SUPPLEMENTARY MATERIAL: Spectral Sensitivity Estimation Without a Camera

Grigory Solomatov, *Member, IEEE,* and Derya Akkaynak, *Member, IEEE*

1 Details for the generation of results in Fig.1

Below, we describe how the results in Fig. 1 of main text were generated using camera spectral sensitivities.

In all examples, the camera RAW images were processed through the Adobe DNG Converter to obtain DNG files, and these were then run through steps 1-4 of the digital camera processing pipeline of Karaimer and Brown [1], which allows users access into the Adobe DNG SDK. This process results in a demosaicked linear tiff image with no additional modifications. The average reflectance spectra of the ColorChecker patches were obtained from Babel Colour and the DGK Kolor Kard patches were obtained from [2].

The error between actual and predicted RGB values were calculated using the angular error:

$$err = cos^{-1}\left(\frac{RGB^{actual} \cdot RGB^{predicted}}{\|RGB^{actual}\|\|RGB^{predicted}\|}\right) \quad (1)$$

and this is used as the objective function for parts **A** and **B**.

1.1 Fig.1 A

The image used is Canon600D_0195.CR2 from the NUS dataset of Cheng et al. [3]. The camera used was a Nikon D40, whose spectral response is made available by Jiang et al. [4]. We discretize this spectral response into $n = 31$ values in the range 400-700 nm, in steps of 10nm.

This dataset also provides the coordinates of the ColorChecker and all its patches, which we used to extract the RGB values of each patch from the linear tiff image. We modeled daylight by the CIE Daylight Model, which is parameterized only by the correlated color temperature (CCT). We formulated an optimization problem in which we minimized the angular RGB between the $m = 24$ actual RGB values (extracted from the ColorChecker) and the predicted RGB values error (Eq. 1 above). The predicted RGB values were obtained using Eq. (1) in main text, with the light spectrum obtained from the CIE daylight model, with the only unknown being CCT. The optimization was implemented in Matlab, using the 'fmincon' function, using a lower bound of 4,000K and an upper bound of 25,000K for CCT. The resulting temperature found for the example shown in Fig. 1A was 5,554 K (approximately CIE D55). The precise optimization problem is

$$\underset{4000\text{K}\leq T\leq 25000\text{K}}{\operatorname{argmin}} \|[\angle(\mathbf{I}_d^{[k,*]}, (\mathbf{R}\mathbf{L}_T\mathbf{S}_d)^{[k,*]})]_{1\leq k\leq m}\|_2 \ ,$$

where the superscript $[k, *]$ means taking the k-th row and $\mathbf{L}_T \in \mathbb{R}^{n\times n}$ is a diagonal illuminant matrix whose diagonal is obtained by discretizing the CIE illuminant series D, i.e.

$$\begin{aligned}
L_T(\lambda) &= L_0(\lambda) + M_1(T)L_1(\lambda) + M_2(T)L_2(\lambda) \ , \text{ where} \\
M_1 &= (-1.3515 - 1.7703x_D(T) + 5.9114y_D(T))/M \ , \\
M_2 &= (0.0300 - 31.4424x_D(T) + 30.0717y_D(T))/M \ , \\
M &= 0.0241 + 0.2562x_D(T) - 0.7341y_D(T) \ , \\
y_D(T) &= -3.000x_D(T)^2 + 2.879x_D(T) - 0.275,
\end{aligned}$$

$$x_D(T) = \begin{cases} 0.244063 + 0.09911\frac{10^3}{T} + 2.9678\frac{10^6}{T^2} - 4.6070\frac{10^9}{T^3} \\ \qquad\qquad \text{if } 4000\text{K} \leq T \leq 7000\text{K} \\ 0.237040 + 0.24748\frac{10^3}{T} + 1.9018\frac{10^6}{T^2} - 2.0064\frac{10^9}{T^3} \\ \qquad\qquad \text{if } 7000\text{K} < T \leq 25000\text{K} \end{cases} .$$

1.2 Fig.1 B

The image used is _DSC0098.tif from the dataset of Akkaynak et al. [2]. The camera used was a Nikon D90, whose spectral response is also made available by Jiang et al. [4].

Note that the color chart used here is not the Macbeth ColorChecker, but the waterproof DGK KolorKard. The $m = 18$ color chart patches were manually extracted and the corresponding RGB values were extracted from the linear tiff image.

Then, the image formation from Eq. (1) of main text was used with the observation that the ambient light will be exponentially attenuated, i.e.,

$$L_{\text{depth}}(\lambda) = L_{\text{surface}}(\lambda)e^{-K_D(\lambda)\times y}, \quad (2)$$

where y is depth and $K_D(\lambda)$ is the diffuse downwelling attenuation coefficient, which is the quantity we would like to estimate. The depth gauge in the photo shows $y = 16.2$ meters.

