



DLT-Net: deep learning transmittance network for single image haze removal

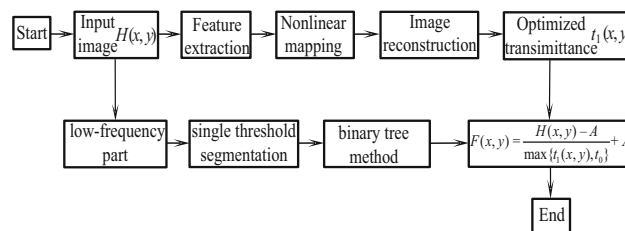
Bin Li¹ · Jingjuan Zhao² · Hui Fu³

Received: 12 September 2019 / Revised: 15 January 2020 / Accepted: 20 February 2020 / Published online: 2 March 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

Outdoor images taken in inclement weather conditions are often contaminated with colloidal particles and droplet in the atmosphere. These captured images are susceptible to low contrast, poor visibility, and color distortion, which is the reason for serious errors in digital image vision systems. Therefore, defogging research has material significance for practical applications. In this paper, image dehazing is regarded as a mathematical inversion and image restoration process on the basis of fog image degradation model. The global atmospheric light A can be approximately estimated by combining Gaussian low-pass filtering with the single-threshold segmentation and binary tree method. And a deep learning transmittance network is adopted to modify transmittance. Comparison experimental results show that our method is effective in dealing with thick fog, complex scenes and multicolor images. In addition, our method is superior to four other state-of-the-art defogging methods in visual impact, universality and running speed.

Graphic abstract The overall framework of our method



Keywords Mathematical inversion · Image restoration · Binary tree · Deep learning · Transmittance

1 Introduction

Currently, many computer vision technologies are required to work in perfect weather condition and high-definition environments, whereas the visibility of each pixel in the object image only relies on the pixel radiance of the scene. In the-

ory, intelligent vehicles, fire monitoring and other image vision applications can only work if the input image is haze-free. Therefore, contaminated images can bring about serious damage. Hence, the study of image defogging is of great value.

The primary target of enhancement-based defogging approaches is to artificially enhance the visual quality of degraded images [1–3]. The traditional enhancement-based defogging approaches contain image histogram equalization and its derivative methods [4]. Other main enhancement-based defogging approaches include Retinex methods [5], prior information-based enhancement methods [6], homomorphic filter dehazing algorithm [7], etc.

In general, the restoration-based defogging approaches perform better than other kinds of approaches. Sulami [8] proposed an algorithm related to Fattal's strategy [9]. This

✉ Bin Li
BinLi1977@163.com

¹ Chongqing Chongyou Information Technology Design Co., Ltd, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

² China Mobile Communications Group Chongqing Co., Ltd, Chongqing 400044, China

³ School of Information Engineering, Southwest University of Science and Technology, Mianyang 621000, China

algorithm assumed that the scene albedo and the transmittance are statistically independent. Sulami's method can not only obtain compelling defogging results but offer a reliable way of transmittance estimation. Nevertheless, this method cannot work effectively when the hypothesis is broken. On the basis of the multiple observation, He et al. [10] put forward the famous theory dark channel prior (DCP) which holds that at least one color channel in the clear outdoor image has very low pixel intensity. This method can obtain quite remarkable defogging results, while it consumes a long processing time because of soft matting strategy. To promote the efficiency, he further introduces guided filter [11] into this algorithm, which has a compelling effect on preserving edge and restraining gradient reversal artifacts along the sharp edges, whereas the DCP theory will be invalid when the scene objects are essentially similar to the global atmospheric light and no shadow is cast on the scene. Presently, various DCP model-based defogging algorithms [12–14] have been developed. Nevertheless, the transmittance map of these algorithms is not smooth and the image noise cannot be avoided. Li et al. [14] put forward a DCP-based deep learning algorithm for calculating the bright region of transmittance map based on inverse tolerance, while the calculation of tolerance needs parameter adjustment for different input haze images. The color aberration cannot be fully eliminated with too small tolerance, whereas restoration errors can be occurred in dark regions with too large tolerance. Through deep analysis of each pixel and training of each haze image, a multi-scale convolution network model is obtained by Ren et al. [15]. However, the processing efficiency of this method decreases with the increase in image size.

In this paper, the deep learning transmittance network is used to get and optimize transmittance. The other parameter, global atmospheric light value, is obtained by combining Gaussian low-pass filtering, simple threshold segmentation method and binary tree method.

2 Theoretical model

On the basis of the theory of atmospheric scattering [16], the fog image degradation model is composed of two parts: one is the attenuation of reflected light between the target and the camera, and another is the scattering process of global atmospheric light from the object to the camera. Therefore, the image degradation mechanism in heavy weather can be depicted by the atmospheric light attenuation model and the atmospheric light scattering model, as shown in Fig. 1. This is the theoretical basis of fog image degradation and the main basis of image restoration from degradation. The basic formula of the atmospheric scattering model [17] is given by Eq. (1).

