Forget Luminance Conversion and Do Something Better

Rang M. H. Nguyen
National University of Singapore
nguyenho@comp.nus.edu.sg

Michael S. Brown
York University
mbrown@eecs.yorku.ca

Supplemental Material

This supplemental material provides additional results in Sec. 1 that we were unable to include in the main paper due to the page limit. In addition, in Sec. 2 we explain how to use *average*-RGB for color conversion for tone-mapping and why *average*-RGB and the HSV transform can preserve color chromaticities.

1. Additional Experiments

Fig. 1 shows the quantitative error of the luminance channel for the Sony $\alpha200$ camera. The proper white-balance is applied, however, instead of the 2.2 sRGB encoding gamma, we used the camera-specific tone-curves from [3] to generate the images. When we linearize the sRGB image, however, we use default 2.2 decoding gamma. These experiments tests the role of the camera-specific tone-curve for converting sRGB back to luminance values.

Fig. 2 shows the quantitative error between the luminance synthesized by the CIE XYZ color matching functions (ground truth) and the real sRGB images from the Nikon D40 camera. The top row shows the comparison between ground truth luminance and the luminance from the linearized sRGB using sRGB gamma correction. The bottom row shows the comparison between the ground truth luminance and the luminance from the linearized sRGB using the camera's tone-curve measured in [3]. The results show that it is very important to use the correct tone-curves to linearize the RGB color values before computing luminance.

In the next experiment, we use a Specim's PFD-CL-65-V10E hyperspectral camera to capture the spectral power distributions of five different scenes. We synthesize sRGB images from these hyperspectral images for the following two cameras: a Canon 1Ds Mark III, and a Nikon D40. As mentioned in the main paper, the sensor sensitivity functions for these cameras were provided by Jiang et al [2]. To establish the ground truth luminance for a scene, we apply the CIE XYZ matching functions directly to its spectral scene to obtain Y.

We compare the ground truth luminance with the luminance obtained using three methods, namely YIQ, HSV and *average*-RGB. The images are rendered with the proper white-balance and an encoding gamma of 2.2. This means the input images are as close to ideal sRGB as possible. We apply these approaches using the proper sRGB decoding gamma and without any linearization. The examples without linearization are referred to as "luma" conversions.

Tab. 1 shows the quantitative error for five different scenes under these two cameras. The table shows that improper conversion (with linearization) results in errors ranging from 1% to 10% for two different cameras. The estimation using luma, however, results in significant errors, with average errors ranging from 20% to over 40%. The scenes and the qualitative errors of each method for two cameras are shown in Figs. 3 and 4.

Fig. 5 shows additional examples of using a simple conversion Y from YIQ and the saliency-preserving decolorization [5] on feature detection such as SIFT [4] and Canny edge detection [1] . As can be seen, the saliency-preserving decolorization [5] helps to preserve the color contrast and allowing SIFT and Canny to obtain better features than the simple conversion Y of YIQ. However when faster processing is not needed, processing all three color channels independently and aggregating the results often give the best performance.

For further evaluation, we also provided a quantitative analysis for edge detection since we can create synthetic images (albeit somewhat unrealistic) that have ground truth edges. Figure 6 show two examples of synthetic images for the task of edge detection. As can be seen, luminance channel is not always the best choice, there are better alternatives such as color-to-gray method proposed by [5] or using all three color channels. Noting that providing quantitative features for something like SIFT is challenging, since there is no way to establish ground truth.
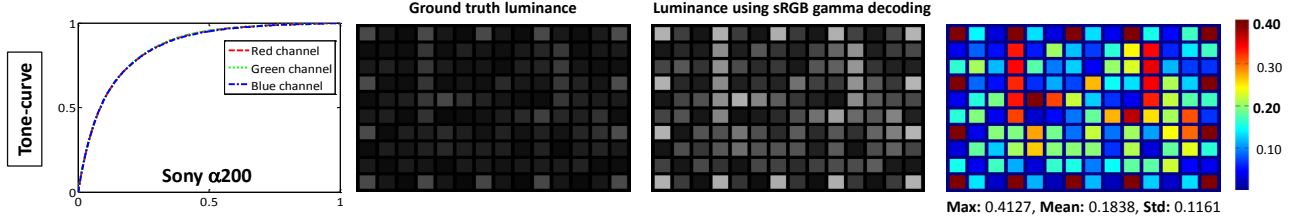
Figure 1. This figure shows the errors that occur when the camera's true tone-curve (Sony α200) is not used to linearize the sRGB values.

