Illuminant Spectra-based Source Separation Using Flash Photography (Supplementary Material)

Zhuo Hui,[†] Kalyan Sunkavalli,[‡] Sunil Hadap,[‡] and Aswin C. Sankaranarayanan[†]
[†]Carnegie Mellon University [‡]Adobe Research

List of Figures

1. Estimating illumination coefficients

We describe in detail the techniques employed to detect the illumination spectra coefficients for varying number of light sources, N. As suggested in the paper, source separation is not interesting for a single light source, i.e., $N = 1$ and not feasible when $N > 4$. So, the discussion below is centered around the two interesting cases of $N = 2$ and 3.

Two light sources ($N = 2$). When there are two light sources in the scene, the set $\mathcal{G}$ lies on an arc on $\mathbb{S}^2$. We use RANSAC to robustly estimate the arc on $\mathbb{S}^2$ with maximum inliers. In particular, we are interested in the end points of this arc, which are associated with the illuminant coefficients; as noted in the paper, this estimate will correspond to the true coefficients if there were "pure pixels" in the no-flash photograph for each of the light sources.

To estimate the parameters of the arc, we first rotate all the pruned points on $\mathcal{G}$ by aligning the center of these points to the north pole. Next, we project the rotated points $\mathcal{G}_{rot}$ to the 2D plane by finding the intersecting on the tangent surface at the north pole (see Figure 1). Finally, we fit the line for all the projected points $\mathcal{G}_{proj}$ via RANSAC and find the start and ending points for the line with maximum inliers. Note that we need to invert the process of rotation and projection to determine the corresponding locations in the space of $\mathcal{G}$. The pseudo code for this is provided Algorithm 1.

Algorithm 1 Two lights coefficients estimation

Input: The set $\mathcal{G}$, the set $\mathcal{H}$ histogram of $\mathcal{G}$
Parameters: cutoff value T
Output: Estimated coefficients $\mathbf{g}_1$ and $\mathbf{g}_2$
Algorithm begin:
Step 1: Pruning the set $\mathcal{G}$
$\mathcal{G} \leftarrow \mathcal{G}(\mathcal{H} > T)$
Step 2: Rotating to north pole
$c_0 \leftarrow$ findCenter$(\mathcal{G})$
$\mathcal{R} \leftarrow$ rotate-to-northPole$(c_0/\|c_0\|)$
$\mathcal{G}_{rot} \leftarrow \mathcal{R} * \mathcal{G}$
Step 3: Projecting on the 2D surface
$\mathcal{G}_z \leftarrow z$ component of $\mathcal{G}_{rot}$
$\alpha \leftarrow 1./\mathcal{G}_z$
$\mathcal{G}_{proj} \leftarrow \alpha.\, \mathcal{G}_{rot}$
Step 4: Fitting a line for the projected points
params $\leftarrow$ fit-line-RANSAC$(\mathcal{G}_{proj})$
Step 5: Finding the ending points $\{\mathbf{p}_1, \mathbf{p}_2\}$
$[\mathbf{p}_1, \mathbf{p}_2] \leftarrow$ find-end-points$(\mathcal{G}_{proj},$ params$)$
Step 6: Converting vertices back to $\mathcal{G}$
$[\mathbf{g}_1,\ \mathbf{g}_2] \leftarrow \mathcal{R}^\top * [\mathbf{p}_1/\|\mathbf{p}_1\|,\ \mathbf{p}_2/\|\mathbf{p}_2\|]$

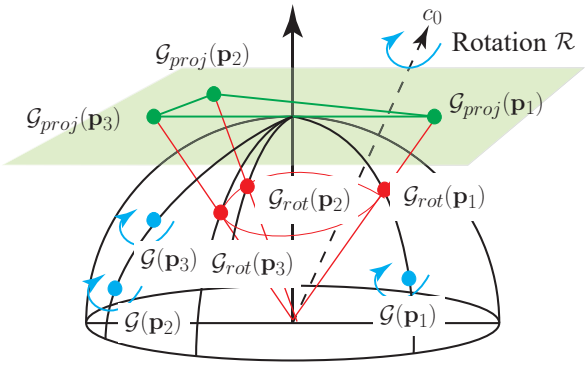


Figure 1. Schematic of rotation and projection of the points in the set $\mathcal{G}$ in estimating the illumination coefficients.

Three light sources ($N = 3$). Similar to two light scenario, we first process the points by rotating the set $\mathcal{G}$ and then projecting them on the tangent plant at north pole (see Figure 1). Next, we use the method in Parvu et al. [9] to determine the minimum triangle enclosing all the points and to find the associated vertices $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$. Finally, we convert the vertices to the corresponding points in the set the set $\mathcal{G}$.

Algorithm 2 Three lights coefficients estimation

Input: The set $\mathcal{G}$, the set $\mathcal{H}$ histogram of $\mathcal{G}$
Parameters: cutoff value T
Output: Estimated coefficients $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$
Algorithm begin:
Step 1: Pruning the set $\mathcal{G}$
$\mathcal{G} \leftarrow \mathcal{G}(\mathcal{H} > T)$
Step 2: Rotating to north pole
$c_0 \leftarrow \text{mean}(\mathcal{G})$
$\mathcal{R} \leftarrow \text{rotate-to-northPole}(c_0/\|c_0\|)$
$\mathcal{G}_{rot} \leftarrow \mathcal{R} * \mathcal{G}$
Step 3: Projecting on the 2D surface
$\mathcal{G}_z \leftarrow z$ component of $\mathcal{G}_{rot}$
$\alpha \leftarrow 1./\mathcal{G}_z$
$\mathcal{G}_{proj} \leftarrow \alpha.\, \mathcal{G}_{rot}$
Step 4: Fitting the minimum triangle
$[\mathbf{p}_1,\ \mathbf{p}_2,\ \mathbf{p}_3] \leftarrow \text{fit-minimum-triangle}(\mathcal{G}_{proj})$
Step 5: Converting vertices back to $\mathcal{G}$
$[\mathbf{g}_1,\ \mathbf{g}_2,\ \mathbf{g}_3] \leftarrow \mathcal{R}^\top * [\mathbf{p}_1/\|\mathbf{p}_1\|,\ \mathbf{p}_2/\|\mathbf{p}_2\|,\ \mathbf{p}_3/\|\mathbf{p}_3\|]$

2. Uniqueness of solution

In this section, we explain why source separation is not feasible for more than three light sources in the scene.

