IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

VOL. E104-A NO. 2
FEBRUARY 2021

A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY
The Institute of Electronics, Information and Communication Engineers
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN

IEICE TRANS. FUNDAMENTALS, VOL.E104–A, NO.2 FEBRUARY 2021

PAPER

Single Image Haze Removal Using Iterative Ambient Light Estimation with Region Segmentation*

Yuji ARAKI[†], **Kentaro MITA**[†], ***Student Members***, ***and*** **Koichi ICHIGE**[†a)], ***Member***

SUMMARY We propose an iterative single-image haze-removal method that first divides images with haze into regions in which haze-removal processing is difficult and then estimates the ambient light. The existing method has a problem wherein it often overestimates the amount of haze in regions where there is a large distance between the location the photograph was taken and the subject of the photograph; this problem prevents the ambient light from being estimated accurately. In particular, it is often difficult to accurately estimate the ambient light of images containing white and sky regions. Processing those regions in the same way as other regions has detrimental results, such as darkness or unnecessary color change. The proposed method divides such regions in advance into multiple small regions, and then, the ambient light is estimated from the small regions in which haze removal is easy to process. We evaluated the proposed method through some simulations, and found that the method achieves better haze reduction accuracy even than the state-of-the art methods based on deep learning.

key words: *haze removal, dark channel, HSV color space, region segmentation*

1. Introduction

Various image-restoration techniques have been developed in recent decades, and single-image haze removal is a key process in image restoration. Single-image haze removal aims to sharpen images by removing haze, fog, clouds, and any other blurring factors. Image-sharpening techniques have gained attention in various fields, such as satellite imaging [1], image classification [2], remote sensing [3], and moving-image analysis [4]. Haze-removal methods often estimate the transmission map, i.e. the attenuation amount of the light reflected from the subject of the photograph, and the ambient light, i.e. light in the atmosphere, and then derive a restored image with linear processing [5], [6]. In general, all the red, green, and blue (RGB) values in colored haze images become equal because haze has a strong white component. However, at least one of the RGB components takes a small value in regions in which no haze exists.

Independent component analysis (ICA) [7] can be used for haze removal, but it requires a long calculation time and is not effective for processing images with a large amount of haze. A haze reduction method using color line [8] is also attracted but it often removes too much haze and may make images with inappropriate contrast. Also this method has a drawback that the ambient light vector should be specified in advance of haze reduction process. The method [9] focuses to accurately estimate the ambient light, however it cannot accurately remove haze from images with sky regions. A method that uses a dark channel [10] creates an image called a dark 0channel that selects the smallest RGB values in each pixel and estimates the transmission map and ambient light. However, there is an inherent problem in this method: the transmission map and the ambient light cannot be accurately estimated if there is white or sky regions in the input image. Therefore, a method to create multiple dark channels with different patch sizes [11] has been proposed, but the haze-amount estimation in sky regions is still insufficient.

Recently the learning-based methods are very much attracted and are already used in many applications [12], [13]. A method that uses a convolutional neural network (CNN) [12] is effective for processing sky regions, but the color of regions without haze also changes when this method is used. Also this kind of CNN-based method often does not work accurately for dense haze images. The method employing generative adversarial networks (GAN) [13] is also effective for processing sky regions but often has a same problem with the CNN-based methods.

Recall that a method using hue, saturation, and value (HSV) color space [14] uses a depth map to estimate the transmission map without approximating the white region. The drawback of this approach is that it is difficult to restore images that have large haze regions. We can improve the accuracy of the transmission map by applying a guided filter [15] to the depth-map processing for the images with a small white area. This method works well overall, but it has difficulty processing white and sky regions, as is the case with the methods using dark channels. A different method divides the images into small regions in advance and employs parallel processing [16]. This method divides sky regions and reduces haze well using a dark channel, but the accuracy of the haze removal is insufficient.

In this paper, we modify the method [14] so that we can estimate the number of iterations prior to haze-removal processing. We also propose a method to estimate optimal ambient light by dividing regions for which it is difficult to estimate the haze amount. The iterative method has an advantage that we can adaptively remove haze depending on how much the given image is hazed. Therefore the method is effective to restore dense haze images. Performance of

Manuscript received April 1, 2020.
Manuscript revised June 22, 2020.
Manuscript publicized August 6, 2020.

[†]The authors are with Department of Electrical and Computer Engineering, Yokohama National University, Yokohama-shi, 240-8501 Japan.

*Preliminary version of this manuscript have been presented in [19].

a) E-mail: koichi@ynu.ac.jp

DOI: 10.1587/transfun.2020EAP1038

the proposed method is evaluated using numerical values for artificially synthesized images and images with actual haze.

The structure of this paper is as follows. Section 2 gives a brief outline of existing methods for haze removal. Section 3 describes our proposed method. In Sect. 4, we evaluate the proposed method through simulation and we state the results in Sect. 5. In Sect. 6, we discuss the haze that remains after haze removal in the simulated results. We conclude with a summary in Sect. 7.

2. Preliminaries

We introduce some existing methods for estimating an image's ambient light and transmission map, which are essential for haze removal. We also discuss the method of dividing white and sky regions in an image for which the haze-removal processing is difficult.

2.1 Observation Model

Figure 1 shows a way of creating haze images. A haze image is generated by synthesizing the transmission map (i.e. how much the reflected light from the image subject reaches the camera lens directly) and ambient light from the light source. Therefore, the input image in the haze-removal process is formulated as

$$I(x) = t(x)J(x) + (1 - t(x))A, \tag{1}$$

As (1) shows, we can linearly derive the desired image J if the transmission map t and ambient light A can be estimated from an input image.

2.2 Estimation of Ambient Light and Transmission Map

To estimate the transmission map, we first transform the input image to the HSV color space and estimate the depth image d on the basis of the following model [14]:

$$d(x) = \theta_0 + \theta_1 v(x) + \theta_2 s(x) + \varepsilon(x), \tag{2}$$

where $\theta_0, \theta_1, \theta_2$ are constants regardless of the input image, ε is the Gaussian function with the standard deviation σ, v is brightness, and s is saturation. The model (2) means that the depth map $d(x)$ increases in proportion to $\{v(x) - s(x)\}$, and therefore we estimate the transmission map using the equation

$$t(x) = e^{-d(x)}. \tag{3}$$

Using (3), we can determine the depth map estimated from (2) within the range of $0 < t(x) < 1$. The ambient light A can be modeled as

$$A = I(x_{dmax}), \tag{4}$$

where x_{dmax} denotes the pixel at which the luminance value of the depth map reaches the maximum. This is because with (1), the pixel with the largest depth gives the smallest transmission map and is the closest to the ambient-light

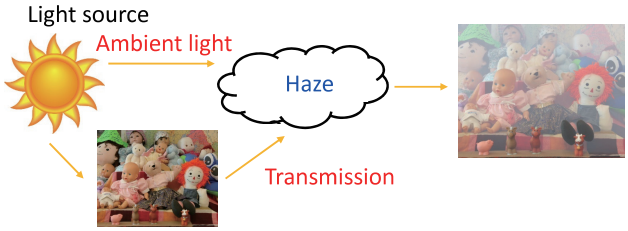


Fig. 1 Example of obtaining haze images.

