

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 1

Near-Infrared Fusion via Color Regularization for Haze and Color Distortion Removals

Chang-Hwan Son and Xiao-Ping Zhang, *Senior Member, IEEE*

Abstract— Different from conventional haze removal methods based on a single image, near-infrared imaging can provide two types of multimodal images: One is the near-infrared image and the other is the visible color image. These two images have different characteristics regarding color and visibility. The captured near-infrared image is haze-free but it is grayscale, whereas the visible color image has colors but it contains haze. There are serious discrepancies in terms of brightness and image structures between the near-infrared image and the visible color image. Due to this discrepancy, the direct use of the near-infrared image for haze removal causes a color distortion problem during near-infrared fusion. The key objective for the near-infrared fusion is therefore to remove the color distortion as well as the haze. To achieve this objective, this paper presents a new near-infrared fusion model that combines the proposed new color and depth regularizations with the conventional haze degradation model. The proposed color regularization sets the color range of the unknown haze-free image based on the combination of the two colors of the colorized near-infrared image and the captured visible color image. That is, the proposed color regularization can provide color information for the unknown haze-free color image. The new depth regularization enables the consecutively estimated depth maps not to be largely deviated, thereby transferring natural-looking colors and high visibility of the colorized near-infrared image into the preliminary dehazed version of the captured visible color image with color distortion and edge artifacts. Experimental results show that the proposed color and depth regularizations can help remove the color distortion and the haze simultaneously. The effectiveness of the proposed color regularization for near-infrared fusion is verified by comparing it with other conventional regularizations.

Index Terms—Coloring, color correction, depth, haze removal, regularization, multi-modal fusion, near-infrared imaging

I. INTRODUCTION

According to Rayleigh physical scattering model [1], the intensity of the scattered light by the particles (e.g., dust or mist) in the atmosphere is inversely proportional to the wavelength of the incident light. Near-infrared lights have relatively long wavelengths between 700 nm and 1100 nm, compared to the visible lights whose wavelength ranges lie

This work was supported in part by the National Research Foundation of Korea (2017R1D1A3B03030853) and in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN239031.

Chang-Hwan Son is with the Department of Software Convergence Engineering, Kunsan National University, South Korea (E-mail: changhwan76.son@gmail.com)

Xiao-Ping Zhang is with the Department of Electrical and Computer Engineering, Ryerson University, Canada (E-mail: xzhang@ryerson.ca).

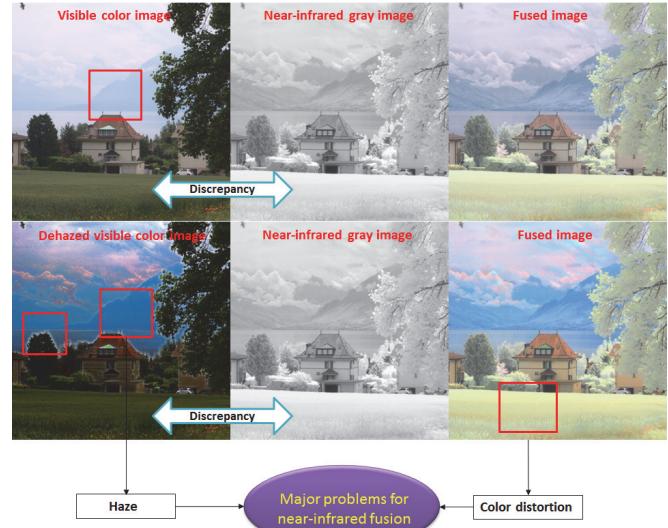


Fig. 1. An example of captured visible color image, near-infrared gray image, and fused image are shown in first row from left to right. The second row illustrates fused image by using the dehazed version of the visible color image and the near-infrared gray image in the first and second image from left, and the main issue for the near-infrared fusion in the last row.

between 400 nm and 700 nm. As particles become denser, the visible lights get more scattered. These scattered lights called *airlight* [2] are blended with the reflected lights from the objects in the scene, thus degrading the visibility, contrast, or color fidelity of the captured visible color images. However, the near-infrared lights are less sensitive to the atmospheric scattering, in other words, near-infrared lights can reach the image sensors directly, and thus *haze-free* near-infrared gray images can be acquired. In addition, useful features for saliency detection, tracking, and shadow removal can be extracted from the near-infrared images [3-5].

Based on this scattering property, near-infrared imaging, which captures the visible color image and the corresponding near-infrared gray image consecutively, was introduced in [6]. In this paper, two filters each of which either passes or blocks the near-infrared lights were used to capture both visible color image and near-infrared image [6,7]. In the first row of Fig. 1, an example of the captured visible color image and near-infrared gray image is provided. In the visible color image, the contrast and details are almost lost in the haze region of the mountain and the sky, whereas the near-infrared gray image is haze-free, due to the fact that near-infrared lights scatter less. However, the captured near-infrared gray image looks unnatural, especially in the grass and the tree regions. This is

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

because molecular structures cause the vegetation to 'glow' when viewed under the near-infrared lights. Also, the intensity of the near-infrared lights reflected by plants varies depending on the type of the plant; please refer to [6] for more details.

A. Main issue for near-infrared fusion

As shown in Fig. 1, *haze-free* near-infrared gray images can be captured via near-infrared imaging. Thus, for haze removal, a naive fusion can be considered to combine the chrominance planes of the visible color image with the near-infrared gray image in an opponent color space (e.g., L*a*b* [8] and decorrelated color space [9]). However, there are discrepancies in the brightness and image structures between the visible color image and near-infrared gray image, thereby resulting in unnatural-looking colors (i.e., color distortion) on the fused image, as shown in the rightmost image of Fig. 1. More sophisticated fusion methods based on multiresolution representation [10, 11] and principal component analysis [6] can be adopted. However, they all lead to a color distortion problem. Another approach is using single image dehazing methods [12-16]. Given only the visible color image, some constraints such as dark prior [12], total variation [13,14], gray world assumption [15], and color attenuation prior [16] can be used to remove the haze. However, the quality of the dehazed images is not satisfactory so that the haze still exists, or dehazed colors are distorted [12,15,16]. An example of the dehazed version of the visible color image with the dark prior method [12] is provided in the second row of Fig. 1 where haze and edge artifacts are shown in the red box. There is also some discrepancy between the dehazed image and the near-infrared gray image, and thus the color distortion problem cannot be avoided in the fused image.

In short, the new issue for the near-infrared fusion is to preserve high visibility of the captured near-infrared image and to remove color distortion that appears during image fusion. The color distortion can also be generated through the haze removal of the captured visible color image [12,15,16]. While fusing the haze-free near-infrared gray image with the captured visible color image (or its dehazed version), the color distortion and haze should be removed simultaneously. Therefore, throughout this paper, a new near-infrared fusion model is presented to remove the haze as well as the color distortion on the captured visible color image.

B. Proposed approach

To handle the main issue related to the near-infrared fusion, this paper introduces a new near-infrared fusion model based on the proposed color and depth regularizations. The concept of the proposed color regularization is simple. If we can generate unknown colors of the captured near-infrared gray image, unknown haze-free image can have color information (or color prior) from the colorized near-infrared image. Also, since the captured near-infrared gray image is haze-free, high visibility and details of the colorized near-infrared image can be utilized to remove the haze on the captured visible color image. As shown in Fig. 2, to model the proposed color regularization, we first generate unknown colors of the captured near-infrared

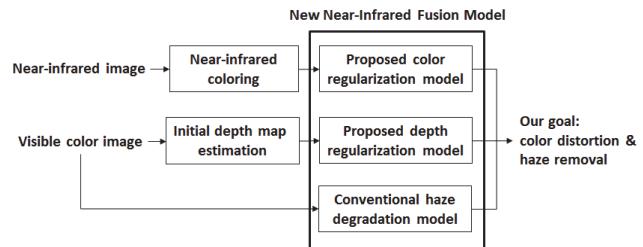


Fig. 2. The proposed approach based on color and depth regularizations.

gray image via near-infrared coloring. Then, we combine a new color regularization with the conventional haze degradation model to provide the color prior for the unknown haze-free color image. However, it should be noted that not only the created near-infrared color image should be natural-looking without color distortion, but it should also preserve fine details and high visibility of the captured near-infrared gray image. The newly added regularization term can provide the color prior for the unknown haze-free image, so we call it *color regularization*.

The proposed depth regularization transfers the natural-looking colors and high visibility of the colorized near-infrared image into the initial dehazed version of the captured visible color image. Different from the conventional depth regularization in [13,14] where smoothness constraint between pixels in the depth map is used, the depth difference constraint that forces the consecutively estimated depth maps not to be largely deviated is adopted in this paper. By combining the proposed color and depth regularizations with the conventional haze degradation model, our goal of removing the color distortion and the haze simultaneously can be achieved for the near-infrared fusion.

C. Our contribution

- This paper is the first work that attempts to address the color distortion problem by combining a new color regularization with the conventional haze degradation model for near-infrared fusion. As shown in Fig. 1, there are two causes that generate color distortion during near-infrared fusion. One is that the used dark prior model [12] is not appropriate for the estimation of unknown colors of the haze-free image. The other is that the discrepancy problem between the captured visible color image and the near-infrared gray image leads to color distortion during image fusion. To avoid this color distortion that comes from these two causes, a color constraint is introduced in this paper to limit the color range of the unknown haze-free image. Although the color attenuation prior [16] and the gray world assumption [15] are used to improve the colors of the dehazed images, the color distortion that comes from the discrepancy problem cannot be removed. In addition, single image dehazing methods [12, 15, 16] have limitations in removing the haze completely, although dark and color priors are used. Therefore, throughout this paper, it is shown how effectively the proposed color regularization can simultaneously remove color distortion and haze. It is also shown that the color quality level of the proposed method is

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

much better than conventional single image dehazing, image fusion, and near-infrared dehazing.

