FWB-NET: FRONT WHITE BALANCE NETWORK FOR COLOR SHIFT CORRECTION IN SINGLE IMAGE DEHAZING VIA ATMOSPHERIC LIGHT ESTIMATION

Cong Wang1, Yan Huang1, Yuexian Zou1,2, Yong Xu2,3

1 ADSLAB, School of ECE, Peking University, Shenzhen, China
2 Peng Cheng Laboratory, Shenzhen, China
3 School of Computer Science and Engineering, South China University of Technology, Guangzhou, China

ABSTRACT

In recent years, single image dehazing deep models based on Atmospheric Scattering Model (ASM) have achieved remarkable results. But the dehazing outputs of those models suffer from color shift. Analyzing the ASM model shows that the atmospheric light factor (ALF) is set as a scalar which indicates ALF is constant for whole image. However, for images taken in real-world, the illumination is not uniformly distributed over whole image which brings model mismatch and possibly results in color shift of the deep models using ASM. Bearing this in mind, in this study, first, a new non-homogeneous atmospheric scattering model (NH-ASM) is proposed for improving image modeling of hazy images taken under complex illumination conditions. Second, a new U-Net based front white balance module (FWB-Module) is dedicatedly designed to correct color shift before generating dehazing result via atmospheric light estimation. Third, a new FWB loss is innovatively developed for training FWB-Module, which imposes penalty on color shift. In the end, based on NH-ASM and front white balance technology, an end-to-end CNN-based color-shift-restraining dehazing network is developed, termed as FWB-Net. Experimental results demonstrate the effectiveness and superiority of our proposed FWB-Net for dehazing on both synthetic and real-world images.

Index Terms— Single image dehazing, atmospheric scattering model, color shift correction, front white balance

1. INTRODUCTION

Single image dehazing aims to recover the clean image from a hazy input, which has received significant attention in the vision community over the past few years. According to the atmospheric scattering model (ASM) [1, 2, 3], the hazing process is usually formulated as:

\[ I(x, y) = J(x, y) t(x, y) + A \left( 1 - t(x, y) \right) \]

(1)

\[ t(x, y) = e^{-\beta d(x, y)} \]

(2)

where \(I(x, y)\) is the hazy image, \(J(x, y)\) is the haze-free image, \(\beta\) is a coefficient which represent the degree of haze. \(d(x, y)\) denotes the distance from object to camera. \(t(x, y)\) is transmission map. \(A\) is global atmospheric light factor, which can be considered as a constant among the whole image. Given a hazy image, most dehazing algorithms try to estimate \(t\) and \(A\), and calculate clean image \(J\) by equation (3):

\[ J(x, y) = \frac{I(x, y) - A}{t(x, y)} + A \]

(3)

However, it has been observed that the atmospheric light factor \(A\) is not globally constant under any conditions, such as there are both shadow and bright areas in one image. This deviation in modeling causes some bias in dehazing results. To modeling more general dehazing scenes with non-homogeneous illumination, we modify the ASM and correcting the atmospheric light factor to be related to pixel position on image. This model is termed as non-homogeneous atmospheric scattering model (NH-ASM).

As for single image dehazing methods, they can be categorized as: image prior-based methods and deep learning-based methods. Image prior-based methods try to find prior information that can be definitive in improving visibility of hazy images. Such as DCP [4] and Berman [5] estimate the transmission map \(t\) based on dark channel prior and haze-line prior, then getting haze-free image by equation (3). Deep learning-based methods try to estimate parameters of ASM by neural networks. DehazeNet [6] use CNN [7] to obtain multi-scale features and output transmission. AOD-Net [8], GFN [9], GDN [10] and ABC-Net [11] are end-to-end neural networks which can output haze-free image directly. In particular, ABC-Net can effectively restrain local color shift of dehazing result via specially designed loss and activation function.

However, all these dehazing algorithms focus on solving the transmission, but using very rough methods to solve the atmospheric light factor. It causes color shift issue in dehazing result due to inaccurate atmospheric light estimation. This kind of color shift diffuses to the whole image and expresses especially serious in real-world dehazing scene. This phenomenon is similar to the visual characteristic of images taken in a colored light environment. The technique, which restores the original color of objects from an image taken in a colored light environment, is called white balance [12]. The core of white balance is to accurately estimate the color of light from an image, and using it to adjust the color of objects in the image. As for dehazing, the atmospheric light factor in ASM reflects the comprehensive influence of light. Therefore, in order to restrain the color shift of dehazing result, it only needs to estimate the color of light by accurately calculating the atmospheric light factor.

