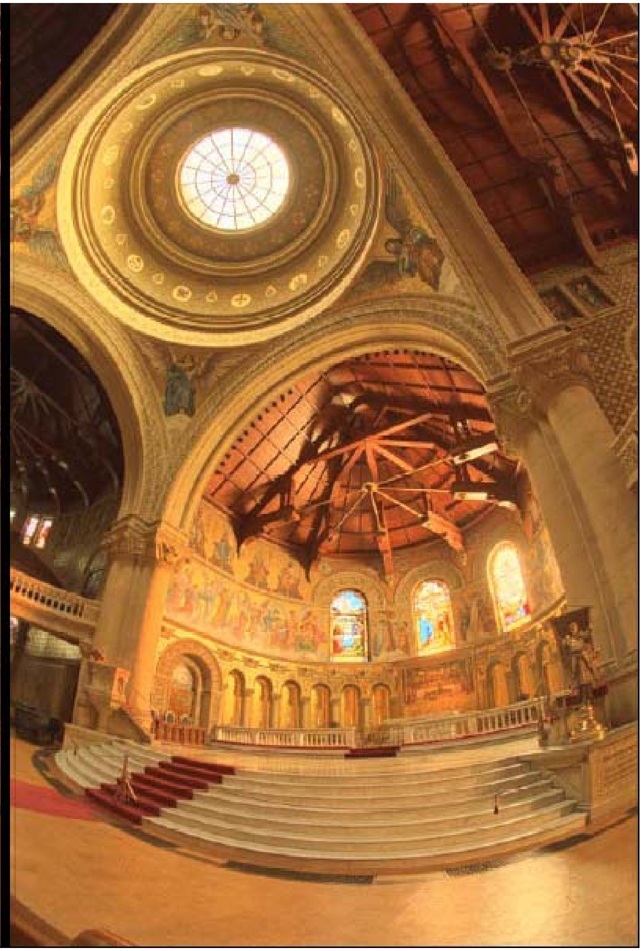
- Can be accelerated using processor-optimized libraries.

Informal comparison



Gradient domain
[Fattal et al.]

Bilateral
[Durand et al.]



Photographic
[Reinhard et al.]

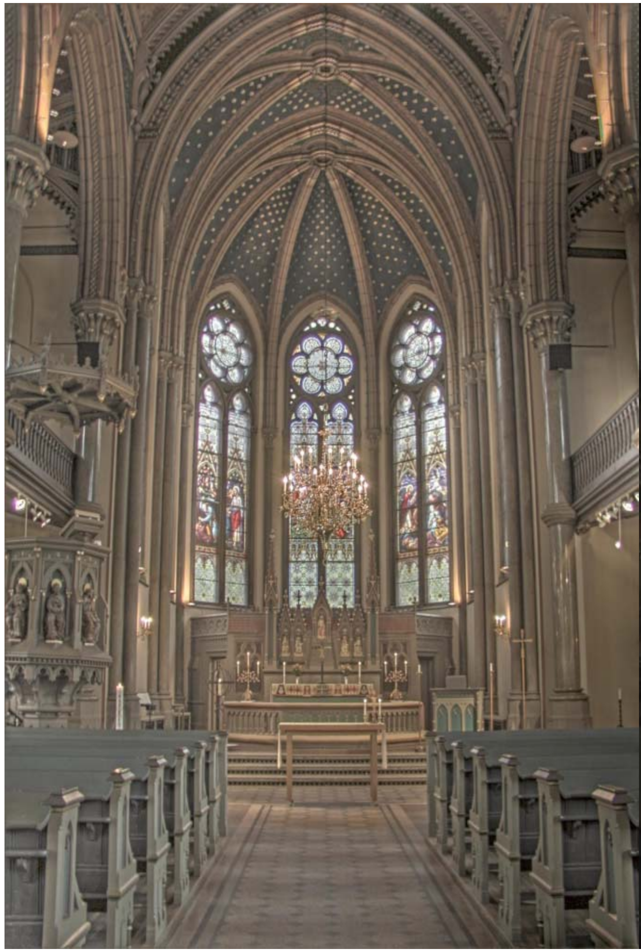
Informal comparison

Gradient domain [Fattal et al.]

Bilateral [Durand et al.]

