

Joint Contrast Enhancement and Exposure Fusion for Real-World Image Dehazing

Xiaoning Liu^{ID}, Hui Li^{ID}, and Ce Zhu^{ID}, *Fellow, IEEE*

Abstract—Due to the complexity of real environment and potential defects of current simulation datasets, either prior-based or deep learning-based single image dehazing methods may not work well in certain scenarios. In this work, we propose an efficient joint contrast enhancement and exposure fusion (CEEF) framework to formulate image dehazing task as a problem of enhancing local visibility and global contrast. In the contrast enhancement stage, several intermediate images are generated through two pre-processing steps. Specifically, gamma correction (GC) is used to adjust local visibility of an input hazy image. To address the issue of applying adaptive histogram equalization (AHE) to each color channel independently, we introduce color-preserving AHE (CP-AHE) to improve global contrast of the input hazy image. In the fusion stage, we develop a fast structural patch decomposition-based fusion strategy with an adaptive kernel size to fuse the inputs obtained by GC and CP-AHE. Extensive experiments on the real-world datasets demonstrate superiority of the proposed method to state-of-the-art methods in terms of visual and quantitative evaluation. Particularly for nighttime hazy scenes, our approach is shown to retain fine details and reduce color artifacts against three latest nighttime defogging methods. Moreover, we discuss potential applications of our CP-AHE in low-light enhancement and image editing.

Index Terms—Image dehazing, multi-exposure fusion, structural patch decomposition, nighttime scene, image enhancement.

I. INTRODUCTION

IMAGE dehazing is aimed at restoring a clean image from an input degraded one captured in foggy scenes, which has improved the performance of visual tasks significantly, such as scene understanding [1] and object detection [2]. The past two decades have witnessed dramatic developments in this area [2]–[4]. Most of the early image dehazing methods were based on atmospheric scattering model [5], *i.e.*,

$$I(x) = J(x)t(x) + A(1 - t(x)), \quad (1)$$

Manuscript received 3 December 2020; revised 22 April 2021 and 11 August 2021; accepted 19 August 2021. Date of publication 14 September 2021; date of current version 9 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grants U19A2052 and 62020106011. The Associate Editor coordinating the review of this manuscript and approving it for publication was Prof. F. Sohel. (*Corresponding Author*: Hui Li.)

Xiaoning Liu and Ce Zhu are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: liuxiaoning2016@sina.com; eczhu@uestc.edu.cn).

Hui Li is with the Imaging Algorithm Research Department, Vivo Mobile Communication Company Ltd., Shenzhen 518101, China (e-mail: li-hui@vivo.com).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMM.2021.3110483>.

Digital Object Identifier 10.1109/TMM.2021.3110483

where $I(x)$ is the observed hazy image, $J(x)$ denotes the ideal haze-free image, A describes the global atmospheric light affected by ambient illumination, and $t(x)$ represents the transmission map. The transmission map can be expressed as $t(x) = e^{-\beta d(x)}$, where $d(x)$ is the scene depth between the camera and the actual object, β is called the atmospheric scattering coefficient, and x represents the pixel coordinate.

It can be obviously seen from the above physical model that restoring J is an ill-posed problem, because both $t(x)$ and A are unknown. To recover the fog-free image J , early prior-based methods [6]–[9] tried to estimate the transmission map $t(x)$ or atmospheric light intensity A . Whether it is dark channel prior [8] or color-lines prior [9], all these assumptions cannot accurately estimate their values, due to the ambiguity of airlight-albedo [10]. Therefore, these methods perform poorly in certain scenes when corresponding prior conditions do not hold [11]. Computational complexity is, in addition, another major problem for such methods [12].

As deep learning technology develops by leaps and bounds in recent years, convolutional neural networks (CNN) [13]–[22] have been hired to estimate the transmission map and atmospheric light in lieu of prior constraints. Some of these methods, however, also have recourse to traditional methods [8], [23], [24] to estimate atmospheric light. Although CNN-based methods have achieved significant performance compared to prior-based approaches, they are also confronted with inaccurate transmission map estimation [25]. To overcome this defect, some end-to-end methods [16], [25]–[34] recover the haze-free image directly from the input fog image. Deep learning-based methods, however, require a large number of image pairs, which is difficult to obtain in practice. Although some datasets [2], [35]–[37] currently provide such image pairs of real scenes, the corresponding fog images are generated from artificial fog environments.

Since foggy images usually have low contrast and visibility, some works have relied on image enhancement techniques to remove haze, such as Retinex [38] and image fusion [39]–[43]. Recently, Galdran *et al.* [38] proved that there is a simple linear relationship between Retinex and image dehazing when assuming that A in Eq. (1) is a specific vector, *i.e.*, $A=(1, 1, 1)$. Nevertheless, restored colors sometimes appear unnatural or over-exposed. This is mainly due to the fact that Retinex performs independently per-channel and such implementation ignores the role of atmospheric light A . Ancuti *et al.* [39] first proposed a multi-scale fusion strategy based on Laplacian pyramid decomposition [44] to remove haze. The strategy includes two input images, namely white balance and contrast enhancement of the foggy image. Follow-up works [41]–[43] have improved the visibility of foggy images. However, most existing

fusion-based methods are pixel-level and weight maps in foggy conditions often contain a lot of noise, which may result in various artifacts (*e.g.*, color cast [39], pseudo-color [41], [42], and glowing effects [43]) in the final fusion result.

From the perspective of image enhancement, we propose an efficient joint contrast enhancement and exposure fusion (CEEF) framework for removing the haze of real-world scenes. More concretely, taking the characteristics of gamma correction (GC) into consideration, we use GC to generate multiple pseudo-exposure images. And through a simple and novel reshaping strategy, we provide a solution for clipped adaptive histogram equalization (AHE) [45] to directly process the input color hazy image. To integrate these derived images, we formulate an adaptive fusion strategy based upon our recent work [46]. Our contributions are summarized as follows:

- In contrast to the operation of AHE processing color hazy images by either channel independence or converting to other color spaces, we provide a simple yet effective reshaping strategy to enhance contrast. The novel manner is able to simultaneously consider the correlation among three channels and vividly preserve the color of fog-free images. In addition, the proposed color-preserving AHE (CP-AHE) is capable of processing images in real time, such as low-light image enhancement and image editing.
- We develop a fast multi-scale structural patch decomposition based fusion strategy for combining derived inputs. The method performs fusion on each image patch, which is conducive to noise reduction on the generated weight map. The adaptive filter kernel size at each scale can better preserve fine details. Furthermore, we consider gradient information when designing the weight of base component for better structure retention.
- CEEF offers flexibility and possibility in the actual fog scene, which can even handle ultra-high-definition or 4K resolution images. Experimental results of daytime and nighttime hazy images in real-world scenes validate the advantages of the proposed method in terms of qualitatively and quantitatively compared to state-of-the-art dehazing methods.

The rest of the paper is organised as follows. Section II briefly reviews the related literature on image dehazing. Section III presents the proposed dehazing method in detail. Experimental results and evaluations are presented in Section IV. The final Section V summarizes the work.

II. RELATED WORK

In this section, we briefly review prior-based, deep learning-based, and fusion-based image dehazing methods.

Prior-based methods. Since atmospheric scattering model of Eq. (1) is an ill-posed problem, most existing methods [6]–[9], [12], [23], [24], [47]–[55] use a variety of priors or assumptions to estimate transmission map and atmospheric light. Specifically, He *et al.* [8] proposed a dark channel prior (DCP) derived from the statistics of outdoor haze-free images. Subsequent variants [12], [52], [54] improve the performance and efficiency of DCP. Fattal [9] used color-lines to recover the scene transmission, which relies on a general assumption that pixels of small image patches usually present a one-dimensional distribution in

the RGB space. Similarly, Berman *et al.* [49] developed a novel non-local prior (NLP), which assumes that pixels in a given cluster are often non-local. Zhu *et al.* [24] proposed a novel color attenuation prior (CAP) and used the prior to build a linear model for scene depth of the hazy image. However, such methods generally have high computational complexity, and their performances will inevitably be degraded since the above assumptions cannot always be true in real scenes [25].

Deep learning-based methods. With the rise of deep learning, CNN-based dehazing methods have achieved significant performance. Such methods can be roughly divided into two categories [56]: model-based and model-free methods. Model-based methods [13]–[19], [21], [22] use deep CNN to directly estimate transmission maps or atmospheric light. Cai *et al.* [13] proposed an end-to-end network, DehazeNet, to estimate medium transmission map. Zhang *et al.* [16] similarly developed a densely connected pyramid dehazing network to jointly learn the transmission map, atmospheric light and haze-free image. Under a reformulation of Eq. (1), Li *et al.* [15] designed AOD-Net to bypass the estimation of above parameters. Model-free methods [16], [25]–[30], [32]–[34] directly learn an end-to-end mapping relationship to recover a haze-free image without atmospheric scattering model. Liu *et al.* [28] proposed an end-to-end GridDehazeNet consisting of three parts: pre-processing, backbone, and post-processing. Qu *et al.* [29] developed an enhanced pix2pix dehazing network (EPDN) that considers image dehazing as an image-to-image translation problem. To reduce the discrepancy between synthetic and real domains, Shao *et al.* [33] recently proposed a domain adaptation framework.

Fusion-based methods. Fusion-based methods mainly consist of two stages: generating pseudo inputs and fusing them. Ancuti *et al.* [39] first proposed a multi-scale fusion framework for removing haze, which uses the results of white balance and contrast enhancement as inputs. Various variants [42], [43] have been proposed to improve the dehazing performance. Guo *et al.* [41] proposed a guided filter-based weighted scheme that combines atmospheric light estimated with Gaussian-based DCP and fined details obtained by the improved adaptive histogram equalization (AHE) [45]. Recently, the combination of feature fusion and deep learning has gradually emerged in the field of image dehazing [25], [26], [57]–[59]. Ren *et al.* [26] proposed a gated fusion network to merge three inputs derived from a hazy image by implementing white balance, contrast enhancement, and gamma correction jointly. In contrast to existing pixel-level-based fusion approaches [39]–[43], the proposed patch-level-based CEEF that takes the correlation of RGB channels into consideration obtains better robustness, more vivid color and lower complexity.

III. PROPOSED METHOD

Fig. 1 shows the proposed framework for image dehazing. First, we employ GC to improve local visibility of input hazy image by generating a set of pseudo-exposure images, and then propose CP-AHE to enhance global contrast of the input hazy image. Finally, we blend these derived inputs via a fast patch-based fusion method for dehazing.

