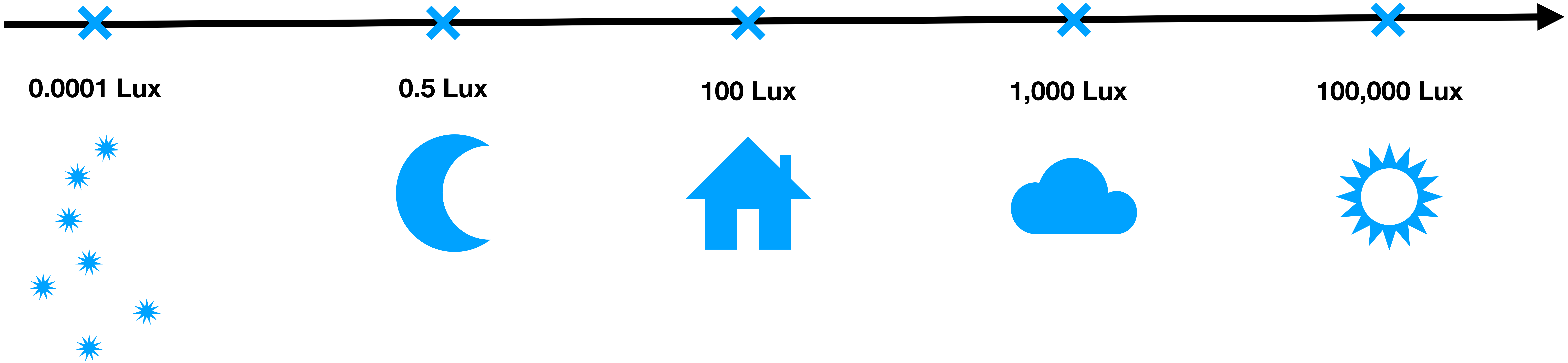
Modern High Dynamic Range Imaging at the Time of Deep Learning

Introduction

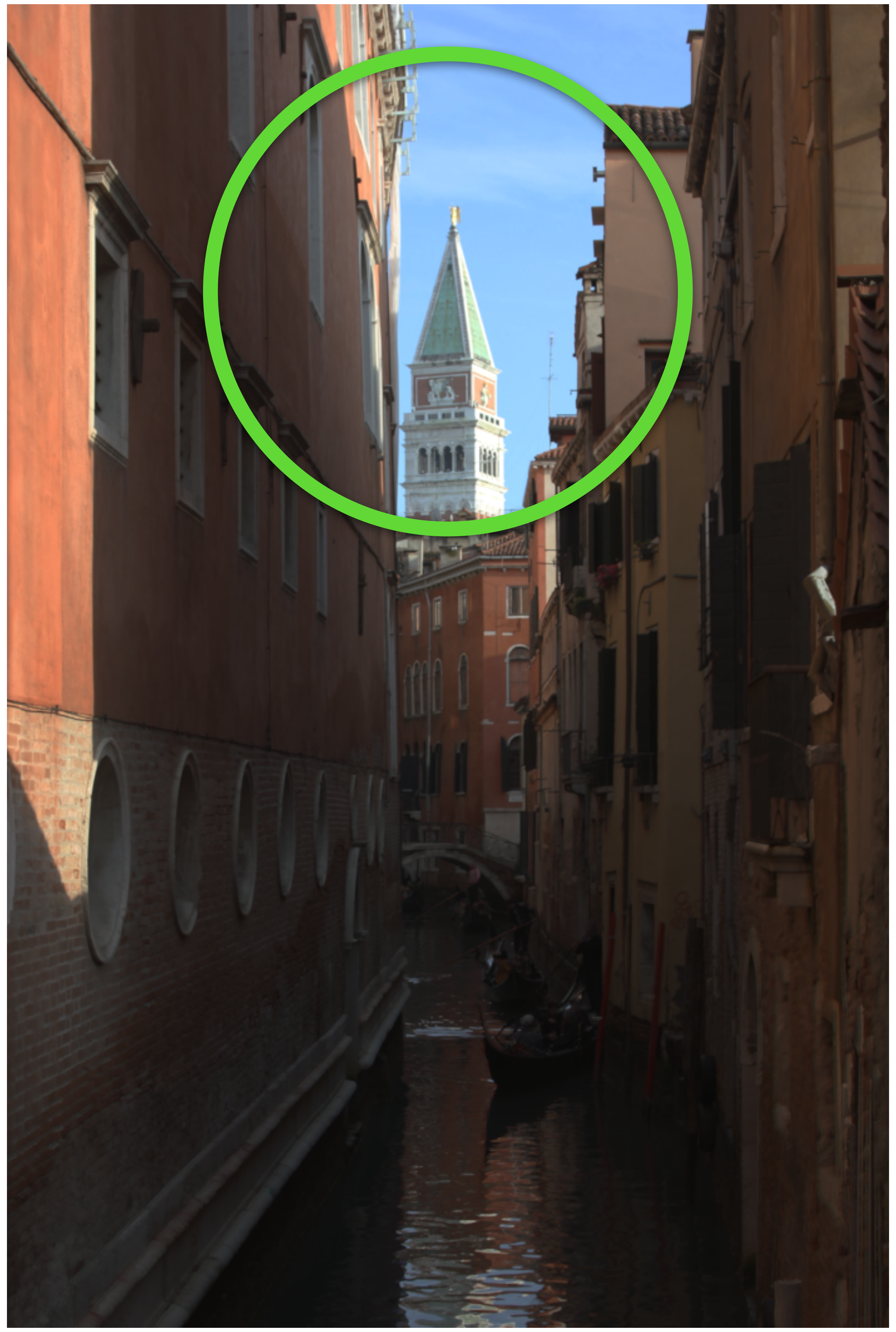
Francesco Banterle and Alessandro Artusi

High Dynamic Range Imaging

HDR Imaging



HDR Imaging



Short Exposure

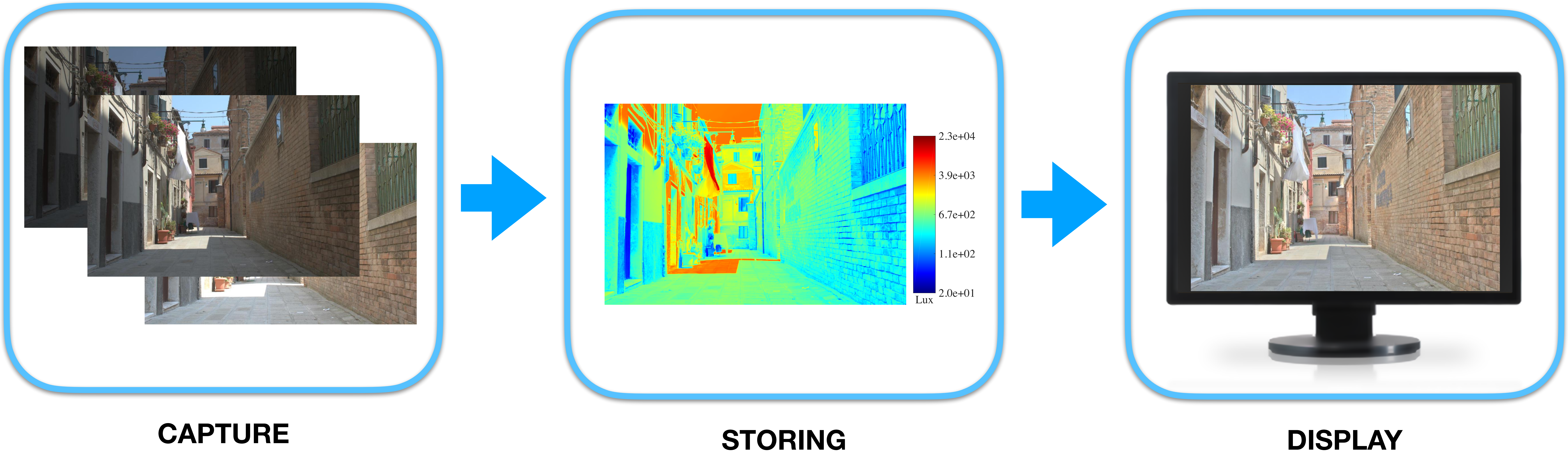
Mid Exposure

Long Exposure

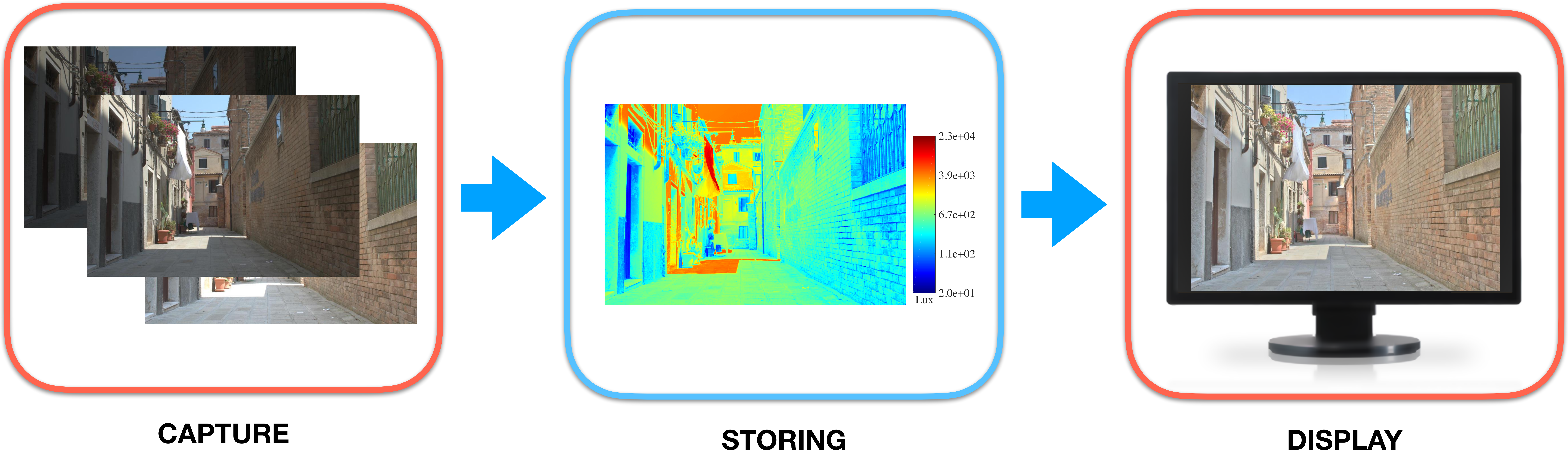
HDR Imaging

Merged Exposures

The HDR Pipeline


2.3e+04
3.9e+03
6.7e+02
1.1e+02
2.0e+01
Lux
CAPTURE
STORING
DISPLAY

HDR Imaging: Merging



HDR Imaging:
Acquisition

HDR Imaging: Merging

- To merge N images, Z_k , at different exposure times, t_k, we sum them up taking into account that they were taken at different shutter speed:

$$E(i,j) = \frac{\sum_{k=1}^{N} w(Z_k(i,j)) \cdot g(Z_k(i,j)) \cdot t_k^{-1}}{\sum_{k=1}^{N} w(Z_k(i,j))}$$

- where $g = f^{-1}$ is the inverse camera response function, and w is a weighting function. Typically, the merge is computed in the log-domain to reduce noise.