TABLE I. NOTATION

Notation	Definition
\mathbf{x}^{nir}	Captured near-infrared gray image
$\mathbf{x}^{v(c)}$	Captured visible color image (or captured haze image)
$\mathbf{x}^{o(c)}$	Colorized near-infrared image
$\mathbf{x}^{s(c)}$	Unknown haze-free color image
\mathbf{x}^a	Atmospheric color light
$\mathbf{u}^{d(c)}$	Depth map
$\mathbf{u}^{v(c)} = \ln(\mathbf{x}^{v(c)} - \mathbf{x}^a)$	Another expression of the captured visible color image in a log domain
$\mathbf{u}^{o(c)} = \ln(\mathbf{x}^{o(c)} - \mathbf{x}^a)$	Another expression of the colorized near-infrared image in a log domain
$\mathbf{u}^{s(c)} = \ln(\mathbf{x}^{s(c)} - \mathbf{x}^a)$	Another expression of the unknown haze-free color image in a log domain

II. NOTATION

In this paper, bold lowercase is used to indicate column vectors. For example, \mathbf{x}^v and \mathbf{x}^{nir} indicate the column vectors that contain the pixel values of the captured visible color image and near-infrared gray image, respectively. The superscripts are used to differentiate between the two images. If the captured images are not grayscale, i.e., there are color channels, the superscript c is additionally used to indicate the color channel such as $\mathbf{x}^{v(c)}$. To indicate vector elements (i.e., pixels), the subscript i is used like \mathbf{x}_i^v . Table I summarizes the used notation throughout this paper.

III. HAZE DEGRADATION MODEL

We first look at the haze degradation model [13], which can help us to understand how the haze in the images is formed. The haze degradation model is given by

$$\mathbf{x}_i^{v(c)} = k_i \mathbf{x}_i^{s(c)} + (1 - k_i) \mathbf{x}_i^{a(c)} \quad (1)$$

where \mathbf{x}_i^a is the atmospheric color light and it is a column vector with the size of 3×1 . Note that \mathbf{x}_i^a is a fixed column vector independent of the pixel location (i), that is, $\mathbf{x}_i^a = \mathbf{x}^a$. k_i is a scalar value indicating the transmission, which describes the portion of the light that is not scattered and reaches the camera. Above equation indicates that the i th pixel value of the visible color image (or haze image) $\mathbf{x}_i^{v(c)}$ is given by blending the i th pixel value of the haze-free image $\mathbf{x}_i^{s(c)}$ with the atmospheric color light $\mathbf{x}_i^{a(c)}$, according to the transmission k_i . When the atmosphere is homogenous, the transmission k_i is expressed by $k_i = e^{-\eta d_i}$, where η is the medium extinction coefficient and d is the depth of scene. Here, it is noticed that the transmission k_i is related with the scene depth d_i .

For simplicity, (1) can be rewritten, as follows [13].

$$\mathbf{x}_i^{v(c)} - \mathbf{x}_i^{a(c)} = k_i (\mathbf{x}_i^{s(c)} - \mathbf{x}_i^{a(c)}) \quad (2)$$

$$\ln(\mathbf{x}_i^{v(c)} - \mathbf{x}_i^{a(c)}) = \ln k_i + \ln(\mathbf{x}_i^{s(c)} - \mathbf{x}_i^{a(c)}) \quad (3)$$

where \ln indicates the natural logarithm and it is used to avoid the product term. (3) is further simplified as

$$\mathbf{u}_i^{s(c)} = \mathbf{u}_i^d + \mathbf{u}_i^{v(c)} \quad (4)$$

where $\mathbf{u}_i^{s(c)}$ and $\mathbf{u}_i^{v(c)}$ are defined as $\ln(\mathbf{x}_i^{s(c)} - \mathbf{x}_i^{a(c)})$ and $\ln(\mathbf{x}_i^{v(c)} - \mathbf{x}_i^{a(c)})$, respectively, and \mathbf{u}_i^d is equal to $-\ln k_i$. The vector form of (4) is expressed as follows.

$$\mathbf{u}^{s(c)} = \mathbf{u}^{d(c)} + \mathbf{u}^{v(c)} \quad (5)$$

where $\mathbf{u}^{s(c)}$, $\mathbf{u}^{v(c)}$, and $\mathbf{u}^{d(c)}$ are the column vectors with the size of $N \times 1$. Here N is the image size. $\mathbf{u}^{d(c)}$ has the pixel value u_i^d at the i th pixel position and it is identical irrespective of color channel. Above equation indicates that the unknown haze-free color image $\mathbf{u}^{s(c)}$ is the addition of the depth map $\mathbf{u}^{d(c)}$ and the captured haze image (or visible color image) $\mathbf{u}^{v(c)}$. In fact, $\mathbf{u}^{v(c)}$ is not the captured visible color image $\mathbf{x}^{v(c)}$, however $\mathbf{u}^{v(c)}$ can be regarded as another expression of the visible color image in a log domain, and thus we will call $\mathbf{u}^{v(c)}$ the visible color image. Similarly, $\mathbf{u}^{s(c)}$ and $\mathbf{u}^{o(c)}$ will be called by the unknown haze-free image and the colorized near-infrared image, respectively. More details about the haze degradation model are provided in [13].

IV. PROPOSED NEAR-INFRARED FUSION

To avoid the color distortion that comes from the two causes, as mentioned in the Introduction, a color constraint is needed to limit the color range of the unknown haze-free image. There is no way of precisely knowing the original colors of the haze-free image given only the haze image, due to the ill-posed problem. Moreover, the colors of the original scenes without the haze depend on illumination and weather conditions. The only way to limit the range of the original colors of the haze image is using the two captured images: the visible color image and the near-infrared gray image. In this paper, the original colors of the haze-free image are limited by the color range, which is created by a combination of the two colors of the colorized near-infrared image and the visible color image.

In this paper, a contrast preserving mapping model is used to estimate the colors of the near-infrared gray image [17]. The contrast preserving mapping model finds the linear mapping relation between the near-infrared gray image and the luminance plane of the visible color image. Therefore, the colors of the near-infrared gray images can be generated, according to the linear mapping relation. Although there is a serious discrepancy in the brightness between the near-infrared gray image and the luminance plane of the visible color image, the colors of the near-infrared gray image can be induced from

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 4

the visible color image. In other words, if the near-infrared gray image is converted into the luminance plane of the visible color image via the linear mapping relation, the colors of the colorized near-infrared gray image become close to the colors of the visible color image. Therefore, we can say that the colors of the near-infrared gray image can be used to set the range of the original colors by combining the two colors of the colorized near-infrared image and the visible color image.

Assuming that the colors of the captured near-infrared gray image are generated via near-infrared coloring [17], the proposed near-infrared fusion for haze and color distortion removals is modeled as follows:

$$\begin{aligned} & \min_{\mathbf{u}_{t+1}^{s(c)}, \mathbf{u}_{t+1}^{d(c)}} \underbrace{\lambda_1 \|\mathbf{u}_{t+1}^{s(c)} - (\mathbf{u}^{v(c)} + \mathbf{u}_{t+1}^{d(c)})\|_2^2}_{\text{Haze Degradation}} + \underbrace{\lambda_2 \|\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{o(c)}\|_2^2}_{\text{Color Regularization}} \\ & + \sum_{j=1}^2 \|\mathbf{f}^j \oplus \mathbf{u}_{t+1}^{s(c)}\|_1 + \underbrace{\lambda_3 \|\mathbf{u}_{t+1}^{d(c)} - \mathbf{u}_t^{d(c)}\|_2^2}_{\text{Depth Regularization}} \end{aligned} \quad (6)$$

$$\begin{aligned} & = \min_{\mathbf{u}_{t+1}^{s(c)}, \mathbf{u}_{t+1}^{m(c)}} \underbrace{\lambda_1 \|\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}_{t+1}^{m(c)}\|_2^2}_{\text{Color Regularization}} + \underbrace{\lambda_2 \|\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{o(c)}\|_2^2}_{\text{Color Regularization}} \\ & + \sum_{j=1}^2 \|\mathbf{f}^j \oplus \mathbf{u}_{t+1}^{s(c)}\|_1 + \lambda_3 \|\mathbf{u}_{t+1}^{m(c)} - \mathbf{u}_t^{m(c)}\|_2^2, \text{ where } \mathbf{u}_{t+1}^{m(c)} = \mathbf{u}^{v(c)} + \mathbf{u}_{t+1}^{d(c)} \end{aligned} \quad (7)$$

The vectors $\mathbf{u}^{s(c)}$, $\mathbf{u}^{v(c)}$, and $\mathbf{u}^{d(c)}$ are the same as the ones defined in (5), but the two vectors $\mathbf{u}^{s(c)}$ and $\mathbf{u}^{d(c)}$ are iteratively updated with an iteration number t . The vector $\mathbf{u}^{o(c)}$ indicates the colorized near-infrared image, as defined in Table I, and it is fixed through the iteration. The symbol \oplus denotes convolution operator and \mathbf{f}^j indicates the vertical and horizontal gradient filters, which are defined as $[1 \ -1]$ and its transposed version, respectively. λ is a penalty parameter.

In (6), the first term indicates the conventional haze degradation model, as defined in (5). This term imposes a constraint that the unknown haze-free color image ($\mathbf{u}^{s(c)}$) is the addition of the captured haze image ($\mathbf{u}^{v(c)}$) and the estimated depth map ($\mathbf{u}^{d(c)}$). The second term is the proposed color regularization. If the unknown colors of the near-infrared gray image can be generated via near-infrared coloring [17], the colors of the unknown haze-free image can be in the color range created by the combination of the two colors of the colorized near-infrared image and the dehazed version of the visible color image, which is modeled by the proposed color regularization term and the haze degradation term, respectively, as shown in (6). In other words, the colors of the haze-free image $\mathbf{u}^{s(c)}$ can exist between the colors of the colorized near-infrared image $\mathbf{u}^{o(c)}$ and the colors of the dehazed version of the visible color image $\mathbf{u}^{v(c)} + \mathbf{u}^{d(c)}$. The third term in (6) is the gradient regularization. It is well-known that the gradient distribution of haze-free images can be modeled by Laplacian or hyper-Laplacian probability distributions in a logarithm domain [18]. The total variation norm in third term can reflect the statistical gradient distribution of haze-free images. The last term in (6) is another proposed depth regularization which

ensures that the currently estimated depth map should not be largely deviated from the previously estimated depth map. Different from the conventional depth regularization [13,14] where smoothness constraint between neighborhood pixels in the depth map is used, the proposed dehazing model adopts the depth difference constraint between the consecutively estimated depth maps. The role of the proposed depth difference regularization in (6) is different from that of the conventional depth regularization, which means that we cannot say that the proposed depth difference regularization is better than the conventional depth regularization. The conventional depth regularization is used to smoothen abrupt gradients on the estimated depth map according to the observed depth prior, which indicates that the depth map has piecewise constant in a local region. The conventional depth regularization is necessary for single image-based dehazing. This is because the dark prior method can produce block artifacts on the estimated depth map. However, in this paper, image pair is used for near-infrared fusion. Therefore, the proposed depth difference regularization is necessary to transfer the natural-looking colors and high visibility of the colorized near-infrared image into the initial dehazed image; This effect will be checked later.