For surface light, it is generally safe to assume any broadband light (i.e., any CIE D-series light), as the water will attenuate all light to the same monochromatic spectrum, so here we assumed CIE D65. Again using Eq. (1) from main text to calculate the predicted RGB values, we minimize the angular error between predicted and observed RGB values using the 'fmincon' function in Matlab using a lower bound of 0 and an upper bound of 1 m^{-1} for $K_D(\lambda)$.

It is important to note here that $K_D(\lambda)$ cannot be solved for using 31 unknowns because this color chart has much fewer linearly independent patches. By trial and error, we ended up

using discretization resolution $\widehat{n} = 10$ to approximate $K_D(\lambda)$ by a linear interpolation $\widehat{K}_D(\lambda)$ of a vector $\boldsymbol{K} \in \mathbb{R}^{\widehat{n}}$. The precise optimization problem is

$$\underset{\substack{\boldsymbol{K}\in\mathbb{R}^{\widehat{n}} \\ \min(\boldsymbol{K})\geq 0 \\ \max(\boldsymbol{K})\leq 1}}{\operatorname{argmin}} \; \|[\angle(\mathbf{I}_d^{[k,*]}, (\mathbf{R}\mathbf{L}_{\boldsymbol{K}}\mathbf{S}_d)^{[k,*]})]_{1\leq k\leq m}\|_2 \;, \tag{3}$$

where $\mathbf{L}_{\boldsymbol{K}} \in \mathbb{R}^{n\times n}$ is the diagonal matrix whose diagonal is obtained by discretizing $L_{\mathrm{D65}}(\lambda)e^{-\widehat{K}_D(\lambda)\times y}$, where $\widehat{K}_D(\lambda)$ is the linear interpolation of $\boldsymbol{K}$.

1.3 Fig.1 C

The cameras used in this example were a Nikon D40, and Canon600D, whose spectral responses were both made available by Jiang et al. [4].

To do the raw-to-raw mapping example here, we followed the illumination-invariant method of [5]. We skipped the pairwise calibration step, since we do not have physical access to the cameras. Instead, we used the spectral responses of both cameras, and published spectral power distributions of 23 standard CIE illuminants (namely CIE F1-12, A,B,C, D40-D75 illuminants) to compute a global mapping between the two cameras as described in [5], using Macbeth ColorChecker reflectances.

Next, we obtained the "white-balanced' mapping by white balancing the ColorChecker values for each illuminant using the white patch of the color chart (even though the white Macbeth patch does not have a perfectly flat reflectance spectrum).

We then picked example images from the NUS dataset of Cheng et al. [3]. The Canon image is Canon600D_0091.CR2, and the Nikon image is NikonD40_0004.NEF. Using gray-world algorithm, we estimated the illuminant in each image and white balanced the source image.

Next, we applied the white-balanced global mapping to the source image, to obtain its projection into the target camera space. Then, we needed to white balance this image to the correct value in the target camera's color space; so we used the global transform to obtain the correct white point, and inverse-white balance the transformed image. We then completed the in-camera processing using the pipeline software of [1] to obtain the photofinished sRGB image.

1.4 Fig.1 D

The camera used in this example was a Nikon D40 whose spectral response was made available by Jiang et al. [4]. The images and the corresponding ground-truth illumination data are from the NUS dataset of Cheng et al. [3]. There were 117 images for the NikonD40 folder of this dataset. We manually scored whether each photo was taken indoors or outdoors.

Next, we calculated the "white-point' of every daylight illuminant between 4,000 and 25,000 K, using the CIE daylight model, in the color space of the Nikon D40 camera. We then plotted the locus of "daylight chromaticities" for this camera, where chromaticity r is given as:

$$r = \frac{R}{(R+G+B)}, \tag{4}$$

and the b chromaticity is given similarly. Finally, we computed the r and b chromaticities of the provided ground-truth illuminations plotted them in the same coordinates.

2 Results on our ground-truth dataset

Here, we present all the results on our ground-truth dataset in Fig. 1 through Fig. 4. As in Fig. 3 and Fig. 5 from the main paper, the horizontal and the vertical axes are respectively wavelength (nm) and relative sensitivity.