HDR Imaging: Merging

- The result $E(i,j)$ is a **radiance map**:
  - Note E is the irradiance symbol; the radiance symbol is L:
    - Technically speaking we should taking into account that:

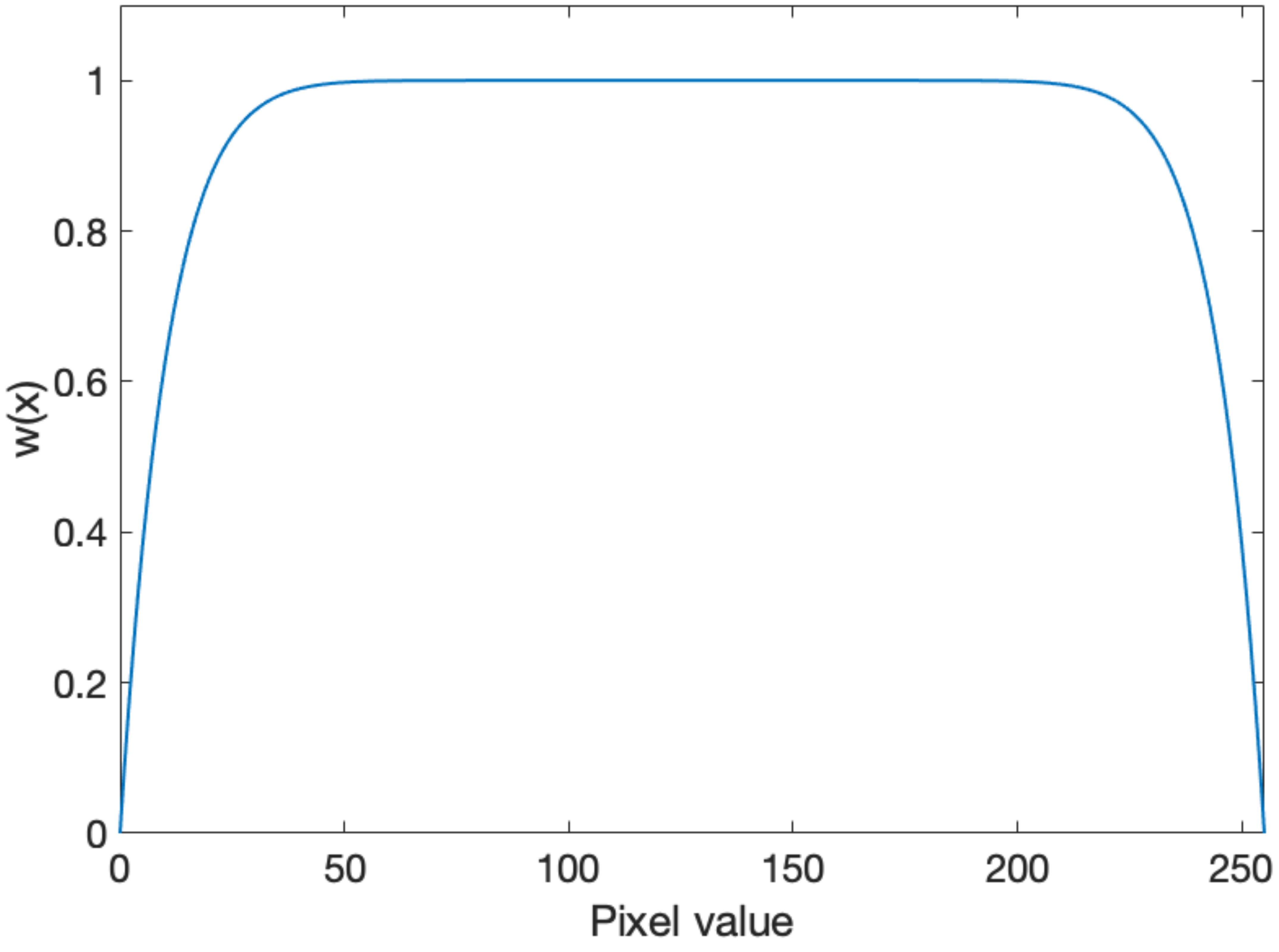
$$E(i,j) = L(i,j)\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

    - But… Most lenses already compensate for this!

HDR Imaging: The Weighting Function

- The weighting function selects well-exposed pixels from the input image to avoid noisy and saturated pixels:
  - Such value increase noise or bias in the final HDR image.
- For example:

$$w(x) = 1 - (2x - 1)^{12}$$



HDR Imaging: Camera Response Function

- A Camera Response Function (CRF), f, is a non-linear function of image irradiance:
  - It is a solution for compressing the irradiance values large dynamic range into a fixed range of recordable values; i.e., 8-bit of a JPEG image.
    - RAW images (stored in 10-14 bits) have mostly a linear behavior.
  - It is typically not known, but it can be estimated.

HDR Imaging: Camera Response Function

- Knowing:

$$Z_k(i,j) = f\big(E(i,j) \cdot t_k\big)$$

- We obtain:

$$f^{-1}\big(Z_k(i,j)\big) = E(i,j) \cdot t_k$$

- Finally:

$$\log f^{-1}\big(Z_k(i,j)\big) = \log E(i,j) + \log t_k$$

Classic Debevec and Malik 1997

HDR Imaging: Camera Response Function

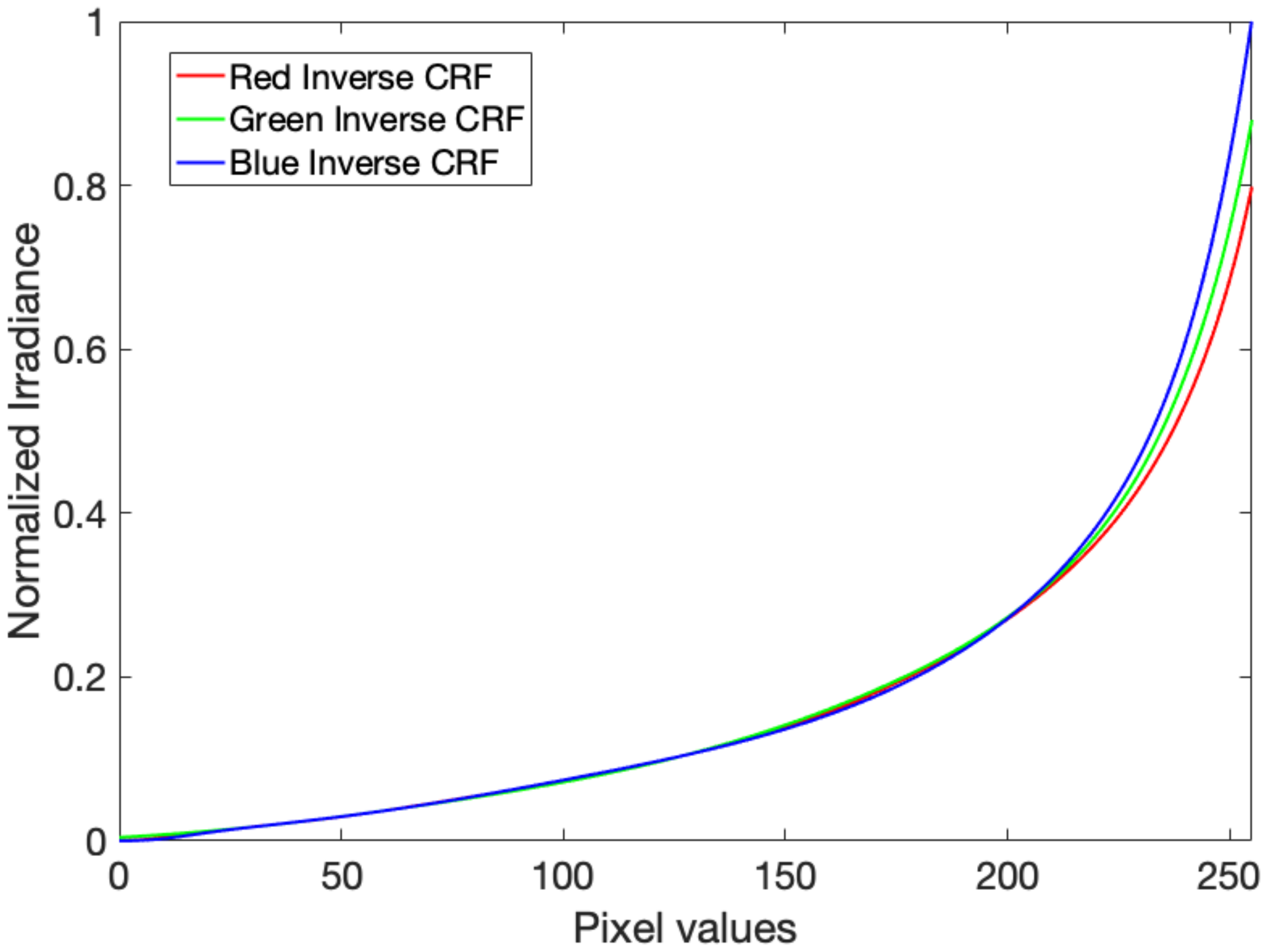
- A typical estimation method is based on optimization:

$$\mathcal{O} = \sum_{k=1}^{N} \sum_{i,j} \left(\log g(Z_k(i,j)) - \log E(i,j) - \log t_j \right)^2 + \lambda \sum_{x} g''(x)^2$$

where $g(x) = f^{-1}(x)$.