2. Conversion for Contrast Adjustment

Average-RGB As discussed in the main paper, the *average*-RGB defines a single brightness channel. Therefore, two more channels have to be used to reconstruct back to RGB color space. This can be done using two additional variables, c and d, as follows:

$$
\begin{aligned}
I &= (R + G + B)/3 \\
c &= R/I \\
d &= G/I
\end{aligned}
\qquad . \tag{1}
$$

After contrast adjustment, the new brightness value I' is obtained, and the new RGB image is reconstructed as follows:

$$
\begin{aligned}
R' &= cI' \\
G' &= dI' \\
B' &= 3I' - R' - G'
\end{aligned}
\qquad . \tag{2}
$$

The new RGB image is normalized (e.g. the maximum value is equal to 1). This formulation will preserve the chromaticities of all colors in image after tone-mapping process.

Proof: Consider a pixel in the input image (R_i, G_i, B_i). Using *average*-RGB conversion, we have:

$$
\begin{aligned}
I_i &= (R_i + G_i + B_i)/3 \\
c_i &= R_i/I_i \\
d_i &= G_i/I_i
\end{aligned}
\qquad . \tag{3}
$$

After contrast adjustment, the new brightness value I_i' is obtained. Let $\alpha = I_i'/I_i$ or $I_i' = \alpha I_i$. After reconstructing back to RGB color space, we have:

$$
\begin{aligned}
R_i' &= c_i I_i' = R_i/I_i \times \alpha I_i = \alpha R_i \\
G_i' &= d_i I_i' = G_i/I_i \times \alpha I_i = \alpha G_i \\
B_i' &= 3I_i' - R_i' - G_i' = \alpha R_i + \alpha G_i + \alpha B_i - \alpha R_i - \alpha G_i = \alpha B_i
\end{aligned}
\qquad . \tag{4}
$$

The output pixel after contrast adjustment is $(R_i', G_i', B_i') = (\alpha R_i, \alpha G_i, \alpha B_i) = \alpha(R_i, G_i, B_i)$. This shares the same chromaticity with the input color pixel (Q.E.D).

The HSV color space uses a similar technique as described above, and as a result can also preserve the chromaticities of all colors after tone-mapping operation. As such, for the case of contrast adjustment, we advocate the use of *average*-RGB or HSV over attempting other luminance conversations.

References

[1] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986. 1

[2] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *WACV*, pages 168–179, 2013. 1, 3, 4, 5

[3] H. Lin, S. J. Kim, S. Susstrunk, and M. S. Brown. Revisiting radiometric calibration for color computer vision. In *ICCV*, pages 129–136, 2011. 1, 3

Tone-curve

Nikon D40

Ground truth luminance | sRGB gamma decoding | Luminance using sRGB gamma decoding | Max: 0.2029, Mean: 0.1179, Std: 0.0420

