Coefficient Optimization for Dehazed Images Based on Derived Brightness Function and Contrastive Learning

Ping Juei Liu[1*]

ABSTRACT

In this study, we demonstrate a simplified framework to reduce the coefficient optimization tasks, and the outcome is better than the conventional method. The deep-learning-based dehazing algorithms have recently been the most common architecture in developing dehazing algorithms, resulting in many end-to-end dehazing algorithms. However, flexibility is a problem as the effectiveness of the deep-learning-based methods usually falls short in dealing with the thick haze. Moreover, these methods are built without a configurable framework to adapt to the issue. A configurable framework usually involves a coefficient to adjust the performance; however, choosing a suitable coefficient is generally annoying. To address this, we propose and evaluate several coefficient optimization methods for dehazing algorithms, including an optimization method based on contrastive learning. Our method can overcome the inconvenience, improve performance, and eventually extend the application of dehazing algorithms. Our approach is proven superior to the extant method through well-controlled experiments, guaranteeing flexible usage and better performance of the configurable dehazing algorithm.

Keywords: Haze, Dehaze, Haze Removal, Coefficient Optimization, Self-adaptive Coefficient, Contrastive Learning

I. INTRODUCTION

Dehazing algorithms aimed to restore the actual scene from hazy images. Conventionally, these algorithms deal with the contrast and saturation degradation caused by scattering, such as the dark channel prior (DCP) [1], boundary constraint and contextual regularization [2], and non-local image dehazing [3]. However, the goal has been changed recently since more dehazing algorithms emphasized restoration accuracy. For example, the color attenuation prior [4], AOD-Net [5], gated fusion network [6], multi-scale boosted Dehazing network [7], and joint contrast enhancement and exposure fusion [8]. Contemporary dehazing algorithms adopted machine learning or deep learning frameworks, and increasing haze datasets have made deep-learning-based algorithms powerful, resulting in better restoration accuracy and convenience [9].

On the other hand, instead of estimating the transmission, the end-to-end framework became mainstream when developing dehazing algorithms, thanks to the popularization of deep learning algorithms. The end-to-end framework meant that research did not estimate the transmission map like that in [1]-[5]; they generated dehazed images from hazy images directly. The widely-used end-to-end algorithms included the deep residual network (ResNet) [10], ResNeXt [11], densely-connected network (DenseNet) [12], FFA-Net [13], and transformer network [14]. These networks' architecture differed; for example, ResNet included a mechanism re-considering the filtered information discarded by previous layers, DenseNet emphasized connections inside the model, and the transformer was attention-oriented. These networks were beneficial in restoration accuracy; however, emphasis on restoration accuracy sometimes resulted in the need for more enhancement ability and was adverse to enhancing images captured in exceptional cases.

Natural environments were complex to the extent that people had no choice but to acquire images under harsh environments. The common issues were the thick haze, underwater environments, low natural illumination, and high artificial light sources. These issues caused effects similar to the haze due to scattering; therefore, dehazing algorithms with slight changes were still considered standard tools to eliminate the haze-like effects. Unfortunately, deep-learning-based dehazing algorithms had no contingency to adapt to the small changes because they adopted the end-to-end framework; therefore, owing to different ontologies, they usually fell short in dealing with various cases. In other words, a re-training was necessary, not only taking additional time and energy but also inconvenience. Therefore, specific methods aiming at different issues were proposed, such as algorithms for the thick haze [15][16], the underwater image [17][18], the low-visibility image captured in low light environments [19] or at nighttime [20][21]. Nevertheless, the core technology of these methods was usually highly correlated to that of dehazing algorithms.

*Corresponding Author: Ping Juei Liu (E-mail:immich0716@gmail.com).
[1] Department of Artificial Intelligence and Engineer, National Chin-Yi University of Technology, No.57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung 411030, Taiwan (R.O.C.).

As discussed previously, Liu et al. reported that deep-learning-based dehazing algorithms were usually weak when dealing with the thick haze. To overcome the issues, Liu et al. proposed a configurable dehazing algorithm [22] based on the fact that any dehazing algorithm involved a tradeoff between brightness and contrast; therefore, any configuration made to a dehazing algorithm tending to boost contrast inevitably resulted in brightness degradation. Accordingly, using a unique coefficient, their dehazing algorithm looked forward to a favorable tradeoff between brightness and contrast. New research adopted the concept to implement a two-step enhancement to brightness and contrast [19][18]. According to the author's report, the coefficient was proven stable and highly associated with the quality of the dehazed image; thereby, the performance of the new dehazing method outperformed state-of-the-art dehazing algorithms in the experiment. Meanwhile, the coefficient also improved the flexibility to extend the application of dehazing algorithms and was worthy of further study; for example, the coefficient still needed to be completely self-adaptive. We figured out that training a coefficient-optimization model in the conventional deep-learning framework was time-consuming because of the need for an optimal coefficient dataset; this means researchers have to try every possible coefficient to obtain the best one. Therefore, this study aims to simplify the training process, shortening the time of developing an automatic coefficient generator, improving performance and convenience, and eventually extending the application of the dehazing algorithm.

