

Beyond the Pixel: a Photometrically Calibrated HDR Dataset for Luminance and Color Prediction

Supplementary Materials

Christophe Bolduc, Justine Giroux, Marc Hébert, Claude Demers, and Jean-François Lalonde
 Université Laval

In this document, we present the following additional information to complement the main paper:

- A description of how to acquire the dataset;
- A description of the photometric and colorimetric quantities in sec. 2 to accompany sec. 3 to sec. 6 of the paper;
- Additional information on the calibration, including more details on the capture configurations and the calibration uncertainty in sec. 3 to complement sec. 3 of the paper;
- More visualisations to explore the calibrated dataset in sec. 4 to augment sec. 4 of the paper;
- More results for the learning tasks, including visual examples in sec. 5 to add to sec. 5 of the paper;

1. Acquiring the dataset

The dataset, released along with the paper, is available at http://www.hdrdb.com/indoor_hdr_photometric/. Access to the complete dataset for non-profit or educational organization is provided after a license agreement is signed. Additionally, a sample of the photometric dataset is directly available (100 samples at 2048×1024 pixel resolution). The HDR data, stored in the “.exr” file format, can be visualised using an HDR viewer such as TEV¹.

2. Photometric and colorimetric quantities

2.1. Illuminance and luminance computations

Planar illuminance The equation used to compute the illuminance on a plan from the luminance of the hemisphere [2] is

$$E_p = \int_{\Omega} L(p, \omega) \cos(\theta) d\omega, \quad (1)$$

where $L(p, \omega)$ is the luminance of an area (subtended by solid angle ω) of the hemisphere Ω , and θ is the angle the surface normal of the plane.

¹<https://github.com/Tom94/tev>



Figure 1. To compute the illuminance of the image, the geometric calibration of the camera is used to project the captured HDR (left) to an orthographic projection (right).

When projecting the hemisphere on the plane with an orthographic projection (as is shown in fig. 1), the projected solid angle with relation to the hemispherical solid angle corresponds to

$$d\omega_{\perp} = \cos(\theta) d\omega. \quad (2)$$

Eq. (1) then becomes

$$E_p = \int_H L(p, \omega) d\omega_{\perp}. \quad (3)$$

Discretizing this equation and integrating on a planar pixel grid of N pixels, the illuminance becomes

$$E_p = \frac{\pi}{N} \sum_{i \in \Omega} L(i). \quad (4)$$

This is the equation used in sec. 3.1 of the main paper for the dataset calibration as well as sec. 5 and sec. 6 for computing the ground truth of illuminance prediction.

Mean spherical illuminance The mean spherical illuminance (MSI) [2] is used to measure the quantity of light received at a single point in the scene in the analysis in sec. 4.1 and is defined as

$$E_{\text{ms}} = \int_S L(p, \omega) d\omega, \quad (5)$$

where $L(p, \omega)$ is the luminance of an area (subtended by solid angle ω) of the sphere S .

Discretizing this equation over a planar pixel grid (in equirectangular format) of N pixels gives

$$E_{\text{ms}} = 4\pi \frac{\sum_{(i) \in S'} L(i) d\omega(i)}{\sum_{(i) \in S'} d\omega(i)}, \quad (6)$$

where $d\omega(i)$ is the solid angle subtended by pixel i , and S' represents the subset of valid pixels in the image².

Average luminance For each individual light source analyzed in sec. 4.2, the average luminance is computed as

$$\bar{L} = \frac{\int_A L(p, \omega) d\omega}{\int_A d\omega}, \quad (7)$$

where L is the luminance of the pixel, $d\omega$ is its solid angle, and A is the region which corresponds to the segmented light source.

Its discretized version is defined as

$$\bar{L} = \frac{\sum_{(i) \in A} L(i) d\omega(i)}{\sum_{(i) \in A} d\omega(i)}. \quad (8)$$

2.2. Photopic values

The previous photometric quantities are defined independently of any color space. In our work, we apply the planar illuminance (eq. (4)) for the dataset calibration in sec. 3 and mean spherical illuminance (eq. (6)) in sec. 4.1 directly to each of the RGB channel.