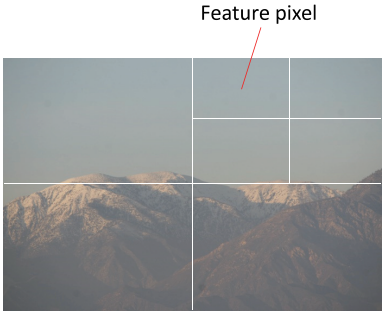


Fig. 2 Determination of feature pixel.

value. We aim to estimate the desired image J using the transmission map t and ambient light A, which are estimated by (3) and (4), respectively.

2.3 Region Segmentation Based on Feature Pixel

White and sky regions must be appropriately processed in haze removal. If these regions are processed in the same way as other regions, the luminance value of the output image decreases and the color of the image may be excessively changed. Therefore, we must divide the image into the regions in which haze removal processing is easy and difficult.

First we specify the pixel called the feature pixel in the region in which haze-removal processing is difficult, as shown in Fig. 2. The input image is divided into four regions, and the averaged pixel value in i-th region I_i^{light} is calculated by

$$I_i^{light} = \frac{1}{N_{\varphi,i}} \sum_{n=1}^{N_{\varphi,i}} \left(\frac{1}{3} \sum_{c \in \{r,g,b\}} I^c(n) \right), \tag{5}$$

$$\varphi_{new} = \max_{I_i^{light}} (\varphi), \tag{6}$$

where φ is a set of divided regions. Then, we select a region satisfying (6) from φ, and set the region as φ_{new}. The parameter i is the set that satisfies

$$i \in \left\{ 1,2,3,4 \,\middle|\, I_i^{grad} \leq \frac{1}{4} \sum_{i=1}^{4} I_i^{grad} \right\}, \tag{7}$$

where I_i^{grad} represents the contour (averaged high-frequency components) of the i-th region [16]. We regard the i-th region which satisfies the condition (7) as the region where the haze removal process is difficult, because of lower contour and

greater average pixel values. The rest regions which do not satisfy the condition (7) are regarded as the region where the haze amount can be easily estimated.

The region with lower contour and greater average pixel values is selected from the divided four regions. This is because white and sky regions for which haze removal is difficult often have lower contour and greater average pixel values. Next, we select the region φ_{new} and again divide it into four regions in the same manner. This process is repeated until the regions reach a certain small size. The pixel at the center of the selected region in the final process is chosen as the feature pixel. The pixels similar to the feature pixel are selected using mean shift filtering [17] to determine the regions to be divided.

3. Proposed Method

We present an overview followed by the details of our proposed method. First, we estimate the optimal number of iterations in haze removal. Next, we describe why there are regions in which haze removal is difficult. Then, we develop a method of using region segmentation in estimating ambient light and the number of iterations, after which we explain the method of adjusting the amount by which haze is reduced at the last iteration.

3.1 Optimal Number of Iterations

First, we estimate the optimal number of iterations for us to repeat the haze-removal process. A flowchart of this process is shown in Fig. 3.

Initially, it is desirable to terminate the iteration when the ambient-light location changes because the ambient light can be regarded as a pixel value that has the maximum haze. However, the location of the ambient light may change in each iteration when haze exists uniformly in the image. Therefore, we repeat the haze-removal process twice and see whether or not the ambient light changes. We regard the ambient light as unchanged when the change in the location of the ambient light satisfies

$$||A_a - A_b|| \leq A_{th}, \tag{8}$$

where A_a, A_b, and A_{th} are the ambient light of the output image, that of the input image, and the threshold of the change in the location of the ambient light, respectively. We regard the ambient light as changed when (8) is not satisfied. We judge if the haze-removal process should be repeated or terminated by calculating the difference Δt of the transmission maps:

$$\Delta t = \frac{1}{N}\sum_{x=1}^{N} |t_a(x) - t_b(x)|, \tag{9}$$

where N, t_a, and t_b are the number of pixels, the transmission map of the output image, and that of the input image, respectively. If Δt is sufficiently small, it can be considered that there are many pixels without change. Therefore, the threshold τ is set, and recursive processing ends, when $\Delta t < \tau$. A pixel whose transmission map exceeds t_{max} is considered as having a small change. We can avoid unnatural color change caused by too many iterations by terminating the iteration when the number of small-change pixels t_{ec} satisfies

$$\frac{t_{ec}}{N} - \delta > 0, \tag{10}$$

where δ is the parameter that judges if the iteration should be terminated or not. Using (10), the iteration is terminated when the ratio of pixels whose transmission map is larger than t_{max} exceeds δ. If the iteration is not terminated, we replace the image I with the estimated one J_a and continue the iterative processing.

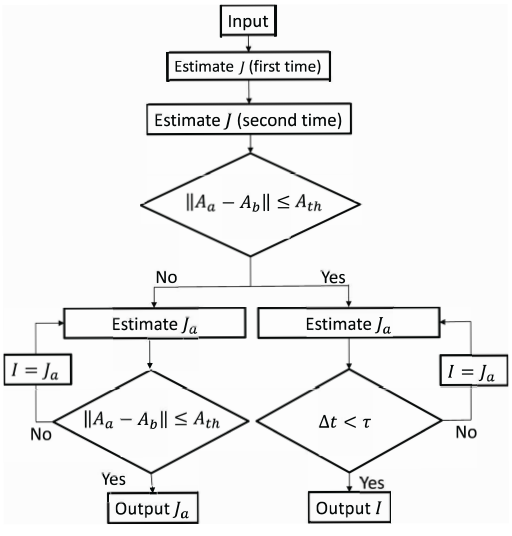
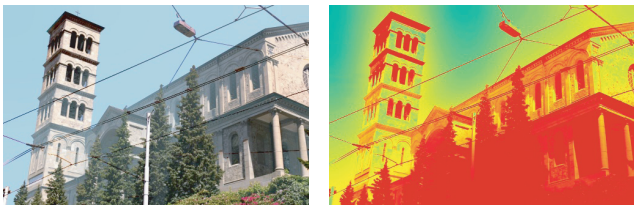


Fig. 3 Flowchart of determining the number of iteration.

