OBJECTIVE, ABSOLUTE AND HUE-AWARE METRICS FOR INTRINSIC IMAGE DECOMPOSITION ON REAL-WORLD SCENES: A PROOF OF CONCEPT

Supplementary materials

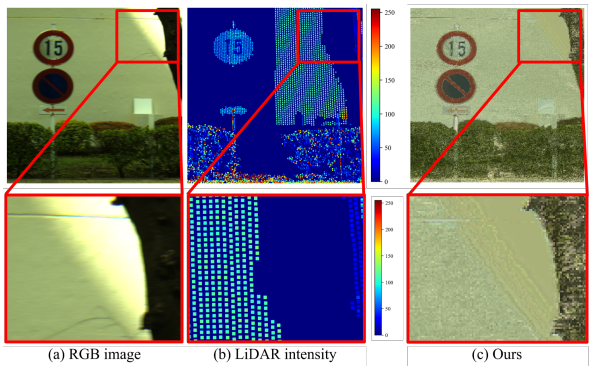

Fig. 8. Application to outdoor scene. (a) RGB image, (b) LiDAR intensity, and (c) our calculated albedo. Albedo was calculated in the same way as in the indoor experiment.

A. APPLICATION TO OUTDOOR SCENE

To demonstrate the effectiveness of our method in real-world applications, we experimented with an outdoor scene. As shown in Fig. 8, we captured hyperspectral images and LiDAR data of a real-world scene. The hyperspectral camera and LiDAR were set up to capture data simultaneously from the same viewpoint. The collected data was then processed to calculate the albedo values with our densification. The RGB image contains a whiteboard (SphereOptics GmbH), a black tree trunk, a yellowish wall, and a signboard. Although the RGB image includes sunlit and shadowed areas, the calculated albedo had no shadows, capturing subtle colors in the scene. This experiment suggests the possibility of our method working outdoors as well.

B. ABLATION STUDY FOR ALBEDO DENSIFICATION

B.1. Number of selected similar pixels

To prevent quantisation errors due to the small number of elements in the dictionary, the three most similar points are selected in densification process. When the selected number is one, the quantisation error may degrade the image quality. Conversely, when the number is large, image quality may also be degraded by the increased probability of including values

model	CIE76 (↓)	CIEDE2000 (↓)	Co-effs(↑)
$n = 1$	13.5±3.28	6.80±1.49	0.979
$n = 3$	13.4±3.33	6.73±1.53	0.978
$n = 5$	13.4±3.32	6.74±1.52	0.978
$n = 7$	13.5±3.34	6.77±1.52	0.977
$n = 10$	13.8±3.42	6.95±1.52	0.975
$n = 30$	22.1±4.18	12.7±1.81	0.936
$n = 50$	25.8±4.56	15.5±2.07	0.864
$n = 70$	27.3±5.09	16.6±2.35	0.841

Table 2. Quantitative evaluation results for different number of selected pixels n.

