

RESEARCH ARTICLE

Singe Image Dehazing With Unsharp Masking and Color Gamut Expansion

DAT NGO¹, GI-DONG LEE^{1,2}, AND BONGSOON KANG^{1,2}¹Technical Support and Development Center for Display Device Convergence Technology, Busan 49315, South Korea²Department of Electronic Engineering, Dong-A University, Busan 49315, South Korea

Corresponding author: Bongsoon Kang (bongsoon@dau.ac.kr)

This work was supported by the Dong-A University, Busan, South Korea.

ABSTRACT Image dehazing is a fundamental problem in computer vision and has hitherto engendered prodigious amounts of studies. Recently, with the well-recognized success of deep learning techniques, this field has been dominated by deep dehazing models. However, deep learning is not always a panacea, especially for the practicalities of image dehazing, because high computational complexity, expensive maintenance costs, and high carbon emission are three noticeable problems. Computational efficiency is, therefore, a decisive factor in real-world circumstances. To cope with this growing demand, we propose a linear time algorithm tailored to three primitive parts: unsharp masking (pre-processing), dehazing, and color gamut expansion (post-processing). The first enhances the sharpness according to the local variance of image intensities. The second removes haze based on the improved color attenuation prior, and the third addresses a residual effect of color gamut reduction. Extensive experimental results demonstrated that the proposed method performed comparatively with popular benchmarks, notably deep dehazing models. With such a comparative performance, the proposed method is still fast and efficient, favoring real-world computer vision systems.

INDEX TERMS Image dehazing, unsharp masking, color gamut expansion, linear time complexity.

I. INTRODUCTION

Atmospheric scattering occurs when the sunlight enters the atmosphere and is diffused in all directions due to particle-particle collision. Scattering, coupled with absorption, decreases the quality of digital images, resulting in various types of degradation, such as faint color, low contrast, and detail loss. Although the degradation degree depends on the size of atmospheric particles, which varies according to weather conditions [1], this phenomenon is widely referred to as *haze*. On the one hand, haze obscures distant objects and affects the visibility perceived by human visual systems. On the other hand, it also affects high-level computer vision applications that assume clean input image/video data, as pointed out in [2]. Hence, *image dehazing* is a research branch focusing on alleviating the adverse effects of haze. Fig. 1 demonstrates the dehazing results of a real-world hazy image using a deep learning model [3] and the

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.



FIGURE 1. A real-world hazy image and its corresponding dehazing results by Ren *et al.* [3] and the proposed method. MS-CNN stands for the multi-scale convolutional neural network.

proposed method. Our result is more favorable to human visual systems because haze has been removed efficiently while fine details have been satisfactorily recovered (in the blue-cropped region).

A. DEGRADATION MODEL

Reference [4] formalizes the haze-induced degradation by a model comprising the direct attenuation and the airlight scattering, denoted as blue and red in Fig. 2, respectively. The former occurs when the reflected light at a particular

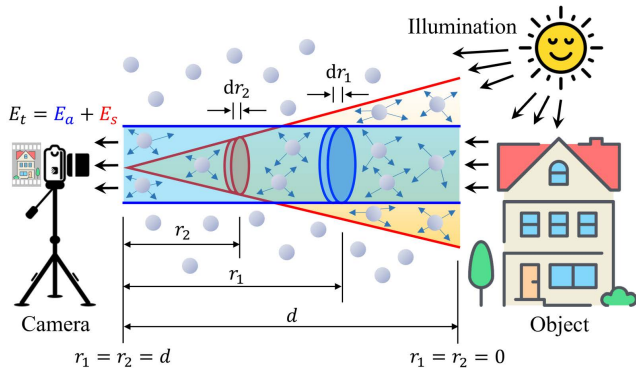


FIGURE 2. Illustration of the atmospheric scattering phenomenon.

wavelength λ (in m) propagates through suspended particles, where each of them causes an angular scattering that attenuates the light energy. For a unit volume of differential width dr_1 , the attenuated energy is formulated by:

$$\frac{dE_a(r_1, \lambda)}{E_a(r_1, \lambda)} = -\beta_{sc} dr_1, \quad (1)$$

where $E_a(r_1, \lambda)$ (in Wm^{-2}) denotes the irradiance at the unit volume dr_1 , and β_{sc} (in m^{-1}) denotes the scattering coefficient. The captured irradiance, $E_a(d, \lambda)$ at $r_1 = d$, is then obtained by integrating both sides of (1) as follows:

$$\int_{E_a(0, \lambda)}^{E_a(d, \lambda)} \frac{dE_a(r_1, \lambda)}{E_a(r_1, \lambda)} = \int_0^d -\beta_{sc} dr_1. \quad (2)$$

At $r_1 = 0$, the irradiance $E_a(0, \lambda) = \Omega_\lambda S_0 F_\lambda$ along the optical path is calculated from the mean radiance of the sky S_0 (in $\text{Wm}^{-2}\text{sr}^{-1}$), the solid angle Ω_λ (in sr), and the dimensionless reflectance coefficient F_λ . Hence,

$$E_a(d, \lambda) = \Omega_\lambda S_0 F_\lambda \exp(-\beta_{sc} d). \quad (3)$$

The other part, the airlight, indicates a portion of light reflected from the terrain surface or scattered directly to the camera. The irradiance $E_s(r_2, \lambda)$, in this case, can be calculated by considering a unit volume dr_2 whose radiance is $S_0 \beta_{sc} dr_2$. Similarly, the irradiance captured by the camera is attenuated, and its value at a distance r_2 from the camera is:

$$dE_s(r_2, \lambda) = \Omega_\lambda S_0 \beta_{sc} \exp(-\beta_{sc} r_2) dr_2. \quad (4)$$

The irradiance $E_s(d, \lambda)$ at $r_2 = d$ is then obtained by integrating (4) over the whole optical path, as follows:

$$\begin{aligned} E_s(d, \lambda) &= \Omega_\lambda S_0 \beta_{sc} \int_0^d \exp(-\beta_{sc} r_2) dr_2 \\ &= \Omega_\lambda S_0 [1 - \exp(-\beta_{sc} d)]. \end{aligned} \quad (5)$$

The total irradiance $E_t(d, \lambda)$ is the sum of $E_a(d, \lambda)$ and $E_s(d, \lambda)$ as:

$$E_t(d, \lambda) = \Omega_\lambda S_0 F_\lambda \exp(-\beta_{sc} d) + \Omega_\lambda S_0 [1 - \exp(-\beta_{sc} d)]. \quad (6)$$

For ease of representation, it is convenient to substitute $E_t(d, \lambda)$ with $\mathbf{I}(x)$, where the boldface indicates the

wavelength-dependent characteristics. Considering general cameras with sensors sensitive to red, green, and blue wavelengths, $\mathbf{I}(x) = \{I^R(x), I^G(x), I^B(x)\}$ denotes the x -th hazy intensities. Similarly, $\mathbf{J}(x)$, \mathbf{A} , and $t(x)$ can substitute for $\Omega_\lambda S_0 F_\lambda$, $\Omega_\lambda S_0$, and $\exp[-\beta_{sc} d(x)]$, respectively. These three are referred to as the haze-free intensities, the atmospheric light, and the transmittance. The following is, therefore, the simplified form of (6), and it commonly serves as the degradation model in image dehazing:

$$\mathbf{I}(x) = \mathbf{J}(x)t(x) + \mathbf{A}[1 - t(x)]. \quad (7)$$

B. ILL-POSEDNESS

The concept of ill-posedness dates back to [5], and a mathematical problem is called ill-posed (or incorrectly posed) if at least one of the following conditions for its solution fails:

- The existence
- The uniqueness
- The stability

In (7), the hazy intensities \mathbf{I} captured by the camera are the only data available, whereas the remainder is unknown. Accordingly, for a given set of \mathbf{A} and t estimated, (7) yields a different solution of \mathbf{J} and violates the uniqueness condition. This issue renders image dehazing ill-posed and brings about prodigious amounts of relevant studies. Recently, researchers have adopted deep learning techniques to address the ill-posedness, as witnessed by [6], [7], [8], [9], [10], and [11]. Despite such excellent performances as those deep dehazing models have delivered, they have been linked with several problems in real-world execution, such as high power consumption, high carbon emission, and expensive maintenance costs [12].

Furthermore, for a low-level vision task such as image dehazing, deep neural networks (DNNs) are often overkill, as discussed in [13] about deep learning and traditional computer vision techniques. In fact, they fit well with high-level cognitive tasks, such as object classification, recognition, and localization. The data-driven performance of DNNs is also more of a hindrance than a help because abstract features learned by DNNs are specific to the training dataset, whose construction is highly cumbersome for statistical reliability. Thus, the learned features may be inappropriate for images different from those in the training set, lowering the performance in general.

II. RELATED WORKS

This section briefly reviews influential works in the literature based on the categorization in [14], where algorithms have been divided into three categories according to their data exploitation. The first two, *image processing* and *machine learning*, were typified by low-level hand-engineered image features discovered through statistical analysis of real-world images. The last category, *deep learning*, exploited the powerful representation capability of DNNs to learn high-level data-driven image features. This categorization could give useful insights into (i) the complexity of dehazing algorithms

and (ii) subjective/objective preferences for dehazed images. Generally, image processing and machine learning-based methods possess low complexity and favor human perception. Deep learning-based methods, on the contrary, are computationally costly and favor image quality assessment metrics.