We use three methods to optimize a configurable dehazing framework proposed in [19] to achieve our goals. In this framework, manipulation of a coefficient can directly affect the contrast and brightness of the dehazed image. To exert the advantage, we proposed a polynomial fitting model (PFM), ResNet network (RNet), and ResNet model based on contrastive regularization term (CLNet) to optimize the coefficient. The three methods use different training strategies, and the advantages differ. Corresponding experimental results show that CLNet improves the restoration accuracy compared to the origin algorithms, which use a fixed coefficient. On the other hand, CLNET also reduces the complexity of the training process because computing the best coefficients is unnecessary in CLNET's training framework.

II. RELATED WORKS

A. *Dark Channel Prior*

Based on the linear atmospheric model obeying Koschmieder's Law [1], an equation involves the hazy image I, transmission t, global atmospheric light A, and dehazed image J is written as:

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

This equation contains three unknown variables, making the widely used dehazing model indeterminate. Generally, the global atmospheric light is estimated individually from the hazy image; thus, when both sides of (1) are divided by A, generating a normalized-form equation, and the result is as follows:

$$\hat{I} = \hat{J}t + (1-t), \tag{2}$$

where the hat symbol indicates a variable divided by the global atmospheric light. After using the minimal operator concerning three RGB color channels and a small patch in each image, (2) can be rewritten as follows:

$$\hat{I}_{min} = \hat{J}_{min}t + (1-t), \tag{3}$$

where $\hat{I}_{min}$ and $\hat{J}_{min}$ respectively denote dark channels of I and J. The dark channel prior (DCP) proposed in [1] is a theory assuming that the darkest response among the RGB color channel in any clear images' small region is close to zero. Applying the DCP assumption to (3) makes $\hat{J}_{min}$ equal to zero, so we obtain the following:

$$\tilde{t} = 1 - \hat{I}_{min}, \tag{4}$$

where $\tilde{t}$ is an initial transmission estimate awaiting a refinement process to compensate for the distinct edge caused by the minimal operator concerning a small patch. The refinement process is essential because the distinct edge between the hazy image and transmission produces unnatural artifacts in the dehazed image. Therefore, the estimated transmission is obtained using the following:

$$\tilde{t}^r = 1 - \hat{I}^r_{min}, \tag{5}$$

where $\tilde{t}^r$ is the estimated transmission, and $\hat{I}^r_{min}$ denotes the refined dark channel. In the DCP framework, we can obtain the dehazed image by substituting the estimated transmission $\tilde{t}^r$ for the transmission t demonstrated in (1).

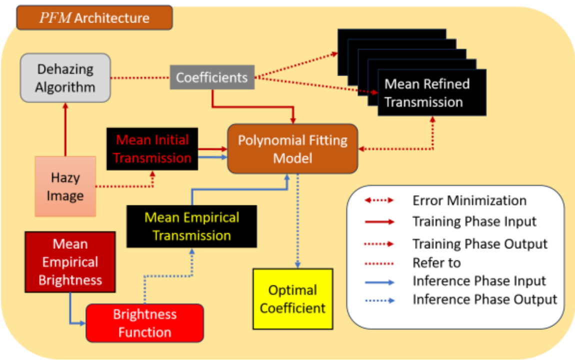
B. *Configurable Dehazing Algorithm*

Liu et al. report that refinement processes significantly correlate to saturation, brightness, and contrast quality [22]. Therefore, they propose a new energy function associated with saturation and contrast in this study, and the energy term is written as follows:

$$\underset{\hat{I}^r_{min}}{arg\,max}\ \delta\left(\hat{I}^r_{min} - \hat{I}_{min}\right)^2 + \left(\nabla \ln \hat{I}^r_{min} - \nabla \ln \hat{I}_{min}\right)^2 \tag{6}$$

where δ represents an ideal monotonically decreasing function ensuring the maximal input results in the zero output. The first and second terms in (6) are called luminance and contrast terms associated with the brightness (or saturation) and contrast of the dehazed image, respectively. Accordingly, they propose a prototype dehazing algorithm to improve the contrast of the dehazed image and suggest using an $\hat{I}^r_{min}$ obtained by subtracting the high-frequency component from $\hat{I}_{min}$ to maximize the contrast term.

The idea is used in the latest research aiming at improving visibility [19], and the author uses a two-step method to enhance the brightness and contrast at each step, resulting in bright and contrasting results. In this framework, the dehazing algorithm used at each step is defined as follows:



(a)

Figure 1. The PFM architecture includes two steps. In the training step (red lines), the polynomial fitting model attempts to fit the mean refined transmission to the mean initial transmission and coefficients. After that, with the brightness chosen empirically, we compute the mean empirical transmission based on the brightness function. In the reference step, the optimal coefficient of the dehazing algorithm can be solved by inputting the initial transmission and empirical brightness to the polynomial fitting model.

$$P_{dc} = ln\,\hat{I}_{min} - \alpha|L_h|, \tag{7}$$

where P_{dc} is an un-refined dark channel associated with the estimated transmission, α is a coefficient, and L_h is a high-frequency component defined as follows:

$$L_h = ln\,\hat{I}_{min} - \varphi(ln\,\hat{I}_{min}), \tag{8}$$

where φ is a smoothing algorithm based on the guided image filter [23]. Afterward, they obtain $\hat{I}^r_{min}$ by refining P_{dc} using φ and use $\hat{I}^r_{min}$ to compute the estimated transmission. Accordingly, the estimated transmission allows the calculation of the dehazed image according to (5) and (2). Except for producing the dehazed image, note that this step results in a relation as follows:

$$ln\,\hat{I}^r_{min} \propto P_{dc} = ln\,\hat{I}_{min} - \alpha|L_h|. \tag{9}$$

Combining (9) and (6), the brightness term can be rewritten as follows:

$$\hat{I}^r_{min} - \hat{I}_{min} \propto \alpha|L_h|. \tag{10}$$

Therefore, α is proportional to the brightness of the dehazed image. On the other hand, the contrast term can be rewritten as follows:

$$\nabla\, ln\,\hat{I}^r_{min} - \nabla\, ln\,\hat{I}_{min} \approx \alpha|L_h|. \tag{11}$$

As a result, α is proportional to the contrast of the dehazed image. The dehazing algorithm is interesting and unique because it generates the dehazed image wherein α controls the brightness and contrast.

