Tone Mapping

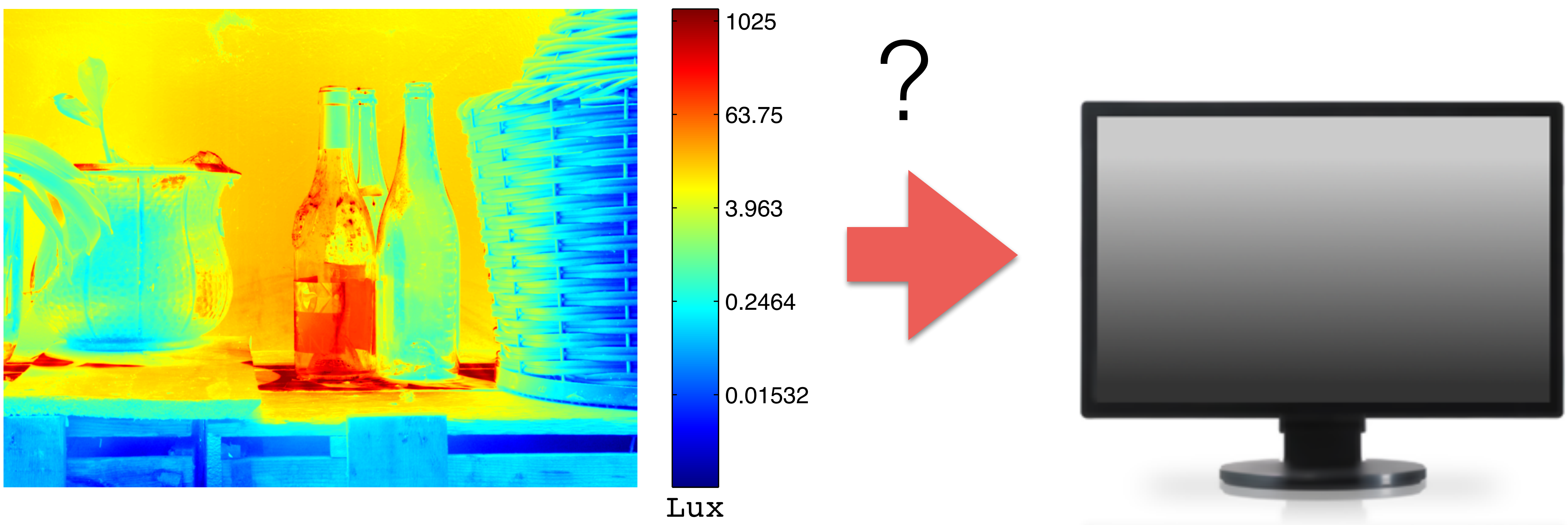
dr. Francesco Banterle

francesco.banterle@isti.cnr.it

Tone Mapping: HDR Visualization on LDR devices

- HDR monitors have started to appear in the market
  - very expensive (full HDR > $28,000 + VAT)
  - not ready for mobile (only TV sets or computer monitors)

Tone mapping



Tone Mapping

$$f(I) : \mathbb{R}^{w \times h \times c} \rightarrow \mathbb{D}^{w \times h \times c}$$

$$\mathbb{D} \subseteq [0, 255]$$

This means to compress the range

Tone Mapping

$$L_\mathrm{d} = f(L_\mathrm{w}) : \mathbb{R}^{w\times h} \to \mathbb{D}^{w\times h}$$

$$\begin{bmatrix} R_\mathrm{d} \\ G_\mathrm{d} \\ B_\mathrm{d} \end{bmatrix} = L_\mathrm{d}\, g\left(\frac{1}{L_\mathrm{w}} \begin{bmatrix} R_\mathrm{w} \\ G_\mathrm{w} \\ B_\mathrm{w} \end{bmatrix} \right)$$

Two steps:

- compress the luminance range
- fix colors

Tone Mapping: gamma encoding

- After tone mapping —> still real values in [0,1]!
- What to do?
  - Apply gamma correction or sRGB:

$$C = C^{\frac{1}{2.2}}$$

or

$$C_{\text{sRGB}} = \begin{cases} 12.92 C_{\text{linear}} & \text{if } C_{\text{linear}} \leq 0.0031308 \\ (1 + 0.055) C_{\text{linear}}^{\frac{1}{2.4}} - 0.055 & \text{otherwise} \end{cases}$$

  - Quantize values in [0,255] for classic 8-bit

Tone Mapping

- There are many tone mapping operators (TMOs)
  - more than 100!
- They can have different goals:
  - to match HVS perception
  - follow photography principles
  - quantization based

Tone Mapping: Taxonomy

- Global operators
- Local operators
- Frequency operators
- Segmentation operators

Global operators

- The same f to all pixels in the luminance channel
- f inputs:
  - current pixel luminance value
  - global statistics of the luminance channel:
    - Maximum (99-th percentile)
    - (Geometric) Average
    - Minimum (1-st percentile)
    - Histogram

Global operators: Geometric Average

$$L_{\mathrm{H}} = \prod_{i=1}^{N} \big(L(\mathbf{x}_i) + \epsilon\big)^{\frac{1}{N}} =$$

$$= \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log\big(L(\mathbf{x}_i) + \epsilon\big)\right) \qquad \epsilon > 0$$

Global operators: linear exposure

- Linear exposure is the simplest method:

$$L_{\mathrm{d}} = e L_{\mathrm{w}}$$

- e can be:
  - the maximum value of the image
  - (geometric) mean
  - a value which maximizing well-exposed pixels

Global operators: linear exposure

Maximum Luminance

Global operators: linear exposure

Mean Luminance

Global operators: linear exposure

Best exposure via histogram

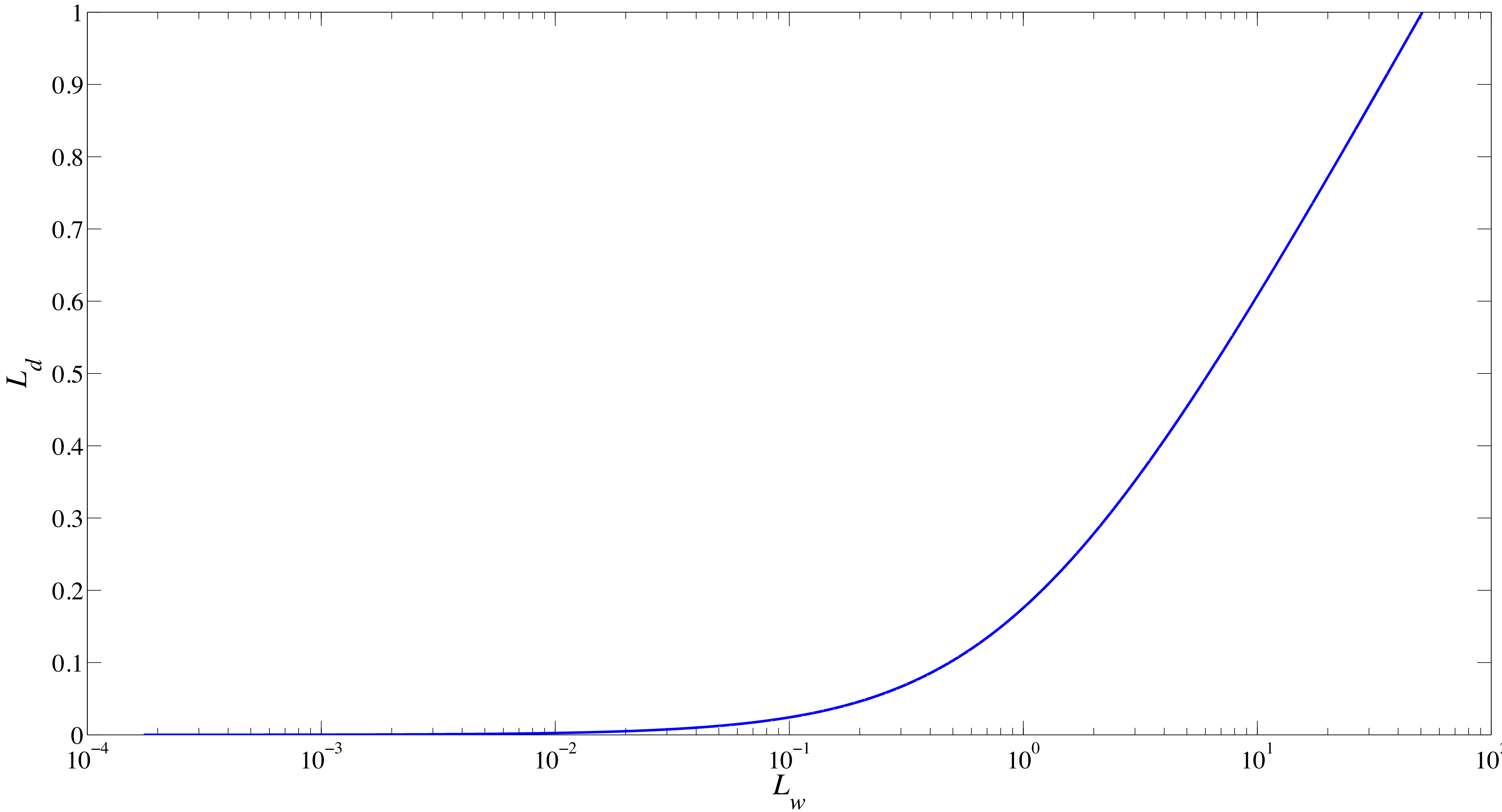
a non-linear mapping
is needed…

Global Operators: Logarithmic Operators

- **Idea**: to apply a logarithm
- which base? To use the maximum value
  - This maps values in [0,1]:

$$L_{\mathrm{d}} = \frac{\log(L_{\mathrm{w}} + 1)}{\log(L_{\mathrm{w,\ max}} + 1)}$$

Global Operators: Logarithmic Operators



Global Operators: Logarithmic Operators

Global Operators: Logarithmic Operators

Global Operators: Logarithmic Operators

- Although values are in [0,1]:
  - dark and mid pixels are pushed down
    - very dark image; better than maximum value linear

Global Operators: Logarithmic Operators

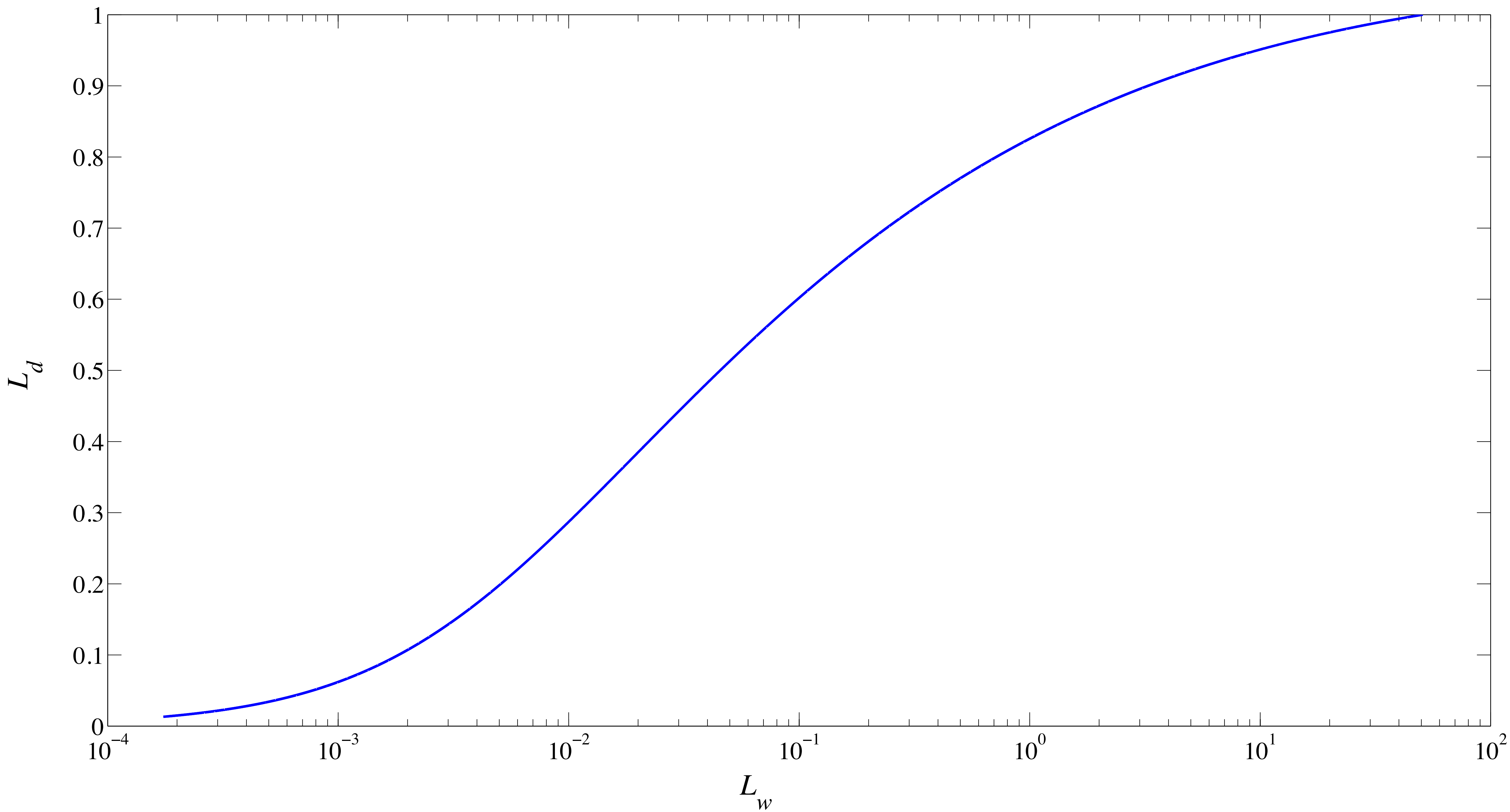
- A better idea: to vary the logarithm base depending on the luminance value —> adaptive logarithmic mapping [Drago et al. 2003]:

$$L_d(\mathbf{x}) = \frac{L_{d,\max}}{100\log_{10}(L_{s,\max})} \cdot \frac{\log(L_s(\mathbf{x})+1)}{\log\left(2+8\left(\frac{L_s(\mathbf{x})}{L_{s,\max}}\right)^{\log_{\frac{1}{2}} b}\right)}$$