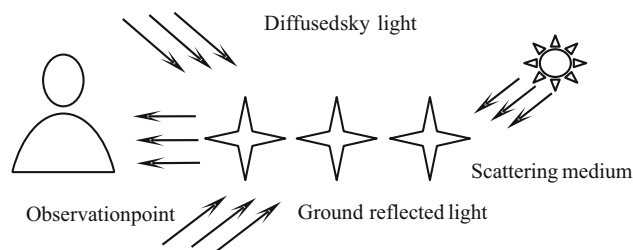


Fig. 1 Atmospheric light scattering model

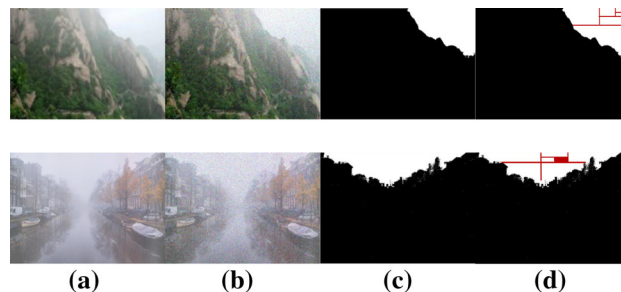


Fig. 2 Estimation of global atmospheric light. **a** Initial fog image, **b** image processed by low-pass filtering, **c** single-threshold segmentation, **d** binary tree method

$$H(x, y) = F(x, y)t(x, y) + A(1 - t(x, y)) \quad (1)$$

where (x, y) means the image pixel coordinates, $H(x, y)$ the fog image, $F(x, y)$ the fog-free image, A the global atmospheric light, and $t(x, y)$ the transmittance.

After estimation of A and $t(x, y)$, the fog-free image can be recovered from the deformation of Eq. (1).

3 Methods

3.1 Estimation of global atmospheric light value A

3.1.1 Obtain the low-frequency part of the image

In Park's theory [18], there are two reasons for image degradation in haze environment. First, the overall image color is gray white due to the expansion of illumination parameters; second, the reflection parameters are weakened, which weakens the image details, edges and other image characters. The interaction of these two influences makes the overall effect of haze image not good.

Therefore, Gaussian low-pass filtering is used to obtain the low-frequency part of the image, and then single-threshold segmentation method and binary tree algorithm are used to acquire the approximate global atmospheric light value, as can be seen in Fig. 2.

3.1.2 Estimation of approximate global atmospheric light A

Through continuous experiments, it can be seen that when the value of global atmospheric light is between 218 and 223, the dehazing effect is good. Besides, manual selection of threshold can save processing time. Therefore, we choose a single-threshold segmentation method [19] to obtain the approximate area of global atmospheric light A. And then global atmospheric light A is determined by binary tree method [20, 21].

3.2 Estimation of transmittance

3.2.1 Feature extraction

1. Dark channel feature map

The dark channel feature map is obtained on the basis of the theory of dark channel prior, i.e., in the non-sky region of a fog-free image, the pixel values of at least one color channel are extremely low and some are even close to zero. Dark channel mainly exists in the shadow of the object or landscape projection, color or black scenery surface. In the case of red objects, the green and blue channels are very low in brightness, so there are dark channels. In this paper, a 5*5 inverse convolution kernel is used to complete the convolution of the input image to obtain the dark channel feature map.

2. Hue difference

Ancuti [22] proposed that in the fog area, except for the depth discontinuous region, the brightness value of other parts changes very little. Based on this assumption, the degraded image and its semi-inverse image are transformed into HSI space. In RGB three channels, the pixel values of the original pixel and the inverse pixel are compared, respectively, and the larger value is taken. The semi-inverse image can be described as Eq. (2).

$$H_{si}^c(x, y) = \max[H^c(x, y), 1 - H^c(x, y)], \quad c \in \{R, G, B\} \tag{2}$$

where $H^c(x, y)$ represents the pixel grayscale values of RGB color channels and $H_{si}^c(x, y)$ the pixel grayscale values of semi-inverse haze images. $H_{si}^c(x, y)$ ranges from 0.5 to 1. Due to the high pixel value in the image area severely affected by haze, the semi-inverse pixel value remains unchanged.

The hue-difference ΔH can be obtained by taking the absolute value through the function $abs()$, as shown in Eqs. (3)–(5).