Classic Debevec and Malik 1997

HDR Imaging: Camera Response Function



Classic Debevec and Malik 1997

HDR Imaging: Camera Response Function

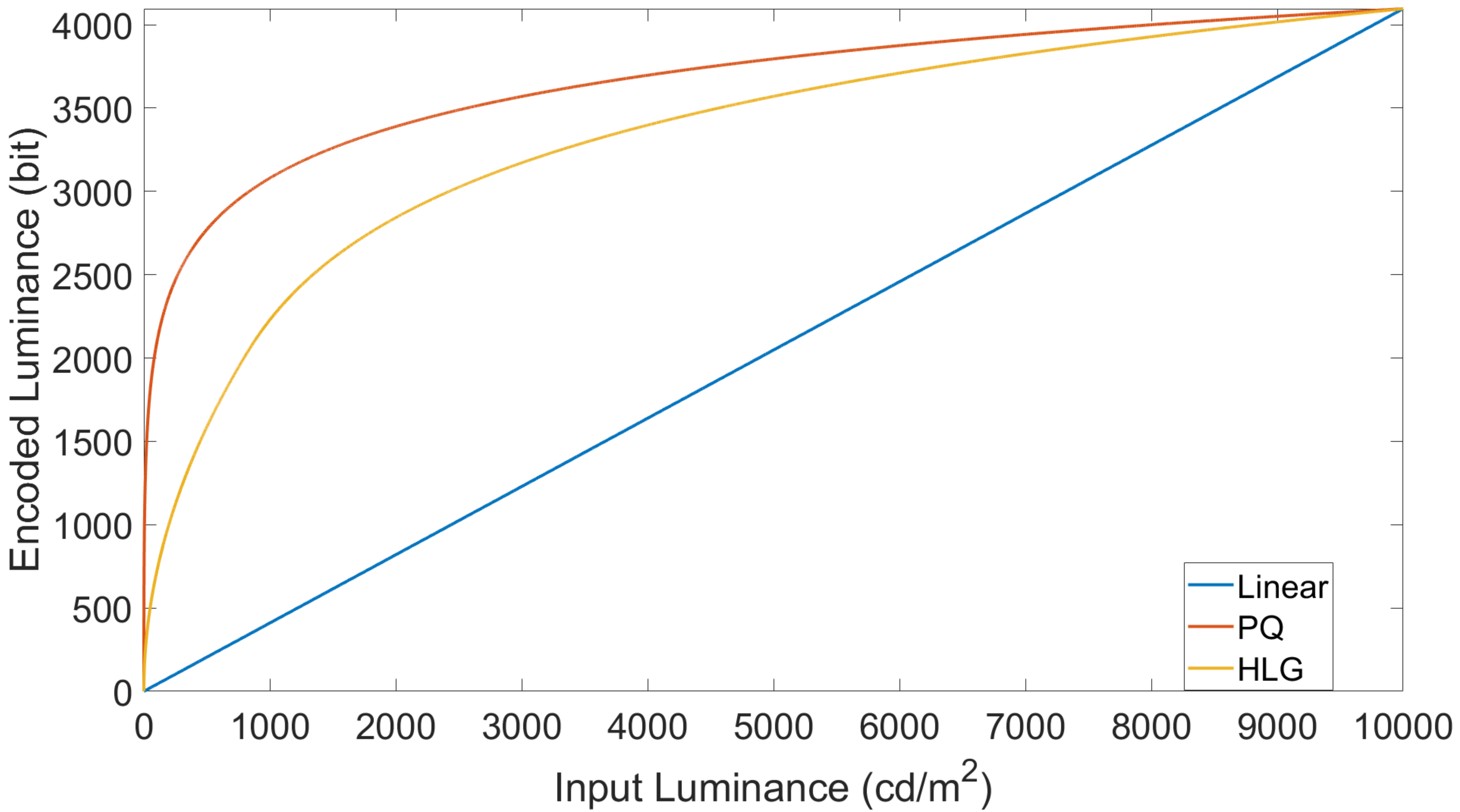
- Many cameras/smartphone manufactures and displays makers have started to agree on some standard CRF or OETF. Most famous examples:
  - PQ:

$$f(Y) = \left(\frac{c_1 + c_2 Y_n^{m_1}}{1 + c_3 Y_n^{m_1}}\right)^{m_2} \text{ where } Y_n = \frac{Y}{10000}$$

  - HLG:

$$f(Y) = \begin{cases} r\sqrt{Y} & Y \in [0,1] \\ a\log(Y - b) + c & Y > 1 \end{cases}$$

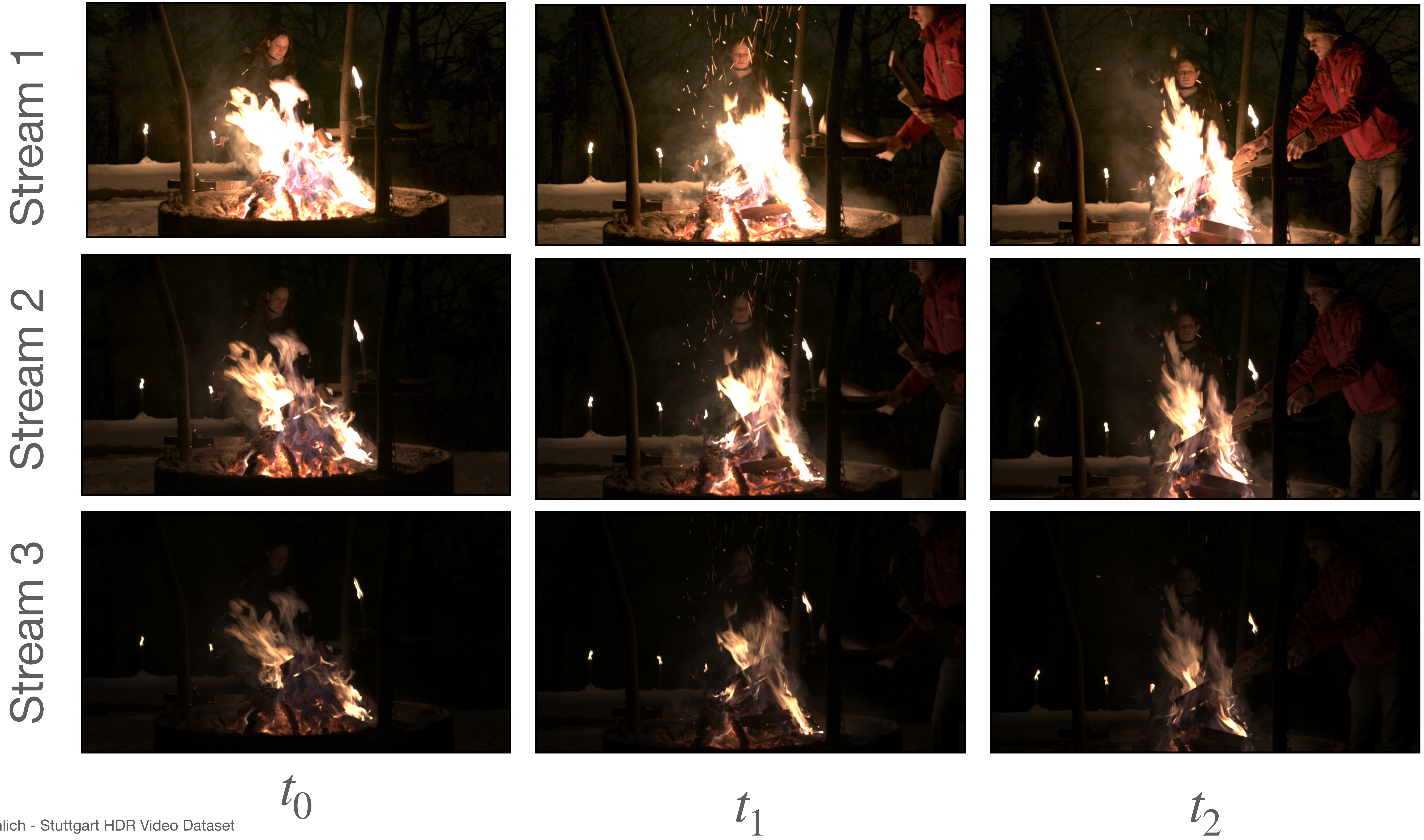
HDR Imaging: Camera Response Function



HDR Videos

- There are different strategies:
    - Multiple sensors combined with beam splitter capturing frames at different exposures time [Tocci+2011].
    - Varying the exposure shutter speed at each frame [Kang+2003].
    - Varying the exposure time in the bayer filter or assorted pixels [Yasuma+2010].

HDR Videos: Multiple Sensors



Video Courtesy of Jan Fröhlich - Stuttgart HDR Video Dataset

HDR Videos: Varying Exposure at Each Frame

Stream



t_0 t_1 t_2

Video Courtesy of Jan Fröhlich - Stuttgart HDR Video Dataset

HDR Videos: Assorted Pixels

Video Courtesy of Jan Fröhlich - Stuttgart HDR Video Dataset

HDR Videos: Assorted Rows

Video Courtesy of Jan Fröhlich - Stuttgart HDR Video Dataset

HDR Imaging:
Tone Mapping - SDR Visualization

Tone Mapping

- A tone mapping operator (TMO) is a function, $f(\cdot)$, that reduces the dynamic range of a HDR image to fit into a SDR display. We have two main classes:
  - **Global operators**: it uses global statistics of the image to be tone mapped:
    - We want to maintain the global contrast of the original image.
  - **Local operators**: it uses both global and local statistics of the image to be tone mapped:
    - We want to maintain both the local and global contrast of the original image.

Tone Mapping

- Most operators work only on the luminance channel:

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

- For the sRGB color space, this is defined as

$$L_w = 0.2126 \cdot R_w + 0.7152 \cdot G_w + 0.0722 \cdot B_w$$

- To avoid color distortions when applying $f(\cdot)$ for each color channel.

Tone Mapping: Global Operators

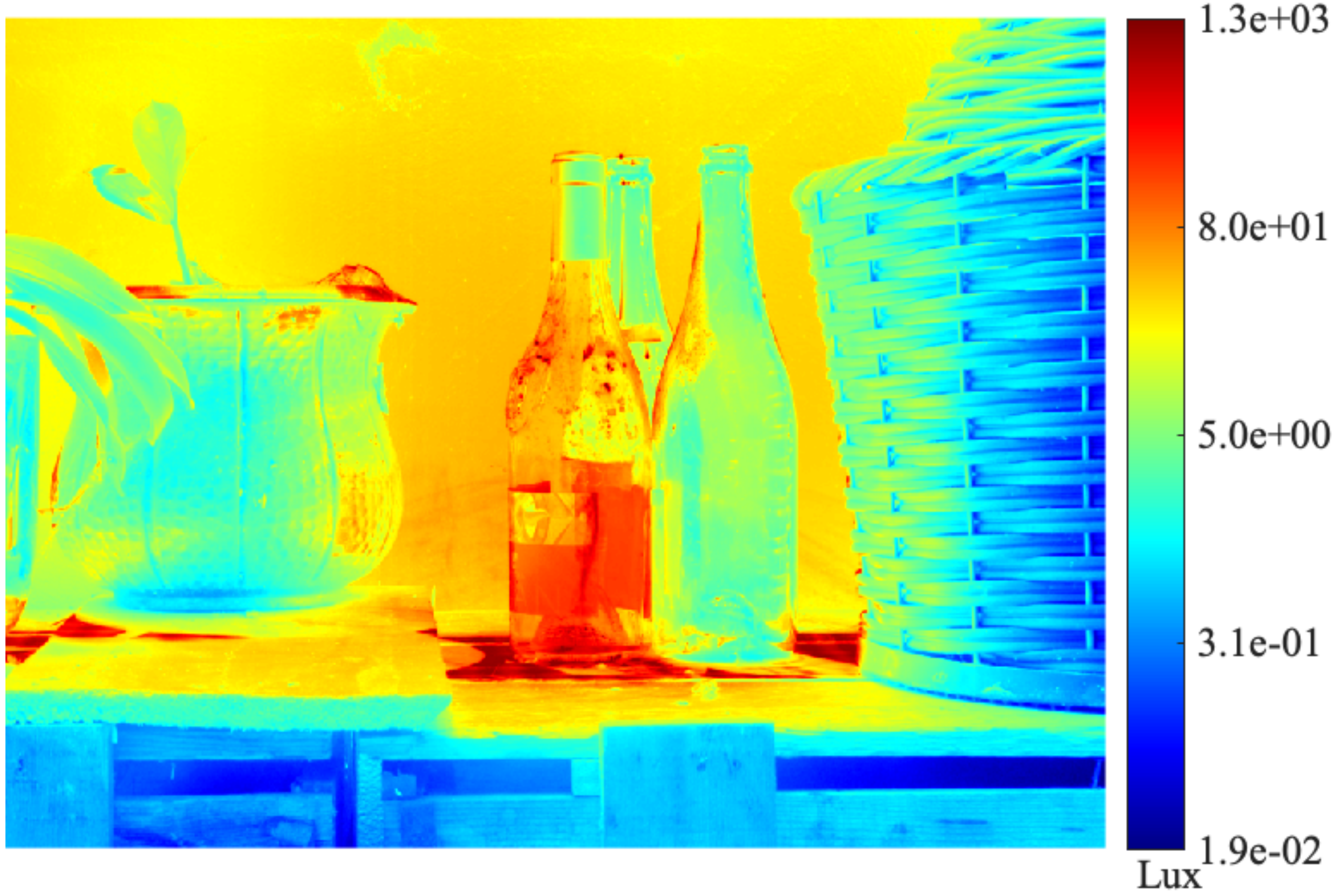
- A classic local TMO is the Reinhard et al.'s operator [Reinhard+2002]:

$$L_d = f(L_w) = \frac{L_m}{1 + L_m} \qquad L_m = \frac{\alpha}{\hat{L}_w} L_w,$$

where α is a user parameter, and $\hat{L}_w$ is the geometric mean of the luminance of the entire image:

$$\hat{L}_w = \exp\left(\frac{1}{n} \sum_{i,j} \log_e\big(L_w(i,j) + \delta\big)\right) \qquad \delta > 0$$