$$L_s(\mathbf{x}) = \frac{L_w(\mathbf{x})}{\overline{L}_w}$$

user parameter ⟶ $b \in [0.75, 1]$

Global Operators: Logarithmic Operators



Global Operators: Logarithmic Operators

Global Operators: Logarithmic Operators

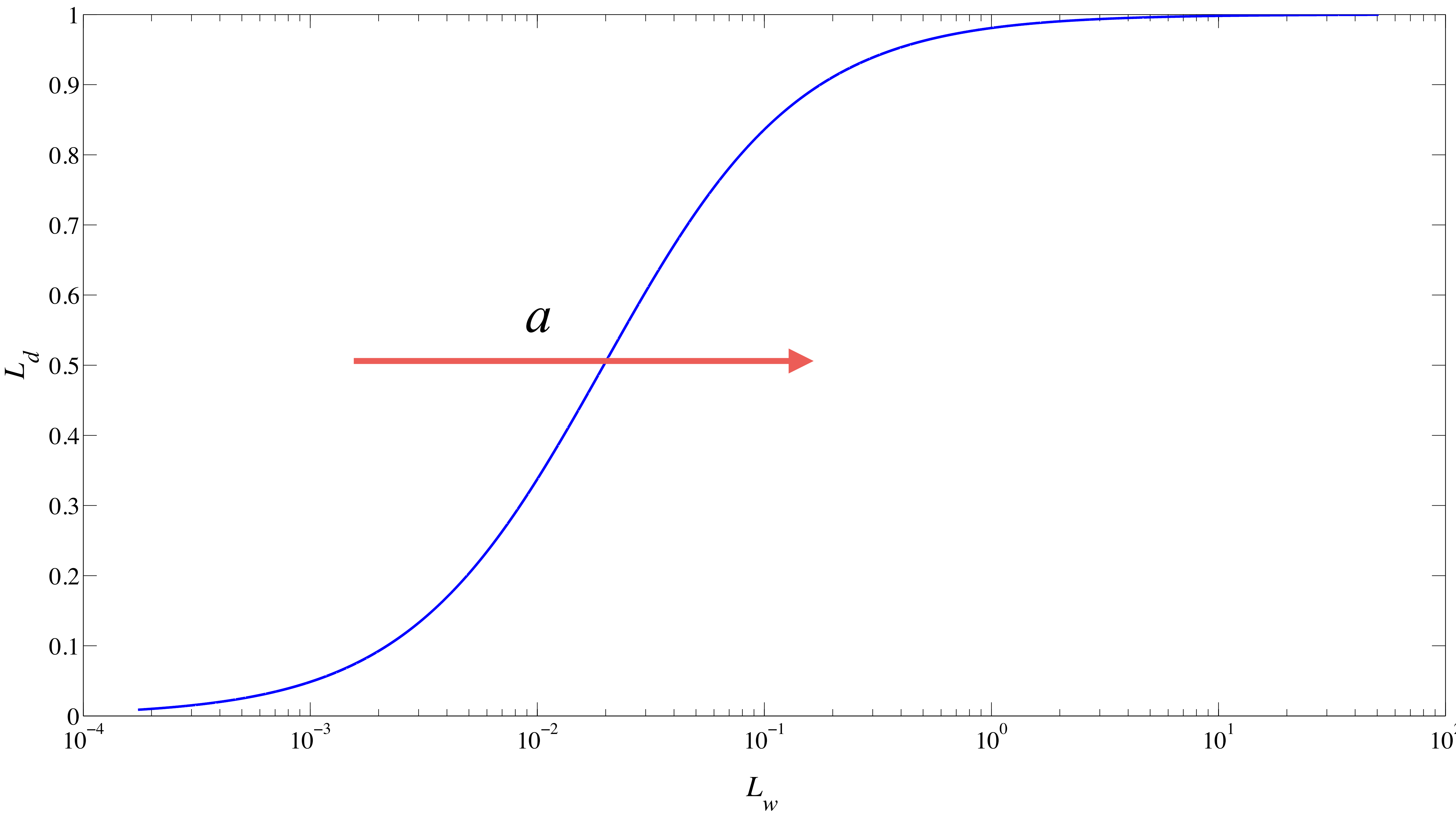
Global Operators: Sigmoid Operators

- Another popular option, which mimics response of rods and cones [Reinhard et al. 2002]:

$$L_d(\mathbf{x}) = \frac{L_m(\mathbf{x})}{L_m(\mathbf{x}) + 1} \qquad L_m(\mathbf{x}) = \frac{aL_w(\mathbf{x})}{\overline{L}_w}$$

$$L_d(\mathbf{x}) = \frac{L_m(\mathbf{x})\left(1 + L_{\text{white}}^{-2} L_m(\mathbf{x})\right)}{L_m(\mathbf{x}) + 1} \qquad L_m(\mathbf{x}) = \frac{aL_w(\mathbf{x})}{\overline{L}_w}$$

Global Operators: Sigmoid Operators



Global Operators: Sigmoid Operators varying a

Global Operators:
Sigmoid Operators varying a

Global Operators: Sigmoid Operators varying white

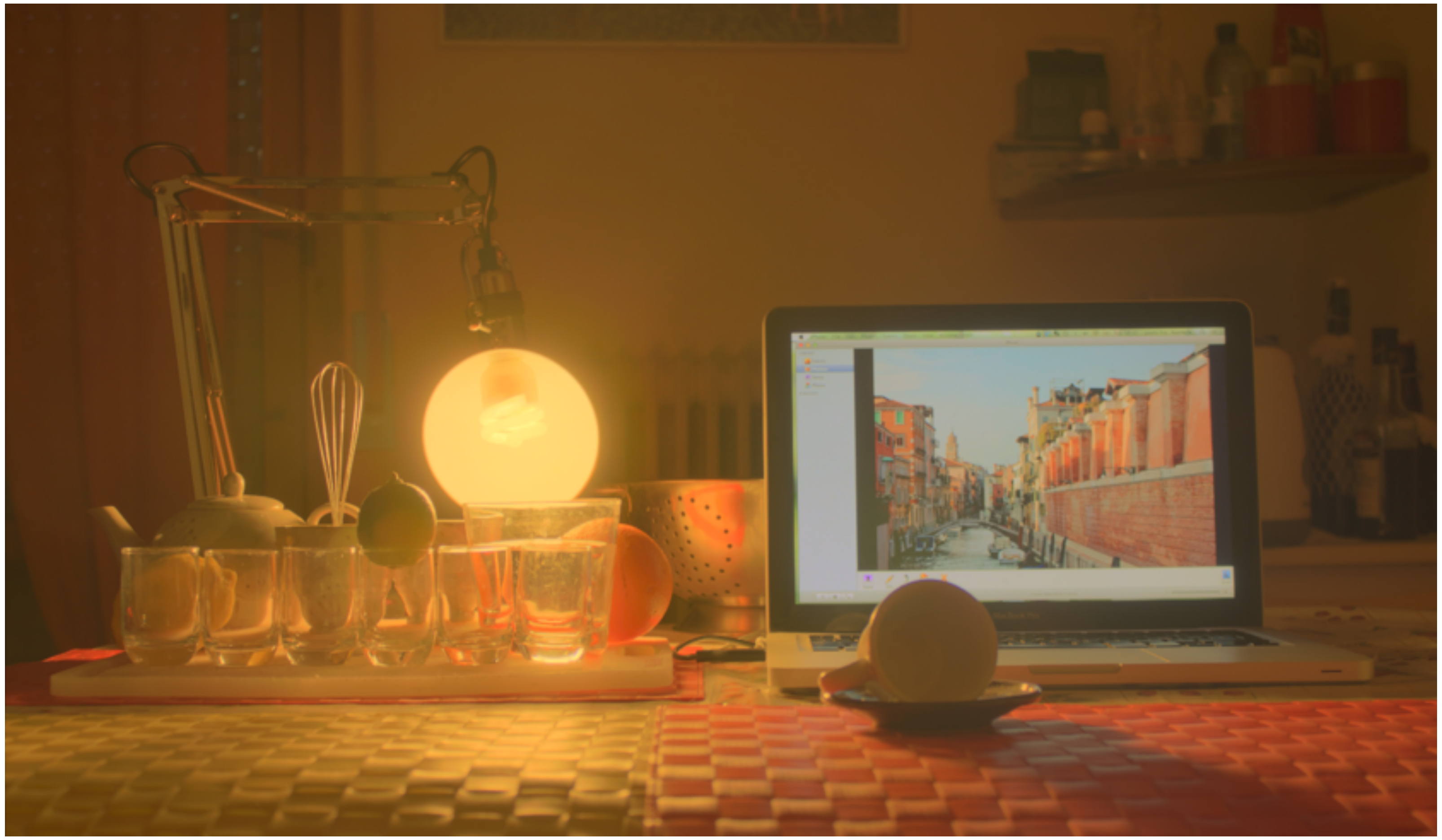
Global Operators: Sigmoid Operators varying white

Global Operators

- There are many global operators:
  - based on the histogram
  - based on power functions
  - mixture of logarithm, linear, power, sigmoid, etc…

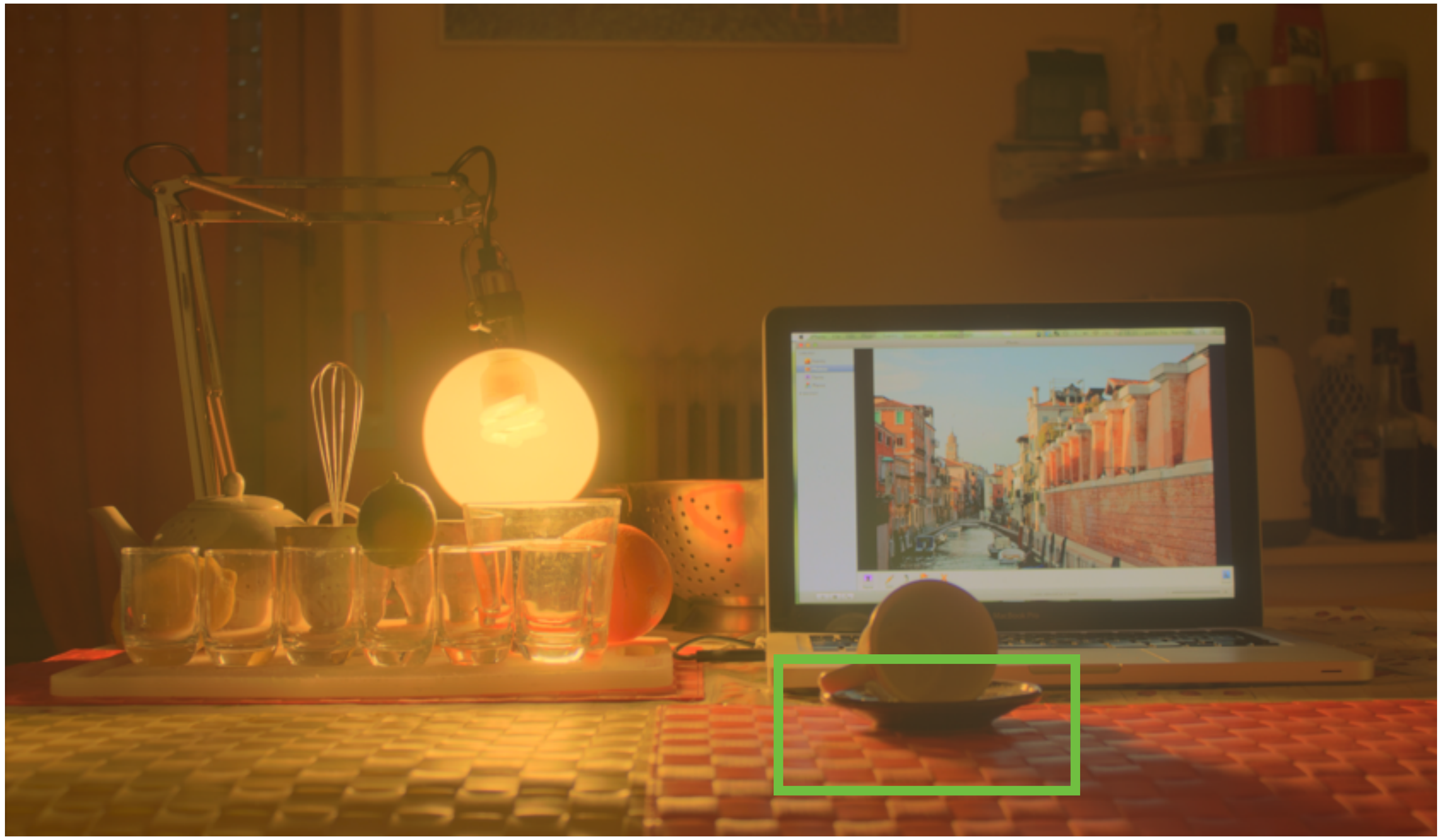
Local Operators

- Global operators preserve well global contrast
- Local contrast may be not preserved!



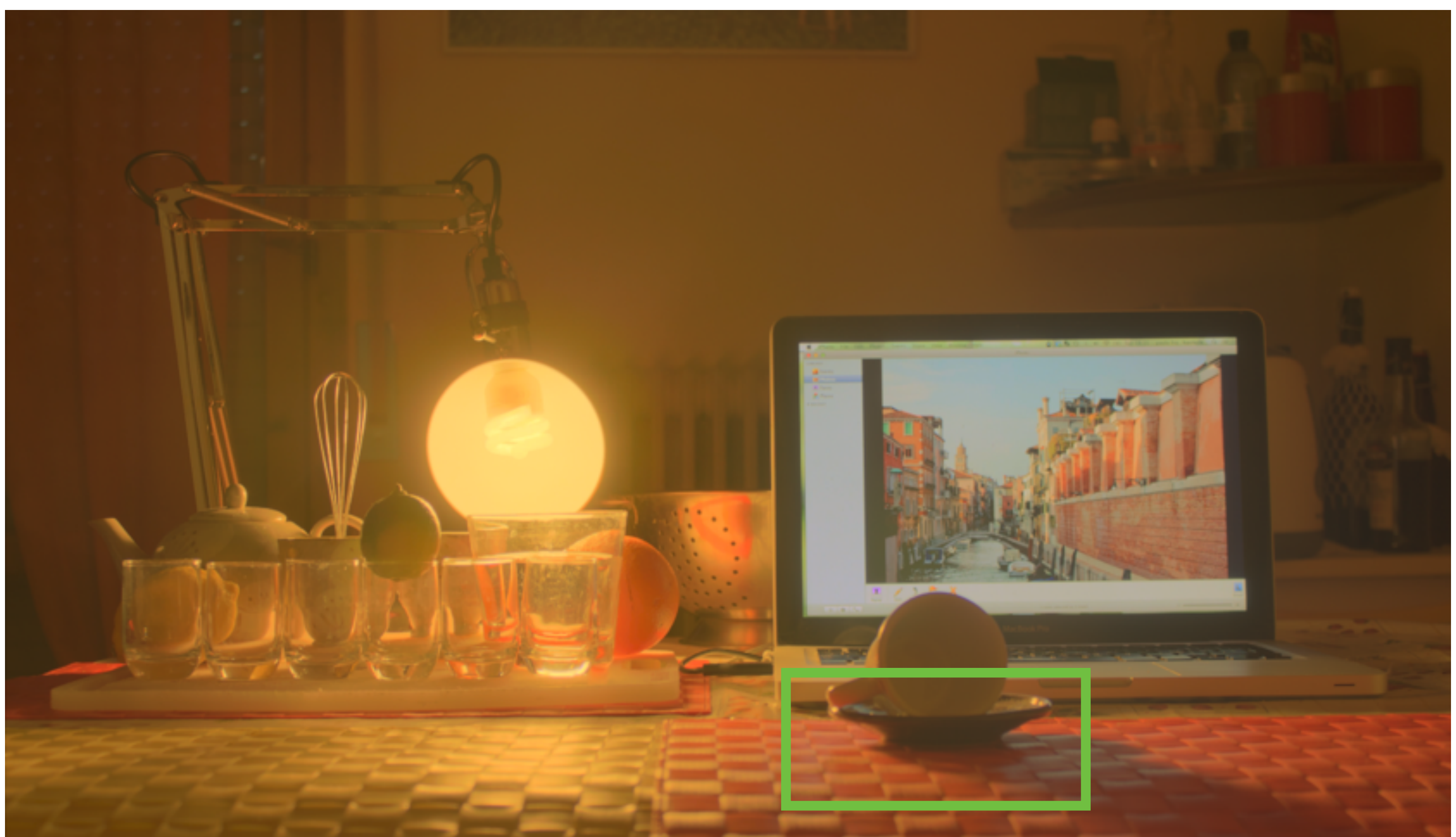
Local Operators

- Global operators preserve well global contrast
- Local contrast may be not preserved!