A. IMAGE PROCESSING

The dark channel prior [15], one of the most influential works in image dehazing, is a prime example of the first category. He *et al.* [15] observed natural haze-free images and discovered that the dark channel—calculated as the local minimum of the minimum channel—tended to approach zero at non-sky image patches. This finding, coupled with the degradation model, offered an efficient means to estimate the transmittance, which required soft-matting [16] for edge-aware smoothing. He *et al.* [15] also proposed locating the atmospheric light as the brightest pixel (in the red-green-blue (RGB) color space) among the top 5% pixels with the highest intensities in the dark channel. Given the estimated transmittance and atmospheric light, they reversed (7) to obtain the haze-free image.

Another influential work is the linear time algorithm by Tarel and Hautiere [17]. They first white-balanced the input image to support the assumption that the atmospheric light was pure white. After that, they inferred the airlight, the term $A[1-t(x)]$ in (7), from image whiteness as a percentage of its local deviation from its local average. This airlight inference required two median filters, hence the widely known name *the median of medians along lines*. Finally, they post-processed the restored image using a simple tone mapping operator to expand the dynamic range. The most computationally heavy part of this algorithm is the median filter, whose implementation is, fortunately, $\mathcal{O}(N)$ [18], therein lying the linear time complexity.

B. MACHINE LEARNING

As image dehazing based on in situ information is highly challenging, the distilled knowledge from relevant image datasets may improve the dehazing performance. Zhu *et al.* [19] developed the color attenuation prior based on extensive observations of natural outdoor images. This prior stated that the scene depth correlated with the difference between image saturation and brightness. Zhu *et al.* [19] then utilized a linear function to model that correlation and estimated the function's parameters by applying supervised learning on a synthetic dataset. Thus, the distilled knowledge was parameters used to estimate the scene depth from image saturation and brightness.

From a more general perspective, Tang *et al.* [20] investigated four haze-relevant features, including the dark channel, hue disparity, locally maximum contrast, and locally maximum saturation, at multiple scales and found the following. Although the dark channel was the most informative feature (as discovered by He *et al.* [15]), other features also contributed in a complementary manner. Hence, Tang *et al.* [20]

devised a framework for inferring the transmittance from different haze-relevant features. In [20], they employed a random forest regressor for ease of analysis and demonstration, albeit with slow inference time. They also discussed the importance of post-processing and presented two post-processing options: adaptive atmospheric light estimation and adaptive exposure scaling.

C. DEEP LEARNING

The aforementioned approaches require significant efforts in seeking (i) a good feature (or a set of features) and (ii) an efficient inference scheme. However, there is no guarantee that they will always perform as intended in all circumstances. As a result, deep learning has been applied to image dehazing to improve flexibility. Given a reliable training dataset, DNNs can estimate the transmittance and atmospheric light with high accuracy because they allow learning and augmenting image features from low to high levels of abstraction. For example, Cai *et al.* [21] designed a convolutional neural network (CNN) to perform the following: low-level feature extraction, multi-scale mapping, augmentation (for spatial invariance), and non-linear transmittance inference.

The powerful learning ability of DNNs or deep CNNs also allows them to infer the dehazed image directly from the hazy input. In this direction, the encoder-decoder network has been proved highly efficient for end-to-end learning [22], [23]. In addition, some well-known image processing schemes can be applied to deep learning to improve performance, as witnessed by multi-scale image fusion [22] and domain adaptation [23]. Also, inspired by the human brain that knowledge learned from doing a particular activity may benefit another activity, joint learning is a promising direction, typified by [24], where image dehazing benefits object detection.

Some state-of-the-art deep dehazing networks developed recently include GridDehazeNet (GDN) [25], multi-scale boosted dehazing network (MSBD) [26], you only look yourself (YOLY) [27], and self-augmented unpaired image dehazing (D^4) [28]. GDN is a supervised network and comprises three modules. The pre-processing module applies different data-driven enhancement processes to the input image. The backbone module then fuses the results based on the grid network, where a channel-wise attention mechanism is adopted to facilitate the cross-scale circulation of information. Finally, the post-processing module remedies residual artifacts to improve the dehazing quality.

MSBD is also a supervised network designed with boosting and error feedback mechanisms. The former successively refines the intermediate dehazing result to reduce the portion of haze (PoH defined by Dong *et al.* [26]), and the latter successively recovers spatial details obscured by haze. YOLY, on the contrary, is an unsupervised and untrained network. Based on the layer disentanglement in [29], YOLY is designed with three sub-networks that decompose the hazy image into three latent layers corresponding to scene radiance, transmittance, and atmospheric light. Thus, YOLY

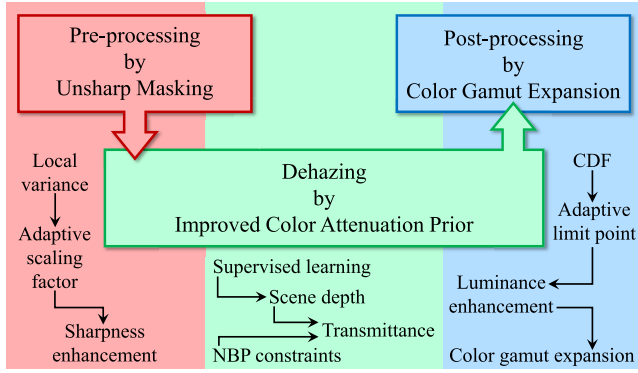


FIGURE 3. Illustration of the proposed method. NBP stands for no-black-pixel, and CDF stands for the cumulative distribution function.

supervises itself to jointly optimize three sub-networks and reconstruct the hazy image from a single input.

Yang *et al.* [28] argued that YOLY lacked knowledge from the clean image domain and developed D^4 as an alternative solution. Unlike other unpaired networks, D^4 takes account of the scattering coefficient and scene depth when carrying out dehazing and rehazing cycles. Consequently, D^4 can benefit from physical-model-based haze removal and generation to improve the performance of unpaired learning.

D. MOTIVATIONS

Image dehazing has undergone approximately five decades of development since the pioneering work in 1972 [30]. It is currently in the mature stage, and the focus is deemed to attain computational efficiency for integrating into low-cost edge devices, which are prevalent in the Industry 4.0 era.

As discussed thus far, although DNNs offer some definite advantages, such as accuracy and flexibility, they are not a preferable option. In contrast, traditional computer vision techniques are more suitable for image dehazing because a hand-engineered method can deliver comparative performance at a cheaper computational cost. This paper, therefore, proposes an $\mathcal{O}(N)$ algorithm that pre-processes, dehazes, and post-processes hazy images to satisfactorily restore clean visibility.

III. PROPOSED METHOD

Fig. 3 illustrates three major steps constituting the proposed method. The first is unsharp masking for enhancing the sharpness, wherein the enhancement is locally adapted to the variance of image intensities lest the out-of-range problem occurs. The second performs image dehazing based on the improved color attenuation prior [31], which estimates the transmittance from saturation and brightness (as by Zhu *et al.* [19]) and applies two no-black-pixel (NBP) constraints. The third performs color gamut expansion by enhancing the luminance and then expanding the color gamut proportionally to avoid color distortion. The following describes these three steps in more detail.

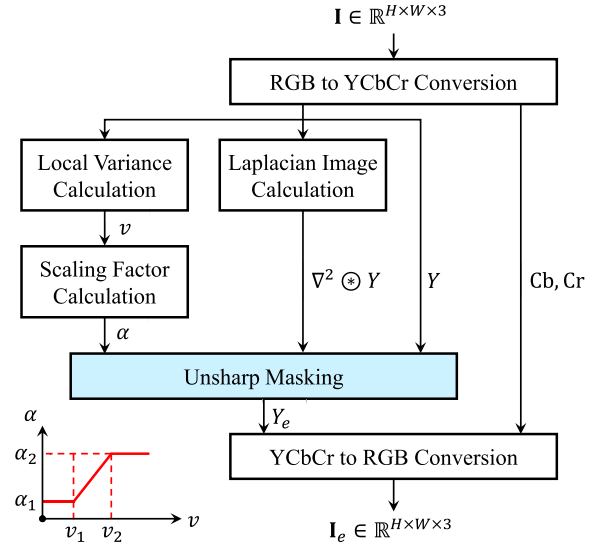


FIGURE 4. Block diagram of the pre-processing step.

A. PRE-PROCESSING

In the beginning, it is worth recalling that the input RGB image of size $H \times W$ is denoted as $\mathbf{I} \in \mathbb{R}^{H \times W \times 3}$, or interchangeably, $\mathbf{I} = \{I^R, I^G, I^B\}$, where $I^c \in \mathbb{R}^{H \times W}$ and $c \in \{R, G, B\}$.