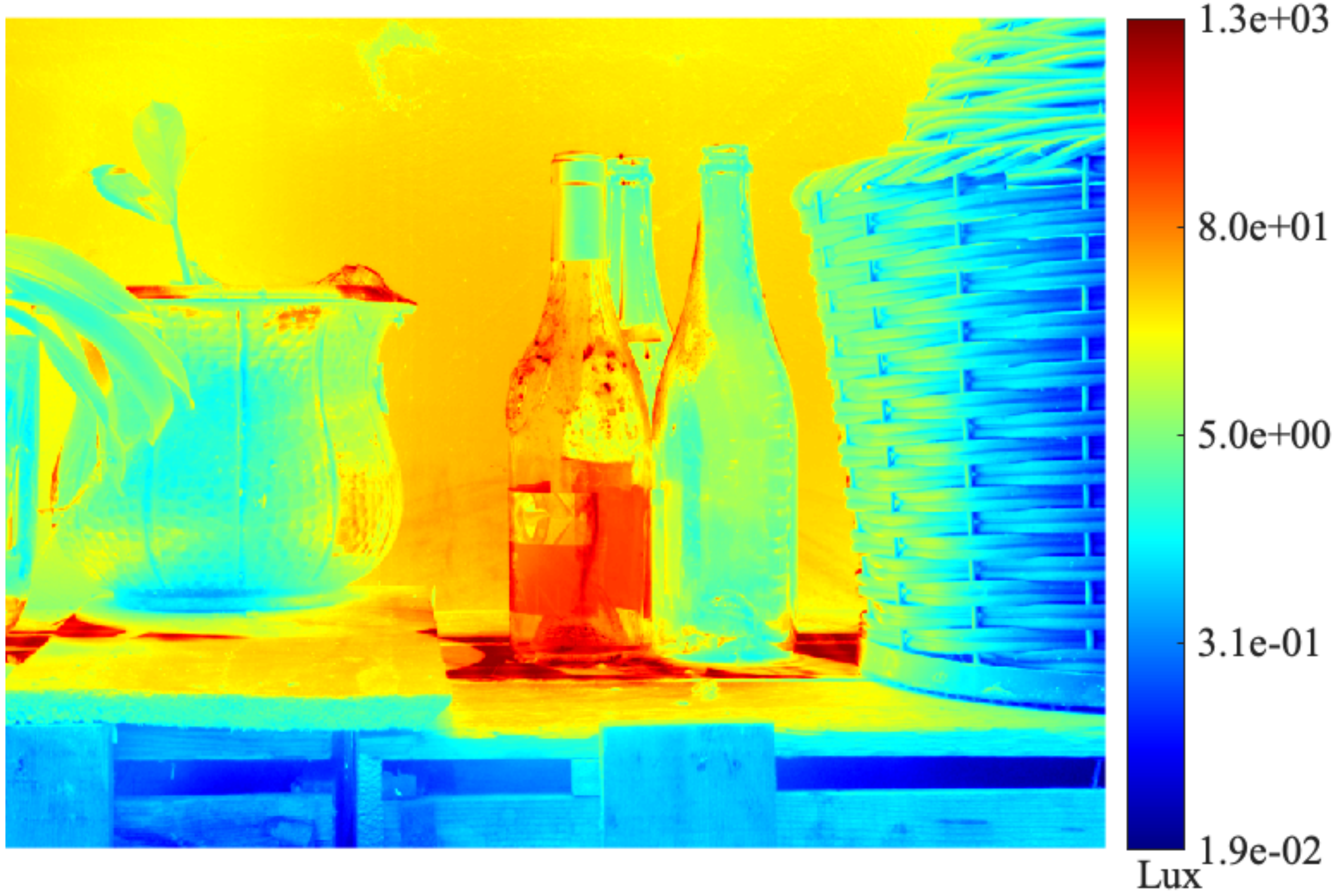
Tone Mapping: Global Operators Example



HDR Image

Reinhard with $\hat{L}_w = 1$

Tone Mapping: Global Operators Example



HDR Image

Reinhard with $\hat{L}_w$ computed

Tone Mapping: Local Operators

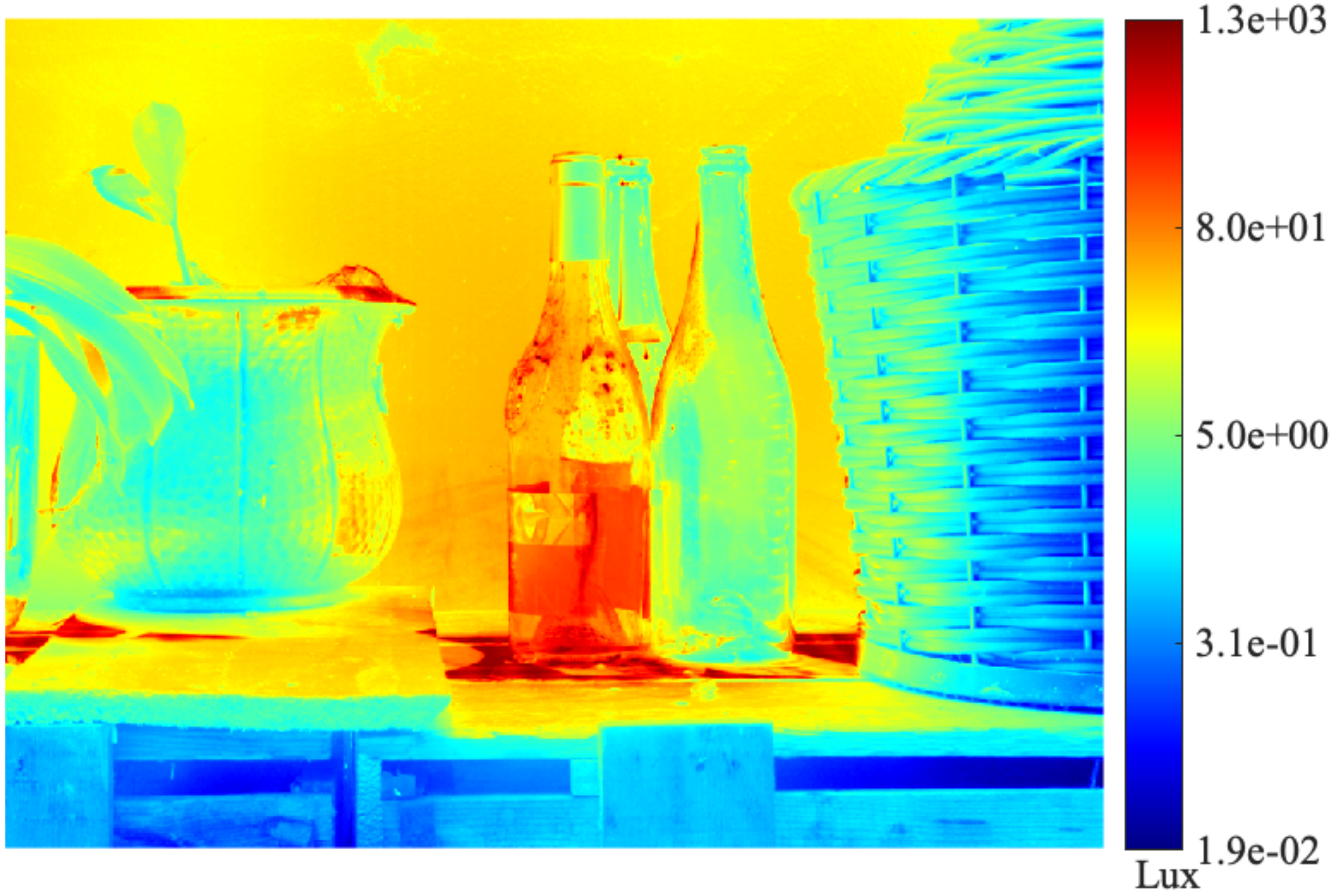
- A classic local TMO is a variant of the Reinhard et al.'s operator [Reinhard+2002]:

$$L_d = f(L_w(i,j)) = \frac{L_m(i,j)}{1 + g(L_m(i,j))} \quad L_m(i,j) = \frac{\alpha}{\hat{L}} L_w(i,j),$$

where $g(\cdot)$ is a function computing the mean around the pixel (i,j).

- However, we need to avoid strong edges that may create halos. So $g(\cdot)$ has to be edge-aware; e.g., the bilateral filter.