3 Relative full scale error versus $\Delta E2000$

In the main text, we calculated reconstruction errors between our predicted curves and the ground truth using the relative full scale error metric. Here, using the CIE $\Delta E2000$ metric, we demonstrate the perceptual projection of these errors relative to the human visual system. We use the case of the Canon 5D Mk II, for which two ground truth values exist (Fig. 3 in main text). We simulated the appearance of a Macbeth ColorChecker under D65 light according to the Kawakami et al [6] ground truth spectral sensitivity curves for the Canon 5D Mk II, and Jiang et al. [4] ground truth. We white balanced the resulting simulations using the white patch. The distributions of the $\Delta E2000$ error resulting from the comparison between our predictions and the two ground truth curves are given in Fig. 5.

The color differences arising from differences in spectral sensitivity curves are difficult to discern when looking at the simulated charts, however they are reflected into the $\Delta E2000$ errors. Considering that $\Delta E2000 = 1$ marks the discriminability threshold below which two solid colors are indiscriminable to the human eye, as expected, large errors in spectral sensitivity reconstruction translates to large color differences that would be visible.

4 A big picture look at the Adobe DNG Converter dataset

Here, we show a breakdown of the statistics of the 1000 camera for which we were able to extract color matrices from the Adobe DNG Converter. The information directly available from the Adobe DNG Converter were the manufacturer name, and the make & model of the camera. We manually compiled the rest of the sensor information including sensor size, format, camera type and release year. While we made every effort to ensure the accuracy of these data and performed checks and validations, some erroneous entries may still be present.

References

[1] H. C. Karaimer and M. S. Brown, "A software platform for manipulating the camera imaging pipeline," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 429–444.

[2] D. Akkaynak, T. Treibitz, T. Shlesinger, Y. Loya, R. Tamir, and D. Iluz, "What is the space of attenuation coefficients in underwater computer vision?" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4931–4940.

[3] D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," *JOSA A*, vol. 31, no. 5, pp. 1049–1058, 2014.

[4] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk, "What is the space of spectral sensitivity functions for digital color cameras?" in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013, pp. 168–179.

[5] R. Nguyen, D. K. Prasad, and M. S. Brown, "Raw-to-raw: Mapping between image sensor color responses," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3398–3405.

[6] R. Kawakami, H. Zhao, R. T. Tan, and K. Ikeuchi, "Camera spectral sensitivity and white balance estimation from sky images," *IJCV*, vol. 105, no. 3, pp. 187–204, 2013.