Based on the above motivation, we propose a pixel level, U-Net-based [13] atmospheric light estimation module, and design an exclusively color-shift-restraining loss function for training this module. This module can accurately estimate the atmospheric light factor and further restrain the color shift of dehazing results.
factor of an input hazy image pixel by pixel, and restrain the diffuse color shift of dehazing result. It plays a similar role in picture editing. And this module has already completed the adjustment for color before creating dehazing results, instead the traditional white balance technology adjusts the color after generating images. So, we term it as front white balance module.

Based on the non-homogeneous atmospheric scattering model (NH-ASM), we propose an end-to-end single image dehazing network with restraining color shift, which applies front white balance technology. Experiments show that our method is capable of keeping original colors of objects in dehazing images. The major contributions of our work are:

1. We propose a non-homogeneous atmospheric scattering model (NH-ASM), which models atmospheric light factor as related to pixel index in image. It can express the imaging law of real-world hazy scene better.
2. We propose a new atmospheric light estimation module, which can effectively limit the global color shift of dehazing. It is termed as front white balance module (FWB-Module), as it plays a similar role to white balance before generating the dehazing result.
3. We propose a new loss function for training FWB-Module, termed as FWB-Loss. It can impose penalty on color shift to keep the original color of objects in dehazing result.
4. We propose an end-to-end CNN-based dehazing network, termed as FWB-Net, which applies FWB-Module and FWB-Loss. The experimental results on both synthetic and real-world images prove the effectiveness of our proposed method.

2. PROPOSED METHOD

2.1. Non-homogeneous atmospheric scattering model

The classic form of atmospheric scattering model (ASM) is showed in equations (1) (2). In ASM, the atmospheric light factor A is a global constant, which is in line with the illumination characteristics of open field under sunlight. However, when there are shadows or multiple light sources in one scene, the optical property of different pixels in the image are obviously distinct. Therefore, we modify the ASM by changing the atmospheric light factor as position dependent. The new model is termed as non-homogeneous atmospheric scattering model (NH-ASM), which can describe the diverse and complex illumination conditions in real world more precisely. And the form of NH-ASM is showed in equations (4):

\[ I(x, y) = f(x, y)t(x, y) + A(x, y)(1 - t(x, y)) \]

2.2. Front white balance module (FWB-Module)

According to NH-ASM (equation 4), there is a pixel-to-pixel and linear mapping relationship between hazy image I and atmospheric light map A. And it is suitable for modeling this mapping by an encoder-decoder model. Specifically, we use a U-Net [13] architecture with multi-scale skip connections between the encoder and decoder. Our framework is shown on Fig. 2, and it consists of two main units: the first is a 4-level encoder unit that is responsible for extracting a multi-scale latent representation of our input image; the second is a 4-level decoder unit, which has a bottleneck and transposed convolutional layers. At the first level of our encoder and decoder, the conv layers have 24 channels. For each subsequent level, the number of channels is doubled (i.e., the fourth level has 192 channels for each conv layer).

This module can learn the mapping between hazy image and atmospheric light factor, and get pixel level atmospheric light maps. Due to the diffuse color shift in an image can be considered as the illumination of hazy image has chromatic color. The atmospheric light estimate module can restrain the global distortion caused by illumination’s color of dehazing result, so it has the similar effect as white balance. Moreover, this module has already corrected color shift before generating dehazing results, instead of traditional white balance technology, which adjusts the color after creating images. So, we term it as front white balance module (FWB-Module).

2.3. Front white balance loss function (FWB-Loss)

Color space is an abstract mathematical model, which describes the specific organization of colors. In color space, colors can be represented as tuples of numbers. Except the well-known RGB color space, there are CMYK, CIELAB [14] as common color spaces. Unlike RGB and CMYK, CIELAB is designed to approximate human vision. It aspires to perceptual uniformity of color. Therefore, we convert the image data from RGB to CIELAB color space, and taking the minimum deviation between dehazing results and ground truth haze-free images in CIELAB as the optimization goal. It can restrain the color shift of dehazing results in vision better.