3.2 Effect of False Detection Region of Transmission Map

As seen from (1) and Fig. 1, the value of the transmission map becomes 1 if there is no haze because the reflected light from the subject directly reaches the observation point. However, the reflected light can hardly reach the observation point directly when the distance from the subject is very large, as is the case in sky regions. Thus, the transmission map decreases below the actual value, as shown in Fig. 4. The same situation occurs when the desired image contains a white region, as in Fig. 5. This is because the transmission map is calculated under the assumption that the haze is white. In addition, the ambient light is selected as the pixel with the smallest transmission map, i.e. the pixel value with the greatest amount of haze. If an erroneously detected region exists in the transmission map, we may select the ambient light from the erroneously detected region, which will degrade the output-image quality.

For images that contain a large sky or white region, there are many regions in which the haze amount is incorrectly estimated. Because of this, the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and the contrast will



(a) Input image (b) Transmission

Fig. 4 False detection example of sky region.

(a) Input image (b) Transmission

Fig. 5 False detection example of white region.

Fig. 6 Example with too large number of iteration (SSIM=0.49, PSNR=8.63 dB).

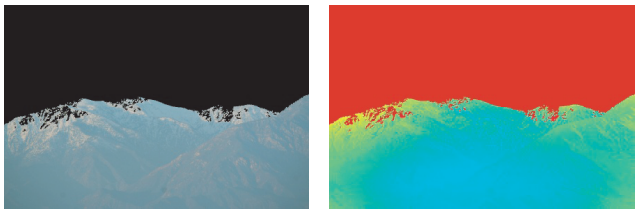
decrease, as shown in Fig. 6, because of excessive iterations.

3.3 Selection Region of Ambient Light

The existing method [16] regards the pixel value with the greatest amount of haze as representative of the ambient light, as discussed in Sect. 2. However, as described in Sect. 3.2, it may not be optimal to select the ambient light from the input image because there may exist regions for which estimation of the haze amount is difficult depending on the input-image characteristics. For this reason, we develop a method to determine the ambient light from the region in which the haze amount is easily estimated using the region segmentation described in Sect. 2.3. Here, the ambient light is given by

$$A = I_{main}(x_{dmax}), \tag{11}$$

where I_{main} is the region in which the amount of haze can be easily estimated, and the value of the ambient light is given as the pixel value at which the depth of I_{main} becomes the maximum. Using (11), we can remove haze without selecting the ambient light from the region in which estimating the haze amount is difficult.



(a) Input image (b) Transmission

Fig. 7 Used region of transmission map in determining iteration number.

3.4 Application of Region Segmentation to Transmission Map Estimation

As described in Sect. 3.1, we use the difference of the transmission maps and the number of pixels for which the value of the $t(x)$ exceeds t_{max} to determine the number of iterations. However, it is difficult to accurately estimate a sufficient number of iterations for images that contain a region very far from the main image subject, such as sky and white regions. Therefore, we propose a method that uses only the region in which the haze amount can be estimated to determine the number of iterations, as shown in Fig. 7. Now, we modify (9) as

$$\Delta t_{main} = \frac{1}{N_{main}} \sum_{x=1}^{N_{main}} |t_{main,a}(x) - t_{main,b}(x)|, \tag{12}$$

where $t_{main,a}$ and $t_{main,b}$ are the images in which the transmission map of the output matches the selected ambient-light region. Using (12), we can calculate the average difference of the transmission maps only for the selected ambient-light region. The existing method of determining the number of iterations also uses the difference of the transmission maps in the region in which the amount of haze cannot be accurately estimated. With the proposed method, we can determine the number of iterations for the images that were difficult in the existing method because we only use the regions in which haze is easily estimated for calculating the transmission-map difference.

3.5 Weighted Transmission Map

Removing haze from images in which the haze amount differs by location is often difficult. For this reason, we adjust the haze-removal intensity with the weighted transmission map used in the last iteration. First, we extract the white component from the image that is output by the automatically determined number of iterations using [18]

$$\boldsymbol{p} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}} \boldsymbol{a}, \tag{13}$$

where $\boldsymbol{p}, \boldsymbol{b}$, and $\boldsymbol{a} = [1, 1, 1]$ are the vectors indicating the extracted white component, the pixel value of the output image, and the direction of the white component, respectively. Then, we measure the error map $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$, which denotes the difference between the extracted white component $\boldsymbol{p}$ and

the pixel value $\boldsymbol{b}$. This indicates the degree to which the pixels in the desired image contain vivid color. Haze removal becomes easier as the error map $\boldsymbol{e}$ increases.

However, it becomes more difficult as the error map $\boldsymbol{e}$ decreases. There are image regions for which it is difficult to remove haze with the transmission map obtained from the depth map because often, only a small amount of haze remains in the last iteration. For these reasons, we apply weights to the estimated depth map and the transmission map used in the last iteration, i.e.

$$d_w(x) = d(x) \cdot \left\{ \frac{\max(w(x))}{w(x)} \right\}^{\frac{1}{m}}, \tag{14}$$

$$t_w(x) = \exp[-\beta d_w(x)], \tag{15}$$

where m is an adjustment factor, and $w(x)$ denotes the weight given by

$$w(x) = \sqrt{e_r^2(x) + e_g^2(x) + e_b^2(x)}, \tag{16}$$

which represents the length of the error vector in each pixel. The weighted depth map $d_w(x)$ improves the haze-removal processing in the regions in which the white component is large and hinders it in the regions in which the white component is small. The factor m is empirically determined as

$$m = \frac{1}{3} e_{sum}, \tag{17}$$

where e_{sum} is obtained from error map $\boldsymbol{e}$ by

$$e_{sum} = \frac{1}{N} \sum_{x=1}^{N} \left(\frac{1}{3} \sum_{c \in \{r,g,b\}} \boldsymbol{e}(\boldsymbol{x}) \right). \tag{18}$$

The depth map must be reduced by weighting the image pixels that have large regions with vivid colors in the desired image. We can adjust the vividness of the whole image and of the weights at the same time using (15).

4. Simulation

The performance of the proposed method was evaluated using computer simulation.

4.1 Simulation Specifications

The simulation parameters are listed in Table 1 and the ground-truth images are shown in Fig. 8, where the parameters θ_0 to σ are the parameters used in (2) and adopted the same values as those in [14]. We used the values $\tau = 0.1$ and $\delta = 0.8$ as thresholds. The block size is the minimum value of φ_{new} when determining feature pixels to be used in region segmentation. Also, the proposed algorithm for determining the number of iteration is called "Proposed 1" when not using region segmentation and "Proposed 2" when region segmentation is used described in Sect. 3.3 and Sect. 3.4.

Table 1 Simulation parameters.

β	1.0
θ_0	0.122
θ_1	0.959
θ_2	−0.780
σ	0.0413
τ	0.1
δ	0.8
A_{th}	4
t_{max}	0.9
Block size	50

(a) couch, 1024 × 680 (b) doll, 512 × 409

(c) fruits, 768 × 480 (d) pizza, 640 × 400

(e) church, 1000 × 655 (f) road, 960 × 540

Fig. 8 Tested Images (ground truth).

