

# Tone mapping

Digital Visual Effects, Spring 2009

*Yung-Yu Chuang*

2009/3/5

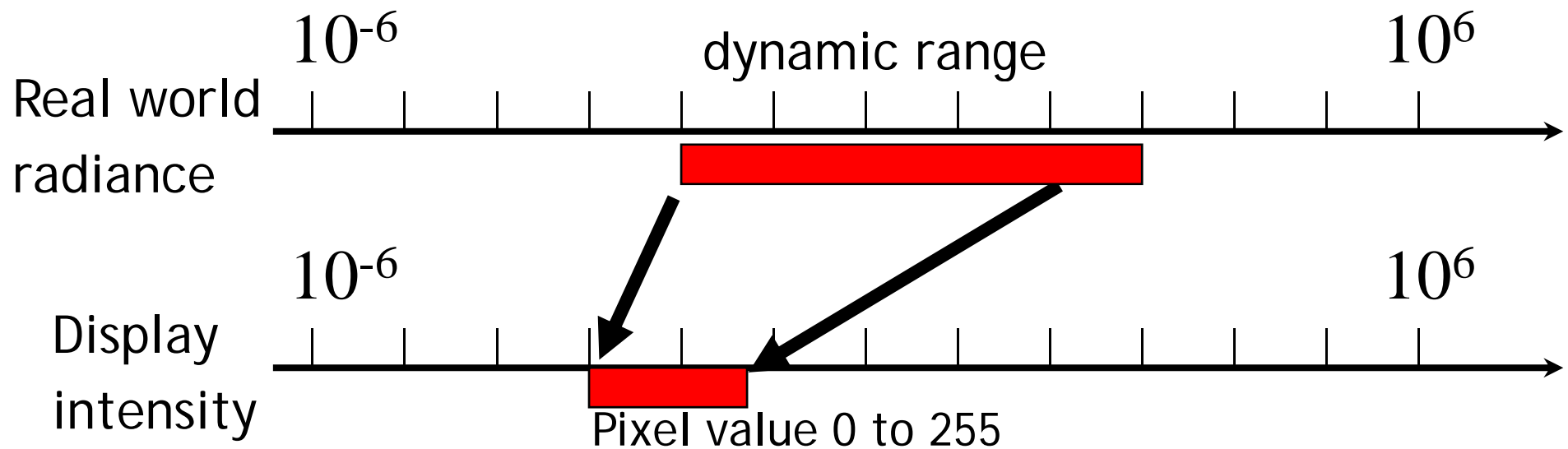
*with slides by Fredo Durand, and Alexei Efros*

# 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?

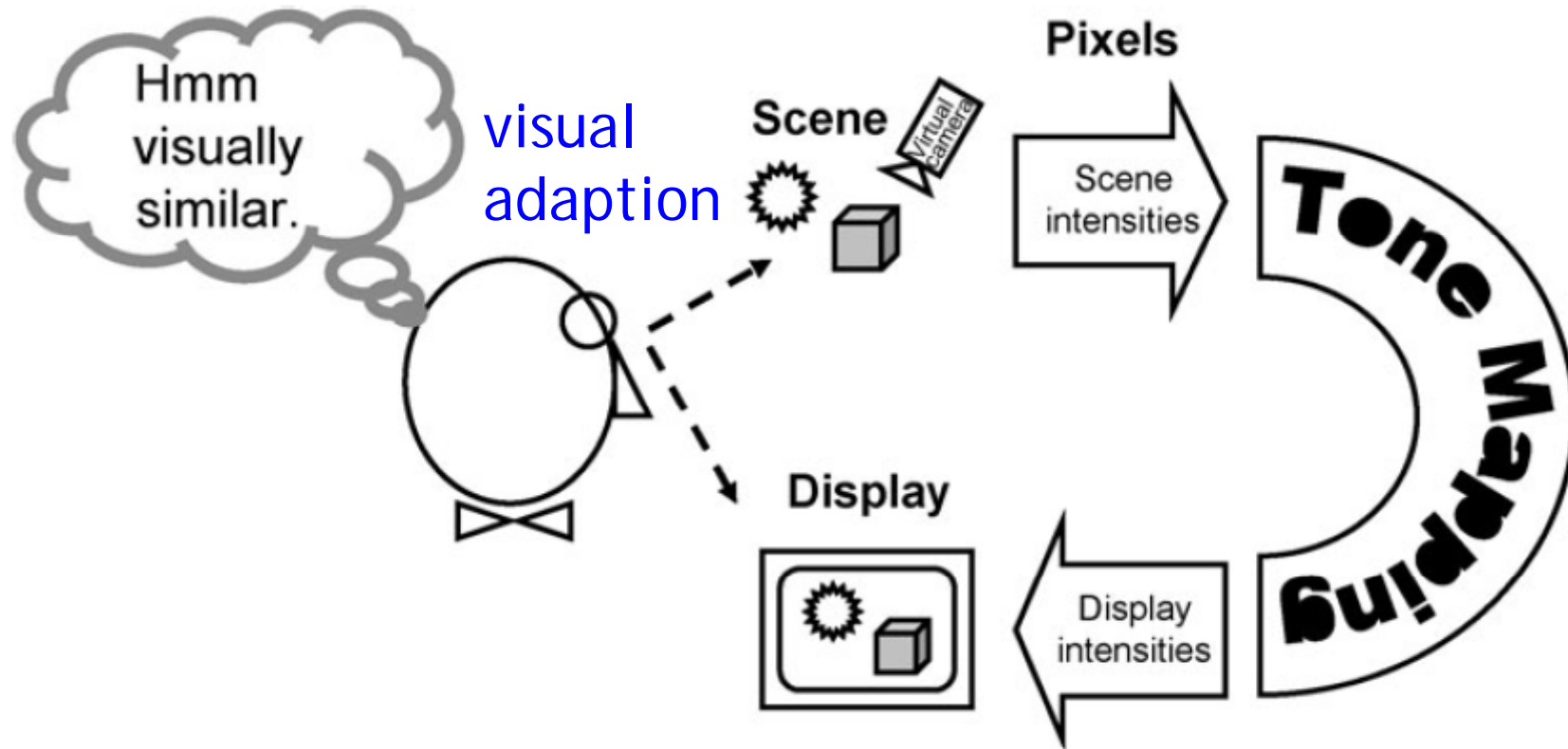
Linear scaling?, thresholding?



CRT has 300:1 dynamic range

# The ultimate goal is a visual match

---



We do not need to reproduce the true radiance as long as it gives us a visual match.

# Eye is not a photometer!

---



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



# Eye is not a photometer!

---



Are the headlights different in two images? Physically, they are the same, but perceptually different.

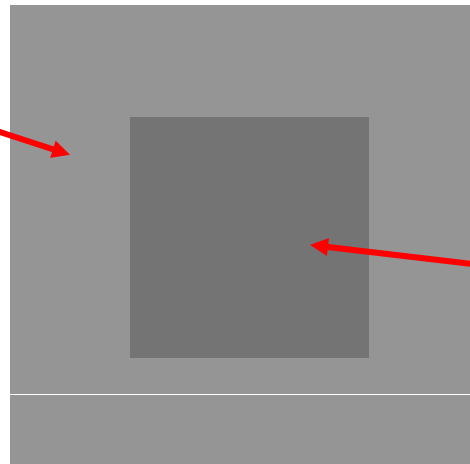
# We are more sensitive to contrast

- Weber's law

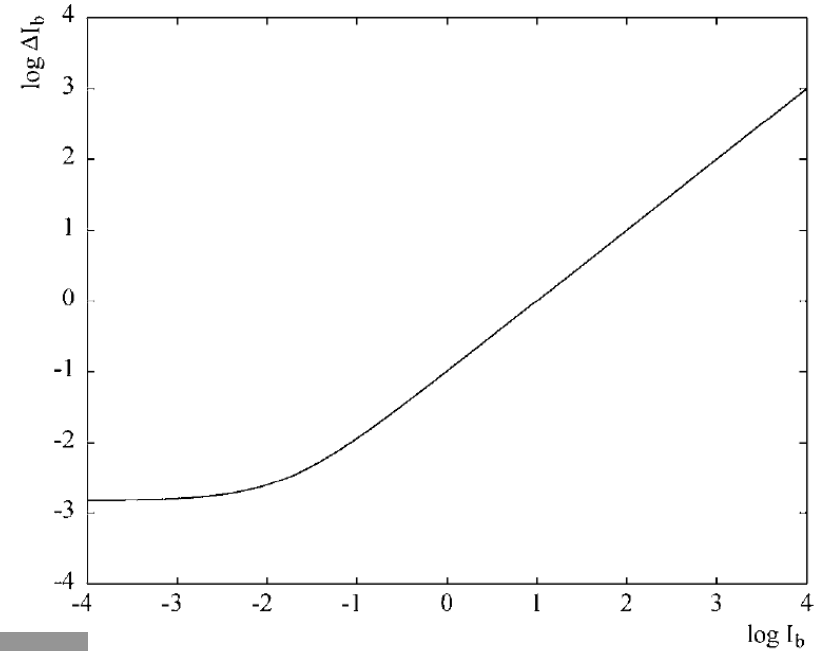
Just-noticeable  
Difference (JND)

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

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

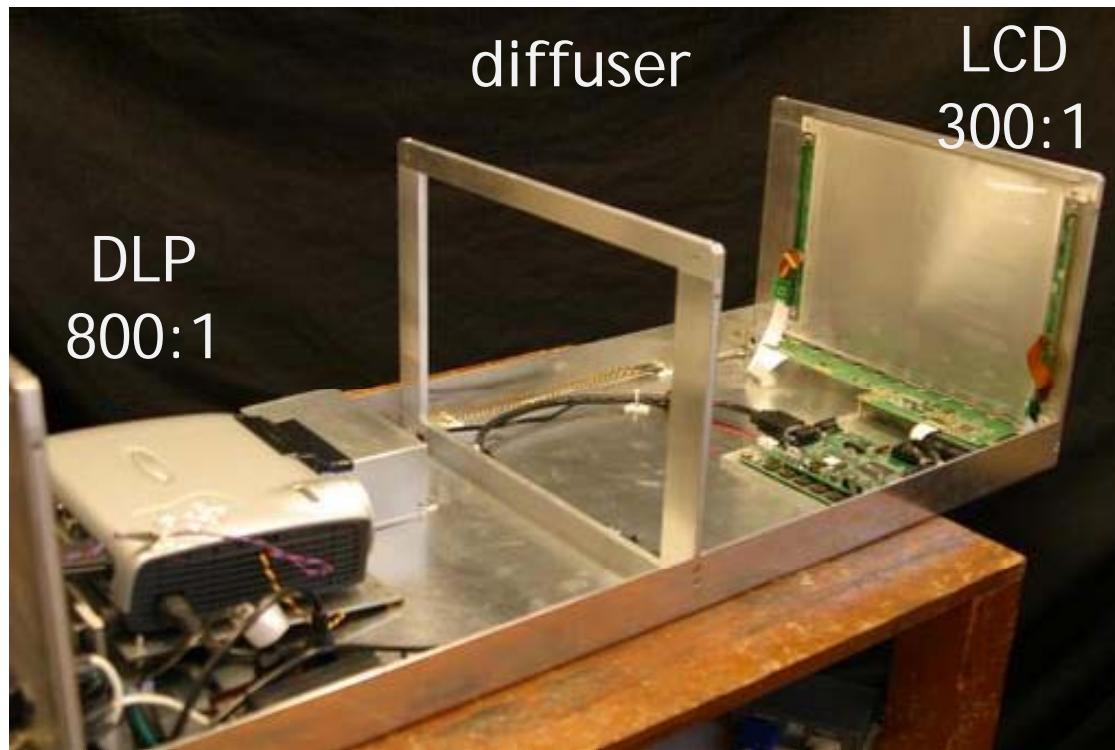
$$\begin{bmatrix} R_d \\ G_d \\ B_d \end{bmatrix} = \begin{bmatrix} L_d \frac{R_w}{L_w} \\ L_d \frac{G_w}{L_w} \\ L_d \frac{B_w}{L_w} \end{bmatrix}$$

- Log domain is usually preferred.

# HDR Display

---

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



Theoretically,  
240,000:1.