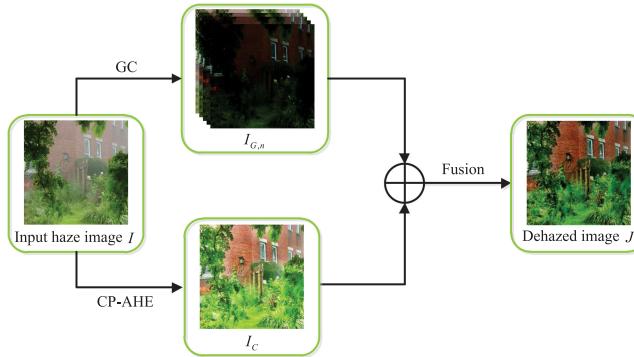


Fig. 1. The framework of CEEF.

A. Derived Inputs

Gamma Correction. Due to its mathematical tractability, GC is used frequently for dynamic range adjustment. Although GC is capable of improving local visibility of hazy images to a certain extent [26], [42], [54], it is difficult to strike a balance between over-saturation of narrow-range regions and haze removal of wide-range regions. Fig. 2 shows that as gamma increases ($\gamma > 1$), more fine detail of trees in the wide-range regions becomes visible (blue zoomed-in area), while the grass in the narrow-range regions begins to reduce contrast (red zoomed-in area). Mathematically GC is expressed as

$$I_G = \alpha I^\gamma \quad (2)$$

where I is input hazy image and I_G is the corrected image. Typically, α defaults to 1. According to Fig. 2 and fusion efficiency, in this paper, we set $\gamma = \{2, 3, 4, 5\}$ to generate a corrected sequence $\{I_{G,n}\}$.

Color-Preserving AHE. Motivated by the success of AHE [45] in image dehazing [41], [42], [60], we also employ it in our framework. However, current AHE-based dehazing approaches almost ignore the dependency among color channels, and inevitably introduce various artifacts. Because there are usually two ways to process color images: 1) Three channels are processed separately; 2) Converting colors from RGB to other color models (e.g., HSV or CIELab), and then performing AHE on the intensity component of conversion results. Here, we consider processing three color channels jointly.

An example is given in Fig. 3, where two slight and heavy fog images are selected. From Fig. 3(a), we can observe that as the density of fog increases, the distribution of histograms in the three channels of foggy images becomes more overlapped. Concretely, we notice that in the slight haze images, the components of histograms cover a wide range of the intensity scale, whereas the distribution of three-channel pixels is far from being uniform. The heavy haze images have a narrow histogram located typically in the middle of the intensity levels. For an image with heavy haze this means washed-out gray tones. Intuitively, a contrast-enhanced image is expected to appear with high contrast and vivid color, of which the pixels tend to be evenly distributed at as many intensity levels as possible. Accordingly we consider a cascade of RGB channels for the AHE processing, also in view that the operation of concatenation is common in image processing tasks, such as image denoising [61]–[63], multi-exposure image fusion [64], and tone mapping [65]. To

this end, we can reshape the color image $\mathbf{I}^{H \times W \times 3}$. That is, the color image $\mathbf{I}^{H \times W \times 3}$ is first converted to two-dimensional form $\mathbf{I}_{RT}^{H \times W}$ or $\mathbf{I}_{RT}^{H' \times W}$, where $W = 3 \times W$ and $H' = 3 \times H$. More concretely, gray-level values of R, G, and B channels for each pixel are arranged in the order of R, G, and B, respectively, and in this way, each pixel of the color image \mathbf{I} is traversed. We denote this process as **Reshaping Transformation** (RT), i.e.,

$$\mathbf{I}_{RT} = \mathbf{RT}(\mathbf{I}), \quad (3)$$

where $\mathbf{RT}(\cdot)$ denotes the transformation. Next, we perform the operation of AHE [45] to \mathbf{I}_{RT} ,

$$\mathbf{I}_{C,RT} = \mathbf{AHE}(\mathbf{I}_{RT}), \quad (4)$$

where parameters of $\mathbf{AHE}(\cdot)$ take the default settings. Finally, we implement an inverse reshaping transformation on the enhanced result $\mathbf{I}_{C,RT}$, i.e., converting back to the RGB color space:

$$\mathbf{I}_C = \mathbf{RT}^{-1}(\mathbf{I}_{C,RT}). \quad (5)$$

We refer to the whole process above as color-preserving AHE (CP-AHE), as it preserves vivid color. Through extensive experiments, the proposed CP-AHE performs well for most real-world hazy images. As to the slight haze image, CP-AHE may produce a transformed image (see Fig. 3(e)) with a higher dynamic range and less color distortion, while other methods (see Fig. 3(b)–(d)) suffer from noticeable gray or purplish cast. However, when it comes to heavy haze, the performance of our method may also deteriorate.

Through the above operations of GC and CP-AHE, we can obtain several derived inputs, i.e., $\{\mathbf{I}_{G,n}\} \cup \mathbf{I}_C$.

B. Fusion for Dehazing

The net result of image dehazing depends on the quality of derived inputs and the specific fusion strategy. Inspired by the success of structural patch decomposition for multi-exposure fusion (SPD-MEF) [64], we carefully explore to fuse the pseudo-exposure hazy images using SPD.

1) **SPD-MEF for Dehazing:** SPD-MEF aims to decompose an image patch $\mathbf{x}_k \in \mathbb{R}^{CN^2}$ (where C refers to the number of color channels, and N denotes to the spatial size of the image patch) into three independent components: signal strength, signal structure, and mean intensity

$$\begin{aligned} \mathbf{x}_k &= \|\mathbf{x}_k - \mu_{\mathbf{x}_k}\| \cdot \frac{\mathbf{x}_k - \mu_{\mathbf{x}_k}}{\|\mathbf{x}_k - \mu_{\mathbf{x}_k}\|} + \mu_{\mathbf{x}_k} \\ &= \|\tilde{\mathbf{x}}_k\| \cdot \frac{\tilde{\mathbf{x}}_k}{\|\tilde{\mathbf{x}}_k\|} + \mu_{\mathbf{x}_k} \\ &= c_k \cdot s_k + l_k, \end{aligned} \quad (6)$$

where $\|\cdot\|$ denotes l_2 norm, $\mu_{\mathbf{x}_k}$ represents the mean of the patch \mathbf{x}_k , and $\{\mathbf{x}_k\} = \{\mathbf{x}_k | 1 \leq k \leq K\}$ is a set of color image patches extracted from the same position in the exposure sequence including K input images. $c_k = \|\tilde{\mathbf{x}}_k\|$ and l_k are two scalars, representing signal strength and mean intensity, respectively. $s_k = \frac{\tilde{\mathbf{x}}_k}{\|\tilde{\mathbf{x}}_k\|}$ is a unit vector that refers to signal structure. Since the patch decomposition is invertible, any ideal output image patch can be merged by fusing three independent components and then performs inverse reconstruction. The fusion strategy for each component is described in the following.

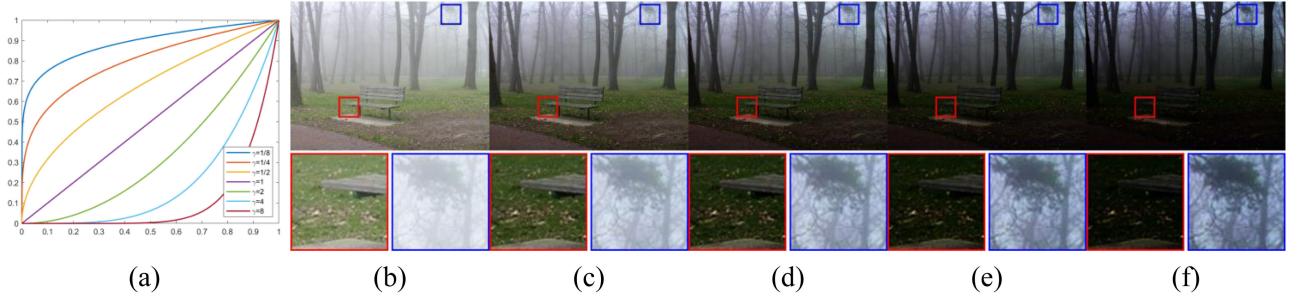


Fig. 2. Plots of gamma transformation curves and gamma correction of the hazy image with different γ values. (a) Plots of the Eq. (2) for various values of γ ($\alpha = 1$ in all cases); (b) Input hazy image; (c)-(f) Results of applying the transformation in Eq. (2) with $\alpha = 1$ and $\gamma = 2, 3, 4$, and 5, respectively.

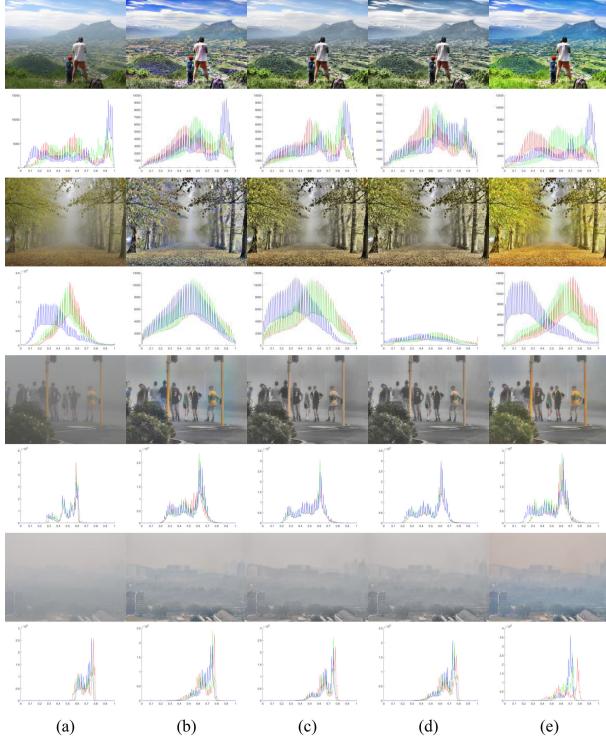


Fig. 3. Visual comparisons between AHE [45] and the proposed CP-AHE on hazy images, and their corresponding RGB histograms. (a) Input hazy images and corresponding histograms; (b) AHE applied to each color channel independently, and corresponding histograms; (c) Enhanced in HSV space with three steps: RGB to HSV transform, AHE of the intensity channel, then HSV to RGB transform, and corresponding histograms; (d) Enhanced in CIELab space with three steps: RGB to CIELab transform, AHE of the intensity channel, then CIELab to RGB transform, and corresponding histograms; and (e) Proposed CP-AHE and corresponding histograms.