Tone Mapping: Local Operators Example



HDR Image

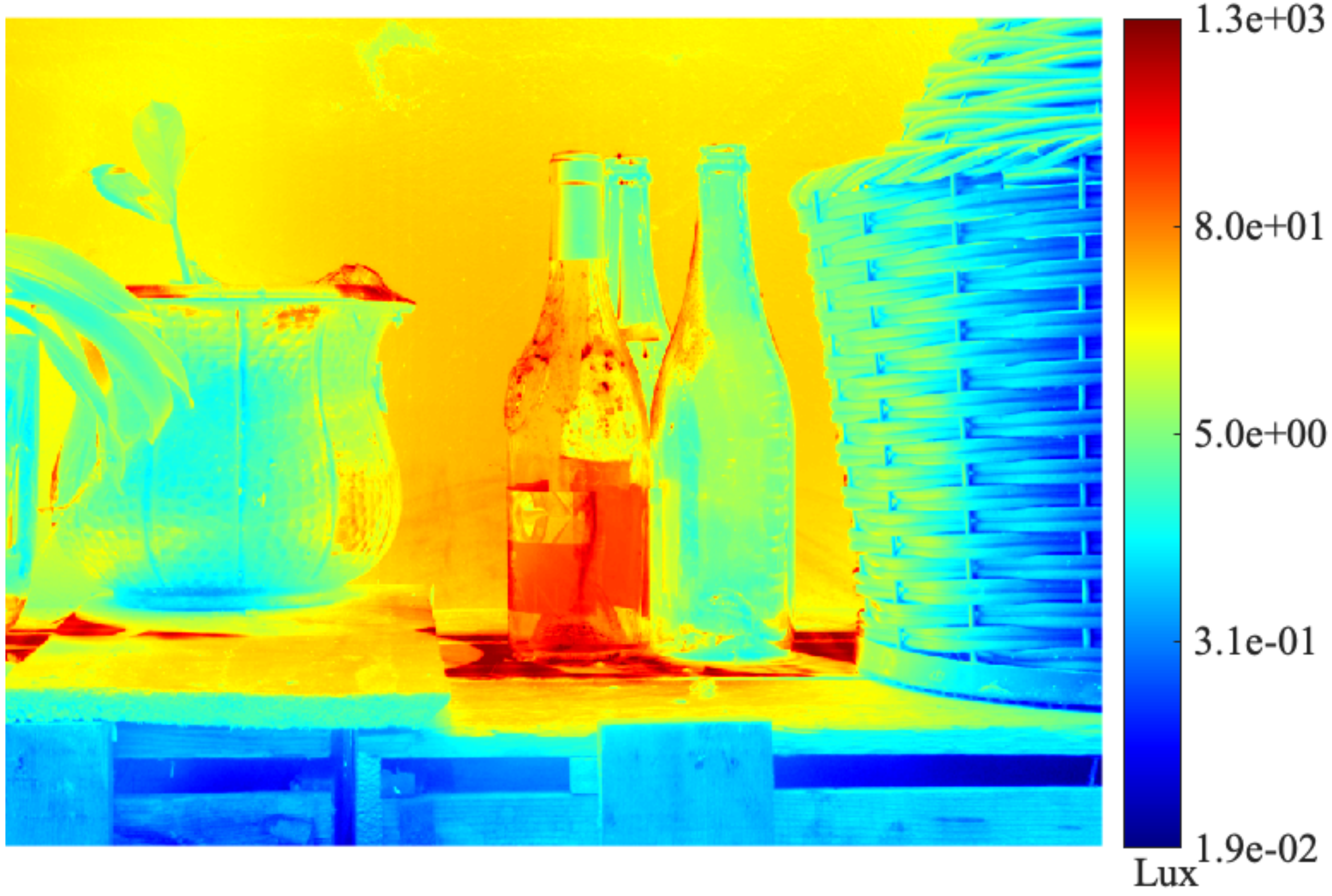
Reinhard without an edge-preserving filter

Tone Mapping: Local Operators Example

Reinhard without an edge-preserving filter

Zoom on a halo

Tone Mapping: Local Operators Example



HDR Image

Reinhard with an edge-preserving filter

Tone Mapping: Local Operators Example

Reinhard without an edge-preserving filter

Reinhard with an edge-preserving filter

Color Distortions

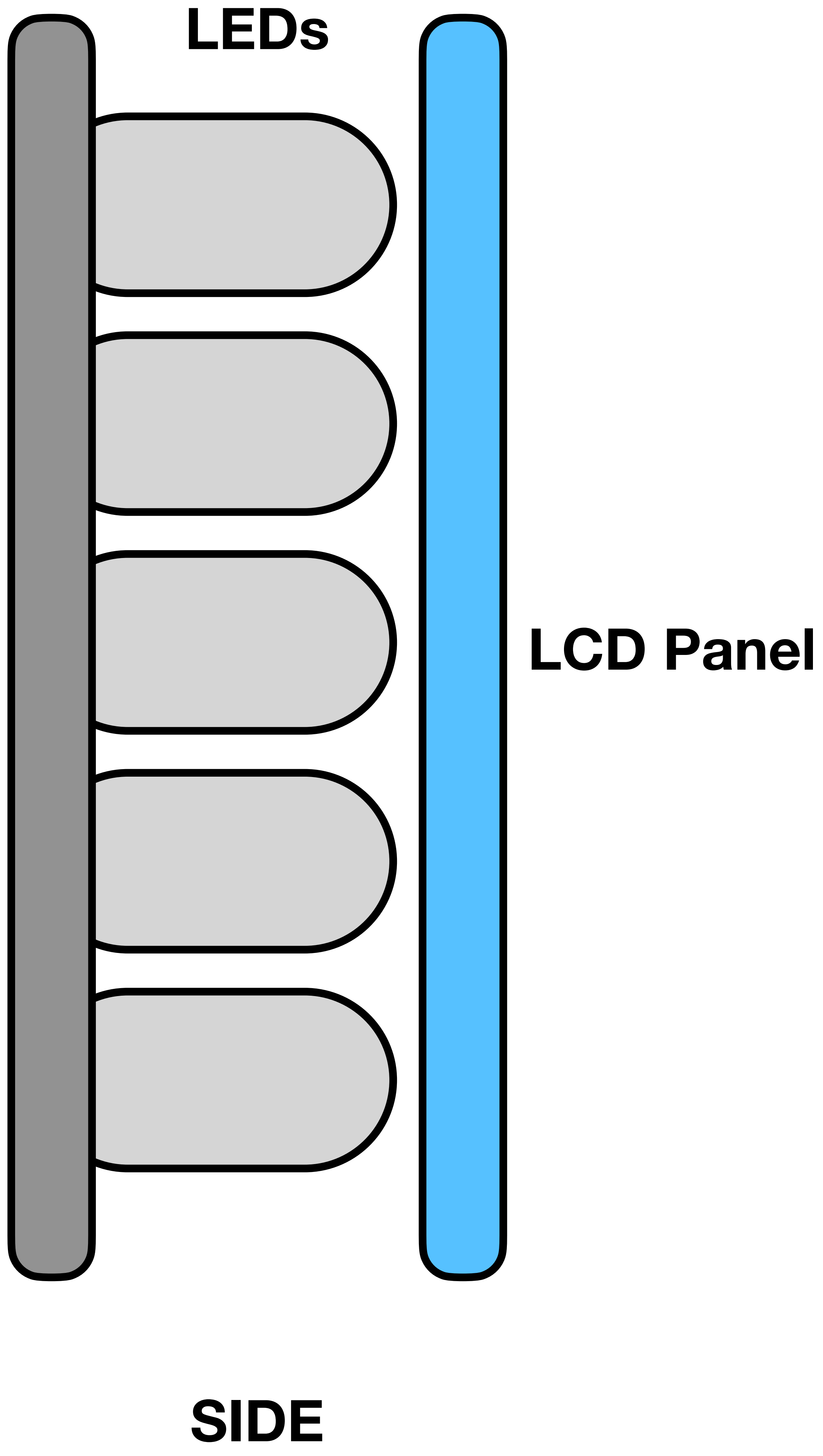
- The problem with processing only the luminance is that we have the following problems after applying $f(\cdot)$:
  - $L_d < L_w$ the saturation of the pixel increases.
  - $L_d > L_w$ the saturation of the pixel decreases.

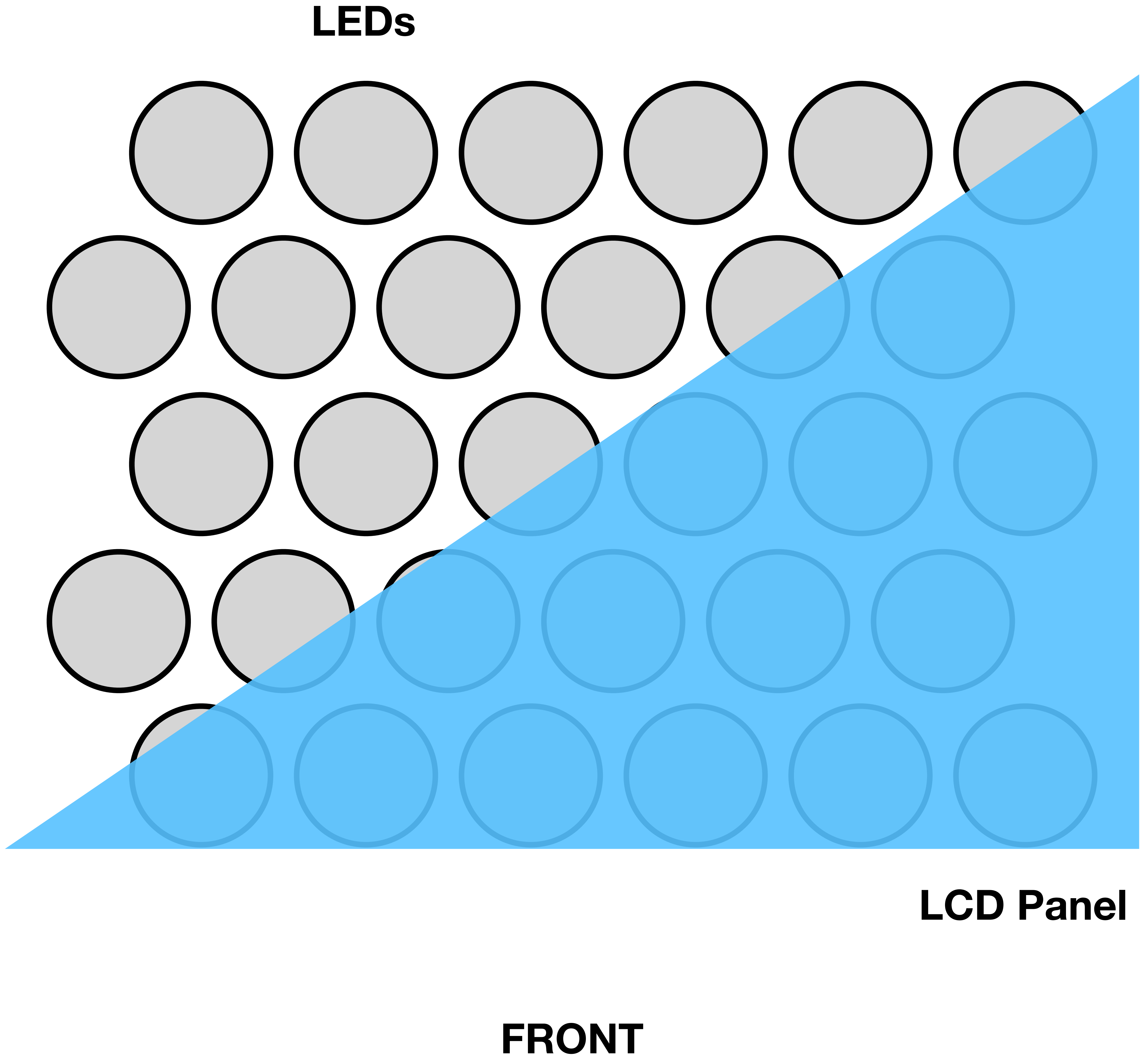
Color Solutions

- Main solutions:
  - Desaturate $[R_w, G_w, B_w]/L_w$ by applying a power with coefficient in $(0,1]$ [Schlick+1995].
  - Linear desaturation taking into account the TMO derivate [Mantiuk+2009].
  - Hue reset and saturation scale in the LCh color space [Pouli+2013, Pouli+2017].