However, we also work with photopic luminance and illuminance, where the equations are applied to the photopic luminance, defined as:

$$L = 0.212671L_R + 0.715160L_G + 0.072169L_B. \quad (9)$$

This is the case for the average source luminance (eq. (8)), luminance visualization (sec. 4) and planar illuminance prediction (sec. 5 and sec. 6).

2.3. Color spaces conversions

CIE Yxy to CIE XYZ The equations allowing the transformation from Yxy to XYZ color spaces [5] are

$$X = \frac{xY}{y}, \quad Y = Y, \quad \text{and} \quad Z = \frac{(1-x-y)Y}{y}. \quad (10)$$

CIE XYZ to CIE Yxy The inverse transformation corresponds to [5]

$$x = \frac{X}{X+Y+Z}, \quad y = \frac{Y}{X+Y+Z}, \quad \text{and} \quad Y = Y. \quad (11)$$

²In practice, not all pixels are valid in the panoramas and a region at the nadir, corresponding to where the tripod was at the time of capture, is all-black.

RGB to CIE XYZ The relation between linear sRGB under reference white D65 and XYZ is given by the following matrix multiplication [3]

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (12)$$

CIE XYZ to RGB The inverse transformation of eq. (12) is approximated as

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (13)$$

Chroma meter RGB conversion To convert the xyY color value captured by the chroma meter to RGB as is done in sec. 3.4 of the paper, the equations eq. (10) and eq. (12) are applied subsequently.

2.4. Color temperature from photometric HDR

We use McCamy's approximation to compute the correlated color temperature (CCT) from the chromaticity in CIE xy format [4] in secs. 4.1, 5 and 6 of the main paper, defined as

$$T = 449n^3 + 3525n^2 + 6823.3n + 5518.87, \quad (14)$$

where

$$n = \frac{x - 0.3320}{0.1858 - y}. \quad (15)$$

The CCT is applied per-pixel to the photometric HDR by first using eq. (12) to convert from RGB to CIE XYZ, then using eq. (11) to convert to CIE xy and finally using eq. (14) to obtain the CCT.

Average CCT The average CCT of a source used in sec. 4.2 is computed on the per-pixel CCT image as

$$\bar{T} = \frac{\int_A T d\omega}{\int_A d\omega}, \quad (16)$$

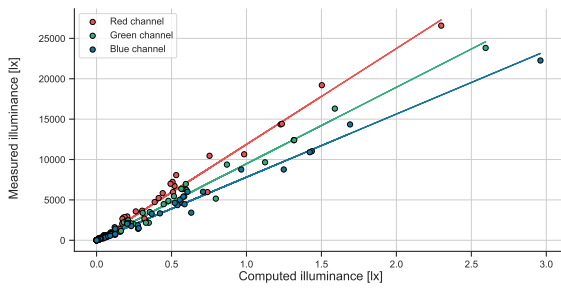
and its discretized version

$$\bar{T} = \frac{\sum_{(i) \in A} T(i) d\omega(i)}{\sum_{(i) \in A} d\omega(i)}. \quad (17)$$

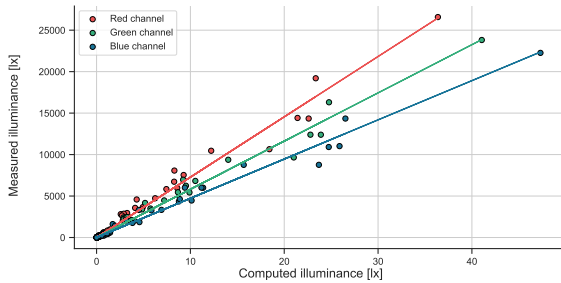
3. Calibration

3.1. Coefficients regressions

The calibration coefficients identified in sec. 3.4 of the paper are computed for each capture configuration. The regressions for each channel of the two main configurations



(a) f/14



(b) f/4

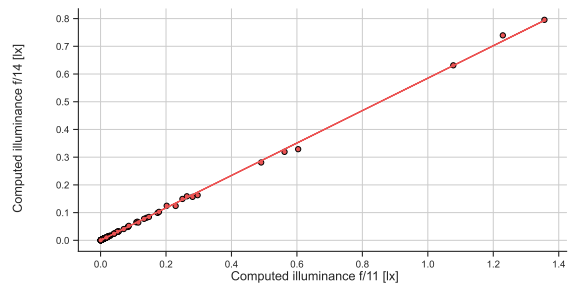
Figure 2. The resulting regression for the measured illuminance with the chroma meter over the integrated illuminance from the HDR images for the aperture of (a) f/14, and (b) f/4. (a) The resulting correction factors (slopes) are (11 872.8, 9472.0, 7814.3) for (R, G, B), with R^2 regression coefficients of determination of (0.985, 0.987, 0.989) respectively. (b) The resulting correction factors (slopes) are (727.5, 581.3, 472.8) for (R, G, B), with R^2 regression coefficients of determination of (0.982, 0.984, 0.982) respectively.