To recall, the problem of interest here is to estimate the relative shading terms $\{z_i(\mathbf{p}), i = 1, \ldots, N\}$ at a pixel $\mathbf{p}$ given the reflectance invariant $\Gamma(\mathbf{p}) \in \mathbb{R}^3$ and the lighting coefficients $\{\widehat{\mathbf{b}}_i, i = 1, \ldots, N\}$. This can be obtained by solving the following system:

$$\Gamma(\mathbf{p}) = \sum_{i=1}^{N} z_i(\mathbf{p})\widehat{\mathbf{b}}_i \quad \text{s.t} \quad \forall i, z_i(\mathbf{p}) \geq 0.$$

The expression above is simply the hull constraint and suggests that $\Gamma(\mathbf{p})$ lies in the cone defined by the lighting coefficients. Note that the non-negativity of the relative shading implies there is no feasible solution if $\Gamma(\mathbf{p})$ lies outside the cone of the lighting coefficients. Finally, observe that since $\Gamma(\mathbf{p}) \in \mathbb{R}^3$, there are three constraints on the unknowns; however, the number of unknowns is equal to N, the number of light sources. Clearly, the system is under-determined if $N > 3$ and as a consequence there are multiple solutions and we have no way to identify the correct relative shading among these. This implies that source separation is not feasible — barring the introduction of additional constraints — for scenes with more than three light sources.

3. Light source separation

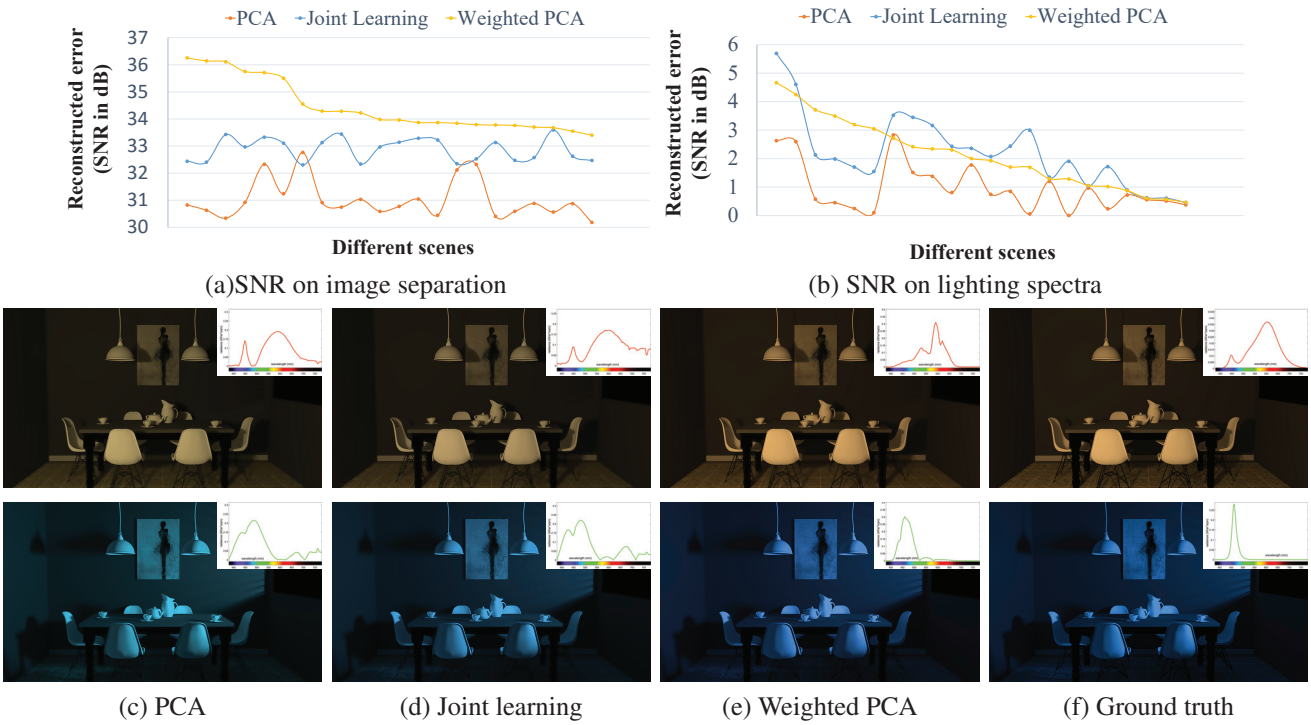
In this section, we describe in detail our evaluation of the lighting separation technique. We start with an evaluation of various choices in the reflectance and illumination bases and how they affect the performance of estimated illumination spectra as well as source separation. Subsequently, we provide synthetic evaluation of source separation as well as additional real results.

3.1. Selection of reflectance and illumination bases

We used the measured database for reflectance [5] and illumination [1] to learn two three-dimensional subspaces, one each for reflectance and illumination. There are many suitable approaches for learning such bases.

- *PCA*. PCA learns the best 3-dimensional subspaces on the given input spectra. This is the classical choice that maximizes representation of the spectra using a 3-dimensional subspaces. However, it does not take camera responses into account.

- *Weighted PCA*. To account for camera spectral response, we use a weighted PCA approach. We define the collection of reflectance and lighting spectra as $\mathcal{R}$ and $\mathcal{L}$, respectively, with one each column denoting a measured spectra in the database. We perform SVD on the term defined as $\mathcal{M}^k_{\mathcal{R}} = \mathcal{R}S^k$, where S^k is a diagonal matrix with the spectra of the camera response in the k-channel. Given the rank 1 approximation for each color channel in $\mathcal{M}^k_{\mathcal{R}}$, we concatenate them with column-wise as a matrix and compute the orthogonal bases. Similarly, we return the bases for the lighting spectra for the measurements defined as $\mathcal{M}^k_{\mathcal{L}} = \mathcal{L}S^k$.



(a)SNR on image separation (b) SNR on lighting spectra

(c) PCA (d) Joint learning (e) Weighted PCA (f) Ground truth

Figure 2. Evaluation of the choice in the reflectance and illumination bases. For the rendered 22 images with two illuminants, we plot the approximation accuracy along with the descending order of weighted PCA in showing the results in (a) and (b). We also visualize comparison of PCA, joint learning method and weighted PCA for one selected scene against ground truth.

- *Joint learning of subspaces.* A third alternative to consider is to jointly learn illumination and reflectance subspaces, under the camera spectral response [4]. Intuitively, the idea behind this method is that we want subspaces that allow for best rendition of image intensities. Given a reflectance spectra $\mathbf{r}$ and an illumination spectra $\mathbf{l}$ — both column vectors — the image intensity under a camera response S^k is given as

$$\mathbf{r}^\top S^k \mathbf{l}.$$

We can now learn bases B_R and B_L such that the image intensity value are preserved.

In Figure 2, we characterize the performance on the synthetic database for different basis choices. We showcase the approximation accuracy for 22 rendered images with two different illuminants in Signal-to-Noise Ratio (SNR) on both the estimation of separated images as well as the lighting spectra. As can be seen here, the weighted PCA method returns best performance on the separated images while produces results close to the joint estimation, which is one of the most accurate techniques for optimal basis on color constancy [4]. We also evaluate the choice of basis on the real world scenes with ground truth in Figure 3. The performance of the different basis on this dataset closely parallels the results we observed in the synthetic dataset. Given these observations, we use the weighted PCA basis for all results in the paper.