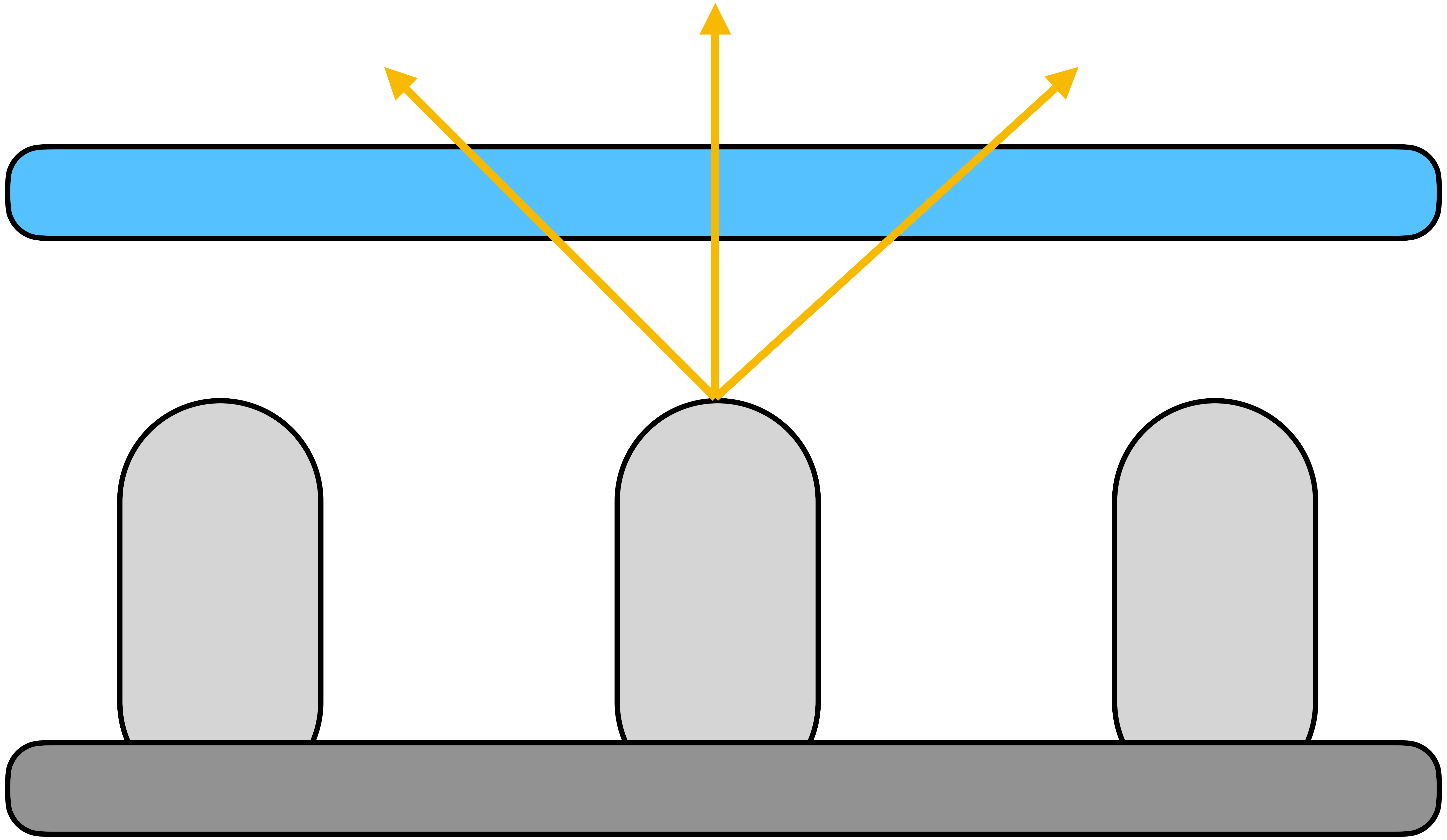
HDR Imaging:
Native Visualization - HDR Monitors

Native Visualization: LEDs HDR Monitors

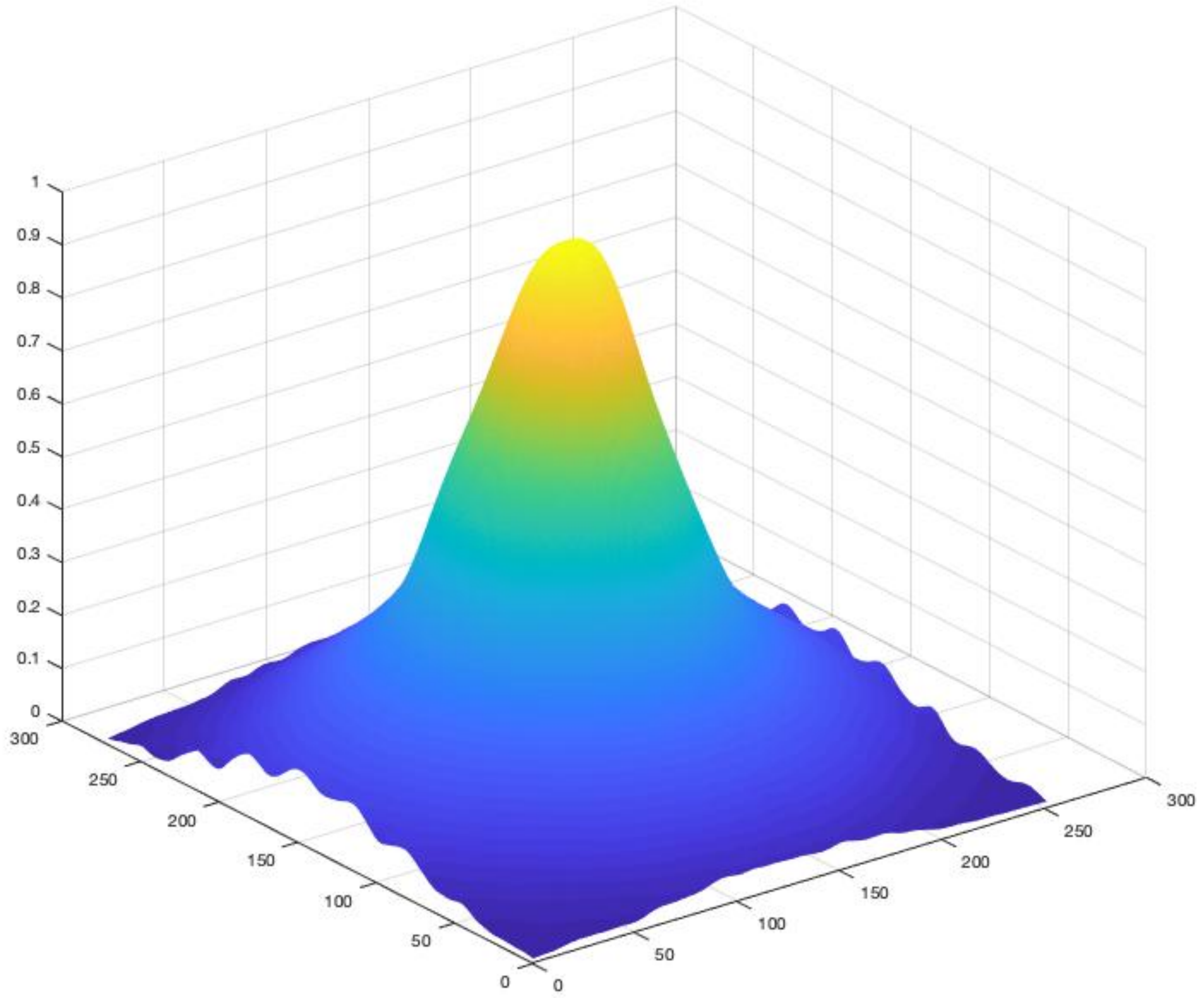




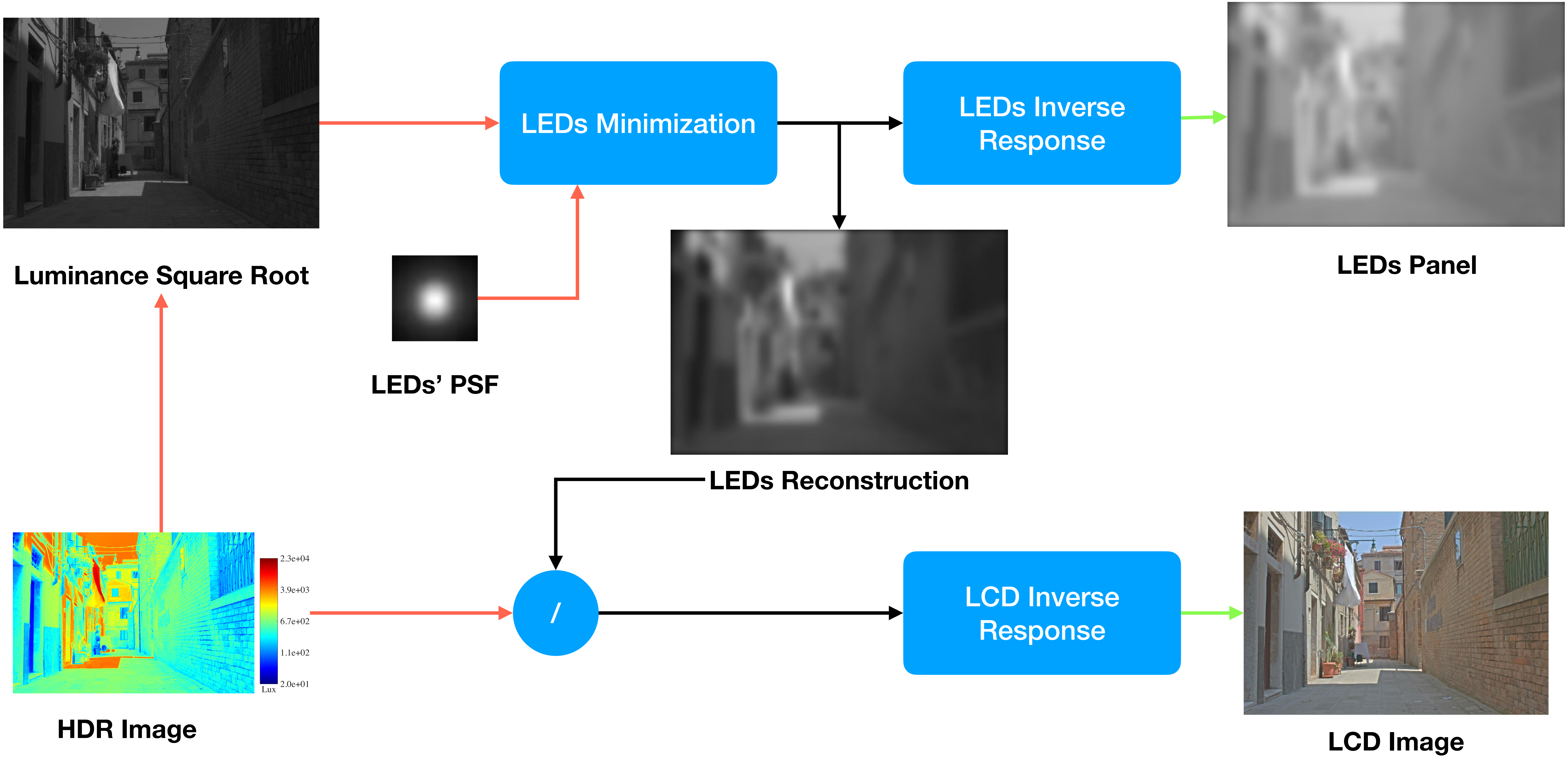
LED-based HDR Monitors: PSF



LED-based HDR Monitors: PSF



Native Visualization: HDR Monitors



HDR Imaging:
Metrics

Image Quality Metrics: with Reference

- A probability map; each pixel has the probability of being detected when compared to the reference by a viewer.
- Q predictor value in the range [0,100]; the higher the better.