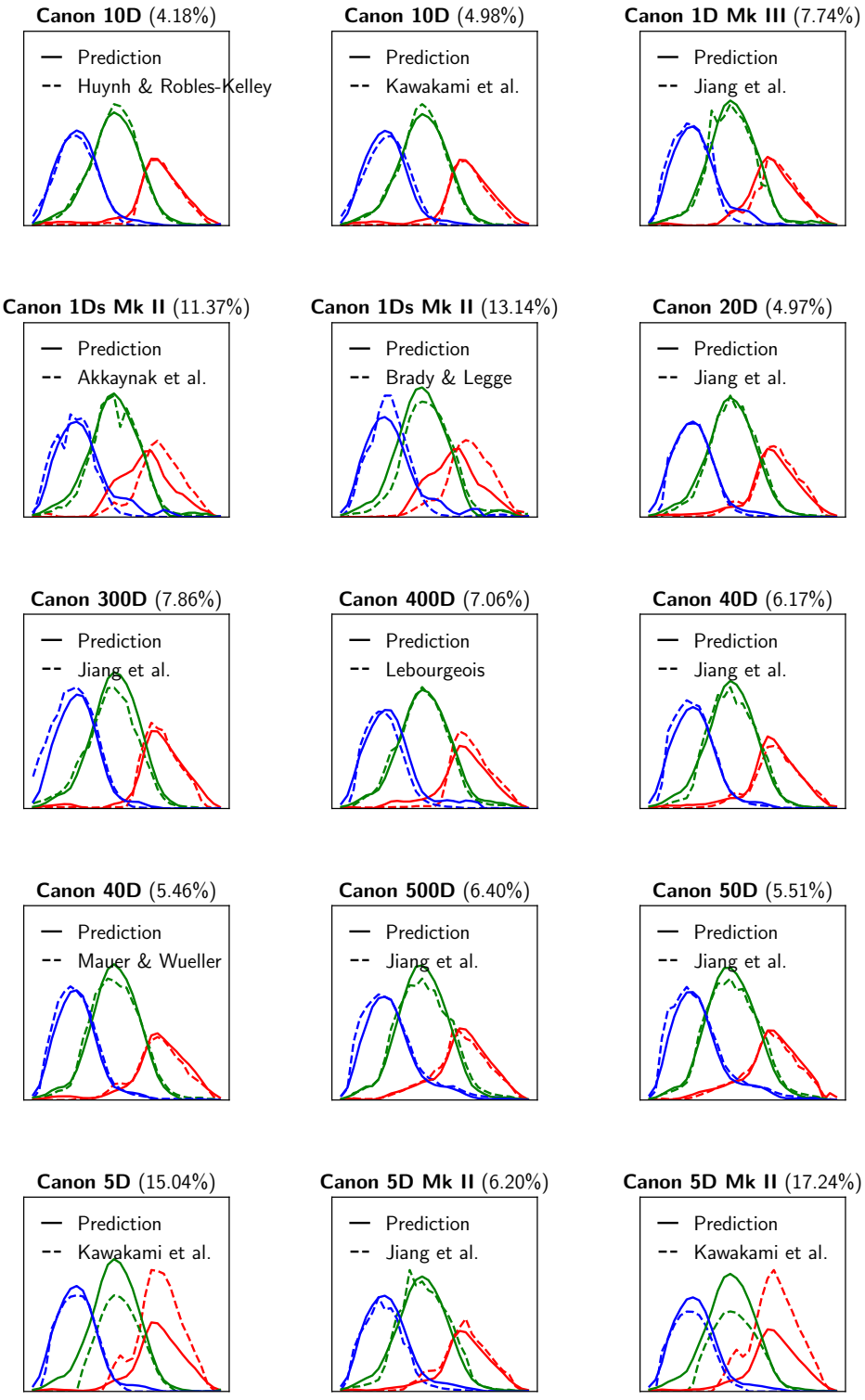

Fig. 1. The results of our method on the ground-truth dataset (Part 1).