Due to imperfect  
optical depth,  
54,000:1 measured

HDR display system, Sunnybrook Technology, SIGGRAPH2004

# Sunnybrook HDR display

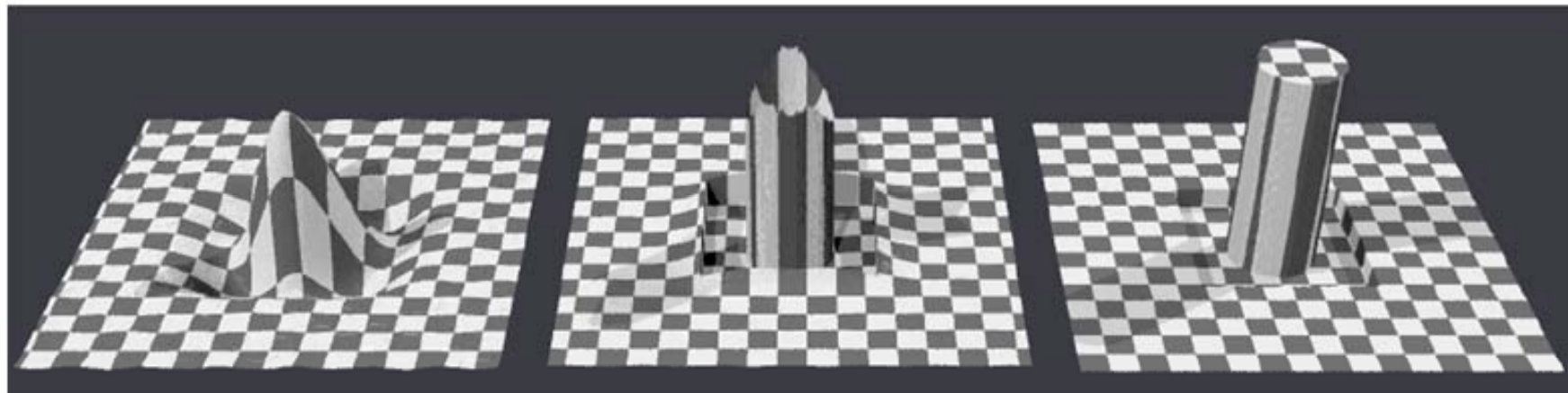
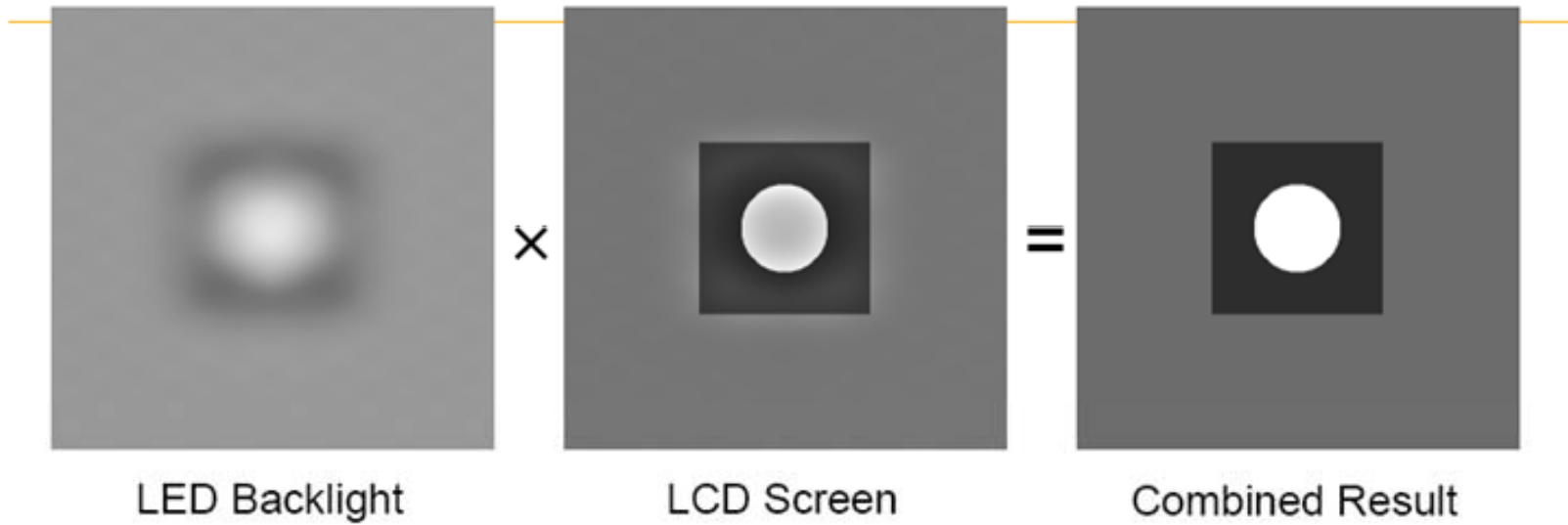
---

- Use Bright Source + Two 8-bit Modulators
  - Transmission multiplies together
  - Over 10,000:1 dynamic range possible



# How it works

---



# Brightside HDR display

---



**BRIGHTSIDE™**  
Technologies Inc

**World's First Extreme Dynamic Range Display**

37"

200000:1

Acquired  
by [Dolby](#)



# Tone mapping operators

---

- Spatial (global/local)
- Frequency domain
- Gradient domain
  
- 3 papers from SIGGRAPH 2002
  - ◆ Photographic Tone Reproduction for Digital Images
  - ◆ Fast Bilateral Filtering for the Display of High-Dynamic-Range Images
  - ◆ Gradient Domain High Dynamic Range Compression

# Photographic Tone Reproduction for Digital Images

Erik Reinhard   Mike Stark

Peter Shirley   Jim Ferwerda

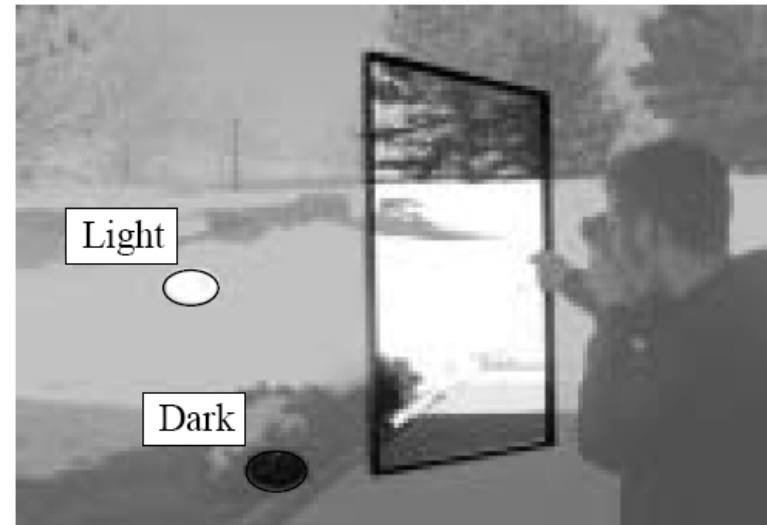
SIGGRAPH 2002

# 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

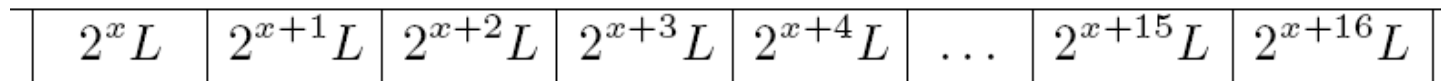
# Zone system



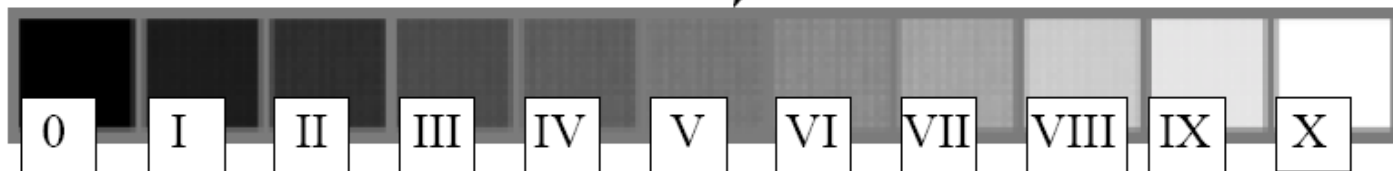
Darkest  
textured  
shadow

Dynamic range = 15 scene zones

Brightest  
textured  
highlight



Middle grey maps to Zone V



Print zones

# The Zone system

---

- Formalism to talk about exposure, density
- Zone = intensity range, in powers of two
- In the scene, on the negative, on the print



Source: Ansel Adams

# The Zones

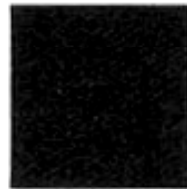
---

## The Zones

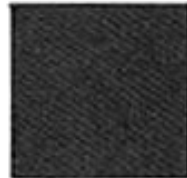
**0** Solid black; the same as the film rebate



**I** Nearly black; just different from Zone 0



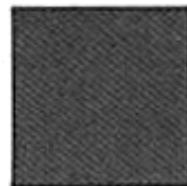
**II** The first hint of texture



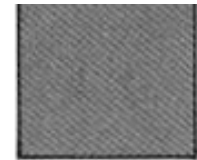
**III** Textured shadow; the first recognizable shadow detail



**IV** Average shadow value on Caucasian skin, foliage and buildings



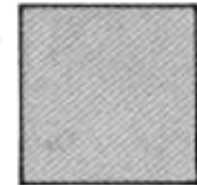
**V** Middle grey: the pivot value; light foliage, dark skin



**VI** Caucasian skin, textured light grey; shadow on snow



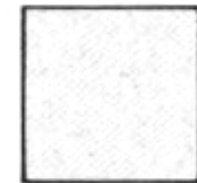
**VII** Light skin; bright areas with texture, such as snow in low sunlight



**VIII** Highest zone with any texture



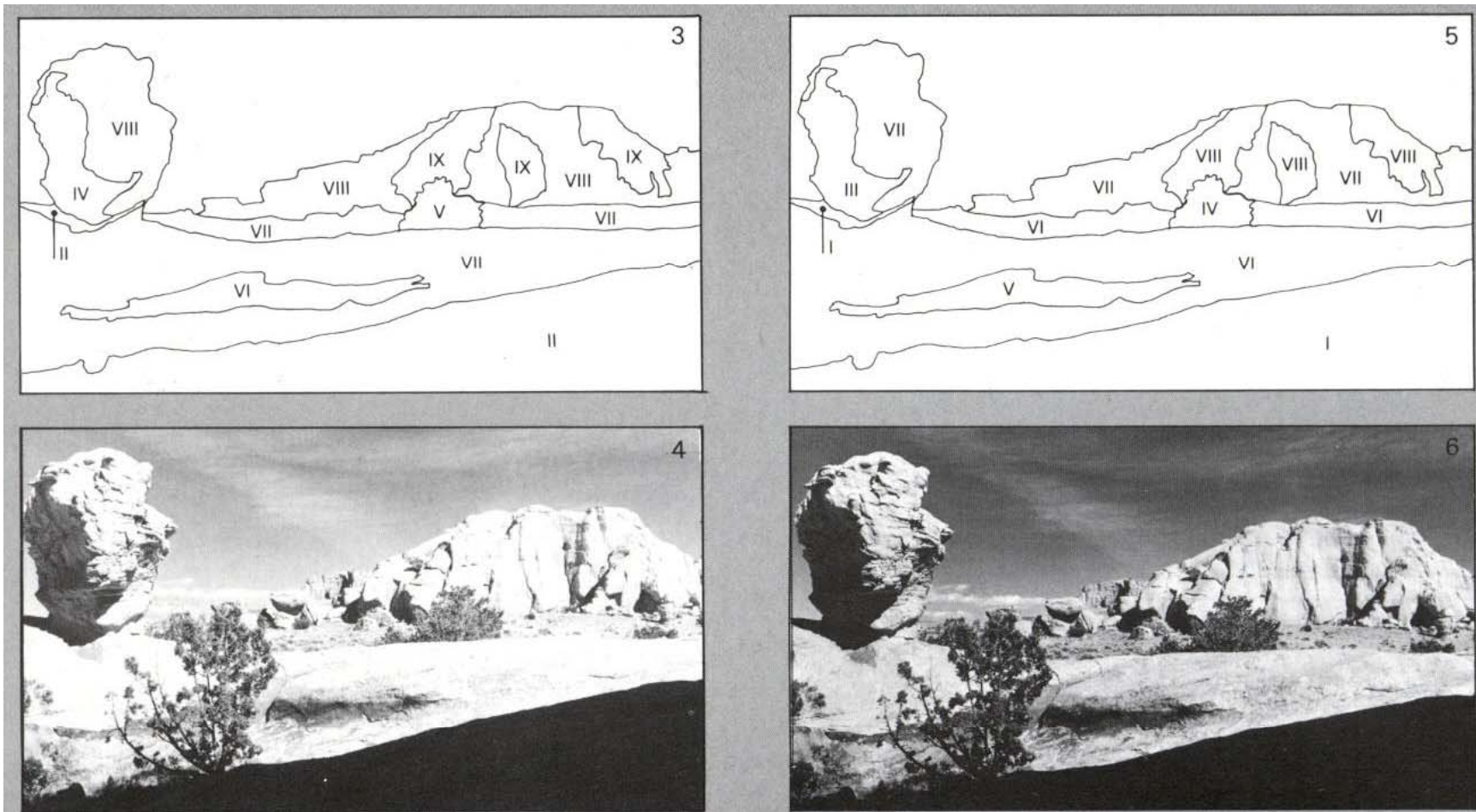
**IX** Pure untextured white



# The Zone system

---

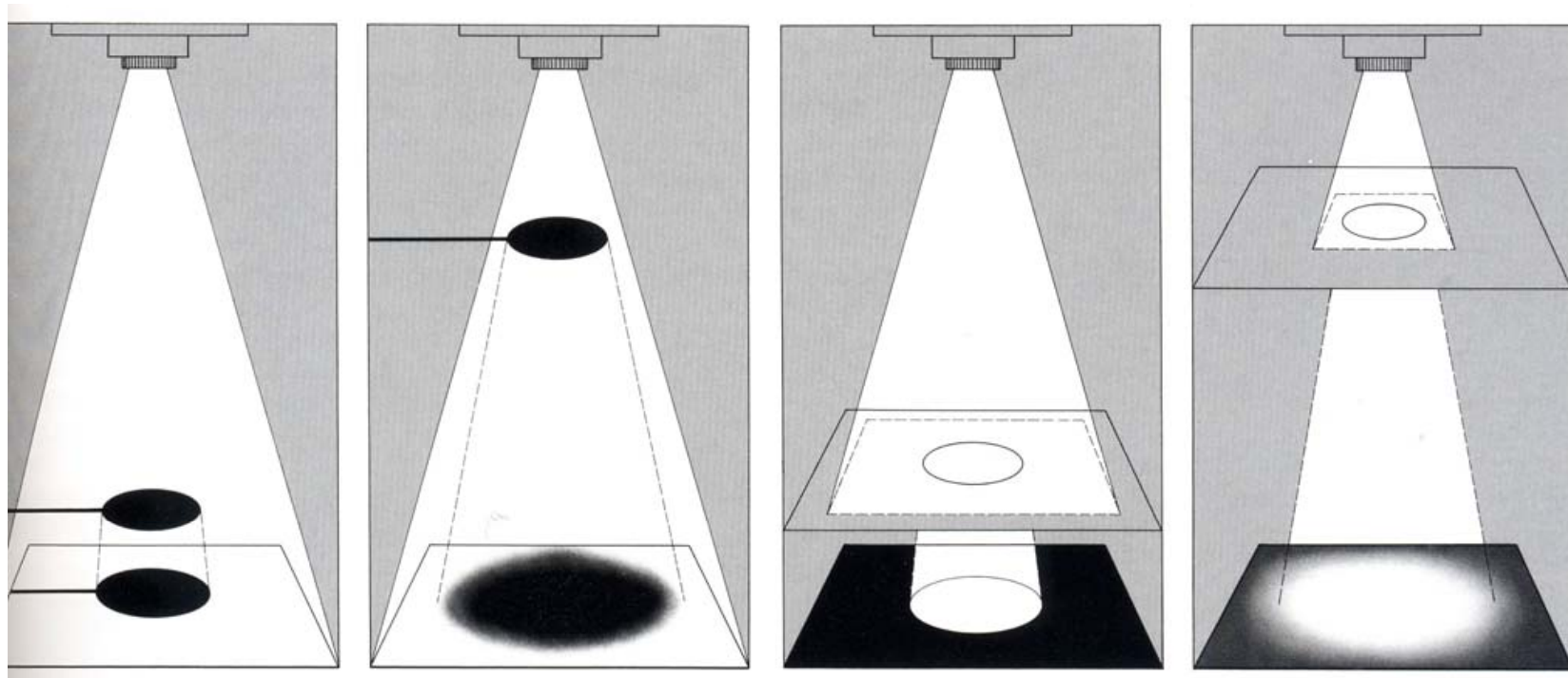
- You decide to put part of the system in a given zone