C. *Derivations Associated with Brightness*

Replacing the minima operator with the maxima operator in (3) and deriving a formula associated with $\tilde{t}^r$ generates the following:

$$\tilde{t}^r = \frac{1-\hat{I}_{max}}{1-\hat{J}_{max}} = \frac{S_{\hat{I}}}{S_{\hat{J}}}, \tag{12}$$

where $S_{\hat{I}}$ and $S_{\hat{J}}$ respectively denotes the negative image of $\hat{I}_{max}$ and $\hat{J}_{max}$. Note that $\hat{I}_{max}$ is the value channel of HSV color space, thereby representing brightness of $\hat{I}$. Accordingly, a relation between transmission and brightness can be written as follows:

$$S_{\hat{J}} = S_{\hat{I}}/\tilde{t}^r. \tag{13}$$

The above equation results in the following:

$$\mu_{g,S_{\hat{J}}} = \mu_{g,S_{\hat{I}}}/\mu_{g,\tilde{t}^r}, \tag{14}$$

where $\mu_{g,S_{\hat{J}}}$, $\mu_{g,S_{\hat{I}}}$, and $\mu_{g,t}$ respectively denote the geometric means of $S_{\hat{J}}$, $S_{\hat{I}}$, and t. The equation helps derive our approaches, indicating that the brightness of a hazy image and the estimated transmission ensure the dehazed image's brightness. For convenience, we call (14) the brightness function.

III. THE PROPOSED METHOD

In this study, we derive and discuss several feasible approaches to automatically generate the optimal coefficient α based on the information within the input image. We use a dehazing algorithm proposed in [19] rather than that offered in [22] owing to the better performance.

A. *Polynomial Fitting Model*

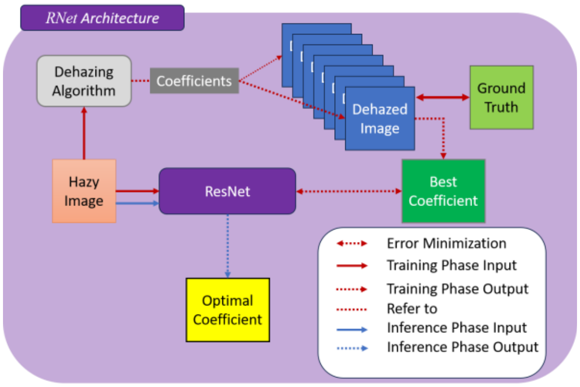
Liu et al. mention a simple coefficient-optimizing method based on polynomial fitting [22]. They compiled statistics on images in the synthetic objective testing set (SOTS) [9]. They found a solid correlation: the arithmetic means linearly correlated to the initial and estimated transmission with a correlation coefficient >0.98. Thus, they concluded that this should result from the coefficient. Accordingly, they conducted an advanced experiment and found that the estimated transmission was exponentially proportional to the coefficient α. Accordingly, they conducted another experiment and found that the estimated transmission was exponentially proportional to the coefficient α. Hence, they suggested using the correlation as follows:

$$\mu_{g,t^r} \approx \sum_{m=0}^{1}\sum_{n=0}^{3} c_p \alpha^n \mu^m_{g,t}, \tag{15}$$

where c_p indicates a serial constant; μ_{g,t^r} and $\mu_{g,t}$ are the geometric means of $\tilde{t}^r$ and $\tilde{t}$, respectively; m and n are two parameters respectively denoting the order of α and $\mu_{g,t}$. This model helps choose an optimal coefficient to generate the quality dehazed image from the initial transmission $\mu_{g,t}$, which can be computed from the hazy image. After repeated tests, we find that the higher order terms of α appear not helpful to overall performance for the dehazing image; thus, we use a correlation demonstrated as follows:

$$\mu_{g,t^r} \approx \sum_{m=0}^{1}\sum_{n=0}^{1} c_p \alpha^n \mu^m_{g,t}. \tag{16}$$

Fig. 1(a) illustrates the flowchart of fitting our polynomial fitting model (PFM) architecture. Red lines



(a)

Figure 3. The RNet architecture involves the ResNet to directly fit the optimal coefficient to the hazy image, so, in the reference step, the optimal coefficient of the dehazing algorithm can be solved easily; however, the training phase is time-consuming.