(f/14 and f/4) are shown in fig. 2. Since the other 3 configurations represent but a small minority (2%) of the total number of images (tab. 1), they were not captured for all of the 135 calibration dataset scenes. Instead, they were only captured on a subset (43) of the scenes, and directly compared with the f/14 configuration instead of the chroma meter to compute the relationship with the RGB coefficients at aperture f/14. Since the change in aperture affects all three channels simultaneously, a single coefficient is computed from the three channels. The coefficients regression presented in fig. 3 bring the f/11, f/13 and f/18 configurations respectively to their f/14 equivalent.

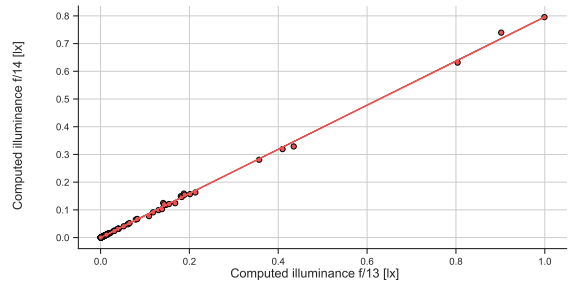
To calibrate the corresponding panoramas, the HDR is first multiplied by the factor correcting to obtain the f/14 equivalent, and the coefficients for each channel of the f/14 configuration are then applied afterwards.

3.2. Uncertainty on calibration

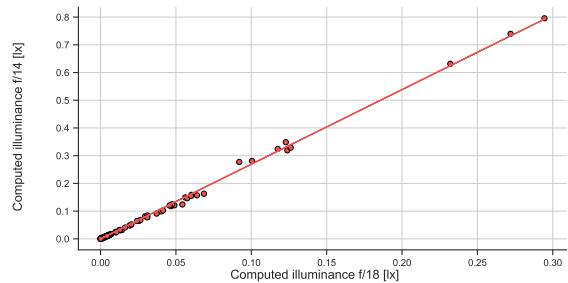
The uncertainty on the calibrated dataset of sec. 3 of the main paper depends on the configuration of the capture for a given panorama. Tab. 1 lists the different configurations



(a) f/14 over f/11



(b) f/14 over f/13



(c) f/14 over f/18

Figure 3. The resulting regression for the integrated illuminance from the HDR images for the aperture of (a) f/14 over the integrated illuminance from the HDR images for the aperture of f/11, (b) f/14 over f/13, and (c) f/14 over f/18. (a) The resulting correction factors (slope) is 0.585, with R^2 regression coefficients of determination of 0.999. (b) The resulting correction factors (slope) is 0.796, with R^2 regression coefficients of determination of 0.999. (c) The resulting correction factors (slope) is 2.69, with R^2 regression coefficients of determination of 0.998.

along with the standard deviation of the linear regression. In all, we achieve very low (less than 1.5%) uncertainty in the recovered luminance values across all configurations and three color channels.

4. Visualisation

To complement fig. 3 of the paper, more examples of scenes contained in the dataset are presented in fig. 4, sorted by their mean spherical illuminance (MSI), with their value close to the quantile indicated. Below are shown the log-

#panos	Aperture	Shutter speed [s±evstop]	R STD [%]	G STD [%]	B STD [%]
540	f/4	1/30 ± 2 2/3	1.43	1.35	1.43
7	f/11	1/30 ± 2 2/3	1.24	1.18	1.10
3	f/13	1/30 ± 2 2/3	1.23	1.17	1.08
1759	f/14	1/30 ± 2 2/3	1.18	1.12	1.03
53	f/18	1/60 ± 2 2/3	1.27	1.22	1.13

Table 1. Uncertainty on the calibration process for each of the 5 capture configurations (aperture and shutter speed) in the dataset. The ISO for each configuration is 100.

luminance maps associated to the scene.