Meanwhile, the proposed model of (6) is equally expressed by (7) by replacing $\mathbf{u}_{t+1}^{d(c)} + \mathbf{u}^{v(c)}$ by $\mathbf{u}_{t+1}^{m(c)}$. Here, $\mathbf{u}_{t+1}^{m(c)}$ indicates the dehazed version of the captured visible color image (or hazy image). Therefore, (7) shows that the proposed depth regularization is equal to the image difference constraint using the consecutively dehazed images.

A. Advantages of the proposed color regularization

- If λ_2 is set with zero, (6) becomes the single image dehazing model. In this case, the only available information is the captured haze image ($\mathbf{u}^{v(c)}$). The prior information is the gradient distribution, as shown in the third term [18]. Thus, it is hard to estimate the unknown haze-free color image, atmospheric color light, and depth map at the same time. Even though dark channel prior [12] can be used to estimate depth map and atmospheric color light, in most cases, it fails to provide satisfactory results. More specifically, the color distortion appears on the dehazed images. However, the proposed color regularization can provide color prior for the unknown haze-free color image, thereby leading to reach a good solution.
- Instead of the proposed color regularization, other regularizations can be considered. For image-pair-based restoration, gradient difference regularization has been widely used [19,20]. Recently, this regularization was also used for the near-infrared fusion [21]. The gradient difference regularization is defined as $\sum_{j=1}^2 \|\mathbf{f}^j \oplus \mathbf{u}_{t+1}^{s(c)} - \mathbf{f}^j \oplus \mathbf{u}_t^{nir}\|_1$

where \mathbf{u}_t^{nir} is given by $\ln(\mathbf{x}_t^{nir} - \mathbf{x}_t^a)$, according to the haze degradation model in (5). This gradient difference regularization forces the gradients of the unknown haze-free color image to be similar to those of the captured near-infrared gray image, thereby increasing visibility on the dehazed images. However, the used near-infrared image

\mathbf{u}^{nir} is grayscale, and thus color information cannot be provided. This means that the color distortion can appear on the dehazed images. In contrast, the proposed color regularization can provide color and gradient information, which leads to improvement in visual color appearance. In summary, different from the conventional gradient [18] and gradient difference regularization [21] which attempt to model gradient distribution for the unknown haze-free color image, the proposed color regularization models a color constraint for the unknown haze-free color image. This proposed color regularization is the first attempt to solve the color distortion that comes from the discrepancy problem and the inaccurate color estimation of the initial dehazing [12] while removing the haze on the captured visible color image.

V. IMPLEMENTATION

The equation (6) or (7) can be minimized by using variable splitting technique in [22] as follows,

$$\begin{aligned} \min_{\mathbf{u}_{t+1}^{s(c)}, \mathbf{u}_{t+1}^{d(c)}, \mathbf{w}} & \lambda_1 \left\| \mathbf{u}_{t+1}^{s(c)} - (\mathbf{u}^{v(c)} + \mathbf{u}_{t+1}^{d(c)}) \right\|_2^2 + \underbrace{\lambda_2 \left\| \mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{o(c)} \right\|_2^2}_{\text{Color Regularization}} \\ & + \sum_{j=1}^2 \left\| \mathbf{w}^j \right\|_1 + \frac{\beta}{2} \sum_{j=1}^2 \left\| \mathbf{w}^j - \mathbf{f}^j \oplus \mathbf{u}_{t+1}^{s(c)} \right\|_2^2 + \underbrace{\lambda_3 \left\| \mathbf{u}_{t+1}^{d(c)} - \mathbf{u}_t^{d(c)} \right\|_2^2}_{\text{Depth Regularization}} \end{aligned} \quad (8)$$

where \mathbf{w} is an auxiliary variable used to optimize (6). In (8), if the value of the parameter β approaches to infinity, the solution of (8) converges to that of (6). The proposed model in (8) has the same form as (1.3) in [22] except that there is another variable ($\mathbf{u}_{t+1}^{d(c)}$) to be estimated. Therefore, (8) can be solved in a similar way, as provided in [22].

A. Initializations

To solve (8), we first need to initialize three vectors: depth map $\mathbf{u}_{t=0}^{d(c)}$, atmospheric color light \mathbf{x}^a , and colorized near-infrared image $\mathbf{u}^{o(c)}$. First, to provide the $\mathbf{u}_{t=0}^{d(c)}$ and \mathbf{x}^a , dark prior [12,13] is utilized. The dark prior is a kind of statistics on the haze-free outdoor images. In most non-sky patches of the haze-free outdoor images, at least one color channel has very low intensity value. Based on this observation, the dark channel is defined as

$$\mathbf{x}_i^{dark} = \min_c \min_{j \in \Omega_i} \mathbf{x}_j^{s(c)} \approx 0 \quad (9)$$

where j indicates the neighborhood pixel index that belongs to a local patch Ω centered at pixel index i . The above equation indicates that the intensity values of the dark channel vector \mathbf{x}_i^{dark} are distributed around zero. Based on this dark prior, the transmission k_i can be derived, as follows,

$$k_i = 1 - \min_c \min_{j \in \Omega_i} \left(\frac{\mathbf{x}_j^{v(c)}}{\mathbf{x}_i^{a(c)}} \right) \quad (10)$$

where the atmospheric color light \mathbf{x}^a is given by picking and averaging the top 0.1% brightest pixels in \mathbf{x}_i^{dark} ; please refer to [13,12] for more details about the transmission and the atmospheric color light. Given the transmission k_i , the initial depth map $\mathbf{u}_{t=0}^{d(c)}$ can be calculated based on the relation between the transmission and scene depth, i.e., $\mathbf{u}_{t=0}^{d(c)} = -\ln k_i$. However, as discussed in [13,12], the transmission k_i is assumed to be a constant in a local patch, and thus this leads to the edge artifacts on initial dehazed images. To remove this edge artifacts, alpha matting is needed [23], which is computationally complex, as indicated in [13]. Therefore, in this paper, the alpha matting is excluded during the dark prior method [12] for fast initial dehazing.

Second, to set the $\mathbf{u}^{o(c)}$, near-infrared coloring [17] is needed. As mentioned in Introduction, the near-infrared coloring should generate natural-looking colors and preserve high visibility of the captured near-infrared gray image. To satisfy this, we used recently introduced near-infrared coloring method based on the contrast-preserving mapping model [17]. Given the colorized near-infrared image $\mathbf{x}^{o(c)}$, $\mathbf{u}^{o(c)}$ is calculated, according to the definition in Table I.

B. Numerical solution

Equation (8) is well-known as total variation problem. In this paper, alternating minimization technique, also known as block coordinate descent, is adopted [22,24-26]. In this paper, $\mathbf{u}_{t+1}^{s(c)}$ is first estimated with the auxiliary variable \mathbf{w} , and then $\mathbf{u}_{t+1}^{d(c)}$ is estimated as follows,

$$\begin{aligned} \min_{\mathbf{u}_{t+1}^{s(c)}, \mathbf{w}} & \frac{\lambda_0}{2} \left\{ w_1 \left\| \mathbf{u}_{t+1}^{s(c)} - (\mathbf{u}^{v(c)} + \mathbf{u}_t^{d(c)}) \right\|_2^2 + w_2 \left\| \mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{o(c)} \right\|_2^2 \right\} \\ & + \sum_{j=1}^2 \left\| \mathbf{w}^j \right\|_1 + \frac{\beta}{2} \sum_{j=1}^2 \left\| \mathbf{w}^j - \mathbf{f}^j \oplus \mathbf{u}_{t+1}^{s(c)} \right\|_2^2 \end{aligned} \quad (11)$$

where $\lambda_0 w_1 / 2$ and $\lambda_0 w_2 / 2$ correspond to λ_1 and λ_2 in (8), respectively. w_1 and w_2 are weighting values and their sum should be equal to one. In (11), there are two unknown vectors $\mathbf{u}_{t+1}^{s(c)}$ and \mathbf{w} , which are iteratively updated via alternating minimization. In other words, given a fixed \mathbf{w} , $\mathbf{u}_{t+1}^{s(c)}$ is minimized with a close-form solution, and then \mathbf{w} is minimized with a shrinkage operation for a fixed $\mathbf{u}_{t+1}^{s(c)}$. This processing continues until a stop criterion is satisfied. During the iteration, the penalty parameter β should be increased for convergence. Please refer to [22,24] for more details.

It is worth noting that the minimization of (11) fuses the two images ($\mathbf{u}^{v(c)} + \mathbf{u}_t^{d(c)}$ and $\mathbf{u}^{o(c)}$), thereby resulting in $\mathbf{u}_{t+1}^{s(c)}$. Therefore, the minimization of (11) can be regarded as the multimodal image fusion [27-29]. However, compared to the conventional multimodal image fusion, in this paper, the haze degradation model is used for the near-infrared fusion. In addition, as shown in [27], the guided image filtering [29,30] cannot remove the color distortion that comes from the discrepancy problem.