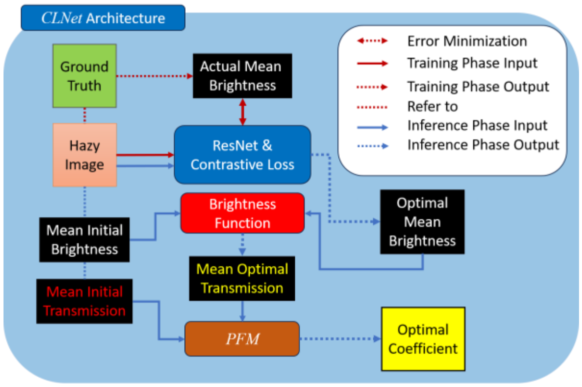
denote steps associated with the training phase, and the blue lines indicate the reference steps. Accordingly, once the optimal brightness of the dehazed image is chosen, the corresponding μ_{g,t^r} can be obtained according to the brightness function. Afterward, an optimal α can be calculated by solving (16). We give μ_{g,s_j} a fixed value of 0.45, as the author's suggestion in [22]; for your reference, the optimal brightness comes from the statistics proposed in [23]. Note that a compromise has been granted since the statistics are associated with arithmetic means while (14) uses geometric means; fortunately, the performance of PFM is acceptable in practice. Generally, the fast execution time is the most crucial advantage of this method, but the performance is limited because PFM uses a fixed empirical brightness.

B. *Self-adaptive Coefficient Based on ResNet*

Conventionally, a neural network can simulate any relation between factors; therefore, it can generate an utterly optimal coefficient from the input image, achieving aself-adaptive process. A ResNet defined as $\Psi_r: R^{H\times W\times N} \to R^1$ and based on the L2-norm loss can be written as follows:

$$\underset{\Psi_r}{arg\,min}\ (\Psi_r(X,\omega_r) - \alpha_t)^2 + \beta_r\rho_r(\Psi_r(X,\omega_r)), \tag{17}$$

where H, W, and N respectively denote the height, width, and number of the hazy image; $X \in R^{H\times W\times N}$ is the input of Ψ_r; ω_r denotes parameters used in Ψ_r; $\alpha_t \in R^1$ represents the optimal α obtained through experiments based on X, and β_r is a coefficient. The first term is the reconstruction loss to force the estimated α to be close to α_t, and the second term is a regularization term including a function ρ_r used to prevent extreme situations or smoothen the output. We use ResNet-18, ResNet-34, and ResNet-50 in this study. These networks are designed for image recognition, thereby naturally generating the result of R^1; we refer to [10] for detailed information on



(a)

Figure 2. The CLNet architecture involves the ResNet based on contrastive loss to fits he actual mean brightness to the hazy image. In the reference step, the ResNet generates the optimal mean brightness from the hazy image; after that, we estimate the mean optimal transmission based on the brightness function. Finally, PFM generates an optimal coefficient for the dehazing algorithm.

these networks. Note that the regularization term is ignored since the widely-used terms, such as TV-norm [25], DCP terms, or L1-norm, are meaningless while the output of Ψ_r is a constant.

Fig. 2(a) illustrates the architecture of our network (RNet). Similarly, red lines denote steps associated with the training phase, and the blue lines indicate the reference steps. The time-consuming part lies in the dehazing process and the benchmark computing between the dehazed images and ground truth. We use mean square error (MSE) to measure the quality of dehazed images in this study. The advantage of RNet is its theoretical soundness and architectural intuition, making RNet perform well. However, when the number of training and validating datasets is large, the cost of training such a network is considerably huge owing to a time-consuming calculation of the best coefficient α_t.

C. *Self-adaptive Coefficient Based on Contrastive Learning*

When conducting the above experiments, we observed that the coefficient holds a solid relation to the dazed image's brightness; this means that even with a limited training set, PFM remains accurate. On the contrary, the performance of RNet strongly relies on the number of the training set. The conclusion inspires us with a new framework capable of generating the optimal coefficient with a simplified process.

As previously discussed, computing the optimal coefficient for each image in the dataset across all coefficients can enhance visual quality significantly. However, this approach is time-consuming. Hence, we seek an alternative benchmark. Leveraging the brightness function proves highly beneficial in simplifying the training phase of the coefficient optimization framework; this is because computing the brightness of both hazy and dehazed images is comparatively easier than refining transmission, suggesting that we can estimate the optimal brightness of the dehazed image from the hazy image instead of directly estimating the optimal coefficient. Consequently, we can train a network to accurately estimate $\mu_{g,s_{\tilde{J}}}$ from the hazy image, enabling rapid generation of the optimal α based on the brightness function and PFM.

Fig. 3(a) illustrates the architecture of our contrastive learning network (CLNet). The red lines show that the training phase is more straightforward than that of RNet; on the other hand, the blue lines show that the inference process can be done rapidly, too. Summarily, CLNet estimates $\mu_{g,s_{\tilde{J}}}$, so we can quickly obtain corresponding $\mu_{g,\tilde{t}^r}$ according to the brightness function since $\mu_{g,s_{\tilde{I}}}$ is known. After that, PFM rapidly generates an optimal α. The preceding discussion highlights several advantages of CLNet, which include:

- Individual brightness estimation for each image in the dataset, as opposed to a fixed estimation derived from statistics.
- Avoidance of compromises by utilizing arithmetic means instead of geometric means.
- Streamlining the training process by circumventing time-consuming calculations required to obtain the optimal coefficient for each image.