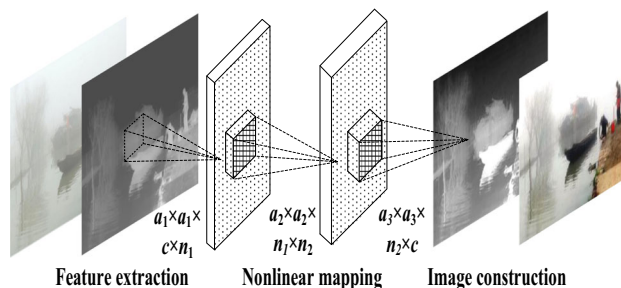


Fig. 3 Three modules of deep convolutional neural network

$$\theta = \left\{ \frac{1/2[(R - G) + (G - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \tag{3}$$

$$\mathbf{H} = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \quad \theta \in [0, 360] \tag{4}$$

$$\Delta H = abs(\mathbf{H} - \mathbf{H}_{si}) \tag{5}$$

where \mathbf{H} and \mathbf{H}_{si} mean hue value of $H^c(x, y)$ and $H_{si}^c(x, y)$. Then we manually set a threshold value t as 10° . The image part with ΔH less than this threshold is set as fog zone, and the image part with ΔH greater than this threshold is set as non-fog zone.

3.2.2 Mathematical model of DLT-Net

The operations of DLT-Net can be divided into three categories: feature extraction, nonlinear mapping and image reconstruction. And the overall structure is shown in Fig. 3.

1. The operation of feature extraction

Feature extraction $V_1(*)$ is an important part of DLT-Net. Specifically, it is the process of representing the image block of the initial transmittance map as high-dimensional variable, which can be expressed by Eq. (6).

$$V_1(H_e(x, y)) = f_1(\mathbf{W}_1 * H_e(x, y) + B_1) \tag{6}$$

where $H_e(x, y)$ denotes feature maps, which are composed of dark channel, hue-difference feature and grayscale feature maps, * convolution operation, $f_1(*)$ means activation function, and \mathbf{W}_1 and B_1 weights and deviations, respectively. Figure 3 shows an example, with c and n_1 representing the number of color channels and convolution kernel, a_1 representing the size of convolution kernel.

2. The operation of nonlinear mapping

The n_1 -dimensional image feature block extracted from the input transmittance is nonlinearly mapped to another

n_2 -dimensional vector space. The object of the nonlinear mapping operation is not the input transmittance block, but the feature image block. This process can be modeled as Eq. (7).

$$\begin{cases} V_2(x_1) = f_2(\mathbf{W}_2 * x_1 + B_2) \\ x_1 = V_1(H_e(x, y)) \end{cases} \quad (7)$$

Here, $f_2(*)$ means activation function, x_1 the intermediate variables, and \mathbf{W}_2 and B_2 weights and deviations, respectively. The output of the nonlinear mapping is close to the transmittance block of haze-free image in theory. Although increasing the number of convolutional layers is beneficial to feature extraction, the parameters of the network model will increase rapidly, leading to the difficulty of convergence and the increase in training time. Therefore, network design is very important.

3. The operation of image reconstruction

The image reconstruction uses a set of trained filters to average the overlapping areas of the feature image blocks in the previous layer to obtain the complete transmittance map. This process can be depicted as Eq. (8).

$$\begin{cases} V_3(x_2) = f_3(\mathbf{W}_3 * x_2 + B_3) \\ x_2 = V_2(x_1) \end{cases} \quad (8)$$

where $f_3(*)$ means activation function, x_2 the intermediate variables, \mathbf{W}_3 and B_3 weights and deviations, respectively.

3. Activation function

In the first group of DLT-Net, we used limited rectified linear unit (L-ReLU) and hyperbolic tangent (\tanh) as activation functions to adapt feature extraction process and ensure the transmittance between 0 and 1. The function of limited rectified linear unit (L-ReLU) and its derivative are Eqs. (9) and (10), respectively.

$$L - \text{ReLU} = \ln\left(1 + e^x/2\right) + 0.1x \quad (9)$$

$$L - \text{ReLU}'(x) = \frac{e^x}{1 + e^x} + 0.1 \quad (10)$$

It can be seen from the formula that L-ReLU has four basic properties necessary for activation function: (1) The function is nonlinear, which can play a good nonlinear mapping role in CNN; (2) since the derivative of the function (9) is greater than 0, the function is monotone decreasing, which ensures that every layer of network in CNN is a convex function; (3) when $x > 0$, the mathematical model of L-ReLU is equivalent to $f(x) = x$; (4) The output value of the function is infinite,

and the model can achieve higher training efficiency when training at a lower learning speed.

Compared with the traditional activation function, L_ReLU activation function does not disappear in the gradient descent method. As x approaches positive infinity and negative infinity, the limit of the derivative of ReLU function is 1.1 and 0.1, respectively. When the derivative of L_ReLU function is too large, its value is close to 1.1. While if x is too small, it's going to be close to 0.1 and it's not going to be 0. Therefore, L_ReLU activation function can be used for effective gradient descent training in CNN.

Besides, since the pixel value of the feature graph is greater than zero, the range of the tanh function is between 0 and 1, as shown in Eq. (11).

$$0 < \frac{(e^{H_e(x,y)} - e^{-H_e(x,y)})}{(e^{H_e(x,y)} + e^{-H_e(x,y)})} < 1 \quad H_e(x, y) > 0 \quad (11)$$

In the other groups of DLT-Net, rectified linear unit (ReLU) is selected as activation functions to improve processing efficiency.