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

After minimizing (11), $\mathbf{u}_{t+1}^{d(c)}$ is then updated with a closed-form solution:

$$\mathbf{u}_{t+1}^{d(c)} = \frac{1}{\lambda_3 + 1} \{(\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{v(c)}) + \lambda_3 \mathbf{u}_t^{d(c)}\} \quad (12)$$

This equation shows that the depth map $\mathbf{u}_{t+1}^{d(c)}$ to be updated is the weighted average of the previous depth map $\mathbf{u}_t^{d(c)}$ and the currently estimated depth map $(\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{v(c)})$, as defined in (5). If $\mathbf{u}_{t+1}^{d(c)}$ is replaced by $\mathbf{u}_{t+1}^{m(c)}$, as shown in (7), the closed-form solution of (12) is changed like this,

$$\mathbf{u}_{t+1}^{m(c)} = \frac{1}{\lambda_3 + 1} \{\mathbf{u}_{t+1}^{s(c)} + \lambda_3 \mathbf{u}_t^{m(c)}\} \quad (13)$$

This equation indicates that the dehazed version ($\mathbf{u}_{t+1}^{m(c)}$) of the captured visible color image (or haze image) is also updated via the weighted average of the estimated haze-free image $\mathbf{u}_{t+1}^{s(c)}$ and the dehazed image $\mathbf{u}_t^{m(c)}$.

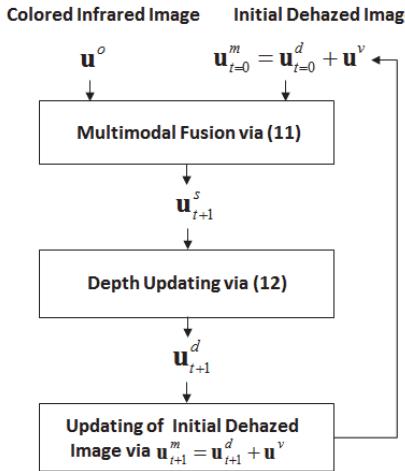


Fig. 3. Block-diagram of the proposed method.

To understand how the proposed method using (11) and (12) or (11) and (13) can update the initial dehazed image (or initial depth map), the block-diagram is provided in Fig. 3. Given the initial depth map $\mathbf{u}_{t=0}^{d(c)}$ with the dark prior [12], initial dehazed image $\mathbf{u}_{t=0}^{m(c)}$ is obtained by the addition of the initial depth map $\mathbf{u}_{t=0}^{d(c)}$ and the captured visible color image $\mathbf{u}^{v(c)}$, according to the haze degradation model in (5). The initial dehazed image $\mathbf{u}_{t=0}^{m(c)}$ has poor quality. More specifically, edge artifact and color distortion appear on the initial dehazed image. To improve the initial dehazed image, the colorized near-infrared image $\mathbf{u}^{o(c)}$ is used during the multimodal fusion. The colorized near-infrared image $\mathbf{u}^{o(c)}$ is haze-free and it has also natural-looking colors. Therefore, the initial dehazed image ($\mathbf{u}_{t=0}^{m(c)}$) can be improved via the multimodal image fusion in (11), producing the haze-free color image $\mathbf{u}_{t+1}^{s(c)}$. Then, the initial

depth map is updated via (12), which utilizes the estimated haze-free image $\mathbf{u}_{t+1}^{s(c)}$. Given the updated depth map $\mathbf{u}_{t+1}^{d(c)}$, the initial dehazed image $\mathbf{u}_{t=0}^{m(c)}$ is updated by the addition of the updated depth map $\mathbf{u}_{t+1}^{d(c)}$ and the visible color image $\mathbf{u}^{v(c)}$. By doing this iteratively, high visibility and natural-looking colors of the colorized near-infrared image can be transferred into the initial dehazed image $\mathbf{u}_{t=0}^{m(c)}$ (or initial depth map $\mathbf{u}_{t=0}^{d(c)}$). As a result, the edge and color artifacts on the initial dehazed image can be removed simultaneously. From this block-diagram, it is concluded that the role of the proposed color regularization is to provide the color prior for the unknown haze-free image and the role of the proposed depth regularization is to transfer high visibility and natural-looking colors of the estimated haze-free images $\mathbf{u}_{t+1}^{s(c)}$ into the initial dehazed image $\mathbf{u}_{t=0}^{m(c)}$. This is the key reason to success of our method.

The proposed near-infrared fusion method is summarized in Algorithm I. Given the colorized near-infrared image \mathbf{x}^o , atmospheric color light \mathbf{x}^a , and transmission k_i , the unknown vectors $\mathbf{u}_{t+1}^{s(c)}$ and $\mathbf{u}_{t+1}^{d(c)}$ are iteratively updated for each color channel c . After reaching the maximum iteration of the inner loop t_{max} , the dehazed image $\mathbf{x}^{s(c)}$ for the channel c is obtained by manipulating $\mathbf{u}^{s(c)}$. When outer loop finishes, the final dehazed color image \mathbf{x}^s is obtained. In Algorithm I, c_{max} and t_{max} are given by 3 and 7, respectively. The edge artifacts on initial dehazed images can be removed after $t_{max} = 7$. The parameters in (11) and (12) are set as $\lambda_0 = 10^5$, $w_1 = 0.8$, $w_2 = 0.2$, and $\lambda_3 = 1$.

Algorithm I: Proposed near-infrared fusion method

Input: Near-infrared gray image \mathbf{x}^{nir} and visible color image (or haze image) \mathbf{x}^v

Output: Dehazed visible color image \mathbf{x}^s

Initialization:

- Generate the colorized near-infrared image \mathbf{x}^o via near-infrared coloring [17]
- Find the transmission k_i and the atmospheric color light \mathbf{x}^a using (9) and (10)
- Define $\mathbf{u}^{o(c)}$, $\mathbf{u}^{v(c)}$, and $\mathbf{u}_{t=0}^{d(c)}$ from \mathbf{x}^o , \mathbf{x}^v , and k_i , respectively, according to the haze degradation model in (5)
- Initialize the parameter w_1 , w_2 , λ_0 , λ_3

Near-infrared fusion:

```

for c = 1,2,...,cmax
  for t = 1,2,...,tmax
    • Update  $\mathbf{u}_{t+1}^{s(c)}$  via (11)
    • Update  $\mathbf{u}_{t+1}^{d(c)}$ , according to (12)
  end
  • Calculate  $\mathbf{x}^{s(c)} = \mathbf{x}^{a(c)} + e^{\mathbf{u}^{s(c)}}$ 
end
Return  $\mathbf{x}^s$ 

```

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

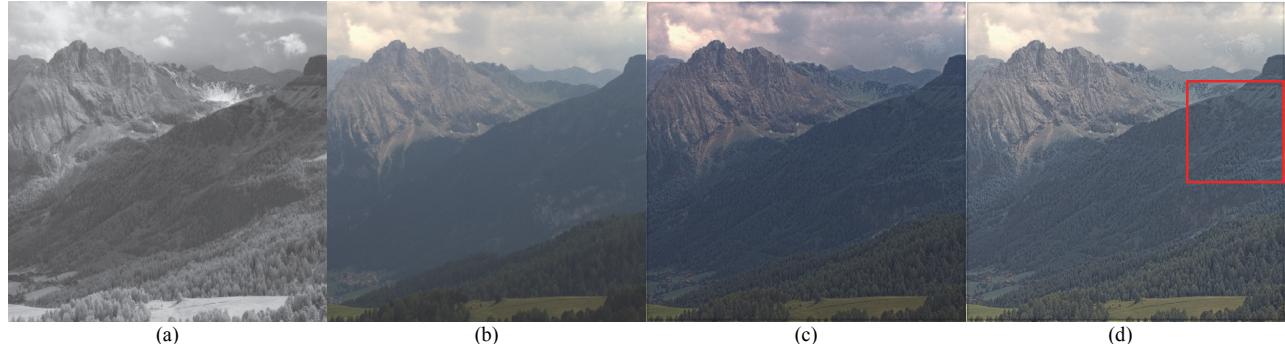


Figure 4. Proposed near-infrared fusion vs. coloring for 'Mountain' image; (a) captured near-infrared gray image, (b) captured visible color image, (c) dehazed visible image with the proposed method, and (d) colored near-infrared image [17] (Please zoom in to the images to check details and colors).

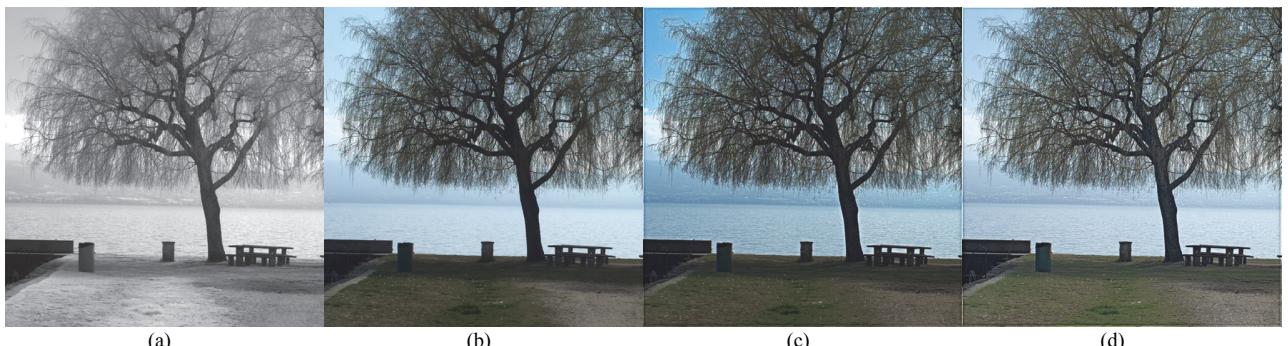


Figure 5. Proposed near-infrared fusion vs. coloring for 'Lake' image; (a) captured near-infrared gray image, (b) captured visible color image, (c) dehazed visible image with the proposed method, and (d) colored near-infrared image [17] (Please zoom in to the images to check details and colors).