To add to fig. 4 of the paper, fig. 5 shows more examples of scenes, this time sorted by their CCT value. The map below corresponds their associated CCT.

Fig. 6 shows the correlation between the CCT and the average luminance for the individual light sources detected discussed in sec. 4.2 of the paper (10 289 out of 11 060 sources are included in the figure). The distribution of the values are also shown on the edge of the figure. It is possible to see that cooler sources are more frequent than the warmer (which tend to correspond to windows). However, the distribution in average luminance is quite symmetrical.

The average value of the average luminance for all the light sources included in fig. 6 (values in $[50 \text{ cd m}^{-2}, 600\,000 \text{ cd m}^{-2}]$) is $18\,029 \text{ cd m}^{-2}$, with a median of 3991 cd m^{-2} . The average value of the average luminance for all the light sources included in the dataset is $27\,874 \text{ cd m}^{-2}$, with a median of 3854 cd m^{-2} . The average value of the CCT for all the light sources included in fig. 6 (values in $[2000 \text{ K}, 9000 \text{ K}]$) is 3633 K , with a median of 3404 K . The average value of the CCT for all the light sources included in the dataset is 3648 K , with a median of 3380 K .

5. Learning tasks

5.1. Input data

We apply different transformations to the input given to the networks in sec. 5 and sec. 6 of the paper. For per-pixel luminance prediction, random noise is added to the input and gamma and quantization are applied to the image. For per-pixel color prediction, a random WB augmenter [1] is applied the input. Those transformations are visualized in fig. 7.

5.2. Experiments

The following results complement the experiments of sec. 5.3 of the paper.

Per-pixel luminance Fig. 8 shows qualitative predictions, comparing the ground truth luminance to the predicted lu-

minance. Observe that most of the errors are due to incorrect scale prediction.

Per-pixel color Fig. 9 shows the effect of different white balance augmentations. The network is trained and tested on all augmentation settings independently.

Planar illuminance Fig. 10 shows illuminance predictions, along with their given inputs and the full hemispheres used for ground truth illuminance. Light sources being slightly outside the FOV and the unknown camera exposure make illuminance prediction from a single image a very difficult task. We hope our dataset will provide a useful resource to the community to tackle these challenging new problems.

Additionally, the effect of modifying the photometric information of the input is visualized in fig. 11

References

- [1] Mahmoud Afifi and Michael S. Brown. What else can fool deep learning? addressing color constancy errors on deep neural network performance. In *Int. Conf. Comput. Vis.*, 2019. 4, 7
- [2] Commission Internationale de L’Eclairage (CIE). Ilv: International lighting vocabulary, 2nd edition. Technical report, 2020. 1
- [3] Bruce Lindbloom. XYZ to RGB. http://www.brucelindbloom.com/Eqn_XYZ_to_RGB.html, 2017. 2
- [4] C. S. McCamy. Correlated color temperature as an explicit function of chromaticity coordinates. *Color Research & Application*, 17(2):142–144, 1992. 2
- [5] Charles Poynton. *Digital Video and HD: Algorithms and Interfaces*, page 275. Morgan Kaufmann, 2 edition, 2012. 2