4.2 Optimal Number of Iteration

Tables 2 and 3 gives the simulated results to test if the estimated number of iterations is optimal or not. The tested images are as shown in Fig. 9. Tables 2 and 3 indicates the determined number of iterations in bold. Note that the best number of iteration for the image "doll" is different for PSNR an SSIM, i.e., 1 for PSNR but 2 for SSIM. To determine which value is better, we evaluated which is better not only from the objective evaluation in Tables 2 and 3 but also the subjective evaluation. We found that the image becomes a bit dark in case of 2 iterations, therefore we adopt 1 iteration as the best number of iteration for the image "doll". As can be seen from the values of PSNR and SSIM, the proposed method can estimate the appropriate number of iterations well overall, but it does not work well for images that have large sky and white regions, as in Fig. 6.

Table 2 Relationship between iteration and PSNR.

Image	number of iteration			
	1	2	3	4
couch	20.03	**26.39**	24.80	19.43
doll	**17.38**	18.52	14.82	12.58
fruits	9.84	15.36	**19.52**	18.93
pizza	17.20	**22.75**	16.54	13.12

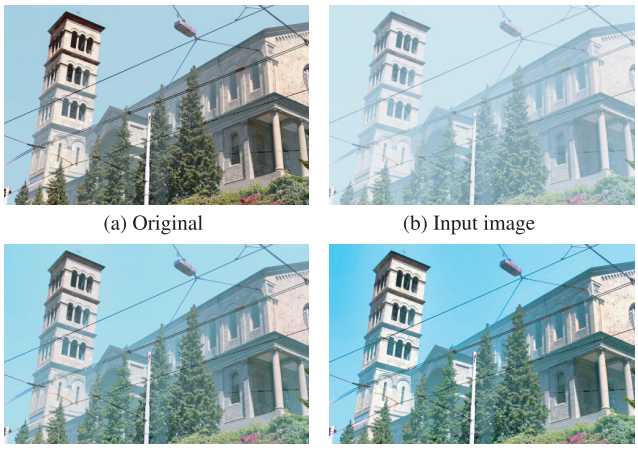
Table 3 Relationship between iteration and SSIM.

Image	number of iteration			
	1	2	3	4
couch	0.83	**0.92**	0.91	0.51
doll	**0.84**	0.79	0.65	0.50
fruits	0.63	0.73	**0.77**	0.72
pizza	0.89	**0.95**	0.89	0.75

(a) Original (b) Input image

(c) Case of 2 iterations (d) Case of 3 iterations

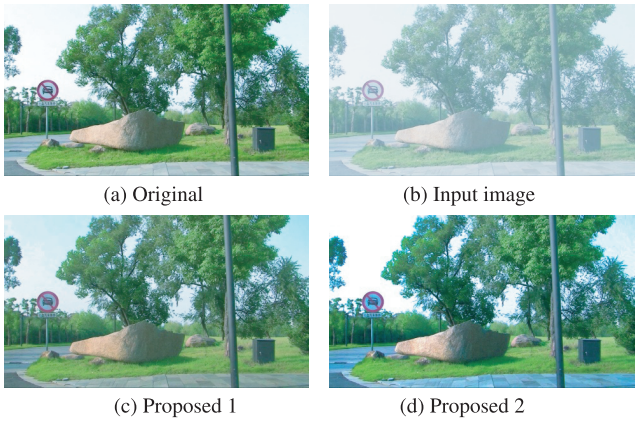
Fig. 9 Comparison of 2 iterations and 3 iterations (fluits).



(a) Original (b) Input image

(c) Proposed 1 (d) Proposed 2

Fig. 10 Comparison of Proposed 1 and Proposed 2 (church).



(a) Original (b) Input image

(c) Proposed 1 (d) Proposed 2

Fig. 11 Comparison of Proposed 1 and Proposed 2 (road).

(a) church (b) road

Fig. 12 Ambient light selection region.

4.3 Application of Region Segmentation

The original and input images are shown in Figs. 10(a), 10(b), 11(a), and 11(b). The output images are as shown in Figs. 10(c), 10(d), 11(c), and 11(d). We can visually recognize that the Proposed 2 expresses color better in both images. In particular, more natural color can be expressed in the sky region, and contrast reduction can be suppressed.

In the without segmentation method, ambient light is selected from the sky or white region, where the amount of haze is detected to be greater than its actual value. As a result, the selected ambient light is deemed to have a higher luminance than it actually does. This is clear from the region where the ambient light is selected, as in Fig. 12. The segmentation method also divides the non-sky regions in the image of the church. This is because that region appears close to white in the input image; thus, these regions do not affect the output image.

Tables 4 and 5 show the values of PSNR and SSIM for each iteration in the proposed 1 and proposed 2, and the value at the terminated iteration is indicated in bold. The image of the church in Fig. 6 looks over-processed; the haze-removal accuracy decreased because the Proposed 1 could not sufficiently estimate a suitable number of iterations. In contrast, the Proposed 2 can determine the number of iterations appropriately. However, iteration 1 in the Proposed 1 is better than iteration 2 in the Proposed 2 in terms of the values of SSIM and PSNR. This is because the Proposed 1 represents the color of the sky region similarly to the original image, as seen from Fig. 10. The Proposed 2, though, can express a more vivid sky color than the original image. Thus the Proposed 2 is more effective than the Proposed 1.

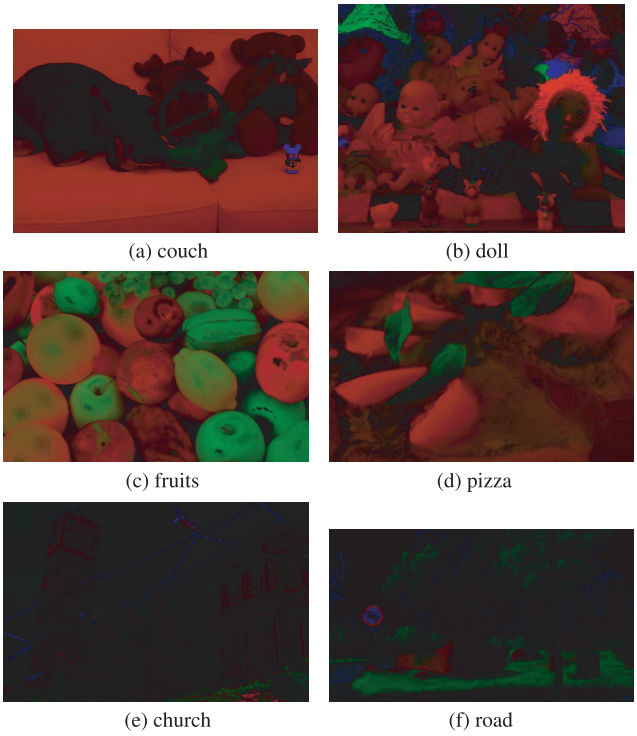
We also tested the with segmentation and without segmentation methods for another image (the road image). It

Table 4 Relationship between iteration and PSNR.