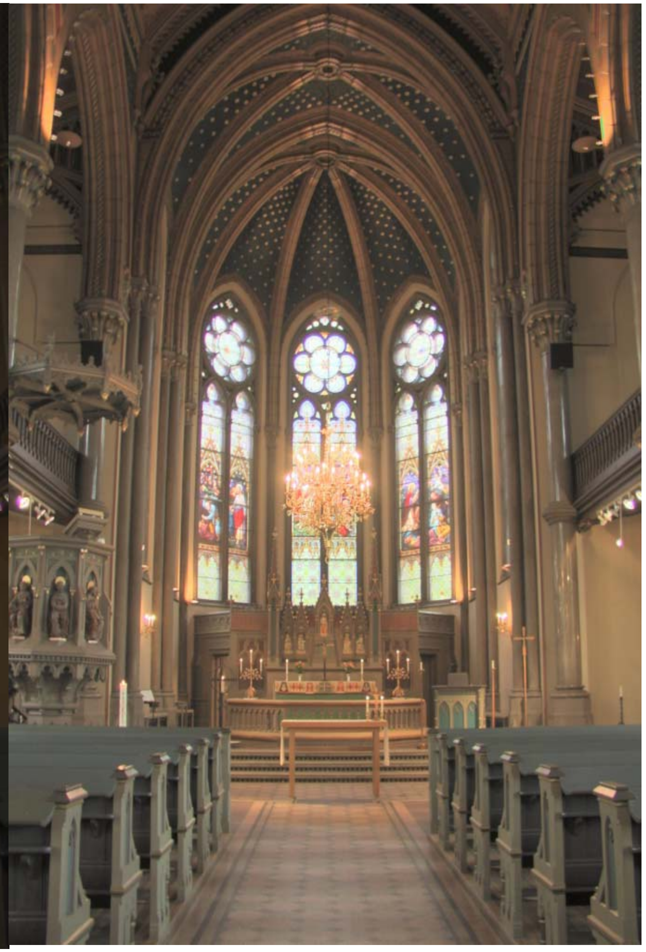
Photographic [Reinhard et al.]