Nikon D40 inverse tone-curve | Luminance using the inverse tone-curve | Max: 0.0792, Mean: 0.0258, Std: 0.0191
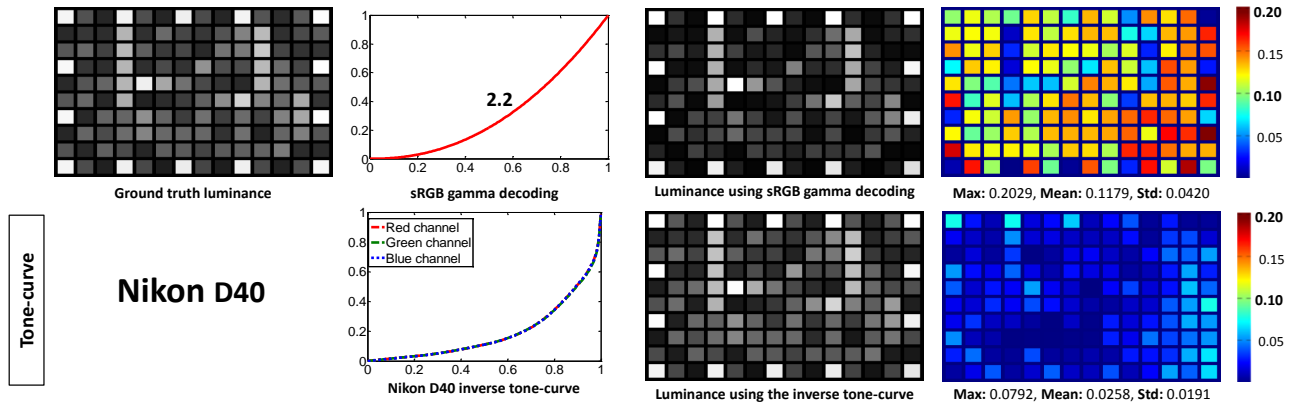
Figure 2. This figure shows the quantitative error between the luminance synthesized by CIE XYZ color matching functions (ground truth) and real sRGB image from the camera Nikon D40. The top row shows the comparison between ground truth luminance and the luminance from the linearized sRGB using sRGB gamma correction. The bottom row shows the comparison between ground truth luminance and the luminance from the linearized sRGB using the camera's tone-curve measured in [3].

Luminance Conversion - Canon 1Ds Mark III

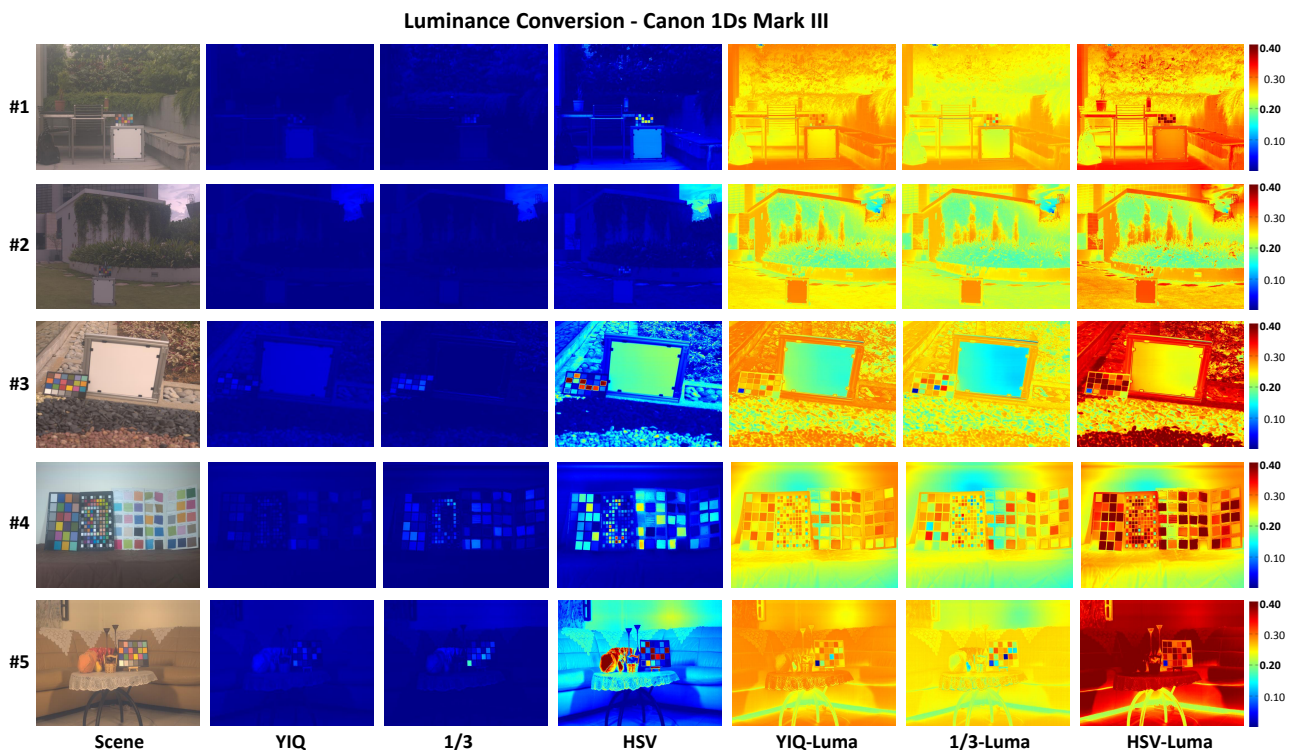Scene | YIQ | 1/3 | HSV | YIQ-Luma | 1/3-Luma | HSV-Luma