3.2. Evaluation on two-light scenario

Synthetic experiment. We evaluate the source separation technique on realistically-rendered scenes using the MIT-SUBA rendering engine [8]. Specifically, when we simulate the scene, we select two light spectra from [1] and compute the errors for both separated images, as well as the light coefficients, against the ground truth. For the scene with two lights, we report these errors as a function of varying angular difference for the ground truth spectra in Figure 4. We observe that the SNR values of the source separation are larger than 30dB for most of the lighting spectra, even for the worst case, i.e. (1° in the angular difference), the SNR value can still be achieved at 16dB, suggesting the robustness of our technique. We also showcase the angular error against the ground truth coefficients. Note that the angular error increases with the difference between the lighting spectra. This is due to the fact that there is a decrease of the conic hull characterized by Γ as we moved one lighting spectra away from the other. In particular, the potential region characterized the estimation errors also shrinks with the increase similarity in lighting spectra.

(a) No-flash/flash images (b) PCA (c) Joint learning (d) Weighted PCA (e) Ground truth

Figure 3. Evaluation of different bases choice for the source separation on the real-world scenes. We compare the results for different bases choices against ground truth capture. We also show the close-up appearance for the separated images. Note that all the methods produce similar shading estimates in the resulting images. However, the method of PCA produces inaccuracy illuminant color estimates for the separated images, while joint learning method induces the artifacts in the scene appearance (see the close-up in (c)). In contrast, weighted PCA is able to return the faithful results that are most close to the ground truth images.

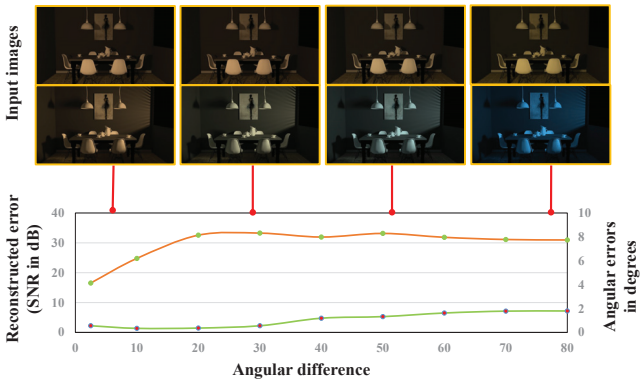


Figure 4. Evaluation of the two-light source separation on synthetic dataset. For each scene (source credit to [2]), we pick two measured illuminant spectra from LSPDD database. We plot the errors measured against the ground truth constituent images as a function of angular difference between ground truth lighting coefficients.

Real world results. We have also tested our algorithm on real-world scenes with two illuminants. In Figures 5 and 6, we demonstrate our technique on the scene with two lights sources and compare with ground truth captures. We compare against a simple non-negative matrix factorization (NNMF) as well as the technique proposed in Hsu et al. [6]. Note that the scene in Figures 5 exhibits complex scene geometries as well as intricate interactions of objects and light rays. As can be seen, naively applying NNMF to the no-flash image leads to the loss of the colors. Hsu et al. [6] focused on estimating relative contribution of the light source from no-flash image at the expense of introducing restrictive assumptions on the scene as well as the colors of the illuminants. This naturally leads to the visual artifacts in the estimation results by violating these assumptions. In contrast, our method showcases the robustness and effectiveness by returning the results that closely resemble to the ground truth.

We also include the additional results on lighting separation in Figures 7 and 8. Our technique is able to attribute both the color and the shading variations to the corresponding manner of the lighting sources.

3.3. Evaluation on three-light scenario

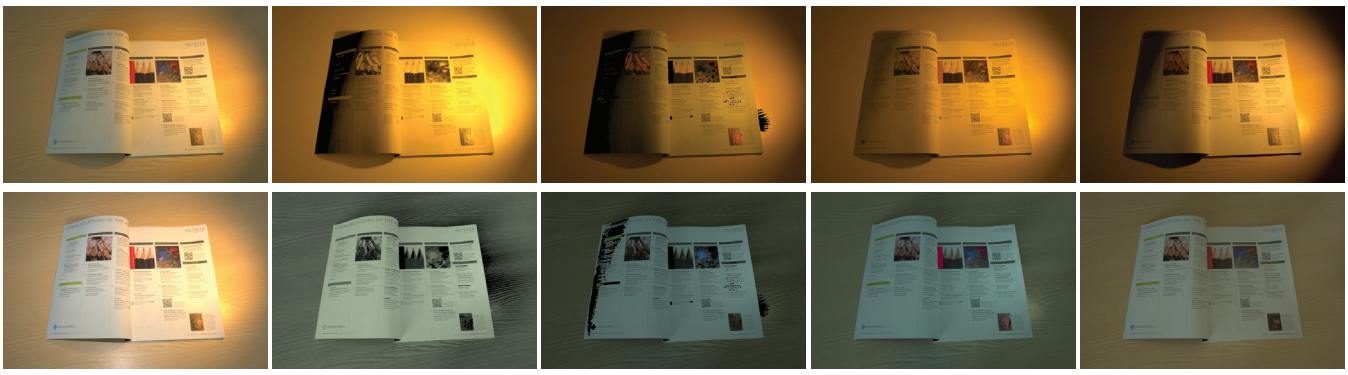
We now discuss evaluation of three-light sources scenario using synthetic and real data.

Synthetic experiment. To evaluate the performance on the scene with three lights, we characterize the performance on the synthetic dataset by using the MITSUBA rendering engine [8]. Specifically, we select three measured lighting spectra from the dataset [1] and make sure the smallest angular difference of the illumination coefficients for these selected lighting spectra is larger than 20° in degrees. In Figure 9, we report the errors against the ground truth for 26 rendered flash/no-flash image pairs. As can be seen, our algorithm is able to return the results larger than 20 dB for all the datasets. We also include the visual results for two selected samples in Figure 9. As can be seen, our algorithm is able to capture the color and shadings for each illuminant as well as produce the results very close to the ground truth.

Real world results. In Figure 10, we characterize the performance of our technique on the real-world scenes against the ground truth captures. Our lighting separation scheme produces visually pleasing results with shadows and shadings that are consistent with those observed in the ground truth. In Figure 11, we showcase our separation results for

(a) Input images | (b) Matrix factorization SNR 5.96 dB | (c) Hsu et al. [6] SNR 3.13 dB | (d) Our results SNR 16.78 dB | (e) Ground truth

Figure 5. We separate a no-flash image (a) into two components and compare with non-negative matrix-factorization (b) and and Hsu et al. [6] (c). Compared to the ground truth images (e), we can see that matrix factorization produces noisy colors loss (see the top row (b)), while Hsu et al. [6] produce an incorrect estimate of light color and shading (c). Our result (d) closely mimics the actual captured results.