Signal Strength. The visibility of a local image patch depends on its contrast, which is directly related to signal intensity. The greater the contrast, the better the visibility. Since visibility is an important measure of image dehazing, signal strength can reflect the degree of hazy images. Therefore, the ideal image patch signal strength can be determined by the largest one

$$\hat{c} = \max_{1 \leq k \leq K} c_k = \max_{1 \leq k \leq K} \|\tilde{\mathbf{x}}_k\|. \quad (7)$$

Signal Structure. The unit structure strength s_k describes the specific direction of CN^2 -dimensional space. The ideal local

signal structure is given by

$$\hat{\mathbf{s}} = \frac{\bar{\mathbf{s}}}{\|\bar{\mathbf{s}}\|}, \quad \text{and} \quad \bar{\mathbf{s}} = \sum_{k=1}^K \beta_k \mathbf{s}_k, \quad (8)$$

where the weights $\beta_k \geq 0$ for $1 \leq k \leq K$, and $\sum_{k=1}^K \beta_k = 1$. β_k is defined as follows

$$\beta_k = \frac{\|\tilde{\mathbf{x}}_k\|^p}{\sum_{k=1}^K \|\tilde{\mathbf{x}}_k\|^p}. \quad (9)$$

where $p \geq 0$ an exponent parameter. β_k is proportional to signal strength $\|\tilde{\mathbf{x}}_k\|$ and normalized power function form. The higher the p value, the greater the weight given to the local patches that have relatively high strength.

Mean Intensity. The desired mean intensity of the local image patch \mathbf{x}_k is computed by

$$\hat{l} = \sum_{k=1}^K \alpha_k l_k, \quad (10)$$

where $\alpha_k \geq 0$ for $1 \leq k \leq K$ is a weighting function of the global mean μ_k of the k -th image X_k and the local mean l_k of the k -th local patch \mathbf{x}_k . A two-dimensional Gaussian profile is employed to quantify

$$\alpha_k = \frac{\exp\left(-\frac{(\mu_k - \mu_c)^2}{2\sigma_g^2} - \frac{(l_k - l_c)^2}{2\sigma_l^2}\right)}{\sum_{k=1}^K \exp\left(-\frac{(\mu_k - \mu_c)^2}{2\sigma_g^2} - \frac{(l_k - l_c)^2}{2\sigma_l^2}\right)}. \quad (11)$$

where σ_g and σ_l control the spreads of the Gaussian profile. Similar to β_k , α_k is also a normalized form, i.e., $\sum_{k=1}^K \alpha_k = 1$. When the input image sequence is normalized to $[0, 1]$, μ_c and l_c are set to 0.5.

Eventually, a desired image patch can be obtained as

$$\hat{\mathbf{x}} = \hat{l} + \hat{c} \cdot \hat{\mathbf{s}}. \quad (12)$$

The final output image is generated by operating aggregation process on the overlapped patches.

SPD-MEF can be directly applied to pseudo-exposure images, but the fused result often suffers from high computational cost, low contrast and inferior sharpness as shown in Fig. 4(b). We propose a fast multi-scale method based on SPD for reducing computational cost and improving dehazing performance. It should be noted that the multi-scale framework is based on

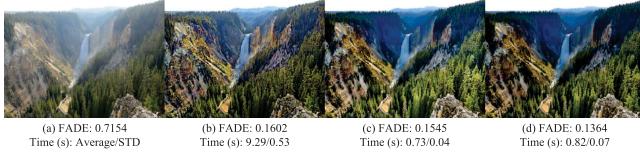


Fig. 4. Effectiveness of the proposed fusion method in terms of visual and quantitative comparison. (a) Input hazy image taken from [31]; (b) SPD-MEF with stride 1 [64]; (c) Fast SPD-MEF [46]; and (d) Proposed fusion method. The second row below the images lists the average time and standard deviation (STD) of three methods running five times on the input hazy image of size 768×432 pixels.

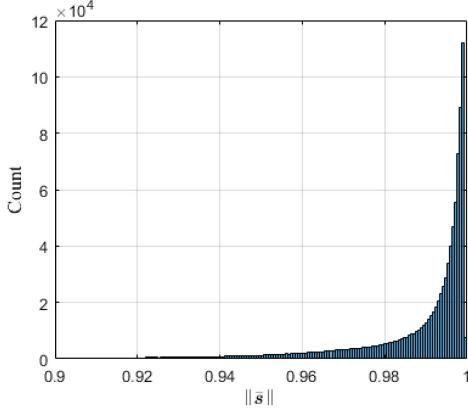


Fig. 5. Randomly select ten real-world hazy images from [31] to count the histogram of $\|\bar{s}\|$.

our recent work [46], but with characteristic weight design and parameters setting.

2) Our Fast Multi-Scale Fusion for Dehazing: As can be seen from Eq. (6), the complexity of SPD-MEF is $\mathcal{O}(KMN^2)$, where K represents the number of input images, M denotes the number of pixels of the input image, and N is the spatial size of the local patch. Next, we show that this complexity can be reduced to $\mathcal{O}(KM)$. It is verified that $\|\bar{s}\|$ calculated by SPD-MEF approximates 1 in [46] via a detailed analysis. The assumption that $\|\bar{s}\| \approx 1$ is more reasonable for pseudo multiple exposure images than real multi-exposure ones as shown in Fig. 5. This is because simulated images are free of pixel misalignment and moving objects. It indicates that whether the signal structure is normalized or not has a negligible effect on the final dehazing result.

Since $\|\bar{s}\| \approx 1$ and $s_k = \frac{\tilde{x}_k}{\|\tilde{x}_k\|}$, we can substitute \hat{l} in Eq. (10) and \bar{s} in Eq. (8) into Eq. (12), i.e.,

$$\begin{aligned} \hat{x} &\approx \sum_{k=1}^K (\alpha_k l_k \cdot e + \hat{c} \beta_k \cdot s_k) \\ &= \sum_{k=1}^K \left(\alpha_k l_k \cdot e + \hat{c} \frac{\beta_k}{\|\tilde{x}_k\|} \cdot \tilde{x}_k \right) \\ &= \sum_{k=1}^K (\alpha_k l_k \cdot e + \gamma_k \cdot (x_k - l_k)), \end{aligned} \quad (13)$$

where $\gamma_k = \hat{c} \frac{\beta_k}{\|\tilde{x}_k\|}$ and e is a unit vector with the same dimension as s_k . Let $\gamma_k = \hat{c} \frac{\beta_k}{\|\tilde{x}_k\|}$, we can perform structural patch decomposition implicitly, and the final fused image can be obtained

by

$$\hat{X} = \sum_{k=1}^K \left(\mathbf{F}(\alpha_k \odot \mathbf{F}(X_k)) + \mathbf{F}(\gamma_k) \odot X_k - \mathbf{F}(\gamma_k \odot \mathbf{F}(X_k)) \right), \quad (14)$$

where $\mathbf{F}(\cdot)$ represents the mean filter with kernel size N , and \odot is the product of the Hadamard. Since the transition from Eq. (14) to Eq. (15) is from patch-level to image-level realized by the mean filtering $\mathbf{F}(\cdot)$, l_k becomes $l_k = \mathbf{F}(X_k)$. Similarly, we use the filter to calculate weight maps α_k and γ_k . The implementation of the mean filter only needs linear time through box filter [66], [67]. Therefore, the complexity of SPD-MEF is reduced to $\mathcal{O}(KM)$, regardless of patch size N .

The kernel size N of the filter $\mathbf{F}(\cdot)$ or equivalent to the size of the patch has a significant impact on the final fusion result. When the kernel size is small, more detail is captured, but images in foggy scenes are often corrupted by various noise [11]. This makes it easy to generate weight maps with visible noise, causing spatial inconsistency in the final result. In contrast, a large-size kernel would reduce noise disturbances, but some fine detail may be lost. Although intermediate kernel sizes can balance spatial consistency and detail preservation, they may induce noticeable halo artifacts to occur near strong edges.

To reduce halo artifacts and preserve fine detail, multi-scale decomposition of Eq. (14) is implemented. When we take the original hazy image sequence as scale 1, we can get rich high-frequency information from Eq. (14)

$$\hat{H}^{(1)} = \sum_{k=1}^K \left(\mathbf{F}(\gamma_k^{(1)}) \odot X_k^{(1)} - \mathbf{F}(\gamma_k^{(1)} \odot \mathbf{F}(X_k^{(1)})) \right). \quad (15)$$

We can obtain the second scale base component by down-sampling

$$X_k^{(2)} = \mathbf{D}(\mathbf{F}(X_k^{(1)})), \quad (16)$$

where $\mathbf{D}(\cdot)$ represents the down-sampling operator with a factor of 2. Similar to Eq. (15), the detailed information of the second scale is

$$\hat{H}^{(2)} = \sum_{k=1}^K \left(\mathbf{F}(\gamma_k^{(2)}) \odot X_k^{(2)} - \mathbf{F}(\gamma_k^{(2)} \odot \mathbf{F}(X_k^{(2)})) \right). \quad (17)$$

Recursively through the above process, we can obtain the detail component of j -th scale until the coarsest scale

$$J = \lfloor \log_2 \min(H, W) \rfloor - T, \quad (18)$$

where $\lfloor \cdot \rfloor$ means round down, H and W represent the height and width of the hazy image sequence, respectively. The value of T is related to the size of kernel size N . The setting of Eq. (18) is inspired by [68], [69]. The kernel size should be smaller than the size of a certain layer image. Similarly, the base component at J -th scale can also be obtained by

$$\hat{B}^{(J)} = \sum_{k=1}^K \mathbf{F}(\alpha_k^{(J)} \odot \mathbf{F}(X_k^{(J)})), \quad (19)$$

As the image becomes smoother at coarser scales with the recursive filtering and down-sampling, there should be a corresponding smaller kernel size. Hence, the filter size at each scale

is adaptive instead of a fixed constant as

$$N^{(j+1)} = \max(N^{(j)} - 2, 3), \quad j = 1, \dots, J-1. \quad (20)$$

where $N^{(j)}$ represents the kernel size at j -th scale. With a scale-aware filter size, more decomposition levels can be attained to finely maintain the detail at different scales.

Finally, the haze-free image can be obtained from

$$\hat{X} = \hat{\mathbf{B}}^{(1)} + \hat{\mathbf{H}}^{(1)}, \quad (21)$$

where $\hat{\mathbf{B}}^{(1)}$ can be obtained by recursively up-sampling and adding the detail layer to the middle base layer

$$\hat{\mathbf{B}}^{(j)} = \mathbf{F} \left(\mathbf{U} \left(\hat{\mathbf{B}}^{(j+1)} + \hat{\mathbf{H}}^{(j+1)} \right) \right), \quad j = J-1, \dots, 1. \quad (22)$$

where $\mathbf{U}(\cdot)$ denotes the up-sampling operator with a fact of 2.

3) *Weight Calculation*: We use an improved arctan function [65] to calculate $\alpha_k^{(J)}$, *i.e.*,

$$\alpha_k^{(J)} = \frac{\arctan \left(0.5\lambda - \left| 0.5 - X_k^{(J)} \right| \lambda \right)}{\sum_{k=1}^K \arctan \left(0.5\lambda - \left| 0.5 - X_k^{(J)} \right| \lambda \right)} \nabla X_k^{(J)}, \quad (23)$$

where λ is a constant and ∇ denotes the gradient operator. Since the coarsest layer still contains a certain amount of edge and contour information in the generated hazy images, $\nabla X_k^{(J)}$ is added to the design of $\alpha_k^{(J)}$. The weight map $\gamma_k^{(j)}$ for $X_k^{(j)}$ is calculated the same as SPD-MEF [64], except that it is calculated on each scale j .