As haze is depth-dependent, it is generally smooth except at discontinuities. Hence, it can be viewed as a low-frequency component that obscures fine details in the captured image. This pre-processing step then enhances these obscured details by adding the scaled Laplacian image to the original, as Fig. 4 shows. Because the sharpness enhancement only applies to the luminance channel, it is necessary to convert between RGB and YCbCr color spaces using (8) and (9) from [32]. In (8), Y , Cb , and Cr are the luminance, blue-difference chroma, and red-difference chroma components of the input image \mathbf{I} , and Y_e denotes the output luminance with sharpness enhanced. In (9), $\mathbf{I}_e \in \mathbb{R}^{H \times W \times 3}$, or $\mathbf{I}_e = \{I_e^R, I_e^G, I_e^B\}$, is the output RGB image corresponding to $\{Y_e, Cb, Cr\}$.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.183 & 0.614 & 0.062 \\ -0.101 & -0.338 & 0.439 \\ 0.439 & -0.399 & -0.040 \end{bmatrix} \begin{bmatrix} I^R \\ I^G \\ I^B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}, \quad (8)$$

$$\begin{bmatrix} I_e^R \\ I_e^G \\ I_e^B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.793 \\ 1.164 & -0.213 & -0.534 \\ 1.164 & 2.115 & 0 \end{bmatrix} \begin{bmatrix} Y_e - 16 \\ Cb - 128 \\ Cr - 128 \end{bmatrix}. \quad (9)$$

Next, the Laplacian image is obtained by convolving the input luminance Y with the Laplacian operator ∇^2 , whose definition is in (10). Meanwhile, the local variance v of luminance intensities is calculated as the expected value of the squared deviation from the mean, as (11) illustrates. The symbol \otimes denotes the convolution operator, and U_k is an all-ones square matrix of size $k \times k$, where $k = \{2n + 1 | n \in \mathbb{Z}^+\}$

is an odd integer.

$$\nabla^2 \triangleq \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad (10)$$

$$v = Y^2 \circledast \left(\frac{U_k}{k^2} \right) - \left[Y \circledast \left(\frac{U_k}{k^2} \right) \right]^2. \quad (11)$$

As demonstrated at the bottom-left of Fig. 4, the scaling factor α is a piece-wise linear function of the local variance v . The function definition is in (12), where $\{\alpha_1, \alpha_2, v_1, v_2\}$ are user-defined parameters for fine-tuning. Hence, the output luminance Y_e is obtained by (13), which scales the Laplacian image and adds it back to the input luminance. The YCbCr-to-RGB conversion in (9) then yields the output RGB image \mathbf{I}_e .

$$\alpha = \begin{cases} \alpha_1 & v < v_1 \\ \left(\frac{\alpha_2 - \alpha_1}{v_2 - v_1} \right) v + \frac{\alpha_1 v_2 - \alpha_2 v_1}{v_2 - v_1} & v_1 \leq v \leq v_2 \\ \alpha_2 & v > v_2, \end{cases} \quad (12)$$

$$Y_e = Y + \alpha \cdot (\nabla^2 \circledast Y). \quad (13)$$

Unsharp masking can be loosely viewed as a ‘‘mildly dehazing’’ step because it partially relieves the impact of haze on image sharpness. The following, conversely, is a haze-removal-dedicated step developed from the improved color attenuation prior [31].

B. DEHAZING

Two important parts of this step are (i) scene depth estimation and (ii) NBP constraint derivation. The former is based on the color attenuation prior [19] with several improvements in the learning scheme and the dataset preparation. Meanwhile, the latter is inspired by [33] to constrain the transmittance lest black pixels occur. Fig. 5 shows the overall block diagram, where the input image \mathbf{I}_e is from the previous step, and the dehazed image is denoted as $\mathbf{J} \in \mathbb{R}^{H \times W \times 3}$.

1) SCENE DEPTH ESTIMATION

The scene depth d is inferred from saturation S and brightness V using a linear function below:

$$d = \theta_0 + \theta_1 S + \theta_2 V + \varepsilon, \quad (14)$$

where $\{\theta_0, \theta_1, \theta_2\}$ are the function’s parameters, and ε denotes the error associated with the inference. Zhu *et al.* [19] assumed that ε followed the normal distribution $\mathcal{N}(0, \sigma^2)$ with zero mean and σ^2 variance. Hence, it was derived from (14) that d also followed a normal distribution $\mathcal{N}(\theta_0 + \theta_1 S + \theta_2 V, \sigma^2)$. Given an annotated dataset of hazy images and their corresponding scene depths, maximum likelihood estimates can be applied to learn the parameters. However, as it is virtually impossible to obtain such a dataset in practice, a synthetic dataset with depth information drawn from a probability distribution is a viable alternative.

Zhu *et al.* [19] utilized the standard uniform distribution (SUD) to generate a synthetic training dataset. After that, they adopted the stochastic gradient ascent (SGA) to find

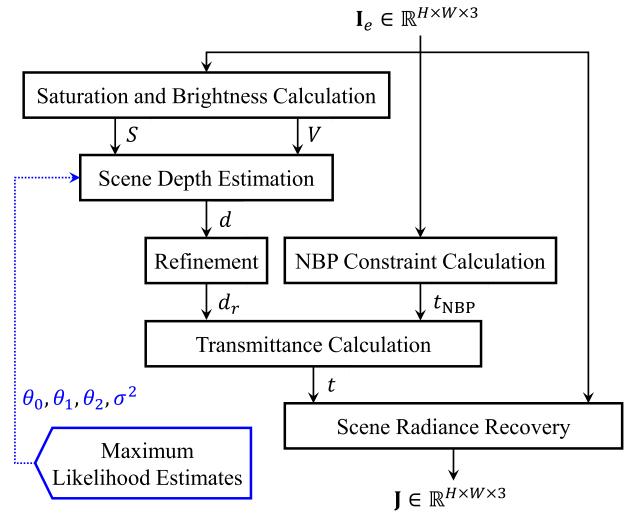


FIGURE 5. Block diagram of the dehazing step.

the parameters that maximized the log-likelihood function. In the proposed method, the enhanced equidistribution [31] supersedes SUD to improve the statistical reliability of the synthetic dataset. Additionally, the mini-batch gradient ascent with an adaptive learning rate [34] replaces SGA to reduce the convergence time.

The scene depth d is now available, but there is no guarantee that it will be mostly smooth except at discontinuities. Consequently, the refinement block applies a modified hybrid median filter [35] to the scene depth to impose edge-aware smoothness. Given the refined scene depth d_r , the transmittance is obtained through $t = \exp[-\beta_{sc} d_r]$, with $\beta_{sc} = 1$. Generally, most image dehazing algorithms in the literature adopted two fixed limits to constrain t , expressed as $t_0 \leq t \leq 1$, with t_0 being a small positive number. The following, on the contrary, describes two NBP constraints for posing an adaptive lower limit on t .

2) NBP CONSTRAINTS

From (7), the dehazed image (or, equivalently, scene radiance) \mathbf{J} can be obtained as:

$$\mathbf{J} = \frac{\mathbf{I}_e - \mathbf{A}}{t} + \mathbf{A}. \quad (15)$$

The first NBP constraint, $\mathbf{J} \geq 0$, is relatively evident because it can reduce the number of black pixels that occur after dehazing due to underflows. Hence, it is derived from (15) that:

$$t \geq 1 - \min_{c \in \{R, G, B\}} \left(\frac{I_c^c}{A^c} \right), \quad (16)$$

where $\min_{c \in \{R, G, B\}}(\cdot)$ denotes a channel-wise minimum operation.

The second NBP constraint is inspired by [33] that the local mean intensity of \mathbf{J} must be greater than or equal to its local standard deviation, as expressed by:

$$\text{mean}_{\forall y \in \Omega(x)} [Y_p(y)] \geq q \cdot \text{std}_{\forall y \in \Omega(x)} [Y_p(y)], \quad (17)$$

where Y_p represents the luminance channel of \mathbf{J} , q is a positive number to adjust the strictness, and $\text{mean}_{y \in \Omega(x)}(\cdot)$ and $\text{std}_{y \in \Omega(x)}(\cdot)$ denote the mean and standard deviation filters, respectively, with $\Omega(x)$ being a square patch centered at x . It is worth noting that Y_p is related to Y_e through (15), and this relation can be exploited to approximate the two terms of (17) as follows:

$$\text{mean}_{y \in \Omega(x)} [Y_p(y)] \approx \frac{1}{t} \left[Y_e \otimes \left(\frac{U_k}{k^2} \right) - \bar{A} \right] + \bar{A}, \quad (18)$$

$$\text{std}_{y \in \Omega(x)} [Y_p(y)] \approx \frac{1}{t} \sqrt{Y_e^2 \otimes \left(\frac{U_k}{k^2} \right) - \left[Y_e \otimes \left(\frac{U_k}{k^2} \right) \right]^2}, \quad (19)$$

where $\bar{A} = (A^R + A^G + A^B)/3$ is the average intensity of atmospheric light \mathbf{A} . Hence, (18) and (19) are substituted back into (17) to obtain the second NBP constraint, as below:

$$t \geq 1 - (\bar{A})^{-1} \left\{ Y_e \otimes \left(\frac{U_k}{k^2} \right) - q \sqrt{Y_e^2 \otimes \left(\frac{U_k}{k^2} \right) - \left[Y_e \otimes \left(\frac{U_k}{k^2} \right) \right]^2} \right\}. \quad (20)$$

Let t_{NBP_1} and t_{NBP_2} denote expressions on the right-hand side of (16) and (20). The NBP constraint t_{NBP} is then expressed as:

$$t_{\text{NBP}} = \max(t_{\text{NBP}_1}, t_{\text{NBP}_2}), \quad (21)$$

where $\max(a, b)$ returns the greater number between a and b . Thus, the transmittance t is constrained between t_{NBP} and unity; that is,

$$t_{\text{NBP}} \leq t \leq 1, \quad (22)$$

and the scene radiance \mathbf{J} is recovered using (15).

As underflows and overflows are inevitable in digital computations, the recovered image suffers from color gamut reduction, rendering a post-processing step highly relevant. The following describes an efficient method for luminance enhancement and color gamut expansion [36]. This method also produces a positive ramification that eases the atmospheric light estimation. More precisely, it can be observed from (15) that \mathbf{A} is proportional to the dehazing power, and $\mathbf{A} = \{255, 255, 255\}$ corresponds to the maximum. Dehazing at this extreme level may worsen the color gamut reduction. It is, however, feasible to use $\mathbf{A} = \{255, 255, 255\}$ in the proposed method because the post-processing step will compensate for the looming problem. Therefore, we adopted $\mathbf{A} = \{255, 255, 255\}$ as the atmospheric light.