To improve the performance of ResNet, we introduce a network based on contrastive regularization, denoted as $\Psi_{rc}: R^{H\times W\times N} \to R^1$; the architecture is inspired by the loss function proposed in [26]-[29], as follows:

$$\underset{\Psi_{rc}}{arg\,min}\ (\Psi_{rc}(X,\omega_{rc}) - \mu_{g,s_{\tilde{J}}})^2 + \beta_{rc}\rho_{rc}\left(\Psi_{rc}(X,\omega_{rc}), \mu_{g,s_{\tilde{J}}}, \mu_{g,s_{\tilde{I}}}\right), \quad (18)$$

where Ψ_{rc} is a ResNet based on the contrastive term ρ_{rc} controlled by a parameter ρ_{rc}, and ω_{rc} indicates parameters used in Ψ_{rc}. More specifically, ρ_{rc} is a function defined as follows:

$$\rho_{rc}(.) = \frac{\mu_{g,s_{\tilde{J}}} - \Psi_{rc}(X,\omega_{rc})}{\mu_{g,s_{\tilde{I}}} - \Psi_{rc}(X,\omega_{rc})}. \quad (19)$$

We use β_{rc} to balance the reconstruction and contrastive terms. Note that the contrastive term pulls the output to ideal brightness (based on the clear image) and pushes the output far from unfavorable brightness (based on the hazy image), so it should improve the performance of this model.

IV. EXPERIMENTAL ANALYSIS

The experiment is implemented by PyTorch 2.01 with Python 3.9.16. We use NVIDIA GTX 3060 12G under an environment based on CUDA 11.8 and CUDNN 8.6.0. We train our networks using Adam optimizer with β_1 and β_2 respectively equal to 0.9 and 0.99. Besides, the initial learning rate and batch size are set to 0.0002 and 16, respectively, and the learning rate is adjusted using the cosine annealing method [30]. Empirically, the balancing coefficient β_r and β_{rc} are set to 0.1 and 0.15, and the total number of the epoch is 100. As for training sets, the RESIDE dataset [9] is a commonly used dataset in which the indoor training set (ITS) and outdoor training set (OTS) are widely used in training networks. The details of our experiment are as follows.

First, we calculate the geometric mean of the initial ($\mu_{g,t}$) and estimated transmission (μ_{g,t^r}), based on α from 0 to 10 with a fixed increasing step of 0.25. After that, we mark every image's optimal coefficient (α_t) in ITS and OTS as labels. Meanwhile, we also calculate the geometric means of the negative image associated with the clear ($\mu_{g,s_{\tilde{J}}}$) and hazy images ($\mu_{g,s_{\tilde{I}}}$) based on the same α. After the preprocessing, we fit μ_{g,t^r} to a function composed of α and $\mu_{g,t}$ to create PFM, as demonstrated in (16); this results in a serial coefficient c_p of (0.0005926, 0.2795, 1.051, −0.1841). As for RNet, it is trained based on α_t, ITS, and OTS. Besides, we train CLNet with the same datasets except for using $\mu_{g,s_{\tilde{J}}}$ instead of α_t. With the help of CLNet, we obtain μ_{g,t^r} according to the brightness function in (14), so an optimal α is available by solving PFM in (16).

We choose Several state-of-the-art dehazing algorithms for comparison, including the dark channel prior (DCP) [1], boundary constraint and contextual regularization (BCCR) [2], non-local image dehazing (NLD) [3], color attenuation prior (CAP) [4], AOD-net (AOD) [5], gated fusion network (GFN) [6], contrast enhancement and exposure fusion algorithm (CEE) [8], contrast in haze removal algorithm (CIH) [22], and Low Visibility Image Enhancement (LVE) [19]. The benchmarks used in our experiments include:

- The mean square error (MSE).
- The peak signal-to-noise ratio (PSNR).
- The structural similarity index (SSIM) [31].
- The CIEDE2000 [32].
- The F&T (including FSITM [33] and TMQI [34]).
- The regular DehazeFR [35].
- The image quality evaluation (PIQE) [36].
- The image quality assessment (NRIQA) [37].
- The blind image quality evaluation (BIQE) [38].

MSE, PSNR, SSIM, and CIEDE2000 can evaluate image quality in the spatial domain concerning the square error, signal-to-noise ratio, structure similarity, and color bias. Besides, F&T assesses features in the frequency domain; we use this for comparison with spatial-domain benchmarks. Moreover, we use the regular DehazeFR [35] to evaluate the performance regarding structure recovery, color renditions, and suppression of over-enhancement. We also test our method using non-reference benchmarks; these benchmarks can assess the quality of contrast, brightness, and naturalness. The non-reference benchmarks are beneficial to compare with reference benchmarks, ensuring algorithms generate accurate and quality results. Eventually, we conduct the evaluations on two datasets, including the synthetic objective testing set (SOTS) [9] in the RESIDE dataset and the HazeRD dataset (HRD) [39].

A. Evaluation of Restoration Accuracy

Our approach aims to generate optimal coefficients for a dehazing algorithm proposed in [19]; we adopt LVE as the dehazing algorithm. Note that our methods provide optimal coefficients based on PFM, RNet, and CLNet, whereas LVE employs a fixed α, We conducted experiments to assess the performance of our methods alongside several state-of-the-art dehazing techniques using hazy images from SOTS and HRD datasets. Notably, in this experiment, our methods are based on ResNet-34, as depicted in Table I. SOTS encompasses both indoor and outdoor scenes. At the same time, HRD includes countryside scenes, each exhibiting synthesized haze. Empirically, the synthesized haze in SOTS tends to be relatively thin, whereas some HRD images feature thicker haze. Conventional dehazing algorithms generally yield average results across both test sets, whereas deep learning-based methods exhibit distinct performance, particularly excelling in scenarios with thin haze layers.