Figure 3. This figure shows qualitative error for the synthetic images of five different scenes using camera sensitivity functions of camera Canon 1Ds Mark III in [2]. The gamma of 2.2 is applied to obtain the sRGB images.

[4] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004. 1

[5] C. Lu, L. Xu, and J. Jia. Contrast preserving decolorization with perception-based quality metrics. *IJCV*, 110(2):222–239, 2014. 1, 6
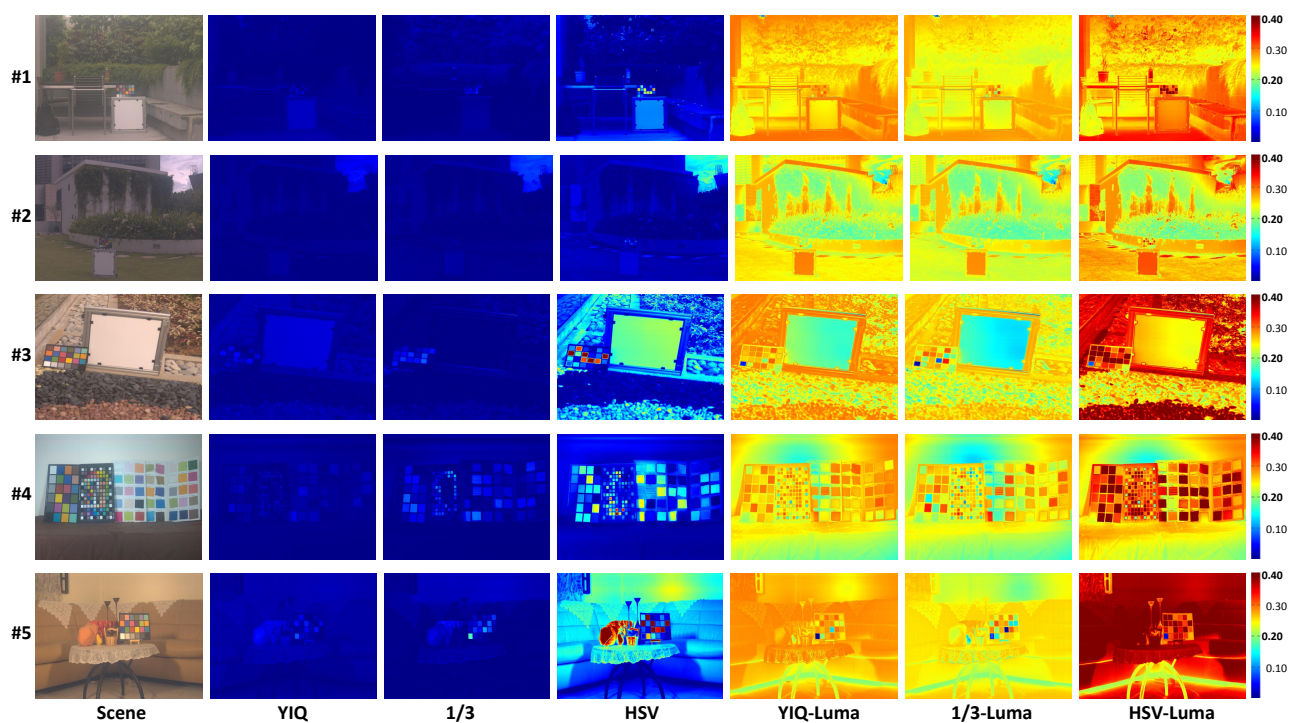
Figure 4. This figure shows qualitative error for the synthetic images of five different scenes using camera sensitivity functions of camera Nikon D40 in [2]. An encoding gamma of 2.2 is applied to obtain the sRGB images.