C. POST-PROCESSING

Fig. 6 shows the overall block diagram, where the input image is the recovered scene radiance $\mathbf{J} \in \mathbb{R}^{H \times W \times 3}$, and the final output image is denoted as $\mathbf{J}_f \in \mathbb{R}^{H \times W \times 3}$.

1) LUMINANCE ENHANCEMENT

Existing enhancement methods generally operate on the entire luminance range, which may result in over-enhancement. Accordingly, the method in [36] adopted an

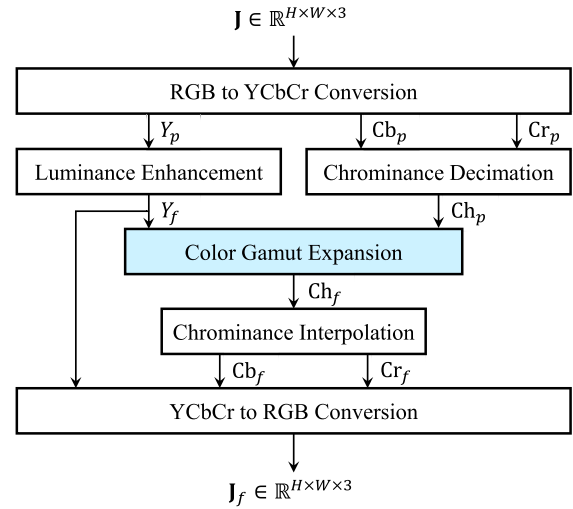


FIGURE 6. Block diagram of the post-processing step.

adaptive limit point (ALP) to constrain the range scene-wisely. Given the luminance channel Y_p of \mathbf{J} , ALP is calculated from the mean \bar{Y}_p and the cumulative distribution function CDF of Y_p as follows:

$$\text{ALP} = \begin{cases} 0.04 + \frac{0.02}{255} (L_{\text{CDF}_{0.9}} - L_{\text{CDF}_{0.1}}) & \bar{Y}_p > 128 \\ 0.04 - \frac{0.02}{255} (L_{\text{CDF}_{0.9}} - L_{\text{CDF}_{0.1}}) & \bar{Y}_p \leq 128, \end{cases} \quad (23)$$

where L_{CDF_k} denotes the luminance value at which $\text{CDF}(L_{\text{CDF}_k}) = k$, with $k \in \mathbb{R}$ and $0 \leq k \leq 1$.

It is worth noting that over-enhancement is avoidable by assigning higher gains to smaller luminance values, and ALP can be exploited for that purpose, as (24) shows:

$$g_1(Y_p) = \frac{Y_p}{2^{21}} \left[255 \left(1 - \frac{Y_p - \text{ALP}}{255} \right)^\theta \left(\frac{255 - Y_p}{255} \right) \right]^2, \quad (24)$$

$$\theta = \frac{1.5 (L_{\text{CDF}_{0.4}} - L_{\text{CDF}_{0.1}})}{\bar{Y}_p - L_{\text{CDF}_{0.1}}} - 0.55, \quad (25)$$

where g_1 is the non-linear luminance gain, 2^{-21} is a normalization factor, and the exponent θ is empirically determined to maximize the tone-mapped image quality index (TMQI) [37]. A linear weight g_2 in (26) is also adopted, where SL and IN are user-defined parameters to adjust the slope and intercept, respectively. The enhanced luminance Y_f is then obtained using (27).

$$g_2(Y_p) = \frac{\text{SL}}{255} Y_p + \text{IN}, \quad (26)$$

$$Y_f = Y_p + g_1(Y_p) \cdot g_2(Y_p). \quad (27)$$

2) COLOR GAMUT EXPANSION

The first block of color space conversion in Fig. 6 produces Y_p , Cb_p , and Cr_p as the luminance, blue-difference chroma,

and red-difference chroma components of \mathbf{J} . However, processing Cb_p and Cr_p separately is unnecessary because the human visual system is less sensitive to color differences than luminance differences [38]. Thus, chrominance subsampling with a 4:2:2 ratio is adopted to combine these two into Ch_p for computational efficiency, as (28), (29), and (30) demonstrate, where $D_{\text{dec}} = [1, 2, 1]/4$ is the decimation filter, and $i \in \{1, 2, \dots, H\}$ and $j \in \{1, 2, \dots, W\}$ are pixel coordinates. Also, as Cb_p , Cr_p , and Ch_p are two-dimensional data, cb_{ij} , cr_{ij} , and ch_{ij} are adopted to denote their component in the i -th row and j -th column, respectively.

$$\text{Cb}_d = \text{Cb}_p \otimes D_{\text{dec}} = \{\text{cb}_{ij} \in \mathbb{R}\}, \quad (28)$$

$$\text{Cr}_d = \text{Cr}_p \otimes D_{\text{dec}} = \{\text{cr}_{ij} \in \mathbb{R}\}, \quad (29)$$

$$\begin{aligned} \text{Ch}_p &= \{\text{ch}_{ij} \in \mathbb{R} \mid \text{ch}_{ij} = \text{cb}_{ij}, \\ &\quad \forall i, j \text{ s.t. } j = \{2n + 1 \mid n \in \mathbb{Z}_0^+\}, \\ &\quad \text{otherwise } \text{ch}_{ij} = \text{cr}_{ij}\}. \end{aligned} \quad (30)$$

According to the Helmholtz-Kohlrausch effect [39], the luminance Y_p is related to the chrominance Ch_p , and luminance enhancement narrows the color gamut in the chromaticity coordinates. Hence, the expansion should be proportional to the ratio between Y_f and Y_p , as expressed by the color gain g_3 below:

$$g_3(Y_p, \text{Ch}_p) = \frac{Y_f}{Y_p} \text{Ch}_p. \quad (31)$$

Moreover, an additional weight g_4 is adopted to maximize the TMQI, and its expression in (32) is determined through experiments.

$$g_4(Y_p) = \begin{cases} 0.7 & Y_p < \text{TH}_1 \\ 0.7 - 0.26 \frac{Y_p - \text{TH}_1}{\text{TH}_2 - \text{TH}_1} & \text{TH}_1 \leq Y_p \leq \text{TH}_2 \\ 0.44 & Y_p > \text{TH}_2, \end{cases} \quad (32)$$

where TH_1 and TH_2 are user-defined parameters to adjust the expansion range. The expanded chrominance Ch_f is then obtained using the following:

$$\text{Ch}_f = \text{Ch}_p + g_3(Y_p, \text{Ch}_p) \cdot g_4(Y_p). \quad (33)$$

Next, the chrominance interpolation block separates Ch_f into two temporary variables, Cb_t and Cr_t , for the final block of color space conversion. To describe chrominance interpolation, we reused ch_{ij} , cb_{ij} , and cr_{ij} to denote a component of Ch_f , Cb_t , and Cr_t in the i -th row and j -th column, respectively. Given $\text{Ch}_f = \{\text{ch}_{ij} \in \mathbb{R}\}$, Cb_t and Cr_t are obtained by interlacing Ch_f with zeros, as (34) and (35) show.

$$\begin{aligned} \text{Cb}_t &= \{\text{cb}_{ij} \in \mathbb{R} \mid \text{cb}_{ij} = \text{ch}_{ij}, \\ &\quad \forall i, j \text{ s.t. } j = \{2n + 1 \mid n \in \mathbb{Z}_0^+\} \\ &\quad \text{otherwise } \text{cb}_{ij} = 0\}, \end{aligned} \quad (34)$$

$$\begin{aligned} \text{Cr}_t &= \{\text{cr}_{ij} \in \mathbb{R} \mid \text{cr}_{ij} = \text{ch}_{ij}, \\ &\quad \forall i, j \text{ s.t. } j = \{2n \mid n \in \mathbb{Z}^+\} \\ &\quad \text{otherwise } \text{cr}_{ij} = 0\}. \end{aligned} \quad (35)$$

After that, Cb_t and Cr_t are convolved with the interpolation filter $D_{\text{int}} = [1, 2, 1]/2$ to get Cb_f and Cr_f . Finally, the image information in the YCbCr color space $\{Y_f, \text{Cb}_f, \text{Cr}_f\}$ is converted back to the RGB color space using (9), yielding the final image \mathbf{I}_f .

IV. RESULTS AND DISCUSSIONS

This section presents a comparative evaluation of the proposed method against nine state-of-the-art benchmarks selected from the three image dehazing categories discussed in Section II. These nine are proposed by He *et al.* [15], Tarel and Hautiere [17], Zhu *et al.* [19], Cai *et al.* [21], Ren *et al.* [3], Liu *et al.* [25], Dong *et al.* [26], Li *et al.* [27], and Yang *et al.* [28], respectively.

A. QUALITATIVE EVALUATION ON REAL-WORLD HAZY IMAGES

Fig. 7 demonstrates a qualitative comparison of ten methods on a real-world hazy image from the IVC dataset [40]. Results by He *et al.* [15], Cai *et al.* [21], Liu *et al.* [25], Dong *et al.* [26], Li *et al.* [27], and Yang *et al.* [28] exhibit a good dehazing performance as the scene radiance has been recovered without any unpleasant artifacts. The result by Zhu *et al.* [19] is slightly over-dehazed, losing dark and distant details. Results by Tarel and Hautiere [17] and Ren *et al.* [3] are less favorable than others because a portion of haze persists.

Above all, it can be observed that nine benchmark methods are ineffective in recovering image details, as witnessed by the traffic light and the man's face in the red-cropped and blue-cropped regions. This common drawback can be explained as follows. Dehazing is fundamentally the subtraction of haze from the input image, and the subtraction degree depends on the transmittance. However, estimating a transmittance with rich details is challenging because spatial filtering usually attenuates high-frequency information. Although an outstanding guided filter [41] has been adopted to refine the transmittance estimate, it is noted that the best guidance image in single image dehazing methods is the input image itself. Accordingly, the lack of an informative guidance image constrains the refinement.