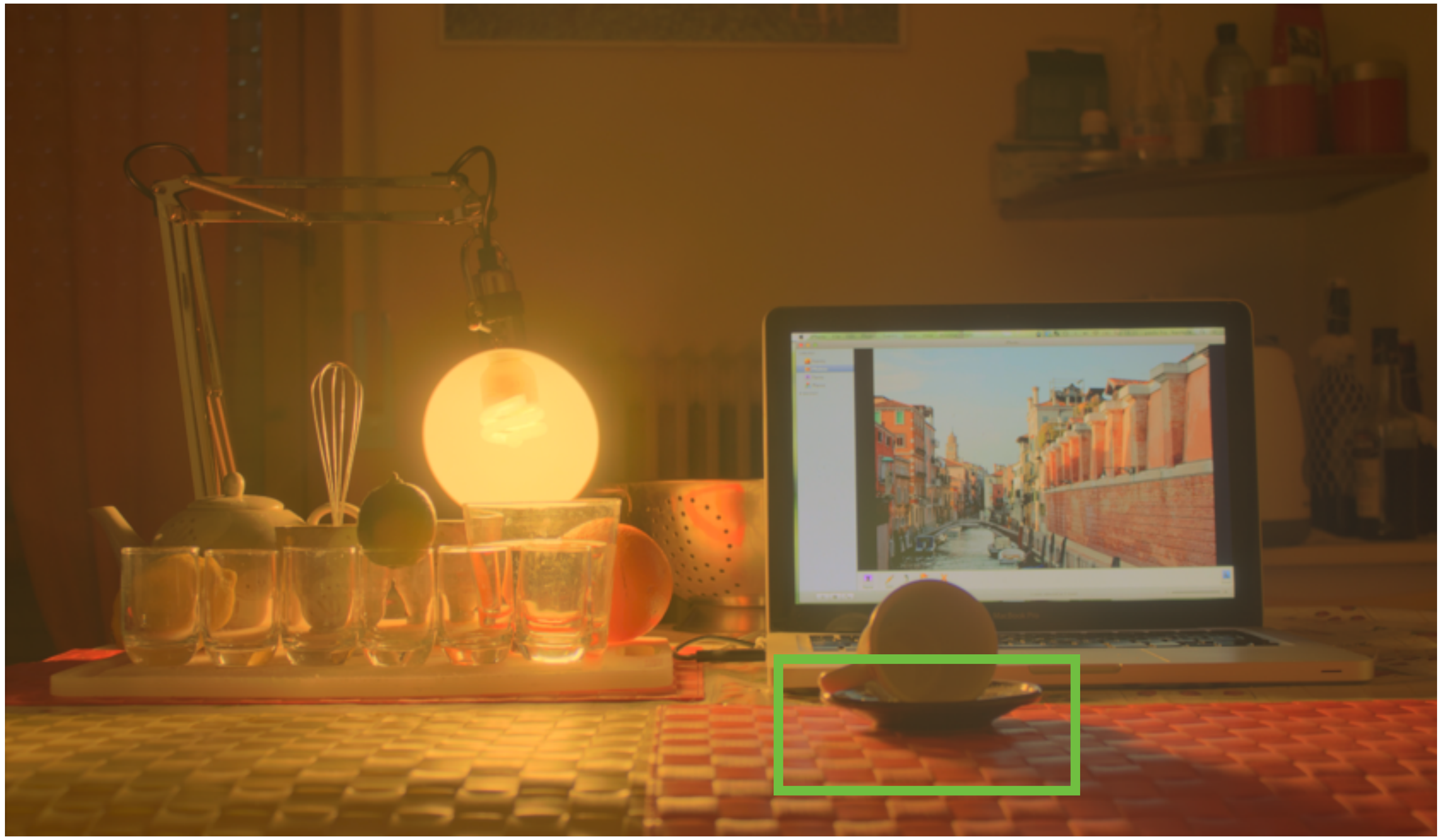
Local Operators

- Global operators preserve well global contrast
- Local contrast may be not preserved!



Local Operators

- Global operators preserve well global contrast
- Local contrast may be not preserved!



Local Operators

- f varies for each pixel
- f inputs:
  - current pixel luminance value
  - global statistics of the luminance channel
  - local statistics of the luminance channel:
    - computed around a neighborhood of the current pixel

Local Operators

- Local sigmoid:

$$L_d(\mathbf{x}) = \frac{L_m(\mathbf{x})}{L_m(\mathbf{x}) + 1} \qquad L_m(\mathbf{x}) = \frac{a L_w(\mathbf{x})}{\overline{L}_w}$$

$$\hat{L}_m(\mathbf{x}) = \frac{\sum_{i=-n}^{n} \sum_{j=-n}^{n} w(\Delta_{ij}) L_m(\mathbf{x} + \Delta_{ij})}{\sum_{i=-n}^{n} \sum_{j=-n}^{n} w(\Delta_{ij})}$$

$$w(\Delta_{ij}) = 1 \qquad w(\Delta_{ij}) = e^{-\frac{\|\Delta_{ij}\|^2}{2\sigma^2}}$$

Local Operators

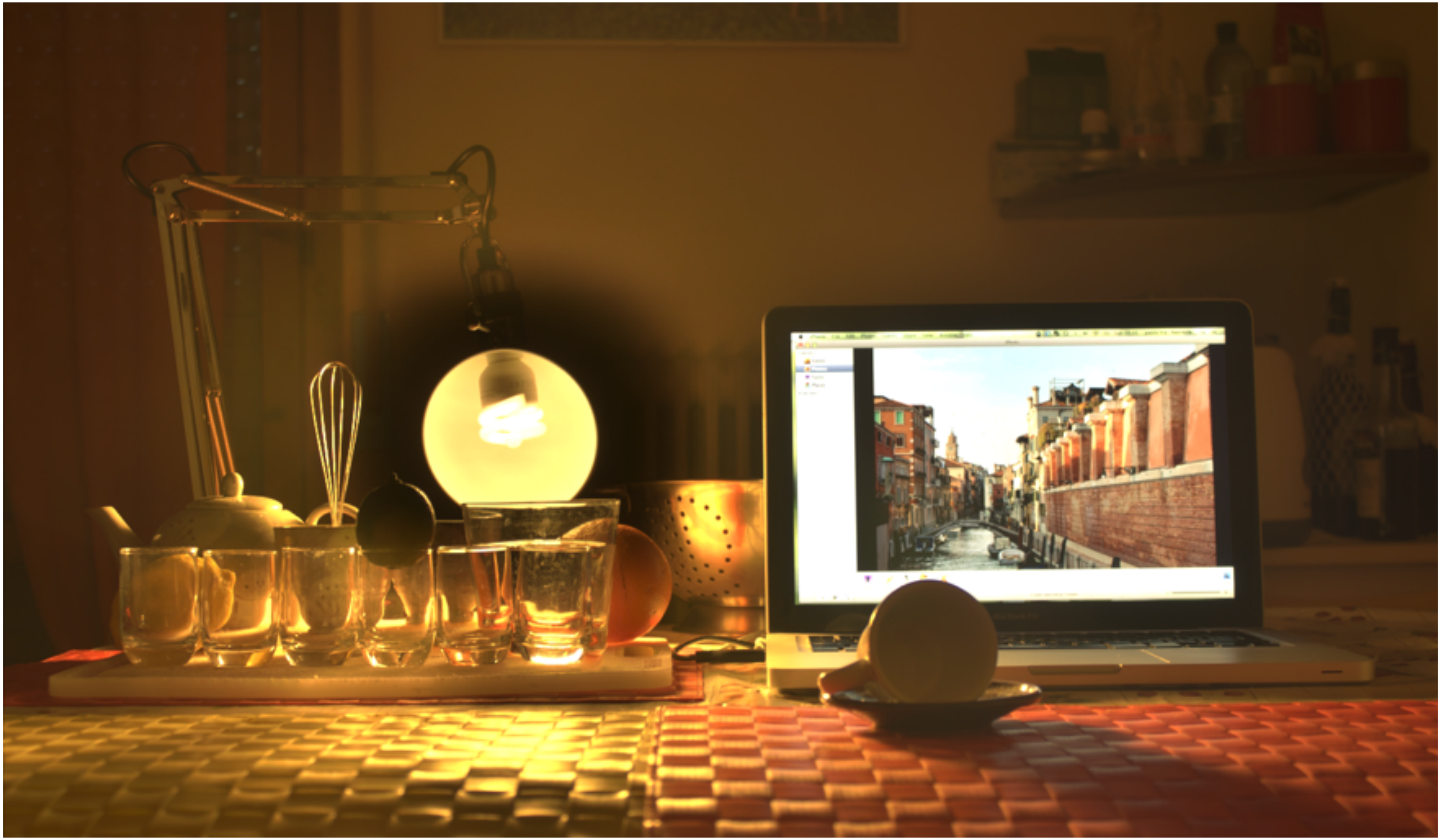
- Local sigmoid:

$$L_d(\mathbf{x}) = \frac{L_m(\mathbf{x})}{L_m(\mathbf{x}) + 1} \qquad L_m(\mathbf{x}) = \frac{aL_w(\mathbf{x})}{\overline{L}_w}$$

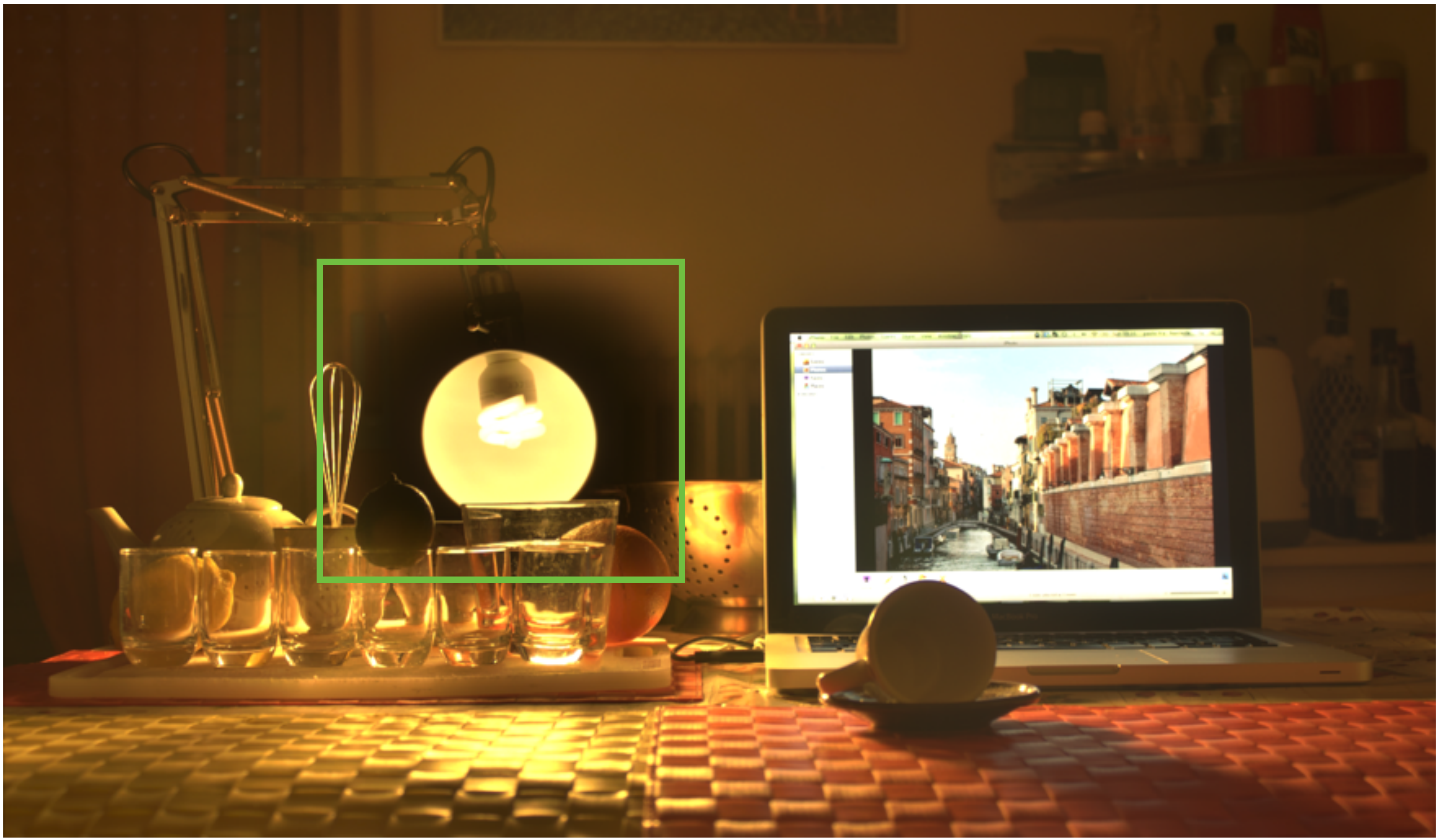
$$\hat{L}_m(\mathbf{x}) = \frac{\sum_{i=-n}^{n} \sum_{j=-n}^{n} w(\Delta_{ij}) L_m(\mathbf{x} + \Delta_{ij})}{\sum_{i=-n}^{n} \sum_{j=-n}^{n} w(\Delta_{ij})}$$

$$w(\Delta_{ij}) = 1 \qquad w(\Delta_{ij}) = e^{-\frac{\|\Delta_{ij}\|^2}{2\sigma^2}}$$

Local Operators



Local Operators



Local Operators

- Where halos?
  - around strong edges, e.g. proximity light sources
- Why halos?
  - There is bias in the statistics computations:
    - mixing areas with high and low luminance values

Local Operators

- How to avoid halos, i.e. bias?
  - Avoid linear filters: box, Gaussian, etc…
  - Use: edge-aware filters!

Local Operators: Bilateral Filter

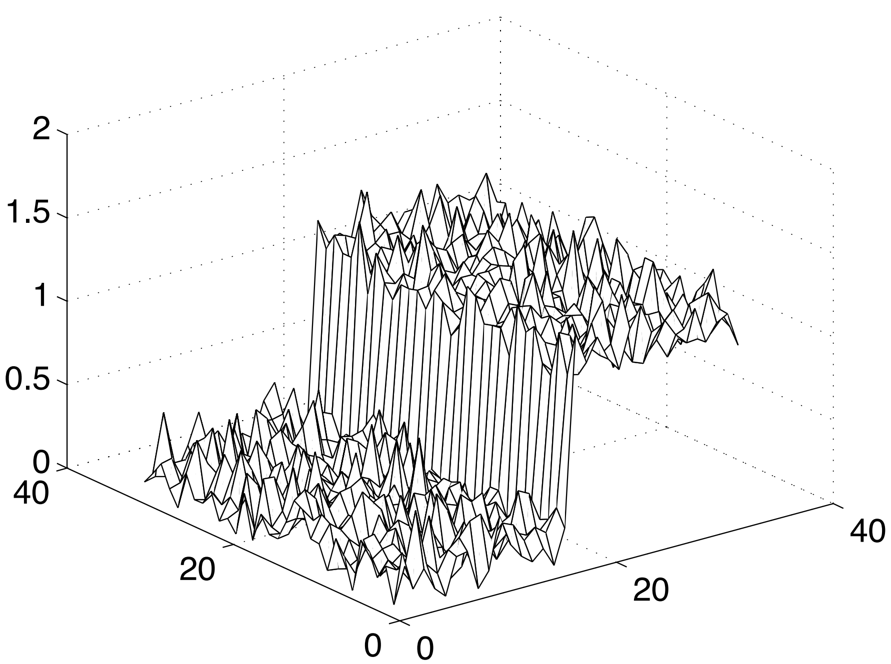
- There are many edge-aware filters. A popular and straightforward to implement is the Bilateral filter [Durand and Dorsey 2002]:

$$\hat{L}_m(\mathbf{x}) = \frac{\sum_{i=-n}^{n}\sum_{j=-n}^{n} f_s(\Delta_{ij}) g_r(\|L_m(\mathbf{x}) - L_m(\mathbf{x}+\Delta_{ij})\|) L_m(\mathbf{x}+\Delta_{ij})}{\sum_{i=-n}^{n}\sum_{j=-n}^{n} f_s(\Delta_{ij}) g_r(\|L_m(\mathbf{x}) - L_m(\mathbf{x}+\Delta_{ij})\|)}$$

