Intrinsic Image Decomposition Under Multiple Colored Lighting Conditions Using a Single Image

Kakeru Hara
Graduate School of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
Email : m1m20a31@oit.ac.jp

Hiromitsu Isobe
Graduate School of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
Email : m1m21a06@oit.ac.jp

Takao Jinno
Faculty of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
Email : takao.jinno@oit.ac.jp

***Abstract*—We propose a framework for intrinsic image decomposition using a single image captured under multiple colored lights. Our method first estimates the illumination, and then calculates the reflectance based on the retinex theory. For estimating the illumination, we use the depth, illuminant color, and illuminance estimated from the input image. We then demonstrate the validity of our method on CG and real image datasets.**

Keywords*—*intrinsic image decomposition, illuminant color estimation, illumination, shading, reflectance.

I. Introduction

Illumination is composed of illuminant color and shading, and reflectance is an intrinsic color that is invariant to illumination conditions. Both illumination and reflectance are useful for many computer vision algorithms, such as those for segmentation, object detection, shape restoration, color constancy, relighting, and recoloring. Several intrinsic image decomposition methods that can divide an input image into reflectance and illumination components have been proposed.

Conventional intrinsic image decomposition methods can be classified into three approaches: methods based on deep learning [1,2,3,4,5,6], methods that require multiple image inputs [1,7,8], and methods that require user assistance [9]. Because these methods assume that input images are captured under white light or sunlight, they are not suitable for scenes under complex lighting conditions such as multiple colored lights.

This study proposes a framework for estimating illumination and reflectance from a single RGB image captured under multiple colored lights. In recent years, because modulable color LEDs, such as Philips Hue, have become widespread, it is necessary to use computer vision techniques for estimating illumination and reflectance in scenes with temporal and spatial lighting color changes. Our method is valid in these cases because it estimates the intrinsic image from a single RGB image.

To estimate the illumination, our method uses depth, illuminant color, and illuminance, which are estimated from an input image. The depth was estimated using the boosting monocular depth [10] method. It is a method for estimating the depth of a scene from a single image based on the difference in the output characteristics of the depth estimations according to the resolution of the input image.

The illuminant color was estimated by spreading the color of the sparse illuminant color map estimated by gray pixels [11] according to the estimated depth. We used a method developed in a previous study [11] for estimating the illuminant color using the information of achromatic pixels detected according to the illuminant-invariant measurement in color-biased images.

The illuminance was estimated by merging the mean illuminance value with the estimated shading value. The illuminance map was estimated based on the assumption that the maximum value of the RGB channels represents the illuminance in [12]. Subsequently, [13] estimated the illuminance map by reducing the shading component from the results of [12] with structural information. In this study, we estimated the shading from the difference between Max-RGB [12] and the illuminance estimation in Low-light Image Enhancement via Illumination Map Estimation (LIME) [13].

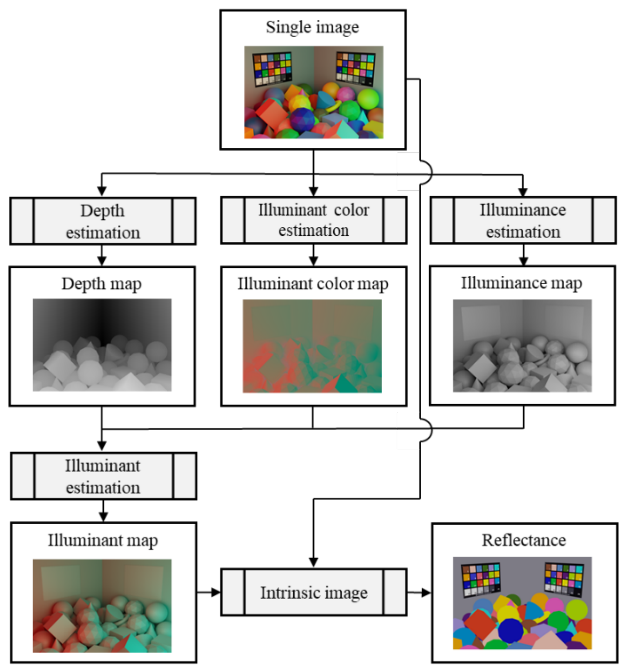

Fig.1. Workflow of our method. The image is the correct CG image.