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

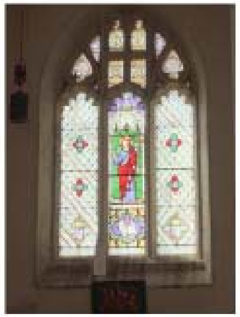

Scene 1


Scene 2

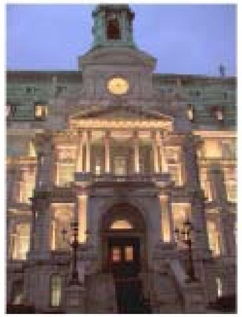

Scene 3


Scene 4

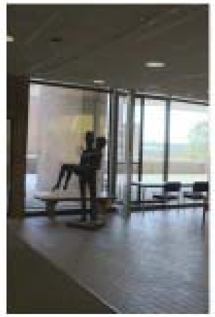

Scene 5


Scene 6

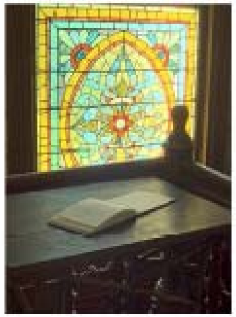

Scene 7


Scene 8


Scene 9


Scene 10


Scene 11

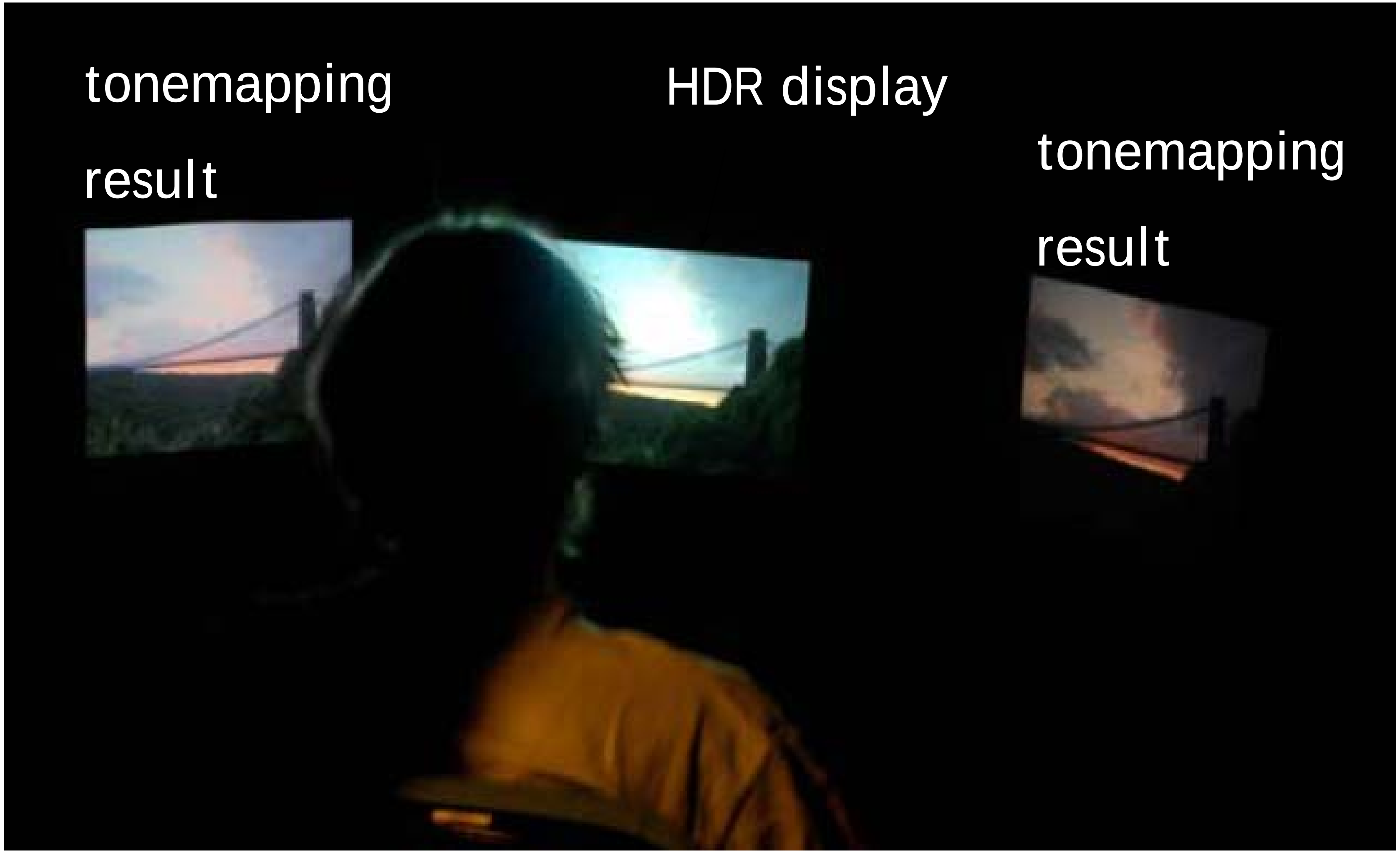

Scene 12


Scene 13


Scene 14


Scene 15

Experiment setting



Preference matrix

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

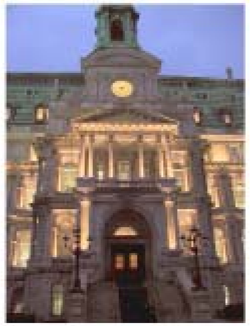
	tmo_1	tmo_2	tmo_3	tmo_4	tmo_5	tmo_6	Score
tmo_1	-	1	0	0	1	1	3
tmo_2	0	-	0	1	1	0	2
tmo_3	1	1	-	1	1	1	5
tmo_4	1	0	0	-	0	0	1
tmo_5	0	0	0	1	-	1	2
tmo_6	0	1	0	1	0	-	2

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 droped.

Overall similarity

- Scene 8 

	P	H	B	L	I	A	Total
P	-	24	46	42	10	32	154
H	24	-	44	32	8	12	120
B	2	4	-	8	2	4	20
L	6	16	40	-	4	12	78
I	38	40	46	44	-	38	206
A	16	36	44	36	10	-	142

Summary

Overall Similarity: Color

I	*P*	*H*	*A*	*L*	*B*
3712	3402	2994	2852	1902	1696

Bright Detail

I	*A*	*P*	*H*	*B*	*L*
823	688	569	549	474	347

Dark Detail

P	*A*	*I*	*L*	*H*	*B*
815	793	583	491	485	283

Not settled down 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

- Raanan Fattal, Dani Lischinski, Michael Werman, Gradient Domain High Dynamic Range Compression, SIGGRAPH 2002.
- Fredo Durand, Julie Dorsey, Fast Bilateral Filtering for the Display of High Dynamic Range Images, SIGGRAPH 2002.
- Erik Reinhard, Michael Stark, Peter Shirley, Jim Ferwerda, Photographics Tone Reproduction for Digital Images, SIGGRAPH 2002.
- Patrick Ledda, Alan Chalmers, Tom Troscianko, Helge Seetzen, Evaluation of Tone Mapping Operators using a High Dynamic Range Display, SIGGRAPH 2005.
- Jiangtao Kuang, Hiroshi Yamaguchi, Changmeng Liu, Garrett Johnson, Mark Fairchild, Evaluating HDR Rendering Algorithms, ACM Transactions on Applied Perception, 2007.