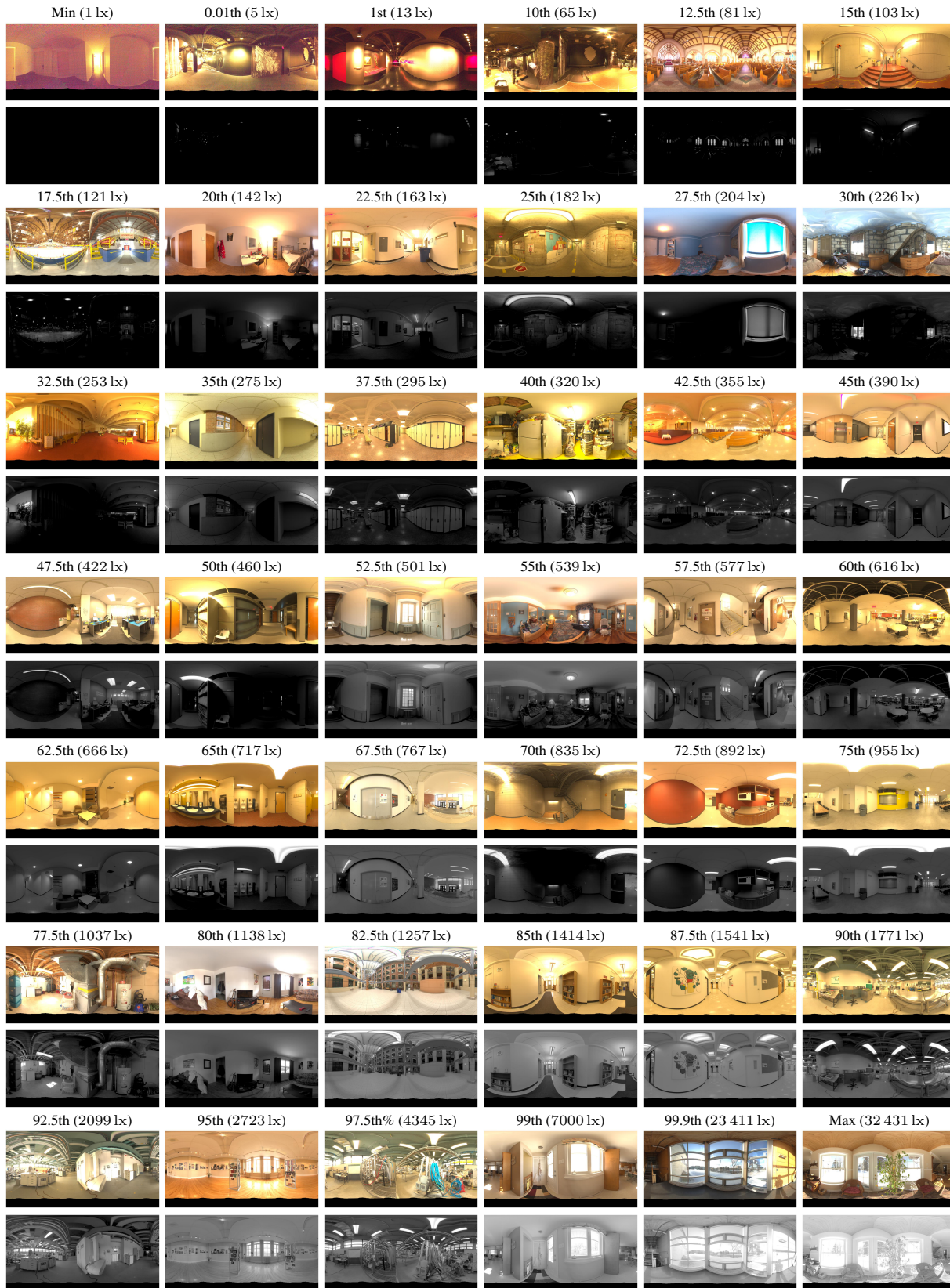
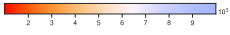


Figure 4. Example scenes with mean spherical illuminance (MSI) close to the quantile values to complement fig. 3 from the main paper. Greyscale images below show the corresponding log-luminance maps. The percentiles and corresponding measured MSI are indicated above the images. Images are reexposed and tonemapped ($\gamma = 2.2$) for display. Luminance color map:



Figure 5. Example scenes with CCT close to the quantile values to complement fig. 4 from the main paper. Colored images below show the CCT map of the scenes. The percentiles and corresponding measured scene CCT are indicated above the images. Images are reexposed and tonemapped ($\gamma = 2.2$) for display. CCT map: 