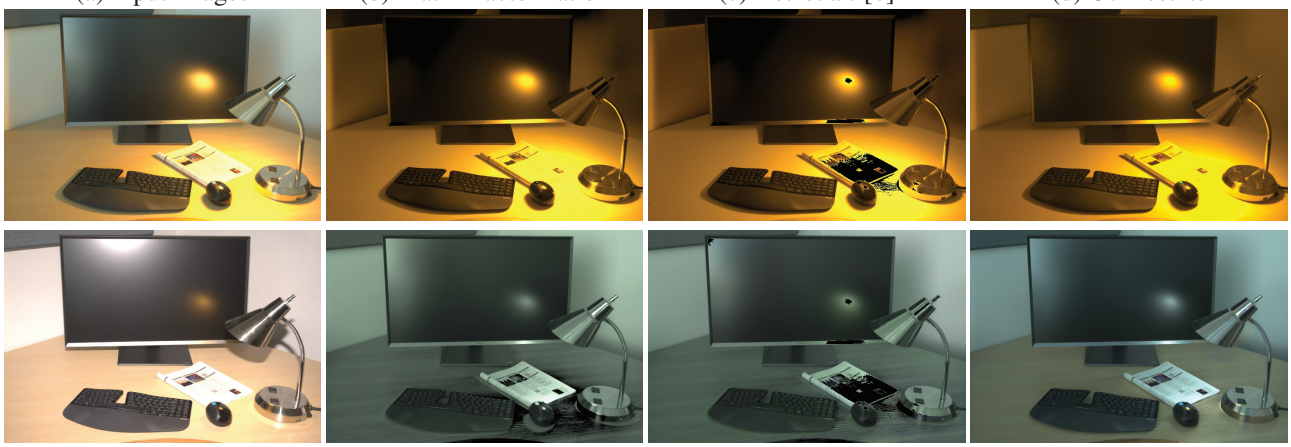


(a) Input images | (b) Matrix factorization SNR 7.68 dB | (c) Hsu et al. [6] SNR 4.20 dB | (d) Our results SNR 18.15 dB | (e) Ground truth

Figure 6. We compare our results against ground truth on the real-world scenes. As can be seen here, naive applying that matrix factorization leads to the loss of the colors (b). The method of Hsu et al. [6] produces artifacts due to the restrictive assumptions on the scene (c). In contrast, our technique produces results that are very close to the ground truth, and significantly better than all other techniques.

the scene under complex light transport. The scene is illuminated by the outdoor illuminant but the curtain on the window introduces a second light source that is tinged with red. In addition to these two sources, the ceiling light is also turned on in the room. While the scene features complex shadows (under the chair, around the small table, near

(a) Input images (b) Matrix factorization (c) Hsu et al. [6] (d) Our results



(a) Input images (b) Matrix factorization (c) Hsu et al. [6] (d) Our results

(a) Input images (b) Matrix factorization (c) Hsu et al. [6] (d) Our results

Figure 7. We include more results on real-world scenes with two illuminants. As can be seen here, the performance of our technique on these datasets closely parallels the results we observed in previous synthetic and real-world dataset.

the red curtain, to name a few) caused by different type of illuminants, our lighting separation scheme is able to group these regions into the associated separated photographs and produce visually pleasing results.

4. White balancing

One of the applications enabled by our technique is white balancing under mixed illumination. The vast majority of white-balance algorithms assume that the scene is lit by a single dominant light source. In contrast, we are able to

(a) Input images (b) Matrix factorization (c) Hsu et al. [6] (d) Our results

Figure 8. Additional results on real-world scenes with two illuminants. Note that the colors of the illuminants in the scene are very close to the colors of the reflectance (see the chair and couch), making both the matrix factorization and the method of Hsu et al. [6] failed. In contrast, our method relies on the Hull constraint, which is reflectance invariant, and is still able to isolate the reflectance and produce visually appealing results.

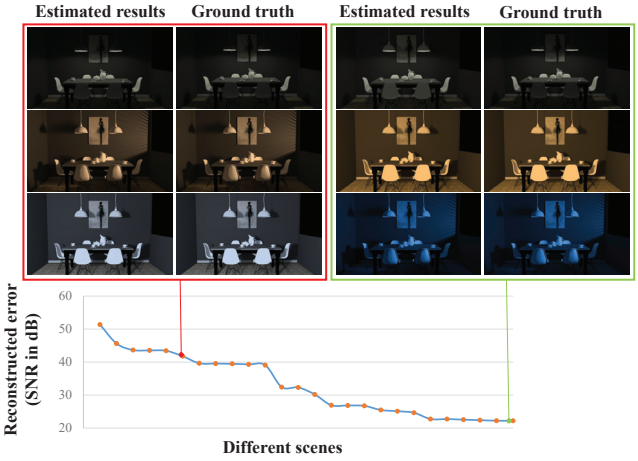


Figure 9. Evaluation of the three-light source separation. For each generated scene, we illuminated it by three lighting spectra picked from LSPDD database. We plot the measured errors against the ground truth constituent images for 26 rendered scenes.

estimate and remove the effect of spatially-varying lighting by using the Hull constraint. There are two approaches to achieve this.

- **Approach I.** We can simply adjust the illumination coefficients in each separated images. In particular, we substitute the estimated coefficients $\widehat{\mathbf{b}}_j$ in (3) with the coefficients corresponding to the neutral light spectra. However, this approach requires us to estimate the light source coefficients and their relative shading, which can only be performed for 3 or fewer light sources.

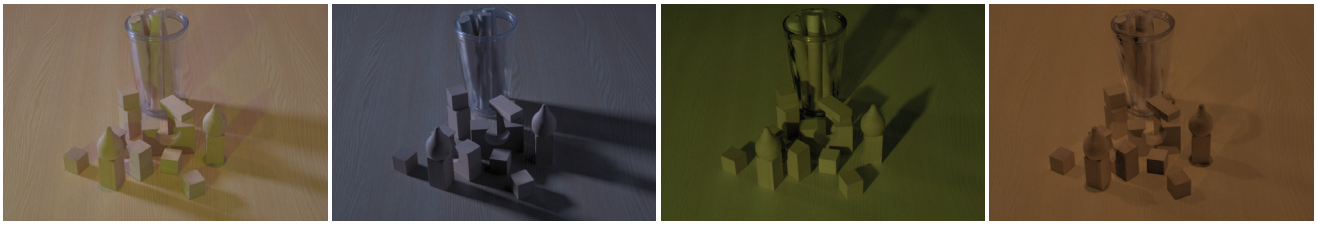
- **Approach II.** We provide an alternative solution that provides the ability to handle any number of light sources in the scene, albeit under some assumptions on their colors. Specifically, we assume that $\|\sum_{i=1}^{N}\eta_i\mathbf{b}_i\|^2 \approx \|\sum_{i=1}^{N}\eta_i\|^2$. That is, $\sum_{i\neq j}\eta_i\eta_j\mathbf{b}_i\mathbf{b}_j \approx \sum_{i\neq j}\eta_i\eta_j$, or equivalently, $\mathbf{b}_i\mathbf{b}_j \approx 1$. In essence, we have constrained the lighting spectra close to each other in the low-dimensional model.