5. Network training

Based on the objective of DLT-Net, Eq. (12) is used as the loss function [23] to compare the predictive transmittance map with the reference transmittance map.

$$C(\mathbf{W}, b) = \frac{1}{2n} \sum_i ||t_1(x, y) - t'(x, y)||^2 \quad (12)$$

Here, $t_1(x,y)$ and $t'(x,y)$ represent the predictive transmittance and the reference transmittance, respectively. \mathbf{W} and b mean weights and deviations, separately. n is the sum of transmittance map pixels in the training set.

In the process of training parameters, not only the synthetic haze images, but also the outdoor real scene haze images were selected. For synthesis haze images, $F^c(x, y)$ is the pixel value of ground-truth image, and for the real outdoor fog images, $F^c(x, y)$ is set to 0.015 on the basis of dark channel prior and repeated experiments. The reference transmittance can be obtained by combining Eq. (1), as shown in Eq. (13).

$$\begin{aligned} C(\mathbf{W}, b) &= \frac{1}{2n} \sum_i ||t_1(x, y) - t'(x, y)||^2 \\ &= \frac{1}{2n} \sum_i \left\| V_3(x_2) - \frac{||A - H^c(x, y)||}{||A - F^c(x, y)||} \right\|^2 \end{aligned} \quad (13)$$

Stochastic gradient descent algorithm (SGD) [24] is used to constantly adjust the weight and bias to get the minimum cost function. All initial weights of DLT-Net are randomly

initialized with a Gaussian distribution with standard deviation of 0.001, and all initial biases of DLT-Net are set to 0. The process of updating weights and biases by SGD method is given in Eqs. (14) and (15).

$$\mathbf{W} = \mathbf{W}_{(old)} - \eta \frac{\partial}{\partial \mathbf{W}_{(old)}} C(\mathbf{W}, b) \tag{14}$$

$$b = b_{(old)} - \eta \frac{\partial}{\partial b_{(old)}} C(\mathbf{W}, b) \tag{15}$$

In Eqs. (14) and (15), \mathbf{W} and b express weights and biases after the update, respectively, $\mathbf{W}_{(old)}$ and $b_{(old)}$ weights and biases before the update, separately.

3.2.3 Structure of DLT-Net

Before entering DLT-Net, preprocess the image and adjust its size to 483*483. This process can ensure that the image size after convolution, pooling or up-sampling is an integer.

1. The first group of DLT-Net

In the first group of DLT-Net, there are convolutional layers, pooling layers and fully connected layers. The input image is processed by $3 \times 5 \times 5$ inverse convolution kernel to obtain dark channel feature map. In the following two pair of convolution and pooling processes, with convolutional and pooling kernels $128 \ 3 \times 3 \times 3$, $128 \ 128 \times 3 \times 3$, $256 \ 128 \times 3 \times 3$ and $258 \ 256 \times 3 \times 3$, we can get $128 \times 239 \times 239$, $128 \times 239 \times 239$, $256 \times 59 \times 59$ and $256 \times 3 \times 3$ feature maps, respectively. Besides, we also adopt max-pooling process with a pooling kernel of $512 \ 256 \times 5 \times 5$ to obtain local extremum. The essence of this operation is to select the maximum value of the neighborhood, which not only has spatial invariance, but also satisfies the local constant hypothesis of hazy medium.

Finally, $1024 \ 1 \times 1$ feature map can be gotten through two sets of fully connected operations. And transmittance between 0 and 1 can be obtained by activation function tanh.

2. The second group of DLT-Net

In first group of DLT-Net, there are convolutional layers, pooling layers and up-sampling operation. The hue-difference feature map obtained by preprocessing is taken as the input image. In the following two pairs of convolution and pooling processes, with convolutional and pooling kernels $16 \ 3 \times 3$, $16 \ 16 \times 5 \times 5$, $32 \ 16 \times 3 \times 3$ and $32 \ 32 \times 3 \times 3$, we can get $16 \times 241 \times 241$, $16 \times 119 \times 119$, $32 \times 59 \times 59$, $32 \times 29 \times 29$ feature maps, respectively. Then we use convolutional kernel of $1 \ 32 \times 5 \times 5$ to obtain two-dimensional feature map. And the image size is linearly enlarged to 483×483 by the up-sampling operation.