The proposed method, in contrast, effectively removes haze while enhancing the sharpness and the color gamut, as witnessed by the man's face and the facial skin color in the blue-cropped region. This definite advantage is attributed to the pre-processing (unsharp masking) and post-processing (color gamut expansion) steps. The intermediate results in Fig. 8 show that the former has improved image details to such an extent that the contours of distant objects have become noticeable. Meanwhile, the latter, as claimed, has successfully remedied the post-dehazing problem of color gamut reduction.

Fig. 9 shows more qualitative comparison results on real-world hazy images. It can be observed from the first row that the result by He *et al.* [15] is satisfactory, albeit with the post-dehazing false enlargement of the train's headlight.

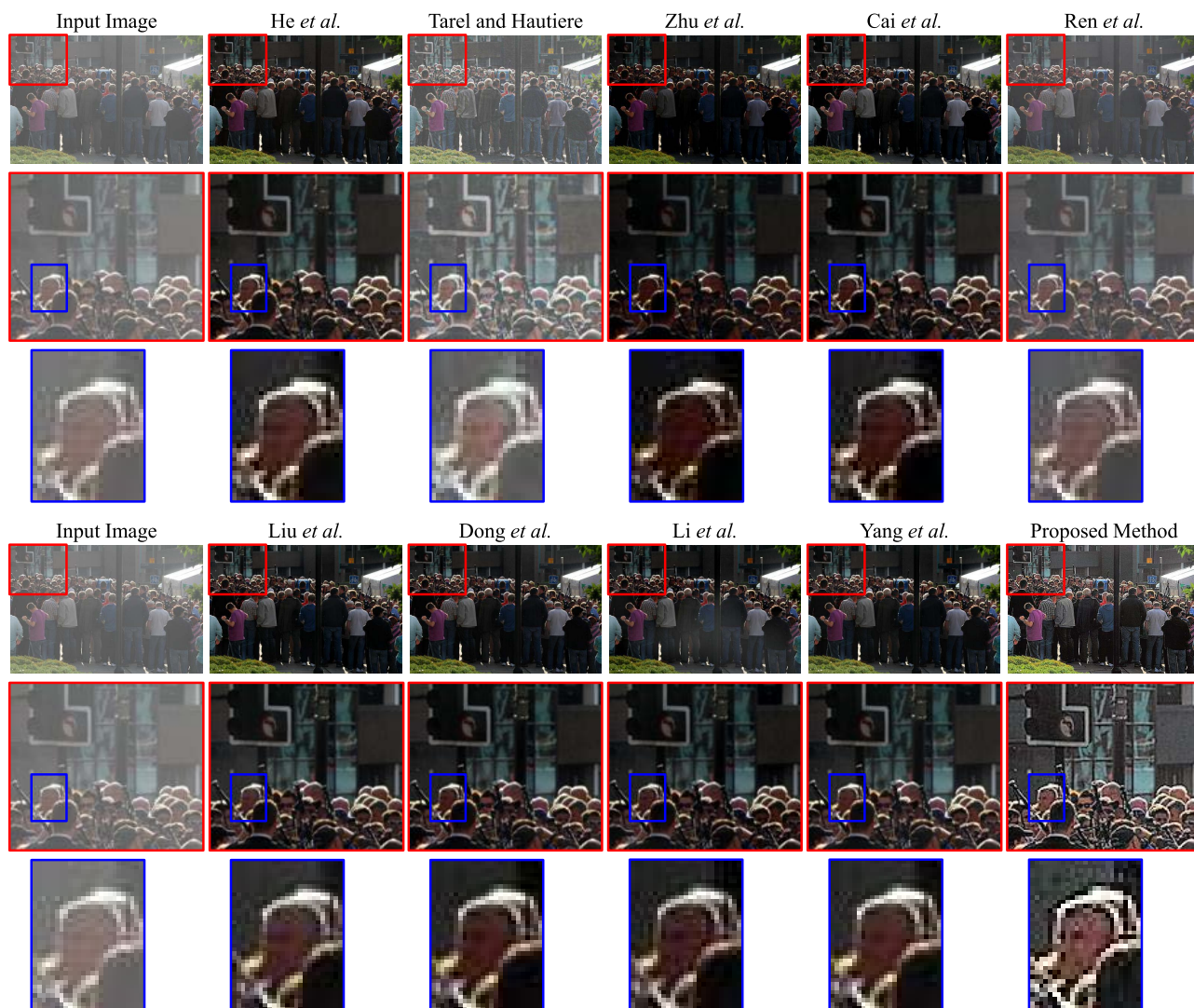


FIGURE 7. Dehazing results of ten methods on a real-world hazy image. From left to right: input image and results by He et al. [15], Tarel and Hautiere [17], Zhu et al. [19], Cai et al. [21], Ren et al. [3], Liu et al. [25], Dong et al. [26], Li et al. [27], Yang et al. [28], and the proposed method. The input image was duplicated for ease of comparison.

This problem occurs when the atmospheric light is less than pixel intensities around the headlight, as discussed in [42]. Accordingly, the maximum value of $\mathbf{A} = \{255, 255, 255\}$ adopted in the proposed method ensures that it is free of that problem. Next, the method of Tarel and Hautiere [17] demonstrates an acceptable performance, but halo artifacts arise at image edges due to the use of large median filters. The result by Zhu et al. [19] suffers from a loss of dark details owing to excessive haze removal. The remaining six deep-learning-based methods perform relatively well, in which results by Cai et al. [21] and Yang et al. [28] are more favorable than others. Our result, as expected, exhibits three desirable outcomes: haze removal, sharpness enhancement, and color gamut expansion.

Similar observations also emerge from the second to the fourth rows of Fig. 9. The dark channel assumption of the method of He et al. [15] does not hold for the sky region,

causing severe color distortion in the second row. As with the interpretation of results in the first row, the method of Tarel and Hautiere [17] suffers from halo artifacts, and the method of Zhu et al. [19] suffers from a loss of dark details. Results by deep-learning-based methods, on the contrary, do not exhibit any unpleasant artifacts, which is attributed to the powerful representation capability and the flexibility of CNNs. Compared with these benchmarks, the proposed method exhibits an almost comparative or even better performance.

B. QUANTITATIVE EVALUATION ON PUBLIC DATASETS

This section presents an objective assessment of the proposed method against nine benchmarks on public image datasets. It is worth noting that there are numerous metrics for image quality assessment, such as the conventional peak signal-to-noise ratio (PSNR), the structural similarity

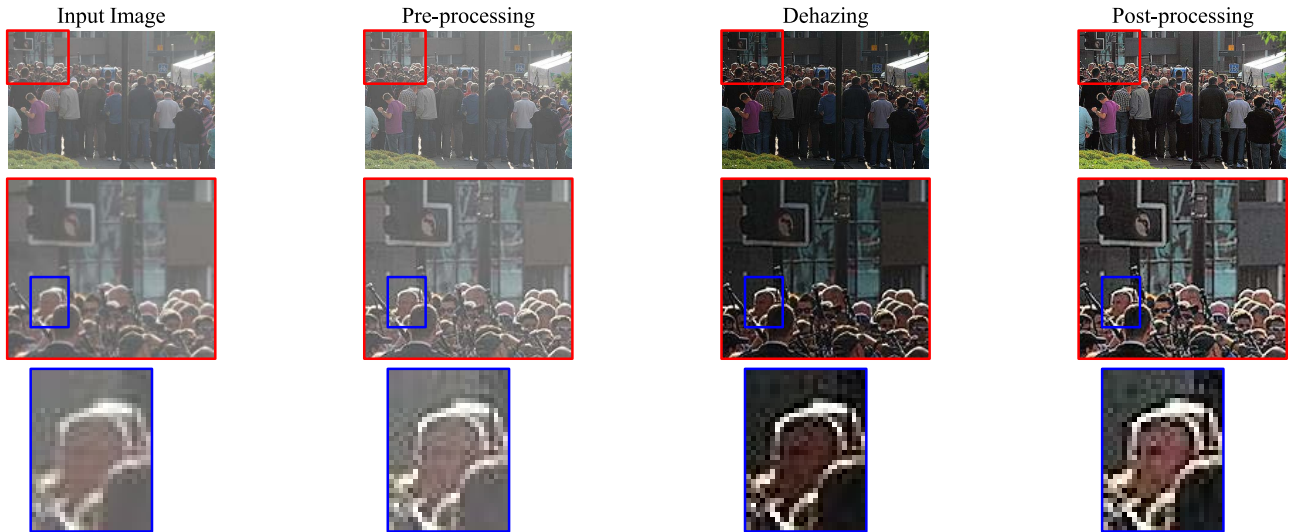


FIGURE 8. Intermediate results of the proposed method on a real-world hazy image. From left to right: input image and results after pre-processing, dehazing, and post-processing steps.