# Dodging and burning

---

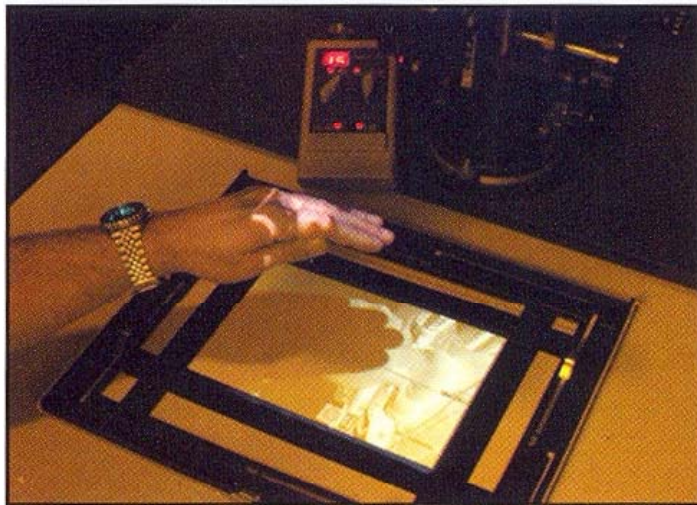
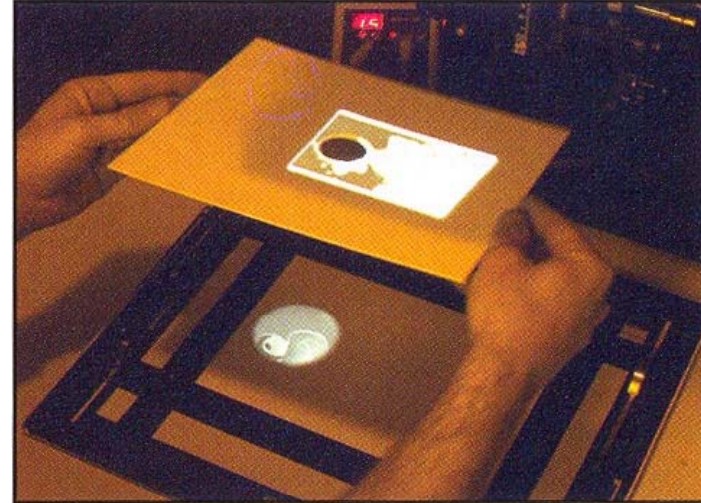
- During the print
- Hide part of the print during exposure
  - Makes it brighter





# Dodging and burning

---



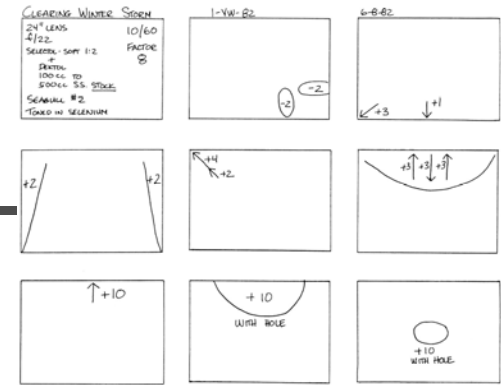
dodging

burning

From Photography by London et al.

# Dodging and burning

- Must be done for every single print!



Straight print



After dodging and burning

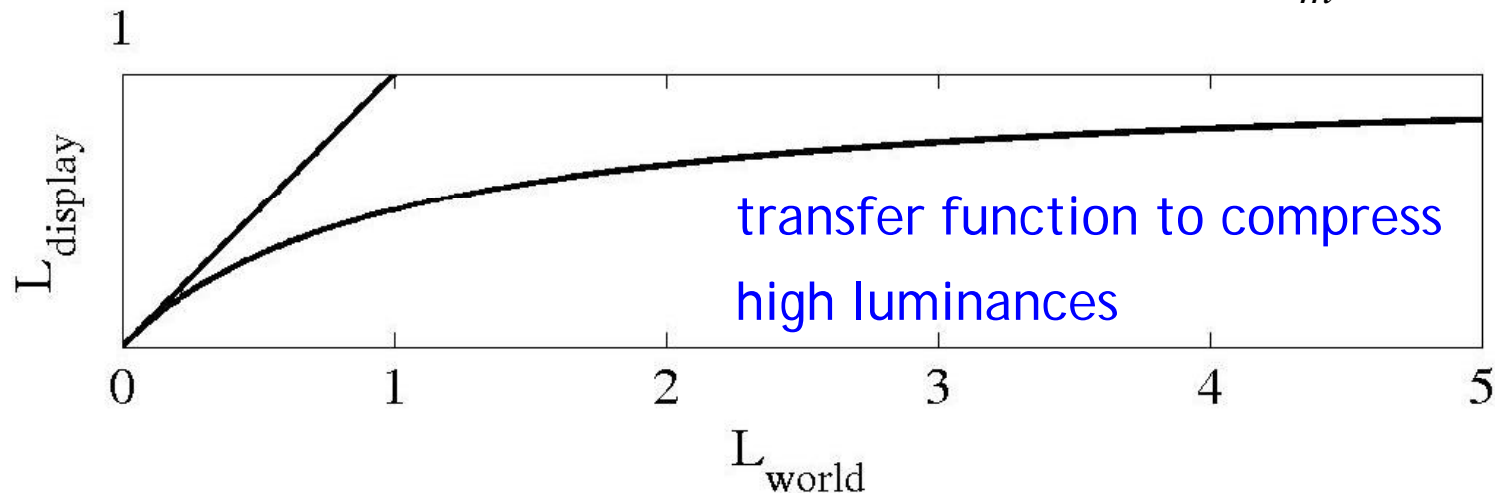
# Global operator

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

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

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

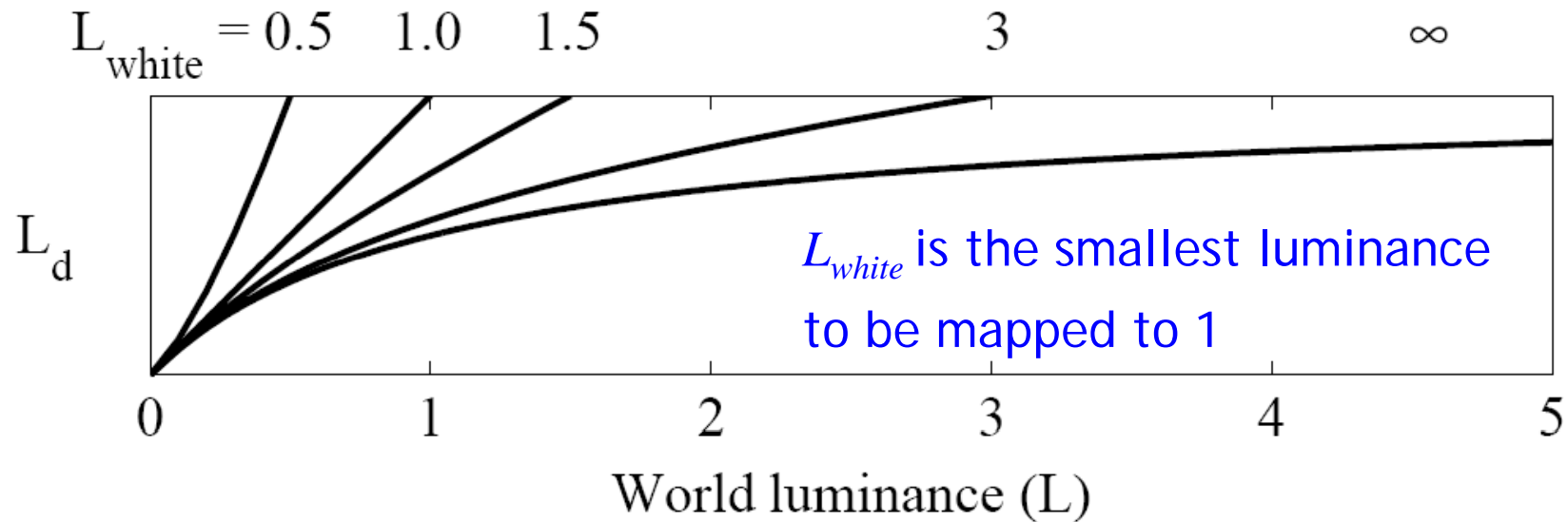


# Global operator

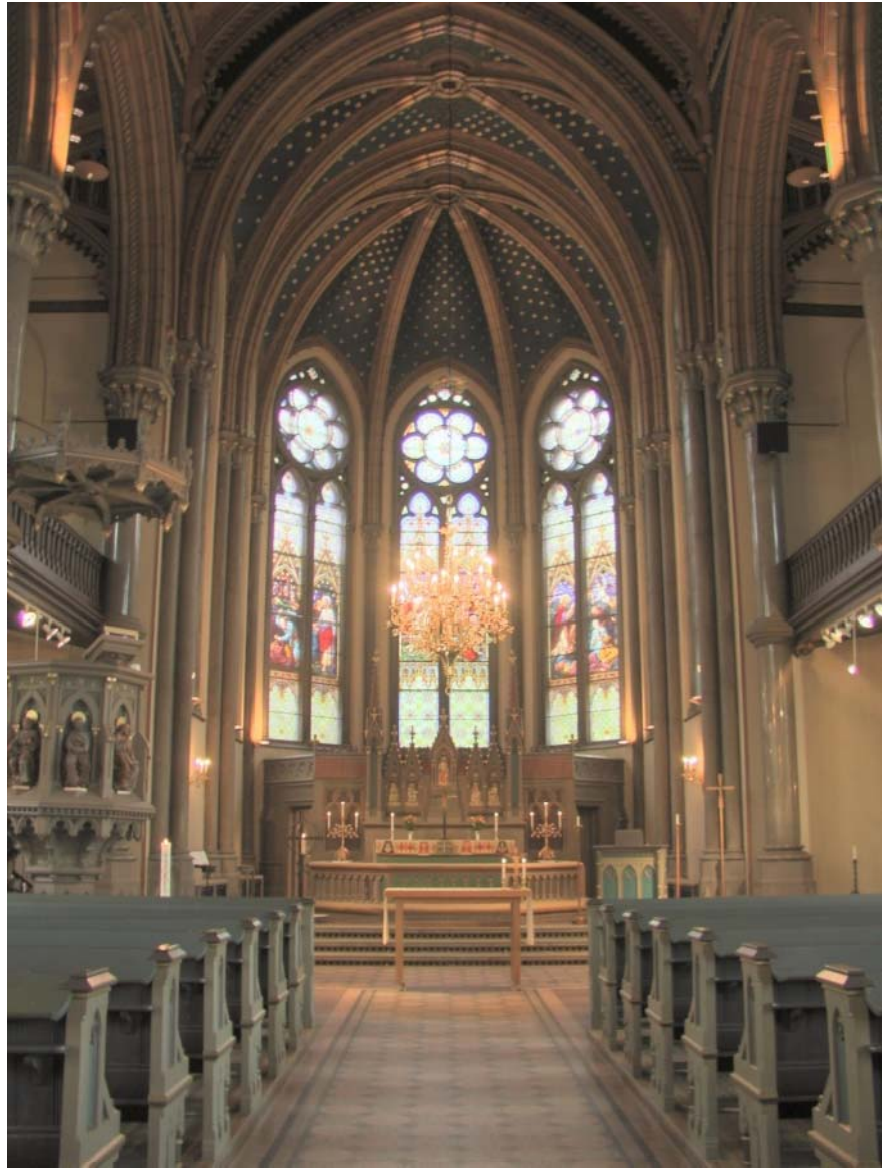
---

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

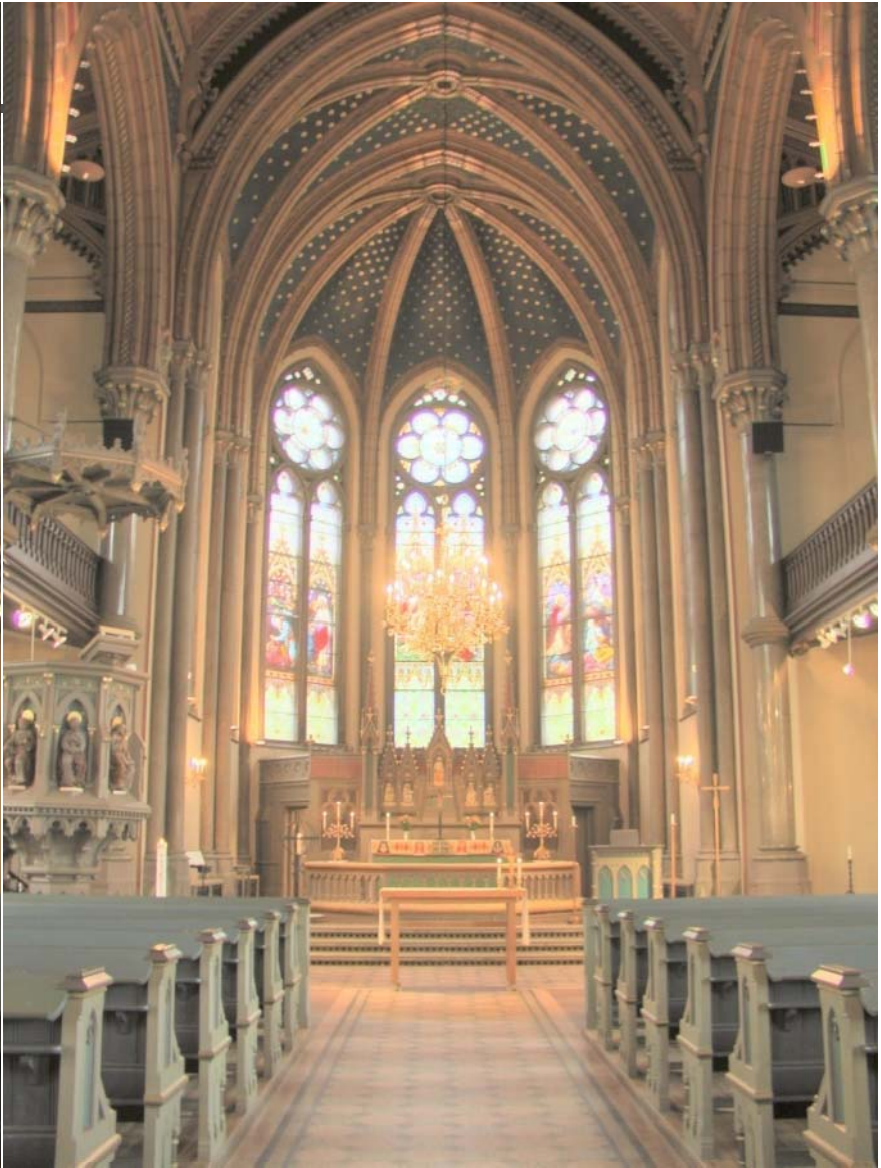
$$L_d(x, y) = \frac{L_m(x, y) \left( 1 + \frac{L_m(x, y)}{L_{white}^2(x, y)} \right)}{1 + L_m(x, y)}$$