According to Table I, our methods and LVE outperform other dehazing algorithms. PFM improves the performance of LVE when dealing with HRD. However, the performance slightly degrades in dehazing the validation image of SOTS; this should be ascribed to the simplification from (15) to (16), a compromise that uses arithmetic mean instead of geometric mean, and a fixed target brightness according to the rough statistics. On the other hand, RNet significantly outperforms PFM because of a direct transformation from the input image to the optimal coefficient. The experimental results show that RNet can generate beneficial coefficients, so all the non-reference benchmark results show a complete upgradation compared with LVE and PFM. Besides, the performance of CLNet is very competitive, outperforming RNet concerning almost all non-reference benchmarks based on SOTS and being comparable to RNet in tests based on HRD. We ascribe the superior performance to the contrastive regularization and the reasonable designation; i.e., this method estimates the brightness of the input image instead of the fixed optimal brightness based on statistics and improves the defection of using the arithmetic mean.

B. Evaluation of Single Image Quality

The above experiments show that the proposed method improves dehazing performance by automatically choosing the optimal coefficient. However, high restoration accuracy does not guarantee high image quality; therefore, we use PIQE, NRIQA, and BIQE to assess the image quality concerning low-level features and image naturalness and demonstrate the experimental results in Table II. PIQE, NRIQA, and BIQE indicate high image quality with lower numbers. The experimental results show that CLNet is superior to state-of-the-art methods again; meanwhile, it also outperforms PMF and RNet. Combining experimental results demonstrated in Table I and Table II, we prove that the optimal coefficient generated by CLNet benefits restoration accuracy and image quality.

TABLE I

TABLE II

NON-REFERENCE BENCHMARK RESULTS

Benchmark	Dataset	DCP	BCCR	NLD	CAP	AOD	GFN	CEE	CIH	LVE	Ours PFM	Ours RNet	Ours CLNet
PIQE	SOTS	40.701	35.934	39.317	42.451	43.863	40.701	42.428	39.870	36.262	35.925	34.802	**34.173**
NRIQA	SOTS	26.815	24.636	26.354	32.527	30.522	26.986	25.904	25.728	24.527	24.671	24.119	**23.991**
BIQE	SOTS	2.8782	2.9021	2.8308	3.0936	3.1258	2.9992	2.9066	2.9582	2.9129	2.8999	2.8629	**2.8015**
PIQE	HRD	49.126	43.682	46.419	53.612	52.167	48.727	51.584	45.683	44.614	43.892	43.262	**42.612**
NRIQA	HRD	35.125	32.581	35.175	38.761	38.596	39.681	40.174	34.982	33.001	33.252	33.016	**32.828**
BIQE	HRD	3.5126	3.3117	3.3671	3.8011	3.8191	3.4100	3.7851	3.6199	3.2162	3.0991	2.9971	**2.8671**

Bold: the best result among the tested algorithms

[illegible]	SOTS	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
SSIM	HRD	0.6055	0.6796	0.6393	0.6385	0.6005	0.5168	0.4888	0.6680	0.7211	0.7384	**0.7528**	0.7514
-Luminance	HRD	0.8750	0.9215	0.9013	0.9172	0.8961	0.7448	0.8062	0.9081	0.9169	0.9191	**0.9295**	0.9288
-Contrast	HRD	0.7353	0.7780	0.7818	0.7265	0.7142	0.7589	0.6734	0.7922	0.8012	0.8135	**0.8217**	0.8199
-Structure	HRD	0.9628	0.9495	0.8739	0.9487	0.9463	0.9222	0.9312	0.9232	0.9710	0.9730	0.9780	**0.9782**
F&T	HRD	0.7969	0.7862	0.8165	0.7898	0.7938	0.8101	0.8057	0.8105	0.8224	0.8276	**0.8365**	0.8347
DehazeFR	HRD	0.8482	0.8950	0.9045	0.8128	0.8075	0.8913	0.8577	0.9105	0.9160	0.9183	**0.9257**	0.9254
MSE	HRD	0.0388	0.0358	0.0443	0.0395	0.0339	0.0513	0.0390	0.0294	0.0312	0.0291	0.0285	**0.0282**
PSNR	HRD	14.661	15.147	14.576	15.506	15.473	13.813	15.832	16.113	16.203	16.496	17.295	**17.305**
CieDE2000	HRD	14.967	13.853	15.046	14.229	14.395	16.277	16.634	12.058	11.900	11.757	11.254	**11.015**

Bold: the best result among the tested algorithms

C. Evaluation of Execution Time

We conducted experiments to assess the performance of various network models and quantify the average inference time. Table III showcases the experimental findings undertaken in the scope of SOTS. Generally, lightweight networks meet the requirements of RNet and CLNet due to the comparatively more straightforward complexity involved in optimizing coefficients or brightness than generating an image. However, our experimental results indicate that RNet's performance is more dependent on network depth, steadily increasing from ResNet-18 to ResNet-50. Conversely, increasing network depth only marginally enhances CLNet's performance