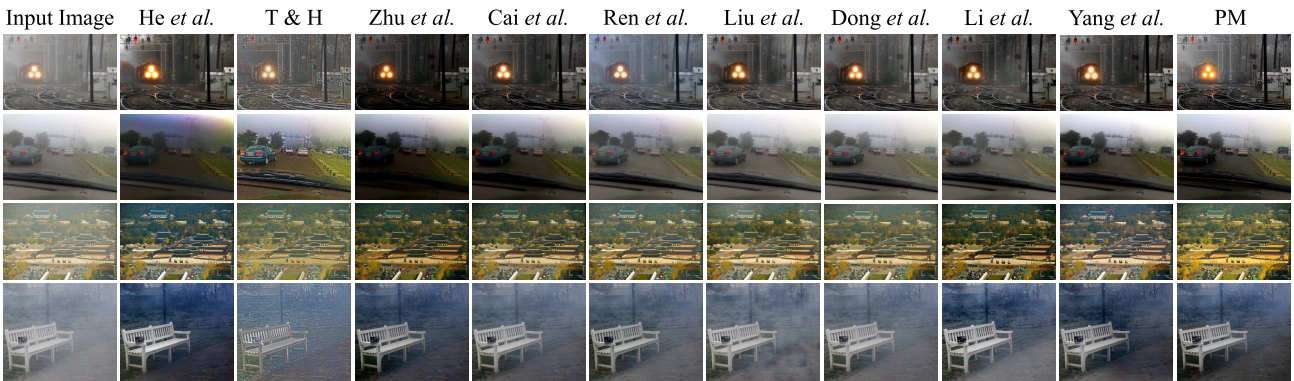


FIGURE 9. Qualitative evaluation of ten methods on real-world hazy images. From left to right: input images and results by He et al. [15], Tarel and Hautiere [17], Zhu et al. [19], Cai et al. [21], Ren et al. [3], Liu et al. [25], Dong et al. [26], Li et al. [27], Yang et al. [28], and the proposed method. We abbreviated the method of Tarel and Hautiere [17] to T & H and the proposed method to PM.

(SSIM) [43], the feature similarity index extended to color images (FSIMc) [44], and the TMQI [37]. The first is pixel-based and thus does not correlate well with subjective ratings. The second, in contrast, is structure-based and can better quantify the perceived image quality. However, it has a drawback in that it utilizes a uniform weight to pool a single quality score. Accordingly, the third improves the SSIM by adopting an adaptive pooling weight and taking account of chrominance information. The fourth improves the SSIM in another direction by considering multi-scale structural similarity and naturalness. Therefore, we selected the FSIMc and the TMQI for our quantitative evaluation due to their high correlation with subjective assessment. These two metrics vary from zero to unity, and higher scores are more favorable in image dehazing. Also, as they are full-reference, their computation requires datasets comprising pairs of hazy and haze-free images.

Table 1 summarizes five public datasets utilized in this evaluation, including FRIDA2 [33], D-HAZY [45], O-HAZE [46], I-HAZE [47], and DENSE-HAZE [48]. The FRIDA2 consists of 66 graphics-generated images depicting

TABLE 1. Summary of image datasets for quantitative evaluation. The symbol # represents quantities.

Dataset	Number of Images (#)		Remark
	Hazy	Haze-free	
FRIDA2	264	66	Synthetic road scenes
D-HAZY	23	23	Synthetic indoors
O-HAZE	45	45	Real-world outdoors
I-HAZE	30	30	Real-world indoors
DENSE-HAZE	55	55	Real-world outdoors/indoors

road scenes, from which four corresponding sets of hazy images are generated, thus a total of 66 haze-free and 264 hazy images. The D-HAZY is generated from the Middlebury [49] and NYU Depth [50] datasets according to the degradation model in (7) with scene depths captured by a Microsoft Kinect camera. It is composed of 1472 pairs of synthetic indoor images, but this evaluation only utilizes 23 pairs from Middlebury due to their substantial variation in image scenes. The other 1449 pairs from NYU Depth portray relatively similar scenes and thus may bias the evaluation results. In contrast, the O-HAZE, I-HAZE, and

TABLE 2. Average evaluation scores in terms of the feature similarity index extended to color images (FSIMc) on five public datasets. Best results are boldfaced, and second-best results are underlined.

Method \ Dataset	FRIDA2	D-HAZY	O-HAZE	I-HAZE	DENSE-HAZE	Overall
He <i>et al.</i> [15]	0.7746	<u>0.9002</u>	0.8423	0.8208	0.6419	0.7746
Tarel and Hautiere [17]	0.7807	0.8703	0.7733	0.8055	0.5598	0.7573
Zhu <i>et al.</i> [19]	0.7918	0.8880	0.7738	0.8252	0.5773	0.7693
Cai <i>et al.</i> [21]	0.7963	0.8874	0.7865	0.8482	0.5573	0.7725
Ren <i>et al.</i> [3]	<u>0.8009</u>	0.8822	0.8553	0.8634	<u>0.6029</u>	<u>0.7896</u>
Liu <i>et al.</i> [25]	0.8003	0.8747	0.8030	0.7416	0.5564	0.7552
Dong <i>et al.</i> [26]	0.7986	0.9059	0.8077	<u>0.8638</u>	0.5650	0.7882
Li <i>et al.</i> [27]	0.7849	0.7383	0.6997	0.7564	0.5763	0.7111
Yang <i>et al.</i> [28]	0.7829	0.8587	<u>0.8459</u>	0.8702	0.5919	0.7899
Proposed Method	0.8016	0.8763	0.8112	0.8586	0.5728	0.7807

TABLE 3. Average evaluation scores in terms of the tone-mapped image quality index (TMQI) on five public datasets. Best results are boldfaced, and second-best results are underlined.

Method \ Dataset	FRIDA2	D-HAZY	O-HAZE	I-HAZE	DENSE-HAZE	Overall
He <i>et al.</i> [15]	0.7291	0.8631	0.8403	0.7319	0.6383	0.7357
Tarel and Hautiere [17]	0.7314	0.8000	0.8416	0.7740	0.5627	0.7294
Zhu <i>et al.</i> [19]	0.7385	<u>0.8206</u>	0.8118	0.7512	0.5955	0.7336
Cai <i>et al.</i> [21]	<u>0.7366</u>	0.7966	0.8413	0.7598	0.5723	0.7312
Ren <i>et al.</i> [3]	0.7232	0.8023	<u>0.8737</u>	<u>0.7819</u>	<u>0.6176</u>	0.7341
Liu <i>et al.</i> [25]	0.6970	0.7938	0.8267	0.6107	0.5196	0.6896
Dong <i>et al.</i> [26]	0.7228	0.7873	0.8342	0.7346	0.5679	0.7294
Li <i>et al.</i> [27]	0.7176	0.6817	0.6566	0.6936	0.5107	0.6520
Yang <i>et al.</i> [28]	0.7158	0.7482	0.8353	0.7322	0.6056	0.7274
Proposed Method	0.7242	0.7841	0.8951	0.8204	0.5921	<u>0.7354</u>

DENSE-HAZE comprise 45, 30, and 55 pairs depicting real-world indoors, outdoors, and both, respectively.

Tables 2 and 3 demonstrate average evaluation scores in FSIMc and TMQI on five public datasets. It can be observed from Table 2 that the proposed method is ranked fourth overall, below the deep dehazing models of Yang *et al.* [28], Ren *et al.* [3], and Dong *et al.* [26]. However, it is worth noting that the difference between its FSIMc score and the best is subtle. Specifically, it outperforms other benchmarks on FRIDA2 and is within the top four methods for dehazing real-world images in O-HAZE and I-HAZE. Nevertheless, its performance is slightly under-par on D-HAZY and DENSE-HAZE. This observation can be related to the fact that FSIMc quantifies the degradation rather than the enhancement. Thus, it is unsurprising that two models of Ren *et al.* [3] and Dong *et al.* [26], trained on fully annotated datasets to minimize the difference between their predictions and corresponding ground-truth references, have achieved the top scores. This interpretation can be further supported by the unimpressive score of the unsupervised model of Li *et al.* [27], which does not require ground-truth references. However, this and the best-performing model of Yang *et al.* [28] have shown the great potential of unsupervised and unpaired learning in computer vision.

According to Table 3, the proposed method is ranked second overall, and its TMQI score only differs from the best at the fourth decimal place. More specifically, the proposed method exhibits a comparative performance on FRIDA2 and an under-par performance on D-HAZY. In contrast, it outperforms benchmark methods on real-world datasets, such

as O-HAZE and I-HAZE, as witnessed by a significant difference in TMQI scores. Hence, unsharp masking and color gamut expansion appear to benefit real-world images. However, such benefits as these two steps offer do not suffice for handling densely hazy images owing to the under-performance of the dehazing step.

As a result, it can be concluded that the proposed method demonstrates a comparative performance with state-of-the-art benchmarks, notably the deep learning models of Yang *et al.* [28], Ren *et al.* [3], and Dong *et al.* [26].

C. PROCESSING TIME COMPARISON

Notwithstanding a comparative performance, the proposed method possesses a linear time complexity, $\mathcal{O}(N)$ or $\mathcal{O}(H \times W)$, where H and W denote the image's height and width. According to Section III, the most computationally intensive operations are the mean filter in the pre-processing step and the modified hybrid median filter in the dehazing step. The implementation of these two filters affects the entire algorithm's complexity directly. Let us assume that the kernel size is $S_h \times S_w$. Naive implementations result in $\mathcal{O}(H \times W \times S_h \times S_w)$ complexity. Fortunately, $\mathcal{O}(H \times W)$ implementations of those two filters are available in [51] and [18]; therein lies the proposed method's linear time complexity.

Table 4 summarizes the processing time of ten methods on different image resolutions, ranging from VGA (640×480) to 8K UHD (7680×4320). As source codes of nine benchmarks are publicly available, we used them and adopted the parameter configuration provided by their authors. This measurement was conducted in MATLAB R2019a and Python 3.9.9

TABLE 4. Processing time in seconds of ten methods on different image resolutions. Best results are boldfaced, second-best results are underlined, and NA stands for not available with the underlying cause, REx (RAM Exhaustion), MEx (Memory Exhaustion), or RTE (Run-Time Error), in parentheses.