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

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

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

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

# Dodging and burning (local operators)

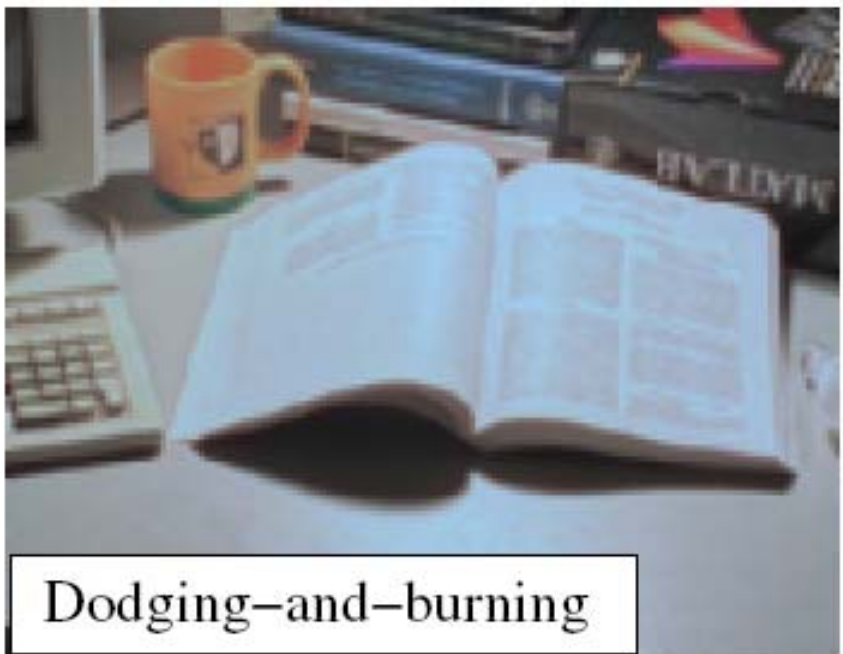
---

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

- 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

---





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

Frédo Durand & Julie Dorsey

SIGGRAPH 2002

# Frequency domain

---

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

$$\text{image} = \text{illuminance} * \text{reflectance}$$

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

# A typical photo

---

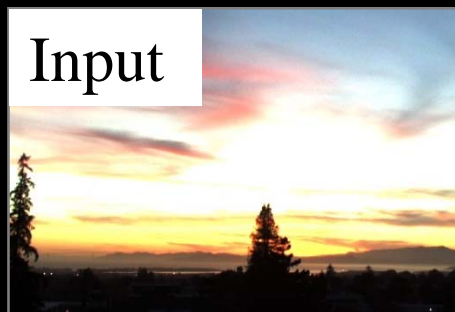
- Sun is overexposed
- Foreground is underexposed



# Gamma compression

---

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



# Gamma compression on intensity

---

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

Intensity



Gamma on intensity



Color

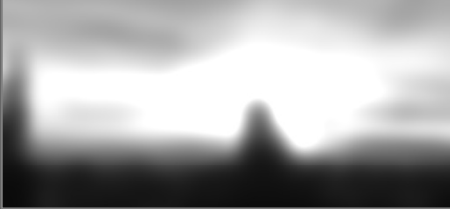


# Chiu et al. 1993

---

- Reduce contrast of low-frequencies
- Keep high frequencies

Low-freq.



High-freq.



Color



Reduce low frequency

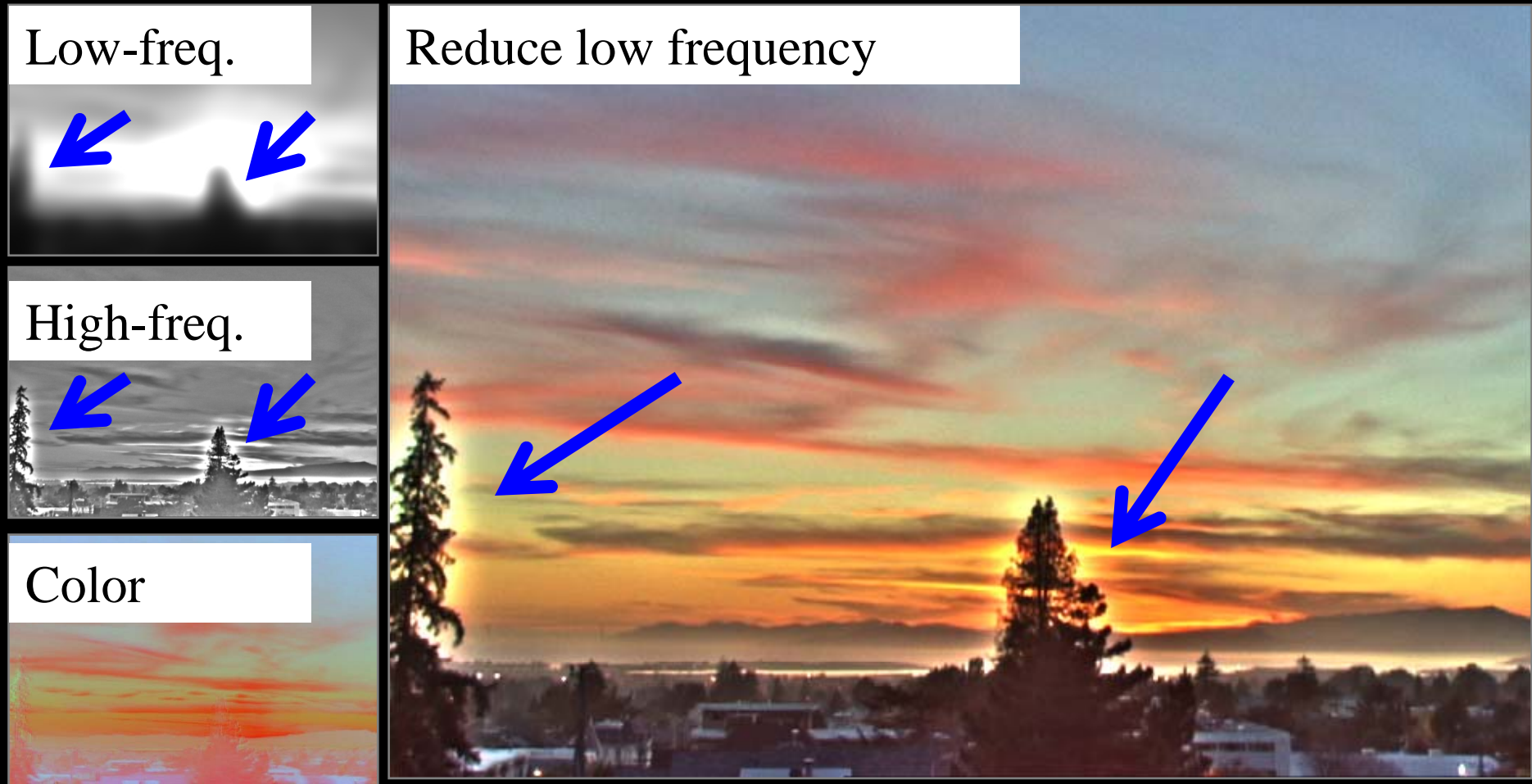




# The halo nightmare

---

- For strong edges
- Because they contain high frequency





# Durand and Dorsey

---

- Do not blur across edges
- Non-linear filtering

Large-scale



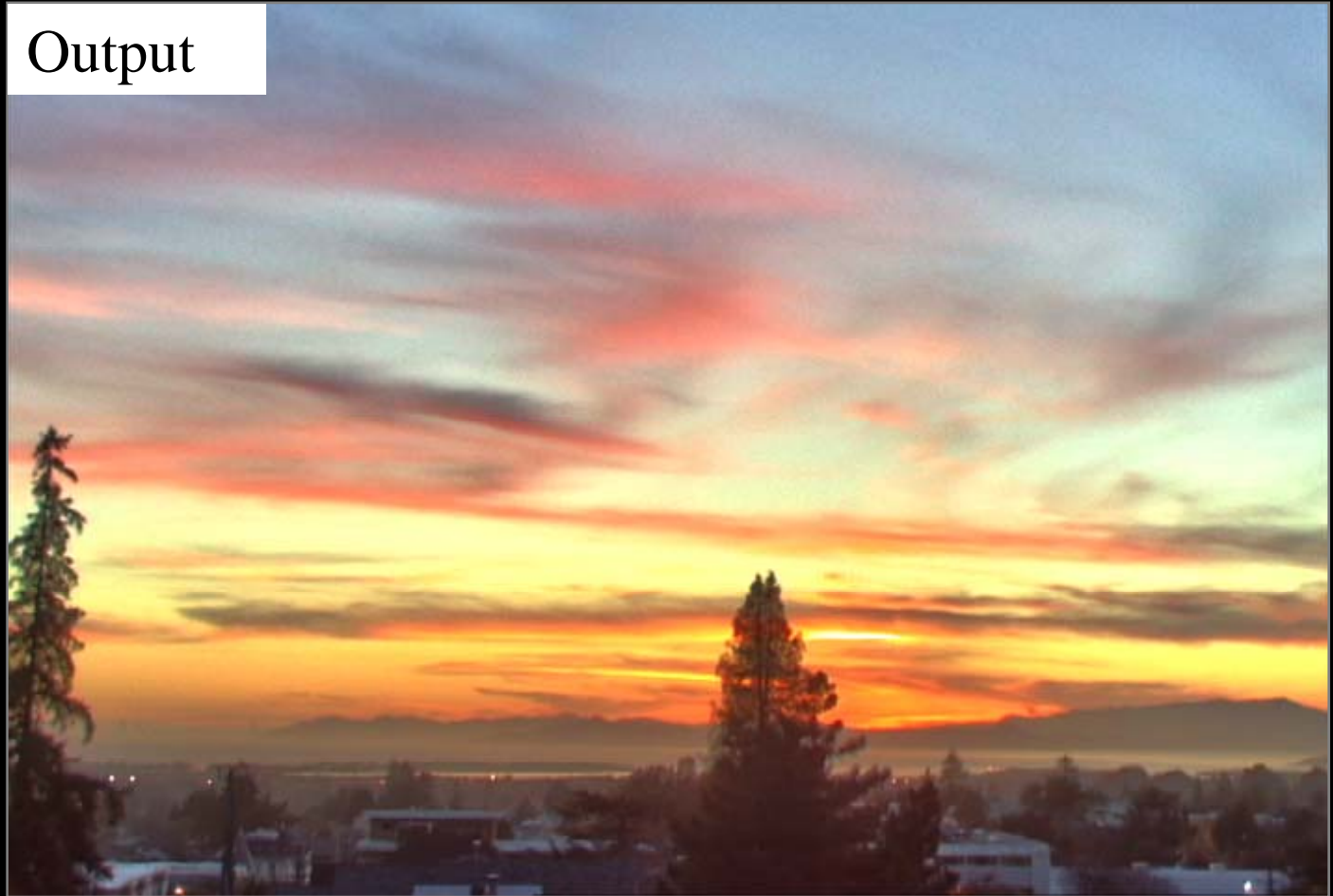
Detail



Color



Output



# Edge-preserving filtering

---

- Blur, but not across edges

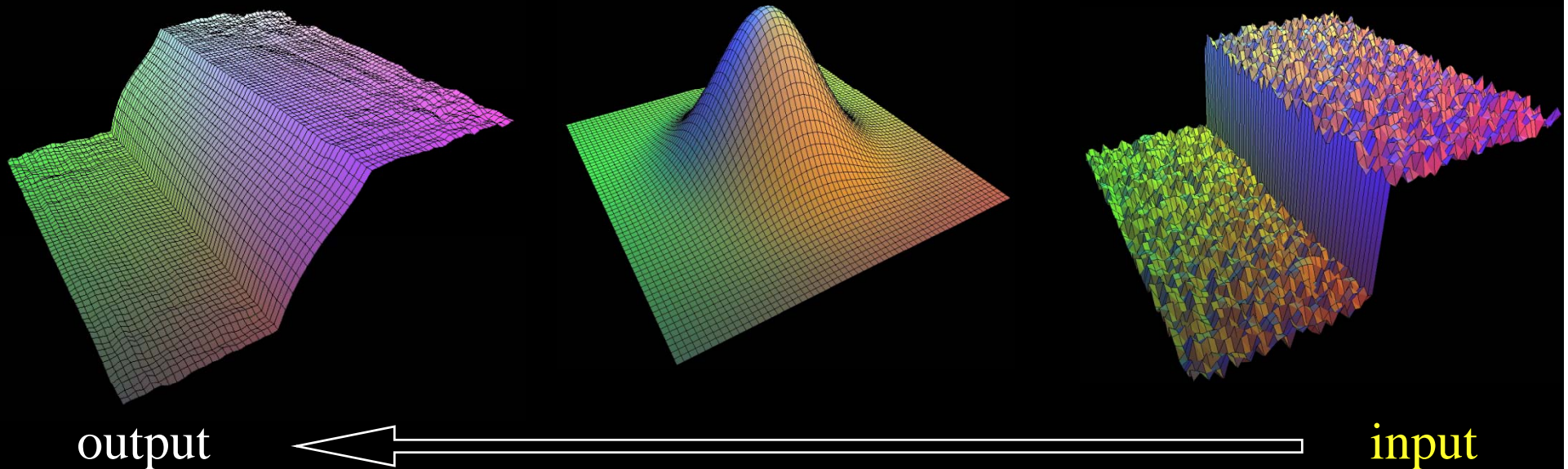


- Anisotropic diffusion [Perona & Malik 90]
  - Blurring as heat flow
  - LCIS [Tumblin & Turk]
- **Bilateral filtering [Tomasi & Manduci, 98]**