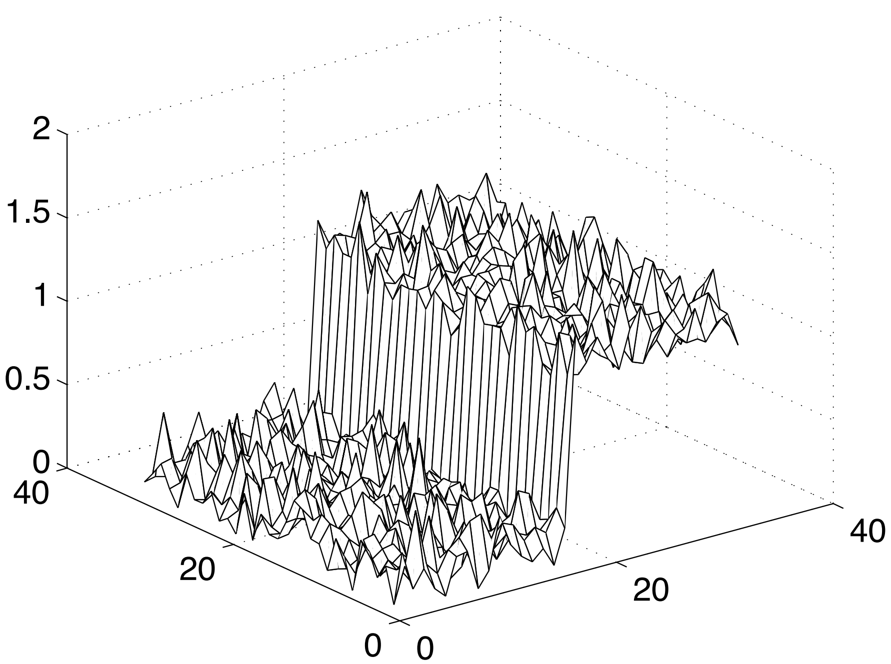
f_s spatial function

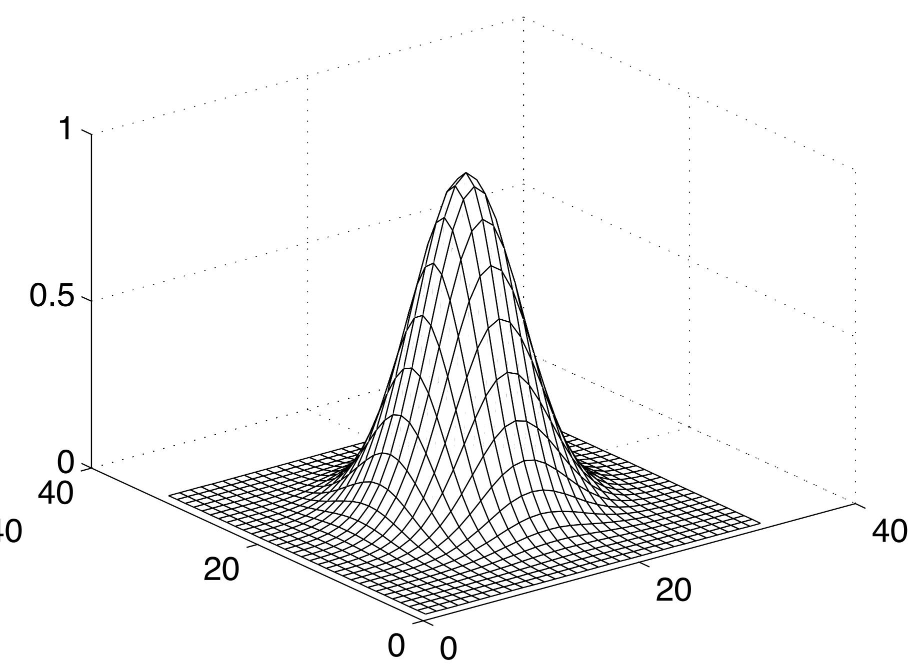
g_r range function

Local Operators: Bilateral Filter

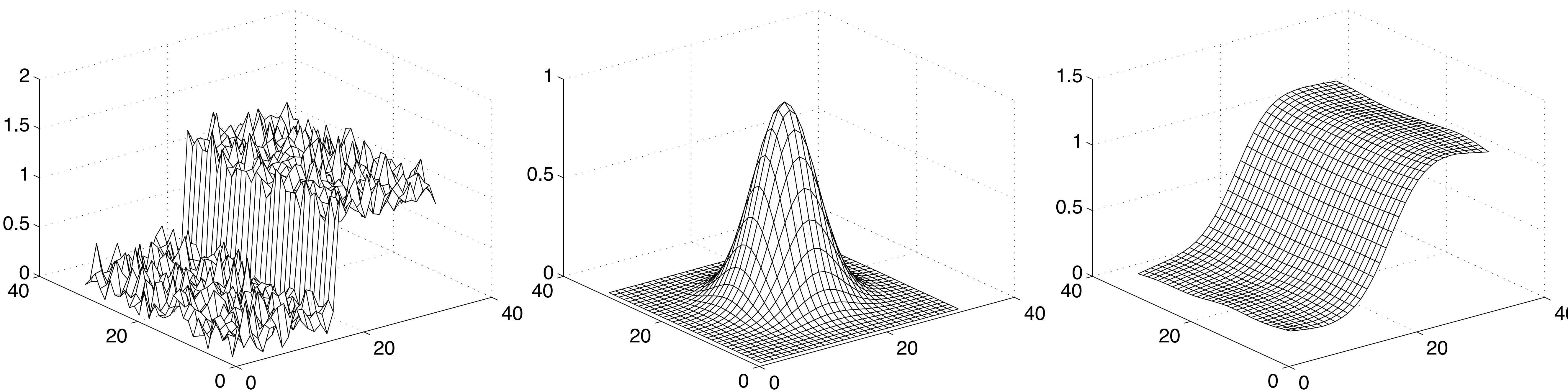


Local Operators: Bilateral Filter

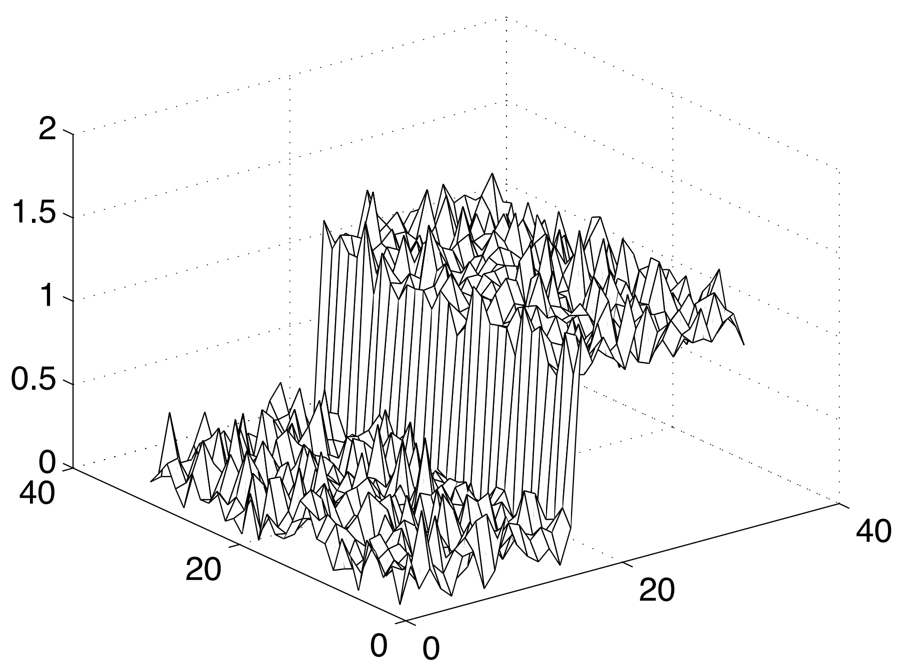




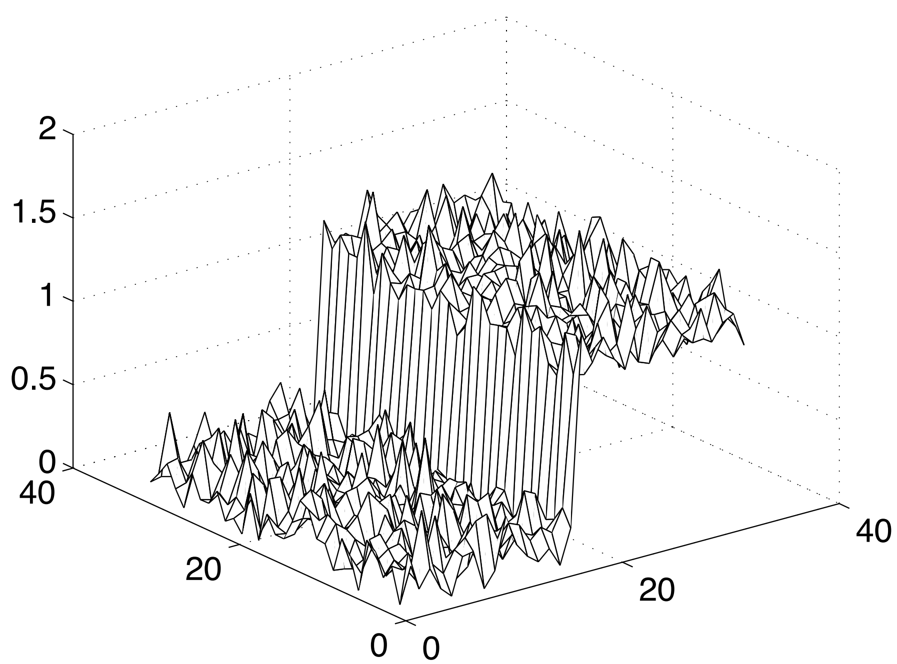
Local Operators: Bilateral Filter

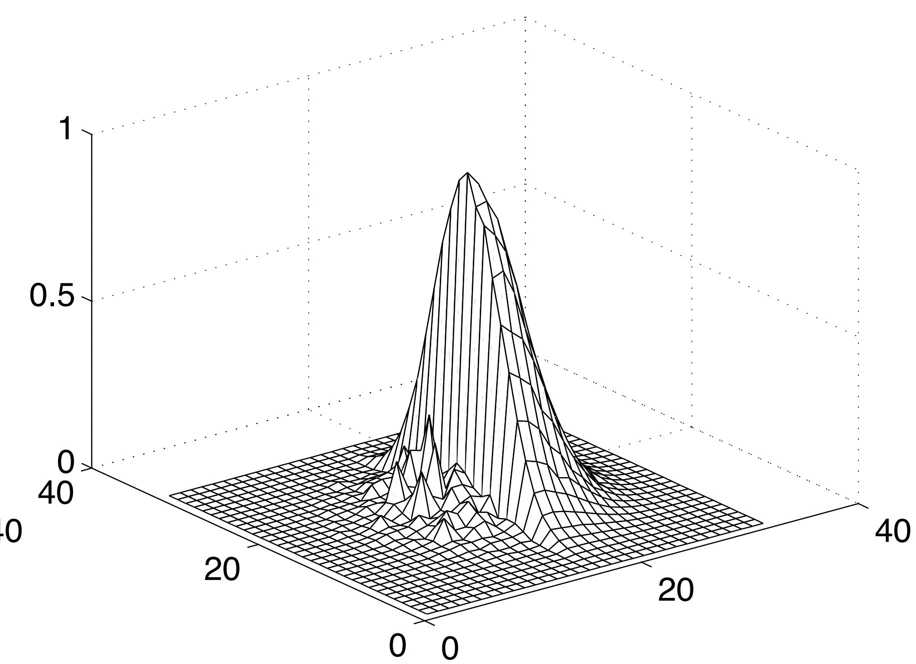


Local Operators: Bilateral Filter

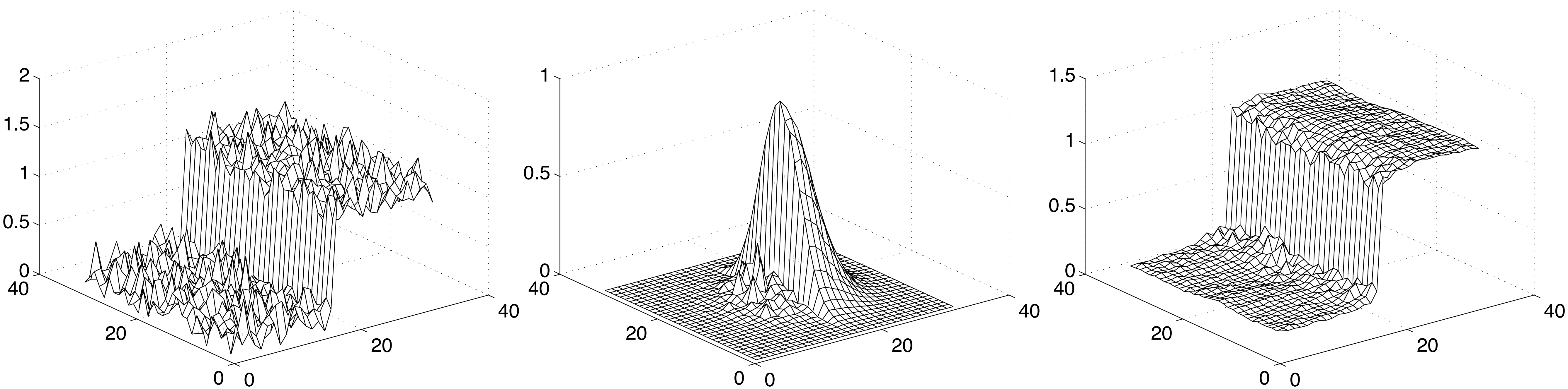


Local Operators: Bilateral Filter



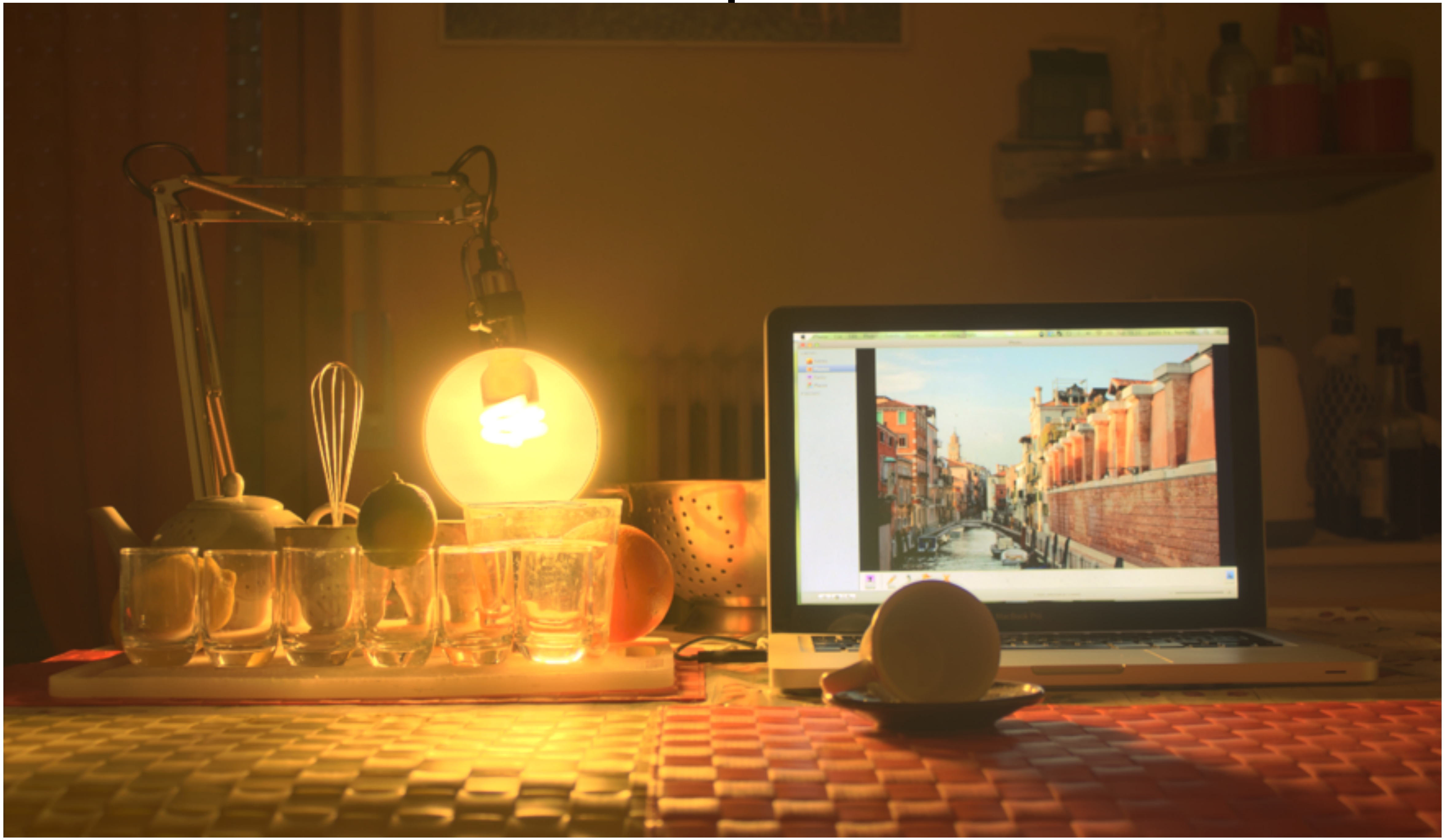


Local Operators: Bilateral Filter



Local Operators: Final Operator

Local Operators: Final Operator



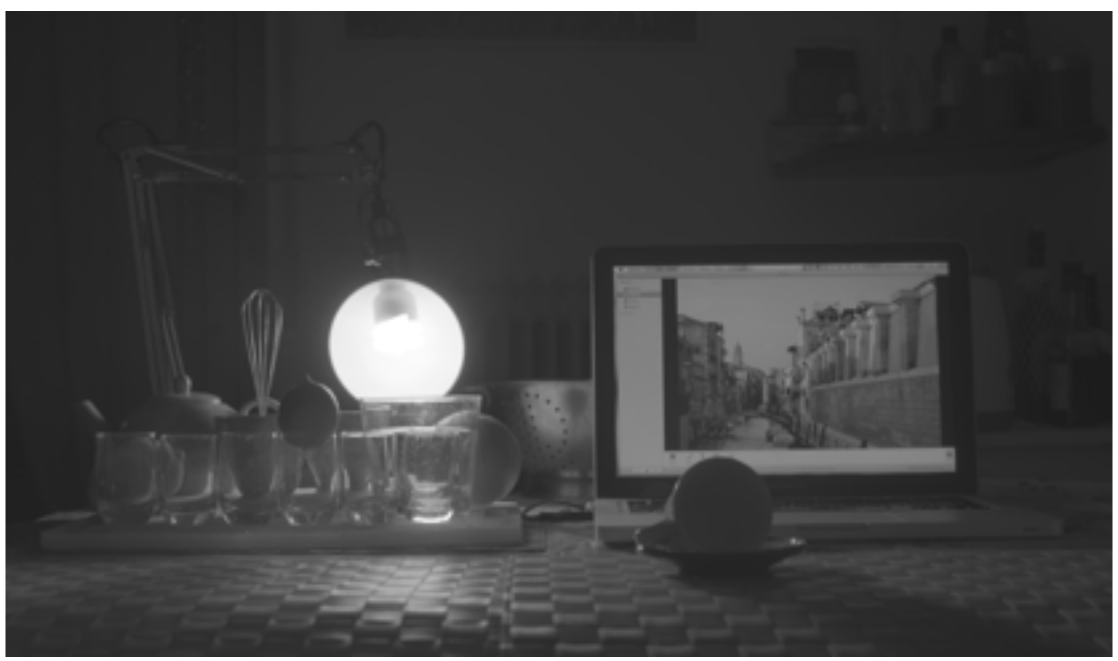
Segmentation Operators

- Segmentation operators:
  - The image is segmented into areas of uniform luminance
  - A TMO for each different area
  - Important to blend with weights different areas to avoid seams!

Frequency Operators

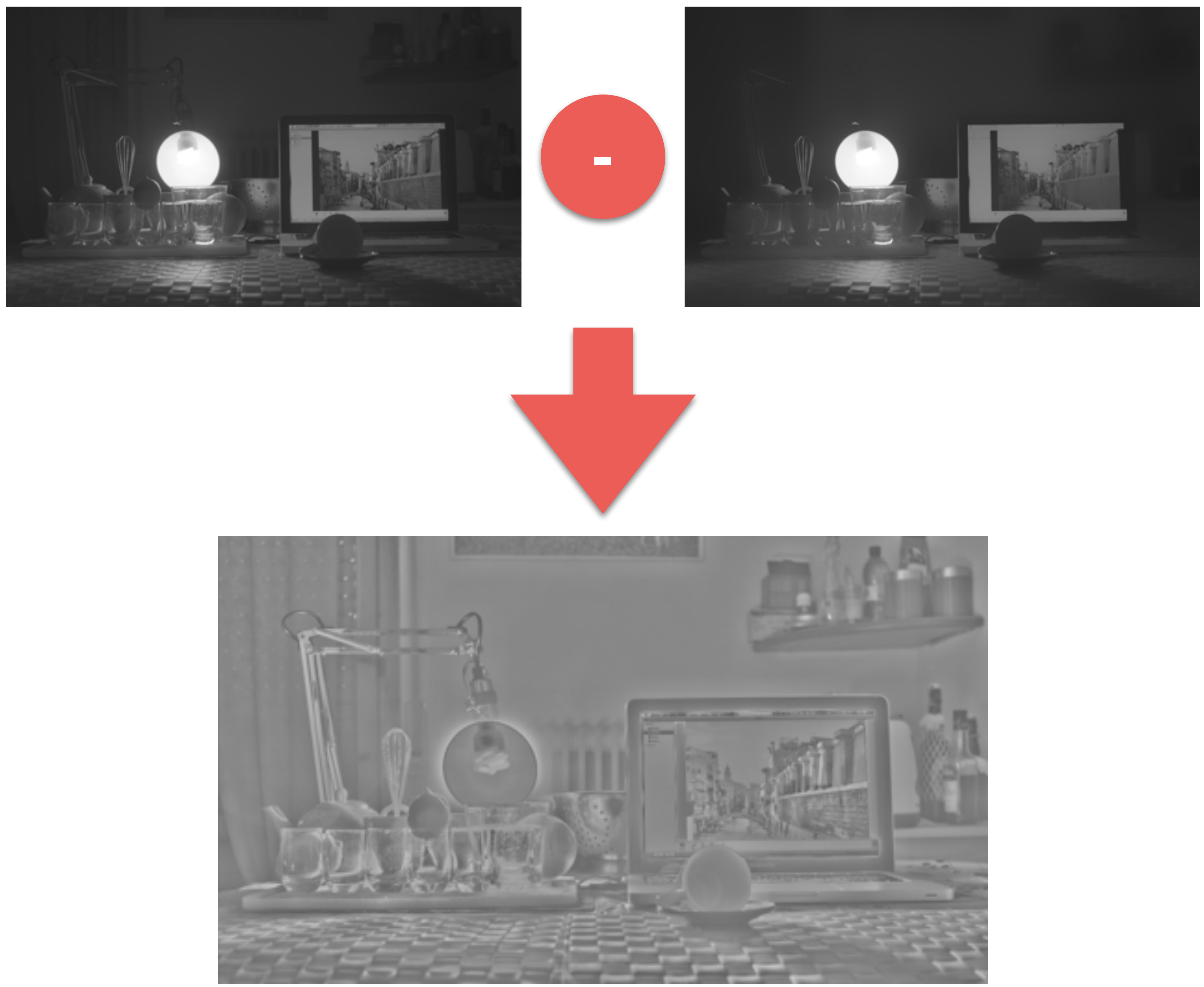
- Decompose the signal into different frequency
- Each frequency is appropriate scaled/tone mapped
- The signal is reconstructed

Frequency Operators



Input Luminance Filtering Filtered Image

Frequency Operators



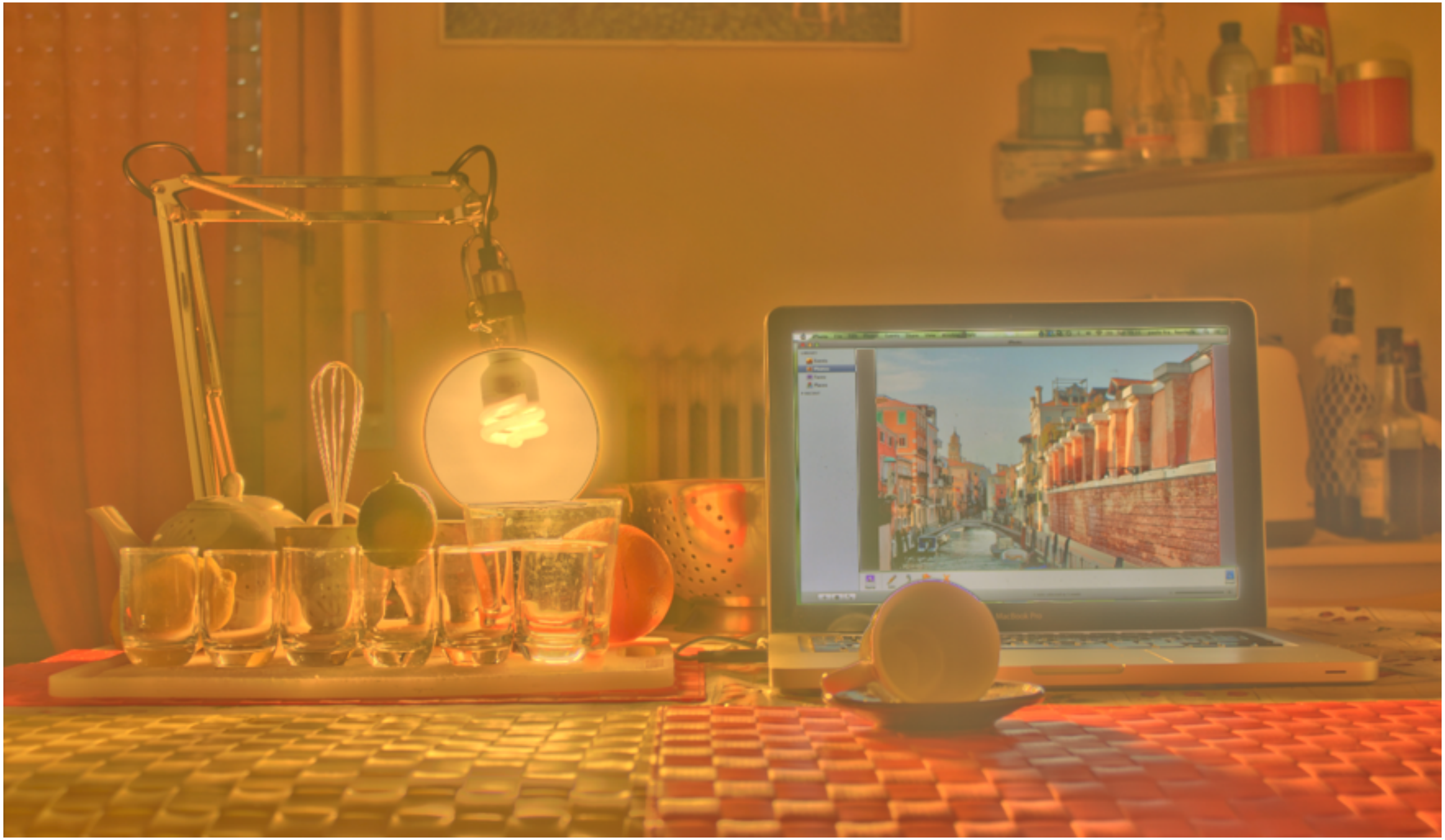
Frequency Operators

Base

Detail

Different TMOs for each layer

Frequency Operators



Messing with Colors

Color Reproduction in TMO

- When an HDR image is tone mapped colors change; they are more saturated
  - Why? Only luminance was reduced
    - Not really, if gamut has changed

Color Reproduction in TMO

- Basic idea is to desaturate colors; typically [Schlick 1994]:

$$\begin{bmatrix} R_d \\ G_d \\ B_d \end{bmatrix} = L_d \left(\frac{1}{L_w} \begin{bmatrix} R_w \\ G_w \\ B_w \end{bmatrix} \right)^s \qquad s \in (0, 1]$$

- s depends on the image content

- **Issues**: it needs manual tweaking and it is a hack

Color Reproduction in TMO

- Better approaches?
- A different desaturation [Mantiuk et al. 2009]:

$$\left(\left(\frac{1}{L_w}\begin{bmatrix}R_w\\G_w\\B_w\end{bmatrix}-1\right)p+1\right)L_d \qquad p=0.5 \qquad p\in[0,1]$$

- To work in color spaces such as IPT and LCh and restore saturation values; given original HDR and tone mapped images [Pouli et al. 2013]

Color Reproduction in TMO

Color Reproduction in TMO

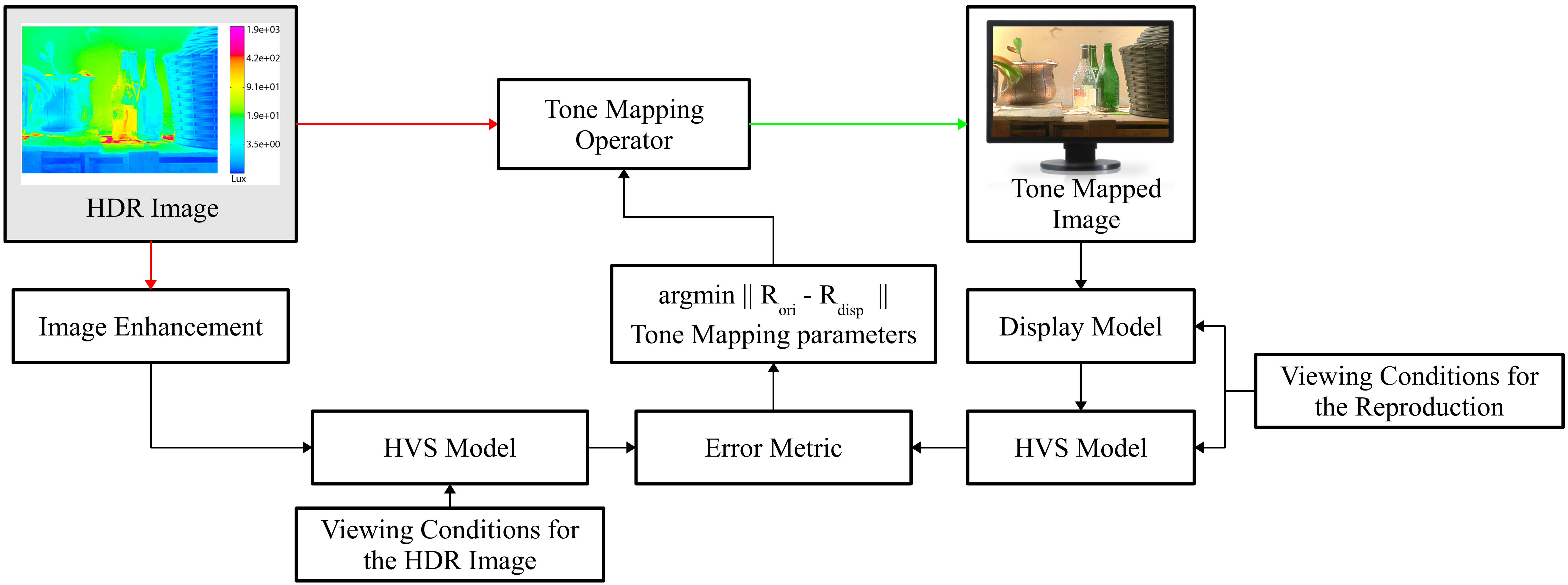
Generic TMO

- **Idea**: a complete TMO has three main steps which can be generalized [Mantiuk and Seidel 2008]
- Tone curve: $$L_d = TC(L_w) = \begin{cases} 0 & \text{if } L' \le b - d_l \\ 0.5c\frac{L'-b}{1-a_l(L'-b)} + 0.5 & \text{if } b - d_l < L' \le b \\ 0.5c\frac{L'-b}{1+a_h(L'-b)} + 0.5 & \text{if } b < L' \le b + d_h \\ 1 & \text{if } L' > b + d_h \end{cases} \quad \text{where } L' = \log_{10} L_w$$
- Modulation Transfer Function: work in the FFT domain (particularly Cortex Transform) and select certain frequencies
- Color Correction: classic power transform
- Parameters? via minimization! $$\underset{b,c,d_l,d_h,s}{\arg\max} \sum_{k=1,2,3} \left| C_{LDR} - TC(C_{HDR}; b, c, d_l, d_h) \cdot R^s \right|^2$$

Adaptive TMO

- **Idea**: to compensate for [Mantiuk et al. 2008]:
  - type of displaying technology: paper, LCD, high-contrast LED+LCD, etc.
  - viewing conditions: dark room, bright office, daylight, etc.