  Now, recall that the no-flash intensity is

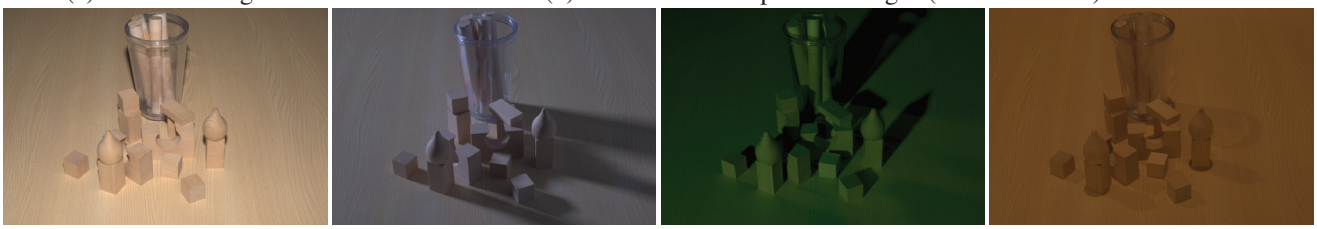
$$I^k_{\text{nf}}(\mathbf{p}) = \|\mathbf{a}_\mathbf{p}\| \left(\frac{\alpha^T_\mathbf{p}}{\|\alpha_\mathbf{p}\|}\right) E^k \sum_{i=1}^{N}\eta_i(\mathbf{p})\mathbf{b}_i, \qquad (1)$$

  the white balancing results at pixel $\mathbf{p}$ can be expressed as

$$I^k_{\text{wb}}(\mathbf{p}) = \|\mathbf{a}_\mathbf{p}\| \left(\frac{\alpha^T_\mathbf{p}}{\|\alpha_\mathbf{p}\|}\right) E^k\, \mathbf{b}_{\text{wb}} \sum_{i=1}^{N}\eta_i(\mathbf{p}), \qquad (2)$$



(a) No-flash image (b) Our estimated separated images (SNR: 15.13dB)



(c) Flash image (d) Captured photograph

(a) No-flash image (b) Our estimated separated images (SNR: 12.46dB)

(c) Flash image (d) Captured photograph

(a) No-flash image (b) Our estimated separated images (SNR: 11.21dB)

(c) Flash image (d) Captured photograph

Figure 10. We evaluate our technique on scenes with mixtures of three lights and compare with the ground truth image. Our technique is able to capture both the color and the shading for each of these sources and produce results similar to the ground truth.

(a) No-flash image (b) Flash image (c) Our estimated separated images

Figure 11. We evaluate our technique on scenes with mixtures of three lights. In the top row, we capture an image under warm indoor LED lights and outdoor lighting that percolates through a red curtain. The outdoor light that is transmitted through the curtain gets colored red, while some of it bounces off the curtains white backing and diffuses into the scene as blue light. Our technique is able to estimate separated results that capture this complex light transport.

where $\mathbf{b}_{\text{wb}}$ is the neutral light coefficients.

Given $\| \sum_{i=1}^{N} \eta_i \mathbf{b}_i \|^2 \approx \| \sum_{i=1}^{N} \eta_i \|^2$, we can substitute $\|\beta(\mathbf{p})\|$ to express (2) as

$$I^k_{\text{wb}}(\mathbf{p}) = \left(\frac{\alpha^T_{\mathbf{p}}}{\|\alpha_{\mathbf{p}}\|} \right) E^k \|\beta(\mathbf{p})\| \mathbf{b}_{\text{wb}}$$

Our results are shown in Figures 12. We compare our results with those from two algorithms that are designed to handle spatially-varying mixed illumination — Hsu et al. [6] and Hui et al. [7]. Hsu et al. require that the color of the illuminants and assume that only two light sources present in the scene. While we manually specified this as input to their technique, their result is not able to deal with extreme illumination (see Fig. 12). Similar to our work, Hui et al. use a flash camera and can generalize to an arbitrary number of scene illuminants. While Hui et al. [7] produces the results of similar quality, their underlying image formation model is not physically accurate; in particular, their method completely ignores the image formation model in (1) of the main paper and instead uses an empirical model that does not account for camera spectral response.

5. Camera editing

Another unique capability of the proposed method is its ability to edit camera spectral response. We can achieve this as follows. Given the estimate of the separated image as

$$\widehat{I}^k_{sep,j}(\mathbf{p}) = \|\beta\| \alpha^{\top}_{\mathbf{p}} E^k \widehat{z}_j(\mathbf{p}) \widehat{\mathbf{b}}_j. \tag{3}$$

and

$$E^k(i,j) = \int_{\lambda} \widetilde{\rho}_i(\lambda) S^k(\lambda) \widetilde{\ell}_j(\lambda) d\lambda,$$

we are able to change the camera spectral response $S^k(\lambda)$ with a novel spectral distribution function $\widehat{S}^k(\lambda)$. Specifically, we change the captured no-flash image with novel camera response function and show the results in Figure 13.

6. Additional results on photometric stereo

In Figure 14, we showcase more results on the flash/no-flash photometric stereo. For each object, we compute average angular error (compared to the ground truth) as indicated below the results. We evaluate our method against single shot photometric stereo method [3]. While our method relies on extra captured flash image, it is always one shot effort for most of the consumer cameras but producing the results better in orders of magnitude.

7. Limitations

Inter-reflections In Figure 15, we showcase the lighting separation results for two-light on the synthetic scene with different number of light bounces. To do that, we use the MITSUBA rendering engine [8] by changing the longest path depth in the generated output image, which characterizes the number of bounces for the illumination in the scene. As shown in Figure 15, the inter-reflection effects can be seen as indicated by the orange boxes. As we increase the number of light bounces, the shadows on the side of the boxes become brighter due to the scattering interactions with the walls. We showcase the approximation accuracy for five depths (i.e. $2, 10, 20, 50, 100, \infty$) with two different illuminants in Signal-to-Noise Ratio (SNR) on the estimation of separated images against the ground truth. We also include the and relative estimation errors in the bottom row. As can be seen here, the performance degrades with the increased number of light bounces, since each interreflection can be treated as a new light source potentially

of a new color since the albedo of an object multiplies with spectrum. To this end, interrelections produce little of diffuse shadows, which are hard to latch on via our conic hull methods.