# Start with Gaussian filtering

---

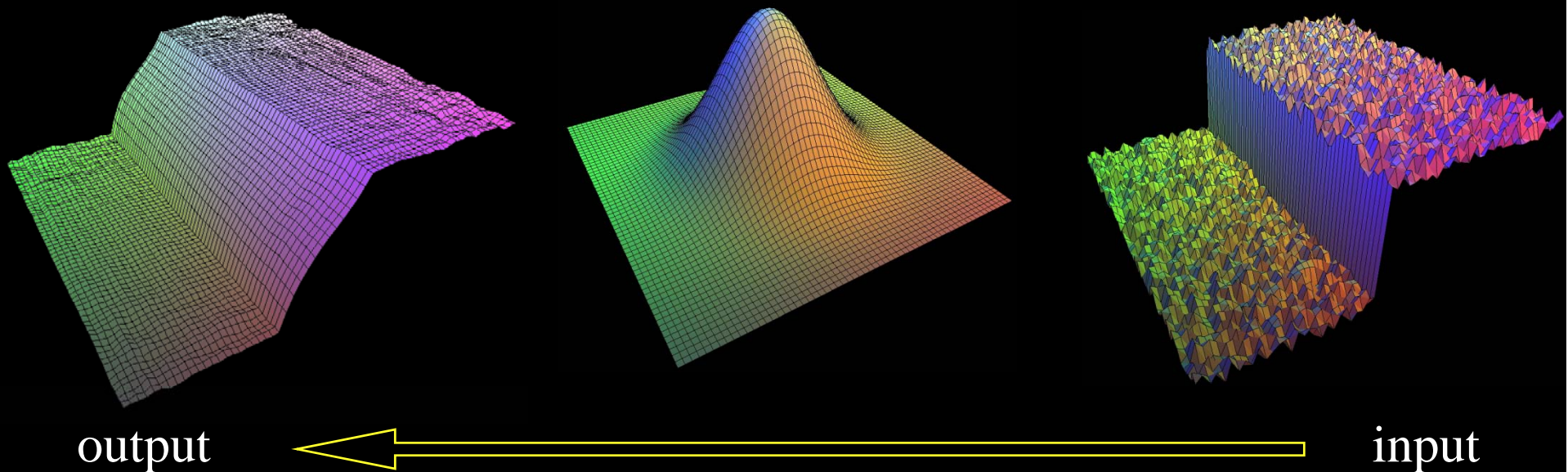
- Here, input is a step function + noise