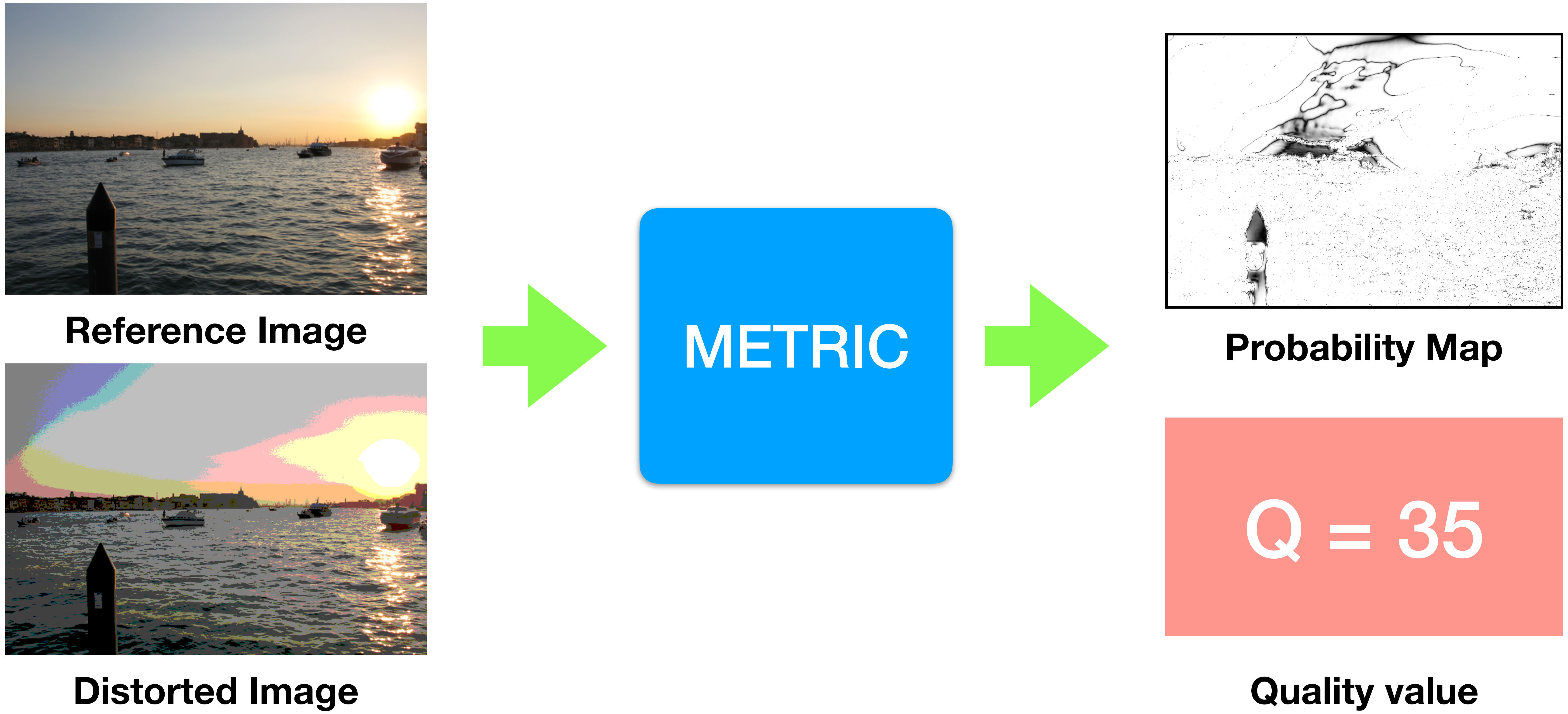
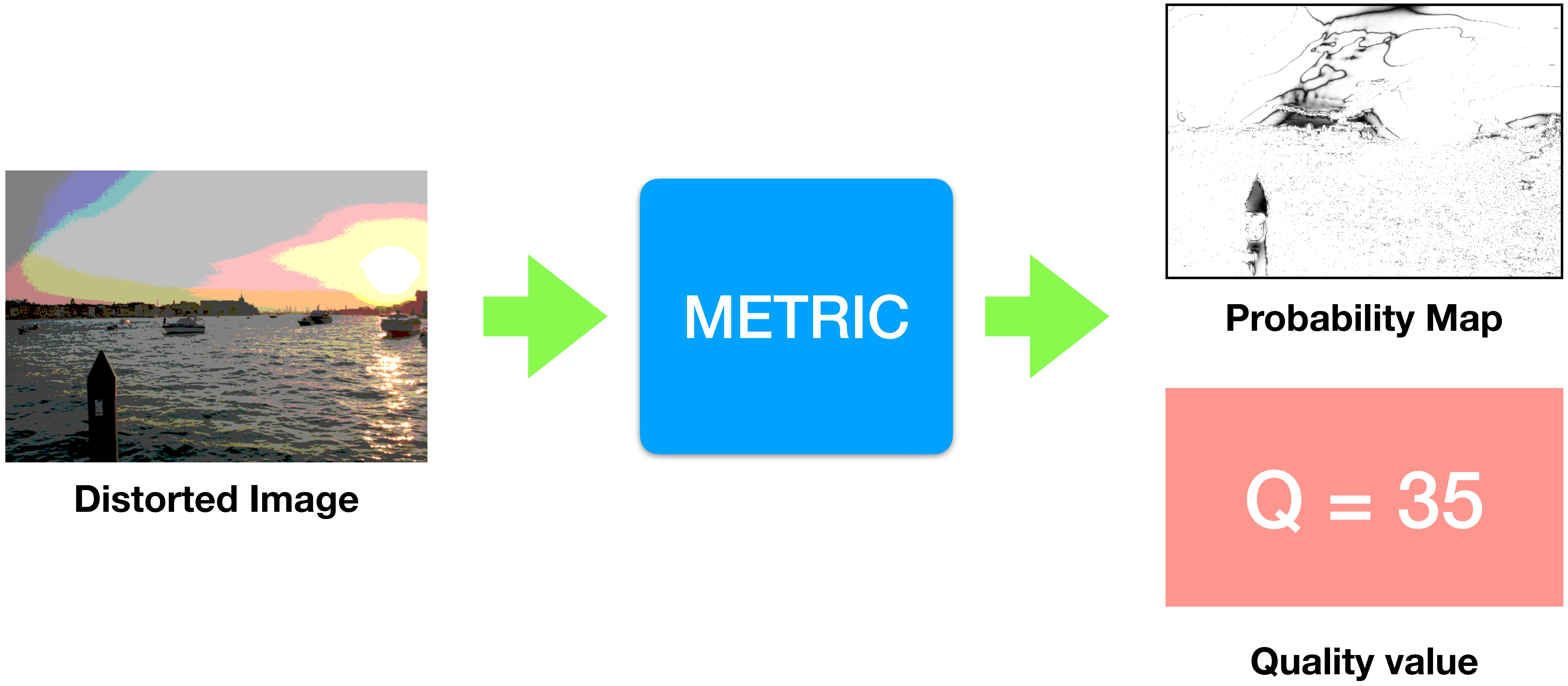


Image Quality Metrics: No Reference

- A probability map; each pixel has the probability of being detected when compared to the reference by a viewer.
- Q predictor value in the range [0,100]; the higher the better.



Metrics for HDR Applications

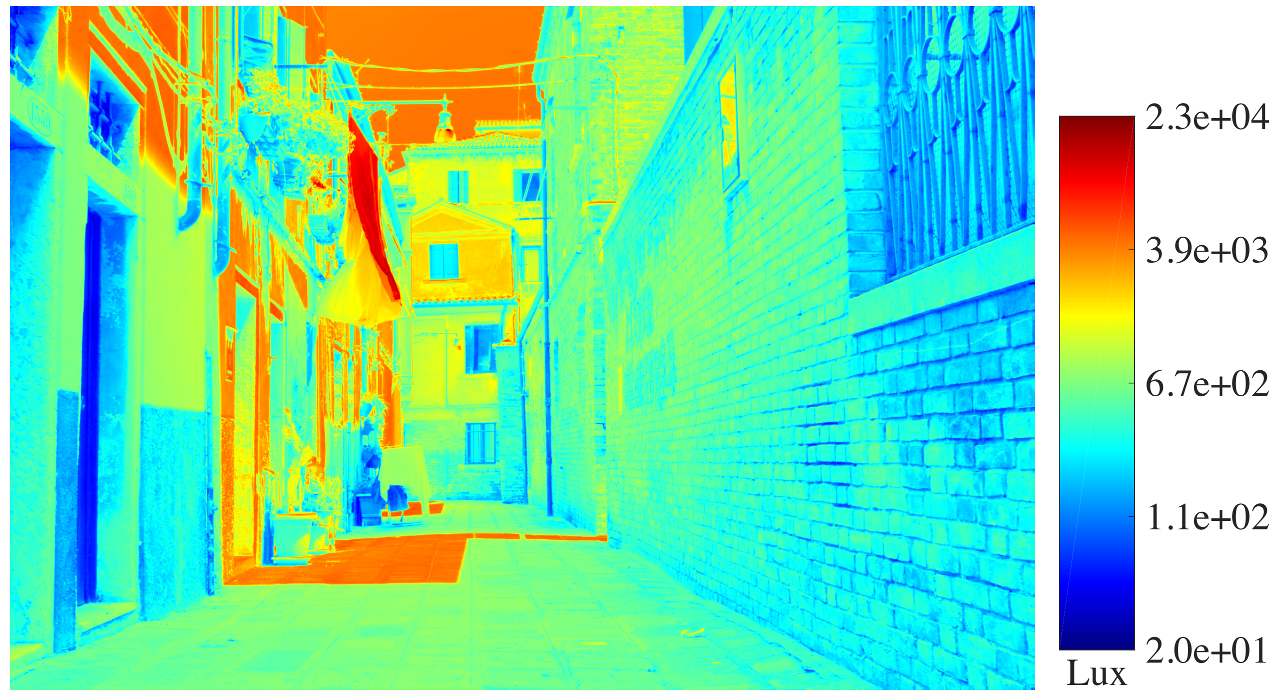
- HDR-VDP 2.2/3.0.6/FoVVDP/DRIM:
    - They are reliable metrics for the general case:
        - HDR vs HDR; HDR vs SDR; etc.
    - Computational cost is demanding.
    - **A reference is required!**
- TMQI and TQMI-II:
    - Limited for comparing HDR vs SDR for tone mapping.
    - **A reference is required!**