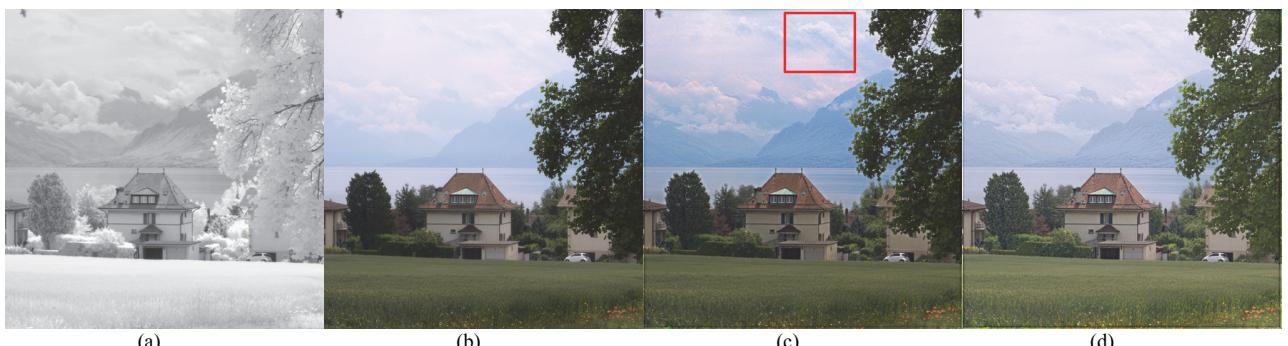


Figure 6. Proposed near-infrared fusion vs. coloring for 'House' image; (a) captured near-infrared gray image, (b) captured visible color image, (c) dehazed visible image with the proposed method, and (d) colored near-infrared image [17] (Please zoom in to the images to check details and colors).



Figure 7. Proposed near-infrared fusion vs. coloring for 'Village' image; (a) captured near-infrared gray image, (b) captured visible color image, (c) dehazed visible image with the proposed method, and (d) colored near-infrared image [17] (Please zoom in to the images to check details and colors).

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org



Figure 8. Experimental results for the used regularizations; (a) initial dehazed images using dark prior where alpha matting was not applied (first row), (b) dehazed images using gradient regularization (second row), and (c) dehazed images using gradient difference regularization (last row).

VI. EXPERIMENTAL RESULTS

In this paper, three types of experiments are conducted. First, visual effects, according to the proposed near-infrared fusion and the conventional near-infrared coloring [17], will be compared. This experiment will give us the difference between the proposed near-infrared fusion and the conventional near-infrared coloring [6,17]. Next, different regularizations will be tested using our fusion model to confirm how the proposed color regularization is more effective than other regularizations for color distortion and haze removals. Finally, visual quality comparison for different fusion methods [21,10] and the dehazing method [12] will be followed.

A. Proposed near-infrared fusion vs. coloring

The proposed color regularization additionally uses the colorized near-infrared image. Therefore, we have a question whether there is a difference between the proposed near-infrared fusion and near-infrared coloring [6,17]. As shown in (6), the proposed fusion model tries to satisfy both haze degradation model in the first term and the color regularization in the second term. On the other hand, near-infrared coloring [6,17] excludes the haze degradation model. Therefore, the visual appearances of the estimated haze-free images are different from those of the colorized

near-infrared images. As shown in Figs. 4-7, the colors of the dehazed images with the proposed method are vividly rendered, however the colorized near-infrared images are slightly diluted. In addition, the cloud, sky, and shadow descriptions of the dehazed images are better than those of the colorized near-infrared images: see the red boxes.

These results come from the difference between the near-infrared coloring and the near-infrared fusion based on the haze degradation model. The purpose of the near-infrared coloring is to add colors to the captured near-infrared gray images, and thus the visual appearances of the colorized near-infrared images are mainly influenced by the overall brightness of the captured near-infrared gray images. In contrast, the purpose of the proposed near-infrared fusion based on the haze degradation model is to remove the haze on the captured visible color images. Even though the proposed color regularization additionally utilizes the colorized near-infrared images, the value of w_1 is set larger than that of w_2 , which means that the visual color appearances of the dehazed images are dominantly influenced by the visible color images.

There are mainly two approaches of removing the haze. One is to use the near-infrared coloring and the other is to use the near-infrared fusion based on the haze degradation model. The near-infrared coloring can be a good solution to remove the haze, as shown in Figs. 4-7. The main difference between the

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

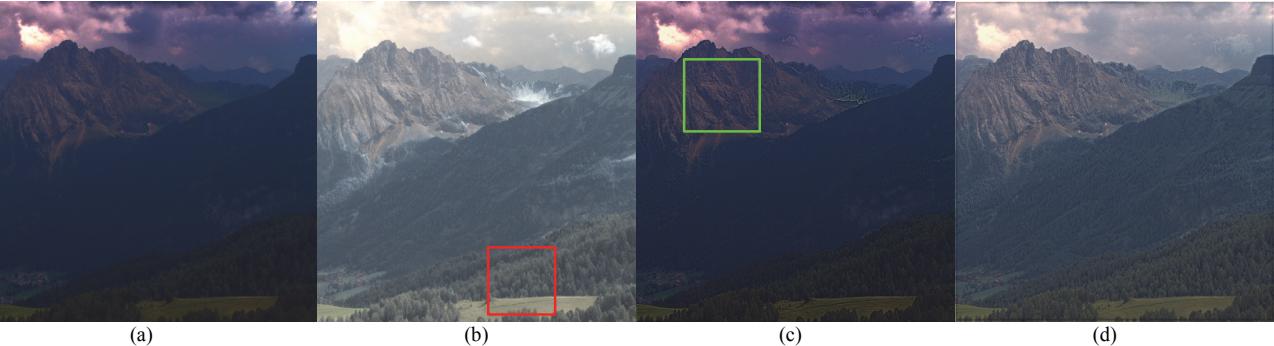


Figure 9. Dehazed 'Mountain' images; (a) single image dehazing using dark prior [12] where alpha matting was applied, (b) image-pair-based near-infrared dehazing using multiresolution representation [10], (c) image-pair-based near-infrared dehazing using gradient difference regularization [21], and (d) proposed method (Please zoom in the images to check details and colors).

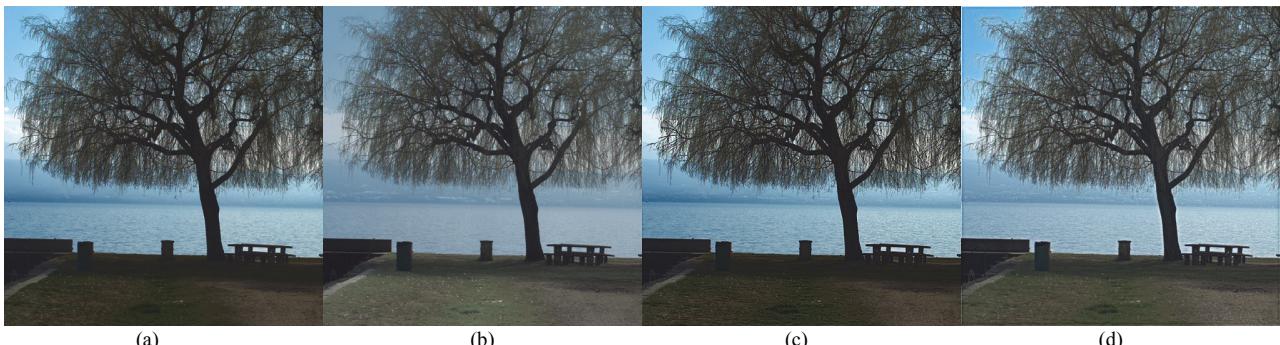


Figure 10. Dehazed 'Lake' images; (a) single image dehazing using dark prior [12] where alpha matting was applied, (b) image-pair-based near-infrared dehazing using multiresolution representation [10], (c) image-pair-based near-infrared dehazing using gradient difference regularization [21], and (d) proposed method (Please zoom in the images to check details and colors).

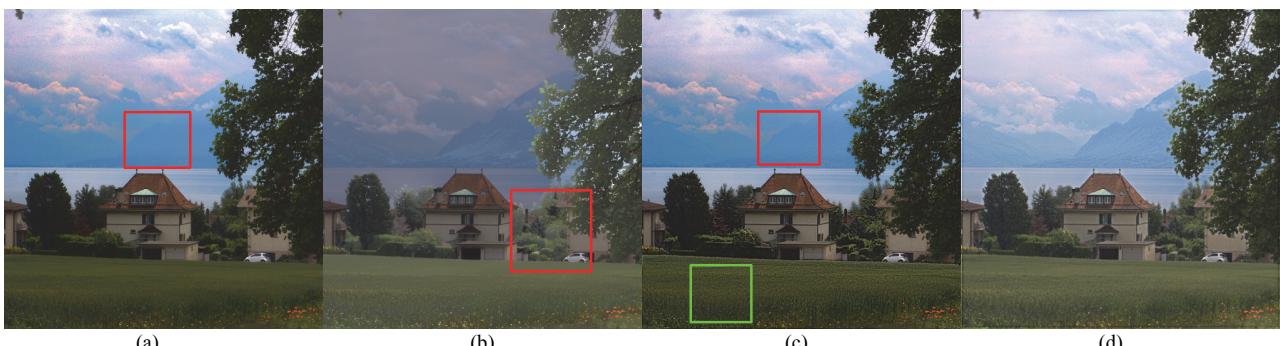


Figure 11. Dehazed 'House' images; (a) single image dehazing using dark prior [12] where alpha matting was applied, (b) image-pair-based near-infrared dehazing using multiresolution representation [10], (c) image-pair-based near-infrared dehazing using gradient difference regularization [21], and (d) proposed method (Please zoom in the images to check details and colors).



Figure 12. Dehazed 'Village' images; (a) single image dehazing using dark prior [12] where alpha matting was applied, (b) image-pair-based near-infrared dehazing using multiresolution representation [10], (c) image-pair-based near-infrared dehazing using gradient difference regularization [21], and (d) proposed method (Please zoom in the images to check details and colors).

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 10

near-infrared coloring and the proposed near-infrared fusion depends on the use of the haze degradation model, which makes a difference in rendered visual color appearances.