Shinny objects In Figure 16, We showcase the performance of our source separation technique on synthetic objects when imaged for the materials with different roughness. We use the MITSUBA rendering engine [8] to simulate the materials and adjust the parameters to change the roughness of the materials (as visualized in the bottom of the plot). In particular, we render the object under two illuminants and gradually increase the roughness parameter defined from .1 to 1. We showcase the approximation accuracy in SNR on the estimation of separated images against the ground truth. As can be seen here, the performance of our algorithm varies as the roughness; this is because the violation of our Lambertian assumption makes about the scene. However, the incorrect results will be localized to the objects since the processing is largely per-pixel and the conic hull processing is inherently robust to outliers via theuse of RANSAC and other pre-processing techniques, as indicated by the visualized errors in the bottom row. Hence, our technique is still able to return visually pleasing results even for very shinny objects.

References

[1] Light spectral power distribution database. *URL: lspdd.com/g*. 2, 3, 4

[2] B. Bitterli. Rendering resources, 2016. https://benedikt-bitterli.me/resources/. 4

[3] A. Chakrabarti and K. Sunkavalli. Single-image rgb photometric stereo with spatially-varying albedo. In *IEEE International Conference on 3D Vision*, 2016. 9, 13

[4] H. Y. Chong, S. J. Gortler, and T. Zickler. The von kries hypothesis and a basis for color constancy. In *ICCV*, 2007. 3

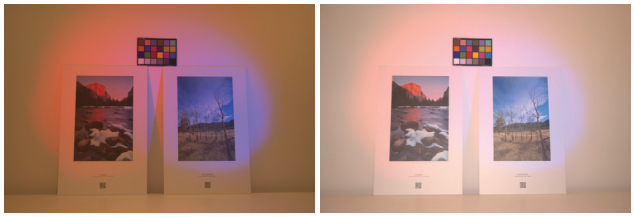
[5] M. Hauta-Kasari, K. Miyazawa, S. Toyooka, and J. Parkkinen. Spectral vision system for measuring color images. *JOSA A*, 16(10):2352–2362, 1999. 2

[6] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand. Light mixture estimation for spatially varying white balance. In *TOG*, volume 27, page 70, 2008. 4, 5, 6, 7, 9, 11

[7] Z. Hui, A. C. Sankaranarayanan, K. Sunkavalli, and S. Hadap. White balance under mixed illumination using flash photography. In *ICCP*, 2016. 9, 11

[8] W. Jakob. Mitsuba renderer, 2010. *URL: http://www. mitsuba-renderer. org*, 3:10, 2015. 3, 4, 9, 10

[9] O. Pârvu and D. Gilbert. Implementation of linear minimum area enclosing triangle algorithm. *Computational and Applied Mathematics*, pages 1–16, 2014. 2



(a) No-flash image (b) Flash image

(c) Hsu et al. [6] (Mean error = 3.58°) (d) Hui et al. [7] (Mean error = 0.88°) (e) Our result (I) (Mean error = 0.85°) (f) Our result (II) (Mean error = 0.92°)

(a) No-flash image (b) Flash image

(c) Hsu et al. [6] (Mean error = 1.76°) (d) Hui et al. [7] (Mean error = 1.07°) (e) Our result (I) (Mean error = 0.87°) (f) Our result (II) (Mean error = 0.90°)

(a) No-flash image (b) Flash image

(c) Hsu et al. [6] (Mean error = 1.12°) (d) Hui et al. [7] (Mean error = 0.12°) (e) Our result (I) (Mean error = 0.21°) (f) Our result (II) (Mean error = 0.36°)

Figure 12. We evaluate our white balance method on a no-flash/flash pair (a/b) from Hui et al. [7]. We show the white balance results and the associated kernels (insets). (c) Hsu et al. [6] require the light colors to be manually specified but fail on the extreme illumination in this scene. (d) Hui et al. use a flash image to improve results but rely on inaccurate physical model. (e) Our method (both approach I (e) and approach II (f)) produces the result which can achieve the same performance in terms of both visual quality and angular error. Note that approach (II) is able to produces visual appealing results as well as similar angular error measurements, making our method applicable to arbitrary number of light sources in the scene.