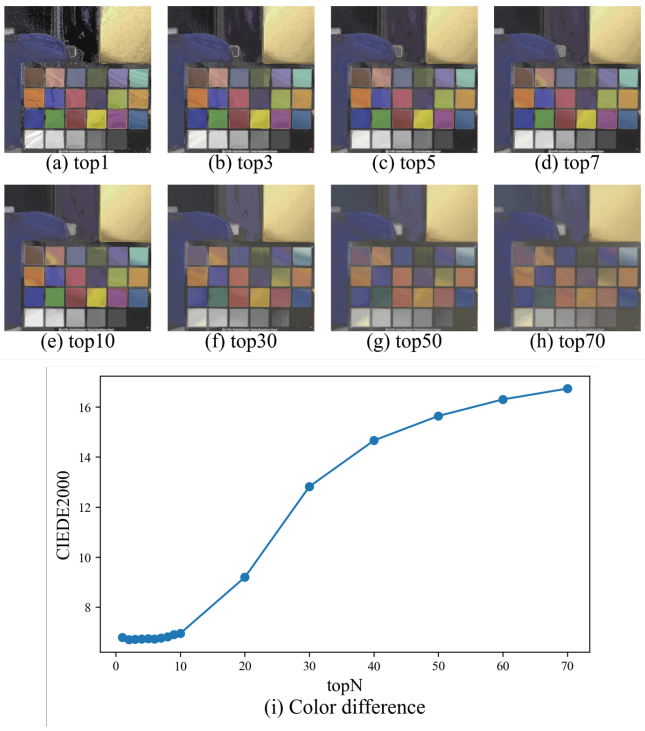

Fig. 9. Visual evaluation results for different number of selected pixels n. (a)-(h) Calculating results of each selected number. (i) Variation of CIEDE2000 evaluation for each number n.

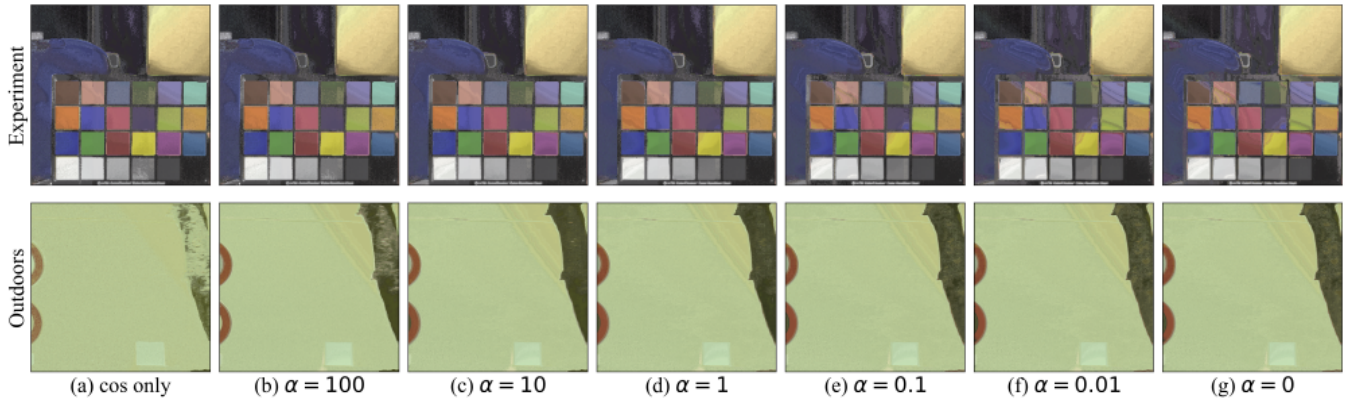

Fig. 10. Visual evaluation results for different values of α. The top row shows the original sparse albedo maps, while the bottom row shows the densified albedo maps using $\alpha = 100, 10, 1, 0.1, 0.01$ from left to right.

for pixels with a non-similar spectrum. Thus, we quantify the color-difference variation by changing the number of selected pixels from one-to-seventy as shown in Tab. 2 and Fig. 9. Following to the results, the best value on the number of selected pixels n is considered to be three. Thus, we selected $n = 3$ in main paper.

B.2. Weight parameter

To evaluate the impact of the weight parameter α on the similarity calculation for albedo densification, we performed an ablation study. The similarity between spectral signatures was calculated using the following equation:

$$\hat{i} = \arg\min_i \Big[\|f(x_j, y_j) - f(x_i, y_i)\| - \alpha \cos(f(x_j, y_j), f(x_i, y_i)) \Big], \quad (1)$$

Here, $\cos(a, b)$ represents the cosine similarity of a and b, while α is the weight parameter balancing the L2 norm and cosine similarity. The cosine similarity helps identify pixels with similar spectral shapes, whereas the L2 norm distinguishes different intensity values with similar spectral shapes, such as white and black walls.

To identify the optimal α, we varied the parameter values and conducted both quantitative and visual evaluations. The results are presented in Tab. 3 and Fig. 10. This ablation study demonstrates how varying α influences the densification quality, providing insights into selecting an appropriate value for accurate albedo estimation. Consequently, $\alpha = 1.0$ or 10.0 are the best for our data, and $\alpha = 1.0$ is selected for this experiment.