B. Visual effects according to the used regularizations

Now let us look at how the used different regularizations can influence the visual quality. Fig. 8 shows the dehazed images using different regularizations: dark prior [12], gradient regularization [18], and gradient difference regularization [21]. In Fig. 8, the first row shows the initial dehazed images using the single image dehazing method where the dark prior was used. However, the alpha matting algorithm [23] was not applied to check pure visual effects, according to the used regularizations. As shown in the figures, the visual qualities of the dehazed images are not satisfactory. The overall brightness of the dehazed images is very dark and the edge artifacts arise, as shown in the yellow boxes. The reasons are already discussed in [13,12]. The second row shows the dehazed images using gradient regularization. In other words, λ_2 is set to zero in (6), which indicates that the proposed fusion model becomes equal to the single image dehazing model. In this case, the gradient regularization, that is, only the third term is considered. The use of the gradient regularization makes the initial dehazed images smooth, and the strength of the edge artifacts can be further reduced. However, overall sharpness can be decreased as well. The original purpose of using the gradient regularization is to suppress the noise amplified after image dehazing [13]. The last row shows the dehazed images using gradient difference regularization. In (6), the proposed color regularization term was replaced by the gradient difference regularization, which is expressed as $\sum_{j=1}^2 \|\mathbf{f}_j \oplus \mathbf{u}_{t+1}^{(c)} - \mathbf{f}_j \oplus \mathbf{u}^{nir}\|_1$.

This regularization term can force the gradients of the initial dehazed images to be close to those of the captured near-infrared gray images. Thus, the edge strength of the initial dehazed images can be increased, as shown in the red boxes. However, the overall color appearances cannot be changed significantly. In other words, both gradient and gradient difference regularizations are highly dependent on the visual quality levels of the initial dehazed images. On the other hand, the proposed color regularization can provide the color prior for the unknown haze-free images. This can significantly improve the visual appearances of the dehazed images. This is why the proposed color regularization can provide better results, as shown in Figs. 4(c), 5(c), 6(c), and 7(c). Particularly, the edge artifacts can be removed without using the alpha matting [23], which has high computation complexity, as indicated in [13]. In addition, the color and details can be improved significantly.

C. Visual quality comparison

Figs. 9-12 show the dehazed images with the conventional single image dehazing method [12], image-pair-based dehazing methods [10,21], and the proposed method. The dehazed images with the proposed fusion method are the same as the ones in Figs. 4-7. The single image dehazing method [12] using the dark prior is the same as the one used in the previous section, however the alpha matting [23] was applied to refine the initial

dehazed images in the first row of Fig. 8. The use of the alpha matting can improve the visual quality of the dehazed images, as shown in Figs. 9(a), 10(a), 11(a), and 12(a). Particularly, the edge artifacts can be significantly reduced. However, even though the single image dehazing method is one of the state-of-the-art methods, it cannot remove haze completely, as shown in the red boxes of Figs. 11(a) and 12(a).

Figs. 9(b), 10(b), 11(b), and 12(b) show the dehazed (or fused) images using the image-pair-based dehazing method [10] where the near-infrared gray images and visible color images are both utilized. In this method, the details of the near-infrared gray images are combined with the visible color images using multiresolution representation, which means that this method does not reflect the haze degradation model. Based on the multiresolution representation, haze is well removed and fine details are preserved. However, the produced colors are unnatural and diluted, as shown in red boxes of Figs. 9(b), 11(b), and 12(b). This is caused by the discrepancy problem between the near-infrared gray images and visible color images. Please note that the objective of the near infrared fusion method is to remove not only the haze, but the color distortion as well.

Figs. 9(c), 10(c), 11(c), and 12(c) show the dehazed image using another image-pair-based dehazing method [21] where the gradient difference regularization is used to transfer the edges of the near-infrared gray images to the initially dehazed images with the dark prior method [12], which are shown in Figs. 9(a), 10(a), 11(a), and 12(a). Note that the initial dehazed images are already refined via alpha matting, which is totally different from the previous section. As expected, the use of the gradient difference regularization can produce sharp edges while preserving the visual color appearances of the initial dehazed images. However, the visual quality of the dehazed images with the gradient difference regularization is highly subject to the visual quality level of the initial dehazed images. In other words, the gradient difference regularization can strengthen the edges of the initial dehazed images, however it cannot change the overall color appearances of the initial dehazed images. For example, the edges in the green boxes of the Figs. 9(c) and 11(c) are more strengthened. However, the dehazed images are still so dark, which indicates that the gradient difference regularization cannot change the overall color appearances of initial dehazed images. For these reasons, haze remains on the dehazed images; see the red boxes of Figs. 11(c) and 12(c). In comparison, the proposed method can provide natural-looking colors and fine details. In the supplementary material, additional resulting images are provided to show that similar effects can be found on the dehazded images. From these results, it is confirmed that the proposed color regularization is effective to remove the color distortion and haze simultaneously.

D. Quantitative evaluation

Two types of measures: image structure similarity (ISS) [31] and color difference (CD) [8] are used to evaluate the near-infrared dehazing and fusion methods. The two types of the measures are defined in Table II. As mentioned in Introduction, color distortion and haze degradation are the main

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 11

TABLE II. DEFINITION OF THE TWO MEASURES

• ISS = $(\sigma_{x^s x^r} + c_3) / (\sigma_{x^s} \sigma_{x^r} + c_3)$ where σ_{x^s} , σ_{x^r} , and $\sigma_{x^s x^r}$ indicate the standard deviations and correlation coefficient for dehazed and reference images \mathbf{x}^s and \mathbf{x}^r , respectively. c_3 is a constant value to avoid numerical instability [31].
• CD = $\sum_{i=1}^N \sqrt{\sum_{c=1}^3 (\mathbf{x}_i^{s(c)} - \mathbf{x}_i^{r(c)})^2} / N$ where \mathbf{x}^s and \mathbf{x}^r indicates the dehazed and reference color images, respectively. The color channel c indicates one of the $L^*, a^*,$ or b^* color spaces [8].

TABLE III. QUANTITATIVE EVALUATION

Method	Images Measures		Mountain	Lake	House	Village	AVG.
	ISS	CD					
Dark Prior [12]	ISS	0.8619	0.7306	0.8331	0.9106	0.8341	
	CD	18.3315	10.2772	10.1315	15.3934	13.5334	
Multiresolution Representation [10]	ISS	0.9482	0.9447	0.9760	0.9738	0.9607	
	CD	13.3806	5.4886	9.3962	6.0070	8.5681	
Gradient Difference Regularization [21]	ISS	0.9066	0.8521	0.8719	0.8538	0.8711	
	CD	18.5545	11.0302	10.6415	15.9251	14.0378	
Near-Infrared Coloring [17]	ISS	0.9454	0.9308	0.9496	0.9700	0.9489	
	CD	7.9292	7.5664	12.135	9.6295	9.3150	
Proposed Method	ISS	0.9388	0.9037	0.9296	0.9567	0.9322	
	CD	4.2215	3.5201	4.6333	14.7750	6.7875	



Fig. 13. Masked images to indicate haze regions, which are filled with black colors.

issue for near-infrared fusion. To reflect this point, the two types of measures are adopted. The ISS is used to measure how well the haze is removed on the dehaze images. The CD is used to examine how accurately the dehazed colors are produced. The ISS and CD are the full-reference image quality evaluations, in other words, reference images should be provided to measure the ISS and CD. However, the captured visible color images and near-infrared gray images cannot be used as the reference images. This is because the captured visible color images contain the haze degradation and the captured near-infrared gray images have no colors. To solve this issue, masked images to indicate the haze regions are used in this paper. The examples of the masked images are provided in Fig. 13, where the haze regions are filled with black colors. The haze regions are manually specified by users. As for the haze regions, the captured near-infrared gray images could be used as the reference images to check how well the haze is removed on dehazed images. As for the non-haze regions, the captured visible color images could be used for the reference images to check how the dehazed colors deviate from the reference images. In other words, for the haze regions, the ISS will be measured using near-infrared gray images (i.e.,

reference images) and dehazed images (i.e., test images), whereas for non-haze regions, the CD will be measured using visible color images (i.e., reference images) and dehazed images (i.e., test images).

In Table II, the ISS value ranges between 0 and 1. Here, '1' indicates the perfect match in the image structures between input images, whereas '0' indicates perfect mismatch. Actually, the ISS defined in Table II is the same as the structure function in SSIM (Structural SIMilarity) measure [31]. There is serious discrepancy in the brightness between the captured visible color image and the near-infrared gray image. This indicates that the luminance function of the SSIM should be excluded for visibility (or haze removal) evaluation. In other words, the brightness of the captured near-infrared gray image cannot be a reference. In addition, the contrast function is also excluded because the local contrast is related to the brightness [8]. Thus, only structure function (i.e., ISS) is used in this paper. The CD value is greater than or equal to zero. When the colors of input images are identical, the CD value will be zero and it will increase as the colors become more different. The CD value is calculated in the uniform color space (i.e., $L^* a^* b^*$ color space), which means that the perceived color difference is measured by

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 12

the Euclidean distance. Therefore, the CD can measure how the colors of the dehazed image are different from those of the captured visible color image.

As shown in the last column of Table III, the single dehazing method using dark prior [12] has the lowest average ISS value, which means that the single dehazing method fails to remove the haze. The dark prior has a limitation in that it is difficult to provide good initial points for the unknown haze-free images. In contrast, the image-pair-based dehazing method using multiresolution representation [10] results in the highest average ISS value, along with good CD values. However, for the additional test images, which are provided in the supplementary material, its average CD score is the worst among all methods, indicating that the produced colors are highly distorted. This method simply combines the visible color images with the details of the near-infrared gray images based on multiresolution representation. Additionally, the haze degradation model is excluded; thus, it obtains the highest ISS values. However, this method is not free from the discrepancies in brightness and image structure that can be found between the visible color image and near-infrared image. Thus, the produced colors are inevitably distorted. Moreover, it should be noted that the colors produced are unnatural and diluted. However, this unnaturalness cannot be evaluated using the CD measure. The unnatural and diluted colors, as shown in Fig. 9(b), 10(b), 11(b), and 12(b), degrade the overall visual quality. Another image-pair-based dehazing method using gradient difference regularization [21] has the highest (or the worst) average CD value, which is caused by the dark colors and enhanced colors in the dehazed images. Compared to the single image dehazing [12], the ISS values can be improved due to the use of the gradient difference regularization. In contrast to the methods discussed above, both the near-infrared coloring method [17] and the proposed near-infrared fusion method can obtain good ISS and CD values. However, the proposed method has smaller average CD value than the near-infrared coloring. This is caused by different goals of the near-infrared coloring and the proposed near-infrared fusion. As discussed in the previous section, the near-infrared coloring adds colors to near-infrared gray images, and thus the visual appearances of the colored near-infrared images are influenced by the overall brightness of the captured near-infrared gray images. On the other hand, the proposed method includes the haze degradation model, thus the visual color appearances of the dehazed images are derived from the captured visible color images. As a result, the proposed method can have smaller CD values than the near-infrared coloring. This is the main difference between the near-infrared coloring and the proposed near-infrared fusion. Also, the sky, cloud, and shadow descriptions of the proposed method are better than those of the near-infrared coloring, which cannot be evaluated by the two measures. The proposed method has the smallest average CD value and obtains high ISS values. This indicates that the proposed color regularization is effective to remove the haze and color distortion problems. The conventional and proposed methods have their own merits and demerits. However, the conventional methods, except near-infrared coloring, have the worst values for ISS and CD,

depending on the test image sets, which significantly degrade the overall visual quality. Even though the multiresolution representation [10] seems to obtain good scores for ISS and CD, it has unnatural coloring. Moreover, the multiresolution representation obtained the worst CD score for the additional test set in the supplementary material. In contrast, the proposed method can obtain good ISS and CD values. In other words, the proposed method can remove haze and preserve high visibility while avoiding color distortion. This is the advantage of the proposed method.