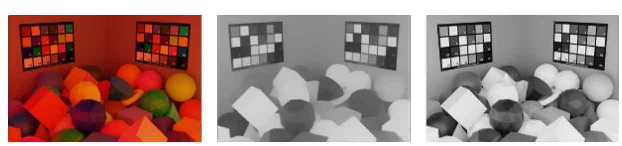

Fig.2. Comparison of illuminance estimation results by LIME and Max-RGB. **Left:** input image. **Middle:** illuminance estimation result by LIME. **Right:** illuminance estimation result by Max-RGB.

ISBN: 978-1-6654-6798-8

II. METHOD

Retinex theory [14] indicates that pixel values I are composed of illumination L and reflectance R as follows:

$$I_c(x) = L_c(x) \cdot R_c(x), \qquad c \in \{r, g, b\}, \tag{1}$$

where $I_c(x)$ is the pixel value at location x in color channel c. Our method calculates the reflectance by dividing the input image by the illumination as

$$R_c(x) = I_c(x)/(L_c(x) + \varepsilon), \qquad c \in \{r, g, b\}, \tag{2}$$

where ε is a very small constant to avoid a zero denominator.

The flow of the proposed method is shown in Fig. 1. The illumination map was estimated using depth, illuminant colors, and illuminance, which were estimated from a single input image.

Because conventional illuminant color estimations estimate sparse illuminant color maps, a spreading method is necessary to estimate the illuminant color of all pixels. Our method uses the conjugate gradient method to spread the colors of the sparse illuminant color map to all the pixels along the structure of the depth map.

We estimate the gradient of the illuminance using the difference between the illuminance estimation in [12] and [13], and then calculate the illuminance by adding the gradient and the mean of the illuminance that is set experientially.

The Max-RGB [12] method uses the maximum RGB at each pixel as illuminance, as shown in (3). Therefore, it is not possible to distinguish between black objects and dark areas in the scene.

$$Y_{MaxRGB}(x) = max(I_c(x)), \qquad c \in \{r, g, b\}. \tag{3}$$

The illuminance estimation in LIME [13] improves the illuminance estimated by Max-RGB [12] using the structure of the scene. Because its result is balanced in the local region, it is similar to the illuminance estimated by [12] for each pixel (Fig. 2). Therefore, by considering these differences, the relative shading of the scene can be obtained as the gradient of the illuminance. Because the relative shading values are distributed around zero, the illuminance map is calculated by adding the mean illuminance value to the relative shading, as shown in (4).

$$Y(x) = Y_{MaxRGB}(x) - Y_{LIME}(\mathrm{x}) + Y_P, \tag{4}$$

where $Y(x)$ is the illuminance at location x, and the mean illuminance value Y_P is set experientially. In this study, it is the mean value of the luminance of the input images in the experiments in the CG image dataset and a fixed value (50% of the range) in the real image dataset.

The illuminant color map is smooth in the local regions, regardless of texture, such as the depth map representing the structure of the scene. In the YUV color space of the illuminant color map, the color difference V of a pixel is expressed as the weighted mean of the color differences U of the surrounding pixels. Therefore, we minimized the difference between the color difference of a pixel and the color difference of the surrounding pixels as follows:

$$J(U) = \sum_x \left(U(x) - \sum_{s \in N(x)} w_{xs} U(s) \right)^2, \tag{5}$$

where x and $N(x)$ are the locations of the target pixel and its neighboring pixels, respectively. $J(V)$ was calculated using the same equation. w_{xs} is the weight whose sum is 1, and is expressed in (6).

$$w_{xs} = \frac{1 + (D(r) - \mu_x)(D(s) - \mu_x)/\sigma_x^2}{W_x}, \tag{6}$$

where μ_r and σ_r represent the mean and variance of the depth map at the pixels surrounding location x, D is the depth, and W_x is the sum of w_{xs} at location x for normalization. The difference between the different signals of different pixels with similar depths is also reduced by these weights.

Solving (5) with the conjugate gradient method, the given illuminant colors are spread to areas where the depth gradient is small and are stopped at boundaries where the depth gradient is large, such as the edges of structures. The illumination map is obtained by assigning the illuminance map calculated by (4) as the luminance of the illuminant color that is colored along the structure.

III. RESULTS

In the evaluation experiment, we compared the estimated illumination map and separated reflectance with the ground truth on the CG image dataset and the real image dataset. The dataset was developed in-house.