# Start with Gaussian filtering

---

- Spatial Gaussian  $f$

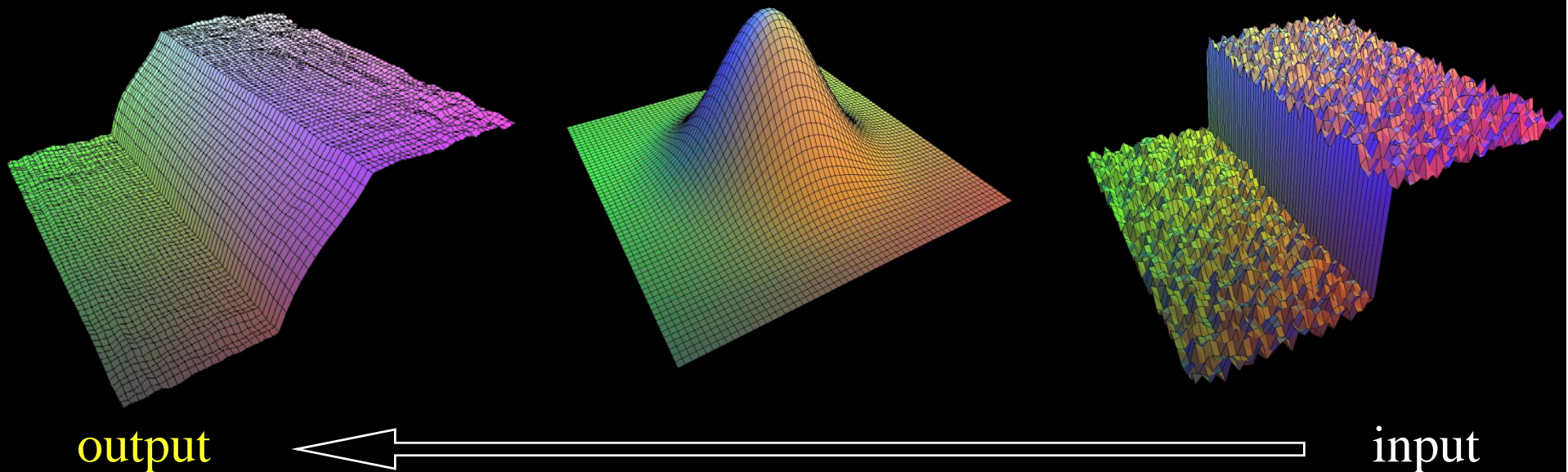




# Start with Gaussian filtering

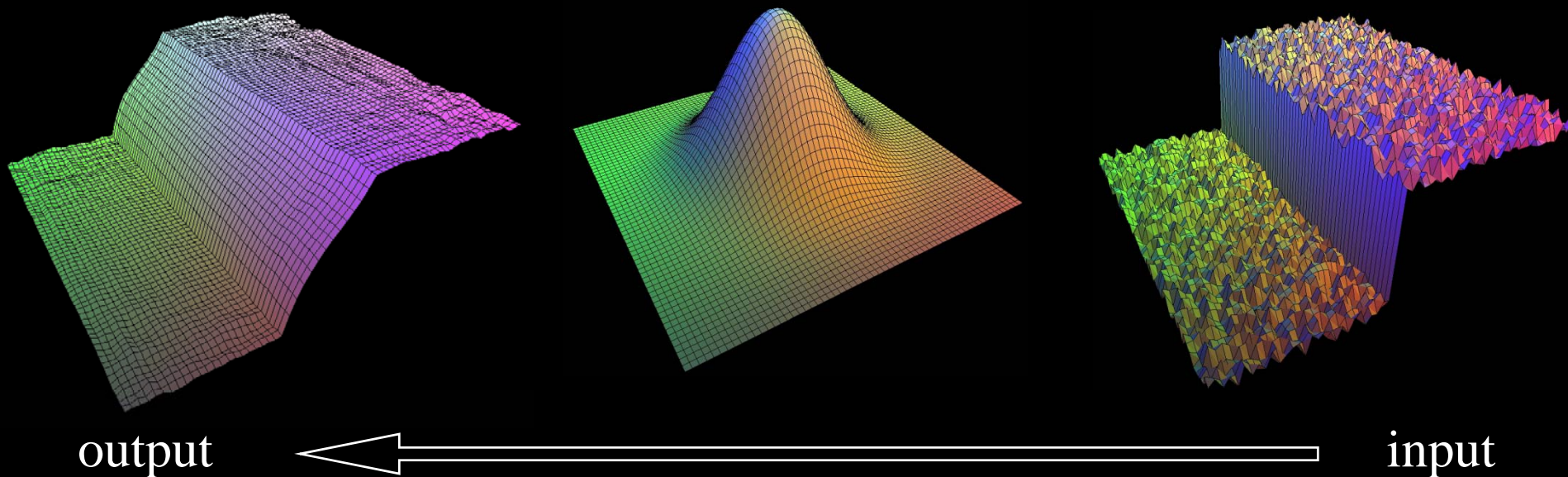
---

- Output is blurred



# Gaussian filter as weighted average

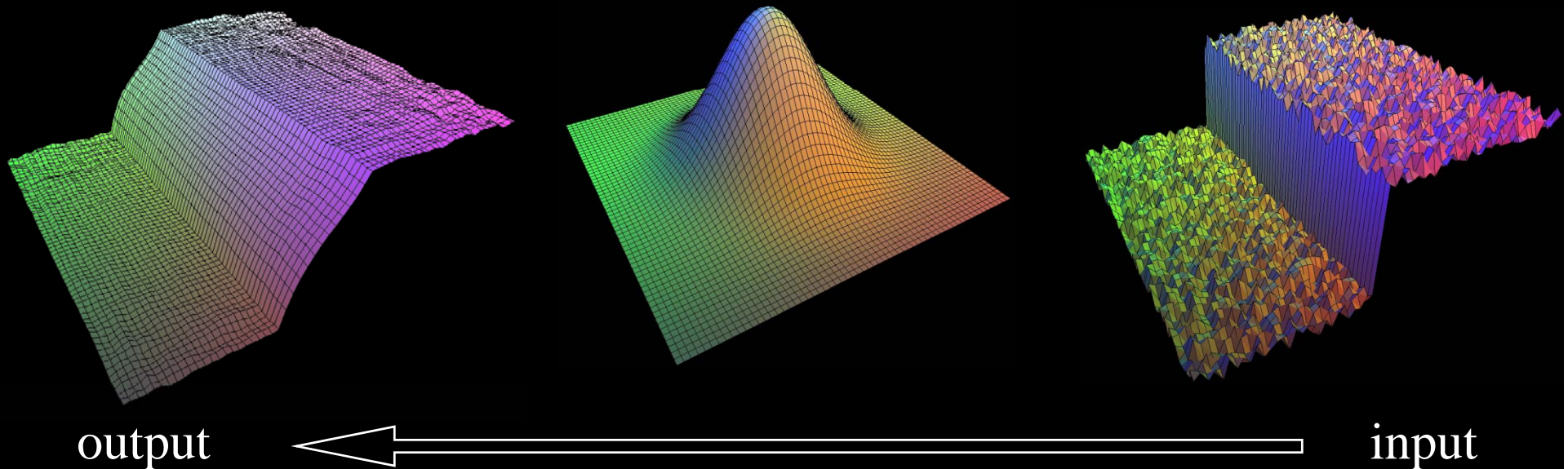
---



# The problem of edges

---

- Here, “pollutes” our estimate  $J(x)$
- It is too different

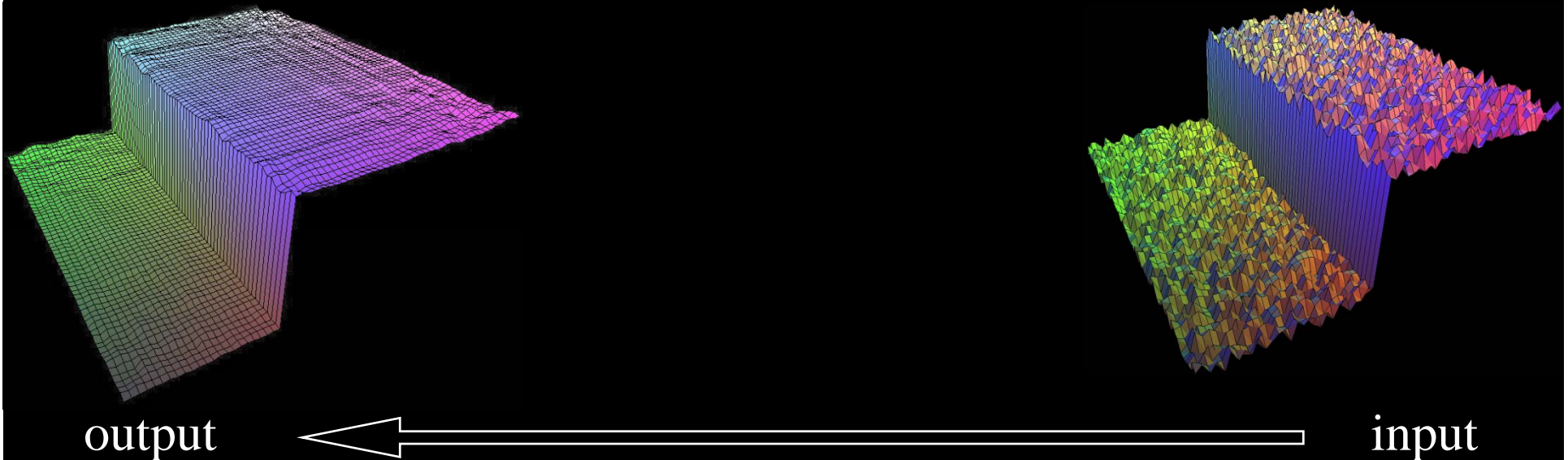




# Principle of Bilateral filtering

---

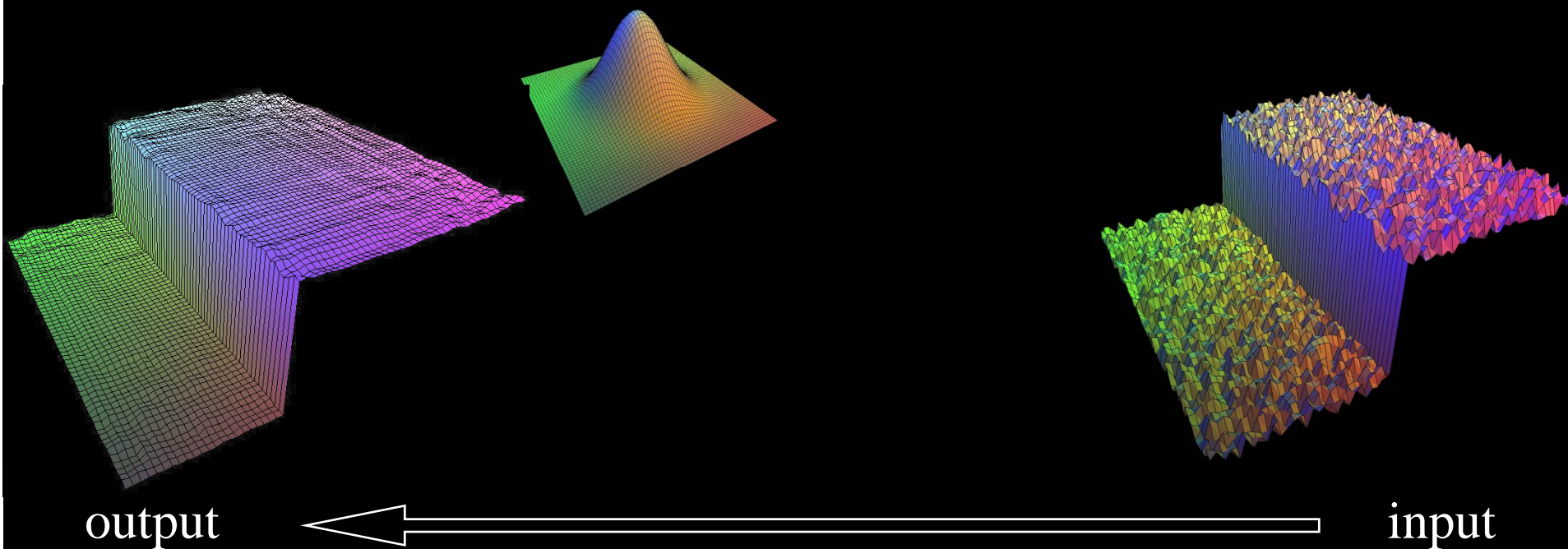
- [Tomasi and Manduchi 1998]
- Penalty  $g$  on the intensity difference



# Bilateral filtering

---

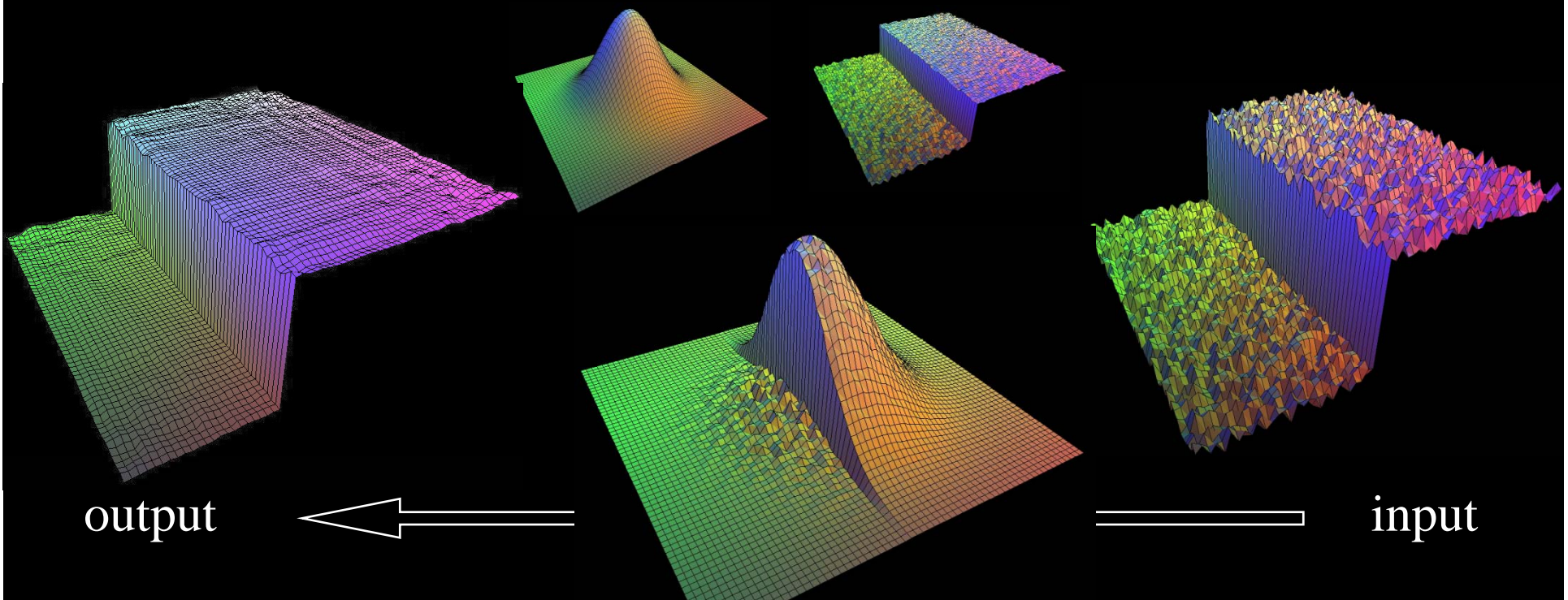
- [Tomasi and Manduchi 1998]
- Spatial Gaussian  $f$



# Bilateral filtering

---

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



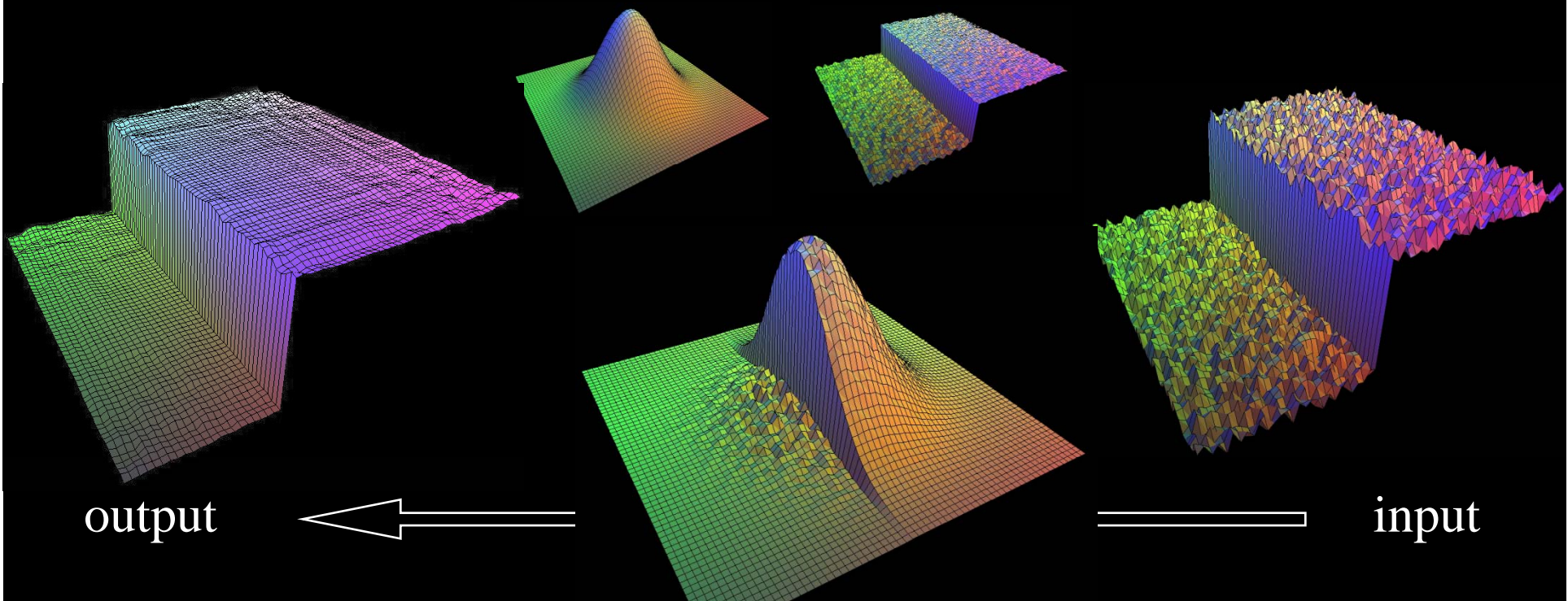


# Normalization factor

---

- [Tomasi and Manduchi 1998]

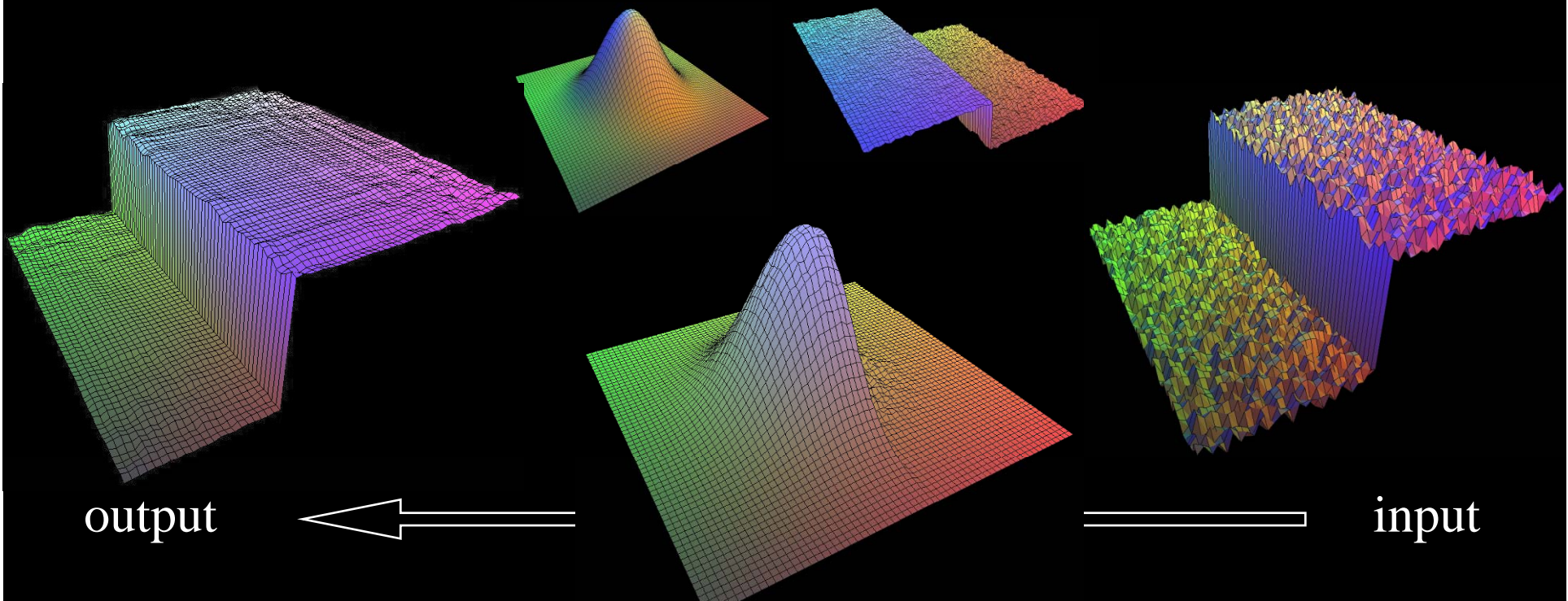
- $k(x) =$



# Bilateral filtering is non-linear

---

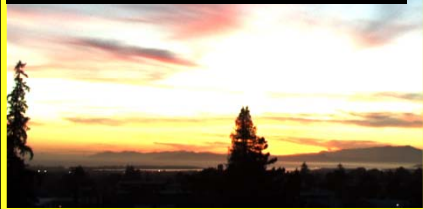
- [Tomasi and Manduchi 1998]
- The weights are different for each output pixel



# Contrast reduction

---

Input HDR image



Contrast  
too high!