In Table III, the multiresolution representation has higher average ISS than the proposed method, whereas the proposed method has smaller average CD than the multiresolution representation. However, the visual quality of the proposed method is better than the multiresolution representation, as shown in Figs. 9-12 and in the supplementary material. This indicates that the reduction in CD value seems more important than the decrease in the ISS value. Thus, the scale of the CD measure is not the same as that of the ISS measure. Therefore, in the future, a new measure for near-infrared fusion must be developed. More specifically, uniform scale in the CD and ISS domains will be considered via a psychophysical experiment or a new measure will be developed to connect the CD and ISS measures to the subjective scores.

VII. DISCUSSIONS

A. Visual effects, according to initial colorization methods

The recently introduced state-of-the-art colorization [32] based on deep learning is tested to evaluate how the initialization of the colorized near-infrared image influences the final dehazed images. The colorized images via deep learning are shown in Fig. 14. As shown in the middle row, the colorized near-infrared images have serious color distortion. If the colorized images are used as the initial colorized images to model the proposed color regularization, the dehazed images have unnatural-looking colors, as shown in the red boxes of the third row images. This is because the distorted colors of the colorized near-infrared images are transferred into the initial dehazed version of the captured visible color images. Although the color distortion can be reduced on the dehazed images owing to the setting of $w_2 = 0.2$ in (11), which means that the colors of the initial dehazed images are fused with the colors of the initial colorized near-infrared images, there are still unnatural looking colors on the dehazed images as shown in the third row images. In addition, visibility is reduced on the dehazed image as shown in the yellow box.

The performance of the proposed method using two different colorization methods is compared. One is the near-infrared coloring [17] and the other is the state-of-the-art colorization based on deep learning [32]. As shown in Table IV, the CD value of the proposed method using the colorization based on deep learning [32] is larger than the CD value of the proposed method using the near-infrared coloring [17]. This indicates that larger color errors are generated when the near-infrared coloring is replaced by colorization based on deep learning. The colorization method based on deep learning [32] uses only

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 13



Fig. 14. Experimental results: near-infrared gray images (upper row), colorized near-infrared images with deep learning [32] (middle row), and dehazed images with the proposed method where the colorized near-infrared images, as shown in the middle row, were used as the initial colorized near-infrared images (third row).

TABLE IV. PERFORMANCE COMPARISON OF THE PROPOSED METHOD ACCORDING TO THE INITIALIZATION OF COLORIZATION.

Colorization Methods	Images Measures	Mountain	Lake	House	Village	AVG.
		ISS	0.939	0.9186	0.9440	0.9146
Deep learning approach [32]	CD	7.8725	7.1652	20.338	17.9	13.319
	ISS	0.9388	0.9037	0.9296	0.9567	0.9322
Near-infrared coloring [17]	CD	4.2215	3.5201	4.6333	14.7750	6.7875

captured near-infrared gray image. It is difficult to estimate the original colors from the given near-infrared gray image. In contrast, the near-infrared coloring method [17] uses two images: the captured visible color image and the near-infrared gray image. Thus, the near-infrared coloring [17] can find the linear mapping relation between the near-infrared gray image and the luminance plane of the visible color image. This means that the colors of the near-infrared gray images can be generated according to the linear mapping relation. Although there is a serious discrepancy in the brightness between the near-infrared gray image and the luminance plane of the visible color image, natural-looking colors of the near-infrared gray image can be induced from the visible color image. This can significantly reduce color errors on the dehazed images.

B. Visual effects, according to initial dehazing methods

In the proposed method, the colorized near-infrared image is

fused with the initial dehazed image, as shown in Fig. 3. Therefore, overall color tones of the dehazed image are affected by those of the initial dehazed image. This means that the dehazed images with the proposed method cannot be same if different methods are used for initial dehazing. However, even though the visual quality of the initial dehazed images are poor regarding visibility and colors, the proposed method can guarantee high visibility and natural-looking colors on the dehazed images, irrespective of initial dehazing method applied. This is thanks to the proposed color and depth regularizations, which can transfer high visibility and natural-looking colors of the colorized near-infrared image into the initial dehazed image. In the supplementary material, another dehazing method [14] is tested to show the visual effects, according to the used initial dehazing methods [12, 14]. In the supplementary material, it is shown that the proposed method can produce high visibility and

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org



Fig. 15. Experimental results: $w_1 = 0.01$, $w_1 = 0.3$, $w_1 = 0.8$, and $w_1 = 0.95$ (left to right).



Fig. 16. Experimental results: initial depth map, estimated depth map with $\lambda_3 = 100$, estimated depth map with $\lambda_3 = 1$, fused image with $\lambda_3 = 100$, and fused image with $\lambda_3 = 1$ (left to right).

natural-looking colors on the dehazed images, irrespective of the used initial dehazing methods [12,14].

C. Parameter setting

First, λ_1 and λ_2 in (6) are also expressed by $\lambda_0 w_1 / 2$ and $\lambda_0 w_2 / 2$ as in (11). Here, w_2 is equal to $1 - w_1$. Therefore, w_1 and w_2 (or λ_1 and λ_2) determine which images between $(\mathbf{u}^{v(c)} + \mathbf{u}_t^{d(c)})$ and $\mathbf{u}^{o(c)}$ should be weighted more heavily. Note that these two images ($\mathbf{u}^{v(c)} + \mathbf{u}_t^{d(c)}$ and $\mathbf{u}^{o(c)}$) are fused, as already shown in the block diagram of Fig. 3. To check the visual effects according to w_1 and w_2 , the value of w_1 is varied. The resulting images are shown in Fig. 15. The resulting images are the fusion of the two images ($\mathbf{u}^{v(c)} + \mathbf{u}_t^{d(c)}$ and $\mathbf{u}^{o(c)}$) based on the assigned weight value of w_1 . In other words, if the value of w_1 approaches 0, the haze degradation model (i.e., the first term in (6) or (11)) disappears, therefore, the resulting image becomes close to the colorized near-infrared image $\mathbf{u}^{o(c)}$. On the other hand, if the value of w_1 approaches 1, the proposed color regularization model (i.e., the second term in (6) or (11)) disappears, therefore, the resulting image becomes close to the dehazed image using the dark prior method. This indicates that the main difference between the near-infrared fusion and the proposed near-infrared fusion depends on the use of the haze degradation model, which makes a difference in rendered visual color appearances. In this paper, w_1 was set to 0.8 to give more weight to the haze degradation term.

Second, λ_3 is related to the updating of the depth map as described in (12), which indicates that the depth map $\mathbf{u}_{t+1}^{d(c)}$ to be updated is the weighted average of the previous depth map $\mathbf{u}_t^{d(c)}$ and the currently estimated depth map $(\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{v(c)})$. Therefore, if the value of λ_3 is large, the previous depth map $\mathbf{u}_t^{d(c)}$ significantly influences the depth map $\mathbf{u}_{t+1}^{d(c)}$ to be updated,

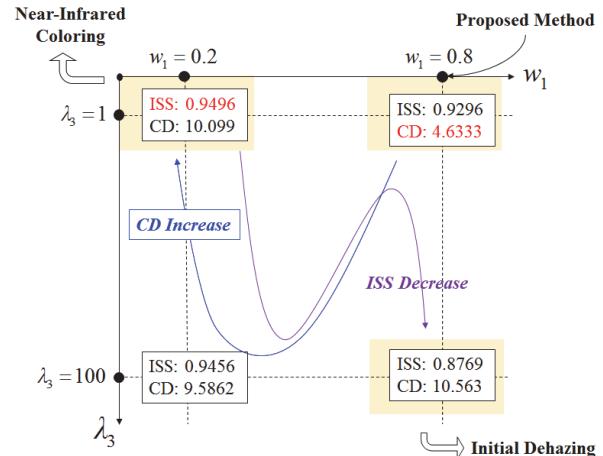


Fig. 17. Relation between the ISS and CD values and their parameters.

whereas if the value of λ_3 is small, the currently estimated depth map $(\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{v(c)})$ influences the depth map $\mathbf{u}_{t+1}^{d(c)}$. The visual effects based on λ_3 are shown in Fig. 16. As shown in the resulting images, the initial depth map includes serious depth errors, especially for the house, as indicated by the red box. However, the use of the proposed depth difference regularization can reduce the depth errors, as shown in the middle image where $\lambda_3 = 1$. If the value of λ_3 is quite high, the estimated depth map appears similar to the initial depth map. This is because the previous depth map $\mathbf{u}_t^{d(c)}$ significantly influences the depth map $\mathbf{u}_{t+1}^{d(c)}$ to be updated. Also, since the estimated depth map appears similar to the initial depth map, the fused image appears similar to the initial dehazed image with the dark prior method, as shown in the fourth image. If the λ_3 is set to zero, the proposed depth difference regularization term disappears in (6). In this case, the dehazed version ($\mathbf{u}_{t+1}^{m(c)}$)

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 15

of the captured visible color image is the same as the estimated haze-free image $\mathbf{u}_{t+1}^{s(c)}$, as described in (13). Therefore, the initial depth map (or the dehazed version of the visible color image) is not updated anymore. In other words, the fused image is just a weighted average of the two input images ($\mathbf{u}^{v(c)} + \mathbf{u}_{t=0}^{d(c)}$ and $\mathbf{u}^{o(c)}$), which is not considered in updating the initial depth map. In this study, λ_3 is determined heuristically and empirically. To the best of our knowledge, it is difficult to determine the value of λ_3 based on mathematical analysis. However, it should be noted that the visual effects according to the weighted averaging of the two images ($\mathbf{u}_t^{d(c)}$ and $(\mathbf{u}_{t+1}^{s(c)} - \mathbf{u}^{v(c)})$ or $\mathbf{u}_t^{m(c)}$ and $\mathbf{u}_{t+1}^{s(c)}$) based on the fixed λ_3 are the same for all test images.