Denote the value of each channel in RGB color space at pixel position \((x, y)\) on the dehazing image and ground truth haze-free image as \(f_{est}(x, y, c)\) and \(f_{gt}(x, y, c)\), where \(c \in \{R, G, B\}\). In order to complete the conversion from RGB to CIELAB color space, we should transform the value of RGB into an auxiliary XYZ color space first, and then converting the value from XYZ to CIELAB. This process can be expressed as equation (5, 6, 7) [15]:

\[
\begin{align*}
J_{type}(x, y, X) &= 0.412 \times 0.358 + 0.180 \times 0.192 \\
J_{type}(x, y, Y) &= 0.213 \times 0.715 + 0.072 \times 0.299 \\
J_{type}(x, y, Z) &= 0.019 \times 0.110 + 0.950 \times 0.168 \\
J_{type}(x, y, L) &= [f_{est}(x, y, X) - 500] \quad \text{if } t > 950 \\
J_{type}(x, y, a) &= [f_{est}(x, y, Y) - 500] \quad \text{if } t > 450 \\
J_{type}(x, y, b) &= [f_{est}(x, y, Z) - 500] \quad \text{if } t > 950 \\
f(t) &= \begin{cases} 
1 \quad \text{if } t > \frac{950}{29} \\
\frac{3}{2} \times \left(\frac{t}{16} + \frac{29}{29}\right) \quad \text{otherwise}
\end{cases}
\end{align*}
\]

where type \(\in \{est, gt\}\). And as our experiments show that 99.9999% of the color in CIELAB color space are greater than \(\frac{81}{29}\), the equation (6) and (7) can be simplified as follow:

\[
\begin{align*}
J_{type}(x, y, L) &= 0 \quad \text{if } t > 950 \\
J_{type}(x, y, a) &= [f_{est}(x, y, X) - 500] \quad \text{if } t > 450 \\
J_{type}(x, y, b) &= [f_{est}(x, y, Z) - 500] \quad \text{if } t > 950 \\
J_{type}(x, y, X) &= 0 \quad \text{if } t > 950 \\
J_{type}(x, y, Y) &= 0 \quad \text{if } t > 450 \\
J_{type}(x, y, Z) &= 0 \quad \text{if } t > 950
\end{align*}
\]
As the optimization goal is to minimize the distance between the estimated dehazed image $f_{est}$ and ground truth haze-free image $f_{gt}$ in CIELAB color space, so the loss function can be defined as:

$$L_{sf} = \frac{1}{3MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \sum_{k} [f_{est}(x,y,k) - f_{gt}(x,y,k)]^2$$  \hspace{1cm} (9)

where $k \in \{L, a, b\}$ means the data channel of CIELAB color space.

From NH-ASM (equation 4), we have:

$$J(x,y,k) = \frac{t(x,y) - 1}{t(x,y)} \cdot A(x,y,k) + \frac{1}{t(x,y)}$$  \hspace{1cm} (10)

Substituting equation (10) into equation (9), we can get the following equation:

$$L_{sf} = \frac{1}{3MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \left( \frac{t(x,y)}{1-t(x,y)} \right)^2 \sum_{k} [A(x,y,k) - A_{est}(x,y,k)]^2$$  \hspace{1cm} (11)

Since $(1 - t(x,y))^2$ is irrelevant to atmospheric light factor $A$ as well as always positive, we omit it and get a simplified loss function:

$$L_{sf} = \frac{1}{3MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \sum_{k} [A_{est}(x,y,k) - A_{gt}(x,y,k)]^2$$  \hspace{1cm} (12)

Using equation (5) and (8) to convert the values of 3 channels of atmospheric light from RGB into CIELAB color space, then penalize the deviation between estimated and ground truth atmospheric light in CIELAB by equation (12). This operation can correct the color shift during dehazing. Since the loss function is used to optimize the front white balance module, we term it as front white balance loss function (FWB-Loss).

### 2.4. Front white balance network (FWB-Net)

Following ABC-Net [11], we design an end-to-end dehazing network based on NH-ASM. For restraining color shift, we apply FWB-Module and FWB-Loss to our dehazing network. That is the reason why the whole network is termed as FWB-Net. The architecture of FWB-Net is shown on Fig. 3 (a). This network has 3 modules: 1) front white balance module (FWB-Module); 2) estimating transmission module (ETM); 3) estimating haze-free image module (EHIM). The detail of FWB-Net is described in below.