In conclusion, we articulate the difference of our fusion method with fast SPD-MEF [46]. Considering that the pseudo multiple exposure images derived from GC and CP-AHE contain sufficient detail, we use more decomposition levels and scale-aware kernel size to extract and blend finer detail. Besides, we leverage gradient information when fusing the base layer. The modified fusion method can obtain sharper dehazing result compared with fast SPD-MEF [46] as shown in Fig. 4(c) and (d). The proposed fusion-based image dehazing framework is presented in Algorithm 1.

IV. EXPERIMENTS

A. Experiment Settings

Dataset. To evaluate our method, we use 32 well-known real-world scenarios¹ as test set used widely by previous approaches [9], [12]–[14], [18], [21], [22], [29], [49]. In addition to these real-world images, we also evaluate the proposed method on the synthetic benchmark datasets, *i.e.*, **RESIDE** [2], **HazeRD** [35] and **O-HAZE** [36]. For the **RESIDE** dataset, **SOTS** and 10 synthetic hazy images of **HSTS** are adopted for testing.

Implementation Details. The parameters of the proposed fusion method are set as follows: the patch size $N^{(1)} = 11$, $T = 1$ to have a trade-off between decomposition levels and low-frequency intensity preserving, $\lambda = 20$ that reflects the arctan curve [65], and the exponent parameter $p = 6$. For benchmark datasets dehazing, we set $\gamma = \{1, 2, 3\}$.

¹[Online]. Available: https://www.cs.huji.ac.il/w~raananf/projects/dehaze_cl/results/

Algorithm 1: The Proposed CEEF for Image Dehazing

Require: Input hazy image \mathbf{I} .

- 1: Set the parameters of GC, CP-AHE, and fast multi-scale SPD-MEF.
 - 2: Calculate $\{\mathbf{I}_{G,n}\}, \mathbf{I}_C$ by Eqs. (2) and (5) respectively, and then generate a multi-exposure sequence $\{\mathbf{I}_{G,n}\} \cup \mathbf{I}_C$ with different fog density.
 - 3: **for** $j = 1$ to $J-1$ **do**
 - 4: Compute the detail component $\hat{\mathbf{H}}^{(j)}$ (refer to Eqs. (15) and (17)).
 - 5: Down-sample $\mathbf{X}_k^{(j)}$ (refer to Eq. (16)).
 - 6: **end for**
 - 7: Compute the detail component $\hat{\mathbf{H}}^{(J)}$ (refer to Eqs. (15) and (17)).
 - 8: Compute the base component $\hat{\mathbf{B}}^{(J)}$ by Eq. (19).
 - 9: Reconstruct the resulting image \hat{X} by Eqs. (21) and (22).
 - 10: **Output** : The dehazed image $\mathbf{J} = \hat{X}$.
-

Baselines. The proposed method is compared with the 19 state-of-the-art methods that include four prior-based methods: DCP [8] (TPAMI'10), CAP [24] (TIP'15), GRM [48] (ECCV'16) and NLP [49] (TPAMI'20); two fusion-based methods: DEFADE [40] (TIP'15) and AMEF [42] (SP'18); and thirteen deep learning-based methods: DehazeNet [13] (TIP'16), MSCNN [14] (ECCV'16), AOD-Net [15] (ICCV'17), PQC [20] (TIP'18), EPDN [29] (CVPR'19), GridDehazeNet [28] (ICCV'19), Deep-DCP [21] (TIP'19), PGAN [22] (TPAMI'20), DA [33] (CVPR'20), FFA-Net [32] (AAAI'20), GFN [26] (CVPR'18), MSBDN [25] (CVPR'20) and RefineDNet [59] (TIP'21). Among them, the last three methods are the combination of deep learning and feature fusion. Due to memory resource limitations, GFN [26] is implemented only on the **SOTS** and **HSTS** datasets. For the sake of fairness, all methods are compared with the corresponding default parameters. Our code is available at: <https://github.com/xiaohuibin/CEEF-TMM-2021>.

Evaluation Metric. For 32 real-world scenes, we select a lately developed no-reference image quality assessment (NRIQA) metric, Fog Aware Density Evaluator (FADE)² [40], which is specially designed for evaluating image dehazing and widely used to evaluate the performance of dehazing algorithms in real scenes [4], [38], [40], [42], [52]–[54]. FADE derived from the statistical observation of a large number of real foggy images and fog-free images not only predicts the visibility of a foggy scene, but also accurately evaluates the performance of the dehazing algorithm. The smaller the value of FADE, the better the quality of image dehazing. We also use three well-known metrics, PSNR, SSIM and CIEDE2000 [35], to evaluate the benchmark datasets.

Experimental Environment. The experimental results of deep learning-based methods are carried out on an NVIDIA GeForce GTX 1080 Ti GPU in PyTorch and TensorFlow. Other experiments are conducted on MATLAB R2016b installed on Win 7 64-bit operating system and a PC with Intel Core i5-3210M CPU (2.50 GHz), 6 GB RAM.

²[Online]. Available: <http://live.ece.utexas.edu/research/fog/index.html>

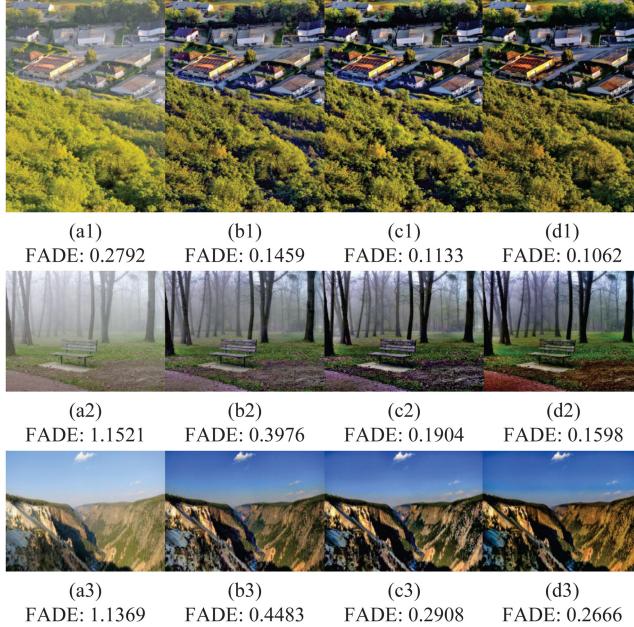


Fig. 6. Ablation study. (a1)-(a3) Input images taken from [31]; (b1)-(b3) AMEF [42]; (c1)-(c3) Replacing the fusion method [44] in AMEF [42] with the proposed fusion method; and (d1)-(d3) Proposed method.

B. Ablation Study

Compared with AMEF [42], the main difference in our method lies in the two aspects: derived inputs and multi-exposure fusion method. Therefore, we conduct an ablation study to verify the effectiveness of the proposed method. There are mainly the following three variants for comparison.

Effectiveness of Proposed Fusion Method. It has been demonstrated in [46] that our previously proposed fast multi-scale SPD-MEF is superior to the traditional Laplacian pyramid decomposition-based method [44]. To illustrate the effectiveness of the proposed method in image dehazing, we only replace the fusion method in AMEF [42] with the proposed fusion method, and the rest remains the same. As shown in Fig. 6(c1)-(c3), our method causes less color distortion and brings richer detail compared to AMEF (see Fig. 6(b1)-(b3)). Quantitative results FADE [40] also demonstrate the effectiveness of our method.

CP-AHE. Figs. 3(e) and 6(d1)-(d3) show that the proposed CP-AHE is wise in the implementation of AHE for image dehazing. This also validates that the proposed fusion method and CP-AHE are capable of improving the performance of image dehazing in real-world scenes.

C. Real-World Images Dehazing

Qualitative Evaluation. Taking into account the length of the space, we divide the state-of-the-art methods into three groups for visual comparisons as shown in Figs. 7-9.

Fig. 7 shows dehazing results of prior-based and fusion-based methods. As can be seen from Fig. 7(b2), (d2), (e2), and (f2), DCP [8], GRM [48], NLP [49], and DEFADE [40] generate clear dehazed images. However, DCP and NLP suffer from serious color artifacts in the sky area (see Fig. 7(b1) and (e1)). CAP [24], GRM and DEFADE perform better on this issue, but

some areas have lower contrast, such as nearby trees (see the second close-up in Fig. 7(c1), (d1) and (f1)). We can observe that AMEF [42] introduces color artifacts (see Fig. 7(g2)), leading to whitish results due to the implementation of AHE [45] on each color channel independently. On the contrary, the proposed CEEP works well, and achieves visually pleasing results with vivid color (see Fig. 7(h1) and (h2)).

MSCNN [14] and EPDN [29] incur varying degrees of color cast, as Fig. 8 (f1) and (c2) shows. Besides, some regions of EPDN tend to reduce contrast (see the second zoomed-in area of Fig. 8). From Fig. 8(b1), (g1), (b2), and (g2), we can see that DehazeNet [13] and GridDehazeNet [28] only remove a small amount of fog. MSCNN, AOD-Net [15], PQC [20] and our method restore images with fine detail and vivid color (see Fig. 8(c1), (d1), (e1), (h1), (d2), (e2), and (h2)). However, the proposed method introduces slight halo artifacts at the cliffs (see Fig. 8(h1)).

We see in Fig. 9 that Deep-DCP [21], FFA-Net [32], MS-BDN [25], and RefineDNet [59] on removing fog far from the camera is relatively poor (see Fig. 9(b1), (e1), (d1), (b2), (e2), and (f2)). PGAN [22], DA [33] and the proposed method are capable of recording more pleasant colors and capturing richer detail than other methods (see Fig. 9(c1), (d1), (h1), (c2), (d2), and (h2)). Compared to PGAN and DA, our approach is superior in terms of detail extraction and color retention in some regions (see the red zoomed-in area of Fig. 9(h1) and the blue zoomed-in area of Fig. 9(h2), respectively). For more visual results, please see supplementary materials.

Quantitative Evaluation. The comparisons results with 19 approaches on 32 real-world scenes are shown in Table I where the values of FADE [40] for each method and the corresponding average score are listed. As shown, our approach outperforms other competitive ones on the real dataset and achieves the gain with 0.02 in terms of the average value of FADE compared to the second place method NLP [49]. From Table I, we also have other notable observations: 1) In the dehazing performance of real-world scenes, the top four are traditional methods, which further illustrates the drawbacks of current simulated datasets [3], [36], [37] for deep learning-based methods; 2) Our method achieves lower FADE than AMEF [42]. We attribute the improved performance to the proposed pre-processing (CP-AHE) and fusion method; 3) Compared to the latest MSBDN [25], DA [33], PGAN [22], GridDehazeNet [28], FFA-Net [32], EPDN [29] and RefineDNet [59], the proposed method has achieved more impressive performance.