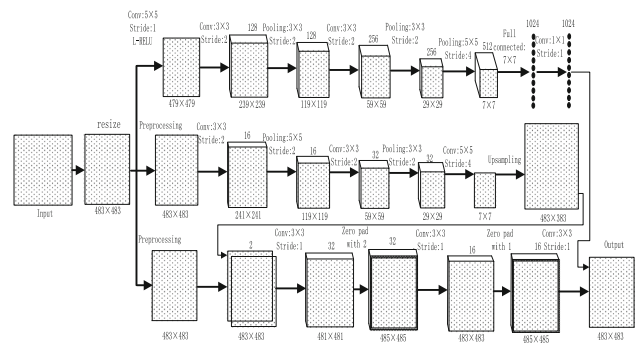


Fig. 4 Estimation of transmittance by DLT-Net

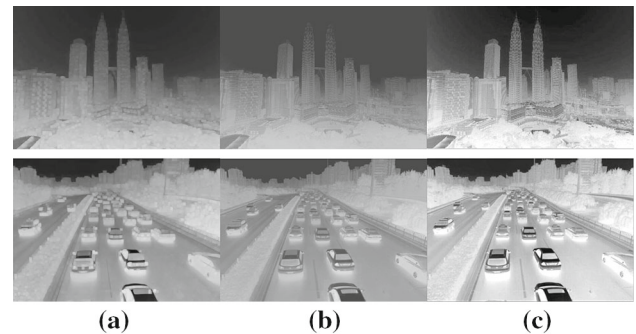


Fig. 5 Transmittance map of our method. **a** A simple dark channel of haze image, **b** initial transmittance, **c** transmittance of our method

3. The third group of DLT-Net

In third group of DLT-Net, we used convolution operation to deal with the fuzzy and noise problems caused by the output by the up-sampling operation of the second group, and optimize the detail and clarity of the transmittance map. The grayscale feature map obtained by preprocessing combined with the output image of the second group is taken as the input image. In addition, we use 0 padding to resize feature maps.

As an illustration, the DLT-Net is applied to estimate the transmittance map of a haze image. Besides, estimation of transmittance by DLT-Net and transmittance map of our method are shown in Fig. 4 and Fig. 5, respectively.

4 Comparison and analysis of experimental results

In this paper, the image defogging effectiveness is compared with other typical defogging methods. All the comparison experiments are performed on MATLAB 2017a under Windows 10 and run on a desktop computer in configuration with Intel(R) i7-6700U 3.4 GHz processor and 16 GB RAM. We created our own image set, which contains 5000 outdoor fog images, and completed the creation of synthetic fog images.

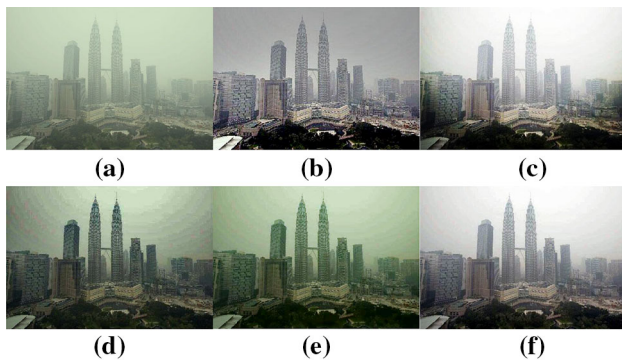


Fig. 6 Comparison of our approach and four advanced dehazing methods applied to haze images. **a** Initial haze image, **b–e** dehazing result by methods [2, 12, 14, 15], **f** our result

These images cover different outdoor natural scenes, including various natural landscapes, buildings, lake views, aerial photography images, distant and close views, etc.

This paper adopts subjective and objective evaluation strategies. Subjective evaluation strategy is based on human visual perception, while objective evaluation strategy is on the basis of experimental data. The objective evaluation strategy in this paper includes full-reference image quality assessment metrics, no-reference image quality assessment metrics, and the algorithm efficiency. And the comparison algorithms are method [2, 12, 14, 15].

4.1 Subjective evaluation

In this paper, haze images are selected to complete the experiment, including the test images of buildings, roads and lake, etc. The resolutions of the images were also set as 280×380 , 380×580 and 580×680 , respectively.

The scene in Fig. 6a is relatively simple, mainly buildings and trees. Here, we compared our results with four excellent dehazing algorithms: algorithm [2, 12, 14, 15], as shown in Fig. 6b–e. It can be seen from Fig. 6b that the method [2] has achieved good visual effect after image enhancement. Although method [12] improves the contrast, the color of the image is obviously distorted, especially in the sky area and the nearby leaves, as shown in Fig. 6c. Similar experimental results can also be seen in Fig. 6d. The result shows that method [14] has a low color fidelity and there are a lot of halo artifacts in a large area of sky. In Fig. 6e, there is much noise, and the image effect is slightly dark. And it can be easily seen in the canopy part of the image. Besides, the effect of our method is bright and clear.