**Front white balance module (FWB-Module):** Its details are described in section 2.2. This module estimates the color of illumination by calculating atmospheric light factor, and realizes the pre-correction of dehazing results’ color shift by this way.

**Estimate transmission module (ETM):** It reuses the ETM of ABC-Net. The module directly learns the mapping between haze image and transmission map through a multi-scale CNN network. In particular, the module also has excellent ability to restrain the color shift caused by inaccurate transmission estimation through specially designed loss function and activation function. It also helps the whole network to achieve color shift correction. The detail of ETM shows in Fig. 3 (b).

**Estimate haze-free image module (EHIM):** After obtaining transmission $t$ and atmosphere light factor $A$, the equation (3) is employed to figure out the haze-free image $J$. In order to refine the output and remove noise, we process $J$ with a convolutional layer followed by a BReLU activation function. The detail of EHIM shows in Fig. 3 (a).

Our proposed FWB-Net is delicately designed for single image dehazing with restraining color shift.

### 3. EVALUATIONS

#### 3.1. Datasets and training details

Adequate training data is essential for CNN-based methods, but since it is very difficult to collect both clean and hazy images of a same scene, real world dehazing databases are very scarce. Lots of methods, such as AOD-Net [8], ABC-Net [11], use ASM [1] to synthesize dehazing dataset. By convention, we use NYU-Depth-v2 [16] to synthesize our dataset. NYU-Depth-v2 provide 1449 clean images and theirs corresponding depth information. Referring to NH-ASM and setting $\beta=0.35$, we use 1000 images and the rest 449 images to create training and testing set, respectively.

At the same time, we randomly set the basic atmospheric light factor between 0.3 and 1.5 of each image, and apply a random disturbance of less than 20% to the atmospheric light factor of each pixel. Using this method, we synthesize the non-homogeneous atmospheric light map of each synthetic hazy image.

In training session, we use FWB-Loss and MSE loss to train the FWB-Module and ETM under the supervision of ground truth atmospheric light maps and transmission maps, respectively. Then, we use MSE loss to train the whole FWB-Net, under the supervision of ground truth haze-free images. With the help of SGD algorithm [17], we take 100 iterations to train FWB-Module and ETM in parallel, and take another 100 iterations to train the whole network, then repeat the cycle until all the FWB-Module, ETM and FWB-Net converge. While in testing session, both synthetic hazy images (NYU-Depth) and real-world hazy images (O-Haze) are input into the model and clean images are used for evaluation.

#### 3.2. Evaluations on synthetic hazy image dataset

We evaluate ABC-Net on a synthetic hazy dataset, which is obtained by setting $\beta=0.35$ and $A$ in the range of $[0.3, 1.5]$, randomly. Another seven high-performance methods [4-6, 8-11] are reimplemented for comparing. The results are shown in Fig. 5(a) and Table 1.

The results of image prior-based methods suffer from color shift in heavy hazy area or white image patches, and five deep learning-based methods avoid color shift but produce poor-effect dehazing outputs. However, the results of our FWB-Net are haze-free as well as avoiding color shift. It shows our method has the most outstanding performance of all the compared methods.
In this paper, we propose a new non-homogeneous atmospheric scattering model (NH-ASM) to precisely model the complex illumination conditions in real hazy scenes. Then, we design a new front white balance module (FWB-Module) and its exclusively training loss function (FWB-Loss), respectively. FWB-Module can correct the color shift of dehazing result by estimating atmospheric light factor. And FWB-Loss imposes penalty on color shift while training FWB-Module. We further incorporate the proposed FWB-Module and FWB-Loss to form a unified dehazing network called FWB-Net, which can restrain the color shift as well as dehazing. Experimental results show our method can get outstanding clean dehazing outputs without color shift.

**4. CONCLUSIONS**

In this paper, first, we propose a new non-homogeneous atmospheric scattering model (NH-ASM) to precisely model the complex illumination conditions in real hazy scenes. Then, we design a new front white balance module (FWB-Module) and its exclusively training loss function (FWB-Loss), respectively. FWB-Module can correct the color shift of dehazing result by estimating atmospheric light factor. And FWB-Loss imposes penalty on color shift while training FWB-Module. We further incorporate the proposed FWB-Module and FWB-Loss to form a unified dehazing network called FWB-Net, which can restrain the color shift as well as dehazing. Experimental results show our method can get outstanding clean dehazing outputs without color shift.
5. REFERENCES