Scene	Method	Canon 1Ds Mark III			Nikon D40		
		Max	Mean	Std	Max	Mean	Std
#1	YIQ	0.0431	0.0052	0.0053	0.0449	0.0055	0.0054
	1/3	0.0911	0.0058	0.0061	0.0893	0.0056	0.0058
	HSV	0.2861	0.0216	0.0267	0.2917	0.0223	0.0272
	YIQ-Luma	0.3161	0.2753	0.0201	0.3162	0.2758	0.0200
	1/3-Luma	0.3336	0.2569	0.0196	0.3341	0.2582	0.0194
	HSV-Luma	0.5014	0.2857	0.0310	0.5067	0.2865	0.0314
#2	YIQ	0.0569	0.0058	0.0089	0.0573	0.0059	0.0090
	1/3	0.0773	0.0080	0.0130	0.0819	0.0084	0.0138
	HSV	0.2243	0.0217	0.0300	0.2263	0.0230	0.0320
	YIQ-Luma	0.3157	0.2430	0.0297	0.3170	0.2434	0.0297
	1/3-Luma	0.3277	0.2338	0.0301	0.3342	0.2348	0.0299
	HSV-Luma	0.4096	0.2586	0.0366	0.4321	0.2603	0.0367
#3	YIQ	0.0519	0.0139	0.0100	0.0515	0.0142	0.0102
	1/3	0.1046	0.0050	0.0082	0.1118	0.0052	0.0081
	HSV	0.5050	0.0959	0.0691	0.5227	0.0992	0.0716
	YIQ-Luma	0.3164	0.2578	0.0425	0.3171	0.2580	0.0422
	1/3-Luma	0.3433	0.2369	0.0450	0.3428	0.2371	0.0446
	HSV-Luma	0.5939	0.3207	0.0639	0.6024	0.3232	0.0647
#4	YIQ	0.0476	0.0073	0.0055	0.0472	0.0069	0.0052
	1/3	0.1281	0.0102	0.0155	0.1273	0.0105	0.0154
	HSV	0.4204	0.0456	0.0489	0.4115	0.0474	0.0490
	YIQ-Luma	0.3219	0.2482	0.0337	0.3232	0.2476	0.0336
	1/3-Luma	0.3531	0.2333	0.0380	0.3584	0.2329	0.0380
	HSV-Luma	0.6060	0.2804	0.0629	0.6019	0.2819	0.0632
#5	YIQ	0.0826	0.0123	0.0073	0.0864	0.0128	0.0076
	1/3	0.1852	0.0094	0.0118	0.1921	0.0092	0.0121
	HSV	0.5714	0.1190	0.0764	0.5883	0.1246	0.0799
	YIQ-Luma	0.3129	0.2782	0.0236	0.3131	0.2780	0.0237
	1/3-Luma	0.3274	0.2483	0.0233	0.3253	0.2481	0.0236
	HSV-Luma	0.6066	0.3680	0.0582	0.6192	0.3723	0.0601

Table 1. This table shows quantitative error for the synthetic images of five different real scenes (shown in Figs. 3 and 4) using camera sensitivity functions of two different cameras Canon 1Ds Mark III and Nikon D40 in [2]. An encoding gamma of 2.2 is applied to synthesize the sRGB images.