Adaptive TMO



Evaluation of TMOs

TMOs Evaluation

- There are many TMOs, more than 100!
  - which’s the best overall?
  - which’s the best for certain viewing conditions?
  - which’s the best for certain images?

TMOs Evaluation

- Subjective evaluation:
  - running psychophysical experiments
- Objective evaluation:
  - running computational metrics based on the HVS

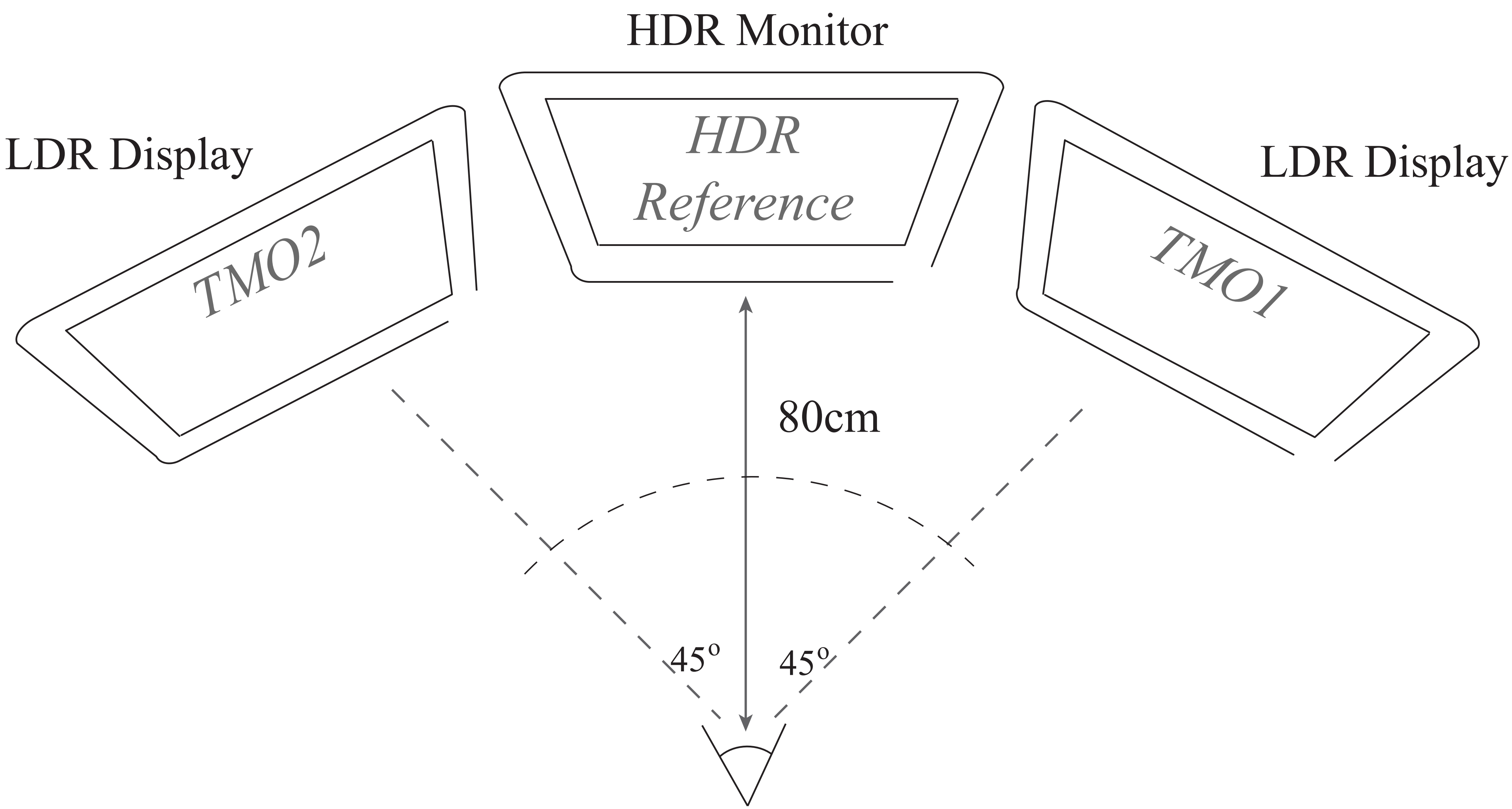
TMOs Evaluation: Subjective Evaluation

- Choose a methodology:
  - paired comparisons
  - ordering
  - ranking —> require learning

TMOs Evaluation: Subjective Evaluation

- Determine the number of subjects given the methodology:
  - ranking typically more subjects to reduce variance
  - a 20-30 subjects are typically OK
- Determine the number of images and type:
  - a good samples 8-10 images
  - covering different lighting conditions

TMOs Evaluation: Subjective Evaluation



TMOs Evaluation: Subjective Evaluation

- After data collection part:
  - determine if the data is statistically significant
  - determine coherency in the data
  - determine trends

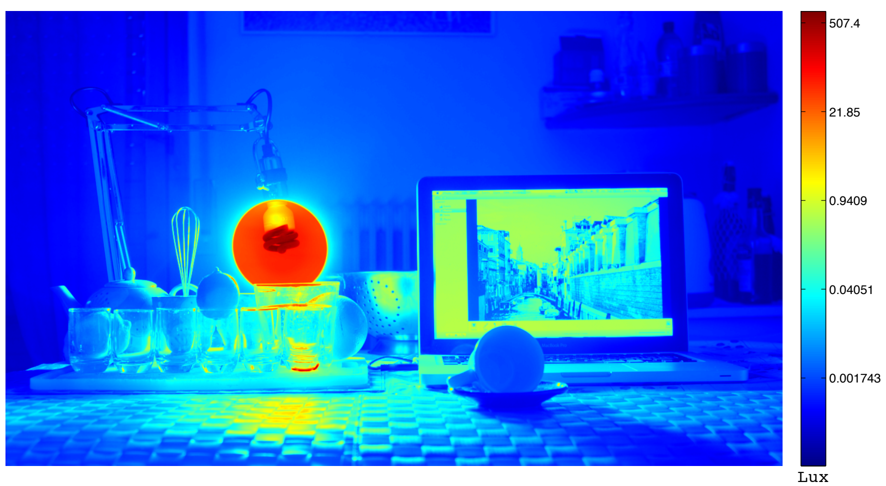
TMOs Evaluation: Objective Evaluation

- To use metrics, based on how HVS behaves and data acquired during experiments
- Typically PSNR and RMSE do not provide meaningful results!

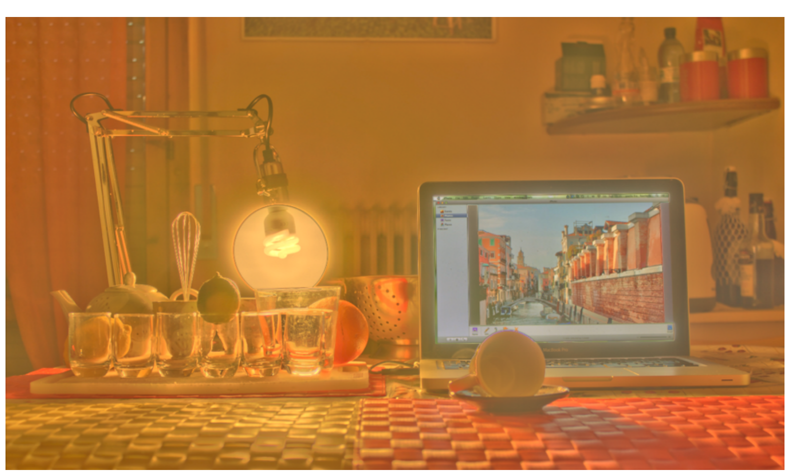
TMOs Evaluation: Objective Evaluation

- Perceptual HDR metric:
  - HDR VDP 2.2:
    - http://hdrvdp.sourceforge.net/wiki/
  - DRIIQM:
    - http://driiqm.mpi-inf.mpg.de
  - TMQI:
    - https://ece.uwaterloo.ca/~z70wang/research/tmqi/

TMOs Evaluation: Objective Evaluation

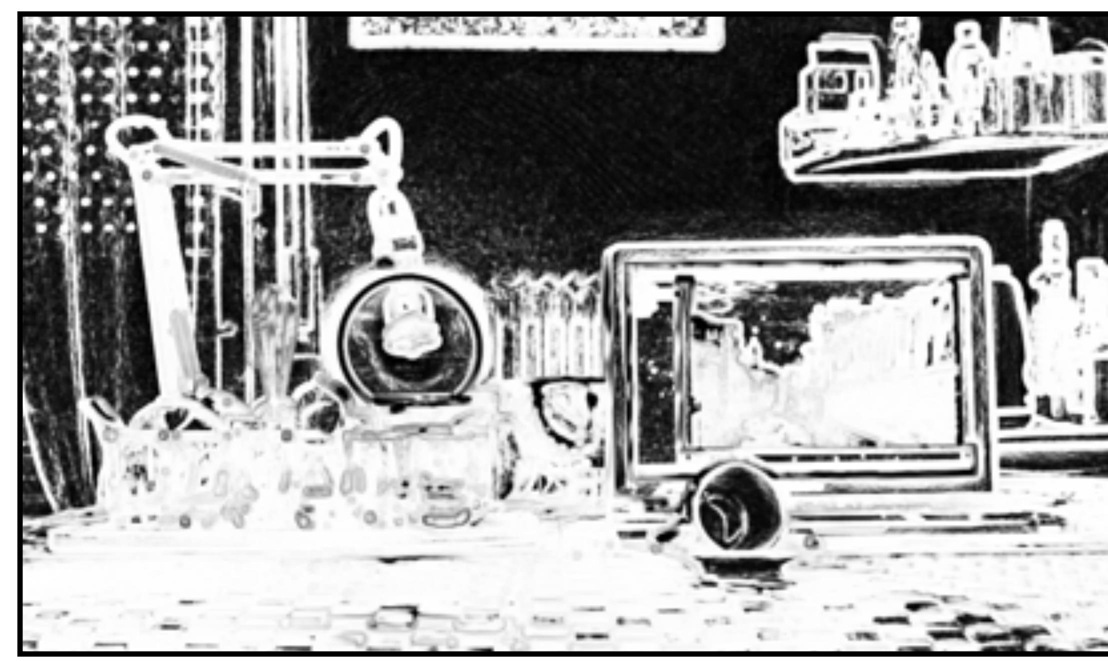


Reference Image



Test Image

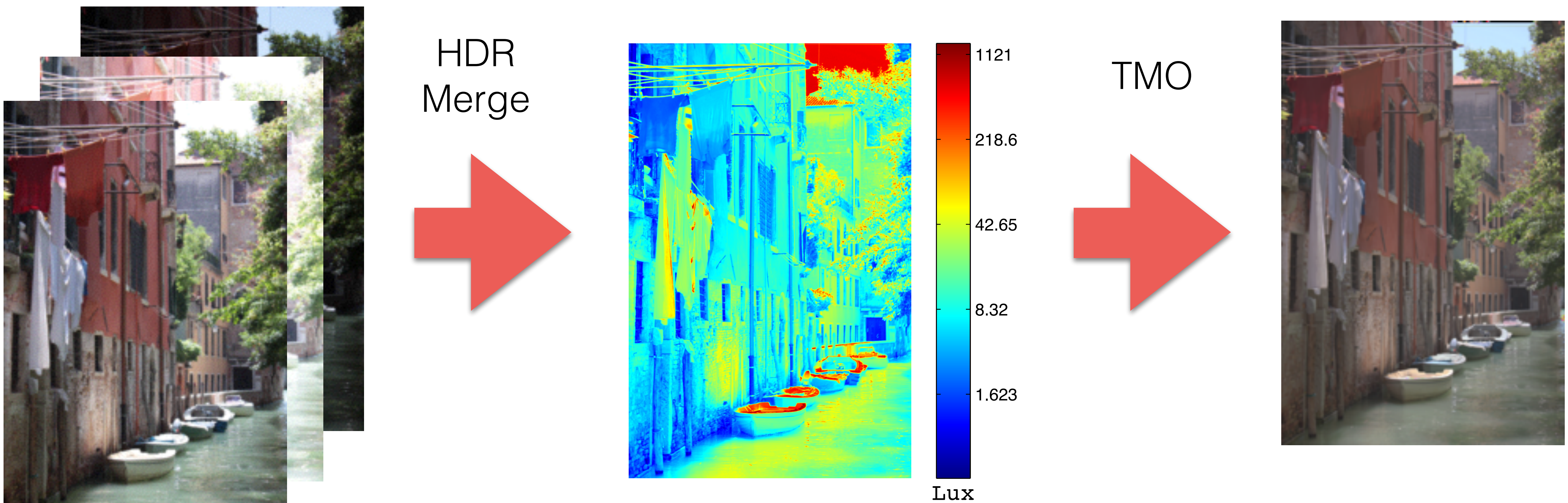
Metric/Index



Distortion Map

a different approach…

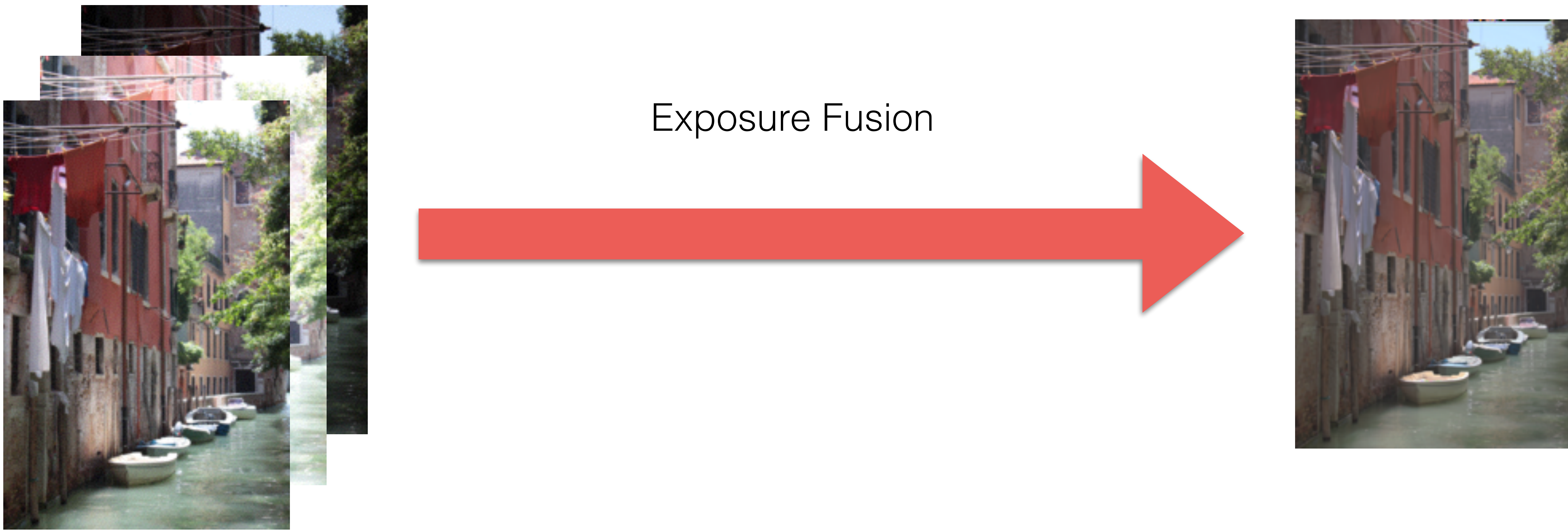
Exposure Fusion



Exposure Fusion

- In some cases:
  - Display only nice images that look good
  - No need to recover real-world luminance and colors:
    - no camera response function
    - no measurements