Image	number of iteration			
	1	2	3	4
church (Proposed 1)	25.42	16.28	10.73	**8.63**
church (Proposed 2)	23.00	**23.20**	17.0	13.00
road (Proposed 1)	21.73	**22.61**	17.77	13.94
road (Proposed 2)	17.49	23.41	**23.54**	18.90

Table 5 Relationship between iteration and SSIM.

Image	number of iteration			
	1	2	3	4
church (Proposed 1)	0.97	0.91	0.70	**0.49**
church (Proposed 2)	0.96	**0.96**	0.90	0.77
road (Proposed 1)	0.93	**0.95**	0.89	0.68
road (Proposed 2)	0.90	0.93	**0.95**	0.85



(a) couch (b) doll (c) fruits (d) pizza (e) church (f) road

Fig. 13 Output result of error map.

is clear from Tables 4, 5 and Fig. 11 that the Proposed 2 is more effective in both visual and numerical evaluations.

4.4 Weighted Transmission Map

The output results for the error maps are shown in Fig. 13. We see from Fig. 13 that the values of the error map increase in the regions with vivid color and decrease in the regions near the sky or white areas. We calculate e_{sum} using this error map and put the weights to the depth map, as described in Sect. 3.5. Figure 14 is an example of processing when the number of iterations is set to 2. First, the number of iterations is determined, and then, the reference image is created. Next, weights are applied to the depth map in the first iteration using the reference image. In this manner, we

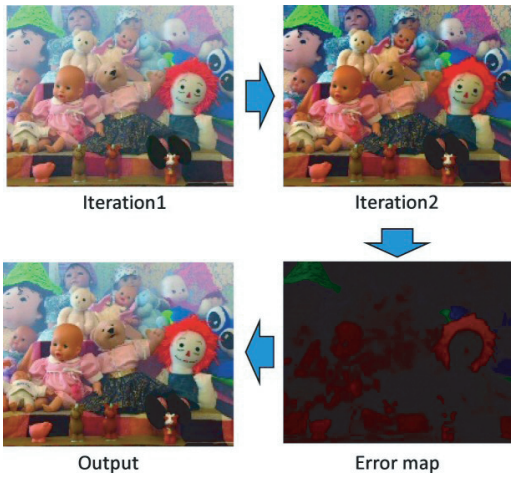


Fig. 14 Example of the haze reduction process with two iterations.

Table 6 Numerical change due to weight.

Image		w/o Weight	w/ Weight	e_{sum}
couch	SSIM	0.92	0.97	27.69
	PSNR	26.39	29.83	
doll	SSIM	0.79	0.83	19.15
	PSNR	18.03	19.55	
fruits	SSIM	0.77	0.79	30.95
	PSNR	19.52	22.07	
pizza	SSIM	0.95	0.94	21.66
	PSNR	22.75	23.71	
church	SSIM	0.96	0.98	3.88
	PSNR	23.20	28.46	
road	SSIM	0.95	0.96	3.27
	PSNR	23.54	25.67	

can apply weights only to the last iteration.

Table 6 shows the numerical results of the simulation. The weight changes in proportion to e_{sum} from (18). The weight increases for the church and road images that require region segmentation and decreases for the images that do not require it. The accuracy of the output result is improved by increasing the weight values at the last iteration for the images that have regions close to white, for which it is difficult to remove the haze. Hence, we can say that the white region, for which it is difficult to remove haze in the last iteration, is also necessary to strengthen the haze-removal processing. Overall, the weight is effective for images with large regions in which haze removal is difficult.

Tables 7–10 and Figs. 15–16 show the output images and numerical evaluation of the conventional methods and the proposed methods. The method of Ref. [10] is effective for images that do not include the sky region, but weak for images containing sky region as shown in Fig. 16(a), The method of Ref. [12] is effective for images that containing sky region as shown in Fig. 16(b), but weak for large haze such the image of fruits as shown in Fig. 15(b). The method [8] assumes that the ambient light is estimated in advance, and has a similar property with [10]. Therefore we employ the ambient light estimated by the first iteration of the proposed method. Besides, the method [9] has a property that it is

Table 7 Comparison of PSNRs (no sky images).

Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed 1	
						w/o Weight	w/ Weight
couch	18.29	12.68	16.46	23.60	17.02	26.39	**29.83**
doll	14.44	16.24	17.34	14.88	13.90	18.03	**19.55**
fruits	20.77	9.27	16.20	8.74	11.55	19.52	**22.07**
pizza	16.57	15.35	13.54	11.06	24.48	22.75	**23.71**

Table 8 Comparison of SSIMs (no sky images).

Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed 1	
						w/o Weight	w/ Weight
couch	0.75	0.70	0.64	0.92	0.66	0.92	**0.97**
doll	0.65	0.82	0.78	0.73	0.81	0.79	**0.83**
fruits	0.77	0.61	0.71	0.51	0.61	0.77	**0.79**
pizza	0.86	0.86	0.69	0.73	0.89	**0.95**	0.94

Table 9 Comparison of PSNRs (sky images).

Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed 1		Proposed 2	
						w/o Weight	w/ Weight	w/o Weight	w/ Weight
church	10.85	26.32	11.67	12.28	20.33	25.42	15.07	23.20	**28.46**
road	13.42	25.55	12.17	19.64	17.91	22.61	25.49	23.54	**25.67**

Table 10 Comparison of SSIMs (sky images).

Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed 1		Proposed 2	
						w/o Weight	w/ Weight	w/o Weight	w/ Weight
church	0.62	0.97	0.61	0.84	0.91	0.97	0.83	0.96	**0.98**
road	0.61	**0.97**	0.47	0.94	0.86	0.93	0.96	0.95	0.96

effective to dense haze images as to the proposed method, however it often dehaze too much. The method [13] usually makes too bright dehazed images. Also this method fixes the image size 500×500, therefore we first resize the given image into the size 500×500, dehaze the image and again resize the image into the original size.

On the other hand, our proposed method not using region segmentation is effective for images that do not contain sky regions. Also, the proposed method using region segmentation is effective for images including sky region. Weighted transmission map is effective for all images and these are the best results compared with other conventional methods in Table 7–Table 10.

4.5 Actual Haze Image

We further evaluated the methods using two images with actual haze, Tiananmen and Snow, shown in Fig. 17. We see from Fig. 18 that the Proposed 2 can generate a more natural sky color in the sky region than the Proposed 1 method, as shown in Figs. 18(g)–(i), while the conventional methods generate unnatural sky colors as in Figs. 18(b)–(f). Also we see from Fig. 19 that the snow region (which is close to white) is separated along with the sky region. Moreover, the Proposed 2 can express a more vivid sky color in the sky region than the Proposed 1 method can, as shown in Figs. 19(g)–(i).