TABLE III
EXECUTION TIME OF PROPOSED METHOD

Methods	Networks	SSIM	Luminance	Contrast	Structure	F&T	DehazeFR	MSE	PSNR	CieDE2000	Inference Time
RNet	ResNet18	0.8912	0.9226	0.9652	0.9861	0.8878	0.9831	0.0105	22.115	6.0525	**4.51**
-	ResNet34	0.9026	0.9286	0.9691	0.9899	0.8911	0.9895	0.0103	22.209	6.0217	7.62
-	ResNet50	0.9078	0.9297	**0.9704**	**0.9905**	0.8932	**0.9903**	0.0102	22.275	6.0185	10.37
CLNet	ResNet18	0.9017	0.9292	0.9697	0.9872	**0.8936**	0.9867	0.0106	22.369	6.0125	4.58
-	ResNet34	**0.9110**	**0.9307**	**0.9704**	0.9898	0.8924	0.9871	**0.0101**	**22.373**	**6.0107**	7.91
-	ResNet50	**0.9156**	**0.9353**	**0.9721**	0.9904	**0.8956**	0.9873	**0.0101**	**22.375**	**6.0105**	11.14

Bold: the best RNet reults and CLNet results better than that
Inference Time: milliseconds (ms.)

when the depth remains below 34, as evidenced by the limited improvement between ResNet-34 and ResNet-50 results. These experiments highlight the intricacy of solving the coefficient optimization problem compared to brightness estimation, necessitating a more profound architecture for RNet to excel. On the contrary, CLNet consistently outperforms RNet, showcasing significant superiority. We denote the best RNet results concerning dehazing algorithms with bold; meanwhile, we also indicate LNet results surpassing the best RNEt results in bold. The superiority of CLNet, based on ResNet34, over RNet based on ResNet50, is unmistakable.

Finally, we compute the average inference time based on 100 executions using images of identical sizes. The results, denoted in milliseconds (ms.) and presented on the right side of Table II, indicate that the average inference times for RNet and CLNet are nearly identical, with slight discrepancies attributed to additional calculations.

V. CONCLUSIONS

In this manuscript, we present several novel coefficient optimization frameworks for dehazing algorithms, encompassing frameworks based on regression, deep-learning networks, and contrastive learning. Despite the widespread adoption of deep-learning-based dehazing algorithms, our findings reveal their limitations in addressing thick haze conditions. Moreover, the end-to-end framework exacerbates these challenges. In contrast, a configurable dehazing algorithm offers adjustability through coefficients to enhance brightness and contrast. However, selecting an appropriate coefficient poses difficulties. Our approach aims to assist configurable dehazing algorithms in determining optimal coefficients, thus overcoming the rigidity inherent in deep-learning-based approaches. By mitigating inconvenience and enhancing performance, our method extends the capabilities of configurable dehazing algorithms. Ultimately, our approach is evaluated and demonstrated to outperform existing methods, ensuring flexible usage and superior performance.

ACKNOWLEDGMENTS

The National Science and Technology Council partially supported this work under contract numbers 112-2222-E-167-003-; meanwhile, this work was also supported in part by the National Chin Yi University of Technology.

REFERENCES

[1] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 33, no. 12, pp. 2341-2353, Dec 2011.

[2] 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., Dec. 2013, pp. 617–624.

[3] D. Berman, T. Treibitz, and S. Avidan, “Non-local image dehazing,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 1674–1682

[4] Q. Zhu, J. Mai and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," in IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3522-3533, Nov. 2015, doi: 10.1109/TIP.2015.2446191.

[5] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “Aod-net: All-in-one dehazing network,” Proceedings of IEEE International Conference on Computer Vision, Oct. 2017, pp. 4780–4788.

[6] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W Liu, and M. Yang, “Gated fusion network for single image dehazing,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[7] H. Dong et al., "Multi-Scale Boosted Dehazing Network With Dense Feature Fusion," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2154-2164, doi: 10.1109/CVPR42600.2020.00223.

[8] X. Liu, H. Li and C. Zhu, "Joint Contrast Enhancement and Exposure Fusion for Real-World Image Dehazing," in IEEE Transactions on Multimedia, doi: 10.1109/TMM.2021.3110483.

[9] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, “Benchmarking single-image dehazing and beyond,” in IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 492-505, Jan. 2019.