The roads, cars and buildings in Fig. 7a contain very complex scenery, and it is difficult to restore image. The dehazing effect shown in Fig. 7b is excessively enhanced, and the color deviation of the sky area and the square is obvious. As can be seen from Fig. 7c, the brightness of recovery effect is

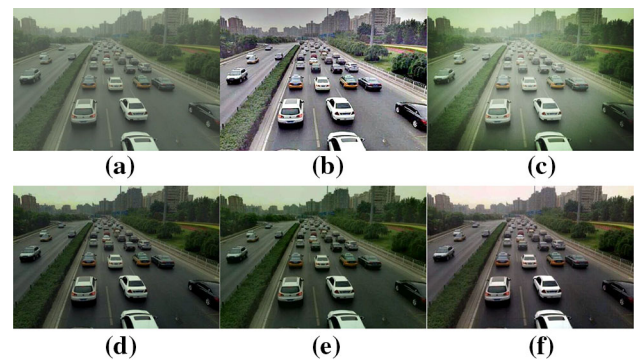


Fig. 7 Comparison of our approach and four advanced dehazing methods applied to haze images. **a** Initial haze image, **b–e** dehazing result by methods [2, 12, 14, 15], **f** our result

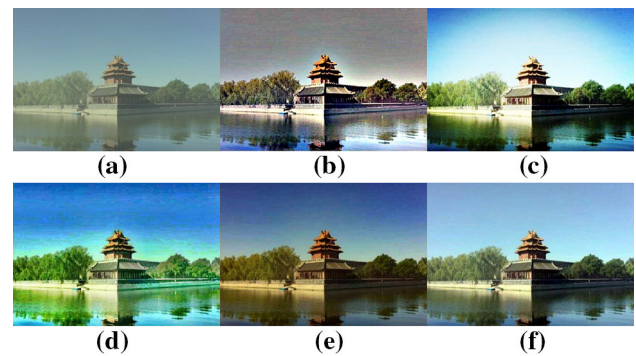


Fig. 8 Comparison of our approach and four advanced dehazing methods applied to haze images. **a** Initial haze image, **b–e** dehazing result by methods [2, 12, 14, 15], **f** our result

improved, but the residual amount of fog is relatively high. Halo effect exists in the dehazing effects of Fig. 7d and e, especially in Fig. 7e. In this paper, the close-up and distant scenes are restored well by our method, and a certain amount of fog is retained to make the image more realistic.

As regards the fog image contains large areas of sky and lake in Fig. 8a, the visual effect is greatly proposed by methods [14, 15] and our method, as shown in Fig. 8b–f. The output image of method [2] is obviously over-saturated, resulting in unnatural scene color distortion, which means that this method stretches local contrast and generates false edges. In the experiment of method [12], the noise variance is set to a relatively large value to restrain noise amplification. As the scene depth increases, the image noise increases sharply. Although the method [14] preserves image details of tree's areas, it leads to unnatural scene color. Although the restoration effect in Fig. 6e is close to the real scene, the details are somewhat fuzzy. Figure 8f displays the defogging result of our algorithm. It can be seen that fog most objects in the scene have been restored, proving that the haze has been almost cleared.

Table 1 The full-reference evaluation of our method and comparison methods

	MSE	SSIM	PSNR
Method [2]	12.8535	0.7732	16.1894
Method [12]	10.1229	0.7789	17.8603
Method [14]	10.0121	0.7851	17.1572
Method [15]	9.2126	0.8369	19.7163
Improved method	8.8915	0.8235	20.3895

4.2 Objective evaluation

4.2.1 Full-reference evaluation metrics

We also adopt full-reference evaluation metrics to compare the dehazing effect of our method and methods [25, 26], including mean square error (MSE), peak signal–noise ratio (PSNR) and structural similarity image measure (SSIM).

The full-reference evaluation of our method and comparison methods is shown in Table 1.

In general, a smaller MSE indicates that the defogging algorithm has good performance, and higher values of SSIM and PSNR imply a better performance of dehazing. Multi-scale analysis is used to process each pixel, so the SSIM value of method [15] is relatively high. The dehazing effect of method [2] has a lot of image noise, distortion and halo phenomenon, and the three indicators are not ideal. In the process of image partitioning by method [14], it is difficult to avoid image noise, so PSNR index is not very good. Halo effect and unnatural scene color distortion problems in method [12] result in poor MSE and PSNR indexes.

Table 2 The no-reference evaluation of comparison methods

Indicator	Images	e	r	σ	IIE
Fog image	Figure 6a	0.5398	1.2539	0.3131	7.2979
	Figure 7a	0.7352	1.5887	0.1979	7.5732
	Figure 8a	1.0601	1.7569	0.1023	6.9031
The algorithm [2]	Figure 6b	0.4815	2.2022	0.3685	7.7035
	Figure 7b	0.5396	2.2359	0.1359	7.2979
	Figure 8b	1.1368	2.3571	0.0889	7.1079
The algorithm [12]	Figure 6c	1.1382	1.4889	0.1856	7.3989
	Figure 7c	0.6015	2.1578	0.3969	7.5515
	Figure 8c	1.2812	2.2869	0.1339	6.9918
The algorithm [14]	Figure 6d	0.6541	1.8973	0.1049	7.4021
	Figure 7d	0.9482	1.7981	0.2621	7.6375
	Figure 8d	1.4051	1.9678	0.0886	7.2912
The algorithm [15]	Figure 6e	0.9842	1.8792	0.1531	7.7012
	Figure 7e	1.2165	1.9746	0.1668	7.8816
	Figure 8e	1.4608	1.8531	1.2885	7.0941
The improved algorithm	Figure 6f	1.3965	2.1915	0.0678	7.9031
	Figure 7f	1.1754	2.3979	0.0896	7.6986
	Figure 8f	1.4862	2.3665	0.0521	7.4033

4.2.2 No-reference evaluation metrics

Four famous no-reference evaluation metrics [27–29] are adopted to evaluate our method and four outstanding defogging methods, which contain e , r , ϵ and information entropy (IIE). The ratio e represents the new visible edges of restored image.