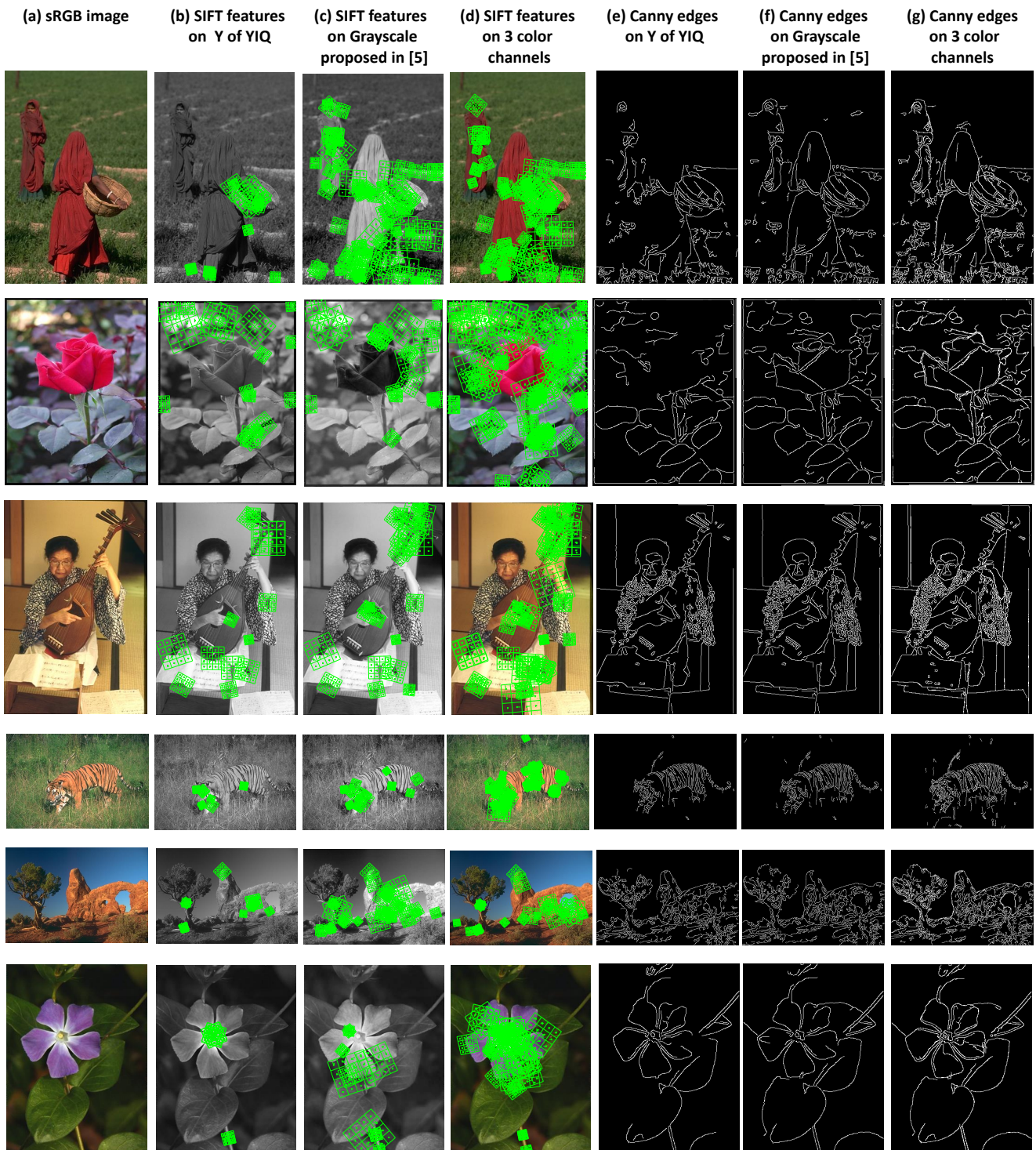| (a) sRGB image	(b) SIFT features on Y of YIQ	(c) SIFT features on Grayscale proposed in [5]	(d) SIFT features on 3 color channels	(e) Canny edges on Y of YIQ	(f) Canny edges on Grayscale proposed in [5]	(g) Canny edges on 3 color channels