D. Nighttime Images Dehazing

Although the foregoing various defogging methods have made significant progress in daytime scenes, it is still a huge challenge for night scenes [43]. There are two main reasons: 1) Due to the presence of active artificial light sources such as street lamp, automotive lamp and decorative lamp, this results in varicolored uneven illumination. As a result, this not just directly affects the accurate estimation of atmospheric light A in Eq. (1), but also invalidates some of the previously proposed priors or assumptions; 2) The lack of large-scale nighttime hazy datasets has hampered progress in this area [70].

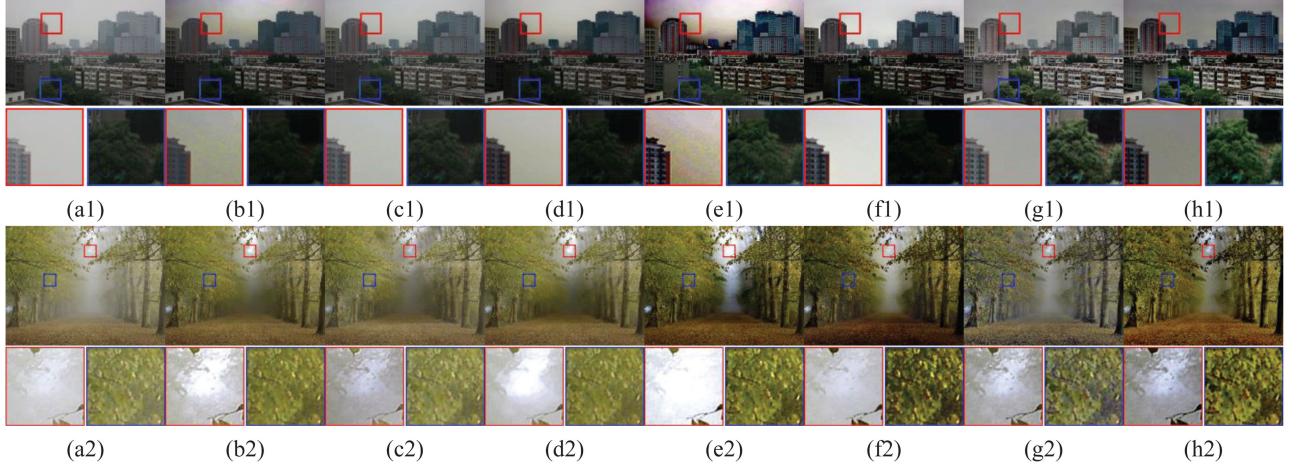


Fig. 7. Visual comparisons in real-world scenes. (a1)-(a2) Input images: “Herzeliya” and “Forest”; (b1)-(b2) DCP [8]; (c1)-(c2) CAP [24]; (d1)-(d2) GRM [48]; (e1)-(e2) NLP [49]; (f1)-(f2) DEFADE [40]; (g1)-(g2) AMEF [42]; (h1)-(h2) Proposed method.

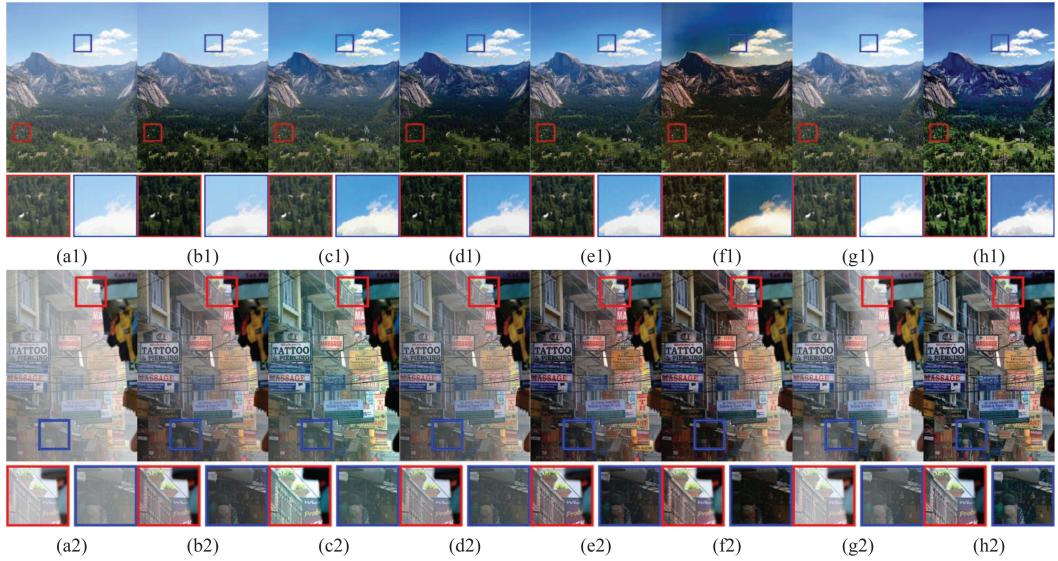


Fig. 8. Visual comparisons in real-world scenes. (a1)-(a2) Input images: “Y1” and “Flags”; (b1)-(b2) DehazeNet [13]; (c1)-(c2) MSCNN [14]; (d1)-(d2) AOD-Net [15]; (e1)-(e2) PQC [20]; (f1)-(f2) EPDN [29]; (g1)-(g2) GridDehazeNet [28]; and (h1)-(h2) Proposed method.

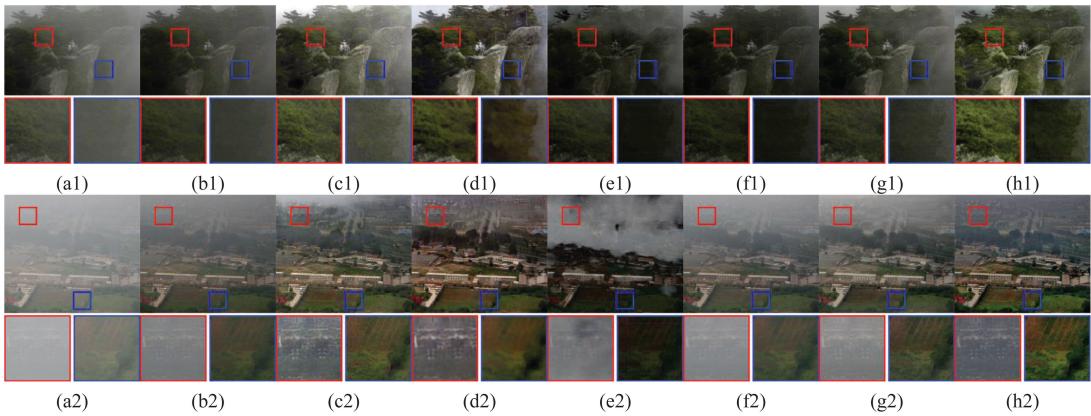


Fig. 9. Visual comparisons in real-world scenes. (a1)-(a2) Input images: “Cliff” and “Canon”; (b1)-(b2) Deep-DCP [21]; (c1)-(c2) PGAN [22]; (d1)-(d2) DA [33]; (e1)-(e2) FFA-Net [32]; (f1)-(f2) MSBDN [25]; (g1)-(g2) RefineDNet [59]; and (h1)-(h2) Proposed method.

TABLE I

FADE (LOW IS BETTER) COMPARISONS OF DCP [8], CAP [24], GRM [48], NLP [49], DEFADE [40], AMEF [42], DEHAZENET [13], MSCNN [14], AOD-NET [15], PQC [20], EPDN [29], GRIDDEHAZENET [28], DEEP-DCP [21], PGAN [22], DA [33], FFA-NET [32], MSBDN [25], REFINEDNET [59] AND OUR METHOD ON 32 REAL-WORLD SCENES. EACH ROW IS HIGHLIGHTED WITH RED, GREEN, AND BLUE TO INDICATE THE BEST THREE RESULTS, RESPECTIVELY. THE BOTTOM ROW REPRESENTS THE AVERAGE VALUE