Fig. 17 shows the ISS and CD values, according to the combination of the two parameters for the 'house' image. In Fig. 17, w_1 is set with 0.2 or 0.8, and λ_3 is set with 1 or 100. This combination can also show how the ISS and CD values are clearly different. As shown in Fig. 17, small w_1 and small λ_3 improve the ISS value, however, the CD value also increased. This is because this setting makes the proposed model closer to the near-infrared coloring. On the other hand, a higher w_1 and a small λ_3 can significantly decrease the CD value though the ISS value is slightly reduced. The setting of a high w_1 and a high λ_3 makes the proposed model appear closer to the single image dehazing. Therefore, the resulting image becomes a close resemblance of the initial dehazed image, which has a small ISS value and a high CD value. Compared to the setting with a small w_1 and a small λ_3 , the increment in the λ_3 resulted in the decrease in the ISS and CD values. This is because the previous depth map significantly influences the updated depth map, as already mentioned in Fig. 16.

VIII. CONCLUSIONS

In this paper, a new near-infrared fusion model based on the proposed color and depth regularizations is proposed for haze and color distortion removals. The discrepancies in the brightness and image structures between captured near-infrared images and visible color images generate the color distortion on the fused images. Therefore, the color distortion and haze degradation should be considered during image fusion. This motivates us to develop the new near-infrared fusion model that combines the proposed color and depth regularizations with the conventional haze degradation model. The proposed color regularization models the unknown colors of the haze-free image and the proposed depth regularization transfers the details and natural-looking colors of the colorized near-infrared image into the initial dehazed image. Experimental results showed that the proposed color and depth regularizations can remove the color distortion and haze simultaneously. Also, it is shown that the proposed color regularization can provide better visual appearances in colors and details than the conventional regularizations: dark prior, gradient regularization, and gradient difference regularization. Furthermore, visual effects according to the haze degradation model are provided to help understand

the differences between the proposed near-infrared fusion and the near-infrared coloring.

REFERENCES

- [1] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles*, Wiley, 2007.
- [2] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sensing and Environment*, vol. 24, no. 3, pp. 450-479, April 1988.
- [3] Q. Wang, P. Yan, Y. Yuan, X. Li, "Multi-spectral saliency detection," *Pattern Recognition Letters*, vol. 34, no. 1, pp. 34-41, Jan. 2013.
- [4] Q. Wang, J. Fang, Y. Yuan, "Multi-cue based tracking," *Neurocomputing*, vol. 131, pp. 227-236, May 2014.
- [5] N. Salamat, A. Germain and S. Süstrunk, "Removing shadows from images using color and near-infrared," in *Proc. IEEE International Conference on Image Processing*, Brussels, Belgium, 2011, pp. 1713-1716.
- [6] C. Fredembach and S. Susstrunk, "Colouring the near infrared," in *Proc. IS&T Color Imaging Conference*, Los Angeles, California, U.S.A., 2008, pp. 176-182.
- [7] T. Ogawa, Y. Igarashi, and M. Haseyama, "NMF-based spectral reflectance estimation from image pairs including near-infrared components," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 5, pp. 855-867, May 2016.
- [8] M. D. Fairchild, *Color appearance models*, Addison-Wesley, 1997.
- [9] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Application*, vol. 21, no. 5, pp. 34-41, Sep./Oct. 2001.
- [10] L. Schaul, C. Fredembach, and S. Susstrunk, "Color image dehazing using the near-infrared," in *Proc. IEEE International Conference on Image Processing*, Cairo, Egypt, 2009, pp. 1629-1932.
- [11] V. S. Petrovic and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228-237, Feb. 2004.
- [12] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 2009, pp. 1956-1963.
- [13] F. Faming, F. Li, and T. Zeng, "Single image dehazing and denoising," *SIAM Journal on Imaging Sciences*, vol. 7, no. 3, pp. 969-996, 2014.
- [14] Q. Yan, L. Xu, and J. Jia, "Dense scattering layer removal," in *Proc. 6th ACM SIGGRAPH Asia*, Hong Kong, 2013, pp. 1-10.
- [15] S.-C. Huang, B.-H. Chen, and W.-J. Wang, "Visibility restoration of single hazy images captured in real-world weather conditions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 10, pp. 1814-1824, Oct. 2014.
- [16] Q. Zhu, J. Mai, and L. Shao, "Fast single image haze removal algorithm using color attenuation prior," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3522-3533, Nov. 2015.
- [17] C.-H. Son, K. Lee, and X.-P. Zhang, "Near-infrared coloring via a contrast-preserving mapping model," in *Proc. IEEE Global Conference on Signal & Information Processing*, Orlando, Florida, USA, 2015.
- [18] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," in *Proc. Neural Information Processing Systems*, Vancouver, Canada, 2009, pp. 1-9.
- [19] W. Lia, J. Zhang, and Q.-H. Daib, "Robust blind motion deblurring using near-infrared flash image," *Journal of Visual Communication and Image Representation*, vol. 24, no. 8, pp. 1394-1413, Nov. 2013.
- [20] S. Zhuo, D. Guo, and T. Sim, "Robust flash deblurring," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, U.S.A., 2010, pp. 2440-2447.
- [21] C. Feng, S. Zhuo, X. Zhang, L. Shen, and S. Susstrunk, "Near-infrared guided color image dehazing," in *Proc. IEEE International Conference on Image Processing*, Melbourne, Australia, 2013, pp. 2363-2367.
- [22] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 248-272, 2008.
- [23] A. Levin, D. Lischinski, and Y. Weiss, "A closed form solution to natural image matting," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, New York, NY, USA, 2006, pp. 61-68.
- [24] C.-H. Son, H. Choo, and H.-M Park, "Image-pair-based deblurring with spatially varying norms and noisy image updating," *Journal of Visual Communication and Image Representation*, vol. 24, no. 8, pp. 1303-1315, Nov. 2013.

Copyright © 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org 16

- [25] M. V. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681-695, 2011.
- [26] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323-343, 2009.
- [27] C.-H. Son and X.-P. Zhang, "Layer-based approach for image pair fusion," *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2866-2881, June 2016.
- [28] V. Maik, D. Cho, J. Shin, and J. Paik, "Regularized restoration using image fusion for digital auto-focusing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1360-1369, Oct. 2007.
- [29] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, Jun. 2013.
- [30] Y. Cao, M. Yang, and C.-L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1-D guided filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 12, pp. 2176-2188, Dec. 2016.
- [31] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004.
- [32] G. Larsson, M. Maire, and G. Shakhnarovich, "Learning representations for automatic colorization," in *Proc. European Conference on Computer Vision*, Amsterdam, Netherlands, 2016, pp 577-593.



Chang-Hwan Son received the B.S., M.S., and Ph.D. degrees in Electronic Engineering from Kyungpook National University, Daegu, Korea, in 2002, 2004, and Aug. 2008, respectively. He was a National Scholarship Student in 2006. From 2008 to 2009, he was with the Advanced Research and Development Group 2, Digital Printing Division, Samsung Electronics Company, Ltd., Suwon, Korea. As a Senior Research Engineer, he was involved in the development of color reproduction algorithms in digital printers and copiers. From Mar. 2015 to Feb. 2017, he was a postdoctoral researcher with the Department of Electrical and Computer Engineering, Ryerson University, Canada. He is now an Assistant Professor of the Department of the Software Convergence Engineering, Kunsan National University, South Korea. His main research interests include color imaging, computational photography, representation learning, image processing, sparse representation, printing, and hardcopy watermarking. He was a reviewer of the *IEEE Transactions on Image Processing*, *IEEE Transactions on Multimedia*, *IEEE Transactions on Circuits and Systems for Video Technology*, *SIAM Journal on Imaging Sciences*, and *IEEE Signal Processing Letters*.



Xiao-Ping Zhang (M'97, SM'02) received B.S. and Ph.D. degrees from Tsinghua University, in 1992 and 1996, respectively, both in Electronic Engineering. He holds an MBA in Finance, Economics and Entrepreneurship with Honors from the University of Chicago Booth School of Business, Chicago, IL. Since Fall 2000, he has been with the Department of Electrical and Computer Engineering, Ryerson University, where he is now Professor, Director of Communication and Signal Processing Applications Laboratory (CASPAL). He has served as Program Director of Graduate Studies. He is cross appointed to the Finance Department at the Ted Rogers School of Management at Ryerson University. His research interests include image and multimedia content analysis, statistical signal processing, sensor networks and electronic systems, machine learning, and applications in bioinformatics, finance, and marketing. He is a frequent consultant for biotech companies and investment firms. He is cofounder and CEO for EidoSearch, an Ontario based company offering a content-based search and analysis engine for financial data.

Dr. Zhang is a registered Professional Engineer in Ontario, Canada, and a member of Beta Gamma Sigma Honor Society. He is the general co-chair for ICASSP2021. He is the general co-chair for 2017 GlobalSIP Symposium on Signal and Information Processing for Finance and Business. He is an elected

member of ICME steering committee. He is the general chair for MMSP'15. He is the publicity chair for ICME'06 and program chair for ICIC'05 and ICIC'10. He served as guest editor for Multimedia Tools and Applications, and the International Journal of Semantic Computing. He is a tutorial speaker in ACM2011, ISCAS2013, ICIP2013, ICASSP2014, IJCNN2017. He is a Senior Area Editor for IEEE Transactions on Signal Processing. He is/was an Associate Editor for IEEE Transactions on Image Processing, IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Signal Processing, and IEEE Signal Processing Letters.