From those results, the proposed 2 can be considered to have removed haze well and have made better output images. Also, the weighted transmission map is visually good as compared with the conventional methods Ref. [10] and Ref. [12]. Especially, the proposed method in Fig. 19(h) have less degradation of luminance than the conventional method in the sky region. So, this method shows best results.

4.6 Evaluation Using Benchmark Datasets

We also evaluate the performance of the proposed method quantitatively using the public datasets [20] and [21]. The dataset [20] contains relatively light haze images in indoor and outdoor environments, we tested total 100 images (50 indoor and 50 outdoor images). The dataset [21] contains dense haze images in natural environments, we tested all the 45 images in the original dataset. Each dataset contains images with sky region and those without sky region. We employed Proposed 1 (w/ Weight) for the images without sky region, and Proposed 2 (w/ Weight) for the images with sky region.

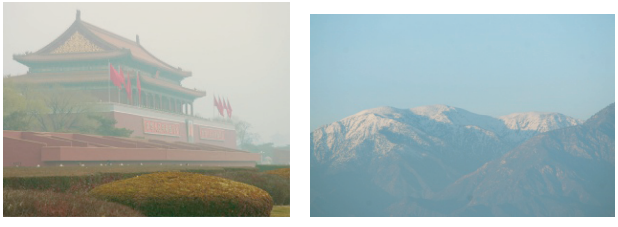
Tables 11–12 show the obtained results of PSNR and SSIM, and Figs. 20–21 show example output images. We see from Tables 11–12 and Figs. 20–21 that the method [12] gives the best performance for the dataset [20] but the proposed method gives the best performance for the dataset [21], which confirms that the proposed method is more effective

(a) Input Image (b) Ref.[10]
(c) Ref.[12] (d) Ref.[8]
(e) Ref.[9] (f) Ref.[13]
(g) Proposed 1 (w/o Weight) (h) Proposed 1 (w/ Weight)

Fig. 15 Comparison of haze reduction results (fruits).

(a) Input image (b) Ref.[10]
(c) Ref.[12] (d) Ref.[8]
(e) Ref.[9] (f) Ref.[13]
(g) Proposed 1 (w/o Weight) (h) Proposed 1 (w/ Weight)
(i) Proposed 2 (w/o Weight) (j) Proposed 2 (w/ Weight)

Fig. 16 Comparison of haze reduction results (church).



(a) Tiananmen (b) Snow

Fig. 17 Input images with actual haze.

for dense haze images.

4.7 Summary

We proposed an iterative haze-removal method and evaluated its accuracy. We also proposed a method to improve the accuracy of ambient-light estimation by segmenting the sky and white regions. As a result, the optimum number of iteration could be determined regardless of the presence or absence of the sky regions. We also found that PSNR can be further improved by weighting the transmission map used for the last iteration. Furthermore, it was found that the proposed method visually showed good results with respect to the actual haze image.

5. Discussion on Remained Haze

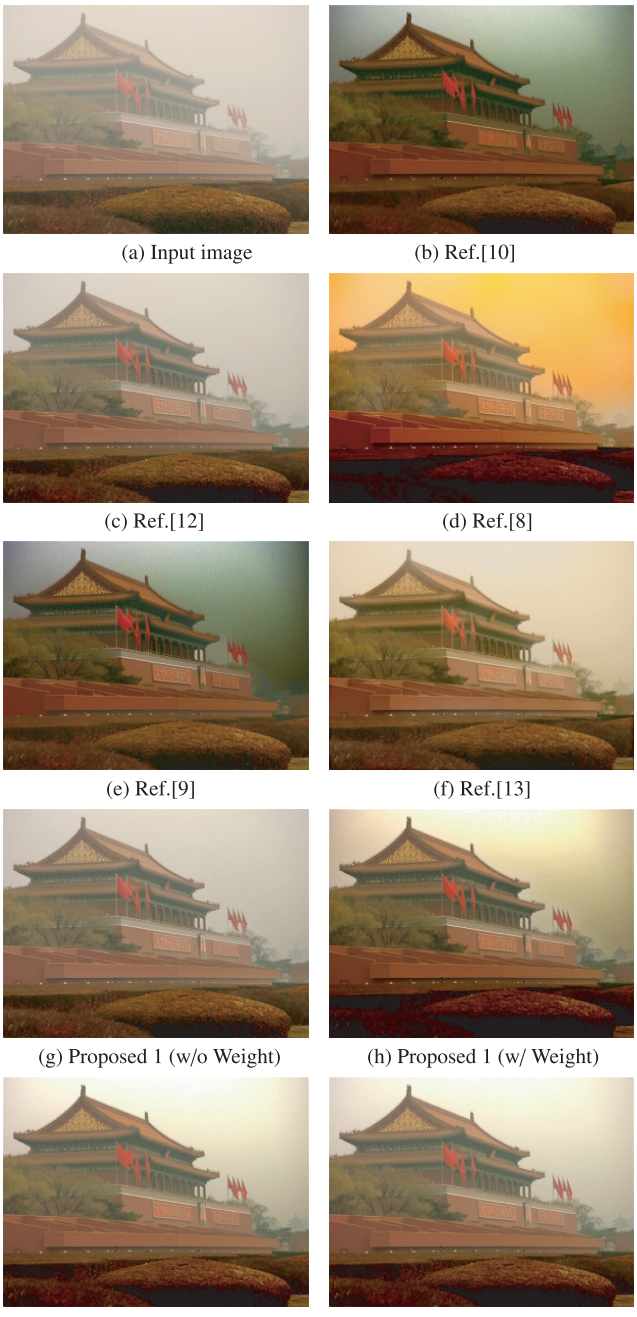
In this section, we analyze and evaluate the haze-removal quality by examining the haze that remained in the images.

5.1 Quantitative Evaluation

First we quantitatively evaluate the remaining haze in each output image. The remaining haze t_{sum} is calculated by (19) as a cumulative sum of the estimated transmission map in each pixel, i.e.

$$t_{sum} = \frac{1}{N}\sum_{x=1}^{N} t(x), \tag{19}$$

where N and $t(x)$ are the number of output-image pixels and



(a) Input image (b) Ref.[10]
(c) Ref.[12] (d) Ref.[8]
(e) Ref.[9] (f) Ref.[13]
(g) Proposed 1 (w/o Weight) (h) Proposed 1 (w/ Weight)
(i) Proposed 2 (w/o Weight) (j) Proposed 2 (w/ Weight)

Fig. 18 Comparison of haze reduction results (Tiananmen).

the transmission map estimated from the output image, respectively. The remaining haze decreases as t_{sum} approaches 1 and increases as it approaches 0.

The remaining haze t_{sum} can also be calculated from the region used in estimating the ambient light for the images that require region segmentation. This is because it is sometimes difficult to accurately estimate the amount of haze in white or sky regions.