Image	DCP	CAP	GRM	NLP	DEF-ADE	AMEF	Dehaze	MSCNN	AOD-Net	PQC	EPDN	GridDehazeNet	Deep-DCP	PGAN	DA	FFA-Net	MSBDN	RefineDNet	Ours
Aerial	0.3691	0.4956	0.3099	0.2441	0.1880	0.2602	0.3218	0.3281	0.2260	0.2227	0.2104	0.4808	0.5442	0.4611	0.3726	0.821	0.5344	0.2946	0.1289
Buildings	0.5735	0.8164	0.5153	0.2143	0.2696	0.4277	0.3967	0.4514	0.2841	0.3506	0.3929	0.6515	0.7620	0.7762	0.5847	0.4115	0.6577	0.5288	0.2113
Canon	0.4554	1.3614	0.8981	0.2729	0.1024	0.8307	1.0031	0.8448	0.947	0.4128	0.3710	1.0355	1.2288	0.4255	0.3808	0.5190	1.1989	0.6547	0.3165
Castle	0.3881	0.6428	0.4895	0.4552	0.3791	0.4779	0.3330	0.7412	0.3422	0.3952	0.3464	0.4987	1.4978	0.6111	0.6289	0.4563	0.5503	0.4180	0.1830
Ctyscape	0.5079	1.4108	1.2927	0.2089	0.5720	0.8685	1.1299	0.7693	0.9407	0.4672	0.4127	1.2751	1.2557	0.7483	0.6108	0.5057	0.9649	0.5613	0.4201
Cliff	0.6839	1.1144	0.5570	0.2256	0.3989	0.6900	0.4650	0.6456	0.4420	0.3055	0.3801	0.6394	1.1458	0.6516	0.4362	0.5326	0.5552	0.6759	0.3255
Cones	0.2212	0.4140	0.2603	0.1269	0.1351	0.2291	0.2533	0.3012	0.2141	0.1509	0.1619	0.2756	0.3265	0.1939	0.3133	0.1567	0.3423	0.1886	0.0922
Dolls	0.3003	0.5889	0.4067	0.3779	0.4392	0.4453	0.5242	0.4924	0.5394	0.2976	0.2504	0.3934	0.6528	0.2451	0.2883	0.2140	0.2632	0.3304	0.2331
Dubai	0.2262	0.3801	0.4325	0.2279	0.1536	0.2925	0.3525	0.2865	0.3388	0.2409	0.2472	0.5387	0.5091	0.3835	0.4189	0.1975	0.7198	0.3035	0.1387
Flags	0.2887	0.3620	0.3558	0.1181	0.2507	0.2517	0.2951	0.2349	0.2750	0.1714	0.1705	0.3500	0.6791	0.3330	0.2636	0.2630	0.4646	0.2433	0.1347
Florence	0.3126	0.6048	0.4911	0.1515	0.1999	0.2936	0.3608	0.3642	0.2675	0.2372	0.3580	0.5695	0.6315	0.4737	0.5709	0.2727	0.5960	0.4203	0.1546
Forest	0.2355	0.2918	0.2660	0.1276	0.1349	0.1953	0.2140	0.2033	0.1619	0.1417	0.1861	0.2967	0.3209	0.2896	0.3210	0.3557	0.2913	0.1591	0.1050
Herzeliya	0.2654	0.3768	0.3477	0.1824	0.2963	0.3235	0.4402	0.2137	0.1965	0.2299	0.1926	0.3767	0.6345	0.2940	0.3684	0.4396	0.3370	0.2205	0.1328
Hongkong	0.3447	0.5272	0.4459	0.2614	0.2842	0.3640	0.4855	0.4461	0.2498	0.3569	0.3552	0.5439	0.6471	0.4908	0.6937	0.4902	0.5879	0.3386	0.2094
House	0.1644	0.1940	0.1810	0.1455	0.1160	0.1450	0.1636	0.1250	0.1335	0.1075	0.1485	0.2236	0.3145	0.2075	0.3388	0.2560	0.2554	0.1589	0.1062
Lviv	0.3640	0.4787	0.5475	0.1929	0.3179	0.2920	0.3894	0.3663	0.2813	0.2882	0.2715	0.5519	0.7060	0.7190	0.5533	0.3901	0.6240	0.3927	0.1732
Mountain	0.2924	0.4181	0.3253	0.2378	0.2916	0.2970	0.3931	0.2340	0.2702	0.2612	0.2547	0.4395	0.6180	0.4580	0.3709	0.2105	0.4631	0.2684	0.1653
Night	0.2493	0.2544	0.2441	0.2297	0.2099	0.1815	0.2511	0.2207	0.2165	0.2053	0.289	0.2435	0.5503	0.3828	0.3945	0.2157	0.2321	0.2100	0.1748
Ny12	0.2409	0.3163	0.2773	0.1134	0.1700	0.2066	0.2504	0.2471	0.1343	0.1881	0.2099	0.2937	0.3828	0.5190	0.4116	0.5136	0.4884	0.1778	0.1051
Ny17	0.2566	0.4787	0.4227	0.1272	0.1978	0.2552	0.3105	0.2811	0.1987	0.2076	0.2496	0.4596	0.4935	0.3352	0.4458	0.4209	0.4039	0.3229	0.1457
Pumpkins	0.2516	0.3908	0.2903	0.1632	0.1689	0.2498	0.2843	0.2478	0.2251	0.1959	0.1969	0.3311	0.5710	0.2212	0.4560	0.1492	0.3502	0.2523	0.1228
Road	0.3185	0.3943	0.3392	0.2348	0.2415	0.2865	0.4007	0.3976	0.2934	0.2192	0.2329	0.5547	0.653	0.5537	0.5165	0.4299	0.6041	0.2170	0.1274
Schechner	0.3932	0.5950	0.4616	0.2774	0.3453	0.4462	0.7879	0.4463	0.3773	0.4912	0.5359	0.6381	0.6365	0.4351	0.6811	0.8246	0.6147	0.4540	0.2776
Snow	0.5963	0.2387	0.5302	0.4974	0.4616	1.5844	1.0617	1.1192	1.5504	0.7134	0.6993	1.1828	1.4350	0.6586	0.8222	0.5414	1.8505	1.5979	0.6389
Stadium	0.1858	0.2920	0.2634	0.1752	0.3198	0.1971	0.1987	0.1439	0.1890	0.1683	0.1582	0.2446	0.4589	0.1953	0.3079	0.2262	0.2640	0.1789	0.1127
Swan	0.4183	0.6323	0.4249	0.1670	0.2283	0.3358	0.4002	0.3955	0.2681	0.2782	0.3074	0.6580	0.7741	0.4849	0.6102	0.7937	0.7354	0.3869	0.1513
Tiananmen	0.5090	0.8967	0.6613	0.3572	0.3619	0.6167	0.7371	0.6793	0.6002	0.5116	0.6201	0.6615	1.0218	0.5515	0.6097	0.3597	0.6282	0.7418	0.3532
Train	0.3487	0.4458	0.3881	0.1375	0.2448	0.3965	0.2772	0.7868	0.3489	0.2327	0.2742	0.4789	0.8037	0.5423	0.4313	0.2213	0.4536	0.2824	0.1571
Tree	0.5483	0.9944	0.6328	0.4232	1.0401	0.8495	0.9217	1.0538	1.0048	0.3765	0.5110	1.0129	0.7031	0.4895	0.5796	1.1554	1.2484	0.7720	0.4426
Urbino	0.4915	0.6409	0.5100	0.2322	0.3254	0.3460	0.2994	0.3255	0.2491	0.2883	0.4659	0.3952	1.9956	0.5346	0.7141	0.4549	0.4543	0.4984	0.2155
Y16	0.3332	0.4210	0.3827	0.2255	0.2379	0.2874	0.4932	0.2462	0.2062	0.2447	0.2763	0.4973	0.5713	0.4008	0.4810	0.4355	0.4809	0.3476	0.1814
Y1	0.3399	0.4730	0.4196	0.2854	0.2638	0.3144	0.4314	0.4549	0.2279	0.3170	0.2508	0.6120	0.5015	0.5028	0.5112	0.4789	0.6119	0.3163	0.1846
Average	0.3595	0.6277	0.4803	0.2318	0.3522	0.4115	0.4571	0.4404	0.3812	0.2899	0.3121	0.5439	0.7508	0.4553	0.4840	0.4285	0.5883	0.4036	0.2069



Fig. 10. Visual and quantitative comparison (FADE) of nighttime hazy images. (a1)-(a5) Input images taken from [51]; (b1)-(b5) GMLC [50]; (c1)-(c5) MRP [51]; (d1)-(d5) MRPF [51]; and (e1)-(e5) Proposed method.

To verify that the proposed method is also suitable for removing haze from the night, we compare the proposed method with the three latest nighttime dehazing methods, *i.e.*, GMLC [50], MRP [51] and MRPF [51] (a fast variant of MRP) on the night scene dataset³ that consists of 20 hazy images captured in different nighttime scenes. As shown in Fig. 10, GMLC, MRP and MRPF not only change color distribution of the original night scene, and also introduce amplified noise and visible color artifacts in some regions (*e.g.*, the night sky, lawn, path, tree and

³[Online]. Available: <https://github.com/chaimi2013/MRP>

TABLE II
QUANTITATIVE COMPARISONS (FADE) ON THE NIGHT SCENE DATASET

Dataset	Metric	MRP	MRPF	GMLC	Ours
[51]	FADE	0.3018	0.2771	0.2373	0.1739

neon lamp). They all have different degrees of overexposure effect near the light source, especially for MRPF. The exposure fusion strategy of the proposed method can effectively reduce the over-exposed effects. In addition, they also lead to color cast in the night sky. It is obvious that the proposed method outperforms other nighttime dehazing methods in detail restoration, color distribution, noise suppression and exposure control. As shown in Table II, our method has achieved the best FADE score and a gain of at least 0.06 compared to the second-place GMLC.

E. Benchmark Datasets Dehazing

Table III lists quantitative results of all methods on synthetic datasets. Due to the high resolution of **HazeRD** and **O-HAZE** datasets, most existing deep learning-based and prior-based methods fail to run directly on an NVIDIA GTX 1080 Ti GPU with 12G RAM. For fair comparisons, we have down-sampled the two datasets. Note that we could not generate the results of GFN [26] on the **HazeRD** and **O-HAZE** datasets due to limited computational resource. In fact, the maximum image size that GFN can handle is 640×480 on an NVIDIA GTX 1080 Ti GPU with 12G RAM. As can be seen from Table III, GridDehazeNet [28] obtains the highest PSNR, SSIM and CIEDE2000 on **SOTS** (Outdoor) and **HSTS** (Synthetic scenes). For the **SOTS** (Indoor) dataset, the PSNR, SSIM and CIEDE2000 [35] produced by FFA-Net [32] are higher

TABLE III
QUANTITATIVE RESULTS (PSNR, SSIM, CIEDE2000) OF STATE-OF-THE-ART METHODS AND OURS ON THE **RESIDE (SOTS AND HSTS)**, **HAZERD**, AND **O-HAZE** DATASETS. THE RESULTS OF THE TOP THREE IN EACH ROW ARE HIGHLIGHTED WITH **RED**, **GREEN**, AND **BLUE**, RESPECTIVELY

Datasets	Metrics	DCP	CAP	GRM	NLP	DEF-ADE	AMEF	Dehaze-Net	MSCNN	AOD-Net	PQC	EPDN	GridDehazeNet	Deep-DCP	PGAN	DA	GPN	FFA-Net	MSBDN	Refine-DNet	Ours
Indoor	PSNR	20.03	19.05	18.86	17.29	17.25	17.52	21.35	17.13	18.73	20.25	25.06	32.16	19.28	28.61	25.72	22.43	36.36	32.00	20.48	16.94
	SSIM	0.867	0.836	0.855	0.749	0.782	0.876	0.807	0.846	0.871	0.923	0.984	0.836	0.966	0.932	0.903	0.994	0.986	0.866	0.772	
	CIEDE2000	7.43	8.20	8.67	11.35	11.14	10.87	6.28	10.40	8.22	6.99	4.60	1.89	7.85	2.93	4.70	5.51	1.11	1.71	7.77	11.77
Outdoor	PSNR	17.49	22.31	17.47	18.07	19.02	17.92	22.94	19.50	19.08	20.32	30.86	19.31	19.99	26.38	21.55	20.23	30.77	20.67	19.12	
	SSIM	0.821	0.914	0.819	0.802	0.831	0.832	0.889	0.889	0.861	0.858	0.838	0.982	0.828	0.649	0.899	0.844	0.890	0.955	0.886	0.792
	CIEDE2000	10.38	5.95	10.38	10.27	9.45	10.34	5.83	7.96	7.95	8.94	8.59	2.32	8.60	7.71	4.71	6.48	6.91	2.57	7.43	9.22
HazeRD	PSNR	14.38	15.73	14.58	14.43	14.87	16.86	15.15	15.65	15.49	14.78	15.21	14.90	15.48	17.33	16.62	\	14.13	14.60	17.37	15.90
	SSIM	0.590	0.630	0.587	0.594	0.568	0.611	0.594	0.624	0.573	0.572	0.469	0.652	0.591	0.585	0.631	\	0.610	0.626	0.657	0.589
	CIEDE2000	15.52	14.09	15.11	15.71	15.98	13.19	14.88	13.82	14.76	15.30	15.76	14.28	14.33	11.98	12.35	\	15.67	14.57	11.77	13.80
O-HAZE	PSNR	15.84	16.43	16.76	15.30	15.51	17.42	15.30	17.14	15.06	16.57	16.89	16.55	16.74	16.53	17.83	\	15.12	16.75	16.58	14.92
	SSIM	0.414	0.431	0.421	0.469	0.330	0.495	0.411	0.437	0.307	0.430	0.544	0.405	0.410	0.459	0.536	\	0.361	0.395	0.465	0.287
	CIEDE2000	17.73	15.41	15.91	17.35	19.83	13.83	16.87	15.08	19.76	16.08	13.92	15.57	15.32	13.76	13.08	\	18.92	15.56	15.55	20.48
HSTS	PSNR	17.21	21.63	18.55	17.62	17.32	17.14	24.51	18.66	19.59	19.63	21.67	32.64	19.24	20.33	27.43	22.98	20.76	31.06	21.03	18.77
	SSIM	0.796	0.872	0.818	0.792	0.784	0.812	0.922	0.817	0.853	0.857	0.884	0.986	0.820	0.675	0.907	0.882	0.900	0.954	0.881	0.793
	CIEDE2000	10.38	6.72	9.55	10.48	11.60	11.30	4.97	9.10	7.97	8.53	7.45	1.93	8.67	7.39	4.06	5.75	6.25	2.45	6.82	9.63