Method	Resolution					
	VGA 640 × 480	SVGA 800 × 600	HD 1024 × 768	FHD 1920 × 1080	DCI 4K 4096 × 2160	8K UHD 7680 × 4320
He <i>et al.</i> [15]	12.64	19.94	32.37	94.25	470.21	NA (REx)
Tarel and Hautiere [17]	0.28	0.59	0.76	1.51	9.02	36.54
Zhu <i>et al.</i> [19]	<u>0.22</u>	0.34	0.64	1.51	6.39	25.20
Cai <i>et al.</i> [21]	1.53	2.39	3.88	10.68	47.35	178.81
Ren <i>et al.</i> [3]	0.54	0.88	1.53	3.43	17.90	245.28
Liu <i>et al.</i> [25]	12.05	0.02	0.03	0.07	0.13	NA (MEx)
Dong <i>et al.</i> [26]	18.69	NA (RTE)	0.36	1.23	<u>2.52</u>	<u>20.48</u>
Li <i>et al.</i> [27]	188.03	398.28	728.83	1875.56	NA (MEx)	NA (MEx)
Yang <i>et al.</i> [28]	16.15	5.14	5.71	7.51	8.67	24.00
Proposed Method	0.10	<u>0.17</u>	<u>0.27</u>	<u>0.68</u>	2.95	11.16

(with PyTorch 1.12.0+cu116), both running on a computer with an Intel Core i9-9900K (3.6 GHz) CPU, 64 GB RAM, and an Nvidia TITAN RTX.

It emerges that the method of He *et al.* [15] is the least efficient in terms of time and memory. This finding is consistent with its widely known drawback rooted in soft-matting. Notably, RAM was exhausted when invoking this algorithm on an 8K UHD image, and the processing time, in this case, was denoted as not available. Similarly, the unsupervised model of Li *et al.* [27] could not process DCI 4K and 8K UHD images owing to memory exhaustion. This model progressively refines the dehazing result, and its default configuration is to run through 800 iterations. Therefore, its processing time is significantly larger than those of other methods.

Next on the list are two models of Cai *et al.* [21] and Ren *et al.* [3]. As discussed earlier, they are deemed to be overkill due to the high computational cost inherent in them. Measurements in Table 4 then verified that claim. However, recent models of Liu *et al.* [25], Dong *et al.* [26], and Yang *et al.* [28] benefited from batch processing and parallel computing. Under these mechanisms, PyTorch needs to initialize the GPU, for example, making replicas of the model on each GPU worker. Accordingly, the execution time of the first image was prolonged, whereas those of the remaining images were substantially shortened. Notably, the model of Liu *et al.* [25] utilized four GPU workers and thus consumed the least processing time for SVGA, HD, FHD, and DCI 4K resolutions. It is also worth noting that batch processing and parallel computing generally cause a jump in memory consumption, proportional to the number of GPU workers. Thus, the model of Liu *et al.* [25] suffered from memory exhaustion when handling an 8K UHD image. Conversely, the model of Dong *et al.* [26] was free of that problem but returned a run-time error when processing an SVGA image.

The two methods of Tarel and Hautiere [17] and Zhu *et al.* [19] are well-recognized for their computational efficiency, thus accounting for their fast processing speed recorded in Table 4. Additionally, it is noteworthy that Zhu *et al.* [19] adopted the fast implementation of the guided filter, which downscaled the input image to ease the computational burden. If they utilize the standard guided filter, their method would be slower than that of Tarel and Hautiere [17].

TABLE 5. Comparisons on five public datasets for different variants of the proposed method. Best results are boldfaced.

Variant	Five public datasets	
	FSIMc	TMQI
Without pre-processing	0.7765	0.7197
Without post-processing	0.7610	0.7131
Without both	0.7597	0.6979
Our full algorithm	0.7807	0.7354

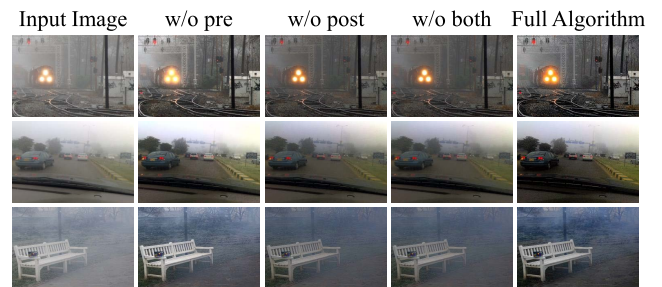


FIGURE 10. Comparisons on real-world hazy images for different variants of the proposed method.

The proposed method is ranked second overall, notably without batch processing or parallel computing but simply sequential computing. Although it is slower than the fastest model of Liu *et al.* [25], Tables 2 and 3 demonstrate that it outperforms this model under FSIMc and TMQI. Also, compared with the fast sequential method of Zhu *et al.* [19], it achieved approximately 2.2× speedup for two main reasons. Firstly, the proposed method skips the atmospheric light estimation. Secondly, it only makes six calls to three different $\mathcal{O}(N)$ spatial filters, including a call to a 3×3 Laplacian filter in (13), four calls to the box filter in (11) and (20), and a call to the modified hybrid median filter in the scene depth refinement step. On the contrary, the method of Zhu *et al.* [19] needs to estimate the atmospheric light and make eighteen calls to the box filter inside the fast-guided filter. This difference accounts for a big gap between the processing time of the two methods. Hence, the definite advantage of a low computational cost is attributed to the elegant partition of image dehazing into three essential steps: pre-processing, dehazing, and post-processing, where each can be implemented using traditional computer vision techniques.

D. ABLATION STUDY

The proposed method consists of three steps that operate in a complementary manner. To verify the individual contribution of each step, we conduct ablation studies by considering three variants of our algorithm. They are created by dropping the pre-processing step, the post-processing step, and both, respectively. Table 5 summarizes the evaluation results on five public datasets under FSIMc and TMQI. It can be observed that the pre-processing step (unsharp masking) contributes more to the structural information, while the post-processing step (color gamut expansion) contributes more to the naturalness. Hence, each of these two steps plays an essential role in the proposed method, which justifies its significance.

Moreover, the contributions of the pre-processing and post-processing steps can also be verified by qualitative results in Fig. 10. Excluding the former causes a loss of image details, and excluding the latter gives rise to color gamut reduction. Image quality further worsens when neither of them is included. Accordingly, these observations verify the whole algorithm with the pre-processing, dehazing, and post-processing steps.

V. CONCLUSION

This paper presented an efficient method for single image dehazing in linear time. It began with a literature review arguing that deep learning models were overkill, and dehazing by traditional computer vision techniques could achieve a comparative performance at a much lower computational cost. After that, a detailed description of the proposed method was followed. The pre-processing step enhances image sharpness, and the dehazing step recovers the scene radiance according to the improved color attenuation prior. Finally, the post-processing step compensates for the color gamut reduction. Subjective and objective evaluation against nine benchmarks demonstrated that the proposed method was substantially fast while achieving a comparative performance.

Nonetheless, under-performance was observed on densely hazy images. This drawback is common for methods developed from hand-engineered features, which are not abstract enough to reflect how human visual systems recognize dense haze. In this case, image inpainting and conditional image generation may be viable alternatives, but it appears challenging to implement them computationally efficiently. Thus, this task is left for future research.

REFERENCES

- [1] E. J. McCartney, "Optics of the atmosphere: Scattering by molecules and particles," *Phys. Today*, vol. 30, no. 5, p. 76, May 1977, doi: 10.1063/1.3037551.
- [2] Y. Pei, Y. Huang, Q. Zou, X. Zhang, and S. Wang, "Effects of image degradation and degradation removal to CNN-based image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 4, pp. 1239–1253, Apr. 2021, doi: 10.1109/TPAMI.2019.2950923.
- [3] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *Proc. ECCV*, Amsterdam, The Netherlands, 2016, pp. 154–169.
- [4] J. P. Oakley and B. L. Satherley, "Improving image quality in poor visibility conditions using a physical model for contrast degradation," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 167–179, Feb. 1998, doi: 10.1109/83.660994.
- [5] J. Hadamard, "Lectures on Cauchy's problem in linear partial differential equations," *Phys. Today*, vol. 6, no. 8, p. 18, Aug. 1953, doi: 10.1063/1.3061337.
- [6] T. Liang, Z. Li, Y. Ren, Q. Mao, and W. Zhou, "A progressive single-image dehazing network with feedback mechanism," *IEEE Access*, vol. 9, pp. 158091–158097, 2021, doi: 10.1109/ACCESS.2021.3130468.
- [7] S. A. Hovhannisyanyan, H. A. Gasparyan, S. S. Agaian, and A. Ghazaryan, "AED-Net: A single image dehazing," *IEEE Access*, vol. 10, pp. 12465–12474, 2022, doi: 10.1109/ACCESS.2022.3144402.
- [8] C. Li, Y. He, and X. Li, "Feature attention parallel aggregation network for single image haze removal," *IEEE Access*, vol. 10, pp. 15322–15335, 2022, doi: 10.1109/ACCESS.2022.3148167.
- [9] X. Jiang, L. Lu, M. Zhu, Z. Hao, and W. Gao, "Haze relevant feature attention network for single image dehazing," *IEEE Access*, vol. 9, pp. 106476–106488, 2021, doi: 10.1109/ACCESS.2021.3100604.
- [10] H. Liu and J. Chen, "Indirect domain shift for single image dehazing," *IEEE Access*, vol. 9, pp. 122959–122970, 2021, doi: 10.1109/ACCESS.2021.3110428.
- [11] D. Yang and J. Sun, "A model-driven deep dehazing approach by learning deep priors," *IEEE Access*, vol. 9, pp. 108542–108556, 2021, doi: 10.1109/ACCESS.2021.3101319.
- [12] J. Lin, W. M. Chen, H. Cai, C. Gan, and S. Han, "MCUNetV2: Memory-efficient patch-based inference for tiny deep learning," in *Proc. NIPS*, 2021.
- [13] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Advances in Computer Vision*, vol. 943, K. Arai and S. Kapoor, Eds. Cham, Switzerland: Springer, 2020, pp. 128–144.
- [14] D. Ngo, S. Lee, T. M. Ngo, G.-D. Lee, and B. Kang, "Visibility restoration: A systematic review and meta-analysis," *Sensors*, vol. 21, no. 8, p. 2625, Apr. 2021, doi: 10.3390/s21082625.
- [15] 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, doi: 10.1109/TPAMI.2010.168.
- [16] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 228–242, Feb. 2008, doi: 10.1109/TPAMI.2007.1177.
- [17] J. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *Proc. ICCV*, Kyoto, Japan, 2009, pp. 2201–2208.
- [18] S. Perreault and P. Hebert, "Median filtering in constant time," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2389–2394, Sep. 2007, doi: 10.1109/TIP.2007.902329.
- [19] 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, doi: 10.1109/TIP.2015.2446191.
- [20] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. CVPR*, Columbus, OH, USA, Jun. 2014, pp. 2995–3002.
- [21] 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, doi: 10.1109/TIP.2016.2598681.
- [22] J. Zhang and D. Tao, "FAMED-Net: A fast and accurate multi-scale end-to-end dehazing network," *IEEE Trans. Image Process.*, vol. 29, pp. 72–84, Jul. 2019, doi: 10.1109/TIP.2019.2922837.
- [23] Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang, "Domain adaptation for image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 2805–2814.
- [24] S.-C. Huang, T.-H. Le, and D.-W. Jaw, "DSNet: Joint semantic learning for object detection in inclement weather conditions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2623–2633, Aug. 2021, doi: 10.1109/TPAMI.2020.2977911.
- [25] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. ICCV*, Seoul, South Korea, 2019, pp. 7313–7322.
- [26] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in *Proc. CVPR*, Seattle, WA, USA, Jun. 2020, pp. 2154–2164.