Exposure Fusion



Exposure Fusion

- **Idea**: for each *i-th* image create a per pixel weight, and use it during merge [Mertens et al. 2007]

$$C_i(\mathbf{x}) = \nabla^2 L((\mathbf{x})$$

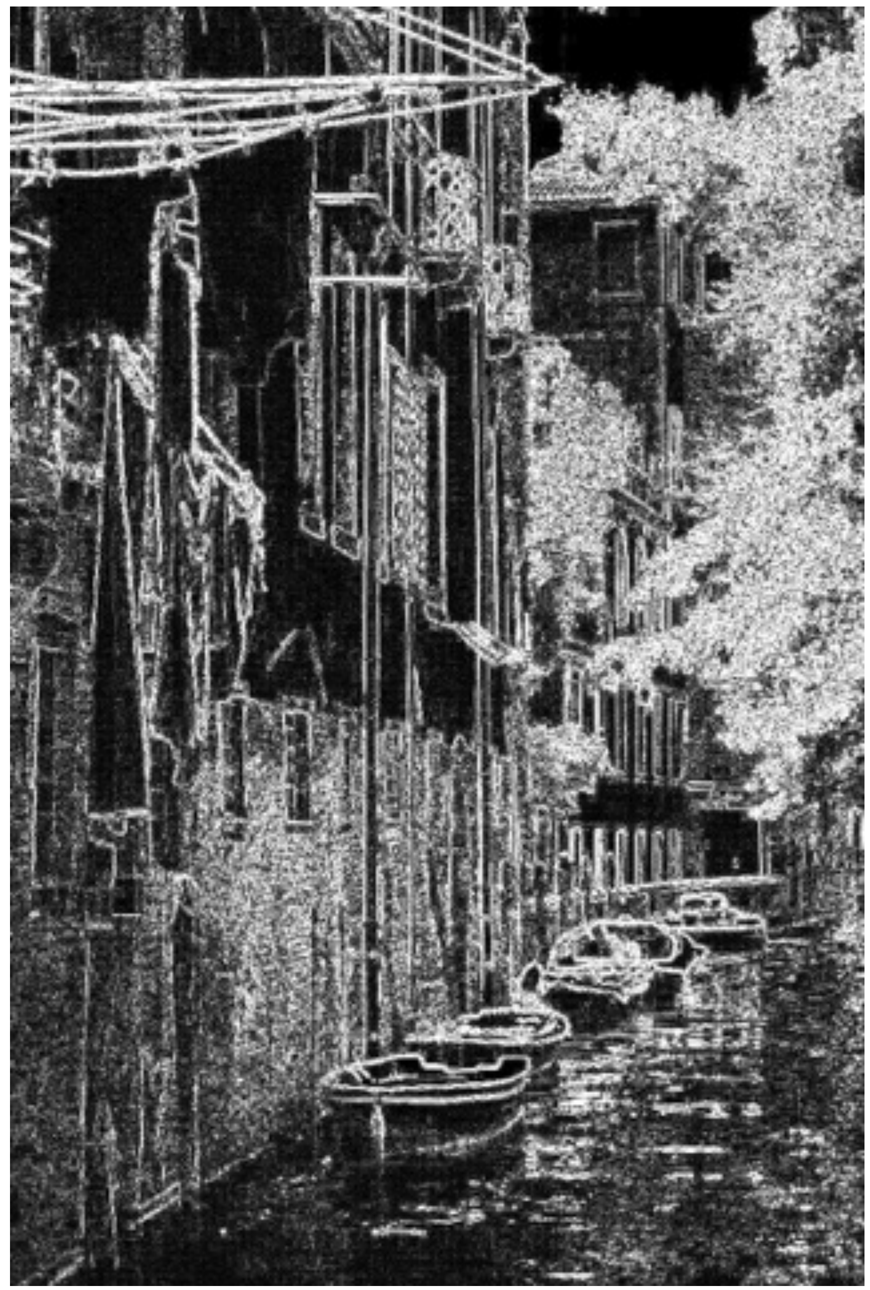
$$S_i(\mathbf{x}) = \mathrm{Var}(I(\mathbf{x}))$$

$$E_i(\mathbf{x}) = e^{\frac{-(L(\mathbf{x})-0.5)^2}{2\sigma^2}}$$

$$W_i(\mathbf{x}) = C_i(\mathbf{x})^{\omega_C} \times S_i(\mathbf{x})^{\omega_S} \times E_i(\mathbf{x})^{\omega_E}$$

Exposure Fusion

E_i



C_i

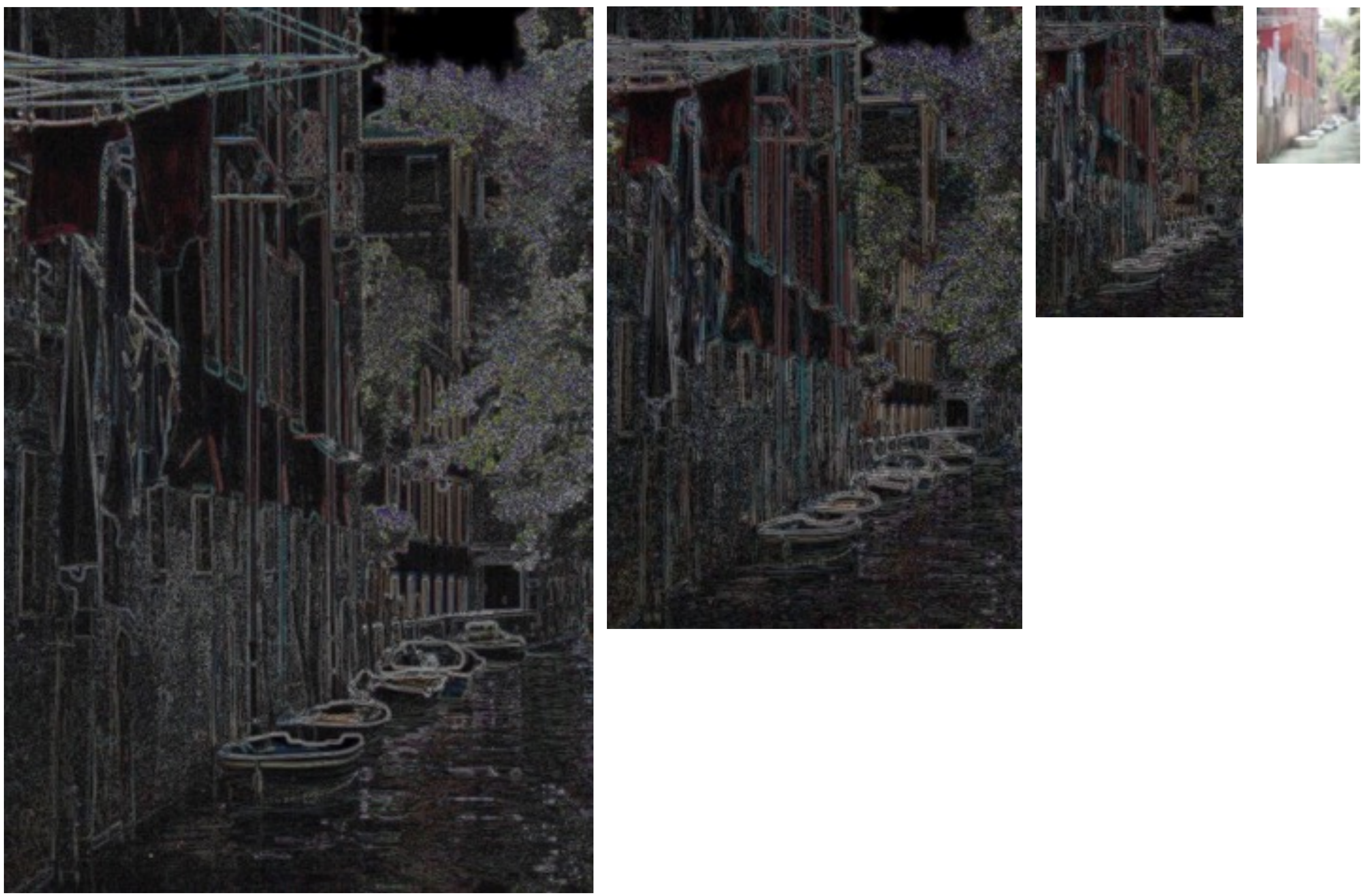
S_i

Exposure Fusion

Exposure Fusion: Blending

- Blending in spatial domain can lead to seams.
  - Blending using Laplacian Pyramids:
    - A multi-resolution tool
    - Gaussian Pyramid: downsample the image
    - Laplacian Pyramid: downsample the image + compute difference with the previous level

Exposure Fusion: Laplacian Pyramids



Exposure Fusion: Gaussian Pyramids

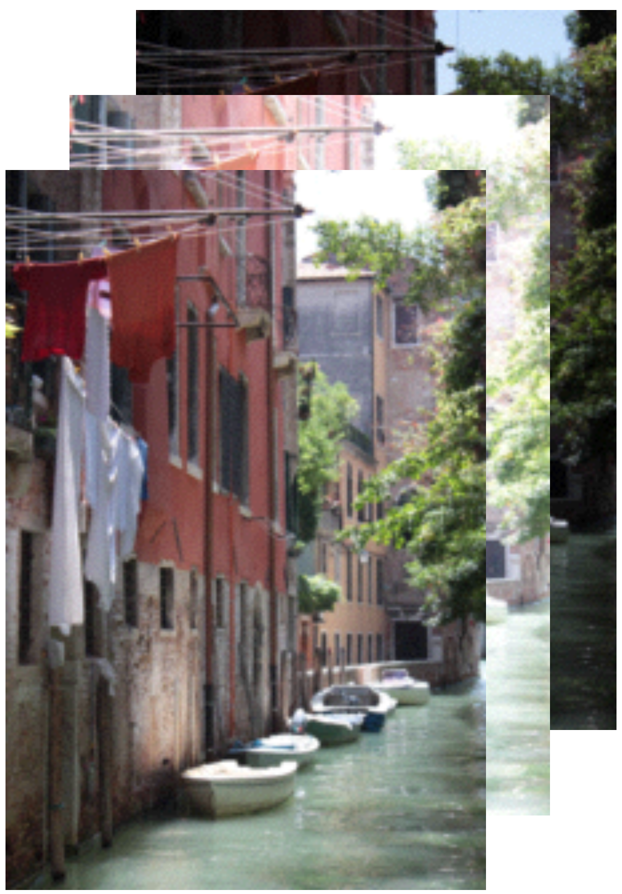
Exposure Fusion: Blending

$$\mathbf{L}^l\{I_\mathrm{d}\}(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{L}^l\{I_i\}(\mathbf{x})\mathbf{G}^l\{W_i\}(\mathbf{x})$$

Each W_i needs to be normalized $\frac{W_i(\mathbf{x})}{\sum_{j=1}^{n} W_j(\mathbf{x})}$