The illumination map was evaluated using angular error, and the reflectance was evaluated using both angular error and color difference in CIE DE2000 [15]. Angular error is a measurement used to evaluate illuminant color estimation against the ground truth, where a smaller angle indicates higher accuracy. The angular error is defined in (7).

$$AngularError = arccos\left(\frac{e_{est} \cdot e_{gt}}{\| e_{est} \| \| e_{gt} \|}\right), \tag{7}$$

where e_{est} is the RGB vector of the estimated color and e_{gt} is the RGB vector of the ground truth.

CIE DE2000 is a criterion based on the Euclidean distance in the $L^*a^*b^*$ color space with a correction to bring it closer to human visual evaluation.

Because CG images have ground truth, we used all pixels for evaluation. In real images, the illuminant colors were evaluated using the average colors of the six achromatic patches of the color chart in the image as the ground truth, and the reflectance was evaluated using the reflectance of all patches of the color chart as the ground truth.

A. Dataset

The CG image dataset (Fig.3) consists of 48 scenes with different lighting conditions, combining 12 illuminant colors, their correct illumination maps, and the correct reflectance. These were created through ray-tracing rendering using Blender's cycles renderer. The reflectance of the objects in the scene are coated with the color of the chromatic patches

Fig.3. Examples of images in the CG image dataset. **Top:** scenes under multiple colored lights. **Bottom:** illumination maps.

Fig.4. Examples of images in the real image dataset.

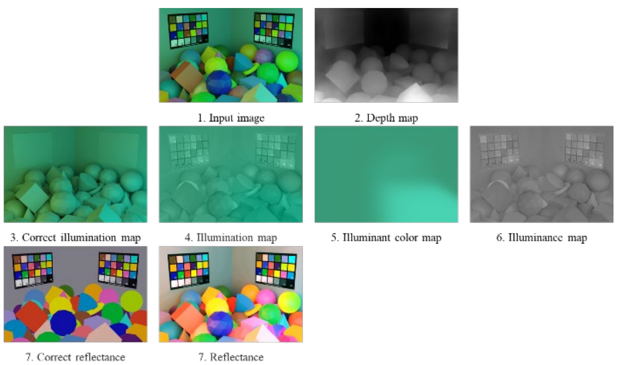

Fig.5. Example of the experimental results for the CG image dataset. Fig. 5(3) is the correct illumination map to compare with Fig. 5(4). Fig. 5(7) is the correct reflectance to compare with Fig. 5(8).

Fig.6. Examples of the result of extracting reflectance with the correct illumination map. **Top:** input image. **Middle:** correct illumination maps. **Bottom:** reflectance extracted with the correct illumination map in the middle row.

of the Macbeth color chart, in which their colors are evenly distributed to all objects. The wall color in the scene was set to 18% gray to make the illuminant colors easier to visualize.

The real image dataset (Fig. 4) consists of 12 images containing the scenes under multi-colored lights. We used Philips Hue Go as the colored lights. The scene has color charts on the left and right and a ball and building blocks at the center. A color chart was placed in the scene because it had the correct reflectance used in the evaluation. The scene was photographed without any external light.

Table 1. Mean statistics of angular error at each pixel for the estimated and the correct illumination map.

	median	mean	trim-mean	top 25% mean	bottom 25% mean
Angular error	5.8757	7.6560	6.2382	2.4869	15.6608

Table 2. Mean statistics of angular error and color difference at each pixel for the extracted correct reflectance.

	median	mean	trim-mean	top 25% mean	bottom 25% mean
Angular error	9.2039	11.9193	9.6705	3.4618	24.8745
Color difference	21.4909	22.6873	21.3231	9.0581	39.0451

Table 3. Mean statistics of angular error and color difference at each pixel for the reflectance extracted with the correct illumination map (Fig. 6) instead of the estimated illumination map and the correct reflectance as a comparison to Table 2.

	median	mean	trim-mean	top 25% mean	bottom 25% mean
Angular error	5.6971	7.8710	6.1619	1.6021	17.5582
Color difference	20.4332	21.7530	20.8906	11.6209	33.6099

B. Results of the CG image dataset

Fig. 5 shows an example of the experimental results for the CG image dataset. Fig. 5(3) shows the correct illumination map for comparison with Fig. 5(4). Fig. 5(7) shows the correct reflectance for comparison with Fig. 5 (8).

The statistics for the results of the evaluation experiment are presented in Tables 1–3. Table 1 lists the mean statistics of the angular error at each pixel for the estimated and correct illumination maps. Table 2 presents the mean statistics of the angular error and color difference at each pixel for the extracted and correct reflectance. Table 3 shows the mean statistics of the angular error and color difference at each pixel for the reflectance extracted with the correct illumination map instead of the estimated illumination map and the correct reflectance as a comparison to Table 2. The angular error values in Table 3 are the lower limit values for this method based on the retinex theory.