In general, a bigger IIE indicates that the defogging algorithm has good performance, making the image details clear and the contrast high. Moreover, higher values of e and r imply a better performance of image levers, details and contrast, whereas a value of ϵ approaches to zero reflects better visual effect.

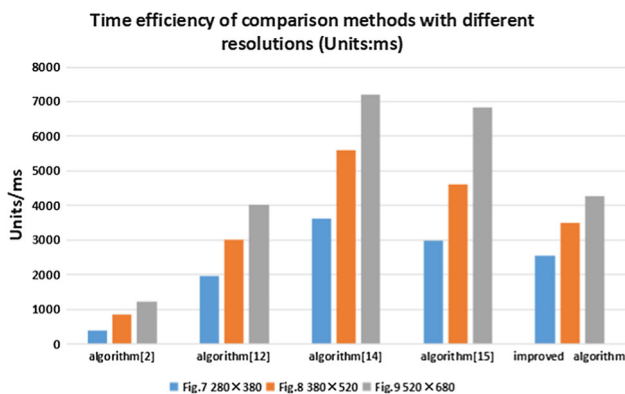
The results of the evaluations (e , r and ϵ) are shown in Table 2, and the bold numbers represent the best index results obtained by our algorithm or the algorithms [2, 12, 14, 15]. From the experimental data, our method gets most of the highest values of IIE, e and r , and the lowest value of ϵ . In short, the images recovered by our method have rich image details and layers, high image contrast and brightness. The r index of Fig. 6b processed by method [2] is the highest; this is because the image is over-enhanced and color cast phenomenon appears. The IIE index of Fig. 7e processed by method [15] is the highest; this is because the dehazing image contains a lot of interference information.

4.2.3 Time efficiency

The processing time [26] of our method and comparison methods in this paper is shown in Table 3. Taking an image with a resolution of 380*520 as an example, the processing time of our method is 3519 ms, which is relatively fast. Combined with the independent component analysis algo-

Table 3 Time efficiency of comparison methods with different resolutions (unit: ms)

Indicator	Resolution ratio	Algorithm [2]	Algorithm [12]	Algorithm [14]	Algorithm [15]	Improved algorithm
Figure 5	280 × 380	398	1963	3625	2980	2541
Figure 6	380 × 520	857	3015	5592	4621	3519
Figure 7	520 × 680	1235	4018	7196	6839	4278

**Fig. 9** The histogram comparison of MSE, SSIM and PSNR indexes among our method and methods [2, 12, 14, 15]

rithm, method [2] took 857 ms. Method [12] used the fast Fourier transform to complete image processing, which took 3015 ms. Besides, method [14] uses an adaptive wiener filter to process transmittance. By calculating the local variance of haze image, the image is divided into flat area, texture area and edge area. This process takes some processing time. The multi-scale analysis method was conducted on each pixel [15] to train different haze images, so the parameters training process took a long time (Fig. 9; Table 3).

5 Conclusions

In this paper, a deep learning transmittance network is designed to complete single image dehazing. Firstly, Gaussian low-pass filtering method is used to obtain the low-frequency part of the image, and then single-threshold segmentation method and binary tree algorithm are used to acquire the approximate global atmospheric light value. After that, a deep learning transmittance network DLT-Net is designed to estimate the transmittance map of haze images. Using the approximate global atmospheric light value and the optimized transmittance, the fog-free image can be recovered from the deformation formula of the atmospheric scattering model.

The main innovations of this paper are as follows.

1. The image captured in dense haze environment and the interference of white objects in haze image scene pose

challenges to the process of haze removal. The global atmospheric light A is usually located in the densest part of haze in the image, which belongs to the low-frequency region. Therefore, the approximate interval of the global atmospheric light value can be obtained by using Gaussian low-frequency filter.

2. A deep learning transmittance network DLT-Net is designed to estimate the transmittance map of haze images. The three-group network structure can output the transmittance map which is close to the real state. The effect and efficiency of our algorithm can be improved by extracting different image features and adjusting the image size through preprocessing, respectively.