HDR Open Problems: Acquisition

HDR Problems:
Merging Exposures in Dynamic Scenes

Stack of 8-bit images

MERGE

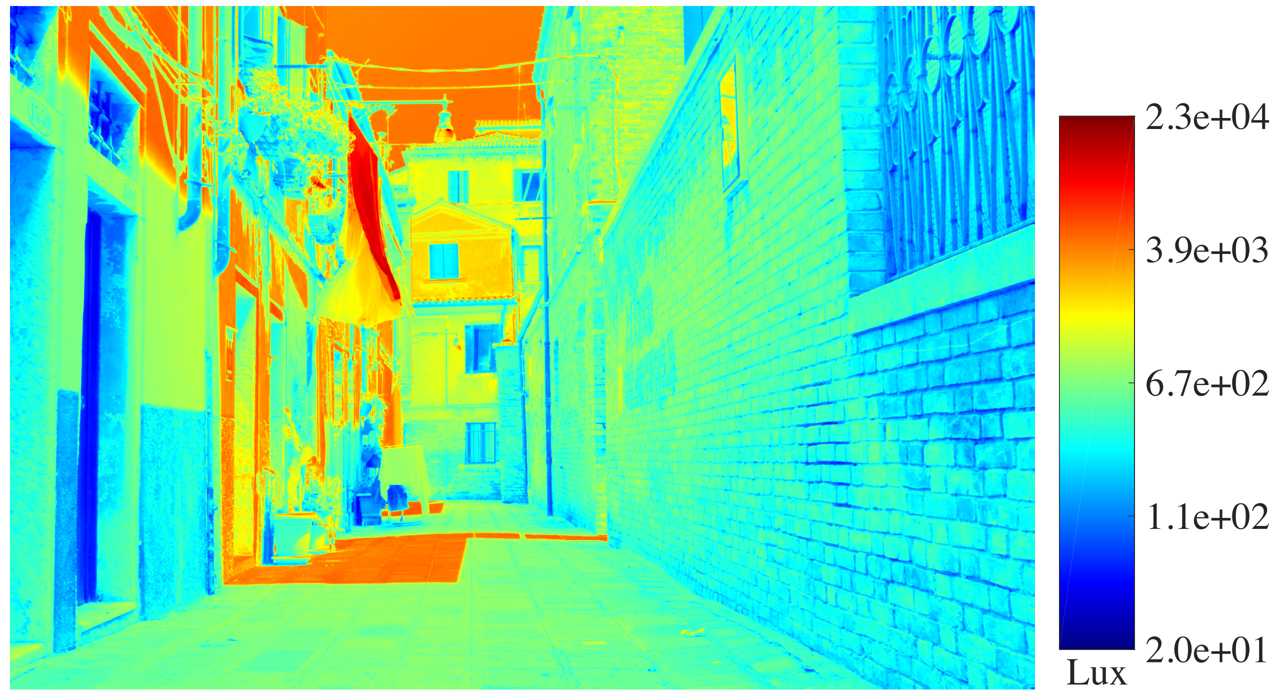


Scene-referred HDR image

HDR Problems:
Merging Exposures in Dynamic Scenes

Stack of 8-bit images

MERGE

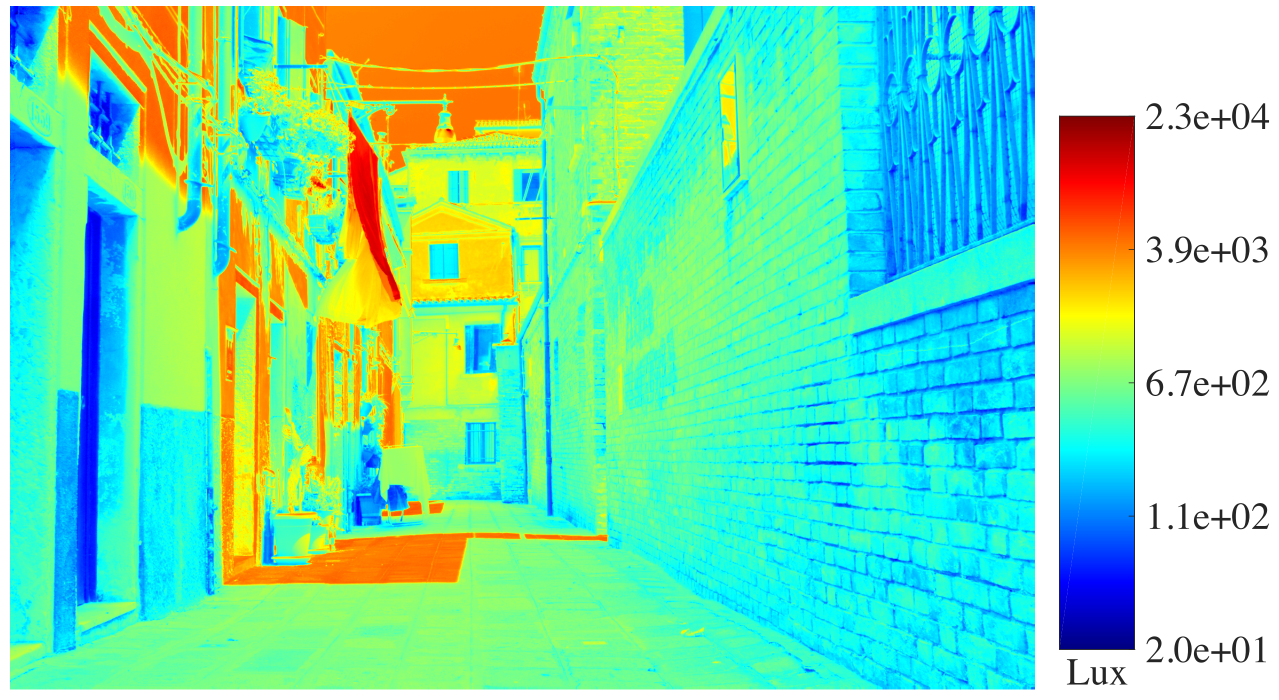


Scene-referred HDR image

HDR Problems:
Merging Exposures in Dynamic Scenes

Stack of 8-bit images

MERGE



Scene-referred HDR image

HDR Problems:
Single-Image Acquisition / Inverse Tone Mapping

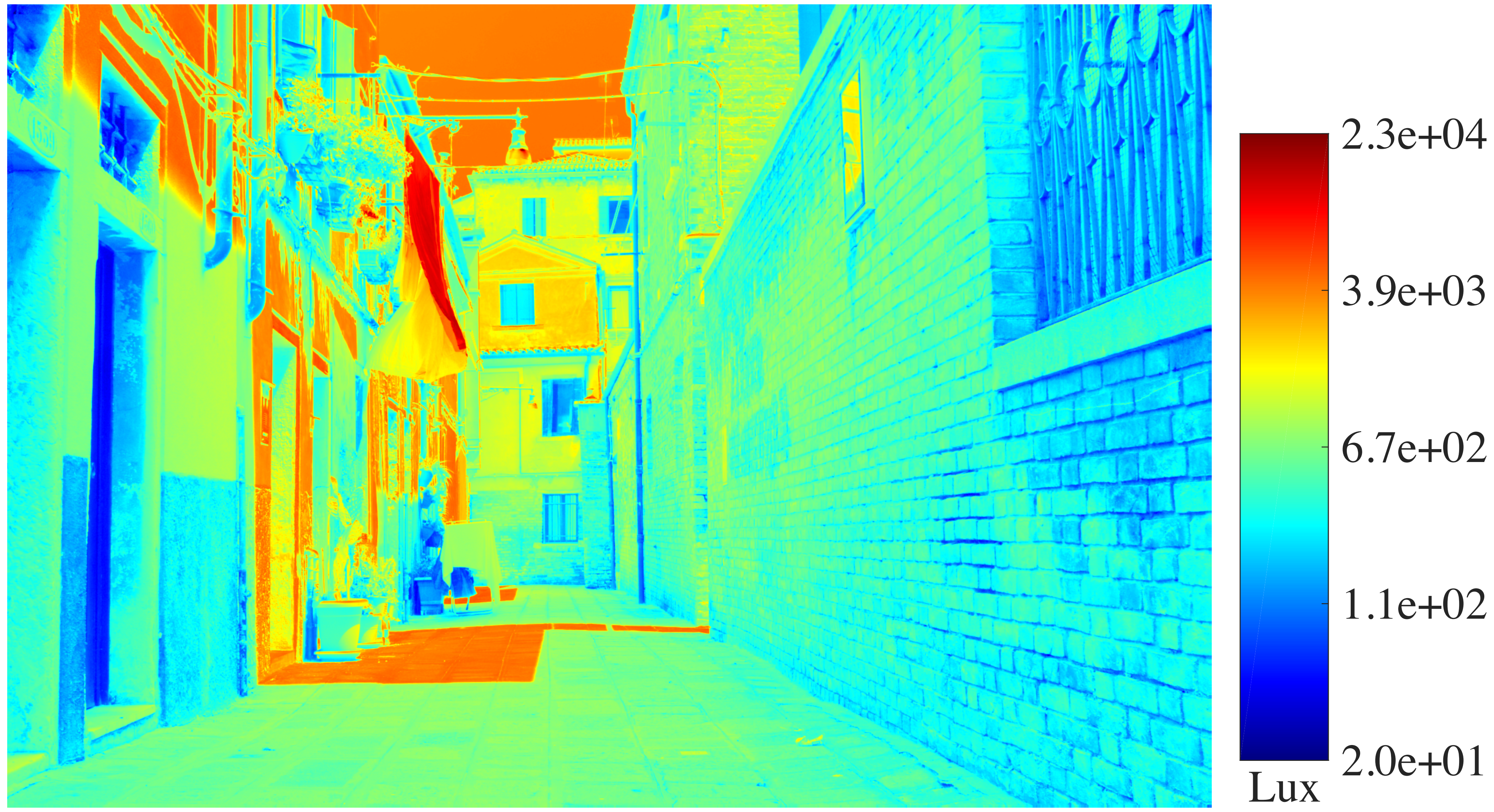
Single 8-bit images

Stack of 8-bit images

Scene-referred HDR image

HDR Open Problems: Visualization

HDR Problems: Tone Mapping



Scene-referred HDR image

TMO

8-bit Tone Mapped Image

HDR Open Problems:
How to Measure the Performance?

How to Measure Performance?

- How do we convert large experiments into metrics?
- Can we speed-up high quality but computationally expensive metrics?
- Can we have no-reference metric?

To Recap

To Recap

- In this tutorial, we will address how to use Deep Learning methods for:
  - Acquiring HDR content;
  - Display HDR images and videos;
  - Metrics for comparing HDR content.

Questions?