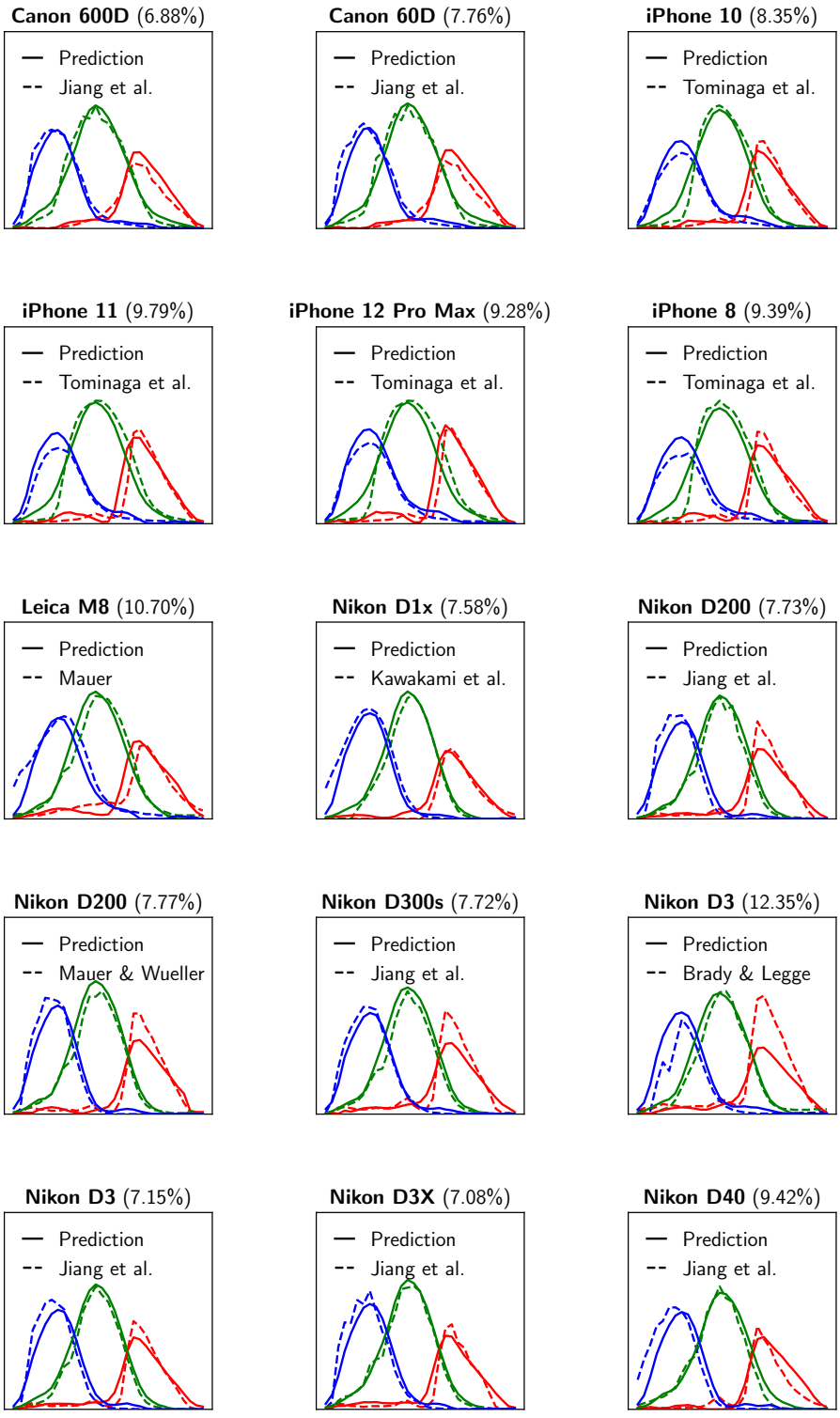

Fig. 2. The results of our method on the ground-truth dataset (Part 2).

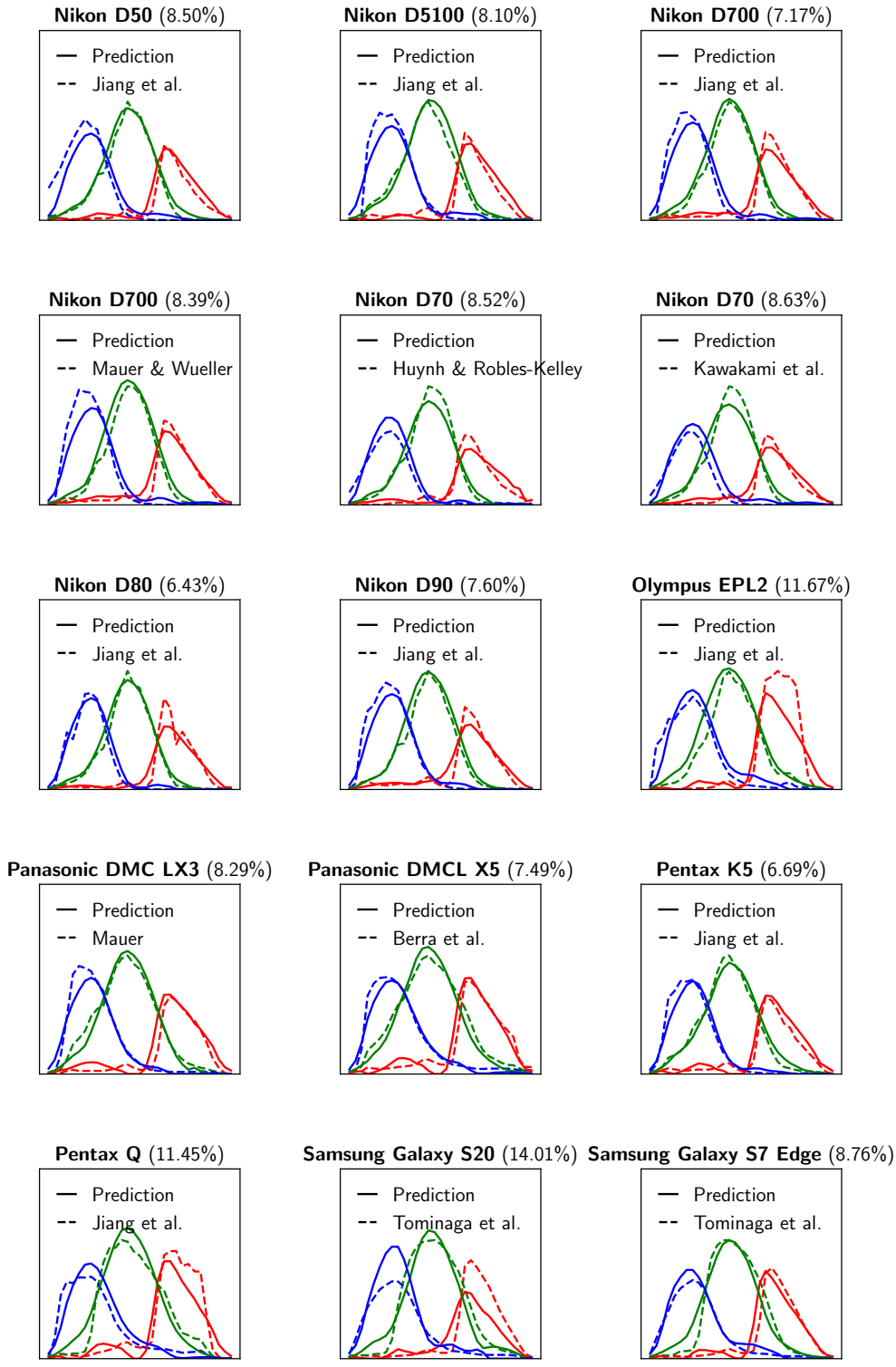

Fig. 3. The results of our method on the ground-truth dataset (Part 3).