The values in the tables are the means of the statistics of the criteria for all pixels for the 48 CG image dataset. The trim-mean is the mean of the middle 50% of the results, sorted results of the evaluation on all pixels of each image, and excluding the results of the top 25% and bottom 25%. The means of the excluded top 25% and bottom 25% are shown in the top 25% and bottom 25% means, respectively.

C. Results of the real image dataset

Fig. 7 shows an example of the experimental results for a real image dataset. Fig. 7(3) shows the average colors of the six achromatic patches of the color chart in the input image. Fig. 7(8) shows the average color of the color patches in Fig. 7(7). Fig. 7(9) shows the correct colors for the color patches.

The statistics for the results of the evaluation experiment are presented in the Tables 4–5. Table 4 shows the mean statistics of the angular error and color difference for each color patch for the extracted reflectance and correct

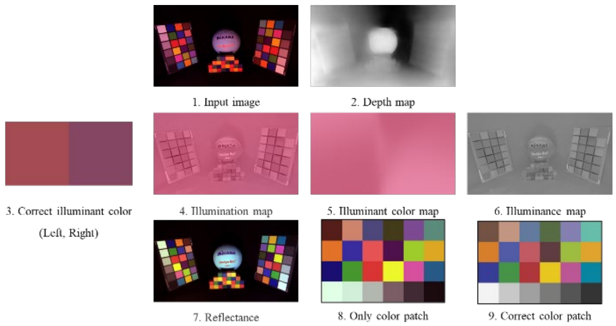

Fig.7. Example of the experimental results for the real image dataset. Fig. 7(3) is the average color of the six achromatic patches of the color chart in the input image. Fig. 7(8) is the average color of the color patches of Fig. 7(7). Fig. 7(9) is the correct color of the color patches.

Table 4. Mean statistics of angular error and color difference of each color patch for the extracted reflectance and the correct reflectance.

	median	mean	trim-mean	top 25% mean	bottom 25% mean
Angular error	12.0145	13.1008	12.2855	3.8371	23.9951
Color difference	14.2243	14.2424	14.2353	6.5602	21.9386

Table 5. Mean statistics of angular error and color difference of each color patch for the reflectance extracted with the correct illumination map instead of the estimated illumination map and the correct reflectance for the CG image dataset as a comparison to Table 4.

	median	mean	trim-mean	top 25% mean	bottom 25% mean
Angular error	10.0341	11.4563	10.1836	2.2343	23.2240
Color difference	17.8175	17.6386	17.6122	10.0737	25.2565

reflectance. Table 5 shows the mean statistics of the angular error and color difference for each color patch for the reflectance extracted with the correct illumination map instead of the estimated illumination map and the correct reflectance for the CG image dataset, as a comparison to Table 4.

D. Consideration

We considered the trends of images with high or low evaluation values through an evaluation experiment.

Fig. 8 and Fig. 9 show examples with high and low evaluation values of the CG image in terms of angular error, respectively. When the hues of the left and right lighting colors were close, images of the scenes illuminated by high perception sensitivity illuminant colors, such as green, tended to have high evaluation values. Conversely, images of scenes illuminated by low-perception sensitivity illuminant colors such as purple tended to have low evaluation. This is because of the small difference between the color of the background area of the reflectance and the correct reflectance. Note that high perception sensitivity indicates high luminance in the images. In addition, when the hues of the left and right illuminant colors were far, the evaluation values tended to be higher because the color mixing decreased saturation and increased the luminance of the illuminance colors in the images, which increased the mean illuminance value added in (4).

Fig. 10 and Fig. 11 show examples with high and low evaluation values of the CG image in the CIE DE2000, respectively. When the hues of the left and right lighting colors were similar, images of the scenes illuminated by illuminant colors with hues ranging from red to purple tended to have high evaluation values. Conversely, images of scenes illuminated by high-luminance illuminant colors, such as green, tended to have low evaluation values. This is because of the small color difference between the color of the background area of the reflectance and that of the background area of the correct reflectance. In addition, when the hue distance between the left and right illuminant colors was large, the evaluation values tended to be lower because the color mixing decreased saturation and increased the luminance of the illuminance colors in the images, which increased the mean illuminance value added in (4).