(a) Input image (b) Ref.[10]
(c) Ref.[12] (d) Ref.[8]
(e) Ref.[9] (f) Ref.[13]
(g) Proposed 1 (w/o Weight) (h) Proposed 1 (w/ Weight)
(i) Proposed 2 (w/o Weight) (j) Proposed 2 (w/ Weight)

Fig. 19 Comparison of haze reduction results (snow).

Table 11 Comparison of Numerical Evaluation (Dataset [20] images).

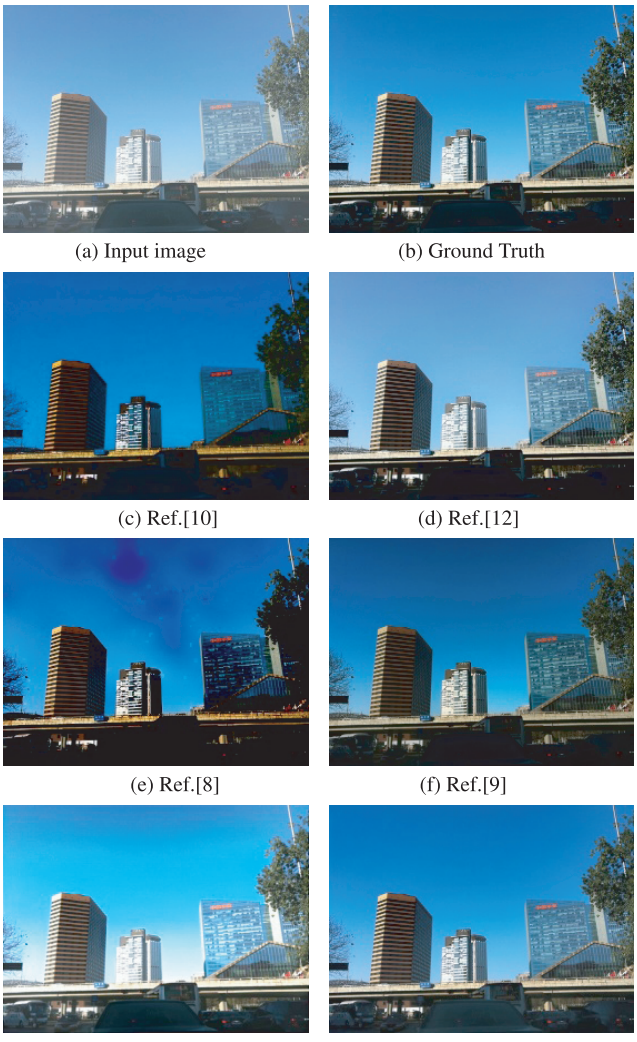
Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed
PSNR	12.37	**21.64**	13.78	12.43	17.52	18.11
SSIM	0.54	**0.88**	0.59	0.73	0.84	0.80

Table 12 Comparison of Numerical Evaluation (Dataset [21] images).

Image	Ref. [10]	Ref. [12]	Ref. [8]	Ref. [9]	Ref. [13]	Proposed
PSNR	12.09	14.27	11.60	14.86	12.81	**15.19**
SSIM	0.32	0.59	0.38	0.61	0.58	**0.62**

5.2 Remained Haze in Each Iteration

Tables 13 and 14 show the behavior of the remaining haze



(a) Input image (b) Ground Truth

(c) Ref.[10] (d) Ref.[12]

(e) Ref.[8] (f) Ref.[9]

(g) Ref.[13] (h) Proposed

Fig. 20 Comparison of haze reduction results (Dataset [20]).

(a) Input image (b) Ground Truth

(c) Ref.[10] (d) Ref.[12]

(e) Ref.[8] (f) Ref.[9]

(g) Ref.[13] (h) Proposed

Fig. 21 Comparison of haze reduction results (Dataset [21]).

t_{sum} for the artificially created haze images, and the bold numbers give the value at the moment when we terminated the iterations. We see from Table 13 that the remaining haze t_{sum} increases with the number of iterations for the images that do not require region segmentation. The PSNRs have the largest values when the iteration is terminated in bold numbers. If we continue iterations, the remaining haze t_{sum} still increases, but the PSNR decreases as we have already seen in Fig. 22. The remaining haze t_{sum} becomes almost equal to its ground truth estimated from the original image at the estimated number of iterations, which means that the estimated number of iterations is almost optimal.

The remaining haze t_{sum} of the original image exceeds 0.7 in all the tested images; this means that the value of t_{sum} under no haze exceeds 0.7. In addition, the remaining haze t_{sum} tends to decrease in images that have pixels and regions close to white, such as the pizza image, but it remained around 0.7, as in Fig. 8(a) and (d). From this, the lower limit of t_{sum} in a no-haze situation is considered to be 0.7.

However, the remaining haze t_{sum} of the original image is often largely estimated for the images with vivid colors, such as the couch image. From this, we say that the optimal value of t_{sum} is not always the same but rather depends on the color trend of the desired image. This can also be seen from Table 14, which is the result of calculating t_{sum} from the region in which the haze amount can be estimated with region segmentation. Moreover, we see from Table 14 that the remaining haze t_{sum} meets its ground truth well in cases with region segmentation; however, it does not meet it in cases without region segmentation.

Table 13 Behavior of the remained haze t_{sum} (No region segmentation).

Image	number of iteration				Ground Truth
	1	2	3	4	
doll	0.44	0.57	**0.73**	0.85	0.78
fruits	0.53	0.70	**0.82**	0.89	0.85
pizza	0.55	**0.72**	0.82	0.90	0.70
couch	0.82	**0.91**	0.95	0.97	0.92

Table 14 Behavior of the remained haze t_{sum} (Region segmentation).

Image	number of iteration				Ground Truth
	1	2	3	4	
church (Proposed 1)	0.61	0.74	0.90	**0.96**	0.82
church (Proposed 2)	0.73	**0.83**	0.90	0.95	0.82
road (Proposed 1)	0.67	**0.76**	0.83	0.88	0.84
road (Proposed 2)	0.68	0.78	**0.86**	0.92	0.84

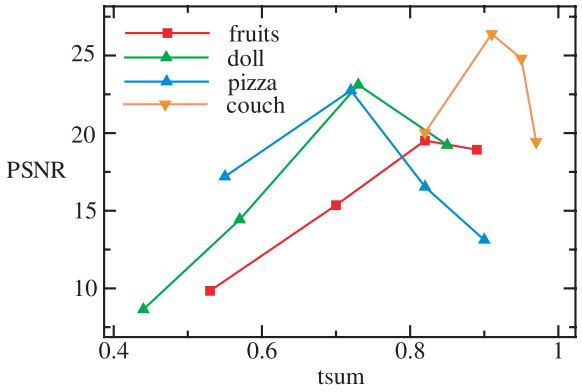


Fig. 22 Relationship between t_{sum} and PSNR.