- [27] B. Li, Y. Gou, S. Gu, J. Z. Liu, J. T. Zhou, and X. Peng, "You only look yourself: Unsupervised and untrained single image dehazing neural network," *Int. J. Comput. Vis.*, vol. 129, no. 5, pp. 1754–1767, May 2021, doi: [10.1007/s11263-021-01431-5](https://doi.org/10.1007/s11263-021-01431-5).
- [28] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, and D. Tao, "Self-augmented unpaired image dehazing via density and depth decomposition," in *Proc. CVPR*, New Orleans, LA, USA, Jun. 2022, pp. 2037–2046.
- [29] B. Li, Y. Gou, J. Z. Liu, H. Zhu, J. T. Zhou, and X. Peng, "Zero-shot image dehazing," *IEEE Trans. Image Process.*, vol. 29, pp. 8457–8466, 2020, doi: [10.1109/TIP.2020.3016134](https://doi.org/10.1109/TIP.2020.3016134).
- [30] R. K. Vincent, "An ERTS multispectral scanner experiment for mapping iron compounds," in *Proc. ISRSE*, Ann Arbor, MI, USA, 1972, pp. 1239–1247.
- [31] D. Ngo, G.-D. Lee, and B. Kang, "Improved color attenuation prior for single-image haze removal," *Appl. Sci.*, vol. 9, no. 19, p. 4011, Sep. 2019, doi: [10.3390/app9194011](https://doi.org/10.3390/app9194011).
- [32] K. Jack, "Color spaces," in *Video Demystified: A Handbook for the Digital Engineer*, 3rd ed. Eagle Rock, VA, USA: LLH Technology Publishing, 2001, p. 19, ch. 3.
- [33] J. P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," *IEEE Intell. Transp. Syst. Mag.*, vol. 4, no. 2, pp. 6–20, Apr. 2012, doi: [10.1109/MITS.2012.2189969](https://doi.org/10.1109/MITS.2012.2189969).
- [34] D. Ngo, S. Lee, U.-J. Kang, T. M. Ngo, G.-D. Lee, and B. Kang, "Adapting a dehazing system to haze conditions by piece-wisely linearizing a depth estimator," *Sensors*, vol. 22, no. 5, p. 1957, Mar. 2022, doi: [10.3390/s22051957](https://doi.org/10.3390/s22051957).
- [35] D. Ngo, G.-D. Lee, and B. Kang, "A 4K-capable FPGA implementation of single image haze removal using hazy particle maps," *Appl. Sci.*, vol. 9, no. 17, p. 3443, Aug. 2019, doi: [10.3390/app9173443](https://doi.org/10.3390/app9173443).
- [36] H. Cho, G.-J. Kim, K. Jang, S. Lee, and B. Kang, "Color image enhancement based on adaptive nonlinear curves of luminance features," *J. Semicond. Technol. Sci.*, vol. 15, no. 1, pp. 60–67, Feb. 2015, doi: [10.5573/jsts.2015.15.1.060](https://doi.org/10.5573/jsts.2015.15.1.060).
- [37] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013, doi: [10.1109/TIP.2012.2221725](https://doi.org/10.1109/TIP.2012.2221725).
- [38] S. Winkler, C. J. van den Branden Lambrecht, and M. Kunt, "Vision and video: Models and applications," in *Vision Models and Applications to Image and Video Processing*. Berlin, Germany: Springer, 2001, p. 209, ch. 10.
- [39] M. D. Fairchild, "Color appearance phenomena," in *Color Appearance Models*, 3rd ed. West Sussex, U.K.: Wiley, 2013, pp. 123–125, ch. 6.
- [40] K. Ma, W. Liu, and Z. Wang, "Perceptual evaluation of single image dehazing algorithms," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Quebec City, QC, Canada, Sep. 2015, pp. 3600–3604.
- [41] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013, doi: [10.1109/TPAMI.2012.213](https://doi.org/10.1109/TPAMI.2012.213).
- [42] D. Ngo, S. Lee, G.-D. Lee, and B. Kang, "Single-image visibility restoration: A machine learning approach and its 4K-capable hardware accelerator," *Sensors*, vol. 20, no. 20, p. 5795, Oct. 2020, doi: [10.3390/s20205795](https://doi.org/10.3390/s20205795).
- [43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [44] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011, doi: [10.1109/TIP.2011.2109730](https://doi.org/10.1109/TIP.2011.2109730).
- [45] C. Ancuti, C. O. Ancuti, and C. De Vleeschouwer, "D-HAZY: A dataset to evaluate quantitatively dehazing algorithms," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 2226–2230.
- [46] C. O. Ancuti, C. O. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Salt Lake City, UT, USA, Jun. 2018, pp. 867–8678.
- [47] C. Ancuti, C. O. Ancuti, and R. Timofte, "NTIRE 2018 challenge on image dehazing: Methods and results," *Proc. CVPRW*, Salt Lake City, UT, USA, Jun. 2018, pp. 1004–100410.
- [48] C. O. Ancuti, C. Ancuti, M. Sbert, and R. Timofte, "Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 1014–1018.
- [49] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nešić, X. Wang, and P. Westling, "High-resolution stereo datasets with subpixel-accurate ground truth," in *Pattern Recognition (Lecture Notes in Computer Science)*, vol. 8753, X. Jiang, J. Hornegger, and R. Koch, Eds. Cham, Switzerland: Springer, 2014, pp. 31–42.
- [50] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Computer Vision (Lecture Notes in Computer Science)*, vol. 7576, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin, Germany: Springer, 2012, pp. 31–42.
- [51] P. Viola and M. Jones, "Robust real-time face detection," in *Proc. 8th IEEE Int. Conf. Comput. Vis. ICCV*, Vancouver, BC, Canada, Oct. 2001, p. 747.



DAT NGO received the B.S. degree in computer engineering from The University of Danang—University of Science and Technology, Danang, Vietnam, in 2016, and the M.S. and Ph.D. degrees in electronic engineering from Dong-A University, Busan, South Korea, in 2018 and 2021, respectively.

He is a Postdoctoral Researcher with the Technical Support and Development Center for Display Device Convergence Technology, Busan. His research interests include image/video processing, machine learning, deep learning, and FPGA prototypes.



GI-DONG LEE received the B.S., M.S., and Ph.D. degrees in electronic engineering from Busan National University, Busan, South Korea, in 1989, 1991, and 2000, respectively.

He was a Postdoctoral Researcher at the Liquid Crystal Institute, Kent State University, OH, USA, until 2003. Since 2003, he has been with the Department of Electronic Engineering, Dong-A University, Busan, where he is a Full Professor. He is also the Director of the Technical Support and Development Center for Display Device Convergence Technology, Busan. His research interests include 3D stereoscopic, auto-stereoscopic, transparent, public, advanced/high-technology displays, and image processing for high-performance displays.



BONGSOON KANG received the B.S. degree in electronic engineering from Yonsei University, Seoul, South Korea, in 1985, the M.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, USA, in 1987, and the Ph.D. degree in electrical engineering from Drexel University, Philadelphia, USA, in 1990.

He was a Senior Staff Researcher at Samsung Electronics, Suwon, South Korea, from 1989 to 1999. Since 1999, he has been with the Department of Electronic Engineering, Dong-A University, Busan, South Korea, where he is a Full Professor. He is also a member with the Technical Support and Development Center for Display Device Convergence Technology, Busan. His research interests include image/video processing, machine learning, deep learning, FPGA prototypes, and SoC/VLSI designs.

...