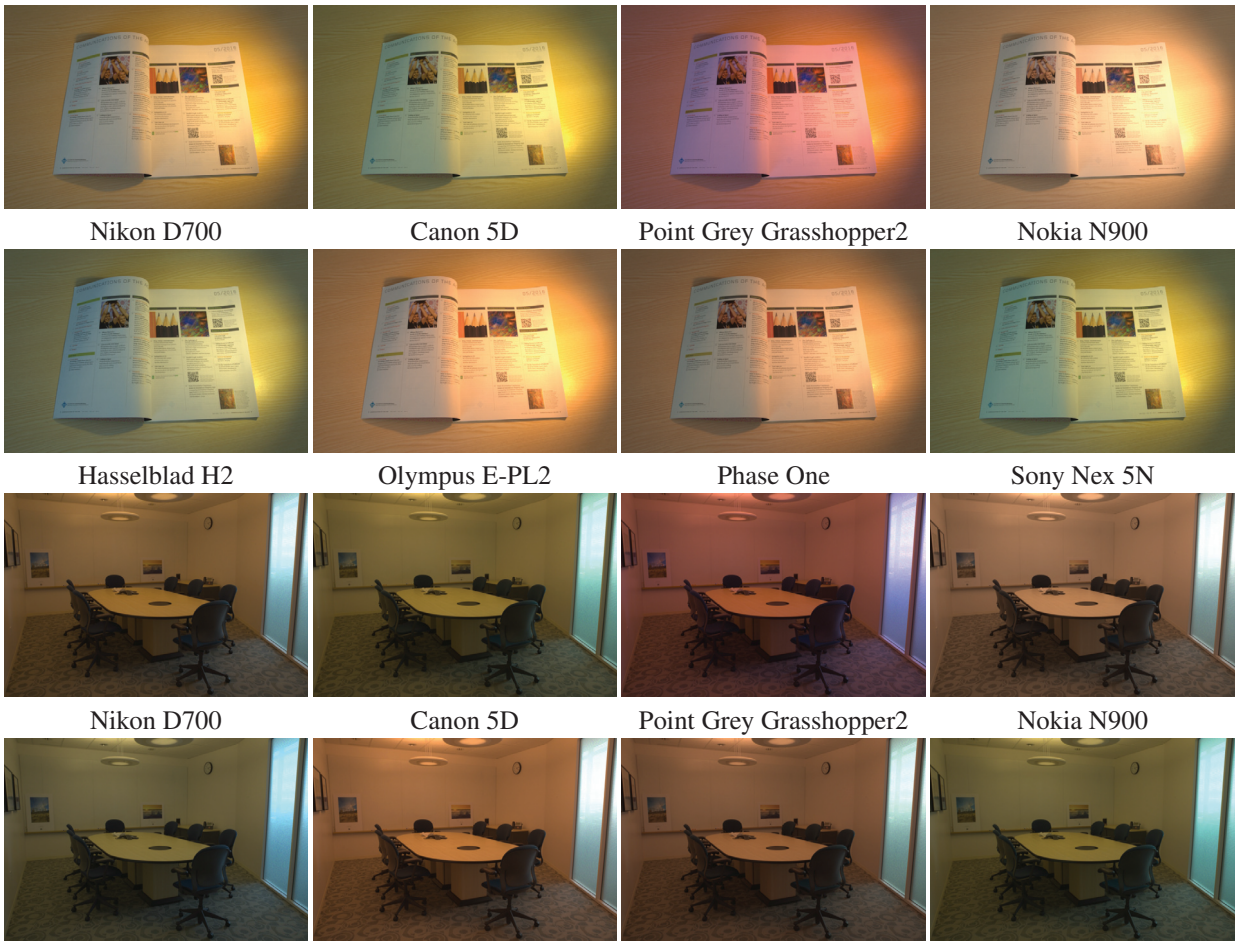
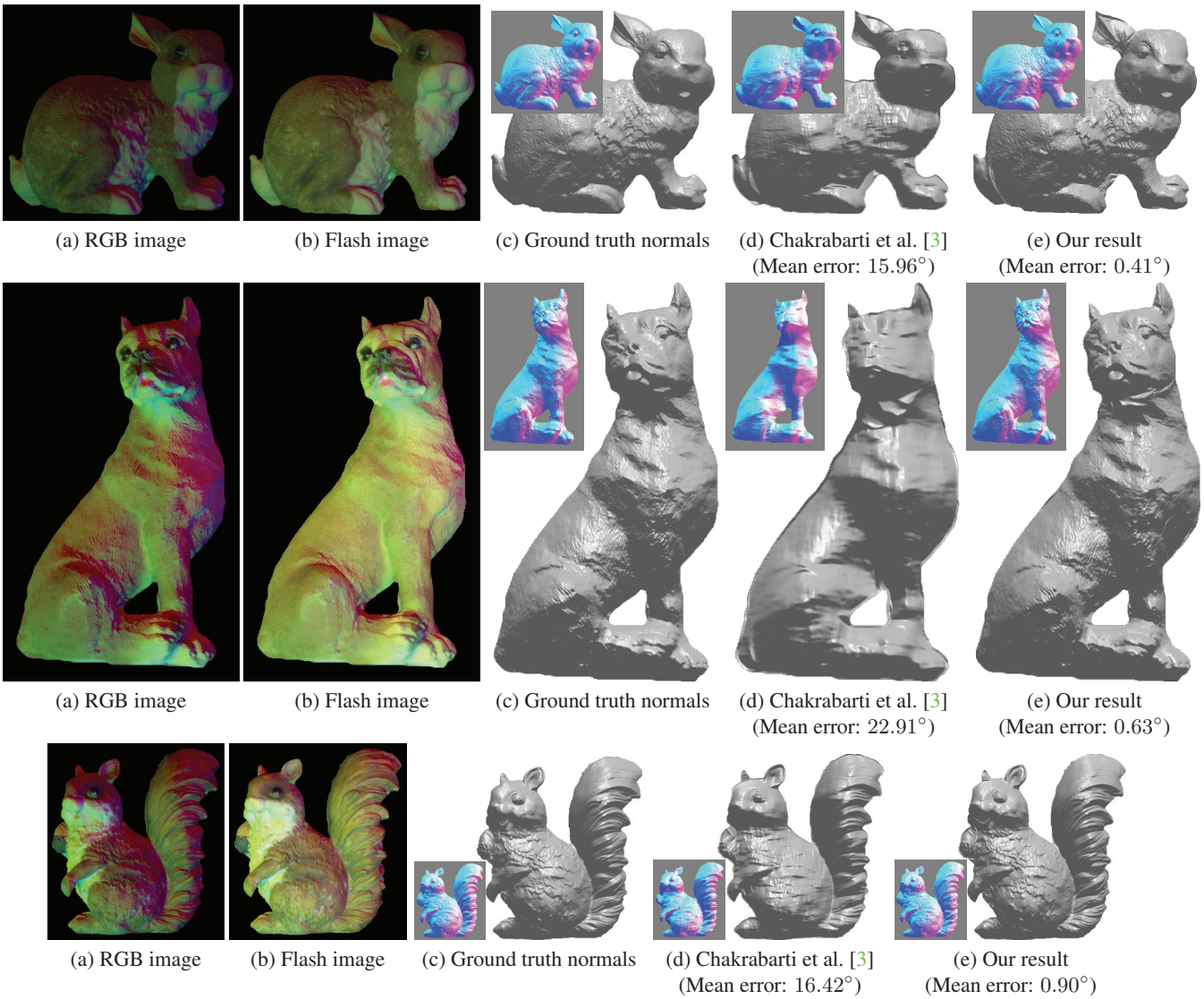


Figure 13. Results on camera response editing. We show estimated rendering results for different camera models.



(a) RGB image (b) Flash image (c) Ground truth normals (d) Chakrabarti et al. [3] (Mean error: 15.96°) (e) Our result (Mean error: 0.41°)

(a) RGB image (b) Flash image (c) Ground truth normals (d) Chakrabarti et al. [3] (Mean error: 22.91°) (e) Our result (Mean error: 0.63°)

(a) RGB image (b) Flash image (c) Ground truth normals (d) Chakrabarti et al. [3] (Mean error: 16.42°) (e) Our result (Mean error: 0.90°)

Figure 14. Results on two-shot captured photometric stereo of real objects. We show both estimated normal map and recovered 3D surfaces for our technique as well as those of single-shot method of Chakrabarti et al. [3]. We also include the mean of the angular errors for the estimated surface normals. Our technique produces surface normals with very low angular errors.