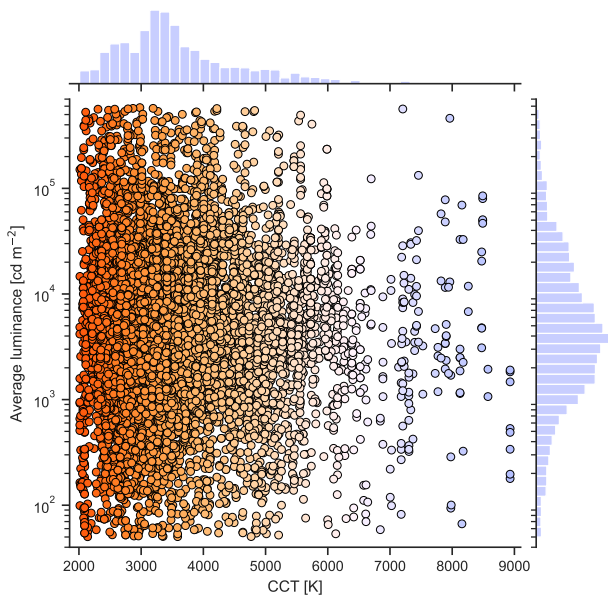


Figure 6. Correlation between the CCT and the average luminance for each light source in our calibrated dataset. The distributions of the CCT (top) and average luminance (right) of the light sources are also displayed. Only the light sources with a CCT in [2000 K, 9000 K] and an average luminance in [50 cd m^{-2} , 600 000 cd m^{-2}] are included to better see the trends (10 289 out of 11 060 sources).



Figure 7. Examples of panorama inputs given to the networks. For visualization, each image is split to show 6 transformations from left to right: Linear, Gamma, Noise, Quantization, Hue. Each input is reexposed and clipped in the range [0, 1]. “Linear” applies no further modification. “Gamma” applies a gamma of $\gamma = 2.2$. “Noise” applies additive Gaussian noise of variance uniformly drawn in the range [0, 0.03]. “Quantization” constrains the input in 255 values. “WB” applies a random WB augmenter [1].

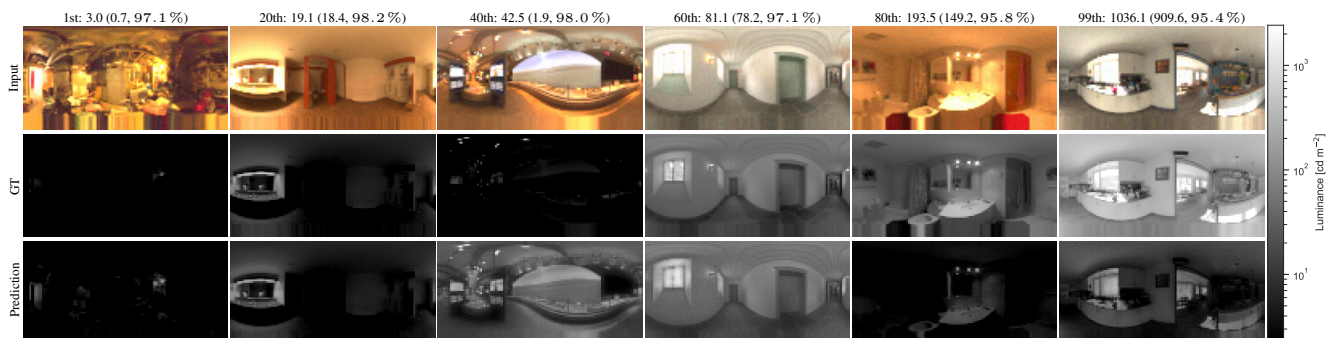


Figure 8. Examples of per-pixel luminance prediction. The first row indicates the RMSE percentile: the RMSE (relative error). The “input” is the calibrated HDR reexposed and clipped. Other rows show the ground truth and predicted photopic luminance maps. The colormap for the luminance is shown at the right.

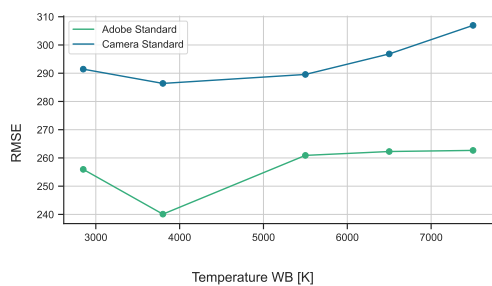


Figure 9. Test scores of color prediction with inputs at different white balance corrections with two different photofinishing profiles. The network is trained on all input corrections.