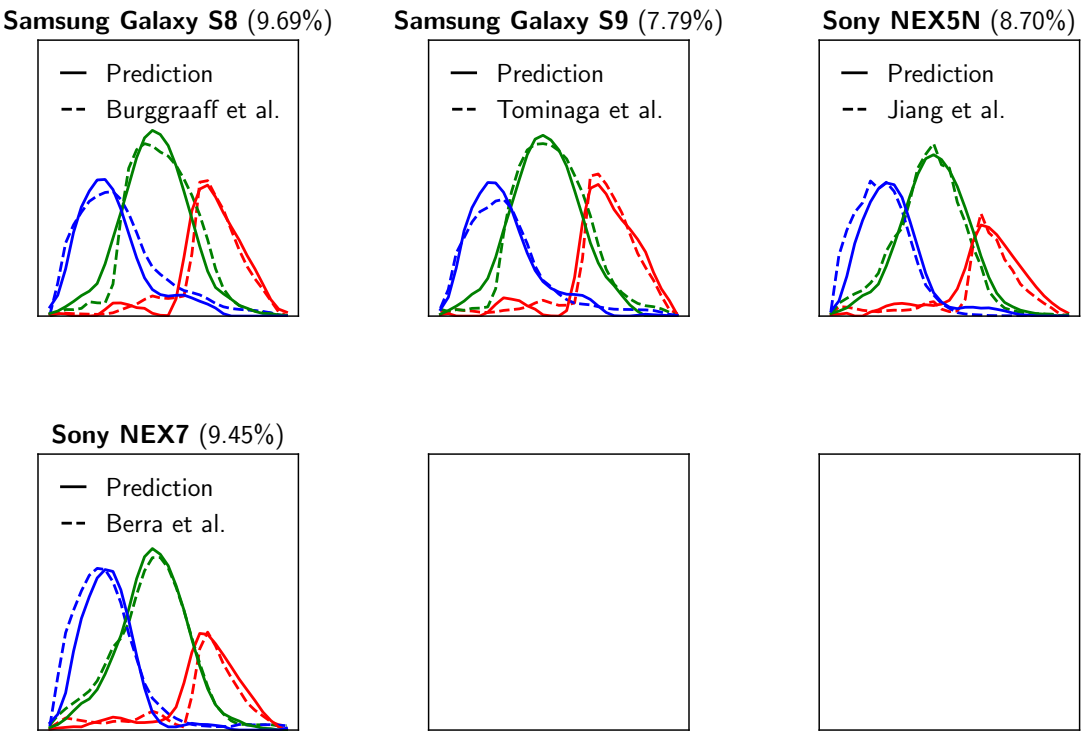

Fig. 4. The results of our method on the ground-truth dataset (Part 4).