TABLE IV
RUN TIME COMPARISON (SECONDS) OF STATE-OF-THE-ART METHODS AND OUR METHOD ON THE REAL-WORLD IMAGE “FOREST” OF SIZE 1024×768 PIXELS

Method	DCP	CAP	GRM	NLP	DEF-ADE	AMEF	Dehaze-Net	MSCNN	AOD-Net	PQC	EPDN	GridDehazeNet	Deep-DCP	PGAN	DA	FFA-Net	MSBDN	Refine-DNet	Ours
Average	53.07	2.66	296.09	10.43	44.74	2.30	7.87	5.92	0.19	19.66	1.67	0.67	3.43	0.61	3.67	1.57	1.10	3.35	1.96
STD	1.34	0.11	175.53	0.30	0.25	0.18	0.85	0.40	0.03	0.15	0.19	0.05	0.02	0.02	0.05	0.02	0.04	0.62	0.06
Platform	Matlab(C) & mex	Matlab(C) & mex	PyTorch(G)	Keras(G)	PyTorch(G)	PyTorch(G)	TensorFlow(G)	PyTorch(G)	PyTorch(G)	PyTorch(G)	PyTorch(G)	PyTorch(G)	Matlab(C)						

than MSBDN [25] by up to 4.36db, 0.008 and 0.6, respectively. On the **HazeRD** dataset, RefineDNet [59] outperforms other state-of-the-art methods. DA [33] achieves the highest PSNR and CIEDE2000 on the **O-HAZE** dataset. The proposed method attains uncompetitive objective scores especially on the **O-HAZE** dataset. The main reason for the low scores of the proposed method is that our method tends to generate a bright and saturated dehazed image, which is quite different from ground-truth.

F. Run-Time Analysis

Table IV lists average run-time and the corresponding standard deviation (STD) of different methods implementing ten times on the real-world image “Forest” of size 1024×768 pixels. The experiment is deployed on a desktop computer with Intel Core i7-6700K CPU (4.0 GHz), 32 GB RAM, and NVIDIA GTX 1080 Ti GPU. C/G in the platform refers to CPU/GPU modes. It can be seen from Table IV that our method is competitive among all approaches. More concretely, our method is faster than almost all prior-based methods and most deep learning-based methods. In particular, the proposed method can be compared with EPN [29]. This indicates that the proposed approach may meet the practical application requirements if implemented with more effective programming language such as C++.

G. The Application of CP-AHE

Low-Light Image Enhancement. Fig. 11 shows visual comparisons of PBS [71] and our CP-AHE on low-light images. As can be seen from Fig. 11(a2)-(e2), PBS produces brighter images but with amplified noise around light sources and night sky. While the proposed CP-AHE is not enhanced bright enough compared to PBS, our method generates visually realistic results, striking a good balance between naturalness and noise suppression. Moreover, our method is faster than PBS and capable of real-time processing.

Image Editing. Fig. 12 shows the performance of the proposed CP-AHE in image editing and aesthetic evaluation.

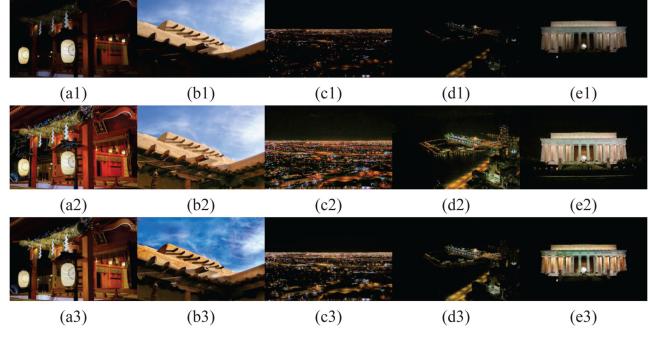


Fig. 11. Visual comparisons on low-light images. (a1)-(e1) Input images taken from [71] and [72]; (a2)-(e2) PBS [71]; and (a3)-(e3) Proposed CP-AHE.



Fig. 12. Aesthetic comparison and users’ aesthetic evaluation. (a) Sample image before being edited by vivo (Smartphone manufacturer); (b) Sample image edited by vivo; (c) Result after being edited by CP-AHE; and (d) Bar chart showing user’s aesthetic vote. Note that the watermark is on the bottom right corner of images (a)-(c).

Fig. 12(a) and (b) are images before and after editing by vivo⁴ (Smartphone manufacturer). In order to make an aesthetic comparison between the sample image Fig. 12(b) and edited image Fig. 12(c), a user study was carried out where a total of 112 people participated in an aesthetic vote, including 43 males and 69 females. Our aesthetic voting scheme is simple. All participating users can only choose one image from Fig. 12(a)-(c)

⁴[Online]. Available: <http://gallery.vivo.com.cn/>

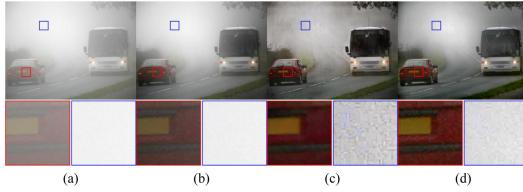


Fig. 13. An example of failure case with dense fog on the Foggy Driving-dense dataset. (a) Input image with dense fog; (b) MSBDN [25]; (c) DA [33]; and (d) Proposed method.

according their aesthetic habits. Compared with the sample image Fig. 12(b) provided by vivo, the image Fig. 12(c) edited with our method has achieved more vivid color and higher saturation in the background of the portrait. Fig. 12(d) illustrates that Fig. 12(c) edited with our method gains more popularity than images Fig. 12(a) and (b).

H. Failure Case

Fig. 13 shows an example of failure case of MSBDN [25], DA [33] and our methods on dense fog on the Foggy Driving-dense dataset⁵. MSBDN just removes the slight fog (see Fig. 13(b)), which is less different from the original foggy image. Compared to MSBDN, DA and the proposed method generate high contrast dehazed images but with diverse color/noise artifacts. Our method has difficulty removing the fog with dense density due to pre-processing steps (GC and CP-AHE) with limited visibility and contrast adjustment, but this is a major challenge for most current algorithms, which may be addressed in the future work by taking depth information of the scenario into account [1].

V. CONCLUSION AND FUTURE WORK

We proposed a fast multi-scale patch-based fusion framework for real-world image dehazing. It does not rely on atmospheric physical models and not require expensive training resources. This is achieved by formulating the single image dehazing task as an image enhancement problem, and blending a set of pre-processing images that are derived from GC and CP-AHE. Further, we developed a fast multi-scale structural patch decomposition-based fusion strategy for fusing derived inputs. Experiments demonstrate the superiority of the proposed method over current dehazing methods.

In future work, there are following three aspects: 1) To remove dense fog mentioned in the failed case and facilitate real-world semantic segmentation and object detection, we will integrate depth maps into our framework [1]; 2) The proposed CP-AHE reduces color/noise artifacts and increases saturation of the color fog-free image compared to AHE [45]. However, there are still some visible artifacts in the net result, especially in the dense fog (see Fig. 13(d)), which may be addressed by further studying channel-aware AHE; 3) We will explore the possibility of applying the proposed framework to other harsh conditions, such as low-light, underwater and sandstorm.

⁵[Online]. Available: http://people.ee.ethz.ch/~csakarid/SFSU_synthetic/

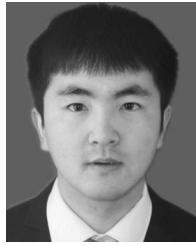
ACKNOWLEDGEMENT

The authors would like to thank Dr. Z. Li of the Institute for Infocomm Research, Singapore, for fruitful discussions and constructive suggestions.

REFERENCES

- [1] C. Sakaridis, D. Dai, and L. Van Gool, “Semantic foggy scene understanding with synthetic data,” *Int. J. Comput. Vis.*, vol. 126, no. 9, pp. 973–992, 2018.
- [2] B. Li *et al.*, “Benchmarking single-image dehazing and beyond,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.
- [3] Y. Li, S. You, M. S. Brown, and R. T. Tan, “Haze visibility enhancement: A survey and quantitative benchmarking,” *Comput. Vis. Image Understanding*, vol. 165, pp. 1–16, 2017.
- [4] S. Zhao, L. Zhang, S. Huang, Y. Shen, and S. Zhao, “Dehazing evaluation: Real-world benchmark datasets, criteria and baselines,” *IEEE Trans. Image Process.*, vol. 29, pp. 6947–6962, 2020.
- [5] S. G. Narasimhan and S. K. Nayar, “Contrast restoration of weather degraded images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 713–724, Jun. 2003.
- [6] R. Fattal, “Single image dehazing,” *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
- [7] R. T. Tan, “Visibility in bad weather from a single image,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [8] 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.
- [9] R. Fattal, “Dehazing using color-lines,” *ACM Trans. Graph.*, vol. 34, no. 1, pp. 1–14, 2014.
- [10] Z. Li *et al.*, “Simultaneous video defogging and stereo reconstruction,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 4988–4997.
- [11] Q. Wu, J. Zhang, W. Ren, W. Zuo, and X. Cao, “Accurate transmission estimation for removing haze and noise from a single image,” *IEEE Trans. Image Process.*, vol. 29, pp. 2583–2597, 2019.
- [12] S. Salazar-Colores *et al.*, “A fast image dehazing algorithm using morphological reconstruction,” *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2357–2366, May 2019.
- [13] 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.
- [14] W. Ren *et al.*, “Single image dehazing via multi-scale convolutional neural networks,” in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 154–169.
- [15] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “AOD-Net: All-in-one dehazing network,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 4770–4778.
- [16] H. Zhang and V. M. Patel, “Densely connected pyramid dehazing network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3194–3203.
- [17] J. Pan *et al.*, “Learning dual convolutional neural networks for low-level vision,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3070–3079.
- [18] W. Ren, J. Pan, H. Zhang, X. Cao, and M.-H. Yang, “Single image dehazing via multi-scale convolutional neural networks with holistic edges,” *Int. J. Comput. Vis.*, vol. 128, no. 1, pp. 240–259, 2020.
- [19] X. Yang, Z. Xu, and J. Luo, “Towards perceptual image dehazing by physics-based disentanglement and adversarial training,” in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 7485–7492.
- [20] S. Santra, R. Mondal, and B. Chanda, “Learning a patch quality comparator for single image dehazing,” *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4598–4607, Sep. 2018.
- [21] A. Golts, D. Freedman, and M. Elad, “Unsupervised single image dehazing using dark channel prior loss,” *IEEE Trans. Image Process.*, vol. 29, pp. 2692–2701, 2019.
- [22] J. Pan *et al.*, “Physics-based generative adversarial models for image restoration and beyond,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 7, pp. 2449–2462, Jul. 2021.
- [23] M. Sulami, I. Glatz, R. Fattal, and M. Werman, “Automatic recovery of the atmospheric light in hazy images,” in *Proc. IEEE Int. Conf. Comput. Photography*, 2014, pp. 1–11.
- [24] 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.