References

1. Narasimhan, S.G., Nayar, S.K.: Chromatic framework for vision in bad weather. In: Proceedings of the IEEE CVPR00, pp. 598–605. IEEE Computer Society, Washington (2000)
2. Ru, Y., Tanaka, G.: Proposal of multiscale retinex using illumination adjustment for digital images. *Ieice Trans. Fundam. Electron. Commun. Comput. Sci.* **E99**(11), 2003–2007 (2016)
3. Raffei, A.M., Asmuni, H., Hassan, R., Othman, R.M.: A low lighting or contrast ratio visible iris recognition using iso-contrast limited adaptive histogram equalization. *Knowl. Based Syst.* **74**(1), 40–48 (2015)
4. Wang, Z., Feng, Y.: Fast single haze image enhancement. *Comput. Electr. Eng.* **40**(3), 785–795 (2014)
5. Babu, P., Rajamani, V.: Contrast enhancement using real coded genetic algorithm based modified histogram equalization for gray scale images. *Int. J. Imaging Syst. Technol.* **25**(1), 24–32 (2015)
6. Dippel, S., Stahl, M., Wiemker, R., Blaffert, T.: Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform. *IEEE Trans. Med. Imaging* **21**(4), 343–353 (2002)
7. Seow, M.J., Asari, V.K.: Ratio rule and homomorphic filter for enhancement of digital colour image. *Neurocomputing* **69**(7), 954–958 (2006)
8. Sulami, M., Glatzer, I., Fattal, R., Werman, M.: Automatic recovery of the atmospheric light in hazy images. In: IEEE International Conference on Computational Photography, pp. 1–11 (2014)
9. Fattal, R.: Single image dehazing. *ACM Trans. Graph.* **27**(3), 1–9 (2008)
10. He, K., Sun, J., Tang, X.O.: Single image haze removal using dark channel prior. In: Proceedings of the IEEE CVPR09, pp. 1956–1963. IEEE Computer Society, Washington (2009)
11. He, K., Sun, J., Tang, X.: Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(6), 1397–1409 (2013)
12. Fu, H., Wu, B., Shao, Y., Zhang, H.: Perception oriented haze image definition restoration by basing on physical optics model. *IEEE Photonics J.* **10**(3), 1–16 (2018)

13. Fu, H., Wu, B., Zhang, H.: De-hazing of atmosphere veil haze image. *Opt. Precis. Eng.* **24**(8), 2018–2026 (2016)
14. Li, J., Li, G., Fan, H.: Image dehazing using residual-based deep CNN. *IEEE Access*. **6**, 26831–26842 (2015)
15. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.H.: Single Image Dehazing via Multi-scale Convolutional Neural Networks. In: *European Conference on Computer Vision*, pp. 154–169 (2016)
16. Koschmider, H.: Theorie der horizontalen sichtweite *Contrib. Beitr. Phys. Freien Atmos.* **12**, 171–181 (1924)
17. Li, Z., Zheng, J.: Edge-preserving decomposition based single image haze removal. *IEEE Trans. Image Process.* **24**(12), 5432–5441 (2015)
18. Seonhee, P., Soohwan, Y., Minseo, K., Kwanwoo, P., Joonki, P.: Dual autoencoder network for retinex-based low-light image enhancement. *IEEE Access* **6**, 22084–22093 (2018)
19. Fu, H., Wu, B., Shao, Y., Zhang, H.: Scene-awareness based single image Dehazing technique via automatic estimation of sky area. *IEEE Access* **7**, 1829–1839 (2019)
20. Fu, H., Wu, B., Zhang, H., Xu, J.: Fast image dehazing based on dark-Chanel physical model. *Opto-Electron. Eng.* **2**, 82–88 (2016)
21. Veelaert, P.: Image segmentation with adaptive region growing based on a polynomial surface model. *J. Electron. Imaging* **22**(4), 3004 (2013)
22. Ancuti, C.O., Ancuti, C., Hermans, C., Philippe, B.: A fast semi-inverse approach to detect and remove the haze from a single image. In: *2010-10th Asian Conference on Computer Vision. Queenstown (ACCV)*, pp. 501–514(2010)
23. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **25**(11), 5187–5198 (2016)
24. Li, X.: Preconditioned stochastic gradient descent. *IEEE Trans. Neural Netw. Learn. Syst.* **29**(5), 1454–1466 (2018)
25. Wang, K., Wang, H., Li, Y., Hu, Y., Li, Y.: Quantitative performance evaluation for Dehazing algorithms on synthetic outdoor hazy images. *IEEE Access* **6**(99), 20481–20496 (2018)
26. Wang, W., Yuan, X., Wu, X., Liu, Y.: Fast image dehazing method based on linear transformation. *IEEE Trans. Multimed.* **19**(6), 1142–1154 (2017)
27. Luan, Z., Shang, Y., Zhou, X., Shao, Z., Guo, D., Liu, X.: Fast single image dehazing based on a regression model. *Neurocomputing* **245**, 10–22 (2017)
28. Hautiere, N., Tarel, J.P., Aubert, D.: Blind contrast enhancement assessment by gradient ratioing at visible edges. *Image Anal. Stereol.* **27**(2), 87–95 (2011)
29. Ju, M., Gu, Z., Zhang, D.: Single image haze removal based on the improved atmospheric scattering model. *Neurocomputing* **260**, 180–191 (2017)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.