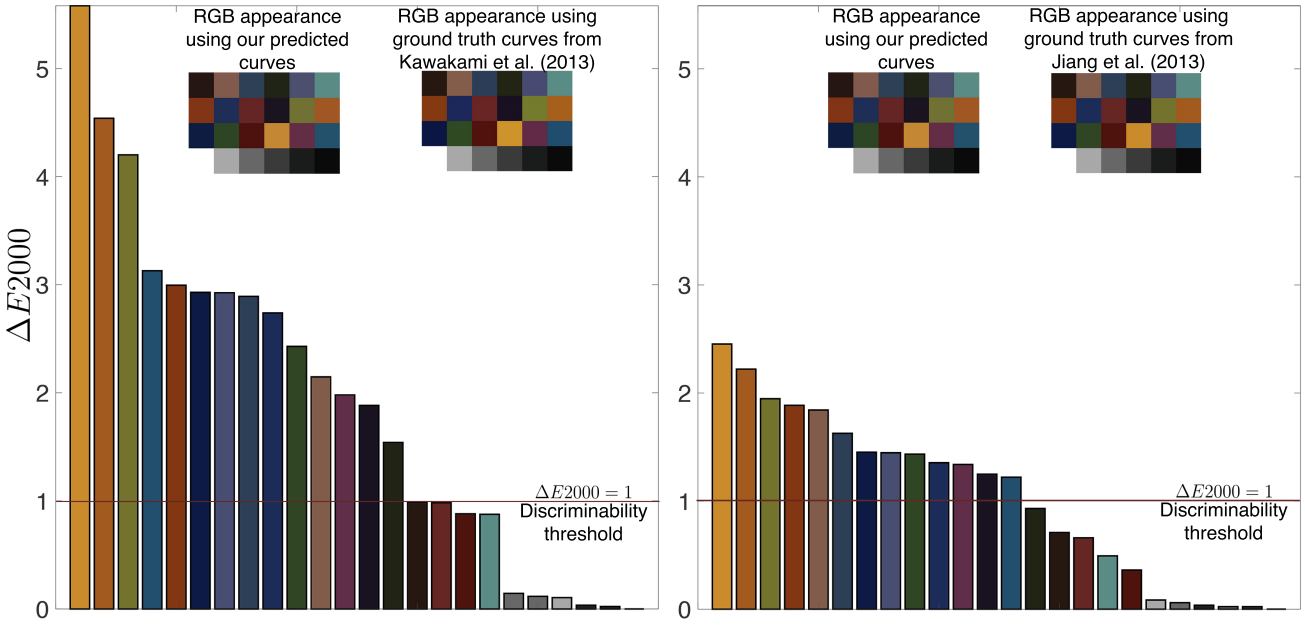

Fig. 5. Perceptual ($\Delta E2000$) errors between high-error (Kawakami et. al ground truth) and low-error (Jiang et al. ground truth) spectral sensitivity reconstructions. While the differences in the simulated color charts are hardly visible to the eye, the perceptual differences can be quantified.

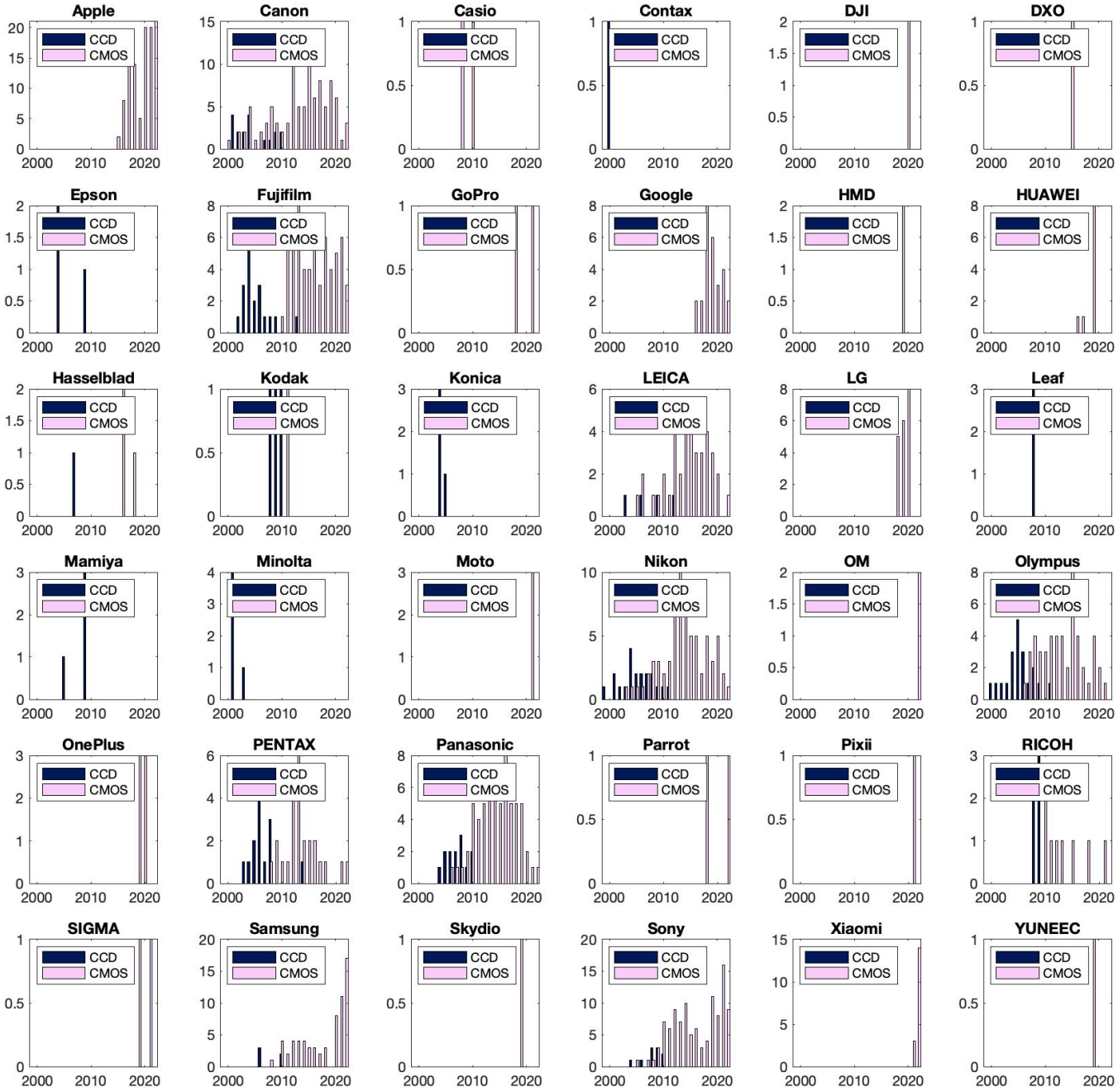

Fig. 6. Sensor types of the cameras in the Adobe dataset.