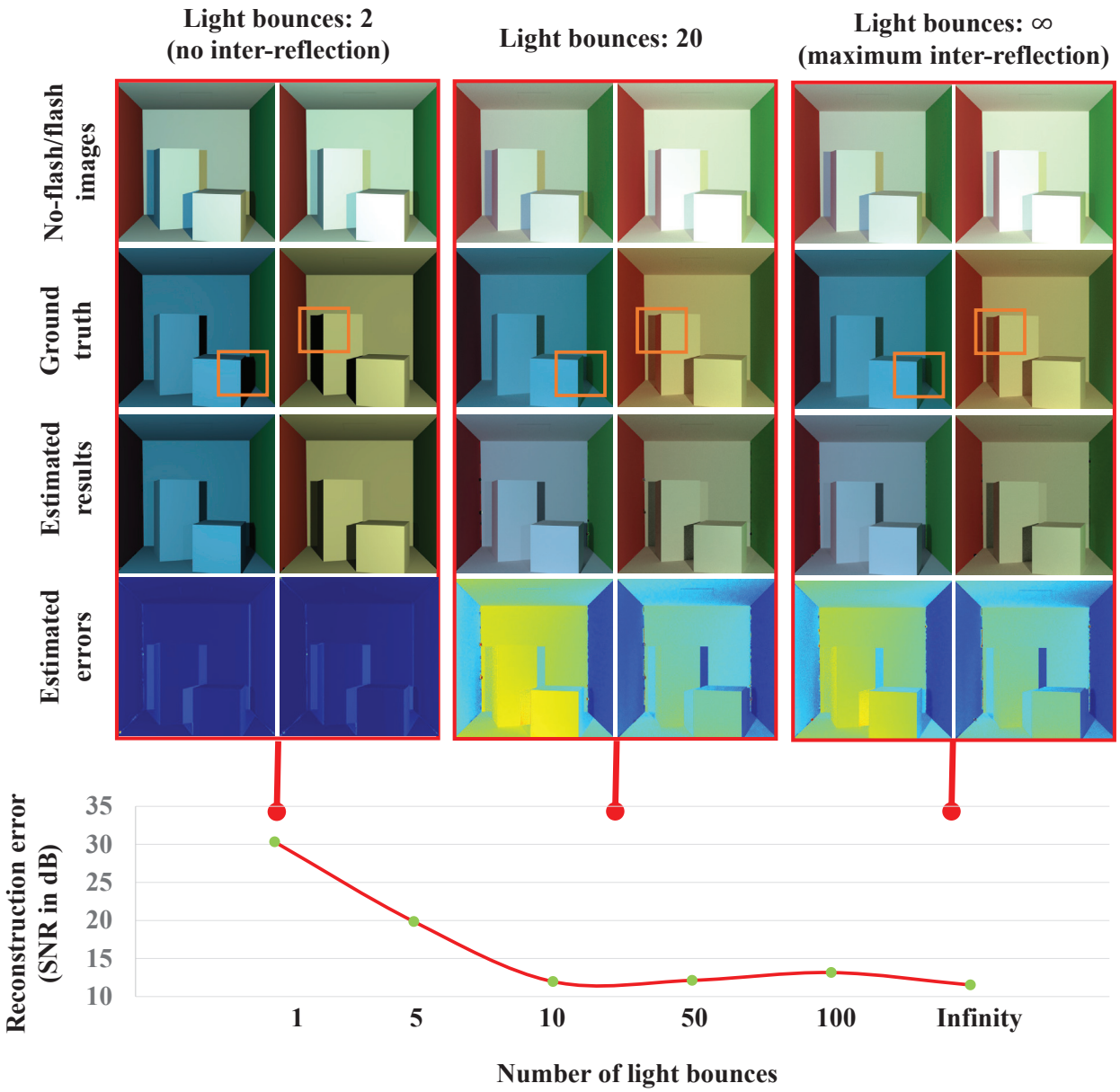


Figure 15. Evaluation of the inter-reflection effects on our two-light source separation technique. For the rendered images with two illuminants, we gradually increase the number of light paths to simulate different inter-reflection effects as can be seen from the orange boxes. We plot the accuracy for different simulated image pairs and compare the estimated source separation results with the ground truth.

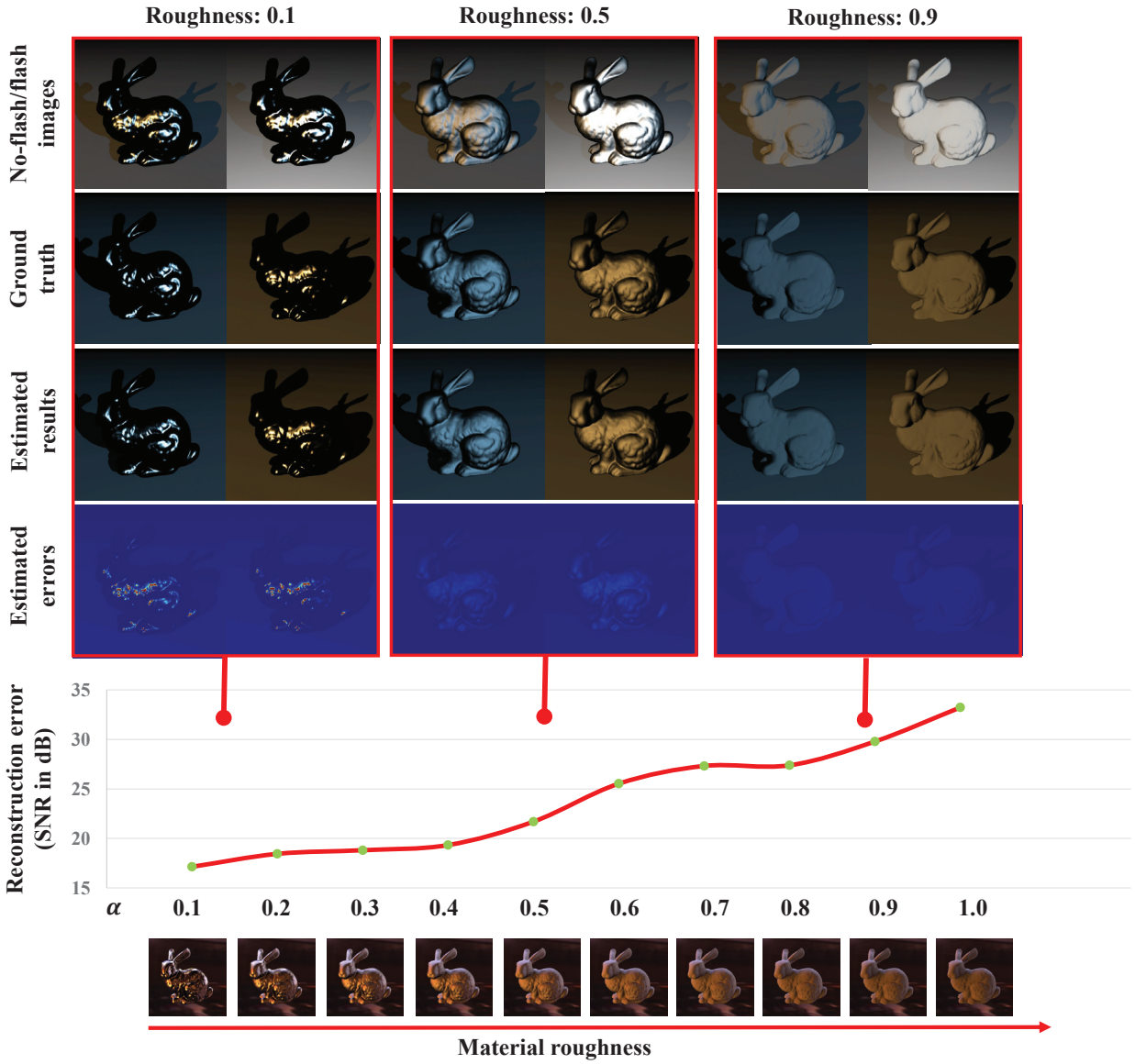


Figure 16. Evaluation of the effects of non-Lambertian material roughness on our two-light source separation technique. For the rendered images with two illuminants, we gradually increase the roughness of the materials (visualized in the bottom of the plot) to simulate different shinny objects effects. We plot the accuracy for different simulated image pairs and compare the estimated source separation results with the ground truth.