[10] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual

Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
[11] S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated Residual Transformations for Deep Neural Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5987-5995, doi: 10.1109/CVPR.2017.634.
[12] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.
[13] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, “FFA-Net: Feature fusion attention network for single image dehazing,” Proceedings of the AAAI Conference on Artificial Intelligence, 2020.
[14] P. L. Suárez, D. Carpio, A. D. Sappa and H. O. Velesaca, "Transformer based Image Dehazing," 2022 16th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Dijon, France, 2022, pp. 148-154, doi: 10.1109/SITIS57111.2022.00037.
[15] S. C. Agrawal and A. S. Jalal, "Dense Haze Removal by Nonlinear Transformation," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 2, pp. 593-607, Feb. 2022, doi: 10.1109/TCSVT.2021.3068625.
[16] P. Narayanan et al., "A Multi-Purpose Realistic Haze Benchmark With Quantifiable Haze Levels and Ground Truth," in IEEE Transactions on Image Processing, vol. 32, pp. 3481-3492, 2023, doi: 10.1109/TIP.2023.3245994.
[17] J. Liu, F. Yuan, C. Xue, Z. Jia and E. Cheng, "An Efficient and Robust Underwater Image Compression Scheme Based on Autoencoder," in IEEE Journal of Oceanic Engineering, vol. 48, no. 3, pp. 925-945, July 2023, doi: 10.1109/JOE.2023.3249243.
[18] X. Sun, Y. Zhu and X. Fu, "RGB and Optimal Waveband Image Fusion for Real-Time Underwater Clear Image Acquisition," in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-17, 2023, Art no. 5019517, doi: 10.1109/TIM.2023.3290366.
[19] P. J. Liu, "Enhance Low Visibility Image Using Haze-Removal Framework," in IEEE Access, doi: 10.1109/ACCESS.2023.3322041.
[20] W. Yan, R. T. Tan, and D. Dai, "Nighttime Defogging Using High-Low Frequency Decomposition and Grayscale-Color Networks," European Conference on Computer Vision, ECCV, 2020.
[21] Y. Liu, Z. Yan, T. Ye, A. Wu, T. Li, "Single nighttime image dehazing based on unified variational decomposition model and multi-scale contrast enhancement," Engineering Applications of Artificial Intelligence, Volume 116, Issue C, 2022, Nov.
[22] P.J. Liu, S.J. Horng, J.S. Lin, and T. Li, “Contrast in haze removal: configurable contrast enhancement model based on dark channel prior,” Image Processing, IEEE Tans. on. vol. 28, issue 5, pp. 2212-2227, 2019.
[23] H. Yeganeh and Z. Wang, "Objective Quality Assessment of Tone-Mapped Images," in IEEE Transactions on Image Processing, vol. 22, no. 2, pp. 657-667, Feb. 2013, doi: 10.1109/TIP.2012.2221725.
[24] K. He, J. Sun and X. Tang, "Guided Image Filtering," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 6, pp. 1397-1409, June 2013, doi: 10.1109/TPAMI.2012.213.
[25] Y. Shao, L. Li, W. Ren, C. Gao and N. Sang, "Domain Adaptation for Image Dehazing," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 2805-2814, doi: 10.1109/CVPR42600.2020.00288.
[26] R. Hadsell, S. Chopra and Y. LeCun, "Dimensionality Reduction by Learning an Invariant Mapping," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 2006, pp. 1735-1742, doi: 10.1109/CVPR.2006.100.
[27] Y. Yuan, W. Chen, Y. Yang and Z. Wang, "In Defense of the Triplet Loss Again: Learning Robust Person Re-Identification with Fast Approximated Triplet Loss and Label Distillation," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1454-1463, doi: 10.1109/CVPRW50498.2020.00185.
[28] L. Li et al., "Semi-Supervised Image Dehazing," in IEEE Transactions on Image Processing, vol. 29, pp. 2766-2779, 2020, doi: 10.1109/TIP.2019.2952690
[29] H. Wu et al., "Contrastive Learning for Compact Single Image Dehazing," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 10546-10555, doi: 10.1109/CVPR46437.2021.01041.
[30] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie and M. Li, "Bag of Tricks for Image Classification with Convolutional Neural Networks," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 558-567, doi: 10.1109/CVPR.2019.00065.
[31] W. Zhou, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” in IEEE Transactions on Image Processing, vol. 13, issue 4, pp. 600-612, 2004.
[32] Sharma, Gaurav, Wencheng Wu, and Edul N. Dalal, "The CIEDE2000 Color-Difference Formula:

Implementation Notes, Supplementary Test Data, and Mathematical Observations". Color Research and Application 30, no. 1 (February 2005): 21–30.
[33] H. Ziaei Nafchi, A. Shahkolaei, R. Farrahi Moghaddam and M. Cheriet, "FSITM: A Feature Similarity Index for Tone-Mapped Images," in IEEE Signal Processing Letters, vol. 22, no. 8, pp. 1026-1029, Aug. 2015, doi: 10.1109/LSP.2014.2381458.
[34] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," Image Processing, IEEE Tans. On. Volume: 22, Issue: 2, Feb. 2013.
[35] X. Min et al., "Quality Evaluation of Image Dehazing Methods Using Synthetic Hazy Images," in IEEE Transactions on Multimedia, vol. 21, no. 9, pp. 2319-2333, Sept. 2019, doi: 10.1109/TMM.2019.2902097.
[36] N. Venkatanath, D. Praneeth, Bh. Maruthi Chandrasekhar, S. S. Channappayya and S. S. Medasani, "Blind image quality evaluation using perception based features," 2015 Twenty First National Conference on Communications (NCC), 2015, pp. 1-6, doi: 10.1109/NCC.2015.7084843.
[37] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," in IEEE Transactions on Image Processing, vol. 21, no. 12, pp. 4695-4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
[38] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a "Completely Blind" Image Quality Analyzer," in IEEE Signal Processing Letters, vol. 20, no. 3, pp. 209-212, March 2013, doi: 10.1109/LSP.2012.2227726.
[39] Y. Hang, L. Ding, G. Sharma, "HazeRD: An outdoor scene dataset and benchmark for single image dehazing," Proc. IEEE Intl. Conf. Image Proc., pp. 3205–3209, 2017.

Ping Juei Liu received a Ph.D. degree in C Computer Science and Information Engineering from the National Taiwan University of Science and Technology. He is an Assistant Professor in the Department of Artificial Intelligence and Computer Engineering at the National Chin-Yi University of Technology in Taiwan. His research interests include computational photography, image processing, and machine learning.