# Contrast reduction

---

Input HDR image



Intensity



Color

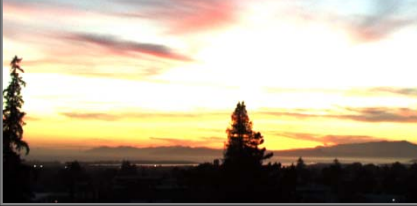




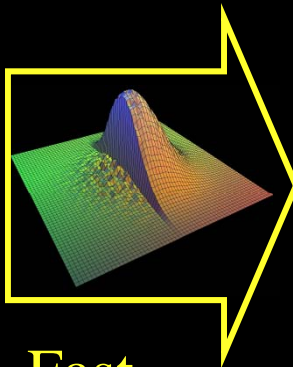
# Contrast reduction

---

Input HDR image



Intensity



Large scale



Fast  
Bilateral  
Filter

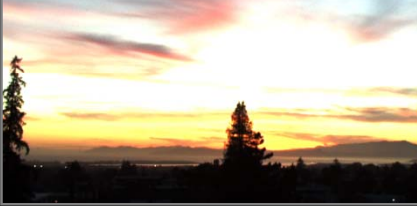
Color



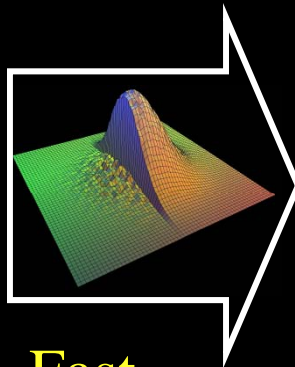
# Contrast reduction

---

Input HDR image



Intensity



Fast  
Bilateral  
Filter

Large scale



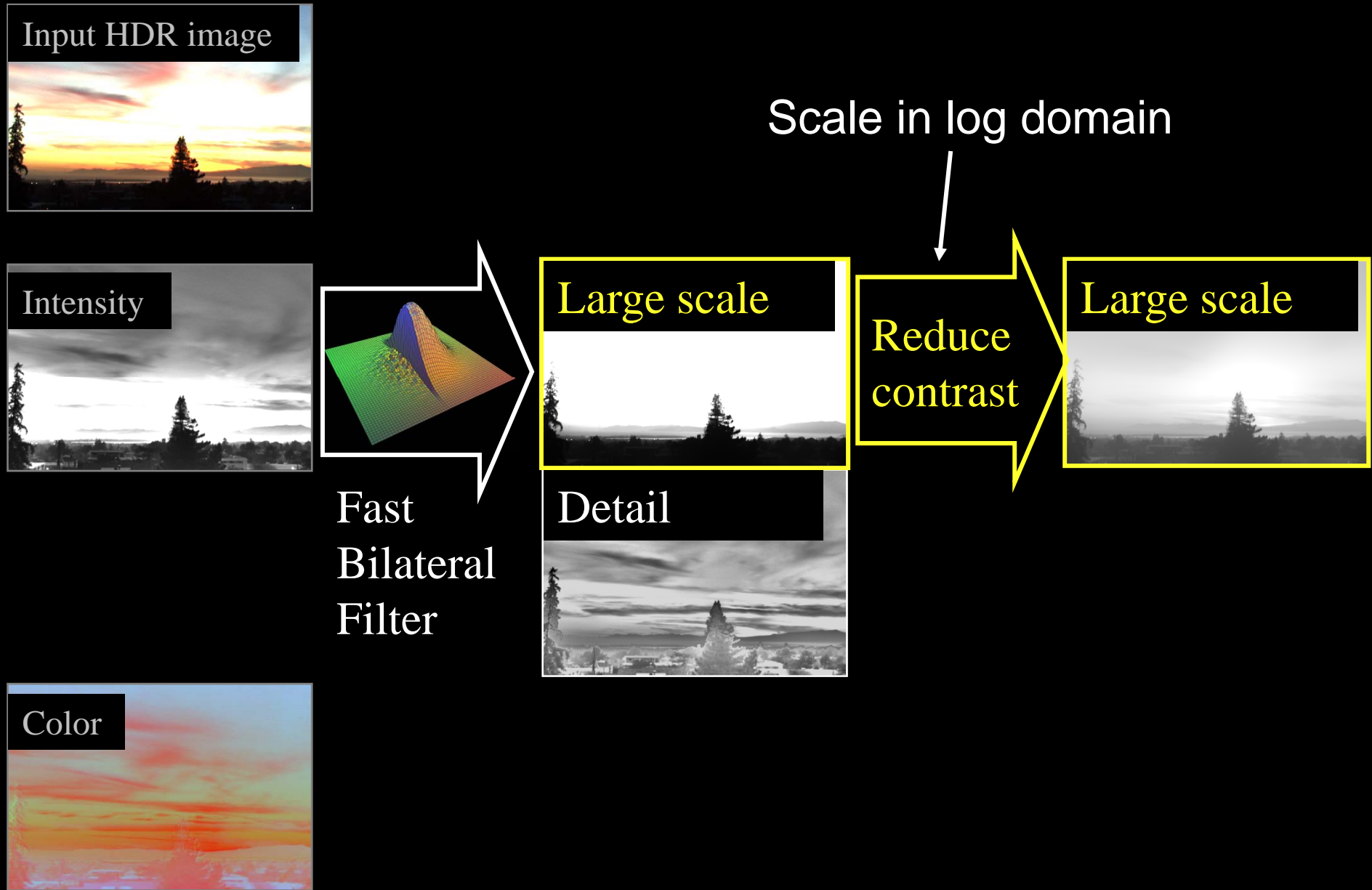
Detail



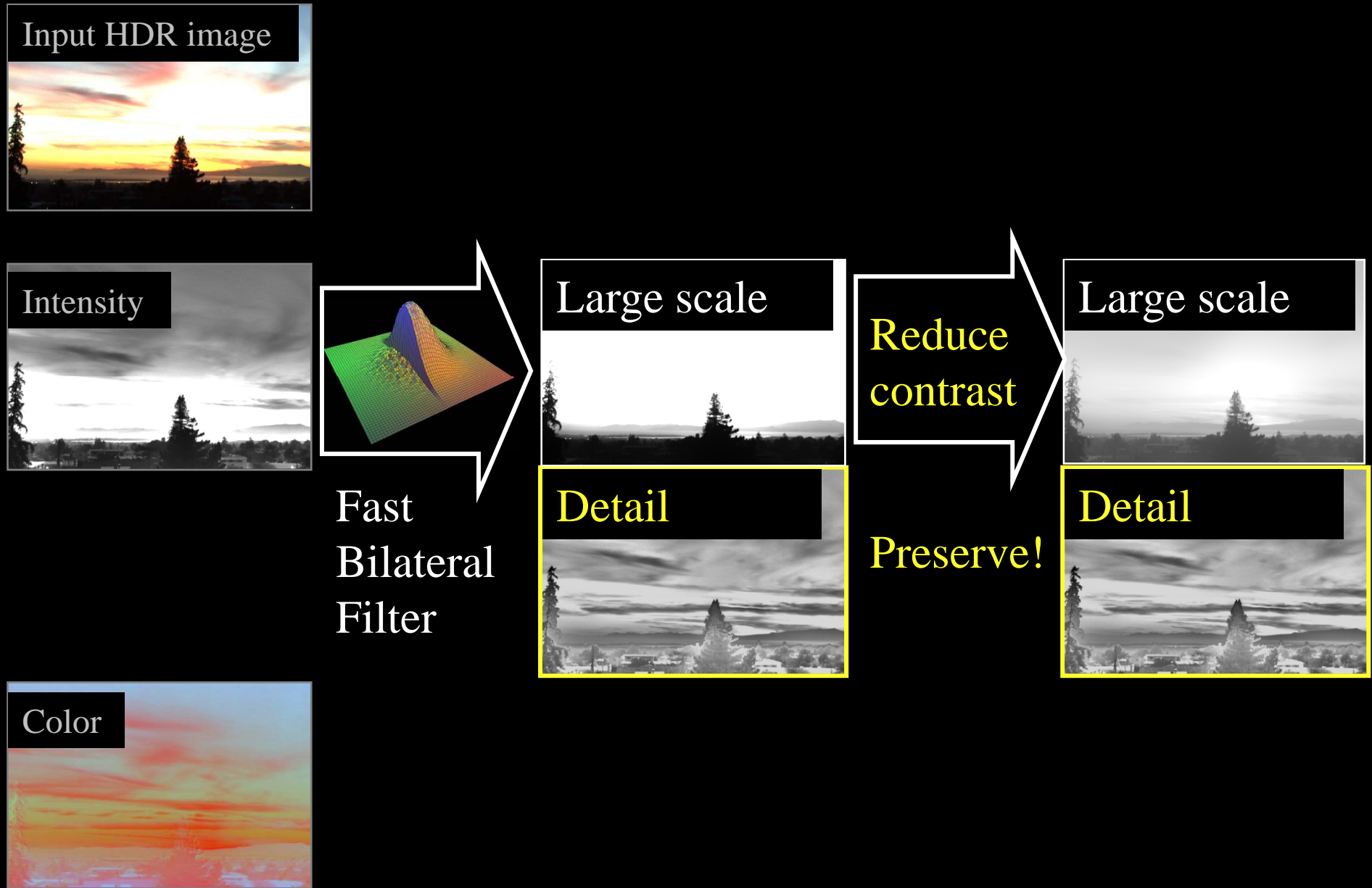
Color



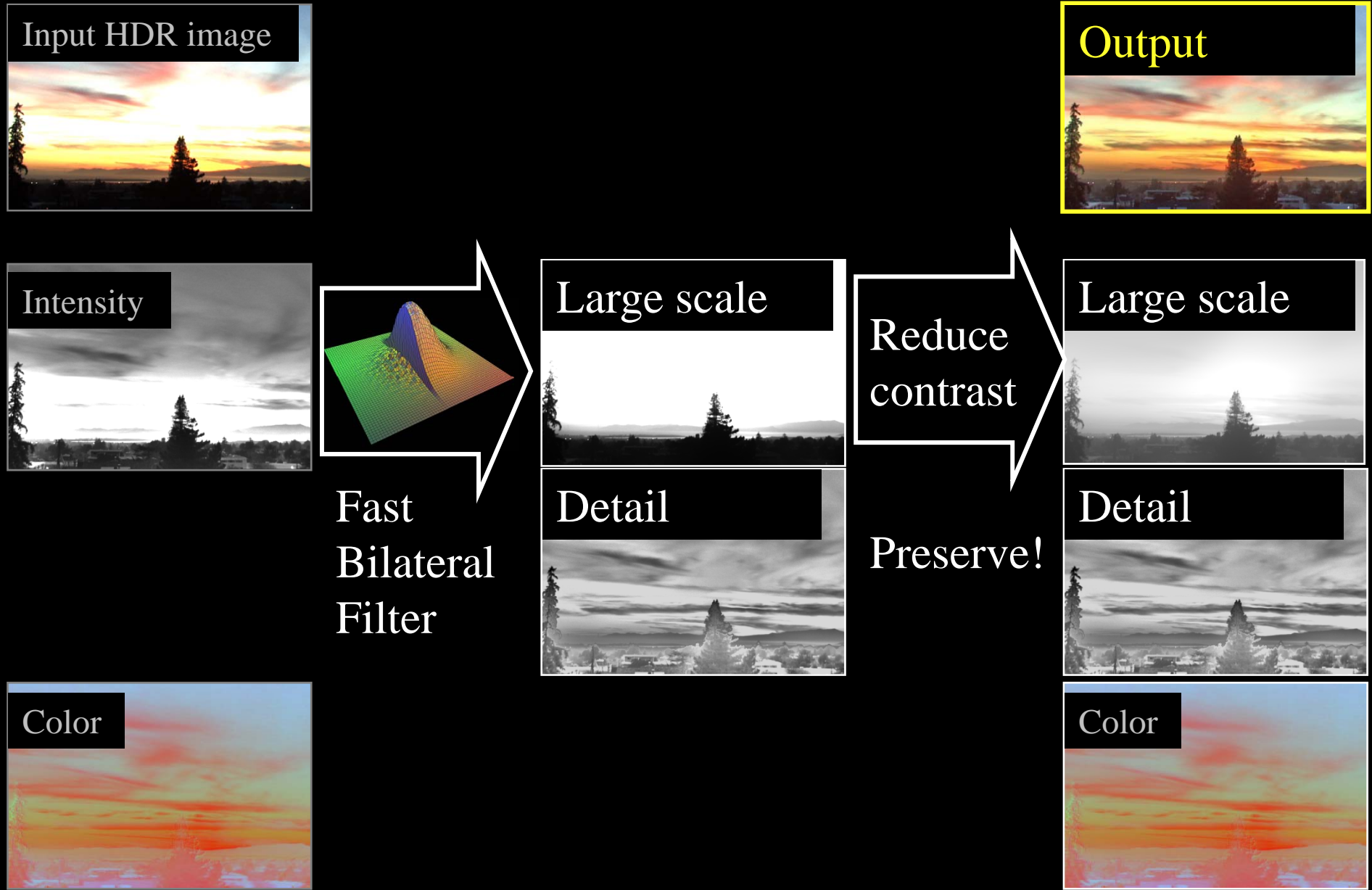
# Contrast reduction



# Contrast reduction



# Contrast reduction

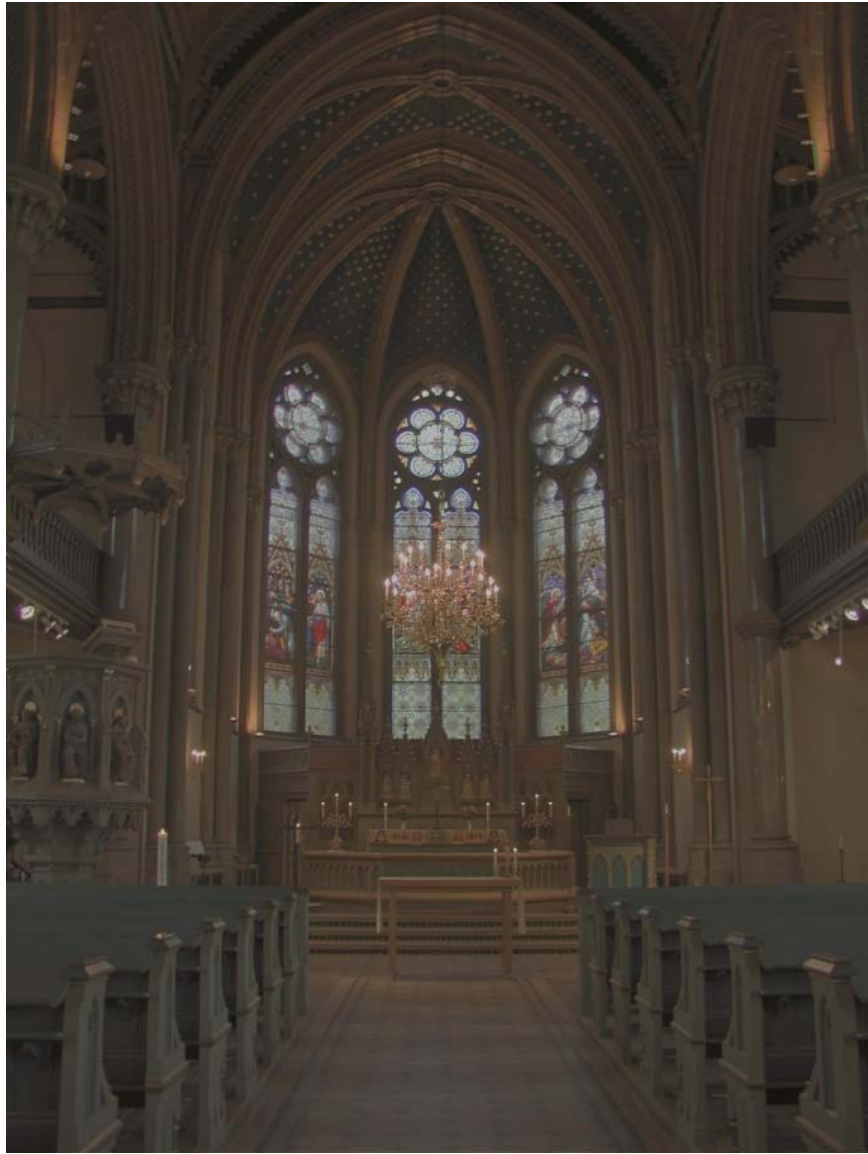


# Bilateral filter is slow!

---

- 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 photography.