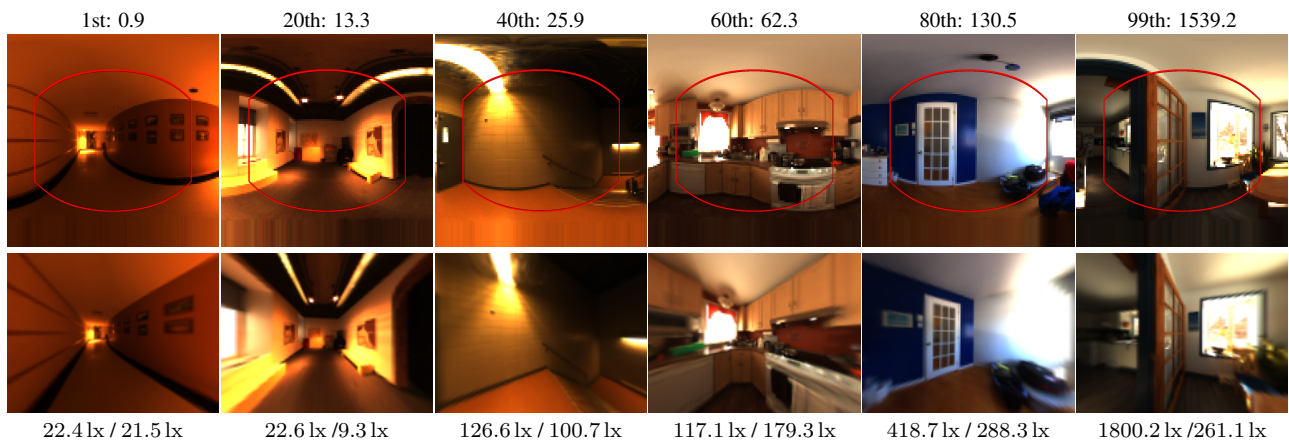
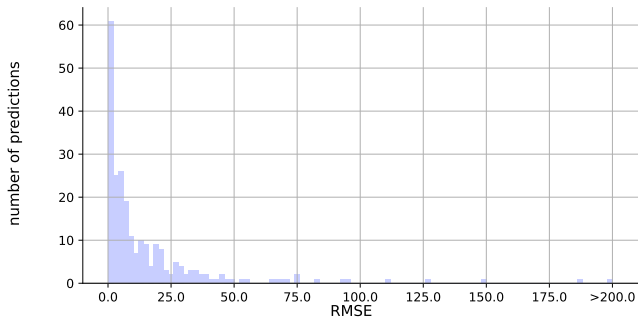
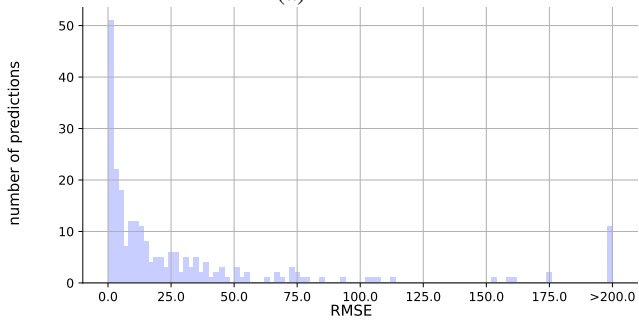


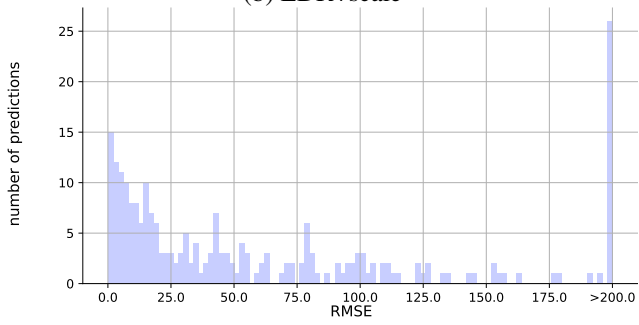
Figure 10. Examples planar illuminance prediction with FOV of 120° . The first row indicates the RMSE percentile: the RMSE. Below are the calibrated HDR hemispheres reexposed and clipped, with the field of view of the image below outlined in red. Below is the projected HDR hemisphere reexposed and clipped given to the network. The last row shows the ground truth and predicted scalars planar illuminance respectively.



(a) HDR



(b) LDR+scale



(c) LDR

Figure 11. The distribution of RMSE scores with different levels of photometric information in the input. The 180° hemisphere image is given as (a) HDR, (b) LDR with photometric scale, and (c) LDR without photometric scale.