Table 15 Behavior of the remained haze t_{sum} (Actual haze image).

Image	number of iteration			
	1	2	3	4
Tienanmen (Proposed 1)	0.84	0.88	**0.91**	0.91
Tienanmen (Proposed 2)	**0.94**	0.96	0.97	0.97
Snow (Proposed 1)	0.71	0.83	**0.91**	0.96
Snow (Proposed 2)	0.86	**0.95**	0.97	0.99

5.3 Case of Images with Actual Haze

Table 15 shows the behavior of the remaining haze t_{sum} in each iteration for images with actual haze (i.e. images without a ground truth). Again, the numbers in bold in Table 15 indicate when the iteration terminated. Similarly to Tables 13 and 14, we see from Table 15 that the remaining haze t_{sum} increases as the number of iterations increases. According to the difference of t_{sum} for each iteration, it decreases as the number of iterations increases. This can be considered as convergence, also confirmed in Tables 13 and 14. Again, the haze is well removed after the estimated number of iterations because the remaining haze t_{sum} when the iteration is terminated again exceeds 0.7 and approaches one.

6. Conclusion

We proposed an iterative haze-removal method and evaluated its accuracy, which can remove haze regardless of the size of haze. This result is very effective when removing the haze contained in the acquired image and video, as the proposed method outperformed the state-of-the-art methods based on deep learning. The proposed method is not a learning-based approach and therefore it works even when only a hazed image is given.

The drawback of the proposed method is that the method has many parameters as in Table 1, and the number of iterations in Tables 2–5 as well. The values of the parameters in Table 1 has been determined to be the same with [14], however they might not be optimum for the proposed method because the algorithm has been modified from [14]. The further optimization of those values remains as one of future studies.

References

[1] G.A. Woodell, D.J. Jobson, Z.-U. Rahman, and G. Hines, "Advanced image processing of aerial imagery," Proc. SPIE, vol.6246, p.62460E, May 2006.
[2] Y. Luo, T. Liu, D. Tao, and C. Xu, "Decomposition-based transfer distance metric learning for image classification" IEEE Trans. Image Process., vol.23, no.9, pp.3789–3801, Sept. 2014.
[3] J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning," IEEE Trans. Geosci. Remote Sens., vol.53, no.6, pp.3325–3337, June 2015.
[4] L. Liu and L. Shao, "Learning discriminative representations from RGB-D video data," Proc. Int. Joint Conf. Artif. Intell., pp.1493–1500, Aug. 2013.
[5] R.T. Tan, "Visibility in bad weather from a single image," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), no.1063-6919, June 2008.
[6] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.2995–3002, June 2014.
[7] R. Fattal, "Single image dehazing," ACM Trans. Graph., vol.27, no.3, p.72, Aug. 2008.
[8] R. Fattal, "Dehazing using color-lines," ACM Trans. Graphics, vol.34, no.1, article no.13, Dec. 2014.
[9] M. Sulami, I. Geltzer, R. Fattal, and M. Werman, "Automatic recovery of the atmospheric light in hazy images," Proc. Int. Conf. Computational Photography (ICCP), pp.1–11, May 2014.
[10] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE Trans. Pattern Anal. Mach. Intell., vol.33, no.12, pp.2341–2353, Dec. 2011.
[11] Y. Cheng, B. Chen, S. Huang, S. Kuo, A. Kopylov, O. Seredint, L. Mestetskiy, B. Vishnyakov, Y. Viziter, O. Vygolov, C. Lian, and C. Wu, "Visibility enhancement of single hazy images using hybrid dark channel prior," IEEE International Conference on Systems, Man, and Cybernetics, pp.3627–3632, Oct. 2013.
[12] B. Cai, X. Xu, K, Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," IEEE Trans. Image Process., vol.25, no.11, pp.5187–5198, Nov. 2016.
[13] H. Zhang and V.M. Patel, "Densely connected pyramid dehazing network," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), June 2018.
[14] Q. Zhu, J. Mai, and L, Shao, "A fast single image haze removal

algorithm using color attenuation prior," IEEE Trans. Image Process., vol.24, no.11, pp.3522–3533, Nov. 2015.
[15] K. He, J. Sun, and X. Tang, "Guided image filtering ", IEEE Trans. Pattern Anal. Mach. Intell., vol.35, no.6, pp.1397–1409, June 2013.
[16] H. Yuan, C. Liu, Z, Guo, and Z. Sun, "A region-wised medium transmission based image dehazing method," IEEE Access, vol.5, pp.1735–1742, Jan. 2017.
[17] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.5, pp.603–619, May 2002.
[18] G. Strang, Introduction to Linear Algebra, 4th ed., Wellesley-Cambridge Press, USA, 2009.
[19] Y. Araki and K. Ichige, "Single image haze removal using ambient light estimation and region segmentation," Proc. Int. Sympo. Intelligent Signal Processing and Communication Systems, pp.70–75, Nov. 2018.
[20] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single image dehazing and beyond," IEEE Trans. Image Process., vol.28, no.1, pp.492–505, Jan. 2019.
[21] C.O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images," Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), June 2018.

Yuji Araki received B.E. and M.E. degrees in Electrical and Computer Engineering from Yokohama National University in 2017 and 2019, respectively. Currently he is with NEC Corporation. During his masters' study, he engaged with the research of digital image processing techniques like restoration or denoising. He received Best Paper Award in Int. Sympo. Intelligent Signal Processing and Communication Systems (ISPACS) in 2018.

Kentaro Mita received B.E. and M.E. degrees in Electrical and Computer Engineering from Yokohama National University in 2018 and 2020, respectively. Currently he is with Yokogawa Electric Corporation. During his masters' study, he engaged with the research of machine learning approach to digital image processing techniques like restoration or separation.

Koichi Ichige received B.E., M.E. and Dr. Eng. degrees in Electronics and Computer Engineering from the University of Tsukuba in 1994, 1996 and 1999, respectively. He joined the Department of Electrical and Computer Engineering, Yokohama National University as a research associate in 1999, where he is currently a professor. He has been on leave to Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland as a visiting researcher in 2001–2002. His research interests include digital signal processing, approximation theory and their applications to image processing and mobile communication. He served as an associate editor of IEEE Transactions on Industrial Electronics in 2004–2008, Journal of Circuits, Systems and Computers (JCSC) in 2012–2014, and IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (IEICE-EA) in 2015–2018. Currently he serves as an editor of IEICE-EA. He received "Meritorious Award on Radio" from the Association of Radio Industries and Businesses (ARIB) in 2006, Best Letter Award from IEICE Communication Society in 2007, and Best Paper Award in Int. Sympo. Intelligent Signal Processing and Communication Systems (ISPACS) in 2018. He is a member of IEEE.