Oppenheim



bilateral

# Gradient Domain High Dynamic Range Compression

Raanan Fattal   Dani Lischinski   Michael Werman

SIGGRAPH 2002

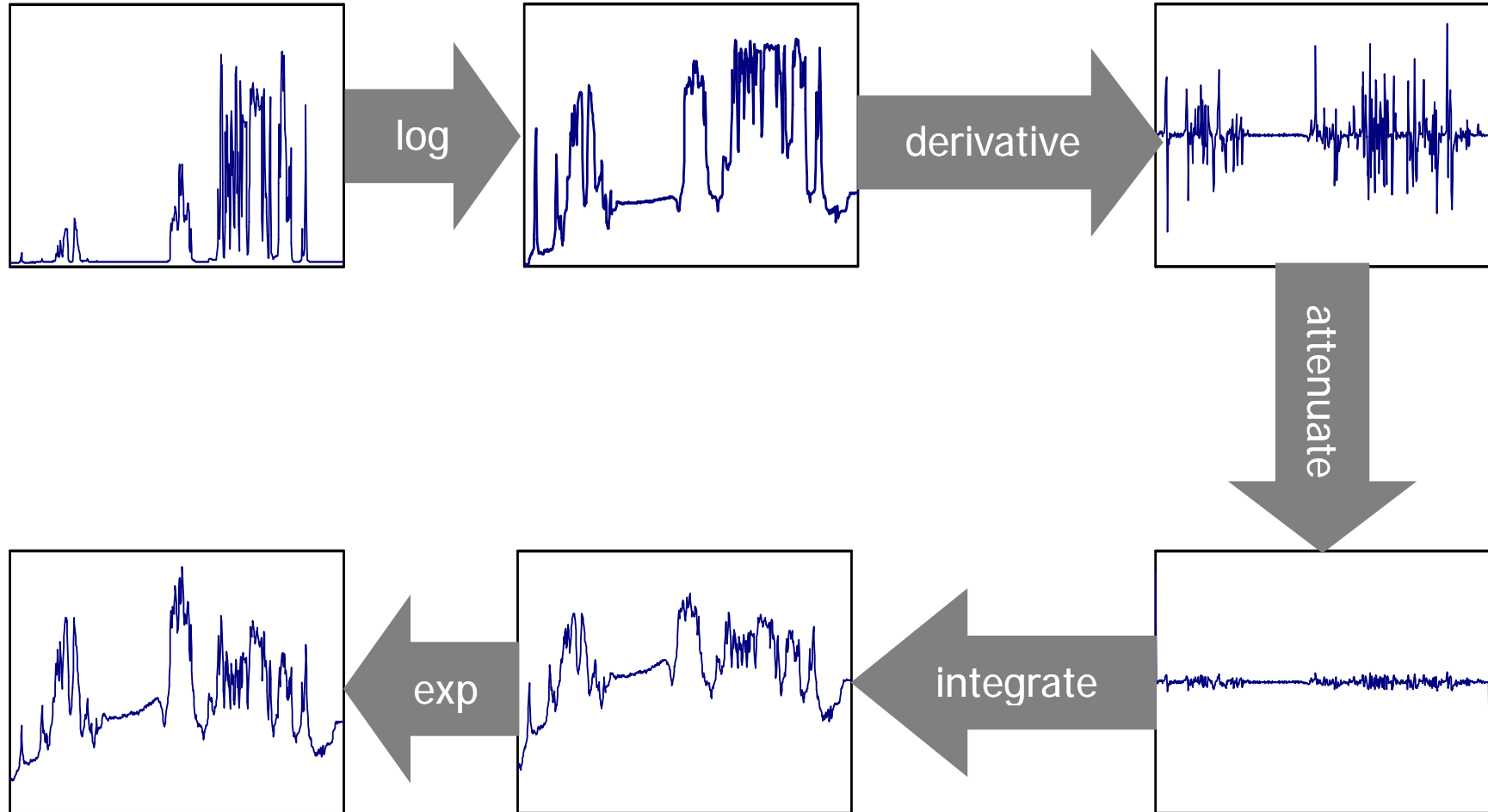
# Log domain

---

- 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*  $\Phi(\|\nabla H\|)$
- Compute an attenuated gradient field  $\mathbf{G}$ :

$$\mathbf{G}(x, y) = \nabla H(x, y) \cdot \Phi(\|\nabla H\|)$$

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


# Solution

---

- Look for image  $I$  with gradient closest to  $G$  in the least squares sense.

- $I$  minimizes the integral:  $\iint F(\nabla I, G) dx dy$

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

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



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

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

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

$$\begin{bmatrix} \dots & 1 & \dots & 1 & -4 & 1 & \dots & 1 & \dots \end{bmatrix} \begin{bmatrix} \mathbf{I} \end{bmatrix} = \begin{bmatrix} \end{bmatrix}$$

# Solving Poisson equation

---

- No analytical solution
- Multigrid method
- Conjugate gradient method

# Attenuation

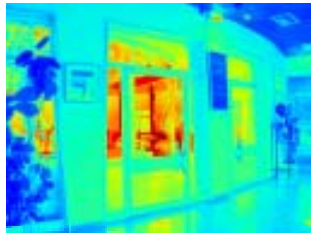
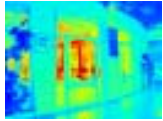
---

- 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

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

$\beta \sim 0.8$   
 $\alpha = 0.1 \overline{\nabla H}$



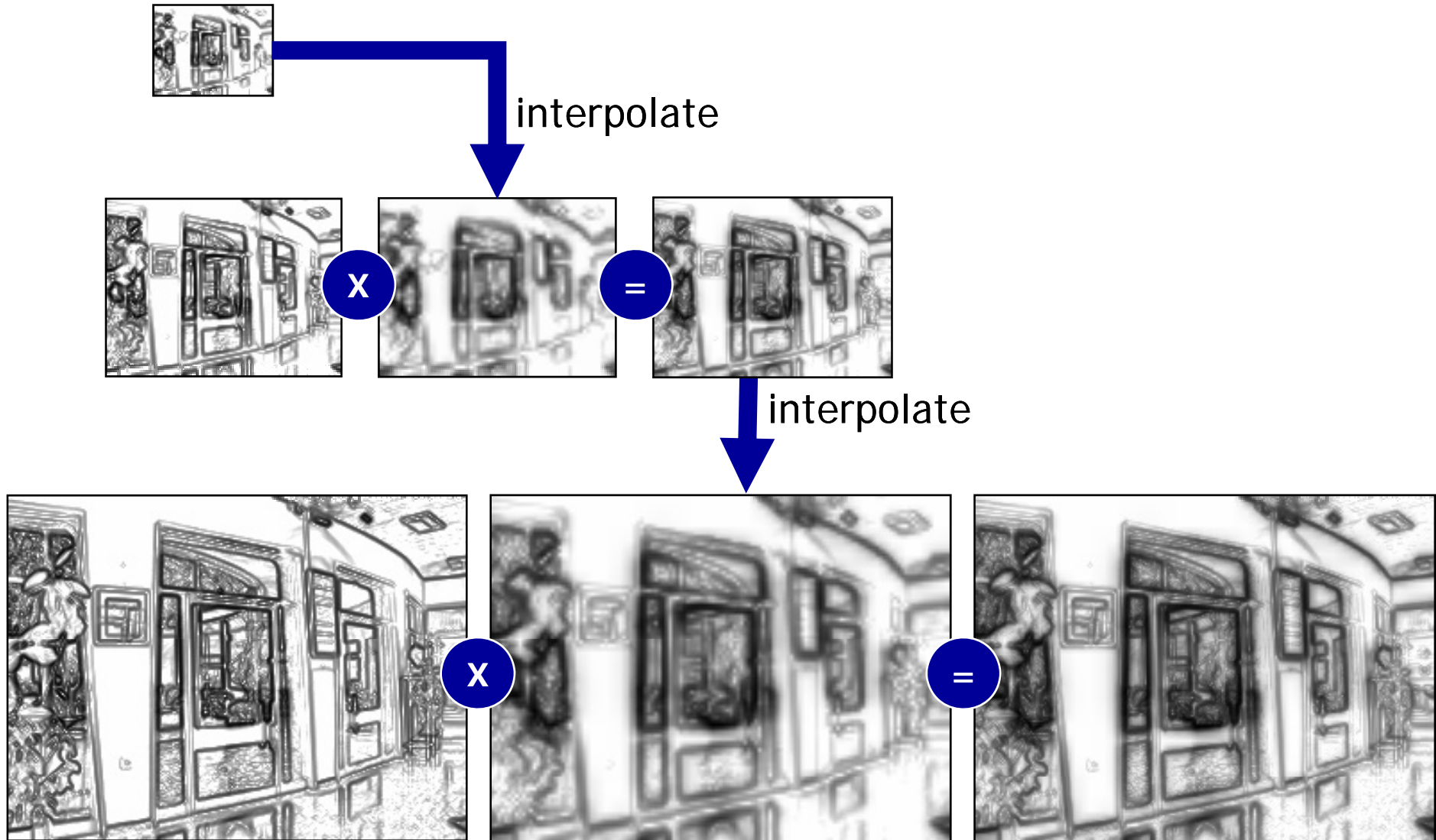
log(Luminance)

gradient magnitude

attenuation map

# Multiscale gradient attenuation

---



# Final gradient attenuation map

---



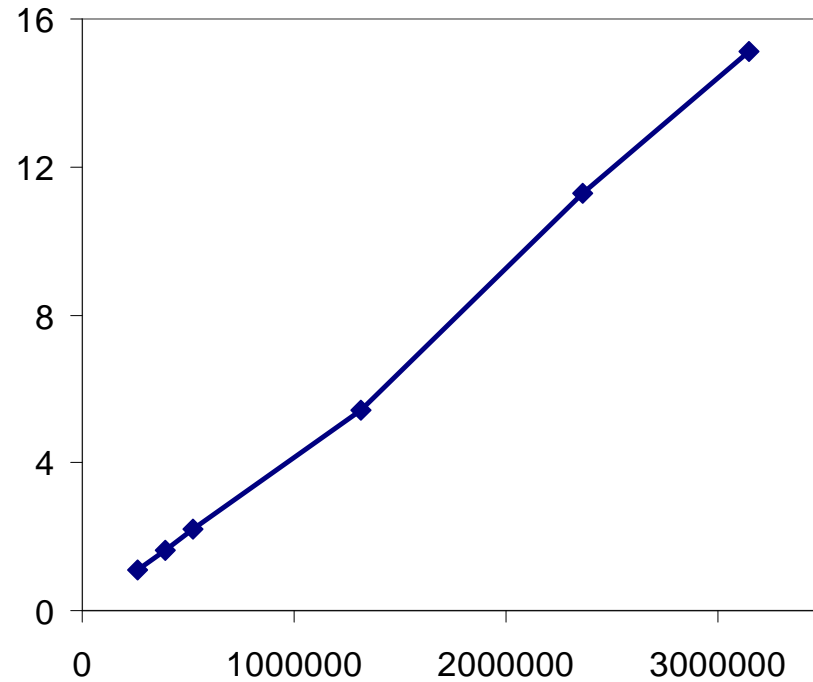


# Performance

---

- 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.

	<i>tmo</i> <sub>1</sub>	<i>tmo</i> <sub>2</sub>	<i>tmo</i> <sub>3</sub>	<i>tmo</i> <sub>4</sub>	<i>tmo</i> <sub>5</sub>	<i>tmo</i> <sub>6</sub>	Score
<i>tmo</i> <sub>1</sub>	-	1	0	0	1	1	3
<i>tmo</i> <sub>2</sub>	0	-	0	1	1	0	2
<i>tmo</i> <sub>3</sub>	1	1	-	1	1	1	5
<i>tmo</i> <sub>4</sub>	1	0	0	-	0	0	1
<i>tmo</i> <sub>5</sub>	0	0	0	1	-	1	2
<i>tmo</i> <sub>6</sub>	0	1	0	1	0	-	2

preference matrix (*tmo*<sub>2</sub>->*tmo*<sub>4</sub>, *tmo*<sub>2</sub> is better than *tmo*<sub>4</sub>)

# Statistical measurements

---

- Statistical measurements are used to evaluate:
  - Agreement: whether most agree on the ranking between two tone mapping operators.
  - Consistency: no cycle in ranking. If all are confused in ranking some pairs, it means they are hard to compare. If someone is inconsistent alone, his ranking could be dropped.

# Overall similarity

---

- Scene 8



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

# Summary

---

## *Overall Similarity: Color*

<i>I</i>	<i>P</i>	<i>H</i>	<i>A</i>	<i>L</i>	<i>B</i>
<u>3712</u>	<u>3402</u>	<u>2994</u>	<u>2852</u>	<u>1902</u>	<u>1696</u>

## *Bright Detail*

<i>I</i>	<i>A</i>	<i>P</i>	<i>H</i>	<i>B</i>	<i>L</i>
<u>823</u>	<u>688</u>	<u>569</u>	<u>549</u>	<u>474</u>	<u>347</u>

## *Dark Detail*

<i>P</i>	<i>A</i>	<i>I</i>	<i>L</i>	<i>H</i>	<i>B</i>
<u>815</u>	<u>793</u>	<u>583</u>	<u>491</u>	<u>485</u>	<u>283</u>



# Not settled yet!

---

- Some other experiment said bilateral are better than others.
- For your reference, photographic reproduction performs well in both reports.
- There are parameters to tune and the space could be huge.

# References

---

- 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.