C. WHDR ANNOTATION FOR OUR IMAGE

This section provides the WHDR annotation for our image. Initially, we calculated RGB image from hyperspectral images. Then, we sampled 24 points from the color board region. Subsequently, we extracted the neighboring pixels of the sampled points and annotated the darker points and confidence level as shown in Fig. 11. Five annotators, who understand the albedo and WHDR concept, annotated all neighboring pixels of the sampled points. The annotation results were weighted averaged to calculate the WHDR value, where the weight was determined by the annotator's confidence level as described in the existing paper [1].

model	CIE76 (↓)	CIEDE2000 (↓)	Co-effs(↑)
cos only	13.4±1.91	6.78±0.88	0.979
$\alpha = 100.0$	13.3±1.93	6.70±0.89	0.979
$\alpha = 10.0$	13.3±1.93	6.70±0.89	0.979
$\alpha = 1.0$	13.3±1.92	6.71±0.88	0.978
$\alpha = 0.1$	13.5±1.92	6.79±0.88	0.977
$\alpha = 0.01$	13.7±1.93	6.90±0.88	0.974
$\alpha = 0.0$	13.7±1.93	6.92±0.88	0.974

Table 3. Quantitative evaluation results for different values of α. "cos only" represents cosine similarity without l2 norm.

D. VISUAL VERIFICATION FOR COMPLEX SCENE

We focused on the color board due to its true color availability for quantitative evaluation. To verify the applicability for more complex situations, we conducted another experiment with a white cylinder and a doll with detailed undulations. Fig. 12 (top) represents the RGB image and estimated albedos. Compared to the input and Bell's output, ours successfully removes cast shadows. Since the accuracy of our method depends on surface-normal precision, increasing the LiDAR density would enhance our albedo calculation, especially for curved surfaces.

In alignment with previous papers, we primarily focused

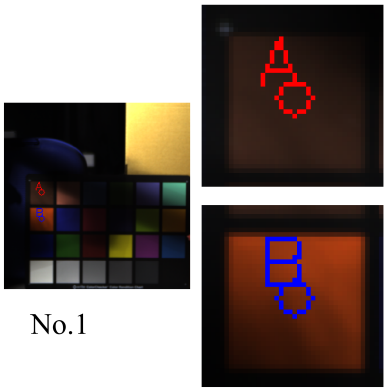

Fig. 11. WHDR annotation for our image. RGB image from hyperspectral images and sampled points from the color board region are depicted.

on albedo. However, we have computed the shading by dividing the RGB image by the calculated albedos. Fig. 12 (bottom) illustrates the calculated shadings including Bell's and ours. In the cast-shadow area highlighted by the red circle, the shadow in Bell's method appears weaker due to the shadow remaining in Bell's output albedo.

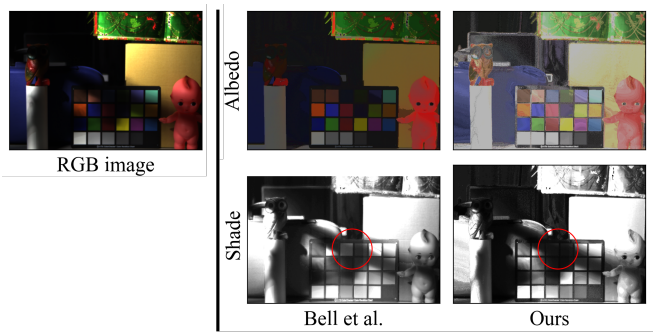

Fig. 12. Visual results of calculated albedos (top) and shades (bottom) for more complex scenarios including input image, Bell [1], and ours.

E. REFERENCES

[1] Sean Bell, Kavita Bala, and Noah Snavely, "Intrinsic Images in the Wild," *ACM TOG*, vol. 33, no. 4, pp. 1–12, 2014.