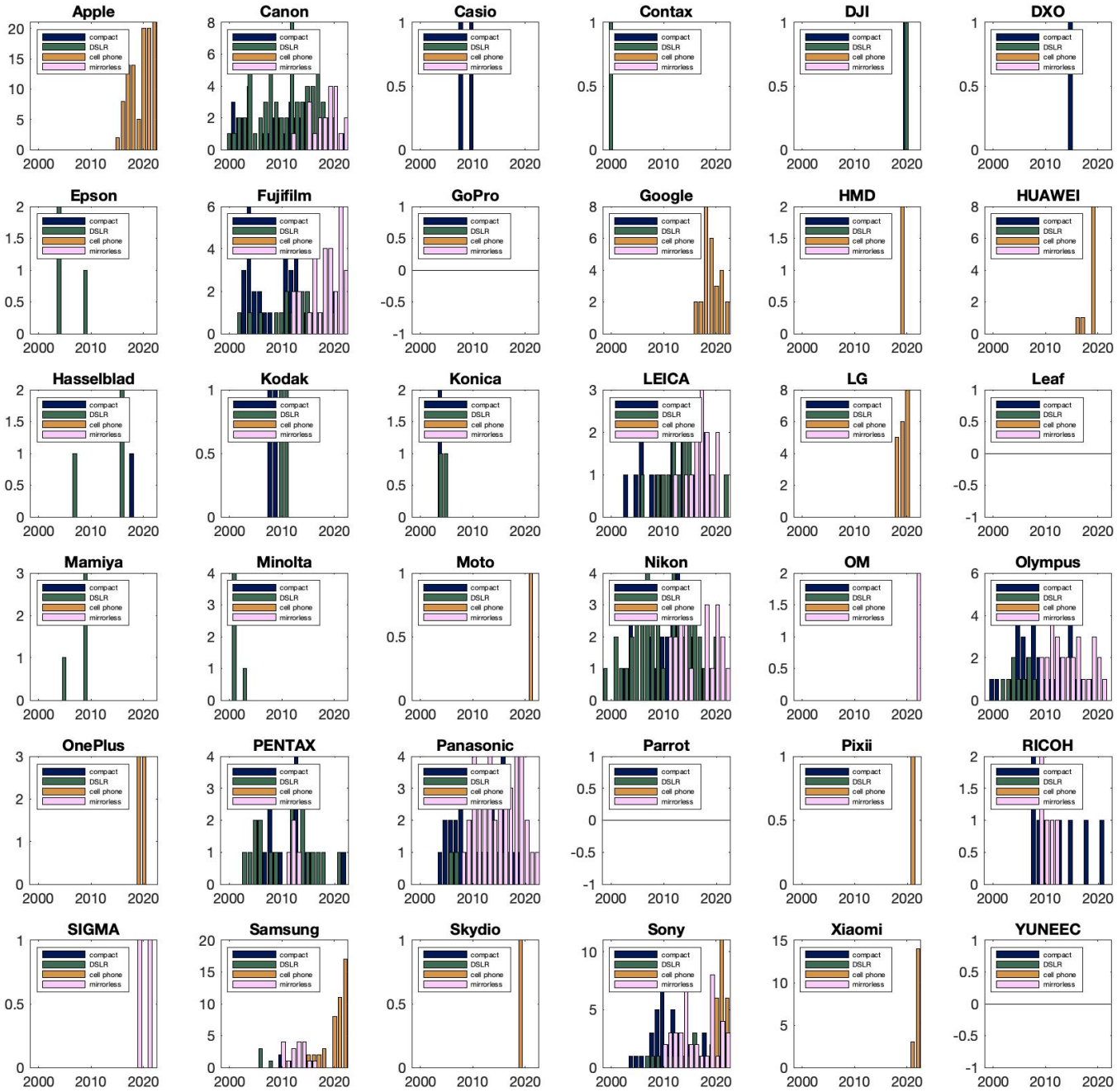

Fig. 7. Types of the cameras in the Adobe dataset according to release year of the camera.