- [25] H. Dong *et al.*, “Multi-scale boosted dehazing network with dense feature fusion,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 2157–2167.
- [26] W. Ren *et al.*, “Gated fusion network for single image dehazing,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3253–3261.
- [27] R. Li, J. Pan, Z. Li, and J. Tang, “Single image dehazing via conditional generative adversarial network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8202–8211.
- [28] X. Liu, Y. Ma, Z. Shi, and J. Chen, “GriddehazeNet: Attention-based multi-scale network for image dehazing,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 7314–7323.
- [29] Y. Qu, Y. Chen, J. Huang, and Y. Xie, “Enhanced pix2pix dehazing network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8160–8168.
- [30] X. Liu, M. Suganuma, Z. Sun, and T. Okatani, “Dual residual networks leveraging the potential of paired operations for image restoration,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7007–7016.
- [31] 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, 2019.
- [32] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, “FFA-Net: Feature fusion attention network for single image dehazing,” in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 11908–11915.
- [33] 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.*, 2020, pp. 2808–2817.
- [34] A. Dudhane, K. M. Biradar, P. W. Patil, P. Hambarde, and S. Murala, “Varicolored image de-hazing,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 4564–4573.
- [35] Y. Zhang, L. Ding, and G. Sharma, “Hazerd: An outdoor scene dataset and benchmark for single image dehazing,” in *Proc. IEEE Int. Conf. Image Process.*, 2017, pp. 3205–3209.
- [36] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, “O-haze: A dehazing benchmark with real hazy and haze-free outdoor images,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 754–762.
- [37] C. O. Ancuti, C. Ancuti, and R. Timofte, “NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2020, pp. 1798–1805.
- [38] A. Galdran, A. Alvarez-Gila, A. Bria, J. Vazquez-Corral, and M. Bertalmío, “On the duality between retinex and image dehazing,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8212–8221.
- [39] C. O. Ancuti and C. Ancuti, “Single image dehazing by multi-scale fusion,” *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3271–3282, Aug. 2013.
- [40] L. K. Choi, J. You, and A. C. Bovik, “Referenceless prediction of perceptual fog density and perceptual image defogging,” *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3888–3901, Nov. 2015.
- [41] J.-M. Guo, J.-Y. Syue, V. R. Radzicki, and H. Lee, “An efficient fusion-based defogging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4217–4228, Sep. 2017.
- [42] A. Galdran, “Image dehazing by artificial multiple-exposure image fusion,” *Signal Process.*, vol. 149, pp. 135–147, 2018.
- [43] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik, “Day and night-time dehazing by local airlight estimation,” *IEEE Trans. Image Process.*, vol. 29, pp. 6264–6275, 2020.
- [44] T. Mertens, J. Kautz, and F. Van Reeth, “Exposure fusion: A simple and practical alternative to high dynamic range photography,” in *Comput. Graph. Forum*, vol. 28, no. 1. Wiley Online Library, 2009, pp. 161–171.
- [45] S. M. Pizer *et al.*, “Adaptive histogram equalization and its variations,” *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.
- [46] H. Li, K. Ma, H. Yong, and L. Zhang, “Fast multi-scale structural patch decomposition for multi-exposure image fusion,” *IEEE Trans. Image Process.*, vol. 29, pp. 5805–5816, 2020.
- [47] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, “Efficient image dehazing with boundary constraint and contextual regularization,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 617–624.
- [48] C. Chen, M. N. Do, and J. Wang, “Robust image and video dehazing with visual artifact suppression via gradient residual minimization,” in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 576–591.
- [49] D. Berman, T. Treibitz, and S. Avidan, “Single image dehazing using hazelines,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 720–734, Mar. 2020.
- [50] Y. Li, R. T. Tan, and M. S. Brown, “Nighttime haze removal with glow and multiple light colors,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 226–234.
- [51] J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. Wen Chen, “Fast haze removal for nighttime image using maximum reflectance prior,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7418–7426.
- [52] H.-M. Hu, H. Zhang, Z. Zhao, B. Li, and J. Zheng, “Adaptive single image dehazing using joint local-global illumination adjustment,” *IEEE Trans. Multimedia*, vol. 22, no. 6, pp. 1485–1495, Jun. 2020.
- [53] F. Fang, T. Wang, Y. Wang, T. Zeng, and G. Zhang, “Variational single image dehazing for enhanced visualization,” *IEEE Trans. Multimedia*, vol. 22, no. 10, pp. 2537–2550, Oct. 2020.
- [54] M. Ju, C. Ding, Y. J. Guo, and D. Zhang, “IDGCP: Image dehazing based on gamma correction prior,” *IEEE Trans. Image Process.*, vol. 29, pp. 3104–3118, 2019.
- [55] Z. Li and J. Zheng, “Single image de-hazing using globally guided image filtering,” *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 442–450, Jan. 2018.
- [56] Y. Pang, J. Nie, J. Xie, J. Han, and X. Li, “BidNet: Binocular image dehazing without explicit disparity estimation,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 5931–5940.
- [57] Z. Deng *et al.*, “Deep multi-model fusion for single-image dehazing,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 2453–2462.
- [58] C.-H. Yeh, C.-H. Huang, and L.-W. Kang, “Multi-scale deep residual learning-based single image haze removal via image decomposition,” *IEEE Trans. Image Process.*, vol. 29, pp. 3153–3167, 2019.
- [59] S. Zhao, L. Zhang, Y. Shen, and Y. Zhou, “RefinedNet: A weakly supervised refinement framework for single image dehazing,” *IEEE Trans. Image Process.*, vol. 30, pp. 3391–3404, 2021.
- [60] S. E. Kim, T. H. Park, and I. K. Eom, “Fast single image dehazing using saturation based transmission map estimation,” *IEEE Trans. Image Process.*, vol. 29, pp. 1985–1998, 2019.
- [61] J. Mairal, M. Elad, and G. Sapiro, “Sparse representation for color image restoration,” *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.
- [62] M. Lebrun, M. Colom, and J.-M. Morel, “Multiscale image blind denoising,” *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3149–3161, Oct. 2015.
- [63] F. Zhu, G. Chen, and P.-A. Heng, “From noise modeling to blind image denoising,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 420–429.
- [64] K. Ma *et al.*, “Robust multi-exposure image fusion: A structural patch decomposition approach,” *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, May 2017.
- [65] H. Li, X. Jia, and L. Zhang, “Clustering based content and color adaptive tone mapping,” *Comput. Vis. Image Understanding*, vol. 168, pp. 37–49, 2018.
- [66] 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.
- [67] Z. Li and J. Zheng, “Edge-preserving decomposition-based single image haze removal,” *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5432–5441, Dec. 2015.
- [68] Z. Li, Z. Wei, C. Wen, and J. Zheng, “Detail-enhanced multi-scale exposure fusion,” *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1243–1252, Mar. 2017.
- [69] F. Kou, Z. Li, C. Wen, and W. Chen, “Edge-preserving smoothing pyramid based multi-scale exposure fusion,” *J. Vis. Commun. Image Representation*, vol. 53, pp. 235–244, 2018.
- [70] J. Zhang, Y. Cao, Z.-J. Zha, and D. Tao, “Nighttime dehazing with a synthetic benchmark,” in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 2355–2363.
- [71] Q. Zhang, G. Yuan, C. Xiao, L. Zhu, and W.-S. Zheng, “High-quality exposure correction of underexposed photos,” in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 582–590.
- [72] C. Guo *et al.*, “Zero-reference deep curve estimation for low-light image enhancement,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 1780–1789.



Xiaoning Liu is currently working toward the Ph.D. degree with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. His research interests include image processing, computational imaging, and computer vision.



Hui Li received the Ph.D. degree in computer science from The Hong Kong Polytechnic University, Hong Kong, in 2019. He is currently a Senior Imaging Algorithm Engineer with vivo Mobile Communication Co. Ltd., Shenzhen, China. Prior to that, he was a Senior Engineer with Huawei 2012 Labs, Shenzhen, China.

His research interests include image processing, computational photography, and computer vision.



Ce Zhu (Fellow, IEEE) received the B.S. degree in electronic and information engineering from Sichuan University, Chengdu, China, in 1989, and the M.Eng. and Ph.D. degrees in electronic and information engineering from Southeast University, Nanjing, China, in 1992 and 1994, respectively. He held a Postdoctoral Research Position with the Chinese University of Hong Kong, Hong Kong, in 1995, the City University of Hong Kong, Hong Kong, and the University of Melbourne, Melbourne, VIC, Australia, from 1996 to 1998. For 14 years from 1998 to 2012, he was with Nanyang Technological University, Singapore, where he was a Research Fellow, a Program Manager, an Assistant Professor, and then promoted to an Associate Professor in 2005. Since 2012, he has been with the University of Electronic Science and Technology of China, Chengdu, China, as a Professor. His research interests include video coding and communications, video analysis and processing, 3-D video, visual perception and applications. Prof. Zhu is an APSIPA Distinguished Lecturer from 2021 to 2022, and an IEEE Distinguished Lecturer of Circuits and Systems Society from 2019 to 2020. He was/is on the Editorial Boards of a few journals, including an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON BROADCASTING, and IEEE Signal Processing Letters, the Editor of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, and an Area Editor of the *Signal Processing: Image Communication*. He was also a Guest Editor of a few special issues in international journals, including as the Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.