Exposure Fusion: Blending

Exposure Fusion

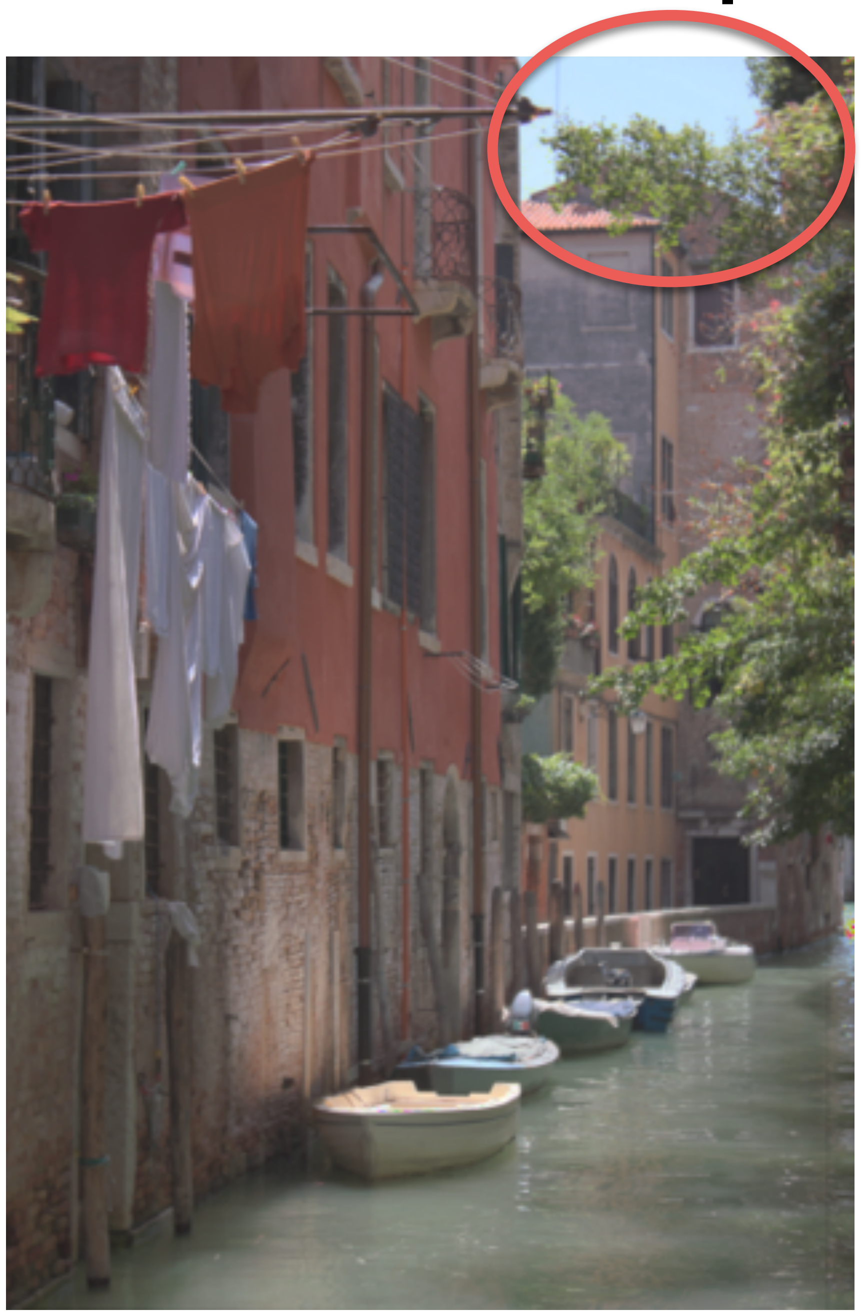


Exposure Fusion: Comparisons

Sigmoid TMO

Exposure Fusion

Exposure Fusion: Comparisons



Sigmoid TMO

Exposure Fusion

Exposure Fusion: Comparisons

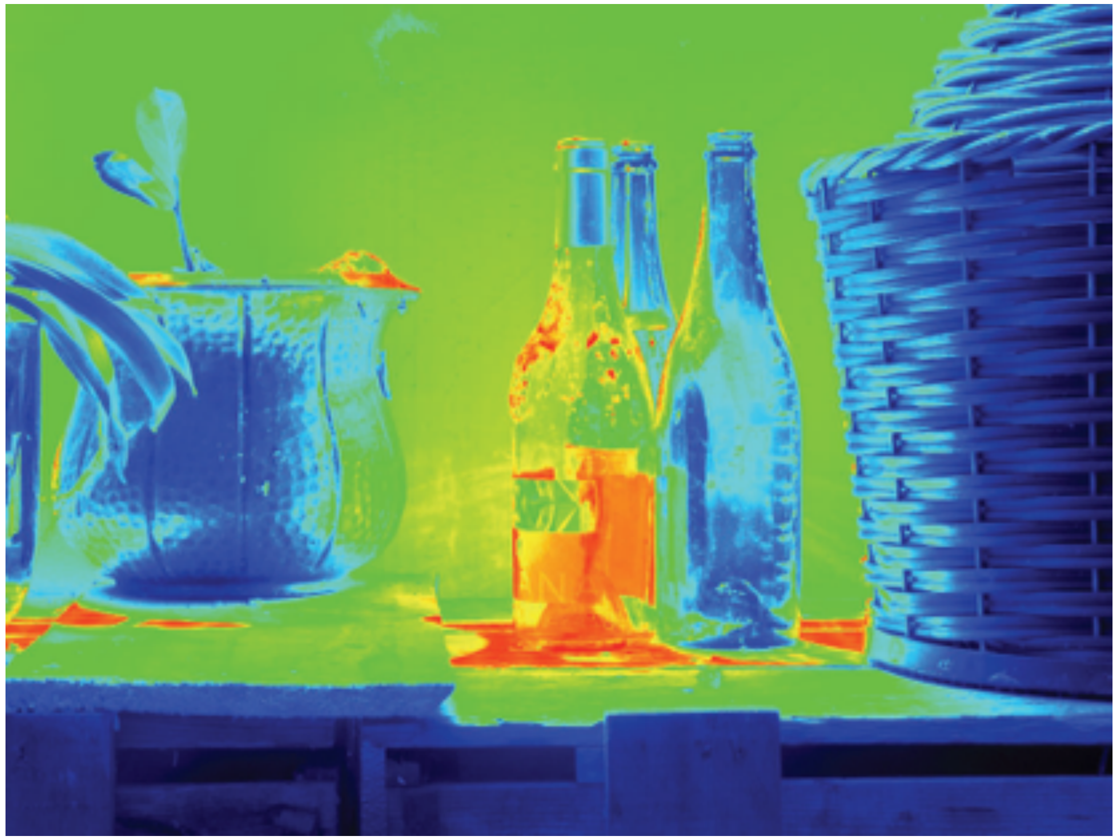
Sigmoid TMO

Exposure Fusion

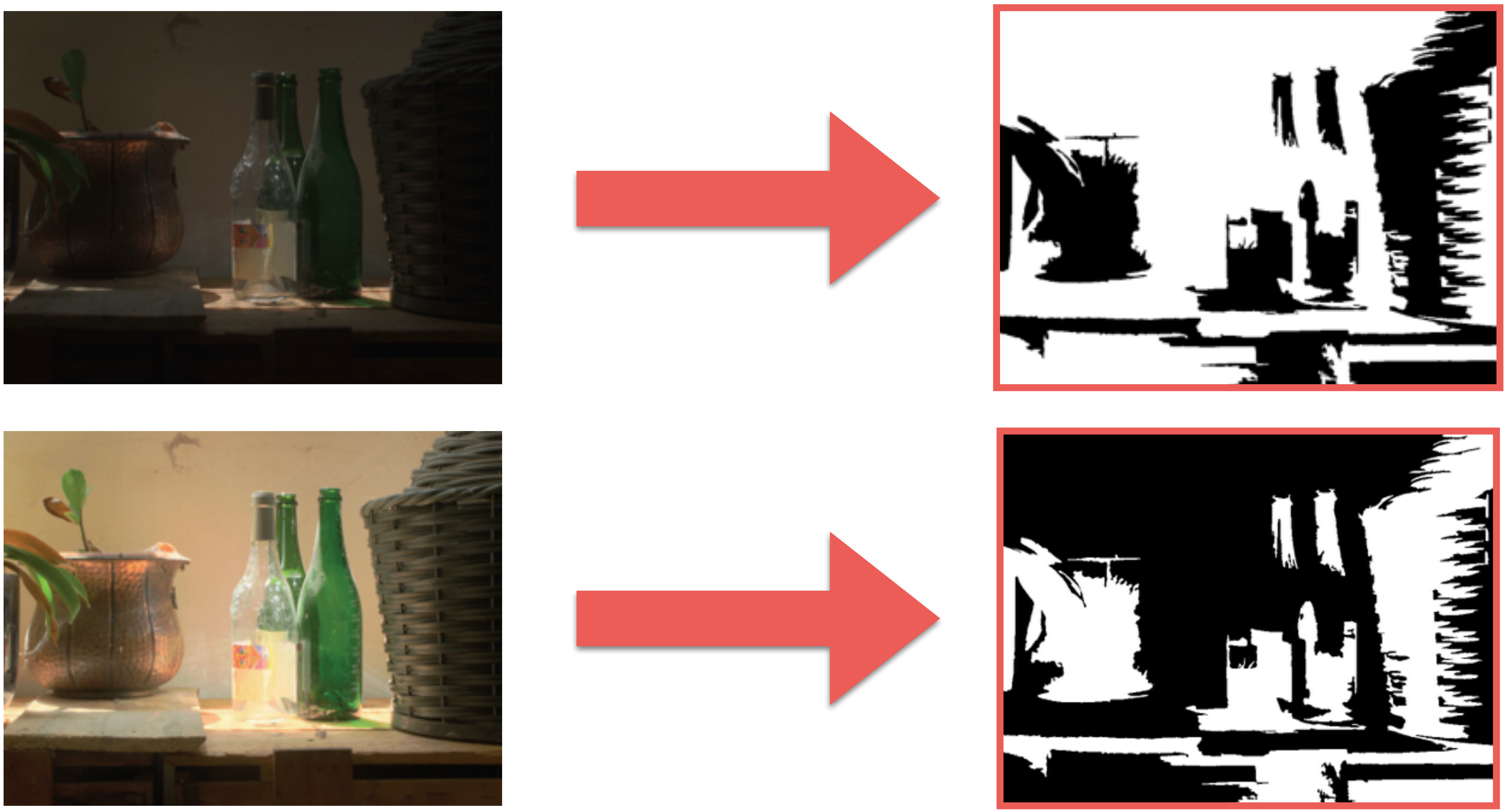
Mixing TMOs

- There are more than 100 TMOs… one or more should be good!
- **Idea**: apply many TMOs to the same HDR image and merge all results
- How is merge carried out?
  - Weights from psychophysical experiments [Banterle et al. 2012]
  - Weights from a perceptual metric [Yeganeh and Wang 2013]

Mixing TMOs



Mixing TMOs

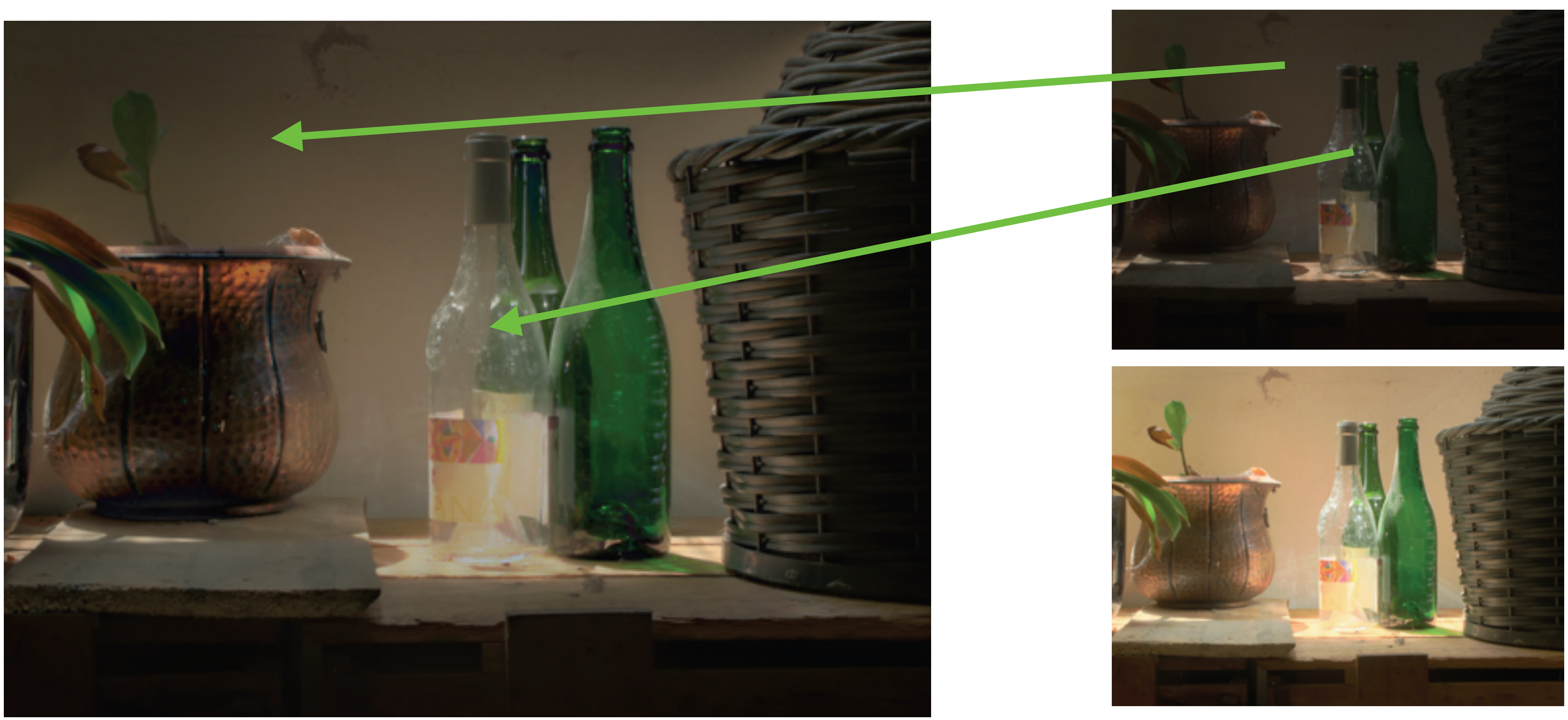


Mixing TMOs

$$\mathbf{L}^l\{I_\mathrm{d}\}(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{L}^l\{I_i\}(\mathbf{x})\mathbf{G}^l\{W_i\}(\mathbf{x})$$

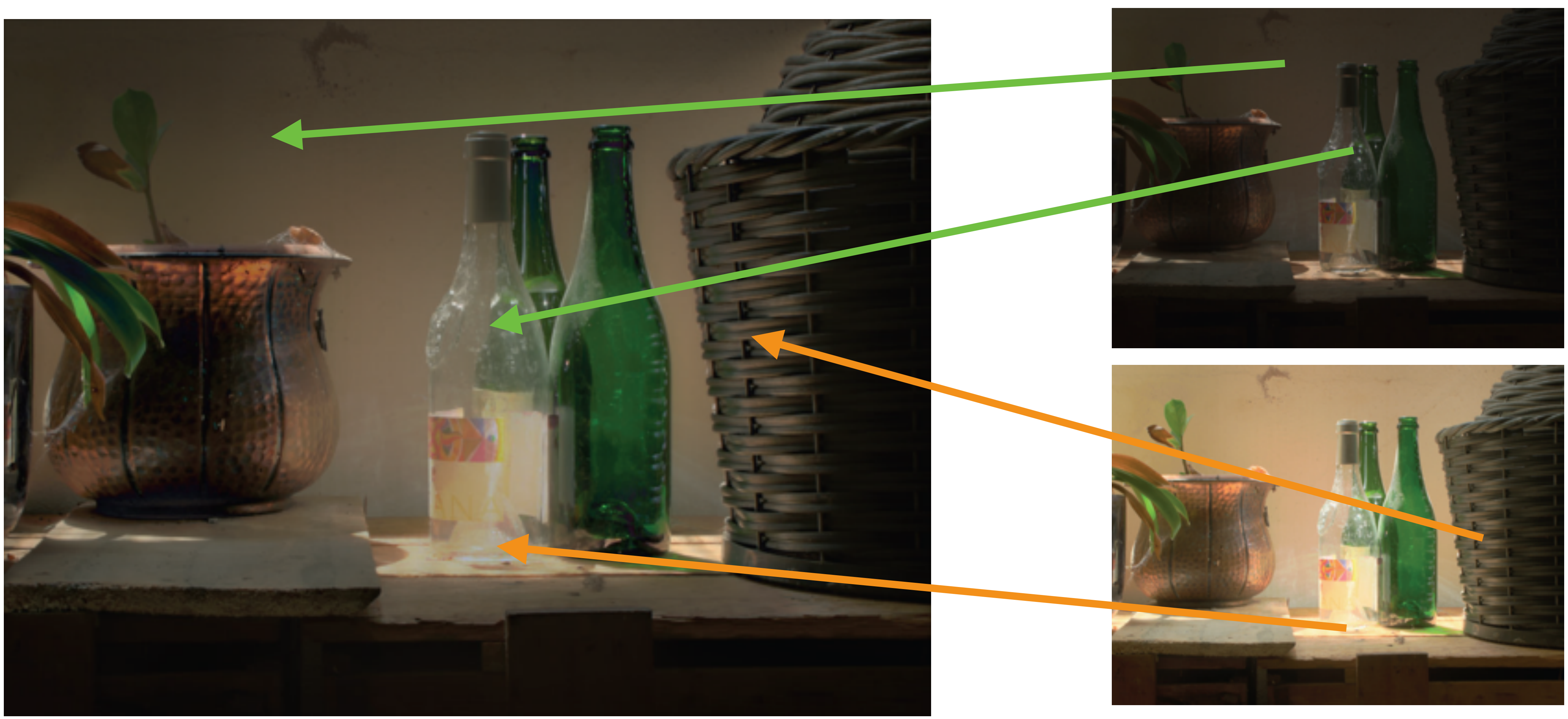
Mixing TMOs

$$\mathbf{L}^l\{I_\mathrm{d}\}(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{L}^l\{I_i\}(\mathbf{x})\mathbf{G}^l\{W_i\}(\mathbf{x})$$



Mixing TMOs

$$\mathbf{L}^l\{I_\mathrm{d}\}(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{L}^l\{I_i\}(\mathbf{x})\mathbf{G}^l\{W_i\}(\mathbf{x})$$



Mixing TMOs

- Advantages:
  - The best from every TMO
- Disadvantages:
  - Computationally expensive —> no real-time

Questions?