Figure 5. This figure shows several additional examples of feature detection. (a) shows the sRGB input images. (b), (c) and (d) show the results of SIFT features using Y channel from YIQ, grayscale images obtained from [5] and three color channels, respectively; while (e), (f) and (g) show the corresponding Canny edges. All the sRGB images used here are in Lu et al.'s dataset [5].
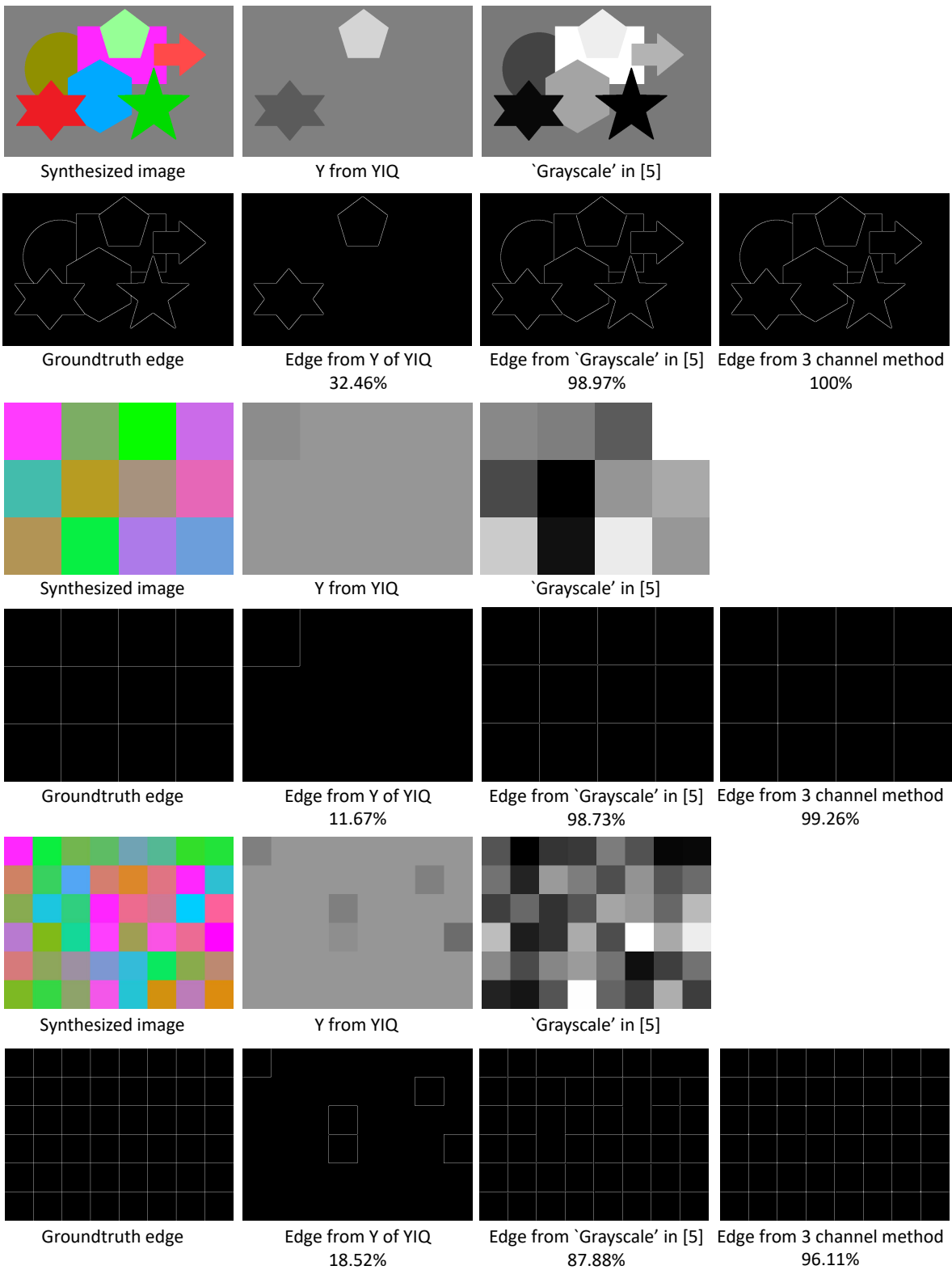
Figure 6. Ground truth example for edge detection. The first column is two synthetic images with known edges. The percentage of correctly labeled edges are shown. As can be seen, luminance channel is not always the best choice, there are better alternatives.