There is an inverse trend between the angular error and color difference in the evaluation value for the hue distance of the two colored lights. When there is a difference in luminance of the reflectance, the color difference also becomes large because the color difference evaluates not only colors, such as angular error, but also the luminance of colors. However, the angular error is not affected. To avoid increasing the color difference, we should modify the calculation of the mean illuminance Y_P. Specifically, it is necessary to apply not only the offset, such as T_P but also the linear conversion such that the gradient of the illuminance matches the shade of the input image to improve the accuracy.

In the case of misestimation of the illuminant color, the evaluation was low in terms of both angular error and color difference. The accuracy of the estimated illuminant colors significantly affected the results. This was also the case in the experiment that used real images.

IV. CONCLUSION

Illumination is composed of illuminant color and shading, and reflectance is an intrinsic color that is invariant to illumination conditions. Conventional intrinsic image estimations can be classified into three approaches: methods based on deep learning, methods that require multiple image inputs, and methods that require user assistance. Because these methods assume that input images are captured under white light or sunlight, they are not suitable for scenes under complex lighting conditions such as multiple colored lights.

We proposed a framework for intrinsic image estimation from a single image captured under multiple colored lights. Our method first estimated the illumination and then calculated the reflectance based on retinex theory. To estimate the illumination, our method used the depth, illuminant color, and illuminance, which were estimated from an input image. We then demonstrated the validity of our method on the CG and real image datasets.

The results of the evaluation experiment show that there is an inverse trend between the angular error and color difference in the evaluation value for the hue distance of the two colored lights. To avoid increasing the color difference, we modified the calculation of the mean illuminance, T_P. To improve the accuracy, it is necessary to apply not only the offset, such as T_P but also the linear conversion so that the gradient of the illuminance matches the shade of the input image.

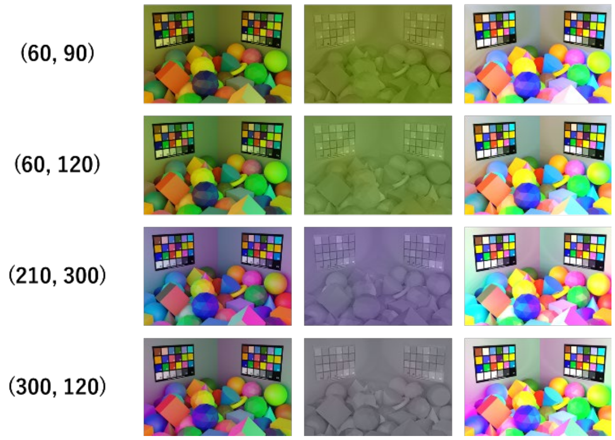

Fig.8. Examples with a high evaluation in angular error. Numbers in parentheses are hues of illuminant colors. **Left:** input image. **Middle:** illumination map. **Right:** reflectance.

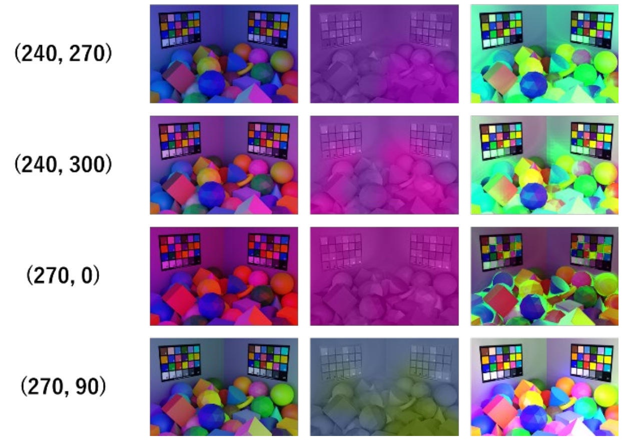

Fig.9. Examples with a low evaluation in angular error. Numbers in parentheses are hue of illuminant colors. **Left:** input image. **Middle:** the illumination map. **Right:** the reflectance.

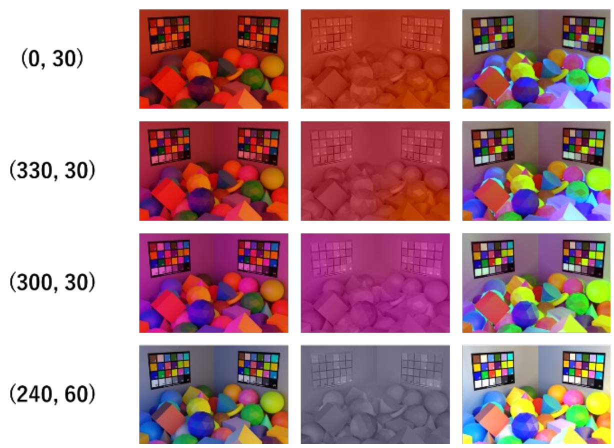

Fig.10. Examples with a high evaluation in color difference. Numbers in parentheses are hues of illuminant colors. **Left:** input image. **Middle:** illumination map. **Right:** reflectance.

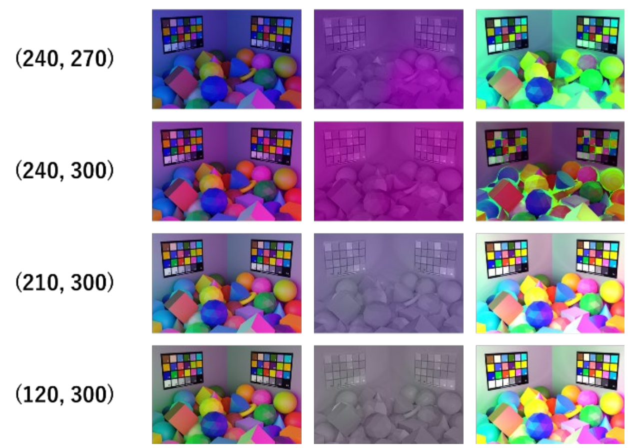

Fig.11. Examples with a low evaluation in color difference. Numbers in parentheses are hue of illuminant colors. **Left:** input image. **Middle:** the illumination map. **Right:** the reflectance.

REFERENCES

[1] W. Ma , H. Chu, B. Zhou, R. Urtasun, and A. Torralba, "Single image intrinsic decomposition without a single intrinsic image," ECCV, 2018.

[2] J. Chang, R. Cabezas, and J. W. Fisher III, "Bayesian nonparametric intrinsic image decomposition," ECCV, 2014.

[3] J. Jeon, S. Cho, X. Tong, and S. Lee, "Intrinsic image decomposition using structure-texture separation and surface normals," ECCV, 2014.

[4] Y. Liu, Y. Li, S. You, and F. Lu, "Unsupervised learning for intrinsic image decomposition from a single image," CVPR, 2020.

[5] M. Janner, J. Wu, T. D. Kulkarni, I. Yildirim, and J. B. Tenenbaum, "Self-supervised intrinsic image decomposition," NIPS, 2017.

[6] T. Zhou, P. Krahenbuhl, and A. A. Efros, "Learning data-driven reflectance priors for intrinsic image decomposition," ICCV, 2015.

[7] P. Y. Laffont, A. Bousseau, and G. Drettakis, "Rich intrinsic image decomposition of outdoor scenes from multiple views," IEEE Trans. Vis. Comput. Graph., vol.19, no.2, pp.210-224, Feb. 2013.

[8] K. J. Lee, Q. Zhao, X. Tong, M. Gong, S. Izadi, S. U. Lee, P. Tan, and S. Lin, "Estimation of intrinsic image sequences from image+depth video," ECCV, 2012.

[9] J. Shen, X. Yang, X. Li and Y. Jia, "Intrinsic image decomposition using optimization and user scribbles," IEEE Trans. Cybernetics, vol.43, no.2, pp.425-436, Apr. 2013.

[10] S. M. H. Miangoleh, S. Dille, L. Mai, S. Paris, and Y. Aksoy, "Boosting monocular depth estimation models to high-resolution via content-adaptive multi-resolution merging," CVPR, 2021.

[11] K. F. Yang, S. B. Gao, and Y. J. Li, "Efficient illuminant estimation for color constancy using grey pixels," CVPR, 2015.

[12] E. Land, "The retinex theory of color vision," Sci. Am., vol.237, no.6, pp.108-128, Dec. 1977.

[13] X. Guo, Y. Li, and H. Ling, "LIME: Low-light Image Enhancement via Illumination Map Estimation," IEEE Trans. Image Process, vol.26, no.2, pp.982-993, Feb. 2017.

[14] E. H. Land, and J. J. McCann, "Lightness and retinex theory," J Opt. Soc. Am., vol.61, no.1, pp.1–11, 1971.

[15] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations, " Color Res., vol.30, no.1, pp.21-30, Appl. 2005.